



دانشگاه گیلان

۱۳۵۳_۱۹۷۴

چاپ اول

اکسل ۲۰۱۳

برای آمار کسب و کار
راهنمای حل مسائل کاربردی

تألیف:
توماس جی. کوئرک

ترجمه:
دکتر محمد کاووسی کلاشمی
استاد یار دانشکده علوم کشاورزی دانشگاه گیلان

اداره چاپ و انتشارات دانشگاه گیلان

Excel 2013

for Business Statistics a Guide
to Solving Practical Problems

By:
Thomas J. Quirk
Translated by:
Mohammad Kavoosi-Kalashami, Ph. D

University of Guilan Press

تألیف: توماس جی. کوئرک
ترجمه: دکتر محمد کاووسی کلاشمی



9 786001 532313

ISBN: 978-600-153-231-3



اکسل ۲۰۱۳

برای آمار کسب و کار راهنمای حل مسائل کاربردی

تألیف:
توماس جی. کوئرک

ترجمه:
دکتر محمد کاووسی کلاشی
استادیار دانشکده علوم کشاورزی

انتشارات دانشگاه گیلان
۱۳۹۹



شابک: ۹۷۸-۲۳۱-۱۵۳-۶۰۰-۹۷۸

سرشناسه : کوئرک، تامس ج.
Thomas J, Quirk

عنوان و نام پدیدآور : اکسل ۲۰۱۳ برای آمار کسب و کار/تالیف توماس جی کوئرک؛ ترجمه محمد کاووسی کلاشی؛ ویراستار علمی؛ رضا اسفنجاری کناری؛ ویراستار ادبی؛ فرشته گلچین راد.

مشخصات نشر : رشت: دانشگاه گیلان، انتشارات، ۱۳۹۹.

مشخصات ظاهری : ص: ۳۰۴؛ جدول: ۱۷×۲۴ س. م.

شابک : 978-600-153231-3

وضعیت فهرست نویسی : فیبا

یادداشت : عنوان اصلی: Excel 2013 for Business Statistics: A Guide to Solving Practical Problems

موضوع : اکسل مایکروسافت (فایل کامپیوتر)

موضوع : Microsoft Excel (Computer file)

موضوع : آمار ریاضی

موضوع : Mathematical statistics

موضوع : آمار ریاضی -- نرم افزار

موضوع : Mathematical statistics-- Software

موضوع : ریاضیات بازرگانی

موضوع : Business mathematics

شناسه افزوده : کاووسی کلاشی، محمد، -۱۳۶۲، مترجم

شناسه افزوده : اسفنجاری کناری، رضا، ویراستار

شناسه افزوده : دانشگاه گیلان، انتشارات

ردیه بندی کنگره : QA276

ردیه بندی دیجیتی : ۵۱۹/۵

شماره کتابشناسی ملی : ۷۲۸۹۵۶۴

اداره چاپ و انتشارات دانشگاه گیلان

نام کتاب :	اکسل ۲۰۱۳ برای آمار کسب و کار: راهنمای حل مسائل کاربردی
مؤلف :	توماس جی. کوئرک
مترجم :	دکتر محمد کاووسی کلاشی
ویراستار علمی :	دکتر رضا اسفنجاری کناری
ویراستار ادبی :	فرشته گلچین راد
نوبت چاپ :	اول، ۱۳۹۹
ناشر :	انتشارات دانشگاه گیلان
شماره گان :	۱۰۰ جلد

* هر گونه چاپ و تکثیر فقط در اختیار انتشارات دانشگاه گیلان است.*

سپاسگزاری

اکسل ۲۰۱۳ برای آمار کسب و کار: راهنمای حل مسائل کاربردی، نتیجه الهام از سه فرد مهم است: دو دختر و همسرم. Jennefer Quirk McLaughli چندین بار از من برای بازدید از کلاس‌های دوره مدیریت اجرایی خود در دانشگاه Witwatersrand در جوهانسبورگ، آفریقای جنوبی دعوت کرد. این بازدیدها از دوره مدیریت اجرایی برتر من را قانع کرد که نیاز به یک کتاب برای آموزش چگونگی حل مسائل عملی با استفاده از اکسل به دانشجویان وجود دارد. تعهد دائمی Meghan Quirk-Horton به یادگیری بسیاری از روش‌های آماری مورد نیاز برای تکمیل پایان نامه دکتری خود، نشان دهنده نیاز به یک کتاب آمار است که این کار دشوار را کاربر پسندتر کند و نفر اول مشوق این پروژه از آغاز بود، همیشه مرا تشویق کرد و به من کمک کرد تا برای تکمیل آن کوشباشم.

Marc Strauss، ویراستار ما در Springer، روح این ایده را در اولین مکالمه تلفنی ما درک کرد و این کتاب را از مرحله ایده به چالش کشید تا آن را به فرم نهایی خود برساند. تشویق و حمایت او برای این کتاب حیاتی بود. ما از وی برای اینکه قهرمان Bracken Hannah برجسته تولید محصول در تمام این فرایندها است، سپاسگزاریم و در Springer کار درجه یک خود را در هماهنگ کردن ویرایش و تولید این کتاب انجام داد؛ کار کردن با او همیشه لذت بخش است.

Thomas J. Quirk

سخن مترجم

گسترش پژوهش‌های میدانی و استفاده از روش‌های کمی در گرایش‌های مختلف رشته‌های اقتصاد و مدیریت، یادگیری، آموزش و به کارگیری ابزارهای آماری را برای پژوهشگران و دانشجویان ضروری ساخته است. علم آمار مجموعه‌ای از مفاهیم و روش‌های تحلیل و استنتاج داده‌هاست که می‌تواند عینیت تحقیق و دقت پژوهشگر را در توصیف و تفسیر پدیده‌ها افزایش دهد. از آنجا که بسیاری از پژوهشگران حوزه اقتصاد و مدیریت پاسخ پرسش‌های پژوهش خود را با روش‌های آماری به دست می‌آورند، دانستن مفاهیم و روش‌های آماری با رویکردی مناسب برای آن‌ها مفید و حتی لازم است.

مایکروسافت اکسل (Microsoft Excel) یک برنامه صفحه گسترده می‌باشد که توسط شرکت مایکروسافت توسعه یافته است. امروزه اکسل پرکاربردترین برنامه صفحه گسترده در دنیا محسوب می‌شود. این نرم‌افزار محبوبیت زیادی در ایران دارد و بسیاری از ما برای محاسبات ریاضی، آماری، متنی و ترسیم نمودار، آن را به کار می‌بریم. بعد از عرضه نسخه پنجم این برنامه صفحه گسترده در سال ۱۹۹۳ میلادی، این نرم‌افزار جزئی از مجموعه مایکروسافت آفیس (Microsoft Office) شد. کتاب حاضر کاربرد مفاهیم بنیادی آمار و روش‌های آماری را به زبانی ساده در قالب مسائل عملی و استفاده از نرم‌افزار اکسل، ارائه کرده است. فراهم آوردن شرایط کار عملی و چگونگی برخورد با مسائل در شرایط واقعی از تعريف مسئله تا نتیجه‌گیری و پیشنهاد سیاستی یکی از مزیت‌های اصلی این کتاب است. تغییرهای رخ داده در نسخه ۲۰۰۷ نسبت به نسخه ۲۰۰۳، بزرگ‌ترین دگرگونی نرم‌افزار اکسل محسوب می‌شود. بعد از نسخه ۲۰۰۷ در این نرم‌افزار تغییرات جزئی رخ داد که بیشتر در راستای افزایش ظرفیت و قابلیت الگو سازی داده بود و کاربران که به نسخه‌های قبلی این نرم‌افزار عادت کرده بودند، مشکل چندانی با این تغییرات نداشتند. بین نسخه‌های ۲۰۱۳، ۲۰۱۶ و ۳۶۵ یا ۲۰۱۹ تفاوت چندانی وجود ندارد و محیط کار نرم‌افزار در این نسخه‌ها بسیار شبیه به هم می‌باشد، تنها در نسخه‌های ۲۰۱۶ و ۲۰۱۹، تعدادی توابع، نمودار و امکانات گرافیکی جدید به محیط کار اضافه شده است. مسیرهای مورد استفاده برای تجزیه و تحلیل داده‌ها در این کتاب، در تمام نسخه‌های یادشده ثابت و بدون تغییر می‌باشد. از این‌رو، توضیحات ارائه شده در قالب نسخه ۲۰۱۳ اکسل، برای نسخه‌های بعدی نیز قابل تعمیم است.

امید است استفاده از این منبع آموزشی در یادگیری و تدریس درس‌های کاربرد آمار بتواند کمبود موجود در خصوص رویکرد عملی به حل مسائل آماری را جبران کند.

این کتاب را به همسر فداکار و فرزند عزیزم که همواره حامی و مشوقم بودند، تقدیم می‌کنم.

صفحه	عنوان
.....	سخن مترجم
.....	پیشگفتار
فصل اول: اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین	
۱	۱-۱- میانگین.....
۲	۱-۲- انحراف معیار.....
۳	۱-۳- خطای معیار میانگین.....
۴	۱-۴- اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین.....
۵	۱-۴-۱- استفاده از دستورهای Fill/Series/Columns
۶	۱-۴-۲- تغییر عرض یک ستون.....
۷	۱-۴-۳- انتقال اطلاعات به مرکز بازه‌ای از سلول‌ها.....
۸	۱-۴-۴- نام‌گذاری بازه‌ای از سلول‌ها.....
۹	۱-۴-۵- به دست آوردن اندازه نمونه با تابع =COUNT
۱۰	۱-۴-۶- یافتن مقدار میانگین با تابع =AVERAGE
۱۱	۱-۴-۷- یافتن انحراف معیار با تابع =STDEV
۱۲	۱-۴-۸- یافتن خطای معیار میانگین.....
۱۳	۱-۴-۸-۱- قالب‌بندی اعداد در Number Format (دو رقم اعشار)
۱۴	۱-۴-۸-۲- ذخیره کردن صفحه گسترده
۱۵	۱-۴-۸-۳- چاپ یک صفحه گسترده
۱۶	۱-۴-۸-۴- قالب‌بندی اعداد در حالت پولی (دو رقم اعشار)
۱۷	۱-۴-۸-۵- تغییر قالب اعداد در Number Format (سه رقم اعشار)
۱۸	۱-۴-۸-۶- تمرین‌های عملی پایان فصل
۱۹	۱-۴-۸-۷- منابع

فصل دوم: تولید عدد تصادفی

۳-۲- مرتب کردن اعداد چارچوب تکراری به صورت یک توالی تصادفی ۳۶
۴-۲- چاپ فایل اکسل به طوری که تمامی اطلاعات آن در یک صفحه جای گیرد. ۴۰
۵-۲- تمرین های عملی پایان فصل ۴۵
منابع ۴۸

فصل سوم: فاصله اطمینان میانگین با استفاده ازتابع $TINV$ و آزمون فرضیه

۱-۳- فاصله اطمینان میانگین ۵۰
۱-۱-۱- چگونه می توان میانگین جامعه را برآورد کرد ۵۰
۱-۱-۲- برآورد حد بالای فاصله اطمینان ۹۵ درصد میانگین ۵۰
۱-۱-۳- برآورد فاصله اطمینان برای مایل طی شده به ازای هر گالن در خودرو چوی ایمپala ۵۲
۱-۱-۴- عدد ۱/۹۶ از کجا آمده است؟ ۵۳
۱-۱-۵- پیدا کردن مقدار t در رابطه فاصله اطمینان ۵۴
۱-۱-۶- استفاده ازتابع $TINV$ اکسل برای پیدا کردن فاصله اطمینان میانگین ۵۵
۱-۱-۷- استفاده از اکسل برای پیدا کردن فاصله اطمینان ۹۵٪ برای ادعای mpg یک خودرو ۵۶
۲-۳- آزمون فرضیه ۶۲
۲-۱-۲-۳- فرضیه ها همواره به جامعه افراد یا رخدادهایی که مطالعه می کنید، مربوط می شوند. ۶۳
۲-۲-۳- فرضیه صفر و فرضیه تحقیق (جای گزین) ۶۴
۲-۲-۳-۱- تعیین فرضیه صفر و فرضیه تحقیق هنگام استفاده از مقیاس های نمره دهی ۶۵
۲-۲-۳-۲- هفت مرحله برای آزمون فرضیه با استفاده از فاصله اطمینان میانگین ۶۸
۲-۲-۳-۳- مرحله اول: فرضیه صفر و فرضیه تحقیق را بیان کنید ۶۸
۲-۲-۳-۴- مرحله دوم: آزمون آماری مناسب را انتخاب کنید ۶۸
۲-۲-۳-۵- مرحله سوم: رابطه آزمون آماری را محاسبه کنید ۶۸
۲-۲-۳-۶- مرحله چهارم: یک تصویر از فاصله اطمینان میانگین ۶۹
۲-۲-۳-۷- مرحله پنجم: در خصوص قاعده تصمیم، تصمیم گیری کنید ۶۹
۲-۲-۳-۸- مرحله ششم: نتایج آزمون آماری خود را بیان کنید ۶۹
۲-۲-۳-۹- مرحله هفتم: نتیجه گیری آزمون آماری را به زبان ساده بیان کنید! ۷۰

۳-۳- راههای جایگزین برای خلاصه کردن نتیجه یک آزمون فرضیه.....	۷۵
۱-۳-۳- روش‌های مختلف برای پذیرفتن فرضیه صفر.....	۷۵
۲-۳-۳- روش‌های مختلف برای رد فرضیه صفر.....	۷۶
۴-۳- تمرین‌های عملی پایان فصل	۷۷
منابع.....	۸۴

فصل چهارم: آزمون یک گروهی t برای میانگین

۱-۴- مراحل هفت‌گانه آزمون فرضیه با استفاده از آزمون یک گروهی t	۸۶
۱-۱-۱- مرحله اول: بیان کردن فرضیه صفر و فرضیه تحقیق.....	۸۶
۱-۱-۲- مرحله دوم: آزمون آماری مناسب را انتخاب کنید.....	۸۷
۱-۱-۳- مرحله سوم: تصمیم‌گیری در مورد یک قاعده تصمیم (نتیجه‌گیری) برای آزمون یک گروهی t	۸۷
۱-۱-۳-۱- پیدا کردن مقدار قدر مطلق یک عدد.....	۸۷
۱-۱-۴- مرحله چهارم: محاسبه رابطه آزمون t یک گروهی	۸۸
۱-۱-۵- مرحله پنجم: مقدار t بحرانی را از جدول t در ضمیمه E پیدا کنید.....	۸۹
۱-۱-۶- مرحله ششم: نتایج آزمون آماری خود را بیان کنید.....	۹۰
۱-۱-۷- مرحله هفتم: نتیجه آزمون آماری خود را به صورت ساده بیان کنید.....	۹۱
۱-۲- آزمون t یک گروهی برای میانگین	۹۱
۱-۳- زمانی که آزمون فرضیه انجام می‌شود، آیا می‌توان از فاصله اطمینان ۹۵٪ میانگین یا آزمون t یک گروهی، استفاده کرد؟.....	۹۸
۱-۴- تمرین‌های عملی پایان فصل	۹۸
منابع.....	۱۰۵

فصل پنجم: آزمون t دو گروهی اختلاف میانگین گروههای مستقل

۱-۵- نه مرحله برای آزمون فرضیه با استفاده از آزمون t دو گروهی.....	۱۰۷
۱-۱-۱- مرحله اول: یک گروه را گروه ۱ و دیگری را گروه ۲ نام‌گذاری کنید.....	۱۰۷
۱-۱-۲- مرحله دوم: یک جدول ایجاد کنید که اندازه نمونه، میانگین و انحراف معیار هر گروه را بهطور خلاصه نشان دهد.....	۱۰۸

۱-۵-۳- مرحله سوم: فرضیه صفر و فرضیه تحقیق برای آزمون t دو گروهی را بیان کنید ...	۱۱۰
۱-۵-۴- مرحله چهارم: انتخاب آزمون آماری مناسب.....	۱۱۰
۱-۵-۵- مرحله پنجم: تصمیم درباره قاعده تصمیم‌گیری برای آزمون t دو گروهی.....	۱۱۰
۱-۵-۶- مرحله ششم: محاسبه رابطه برای آزمون t دو گروهی.....	۱۱۰
۱-۵-۷- مرحله هفتم: مقدار t بحرانی را از جدول t در ضمیمه E پیدا کنید	۱۱۱
۱-۵-۸- پیدا کردن درجه آزادی (df) برای آزمون t دو گروهی	۱۱۱
۱-۵-۹- مرحله هشتم: نتیجه آزمون آماری خود را تفسیر کنید	۱۱۲
۱-۵-۱۰- مرحله نهم: نتیجه آزمون آماری خود را تفسیر کنید.....	۱۱۲
۱-۵-۱۱- نتیجه‌گیری آزمون t دو گروهی زمانی که فرضیه صفر پذیرفته می‌شود.....	۱۱۲
۱-۵-۱۲- نتیجه‌گیری آزمون t دو گروهی زمانی که فرضیه صفر رد و فرضیه تحقیق پذیرفته می‌شود.....	۱۱۴
۱-۵-۱۳- رابطه شماره ۱: هر دو گروه دارای جمعیتی بیش از ۳۰ نفر باشند.....	۱۱۷
۱-۵-۱۴- مثالی از رابطه شماره ۱ برای آزمون t دو گروهی.....	۱۱۹
۱-۵-۱۵- رابطه شماره ۲: یکی یا هر دو گروه دارای جمعیتی کمتر از ۳۰ نفر باشند.....	۱۲۶
۱-۵-۱۶- تمرین‌های عملی پایان فصل	۱۳۴
۱-۵-۱۷- منابع.....	۱۳۹

فصل ششم: همبستگی و رگرسیون خطی ساده

۱-۶- همبستگی چیست؟	۱۴۱
۱-۶-۱- درک رابطه محاسبه همبستگی	۱۴۹
۱-۶-۲- درک مراحل نه گانه محاسبه همبستگی	۱۵۰
۱-۶-۳- استفاده از اکسل برای محاسبه همبستگی بین دو متغیر	۱۵۳
۱-۶-۴- ایجاد نمودار و رسم خط رگرسیون بر روی نمودار	۱۵۹
۱-۶-۵- استفاده از اکسل برای ایجاد نمودار و رسم خط رگرسیون از طریق نقاط داده‌ها	۱۶۰
۱-۶-۶- رسم خط رگرسیون از طریق نقاط داده‌ها در نمودار	۱۶۶
۱-۶-۷- جابجایی نمودار در زیر جدول صفحه گستردگی	۱۶۹
۱-۶-۸- ساخت نمودار "طویل‌تر" به‌طوری که بلندتر باشد	۱۷۰

۱۷۱.....	۶-۳-۱-۴- ساخت نمودار "پهن تر"
۱۷۲.....	۶-۴- چاپ کردن صفحه گسترده به طوری که جدول و نمودار در یک صفحه باشند
۱۷۴.....	۶-۵- یافتن معادله رگرسیون
۱۷۴.....	۶-۵-۱- نصب "بسته ابزار تجزیه و تحلیل داده" بر روی اکسل
۱۷۴.....	۶-۵-۱-۱- نصب بسته ابزار تجزیه و تحلیل داده بر روی اکسل ۲۰۱۳
۱۷۵.....	۶-۵-۲- نصب بسته ابزار تجزیه و تحلیل داده بر روی اکسل ۲۰۱۰
۱۷۶.....	۶-۵-۳- نصب بسته ابزار تجزیه و تحلیل داده بر روی اکسل ۲۰۰۷
۱۷۷.....	۶-۵-۴- نصب بسته ابزار تجزیه و تحلیل داده بر روی اکسل ۲۰۰۳
۱۷۸.....	۶-۵-۲- استفاده از اکسل برای پیدا کردن خلاصه خروجی رگرسیون
۱۸۱.....	۶-۴-۱- یافتن عرض از مبداء y ، برای خط رگرسیون
۱۸۱.....	۶-۴-۲- پیدا کردن شیب، a ، برای خط رگرسیون
۱۸۲.....	۶-۴-۳- نوشتن معادله برای خط رگرسیون
۱۸۲.....	۶-۴-۴- استفاده از خط رگرسیون برای پیش‌بینی ارزش y به ازای ارزش x داده شده
۱۸۳.....	۶-۴-۵- اضافه کردن معادله رگرسیون به نمودار
۱۸۷.....	۶-۴-۶- نحوه تشخیص همبستگی منفی در جدول خروجی خلاصه
۱۸۷.....	۶-۴-۷- چاپ تنها بخشی از صفحه گسترده به جای تمام آن
۱۸۷.....	۶-۴-۸- چاپ تنها جدول و نمودار بر روی صفحه‌ای جداگانه
۱۸۸.....	۶-۴-۹- چاپ تنها نمودار بر روی صفحه‌ای جداگانه
۱۸۹.....	۶-۴-۱۰- چاپ تنها خروجی خلاصه تحلیل رگرسیون بر روی صفحه‌ای جداگانه
۱۸۹.....	۶-۴-۱۱- تمرین‌های عملی پایان فصل
۱۹۶.....	۶-۴-۱۲- منابع

فصل هفتم: همبستگی و رگرسیون چندگانه

۱۹۸.....	۷-۱- معادله رگرسیون چندگانه
۲۰۱.....	۷-۲- پیدا کردن همبستگی چندگانه و معادله رگرسیون چندگانه
۲۰۵.....	۷-۳- استفاده از معادله رگرسیون برای پیش‌بینی فروش سالانه
۲۰۵.....	۷-۴- استفاده از اکسل برای ایجاد یک ماتریس همبستگی در رگرسیون چندگانه

۵-۷- تمرین‌های عملی پایان فصل ۲۰۹
منابع ۲۱۶

فصل هشتم: تجزیه واریانس یک‌طرفه (ANOVA)

۱-۸- استفاده از اکسل بهمنظور انجام تجزیه واریانس یک‌طرفه (ANOVA) ۲۲۰
۲-۸- چگونه جدول ANOVA را بهدرستی تفسیر کنیم ۲۲۳
۳-۸- استفاده از قاعدة تصمیم برای آزمون F در ANOVA ۲۲۴
۴-۸- آزمون اختلاف بین دو گروه با آزمون t در ANOVA ۲۲۵
۱-۴-۸- مقایسه قیمت دائرگ در مقابل شاپ انسیو با آزمون t در ANOVA ۲۲۵
۱-۱-۴-۸- تعیین درجه آزادی آزمون t در ANOVA ۲۲۶
۲-۱-۴-۸- بیان توضیح قاعده تصمیم برای آزمون t در ANOVA ۲۲۷
۳-۱-۴-۸- انجام آزمون t در ANOVA با استفاده از دستورهای اکسل ۲۲۷
۵-۸- تمرین‌های عملی پایان فصل ۲۳۰
منابع ۲۳۹

پیوست‌ها

پیوست A: پاسخ تمرین‌های عملی پایان هر فصل ۲۴۲
پیوست B: آزمون عملی ۲۷۶
پیوست C: پاسخ آزمون عملی ۲۹۰
پیوست D: رابطه‌های آماری ۳۰۲
پیوست E: جدول t ۳۰۴

پیشگفتار

کتاب "اکسل ۲۰۱۳ برای آمار کسب و کار؛ راهنمایی برای حل مسائل کاربردی" مناسب افرادی است که به دنبال یادگیری اصول استفاده از ابزارهای آماری قدرتمند اکسل در دوره‌های آموزشی کسب و کار یا فعالیت‌های حرفه‌ای هستند. اگر در ک آمار نقطه قوت شما نیست یا شما به طور خاص علاقه‌ای به ریاضیات ندارید یا اگر در کار با رایانه محتاط هستید، در این صورت، کتاب حاضر برای شما مناسب است.

در این کتاب خواهید آموخت که چگونه از آزمون‌های آماری کلیدی در اکسل استفاده کنید بدون اینکه بیش از حد تحت تأثیر تئوری آماری قرار بگیرید. این کتاب به طور واضح و با روش‌شناسی مناسب نشان می‌دهد که چگونه می‌توانید برای حل مسائل عملی در کسب و کار، آزمون‌های آماری را به وجود آورید و از آن‌ها استفاده کنید.

اکسل برنامه رایانه‌ای است که به راحتی در دسترس دانشجویان، مدرسین و مدیران می‌باشد. همچنین، این برنامه یک ابزار آموزشی و یادگیری مؤثر برای تجزیه و تحلیل کمی در دوره‌های کسب و کار است. وجود توانایی محاسبات عددی قدرتمند و توابع گرافیکی در اکسل، یادگیری آمار را در قیاس با سال‌های گذشته بسیار آسان تر کرده است. اگرچه، این اولین کتابی است که قابلیت‌های اکسل برای آموزش مؤثرتر آمار کسب و کار را نشان می‌دهد؛ اما به صورت انحصاری بر روی این موضوع متمرکز شده تا نه تنها تلاشی برای ارائه کاربردی و عملیاتی آمار باشد بلکه درک و استفاده آسان از آن را نیز در بر داشته باشد.

ویژگی‌های منحصر به فرد این کتاب به قرار زیر است:

- گام به گام، طریقه انجام کار به شما گفته می‌شود، نه تنها چگونگی استفاده از اکسل بلکه چرایی انجام هر گام آموزش داده می‌شود تا اینکه آنچه انجام می‌دهید را درک کنید و نه اینکه صرفاً نحوه استفاده از آزمون‌های آماری را یاد بگیرید.
- ذکر اهداف برای هر مفهوم اساسی در متن کتاب، تا بتوانید هدف از اجرای مراحل مختلف در اکسل را درک کنید.
- ارائه ۱۶۷ عکس رنگی در متن کتاب تا اطمینان یابید مراحل اکسل را به درستی اجرا می‌کنید.
- این کتاب ابزاری است که می‌تواند به تنها یکی یا به همراه هر کتاب خوب آماری مورد استفاده قرار گیرد.
- مثال‌های عملی و مسائل از دنیای کسب و کار اقتباس شده است.
- مبانی نظری و رابطه‌های آماری به زبان ساده بدون اینکه در نکته‌های ظرفی ریاضی غرق شوید، توضیح داده می‌شود.

- چگونگی نوشتمن رابطه‌های آماری در اکسل و استفاده از لیست‌های کشویی اکسل که روابط را ایجاد می‌کنند، خواهید آموخت.
- این کتاب همراه با سی دی فایل‌های اکسل که شما بتوانید در رایانه خود آپلود کنید، نیست. در عوض، به شما نشان خواهد داد تا چگونه هر فایل اکسل را خودتان ایجاد کنید. در محیط کاری، همکاران شما فایل اکسل را در اختیار شما قرار نمی‌دهند؛ از شما انتظار می‌رود که خودتان فایل اکسل را ایجاد کنید. کتاب حاضر برای توسعه این مهارت تمرين‌های زیادی را فراهم آورده است.
- هر فصل گام‌های مورد نیاز برای حل یک مسئله عملی کسب و کار را با استفاده از اکسل ارائه می‌دهد. به علاوه، در پایان هر فصل سه مسئله عملی وجود دارد، بنابراین شما می‌توانید دانش آماری خود را امتحان کنید. پاسخ به این مسائل عملی کسب و کار در پیوست A ارائه شده است.
- به منظور ارزیابی دانش شما یک "آزمون عملی" در پیوست B در پایان کتاب ارائه شده است. پاسخ به این مسائل عملی کسب و کار در پیوست C موجود است.

این کتاب برای استفاده در هر دوره آموزشی آمار کسب و کار (سطوح کارشناسی و تحصیلات تكمیلی) و همچنین برای مدیرانی که مایل به بهبود فواید مهارت‌های اکسل خود هستند، مناسب است.

پروفسور Tom Quirk، استاد فعلی بازاریابی در دانشکده تجارت و فناوری George Herbert Walker در دانشگاه Missouri St. Louis در Webster (ایالات متحده آمریکا) است و آمار بازاریابی، تحقیقات بازاریابی و راهبردهای قیمت‌گذاری را تدریس می‌کند. در آغاز حرفه آکادمیک خود، پروفسور Quirk شش سال در پژوهش‌های آموزشی در مؤسسات آمریکایی حوزه خدمات پژوهش و آزمون آموزشی صرف کرد. وی مقالاتی را در مجله روانشناسی آموزشی، مجله تحقیقات آموزشی، مرور تحقیقات آموزشی، مجله سنجش آموزشی، فناوری آموزشی، مجله مدرسه ابتدایی، مجله آموزش متوسطه، افق‌های آموزش و Phi Delta Kappan منتشر کرد. به علاوه، پروفسور Quirk بیش از ۶۰ کتاب کمک درسی در حوزه مدیریت و بازاریابی نوشته است، بیش از ۲۰ مقاله در مجلات تخصصی منتشر کرده است و بیش از ۲۰ مقاله در جلسات حرفه‌ای شامل انجمن تحقیقات آموزشی آمریکا، انجمن روان‌شناسی آمریکا، و شورای ملی سنجش آموزشی، ارائه کرده است. او دارای مدرک کارشناسی ریاضی از دانشگاه John Carroll، کارشناسی ارشد آموزش و دکتری تخصصی روان‌شناسی آموزشی از دانشگاه Stanford و کارشناسی ارشد مدیریت کسب و کار از دانشگاه Missouri-St. Louis است.

St. Louis, MO, USA
Thomas J. Quirk

فصل اول

اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین

این فصل به بررسی روش استفاده از اکسل برای به دست آوردن مقدار متوسط (یعنی "میانگین") مجموعه‌ای از اعداد، انحراف معیار (STDEV) و خطای معیار میانگین (S.E.) این اعداد می‌پردازد. هر سه این موارد آماری اغلب به کار برده می‌شوند و مبنای آزمون‌های آماری بیشتری را در کسب و کار تشکیل می‌دهند.

۱-۱- میانگین

میانگین «مقدار متوسط حسابی» مجموعه‌ای از اعداد است. وقتی دخترم کلاس پنجم بود، با چهره‌ای غمگین از مدرسه به خانه آمد و گفت که نتوانسته مقدار متوسط را به دست آورد. کتابی که از آن استفاده می‌کرد نحوه محاسبه میانگین مجموعه‌ای از اعداد را توضیح می‌داد، بنابراین به او گفتم:

«جنیفر، تمام اعداد را باهم جمع کن و بر تعداد اعدادی که داری، تقسیم کن.»
نگاهی به من انداخت و گفت: "پدر، موضوع جدی است!" فکر کرد قصد آزار او را دارد. بنابراین گفتم:

"این اعداد را در کتاب خود ببین، آنها را باهم جمع کن. جواب چیست؟" (پاسخ را گفت)
"حالا چند عدد داری؟" (پاسخ سؤال را گفت)
"حاصل جمع اعداد را بر تعداد آنها تقسیم کن."

این کار را کرد و پاسخ صحیح را به دست آورد. حالا شما هم همین استدلال را به کار خواهید برد، هرچند این کار برای شما بسیار راحت‌تر خواهد بود زیرا اکسل تمام این مراحل را برایتان انجام می‌دهد.

مقدار متوسط نمرات را "میانگین" می‌نامیم که نماد آن: \bar{X} بوده و "ایکس بار" تلفظ می‌نماییم.
رابطه محاسبه میانگین با ماشین حساب به صورت زیر است:

$$\bar{X} = \frac{\Sigma X}{n} \quad (1-1)$$

نماد Σ حرف یونانی سیگما به مفهوم "حاصل جمع" است و بیانگر حاصل جمع تمامی اعدادی است که با حرف X مشخص شده‌اند. سپس حاصل جمع را بر n (تعداد اعدادی که دارید) تقسیم کنید.

یک مثال ساده: فرض کنید که این شش عدد را دارید:

۶
۴
۵
۳
۲
۵

برای به دست آوردن میانگین این اعداد، در ابتدا آنها را با هم جمع کنید؛ سپس بر تعداد اعداد تقسیم نمایید. بنابراین، میانگین آن‌ها برابر است با:

$$\frac{25}{6} = 4.17$$

۲-۱- انحراف معیار

انحراف معیار "میزان نزدیکی اعداد به میانگین" را بیان می‌کند. اگر انحراف معیار عددی کوچک باشد بیانگر آن است که اعداد "به صورت دسته شده" و نزدیک به میانگین هستند. اگر انحراف معیار عددی بزرگ باشد؛ بیانگر آن است که اعداد "به صورت پراکنده" بوده و فاصله زیادی از میانگین دارند. رابطه انحراف معیار (STDEV) که با حرف S نشان داده می‌شود، عبارت است از:

$$STDEV = S = \sqrt{\frac{\sum(X - \bar{X})^2}{n-1}} \quad (2-1)$$

رابطه بالا پیچیده به نظر می‌رسد اما آنچه این رابطه از شما می‌خواهد به قرار زیر است:

۱. میانگین را از تک‌تک اعداد کم کنید $(\bar{X} - X)$.
۲. سپس عدد حاصل را به توان دو برسانید تا تبدیل به عددی مثبت شود.
۳. مربع اعداد را با هم جمع نمایید تا حاصل جمع کل به دست آید.

۴. سپس، حاصل جمع کل را بر $n-1$ تقسیم نمایید (n نشان‌دهنده تعداد اعدادی است که داریم).

۵. در گام آخر، از عدد به دست آمده در مرحله چهار جذر بگیرید.

در این کتاب از شما خواسته نمی‌شود که انحراف معیار را با کمک ماشین حساب به دست آورید، اما مثال‌هایی از نحوه محاسبه آن، در هر کتاب پایه آماری، می‌توانید مشاهده کنید. در عوض، ما از اکسل برای به دست آوردن انحراف معیار مجموعه‌ای از اعداد استفاده می‌کنیم. زمانی که از اکسل برای محاسبه انحراف معیار شش عدد مورد استفاده در مثال بخش میانگین استفاده شود، متوجه می‌شویم که $STDEV$ این اعداد، S ، برابر با $1/47$ است.

۱-۳-۱- خطای معیار میانگین

رابطه خطای معیار میانگین ($s.e.$) که نماد آن $S_{\bar{X}}$ است) به صورت زیر می‌باشد:

$$s.e. = S_{\bar{X}} = \frac{S}{\sqrt{n}} \quad (3-1)$$

برای به دست آوردن خطای معیار میانگین، کافی است انحراف معیار $STDEV$ را بر جذر n که همان "تعداد اعداد" مجموعه داده‌ها است، تقسیم نمایید. در مثال توصیف انحراف معیار که در بالا آمد، خطای معیار میانگین برابر با $1/47$ خواهد بود (می‌توانید این مقدار را با ماشین حساب بررسی نمایید).

در صورت تمایل به فراگیری بیشتر در خصوص انحراف معیار و خطای معیار میانگین به وايرز^۱ (۲۰۱۱) مراجعه نمایید.

اکنون چگونگی استفاده از اکسل جهت محاسبه اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین را با استفاده از یک مسئله فروش بیاموزیم:

فرض کنید می‌خواهید فروش سال اول محصول جدیدی که شرکت شما قصد عرضه آن به بازار را دارد را برآورد نمایید. ابتدا به فروش سال اول محصولات مشابهی که شرکت شما قبلًا عرضه نموده، نگاهی بیاندازید، تا ایده‌ای از خصوصیات فروش محصولاتی که اولین بار به بازار عرضه شده را برای عرضه محصول جدید به دست آورید.

^۱ Weiers

از فروش سال اول محصول مشابه طی هشت سال اخیر بهره برده و جدول زیر را مشابه شکل ۱-۱ تشکیل می‌دهید:

سال	فروش سال اول (هزار دلار)
Year	First-year sales (\$000)
۱	۱۰
۲	۱۰
۳	۱۲
۴	۱۶
۵	۲۲
۶	۲۹
۷	۳۹
۸	۴۷

شکل ۱-۱: داده صفحه کاری برای فروش سال اول (مثال عملی)

دقت کنید که اعداد فروش سال اول به هزار دلار (\$000) است. بنابراین، عدد ۱۰ بیانگر فروش ۱۰۰۰۰ دلاری می‌باشد.

۱-۴-۱- اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین

هدف: یافتن اندازه نمونه (n)، میانگین، انحراف معیار (STDEV) و خطای معیار میانگین (s.e.)
این داده‌ها

رایانه خود را روشن نموده و روی نماد اکسل ۲۰۱۳ کلیک نمایید تا یک صفحه گسترده خالی اکسل باز شود.

داده‌ها را به این شکل وارد نمایید:

سال : A3

فروش سال اول (هزار دلار) : B3

۱ : A4

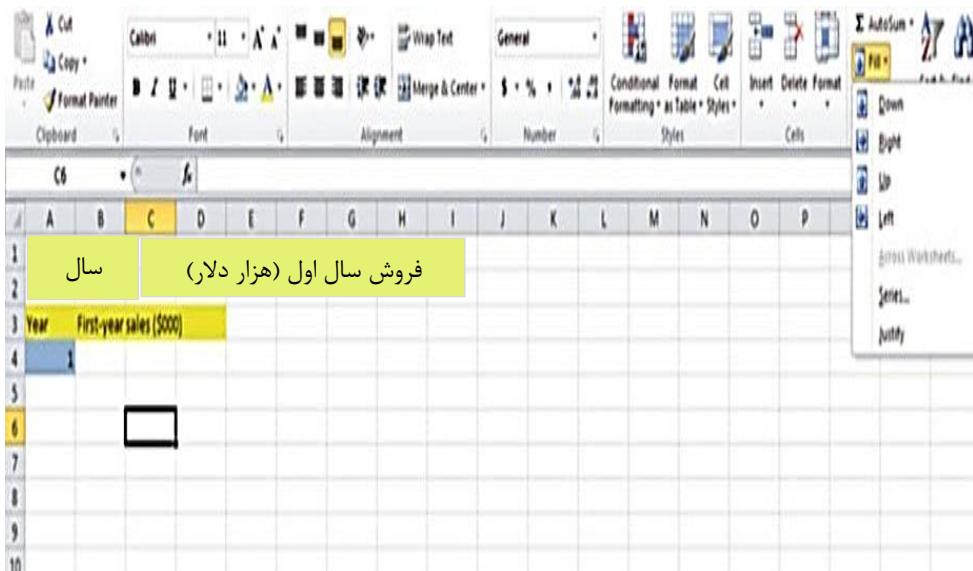
۴-۱- استفاده از دستورهای Fill/Series/Columns

هدف: افزودن سال‌های ۲-۸ در یک ستون در زیر سال اول

نشانگر را روی A4 قرار دهید

(بالا سمت چپ صفحه) Home

(بالا سمت راست: روی فلش پایین کلیک نمایید؛ به شکل ۲-۱ مراجعه شود)



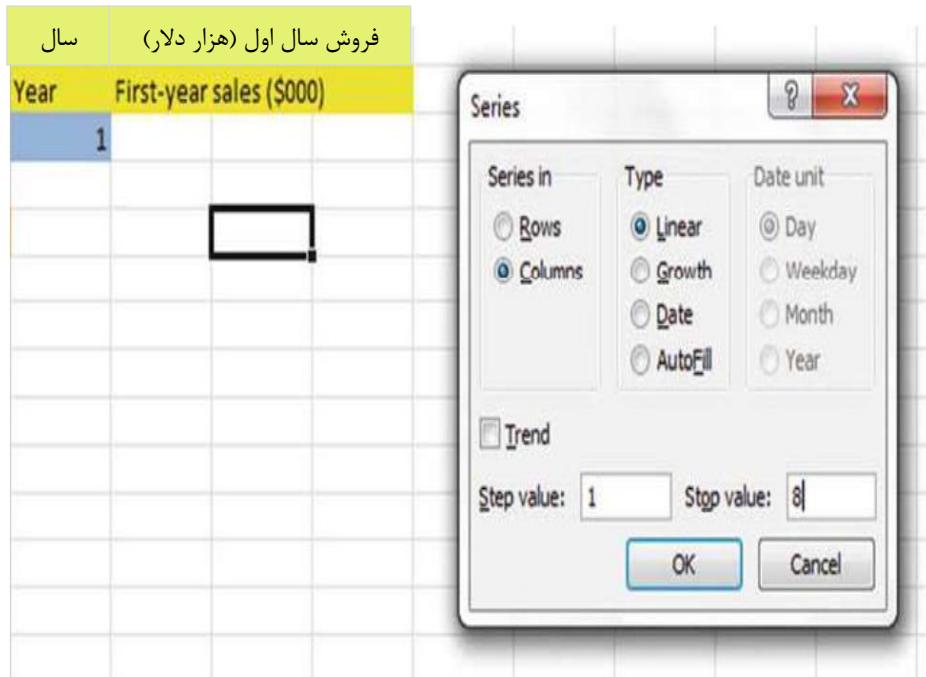
شکل ۱-۲: دستورهای Fill/Series/Columns

Series

Columns

۱ : Step value

۸ (به شکل ۱-۳ مراجعه شود) : Stop Value



شکل ۱-۳: مثال جعبه محاوره‌ای برای دستورهای Fill/Series/Columns/Step Value/Step Value

را انتخاب کنید

سال‌های ۱-۸ با عدد ۸ در سلول A11 مشخص می‌شوند.

حالا ارقام فروش سال اول را به کمک جدول بالا در سلول‌های B4:B11 وارد کنید.

از آنجایی که صفحه نمایش اطلاعات را به صورت حرفه‌ای نشان نمی‌دهد بایستی نحوه «عریض کردن پهنه‌ای ستون» و چگونگی «در مرکز سلول قرار دادن اطلاعات» در یک گروه از سلول‌ها را بیاموزید.

۲-۴-۱- تغییر عرض یک ستون

هدف: عریض کردن پهنه‌ای یک ستون به نحوی که تمامی اطلاعات در آن قرار گیرند

اگر به صفحه نمایش رایانه خود نگاه کنید، می‌بینید که ستون B به اندازه کافی عریض نیست تا تمام اطلاعات در آن جا گیرد. برای عریض کردن ستون B باید:

روی حرف B در بالای صفحه نمایش رایانه کلیک نمایید.

نشانگر ماوس را روی منتهی‌الیه سمت راست B قرار دهید تا یک «علامت متقطع» در گوشه آن ایجاد شود.

کلیک چپ ماوس را انتخاب نموده، نگه دارید و گوشه آن را به سمت راست حرکت دهید تا جایی که ستون به اندازه کافی عرض شده تا تمام داده‌ها جا شوند.

برای تنظیم عرض جدید ستون، انگشت خود را از ماوس بردارید (به شکل ۴-۱ مراجعه شود)

A	B	C
سال	فروش سال اول (هزار دلار)	
Year	First-year sales (\$000)	
1	10	
2	10	
3	12	
4	16	
5	22	
6	29	
7	39	
8	47	

شکل ۴-۱: مثالی از چگونگی عرض کردن پهنای ستون

سپس، روی هر سلول خالی کلیک کنید تا ستون B از حالت انتخاب خارج شود، بهنحوی‌که رنگ تیره‌تر آن از صفحه نمایش حذف شود.

هنگامی که یک ستون را عرض می‌کنید، تمام سلول‌ها را در تمام ردیف‌های آن به همان اندازه عرض می‌کنید.

اکنون وارد مراحل انتقال اطلاعات به مرکز سلول در هر دو ستون A و B می‌شویم.

۴-۳-۱- انتقال اطلاعات به مرکز بازه‌ای از سلول‌ها

هدف: انتقال دادن اطلاعات به مرکز گروهی از سلول‌ها

در راستای تخصصی‌تر نمودن اطلاعات در سلول‌ها، می‌توانید طی مراحل زیر، اطلاعات را به مرکز سلول انتقال دهید:

کلیک چپ ماوس را نگه‌داشته و به سمت راست و پایین بکشید تا سلول‌های A3:B11 انتخاب (هایلایت) شوند به طوری که این سلول‌ها تیره‌تر به نظر آیند.

در بالای صفحه‌نمایش، مجموعه‌ای از خطوط را می‌بینید که تمامی آن‌ها به همان عرض در زیر Alignment box انتقال یافته‌اند (دومین نماد در پایین سمت چپ Alignment box؛ به شکل ۵-۱ مراجعه شود)

Year	First-year sales (\$000)
1	10
2	10
3	12
4	16
5	22
6	29
7	39
8	47

شکل ۱-۵: مثالی از چگونگی انتقال اطلاعات به مرکز سلول‌ها

برای انتقال اطلاعات به مرکز سلول‌های انتخاب شده، روی این نماد کلیک کنید (به شکل ۱-۶). مراجعه شود).

A	B	C	D	E	F
سال	فروش سال اول (هزار دلار)				
Year	First-year sales (\$000)				
1	10				
2	10				
3	12				
4	16				
5	22				
6	29				
7	39				
8	47				

شکل ۱-۶: نتیجه نهایی انتقال اطلاعات به مرکز سلول‌ها

باتوجه به اینکه نیاز به رجوع به ارقام فروش سال اول در رابطه‌ها دارد، به جای به خاطر سپردن محل دقیق سلول‌های (B4: B11) که اعداد در آن قرار گرفته، به راحتی نامی برای این بازه داده‌ها انتخاب نمایید. این گروه سلول‌ها را Product بنامید، اما می‌توانید هر نام دیگری را هم بر آن بگذارید.

۴-۴-۱- نام‌گذاری بازه‌ای از سلول‌ها

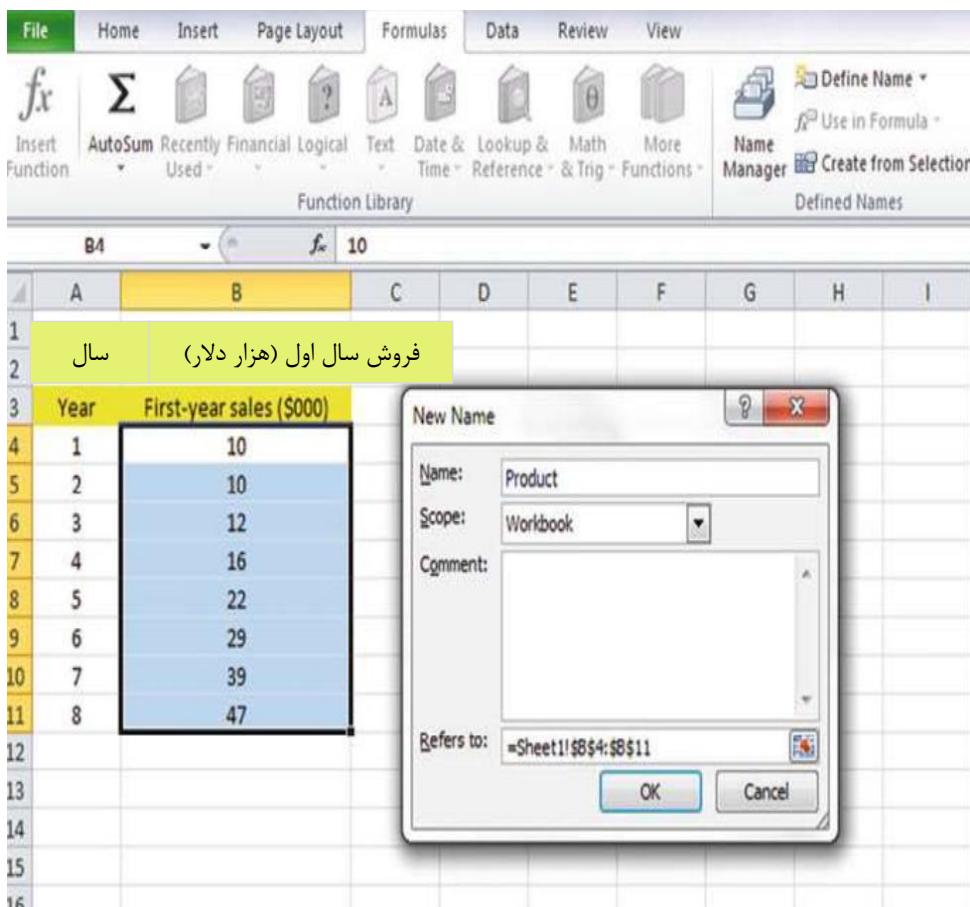
هدف: نام‌گذاری بازه‌ای از داده‌های ارقام فروش سال اول با عنوان: Product

با کلیک چپ ماوس روی B4 و کشیدن آن به سمت پایین تا B11 سلول‌های B4:B11 را انتخاب(هایلایت) نمایید.

بالا سمت چپ صفحه‌نمایش) را انتخاب کنید Formulas

(بالا وسط صفحه‌نمایش) را انتخاب کنید Define Name

(این نام را در جعبه بالایی تایپ کنید؛ شکل ۱-۷)



شکل ۱-۷: جعبه محاوره‌ای نام‌گذاری بازه سلول‌ها به نام Product

OK را انتخاب کنید

سپس روی یک سلول از صفحه گسترده که اطلاعاتی در آن موجود نیست (سلول خالی) کلیک نمایید تا حالت انتخاب حذف شود.

حالا موارد زیر را به صفحه گسترده اضافه کنید:

n : E6

Mean : E9

STDEV : E12

(به شکل ۸-۱ مراجعه شود) s.e. : E15

	A	B	C	D	E	F
1						
2	سال	فروش سال اول (هزار دلار)				
3	Year	First-year sales (\$000)				
4	1	10				
5	2	10				
6	3	12	تعداد	n		
7	4	16				
8	5	22	میانگین	Mean		
9	6	29				
10	7	39				
11	8	47	انحراف معیار	STDEV		
12						
13						
14						
15			خطای معیار	s.e.		
16						

شکل ۸-۱: مثال وارد کردن برچسب تعداد، میانگین، انحراف معیار و خطای معیار میانگین

توجه نمایید: هرگاه از یک رابطه استفاده می کنید، بایستی علامت مساوی (=) را در ابتدای نام تابع اضافه نمایید تا اکسل متوجه شود که قصد استفاده از رابطه را دارد.

=COUNT - به دست آوردن اندازه نمونه با تابع

هدف: به دست آوردن اندازه نمونه (n) داده ها با تابع =COUNT

=COUNT (Product) : F6

کلید اینتر را بزنید، از آنجایی که هشت رقم برای فروش سال اول داریم، خروجی دستور باید عدد هشت را وارد سلول F6 نماید.

=AVERAGE - یافتن مقدار میانگین با تابع

هدف: یافتن میانگین فروش با تابع =AVERAGE

=AVERAGE (Product) : F9

خروجی این دستور بایستی عدد 23.125 را وارد سلول F9 نماید.

=STDEV - یافتن انحراف معیار با تابع

هدف: یافتن انحراف معیار با تابع =STDEV

=STDEV(Product) : F12

خروجی این دستور بایستی عدد 14.02485 را وارد سلول F12 نماید.

=SQRT - یافتن خطای معیار میانگین

هدف: یافتن خطای معیار میانگین با رابطه برای این هشت داده

=F12/SQRT(8) :F15

خروجی این دستور باید عدد ۴/۹۵۸۵۳۳ F15 را وارد سلول ۹-۱ نماید (به شکل ۱-۹ مراجعه شود).

	A	B	C	D	E	F	G
1	سال	فروش سال اول (هزار دلار)					
3	Year	First-year sales (\$000)					
4	1	10					
5	2	10					
6	3	12	تعداد	n		8	
7	4	16					
8	5	22					
9	6	29	میانگین	Mean	23.125		
10	7	39					
11	8	47					
12			انحراف معیار	STDEV	14.02485		
13							
14							
15			خطای معیار	s.e.	4.958533		
۱۶							

شکل ۱-۹: مثالی از کاربرد توابع و رابطه‌های اکسل برای تعیین اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین

نکته مهم: برای اطمینان در این کتاب تمام رقم‌ها را در صفحه گستردگی دو بار کنترل نمایید تا خاطرجمع شوید که در سلول صحیح خود قرار گرفته‌اند؛ در غیر این صورت رابطه‌ها به درستی عمل نخواهند کرد.

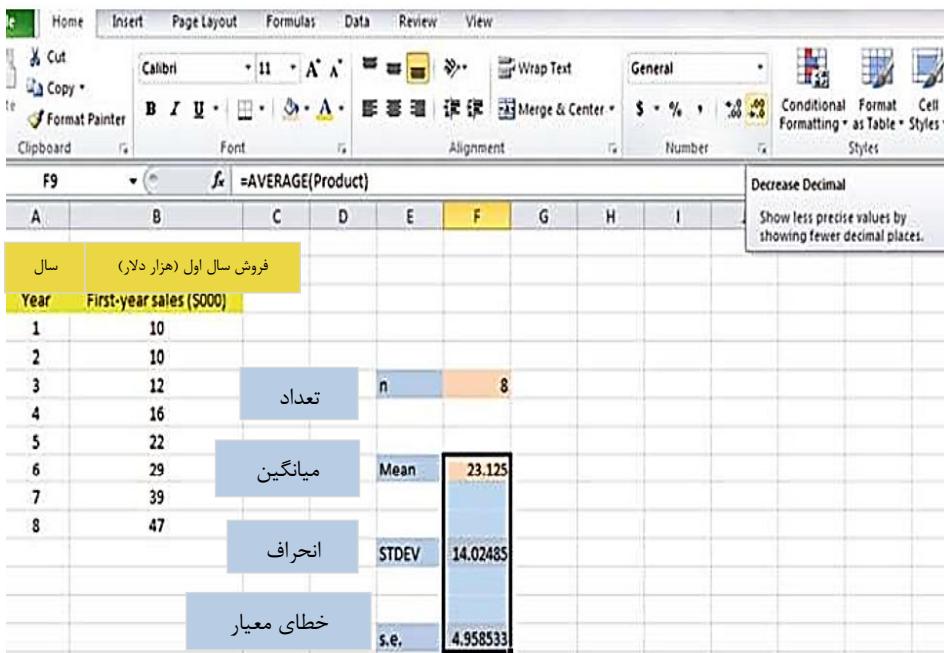
۱-۸-۴-۱- قالب‌بندی اعداد در Number Format (دو رقم اعشار)

هدف: تبدیل میانگین، انحراف معیار و خطای معیار میانگین به دو رقم اعشار

سلول‌های F9:F15 را انتخاب (هایلایت) نمایید

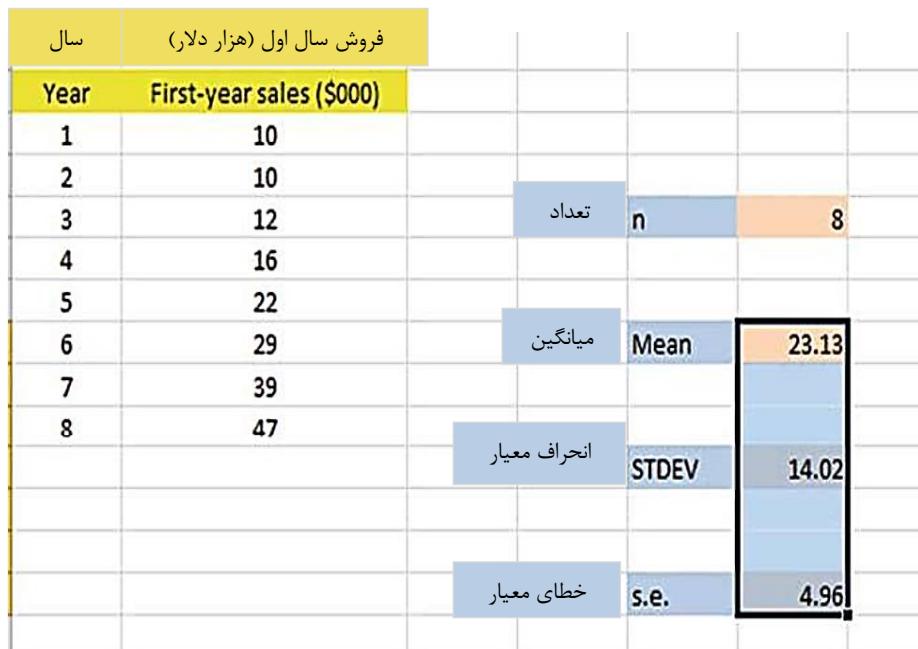
(بالا سمت چپ صفحه‌نمايش) را انتخاب نمایید Home

به زیر "Number" در بالا وسط صفحه نمایش بنگرید. در گوشه راست پایین، نشانگر ماوس را به آرامی در زیر ۰.۰۰.۰ قرار دهید تا نوشه "Decrease Decimal" را نشان دهد (به شکل ۱۰-۱ مراجعه شود).



شکل ۱۰-۱: استفاده از نماد "Decrease Decimal" جهت تبدیل اعداد به رقم با اعشار کمتر

یکبار روی این نماد کلیک نمایید و دقت کنید که اکنون اعداد سلول‌های F9:F15 دقیقاً به دو رقم اعشار است (شکل ۱۱-۱).



شکل ۱۱-۱: مثالی از تبدیل اعداد به دو رقم اعشار

نکته مهم: ارقام فروش به هزار دلار می‌باشند، بنابراین، میانگین 23130 دلار، انحراف معیار 14020 دلار و خطای معیار میانگین 4960 دلار می‌باشد.
اکنون بر روی یک "سلول خالی" در صفحه گسترده کلیک نمایید تا حالت انتخاب سلول‌های F9:F15 حذف گردد.

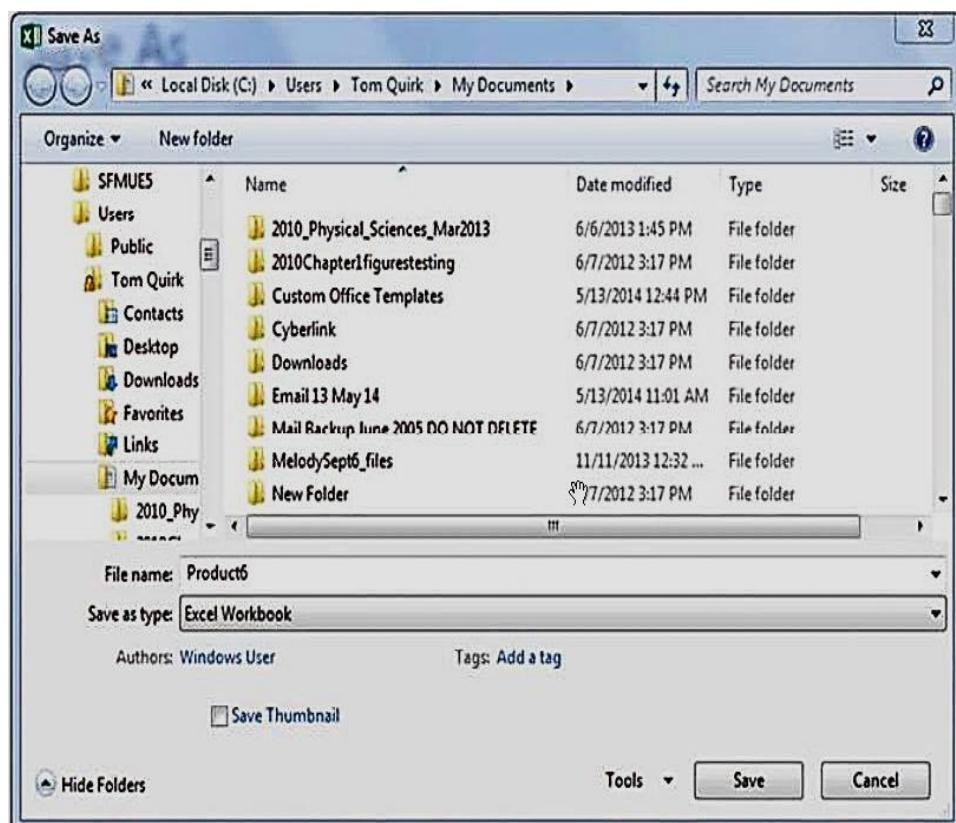
۱-۵-۱- ذخیره کردن صفحه گسترده

هدف: ذخیره کردن این صفحه گسترده با نام: Product6

برای ذخیره کردن صفحه گسترده به نحوی که بتوان آن را در آینده بازیابی کرد، ابتدا باید « محل » ذخیره کردن آن را انتخاب نمایید. تصمیم با شما است و چند گزینه دارد. اگر (هدف ذخیره در) رایانه شخصی‌تان باشد، می‌توانید آن را روی هارد دیسک ذخیره نمایید (نحوه این کار را سؤال نمایید). یا اینکه آن را روی لوح فشرده یا فلاش درایو ذخیره نمایید. سپس بایستی مراحل زیر را تکمیل نمایید:

بالای صفحه نمایش، نماد سمت چپ) را انتخاب کنید
Save as را انتخاب کنید

(محل ذخیره فایل را انتخاب نمایید. برای مثال: Computer: My Documents Location (این نام را در قسمت File name وارد نمایید؛ به شکل ۱۲-۱ مراجعه شود)



شکل ۱۲-۱: جعبه محاوره‌ای ذخیره کردن یک فایل Excel Workbook به نام Product6 در My Documents Location

Save (پایین سمت راست جعبه محاوره‌ای) را انتخاب کنید

نکته مهم: مراقب باشید تا صفحه گسترده فایل اکسل را هر چند دقیقه یکبار ذخیره نمایید تا اطلاعات خود را از دست ندهید.

۶-۶- چاپ یک صفحه گسترده

هدف: چاپ کردن یک صفحه گسترده

مراحل زیر را حین چاپ یک صفحه گسترده دنبال نمایید.

(بالای صفحه، نماد سمت چپ) را انتخاب کنید.

Print (به شکل ۱۳-۱ مراجعه شود) را انتخاب کنید.



شکل ۱۳-۱: مثالی از نحوه چاپ یک صفحه گسترده اکسل با استفاده از دستورهای File/Print صفحه گسترده نهایی در شکل ۱۴-۱ نشان داده شده است.

سال Year	فروش سال اول (هزار دلار) First-year sales (\$000)	تعداد n	میانگین Mean	انحراف معیار STDEV	خطای معیار S.E.
1	10				
2	10				
3	12				
4	16				
5	22				
6	29				
7	39				
8	47				

شکل ۱۴-۱: نتیجه نهایی چاپ یک صفحه گسترده اکسل

پیش از پایان این فصل تغییر قالب (فرمت) ارقام در یک صفحه گسترده را با دو مثال انجام دهید:

(۱) استفاده از دو رقم اعشار برای مقادیر دلار و (۲) استفاده از سه رقم اعشار برای اعداد.

صفحه گسترده را با دستور File/Save ذخیره کنید. سپس آن را با دستور File/Close ببندید و پس از آن یک صفحه خالی اکسل را با استفاده از نماد File/New/Blank Worksheet باز نمایید (بالا سمت چپ صفحه‌نمایش).

۷-۱- قالب‌بندی اعداد در حالت پولی (دو رقم اعشار)

هدف: تغییر ارقام با قالب دلار تا دو رقم اعشار

A3 : قیمت

1.25 : A4

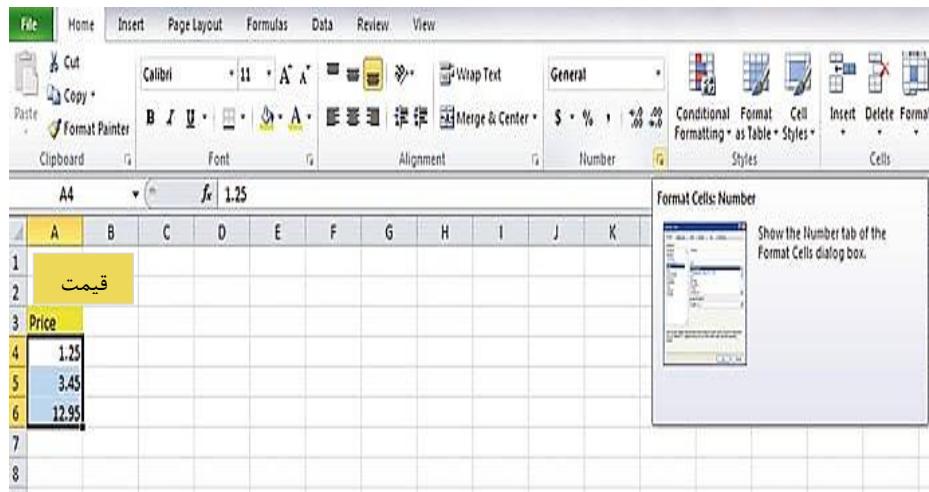
3.45 : A5

12.95 : A6

Home را انتخاب نمایید.

ابتدا با کلیک چپ ماوس روی سلول A4 و کشیدن آن به پایین، سلول های A4:A6 را انتخاب (هایلایت) نمایید، به نحوی که رنگ هر سه سلول تیره تر گردد.

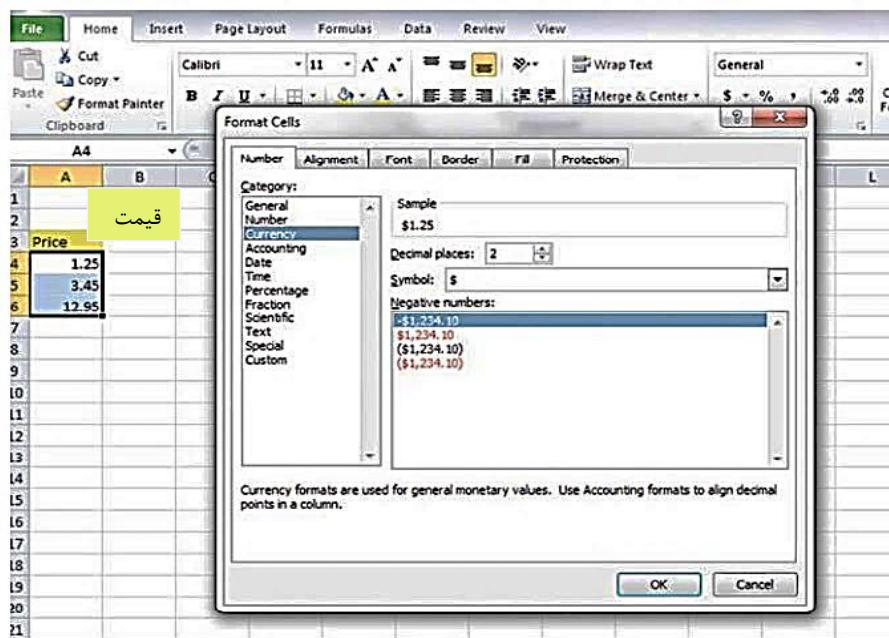
Number (بالای صفحه نمایش: روی فلاش پایین سمت راست کلیک نمایید؛ به شکل ۱۵-۱ مراجعه شود) را انتخاب کنید.



شکل ۱۵-۱: جعبه محاوره‌ای گزینه‌های قالب اعداد

Currency را انتخاب کنید.

Category ۲ را انتخاب کنید (به شکل ۱۶-۱ مراجعه شود).



شکل ۱۶-۱: جعبه محاوره‌ای قالب پولی اعداد (دو رقم اعشار)

را انتخاب کنید.

هر سه سلول دارای علامت دلار شد و تا دو رقم اعشار می‌باشند.

در مرحله بعد تغییر قالب اعداد به سه رقم اعشار را تمرین خواهید کرد.

۱-۸-۱- تغییر قالب اعداد در Number Format (سه رقم اعشار)

هدف: تغییر قالب‌بندی اعداد به سه رقم اعشار

را انتخاب کنید.

سلول‌های A4:A6 را روی صفحه نمایش انتخاب (هایلایت) نمایید.

روی فلش پایین سمت راست کلیک کنید) را انتخاب کنید.
Category : اعداد.

در سمت راست جعبه، با یکبار کلیک روی فلش بالا، دو رقم اعشار را به سه رقم اعشار تغییر دهید.

را انتخاب کنید.

هر سه عدد اکنون با سه رقم اعشار نمایش داده می‌شوند.

حالا روی یک سلول خالی کلیک کنید تا حالت انتخاب سلول‌های A4:A6 حذف شود. سپس، این فایل را با دستور File/Close/Don't Save بیندید (نیازی به ذخیره نیست چون این مراحل فقط برای تمرین بود).

می‌توانید از همین دستور برای تغییر قالب بازه‌ای از اعداد به درصد (و قالب‌های دیگر) به هر تعداد رقم اعشار استفاده نمایید.

۹-۱- تمرین‌های عملی پایان فصل

۱. فرض نمایید یک نمونه تصادفی از مشتریان هفته گذشته فروشگاه وال-مارت^۱ را انتخاب کرده‌اید.

جدول شکل ۱۷-۱ را تشکیل می‌دهید:

فروش به ازاء هر مشتری در هفته گذشته (بر حسب دلار)
DOLLAR SALES PER CUSTOMER LAST WEEK
127.12
140.45
104.64
80.06
114.07
109.35
117.28
72.84
67.67
79.85
109.96
117.13
85.25
149.36
147.57
153.54
118.76
69.86
154.47
154.88
109.44
97.36
87.55

شکل ۱۷-۱: داده‌های صفحه کاری برای فصل ۱: تمرین عملی شماره ۱

¹ Wal-Mart

(الف) برای بهدست آوردن میانگین، انحراف معیار و خطای معیار میانگین این داده‌ها مطابق جدول، از اکسل استفاده نمایید. جواب‌ها را نام‌گذاری کنید و میانگین، انحراف معیار و خطای معیار میانگین را تا دو رقم اعشار گرد نمایید؛ از قالب پولی برای این سه رقم استفاده کنید.

ب) نتیجه را روی صفحه جداولهای چاپ کنید.

ج) فایل را با نام WAL6 ذخیره نمایید.

۲. فرض کنید دایره منابع انسانی یک شرکت، پژوهشی جهت تعیین روحیه از تمامی مدیران میانی انجام داده و از شما خواسته تا نتایج این برآورد را خلاصه نمایید. قصد دارید مهارت‌های کار با اکسل را روی این مورد امتحان کنید تا معلوم شود که قادر به انجام صحیح این مأموریت هستید، و گزینه شماره ۲۱ را برای ارزیابی مهارت‌هایتان برمی‌گزینید. داده‌ها در شکل ۱۸-۱ آمده است.



شکل ۱۸-۱: داده‌های صفحه کاری فصل ۱: تمرین عملی شماره ۲

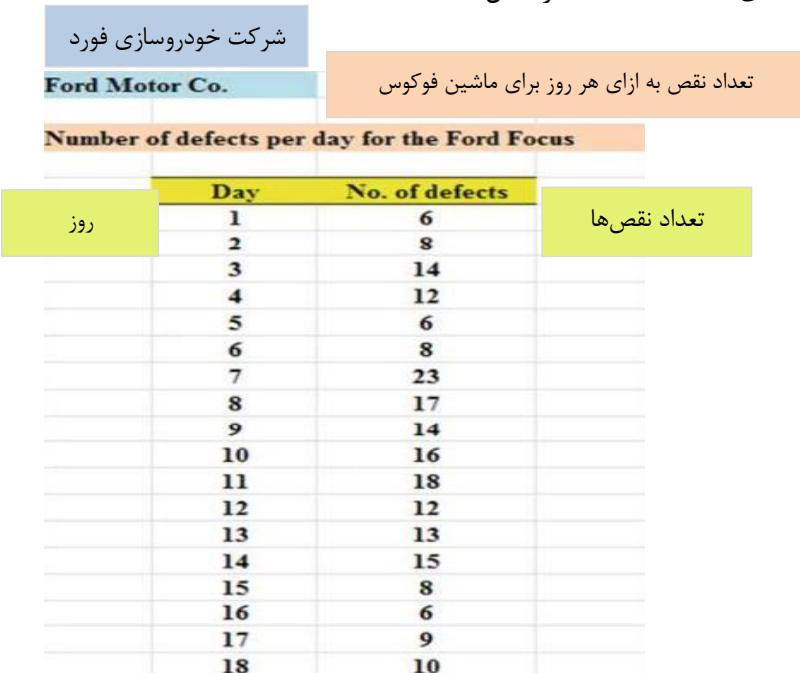
الف) از اکسل جهت ایجاد جدول رتبه‌بندی استفاده نمایید و با توجه به جدول از اکسل برای به دست آوردن اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین این داده‌ها استفاده کنید. جواب‌ها را نام‌گذاری نمایید و میانگین، انحراف معیار و خطای معیار میانگین را تا دو رقم اعشار گرد نمایید.

ب) نتیجه را در صفحه جداگانه‌ای چاپ کنید.

ج) فایل را با عنوان MORALE4 ذخیره کنید

۳. فرض کنید برای تحلیل داده‌های ۱۸ روز گذشته خودروسازی فورد^۱ که خودروهای مدل فوکوس^۲ را تولید می‌کند، استخدام شده‌اید.

مدیر کارخانه از شما خواسته تا خلاصه‌ای از بروز تعداد نواقص روزانه این خودرو را طی دوره سه هفته‌ای تهیه نمایید. منظور از نقص، هر نوع کارکرد نادرست خودرو در انتهای خط تولید است که خودرو را پیش از تحویل به فروشنده، از خط تولید خارج کرده و روانه تعمیر می‌کند. داده‌های سه هفته گذشته در شکل ۱۹-۱ آمده است:



شکل ۱-۹: داده‌های صفحه کاری فصل ۱: تمرین عملی شماره ۳

¹ Ford

² Focus

الف) از اکسل برای تشکیل جدول این داده‌ها استفاده نمایید و با توجه به جدول برای به دست آوردن اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین این داده‌ها از اکسل استفاده نمایید. جواب‌ها را نام‌گذاری کرده و میانگین، انحراف معیار و خطای معیار میانگین را تا سه رقم اعشار گرد نمایید.

ب) نتیجه را روی صفحه جداگانه‌ای چاپ کنید.

ج) فایل را با نام DEFECTS4 ذخیره کنید.

منابع

Weiers RM. Introduction to business statistics. 7th ed. Mason: South-Western Cengage Learning; 2011.

فصل دوم

تولید عدد تصادفی

فرض کنید می خواهید یک نمونه تصادفی شامل پنج نفر از ۳۲ فروشنده شرکت خود را با استفاده از اکسل انتخاب نمایید تا از این پنج فروشنده درباره رضایت شغلی آنها در شرکت مصاحبه کنید. برای انجام این کار، شما نیاز به تعریف چارچوب نمونه گیری خواهید داشت. چارچوب نمونه گیری لیستی از افراد است که می خواهید از آن یک نمونه تصادفی را انتخاب کنید. این چارچوب با کد شناسایی (ID) از شماره ۱ آغاز می شود که این کد به فروشنده اول در لیست شما از ۳۲ فروشنده در شرکت، تعلق می گیرد. به دو میان فروشنده کد شماره ۲، سومین نفر کد شماره ۳ و به همین ترتیب به آخرین فروشنده کد شماره ۳۲ تعلق می گیرد.

از آنجایی که شرکت شما دارای ۳۲ فروشنده می باشد، چارچوب نمونه گیری شما از ۱ تا ۳۲ به ازای هر فروشنده، دارای شماره شناسایی منحصر به فرد است.

ابتدا باید چارچوب اعداد را به صورت زیر در یک صفحه جدید اکسل ایجاد کنیم:

۱-۲- ایجاد چارچوب اعداد برای تولید اعداد تصادفی

هدف: ایجاد چارچوب اعداد برای تولید اعداد تصادفی

A3 : شماره چارچوب

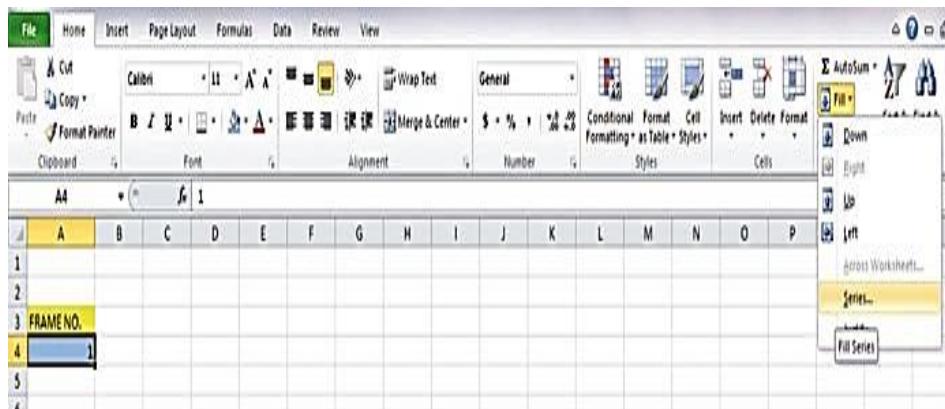
۱ : A4

اکنون، شماره های کادر را در ستون A از طریق گزینه HOME/FILL commands که در فصل اول این کتاب توضیح داده شد، ایجاد می کنیم (به بخش ۱-۴-۱ مراجعه شود) به طوری که اعداد چارچوب از ۱ تا ۳۲، با شماره ۳۲ در سلول A35 قرار می گیرند. اگر شما در مورد نحوه انجام این کار نیاز به یادآوری دارید، مراحل زیر را انجام دهید:

در سلول A4 کلیک کرده تا این سلول را انتخاب کنید.

گزینه HOME را انتخاب کنید

(روی گزینه Fill جهت به سمت پایین کنار این دستور را انتخاب کرده و series (سری ها) را مطابق شکل ۱-۲ انتخاب کنید).

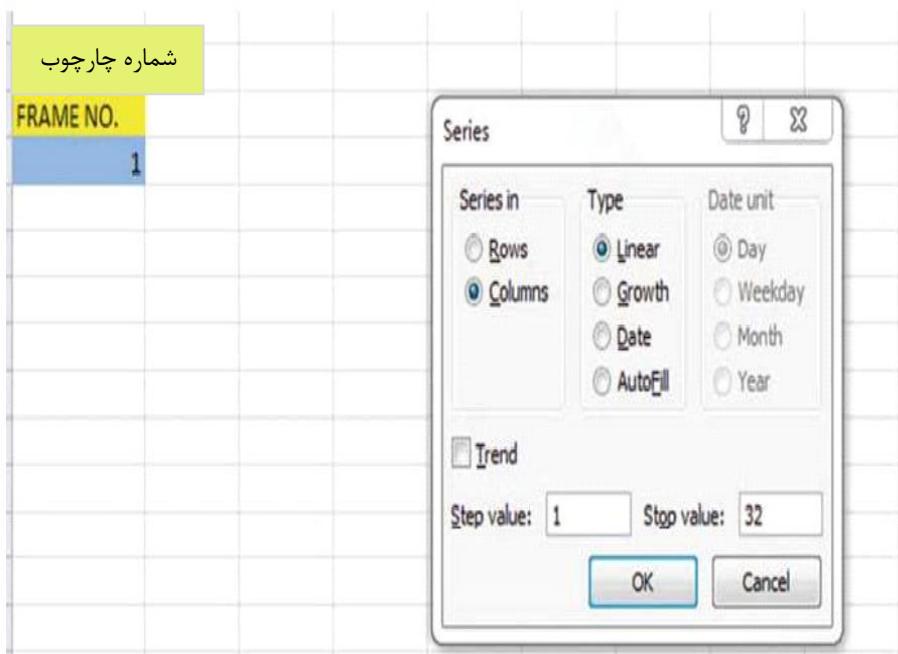


شکل ۲-۱: جعبه محاوره‌ای برای دستورهای Fill/series (پر کردن / سری‌ها)

انتخاب ستون‌ها در Series in

ارزش گام (Step value)

ارزش توقف (Stop value): شکل ۲-۲ مراجعه شود.



شکل ۲-۲: جعبه محاوره‌ای برای دستورهای پر کردن (دادن فرمان) / سری‌ها / ارزش گام / ارزش توقف (Fill/Series/Columns/Step value/Stop value)

را انتخاب کنید. Ok

سپس این فایل را با عنوان: RANDOM 2 ذخیره کنید. شما باید نتیجه را مطابق شکل ۳-۲ به دست آورید.

شماره چارچوب	FRAME NO.		
	1		
	2		
	3		
	4		
	5		
	6		
	7		
	8		
	9		
	10		
	11		
	12		
	13		
	14		
	15		
	16		
	17		
	18		
	19		
	20		
	21		
	22		
	23		
	24		
	25		
	26		
	27		
	28		
	29		
	30		
	31		
	32		

شکل ۳-۲: اعداد چارچوب از ۱ تا ۳۲

اکنون، به روش زیر یک ستون کنار این اعداد چارچوب ایجاد کنید:

B3 : شماره چارچوب تکراری

1 : B4

سپس مجدداً، با استفاده از دستور Home/Fill اعداد چارچوب ۳۲ تایی را ایجاد کنید به نحوی که از سلول B4 آغاز شود و در سلول B35 خاتمه یابد. مطمئن شوید که ستون‌های A و B را عریض کرده‌اید به نحوی که تمامی اطلاعات در این ستون‌ها در داخل آنها جای بگیرد. سپس، اطلاعات را در وسط ستون‌های A و B مرتب کنید. شما باید اطلاعات مندرج در شکل ۴-۲ را به دست آورید.

۲-۲- ایجاد اعداد تصادفی در صفحه کاری اکسل

شماره چارچوب	شماره چارچوب تکراری
FRAME NO.	DUPLICATE FRAME NO.
1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32

شکل ۴-۲: اعداد چارچوب تکراری از ۱ تا ۳۲

این فایل را با نام: RANDOM 3 ذخیره کنید.

شما احتمالاً تعجب می‌کنید که چرا اطلاعات مشابهی را در یک ستون A و B صفحه گستردۀ خود به وجود آورده‌اید (این کار برای این است که قبل از اینکه شما ۳۲ عدد چارچوب را مرتب کنید، مطمئن شوید دقیقاً ۳۲ عدد به هنگام اتمام مرتب کردن آن‌ها در یک تناوب تصادفی از ۳۲ عدد در اختیار دارید).

