DATA VISUALIZATION

ASSIGNMENT - 9

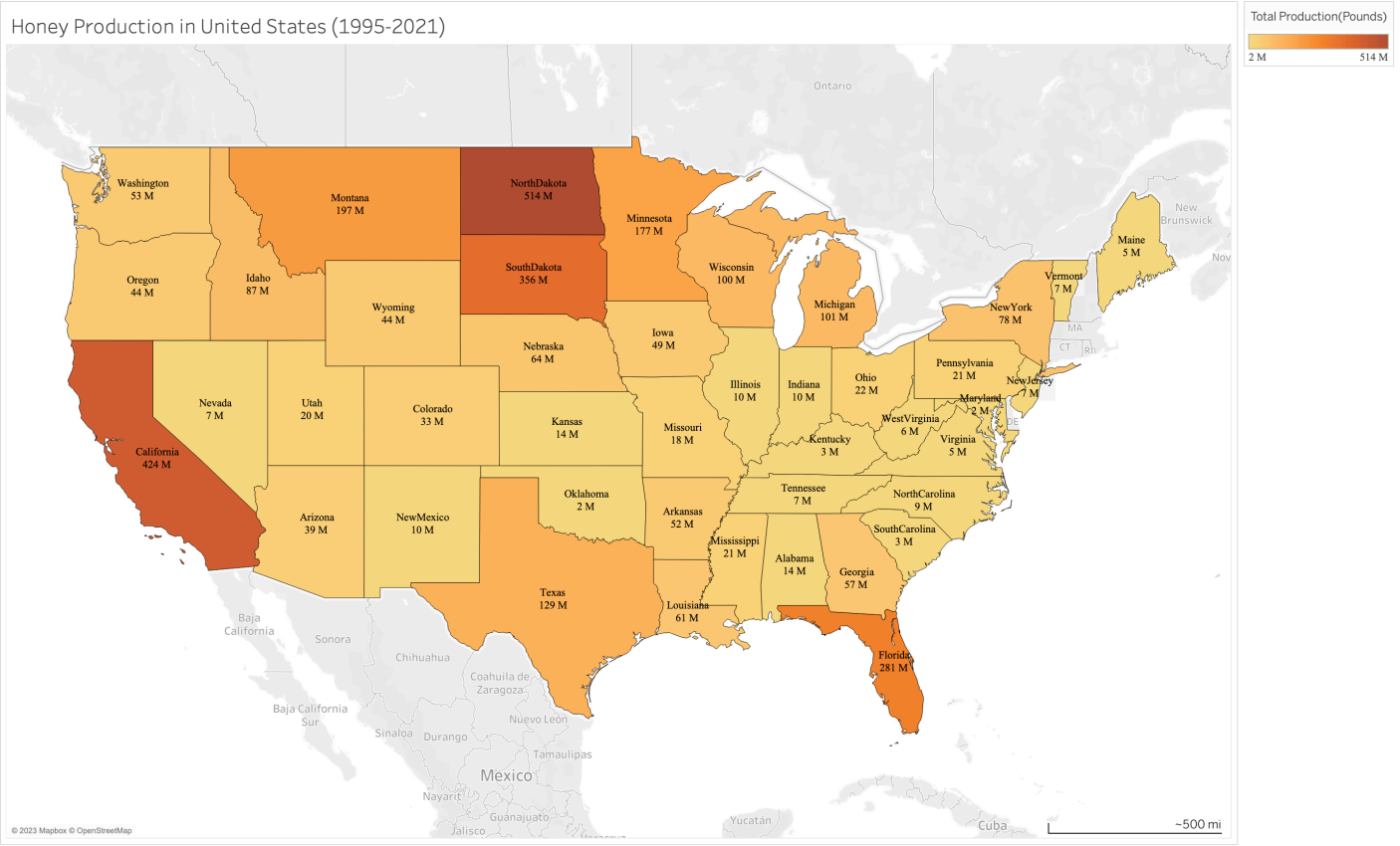
- Rohith Reddy Vangala (016762109)

Question :

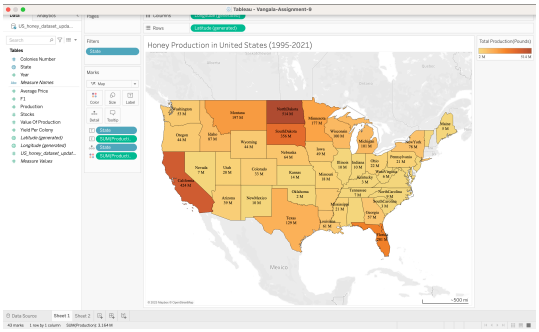
Using the data and tool of your choice, create and submit two charts:
- A geospatial chart
- A time-series chart

Geospatial chart :

- For creating a geospatial chart I have used Tableau and the dataset is US honey production (1995-2021) from kaggle (https://www.kaggle.com/datasets/mohitpoudel/us-honey-production-19952021).



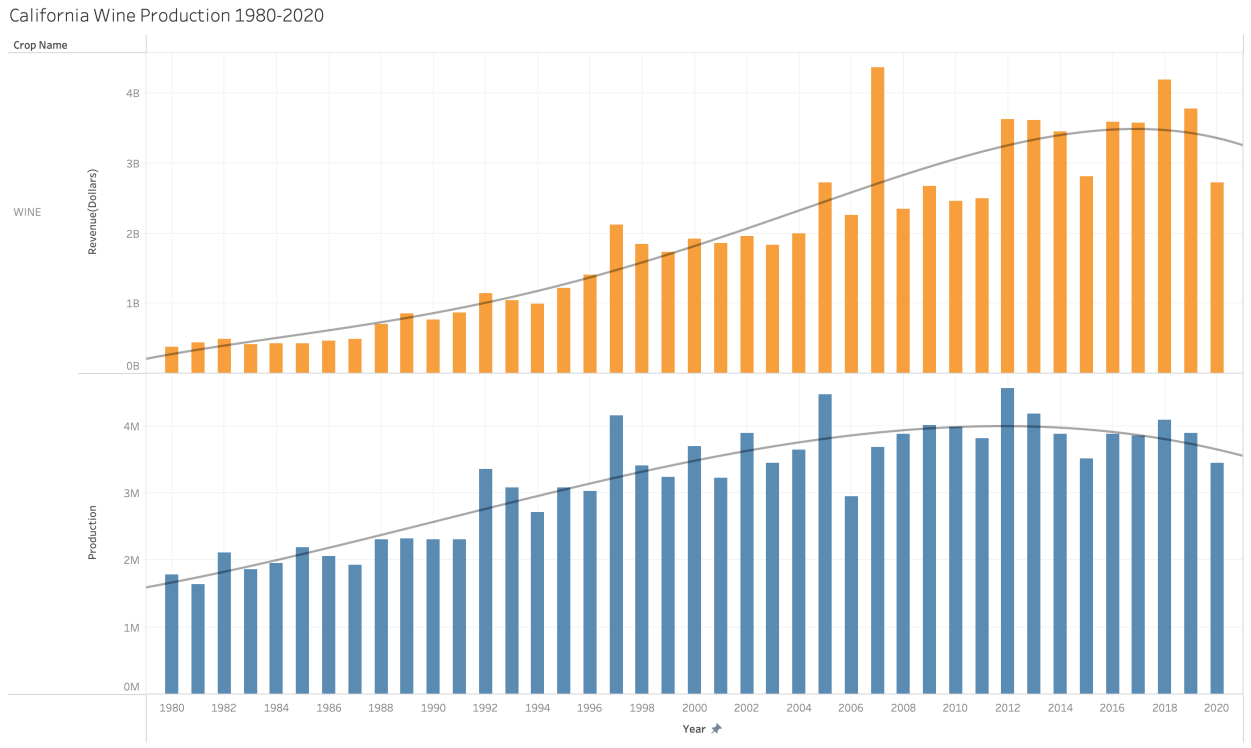
- The above chart shows the honey production (in pounds) in united states in the year 1995-2021.
- The use of a geospatial chart is appropriate for this type of data, as it allows viewers to quickly and easily understand the differences in honey production between regions and identify any patterns or trends in the data.
- The color palette used in the chart has been chosen with consideration for people who are color blind. By using colors that are distinguishable for both color blind and non-color blind viewers, the chart is more accessible and can be understood by a wider audience.
- The chart follows the Gestalt principle of similarity by using consistent colors to represent honey production in different regions. This helps viewers to quickly and easily understand the patterns in the data.
- The use of color to highlight the differences in honey production between regions also follows the Gestalt principle of foreground. By using colors that are distinct from one another, the chart helps to emphasize the differences in the data and draw the viewer's attention to the most important information.



- The chart also follows the Gestalt principle of closure, as the use of color and shading to fill in the regions creates a sense of completeness and makes it easier for the viewer to interpret the data.
- The use of the #,##0,, "M" format to change the revenue values to millions of dollars is a useful technique that helps to make the chart easier to read and understand. By presenting the data in this way, the chart avoids cluttering the visualization with many zeros, which can be confusing and difficult to interpret. The use of the "M" unit also helps to make it clear to viewers that the revenue values are in millions, which reduces the risk of misunderstandings or misinterpretations of the data.
- The use of color to highlight the differences in honey production is effective at conveying the information, as it is a visually striking and attention-grabbing way to present the data.
- The chart is well-organized and easy to read, with clear labels and a logical layout that helps to guide the viewer's eye through the data.
- Overall, the chart is a clear and effective way to present the honey production data, and the use of color, font, and layout has been carefully considered to ensure that the information is accessible and easily understood by all viewers.

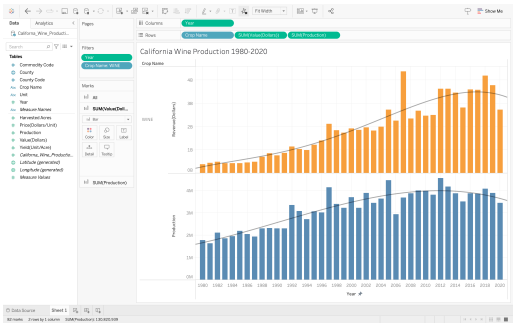
Time-series chart :

- For creating a time series chart I have used Tableau and the dateset is California wine production (1980-2020) from kaggle (https://www.kaggle.com/datasets/jarredpriester/california-wine-production-19802020).
- The below chart shows the California wine production and revenue (in dollars) in the year 1980-2020.



- The use of a dual bar chart to display both wine production and revenue in dollars is an effective way to compare the two variables within the same chart. This layout allows viewers to easily compare the relative magnitudes of production and revenue over time, and to identify any patterns or trends that may exist between the two variables.
- The choice of colors used in the chart is suitable for color-blind viewers, making the chart accessible to a wider audience.
- The chart follows the principle of proximity, grouping related data points closely together and separating them from unrelated data.
- The chart follows the principle of similarity, using consistent colors and shapes to represent similar data points.

- The chart follows the principle of continuity, using a smooth trendline to represent the overall trend of the data.
- The chart follows the principle of foreground and background, with the data points and trendline placed in the foreground and the background kept neutral to avoid distractions.
- The trendline on the production and revenue bars reveals that there was a significant reduction in both variables in 2020. This can be attributed to the lockdowns and restrictions that were put in place in response to the COVID-19 pandemic, which had a profound impact on the wine industry in California. The inclusion of this insight adds a valuable layer of context to the visualization, helping viewers to understand the factors that may have influenced the trends shown in the data.



THE END