



اکسل

Excel

(مقدمه، تحلیل واریانس، رگرسیون و آزمون میانگین‌ها)

*

*

*

تهیه و گردآوری:

حمید مسعودی

تابستان 1389

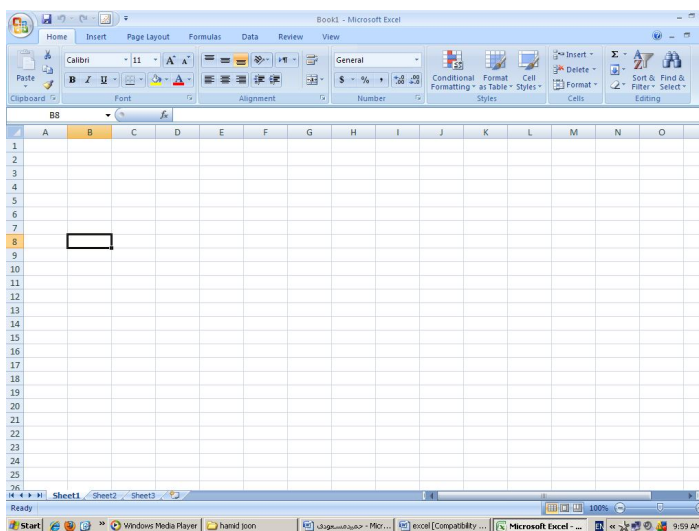
massoudihamid@gmail.com

اکسل: مقدمه، تحلیل واریانس، رگرسیون و آزمون میانگینها.....1

بخش اول:

مقدمه ای در باره ی برنامه اکسل (Excel):

نحوه راه اندازی اکسل: از منوی `start`، `Allprograms`، `excel`، `microsoft` را انتخاب کنید.
اکسل راه اندازی شده و پوشه کار جدید با سه کاربرگ آماده، نمایش داده می شود. که اکسل کاربرگ ها را `sheet1`، `sheet2`، `sheet3` نام گذاری می کند و برای دستیابی به هر کاربرگ فقط روی نام کاربرگ کلیک کنید.



نام قسمت های مهم کاربرگ:

خانه ها: باکس های کوچکی که کاربرگ را ایجاد می کنند.

خانه فعال: خانه ای که مکان نما در آن قرار دارد

ستون: خط عمودی خانه ها از بالا به پایین کاربرگ که از A تا ... نام گذاری شده اند.

ردیف: خط افقی خانه ها که از چپ به راست کاربرگ کشیده شده و از 1 تا 65536 شماره گذاری شده اند.

مرجع خانه: محل یا نشانی خانه در یک کاربرگ که از دو قسمت تشکیل شده است (حروف

ستون، عدد ردیف) بطور مثال `a1`، `b10`، `w90`، ...

Name box: ناحیه ای مستطیلی در بالای سمت چپ کاربرگ که مرجع خانه فعال در آنجا

قابل مشاهده است و با تایپ کردن نام خانه مورد نظر برای فعال نمودن می توان آن را فعال نمود.

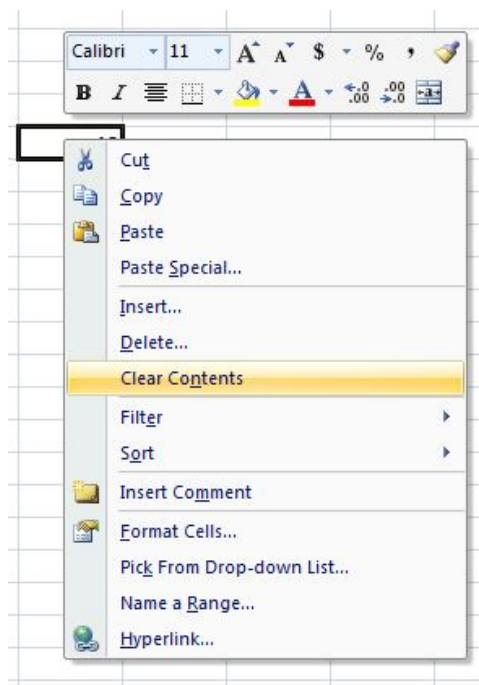
اکسل: مقدمه، تحلیل واریانس، رگرسیون و آزمون میانگینها.....2

وارد کردن عدد یا متن در خانه ها: شما با کلیک بر روی یک خانه (B1) آن را فعال نموده و عدد یا متن مورد نظر را در آنجا تایپ نمایید سپس دکمه Enter را زده مکان نما در خانه پایینی قرار می گیرد البته راه دیگر انتقال مکان نما به خانه های اطراف استفاده از کلیدهای فلشی و یا ماوس می باشد.

ویرایش خانه ها: برای ویرایش متن یا عدد تایپ شده در هر خانه به دو روش می توان عمل کرد: **الف)** روی خانه مورد نظر دو بار کلیک کرده یک مکان نمای چشمک زن در خانه مشاهده می شود که با جابجایی آن توسط کلیدهای حرکتی در مکان مناسب ویرایش را انجام دهید. **ب)** با کلیک بر روی خانه آن را فعال کرده و سپس دکمه F2 را فشار دهید خانه آماده ویرایش است.

حذف اطلاعات خانه ها:

روش اول: روی خانه مورد نظر کلیک کرده و کلید Delete یا Backspace را فشار دهید.
روش دوم: روی خانه کلیک راست کرده و گزینه clear contents را انتخاب و دستور delete را اجرا نمایید.



توجه: زمان وارد کردن اطلاعات اکسل فقط بعد از فشار دادن کلید Enter آن را اعمال می کند اما برای حذف آن نیازی به زدن کلید فوق نیست.

در اکسل دو نوار ابزار به طور پیش فرض قابل مشاهده است:

1- نوار استاندارد: که شامل دکمه هایی برای تغییر ظاهر متن و عدد خانه ها می باشد.

2-نوار فرمت: که شامل دکمه هایی برای مدیریت فایلها و کار با اعداد در خانه ها می باشد.

اگر نوار ابزارهای فوق در دسترس نبودند از منوی tools زیر منوی customize را انتخاب کرده و از پنجره ای که باز می شود در tab, toolbars منوی مورد نیاز را فعال کرده و کلید ok را فشار دهید.

Undo: اگر در خانه ای دکمه ای را اشتباه زده یا مطلبی را اشتباه تایپ کرده و..... برای برگشتن به حالت قبل از منوی edit گزینه undo و یا از نوار ابزار بر روی کلید آن کلیک نمایید.

ذخیره کردن پوشه کار: از منوی File فرمان Save را اجرا کرده یا روی دکمه Save در نوار ابزار استاندارد کلیک کنید. اگر پوشه را برای اولین بار ذخیره می کنید اکسل از شما می خواهد تا فایل را نامگذاری کرده و مسیر قرار گرفتن آن را مشخص نمایید.

ایجاد پوشه ی کار جدید: از نوار ابزار استاندارد روی دکمه New کلیک نمایید.

باز کردن یک پوشه ی کار: برای باز کردن پوشه ای که قبلا ذخیره نموده اید از منوی File گزینه Open را انتخاب کرده یا روی دکمه Open در نوار ابزار استاندارد کلیک نمایید.

بستن پوشه کار: برای بستن پوشه کار از منوی File گزینه Close را انتخاب کرده یا روی دکمه Close در پنجره کلیک کنید.

ذخیره کردن فایل پوشه کار با نام یا در محلی متفاوت: از منوی File گزینه Save As را انتخاب نمایید و نام و مسیر جدید را مشخص نمایید.

وارد کردن کاربرگ جدید: اگر تعداد کاربرگهای موجود در پوشه کار کافی نبود برای اضافه کردن کاربرگ جدید از منوی Insert گزینه Work sheet را انتخاب کنید. کاربرگ جدید سمت چپ کاربرگ باز شده جاری با نام پیش فرض 4 sheet اضافه می شود.

تغییر نام کاربرگها: برای تغییر نام کاربرگها روی نام آن کلیک راست کرده از منوی میان بر، Rename را انتخاب و نام جدید را وارد کنید.

پاک کردن کاربرگها: با کلیک راست بر روی نام کاربرگ از منوی میان بر، Delete را انتخاب نمایید.

استفاده از فرمول در اکسل:

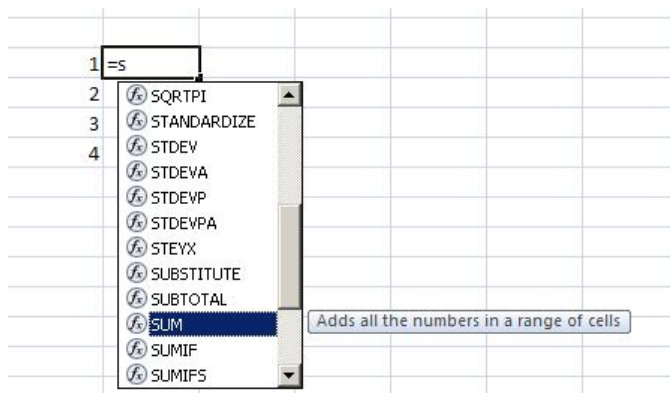
انجام عمل جمع: به طور مثال در خانه B3 عدد 25 و در خانه B4 عدد 170 را وارد نمایید سپس در خانه B5 تایپ کنید $B3+B4 =$ و کلید Enter را فشار دهید خواهد دید که جمع دو خانه در خانه B5 قابل مشاهده است.

توجه: وقتی فرمولی را وارد کردید، اکسل نشانی هر خانه را با رنگی متفاوت نشان می دهد و کادر دور خانه نیز به همان رنگ نشانی آن می شود.

برای انجام عمل تفریق، ضرب و تقسیم نیز از همین راه (با استفاده از علامت -, *, /) به جای + می توانید به نتیجه برسید. ضمناً این اعمال در هر دو حالت لیست افقی و عمودی از اعداد قابل اجرا هستند. شما برای خانه های غیر مجاور که به طور پراکنده یا دور از هم نیز هستند می توانید از فرمولهای فوق استفاده نمایید.

استفاده از توابع در اکسل:

ستونی از اعداد را در نظر بگیرید. شما میتوانید با استفاده از این ستون و توابع اکسل داده های جدید تهیه کنید. به تصویر زیر نگاه کنید، بعد از قرار دادن علامت مساوی "=" و تایپ کردن یک حرف توابع مربوطه در نواری باز میشود. و میتوان از آن استفاده کرد.



تابع SUM: به طور مثال برای جمع کردن 100 خانه، استفاده از فرمول جمع مناسب نیست. در اکسل تابع SUM اجازه می دهد تا کل لیست عمودی یا افقی از اعداد را با وارد کردن سه مورد جمع کنید:

1- نام تابع (SUM) 2- نشانی خانه اول 3- نشانی خانه آخر

مانند نمونه: = SUM(B3:B19)

به دلیل کاربرد فراوان تابع فوق دکمه AUTO SUM در نوار ابزار استاندارد تعبیه شده است و روش کار به این صورت است: فرض میکنیم در خانه های B1 تا B5 اعدادی را وارد کرده برا جمع آنها روی خانه B6 کلیک کرده (فعال می کنیم) سپس از نوار ابزار استاندارد روی فلش رو به پایین AUTO SUM کلیک کرده و گزینه SUM را انتخاب کنید. اکسل سعی می کند که حدس بزند میخواهید کدام خانه ها را با هم جمع کنید سپس کلید Enter را زده و جواب را مشاهده می کنیم.

