

Actividad Módulo 9: Visualización de datos con Matplotlib

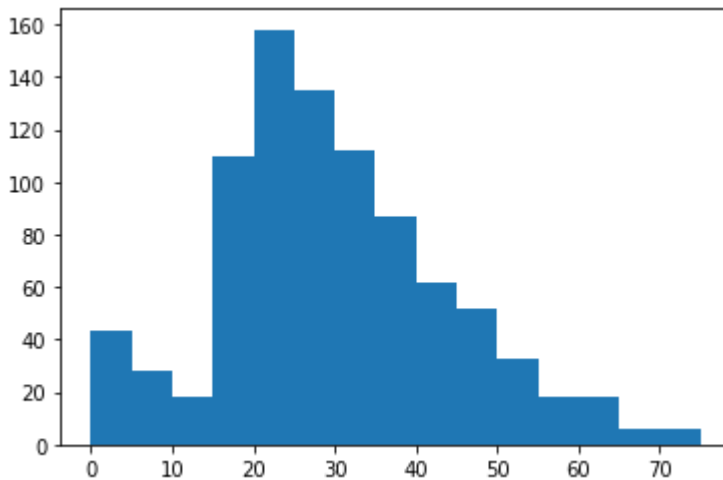
A partir de la columna Age del dataset del Titanic, obtenga lo siguiente mediante Matplotlib:

1. Histograma de frecuencias
2. Diagrama de Caja
3. Un gráfico que muestre el número de pasajeros de cada clase

Interprete sus resultados y guarde las imágenes en formato .jpg

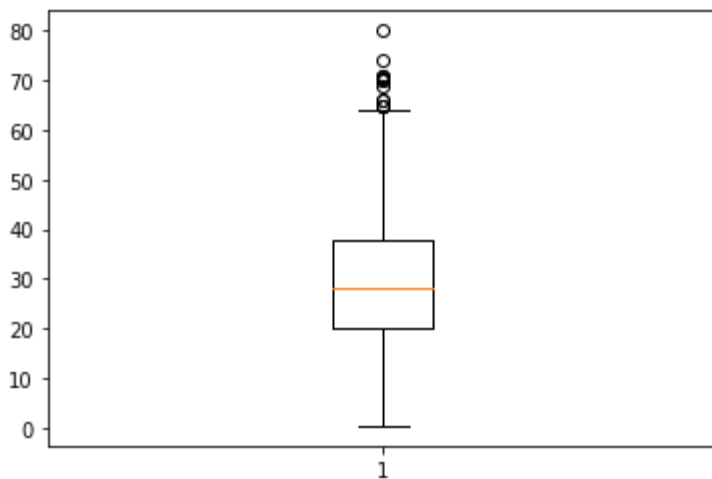
```
In [ ]: import pandas as pd
import matplotlib.pyplot as plt
df = pd.read_csv('titanic.csv')

# Histograma de frecuencias
fig, ax = plt.subplots()
ax.hist(df['Age'], bins=range(0,80,5))
plt.savefig('Ex_AgeHistogram.jpg')
plt.show()
```



La mayoría de los pasajeros eran adultos, entre 15 y 40 años de edad. Había pocos niños y adultos mayores de 60 años

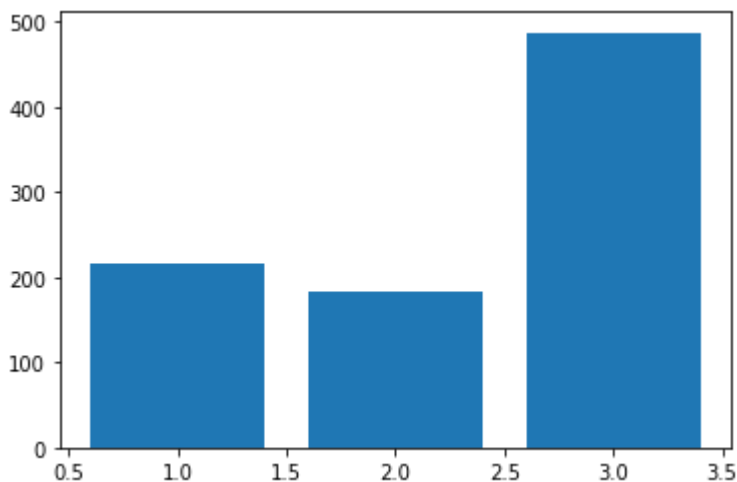
```
In [ ]: fig, ax = plt.subplots()
ax.boxplot(df['Age'])
plt.savefig('Ex_AgeBoxplot.jpg')
plt.show()
```



Confirmando lo que se vió en el histograma, la mayoría de los pasajeros tenían una edad de entre 20 y 40 años. Los valores atípicos son de personas mayores únicamente, a partir de los 62 años de edad aproximadamente. La persona más grande que iba en el barco tenía aproximadamente 80 años.

```
In [ ]: names = df['Pclass'].value_counts().keys()
        values = df['Pclass'].value_counts().values

fig, ax = plt.subplots()
ax.bar(names, values)
plt.savefig('Ex_PassengerClass.jpg')
plt.show()
```



La gran mayoría de los pasajeros eran de clase tres y, sorprendentemente había más pasajeros en primera que en segunda clase.