CASE STUDY

Visualizations in Python on Sales Data

DATA VISUALIZATION:

Data visualization is a general term that describes any effort to help people understand the significance of **data** by placing it in a visual context. Patterns, trends and correlations that might go undetected in text-based **data** can be exposed and recognized easier with **data visualization** software.

It enables decision makers to see analytics presented visually, so they can grasp difficult concepts or identify new patterns.

WHY IS DATA VISUALIZATION IMPORTANT?

Using charts or graphs to visualize large amounts of complex data is easier than poring over spreadsheets or reports because human brain perceives, understands and grasps visuals better and faster than numbers or alphabets. Data visualization is a quick, easy way to convey concepts in a universal manner – and you can experiment with different scenarios by making slight adjustments.

Data visualization can:

- Identify areas that need attention or improvement.

- Clarify which factors influence customer behavior.

- Help you understand which products to place where.

- Predict sales volumes.

CASE STUDY IN Python:

The objective of this dashboard is to understand sales trends for one of the leading pharmaceutical company.

The client would like to come see a dynamic dashboard with different KPI's at different levels (National, Region & Territory etc). Create the below charts as these would help us explore the past data in a better manner and give a good picture of the progress and failures. This in turn would catalyse the decision-making process, making it easier, simpler and accurate.

Visit https://www.sandeepsappal.in

ABOUT DATA:

The data attached is a two-year sales data of a pharma company which talks about sales in 2015 and 2016 across various regions and time frames.

Account Id   : Customer ID

Account Name : Customer Name

Tier   : Customer Segment

Sales 2015   : Sales for the year 2015

Sales 2016   : Sales for the year 2015

Units 2015   : No of Units sold for 2015

Units 2016   : No of Units sold for 2016

IMPORTANT NOTE: When answering the below questions, you would require to reshape the data i.e., convert from wide to long and long to wide (with stack(), unstack() and pivot()) after aggregating it by using groupby(). Please refer to the material at Pandas – Data Understanding.ipynb for details.

1. Compare Sales by region for 2016 with 2015 using bar chart

2. What are the contributing factors to the sales for each region in 2016. Visualize it using a Pie Chart.

3. Compare the total sales of 2015 and 2016 with respect to Region and Tiers

4. In East region, which state registered a decline in 2016 as compared to 2015?

5. In all the High tier, which Division saw a decline in number of units sold in 2016 compared to 2015?

6. Create a new column Qtr using numpy.where() or any suitable utility in the imported dataset. The Quarters are based on months and defined as -