تابع میانگین (Average): این تابع میانگین گروهی از اعداد را حساب می کند.

پس از طی مراحل مانند مثال بالا از نوار ابزار استاندارد روی فلش رو به پایین AUTO SUM کلیک کرده و گزینه Average را انتخاب کنید.

توابع MIN,MAX: از توابع فوق برای یافتن کوچک ترین و بزرگترین مقادیر در خانه ها استفاده می شود. نحوه عمل همانند توابع قبل می باشد.

تابع Count: از این تابع برای شمارش عددی ردیف‌ها یا ستون‌ها در یک محدوده استفاده میشود.

پخش کردن محتوای سلول: اگر می‌خواهید که خانه‌ای به اندازه محتوایش گسترده باشد و به خانه مجاور وارد نشده باشد، آن را انتخاب کرده از منوی **Format** زیرمنوی **Cells** را انتخاب و سپس در بند **Alignment** باکس **warp text** را علامت بزنید.

ادغام خانه‌ها و وسط چین کردن محتوای آنها: خانه‌هایی را که می‌خواهید ادغام کنید انتخاب کرده و از منوی **Format** زیرمنوی **Cells** را انتخاب و سپس در بند **Alignment** باکس **merge** را علامت بزنید.

پنهان کردن ردیف‌ها و ستون‌ها: برای پنهان کردن یک ردیف، روی خانه ردیف کلیک کرده و در منوی **Format**، روی زیرمنوی **Row**، کلیک و گزینه **Hide** را انتخاب کنید.

برای پنهان کردن یک ستون، روی خانه ستون کلیک کرده و در منوی **Format**، روی زیرمنوی **Column**، کلیک و گزینه **Hide** را انتخاب کنید.

برای آشکار کردن ردیف، روی خانه‌های ردیف‌های همجوار کلیک کرده و از منوی **Format**، **Row**، **Un hide** را انتخاب کنید.

برای آشکار کردن ستون، روی خانه‌های ستون‌های همجوار کلیک کرده و از منوی **Format**، **Column**، **Un hide** را انتخاب کنید.

Freeze panes: اگر کاربرگ شما به صفحه دیگر برود شما دیگر قادر به دیدن عناوین ستونی نیستید و زمانی که بخواهید عناوین ستون یا ردیفی همیشه در صفحه ظاهر شوند به روش زیر عمل می‌کنید:

مکان نما را در خانه‌ای زیر ردیف و در جلوی ستونی که می‌خواهید همیشه ظاهر باشند قراردادده و از منوی **View** گزینه **Freeze panes** را انتخاب نمایید در این حالت شما در هر موقعیتی از کاربرگ باشید می‌توانید عناوین فوق را مشاهده نمایید. برای از بین بردن این حالت از منوی **View** گزینه **Un Freeze panes** را انتخاب نمایید.

تنظیم عرض ستون و طول ردیف:

عرض ستون: برای تغییر عرض ستون ماوس را به عنوان ستون ببرید و سپس مرز ستون را به سمت راست ستون بکشید تا ستون بزرگ شود.

طول ردیف: برای تغییر طول ردیف ماوس را روی عنوان ردیف ببرید و سپس مرز زیر عنوان ردیف را بکشید تا ردیف به ارتفاع مورد نظر برسد.

چیدمان محتویات: و از منوی **Format** زیرمنوی **Cells** را انتخاب و در پنجره **Format Cells** دومین تب (**Alignment**) را کلیک کنید در قسمت **Horizontal** شما می توانید نحوه قرار گرفتن طولی محتویات هر خانه (راست چین، وسط چین و...) را تعیین نمایید. در قسمت **Vertical** نیز نحوه قرارگرفتن عرضی محتویات قابل تنظیم می باشند.(بالا، وسط، پایین و...)

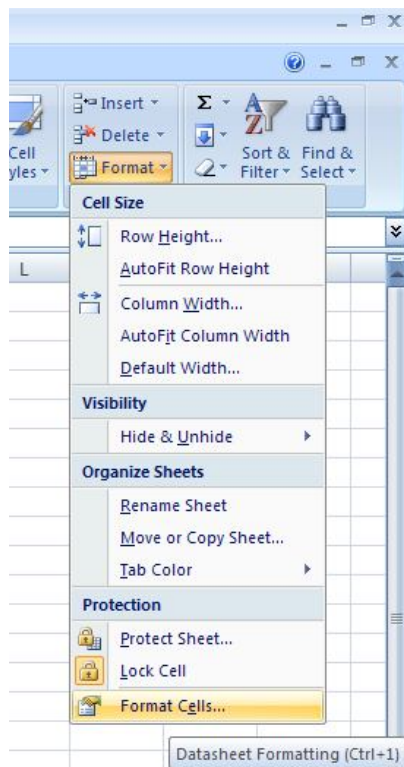
در قسمت **Orientation** شما می توانید میزان چرخش متن و... را در هر سلول تنظیم نمایید.

در تب **Font** می توانید نوع، سایز و... قلم نوشتاریتان را تنظیم نمایید.

Borders: در این قسمت می توانید خانه های مورد نظرتان را کادر بندی نمایید.

وارد کردن ردیفها و ستونها: گاهی اوقات لازم می شود تا ردیف یا ستونی را برای نگهداری داده

های جدید در محدوده از خانه های متنی و عددی وارد کنید.



وارد کردن و حذف ردیف جدید: عنوان ردیف را زیر محلی که می خواهید ردیف جدید وارد

کنید انتخاب کرده و از منوی **Insert** گزینه **Rows** را انتخاب کنید. ردیف جدیدی از خانه های خالی باز می شود.

برای حذف ردیف روی عنوان آن کلیک کرده و از منوی **delete, edit** را انتخاب نمایید.

وارد کردن ستون جدید: عنوان ستونی را که می خواهید در سمت چپ آن ستون جدیدی ایجاد

نمایید انتخاب کرده و از منوی **Insert** گزینه **Columns** را انتخاب کنید. ستون جدیدی از خانه های خالی باز می شود.

برای حذف ستون روی عنوان آن کلیک کرده و از منوی delete, edit را انتخاب نمایید.

کپی و درج محتوای خانه ها: دو راه معمول برای انجام این کار به شرح ذیل می باشد:

روش اول کپی کردن و درج با روش کشیدن و انداختن: خانه ای را که می خواهید کپی کنید انتخاب کرده مکان نما را به گوشه پایین خانه انتخاب شده ببرید و کلید Ctrl را نگه دارید مکان نما به حالت به اضافه در می آید شما با نگه داشتن کلید Ctrl خانه را به خانه مقصد کشیده دکمه ماوس و کلید کنترل را رها کنید.

روش دوم کپی کردن و درج با کلیک راست: خانه مورد نظر را انتخاب کرده و بر روی آن راست کلیک کنید گزینه Copy را از منوی میان بر انتخاب کرده و در خانه مقصد مجدداً کلیک راست کرده و این بار گزینه Paste را کلیک نمایید.

مرتب کردن و ترتیب خانه ها از روی محتوا (Sort): مرتب کردن مجدد ستون های خانه ها بر اساس مقادیر خانه انجام می شود این روش محتوای خانه ها را تغییر نمی دهد و فقط محل آنها را تغییر می دهد ترتیب چیدن می تواند الفبایی یا عددی، صعودی یا نزولی باشد. برای مثال شما لیستی از محصولات یک کارخانه و بسته بندی و قیمت آنها دارید و برای سهولت دسترسی به اطلاعات می خواهید نام محصولات را بر اساس حروف الفبا مرتب نمایید (شما می توانید بر حسب قیمت یا نوع بسته بندی نیز این کار را انجام بدهید) ابتدا محدوده خانه محصولات (از عنوان تا آخرین خانه که شامل اطلاعات می باشد) را انتخاب نمایید. از منوی Sort, Data را انتخاب نمایید پنجره ای باز می شود که شامل دو گزینه است که می خواهد بداند فقط داده های انتخاب شده باید مرتب شود یا داده های کل خانه ها گزینه اول کل خانه ها را مرتب می نماید و گزینه دوم فقط داده های انتخابی و ما در اینجا گزینه اول را انتخاب می نمایم اطلاعات خانه های مجاور مرتبط با اطلاعات فوق است و می بایست ترتیبشان حفظ شود. زمانی که کلید Sort را کلیک نمایید پنجره ای باز می شود که شما می توانید در آنجا تعیین کنید که بر اساس کدامیک از اطلاعات بر حسب اولویت و صعودی یا نزولی اطلاعات مرتب شوند.

این جدول بر اساس نام دارو (حروف الفبا) مرتب شده است.

نام دارو	شکل دارو	قیمت به ریال
استامینوفن	قرص	1,300
اکسیکتورانت	شربت	2,500
ایبوپروفن	قرص	2,000
دیکلن	ژل	8,900
سلکسیب	کپسول	1,100

این جدول بر اساس قیمت مرتب شده است.

نام دارو	شکل دارو	قیمت به ریال
سلکسیب	کپسول	1,100
استامینوفن	قرص	1,300
ایبوپروفن	قرص	2,000
اکسیکتورانت	شربت	2,500
دیکلن	ژل	8,900

این جدول بر اساس قیمت و سپس شکل دارو مرتب شده است (بر حسب دو عامل).

نام دارو	شکل دارو	قیمت به ریال
سلکسیب	کپسول	1,100
استامینوفن	قرص	1,300
ایبوپروفن	قرص	2,000
اکسیکتورانت	شربت	2,500
دیکلن	ژل	8,900

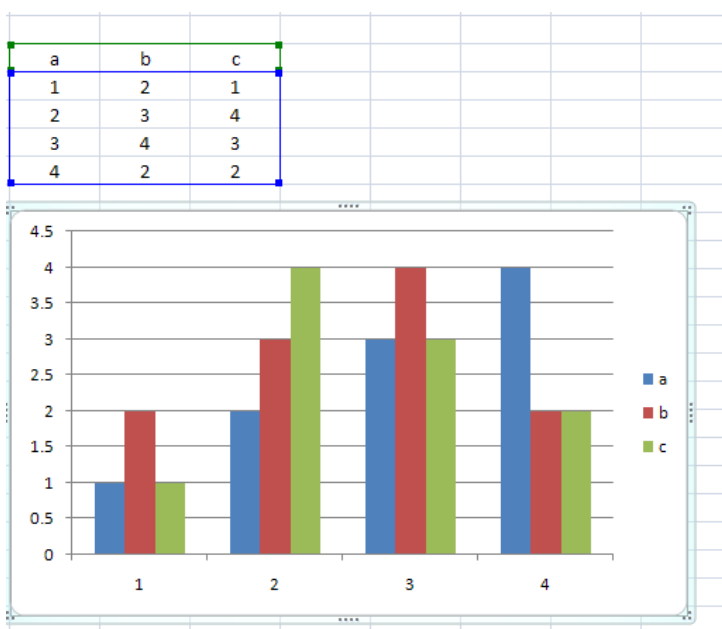
ترسیم نمودار با اکسل:

خانه هایی را که می خواهید نمودار بر اساس محتوای آنها رسم شود انتخاب کنید. در منوی Insert انواع Chart وجود دارد که شما می توانید یکی را به دلخواه و بر حسب روش انتخاب کنید.



- نوع نمودار (chart type): که در این پنجره انواع نمودار ارائه می شود و شما می توانید به دلخواه از میان آنها انتخاب نمایید.

