NAME:SAMYAK SONAWANE

763

202201090097

G3

Code: import pandas as pd

import matplotlib.pyplot as plt

from sklearn.cluster import KMeans

Load the dataset

data = pd.read_csv('Titanic.csv')

1

male_survived_third_class = data[(data['Sex'] == 'male') & (data['Survived'] == 1) & (data['Pclass'] == 3)]

count_male_survived_third_class = len(male_survived_third_class)

print("Count of male passengers who survived and preferred 'third class' ticket:", count_male_survived_third_class)

2)

gender_class_count = data.groupby(['Sex', 'Pclass']).size().reset_index(name='Count')

survived_count = data.groupby('Survived').size().reset_index(name='Count')

print("Count of passengers gender-wise and class-wise:")

print(gender_class_count)

print("Count of survived and non-survived passengers:")

print(survived_count)

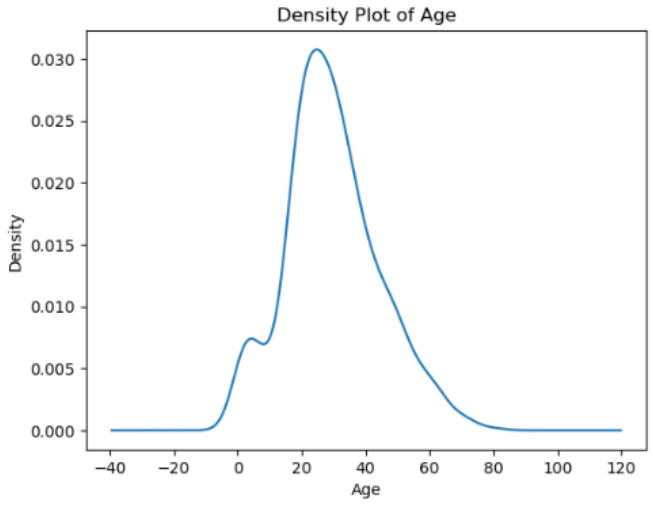
#

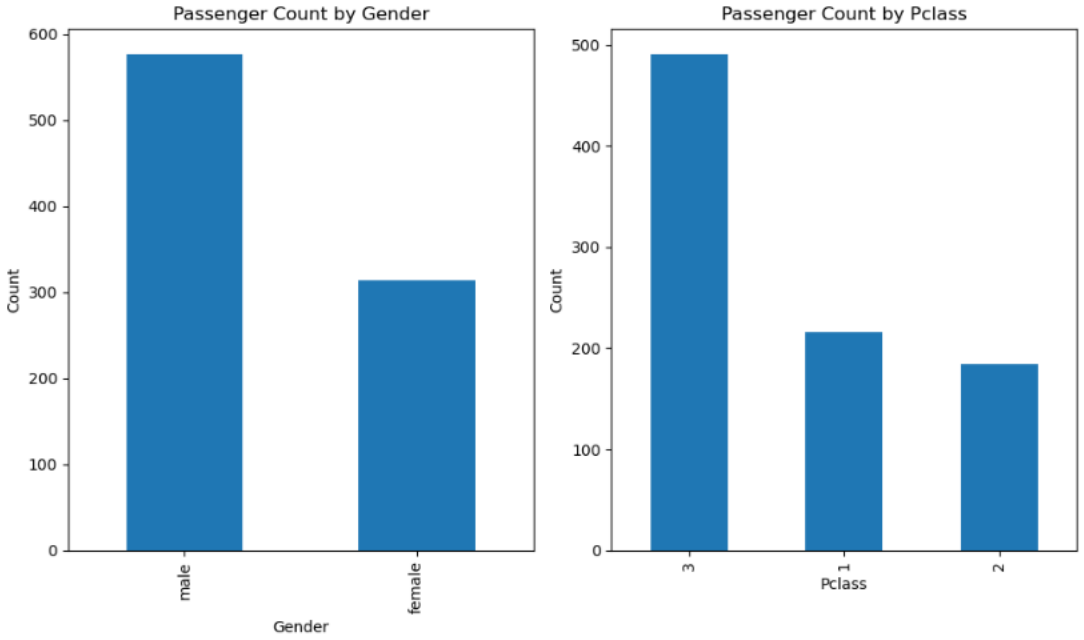
```
data['Age'].plot(kind='density')
plt.xlabel('Age')
plt.title('Density Plot of Age')
plt.show()

plt.figure(figsize=(10, 6))

plt.subplot(1, 2, 1)
data['Sex'].value_counts().plot(kind='bar')
plt.xlabel('Gender')
plt.ylabel('Count')
plt.title('Passenger Count by Gender')

plt.subplot(1, 2, 2)
data['Pclass'].value_counts().plot(kind='bar')
plt.xlabel('Pclass')
plt.ylabel('Count')
plt.title('Passenger Count by Pclass')

plt.tight_layout()
plt.show()

import pandas as pd
from sklearn.cluster import KMeans
import matplotlib.pyplot as plt

# Create a sample dataset
data = {'Age': [22, 38, 26, 35, 35, 28, 54, 2, 27, 14, 30, 42, 35, 58, 20],
        'Pclass': [3, 1, 3, 1, 3, 3, 1, 3, 2, 3, 1, 1, 3, 3, 2]}
```

```
df = pd.DataFrame(data)
kmeans = KMeans(n_clusters=3)  # Define the number of clusters
kmeans.fit(df)
labels = kmeans.labels_
df['Cluster'] = labels
plt.scatter(df['Age'], df['Pclass'], c=df['Cluster'], cmap='viridis')
plt.xlabel('Age')
plt.ylabel('Pclass')
plt.title('K-means Clustering of Age and Pclass')
plt.show()
```

Output:

```
Count of male passengers who survived and preferred 'third class' ticket: 47
Count of passengers gender-wise and class-wise:
      Sex  Pclass  Count
0  female       1     94
1  female       2     76
2  female       3    144
3    male       1    122
4    male       2    108
5    male       3    347
Count of survived and non-survived passengers:
   Survived  Count
0         0    549
1         1    342
```






```
C:\Users\samya\anaconda3\lib\site-packages\sklearn\cluster\_kmeans.py:870: FutureWarning: The default value of `n_init` will ch
ange from 10 to 'auto' in 1.4. Set the value of `n_init` explicitly to suppress the warning
  warnings.warn(
C:\Users\samya\anaconda3\lib\site-packages\sklearn\cluster\_kmeans.py:1382: UserWarning: KMeans is known to have a memory leak
on Windows with MKL, when there are less chunks than available threads. You can avoid it by setting the environment variable OM
P_NUM_THREADS=1.
  warnings.warn(
```