WHEATHER ANALYSIS

Capstone Project

Gaurav Wankhede

Overview

This project involves analyzing weather data to uncover trends, seasonal variations, and correlations between weather attributes. Utilizing Excel and SQL, the dataset encompasses information on city attributes, humidity, pressure, temperature, weather descriptions, wind direction, and wind speed. By examining hourly records, the analysis aims to understand patterns and relationships within the data. Additionally, a Power BI dashboard will be developed to provide comprehensive visualizations for weather monitoring and historical analysis. This integrated approach will offer insights into weather phenomena and their impacts, facilitating informed decision-making and resource allocation in various sectors affected by weather conditions.

The Process

1 – Data acquisition from GitHub:

I acquire the necessary dataset from GitHub repository, specifically for weather analysis. The dataset encompasses a wide range of meteorological data, including temperature, humidity, wind patterns, and spatial information. These data sources are crucial for gaining insights into weather conditions and patterns, enabling comprehensive analysis and decision-making in the field of meteorology.

2 – Data Transformation:

In the context of My weather analysis project, data transformation played a pivotal role in extracting meaningful insights from a vast and complex dataset. The process involved cleaning, structuring, and aggregating raw meteorological data from multiple sources, making it amenable to rigorous analysis. I standardized data formats, handled missing values, and organized information by date, location, and other relevant attributes. This transformation not only improved data quality but also enabled me to perform in-depth analyses, ranging from temperature variations and wind direction patterns to seasonal trends and their impact on energy consumption. Data transformation served as the foundation for my project, allowing me to unravel valuable patterns and correlations within the world of weather data analysis

3 – Connecting with Tools :

Establish connection between Dataset and various Analytical Tools. Interface the dataset with Excel, Mysql workbench (SQL), POWER BI facilitating seamless data integration and processing

4 – MECE in Excel :

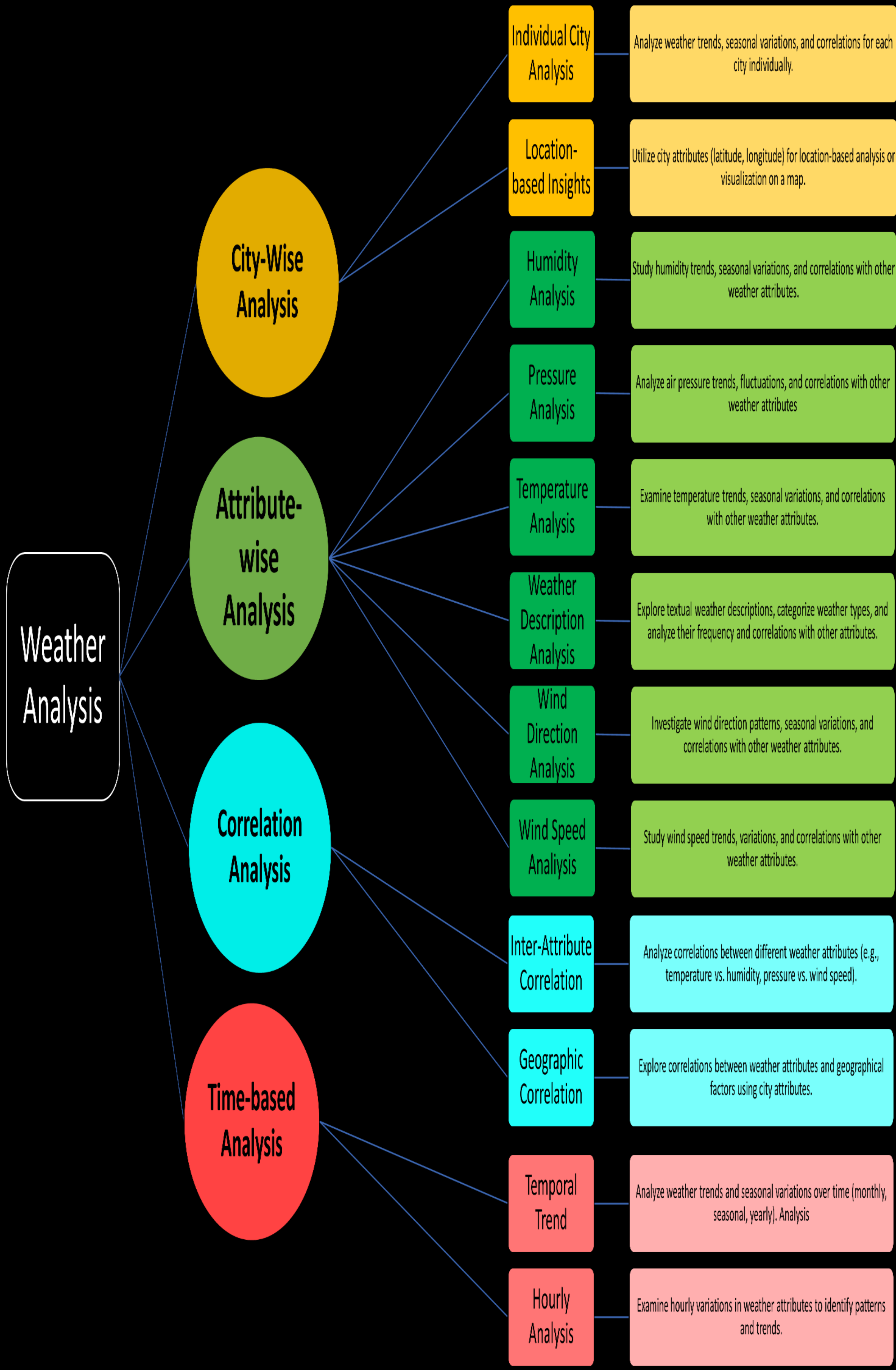
Created MECE Breakdown after seeing the data and created break into different types of analysis

5 - Problem Statement Solution in Power BI:

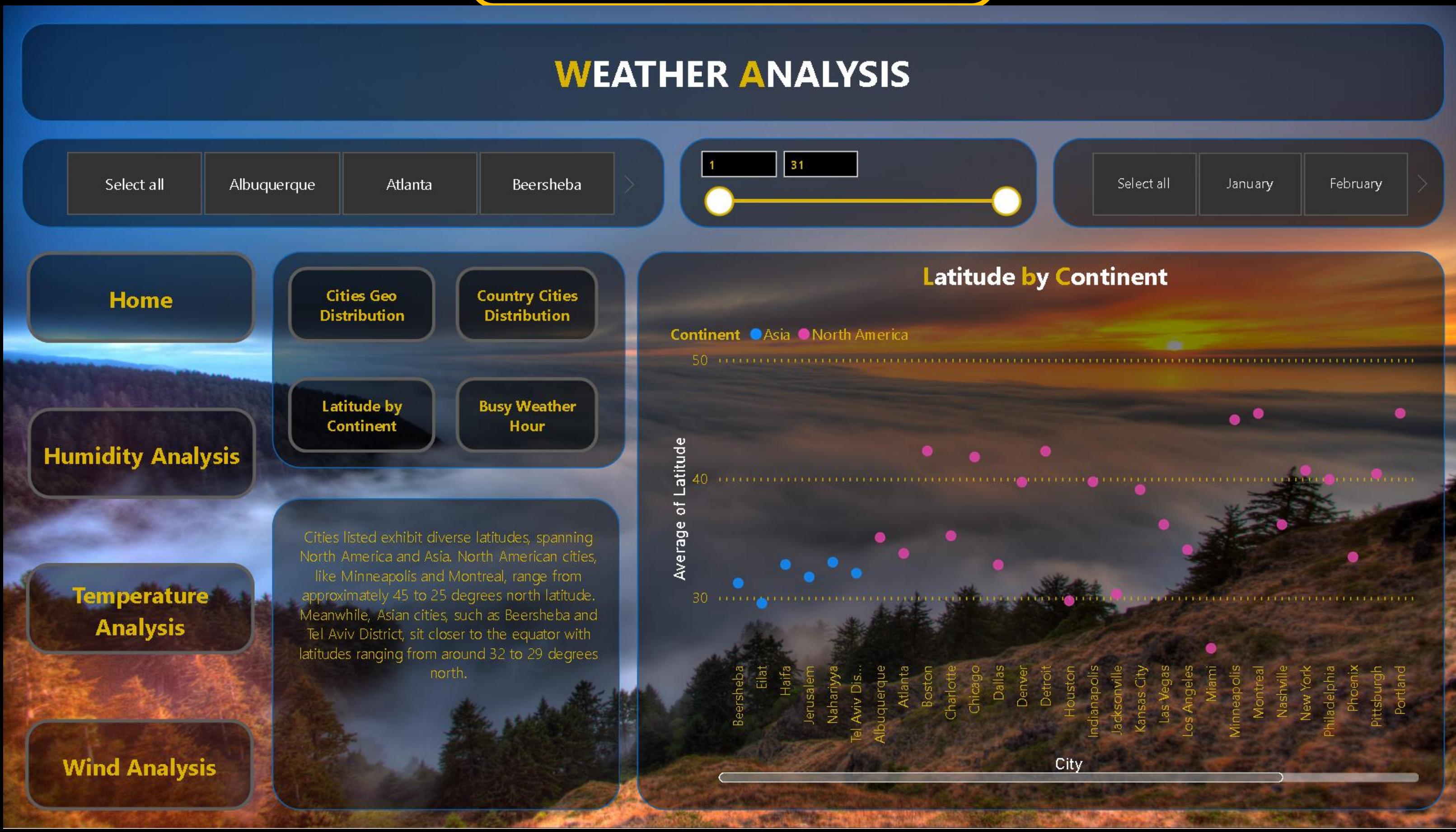
Utilize Power BI to delve into the specified problem statements. Employ its robust features for data visualization, exploration, and analysis, effectively deriving insights and solutions.

6 – EDA Statement Solution in SQL:

Utilize SQL to Solve the specified EDA problem statements. Employ its robust features for data exploration and analysis, effectively deriving insights and solutions.

