

MÓDULO 2: TRATAMIENTO DE DATOS CON HOJA DE CÁLCULO

Tema 6: Análisis Estadístico de Datos

Leire Aldaz, Begoña Eguía y Leire Urcola



Índice del tema

Introducción

Herramientas para Análisis

Complemento EZAnalyze

INTRODUCCIÓN

Las aplicaciones que permiten manejar hojas de cálculo, como Microsoft Excel u OpenOffice.org Calc, ofrecen funciones que ayudan a realizar análisis estadísticos de los datos mediante promedios, modas, medianas, máximos, varianzas...

El cálculo de cada uno de estos estadísticos implica, sin embargo, una tarea un tanto laboriosa, por cuanto debe realizarse de forma individual para cada una de las funciones de interés. No obstante, en la actualidad existen complementos que disponen de herramientas que pueden facilitar en gran medida la obtención “inmediata” de todos estos estadísticos. Ejemplos son:

- Para Microsoft Excel: el complemento *Herramientas para Análisis* desarrollado por Microsoft o el complementos *EZAnalyze* desarrollado por Tim Poynton en 2004-2005, que permiten ampliar las prestaciones del programa para el análisis estadístico de datos.
- Para OpenOffice.org Calc: la macro *OOo Statistics* descargable de la página <http://sourceforge.net/projects/oomacros/files/OOo%20Statistics/>

A continuación se van a describir brevemente los citados dos complementos que pueden emplearse en Microsoft Excel: *Herramientas para Análisis* y *EZAnalyze*.

HERRAMIENTAS PARA ANÁLISIS

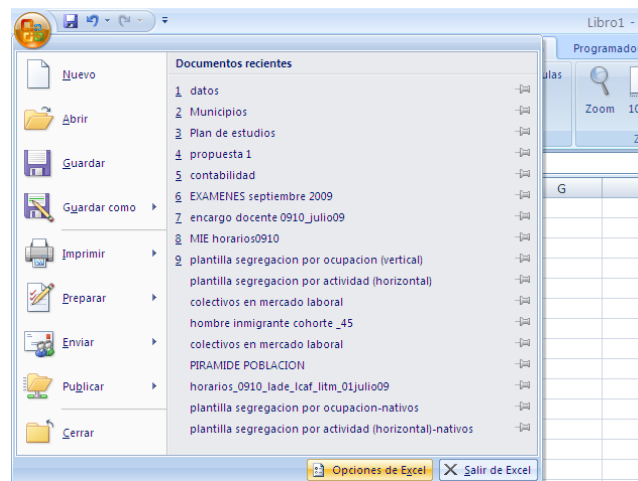
Microsoft Excel ofrece un conjunto de herramientas que facilitan el análisis estadístico de datos. Estas herramientas están disponibles cuando se instala el programa, aunque para usarlas deben ser previamente cargadas. El procedimiento es muy sencillo:



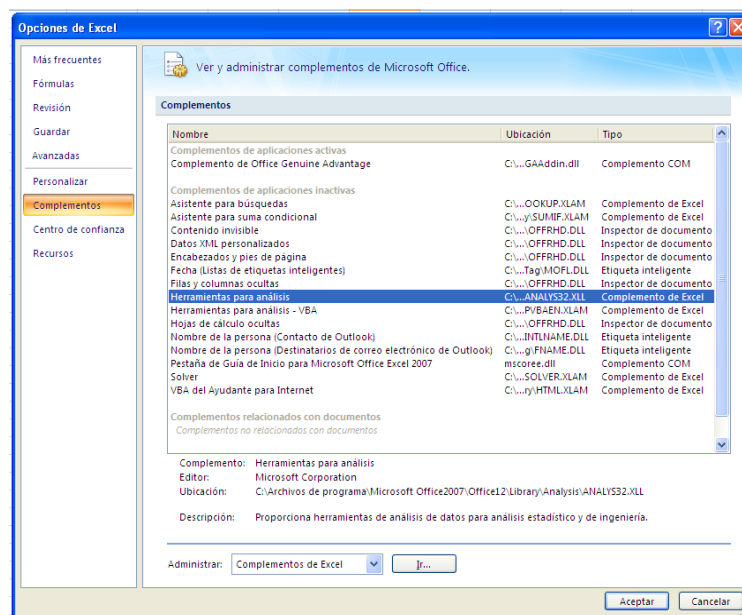
Módulo 2: Tratamiento de Datos con Hoja de Cálculo

Tema 6: Análisis Estadístico de Datos

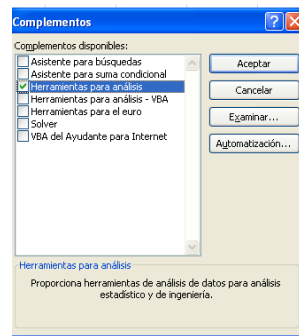
Hacer clic en el botón *Opciones de Excel* que se encuentra al desplegar el *Botón de Office*



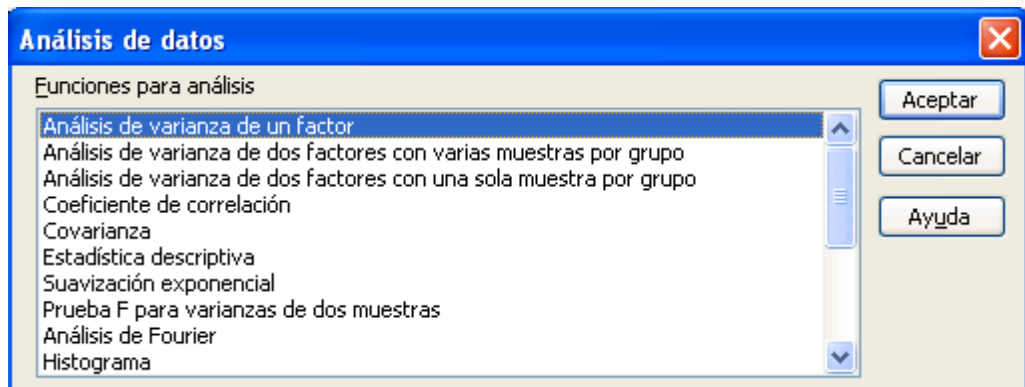
De entre las opciones, elegir *Complementos*. En el cuadro *Administrar* que se encuentra en la parte inferior, seleccionar *Complementos de Excel*, y hacer un clic en *Ir*.



Aparece un nuevo cuadro que muestra los complementos disponibles que pueden ser cargados: debe activarse la casilla de verificación *Herramientas para análisis*. Finalmente debe hacerse un clic en *Aceptar*.



De esta manera el complemento *Herramientas para análisis* queda cargado y está disponible para ser utilizado en el comando *Análisis de datos* del grupo *Análisis* de la ficha *Datos* de la cinta de opciones. Al seleccionarlo aparece la ventana:



donde se encuentran todas las opciones disponibles. En particular, las herramientas de análisis estadístico que proporciona este complemento son las siguientes:

- Análisis de Varianza
 - Varianza de un factor
 - Varianza de dos factores con varias muestras por grupo
 - Varianza de dos factores con una sola muestra por grupo
- Correlación
- Covarianza
- Estadística descriptiva
- Suavización exponencial
- Prueba t para varianzas de dos muestras
- Análisis de Fourier
- Histograma
- Media móvil
- Generación de números aleatorios



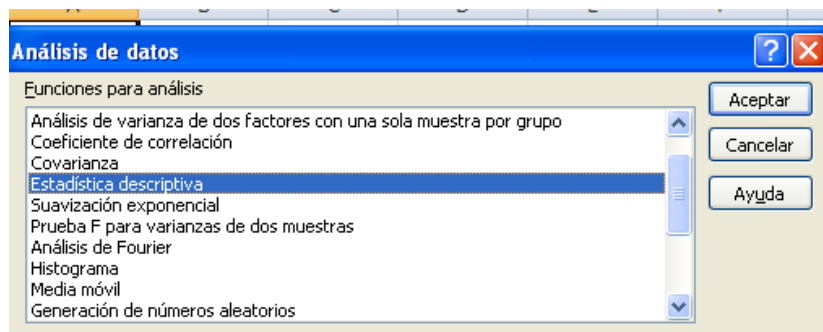
- Jerarquía y percentil
- Regresión
- Muestreo
- Prueba t
 - Prueba t para medias de dos muestras emparejadas
 - Prueba t para dos muestras suponiendo varianzas iguales
 - Prueba t para dos muestras suponiendo varianzas desiguales
- Prueba z

Una breve descripción de cada una de estas herramientas de análisis se puede encontrar [aquí](#).

Ejemplo: Se dispone de observaciones relativas al consumo y renta semanal de diez familias. Se pretende, primero, realizar un análisis descriptivo de las series. Y, segundo, estimar un modelo de regresión donde se trate de explicar el consumo en función de la renta.

	A	B
1	Consumo	Renta
2	165	240
3	264	300
4	270	360
5	240	420
6	354	480
7	360	540
8	435	600
9	405	660
10	435	720
11	525	780

Haciendo uso del comando *Análisis de datos* del grupo *Análisis* de la ficha *Datos* de la cinta de opciones, aparece una ventana:

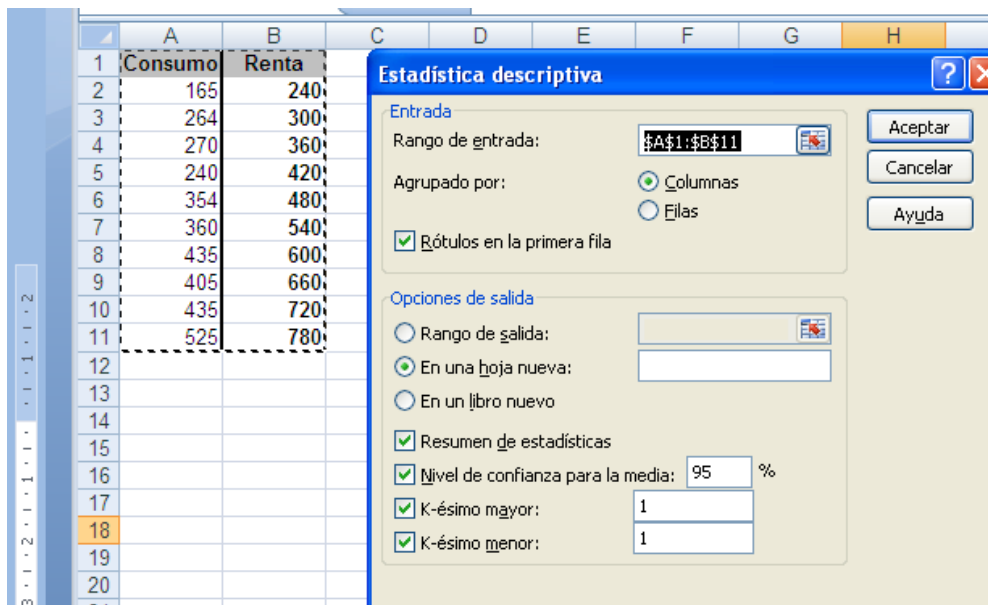


donde habría que elegir *Estadística Descriptiva*:



Módulo 2: Tratamiento de Datos con Hoja de Cálculo

Tema 6: Análisis Estadístico de Datos



Seleccionando el rango de celdas que ocupan las series, indicando su disposición (en columnas), mostrando dónde colocar los resultados y eligiendo las opciones que se quieren ver mostradas, aparece:

	A	B	C	D
1	Consumo		Renta	
2				
3	Media	345,3	Media	510
4	Error típico	34,6596307	Error típico	57,4456265
5	Mediana	357	Mediana	510
6	Moda	435	Moda	#N/A
7	Desviación estándar	109,603376	Desviación estándar	181,659021
8	Varianza de la muestra	12012,9	Varianza de la muestra	33000
9	Curtosis	-0,64816526	Curtosis	-1,2
10	Coefficiente de asimetría	-0,05592265	Coefficiente de asimetría	0
11	Rango	360	Rango	540
12	Mínimo	165	Mínimo	240
13	Máximo	525	Máximo	780
14	Suma	3453	Suma	5100
15	Cuenta	10	Cuenta	10
16	Mayor (1)	525	Mayor (1)	780
17	Menor(1)	165	Menor(1)	240
18	Nivel de confianza(95,0%)	78,4055317	Nivel de confianza(95,0%)	129,951035

que muestra un conjunto de estadísticos para ambas series de datos.

Si además interesase la correlación entre ambas variables, se elegiría *Coefficiente de Correlación* de las funciones de Análisis de datos:



Módulo 2: Tratamiento de Datos con Hoja de Cálculo

Tema 6: Análisis Estadístico de Datos

The screenshot shows an Excel spreadsheet with columns A and B containing data for 'Consumo' and 'Renta' respectively. The 'Análisis de datos' dialog box is open, and 'Coficiente de correlación' is selected in the list of functions for analysis.

	A	B
1	Consumo	Renta
2	165	240
3	264	300
4	270	360
5	240	420
6	354	480
7	360	540
8	435	600
9	405	660
10	435	720
11	525	780

y, tras indicarle el rango de entrada y de salida:

The screenshot shows the 'Coficiente de correlación' dialog box. The 'Entrada' section has 'Rango de entrada' set to '\$A\$1:\$B\$11' and 'Agrupado por' set to 'Columnas'. The 'Opciones de salida' section has 'En una hoja nueva' selected.

	A	B
1	Consumo	Renta
2	165	240
3	264	300
4	270	360
5	240	420
6	354	480
7	360	540
8	435	600
9	405	660
10	435	720
11	525	780

arrojaría como resultado:

	A	B	C
1		Consumo	Renta
2	Consumo	1	
3	Renta	0,95477539	1

Finalmente, la estimación MCO de un modelo que explique el consumo en función de la renta se podría obtener directamente eligiendo la función *Regresión* de la ventana de *Análisis de datos*:

The screenshot shows the 'Análisis de datos' dialog box with 'Regresión' selected in the list of functions for analysis.

	A	B
1	Consumo	Renta
2	165	240
3	264	300
4	270	360
5	240	420
6	354	480
7	360	540
8	435	600
9	405	660
10	435	720
11	525	780

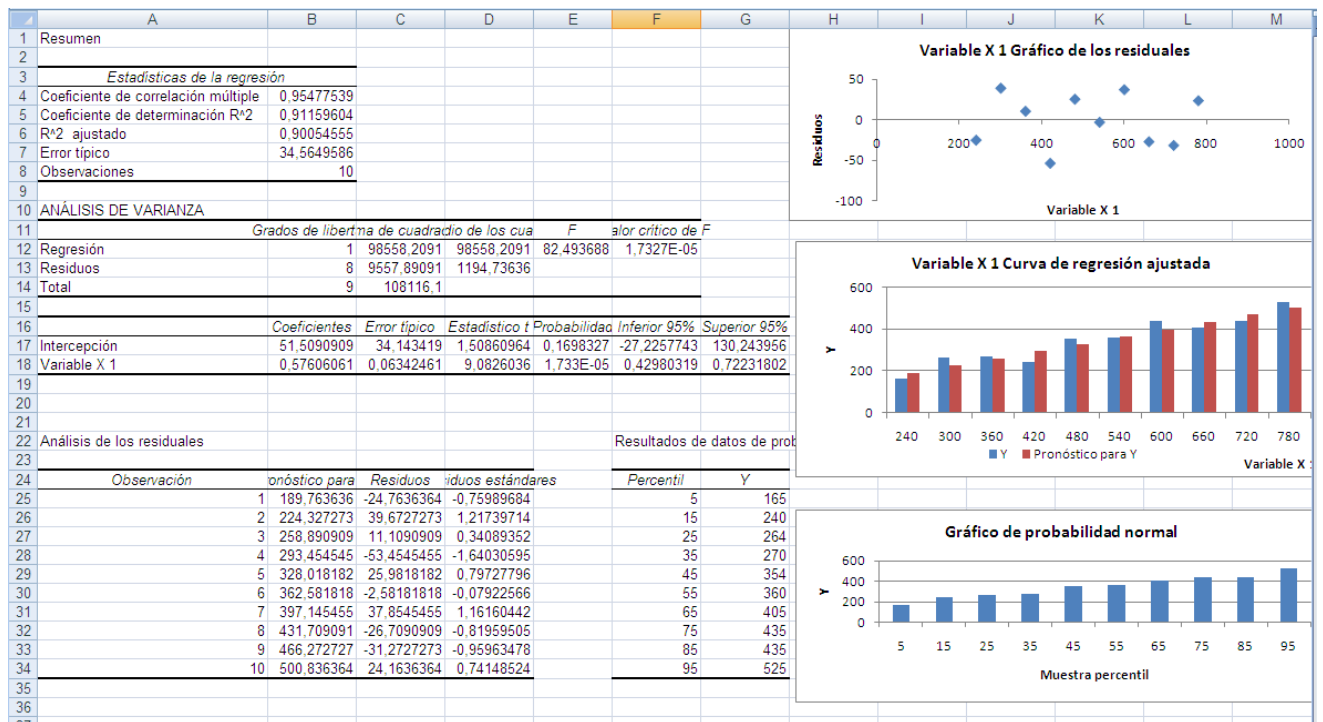


Módulo 2: Tratamiento de Datos con Hoja de Cálculo

Tema 6: Análisis Estadístico de Datos

donde habría que seleccionar el rango que ocupa la variable a explicar (consumo) y la variable explicativa (renta) e indicar si el modelo debe o no incluir una constante. Además da la posibilidad de que en los resultados aparezcan listados los residuos, mostrando el gráfico de éstos, representada la recta de regresión muestral...

Tras elegir las opciones deseadas se obtiene como output:





Este ha sido un ejemplo de uso de tres herramientas de análisis que proporciona este complemento. Sin embargo, la forma de proceder en todas ellas es similar: proporcionando una serie de datos y parámetros para cada análisis y eligiendo una serie de opciones, la herramienta realiza los cálculos y muestra los resultados en una tabla que, en algunas ocasiones, viene acompañada de gráficos.

COMPLEMENTO EZAnalyze

EZAnalyze es un complemento de Microsoft Excel, desarrollado por Tim Poynton en 2004-2005, que permite ampliar las prestaciones del programa para el análisis estadístico de datos y la creación de gráficos.

Este complemento puede descargarse e instalarse directamente de la página de EZAnalyze <http://www.ezanalyze.com/download/>:

EZAnalyze
Data Tools for Educators

Google Custom Search Go

Products Documents Support Donate About

Download Page

The files below are available for free download. If clicking on the link does not download the file, you may need to right-click on the link (or hold the control key and click on the Mac), and select 'Save Link As' or 'Save Target As' from the menu.

EZAnalyze Program Files

[EZAnalyze Version 3.0. Automated Installer \(1.30MB\)](#) - An Excel worksheet file (EZAnalyze3.xls) that automates the installation of EZAnalyze in Excel. Recommended for use on most personal computers. NOT recommended for use in computer labs or on computers you do not have the 'rights' to install new programs on. Installation Instructions are minimal (English Only). Last Updated 9/18/2007.

[EZAnalyze Version 3.0, without Automated Installer \(1.30MB\)](#) - An Excel Add In File (EZAnalyze3.xla) that needs to be 'manually' installed according to the Installation Instructions. Recommended for use on all computers (English Only). Last Updated 9/18/2007.

[EZAnalyze Version 2.5 \(1.10MB\)](#) - An Excel Add In File (ezanalyze.xla) that needs to be 'manually' installed according to the Installation Instructions. Recommended for use by people who speak Spanish, Serbian, or Hungarian as these languages are included.

Other Related Files

[EZAnalyze User Manual and Tutorials \(1.15MB\)](#) - A PDF File (EZAnalyze_version3.pdf) for use with the EZAnalyze program file. (English Only)

[EZAnalyze Sample Data \(30KB\)](#) - An Excel workbook file (EZAsample_data.xls) that can be used in conjunction with the User Manual and Tutorials.

NOTES

You will remain on this page after the file has downloaded. If the program does not work for you, check the file size to make sure your download was complete.

I hope you find EZAnalyze and the related documentation useful. If you would like to give back to the cause, please feel free to make a monetary donation, or donate your time to develop needed documentation.

Solar Energy Charity

Helping Relieve Poverty Through The Provision of Solar Energy. Join Us!

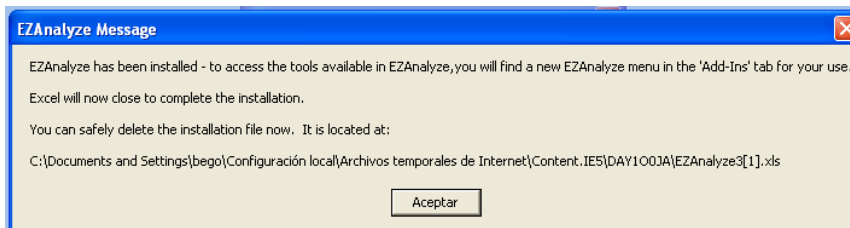
Public Service Ads by Google

Una vez instalado aparece el mensaje:

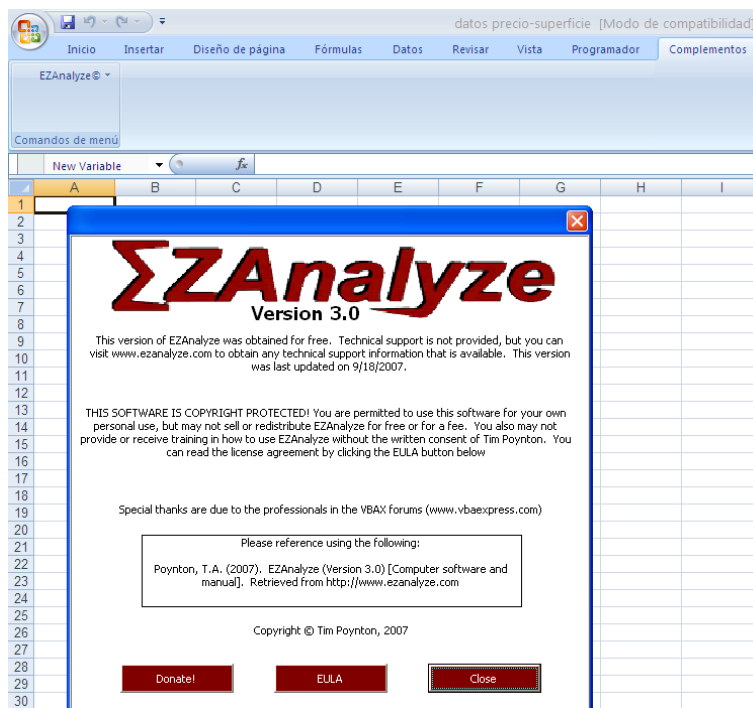


Módulo 2: Tratamiento de Datos con Hoja de Cálculo

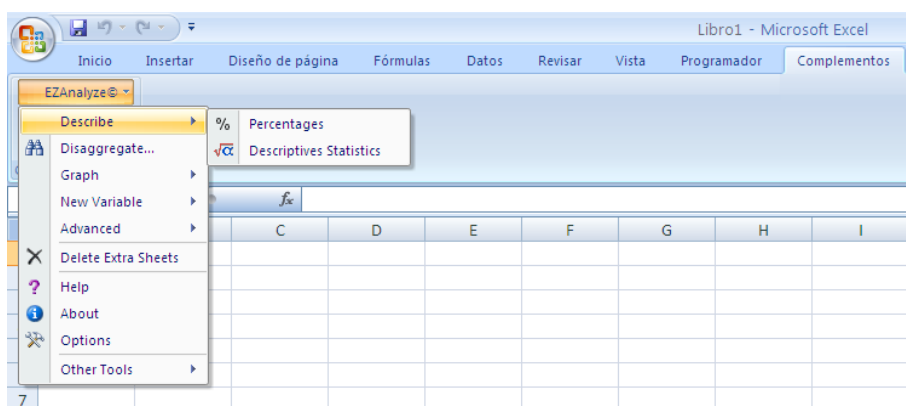
Tema 6: Análisis Estadístico de Datos



Cuando se abra de nuevo Microsoft Excel aparecerá una nueva ficha en la cinta de opciones rotulada como Complementos:



En esta ficha aparece el comando que da acceso a todas las prestaciones que ofrece *EZAnalyze*:



Para poder trabajar con este complemento los datos tienen que tener una distribución concreta: en la primera fila de la hoja deben aparecer los nombres de las



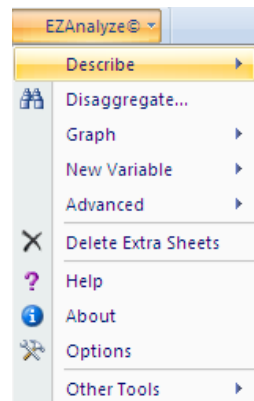
Módulo 2: Tratamiento de Datos con Hoja de Cálculo

Tema 6: Análisis Estadístico de Datos

variables a analizar; y los datos asociados a tales variables para cada elemento analizado se deben encontrar en las filas siguientes. Ejemplo:

	A	B	C	D	E
1	Precio	Superficie	Antigüedad	Huecos	
2	319.840	49.47	16	4	
3	364.800	58.25	73	6	
4	376.000	60.39	48	6	
5	456.000	73.25	1	5	
6	382.400	74.32	38	5	
7	468.800	81.29	68	6	
8	456.000	83.61	12	6	
9	584.000	86.86	68	6	
10	472.000	89.88	32	8	
11	464.000	90.49	31	5	
12	616.000	104.70	68	6	
13	808.000	120.77	28	5	
14	680.000	130.06	8	6	
15	664.000	139.35	22	6	
16					

Una vez que los datos están bien estructurados, se puede proceder a efectuar el análisis estadístico utilizando una de las funciones que ofrece *EZAnalyze*:



- ✓ **Describe:** proporciona porcentajes y estadísticos descriptivos.
- ✓ **Disaggregate:** permite clasificar los resultados en categorías.
- ✓ **Graph:** crea gráficos básicos.
- ✓ **New Variable:** permite crear nuevas variables a partir de las ya existentes (como suma o promedio, diferencias entre ellas.....).
- ✓ **Advanced:** proporciona opciones más avanzadas de estadística, como el cálculo de correlaciones entre las variables, el análisis de varianza o el análisis de regresión lineal, entre otros.
- ✓ **Delete Extra Sheets:** borra todas las hojas creadas por *EZAnalyze* que se han ido creando al analizar los datos. Por defecto, el nombre de estas hojas empieza por EZA.



Módulo 2: Tratamiento de Datos con Hoja de Cálculo

Tema 6: Análisis Estadístico de Datos

- ✓ **Help:** da acceso a la ayuda online de *EZAnalyze*, que se encuentra disponible en <http://www.ezanalyze.com/help/>
- ✓ **About:** contiene información sobre la versión de *EZAnalyze* que se está utilizando y sobre los términos del acuerdo de licencia.
- ✓ **Options:** da la posibilidad de desinstalar *EZAnalyze* y de cambiar la forma de desplegar las opciones de los menús (en cascada o de forma tradicional).
- ✓ **Other Tools:** contiene funciones que permiten fusionar datos y transponer datos.

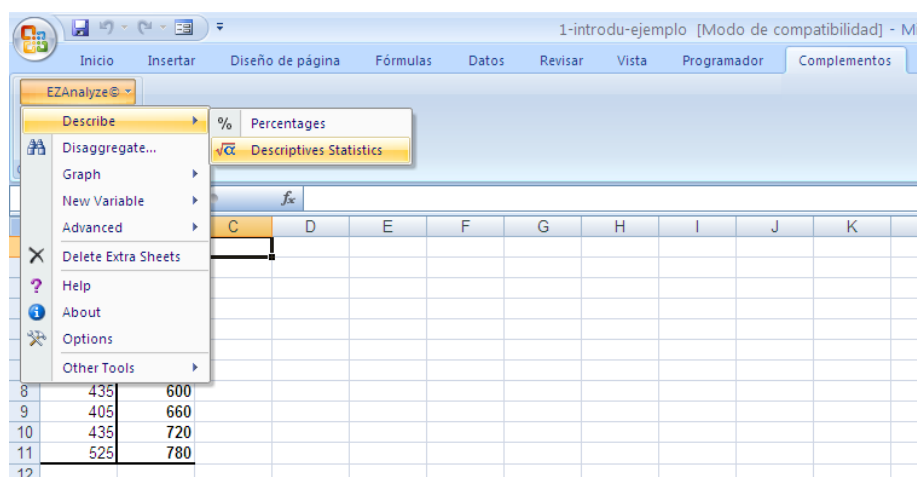
Elegida la función más idónea para el tema tratado, se debe seleccionar la variable o variables implicadas en el análisis. Los resultados de cada función elegida aparecerán en hojas de cálculo nuevas e independientes.

Información detallada sobre este complemento de Microsoft Excel puede obtenerse en el manual de usuario y tutoriales de Tim Poynton ([clic aquí](#)).

Ejemplo: Se dispone de observaciones relativas al consumo y renta semanal de diez familias. Se pretende, primero, realizar un análisis descriptivo de las series. Y, segundo, estimar un modelo de regresión donde se trate de explicar el consumo en función de la renta.

	A	B
1	Consumo	Renta
2	165	240
3	264	300
4	270	360
5	240	420
6	354	480
7	360	540
8	435	600
9	405	660
10	435	720
11	525	780

Haciendo uso del comando *Descriptives Statistics* del grupo *EZAnalyze* de la ficha *Complementos* de la cinta de opciones:

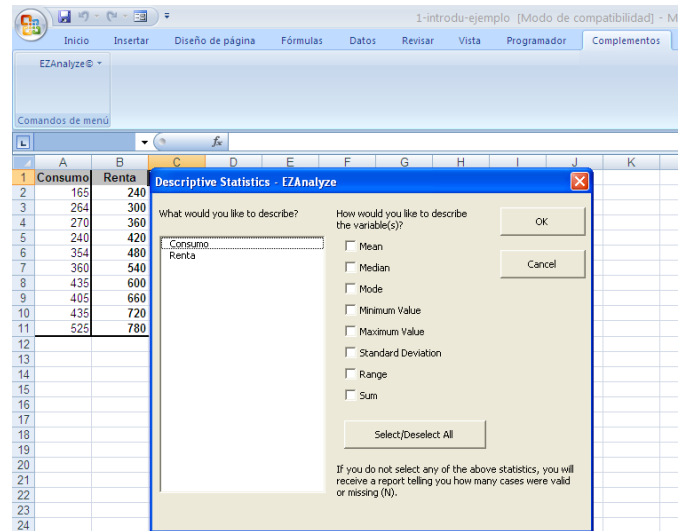




Módulo 2: Tratamiento de Datos con Hoja de Cálculo

Tema 6: Análisis Estadístico de Datos

aparece una ventana:



donde se debe indicar las variables que intervienen en el análisis y los estadísticos que se desean calcular para esas variables. Seleccionado esto, aparece una nueva hoja, denominada EZA1 con los resultados:

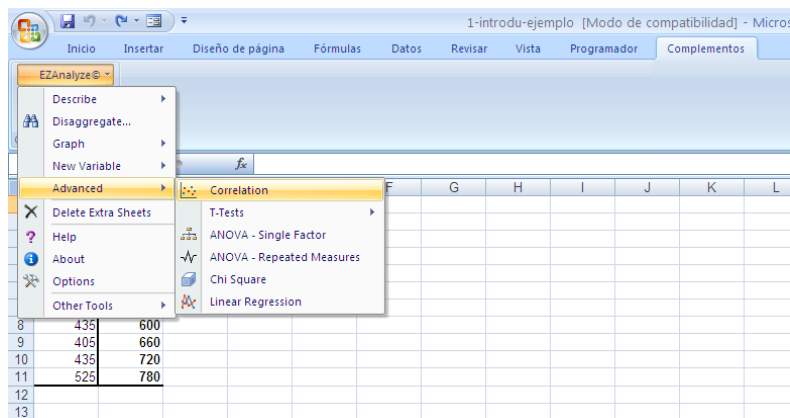
	Consumo	Renta
N Valid:	10	10
N Missing:	0	0
Mean:	345,300	510,000
Median:	357,000	510,000
Mode:	435,000	
Std. Dev:	109,603	181,659
Range:	360,000	540,000
Minimum Value:	165,000	240,000
Maximum Value:	525,000	780,000
Sum:	3453,000	5100,000

Si además interesase la correlación entre ambas variables, se elegiría *Correlation* de la Opción *Advanced* del complemento *EZAnalyze*:

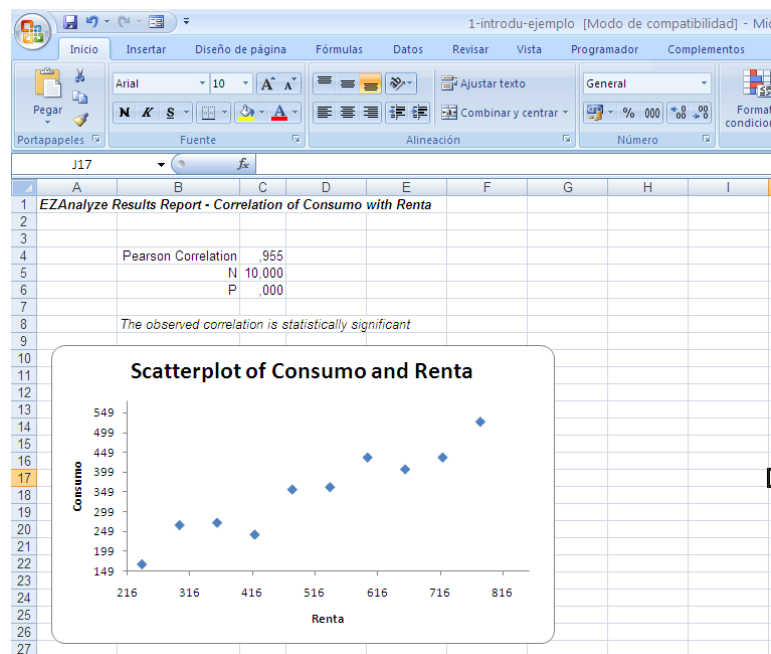


Módulo 2: Tratamiento de Datos con Hoja de Cálculo

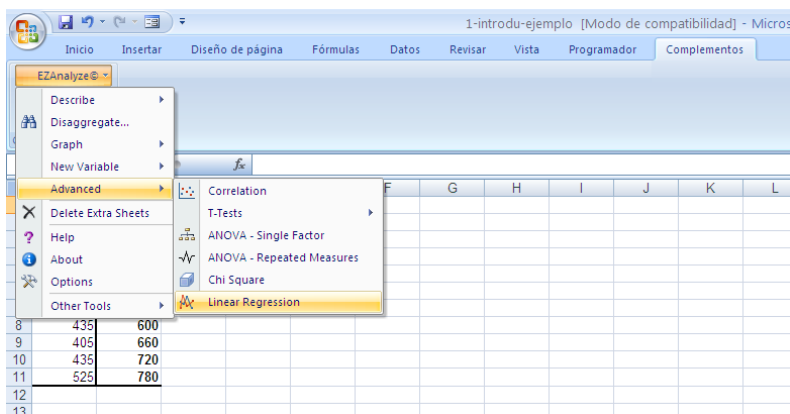
Tema 6: Análisis Estadístico de Datos



Obteniendo como resultado:



Finalmente, la estimación MCO de un modelo que explique el consumo en función de la renta se podría obtener directamente eligiendo *Linear Regresión* del comando *Advanced* del complemento *EZAnalyze*:

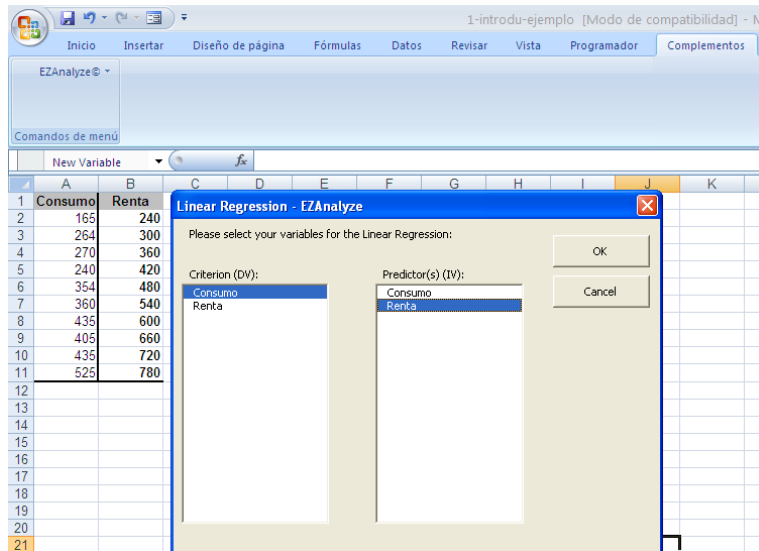




Módulo 2: Tratamiento de Datos con Hoja de Cálculo

Tema 6: Análisis Estadístico de Datos

donde se debe indicar cuál es la variable a explicar (*criterion*) y cuales las variables explicativas (*predictor*):



Y como hoja de resultados se obtiene:

