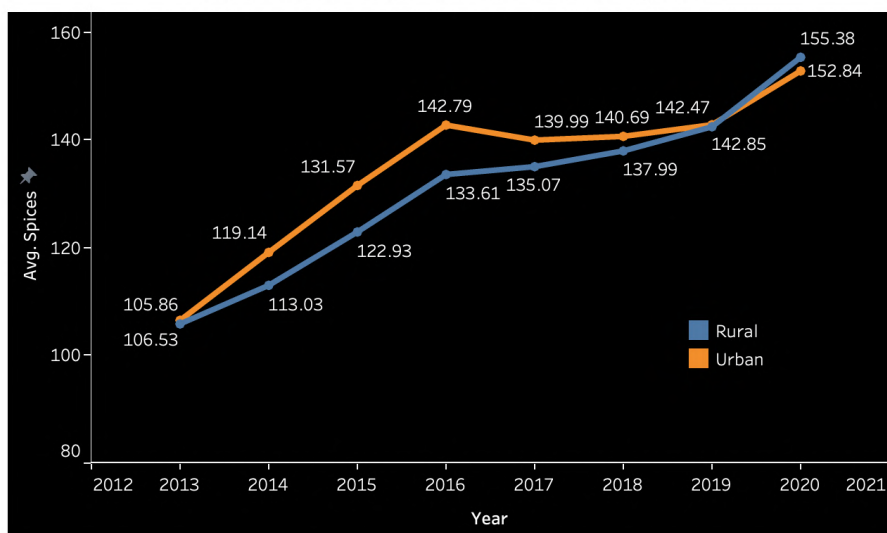


Assignment 2

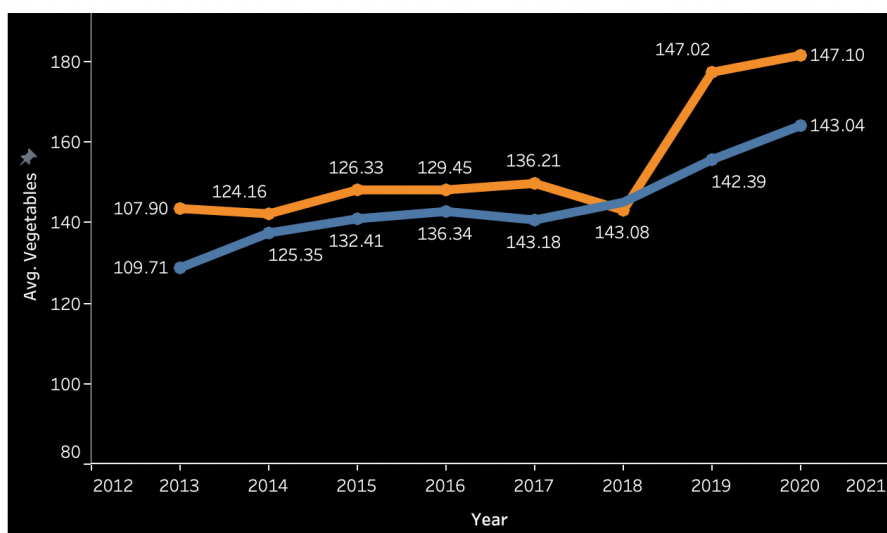
Submitted by Anitha Balachandran (016684486)

Chapter 2: Mapping Data into Aesthetics using Tableau

Trends in Average CPI Values for Spices



Trends in Average CPI Values for Vegetables



The color, size, shape, position, and line type aesthetics are used as below:

- Color is used to differentiate between the different sectors (Rural, Urban).
- Line size is used to make the lines more visible and distinguishable from each other.
- Data labels are added to the points to show the actual CPI values for each sector and year.
- Line type (solid) is used to show the continuous trend in CPI values over the years.

The use of all the above aesthetics helps to highlight the important trends and differences in the CPI values for different categories.

The data used in the above plot consists of a combination of categorical, continuous, and discrete data types.

- Categorical data: The sectors - Rural and Urban - are categorical data as they represent distinct categories with no specific order.
- Continuous data: The CPI values for Spices and Vegetables are continuous data as they represent values on a continuous scale that can take on any value within a range.
- Discrete data: The years are discrete data as they represent distinct, non-continuous values that are ordered in time.