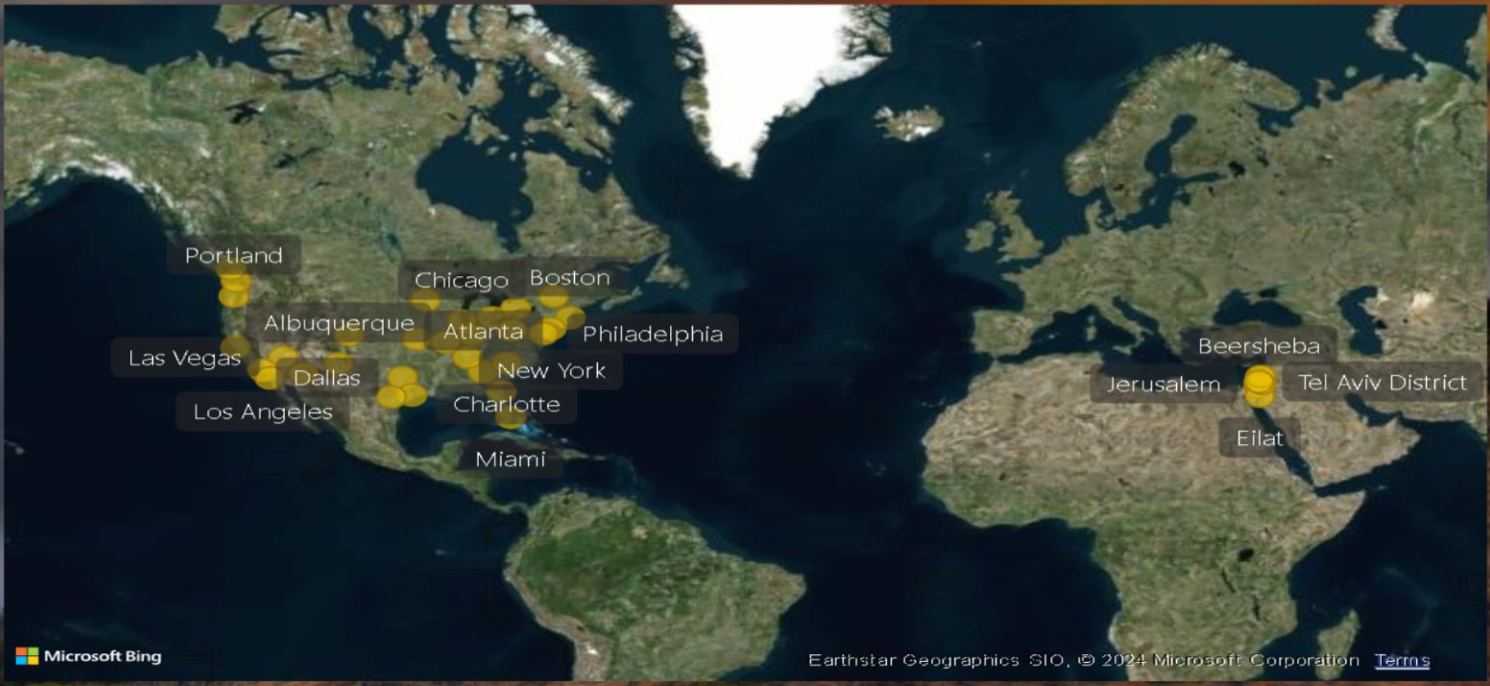

Weather Analysis
City-Wise Analysis
Attribute-wise Analysis
Correlation Analysis
Time-based Analysis
Individual City Analysis
Analyze weather trends, seasonal variations, and correlations for each city individually.
Location-based Insights
Utilize city attributes (latitude, longitude) for location-based analysis or visualization on a map.
Humidity Analysis
Study humidity trends, seasonal variations, and correlations with other weather attributes.
Pressure Analysis
Analyze air pressure trends, fluctuations, and correlations with other weather attributes
Temperature Analysis
Examine temperature trends, seasonal variations, and correlations with other weather attributes.
Weather Description Analysis
Explore textual weather descriptions, categorize weather types, and analyze their frequency and correlations with other attributes.
Wind Direction Analysis
Investigate wind direction patterns, seasonal variations, and correlations with other weather attributes.
Wind Speed Analysis
Study wind speed trends, variations, and correlations with other weather attributes.
Inter-Attribute Correlation
Analyze correlations between different weather attributes (e.g., temperature vs. humidity, pressure vs. wind speed).
Geographic Correlation
Explore correlations between weather attributes and geographical factors using city attributes.
Temporal Trend
Analyze weather trends and seasonal variations over time (monthly, seasonal, yearly). Analysis
Hourly Analysis
Examine hourly variations in weather attributes to identify patterns and trends.

Power BI



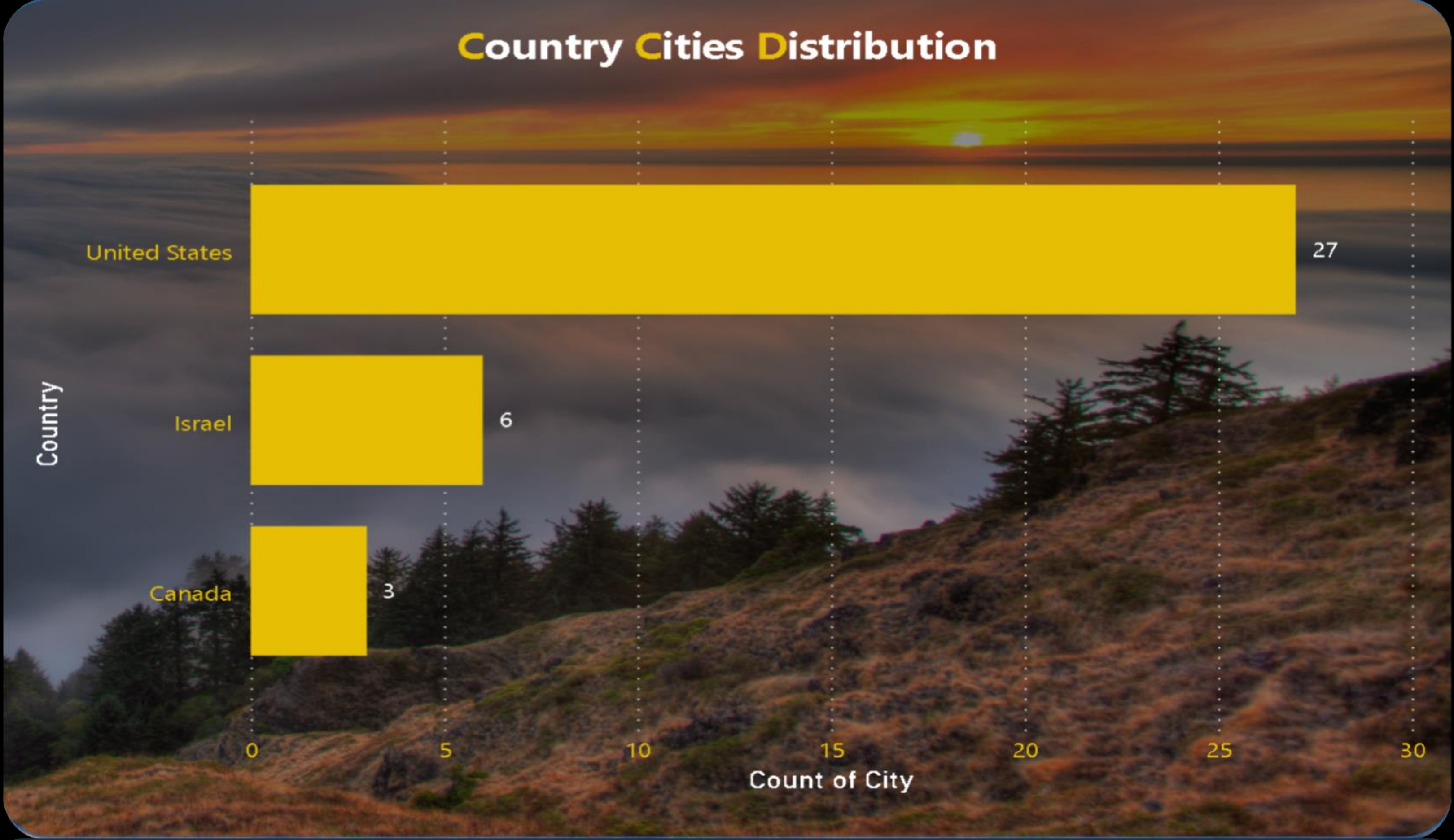
The Home page of our Power BI Weather Analysis Dashboard provides a comprehensive overview of weather data for cities. It includes essential elements such as temperature, precipitation, humidity, and wind speed. Intuitive navigation buttons facilitate seamless exploration, while filters enable users to refine their analysis based on specific criteria. This user-friendly interface empowers users to efficiently analyze and interpret weather patterns, aiding in informed decision-making.

