Question1

February 10, 2024

```python
[1]: import pandas as pd
     import matplotlib.pyplot as plt
```

```python
[2]: data = pd.read_csv('sales.csv')
```

```python
[3]: data['Total Sale'] = data['Quantity'] * data['PricePerUnit']
```

```python
[10]: max_sale = data['Total Sale'].max()
      min_sale = data['Total Sale'].min()
      Product_max_sale = data[data['Total Sale'] == max_sale]
      Product_min_sale = data[data['Total Sale'] == min_sale]
      print('Product with highest sale is:\n', Product_max_sale)
      print('Product with lowest sale is:\n', Product_min_sale)
```

```
Product with highest sale is:
     ID          Product      Category  Quantity  PricePerUnit      Date  \
17   18  Vacuum Cleaner  Electronics        30          20.0  6/3/2024

    Total Sale  Month
17       600.0      6
Product with lowest sale is:
    ID  Product    Category  Quantity  PricePerUnit       Date  Total Sale  \
4    5  Blender  Appliances         5          22.5  2/2/2024       112.5

   Month
4      2
```

```python
[5]: #Mean, median and standard deviation
     mean = data['Total Sale'].mean()
     median = data['Total Sale'].median()
     std_dev = data['Total Sale'].std()
     print('Mean:', mean)
     print('Median:', median)
     print('Standard Deviation:', std_dev)
```

```
Mean: 325.422
Median: 300.0
Standard Deviation: 133.83027970779509
```

```
[6]: # 3.Create a bar graph showcasing total sales data for each month using␣
     ↪matplotlib. What do you observe ?
     data['Month'] = pd.to_datetime(data['Date']).dt.month
     plt.bar(data['Month'], data['Total Sale'])
     plt.xlabel('Month')
     plt.ylabel('Total Sale')
     plt.title('Total Sales for each month')
     plt.show()
```
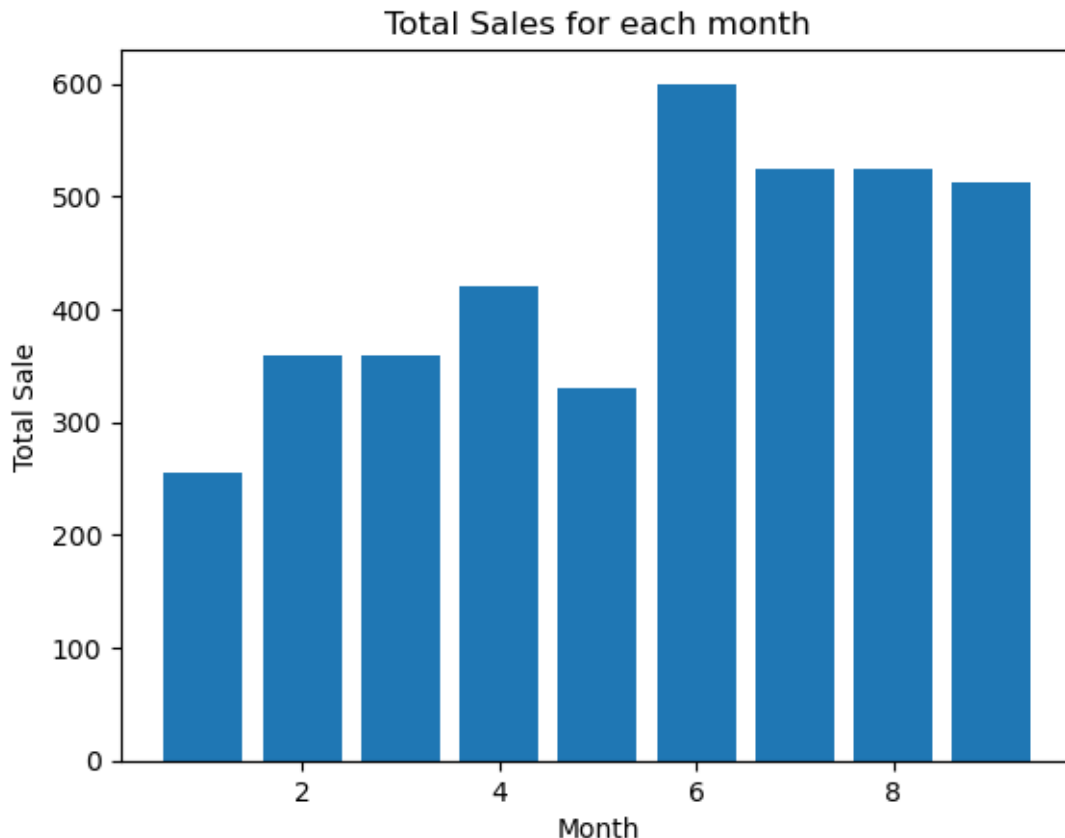


```
[7]: correlation = data['Quantity'].corr(data['PricePerUnit'])
     plt.scatter(data['Quantity'], data['PricePerUnit'])
     plt.xlabel('Quantity')
     plt.ylabel('Price Per Unit')
     plt.title('Scatter plot between Quantity and Price Per Unit')
     plt.show()
     print('Correlation between Quantity and Price Per Unit:' , correlation)
     print('There is a negative correlation between Quantity and Price Per Unit. As␣
     ↪the quantity increases, the price per unit decrease')
```

