

# Estadística

# Contenido

- 1. Introducción
  - ¿Que es la estadística?
  - Modelo estadístico
  - Método estadístico
  - Algunas herramientas de análisis de datos mediante estadística
- 2. Datos
  - Descripción de una sola variable
  - Descripción conjunta de varias variables
- 3. Modelos
  - Probabilidad y variables aleatorias
  - Modelos de distribución de probabilidad
  - Modelos Multivariantes
- 4. Inferencia estadística
  - Estimación puntual
  - Estimación por intervalos
  - Estimación bayesiana
  - Contraste de hipótesis

# Introducción

- **Introducción**
  - ¿Que es la estadística?
  - Modelo estadístico
  - Método estadístico
  - Construir un Modelo estadístico
  - Algunas herramientas de análisis de datos mediante estadística (además de Python): Matlab y SPSS

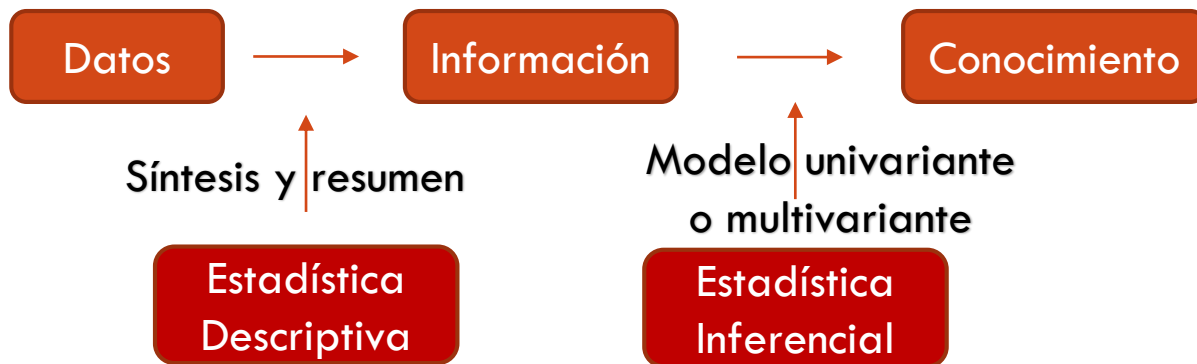
# Introducción

- ¿Que es la estadística?
  - Es la ciencia que estudia como obtener conclusiones de la investigación empírica mediante el uso de modelos matemáticos.
  - Actúa como vínculo en los modelos matemáticos y los fenómenos reales.
  - Un modelo matemático es una abstracción simplificada de algún aspecto de la realidad más compleja.
    - Siempre existen **discrepancias** entre el modelo y lo que representa este en la realidad.
  - La estadística proporciona una metodología para medir esas discrepancias (fundamental en ciencia aplicada, como tecnología, economía, sociología, medicina, y hoy en día en Big Data, etc.).
  - Es la **tecnología** del método científico experimental.

# Introducción

- ¿Que problemas puede resolver?
  - **Descripción de datos:** permite resumir la información contenida en los datos.
  - **Análisis de muestras:** No es posible en algunos caso estudiar todos los elementos de una población. Así dada una muestra representativa se pueden hacer inferencias respecto la población completa. La inferencia se puede realizar mediante un modelo probabilístico obtenido de la muestra. Hoy en día con Big Data, casi se puede tener una población.
  - **Contrastación de hipótesis:** En investigación empírica es fundamental la contrastación de hipótesis. Nos permiten comparar las predicciones resultantes de la hipótesis con los datos observados.
  - **Medición de relaciones:** las relaciones que observamos entre variables físicas, sociales y técnicas son prácticamente siempre estadísticas, por los errores de medición.
  - **Predicción:** el estudio de la historia de variables estadísticas permite inferir valores futuros.

# Introducción: Método estadístico



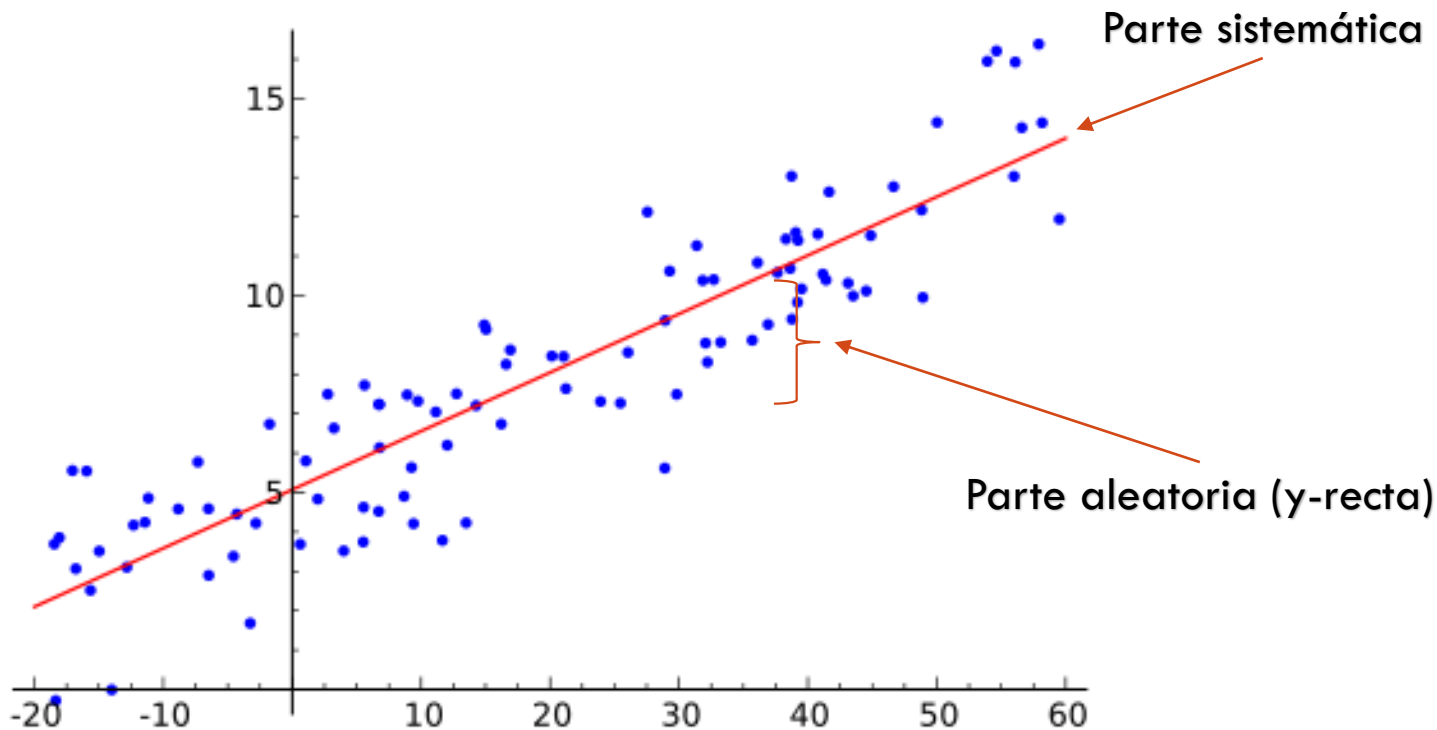
- **Estadística descriptiva:** recolecta, ordena, analiza y representa a un conjunto de datos, con el fin de describir apropiadamente las características de este y así resumirlos.
  - Calcula que parámetros estadísticos el conjunto de datos: gráficos, medidas de frecuencias, centralización, posición, dispersión, etc.
- **La estadística inferencial:** saca conclusiones generales para toda la población a partir del estudio de una muestra (para un grado de fiabilidad o significación de los resultados obtenidos).
  - Utiliza herramientas como el muestreo, la estimación de parámetros, el contraste de hipótesis, etc.

# Introducción: Modelo estadístico

- El Modelo estadístico **descompone** una variable respuesta de estudio en dos valores:
  - Parte predecible o sistemática
  - Parte aleatoria, impredecible o residual.
- Así el modelo estadístico define la forma de la **parte predecible** (respuesta media), y la **variabilidad de la impredecible** respecto a la respuesta media.
  - Observación=parte sistemática (predecible) + parte aleatoria (impredecible).
- En un problema real, en definitiva un variable respuesta  $y$  (o varias  $y_i$  en el caso multivariante), será función de otras variables de número desconocido (incluso no observables), siendo el modelo estadístico un acercamiento aproximación de esta realidad.
  - Esta aproximación tiene en cuenta las variables observables presumiblemente más importantes y engloba en la parte aleatoria los efectos del resto de variables que no se tienen en cuenta.
- Para el cálculo de **modelos estadísticos** se suele utilizar **distribuciones de probabilidad**.

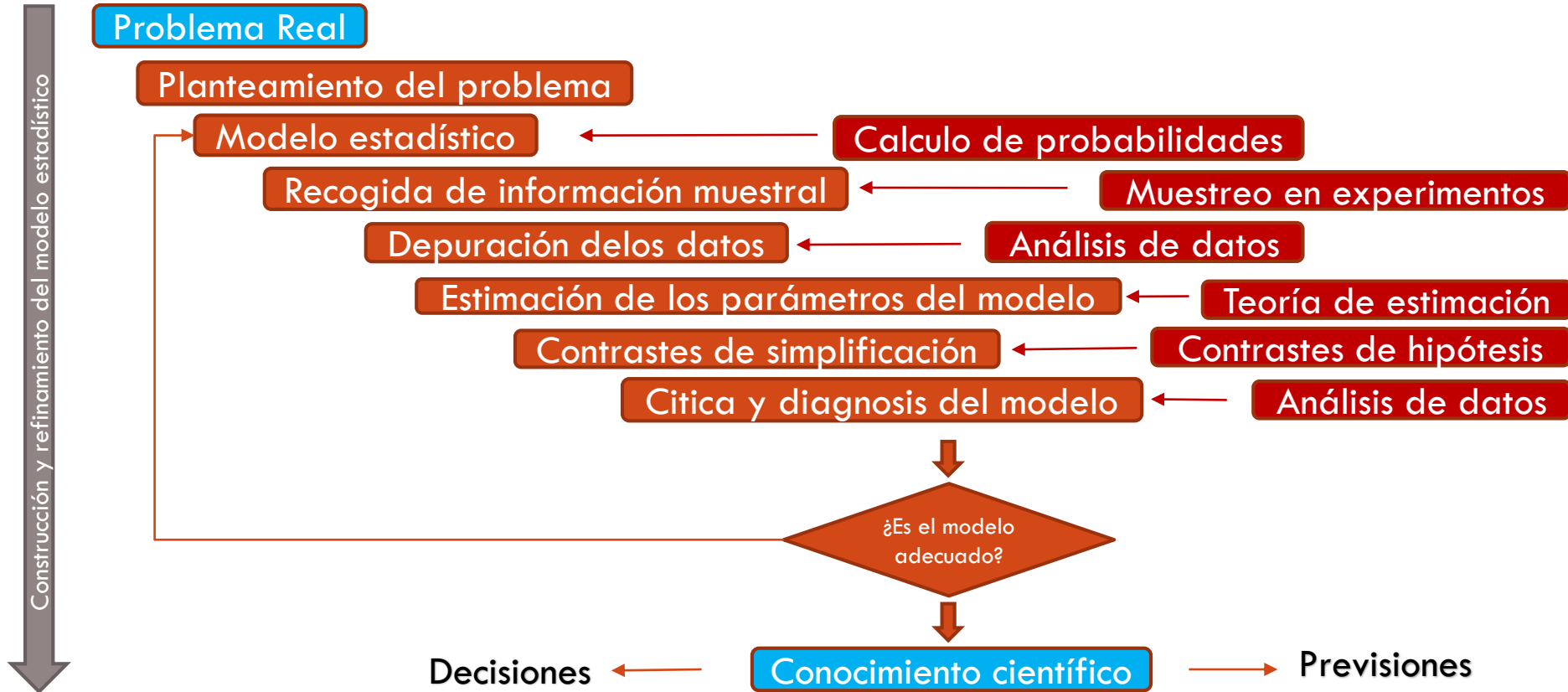
# Introducción: Método estadístico

Parte de la foto sacada de  
[https://en.wikipedia.org/wiki/Linear\\_regression](https://en.wikipedia.org/wiki/Linear_regression)





# Introducción: Etapas para construir un Modelo estadístico



# Introducción: Metodología Estadística y Big Data

- El Big Data empieza a cambiar la metodología vista anteriormente de estadística:

- Según lo que hemos mostrado anteriormente, generalmente tenemos un problema real que queremos resolver, así empezamos a tomar datos mediante un muestreo de una población definida orientados a resolver ese problema. Es decir no tenemos los datos al principio.
- Sin embargo en Big Data, tenemos los datos al principio y empezamos a hacer un análisis exploratorio para aprender sus estructuras.

- Generalmente en estadística proponemos un modelo para hacer inferencia respecto a unos parámetros del modelo que ajusto con la población.
- Con Big Data es muy difícil hacer esto, no sabemos como parametrizar en algunos casos.

- En estadística es muy importante como extraer de la muestra la máxima información (eficiencia estadística) que me pueda dar a conocer la población de la que proviene esa muestra.
- En Big Data no es tan importante la eficiencia estadística ya que en algunos casos tenemos casi la población.

# Introducción: Algunas Herramientas - Matlab

- Introducción
- Tipos de elementos
- Tipos de estructuras de datos
- Importar ficheros cvs
- Histogramas
- Medidas de centralidad y posición
- Medidas de forma
- Diagramas de caja (boxplot)
- Q-Qplots
- Diagramas de dispersión e histogramas

# Introducción

- Matlab (abreviatura de MATrix LABoratory, "laboratorio de matrices") es un software matemático comercial, cuya función principal es análisis de datos:
  - Entorno de desarrollo integrado (IDE).
  - Lenguaje de programación propio (lenguaje M).
- ¿Qué puede hacer Matlab?
  - Manipulación de matrices
  - representación de datos y funciones
  - implementación de algoritmos
  - creación de interfaces de usuario (GUI)
  - Comunicación con programas en otros lenguajes y con otros dispositivos hardware.
- MATLAB dispone de dos herramientas adicionales:
  - Simulink (plataforma de simulación multidominio)
  - GUIDE (editor de interfaces de usuario - GUI)
  - Paquetes para ampliar las capacidades de Matlab:
    - Cajas de herramientas (toolboxes)
  - Y las de Simulink:
    - Paquetes de bloques (blocksets)
- Es un software muy usado en universidades, centros de investigación y desarrollo.

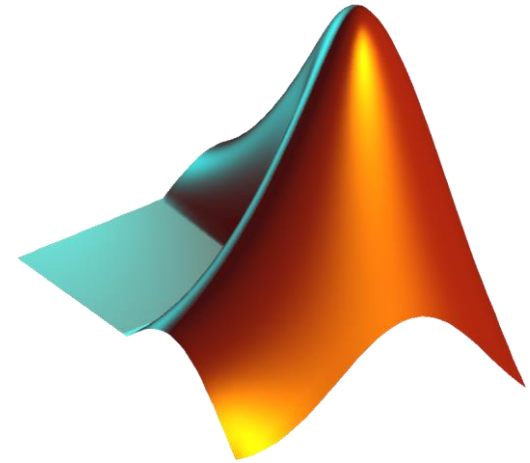


Imagen extraída de  
<https://es.wikipedia.org/wiki/MATLAB>

# Introducción

- Por ejemplo, la Toolbox Statistics proporciona un conjunto completo de herramientas para evaluar e interpretar datos:
  - Organización y gestión de datos
  - Estadística descriptiva
  - Elaboración de gráficos estadísticos y visualización de datos
  - Distribuciones de probabilidades
  - Modelo lineal y no lineal
  - Contrastes de hipótesis
  - Estadística multivariante
  - Etc.

