import pandas as pd  
import numpy as np  
import matplotlib.pyplot as plt  
import seaborn as sns  
  
# upload CSV file  
sales = pd.read\_csv(r"C:\Users\ak240\Desktop\Data Analytics Course\Salesstore.csv")  
print(sales.shape)  
print(sales.head(10))  
  
# visualisation  
  
plot\_1 = sns.histplot(data=sales, x='Ship\_Mode')  
plt.show()  
  
plot\_2 = sns.histplot(data=sales, x='Customer\_Segment')  
plt.show()  
  
plot\_3 = sns.histplot(data=sales, x='Product\_Category')  
plt.show()  
  
plot\_4 = sns.histplot(data=sales, x='Product\_Container')  
plt.show()  
  
plot\_5 = sns.barplot(data=sales, x='Region', y='Profit', hue='Ship\_Mode')  
plt.xticks(rotation=45)  
plt.show()  
  
plot\_6 = sns.barplot(data=sales, x='Region', y='Sales', hue='Ship\_Mode')  
plt.xticks(rotation=45)  
plt.show()  
  
plot\_7 = sns.barplot(data=sales, x='Region', y='Profit', hue='Customer\_Segment')  
plt.xticks(rotation=45)  
plt.show()  
  
plot\_8 = sns.barplot(data=sales, x='Region', y='Profit', hue='Product\_Category')  
plt.xticks(rotation=45)  
plt.show()  
  
# groups most common  
  
from itertools import combinations  
from collections import Counter  
  
count = Counter()  
for row in sales['Grouped']:  
 row\_list = row.split(',')  
 count.update(Counter(combinations(row\_list, 2)))  
  
for key, value in count.most\_common():  
 print(key, value)  
# group by  
print(sales.groupby(['Region']).count())  
  
# duplicates  
  
print(sales[sales['Order\_ID'].duplicated(keep=False)])  
sales['Grouped'] = sales.groupby('Order\_ID')['Product\_Category'].transform(lambda x: ','.join(x))  
sales = sales[['Order\_ID', 'Grouped']].drop\_duplicates()  
print(sales)  
  
# filtering  
print(sales['Ship\_Mode'] == 'Express Air')  
filt = (sales['Ship\_Mode'] == 'Express Air')  
print(sales[filt])  
  
print(sales['Ship\_Mode'] == 'Regular Air')  
filt2 = (sales['Ship\_Mode'] == 'Regular Air')  
print(sales[filt2])  
  
print(sales.loc[sales['Customer\_Segment'] == 'Corporate'])  
print(sales.loc[(sales['Customer\_Segment'] == 'Corporate')  
 & (sales['Ship\_Mode'] == 'Express Air')  
 & (sales['Order\_Quantity'] >= 0)])