Cities Geo Distribution



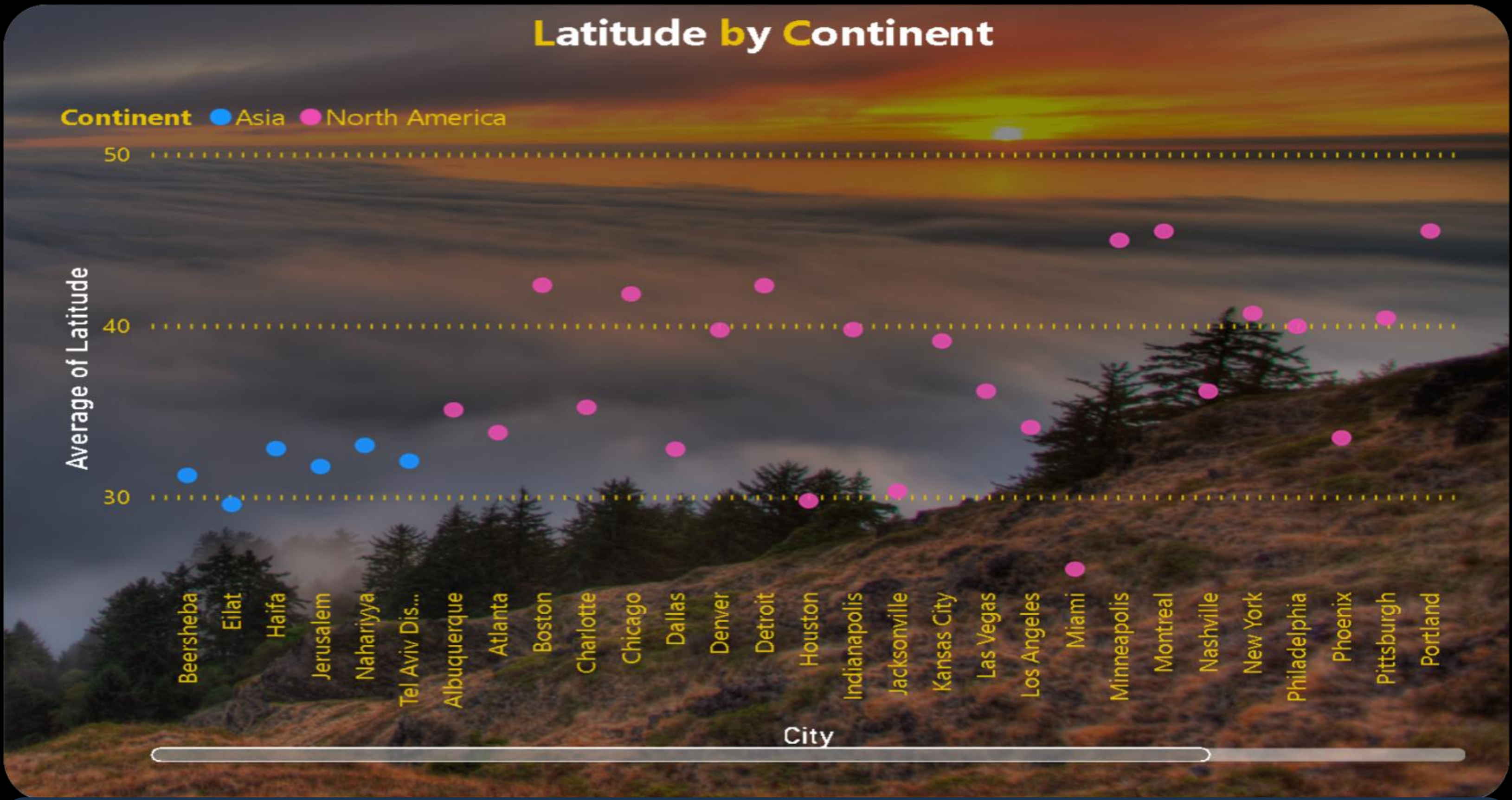
Can you create a geographical map in Power BI showing the distribution of cities in the dataset based on their latitude and longitude?

Each row represents a city, displaying its average longitude, latitude, and count. For instance, Jerusalem stands out with the highest longitude average at 35.22, while Vancouver has the lowest at -123.12. This setup facilitates efficient visualization and analysis.



