Matplotlib

What is matplotlib?

Matplotlib is a low level graph plotting library in python that serves as a visualization utility.

Matplotlib was created by John D. Hunter.

Installation of Matplotlib

!pip install matplotlib

In [39]:
```
import matplotlib.pyplot as plt
import pandas as pd
df=pd.read_csv('C:/Users/user/Downloads/kaggle/titanic (1).csv')
df
```

Out[39]:

	survived	pclass	sex	age	sibsp	parch	fare	embarked	class
0	0.0	3.0	male	22.0	1	0	7.2500	S	Third
1	1.0	1.0	female	38.0	1	0	71.2833	C	First
2	1.0	3.0	female	26.0	0	0	7.9250	S	Third
3	1.0	1.0	female	35.0	1	0	53.1000	S	First
4	0.0	3.0	male	35.0	0	0	8.0500	S	Third
...	...	...	...	...	...	...	...	...	...
886	0.0	870.0	male	27.0	0	0	13.0000	S	Second
887	1.0	871.0	female	19.0	0	0	30.0000	S	First
888	0.0	872.0	female	NaN	1	2	23.4500	S	Third
889	1.0	873.0	male	26.0	0	0	30.0000	C	First
890	0.0	874.0	male	32.0	0	0	7.7500	Q	Third

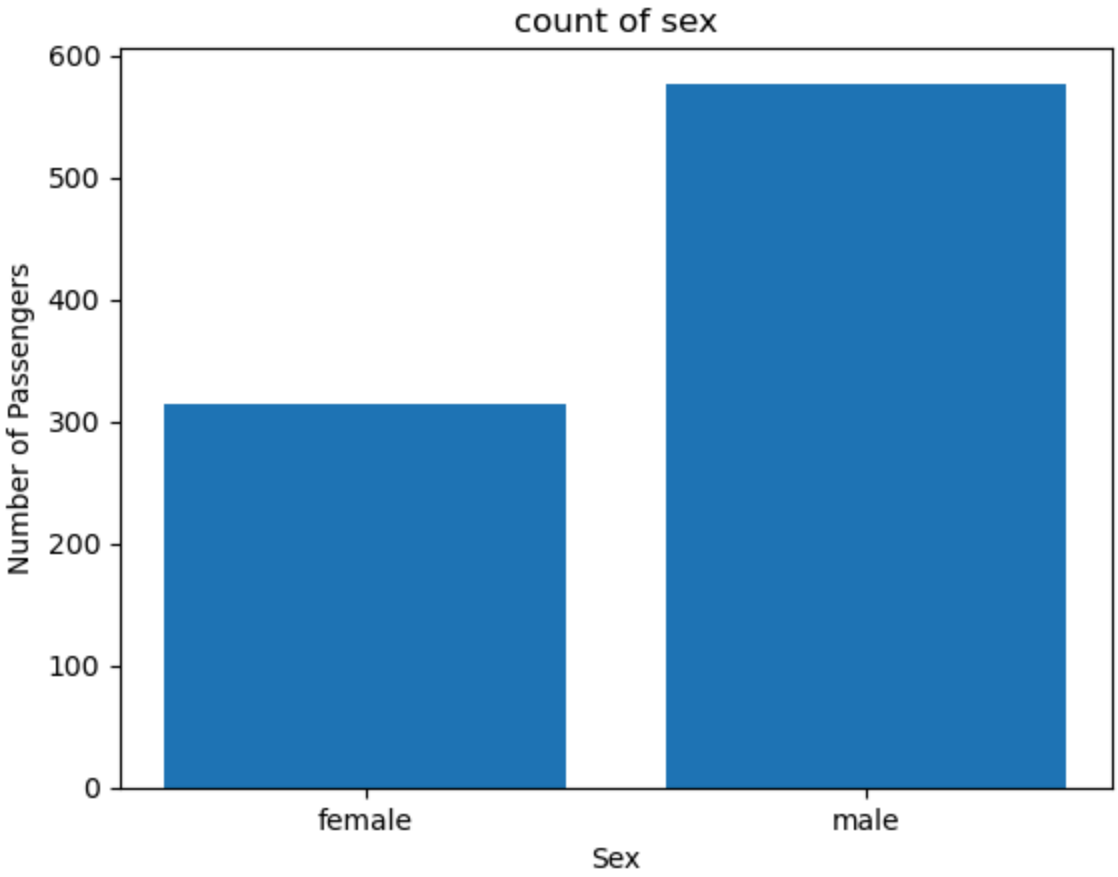
891 rows × 15 columns

Examples:

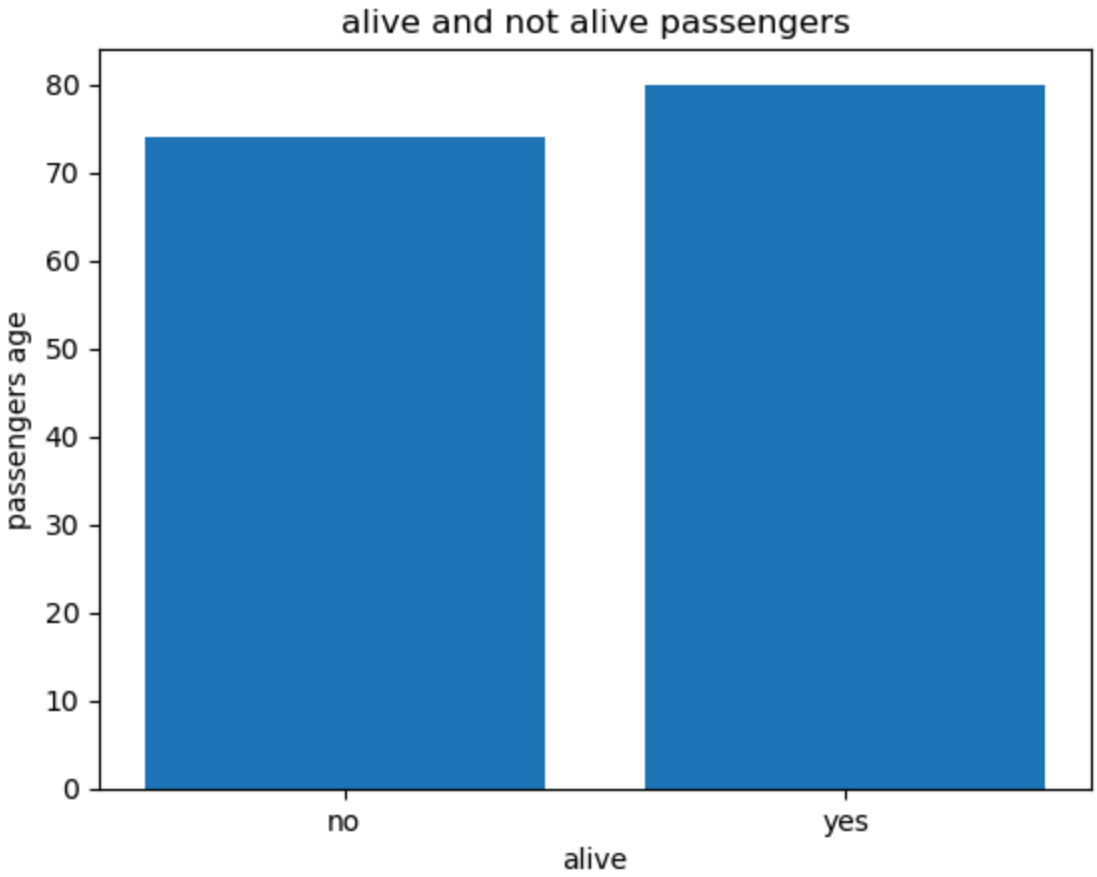
1.Bar graph

A bar graph is a visual representation of data using rectangular bars to show the values of different categories. The length or height of each bar corresponds to the data value it represents.

In [40]:
```
a= df['sex'].value_counts().sort_index()
plt.bar(a.index, a.values)
plt.xlabel('Sex')
plt.ylabel('Number of Passengers')
plt.title('count of sex')
plt.show()
```



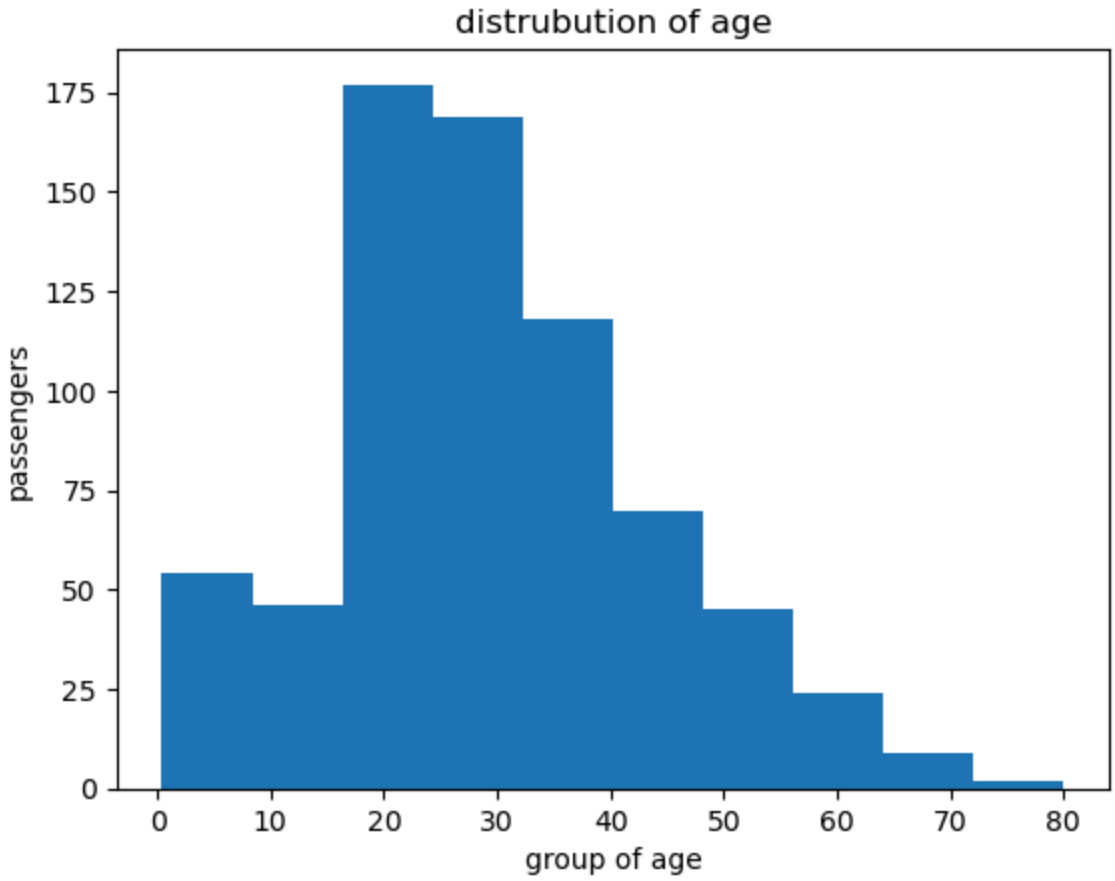
In [41]:
```
plt.bar(df['alive'],df['age'])
plt.title("alive and not alive passengers")
plt.xlabel("alive")
plt.ylabel("passengers age")
plt.show()
```



2.Histogram

A histogram is a graphical representation that shows the distribution of a numerical dataset using bars. Each bar represents the frequency of data points within a specific range (bin).

In [42]:
```
plt.hist(df['age'],bins=10)
plt.title("distrubution of age")
plt.xlabel("group of age")
plt.ylabel("passengers")
plt.show()
```



3.line graph

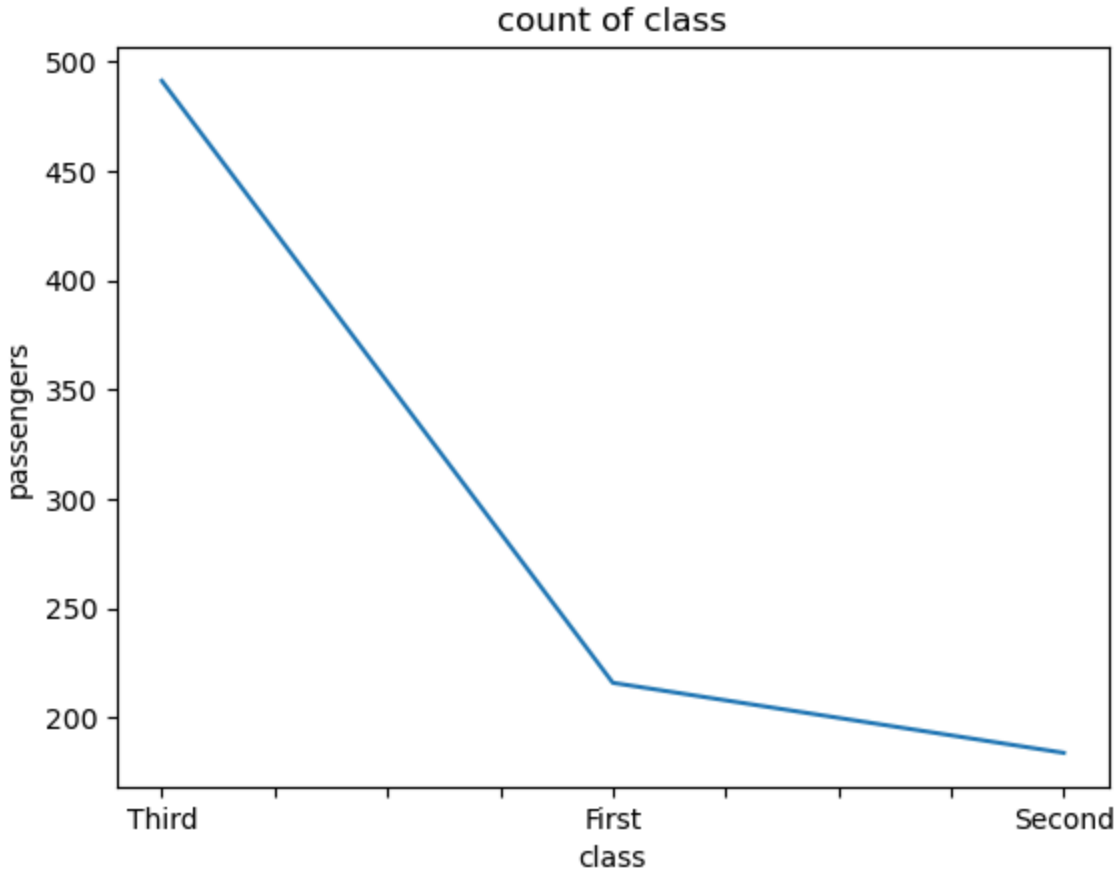
A line graph is a chart that connects data points with a continuous line to show trends over time or ordered categories. It’s useful for visualizing changes and patterns in data.

In [43]:
```
a=df['class'].value_counts()
a
```

Out[43]:
```
class
Third     491
First     216
Second    184
Name: count, dtype: int64
```

In [44]:
```
a=df['class'].value_counts()
plt.title("count of class")
plt.xlabel("class")
plt.ylabel("passengers")
a.plot(kind='line')
```

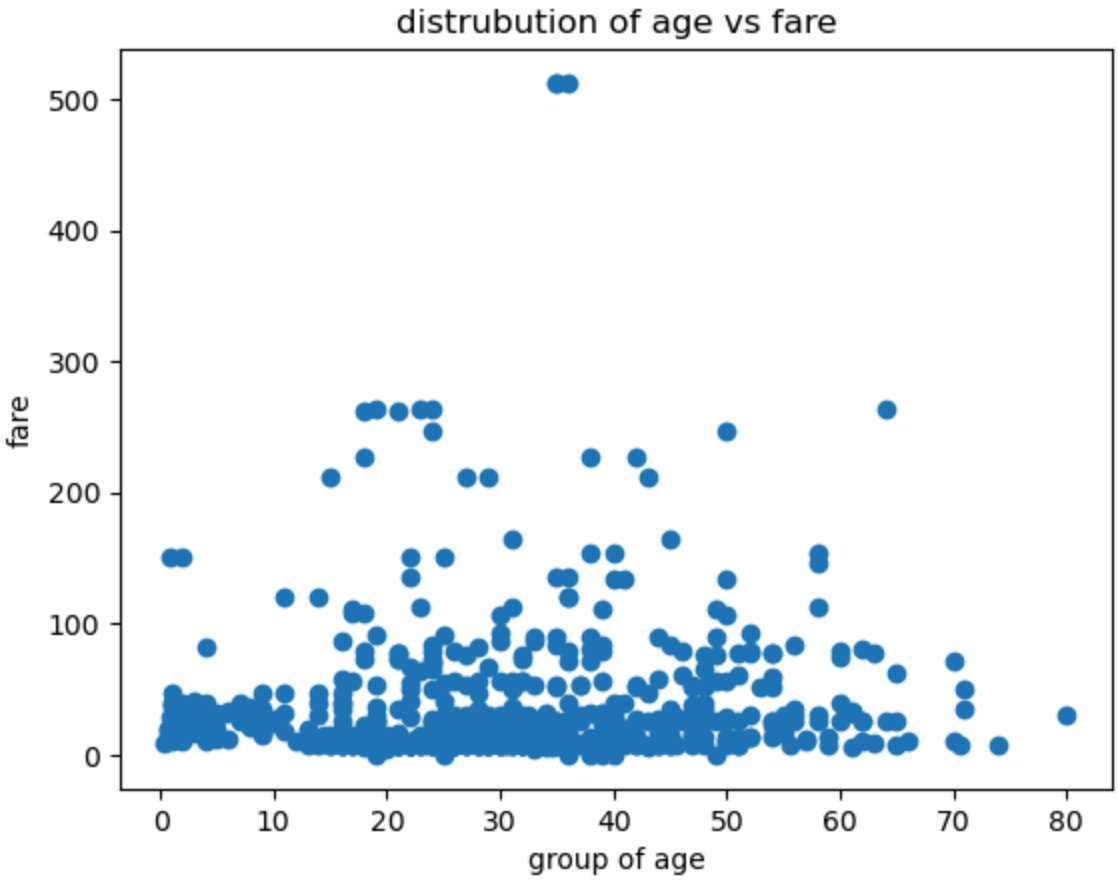
Out[44]:
```
<Axes: title={'center': 'count of class'}, xlabel='class', ylabel='passenge
rs'>
```



4.Scatter plot

A scatter plot is a graph that displays points to show the relationship between two different variables. Each point represents an observation with values on the x- and y-axes.

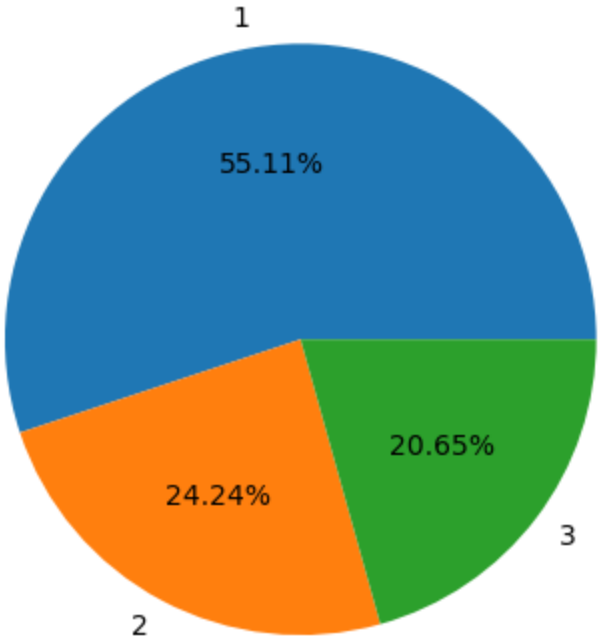
In [45]:
```
plt.scatter(df['age'],df['fare'])
plt.title("distrubution of age vs fare")
plt.xlabel("group of age")
plt.ylabel("fare")
plt.show()
```



5.Piechart

A pie chart is a circular graph divided into slices to illustrate numerical proportions. Each slice represents a category's contribution to the whole.

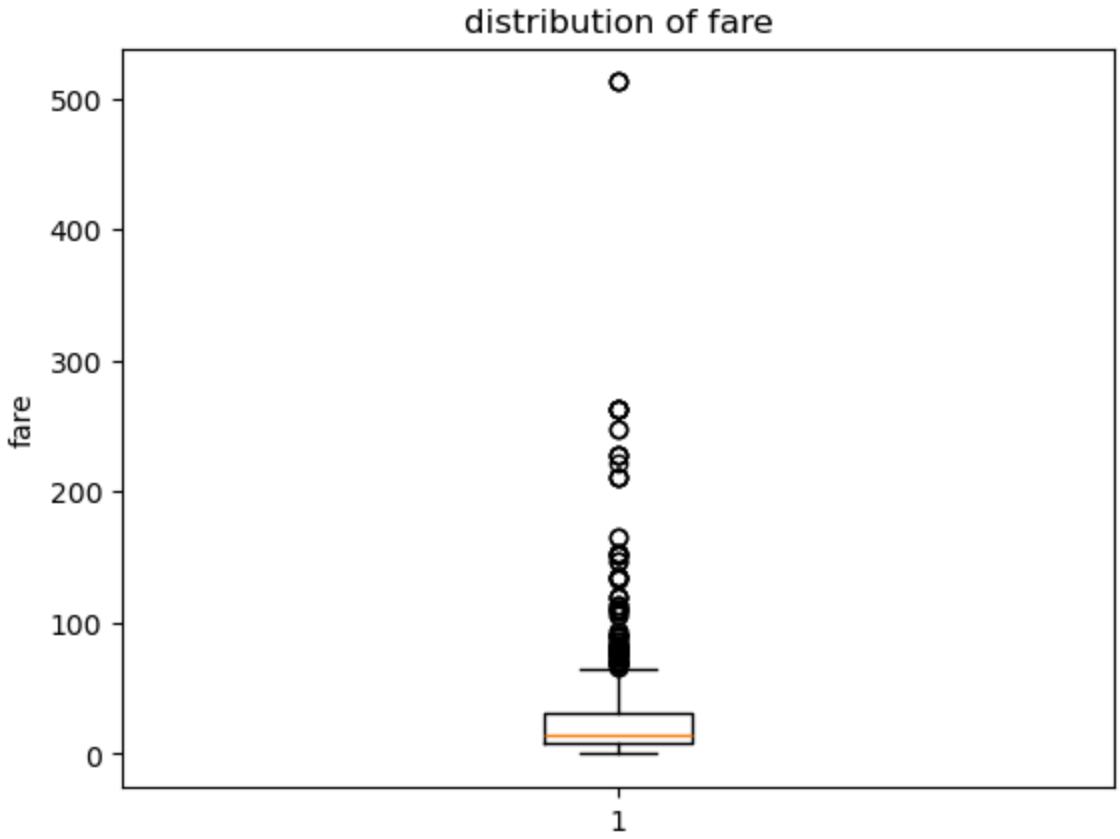
In [46]:
```
plt.pie(df['class'].value_counts(),labels=[1,2,3],autopct='%1.2f%%')
plt.show()
```



6.Boxplot

A boxplot is a graphical display that shows the distribution of a dataset through its quartiles. It highlights the median, spread, and potential outliers in the data.

In [47]:
```
plt.boxplot(df['fare'])
plt.title("distribution of fare")
plt.ylabel("fare")
plt.show()
```

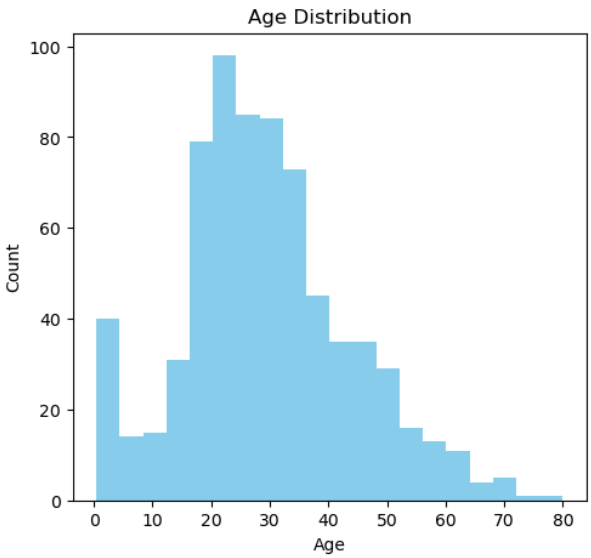


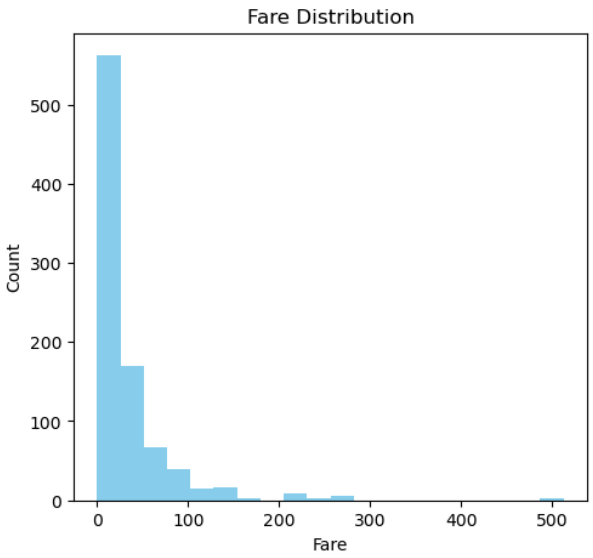
7.Subplots

A subplot is a secondary storyline in a narrative that supports or contrasts the main plot. It adds depth to characters or themes and often intersects with the main story.

In [48]:
```
fig, axs = plt.subplots(1, 2, figsize=(12, 5))
axs[0].hist(df['age'].dropna(),bins=20, color='skyblue')
axs[0].set_title('Age Distribution')
axs[0].set_xlabel('Age')
axs[0].set_ylabel('Count')

axs[1].hist(df['fare'].dropna(),bins=20, color='skyblue')
axs[1].set_title('Fare Distribution')
axs[1].set_xlabel('Fare')
axs[1].set_ylabel('Count')
```

Out[48]:
```
Text(0, 0.5, 'Count')
```





In [49]:
```
df.value_counts()
```

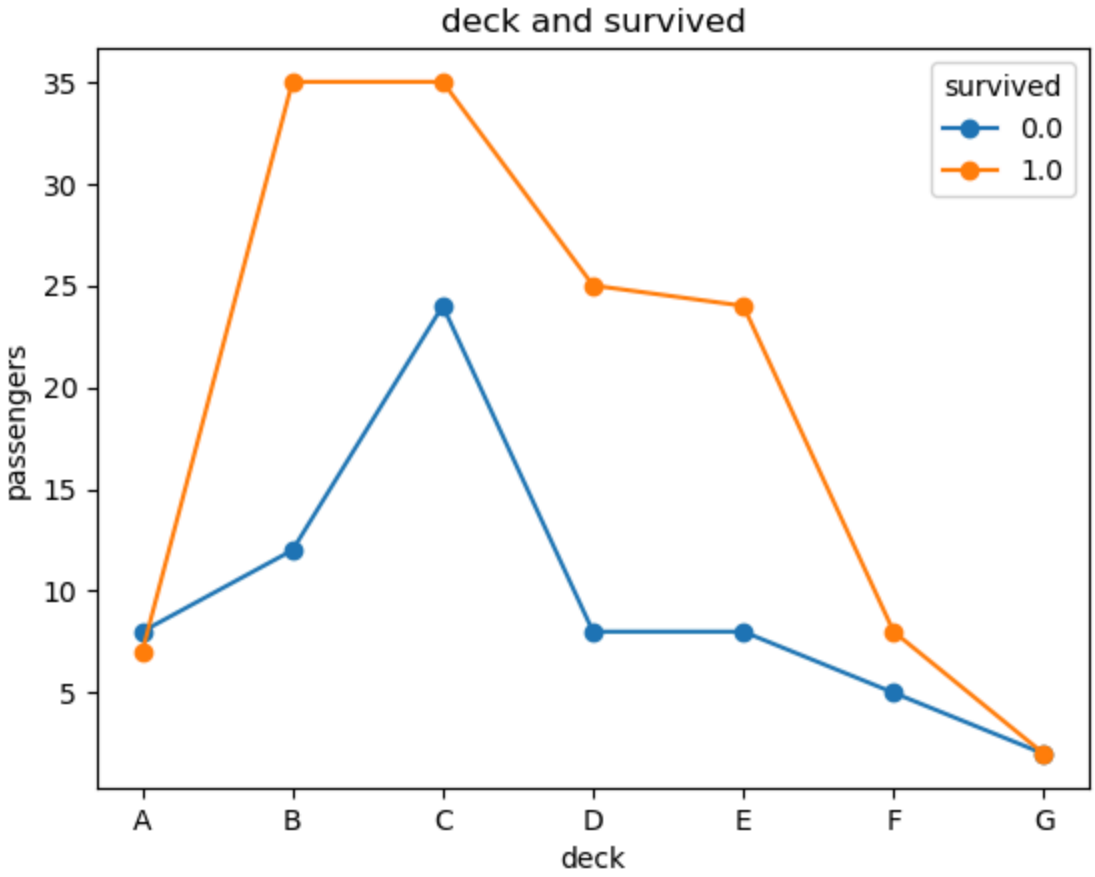
Out[49]:
```
survived  pclass  sex     age   sibsp  parch  fare     embarked  class   wh
o    adult_male  deck  embark_town  alive  alone
0.0       1.0     male    54.0  0      0      51.8625  S         First   ma
n    True        E     Southampton  no     True     1
1.0       433.0   male    52.0  0      0      30.5000  S         First   ma
n    True        C     Southampton  yes    True     1
          444.0   male    48.0  0      0      26.5500  S         First   ma
n    True        E     Southampton  yes    True     1
          457.0   female  23.0  0      0      13.7917  C         Second  wo
man  False       D     Cherbourg    yes    True     1
          468.0   male    25.0  1      0      91.0792  C         First   ma
n    True        B     Cherbourg    yes    False    1

..
          3.0     female  4.0   1      1      16.7000  S         Third   ch
ild  False       G     Southampton  yes    False    1
          5.0     male    34.0  0      0      13.0000  S         Second  ma
n    True        D     Southampton  yes    True     1
          7.0     male    28.0  0      0      35.5000  S         First   ma
n    True        A     Southampton  yes    True     1
          36.0    female  49.0  1      0      76.7292  C         First   wo
man  False       D     Cherbourg    yes    False    1
          873.0   male    26.0  0      0      30.0000  C         First   ma
n    True        C     Cherbourg    yes    True     1
Name: count, Length: 182, dtype: int64
```

8.line graph

A line graph is a chart that displays information as a series of data points connected by straight lines.

It is commonly used to show trends or changes over time or continuous data.

In [55]:
```
a=df[['deck','survived']].value_counts().unstack()
a.plot(kind='line',marker='o')
plt.title("deck and survived")
plt.xlabel("deck")
plt.ylabel("passengers")
plt.show()
```



In [ ]:
```
```