2024 | By: Moumita Biswas

Weather_analysis

[Capstone Project]

Contents

Introduction of data

The Weather Analysis Dataset provides a comprehensive collection of meteorological data encompassing various parameters such as temperature, humidity, wind speed, wind direction, air pressure, weather description, city attributes, and geographical coordinates. This dataset serves as a valuable resource for researchers, meteorologists, urban planners, and decision-makers seeking to understand weather patterns, climate variability, and their impacts on diverse sectors.

The dataset covers a wide range of temporal and spatial scales, offering insights into both short-term weather phenomena and long-term climate trends. With detailed information on temperature variations, humidity levels, wind behaviour, wind directions, air pressure fluctuations, weather descriptions, city attributes such as population density, elevation, and geographical distributions, users can conduct in-depth analyses to address a myriad of research questions and practical applications.

Key Features of the Dataset:

1. **City Attributes:** Additional attributes such as population density, elevation, and urbanization level provide context for understanding how weather affects different urban environments.
2. **Geographical Coordinates:** Latitude and longitude coordinates enable spatial mapping and visualization, aiding in the identification of regional weather patterns, climate zones, and geographical distributions of weather stations.
3. **Temperature Analysis:** Hourly, daily, and monthly temperature records allow for the examination of seasonal variations, heatwaves, and cold spells across different regions.
4. **Humidity Analysis:** Relative humidity data provides insights into atmospheric moisture levels, influencing human comfort, agricultural practices, and disease spread.
5. **Air Pressure Analysis:** Air pressure readings offer crucial information on atmospheric stability, weather system movements, and the likelihood of precipitation events.
6. **Weather Description:** Descriptive information on weather conditions such as clear sky, cloudy, rainy, snowy, etc., enhances the understanding of weather phenomena and their impacts on daily activities.
7. **Wind Speed Analysis:** Wind speed measurements facilitate the understanding of wind patterns, prevailing directions, and their implications for energy generation, transportation, and urban microclimates.
8. **Wind Direction Analysis:** Wind direction data enables the characterization of prevailing wind patterns, local wind systems, and atmospheric circulation patterns.

What should be the insights from this data ?

Geographical Patterns: Need to understand how cities are distributed globally based on latitude and longitude coordinates.

Urbanization Trends: Need to identify countries or regions with the highest concentration of cities, indicating levels of urbanization and population density.

Climate Variability: Need to explore how city distribution varies in terms of latitude across different continents, providing insights into regional climate patterns.

Weather Trends: Need to analyze temperature and humidity trends over time to identify patterns, anomalies, and extreme weather events.

Correlation Analysis: Need to investigate relationships between weather variables such as temperature, humidity, wind speed, and air pressure to understand their interplay.

Spatial Analysis: Need to visualize spatial variations in weather conditions, such as temperature ranges, humidity levels, and wind patterns across different geographical locations.

Temporal Analysis: Need to examine temporal patterns in weather conditions, including diurnal variations, seasonal trends, and long-term climate shifts.

Comparative Analysis: Need to compare weather conditions between cities, countries, or regions to identify differences, similarities, and potential factors influencing climate variability.

Impact Assessment: need to assess the impact of weather conditions on various sectors such as agriculture, energy, transportation, tourism, and public health.

Forecasting and Prediction: Need to use historical weather data to develop predictive models and forecasts for future weather trends, enabling proactive planning and risk mitigation strategies.

By considering these insights, we can effectively approach the analysis of the weather data and derive meaningful conclusions to inform decision-making processes across different domains.

Overview of Weather_analysis dataset

- **City Attributes**: Gain vital context about each city's location (latitude, longitude), country, and name.
- **Humidity:** Dive into hourly humidity levels for each city, uncovering patterns and seasonal shifts.
- **Pressure:** Chart air pressure fluctuations across all cities, potentially predicting weather changes.
- **Temperature:** Unveil temperature trends, delve into heatwaves or cold spells, and correlate temperature with energy consumption.
- **Weather Descriptions:** Understand the qualitative aspects of weather through hourly textual descriptions for each city.
- **Wind Direction:** Track wind patterns, anticipate potential wind events, and explore the link between wind direction and factors like pollution dispersion.
- **Wind Speed:** Uncover crucial insights into wind patterns, predict potential hazards, and study the impact of wind speed on various activities.

Process of Data cleaning

City Lookup Table Transformation:

- Created a new table named City_lookup with columns for city_id and city_name.
- Populated this table with unique city names and corresponding city_id.

Country Table Transformation:

- Created a new table named Country with columns for country_id and country_name.
- Populated this table with unique country names and corresponding country_id.

Date Lookup Table Transformation:

- Created a new table named Date_lookup with columns for date_id and date.
- Populated this table with dates ranging from 2012-10-01 to 2017-11-30, assigning a unique date_id to each date.

Time Lookup Table Transformation:

- Created a new table named Time_lookup with columns for time_id and time.
- Populated this table with hours ranging from 1 to 12, assigning a unique time_id to each hour.

City Attributes Table Transformation:

- Created a new table named City_attributes with columns for city_id, latitude, longitude, country_id.
- Populated this table with data from the raw dataset, replacing city and country names with corresponding city_id and country_id from the City_lookup and Country tables.

Final Fact Table Transformation:

Temperature Transformation:

- To transform temperature data into the Final_fact table, I first joined the Temperature table with the City_lookup and Date_lookup tables to replace city names with city_id and dates with date_id.
- Then, I joined this data with the Time_lookup table to replace time with time_id.
- The resulting table will have columns for city_id, date_id, time_id, and temperature.

Humidity Transformation:

- Similar to temperature transformation, I performed joins with the Humidity table and the lookup tables (City_lookup, Date_lookup, Time_lookup) to replace city names, dates, and times with their respective IDs.
- This process results in a table with columns for city_id, date_id, time_id, and humidity.

Air Pressure Transformation:

- Followed the same steps as temperature and humidity transformation, joining the Air_Pressure table with lookup tables to replace city names, dates, and times with IDs.

- This transformation yields a table with columns for city_id, date_id, time_id, and air_pressure.

Weather Description Transformation:

- To transform weather description data into the Final_fact table, I first joined the Weather_Description table with the City_lookup and Date_lookup tables to replace city names with city_id and dates with date_id. Then, joined this data with the Time_lookup table to replace time with time_id.
- This process results in a table with columns for city_id, date_id, time_id, and weather_description.

Wind Speed and Wind Direction Transformation:

- For wind speed and wind direction, I had separate tables (Wind_Speed and Wind_Direction) with columns for city_name, date, time, and respective values.
- Joined these tables with lookup tables to replace city names, dates, and times with IDs, resulting in tables with columns for city_id, date_id, time_id, wind_speed, and wind_direction.

Combining Data into Final_fact Table:

- After transforming each individual parameter into separate tables with consistent columns (city_id, date_id, time_id, parameter value), combined these tables using SQL UNION or JOIN operations into a single Final_fact table.
- Ensured that each row in the Final_fact table contains values for all parameters (temperature, humidity, air pressure, weather description, wind speed, wind direction) for a given city, date, and time.

Data Cleaning:

- Removed any duplicate or irrelevant rows from the Final_fact table.
- Handled missing values by either removing rows with missing data or imputing missing values using appropriate statistical methods.
- Verified Data Integrity:
- Checked for data integrity by ensuring that each column contains valid and consistent data types and values.
- Performed sanity checks to verify that transformed data aligns with the original dataset.

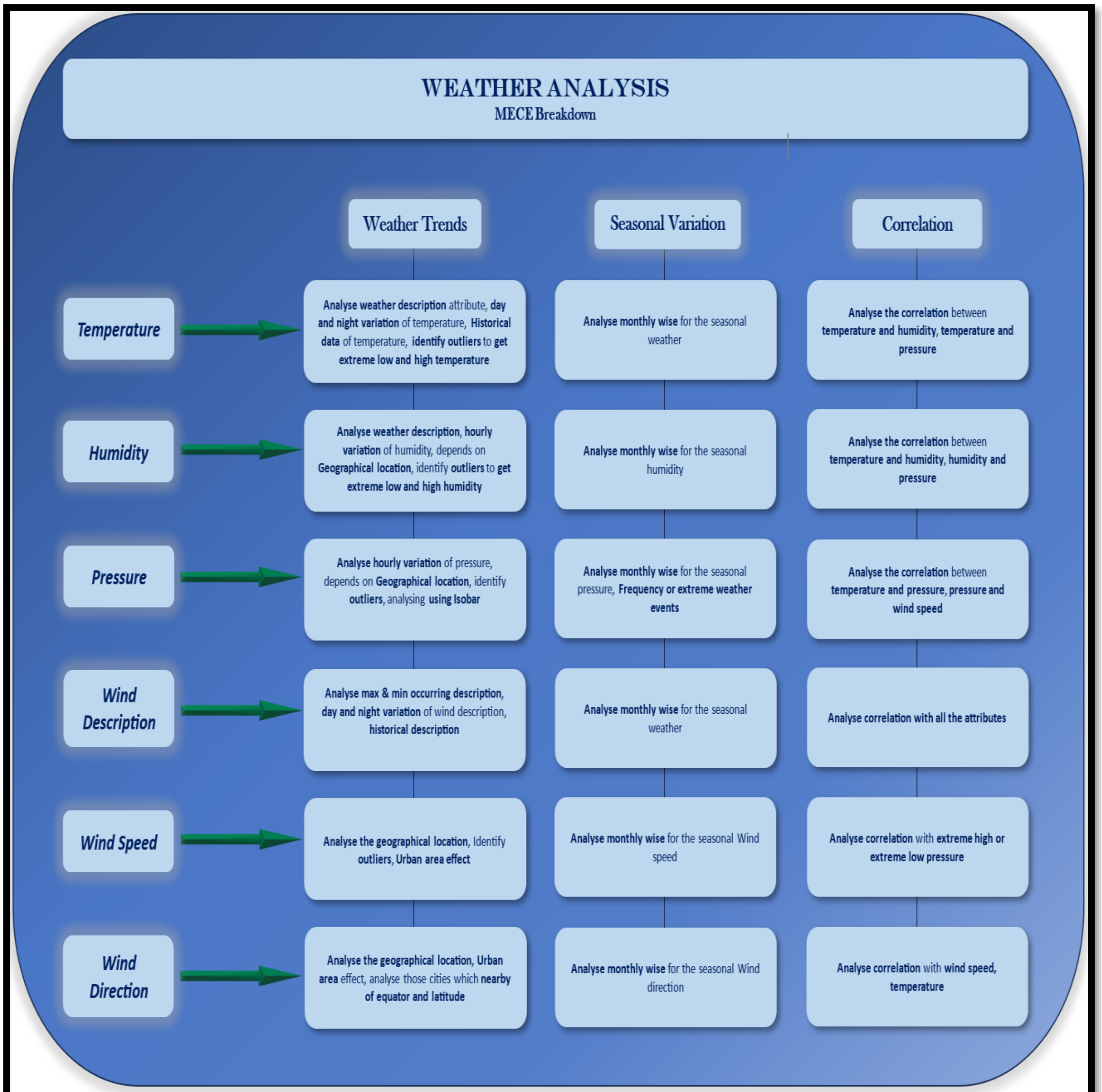
This is the final_fact Table sample I made

City_id	date_id	time_id	humidity	pressure	temperature	weather_description	wind_direction	wind_speed
32	36	16	48	1014	293.23	sky is clear	0	1
32	86	13	81	1014	283.15	sky is clear	0	1
32	132	7	72	1014	288.15	sky is clear	0	1
32	132	8	72	1014	288.15	sky is clear	0	1
32	145	10	81	1014	283.15	sky is clear	0	1
32	163	5	38	1014	297.17	sky is clear	0	1
32	172	18	82	1014	285.16	sky is clear	0	1
32	173	22	24	1014	298.02	sky is clear	0	1
32	185	7	82	1014	290.31	sky is clear	0	1
32	185	10	77	1014	288.48	sky is clear	0	1
32	186	18	71	1014	286.57	sky is clear	0	1
32	189	11	77	1014	290.13	sky is clear	0	1
32	189	12	82	1014	289.45	sky is clear	0	1
32	194	16	72	1014	287.78	sky is clear	0	1
32	194	12	76	1014	287.42	sky is clear	0	1
32	195	13	79	1014	286.97	sky is clear	0	1

Overview of Weather_analysis_transformed dataset

- **City_attributes –** Gain vital context about each city's location (latitude, longitude), country_id, and city_id
- **City_lookup –** Each city’s name & id
- **Country –** Each country’s name & id
- **Date_lookup –** 5 years (2012 -2017) date & date_id
- **Final_fact –** 5 years data of humidity, temperature, air_pressure, wind_speed, wind_direction, weather_description along with city_id, date_id & time_id
- **Time_lookup –** Time in hour and time_id

MECE Breakdown

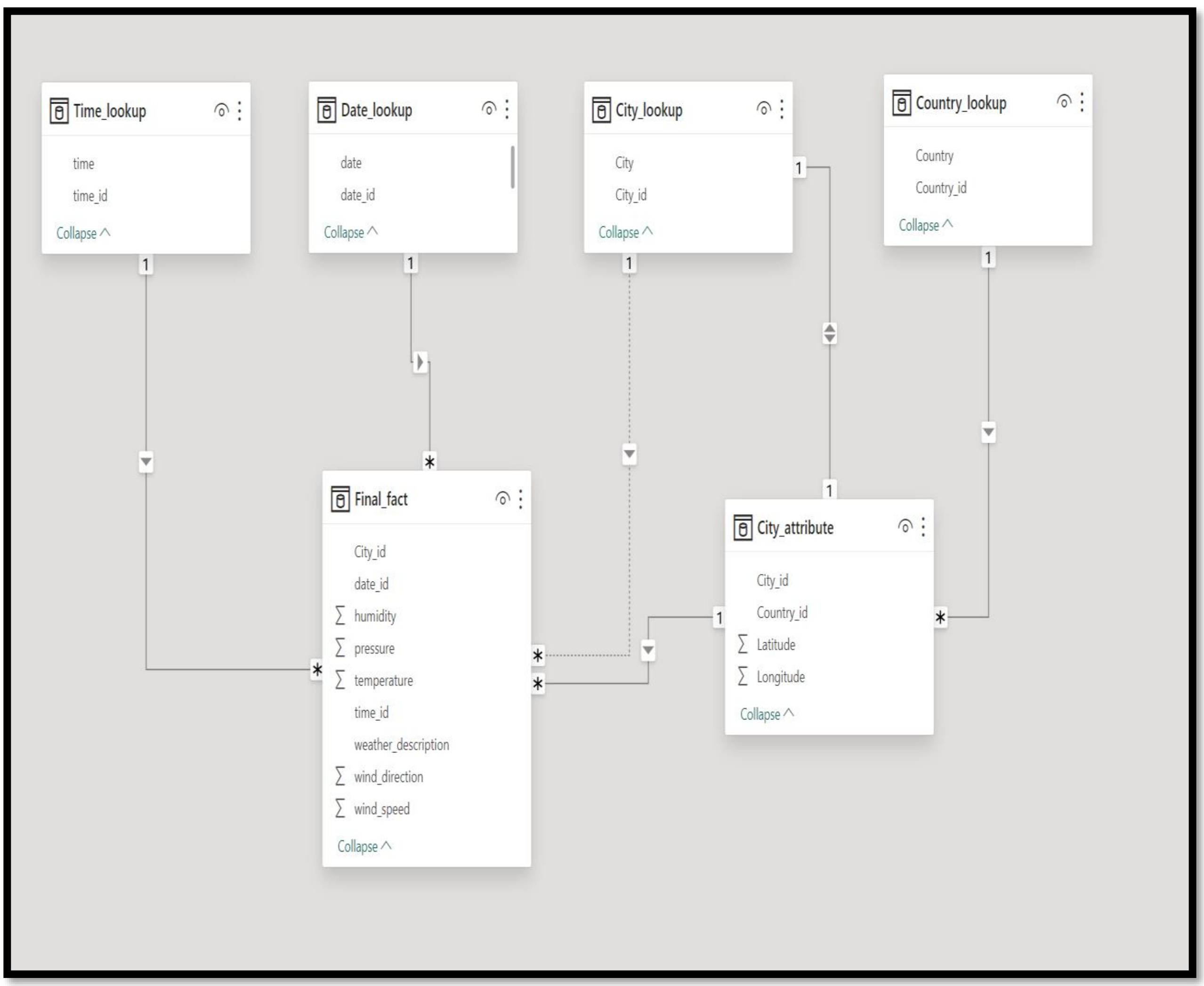


Understanding & solving PowerBI questions

To achieve a profound grasp of the dataset, I delved into 15 questions utilizing PowerBI. Let's navigate through them comprehensively.

- Gaining insight into global city distribution based on latitude and longitude.
- Identifying countries with the highest concentration of cities.
- Explore latitude distribution across continents.
- Analyzing temperature trends and extreme events over time.
- Visualizing humidity variation across cities.
- Understanding the relationship between wind speed and air pressure.
- Identifying overall temperature trends in the dataset.
- Discovering busiest hours for specific weather conditions.
- Exploring diurnal variation in wind speed.
- Comparing temperature variations between cities.
- Visualizing temperature ranges across countries.
- Identifying cities with extreme average temperatures.
- Understanding prevailing wind directions for cities.
- Exploring average wind speeds across cities by month.
- Analyzing the relationship between wind speed and air pressure.

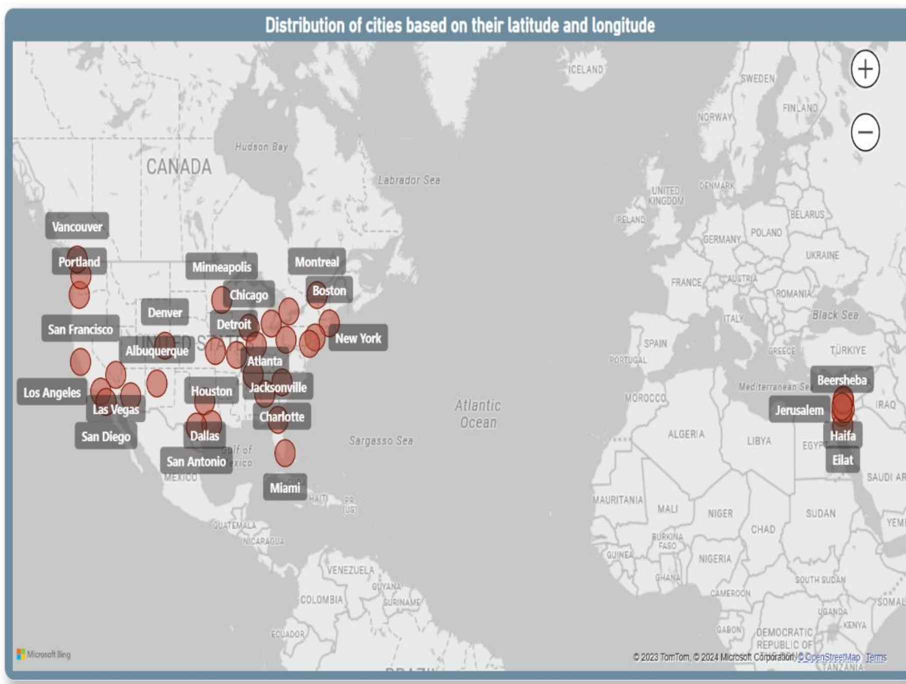
ER-Diagram for weather_analysis_transformed dataset



1. Can you create a geographical map in Power BI showing the distribution of cities in the dataset based on their latitude and longitude?

<u>How will creating a geographical map help achieve the specific goals for this data analysis:</u>

- **Visualize geographic spread:** Get a quick overview of where the cities are located globally or within a specific region.
- **Identify clusters or patterns:** If cities are concentrated in certain areas, then it is revealing potential regional trends or factors influencing their location.
- **Plan future initiatives:** This map can inform decisions about resource allocation or expansion plans based on geographical distribution.



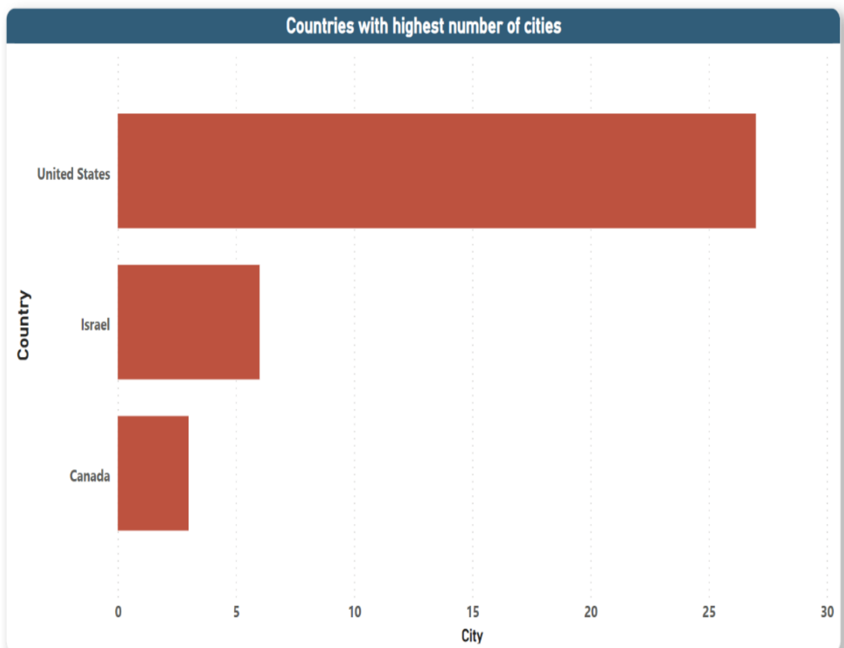
<u>Insights we get:</u>

- **Geographic distribution:** We can see the distribution of United states cities maximum based on their latitude & longitude.

2. In Power BI, can you create a bar chart representing the top 10 countries with the highest number of cities in the dataset?

<u>How will creating a bar chart help achieve the specific goals for this data analysis:</u>

- **Identify major players**: Understanding which countries hold the most significant portion of city data.
- **Focus analysis:** Prioritizing efforts by directing attention to countries with more data points.
- **Compare regional trends:** Exploring potential differences in city distribution across various regions.



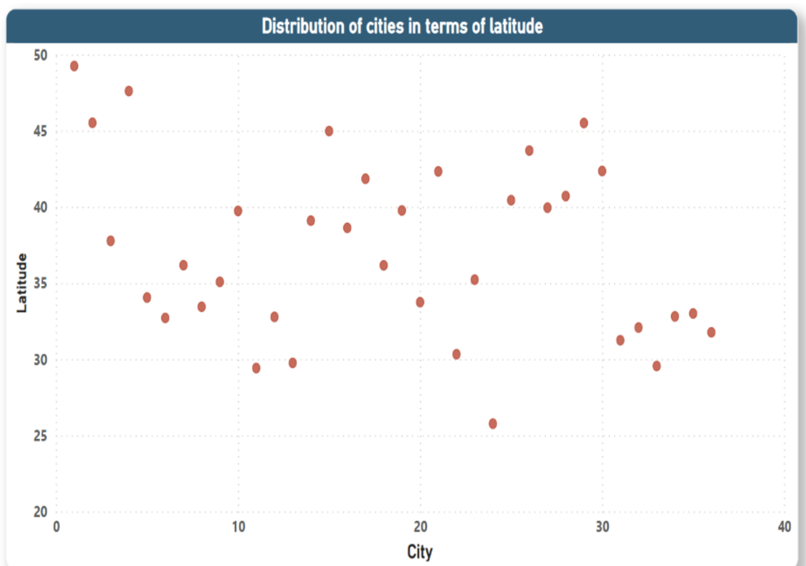
<u>Insights we get:</u>

- **Top countries:** Clearly can see United States has the most cities, potentially indicating historical, economic, or geographical factors in play.

3. How does the distribution of cities in terms of latitude vary across different continents? Create a scatter plot in Power BI to illustrate this.

How can this help us understand the data better:

- **Visualize global patterns:** Gaining a quick overview of where cities are concentrated or dispersed across different latitude ranges within each continent.
- **Identify potential biases:** Uncovering any imbalances in data if cities primarily fall within specific latitude ranges for certain continents.
- **Informs further analysis:** This visualization can guide us towards deeper dives into specific latitude zones or comparisons between continents.



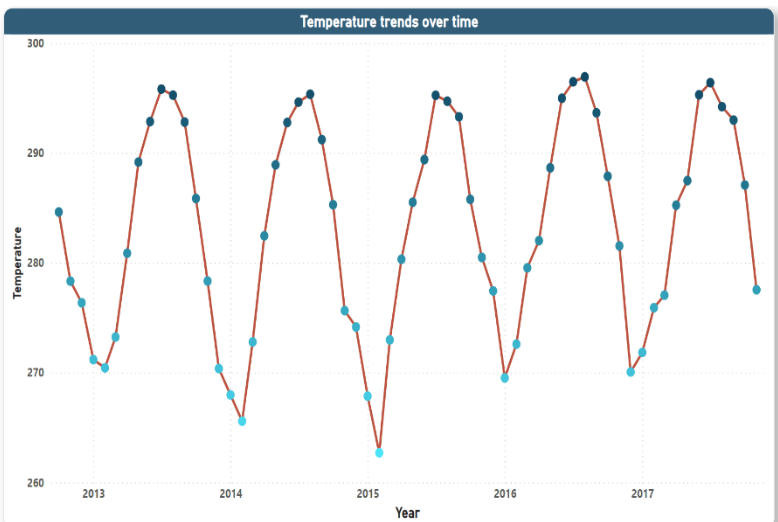
Insights we get:

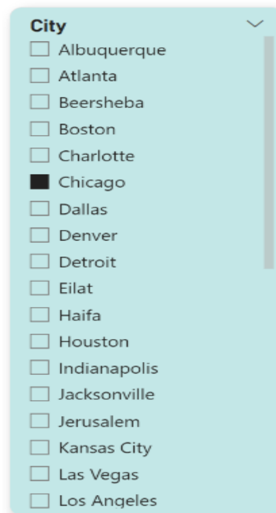
We get the correlation of -0.269 between cities and latitude. It indicates a weak negative correlation. This suggests that as the distribution of city increases, latitude tends to decrease slightly, or vice versa. However, the correlation is weak, meaning that there is not a strong linear relationship between the two variables.

4. Create a line chart in Power BI to display the temperature trends over time for a selected city. Highlight extreme temperature events.

<u>How can this help us understand the city's climate better:</u>

- **Track seasonal variations:** Seeing how temperatures fluctuate throughout the year, identifying warm and cold seasons.





- **Identify trends:** Understanding long-term patterns like warming or cooling trends over time.
- **Spot extreme events:** Spotting the highlight heatwaves, cold spells, or unusual temperature occurrences for further analysis.
- **Inform decision-making:** Help with resource planning, energy usage predictions, or understanding impacts on agriculture or human health based on temperature patterns.

<u>Insights we get:</u>

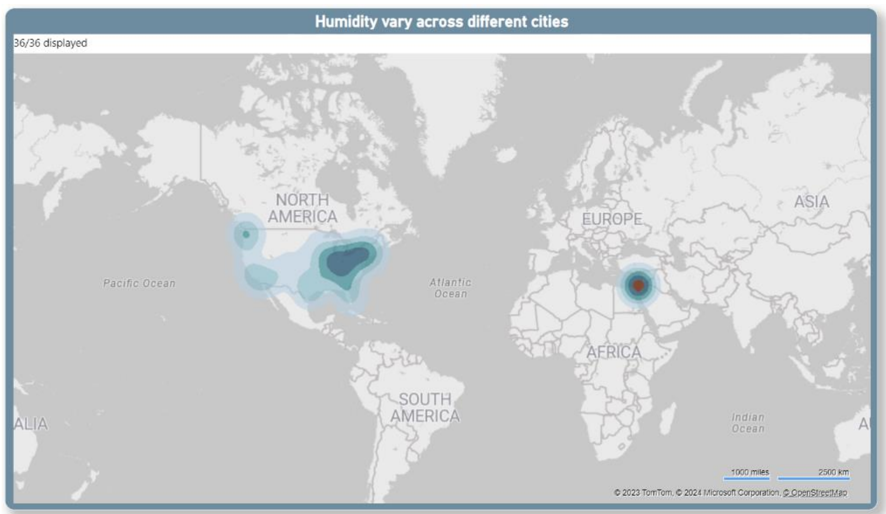
We can see for city Chicago in 2015 February the temperature was extreme lowest around 262.71 degree kelvin & in 2016 August the temperature was extreme highest around 296.93 degree kelvin.

5. How does humidity vary across different cities? Generate a heatmap in Power BI to visualize this variation.

Benefits:

- Heatmaps showcasing differences in magnitude across multiple data points.
- The colour gradient provides an intuitive comparison of high and low humidity levels.
- It's a visually engaging way to present geographical variations.

Insights we get:



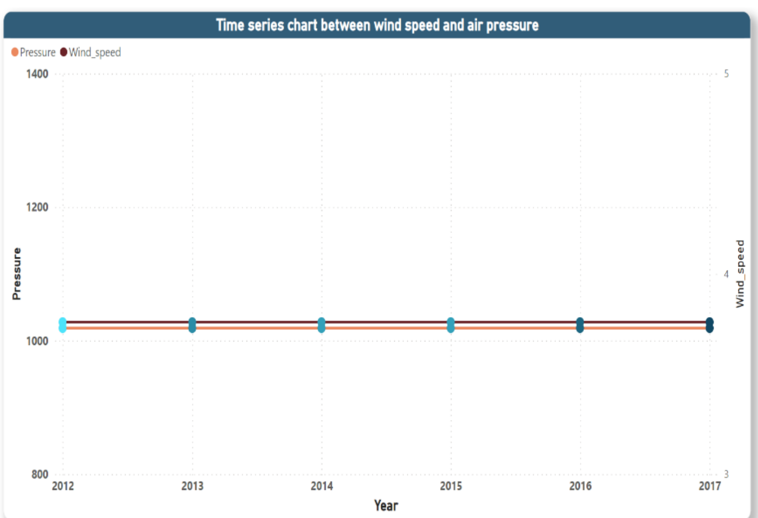
The **deepest brown colour showing the max humidity level**, and **light green to dark green showing here the lowest to 2nd highest humidity level**.

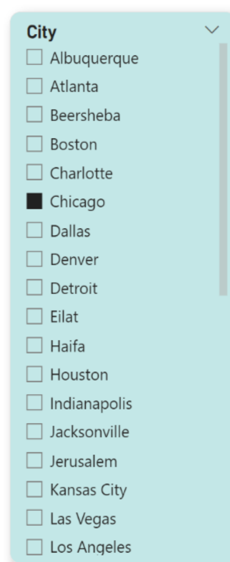
Nearby Europe, we can see a clustered of deep brown colour, means in that area the humidity level is maximum, also in North America some places we can see lightest green colour, which means in that area the humidity level is minimum.

6. Can you create a time-series chart in Power BI showing the relationship between wind speed and air pressure for a specific city?

<u>Why a time-series chart:</u>

- This chart effectively visualizes how both variables change over time, highlighting potential correlations or patterns.
- It allows you to easily see how wind speed reacts to fluctuations in air pressure, and vice versa.
- Observing if the wind speed-air pressure relationship varies across seasons, revealing seasonal patterns or influences.





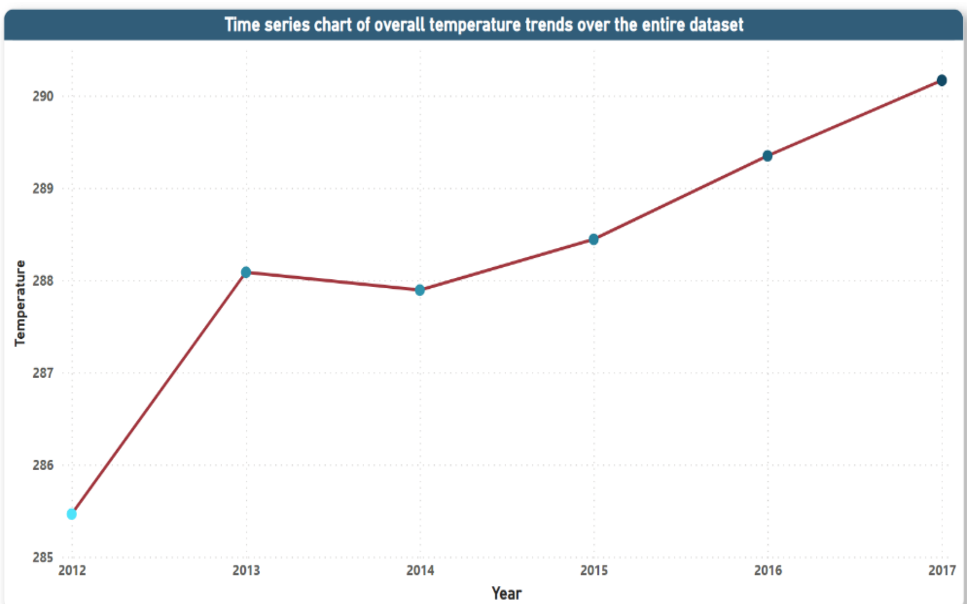
<u>Insights we get:</u>

We can see for the city Chicago these two variables wind_speed & air_pressure moving parallelly within minimal distance, it suggests a strong positive correlation between them. This means that as one variable increases or decreases, the other variable tends to move in the same direction and by a similar amount. Here are some specific interpretations depending on the variables involved.

7. Create a time-series line chart in Power BI to show the overall temperature trends over the entire dataset.

<u>Why a time-series line chart:</u>

- This chart effectively visualizes how temperature changes over time, allowing us to easily see trends and patterns.
- It provides a comprehensive overview of the temperature fluctuations across the entire dataset.



<u>Insights we get:</u>

In 2012 overall temperature was 285.46 degree kelvin, then suddenly in 2013 a big break it became 288.09 degree kelvin. Gradually it increased over the year and in 2017 the temperature became 290.17

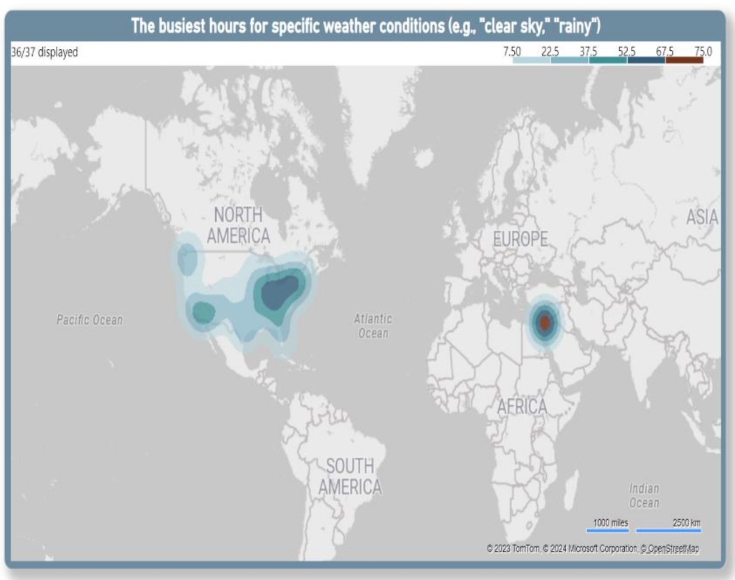
We can understand that our world temperature is increasing year over year, it can cause global-warming.

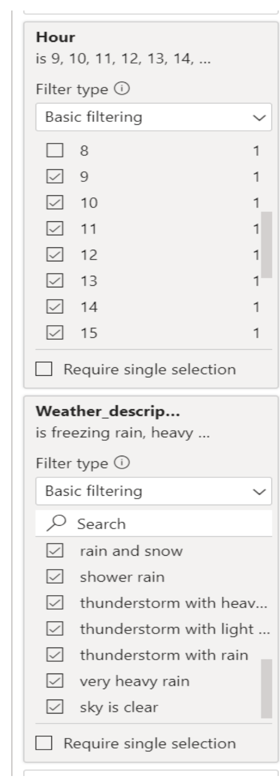
8. Can you create a heatmap in Power BI to visualize the busiest hours for specific weather conditions (e.g., "clear sky," "rainy")?

<u>Heatmap Advantages:</u>

- **Visually engaging:** Colours quickly convey busy and slow periods, making it easy to identify patterns.
- **Easy comparison:** We can compare different weather conditions side-by-side within the same visualization.
- **Flexible:** We can filter or segment the data by season, day of the week, or other relevant factors.

<u>Insights we get:</u>





The **deepest brown colour showing the max busiest hours**, and **light green to dark green showing here the lowest to 2nd highest busiest hours**.

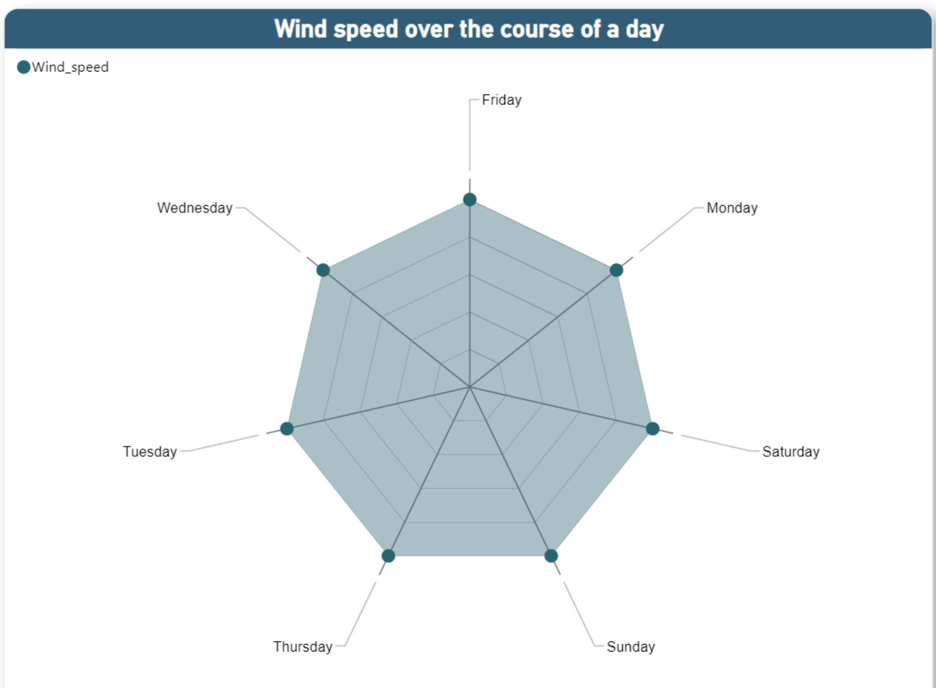
The busiest hours referring the office timing(9 am-5pm), by filtering the specific weather conditions we get the clusters of busiest hours.

9. How does the wind speed change over the course of a day? Create a radial chart in Power BI to represent this.

Why a radial chart:

- This chart effectively showcases cyclical data like time of day, making it ideal for visualizing how wind speed changes throughout a 24-hour period.
- It allows us to easily see the overall trend and specific values for each hour of the day in a visually appealing way.

Insights we get:

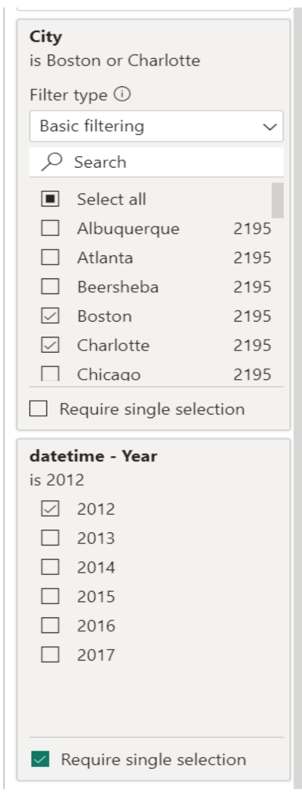


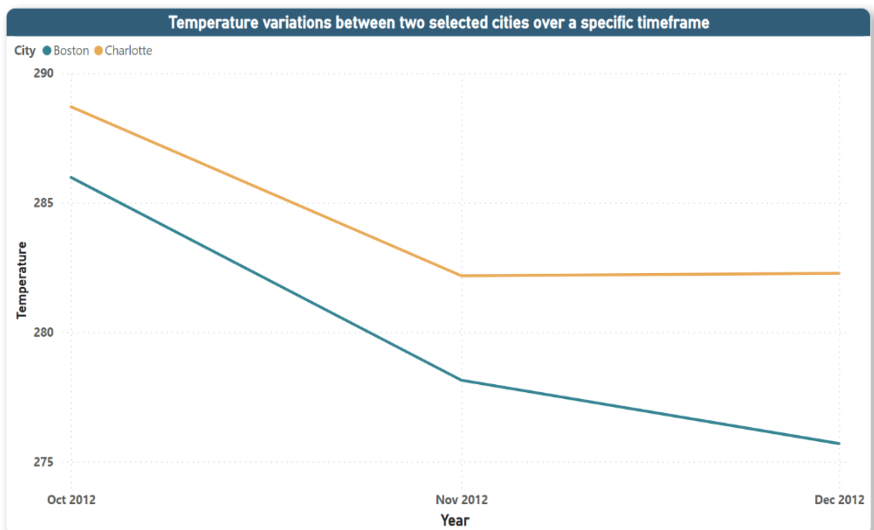
To get the course of a day we need the days , meaningfully days name. So, I extracted the name of the days from date column in power query editor in PowerBI.

By this chart, we can get the wind_speed for each day.

10. Create a Power BI chart comparing the temperature variations between two selected cities over a specific timeframe.

Line chart: This is a classic choice for visualizing trends and fluctuations, offering a clear comparison of temperature changes over time.





Why it's helpful:

By creating this visualization, we gain valuable insights into the comparative temperature profiles of the chosen cities. This can inform various decisions, planning strategies, and a deeper understanding of climatic variations between different locations.

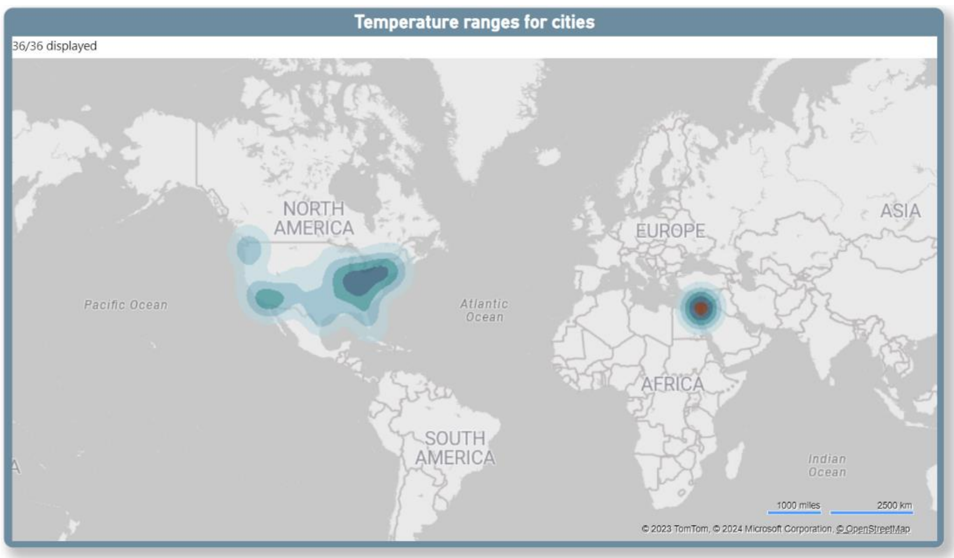
Insights we get:

We can see the temperature variation for the city Boston & charlotte together for the year 2012. We can see how temperature gradually decreased for both of the cities in dec 2012 compare to Oct 2012 . We can call it a seasonal temperature variation.

11. Can you build a heatmap in Power BI to show the temperature ranges for cities across different countries?

Benefits:

- Heatmaps showing comparative values across multiple geographical locations.
- Different colors effectively convey temperature ranges, making it easy to identify patterns and anomalies.
- It provides a visually engaging way to represent global temperature variations.



Insights we get:

The **deepest brown colour showing the max temperature level**, and **light green to dark green showing here the lowest to 2nd highest temperature level**.

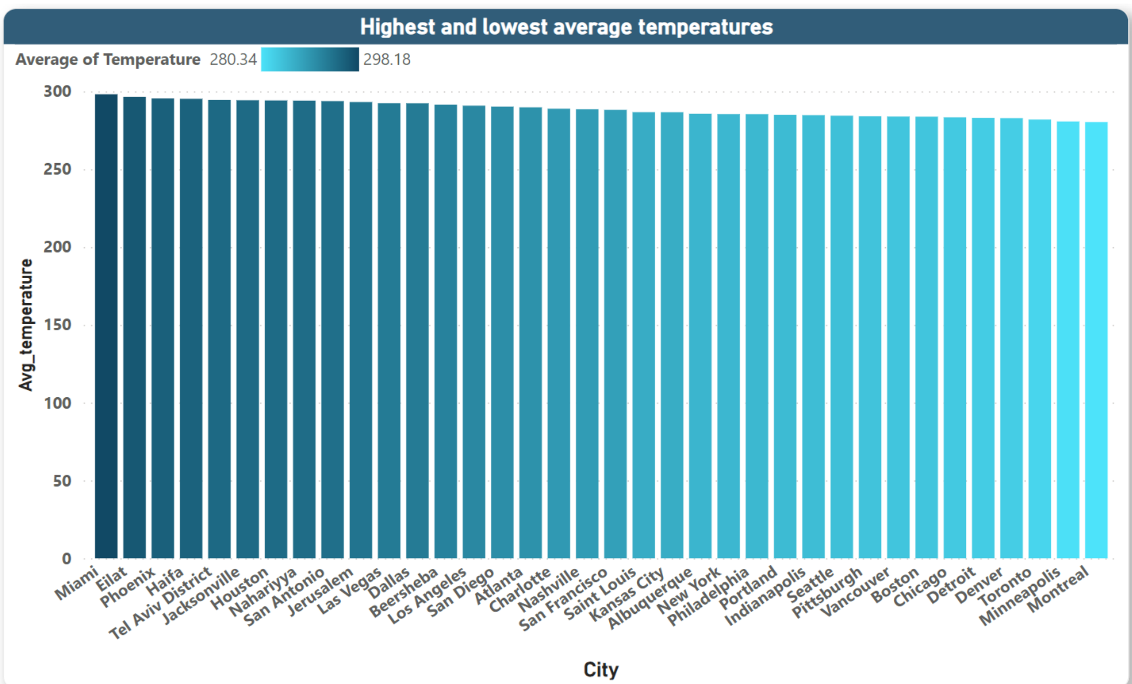
Nearby Europe, we can see a clustered of deep brown colour, means in that area the temperature level is maximum, also in North America some places we can see lighter green colour, which means in that area the temperature level is minimum.

12. Create a bar chart in Power BI to highlight cities with the highest and lowest average temperatures in the dataset.

Why a bar chart:

- This chart effectively displays categorical data like city names, making it clear which cities have the highest and lowest average temperatures.
- The bar lengths provide an intuitive comparison of the average temperatures in different cities.
- It's a simple and widely understood visualization format.

Insights we get:



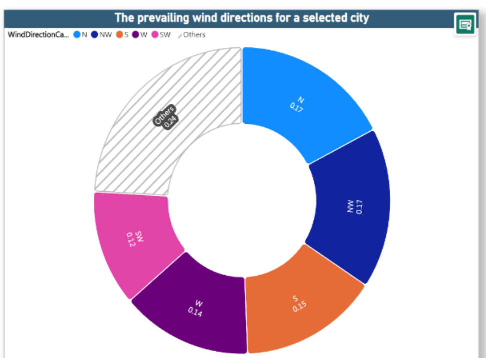
Here, I used the conditional format in bar chart. Darkest green showing the highest temp when the lightest green showing the lowest temp.

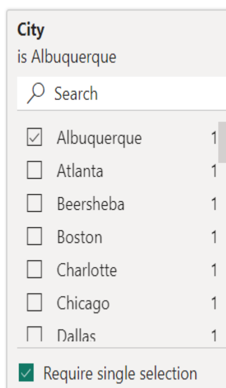
Clearly we can see the city Miami has the highest average temperature 298.18 degree kelvin & the city Montreal has the lowest average temperature 280.34 degree kelvin.

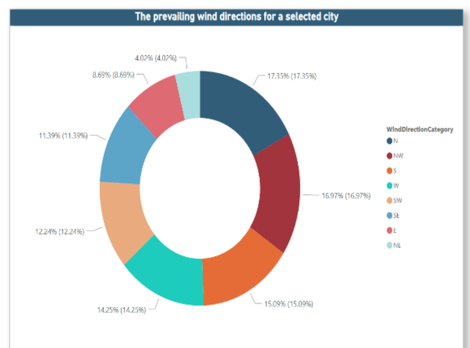
13. Create a wind rose chart in Power BI to visualize the prevailing wind directions for a selected city.

<u>Why a wind rose chart:</u>

- This chart is specifically designed to represent wind data, showcasing both direction and frequency in a visually intuitive way.
- The circular layout and segmented spokes highlight the prevailing wind directions and their relative dominance.
- It provides a compact and informative way to summarize wind patterns for a specific location.







<u>Insights we get:</u>

Unfortunately, the wind rose chart was not available, so, I tried to visualize it with rose donut chart & donut chart. **In the left side, it is rose donut chart & the right side it is donut chart.**

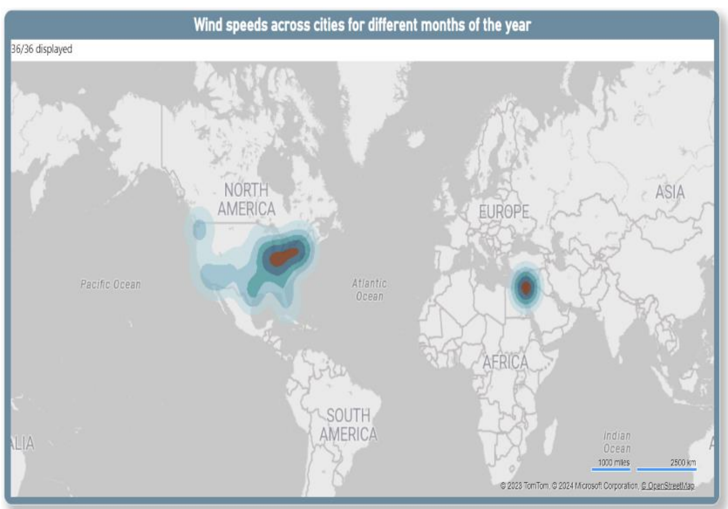
```
WindDirectionCategory =
SWITCH(
    TRUE(),
    wind_direction[Wind_direction] >= 337.5 || wind_direction[Wind_direction] < 22.5, "N",
    wind_direction[Wind_direction] >= 22.5 && wind_direction[Wind_direction] < 67.5, "NE",
    wind_direction[Wind_direction] >= 67.5 && wind_direction[Wind_direction] < 112.5, "E",
    wind_direction[Wind_direction] >= 112.5 && wind_direction[Wind_direction] < 157.5, "SE",
    wind_direction[Wind_direction] >= 157.5 && wind_direction[Wind_direction] < 202.5, "S",
    wind_direction[Wind_direction] >= 202.5 && wind_direction[Wind_direction] < 247.5, "SW",
    wind_direction[Wind_direction] >= 247.5 && wind_direction[Wind_direction] < 292.5, "W",
    wind_direction[Wind_direction] >= 292.5 && wind_direction[Wind_direction] < 337.5, "NW",
    "Invalid")
```

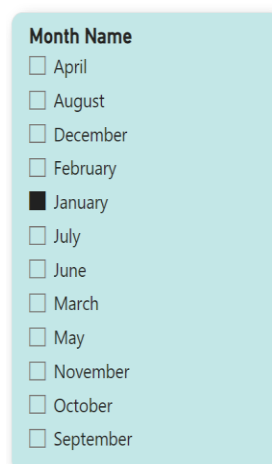
We had wind_direction values in degree unit. So, to show the wind_direction like a compass , I created a conditional column to show the direction like N(North), S(South), W(West), E(East) etc. By the rose donut chart, the insight was not clear for me, as it is not showing all the direction. That's why, I made a donut chart to show the wind_direction for each city properly. In the donut & rose chart both, the colors of slices referring to the wind_direction.

14. Can you generate a Power BI heatmap illustrating the average wind speeds across cities for different months of the year?

Why a heatmap:

- Heatmaps excel at visually representing variations across multiple data points, making it easy to see how wind speeds change for each city and month.
- The colour gradient effectively conveys differences in wind speed magnitude, allowing for quick identification of high and low wind periods.
- It provides a concise and visually engaging way to present complex wind data.





Insights we get:

The deepest brown colour showing the max wind_speed level, and light green to dark green showing here the lowest to 2nd highest wind_speed level.

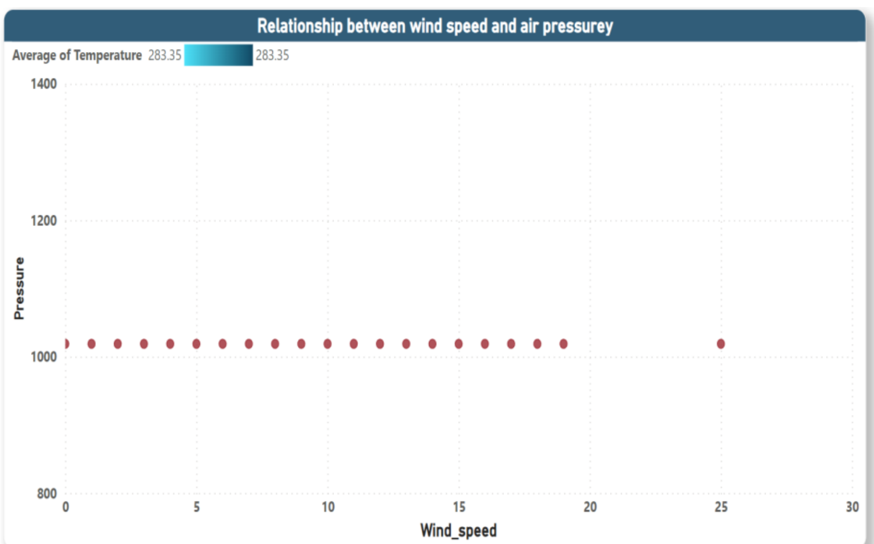
Here, I extracted the name of the month from power query in PowerBI. Nearby Europe, we can see a clustered of deep brown colour, means in that area the average wind_speed level is maximum, also in North America some places we can see lighter green colour, which means in that area the average wind_speed level is minimum.

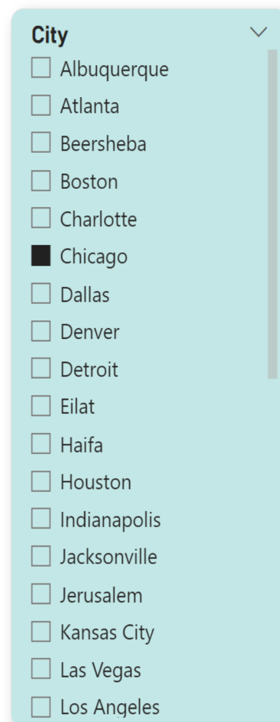
With the slicer of the month, I showed the average wind_speed for only January month.

15. Create a Power BI scatter plot to show the relationship between wind speed and air pressure for a specific city.

Why a scatter plot:

- This chart effectively showcases how two variables change simultaneously, enabling you to easily see potential correlations or patterns.
- It allows us to see the full range of data points, providing a comprehensive picture of the wind speed-air pressure relationship.





Insights we get:

To show a single city I used here slicer, here I presented the relationship between wind_speed & air_pressure for the city Chicago. Their **correlation is 0.0028**.

It suggests a **very weak positive correlation**. In practical terms, this correlation is almost negligible, indicating that there is essentially no meaningful relationship between air pressure and wind speed for Chicago city.

Why to make a dashboard ?

Improved decision-making:

Individuals: Stay informed about upcoming weather events to plan daily activities like choosing outfits, scheduling outdoor events, or deciding whether to travel.

Businesses: Make informed decisions based on weather forecasts, such as adjusting staffing levels, optimizing delivery routes, or managing inventory in weather-sensitive industries like agriculture or construction.

Organizations: Monitor weather patterns and identify trends to improve resource allocation, prepare for potential emergencies, and mitigate weather-related risks.

Enhanced situational awareness:

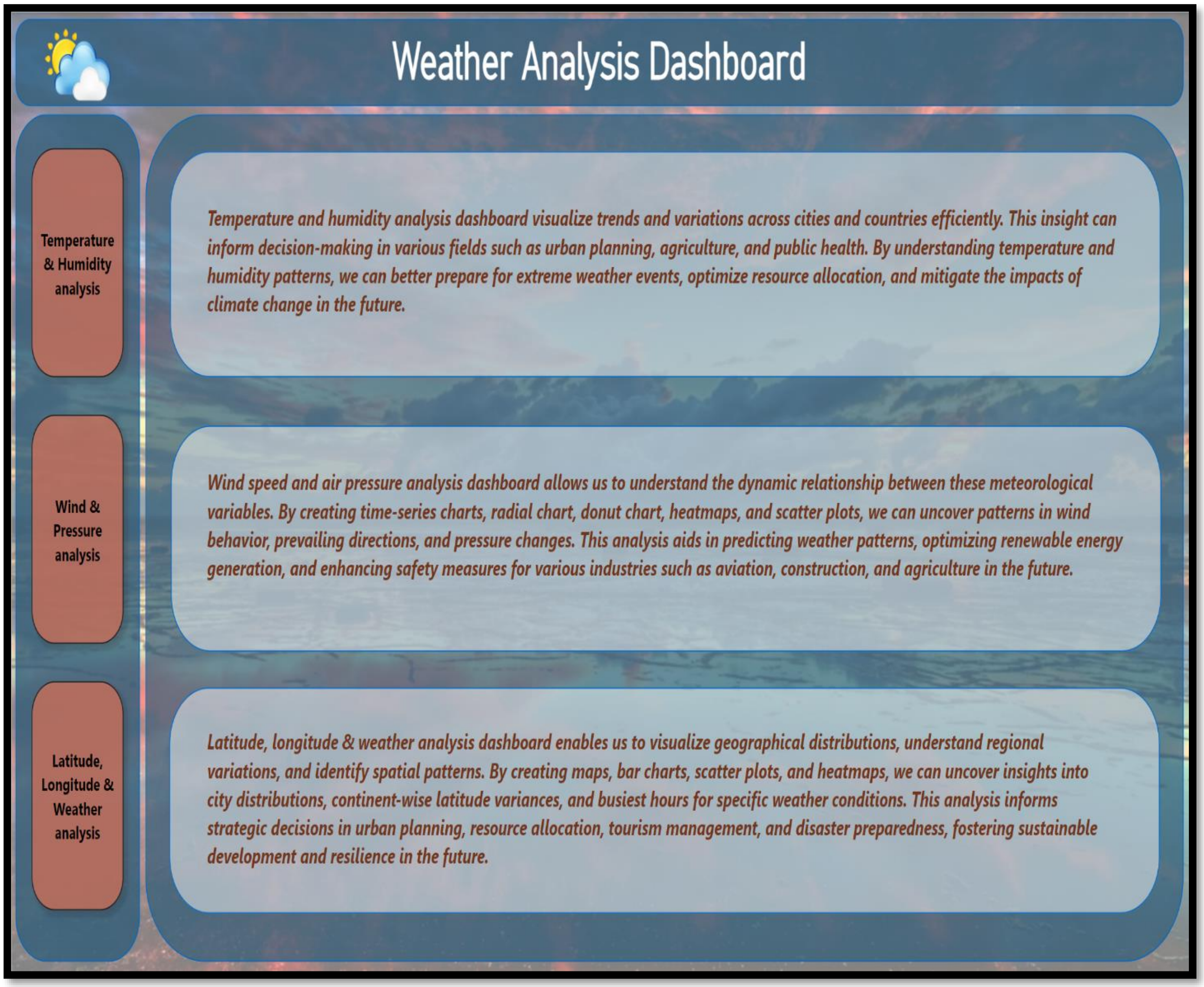
Visualize weather data: Gain a comprehensive understanding of current and past weather conditions through charts, graphs, and maps.

Identify trends and patterns: Analyse historical data to uncover seasonal variations, long-term climate changes, and potential anomalies.

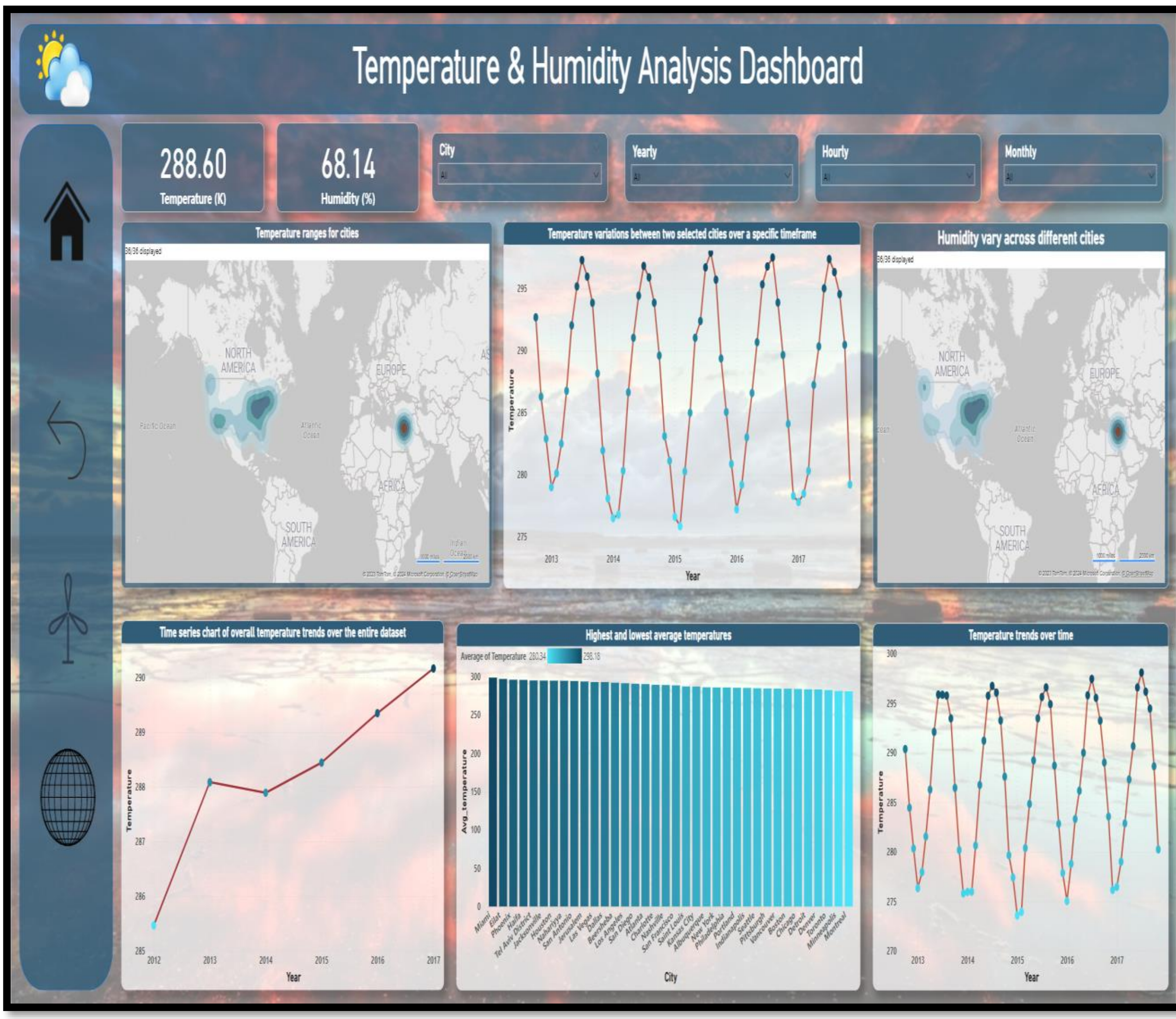
Track severe weather events: Monitor real-time updates on storms, floods, heatwaves, and other hazards to stay informed and prepared.

- **In the home page of Weather Analysis Dashboard , I briefly explained about all the 3 dashboard. By this, we can get a overview of the all analysis.**
- **All the pages of the dashboard I created Page navigator to switch one page to another page easily. Also, I created a reset button to unfilter the charts .**

Weather Analysis Dashboard



Temperature & Humidity Analysis Dashboard



Insights we get from Temperature & Humidity Analysis Dashboard

Comfort Levels: Understanding the relationship between temperature and humidity allows for the assessment of human comfort levels. By analyzing data on temperature and relative humidity, businesses can optimize indoor environments in settings such as offices, retail spaces, and hotels to enhance customer satisfaction and employee productivity.

Health Impact: Temperature and humidity levels influence human health and well-being. Monitoring these variables can help identify conditions conducive to the spread of infectious diseases, respiratory ailments, and heat-related illnesses. Healthcare facilities, public health agencies, and urban planners can use this information to implement appropriate interventions and mitigate health risks.

Energy Efficiency: Analyzing temperature and humidity data can inform energy management strategies for heating, ventilation, and air conditioning (HVAC) systems. By optimizing temperature and humidity levels based on occupancy patterns and outdoor conditions, businesses and building managers can reduce energy consumption, lower utility costs, and minimize environmental impact.

Crop Management: Agriculture relies heavily on temperature and humidity conditions for crop growth and yield. Monitoring these variables can aid farmers in making informed decisions regarding irrigation, pest control, and crop selection. Additionally, analyzing historical temperature and humidity data can facilitate the development of predictive models for crop yield forecasting and climate-smart agriculture practices.

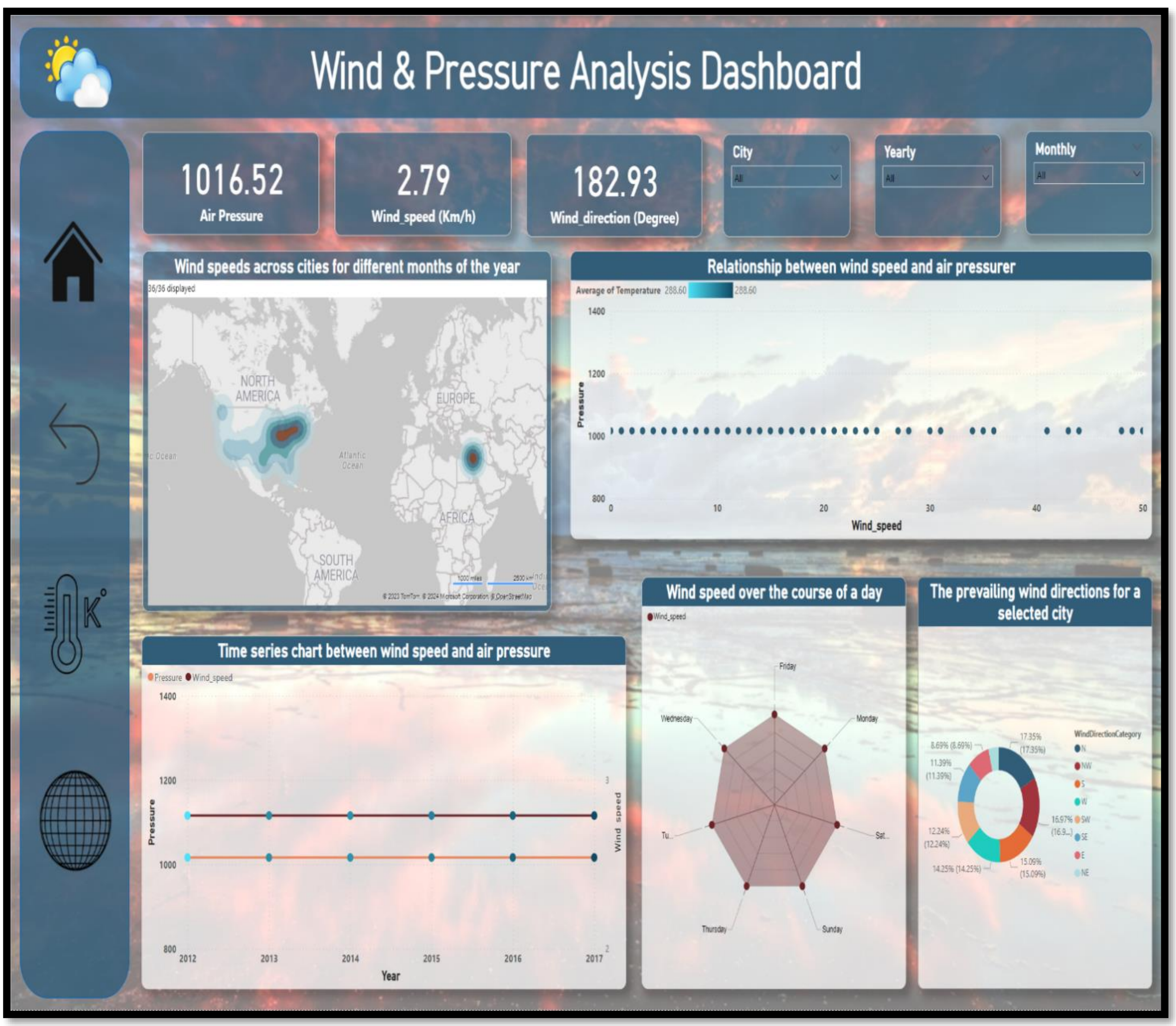
Industrial Processes: Temperature and humidity control are critical in various industrial processes, including manufacturing, food production, and pharmaceuticals. By tracking these variables in real-time and analyzing historical trends, industries can optimize production processes, ensure product quality and safety, and comply with regulatory standards.

Weather Forecasting: Temperature and humidity data are essential inputs for weather forecasting models. Analyzing historical trends and patterns in temperature and humidity can improve the accuracy of short-term and long-term weather predictions. Meteorologists and climate scientists use this information to anticipate weather-related hazards, such as storms, droughts, and heatwaves, and communicate timely warnings to the public.

Building Design and Construction: Architects and engineers consider temperature and humidity conditions when designing buildings and infrastructure. Analyzing temperature and humidity data can inform decisions regarding building materials, insulation, ventilation, and moisture control measures to enhance occupant comfort, indoor air quality, and structural durability.

Tourism and Recreation Planning: Temperature and humidity conditions play a significant role in outdoor recreational activities and tourism. Analyzing historical weather data can help businesses in the tourism industry, such as hotels, resorts, and outdoor recreation providers, anticipate seasonal fluctuations in visitor demand and tailor marketing strategies accordingly.

Wind & Pressure Analysis Dashboard



Insights we get from Wind & Pressure Analysis Dashboard

Weather Prediction and Forecasting: Wind speed and direction, along with atmospheric pressure, are critical variables in weather prediction models. Analyzing historical wind and pressure data enables meteorologists to better understand atmospheric dynamics, improving the accuracy of weather forecasts and predictions of severe weather events such as storms, hurricanes, and tornadoes.

Renewable Energy Generation: Wind speed data is essential for assessing the potential for wind energy generation in a particular area. By analyzing wind patterns and pressure gradients, energy companies and policymakers can identify optimal locations for wind farm development and plan investments in renewable energy infrastructure.

Air Quality Monitoring: Wind patterns play a significant role in dispersing air pollutants and regulating air quality. By integrating wind and pressure data with information on pollutant emissions, urban planners, environmental agencies, and public health officials can assess the impact of atmospheric conditions on local air quality and implement targeted measures to mitigate pollution levels.

Aviation and Transportation: Wind speed and direction, as well as atmospheric pressure, are crucial factors affecting aviation operations, particularly during takeoff, landing, and flight planning. Pilots and air traffic controllers rely on real-time wind and pressure data to ensure safe and efficient air travel. Similarly, wind conditions influence maritime navigation and shipping routes, with vessel operators using wind forecasts to optimize fuel efficiency and voyage planning.

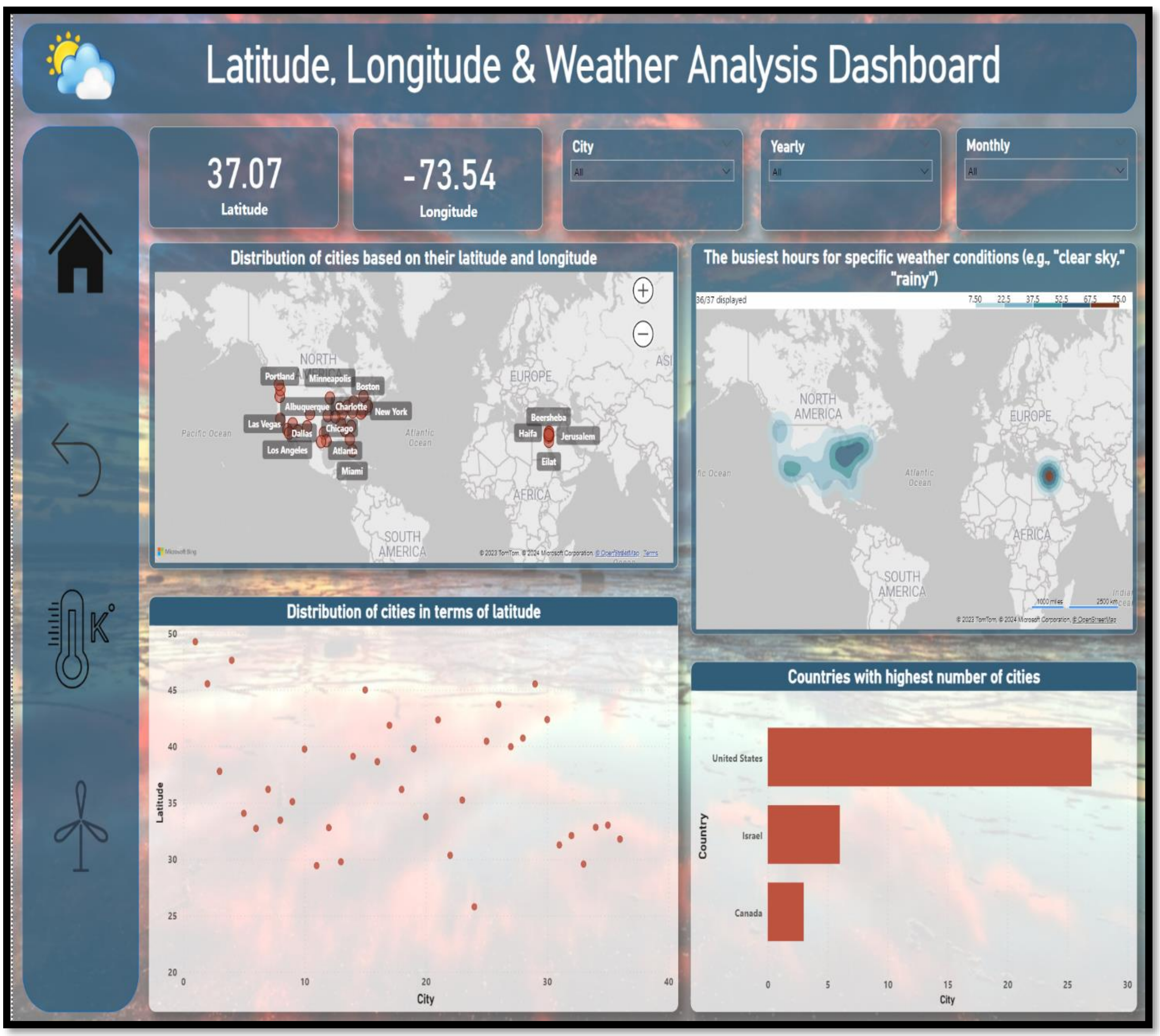
Climate Research and Modeling: Wind patterns and atmospheric pressure variations are key drivers of global climate systems. By analyzing long-term wind and pressure data, climate scientists can study climate trends, variability, and changes over time. This information contributes to climate modeling efforts, helping researchers understand the mechanisms driving climate change and its potential impacts on ecosystems and societies.

Natural Disaster Preparedness: Wind speed and pressure data are essential for assessing the intensity and track of tropical cyclones, hurricanes, and other severe weather phenomena. Emergency management agencies and disaster response organizations use this information to issue timely warnings, evacuate at-risk populations, and prepare for the impacts of extreme weather events, such as storm surges, high winds, and flooding.

Building Design and Structural Engineering: Wind load calculations based on wind speed and pressure data are critical for designing buildings, bridges, and other structures to withstand wind forces. Structural engineers use wind analysis to assess the structural integrity of infrastructure projects and ensure compliance with building codes and safety standards.

Agricultural and Environmental Monitoring: Wind patterns influence various agricultural processes, including pollination, seed dispersal, and crop evapotranspiration. By analyzing wind and pressure data alongside other environmental variables, farmers and agricultural researchers can optimize farming practices, manage soil erosion, and enhance crop yields. Additionally, wind data is valuable for studying natural ecosystems, such as dune formation, desertification, and habitat connectivity for wildlife.

Latitude, Longitude & Weather Analysis Dashboard



Insights we get from Latitude, Longitude & Weather Analysis Dashboard

Geospatial Weather Patterns: By correlating weather data with latitude and longitude coordinates, we can visualize spatial variations in weather patterns. This can help identify regions that are particularly prone to certain weather phenomena such as storms, precipitation, or temperature extremes.

Climate Zones: Latitude and longitude information can be used to categorize regions into different climate zones (e.g., tropical, temperate, polar). Analysing weather data within these zones can provide insights into typical weather conditions, seasonal variations, and climate trends specific to each zone.

Microclimate Analysis: Understanding how weather conditions vary at a local level within specific latitude and longitude coordinates is crucial for various applications, such as urban planning, agriculture, and tourism. This analysis can reveal microclimates influenced by factors like altitude, proximity to water bodies, and urban heat islands.

Weather Extremes: Examining weather data alongside latitude and longitude can help identify areas prone to extreme weather events, such as hurricanes, tornadoes, heatwaves, or droughts. This information is valuable for disaster preparedness, risk assessment, and resilience planning.

Environmental Impact Assessment: Latitude and longitude data can be used to study the environmental impact of weather phenomena on ecosystems, biodiversity, and natural resources. For example, analysing the effects of temperature changes and precipitation patterns on wildlife habitats or vegetation cover in specific geographical areas.

Regional Comparisons: Comparing weather data across different latitudes and longitudes enables insights into regional climatic differences and similarities. This analysis can be useful for businesses operating in multiple locations, policymakers, and researchers studying global climate trends.

Weather-Related Risks: Assessing the relationship between latitude, longitude, and weather data helps identify areas vulnerable to specific weather-related risks, such as flooding, landslides, wildfires, or agricultural pests. This information supports risk management and adaptation strategies for affected communities.

Tourism and Recreation Planning: Latitude and longitude data combined with weather analysis can assist in tourism and recreation planning by identifying destinations with favourable weather conditions for outdoor activities, sightseeing, and leisure travel throughout the year.

Infrastructure Planning: Understanding regional weather patterns based on latitude and longitude coordinates is essential for infrastructure planning and development. This includes considerations for building design, transportation networks, water management systems, and energy infrastructure resilience to withstand local weather conditions.

Understanding & Solving EDA questions

To gain a comprehensive understanding of the dataset, I delved into 15 exploratory data analysis (EDA) questions. Let's explore them thoroughly.

- Understanding the climate implications of extreme latitudes on countries and cities.
- Identifying geographical clusters of cities and exploring factors contributing to their clustering.
- Investigating correlations between geographical location (latitude and longitude) and weather attributes like temperature and humidity.
- Determining the top three cities with frequent rainy weather and discerning seasonal patterns.
- Analyzing the relationship between humidity levels and air pressure and its impact on weather conditions.
- Exploring patterns of temperature fluctuations in coastal cities based on wind direction.
- Identifying significant temperature variations in certain months and examining potential explanations.
- Detecting patterns in extreme weather events over time.
- Comparing temperature trends between northern and southern hemisphere cities and their relation to seasons.
- Assessing the consequences and adaptation strategies for prolonged extreme cold or heat in specific cities.
- Investigating temperature anomalies and their potential correlation with events or environmental factors.
- Analyzing the impact of temperature on energy consumption patterns in urban areas.
- Exploring how wind patterns affect air quality and pollution dispersion in urban environments.
- Identifying cities prone to strong winds and potential consequences such as increased risk of natural disasters.
- Investigating how wind speed and direction influence the frequency and severity of weather-related events in coastal cities.

1. Are there any countries with cities located at extreme latitudes, and how might this impact their climate?

```
create view abc as
select co.Country, cl.City_id, cl.City,
max(ca.Latitude) as max_latitude, ff.weather_description
from City_attributes ca
join city_lookup cl
on cl.City_id = ca.City_id
join country co
on co.Country_id = ca.Country_id
join final_fact ff
on ca.City_id = ff.City_id
group by 1,2,3,5
order by 3 desc;
```

Country	City_id	City	max_latitude	weather_description
Canada	1	Vancouver	49.24966	mist
Canada	1	Vancouver	49.24966	broken clouds
Canada	1	Vancouver	49.24966	sky is clear
Canada	1	Vancouver	49.24966	light rain
Canada	1	Vancouver	49.24966	few clouds
Canada	1	Vancouver	49.24966	fog
Canada	1	Vancouver	49.24966	overcast clouds
Canada	1	Vancouver	49.24966	light intensity shower rain

To achieve this solution, I made a temporary table (view) named 'abc' in MySQL, In this query I fetched the maximum latitude using max function in MySQL along with country, city_id, city & weather description.

- Here I pasted 9 rows including headers for a sample of the data I fetched by the SQL query.
- In this original data I got total 640 rows including headers & 4 cities – Vancouver, Montreal, Portland, Seattle.

The concept of extreme latitudes typically refers to locations that are far from the equator(both North & South), rather than those very close to it. The equator itself is defined as 0 degrees latitude, and as we move away from the equator towards higher latitudes (both north and south), the climate, daylight variations, and environmental conditions tend to change more significantly.

Extreme Northern Latitude:

A latitude above 60 degrees North is often considered to be an extreme northern latitude or Extreme maximum latitude. This includes locations in the Arctic Circle, where the sun can stay above the horizon for an extended period during the summer.

Extreme Southern Latitude:

A latitude below -60 degrees South is considered an extreme southern latitude or Extreme minimum latitude. This includes locations in the Antarctic Circle, where similar to the Arctic, the sun can stay above the horizon for an extended period during the summer.

Though, there are no cities with more than 60 degree(both North & South hemisphere) latitude in this weather_analysis_dataset, we can't consider any of them within extreme latitude. But, I'll try to decode that how might these extreme latitude can impact those cities which's latitude more than 60 degree(both North & South hemisphere). Let's deep dive into it:

North Hemisphere:

Iqaluit, Canada (63.7°N):

1) Averages -21°C in winter and 11°C in summer, showcasing similar stark seasonal variations.
2) Receives more precipitation (510mm) than other examples, contributing to its tundra landscape.
3) Faces challenges like blizzards and extreme cold snaps due to its coastal location.

Troms, Norway (69.6°N):

1) Experiences an average temperature of -5°C in winter and 13°C in summer, showcasing significant seasonal variations.
2) Receives about 900mm of precipitation annually, mostly as snow, contributing to its snowy landscapes.
3) Witnesses "polar night" (24 hours of darkness) for two months and "midnight sun" (24 hours of sunlight) for two months, impacting daily life and ecosystems.
4) Wind speeds can reach over 40 km/h, affecting transportation and infrastructure.

Yakutsk, Russia (62.0°N):

1) Holds the record for the coldest city with inhabited territory, averaging -41°C in winter and 19°C in summer.
2) Receives only 250mm of precipitation annually, contributing to its permafrost presence.
3) Experiences strong continental climate with short growing seasons, limiting agriculture.

Fairbanks, Alaska (64.8° N):

1) Fairbanks experiences a subarctic climate with extremely cold winters, often dropping below -20°C (-4°F).
2) The city has a long winter season with a significant amount of snowfall.

South Hemisphere:

Ushuaia, Argentina (54.8°S):

1) Averages 0°C in winter and 10°C in summer, representing a maritime climate with less extreme temperatures.
2) Receives over 3000mm of precipitation annually, creating a rainforest environment.
3) Experiences strong Patagonian winds that can affect outdoor activities and infrastructure.

McMurdo Station, Antarctica (77.8°S):

1) Averages -28°C year-round, reflecting the extreme cold of the Antarctic plateau.
2) Receives minimal precipitation (200mm) due to its desert-like conditions.
3) Experiences constant darkness in winter and continuous sunlight in summer, posing challenges for human adaptation.
4) Strong katabatic winds, reaching over 100 km/h, pose significant challenges for research and infrastructure.

2. Can you identify any clusters of cities with similar latitude and longitude values? What factors might explain these clusters?

```sql
create view p as
SELECT cl1.City as City1, cl2.City as City2, ca1.Latitude as Lat1,
ca2.Latitude as Lat2, ca1.Longitude as Long1, ca2.Longitude as Long2
from city_lookup as cl1
join city_lookup as cl2
on cl1.City_id <> cl2.City_id
join city_attributes as ca1
on cl1.City_id = ca1.City_id
join city_attributes as ca2
on cl2.City_id = ca2.City_id
where abs(ca1.Latitude-ca2.Latitude) <=1
and abs(ca1.Longitude-ca2.Longitude) <=1
```

```sql
select cl1.City as City1, cl2.City as City2, ca1.Latitude as Lat1,
ca2.Latitude as Lat2, ca1.Longitude as Long1, ca2.Longitude as Long2
from
  city_lookup as cl1
join city_lookup as cl2
  on cl1.City_id <> cl2.City_id
  join city_attributes as ca1
on cl1.City_id = ca1.City_id
join city_attributes as ca2
on cl2.City_id = ca2.City_id
where
  ca1.Latitude between ca2.Latitude - 1 and ca2.Latitude + 1 and
  ca1.Longitude between ca2.Longitude - 1 and ca2.Longitude + 1
```

To identify clusters of cities with exact similar latitude & longitude in this database, not possible. That's why we need to search for those cities whose latitude & longitude range within 1 degree.

So, to fetch this result I created 2 different SQL query, to fetch the same result, we can use any of this query.

- For the 1st query, in where clause I gave the difference between latitude & longitude should be equal or less than 1 degree.
- For the 2nd query, in where clause I gave the latitude should be -1 degree to +1 degree, same for longitude.
- For both queries we will get the same result as below.

City1	City2	Lat1	Lat2	Long1	Long2
Tel Aviv District	Beersheba	32.083328	31.25181	34.799999	34.791302
Jerusalem	Beersheba	31.769039	31.25181	35.216331	34.791302
Beersheba	Tel Aviv District	31.25181	32.083328	34.791302	34.799999
Haifa	Tel Aviv District	32.815559	32.083328	34.98917	34.799999
Nahariyya	Tel Aviv District	33.005859	32.083328	35.09409	34.799999
Jerusalem	Tel Aviv District	31.769039	32.083328	35.216331	34.799999
Tel Aviv District	Haifa	32.083328	32.815559	34.799999	34.98917
Nahariyya	Haifa	33.005859	32.815559	35.09409	34.98917
Tel Aviv District	Nahariyya	32.083328	33.005859	34.799999	35.09409
Haifa	Nahariyya	32.815559	33.005859	34.98917	35.09409
Beersheba	Jerusalem	31.25181	31.769039	34.791302	35.216331
Tel Aviv District	Jerusalem	32.083328	31.769039	34.799999	35.216331

I don't have enough data to fetch the factors which might explain these clusters, Though, I will try to decode these factors which might explain the clusters of the cities with similar Latitude & Longitude. Let's deep dive into it:

Geographical Features:

- Proximity to coastlines, rivers, mountains, or other natural features might influence settlement patterns.
- Cities located in close proximity to each other might share similar environmental conditions based on their geographical features.

Climate Zones:

- Latitude is a key determinant of climate zones. Cities at similar latitudes often experience similar climate patterns, which can influence their development and infrastructure.

Historical Settlement Patterns:

- Historical factors, such as colonization, migration, or trade routes, can influence the location and development of cities.
- Cities that were historically important for trade or served as regional hubs may cluster together.

Economic Activities:

- Similar economic activities, such as agriculture, industry, or technology, may attract cities to specific regions.
- Economic factors can contribute to the clustering of cities with similar latitude and longitude values.

Cultural and Political Influences:

- Cultural and political factors can play a role in the development and clustering of cities.
- Cities within the same cultural or political region may share similar characteristics.
- Shared ethnicities, languages, or religious beliefs could contribute to clustered settlements.

Infrastructure and Transportation Networks:

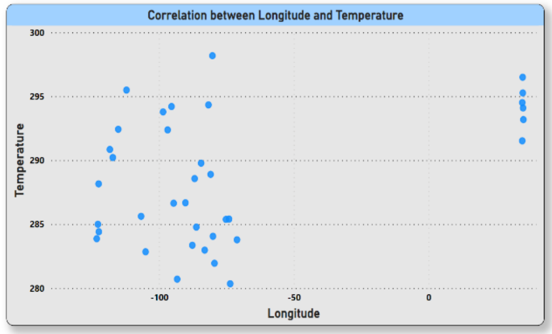
- Proximity to major transportation routes, such as highways, railways, or airports, can influence the development and connectivity of cities.

3. Are there any correlations between a city's geographical location (latitude and longitude) and its weather attributes, such as temperature or humidity?

```sql
WITH avg_temp AS (
        SELECT ROUND(AVG(ff.temperature), 2) AS a_temp
        FROM final_fact ff
),
avg_long AS (
        SELECT ROUND(AVG(ca.Longitude), 2) AS a_long
        FROM city_attributes ca
),
temp_coe AS (
        SELECT
            ff.temperature - (SELECT a_temp FROM avg_temp) AS xi_x,
            ca.Longitude - (SELECT a_long FROM avg_long) AS yi_y
        FROM final_fact ff
        JOIN city_attributes ca ON ca.City_id = ff.City_id
 )
 SELECT SUM((xi_x)*(yi_y)) / SQRT(SUM(POW(xi_x, 2)) * SUM(POW(yi_y, 2)))
 as Correl_bw_longitude_temperature
 FROM temp_coe;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Correl_bw_longitude_temperature
0.20575749576366825

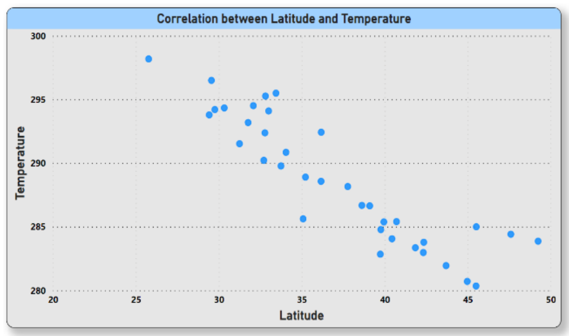


Correlation b/w Longitude & Temperature: (-0.4301449)

```sql
WITH avg_temp AS (
        SELECT ROUND(AVG(ff.temperature), 2) AS a_temp
        FROM final_fact ff
),
avg_lat AS (
        SELECT ROUND(AVG(ca.Latitude), 2) AS a_lat
        FROM city_attributes ca
),
temp_coe AS (
        SELECT
            ff.temperature - (SELECT a_temp FROM avg_temp) AS xi_x,
            ca.Latitude - (SELECT a_lat FROM avg_lat) AS yi_y
        FROM final_fact ff
        join city_attributes ca
        on ca.City_id = ff.City_id
)
SELECT SUM((xi_x)*(yi_y)) / SQRT(SUM(POW(xi_x, 2)) * SUM(POW(yi_y, 2)))
as Correl_bw_latitude_temperature
FROM temp_coe;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Correl_bw_latitude_temperature
-0.4301449468982176

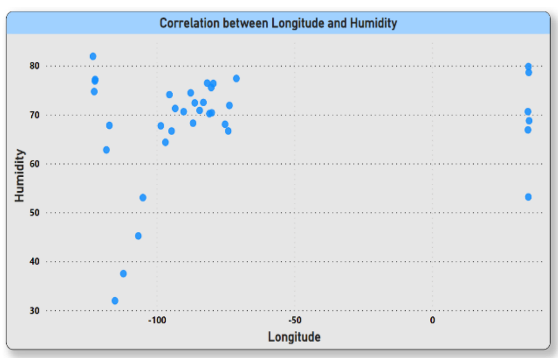


Correlation b/w Latitude & Temperature: (0.20575749)

```sql
WITH avg_hum AS (
        SELECT ROUND(AVG(ff.humidity), 2) AS a_hum
        FROM final_fact ff
 ),
 avg_long AS (
        SELECT ROUND(AVG(ca.Longitude), 2) AS a_long
        FROM city_attributes ca
 ),
 temp_coe AS (
        SELECT
            ff.humidity - (SELECT a_hum FROM avg_hum) AS xi_x,
            ca.Longitude - (SELECT a_long FROM avg_long) AS yi_y
        FROM final_fact ff
        JOIN city_attributes ca ON ca.City_id = ff.City_id
 )
 SELECT SUM((xi_x)*(yi_y)) / SQRT(SUM(POW(xi_x, 2)) * SUM(POW(yi_y, 2)))
 as Correl_bw_longitude_humidity
 FROM temp_coe;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Correl_bw_longitude_humidity
0.06713517937490078

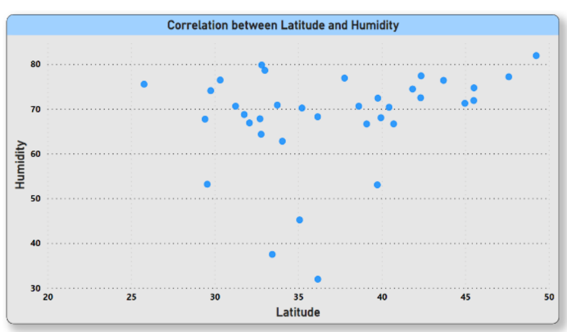


Correlation b/w Longitude & Humidity: (0.11692809)

```sql
WITH avg_hum AS (
        SELECT ROUND(AVG(ff.humidity), 2) AS a_hum
        FROM final_fact ff
 ),
 avg_lat AS (
        SELECT ROUND(AVG(ca.Latitude), 2) AS a_lat
        FROM city_attributes ca
 ),
 temp_coe AS (
        SELECT
            ff.humidity - (SELECT a_hum FROM avg_hum) AS xi_x,
            ca.Latitude - (SELECT a_lat FROM avg_lat) AS yi_y
        FROM final_fact ff
        JOIN city_attributes ca ON ca.City_id = ff.City_id
 )
 SELECT SUM((xi_x)*(yi_y)) / SQRT(SUM(POW(xi_x, 2)) * SUM(POW(yi_y, 2)))
 as Correl_bw_latitude_humidity
 FROM temp_coe;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Correl_bw_latitude_humidity
0.11692809006383399



Correlation b/w Longitude & Humidity: (0.067135179)

Correlation between Longitude and Temperature (-0.4301449):

- **Negative correlation:** This means that as longitude increases (moving eastward), temperature tends to decrease.
- **Explanation:** This negative correlation suggests that cities located farther east experience lower temperatures, while those located farther west experience higher temperatures. This relationship may be influenced by factors such as proximity to oceans or landmasses, altitude, and prevailing wind patterns. For example, cities in eastern regions of continents may experience colder climates due to continental effects, while those in western regions may benefit from maritime climates with milder temperatures.

Correlation between Latitude and Temperature (0.20575749):

- **Positive correlation:** This indicates that as latitude increases (moving towards the poles), temperature tends to increase.
- **Explanation:** This positive correlation aligns with the general understanding of how temperature varies with latitude. As you move towards higher latitudes, away from the equator, temperatures tend to decrease due to the angle of sunlight and the Earth's curvature. This relationship is commonly observed in the latitudinal temperature gradient, where equatorial regions experience warmer temperatures compared to polar regions.

Correlation between Longitude and Humidity (0.11692809):

- **Weak positive correlation:** This suggests a slight tendency for humidity to increase as longitude increases (moving eastward).
- **Explanation:** The weak positive correlation between longitude and humidity implies that there may be some influence of eastward movement on humidity levels, although other factors likely play a more significant role. Humidity is primarily influenced by factors such as proximity to water bodies, vegetation cover, temperature, and prevailing wind patterns. While longitude may have some influence on humidity due to regional climate patterns, its impact is relatively minor compared to other factors.

Correlation between Latitude and Humidity (0.067135179):

- **Weak positive correlation:** This indicates a slight tendency for humidity to increase as latitude increases (moving towards the poles).
- **Explanation:** Similar to the correlation between longitude and humidity, the weak positive correlation with latitude suggests a minor influence on humidity levels. Humidity is influenced by factors such as proximity to oceans, lakes, or rivers, vegetation cover, and temperature. While higher latitudes may experience higher humidity due to factors such as colder temperatures and proximity to water bodies, the correlation is relatively weak, indicating that latitude alone does not strongly determine humidity levels.

4. Identify the top three cities with the most frequent occurrence of rainy weather based on weather descriptions. What are the seasonal patterns?

```sql
with cte as
( select cl.City, ff.weather_description, count(*) rainy_days
from city_lookup cl
join final_fact ff
on cl.City_id = ff.City_id
where ff.weather_description like "%rain%"
group by 1,2)
select cte.City, cte.weather_description from cte
order by rainy_days desc
```

To fetch this result, I made a CTE SQL query in MySQL, here I selected city, weather description along with count of each rainy weather description for those rainy cities.
And, in where clause, I have given the condition should be any type of rainy weather.

Then, I selected only city and weather from the CTE table.

This is the sample output I got from this query, this output giving 359 rows including header, and only rainy based.

From this output I made a pivot table in excel, and fetched the top 3 cities along with the maximum number of rainy weathers. Applied sort in descending order in max count of rainy weather, then filter for top 3 cities.

City	weather_description
Portland	light rain
Seattle	light rain
Miami	light rain
Vancouver	light rain
Philadelphia	light rain
Boston	light rain
Pittsburgh	light rain
Houston	light rain
Montreal	light rain
New York	light rain

Top 3 Cities	Max count of rainy weather
Toronto	15
Montreal	14
Vancouver	13

Again, by using the pivot table in excel I got this 18 types of rainy weather in the top 3 cities.

Rainy weathers for top 3 cities
freezing rain
heavy intensity rain
heavy intensity shower rain
light intensity drizzle rain
light intensity shower rain
light rain
light rain and snow
moderate rain
proximity moderate rain
proximity shower rain
proximity thunderstorm with rain
ragged shower rain
rain and snow
shower rain
thunderstorm with heavy rain
thunderstorm with light rain
thunderstorm with rain
very heavy rain

Here's an overview of the seasonal rainy patterns for Toronto, Montreal, and Vancouver. Let's analyse and explore in details:

Toronto, Canada:

- **Wettest Seasons:** Spring (March-May) and Fall (September-November) average around 75-85mm of rain per month, with April and October typically being the wettest months.
- **Driest Season**: Summer (June-August) receives the least rainfall, averaging around 60-70mm per month.
- **Overall:** Toronto experiences a relatively even distribution of rainfall throughout the year, with no distinct dominant rainy season.

Montreal, Canada:

- **Wettest Seasons:** Spring (March-May) sees the most rain, averaging around 75-90mm per month, with May being the wettest month.
- **Driest Season:** Winter (December-February) receives the least precipitation, averaging around 50-60mm per month.
- **Overall:** Montreal has a more distinct seasonal pattern, with wetter springs and drier winters.

Vancouver, Canada:

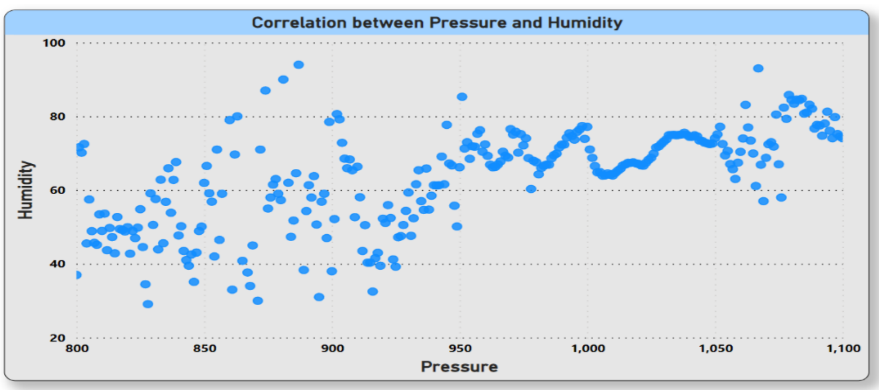
- **Wettest Seasons:** Winter (December-February) brings the most rain, averaging around 150-200mm per month, with December and January typically being the wettest months.
- **Driest Season:** Summer (June-August) is the driest period, averaging around 30-40mm per month.
- **Overall:** Vancouver has a clear rainy season in winter and a dry summer period.

5. Is there a correlation between humidity levels and air pressure? How might this relationship affect weather conditions?

```sql
WITH avg_hum AS (
    SELECT ROUND(AVG(ff.humidity), 2) AS a_hum
    FROM final_fact ff
),
avg_pres AS (
    SELECT ROUND(AVG(ff.pressure), 2) AS a_pres
    FROM final_fact ff
),
coe AS (
    SELECT
        ff.humidity - (SELECT a_hum FROM avg_hum) AS xi_x,
        ff.pressure - (SELECT a_pres FROM avg_pres) AS yi_y
    FROM final_fact ff
)
SELECT sum((xi_x)*(yi_y)) / SQRT(SUM(POW(xi_x, 2)) * SUM(POW(yi_y, 2)))
as Correl_bw_pressure_humidity
FROM coe;
```

Result Grid | Filter Rows: | Export: | Wrap Cell Content:

Correl_bw_pressure_humidity
0.07716152149953016



A correlation of 0.077 between humidity levels and air pressure is a relatively weak positive relationship. Let's see the overview how this relationship could affect weather conditions in details:

- Weak correlation means humidity level and air pressure do not strongly drive each other. Other factors have a bigger direct impact on weather.
- The slight positive correlation indicates humidity tends to be a bit higher when pressure increases, and lower when pressure decreases.
- Higher pressure typically brings more settled, stable weather. Lower pressure indicates storm systems are more likely.
- With higher pressure, subsiding air warms and loses relative humidity. Lower pressure allows rising, cooling air to retain more moisture.
- However, the weak correlation means this relationship between humidity and pressure is easily overwhelmed by other variables like temperature, winds, location, etc.
- Any impact on weather would be modest and localized. The two variables alone would not drive significant weather changes.

- Humidity and pressure may correlate more strongly on short time scales, but other dynamics dominate over long periods.

To fetch the details of weather conditions affected by this weak positive relationship, I wrote a query created a temporary table c (view), where I selected the average of humidity & air pressure along with weather_description,where weather_description is not null.

```
create view c as
select round(avg(humidity),2) avg_humidity,
round(avg(pressure),2) avg_pressure, weather_description
from final_fact ff
where weather_description is not null
group by 3
```

avg_humidity	avg_pressure	
87.35	1015.28	mist
66.18	1017.21	broken clouds
61.17	1015.85	sky is clear
80.27	1017.44	light rain
62.46	1017.09	few clouds
90.46	1016.38	fog
73.33	1015.88	overcast clouds
78.6	1011.91	light intensity shower rain
85.93	1019.6	moderate rain

This is the sample output I got from this above query, here I got total 55 rows including headers.

Weather conditions
broken clouds
drizzle
dust
few clouds
fog
freezing rain
haze
heavy intensity drizzle
heavy intensity rain
heavy intensity shower rain

From this output I made a pivot table in excel and got all the weather conditions , this is a sample output from the pivot table, totally I got 54 types of weather patterns.

In this weather patterns I got the extreme weather events, like:

- Heavy snow
- Thunderstorms
- Heavy rain
- Heatwave
- Volcanic ash

And some times clear sky , fog , smoke.

6. Explore the impact of wind direction on temperature for coastal cities. Are there noticeable patterns?

City ID	Coastal Cities
1	Vancouver
2	Portland
3	San Francisco
4	Seattle
5	Los Angeles
6	San Diego
13	Houston
22	Jacksonville
24	Miami
30	Boston
32	Tel Aviv District
33	Eilat

After exploring and researching this dataset I got 12 coastal cities, here I presented these cities in tabular form.

```
with cte as (
select cl.City, ff.wind_direction, ff.temperature, ff.weather_description
from city_lookup cl
join final_fact ff
on cl.City_id = ff.City_id
where cl.City_id in (1,2,3,4,5,6,13,22,24,30,32,33)
)
select cte.City, cte.weather_description from cte;
```

Now to get the impact of wind direction on temperature for these coastal cities I made this above SQL query using CTE function. Here, inside the CTE I selected city, wind direction, temperature and weather description where the city id should be these coastal city id.

Then, I selected only city & weather_description from the CTE.

Here, I took the sample to show the query output. In all, I got for each coatal cities different type of weather, total 539066 rows including headers.

Coastal City	weather_description
Vancouver	mist
Vancouver	broken clouds
Vancouver	broken clouds
Vancouver	broken clouds
Vancouver	broken clouds
Vancouver	broken clouds
Vancouver	broken clouds
Vancouver	broken clouds
Vancouver	broken clouds
Vancouver	broken clouds

Patterns in weathers
broken clouds
drizzle
dust
few clouds
fog
freezing rain
haze

Made a pivot tabble of that query output to get all the distinct weather conditions. These are the sample patterns of that pivot output. In this weather patterns maximum time rainy weather, cloud, light snow, proximate thunderstorm visible.

The impact of wind direction on temperature for coastal cities can exhibit several noticeable patterns:

Sea Breezes: Coastal cities often experience sea breezes, where cool air from the sea replaces warmer air over the land during the day. This can lead to lower temperatures near the coast compared to inland areas, especially during the warmer months.

Land Breezes: Conversely, at night, land breezes can develop as the land cools more rapidly than the sea, leading to cooler temperatures near the coast compared to inland areas.

Directional Influence: The direction from which the wind is blowing can have a significant impact on temperature. For example, in many coastal cities, winds blowing from the ocean (onshore winds) can bring cooler temperatures, while winds blowing from the land (offshore winds) can bring warmer temperatures.

Temperature Gradient: Depending on the specific geography and local weather patterns, there may be a temperature gradient from the coast to inland areas. Winds can transport air masses of different temperatures, and this can lead to variations in temperature along the coast.

Microclimates: Coastal cities often have microclimates influenced by factors such as elevation, proximity to water bodies, and urbanization. Wind direction can interact with these factors to create localized temperature variations.

Seasonal Variations: The impact of wind direction on temperature may vary seasonally. For example, in some coastal areas, winds blowing from the ocean may have a cooling effect in summer but a warming effect in winter.

Weather Systems: Larger weather systems, such as cold fronts or coastal storms, can also influence the relationship between wind direction and temperature in coastal cities. These systems can bring abrupt changes in temperature depending on their direction and strength.

7. Are there specific months when cities experience significant temperature fluctuations? What might explain these variations?

```sql
create view q as
select t.City,
t.month,
t.avg_temp,
max(t.avg_temp) over(partition by t.City) as max_temp,
max(t.avg_temp) over(partition by t.City) -
 t.avg_temp as max_temp_fluctuation
from(
select cl.City,
monthname(dl.date) as month,
avg(ff.temperature) as avg_temp
from final_fact ff
join city_lookup cl
on cl.City_id = ff.City_id
join date_lookup dl
on dl.date_id = ff.date_id
group by 1,2) t
order by max_temp_fluctuation desc;
```

To fetch significant maximum temperature fluctuations, I made a temporary table(view) q in MySQL. Here I fetched the average temperature and maximum average temperature for each city & month. Then subtracted the maximum average temperature with average temperature to get the maximum temperature fluctuation.

```sql
create view r as
select t.City,
t.month,
t.avg_temp,
min(t.avg_temp) over(partition by t.City) as min_temp,
t.avg_temp - min(t.avg_temp) over(partition by t.City)
 as min_temp_fluctuation
from(
select cl.City,
monthname(dl.date) as month,
avg(ff.temperature) as avg_temp
from final_fact ff
join city_lookup cl
on cl.City_id = ff.City_id
join date_lookup dl
on dl.date_id = ff.date_id
group by 1,2) t
order by min_temp_fluctuation desc;
```

Similarly, To fetch significant minimum temperature fluctuations, I made a temporary table(view) r in MySQL. Here I fetched the average temperature and minimum average temperature for each city & month. Then subtracted the average temperature with minimum average temperature to get the minimum temperature fluctuation.

As, we can see in the month of January, February & December the max_temerature_fluctuations of the city Minneapolis can go from 28 degree K to 31 degree K.
Also for the city Montreal in the of the January & February the max_temerature_fluctuation can go from 28 degree K to 29 degree K

City	month	avg_temp	max_temp	max_temp_fluctuation
Minneapolis	January	264.1889103	295.3295214	31.14061107
Minneapolis	February	265.1646816	295.3295214	30.16483977
Montreal	February	264.8355044	294.3005872	29.46508282
Montreal	January	265.6189056	294.3005872	28.68168168
Minneapolis	December	267.2912517	295.3295214	28.03826963
Detroit	February	268.8679292	295.6821602	26.81423106
Las Vegas	December	279.7797912	306.2685698	26.48877857

City	month	avg_temp	min_temp	min_temp_fluctuation
Minneapolis	July	295.3295214	264.1889103	31.14061107
Minneapolis	August	293.9916383	264.1889103	29.80272804
Montreal	July	294.3005872	264.8355044	29.46508282
Minneapolis	June	293.3294662	264.1889103	29.14055591
Montreal	August	293.0499331	264.8355044	28.21442863
Detroit	July	295.6821602	268.8679292	26.81423106
Minneapolis	September	290.9390717	264.1889103	26.75016145

Yes there are specific months when cities experience significant temperature fluctuations. As, we can see in the month of June, July & August the min_temerature_fluctuations of the city Minneapolis can go from 29 degree K to 31 degree K.
Also for the city Montreal in the of the July & August the min_temerature_fluctuation can go upto 29 degree K.

Cities can experience significant temperature fluctuations for various reasons, depending on several factors. Here's a breakdown of some key influences:

Global and Regional Factors:

Seasonal changes: This is the most common cause, with cities in temperate regions exhibiting significant swings between warm summers and cool winters.
Continental vs. maritime climates: Cities further inland (continental) experience wider fluctuations compared to coastal cities (maritime) due to the moderating effect of large bodies of water.
Synoptic weather patterns: Large-scale weather systems like high- and low-pressure systems can cause rapid temperature changes within a short period.

Local Factors:

Urban heat island effect: Cities tend to be warmer than surrounding areas due to increased heat absorption from structures and reduced vegetation. This effect can amplify temperature fluctuations, especially during heat waves.
Local geography: Mountains, bodies of water, and specific landforms can influence wind patterns and air circulation, impacting local temperature variations.
Microclimates: Within a city, smaller areas can have their own microclimates due to factors like building density, vegetation cover, and proximity to water.

8. Identify periods of extreme weather events, such as storms or heatwaves, by analyzing the time-based data. What patterns emerge?

```
use weather_analysis_transformed;
select dl.date, cl.City, ff.weather_description
from date_lookup dl
join final_fact ff
on ff.date_id = dl.date_id
join city_lookup cl
on cl.City_id = ff.City_id
where ff.weather_description like "%thunderstorm%"
or ff.weather_description like "%tornado%"
or ff.weather_description like "%sand%"
or ff.weather_description like "%ash%";
```

In this SQL query a made a temporary table to fetch this output. Here, I selected date, city & weather description where the weather description are thunderstorm, tornado, sand & ash (heatwaves). These all are extreme weather events.

date	City	weather_description	Year
2012-10-07	San Francisco	proximity thunderstorm	2012
2012-10-09	San Francisco	thunderstorm	2012
2012-10-19	San Francisco	thunderstorm with heavy rain	2012
2012-12-11	San Francisco	thunderstorm	2012
2012-12-11	San Francisco	thunderstorm with heavy rain	2012
2012-10-12	Los Angeles	thunderstorm with rain	2012
2012-10-12	Los Angeles	thunderstorm	2012
2012-10-12	Las Vegas	proximity thunderstorm	2012

This is the sample output which I got from the above query, here I extracted the year using year function to get the period. Then, I created a pivot table in excel with this output.

Top Periods	Count of weather_description
2017	3238
2016	2382
2015	2033
2014	1087
2013	1975

2012	153

Weather patterns
heavy thunderstorm
proximity sand/dust whirls
proximity thunderstorm
proximity thunderstorm with drizzle
proximity thunderstorm with rain
ragged thunderstorm
sand
sand/dust whirls
thunderstorm
thunderstorm with drizzle
thunderstorm with heavy drizzle
thunderstorm with heavy rain
thunderstorm with light drizzle
thunderstorm with light rain
thunderstorm with rain
tornado
volcanic ash

In the pivot, I got the years with the count of weather description, Here, I sorted in descending order the year to get the top year of maximum extreme weather events.

These are the extreme weather events, such as storm or heatwaves, by analyzing time – based data.

9. Are there any notable differences in temperature trends between northern and southern hemisphere cities over the year? How do they relate to seasons?

In this Weather_analysis dataset we have total 36 Cities, which all are coming under Northern hemisphere, unfortunately we don't have any cities under Southern hemisphere in this data.

But, there are many significant differences in temperature trends between northern and southern hemisphere cities over the year, due to the Earth's tilt on its axis and its revolution around the sun, different parts of the planet receive varying amounts of sunlight throughout the year, leading to the changing seasons. Let's dive deeper into :

Northern Hemisphere:

Season	Periods	Impact for situated in the Northern hemisphere
Summer	*June-August*	Longer days and more direct sunlight lead to warmer temperatures. The higher the latitude, the shorter the days and cooler the temperatures become.
Autumn	*September-November*	Days grow shorter, and sunlight becomes less direct, leading to gradual cooling. Temperatures dip more rapidly at higher latitudes.
Winter	*December-February*	Shortest days and least direct sunlight result in coldest temperatures. High-latitude regions experience polar nights
Spring	*March-May*	Days lengthen, and sunlight becomes more direct, leading to warming temperatures. The rate of warming increases at higher latitudes.

Southern Hemisphere:

Season	Periods	Impact for situated in the Southern hemisphere
Summer	*December-February*	Opposite to the north, this is the hottest time with longer days and more direct sunlight. Higher latitudes are cooler.
Autumn	*March-May*	Days shorten, and sunlight becomes less direct, leading to cooling. High-latitude temperatures drop faster.
Winter	*June-August*	Shortest days and least direct sunlight bring coldest temperatures. Polar nights occur in high-latitude regions.
Spring	*September-November*	Days lengthen, and sunlight becomes more direct, causing warming. Higher latitudes experience a faster rise in temperatures.

10. What are the consequences of prolonged periods of extreme cold or heat in specific cities? How do residents adapt to such conditions?

```
SELECT
    cl.City,
    ff.weather_description
FROM final_fact ff
  join city_lookup cl
  on cl.City_id = ff.City_id
  join date_lookup dl
  on dl.date_id = ff.date_id
  join city_attributes ca
  on ca.City_id = ff.City_id

WHERE
     (ff.weather_description LIKE '%heavy%'
    OR ff.weather_description LIKE '%ash%'
    OR ff.weather_description LIKE '%sand%')
    AND ff.weather_description not LIKE '%light%'
    AND ff.weather_description not LIKE "%rain%"
    AND ff.weather_description not LIKE "%drizzle%"
ORDER BY
    1,2
```

To fetch the consequences of prolonged period of extreme cold or heat I created a temporary table (view). There I selected city & Weather description, where weather descriptions are heavy, ash, sand and not light, rain, drizzle. Heavy will cover all type of extreme events like thunderstorms, heavy snow etc.

City	weather_description
Albuquerque	heavy snow
Albuquerque	heavy snow
Albuquerque	heavy snow
Albuquerque	heavy snow
Albuquerque	heavy snow
Albuquerque	heavy snow
Albuquerque	heavy snow
Albuquerque	heavy snow

(left) This is the sample output which I fetched from that SQL query.

(right) created a pivot table of that sample output to fetch the all unique extreme cold or heat weather events in specific cities.

Weather events
heavy shower snow
heavy snow
heavy thunderstorm
proximity sand/dust whirls
sand
sand/dust whirls
volcanic ash

I don't have enough data to show the consequences of prolonged periods of extreme cold or heat. Though, I'll still try to breakdown what can be the consequences and also how residents manage to adapt such situation. In contrast, I can offer a step-by-step explanation of the process :

Consequences of Extreme Cold:

Health Risks	Exposure to extreme cold can lead to health risks such as hypothermia, frostbite, and respiratory issues, increased risk of heart attacks and strokes, respiratory problems like pneumonia.
Infrastructure Challenges	Cold temperatures can lead to frozen & burst pipes, damaged roads, power outages, transportation disruptions and other infrastructure challenges.
Disruption of Services	Snow and ice accumulation can disrupt transportation services, leading to delays and cancellations.
Economy	Extreme cold and snow can cause for business closures, job losses, reduced agricultural yields.
Increased Energy Consumption	Residents may need to use more energy for heating, potentially leading to increased utility bills.
Social	Increased reliance on social services, isolation and mental health issues.

Adaptation Strategies to Extreme Cold:

Winter Clothing	Residents typically dress in layers and wear insulated clothing to stay warm, use snowmobiles for transportation. However, prolonged cold can still strain infrastructure and lead to health risks.
Home Insulation	Adequate insulation and heating systems are essential for maintaining a comfortable indoor environment.
Snow Removal	Cities implement snow removal strategies to keep roads and sidewalks clear for safe transportation.
Emergency Preparedness	Residents may be advised to have emergency kits, including blankets, food, and water, in case of power outages or transportation disruptions.
Social support systems	Ensuring vulnerable populations have access to shelter and healthcare during extreme events.

Consequences of Extreme Heat:

Health Risks	Extreme heat can lead to heat-related illnesses, including heat exhaustion and heatstroke, dehydration, worsening of pre-existing conditions.
Energy Demands	Increased use of air conditioning can strain energy resources and lead to power outages.
Infrastructure	Buckled roads, wildfires, power outages.
Social	Prolonged heatwaves can affect crops and lead to water shortage, increase crimes, displacement of people in some areas.
Urban Heat Island Effect	Cities may experience higher temperatures due to the urban heat island effect, impacting both outdoor and indoor environments.
Economy	Reduced productivity, increased energy costs, crop damage.

Adaptation Strategies to Extreme Heat:

Cooling Centers	Cities may establish cooling centers where residents can seek relief from high temperatures.
Water Conservation	Residents are often encouraged to conserve water during heatwaves to mitigate potential shortages.
Shade and Green Spaces	Increasing shade through trees and creating green spaces can help cool urban areas.
Awareness and Education	Public awareness campaigns educate residents on the risks of extreme heat and provide tips for staying safe, staying hydrated early morning activities.
Social support systems	Ensuring vulnerable populations have access to shelter, cooling centers and healthcare during extreme events.

11. Investigate whether temperature anomalies (unusual deviations from the norm) coincide with certain events or environmental factors in specific cities.

```sql
create view b as
SELECT
    cl.City,
    ff.temperature as Unusual_temp,
    ff.weather_description
FROM
    city_lookup cl
JOIN
    final_fact ff ON cl.City_id = ff.City_id
WHERE
  ff.temperature > (SELECT AVG(ff.temperature) + 3*STDDEV(ff.temperature)
                FROM final_fact ff)
  OR
  ff.temperature < (SELECT AVG(ff.temperature) - 3*STDDEV(ff.temperature)
                FROM final_fact ff)
```

To get the temperature anomalies coincides with certain events or environmental factors in specific cities, I created a temporary table b (view). There I selected city , temperature & weather description where (temperature is more than 3 times more of standard deviation of temperature + average temperature) or (temperature is less than 3 times more of standard deviation of temperature - average temperature). By this query we can get unusual temperature along with specific cities.

City	Unusual_temp	weather_description
Vancouver	245.15	broken clouds
Vancouver	248.9412006	broken clouds
Vancouver	254.9672484	broken clouds
Vancouver	250.15	light snow
Phoenix	319.83	scattered clouds
Phoenix	319.71	scattered clouds
Phoenix	319.8	sky is clear
Phoenix	320.89	sky is clear
Phoenix	321.22	sky is clear

This is the sample output got from the above query. Created a pivot table in excel to fetch all the unique weather events, cities, their average of unusual temperature & count of weather description. This is a sample output of the original query output.

City	Average of Unusual_temp	Count of weather_description	Weather_events
Tel Aviv District	320.93	1	broken clouds
Haifa	320.375	2	few clouds
Phoenix	320.3535714	14	fog
Eilat	320.15	2	haze
Nashville	256.8724722	12	heavy shower snow

The comfortable or ideal temperature for a city around the year between 275-295K. In this 12 cities from weather_analysis dataset has the average anomalies temperature, which is ***less than 275K and more than 295K.***

Here are some potential weather events and environmental factors that could occur when temperatures in a city drop below 275K (-2°C/28°F) or rise above 295K (22°C/72°F) which completely anomalies temperature. Let's analyse this:

Low Temperatures (below 275K):

Snow storms, blizzards	*Heavy snowfall and icy conditions due to cold air masses.*
Frost, freezing rain	*Below freezing temps lead to frost and icy precipitation.*
Cold snaps, polar vortex	*Extended periods of well below normal cold temperatures.*
Hypothermia risks	*Extremely cold weather poses health risks to vulnerable populations.*
Heating demands	*Energy consumption for heating spikes during extreme cold.*
School/transportation closures	*Cold temps disrupt services, commerce and travel.*

High Temperatures (above 295K):

Heat waves	*Sustained periods of exceptionally hot weather above normal.*
Drought conditions	*Hot and dry conditions reduce rainfall and deplete soil moisture.*
Wildfire risk	*Hot, arid weather increases probability of wildfires occurring.*
Air pollution increase	*Higher temps exacerbate ground-level ozone formation.*
Heat stroke risk	*Dangerously hot weather poses health issues, especially for vulnerable groups.*
Cooling energy demands	*Electricity usage for air conditioning surges to counter heat.*

12. Analyze the impact of temperature on energy consumption patterns in cities. Are there noticeable trends or correlations?

Unfortunately, I don't have any real time data about the impact of temperature on energy consumption patterns in cities. Before exploring the complexities of the topic, it might be helpful to start with a fundamental understanding:

The relationship between temperature and energy consumption in cities is complex and multifaceted, with several factors influencing these patterns. Here's an analysis of the major impacts and potential trends:

Increased Consumption During:

Summer	*Higher temperatures lead to increased demand for air conditioning, resulting in higher electricity consumption. This impact is particularly significant in warmer climates or cities with high population density.*
Winter	*Colder temperatures necessitate increased heating needs, both for residential and commercial buildings. This leads to higher consumption of natural gas, electricity, or other heating fuels. The specific fuel used and the impact on consumption vary depending on the infrastructure and heating systems prevalent in the city.*

Decreased Consumption During:

Mild Seasons	*During months with moderate temperatures, both heating and cooling demands decrease, leading to potentially lower overall energy consumption.*

Factors Influencing the Impact:

Building Efficiency	*Buildings with poor insulation and inefficient heating/cooling systems experience greater energy consumption fluctuations with temperature changes. Investing in energy-efficient infrastructure can significantly reduce this impact.*
Urban Planning and Design	*Compact, walkable cities with mixed-use development can encourage alternative transportation modes like cycling and walking, reducing transportation-related energy consumption influenced by temperature.*
Socioeconomic Factors	*Low-income communities are often more vulnerable to energy price fluctuations and may struggle to afford adequate heating or cooling during extreme temperatures.*
Renewable Energy Integration	*Cities with a higher penetration of renewable energy sources like solar or wind power are less susceptible to price fluctuations driven by increased demand during extreme temperatures.*

Notable Trends and Correlations:

*Studies have shown **a strong positive correlation between average annual temperature and electricity consumption** in cities.*
*The impact of temperature on energy consumption is expected to become more significant in the future due to **climate change and increasingly extreme weather events.***
Cities are implementing various strategies to mitigate the impact of temperature on energy consumption, such as promoting energy efficiency, investing in renewable energy, and adopting smart grid technologies.

13. How do specific wind patterns impact air quality and pollution dispersion in urban areas? Analyze wind direction data for insights.

Note: *Well, I will prefer to analyze the wind_speed to get the insights for air quality & pollution instead of Wind_direction*

Unfortunately, I don't have for air quality & pollution dispersion real data for these 36 cities in the dataset. However, I will endeavour to analyze the procedure and its repercussions:
To judge the air quality & pollution, we need to know about the different wind patterns & how to categorize them:

Because, we need ***well-established, simple to understand, qualitative descriptions of wind effects****, I'll go with* ***a common general wind speed classification****:*

a) **Calm breeze:** 0-1.8 km/h (or 0-1 knots)
b) **Light breeze:** 1.9-11.2 km/h (or 2-6 knots)
c) **Moderate breeze:** 11.3-28.8 km/h (or 7-16 knots)
d) **Strong breeze:** 28.9-49.9 km/h (or 17-27 knots)
e) **High wind:** 50 km/h or higher (or 28 knots or higher)

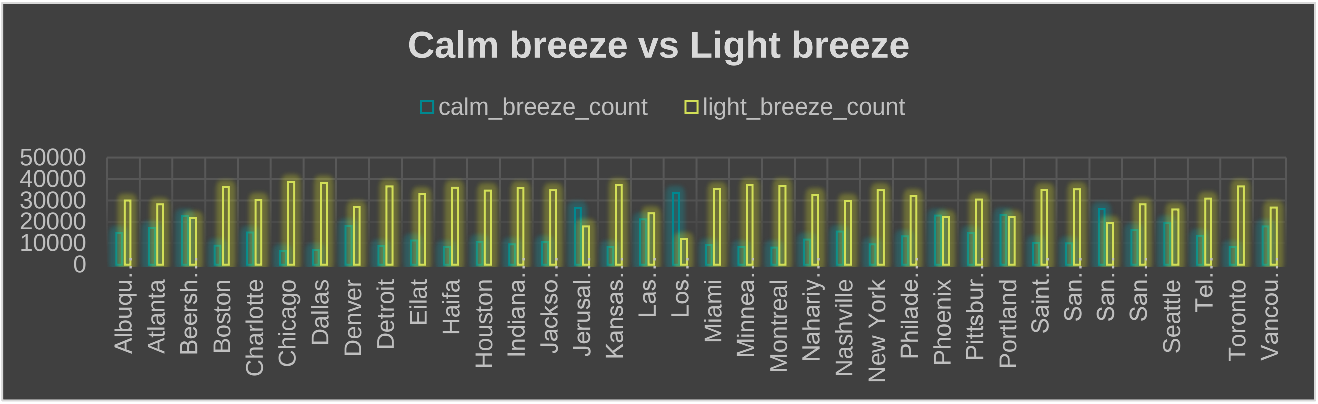
```sql
create view b as
SELECT
    cl.City,
    SUM(CASE WHEN ff.wind_speed >= 0 AND ff.wind_speed <= 1.8 THEN 1 ELSE 0 END) AS calm_breeze_count,
    SUM(CASE WHEN ff.wind_speed >= 1.9 AND ff.wind_speed <= 11.2 THEN 1 ELSE 0 END) AS light_breeze_count,
    SUM(CASE WHEN ff.wind_speed >= 11.3 AND ff.wind_speed <= 28.8 THEN 1 ELSE 0 END) AS moderate_breeze_count,
    SUM(CASE WHEN ff.wind_speed >= 28.9 AND ff.wind_speed <= 49.9 THEN 1 ELSE 0 END) AS strong_breeze_count,
    SUM(CASE WHEN ff.wind_speed >= 50 THEN 1 ELSE 0 END) AS high_wind_count
FROM
    final_fact ff
JOIN
    city_lookup cl ON cl.City_id = ff.City_id
GROUP BY
    cl.City
ORDER BY
    cl.City;
```

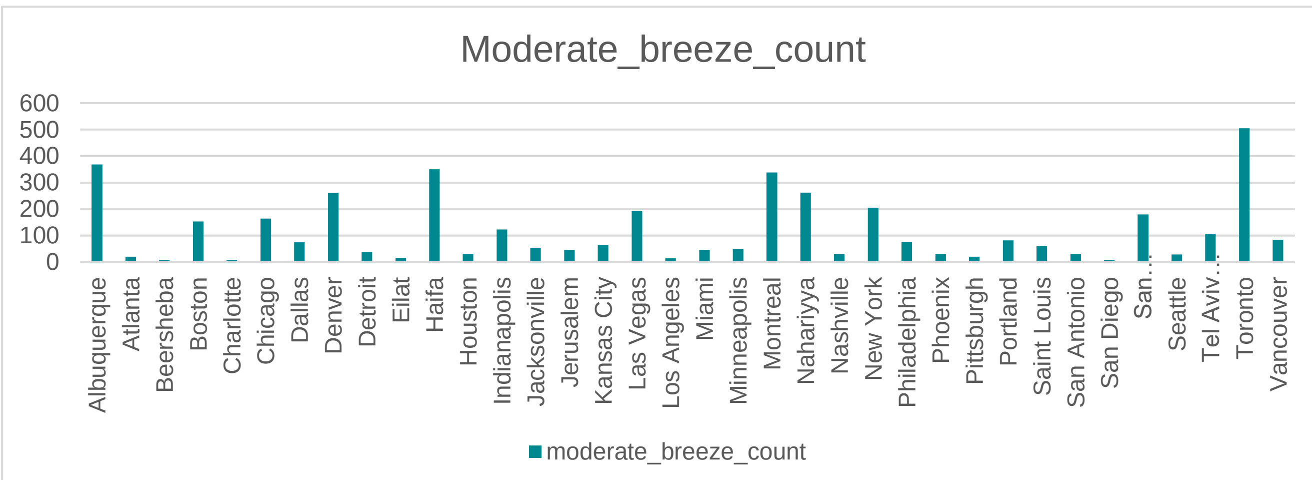
Here, I created a temporary table in MySQL to categorize the wind speed base on the range along with specific city.

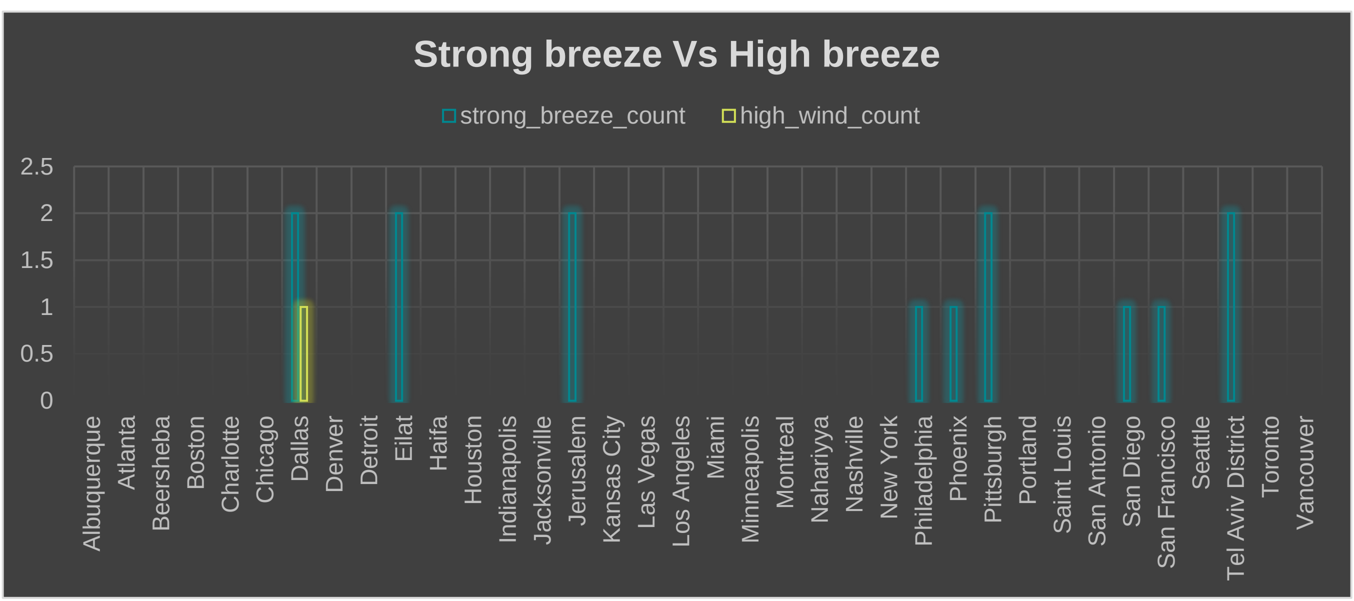
- Calm breeze
- Light breeze
- Moderate breeze
- Strong breeze
- High wind

City	calm_breeze_count	light_breeze_count	moderate_breeze_count	strong_breeze_count	high_wind_count
Albuquerque	**14904**	**29977**	**368**	**0**	**0**
Atlanta	**17069**	**28163**	**20**	**0**	**0**
Beersheba	**22616**	**21835**	**9**	**0**	**0**
Boston	**8843**	**36256**	**153**	**0**	**0**
Charlotte	**14941**	**30302**	**8**	**0**	**0**

This is the sample output of the query output. From this I created 3 column charts to compare them


Calm breeze vs Light breeze
calm_breeze_count
light_breeze_count
50000
40000
30000
20000
10000
0
Albuqu..
Atlanta
Beersh..
Boston
Charlotte
Chicago
Dallas
Denver
Detroit
Eilat
Haifa
Houston
Indiana..
Jackso..
Jerusal..
Kansas..
Las..
Los..
Miami
Minnea..
Montreal
Nahariy..
Nashville
New York
Philade..
Phoenix
Pittsbur..
Portland
Saint..
San..
San..
San..
Seattle
Tel..
Toronto
Vancou..


Moderate_breeze_count
600
500
400
300
200
100
0
Albuquerque
Atlanta
Beersheba
Boston
Charlotte
Chicago
Dallas
Denver
Detroit
Eilat
Haifa
Houston
Indianapolis
Jacksonville
Jerusalem
Kansas City
Las Vegas
Los Angeles
Miami
Minneapolis
Montreal
Nahariyya
Nashville
New York
Philadelphia
Phoenix
Pittsburgh
Portland
Saint Louis
San Antonio
San Diego
San…
Seattle
Tel Aviv…
Toronto
Vancouver
moderate_breeze_count


Strong breeze Vs High breeze
strong_breeze_count
high_wind_count
2.5
2
1.5
1
0.5
0
Albuquerque
Atlanta
Beersheba
Boston
Charlotte
Chicago
Dallas
Denver
Detroit
Eilat
Haifa
Houston
Indianapolis
Jacksonville
Jerusalem
Kansas City
Las Vegas
Los Angeles
Miami
Minneapolis
Montreal
Nahariyya
Nashville
New York
Philadelphia
Phoenix
Pittsburgh
Portland
Saint Louis
San Antonio
San Diego
San Francisco
Seattle
Tel Aviv District
Toronto
Vancouver

We got the cities with the count of different types of wind_speed. Now, let's analyse how can these wind patterns impact the air quality and pollution in urban area:

Calm Breeze (0-1.8 km/h):

Air Quality Impact	*During calm breeze conditions, air pollutants such as particulate matter (PM), nitrogen dioxide (NO2), and volatile organic compounds (VOCs) may accumulate near the ground due to limited dispersion. This can lead to poor air quality, especially in urban areas with high pollution sources.*
Pollution Effects	*Calm breeze conditions can exacerbate pollution levels, contributing to health issues such as respiratory problems, allergies, and cardiovascular diseases. Without sufficient wind to disperse pollutants, concentrations can build up, leading to smog and haze.*

Light Breeze (1.9-11.2 km/h):

Air Quality Impact	*Light breezes help disperse pollutants and improve air quality by enhancing ventilation and mixing of the atmosphere. This can reduce local pollution concentrations and mitigate the effects of emissions from traffic, industry, and other sources.*
Pollution Effects	*Light breeze conditions can still lead to moderate pollution levels, particularly in areas with significant emission sources. However, the impact on air quality is generally less severe compared to calm breeze conditions.*

Moderate Breeze (11.3-28.8 km/h):

Air Quality Impact	*Moderate breezes are highly beneficial for air quality as they facilitate efficient dispersion of pollutants. Increased wind speeds enhance atmospheric mixing, diluting pollution and reducing local concentrations.*
Pollution Effects	*Pollution levels are typically lower during moderate breeze conditions, resulting in improved air quality and reduced health risks for residents. However, sensitive individuals may still experience some adverse effects, especially in areas with persistent pollution sources.*

Strong Breeze (28.9-49.9 km/h):

Air Quality Impact	*Strong breezes provide effective ventilation and dispersion of pollutants, leading to further improvement in air quality. High wind speeds enhance atmospheric mixing, rapidly dispersing pollutants and minimizing their impact on local air quality.*
Pollution Effects	*Pollution levels are generally low during strong breeze conditions, with minimal health impacts on the population. The rapid dispersion of pollutants reduces exposure levels, contributing to better respiratory health and overall well-being.*

High Wind (50 km/h or higher):

Air Quality Impact	*High winds have a significant cleansing effect on the atmosphere, quickly dispersing pollutants and improving overall air quality. Intense mixing of the lower atmosphere with cleaner air from aloft helps reduce pollution concentrations at ground level.*
Pollution Effects	*High wind speeds result in very low pollution levels and excellent air quality. Exposure to harmful pollutants is minimal, leading to improved respiratory health and reduced health risks for the population.*

14. Identify cities prone to strong winds and the potential consequences, such as increased risk of natural disasters or challenges for transportation.

```
use weather_analysis_transformed;
create view a as
select cl.City as City_prone, ff.weather_description
from final_fact ff
join city_lookup cl
on cl.City_id = ff.City_id
where (ff.weather_description like "%thunderstorm%"
or ff.weather_description like "%tornado%"
or ff.weather_description like "%sand%"
or ff.weather_description like "%ash%"
or ff.weather_description like "%heavy%"
or ff.weather_description like "%dizzle%"
or ff.weather_description like "%snow%")
and ff.weather_description not like "%light%"
and ff.weather_description not like "snow"
```

City_prone	weather_description
Vancouver	heavy intensity rain
Vancouver	heavy intensity rain
Vancouver	heavy snow
Vancouver	heavy snow
Vancouver	heavy snow
Vancouver	heavy snow
Vancouver	heavy snow
Vancouver	heavy snow

Created a temporary table in MySQL to identify cities prone to strong winds and the potential consequences. Selected city, weather description where weather description are thunderstorm, tornado, sand & ash(heatwaves), heavy, dizzle, snow and not light & only snow.
In top, I provided the sample output which I've got from this query.
After that, I made a pivot table to fetch the cities & the all unique natural disaster.

After analyzing we saw all the cities are prone area.

From the pivot I got all these natural weather descriptions, this is a sample output.

Natural disaster
heavy intensity drizzle
heavy intensity rain
heavy intensity shower rain
heavy shower snow
heavy snow
heavy thunderstorm
proximity sand/dust whirls
proximity thunderstorm

I don't have enough data to show the potential consequences of such as increased risk of natural disasters or challenges for transportation. Despite this, I can provide a detailed breakdown of the process:

Strong winds can bring significant consequences for cities, impacting various aspects of life and infrastructure. Here are some potential consequences we can explore based on the data we have:

Increased Risk of Natural Disasters:

Damage to buildings and infrastructure	*Strong winds can cause structural damage to buildings, bridges, power lines, and other infrastructure. This can lead to power outages, disruptions in communication, and even building collapses.*
Flying debris	*High winds can cause loose objects like shingles, branches, and even furniture to become airborne, posing a danger to pedestrians and property.*
Flooding	*Strong winds can increase the severity of tropical cyclones like hurricanes for coastal cities in their path. This brings high winds, storm surge and flooding.*
Wildfires	*Areas with hot, dry winds are at higher risk of wildfires spreading rapidly. Winds fan and spread the flames.*
Land Famine	*Dust storms and soil erosion risks go up in desert regions with consistent wind patterns.*
Blizzard Threat	*Blizzards and whiteout conditions become more likely in areas prone to cold, icy winds.*

Challenges for Transportation:

Disruptions to air travel	*Air travel can be disrupted by takeoff and landing restrictions due to crosswinds. Flight delays and cancellations may increase. Also, Takeoff and landing of small aircraft may be prohibited during strong winds due to safety issues.*
Road closures	*High winds can blow over trees, debris, and even vehicles, blocking roads and creating dangerous driving conditions.*
Disruptions to public transportation	*Buses, trains, and trams might be delayed or cancelled due to strong winds.*
Vehicle Wobble	*Driving conditions become hazardous when winds are powerful enough to push vehicles off course. High profile vehicles are especially vulnerable.*
Waves on Hold	*Shipping, boating and other marine transport face restrictions and delays during powerful winds which churn up waves.*
Blindsided Travel	*Windblown dust and debris can severely reduce visibility for all transportation modes.*
Cabled Chaos	*Extreme winds can prevent safe operation of cranes, trams, aerial tramways and other modes relying on cables.*

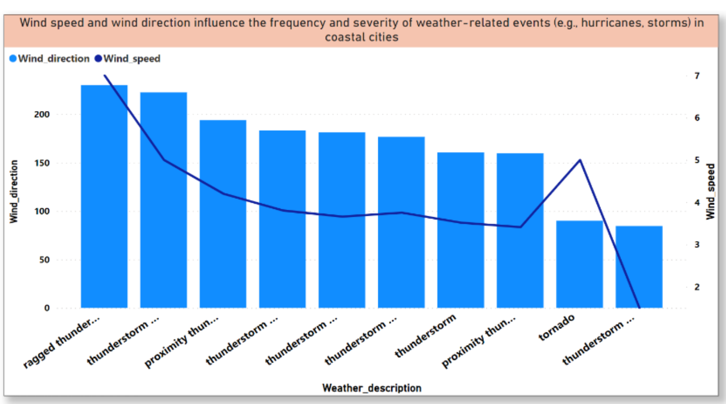
Explore whether wind speed and direction influence the frequency and severity of weather-related events (e.g., hurricanes, storms) in coastal cities.

```
select cl.City_id, cl.City, ff.wind_direction, ff.wind_speed,
ff.weather_description, monthname(dl.date) as Month_Name
from final_fact ff
join city_lookup cl
on ff.City_id = cl.City_id
join date_lookup dl
on ff.date_id = dl.date_id
where  (  ff.weather_description like '%thunderstorm%'
or ff.weather_description like '%tornado%'
    or ff.weather_description like '%storm%'
    or ff.weather_description like '%hurricanes%')
    and cl.City_id in (1,2,3,4,5,6,13,22,24,30,32,33)
    order by 1 asc
```

City ID	Coastal Cities
1	Vancouver
2	Portland
3	San Francisco
4	Seattle
5	Los Angeles
6	San Diego
13	Houston
22	Jacksonville
24	Miami
30	Boston
32	Tel Aviv District
33	Eilat

- After exploring and researching this dataset I got 12 coastal cities, here I presented these cities in tabular form.
- To fetch whether wind speed and direction influence the frequency and severity of weather-related events (e.g., hurricanes, storms) in coastal cities, I made a query where I selected the city id, city, weather direction, wind speed, weather description & extracted month name from the date.
- Where the city id are the coastal city id and the weather descriptions are thunderstorm, tornado, storm, hurricane.
- Then I created a pivot table in excel to fetch the unique weather events like hurricanes, storms.

Weather events(hurricanes, storms)
proximity thunderstorm
proximity thunderstorm with rain
ragged thunderstorm
thunderstorm
thunderstorm with drizzle
thunderstorm with heavy rain
thunderstorm with light drizzle
thunderstorm with light rain
thunderstorm with rain
tornado



I created a combo chart in PowerBI to visualize the relationship of wind speed, wind direction & weather description. Chart reference above.

Here we can see except Tornado, all kind of weather events are in positive relationship with wind direction and wind speed.

Again another pivot I created where I showed the coastal cities along with their average wind direction & speed, Standard deviation of wind direction & speed, MAX & MIN wind direction & speed to analyze the influence deeply.

Coastal city	Average of wind_direction	Average of wind_speed	StdDev of wind_direction	StdDev of wind_speed	Max of wind_direction	Max of wind_speed	Min of wind_direction	Min of wind_speed	Count of weather_description
Miami	**149.3480519**	**3.22**	**87.85**	**1.82**	**360**	**11**	**0**	**0**	**1155**
Jacksonville	**153.944793**	**3.67**	**88.84**	**2.12**	**360**	**14**	**0**	**0**	**797**
Houston	**161.6235294**	**3.58**	**92.42**	**2.07**	**360**	**15**	**0**	**0**	**680**
San Francisco	**206.538961**	**3.83**	**105.34**	**2.43**	**360**	**13**	**0**	**0**	**308**
Boston	**186.4487179**	**3.59**	**84.06**	**2.15**	**350**	**12**	**0**	**0**	**156**
Tel Aviv District	**185.8157895**	**4.39**	**91.58**	**3.01**	**360**	**15**	**0**	**0**	**114**
Los Angeles	**174.4237288**	**3.14**	**105.75**	**2.12**	**340**	**9**	**0**	**0**	**59**
Seattle	**202.7142857**	**3.31**	**80.92**	**1.98**	**359**	**9**	**20**	**1**	**42**
Portland	**204**	**2.67**	**105.12**	**1.81**	**330**	**7**	**0**	**0**	**33**
San Diego	**217.137931**	**2.66**	**94.50**	**1.80**	**340**	**9**	**0**	**1**	**29**
Vancouver	**180.3461538**	**3.96**	**94.88**	**2.09**	**340**	**8**	**0**	**1**	**26**
Eilat	**196.7391304**	**3.65**	**143.93**	**1.90**	**360**	**8**	**0**	**0**	**23**

Conclusion:

After exploring wind speed and direction and severity of weather-related events (e.g., hurricanes, storms) in coastal cities, we can say there are 10 weather events like hurricanes, storms, along with rainfall in some cities. Miami has maximum variation in the change of climate and Eilat has minimum variation in the change of climate.
Also, the mean Wind_direction of these cities lying between 149-218 degree and mean Wind_speed lying between 3.14-4.40 km/h
Also, the Standard_deviation of Wind_direction lying between 80.92-143.93 degree and the Standard_deviation of Wind_speed lying between 1.80-3.01km/h
Also, we get the maximum and minimum wind_direction and wind_speed of each coastal cities.
By this historical data, we can prevent any big disaster in future for coastal cities.

Conclusion about my understanding about this data

Geographic Influences:

- Recognized the impact of longitude and latitude on weather variables like temperature and humidity.
- Understood the negative correlation between longitude and temperature, indicating lower temperatures farther east.
- Acknowledged the positive correlation between latitude and temperature, with temperatures decreasing towards higher latitudes.
- Noted the minor influence of longitude and latitude on humidity, understanding that other factors play a more significant role.

Complex Interplay of Factors:

- Understood the multifaceted nature of weather analysis, considering factors such as proximity to water bodies, altitude, and vegetation cover.
- Recognized that a combination of factors contributes to regional weather patterns.

Critical Thinking and Interpretation:

- Demonstrated analytical skills in interpreting the direction and strength of correlations.
- Awareness of the limitations of correlations and the need to consider other influencing factors.

Application of Knowledge:

- Ability to apply weather analysis data in practical contexts such as agriculture, urban planning, and disaster preparedness.
- Utilized weather analysis data to inform decision-making processes and address challenges related to weather variability and climate change.

Thank you ***for taking the time to review this document. Should you have any further questions, require clarification, or wish to discuss any aspect in more detail, please do not hesitate to reach out. Your feedback and input are highly valued and appreciated.***

Best regards,
Moumita Biswas
DA 5th Batch