   • Jan - Mar : Q1
   • Apr - Jun : Q2
   • Jul - Sep : Q3

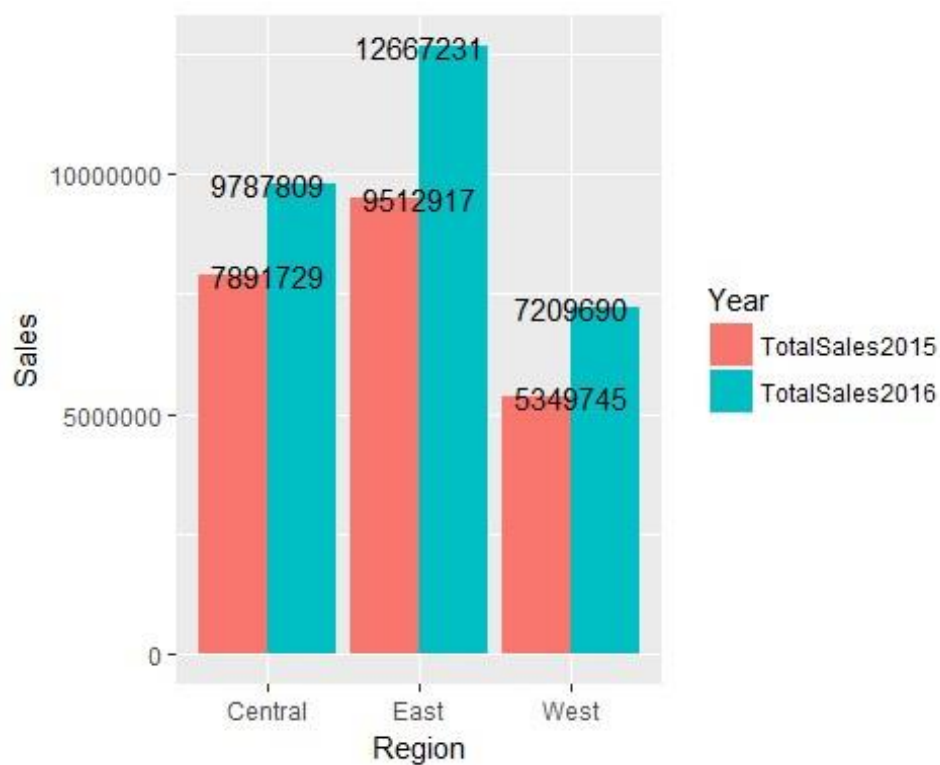Visit https://www.sandeepsappal.in

- Oct - Dec : Q4

7. Compare Qtr wise sales in 2015 and 2016 in a bar plot

8. Determine the composition of Qtr wise sales in and 2016 with regards to all the Tiers in a pie chart.

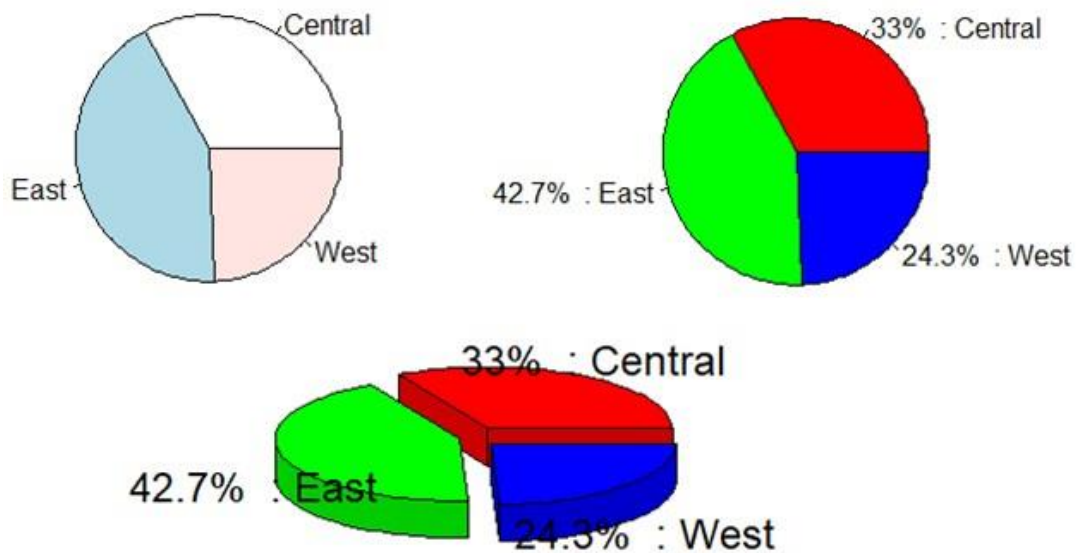   (Draw 4 pie charts representing a Quarter for each Tier)

SAMPLE OUTPUTS:

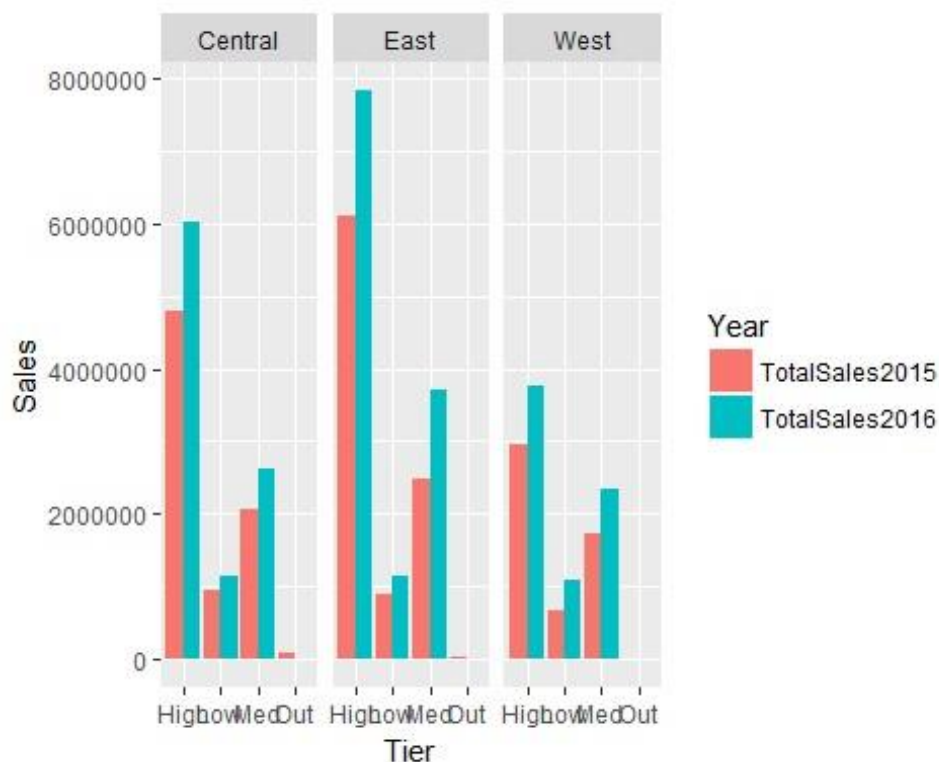1. Compare Sales by region for 2016 with 2015 using bar chart

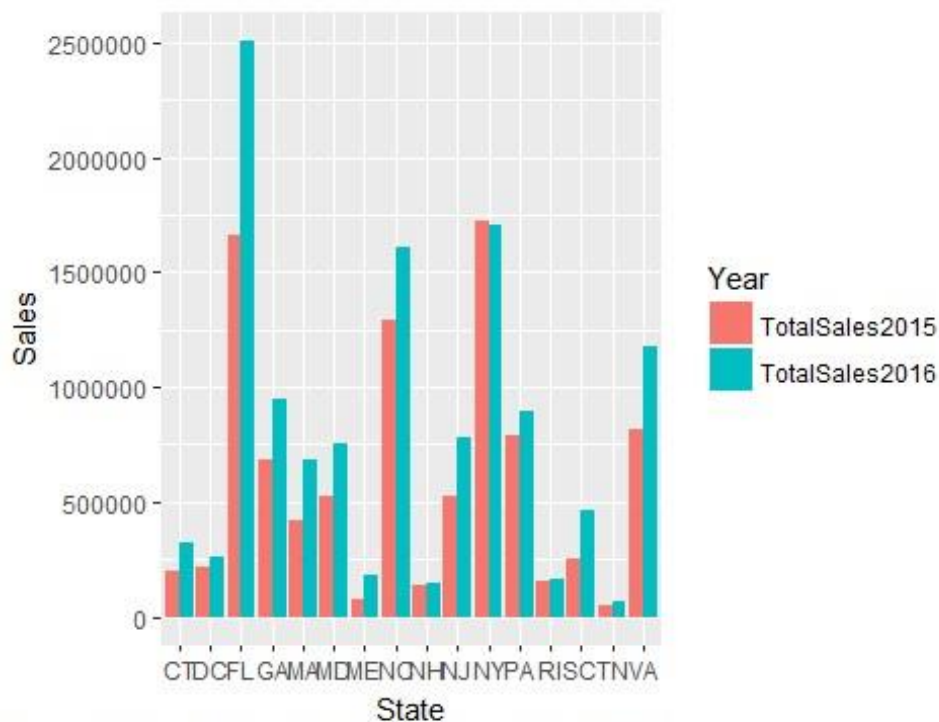2. Pie charts for sales for each region in 2016

