```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

Reading CSV File

```
df=pd.read_csv('Titanic-Dataset.csv')
```

Titanic Dataset

```
df
```

```
     PassengerId  Survived  Pclass  \
0              1         0       3
1              2         1       1
2              3         1       3
3              4         1       1
4              5         0       3
..           ...       ...     ...
886          887         0       2
887          888         1       1
888          889         0       3
889          890         1       1
890          891         0       3

                                                  Name     Sex   Age
SibSp  \
0                              Braund, Mr. Owen Harris    male  22.0
1
1    Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0
1
2                               Heikkinen, Miss. Laina  female  26.0
0
3         Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0
1
4                             Allen, Mr. William Henry    male  35.0
0
..                                                 ...     ...   ...
...
886                              Montvila, Rev. Juozas    male  27.0
0
887                       Graham, Miss. Margaret Edith  female  19.0
0
888           Johnston, Miss. Catherine Helen "Carrie"  female   NaN
1
889                              Behr, Mr. Karl Howell    male  26.0
```

```
0
890                                 Dooley, Mr. Patrick    male  32.0
0

     Parch            Ticket     Fare Cabin Embarked
0        0         A/5 21171   7.2500   NaN        S
1        0          PC 17599  71.2833   C85        C
2        0  STON/O2. 3101282   7.9250   NaN        S
3        0            113803  53.1000  C123        S
4        0            373450   8.0500   NaN        S
..     ...               ...      ...   ...      ...
886      0            211536  13.0000   NaN        S
887      0            112053  30.0000   B42        S
888      2        W./C. 6607  23.4500   NaN        S
889      0            111369  30.0000  C148        C
890      0            370376   7.7500   NaN        Q

[891 rows x 12 columns]
```

Top 5 Dataset

```
df.head(5)
```

```
   PassengerId  Survived  Pclass  \
0            1         0       3
1            2         1       1
2            3         1       3
3            4         1       1
4            5         0       3

                                                 Name     Sex   Age
SibSp  \
0                             Braund, Mr. Owen Harris    male  22.0
1
1  Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0
1
2                              Heikkinen, Miss. Laina  female  26.0
0
3       Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0
1
4                            Allen, Mr. William Henry    male  35.0
0

   Parch            Ticket     Fare Cabin Embarked
0      0         A/5 21171   7.2500   NaN        S
1      0          PC 17599  71.2833   C85        C
2      0  STON/O2. 3101282   7.9250   NaN        S
```

```
3      0            113803  53.1000  C123        S
4      0            373450   8.0500   NaN        S
```

Bottom of 5 Dataset

```
df.tail(5)
```

```
     PassengerId  Survived  Pclass
Name  \
886          887         0       2                      Montvila, Rev.
Juozas
887          888         1       1               Graham, Miss. Margaret
Edith
888          889         0       3  Johnston, Miss. Catherine Helen
"Carrie"
889          890         1       1                      Behr, Mr. Karl
Howell
890          891         0       3                        Dooley, Mr.
Patrick

        Sex   Age  SibSp  Parch      Ticket   Fare Cabin Embarked
886    male  27.0      0      0      211536  13.00   NaN        S
887  female  19.0      0      0      112053  30.00   B42        S
888  female   NaN      1      2  W./C. 6607  23.45   NaN        S
889    male  26.0      0      0      111369  30.00  C148        C
890    male  32.0      0      0      370376   7.75   NaN        Q
```

Dataset Information

```
df.info()
```

```
<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
```

No of Rows and Columns in Dataset

```
print("Number of Rows ==> " ,df.shape[0])
print("Number of Columns==> ",df.shape[1])
```

```
Number of Rows ==>  891
Number of Columns==>  12
```

Describe Method

```
df.describe()
```

```
       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  
```

Checking NULL Values in Dataset

```
df.isnull().sum()
```

```
PassengerId      0
Survived         0
Pclass           0
Name             0
```

```
Sex              0
Age            177
SibSp            0
Parch            0
Ticket           0
Fare             0
Cabin          687
Embarked         2
dtype: int64
```

Male and Female Counts in Dataset

```
male_count = df[df['Sex'] == 'male'].shape[0]
female_count = df[df['Sex'] == 'female'].shape[0]

print(f"Number of Males Travelled in Titanic: {male_count}")
print(f"Number of Females Travelled in Titanic: {female_count}")
```

```
Number of Males Travelled in Titanic: 577
Number of Females Travelled in Titanic: 314
```

Finding Duplicate Values in Dataset

```
df.duplicated().value_counts()
```

```
False    891
Name: count, dtype: int64
```

Finding Name Start with 'S' Letter

```
df.columns
```

```
Index(['PassengerId', 'Survived', 'Pclass', 'Name', 'Sex', 'Age',
'SibSp',
       'Parch', 'Ticket', 'Fare', 'Cabin', 'Embarked'],
      dtype='object')
```

```
df[df['Name'].str.startswith("S")].value_counts()
```

```
PassengerId  Survived  Pclass  Name
Sex     Age   SibSp  Parch  Ticket    Fare      Cabin  Embarked
11           1         3      Sandstrom, Miss. Marguerite Rut
female  4.0   1      1      PP 9549   16.7000   G6     S           1
24           1         1      Sloper, Mr. William Thompson
male    28.0  0      0      113788    35.5000   A6     S           1
```

```
175          0         1       Smith, Mr. James Clinch
male    56.0  0      0      17764     30.6958   A7     C           1
206          0         3       Strom, Miss. Telma Matilda
female  2.0   0      1      347054    10.4625   G6     S           1
252          0         3       Strom, Mrs. Wilhelm (Elna Matilda
Persson)                female  29.0  1      1      347054    10.4625
G6     S           1
253          0         1       Stead, Mr. William Thomas
male    62.0  0      0      113514    26.5500   C87    S           1
320          1         1       Spedden, Mrs. Frederic Oakley
(Margaretta Corning Stone)  female  40.0  1      1      16966
134.5000  E34    C           1
395          1         3       Sandstrom, Mrs. Hjalmar (Agnes
Charlotta Bengtsson)       female  24.0  0      2      PP 9549
16.7000   G6     S           1
435          0         1       Silvey, Mr. William Baird
male    50.0  1      0      13507     55.9000   E44    S           1
578          1         1       Silvey, Mrs. William Baird (Alice
Munger)                 female  39.0  1      0      13507     55.9000
E44    S           1
592          1         1       Stephenson, Mrs. Walter Bertram (Martha
Eustis)            female  52.0  1      0      36947     78.2667   D20
C           1
610          1         1       Shutes, Miss. Elizabeth W
female  40.0  0      0      PC 17582  153.4625  C125   S           1
626          0         1       Sutton, Mr. Frederick
male    61.0  0      0      36963     32.3208   D50    S           1
633          1         1       Stahelin-Maeglin, Dr. Max
male    32.0  0      0      13214     30.5000   B50    C           1
642          1         1       Sagesser, Mlle. Emma
female  24.0  0      0      PC 17477  69.3000   B35    C           1
648          1         1       Simonius-Blumer, Col. Oberst Alfons
male    56.0  0      0      13213     35.5000   A26    C           1
702          1         1       Silverthorne, Mr. Spencer Victor
male    35.0  0      0      PC 17475  26.2875   E24    S           1
716          0         3       Soholt, Mr. Peter Andreas Lauritz
Andersen                male    19.0  0      0      348124    7.6500
F G73  S           1
863          1         1       Swift, Mrs. Frederick Joel (Margaret
Welles Barron)       female  48.0  0      0      17466     25.9292
D17    S           1
Name: count, dtype: int64
```

Drop the Column

```
df_dropped = df.drop(['SibSp'], axis=1)
print(df_dropped)
```

```
     PassengerId  Survived  Pclass  \
0              1         0       3
1              2         1       1
2              3         1       3
3              4         1       1
4              5         0       3
..           ...       ...     ...
886          887         0       2
887          888         1       1
888          889         0       3
889          890         1       1
890          891         0       3

                                                    Name     Sex   Age
Parch  \
0                                Braund, Mr. Owen Harris    male  22.0
0
1    Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0
0
2                                 Heikkinen, Miss. Laina  female  26.0
0
3          Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0
0
4                               Allen, Mr. William Henry    male  35.0
0
..                                                   ...     ...   ...
...
886                               Montvila, Rev. Juozas    male  27.0
0
887                        Graham, Miss. Margaret Edith  female  19.0
0
888            Johnston, Miss. Catherine Helen "Carrie"  female   NaN
2
889                               Behr, Mr. Karl Howell    male  26.0
0
890                                 Dooley, Mr. Patrick    male  32.0
0

                 Ticket     Fare Cabin Embarked
0             A/5 21171   7.2500   NaN        S
1              PC 17599  71.2833   C85        C
2    STON/O2. 3101282   7.9250   NaN        S
3              113803  53.1000  C123        S
4              373450   8.0500   NaN        S
..                ...      ...   ...      ...
886            211536  13.0000   NaN        S
887            112053  30.0000   B42        S
888        W./C. 6607  23.4500   NaN        S
889            111369  30.0000  C148        C
890            370376   7.7500   NaN        Q
```

```
[891 rows x 11 columns]
```

Column Rename

```
df_rname = df.rename(columns={'PassengerId':'Passenger Number'})
print(df_rname)
```

```
     Passenger Number  Survived  Pclass  \
0                   1         0       3
1                   2         1       1
2                   3         1       3
3                   4         1       1
4                   5         0       3
..                ...       ...     ...
886               887         0       2
887               888         1       1
888               889         0       3
889               890         1       1
890               891         0       3

                                                  Name     Sex   Age
SibSp  \
0                              Braund, Mr. Owen Harris    male  22.0
1
1    Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0
1
2                               Heikkinen, Miss. Laina  female  26.0
0
3         Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0
1
4                            Allen, Mr. William Henry    male  35.0
0
..                                                 ...     ...   ...
...
886                            Montvila, Rev. Juozas    male  27.0
0
887                     Graham, Miss. Margaret Edith  female  19.0
0
888           Johnston, Miss. Catherine Helen "Carrie"  female   NaN
1
889                              Behr, Mr. Karl Howell    male  26.0
0
890                                Dooley, Mr. Patrick    male  32.0
0

     Parch            Ticket     Fare Cabin Embarked
0        0         A/5 21171   7.2500   NaN        S
1        0          PC 17599  71.2833   C85        C
```

```
2         0  STON/O2. 3101282   7.9250   NaN        S
3         0            113803  53.1000  C123        S
4         0            373450   8.0500   NaN        S
..      ...               ...      ...   ...      ...
886       0            211536  13.0000   NaN        S
887       0            112053  30.0000   B42        S
888       2        W./C. 6607  23.4500   NaN        S
889       0            111369  30.0000  C148        C
890       0            370376   7.7500   NaN        Q

[891 rows x 12 columns]
```

Finding How many of Passenger age above or equal to 60 using Lambda function

```
df_age = df['Age'].apply(lambda x:x>=60).sum()
print("Passangers Age Above or Equal to 60 ==> ", df_age)
```

```
Passangers Age Above or Equal to 60 ==>  26
```

Replace the Values

```
df['Survived'] = df['Survived'].replace({0 :'Died',1 : 'Live'})
df['Survived']
```

```
0      Died
1      Live
2      Live
3      Live
4      Died
       ...
886    Died
887    Live
888    Died
889    Live
890    Died
Name: Survived, Length: 891, dtype: object
```

How Many of them Live in Titanic

```
df['Survived'].value_counts()
```

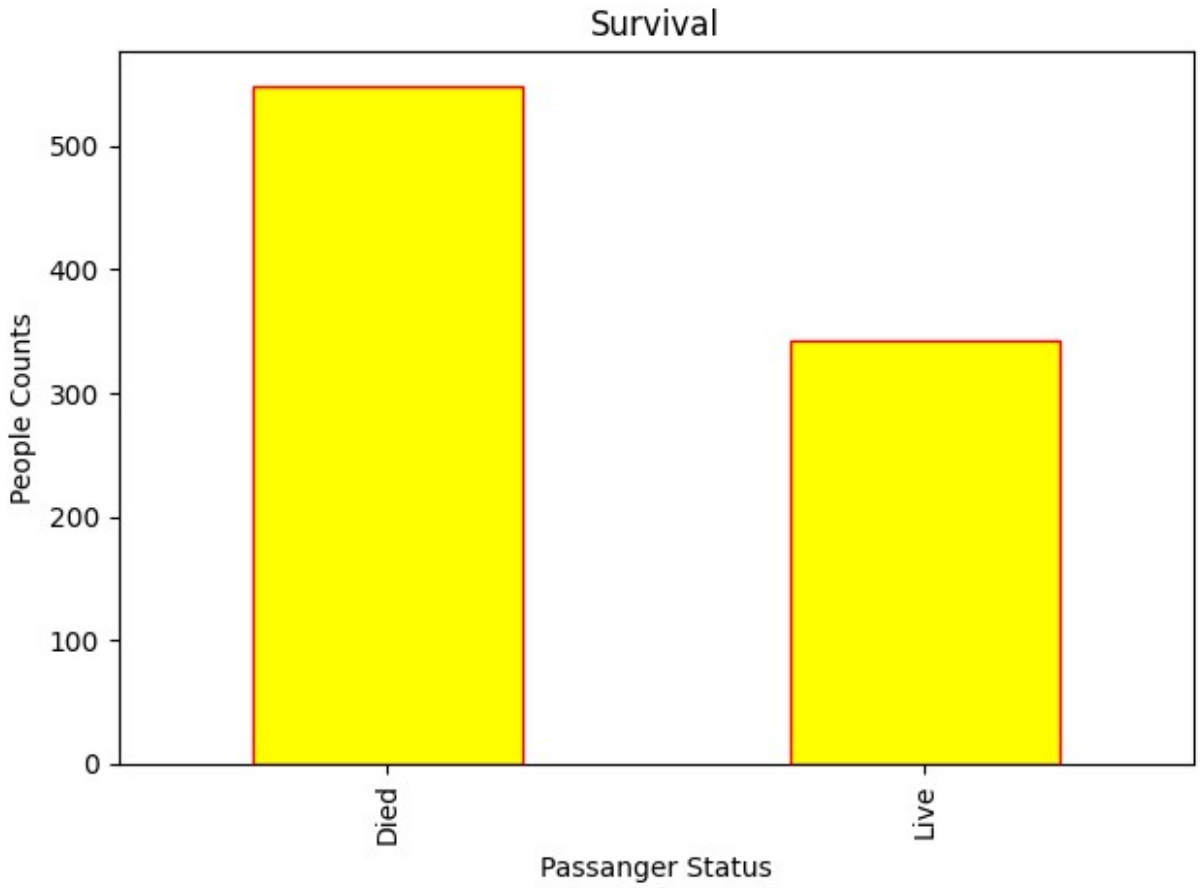
```
Survived
Died    549
```

```
Live    342
Name: count, dtype: int64
```

Pictorized Analysis

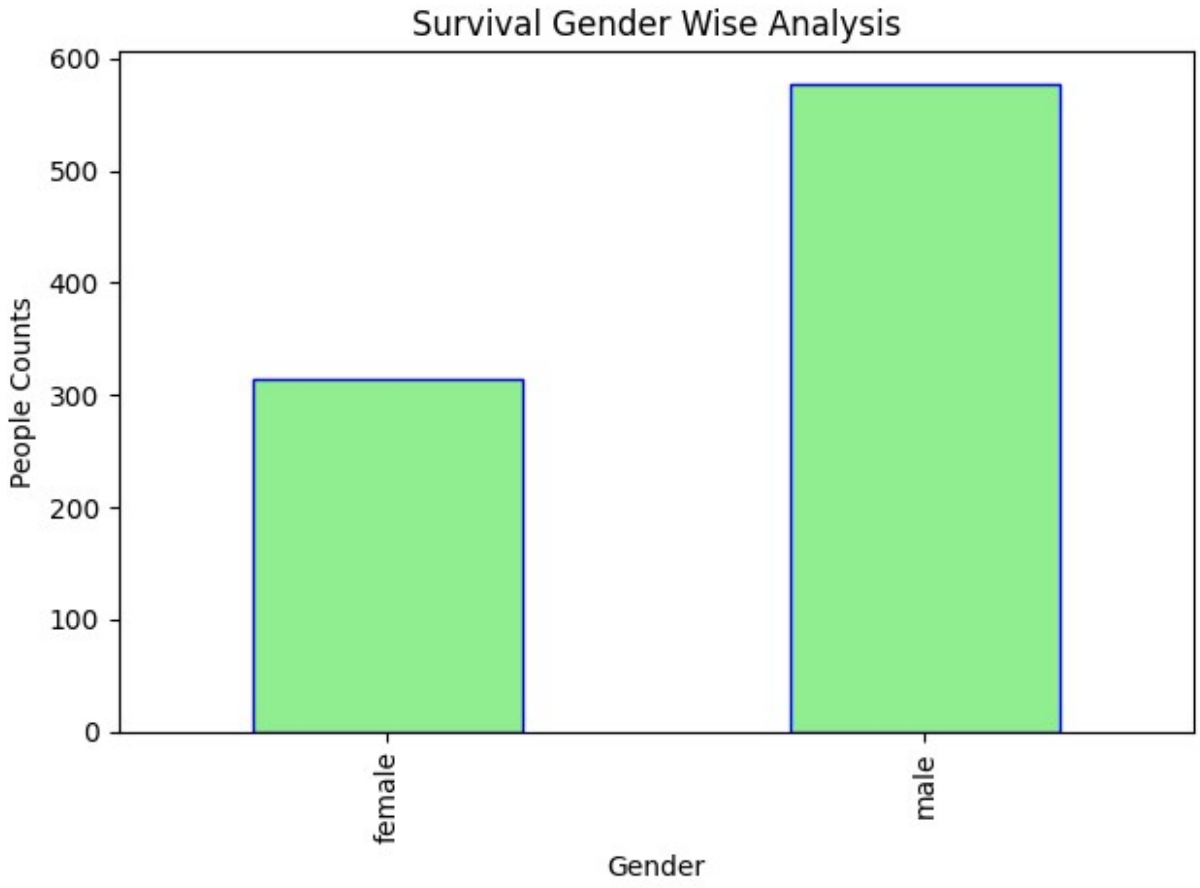
Bar Graph

```
fig = plt.figure()
tn=df.groupby('Survived').size()
tn.plot.bar(color="yellow", edgecolor="red")
plt.title("Survival")
plt.xlabel("Passanger Status")
plt.ylabel("People Counts")
plt.tight_layout()
plt.show()
```




```
fig = plt.figure()
tn=df.groupby('Sex').size()
```

```
tn.plot.bar(color="lightgreen", edgecolor="Blue")
plt.title("Survival Gender Wise Analysis")
plt.xlabel("Gender")
plt.ylabel("People Counts")
plt.tight_layout()
plt.show()
```



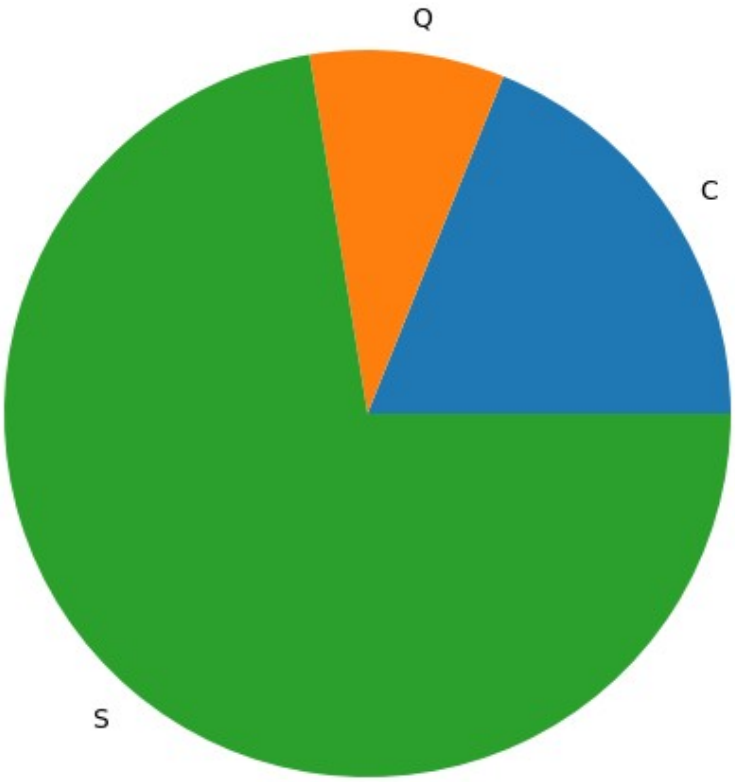

```
df
```

	PassengerId	Survived	Pclass	\
0	1	Died	3	
1	2	Live	1	
2	3	Live	3	
3	4	Live	1	
4	5	Died	3	
..	...	...	...	
886	887	Died	2	
887	888	Live	1	
888	889	Died	3	
889	890	Live	1	
890	891	Died	3	

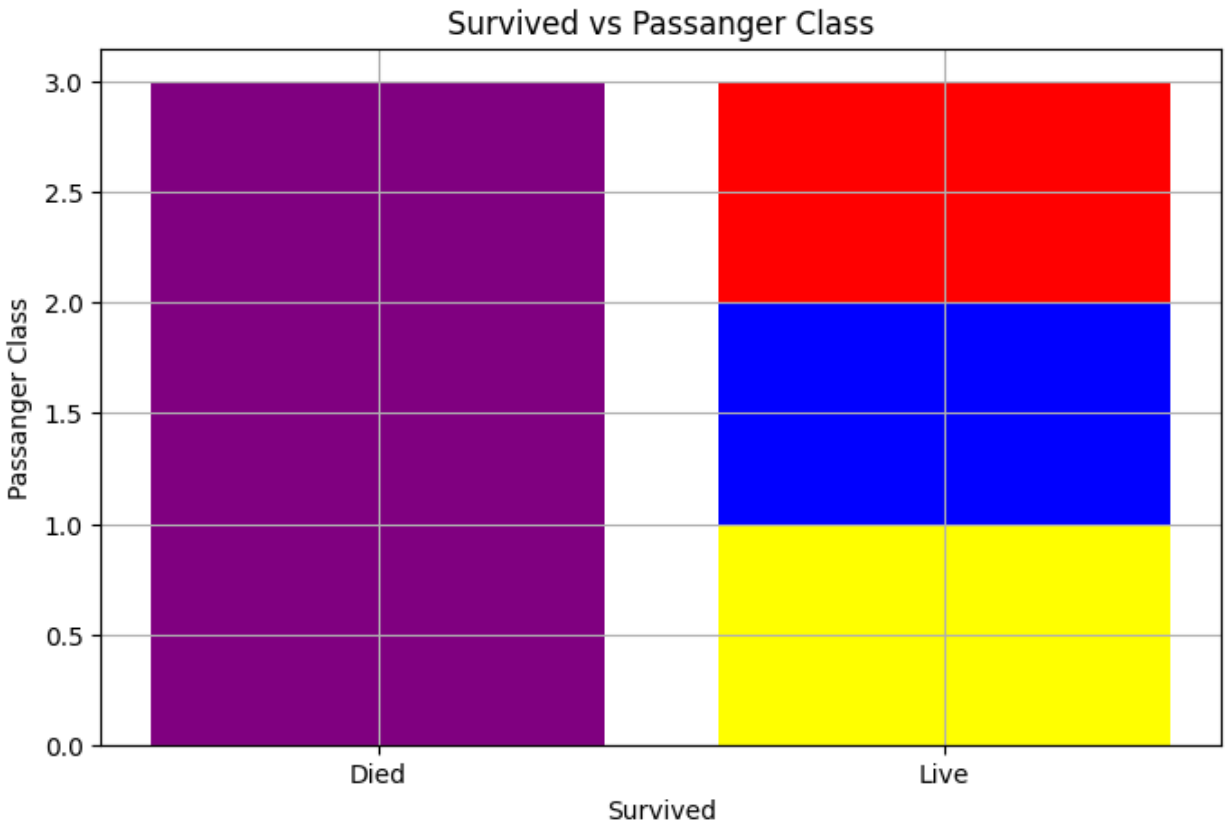
```
                                                     Name     Sex   Age
SibSp  \
0                                Braund, Mr. Owen Harris    male  22.0
1
1    Cumings, Mrs. John Bradley (Florence Briggs Th...  female  38.0
1
2                                 Heikkinen, Miss. Laina  female  26.0
0
3           Futrelle, Mrs. Jacques Heath (Lily May Peel)  female  35.0
1
4                               Allen, Mr. William Henry    male  35.0
0
..                                                   ...     ...   ...
...
886                                Montvila, Rev. Juozas    male  27.0
0
887                         Graham, Miss. Margaret Edith  female  19.0
0
888             Johnston, Miss. Catherine Helen "Carrie"  female   NaN
1
889                                Behr, Mr. Karl Howell    male  26.0
0
890                                  Dooley, Mr. Patrick    male  32.0
0

     Parch            Ticket     Fare Cabin Embarked
0        0         A/5 21171   7.2500   NaN        S
1        0          PC 17599  71.2833   C85        C
2        0  STON/O2. 3101282   7.9250   NaN        S
3        0            113803  53.1000  C123        S
4        0            373450   8.0500   NaN        S
..     ...               ...      ...   ...      ...
886      0            211536  13.0000   NaN        S
887      0            112053  30.0000   B42        S
888      2        W./C. 6607  23.4500   NaN        S
889      0            111369  30.0000  C148        C
890      0            370376   7.7500   NaN        Q

[891 rows x 12 columns]
```

Pie Chart Based on Embarked

```
tn=df.groupby('Embarked').size()
fig=plt.figure()
ax=fig.add_axes([0,0,1,1])
ax.axis("equal")
ax.pie(tn,labels=tn.index)
plt.show()
```




```
plt.figure(figsize=(8, 5))
plt.bar(df['Survived'], df['Pclass'],
color=['Blue','yellow','purple','red'])
plt.title('Survived vs Passanger Class')
plt.xlabel('Survived')
plt.ylabel('Passanger Class')
plt.grid(True)
plt.savefig('bar_chart.png')
```


Survived vs Passanger Class
Passanger Class
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Died
Live
Survived