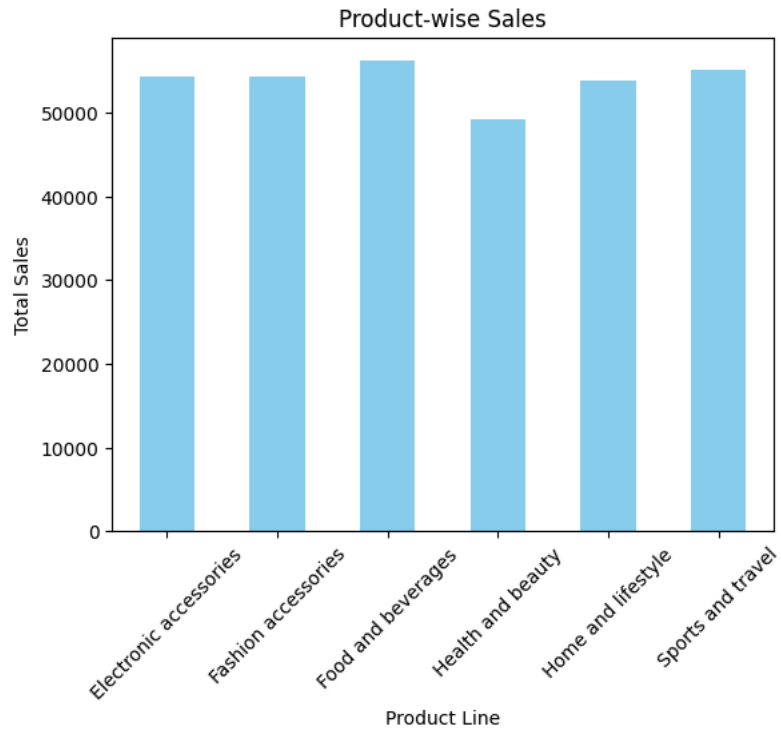
Exp 1

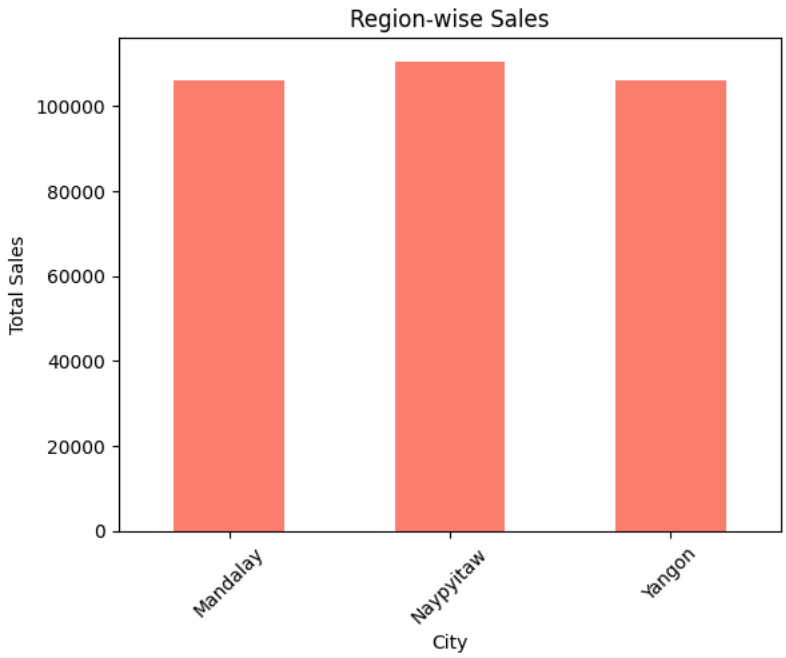
```
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt


df = pd.read_csv('supermarket_sales - Sheet1.csv')
df['Date'] = pd.to_datetime(df['Date'])
df.head()
```

	Invoice ID	Branch	City	Customer type	Gender	Product line	Unit price	Quantity	Tax 5%	Total	Date	Time	Payment	cogs	gross margin percentage	
0	750-67-8428	A	Yangon	Member	Female	Health and beauty	74.69	7	26.1415	548.9715	2019-01-05	13:08	Ewallet	522.83	4.761905	2
1	226-31-3081	C	Naypyitaw	Normal	Female	Electronic accessories	15.28	5	3.8200	80.2200	2019-03-08	10:29	Cash	76.40	4.761905	
2	631-41-3108	A	Yangon	Normal	Male	Home and lifestyle	46.33	7	16.2155	340.5255	2019-03-03	13:23	Credit card	324.31	4.761905	1
3	123-19-1176	A	Yangon	Member	Male	Health and beauty	58.22	8	23.2880	489.0480	2019-01-27	20:33	Ewallet	465.76	4.761905	2
4	373-73-	A	Yangon	Normal	Male	Sports and	86.31	7	30.2085	634.3785	2019-	10:37	Ewallet	604.17	4.761905	1

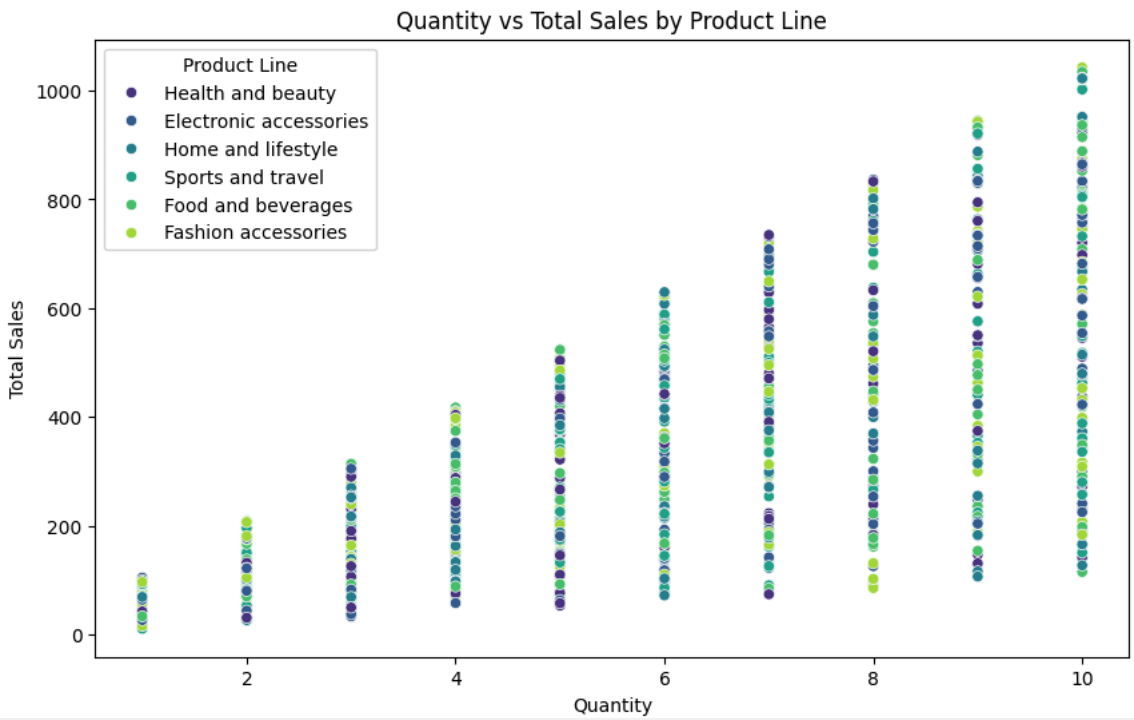
```
product_sales = df.groupby('Product line')['Total'].sum()
product_sales.plot(kind='bar', title='Product-wise Sales', ylabel='Total Sales', xlabel='Product Line', color='skyblue')
plt.xticks(rotation=45)
plt.show()
```



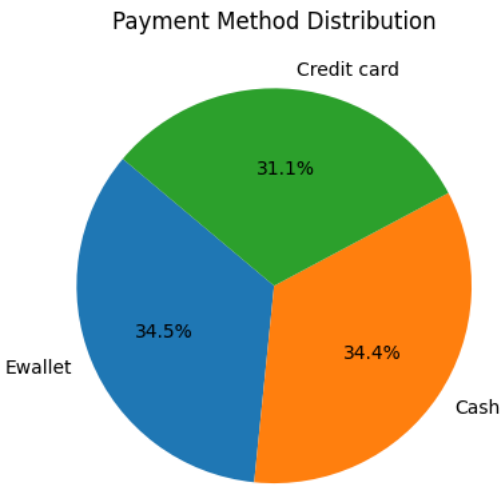

```
# Region-wise sales (City-wise sales)
region_sales = df.groupby('City')['Total'].sum()
region_sales.plot(kind='bar', title='Region-wise Sales', ylabel='Total Sales', xlabel='City', color='salmon')
plt.xticks(rotation=45)
plt.show()
```



