Lab Manual Subject: Foundations of Data Analysis Laboratory (DJ19DSL303)

Semester: III

Experiment 1 (Data Visualization)

Name: Dev Patel Sap ID: 60009200016

Batch: K/K1 Date: 16/11/2021

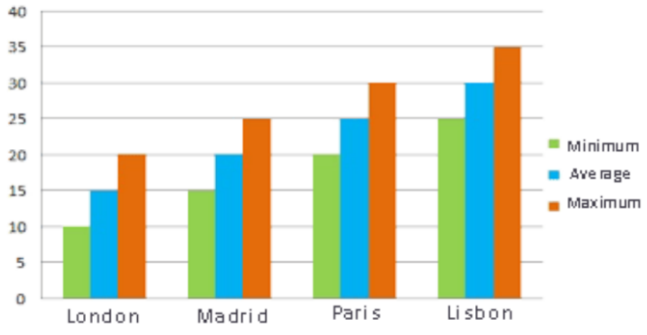
Aim: Create new measures on a given dataset and visualize them using a bar graph.

Theory:

Bar Graphs:

A bar graph is a graphical display of data using bars of different heights. The heights are proportional to the values they represent. Bar graphs are a good way to show relative sizes. A vertical bar chart is sometimes called a column chart.

A bar graph shows comparisons among discrete categories. One axis of the chart shows the specific categories being compared, and the other axis represents a measured value. Bar charts have a discrete domain of categories and are usually scaled so that all the data can fit on the chart. When there is no natural ordering of the categories being compared, bars on the chart may be arranged in any order. Bar charts arranged from highest to lowest incidence are called Pareto charts.



Grouped (clustered) and stacked bar graphs:

In grouped (clustered) bar charts, for each categorical group there are two or more bars color-coded to represent a particular grouping. For example, a

business owner with two stores might make a grouped bar chart with different coloured bars to represent each store: the horizontal axis would show the months of the year and the vertical axis would show revenue. Alternatively, a stacked bar chart stacks bars on top of each other so that the height of the resulting stack shows the combined result. Stacked bar charts are not suited to data sets having both positive and negative values.

Tool Used:

Tableau Public 2021.3

Visualizations:

DATASET 1: This dataset included the figures related to the coronavirus pandemic. We are required to answer the following questions using the visualizations prepared using this dataset.

a) **The United Nations (UN) wants to know what Countries are suffering more deaths due to Covid-19, in order to donate additional vaccines and prevent the spread of Covid, from the additional funds and vaccine budget of this month.**

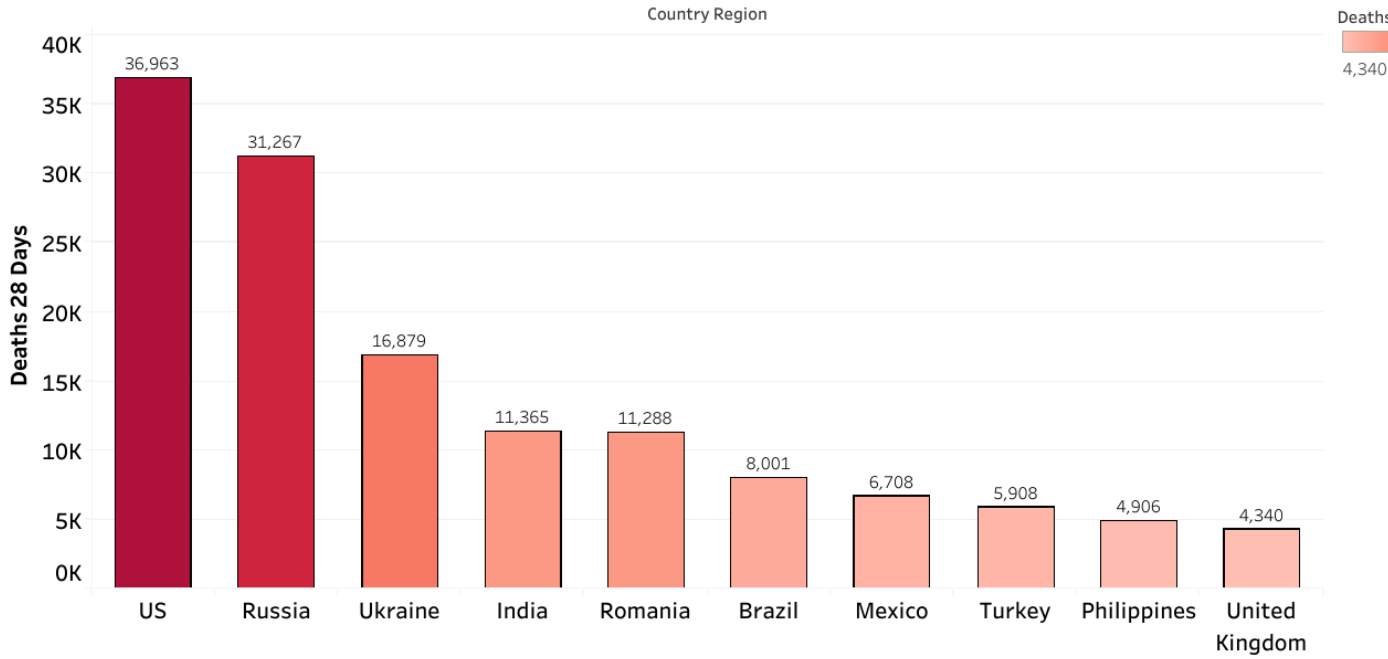
As the United Nations wants to donate vaccines and funds and prepare the vaccine budget of this month for countries suffering more deaths, the below graph shows the top ten countries that have suffered more deaths in the past four weeks. This graph gives a better idea of which countries require the most help, especially in the upcoming month, hence the vaccine budget of the month can be prepared accordingly.

On X axis: Country Regions

On Y axis: Number of deaths in the past 28 days

Ans: The top 10 countries with most deaths in the last 4 weeks are the United States, Russia, Ukraine, Romania, India, Brazil, Mexico, Turkey, Philippines and the United Kingdom.

a. Most deaths due to coronavirus per country



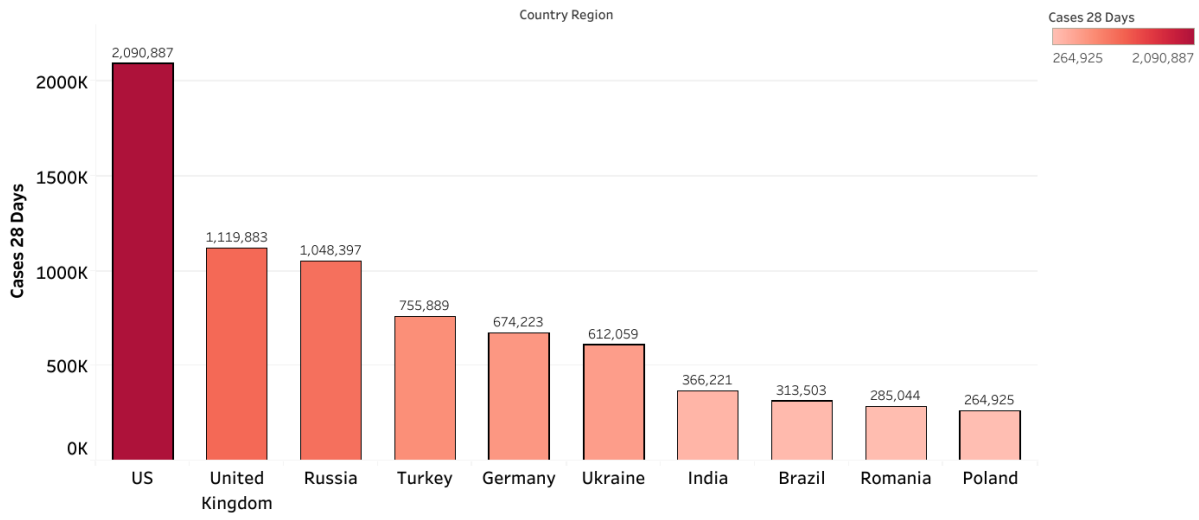
b) Find out what are the top 10 Countries suffering severely due to covid in the past month eligible for additional vaccine donations.

For additional vaccine donations, we need a graph showing the top ten countries with the most coronavirus cases in the last 4 weeks.

On X axis: Country Regions

On Y axis: Number of Coronavirus cases in the past 4 weeks.

b. Top 10 countries with the most COVID cases in the past month



Ans: The top ten countries with most COVID cases in the last 28 days are the United States, United Kingdom, Russia, Turkey, Germany, Ukraine, India, Brazil, Romania and Poland.

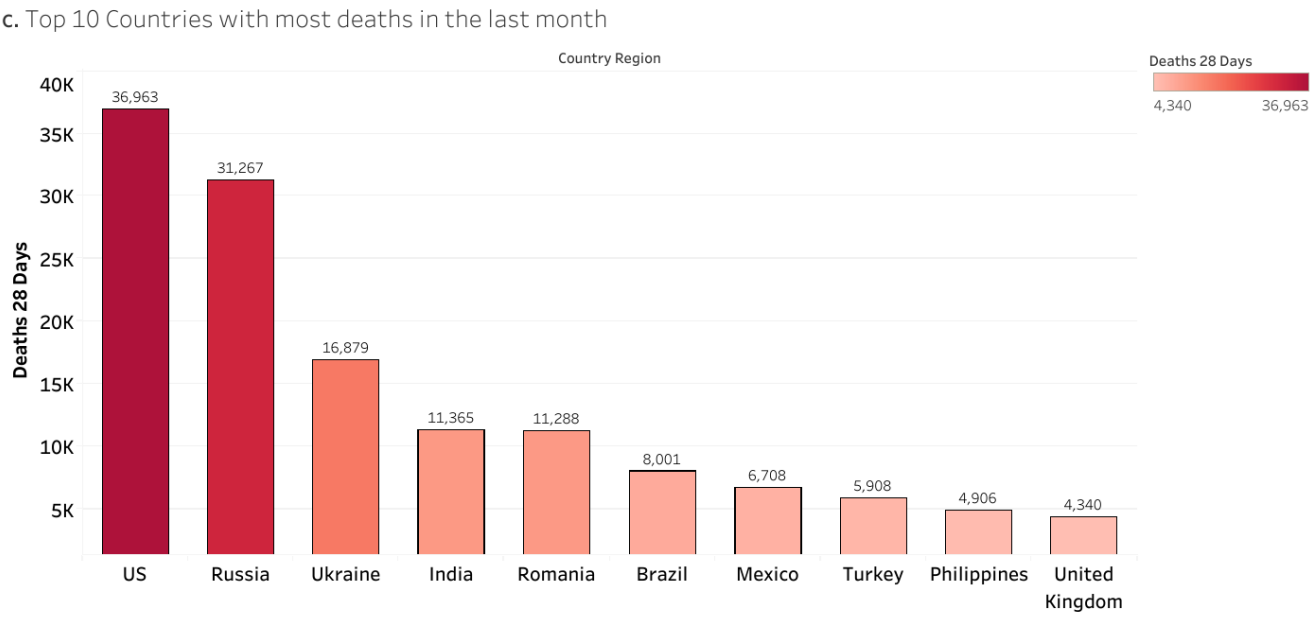
c) Need for vaccines and funding are estimated from the number of deaths in the past month.

The countries need to be arranged in descending order of the number of deaths caused due to Coronavirus in the past month.