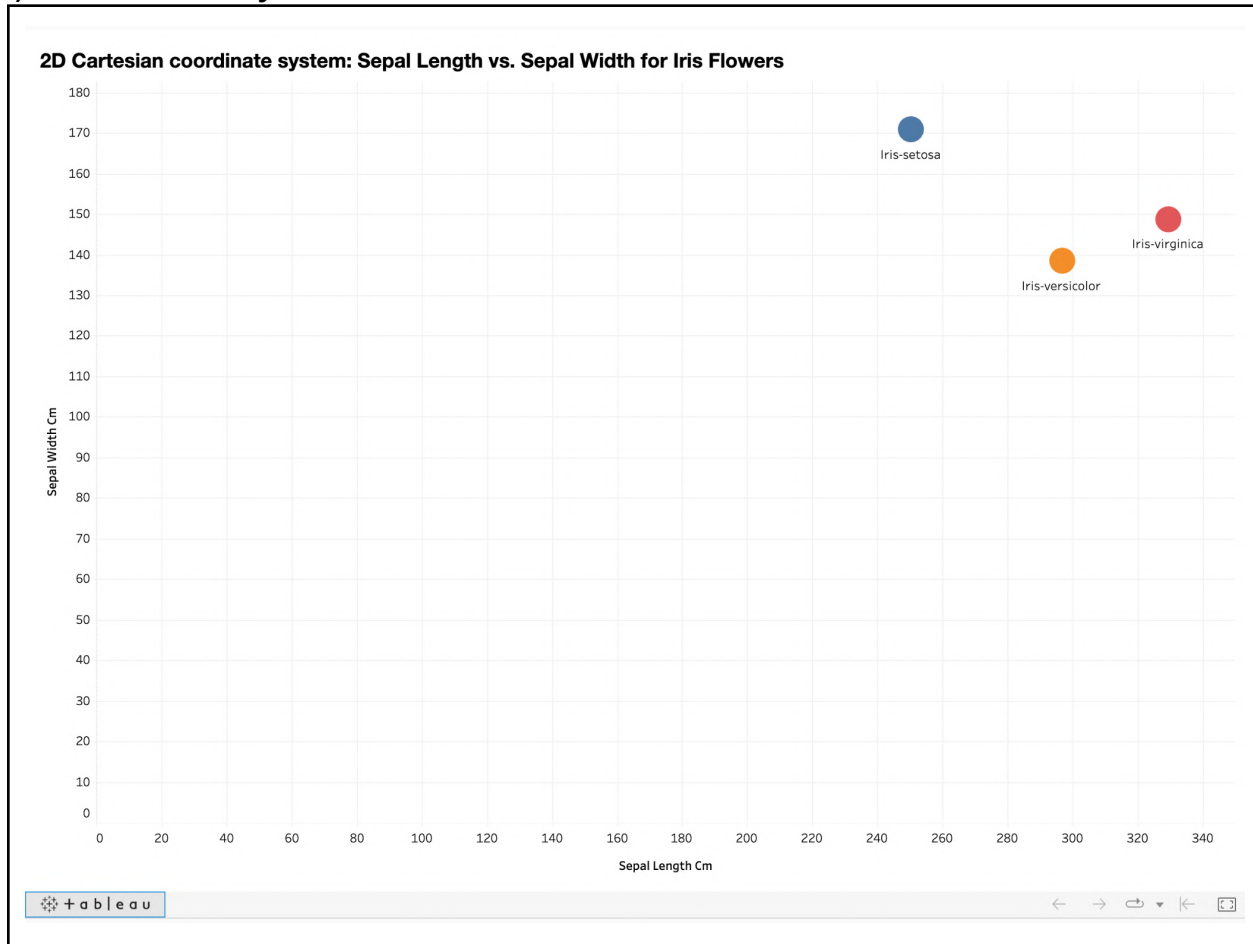
Observations from the above dashboard:

1. The CPI values for both Spices and Vegetables have been increasing over the years.
2. The CPI values for Vegetables are higher than the CPI values for Spices for most years and sectors.
3. The Rural sector tends to have higher CPI values compared to the Urban sector for both Spices and Vegetables.
4. There is a significant increase in the CPI values for both Spices and Vegetables in the year 2020, which can be attributed to the COVID-19 pandemic.

=====

Chapter 3 : Coordinate Systems and Axis using Tableau

i) 2D Coordinate system:



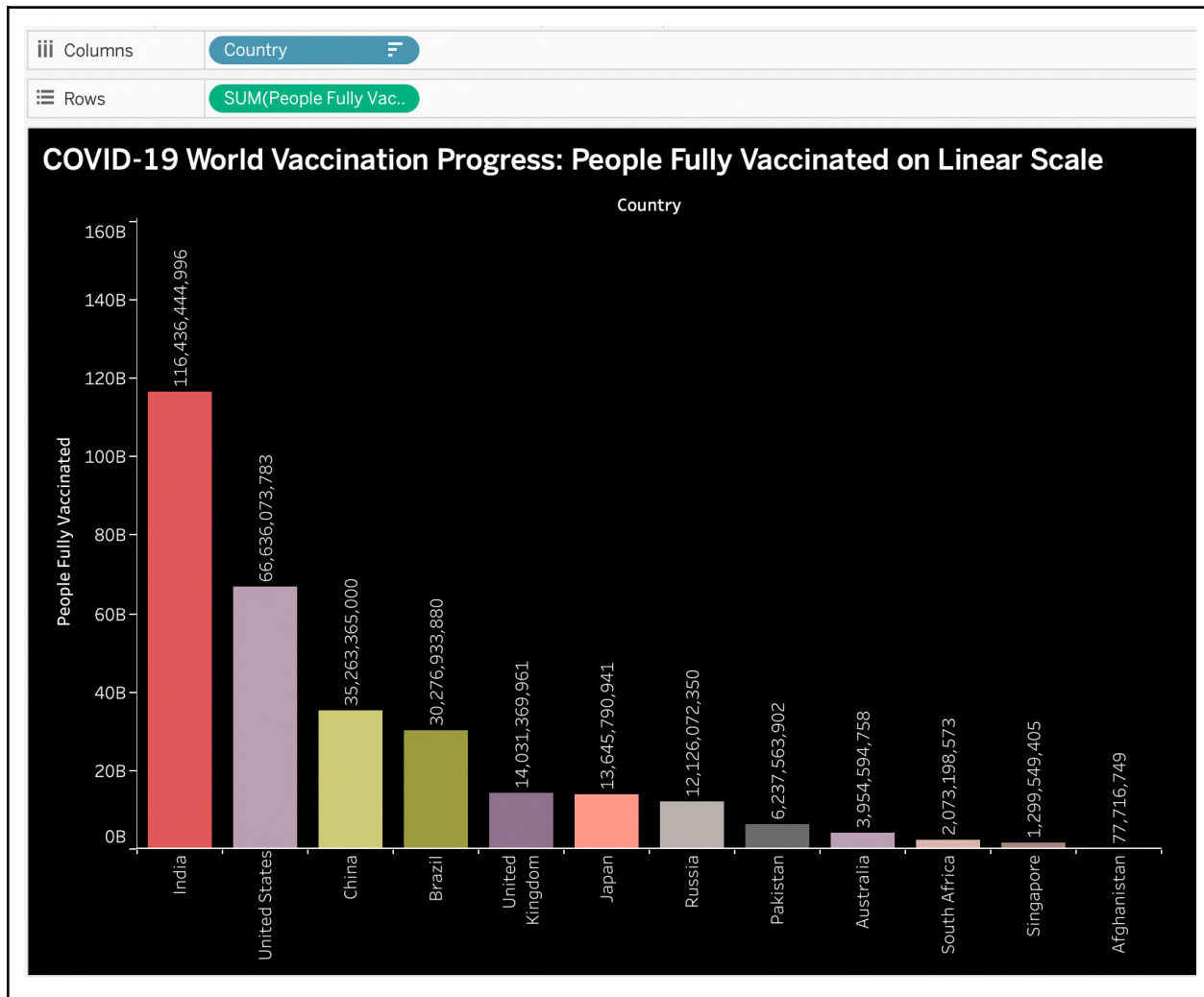
x-axis	sepal length
y-axis	sepal width

- The above 2D Cartesian coordinate system is used to plot the sepal length vs. sepal width measurements for each iris flower.
- By plotting these measurements in a 2D space, we can visualize the relationships between them and by using color to differentiate between the different species of iris flowers, we can see how they are distributed in the 2D coordinate system and identify any patterns or trends that may exist between species.

Observations:

1. Iris-Setosa flowers tend to have shorter sepal lengths and wider sepal widths than the other two species.
2. Iris-Versicolor flowers tend to have intermediate values for both sepal length and sepal width.
3. Iris-Virginica flowers tend to have longer sepal lengths and narrower sepal widths than the other two species.

ii) Linear Scale:



x-axis	Countries
y-axis	Linear scale values for people who are fully vaccinated

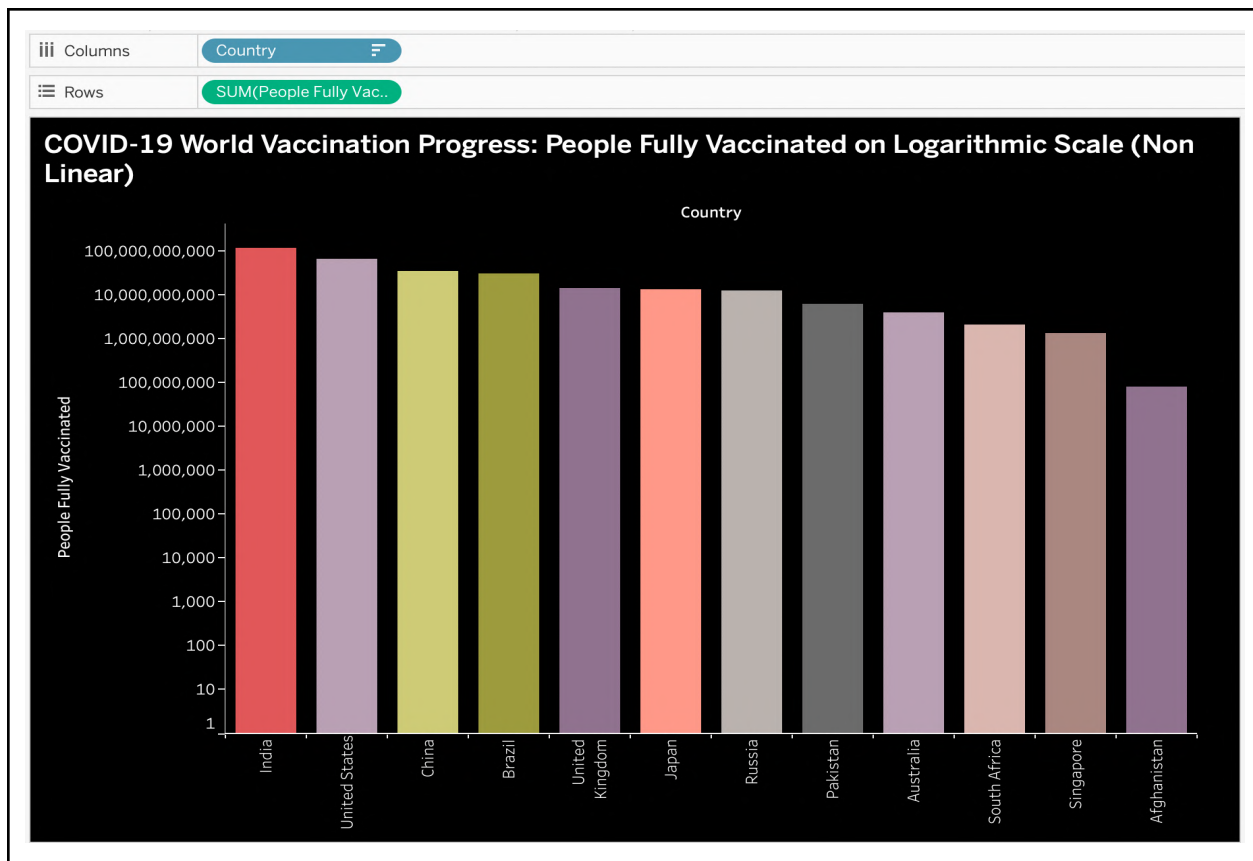
- Linear scale displays data that increases or decreases at a constant rate.
- In the above bar chart, a linear scale is used to display the number of people fully vaccinated, with each mark on the y-axis representing an equal increase in the number of people fully vaccinated.
- This allows us to easily compare the number of people fully vaccinated across different countries and identify any trends or patterns in the data.

Observations:

The above chart displays the number of people fully vaccinated against COVID-19 in various countries around the world.

1. India has the highest number of people fully vaccinated, followed by the United States and China.
2. There is a large variation in the number of people fully vaccinated across different countries.
3. The scale on the y-axis is linear, which means that each mark represents an equal increase in the number of people fully vaccinated.

iii) Non-Linear Scale:



x-axis	Countries
y-axis	Logarithmic scale values for people who are fully vaccinated

