

CARACTERÍSTICAS DE LAS DISTRIBUCIONES

En esta hoja de trabajo vas a consolidar los conceptos relacionados con las distribuciones.

Recuerda que una distribución es cómo se reparte la variable cuantitativa a lo largo de los posibles valores.

Nos sirve para darte cuenta dónde se concentra la mayoría de tu variable o si está muy bien repartida.

Un histograma o un boxplot son herramientas excelentes para visualizar las distribuciones.

Pero además podemos utilizar los conceptos de:

- Centralidad
- Dispersión
- Posición

Para poder entender mejor las variables.

Ahora vas a aplicar los conceptos con Excel. Utilizaremos los tablas de datos “cars”

ABRE LA TABLA DE DATOS CARS

Esta tabla de datos tiene dos variables cuantitativas:

- Speed: velocidad a la que va el vehículo
- Dist: la distancia que tarda en frenar

Vas a describir estas variables con la ayuda de Excel. Te muestro las funciones de Excel para calcular las diferentes características:

Característica		Función de Excel
CENTRALIDAD	Media	=promedio("DATOS")
	Mediana (cuartil 2)	=mediana("DATOS")
	Moda	=moda.uno("DATOS")
	Desviación estándar	=desvesta("DATOS")
	Rango Intercuartílico	=cuartil.EXC("DATOS";3)-cuartil.EXC("DATOS";1)
DISPERSIÓN	Error estándar	= desvesta("DATOS")/raiz(n)
	IC 95%	=INTERVALO.CONFIANZA() El resultado sirve para: Intervalo de confianza para la media de una población. Es decir, el intervalo de confianza para la media de la población de desplazamientos al trabajo es $30 \pm 0,692952$ minutos, o de 29,3 a 30,7 minutos. El resultado de la función es : 0,692951912
	Coeficiente de Variación	=desvesta("DATOS")/promedio("DATOS")*100
POSICIÓN	Cuartil 1	=cuartil.EXC("DATOS";1)
	Cuartil 3	=cuartil.EXC("DATOS";3)
	Percentil 25	=PERCENTIL.EXC("DATOS";25)
	Percentil 75	=PERCENTIL.EXC("DATOS";75)
	Percentil 5	=PERCENTIL.EXC("DATOS";5)
	Percentil 95	=PERCENTIL.EXC("DATOS";95)
FORMA	Coeficiente de Variación	=desvesta("DATOS")/promedio("DATOS")*100
	Curtosis	=CURTOSIS("DATOS")
	Asimetría	=COEFICIENTE.ASIMETRIA("DATOS")

CALCULA LAS CARACTERÍSTICAS NUMÉRICAS

Calcula las características de

- Speed: velocidad a la que va el vehículo
- Dist: la distancia que tarda en frenar

Característica	Variable "Speed"	Variable "Dist"
CENTRALIDAD	Media	15,4
	Mediana (cuartil 2)	15
	Moda	20
DISPERSIÓN	Desviación estándar	5,2876
	Rango Intercuartílico	7
	Error estándar	0.7477858473315656
	IC 95%	(14.170001736833694, 16.629998263166307)
	Coefficiente de Variación	0,33
POSICIÓN	Cuartil 1	12
	Cuartil 3	19
	Percentil 25	12
	Percentil 75	19
	Percentil 5	7
	Percentil 95	24
FORMA	Coefficiente de Variación	0,59
	Curtosis	-0.5771474239437371 platicurtica
	Asimetría	-0.11395477012828319 asimetría negativa

Con estos valores calculados te puedes hacer una idea de la variable y describirla.

A parte es muy importante acompañarla con gráficos de la distribución.

DIBUJA LA DISTRIBUCIÓN DE LAS VARIABLES

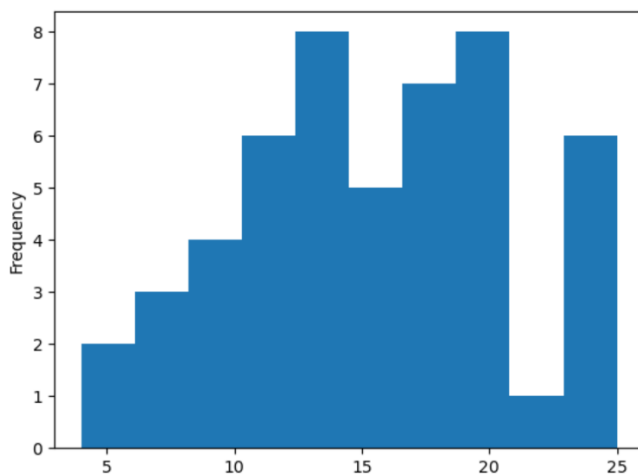
Calcula las características de

- Speed: velocidad a la que va el vehículo
- Dist: la distancia que tarda en frenar