- داده های منبع نمودار (chart source data): در این قسمت محدوده خانه هایی را که می خواهید در قالب نمودار بیاورید انتخاب می نمایید.
- گزینه های نمودار (chart options): در این قسمت گزینه های متعددی وجود دارد که شما بر حسب نیاز و دلخواه آنها را فعال می نمایید.
- محل نمودار (chart location): در این مکان شما می توانید مکان ذخیره نمودار (کاربرگ جاری، پوشه کاری دیگر، پوشه کار جدید) را انتخاب نمایید.
در سه کادر اول روی دکمه Next کلیک کرده و در کادر آخر دکمه finish را کلیک نمایید.
همانطور که ملاحظه میکنید، این تب گزینه های دیگری نیز جهت ورود به صفحه اکسل دارد.



ویرایش نمودار:

نموداری که رسم می شود قابل تغییر می باشد.

تغییر داده های نمودار: شما می توانید محتوای خانه های نمودار را که بر اساس آنها نمودار رسم شده است تغییر دهید همزمان با تغییر داده ها در کاربرگ، اکسل نمودار را تغییر می دهد.

تغییر اندازه نمودار: با کلیک بر روی محوطه نمودار دسته هایی در گوشه های نمودار دیده می شود با کلیک بر روی آنها و نگه داشتن کلید ماوس شما می توانید ابعاد نمودار را کوچک یا بزرگ نمایید.

تغییر عنوان نمودار: در هر محلی روی نمودار کلیک کرده و سپس روی متن عنوان کلیک نمایید تا آن را ویرایش کنید.

اضافه کردن عنوان نمودار: اگر نمودار فاقد عنوان بود برای اضافه کردن آن بر روی نمودار کلیک راست کرده از منوی میان بر chart option را انتخاب نمایید در تب Title قسمت اول (chart title) می توانید عنوان را تایپ نمایید.

برای حذف عنوان نمودار بر روی آن کلیک کرده و delete را بزنید.

اضافه کردن برجسب داده ها: می توان دو نوع برجسب داده ای را به نمودار اضافه کرد.

1- برجسب های مقدار: شامل مقادیر عددی از نقاط داده ای جداگانه است. روش کار: روی محوطه نمودار کلیک راست کرده از منوی میان بر chart option را انتخاب کرده و در تب پنجم Data labels را کلیک نمایید. گزینه Value را انتخاب کرده و ok نمایید.

2- برجسب های متی: که نام نقاط داده ای را نمایش می دهد. روش کار: روی محوطه نمودار کلیک راست کرده از منوی میان بر chart option را انتخاب کرده و در تب پنجم Data labels را کلیک نمایید. گزینه Category name را انتخاب کرده و ok نمایید.

تغییر فرمت برجسب داده ها:

روی برجسب کلیک راست کرده و از منوی میان بر Format Data Labels را انتخاب کنید و گزینه های مورد نظران را از تب های چهار کارد محاوره ای انتخاب کنید. (Alignment, Number, Font, Patterns)

تغییر مقیاس: برای تغییر مقیاس روی محور عمودی نمودار 2 بار کلیک کرده و از کادر Format Axis بند Scale را انتخاب نمایید. در قسمت Minimum شما می توانید تعیین کنید که نمودار از چه عددی شروع شود در قسمت Maximum بیشترین مقدار نمودار را مشخص می نمایید. و در قسمت Major unit میزان فاصله بین اعداد را تعیین می نمایید.

تغییر رنگ نمودار:

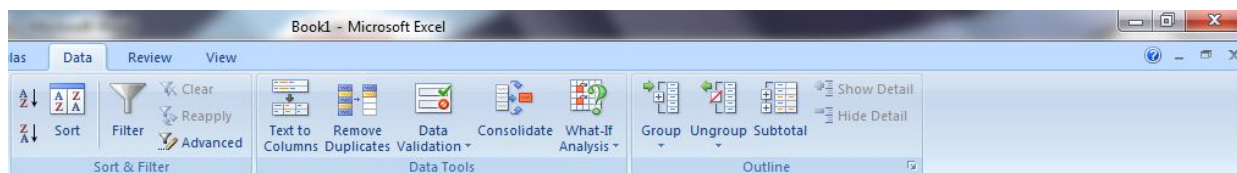
برای تغییر محوطه نمودار روی آن کلیک راست کرده و از منوی میان بر گزینه Format plot Area را انتخاب نمایید از کادر محاوره ای باز شده رنگ یا افکت دلخواه را انتخاب نمایید. برای تغییر رنگ ستون سری داده ها روی خود نمودار روی نمودار کلیک راست کرده از منوی میان بر گزینه Format Data Series را انتخاب و در کادر محاوره ای باز شده بند Patterns را انتخاب کرده و در آنجا رنگ و افکت مورد نظر را برگزینید.

تغییر نوع نمودار:

برای تغییر نوع نمودار روی هر قسمت از نمودار کلیک راست کرده و از منوی میان بر گزینه Chart Type را انتخاب کرده و در کادر باز شده نوع دیگری از نمودار را برگزینید.

بخش دوم

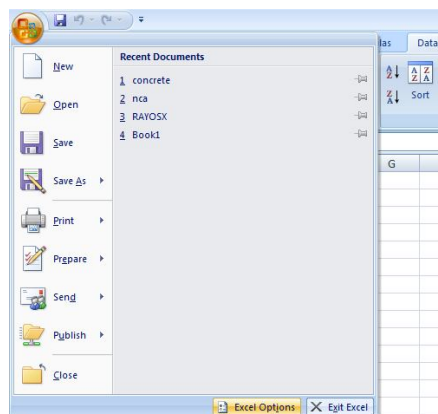
نحوه افزودن گزینه Data Analysis به زیر منوهای تب Data:
در حالت اولیه تب Data به صورت زیر است:



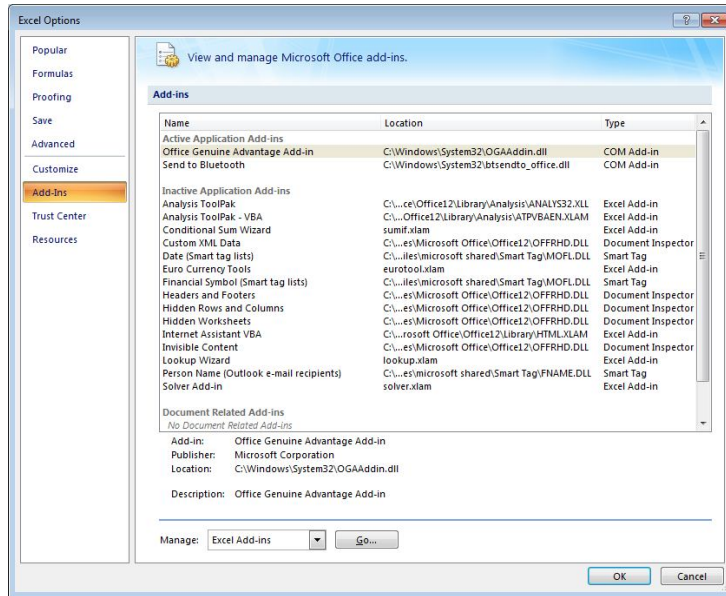
همانطور که ملاحظه میکنید در زیر تب Data خبری از Data Analysis نمی باشد.
برای بازخوانی این زیرگزینه و انجام تحلیل های آماری مربوطه مراحل زیر را دنبال کنید:

1. اکسل را باز کنید.

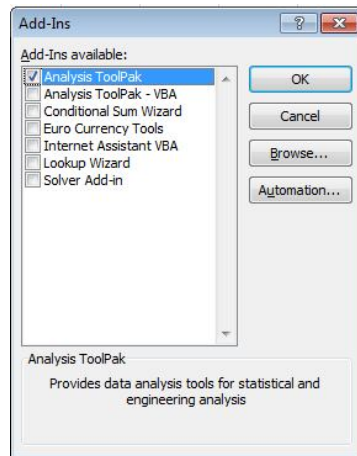
2. گزینه Excel options را که در قسمت گوشه چپ می باشد، کلیک کنید تا باز شود.



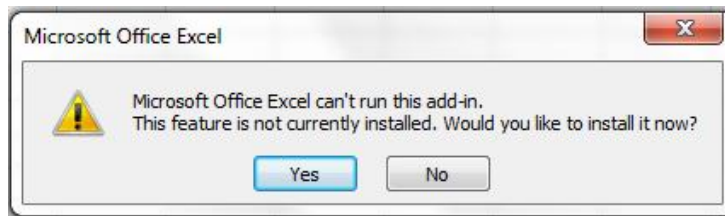
3. در قسمت چپ از زیر منوهای مشاهده شده، گزینه Add-Ins را کلیک کنید.



4. بر روی کلید Go در قسمت پایین کلیک کنید تا وارد برنامه های جانبی اکسل شود. سپس از گزینه هایی که مشاهده میکنید، اولین گزینه یعنی Analysis ToolPak را علامت زده و Ok کنید.

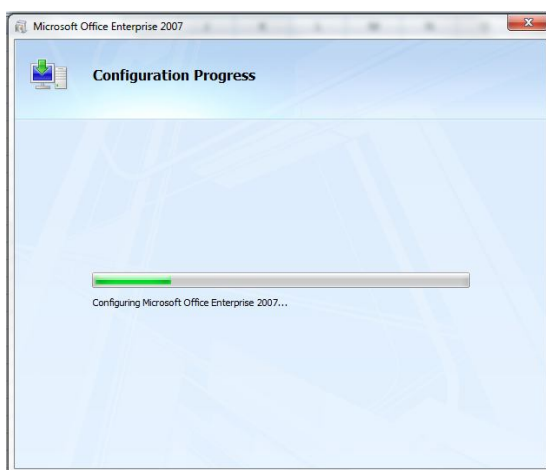


5. در این مرحله پیغامی ظاهر میشود مبنی بر اینکه آیا میخواهید برنامه جدید در اکسل قرار دهید و آن را آرایش کنید؟ شما با جواب مثبت (Yes) ادامه دهید.

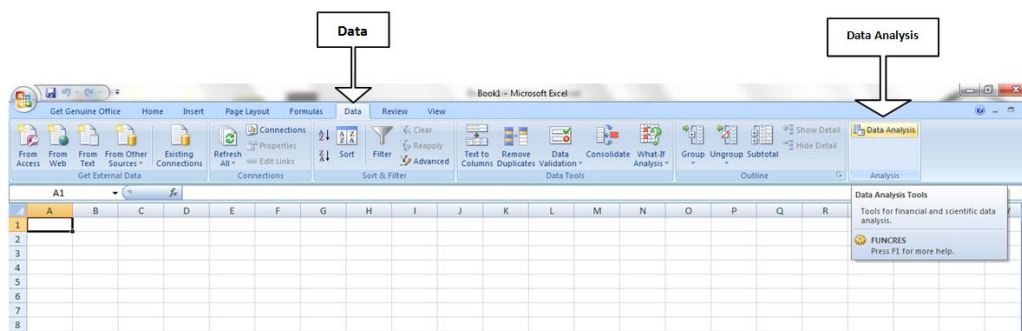


اکسل: مقدمه، تحلیل واریانس، رگرسیون و آزمون میانگینها.....13

باید قدری شکيبا باشيد تا برنامه به طور کامل Run شود. در اين مرحله پنجره جديدي باز ميشود که در حال آرايش برنامه ای جديد در اکسل است.