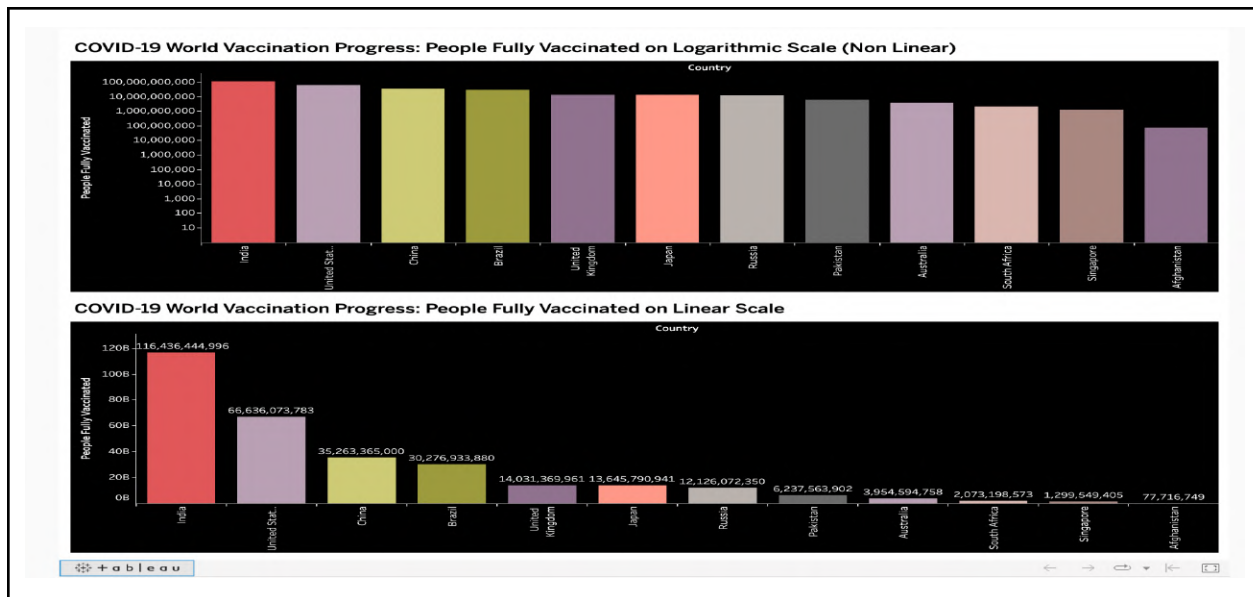
- Non-linear scales, such as logarithmic scales, are used when the data has a wide range of values that vary significantly in magnitude.
- In the above bar chart, a logarithmic scale is used to display the number of people fully vaccinated, with each tick mark on the y-axis representing a multiplication of the previous tick mark by a constant factor of 10.
- This allows us to better visualize the large variation in the number of people fully vaccinated across different countries, as smaller values are displayed more clearly and accurately (**which was not the case with linear scale[Afghanistan]**).

Observations:

The above chart displays the number of people fully vaccinated against COVID-19 in various countries around the world.

1. India has the highest number of people fully vaccinated, followed by the US and China.
2. There is a large variation in the number of people fully vaccinated across different countries.
3. The scale on the y-axis is logarithmic, which means that each tick mark represents a multiplication of the previous tick mark by a constant factor (usually 10).

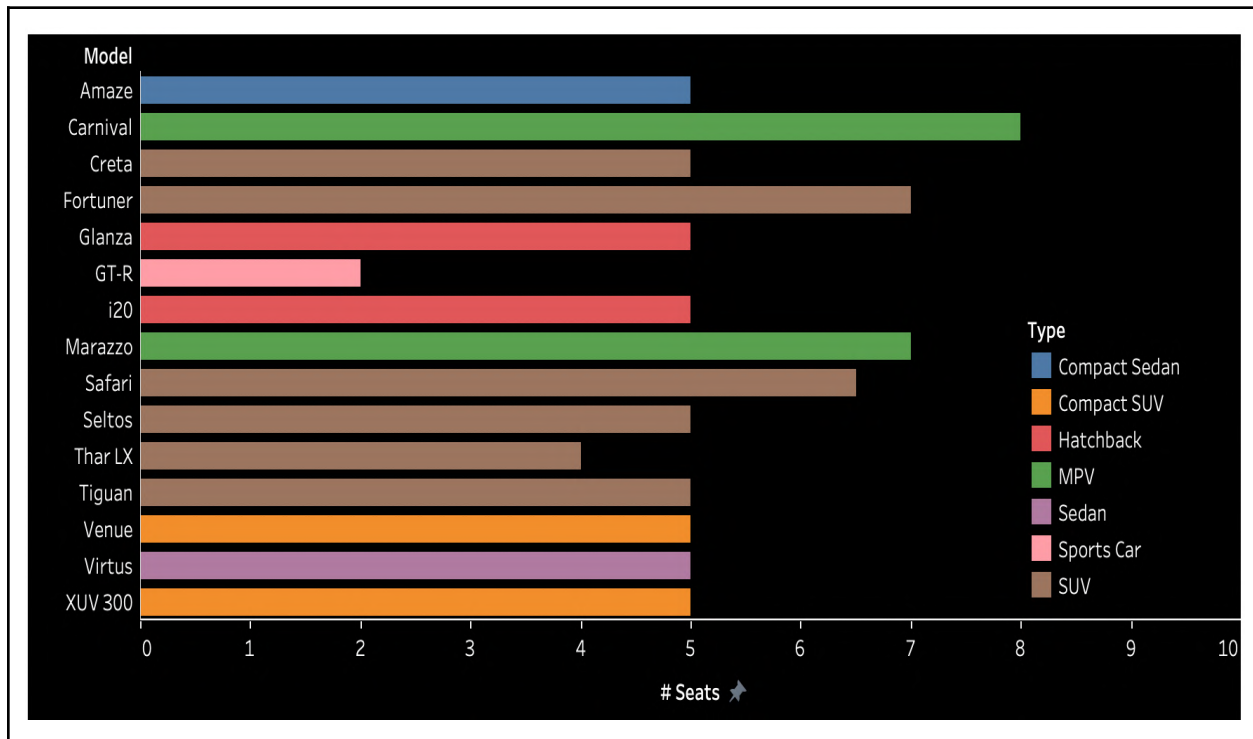
Dashboard: Non-Linear vs Linear



=====

Chapter 4: Color Scales

i) Qualitative Color scales:

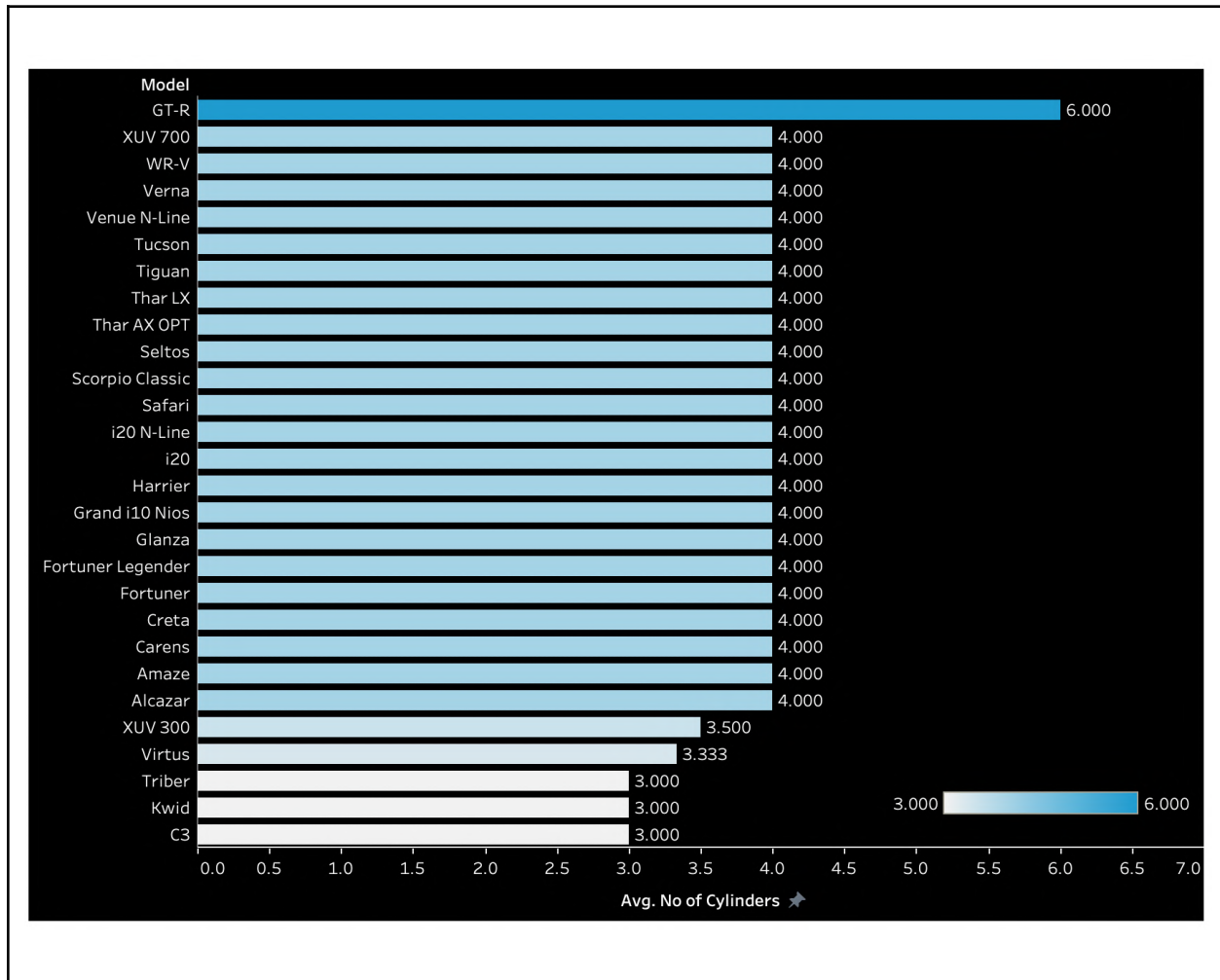


The above plot shows the number of seats of different car models, grouped by their type (e.g., sedan, SUV, etc.). Qualitative color scales are used to differentiate between the different types of cars, making it easy to identify which type each model belongs to.

Observations:

- Most car models have 5 seats, with few exceptions having 6 or more seats.
- Some SUVs tend to have 4 seats, while other Sedan, mid-size and compact SUVs have more seats on average.
- Sports cars tend to have fewer seats compared to other types of cars.

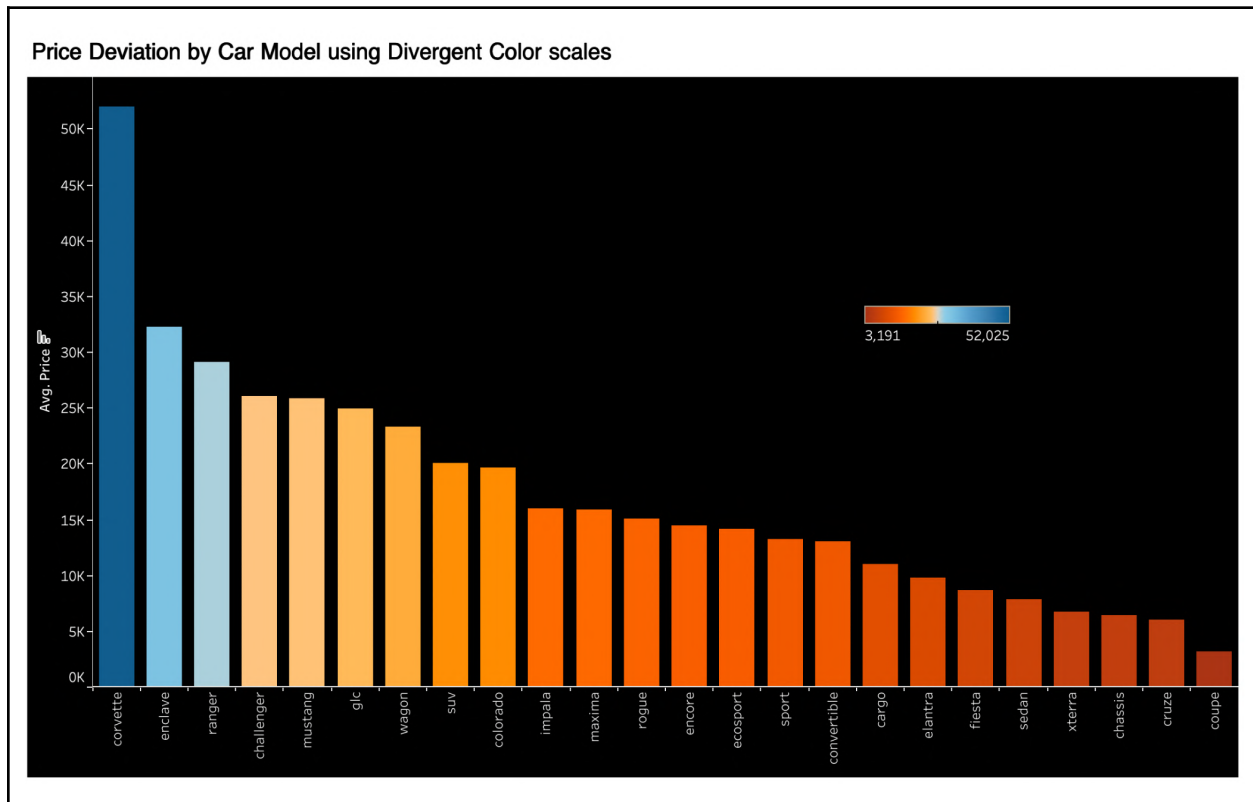
ii) Sequential Color scales:



Observations:

1. The use of a sequential color scale in the above plot helps to show the relative differences in the number of cylinders between different car models, making it easier to see which models have more or fewer cylinders compared to others.
2. Sports cars tend to have more cylinders on average compared to other cars which are represented using a dark color.
3. The most common and least number of cylinders for car models in the dataset is 3-4, which is commonly used in smaller and more economical cars and it is represented using a light color.

iii) Divergent and Accent Color scales:

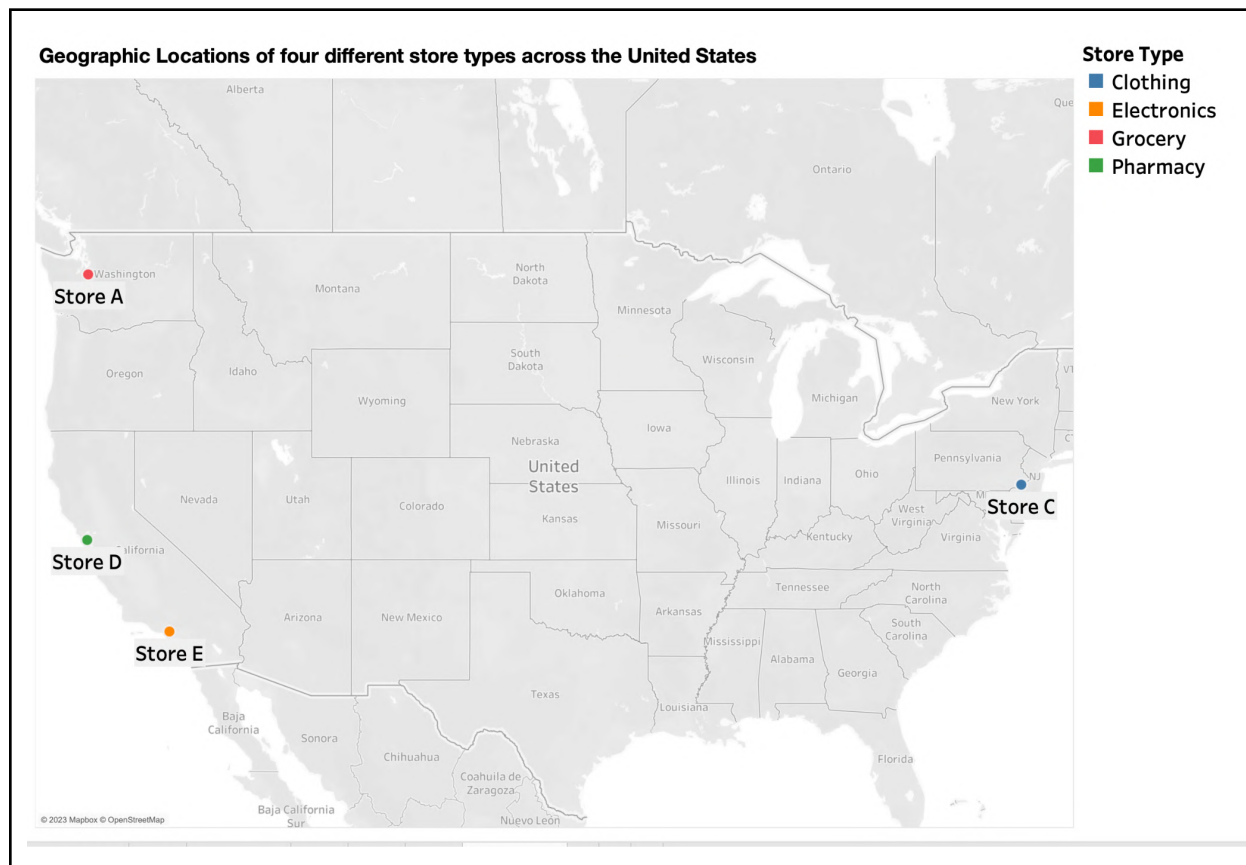


Observations:

1. In the above chart, Some car models have prices that are significantly above or below the mean value, as shown by the size and color of the bars.
2. Using a diverging color scale in the above plot helps to highlight the relative deviations in price for different car models, making it easier to see which models are priced higher or lower compared to the mean value.
3. The higher and lower average prices have similar levels of brightness and intensity which clearly deviates from the central average price and it also helps to **highlight the highest priced car model which is corvette.**

=====

Chapter 5: Geospatial data using Tableau



Observations:

1. In the above Map, four stores are located in different cities across the United States, including Seattle, New York, San Francisco, and Los Angeles that belong to a specific owner.
2. The store types include Grocery, Clothing, Pharmacy, and Electronics.
3. The latitude and longitude coordinates provide the geographic location of each store.

=====