

```
In [6]: import pandas as pd
df = pd.read_csv('titanic.csv')
df
```

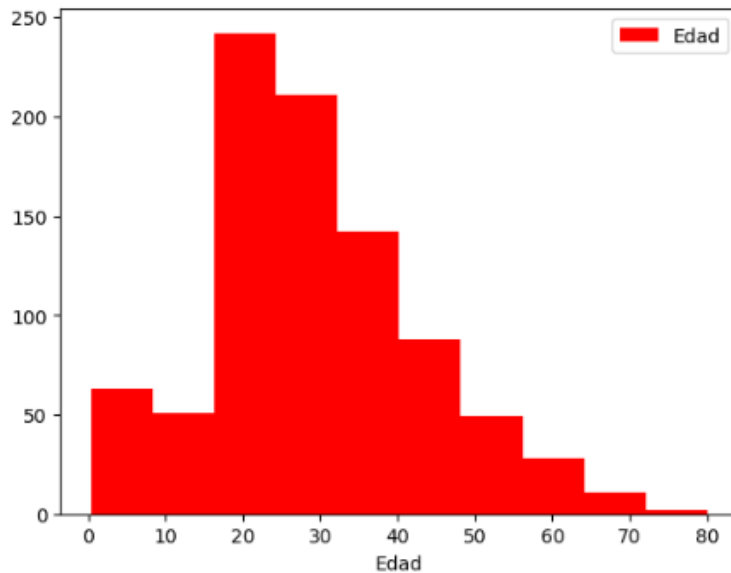
Out[6]:

	Survived	Pclass	Name	Sex	Age	Siblings/Spouses Aboard	Parents/Children Aboard	Fare
0	0	3	Mr. Owen Harris Braund	male	22.0	1	0	7.2500
1	1	1	Mrs. John Bradley (Florence Briggs Thayer) Cum...	female	38.0	1	0	71.2833
2	1	3	Miss. Laina Heikkinen	female	26.0	0	0	7.9250
3	1	1	Mrs. Jacques Heath (Lily May Peel) Futrelle	female	35.0	1	0	53.1000
4	0	3	Mr. William Henry Allen	male	35.0	0	0	8.0500
...	...	...	...	...	...	...	...	...
882	0	2	Rev. Juozas Montvila	male	27.0	0	0	13.0000
883	1	1	Miss. Margaret Edith Graham	female	19.0	0	0	30.0000
884	0	3	Miss. Catherine Helen Johnston	female	7.0	1	2	23.4500
885	1	1	Mr. Karl Howell Behr	male	26.0	0	0	30.0000
886	0	3	Mr. Patrick Dooley	male	32.0	0	0	7.7500