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

In the analysis of city counts across countries, the United States emerged with the highest count at 27, showcasing a significant lead over Canada, which had the lowest count at 3, marking an 800.00% difference. Following the United States, Israel stood at 6 cities. Remarkably, the United States accounted for 75.00% of the total city count, with Canada and Israel contributing 8.33% and 16.67% respectively. This data underscores the prominence of the United States in terms of city distribution among the countries analyzed.



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

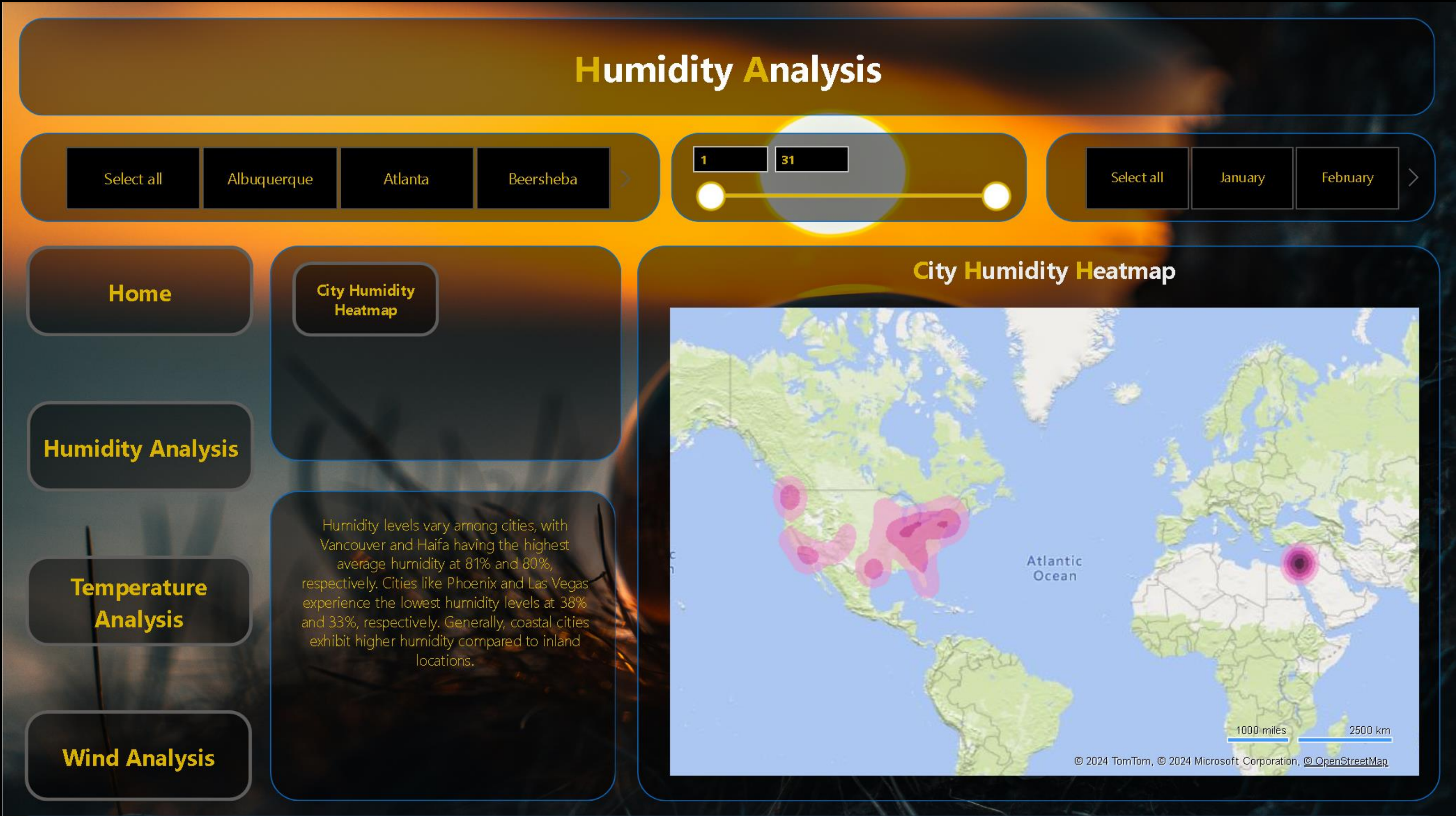
Cities listed exhibit diverse latitudes, spanning North America and Asia. North American cities, like Minneapolis and Montreal, range from approximately 45 to 25 degrees north latitude. Meanwhile, Asian cities, such as Beersheba and Tel Aviv District, sit closer to the equator with latitudes ranging from around 32 to 29 degrees north.

Busy Weather Hours

time	broken clouds	drizzle	dust	few clouds	fog	freezing rain	haze	heavy intensity drizzle	hea
00:00:00	8004	96	41	5899	210	1	969	7	
01:00:00	7705	93	39	5628	268	1	928	8	
02:00:00	7263	113	33	5815	295	1	888	7	
03:00:00	6968	91	37	5808	401	1	704	10	
04:00:00	6858	97	54	5858	539	3	659	12	
05:00:00	6446	101	49	5713	622	3	689	7	
06:00:00	5946	113	54	5469	718	2	708	12	
07:00:00	5951	125	52	4942	868	4	713	10	
08:00:00	5809	123	58	4804	1008	3	680	10	
09:00:00	5667	123	45	4925	1109	1	665	19	
10:00:00	5619	142	61	4991	1326	1	668	15	
11:00:00	5628	125	56	5089	1422	3	721	15	
12:00:00	5836	121	61	5447	1505	3	791	16	
13:00:00	6569	118	67	5691	1321	4	945	8	
14:00:00	6676	113	77	5924	1148	5	1215	10	
15:00:00	7386	111	66	6163	923	5	1471	8	
16:00:00	7753	89	60	6059	719	2	1653	3	
17:00:00	7869	84	48	5933	530	1	1712	10	
18:00:00	8417	82	44	5991	351	2	1603	7	
19:00:00	8356	82	44	5858	287	1	1438	7	

