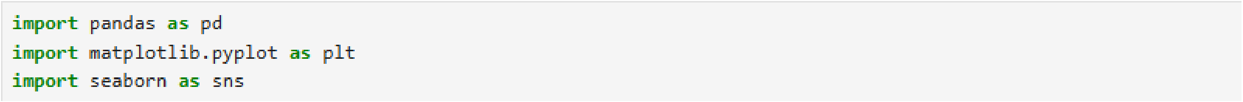
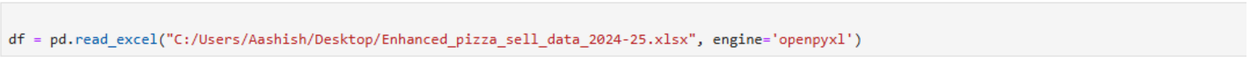
Pizza Data Analysis



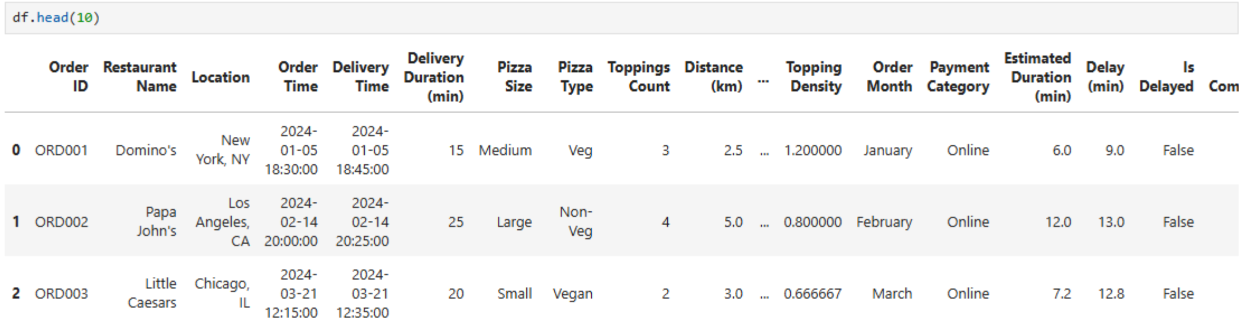
```
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

Load the Data Set ¶



```
df = pd.read_excel("C:/Users/Aashish/Desktop/Enhanced_pizza_sell_data_2024-25.xlsx", engine='openpyxl')
```

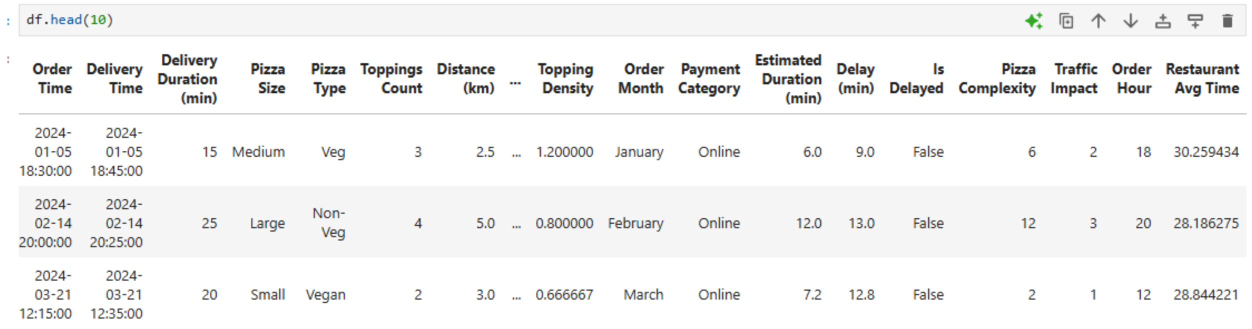
Sample Data



```
df.head(10)
```

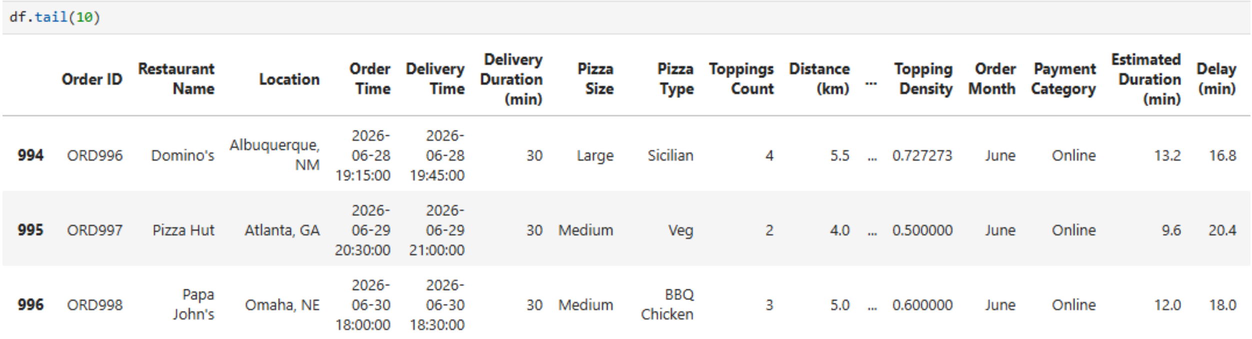
	Order ID	Restaurant Name	Location	Order Time	Delivery Time	Delivery Duration (min)	Pizza Size	Pizza Type	Toppings Count	Distance (km)	...	Topping Density	Order Month	Payment Category	Estimated Duration (min)	Delay (min)	Is Delayed	Com
0	ORD001	Domino's	New York, NY	2024-01-05 18:30:00	2024-01-05 18:45:00	15	Medium	Veg	3	2.5	...	1.200000	January	Online	6.0	9.0	False	
1	ORD002	Papa John's	Los Angeles, CA	2024-02-14 20:00:00	2024-02-14 20:25:00	25	Large	Non-Veg	4	5.0	...	0.800000	February	Online	12.0	13.0	False	
2	ORD003	Little Caesars	Chicago, IL	2024-03-21 12:15:00	2024-03-21 12:35:00	20	Small	Vegan	2	3.0	...	0.666667	March	Online	7.2	12.8	False	

Sample Data



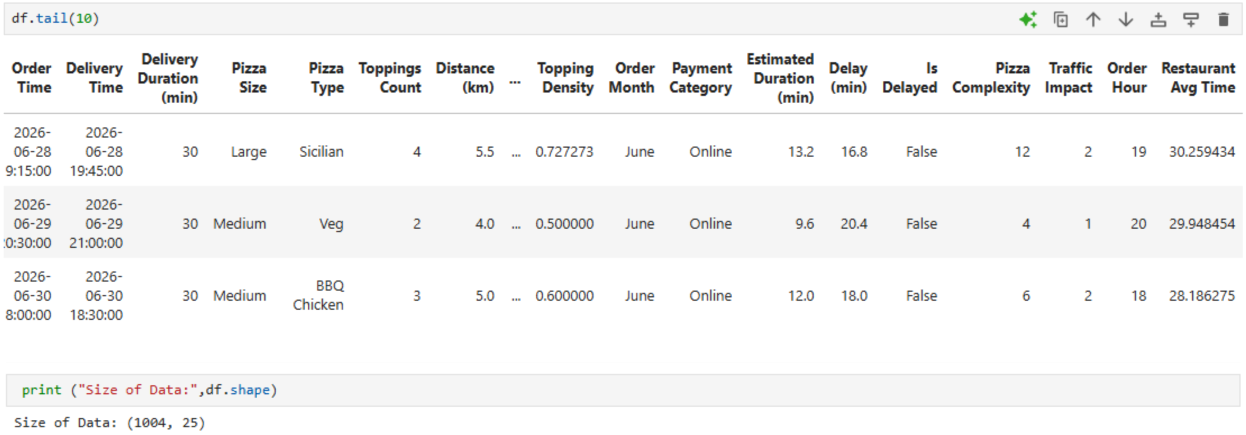
```
df.head(10)
```

Order Time	Delivery Time	Delivery Duration (min)	Pizza Size	Pizza Type	Toppings Count	Distance (km)	...	Topping Density	Order Month	Payment Category	Estimated Duration (min)	Delay (min)	Is Delayed	Pizza Complexity	Traffic Impact	Order Hour	Restaurant Avg Time
2024-01-05 18:30:00	2024-01-05 18:45:00	15	Medium	Veg	3	2.5	...	1.200000	January	Online	6.0	9.0	False	6	2	18	30.259434
2024-02-14 20:00:00	2024-02-14 20:25:00	25	Large	Non-Veg	4	5.0	...	0.800000	February	Online	12.0	13.0	False	12	3	20	28.186275
2024-03-21 12:15:00	2024-03-21 12:35:00	20	Small	Vegan	2	3.0	...	0.666667	March	Online	7.2	12.8	False	2	1	12	28.844221