On X axis: Country Regions

On Y axis: Number of deaths in the past 28 days

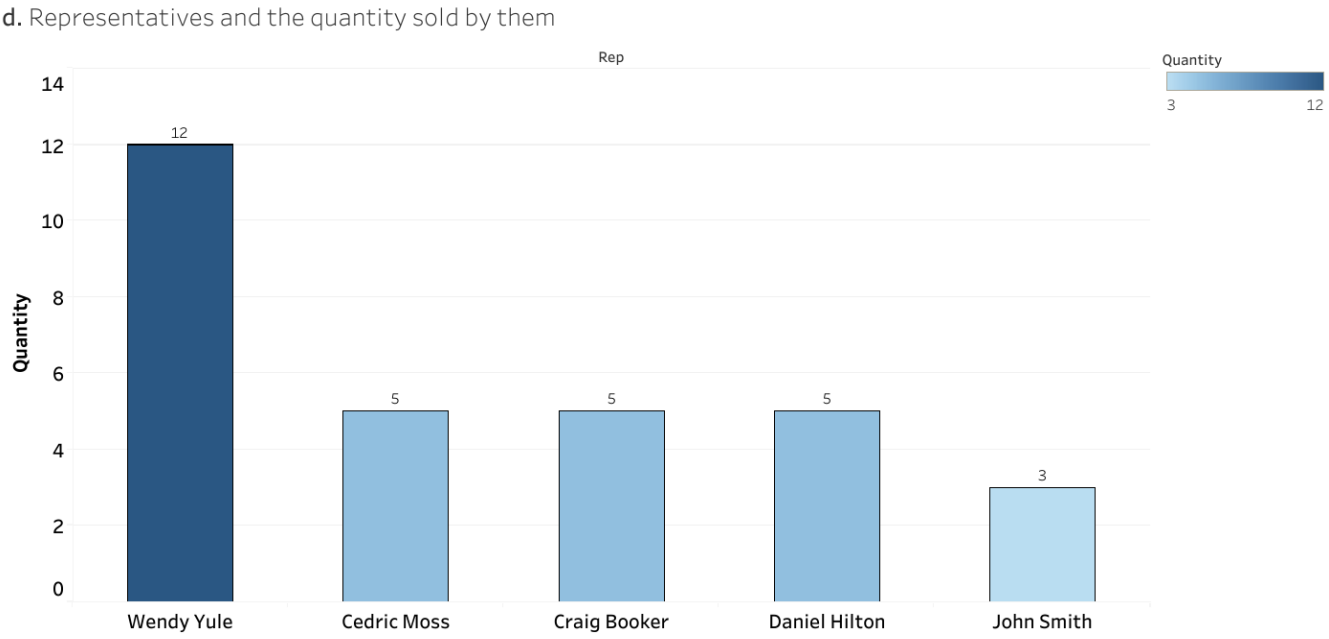
Ans: The top ten countries with the most casualties in last 28 days are the United States, Russia, Ukraine, Romania, India, Brazil, Mexico, Turkey, Philippines and the United Kingdom.



DATASET 2: This dataset shows the sales data of a company which contains the products sold by the company, the representative that carried out the sale, the manager the rep works under, the amount of the sale and other related data. The next questions are to be answered using the visualizations carried out using this dataset.

d) Which representative has sold highest quantity?

The below graph is presented in the descending order of the most quantity of services sold by the representatives.
On X axis: Representatives
On Y axis: The quantity of products and services provided
Ans: We can say that Wendy Yule has sold the highest quantity of products and services, 12 to be precise, significantly more than anyone else.



e) Which representative has sold highest price value?
f) Who is the manager of the representative who sold highest price?

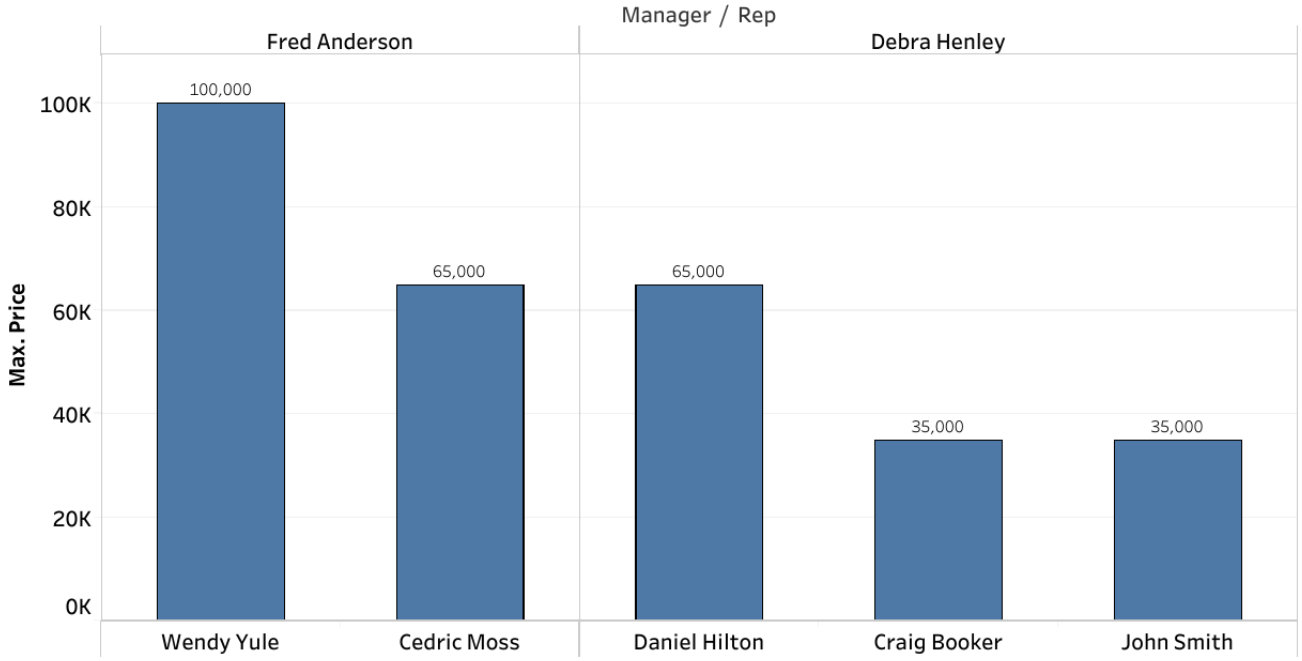
The highest price value is the single highest sale carried out by the representative during the given period of time. So we need the maximum of the prices of the sales done by each of the representatives.
On X axis: Representatives categorized on the basis of the manager they work under.
On Y axis: Maximum value of the price sold

Ans:

e. As per this graph, Wendy Yule has sold the service of the highest price value (100,000).

f. Also, we can see that Wendy Yule, who has the sold the highest price value, works under the manager Fred Anderson.

e and f. Representatives grouped with managers and their highest price



g) Which product is sold the most?

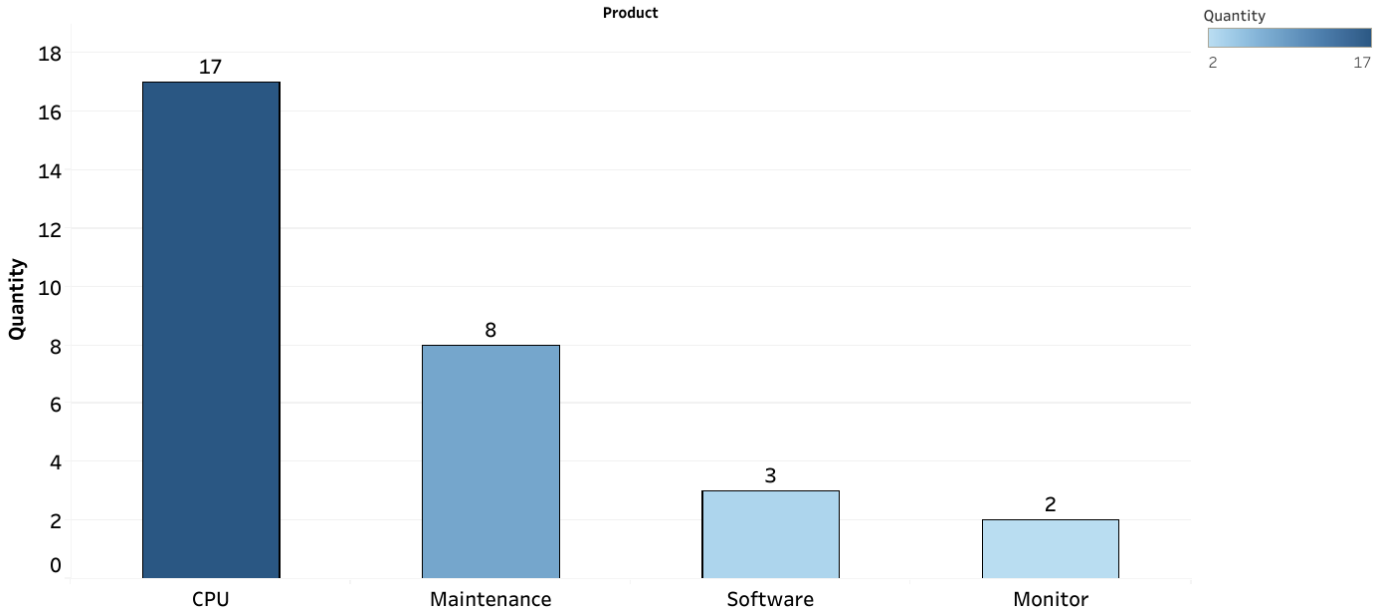
The visualization below shows the number of products that were sold in the given time frame.

On X axis: The name of the product

On Y axis: Quantity

Ans: The CPU, was the best-selling product. 17 CPU were sold in the given time frame, more than the double of any other product.

g. Quantity of products sold



Conclusion:

We learnt how we can create new measures and visualize data using bar graphs in Tableau.