6. شما کار بزرگی انجام داده اید! بر روی تب (گزینه) Data کلیک کنید در قسمت انتهایی این بخش به نرم افزار گزینه ی جدیدی با عنوان "Data Analysis" اضافه شده است. به شکل زیر نگاه کنید.

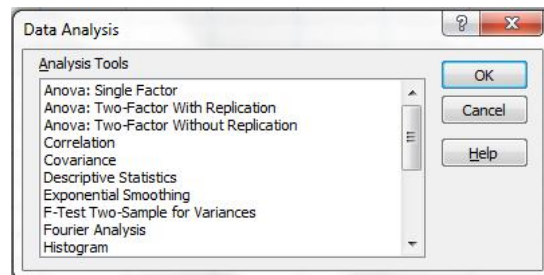


اگر مراحل بالا را به درستی انجام داده باشید با گزینه جدید را ملاحظه کنید. یکی از مواردی که برنامه بازخوانی نمیشود و شما زیر تب Data Analysis را نمیبینید، این است که برنامه office ایراد ریشه ای دارد باید دوباره، آنهم نرم افزار معتبرتر نصب شود. حال با مراجعه به بخش بعد، کارایی این گزینه جدید را بیشتر و بهتر توضیح میدهیم.

بخش سوم

اضافه شدن این گزینه به اکسل به چه معناست؟ اینکه چه استفاده ای از این گزینه میتوان کرد را با نگاه به تصویر مربوطه عنوان میکنیم.

بدون اینکه داده ای در اکسل وارد کنید، بروی مهمان جدید (Data Analysis) کلیک کنید. پنجره ای باز میشود که شامل گزینه های زیر است:



- Anova: Single Factor
- Anova: Two-Factor with Replication
- Anova: Two-Factor Without Replication
- Correlation
- Covariance
- Descriptive Statistics
- Exponential Smoothing
- F-Test Two-Sample for Variances
- Fourier Analysis
- Histogram
- Moving Average:
- Random Number Generation
- Rank and Percentile
- Regression
- Sampling
- t-Test: Paired Two Sample For Means
- t-Test: Two-Sample Assuming Equal Variances
- t-Test: Two-Sample Assuming Unequal Variances
- z-Test

اکسل: مقدمه، تحلیل واریانس، رگرسیون و آزمون میانگینها.....15

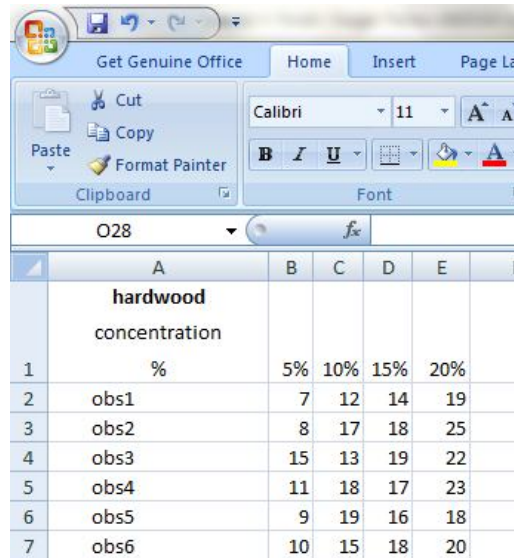
تمامی این موارد برای دانشجویان علوم انسانی که آمار خوانده اند آشناست. ترجیح داده شده است، با مثالهایی ساده تمامی موارد بالا توضیح داده شود.

Anova: Single Factor

تحلیل واریانس: تک فاکتوری

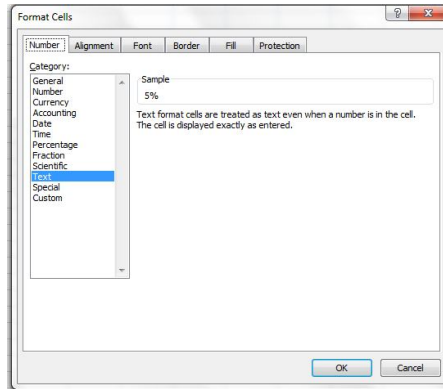
داده های زیر را همانند شکل وارد اکسل کنید:

	Hardwood concentration %	Obs1	Obs2	Obs3	Obs4	Obs5	Obs6
گروه ها	5%	7	8	15	11	9	10
	10%	12	17	13	18	19	15
	15%	14	18	19	17	16	18
	20%	19	25	25	23	18	20



	A	B	C	D	E	F
	hardwood concentration					
1	%	5%	10%	15%	20%	
2	obs1	7	12	14	19	
3	obs2	8	17	18	25	
4	obs3	15	13	19	22	
5	obs4	11	18	17	23	
6	obs5	9	19	16	18	
7	obs6	10	15	18	20	

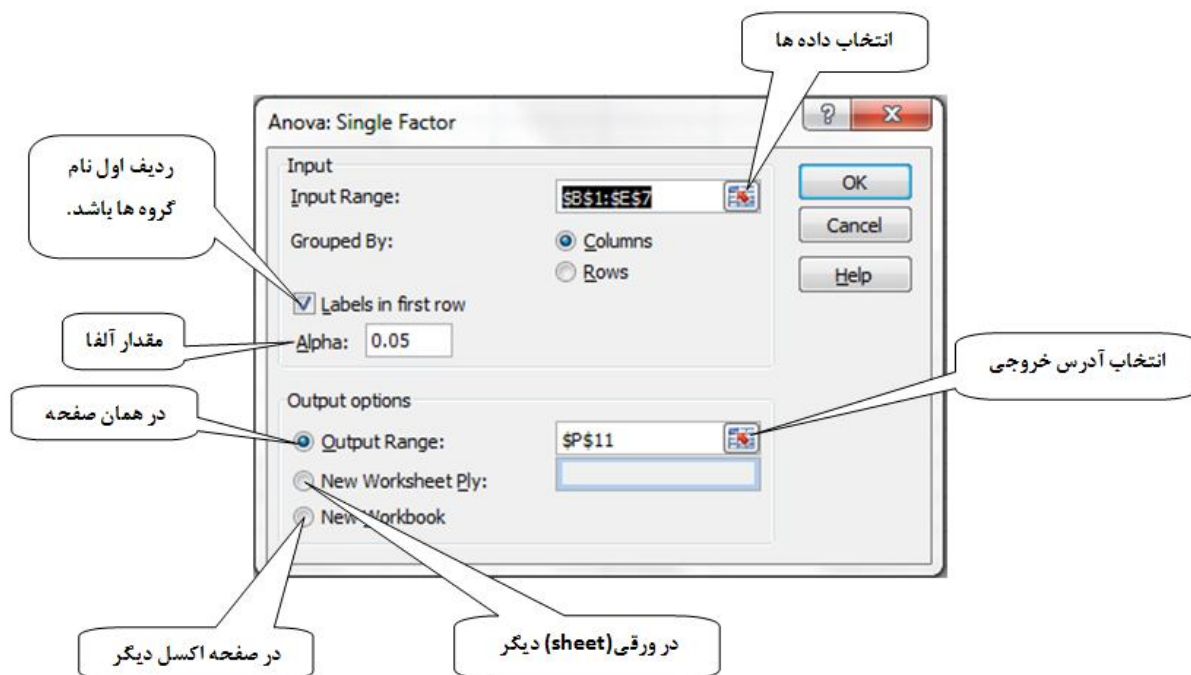
نکته: روی ردیف اول کلیک راست کرده ... Format Cells را انتخاب کنید. بعد در قسمت Number گزینه Text را انتخاب کنید.




چرا؟ به این دلیل که در اینجا اکسل این ردیف را به عنوان نام گروه ها وارد میکند نه مقداری از داده ها. و اگر در حالت عددی باشد و حرف وارد کنیم، هنگام محاسبات خطا میدهد. البته هنگامی که پنجره دستوری باز میشود، میتوان با انجام عملی از بروز این خطا جلوگیری کرد(در بخشهای بعدی بیشتر توضیح داده خواهد شد). تفاوتی که با نرم افزار Spss وجود دارد در این است که نام متغیرها در Spss در بالای ستون ها به صورت جداست، ولی در اکسل خیر. نامهای ... A B C را که در بالا میبینید ثابت بوده و مخصوص آدرس دهی برای اجرای دستورات است، نمیتوان آنها را عوض کرد. در مراحل بعدی این مسئله را بیشتر توضیح میدهم.

بروی کلید (Data Analysis) کلیک کنید، از پنجره باز شده **Anova: Single Factor**

را انتخاب کنید. پنجره زیر باز میشود:




با زدن کلیدهای به شکل  محل ورودی و خروجی را مشخص کنید.

هنگامی که بر روی این کلیدها کلیک می‌کنید پنجره بزرگ بالا باریک می‌شود. در این جا با کلیک کردن بر گوشه بالای گروه اول و کشیدن آن به سمت گوشه آخرین داده، داده‌های مربوطه را انتخاب کنید. در این حالت دور داده‌ها به همراه ردیف اول چین چین می‌شود (فقط دقت کنید).

	A	B	C	D	E	F	G
	hardwood concentration						
1	%	5%	10%	15%	20%		
2	obs1	7	12	14	19		
3	obs2	8	17	18	25		
4	obs3	15	13	19	22		
5	obs4	11	18	17	23		
6	obs5	9	19	16	18		
7	obs6	10	15	18	20		
8							
9							
10							
11							
12							

Anova: Single Factor	
Input Range:	\$B\$1:\$E\$7

بر روی علامت  کلیک کنید تا به پنجره سابق برگردید، حال به سراغ محل خروجی بروید: در محل خروجی سه گزینه وجود دارد:

الف) Output Range: این گزینه شما را قادر می‌سازد تا خروجی‌ها را در همان صفحه مشخص کنید (در اینجا دوباره بر روی علامت مربوطه کلیک کنید در صفحه اکسل کافیست بر روی یک خانه با فاصله دلخواه از داده‌ها کلیک کنید و باز به صفحه پنجره اصلی بازگردید).

ب) New Worksheet ply: این گزینه شما را قادر می‌سازد که خروجی‌ها را در کاربرگ (sheet) دیگر ببینید. (در اکسل می‌توان مجموعه داده‌ها را در تعداد کاربرگ‌های دلخواه وارد کرد).

ج) New Workbook: با زدن این گزینه شما خروجی‌ها را در صفحه جداگانه دیگری مشاهده خواهید کرد.

گروه‌ها همان ستون‌های ما هستند پس باید گزینه Columns تیک خورده باقی بماند. اگر گزینه Labels in first row را تیک بزنید و هنگام انتخاب ورودی‌ها آن را هم انتخاب کرده باشید به جای نام پیش فرض گروه‌ها نام‌هایی که شما انتخاب کرده‌اید ظاهر خواهد شد. مقدار آلفا نیز که پیش فرض 0.05 دارد، سطح معنادار بودن را نشان می‌دهد و قابل تغییر است. در این مرحله ورود داده‌ها پایان یافته است. دستور را OK کنید. اگر تمام مراحل را بدرستی انجام داده باشید، خروجی زیر در صفحه اکسل شما قابل رؤیت است:

Anova: Single Factor						
SUMMARY						
Groups	Count	Sum	Average	Variance		
5%	6	60	10	8		
10%	6	94	15.66667	7.866667		
15%	6	102	17	3.2		
20%	6	127	21.16667	6.966667		
ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	382.7917	3	127.5972	19.60521	3.593E-06	3.098391
Within Groups	130.1667	20	6.508333			
Total	512.9583	23				

تفسیر با خودتان!

