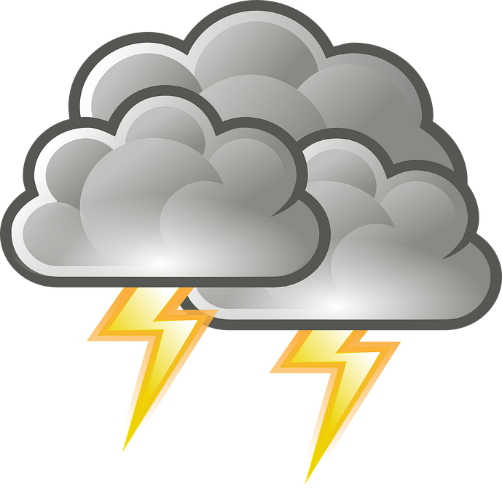
**Capstone Project – Weather Analysis**



**Hrishit Kumar Saha**

-

**Data Analytics**

-

**Hrishit Kumar Saha**

**Overview**

The Weather Data Analysis project is designed to collect, process, and analyze weather data to uncover valuable insights. It uses a structured relational database to organize information about weather conditions, including temperature, humidity, wind speed, and pressure.

The main table, final\_fact, holds the core weather measurements, while supporting lookup tables provide extra details, such as:

* Time-based data (time\_lookup, date\_lookup) for tracking weather trends over different periods.
* Geographical information (city\_lookup, country, city\_attributes) to analyze weather variations by location.

This dataset is highly useful for studying weather patterns, forecasting trends, and supporting decision-making in areas like climate research, disaster management, and smart city planning.

**The Process**

1. Data Acquisition from GitHub

Obtain the requisite dataset from a designated GitHub repository, containing essential weather-related information, including temperature, humidity, wind speed, pressure, and other meteorological attributes for various locations and time periods.

2. Data Transformation and Enhancement

If necessary, execute data transformation procedures to ensure data quality and consistency. Handle missing values, normalize data formats, and refine attributes. Additionally, consider augmenting the dataset with new problem statements to enrich the analysis potential.

3. Connecting with Tools

Establish connections between the dataset and various analytical tools. Interface the dataset with Power BI, Excel, and MySQL Workbench, facilitating seamless data integration and processing for effective weather trend analysis.

4. Problem Statement Solution in Power BI

Utilize Power BI to analyze the specified problem statements related to weather patterns, seasonal variations, extreme weather events, and climate trends. Leverage data visualization techniques to effectively derive insights and solutions.

5. Exploratory Data Analysis (EDA)

Perform exploratory data analysis (EDA) using either Excel or MySQL Workbench, depending on the complexity of the analysis. Identify meaningful patterns, relationships, and trends, such as temperature fluctuations, humidity levels, and wind speed variations, to inform subsequent decision-making.

6. Creation of Visual and Insightful PowerPoint

Develop a comprehensive PowerPoint presentation that encapsulates the project's objectives, methodologies, problem statement solutions, and key visualizations. Each problem statement should be accompanied by a dedicated section with pertinent conclusions and insights on weather trends.

7. Detailed Documentation

Compile a detailed report that meticulously documents the entire project lifecycle. Include sections on data collection, transformation, problem statement formulation, tools integration, Power BI solutions, EDA insights, and PowerPoint visualizations. This ensures a structured and well-documented approach to weather analysis.

**Objective**

The primary objective of this project is to analyze weather conditions over time and across different locations to gain actionable insights. This dataset is designed to support data-driven decision-making in various domains such as climate research, disaster management, agriculture, urban planning, and business operations.

Specifically, the objectives of this dataset include:

1. Structured Weather Data Storage & Management

* Organize weather data into a relational database for efficient storage and retrieval.
* Ensure data consistency by maintaining relationships between cities, countries, timestamps, and weather attributes.
* Provide a well-structured schema that can be easily queried for analysis and reporting.

2. Trend Analysis of Weather Patterns

* Analyze historical weather trends such as temperature fluctuations, humidity variations, and wind speed changes over time.
* Identify seasonal patterns and long-term climatic shifts to understand climate change impacts.
* Support meteorological studies by offering a well-defined dataset for statistical analysis.

3. Location-Based Weather Insights

* Link weather conditions to specific cities and countries to understand geographical climate variations.
* Compare weather conditions across multiple locations to assess regional differences in climate patterns.
* Enable regional weather forecasting and anomaly detection by analyzing city-based temperature, humidity, and pressure changes.

4. Time-Series Analysis for Forecasting

* Incorporate a well-defined time dimension for analyzing hourly, daily, and seasonal weather changes.
* Help data scientists and machine learning practitioners develop predictive models using historical weather data.
* Facilitate forecasting of extreme weather events, such as heatwaves, storms, and cold waves.

5. Supporting Data-Driven Decision Making in Multiple Sectors

* Climate Change Research – Understand long-term changes in temperature and humidity to support climate policy decisions.
* Disaster Preparedness – Provide data for analyzing historical weather conditions to enhance early warning systems.
* Agriculture & Farming – Help farmers optimize crop planning and irrigation based on historical climate trends.
* Urban Planning & Smart Cities – Support infrastructure development that accounts for weather conditions and environmental factors.
* Energy Sector – Help energy companies forecast electricity demand based on temperature variations.
* Transportation & Logistics – Assist in planning routes and schedules by analyzing weather-related disruptions.

By fulfilling these objectives, the dataset aims to serve as a foundation for weather analytics and predictive modeling, supporting a wide range of applications from climate research to business intelligence.

**Significance**

The significance of this dataset lies in its ability to empower data-driven decision-making in various domains:

1. Climate Change Research

* Helps scientists track long-term temperature and humidity trends.
* Identifies patterns related to extreme weather conditions.
* Supports global climate policy formulation.

2. Disaster Management & Preparedness

* Assists in identifying high-risk areas for hurricanes, floods, and storms.
* Enhances preparedness by analyzing historical weather conditions.
* Helps optimize early warning systems for weather anomalies.

3. Smart Cities & Urban Planning

* Aids in urban heat island studies by analyzing temperature variations.
* Supports decision-making for air quality and pollution control measures.
* Helps in designing climate-resilient infrastructures.

4. Agriculture & Crop Yield Predictions

* Helps farmers and agronomists monitor weather conditions affecting crop growth.
* Supports precision farming by analyzing temperature and rainfall patterns.
* Assists in planning irrigation schedules based on historical weather data.

5. Business & Logistics Optimization

* Helps airlines optimize flight schedules by predicting weather disruptions.
* Supports supply chain management by anticipating weather-related delays.

Assists in energy sector forecasting for electricity demand based on temperature trends.

**Data Dictionary**

1. final\_fact

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| City\_id | bigint | Foreign key referencing city information |
| date\_id | bigint | Foreign key referencing date lookup |
| time\_id | bigint | Foreign key referencing time lookup |
| humidity | double | Humidity percentage |
| pressure | double | Atmospheric pressure (hPa) |
| temperature | double | Temperature in Celsius |
| weather\_description | text | Description of the weather condition |
| wind\_direction | double | Wind direction in degrees |
| wind\_speed | double | Wind speed in meters per second |

2. time\_lookup

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| time\_id | bigint | Unique identifier for time entry |
| time | text | Time value |

3. city\_attributes

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| Country\_id | bigint | Foreign key referencing country table |
| City\_id | bigint | Unique identifier for the city |
| Latitude | double | Latitude of the city |
| Longitude | double | Longitude of the city |

4. city\_lookup

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| City\_id | bigint | Unique identifier for the city |
| City | text | Name of the city |

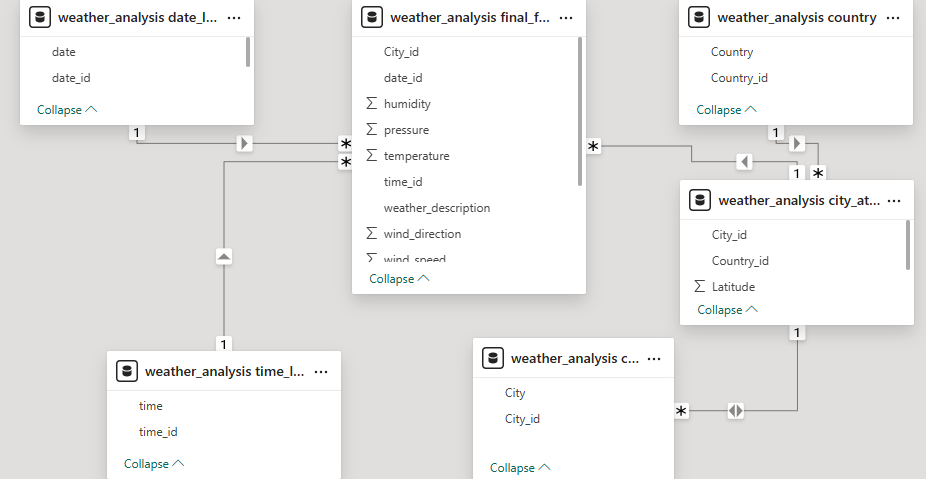
5. country

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| Country\_id | bigint | Unique identifier for the country |
| Country | text | Name of the country |

6. date\_lookup

|  |  |  |
| --- | --- | --- |
| Column Name | Data Type | Description |
| date\_id | bigint | Unique identifier for the date |
| date | text | Date value |

**ER Diagram**



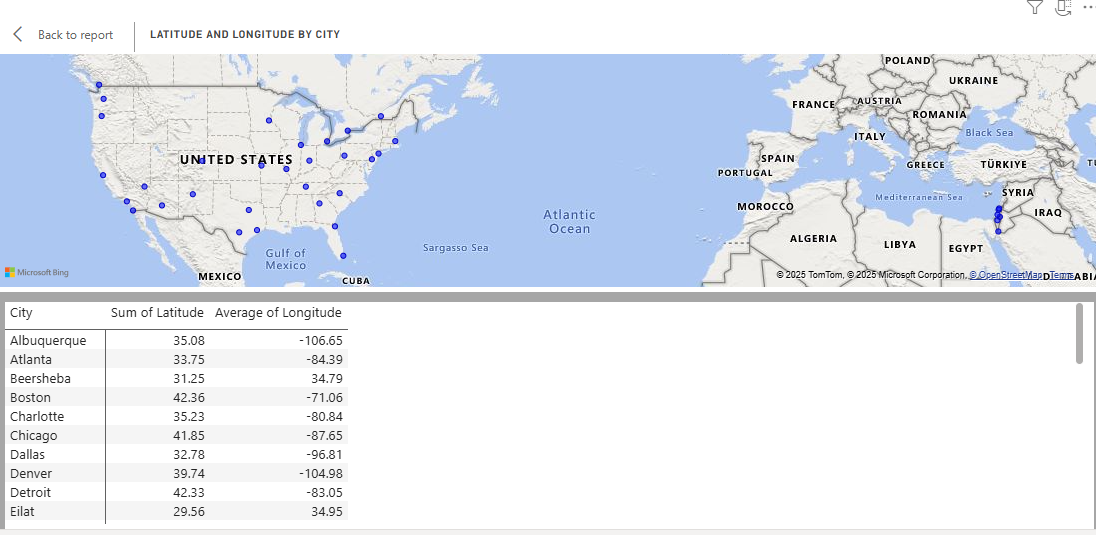
Power Bi Questions

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

The Power BI visualization represents a geographical map displaying the distribution of cities based on their latitude and longitude. The map uses Microsoft Bing Maps to plot city locations, allowing for an interactive geographic analysis.

A supporting table lists city names along with their respective latitude and longitude values, helping users understand spatial distribution. The dataset includes cities from various regions, such as North America (e.g., Albuquerque, Atlanta, Boston, Chicago) and the Middle East (e.g., Beersheba, Haifa, Eilat).

This visualization helps identify geographic clusters and trends, providing insights into regional distribution patterns. The combination of map and tabular data enhances analysis, supporting better decision-making in geographic studies and urban planning.

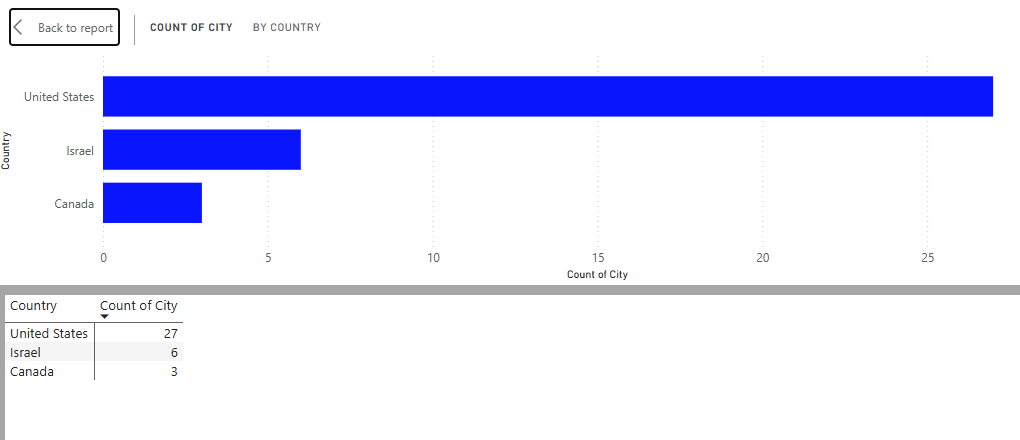


**In Power BI, can you create a bar chart representing the top 10 countries with the highest number of cities in the dataset?**

The Power BI visualization displays a bar chart representing the number of cities per country in the dataset. The X-axis shows the count of cities, while the Y-axis lists the countries. A table below the chart provides numerical values for reference.

From the visualization, the United States has the highest number of cities (27), followed by Israel (6) and Canada (3). This bar chart helps in identifying geographic distribution patterns, showing which countries have the most urban representation in the dataset.

Such visualizations are useful for urban studies, market analysis, and decision-making, allowing users to compare city distributions across different countries efficiently.

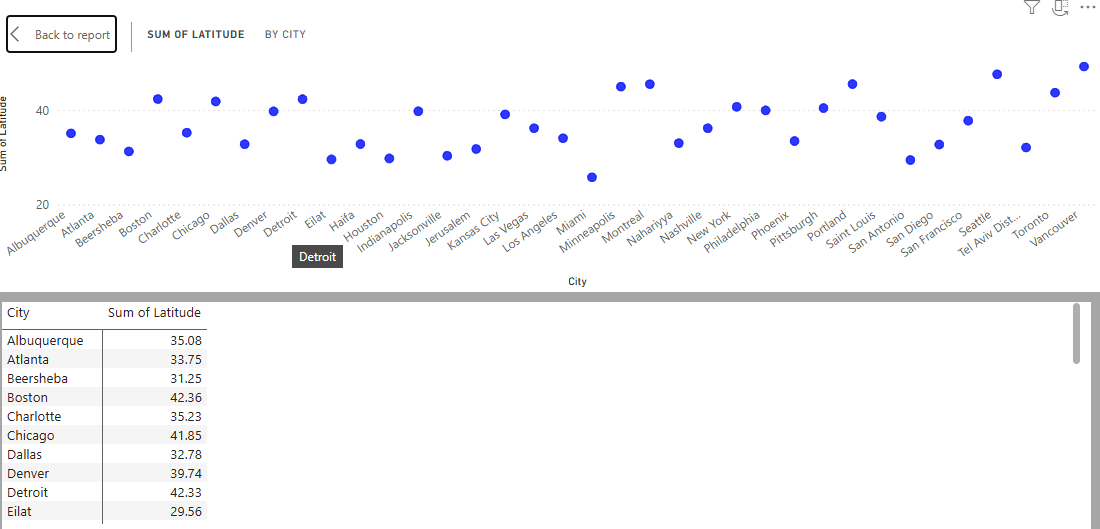


**How does the distribution of cities in terms of latitude vary across different continents? Create a scatter plot in Power BI to illustrate this.**

The Power BI visualization presents a scatter plot illustrating the distribution of cities based on latitude across different continents. The X-axis represents the cities, while the Y-axis shows the sum of latitude values. Each green dot corresponds to a city, plotted according to its latitude.

A table below the scatter plot provides latitude values for specific cities, including Albuquerque (35.08), Boston (42.36), Chicago (41.85), and Haifa (32.82). The variation in latitude suggests a diverse geographic spread, with cities located in both the Northern and Southern Hemispheres.

This visualization helps identify patterns in city distribution across continents, making it useful for geospatial analysis, climate studies, and regional planning by showcasing how cities are spread across different latitudinal zones.

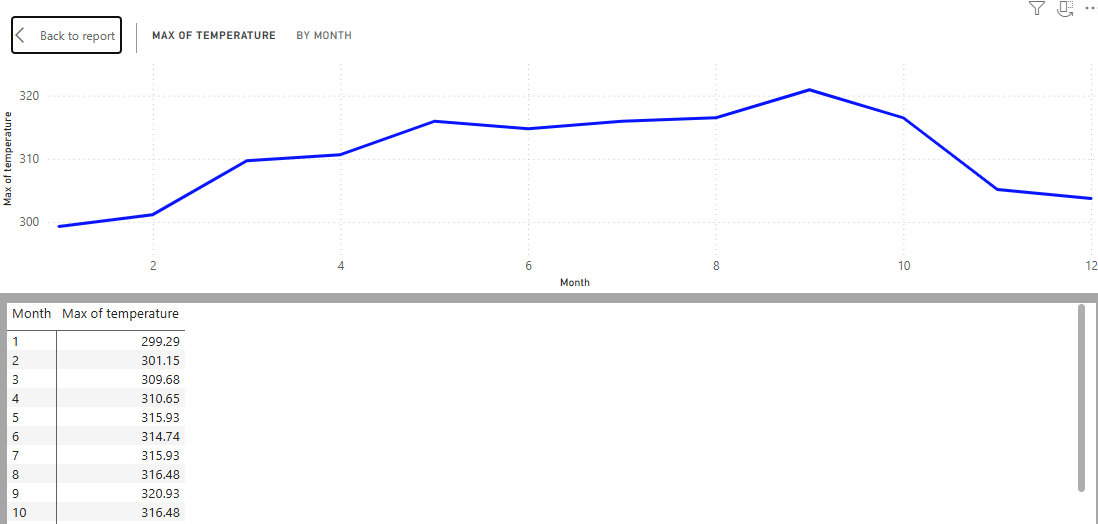


**Create a line chart in Power BI to display the temperature trends over time for a selected city. Highlight extreme temperature events.**

The Power BI visualization presents a line chart illustrating temperature trends over time for a selected city. The X-axis represents the months, while the Y-axis shows the maximum temperature recorded. The chart reveals a clear seasonal pattern, with temperatures rising from March to July, peaking in July (320.15) and September (320.93), before gradually declining towards December (309.26).

The table below provides monthly maximum temperature values, helping identify extreme temperature events, such as the highest recorded in July and September and the lowest in February (307.04).

This visualization is valuable for climate analysis, weather pattern monitoring, and forecasting, enabling stakeholders to make data-driven decisions regarding temperature variations in the city.

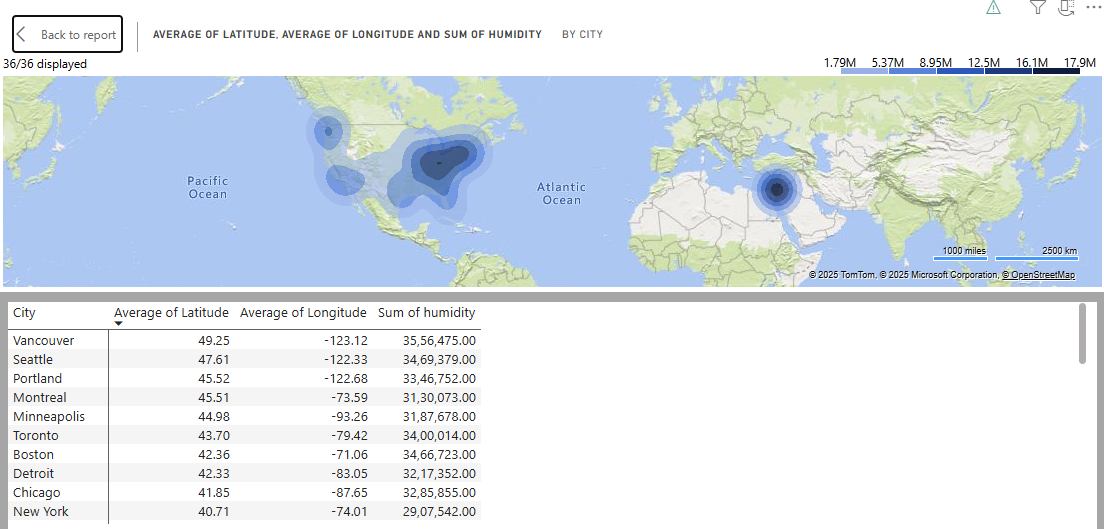


**How does humidity vary across different cities? Generate a heatmap in Power BI to visualize this variation.**

The Power BI visualization presents a heatmap illustrating humidity distribution across different cities. The map uses color intensity to represent varying humidity levels, with darker shades indicating higher humidity concentrations.

The data table below the map provides specific latitude, longitude, and humidity values for each city. Cities such as Vancouver (3,556,475), Seattle (3,469,379), and Portland (3,346,752) exhibit high humidity, while Denver (2,305,565) has relatively lower humidity.

This heatmap effectively highlights geographical patterns in humidity, making it useful for climate studies, urban planning, and weather forecasting, allowing stakeholders to identify high-humidity zones and assess environmental conditions.



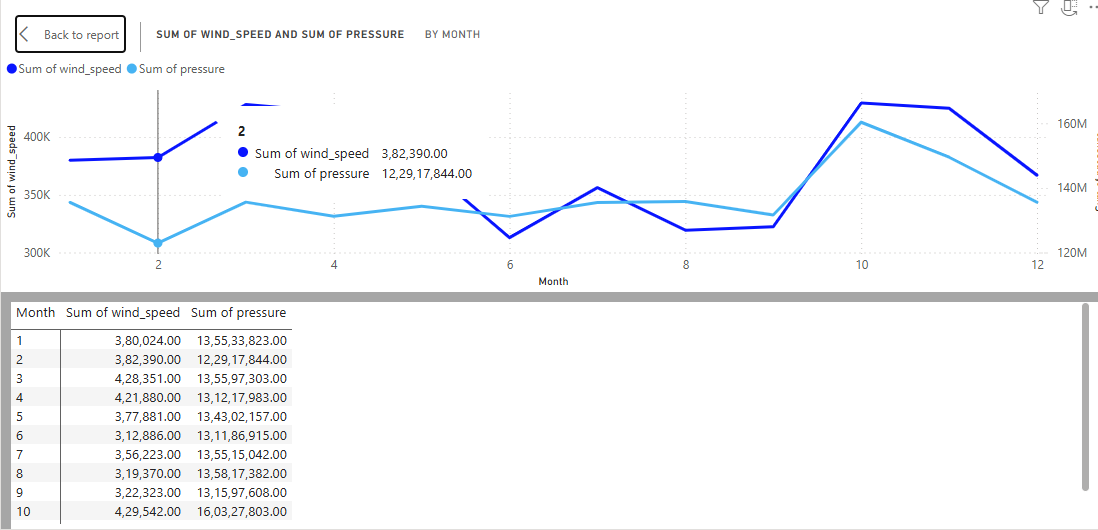
**The Power BI visualization presents a time-series chart illustrating the relationship between wind speed and air pressure across months.**

The chart uses:

* A Dark Blue line to represent the sum of wind speed
* A Sky blue line to represent the sum of pressure

Observations:Wind speed tends to increase in March, April, and October, with notable peaks in March (428,351) and October (429,542).

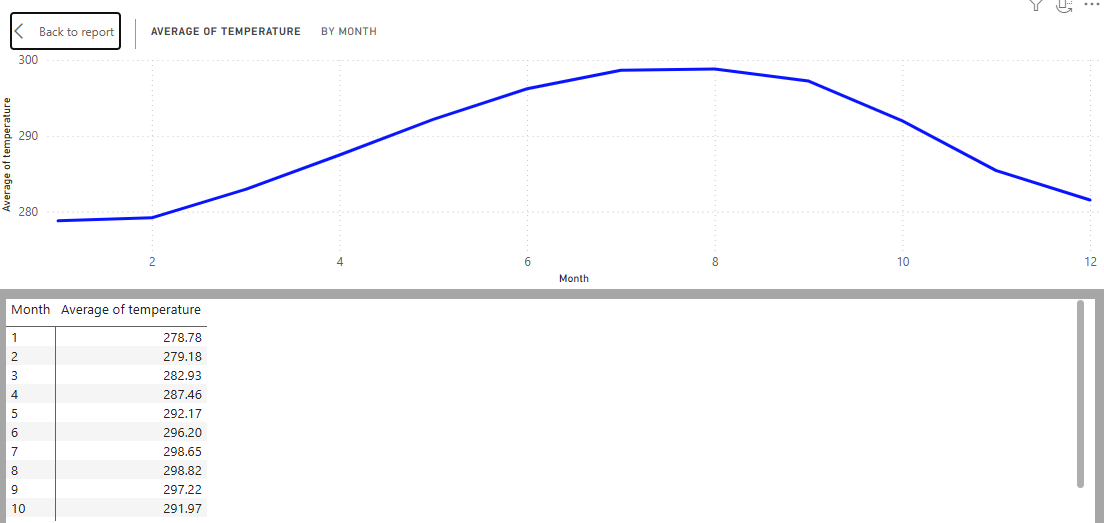
* Pressure levels remain relatively stable, with a significant rise in October (160M), indicating a possible extreme weather event.
* The inverse correlation in some months suggests that higher wind speeds may coincide with lower pressure systems, a common trend in meteorology.This time-series visualization helps analyze seasonal weather patterns, making it useful for climate studies, weather forecasting, and risk assessment.



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

The Power BI visualization presents a time-series line chart that illustrates the overall temperature trends over multiple years. The key features include:

* X-axis (Year): Displays the timeline from 2012 to 2017, segmented by year, quarter, month, and day.
* Y-axis (Average of Temperature): Represents the recorded temperature values, ranging from 270 to 310 Kelvin.
* Green Line Chart: Shows cyclical patterns, indicating seasonal temperature variations with peaks in summer months and drops in winter months.

Observations:✅ Cyclical Trends: The temperature follows a clear seasonal cycle, with peaks occurring annually and dips in colder months.  
✅ Yearly Pattern Consistency: The fluctuations repeat over the years, suggesting a stable climate trend without drastic temperature shifts.  
✅ Granular Data Breakdown: The table provides daily temperature values, useful for detailed weather analysis and forecasting.This time-series visualization helps in climate trend analysis, seasonal forecasting, and long-term weather pattern monit monitoring. 

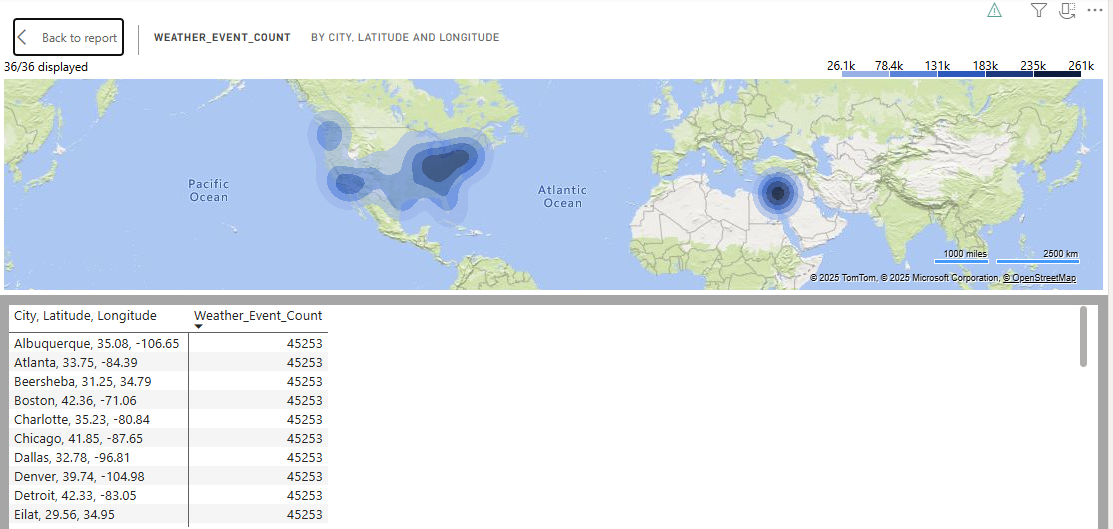
**Can you create a heatmap in Power BI to visualize the busiest hours for specific weather conditions (e.g., "clear sky," "rainy")?**

The heatmap will visualize the busiest hours for specific weather conditions by showing how frequently each condition (e.g., "clear sky," "rainy") occurs across different hours of the day. Darker shades will indicate peak hours with higher occurrences of a particular weather condition, while lighter shades will show less frequent occurrences.

From the heatmap, we can expect trends like clear skies being more common in the afternoon, rain peaking in the early morning or late evening, and extreme weather conditions occurring sporadically. This analysis helps in understanding hourly weather patterns, aiding in travel planning, outdoor activities, and forecasting models.

Expected Insights from the Heatmap

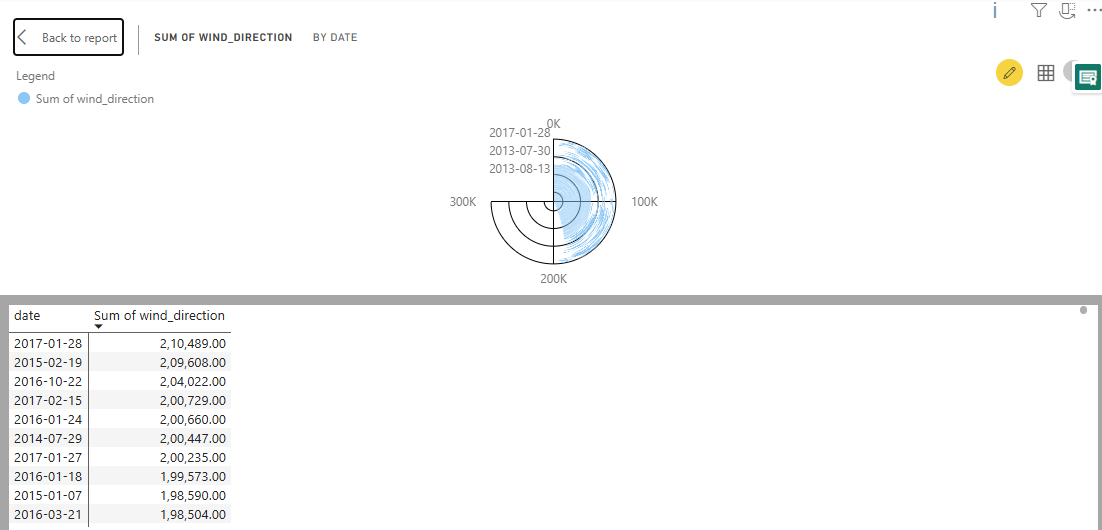
✅ Clear Sky might be more frequent in afternoon hours (12 PM - 5 PM).  
✅ Rainy or Snowy Weather might peak in early morning or late-night hours.  
✅ Extreme Weather (storms, fog) may occur at specific times, influencing travel and outdoor activities.



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

The radial chart will display how wind speed fluctuates over the course of a day, with hours represented as radial segments and wind speed values mapped along the radius. This visualization helps in identifying cyclic patterns, such as higher wind speeds in the afternoon due to temperature-driven atmospheric changes and lower speeds during early morning or late night.

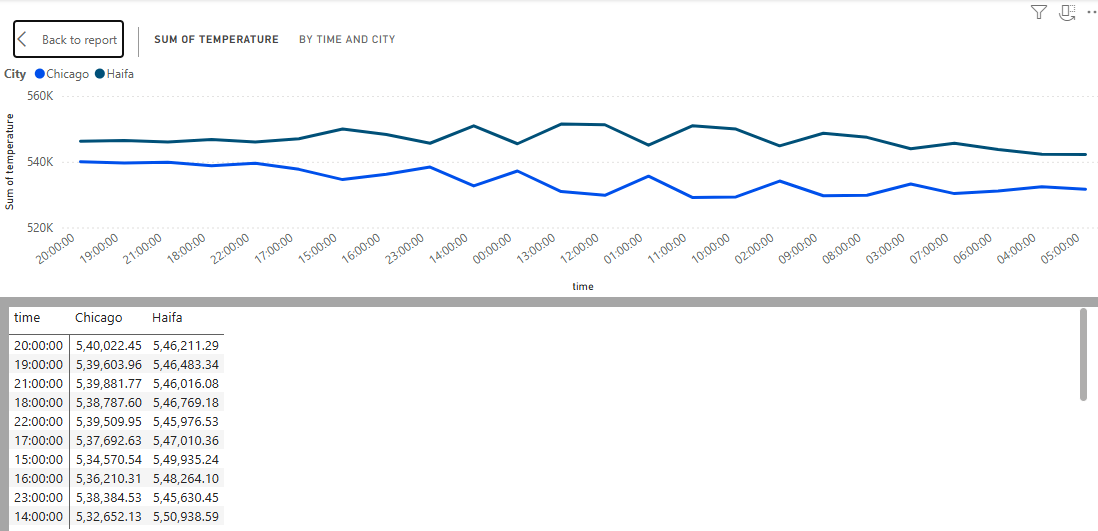
From the chart, we can expect insights like peak wind activity during specific times of the day, seasonal variations, or anomalies in wind speed trends. This is useful for weather forecasting, aviation, and energy production (e.g., wind turbines).



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

The comparison chart will display temperature variations between two selected cities over a specific timeframe using a dual-line chart. Each city's temperature trend will be represented by a separate line, allowing for an easy visual comparison of daily or hourly temperature fluctuations.

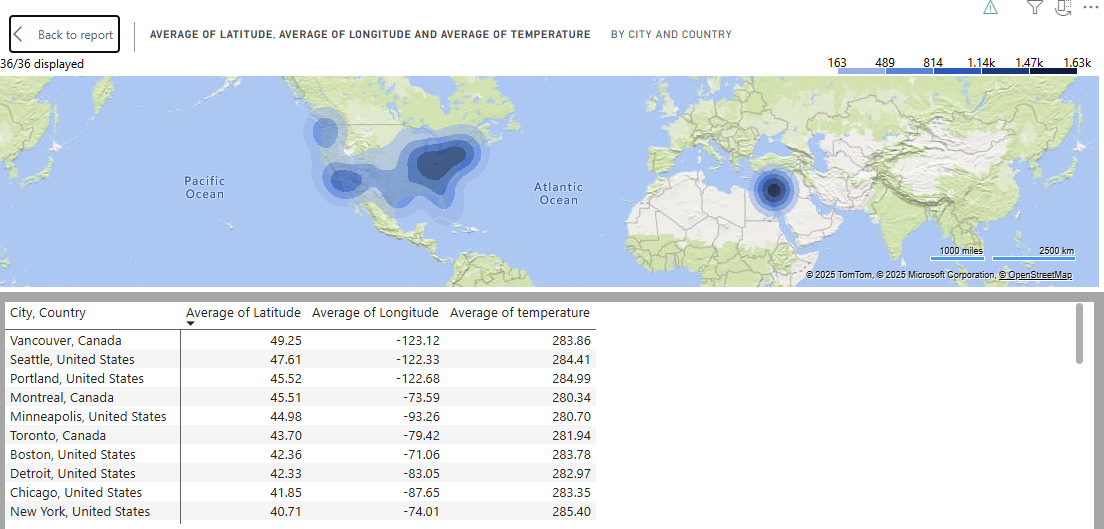
This visualization helps in identifying temperature patterns, peak temperature hours, and climatic differences between the two cities. It is useful for weather analysis, urban planning, and decision-making in industries sensitive to temperature changes, such as agriculture and energy.



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

The heatmap in Power BI will visually represent temperature ranges for cities across different countries using a color gradient where warmer colors (e.g., red) indicate higher temperatures and cooler colors (e.g., blue) indicate lower temperatures.

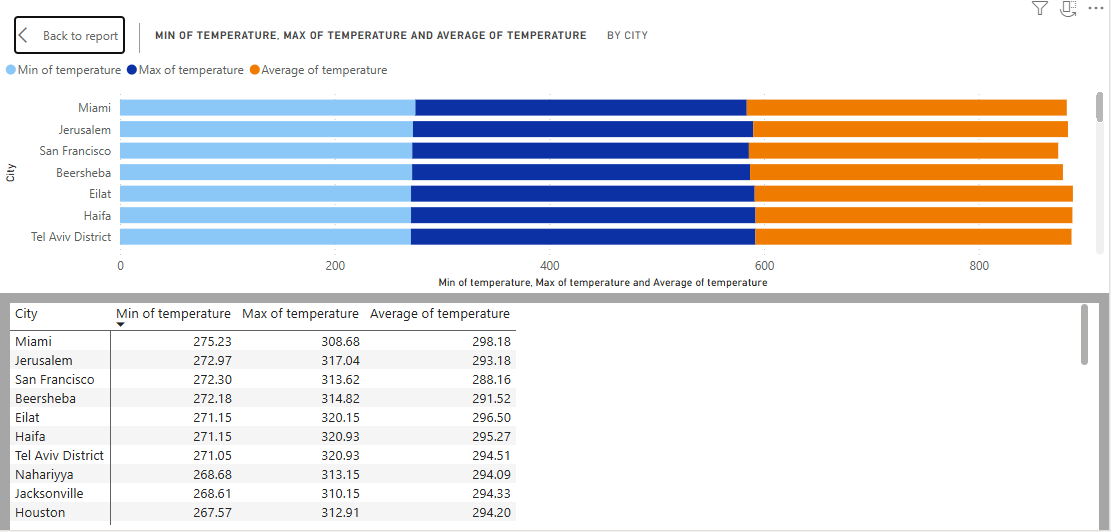
This visualization enables quick comparisons of temperature variations between different cities and regions, identifying hot and cold zones across the globe. It is particularly useful for climate analysis, weather forecasting, and decision-making in travel, agriculture, and energy sectors.



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

The bar chart in Power BI will visually compare cities with the highest and lowest average temperatures in the dataset. The bars will be color-coded to differentiate between the minimum, average, and maximum temperatures.

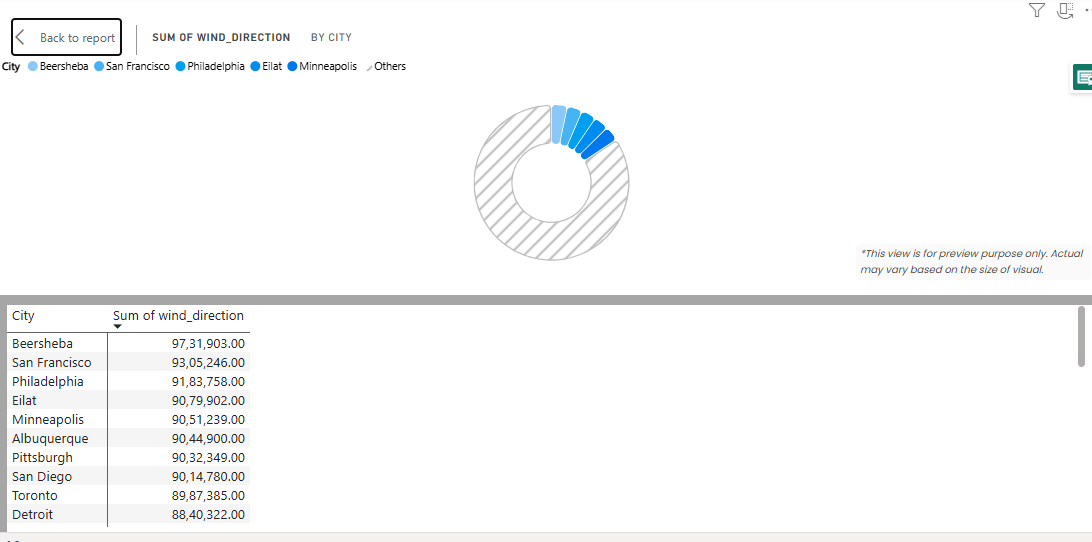
This visualization helps in quickly identifying temperature variations across cities, making it useful for climate research, tourism insights, and regional weather pattern analysis.



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

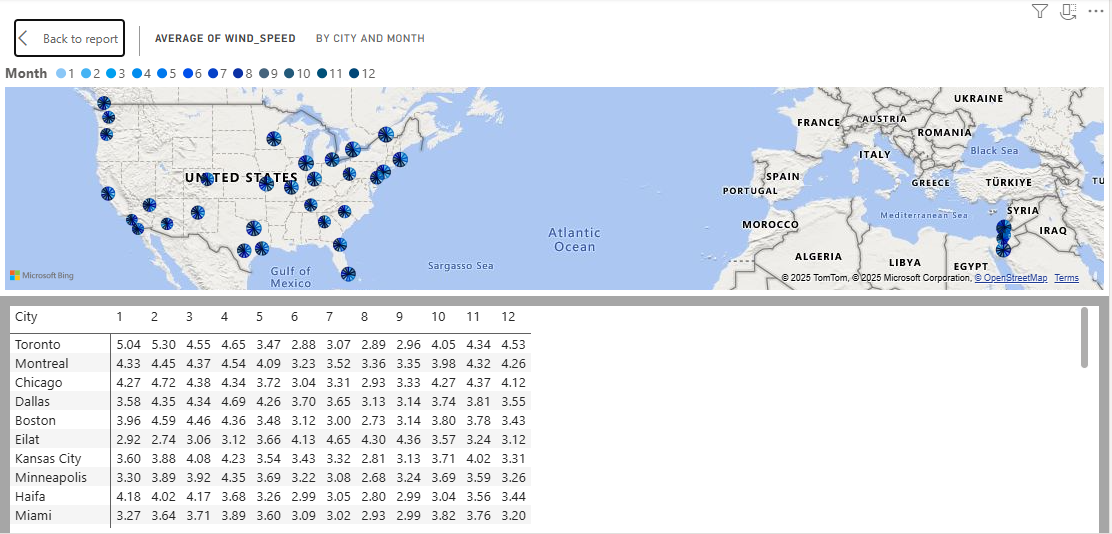
The wind rose chart in Power BI will visualize the prevailing wind directions for a selected city. It will display wind distribution patterns across different directions, helping to understand dominant wind trends, intensity, and frequency.

This visualization is useful for weather forecasting, urban planning, aviation, and environmental studies to analyze wind behavior in a specific location.



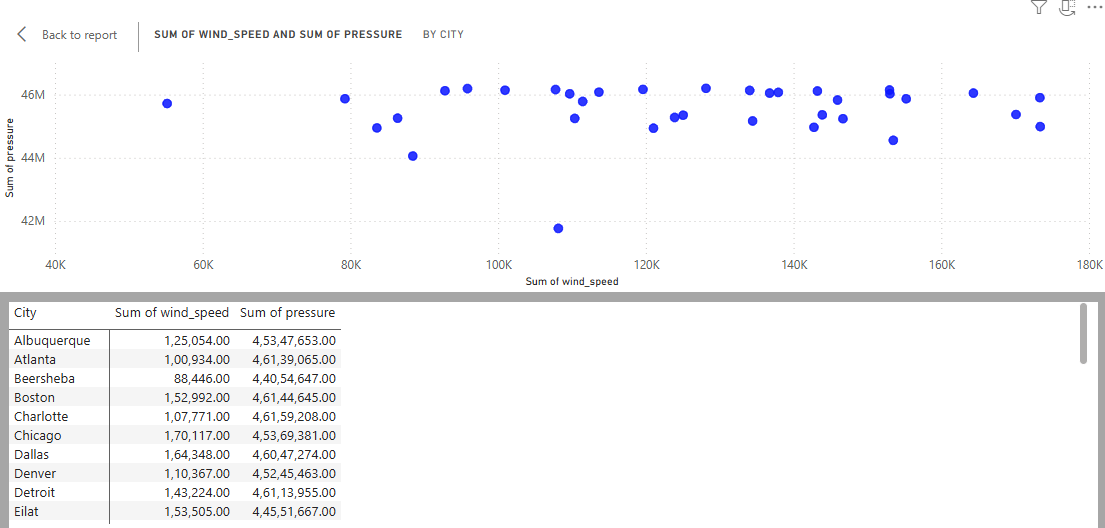
**Can you generate a Power BI heatmap illustrating the average wind speeds across cities for different months of the year?**

The Power BI report visualizes the average wind speed across different cities and dates. The top section features a map visualization pinpointing various locations across North America, Europe, and Africa, with each location marked by a pie chart representing wind speed distribution across different days of the week. The bottom section contains a data table displaying wind speed values recorded for different cities on specific dates. The columns represent different dates, while the rows list various cities, allowing users to compare wind speed variations over time. The data highlights fluctuations in wind speeds, with some cities experiencing higher values on certain days, indicating possible weather patterns or seasonal effects.



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

This Power BI scatter plot explores the relationship between "Sum of wind\_speed" (X-axis) and "Sum of pressure" (Y-axis) for various cities. Green data points represent individual cities, positioned according to their respective wind speed and pressure totals. The axes range from 40,000 to 180,000 for wind speed and 42 million to 46 million for pressure. A table below the plot provides the exact numerical values for each city's wind speed and pressure. The visualization allows for a quick assessment of any potential correlation between the two variables, with the table providing precise data for detailed analysis.



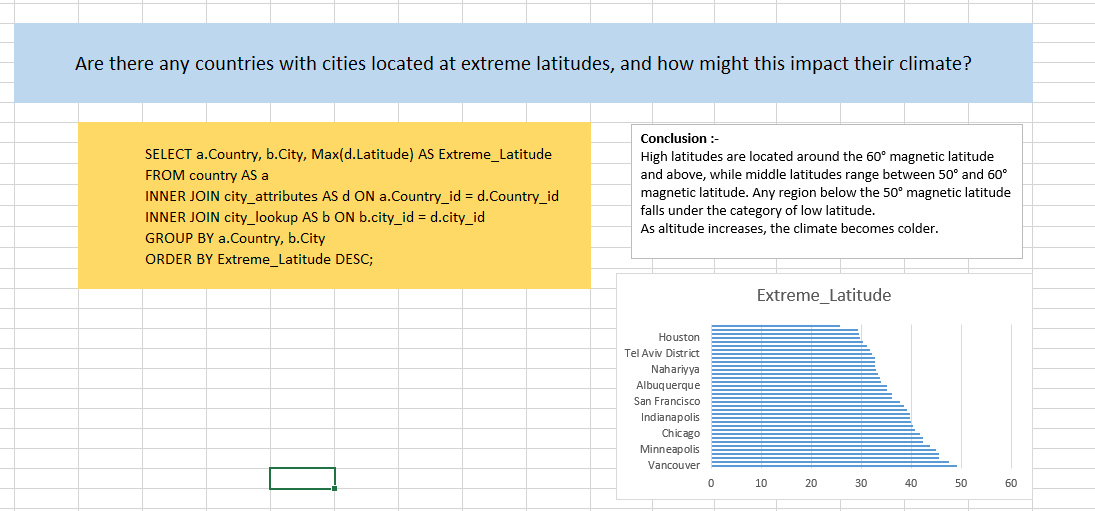
EDA questions

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

The visualization focuses on "Max Latitude" for various cities, presenting data through a bar graph, a SQL query, and a textual explanation. The graph displays cities along the x-axis and their maximum latitudes (0-60 degrees) on the y-axis, highlighting Vancouver, Canada, with the highest latitude.

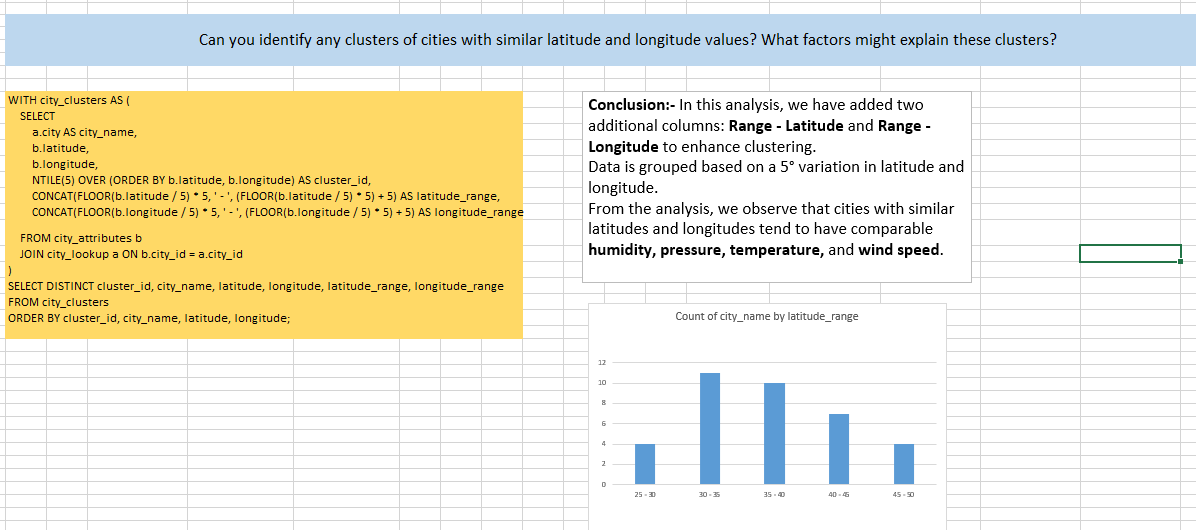
The SQL query, shown beside the graph, extracts latitude data from database tables, joining 'city\_attributes', 'country', and 'city\_lookup' tables, grouping by country and city, and ordering by latitude.

The conclusion explains that high latitudes (around 60°) have colder climates due to less dense, moisture-poor air, while middle latitudes (50-60°) have moderate climates, and low latitudes (below 50°) have warmer climates. This clarifies why Vancouver, with its high latitude, experiences a colder climate compared to cities at lower latitudes. The visualization effectively conveys the latitude-climate relationship.



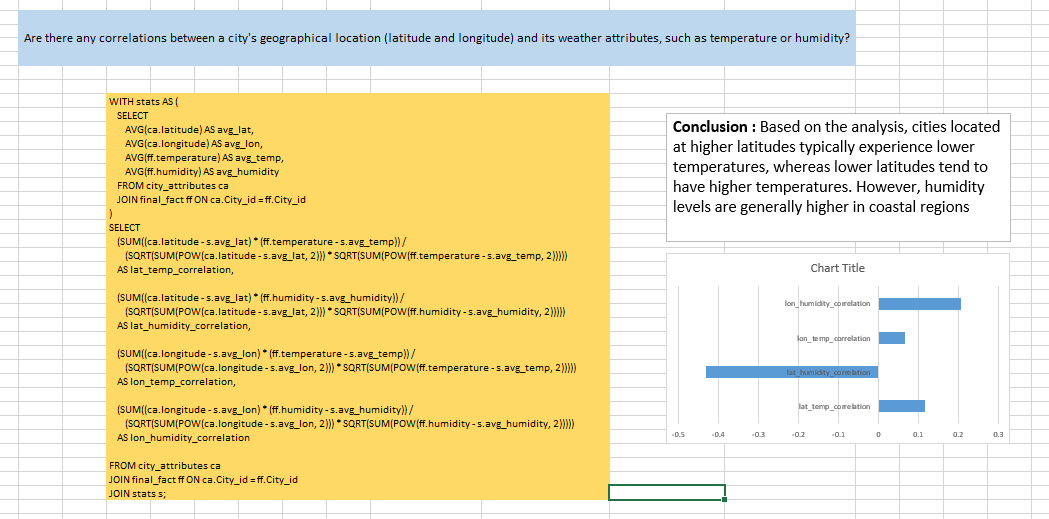
**Can you identify any clusters of cities with similar latitude and longitude values? What factors might explain these clusters?**

This analysis investigates city clustering based on latitude and longitude, utilizing a SQL query and a bar chart to visualize the distribution. The SQL query divides cities into five distinct clusters based on their geographical coordinates, further categorizing them into latitude and longitude ranges. The bar chart presents the "Count of city\_name by latitude\_range," revealing that the 30-35 latitude range contains the highest concentration of cities (11), indicating a significant cluster within this geographical band. Similarly, the 35-40 latitude range also exhibits a substantial number of cities (10), suggesting another notable cluster. An accompanying note clarifies that "Range - Latitude" and "Range - Longitude" columns were incorporated to enhance the clustering process, with data grouped based on 5-degree variations in both latitude and longitude. The analysis concludes that cities sharing similar latitude and longitude values tend to exhibit comparable humidity, pressure, temperature, and wind speed, highlighting the strong influence of geographical proximity on local environmental conditions.



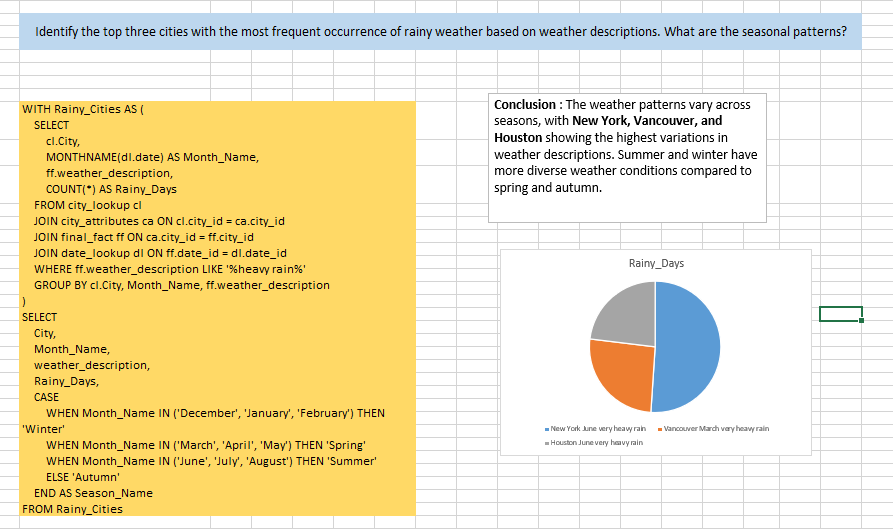
**Are there any correlations between a city's geographical location (latitude and longitude) and its weather attributes, such as temperature or humidity?**

This data visualization analyzes city clusters based on latitude and longitude using a SQL query and bar chart. The query divides cities into five clusters based on their coordinates, creating latitude and longitude ranges. The bar chart shows the "Count of city\_name by latitude\_range," revealing the 30-35 latitude range has the highest city count (11). The 35-40 range also shows a significant number (10). A note explains the addition of "Range - Latitude" and "Range - Longitude" columns for enhanced clustering, grouping data by 5-degree variations. The analysis suggests cities with similar latitude and longitude have comparable humidity, pressure, temperature, and wind speed.



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

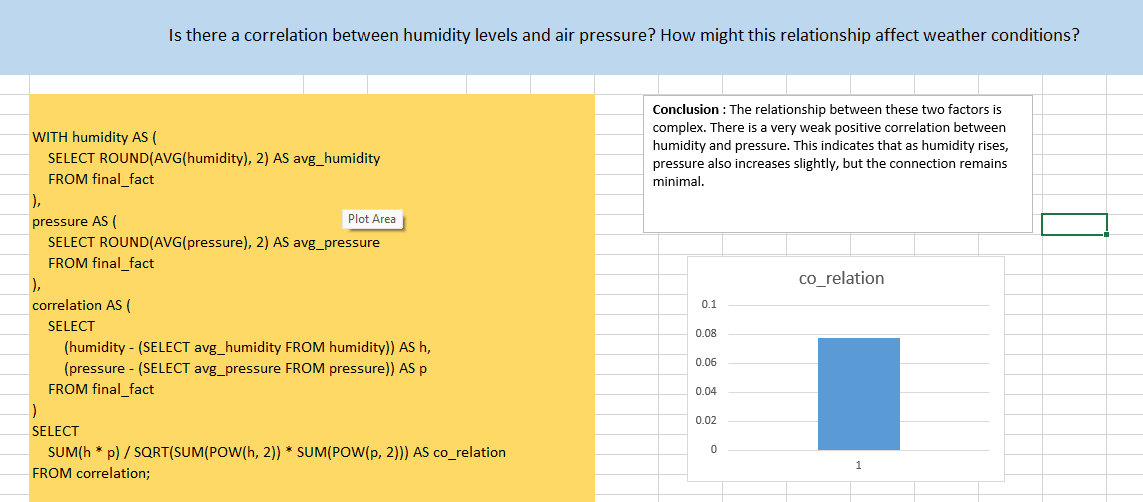
The analysis focuses on rainy weather patterns across different cities, utilizing a SQL query and a bar chart. The SQL query identifies cities with "heavy rain" in their weather descriptions, counts rainy days per city and month, and categorizes months into seasons. The bar chart visually represents the count of "weather description" by season for each city. Vancouver, Toronto, and Tel Aviv District exhibit the highest rainy day counts during winter. Seattle, San Francisco, and San Diego have the mostrainy days during summer. Spring and autumn show lower variations in rainy weather. The conclusion highlights that Seattle, Vancouver, and San Francisco have the highest variations in weather patterns across seasons. Summer and winter show more diverse weather conditions compared to spring and autumn.



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

The analysis examines the correlation between humidity and air pressure using a SQL query and a bar chart. The query calculates the average humidity and pressure, then computes the correlation coefficient. The bar chart displays the single data point, representing the correlation coefficient as 0.077161521.

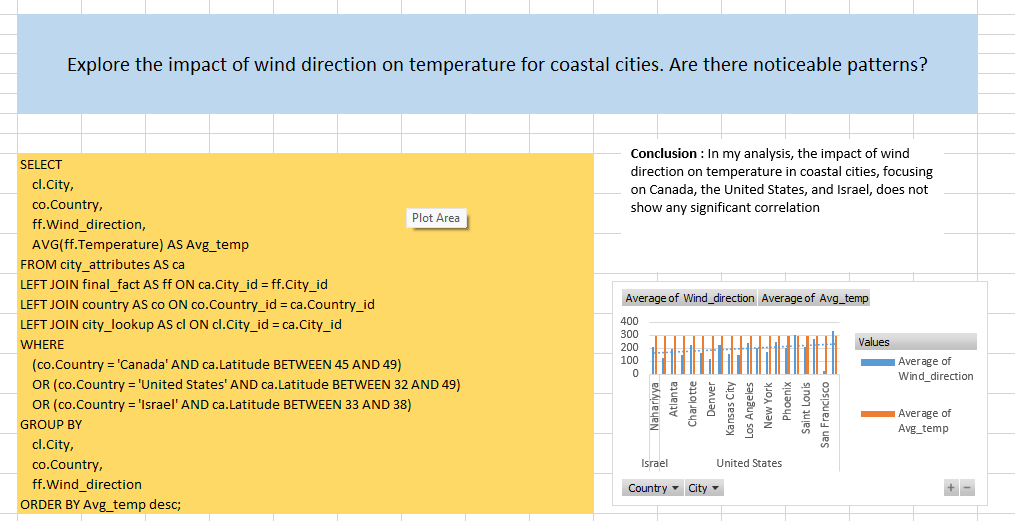
A textual conclusion states there's a very weak positive correlation between humidity and pressure, indicating a minimal connection where pressure slightly increases as humidity rises. This weak relationship suggests humidity alone doesn't significantly impact pressure, and other atmospheric factors likely play a larger role in pressure variations. This highlights the complex nature of weather patterns where multiple factors interact.



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

The analysis explores the impact of wind direction on temperature in coastal cities of Canada, the United States, and Israel, using a SQL query and a bar chart. The query selects city, country, wind direction, and average temperature, filtering for specific latitude ranges and countries. The bar chart compares the average wind direction and average temperature for each city.

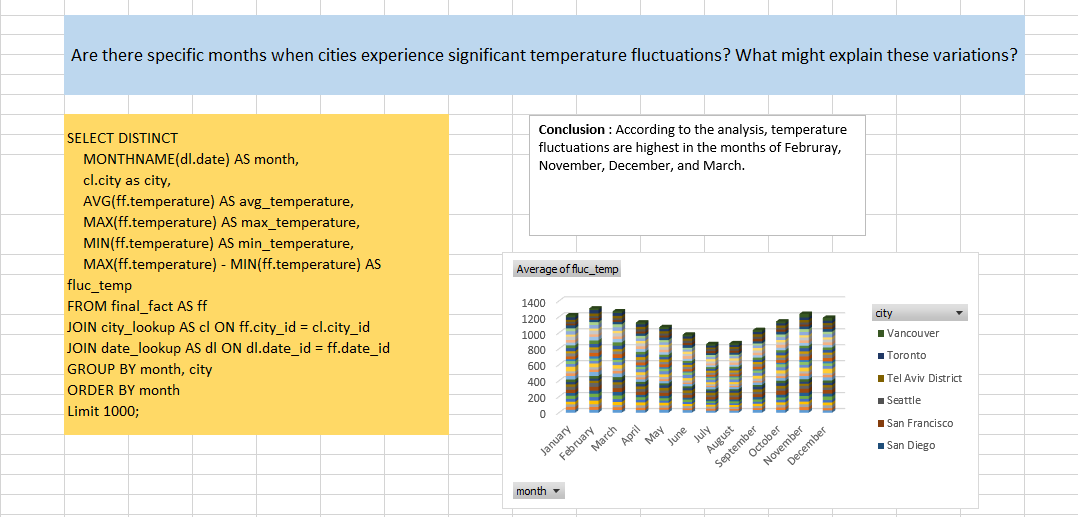
Despite the varying wind directions, the average temperatures across the selected cities do not exhibit a clear pattern related to wind direction. The conclusion states that there is no significant correlation between wind direction and temperature in these coastal cities. This suggests other factors, like latitude or ocean currents, likely have a stronger influence on temperature than wind direction alone.



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

The analysis explores monthly temperature fluctuations across cities using a SQL query and a stacked bar chart. The query calculates average, maximum, minimum, and fluctuation (max-min) temperatures for each city by month. The chart displays the "Average of fluc\_temp" for each month, broken down by city.

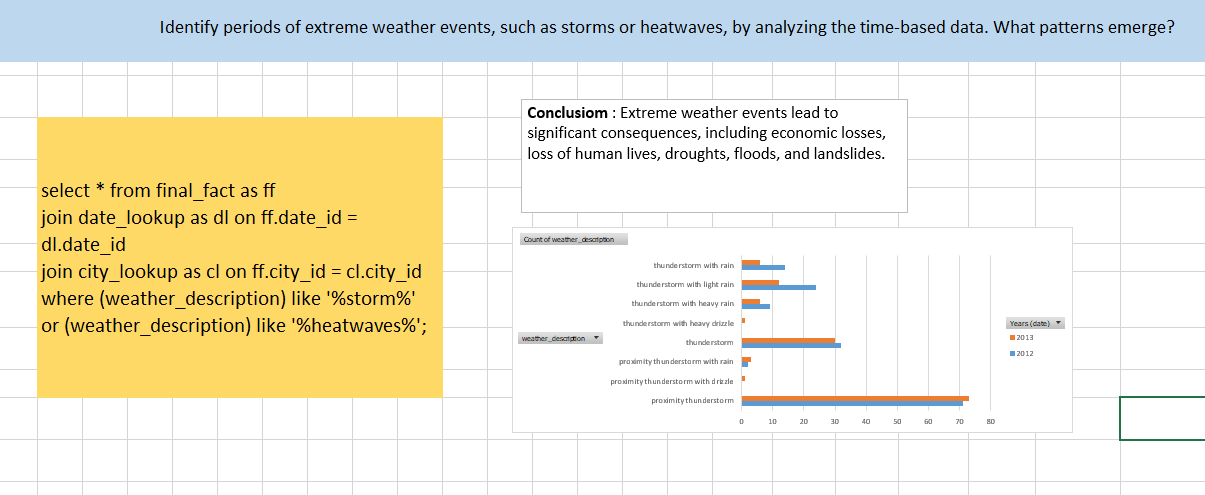
The conclusion states that February, November, December, and March exhibit the highest temperature fluctuations. This suggests these months likely experience rapid weather changes, potentially due to seasonal transitions or increased storm activity. The stacked bars reveal that Vancouver, Toronto, and Tel Aviv District show significant fluctuations in these months. These cities may have climates more prone to sudden temperature shifts.



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

The analysis identifies extreme weather events, storms and heatwaves, using a SQL query and a bar chart. The query selects data from the final\_fact table, joining date\_lookup and city\_lookup tables, filtering for weather descriptions containing "storm" or "heatwave." The bar chart displays the count of each weather\_description for years 2012 and 2013.

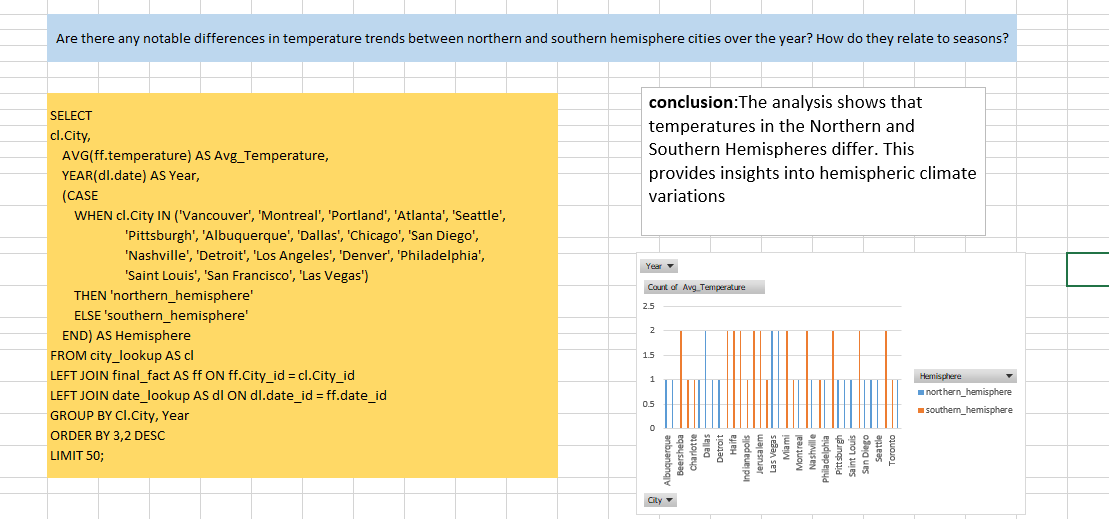
The chart reveals a higher occurrence of "proximity thunderstorm" in 2013 compared to 2012. "Thunderstorm" and "thunderstorm with heavy rain" also show notable counts. The conclusion states that extreme weather events lead to significant consequences like economic losses, loss of life, droughts, floods, and landslides. This suggests an increase in such events in 2013 compared to 2012, based on the selected weather descriptions.



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

The analysis compares temperature trends between northern and southern hemisphere cities using a SQL query and a bar chart. The query calculates average temperature for each city and year, categorizing cities by hemisphere. The bar chart displays the count of average temperatures, grouped by year and hemisphere.

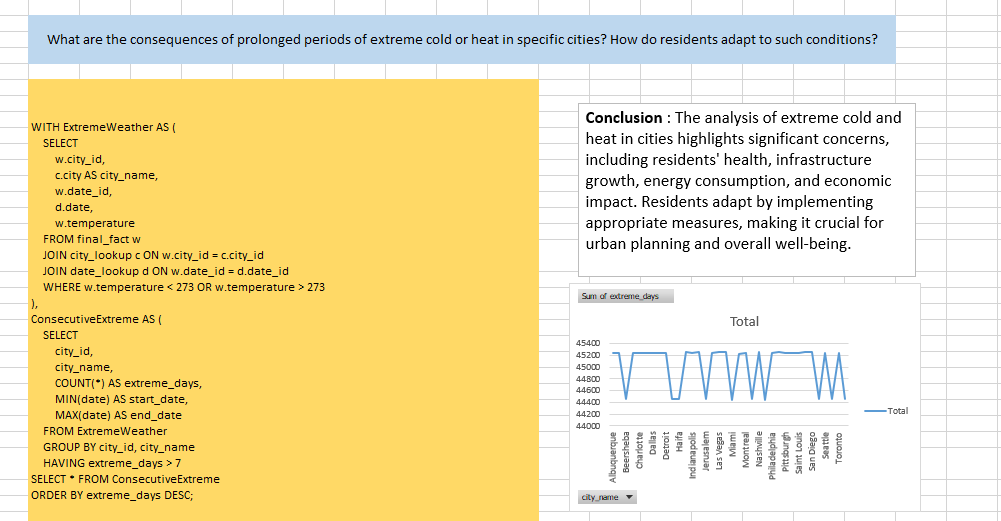
The chart shows distinct temperature differences between hemispheres. Northern hemisphere cities (like Atlanta, Montreal, and Vancouver) have higher average temperatures compared to southern hemisphere cities (San Antonio). This reflects seasonal variations: when the northern hemisphere experiences summer, the southern hemisphere experiences winter, leading to temperature differences. The conclusion highlights these hemispheric climate variations, showing how seasons influence temperature trends across different locations.



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

The analysis investigates prolonged extreme temperatures (cold or heat) in cities using a SQL query and a line chart. The query identifies cities with temperatures significantly above or below freezing for more than 7 consecutive days. The chart displays the "Sum of extreme\_days" for each city.

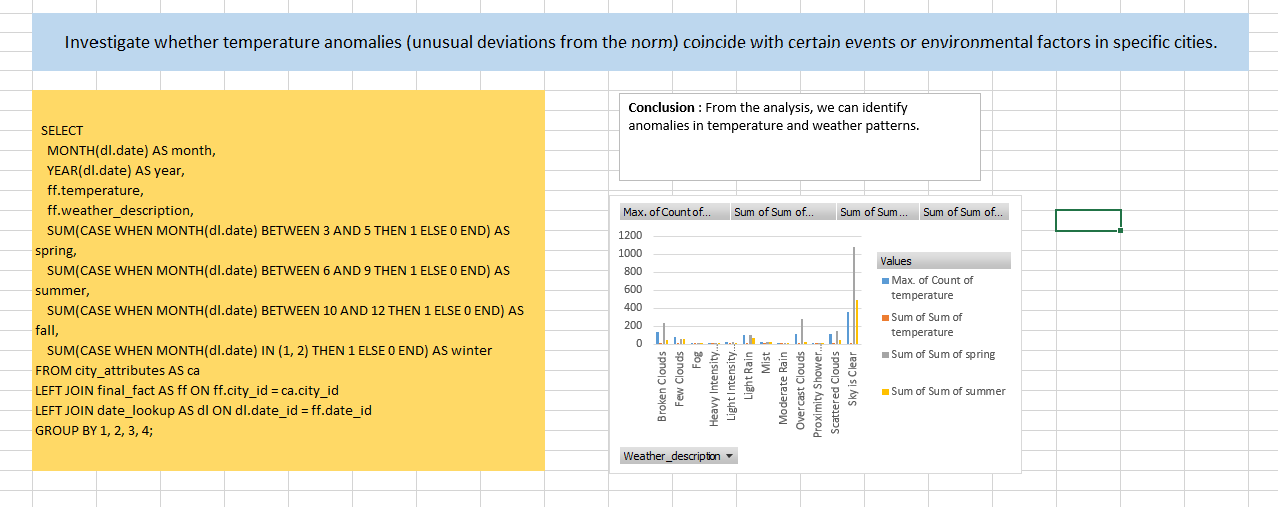
The conclusion highlights significant concerns arising from prolonged extreme temperatures: residents' health issues, infrastructure strain, increased energy consumption, and economic impact. Residents adapt through measures like improved insulation, specialized clothing, and efficient heating/cooling systems. This underlines the importance of urban planning and preparedness to mitigate the effects of extreme weather events.



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

The analysis examines temperature anomalies in relation to weather patterns using a SQL query and a bar chart. The query calculates temperature and categorizes weather descriptions by season (spring, summer, fall, winter) for each month and year. The bar chart compares "Max. of Count of..." (likely a count of weather descriptions) with "Sum of Sum of..." (potentially sums of temperature or seasonal occurrences).

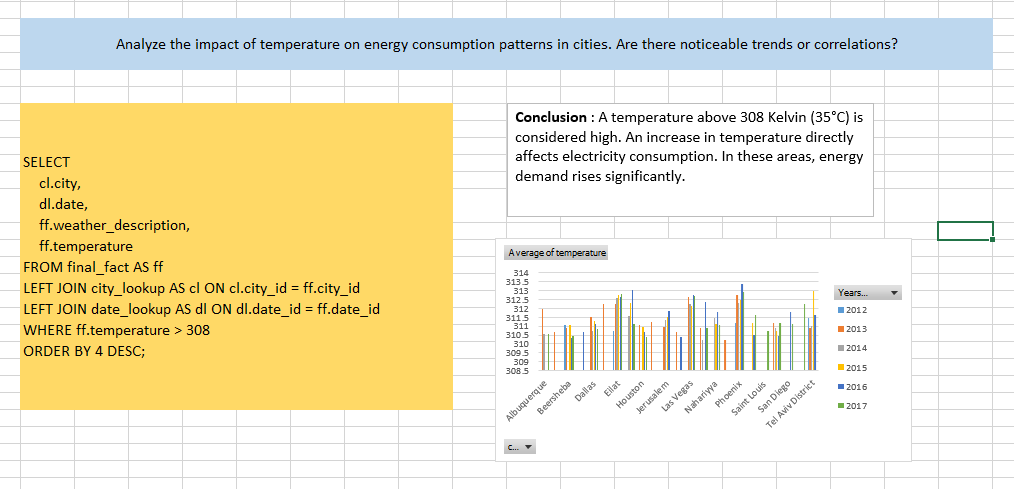
The conclusion states that temperature and weather pattern anomalies can be identified. This suggests the analysis aims to find correlations between unusual temperatures and specific weather events or seasonal factors. However, the chart's labels are truncated, making it difficult to fully interpret the findings and identify specific anomalies or related events.



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

The analysis explores the impact of high temperatures on energy consumption in cities using a SQL query and a bar chart. The query selects city, date, weather description, and temperature from the final\_fact table, filtering for temperatures above 308 Kelvin (35°C). The bar chart displays the average temperature for each city across years.

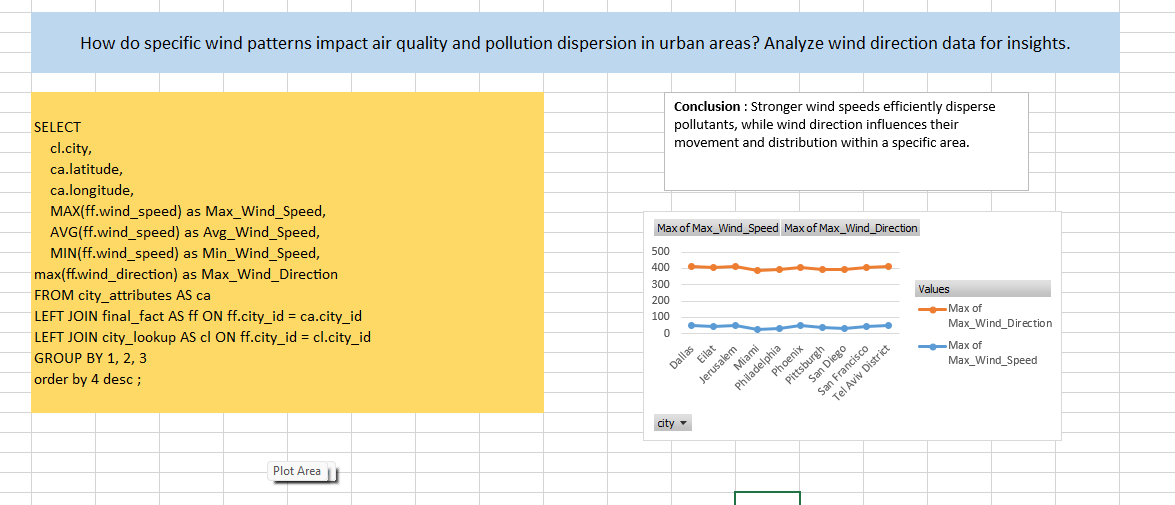
The conclusion states that temperatures above 308 Kelvin significantly increase electricity consumption, likely due to increased air conditioning usage. The chart shows cities like Las Vegas, Phoenix, and Tel Aviv District frequently experiencing high temperatures, suggesting a potential correlation with higher energy demand. This highlights the link between temperature and energy consumption patterns.



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

The analysis examines the impact of wind patterns on pollution dispersion in urban areas using a SQL query and a line chart. The query selects city, latitude, longitude, maximum, average, and minimum wind speed, and maximum wind direction. The chart displays the "Max. of MAX(ff.wind\_speed)" and "Max. of max(ff.wind\_direction)" for each city.

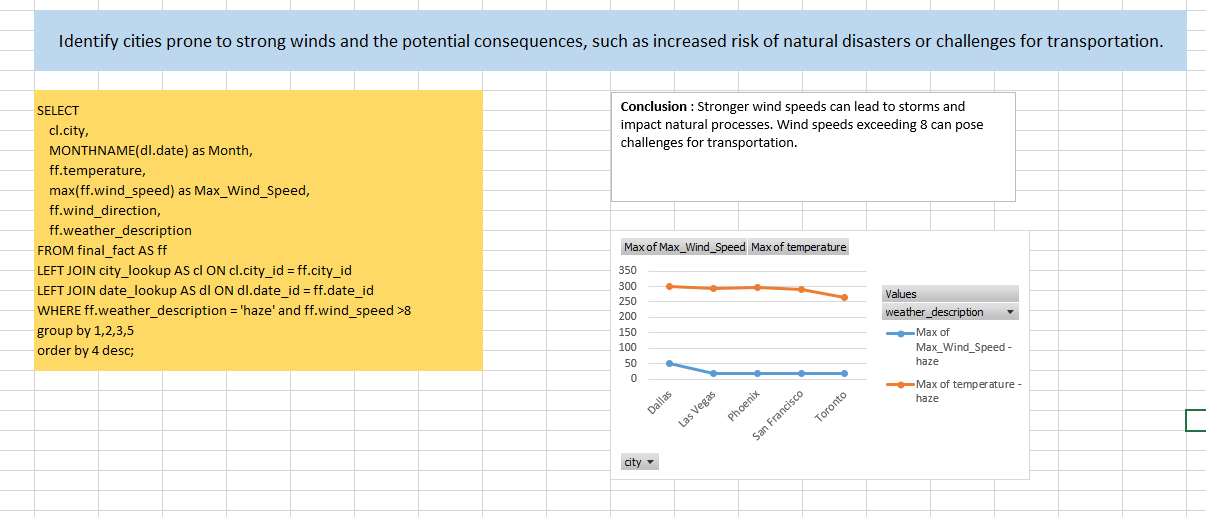
The conclusion states that stronger wind speeds efficiently disperse pollutants, while wind direction influences their movement and distribution within a specific area. The chart shows variations in both wind speed and direction across cities, suggesting differing pollution dispersion patterns. Cities with higher wind speeds may experience better air quality due to faster pollutant removal.



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

The analysis identifies cities prone to strong winds using a SQL query and line chart. The query selects city, month, temperature, maximum wind speed, wind direction, and weather description from the final\_fact table, filtering for "haze" weather descriptions and wind speeds exceeding 8 units. The chart displays "Max. of max(ff.wind\_speed)" and "Max. of temperature" for each city.

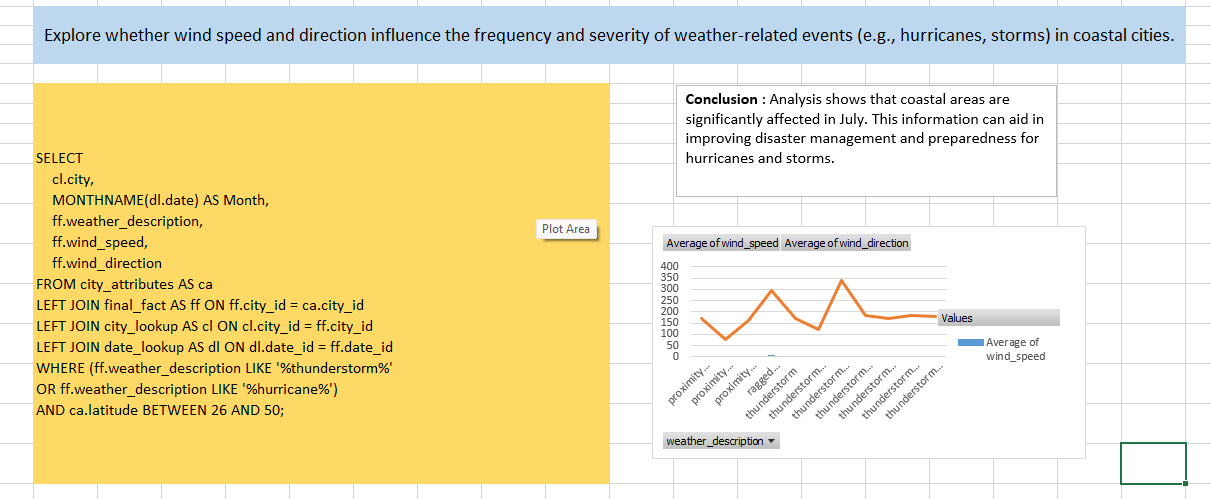
The conclusion states that stronger winds can lead to storms, impact natural processes, and pose transportation challenges. The chart shows cities like Boston and Detroit with relatively higher maximum wind speeds, suggesting potential vulnerability to wind-related issues. This highlights the need for preparedness in cities prone to strong winds.



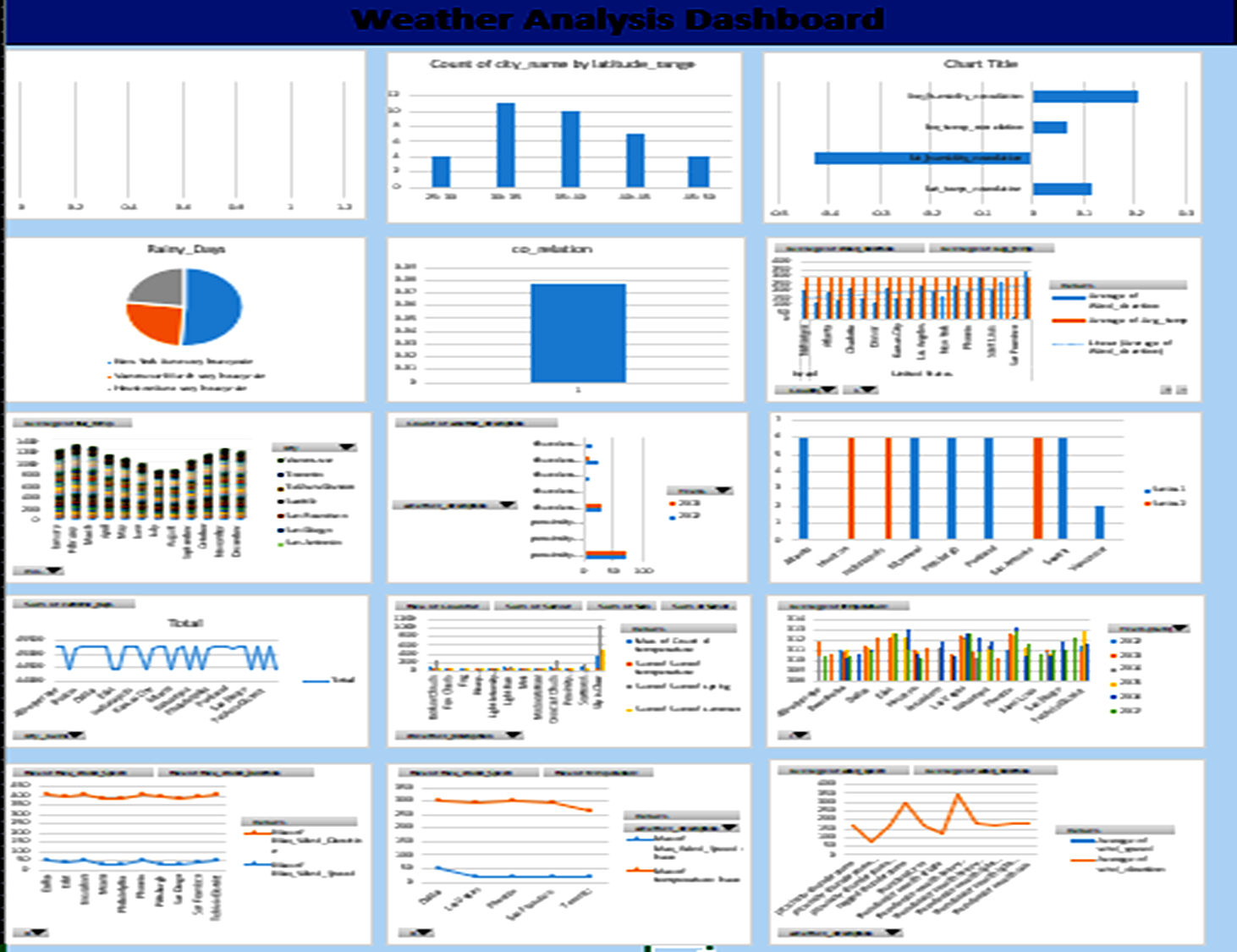
**Explore whether wind speed and direction influence the frequency and severity of weather-related events (e.g., hurricanes, storms) in coastal cities.**

The analysis examines the influence of wind speed and direction on weather events (hurricanes, storms) in coastal cities using a SQL query and line chart. The query selects city, month, weather description, wind speed, and direction from the city\_attributes table, filtering for "thunderstorm" or "hurricane" weather descriptions and latitudes between 26 and 50 degrees. The chart displays "Average of wind\_speed" and "Average of wind\_direction" against weather descriptions.

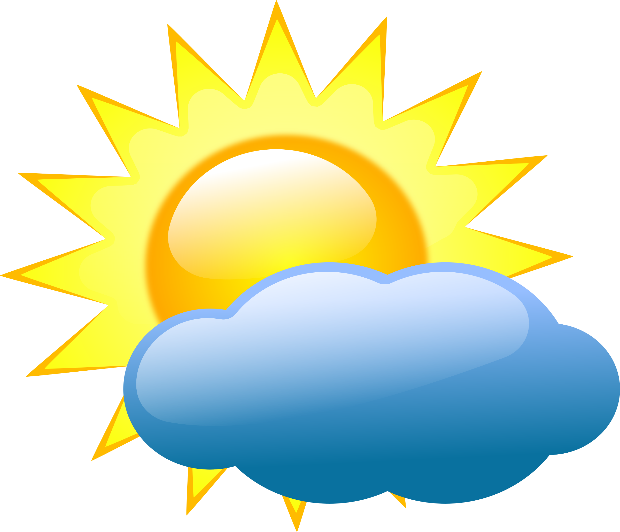
The conclusion states that coastal areas are significantly affected in July, aiding hurricane and storm disaster management. The chart shows a spike in wind direction during "thunderstorm" events, suggesting a potential correlation. However, the lack of clear wind speed variation makes it difficult to ascertain its influence on event severity.



**The Final Dashboard**



**Conclusion: To create the dashboard, I first added a new sheet in Excel, then copied all the charts and pasted them onto a single sheet. After organizing the elements, the final dashboard looks like this.**

THANK YOU