Visit https://www.sandeepsappal.in

Pie Chart of Sales 2016

3. Compare sales of 2015 and 2016 with Region and Tiers
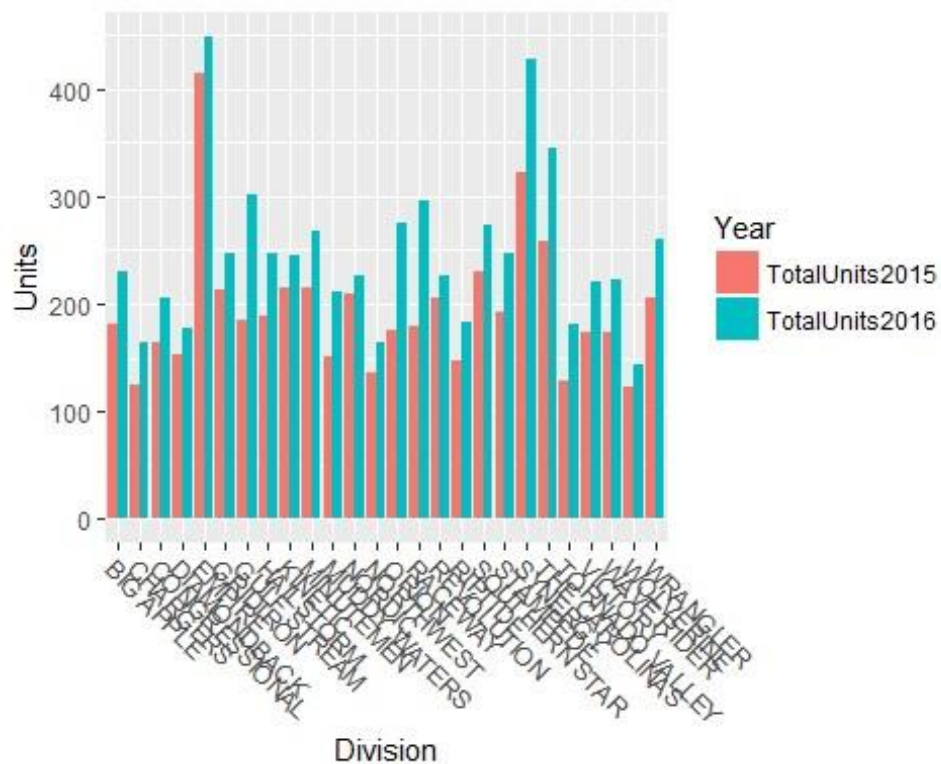

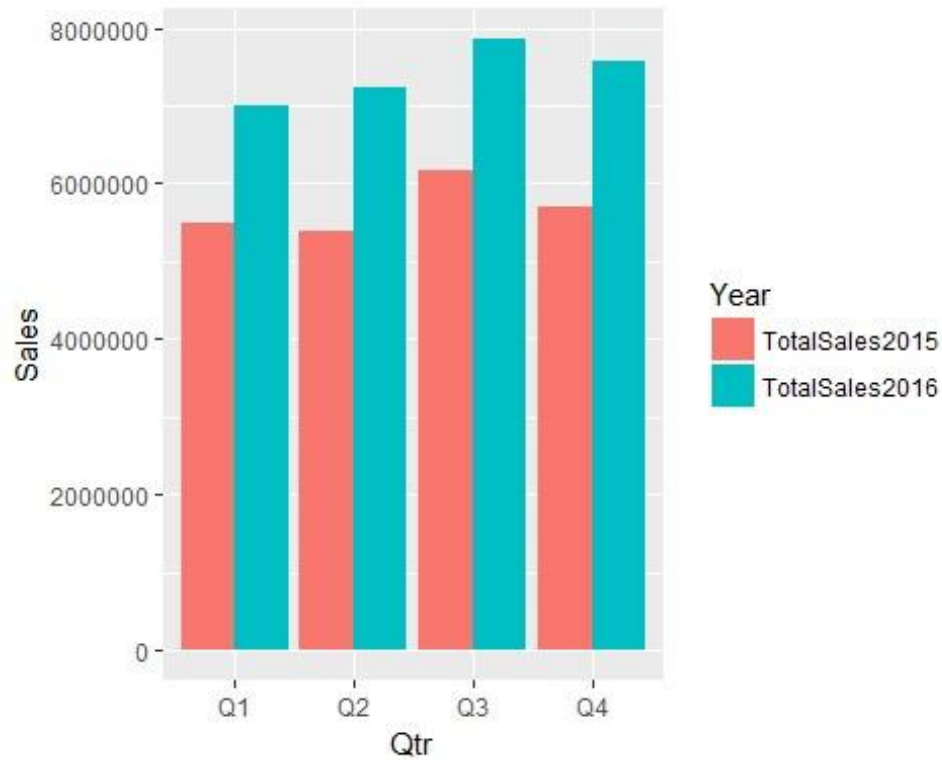4. In East region, which state registered a decline in 2016 as compared to 2015?

In East region, NY registered a decline in sales in 2016

5.   In all the High tier, which Division saw a decline in number of units sold in 2016 compared to
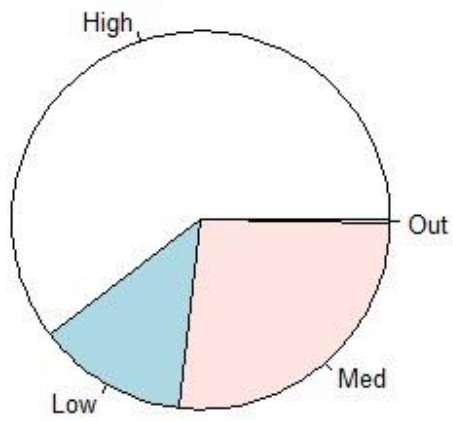

2015?

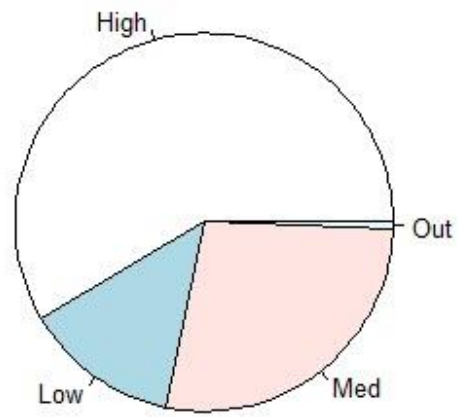7. Compare Qtr wise sales in 2015 and 2016 in a bar plot

8. Determine the composition of Qtr wise sales in and 2016 with regards to all the Tiers in a pie chart. (Draw 4 pie charts representing a Quarter for each Tier)
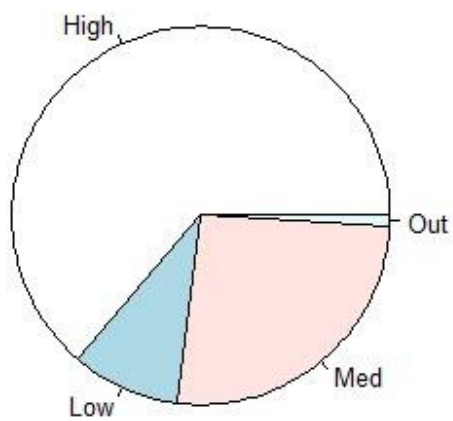
Qtr1

Qtr2

Qtr3

Qtr4

Visit https://www.sandeepsappal.in