Anova: Two-Factor with Replication

تحلیل واریانس: دو فاکتوری با تکرار

مثال: فرض کنید میخواهید تأثیر سه نوع دارو را در مورد دو گروه مرد و زن آزمایش کنید. یعنی دو جمعیت بیمار و سه نوع مختلف دارو.

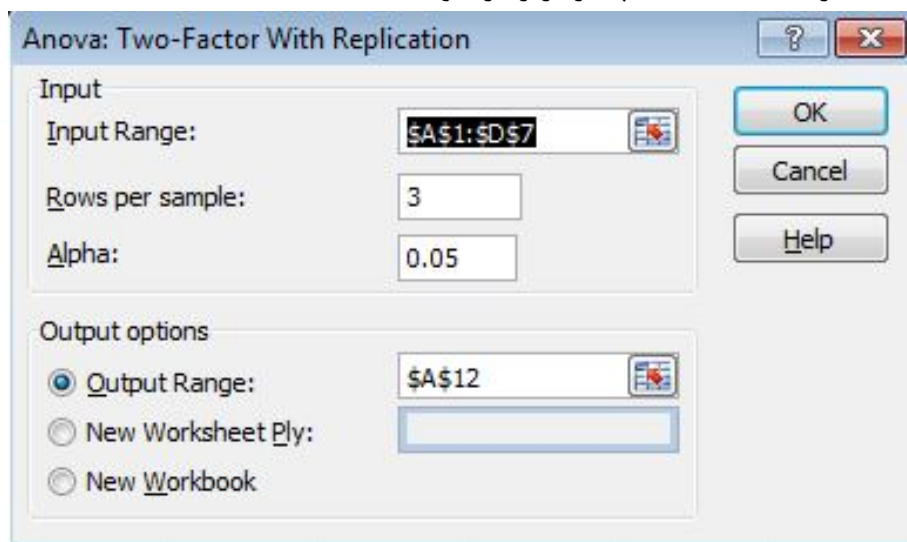
	A	B	C	D
1	Patients	Drug 1	Drug 2	Drug 3
2	Male	8	10	8
3		4	8	6
4		0	6	4
5	Female	14	4	15
6		10	2	12
7		6	0	9

داده ها را وارد اکسل کنید. مانند شکل زیر:

	A	B	C	D
1	patients	Drug1	Drug2	Drug3
2	Male	8	10	8
3		4	8	6
4		0	6	4
5	Female	14	4	15
6		10	2	12
7		6	0	9

چرا بصورت بالا وارد میکنیم؟ ردیف اول نشان از بیماری و نوع داروهاست، ردیف 2، 3 و 4 در مورد مردان و ردیفهای باقی در مورد زنان است. در اینجا نیازی به تغییر فرمت ردیف ها ندارید. همین گونه به حالت پیش فرض باقی گذارید.

حال بر روی گزینه (Data Analysis) کلیک کنید واز پنجره باز شده Anova: Two-Factor with Replication را انتخاب کنید تا پنجره زیر باز شود.



تمامی موارد همان دستور قبلی میباشد ولی فقط باید دو نکته زیر را دقت کنید:
الف) ورودی داده ها باید شامل تمامی جدول شود یعنی از خانه A₁ کلیک کرده تا آخرین خانه میکشیم (مانند شکل زیر).

	A	B	C	D	E
1	patients	Drug1	Drug2	Drug3	
2	Male		8	10	8
3			4	8	6
4			0	6	4
5	Female		14	4	15
6			10	2	12
7			6	0	9
8					

ب) وقتی پنجره را باز میکنید قسمتی با این عنوان Rows per sample: اضافه شده است. در مثال حاضر برای هر کدام از نمونه ها(زن و مرد) سه ردیف داریم. فراموش نکنید که اندازه گیری با تکرار است و یک عدد را باید وارد کنید. بعبارت دیگر سوال این است که برای هر کدام از نمونه ها چند ردیف اختصاص یابد، پاسخ برابر است با:3.

حال بعد از آنکه مکان خروجی را مشخص کردید، دستور را OK کنید.

اگر تمامی مراحل را درست انجام داده باشید باید خروجی زیر را مشاهده کنید.

Anova: Two-Factor With Replication						
SUMMARY	Drug1	Drug2	Drug3	Total		
<i>Male</i>						
Count	3	3	3	9		
Sum	12	24	18	54		
Average	4	8	6	6		
Variance	16	4	4	9		
<i>Female</i>						
Count	3	3	3	9		
Sum	30	6	36	72		
Average	10	2	12	8		
Variance	16	4	9	28.25		
<i>Total</i>						
Count	6	6	6			
Sum	42	30	54			
Average	7	5	9			
Variance	23.6	14	16			
ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Sample	18	1	18	2.037736	0.17894	4.747225
Columns	48	2	24	2.716981	0.106343	3.885294
Interaction	144	2	72	8.150943	0.00581	3.885294
Within	106	12	8.833333			
Total	316	17				

اکسل: مقدمه، تحلیل واریانس، رگرسیون و آزمون میانگینها.....21

تفسیر: در اینجا سه نوع F وجود دارد:

الف) Sample: تفاوت میانگین میان دو گروه یعنی زنان و مردان را میسنجد.

ب) Columns: تفاوت میان سه گروه داروها را میسنجد.

ج) Interaction: تفاوت میانگین تعاملی بین دو گروه (بیماری و جنسیت). که مهمتر نیز میباشد.

با توجه به مقدار p-value در دو تای اول فرض صفر رد نشده ولی در سومی رد میشود. این به

معنای این است که تأثیر داروها برای دو گروه مردان و زنان برابر نیست.

Anova: Two-Factor Without Replication

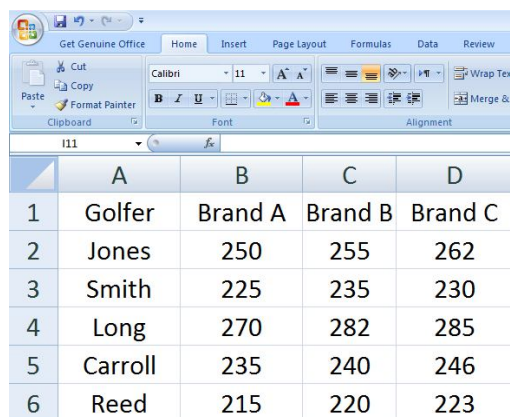
تحلی واریانس: دوفاکتوری بدون تکرار

مثال: فرض کنید 5 بازیکن گلف سه رکورد زده اند. داده های زیر رکوردهای آنها را نشان میدهد.

میخواهیم اختلاف میانگین ها را در رکوردها و گلف بازها بسنجیم.

	A	B	C	D
1	Golfer	Brand A	Brand B	Brand C
2	Jones	250	255	262
3	Smith	225	235	230
4	Long	270	282	285
5	Carroll	235	240	246
6	Reed	215	220	223

داده ها را مانند شکل زیر وارد اکسل کنید:



	A	B	C	D
1	Golfer	Brand A	Brand B	Brand C
2	Jones	250	255	262
3	Smith	225	235	230
4	Long	270	282	285
5	Carroll	235	240	246
6	Reed	215	220	223

دستور Anova: Two-Factor Without Replication را اجرا کنید.

همه مراحل را مانند دستور قبلی بروید. فقط گزینه Labels را تیک بزنید. اگر مراحل را بدرستی

انجام دهید، خروجی زیر را در صفحه اکسل مشاهده خواهید کرد.

Anova: Two-Factor Without Replication						
SUMMARY	Count	Sum	Average	Variance		
Jones	3	767	255.667	36.3333		
Smith	3	690	230	25		
Long	3	837	279	63		
Carroll	3	721	240.333	30.3333		
Reed	3	658	219.333	16.3333		
Brand A	5	1195	239	467.5		
Brand B	5	1232	246.4	552.3		
Brand C	5	1246	249.2	627.7		
ANOVA						
Source of Variation	SS	df	MS	F	P-value	F crit
Rows	6525.73	4	1631.43	203.083	4.49E-08	3.83785
Columns	277.733	2	138.867	17.2863	0.00125	4.45897
Error	64.2667	8	8.03333			
Total	6867.73	14				

تفسیر: p-value برای Row بیشتر از 0.05 میباشد. یعنی فرض صفر رد نشده و میانگینهای بازیکنان با هم برابر نیست. اما در مورد columns این گونه نیست. یعنی فرض صفر رد میشود، میانگین رکوردها برای هر بند با هم برابر است. Error هم نشان از میزان خطاست، که برای آن F تعریف نمیشود.

Correlation

همبستگی

همبستگی، به سادگی بدست می آید. برای مثال قبل با اجرای دستور Correlation همبستگی بین سه برند را بدست آورد.

دستور را اجرا کنید، داده ها را طوری انتخاب کنید که ردیف اول را نیز شامل شود. اما ستون اول لازم نیست. سپس، گزینه Labels in First Row را تیک بزنید تا نام برندها در خروجی ظاهر شود.

اگر مراحل را بدرستی انجام داده باشید خروجی زیر را ملاحظه خواهید کرد:

	Brand A	Brand B	Brand C
Brand A	1		
Brand B	0.992356	1	
Brand C	0.997305	0.983196	1

Covariance

کوواریانس

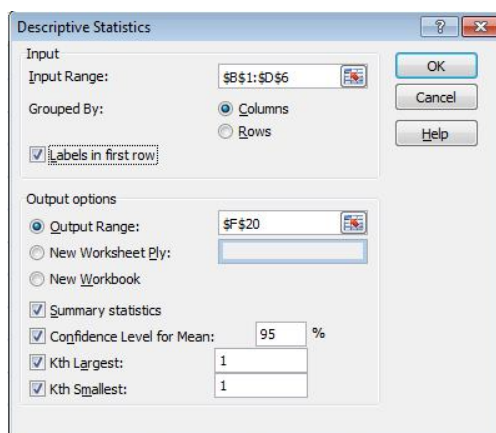
کوواریانس، یعنی هم تغییری همزمان داده ها در گروه های جداگانه باهم. برای مثال سابق همین عمل را انجام میدهیم: دستور را مانند دستور همبستگی اجرا کنید. اگر اشتباه نکرده باشید باید خروجی زیر را بدست آورید:

	Brand A	Brand B	Brand C
Brand A	374		
Brand B	403.4	441.84	
Brand C	432.2	463.12	502.16

Descriptive Statistics

آماره های توصیفی

با اجرای دستور **Descriptive Statistics** شما میتوانید آماره های توصیفی مربوط به داده های خود را بدست آورید. انتخاب داده ها همانند دستور قبلی است. پنجره ای که باز میشود چند گزینه مازاد دارد.