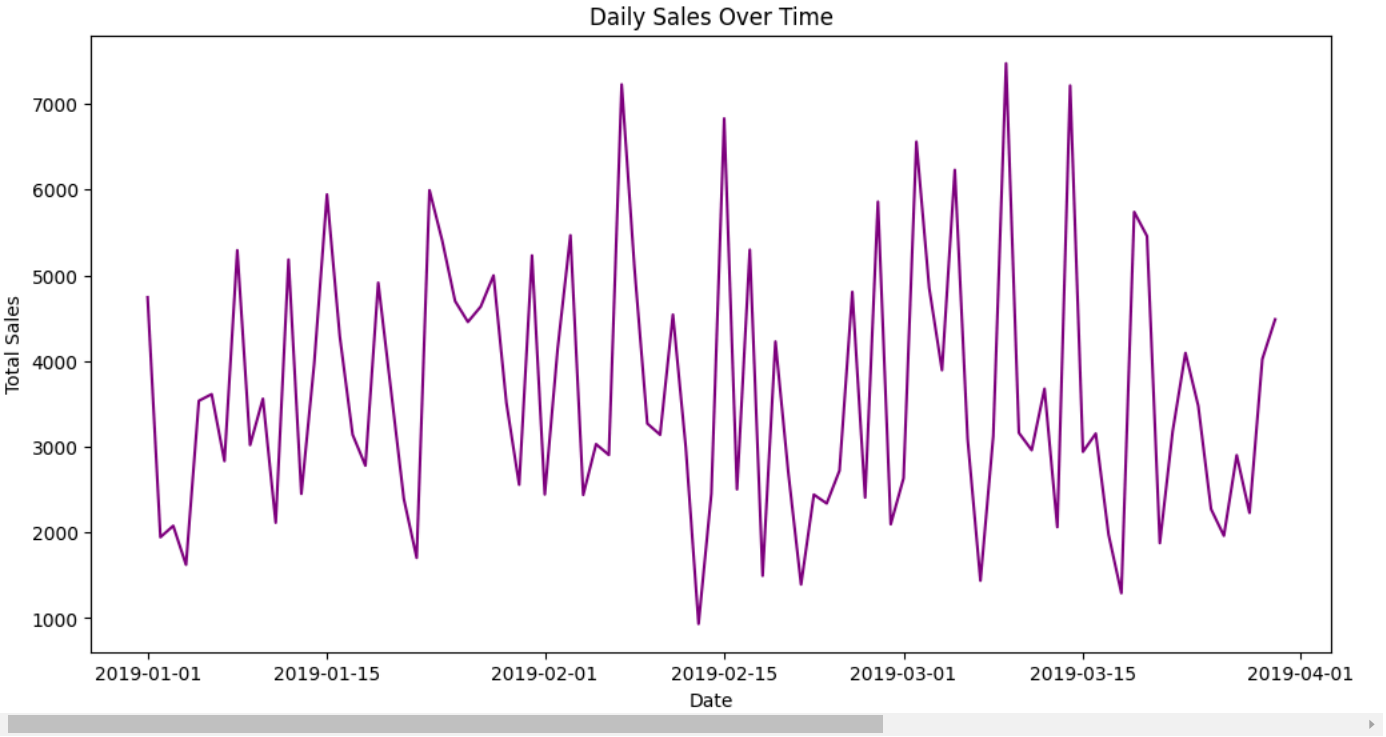

```
# Scatter plot for Quantity vs Total Sales
plt.figure(figsize=(10, 6))
sns.scatterplot(data=df, x='Quantity', y='Total', hue='Product line',
palette='viridis')
plt.title('Quantity vs Total Sales by Product Line')
plt.xlabel('Quantity')
plt.ylabel('Total Sales')
plt.legend(title='Product Line')
plt.show()
```



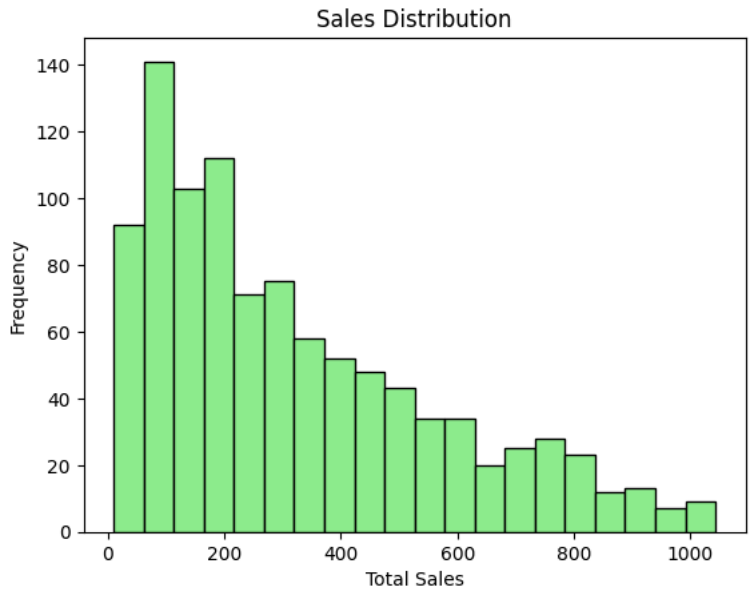

```
# Payment method distribution
payment_counts = df['Payment'].value_counts()
payment_counts.plot(kind='pie', autopct='%1.1f%%', startangle=140,
title='Payment Method Distribution')
plt.ylabel('')  # Removes the y-axis label for clarity
plt.show()
```




```
# Daily sales timeline
daily_sales = df.groupby(df['Date'].dt.date)['Total'].sum()
plt.figure(figsize=(12, 6))
daily_sales.plot(kind='line', title='Daily Sales Over Time', xlabel='Date',
ylabel='Total Sales', color='purple')
plt.show()
```




```
# Sales distribution
df['Total'].plot(kind='hist', bins=20, title='Sales Distribution',
color='lightgreen', edgecolor='black')
plt.xlabel('Total Sales')
plt.ylabel('Frequency')
plt.show()
```


Sales Distribution
Frequency
Total Sales
0
20
40
60
80
100
120
140
0
200
400
600
800
1000