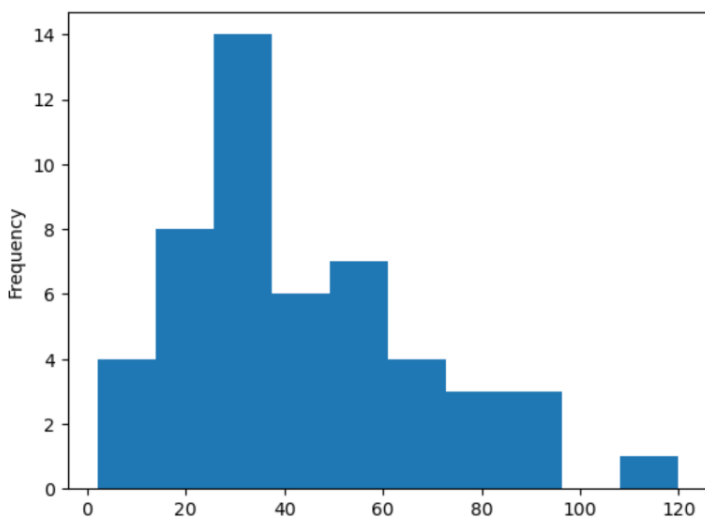
Utiliza el histograma

Copia el histograma que has creado la hoja de trabajo anterior para tener la foto de las variables completa.

```
1 histogram1 = df["speed"].plot.hist()  
2  
3 plt.show()
```



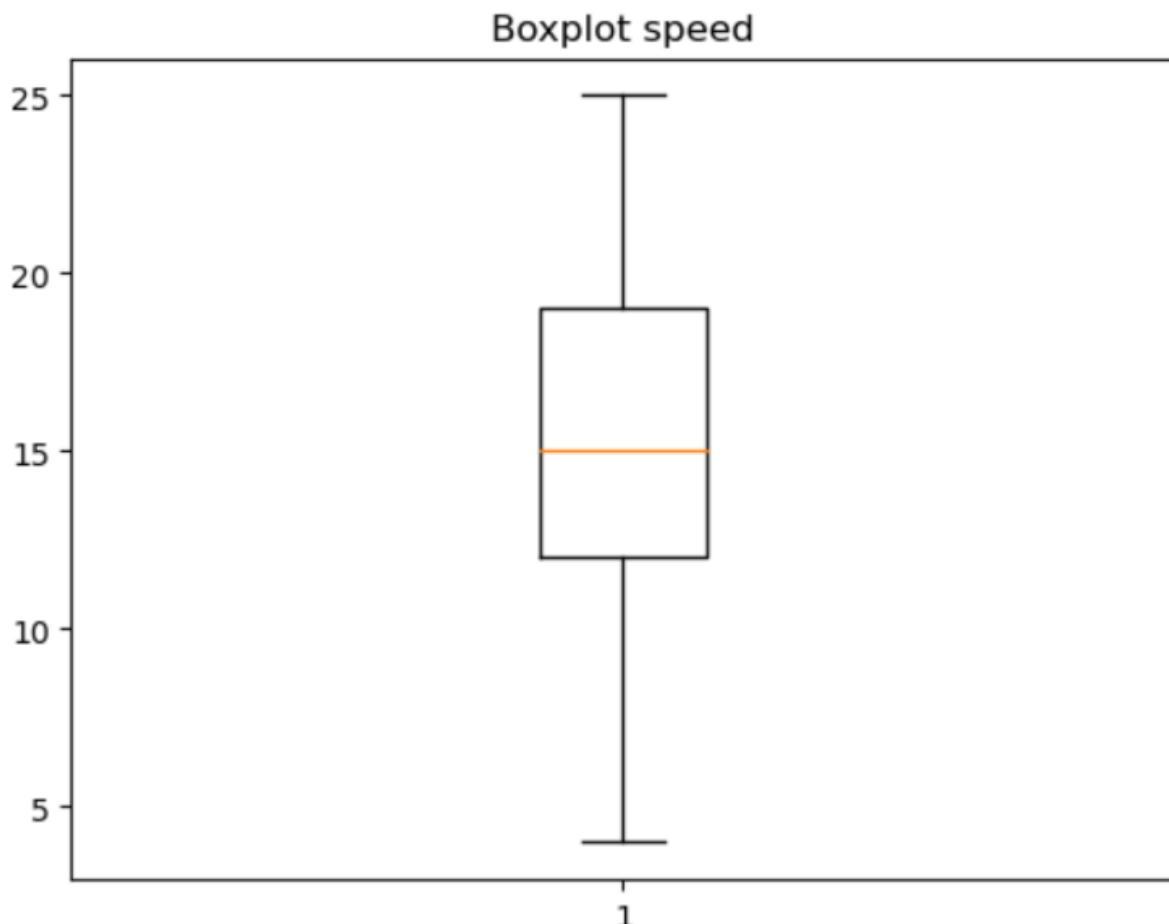
```
1 histogram2 = df["dist"].plot.hist()  
2  
3 plt.show()
```



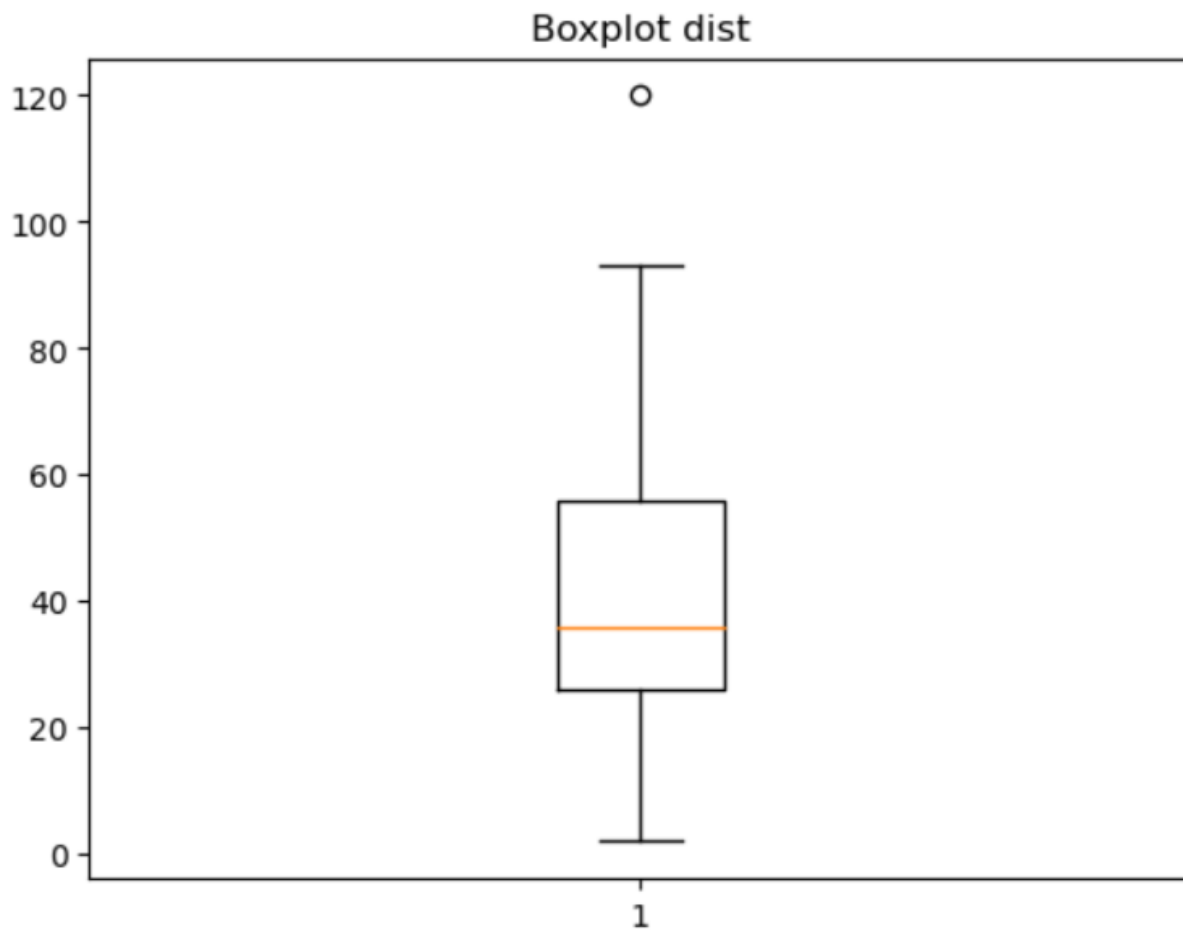
Utiliza el boxplot

Crea dos boxplots. A mano o utilizando Excel u otro programa. Con Excel es un poco difícil de crear pero lo puedes hacer siguiendo el tutorial.

```
1 plt.boxplot(df["speed"])  
2 plt.title("Boxplot speed")  
3 plt.show()
```



```
1 plt.boxplot(df["dist"])
2 plt.title("Boxplot dist")
3 plt.show()
```



¡Buen trabajo!

¡Has consolidado los conceptos de forma práctica!

¡Poco a poco estás asentando las bases! ¡Enhorabuena!