مقدار Kth به چه معناست؟

این مقادیر بزرگترین و کوچکترین مقادیر را نشان میدهند. با این وصف که، قرار دادن عدد یک در کادرهای مقابل آنها به معنای اولین مقدار بزرگ و اولین مقدار کوچک در هر گروه است، که در خروجی دیده میشود. اگر عدد 2 را قرار دهید، دومین مقدار بزرگ و دومین مقدار کوچک را نشان میدهد. اگر مراحل دستوری را بدرستی انجام داده باشید، باید به خروجی زیر برسید:

Brand A		Brand B		Brand C	
Mean	239	Mean	246.4	Mean	249.2
Standard Error	9.669539803	Standard Error	10.50999524	Standard Error	11.2044634
Median	235	Median	240	Median	246
Mode	#N/A	Mode	#N/A	Mode	#N/A
Standard Deviation	21.62174831	Standard Deviation	23.50106381	Standard Deviation	25.05394181
Sample Variance	467.5	Sample Variance	552.3	Sample Variance	627.7
Kurtosis	-0.515599531	Kurtosis	0.671095138	Kurtosis	-0.77727572
Skewness	0.597289281	Skewness	0.822141602	Skewness	0.606319609
Range	55	Range	62	Range	62
Minimum	215	Minimum	220	Minimum	223
Maximum	270	Maximum	282	Maximum	285
Sum	1195	Sum	1232	Sum	1246
Count	5	Count	5	Count	5
Largest(5)	215	Largest(5)	220	Largest(5)	223
Smallest(5)	270	Smallest(5)	282	Smallest(5)	285
Confidence Level(95.0%)	26.84694645	Confidence Level(95.0%)	29.18042485	Confidence Level(95.0%)	31.10857755

در نرم افزار اکسل علامت #N/A نشان میدهد که در این قسمت مقداری وجود ندارد.

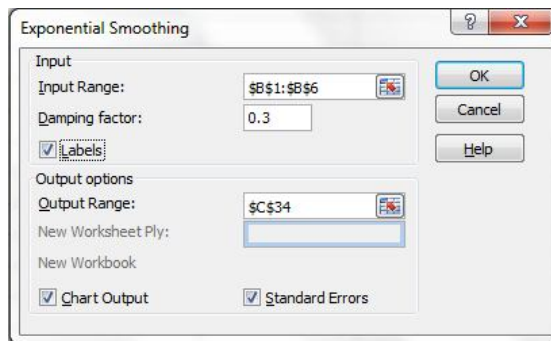
Exponential Smoothing

برازشِ نمایی

این دستور داده ها با تابع نمایی برازش میکند و نموداری را رسم میکند که دو جزء دارد: منحنی طبیعی یا واقعی نمایی (Actual) و منحنی پیش بینی نمایی (Forecast).

دستور Exponential Smoothing را اجرا کنید. در نظر بگیرید که فقط میتوان یک سطر یا ستون را وارد دستور کرد.

در برخی مواقع اعلام میکنند که نتیجه این کار شبیه به میانگین متحرک است. در مثال قبل برای Brand A این دستور را اجرا کنید.



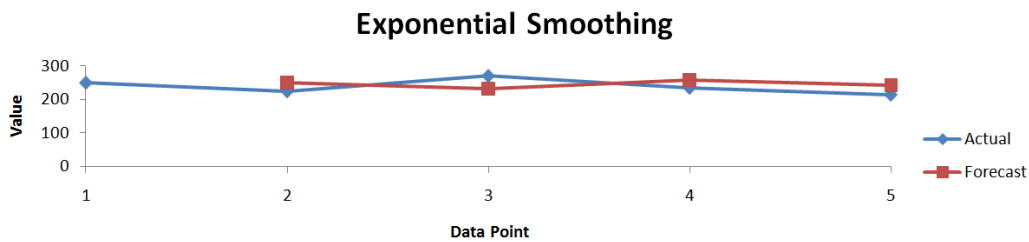
نکته:

الف) **Damping factor** چیست؟ فاکتور تعدیل، این توانایی را به شما میدهد که، مقدار ثابت برازش نمایی را مشخص کنید. باید کمترین مقدار برای آن در نظر بگیرید. این عامل، عاملی است که عدم تعادل داده های جمع آوری شده در بین یک جمعیت را حداقل میکند. پیش فرض برای آن 0.3 است.

ب) **Chart Output**: انتخاب این گزینه باعث میشود، نمودارهای جاسازی شده برای مقادیر واقعی و پیش بینی شده در جدول خروجی ظاهر شود.

ج) **Standard Errors**: با انتخاب این گزینه مقادیر خطای استاندارد در جدول خروجی در ستونی دیگر ظاهر میشود.

با انجام صحیح دستورات خروجی زیر را در صفحه اکسل مشاهده خواهید کرد.



#N/A	#N/A
250	#N/A
232.5	#N/A
258.75	#N/A
242.125	29.4127

با مطالعه توضیحات دستورات میتوان نمودار و جدول بالا را تفسیر کرد.

F-Test Two-Sample for Variances

آزمون F برای مقایسه واریانس دو نمونه

از این آزمون پراکندگی دو جامعه را با مقایسه واریانسها بررسی میکند. این آزمون با نام نسبت واریانس نیز شناخته میشود. آزمون F واریانس دو مجموعه داده را محاسبه کرده، مقدار بیشتر بر کمتر را قرار داده نسبت را در جدول نشان میدهد.

مثال: برای دو مجموعه داده زیر دستور بالا را اجرا می کنیم.

	A	B	C	D
1				
2		Sample 1	Sample 2	
3		11	15	
4		2	8	
5		16	23	
6		12	12	
7		7	21	
8		18	13	
9		15	23	
10		22	10	
11		14	15	
12		19	16	
13				

اگر دستور را بدرستی انجام دهید خروجی زیر حاصل خواهد شد. ذکر این نکته خالی از لطف نیست که دو نوار برای ورود داده وجود دارد.

F-Test Two-Sample for Variances		
	sample 1	sample 2
Mean	13.6	15.6
Variance	34.93333	27.6
Observations	10	10
df	9	9
F	1.2657	
P(F<=f) one-tail	0.365652	
F Critical one-tail	3.178893	

مقدار P-value معناداری را می‌رساند. در ضمن اگر واریانسها برابر باشد مقدار F یک خواهد شد.

Fourier Analysis

تحلیل فوریه

در این تحلیل داده‌هایی مورد تحلیل قرار می‌گیرند که توانی از 2 باشند. به این سری، سری اعداد فوریه گویند. این مبحث تناسبی با موضوع ندارد دوستان برای اطلاعات بیشتر به کتب آماری و ریاضیات مراجعه کنند.

Histogram

هیستوگرام

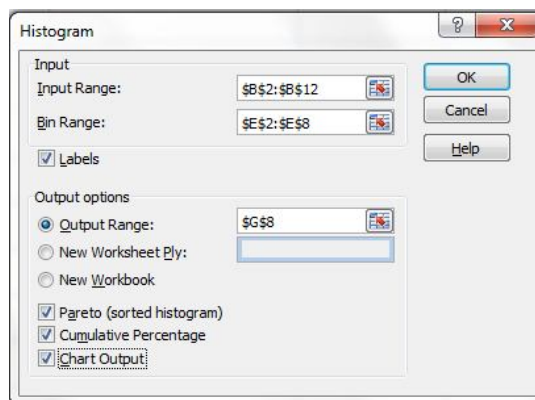
اعداد Bin (اعداد جفت یا دوتایی): این اعداد نشان دهنده فواصلی است که شما می‌خواهید، هیستوگرام برای اندازه‌گیری داده‌های ورودی در آنالیز داده‌ها بکار برد. یعنی از یک داده کم، که می‌تواند کمترین داده باشد، شروع کرده و تقسیم فواصل مساوی یا نامساوی تا داده بزرگتر که می‌تواند آخرین داده باشد، برای اکسل مشخص می‌کنیم که چه رنجی را در نظر گیرد.

در مثال قبل فرض کند بخواهیم برای داده‌های Sample 1 هیستوگرام با داده‌های bin زیر رسم کنیم:
2 – 6 – 10 – 14 – 18 – 22

نکته: باید عدد Bin را به طور جداگانه در ستونی دیگر وارد کرده و هنگام ورود اعداد Bin به آن ستون دلخواه ارجاع دهیم. به شکل زیر که داده‌های وارد شده را نشان می‌دهد بنگرید:

B	C	D	E
sample 1	sample 2		Bin
2	8		2
7	21		6
11	15		10
12	12		14
14	15		18
15	23		22
16	23		
18	13		
19	16		
22	10		

دستور را اجرا می‌کنیم تا پنجره زیر باز شود.



نکته:

مکان قرار گیری داده های Bin مشخص است.

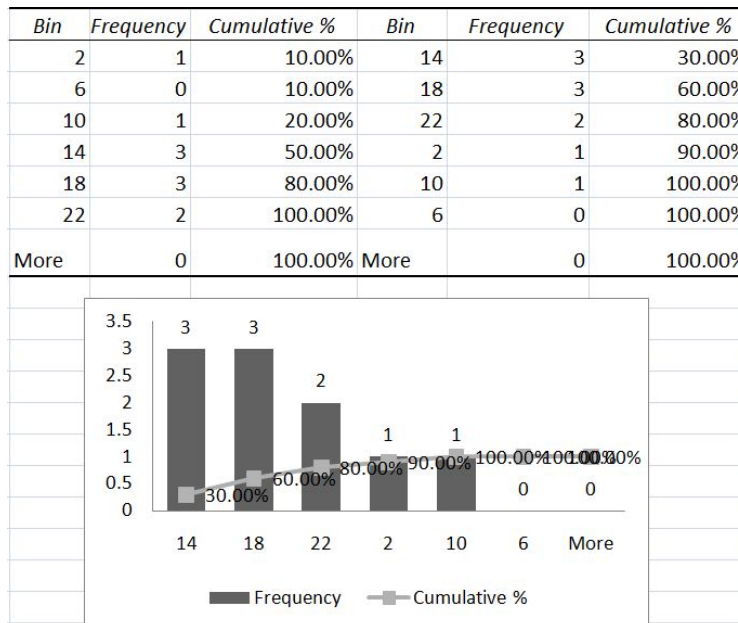
Pareto: انتخاب این گزینه داده ها را در خروجی بصورت نزولی نمایش میدهد و اگر آن را علامت نزینید

بصورت صعودی نشان میدهد.

Cumulative Percentage: جهت نشان دادن داده ها و منحنی تجمعی بکار میرود.

Chart Output: جهت نشان دادن چارت در خروجی اکسل بکار میرود.

اگر دستورات را به درستی انجام داده باشید، خروجی زیر را مشاهده خواهید کرد.



همانطور که در جدول ملاحظه میکنید: یکبار اعداد Bin را که ما مرتب کرده ایم نوشته شده، سپس

براساس مقدار فراوانی دوباره مرتب شده است.

Moving Average

میانگین متحرک

میانگین متحرک به شناسایی روندها در داده ها و فراوانی های ترتیبی (پشت سر هم) کمک میکند. داده ها

بر یک دوره مثلا دوره ی روزانه مبتنی هستند. این میانگین برای شماری از این دوره ها محاسبه میشود. پس،

ممکن است اولین میانگین برای روز اول تا هفتم و دومین میانگین برای روز دوم تا هشتم و ال آخر باشد. این

نوع میانگین مانند سری زمانی جهت پیش بینی نیز کاربرد دارد. لازم به ذکر است هر مقدار پیش بینی کننده بر

فرمولهای دنباله های مبتنی میباشد. برای کاهش اثرات خوشه یا دوره ها از این میانگین استفاده میکنیم.

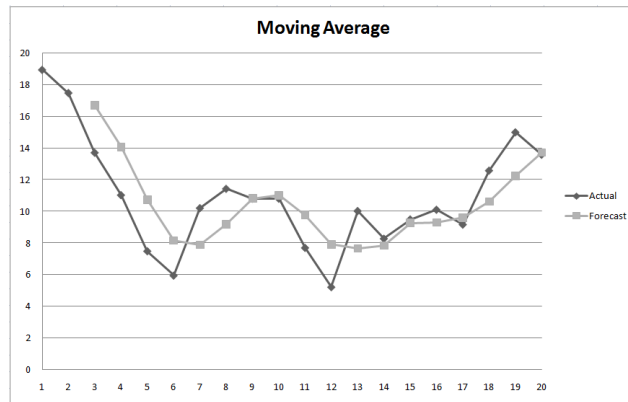
مثال: برای داده های زیر دستور را اجرا میکنیم.