حالا بباید عدد تصادفی را به هر کدام از اعداد چارچوب تکراری اضافه کنیم:

C3: شماره تصادفی

ستون‌های A، B و C را عریض‌تر کنید به‌نحوی که عنوان‌های آن‌ها متناسب با عرض ستون شود؛ سپس اطلاعات را در مرکز سلول‌های A3 تا C35 قرار دهید.

C4: RAND() = (نوشتمن دستور)

سپس: کلید ENTER را برای اضافه کردن عدد تصادفی در سلول C4 بزنید.

توجه کنید که: شما نیاز به هر دو پرانتز باز و پرانتز بسته بعد از دستور RAND دارید. "RAND()" دستور در سمت چپ سلول با دستور RAND() در آن، جای می‌گیرد. دو عدد تصادفی را در آن سلول قرار می‌دهد.

حالا اشاره‌گر را با استفاده از ماوس خود در سلول C4 قرار دهید و سپس اشاره‌گر را به سمت پایین گوشه سمت راست آن سلول منتقل کنید تا زمانی که علامت + را در آن سلول ببینید. سپس، کلیک کنید و اشاره‌گر را تا سلول C35 بکشید تا عدد تصادفی به تمام ID32 اعداد چارچوب اضافه شود (مراجعه شود به شکل ۵-۲).

شماره چارچوب	شماره چارچوب تکراری	شماره تصادفی
FRAME NO.	DUPLICATE FRAME NO.	RANDOM NO.
1	1	0.690332931
2	2	0.022334603
3	3	0.89452184
4	4	0.981573849
5	5	0.698381228
6	6	0.611413628
7	7	0.013551391
8	8	0.036862479
9	9	0.412932328
10	10	0.460808373
11	11	0.533416136
12	12	0.988470378
13	13	0.097821358
14	14	0.881481661
15	15	0.352287507
16	16	0.344014139
17	17	0.084570168
18	18	0.467909507
19	19	0.904917153
20	20	0.232482436
21	21	0.788783634
22	22	0.592964999
23	23	0.946665187
24	24	0.214249616
25	25	0.509340791
26	26	0.439105519
27	27	0.086378662
28	28	0.975489923
29	29	0.120077924
30	30	0.216062043
31	31	0.353995884
32	32	0.558171248

شکل ۲-۵: مثالی از اعداد تصادفی اختصاص یافته به اعداد چارچوب تکراری

سپس، روی سلوی خالی کلیک کنید تا انتخاب C4:C35 از بین برود و رنگ تیره انتخاب شده این سلوول ها نیز محو شود.

این فایل را با نام: RANDOM 3A ذخیره کنید.

اکنون، اعداد چارچوب تکراری (چارچوب) را به صورت یک توالی تصادفی مرتب کنید.

۳-۲- مرتب کردن اعداد چارچوب تکراری به صورت یک توالی تصادفی

هدف: مرتب سازی اعداد چارچوب تکراری به یک توالی تصادفی

سلول های B3 تا C35 را انتخاب کنید (عنوان ها را در بالای ستون های C و B قرار دهید) در قسمت بالای صفحه نمایش، گزینه data را انتخاب کنید.

روی گزینه مرتب سازی (Sort) در بالا و مرکز صفحه نمایشگر خودتان کلیک کنید: به شکل ۶-۲ مراجعه شود.

	A	B	C
1	شماره چارچوب	شماره چارچوب تکراری	عدد تصادفی
2			Show the Sort dialog box to set it based on several criteria i ce.
3	FRAME NO	DUPLICATE FRAME NO	RANDOM
4	1	1	0.690332931
5	2	2	0.022334603
6	3	3	0.89452184
7	4	4	0.981573849
8	5	5	0.698381228
9	6	6	0.611413628
10	7	7	0.013551391
11	8	8	0.036862479
12	9	9	0.412932328
13	10	10	0.460808373
14	11	11	0.533416136
15	12	12	0.988470378
16	13	13	0.097821358
17	14	14	0.881481661

شکل ۶-۲: جعبه محاوره ای برای دستورهای data/sort

مرتب سازی بر اساس ستون عدد تصادفی (random no): بر روی جهت پایین کلیک کنید. از کوچک ترین به بزرگترین مرتب کنید (به شکل ۷-۲ مراجعه شود).

شماره چارچوب		شماره چارچوب تکراری		عدد تصادفی
FRAME NO.	DUPLICATE FRAME NO.	RANDOM NO.		
1	1	0.690332931		
2	2	0.022334603		
3	3	0.89452184		
4	4			
5	5			
6	6			
7	7			
8	8			
9	9			
10	10			
11	11			
12	12			
13	13			
14	14			
15	15			
16	16			
17	17	0.084570168		
18	18	0.467909507		

شکل ۲-۷: جعبه محاوره‌ای برای دستورهای data/sort/random no/smallest to largest

Ok را انتخاب کنید

روی هر سلوخالی به جهت حذف انتخاب C35 تا B3 کلیک کنید

این فایل را با نام 4 random ذخیره کنید.

اکنون این فایل را چاپ کنید.

این مراحل باید شکل ۲-۸ با اعداد چارچوب تکراری مرتب شده بر اساس سری تصادفی را به وجود آورد.

نکته مهم: به دلیل اینکه نرم افزار اکسل به صورت تصادفی این اعداد تصادفی را اختصاص می‌دهد، دستورهای اکسل شما یک سری متفاوت از اعداد تصادفی را برای هر کسی که این کتاب را می‌خواند، تولید می‌کند.

شماره چارچوب	شماره چارچوب تکراری	عدد تصادفی
FRAME NO.	DUPLICATE FRAME NO.	RANDOM NO.
1	7	0.343261283
2	2	0.929607291
3	8	0.914304212
4	17	0.903618324
5	27	0.257228182
6	13	0.456204036
7	29	0.390622986
8	24	0.222210116
9	30	0.432155483
10	20	0.219982266
11	16	0.842461398
12	15	0.3781508
13	31	0.694049089
14	9	0.939764564
15	26	0.075689667
16	10	0.302227714
17	18	0.468687794
18	25	0.148502036
19	11	0.49462371
20	32	0.87719372
21	22	0.413151766
22	6	0.094310793
23	1	0.962115342
24	5	0.528964967
25	21	0.401140496
26	14	0.403327013
27	3	0.865025638
28	19	0.517332393
29	23	0.968085821
30	28	0.647609375
31	4	0.670143403
32	12	0.09483352

شکل ۲-۸: اعداد چارچوب تکراری مرتب شده بر اساس عدد تصادفی

بهدلیل اینکه هدف اولیه در این فصل این بود که بهطور تصادفی پنج نفر از ۳۲ فروشنده‌گان شرکت شما، برای مصاحبه شخصی انتخاب شوند؛ اکنون شما می‌توانید پنج شماره اول شناسه افراد در ستون شماره چارچوب تکراری بعد از مرتبسازی را انتخاب کنید. اگرچه پنج عدد تصادفی شما از مواردی که ما در مرتبسازی تصادفی که در این فصل انجام دادیم متفاوت است، ما این پنج شناسه فروشنده‌گان را برای مصاحبه با استفاده از شکل ۹-۲ انتخاب می‌کنیم (۷-۲-۸-۱۷-۲۷).

شماره چارچوب	شماره چارچوب تکراری	شماره تصادفی
FRAME NO.	DUPLICATE FRAME NO.	RANDOM NO.
1	7	0.343261283
2	2	0.929607291
3	8	0.914304212
4	17	0.903618324
5	27	0.257228182
6	13	0.456204036
7	29	0.390622986
8	24	0.222210116
9	30	0.432155483
10	20	0.219982266
11	16	0.842461398
12	15	0.3781508
13	31	0.694049089
14	9	0.939764564
15	26	0.075689667
16	10	0.302227714
17	18	0.468687794
18	25	0.148502036
19	11	0.49462371
20	32	0.87719372
21	22	0.413151766
22	6	0.094310793
23	1	0.962115342
24	5	0.528964967
25	21	0.401140496
26	14	0.403327013
27	3	0.865025638
28	19	0.517332393
29	23	0.968085821
30	28	0.647609375
31	4	0.670143403
32	12	0.09483352

شکل ۹-۲: پنج فروشنده اول بهصورت تصادفی انتخاب شدند

به یاد داشته باشید که پنج شناسه انتخابی شما بعد از دسته‌بندی تصادفی با پنج عدد اول در شکل ۹-۲ متفاوت خواهد بود زیرا نرم‌افزار اکسل در هر بار استفاده از دستور $=RAND()$ اعداد تصادفی متفاوتی را اختصاص می‌دهد.

قبل از این‌که ما این فصل را به اتمام برسانیم، شما باید چگونگی چاپ فایل به نحوی که تمامی اطلاعات آن در یک صفحه جای بگیرد را بیاموزید.

۴-۲- چاپ فایل اکسل به‌طوری‌که تمامی اطلاعات آن در یک صفحه جای گیرد.

هدف: چاپ فایل اکسل به‌طوری‌که تمامی اطلاعات آن در یک صفحه جای گیرد.

توجه داشته باشید که سه تمرین عملی در انتهای این فصل نیازمند این است که شما مرتب‌سازی کنید. اعداد تصادفی زمانی که فایل‌ها شامل ۶۳ مشتری، ۱۱۴ شهرستان از ایالت میسوری و ۷۶ حساب کلیدی است، ایجاد شده‌اند. زمانی که آن‌ها را چاپ می‌کنید این فایل‌ها بیش از اندازه بزرگ هستند تا در یک صفحه جای گیرند، مگر اینکه زمانی که قصد چاپ آن را دارید. شما فرمت این فایل را به شکلی تنظیم کنید که در هنگام چاپ در یک صفحه قرار بگیرد.

حال شرایطی را ایجاد کنیم که در آن فایل روی یک صفحه چاپ نشود، مگر اینکه اول ساختار آن را ایجاد کنید.

به فایلی که ساخته‌اید برمی‌گردیم **RANDOM 4** و نام جنیفر را در سلول **۵۲** وارد می‌کیم.
اگر این فایل را اکنون چاپ کنید، نام جنیفر در صفحه دوم به چاپ خواهد رسید، زیرا در خارج از محدوده صفحه برای این فایل در ساختار فعلی قرار می‌گیرد.
بنابراین، شما باید ساختار صفحه را تغییر دهید به‌نحوی که تمام اطلاعات، از جمله نام، جنیفر، فقط در یک صفحه جای بگیرد. مراحل انجام این کار را در چند مرحله زیر توضیح می‌دهیم:

طرح‌بندی صفحه page layout (در سمت چپ بالای صفحه رایانه)

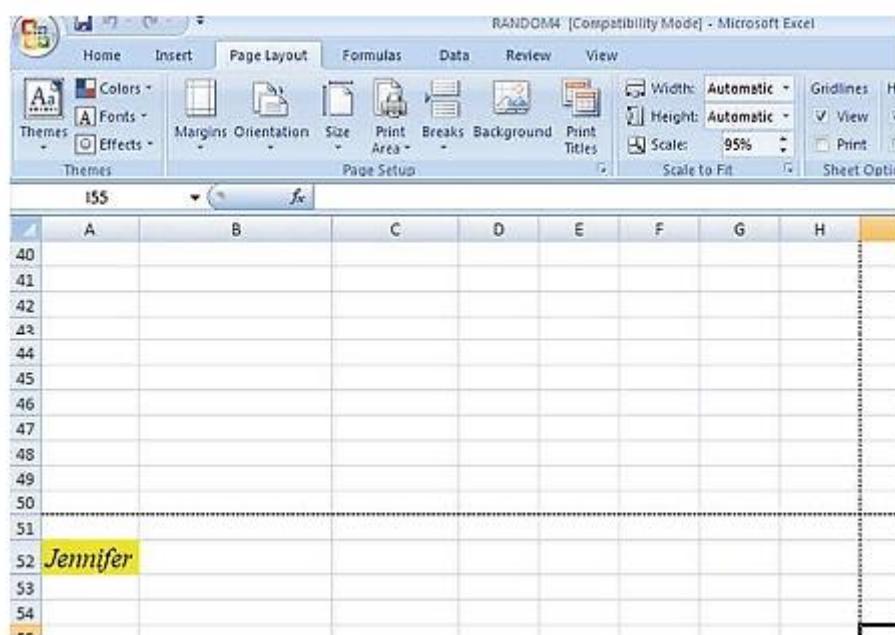
(توجه کنید که بخش مقیاس تناسب (Scale to fit) در مرکز صفحه شما قرار دارد؛ به شکل ۱۰-۲ مراجعه شود)

File	Home	Insert	Page Layout	Formulas	Data	Review	View
	Colors	Margins	Orientation	Size	Print Area	Breaks	Background
Themes	Fonts	Size	Orientation	Width	Print Titles	Scale	Height
	Effects			Width: Automatic		Scale: 100%	
	Themes			Height: Automatic		Scale to Fit	
		A1	f _x				
1	شماره چارچوب		شماره چارچوب تکراری		شماره تصادفی		
2							
3	FRAME NO.		DUPLICATE FRAME NO.		RANDOM NO.		
4	1		7		0.449458386		
5	2		2		0.393795703		
6	3		8		0.429071822		
7	4		17		0.037112211		
8	5		27		0.23465142		
9	6		13		0.89011181		
10	7		29		0.063857181		
11	8		24		0.555967593		
12	9		30		0.88888934		
13	10		20		0.456567636		

شکل ۲-۰۱: جعبه محاوره‌ای برای دستورهای page layout/scale to fit commands

به پایین صفحه بباید و برای کاهش سایز صفحه از ۱۰۰ به ۹۵ درصد، پیکان را زده تا اندازه صفحه کاهش یابد. حالا متوجه می‌شوید که نام جنیفر هنوز در صفحه دوم قرار دارد بهدلیل اینکه این اسم زیرخط چین افقی صفحه در شکل ۱۱-۲ قرار دارد.

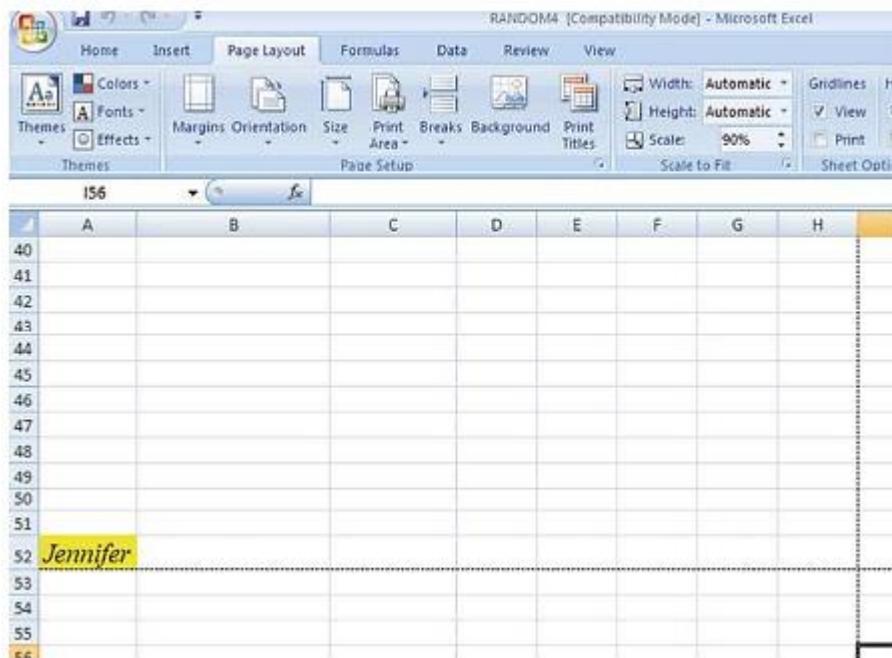
(اگر اکنون فایل را چاپ کنید، خط چین به شما ابعاد خروجی فایل را نشان می‌دهد)



شکل ۱۱-۲: مثال کاهش اندازه به ۹۵٪ با چاپ نام جنیفر در صفحه دوم

سپس، شما نیاز به تکرار «مراحل تغییر اندازه» از راه ضربه زدن به جهت پایین در سمت راست و کاهش صفحه به ۹۰ درصد اندازه عادی آن دارید.

توجه داشته باشید اکنون خطچین روی نمایشگر رایانه شما در شکل ۱۲-۲ زیر نام جنیفر قرار دارد که در این صورت تمامی اطلاعات شما را از جمله این نام را در برمی‌گیرد، اکنون با ساختار جدید به هنگام چاپ این فایل اطلاعات در یک صفحه قرار می‌گیرند.



شکل ۱۲-۲: مثال کاهش اندازه تا ۹۰٪ با چاپ شدن اسم جنیفر (Jennifer) در صفحه اول (توجه کنید که خطچین زیر اسم جنیفر در صفحه نمایشگر قرار می‌گیرد)

فایل را با نام Random4a ذخیره کنید.

فایل را چاپ کنید. آیا تمامی موارد در یک صفحه قرار گرفته است؟ باید این طور باشد (به شکل ۱۳-۲ مراجعه شود).

شماره چارچوب	شماره چارچوب تکراری	شماره تصادفی
FRAME NO.	DUPLICATE FRAME NO.	RANDOM NO.
1	7	0.661660768
2	2	0.408218127
3	8	0.360146461
4	17	0.547997374
5	27	0.821419485
6	13	0.654126828
7	29	0.704348993
8	24	0.687297652
9	30	0.577967707
10	20	0.0981433
11	16	0.609199142
12	17	0.287802372
13	31	0.435789306
14	9	0.104459646
15	26	0.805430237
16	10	0.039516242
17	18	0.734135176
18	25	0.566871050
19	11	0.381795818
20	32	0.11660887
21	22	0.891887278
22	6	0.370903093
23	1	0.109567029
24	5	0.94724966
25	21	0.650861462
26	14	0.678153692
27	3	0.081214079
28	19	0.421424271
29	23	0.817358479
30	28	0.573849656
31	4	0.597010138
32	12	0.853493587

Jennifer

شکل ۱۳-۲: صفحه گسترده نهایی با اندازه ۹۰ درصد برای متناسبسازی

۲-۵- تمرین‌های عملی پایان فصل

۱. فرض کنید می‌خواهید با استفاده از مصاحبه تلفنی رضایتمندی مشتریان را ارزیابی کنید، از ۱۵ نفر از ۶۳ مشتری که حداقل ۱۰۰۰ دلار ارزش خرید کالای آن‌ها از شرکت شما طی ۶۰ روز اخیر بوده، نظرسنجی انجام می‌دهید.
 - (a) یک صفحه گسترده از اعداد چارچوب برای این مشتریان باعنوان: شماره‌های چارچوب با استفاده از دستورهای Home/Fill تنظیم کنید.
 - (b) سپس، یک ستون جدایگانه در سمت راست این اعداد چارچوب ایجاد کنید که این اعداد چارچوب تکراری دارای عنوان "اعداد چارچوب تکراری" است.
 - (c) سپس، یک ستون جدایگانه در سمت راست این اعداد تکراری ایجاد کنید و از تابع $RAND()$ برای اختصاص عدد تصادفی به تمام اعداد چارچوب در ستون اعداد چارچوب تکراری استفاده کنید، و ساختار این ستون را تا سه رقم اعشاری برای هر عدد تصادفی تغییر دهید.
 - (d) اعداد چارچوب تکراری و اعداد تصادفی را به صورت تصادفی مرتب کنید.
 - (e) نتیجه را به طوری که محتوی صفحه گسترده در یک صفحه قرار گیرد، چاپ کنید.
 - (f) دور شماره شناسایی ۱۵ مشتری اول که در بررسی تلفنی با آن‌ها تماس می‌گیرید را در چاپ خروجی دایرہ بزنید.
 - (g) فایل را با نام 9 RAND ذخیره کنید.

نکته مهم: توجه داشته باشید هر کس که این تمرین را انجام می‌دهد، یک ترتیب تصادفی متفاوتی از شماره شناسایی مشتری را به وجود می‌آورد؛ زیرا اکسل هر بار که از دستور RAND استفاده می‌کند، یک عدد تصادفی متفاوت را انتخاب می‌کند. به همین دلیل، پاسخ به این تمرین که در این راهنمای اکسل داده شده است، یک توالی کاملاً متفاوت از اعداد تصادفی را تولید می‌کند. این امری عادی است و انتظار آن می‌رود.

۲. فرض کنید شما می‌خواهید نمونه‌ای تصادفی از ۱۰ شهرستان از ۱۱۴ شهرستان ایالت میسیوری را در پاسخ به درخواستی از طرف یک متخصص علوم سیاسی که می‌خواهد برای یک نظرسنجی تلفنی نظر رأی دهنده‌گان ثبت‌نام کرده در ایالت میسیوری درخصوص اولویت‌های خود در انتخابات بعدی را بررسی کند، انتخاب نمایید. شما می‌دانید که ۱۱۴ شهرستان در میسیوری وجود دارد زیرا شما به پایگاه اطلاعاتی اداره سرشماری ایالات متحده دسترسی

داشته‌اید (اداره آمار ایالات متحده آمریکا، ۲۰۰۰) برای اطلاعات شما، ایالات متحده در مجموع ۳۱۴۰ شهرستان در ۵۰ ایالت دارد (اداره آمار ایالات متحده آمریکا، ۲۰۰۰).

(a) یک صفحه گسترده از اعداد چارچوب برای این شهرستان‌ها با این عنوان تنظیم کنید:

FRAME NO

(b) سپس یک ستون جداگانه در سمت راست این اعداد چارچوب ایجاد کنید که این اعداد چارچوب با این عنوان تکرار می‌شوند:

Duplicate frame numbers

(c) سپس، یک ستون جداگانه در سمت راست این شماره‌های چارچوب تکراری با عنوان شماره تصادفی ایجاد کنید و از تابع $=RAND()$ برای اختصاص عدد تصادفی به تمام اعداد چارچوب در ستون اعداد چارچوب تکراری استفاده کنید. سپس، این فرمت ستون را تغییر دهید تا ۳ عدد اعشاری برای هر عدد تصادفی ظاهر شود.

(d) اعداد چارچوب تکراری و اعداد تصادفی را به صورت تصادفی مرتب کنید.

(e) نتیجه را چاپ کنید به‌طوری‌که محتوی صفحه گسترده در یک صفحه قرار گیرد.

(f) دور شماره شناسایی ۱۰ شهرستان اول را که در نظرسنجی تلفنی متخصص با آن‌ها تماس خواهد گرفت را روی نسخه چاپی خود دایره بزنید.

(g) فایل را با نام 6 RANDOM ذخیره کنید.

۳. فرض کنید که بخش فروش شرکت شما مایل به انجام "نظرسنجی رضایت مشتری" از ۲۰ "حساب کلیدی" شرکت باشد. همچنین، فرض کنید، علاوه بر این، معاون فروش، هر حساب کلیدی را به عنوان یک مشتری تعریف کرده است که در ۹۰ روز گذشته، حداقل ۳۰.۰۰۰ دلار کالا از شرکت شما خریداری کرده است.

(a) تنظیم یک صفحه گسترده از اعداد چارچوب برای این مشتریان با عنوان:

NUMBERS

(b) سپس، یک ستون جداگانه در سمت راست اعداد چارچوب ایجاد کنید که اعداد چارچوب را با این عنوان تکرار می‌کند:

Duplicate frame numbers

(c) سپس، یک ستون جداگانه در سمت راست این شماره‌های چارچوب تکراری با عنوان شماره تصادفی ایجاد کنید و از تابع $=RAND()$ برای اختصاص عدد تصادفی به تمام اعداد چارچوب در ستون اعداد چارچوب تکراری استفاده کنید. سپس، فرمت این ستون را تغییر دهید تا ۳ عدد اعشاری برای هر عدد تصادفی ظاهر شود.

(d) اعداد چارچوب تکراری و اعداد تصادفی را به صورت تصادفی مرتب کنید.

(e) نتیجه را چاپ کنید بهطوری که محتوی صفحه گسترده در یک صفحه قرار گیرد.

(f) روی نسخه چاپی، دور شماره ۲۰ مشتری اول که معاون فروش برای بررسی تلفنی با آنها تماس می‌گیرد را دایره بزنید.

(g) فایل را با نام RAND 5 ذخیره کنید.

منابع

U.S. Census Bureau Census 2000 PHC-T-4.Ranking tables for counties 1990 and 2000.Retrieved from <http://www.census.gov/population/www/cen2000/briefs/phc-t4/tables/tabc01.pdf>.

فصل سوم

فاصله اطمینان میانگین با استفاده از قابع TINV و آزمون فرضیه

این فصل بر دو ایده مرکز است: (۱) یافتن فاصله اطمینان ۹۵٪ برای میانگین و (۲) آزمون فرضیه. بباید ابتدا درباره فاصله اطمینان صحبت کنیم.

۱-۳- فاصله اطمینان میانگین

در آمار، همیشه علاقه‌مند به برآورد میانگین جامعه هستیم. چگونه آن را انجام دهیم؟

۱-۱- چگونه می‌توان میانگین جامعه را برآورد کرد

هدف: برآورد میانگین جامعه، μ

به یاد داشته باشید که میانگین جامعه متوسط تمام افراد در جامعه هدف است. به عنوان مثال، اگر علاقه‌مند (به این موضوع) شدیم که بزرگسالان بین ۲۵ تا ۴۴ سال، یک طعم جدید از بستنی بن و جری را دوست دارند، هرگز نمی‌توانیم این سؤال را از همه مردم در ایالات متحده که در آن گروه سنی هستند، بپرسیم. تکمیل چنین پژوهشی نیاز به زمان زیادی دارد و هزینه انجام این مطالعه بسیار پراهمیت است.

بنابراین، به جای بررسی هر فرد در جامعه، یک نمونه از مردم را در جامعه انتخاب می‌کنیم و از نتایج این نمونه برای برآورد میانگین کل جامعه استفاده می‌کنیم. این موجب صرفه جویی در وقت و پول می‌شود. وقتی از نتایج یک نمونه برای برآورد میانگین جامعه استفاده می‌کنیم، این فرایند «آمار استنباطی» نامیده می‌شود زیرا میانگین جامعه را از روی میانگین نمونه استنباط می‌کنیم. وقتی نمونه‌ای از افراد را در تحقیقات کسب و کار مطالعه می‌کنیم، اندازه نمونه (n)، میانگین نمونه (\bar{X}) و انحراف معیار نمونه ($STDEV$) را می‌دانیم. از این ارقام برای تخمین میانگین جامعه با یک آزمون به نام «فاصله اطمینان میانگین» استفاده می‌کنیم.

۱-۲- برآورد حد پایین و حد بالای فاصله اطمینان ۹۵ درصد میانگین

مبانی نظری این آزمون فراتر از محدوده این کتاب است و شما می‌توانید با مطالعه هر کتاب مرجع آماری خوب بیشتر در مورد آن بیاموزید (مانند لوین^۱ ۲۰۱۱)، اما ایده‌های اساسی به شرح زیر است:

¹ Levine

فرض می کنیم که میانگین جامعه نقطه‌ای در یک فاصله است که دارای "حد پایین" و "حد بالا" است. همچنین، در این کتاب فرض می کنیم که می خواهیم ۹۵٪ مطمئن باشیم که میانگین جامعه در این فاصله قرار دارد. بنابراین، قصد داریم که عبارت زیر را ارائه دهیم: ۹۵٪ اطمینان داریم که میانگین جامعه مایل طی شده به ازای هر گالن (mpg)، برای خودرو چوی ایمپالا^۱ بین ۲۶/۹۲ تا ۲۹/۴۲ mpg است.

اگر می خواهیم یک بیلبورد برای این خودرو ارائه کنیم که ادعا کند این خودرو ۲۸ مایل به ازای هر گالن (mpg) می تواند طی کند، این کار را می توانیم انجام دهیم اگر ۲۸ در فاصله اطمینان ۹۵٪ در این مثال باشد. دقیقاً نمی دانیم که میانگین جامعه چقدر است، فقط در جایی بین ۲۶/۹۲ mpg و ۲۹/۴۲ mpg است و ۲۸ در داخل این فاصله است.

اما فقط ۹۵٪ اطمینان داریم که میانگین جامعه در داخل این فاصله است، و در ۵٪ از موارد فرض ما در خصوص اینکه میانگین جمعیت ۲۸ (mpg) است، غلط خواهد بود.

اما برای اهداف ما در بررسی های کسب و کار، خوشبختانه ۹۵٪ اطمینان داریم که فرض ما دقیق است. همچنین، باید اشاره کرد که به تبع آن سطح اطمینان برای نتایج حاصل، ۹۵ درصد است. می توانیم، ۸۰٪ یا حتی ۹۰٪ یا ۹۹٪ اطمینان را برای انجام این کار، انتخاب نماییم. اما، در این کتاب، همیشه فرض می کنیم که می خواهیم از نتایج، اطمینان ۹۵٪ داشته باشیم. به این ترتیب، نیازی نیست سطح اطمینان را برای هر یک از تمرین های موجود در این کتاب حدس بزنید. همیشه می خواهیم اطمینان ۹۵٪ از نتایج خود در این کتاب داشته باشیم.

بنابراین، چگونه می توانیم فاصله اطمینان ۹۵٪ را برای میانگین داده های خود پیدا کنیم؟

این فاصله را به روش زیر پیدا خواهیم کرد:

برای به دست آوردن حد بالا در فاصله اطمینان، میانگین نمونه (\bar{X}) را به دست آورید و به آن ۱/۹۶ برابر خطای معیار میانگین (S.E) را اضافه کنید. سپس، برای به دست آوردن حد پایین فاصله اطمینان، ۱/۹۶ برابر خطای معیار میانگین را از میانگین نمونه کم کنید.

به یاد آورید که (به بخش ۱-۳ مراجعه شود) خطای معیار میانگین (S.E) از تقسیم انحراف معیار نمونه (STDEV) بر جذر تعداد اعضای نمونه، n به دست می آید.

به صورت ریاضی رابطه فاصله اطمینان ۹۵٪ میانگین عبارتست از:

$$\bar{X} \pm 1.96 \text{ s.e.} \quad (1-3)$$

¹ Chevy Impala

توجه داشته باشید که "علامت \pm " مخفف بعلاوه یا منها است، و این بدان معنی است که به منظور به دست آوردن حد بالای فاصله اطمینان، ابتدا $1/96$ برابر $s.e.$ را به میانگین اضافه کنید و سپس $1/96$ برابر $s.e.$ را از میانگین کم کنید تا حد پایین فاصله اطمینان را به دست آورید. همچنین، نماد $1.96 s.e.$ به معنای این است که شما $1/96$ را در $s.e.$ ضرب می کنید تا این بخش از رابطه برای فاصله اطمینان را به دست آورید.

توجه: در ادامه توضیح کوتاهی خواهیم داد که عدد $1/96$ از کجا آمده است.
بیایید مثال ساده‌ای برای نشان دادن این رابطه بررسی کنیم.

۳-۱-۳- برآورد فاصله اطمینان برای مایل‌طی شده به ازای هر گالن در خودرو چوی ایمپala
بیایید فرض کنیم که شما از صاحبان چوی ایمپala خواسته‌اید مسافت طی شده خود و تعداد گالن‌های مورد استفاده برای دو مخزن بنزین را بررسی کنند. فرض کنیم که 49 مالک، این کار را انجام دادند و به طور میانگین $27/83$ مایل به ازای هر گالن (mpg) با انحراف معیار $3/01$ به دست آمد. خطای معیار با تقسیم $3/01$ بر جذر 49 (یعنی 7) برابر با $0/043$ به دست می‌آید.
فاصله اطمینان 95% برای این داده‌ها به صورت زیر می‌باشد:

$$27.83 \pm 1.96 (0.43)$$

حد بالای این فاصله اطمینان از علامت مثبت در \pm استفاده می‌کند. بنابراین، حد بالا به صورت زیر خواهد بود:

$$27.83 + 1.96 (0.43) = 27.83 + 0.84 = 28.67 \text{ mpg}$$

به همین ترتیب، حد پایین این فاصله اطمینان از علامت منفی در \pm استفاده می‌کند. بنابراین، حد پایین به صورت زیر خواهد بود:

$$27.83 - 1.96 (0.43) = 27.83 - 0.84 = 26.99 \text{ mpg}$$

بنابراین، نتیجه بررسی ما این است که: "۹۵٪ اطمینان داریم که میانگین جامعه برای خودروی چوی ایمپala بین $26/99$ mpg و $28/67$ mpg است".

اگر قصد داشتیم یک بیلبورد آماده کنیم که ادعا کند این خودرو دارای 28 mpg است، می‌توانیم بر اساس اطلاعات حاصل این کار را انجام دهیم، زیرا 28 داخل این فاصله اطمینان 95% برای میانگین جامعه قرار دارد. احتمالاً از خودتان می‌پرسید: که $1/96$ در این رابطه از کجا آمده است؟

۳-۴-۱-۱ از کجا آمده است؟

پاسخ دقیق ریاضی به این پرسش فراتر از محدوده این کتاب است، اما در اینجا، ایده اصلی ارائه می شود. فرض می کنیم که داده ها در جامعه دارای توزیع نرمال هستند، به این معنا که اگر داده های همه افراد جامعه را بررسی کنیم، "شکل منحنی نرمال" را خواهد داشت.

منحنی نرمال شبیه زنگ آزادی است که در جلوی تالار استقلال در فیلادلفیا پنسیلوانیا قرار دارد. منحنی نرمال "متقارن" است؛ به این معنا که اگر آن را از وسط برش دهیم، و یک طرف آن را روی طرف دیگر قرار دهیم، کاملاً هم دیگر را پوشش می دهند. بحث درباره محاسبات انتگرال فراتر از محدوده این کتاب است. اما اساساً می خواهیم حد پایین و حد بالای داده های جامعه را در منحنی نرمال پیدا کنیم، به طوری که ۹۵٪ از سطح زیر این منحنی بین این دو حد قرار گیرد. اگر در تحقیق ما بیش از ۴۰ نفر داشته باشیم، مقدار این دو حد به ترتیب به اضافه و منهای ۱/۹۶ برابر خطای معیار میانگین (S.E.) نمونه ما است. مقدار ۱/۹۶ برابر ۰.۳۸۲ نمونه، حد بالا و حد پایین فاصله اطمینان را می دهد. اگر می خواهید درباره این موضوع بیشتر بدانید، می توانید یک کتاب آمار خوب (همانند سالکیند^۱) را مورد استفاده قرار دهید. اگر بخواهیم سطح اطمینان متفاوتی از ۹۵٪ در نتایج خود داشته باشیم، عدد ۱/۹۶ تغییر خواهد کرد.

برای مثال:

۱. اگر می خواستیم ۸۰٪ از نتایج اطمینان داشته باشیم، این رقم ۱/۲۸۲ است.
۲. اگر می خواستیم ۹۰٪ از نتایج اطمینان داشته باشیم، این رقم ۱/۶۴۵ خواهد بود.
۳. اگر می خواستیم ۹۹٪ از نتایج اطمینان داشته باشیم، این شماره ۲/۵۷۶ خواهد بود.

اما از آنجایی که در این کتاب همیشه می خواهیم از نتایج خود اطمینان ۹۵٪ را داشته باشیم، هر زمان که در بررسی خود بیش از ۴۰ نفر داشته باشیم در این کتاب همواره از ۱/۹۶ استفاده خواهیم کرد.

در حال حاضر، احتمالاً از خودتان می پرسید:

"آیا این مقدار برای فاصله اطمینان میانگین همیشه برابر ۱/۹۶ است؟" پاسخ این است: "خیر"، و اکنون توضیح خواهیم داد که چرا این پاسخ درست است.

¹ Salkind

۱-۳-۵- پیدا کردن مقدار t در رابطه فاصله اطمینان

هدف: پیدا کردن مقدار t در رابطه فاصله اطمینان

رابطه صحیح فاصله اطمینان میانگین برای اندازه‌های مختلف نمونه به شرح زیر است:

$$\bar{X} \pm t * s.e.$$

برای استفاده از این رابطه، میانگین نمونه را پیدا می‌کنید. \bar{X} ، و مقدار t برابر $s.e.$ را برای به دست آوردن حد بالای فاصله اطمینان 95% ، به آن اضافه می‌کنیم. همچنین، برای به دست آوردن حد پایین فاصله اطمینان 95% ، از میانگین نمونه، \bar{X} ، مقدار t برابر $s.e.$ را کم می‌کنیم. مقدار t را از جدولی که در ضمیمه E این کتاب است، به صورت زیر پیدا کنید:

هدف: پیدا کردن مقدار t در جدول t در ضمیمه E

قبل از توضیح آنچه به معنای "مقدار t " است، بیایید به شما تمرینی برای پیدا کردن مقدار t با استفاده از جدول t در ضمیمه E دهیم.

به ضمیمه E مراجعه کنید تا شرح دهیم که چگونه باید آن جدول را "بخوانید".

از آنجا که آزمون در این فصل "آزمون فاصله اطمینان میانگین" نامیده می‌شود، "شما از ستون اول در سمت چپ در ضمیمه E برای پیدا کردن ارزش بحرانی برای بررسی خود استفاده می‌کنید (توجه داشته باشید که این ستون با عنوان "اندازه نمونه n " نامیده می‌شود).

برای پیدا کردن مقدار t ، شما در ستون اول پایین می‌آید تا اندازه نمونه را در بررسی خود پیدا کنید و سپس به سمت راست بروید و در "ستون مقدار t بحرانی" این جدول، مقدار t را برای اندازه نمونه یاد شده پیدا کنید (توجه داشته باشید که این ستون، ستونی است که شما برای فاصله اطمینان 95% میانگین استفاده می‌کنید).

به عنوان مثال، اگر ۱۴ نفر در بررسی خود دارید، مقدار t برابر با $2/160$ است. اگر شما در تحقیقات خود ۲۶ نفر دارید، مقدار t برابر با $2/060$ است، اگر در مطالعه خود بیش از ۴۰ نفر داشته باشید، مقدار t همیشه برابر $1/96$ است.

توجه داشته باشید که "ستون t بحرانی" در ضمیمه E نشان‌دهنده مقدار t است که شما باید از آن استفاده کنید تا از نتایج خود به عنوان نتایج "معنی دار" 95% اطمینان داشته باشید.

در سرتاسر این کتاب، فرض می‌کنیم که می‌خواهید به نتیجه آزمون‌های آماری خود، 95% اطمینان داشته باشید. بنابراین، مقدار t در جدول t در ضمیمه E به شما می‌گوید که زمانی که از رابطه فاصله اطمینان 95% میانگین استفاده می‌کنید از کدام مقدار t باید استفاده کنید.

حالا که می‌دانید چگونه مقدار t را در رابطه فاصله اطمینان میانگین پیدا کنید، بیایید بررسی کنیم که چگونه این فاصله اطمینان را با استفاده از اکسل پیدا کنید.

۳-۱-۶- استفاده از تابع **TINV** اکسل برای پیدا کردن فاصله اطمینان میانگین

هدف: استفاده از تابع **TINV** در اکسل برای پیدا کردن فاصله اطمینان میانگین

هنگامی که از اکسل استفاده می‌کنید، رابطه‌هایی برای پیدا کردن فاصله اطمینان وجود دارد:

$$\bar{X} - TINV(1 - 0.95, n-1) * s.e. = \text{حد پایین} \quad (3-3)$$

$$\bar{X} + TINV(1 - 0.95, n-1) * s.e. = \text{حد بالا} \quad (4-3)$$

توجه داشته باشید که "نماد*" در این رابطه به اکسل می‌گوید که از ضرب استفاده کند و بهجای "چند برابر" از این علامت استفاده می‌شود.

از فصل ۱ به یاد می‌آورید که n بجای اندازه نمونه استفاده می‌شود و بنابراین $n-1$ جایگزین اندازه نمونه منهای یک است. همچنین شما در فصل اول یاد گرفتید که خطای معیار میانگین، یعنی $s.e.$ ، برابر با $STDEV$ است که بر جذر اندازه نمونه، n ، تقسیم شده است (به بخش ۳-۱ مراجعه شود) بیایید یک تمرین با استفاده از اکسل را برای پیدا کردن فاصله اطمینان ۹۵٪ میانگین را امتحان کنیم.

فرض کنید شرکت سازنده می‌خواهد ادعا کند که خودرو چوی ایمپالا ۲۸ مایل را به ازای هر گالن طی می‌کند و می‌خواهد روی بیلبوردها در سنت لوئیس در رودی وان دونتر^۱ به جاده ۴۴ تبلیغ کند: "خودرو چوی ایمپالا جدید ۲۸ مایل را به ازای هر گالن طی می‌کند"

بیایید mpg را "مقدار مرجع" برای این خودرو نامگذاری کنیم. فرض کنید که شما برای شرکت فورد موتور کار می‌کنید و می‌خواهید این ادعا را بررسی کنید تا ببیند آیا بر اساس برخی شواهد این ادعا درست است یا خیر. شما تصمیم به جمع‌آوری برخی از داده‌ها و استفاده از یک فاصله اطمینان ۹۵٪ دوطرفه میانگین برای بررسی نتایج خود دارید:

¹ Vandeventer

۳-۱-۷- استفاده از اکسل برای پیدا کردن فاصله اطمینان ۹۵٪ برای ادعای mpg یک خودرو

هدف: تجزیه و تحلیل داده‌ها با استفاده از فاصله اطمینان ۹۵٪ دو طرفه میانگین

شما نمونه‌ای از صاحبان خودرو جدید را برای این مدل انتخاب می‌کنید و آن‌ها موافقت می‌کنند که مسافت پیموده شده با دو مخزن سوخت را ثبت کنند و میانگین مایل به ازای هر گالن را برای این دو مخزن سوخت به دست آورند. بررسی شما نتایج ارائه شده در شکل ۱-۳ را حاصل کرد.



شکل ۱-۳: داده صفحه کاری برای خودرو چوی ایمپلا (مثال عملی)

یک صفحه گسترده با این داده‌ها ایجاد کنید و اکسل را برای پیدا کردن اندازه نمونه (n ، میانگین)، انحراف معیار (STDEV) و خطای معیار میانگین (S.E.) برای این داده‌ها با استفاده از سلول‌های مرجع زیر به کار ببرید.

A3: Chevy Impala

A5: Miles per gallon

A6: ۳۰/۹

داده‌های دیگر mpg را در سلول‌های A7:A30 وارد کنید
اکنون، سلول‌های A6:A30 را انتخاب کنید و این عده‌ها را با ساختار عددی یک رقم اعشار بازسازی کنید. این اعداد را در مرکز سلول‌های ستون A قرار دهید. سپس ستون‌های A و B را با

دو برابر کردن عرض آنها در قیاس با عرض اولیه ستون A، عریض تر کنید. سپس، ستون C را عریض تر کنید، به نحوی که عرض آن سه برابر عرض اولیه ستون A باشد. این فرایند موجب می‌شود جدول شما به نظر حرفه‌ای‌تر برسد.

C7: n اندازه نمونه

C10: mean میانگین

C13: STDEV انحراف معیار

C16: s.e. خطای معیار

C19: 95% confidence interval فاصله اطمینان ۹۵٪

D21: Lower limit حد پایین

D23: Upper limit حد بالا (به شکل ۲-۳ مراجعه شود)



شکل ۲-۳: مثال برای ایجاد برچسب‌ها در فاصله اطمینان میانگین برای تمرین خودرو چوی‌ایمپالا

B26: یک تصویر زیر این فاصله اطمینان بکشید:

B28: 26.92

B29: پایین (سپس این کلمه را راستچین کنید):

B30: حد (سپس این کلمه را راستچین کنید):

C28: "..... 28..... 28.17.....

(توجه داشته باشید که شما باید سلول C28 را با یک علامت نقل قول ("") شروع کنید تا به اکسل

بگویید که این یک عنوان است، و نه یک عدد)

D28: "..... (در ابتدا به علامت نقل قول توجه داشته باشید)

E28: " 29.42 (به علامت نقل قول توجه کنید) "

C29: میانگین مرجع

C30: مقدار

E29: بالا

E30: حد

B33: نتیجه‌گیری:

اکنون، عنوان‌ها را در زیر تصویر فاصله اطمینان مرتب کنید تا به صورت شکل ۳-۳ ظاهر شوند.



شکل ۳-۳: مثال برای رسم یک تصویر از فاصله اطمینان میانگین

سپس نام دامنه‌ای از داده‌ها شامل A30: A6 را miles بنامید.

استفاده از اکسل برای پیدا کردن اندازه نمونه: D7:

استفاده از اکسل برای پیدا کردن میانگین: D10:

استفاده از اکسل برای پیدا کردن انحراف معیار: D13:

استفاده از اکسل برای پیدا کردن خطای معیار: D16:

اکنون باید حد پایین و حد بالای فاصله اطمینان ۹۵٪ را برای این مطالعه پیدا کنید.

برای انجام این کار از تابع TINV اکسل استفاده خواهیم کرد. ما فرض می‌کنیم که می‌خواهید

اطمینان ۹۵٪ از نتایج خود داشته باشیم.

$$F21 := D10 - TINV(1 - 0.95, 24) * D16$$

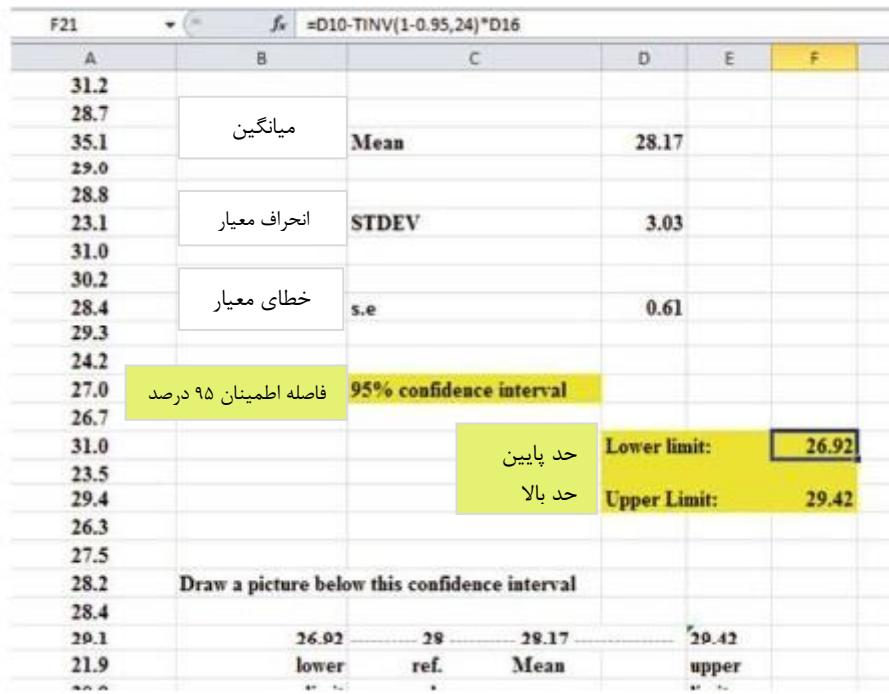
توجه داشته باشید که این رابطه $TINV$ از ۲۴ استفاده می‌کند از آنجاکه ۲۴ کمتر از اندازه نمونه یعنی ۲۵ است (یعنی، ۲۴ معادل $n-1$ است). توجه داشته باشید که D10 میانگین است، در حالی که D16 خطای معیار میانگین است. رابطه بالا به حد پایین فاصله اطمینان، مقدار ۲۶/۹۲ را ارائه می‌دهد.

$$F23 := D10 + TINV(1 - 0.95, 24) * D16$$

رابطه بالا، برای حد بالای فاصله اطمینان مقدار ۲۹/۴۲ را ارائه می‌دهد. اکنون، از قالب عددی (دو رقم اعشار) در صفحه اکسل خود برای میانگین، انحراف معیار، خطای معیار میانگین و برای هر دو حد پایین و حد بالای فاصله اطمینان خود استفاده کنید. اگر اکنون این صفحه گسترده را چاپ کنید، حد پایین فاصله اطمینان (۲۶/۹۲) و حد بالای فاصله اطمینان (۲۹/۴۲) روی صفحه دوم چاپ می‌شود؛ زیرا اطلاعات در صفحه گسترده بیش از حد بزرگ است تا در یک صفحه در قالب فعلی قرار گیرد.

بنابراین، شما باید از دستورهای "تفییر مقیاس برای متناسب‌سازی" اکسل برای کاهش اندازه صفحه گسترده به ۹۵٪ اندازه فعلی آن با استفاده از تابع Page Layout / Scale to Fit استفاده کنید که ما در فصل دوم مورد بحث قرار دادیم (به بخش ۴-۲ مراجعه شود).

اکنون این کار را انجام دهید، و توجه کنید که خط نقطه‌چین سمت راست ۲۶/۹۲ و ۲۹/۴۲ نشان می‌دهد که این شماره‌ها در زمانی که صفحه گسترده چاپ می‌شوند، در یک صفحه قرار می‌گیرند (به شکل ۴-۳ مراجعه کنید)



شکل ۳-۴: نتیجه استفاده از تابع TINV برای یافتن فاصله اطمینان میانگین

توجه داشته باشید که شما یک تصویر از فاصله اطمینان ۹۵٪ زیر سلول B26 کشیده‌اید، که حد پایین، حد بالا، میانگین و مقدار مرجع mpg ۲۸ است که ادعای شرکت در خصوص مایل طی شده به ازای هر گالن بوده را در برمی‌گیرد.

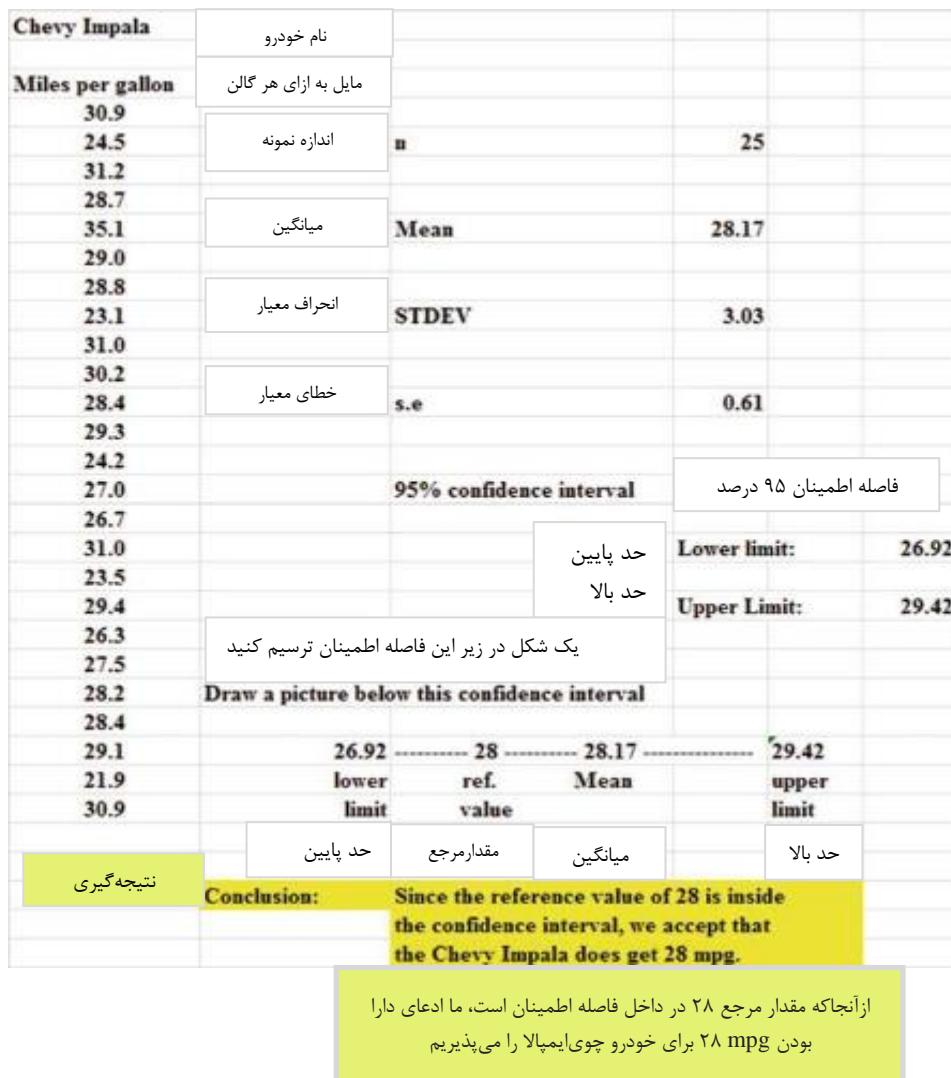
اکنون، ببایید نتیجه‌گیری مطالعه خود را روی صفحه گسترده اضافه کنیم:

از آنجاکه مقدار مرجع ۲۸ در داخل C33:

فاصله اطمینان است ما قبول می‌کنیم که

خودرو چوی‌ایمپالا می‌تواند mpg ۲۸ را دارا باشد. C35:

بررسی شما این ادعا را پذیرفت که خودروی چوی‌ایمپالا ۲۸ مایل به ازای هر گالن را طی می‌کند. میانگین مایل طی شده به ازای هر گالن در مطالعه شما ۲۸/۱۷ بود. (به شکل ۳-۵ مراجعه شود) صفحه گسترده را با عنوان CHEVY7 ذخیره کنید.



شکل ۳-۵: صفحه گسترده نهاي برای فاصله اطمینان ميانگين خودرو چوی ايمپالا

۲-۳- آزمون فرضيه

يکی از فعالیت‌های مهم پژوهشگرانی که در تحقیقات کسب و کار، تحقیقات بازاریابی، تحقیقات روان‌شناسی، تحقیقات آموزشی، و یا در هر یک از علوم اجتماعی فعال می‌باشند، این است که آن‌ها تلاش می‌کنند تا حدهای خود در مورد جهان پیرامون را به صورت "فرضیه‌ها" بررسی کنند.

یک فرضیه رایج بهصورت: "اگر X , سپس y ." است.

برخی از نمونه‌ها عبارت‌اند از:

۱- "اگر قیمت را پنج درصد افزایش دهیم، در آن صورت فروش محصول ما هشت درصد کاهش خواهد یافت."

۲- "اگر بودجه تبلیغات را ۴۰۰۰۰۰ دلار افزایش دهیم، در آن صورت سهم ما از بازار، دو رتبه افزایش خواهد یافت."

۳- "اگر از روش جدید تدریس ریاضیات در کلاس نهم جبر استفاده کنیم، در آن صورت امتیاز موققیت محاسبات ما ۱۰ درصد افزایش خواهد یافت."

۴- "اگر مواد خام محصول را تغییر دهیم، در آن صورت هزینه تولید به ازای هر واحد تولید، پنج درصد کاهش خواهد یافت."

یک فرضیه برای یک محقق علوم اجتماعی یک "حدس" در مورد آنچه ما فکر می‌کنیم در دنیای واقعی درست است، می‌باشد. می‌توانیم این حدس‌ها را با استفاده از رابطه‌های آماری آزمون کنیم تا ببینیم آیا پیش‌بینی‌های ما در دنیای واقعی درست است یا خیر. بنابراین، برای انجام این آزمون‌های آماری، ابتدا باید فرضیه‌های خود را مطرح کنیم تا بتوانیم نتایج خود را در مقابل فرضیه‌هایمان آزمون کنیم تا ببینیم فرضیه‌های ما با واقعیت مطابقت دارد یا خیر.
بنابراین، فرضیه‌ها را در دنیای کسب‌وکار چگونه ایجاد می‌کنیم؟

۱-۲-۳- فرضیه‌ها همواره به جامعه افراد یا رخدادهایی که مطالعه می‌کنید، مربوط می‌شوند

گام اول است که در کنیم فرضیه‌ها همیشه به جامعه افراد تحت مطالعه مربوط می‌شود. برای مثال، اگر ما علاقه‌مند به مطالعه افراد ۱۸-۲۴ سال در سنت لوئیس به عنوان بازار هدف خود هستیم و یک نمونه از افراد در این گروه سنی در سنت لوئیس انتخاب کردہ‌ایم، بسته به نحوه انتخاب نمونه امیدواریم نتایج حاصل از این مطالعه در تعمیم یافته‌ها برای همه افراد ۱۸ تا ۲۴ سال در سنت لوئیس و نه فقط برای افراد خاصی در نمونه ما، مفید باشد.

کل افراد گروه ۱۸ تا ۲۴ سال در سنت لوئیس جامعه‌ای است که ما علاقه‌مند به مطالعه آن بوده‌ایم، در حالی که گروه خاصی از افراد در مطالعه ما، نمونه‌ای از این جامعه نامیده می‌شوند. از آنجایی که اندازه نمونه به طور معمول تنها تعداد کمی از افراد جامعه را شامل می‌شود، لذا می‌توان نتایج نمونه را به جامعه‌ای که واقعاً علاقه‌مند به بررسی آن هستیم، تعمیم داد.

به همین دلیل است که فرضیه‌ها همیشه به جامعه مربوط می‌شود، و هرگز به نمونه‌ای از افراد در مطالعه اشاره نمی‌کند.

از فصل اول به یاد دارید که از نماد \bar{X} برای اشاره به میانگین نمونه مورداستفاده در تحقیق استفاده می‌کردیم (به بخش ۱-۱ مراجعه شود).

ما از نماد μ (حروف یونانی "mu") برای اشاره به میانگین جامعه استفاده می‌کنیم. در آزمون فرضیه‌ها ما تصمیم می‌گیریم که کدامیک از دو فرضیه رقیب در مورد میانگین جامعه را باید با توجه به مجموعه داده‌های خود پذیریم.

۳-۲-۲- فرضیه صفر و فرضیه تحقیق (جای‌گزین)

این دو فرضیه، فرضیه صفر و فرضیه تحقیق نامیده می‌شوند.

در کتاب‌های آماری معمولاً به فرضیه صفر با علامت H_0 اشاره می‌کنند:

فرضیه تحقیق به طور معمول با علامت H_1 نشان داده می‌شود و گاهی اوقات به عنوان فرضیه جای‌گزین نامیده می‌شود.

اجازه دهید نخست توضیح دهیم که منظور از فرضیه صفر و فرضیه تحقیق چیست (۱) فرضیه صفر چیزی است که ما به عنوان گویه درست می‌پذیریم، مگر اینکه قانون شویم بر اساس شواهدی درست نیست. (۲) فرضیه تحقیق چیزی است که به هنگام رد فرضیه صفر به عنوان گویه درست مورد پذیرش قرار می‌دهیم.

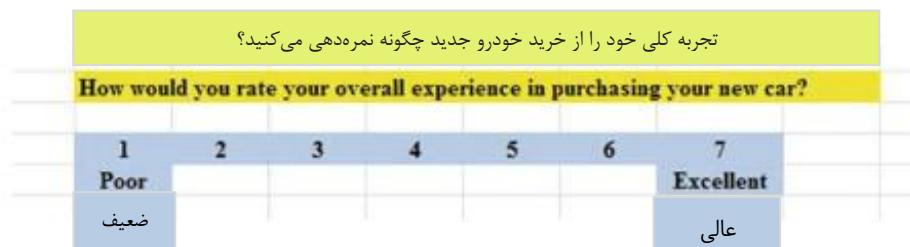
این ساز و کار مشابه نظام حقوقی امریکا است که در آن فرض می‌کنیم یک جنایتکار ظاهرًا بی‌گناه است مگر اینکه در مقابل چشم هیئت‌منصفه گناه او ثابت شود. فرضیه صفر ما این است که این متهم بی‌گناه است، در حالی که فرضیه تحقیق این است که او گناهکار است.

در ایالت بزرگ میسوری، هر اسکناس دارای شعار دولتی است: "به من نشان بده" این بدان معنا است که مردم در میسوری هر چیزی را که فردی بگوید، اگرچه به درستی، باور نمی‌کنند، مگر آن‌که اقدامات این شخص حقیقت ادعای او را نشان دهد. به عبارت دیگر، مردم در میسوری قویاً معتقدند که اقدام‌های یک فرد با صدای بلندتری از کلمات او صحبت می‌کنند.

از آنجاکه هر دو فرضیه صفر و فرضیه تحقیق نمی‌تواند درست باشد، وظیفه آزمون فرضیه‌ها استفاده از رابطه‌های آماری برای تصمیم‌گیری در مورد این است که کدام مورد را به عنوان گویه درست می‌پذیرید و کدامیک را به عنوان گویه درست رد می‌کنید. گاهی اوقات در تحقیقات کسب و کار مجموعه‌ای از مقیاس‌های نمره‌دهی برای سنجش نگرش مردم نسبت به یک شرکت، یکی از محصولات آن شرکت یا به منظور قصد خرید محصولات آن شرکت، استفاده می‌شود. این

مقیاس‌های نمره‌دهی به‌طورمعمول مقیاس‌های پنج نقطه‌ای، هفت نقطه‌ای یا ۱۰ نقطه‌ای است؛ اگرچه مقیاس‌های دیگر نیز غالباً استفاده می‌شود.

۳-۲-۱- تعیین فرضیه صفر و فرضیه تحقیق هنگام استفاده از مقیاس‌های نمره‌دهی
در اینجا نمونه‌ای از مقیاس هفت نقطه‌ای در تحقیقات نگرشی مطالعات رضایتمندی مشتری ارائه شده است (به شکل ۳-۶ مراجعه کنید).



شکل ۳-۶: مثال از یک مقیاس نمره‌دهی برای تجربه خرید یک خودرو جدید (مثال عملی)

بنابراین، در زمانی که از مقیاس نمره‌دهی استفاده می‌شود، چگونه تصمیم می‌گیریم که چه چیزی بهعنوان فرضیه صفر و فرضیه تحقیق بکار رود؟

هدف: تصمیم‌گیری در مورد فرضیه صفر و فرضیه تحقیق زمانی که از مقیاس نمره‌دهی استفاده می‌شود.

برای این تشخیص، از یک قانون ساده استفاده خواهیم کرد.
قانون: هر بار که از مقیاس نمره‌دهی استفاده شود، ما از "وسط" مقیاس بهعنوان فرضیه صفر و فرضیه تحقیق استفاده می‌کنیم.

در مثال بالا، از آنجاکه عدد چهار در وسط مقیاس است (یعنی سه عدد زیر آن است، و سه عدد بالاتر از آن می‌باشد)، فرضیه‌های ما بهقرار زیر است:

$$\text{فرضیه صفر: } \mu = 4$$

$$\text{فرضیه تحقیق: } \mu \neq 4$$

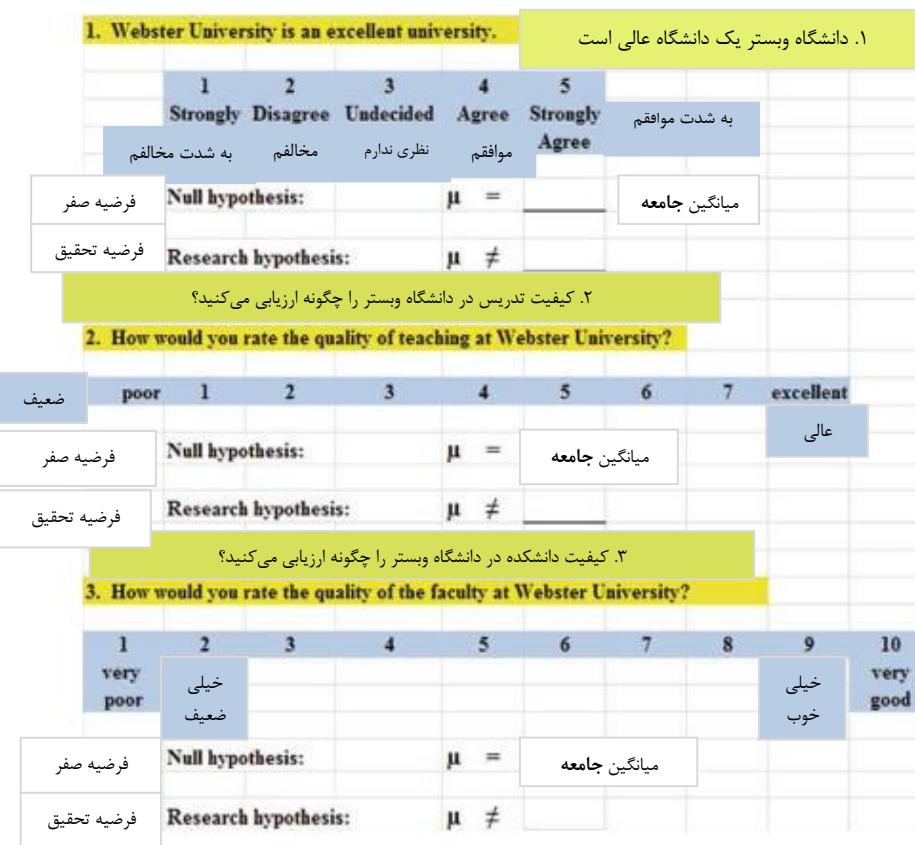
در مثال مقیاس نمره‌دهی فوق، اگر نتیجه آزمون آماری ما برای این مورد مقیاس نگرشی نشان دهد که میانگین جامعه ما "نژدیک به چهار" است، می‌گوییم که ما فرض صفر مبنی بر نه مثبت و نه منفی بودن تجربه خرید خودرو جدید را می‌پذیریم.

در مثال بالا، اگر نتیجه آزمون آماری ما نشان می‌دهد که میانگین جامعه به‌طور معنی‌داری متفاوت از چهار است. ما فرضیه صفر را رد می‌کنیم و فرضیه تحقیق را با بیان اینکه:

"تجربه خرید خودرو جدید به‌طور قابل توجهی مثبت بود" (هر زمان که میانگین نمره نمونه ما به‌طور معنی‌داری بیشتر از میانگین مورد انتظار جامعه ما یعنی ۴ است، این به‌عنوان گوییه درست پذیرفته می‌شود)، یا "تجربه خرید خودرو جدید به‌طور قابل توجهی منفی بود" (هر زمان که میانگین نمونه ما به‌طور معنی‌داری کمتر از میانگین مورد انتظار جامعه ما یعنی ۴ است، این به‌عنوان گوییه درست پذیرفته می‌شود).

همزمان هر دو نتیجه‌گیری نمی‌تواند درست باشد. یکی از فرضیه‌ها بر اساس اطلاعات داده شده در تحقیق ما به‌عنوان گوییه "درست" و دیگری بر اساس مجموعه داده‌های ما به‌عنوان گوییه "نادرست" پذیرفته می‌شود. پس از آن، وظیفه پژوهشگر کسب و کار، این است که تصمیم بگیرد کدام یک از این دو فرضیه را انتخاب کند، وی فرضیه صفر یا فرضیه تحقیق، را بر اساس داده‌های داده شده در تحقیق، به‌عنوان گوییه درست می‌پذیرد.

باییید نمونه‌هایی از مقیاس‌های نمره‌دهی را امتحان کنیم تا بتوانید تصور کنید که بر اساس هر مقیاس نمره‌دهی چه فرضیه‌ای صفر و چه فرضیه‌ای تحقیق است.
در فضای خالی شکل ۷-۳، فرضیه‌های صفر و تحقیق را برای مقیاس‌های نمره‌دهی بنویسید:



شكل ۳-۷: نمونه هایی از مقیاس های نمره دهی برای تعیین فرضیه صفر و فرضیه تحقیق

چگونه آن را انجام دادی؟

در اینجا پاسخ به این سه پرسش مطرح می شود:

- ۱- فرضیه صفر برابر بودن با عدد ۳ است و فرضیه تحقیق در مورد مقیاس پنج نقطه ای برابر نبودن با عدد ۳ است (یعنی وسط مقیاس، عدد ۳ است).
- ۲- فرضیه صفر برابر بودن با عدد ۴ است و فرضیه تحقیق در مورد مقیاس هفت نقطه ای برابر نبودن با عدد ۴ است (یعنی وسط مقیاس، عدد ۴ است).

۳- فرضیه صفر برابر بودن با $5/5$ است و فرضیه تحقیق در مورد مقیاس 10 نقطه‌ای برابر نبودن با عدد $5/5$ است (یعنی "وسط" مقیاس، عدد $5/5$ است، زیرا پنج عدد زیر $5/5$ و پنج عدد بالاتر از آن وجود دارد).

به عنوان مثال دیگر، مسافرخانه هالیدی تجربه اقامت مسافران خود را با استفاده از یک مقیاس چهار نقطه‌ای مورد ارزیابی قرار می‌دهد که در آن: $1 =$ نه خیلی خوب، $2 =$ متوسط، $3 =$ بسیار خوب، $4 =$ عالی، است.

در این مقیاس، فرضیه صفر برابر بودن با عدد $2/5$ است و فرضیه تحقیق $2/5 \neq \mu$ می‌باشد. دو عدد زیر $2/5$ و دو عدد بالاتر از $2/5$ در این مقیاس نموده‌هی وجود دارد.

اکنون هفت مرحله برای آزمون فرضیه با استفاده از فاصله اطمینان میانگین مورد بحث قرار خواهد گرفت.

۳-۲-۳- هفت مرحله برای آزمون فرضیه با استفاده از فاصله اطمینان میانگین

هدف: یادگیری هفت مرحله آزمون فرضیه با استفاده از فاصله اطمینان میانگین

هفت مرحله اصلی برای آزمون فرضیه وجود دارد.

۳-۲-۱- مرحله اول: فرضیه صفر و فرضیه تحقیق را بیان کنید

اگر در نظرسنجی خود از مقیاس‌های عددی استفاده می‌کنید، باید این را به یاد داشته باشید که این فرضیه‌ها به "وسط" مقیاس عددی اشاره دارند. به عنوان مثال، اگر شما از مقیاس 7 نقطه‌ای با $1 =$ ضعیف و $7 =$ عالی استفاده می‌کنید، این فرضیه‌ها به وسط این مقیاس‌ها اشاره می‌کنند و خواهیم داشت:

$\mu = 4$: فرضیه صفر

$\mu \neq 4$: فرضیه تحقیق

۳-۲-۲- مرحله دوم: آزمون آماری مناسب را انتخاب کنید

در این فصل، در حال بررسی فاصله اطمینان میانگین هستیم، بنابراین این آزمون را انتخاب خواهیم کرد.

۳-۲-۳- مرحله سوم: رابطه آزمون آماری را محاسبه کنید

به یاد می‌آورید (به بخش ۳-۵-۱ مراجعه شود) که رابطه فاصله اطمینان میانگین به صورت زیر است:

$$\bar{X} \pm t * s.e. \quad (2-3)$$

در مورد روش محاسبه این رابطه برای فاصله اطمینان میانگین با استفاده از اکسل در ابتدای این فصل بحث کردیم و مراحل مربوط به استفاده از آن رابطه عبارتند از:

۱. از تابع COUNT = اکسل برای یافتن اندازه نمونه، استفاده کنید.

۲. از تابع AVERAGE = اکسل برای پیدا کردن میانگین نمونه، \bar{X} استفاده کنید.

۳. از تابع STDEV = اکسل برای کردن انحراف معیار، STDEV استفاده کنید.

۴. خطای معیار میانگین (S.E) با تقسیم انحراف معیار(STDEV) بر جذر اندازه نمونه، n بدست آورید.

۵. از تابع TINV اکسل برای حد پایین فاصله اطمینان استفاده کنید.

۶. از تابع TINV اکسل برای حد بالای فاصله اطمینان استفاده کنید.

۳-۲-۴- مرحله چهارم: یک تصویر از فاصله اطمینان میانگین

یک تصویر شامل میانگین، حد پایین فاصله، حد بالای فاصله و مقدار مرجع داده شده در فرضیه صفر، H_0 ترسیم کنید.

۳-۲-۵- مرحله پنجم: در خصوص قاعده تصمیم، تصمیم‌گیری کنید

الف: اگر مقدار مرجع داخل فاصله اطمینان باشد، H_0 را قبول کنید

ب: اگر مقدار مرجع خارج از فاصله اطمینان باشد، فرضیه صفر، H_0 را رد و فرضیه تحقیق، H_1 را پذیرید.

۳-۲-۶- مرحله ششم: نتایج آزمون آماری خود را بیان کنید

هنگامی که شما از فاصله اطمینان میانگین استفاده می‌کنید، تنها یکی از فرضیه‌ها می‌تواند به عنوان گوییه "درست" پذیرفته شود. بنابراین، نتیجه شما یکی از موارد زیر است:

از آنجاکه مقدار مرجع داخل فاصله اطمینان است، فرضیه صفر، H_0 را قبول می‌کنیم

یا

از آنجاکه مقدار مرجع خارج از فاصله اطمینان است، فرضیه صفر، H_0 را رد می‌کنیم و فرضیه تحقیق، H_1 را می‌پذیریم.

۳-۲-۷- مرحله هفت: نتیجه‌گیری آزمون آماری را به زبان ساده بیان کنید!

در عمل، ارائه نتیجه‌گیری دشوارتر از آن است که به نظر می‌رسد، زیرا سعی می‌کنید نتیجه آزمون آماری خود را به صورت مختصر و دقیق و به زبان ساده خلاصه کنید.

به طوری که کسی که هرگز یک دوره آمار را نگذرانده است (شاید مانند رئیس شما) بتواند نتیجه آزمون شما را درک کند. این یک وظیفه دشوار است و ما به شما تمرین‌های زیادی در سرتاسر این کتاب می‌دهیم که این آخرین و مهم‌ترین گام را انجام دهید.

هدف: نوشتمن نتیجه‌گیری آزمون فاصله اطمینان میانگین

باید تعدادی از قوانین اساسی را برای بیان نتیجه‌گیری یک آزمون فرضیه تنظیم کنیم.

قانون شماره ۱: هر زمان که H_0 را رد کنید و H_1 را قبول کنید، شما باید از کلمه "به طور معنی‌داری" در نتیجه‌گیری استفاده کنید تا به خواننده اطلاع دهید که این آزمون یک نتیجه مهم را به دست آورده است.

قانون شماره ۲: یک طرح کلی برای «اصطلاحات کلیدی» که می‌خواهید در نتیجه‌گیری خود وارد کنید، ایجاد نمایید، طوری که، فراموش نکنید تا برخی از آن‌ها را استفاده کنید.

قانون شماره ۳: نتیجه‌گیری را به زبان ساده بنویسید تا خواننده بتواند آن را درک کند حتی اگر خواننده هرگز دوره آمار را نگذرانده باشد.

باید این قوانین را با استفاده از صفحه گسترده اکسل برای مثال خودرو چوی ایمپالا که قبلًا در این فصل ایجاد کرداید، تمرین کنید اما ابتدا باید فرضیه‌ها را بیان کنیم. از آنجاکه شرکت سازنده (در بیلبورد تبلیغاتی) می‌خواهد ادعا کند که خودرو چوی ایمپالا، ۲۸ مایل به ازای هر گالن را طی می‌کند، فرضیه‌های زیر را خواهیم داشت:

$$H_0: \mu = 28 \text{ mpg}$$

$$H_1: \mu \neq 28 \text{ mpg}$$

به یاد دارید که مقدار مرجع ۲۸ mpg ۲۸ داخل فاصله اطمینان ۹۵٪/میانگین داده‌ها بود، بنابراین، H_0 را برای خودرو چوی ایمپالا قبول می‌کنیم، یعنی، این خودرو ۲۸ mpg را دارا می‌باشد.

هدف: هنگامی که H_0 را قبول می‌کنید، نتیجه را بیان کنید

نتیجه: از آنجاکه مقدار مرجع ۲۸ mpg ۲۸ داخل فاصله اطمینان است، فرضیه صفر، H_0 را می‌پذیریم،

اکنون سه قانون خود را امتحان کنید:

هدف: نوشتمن نتیجه‌گیری وقتی که H_0 را قبول می‌کنید

قانون شماره ۱: ازآنجاکه مقدار مرجع داخل فاصله اطمینان بود، ما نمی‌توانیم از کلمه "به طور معنی‌داری" در نتیجه‌گیری استفاده کنیم. این قانون اصلی می‌باشد که ما در این فصل برای هر تمرین از آن استفاده می‌کنیم.

قانون شماره ۲: اصطلاحات کلیدی در نتیجه‌گیری عبارتند از:

- خودرو چوی ایمپالا

- مقدار مرجع ۲۸ mpg

قانون شماره ۳: خودرو چوی ایمپالا، mpg ۲۸ را دارا می‌باشد

فرایند نوشتمن نتیجه‌گیری زمانی که شما H_0 را قبول می‌کنید نسبتاً سرراست است، زیرا شما آنچه را که می‌گویید با فرضیه صفر، مطابقت دارد.

باین حال، فرایند بیان نتیجه‌گیری زمانی که شما H_0 را رد می‌کنید و H_1 را قبول می‌کنید، سخت‌تر است، بنابراین، بیایید این نوع نتیجه‌گیری را با سه مثال عملی انجام دهیم:

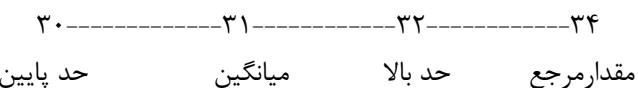
هدف: نوشتمن نتیجه و نتیجه‌گیری زمانی که H_0 را رد می‌کنید

مورد شماره ۱: فرض کنید یک تبلیغ در مجله بیزینس ویک^۱ ادعا کرد که خودرو فورد اسکیپ هیبرید^۲ ۳۴ مایل به ازای هر گالن طی می‌کند. فرضیه‌ها عبارتند از:

$$H_0: \mu = 34 \text{ mpg}$$

$$H_1: \mu \neq 34 \text{ mpg}$$

فرض کنید که تحقیقات شما فاصله اطمینان زیر را ایجاد کند:



نتیجه: ازآنجاکه مقدار مرجع خارج از فاصله اطمینان است، ما فرضیه صفر را رد می‌کنیم و فرضیه تحقیق را می‌پذیریم.

سه قاعده برای بیان نتیجه‌گیری عبارت‌اند از:

¹ Business Week

² Ford Escape Hybrid

قانون شماره یک: ما باید عبارت "به طور معنی داری" مقدار مرجع (۳۴) خارج از فاصله اطمینان است را به کار ببریم.

قانون شماره دو: اصطلاحات کلیدی عبارت اند از:

- خودرو فورد اسکیپ هیبرید
- به طور معنی داری
- "بیش از" یا "کمتر از"
- و احتمالاً نزدیک به

قانون شماره سه: خودرو فورد اسکیپ هیبرید به معنی داری کمتر از ۳۴ mpg و احتمالاً نزدیک به ۳۱ mpg را دارا می باشد

توجه داشته باشید که این نتیجه گیری می گوید که mpg کمتر از ۳۴ بود، زیرا میانگین نمونه فقط ۳۱ mpg بود. همچنین، توجه داشته باشید که هنگامی که شما با رد فرضیه صفر نتیجه معنی داری پیدا کردید، تنها کافی نیست که بگوییم: "به طور معنی داری کمتر از ۳۴ mpg" به این دلیل که این به خواننده نمی گوید "چه مقدار کمتر از ۳۴ mpg" میانگین نمونه بوده است.

برای نتیجه گیری روشن، شما باید اضافه کنید که: "احتمالاً نزدیک به ۳۱ mpg" زیرا میانگین نمونه فقط ۳۱ mpg است.

مورد شماره دو: فرض کنید که شما به عنوان یک مشاور توسط ارکستر سمفونی سنت لوئیس (SLSO) برای تحلیل داده های یک مطالعه اینترنتی از شرکت کنندگان در کنسرت سالن سمفونیک پائول^۱ در سنت لوئیس در ماه گذشته، استخدام شده اید. شما تصمیم گرفتید مهارت های تجزیه و تحلیل داده خود را در سؤال شماره هفت که در شکل ۳-۸ ارائه شده، بکار ببرید:



شکل ۳-۸: مثال بررسی میدانی مورداستفاده توسط ارکستر سمفونی سنت لوئیس (SLSO)

¹ Powell

فرضیه‌های این مورد عبارت‌اند از:

$$H_0: \mu = 4$$

$$H_1: \mu \neq 4$$

اساساً، فرضیه صفر برابر بودن با ۴ را بیان می‌کند که اگر نمره میانگین به دست‌آمده برای این پرسش به‌طور معنی‌داری از ۴ در مقیاس نمره‌دهی تفاوت نداشته باشد، درمجموع می‌توان نتیجه گرفت که شرکت‌کنندگان از کنسرت‌های SLSO نه راضی بودند و نه ناراضی. فرض کنید که تجزیه و تحلیل شما فاصله اطمینان زیر را برای این مورد در نظرسنجی ایجاد کرد.

$$4 - 3/8 - 2/8 - 1/8 -$$

حد پایین	میانگین	حد بالا
----------	---------	---------

نتیجه: از آنجایی که مقدار مرجع خارج از فاصله اطمینان است، فرضیه صفر را رد می‌کنیم و فرضیه تحقیق را می‌پذیریم.

قانون شماره یک: شما باید کلمه "به‌طور معنی‌داری" را وارد کنید زیرا مقدار مرجع خارج از فاصله اطمینان است.

قانون شماره دو: اصطلاحات کلیدی عبارت‌اند از:

- شرکت‌کنندگان
- بررسی اینترنتی SLSO
- به‌طور معنی‌داری
- در ماه گذشته
- راضی یا ناراضی (از آنجاکه نتیجه معنی‌دار است)
- تجربه در کنسرت
- به‌طور کلی

قانون شماره سه: بررسی اینترنتی تجربه کلی شرکت‌کنندگان در کنسرت SLSO در ماه گذشته نشان داد که شرکت‌کنندگان به‌طور معنی‌داری ناراضی بودند. توجه داشته باشید که نیاز است تا از کلمه "ناراضی" استفاده کنید؛ زیرا میانگین نمونه (۲/۸)، در بخش ناراضی، نسبت به وسط مقیاس نمره‌دهی قرار دارد.