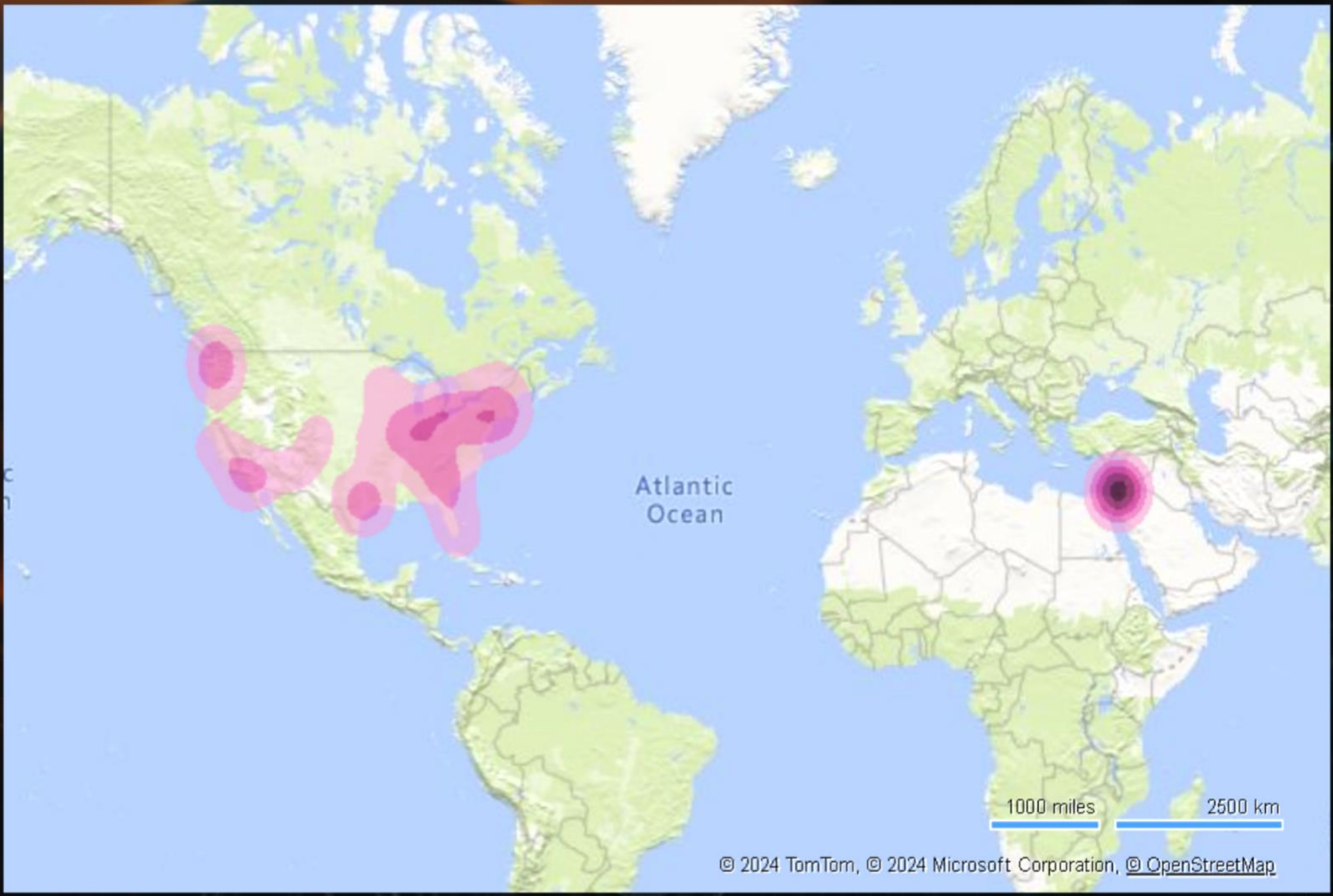
Can you create a heatmap in Power BI to visualize the busiest hours for specific weather conditions (e.g., "clear sky," "rainy")?

The data represents weather conditions recorded at different hours. The most frequent condition is 'sky is clear' with 659,181 occurrences. The least common are 'freezing rain', 'proximity sand/dust whirls', 'rain and snow', 'sand', 'squalls', and 'tornado' with only one occurrence each. The total count is 1,629,108.