	A	B	C
1			
2			
3		1	18.15
4		2	18.93
5		3	17.47
6		4	13.7
7		5	11.02
8		6	7.48
9		7	5.94
10		8	10.2
11		9	11.42
12		0	10.8
13		10	10.8
14		11	7.69
15		12	5.21
16		13	10.02
17		14	8.27
18		15	9.49
19		16	10.1
20		17	9.16
21		18	12.57
22		19	14.99
23		20	13.57

نکته: در قسمت Interval عددی را وارد میکنیم که میخواهیم تعداد داده های هر گروه را تشکیل دهیم پیش فرض عدد 3 می باشد. در ضمن دو ستون داده که ایجاد میشود به دلیل تیک زدن گزینه آخر پنجره یعنی، Standard Error میباشد.

در صورت اجرای درست دستور با خروجی زیر روبرو میشویم:

F	G
#N/A	#N/A
#N/A	#N/A
16.7	#N/A
14.0633	#N/A
10.7333	3.10086
8.14667	2.87027
7.87333	2.63736
9.18667	2.25614
10.8067	1.86201
11.0067	1.29493
9.76333	1.20298
7.9	1.96448
7.64	2.39438
7.83333	2.08895
9.26	1.40333
9.28667	0.54927
9.58333	0.54578
10.61	1.24931
12.24	1.96497
13.71	1.95138



Random Number Generation

ساخت اعداد تصادفی

اکسل شما را قادر میسازد تا اعداد تصادفی بسازید. با استفاده از این دستور ساده شما میتوانید با در نظر

گرفتن موارد زیر اعمال منحصر بفردی انجام دهید.

تعداد متغیرها: همان تعداد ستون هاست.

تعداد شماره های تصادفی: چگونگی بازگشت مقادیر در هر ستون را مشخص میکند.

مورد تصادفی/انتخاب: با استفاده از این عدد تعداد بازگشت به همان اعداد را مشخص میکنیم.

انواع توزیع:

Uniform: یکنواخت، بر اساس مقادیر کم و زیاد.

Normal: نرمال، میانگین و انحراف استاندارد.

Bernoulli: برنولی، P-value

Binominal: دو جمله ای، P-value و تعداد آزمونها.

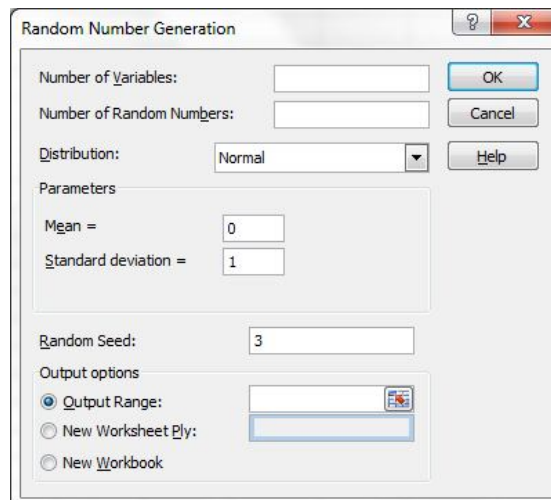
Poisson: پواسون، لاند.

Patterned: الگویی، کم، زیاد، مرحله ای، بدون تکرار.

Discrete: مجزا، مقدار و احتمال.

این مبحث را با مثالی ساده توضیح میدهیم، چرا که جزء دستوراتی است که در رشته های علوم انسانی خیلی کاربرد ندارد.

مثال: فرض کنید تعدادی کد (مثلا شماره 20000 نفر که میخواهید از آنها نمونه تصادفی ساده بگیرید). شماره ها را در یک ستون وارد کنید. سپس دستور را فراخوانی کنید در پنجره زیر میتوانید با وارد کردن مقادیر مربوطه و دلخواه اعداد تصادفی یا شماره های تصادفی را در خروجی ببینید.



تعداد متغیرها یا همان ستون هایی که میخواهید در خروجی ببینید را در Number Of variable وارد کنید.

تعداد اعداد تصادفی که نیاز دارید (مثلا 250 تا از 20 هزارتا) را در قسمت Number of Random Number وارد کنید.

با توجه به نوع توزیع (بیشتر نرمال) توزیع مربوطه را با نگاه به تعاریف بالا انتخاب کنید. هر توزیع برای خود در قسمت پارامترهایش پیش فرض دارد میتوانید آن را عوض کنید.

عددی را به دلخواه در قسمت Random Seed انتخاب کرده و در محل آن قرار دهید. شما از این عدد برای انتخاب های بعدی میتوانید استفاده کنید. به عبارت بهتر این عدد نوعی معیار برای استاندارد کردن همه انتخاب ها می باشد.

Rank and Percentile

رتبه و درصدگیری

این دستور میتواند برای داده هایی که وارد میشود ابته رتبه بندی کرده و سپس درصد گیری کند.

برای داده های زیر این دستور را اجرا میکنیم.

40 - 12 - 12 - 55 - 23- 10 - 12 - 32 - 32 - 54 - 23 - 24 - 10

با اجرای درست دستورات خروجی زیر را در صفحه اکسل خواهید دید.

Point	Case	Rank	Percent
10	55	1	100.00%
4	54	2	91.60%
13	40	3	83.30%
5	32	4	66.60%
6	32	4	66.60%
2	24	6	58.30%
3	23	7	41.60%
9	23	7	41.60%
7	12	9	16.60%
11	12	9	16.60%
12	12	9	16.60%
1	10	12	0.00%
8	10	12	0.00%

ستون اول (Point) رتبه های ساختگی یا مصنوعی را نشان میدهد. یعنی اگر دو عدد مشابه داشته باشیم (مثلا 32 یکی 5 و دیگری 6 رتبه گرفته است)، هر کدام رتبه جدا میگیرند. ستون دوم (Case) داده های خام را نشان میدهد. ستون سوم (Rank) رتبه های اولیه و خام را که واقعی هستند را نشان میدهد (یعنی دو عدد 32 هر کدام رتبه 4 دارند). ستون چهارم (Percent) نیز درصدها را نشان میدهد.

Regression

رگرسیون

از مباحث پیچیده و جالب آماری مبحث رگرسیون است. این که نرم افزار اکسل چقدر در این حیطة قابلیت دارد، مشخص نیست ولی میتوان از اکسل در حد ساده رگرسیون گیری انجام داد و گاهی مشکلات را از پیش رو برداشت.

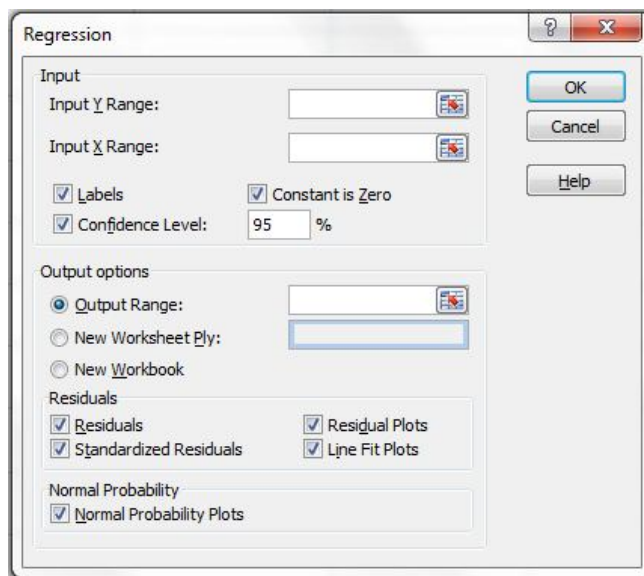
واژه رگرسیون در فرهنگ لغت به معنی بازگشت است و اغلب جهت رساندن مفهوم "بازگشت به یک مقدار متوسط یا میانگین" به کار می رود. بدین معنی که برخی پدیده ها به مرور زمان از نظر کمی به طرف یک مقدار متوسط میل می کنند.

وقتی تعدادی متغیر مستقل برای یک متغیر وابسته داریم، میخواهیم ببینیم که آیا میتوانیم با داشتن متغیرهای مستقل با مقادیر متفاوت، متغیر وابسته را پیش بینی کنیم. این مسئله به کرات در تحلیل‌های علوم انسانی و سایر علوم کاربرد دارد. براحتی میتوان از نرم افزارهای تخصصی همانند SPSS این نوع تحلیلها را انجام داد.

در این مرحله فقط میخواهیم عنوان کنیم که اکسل هم تا حدودی از عهده رگرسیون بر می آید. مثال: داده های زیر را با عطف به اینکه متغیر a وابسته و b ، c و d مستقل باشند، در نظر بگیرید.

	A	B	C	D	E
1					
2		a	b	c	d
3		1	3	2	1
4		2	4	1	2
5		3	2	3	3
6		4	3	2	4
7		5	2	4	2

دستور رگرسیون را اجرا کنید، تا پنجره زیر باز شود.



همانطور که ملاحظه می فرمایید:

Input Y Range: این پنجره جهت ورود متغیر وابسته است.

Input X Range: این پنجره جهت ورود متغیر یا متغیرهای وابسته تا حداکثر 16 متغیر میباشد.

Labels: اگر این گزینه را علامت بزنی ردیف اول یعنی نامها به درستی جای خود قرار گرفته و

اکسل خطا نمیدهد..

Confidence Level: سطح معناداری با پیش فرض 0.95 میباشد.

Constant is Zero: مقدار ثابت را صفر در نظر میگیرد.

Output Options: مکان خروجی را مشخص میکند.

Residuals: مقادیر باقیمانده را در خروجی نشان میدهد

Standardized Residuals: مقادیر باقیمانده استاندارد را در خروجی نشان میدهد

Residual Plots: نمودار مقادیر باقیمانده را رسم میکند.

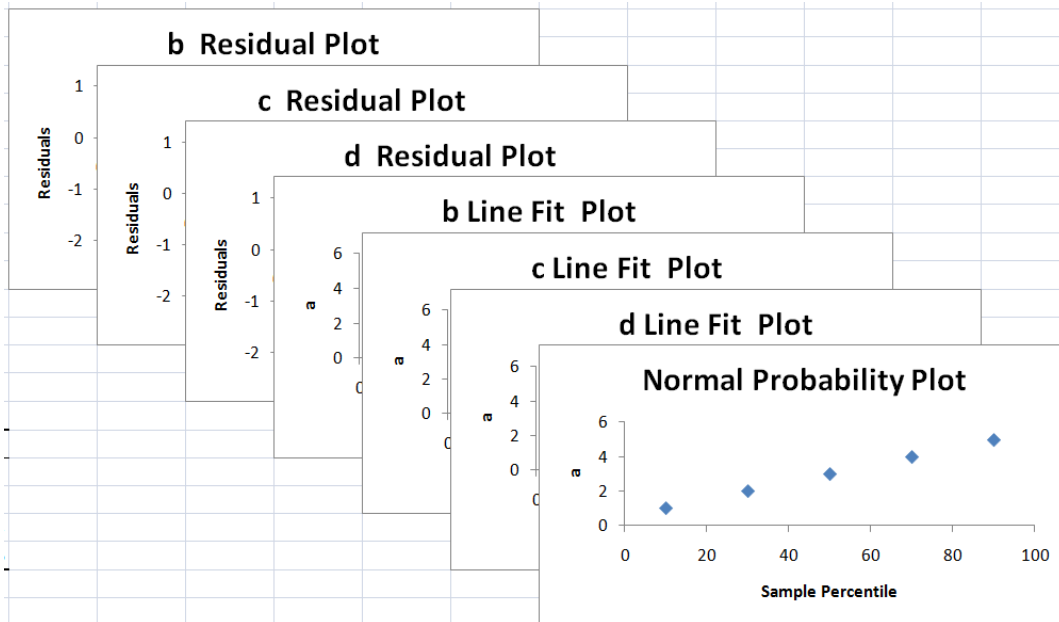
Line Fit Plots: جهت رسم نموداری برای مقادیر پیش بینی شده در برار مقادیر واقعی استفاده

میشود.

Normal Probability: نمودار احتمال نرمال را رسم میکند.

بعد از این که داده ها را وارد کردید، دستور را اجرا کنید باید خروجی زیر را ملاحظه کنید.

SUMMARY OUTPUT								
Regression Statistics								
Multiple R	0.975345427							
R Square	0.951298701							
Adjusted R Square	0.402597403							
Standard Error	1.157275125							
Observations	5							
ANOVA								
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>			
Regression	3	52.32142857	17.44047619	13.02222	0.200309405			
Residual	2	2.678571429	1.339285714					
Total	5	55						
	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>	<i>Lower 95.0%</i>	<i>Upper 95.0%</i>
Intercept	0	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A	#N/A
b	-0.125	0.37350894	-0.334664011	0.769717	-1.732079262	1.482079262	-1.732079262	1.482079262
c	0.767857143	0.387476684	1.981686061	0.186019	-0.899320467	2.435034753	-0.899320467	2.435034753
d	0.642857143	0.483573715	1.329388103	0.315083	-1.43779262	2.723506906	-1.43779262	2.723506906
RESIDUAL OUTPUT					PROBABILITY OUTPUT			
<i>Observation</i>	<i>Predicted a</i>	<i>Residuals</i>	<i>Standard Residuals</i>		<i>Percentile</i>	<i>a</i>		
1	1.803571429	-0.803571429	-1.097887582		10	1		
2	1.553571429	0.446428571	0.609937546		30	2		
3	3.982142857	-0.982142857	-1.3418626		50	3		
4	3.732142857	0.267857143	0.365962527		70	4		
5	4.107142857	0.892857143	1.219875091		90	5		

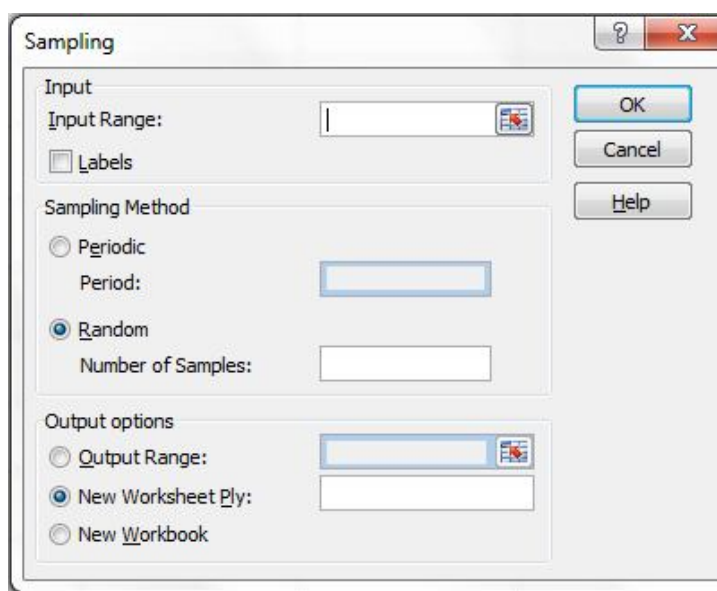


Sampling

نمونه گیری

این دستور شما را قادر میسازد تا از جمعیت نمونه ی تصادفی ساده بگیریید. وقتی حجم جامعه بزرگ است برای انجام برخی اعمال مثل رسم نمودار شما میتوانید با استفاده از این دستور نمونه ای گرفته و اعمال را بر روی آن انجام دهید.

داده های جمعیتی بزرگ را وارد اکسل کنید. دستور را اجرا کنید تا پنجره زیر باز شود.



داده ها را فراخوانی کنید. در قسمت Sampling Method نوع نمونه گیری را مشخص کنید. شما دو انتخاب دارید و فقط یکی را میتوانید انتخاب کنید: الف) با انتخاب گزینه Periodic (دوره ای) تعداد دوره ها را انتخاب میکنید. دوره همان دسته های کوچک نمونه در کل جمعیت است. ب) با انتخاب گزینه Random (تصادفی)، شما کفیس تعداد نمونه ای را که میخواهید انتخاب کنید. گاهی این دو روش اعداد متفاوتی را انتخاب میکنند.

t-Test: Paired Two Sample For Means

آزمون t: نمونه های جفت

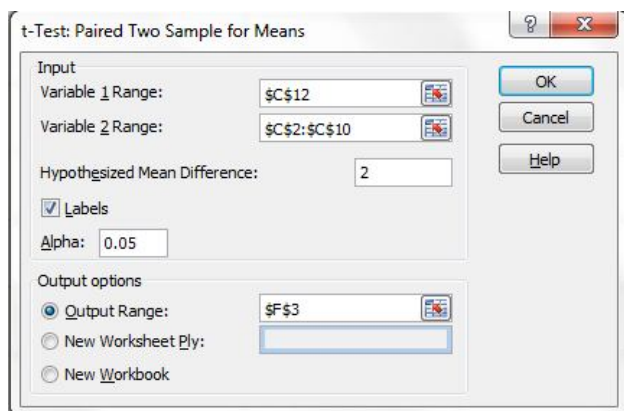
آزمونهای t زمانی استفاده میشود که میخواهیم دو میانگین را در دو جامعه باهم مقایسه کنیم. نوع اول زمانی است که گروه ها یا جوامع ما وابسته باشند. عبارت بهتر دو میانگینی که قرار است با هم مقایسه شوند از مشاهدات یک جامعه باشند. مثلاً، میانگین نمره آمار یک کلاس با میانگین نمره روش تحقیق همان کلاس.

36.....اکسل: مقدمه، تحلیل واریانس، رگرسیون و آزمون میانگینها

مثال: داده های زیر را وارد اکسل کنید. و با فرض این که، b و c مقادیر مشاهده شده وابسته برای افراد باشند را با هم مقایسه کنید.

a	b	c
1	2	1
2	3	3
1	1	2
3	4	5
1	5	4
4	2	3
5	1	2
3	4	1

پنجره زیر مربوط به این دستور است.



در قسمت Hypothesized Mean Difference میانگینی را که میخواهید با آن مقایسه شود وارد کنید.

با اجرای درست دستور خروجی زیر را مشاهده خواهید کرد.

t-Test: Paired Two Sample for Means		
	<i>a</i>	<i>b</i>
Mean	2.5	2.75
Variance	2.285714286	2.214285714
Observations	8	8
Pearson Correlation	-0.254000254	
Hypothesized Mean Difference	2	
df	7	
t Stat	-2.67903251	
P(T<=t) one-tail	0.015791752	
t Critical one-tail	1.894578604	
P(T<=t) two-tail	0.031583504	
t Critical two-tail	2.364624251	

با نظر اجمالی به خروجی ملاحظه میکنید که تمامی عناوین آشنا هستند.

t-Test: Two-Sample Assuming Equal Variances

آزمون t با فرض برابری واریانسها

حال اگر دو جامعه دارای واریانسهای برابر باشند از این آزمون استفاده میکنیم. برای همان داده ها با فرض (فقط فرض، یادگیری روش مهم است!) برابری واریانس این دستور را اجرا کنید. خروجی درست به صورت زیر است:

t-Test: Two-Sample Assuming Equal Variances		
	<i>a</i>	<i>b</i>
Mean	2.5	2.75
Variance	2.285714286	2.214285714
Observations	8	8
Pooled Variance	2.25	
Hypothesized Mean Difference	2	
df	14	
t Stat	-3	
P(T<=t) one-tail	0.004775756	
t Critical one-tail	1.761310115	
P(T<=t) two-tail	0.009551513	
t Critical two-tail	2.144786681	

t-Test: Two-Sample Assuming Unequal Variances

آزمون t با فرض عدم برابری واریانسها

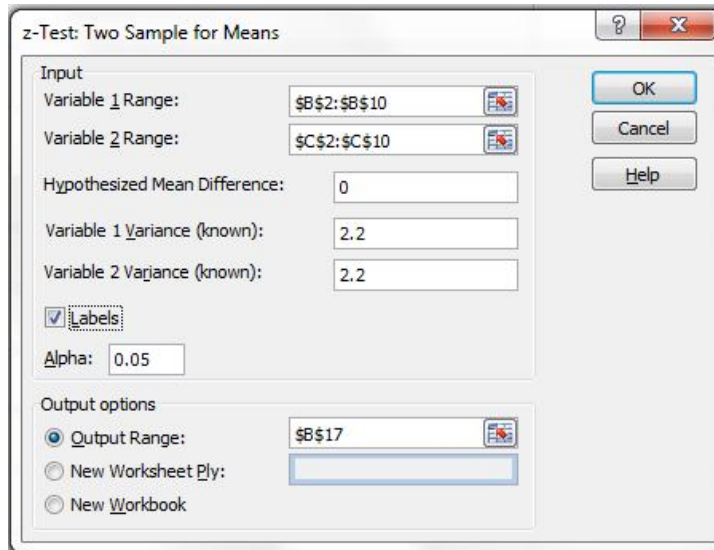
و اگر واریانسها برابر نباشند هم از این آزمون برای مقایسه میانگینها سود می جوییم. برای مثال بالا دستور را با فرض نابرابری واریانس اجرا کنید. اگر دستورات را به درستی انجام دهید با خروجی زیر روبرو خواهید شد:

t-Test: Two-Sample Assuming Unequal Variances		
	<i>a</i>	<i>b</i>
Mean	2.5	2.75
Variance	2.285714	2.214286
Observations	8	8
Hypothesized Mean Difference	2	
df	14	
t Stat	-3	
P(T<=t) one-tail	0.004776	
t Critical one-tail	1.76131	
P(T<=t) two-tail	0.009552	
t Critical two-tail	2.144787	

z-Test

آزمون z

این آزمون زمانی است که حجم دو جامعه بزرگ (بیشتر از 30) باشد. با فرض کنید که دو جامعه *a* و *b* حائز این شرط باشند. با اجرای این دستور پنجره زیر باز میشود:



در قسمت های Variable 1 & 2 Variance(known) واریانسی را که مربوط به جمعیت معلوم است، را وارد کنید.
خروجی به صورت زیر است:

z-Test: Two Sample for Means		
	<i>a</i>	<i>b</i>
Mean	2.5	2.75
Known Va	2.2	2.2
Observatic	8	8
Hypothesi	0	
z	-0.3371	
P(Z<=z) on	0.368021	
z Critical o	1.644854	
P(Z<=z) tw	0.736042	
z Critical tv	1.959964	

پایان

موفق و پیروز باشید