```
df.tail(10)
```

	Order ID	Restaurant Name	Location	Order Time	Delivery Time	Delivery Duration (min)	Pizza Size	Pizza Type	Toppings Count	Distance (km)	...	Topping Density	Order Month	Payment Category	Estimated Duration (min)	Delay (min)
994	ORD996	Domino's	Albuquerque, NM	2026-06-28 19:15:00	2026-06-28 19:45:00	30	Large	Sicilian	4	5.5	...	0.727273	June	Online	13.2	16.8
995	ORD997	Pizza Hut	Atlanta, GA	2026-06-29 20:30:00	2026-06-29 21:00:00	30	Medium	Veg	2	4.0	...	0.500000	June	Online	9.6	20.4
996	ORD998	Papa John's	Omaha, NE	2026-06-30 18:00:00	2026-06-30 18:30:00	30	Medium	BBQ Chicken	3	5.0	...	0.600000	June	Online	12.0	18.0



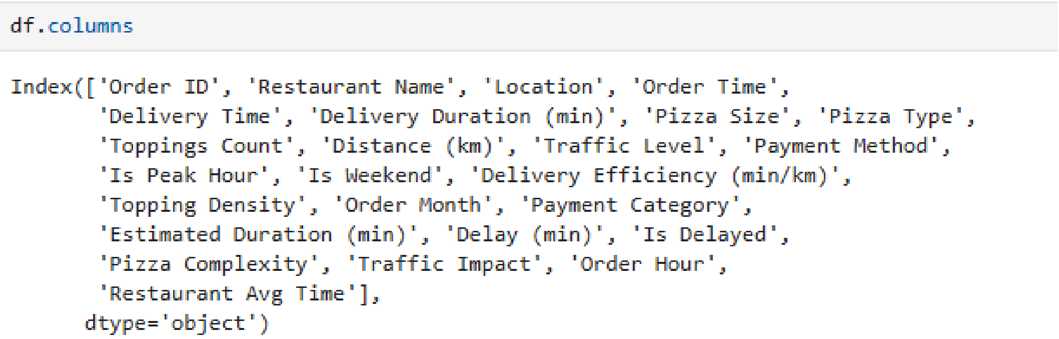
```python
df.tail(10)
```

Order Time	Delivery Time	Delivery Duration (min)	Pizza Size	Pizza Type	Toppings Count	Distance (km)	...	Topping Density	Order Month	Payment Category	Estimated Duration (min)	Delay (min)	Is Delayed	Pizza Complexity	Traffic Impact	Order Hour	Restaurant Avg Time
2026-06-28 9:15:00	2026-06-28 19:45:00	30	Large	Sicilian	4	5.5	...	0.727273	June	Online	13.2	16.8	False	12	2	19	30.259434
2026-06-29 0:30:00	2026-06-29 21:00:00	30	Medium	Veg	2	4.0	...	0.500000	June	Online	9.6	20.4	False	4	1	20	29.948454
2026-06-30 8:00:00	2026-06-30 18:30:00	30	Medium	BBQ Chicken	3	5.0	...	0.600000	June	Online	12.0	18.0	False	6	2	18	28.186275

```python
print ("Size of Data:",df.shape)
```

```
Size of Data: (1004, 25)
```

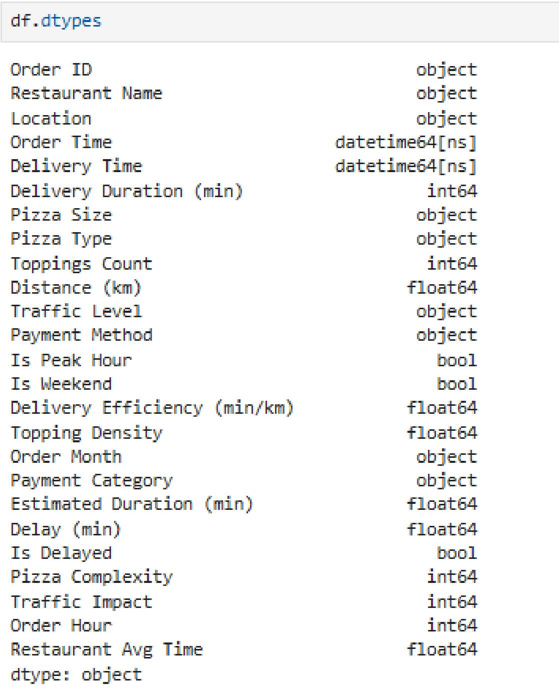
Field Info



```python
df.columns
```

```
Index(['Order ID', 'Restaurant Name', 'Location', 'Order Time',
       'Delivery Time', 'Delivery Duration (min)', 'Pizza Size', 'Pizza Type',
       'Toppings Count', 'Distance (km)', 'Traffic Level', 'Payment Method',
       'Is Peak Hour', 'Is Weekend', 'Delivery Efficiency (min/km)',
       'Topping Density', 'Order Month', 'Payment Category',
       'Estimated Duration (min)', 'Delay (min)', 'Is Delayed',
       'Pizza Complexity', 'Traffic Impact', 'Order Hour',
       'Restaurant Avg Time'],
      dtype='object')
```

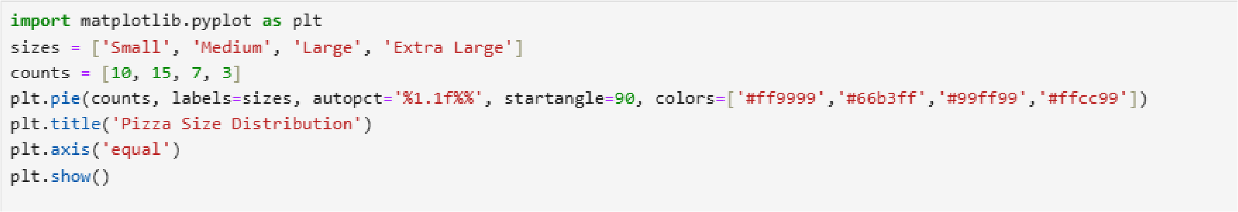
Datatypes



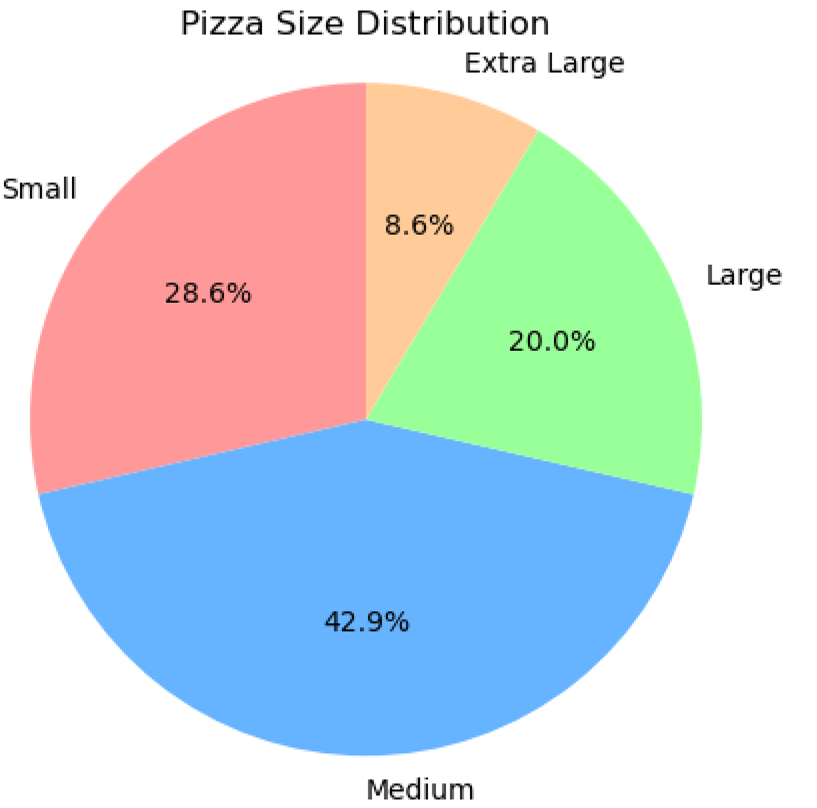
```python
df.dtypes
```

