



Estadística Descriptiva



Medidas de Tendencia Central

Media

La media nos permite resumir básicamente la información.

$$\frac{1}{n} \sum_{i=1}^n x_i$$

mean = np.mean(x)

Mediana:

Valor que tiene el 50% de los valores que son anteriores y posteriores a él. Es el centil 50.

$$\frac{x_n + x_{n/2+1}}{2}$$

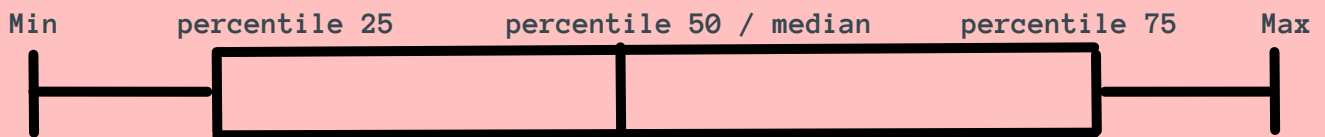
$$x_{\text{int}(\frac{n}{2})+1}$$

n impar

n par

median = np.median(x)

Percentil: El percentil n representa el valor de la muestra que tiene el n% del conjunto de datos anteriores a él.



p es la proporción de valores anteriores al percentil (ex: 0.25, 0.5, 0.99).

percentile = np.quantile(x, p)



Medidas de Dispersión

Var:

Explica la dispersión de los valores en torno a la media.

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2$$

var = np.var(x)

Std:

Nos permite calcular la volatilidad de un activo o dato numérico.

$$\sqrt{\text{var}}$$

volatility = np.std(x)

Sesgo: Nos permite calcular la asimetría de un activo.

skw = scipy.stats.skew(x)

Curtosis: Da información sobre la acumulación de valores alrededor de la media.

kurto = scipy.stats.kurtosis(x)



Medidas de Relación

Cov: Cuantifica la relación entre dos muestras (o variables aleatorias).

$$\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

```
cov=np.cov(array, rowvar=False)[i][j]
```

Matriz de Var/Cov: Matriz que organiza la relación entre varias muestras.

	x1		xn
x1	$\begin{pmatrix} Var(x_1) & \cdots & Cov(x_1, x_n) \\ \vdots & \ddots & \vdots \\ Cov(x_n, x_1) & \cdots & Var(x_n) \end{pmatrix}$		
xn			

```
cov=np.cov(array, rowvar=False)
```

Matriz Correlación: La Correlación normaliza las medidas de relación entre -1 y 1.

	x1		xn
x1	$\begin{pmatrix} Corr(x_1, x_1) = 1 & \cdots & Corr(x_1, x_n) \\ \vdots & \ddots & \vdots \\ Corr(x_n, x_1) & \cdots & Corr(x_n, x_n) = 1 \end{pmatrix}$		
xn			

```
corr = df.corr()
```