Scatter plot between Quantity and Price Per Unit
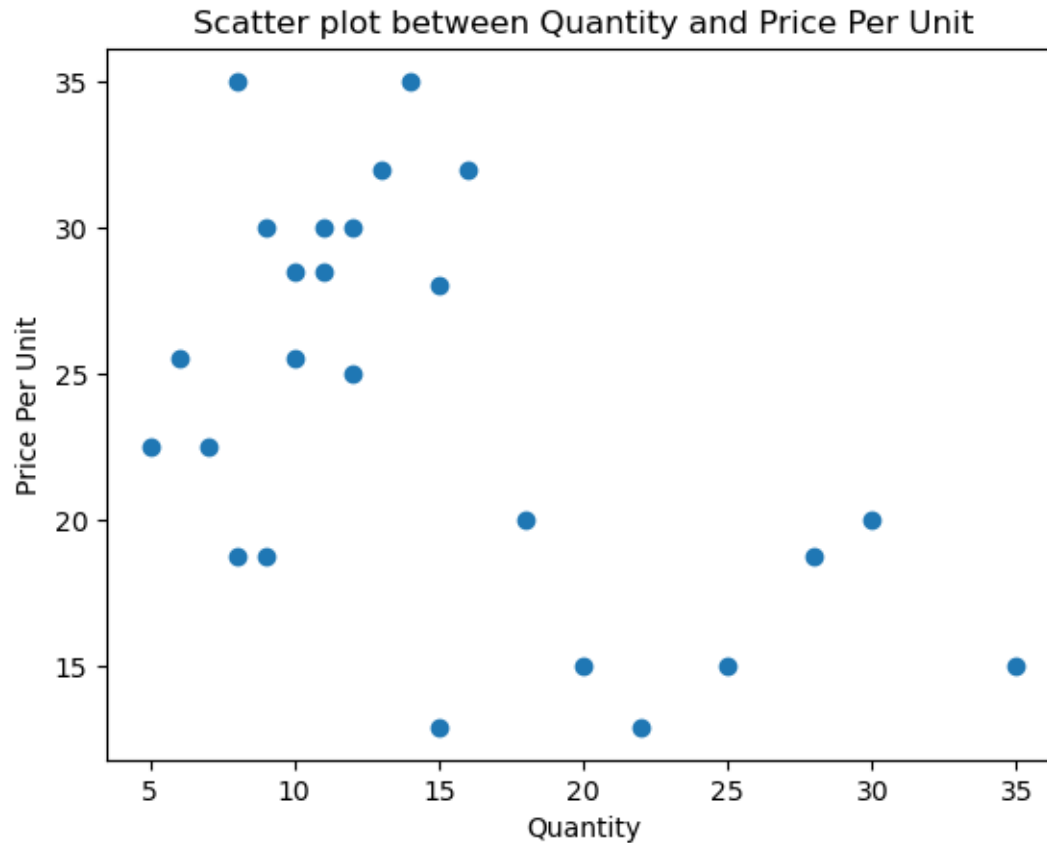

```
Correlation between Quantity and Price Per Unit: -0.5163718700038966
There is a negative correlation between Quantity and Price Per Unit. As the
quantity increases, the price per unit decrease
```

[8]:
```python
# 5. Filter the data whose quantity is greater than 10 and produced after month
 ↪of may and
#    save the filtered data into a CSV file.

filtered_data = data[(data['Quantity'] > 10) & (data['Month']> 5)]
filtered_data.to_csv('filtered_data.csv', index = False)
print(filtered_data)
```

```
    ID         Product      Category  Quantity  PricePerUnit      Date  \
15  16        Air Fryer    Appliances        13         32.00  6/1/2024
17  18   Vacuum Cleaner   Electronics        30         20.00  6/3/2024
19  20         Microwave    Appliances        12         25.00  7/2/2024
20  21             Drone   Electronics        28         18.75  7/3/2024
21  22            Camera   Electronics        14         35.00  8/1/2024
22  23              Iron   Electronics        11         28.50  8/2/2024
23  24          Soundbar    Appliances        35         15.00  8/3/2024
24  25      Laptop Stand    Appliances        16         32.00  9/1/2024
```

```
       Total Sale   Month
   15        416.0       6
   17        600.0       6
   19        300.0       7
   20        525.0       7
   21        490.0       8
   22        313.5       8
   23        525.0       8
   24        512.0       9
```

[ ]: