



**International Journal of Advanced Research in
Education and Technology (IJARETY)**

Volume 11, Issue 4, July-August 2024

Impact Factor: 7.394



Data Analytics and Visualization using Tableau on Weather Dataset

Mohammed Mohiuddin Khan, Mohammed Munawar, Syed Abdus Sami Mujtaba, Dr Diana Moses

Bachelor of Engineering Student, Department of Artificial Intelligence and Data Science, Methodist College of Engineering and Technology, Hyderabad, India

Guide, Department of Artificial Intelligence and Data Science, Methodist College of Engineering and Technology, Hyderabad, India

ABSTRACT: The (“Weather Dataset”) is a comprehensive dataset containing weather-related information for various cities around the world. The dataset comprises 100 rows and 7 columns, each representing different attributes of weather data.

This dataset provides a valuable resource for analyzing weather patterns, understanding geographical climate variations, and conducting various meteorological studies.

I. INTRODUCTION

- **Geographical Data Analysis:** The geographical data analysis focused on latitude and longitude ranges, providing insights into the spatial distribution of the dataset
- **Temperature Data Analysis:** The temperature data analysis aimed to determine the average, minimum, and maximum temperatures, offering a snapshot of the thermal conditions within the dataset.
- **Wind Speed Data Analysis:** The wind speed data analysis calculated the average, minimum, and maximum wind speeds, contributing to an understanding of the wind dynamics present in the dataset.

II. DATASET DESCRIPTION

This dataset provides temperature data for the top 100 cities worldwide as of April 28th, 2024. It consists of seven columns, each offering specific information to understand the weather conditions in different cities. The dataset includes the following columns:

1.City:

The name of the city in English language. This column provides the official or commonly used name of each city. City names are essential for identifying the geographical location associated with the weather data.

2.Temperature(°C):

The temperature recorded for each city on April 28th, 2024. Temperature is measured in degrees Celsius (°C) and represents the average atmospheric temperature at the specified location on the given date. It is an essential parameter in weather analysis, indicating the warmth or coldness of the air.

3.WindSpeed(m/s):

The average wind speed measured at each location in meters per second (m/s). Wind speed is a measure of the movement of air molecules and is an important factor in weather prediction and analysis.

4.Latitude(°):

The latitude coordinates of each city, expressed in degrees. Latitude measures the north-south position of a location on the Earth's surface relative to the equator. Positive values represent locations in the northern hemisphere, while negative

values represent locations in the southern hemisphere. Latitude ranges from -90° (South Pole) to $+90^{\circ}$ (North Pole), with 0° at the equator.

5.Longitude($^{\circ}$):

The longitude coordinates of each city, expressed in degrees. Longitude measures the east-west position of a location on the Earth's surface relative to the Prime Meridian. Positive values represent locations east of the Prime Meridian, while negative values represent locations west of it. Longitude ranges from -180° (West) to $+180^{\circ}$ (East), with 0° at the Prime Meridian.

6.Description:

This column provides details about the prevailing weather conditions in each city on April 28th, 2024. It includes descriptions such as clear sky, scattered clouds, broken clouds, overcast clouds, and few clouds, indicating the level of cloud cover and sky visibility at the time of observation.

7.Country:

This column contains the name of the country to which each city belongs. It provides information about the city with its respective country.

III. INTRODUCTION TO TABLEAU

Tableau is a powerful tool used for data analysis and visualization. It allows the creation of amazing and interactive visualization and that too without coding. Tableau is very famous as it can take in data and produce the required data visualization output in a very short time. Basically, it can elevate your data into insights that can be used to drive your action in the future.

Tableau is a visual analytics platform that is revolutionizing the way we use data to solve problems by enabling individuals and organizations to make the most of their data.

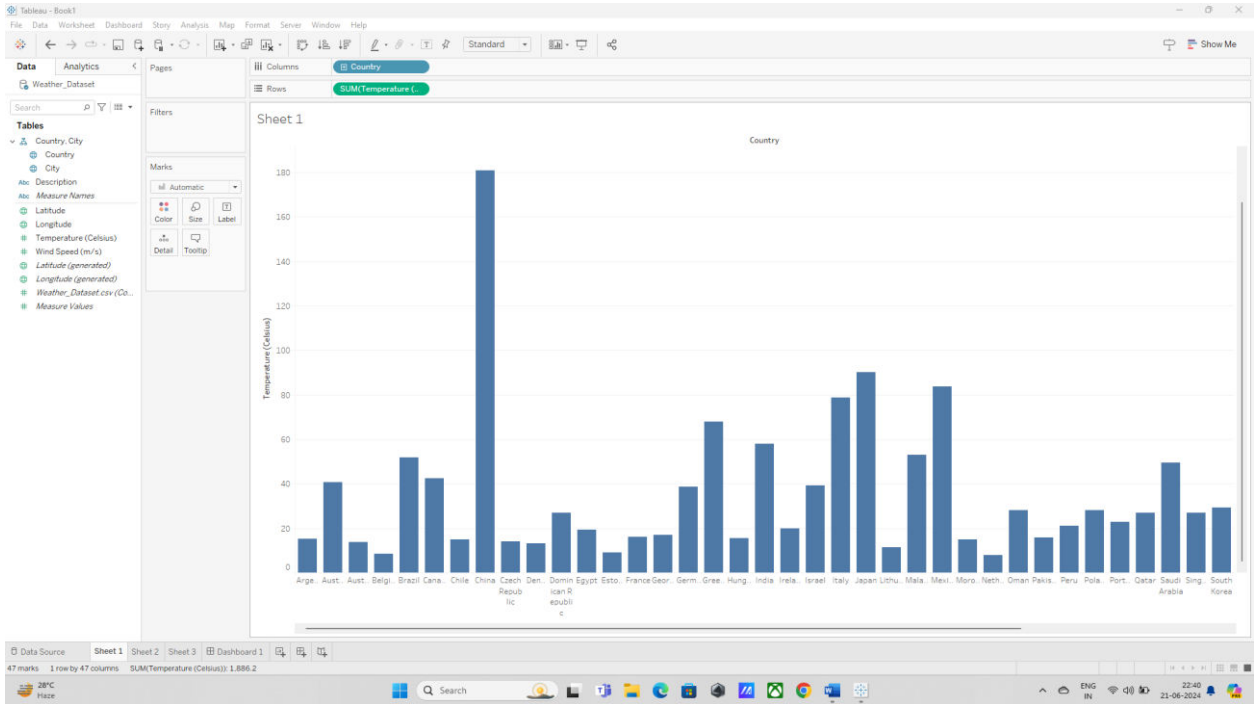
Tableau is a great data visualization and business intelligence application that can be used to report and analyze massive amounts of data. Salesforce purchased Tableau in June 2019, an American firm founded in 2003. It enables users to build various charts, graphs, maps, dashboards, and stories for visualizing and analyzing data in order to aid in business choices. Tableau offers several unique and fascinating features that make it one of the most popular business intelligence (BI) applications.

Tableau is the fastest and most powerful visualization tool. It is very easy to use. There are no complex formulas like Excel and other visualization tools. It provides the features like cleaning, organizing, and visualizing data, it easier to create interactive visual analytics in the form of dashboards. These dashboards make it easier for non-technical analysts and end-users to convert data into understandable ones.

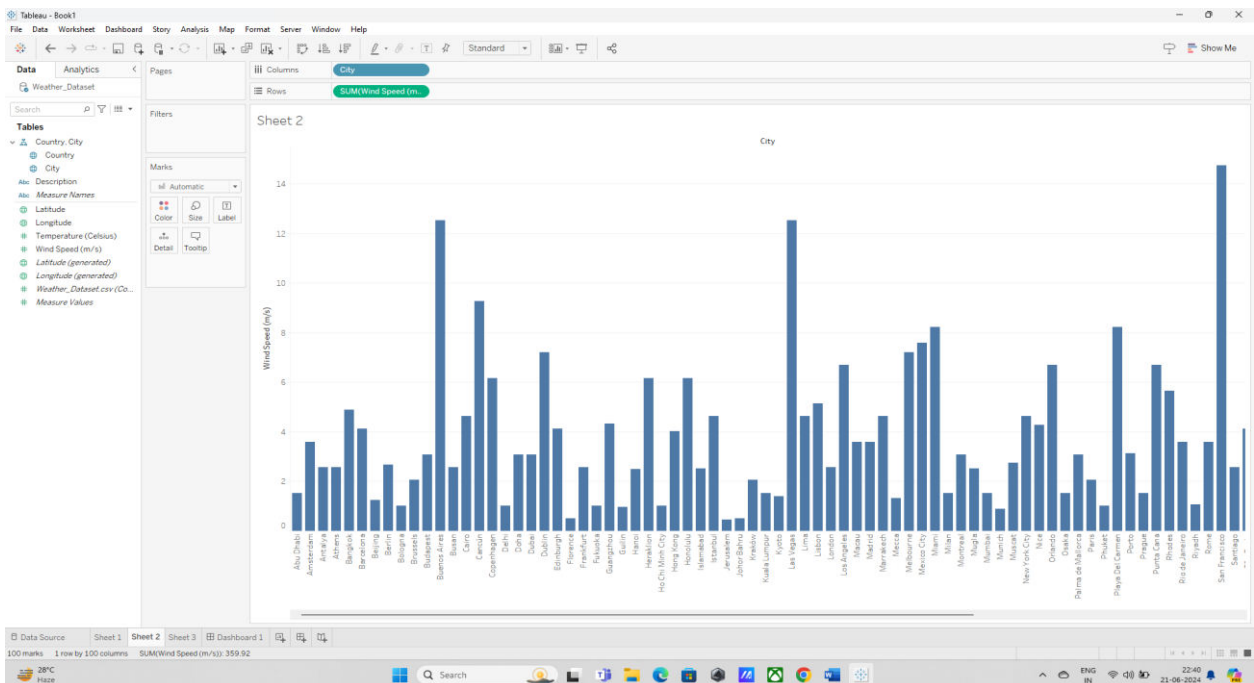
Tableau Features

- Tableau supports powerful data discovery and exploration that enables users to answer important questions in seconds
- No prior programming knowledge is needed; users without relevant experience can start immediately with creating visualizations using Tableau
- It can connect to several data sources that other BI tools do not support. Tableau enables users to create reports by joining and blending different datasets
- Tableau Server supports a centralized location to manage all published data sources within an organization

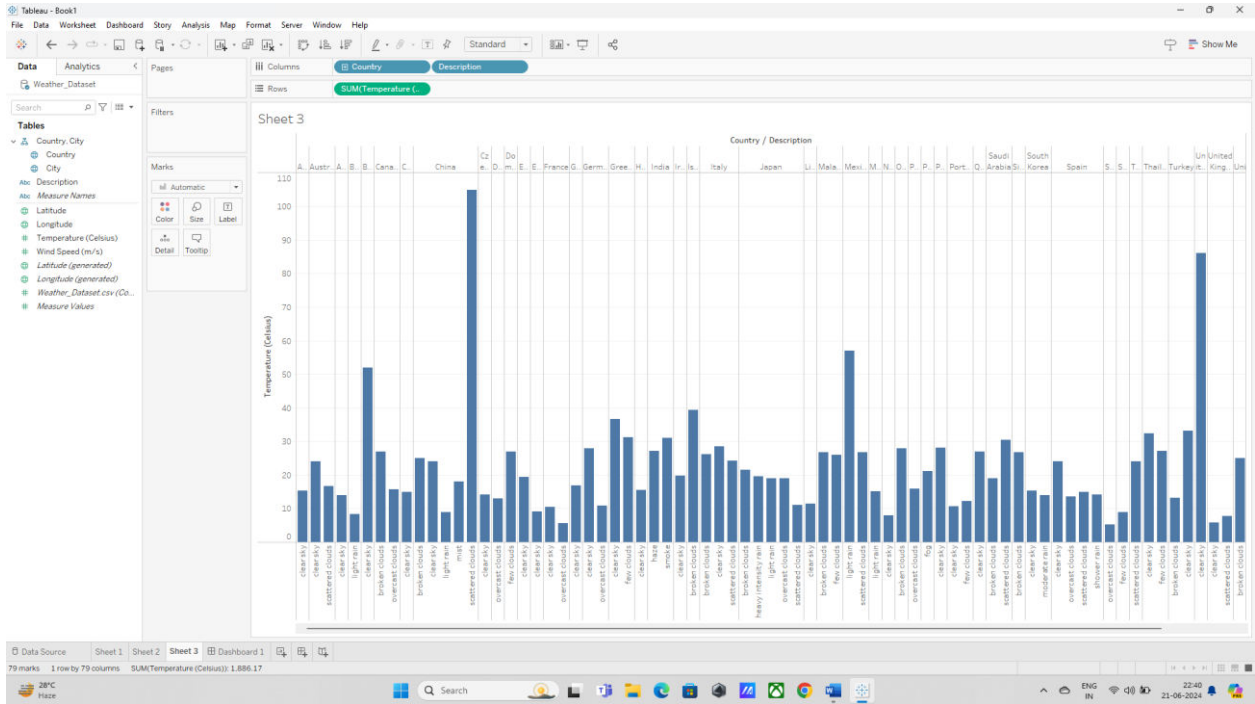
IV. DATA ANALYSIS



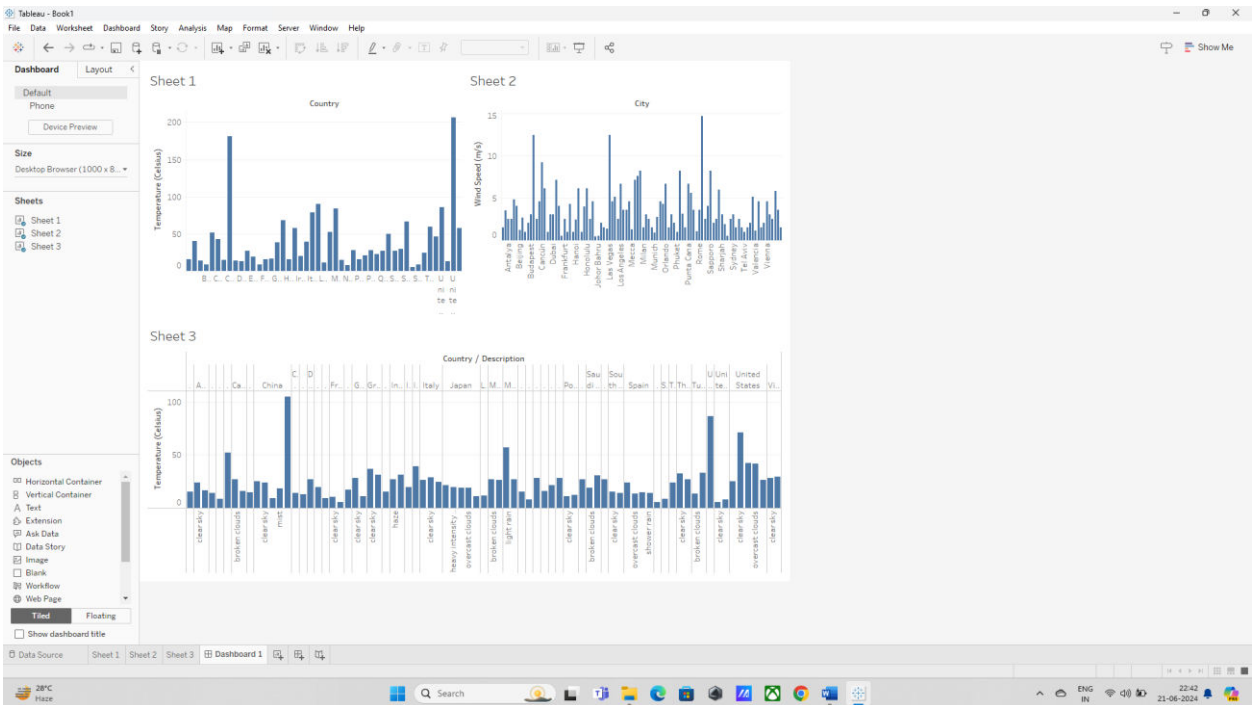
In Sheet-1, We have added Country to columns and Temperature(Celsius) to rows. We have chosen a Stacked Bars chart.



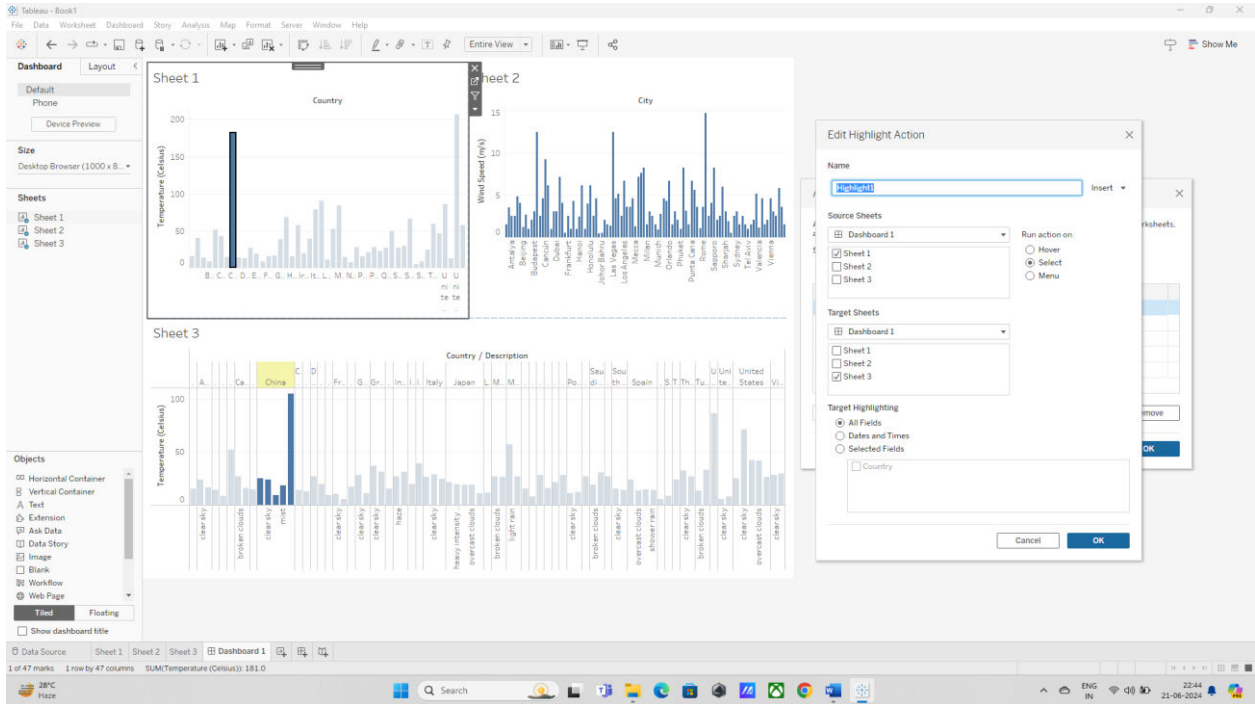
In Sheet-2, We have added City to columns and Wind Speed(m/s) to rows. We have Chosen a Stacked Bars chart.



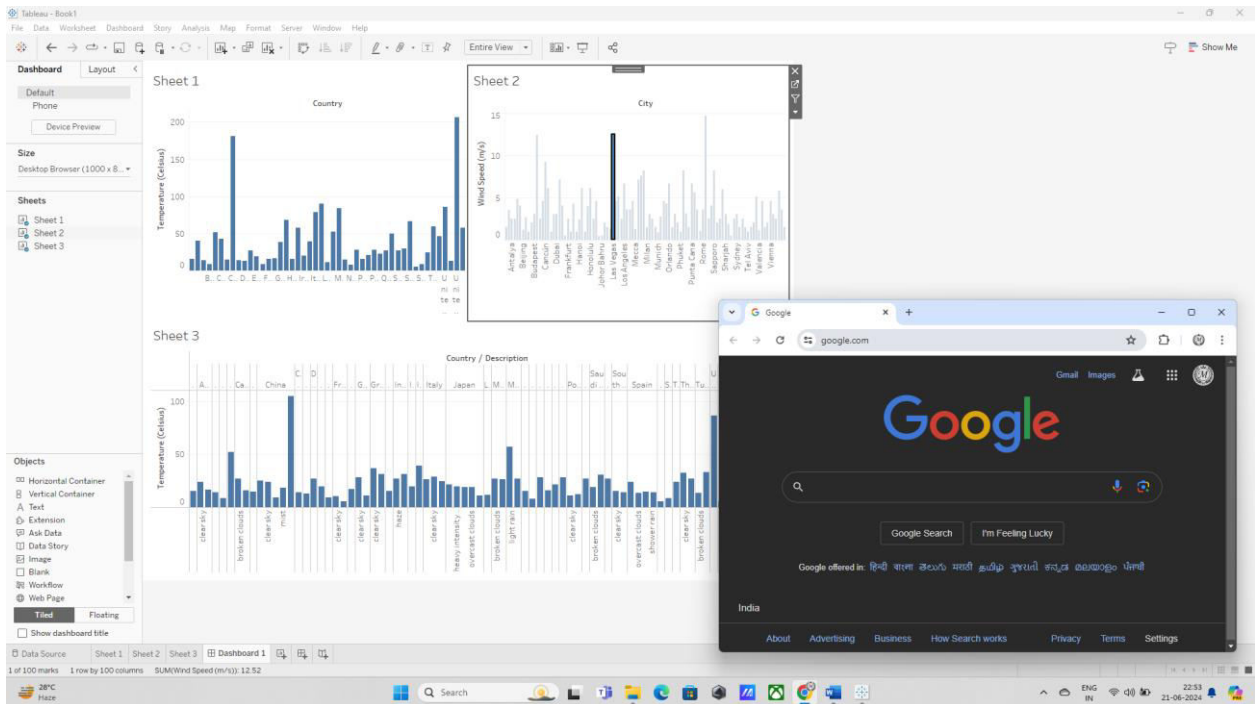
In Sheet-3, We have added Country and Description to columns and Temperature (Celsius) to rows we have Chosen a Stacked Bars chart.



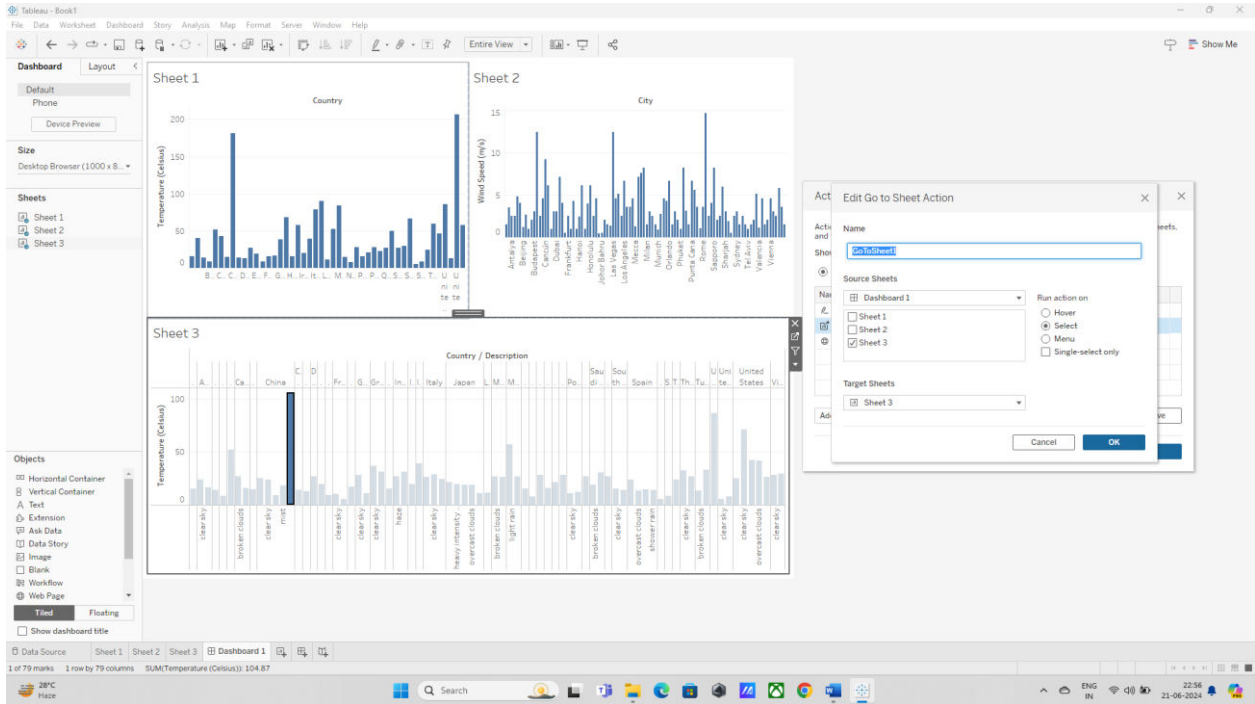
We have Created a Dashboard by adding all three sheets to the new dashboard.



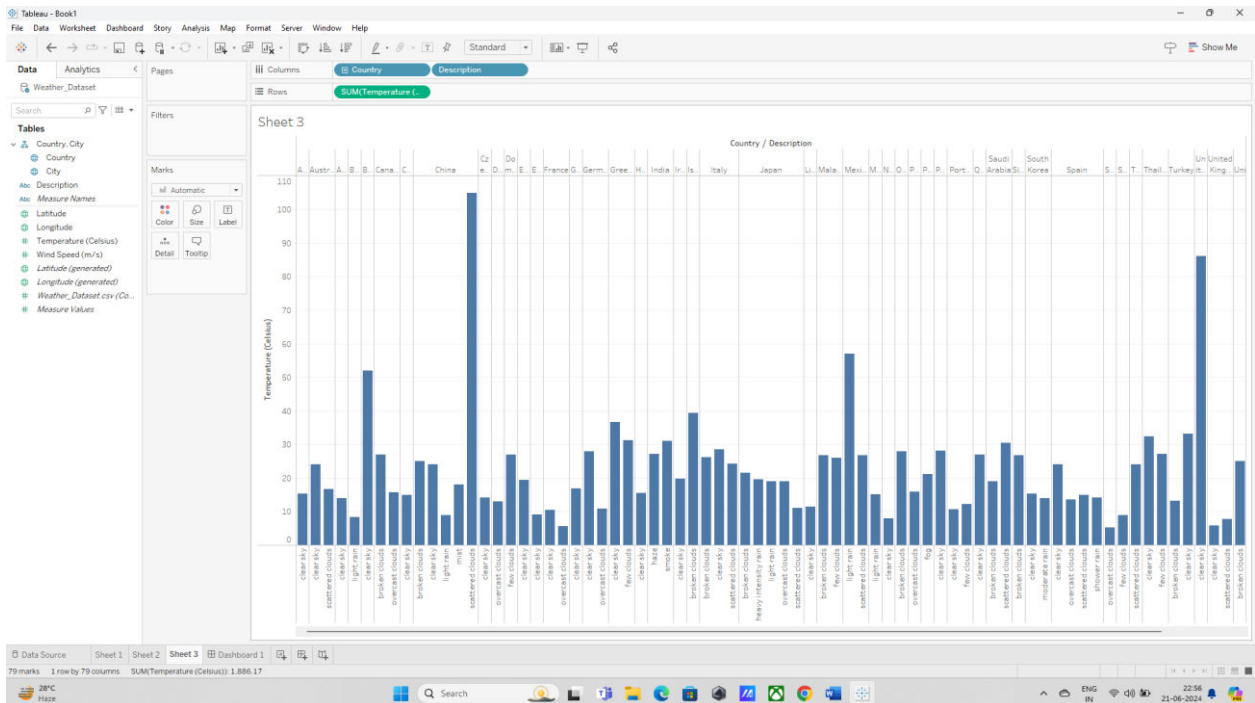
We have added the Action (Highlight) to Sheet-1 and Sheet-3, Action runs when we select any bar in Sheet-1.

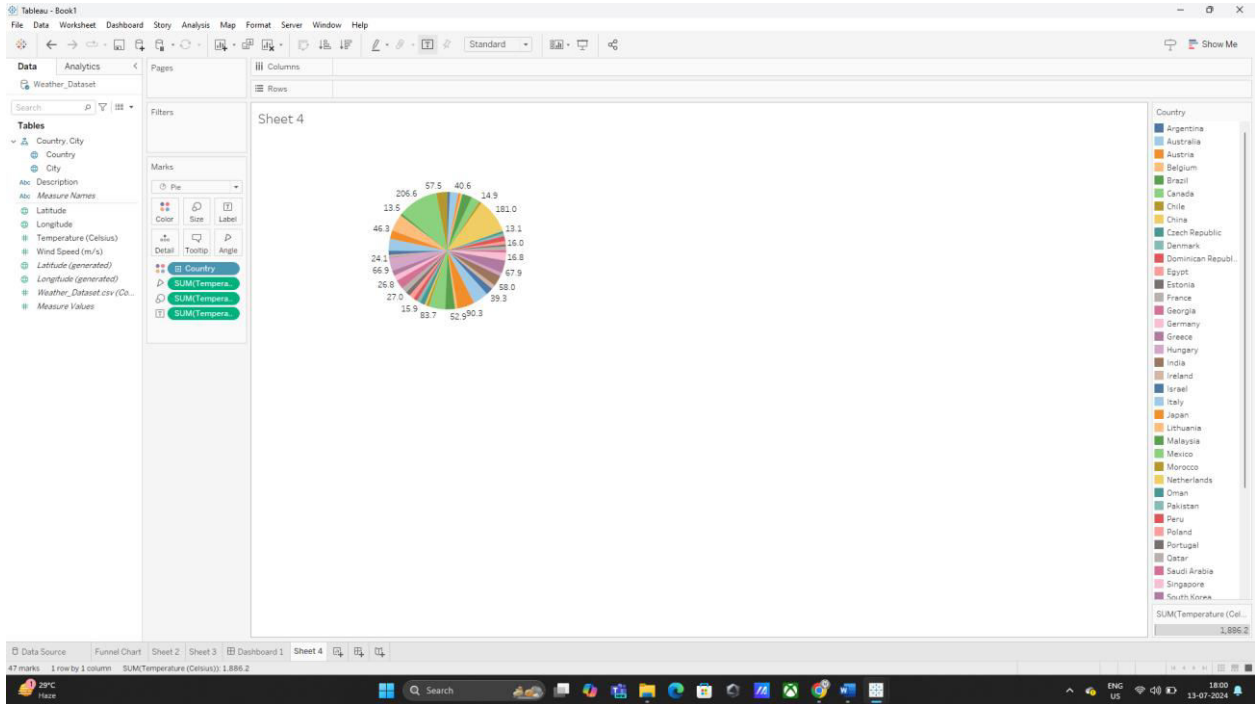


We have added the Action (URL) to Sheet-2, Action runs when we select any bar in Sheet-2.

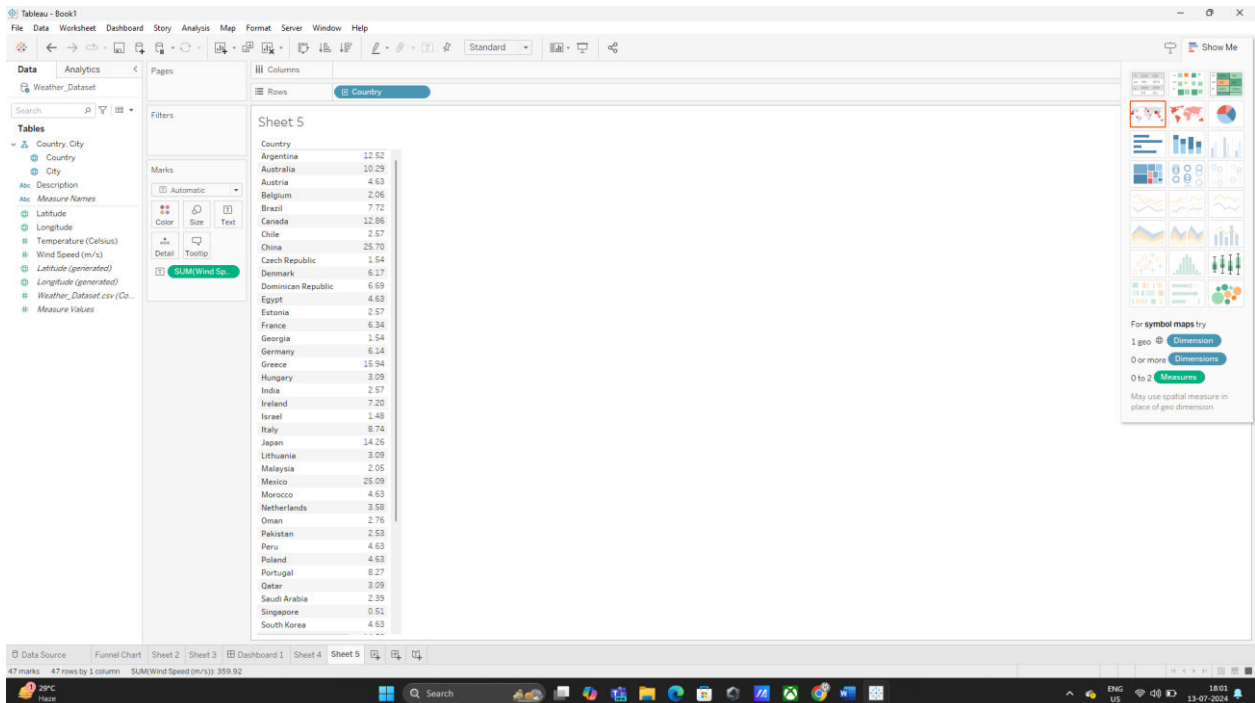


We have added an Action (Go ToSheet) on Sheet-3, If we click on any bar or we select a bar we go to a target sheet-3. As shown in the below Image.

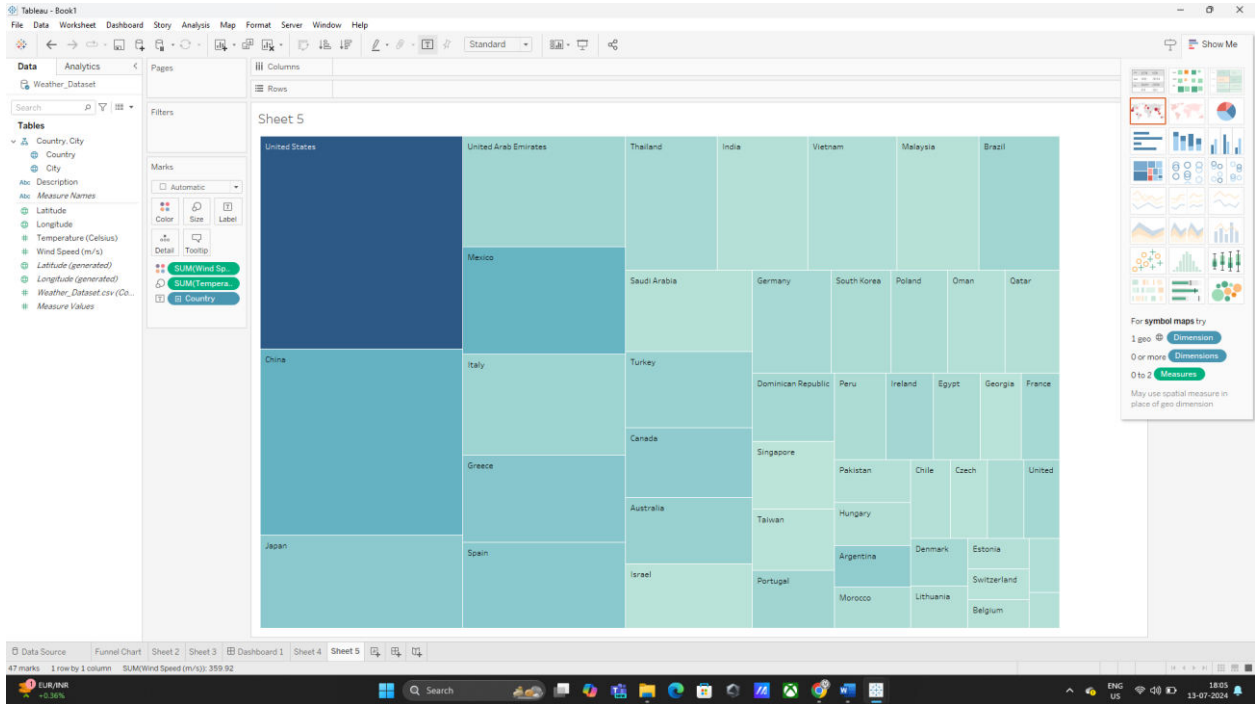




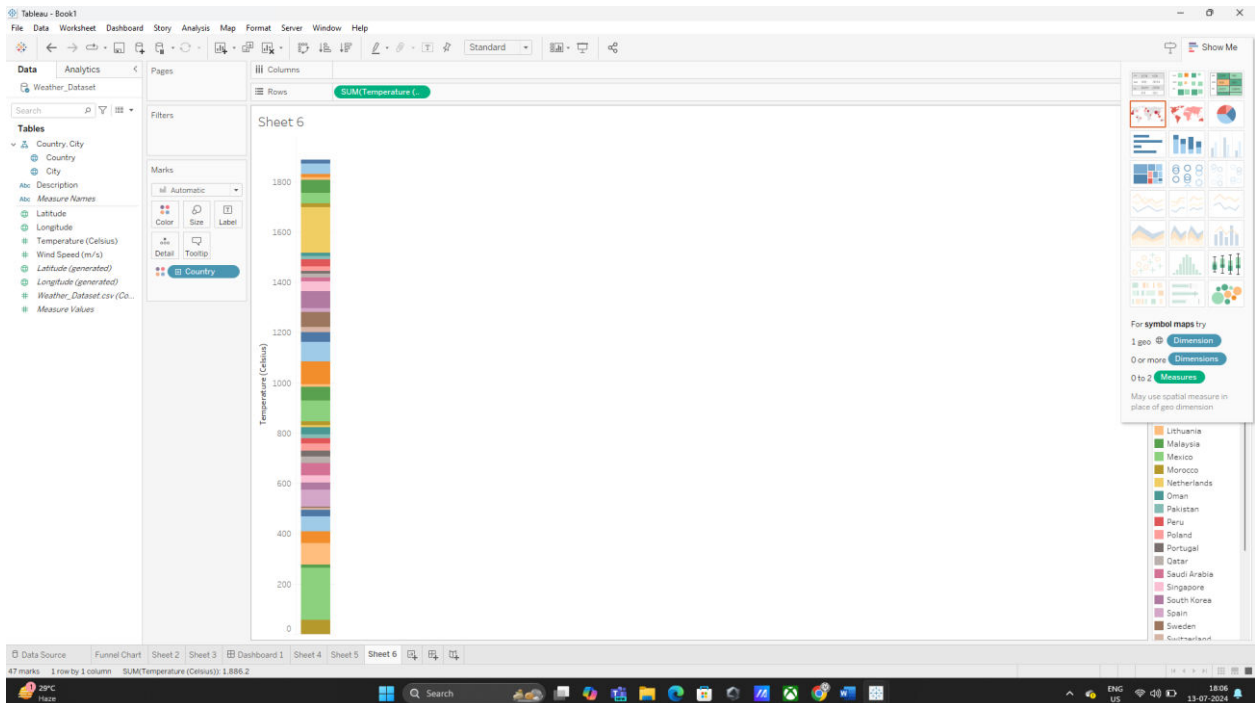
We have created a pie chart by adding Country to Marks (Color) and Temperature (Celsius) to angle, Marks (Size) and (Label).



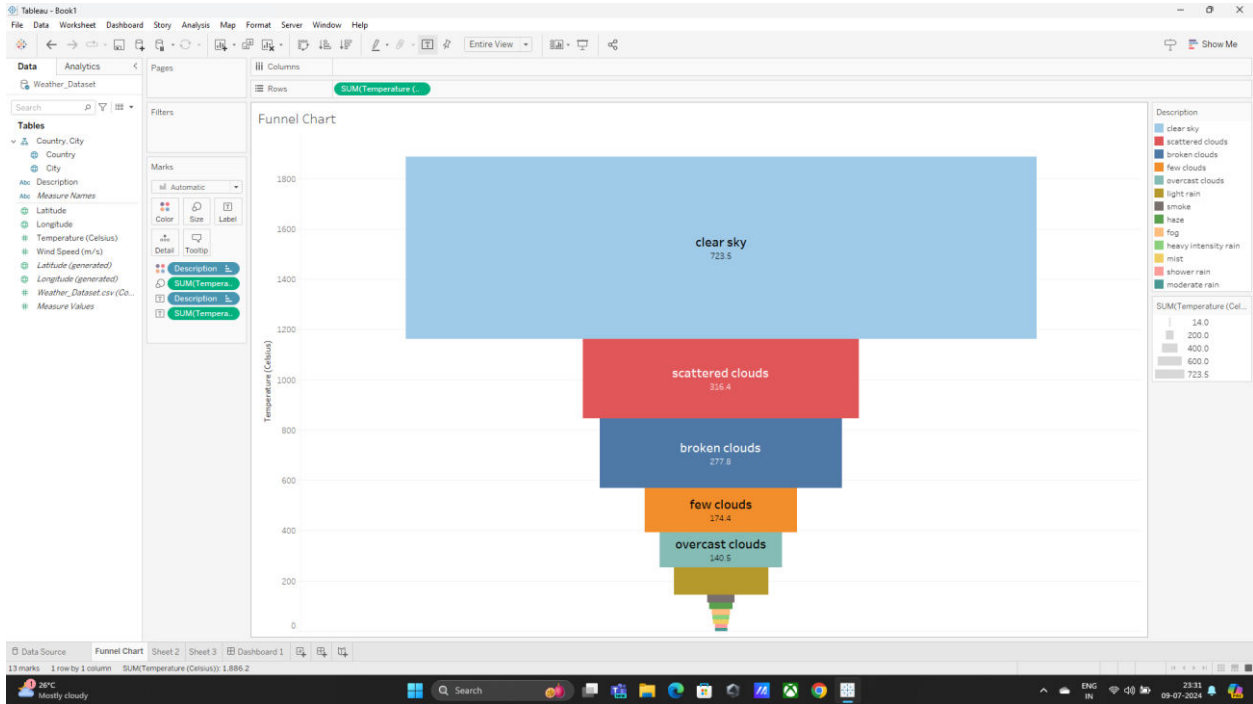
We have created a Text Table by adding Country to rows and Wind Speed(m/s) to Marks (Label).



We have created a Tree Maps by adding Wind Speed(m/s) to Marks (Color), Temperature (Celsius) to Marks (Size) and Country to Marks (Label).



We have Created a Stacked Bars by adding Temperature (Celsius) to rows and Country to Marks(color).



We have created a Funnel Chart by adding Temperature (Celsius) to rows, Description to Marks (Color), Temperature (Celsius) to Marks (Size), Temperature (Celsius) and Description to Marks (Label).

V. RESULTS & DISCUSSIONS

The dataset provides a comprehensive overview of weather conditions across different cities. The temperature and wind speed data show significant variability, indicating diverse climatic conditions. The geographical coordinates (latitude and longitude) cover a wide range, suggesting a global dataset.

VI. CONCLUSION

- **Temperature Distribution:**

The **mean temperature** across various cities is **18.86°C**, with a **standard deviation** of **7.02°C**, indicating moderate variability in temperature.

 - The **minimum recorded temperature** is **5.29°C**, and the **maximum** is **32.35°C**.
 - The **median temperature** (50% percentile) is **18.005°C**, suggesting that half of the cities have temperatures below this value and half above.
- **Wind Speed Distribution:**
 - The **mean wind speed** is **3.60 m/s** with a higher **standard deviation** of **2.66 m/s**, pointing to a wider spread in wind speed measurements.
 - The **lowest wind speed** is **0.45 m/s**, and the **highest** is **14.75 m/s**.
 - The **median wind speed** (50% percentile) is **2.925 m/s**.
- **Temperature and Wind Speed Relationship:**
 - The dataset does not provide a direct correlation value between temperature and wind speed. Additional analysis would be required to determine the relationship.

REFERENCES

1. COBE SST
Ishii, M., A. Shouji, S. Sugimoto, and T. Matsumoto (2005) Objective Analyses of Sea-Surface Temperature and Marine Meteorological Variables for the 20th Century using ICOADS and the Kobe Collection. *Int. J. Climatol.*, 25, 865-879.
2. COBE2 SST
Hirahara, S., Ishii, M., and Y. Fukuda, (2014) Centennial-scale sea surface temperature analysis and its uncertainty. *J of Climate*, 27, 57-75. doi: 10.1175/JCLI-D-12-00837.1
3. GHCN CAMS
Fan, Y., and H. van den Dool (2008) A global monthly land surface air temperature analysis for 1948-present, *J. Geophys. Res.*, 113, D01103, doi:10.1029/2007JD008470
4. GPCC
Schneider et al (2017) Evaluating the Hydrological Cycle over Land Using the Newly-Corrected Precipitation Climatology from the Global Precipitation Climatology Centre (GPCC). *Atmosphere* 8(3), 52; doi:10.3390/atmos8030052
5. GPCP
Adler et al. (2016) An Update (Version 2.3) of the GPCP Monthly Analysis. (in Preparation). Huffman, G.J., R.F. Adler, P. Arkin, A. Chang, R. Ferraro, A. Gruber, J. Janowiak, A. McNab, B. Rudolf, U. Schneider, 1997: The Global Precipitation Climatology Project (GPCP) Combined Precipitation Dataset. *Bull. Amer. Meteor. Soc.*, 78(1), 5-20.
6. HadCRUT5 Analysis
Osborn, T.J., Jones, P.D., Lister, D.H., Morice, C.P., Simpson, I.R., Winn, J.P., Hogan, E., and Harris, I.C., (2021) Land surface air temperature variations across the globe updated to 2019: the CRUTEM5 dataset. *Journal of Geophysical Research: Atmospheres*. 126, e2019JD032352, doi: 10.1029/2019JD032352
7. HadISST
Rayner, N. A.; Parker, D. E.; Horton, E. B.; Folland, C. K.; Alexander, L. V.; Rowell, D. P.; Kent, E. C.; Kaplan, A. (2003) Global analyses of sea surface temperature, sea ice, and night marine air temperature since the late nineteenth century *J. Geophys. Res.* Vol. 108, No. D14, 4407 doi: 10.1029/2002JD002670
8. NASA GISTEMP
Lenssen, N., G. Schmidt, J. Hansen, M. Menne, A. Persin, R. Ruedy, and D. Zyss, (2019) Improvements in the GISTEMP uncertainty model. *J. Geophys. Res. Atmos.*, 124, no. 12, 6307-6326, doi: 10.1029/2018JD029522
9. NOAA ERSST V3b
Smith, T.M., R.W. Reynolds, T.C. Peterson, and J. Lawrimore, (2008) Improvements NOAAs Historical Merged Land–Ocean Temp Analysis (1880–2006). *Journal of Climate*, 21, 2283–2296
10. NOAA ERSST V5
Boyin Huang, Peter W. Thorne, Viva F. Banzon, Tim Boyer, Gennady Chepurin, Jay H. Lawrimore, Matthew J. Menne, Thomas M. Smith, Russell S. Vose, and Huai-Min Zhang (2017) NOAA Extended Reconstructed Sea Surface Temperature (ERSST), Version 5. subset used. NOAA National Centers for Environmental Information. doi: 10.7289/V5T72FNM date
11. NOAA OISST v2.1
Reynolds, R.W., N.A. Rayner, T.M. Smith, D.C. Stokes, and W. Wang, (2002) An improved in situ and satellite SST analysis for climate. *J. Climate*, 15, 1609-1625
12. NOAA PRECL
Chen, M., P. Xie, J. E. Janowiak, and P. A. Arkin (2002) Global Land Precipitation: A 50-yr Monthly Analysis Based on Gauge Observations, *J. of Hydrometeorology*, 3, 249-266
13. U Delaware 5.01
Willmott, C. J. and K. Matsuura (2001) Terrestrial Air Temperature and Precipitation: Monthly and Annual Time Series (1950 - 1999), http://climate.geog.udel.edu/~climate/html_pages/README.ghcn_ts2.html

International Journal of Advanced Research in Education and Technology

ISSN: 2394-2975

Impact Factor: 7.394