887 rows x 8 columns

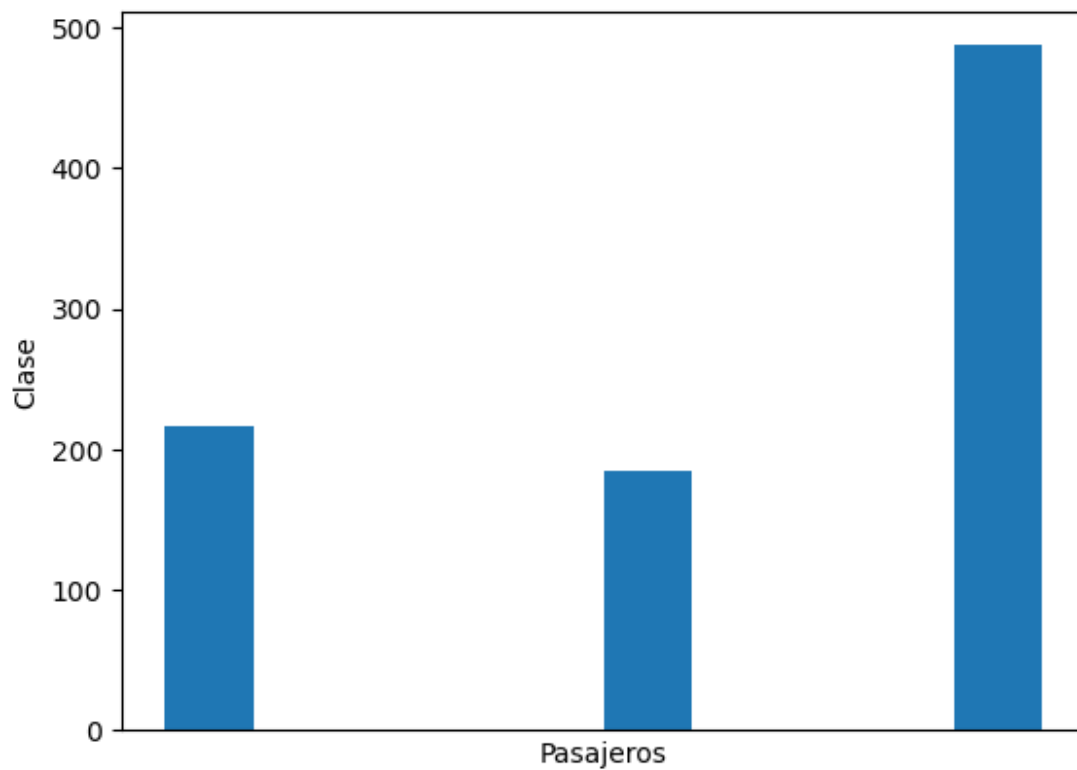
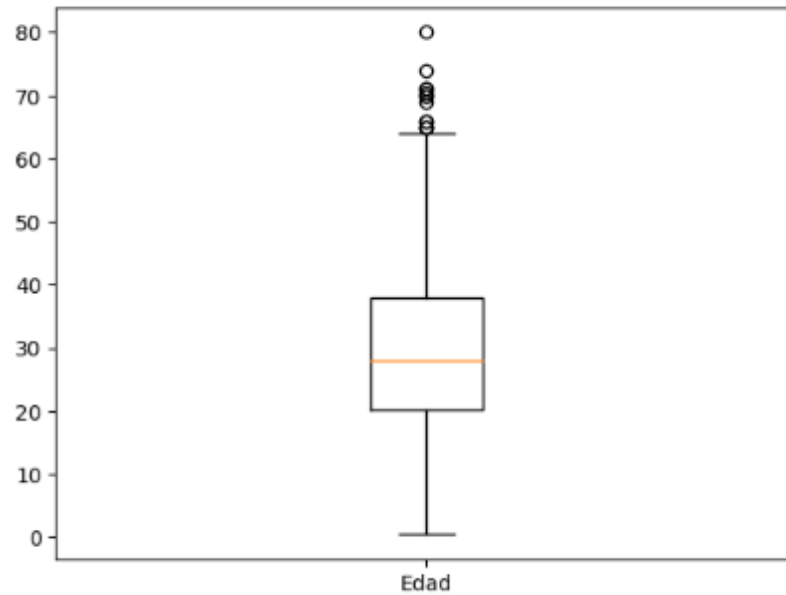
## a) El Histograma de frecuencias para la variable edad

```
In [29]: import matplotlib.pyplot as plt
fig, ax = plt.subplots()
ax.hist(df['Age'], color = 'red')
ax.set_xlabel('Edad')
ax.legend(['Edad'])
fig.savefig('edad.jpg')
plt.show()
```



## b) El diagrama de caja para la variable edad

```
In [33]: import matplotlib.pyplot as plt
fig, ax = plt.subplots()
ax.boxplot(df['Age'])
ax.set_xlabel('Edad')
fig.savefig('DcajaEdad.jpg')
plt.show()
```



## Interprete sus resultados en cada caso y guarde dichas imágenes en formato .jpg

En el histograma de Edad, nos muestra que existe una mayor cantidad de pasajeros entre 20 años de edad y unos cuantos pasajeros entre 70 y 79 años

El diagrama de Caja nos muestra una mayor concentracion en el rango de 20 a 35 años de los pasajeros

Para el segundo grafico use un diagrama de dispersion lo que muestra son 3 tipos de clase(1,2,3) teniendo la mayor cantidad de pasajeros la 3ra clase, le sigue 1ra clase y por ultimo 2da clase

```
In [31]: # Guardamos el resultado en formato jpg
fig.savefig('clasepass.jpg')
```