مورد شماره سه: فرض کنید هتل ماریوت^۱ در محل فرودگاه سنت لوئیس، نتایج بررسی رضایت مشتریان از سرعت پذیرش طی هفته گذشته را در اختیار دارد (به شکل ۹-۳ مراجعه شود):



شکل ۳-۹: نمونه‌ای از بررسی میدانی در هتل ماریوت

برای این مورد، فرضیه‌های زیر ارائه شد:

$$H_0: \mu = 5.5$$

$$H_1: \mu \neq 5.5$$

فرض کنید که تحقیقات شما فاصله اطمینان زیر را برای این مورد در نظرسنجی ایجاد کرد:

Δ/Δ ----- Δ/∇ ----- Δ/Λ ----- Δ/Ψ

حد بالا میانگین حد پایین مقدار مرجع

نتیجه: از آنجاکه مقدار مرجع خارج از فاصله اطمینان است، فرضیه صفر را رد و فرضیه تحقیق را قبول می‌کنیم.

سه قانون برای بیان نتیجه‌گیری عبارتند از:

قانون شماره یک: شما باید کلمه "بهطور معنی داری" را وارد کنید؛ زیرا مقدار مرجع خارج از فاصله اطمینان است.

قانون شماره دو: اصطلاحات کلیدی عبارتند از:

- هتل ماریوت

فروڈگاہ سنت لوئیس -

- به طور معنی داری

سرعت پذیرش -

بررسی (مطالعه میدانی)

- هفته گذشته

¹ Marriott

- مشتریان

- "مثبت" یا "منفی" (ما این را توضیح خواهیم داد)

قانون شماره سه: مشتریان هفتاه گذشته هتل ماریوت در فرودگاه سنت لوئیس، سرعت پذیرش خود را به طور معنی داری مثبت اظهار کردند. به دو مورد مهم در مورد نتیجه‌گیری فوق توجه داشته باشد: (۱) مردم در هنگام صحبت کردن به طور معمول نمی‌گویند: "به طور معنی داری عالی است" از آنجاکه نمره $5/8$ به طور معنی داری بیشتر از $5/5$ در سمت مثبت مقیاس بود، ما برای نشان دادن این واقعیت می‌گوییم "به طور معنی داری مثبت".

سه تمرین پایانی این فصل، تمرین‌های بیشتری را برای بیان نتیجه‌گیری مربوط به نتایج فراهم می‌آورد و این کتاب در بردارنده نمونه‌های بسیاری است که به شما کمک می‌کند تا نتیجه‌گیری دقیق و شفاف، برای یافته‌های پژوهش خود بنویسید.

۳-۳- راه‌های جای‌گزین برای خلاصه کردن نتیجه یک آزمون فرضیه

مهم است بدانید که در این کتاب، آزمون فرضیه به یکی از دو روش زیر خلاصه می‌شود: (۱) فرضیه صفر را قبول می‌کنیم؛ (۲) فرضیه صفر را رد می‌کنیم و فرضیه تحقیق را می‌پذیریم. به صورت مداوم از این واژه‌ها استفاده می‌کنیم تا بتوانید مفهوم آزمون فرضیه را درک کنید. با وجود این، بسیاری از روش‌های دیگر برای خلاصه کردن نتیجه آزمون فرضیه وجود دارد، و همه آن‌ها به لحاظ نظری درست است، هرچند اصطلاحات به کار رفته متفاوت است. اگر دوره آموزشی را می‌گذرانید که از شما خواسته می‌شود نتیجه یک آزمون فرضیه را به زبان متفاوت از زبان این کتاب خلاصه کنید، وحشت نکنید! اگر مفهوم آزمون فرضیه را همان‌گونه که در این کتاب توضیح داده شده است، درک کرده‌اید، می‌توانید جهت نتیجه‌گیری مشابه از آزمون فرضیه، درک خود را برای استفاده از شرایطی که استاد شما می‌خواهد، تغییر دهید. متخصصان آمار و استادان کسب و کار هر کدام زبان مورد علاقه خود را برای خلاصه کردن نتایج آزمون فرضیه دارند. هرگز مجموعه‌ای از کلمات وجود ندارد که آنها در مورد آن توافق داشته باشند، بنابراین، موردنی را که از لحاظ درک مفهوم آزمون فرضیه آسان‌تر است، انتخاب می‌کنیم. برای متقاعد کردن شما که روش‌های مختلفی برای خلاصه کردن نتایج یک آزمون فرضیه وجود دارد، نقل قول‌های زیر را از کتاب‌های آماری و روش تحقیق برجسته ارائه می‌دهیم تا به شما ایده‌هایی از راه‌های مختلف ممکن ارائه دهیم.

۳-۱- روش‌های مختلف برای پذیرفتن فرضیه صفر

زمانی که فرضیه صفر پذیرفته می‌شود، نقل قول‌های زیر، جمله‌های مورد استفاده در کتاب‌های آماری و روش تحقیق است:

"فرضیه صفر رد نشده است." (بلاک^۱، ۲۰۱۰، ص. ۳۱۰)

"فرضیه صفر را نمی‌توان رد کرد." (مکدانیل و گیتس^۲، ۲۰۱۰، ص. ۵۴۵)

"فرضیه صفر ... ادعا می‌کند که بین گروه‌ها تفاوت وجود ندارد." (سالکیند ۲۰۱۰، ص. ۱۹۳)

"تفاوت به لحاظ آماری معنی دار نیست." (مکدانیل و گیتس^۳، ۲۰۱۰، ص. ۵۴۵)

"... مقدار به دست آمده به اندازه کافی بزرگ نیست که بگوییم تفاوت بین گروه ۱ و ۲ توسط چیزی غیر از شанс تبیین می‌شود." (سالکیند ۲۰۱۰، ص. ۲۲۵)

"اگر فرضیه صفر را رد نکنیم، نتیجه‌گیری می‌کنیم که شواهد آماری کافی برای تغییر این نکته وجود ندارد که جای‌گزین (فرضیه) درست باشد." (کلر^۴، ۲۰۰۹، ص. ۳۵۸)

"فرضیه تحقیق پشتیبانی نمی‌شود." (زیکماند و بابین^۵، ۲۰۱۰، ص. ۵۵۲)

۳-۲-۳- روش‌های مختلف برای رد فرضیه صفر

زمانی که فرضیه صفر رد می‌شود، جمله‌های زیر، نقل قول‌های بارز مورد استفاده در کتاب‌های آماری و روش تحقیق است:

"فرضیه صفر رد شده است." (مکدانیل و گیتس^۶، ۲۰۱۰، ص. ۵۴۶)

"اگر ما فرضیه صفر را رد کنیم، نتیجه‌گیری می‌کنیم که شواهد آماری کافی برای برقراری فرضیه جای‌گزین وجود دارد." (کلر^۷، ۲۰۰۹، ص. ۳۵۸)

"اگر مقدار آماره آزمون با فرضیه صفر ناسازگار باشد، فرضیه صفر را رد می‌کنیم و می‌پذیریم که فرضیه جای‌گزین درست است." (کلر^۸، ۲۰۰۹، ص. ۳۴۸)

"از آنجاکه مقدار مشاهده شده ... بزرگتر از مقدار بحرانی است...، تصمیم این است که فرضیه صفر را رد کنیم." (بلاک^۹، ۲۰۱۰، ص. ۳۵۹)

"اگر مقدار به دست آمده از مقدار بحرانی بسیار بزرگتر باشد، فرضیه صفر نمی‌تواند پذیرفته شود"

(سالکیند ۲۰۱۰، ص. ۲۴۳)

"اگر آزمون فرضیه به لحاظ آماری معنی دار باشد، نباید مقدار t بحرانی از مقدار مشاهده شده بیشتر باشد" (زیکماند و بابین^{۱۰}، ۲۰۱۰، ص. ۵۶۷)

"آماره محاسباتی آزمون فراتر از حد بالا بوده و به ناحیه رد وارد می‌شود. پس، فرضیه صفر رد می‌شود" (ویرس^{۱۱}، ۲۰۱۱، ص. ۳۳۰)

¹ Black

² McDaniel and Gates

³ Keller

⁴ Zikmund and Babin

⁵ Black

باید توجه داشته باشید که تمام نقل قول های بالا در هنگام تفسیر نتایج یک آزمون فرضیه، به وسیله متخصصان آمار و اساتید استفاده می شود و بنابراین، نباید تعجب کنید اگر کسی از شما بخواهد که نتایج یک آزمون آماری را با استفاده از زبان متفاوت از آنچه در این کتاب استفاده می کنیم، خلاصه کنید.

۴-۳- تمرین های عملی پایان فصل

۱. فرض کنید که مدیر روزنامه پست-دیسپچ^۱ سنت لوئیس از شما خواسته است تا تجزیه و تحلیل داده های نظرسنجی اخیر از مشترکین قدیمی که اخیراً اشتراک روزنامه خود را طی سه ماه گذشته لغو کرده اند، انجام دهید. یک نمونه تصادفی از این گروه از طریق تلفن فراخوانده شدن و مجموعه-ای از پرسش ها درباره روزنامه از آن ها پرسیده شد. داده های فرضی برای پرسش شماره ۴ نظرسنجی، در شکل ۳-۱۰ نشان داده شد:

¹ Weiers

² Post-Dispatch

بررسی تلفنی روزنامه پست-دیسپچ سنت لوئیس

St. Louis Post-Dispatch Phone Survey

Question #4: "How much would you be willing to pay per week for a six-month weekday/weekend subscription to the Post-Dispatch?"

حق اشتراک	Subscription Price (\$)
	4.15
	3.75
	3.80
	4.10
	3.60
	3.60
	3.65
	4.40
	3.15
	4.00
	3.75
	4.00
	3.25
	3.75
	3.30
	3.75
	3.65
	4.00
	4.10
	3.90
	3.50
	3.75

پرسش شماره ۴: "تمایل به پرداخت
شما به ازای هر هفته برای حق اشتراک
شش ماهه روزهای هفته/آخر هفته
روزنامه پست-دیسپچ چقدر است؟"

شکل ۱۰-۳: داده‌های صفحه کاری برای فصل ۳: تمرین شماره ۱

همچنین، فرض کنید، مدیریت ارشد می‌خواهد هزینه $\frac{3}{80}$ دلار بابت حق اشتراک جدید در نظر بگیرد. آیا این ($\frac{3}{80}$ دلار) قیمت مناسب برای حق اشتراک بر اساس نتایج این پرسش نظرسنجی است؟ (نکته: $\frac{3}{80}$ دلار فرضیه صفر برای این حق اشتراک است)

- (a) در سمت راست این جدول، از اکسل برای پیدا کردن اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین برای ارقام قیمت استفاده کنید. عنوان‌های پاسخ خود را با استفاده از قالب پولی (با دو رقم اعشار) برای میانگین، انحراف معیار و خطای معیار میانگین بکار گیرید.
- (b) فرضیه صفر و فرضیه تحقیق را روی صفحه گسترده وارد کنید.
- (c) استفاده از تابع TINV اکسل برای پیدا کردن فاصله اطمینان ۹۵٪ میانگین برای این اعداد. عنوان‌های پاسخ خود را با استفاده از قالب پولی (با دو رقم اعشار) بیان کنید.
- (d) نتیجه خود را بر روی صفحه گسترده وارد کنید.
- (e) نتیجه‌گیری خود را به زبان ساده روی صفحه گسترده وارد کنید.
- (f) صفحه گسترده نهایی را به نحوی که نتایج روی یک صفحه قرار گرفته باشد، چاپ کنید (اگر نیاز به کمک دارید تا چگونگی انجام این کار را به یاد آورید، به قسمت اهداف را در پایان فصل ۲ در بخش ۴-۲ مراجعه کنید).
- (g) در چاپ خود، نمودار این فاصله اطمینان ۹۵٪ را به صورت دستی بکشید.
- (h) فایل را با عنوان POST9 ذخیره کنید

۲. فرض کنید بخش منابع انسانی (HR) شرکت از شما خواسته است تا تجزیه و تحلیل داده‌های نظرسنجی نگرشی با عنوان چگونگی تفکر مدیران درخصوص کار در شرکت را انجام دهید. می‌خواهید مهارت‌های اکسل خود را در یک نمونه کوچک از مدیران با یک مورد پرسش موجود در نظرسنجی آزمون کنید. فرض کنید یک نمونه تصادفی از مدیران را انتخاب می‌کنید و داده‌های مربوط به مورد شماره ۲۴ که در شکل ۱۱-۳ آورده شده است، را در اختیار دارید.

HUMAN RESOURCES DEPARTMENT							دایره منابع انسانی
MORALE SURVEY OF MANAGERS							بررسی نگرش مدیران
Item #24: "How would you rate the quality of leadership shown by top management in this company?"							مورد شماره ۲۴: "کیفیت رهبری مدیر ارشد شرکت را چگونه ارزیابی می‌کنید؟"
1 very weak	2	3	4	5	6	7 very strong	خیلی قوی
	خبلی	ضعیف		Rating	نمره دهی		
			5				
			6				
			3				
			4				
			7				
			2				
			3				
			4				
			2				
			5				
			3				
			4				
			2				
			3				
			6				
			5				
			7				
			4				
			6				
			4				
			3				
			4				
			2				
			3				
			5				
			4				
			2				
			3				
			5				
			4				

شکل ۱۱-۳: داده‌های صفحه کاری برای فصل ۳: تمرین عملی شماره ۲

صفحه گسترده اکسل را با این داده‌ها ایجاد کنید.

- (a) از سمت راست جدول اکسل برای پیدا کردن اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین برای این داده‌ها استفاده کنید. عنوان پاسخ را ارائه کنید و از دو رقم اعشار برای میانگین، انحراف معیار و خطای معیار میانگین استفاده کنید.
- (b) فرضیه صفر و فرضیه تحقیق برای این مورد را در صفحه گسترده خود وارد کنید.
- (c) از تابع **TINV** اکسل برای پیدا کردن فاصله اطمینان ۹۵٪ میانگین این داده‌ها استفاده کنید. عنوان پاسخ‌های خود را در صفحه گسترده بنویسید و از دو رقم اعشار برای حد پایین و حد بالا فاصله اطمینان استفاده کنید.
- (d) نتیجه آزمون را در صفحه گسترده خود وارد کنید.
- (e) نتیجه‌گیری آزمون را به زبان ساده در صفحه گسترده وارد کنید. (اگر نیاز به کمک دارد تا چگونگی انجام این کار را به یادآورید، اهداف را در پایان فصل ۲ (بخش ۴-۲ ببینید)
- (f) یک تصویر از فاصله اطمینان، شامل مقدار مرجع، بر روی صفحه گسترده ترسیم نمایید.
- (g) صفحه گسترده نهایی را با عنوان top8 ذخیره کنید.

۳. فرض کنید که از شما خواسته شده است که به منظور تعیین میزان علاقه‌مندی مشتریان به یک لباس جدید که توسط طراح مشهور عرضه شده است، ارزیابی سه گروه مخاطب از زنان بالغ (۲۵-۴۴ ساله) را در شهرهای مختلف بر عهده بگیرید. طراح امیدوار است که این لباس در فروشگاه‌های بزرگ با قیمت خردۀ فروشی ۶۸/۰۰ دلار، فروخته شود. یک بحث گروهی یک ساعته با سه گروه از زنان بزرگسال مخاطب در این محدوده سنی را هدایت می‌کنید.
داده‌های حاصل برای آخرین پرسشن این بررسی در شکل ۱۲-۳ ارائه شده است:



شکل ۱۲-۳: داده‌های صفحه کاری برای فصل ۳: تمرین عملی شماره ۳

صفحه گسترده اکسل را با این داده‌ها ایجاد کنید.

- (a) از سمت راست جدول اکسل برای پیدا کردن اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین برای این داده‌ها استفاده کنید. پاسخ‌های خود را عنوان گذاری کنید و از دو رقم اعشار و قالب پولی برای میانگین، انحراف معیار و خطای معیار میانگین استفاده کنید.
- (b) فرضیه صفر و فرضیه تحقیق برای این مورد را بر روی صفحه گسترده خود وارد کنید.
- (c) از تابع TINV اکسل برای پیدا کردن فاصله اطمینان ۹۵٪ میانگین برای این داده‌ها استفاده کنید پاسخ‌های خود را در صفحه گسترده عنوان‌بندی کنید. با استفاده از دو رقم اعشار در قالب پولی حد پایین و حد بالای فاصله اطمینان را بنویسید.
- (d) نتیجه آزمون را در صفحه گسترده خود وارد کنید.
- (e) نتیجه‌گیری آزمون را به زبان ساده در صفحه گسترده خود وارد کنید.
- (f) صفحه نهایی خود را به‌طوری‌که بر روی یک صفحه قرار گیرد چاپ کنید (اگر شما نیاز به کمک در به یادآوردن چگونگی انجام این کاردارید، اهداف را در پایان فصل ۲ بخش ۴-۲ ببینید).
- (g) تصویر فاصله اطمینان، با لحاظ مقدار مرجع، را بر روی صفحه گسترده بکشید.
- (h) صفحه گسترده نهایی را با عنوان blouse9 ذخیره کنید.

منابع

- Black K. Business statistics: for contemporary decision making. 6th ed. Hoboken: Wiley; 2010.
- Keller G. Statistics for management and economics. 8th ed. Mason: South-Western Cengage learning; 2009.
- Levine DM. Statistics for managers using microsoft excel. 6th ed. Boston: Prentice Hall/Pearson; 2011.
- McDaniel C, Gates R. Marketing research. 8th ed. Hoboken: Wiley; 2010.
- Salkind NJ. Statistics for people who (think they) hate statistics (2nd Excel 2007 ed.). Los Angeles: Sage Publications; 2010.
- Weiers RM. Introduction to business statistics. 7th ed. Mason: South-Western Cengage Learning; 2011.
- Zikmund WG, Babin BJ. Exploring marketing research. 10th ed. Mason: South-Western Cengage learning; 2010.

فصل چهارم

آزمون یک گروهی t برای میانگین

در این فصل خواهید آموخت که چگونه از یکی از عمومی‌ترین و مفیدترین آزمون‌های آماری در تحقیقات کسب و کار استفاده کنید: آزمون یک گروهی t برای میانگین. رابطه آزمون یک گروهی t به شرح زیر است:

$$t = \frac{\bar{X} - \mu}{S_{\bar{X}}} \quad (1-4)$$

$$s.e. = S_{\bar{X}} = \frac{s}{\sqrt{n}} \quad (2-4)$$

این رابطه از شما می‌خواهد میانگین نمونه (\bar{X}) را به دست آورید و آن را از میانگین جامعه (μ) کم کنید و سپس حاصل را بر خطای معیار میانگین ($s.e.$) تقسیم کنید. خطای استاندارد میانگین برابر است با انحراف معیار تقسیم بر جذر n (اندازه نمونه). بباید ۷ مرحله آزمون فرضیه با استفاده از آزمون یک گروهی t را مورد بحث قرار دهیم تا بتوانید درک کنید که چگونه این آزمون مورد استفاده قرار می‌گیرد.

۱-۴-مراحل هفت‌گانه آزمون فرضیه با استفاده از آزمون یک گروهی t

هدف: پادگیری مراحل هفت گانه آزمون فرضیه با استفاده از آزمون یک گروهی t

قبل از این که قادر باشید مهارت‌های اکسل خود را برای آزمون یک گروهی t ارزیابی کنید، باید مراحل اساسی آزمون فرضیه را برای این آزمون آماری یاد بگیرید. در این فرایند ۷ مرحله وجود دارد:

۱-۱-۱- مرحله اول: بیان کردن فرضیه صفر و فرضیه تحقیق

اگر شما در بررسی خود از مقیاس‌های عددی یا کمی استفاده می‌کنید باید به خاطر داشته باشید که این فرضیه‌ها به وسط مقیاس‌های عددی اشاره دارند. به عنوان مثال، اگر شما از مقیاس ۷ نقطه-ای با $=1$ ضعیف و $=7$ عالی استفاده می‌کنید، این فرضیه‌ها نشان‌دهنده وسط این مقیاس است و خواهیم داشت:

$H_0: \mu = 4$ فرضیه صفر

$H_1: \mu \neq 4$ فرضیه تحقیق

به عنوان مثال دوم، فرض کنید که شما برای شرکت هوندا موتور کار می‌کنید و در نظر دارید تبلیغی در یک مجله ارائه دهید که ادعا می‌کند هوندا به ازای هر گالن، ۳۵ مایل را طی می‌کند. فرضیه‌ها برای آزمون این ادعا بر اساس داده‌های واقعی عبارتند از:

$H_0: \mu = 35 \text{ mpg}$

$H_1: \mu \neq 35 \text{ mpg}$

۴-۱-۲- مرحله دوم: آزمون آماری مناسب را انتخاب کنید.

در این فصل آزمون یک گروهی t را مطالعه خواهیم کرد؛ بنابراین بررسی این آزمون مدنظر است.

۴-۱-۳- مرحله سوم: تصمیم‌گیری در مورد یک قاعده تصمیم (نتیجه‌گیری) برای آزمون یک گروهی t

(a) اگر قدر مطلق t محاسباتی کمتر از مقدار t بحرانی باشد، فرضیه صفر قبول می‌شود.

(b) اگر قدر مطلق t محاسباتی بزرگتر از مقدار t بحرانی باشد، فرضیه صفر رد و فرضیه تحقیق پذیرفته می‌شود.

احتمالاً می‌گویید: خوب به نظر می‌رسد، اما چگونه می‌توان مقدار قدر مطلق آماره t محاسباتی را پیدا کرد؟

۴-۱-۳-۱- پیدا کردن مقدار قدر مطلق یک عدد

برای انجام این کار، ما به یک هدف دیگر نیاز داریم:

هدف: پیدا کردن مقدار قدر مطلق یک عدد

اگر یک دوره پایه جبری را در دبیرستان گذرانده باشید، ممکن است مفهوم «مقدار قدر مطلق» را به خاطر داشته باشید: در اصطلاح ریاضی قدر مطلق هر عدد همیشه همان عدد است که به صورت مشبّت بیان می‌شود.

به عنوان مثال، قدر مطلق $2/35$ برابر $2/35 +$ است.

و قدر مطلق منفی $2/35$ (یعنی $-2/35$) نیز $2/35 +$ است.

این امر زمانی حائز اهمیت است که شما از جدول t در ضمیمه E این کتاب استفاده می‌کنید. در ادامه در مورد این جدول بحث خواهیم کرد و در مرحله ۵ آزمون یک گروهی t توضیح داده می‌شود که چگونه می‌شود با استفاده از ضمیمه E مقدار t بحرانی را پیدا کرد.

۴-۱-۴- مرحله چهارم: محاسبه رابطه آزمون t یک گروهی

هدف : یادگیری نحوه استفاده از رابطه آزمون t یک گروهی

رابطه آزمون t یک گروهی به شرح زیر است:

$$t = \frac{\bar{X} - \mu}{S_{\bar{X}}} \quad (1-4)$$

که در آن

$$S.e. = S_{\bar{X}} = \frac{S}{\sqrt{n}} \quad (2-4)$$

این رابطه فرضیه‌های زیر را در مورد داده‌ها لحاظ می‌کند (فاستر و همکاران^۱، ۱۹۹۸).

(۱) داده‌ها مستقل از یکدیگر هستند (یعنی هر فرد تنها یک نمره دریافت می‌کند).

(۲) جامعه آماری مرتبط با داده دارای توزیع نرمال است و (۳) داده‌ها واریانس ثابت دارند (توجه داشته باشید که انحراف معیار، جذر واریانس است).

برای استفاده از این رابطه نیاز است این مراحل را دنبال کنید:

۱. میانگین نمونه مورد مطالعه خود را به دست آورده و میانگین جامعه μ را از آن کم کنید (به یاد داشته باشید که میانگین جامعه برای مطالعه‌ای که شامل مقیاس‌های رتبه‌بندی عددی است، مقدار وسط، در آن مقیاس می‌باشد).

۲. پس از اینکه جواب مرحله بالا را به دست آورده‌ید، پاسخ خود را بر خطای معیار میانگین، تقسیم کنید (در فصل اول یاد گرفتید چگونه خطای معیار میانگین را پیدا کنید، برای پیدا کردن خطای معیار میانگین فقط انحراف معیار مطالعه خود را بیابید و آن را بر جذر n تقسیم کنید که در آن n تعداد افرادی است که در مطالعه شما مورد ارزیابی قرار می‌گیرند).

۳. عددی که شما بعد از تکمیل مرحله بالا به دست می‌آورید، مقدار t محاسباتی است که به هنگام استفاده از رابطه یادشده حاصل می‌شود.

¹Foster et al.

۴-۱-۵- مرحله پنجم: مقدار t بحرانی را از جدول t در ضمیمه E پیدا کنید.

هدف: پیدا کردن مقدار t بحرانی از جدول t در ضمیمه E

قبل از توضیح معنای مقدار " t بحرانی"، اجازه دهید بهمنظور پیدا کردن مقدار t بحرانی با استفاده از جدول t در ضمیمه E تمرين ارائه شود.

به ضمیمه E مراجعه کنید، در ادامه توضیح می‌دهیم که چطور باید آن جدول را بخوانید.

از آنجاکه آزمون موردبحث در این فصل آزمون t یک گروهی نامیده شد، برای پیدا کردن مقدار t بحرانی در تحقیق خود باید از اولین ستون سمت چپ در ضمیمه E استفاده کنید (توجه داشته باشید که عنوان این ستون «تعداد نمونه» است).

برای پیدا کردن مقدار t بحرانی، در ستون اول پایین بیایید تا تعداد نمونه موردبررسی در تحقیق خود را پیدا کنید و سپس به سمت راست بروید و مقدار t بحرانی را برای این تعداد نمونه در ستون t بحرانی در جدول بخوانید (توجه داشته باشید ستونی که شما استفاده خواهید کرد برای آزمون t یک گروهی و فاصله اطمینان ۹۵٪ برای میانگین است).

به عنوان نمونه، اگر در پژوهش خود ۲۷ نفر داشته باشید، مقدار t بحرانی ۲/۰۵۶ است.

اگر افراد پژوهش خود را به ۳۸ نفر برسانید مقدار t بحرانی ۲/۰۲۶ است.

اگر در پژوهش خود بیش از ۴۰ نفر داشته باشید، مقدار t بحرانی همیشه برابر ۱/۹۶ است.

توجه داشته باشید که ستون t بحرانی در ضمیمه E نشان‌دهنده مقدار t است که باید به منظور اطمینان از معنی‌داری نتایج با اطمینان ۹۵ درصد به دست آورید.

مقدار t بحرانی، مقداری است که به شما می‌گوید آیا در آزمون آماری خود نتیجه معنی‌داری پیدا کرده‌اید یا خیر.

جدول t در ضمیمه E یک سری از منحنی‌های نرمال زنگوله‌ای شکل را نشان می‌دهد (منحنی‌ها زنگوله‌ای نامیده می‌شوند زیرا آن‌ها شبیه ناقوس‌های آزادی هستند که شما می‌توانید در فیلادلفیا بیرون سالن استقلال ببینید).

وسط منحنی نرمال، جایی روی محور X وجود دارد که نقطه صفر در نظر گرفته می‌شود (توضیحات فنی این مطلب فراتر از محدوده این کتاب است، اما اگر علاقه‌مند به یادگیری بیشتر در این مورد هستید، هر کتاب آمار خوب (به عنوان مثال زیکموند و بابین^۱، ۲۰۱۰) این مفهوم را به شما توضیح خواهد داد).

بنابراین، مقادیر t که در سمت راست صفر قرار دارند مقادیر مثبتی هستند که قبل از آن‌ها از علامت (+) استفاده می‌کنند، و مقادیر t که در سمت چپ صفر قرار دارند، مقادیر منفی هستند که

¹ Zikmund and Babin

از علامت منفی (-) قبل از آن‌ها استفاده می‌کنند. بنابراین، برخی مقادیر t مثبت و برخی دیگر منفی هستند.

با وجود این، هر کتاب آمار که جدول t را در بر دارد، تنها بخش مثبت منحنی t را چاپ می‌کند زیرا طرف منفی تصویری از سمت مثبت است؛ این بدان معنی است که طرف منفی حاوی اعدادی است که دقیقاً همانند طرف مثبت هستند، اما علامت منفی برای همه این اعداد لحاظ می‌شود.

بنابراین، برای استفاده از جدول t در ضمیمه E، شما باید قدر مطلق مقدار t که از فرمول t محاسبه کرده‌اید را به دست آوردید. زیرا جدول t در ضمیمه E فقط دارای مقادیر t مثبت است.

در سرتاسر این کتاب، ما فرض می‌کنیم که می‌خواهید نتایج آزمون‌های آماری خود را با ۹۵٪ اطمینان داشته باشید. بنابراین، مقدار t در جدول t ضمیمه E به شما می‌گوید که آیا مقدار t که با استفاده از رابطه آزمون t یک گروهی به دست آورده‌اید در محدوده ۹۵٪ زیر منحنی t قرار دارد یا

خیر، که بتوان انتظار داشت مقدار t محاسبه شده با اطمینان ۹۵٪ رخ داده است.

اگر مقدار t که با استفاده از رابطه آزمون t یک گروهی به دست آوردید در فاصله اطمینان ۹۵٪ قرار داشت، می‌گوییم نتیجه‌ای که به دست آورده‌اید معنی دار نیست (توجه داشته باشید که این معادل پذیرش فرضیه صفر است).

اگر مقدار t که با استفاده از رابطه آزمون t یک گروهی به دست آوردید در خارج از فاصله اطمینان ۹۵٪ قرار داشت، می‌گوییم که نتیجه معنی دار است که انتظار می‌رود در ۵٪ از کل حالات اتفاق بیفتد (توجه داشته باشید که این معادل رد فرضیه صفر و پذیرش فرضیه تحقیق است).

۴-۶-۶- مرحله ششم: نتایج آزمون آماری خود را بیان کنید

وقتی از آزمون t یک گروهی استفاده می‌کنید دو نتیجه ممکن است حاصل شود، و فقط یکی از آن‌ها می‌تواند به عنوان نتیجه "درست" پذیرفته شود.

اگر قدر مطلق t که از طریق رابطه t به دست آورده‌اید از مقدار t بحرانی در ضمیمه E کمتر باشد، فرضیه صفر را می‌پذیرید.

یا اگر مقدار قدر مطلق t محاسباتی که از طریق رابطه t به دست آورده‌اید از مقدار t بحرانی در ضمیمه E بیشتر باشد، فرضیه صفر را رد کرده و فرضیه تحقیق را می‌پذیرید.

۴-۱-۷- مرحله هفتم: نتیجه آزمون آماری خود را به صورت ساده بیان کنید

در عمل، این کار سخت‌تر از آن است که به نظر می‌رسد، زیرا سعی می‌کنید نتیجه آزمون آماری خود را به زبان ساده که مختصر و دقیق است خلاصه کنید تا کسی که هرگز دوره‌ی آماری را ندیده است (شاید مانند رئیس شما) بتواند نتایج آزمون را درک کند. این یک وظیفه دشوار است و در این خصوص به شما تمرین‌های زیادی ارائه می‌دهیم. این آخرین و مهم‌ترین مرحله در سرتاسر این کتاب است.

اگر این مطالب را به خوبی مطالعه کرده باشید، آماده هستید تا پشت رایانه خود بشینید و با استفاده از اکسل آزمون t مربوط به اطلاعات فرضی در مورد نظرسنجی از میزان رضایتمندی مهمانان هتل ماریوت را انجام دهید. اجازه دهید امتحان کنیم.

۴-۲- آزمون t یک گروهی برای میانگین

فرض کنید به عنوان مشاور آماری توسط هتل ماریوت^۱ در سنت لوئیس^۲ استخدام شده‌اید تا داده‌هایی که در آن‌ها مشتریان درجه رضایتمندی خود را از فعالیت‌های مختلف هتل بیان کرده‌اند، تجزیه و تحلیل کنید. این نظرسنجی شامل تعدادی از موارد است، اما فرض کنید مورد شماره ۷ که یکی از آن‌ها می‌باشد، مدنظر بوده و در شکل ۱-۴ آمده است.



شکل ۱-۴: نمونه مورد بررسی برای هتل ماریوت (مثال عملی)

فرض کنید که تصمیم گرفته‌اید داده‌های مربوط به مشتریان هفته گذشته را با استفاده از آزمون یک گروهی t تجزیه و تحلیل کنید. نکته مهم: باید از این آزمون برای بررسی هر یک از موارد به طور جداگانه استفاده کنید.

¹ Marriott Hotel

² St. Louis

فرض کنید داده‌های فرضی مورد شماره ۷ برای هفته گذشته در هتل ماریوت سنت لوئیس شامل نمونه‌ای از ۱۲۴ میهمان است که دارای میانگین ۶/۵۸ و انحراف معیار ۲/۴۴ است.

هدف: تجزیه و تحلیل داده‌ها برای هر پرسش به صورت جداگانه با استفاده از آزمون t یک گروهی.

یک فایل صفحه گسترده اکسل را با اطلاعات زیر ایجاد کنید:

فرضیه صفر: B11

فرضیه تحقیق: B14

توجه: به یاد داشته باشید زمانی که از یک مورد با مقیاس رتبه‌ای استفاده می‌کنید هر دو فرضیه صفر و تحقیق به "وسط مقیاس" اشاره دارند. در مقیاس ۱۰ نقطه‌ای این مثال، وسط مقیاس ۵/۵ است، زیرا پنج عدد زیر ۵/۵ (یعنی ۱ تا ۵) و پنج عدد بالاتر از ۵/۵ (یعنی ۶ تا ۱۰) می‌باشند. بنابراین، فرضیه‌ها برای این مقیاس رتبه‌ای عبارت‌اند از:

$$H_0: \mu = 5.5$$

$$H_1: \mu \neq 5.5$$

n: تعداد B17

میانگین: B20

انحراف معیار: B23

خطای معیار: B26

t بحرانی: B29

t آزمون: B32

نتیجه: B36

جمع‌بندی: B41

اکنون از اکسل استفاده کنید:

اندازه نمونه (تعداد) را وارد کنید: D17

میانگین را وارد کنید: D20

انحراف معیار را وارد کنید (به شکل ۲-۴ مراجعه شود): D23

Null hypothesis:	فرضیه صفر
Research hypothesis:	فرضیه تحقیق
n	تعداد 124
mean	میانگین 6.58
STDEV	انحراف معیار 2.44
s.e.	خطای معیار
critical t	مقدار t بحرانی
t-test	مقدار t محاسباتی
Result:	نتیجه
Conclusion:	نتیجه‌گیری

شکل ۴-۴: جدول داده‌های اصلی برای برخورد صمیمانه کارکنان پذیرش

D26: خطای معیار را با استفاده از رابطه ارائه شده در فصل اول محاسبه کنید.

D29: مقدار t بحرانی را از جدول t در ضمیمه E پیدا کنید.

اکنون رابطه زیر را در سلول D32 وارد کنید تا نتیجه آزمون t را پیدا کند.

$$= (D20-5.5)/ D26$$

این رابطه میانگین نمونه (D20) را انتخاب می‌کند و میانگین جامعه فرضی یعنی ۵/۵ را از میانگین نمونه کم می‌کند و سپس حاصل را بر خطای معیار میانگین (D26) تقسیم می‌کند. توجه داشته باشید که شما نیاز به وارد کردن $5/5 - D20$ با یک پرانتر باز قبل از D20 و یک پرانتر بسته بعد از $5/5$ دارید تا پاسخ یعنی $1/0.8$ را بر خطای معیار یعنی 0.22 تقسیم کند تا نتیجه آزمون t برابر $4/93$ حاصل شود.

اکنون از دو رقم اعشار برای خطای معیار و نتیجه آزمون t استفاده کنید (به شکل ۳-۴ مراجعه شود)

Null hypothesis:	فرضیه صفر
فرضیه تحقیق	
Research hypothesis:	
n	تعداد 124
mean	میانگین 6.58
STDEV	انحراف معیار 2.44
s.e.	خطای معیار 0.22
مقدار t بحرانی	
critical t	1.96
مقدار t محاسباتی	
t-test	4.93
Result:	نتیجه
Conclusion:	نتیجه‌گیری

شکل ۴-۳: نتیجه رابطه آزمون t برای برخورد صمیمانه کارکنان پذیرش

حالا، جملات زیر را در D36 تا D39 بنویسید تا نتیجه آزمون t را خلاصه کنید:

D36: از آنجاکه قدر مطلق t محاسباتی، $4/93$ است

D37: بزرگ‌تر از t بحرانی $1/96$ است،

D38: فرضیه صفر را رد می‌کنیم و

D39: فرضیه تحقیق را می‌پذیریم.

درنهایت جمله زیر را در D43 تا D41 بنویسید تا جمع‌بندی مربوط به مورد شماره ۷ یعنی بررسی رضایتمندي مهمانان هتل ماریوت را به صورت خلاصه ارائه کنید.

D41: بر اساس امتیازدهی مهمانان هتل ماریوت سنت لوئیس

D42: برخورد کارکنان پذیرش به صورت معنی‌داری

D43: در هفته گذشته صمیمانه بوده است.

فایل خود را با عنوان MARRIOTT3 ذخیره کنید.

صفحه گسترده نهایی را مطابق شکل ۴-۴ به نحوی چاپ کنید که بر روی یک صفحه تنظیم شده باشد و فرضیه صفر و فرضیه تحقیق را به صورت دستی روی صفحه گسترده بنویسید.

فرضیه صفر		
Null hypothesis:	$\mu = 5.5$	
فرضیه تحقیق		
Research hypothesis:	$\mu \neq 5.5$	
n	تعداد	124
mean	میانگین	6.58
STDEV	انحراف معیار	2.44
s.e.	خطای معیار	0.22
critical t	آماره بحرانی	1.96
t-test	آماره محاسباتی	4.93
Result:	<p>Since the absolute value of t of 4.93 is greater than the critical t of 1.96, we reject the null hypothesis and accept the research hypothesis.</p>	
Conclusion:	<p>St. Louis Marriott Hotel guests rated the Front Desk Clerks as significantly friendly last week.</p>	
نتیجه گیری	<p>در هفته گذشته میهمانان هتل ماریوت به طور معنی داری رفتار کارکنان پذیرش را صمیمانه ارزیابی کردند.</p>	

شکل ۴-۴: صفحه گسترده نهایی برای برخورد صمیمانه کارکنان پذیرش

نکته مهم: برای شما اهمیت دارد که بدانید "از لحاظ فنی"، نتیجه گیری فوق باید بر اساس اصطلاحات آماری به صورت زیر بیان شود:
 میهمانان هتل ماریوت سنت لوئیس در هفته گذشته، برخورد کارکنان پذیرش را صمیمانه ارزیابی کردند، و این نتیجه احتمالاً از روی شانس به دست نیامده است.

با این حال، در سرتاسر این کتاب، از عبارت "به طور معنی دار" در نوشتن جمع‌بندی آزمون‌های آماری استفاده شده است تا خوانندگان متوجه شوند که نتیجه آزمون آماری احتمالاً این نتیجه شناسی نبوده است. اما بجای نوشتن مکرر تمام این کلمات، ما از کلمه "به طور معنی دار" به عنوان یک اصطلاح مختصر به جای یک توضیح طولانی استفاده می‌کنیم. زمانی که نتایج بجای زبان فنی آماری با زبان ساده نوشته شود، فهمیدن نتایج برای خواننده آسان‌تر می‌شود.

۴-۳-۴- زمانی که آزمون فرضیه انجام می‌شود، آیا می‌توان از فاصله اطمینان ۹۵٪ میانگین یا آزمون t یک گروهی، استفاده کرد؟

احتمالاً از خودتان پرسیده‌اید:

به نظر می‌رسد که می‌توانید از فاصله اطمینان ۹۵٪ میانگین یا آزمون t یک گروهی برای تجزیه و تحلیل نتایج انواع مسائل که تاکنون در این کتاب مطرح شده است، استفاده کنید؟ آیا این جمله صحیح است؟

پاسخ آشکار است: بله! هم فاصله اطمینان میانگین و هم آزمون t یک گروهی اغلب در تحقیقات کسب و کار برای انواع مسائلی که تاکنون در این کتاب شرح داده شده است، مورد استفاده می‌باشد. هردو این آزمون‌ها نتایج مشابهی به دست می‌آورند و نتیجه‌گیری یکسانی را برای سری داده‌ها عاید می‌کنند.

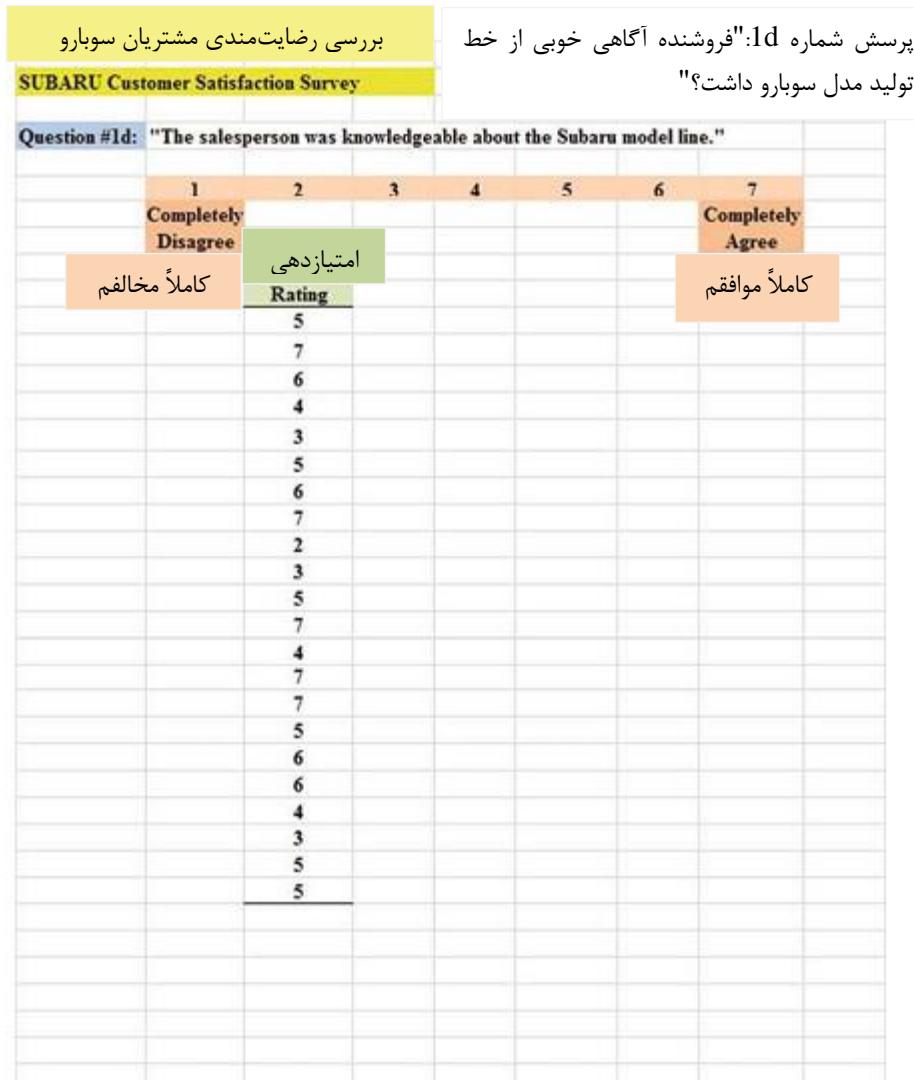
هردو این آزمون‌ها در این کتاب توضیح داده شده است، زیرا برخی از مدیران ترجیح می‌دهند که فاصله اطمینان میانگین را استفاده کنند و برخی دیگر آزمون t را ترجیح می‌دهند و عده‌ای نیز ترجیح می‌دهند از هر دو آزمون برای بررسی داده‌های مشابه استفاده کنند تا نتایج برای خوانندگان گزارش‌های پژوهشی شفاف‌تر باشد. از آنجایی که نمی‌دانیم کدامیک از این آزمون‌ها را مدیر شما ترجیح می‌دهد، هر دو را توضیح دادیم تا قادر به استفاده از هر دو آزمون در تجزیه و تحلیل داده‌های آماری باشید.

اکنون باید مهارت‌های اکسل خود را در آزمون یک گروهی t روی سه تمرین پایانی این فصل امتحان کنید.

۴-۴- تمرین‌های عملی پایان فصل

۱. شرکت سوبارو آمریکا میزان رضایتمندی مشتریان از تجربه خرید از نمایندگان فروش را به صورت هفتگی بررسی می‌کند، و در پی این است که ارزیابی کند آیا نمایندگان فروش به رضایتمندی ۹۳٪ مشتریان دست یافته‌اند، یا آنها نیازمند آموزش‌های بیشتری برای بهبود

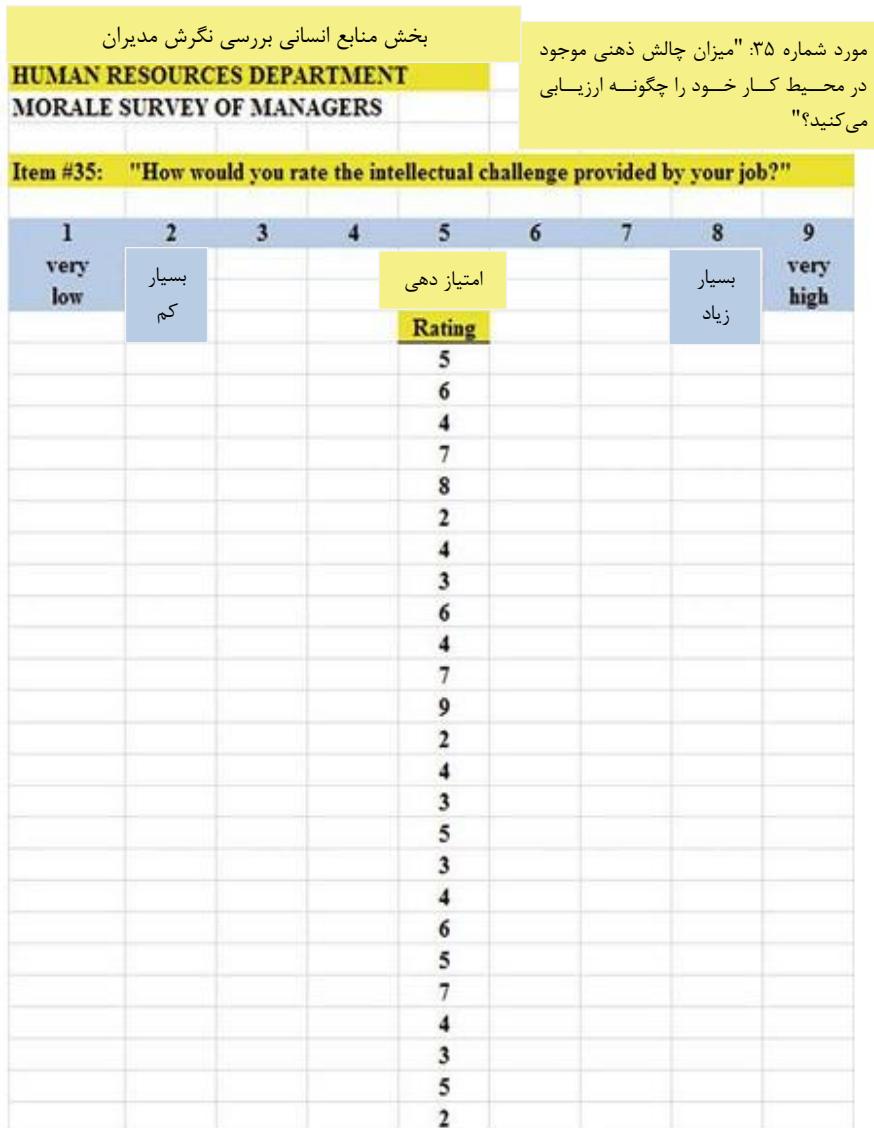
رضایتمندی مشتریان خود هستند. فرض کنید شما یک نمونه تصادفی از فرم‌های ارزیابی ارائه شده به وسیله خریداران خودرو جدید (برخط یا از راه پست الکترونیک) برای نمایندگی سوبارو سنت لوئیس طی یک هفته اخیر را انتخاب کرده‌اید و جدول فرضی مطابق شکل ۴-۵ را برای پرسش شماره 1d تهیه کرده‌اید:



شکل ۴-۵: صفحه گسترده داده‌ها برای فصل ۴: تمرین عملی شماره ۱

- (a) فرضیه صفر و فرضیه تحقیق را در صفحه گسترده بنویسید.
- (b) با استفاده از اکسل اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین را در سمت راست سری داده‌ها پیدا کنید. با استفاده از گزینه فرمت عدد تا دو رقم اعشار را برای میانگین، انحراف معیار و خطای معیار میانگین ارائه کنید.
- (c) از جدول t در ضمیمه E مقدار t بحرانی را در صفحه گسترده وارد کنید و آن را نام‌گذاری نمایید.
- (d) با استفاده از اکسل مقدار t را برای این داده‌ها محاسبه کنید (از دو رقم اعشار استفاده کنید) و روی صفحه گسترده خود نام‌گذاری نمایید.
- (e) نتیجه را در صفحه گسترده خود تایپ کنید و سپس نتیجه‌گیری یا جمع‌بندی را به زبان ساده در صفحه تایپ کنید.
- (f) فایل را با عنوان 4 subaru ذخیره کنید.

۲. فرض کنید که در بخش منابع انسانی شرکت خود کار می‌کنید و مدیریت ارشد بخش، از شما خواسته است تا یک نظرسنجی در مورد نگرش مدیران نسبت به کار در این شرکت، انجام دهید. برای بررسی مهارت‌های اکسل خود، یک نمونه تصادفی از نتایج نظرسنجی مدیران و پاسخ آنها به یک پرسش را در اختیار دارید. داده‌های مربوط به مورد شماره ۳۵ مطابق شکل ۶-۴ در اختیار شما است:



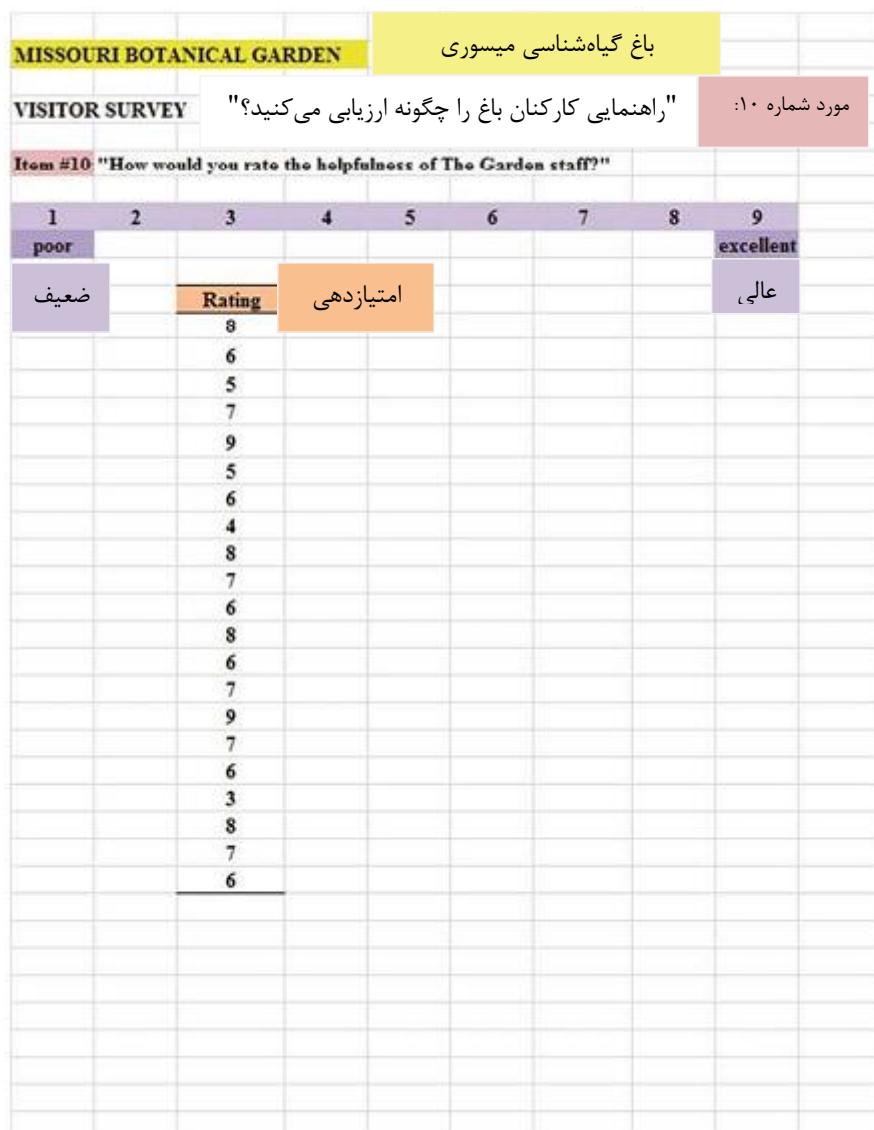
شکل ۴-۶: صفحه گسترده داده‌های فصل ۴: تمرین عملی شماره ۲

- (a) در صفحه گسترده اکسل، فرضیه صفر و فرضیه تحقیق برای این داده‌ها را بنویسید.
- (b) از اکسل برای یافتن اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین برای این داده‌ها استفاده کنید (دو رقم اعشار برای میانگین، انحراف معیار و خطای معیار از میانگین).

- (c) از اکسل برای انجام دادن آزمون t یک گروهی بر روی داده‌ها استفاده کنید (دو رقم اعشار).
- (d) در نتیجه چاپ شده خود، مقدار t بحرانی (در سطح ۰/۰۵) را که در جدول t در ضمیمه E ارائه شده است، تایپ کنید.
- (e) در صفحه گسترده نتیجه آزمون t را تایپ کنید.
- (f) در صفحه گسترده نتیجه گیری را به زبان ساده تایپ کنید.
- (g) فایل را با عنوان challenge4 ذخیره کنید.

۳. فرض کنید به عنوان مشاور بازاریابی توسط باغ گیاه‌شناسی میسوری استخدام شده‌اید و از شما خواسته شده است که کارت‌های نظرسنجی که بازدیدکنندگان بعد از بازدید باغ پر می‌کنند را دوباره طراحی کنید. مدیریت باغ از یک مقیاس ۵ نقطه‌ای از ۱ = ضعیف تا ۵ = عالی استفاده می‌کرد.

فرض کنید، علاوه براین شما کارکنان باغ را متقادع کرده‌اید تا امتیازدهی، به مقیاس ۹ نقطه‌ای از ۱ = ضعیف تا ۹ = عالی تغییر کند، بطوریکه داده‌ها، انحراف معیار بزرگ‌تری داشته باشند. نتایج فرضی برای پرسش شماره ۱۰ از فرم نظرسنجی بازنگری شده شما طی یک هفته اخیر در شکل ۴-۷ نشان داده شده است.



شکل ۴-۷: صفحه گستردۀ داده‌های مربوط به فصل ۴: تمرین عملی شماره ۳

- (a) فرضیه صفر و فرضیه تحقیق را در صفحه گستردۀ خود بنویسید.
- (b) با استفاده از اکسل، تعداد نمونه، میانگین، انحراف معیار و خطای معیار میانگین را به دست آورید. با استفاده از گزینه فرمت عدد، مقادیر میانگین، انحراف معیار و خطای معیار میانگین، را تا دو رقم اعشار ارائه کنید.

- (c) مقدار t بحرانی را از جدول t در ضمیمه E پیدا کرده و در صفحه گسترده وارد کنید و آن را نام‌گذاری نمایید.
- (d) از اکسل برای محاسبه مقدار t برای این داده‌ها استفاده کنید (از دو رقم اعشار استفاده کنید) و آن را روی صفحه گسترده نام‌گذاری نمایید.
- (e) نتیجه را در صفحه گسترده تایپ کنید و سپس نتیجه‌گیری را به زبان ساده در صفحه گسترده تایپ نمایید.
- (f) فایل را با نام Garden5 ذخیره کنید.

منابع

- Foster DP, Stine RA, Waterman RP. Basic business statistics: a casebook. New York: Springer; 1998.
- Zikmund WG, Babin BJ. Exploring marketing research. 10th ed. Mason: South-Western Cengage learning; 2010.

فصل پنجم

آزمون t دو گروهی اختلاف میانگین گروههای مستقل

تاکنون در این کتاب، با وضعیتی که در آن فقط یک گروه از افراد و تنها یک عدد مورد اندازه‌گیری برای هر یک از این افراد وجود داشت، مواجه بوده‌اید. ما تغییر رویه داده و با وضعیتی که در آن دو گروه از افراد را به جای یک گروه از افراد موردنظرسی قرار داده، مواجه می‌شویم. هر زمان که دو گروه کاملاً متفاوت از افراد داشته باشید، (یعنی یک فرد در هر دو گروه نباشد، اما هر فرد تنها با یک متغیر اندازه‌گیری می‌شود که به تولید یک عدد برای هر فرد منجر شود)، می‌گوییم که دو گروه «مستقل از یکدیگر» هستند. این فصل فقط به این وضعیت می‌پردازد. از این رو، به آن، آزمون t دو گروهی برای گروههای مستقل اطلاق می‌شود.

فرضیه‌های آزمون t دو گروهی به شرح ذیل است (زیکماند و بابین، ۲۰۱۰): (۱) هر دو گروه از یک جامعه نرمال نمونه‌برداری شده‌اند و (۲) واریانس دو جامعه تقریباً برابر است. توجه داشته باشید که انحراف معیار فقط جذر واریانس است (رابطه‌های مختلفی برای زمانی که هر فرد دو بار برای ایجاد دو گروه داده اندازه‌گیری شود، وجود دارد و این وضعیت "وابسته" نامیده می‌شود، اما این روابط فراتر از محدوده این کتاب است). این کتاب فقط با دو گروه سروکار دار که مستقل از یکدیگر هستند به طوری که هیچ‌کس در هر دو گروه از داده‌ها وجود ندارد.

هنگامی که اختلاف میانگین دو گروه را آزمون می‌کنید، مهم است که به یاد داشته باشید که دو رابطه مختلف وجود دارد که باید با توجه به اندازه نمونه دو گروه استفاده شوند:

- اگر، هر دو گروه بیش از ۳۰ نفر داشته باشند، از رابطه اول در این فصل استفاده کنید.

- اگر یک گروه یا هر دو نمونه‌ای کمتر از ۳۰ نفر داشته باشند، از رابطه دوم استفاده کنید هر دو وضعیت را در این فصل تشریح می‌نماییم. قبل از ورود به روابط این آزمون، نخست، نیاز است تا گام‌های موجود در آزمون فرضیه را زمانی که دو گروه از افراد وجود دارند موردنظرسی قرار دهیم.

۵-۱- نه مرحله برای آزمون فرضیه با استفاده از آزمون t دو گروهی

هدف: یادگیری نه مرحله آزمون فرضیه با استفاده از دو گروه از افراد و آزمون t دو گروهی

خواهید دید که این مراحل مشابه مراحل استفاده شده در فصل قبل بوده که مربوط به آزمون t یک گروهی است، اما تفاوت‌های مهمی بین این مراحل وجود دارد که قبل از این که وارد رابطه‌ها برای آزمون t دو گروهی شویم، به وضوح آن‌ها را درک می‌کنید.

۵-۱-۱- مرحله اول: یک گروه را گروه ۱ و دیگری را گروه ۲ نام‌گذاری کنید

رابطه‌های مورداستفاده در این فصل از اعداد ۱ و ۲ برای تمایز کردن دو گروه استفاده می‌کنند. اگر پی ببرید که کدام، گروه ۱ و کدام گروه ۲ است، می‌توانید از این اعداد در محاسبات استفاده کنید بدون اینکه مجبور باشید نام گروه را بنویسید.

برای مثال، اگر شما ترجیحات پسروان نوجوان را برای مزه نوشابه‌های کوکا^۱ یا پپسی^۲ مورد بررسی قرار دهید، شما می‌توانید گروه‌ها را: "کوکا" و "پپسی" بنامید. این امر نیاز به نوشتمن کلمه‌های "کوکا" یا "پپسی" در هر زمان که بخواهید به یکی از این گروه‌ها اشاره کنید، دارد. اگر شما گروه کوکا را، گروه ۱ و گروه پپسی را گروه ۲ بنامید، این باعث می‌شود که خیلی راحت‌تر به گروه‌ها اشاره کنید، زیرا موجب صرفه جویی در زمان نوشتمن می‌شود.

در مثال دوم، نتایج آزمون بازار را برای کانزاس سیتی در مقابل ایندیاناپولیس مقایسه می‌کنید، اگر مجبور شوید نام این شهرها را هر زمان که بخواهید به آن‌ها اشاره کنید، بنویسید، زمان بیشتری را در قیاس با حالتی که یک شهر را گروه ۱ و دیگری را گروه ۲ می‌نامید، صرف می‌کنید. همچنین، توجه داشته باشید که این موضوع کاملاً اختیاری است که شما کدام را گروه ۱ و کدام را گروه ۲ می‌نامید. جدای از چگونگی تصمیم شما در خصوص نحوه نام‌گذاری، نتیجه و نتیجه‌گیری مشابهی از این رابطه‌ها به دست خواهد آورد.

۵-۱-۲- مرحله دوم: یک جدول ایجاد کنید که اندازه نمونه، میانگین و انحراف معیار هر گروه را به‌طور خلاصه نشان دهد

این مرحله، اطمینان یافتن از این‌که اعداد درستی را در رابطه آزمون t دو گروهی به کار می‌برید، آسان می‌کند. اگر اعداد اشتباه را به کار ببرید، تمام کار شما اشتباه خواهد شد و مسئله را به سختی حل خواهید کرد.

برای مثال، فرض کنید که شما ترجیحات پسروان نوجوان را برای مزه کوکا در مقابل پپسی بررسی می‌کنید که در آن پسرها به‌طور تصادفی برای قضاوت در مورد مزه یکی از این برندها انتخاب شده‌اند و مزه آن را در مقیاس ۰ تا ۱۰۰ برابر ضعیف تا عالی اظهار می‌کنند.

پس از اتمام بررسی، فرض کنید که گروه کوکا ۵۲ نوجوان را شامل می‌شود، میانگین امتیاز مزه ۵۵، با انحراف معیار ۷ است، در حالی که گروه پپسی دارای ۵۷ نوجوان با میانگین امتیاز مزه اظهار شده ۶۴ و انحراف معیار ۱۳ است.

رابطه تجزیه و تحلیل این داده‌ها برای تعیین این‌که آیا تفاوت معنی‌داری در امتیاز مزه اظهار شده پسروان نوجوان برای این دو برنده وجود دارد، نیازمند استفاده صحیح از شش عدد است. اگر در پی تجزیه و تحلیل درست داده‌ها هستید باید: اندازه نمونه، میانگین و انحراف معیار هر یک از دو گروه را در اختیار داشته باشید و تمام این ۶ عدد به درستی در رابطه‌ها استفاده شوند.

قالب ارائه شده در شکل ۱-۵ مثال خوبی از یک جدول داده‌ها برای مرحله ۱ است:

¹ Coke

² Pepsi

	میانگین	انحراف معیار	
Group	n	Mean	STDEV
نام‌گذاری کنید	1 (name it)		
نام‌گذاری کنید	2 (name it)		

شکل ۵-۱: قالب جدول پایه برای آزمون t دو گروهی

برای مثال، اگر تصمیم بگیرید که گروه ۱ را گروه کوکا و گروه ۲ را گروه پیسی بنامید، همانطور که در شکل ۵-۲ نشان داده شد، شش عدد یادشده در سلول‌های مربوطه این جدول قرار داده شده است:

	میانگین	انحراف معیار	
Group	n	Mean	STDEV
نام‌گذاری کنید	1 (name it)	52	55
نام‌گذاری کنید	2 (name it)	57	64

شکل ۵-۲: نتایج واردکردن داده‌های مورد نیاز برای آزمون t دو گروهی

اکنون می‌توانید با اطمینان بیشتری رابطه‌های آزمون t دو گروهی را استفاده کنید، شش عدد در محل مناسب در رابطه‌ها قرار گرفته‌اند. توجه داشته باشید که می‌توانستید به آسانی گروه ۱ را گروه پیسی و گروه ۲ را گروه کوکا بنامید، این که شما چه تصمیمی برای نامیدن این گروه‌ها بگیرید، تفاوتی در نتیجه و نتیجه‌گیری ایجاد نخواهد کرد: این تصمیم‌گیری به عهده شما است.

۳-۱-۵- مرحله سوم: فرضیه صفر و فرضیه تحقیق برای آزمون t دو گروهی را بیان کنید اگر مرحله ۱ را کامل کرده‌اید، این مرحله بسیار آسان است، زیرا فرضیه صفر و فرضیه تحقیق همیشه برای آزمون t دو گروهی به صورت یکسان بیان می‌شوند. فرضیه صفر بیان می‌کند که میانگین‌های جامعه در دو گروه برابر است، درحالی که فرضیه تحقیق بیان می‌کند که میانگین جامعه در دو گروه برابر نیست. این آزمون فرضیه، به صورت زیر می‌باشد:

$$H_0: \mu_1 = \mu_2$$

$$H_1: \mu_1 \neq \mu_2$$

اگرچه می‌توانید بینید که استفاده از این نمادها بسیار ساده‌تر از نوشتن نام دو گروه در تمام رابطه‌ها است.

۴-۱-۵- مرحله چهارم: انتخاب آزمون آماری مناسب

از آنجاکه در این فصل وضعیتی مورد بررسی است که در آن با دو گروه از افراد سروکار دارید، اما فقط یک مورد اندازه‌گیری برای هر فرد در هر گروه وجود دارد، از آزمون t دو گروهی مستقل در این فصل استفاده خواهیم کرد.

۵-۱-۵- مرحله پنجم: تصمیم درباره قاعده تصمیم‌گیری برای آزمون t دو گروهی
قاعده تصمیم‌گیری دقیقاً همان چیزی است که در فصل گذشته زمانی که با آزمون t یک گروهی
مواجه بودیم، ارائه شد (به بخش ۳-۱-۴ مراجعه شود).

(الف) اگر مقدار قدر مطلق آماره t (محاسباتی) کمتر از مقدار بحرانی t باشد، فرضیه صفر را بپذیرید.

(ب) اگر مقدار قدر مطلق آماره t (محاسباتی) بزرگ‌تر از مقدار بحرانی t باشد، فرضیه صفر را رد کنید و فرضیه تحقیق را بپذیرید.

از آنجاکه در فصل قبل یاد گرفتید که چگونه مقدار قدر مطلق t را پیدا کنید (به بخش ۱-۳-۱-۴ مراجعه شود)، می‌توانید از آن آموخته‌ها در این فصل نیز استفاده کنید.

۶-۱-۵- مرحله ششم: محاسبه رابطه برای آزمون t دو گروهی

از آنجایی که بسته به اندازه نمونه افراد، در دو گروه، از دو رابطه مختلف در این فصل برای آزمون t دو گروهی استفاده می‌کنیم؛ در ادامه توضیح خواهیم داد که چگونه از این رابطه‌ها در این فصل استفاده شود.

۵-۷-۱-۵- مرحله هفتم: مقدار t بحرانی را از جدول t در ضمیمه E پیدا کنید

در فصل قبل که با آزمون t یک گروهی مواجه بودیم، ارزش بحرانی t از جدول t در پیوست E را با پیدا کردن اندازه نمونه برای یک گروه از افراد در ستون اول جدول و سپس خواندن مقدار بحرانی t در مقابل آن در ستون سمت راست "ستون بحرانی t " جدول، مشخص شد (مراجعةه کنید به بخش ۴-۱-۵). زمانی این فرایند نسبتاً ساده خواهد شد که شما این مرحله را تمرین کرده باشید. با این حال، برای آزمون t دو گروهی، روش پیدا کردن مقدار بحرانی t پیچیده تر است، زیرا شما دو گروه مختلف از افراد در مطالعه خود دارید، و اغلب آن‌ها دارای اندازه‌های مختلف نمونه در هر گروه هستند.

برای استفاده درست از ضمیمه E در این فصل، باید یاد بگیرید که چگونه درجه آزادی مربوط به مطالعه خود را پیدا کنید. اکنون این روند موربدبخت قرار می‌گیرد.

۵-۷-۱-۵- پیدا کردن درجه آزادی (df) برای آزمون t دو گروهی

هدف: پیدا کردن درجه آزادی برای آزمون t دو گروهی و استفاده از آن برای پیدا کردن مقدار بحرانی t از جدول t در ضمیمه E.

توضیح ریاضی مفهوم "درجه آزادی" فراتر از محدوده این کتاب است، اما شما می‌توانید اطلاعات بیشتری درباره این مفهوم با خواندن هر کتاب آماری خوب (مانند کلر ۲۰۰۹) کسب کنید. براساس اهداف تعریف شده، به راحتی می‌توانید در ک کنید که چگونه می‌توان درجه آزادی را پیدا کرده و از آن برای پیدا کردن مقدار بحرانی t در ضمیمه E استفاده کنید. رابطه درجه آزادی (df) برابر با:

$$df = n_1 + n_2 - 2 \quad (1-5)$$

به عبارت دیگر، اندازه نمونه گروه ۱ را با اندازه نمونه گروه ۲ جمع می‌کنید و سپس ۲ را از این مجموع کم می‌کنید تا عدد درجه آزادی را برای استفاده در ضمیمه E به دست آورید. به ضمیمه E نگاهی بیندازید.

به جای استفاده از ستون اول که در آزمون یک گروهی t بر اساس اندازه نمونه، n ، برای یک گروه از افراد، مقدار بحرانی t مشخص می‌شود، باید از ستون دوم این جدول (df) برای پیدا کردن مقدار بحرانی برای آزمون t دو گروهی استفاده کنیم.

به عنوان مثال، اگر شما ۱۳ نفر در گروه ۱ و ۱۷ نفر در گروه ۲ داشته باشید، درجه آزادی عبارت است از: $t = \frac{17+13-2}{20+48} = 2.8$ و مقدار بحرانی t معادل ۲/۰۴۸ خواهد بود. در جدول برای پیدا کردن مقدار بحرانی t وقتی $t = 2.8$ باشد، به اعداد ستون دوم جدول که در بردارنده درجه آزادی است نگاه می‌کنید تا زمانی که به عدد ۲.۸ برسید و سپس عدد ۲/۰۴۸ را در ستون مقدار بحرانی t بخوانید.

به عنوان مثال دوم، اگر شما ۵۲ نفر در گروه ۱ و ۵۷ نفر در گروه ۲ داشته باشید، درجه آزادی: $t = \frac{52+57-2}{107} = 1.07$ خواهد بود.

هنگامی که به پایین ستون دوم در ضمیمه E برای پیدا کردن درجه آزادی می‌روید، متوجه می‌شوید که بیشتر از درجه آزادی ۳/۹، مقدار بحرانی t همیشه برابر با ۱/۹۶ بوده و این ارزشی است که می‌توانید برای مقدار t بحرانی در این مثال استفاده کنید.

۱-۸- مرحله هشتم: نتیجه آزمون آماری خود را بیان کنید

ارائه نتیجه این آزمون از قالب نتایج آزمون t تک گروهی در فصل قبل تبعیت می‌کند (به بخش ۴-۶ مراجعه شود):

اگر مقدار قدر مطلق t محاسباتی که در رابطه t تعیین شد، کمتر از مقدار بحرانی t در ضمیمه E باشد، فرضیه صفر را قبول می‌کنید.

یا: اگر مقدار قدر مطلق t که در رابطه t تعیین شد، بیشتر از مقدار بحرانی t در ضمیمه E باشد، فرضیه صفر را رد می‌کنید و فرضیه تحقیق را می‌پذیرید.

۱-۹- مرحله نهم: نتیجه آزمون آماری خود را تفسیر کنید

نوشتن نتیجه‌گیری برای آزمون t دو گروهی دشوارتر از نوشتن نتیجه‌گیری برای آزمون t تک گروهی است زیرا شما باید تصمیم بگیرید که تفاوت بین دو گروه چیست.

زمانی که فرضیه صفر قبول می‌شود، نوشتن نتایج ساده است: تفاوتی بین دو گروه در متغیری که اندازه‌گیری شده وجود ندارد.

اما هنگامی که فرضیه صفر را رد می‌کنید و فرضیه تحقیق را می‌پذیرید، لازم است در نوشتن نتیجه‌گیری محاط باشید، به طوری که تفسیر دقیق و مختصر باشد.

اجازه دهید تمرين‌هایی را برای نوشتن نتیجه‌گیری آزمون t دو گروهی انجام دهیم.

۱-۹-۱- نتیجه‌گیری آزمون t دو گروهی زمانی که فرضیه صفر پذیرفته می‌شود

هدف: نوشتن نتیجه‌گیری آزمون t دو گروهی زمانی که فرضیه صفر پذیرفته شده است.

فرض کنید که به عنوان مشاور آماری توسط هتل ماریوت در سنت لوئیس استخدام شده اید تا داده های حاصل از نظرسنجی برای تعیین میزان رضایت مشتریان از فعالیت های مختلف هتل را تجزیه و تحلیل کنید. این نظرسنجی شامل تعدادی موارد است، اما فرض کنید که مورد شماره ۷ به شرح ارائه شده در شکل ۵-۳، مورد بررسی است:



شکل ۵-۳: نظرسنجی رضایت میهمانان هتل ماریوت، مورد شماره ۷

در ادامه فرض کنید که تصمیم گرفته اید داده های مشتریان هفت هفته گذشته را برای مقایسه نظر مردان در مقابل زنان با استفاده از آزمون t دو گروهی تحلیل کنید.

نکته مهم: برای هر یک از موارد مورد نظر (مانند مورد ۷) به طور جداگانه به این آزمون نیاز دارد.

فرض کنید داده های فرضی برای مورد ۷ از مشتریان هفت هفته گذشته هتل ماریوت سنت لوئیس شامل نمونه ای از ۱۲۴ مرد که دارای میانگین نمره ۶/۵۸ بود و انحراف معیار ۲/۴۴ بوده اند. همچنین، فرض کنید نظر ۸۶ مشتری زن هفته گذشته را در اختیار دارید که میانگین نمره آن ها ۶/۴۵ با انحراف معیار ۱/۸۶ است.

در ادامه این فصل توضیح خواهیم داد که چگونه نتایج آزمون t دو گروهی را با استفاده از رابطه های مربوطه تولید کنیم، اما در حال حاضر، بباید این کار را انجام نداده و نشان دهیم چگونه این رابطه ها براساس عده های موجود در شکل ۵-۴ تولید می شوند:

گروه	میانگین	انحراف معیار	
Group	n	Mean	STDEV
مردان	124	6.58	2.44
زنان	86	6.45	1.86

شکل ۵-۴: صفحه کاری داده‌ها برای مقایسه نظر میهمانان مرد در مقابل زن برای مثال هتل ماریوت سنت لوییس

درجه آزادی ۲۰۸
 مقدار بحرانی t ۱/۹۶ (در ضمیمه E)
 رابطه آزمون t : ۰/۴۴ (زمانی که شما از ماشین حساباتان استفاده کنید)
 نتیجه: مقدار قدر مطلق آماره t محاسباتی، ۰/۴۴ کمتر از مقدار t بحرانی (۱/۹۶) است، پس، فرضیه صفر را قبول می‌کنیم.
 نتیجه‌گیری: در خصوص رفتار صمیمانه کارمندان پذیرش هتل ماریوت سنت لوییس، تفاوتی بین نظر مردان و زنان میهمان در هفته گذشته وجود نداشت.
 حالا بباییید ببینیم زمانی که فرضیه صفر (H_0) را رد و فرضیه تحقیق (H_1) را قبول می‌کنید، چه اتفاقی می‌افتد.

۵-۱-۲-۹-۵- نتیجه‌گیری آزمون t دو گروهی زمانی که فرضیه صفر رد و فرضیه تحقیق پذیرفته می‌شود

هدف: نوشتن نتیجه‌گیری آزمون t دو گروهی زمانی که فرضیه صفر رد و فرضیه تحقیق پذیرفته می‌شود

اجازه دهید با مثال هتل ماریوت ادامه دهیم، اما با نتیجه‌های که ما فرضیه صفر را رد می‌کنیم و فرضیه تحقیق را می‌پذیریم.
 فرض کنیم که این بار اطلاعات مربوط به ۸۵ میهمان مرد هفته گذشته را در اختیار دارد و میانگین نمره آن‌ها در این مورد ۷/۲۶ با انحراف معیار ۲/۳۵ است. اطلاعات مربوط به ۴۸ میهمان زن هفته گذشته را نیز در اختیار دارید و میانگین نمره آن‌ها در این مورد ۴/۳۷ با انحراف معیار

۳/۲۶ است. بدون در نظر گرفتن جزئیات رابطه‌های آزمون t دو گروهی، داده‌های مربوط به این مثال را در شکل ۵-۵ ارائه می‌کنیم:

گروه		میانگین	انحراف معیار
Group	n	Mean	STDEV
مردان	1 Males	85	7.26
زنان	2 Females	48	4.37

شکل ۵-۵: صفحه کاری داده مثال هتل سنت لوئیس برای به دست آوردن تفاوت معنی‌دار بین نظر میهمانان مرد و زن

فرض صفر: $\mu_1 = \mu_2$

فرضیه تحقیق: $\mu_1 \neq \mu_2$

درجه آزادی ۱۳۱

مقدار t بحرانی: ۱/۹۶

رابطه آزمون t : ۵/۴۰ (زمانی که از ماشین حساب استفاده کنید)

نتیجه: از آنجایی که مقدار قدر مطلق، یعنی ۵/۴۰ (آماره محاسباتی) بیشتر از مقدار بحرانی ۱/۹۶ است، فرضیه صفر را رد می‌کنیم و فرضیه تحقیق را می‌پذیریم.

اکنون با استفاده از قانون زیر، باید نظر میهمانان مرد و زن را مقایسه کنید تا بدانید که کدام گروه دارای امتیاز بیشتر در خصوص برخورد صمیمانه کارکنان پذیرش هستند.

قانون: اگر فرضیه صفر را رد کرده و فرضیه تحقیق را پذیرفته بودیم، برای ارائه نتیجه‌گیری آزمون t دو گروهی، فقط میانگین دو گروه را مقایسه و اطمینان حاصل کنید که از کلمه "معنی‌دار" در نتیجه‌گیری استفاده کردۀ‌اید.

یک روش خوب جهت آماده شدن برای نوشتن نتیجه‌گیری آزمون t دو گروهی در هنگام استفاده از یک مقیاس رتبه‌بندی، قرار دادن میانگین نمرات دو گروه در تصویر مقیاس است تا بتوانید تفاوت میانگین نمرات را ببینید. برای نمونه، در مورد مثال هتل ماریوت، می‌توانید همانند تصویر ارائه شده در شکل ۵-۶، میانگین نمره گروه‌ها را رسم کنید:



شکل ۵-۶: مثال ایجاد یک "تصویر" از میانگین‌های دو گروه در مقیاس رتبه‌بندی

این نقاشی به صورت نشان می‌دهد که مردان دارای رتبه مثبت بیشتری نسبت به زنان در این مورد هستند (۷/۲۶ در مقابل ۴/۳۷)، و از آنجایی که شما فرضیه صفر را رد کردید و فرضیه تحقیق را قبول کردید، می‌دانید که بین دو نمره میانگین تفاوت معنی‌داری وجود دارد. بنابراین، نتیجه‌گیری ما باید شامل کلمات کلیدی زیر باشد:

- میهمانان مرد
- میهمانان زن
- هتل ماریوت
- سنت لوییس
- هفته گذشته
- معنی‌داری
- کارکنان پذیرش
- رفتار صمیمانه بیشتر یا رفتار صمیمانه کمتر
- (۷/۲۶ در مقابل ۴/۳۷) یا (۴/۳۷ در مقابل ۷/۲۶)

ما می‌توانیم از این کلمات کلیدی برای نوشتن هر یک از دو نتیجه‌گیری ذیل که از نظر منطقی یکسان هستند استفاده کنیم:

در هفته گذشته میهمانان مرد هتل ماریوت سنت لوییس رفتار کارکنان پذیرش را به صورت معنی‌داری صمیمانه‌تر در مقایسه با میهمانان زن، رتبه‌بندی کردند (۷/۲۶ در مقابل ۴/۳۷). یا: در هفته گذشته میهمانان زن هتل ماریوت سنت لوییس در مقایسه با میهمانان مرد، رفتار کارکنان پذیرش را به صورت معنی‌داری کمتر صمیمانه رتبه‌بندی کردند (۴/۳۷ در مقابل ۷/۲۶). هر دو نتیجه‌گیری صحیح است، بنابراین می‌توانید تصمیم بگیرید که کدام نتیجه‌گیری را بنویسید. این انتخاب با شما است.

همچنین، توجه داشته باشید که میانگین نمرات در داخل پرانتز در پایان این نتیجه‌گیری باید با ترتیب ارائه دو گروه درنتیجه‌گیری شما منطبق باشد. به عنوان مثال، اگر شما می‌گویید: میهمانان مرد در مقایسه با میهمانان زن، رفتار کارکنان پذیرش را به طور معنی‌داری صمیمانه‌تر ارزیابی کردند-

اند. پایان این نتیجه باید این گونه باشد: (۷/۲۶ در مقابل ۴/۳۷) زیرا شما اول به مردان و بعد به زنان اشاره کردید.

به طریق مشابه، اگر شما می‌گویید: میهمانان زن در قیاس با میهمانان مرد، رفتار کارکنان پذیرش را به طور معنی‌داری کمتر صمیمانه ارزیابی کرده‌اند، باید بگویید: (۴/۳۷ در مقابل ۷/۲۶) زیرا اول به زنان و بعد به مردان اشاره کردید.

قرار دادن میانگین نمرات مردان و زنان در پایان نتیجه‌گیری، باعث می‌شود که خواننده برای پیدا کردن این‌که تا چه حد این میانگین‌ها از هم دور هستند به جدول نتایج در تحقیق رجوع نکند. اکنون باید درباره رابطه ۱ که مربوط به وضعیتی است که در آن هر دو گروه دارای جمعیتی بیشتر از ۳۰ نفر هستند، بحث کنیم.

هدف: استفاده از رابطه شماره ۱ برای آزمون t دو گروهی، زمانی که هر دو گروه دارای اندازه نمونه بزرگ‌تر از ۳۰ نفر باشند

۲-۵- رابطه شماره ۱: هر دو گروه دارای جمعیتی بیش از ۳۰ نفر باشند

رابطه نخست که ما در مورد آن بحث می‌کنیم، زمانی استفاده می‌شود که ما دو گروه از افراد را داشته باشیم که هر کدام دارای جمعیتی بیشتر از ۳۰ نفر بوده و یک اندازه‌گیری برای هر فرد در هر گروه وجود داشته باشد. این رابطه برای آزمون t دو گروهی به قرار زیر است:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}} \quad (2-5)$$

$$\text{که} \quad S_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}} \quad (3-5)$$

$$\text{و درجه آزادی برابر است با} \quad df = n_1 + n_2 - 2 \quad (1-5)$$

زمانی که برای بار اول این رابطه را می‌بینید به نظر هولناک می‌رسد، اما باید بخش‌هایی از این رابطه را توضیح دهیم:

پیشتر مفهوم درجه آزادی را در این فصل توصیح دادیم، و در نتیجه باید قادر باشید که درجه آزادی را که برای رابطه مورد نیاز است را بیابید تا مقدار بحرانی t را در ضمیمه E پیدا کنید. در فصل گذشته، رابطه برای آزمون t یک گروهی به صورت زیر بود:

$$t = \frac{\bar{X} - \mu}{S_{\bar{X}}} \quad (1-4)$$

$$s.e. = S_{\bar{X}} = \frac{S}{\sqrt{n}} \quad (2-4)$$

برای آزمون t یک گروهی، میانگین نمرات را پیدا نموده و از میانگین جامعه کم کردید، و بر خطای معیار میانگین تقسیم کردید تا نتایج آزمون t را به دست آورید. سپس نتایج آزمون t (محاسباتی) را با مقدار بحرانی t به جهت پذیرش فرضیه صفر، یا رد فرضیه صفر و پذیرش فرضیه تحقیق، مقایسه کردید.

آزمون t دو گروهی نیازمند رابطه متفاوتی است به این خاطر که دو گروه از افراد را دارید، که هر کدام میانگین متفاوتی دارند. سعی در این دارید که مشخص کنید آیا فرضیه صفر را که مبنی بر این است که میانگین جامعه در دو گروه مساوی است را قبول کنید (به عبارت دیگر، تفاوت معنی داری بین میانگین دو گروه وجود ندارد)، یا اینکه تفاوت بین دو میانگین دو گروه به اندازه کافی زیاد است که باعث می شود شما بپذیرید که تفاوت معنی دار بین میانگین دو گروه وجود دارد. در صورت کسر آزمون t دو گروهی از شما خواسته می شود که تفاوت بین میانگین دو گروه را بدست آورید:

$$\bar{X}_1 - \bar{X}_2 \quad (4-5)$$

قدم بعدی در این رابطه برای آزمون t دو گروهی این است که جوابی را که از کم کردن دو میانگین به دست می آید بر خطای معیار تفاوت میان میانگین دو گروه، تقسیم کنید و این خطای معیار متفاوتی است از آنچه شما در آزمون t یک گروهی به دست آورید به این خاطر که در آزمون t دو گروهی، دو میانگین وجود دارد.

خطای استاندارد میانگین زمانی که شما دو گروه از افراد را دارید "خطای استاندار تفاوت میانگین ها" بین دو گروه نامیده می شود. این رابطه زمانی که به چهار گام تجزیه شود کمتر ترسناک به نظر می رسد:

- ۱- انحراف معیار گروه ۱ را به توان دوم برسانید، سپس نتیجه را بر اندازه نمونه گروه ۱ (n_1) تقسیم کنید.
- ۲- انحراف معیار گروه ۲ را به توان دوم برسانید، سپس نتیجه را بر اندازه نمونه گروه ۲ (n_2) تقسیم کنید.
- ۳- نتایج بالا را به یکدیگر اضافه کنید تا مقدار کلی را به دست آورید.

-۴- از این مقدار نهایی برای یافتن خطای معیار اختلاف میانگین جذر بگیرید.

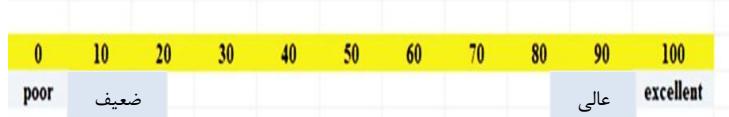
$$S_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

برای دانشجویان مرحله آخر بیشترین مشکل را زمانی که آنها به دنبال یافتن این خطای معیار توسط ماشین حساب هستند، به وجود می‌آورد، زیرا آنها عجله زیادی برای یافتن پاسخ داشته و فراموش می‌کنند ریشه دوم (جذر) را که در گام آخر به آن اشاره شد، را انجام دهند. بنابراین، عدد بزرگتری را برای خطای معیار نسبت به آن چیزی که باید به دست می‌آمد، به دست می‌آورند.

۵-۱-۲-۵- مثالی از رابطه شماره ۱ برای آزمون t دو گروهی

اجازه دهید رابطه ۱ را برای وضعیتی که هر دو گروه دارای اندازه نمونه بیشتر از ۳۰ نفر هستند مورداستفاده قرار دهیم.

تصور کنید که شما توسط شرکت پیسی استخدام شده‌اید تا مزه نوشابه‌ها را در بین نوجوانان (بین ۱۳ تا ۱۸ سال) به منظور مشخص کردن این موضوع که آیا آنها مزه پیسی را همانند کوکا دوست دارند یا نه، بررسی کنید. پسرها برنده نوشیدنی را که مزه کردند را نگفته‌اند. گروهی از پسران که در این بازه سنی قرار دارند را انتخاب می‌کنید و به طور تصادفی در هر یک از دو گروه قرار می‌دهید: (۱) گروه ۱ کوکا را مزه می‌کنند و (۲) گروه ۲ پیسی را مزه می‌کنند. هر گروه به مزه نوشیدنی خود در یک مقیاس ۱۰۰ نمره ای مانند مقیاس شکل (۷-۵) نمره می‌دهند:



شکل ۷-۵: مثالی از رتبه‌دهی برای آزمون مزه نوشیدنی غیر الکلی (مثال عملی)

فرض کنید که شما نمره‌ها را جمع‌آوری و مشخص گردید (از مهارت‌های جدید اکسل خود بهره بگیرید) ۵۲ نفر از پسرها در گروه کوکا، دارای میانگین نمرات ۵۵ و انحراف معیار ۷ هستند، در حالی که ۵۷ نفر از پسرها در گروه پیسی، دارای میانگین نمرات ۶۴ و انحراف معیار ۱۳ می‌باشند. توجه کنید که در آزمون t دو گروهی، نیازی نیست که هر دو گروه اندازه نمونه مشابهی داشته باشند. به بیان دیگر، آزمون t دو گروهی "قدرمند است" (یک اصطلاح خیالی که متخصصان آمار علاقه به استفاده از آن دارند).

سپس داده‌های جدول زیر را مشابه شکل (۸-۵) ایجاد کنید:

	گروه	میانگین	انحراف معیار	
	Group	n	Mean	STDEV
کوکا	1 Coke	52	55	7
پپسی	2 Pepsi	57	64	13

شکل ۸-۵: صفحه کاری برای آزمون مزه نوشیدنی غیر الکلی

یک صفحه گسترده اکسل ایجاد کرده و اطلاعات زیر را وارد کنید:

گروه:B3

کوکا:B4

پپسی:B5

اندازه:C3

میانگین:D3

انحراف معیار:E3

۵۲ :C4

۵۵ :D4

۷ :E4

۵۷ :C5

۶۴ :D5

۱۳ :E5

اکنون، ستون B را عرض تر کنید به طوری که دو برابر ستون A باشد، و ۶ عدد و نماد هر کدام را در جدول خود قرار دهید (به شکل ۹-۵ مراجعه شود)

A	B	C	D	E	F
۱	گروه		میانگین	انحراف معیار	
۲		Group	n	Mean	STDEV
کوکا	1 Coke	52	55	7	
پپسی	2 Pepsi	57	64	13	

شکل ۹-۵: نتایج عرض کردن ستون B و گذاشتن اعداد در مرکز سلولها

فرضیه صفر:B8

B10: فرضیه تحقیق

از آنجایی که هر دو گروه دارای اندازه نمونه بالای ۳۰ نفر می‌باشند، شما باید از رابطه شماره ۱ برای آزمون t اختلاف میانگین دو گروه استفاده کنید.

باید در ادامه این رابطه را به اجزایی تجزیه کرده تا شанс اشتباه کردن را کاهش دهیم.

$\frac{STDEV1}{n_1}$: B13 (توجه داشته باشید که شما انحراف معیار گروه یک را به توان دو می‌رسانید و نتیجه را بر اندازه نمونه گروه ۱ تقسیم می‌کنید).

$\frac{STDEV2}{n_2}$: B16

D13+D16 : B19

B22: خطای معیار

B28: مقدار بحرانی t

B31: نتیجه

B36: نتیجه‌گیری (به شکل ۵-۱۰ مراجعه شود)

گروه		میانگین		انحراف معیار
کوکا	پپسی	n	Mean	STDEV
1 Coke	52	55	7	
2 Pepsi	57	64	13	
Null hypothesis:		فرضیه صفر		
Research hypothesis:		فرضیه تحقیق		
STDEV1 squared / n1		انحراف معیار گروه اول به توان 2 اندازه نمونه اول		
STDEV2 squared / n2		انحراف معیار گروه دوم به توان 2 اندازه نمونه دوم		
D13 + D16				
s.e.		خطای معیار		
critical t		آماره t بحرانی		
t-test		آماره t محاسباتی		
Result:		نتیجه		
Conclusion:		نتیجه‌گیری		

شکل ۱۰-۵: برچسب‌های رابطه برای آزمون t دو گروهی

اکنون شما باید مقادیر فرمول‌های بالا در سلول‌های زیر محاسبه کنید:

D13: نتیجه فرمول برای محاسبه سلول B13 (از دو رقم اعشار استفاده کنید)

D16: نتیجه فرمول برای محاسبه سلول B16 (از دو رقم اعشار استفاده کنید)

D19: نتیجه فرمول برای محاسبه سلول B19 (از دو رقم اعشار استفاده کنید)

D22: $=SQRT(D19)$ (از دو رقم اعشار استفاده کنید)

این رابطه باید خطای معیاری برابر با ۱/۹۸ ارائه دهد.

t آزمون ۱/۹۶ (از آنجایی که $df = n_1 + n_2 - 2$ که حاصل آن $107 = 109 - 2$ و مقدار $D25$ بحرانی طبق ضمیمه E برابر ۱/۹۶ خواهد بود)

$$= \frac{(D4 - D5)}{D22} \quad D28$$

این رابطه باید مقدار آماره (محاسباتی) آزمون t را معادل ۴/۵۵- ارائه دهد.
بررسی کنید که همهی اعداد را در D28 تا D13 تا دو رقم اعشار گرد کرده‌اید (به شکل ۱۱-۵ مراجعه شود)

A	B	C	D	E
11				
انحراف معیار گروه اول به توان ۲ اندازه نمونه اول	STDEV1 squared / n1		0.94	
انحراف معیار گروه دوم به توان ۲ اندازه نمونه دوم	STDEV2 squared / n2		2.96	
19	D13 + D16		3.91	
20				
خطای معیار	s.e.		1.98	
23				
آماره t بحرانی	critical t		1.96	
24				
آماره t محاسباتی	t-test		-4.55	
25				
26				
27				

شکل ۱۱-۵: نتایج رابطه آزمون t برای بررسی مزه نوشیدنی غیر الکلی

اکنون، جملات زیر را در D34 تا D31 برای خلاصه کردن نتایج مطالعه بنویسید:

از آنجایی که مقدار قدر مطلق $-4/55$ D31

t : بزرگتر از مقدار بحرانی D32

۱/۹۶ بوده، فرضیه صفر را رد می‌کنیم. D33

فرضیه تحقیق را موردنظر پذیرش قرار می‌دهیم D34

پسران نوجوان مزه D36

D37: پیسی را به طور معنی‌داری بهتر امتیازگذاری کردنD

D38: در قیاس با کوکا (۶۴ در مقابل ۵۵)

در آخر، جملات زیر را برای D38 تا D36 جهت خلاصه کردن نتیجه‌گیری مطالعه بنویسید.

فایل خود را با نام COKE4 ذخیره کنید.

این فایل را طوری چاپ کنید که دقیقاً در یک صفحه جا شود، و فرضیه‌های صفر و تحقیق را روی برگ چاپ شده خود دستی بنویسید.

صفحه گسترده نهایی در شکل (۱۲-۵) نشان داده می‌شود.

گروه	میانگین	انحراف معیار
کوکا		
پپسی		
فرضیه صفر	$\mu_1 = \mu_2$	
فرضیه تحقیق	$\mu_1 \neq \mu_2$	
انحراف معیار گروه اول به توان 2 اندازه نمونه اول	STDEV1 squared / n1	0.94
انحراف معیار گروه دوم به توان 2 اندازه نمونه دوم	STDEV2 squared / n2	2.96
خطای معیار	D13 + D16	3.91
آماره t بحرانی	s.e.	1.98
آماره t محاسباتی	critical t	1.96
نتیجه	t-test	-4.55
نتیجه‌گیری	Result:	Since the absolute value of - 4.55 is greater than the critical t of 1.96, we reject the null hypothesis and accept the research hypothesis.
نتیجه‌گیری	Conclusion:	Teenage boys rated the taste of Pepsi as significantly better than the taste of Coke (64 vs. 55)
پسران نوجوان مزه پپسی را به طور معنی‌داری بهتر از کوکا رتبه‌بندی کردند (۶۴ در مقابل ۵۵)		

شکل ۱۲-۵: صفحه‌کاری نهایی برای آزمون مزه کوکا در مقابل پپسی

اکنون اجازه دهید، رابطه دوم آزمون t دو گروهی را استفاده کنیم، که می‌توانیم آن را هر زمانی که هر دو گروه یا یکی از آن‌ها دارای جمعیتی کمتر از ۳۰ نفر باشد، مورداستفاده قرار دهیم.

هدف: استفاده از رابطه ۲ برای آزمون t دو گروهی زمانی که یکی یا هر دو گروه دارای جمعیتی زیر ۳۰ نفر باشند.

اکنون باید نگاهی به موردی بیندازیم که در آن یکی یا هر دو گروه دارای جمعیتی کمتر از ۳۰ نفر هستند.

۳-۵- رابطه شماره ۲: یکی یا هر دو گروه دارای جمعیتی کمتر از ۳۰ نفر باشند

فرض کنید که شما برای تولید کننده دستگاههای پخش MP3 کار می‌کنید و از شما خواسته شده است که یک ارزیابی قیمت‌گذاری انجام دهید تا امکان فروش واحدهای بیشتر را با کاهش قیمت بررسی کنید.

علاوه براین، فرض کنید به طور تصادفی ۷ عدد فروش را برای خرید محصول در قیمت رایج انتخاب کرده‌اید، که میانگین خرید آن‌ها ۱۱۷/۷ واحد با انحراف معیار ۱۹/۹ است. همچنین، به صورت تصادفی یک گروه دیگر از ۸ عدد فروش را برای خرید محصول با ۱۰ درصد تخفیف قیمت انتخاب کرده‌اید و میانگین خرید آن‌ها ۱۲۵/۱ واحد با انحراف معیار ۱۵/۱ واحد است.

می‌خواهید آزمون کنید که آیا دو قیمت مختلف باعث ایجاد اختلاف معنی‌داری در تعداد واحدهای MP3 فروخته شده، گردیده است یا خیر.

تصمیم گرفتید که آزمون t دو گروهی را برای نمونه‌های مستقل استفاده کنید، و داده‌های زیر در قالب شکل (۱۳-۵) حاصل شده است:

	گروه	میانگین	انحراف معیار
قیمت رایج	Group	n	STDEV
قیمت کاهش‌یافته	1 Regular Price	7	19.9
	2 Reduced price	8	15.1

شکل ۱۳-۵: صفحه کاری برای مقایسه قیمت عمده‌فروشی (مثال عملی)

فرضیه صفر: $\mu_1 = \mu_2$

فرضیه تحقیق: $\mu_1 \neq \mu_2$

توجه: از آنجایی که هر دو گروه دارای اندازه نمونه کم تر از ۳۰ نفر هستند، شما باید از رابطه ۲ مطابق گام های زیر استفاده کنید:

یک صفحه گسترده اکسل ایجاد کرده و اطلاعات زیر را وارد کنید:

B3: گروه

B4: ۱ قیمت معمول

B5: ۲ قیمت کاهش یافته

(تعداد) C3

D3: میانگین

E3: انحراف معیار

اکنون، ستون B را عریض تر کنید به طوری که سه برابر ستون A عرض داشته باشد.

برای این کار بر روی B در گوشه چپ صفحه گسترده خود کلیک کرده تا تمام سلولهای موجود در ستون B پرنگ شوند. سپس، اشاره گر ماوس را به سمت راست و انتهای سلول B برد و تا زمانی که علامت "صلیب" ظاهر شود: سپس روی این علامت صلیب کلیک کنید و علامت را به سمت راست بکشید تا بتوانید تمام کلمات را در صفحه خود بخوانید. سپس کلیک کردن را متوقف کنید.

۷ :C4

۱۱۷/۷ :D4

۱۹/۹ :E4

۸ :C5

۱۲۵/۱ :D5

۱۵/۱ :E5

سپس، اطلاعات را در سلولهای C3 تا E5، با انتخاب کردن این سلولها و سپس با استفاده از این گام، در مرکز سلول قرار دهید:

روی خط پایین، از نماد سمت چپ، در زیر "تراز" در مرکز بالای صفحه اصلی کلیک کنید.

B8: فرضیه صفر

B10: فرضیه تحقیق: (به شکل ۱۴-۵ مراجعه شود)

	A	B	C	D	E	F	G
1							
2		گروه		میانگین	انحراف معیار		
3		Group	n	Mean	STDEV		
4	قیمت رایج	1 Regular Price	7	117.7	19.9		
5	قیمت کاهش‌یافته	2 Reduced Price	8	125.1	15.1		
6							
7	فرضیه صفر	Null hypothesis:					
8							
9	فرضیه تحقیق	Research hypothesis:					
10							
11							

شکل ۱۴-۵: صفحه کاری مقایسه قیمت عمده فروشان برای آزمون فرضیه

از آنجایی که هر دو گروه دارای اندازه نمونه کمتر از ۳۰ نفر می‌باشند، باید از رابطه شماره ۲ برای آزمون t اختلاف میانگین دو نمونه مستقل استفاده کنید. رابطه شماره ۲ برای آزمون t دو گروهی به قرار زیر است:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}} \quad (2-5)$$

$$\text{که در آن} \quad S_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}} \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \quad (5-5)$$

$$\text{و درجه آزادی آن} \quad = df = n_1 + n_2 - 2 \quad (1-5)$$

رابطه بالا پیچیده است، اگر، به جای اینکه آن را به صورت یک سلول ورودی بنویسید آن را به اجزاء کوچکتر تجزیه کنید، شناس ارتکاب اشتباه در نوشتن رابطه را کاهش خواهد داد.

اکنون، این کلمات را در صفحه گستردۀ خود وارد کنید:

۱: $B13 \times (n_1 - 1)$ توان دوم انحراف معیار گروه ۱

۲: $B16 \times (n_2 - 1)$ توان دوم انحراف معیار گروه ۲

$n_1 + n_2 - 2 : B19$

$\frac{1}{n_1} + \frac{1}{n_2} : B22$

(s.e.): $B25$ خطای معیار

B28: مقدار بحرانی t B31: آزمون t

B34: نتیجه

B39: نتیجه‌گیری (به شکل ۱۵-۵ مراجعه شود)

گروه	میانگین	انحراف معیار
قیمت رایج	Group 1 Regular Price 2 Reduced Price	n Mean STDEV
قیمت کاهش‌یافته		7 117.7 19.9 8 125.1 15.1
فرضیه صفر	Null hypothesis:	
فرضیه تحقیق	Research hypothesis:	
	(n1 - 1) x STDEV1 squared	(۱- تعداد اعضای گروه اول) * توان دوم انحراف معیار گروه اول
	(n2 - 1) x STDEV2 squared	(۱- تعداد اعضای گروه دوم) * توان دوم انحراف معیار گروه دوم
	n1 + n2 - 2	
	1/n1 + 1/n2	
خطای معیار	s.e.	
آماره t بحرانی	critical t	
آماره t محاسباتی	t-test	
نتیجه	Result:	
نتیجه‌گیری	Conclusion:	

شکل ۱۵-۵: برچسب‌های رابطه مقایسه قیمت عمده‌فروشان برای آزمون t دو گروهی

باید مقادیر بالا که در رابطه آمده است را در سلول های زیر محاسبه کنید:

E13: نتیجه رابطه نیازمند محاسبه سلول B13 می باشد (از دو رقم اعشار استفاده کنید)

E16: نتیجه رابطه نیازمند محاسبه سلول B16 می باشد (از دو رقم اعشار استفاده کنید)

E19: نتیجه رابطه نیازمند محاسبه سلول B19 می باشد

E22: نتیجه رابطه نیازمند محاسبه سلول B22 می باشد (از دو رقم اعشار استفاده کنید)

$$= \text{SQRT} (((E13 + E16) / E19) * E22) \quad :E25$$

به سه پرانتز باز شده پس از SQRT، و سه پرانتز بسته در سمت راست این رابطه توجه داشته باشید.

شما به سه پرانتز باز و سه پرانتز بسته در این فرمول نیاز دارید، در غیر این صورت رابطه به درستی عمل نمی کند.

رابطه بالا یک خطای اختلاف میانگین ها معادل ۹/۰۵ را عاید می کند (دو رقم اعشار).

E28: مقدار t بحرانی را که در جدول t ضمیمه E قرار دارد را در این سلول وارد کرده و برای یافتن مقدار بحرانی t از $df = n1+n2-2$ استفاده کنید.

$$\frac{(D4-D5)}{E25} \quad :E31$$

توجه داشته باشید که شما نیاز به یک پرانتز باز قبل از D4 و یک پرانتز بسته پس از D5 دارید، به طوری که این محاسبه پاسخ ۷/۴۰ - را ارائه می دهد و سپس این مقدار تقسیم بر خطای معیار تفاوت میانگین معادل ۹/۰۵ شده، تا مقدار t یعنی ۰/۸۲ - را عاید کند (اینجا به نشانه منفی توجه کنید). از دو رقم اعشار برای نتیجه آماره t استفاده کنید (به شکل ۱۶-۵ مراجعه شود).

گروه		میانگین	انحراف معیار
قیمت رایج	Group		
قیمت کاهش یافته	1 Regular Price	7	117.7 19.9
	2 Reduced Price	8	125.1 15.1
فرضیه صفر	Null hypothesis:		
فرضیه تحقیق	Research hypothesis:		
(۱) - تعداد اعضای گروه اول) *	(n1 - 1) x STDEV1 squared		2376.06
توان دوم انحراف معیار گروه اول			
(۲) - تعداد اعضای گروه دوم) *	(n2 - 1) x STDEV2 squared		1596.07
توان دوم انحراف معیار گروه دوم			
خطای معیار	n1 + n2 - 2		13
	1/n1 + 1/n2		0.27
آماره t بحرانی	s.e.		9.05
آماره t محاسباتی	critical t		2.160
نتیجه	t-test		-0.82
نتیجه گیری	Result:		
	Conclusion:		

شکل-۵: نتایج رابطه آزمون t دو گروهی مقایسه قیمت عمدۀ فروشان

اکنون جملات زیر را برای خلاصه نتایج تحقیق در ستون‌های D37 تا D34 بنویسید:

D34: از آنجایی که مقدار قدر مطلق

D35: t یعنی -0.82 کمتر از

D36: مقدار بحرانی t یعنی 2.160 می‌باشد

D37: فرضیه صفر را قبول می‌کنیم.

در آخر، جملات زیر را برای خلاصه کردن نتایج تحقیق در ستون‌های D39 تا D43 بنویسید:

D39: تفاوتی بین

D40: تعداد واحدهای

D41: MP3 های فروخته شده

D42: در دو قیمت وجود ندارد. بنابراین، شما نباید

D43: قیمت را کاهش دهید.

فایل خود را با عنوان MP4 ذخیره کنید.

صفحه گسترده نهایی را طوری چاپ کنید که در یک صفحه قرار می‌گیرد.

فرضیه صفر و فرضیه تحقیق را بر روی برگه چاپ شده خود به صورت دستی بنویسید.

صفحه گسترده نهایی در شکل (۵-۱۷) نمایش داده شده است.

		گروه	میانگین	انحراف معیار
قیمت رایج	Group	n	Mean	STDEV
	1 Regular Price	7	117.7	19.9
قیمت کاهش یافته	2 Reduced Price	8	125.1	15.1
فرضیه صفر	Null hypothesis:		$\mu_1 = \mu_2$	
فرضیه تحقیق	Research hypothesis:		$\mu_1 \neq \mu_2$	
	(n1 - 1) x STDEV1 squared		2376.06	
	(n2 - 1) x STDEV2 squared		1596.07	
	n1 + n2 - 2		13	
	1/n1 + 1/n2		0.27	
خطای معیار	s.e.		9.05	
آماره t بحرانی	critical t		2.160	از آنجا که مقدار قدر مطلق آماره t محاسباتی یعنی -0.82 کمتر از مقدار بحرانی یعنی 2.16 است، فرضیه صفر پذیرفته می شود
آماره t محاسباتی	t-test		-0.82	
نتیجه	Result:		Since the absolute value of t of -0.82 is less than the critical t of 2.160, we accept the null hypothesis.	
نتیجه گیری	Conclusion:		There was no difference in the number of units of MP3 players sold at the two prices. So, you should not reduce the price!	تفاوت معنی داری در تعداد های فروخته شده در MP3 دو قیمت وجود ندارد. پس، نباید قیمت را کاهش داد.

شکل ۱۷-۵: صفحه گستردگی مقایسه قیمت عمده فروشان

۴-۵- تمرین‌های عملی پایان فصل

۱. فرض کنید شرکت بوئینگ شما را استخدام کرده است تا تجزیه و تحلیل داده‌های میدانی نظرسنجی‌هایی با عنوان نگرشی را انجام دهید. این نظرسنجی‌ها از مدیران این شرکت در طول ماه گذشته صورت گرفته است. موارد باهم جمع شدند تا نمره کلی به دست آید که در آن نمره بالا نشان‌دهنده رضایت شغلی بالا است، در حالی که نمره پایین نشان‌دهنده رضایت شغلی پایین کارکنان است.

شما نمونه‌ای تصادفی از ۲۰۲ زن با میانگین نمرات ۸۴/۸۰ و انحراف معیار ۵/۱۰ را انتخاب می‌کنید. همچنین، نمونه‌ای دیگر از ۲۴۱ مرد حاضر در این بررسی میدانی که دارای میانگین نمره ۸۸/۲۰۸ با انحراف معیار ۴/۳۰ بوده را برمی‌گزینید.

(a) فرضیه صفر و فرضیه تحقیق را در صفحه گسترده اکسل بیان کنید.

(b) خطای معیار تفاوت بین میانگین‌ها را توسط اکسل بیابید.

(c) مقدار آماره t بحرانی را با استفاده از ضمیمه E بیابید، و در داخل صفحه گسترده خود وارد کنید.

(d) آزمون t را بر روی این داده‌ها توسط اکسل انجام دهید. مقدار آماره t محاسباتی چقدر است؟ از سه رقم اعشار برای تمام ارقام در بخش رابطه صفحه گسترده استفاده کنید.

(e) نتایج را بر روی صفحه گسترده بیاورید.

(f) نتیجه‌گیری را به زبان ساده بر روی صفحه گسترده بیاورید.

(g) فایل را با نام Boeing3 ذخیره کنید.

۲. گروه مالی ماساچوست یک تبلیغ تمام صفحه رنگی را در مجله وال استریت انجام داد که در آن یک مدل مرد دو دختر دو ساله را بغل کرده بود. این تبلیغ دارای یک عنوان و یک زیر عنوان بود: چه چیزی نشانه یک تصمیم خوب است؟

دانستن اینکه زندگی شما بیمه شده است و می‌توانید برای دوران بازنشستگی درآمدی را تأمین کنید تا آرامش ذهنی در آن مقطع زمانی داشته باشید.

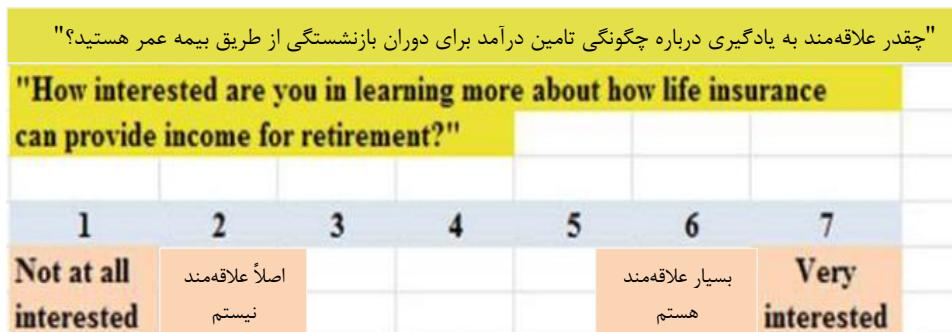
از آنجایی که اکثر مشترکان مجله وال استریت مرد هستند، پرسش جالب تحقیق به صورت زیر خواهد بود:

پرسش تحقیق: "آیا مدل مردی که در مجله بود توانست اثری بر روی تمایل مردان و زنان بالغ برای یادگیری بیشتر درباره اینکه چگونه بیمه عمر می‌تواند درآمدی برای بازنشستگی تامین کند، داشته باشد؟"

تصور کنید که شما گروهی از مردان (در سن ۲۵-۳۹) و گروهی از زنان (در سن ۲۵-۳۹) را به عنوان مخاطب تبلیغ در اختیاردارید به طوری که هر دو گروه تبلیغ را با مدل مرد دیده باشند کپی تبلیغ نشان داده شده به این افراد کاملاً واضح بود.

دو گروه هنگام انجام آزمایش از یکدیگر جدا نگه داشته شدند به طوری که نتوانند بر یکدیگر اثر بگذارند.

بعد از یک ساعت بحث درباره مدل تبلیغ، از پاسخدهندگان پرسشی که در جدول ۱۸-۵ آمده است، پرسیده شد:



شکل ۱۸-۵: مقیاس امتیازدهی برای شاخص علاقه‌مندی به تبلیغ مجله (مثال عملی)

نتیجه داده‌های حاصل برای این پرسش در شکل ۱۹-۵ آرائه شده است:

تبلیغ مجله: مدل مرد			
Magazine ad: Male model			
مردان	Men	Women	زنان
	5	3	
	6	4	
	4	6	
	7	5	
	5	2	
	6	3	
	5	1	
	4	3	
	3	2	
	6	4	
	7	3	
	5	5	
	6	6	
	4	3	
	7	4	
	5	2	
	4	5	
	6	3	
	3	4	
	7	5	
	5	4	
	6	3	
	2	2	
	6	4	
	1	3	
	7	5	
	6	1	
	5	3	
	4	2	
	6	3	
	5	2	
	7	5	
	3		
	4		

شکل ۱۹-۵: داده صفحه کاری برای فصل ۵: تمرین شماره ۲

- (a) بر روی صفحه گسترده اکسل خود، فرضیه صفر و فرضیه تحقیق را بنویسید.
- (b) جدولی تشکیل دهید که این دادهها را بر روی صفحه گسترده تان خلاصه کند و از اکسل برای یافتن اندازه نمونه‌ها، میانگین‌ها، و انحرافات معیار برای دو گروه در جدول استفاده کنید.
- (c) از اکسل برای پیدا کردن خطای معیار اختلاف میانگین‌ها استفاده کنید.
- (d) از اکسل برای اجرای آزمون t دو گروهی استفاده کنید. مقدار t محاسباتی را چقدر به دست آورده‌اید؟ (از دو رقم اعشار استفاده کنید)
- (e) در صفحه گسترده خود، مقدار t بحرانی را تایپ کرده و برای این کار از جدول t در ضمیمه E استفاده کنید.
- (f) نتایج را بر روی صفحه گسترده تایپ کنید.
- (g) نتیجه‌گیری خود را به صورت ساده روی صفحه گسترده بنویسید.
- (h) فایل را با عنوان lifeinsurance12 ذخیره کنید.

۳. هواپیمایی آمریکایی یک وعده غذایی در پرواز ارائه می‌دهد که مسافران می‌توانند با پرداخت ۸ دلار آن را خریداری کنند. از مشتریان در قالب یک نظرسنجی خواسته شد تا نظراتشان را درباره غذا ارائه کنند. از مسافران خواسته شد که احتمال خرید این وعده غذایی را در پرواز آتی در یک مقیاس ۵ نقطه‌ای رتبه‌بندی کنند. اما تصور کنید که شما شرکت هواپیمایی را قانع کردید که ارزیابی میدانی قصد خرید را به مقیاس ۷ نقطه‌ای تغییر دهد. پاسخ به پرسش قصد خرید در شکل (۲۰-۵) به تصویر کشیده شده است:



شکل ۲۰-۵: مقیاس امتیازدهی در نظرسنجی برای وعده غذایی در پرواز شرکت هواپیمایی آمریکایی (مثال عملی)

از مسافران در این نظرسنجی خواسته شد تا بیان کنند که آیا مسافران کلاس تجاری هستند یا عادی؟

فرض کنید میانگین امتیاز ۶۴ مسافر کلاس تجاری در ماه گذشته ۳/۲۳ با انحراف معیار ۱/۰۴ بوده است، در حالی که مسافران عادی دارای میانگین رتبه‌دهی ۲/۳۶ با انحراف معیار ۱/۳۵ می‌باشد.

(a) فرضیه صفر و فرضیه تحقیق را بر روی صفحه گسترده اکسل بیان کنید.

(b) خطای معیار اختلاف میانگین‌ها را با استفاده از اکسل بیابید.

(c) مقدار t بحرانی را با استفاده از ضمیمه E یافته، و در صفحه گسترده وارد کنید.

(d) آزمون t را بر روی این داده‌ها با استفاده از اکسل انجام دهید. مقدار t محاسباتی را چقدر بهدست آورده‌اید؟

(e) نتایج را بر روی صفحه گسترده بیان کنید.

(f) نتیجه‌گیری را به صورت ساده بر روی صفحه گسترده بیان کنید.

(g) فایل را با عنوان Ameal3 ذخیره کنید.

منابع

- Keller G. Statistics for management and economics. 8th ed. Mason: South-Western Cengage Learning; 2009.
- Mass Mutual Financial Group. What is the sign of a good decision? (advertisement). The Wall Street Journal. 2010; September 29:A22.
- Zikmund WG, Babin BJ. Exploring marketing research. 10th ed. Mason: South-Western Cengage Learning; 2010.

فصل ششم

همبستگی و رگرسیون خطی ساده

انواع مختلفی از ضرایب همبستگی وجود دارند، اما یکی از آن‌ها که در این کتاب استفاده خواهیم کرد، ضریب همبستگی است که در زمان پیرسون به وجود آمد که آن را r می‌نامیم.

۶-۱- همبستگی چیست؟

اساساً، همبستگی، عددی است بین -1 و $+1$ که رابطه بین دو متغیر را که X و Y خواهیم نامید، خلاصه می‌کند.

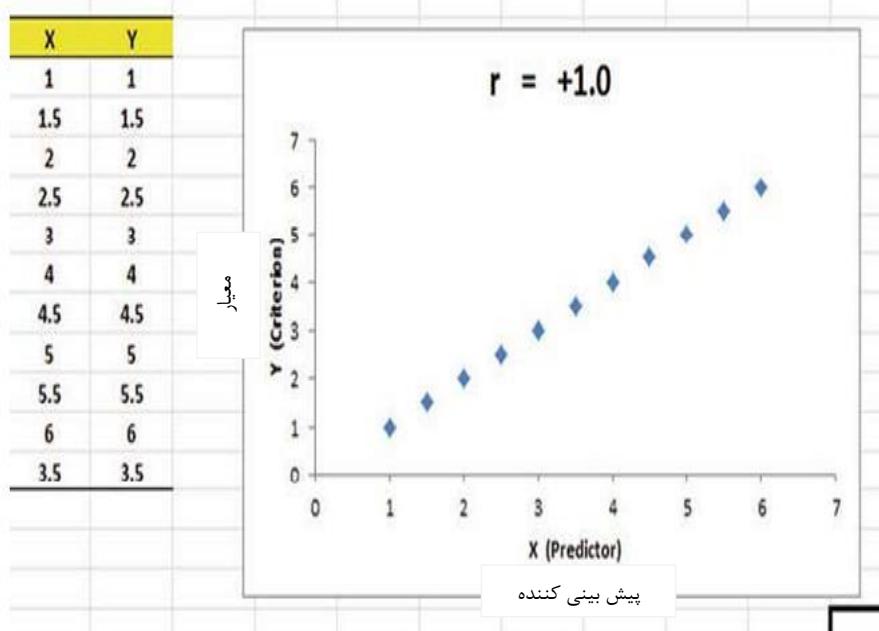
همبستگی می‌تواند مثبت یا منفی باشد. همبستگی مثبت به این معناست که با افزایش X ، Y افزایش می‌یابد. همبستگی منفی به این معناست که با افزایش X ، Y کاهش می‌یابد. در کتاب‌های آمار، علامت همبستگی، جهت رابطه (یعنی مثبت یا منفی) نامیده می‌شود.

همچنین، همبستگی مقدار رابطه بین X و Y را به ما می‌گوید. وقتی همبستگی به $+1$ نزدیک می‌شود، می‌گوییم که رابطه قوی و مثبت است.

وقتی همبستگی به -1 نزدیک می‌شود، می‌گوییم که رابطه قوی و منفی است. همبستگی صفر به این معناست که رابطه‌ای بین X و Y وجود ندارد. این بدان معنی است که نه X و نه Y نمی‌توانند به عنوان پیش‌بینی کننده (تعیین‌کننده) دیگری استفاده شوند.

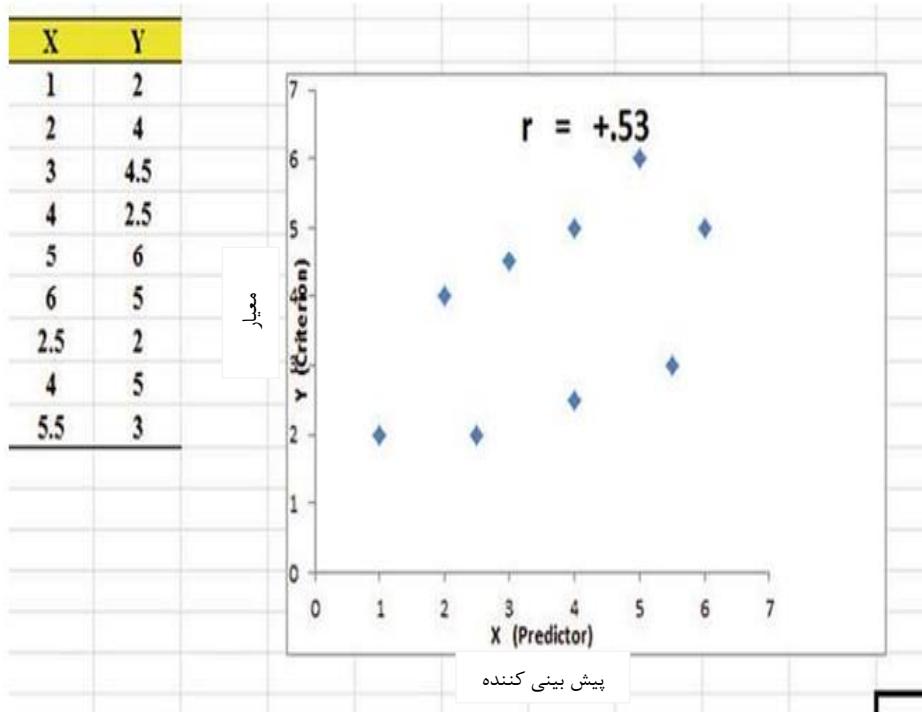
یک روش خوب برای درک اینکه همبستگی به چه معنی بوده، این است که تصویری از پراکنش نقاط داده را در یک نمودار ببینید. فرض کنید می‌خواهید بدانید که آیا متغیر X می‌تواند به عنوان پیش‌بینی کننده متغیر Y استفاده شود یا خیر. متغیر پیش‌بینی کننده X را در محور X (محور افقی یک نمودار) و متغیر معیار Y را روی محور Y (محور عمودی یک نمودار) قرار خواهیم داد. همچنین، فرض کنید که شما داده‌های مفروض مورداستفاده در نمودار پراکنش را جمع‌آوری کرده‌اید (به شکل ۶-۱ تا شکل ۶-۶ مراجعه شود).

شکل ۱-۶ نشان‌دهنده یک نمودار پراکنش برای یک همبستگی کامل مثبت $r = +1$ می‌باشد. این بدان معنی است که شما می‌توانید به طور کامل هر مقدار Y را از هر مقدار X پیش‌بینی کنید زیرا نقاط داده‌ها "به سمت بالا و راست" را در امتداد یک خط راست کاملاً برازش کننده حرکت می‌کنند. (به شکل ۱-۶ مراجعه شود).



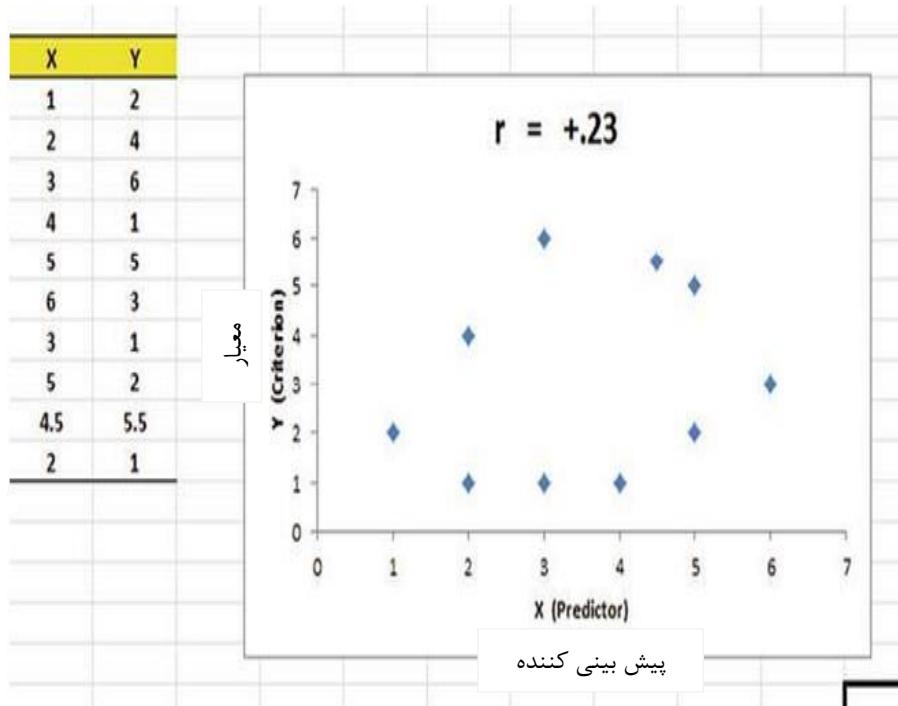
شکل ۱-۶: مثالی از نمودار پراکنش برای یک همبستگی مثبت و کامل ($r = +1$)

شکل ۲-۶ نشان‌دهنده نمودار پراکنش برای همبستگی مثبت و متوسط است $r = +0.53$. این بدان معنی است که هر مقدار از X می‌تواند هر مقدار Y را به طور نسبتاً خوبی پیش‌بینی کند، زیرا می‌توانید تصویری بیضی شکل اما نه در امتداد یک خط راست در اطراف نقاط داده که به سمت بالا و راست حرکت می‌کنند، را رسم کنید، (به شکل ۲-۶ مراجعه شود).



شکل ۲-۶: مثالی از نمودار پراکنش برای همبستگی مثبت و متوسط ($r = +0.53$)

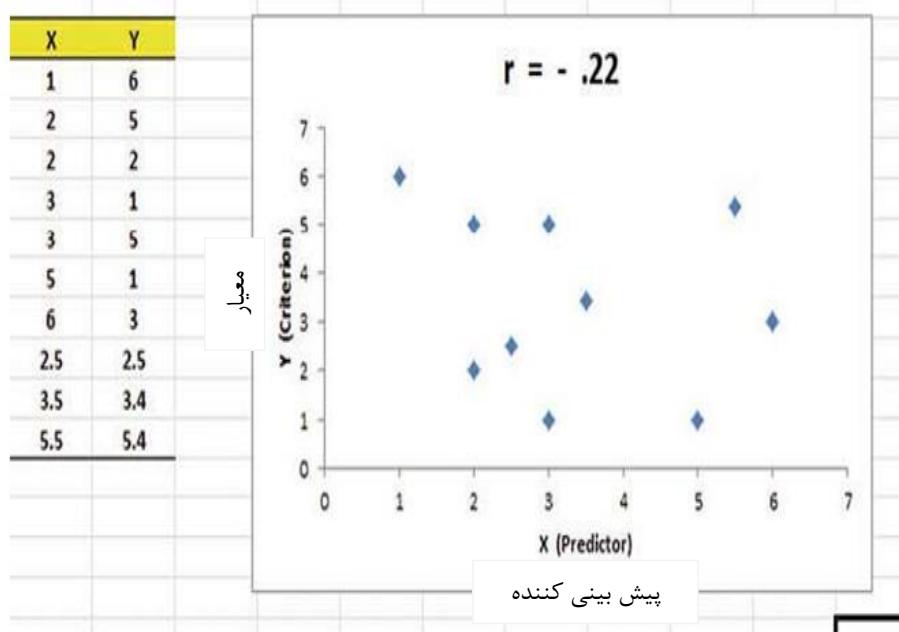
شکل ۳-۶ نشان‌دهنده نمودار پراکنش برای همبستگی مثبت و ضعیف می‌باشد $r = +0.23$. این بدان معنی است که هر مقدار X یک پیش‌بینی کننده ضعیف برای هر مقدار Y است، زیرا «تصویری» که می‌توانید در اطراف نقاط داده شده بکشید، نزدیک به یک دایره است (به شکل ۳-۶ مراجعه شود).



شکل ۳-۶: مثالی از نمودار پراکنش برای همبستگی مثبت و ضعیف ($r = +0.23$)

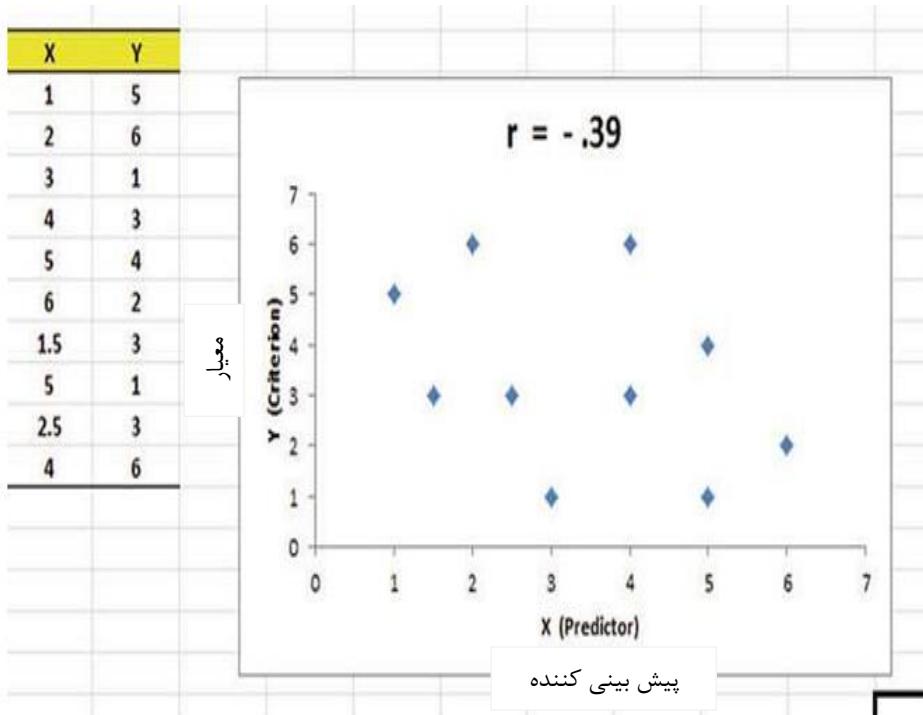
ما تصویری از همبستگی صفر را ارائه ندادیم زیرا تصور اینکه نمودار پراکنش آن به چه شکلی شبیه بوده آسان است. همبستگی صفر ($r = 0$) به این معنی است که بین X و Y ارتباطی وجود ندارد و "تصویری" که در اطراف نقاط داده شده تشکیل می‌شود یک دایره کامل است که نشان می‌دهد نمی‌توانید از X برای پیش‌بینی Y استفاده کنید، زیرا این دو متغیر با یکدیگر همیستگی ندارند.

شکل ۶-۶ نشان‌دهنده نمودار پراکنش برای همبستگی منفی و ضعیف $r = -0.22$ می‌باشد که بدان معنی است که هر X یک پیش‌بینی کننده ضعیف برای هر مقدار Y در یک رابطه معکوس است، به این معنی که با افزایش X , Y کاهش می‌یابد (به شکل ۶-۶ مراجعه شود). در این مورد همبستگی منفی وجود دارد؛ زیرا می‌توانید شکل "بیضی" در اطراف نقاط داده رسم کنید که دارای شبیه به سمت پایین و راست است.



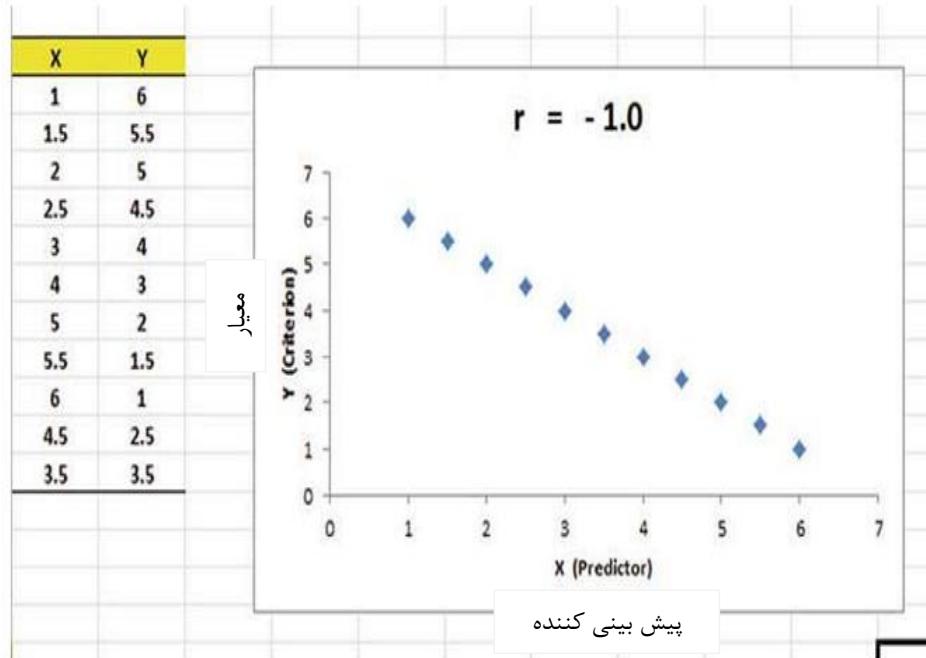
شکل ۶-۶: مثالی از نمودار پراکنش برای همبستگی منفی و ضعیف ($r = -0.22$)

شکل ۶-۵ نشان‌دهنده نمودار پراکنش برای همبستگی متوسط و منفی $r = -0.39$ می‌باشد که بدان معنی است که X یک پیش‌بینی کننده نسبتاً خوب برای Y است، اگرچه رابطه‌ای معکوس بین X و Y وجود دارد. (یعنی با افزایش X , Y کاهش می‌یابد؛ به شکل ۵-۶ مراجعه شود). در این مورد همبستگی منفی وجود دارد زیرا بیضی که شما می‌توانید در اطراف نقاط داده‌ها رسم می‌کنید به سمت پایین و راست می‌باشند.



شکل ۶-۵: مثالی از نمودار پراکنش برای همبستگی منفی و متوسط ($r = -0.39$)

شکل ۶-۶ نشان‌دهنده همبستگی منفی و کامل است $r = -1$ که نشان می‌دهد X یک پیش‌بینی کننده کامل برای Y است، اگرچه این یک رابطه معکوس می‌باشد، به طوری که با افزایش X ، Y کاهش می‌یابد. نقاط داده‌ها کاملاً در امتداد یک خط راست به سمت پایین و راست قرار دارند (به شکل ۶-۶ مراجعه شود).



شکل ۶-۶: مثالی از نمودار پراکنش برای همبستگی منفی کامل ($r = -1$)

بیایید رابطه محاسبه همبستگی r را توضیح دهیم به‌نحوی که بتوانید درک کنید عدد خلاصه‌کننده همبستگی از کجا آمده است.

برای کمک به اینکه درک کنید عدد همبستگی که در بازه $-1 \leq r \leq +1$ است از کجا آمده است، می‌خواهیم مراحلی را که متناسبه حالت استفاده از یک ماشین حساب جیبی است را ارائه کنیم. این تنها دفعه‌ای است که از شما می‌خواهیم تا از ماشین حساب جیبی خود برای پیدا کردن همبستگی استفاده کنید، اما دانستن اینکه همبستگی چگونه مرحله به مرحله محاسبه می‌شود، به شما این امکان را می‌دهد تا بفهمید چطور از این رابطه در عمل استفاده کنید.

برای انجام این کار موقعیتی ایجاد می‌کنیم که شما در آن نیاز به پیدا کردن همبستگی بین دو متغیر داشته باشید.

فرض کنید که توسط مدیر یک فروشگاه زنجیره‌ای استخدام شده‌اید تا رابطه بین مقدار پولی که به صورت هفتگی توسط فروشگاه برای تبلیغات تلویزیونی هزینه می‌شود و فروش هفتگی فروشگاه زنجیره‌ای در شهر سنت لوئیس را پیدا کنید. شما داده‌هایی را از ۸ هفته گذشته جمع‌آوری کرده‌اید که در شکل ۷-۶ ارائه شده است.

هفته	هزینه تبلیغات تلویزیونی (هزار دلار)	فروش هفتگی (هزار دلار)
Week	TV ad cost (\$000)	Weekly Sales (\$000)
1	4.8	94
2	1.9	87
3	3.8	93
4	2.3	89
5	2.9	92
6	3.3	92
7	2.4	93
8	2.8	92

n	تعداد نمونه	8	8
MEAN	میانگین	3.03	91.50
STDEV	انحراف معیار	0.93	2.33

شکل ۷-۶: صفحه کاری داده‌ها برای فروشگاه زنجیره‌ای (مثال عملی)

برای توضیح، اجازه دهید هزینه هفتگی تبلیغات تلویزیونی را به عنوان متغیر پیش‌بینی کننده X و فروش هفتگی را به عنوان متغیر معیار Y بنامیم. توجه داشته باشید که داده‌های هزینه تبلیغات تلویزیونی برای هر هفته به هزار دلار (۱۰۰۰ دلار) است. به عنوان مثال، تبلیغات تلویزیونی برای هفته ششم ۳۳۰۰ دلار هزینه داشته و زمانی که اعشار را سه رقم به سمت چپ تغییر می‌دهیم تا

مقدار متغیر به هزار دلار بیان شود، این مقدار $\frac{3}{3}$ می‌شود. به طور مشابه، فروش هفتگی برای هفته ششم درواقع ۹۲۰۰۰ دلار می‌باشد، زیرا این داده هم در قالب هزار دلار (۱۰۰۰ دلار) است. همچنین، توجه داشته باشید که ما از اکسل برای پیدا کردن اندازه نمونه برای هر دو متغیر X و Y و میانگین و انحراف معیار هر دو متغیر استفاده کردیم (می‌توانید مهارت اکسل خود را با دیدن اینکه آیا نتایج مشابه را در هنگام ایجاد یک صفحه گسترده اکسل برای این داده‌ها به دست آورده‌اید یا خیر، بیازمایید).

حال اجازه دهید که از جدول بالا برای محاسبه همبستگی r بین هزینه هفتگی تبلیغات تلویزیونی و فروش هفتگی این فروشگاه زنجیره‌ای با استفاده از ماشین حساب جیبی، استفاده کنیم.

۱-۱-۶- درک رابطه محاسبه همبستگی

هدف: درک رابطه محاسبه همبستگی r

رابطه محاسبه همبستگی r به شرح زیر است:

$$r = \frac{\frac{1}{n-1} \sum (X - \bar{X})(Y - \bar{Y})}{S_x S_y} \quad (1-6)$$

این رابطه در نگاه اول دلهره آور است، اما اجازه دهید برای درک چگونگی محاسبه همبستگی r آن را به مراحل تشکیل دهنده‌اش تقسیم کنیم.

۶-۱-۲- درک مراحل نه گانه محاسبه همبستگی r

هدف: درک مراحل نه گانه محاسبه همبستگی r

این ۹ مرحله به شرح زیر است:

مرحله	محاسبه	نتیجه
۱	تعداد نمونه را با توجه به تعداد هفته‌ها پیدا کنید	۸
۲	یک را تقسیم بر تعداد نمونه منهای ۱ کنید (یعنی $1/7$)	$0/14286$
۳	برای هر هفته، هزینه‌ای که بابت تبلیغات تلویزیونی پرداخت شده را در نظر گرفته و از میانگین هزینه ۸ هفته کم کنید و آن را $\bar{X} - X$ بنامید (برای مثال، برای هفته ۶ این مقدار برابر: $3/03 - 3/03 = 0/00$). توجه: با ماشین حساب شما این تفاوت برابر $0/27$ است، اما زمانی که اکسل از ۱۶ رقم اعشار برای هر محاسبه استفاده می‌کند نتیجه $0/28$ بجای $0/27$ خواهد بود.	$0/27$
۴	برای هر هفته، فروش هفتگی را از میانگین فروش ۸ هفته کم کنید و آن را $\bar{Y} - Y$ بنامید (برای مثال، برای هفته ۶ این مقدار برابر: $91/5 - 92/5 = -0/4$).	$0/50$
۵	سپس، برای هر هفته، $(X - \bar{X})$ را در $(Y - \bar{Y})$ ضرب کنید (به عنوان مثال، برای هفته ۶ این مقدار برابر $0/27 \times 0/50 = 0/0135$)	$0/135$
۶	نتایج $(X - \bar{X})$ ضربدر $(Y - \bar{Y})$ را برای ۸ هفته باهم جمع کنید.	$11/50$

مراحل ۱ تا ۶ جدول اکسل داده شده، شکل ۶-۸ را ایجاد می‌کند.

هفته	هزینه تبلیغات تلویزیونی (هزار دلار)		فروش هفتگی (هزار دلار)		
	X	Y	$X - \bar{X}$	$Y - \bar{Y}$	$(X - \bar{X})(Y - \bar{Y})$
Week	TV ad cost (\$000)	Weekly Sales (\$000)			
1	4.8	94	1.78	2.50	4.44
2	1.9	87	-1.13	-4.50	5.06
3	3.8	93	0.78	1.50	1.16
4	2.3	89	-0.73	-2.50	1.81
5	2.9	92	-0.13	0.50	-0.06
6	3.3	92	0.28	0.50	0.14
7	2.4	93	-0.63	1.50	-0.94
8	2.8	92	-0.23	0.50	-0.11
تعداد نمونه	8	8		Total	11.50
میانگین	3.03	91.50		مجموع	
انحراف معیار	0.93	2.33			

شکل ۸-۶: صفحه گسترده برای محاسبه همبستگی، r

توجه داشته باشید هنگامی که اکسل یک عدد منفی را در یک عدد منفی دیگر ضرب می کند نتیجه یک عدد مثبت است (به عنوان مثال، برای هفته ۲: $+5/0.6 = +8/13 \times -4/50 = -4/13$) و هنگامی که اکسل یک عدد منفی را در یک عدد مثبت ضرب کند، نتیجه یک عدد منفی است (مثلاً، برای هفته ۵: $-0/13 \times +0/50 = -0/0.6$)

توجه: اکسل تمام محاسبات را با ۱۶ رقم اعشار محاسبه می کند. بنابراین، هنگامی که کار خود را با یک ماشین حساب انجام می دهید، اغلب پاسخ هایی با کمی اختلاف از پاسخ اکسل دریافت می کنید. به عنوان مثال، هنگامی که شما در بالا محاسبه می کنید:

$$(X - \bar{X}) \times (Y - \bar{Y}) \text{ برای هفته دوم، ماشین حساب شما پاسخ زیر را خواهد داد:}$$

$$(-1/13) \times (-4/50) = +5/0.85$$

اما همانطور که در جدول می بینید، پاسخ اکسل یعنی ۵/۰۶ دقیق‌تر است، زیرا اکسل از ۱۶ رقم اعشار برای هر عدد استفاده می کند.

همچنین، باید توجه داشته باشید هنگامی که مرحله ۶ را انجام می‌دهید، ابتدا باید تمام اعداد مثبت را به $+12/61$ + اضافه و سپس، تمام اعداد منفی را به $-1/11$ - اضافه کنید تا وقتی این دو عدد را از

مرحله	محاسبه	نتیجه
۷	پاسخ مرحله ۲ را در پاسخ مرحله ۶ ضرب کنید ($0/16286 \times 11/5$)	۱/۶۴۲۹
۸	انحراف معیار X را در انحراف معیار Y ضرب کنید ($0/93 \times 2/33$)	۲/۱۶۶۹
۹	درنهایت پاسخ مرحله ۷ را تقسیم بر پاسخ مرحله ۸ کنید. ($1/6829 / 2/1669$) تقسیم بر $2/1669$	$+0/76$

هم کم می‌کنید، عدد $11/50$ + را که پاسخ مرحله ۶ است را به دست آورید.

عدد $0/76$ همبستگی بین هزینه هفتگی تبلیغات تلویزیونی (X) و فروش هفتگی در این فروشگاه زنجیره‌ای (Y) طی این دوره ۸ هفته‌ای است. عدد $+0/76$ به معنای این است که بین این دو متغیر همبستگی مثبت و قوی وجود دارد. از این رو، زمانی که فروشگاه زنجیره‌ای هزینه‌های خود را در تبلیغات تلویزیونی افزایش می‌دهد، فروش آن هفته افزایش می‌یابد. برای بحث دقیق‌تر درباره همبستگی، به زیکموند و بابیان (۲۰۱۰) مراجعه کنید.

همچنین، می‌توانید از نتایج محاسبات بالا در رابطه همبستگی r به روش زیر استفاده کنید:

$$r = \left[\frac{1}{(n-1)} \times \sum_{y} (X - \bar{X})(Y - \bar{Y}) \right] / \left(\sqrt{\sum_{x} (X - \bar{X})^2} \times \sqrt{\sum_{y} (Y - \bar{Y})^2} \right)$$

$$r = [(1/7) \times 11.5] / [(0.93) \times (2.33)]$$

$$r = 0.76$$

حال بیایید بحث کنیم که چگونه می‌توان از اکسل برای پیدا کردن همبستگی بین دو متغیر به روش بسیار ساده‌تر و سریع‌تر در مقایسه با کاربرد ماشین حساب، استفاده کرد.

۶-۲- استفاده از اکسل برای محاسبه همبستگی بین دو متغیر

هدف: استفاده از اکسل برای محاسبه همبستگی بین دو متغیر

فرض کنید که توسط صاحب یک فروشگاه زنجیره‌ای در شهر سنت لوئیس استخدام شده‌اید تا توصیه کنید که چه تعداد قفسه در این فروشگاه برای ذرت کللوگ^۱ استفاده شود. در هر «قفسه» جعبه‌های غلات کنار یکدیگر قرار داده شده است. به این ترتیب، عبارت «۳ مخصوص روی قفسه» به این معنی است که ۳ جعبه ذرت کللوگ در کنار یکدیگر در قفسه فروشگاه در بخش غلات قرار می‌گیرند.

شما بطور تصادفی مکان‌هایی از فروشگاه را برای بررسی خود اختصاص داده‌اید و به صورت تصادفی تعدادی از قفسه‌های استفاده شده در هر قسمت فروشگاه را انتخاب می‌کنید، جایی که تعداد قفسه‌ها در محدوده ۱ تا ۳ قفسه بود. شما فروش هفته‌ای (به هزار دلار) این غلات را طی ۱۰ هفته دنبال می‌کنید و رقم فروش حاصل در شکل ۹-۶ نشان داده شده است.

هفته	تعداد قفسه‌ها	فروش (هزار دلار)
Week	No. of facings	Sales (\$000)
1	1	1.1
2	2	2.2
3	3	2.1
4	1	1.2
5	2	2.3
6	3	5.2
7	3	4.6
8	2	2.3
9	2	1.9
10	3	4.5

شکل ۹-۶: صفحه کاری داده‌ها برای تعداد قفسه‌ها و مقدار فروش (مثال عملی)

¹ Kellogg

شما می‌خواهید تعیین کنید که آیا رابطه‌ای بین تعداد قفسه‌های ذرت کللوگ و فروش هفتگی این غلات وجود دارد یا خیر و تصمیم دارید تا برای تعیین این رابطه از همبستگی استفاده کنید. اجازه دهید تعداد قفسه‌ها را X و رقم فروش را Y بنامیم.

صفحه گسترده اکسل را با اطلاعات زیر ایجاد کنید.

A2: هفته

B2: تعداد قفسه‌ها

C2: فروش (هزار دلار)

1: A3

اکنون، عرض ستون‌های B و C را تغییر دهید تا مناسب با اطلاعات درون سلول‌ها شود.
حالا ارقام باقیمانده که در جدول بالا داده شده است را کامل کنید، بطوریکه A12، B12، ۳، C12 ۴/۵ است (دوباره اعداد را بررسی کنید تا مطمئن شوید که درست هستند). سپس، اطلاعات همه سلول‌ها را در مرکز سلول تنظیم کنید.

A14: تعداد نمونه

A15: میانگین

A16: انحراف معیار

بعد نام محدوده اطلاعات B3 تا B12 را به عنوان "تعداد قفسه‌ها" تعریف کنید.
ما قبلاً در این کتاب در مورد نحوه "نام‌گذاری بازه‌ای از سلول‌ها" بحث کردیم (به بخش ۱-۴-۱ مراجعه شود)، اما اینجا نحوه انجام این کار یادآوری می‌شود:
نام‌گذاری بازه‌ای از سلول‌ها:

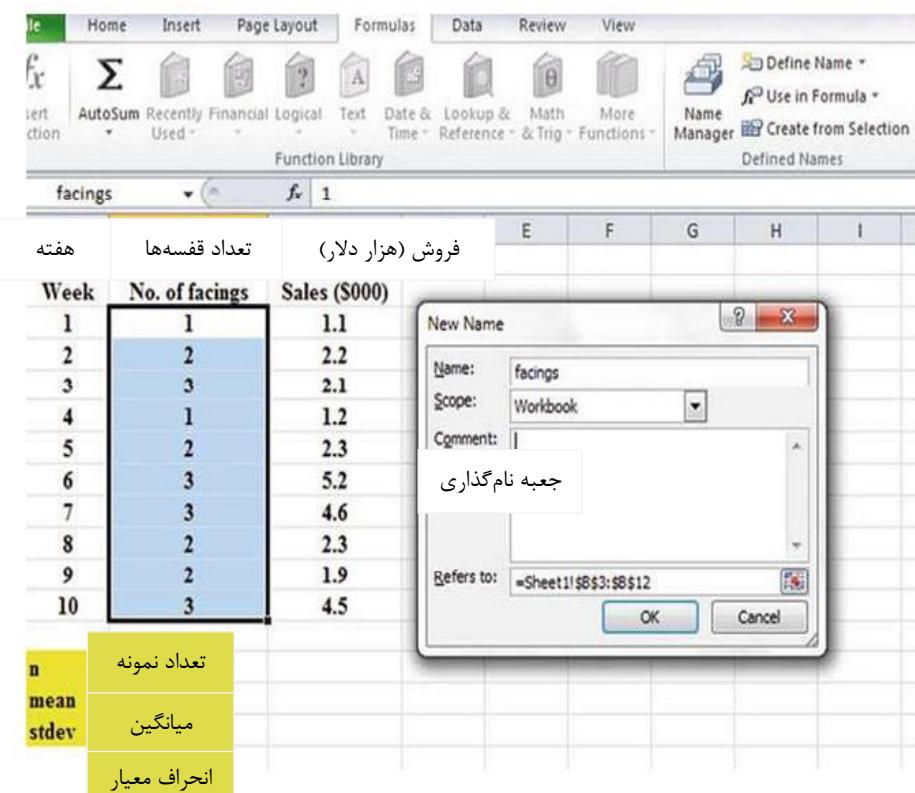
روی بالاترین داده در بازه داده‌ها کلیک کرده، سپس، ماوس را تا پایین‌ترین داده در محدوده بکشید.

به عنوان مثال، برای دادن نام: "تعداد قفسه‌ها" به سلول‌های B3 تا B12، روی B3 کلیک کنید و سپس نشانگر ماوس را به سمت پایین تا B12 بکشید تا سلول‌های B3 تا B12 بر روی صفحه رایانه شما انتخاب شوند. سپس، کلیک کنید روی:

(Formulas)

تعریف نام (Dedine name) (وسط بالای صفحه شما)

تعداد قفسه‌ها (در جعبه نام، به شکل ۶-۱۰ مراجعه شود)



شکل ۶-۱۰: جعبه محاوره‌ای برای نام‌گذاری بازه‌ای از سلول‌ها با عنوان "تعداد قفسه‌ها"

را انتخاب کنید OK

اکنون، این مراحل را تکرار کنید تا بازه C3 تا C12 را "فروش" نام‌گذاری کنید.
درنهایت، بر روی هر سلول خالی در صفحه گسترده کلیک کنید تا سلول C3 تا C12 بر روی صفحه رایانه از حالت انتخاب خارج شود.

اکنون، داده‌ها برای تعداد نمونه، میانگین و انحراف معیار را در ستون‌های B و C کامل کنید، بطوریکه، B16، ۰/۷۹ است و C16، ۱/۴۷ است. از دو رقم اعشار برای میانگین و انحراف معیار استفاده کنید، (به شکل ۱۱-۶ مراجعه شود).

هفته	تعداد قفسه‌ها	فروش (هزار دلار)
Week	No. of facings	Sales (\$000)
1	1	1.1
2	2	2.2
3	3	2.1
4	1	1.2
5	2	2.3
6	3	5.2
7	3	4.6
8	2	2.3
9	2	1.9
10	3	4.5

تعداد نمونه	n	10	10
میانگین	mean	2.20	2.74
انحراف معیار	stdev	0.79	1.47

شکل ۱۱-۶: مثالی از کاربرد اکسل برای پیدا کردن تعداد نمونه، میانگین و انحراف معیار

هدف: پیدا کردن همبستگی بین تعداد قفسه‌ها و مقدار فروش هفتگی

B18: همبستگی

. (به شکل ۱۲-۶ مراجعه شود).

= correl(facings,sales) :C18

	A	B	C	D	E
1	هفته	تعداد قفسه‌ها	فروش (هزار دلار)		
2	Week	No. of facings	Sales (\$000)		
3	1	1	1.1		
4	2	2	2.2		
5	3	3	2.1		
6	4	1	1.2		
7	5	2	2.3		
8	6	3	5.2		
9	7	3	4.6		
10	8	2	2.3		
11	9	2	1.9		
12	10	3	4.5		
13					
14	n	تعداد نمونه	10	10	
15	mean	میانگین	2.20	2.74	
16	stdev	انحراف معیار	0.79	1.47	
17					
18		correlation	=correl(facings,sales)		
19		همبستگی			
۲۰					

شکل ۶-۱۲: مثالی از کاربرد دستور همبستگی اکسل برای محاسبه ضریب همبستگی

برای محاسبه همبستگی کلید اینتر را فشار دهید.

قالب این سلول را به دو رقم اعشار تنظیم کنید.

توجه داشته باشید که علامت مساوی به اکسل می‌گوید که شما از یک دستور استفاده می‌کنید. همبستگی بین تعداد قفسه‌ها (X) و فروش هفتگی (Y) برابر با $+0.83$ است. این همبستگی مثبت و بسیار قوی می‌باشد. این بدان معنی است که شما شواهدی دارید که رابطه‌ای قوی بین این دو

متغیر وجود دارد. درواقع، قفسه‌های بیشتر (زمانی که ۱ و ۲ و ۳ قفسه استفاده می‌شد)، فروش هفتگی بالاتری برای این غلات به وجود می‌آورد. این فایل را به عنوان: FACINGS5 ذخیره کنید. صفحه گسترده نهایی در شکل ۱۳-۶ نشان داده می‌شود.

		C18	f(x)	=CORREL(facings,sales)
A	B	C	D	E
Week	No. of facings	Sales (\$000)		
1	1	1.1		
2	2	2.2		
3	3	2.1		
4	1	1.2		
5	2	2.3		
6	3	5.2		
7	3	4.6		
8	2	2.3		
9	2	1.9		
10	3	4.5		
n		تعداد نمونه	10	10
mean		میانگین	2.20	2.74
stdev		انحراف معیار	0.79	1.47
correlation		همبستگی	0.83	

شکل ۱۳-۶: نتایج نهایی برای استفاده از تابع =correl برای محاسبه ضریب همبستگی

۶-۳- ایجاد نمودار و رسم خط رگرسیون بر روی نمودار

این بخش با مفهوم "رگرسیون خطی" سروکار دارد. از لحاظ فنی، استفاده از رگرسیون خطی ساده (کلمه "ساده" به این معنی است که فقط یک پیش‌بینی کننده X برای پیش‌بینی معیار Y استفاده می‌شود) هنگام استفاده از الگوی آماری لازم است که داده‌ها از چهار فرضیه زیر پیروی کنند:

۱- رابطه پایه‌ای بین دو متغیر مورد مطالعه (X و Y)، خطی است، به این معنی که یک خط

مستقیم و نه یک منحنی می‌تواند برازش مناسبی برای داده‌های در نمودار باشد.

۲- خطاهای اندازه‌گیری مستقل از یکدیگر هستند (به عنوان مثال، خطاهای یک دوره زمانی خاص، گاهی با خطاهای در دوره زمانی قبل، همبستگی ندارند).

۳- خطاهای دارای توزیع نرمال در مقادیر Y به ازای هر یک از مقادیر X می‌باشند.

۴- واریانس خطاب برای تمام مقادیر X یکسان است (یعنی، نوسان مقادیر Y برای مقادیر بالا و پایین X یکسان است).

توضیحات کامل این مفروضات فراتر از محدوده این کتاب است، خواننده علاقه‌مند می‌تواند جزیبات کامل این مفروضات را در لوین و همکاران پیدا کند (۲۰۱۱، ص ۵۲۹-۵۳۰). اکنون اجازه دهید یک نمودار برای خلاصه کردن این داده‌ها ایجاد کنیم.

نکته مهم: هنگام رسم نمودار، ضروری است که متغیر پیش‌بینی کننده (X) در سمت چپ و متغیر معیار (Y) را در سمت راست صفحه گسترده اکسل قرار دهید تا بدانید کدام متغیر، متغیر پیش‌بینی کننده و کدام متغیر، متغیر معیار است. اگر این کار را انجام دهید، هر زمان مشکلی با همبستگی و رگرسیون خطی در کاربرد اکسل ایجاد شد، رحمت خود را کم می‌کنید.

نکته مهم: باید درک کنید که در هر نموداری که یک (متغیر) پیش‌بینی کننده و یک (متغیر) معیار است، در واقع دو خط وجود دارند که می‌توانند بین نقاط داده‌ها کشیده شوند.

(۱) یک خط از X به عنوان متغیر پیش‌بینی کننده و از Y به عنوان متغیر معیار استفاده می‌کند.

(۲) خط دوم از Y به عنوان متغیر پیش‌بینی کننده و از X به عنوان متغیر معیار استفاده می‌کند.

این بدان معنی است که باید در نوشتن داده‌های ورودی خود در سلول‌هایی که شامل X به عنوان متغیر پیش‌بینی کننده و Y به عنوان متغیر معیار هستند، دقت داشته باشید. اگر اعداد این سلول‌ها را جابجا وارد کنید، خطوط اشتباہی برای داده‌ها ایجاد خواهد کرد.

به همین دلیل است که در این کتاب به شدت توصیه می‌کنیم که همیشه داده X (یعنی متغیر پیش‌بینی کننده) را در سمت چپ جدول و داده‌های Y (یعنی متغیر معیار) را در سمت راست جدول در صفحه گسترده قرار دهید به طوری که این متغیرها را اشتباہ نگیرید.

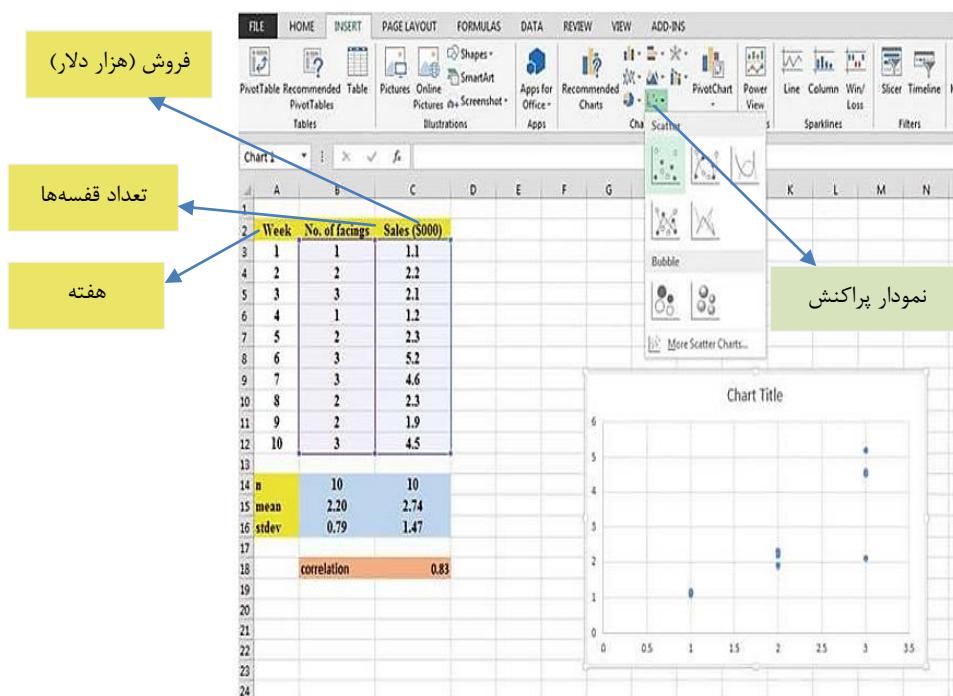
همچنین باید توجه داشته باشید که بدون در نظر گرفتن اینکه شما کدام متغیر را پیش‌بینی کننده و کدام یک را متغیر معیار بنامید همبستگی، دقیقاً برابر خواهد بود. ضریب همبستگی فقط رابطه بین دو متغیر را خلاصه می‌کند و مهم نیست کدام یک پیش‌بینی کننده و کدام یک معیار است.

باید فرض کنیم که شما مایل به استفاده از تعداد قفسه‌های ذرت به عنوان متغیر پیش‌بینی کننده بوده و از آن برای پیش‌بینی مقدار فروش هفتگی این غله استفاده کنید. از آنجاکه همبستگی بین این دو متغیر برابر با $+0.83$ است، این نشان می‌دهد که یک رابطه مثبت و قوی وجود دارد و تعداد قفسه‌ها پیش‌بینی کننده خوبی برای فروش هفتگی این غله است.

۶-۳-۱- استفاده از اکسل برای ایجاد نمودار و رسم خط رگرسیون از طریق نقاط داده‌ها

هدف: ایجاد نمودار و رسم خط رگرسیون خلاصه کننده رابطه بین تعداد قفسه‌ها و فروش هفتگی (هزار دلار)

- ۱- فایل FACINGS5 که قبلاً در این فصل ذخیره کرده‌اید را باز کنید.
- ۲- هر دو ستون اعداد (B3:C12) را با کلیک کردن و کشیدن ماوس انتخاب کنید، اما نام بالای ستون B و C را انتخاب نکنید.
سری داده‌های B3:C12 را انتخاب کنید.
- منوی Insert (در بالای صفحه سمت چپ) را انتخاب کنید
آیکون نمودار پراکنش (در بالای کلمه "نمودارها" در وسط صفحه) را انتخاب کنید
روی فلاش پایین سمت راست نماد نمودار کلیک کنید.
نماد نمودار پراکنش در بالا و سمت چپ را انتخاب کنید (به شکل ۱۴-۶ مراجعه شود)

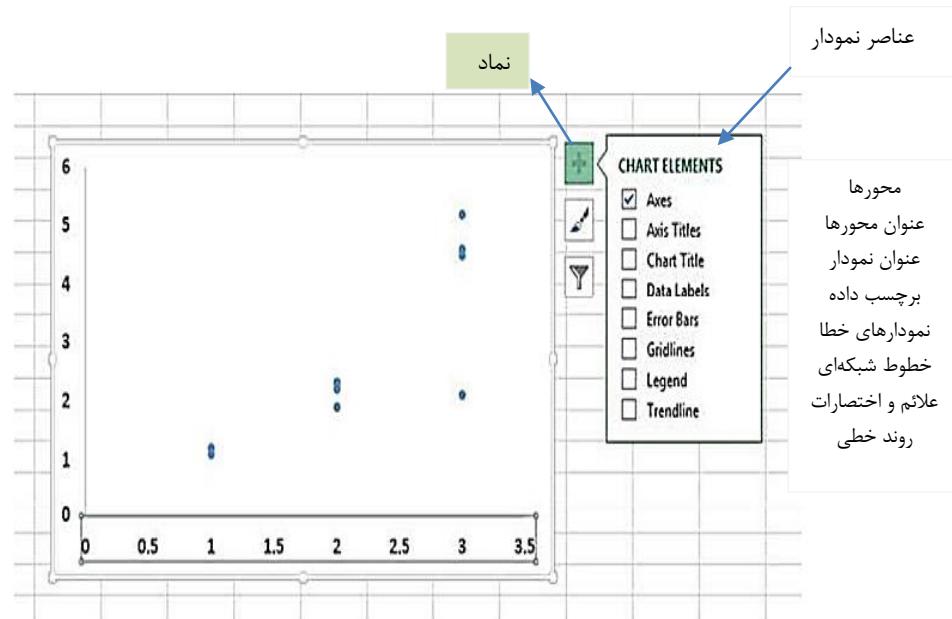


شکل ۱۴-۶: مثالی از انتخاب نمودار پراکنش

برای انتخاب آن روی نمودار بالا سمت چپ کلیک کنید.

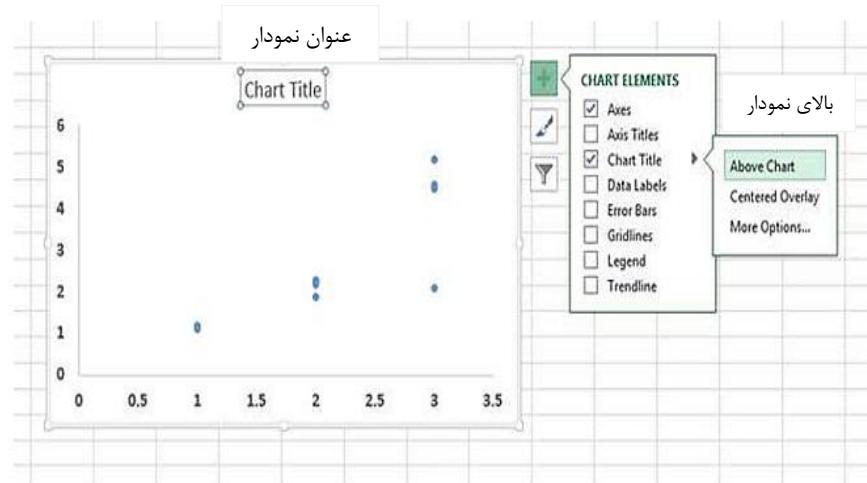
روی علامت "نماد +" در سمت راست نمودار (اجزاء نمودار) کلیک کنید.

روی علامت چک مارک در کنار "عنوان نمودار" و همچنین "خطوط شبکه‌ای" کلیک کنید تا این علامت‌های تیک را حذف کنید (به شکل ۱۵-۶ مراجعه شود).



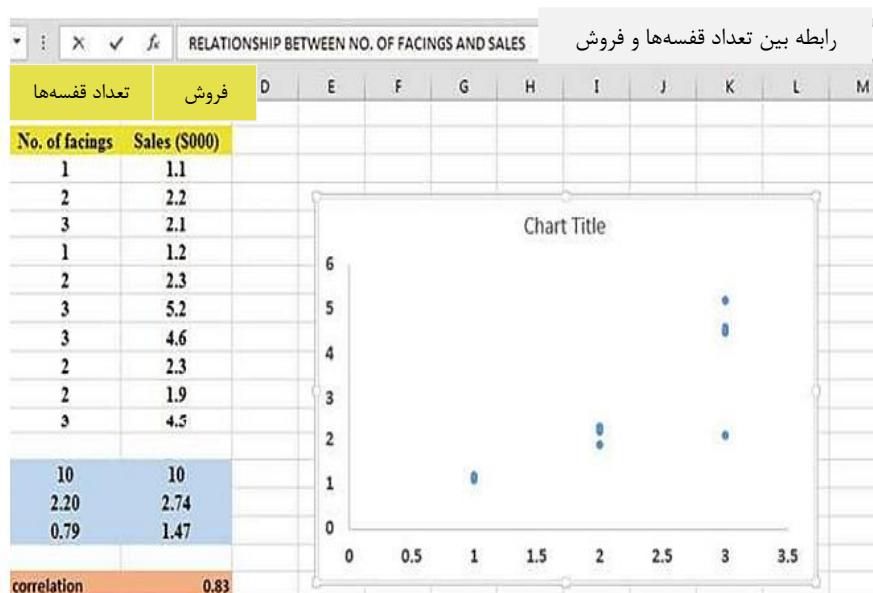
شکل ۱۵-۶: مثالی از اجزای نمودار انتخاب شده

روی جعبه کنار "عنوان نمودار" کلیک کنید و سپس روی فلش سمت راست کلیک کنید. سپس، بر روی "بالای نمودار" کلیک کنید.
توجه داشته باشید که کلمات "عنوان نمودار" در حال حاضر در جعبه بالای نمودار هستند (به شکل ۱۶-۶ مراجعه شود)



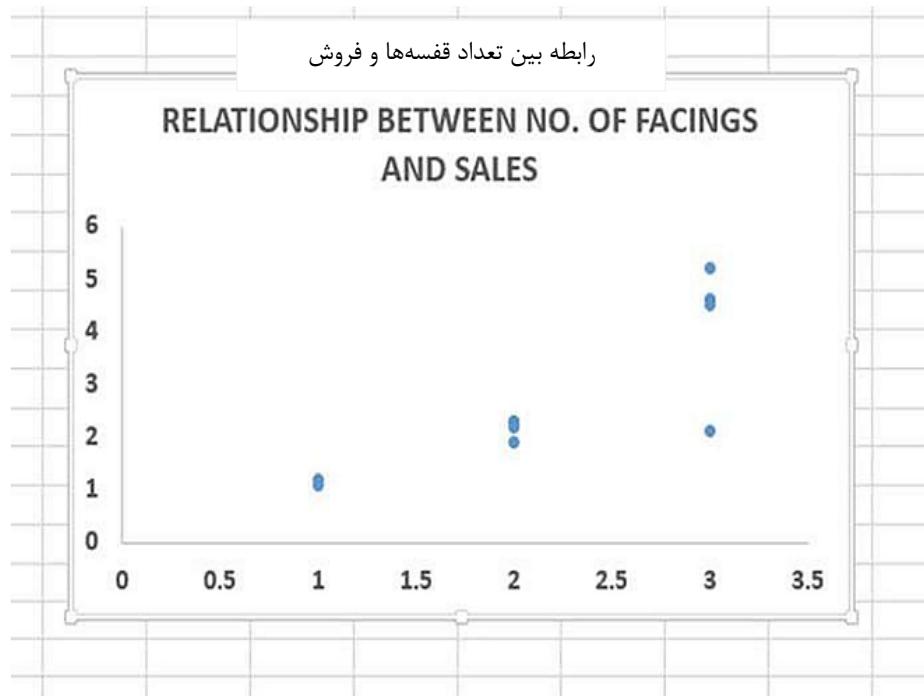
شکل ۱۶-۶: مثالی از عنوان نمودار انتخاب شده

عنوان نمودار زیر را در سمت راست در بالای صفحه خود وارد کنید: رابطه بین تعداد قفسه‌ها و فروش (به شکل ۱۷-۶ مراجعه شود).



شکل ۱۷-۶: مثالی از ایجاد عنوان برای نمودار

کلید اینتر را فشار دهید تا این عنوان روی نمودار قرار گیرد.
داخل نمودار در گوشه بالا سمت چپ نمودار کلیک کنید تا عنوان نمودار از حالت انتخاب خارج شود (به شکل ۱۸-۶ مراجعه شود)



شکل ۱۸-۶: مثالی از عنوان قرار داده شده در داخل نمودار

روی جعبه در سمت راست نمودار کلیک کنید.

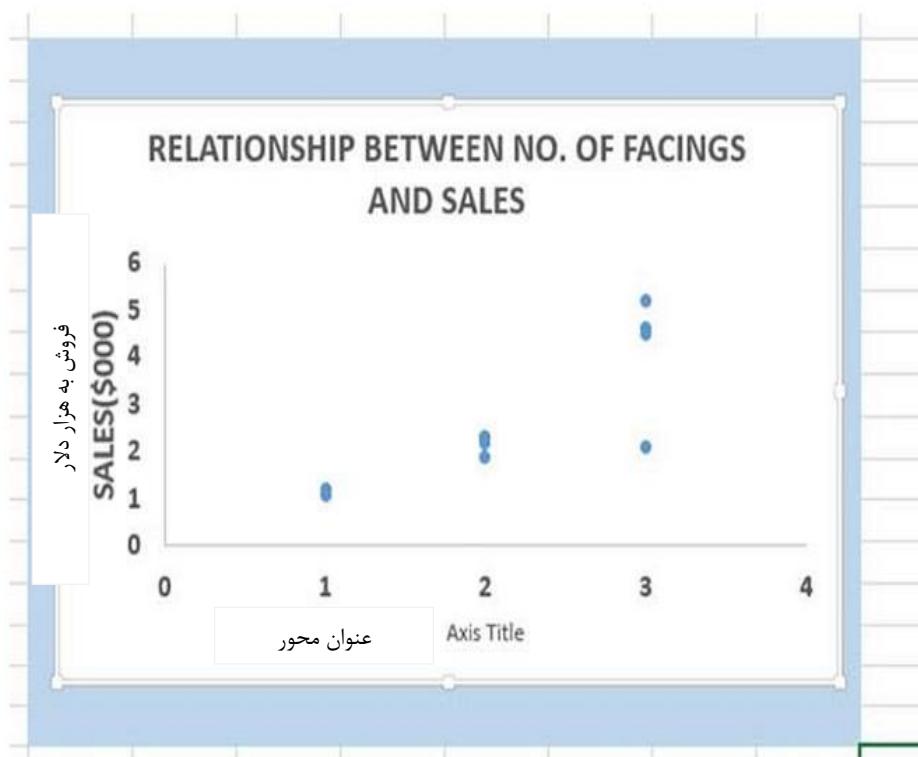
علامت تیک را به سمت چپ قسمت «عنوان محور» اضافه کنید. (این کار یک کادر "عنوان محور" در محور Y در نمودار ایجاد می‌کند)

روی فلاش سمت راست کلیک کنید: "عنوان محورها" و سپس روی "نخست افقی" کلیک کنید تا تیک داخل جعبه برداشته شود (این کار یک عنوان برای محور Y ایجاد خواهد کرد) عنوان محور Y زیر را در سمت راست f/x در بالای صفحه خود وارد کنید:

فروش (هزار دلار)

سپس کلید اینتر را فشار دهید تا این عنوان محور Y در نمودار وارد شود.

در داخل نمودار در گوشه بالا سمت چپ نمودار، کلیک کنید تا جعبه اطراف عنوان محور Y را از حالت انتخاب خارج کنید. (به شکل ۱۹-۶ مراجعه شود).



شکل ۱۹-۶: مثالی از اضافه کردن عنوان برای محور Y در نمودار

روی جعبه در سمت راست نمودار کلیک کنید.

"عنوان محورها" را مشخص کنید و روی فلش سمت راست آن کلیک کنید.

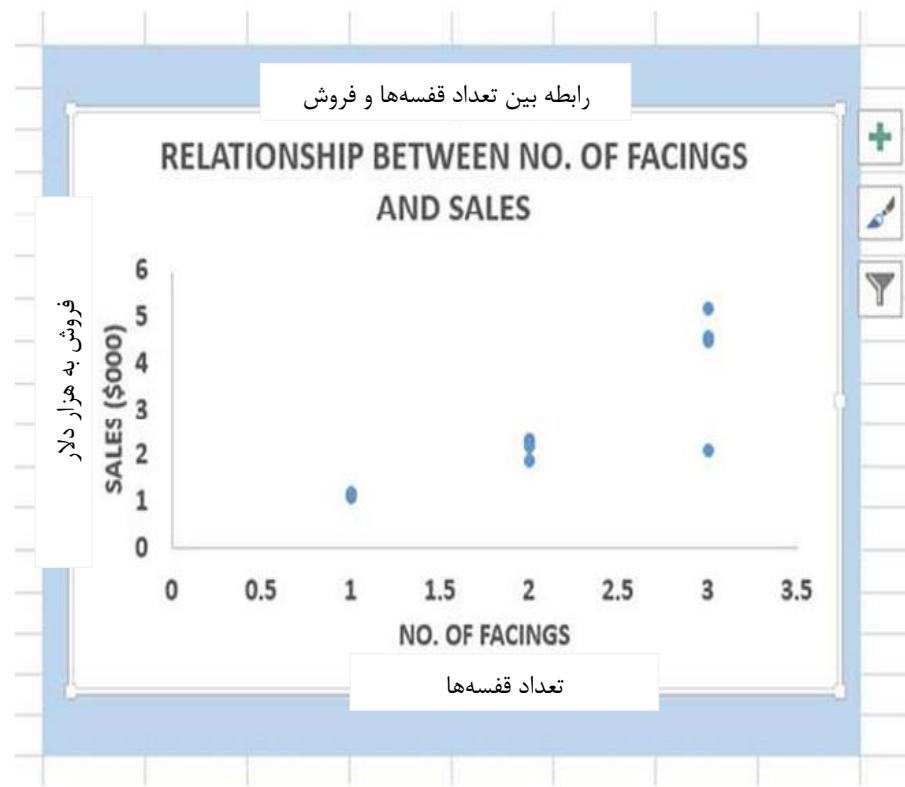
روی کلمات "نخست افقی" کلیک کنید تا یک تیک در جعبه آن اضافه شود. (این کار یک جعبه

"عنوان محور" برای محور X در نمودار ایجاد می‌کند).

عنوان محور X را در سمت راست $f(x)$ در بالای صفحه خود وارد کنید:
تعداد قفسه‌ها

سپس کلید اینتر را بزنید تا این عنوان محور X را به نمودار اضافه کنید.

در داخل نمودار در گوشه بالا سمت چپ نمودار کلیک کنید تا کادر دور عنوان محور X از حالت انتخاب خارج شود (به شکل ۲۰-۶ مراجعه شود).

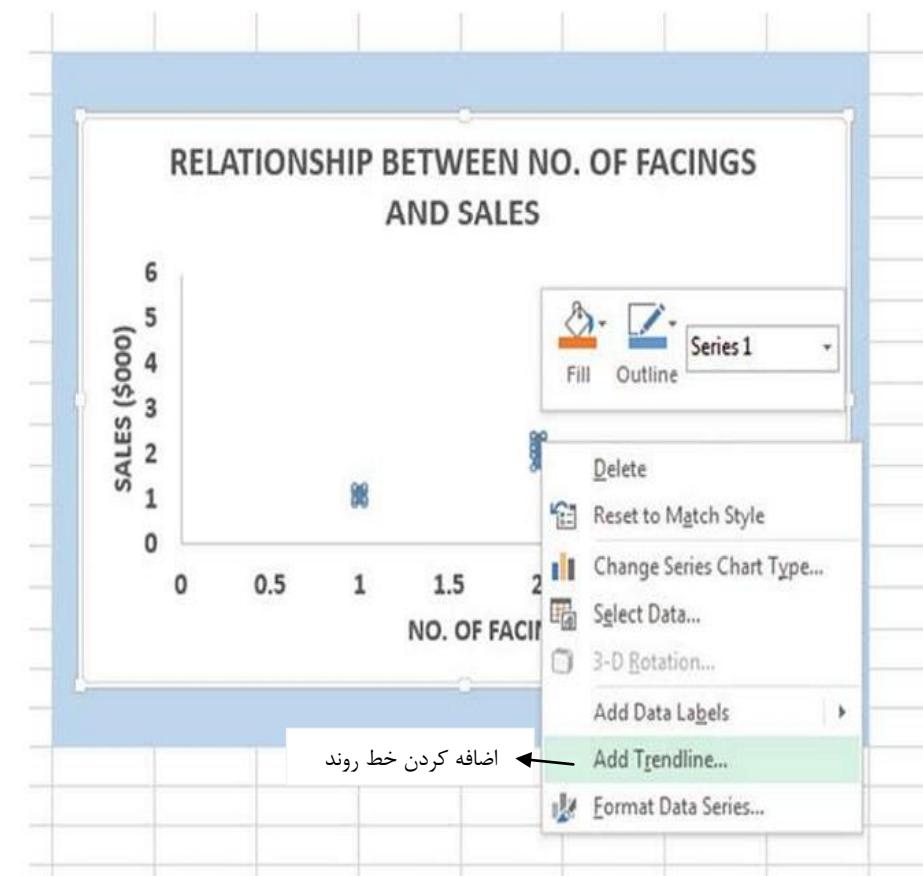


شکل ۲۰-۶: مثالی از یک عنوان نمودار، عنوان محور X و عنوان محور Y

۱-۱-۳-۶- رسم خط رگرسیون از طریق نقاط داده‌ها در نمودار

هدف: رسم خط رگرسیون از طریق نقاط داده‌ها در نمودار

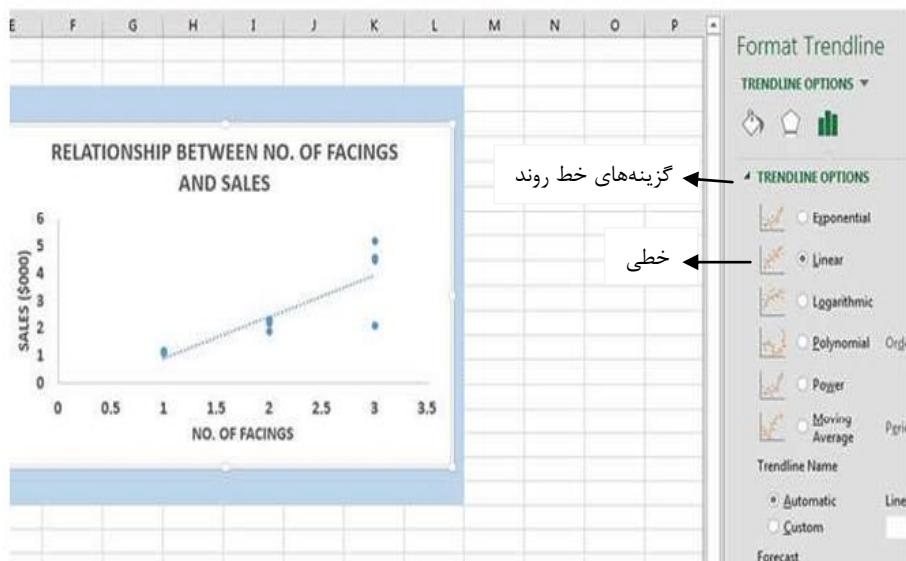
بر روی هر یک از نقاط داده در داخل نمودار کلیک راست کنید
گزینه اضافه کردن خط روند (Add Trendline) را انتخاب کنید (به شکل ۲۱-۶ مراجعه شود).



شکل ۲۱-۶: جعبه محاوره‌ای برای اضافه کردن خط روند به نمودار

روی Add Trendline کلیک کنید.

انتخاب حالت خطی (linear) (مطمئن شوید گزینه «خطی» نزدیک به بالا در جعبه محاوره‌ای انتخاب شده است؛ به شکل ۲۲-۶ مراجعه شود).

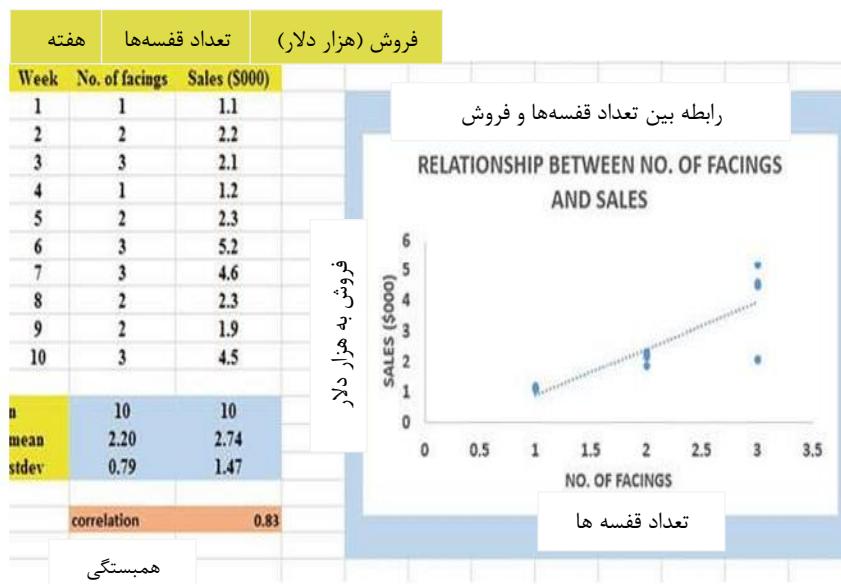


شکل ۲۲-۶: جعبه محاوره‌ای برای ایجاد روند خطی

برای بستن این کادر محاوره‌ای، روی X در سمت راست بالای کادر "Format Trendline" کلیک کنید.

بر روی هر سلوی خالی خارج از نمودار کلیک کنید تا نمودار را از حالت انتخاب خارج کنید. این فایل را به عنوان FACINGS7 ذخیره کنید.

صفحه گسترده شما باید مانند صفحه گسترده در شکل ۲۳-۶ ظاهر شود.

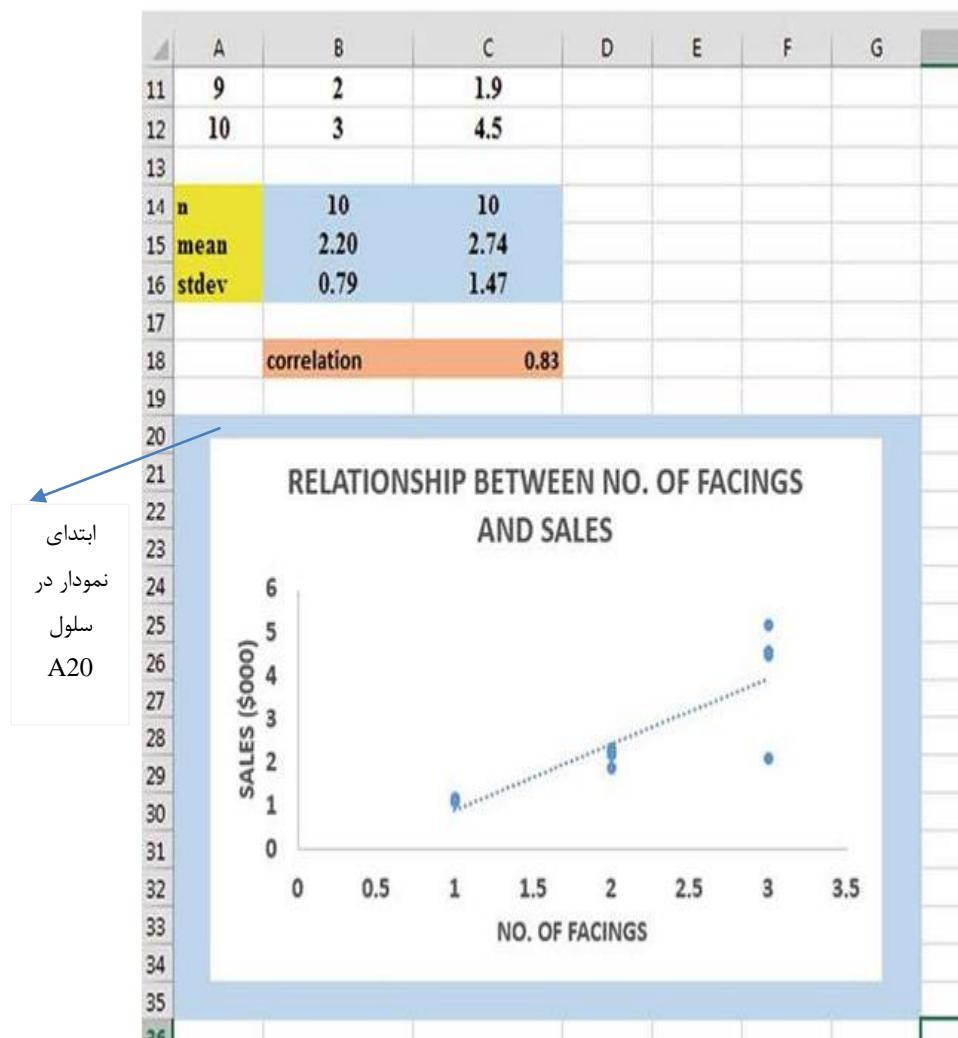


شکل ۲۳-۶: نمودار نهایی، خط روند که از طریق نقاط داده‌های نمودار پراکنش رسم شده است.

۲-۱-۳-۶- جابجایی نمودار در زیر جدول صفحه گسترده

هدف: جابجایی نمودار در زیر جدول

با ماوس خود روی هر فضای سفید در سمت راست بالای عنوان در داخل نمودار کلیک چپ کنید. کلیک چپ را پایین نگهدارید و نمودار را به پایین و چپ بکشید تا گوشه بالایی سمت چپ نمودار در سلول A20 باشد. سپس انگشت خود را از کلیک چپ ماوس بردارید (به شکل ۲۴-۶ مراجعه شود)



شکل ۲۴-۶: مثالی از جابجایی نمودار به زیر جدول

۳-۱-۳-۶- ساخت نمودار "طویل تر" به طوری که بلندتر باشد

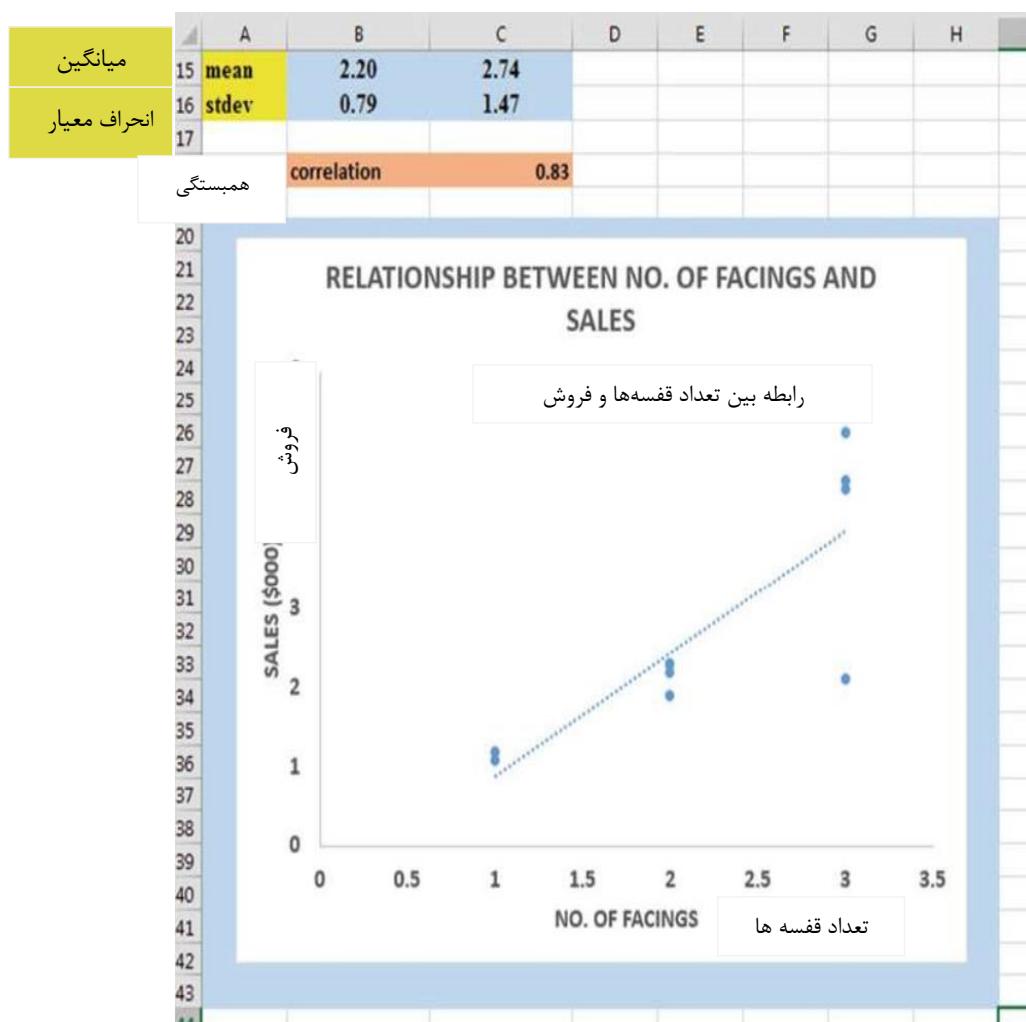
هدف: ساخت نمودار "طویل تر" به طوری که بلندتر باشد

نشانگر ماوس خود را در قسمت پایین و مرکز نمودار قرار دهید تا علامت نماد بالا-پایین ایجاد شود، به منظور ساخت نمودار بلندتر کلیک چپ ماوس را نگهدارید و نمودار را تا ردیف ۴۲ پایین بکشید و سپس انگشت خود را از روی ماوس بردارید.

"۶-۳-۱-۴- ساخت نمودار "پهن تر"

هدف: ساخت نمودار "پهن تر"

برای عریض‌تر کردن نمودار، اشاره‌گر را در وسط مرز راست نمودار قرار دهید تا یک نماد "پیکان چپ به راست" ایجاد شود، سپس کلیک چپ ماوس خود را فشار داده و درحالی‌که آن را نگه داشته‌اید، مرز راست نمودار را تا وسط ستون H بکشید (شکل ۶-۲۵ را مشاهده کنید).



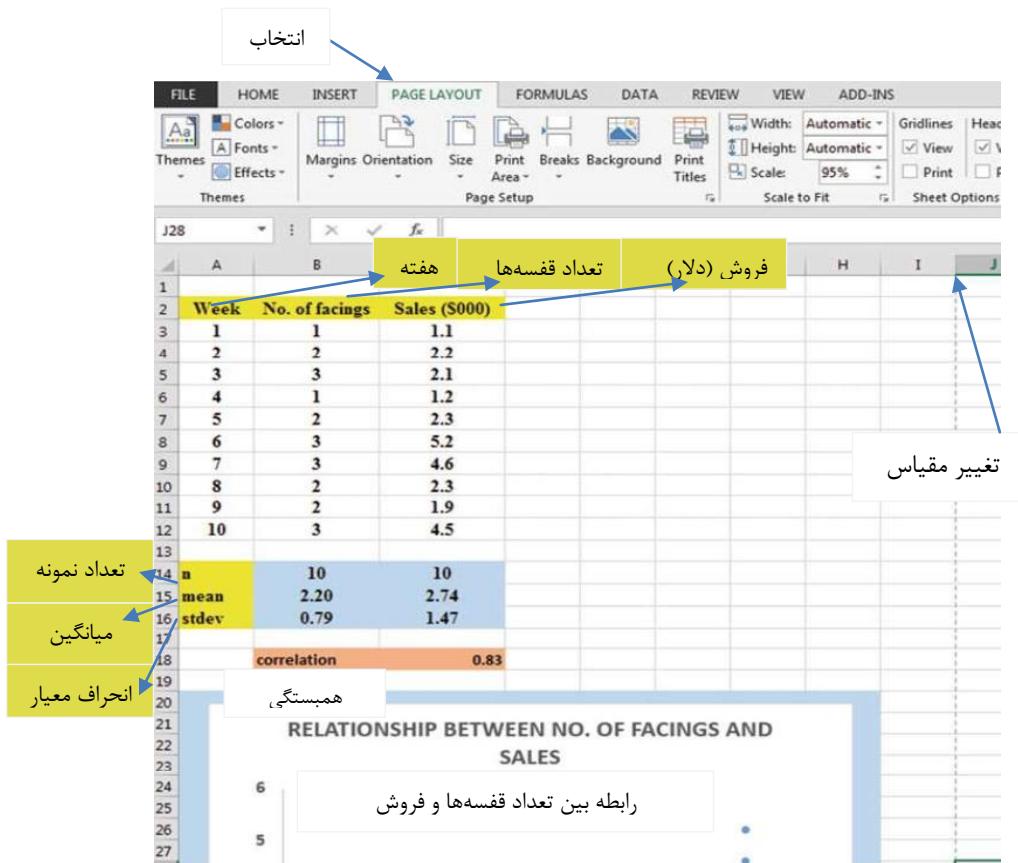
شکل ۶-۲۵: مثال از نموداری که بزرگتر شده تا متناسب با سلول‌ها شود: A20:H24

اکنون، روی هر سلول خالی خارج از نمودار کلیک کنید تا نمودار از حالت انتخاب خارج شود.

۴-۶- چاپ کردن صفحه گسترده به طوری که جدول و نمودار در یک صفحه باشند

هدف: چاپ کردن صفحه گسترده به طوری که جدول و نمودار در یک صفحه باشند.

مقیاس (بالای صفحه‌نمایش) را انتخاب کنید Page Layout
مقیاس را با استفاده از آیکون "Scale to Fit" نزدیک به بالای صفحه‌نمایش، به وسیله کلیک بر روی پیکان پایین تا زمانی که ۹۵٪ را بخواند، تغییر دهید. بدین ترتیب نمودار و جدول بر روی یک صفحه در صفحه‌نمایش شما جای خواهد گرفت. (به شکل ۲۶-۶ مراجعه شود)

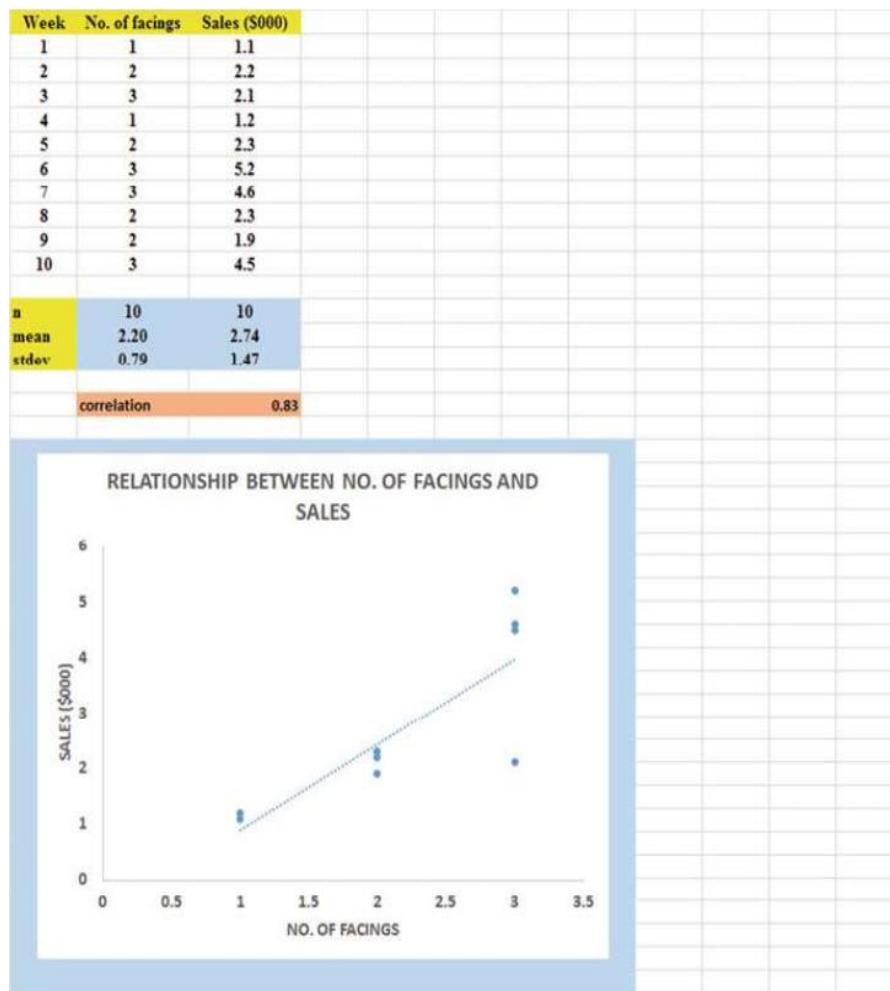


شکل ۲۶-۶: مثال ۲۶-۶ برای کاهش مقیاس نمودار به ۹۵٪ اندازه عادی

(فایل) را انتخاب کنید

Print (چاپ) را انتخاب کنید

چاپ (به شکل ۲۷-۶ مراجعه شود)



شکل ۲۷-۶: صفحه گسترده نهایی از نمودار و جدول (مقیاس ۹۵٪ برای متناسبسازی اندازه)

فایل خود را با عنوان **FACINGS8** ذخیره کنید.

۶-۵-۶- یافتن معادله رگرسیون

هدف اصلی از کشیدن نمودار برای رابطه‌ی بین X و Y (یعنی تعداد قفسه‌ها به عنوان X و فروش هزار دلار) به عنوان Y در مثال ما) این بود که ببینیم آیا یک رابطه قوی بین X و Y وجود دارد به طوری که معادله رگرسیون که می‌تواند این رابطه را خلاصه کند، در پیش‌بینی Y براساس ارزش X مفروض، استفاده شود.

می‌دانیم که ضریب همبستگی بین تعداد قفسه‌ها و فروش، $+0.83$ است، این به ما می‌گوید که استفاده از قفسه‌ها برای پیش‌بینی فروش هفتگی منطقی به نظر می‌رسد که بر اساس داده‌های قبلی ما چنین انتظاری می‌رفت.

اکنون ما باید معادله رگرسیون را پیدا کنیم که معادله "بهترین برازش کننده خطی" از طریق نقاط داده شده است.

هدف: برآورد معادله رگرسیون که رابطه بین X و Y را خلاصه کند

برای پیدا کردن این معادله، باید بررسی کنید که آیا نسخه اکسل شما "بسته‌ابزار تجزیه و تحلیل داده"^۱ که برای اجرای تحلیل رگرسیون لازم است را داشته باشد.

۶-۵-۱- نصب "بسته‌ابزار تجزیه و تحلیل داده" بر روی اکسل

هدف: نصب "بسته‌ابزار تجزیه و تحلیل داده" بر روی اکسل

از آنجایی که در حال حاضر چهار نسخه از اکسل در بازار وجود دارد^۲ (۲۰۱۰، ۲۰۰۷، ۲۰۰۳، ۲۰۱۳)، توضیح مختصری در مورد چگونگی نصب بسته‌ابزار تجزیه و تحلیل داده در هر یک از این نسخه‌های اکسل خواهیم داد.

۶-۵-۱-۱- نصب بسته‌ابزار تجزیه و تحلیل داده بر روی اکسل ۲۰۱۳

یک صفحه‌گسترده اکسل جدید باز کنید.
کلیک کنید بر روی: Data (بر روی بالای صفحه‌نمایش)

¹ Data Analysis Toolpack

². نرم‌افزار کاربردی اکسل دارای ویرایش‌های گوناگونی است که در حال حاضر نسخه‌های جدیدتر از آن در بازار موجود است اما از آنجایی که کتاب مورد نظر ترجمه است و مترجم می‌بایست به متن اصلی کتاب وفادار باشد، در متن اصلی نسخه ۲۰۱۳ به عنوان آخرین ویرایش نرم‌افزار اکسل معرفی شد.

به قسمت بالای صفحه‌نمایش رایانه خود نگاه کنید. آیا Data Analysis را در سمت راست صفحه‌نمایش خود می‌بینید؟ اگر می‌بینید پس Data Analysis برای اکسل ۲۰۱۳ زمانی که office 2013 را نصب کرده‌اید، به درستی نصب شده است و شما باید به بخش ۶-۵-۶ بروید. اگر کلمات "Data Analysis" در بالای صفحه‌نمایش شما نیست، پس قسمت ToolPack اکسل ۲۰۱۳ زمانی که Office 2013 را نصب می‌کردید بر روی کامپیوتر شما نصب نشده است. اگر چنین اتفاقی برای شما رخ داد، شما باید مراحل زیر را دنبال کنید:

File (فایل) را انتخاب کنید

گزینه Option (دکمه سمت چپ صفحه‌نمایش) را انتخاب کنید

توجه: این کار منجر به باز شدن جعبه محاوره‌ای "Excel Option" در سمت چپ بالای صفحه می‌شود.

(در سمت چپ صفحه) Add-Ins (اضافه) را انتخاب کنید

(در پایین کادر محاوره‌ای) Manage:Excel Add-Ins را انتخاب کنید

(پایین و مرکز کادر محاوره‌ای) GO (برو) را انتخاب کنید

انتخاب کنید: Analysis ToolPack (بسته‌ابزار تجزیه و تحلیل) (در کادر محاوره‌ای Add-Ins

گزینه Analysis ToolPack را علامت‌گذاری کنید

(در سمت راست کادر محاوره‌ای) OK را انتخاب کنید

Data (داده) را انتخاب کنید

برای اینکه ببینید این ویژگی به درستی نصب شده است، باید عبارت "Data Analysis" را در سمت راست بالای صفحه‌نمایش خود مشاهده کنید.

اگر پیام درخواست استفاده از لوح فشرده نصبی را دریافت کردید، لوح فشرده را درون رایانه قرار داده و بر روی گزینه OK کلیک کنید.

توجه: اگر این اقدامات به ثمر ننشست، گام‌های زیر را دنبال کنید: Options / File (سمت چپ) و پایین / Analysis / Toolpack / Add-ins / GO / Analysis گزینه

Toolpack را کلیک کنید تا علامت چک را به آن اضافه کنید / OK

اگر برای انجام این کار به کمک احتیاج دارید، از متخصص رایانه کمک بگیرید.
اکنون آماده هستید که به بخش ۶-۵-۶ بروید.

۶-۱-۵-۶- نصب بسته‌ابزار تجزیه و تحلیل داده بر روی اکسل ۲۰۱۰

یک صفحه گسترده جدید اکسل را باز کنید

بر روی Data (داده) (در بالای صفحه‌نمایش کلیک کنید)

به بالای صفحه‌نمایش نگاه کنید. آیا کلمات "Data Analysis" در سمت راست بالای صفحه‌نمایش می‌بینید؟ اگر آن را می‌بینید پس Data Analysis ToolPack برای اکسل ۲۰۱۰ زمانی که Office ۲۰۱۰ را نصب می‌کردید برای شما به درستی نصب شده است و باید به بخش ۲-۵-۶ بروید.

اگر کلمات "Data Analysis" در سمت راست بالای صفحه‌نمایش نیست، پس قسمت ToolPack اکسل ۲۰۱۰ زمانی که Office ۲۰۱۰ را نصب می‌کردید بر روی رایانه شما نصب نشده است. اگر چنین اتفاقی برای شما رخ داد، شما باید مراحل زیر را دنبال کنید:

File (فایل) را انتخاب کنید.

Option (گزینه) را انتخاب کنید.

Excel Option (یک کادر محاوره‌ای ایجاد می‌کند) را انتخاب کنید.

Add-Ins (اضافه) را انتخاب کنید.

(در پایین کادر محاوره‌ای) Manage:Excel Add-Ins (مدیریت) را انتخاب کنید.

GO را انتخاب کنید.

انتخاب کنید: Analysis ToolPack (بسته ابزار تحلیل) (در کادر محاوره‌ای Add-Ins

OK را انتخاب کنید.

Data (داده) را انتخاب کنید.

(باید عبارت "Data Analysis" در سمت راست و بالای صفحه خود داشته باشد).

اگر پیام استفاده از لوح فشرده را دریافت گردید، لوح فشرده را درون رایانه قرار داده و بر روی گزینه OK کلیک کنید.

توجه: اگر این اقدامات موثر واقع نشد، گام‌های زیر را دنبال کنید: Options / File (پایین سمت Analysis / GO / Analysis Toolpack / Add-ins) / چپ) / OK

Toolpack کلیک کنید تا علامت چک را به آن اضافه نمایید / OK.

اگر برای انجام این کار به کمک احتیاج دارید، از متخصص رایانه کمک بگیرید.

اکنون شما آماده هستید که به بخش ۲-۵-۶ بروید.

۲۰۰۷ - نصب بسته ابزار تجزیه و تحلیل داده بر روی اکسل ۲۰۰۷

یک صفحه گسترده جدید اکسل را باز کنید.

بر روی: Data (داده) (در بالای صفحه‌نمایش تان) کلیک کنید.

اگر کلمات "Data Analysis" در سمت راست بالای صفحه‌نمایش، نمایان نشد، باید برای نصب Data Analysis ToolPack مراحل زیر را دنبال کنید:

گزینه Microsoft Office (سمت چپ و بالای صفحه‌نمایش) را انتخاب کنید.
 Excel Option (پایین کادر محاوره‌ای) را انتخاب کنید.
 Add-Ins (گوشه سمت چپ کادر) را انتخاب کنید.
 GO (برای ایجاد جعبه محاوره‌ای برای Add-Ins) را انتخاب کنید.
 Analysis ToolPack انتخاب کنید.

OK (اگر اکسل از شما برای ادامه روند اجازه خواست، بر روی Yes کلیک کنید) را انتخاب کنید.
 Data (شما اکنون باید عبارت "Data Analysis" را در بالا سمت راست صفحه‌نمایش خود داشته باشد) را انتخاب کنید.
 اگر برای انجام این کار به کمک احتیاج دارید، از متخصص رایانه کمک بگیرید.
 اکنون آمده هستید که به بخش ۶-۵-۲ بروید.

۶-۱-۴- نصب بسته‌ابزار تجزیه و تحلیل داده بر روی اکسل ۲۰۰۳

یک صفحه گسترده جدید اکسل را باز کنید.

بر روی Tools (در بالای صفحه‌نمایش) کلیک کنید.

اگر کادر باز شده توسط Tools "Data Analysis" نمایش داد پس ToolPack هم اکنون بر روی این نسخه اکسل شما نصب است و آمده‌اید که معادله رگرسیون را برآورد کنید. اگر گزینه "Data Analysis" برای شما نمایش داده نشد، باید ToolPack را طبق دستور زیر نصب کنید:

بر روی File کلیک کنید.

Excel Option (پایین کادر محاوره‌ای) را انتخاب کنید.
 Add-Ins را انتخاب کنید.

Analysis ToolPack (این گزینه زیر افزونه غیرفعال Add-Ins بالای جعبه محاوره‌ای قرار دارد) را انتخاب کنید.
 GO را انتخاب کنید.

(برای اضافه کردن علامت چک به سمت چپ) Analysis ToolPack را کلیک کنید.
 OK را انتخاب کنید.

توجه: اگر مراحل بالا موثر واقع نشد، مراحل جایگزین زیر را امتحان کنید: Tools / Add-Ins / بر روی سمت چپ گزینه ToolPack کلیک کنید تا علامت چک را به آن اضافه کنید / OK / اکنون آمده‌اید که به بخش ۶-۵-۲ بروید.

۶-۵-۲- استفاده از اکسل برای پیدا کردن خلاصه خروجی رگرسیون

اکنون بسته‌ابزار را نصب کرده‌اید و آماده‌ی یافتن معادله رگرسیون برای بهترین برآورد کننده خطی، برای نقاط داده‌ها با استفاده از مراحل زیر هستید:

فایل اکسل FACINGS8 را باز کنید (اگر هم‌اکنون بر روی صفحه‌نمایش باز نیست)

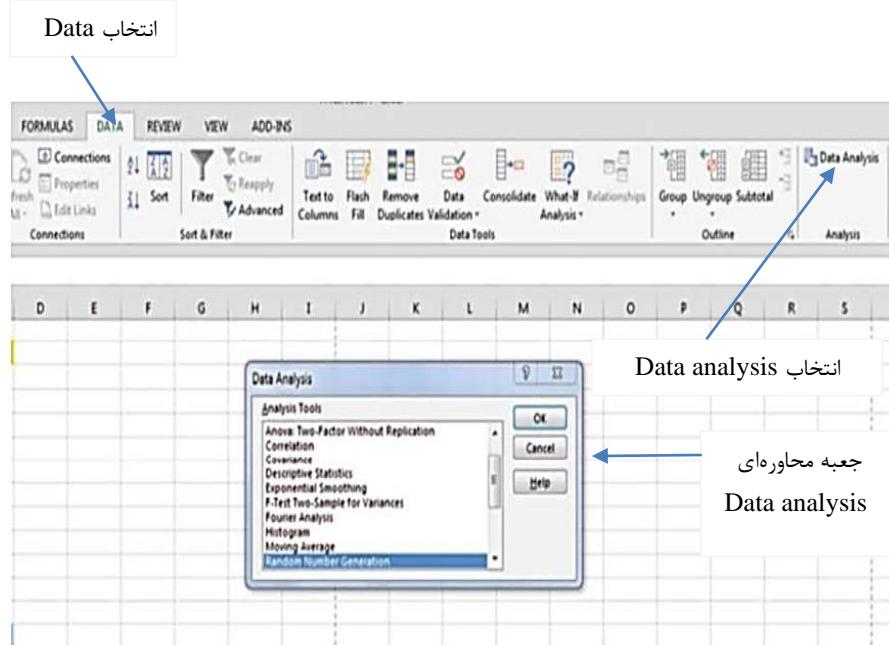
توجه: اگر این فایل در حال حاضر باز است و حاشیه خاکستری اطراف نمودار وجود دارد، شما باید بر روی هر سلول خالی خارج از نمودار کلیک کنید تا نمودار از حالت انتخاب خارج شود.

اکنون که شما بسته‌ابزار را نصب کردید، آماده‌اید تا معادله رگرسیون خلاصه کننده رابطه بین تعداد قفسه‌های ذرت کللوگ و فروش به دلار را برای داده‌ها به دست آورید.

به یاد داشته باشید که شما به داده X (پیش‌بینی کننده)، نام تعداد قفسه‌ها و به داده Y (معیار)، نام فروش را در قسمت قبلی این فصل اختصاص دادید (به بخش ۲-۶ مراجعه شود)

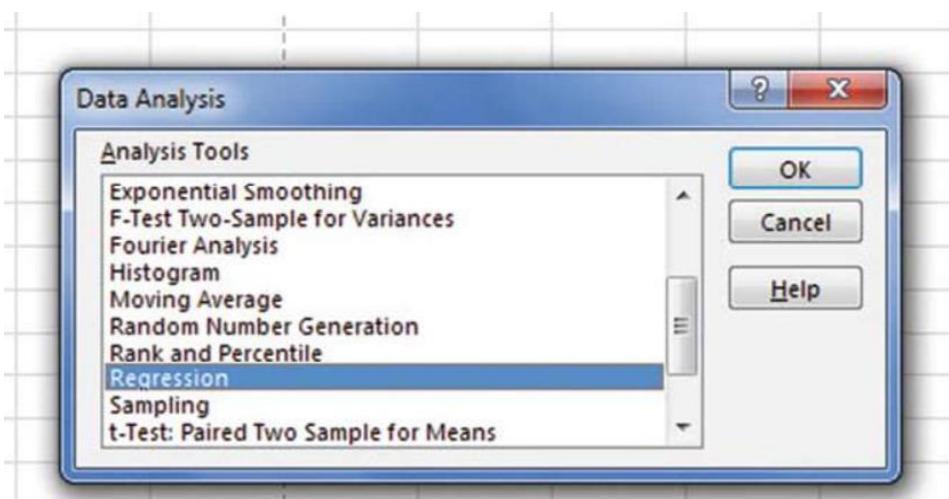
Data (داده) (بالای صفحه‌نمایش) را انتخاب کنید

Data analysis (گوشه بالا و سمت راست صفحه‌نمایش؛ شکل ۲۸-۶ را ببینید) را انتخاب کنید.



شکل ۲۸-۶: مثالی برای استفاده از تابع Data / Data analysis در اکسل

با استفاده از جهت پایین پیکان به پایین جعبه محاوره‌ای بروید و گزینه Regression (رگرسیون) را انتخاب کنید (شکل ۲۹-۶ را ببینید)



شکل ۶-۲۹: جعبه محاوره‌ای برای ایجاد رگرسیون در اکسل

OK را انتخاب کنید

وارد کردن بازه داده‌های Y: فروش

وارد کردن بازه داده‌های X: تعداد قفسه‌ها

بر روی گزینه سمت چپ Output Range (بازه خروجی) کلیک کنید تا انتخاب شود، و در جعبه باز شده A44 را به عنوان سلولی که خروجی تحلیل رگرسیون از آنجا در صفحه گسترده وارد شود، بنویسید

OK را انتخاب کنید

خروجی خلاصه اکنون باید در سلول‌های I61 تا A44 باشد

ستون A را عریض‌تر کنید به‌طوری‌که تمام اعداد موجود در خروجی خلاصه قابل خواندن باشند.

اکنون، با یکبار کلیک بر روی "Home" در قسمت بالا و سمت چپ صفحه‌نمایش‌تان داده‌های سه سلول زیر را به فرمت عددی تغییر دهید (تا ۲ رقم اعشار) :

B47

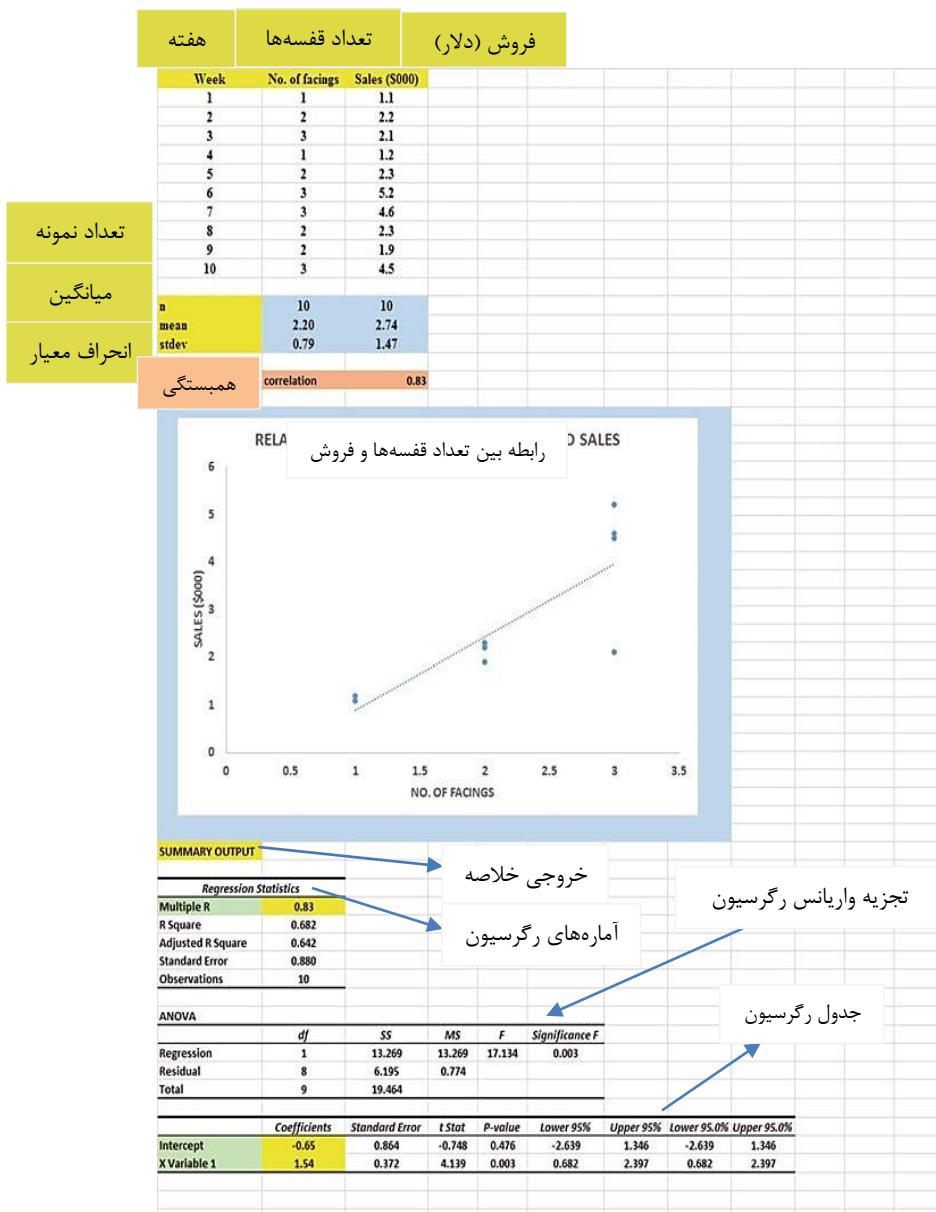
B60

B61

اکنون، فرمت تمام اعداد دیگری که به صورتی اعشاری هستند را تا سه رقم اعشار به صورت عددی تغییر دهید.

سپس، تمام ستون‌ها را به‌طوری عریض کنید که تمام برقسپ‌ها در داخل ستون‌ها متناسب با عرض ستون باشد. سپس تمام اعداد را در مرکز سلول‌ها قرار دهید.

فایل را طوری چاپ کنید که در یک صفحه باشد. (نکته: مقیاس را توسط "Page Layout" به ۷۰٪ برای متناسب کردن آن تعییر دهید). فایل شما باید شبیه فایل موجود در شکل ۳۰-۶ باشد.



شکل ۳۰-۶: شکل نهایی صفحه گسترده همبستگی و رگرسیون ساده خطی
 که شامل خروجی خلاصه برای داده‌ها می‌باشد

فایل نهایی را با عنوان FACINGS9 ذخیره کنید

به مشکل خروجی خلاصه زیر توجه کنید.

کسی که این نسخه از اکسل را برنامه‌نویسی کرده، مرتکب اشتباهی شده و به سلول A47 نام R (چند گانه) را داده است.

این صحیح نمی‌باشد. به جای این عبارت، باید در سلول "correlation r" A47 (همبستگی) نوشته شود، زیرا این نمادی است که ما برای همبستگی بین X و Y استفاده می‌کنیم. شما اکنون می‌توانید از صفحه چاپ شده تحلیل رگرسیونی خود برای پیدا کردن معادله رگرسیون که بهترین برازش‌کننده خطی داده‌ها است، بهره بگیرید. اما قبل از آن باید چند اصطلاح پایه‌ای را مرور کنیم.

۶-۵-۱- یافتن عرض از مبدأ y ، برای خط رگرسیون

نقاطهای بر روی محور y ها که خط رگرسیون در صورت امتداد، محور y را در آن نقطه قطع می‌کند عرض از مبدأ y نامیده می‌شود و ما از حرف a برای نشان دادن عرض از مبدأ خط رگرسیون استفاده می‌کنیم. عرض از مبدأ y در خروجی خلاصه در صفحه قبل ۰/۶۵ بود که در سلول B60 نمایش داده شد (به علامت منفی توجه کنید). این بدان معناست که اگر شما خطی فرضی از پایین خط رگرسیون به سمت محور y می‌کشیدید، این خط فرضی محور y را در ۰/۶۵ قطع می‌کرد.

۶-۵-۲- پیدا کردن شبیب، b ، برای خط رگرسیون

"کجی" خط رگرسیون را "شبیب" آن می‌نامند. این بیان می‌دارد که خط رگرسیون چند درجه بالا یا پایین خط افقی میان نقاط داده می‌باشد. اگر همبستگی بین X و Y صفر باشد، خط رگرسیون موازی محور X خواهد بود و شبیب آن برابر صفر است. اگر همبستگی بین X و Y مشتبث باشد، خط رگرسیون دارای شبیب به سمت بالا و راست در بالای محور X است.

از آنجایی که شبیب خط رگرسیون در شکل ۳۰-۶ دارای شبیب به سمت بالا و راست است، شبیب خط رگرسیون $+1/54$ می‌باشد همانطور که در سلول B61 آمده است. ما از علامت "b" برای نشان دادن شبیب خط رگرسیون استفاده می‌کنیم (توجه کنید که اکسل شبیب خط را در برگه چاپ "XVariable 1" می‌نامد). از آنجایی که همبستگی بین تعداد قفسه‌ها و فروش هفتگی به دلار بود، شما می‌توانید مشاهده کنید که خط رگرسیون برای این داده شبیب رو به بالا و راست

دارد. توجه کنید که خروجی خلاصه خط رگرسیون در شکل ۲۸-۶، همبستگی را با نشان داده و مقدار آن $0.83/0$ بوده و در سلول B47 می‌باشد.

اگر همبستگی بین X و Y منفی بود، خط رگرسیون شبیه روبه پایین و راست در بالای محور X ها خواهد داشت. این اتفاق ممکن است هر زمانی که همبستگی بین X و Y منفی باشد، رخ دهد که باید بین 0 و -1 باشد (0 و -1).

۳-۵-۶- نوشتن معادله برای خط رگرسیون

بهمنظور یافتن معادله رگرسیون برای خط مستقیمی که بتواند برای پیش‌بینی فروش هفتگی توسط تعداد قفسه‌ها استفاده شود، ما فقط به دو عدد در برگه خروجی خلاصه در شکل (۲۸-۶) نیاز داریم: B60 و B61.

$$\text{فرمت خط رگرسیون به صورت } Y = a + bX \text{ می‌باشد.} \quad (2-6)$$

به طوری که a = عرض از مبدأ ($0/65$) - در مثال ما که در سلول B60 قرار داشت) و b = شیب خط ($1/54$) در مثال ما که در سلول B61 قرار داشت)

بنابراین، معادله بهترین رگرسیون برازش کننده خطی برای مثال ما به صورت زیر است:

$$Y = a + bX$$

$$Y = -0.65 + 1.54X$$

به یاد داشته باشید که Y فروش هفتگی هزار دلار است که ما قصد پیش‌بینی آن را با استفاده از تعداد قفسه‌ها به عنوان پیش‌بینی کننده، X ، را داریم. باید با یک مثال و با استفاده از این رابطه فروش هفتگی را پیش‌بینی کنیم.

۴-۵-۶- استفاده از خط رگرسیون برای پیش‌بینی ارزش y به ازای ارزش x داده شده

هدف: یافتن فروش هفتگی پیش‌بینی شده توسط یک قفسه دانه ذرت کللوگ در فروشگاه

از آنجایی که تعداد قفسه‌ها برابر یک است (یعنی $1 = X$)، جایگذاری این عدد در معادله رگرسیون به صورت زیر خواهد بود:

$$Y = -0.65 + 1.54 \quad (1)$$

$$Y = -0.65 + 1.54$$

$$Y = 0.89$$

نکته مهم: اگر به نمودار نگاه کنید و مستقیماً از قفسه یک به سمت بالا بروید تا جایی که خط رگرسیون را قطع کنید، خواهید دید که این خط در زیر عدد ۱ بر روی محور y در سمت چپ

است (درواقع در ۰/۸۹) که همان نتیجه‌ای است که ما در بالا برای پیش‌بینی فروش از یک قفسه به دست آورده‌یم.

اما از آنجایی که فروش هفتگی به هزار دلار ثبت می‌شود، ما باید پاسخ بالای خود را در ۱۰۰۰ ضرب کنیم تا رقم فروش هفتگی را پیدا کنیم.

زمانی که ما این کار را انجام می‌دهیم، تخمین فروش هفتگی ۸۹۰ دلار (۰/۸۹*۱۰۰۰) از یک قفسه غلات حاصل می‌شود.

اکنون باید مثال دوم را انجام دهیم و پیش‌بینی کنیم که اگر ما ۳ قفسه ذرت کللوگ را استفاده کنیم مقدار فروش هفتگی در فروشگاه چقدر خواهد شد.

$$Y = -0.65 + 1.54 X$$

$$Y = -0.65 + 1.54 (3)$$

$$Y = -0.65 + 4.62$$

$$Y = 3.97$$

نکته مهم: اگر به نمودار نگاه کنید، مستقیماً از ۳ قفسه بالا بروید تا جایی که خط رگرسیون را قطع کنید، شما خواهید دید که این خط در زیر عدد ۴ بر روی محور y در سمت چپ است (درواقع ۳/۹۷) که همان نتیجه‌ای است که ما در بالا برای پیش‌بینی فروش از ۳ قفسه به دست آورده‌یم.

اما از آنجایی که فروش هفتگی به هزار دلار ثبت می‌شود، ما باید پاسخ بالای خود را در ۱۰۰۰ ضرب کنیم تا رقم فروش هفتگی را پیدا کنیم.

زمانی که ما این کار را انجام می‌دهیم، برای ما فروش هفتگی تخمینی ۳۹۷۰ دلار به ازای سه قفسه این غله حاصل می‌شود. برای دیدن جزئیات بیشتر در مورد بحث رگرسیون به بلاک (۲۰۱۰) رجوع کنید.

۶-۶- اضافه کردن معادله رگرسیون به نمودار

هدف: اضافه کردن معادله رگرسیون به نمودار

اگر می‌خواهید معادله رگرسیون را در کنار خط رگرسیون در نمودار قرار دهید، می‌توانید این کار را انجام دهید، اما مراقب یک مسئله باشید.

در این کتاب، از معادله رگرسیون برای یک پیش‌بینی کننده و یک معیار به صورت زیر استفاده می‌کنیم:

$$Y = a + bX$$

(۳-۶)

به طوری که $a = \text{عرض از مبدأ} y$ و $b = \text{شیب خط برای مثال معادله رگرسیون در بخش ۵-۶}$ ،
جایی که عرض از مبدأ y یعنی a برابر $-۰/۶۵$ بود و شیب خط یعنی b برابر $۱/۵۴$ بود منجر به
تولید معادله رگرسیون زیر شد:

$$Y = -0.65 + 1.54 X$$

با این حال، اکسل ۲۰۱۳ از یک معادله رگرسیون کمی متفاوت استفاده می‌کند (که اساساً با آنچه در
این کتاب استفاده شده، یکسان است) وقتی معادله رگرسیون را به نمودار اضافه می‌کنید.

$$Y = b X + a \quad (۴-۶)$$

به طوری که $a = \text{عرض از مبدأ} y$ و $b = \text{شیب خط}$

توجه داشته باشید که این همان معادله‌ای است که ما در این کتاب با اصطلاحات مرتب شده در یک
چیدمان متفاوت استفاده کردیم.

برای مثالی که ما در بخش ۵-۳ استفاده کردیم، اکسل ۲۰۱۳ ممکن است معادله رگرسیون در
نمودار را به صورت زیر نشان دهد:

$$Y = 1.54X - 0.65$$

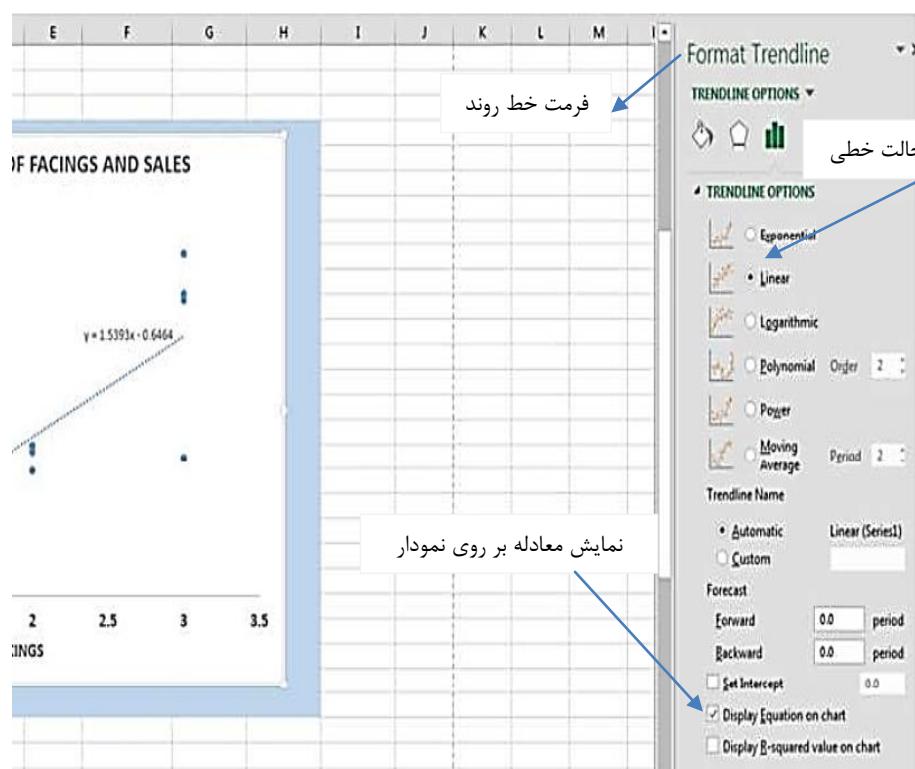
این فرمت زمانی که معادله رگرسیون را به نمودار در اکسل ۲۰۱۳ اضافه می‌کنید و مراحل زیر را
دنبال می‌کنید، به دست می‌آید:

فایل FACINGS9 را باز کنید (در بخش ۵-۲ آن را ذخیره کرده بودید)

برای اضافه کردن "حاشیه خاکستری" در اطراف نمودار فقط به داخل مرز بیرونی نمودار در گوشه
بالا سمت راست کلیک کنید تا نمودار انتخاب شود. برای تغییراتی که در پی آن هستید، بر روی
یکی از نقاط داده روی نمودار راست کلیک کنید

Add Trendline (اضافه کردن خط روند) را برجسته کنید و بر روی آن کلیک کنید تا دستور
اجرا شود.

گزینه Linear (خطی) در نزدیک بالای جعبه محاوره‌ای انتخاب شده است. (در سمت چپ آن)
بر روی Display Equation (نمایش معادله) در نمودار کلیک کنید (نزدیک به پایین جعبه
محاوره‌ای؛ شکل ۳۱-۶ را ببینید)



شکل ۳۱-۶: جعبه محاوره‌ای برای افزودن معادله رگرسیون به نمودار در کنار خط رگرسیون بر روی نمودار

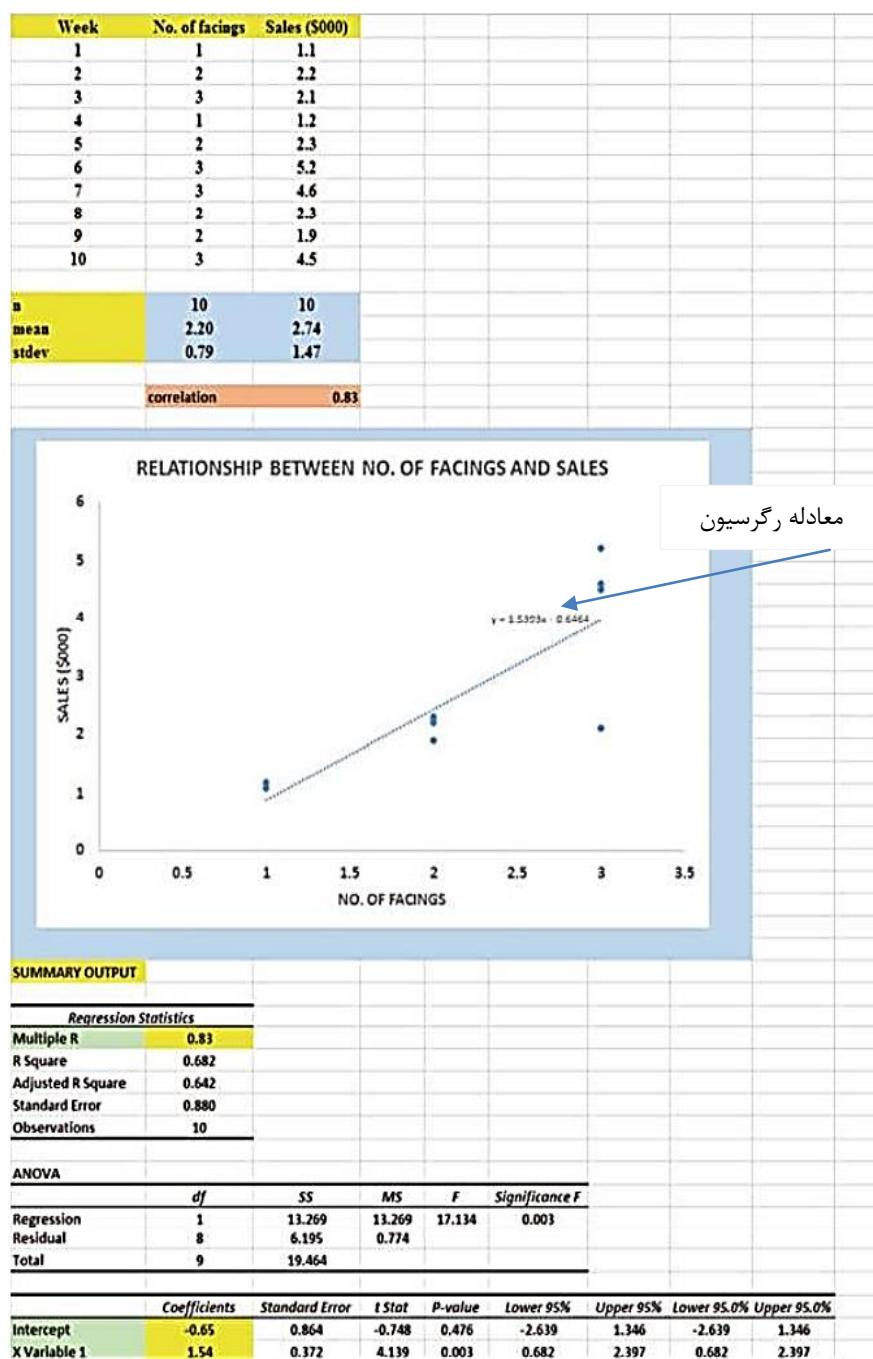
بر روی X در بالا و سمت راست جعبه محاوره‌ای Format Trendline (فرمت خط روند) برای حذف این کادر، کلیک کنید.

توجه کنید که معادله رگرسیون بر روی نمودار به شکل زیر در کنار خط رگرسیون بر روی نمودار می‌باشد. (شکل ۳۲-۶ را ببینید)

بر روی هر سلول خالی دیگری خارج از نمودار کلیک کنید تا نمودار از حالت انتخاب خارج شود.

$$Y = 1.54 \times -0.65$$

(فایل را با نام FACINGS10 ذخیره و چاپ کنید، بهطوری که در یک صفحه قرار گیرد)



شکل ۳۲-۶: مثالی از نمودار به همراه معادله رگرسیون نمایش داده شده در کنار خط رگرسیون

۶-۷- نحوه تشخیص همبستگی منفی در جدول خروجی خلاصه

نکته مهم: از آنجایی که اکسل همبستگی منفی را در خروجی خلاصه تشخیص نمی‌دهد، و تمام همبستگی‌ها را مثبت در نظر می‌گیرد (این مشکل از برنامه‌نویسی این برنامه می‌باشد)، باید مراقب باشد که شاید همبستگی بین X و Y منفی باشد، اگرچه خروجی خلاصه آن را مثبت نشان می‌دهد.

زمانی که یکی از دو حالت زیر اتفاق بیافتد، می‌توانید مطمئن باشید که همبستگی بین X و Y منفی است:

- (۱) شیب، b، عددی منفی باشد. این فقط زمانی اتفاق می‌افتد که همبستگی منفی باشد.
- (۲) نمودار به‌طور صریح شیب نزولی در خط رگرسیون داشته باشد، که این تنها زمانی رخ می‌دهد که همبستگی بین X و Y منفی باشد.

۶-۸- چاپ تنها بخشی از صفحه گسترده به جای تمام آن

هدف: چاپ قسمتی از صفحه گسترده به صورت جداگانه به جای چاپ تمام آن

موقعیت‌های بسیار زیادی وجود دارد که در صفحه گسترده‌تان از تعداد زیادی از سلول‌های اعداد برای داده‌هایتان و نمودارها استفاده کردۀاید و می‌خواهید تنها بخشی از آن را به‌طور جداگانه چاپ کنید بنابراین صفحه چاپ شده آنقدر کوچک نخواهد بود که بتوانید به راحتی آن را بخوانید. توضیح خواهیم داد که چگونه بخشی از صفحه گسترده را به‌طور جداگانه با استفاده از سه مثال و فایل FACINGS10 که در بخش ۶-۶ ایجاد کردید: (۱) چاپ تنها جدول و نمودار بر روی صفحه‌ای جداگانه، (۲) چاپ تنها نمودار در صفحه‌ای جداگانه، و (۳) چاپ تنها خروجی خلاصه تحلیل رگرسیون بر روی صفحه‌ای جداگانه.

توجه: اگر فایل FACINGS10 بر روی صفحه‌نمایش شما باز نیست، باید اکنون آن را باز کنید. اگر مرز خاکستری اطراف نمودار وجود دارد، روی هر فضای سفیدی خارج از نمودار کلیک کرده تا نمودار را از حالت انتخاب خارج کنید. اکنون باید چگونگی انجام این سه مورد را با سه هدف جداگانه توصیف کنیم.

۶-۸-۱- چاپ تنها جدول و نمودار بر روی صفحه‌ای جداگانه

هدف: چاپ تنها جدول و نمودار بر روی صفحه‌ای جداگانه

- (۱) با کلیک چپ ماوس و آغاز از بالا و سمت چپ جدول در سلول A2 شروع کرده و تا پایین آن بکشید به‌طوری که تمامی نمودار و جدول با رنگ آبی ملایم از سلول A2 تا

سلول I43 بر روی صفحه‌نمايش انتخاب شوند (سلول‌های برجسته شده، سلول‌های "انتخابی" نامیده می‌شوند).

(۲) File (فایل) را انتخاب کنید
 Print (چاپ) را انتخاب کنید
 Print Active Sheets (چاپ صفحه‌های فعال) را انتخاب کنید
 Print Selection (چاپ انتخابی) را انتخاب کنید
 Print (چاپ) را انتخاب کنید

خروجی نهایی باید شامل فقط جدول و نمودار حاصل از داده‌ها باشد.
 سپس، بر روی هر سلول خالی روی صفحه‌گسترده کلیک کنید تا نمودار از حالت انتخاب خارج شود.

۶-۸-۲- چاپ تنها نمودار بر روی صفحه‌ای جداگانه

هدف: چاپ تنها نمودار بر روی صفحه‌ای جداگانه

(۱) برای خاکستری شدن تمام حاشیه نمودار به منظور انتخاب آن، بر روی هر فضای سفیدی در داخل مرز بیرونی نمودار در گوشه سمت چپ نمودار کلیک کنید.

(۲) File (فایل) را انتخاب کنید
 Print (چاپ) را انتخاب کنید
 Print Selected chart (چاپ جدول انتخابی) را انتخاب کنید
 Print Selected chart (مجددأً)
 Print (چاپ) را انتخاب کنید
 Print (چاپ) را انتخاب کنید

خروجی نهایی باید فقط شامل نمودار حاصل از داده‌ها باشد.

نکته مهم: پس از هر بار چاپ نمودار به تنها‌ی در یک صفحه، شما باید فوراً بر روی هر فضای سفیدی خارج از نمودار به منظور حذف حاشیه خاکستری از حاشیه نمودار، کلیک کنید. زمانی که حاشیه‌های خاکستری بر روی مرزهای نمودار است، این کار به اکسل می‌گوید که شما فقط قصد چاپ نمودار به تنها‌ی را دارید.

۶-۳-۸-۳- چاپ تنها خروجی خلاصه تحلیل رگرسیون بر روی صفحه‌ای جداگانه

هدف: چاپ تنها خروجی خلاصه تحلیل رگرسیون بر روی صفحه‌ای جداگانه

(۱) بر روی سلول A43 که در بالای خروجی خلاصه در سمت چپ صفحه گسترده شما قرار دارد کلیک کرده و تا پایین و سمت راست ادامه دهید، تا جایی که تمام خروجی خلاصه با رنگ آبی تیره از سلول A43 تا I62 بر روی صفحه‌نمایش شما انتخاب شود.