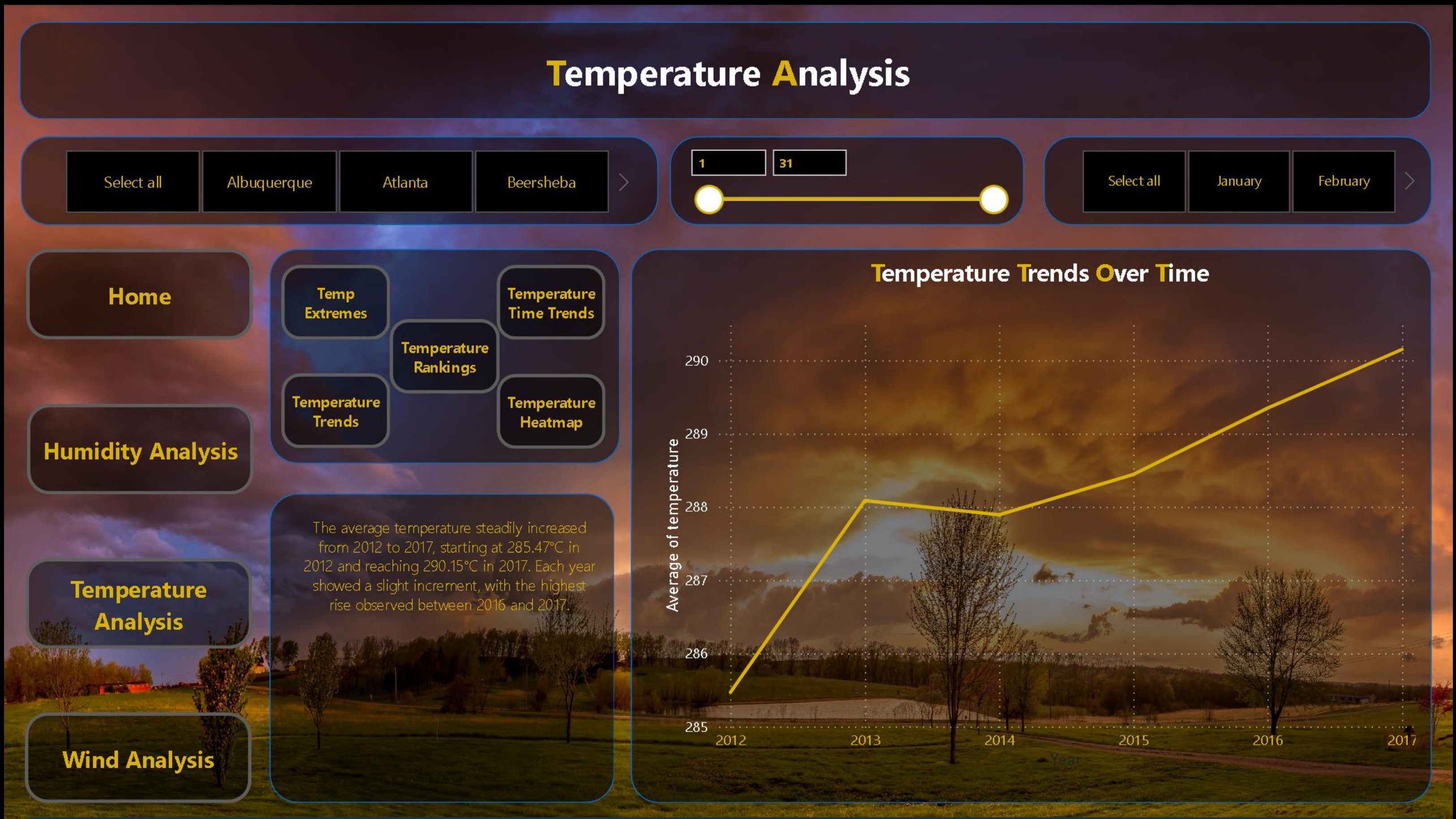
The Humidity Analysis Page on our Power BI Weather Dashboard offers detailed insights into humidity levels across various cities. Users can explore humidity trends over time and compare humidity levels between different locations effortlessly. Interactive visualizations allow for in-depth examination of humidity fluctuations, aiding in identifying patterns and anomalies. With intuitive navigation and user-friendly filters, users can tailor their analysis to focus on specific time periods or regions. This page serves as a valuable tool for understanding humidity dynamics, supporting informed decision-making in various sectors reliant on weather data.

