Fathima R

# Simple bar chart using plotly to analyze sales

October 13 2023 -part 1

## PLOTLY:

Plotly is a data visualization that helps in creating interactive visuals.It helps in creating different visuals including scatter plots,bar charts,pie charts.It is being interactive by allowing users to zoom in, zoom out, and hover over data points.It allows creating visuals by updating layout,traces,axes and also do styling that results in creating professional - looking visualizations.

## OBJECTIVE:

To analyze the sales of products of a particular company based on month, quarter and time range.

To find the top 5 selling products and 5 least selling products based on month,quarter and time.

## ABOUT THE DATASET:

Dataset variables:

```
------             --------------   -----
Order ID           185950 non-null  int64
Product            185950 non-null  object
Quantity Ordered   185950 non-null  int64
Price Each         185950 non-null  float64
Order Date         185950 non-null  object
Purchase Address   185950 non-null  object
Month              185950 non-null  int64
Sales              185950 non-null  float64
City               185950 non-null  object
Hour               185950 non-null  int64
```

**DATA PREPROCESSING:**

Data preprocessing involves cleaning, structuring the raw data into the format where it can be used in analyzing to make data driven timely decisions.

**DATA CLEANING:**

**Check for the duplicates** as duplicates misleads in analyzing the data.

```
]: df.duplicated().value_counts()
```

```
]: False    185686
   True        264
   dtype: int64
```

**Remove the duplicates**

```
df.drop_duplicates(inplace=True)
```

```
df.duplicated().value_counts()
```

```
False    185686
dtype: int64
```

**FEATURE ENGINEERING**

Feature engineering is one of the steps in preparing data that can be used in analyzing the data and also used in machine learning. It also involves creating new features or involves transforming the existing ones.

Creating separate variables month and day from the order date column.

Converting the datatype of order date from int 64 to datetime

```
df["Order Date"]=pd.to_datetime(df["Order Date"])
```

**month column:**

```
df["month"]=df["Order Date"].dt.month_name()
```

```
df["month"].sample(5)
```

```
109296     November
60190         March
92697           May
154559    September
123745      October
Name: month, dtype: object
```

**Date column:**

```
df["date"]=df["Order Date"].dt.day
```

```
df["date"].sample(5)
```

```
72497    27
97251    30
14265     1
52014    22
4210     20
Name: date, dtype: int64
```

**Creating quarter column based on month column:**

```
def func_quar(var):
    if var in [1,2,3]:
        return "Qtr1"
    elif var in [4,5,6]:
        return "Qtr2"
    elif var in [7,8,9]:
        return "Qtr3"
    else:
        return "Qtr4"
df["Quarter"]=np.vectorize(func_quar)(df["Month"])
```

```
df.sample()
```

| | Order ID | Product | Quantity Ordered | Price Each | Order Date | Purchase Address | Month | Sales | City | Hour | date | month | Quarter |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **67631** | 173871 | USB-C Charging Cable | 1 | 11.95 | 2019-03-30 10:49:00 | 797 Walnut St, Boston, MA 02215 | 3 | 11.95 | Boston | 10 | 30 | March | Qtr1 |

### Creation of time-range column from order date

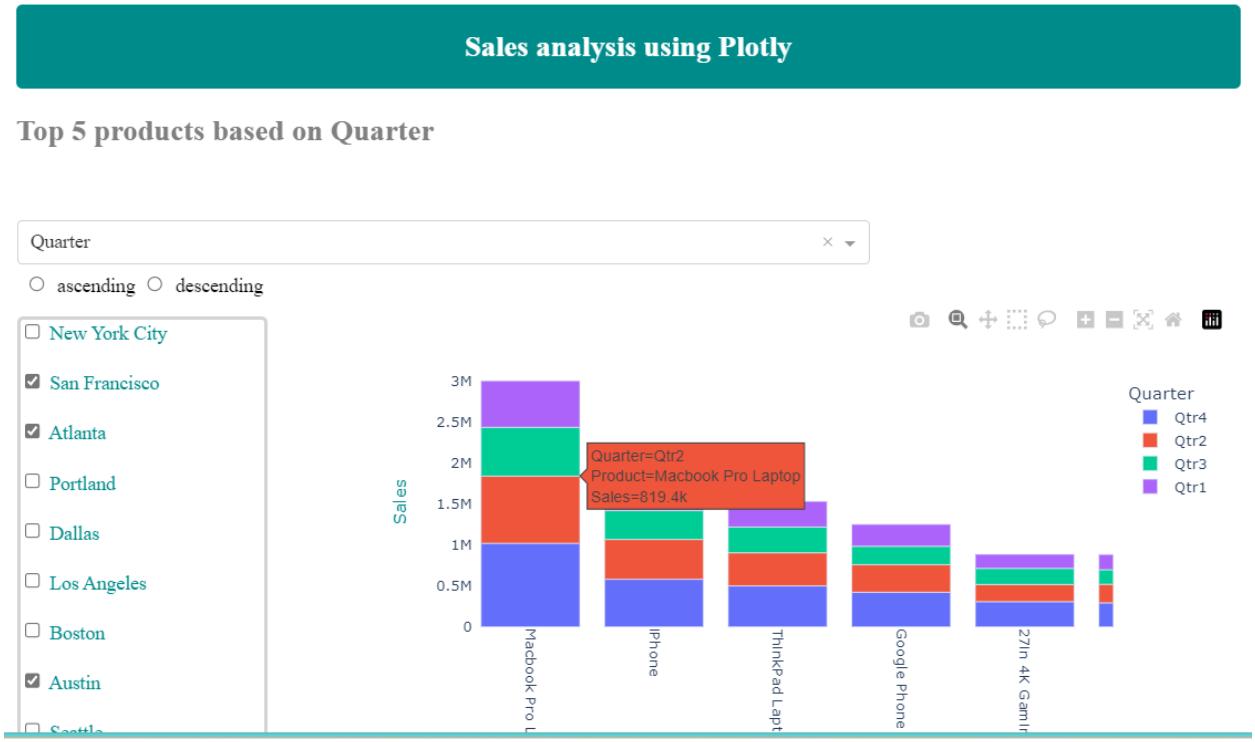
```
def func_time(hour):
    if hour in range(3,7):
        return "early morning"
    elif hour in range(7,12):
        return "morning"
    elif hour in range(12,16):
        return "afternoon"
    elif hour in range(16,20):
        return "evening"
    else:
        return "night"
df["time-range"]=np.vectorize(func_time)(df["Hour"])
```

```
df.sample()
```

| | Order ID | Product | Quantity Ordered | Price Each | Order Date | Purchase Address | Month | Sales | City | Hour | date | month | Quarter | time-range |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **61733** | 168212 | AA Batteries (4-pack) | 1 | 3.84 | 2019-03-03 10:43:00 | 175 10th St, Dallas, TX 75001 | 3 | 3.84 | Dallas | 10 | 3 | March | Qtr1 | morning |

**PRODUCT SALES ANALYSIS USING BAR CHART IN PLOTLY:**

**Based on quarter**




This plotly chart helps in analyzing  the first five products based on quarter.

Filter helps in analyzing **Top 5 products** for one or more cities together.

By zoom in can find the first n  products based on user preference.

With the help of radiobutton can design the bar chart in ascending or descending order.

**Based on time-range**