# Tipos de elementos

## ➤ Los tipos de elementos en Matlab para generar expresiones:

### ➤ Números

#### ➤ Decimales

#### ➤ Notación científica

#### ➤ Números imaginarios

`% Números decimales`

`d1=10, d2=+2, d3=-15, d4=2.1, d5=-2.5`

`% Notación científica`

`c1=0.416e+1, c2=0.0000416e5, c3=4.16e-2`

`% Números imaginarios`

`i1=3i, i2=3j, i3=-1.6*i, i4=2.1e4*i`

`d1 =`

`10`

`d2 =`

`2`

`d3 =`

`-15`

`d4 =`

`2.1000`

`d5 =`

`-2.5000`

`c1 =`

`4.1600`

`c2 =`

`4.1600`

`c3 =`

`0.3416`

`i1 =`

`0.0000 + 3.0000i`

`i2 =`

`0.0000 + 3.0000i`

`i3 =`

`0.0000 - 1.6000i`

`i4 =`

`2.0000e+00 + 2.1000e+04`

# Tipos de elementos

- Los tipos de elementos en Matlab para generar expresiones:
  - Variables: no requiere declaración del tipo de variable, ya que son consideradas como matrices de tamaño variable.
    - Por defecto se inicializan a 0.
    - El nombre de la variable siempre empieza por una letra seguida por letras, dígitos y “\_”, máximo 31 (son los que se leen).
    - Se distingue entre mayúsculas y minúsculas.
    - También se pueden usar como variables lógicas.

```
% Variables lógicas
if (0), disp('0 equivale a False'), end
if (1), disp('1 equivale a True'), end
a_1=21; if (a_1), disp('a_1 distinto 0 equivale a True'), end
if (~0), disp('~ 0 equivale a True'), end
```

Salida en la terminal:

```
1 equivale a True
a_1 distinto 0 equivale a True
~ 0 equivale a True
```

# Tipos de elementos

## ➤ Los tipos de elementos en Matlab para generar expresiones:

### ➤ Operadores

#### ➤ Operadores aritméticos: + , - , \* , / , ^

```
% Operadores aritméticos + , - , * , / , ^  
a = 10  
b = 10.1  
s=a + b           % Suma  
r=a - 3           % Resta  
p=b*3             % Multiplicación  
d=3.5/7           % División  
e=a^b             % Exponenciación
```

```
a =  
    10  
b =  
   10.1000  
s =  
   20.1000  
r =  
     7  
p =  
   30.3000  
d =  
    0.5000  
e =  
  1.2589e+10
```



# Tipos de elementos

- Los tipos de elementos en Matlab para generar expresiones:
  - Funciones: realizan un conjunto de operaciones sobre los parámetros de las funciones. Las funciones están vectorizadas:

```
% Ayuda sobre funciones  
elementales  
help elfun
```

```
>> help elfun  
Elementary math functions.
```

```
Trigonometric.
```

```
sin      - Sine.  
sind     - Sine of argument in degrees.  
sinh     - Hyperbolic sine.  
asin     - Inverse sine.  
asind    - Inverse sine, result in degrees.  
asinh    - Inverse hyperbolic sine.  
cos      - Cosine.  
cosd     - Cosine of argument in degrees.
```

```
.....
```

# Tipos de estructuras de datos

- Las estructuras fundamentales son las matrices:
  - Matrices de dimensión  $n \times 1$  (vectores), por ejemplo para organizar  $n$  observaciones de variables unidimensionales

```
% Vectores
```

```
vf1=[2 3 4 5 6] % vector fila (1,5)
```

```
vf2=[2, 3, 4, 5, 6] % vector fila (1,5)
```

```
vc=[2; 3; 4; 5; 6] % vector columna (5,1)
```

```
vc =
```

```
vf1 =
```

```
2 3 4 5 6
```

```
vf2 =
```

```
2 3 4 5 6
```

```
2
```

```
3
```

```
4
```

```
5
```

```
6
```

# Tipos de estructuras de datos

- Las estructuras fundamentales son las matrices:
  - Matrices de dimensión  $n \times m$  (matrices), por ejemplo para organizar  $n$  observaciones multivariantes en filas y columnas

```
% Matrices
```

```
a=[2 3 4 5 6;2 3 4 5 6] % Matriz (2,5)
```

```
b=[1 2 3 4 5;1 2 3 4 5] % Matriz (2,5)
```

```
c1=[a;b] % Matriz (4,5)
```

```
c2=[a b] % Matriz (2,10)
```

```
a =  
 2 3 4 5 6  
 2 3 4 5 6  
b =  
 1 2 3 4 5  
 1 2 3 4 5
```

```
c1 =  
 2 3 4 5 6  
 2 3 4 5 6  
 1 2 3 4 5  
 1 2 3 4 5  
c2 =  
 2 3 4 5 6 1 2 3 4 5  
 2 3 4 5 6 1 2 3 4 5
```

- Cells para organizar datos heterogéneos con diferentes tamaños, unidades, etc,...

# Tipos de estructuras de datos

- La generación y manipulación de matrices se hace con el operador colon ":"

```
%Generación y manipulación de matrices  
":"
```

```
a=1:10 %vector (1,10)  
b=1:0.5:10 %vector (1,19)  
c=[1:10; 11:20; 21:30] %matriz (3,10)  
ct=c' %matriz (10,3)  
a=[1 2 3; 4 5 6; 7 8 9] %matriz (3,3)  
f1a=a(1,:) % fila 1 de la matriz a  
f3a=a(3,:) % fila 3 de la matriz a  
c1a=a(:,1) % columna 1 de la matriz a  
c2a=a(:,2) % columna 2 de la matriz a  
v=[] % matriz vacía  
a=[1 2 3; 4 5 6; 7 8 9] %matriz (3,3)  
r1=a(1:3,2:3) %matriz (3,2)  
r2=a(1:3,[1 3]) %matriz (3,2)
```

a =

1 2 3 4 5 6 7 8 9 10

b =

Columns 1 through 12

1.0000 1.5000 2.0000 2.5000 3.0000  
3.5000 4.0000 4.5000 5.0000 5.5000  
6.0000 6.5000

Columns 13 through 19

7.0000 7.5000 8.0000 8.5000 9.0000  
9.5000 10.0000

# Tipos de estructuras de datos

- La generación y manipulación de matrices se hace con el operador colon ":"

```
%Generación y manipulación de matrices  
":"
```

```
a=1:10 %vector (1,10)  
b=1:0.5:10 %vector (1,19)  
c=[1:10; 11:20; 21:30] %matriz (3,10)  
ct=c' %matriz (10,3)  
a=[1 2 3; 4 5 6; 7 8 9] %matriz (3,3)  
f1a=a(1,:) % fila 1 de la matriz a  
f3a=a(3,:) % fila 3 de la matriz a  
c1a=a(:,1) % columna 1 de la matriz a  
c2a=a(:,2) % columna 2 de la matriz a  
v=[] % matriz vacía  
a=[1 2 3; 4 5 6; 7 8 9] %matriz (3,3)  
r1=a(1:3,2:3) %matriz (3,2)  
r2=a(1:3,[1 3]) %matriz (3,2)
```

```
ct =  
 1 11 21  
 2 12 22  
 3 13 23  
 4 14 24  
 5 15 25  
 6 16 26  
 7 17 27  
 8 18 28  
 9 19 29  
10 20 30  
a =  
 1 2 3  
 4 5 6  
 7 8 9  
f1a =  
 1 2 3  
f3a =  
 7 8 9  
c1a =  
 1  
 4  
 7  
c2a =  
 2  
 5  
 8  
v =  
 []  
a =  
 1 2 3  
 4 5 6  
 7 8 9  
r1 =  
 2 3  
 5 6  
 8 9  
r2 =  
 1 3  
 4 6  
 7 9
```

# Importar ficheros cvs

- Para leer un fichero cvs, se hace mediante la función `csvread()`

```
%Importando datos en cvs  
data = csvread('spamND.csv', 1);
```

- El parámetro 1 sirve para eliminar la fila de etiquetas.
- Podemos eliminar columnas también, por ejemplo las 4 primera columnas:
  - `data = csvread('spamND.csv', 1, 4);`

# Histogramas

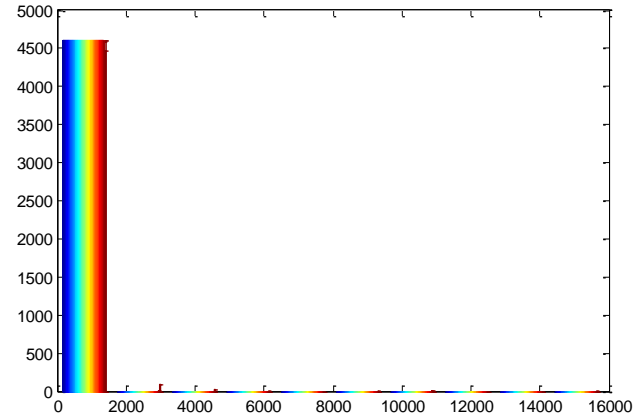
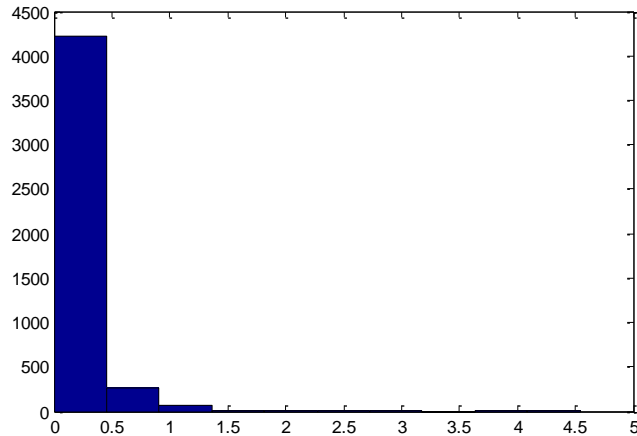
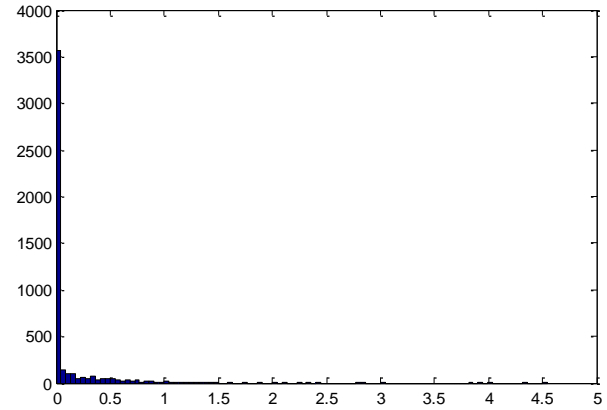
## ➤ Haciendo histogramas por defecto

```
%Calculando histogramas
size(data) %filas columnas
hist(data(:,1)) % histograma de la columna
hist(data(:,1),10) % igual que el anterior
hist(data(:,1),100) % con 100 bins
[N,X] = hist(data(:,1),10) % no pinta nada
hist(data) % pinta todos a la vez
```

```
ans =
    100    16
N =
     4     9    10    15    22    17    13     7     2     1
X =
-1.9690 -1.4932 -1.0174 -0.5416 -0.0658  0.4100  0.8857  1.3615  1.8373  2.3131
```

# Histogramas

## ➤ Haciendo histogramas por defecto



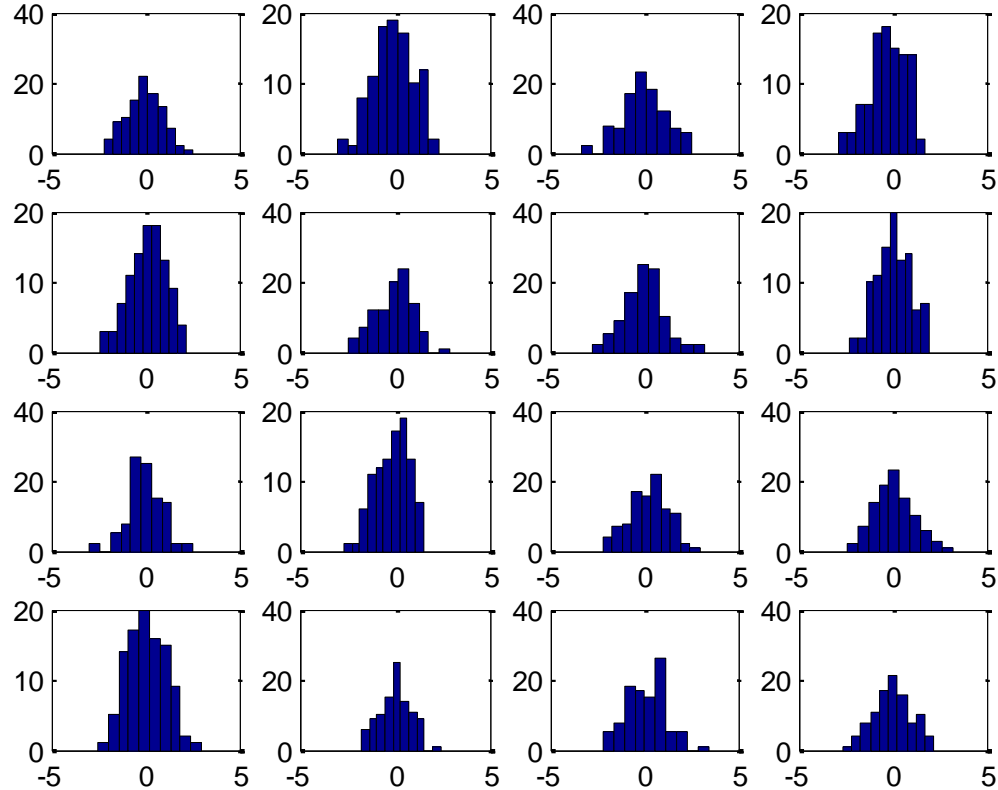


# Histogramas

## ➤ Pintando varios histogramas en una matriz:

```
%Pintando varios histogramas en una matriz
data=randn(100,16); % matriz de datos
aleatorios gaussianos (100,16)
for(i=1:16)
    subplot(4,4,i)% matriz de 4x4 subfiguras
    hist(data(:,i))
end
```

# Matriz de Histogramas



# Medidas de centralidad y posición

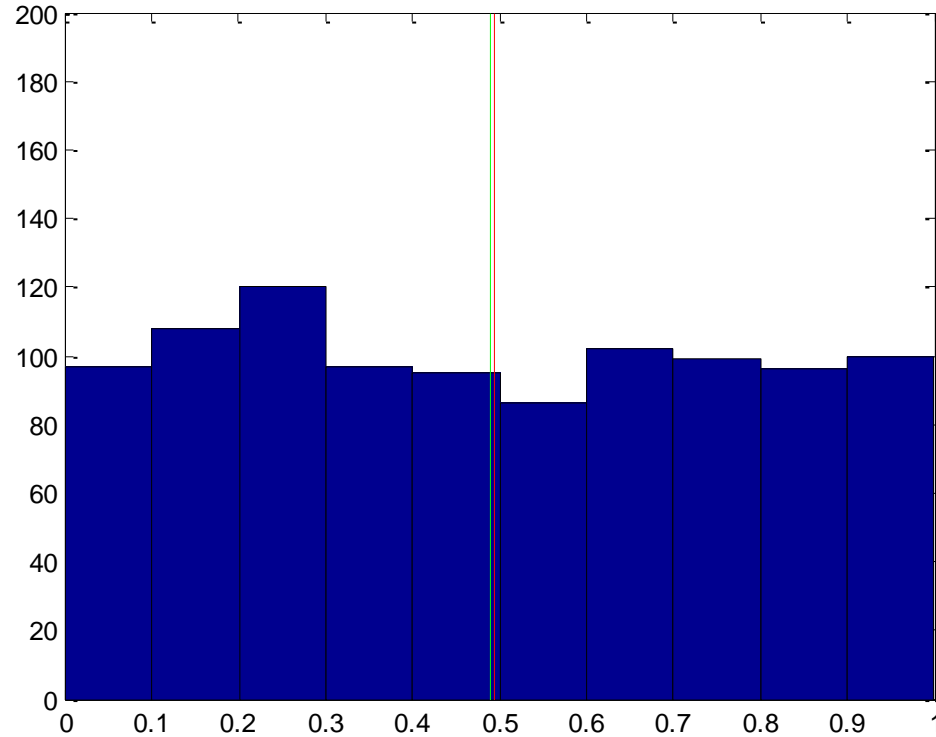
- Para el cálculo de la media aritmética podemos utilizar `mean()`
- Para el cálculo de la mediana podemos utilizar `median()`
- Para el cálculo de percentiles podemos utilizar `prctile()`
  - `prctile(x,p)` devuelve un valor que sería mayor que el  $p\%$  de los valores del vector  $x$ .
  - Así la mediana es el percentil 50%  $\text{median}(x)=\text{prctile}(x,50)$
  - Los cuartiles se calculan como `prctile(x,[25,50,75])`

# Medidas de centralidad y posición

## ➤ Distribución uniforme de datos

```
%Medidas de centralidad y posición de datos uniformes
x_uniforme=rand(1000,1); % 1000 datos distribuidos
uniformemente entre 0 y 1
m=mean(x_uniforme)
med=median(x_uniforme)
hist(x_uniforme)
hold on
plot([m m],[0 200],'r')
plot([med med],[0 200],'g')
hold off
pqr=prctile(x_uniforme,50)
q=prctile(x_uniforme,[25,50,75])
```

# Medidas de centralidad y posición



m =

0.4932

med =

0.4892

pcr =

0.4892

q =

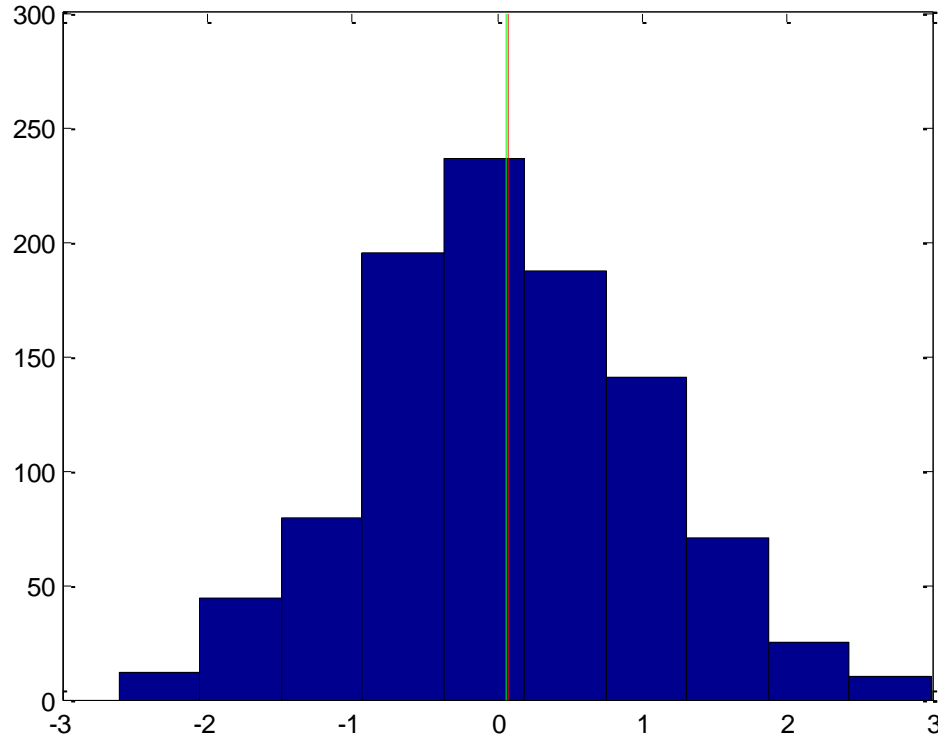
0.2384 0.4892 0.7532

# Medidas de centralidad y posición

## ➤ Distribución gaussiana de datos

```
%Medidas de centralidad y posición de datos distribuidos
según una gaussiana
x_g=randn(1000,1); % 1000 datos distribuidos
uniformemente entre 0 y 1
m=mean(x_g)
med=median(x_g)
hist(x_g)
hold on
plot([m m],[0 300],'r')
plot([med med],[0 300],'g')
hold off
pqr=prctile(x_g,50)
q=prctile(x_g,[25,50,75])
```

# Medidas de centralidad y posición



m =

0.0745

med =

0.0621

pcr =

0.0621

q =

-0.5949    0.0621    0.7452

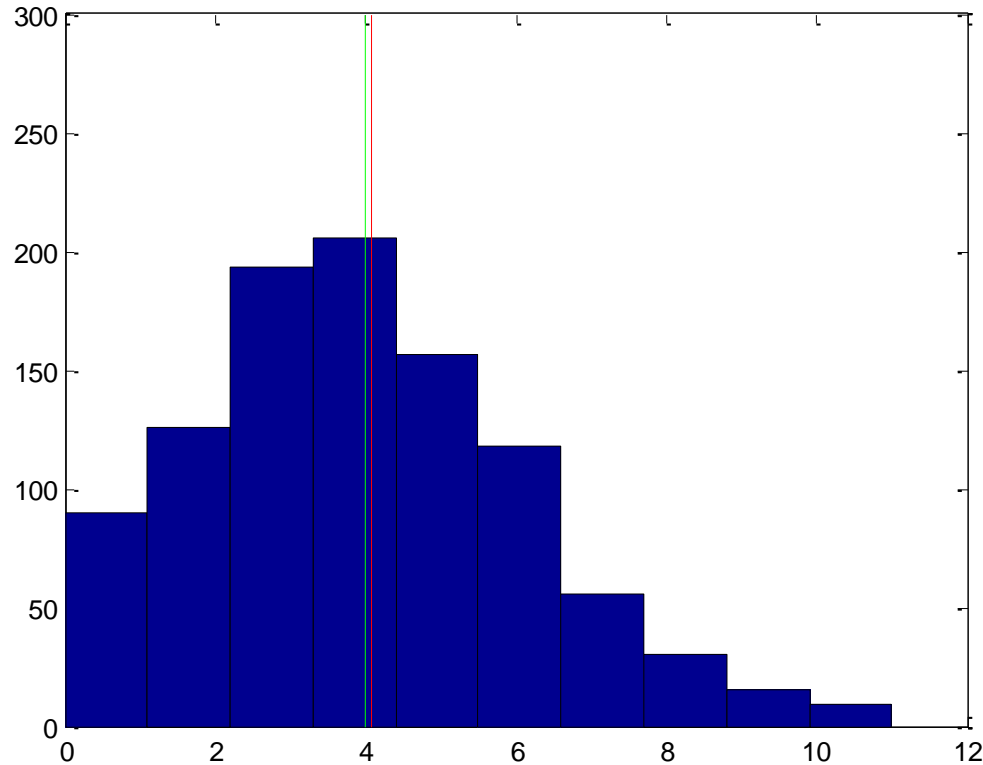
# Medidas de centralidad y posición

## ➤ Distribución de Poisson para los datos

```
%Medidas de centralidad y posición de datos
distribuidos según una poisson
x_p=poissrnd(4,1000,1); % 1000 datos distribuidos
seún una poisson de lambda 4
m=mean(x_p)
med=median(x_p)
hist(x_p)
hold on
plot([m m],[0 300],'r')
plot([med med],[0 300],'g')
hold off
pqr=prctile(x_p,50)
q=prctile(x_p,[25,50,75])
```



# Medidas de centralidad y posición



m =  
4.0800

med =  
4

pcr =  
4

q =  
3 4 5

# Medidas de dispersión

## ➤ Rango

- `range(x)`
- `max(x)-min(x)`

## ➤ Rango intercuartílico

- `iqr(x)`
- `prctile(x,75)-prctile(x,25)`

## ➤ Varianza

- `sum((x-mean(x)).^2)/length(x)`
- `var(x)` %normaliza por n - 1
- `var(x,1)` %normaliza por n

## ➤ Desviación típica

- `sqrt(sum((x-mean(x)).^2)/ length(x))`
- `sqrt(var(x))` %normaliza por n - 1
- `std(x)` %normaliza por n - 1
- `sqrt(var(x,1))` %normaliza por n
- `std(x,1)` %normaliza por n

# Medidas de dispersión

```
%Medidas de dispersión de datos distribuidos según
una gaussiana
x_g=randn(1000,1); % 1000 datos distribuidos según
una gaussiana
r1=range(x_g) %Rango
r2=max(x_g)-min(x_g)
i1=iqr(x_g) %Rango intercuartílico
i2=prctile(x_g,75)-prctile(x_g,25)
v1=sum((x_g-mean(x_g)).^2)/length(x_g) %Varianza
v2=var(x_g) % normaliza por n-1
v3=var(x_g,1) % normaliza por n
s1=sqrt(sum((x_g-mean(x_g)).^2)/length(x_g))
%Desviación típica
s2=sqrt(var(x_g,1)) % normaliza por n
s3=std(x_g,1) % normaliza por n
s4=sqrt(var(x_g,1))% normaliza por n
s5=std(x_g,1) % normaliza por n
```

```
r1 =
    6.8901

r2 =
    6.8901

i1 =
    1.4083

i2 =
    1.4083

v1 =
    1.0217

v2 =
    1.0227

v3 =
    1.0217

s1 =
    1.0108

s2 =
    1.0108

s3 =
    1.0108

s4 =
    1.0108

s5 =
    1.0108
```

# Medidas de forma

- **Coefficiente de asimetría:**

- $CA = \sum (x_i - \langle x \rangle)^3 / ns^3$

- $skewness(x)$

- **Coefficientes de apuntamiento o curtosis:**

- $CA_p = \sum (x_i - \langle x \rangle)^4 / ns^4$

- $kurtosis(x)$

# Medidas de forma

```
%Medidas de forma
x_u=rand(1000,1); % 1000 datos
distribuidos uniformemente entre 0 y 1
figure(1); hist(x_u)
x_g=randn(1000,1); % 1000 datos
distribuidos según una gaussiana
figure(2);hist(x_g)
x_p=poissrnd(4,1000,1); % 1000 datos
distribuidos según una Poisson de lambda
4
figure(3);hist(x_p)
CA_u=skewness(x_u), CA_g=skewness(x_g),
CA_p=skewness(x_p) % Coeficientes de
asimetría
CAp_u=kurtosis(x_u), CAp_g=kurtosis(x_g),
CAp_p=kurtosis(x_p) % Coeficientes de
curtosis
```

CA\_u =  
0.0281

CA\_g =  
-0.0526

CA\_p =  
0.6886

CAp\_u =  
1.8104

CAp\_g =  
2.9335

CAp\_p =  
3.8057

# Medidas de forma

$$CA_u = 0.0281$$

$$CAp_u = 1.8104$$

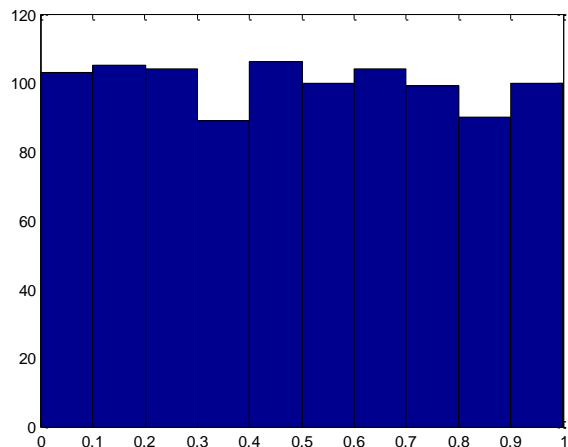
$$CA_g = -0.0526$$

$$CAp_g = 2.9335$$

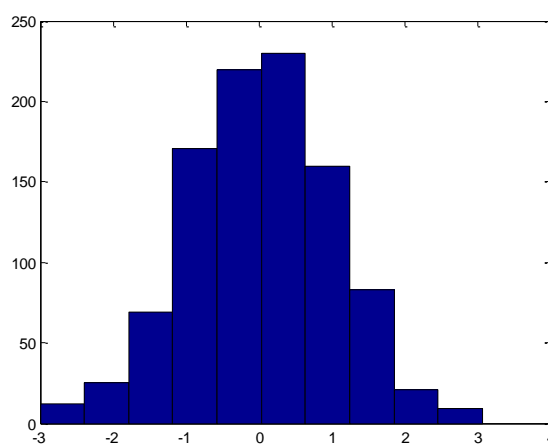
$$CA_p = 0.6886$$

$$CAp_p = 3.8057$$

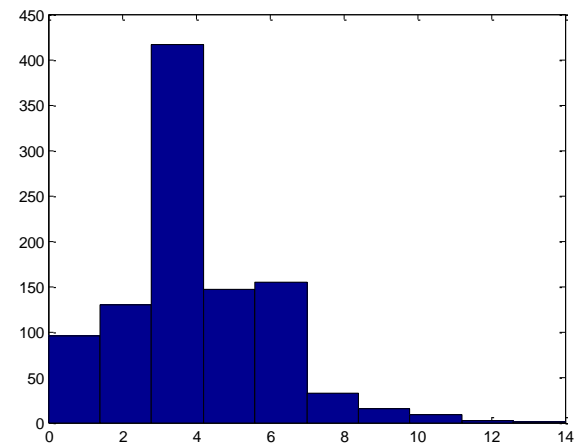
figure(1); hist(x\_u)



figure(2); hist(x\_g)



figure(3); hist(x\_p)



# Diagramas de caja (boxplot)

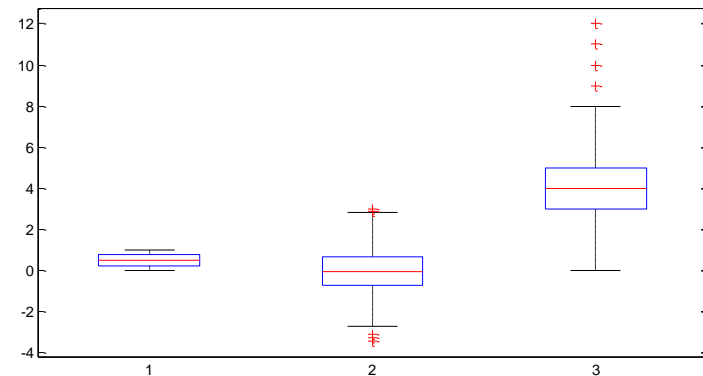
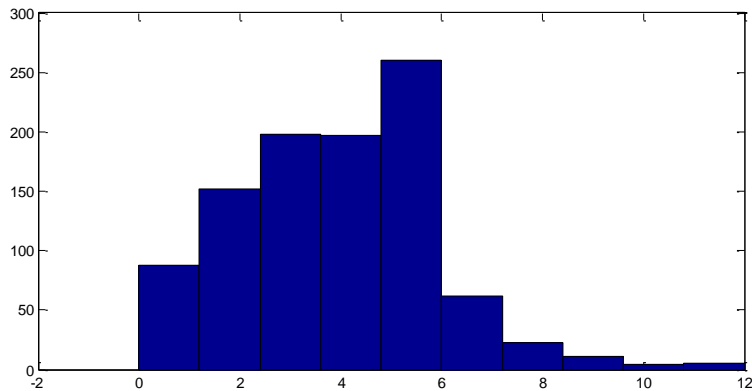
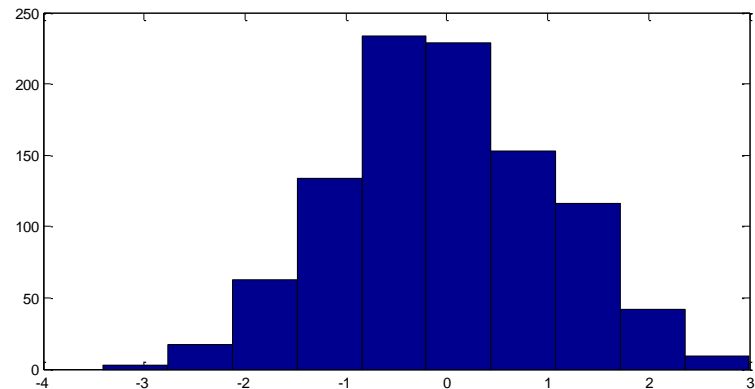
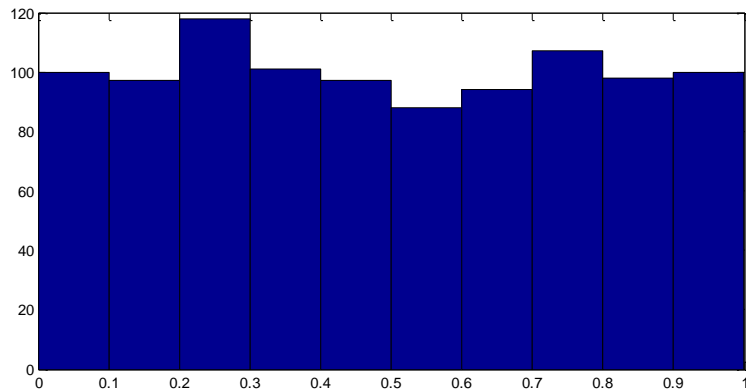
- La información obtenida a partir de las medidas de centralización, dispersión y forma se pueden resumir en un único diagrama: boxplots.
- El diagrama de caja consta de una caja central que está delimitada por la posición de los cuartiles Q3 y Q1.
- La caja está dividida por la mediana.
- De los extremos de la caja salen unas líneas que se extienden hasta los punto inferior y superior:
  - $LI = \max\{ \min(x_i) , Q1 - 1.5 RIC \}$
  - $LS = \min\{ \max(x_i) , Q3 + 1.5 RIC \}$
- Los datos fuera del intervalo (LI,LS) se consideran atípicos y se pintan en el gráfico.
- La función de Matlab es `boxplot(x)`.

# Diagramas de caja (boxplot)

```
%Boxplots
x_u=rand(1000,1); % 1000 datos distribuidos
uniformemente entre 0 y 1
subplot(2,2,1); hist(x_u)
x_g=randn(1000,1); % 1000 datos distribuidos según
una gaussiana
subplot(2,2,2);hist(x_g)
x_p=poissrnd(4,1000,1); % 1000 datos distribuidos
según una Poisson de lambda 4
subplot(2,2,3);hist(x_p)
subplot(2,2,4);boxplot([x_u x_g x_p])
```



# Diagramas de caja (boxplot)



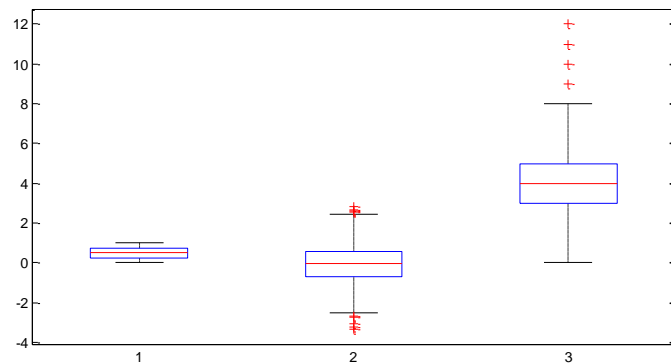
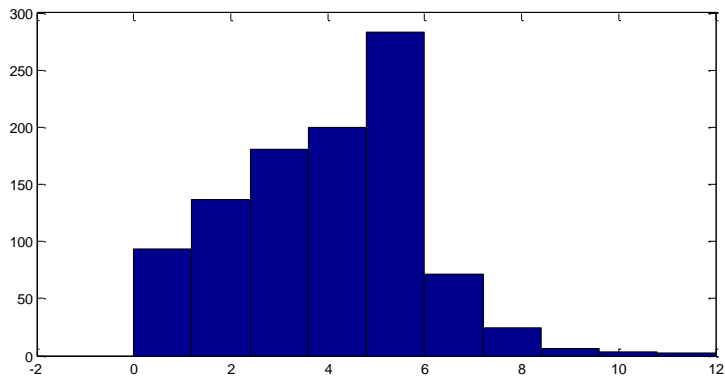
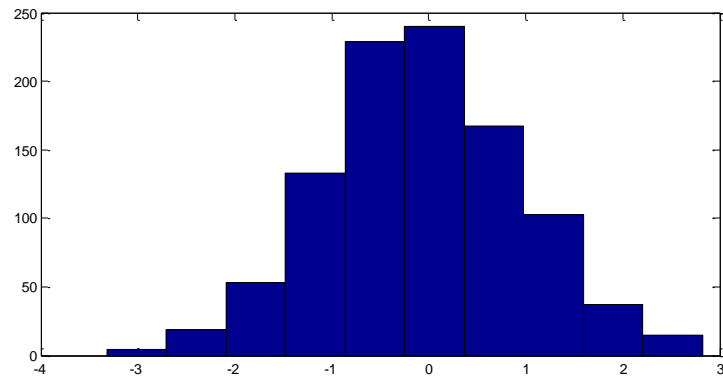
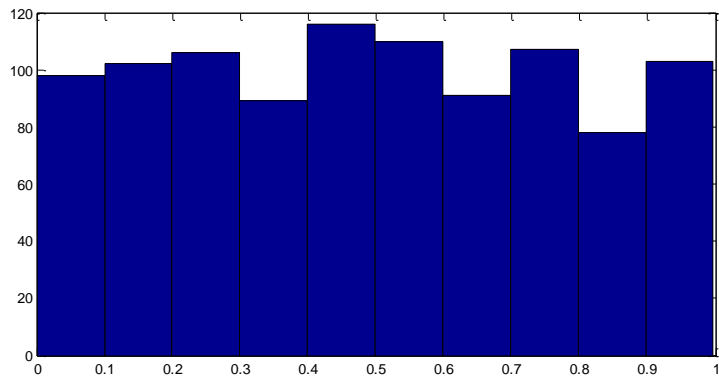
# Visualización datos con medidas de posicionamiento: Gráficos Q-Q

- Los gráficos Q-Q (Q-Q plots) se caracterizan por visualizar de una manera muy rápida y sencilla como se diferencian los datos de dos distribuciones de observaciones.
- Se basan en representar enfrentados en un gráfico x-y los cuantiles de ambas distribuciones. El “Q” viene de cuantil en inglés.
- Si todos los cuantiles son iguales aparecerá la recta  $x=y$  en el gráfico, y significará los dos conjuntos de datos se distribuyen de manera idéntica.
- Generalmente una de la distribuciones es conocida (por ejemplo una normal en Matlab por defecto), para contrastar si los datos observados se ajustan a la distribución conocida.
- La función en Matlab es `qqplot()`.

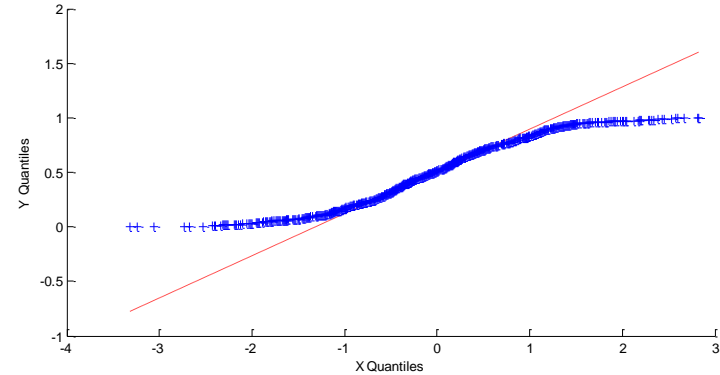
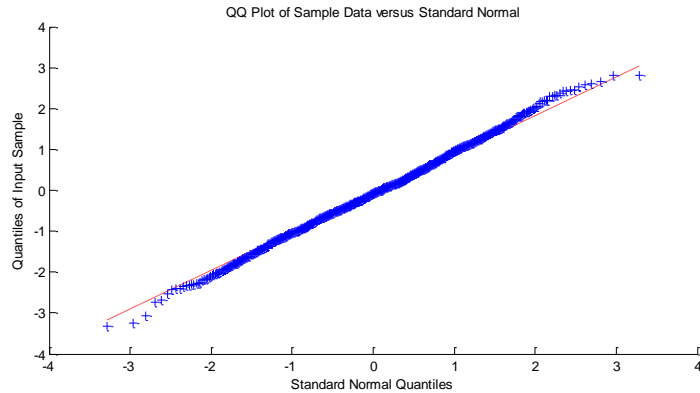
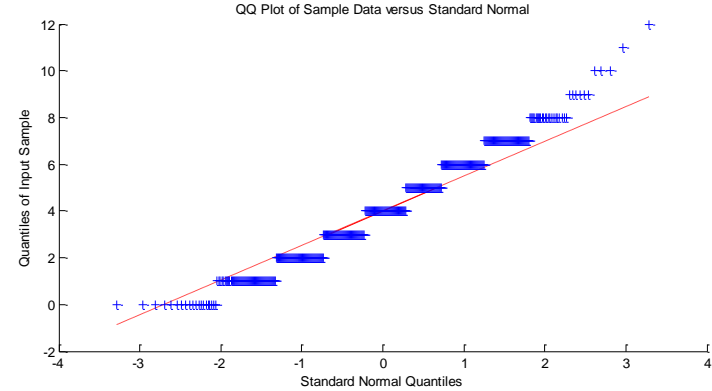
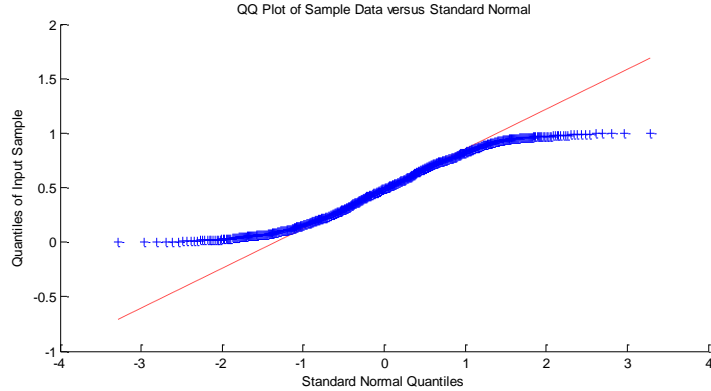
# Visualización datos con medidas de posicionamiento: Gráficos Q-Q

```
%QQplots
x_u=rand(1000,1); % 1000 datos distribuidos uniformemente entre
0 y 1
subplot(2,2,1); hist(x_u)
x_g=randn(1000,1); % 1000 datos distribuidos según una
gaussiana
subplot(2,2,2);hist(x_g)
x_p=poissrnd(4,1000,1); % 1000 datos distribuidos según una
Poisson de lambda 4
subplot(2,2,3);hist(x_p)
subplot(2,2,4);boxplot([x_u x_g x_p])
figure(2)
subplot(2,2,1); qqplot(x_u)
subplot(2,2,2);qqplot(x_p)
subplot(2,2,3);qqplot(x_g)
subplot(2,2,4);qqplot(x_g,x_u)
```

# Visualización datos con medidas de posicionamiento: Gráficos Q-Q



# Visualización datos con medidas de posicionamiento: Gráficos Q-Q



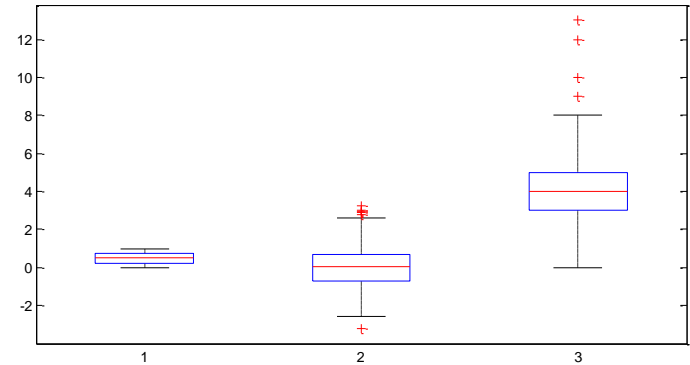
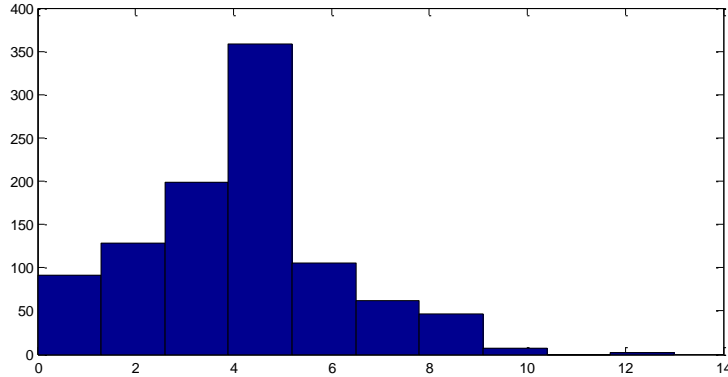
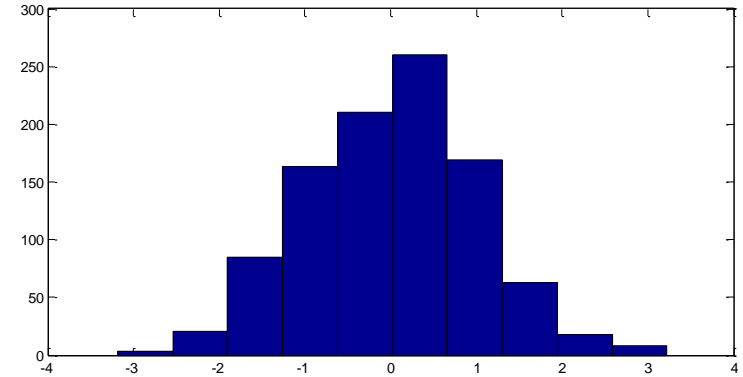
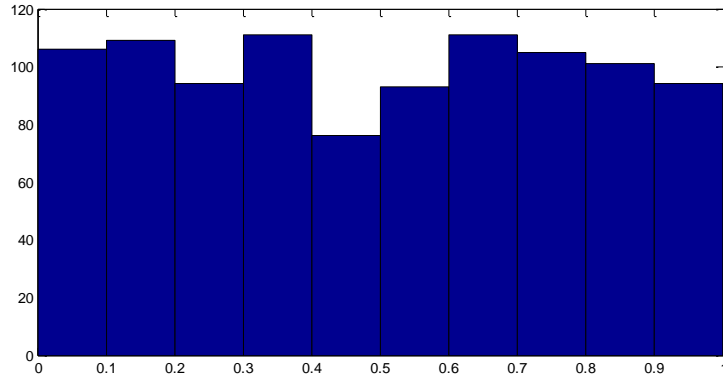
# Diagramas de dispersión e histogramas

- Se suelen combinar en estadística multivariante los diagramas de dispersión de las diferentes variables junto con los histogramas.
- Un diagrama de dispersión es la representación de las observaciones de dos o tres variables enfrentadas.
- Existen varios tipos en matlab:
  - `scatterhist()`
  - `gplotmatrix()`

# Diagramas de dispersión e histogramas

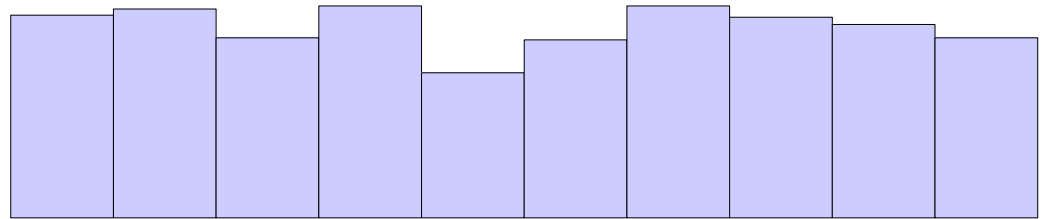
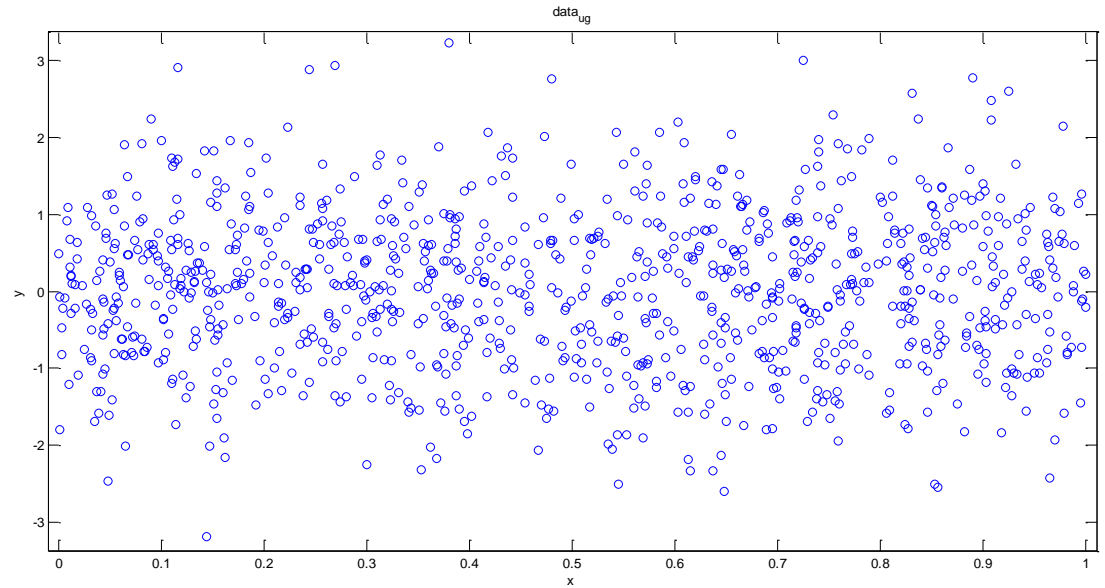
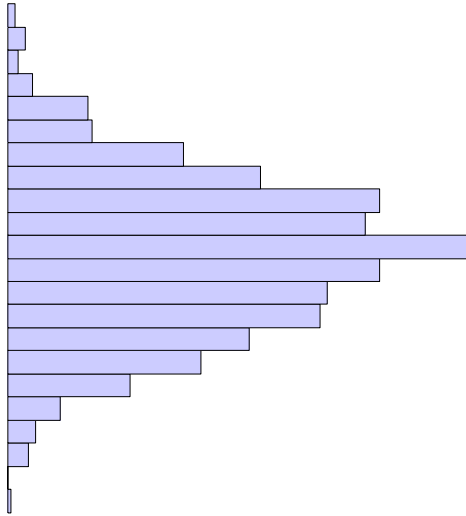
```
%Diagramas de dispersión e histogramas
figure(1);
x_u=rand(1000,1); % 1000 datos distribuidos uniformemente entre 0 y 1
subplot(2,2,1); hist(x_u)
x_g=randn(1000,1); % 1000 datos distribuidos según una gaussiana
subplot(2,2,2);hist(x_g)
x_p=poissrnd(4,1000,1); % 1000 datos distribuidos según una Poisson de
lambda 4
subplot(2,2,3);hist(x_p)
subplot(2,2,4);boxplot([x_u x_g x_p])
data_ug=[x_u x_g];
data_gp=[x_g x_p];
data_pu=[x_p x_u];
figure(2); scatterhist(data_ug(:,1),data_ug(:,2)); title('data_{ug}')
figure(3); scatterhist(data_gp(:,1),data_gp(:,2)); title('data_{gp}')
figure(4); scatterhist(data_pu(:,1),data_pu(:,2)); title('data_{pu}')
data=[x_u x_g x_p];
figure(5); gplotmatrix(data);
```

# Diagramas de dispersión e histogramas

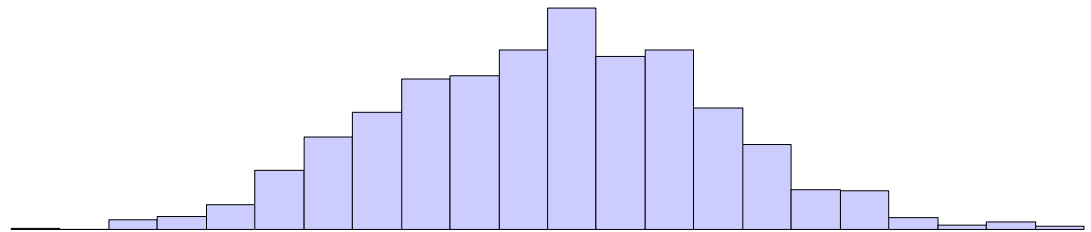
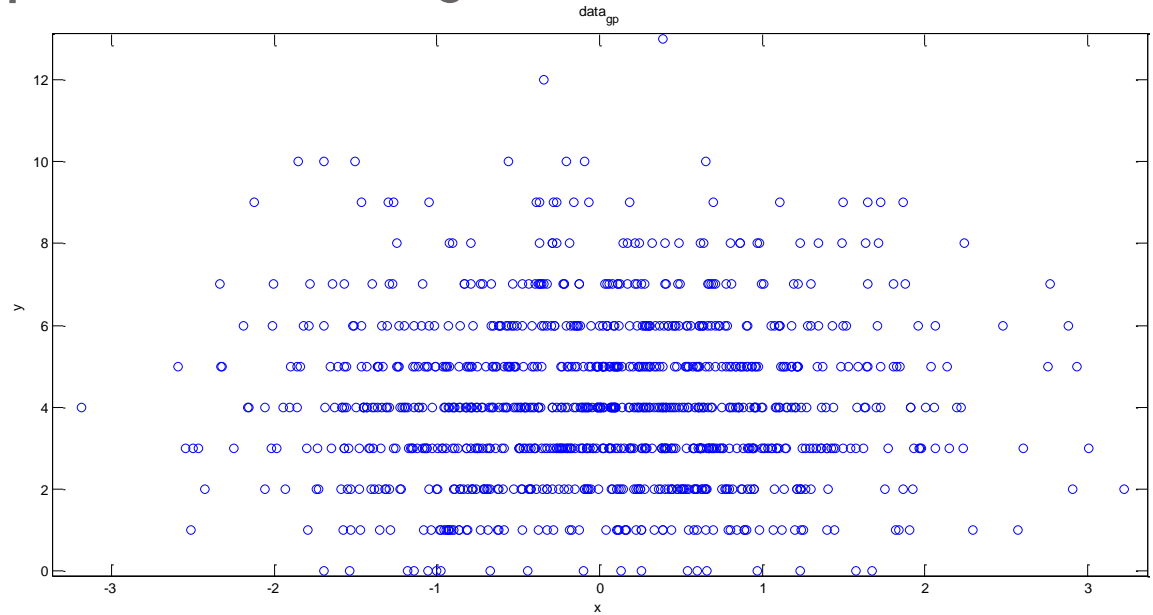
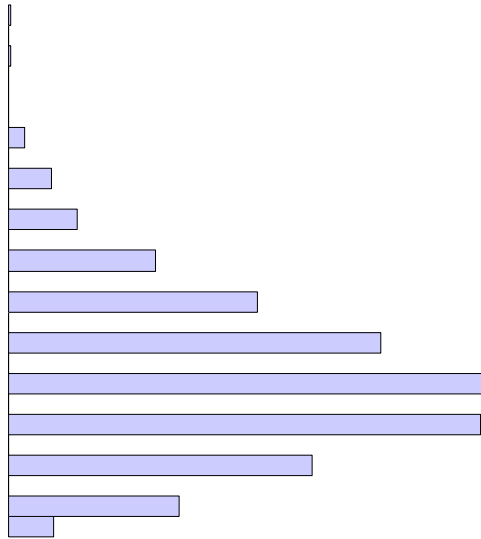




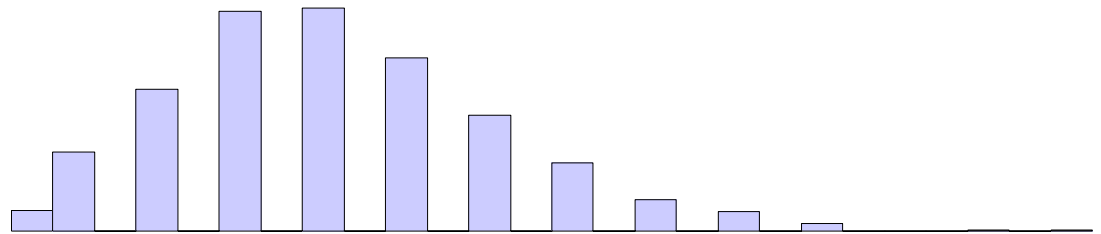
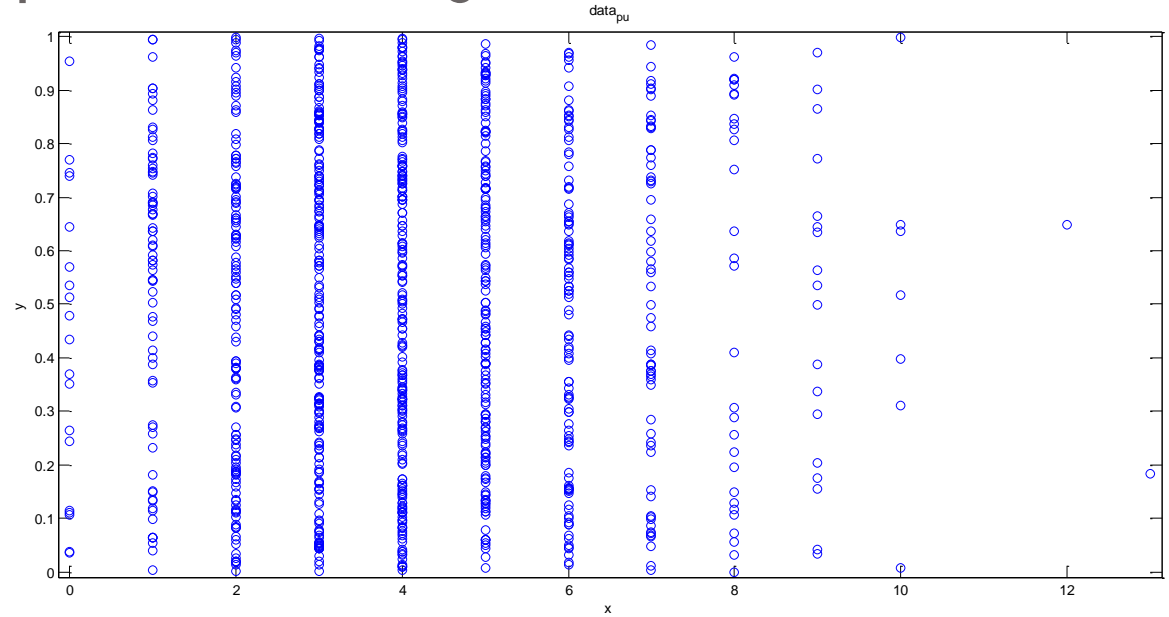
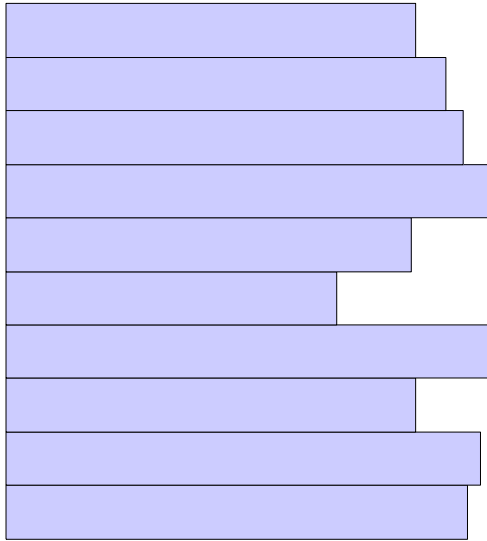
# Diagramas de dispersión e histogramas



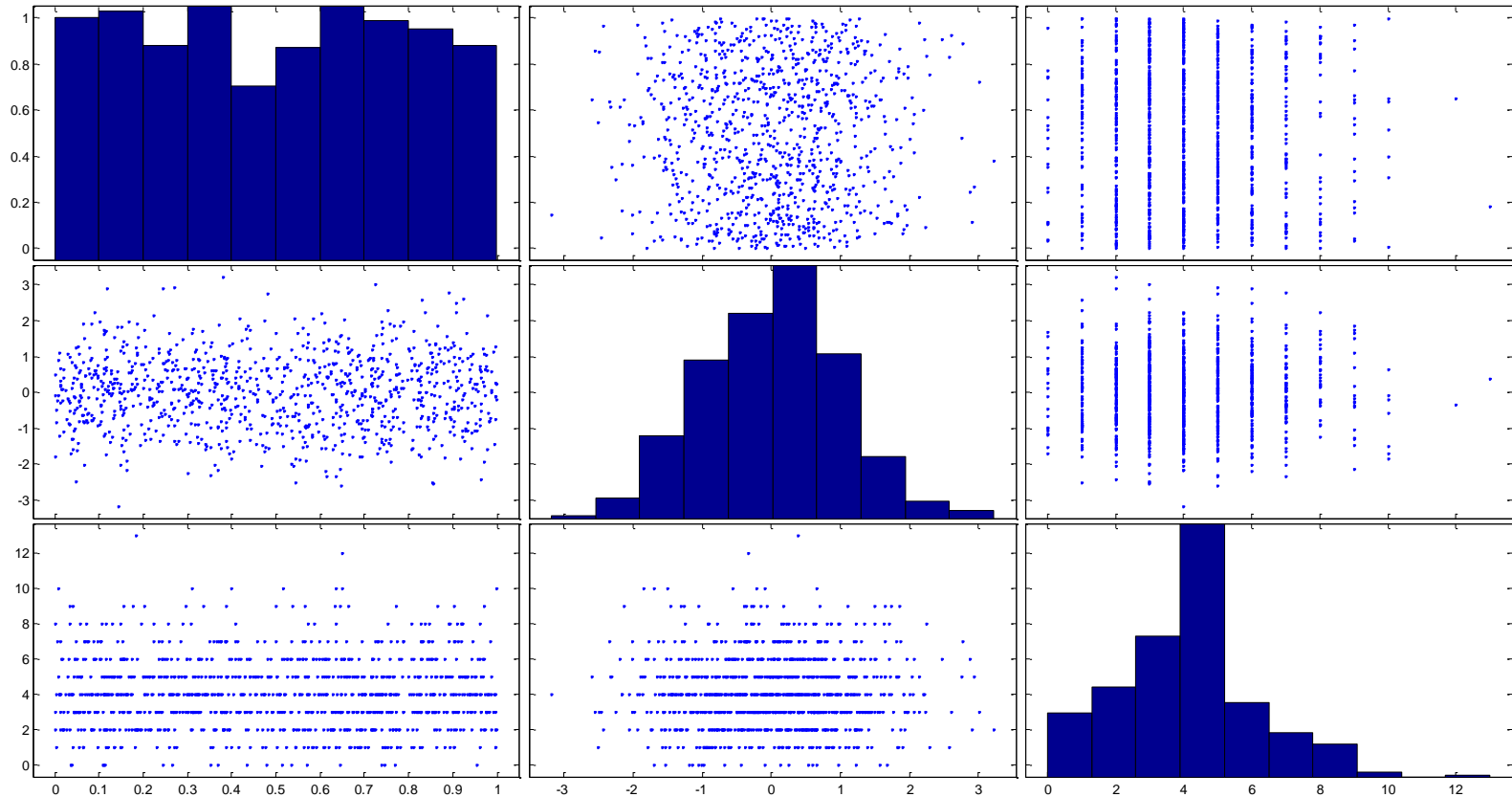
# Diagramas de dispersión e histogramas



# Diagramas de dispersión e histogramas



# Diagramas de dispersión e histogramas



## Bibliografía y lecturas relacionadas:

- <http://es.mathworks.com/help/>
- <http://es.mathworks.com/products/statistics/>
- <http://www.pi.ingv.it/~longo/CorsoMatlab/OriginalManuals/stats.pdf>
- [Fundamentos de estadística. Daniel Peña Sánchez Ribera. Alianza Editorial, 2001 o 2008.](#)
- [Computational statistics handbook with MATLAB. Martinez, Wendy L. Chapman & Hall/CRC, 2008.](#)
- [Statistics in MATLAB A Primer. Cho, MoonJung. Chapman and Hall/CRC, 2014.](#)

# Introducción: Algunas Herramientas - SPSS

- **Introducción**
- **Las cuatro ventanas**
- **Análisis Básico**
- **Frecuencias**
- **Descriptivos**
- **Análisis de regresión lineal**
- **Varias representaciones Visuales**

# Introducción

- Originalmente es un acrónimo del Paquete Estadístico para las Ciencias Sociales (Statistical Package for the Social Science), Sin embargo, en la actualidad la parte SPSS del nombre completo del software (IBM SPSS) no es acrónimo de nada.
- Es uno de los paquetes estadísticos hoy en día más populares que puede realizar la manipulación de datos de gran complejidad y un análisis de los mismos con instrucciones muy simples.



Imagen extraída de <https://es.wikipedia.org/wiki/SPSS>

# Las cuatro ventanas

- En SPSS puedes manejar lo básico con cuatro ventanas.
- Las cuatro ventanas:
  - Editor de datos
  - Una ventana de salida de resultados
  - Un editor de sintaxis
  - Y una ventana para ejecutar secuencias de comandos
- Estas cuatro ventanas pueden ser invocadas desde
  - Si son nuevas: `archivo>nuevo` y aparecen las cuatro opciones.
  - Si ya existen como fichero: `archivo>abrir` y aparecen las cuatro opciones



# Las cuatro ventanas: Editor de datos

- Editor de datos: Sistema de hoja de cálculo para definir, introducir, editar y visualizar datos. La extensión del archivo guardado será "SAV", esto es lo que se llama archivo muestral.
- Vamos a utilizar demo.sav: es un estudio ficticio de varios miles de personas que contiene información básica demográfica y de consumo (viene en el directorio de "samples" de SPSS con otros muchos ejemplos).
- demo.sav es un fichero binario, y para abrirlo `archivo>abrir>datos....`
- [Descripción de todos los archivos muestrales de SPSS](#)

# Las cuatro ventanas: Editor de datos

- La ventana por defecto tendrá el editor de datos
- Hay dos hojas en la ventana:
  - Vista de datos
  - Vista de variables
- La ventana Vista de datos es visible cuando se abre por primera vez el Editor de datos
  - Esta hoja contiene los datos que se han cargado
- Se pueden marcar datos con el ratón y hacer análisis estadísticos.

# Las cuatro ventanas: Editor de datos

demo.sav [Conjunto\_de\_datos2] - IBM SPSS Statistics Editor de datos

Archivo Editar Ver Datos Transformar Analizar Gráficos Utilidades Ventana Ayuda

15 - cochecat 3.00 Visible: 29 de 29 variables

	edad	marital	direcc	ingres	ingcat	coche	cochecat	educ	empleo	retirado	empcat	satlab	genero	residen	inalam	multine	voz	busca	internet	idilam	espera	tv
1	55	1	12	72.00	3.00	36.20	3.00	1	23	0	3	5 m	4	0	0	1	0	0	0	0	0	1
2	56	0	29	153.00	4.00	76.90	3.00	1	35	0	3	4 h	1	1	0	0	1	0	0	1	1	1
3	28	1	9	28.00	2.00	13.70	1.00	3	4	0	1	3 m	3	1	0	1	0	0	0	1	1	1
4	24	1	4	26.00	2.00	12.50	1.00	4	0	0	1	1 h	3	1	1	1	0	0	0	0	1	1
5	25	0	2	23.00	1.00	11.30	1.00	2	5	0	2	2 h	2	0	0	0	0	0	0	1	0	1
6	45	1	9	76.00	4.00	37.20	3.00	3	13	0	2	2 h	2	0	1	1	1	0	0	0	0	1
7	42	0	19	40.00	2.00	19.80	2.00	3	10	0	2	2 h	1	1	1	1	1	0	1	1	0	1
8	35	0	15	57.00	3.00	28.20	2.00	2	1	0	1	1 m	1	0	0	0	0	0	0	0	0	1
9	46	0	26	24.00	1.00	12.20	1.00	1	11	0	2	5 m	2	0	0	1	0	0	0	1	0	1
10	34	1	0	89.00	4.00	46.10	3.00	3	12	0	2	4 h	6	1	0	0	1	0	0	1	1	1
11	55	1	17	72.00	3.00	35.50	3.00	3	2	0	1	3 m	2	1	0	0	0	0	1	0	0	1
12	28	0	3	24.00	1.00	11.80	1.00	4	4	0	1	5 h	1	1	0	0	1	1	0	0	1	1
13	31	1	9	40.00	2.00	21.30	2.00	4	0	0	1	2 m	4	0	0	0	0	0	0	0	0	1
14	42	0	8	137.00	4.00	68.90	3.00	3	3	0	1	1 m	1	1	0	0	0	0	0	1	0	1
15	35	0	8	70.00	3.00	34.10	3.00	3	9	0	2	4 h	3	0	0	0	0	0	1	0	0	1
16	52	1	24	159.00	4.00	78.90	3.00	4	16	0	3	5 h	2	1	1	1	1	0	0	1	1	1
17	21	1	1	37.00	2.00	18.60	2.00	3	0	0	1	1 h	7	1	1	0	0	0	0	1	1	1
18	32	0	0	28.00	2.00	13.70	1.00	1	2	0	1	4 m	2	0	0	0	0	0	0	1	1	1
19	42	0	9	109.00	4.00	54.70	3.00	3	20	0	3	3 m	1	1	0	0	0	0	1	0	0	1
20	40	1	12	117.00	4.00	58.30	3.00	2	19	0	3	5 m	4	1	1	1	1	0	1	0	0	1
21	30	0	3	23.00	1.00	11.80	1.00	1	3	0	1	3 h	1	0	0	0	0	0	0	0	0	1
22	48	0	14	21.00	1.00	9.50	1.00	3	2	0	1	3 h	1	0	0	1	0	0	0	0	0	1
23	39	1	17	17.00	1.00	8.50	1.00	4	2	0	1	3 h	5	1	1	1	0	0	0	1	0	1
24	42	1	5	34.00	2.00	16.60	2.00	2	13	0	2	3 m	4	0	0	0	0	0	0	1	0	1
25	45	1	12	115.00	4.00	57.40	3.00	1	27	0	3	4 m	5	0	0	1	0	0	0	1	1	1
26	51	1	10	47.00	2.00	23.00	2.00	1	9	0	2	3 h	3	0	0	0	1	0	0	0	0	1
27	39	1	9	33.00	2.00	16.30	2.00	3	1	0	1	1 h	4	1	0	0	0	0	0	0	0	1
28	49	0	29	135.00	4.00	68.40	3.00	2	14	0	2	5 m	1	0	1	0	0	0	0	0	0	1
29	52	0	20	272.00	4.00	74.90	3.00	1	35	0	3	5 h	1	1	1	1	0	0	0	1	0	1
30	53	1	29	41.00	2.00	19.90	2.00	1	9	0	2	4 h	2	1	0	1	0	0	0	1	1	1
31	34	0	10	20.00	1.00	10.00	1.00	3	0	0	1	1 m	1	1	1	1	1	9	1	1	1	1
32	47	1	6	22.00	1.00	11.40	1.00	3	7	0	2	4 h	2	0	0	0	0	0	0	0	0	1
33	58	0	2	60.00	3.00	29.70	2.00	4	1	0	1	1 h	1	1	0	0	0	0	0	1	1	1
34	25	1	0	58.00	3.00	28.40	2.00	3	4	0	1	2 m	5	1	1	1	0	0	0	1	1	1
35	57	1	28	92.00	4.00	45.50	3.00	2	25	0	3	5 m	4	0	1	1	1	0	0	0	1	1
36	30	1	7	21.00	1.00	10.50	1.00	4	4	0	1	1 h	2	0	0	0	0	0	1	0	1	1
37	21	0	0	13.00	1.00	6.30	1.00	3	0	0	1	1 m	2	0	0	0	0	0	1	0	0	1

Vista de datos Vista de variables

IBM SPSS Statistics Processor está listo Unicode ON

13:50 30/09/2015

# Las cuatro ventanas: Editor de datos

- Si hacemos clic en la pestaña Vista de variables, pasamos a vista de variables.
- Esta hoja contiene información sobre el conjunto de datos que se almacena en el conjunto de datos: nombre, tipo, anchura, ...
- Nombre: el primer carácter del nombre de la variable debe ser alfabético
  - Los nombres de variables deben ser únicos, y tiene que ser inferior a 64 caracteres.
  - No se permiten espacios.
- Tipo: indica el tipo de variable.
  - Haga clic en la casilla de «tipo». Los dos tipos básicos de las variables que se utilizarán son numéricas y de cadena. Esta columna le permite especificar el tipo de variable.

# Las cuatro ventanas: Editor de datos

- **Anchura:** el número de dígitos para valores numéricos o la longitud de una variable de cadena.
- **Decimales:** número de decimales
  - Tiene que ser menor o igual a 16
- **Etiqueta:** puede especificar los detalles de la variable
  - Se puede escribir caracteres hasta 256.
- **Valores:** esto se utiliza y sugerir que los números representan la categoría de la variable, cuando esta representa una categoría.
  - **Definición de las etiquetas de valor:**
    - Haga clic en la celda de la columna los valores.
    - Para el valor, y la etiqueta, puede poner hasta 60 caracteres.
    - Después de definir los valores haga clic en Agregar y, a continuación, haga clic en Aceptar.

# Las cuatro ventanas: Editor de datos

demo.sav [Conjunto\_de\_datos2] - IBM SPSS Statistics Editor de datos

Archivo Editar Ver Datos Transformar Analizar Gráficos Utilidades Ventana Ayuda

	Nombre	Tipo	Anchura	Decimales	Etiqueta	Valores	Perdidos	Columnas	Alineación	Medida	Rol
4	ingres	Númerico	8	2	Ingresos del ho...	Ninguna	Ninguna	8	Derecha	Escala	Entrada
5	ingcat	Númerico	8	2	Categoría de in...	{1,00, Meno...	Ninguna	8	Derecha	Ordinal	Entrada
6	coche	Númerico	8	2	Precio del coch...	Ninguna	Ninguna	8	Derecha	Escala	Entrada
7	cochecat	Númerico	8	2	Categoría del pr...	{1,00, Econ...	Ninguna	8	Derecha	Ordinal	Entrada
8	educ	Númerico	4	0	Nivel educativo	{1, No comp...	Ninguna	8	Derecha	Escala	Entrada
9	empleo	Númerico	4	0	Años con la em...	Ninguna	Ninguna	8	Derecha	Escala	Entrada
10	retrato	Númerico	4	0	Retrato	{0, No}...	Ninguna	8	Derecha	Escala	Entrada
11	empcat	Númerico	4	0	Años con la em...	{1, Menos d...	Ninguna	8	Derecha	Ordinal	Entrada
12	satlab	Númerico	4	0	Satisfacción la...	{1, Muy ins...	Ninguna	8	Derecha	Escala	Entrada
13	genero	Cadena	1	0	Género	{h, Hombre}...	Ninguna	8	Izquierda	Nominal	Entrada
14	residen	Númerico	4	0	Número de per...	Ninguna	Ninguna	8	Derecha	Escala	Entrada
15	inalam	Númerico	4	0	Servicio inalám...	{0, S}...	Ninguna	8	Derecha	Escala	Entrada
16	multiline	Númerico	4	0	Múltiples líneas	{0, S}...	Ninguna	8	Derecha	Escala	Entrada
17	voz	Númerico	4	0	Buzón de voz	{0, S}...	Ninguna	8	Derecha	Escala	Entrada
18	busca	Númerico	4	0	Buscapersonas	{0, S}...	Ninguna	8	Derecha	Escala	Entrada
19	internet	Númerico	4	0	Internet	{0, S}...	8, 9	8	Derecha	Escala	Entrada
20	idllam	Númerico	4	0	ID llamadas	{0, S}...	Ninguna	8	Derecha	Escala	Entrada
21	espera	Númerico	4	0	Llamada en es...	{0, S}...	Ninguna	8	Derecha	Escala	Entrada
22	tv	Númerico	4	0	Tiene TV	{0, S}...	Ninguna	8	Derecha	Escala	Entrada
23	video	Númerico	4	0	Tiene Vídeo	{0, S}...	Ninguna	8	Derecha	Escala	Entrada
24	cd	Númerico	4	0	Tiene Hi-fi/CD	{0, S}...	Ninguna	8	Derecha	Escala	Entrada
25	pda	Númerico	4	0	Tiene PDA	{0, S}...	Ninguna	8	Derecha	Escala	Entrada
26	pc	Númerico	4	0	Tiene ordenador	{0, S}...	Ninguna	8	Derecha	Escala	Entrada
27	fax	Númerico	4	0	Tiene fax	{0, S}...	Ninguna	8	Derecha	Escala	Entrada
28	noticias	Númerico	4	0	Suscrito a un p...	{0, S}...	Ninguna	8	Derecha	Escala	Entrada
29	respuest	Númerico	4	0	Respuesta	{0, S}...	Ninguna	8	Derecha	Escala	Entrada
30											
31											
32											
33											
34											
35											
36											
37											
38											
39											
40											
41											
42											
43											
44											
45											
46											
47											
48											
49											
50											
51											
52											
53											
54											
55											
56											
57											
58											
59											
60											
61											
62											
63											
64											
65											
66											
67											
68											
69											
70											
71											
72											
73											
74											
75											
76											
77											
78											
79											
80											
81											
82											
83											
84											
85											
86											
87											
88											
89											
90											
91											
92											
93											
94											
95											
96											
97											
98											
99											
100											

Vista de datos Vista de variables

IBM SPSS Statistics Processor está listo Unicode: ON

14:57 30/09/2015

## Las cuatro ventanas: Salida de resultados

- Una ventana de salida de resultados: visualiza la salida de los resultados.
- Puede grabar los resultados en ficheros de extensión “spv”.
- Aparece los comandos que se introducen a SPSS en la parte de arriba.
- Aparece un árbol de operaciones a la izquierda.
- Puedes modificar las etiquetas de la salida.
- p. ej. marcar columnas con el ratón>botón derecho>estadísticos descriptivos.

# Las cuatro ventanas: Salida de resultados

The screenshot shows the IBM SPSS Statistics Editor de datos interface. The main window displays a data view with columns: edad, marital, direc, and inc. The results window, titled '\*Resultado1 [Documento1] - IBM SPSS Statistics Visor', shows the following command and output:

```
FRECUENCIAS VARIABLES=edad marital
/STATISTICS=RANGE MINIMUM MAXIMUM STDDEV MEAN MEDIAN
/FORMAT=LIMIT(50)
/ORDER=ANALYSIS.
```

**+ Frecuencias**

**Estadísticos**

		Edad en años	Estado civil
N	Válido	6400	6400
	Perdidos	0	0
Media		42,06	,50
Mediana		41,00	,00
Desviación estándar		12,290	,500
Rango		59	1
Mínimo		18	0
Máximo		77	1

**Tabla de frecuencia**

**Estado civil**

		Frecuencia	Porcentaje	Porcentaje válido	Porcentaje acumulado
Válido	Sin casar	3224	50,4	50,4	50,4
	Casado	3176	49,6	49,6	100,0
	Total	6400	100,0	100,0	



# Las cuatro ventanas: Editor de sintaxis

- Es un editor para composición de comandos y lenguaje de SPSS y luego ejecutarlo.
- Se pueden grabar las hojas de composición de expresiones en SPSS con la extensión “sps”. Son archivos de texto.
- Puede ahorrar y automatizar muchas tareas comunes mediante el eficaz lenguaje de comandos.
- El lenguaje de comandos también proporciona algunas funcionalidades no incluidas en los menús y cuadros de diálogo.
- El lenguaje de comandos también permite guardar los trabajos en un archivo de sintaxis, con lo que podrá repetir los análisis en otro momento.

# Las cuatro ventanas: Editor de sintaxis

The screenshot displays the IBM SPSS Statistics Editor de datos interface. The main window shows a data grid with columns for variables like 'edad', 'marital', 'ingres', 'ingcat', 'coche', 'cochecat', 'educ', 'empleo', 'retirado', 'empcat', 'satlab', 'genero', 'residen', 'inalam', 'multiline', 'voz', 'busca', 'internet', 'idllam', 'espera', and 'tv'. The 'ingres' variable is selected for analysis.

The 'SintaxisCursoDemo.sps - IBM SPSS Statistics Editor de sintaxis' window shows the following syntax code:

```
GET
DATASET NAME
FREQUENCIES
1 GET
2 FILE="C:\Users\vrodrig\Documents\SPSS\demo.sav".
3 DATASET NAME Conjunto_de_datos1 WINDOW=FRONT.
4
5 FREQUENCIES VARIABLES=ingres
6 /STATISTICS=RANGE MINIMUM MAXIMUM STDDEV MEAN MEDIAN
7 /FORMAT=LIMIT(50)
8 /ORDER=ANALYSIS.
```

The 'Resultado' window displays the output for 'FRECUENCIAS VARIABLES=ingres'. The 'Estadísticos' section shows the following summary statistics for 'Ingresos del hogar en miles':

Estadísticos		
Ingresos del hogar en miles		
N	Válido	6400
	Perdidos	0
Media		69,4748
Mediana		45,0000
Desviación estándar		78,71856
Rango		1107,00
Mínimo		9,00
Máximo		1116,00

# Las cuatro ventanas: Ventana de Script

- Brinda la oportunidad de escribir programas en toda regla, en un lenguaje muy similar al BASIC.
- La ventana es un simple editor de texto para la composición de sintaxis. La extensión del archivo guardado será "sbs".
- El fichero se puede ejecutar en el SPSS.
- Esta ventana es para usuarios más avanzados.
- Algunos ejemplos en:
  - [http://www.ibm.com/support/knowledgecenter/es/SSLVQG\\_7.0.1/datacollection\\_cads\\_ddita/datacollection/mrstudio/xml/dmgr\\_cads\\_script\\_example.html](http://www.ibm.com/support/knowledgecenter/es/SSLVQG_7.0.1/datacollection_cads_ddita/datacollection/mrstudio/xml/dmgr_cads_script_example.html)

# Las cuatro ventanas: Ventana de Script

The screenshot displays the IBM SPSS Statistics Editor de datos interface. The main window shows a data table with 29 variables and 37 rows. A 'Script1 \* (script) - SPSS Statistics Basic Script Editor (design)' window is overlaid on top, showing a macro script. The script includes comments and a sub-procedure named 'Main' that handles file paths and directory operations.

Visible: 29 de 29 variables

	edad	marital	direcc	ingres	ingcat	coche	cochecat	educ	empleo	retirado	empcat	satlab	genero	residen	inalam	multiline	voz	busca	internet	idllam	espera	tv
1	55	1	12	72.00	3.00	36.20	3.00	1	23	0	3	5 m		4	0	0	1	0	0	0	0	1
2	56	0	29	153.00	4.00	76.90	3.00	1	35	0	3	4 h		1	1	0	1	1	0	1	1	1
3	28	1	9	28.00	2.00	13.70	1.00	3	4	0	1	3 m		3	1	0	1	0	0	1	1	1
4	24	1	4	26.00	2.00	12.50	1.00	4	0	0	1	1 h		3	1	1	1	0	0	0	1	1
5	25	0	2	23.00	1.00	11.30	1.00	2	5	0	2	2 h		2	0	0	0	0	0	1	0	1
6	45	1	9	76.00	4.00	37.20	3.00	3	13	0	2	2 h		2	0	1	1	1	0	0	0	1
7	42																					
8	35																					
9	46																					
10	34																					
11	55																					
12	28																					
13	31																					
14	42																					
15	35																					
16	52																					
17	21																					
18	32																					
19	42																					
20	40																					
21	30																					
22	48																					
23	39																					
24	42																					
25	45																					
26	51																					
27	39																					
28	49																					
29	52																					
30	53																					
31	34																					
32	47																					
33	58																					
34	25																					
35	57	1	28	92.00	4.00	45.50	3.00	2	25	0	3	5 m		4	0	1	1	0	0	1	1	1
36	30	1	7	21.00	1.00	10.50	1.00	4	4	0	1	1 h		2	0	0	0	0	1	0	1	1
37	21	0	0	13.00	1.00	6.30	1.00	3	0	0	1	1 m		2	0	0	0	0	1	0	0	1

```
Script1 * (script) - SPSS Statistics Basic Script Editor (design)
Archivo Edición Ver Macro Depurar Hoja Ayuda
Object: [General] Proc: [Main]
1 'Add file date to file name.
2
3
4 'Adds file date to file names meeting given file path and file mask.
5 'Posted to SPSS-L list by Reynald Levesque on 2003/05/21
6 'levesque@videotron.ca http://pages.infinit.net/levesqu/index.htm
7
8
9
10
11
12
13
14
15
16 Sub Main
17   Path$="c:\test\"
18   F$ = Dir$(Path$ & "*.spo")
19   Debug.Print F$
20   While F$ <> ""
21     T$=Str$(FileDate$(Path$ & F$))
22     Debug.Print F$:T$
23     T$=Replace(T$,"/","-")
24     Name Path$ & F$ As Path$ & Left(F$,InStr(F$,"-")-1) & "-" & Left(T$,InStr(T$,"-")-1) & ".spo"
25     F$ = Dir$(Path$)
26   End While
27 End Sub
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
NUM 19
```

# Análisis Básico

## ➤ Frecuencias

- Este análisis genera tablas de frecuencias que muestran recuentos de frecuencias y porcentajes de los valores de las variables individuales.

## ➤ Descriptivos

- Este análisis muestra el máximo, mínimo, media y desviación estándar de las variables

## ➤ Análisis de regresión lineal

- La regresión lineal estima los coeficientes de la ecuación lineal

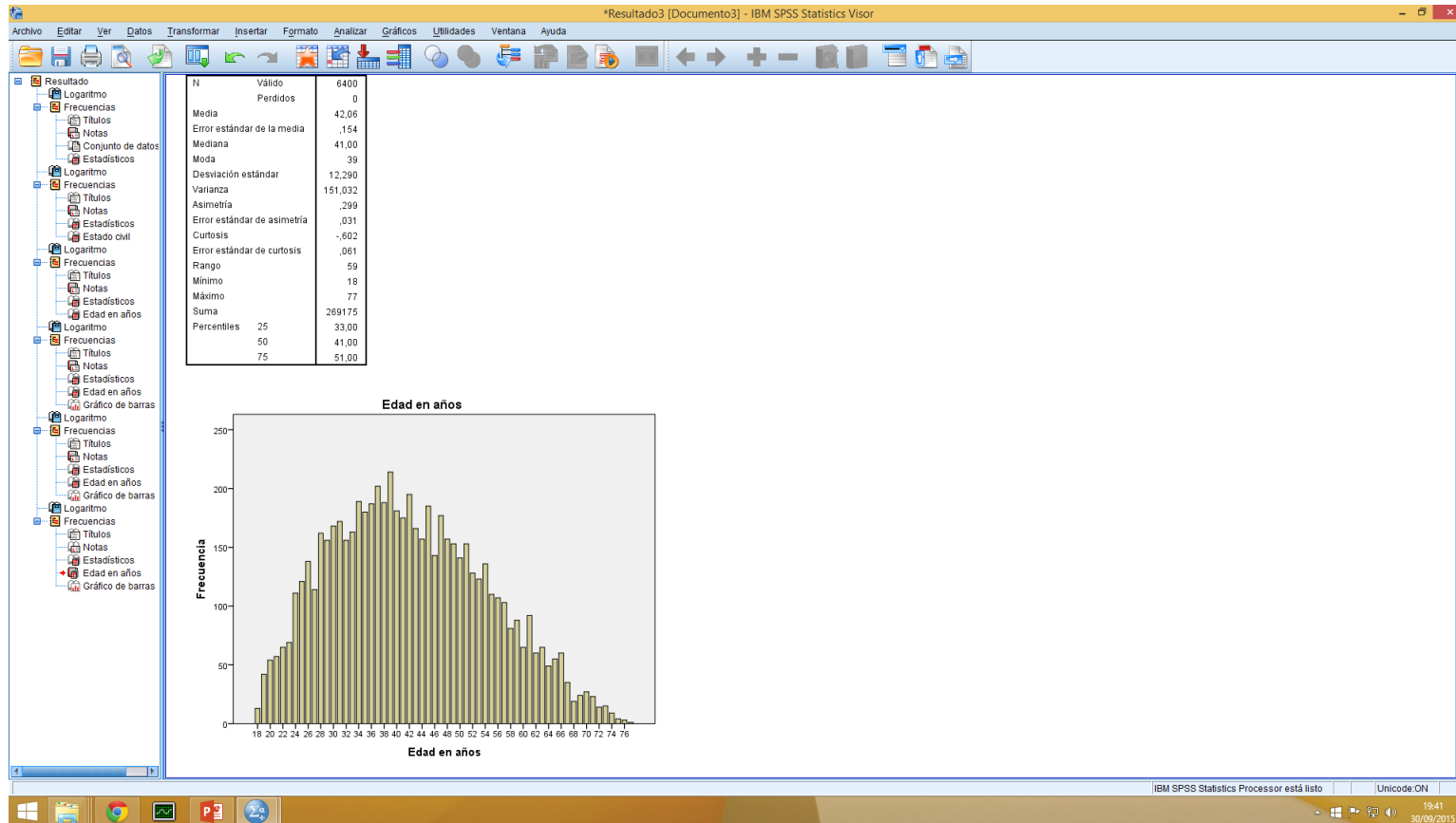
# Análisis Básico: Frecuencias

- Haga clic en Analizar>estadísticos descriptivos>Frecuencias
- Haga clic en la variable de estudio y muévela a la derecha
  - En estadísticos: se puede seleccionar lo que quieres presentar.
  - En gráficos puedes elegir barras, histograma, etc.
- En aceptar presenta los cálculos y gráficas.

The image displays four screenshots from the SPSS software interface, illustrating the steps to generate frequency statistics and charts for a variable.

- Top Left:** A partial view of the 'Frecuencias' dialog box, showing the 'Mostrar tablas de frecuencias' section with buttons for 'Aceptar', 'Pegar', 'Restablecer', and 'Cancelar'.
- Center:** The 'Frecuencias: Estadísticos' dialog box. The 'Mostrar' section is checked. Under 'Tendencia central', 'Media', 'Mediana', 'Moda', and 'Suma' are selected. Under 'Dispersión', 'Desviación estándar', 'Varianza', 'Rango', 'Mínimo', 'Máximo', and 'Error estándar media' are selected. Under 'Distribución', 'Asimetría' and 'Curtosis' are selected. A 'Puntos de corte para' field is set to 10 'grupos iguales'. A 'Percentiles' field is set to 25,0. Buttons at the bottom include 'Continuar', 'Cancelar', and 'Ayuda'.
- Center Right:** The 'Frecuencias' dialog box. The variable 'Edad en años [edad]' is moved to the 'Mostrar' list. The 'Mostrar tablas de frecuencias' checkbox is checked. Buttons at the bottom include 'Aceptar', 'Pegar', 'Restablecer', 'Cancelar', and 'Ayuda'.
- Far Right:** The 'Frecuencias: Gráficos' dialog box. The 'Tipo de gráfico' section has 'Gráficos de barras' selected. The 'Mostrar curva normal en el histograma' checkbox is checked. Buttons at the bottom include 'Continuar', 'Cancelar', and 'Ayuda'.

# Análisis Básico: Frecuencias



# Análisis Básico: Frecuencias

- Todos procedimientos utilizando la GUI de SPSS, se pueden hacer mediante el editor de sintaxis, copiando los comandos que aparecen al principio de los cálculos del visor de resultados:

The screenshot illustrates the process of copying SPSS commands from the Results Viewer to the Syntax Editor. The Results Viewer window, titled '\*Resultado3 [Documento3] - IBM SPSS Statistics Visor', shows the following statistics for 'Edad en años':

Estadísticos	
N	Válido 6400
	Perdidos 0
Media	42,06
Error estándar de la media	,154
Mediana	41,00
Moda	39
Desviación estándar	12,290
Varianza	151,032
Asimetría	,299
Error estándar de asimetría	,031
Curtoisía	-,602
Error estándar de curtoisía	,081
Rango	59
Mínimo	18
Máximo	77
Suma	269175
Percentiles	25 33,00
	50 41,00
	75 51,00

The Syntax Editor window, titled '\*SintaxisCursoDemo.sps - IBM SPSS Statistics Editor de sintaxis', shows the following commands:

```
GET  
DATASET NAME  
FILE='C:\Users\fooding\Documents\SPSS\demo.sav'  
FRECUENCIAS  
FRECUENCIAS  
FRECUENCIAS  
/STATISTICS=RANGE MINIMUM MAXIMUM STDDV MEAN MEDIAN  
/FORMAT=HLIMIT(50)  
/ORDER=ANALYSIS.  
FRECUENCIAS VARIABLES=edad  
/NTILES=4  
/PERCENTILES=25,0  
/STATISTICS=STDDV VARIANCE RANGE MINIMUM MAXIMUM SEMEAN MEAN MEDIAN MODE SUM SKEWNESS SESKEW KURTOSIS SEK  
/BARCHART FREQ  
/ORDER=ANALYSIS.
```

An orange arrow points from the first command in the Syntax Editor to the corresponding command in the Results Viewer's command window.



# Análisis Básico: Descriptivos

- Este análisis muestra el máximo, mínimo, media y desviación estándar de las variables.
- Haga clic en Analizar > estadísticos descriptivos > Descriptivos
- Haga clic en la variable de estudio y muévela a la derecha
- En aceptar presenta los cálculos y gráficas.

DESCRIPTIVES VARIABLES=edad  
/STATISTICS=MEAN STDDEV MIN MAX.

➤ **Descriptivos**

Estadísticos descriptivos					
	N	Mínimo	Máximo	Media	Desviación estándar
Edad en años	6400	18	77	42,06	12,290
N válido (por lista)	6400				

# Análisis Básico: Análisis de regresión lineal

- La regresión lineal estima los coeficientes de la ecuación lineal
- Haga clic en Analizar>Regresiones>Lineales y por ejemplo:

REGRESIÓN

```
REGRESIÓN  
/MISSING=LISTWISE  
/STATISTICS=COEFF OUTS R ANOVA  
/CRITERIA=PIN(.05) POUT(.10)  
/NOORIGIN  
/DEPENDENT=edad  
/METHOD=ENTER ingresos.
```

### Regresión

Variables entradas/eliminadas<sup>a</sup>

Modelo	Variables introducidas	Variables eliminadas	Método
1	Ingresos del hogar en miles <sup>b</sup>		Intro

a. Variable dependiente: Edad en años  
b. Todas las variables solicitadas introducidas.

### Resumen del modelo

Modelo	R	R cuadrado	R cuadrado ajustado	Error estándar de la estimación
1	.335 <sup>a</sup>	.112	.112	11.579

a. Predictores: (Constante), Ingresos del hogar en miles

### ANOVA<sup>a</sup>

Modelo	Regresión	Suma de cuadrados	gl	Media cuadrática	F	Sig.
1	Regresión	108590.209	1	108590.209	809.871	.000 <sup>b</sup>
	Residuo	857864.818	6398	134.083		
	Total	966455.027	6399			

a. Variable dependiente: Edad en años  
b. Predictores: (Constante), Ingresos del hogar en miles

### Coefficientes<sup>a</sup>

Modelo		Coeficientes no estandarizados		Coeficientes estandarizados		Sig.
		B	Error estándar	Beta	t	
1	(Constante)	38.423			199.020	.000
	Ingresos del hogar en miles	.052	.002	.335	28.458	.000

# Análisis Básico: Análisis de regresión lineal

## ➤ Pintar la regresión: Gráficos > Gráficos de variables de regresión.

The screenshot displays the IBM SPSS Statistics interface. The main window shows the 'Gráficos de variables de regresión' dialog box, which is used to configure regression plots. The dialog includes options for selecting variables for the horizontal and vertical axes, and for applying various styles and options to the plots.

The 'Gráficos de variables de regresión' dialog box shows the following configuration:

- Variables de eje vertical:** Edad en años [edad]
- Variables de eje horizontal:** Precio del coche principal [coche]
- Variables:** Categoría del precio del coche principal [coche] (selected), Nivel educativo [educ], Años con la empresa actual [empleo], Retirado [retirado], Años con la empresa actual [empcat], Satisfacción laboral [satisfac], Género [genero], Número de personas en el hogar [residen], Servicio inalámbrico [inalam], Múltiples líneas [multiline], Buzón de voz [voz], Buscapersonas [busca], Internet [internet], ID llamadas [idllam], Llamada en espera [espera], Tiene TV [tv], Tiene Video [vide], Tiene Hi-Fi/CD [cd], Tiene PDA [pda], Tiene ordenador [pc], Tiene fax [fax].
- Estilo:** LINEAR (selected)
- Opciones:** Agrupar puntos en diagramas de dispersión (desselected), Diagramas de caja con bordes para diagramas de dispersión, si hay uno por fila (selected).

The 'GGraph' plot shows a scatter plot of 'Edad en años' (Y-axis) versus 'Precio del coche principal' (X-axis). The plot includes a linear regression line and a box plot for the Y-axis variable. The regression line is shown as a solid black line, and the box plot is shown as a vertical bar with a horizontal line indicating the median.

The 'Opciones...' dialog box shows the following configuration:

- Título de gráfico:** (empty)
- Líneas de ajuste de diagrama de dispersión:** Lineal (selected), Cuadrático, Cúbico, LQESS.
- Agrupación:** Línea de ajuste total (selected), Ajustar línea para cada grupo de colores categóricos.
- Gráficos de variables categóricas:** Gráficos de barras de medios (selected), Gráficos de líneas de medios, Diagramas de caja.
- Parámetros de diseño del gráfico:** Número de gráficos por fila: 1.
- Parámetros de múltiples gráficos por fila:** Sangría del primer gráfico (%): 15, Escalamiento de la dimensión Y (%): 75.

# Análisis Básico: Análisis de regresión lineal

## ➤ Pintar la regresión: Gráficos > Gráficos de variables de regresión.

The screenshot displays the IBM SPSS Statistics Visor interface. The main window shows a regression plot titled "STATS REGRESS" with the following command: `STATS REGRESS PLOT YVARS=edad XVARS=coche COLOR=educ /OPTIONS CATEGORICAL=BAR GROUP=1 BOXPLOTS INDEXT=15 YSCALE=75 /FITLINES LINEAR APPLYTO=TOTAL.` The plot shows a scatter of data points colored by education level (educ) with a linear regression line. The y-axis is labeled "Edad en años" and the x-axis is "Precio del coche principal".

Overlaid on the plot is a dialog box titled "Gráficos de variables de regresión". It contains the following information:

- Variables de eje vertical:  Edad en años [edad]
- Variables de eje horizontal:  Precio del coche principal [coche]
- Variables list: Categoría del precio del coche principal [cochecat], Nivel educativo [educ], Años con la empresa actual [empleo], Retirado [retirado], Años con la empresa actual [empcat], Satisfacción laboral [satiab], Género [genero], Número de personas en el hogar [residen], Servicio inalámbrico [inalam], Múltiples líneas [multiline], Buzón de voz [voz], Buscapersonas [busca], Internet [internet], ID llamadas [idllam], Llamada en espera [espera], Tiene TV [tv], Tiene Video [video], Tiene Hi-KCD [pcd], Tiene PDA [pda], Tiene ordenador [pc], Tiene fax [fax].

Below the main dialog is an "Opciones..." dialog box with the following settings:

- Título de gráfico: (empty)
- Agrupar puntos en diagramas de dispersión
- Diagramas de caja con bordes para diagramas de dispersión, si hay uno por fila
- Parámetros de diseño del gráfico:
  - Número de gráficos por fila: 1
  - Tamaño de página vertical (pulgadas): (empty)
  - Tamaño de página horizontal (pulgadas): (empty)
- Parámetros de múltiples gráficos por fila:
  - Sangría del primer gráfico (%): 15
  - Escalamiento de la dimensión Y (%): 75

On the left side of the main window, there is a "STATS REGRESS" panel with a table titled "Información de la leyenda del gráfico":

Configuración	Valor
Colorear	Nivel educativo
Aplicar tamaño	---
Aplicar forma	---
Etiquetar	---
Lineas de ajuste	LINEAR (solid)

Below the table is a "GGraph" section with a legend for the regression line types:  Lineal,  Cuadrático,  Cúbico,  LOESS. Under "Agrupación",  Línea de ajuste total and  Ajustar línea para cada grupo de colores categóricos. Under "Gráficos de variables categóricas",  Gráficos de barras de medios,  Gráficos de líneas de medios, and  Diagramas de caja.

# Análisis Básico: Análisis de regresión lineal

- Pintar la regresión: Gráficos > Gráficos de variables de regresión.

The screenshot displays the IBM SPSS Statistics interface. The main window shows a 'STATS REGRESS' plot with a scatter plot of 'Edad en años' (Age in years) on the y-axis and 'Precio del coche principal' (Main car price) on the x-axis. A regression line is fitted to the data points, which are color-coded by 'Nivel educativo' (Education level). The plot includes marginal boxplots for both variables. The 'Gráficos de variables de regresión' dialog box is open, showing the configuration for the regression plot. The 'Variables de eje vertical' (Vertical axis variables) list includes 'Edad en años [edad]'. The 'Variables de eje horizontal' (Horizontal axis variables) list includes 'Precio del coche principal [coche]'. The 'Colorear' (Color) section is set to 'Nivel educativo [educ]'. The 'Aplicar tamaño' (Apply size) section is set to 'Tiene fax [fax]'. The 'Aplicar forma' (Apply shape) section is set to 'Tiene fax [fax]'. The 'Etiquetar' (Label) section is empty. The dialog box also contains a list of variables for the plot, including 'Edad en años [edad]', 'Estado civil [mantat]', 'Años en la dirección actual [direcc]', 'Ingresos del hogar en miles [ingres]', 'Categoría de ingresos en miles [ingcat]', 'Precio del coche principal [coche]', 'Categoría del precio del coche principal [cocheca]', 'Nivel educativo [educ]', 'Años con la empresa actual [empleo]', 'Retraido [retraido]', 'Años con la empresa actual [empcat]', 'Satisfacción laboral [satsat]', 'Género [genero]', 'Número de personas en el hogar [residen]', 'Servicio inalámbrico [inalam]', 'Múltiples líneas [multiline]', 'Buzón de voz [voz]', 'Buscaperosnas [busca]', 'Internet [internet]', 'ID llamadas [diliam]', and 'Llamada en espera [espera]'. The status bar at the bottom indicates 'IBM SPSS Statistics Processor está listo' and 'Unicode ON'.

STATS REGRESS PLOT YVARS=edad XVARS=coche COLOR=educ SIZE=fax  
/OPTIONS CATEGORICAL=BAR GROUP=1 BOXPLOTS INDENT=15 YSCALE=75  
/FILLINES LINEAR APPLYTO=TOTAL.

**STATS REGRESS**

Información de la leyenda del gráfico

Configuración	Valor
Colorear	Nivel educativo
Aplicar tamaño	Tiene fax
Aplicar forma	---
Etiquetar	---
Líneas de ajuste	LINEAR (solid)

Configuración de la descripción para los gráficos que siguen  
Alguna configuración no se aplica a los gráficos categóricos.

**GGraph**

Edad en años

Precio del coche principal

Este cuadro de diálogo requiere Python Essentials

Gráficos de variables de regresión

Utilice este cuadro de diálogo para trazar cada variable de eje vertical frente a cada variable de eje horizontal.  
Para las variables categóricas, se crean gráficos de barras, de líneas o de diagramas de caja.

Variables de eje vertical:

- Edad en años [edad]

Variables de eje horizontal:

- Precio del coche principal [coche]

Las especificaciones siguientes se aplican únicamente a gráficos de predictores continuos excepto Etiquetar, que se aplica también a diagramas de caja

Colorear:

- Nivel educativo [educ]

Aplicar tamaño:

- Tiene fax [fax]

Aplicar forma:

- Tiene fax [fax]

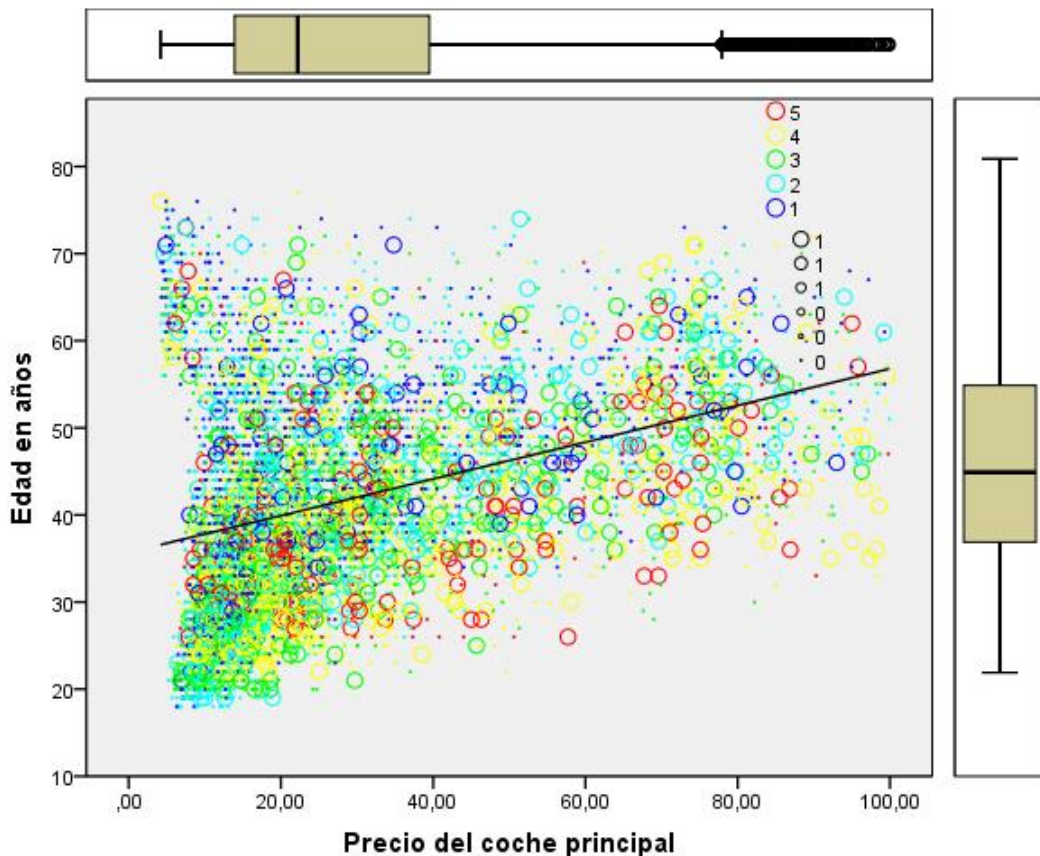
Etiquetar:

- 

Este cuadro de diálogo requiere Python Essentials

Aceptar Pegar Restablecer Cancelar Ayuda

# Análisis Básico: Análisis de regresión lineal



## Información de la leyenda del gráfico

Configuración	Valor
Colorear	Nivel educativo
Aplicar tamaño	Tiene fax
Aplicar forma	---
Etiquetar	---
Líneas de ajuste	LINEAR (solid)

Configuración de la descripción para los gráficos que siguen. Algunas configuraciones no se aplican a los gráficos categóricos.

# Representaciones Visuales: Histogramas

IBM SPSS Statistics Visor - \*Resultado3 [Documento3]

Archivo Editar Ver Datos Transformar Insertar Formato Analizar Gráficos Utilidades Ventana Ayuda

Edad en años

```
* Generador de gráficos.
GGRAPH
  /GRAPHDATASET NAME="graphdataset" VARIABLES=edad MISSING=LISTWISE REPORTMISSING=NO
  /GRAPHSPEC SOURCE=INLINE.
BEGIN GPL
  SOURCE: s=userSource(id("graphdataset"))
  DATA: edad=col(source(s), name("edad"))
  GUIDE: axis(dim(1), label("Edad en años"))
  GUIDE: axis(dim(2), label("Frecuencia"))
  ELEMENT: interval(position(summary.count(bin.rect(edad))), shape.interior(shape.square))
END GPL.
```

➔ GGraph

Media = 42,06  
Desviación estándar = 12,29  
N = 6.400

Generador de gráficos

Variables: Edad en años [e...]

La vista previa del gráfico utiliza datos de ejemplo

No categorías (scale variable)

Galería Elementos básicos Grupos/ID de puntos Títulos/notas al pie

Elija entre:

- Favoritos
- Barras
- Líneas
- Áreas
- Circular/Polár
- Dispersión/Puntos
- Histograma
- Máximos-mínimos
- Diagramas de cajas
- Ejes dobles

Aceptar Pegar Restablecer Cancelar Ayuda

# Representaciones Visuales: Matriz de dispersión

\*Resultado3 [Documento3] - IBM SPSS Statistics Visor

```
GGRAPH
  /GRAPHDATASET NAME="graphdataset" VARIABLES=edad coche direcc MISSING=LISTWISE REPORTMISSING=NO
  /GRAPHSPEC SOURCE=INLINE.
BEGIN GPL
  SOURCE: s=userSource(id("graphdataset"))
  DATA: edad=col(source(s), name("edad"))
  DATA: coche=col(source(s), name("coche"))
  DATA: direcc=col(source(s), name("direcc"))
  GUIDE: axis(dim(1,1), ticks(null()))
  GUIDE: axis(dim(2,1), ticks(null()))
  GUIDE: axis(dim(1), gsp(0px))
  GUIDE: axis(dim(2), gsp(0px))
  TRANS: edad_label = eval("Edad en años")
  TRANS: coche_label = eval("Precio del coche principal")
  TRANS: direcc_label = eval("Años en la dirección actual")
  ELEMENT: point(position(edad/edad_label+coche/coche_label+direcc/direcc_label)*(ed
```

**Generador de gráficos** (Efectúa una doble pulsación para activar)

Variables:

- Edad en años [e...]
- Estado civil [marit...]
- Años en la direcc...
- Ingresos del hog...
- Categoría de ingr...
- Precio del coche ...
- Categoría del pre...
- Nivel educativo [e...]
- Años con la empr...
- Retirado [retirado]
- Años con la empr...
- Satisfacción labo...
- Género [genero]
- Número de pers...

No categorías (scale variable)

Edad en años; Precio del coche principal; Años en la dirección actual

Galería | Elementos básicos | Grupos/ID de puntos | Títulos/notas al pie

Elija entre:

- Favoritos
- Barras
- Líneas
- Áreas
- Circular/Polár
- Dispersión/Puntos
- Histograma
- Máximos-mínimos
- Diagramas de cajas
- Ejes dobles

Propiedades de elemento... Opciones...

Aceptar | Pegar | Restablecer | Cancelar | Ayuda

**GGraph**

Edad en años

Precio del coche principal

Años en la dirección actual

Edad en años | Precio del coche principal | Años en la dirección actual



# Representaciones Visuales: Boxplots

\*Resultado3 [Documento3] - IBM SPSS Statistics Visor

Archivo Editar Ver Datos Transformar Insertar Formato Analizar Gráficos Utilidades Ventana Ayuda

0.0 20 40 60

Años en la dirección actual

```
* Generador de gráficos.  
GGRAPH  
  /GRAPHDATASET NAME="graphdataset" VARIABLES=direcc MISSING=LISTWISE REPORTMISSING=NO  
  /GRAPHSPEC SOURCE=INLINE.  
BEGIN GPL  
  SOURCE: s=userSource(id("graphdataset"))  
  DATA: direcc=col(source(s), name("direcc"))  
  DATA: id=col(source(s), name("$CASENUM"), unit.category(1))  
  GUIDE: axis(dim(2), label("Años en la dirección actual"))  
  ELEMENT: schema(position(bin.quantile.letter(1*direcc)), label(id))  
END GPL.
```

Etchete una doble pulsación para activar

Generador de gráficos

La vista previa del gráfico utiliza datos de ejemplo

Variables:

- Edad en años [e...]
- Estado civil [mant...]
- Años en la direcc...
- Ingresos del hog...
- Categoría de ingr...
- Precio del coche ...
- Categoría del pre...
- Nivel educativo [e...
- Años con la empr...
- Retirado [retirado]
- Años con la empr...

No categorías (scale variable)

Eje X?

Galería Elementos básicos Grupos/ID de puntos Títulos/notas al pie

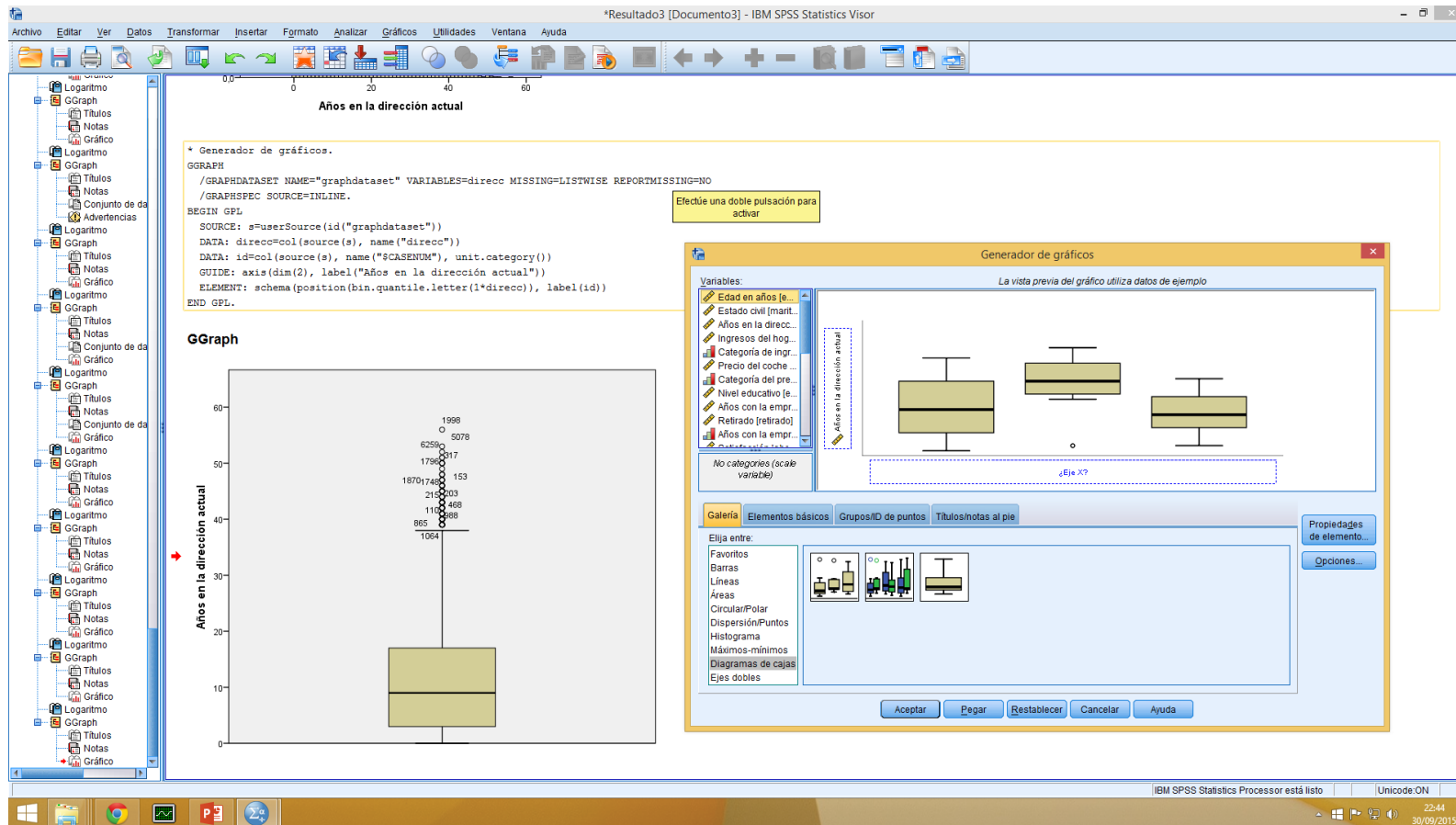
Elija entre:

- Favoritos
- Barras
- Líneas
- Áreas
- Circular/Polár
- Dispersión/Puntos
- Histograma
- Máximos-mínimos
- Diagramas de cajas
- Ejes dobles

Aceptar Pegar Restablecer Cancelar Ayuda

Propiedades de elemento... Opciones...

IBM SPSS Statistics Processor está listo Unicode ON 22:44 30/09/2015



# Representaciones Visuales: Histogramas

\*Resultado3 [Documento3] - IBM SPSS Statistics Visor

Archivo Editar Ver Datos Transformar Insertar Formato Analizar Gráficos Utilidades Ventana Ayuda

```

* Generador de gráficos.
GGRAPH
  /GRAPHDATASET NAME="graphdataset" VARIABLES=direcc MISSING=LISTWISE REPORTMISSING=NO
  /GRAPHSPEC SOURCE=INLINE.
BEGIN GPL
  SOURCE: s=userSource(id("graphdataset"))
  DATA: direcc=col(source(s), name("direcc"))
  GUIDE: axis(dim(1), label("Años en la dirección actual"))
  GUIDE: axis(dim(2), label("Frecuencia"))
  ELEMENT: interval(position(summary.count(bin.rect(direcc))), shape.interior(shape.square))
END GPL.

```

**GGraph**

Media = 11,56  
Desviación estándar = 9,938  
N = 6.400

**Generador de gráficos**

Efectúe una doble pulsación para activar

La vista previa del gráfico utiliza datos de ejemplo

Edad en años (e...)  
Estado civil (marit...)  
Años en la direcc...  
Ingresos del hog...  
Categoría de Ingr...  
Precio del coche ...  
Categoría del pre...  
Nivel educativo [e...  
Años con la empr...  
Retirado (retirado)  
Años con la empr...

No categorías (scale variable)

Histograma

Años en la dirección actual

Galería Elementos básicos Grupos/ID de puntos Títulos/notas al pie

Elija entre:

- Favoritos
- Barras
- Líneas
- Áreas
- Circular/Polar
- Dispersión/Puntos
- Histograma**
- Máximos-mínimos
- Diagramas de cajas
- Ejes dobles

Propiedades de elemento...  
Opciones...

Aceptar Pegar Restablecer Cancelar Ayuda

# Representaciones Visuales: QQplot

The image shows the IBM SPSS Statistics interface. The main window displays a list of variables on the left and a central area with a QQ plot titled "Gráfico Q-Q Normal de Edad en años". The plot shows "Valor observado" on the x-axis and "Valor Normal esperado" on the y-axis, both ranging from 0 to 100. The data points follow a diagonal line, indicating a normal distribution. A "Gráficos Q-Q" dialog box is open on the right, showing the variable "Edad en años [edad]" selected. The dialog includes options for "Distribución de prueba" (Normal), "Estimar a partir de los datos", and "Rango asignado a empates" (Media).

IBM SPSS Statistics \*Resultado3 [Documento3] - IBM SPSS Statistics Visor

Archivo Editar Ver Datos Transformar Insertar Formato Analizar Gráficos Utilidades Ventana Ayuda

Informes

- Estadísticos descriptivos
  - Frecuencias...
  - Descriptivos...
  - Explorar...
  - Tablas cruzadas...
  - Análisis TURF
  - Razón...
  - Gráficos P-P...
  - Gráficos Q-Q...
- Tablas
- Comparar medias
- Modelo lineal general
- Modelos lineales generalizados
- Modelos mixtos
- Correlaciones
- Regresión
  - Loglineal
- Clasificar
- Reducción de dimensiones
- Escala
- Pruebas no paramétricas
  - Previsiones
  - Supervivencia
  - Respuesta múltiple
- Simulación...
- Control de calidad
  - Cunya COR...

Resumen de procedimiento

Longitud de serie o secuencia

Número de valores perdidos en el gráfico

Los casos no están ponderados.

Parámetros de distribución

Distribución normal Ubicación

Escala

Los casos no están ponderados.

Edad en años

Gráfico Q-Q Normal de Edad en años

Valor Normal esperado

Valor observado

Gráficos Q-Q

Variables:

Edad en años [edad]

Distribución de prueba

Normal

gl: 1

Parámetros de distribución

Estimar a partir de los datos

Posición: 0

Escala: 1

Transformar

Transformación log natural

Estandarizar valores

Diferencia: 1

Diferenciar cdo: 1

Periodicidad actual: Ninguna

Fórmula de estimación de la proporción

De Blom  Rankit  De Tukey

De Van der Waerden

Rango asignado a empates

Media  Mayor  Menor

Romper los empates arbitrariamente

Aceptar Pegar Restablecer Cancelar Ayuda

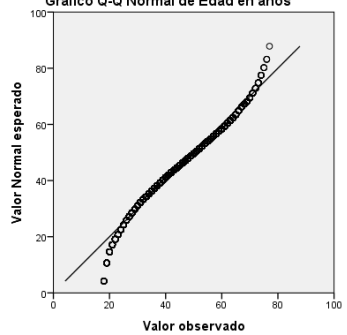
# Representaciones Visuales: QQplot

\*Resultado3 [Documento3] - IBM SPSS Statistics Visor

Archivo Editar Ver Datos Transformar Insertar Formato Analizar Gráficos Utilidades Ventana Ayuda

Edad en años

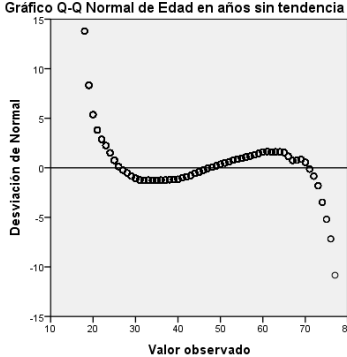
Gráfico Q-Q Normal de Edad en años



Valor Normal esperado

Valor observado

Gráfico Q-Q Normal de Edad en años sin tendencia



Desviación de Normal

Valor observado

Gráficos Q-Q

Variables: Edad en años [edad]

Distribución de prueba: Normal

gt: 1

Estimar a partir de los datos

Posición: 0

Escala: 1

Fórmula de estimación de la proporción: De Blom, Rankit, De Tukey, De Van der Waerden

Rango asignado a empates: Media, Mayor, Menor, Romper los empates arbitrariamente

Transformar: Transformación log natural, Estandarizar valores, Diferencia: 1, Diferenciar cido: 1, Periodicidad actual: Ninguna

Aceptar Pegar Restablecer Cancelar Ayuda

# Bibliografía y lecturas relacionadas:

- [SPSS White Papers](#)
- [SPSS Demos and Tutorials](#)
- Repositorio de IBM: <ftp://public.dhe.ibm.com/>
  - <ftp://public.dhe.ibm.com/software/analytics/spss/>
- [Descripción de todos los archivos muestrales de SPSS](#)
- [Discovering Statistics Using IBM SPSS Statistics. Field, Andy. Sage Publications Ltd, 2012.](#)