```
Order ID                                object
Restaurant Name                         object
Location                                object
Order Time                      datetime64[ns]
Delivery Time                   datetime64[ns]
Delivery Duration (min)                  int64
Pizza Size                              object
Pizza Type                              object
Toppings Count                           int64
Distance (km)                          float64
Traffic Level                           object
Payment Method                          object
Is Peak Hour                              bool
Is Weekend                                bool
Delivery Efficiency (min/km)           float64
Topping Density                        float64
Order Month                             object
Payment Category                        object
Estimated Duration (min)               float64
Delay (min)                            float64
Is Delayed                                bool
Pizza Complexity                         int64
Traffic Impact                           int64
Order Hour                               int64
Restaurant Avg Time                    float64
dtype: object
```

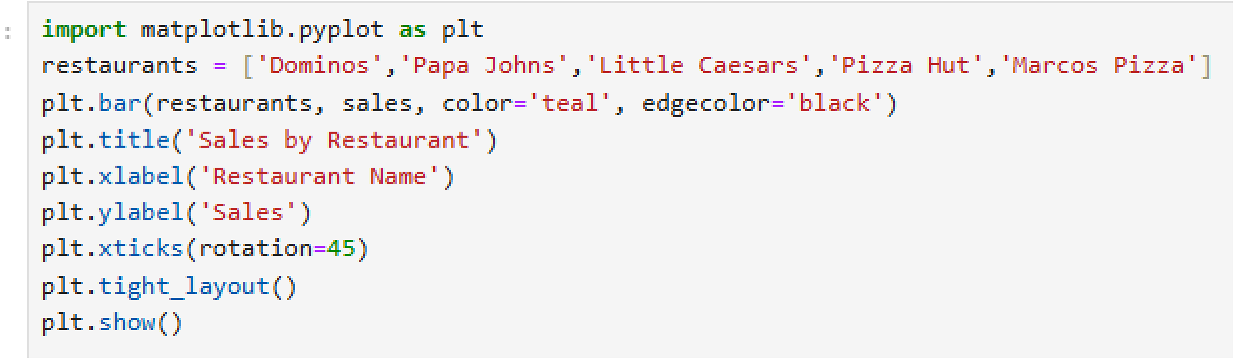
Pie chart in Pizza size



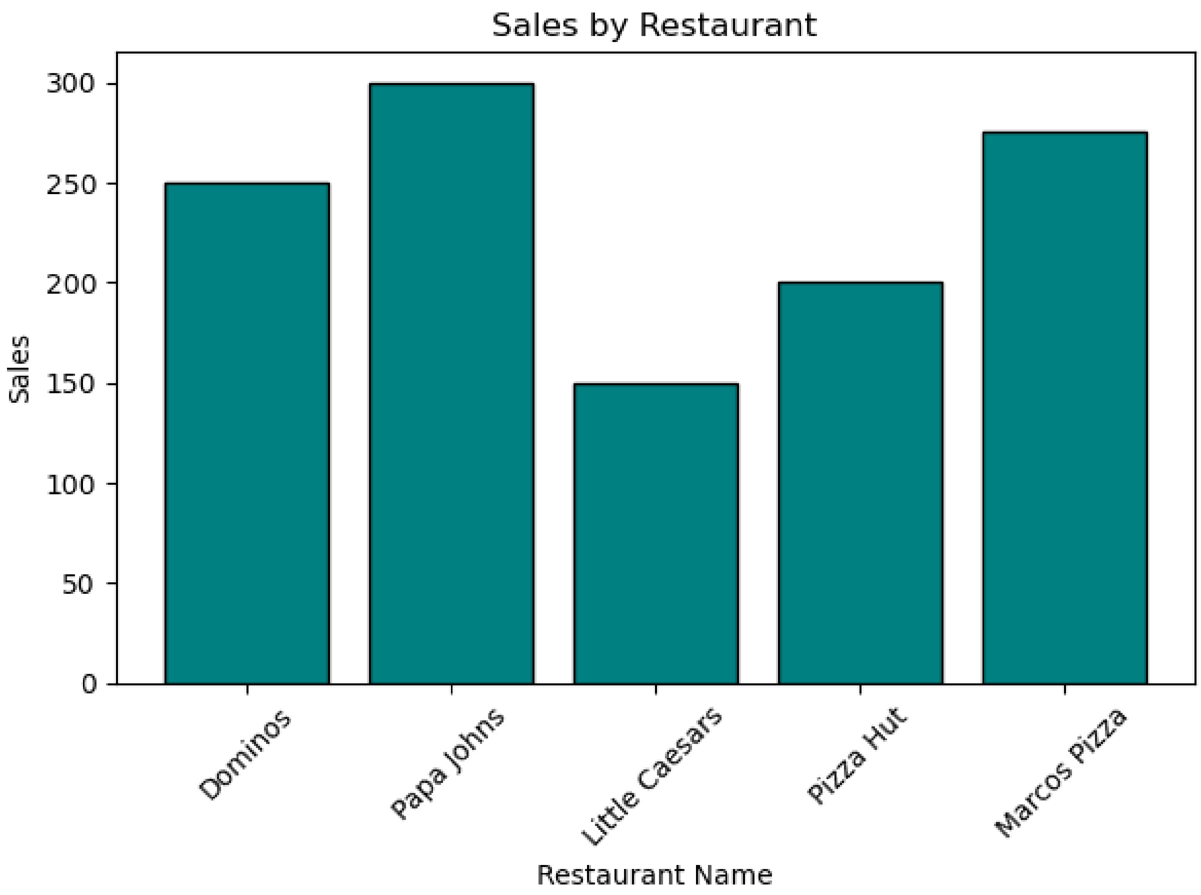
```
import matplotlib.pyplot as plt
sizes = ['Small', 'Medium', 'Large', 'Extra Large']
counts = [10, 15, 7, 3]
plt.pie(counts, labels=sizes, autopct='%1.1f%%', startangle=90, colors=['#ff9999','#66b3ff','#99ff99','#ffcc99'])
plt.title('Pizza Size Distribution')
plt.axis('equal')
plt.show()
```



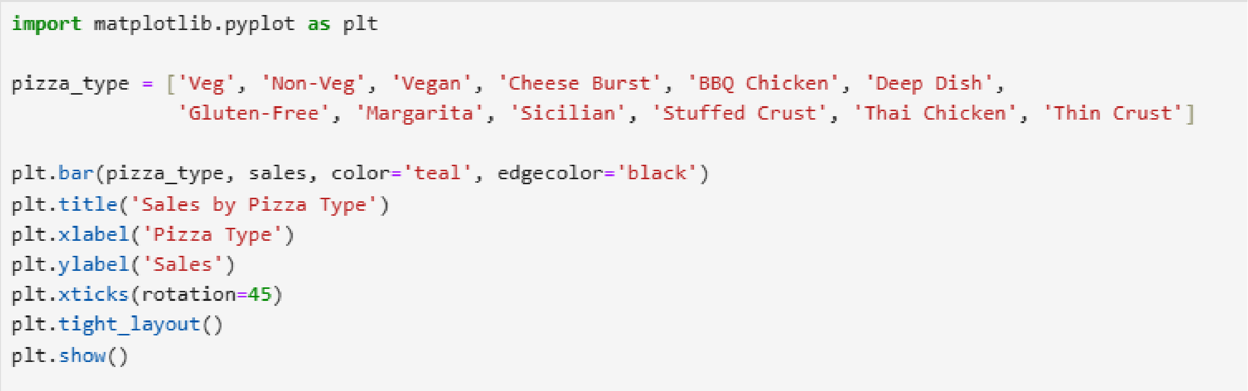
Bar chart in sales by Restaurant



```
import matplotlib.pyplot as plt
restaurants = ['Dominos','Papa Johns','Little Caesars','Pizza Hut','Marcos Pizza']
plt.bar(restaurants, sales, color='teal', edgecolor='black')
plt.title('Sales by Restaurant')
plt.xlabel('Restaurant Name')
plt.ylabel('Sales')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```



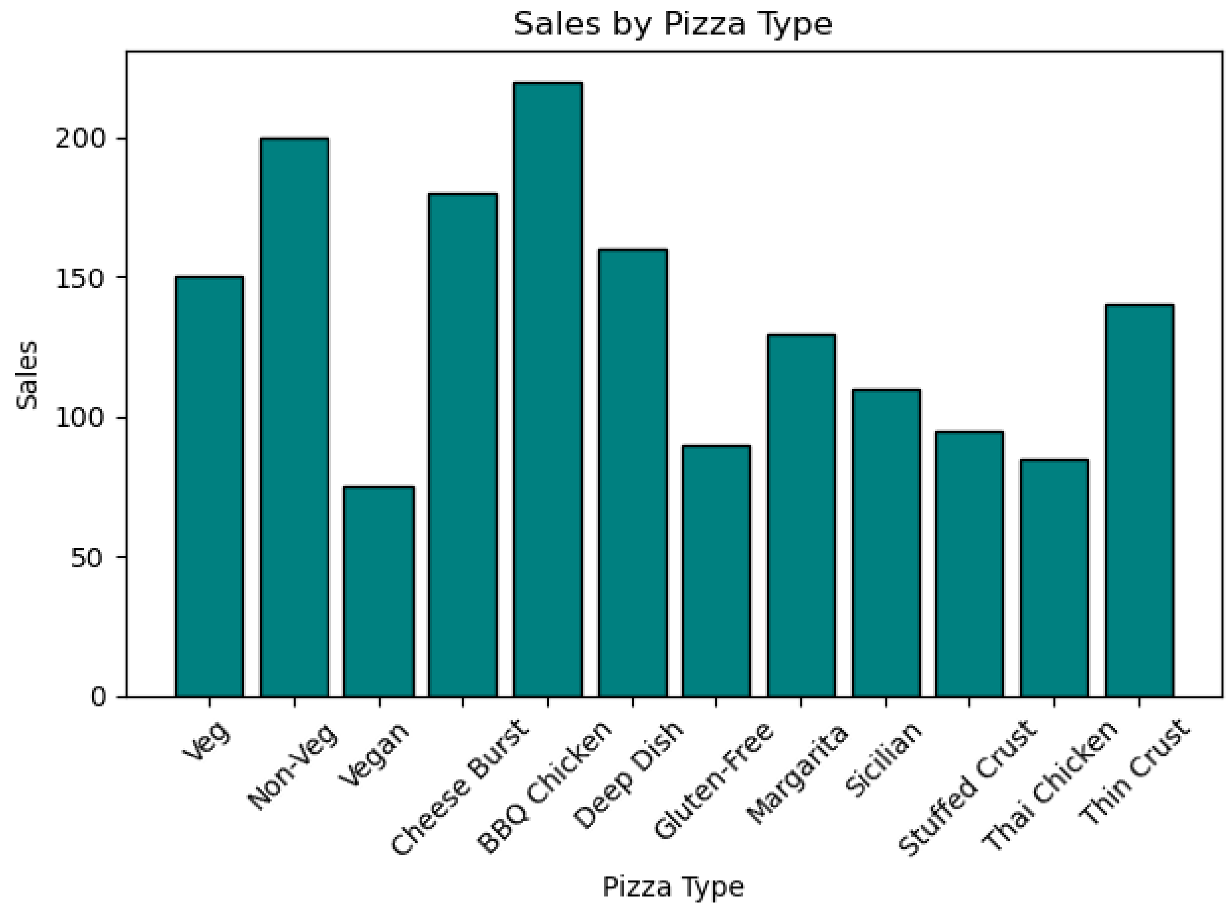
Bar chart in sales by pizza type ¶



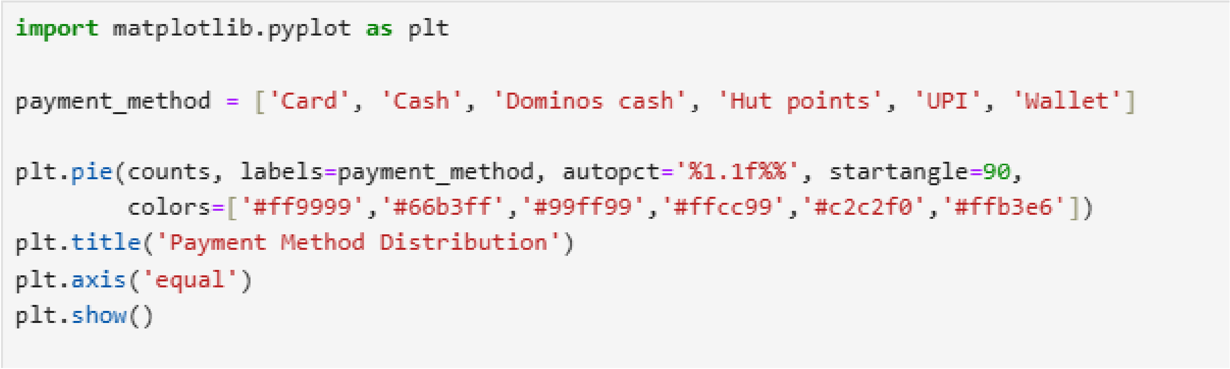
```
import matplotlib.pyplot as plt

pizza_type = ['Veg', 'Non-Veg', 'Vegan', 'Cheese Burst', 'BBQ Chicken', 'Deep Dish',
              'Gluten-Free', 'Margarita', 'Sicilian', 'Stuffed Crust', 'Thai Chicken', 'Thin Crust']

plt.bar(pizza_type, sales, color='teal', edgecolor='black')
plt.title('Sales by Pizza Type')
plt.xlabel('Pizza Type')
plt.ylabel('Sales')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
```



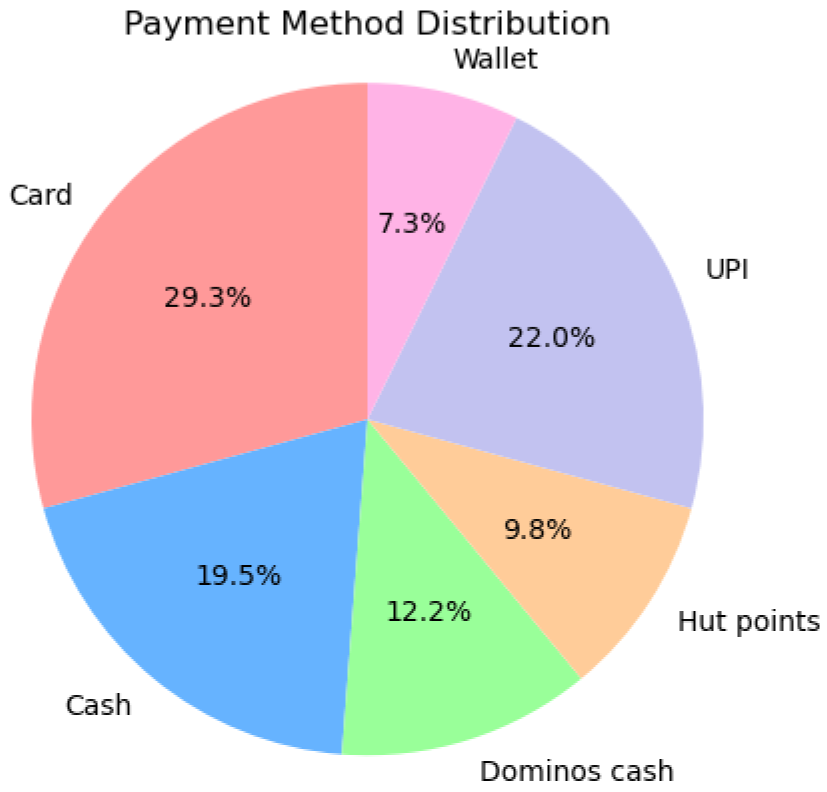
Pie Chart in payment method



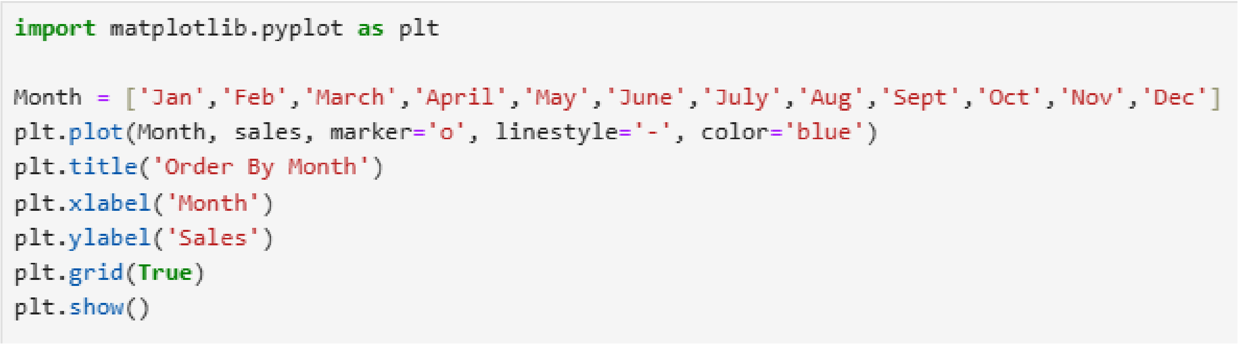
```
import matplotlib.pyplot as plt

payment_method = ['Card', 'Cash', 'Dominos cash', 'Hut points', 'UPI', 'Wallet']

plt.pie(counts, labels=payment_method, autopct='%1.1f%%', startangle=90,
        colors=['#ff9999','#66b3ff','#99ff99','#ffcc99','#c2c2f0','#ffb3e6'])
plt.title('Payment Method Distribution')
plt.axis('equal')
plt.show()
```

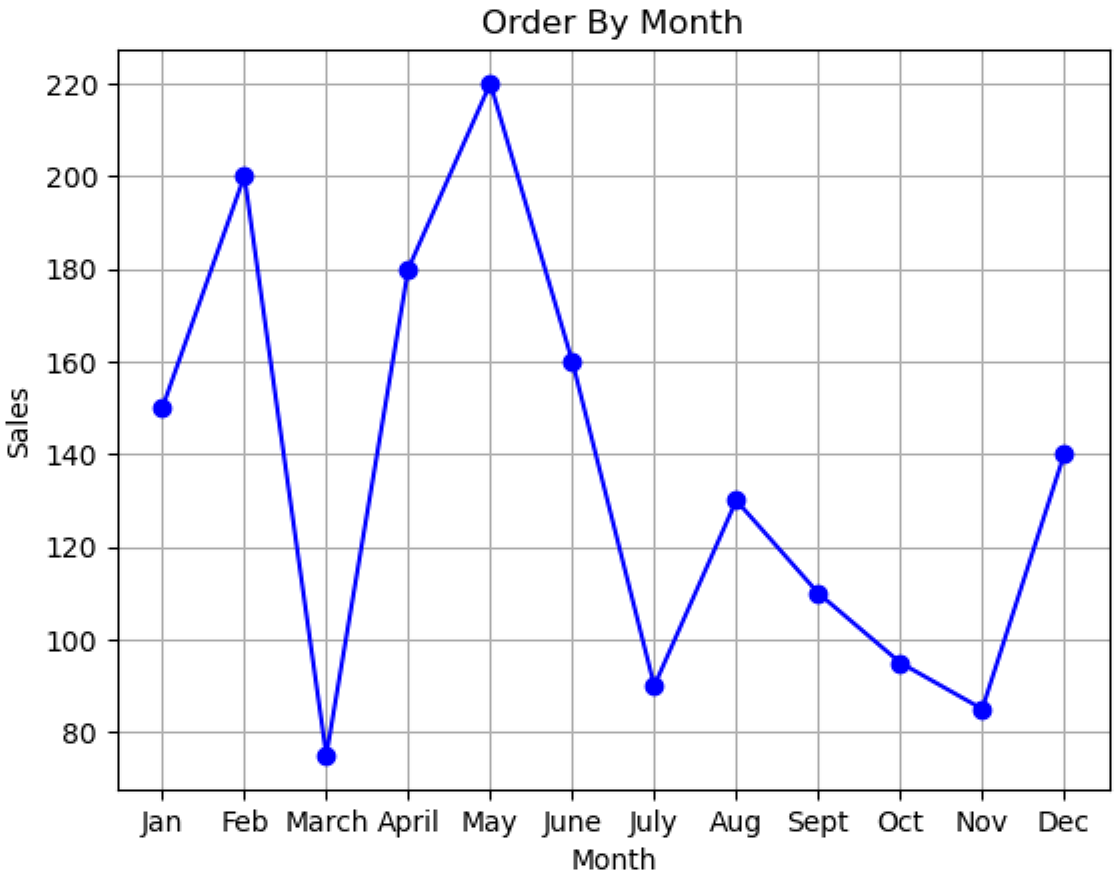


Line Chart in payment method



```
import matplotlib.pyplot as plt

Month = ['Jan','Feb','March','April','May','June','July','Aug','Sept','Oct','Nov','Dec']
plt.plot(Month, sales, marker='o', linestyle='-', color='blue')
plt.title('Order By Month')
plt.xlabel('Month')
plt.ylabel('Sales')
plt.grid(True)
plt.show()
```

Order By Month
Sales
220
200
180
160
140
120
100
80
Jan
Feb
March
April
May
June
July
Aug
Sept
Oct
Nov
Dec
Month