(مقیاس را به ۷۵٪ تغییر دهید تا زمانی که خروجی خلاصه که آن را چاپ می‌کنید، بر روی یک صفحه باشد.)

(۲) File (فایل) را انتخاب کنید
Print (چاپ) را انتخاب کنید

Print Active Sheets (چاپ صفحه‌های فعال) را انتخاب کنید
Print Selection (چاپ انتخابی) را انتخاب کنید
Print (چاپ) را انتخاب کنید

خروجی نهایی باید شامل فقط خروجی خلاصه تحلیل رگرسیون بر روی یک صفحه جداگانه باشد. سرانجام، بر روی هر سلول خالی در صفحه گسترده برای خارج شدن از حالت انتخاب جدول رگرسیون کلیک کنید.

۶-۹- تمرین‌های عملی پایان فصل

۱. فرض کنید که توسط بلکباستر ویدئو^۱ استخدام شده‌اید تا یک معادله رگرسیون برای پیش‌بینی میانگین تعداد اجاره‌ها در روز از فروشگاه‌ها بر اساس درآمد متوسط خانواده برای خانواده‌هایی که در شعاع دو مایلی از فروشگاه‌های فعلی بلکباستر در ایالت میسوری هستند، ایجاد کنید. بلکباستر قصد دارد از این معادله برای پیش‌بینی فروش فروشگاه‌های برای فروشگاه‌های جدید که در حال بازگشایی آن‌ها در میسوری است، استفاده کند. داده‌های فرضی در شکل ۳۳-۶ را برای آزمون مهارت‌های اکسل خود، توسعه دهید.

^۱ Blockbuster Video

درآمد متوسط خانواده(هزار دلار)	اجاره‌ها (روزانه)
Average Family Income (\$000)	Rentals (per day)
62	705
41	525
27	309
45	498
50	623
47	425
44	314
28	203
30	465
41	540
47	605
62	690

شکل ۶-۳۳: صفحه کاری داده برای فصل ۶: تمرین عملی شماره ۱

یک صفحه گسترده اکسل ایجاد کنید و داده‌ها را وارد کنید، از درآمد به عنوان متغیر مستقل (پیش‌بینی کننده) و از تعداد اجاره‌های روزانه به عنوان متغیر وابسته (معیار) استفاده کنید.

(تذکر: به خاطر داشته باشید که متغیر مستقل X باید در ستون سمت چپ جدول باشد و متغیر وابسته Y باید در سمت راست جدول باشد.)

نکته مهم؛ زمانی که شما قصد یافتن همبستگی بین دو متغیر را دارید، مهم است که متغیر پیش‌بینی کننده X را در سمت چپ صفحه گسترده خود قرار دهید و متغیر وابسته Y را فوراً در سمت راست ستون X جایگذاری کنید. شما باید هر نوبت که می‌خواهید از اکسل برای یافتن همبستگی بین دو متغیر استفاده کنید، این کار را انجام دهید تا تفکر خود را بررسی کنید.

(a) از تابع correl = اکسل برای یافتن همبستگی بین این دو متغیر استفاده کرده و نتیجه را تا دو رقم اعشار گرد کنید.

(b) یک نمودار پراکنش XY برای این دو سری داده به صورت زیر ایجاد کنید:

- عنوان: رابطه بین درآمد و اجاره روزانه
- عنوان محور x: متوسط درآمد خانواده (هزار دلار)
- عنوان محور y: اجاره‌ها (روزانه)

- نمودار را مجدداً اندازه‌گیری کنید به طوریکه ۸ ستون و ۲۵ ردیف داشته باشد.
- نمودار را به زیر جدول منتقل کنید.

(c) رگرسیون حداقل مربعات را برای این داده‌ها بر روی نمودار پراکنش ایجاد کنید.

(d) از اکسل برای اجرای آماره‌های رگرسیون و پیدا کردن معادله خط رگرسیون حداقل مربعات برای این داده‌ها استفاده کنید و نتایج را زیر نمودار در صفحه گسترده نمایش دهید. از فرمت عددی (دو رقم اعشار) برای همبستگی و ضرایب استفاده کنید.

(e) داده‌های ورودی و نمودار را طوری چاپ کنید که اطلاعات آن‌ها در یک صفحه باشد. سپس، جدول خروجی رگرسیون را در یک صفحه مجزا طوری چاپ کنید که در همان یک صفحه جا شود.

(f) فایل را با نام RENTAL10 ذخیره کنید.

اکنون با استفاده از برگه چاپی اکسل به پرسش‌های زیر پاسخ دهید:

(۱) عرض از مبدأ y کدام است؟

(۲) شبی خط چقدر است؟

(۳) معادله رگرسیون برای این داده‌ها چیست (از دو رقم اعشار برای شبی و عرض از مبدأ استفاده کنید)؟

(۴) از معادله رگرسیون برای پیش‌بینی میانگین تعداد اجاره‌های روزانه که می‌توانید برای یک منطقه خردمند فروشی با متوسط درآمد خانوار ۵۰ هزار دلار، استفاده کنید.

۲. در یک شرکت مهندسی بزرگ، رابطه بین حقوق و دستمزد مهندسان به عنوان درصدی از حقوق متوسط مهندسان (موقعیت در بازه) و افزایشی که به مهندسان در قرارداد قبلی داده شده است، چیست؟ بازه حقوق مهندسان با ۱۰۰ امتیاز رتبه‌بندی شده، و حقوق هر مهندس با آن مقایسه می‌شود تا مشخص شود که چه درصدی از آن حقوق هر مهندس است. عدد حاصله "موقعیت در بازه" نامیده می‌شود. مهندسانی که حقوقی زیر نقطه میانه دارند، امتیازشان کمتر از ۱۰۰ و مهندسانی که امتیازشان بالای نقطه میانه می‌باشد، امتیازی بیشتر از ۱۰۰ دارند. فرض کنید که می‌خواهید این پرسش را مطالعه کنید. داده‌های فرضی که در شکل ۳۴-۶ آمده است را تحلیل کنید.



شکل ۳۴-۶: صفحه کاری داده‌ها برای فصل ۶: تمرین شماره ۲

یک صفحه گسترده اکسل ایجاد کرده و داده‌ها را وارد کنید.

(a) یک نمودار پراکنش XY برای این دو دسته داده ایجاد کرده به‌طوری که به شکل زیر باشد:

- عنوان: رابطه بین موقعیت در بازه و درصد افزایش برای مهندسین
- عنوان محور x: موقعیت در بازه
- عنوان محور y: درصد افزایش
- نمودار را به زیر جدول انتقال دهید.
- نمودار را دوباره اندازه‌گیری کنید به‌طوری که ۷ ستون و ۲۵ ردیف داشته باشد.

(b) خط رگرسیون حداقل مربعات را برای این داده‌ها در نمودار پراکنش ایجاد کنید.

(c) از اکسل برای اجرای آماره‌های رگرسیون به‌منظور پیدا کردن معادله خط رگرسیون حداقل مربعات برای این داده‌ها استفاده کنید و نتایج را زیر نمودار در صفحه گسترده نمایش دهید. معادله رگرسیون را به نمودار اضافه کنید. از فرمت عددی (۲ رقم اعشار) برای همبستگی و فرمت عددی (۳ رقم اعشار) برای ضرایب استفاده کنید.

فقط داده‌های ورودی و نمودار را به نحوی چاپ کنید که اطلاعات آن بر روی یک صفحه باشد.

سپس، فقط خروجی جدول رگرسیون را در صفحه‌ای جداگانه به‌نحوی که در همان یک صفحه باشد، چاپ کنید.

موارد زیر را به صورت دستی انجام دهید:

- (d) دور مقدار عرض از مبدأ و شبیه خط رگرسیونی در برگه خروجی خود دایره بکشید و نام‌گذاری کنید.
- (e) معادله رگرسیونی را به صورت دستی بر روی برگه خروجی خود برای این داده‌ها بنویسید (از سه رقم اعشار برای عرض از مبدأ و شبیه خط استفاده کنید).
- (f) دور همبستگی بین دو گروه از امتیازات در جدول خروجی خلاصه رگرسیون بر روی برگه چاپی خود دایره بکشید و نام‌گذاری کنید.
- (g) در زیر معادله رگرسیون در برگه چاپی خود به صورت دستی خط بکشید، با استفاده از معادله رگرسیونی درصد افزایشی که برای یک مهندس با امتیاز ۹۰ در موقعیت بازه قرار دارد را پیش‌بینی کنید.
- (h) از روی نمودار درصد افزایشی را که برای یک مهندس امتیاز ۱۱۰ در بازه موقعیت انتظار دارید را بخوانید و پاسخ خود را فوراً در فضای زیر آن بنویسید:
- (i) فایل را با عنوان ENGINE3 ذخیره کنید.

۳. آیا بین تعداد تماس‌هایی که یک کارمند فروش به قصد فروش محصولات با مشتریان بالقوه برقرار می‌کند و تعداد ماشین‌های کپی که توسط فروشنده‌ها در همان ماه به فروش می‌رسد، رابطه‌ای وجود دارد؟ فرض کنید داده‌های فرضی را که برای ماه قبل برای کارکنان فروش در زیر ارائه شده است را جمع‌آوری کرده‌اید.
داده‌های حاصل در شکل ۳۵-۶ ارائه شده است.

تعداد تماس‌های فروش No. of sales calls	تعداد ماشین‌های کپی فروخته شده No. of copiers sold
25	40
30	55
18	30
22	35
14	18
18	23
22	28
24	38
12	15
13	16
18	25
22	28
25	36

شکل ۶-۳۵: صفحه کاری داده‌ها برای فصل ۶: تمرین شماره ۳

یک صفحه‌گسترده اکسل ایجاد کرده و داده‌ها را وارد کنید، از تعداد تماس‌های فروش به عنوان متغیر مستقل (پیش‌بینی کننده) و از تعداد ماشین‌های کپی فروخته شده در ماه توسط هر فروشنده به عنوان متغیر وابسته (معیار) استفاده کنید.

(a) از تابع $=\text{correl}$ اکسل برای یافتن همبستگی بین این دو سری استفاده کنید و مقدار حاصل را تا دو رقم اعشار گرد کنید.

(b) یک نمودار پراکنش XY برای این دو سری از داده‌ها به صورت زیر ایجاد کنید:

- عنوان: رابطه بین تعداد تماس‌های فروش و ماشین‌های کپی فروخته شده
- عنوان محور X: تعداد تماس‌های فروش
- عنوان محور Y: تعداد ماشین‌های کپی فروخته شده
- نمودار را به زیر جدول انتقال دهید
- نمودار را مجدداً اندازه‌گیری کرده به طوری که ۷ ستون و ۲۵ ردیف داشته باشد

(c) خط رگرسیون حداقل مربعات برای این داده‌ها را در نمودار پراکنش ایجاد کنید.

(d) از اکسل برای اجرای آماره‌های رگرسیون به منظور پیدا کردن معادله خط رگرسیون حداقل مربعات برای این داده‌ها استفاده کنید و نتایج را زیر نمودار در صفحه‌گسترده نمایش دهید. از فرمت عددی (۲ رقم اعشار) برای همبستگی و ضرایب استفاده کنید.

(e) فقط داده‌های ورودی و نمودار را به طوری که در یک صفحه باشند، چاپ کنید. سپس، جدول خروجی رگرسیون را در صفحه‌ای جداگانه طوری چاپ کنید به طوری که در یک صفحه باشد.

(f) فایل را با عنوان copier4 ذخیره کنید

پرسش‌های زیر را با استفاده از برگه چاپی اکسل خود پاسخ دهید:

۱- رابطه بین تعداد تماس‌های فروش و تعداد ماشین‌های کپی فروخته شده چیست؟

۲- عرض از مبدا y چقدر است؟

۳- شبیه خط چقدر است؟

۴- معادله رگرسیونی چیست؟

۵- از معادله رگرسیون برای پیش‌بینی تعداد ماشین‌های کپی فروخته شده که انتظار دارید

از فروشنده‌ای که ۲۵ تماس در ماه گذشته گرفته بود، استفاده کنید. کار خود را در یک

صفحه جداگانه نمایش دهید.

منابع

- Black K. Business statistics: for contemporary decision making. 6th ed. Hoboken: Wiley; 2010.
- Levine DM, Stephan DF, Krehbiel TC, Berenson ML. Statistics for managers using Microsoft excel. 6th ed. Boston: Prentice Hall/Pearson; 2011.
- Zikmund WG, Babin BJ. Exploring marketing research. 10th ed. Mason: South-Western Cengage Learning; 2010.

فصل هفتم

همبستگی و رگرسیون چندگانه

بسیاری از موقع در کسب و کار می خواهید معیاری مانند Y را پیش بینی کنید، برای پیدا کردن یک الگوی پیش بینی بهتر، می توان بجای یک پیش بینی کننده مانند X ، ترکیب چندین پیش بینی کننده (به عنوان مثال X_1 ، X_2 ، X_3 وغیره) را به کار برد.

روش آماری حاصل "همبستگی چندگانه" نامیده می شود، زیرا از دو یا چند پیش بینی کننده به صورت ترکیبی بجای یک پیش بینی کننده مانند X برای پیش بینی Y استفاده می کند. هر پیش بینی کننده با توجه به همبستگی جداگانه با Y و سایر متغیرهای پیش بینی کننده، وزن متفاوتی دارد. وظیفه همبستگی چندگانه، تولید یک معادله رگرسیون است که هر پیش بینی کننده را به طور متفاوت وزن می دهد، به طریقی که ترکیبی از پیش بینی کننده ها عملکرد بهتری نسبت به یک پیش بینی کننده برای پیش بینی کردن Y دارند که ما آن را همبستگی چندگانه R_{xy} می نامیم.

نکته مهم: از فصل ۶ بخارطه دارید (به بخش ۱-۶ مراجعه شود) که همبستگی r در محدوده $-1 \leq r \leq 1$ است. بنابراین همبستگی می تواند یک عدد منفی نیز باشد.

بسیار مهم است که این واقعیت را به یاد داشته باشید که همبستگی چندگانه، R_{xy} تنها در محدوده $-1 \leq R_{xy} \leq 1$ است و هرگز نمی تواند یک عدد منفی باشد.

به یاد دارید (به بخش ۶-۵-۳ مراجعه کنید) که معادله رگرسیونی که Y را پیش بینی می کند زمانی که تنها از یک پیش بینی کننده مانند X استفاده کند، به صورت زیر است:

$$Y = a + bX \quad (1-7)$$

۱-۷ - معادله رگرسیون چندگانه

معادله رگرسیون چندگانه ساختار مشابه رگرسیون ساده را دنبال می کند و عبارت است از:

$$Y = a + b_1X_1 + b_2X_2 + b_3X_3 + \dots \quad (2-7)$$

بسته به تعداد پیش بینی کننده که مورد استفاده است این معادله ادامه خواهد داشت.

"وزن" داده شده به هر یک از پیش بینی کننده ها در معادله با فرمت b نمایش داده می شود و این حرف با یک زیرنویس به متغیر پیش بینی کننده مربوطه مرتبط می شود.

نکته مهم: برای انجام رگرسیون چندگانه، شما باید "بسته ابزار تجزیه و تحلیل داده" که در فصل ۶ توضیح داده شد (به فصل ۶-۵-۱ مراجعه شود) را نصب کرده باشید. اگر این بسته را نصب نکردید، باید این کار را اکنون انجام دهید.

باید یک مسئله عملی انجام دهیم.

فرض کنید که توسط یک شرکت اجاره خودرو استخدام شده‌اید تا ببینید آیا می‌توانید فروش سالانه را بر اساس تعداد خودروهایی که شرکت در ناوگان خود دارد و تعداد مکان‌هایی که می‌توان خودروهای شرکت را از آنجا اجاره کرد، پیش‌بینی کنید:

Y فروشن سالانه (بر حسب میلیون دلار)

X_1 تعداد خودروها در ناوگان شرکت (بر حسب هزار خودرو)

X_2 تعداد مکان‌های اجاره در ایالات متحده آمریکا

علاوه بر این، فرض کنید، شرکت اجاره خودرو، داده‌های زیر را که عملکرد شرکت و سایر رقبا را به طور خلاصه نشان می‌دهد در اختیار شما قرار داده است (به شکل ۷-۱ مراجعه شود):

شرکت‌های اجاره خودرو		
CAR RENTAL COMPANIES		
فروش (میلیون دلار)	تعداد خودروها (هزار خودرو)	تعداد مکان‌های اجاره
Y	X₁	X₂
SALES (\$millions)	NO. OF CARS (000)	NO. OF LOCATIONS
1070	120	152
1460	180	1120
1480	85	1032
552	92	440
2105	315	2587
308	71	1697
2380	221	1153
1140	142	922
43	25	105
154	35	1483
72	15	442
81	18	251
333	42	465
91	15	492
147	18	44

شکل ۷-۱: داده‌های صفحه کاری شرکت‌های اجاره خودرو (مسئله عملی)

معادله رگرسیون چندگانه

یک صفحه گسترده اکسل را برای این داده‌ها با استفاده از مرجع سلول زیر ایجاد کنید:

A3: شرکت‌های اجاره خودرو

Y: A5

A6: فروش (میلیون دلار)

۱۰۷۰ :A7

X1 :B5

B6: تعداد خودروها (هزار ماشین)

۱۲۰ :B7

X2 :C5

C6: تعداد مکان‌ها

۱۵۲ :C7

سپس، عرض ستون را برای مطابقت با جدول بالا تغییر دهید و تمام عدها را به قالب عدد تغییر دهید. (تعداد رقم اعشاری صفر).

اکنون، داده‌های اضافی زیر در نمودار را وارد کنید

۱۴۷ :A21

۱۸ :B21

C21: (سپس اطلاعات را در مرکز تمام سلول‌های جدول خود قرار دهید)

نکته مهم: مطمئن شوید که تمام عدها در جدول را دوبار بررسی کرده‌اید تا از صحت آن‌ها اطمینان حاصل شود در غیر این صورت صفحه گسترده شما نادرست خواهد بود.

این فایل را با عنوان RENTAL5 ذخیره کنید.

پیش از آنکه تجزیه و تحلیل رگرسیون چندگانه را انجام دهیم، نیاز است که یک نکته مهم را بررسی کنیم:

نکته مهم: هنگام استفاده از یک متغیر پیش‌بینی‌کننده مانند X، برای پیش‌بینی یک متغیر معیار مانند Y، باید مطمئن شوید که متغیر X در سمت چپ جدول شما و متغیر Y در سمت راست قرار دارد. با این روش، متوجه می‌شوید کدام متغیر پیش‌بینی‌کننده و کدام یک معیار است (به بخش ۳-۶ مراجعه شود).

اما، در رگرسیون چندگانه، باید این قاعده را دقیقاً به‌طور معکوس اجرا کنید:

هنگامی که از چندین پیش‌بینی‌کننده در رگرسیون چندگانه استفاده می‌کنید، ضروری است که معیاری که می‌خواهید پیش‌بینی کنید، مانند Y باشد و در منتهی‌الیه سمت چپ قرار بگیرید، و همه پیش‌بینی‌کننده‌ها سمت راست معیار Y در جدول قرار گیرند، در این صورت می‌دانید که کدام متغیر معیار Y و کدام متغیرها پیش‌بینی‌کننده‌ها هستند.

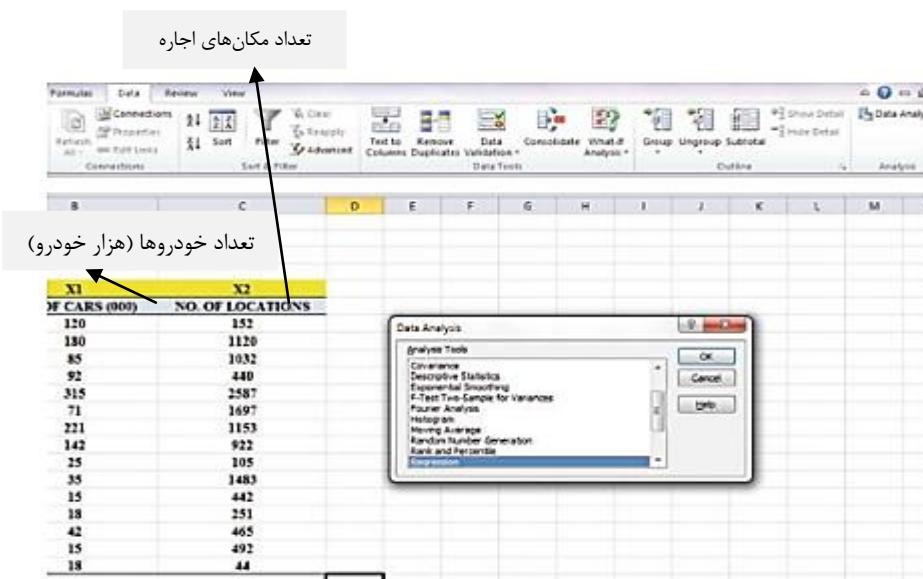
در جدول بالا توجه کنید که متغیر معیار Y (فروش) در منتهی‌الیه سمت چپ جدول قرار دارد و دو پیش‌بینی‌کننده (تعداد خودرو) و (تعداد مکان‌ها) در سمت راست متغیر معیار قرار دارند. شما باید از این قانون پیروی کنید، در غیر این صورت، معادله رگرسیون شما کاملاً اشتباه خواهد بود.

۲-۷- پیدا کردن همبستگی چندگانه و معادله رگرسیون چندگانه

هدف: پیدا کردن همبستگی چندگانه و معادله رگرسیون چندگانه با استفاده از اکسل.

شما این کار را با دستورات زیر انجام می‌دهید:
داده‌ها را انتخاب کنید.

بر روی Data Analysis کلیک کنید (منتهی‌الیه سمت راست بالای صفحه)
رگرسیون را انتخاب کنید (به سمت پایین در این جعبه بروید؛ به شکل ۲-۷ مراجعه شود)



شکل ۲-۷: جعبه محاوره‌ای برای رگرسیون

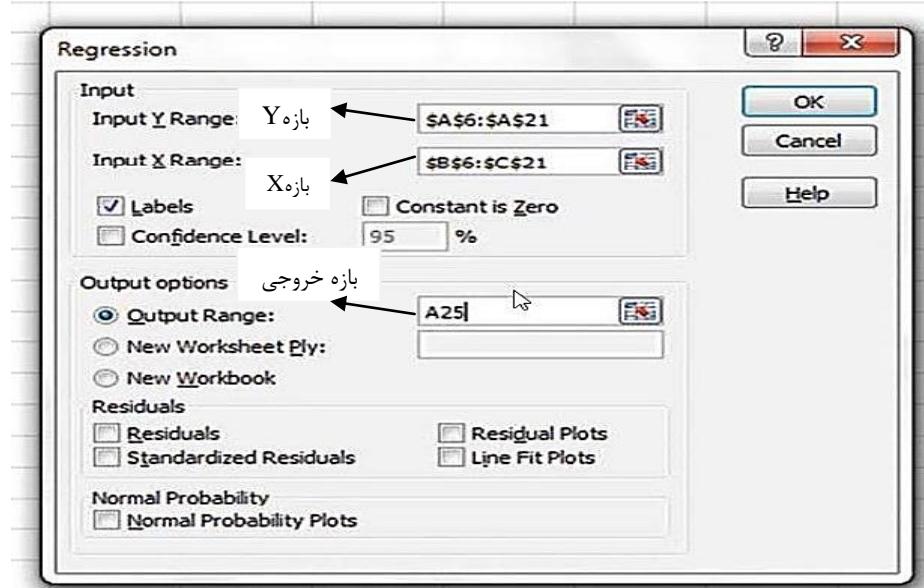
OK را انتخاب کنید

بازه داده‌های Y را وارد کنید:

B6:C21 بازه داده‌های X را وارد کنید:

بر روی جعبه برچسب (عنوان) کلیک کنید تا علامت چک را به آن اضافه کنید (چون شما برچسب (عنوان) ستون را در ردیف ۶ اضافه کرده‌اید)
بازه خروجی را انتخاب کنید (روی دکمه سمت چپ آن کلیک کنید و وارد کنید): A25 (به شکل ۳-۷ مراجعه شود)

پیدا کردن همبستگی چندگانه و معادله رگرسیون چندگانه
نکته مهم: اکسل به‌طور خودکار علامت دلار \$ را در مقابل هر حرف ستون و هر شماره ردیف قرار می‌دهد به‌طوری‌که می‌توانید این بازه داده‌ها را برای تجزیه و تحلیل رگرسیون، ثابت نگه دارید.



شکل ۳-۷: جعبه محاوره‌ای برای انجام رگرسیون کرایه خودرو را انتخاب کنید (برای دیدن خلاصه نتایج خروجی، به شکل ۴-۷ مراجعه شود)

خروجی خلاصه	A	B	C	D	E	F	G	H	I
	147	18	44						
	23	22	24	25	26	27	28	29	30
SUMMARY OUTPUT									
Regression Statistics									
Multiple R			0.93						
R Square			0.86						
Adjusted R Square			0.83						
Standard Error			321.49						
Observations			15						
ANOVA									
	df	SS	MS	F	Significance F				
Regression	2	7510945.33	3755473	36.33	8.10477E-06				
Residual	12	1240299.61	103358						
Total	14	8751244.93							
Coefficients									
Intercept	53.55	133.20	0.40	0.69	-236.66	343.76	-236.66	343.76	
NO. OF CARS (000)	9.09	1.34	6.78	0.00	6.17	12.01	6.17	12.01	
NO. OF LOCATIONS	-0.17	0.17	-0.98	0.34	-0.53	0.20	-0.53	0.20	

شکل ۷-۴: خروجی خلاصه رگرسیون کرایه خودرو

سپس، چهار سلول زیر را در فرمت عدد قالب‌بندی کنید (با ۲ رقم اعشار):

B28

B41

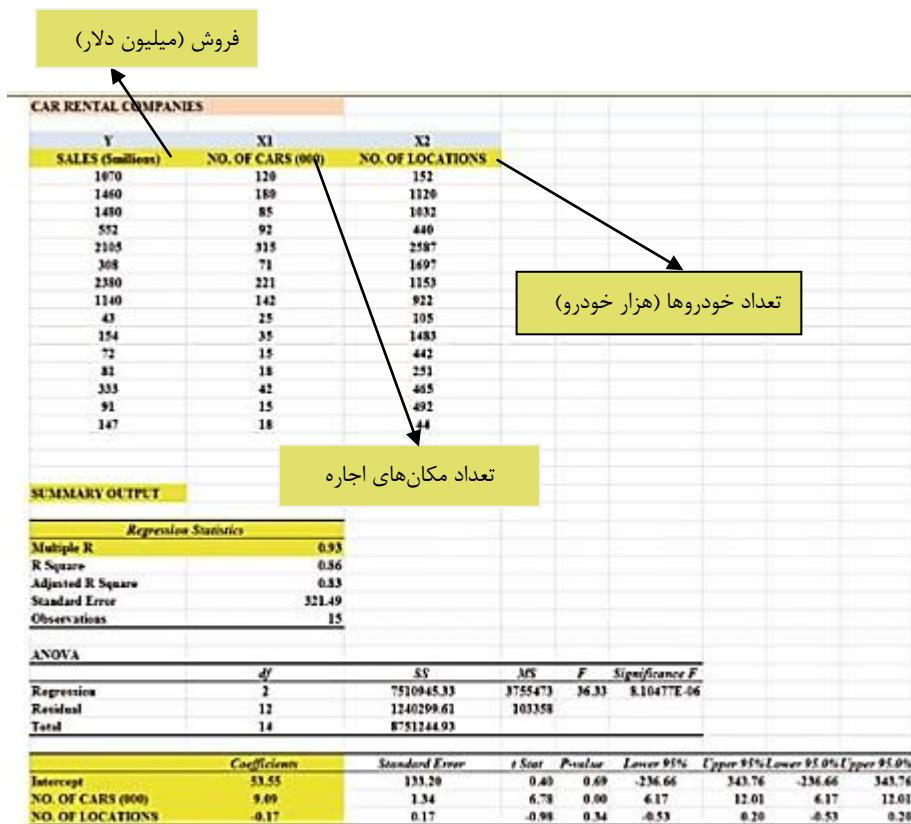
B42

B43

توجه داشته باشید که برای هر دو بازه ورودی داده‌های Y و X برچسب (عنوان) در بالای ستون‌ها وجود دارد.

دوباره فایل را با عنوان RENTAL5 ذخیره کنید

اکنون، فایل را به‌طوری‌که در یک صفحه جا شود از طریق تغییر مقیاس به اندازه ۶۰٪ چاپ کنید.
تجزیه و تحلیل رگرسیون حاصل در شکل ۵-۷ ارائه شده است.



شکل ۵-۷: صفحه گستردۀ نهایی برای تجزیه و تحلیل رگرسیون شرکت کرایه خودرو

هنگامی که خروجی خلاصه را دارید، می‌توانید همبستگی چندگانه و معادله رگرسیون را که بهترین خط برازش شده برای نقاط داده‌ها است را با استفاده از تعداد خودروها (هزار خودرو) و تعداد مکان‌های اجاره به عنوان دو متغیر پیش‌بینی‌کننده و فروش (میلیون دلار) به عنوان متغیر معیار، تعیین کنید.

توجه داشته باشید در خروجی خلاصه وقتی که گفته می‌شود: "R چندگانه"^۱، این اصطلاح صحیح است زیرا این عبارت توسط اکسل برای همبستگی چندگانه که برابر با $+0.93$ است، به کار می‌رود. این مقدار نشان می‌دهد که ترکیب تعداد خودروها و تعداد مکان‌های اجاره باهم یک رابطه بسیار قوی مشبت را برای پیش‌بینی فروش سالانه تشکیل می‌دهند.

¹ Multiple R

برای پیدا کردن معادله رگرسیون، به مقادیر ضرایب در پایین خروجی خلاصه توجه کنید:

۵۳/۵۵

مقدار ثابت: a (این مقدار ثابت y است)

۹/۰۹

تعداد خودروها (هزار ماشین): b_1

-۰/۱۷

تعداد مکان‌ها: b_2

از آنجاکه رابطه کلی معادله رگرسیون چندگانه به صورت زیر است:

$$Y = a + b_1 X_1 + b_2 X_2$$

اکنون می‌توانیم معادله رگرسیون چندگانه را برای این داده‌ها به قرار زیر، ارائه کنیم:

$$Y = 53.55 + 9.09 X_1 - 0.17 X_2$$

۳-۷- استفاده از معادله رگرسیون برای پیش‌بینی فروش سالانه

هدف: پیدا کردن فروش سالانه پیش‌بینی شده برای یک شرکت کرایه خودرو که دارای ۸۰,۰۰۰ خودرو و ۹۰۰ مکان اجاره است.

توجه داشته باشید که X_1 (تعداد خودروها) بر حسب هزار خودرو در داده‌های اصلی اندازه‌گیری شده است. در مثال ما، ۸۰۰۰۰ خودرو فقط ۸۰ خواهد بود، از آنجایی که ۸۰۰۰۰ معادل ۸۰ خودرو بوده و بر حسب هزار خودرو اندازه‌گیری شده است. پیوستن این دو عدد به معادله رگرسیون ما مقدار زیر را محاسبه می‌کند:

$$Y = 53.55 + 9.09 (80) - 0.17 (900)$$

$$Y = 53.55 + 727.2 - 153$$

$$Y = 627.75$$

اما، از آنجاکه فروش سالانه بر حسب میلیون دلار در داده‌های اصلی اندازه‌گیری شده است، باید این رقم را به میلیون دلار تبدیل کنیم. بنابراین، پیش‌بینی فروش سالانه شرکت کرایه خودرو که مشتریان می‌توانند خودروها را از آن اجاره کنند و دارای ۸۰۰۰۰ خودرو و ۹۰۰ مکان اجاره است، عبارت از:

$$627750000 \text{ دلار} \text{ یا } 627.75 \text{ میلیون دلار}$$

اگر می‌خواهید درباره نظریه رگرسیون چندگانه بیشتر بدانید، به کلر (۲۰۰۹) مراجعه کنید.

۴-۷- استفاده از اکسل برای ایجاد یک ماتریس همبستگی در رگرسیون

چندگانه

مرحله نهایی در رگرسیون چندگانه، پیدا کردن همبستگی بین تمام متغیرهایی است که در معادله رگرسیون ظاهر می‌شوند.

در این مثال، نیاز به پیدا کردن همبستگی بین هر یک از سه جفت متغیر زیر داریم:

- (۱) تعداد خودروها و فروش
- (۲) تعداد مکان اجاره و فروش
- (۳) تعداد خودروها و تعداد مکان اجاره

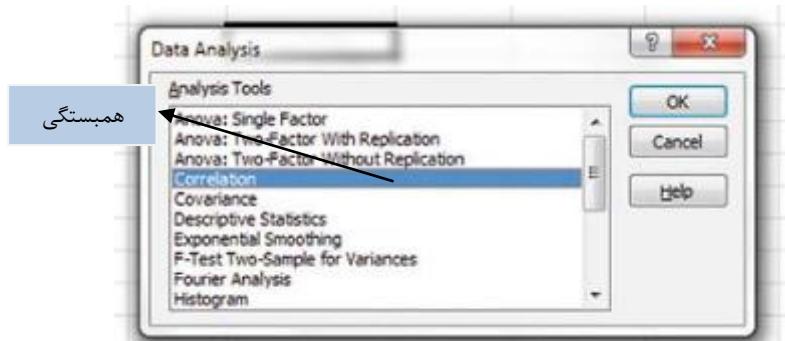
برای انجام این کار، ما نیاز به استفاده از اکسل برای ایجاد "ماتریس همبستگی" داریم. این ماتریس، سه همبستگی بالا را خلاصه می‌کند.

هدف: استفاده از اکسل برای ایجاد یک ماتریس همبستگی بین سه متغیر در این مثال.

برای استفاده از اکسل برای انجام این کار، از این مراحل استفاده کنید:

با (بالای صفحه زیر "Home" در سمت چپ بالای صفحه) Data را انتخاب کنید. Data Analysis را انتخاب کنید.

همبستگی (Correlation) (به بالا بروید تا این رابطه را های لایت کنید: به شکل ۶-۷ مراجعه شود) را انتخاب کنید.

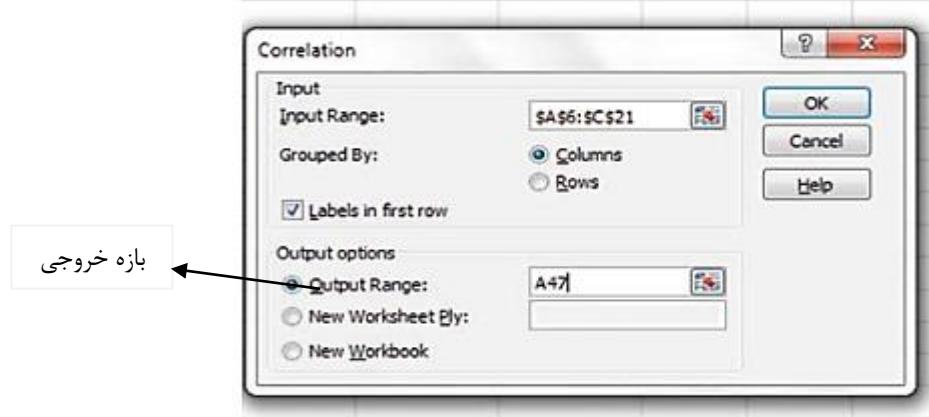


شکل ۶-۷: جعبه محاوره‌ای برای ماتریس همبستگی شرکت‌های اجاره خودرو

بازه ورودی داده‌ها: A6:C21
OK را انتخاب کنید

(توجه داشته باشید که بازه ورودی داده‌ها شامل برچسب‌هایی (عنوان‌هایی) در بالای سه متغیر مورد بررسی (فروش، تعداد خودروها، و تعداد مکان‌ها) و همچنین تمام ارقام موجود در داده‌های اصلی است).

گروه بندی شده بر اساس: ستون علامت چک را در جعبه برای برچسب‌ها (عنوان‌ها) در ردیف اول قرار دهید (از آنجاکه برچسب‌ها (عنوان‌ها) را در بالای ستون‌ها در بازه ورودی داده‌های بالا لحاظ کردۀاید) بازه خروجی (روی دکمه سمت چپ آن کلیک کنید و وارد کنید: A47) (به شکل ۷-۷ مراجعه شود)



شکل ۷-۷- جعبه محاوره‌ای برای بازه ورودی / خروجی ماتریس همبستگی

۸-۷ را انتخاب کنید. ماتریس همبستگی حاصل در A47: D50 ظاهر می‌شود (به شکل ۷-۷ مراجعه شود).

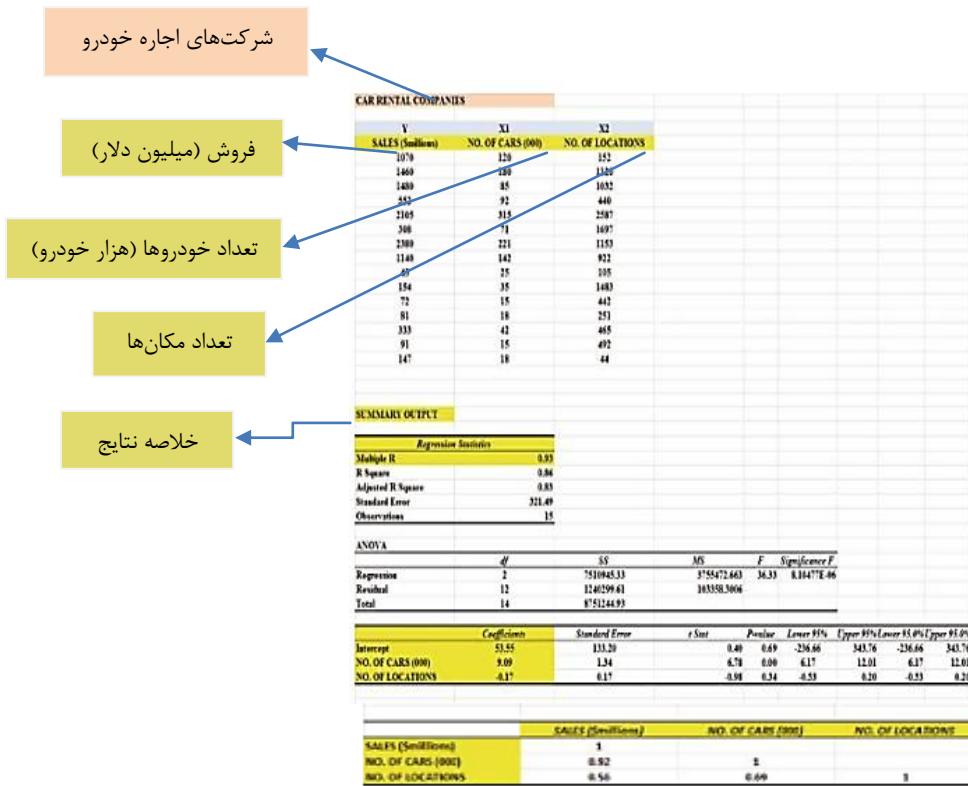
فروش(میلیون دلار)	فروش(میلیون دلار)	تعداد خودروها(هزار خودرو)	تعداد مکان‌های اجاره
SALES (\$millions)	SALES (\$millions)	NO. OF CARS (000)	NO. OF LOCATIONS
47	1		
48 SALES (\$millions)	0.920235314	1	
49 NO. OF CARS (000)	0.562140716		1
50 NO. OF LOCATIONS	0.694488326		

شکل ۸-۷: ماتریس همبستگی حاصل برای مثال شرکت کرایه خودرو

سپس، سه عدد موجود در ماتریس همبستگی را تا دو رقم اعشار اصلاح کنید. همچنین، ستون D را عریض‌تر کنید تا برچسب تعداد مکان‌ها در داخل سلول D47 جا شود.

این فایل اکسل را با عنوان RENTAL6 ذخیره کنید

صفحه گسترده نهایی برای شرکت کرایه خودرو در شکل ۹-۷ نمایش داده شده است.



شکل ۹-۷: صفحه گسترده نهایی برای رگرسیون شرکت‌های کرایه خودرو و ماتریس همبستگی

توجه داشته باشید که عدد "۱" در امتداد قطر اصلی ماتریس همبستگی بدان معنی است که همبستگی هر متغیر با خودش، یک همبستگی کامل و مثبت برابر ۱/۰ است.

ضریب همبستگی همیشه فقط تا دو رقم اعشار ارائه می‌شود.

شما اکنون آمده‌هستید تا همبستگی بین سه جفت متغیر را بخوانید:

همبستگی بین تعداد خودروها و فروش: +۰/۹۲

همبستگی بین تعداد مکان‌ها و فروش: +۰/۵۶

همبستگی بین تعداد خودروها و تعداد مکان‌ها: +۰/۶۹

این مقادیر نشان می‌دهد که تعداد خودروها پیش‌بینی کننده بهتر فروش، با همبستگی $+0/92$ است. با اضافه کردن متغیر پیش‌بینی کننده دوم، تعداد مکان‌ها، پیش‌بینی را به مقدار $+0/01$ تا $+0/93$ بهبود می‌دهد. بنابراین، ارزش تلاش اضافی را ندارد. تعداد خودروها پیش‌بینی بهتری برای فروش سالانه است.

اگر می‌خواهید درباره ماتریس همبستگی بیشتر بدانید، به لوین و همکاران (۲۰۱۱) مراجعه کنید.

۷-۵- تمرین‌های عملی پایان فصل

۱. آزمون استعداد تحصیلی (GRE) اغلب برای پیش‌بینی GPA سال اول دانشجویان در یک برنامه MBA استفاده می‌شود.

آزمون استعداد تحصیلی (GRE) یک آزمون استاندارد است که پیش شرط پذیرش برای بسیاری از کالج‌های ایالات متحده است که مدرک MBA را ارائه می‌دهند. آزمون GRE در پی اندازه-گیری آمادگی علمی عمومی، بدون در نظر گرفتن زمینه تخصصی است. در آزمون GRE سه نمره زیرآزمون به شرح ذیل محاسبه می‌شود:

(۱) GRE استدلال کلامی (مقیاس ۱۳۰-۱۷۰)

(۲) GRE استدلال کمی (مقیاس ۱۳۰-۱۷۰)

(۳) GRE نوشتاری تحلیلی (مقیاس ۶۰-۶۴).

فرض کنید که مدیر برنامه MBA از شما خواسته تا ارتباط بین این متغیرها را بر اساس کلاس ورودی سال گذشته و توانایی GRE برای پیش‌بینی میانگین نمره سال اول (GPA) به دست آورید.

شما تصمیم گرفتید از سه نمره زیرآزمون به عنوان پیش‌بینی کننده $X1$, $X2$ و $X3$ و میانگین نمره سال اول (GPA) سال اول) به عنوان متغیر معیار Y، استفاده کنید. برای آزمون مهارت‌های اکسل، شما به صورت تصادفی یک گروه کوچک از دانشجویانی که از سال گذشته وارد MBA شده‌اند، را انتخاب کردید. نمرات آن‌ها را برای این متغیرها ثبت شده است.

اما، فرض کنید که می‌خواهید بدانید چه اتفاقی می‌افتد اگر شما GPA دوره کارشناسی را به عنوان پیش‌بینی کننده چهارم اضافه کنید. همبستگی چندگانه چقدر خواهد بود؟ بیایید ببینیم وقتی که از داده‌های GPA دوره کارشناسی به عنوان پیش‌بینی کننده چهارم برای GPA سال اول دانشجویان دوره MBA، استفاده کنیم، چه اتفاقی می‌افتد. داده‌ها در شکل ۷-۷ آرائه شده است.

آزمون سابقه تحصیلی (GRE)

GRADUATE RECORD EXAMINATIONS (GRE)				
FIRST-YEAR GPA	GRE VERBAL	GRE QUANTITATIVE	GRE WRITING	UNDERGRAD GPA
3.25	160	161	5	3.40
3.42	156	158	4	3.15
2.85	156	157	2	3.05
2.65	154	153	1	2.55
3.65	166	166	6	3.25
3.16	159	160	3	3.20
3.56	166	163	4	3.66
2.35	155	154	2	2.55
2.86	153	154	3	2.85
2.95	158	157	4	2.80
3.15	158	159	4	3.05
3.45	160	160	5	3.44

شکل ۱۰-۷: داده‌های صفحه کاری برای فصل ۷: تمرین عملی شماره ۱

- (a) ایجاد یک صفحه گسترده اکسل با استفاده از GPA سال اول به عنوان متغیر معیار (Y) و سایر متغیرها به عنوان چهار پیش‌بینی‌کننده این متغیر.
- (b) استفاده از رگرسیون چندگانه اکسل برای پیدا کردن رابطه بین این متغیرها و قرار دادن آن در زیر جدول.
- (c) از فرمت عدد (با ۲ رقم اعشار) برای گزارش همبستگی چندگانه در خروجی خلاصه استفاده کنید، از فرمت عدد (با سه رقم اعشار) برای ضرایب استفاده کنید و از چهار رقم اعشار برای تمام ارقام اعشاری دیگر در خروجی خلاصه استفاده کنید.
- (d) جدول و نتایج رگرسیون زیر جدول را پس از تنظیم بر روی یک صفحه، چاپ کنید.
- (e) به صورت دستی بر روی این نسخه چاپی، دور موارد زیر را دایره بکشید و نام‌گذاری کنید: R_{xy} (1a)
- (f) عرض از مبداء Y، GRE کلامی، GRE کمی، GRE نوشتاری، GPA کارشناسی (2b)
- (f) این فایل را با عنوان GRE24 ذخیره کنید

(g) حالا، به فایل اکسل خود بروید و یک ماتریس همبستگی برای این پنج متغیر ایجاد کنید و در زیر خروجی خلاصه قرار دهید. هر مقدار همبستگی را فقط تا دو رقم اعشار تغییر دهید. این فایل را دوباره با عنوان GRE24 ذخیره کنید.

(h) اکنون، فقط این ماتریس همبستگی را در حالت عمودی بر روی یک برگه جداگانه چاپ کنید. با استفاده از نسخه چاپی اکسل خود به سوالات زیر پاسخ دهید:

۱- همبستگی چندگانه R_{XY} چقدر است؟

۲- مقدار عرض از مبدأ y یعنی a چقدر است؟

۳- ضریب GRE کلامی یعنی b_1 چقدر است؟

۴- ضریب GRE کمی یعنی b_2 چقدر است؟

۵- ضریب GRE نوشتاری یعنی b_3 چقدر است؟

۶- ضریب GPA یعنی b_4 چقدر است؟

۷- معادله رگرسیون چندگانه را ارائه کنید؟

۸- در زیر این معادله رگرسیونی به صورت دستی، پیش‌بینی GPA سال اول را به دست آورید مقدار انتظاری نمره GRE کلامی برابر ۱۵۹، نمره GRE کمی برابر ۱۵۴، نمره GRE نوشتاری برابر ۴ و نمره GPA کارشناسی برابر ۳۰۵ است.

با استفاده از نسخه چاپی اکسل خود به سوالات زیر پاسخ دهید. اطمینان حاصل کنید که علامت مثبت یا منفی برای هر همبستگی لحاظ شده است:

۹- همبستگی بین GPA کارشناسی و GPA سال اول چقدر است؟

۱۰- همبستگی بین GPA کارشناسی و GPA کلامی چقدر است؟

۱۱- همبستگی بین GPA کارشناسی و GPA کمی چقدر است؟

۱۲- همبستگی بین GPA کارشناسی و GPA نوشتاری چقدر است؟

۱۳- بحث کنید که کدام‌یک از چهار پیش‌بینی کننده، بهترین پیش‌بینی کننده GPA سال اول است.

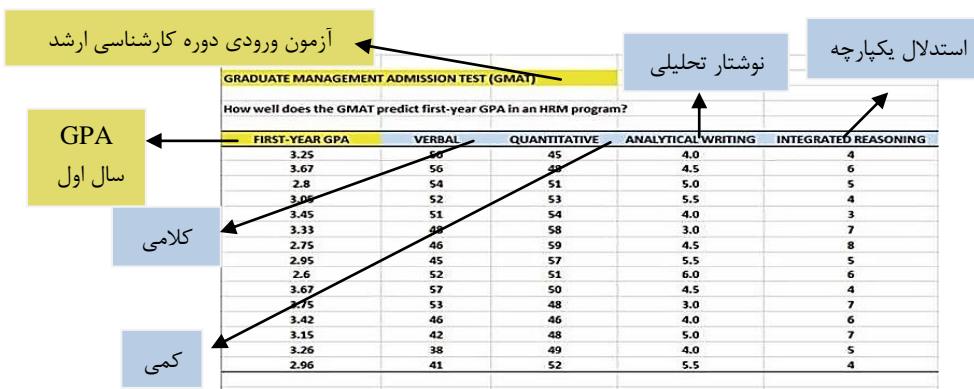
۱۴- در چند کلمه توضیح دهید که ترکیب چهار متغیر پیش‌بینی کننده GPA سال اول چقدر از بهترین پیش‌بینی کننده منفرد، بهتر است.

۲. آزمون ورودی دوره کارشناسی ارشد^۱ (GMAT) یک آزمون سه و نیم ساعته است که توسط تقریباً ۶۰۰۰ برنامه تجاری و مدیریتی در بیش از ۸۰ کشور به عنوان بخشی از درخواست پذیرش

¹ Graduate Management Admission Test

برای افرادی که می‌خواهند مدرک فارغ‌التحصیلی را به دست آورند، پذیرفته شده است. این آزمون توسط بیش از ۲۰۰،۰۰۰ متقدضی در هر سال انجام می‌شود. فرض کنید یک دانشگاه بزرگ که مدرک کارشناسی ارشد مدیریت منابع انسانی را ارائه می‌دهد، نمره GMAT را به عنوان بخشی از فرایند درخواست برای این دوره قرار داده است، دانشگاه می‌خواهد بداند که آیا نمره GMAT متقارضیان، میانگین نمرات (GPA) پایان سال اول دوره تحصیلات تکمیلی را پیش‌بینی می‌کند یا خیر.

GMAT دارای چهار نمره زیرمجموعه است: (۱) کلامی (با زه نمره ۶۰-۶۰)، (۲) کمی (با زه نمره ۶۰-۶۰)، (۳) نوشتار تحلیلی (با زه نمره ۶-۰ با فواصل ۵/۰) و (۴) استدلال یکپارچه (با زه نمره ۸-۸)، شما تصمیم گرفتید از این چهار نمره زیرمجموعه به عنوان پیش‌بینی کننده GPA سال اول استفاده کنید و برای بررسی مهارت‌های خود در اکسل، داده‌های فرضی را که در شکل ۱۱-۷ ارائه شده است، ایجاد کرده‌اید.



شکل ۱۱-۷: داده صفحه کاری برای فصل ۷: تمرين عملی شماره ۲

- (a) ایجاد یک صفحه گسترده اکسل با استفاده از GPA سال اول به عنوان متغیر معیار (Y)، و سایر متغیرها به عنوان چهار متغیر پیش‌بینی کننده این معیار $X_1 = \text{کلامی}$ ، $X_2 = \text{کمی}$ ، $X_3 = \text{نوشتار تحلیلی}$ ، $X_4 = \text{استدلال یکپارچه}$.
- (b) استفاده از رگرسیون چندگانه اکسل برای یافتن رابطه بین این پنج متغیر و قرار دادن خروجی خلاصه در زیر جدول.
- (c) استفاده از فرمت عدد (با ۲ رقم اعشار) برای همبستگی چندگانه در خروجی خلاصه، و استفاده از سه رقم اعشار برای ضرایب در خروجی خلاصه.
- (d) فایل را با عنوان GMAT26 ذخیره کنید.

(e) جدول و نتایج رگرسیون زیر جدول را چاپ کنید به طوری که آن‌ها در یک صفحه قرار بگیرند.
با استفاده از نسخه چاپی اکسل خود به سوالات زیر پاسخ دهید:

- ۱ همبستگی چندگانه R_{xy} چقدر است؟
- ۲ عرض از مبداء \mathbb{Y} یعنی a چقدر است؟
- ۳ ضریب متغیر کلامی، یعنی b_1 چقدر است؟
- ۴ ضریب متغیر کمی، یعنی b_2 چقدر است؟
- ۵ ضریب متغیر نوشتار تحلیلی، یعنی b_3 چقدر است؟
- ۶ ضریب متغیر استدلال یکپارچه، یعنی b_4 چقدر است؟
- ۷ معادله رگرسیون چندگانه را ارائه کنید؟
- ۸ GPA سال اول را پیش بینی کنید، انتظار دارید نمره متغیر کلامی ۵۲، نمره متغیر کمی ۴۸، نمره متغیر استدلال یکپارچه ۶ است.

(f) اکنون، به فایل اکسل خود بازگردید و یک ماتریس همبستگی برای این پنج متغیر ایجاد کنید، و آن را زیر خروجی خلاصه قرار دهید.

(g) این فایل را دوباره با عنوان GMAT26 ذخیره کنید.

(h) اکنون فقط این ماتریس همبستگی را بر روی یک صفحه جداگانه چاپ کنید. با استفاده از نسخه چاپی اکسل خود به سوالات زیر پاسخ دهید. (اطمینان حاصل کنید که علامت مشتبه یا منفی برای هر همبستگی لحاظ می‌شود):

- ۹ همبستگی بین متغیرهای کلامی و GPR سال اول چقدر است؟
- ۱۰ همبستگی بین متغیرهای کمی و GPA سال اول چقدر است؟
- ۱۱ همبستگی بین متغیرهای نوشتار تحلیلی و GPA سال اول چقدر است؟
- ۱۲ همبستگی بین متغیرهای استدلال یکپارچه و GPA سال اول چقدر است؟
- ۱۳ همبستگی بین متغیرهای کلامی و کمی چقدر است؟
- ۱۴ همبستگی بین متغیرهای کمی و نوشتار تحلیلی چقدر است؟
- ۱۵ همبستگی بین متغیرهای نوشتار تحلیلی و استدلال یکپارچه چقدر است؟
- ۱۶ همبستگی بین متغیرهای کمی و استدلال یکپارچه چقدر است؟
- ۱۷ بحث کنید که کدامیک از چهار پیش‌بینی‌کننده، بهترین پیش‌بینی‌کننده GPA سال اول می‌باشد.
- ۱۸ در چند کلمه توضیح دهید که چقدر ترکیب چهار متغیر پیش‌بینی‌کننده GPA سال اول از بهترین پیش‌بینی‌کننده منفرد، بهتر است.

۳. فرض کنید که شما مدیر بازاریابی فروشگاه‌های ۷یازده^۱ در میسوری هستید و می‌خواهید ببینید که آیا مکان پیشنهادی برای ایجاد فروشگاه جدید، با توجه به معیارهای مدنظر مناسب است یا خیر. داده‌های موجود برای یک نمونه تصادفی ۲۰ تایی از فروشگاه‌های ۷یازده در میسوری بر اساس اطلاعات سال گذشته در شکل ۷-۱۲ ارائه شده است.

Store ID	Y	X ₁	X ₂	X ₃
	Annual Sales (\$000)	Average Daily Traffic	Population (2-mile radius)	Average Income in Area
1	1,121	61,655	17,880	\$28,991
2	766	35,236	13,742	\$14,731
3	595	35,403	19,741	\$8,114
4	899	52,832	23,246	\$15,324
5	915	40,809	24,485	\$11,438
6	782	40,820	20,410	\$11,730
7	833	49,147	28,997	\$10,589
8	571	24,953	9,981	\$10,706
9	692	40,828	8,982	\$23,591
10	1,005	39,195	18,814	\$15,703
11	589	34,574	16,941	\$9,015
12	671	26,639	13,319	\$10,065
13	903	55,083	21,482	\$17,365
14	703	37,892	26,524	\$7,532
15	556	24,019	14,412	\$6,950
16	657	27,791	13,896	\$9,855
17	1,209	53,438	22,444	\$21,589
18	997	54,835	18,096	\$22,659
19	844	32,919	16,458	\$12,660
20	883	29,139	16,609	\$11,618

شکل ۷-۱۲: داده‌های صفحه کاری برای فصل ۷: تمرین عملی شماره ۳

- (a) ایجاد یک صفحه گسترده اکسل با استفاده از آمار فروش سالانه به عنوان متغیر معیار و میانگین روزانه ترافیک، جمعیت و متوسط درآمد به عنوان متغیرهای پیش‌بینی کننده.
- (b) استفاده از رگرسیون چندگانه اکسل برای یافتن ارتباط بین این چهار متغیر و قرار دادن خروجی خلاصه در زیر جدول.
- (c) استفاده از فرمت عدد (با ۲ رقم اعشار) برای همبستگی چندگانه در خروجی خلاصه، و استفاده از همین فرمت عدد برای ضرایب در خروجی خلاصه.
- (d) فایل را با عنوان multiple2 ذخیره کنید.
- (e) جدول و نتایج رگرسیون زیر جدول را طوری چاپ کنید که آن‌ها در یک صفحه قرار بگیرند.
با استفاده از نسخه چاپی اکسل خود به سوالات زیر پاسخ دهید:

۱- همبستگی چندگانه R_{xy} چقدر است؟

¹ 7Eleven

- ۲- مقدار عرض از مبدأ Y یعنی a چقدر است؟
- ۳- ضریب میانگین ترافیک روزانه یعنی b_1 چقدر است؟
- ۴- ضریب جمعیت یعنی b_2 چقدر است؟
- ۵- ضریب میانگین درآمد یعنی b_3 چقدر است؟
- ۶- معادله رگرسیون چندگانه را ارائه کنید؟
- ۷- فروش سالانه را پیش بینی کنید، انتظار دارید مقدار میانگین ترافیک روزانه ۴۲۰۰۰، جمعیت ۲۳۰۰۰ و درآمد ۲۲۰۰۰ دلار باشد.

(f) اکنون، به فایل اکسل خود بازگردید و یک ماتریس همبستگی برای این چهار متغیر ایجاد کنید و آن را در زیر خروجی خلاصه در صفحه گسترده خود قرار دهید.

(g) این فایل را با عنوان `multiple3` ذخیره کنید.

(h) اکنون، فقط این ماتریس همبستگی را بر روی یک صفحه جداگانه چاپ کنید. با استفاده از نسخه چاپی اکسل خود به سوالات زیر پاسخ دهید. اطمینان حاصل کنید که علامت مثبت یا منفی برای هر همبستگی لحاظ شده است:

- ۸- همبستگی بین ترافیک و فروش چقدر است؟
- ۹- همبستگی بین جمعیت و فروش چقدر است؟
- ۱۰- همبستگی بین درآمد و فروش چقدر است؟
- ۱۱- همبستگی بین ترافیک و جمعیت چقدر است؟
- ۱۲- همبستگی بین جمعیت و درآمد چقدر است؟
- ۱۳- بحث کنید که کدامیک از سه پیش‌بینی کننده فروش سالانه است.
- ۱۴- در چند کلمه توضیح دهید که ترکیب سه متغیر پیش‌بینی کننده فروش سالانه چقدر از بهترین پیش‌بینی کننده منفرد، بهتر است.

منابع

- Keller G. Statistics for management and economics. 8th ed. Mason: South-Western Cengage Learning; 2009.
- Levine DM, Stephan DF, Krehbiel TC, Berenson ML. Statistics for managers using Microsoft excel. 6th ed. Boston: Prentice Hall/Pearson; 2011.

فصل هشتم

تجزیه واریانس یک طرفه (ANOVA)

تاکنون در این راهنمای اکسل ۲۰۱۳، شما چگونگی استفاده از آزمون t گروهی برای مقایسه میانگین نمونه با میانگین جامعه و آزمون t دو گروهی برای مقایسه تفاوت میانگین‌های دو نمونه را آموخته‌اید. اما وقتی بیش از دو گروه دارید و می‌خواهید اختلاف معنی‌دار بین میانگین‌گروه‌ها را تعیین کنید، چه می‌کنید؟

جواب این پرسش: تجزیه واریانس^۱ (ANOVA).

آزمون ANOVA به شما امکان می‌دهد تا تفاوتات بین میانگین‌ها را زمانی که ۳ یا بیش از ۳ گروه در مطالعه خود دارید مورد بررسی قرار دهید.

نکته مهم: به منظور انجام تجزیه واریانس یک طرفه به نصب بسته‌ابزار تجزیه و تحلیل داده^۲ دارید که در فصل ۶ توضیح داده شده است، (به بخش ۱-۵-۶ مراجعه شود). اگر این مورد را نصب نکرده‌اید اکنون لازم است این کار را انجام دهید.

فرض کنید علاقه‌مند هستید قیمت‌های سه فروشگاه زنجیره‌ای در سنت لویس به نام‌های (۱) دائربرگ^۳ (فروشگاه اول)، (۲) اشناک^۴ (فروشگاه دوم) و (۳) شاپ ان سیو^۵ (فروشگاه سوم) را مورد مقایسه قرار دهید. همچنین، فرض کنید که شما ۲۸ قلم جنس فهرست شده در جدول زیر را به عنوان "سبد بازاری محصولات" برای مقایسه قیمت‌ها در این سه فروشگاه انتخاب کرده‌اید. اندازه بسته هر کدام از اجنباس در فهرست را مشخص کرده‌اید. برای مثال، جنس شماره ۱۴: مایع شوینده با وزن ۱۶ اونس است.

فرض کنید که جغرافیایی ۶۳۱۱۹ در سنت لویس را انتخاب کرده‌اید این کد یک شعبه از هر سه فروشگاه زنجیره‌ای را دارد. شما به سه شعبه فروشگاه زنجیره‌ای در این منطقه جغرافیایی می‌روید و اطلاعات فرضی در شکل ۱-۸ را که قیمت‌های اقلام سبد بازاری محصولات است، را به دست می‌آورید.

¹ Analyse of Variance

² Data Analysis Toolpak

³ Dierberg

⁴ Schnuck

⁵ Shop'n Save

ITEM	DIERBERG'S	SCHNUCK'S	SHOP 'n SAVE
1	1.85	1.45	1.25
2	3.95	3.35	3.04
3	2.25	1.75	1.45
4	2.85	2.35	2.25
5	1.65	1.10	0.85
6	3.65	2.95	2.45
7	2.45	1.85	1.45
8	1.95	1.56	1.44
9	1.83	1.25	1.15
10	2.64	2.14	2.04
11	2.84	2.25	2.15
12	1.84	1.20	0.55
13	1.65	1.25	1.15
14	2.75	2.10	2.04
15	2.71	1.86	1.75
16	1.55	0.94	0.85
17	1.85	1.30	1.01
18	0.95	0.55	0.45
19	1.55	1.28	1.06
20	1.44	0.85	0.74
21	1.65	1.25	1.15
22	1.64	1.28	1.04
23	4.21	3.75	3.36
24	1.20	0.71	0.61
25	4.55	3.90	3.25
26	3.45	2.84	2.65
27	5.85	5.30	5.14
28	1.65	1.25	1.04

شکل ۸-۱: داده‌های صفحه کاری برای مثال مقایسه قیمت فروشگاه‌ها (مثال عملی)

برای این اطلاعات یک صفحه گسترده اکسل به صورت زیر ایجاد کنید:

B1: مقایسه قیمت فروشگاهها

A3: جنس (قلم)

B3: دائربرگ (فروشگاه اول)

C3: اشناك (فروشگاه دوم)

D3: شاپ آن سیو (فروشگاه سوم)

1 : A4

1/۸۵ : B4

بقيه اطلاعات را وارد جدول صفحه گسترده (اکسل) کنيد. وقتی ورود اطلاعات را تمام کردید، آخرین سلول سمت چپ باید مقدار ۲۸ در سلول A31 را دارا باشد، و آخرین سلول سمت راست باید ۱/۰۴ را در سلول D31 داشته باشد. اعداد هر کدام از ستون‌ها را به مرکز سلول‌ها انتقال دهيد. از فرمت اعداد دارای دو رقم اعشار استفاده کنيد.

نکته مهم: همه ارقام جدول را دوباره کنترل کنيد تا مطمئن شويد تمامی آن‌ها کاملاً صحیح می‌باشند در غیر این صورت پاسخ صحیحی برای این مسئله به دست نخواهد آورد! این فایل را به نام 5 Supermarket ذخیره کنيد.

۱-۸- استفاده از اکسل به منظور انجام تجزیه واریانس یک‌طرفه (ANOVA)

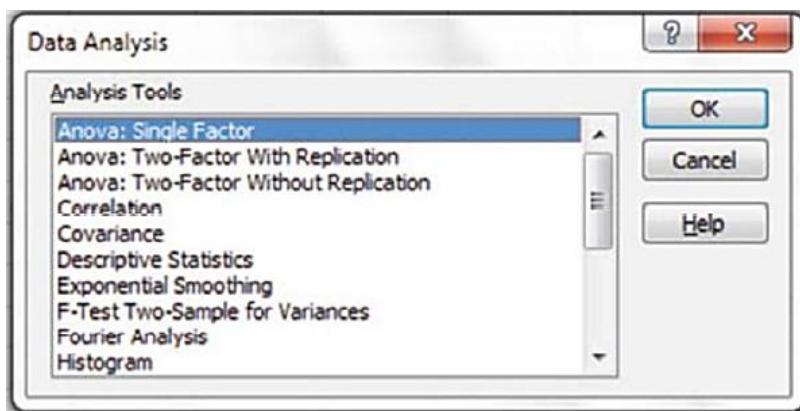
هدف: استفاده از اکسل به منظور انجام آزمون تجزیه واریانس یک‌طرفه (ANOVA)

اکنون با استفاده از مراحل زیر شما آمده انجام آزمون ANOVA بر روی این داده‌ها هستيد:

انتخاب سر برگ Data در نوار ابزار اکسل

انتخاب گزینه Data Analysis در انتهای سمت راست این قسمت

انتخاب گزینه ANOVA:Single Factor (نوار انتخاب را به سمت اين رابطه هدایت کنيد و آن را انتخاب نمایيد، به شکل ۲-۸ مراجعه شود)



شکل ۸-۸: جعبه محاوره‌ای برای گزینه "Data Analysis" انتخاب ANOVA یکطرفه

گزینه OK را انتخاب کنید.

بازه ورودی داده‌ها: B3:D31 (توجه کنید که شما در این بازه عنوان‌های ستون را که در ردیف ۳ قرار دارند، وارد کرده‌اید)

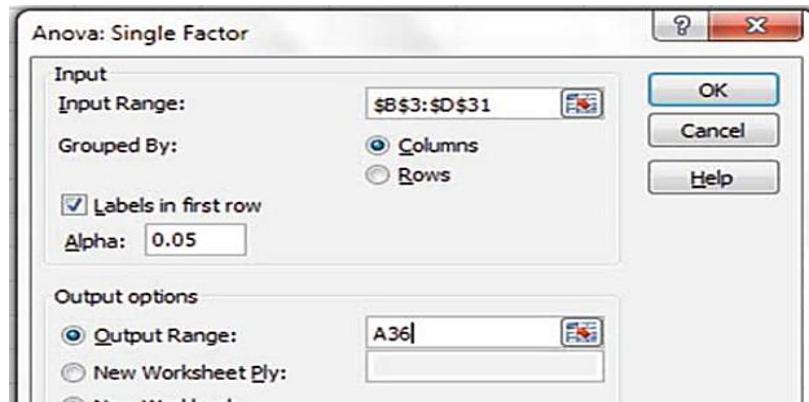
نکته مهم: هرگاه سری داده‌ها دارای اندازه نمونه متفاوت در گروه‌های مورد مقایسه است، بازه ورودی که شما تعریف می‌کنید باید از عنوان ستون گروه اول در سمت چپ آغاز شده و تا آخرین ستون در سمت راست ادامه یابد سپس به سمت پایین ترین ردیف در ماتریس داده که دارای عدد است حرکت کرده به نحوی که دامنه ورودی شکل مستطیل را به هنگام انتخاب آن دارا باشد.

گروه‌بندی شده توسط: ستون‌ها

یک علامت انتخاب بگذارید: عنوان‌ها در ردیف اول

بازه خروجی (بر روی دکمه سمت چپ آن کلیک کنید): سلول A36 را انتخاب کنید یا تایپ نمایید

(به شکل ۳-۸ مراجعه شود)



شکل ۸-۳: جعبه محاوره‌ای برای ANOVA: Single Factor بازه ورودی/خروجی
گزینه OK را انتخاب کنید.
این فایل را با نام SUPER6 ذخیره کنید.

برای ایجاد جدول در شکل ۸-۴ همه ارقام را که به صورت اعشار هستند به دو رقم گرد کنید و همه اعداد سلول‌ها را در مرکز آن‌ها قرار دهید، در این صورت، خواندن جدول، آسان‌تر می‌شود.

تجزیه واریانس یک‌طرفه					
خلاصه	B	C	D	E	F
35					
36 Anova: Single Factor	تعداد	جمع	میانگین	واریانس	
37					
38 SUMMARY	Count	Sum	Average	Variance	
39 Groups	28	68.40	2.44	1.32	
40 DIERBERG'S	28	53.61	1.91	1.22	
41 SCHNUCK'S	28	47.36	1.69	1.13	
42 SHOP 'n SAVE					
43	مجموع مربعات	درجه آزادی	میانگین مربعات	آماره F محاسباتی	آماره F بحرانی
44					
45 ANOVA	SS	df	MS	F	P-value
46 Source of Variation					
47 Between Groups	8.34	2	4.17	3.40	0.04
48 Within Groups	99.23	81	1.23		3.11
49					
50 Total	107.57	83			
51					

تجزیه واریانس
گروه‌ها
مجموع مربعات
درجه آزادی
میانگین مربعات
آماره F محاسباتی
آماره F بحرانی
سطح معنی‌داری
منبع تغییرات
بین گروه‌ها
داخل گروه‌ها
کل

شکل ۸-۴: نتایج ANOVA برای مقایسه قیمت در فروشگاه‌ها

جدول داده‌ها و جدول خلاصه ANOVA را چاپ کنید به‌گونه‌ای که همه اطلاعات داخل یک صفحه باشند (تذکر: گزینه Layout/Fit صفحه را متناسب با اندازه ۸۵٪ تنظیم کنید). کنترل خروجی تجزیه، باید مقادیر زیر در سلول‌ها وجود داشته باشد:

ANOVA: Single Factor A36

۲/۴۴ :D40

۴/۱۷ :D47

۳/۴۰ :E47

۳/۱۱ :G47

اکنون اجازه دهید در خصوص تفسیر جدول بحث کنیم:

۲-۸- چگونه جدول ANOVA را به درستی تفسیر کنیم

هدف: تفسیر جدول ANOVA به صورت درست

ANOVA به شما امکان می‌دهد تا اختلاف بین میانگین‌ها را وقتی سه یا بیش از ۳ گروه داده وجود دارند، مورد آزمون قرار دهید. این آزمون ANOVA آماره آزمون F نامیده می‌شود و با حرف F نشان داده می‌شود.

رابطه آزمون F به صورت زیر است:

$F = \frac{\text{میانگین مربعات بین گروه‌ها}}{\text{میانگین مربعات داخل گروه‌ها}} = \frac{MS_b}{MS_w}$

$$F = MS_b / MS_w \quad (1-8)$$

طرز استخراج و توضیح این رابطه خارج از محدوده این کتاب است. در این راهنمای سعی داریم به شما آموزش دهیم که چگونه از اکسل استفاده کنید، و قصد نداریم، تئوری آماری که در پشت رابطه‌های ANOVA قرار دارند را آموزش دهیم، برای توضیح بیشتر ANOVA به ویرز (۲۰۱۱) مراجعه کنید.

توجه کنید که سلول D47 محتوی $MS_b = ۴/۱۷$ و سلول D48 محتوی $MS_w = ۱/۲۳$ است. وقتی شما این دو عدد را با استفاده از منابع سلول‌ها در اکسل بر هم تقسیم می‌کنید، جواب آزمون F را معادل $۳/۴۰$ به دست می‌آورید که در سلول E47 است. اکنون معنی عدد $۳/۴۰$ را مورد بحث قرار دهیم.

برای تعیین اینکه مقدار F برابر $۳/۴۰$ ، اختلاف معنی‌داری بین میانگین ۳ گروه را نشان می‌دهد یا خیر، اولین گام نوشتن فرضیه صفر و فرضیه تحقیق برای سه گروه از قیمت‌ها است.

در مقایسه قیمت بین فروشگاه‌ها، فرضیه صفر بیان می‌کند که میانگین جامعه برای سه گروه مساوی هستند درحالی‌که فرضیه تحقیق بیان می‌کند که میانگین جامعه در سه گروه مساوی نیست و اختلاف معنی‌داری آماری بین میانگین جامعه در سه گروه وجود دارد. کدامیک از این دو فرضیه بر اساس نتایج ANOVA باید پذیرفته شوند؟

۳-۸- استفاده از قاعده تصمیم برای آزمون F در ANOVA

برای بیان فرضیه‌ها اجازه دهد اثباتگر را گروه ۱، اشناس را گروه ۲ و شاپ آنسیو را گروه ۳ نام‌گذاری کنیم.

$$H_0: \mu_1 = \mu_2 = \mu_3$$

$$H_1: \mu_1 \neq \mu_2 \neq \mu_3$$

جواب این پرسش قابل مقایسه با قاعده تصمیم مورداستفاده در این کتاب برای دو مورد آزمون t یک گروهی و t دو گروهی، است. شما به خاطر دارید که این قاعده به صورت زیر است (به بخش‌های ۴-۶ و ۸-۱-۵ مراجعه شود):

اگر مقدار قدر مطلق t (محاسباتی) کمتر از t بحرانی باشد، فرضیه صفر را پذیرفته می‌شود. اگر مقدار قدر مطلق t (محاسباتی) بیشتر از t بحرانی باشد، فرضیه صفر را رد کرده و فرضیه تحقیق پذیرفته می‌شود.

اکنون قاعده تصمیم برای ANOVA به صورت زیر است:

هدف: یادگیری قاعده تصمیم برای آزمون F در ANOVA

از این‌رو، قاعده تصمیم برای آزمون F در ANOVA برای مثال فروشگاه‌ها به شرح زیر است:

اگر مقدار F (محاسباتی) کمتر از F بحرانی باشد، فرضیه صفر را می‌پذیرید و اگر مقدار F (محاسباتی) بزرگتر از F بحرانی باشد، فرضیه صفر را رد کرده و فرضیه تحقیق را می‌پذیرید.

توجه داشته باشید که مقدار F بحرانی را در سلول G47:3.11 اکسل ارائه شده است.

از آنجاکه مقدار F برابر با $\frac{3}{40}$ بزرگتر از مقدار F بحرانی معادل $\frac{3}{11}$ است، فرضیه صفر را رد می‌کنیم و فرضیه تحقیق را می‌پذیریم.

بنابراین، نتیجه‌گیری (جمع‌بندی) ما به طور ساده بشرح زیر است:

بین میانگین قیمت‌های سه فروشگاه، اختلاف معنی‌داری وجود دارد.

توجه کنید که لازم نیست مقدار قدر مطلق F برابر $\frac{3}{40}$ را به دست آورید. مقدار F هرگز کمتر از یک نیست، بنابراین، این آماره هرگز نمی‌تواند مقدار منفی داشته باشد و نیازی نیست مقدار قدر مطلق آن لحاظ شود تا به عنوان مقداری مثبت مدنظر قرار گیرد.

لازم است توجه کنیم که ANOVA به ما می‌گوید بین میانگین‌های جامعه سه گروه اختلاف معنی‌داری وجود دارد اما تعیین نمی‌کند که کدام جفت از گروه‌ها اختلاف معنی‌داری با یکدیگر داشتند.

۴-۸- آزمون اختلاف بین دو گروه با آزمون t در ANOVA

برای پاسخ به این پرسش نیاز به انجام یک آزمون متفاوت به نام آزمون t در ANOVA داریم.

هدف: آزمون اختلاف میانگین بین دو گروه با استفاده از آزمون t در ANOVA زمانی که نتایج ANOVA اختلاف معنی‌داری بین میانگین‌های جامعه را نشان می‌دهد.

از آنجاکه ما سه گروه داده داریم (یک گروه برای هر یک از ۳ فروشگاه) مجبوریم سه آزمون t در ANOVA مجزا برای تعیین جفت گروه‌های با اختلاف معنی‌دار از یکدیگر، انجام دهیم.

آزمون t در ANOVA مجزا باید برای میانگین جفت گروه‌های زیر صورت گیرد:

(۱) دائربرگ در مقابل اشناك

(۲) دائربرگ در مقابل شاپ^۱ ان سیو^۱

(۳) اشناك در مقابل شاپ^۱ ان سیو

فقط یکی از سه جفت آزمون را انجام می‌دهیم، دائربرگ در مقابل شاپ^۱ ان سیو تا مسیر انجام آزمون t در ANOVA را برای مقایسه دو فروشگاه شرح دهیم. آزمون t در ANOVA برای سایر جفت گروه‌ها نیز به طریق مشابه صورت می‌گیرد.

۴-۱- مقایسه قیمت دائربرگ در مقابل شاپ^۱ ان سیو با آزمون t در ANOVA

هدف: مقایسه قیمت دائربرگ در مقابل شاپ^۱ ان سیو برای ۲۸ قلم کالا در سبد خرید با استفاده از آزمون t در ANOVA

مرحله اول نوشتن فرضیه صفر و فرضیه تحقیق برای این دو فروشگاه است.

برای آزمون t در ANOVA فرضیه صفر بیان می‌کند که میانگین جامعه در دو گروه مساوی هستند، در حالی که فرضیه تحقیق بیان می‌کند که میانگین جامعه در دو گروه مساوی نیستند (یعنی اختلاف معنی‌داری بین میانگین جامعه در دو گروه وجود دارد). چون ما دائربرگ (گروه ۱) را در مقابل شاپ^۱ ان سیو (گروه ۳) مقایسه می‌کنیم، این فرضیه‌ها به قرار زیر است:

$$H_0: \mu_1 = \mu_3$$

$$H_1: \mu_1 \neq \mu_3$$

¹ Shop 'n Save

برای گروه ۱ در مقابل گروه ۳ رابطه آزمون t در ANOVA بشرح زیر است:

$$ANOVA_t = \frac{\overline{X}_1 - \overline{X}_3}{S.e_{ANOVA}} \quad (2-8)$$

که در آن

$$S.e_{ANOVA} = \sqrt{MSw \left(\frac{1}{n_1} + \frac{1}{n_3} \right)} \quad (3-8)$$

مراحل مدنظر در محاسبه آزمون t در ANOVA به شرح زیر است:

۱. پیدا کردن اختلاف بین میانگین نمونه در دو گروه $(0/75 - 0/69 = 0/06)$.

۲. پیدا کردن $1/n_1 + 1/n_3$ (در هر دو گروه ۲۸ قلم کالای فروشگاهی وجود دارد که در نتیجه خواهیم داشت: $1/28 + 1/28 = 0/028 + 0/028 = 0/056$).

۳. ضرب پاسخ حاصل از MS_w در پاسخ حاصل از مرحله ۲ $(0/0878 \times 0/0714 = 0/00878)$.

۴. گرفتن جذر از پاسخ حاصل از مرحله ۳ $(\sqrt{0/00878} = 0/0937)$.

۵. تقسیم مقدار حاصل از مرحله ۱ به مرحله ۴ برای یافتن t در ANOVA $(0/030 / 0/0937 = 0/32)$.

توجه: چون اکسل همه محاسبات را تا ۱۶ رقم اعشار انجام می‌دهد، زمانی که برای محاسبه بالا از اکسل استفاده می‌کنید، به جای جواب $0/32$ برای محاسبه، ماشین حساب جواب معادل $2/54$ را ارائه می‌کند.

نتیجه $0/32$ برای آزمون t در ANOVA چه تفسیری دارد؟ به منظور تفسیر ارزش $2/5$ به طور صحیح، لازم است ارزش بحرانی t برای آزمون t در ANOVA را تعیین کیم. برای انجام آن لازم است درجه آزادی آزمون t در ANOVA را بشرح زیر به دست آوریم:

۸-۱-۴-۱- تعیین درجه آزادی آزمون t در ANOVA

هدف: تعیین درجه آزادی آزمون t در ANOVA

درجه آزادی (df) آزمون t در ANOVA بشرح زیر است:

$$(k = \text{تعداد گروه‌ها}) \quad df = n_{\text{TOTAL}} - k$$

تعداد گروه‌ها در مطالعه شما- اندازه کل نمونه از همه گروه‌ها = df

در مثال ما، اندازه کل نمونه ۳ گروه برابر ۸۴ است، چون ۲۸ قیمت برای هر ۳ فروشگاه وجود دارد و از آنجا که ۳ گروه وجود دارد، درجه آزادی آزمون t در ANOVA $84 - 3 = 81$ برابر ۸۱ می‌شود.

اگر شما $df=81$ را در ستون درجه آزادی (df) در جدول t در ضمیمه E نگاه کنید که ستون دوم سمت چپ از این جدول است، مقدار t بحرانی را معادل $1/96$ پیدا می‌کنید.
نکته مهم: مطمئن باشید که از ستون درجه آزادی (df) در ضمیمه E برای مقدار t بحرانی آزمون t در ANOVA/استفاده می‌کنید.

۴-۱-۲-۲- بیان توضیح قاعده تصمیم برای آزمون t در ANOVA

هدف: یادگیری قاعده تصمیم برای آزمون t در ANOVA

تفسیر نتایج آزمون t در ANOVA از همان قاعده تصمیم استفاده شده برای آزمون t یک گروهی و آزمون t دو گروهی تبعیت می‌کند. (به بخش ۱-۴ و ۱-۵ مراجعه شود):
اگر مقدار قدر مطلق t (محاسباتی) کمتر از t بحرانی باشد، فرضیه صفر را می‌پذیریم.
یا اگر مقدار قدر مطلق t (محاسباتی) بزرگتر از t بحرانی باشد، فرضیه صفر را رد می‌کنیم و فرضیه تحقیق را می‌پذیریم.
از آنجاکه از آزمون t یک طرفه استفاده می‌کنیم، لازم است مقدار قدر مطلق t را به دست آوریم.
چون مقدار قدر مطلق آماره محاسباتی یعنی $2/50$ بزرگتر از t بحرانی که معادل $1/96$ است، فرضیه صفر (میانگین جامعه در دو گروه مساوی هستند) را رد می‌کنیم و فرضیه تحقیق (میانگین جامعه در دو گروه اختلاف معنی‌داری از یکدیگر دارند) را می‌پذیریم.
این بدان معناست که نتیجه‌گیری ما به زبان ساده بشرح زیر است:
میانگین قیمت‌های سبد خرید اجنباس در فروشگاه دائبرگ به‌طور معنی‌داری بیشتر از میانگین قیمت همین اقلام در فروشگاه شاپ آن‌سیو است ($2/44$ دلار در مقابل $1/44$ دلار).
توجه کنید که اختلاف میانگین قیمت‌ها یعنی $0/75$ دلار به نظر زیاد نیست اما در اصطلاح علمی این بدان معناست که میانگین قیمت‌ها در دائبرگ 44% بیشتر از میانگین قیمت‌ها در شاپ آن‌سیو می‌باشد. واضح است که اختلاف در قیمت‌های این دو فروشگاه بر اساس داده‌های فرضی ما حاصل شده است.

۴-۱-۳- انجام آزمون t در ANOVA با استفاده از دستورهای اکسل

اجازه دهید محاسبات آزمون t در ANOVA را با استفاده از اکسل با فایل SUPER6 که قبلاً در این فصل ایجاد کرده‌اید، انجام دهیم.
D52: دائبرگ در مقابل شاپ آن‌سیو
 $1/n$: A54
 $1/n$: A54 + شاپ آن‌سیو

A56: خطای معیار (S.e.) دائربرگ در مقابل خطای معیار (S.e.) شاپ^{ان سیو}

A58: ANOVA t-test

D54: =(1/28+1/28)

D56: =SQRT(D48*D54)

D58: =(D40-D42)/D56

اکنون، زمانی که همه رقمنا را در آزمون t در ANOVA تا دو رقم اعشار گرد می‌کنید، باید نتایج زیر را در این سلول‌ها داشته باشید:

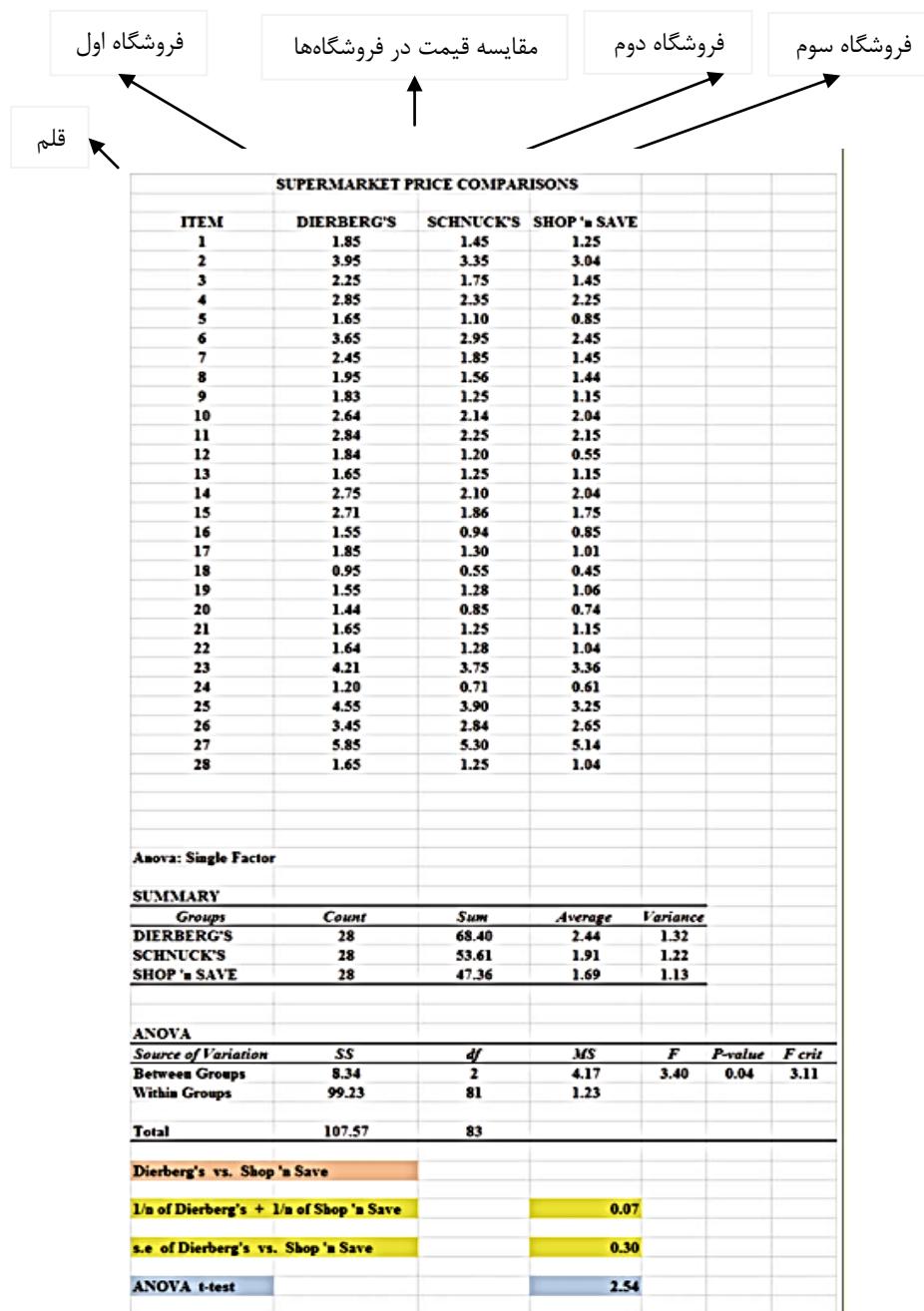
۰/۰۷ : D54

۰/۳۰ : D56

۲/۵۴ : D58

نتیجه نهایی را در قالب فایل با نام SUPER7 ذخیره کنید.

نتایج صفحه گسترده را در یک صفحه که مانند شکل ۵-۸ تنظیم شده است، چاپ کنید (تذکر: کاهش مقیاس از طریق Layout /Scale تا مناسب شدن با ۷۵٪ اندازه واقعی صورت می‌گیرد).



شكل ۸-۵: صفحه گسترده نهایی مقایسه قیمت فروشگاه دیربرگ در مقابل شاپ'ان سیو

برای توضیح بیشتر آزمون t در ANOVA به بلاک (۲۰۱۰) مراجعه شود.

نکته مهم: شما زمانی اجازه دارید آزمون t در ANOVA را برای مقایسه میانگین دو گروه بکار ببرید که آزمون F اختلاف معنی‌داری بین میانگین جامعه در همه گروه‌های مورد مطالعه را نشان می‌دهد.

زمانی که مقدار F (محاسباتی) کمتر از مقدار F بحرانی است انجام آزمون t در ANOVA مناسب نمی‌باشد. وقتی مقدار F (محاسباتی) کمتر از مقدار F بحرانی است به این معنی است که هیچ اختلافی بین میانگین جامعه در گروه‌ها وجود ندارد و بنابراین، شما نمی‌توانید آزمون کنید که آیا اختلاف بین میانگین در هر دو گروه وجود دارد یا خیر، زیرا اختلاف بین دو گروه ناشی از شанс است.

۸-۵-۸- تمرین‌های عملی پایان فصل

۱. فرض کنید می‌خواهید مزایای برنده لاستیک تولیدی شرکت خود را (برند A) را در مقابل دو برنده رقیب (B و C)، مقایسه کنید. شما آزمون آزمایشگاهی برای سه نوع لاستیک اجرا می‌کنید و تعداد مایل‌های طی شده قبل از اینکه قطر لاستیک به مقدار از قبیل تعیین شده برسد، اندازه‌گیری می‌کنید.

آزمون مقاومت لاستیک			داده‌ها بر حسب هزار مایل است
TIRE MILEAGE TEST			
(Data are in thousands of miles)			
Brand A	Brand B	Brand C	
62	61	65	
61	62	67	
62	63	71	
64	60	66	
61	64	65	
	59	64	
	62		
	63		
	62		
	63		

شکل ۸-۶: داده‌های صفحه کاری برای فصل ۸: تمرین عملی شماره ۱

نتایج فرضی در شکل ۸-۶ ارائه شده است. توجه کنید که داده‌ها بر حسب هزار مایل طی شده، است.

بنابراین، برای مثال، ۶۳۰۰۰ معنی دار مایل طی شده است.

(الف) وارد کردن داده‌ها در صفحه گسترده اکسل

(ب) آزمون ANOVA یک طرفه بر روی این داده‌ها انجام دهید و نتایج جدول ANOVA را در زیر داده‌های ورودی برای ۳ برنده لاستیک نشان دهید.

(ج) اگر مقدار آماره F در ANOVA معنی دار است یک رابطه در اکسل به منظور محاسبه آزمون t در ANOVA برای مقایسه میانگین مایل طی شده برنده A در برابر برنده C ایجاد کرده و نتایج را در زیر جدول ANOVA در صفحه گسترده نشان دهید (خطای معیار و مقدار آزمون t در ANOVA را در خطاهای جداگانه صفحه گسترده قرار دهید و برای هر مقدار تا دو رقم اعشار استفاده کنید).

(د) نتایج صفحه گسترده را چاپ کنید به گونه‌ای که همه اطلاعات در یک صفحه قرار گیرند.

(ه) صفحه گسترده را با نام TIRE7 ذخیره کنید.

اکنون جواب پرسش‌های زیر را با استفاده از خروجی اکسل بدهید:

۱. فرضیه صفر و فرضیه تحقیق برای آزمون F در ANOVA چیست؟
۲. در خروجی اکسل MS_b چقدر است؟
۳. در خروجی اکسل MS_w چقدر است؟
۴. با استفاده از ماشین حساب $F = MS_b / MS_w$ را محاسبه کنید.
۵. مقدار بحرانی F در خروجی اکسل چقدر است؟
۶. نتایج آزمون F در ANOVA چیست؟
۷. نتیجه‌گیری آزمون F در ANOVA به زبان ساده چیست؟
۸. اگر آزمون F در اختلاف معنی داری بین ۳ برنده بر اساس مایل‌های طی شده نشان دهد، فرضیه صفر و فرضیه تحقیق برای آزمون F در ANOVA به منظور مقایسه برنده A در برابر برنده C چیست؟
۹. میانگین برنده A در خروجی اکسل چقدر است؟
۱۰. میانگین برنده C در خروجی اکسل چقدر است؟
۱۱. درجه آزادی (df) برای آزمون t در ANOVA برای مقایسه برنده A در برابر برنده C چقدر است؟
۱۲. برای این درجه آزادی مقدار t بحرانی برای آزمون t در ANOVA در ضمیمه E چقدر است؟
۱۳. با استفاده از ماشین حساب s.e.ANOVA (خطای معیار ANOVA) را محاسبه کنید.

۱۴. با استفاده از ماشین حساب مقدار آزمون t در ANOVA برای مقایسه برنده A در برابر برنده C را محاسبه کنید.

۱۵. نتایج آزمون t در ANOVA برای مقایسه برنده A در برابر برنده C چیست؟

۱۶. نتیجه‌گیری آزمون t در ANOVA برای مقایسه برنده A در برابر برنده C به زبان ساده چیست؟

توجه کنید که ۳ برنده لاستیک وجود دارد و شما لازم است سه آزمون t در ANOVA برای تعیین اختلاف معنی‌دار بین لاستیک‌ها انجام دهید. از آنجاکه آزمون t در ANOVA برای مقایسه برنده A در برابر برنده C را انجام داده‌اید باید آزمون t در ANOVA بعدی برای مقایسه برنده A در برابر برنده B را انجام دهید.

۱۷. فرضیه صفر و فرضیه تحقیق برای مقایسه برنده A در برابر برنده B را بیان کنید.

۱۸. میانگین برنده A در خروجی اکسل چیست؟

۱۹. میانگین برنده B در خروجی اکسل چیست؟

۲۰. درجه آزادی (df) برای آزمون t در ANOVA به منظور مقایسه برنده A در برابر برنده B چیست؟

۲۱. برای این درجه آزادی (df) مقدار t بحرانی آزمون t در ANOVA در ضمیمه E چیست؟

۲۲. با استفاده از ماشین حساب s.e.ANOVA (خطای معیار ANOVA) را برای مقایسه برنده A در برابر برنده B محاسبه کنید.

۲۳. مقدار آزمون t در ANOVA برای مقایسه برنده A در برابر برنده B را محاسبه کنید.

۲۴. نتیجه آزمون t در ANOVA برای مقایسه برنده A در برابر برنده B چیست؟

۲۵. نتیجه‌گیری آزمون t در ANOVA برای مقایسه برنده A در برابر برنده B به زبان ساده چیست؟ آخرین آزمون t در ANOVA مقایسه برنده B در برابر برنده C است. این آزمون بشرح زیر انجام می‌شود:

۲۶. فرضیه صفر و فرضیه تحقیق برای مقایسه برنده B در برابر برنده C را بیان کنید.

۲۷. میانگین برنده B در خروجی اکسل چیست؟

۲۸. میانگین برنده C در خروجی اکسل چیست؟

۲۹. درجه آزادی (df) برای آزمون t در ANOVA به منظور مقایسه برنده B در برابر برنده C چیست؟

۳۰. برای این درجه آزادی مقدار t بحرانی برای آزمون t در ANOVA در ضمیمه E چیست؟

۳۱. با استفاده از ماشین حساب s.e.ANOVA (خطای معیار ANOVA) برای مقایسه برندهای A، B و C در برابر برندهای آزمون t در ANOVA برای مقایسه برندهای A، B و C محاسبه کنید.
۳۲. با استفاده از ماشین حساب مقدار آزمون t در ANOVA برای مقایسه برندهای A، B و C در برابر برندهای آزمون t محاسبه کنید.
۳۳. نتیجه آزمون t در ANOVA برای مقایسه برندهای A، B و C در برابر برندهای آزمون t چیست؟
۳۴. نتیجه گیری آزمون t در ANOVA برای مقایسه برندهای A، B و C به زبان ساده و روان چیست؟
۳۵. خلاصه نتایج سه آزمون t در ANOVA به زبان ساده و روان چیست؟
۳۶. بر اساس نتایج تجزیه و تحلیل خود چه پیشنهادی برای شرکت خود در مورد این ۳ برندهای A، B و C دارید؟ علت این پیشنهاد چیست؟

۲. مک دونالد^۱ "برگرهای سه بار کوبیده ۱۰۰٪ از گوشت گاو" را برای رقابت با برگرهای خیلی بزرگ فروخته شده توسط هاردی^۲ به بازار معرفی کرد. فرض کنید شما به عنوان مشاور مک دونالد برای تجزیه و تحلیل داده‌های چهار آزمون بازار شهری، براساس اندازه جمعیت، متوسط درآمد خانوار، متوسط تعداد منابع درآمدی خانوار و تعداد رستوران‌های مک دونالد در هر شهر، استخدام شده‌اید. فرض کنید آزمون بازار ۱۲ هفته طول می‌کشد و هر شهر فقط از یک نوع تبلیغات برای این برگرهای استفاده می‌کند:

(۱) رادیو (۲) تلویزیون محلی (۳) بیلیوردها و (۴) روزنامه محلی. شهرها به طور تصادفی برای یک نوع تبلیغ انتخاب شدند و در هر شهر میزان مشابهی مخارج تبلیغاتی به صورت هفتگی برای نوع تعیین شده تبلیغات صورت گرفت. اطلاعات فرضی برای تعداد واحدهای فروخته شده از برگر آنگوس^۳ در

شکل ۷-۸ ارائه شده است.

¹ McDonald

² Hardee

³ Angus

آزمون مطالعه بازار برگر آنگوس

رادیو	تلوزیون محلی	بیلبورد	روزنامه محلی
ANGUS BURGER TEST MARKET STUDY			
1	2	3	4
Radio	Local TV	Billboards	Local newspaper
300	310	340	280
320	315	330	285
310	320	345	290
290	326	342	275
280	324	341	282
315	318	351	284
326	330	339	291
295	327	337	284
278	328	329	279
289	319	328	274
287	326	332	283
305	328	335	285

شکل ۷-۸: داده‌های صفحه کاری برای فصل ۸: تمرین عملی شماره ۲

- (a) وارد کردن داده‌ها در صفحه گسترده اکسل
 - (b) آزمون ANOVA یک طرفه را روی این داده‌ها انجام دهید و نتایج جدول ANOVA را زیر داده‌های ورودی برای این ۴ نوع تبلیغ ارائه کنید.
 - (c) اگر مقدار F جدول ANOVA معنی‌دار است یک رابطه اکسل برای محاسبه آزمون t ANOVA به منظور مقایسه میانگین تعداد واحدهای فروخته شده برای تبلیغ بیلبورد در مقابل تبلیغ رادیو ایجاد کرده و نتایج را زیر جدول ANOVA در صفحه گسترده نشان دهید (خطای معیار و مقدار آزمون t در ANOVA را روی خطهای جداگانه صفحه گسترده قرار دهید و برای هر مقدار، از دو رقم اعشار استفاده کنید).
 - (d) نتایج صفحه گسترده را چاپ کنید به گونه‌ای که همه اطلاعات در یک صفحه قرار بگیرند.
 - (e) صفحه گسترده را با نام McD4 ذخیره کنید.
- اجازه دهید تبلیغ در رادیو را گروه ۱، تلویزیون محلی را گروه ۲، بیلبوردها را گروه ۳ و روزنامه محلی را گروه ۴ نام‌گذاری کنیم.

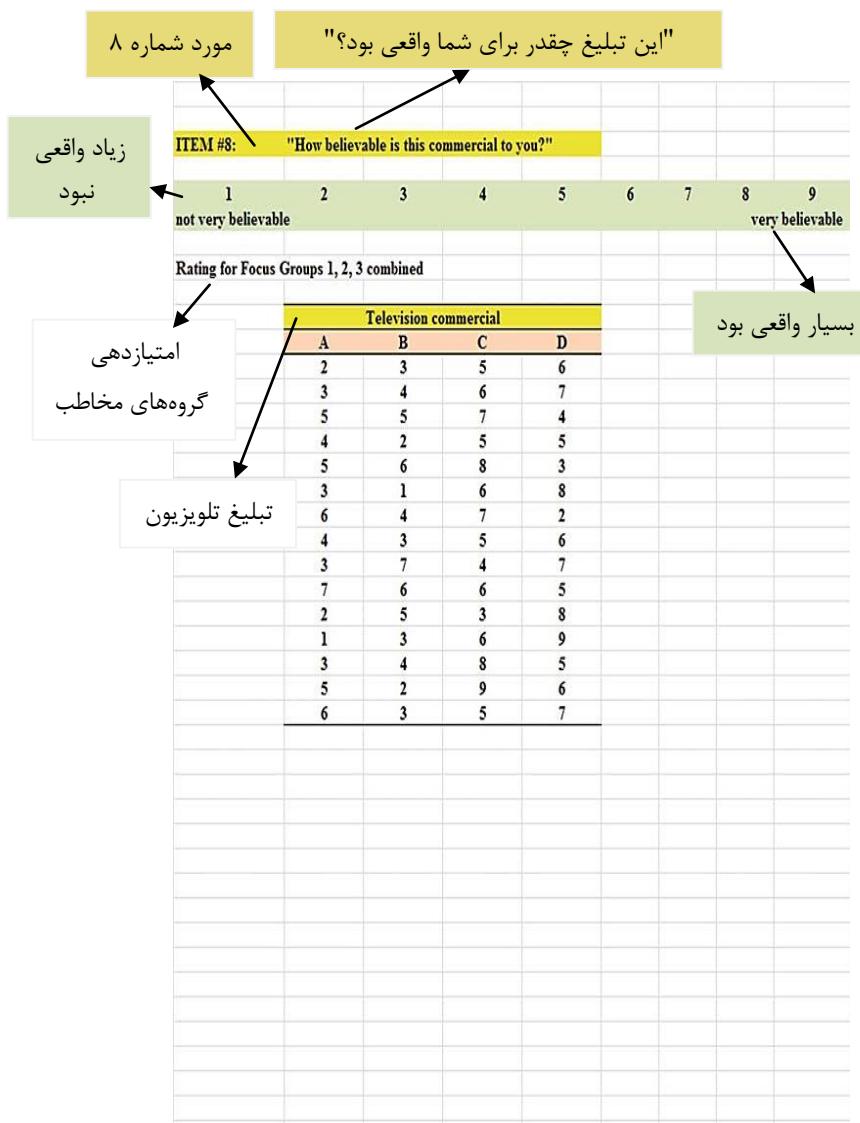
اکنون جواب پرسش‌های زیر را با استفاده از خروجی اکسل بدھید:

۱. فرضیه صفر و فرضیه تحقیق برای آزمون F در ANOVA چیست؟
۲. در خروجی اکسل MS_b چیست؟
۳. در خروجی اکسل MS_w چیست؟
۴. با استفاده از ماشین حساب $F = MS_b / MS_w$ محاسبه کنید.
۵. مقدار بحرانی F در خروجی اکسل چیست؟
۶. نتیجه آزمون F در ANOVA چیست؟
۷. نتیجه گیری آزمون F به زبان ساده و روان چیست؟
۸. اگر آزمون F در ANOVA اختلاف معنی‌داری بین اثر چهار نوع تبلیغات بر تعداد برگرهای Angus فروخته شده در هر هفته را نشان دهد، فرضیه صفر و فرضیه تحقیق برای آزمون t در ANOVA مقایسه تبلیغ بیلبورد (گروه ۳) در برابر تبلیغ رادیو (گروه ۱) چه می‌باشند؟
۹. میانگین تبلیغ بیلبورد در خروجی اکسل چیست؟
۱۰. میانگین تبلیغ رادیو در خروجی اکسل چیست؟
۱۱. درجه آزادی (df) برای آزمون t در ANOVA مقایسه تبلیغ بیلبورد در برابر تبلیغ رادیو چیست؟
۱۲. برای این درجه آزادی مقدار t بحرانی در آزمون t در ANOVA در ضمیمه E چیست؟
۱۳. با استفاده از ماشین حساب s.e.ANOVA (خطای معیار ANOVA) برای مقایسه تبلیغ بیلبورد در برابر تبلیغ رادیو را محاسبه کنید.
۱۴. با استفاده از ماشین حساب مقدار آزمون t در ANOVA برای مقایسه تبلیغ بیلبورد در برابر تبلیغ رادیو را محاسبه کنید.
۱۵. نتیجه آزمون t در ANOVA برای مقایسه تبلیغ بیلبورد در برابر تبلیغ رادیو چیست؟
۱۶. نتیجه گیری آزمون t در ANOVA برای مقایسه تبلیغ بیلبورد در برابر تبلیغ رادیو به زبان ساده و روان چیست؟

۳. فرض کنید شما بعنوان مشاور توسط پروکتر و گامبل^۱ برای تجزیه و تحلیل داده‌های مطالعه آزمایشی ۳ گروه افراد که مخاطب ۴ آگهی تجاری متفاوت بودند که برای یک نوع جدید از خمیر

^۱ Procter and Gamble

دندان کرست^۱ که هنوز در تلویزیون نشان داده نشده است، استخدام شده‌اید. به شرکت کنندگان یک پرسشنامه ۱۰ موردی برای پاسخگویی بعد از دیدن تبلیغ داده می‌شود. اطلاعات آماری پرسش مورد ۸ در شکل ۸-۸ برای ۴ آگهی تجاری داده شده است.



شکل ۸-۸: داده‌های صفحه کاری برای فصل ۸: تمرین عملی شماره ۳

^۱ Crest

- (a) داده‌ها را در صفحه گسترده اکسل وارد کنید
- (b) آزمون ANOVA یکطرفه را روی این داده‌ها انجام داده و نتایج جدول ANOVA را زیر داده‌های ورودی برای ۴ نوع از تبلیغ نشان دهید.
- (c) اگر مقدار F جدول ANOVA معنی‌دار است یک رابطه در اکسل برای محاسبه آزمون t ANOVA جهت مقایسه میانگین تعداد واحدهای فروخته شده برای تبلیغ B در برابر تبلیغ D ایجاد کرده و نتایج را زیر جدول ANOVA در صفحه گسترده نشان دهید (خطای معیار ANOVA و مقدار آزمون t در ANOVA را روی خطوطهای جداگانه صفحه گسترده قرار دهید و برای هر مقدار از دو رقم اعشار استفاده کنید).
- (d) نتایج صفحه گسترده را چاپ کنید به گونه‌ای که همه اطلاعات در یک صفحه قرار گیرد.
- (e) صفحه گسترده را با نام TV6 ذخیره کنید.
- حالا جواب پرسش‌های زیر را با استفاده از خروجی اکسل بدھید:
۱. فرضیه صفر و فرضیه تحقیق برای آزمون F در ANOVA چیست؟
 ۲. در خروجی اکسل MS_b چیست؟
 ۳. در خروجی اکسل MS_w چیست؟
 ۴. با استفاده از ماشین حساب $F = MS_b / MS_w$ محاسبه کنید.
 ۵. مقدار بحرانی F در خروجی اکسل چیست؟
 ۶. نتایج آزمون F در ANOVA چیست؟
 ۷. نتایج آزمون F در ANOVA در به زبان ساده و روان چیست؟
 ۸. اگر آزمون F در ANOVA اختلاف معنی‌داری بین ۴ نوع از تبلیغ را نشان دهد، فرضیه صفر و فرضیه تحقیق برای آزمون F در ANOVA برای مقایسه تبلیغ B در برابر تبلیغ D چیست؟
 ۹. میانگین تبلیغ B در خروجی اکسل چیست؟
 ۱۰. میانگین تبلیغ D در خروجی اکسل چیست؟
 ۱۱. درجه آزادی (df) برای آزمون t در ANOVA برای مقایسه تبلیغ B در برابر تبلیغ D چیست؟
 ۱۲. برای این درجه آزادی مقدار t بحرانی برای آزمون t در ضمیمه E چیست؟

۱۳. با استفاده از ماشین حساب S.e.ANOVA (خطای معیار) برای مقایسه تبلیغ B در برابر تبلیغ D را محاسبه کنید.

۱۴. با استفاده از ماشین حساب مقدار آزمون t در ANOVA برای مقایسه تبلیغ B در برابر تبلیغ D را محاسبه کنید.

۱۵. نتیجه آزمون t در ANOVA برای مقایسه تبلیغ B در برابر تبلیغ D چیست؟

۱۶. نتیجه‌گیری آزمون t در ANOVA برای مقایسه تبلیغ B در برابر تبلیغ D به زبان ساده و روان چیست؟

منابع

- Black K. Business statistics: for contemporary decision making. 6th ed. Hoboken: Wiley; 2010.
- WeiersRM. Introductionto business statistics. 7th ed.Mason: South-WesternCengageLearning; 2011.

پیوست‌ها

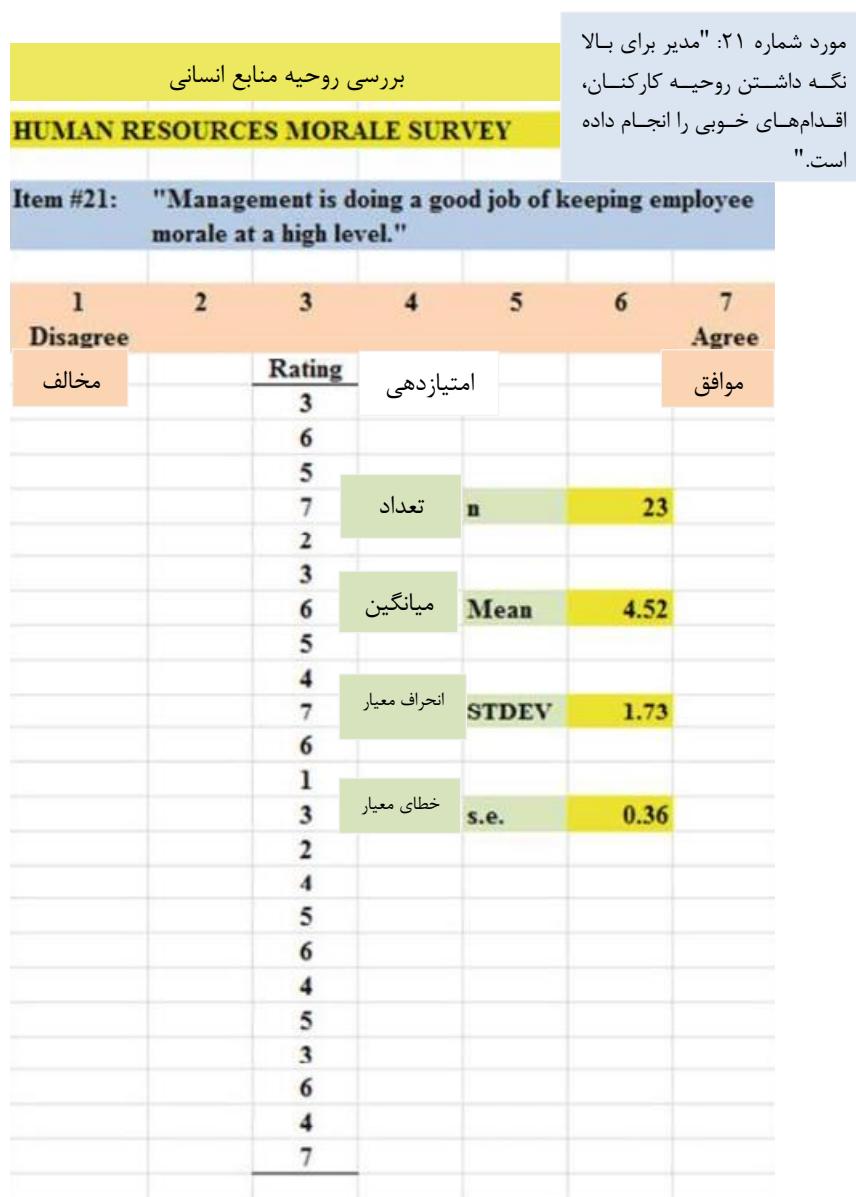
پیوست A: پاسخ تمرین‌های عملی پایان هر فصل

فصل ۱: پاسخ تمرین شماره ۱ (به شکل A.1 مراجعه کنید)

وال مارت سنت لوئیس			
Wal-Mart St. Louis			
فروش به دلار به ازای هر مشتری در هفته گذشته			
DOLLAR SALES PER CUSTOMER LAST WEEK			
127.12			
140.45			
104.64	تعداد	n	29
80.06			
114.07			
109.35	میانگین	Mean	\$ 117.46
117.28			
72.84			
67.67	انحراف معیار	STDEV	\$ 28.36
79.85			
109.96			
117.13	خطای معیار	s.e.	\$ 5.27
85.25			
149.36			
147.57			
153.54			
118.76			
69.86			
154.47			
154.88			
109.44			
97.36			
87.55			
154.85			
143.82			
145.55			
142.33			
122.57			
128.75			

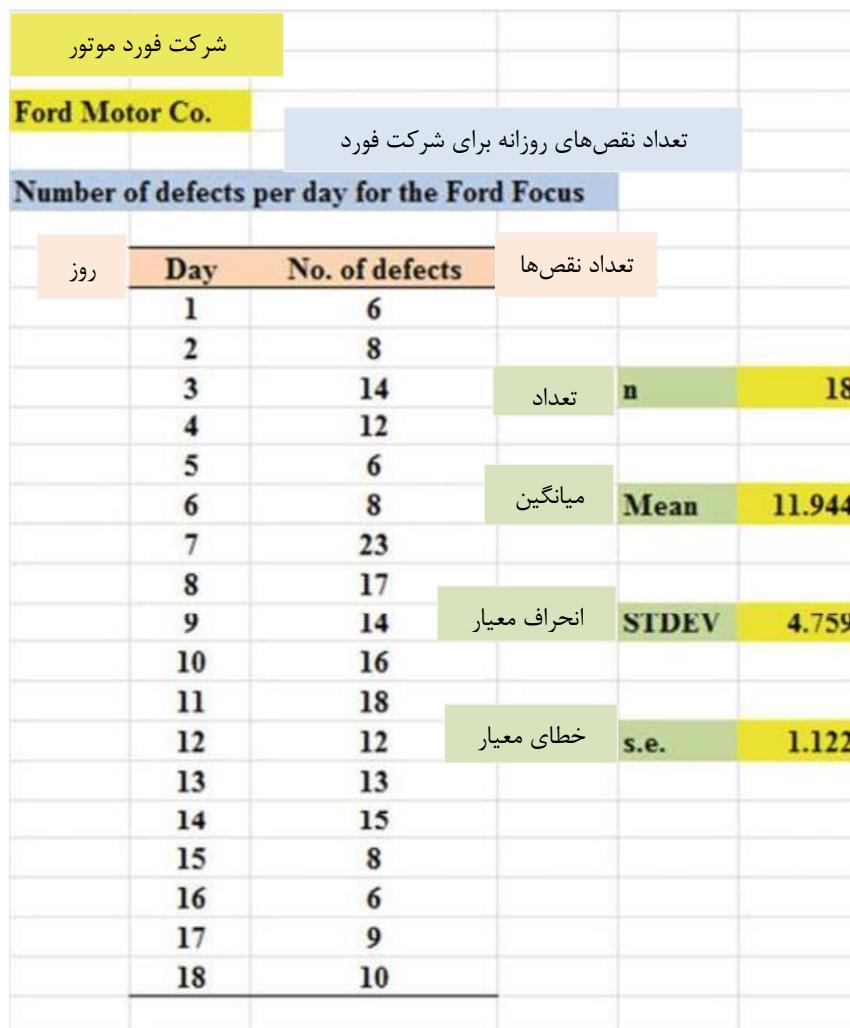
شکل A.1: پاسخ تمرین عملی فصل ۱: تمرین شماره ۱

فصل ۱: پاسخ تمرین شماره ۲ (به شکل A.2 مراجعه کنید)



شکل A.2: پاسخ تمرین عملی فصل ۱: تمرین شماره ۲

فصل ۱: پاسخ تمرین شماره ۳ (به شکل A.3 مراجعه کنید)



شکل A.3: پاسخ تمرین عملی فصل ۱: تمرین شماره ۳

فصل ۲: پاسخ تمرین شماره ۱ (به شکل A.4 مراجعه کنید)

اعداد چارچوب‌ها	اعداد چارچوب‌های تکراری	اعداد تصادفی
FRAME NUMBERS	Duplicate frame numbers	RANDOM NO.
1	44	0.355
2	33	0.311
3	38	0.305
4	43	0.784
5	13	0.569
6	10	0.365
7	50	0.778
8	1	0.851
9	48	0.156
10	61	0.469
11	4	0.708
12	22	0.905
13	40	0.470
14	37	0.093
15	35	0.225
16	60	0.510
17	59	0.173
18	7	0.776
19	17	0.174
20	30	0.999
21	29	0.830
22	47	0.770
23		0.20
51	5	0.5~
52	45	0.961
53	28	0.810
54	24	0.241
55	42	0.888
56	11	0.467
57	56	0.977
58	57	0.610
59	54	0.511
60	9	0.697
61	51	0.884
62	39	0.985
63	53	0.760
	26	0.163

شکل A.4: پاسخ تمرین عملی فصل ۲: تمرین شماره ۱

فصل ۲: پاسخ تمرین شماره ۲ (به شکل A.5 مراجعه کنید)

اعداد چارچوب‌ها	اعداد چارچوب‌های تکراری	اعداد تصادفی
FRAME NO.	Duplicate frame no.	Random number
1	45	0.955
2	102	0.804
3	16	0.995
4	8	0.976
5	109	0.221
6	64	0.580
7	37	0.509
8	31	0.208
9	27	0.475
10	76	0.471
11	9	0.952
12	70	0.330
13	13	0.481
14	32	0.754
15	56	0.816
16	46	0.986
17	3	0.692
18	98	0.634
19	10	0.526
20	100	0.825
21	29	
		0.224
90	101	0.964
91	15	0.901
92	61	0.854
93	90	0.059
94	78	0.451
95	69	0.006
96	93	0.621
97	75	0.764
98	59	0.317
99	2	0.805
100	35	0.984
101	20	0.776
102	73	0.398
103	11	0.747
104	24	0.441
105	82	0.637
106	5	0.152
107	17	0.409
108	34	0.963
109	104	0.072
110	51	0.990
111	6	0.455
112	84	0.508
113	96	0.466
114	67	0.650

شکل A.5: پاسخ تمرین عملی فصل ۲: تمرین شماره ۲

فصل ۲: پاسخ تمرین شماره ۳ (به شکل A.6 مراجعه کنید)

اعداد چارچوب‌ها	اعداد چارچوب‌های تکراری	اعداد تصادفی
FRAME NUMBERS	Duplicate frame numbers	Random number
1	47	0.364
2	68	0.637
3	15	0.217
4	69	0.725
5	67	0.192
6	38	0.577
7	43	0.788
8	50	0.527
9	65	0.040
10	40	0.575
11	57	0.189
12	37	0.648
13	22	0.293
14	3	0.832
15	17	0.819
16	60	0.215
17	5	0.670
18	29	0.112
19	74	0.078
20	72	0.766
21	14	0.972
22	41	0.861
23	53	0.495
24	9	0.004
25	19	0.066
26	26	0.26
60	21	0..
61	26	0.949
62	36	0.241
63	70	0.626
64	39	0.044
65	2	0.683
66	54	0.378
67	44	0.030
68	25	0.941
69	61	0.599
70	23	0.118
70	27	0.166
71	46	0.722
72	35	0.747
73	11	0.368
74	7	0.429
75	12	0.299
76	30	0.110

شکل A.6: پاسخ تمرین عملی فصل ۲: تمرین شماره ۳

فصل ۳: پاسخ تمرین شماره ۱ (به شکل A.7 مراجعه کنید)



شکل A.7: پاسخ تمرین عملی فصل ۳: تمرین شماره ۱

فصل ۳: پاسخ تمرین شماره ۲ (به شکل A.8 مراجعه کنید)



شکل A.8: پاسخ تمرین عملی فصل ۳: تمرین شماره ۲

فصل ۳: پاسخ تمرین شماره ۳ (به شکل A.9 مراجعه کنید)



شکل A.9: پاسخ تمرین عملی فصل ۳: تمرین شماره ۳

فصل ۴: پاسخ تمرین شماره ۱ (به شکل A.10 مراجعه کنید)

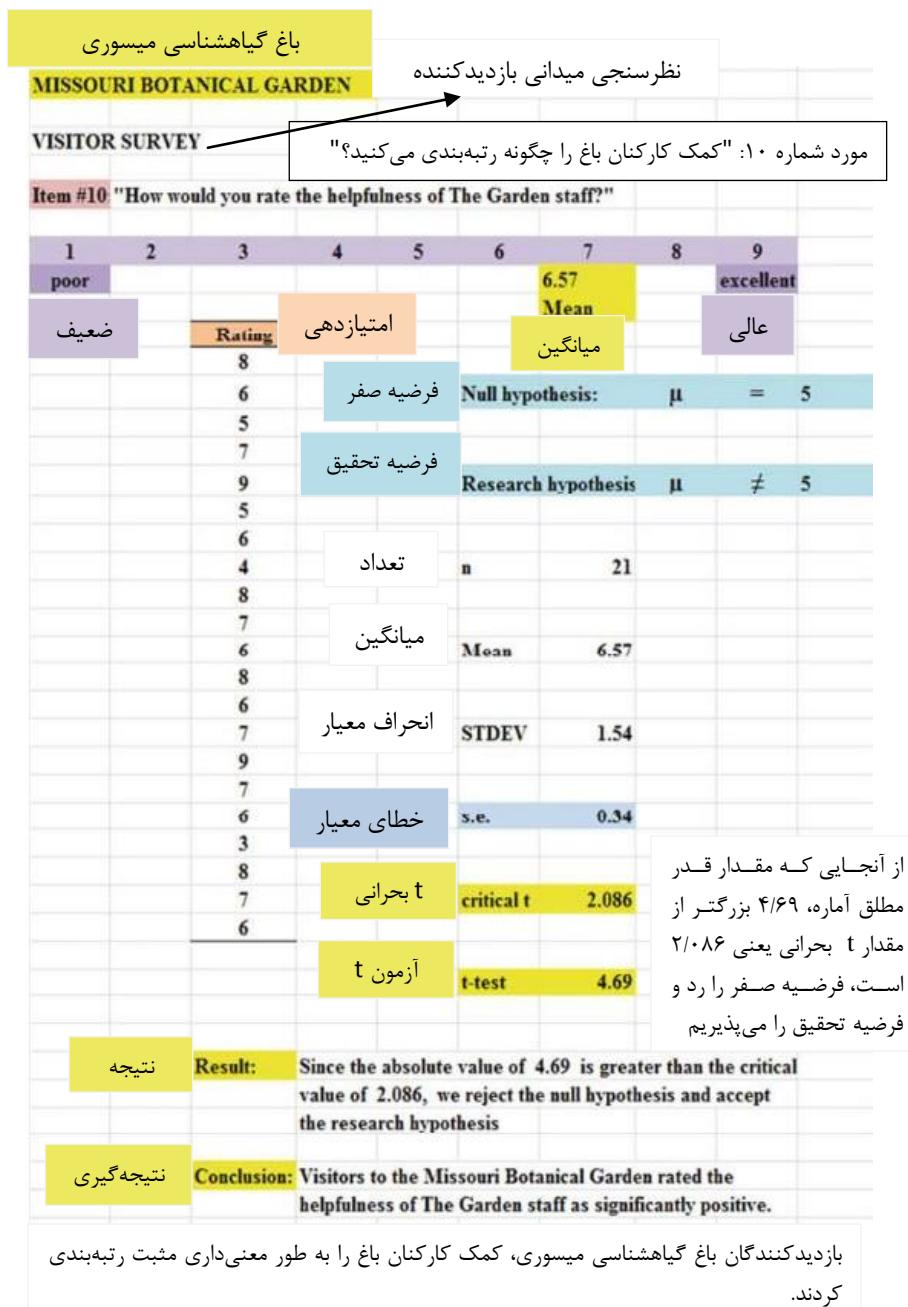


فصل ۴: پاسخ تمرین شماره ۲ (به شکل A.11 مراجعه کنید)



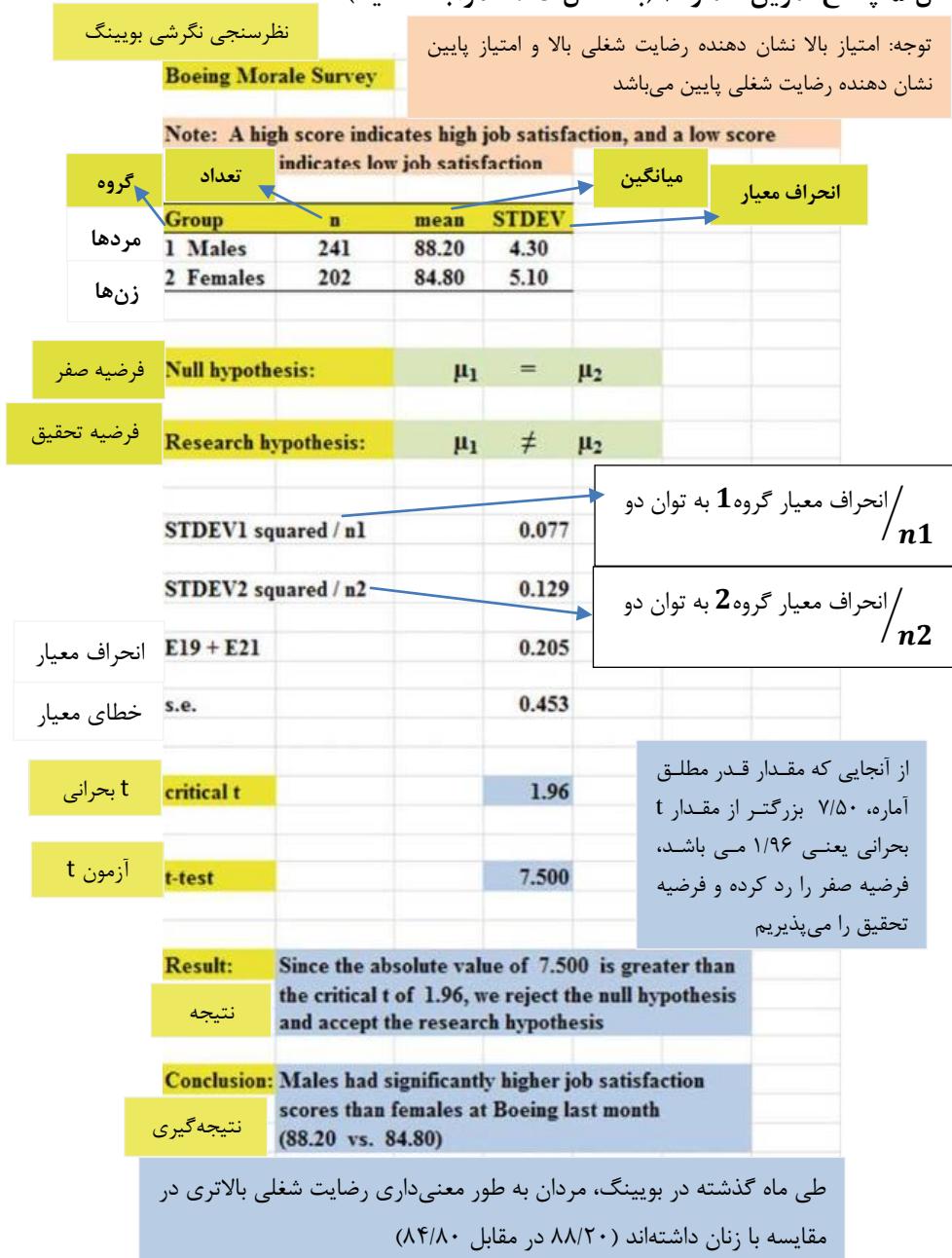
شکل A.11: پاسخ تمرین عملی فصل ۴: تمرین شماره ۲

فصل ۴: پاسخ تمرین شماره ۳ (به شکل A.12 مراجعه کنید)



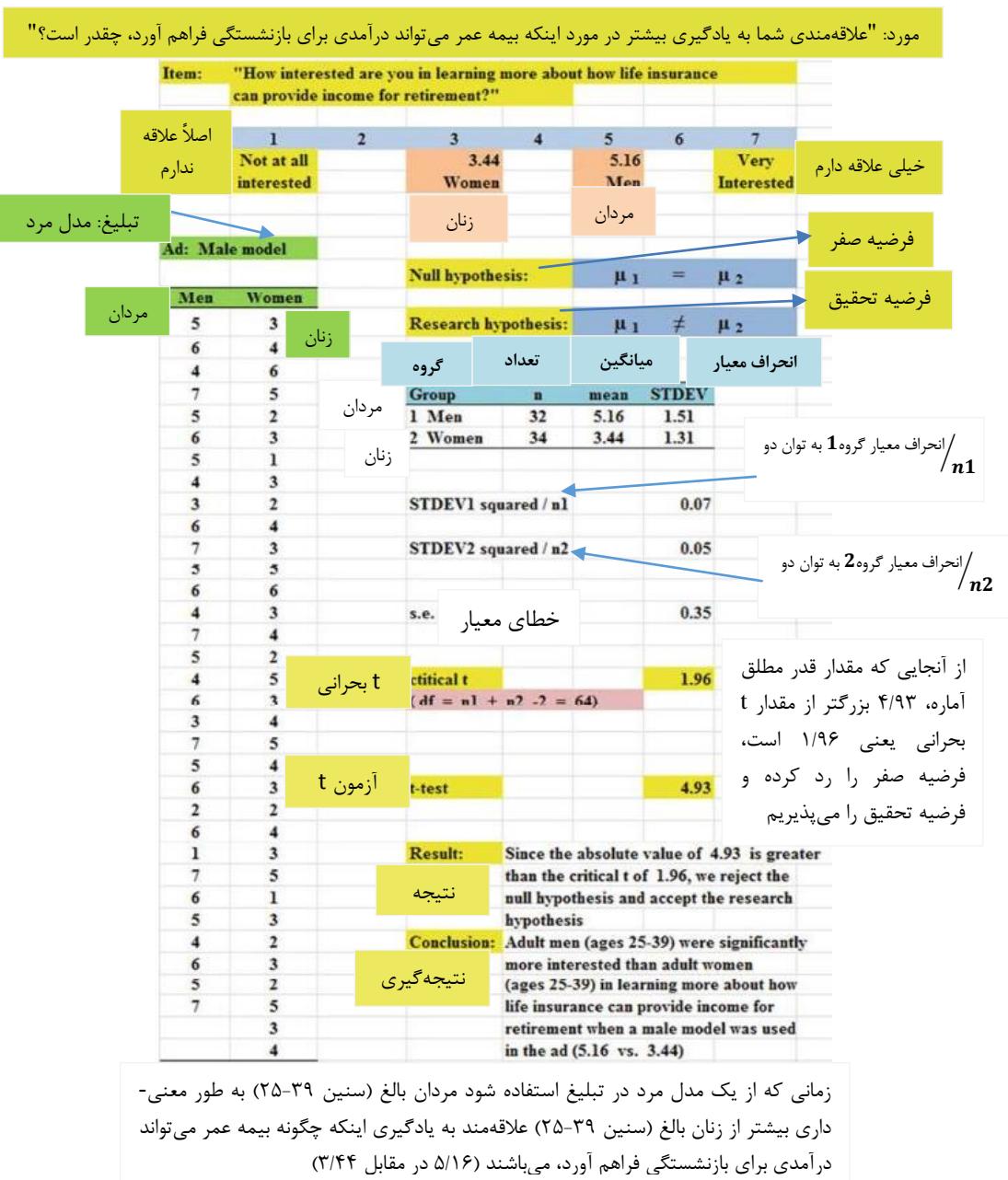
شکل A.12: پاسخ تمرین عملی فصل ۴: تمرین شماره ۳

فصل ۵: پاسخ تمرین شماره ۱ (به شکل A.13 مراجعه کنید)

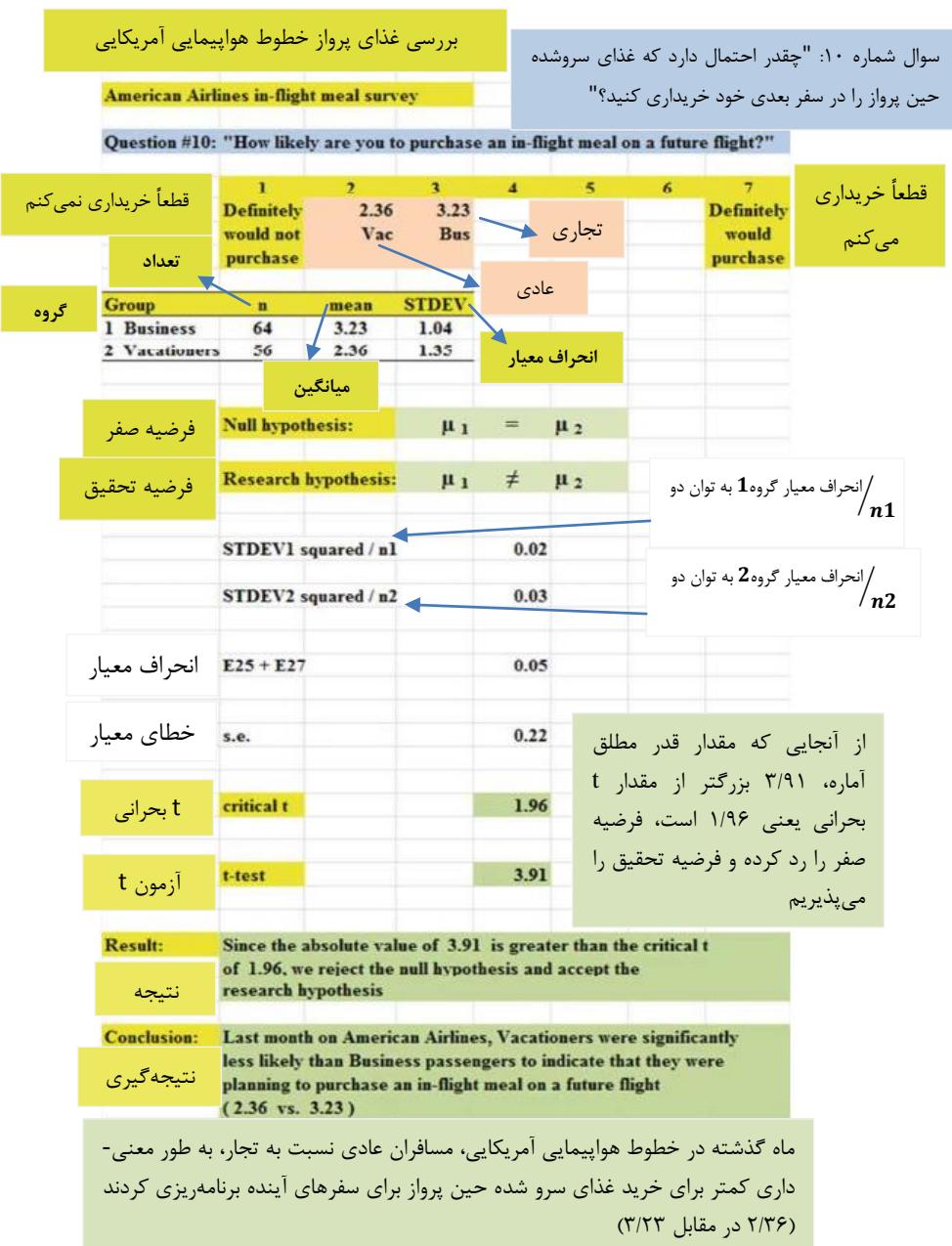


شکل A.13: پاسخ تمرین عملی فصل ۵: تمرین شماره ۱

فصل ۵: جواب تمرین شماره ۲ (به شکل A.14 مراجعه کنید)

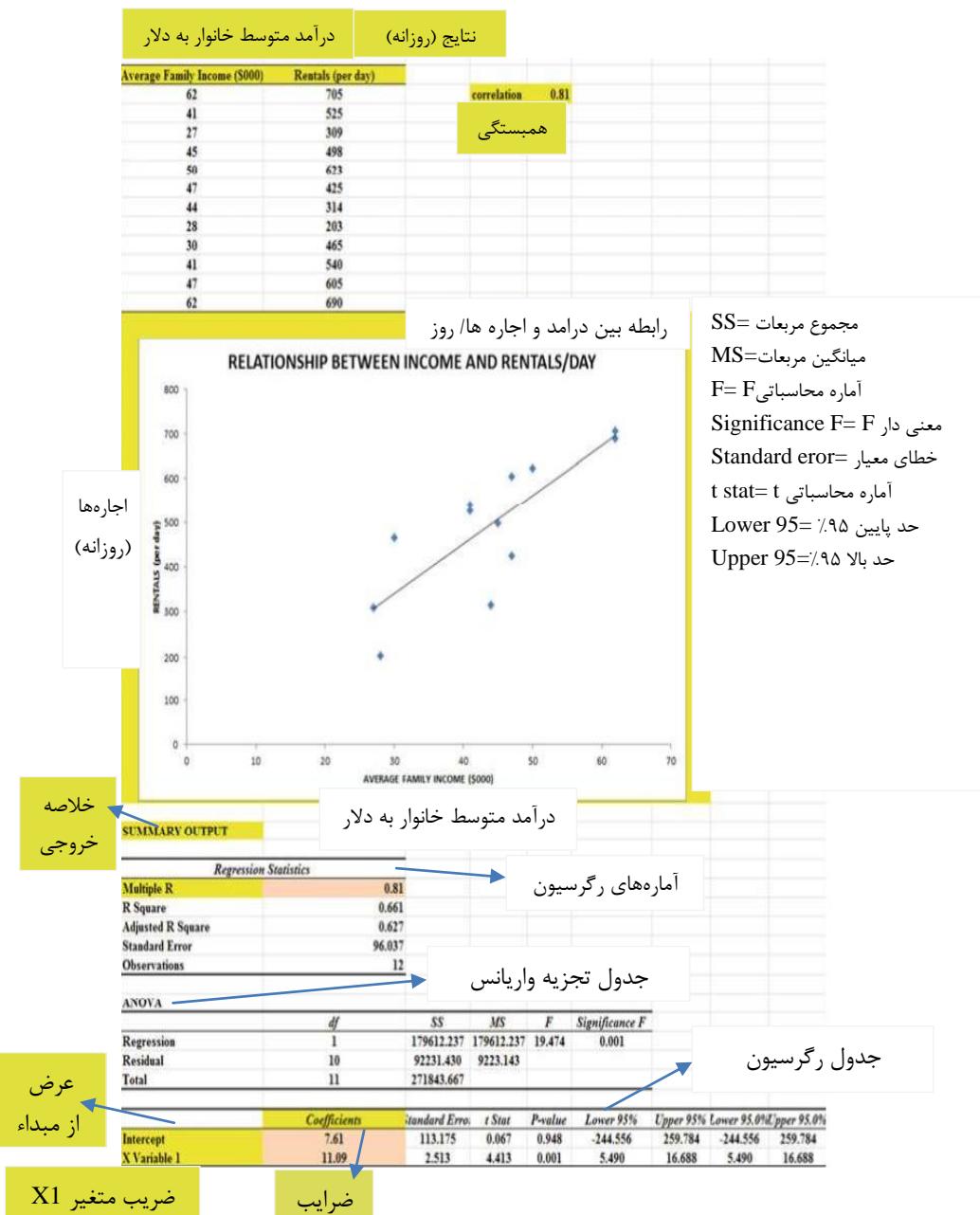


فصل ۵: پاسخ تمرین شماره ۳ (به شکل A.14 مراجعه کنید)



شکل A.15: پاسخ تمرین عملی فصل ۵: تمرین شماره ۳

فصل ۶: پاسخ تمرین شماره ۱ (به شکل A.16 مراجعه کنید)



شکل A.16: پاسخ تمرین عملی فصل ۶: تمرین شماره ۱

فصل ۶: پاسخ تمرین عملی شماره ۱ (ادامه)

$$1- \quad a = y_{\text{مبدأ}} = 7.61 \quad \text{عرض از مبدأ} = y/61$$

$$2- \quad b = 11.09 \quad \text{شیب} = 11/09$$

$$3- \quad Y = a + b X$$

$$Y = 7.61 + 11.09 X$$

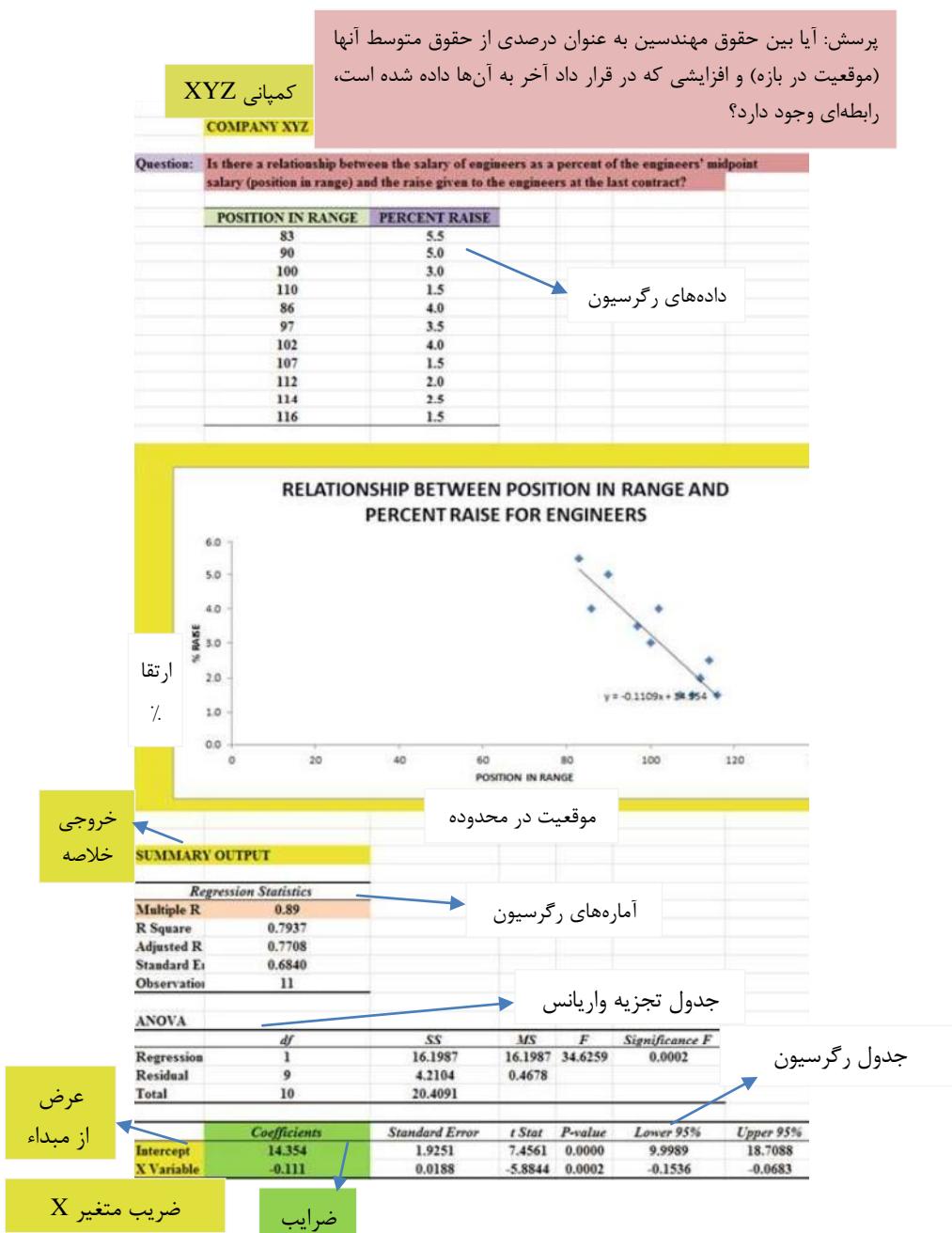
$$4- \quad Y = 7.61 + 11.09 (50)$$

$$Y = 7.61 + 554.5$$

$$Y = 562.11$$

$$Y = 562.11 \quad \text{اجاره به ازای هر روز} = 562$$

فصل ۶: پاسخ تمرین شماره ۲ (به شکل A.17 مراجعه کنید)



شکل A.17: پاسخ تمرین عملی فصل ۶: تمرین شماره ۲

فصل ۶: پاسخ تمرین شماره ۲ (ادامه)

(d) $a = y = 14/354$ عرض از مبدأ(به علامت منفی به عنوان شبیه دقت کنید) $b = -0/111$ شبیه(e) $Y = a + b X$

$$Y = 14.354 - 0.111 X$$

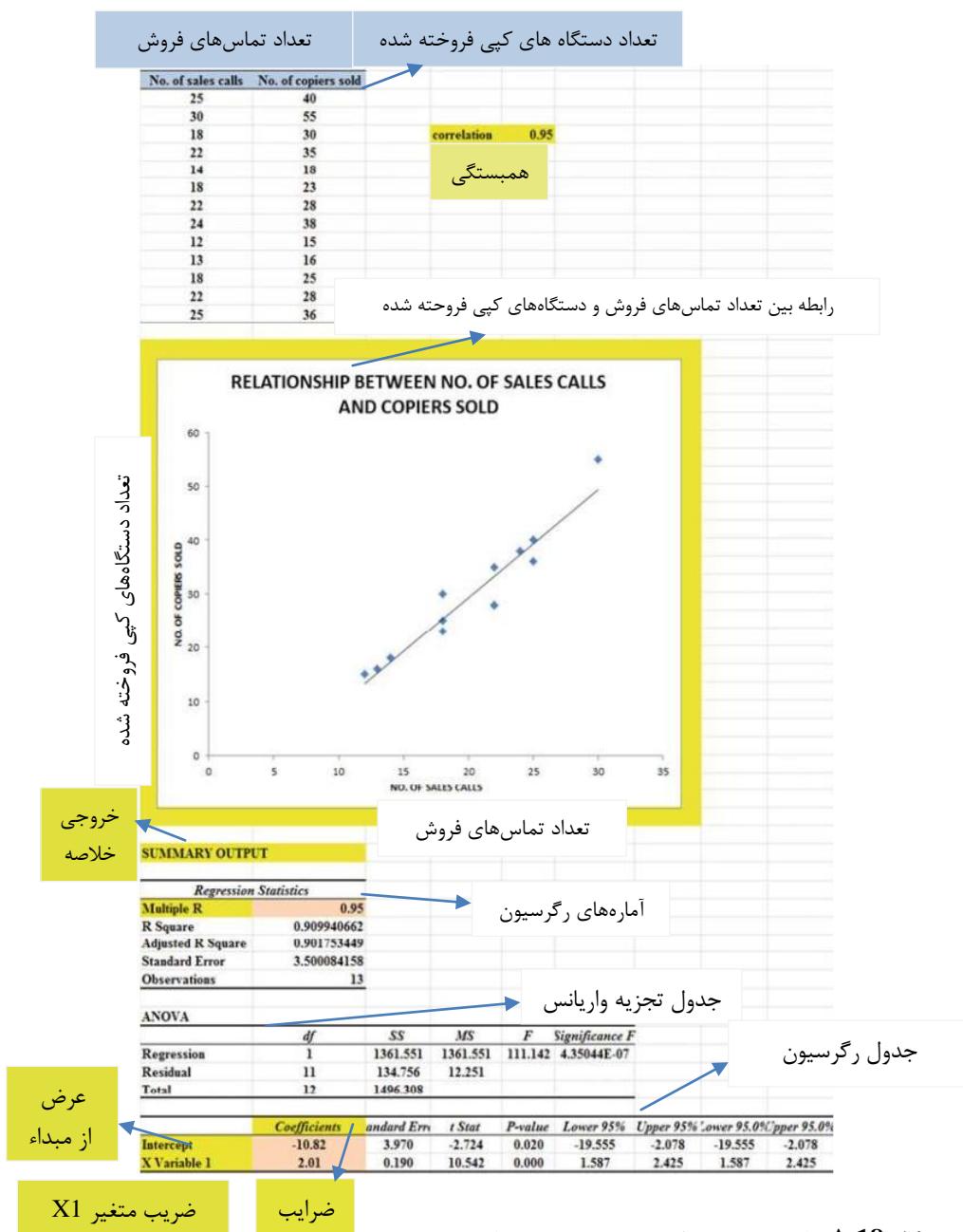
(f) $r = -0.89$ به همبستگی منفی توجه کنید!(g) $Y = 14.354 - 0.111 (90)$

$$Y = 14.354 - 9.99$$

$$Y = 4.4\%$$

(h) حدود دو درصد

فصل ۶: پاسخ تمرین شماره ۳ (به شکل A.18 مراجعه کنید)



شکل A.18: پاسخ تمرین عملی فصل ۶: تمرین شماره ۳

فصل ۶: پاسخ تمرین شماره ۳ (ادامه)

$$1- r = 0.95$$

$$2- a = y_{\text{مبدأ}} = -10/82$$

$$3- b = 2/0.1$$

$$4- Y = a + b X$$

$$Y = -10.82 + 2.01 X$$

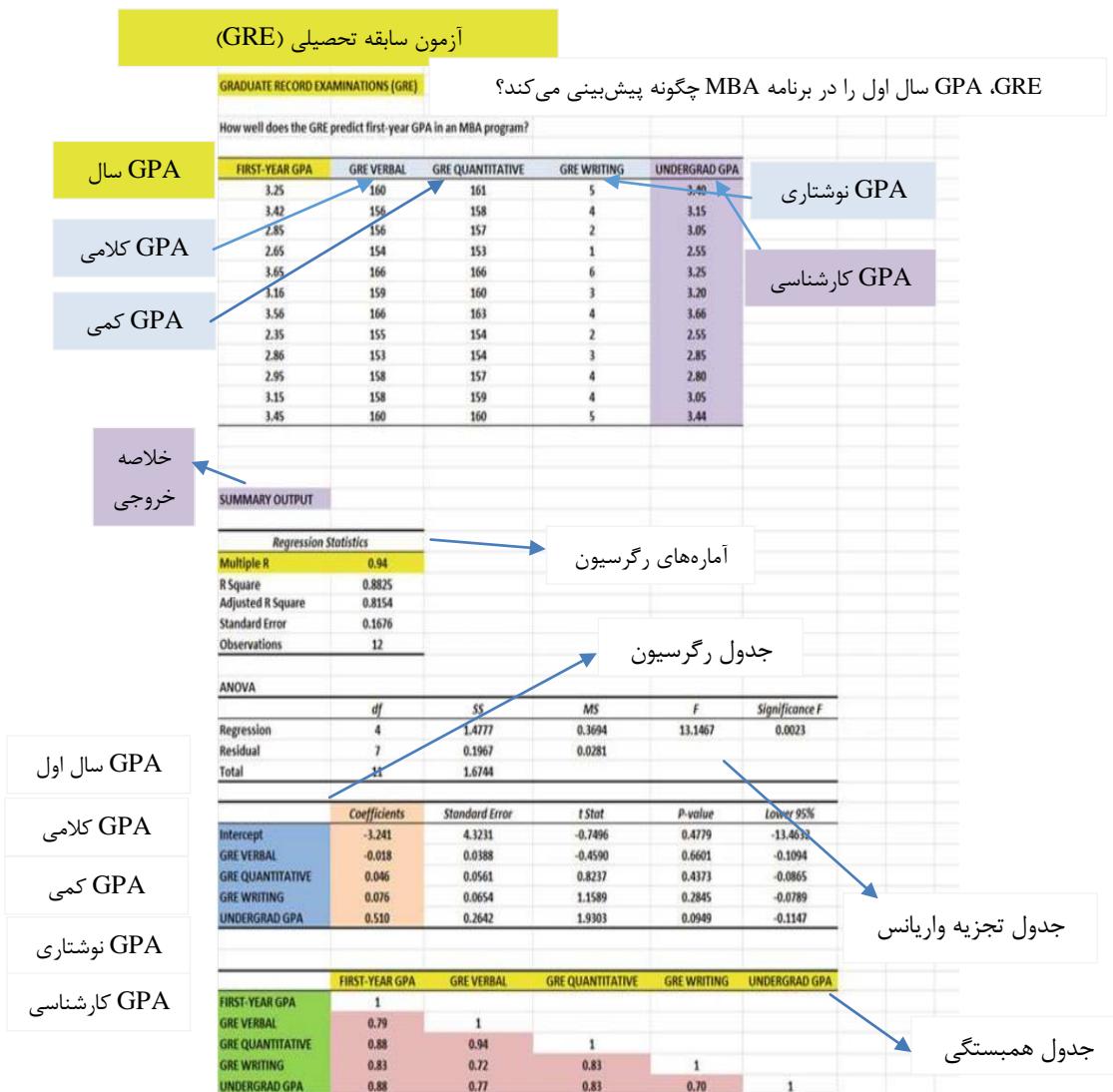
$$5- Y = -10.82 + 2.01 (25)$$

$$Y = -10.82 + 5.25$$

$$Y = 39.43$$

Y = 39 دستگاه کپی به ازای هر ماه فروخته شده است

فصل ۷: پاسخ تمرین شماره ۱ (به شکل A.19 مراجعه کنید)



شکل A.19: پاسخ تمرین عملی فصل ۷: تمرین شماره ۱

فصل ۷: پاسخ تمرین شماره ۱ (ادامه)

۱- همبستگی چندگانه $R_{xy} = 0.94$

۲- $a = y - 0.241$ عرض از مبدأ

۳- ضریب $b_1 = -0.18$

۴- ضریب $b_2 = 0.46$

۵- ضریب $b_3 = 0.76$

۶- ضریب $b_4 = 0.51$

۷- $Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_4$

$$Y = -0.241 - 0.18 X_1 + 0.46 X_2 + 0.76 X_3 + 0.51 X_4$$

۸- $Y = -0.241 - 0.18(159) + 0.46(154) + 0.76(4) + 0.51(305)$

$$Y = -0.241 - 2.862 + 7.084 + 0.304 + 1.556$$

$$Y = 8.944 - 6.103$$

$$Y = 2.84$$

۹- 0.88

۱۰- 0.77

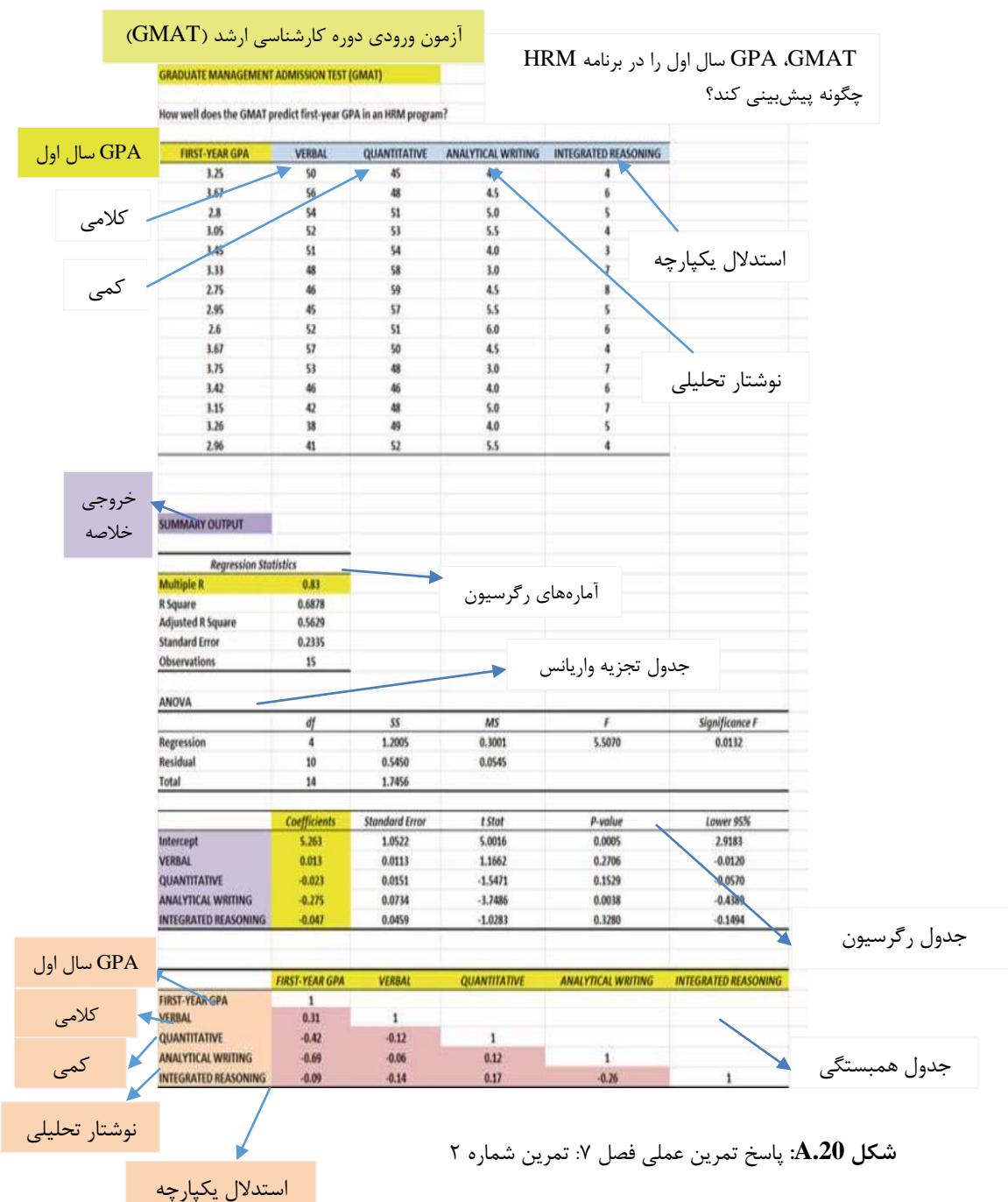
۱۱- 0.83

۱۲- 0.70

۱۳- بهترین پیش‌بینی‌کننده GPA سال اول، اثر متقابل بین GRE کمی و کارشناسی است ($r = 0.88$)

۱۴- ترکیب چهار پیش‌بینی کننده، GPA سال اول را خیلی بهتر ($R_{xy} = 0.94$) از بهترین پیش‌بینی‌کننده‌های منفرد، پیش‌بینی می‌کند ($r = 0.88$)

فصل ۷: پاسخ تمرین شماره ۲ (به شکل A.20 مراجعه کنید)



شکل A.20: پاسخ تمرین عملی فصل ۷: تمرین شماره ۲

فصل ۷: پاسخ تمرین شماره ۲ (ادامه)

۱- همبستگی چندگانه $R_{xy} = +0.83$

۲- $a = y - b_1 X_1 - b_2 X_2 - b_3 X_3 - b_4 X_4$
 $a = 5/263 + +0.13 X_1 - +0.23 X_2 - +0.275 X_3 - +0.47 X_4$

۳- ضریب $b_1 = -0.023$

۴- ضریب $b_2 = +0.275$

۵- ضریب $b_3 = +0.47$

۶- ضریب $b_4 = +0.47$

۷- $Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + b_4 X_4$

$Y = 5/263 + +0.13 (52) - +0.23 (48) - +0.275 (45) - +0.47 (6)$

$Y = 5/263 + +0.676 - 1/10.4 - 1/23.8 - +0.282$

$Y = 5/939 - 2/624$

$Y = 3/32$

۹- -0.31

۱۰- -0.42

۱۱- -0.69

۱۲- -0.09

۱۳- -0.12

۱۴- $+0.12$

۱۵- -0.26

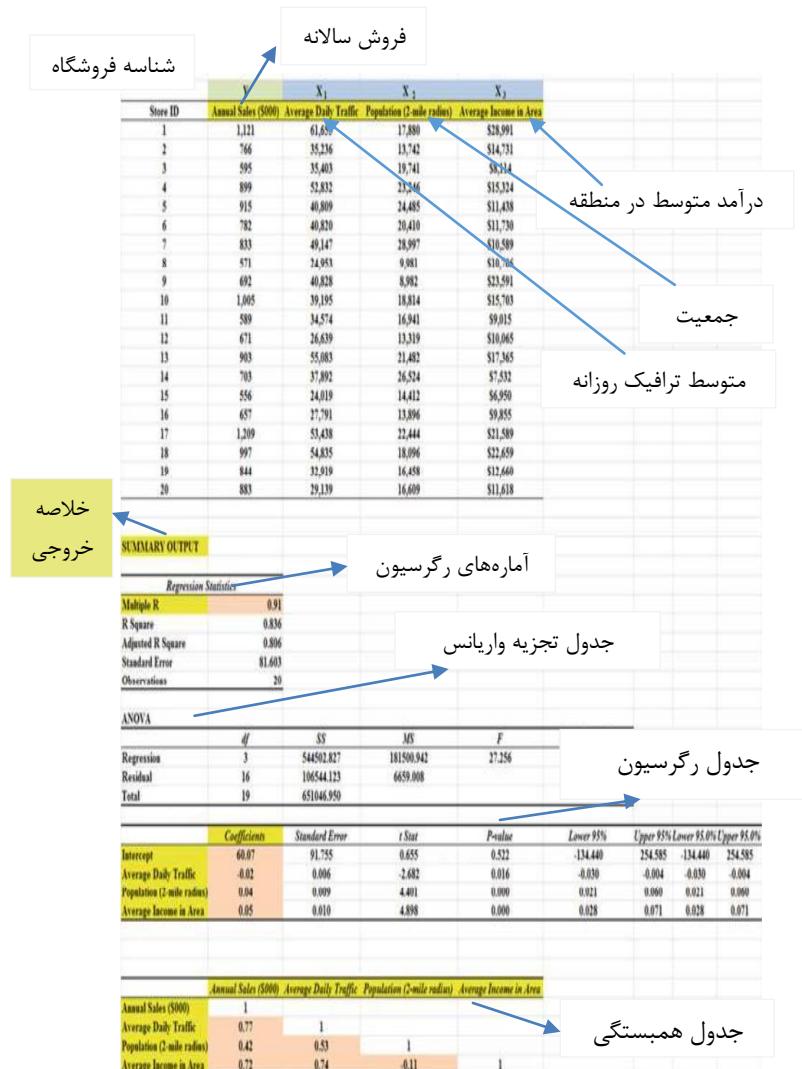
۱۶- $+0.17$

۱۷- بهترین پیش‌بینی‌کننده GPA سال اول، تجزیه و تحلیل نوشتاری بود ($r = -0.69$).

توجه داشته باشد که بهترین پیش‌بینی کننده، بالاترین همبستگی است، صرف نظر از اینکه منفی یا مثبت باشد.

۱۸- ترکیب چهار پیش‌بینی‌کننده، GPA سال اول را خیلی بهتر ($R_{xy} = +0.83$) از بهترین پیش‌بینی‌کننده‌های منفرد، پیش‌بینی می‌کند ($r = -0.69$).

فصل ۷: پاسخ تمرین شماره ۳ (به شکل A.21 مراجعه کنید)



شکل A.21: پاسخ تمرین عملی فصل ۷: تمرین شماره ۳

فصل ۷: پاسخ تمرین شماره ۳ (ادامه)

1- همبستگی چندگانه $R_{xy} = +/91$

2- عرض از مبداء $y = 60/07$

3- متوسط ترافیک روزانه $-0/02$

4- جمعیت $= 0/04$

5- درآمد متوسط $= 0/05$

6- $Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3$

$$Y = 60/07 - 0/02 X_1 + 0/04 X_2 + 0/05 X_3$$

7- $Y = 60/07 - 0/02 X_1 + 0/04 X_2 + 0/05 X_3$

$$Y = 60/07 - 0/02 (42000) + 0/04 (23000) + 0/05 (22000)$$

$$Y = 60/07 - 840 + 920 + 1100$$

$$Y = 1240/07$$

$Y = 1240000$ دلار یا $1/24$ میلیون دلار

8- $0/77$

9- $0/42$

10- $0/72$

11- $0/53$

12- $-0/11$

۱۳- متوسط ترافیک روزانه بهترین پیش‌بینی کننده فروش سالانه است زیرا همبستگی آن با

فروش سالانه $0/77$ می‌باشد، و پیش‌بینی کننده‌های دیگر همبستگی کوچکتر از $0/77$

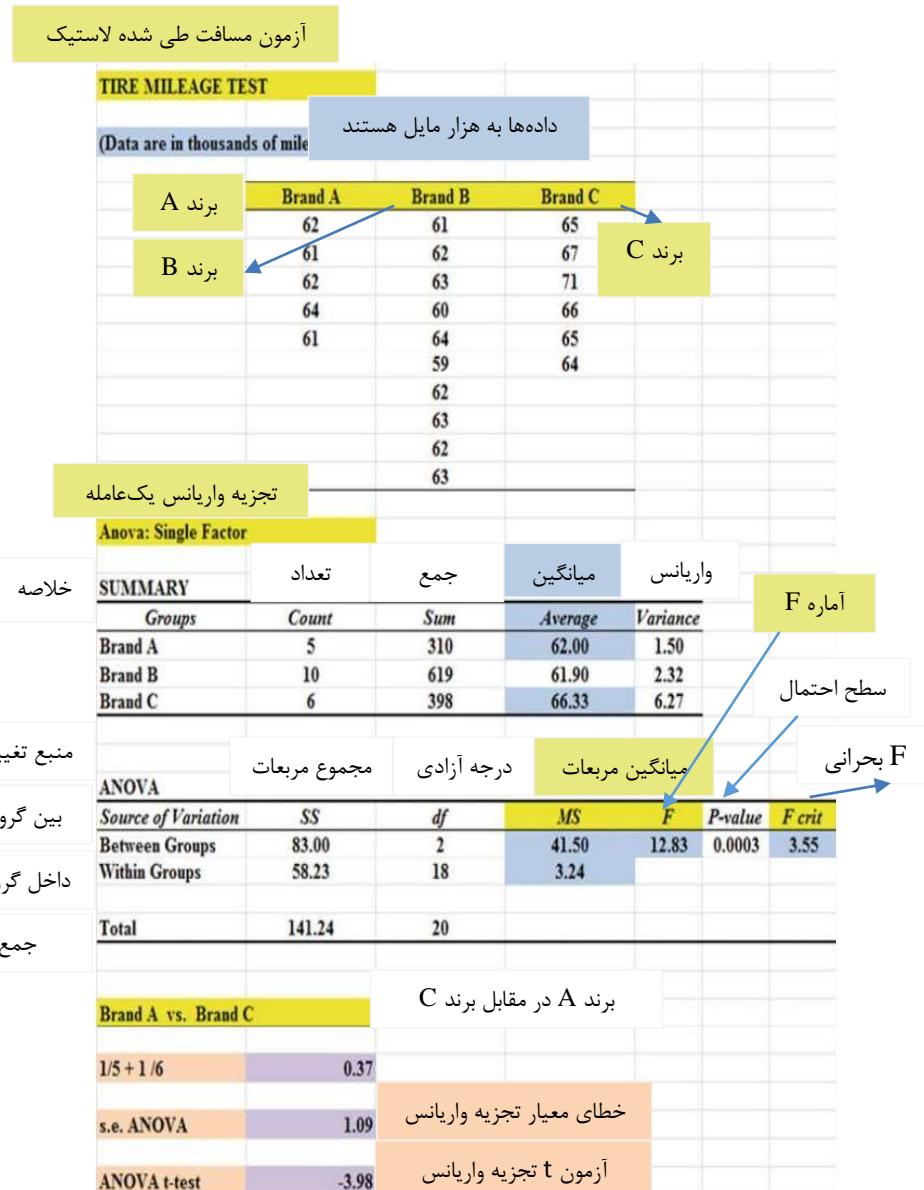
دارند ($0/72$ و $0/42$).

۱۴- ترکیب سه پیش‌بینی کننده 91 درصد فروش سالانه را پیش‌بینی می‌کند و این خیلی

بهتر از بهترین پیش‌بینی کننده‌ی منفرد است که همبستگی آن با فروش سالانه 77

درصد می‌باشد.

فصل ۸: پاسخ تمرین شماره ۱ (به شکل A.22 مراجعه کنید)



شکل A.22: پاسخ تمرین عملی فصل ۸: تمرین شماره ۱

فصل ۸: پاسخ تمرین شماره ۱ (ادامه)

۱- $\mu_A = \mu_B = \mu_C$: فرضیه صفر

$\mu_A \neq \mu_B \neq \mu_C$: فرضیه تحقیق

۲- $MS_b = 41/5$.

۳- $MS_w = 3/24$

۴- $F = 12/81$

۵- $t = 3/55$ بحرانی

۶- از آنجایی که مقدار F محاسباتی یعنی ۱۲/۸۱ از F بحرانی یعنی ۳/۵۵ بزرگتر است،

فرضیه صفر را رد کرده و فرضیه تحقیق را می‌پذیریم.

۷- تفاوت معنی‌داری در مایل‌های طی شده در سه برنده لاستیک وجود دارد.

برند A در مقابل برنده C

۸- $\mu_A = \mu_C$: فرضیه صفر

$\mu_A \neq \mu_C$: فرضیه تحقیق

۹- ۶۲

۱۰- ۶۶/۳۳

۱۱- درجه آزادی = ۲۱-۳ = ۱۸

۱۲- $t = 2/101$ بحرانی

۱۳- s.e. ANOVA = $SQRT(MS_w \times \{1/5 + 1/6\}) = SQRT(3.24 \times \{0.20 + 0.167\}) = SQRT(1.19 \times 1.09)$

۱۴- $t = 1/0.9 / (62 - 66/33) = -3/97$ تجزیه واریانس

۱۵- از آنجایی که مقدار قدر مطلق آماره محاسباتی، ۳/۹۷، بزرگتر از مقدار بحرانی، ۲/۱۰۱، است، فرضیه صفر را رد کرده و فرضیه تحقیق را می‌پذیریم.

۱۶- برنده C به طور معنی‌داری مایل‌های طی شده بیشتری از برنده A دارد (۶۶.۰۰۰ در مقابل ۶۲.۰۰۰).

برند A در مقابل برنده B

۱۷- $\mu_A = \mu_B$: فرضیه صفر

$\mu_A \neq \mu_B$: فرضیه تحقیق

۱۸- ۶۲

۱۹- $61/9$

۲۰- درجه آزادی $= 21 - 3 = 18$

۲۱- بحرانی $t = 2/101$

۲۲- s.e. ANOVA = $SQRT(MS_w \times \{1/5 + 1/10\}) = SQRT(3.24 \times \{0.20 + 0.10\}) = SQRT(0.972) = 0.99$

۲۳- تجزیه واریانس $t = 0/10 / (62 - 61/9) = 0/10$

۲۴- از آنجایی که مقدار قدر مطلق آماره محاسباتی، $10/0$ ، از مقدار بحرانی $10/2$ کمتر است، فرضیه صفر را می‌پذیریم.

۲۵- هیچ تفاوت معنی‌داری بین مایل‌های طی شده برندهای A و B وجود ندارد.

برند B در مقابل برند C

۲۶- فرضیه صفر: $\mu_C = \mu_B$

فرضیه تحقیق: $\mu_C \neq \mu_B$

۲۷- $61/90$

۲۸- $66/33$

۲۹- درجه آزادی $= 21 - 3 = 18$

۳۰- بحرانی $t = 10/1$

۳۱- s.e. ANOVA = $SQRT(MS_w \times \{1/10 + 1/6\}) = SQRT(3.24 \times \{0.10 + 0.167\}) = SQRT(0.87) = 0.93$

۳۲- تجزیه واریانس $t = 0/93 / (66/33 - 66/90) = -4/76$

۳۳- از آنجایی که مقدار قدر مطلق آماره محاسباتی $4/76$ بزرگتر از مقدار بحرانی $10/2$ می‌باشد، فرضیه صفر را رد کرده و فرضیه تحقیق را می‌پذیریم.

۳۴- برند C به طور معنی‌داری مایل‌های طی شده بیشتری نسبت به برند B دارد (66.000 در مقابل 62.000).

۳۵- برند C به طور معنی‌داری از هر دو برند B و A مایل‌های طی شده بیشتری دارد. تفاوت معنی‌داری بین مایل‌های طی شده برند A و B وجود نداشت.

۳۶- از آنجایی که مایل‌های طی شده برند A به طور معنی‌داری کمتر از برند C است، هرگز نباید در تبلیغات خود ادعا کنیم که برند A بیشتر از برند C دوام می‌آورد. از آنجایی که برند A و برند B مایل‌های طی شده مشابهی دارند، هرگز نباید ادعا کنیم که لاستیک برند A بیشتر از برند B دوام می‌آورد.

فصل ۸: پاسخ تمرین شماره ۲ (به شکل A.23 مراجعه کنید)



شکل A.23: پاسخ تمرین عملی فصل ۸: تمرین شماره ۲

فصل ۸: پاسخ تمرین شماره ۲ (ادامه)

۱- $\mu_1 = \mu_2 = \mu_3 = \mu_4$: فرضیه صفر
 $\mu_1 \neq \mu_2 \neq \mu_3 \neq \mu_4$: فرضیه تحقیق

۲- $MS_b = 70.57/47$

۳- $MS_w = 90/12$

۴- $F = 78/31$

۵- $F = 78/31$ بحرانی

۶- از آنجایی که مقدار قدر مطلق آماره محاسباتی $78/31$ بزرگتر از مقدار بحرانی F یعنی

$2/82$ می‌باشد، فرضیه صفر را رد کرده و فرضیه تحقیق را می‌پذیریم.

۷- تفاوت معنی‌داری بین آگنوس برگرهای فروخته شده در ۴ نوع تبلیغ رسانه‌ای وجود داشت.

-۱- فرضیه صفر: $\mu_3 = \mu_1$

فرضیه تحقیق: $\mu_3 \neq \mu_1$

$337/42 - 9$

$299/58 - 10$

۱۱- درجه آزادی $= 48 - 4 = 44$

$t = 1/96$ بحرانی

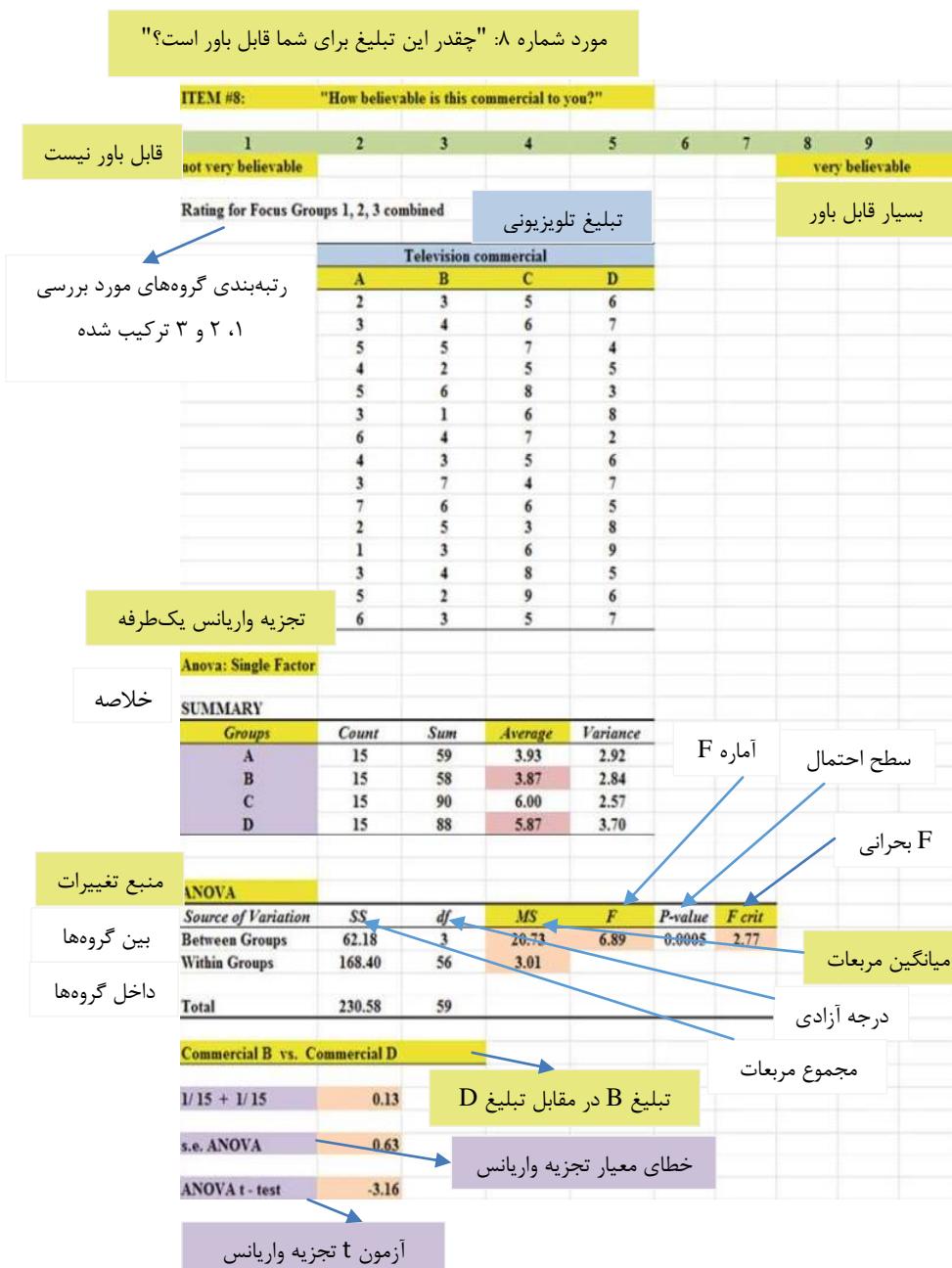
۱۳- s.e. ANOVA = $SQRT(MS_w \times \{1/12 + 1/12\}) = SQRT(90.12 \times \{0.083 + 0.083\}) = SQRT(14.96) = 3.87$

۱۴- $t = (337/42 - 299/58) / 3/87 = 9/78$ تجزیه واریانس

۱۵- از آنجایی که مقدار قدر مطلق آماره محاسباتی، $9/78$ بیشتر از t بحرانی یعنی $1/96$ است، فرضیه صفر را رد کرده و فرضیه تحقیق را می‌پذیریم.

۱۶- تابلوهای تبلیغاتی به‌طور معنی‌داری بیشتر از تبلیغات رادیویی منجر به فروش آگنوس برگرهای شده است (337 در برابر 300).

فصل ۸: پاسخ تمرین شماره ۳ (به شکل A.24 مراجعه کنید)



شکل A.24: پاسخ تمرین عملی فصل ۸: تمرین شماره ۳

فصل ۸: پاسخ تمرین شماره ۳ (ادامه)

۱- $\mu_A = \mu_B = \mu_C = \mu_D$: فرضیه صفر

$\mu_A \neq \mu_B \neq \mu_C \neq \mu_D$: فرضیه تحقیق

۲- $MS_b = ۲۰/۷۳$

۳- $MS_w = ۳/۰۱$

۴- $F = ۶/۸۹$

۵- $F_{بـحرانـي} = ۲/۷۷$

۶- از آنجایی که مقدار آماره محاسباتی F بزرگتر از مقدار بحرانی یعنی $۲/۷۷$ است، فرضیه صفر را رد کرده و فرضیه تحقیق را می‌پذیریم.

۷- تفاوت معنی‌داری بین چهار تبلیغ تلویزیونی وجود دارد.

-۱ فرضیه صفر: $\mu_D = \mu_B$

فرضیه تحقیق: $\mu_D \neq \mu_B$

$۳/۸۷ - ۹$

$۵/۸۵ - ۱۰$

۱۱- درجه آزادی $= ۶۰ - ۴ = ۵۶$

$t_{بـحرانـي} = ۱/۹۶ - ۱۲$

۱۳- s.e. ANOVA = $SQRT(MS_w \times \{1/15 + 1/15\}) = SQRT(3.01 \times \{.067 + .067\}) = SQRT(0.40) = 0.64$

۱۴- $t = (۳/۸۷ - ۵/۸۷) / ۰/۶۴ = -۳/۱۲۵$ تجزیه واریانس

۱۵- از آنجایی که مقدار قدر مطلق آماره محاسباتی، $-۳/۱۲۵$ بزرگتر از مقدار بحرانی F یعنی $۱/۹۶$ می‌باشد، فرضیه صفر را رد کرده و فرضیه تحقیق را می‌پذیریم.

۱۶- تبلیغ D بهطور معنی‌داری از تبلیغ B قابل باورتر بوده است ($۵/۸۷$ در مقابل $۳/۸۷$).

پیوست B: آزمون عملی

فصل ۱: آزمون عملی

تصور کنید که مدیر وبستر سابورا^۱ در سنت لوییس از شما خواسته داده‌های نظرسنجی اخیر از مشتریان را تحلیل کنید. سایر ایام آمریکا یک نظرسنجی تجربه خدمت رسانی برای مشتریانی که اخیراً از اداره خدمات خودرو استفاده کرده‌اند، ارسال کرده است. بیایید مهارت‌های اکسل خود را بر روی مورد شماره ۱۰ این نظرسنجی امتحان کنید (به شکل B.1 مراجعه شود).

Mord Shmara 10: "Amtiyazdehi kli shma be kifiyet kar anjam shde br rovi wosileh nqiliye."									
Item #10e: "Your overall rating of the quality of work performed on your vehicle."									
1 Unacceptable	2	3	4	5	6	7	8	9	10 Extraordinary
غیرقابل قبول	RATING	امتیاز دهنی							خارج العاده
8									
5									
6									
5									
4									
8									
7									
7									
8									
6									
7									
5									
4									
8									
7									
5									
7									
5									
6									

شکل B.1: داده صفحه کاری برای آزمون عملی فصل ۱ (مثال عملی)

(a) یک جدول اکسل برای این داده‌ها ایجاد کنید، و سپس از سمت راست اکسل برای یافتن اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین این داده‌ها استفاده کنید. جواب‌های خود را نام‌گذاری کنید و میانگین، انحراف معیار و خطای معیار میانگین این داده‌ها را تا دو رقم اعشار گرد کنید.

(b) فایل را با نام SUBARA8 ذخیره کنید.

¹ Webster Saboura

فصل ۲: آزمون عملی

فرض کنید که می‌خواهید به صورت تصادفی یک مصاحبه شخصی با ۱۲ نفر از ۴۲ فروشنده یک شرکت به عنوان بخشی از "یک نظرسنجی نگرشی شرکت"، انجام دهید.

(a) یک صفحه گسترده چارچوب اعداد برای این فروشنده‌گان ایجاد کنید با عنوان: چارچوب اعداد

(b) سپس، یک ستون جداگانه در سمت راست این چارچوب اعداد ایجاد کنید که این چارچوب اعداد را تکرار می‌کند با عنوان: چارچوب اعداد تکراری

(c) سپس یک ستون جداگانه در سمت راست این چارچوب اعداد تکراری به نام اعداد تصادفی ایجاد کنید، و از تابع $=RAND()$ برای اختصاص اعداد تصادفی به تمام اعداد چارچوب در ستون چارچوب اعداد تکراری استفاده کنید، و قالب ستون را طوری تغییر دهید که تا ۳ رقم اعشار برای هر عدد تصادفی ظاهر شود.

(d) چارچوب اعداد تکراری و اعداد تصادفی را به صورت تصادفی مرتب‌سازی کنید.

(e) نتایج را طوری چاپ کنید که صفحه گسترده در یک صفحه باشد.

(f) روی برگ چاپ شده شماره شناسایی ۱۲ نفر نخست که مایلید در نظرسنجی نگرشی شرکت با آن‌ها مصاحبه کنید را دایره بکشید.

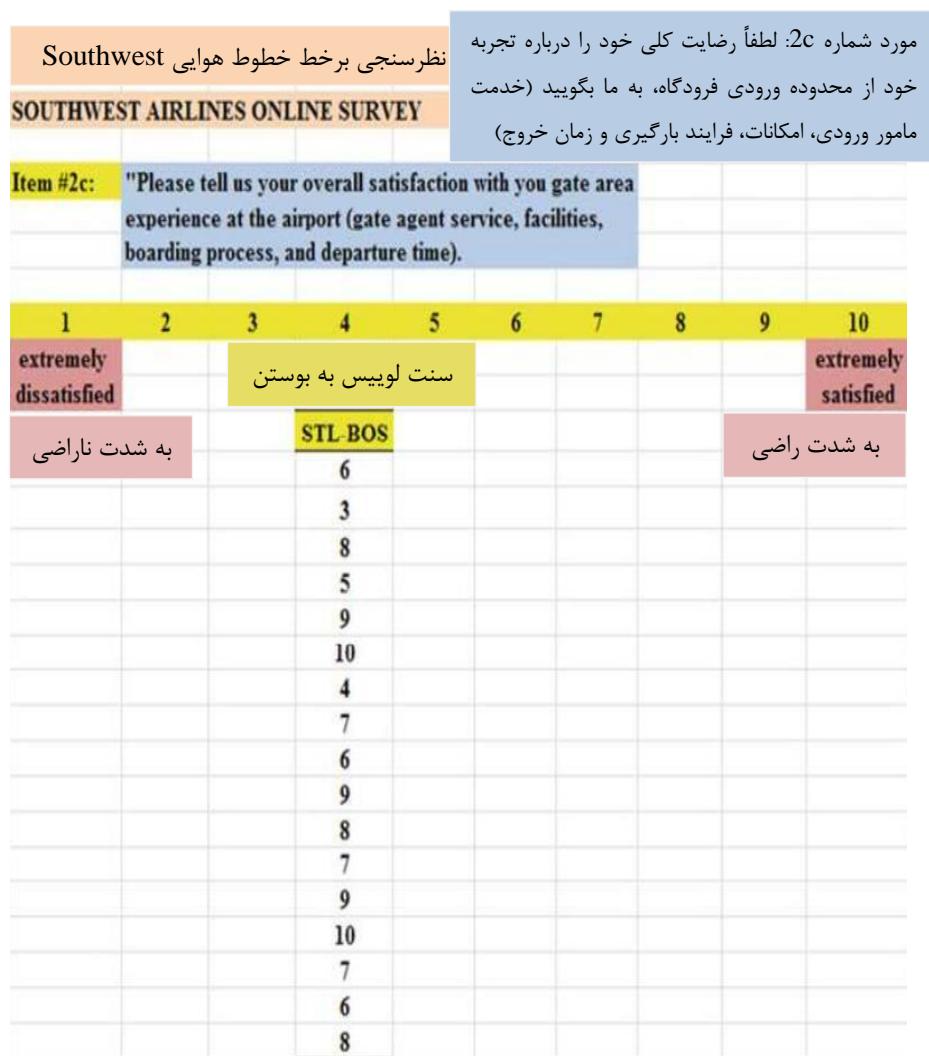
(g) فایل را با نام RAND 15 ذخیره کنید

نکته مهم: توجه داشته باشید هر کسی که این مسئله را انجام دهد اعداد تصادفی متفاوتی در ارتباط با اعداد شماره شناسایی فروشنده‌گان تولید می‌کند زیرا، هر بار که از تابع $=RAND()$ استفاده می‌شود، اکسل اعداد تصادفی متفاوتی را ارائه می‌دهد. به این علت، جوابی که در این راهنمای اکسل ارائه شده است، یک سری کاملاً متفاوت با اعداد تصادفی که شما در سری خود به دست خواهید آورد، حاصل می‌کند. این یک مساله عادی و امری است که انتظار می‌رود.

فصل ۳: آزمون عملی

فرض کنید که از شما خواسته شده، داده‌های پرواز خطوط هوایی شرکت سوس وست^۱ از سنت لوییس به بوستون، را تجزیه و تحلیل کنید. سوس وست یک نظرسنجی رضایتمندی برخط از مشتریان دائمی خود انجام داده است، این نظرسنجی فردای روز پرواز برای آنها ارسال شده و از آن‌ها خواسته شده است که به پرواز خود در یک مقیاس ۱۰ نمره‌ای از ۱ به شدت ناراضی و ۱۰ به شدت راضی، نمره بدنهند. داده‌های حاصل برای مورد شماره ۲۵ این نظرسنجی در شکل B.2 نشان داده شده است.

¹ Southwest



شکل ۲.B.2: داده صفحه کاری برای آزمون عملی فصل ۳ (مثال عملی)

(a) یک جدول اکسل برای این داده‌ها ایجاد کنید، و سپس از سمت راست اکسل برای یافتن اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین این داده‌ها استفاده کنید. جواب‌های خود را نام‌گذاری کنید و میانگین، انحراف معیار و خطای معیار میانگین این داده‌ها را تا دو رقم اعشار گرد کنید.

(b) فرضیه صفر و فرضیه تحقیق را به صورت دستی رو صفحه چاپی خود بنویسید.

(c) ازتابع TINV اکسل برای یافتن فاصله اطمینان ۹۵ درصد برای میانگین این داده‌ها، استفاده کنید. جواب‌های خود را نام‌گذاری کنید. از دو رقم اعشار برای ارقام فاصله اطمینان در قالب عددی استفاده کنید.

(d) در صفحه چاپ شده، بهصورت دستی یک نمودار از فاصله اطمینان ۹۵ درصد که شامل مقدار مرجع هم باشد، رسم کنید.

(e) بر روی صفحه گسترده نتایج را وارد کنید.

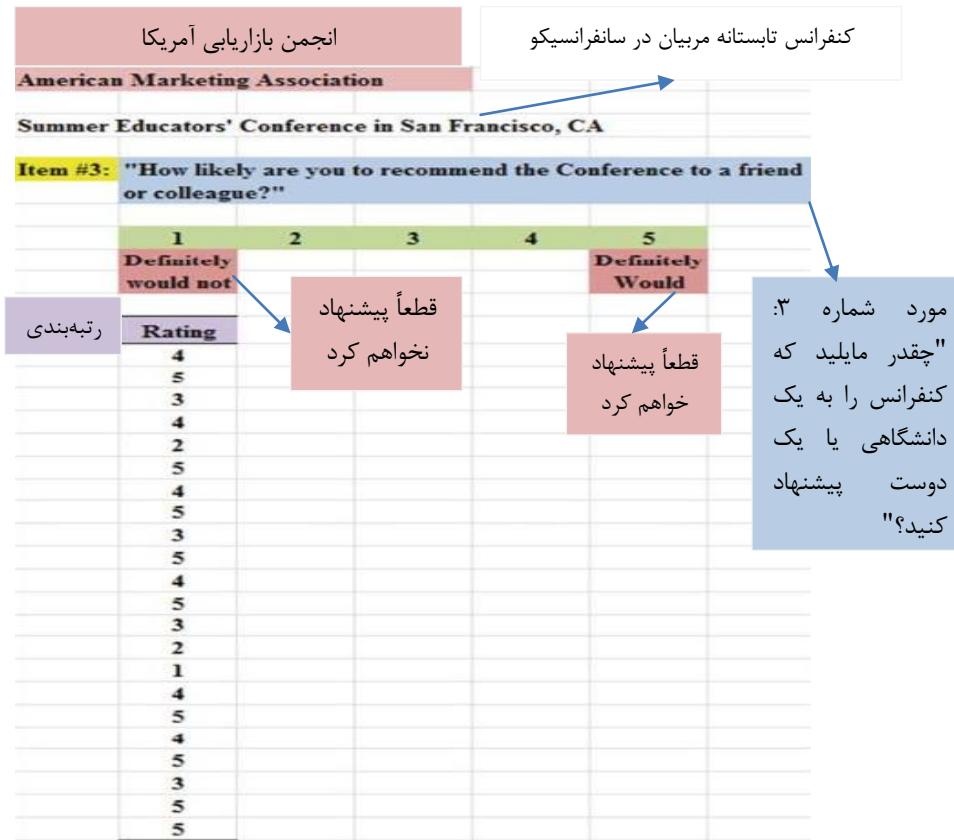
(f) بر روی صفحه گسترده خود نتیجه‌گیری را به زبان ساده بنویسید.

(g) داده‌ها و نتایج را طوری چاپ کنید که در یک صفحه باشند.

(h) فایل را با نام south3 ذخیره کنید

فصل ۴: آزمون عملی

فرض کنید که انجمن بازاریابی آمریکا از شما درخواست کرده تا داده‌های مربوط به کنفرانس تابستانه مربیان در سانفرانسیکو را تجزیه و تحلیل کنید. بهمنظور بررسی فرمول‌های اکسل، قبل از اینکه داده‌های تمام نظرسنجی را مورد به مورد تحلیل کنید، تصمیم گرفتید که داده‌های مربوط به یکی از این پرسش‌ها را تجزیه و تحلیل کنید. کنفرانس از یک مقیاس پنج نمره‌ای ($1 =$ قطعاً مایل تا $=5$ قطعاً مایل نیستم)، استفاده کرده است. نمونه تصادفی از داده‌های فرضی برای یک مورد در شکل B.3 نشان داده شده است.



شکل B.3: داده صفحه کاری برای آزمون عملی فصل ۴ (مثال عملی)

- (a) فرضیه صفر و فرضیه تحقیق را روی صفحه گسترش ده بنویسید.
- (b) برای این داده‌ها صفحه گسترده‌ای ایجاد کرده و سپس از اکسل برای پیدا کردن اندازه نمونه، میانگین، انحراف معیار و خطای معیار میانگین در سمت راست صفحه داده‌ها، استفاده کنید. از فرمت عددی (سه رقم اعشار) برای میانگین، انحراف معیار، و خطای معیار میانگین استفاده کنید.
- (c) t بحرانی را از جدول t در ضمیمه E بر روی صفحه گسترده خود تایپ کرده و نام‌گذاری کنید.
- (d) از اکسل برای محاسبه‌ی ارزش آزمون t برای این داده‌ها (تا سه رقم اعشار) استفاده کنید و بر روی صفحه گسترده نام‌گذاری کنید.
- (e) نتایج را بر روی صفحه گسترده تایپ کرده و سپس نتیجه‌گیری را بنویسید.
- (f) فایل را با نام BOS2ANSWER ذخیره کنید

فصل ۵: آزمون عملی

گروه مالی ماساچوست (۲۰۱۰)، یک تبلیغ تمام صفحه رنگی در مجله وال استریت سفارش داد که در آن یک مدل مرد یک دختر دو ساله را بغل کرده است. تبلیغ یک عنوان و یک زیر عنوان داشت: نشانه یک تصمیم خوب چیست؟

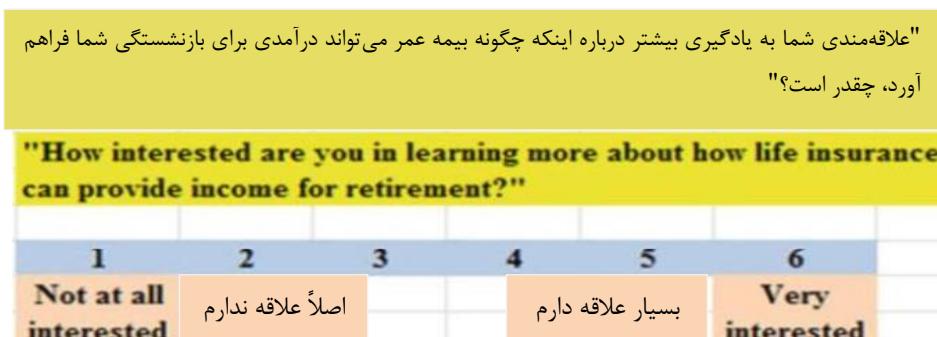
می‌دانید بیمه عمر می‌تواند درآمدی برای بازنشستگی شما فراهم کند و موجب آرامش شما در آن زمان شود.

از آنجایی که بیشتر مشترکان مجله وال استریت مرد هستند، سؤال جالب توجه تحقیق به صورت زیر خواهد بود:

سؤال تحقیق: آیا جنسیت مدل بر تمایل مردان بزرگسال برای فراغیری اینکه بیمه عمر می‌تواند درآمدی برای بازنشستگی فراهم آورد، تاثیرگذار است؟

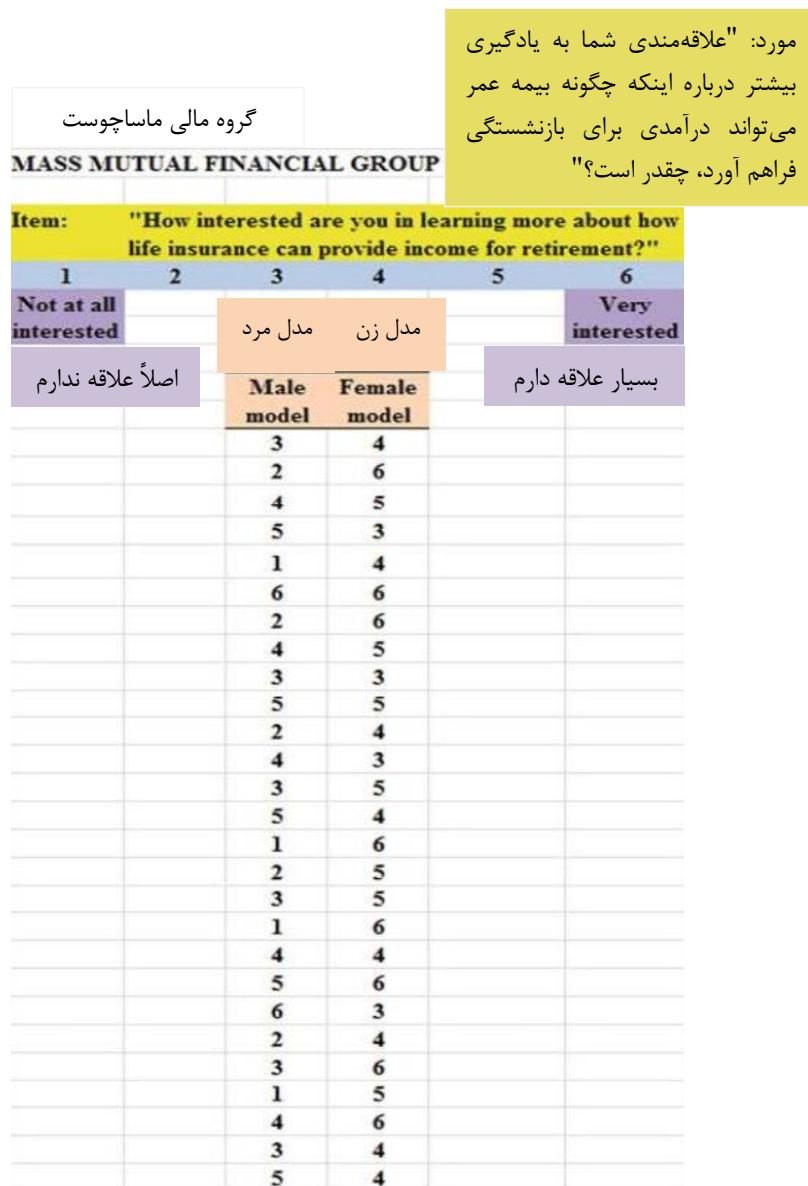
فرض کنید که شما تبلیغ را به دو گروه از مردان بالغ (سنین ۴۴-۲۵) به عنوان مخاطب نشان داده‌اید که یک گروه از مردان تبلیغ را با مدل مرد دیده‌اند در حالی که، گروه دیگر شعار همان تبلیغ را دیده و انتظار مدل زن را در آن داشته‌اند. (شما به طور تصادفی مردان نمونه را به یکی از این دو گروه تجربی اختصاص داده‌اید). دو گروه در طی آزمایش از هم جدا نگه داشته شده‌اند تا هر کدام نتواند بر دیگری اثرگذار باشدند.

بعد از پایان یک ساعت بحث حول محور مدل تبلیغاتی، از پاسخگویان پرسش مطرح شده در شکل B.4، پرسیده شد.



شکل B.4: مورد نظرسنجی برای مدل تبلیغاتی (مثال عملی)

داده‌های حاصل برای این مورد در شکل B.5 نشان داده شده است.



شکل B.5: داده صفحه کاری برای آزمون عملی فصل ۵ (مثال عملی)

- (a) فرضیه صفر و فرضیه تحقیق را بنویسید.
- (b) یک جدول اکسل که این داده‌ها را خلاصه می‌کند، ایجاد کنید.
- (c) از اکسل برای یافتن خطای معیار اختلاف میانگین‌ها استفاده کنید.
- (d) از اکسل برای انجام آزمون t دو گروهی مستقل استفاده کنید. ارزش t که شما به دست آورده‌اید، چقدر است (از دو رقم اعشار استفاده کنید)?
- (e) بر روی صفحه‌گسترده خود مقدار t بحرانی را با استفاده از جدول t در ضمیمه E وارد کنید.
- (f) نتایج آزمون را بر روی صفحه‌گسترده خود وارد کنید.
- (g) نتیجه‌گیری خود را به زبان ساده روی صفحه‌گسترده خود بنویسید.
- (h) فایل را با نام lifeinsur3 ذخیره کنید.
- (i) صفحه‌گسترده نهایی را به‌طوری‌که در یک صفحه باشد چاپ کنید.

فصل ۶: آزمون عملی

فرض کنید که شما می‌خواهید رابطه‌ی بین رژیم (با استفاده از کالری مجاز به ازای هر روز اندازه-گیری می‌شود) و کاهش وزن (به کیلوگرم اندازه-گیری می‌شود، kg) را برای زنان بالغ (سنین ۳۰ و ۴۰) مطالعه کنید که آن‌ها نسبت به قد و ساختار بدن‌شان، اضافه وزن دارند و همه قبل از انجام برنامه کاهش وزن تقریباً همان کیلوگرم وزن مشابه دارند. شما قصد دارید تا مهارت‌های اکسل خود را بر روی نمونه تصادفی از این زن‌ها بر اساس تغییر وزن آن‌ها در طول ۴ ماه گذشته امتحان کنید تا مطمئن شوید که می‌توانید این نوع پژوهش را انجام دهید. داده‌های فرضی در شکل B.6 نشان داده شده است:

رابطه بین رژیم و کاهش وزن		
RELATIONSHIP BETWEEN DIET AND WEIGHT LOSS		
ADULT WOMEN AGES 30-40		زنان بالغ سنین ۳۰ - ۴۰
DIET (calories allowed per day)	WEIGHT LOSS (kg)	کاهش وزن (کیلوگرم)
رژیم (کالری مجاز به ازای هر روز)	900	16.0
	1050	12.0
	1150	8.0
	1275	6.0
	1420	3.0
	1530	5.5
	1610	9.5
	1710	2.5
	1820	6.0
	1875	9.0
	1930	6.0
	2100	3.0

شکل B.6: داده صفحه کاری برای آزمون عملی فصل ۶ (مثال عملی)

یک صفحه‌گسترده اکسل ایجاد کرده و داده‌ها را وارد کنید، از رژیم (کالری مجاز در هر روز) به عنوان متغیر مستقل (توضیحی) و کاهش وزن (کیلوگرم) به عنوان متغیر وابسته (معیار) استفاده کنید. در زیر جدول، از تابع $=correl$ برای یافتن همبستگی بین این دو متغیر استفاده کنید. همبستگی را نام‌گذاری کرده و آن را در زیر جدول قرار دهید؛ سپس همبستگی را تا دو رقم اعشار گرد کنید.

(a) نمودار پراکنش XY را برای این دو دسته داده ایجاد کنید بهطوری که:

- عنوان بالا: رابطه بین رژیم و کاهش وزن
- عنوان محور x: رژیم (کالری مجاز در هر روز)
- عنوان محور y: کاهش وزن (کیلوگرم)
- نمودار را به زیر جدول انتقال دهید

اندازه نمودار را تغییر دهید به نحوی که ۸ ستون عرض و ۲۵ ردیف طول داشته باشد

(b) خط رگرسیون حداقل مربعات را برای این داده‌ها بر روی نمودار پراکنش ایجاد کنید و معادله رگرسیون را به نمودار اضافه کنید.

(c) از اکسل برای اجرای آماره‌های رگرسیون به منظور یافتن معادله خط رگرسیون حداقل مربعات برای این داده‌ها و نمایش نتایج زیر نمودار در صفحه گسترده، استفاده کنید. از قالب عددی دو رقم اعشار برای همبستگی و سه رقم اعشار برای تمام ارقام اعشاری، شامل ضرایب استفاده کنید.

(d) تنها داده‌های ورودی و نمودار را به طوری که در یک صفحه باشند، چاپ کنید. سپس، رگرسیون خروجی را در صفحه‌ای جداگانه به تنها یک صفحه باشند، چاپ کنید.

(e) فایل را با عنوان DIET3 ذخیره کنید.

سوالات زیر را با استفاده از برگه چاپی خود اکسل خود پاسخ دهید:

۱. همبستگی بین رژیم و کاهش وزن چقدر است؟
۲. عرض از مبدأ y چقدر است؟
۳. شبیه خط چقدر است؟
۴. معادله رگرسیون چیست؟

۵. از معادله رگرسیون برای پیش‌بینی کاهش وزن مورد انتظار برای زنی که رژیم ۱۵۰۰ کالری مجاز در روز را اجرا کرده است، استفاده کنید. کار خود را در صفحه‌ای جداگانه نشان دهید.

فصل ۷: آزمون عملی

امتیاز دادن به عملکرد یک مدیر در یک سازمان به طور مداوم یک پایه برای فرصت‌های ارتقای مدیران، درک ارزش برای سازمان، و گاهی حتی معیاری برای افزایش حقوق مدیر خواهد بود. فرض کنید که می‌خواهید رابطه‌ی بین تعداد سال‌های تجربه کسب‌وکار مربوط به یک مدیر، تعداد مدارک دانشگاهی یا غیردانشگاهی که توسط آن مدیر به دست آمده‌اند، و رتبه‌دهی به عملکرد مدیر (در مقیاسی که ۱= ضعیف و ۷= عالی رتبه داده شده است) را در یک شرکت بزرگ با فناوری بالا

مطالعه کنید. شما قصد دارید که مهارت‌های اکسل خود را بر روی یک نمونه کوچک از مدیران سطح متوسط در شرکت را برای مطالعه‌ی این رابطه، آزمون کنید. داده‌های فرضی در شکل 7 نشان داده شده است.

سوال پژوهش		آیا تجربه و تحصیلات پیش‌بینی‌کننده‌های خوبی برای عملکرد هستند؟	
Research question:		"Are experience and education good predictors of performance?"	
PERFORMANCE RATING		EXPERIENCE	NO. DEGREES
امتیازدهی	7	20	3
عملکرد	6	15	2
	4	8	2
	1	5	0
	2	6	1
	6	18	3
	5	6	2
	7	10	3
	4	11	2
	5	12	3
	4	8	4
	6	14	3
	5	9	2

شکل B.7: داده صفحه کاری برای آزمون عملی فصل ۷ (مثال عملی)

(a) یک صفحه‌گسترده اکسل با استفاده از امتیاز عملکرد به عنوان متغیر معیار، و تعداد سال‌های تجربه کاری مرتبط و تعداد مدرک‌های دانشگاهی و غیردانشگاهی به دست‌آمده توسط مدیر به عنوان متغیرهای پیش‌بینی‌کننده ایجاد کنید.

(b) فایل را با عنوان Performance2 ذخیره کنید.

(c)تابع رگرسیون چندگانه اکسل را برای پیدا کردن رابطه بین این سه متغیر به کار بگیرید و خروجی خلاصه را در زیر جدول قرار دهید.

(d) از قالب عددی دو رقم اعشار برای همبستگی چندگانه، و چهار رقم اعشار برای عرض از مبداء y ، ضرایب تجربه و تعداد مدارک بر روی خروجی خلاصه استفاده کنید. از قالب عددی سه رقم اعشار برای سایر ارقام اعشاری در خروجی خلاصه استفاده کنید.

(e) جدول نتایج رگرسیون که زیر جدول داده‌ها است را طوری چاپ کنید که در یک صفحه قرار گیرد.

با استفاده از صفحه چاپی خروجی اکسل خود، به سوالات زیر پاسخ دهید:

۱. همبستگی چندگانه R_{xy} چقدر است؟
 ۲. مقدار عرض از مبدأ y , a , چقدر است؟
 ۳. مقدار ضریب تجربه b_1 چقدر است؟
 ۴. مقدار ضریب تعداد مدارک b_2 چقدر است؟
 ۵. معادله رگرسیون چندگانه چیست؟
 ۶. امتیاز عملکرد را که شما از مدیر با ۱۰ سال تجربه شغلی مرتبط و سه مدرک دانشگاهی و غیر دانشگاهی انتظار دارید را پیش‌بینی کنید.
- (f) اکنون به فایل اکسل خود برگشته و یک ماتریس همبستگی برای این سه متغیر ایجاد کرده و در زیر خلاصه خروجی در صفحه گسترده قرار دهید.
- (g) این فایل را با عنوان Performance3 ذخیره کنید.
- (h) اکنون فقط ماتریس همبستگی را بر روی یک صفحه جداگانه چاپ کنید.
۷. همبستگی بین تجربه و امتیاز عملکرد چقدر است؟
 ۸. همبستگی بین تعداد مدارک و امتیاز عملکرد چقدر است؟
 ۹. همبستگی بین تجربه و تعداد مدارک چقدر است؟
 ۱۰. بحث کنید که کدام‌یک از این دو پیش‌بینی‌کننده، امتیاز عملکرد را بهتر پیش‌بینی می‌کنند.
 ۱۱. توضیح دهید که ترکیب دو متغیر پیش‌بینی کننده در مقایسه با یک پیش‌بینی کننده منفرد چقدر امتیاز عملکرد را بهتر پیش‌بینی می‌کنند.

فصل ۸: آزمون عملی

فرض کنید که در واحد تحقیق و توسعه شرکت پورینا^۱ در سنت لوییس کار می‌کنید و از شما خواسته شده است تا چهار طعم مختلف غذای گربه را به منظور اینکه کدام طعم بیشترین استفاده توسط گربه‌ها را داشته، آزمون کنید. همچنین، فرض کنید که گربه‌ها بر اساس سن، جنسیت و گونه مرتب و به طور تصادفی به چهار گروه تقسیم شده‌اند. نتایج حاصل از مقدار غذای مصرف شده توسط گربه‌ها در قالب داده‌های فرضی در شکل B.8 نشان داده شده است. از شما خواسته شده است که تعیین کنید آیا تفاوت معنی‌داری بین مقدار غذای مصرف شده توسط گربه‌ها در این چهار گروه وجود دارد یا خیر؟

طعم‌های مختلف غذاهای جدید گربه			
FLAVORS OF NEW KITTEN FOOD			
A	B	C	D
12	23	29	38
14	20	27	33
18	17	30	40
11	23	35	34
19	20	33	34
10	28	34	37
17	25	32	43
19	22	35	38
23	28	40	45
16	25	38	39
24			39
15			42

شکل B.8: داده صفحه کاری برای آزمون عملی فصل ۸ (مثال عملی)

- (a) این داده‌ها را در یک صفحه گسترده اکسل وارد کنید.
- (b) بر روی صفحه گسترده، فرضیه‌های صفر و تحقیق را بنویسید.
- (c) تجزیه واریانس یک عامله را برای این داده‌ها اجرا کرده، و جدول تجزیه واریانس حاصل را در زیر جدول داده‌ها برای این چهار نوع غذای گربه، قرار دهید.

¹ Purina

- (d) اگر مقدار F در جدول تجزیه واریانس معنی‌دار بود، یک رابطه اکسل برای محاسبه آزمون t تجزیه واریانس برای مقایسه مقدار غذای مصرف شده در گروه B در مقابل مقدار غذای مصرف شده در گروه D ایجاد کرده، و نتیجه را در زیر جدول تجزیه واریانس در صفحه‌گسترده نمایش دهید (خطای معیار و مقدار آزمون t تجزیه واریانس را در خطوط جداگانه بر روی صفحه‌گسترده قرار داده، و از دو رقم اعشار برای همه مقادیر استفاده کنید)
- (e) نتایج صفحه‌گسترده را طوری چاپ کنید که تمام اطلاعات در یک صفحه قرار گیرد
- (f) بر روی صفحه‌چاپ شده به صورت دستی میانگین مربعات (MS) بین گروه‌ها و میانگین مربعات درون گروه‌ها را نام‌گذاری کنید.
- (g) دور مقدار F بر روی برگه چاپی خود برای تجزیه واریانس داده‌های ورودی، خط کشیده و آن را نام‌گذاری کنید.
- (h) میانگین گروه B و میانگین گروه D را که توسط رابطه تجزیه واریانس تولید شده بود را بر روی صفحه چاپی خود به صورت دستی نام‌گذاری کنید.
- صفحه‌گسترده را با عنوان kitten² ذخیره کنید
- اکنون در یک صفحه جدالگانه ، موارد ذیل را به صورت دستی انجام دهید:
- (i) مقدار بحرانی F را برای تجزیه واریانس یک عاملهای که ایجاد کرده‌اید، بیابید.
- (j) خلاصه‌ای از نتایج آزمون تجزیه واریانس را برای این داده‌های ورودی، بنویسید.
- (k) خلاصه‌ای از نتیجه‌گیری آزمون تجزیه واریانس را به زبان ساده بنویسید.
- (l) فرضیه صفر و فرضیه تحقیق را برای مقایسه گروه B در مقابل D بنویسید.
- (m) مقدار درجات آزادی برای آزمون تجزیه واریانس چهار طعم را به صورت دستی محاسبه کنید.
- (n) مقدار بحرانی t را برای آزمون تجزیه واریانس با استفاده از جدول ضمیمه E، بنویسید.
- (o) خلاصه نتایج آزمون t تجزیه واریانس را بنویسید.
- (p) خلاصه نتیجه‌گیری آزمون t تجزیه واریانس را به زبان ساده بنویسید.

منابع

Mass Mutual Financial Group. What is the Sign of a Good Decision? (Advertisement) The Wall Street Journal, September 29, 2010, p. A22.

پیوست C: پاسخ آزمون عملی

پاسخ آزمون عملی: فصل ۱ (به شکل C.1 مراجعه شود)



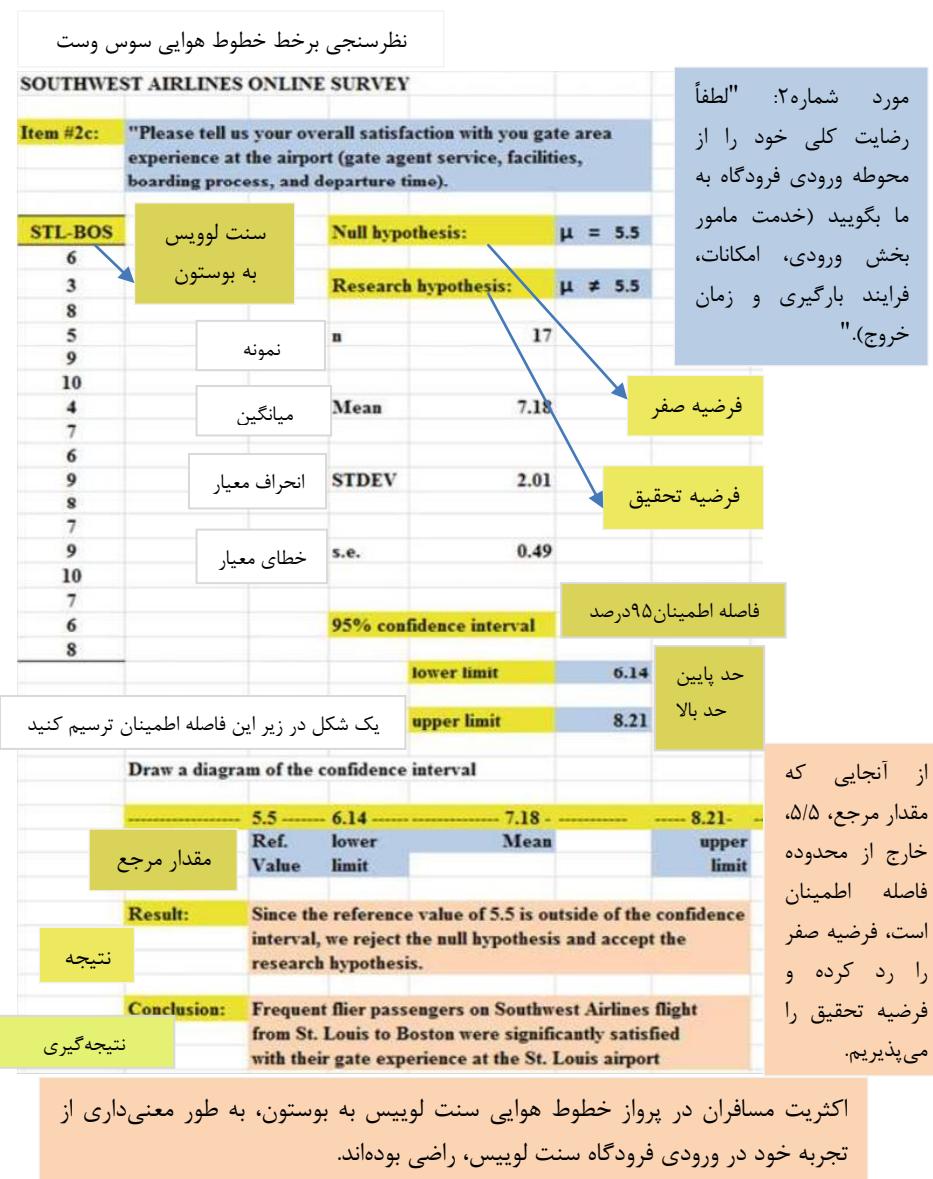
شکل C.1: پاسخ آزمون عملی فصل ۱

پاسخ آزمون عملی: فصل ۲ (به شکل C.2 مراجعه شود)

اعداد چارچوب	اعداد چارچوب تکراری	اعداد تصادفی
FRAME NUMBERS	Duplicate frame numbers	RAND NO.
1	8	0.871
2	22	0.309
3	31	0.658
4	42	0.443
5	4	0.489
6	29	0.370
7	3	0.064
8	21	0.440
9	37	0.026
10	17	0.922
11	34	0.980
12	25	0.930
13	10	0.138
14	41	0.504
15	30	0.884
16	36	0.789
17	13	0.243
18	15	0.250
19	20	0.343
20	14	0.958
21	9	0.779
22	12	0.147
23	38	0.253
24	26	0.476
25	1	0.865
26	5	0.170
27	35	0.410
28	28	0.325
29	24	0.216
30	32	0.439
31	27	0.138
32	19	0.168
33	6	0.326
34	39	0.373
35	2	0.454
36	18	0.777
37	7	0.631
38	11	0.448
39	16	0.412
40	40	0.391
41	33	0.471
42	23	0.865

شکل C.2: پاسخ آزمون عملی فصل ۲

پاسخ آزمون عملی: فصل ۳ (به شکل C.3 مراجعه شود)



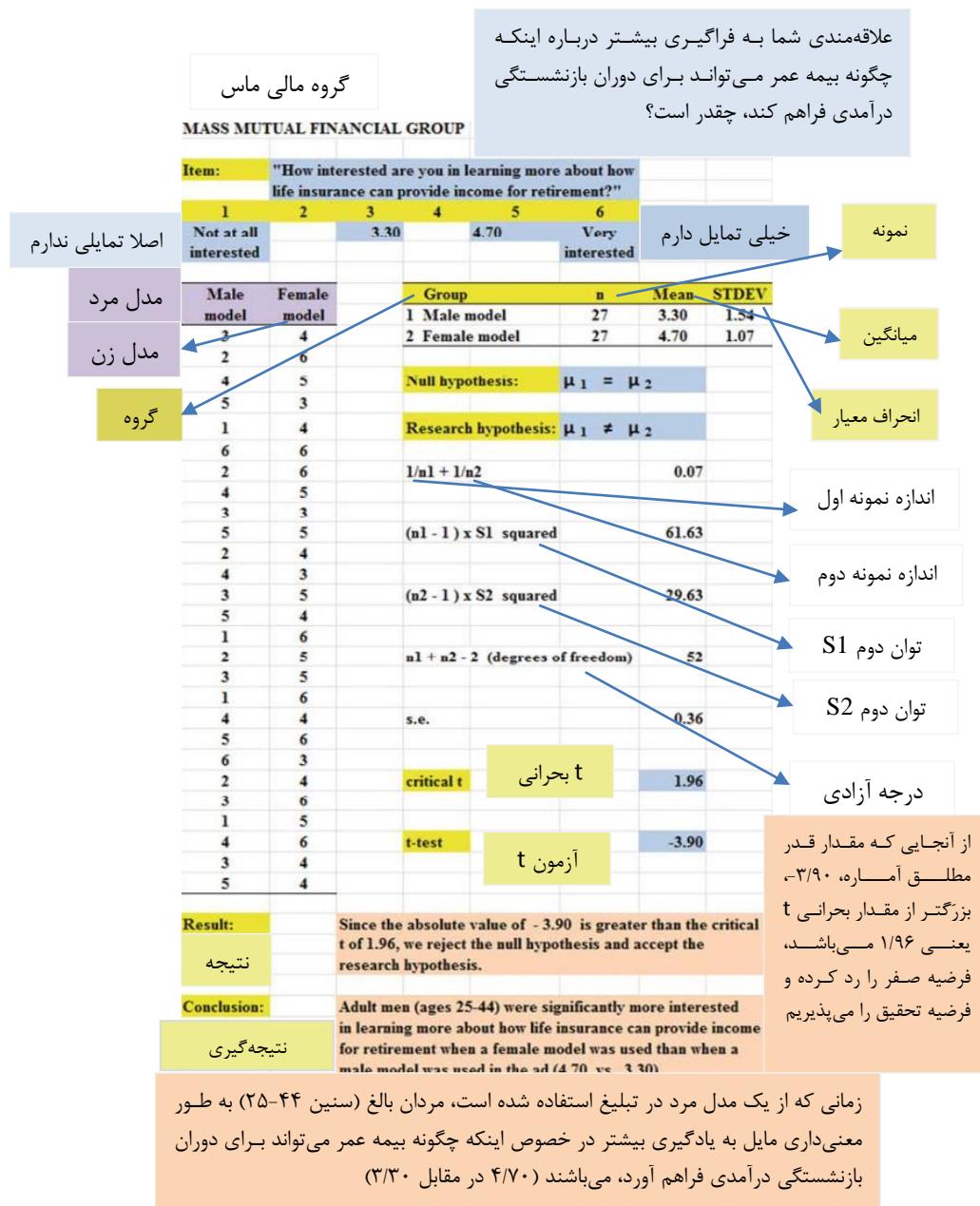
شکل C.3: پاسخ آزمون عملی فصل ۳

پاسخ آزمون عملی: فصل ۴ (به شکل C.4 مراجعه شود)



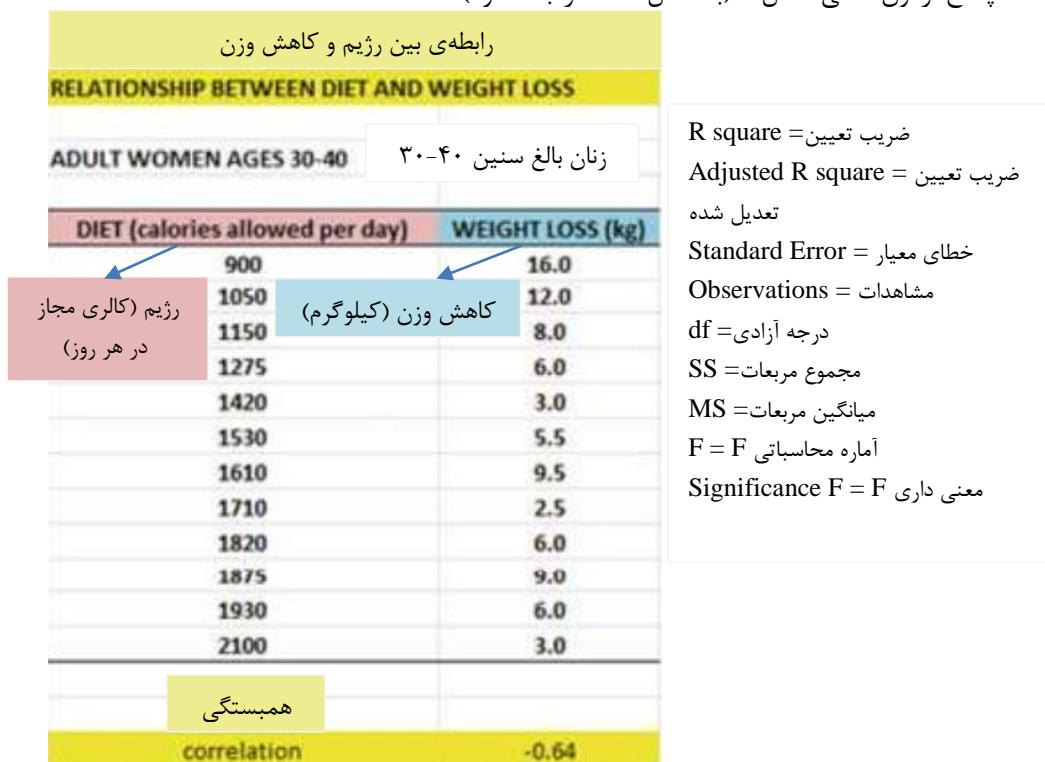
شکل C.4: پاسخ آزمون عملی فصل ۴

پاسخ آزمون عملی: فصل ۵ (به شکل C.5 مراجعه شود)



شکل C.5: پاسخ آزمون عملی فصل ۵

پاسخ آزمون عملی: فصل ۶ (به شکل C.6 مراجعه شود)



R square = ضریب تعیین

Adjusted R square = ضریب تعیین

تعدیل شده

خطای معیار = خطای معيار

Observations = مشاهدات

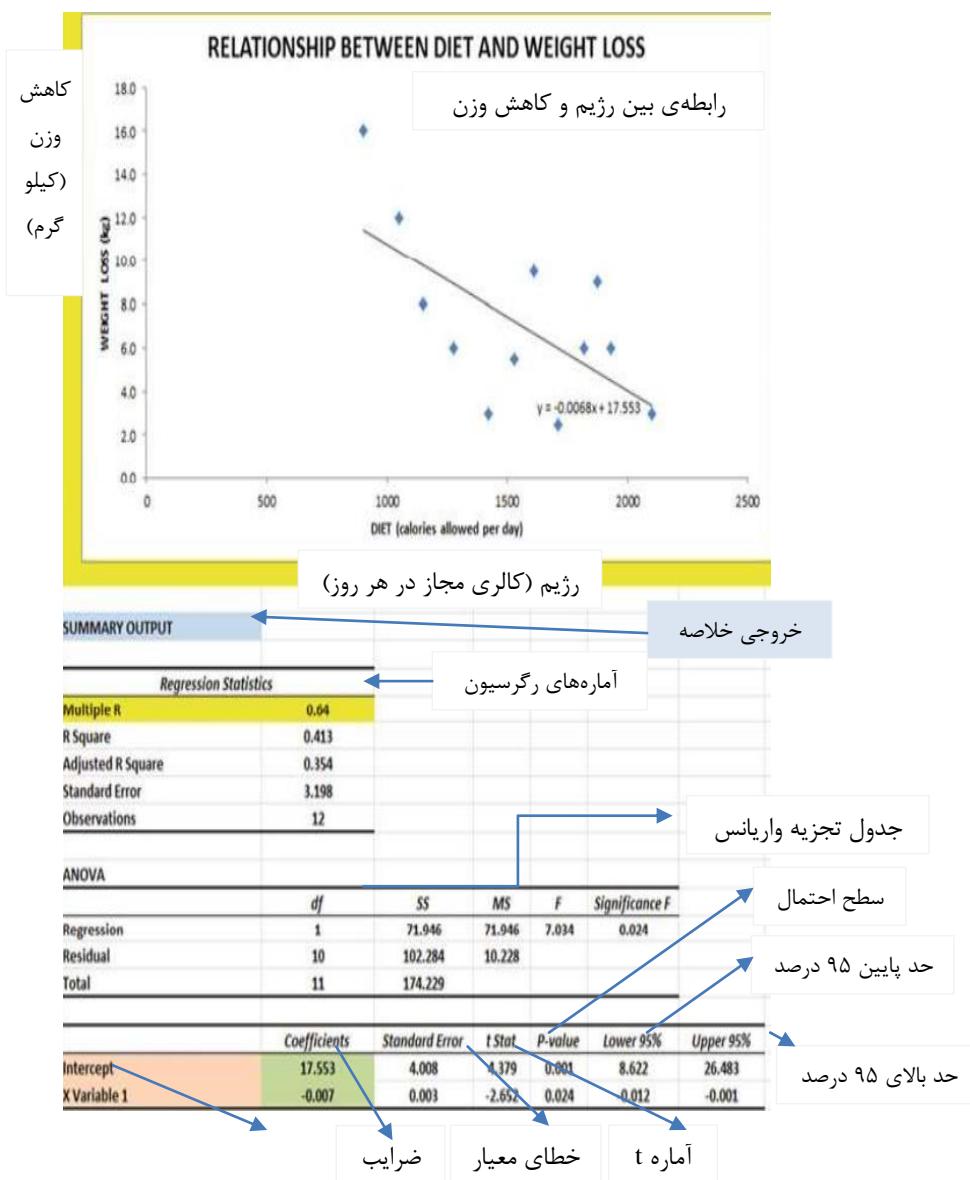
df = درجه آزادی

SS = مجموع مربعات

MS = میانگین مربعات

F = F آماره محاسباتی

Significance F = F معنی داری



شکل C.6: پاسخ آزمون عملی فصل ۶

پاسخ آزمون عملی: فصل ۶ (ادامه)

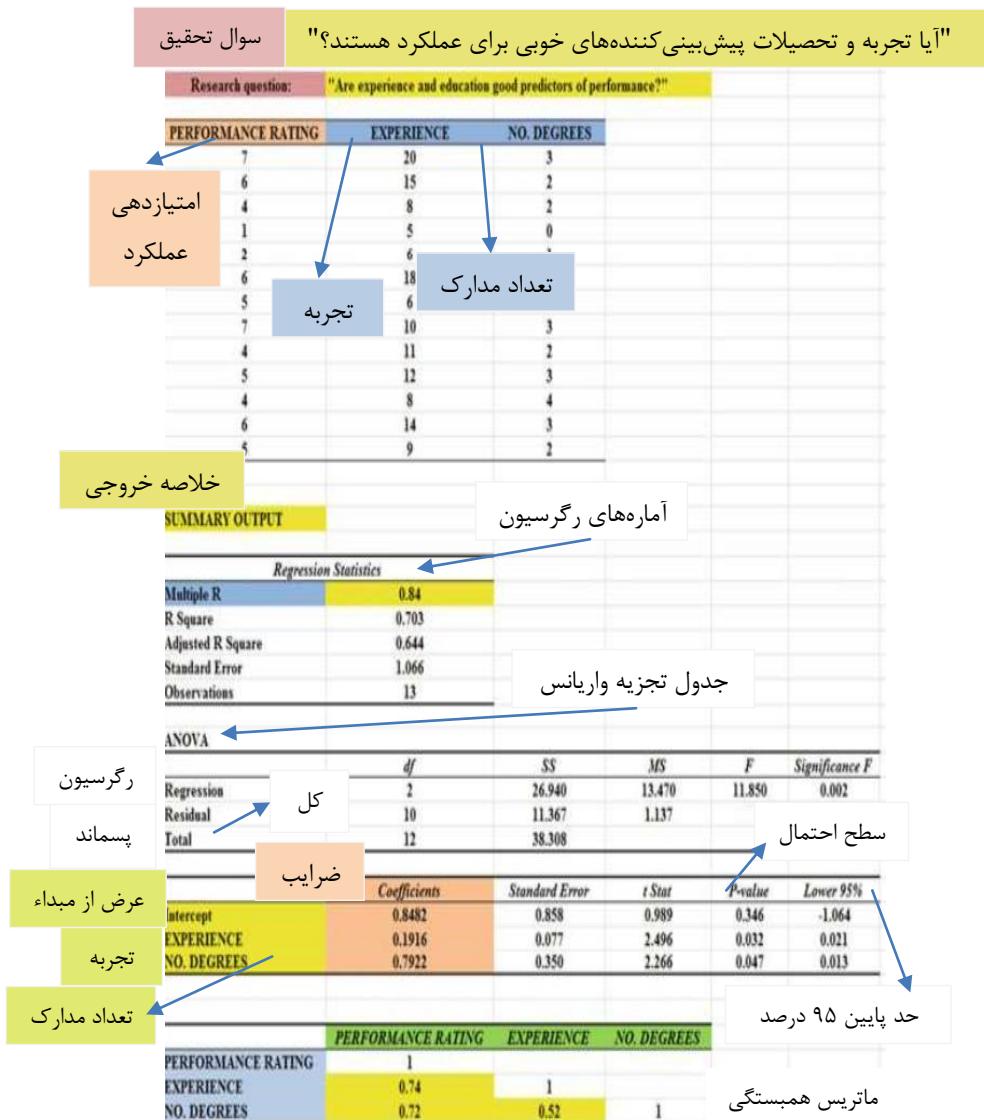
۱. $t = -0.64 / 0.0068 = -94$ (به همبستگی منفی توجه کنید!)۲. عرض از مبداء $y = a = 17.553$

۳. شیب $b = -0.007$ (به شیب منفی توجه کنید که بیان می‌کند همبستگی منفی است!)

$$\begin{aligned} 4. \quad Y &= a + b X \\ Y &= 17.553 - 0.007 X \\ Y &= 17.553 - 0.007 (1500) \\ Y &= 17.553 - 10.5 \end{aligned}$$

$Y = 7.1$ کیلوگرم کاهش وزن

پاسخ آزمون عملی: فصل ۷ (به شکل C.7 مراجعه شود)



شکل C.7: پاسخ آزمون عملی فصل ۷

پاسخ آزمون عملی: فصل ۷ (ادامه)

۱- همبستگی چندگانه = $0/84$

۲- عرض از مبدأ = $y = a - 2/8482$

۳- $b_1 = 1916/0$

۴- $b_2 = 7922/0$

$$5- Y = a + b_1 X_1 + b_2 X_2$$

$$Y = 0.8482 + 0.1916 X_1 + 0.7922 X_2$$

$$6- Y = 0.8482 + 0.1916 (10) + 0.7922 (3)$$

$$Y = 0.8482 + 1.916 + 2.377$$

$$Y = 5$$

+۰/۷۴ -۷

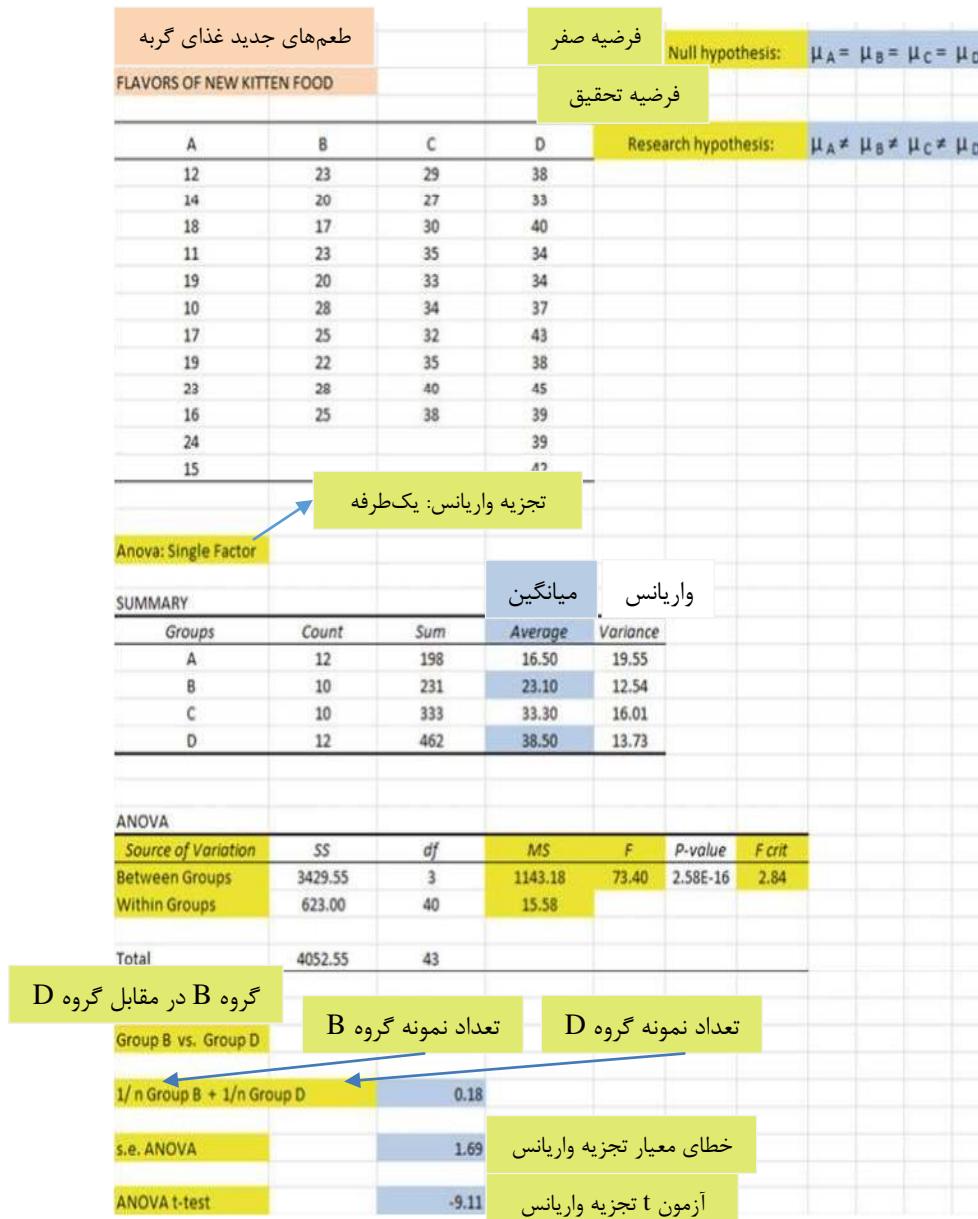
+۰/۷۲ -۸

+۰/۵۲ -۹

۱۰- متغیر تجربه، پیش‌بینی کننده بهتری برای امتیاز عملکرد، بوده است ($r = 0/74$)

۱۱- ترکیب دو پیش‌بینی کننده، امتیاز عملکرد را بهتر پیش‌بینی کردند ($R_{xy} = 0/84$)

پاسخ آزمون عملی: فصل ۸ (به شکل C.8 مراجعه شود)



شکل C.8: پاسخ آزمون عملی فصل ۸

پاسخ آزمون عملی: فصل ۸ (ادامه)

f. $MS_b = 1143.18$ و $MS_W = 15.58$

g. $F = 73/40$

h. میانگین گروه $B = 23/10 = 38/50$ و میانگین گروه $D = 23/10 = 38/50$

i. مقدار F بحرانی $= 2/84$

j. نتیجه: از آنجایی که مقدار آماره محاسباتی، $73/40$ ، از مقدار بحرانی F یعنی $2/84$ بزرگتر است، فرضیه صفر را رد کرده و فرضیه تحقیق را می‌پذیریم.

k. نتیجه‌گیری: تفاوت معنی‌داری بین مقدار غذای مصرف شده توسط گربه‌ها در چهار گروه طعم غذای گربه وجود دارد.

l. فرضیه صفر: $\mu_D = \mu_B$

فرضیه تحقیق: $\mu_B \neq \mu_D$

m. $df = n_{TOTAL} - k = 44 - 4 = 40$

n. بحرانی $t = 1/69$

o. نتیجه: از آنجایی که مقدار قدر مطلق آماره محاسباتی، $9/11$ ، بزرگتر از مقدار t بحرانی یعنی $1/69$ می‌باشد، فرضیه صفر را رد کرده و فرضیه تحقیق را می‌پذیریم.

p. نتیجه‌گیری: گربه‌ها به طور معنی‌داری طعم D را در مقایسه با طعم B ترجیح داده‌اند ($38/50$) در مقابل ($23/10$)

پیوست D: رابطه‌های آماری

$$\bar{X} = \frac{\sum x}{n}$$

میانگین

$$STDEV = S = \sqrt{\frac{\sum(X - \bar{X})^2}{n - 1}}$$

انحراف معیار

$$s.e. = S_{\bar{X}} = \frac{S}{\sqrt{n}}$$

خطای معیار میانگین

$$\bar{X} \pm t S_{\bar{X}}$$

فاصله اطمینان میانگین

$$S_{\bar{X}} = \frac{S}{\sqrt{n}} \quad \text{که در آن}$$

$$t = \frac{\bar{X} - \mu}{S_{\bar{X}}}$$

آزمون t یک گروهی

$$S_{\bar{X}} = \frac{S}{\sqrt{n}} \quad \text{که در آن}$$

آزمون t دو گروهی

(الف) زمانی که هر دو گروه اندازه نمونه‌ای بیشتر از ۳۰ دارند

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}}$$

که در آن

$$S_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

و در آن

$$df = n_1 + n_2 - 2$$

(ب) زمانی که یک یا هر دو گروه اندازه نمونه‌ای کمتر از ۳۰ داشته باشند

$$t = \frac{\bar{X}_1 - \bar{X}_2}{S_{\bar{X}_1 - \bar{X}_2}}$$

که در آن

$$S_{\bar{X}_1 - \bar{X}_2} = \sqrt{\frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

و در آن

$$df = n_1 + n_2 - 2$$

$$r = \frac{1}{n-1} \frac{\sum (X - \bar{X})(Y - \bar{Y})}{S_X S_Y}$$

همبستگی

که در آن

X = انحراف معیار S_X

Y = انحراف معیار S_Y

$$Y = a + b X$$

رگرسیون خطی ساده

که در آن

y = عرض از مبدأ a

و

b = شیب خط

$$Y = a + b_1 X_1 + b_2 X_2 + b_3 X_3 + \dots$$

معادله رگرسیون چندگانه

که در آن

y = عرض از مبدأ a

$$F = MS_b / MS_w$$

آزمون F در تجزیه واریانس یک طرفه

$$ANOVA t = \frac{\bar{X}_1 - \bar{X}_2}{S.e. ANOVA}$$

آزمون t در تجزیه واریانس

که در آن

$$S.e. ANOVA = \sqrt{MS_w \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}$$

و در آن

$$df = n_{TOTAL} - k$$

$$n_{TOTAL} = n_1 + n_2 + n_3 + \dots$$

و در آن

k = تعداد گروهها

پیوست E: جدول t

مقادیر t بحرانی لازم برای رد کردن فرضیه صفر (به شکل E.1 مراجعه شود)

(n) اندازه نمونه	degrees of freedom df	درجات آزادی	بحرانی t
		critical t	
10	9	2.262	
11	10	2.228	
12	11	2.201	
13	12	2.179	
14	13	2.160	
15	14	2.145	
16	15	2.131	
17	16	2.120	
18	17	2.110	
19	18	2.101	
20	19	2.093	
21	20	2.086	
22	21	2.080	
23	22	2.074	
24	23	2.069	
25	24	2.064	
26	25	2.060	
27	26	2.056	
28	27	2.052	
29	28	2.048	
30	29	2.045	
31	30	2.042	
32	31	2.040	
33	32	2.037	
34	33	2.035	
35	34	2.032	
36	35	2.030	
37	36	2.028	
38	37	2.026	
39	38	2.024	
40	39	2.023	
بی‌نهایت	infinity	infinity	1.960

شکل E.1 مقادیر t بحرانی لازم برای رد کردن فرضیه صفر

University of Guilan Press

Excel 2013

**For Business Statistics a Guide
to Solving Practical Problems**

By:

Thomas J. Quirk

Translated by:

Mohammad Kavoosi-Kalashami, Ph. D