Q1. Download the Titanic dataset and perform the Exploratory data analysis using pandas. Read the dataset (df= pd.read_csv(r'..............................\Titanic.csv')

> Display the first and last 10 instances from the dataset

> Acquire the necessary information using the df.info() and df. Describe().

> Retrieve the number of columns and rows. (using shape)

```python
In [ ]:  import pandas as pd
         import numpy as np


         df = pd.read_csv("Titanic.csv")

         # First 10 instances
         print(df.head(10).to_string())
         print()

         # necessary information
         print()
         print("Information")
         print(df.info())
         print(df.describe())

         # Number of columns and rows
         print()
         rows, cols = df.shape
         print(f"Rows : {rows}, Columns : {cols}")
```

```
     PassengerId  Survived  Pclass
Name       Sex    Age  SibSp  Parch                 Ticket      Fare Cabin Embarked
0              1         0       3                                        Braund, Mr. Owen Ha
rris     male  22.0      1       0          A/5 21171    7.2500   NaN        S
1              2         1       1  Cumings, Mrs. John Bradley (Florence Briggs Tha
yer)   female  38.0      1       0          PC 17599   71.2833   C85        C
2              3         1       3                                        Heikkinen, Miss. L
aina   female  26.0      0       0  STON/O2. 3101282    7.9250   NaN        S
3              4         1       1           Futrelle, Mrs. Jacques Heath (Lily May P
eel)   female  35.0      1       0            113803   53.1000  C123        S
4              5         0       3                                        Allen, Mr. William H
enry     male  35.0      0       0            373450    8.0500   NaN        S
5              6         0       3                                             Moran, Mr. J
ames     male   NaN      0       0            330877    8.4583   NaN        Q
6              7         0       1                                        McCarthy, Mr. Timot
hy J     male  54.0      0       0             17463   51.8625   E46        S
7              8         0       3                                   Palsson, Master. Gosta Leo
nard     male   2.0      3       1            349909   21.0750   NaN        S
8              9         1       3  Johnson, Mrs. Oscar W (Elisabeth Vilhelmina B
erg)   female  27.0      0       2            347742   11.1333   NaN        S
9             10         1       2                             Nasser, Mrs. Nicholas (Adele Ac
hem)   female  14.0      1       0            237736   30.0708   NaN        C


Information
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 891 entries, 0 to 890
Data columns (total 12 columns):
 #   Column       Non-Null Count  Dtype
---  ------       --------------  -----
 0   PassengerId  891 non-null    int64
 1   Survived     891 non-null    int64
 2   Pclass       891 non-null    int64
 3   Name         891 non-null    object
 4   Sex          891 non-null    object
 5   Age          714 non-null    float64
 6   SibSp        891 non-null    int64
 7   Parch        891 non-null    int64
 8   Ticket       891 non-null    object
 9   Fare         891 non-null    float64
 10  Cabin        204 non-null    object
 11  Embarked     889 non-null    object
dtypes: float64(2), int64(5), object(5)
memory usage: 83.7+ KB
None
       PassengerId    Survived      Pclass         Age       SibSp  \
count   891.000000  891.000000  891.000000  714.000000  891.000000
mean    446.000000    0.383838    2.308642   29.699118    0.523008
std     257.353842    0.486592    0.836071   14.526497    1.102743
min       1.000000    0.000000    1.000000    0.420000    0.000000
25%     223.500000    0.000000    2.000000   20.125000    0.000000
50%     446.000000    0.000000    3.000000   28.000000    0.000000
75%     668.500000    1.000000    3.000000   38.000000    1.000000
max     891.000000    1.000000    3.000000   80.000000    8.000000

             Parch        Fare
count   891.000000  891.000000
mean      0.381594   32.204208
std       0.806057   49.693429
min       0.000000    0.000000
```

```
25%        0.000000     7.910400
50%        0.000000    14.454200
75%        0.000000    31.000000
max        6.000000   512.329200
```

Rows : 891, Columns : 12

Q2. Create the data visualization using the matplotlib.

> Visualize the Gender of Passengers using the Bar graph

> Visualize the Survival Count of Passengers using the Bar graph.

> Visualize the Age of Passengers using the Bar/Histogram graph.

> Visualize the comparison of Age and Fare of Passengers using the Scatterplot.

In [ ]:
```python
import matplotlib.pyplot as plt

gender_counts = df['Sex'].value_counts()
plt.bar(gender_counts.index, gender_counts.values)
plt.title('Gender of Passengers')
plt.xlabel('Gender')
plt.ylabel('Count')
plt.show()
```
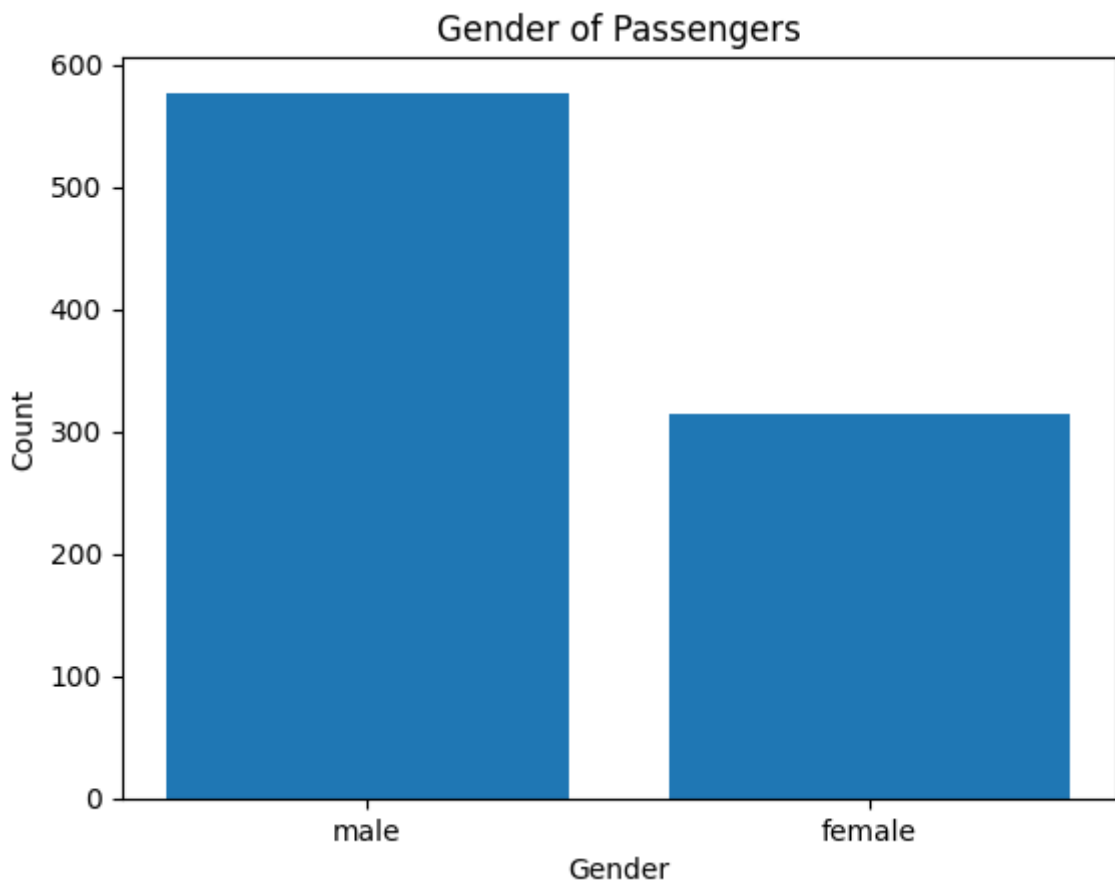


In [ ]:
```python
fig = plt.figure(figsize= (5,4))
survival_counts = df['Survived'].value_counts()
```

```python
x = list(survival_counts.index)
y = list(survival_counts)
# fig, ax = plt.subplots()
# ax.plot(x, y)
plt.bar(x, y)
# plt.title('Survival Count of Passengers')
# plt.xlabel('Survival')
# plt.ylabel('Count')
# plt.show()
```

Out[ ]:   <BarContainer object of 2 artists>
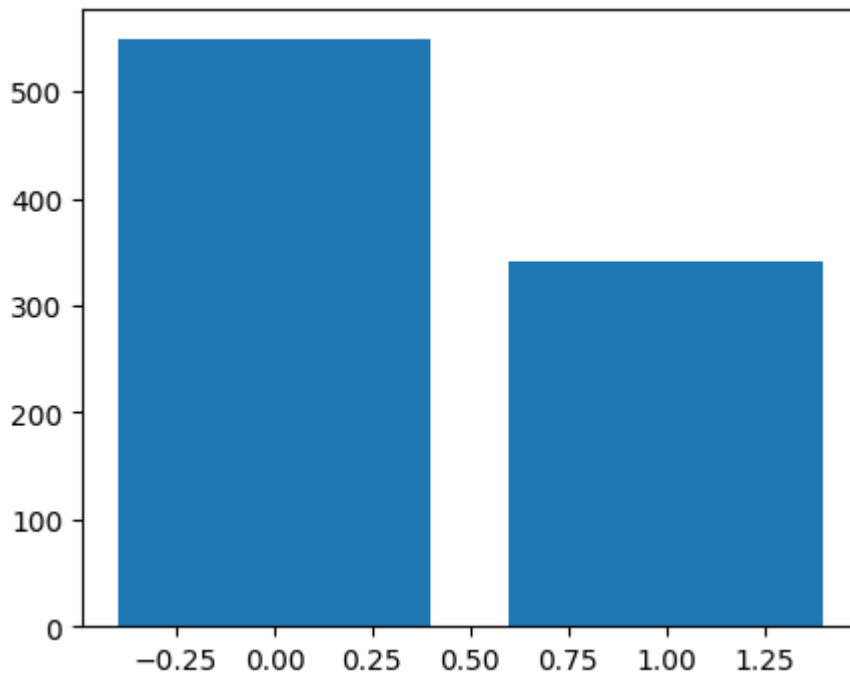

```python
plt.hist(df['Age'].dropna(), bins=20)
plt.title('Age of Passengers')
plt.xlabel('Age')
plt.ylabel('Count')
plt.show()
```
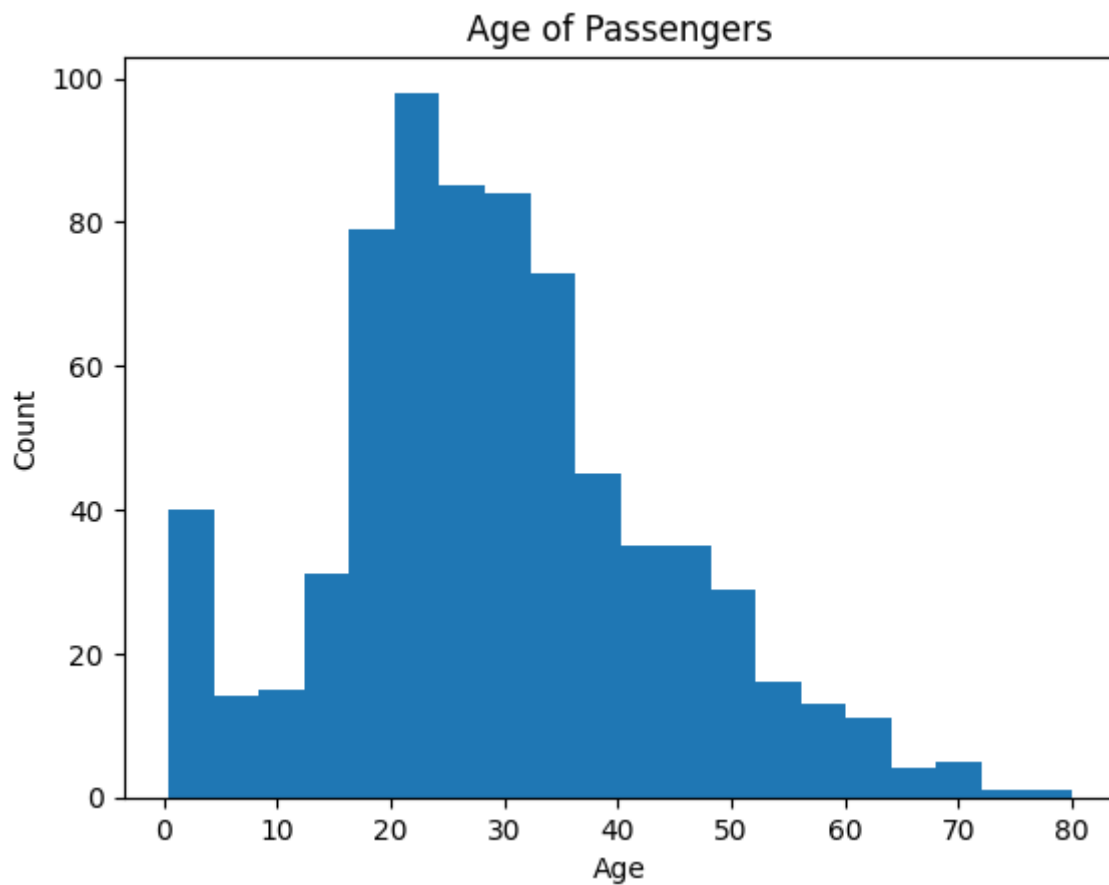
Age of Passengers


```
In [ ]:  plt.scatter(df['Age'], df['Fare'])
         plt.title('Comparison of Age and Fare of Passengers')
         plt.xlabel('Age')
         plt.ylabel('Fare')
         plt.show()
```

Comparison of Age and Fare of Passengers