City Humidity Heatmap

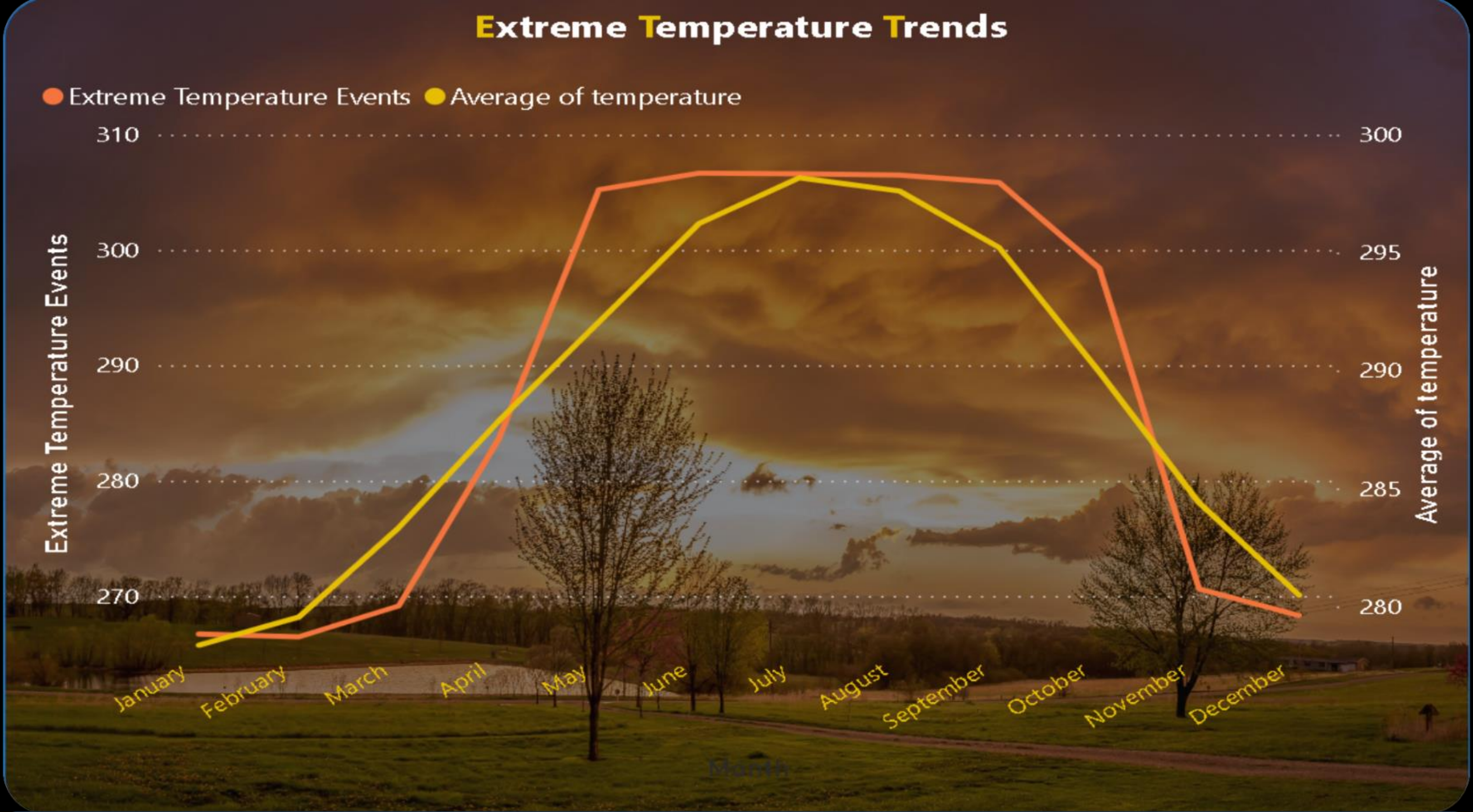


How does humidity vary across different cities? Generate a heatmap in Power BI to visualize this variation.

Humidity levels vary among cities, with Vancouver and Haifa having the highest average humidity at 81% and 80%, respectively. Cities like Phoenix and Las Vegas experience the lowest humidity levels at 38% and 33%, respectively. Generally, coastal cities exhibit higher humidity compared to inland locations.

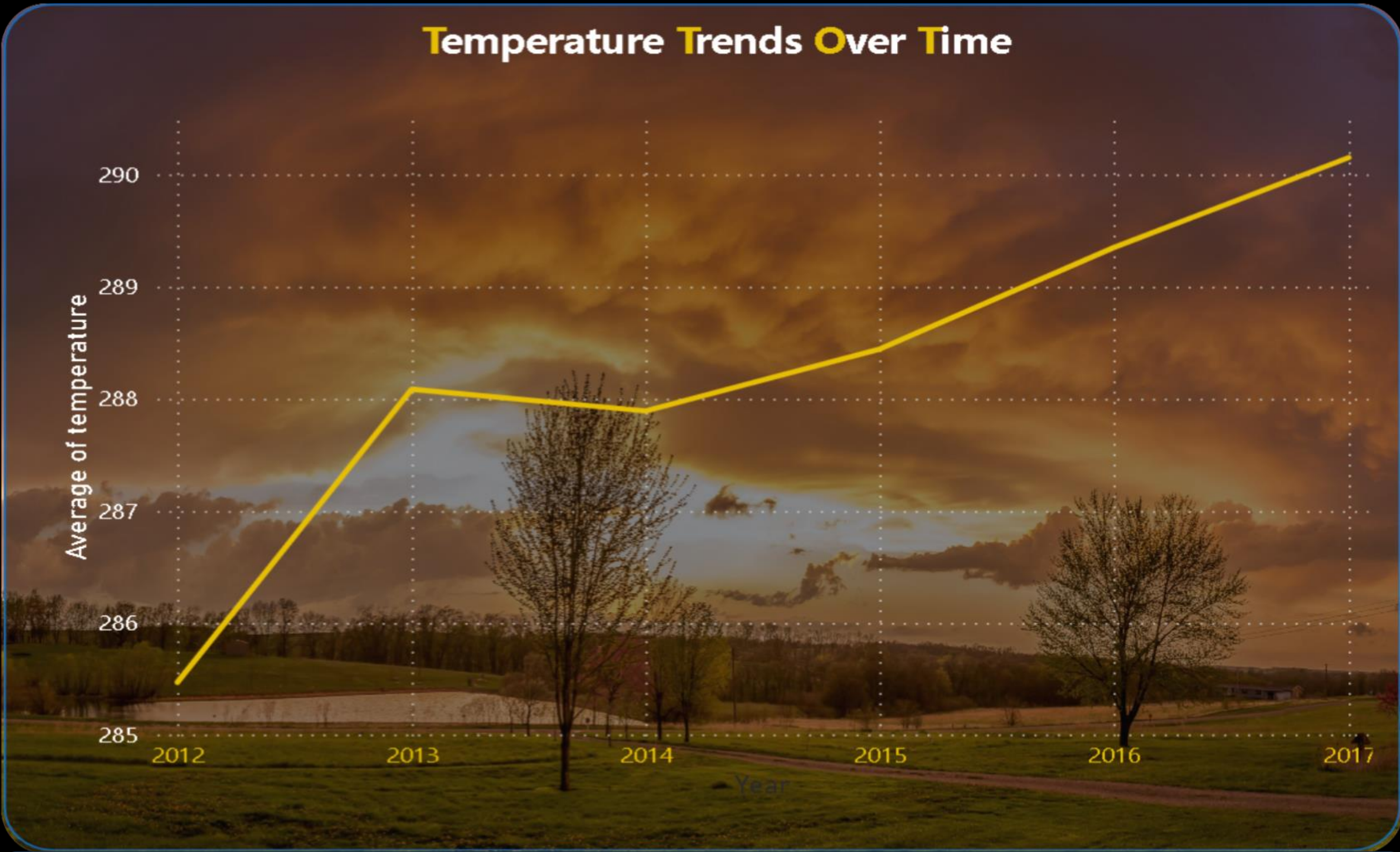


The Temperature Analysis Page within our Power BI Weather Dashboard provides a comprehensive examination of temperature patterns across diverse geographic locations. Users can delve into historical temperature data, analyze trends, and compare temperature variations between cities with ease. Interactive visualizations offer dynamic insights, enabling users to identify temperature fluctuations over time and across different regions. With intuitive navigation features and customizable filters, users can refine their analysis based on specific parameters, facilitating precise decision-making processes. This page serves as a valuable resource for understanding temperature dynamics and their implications across various industries and sectors.



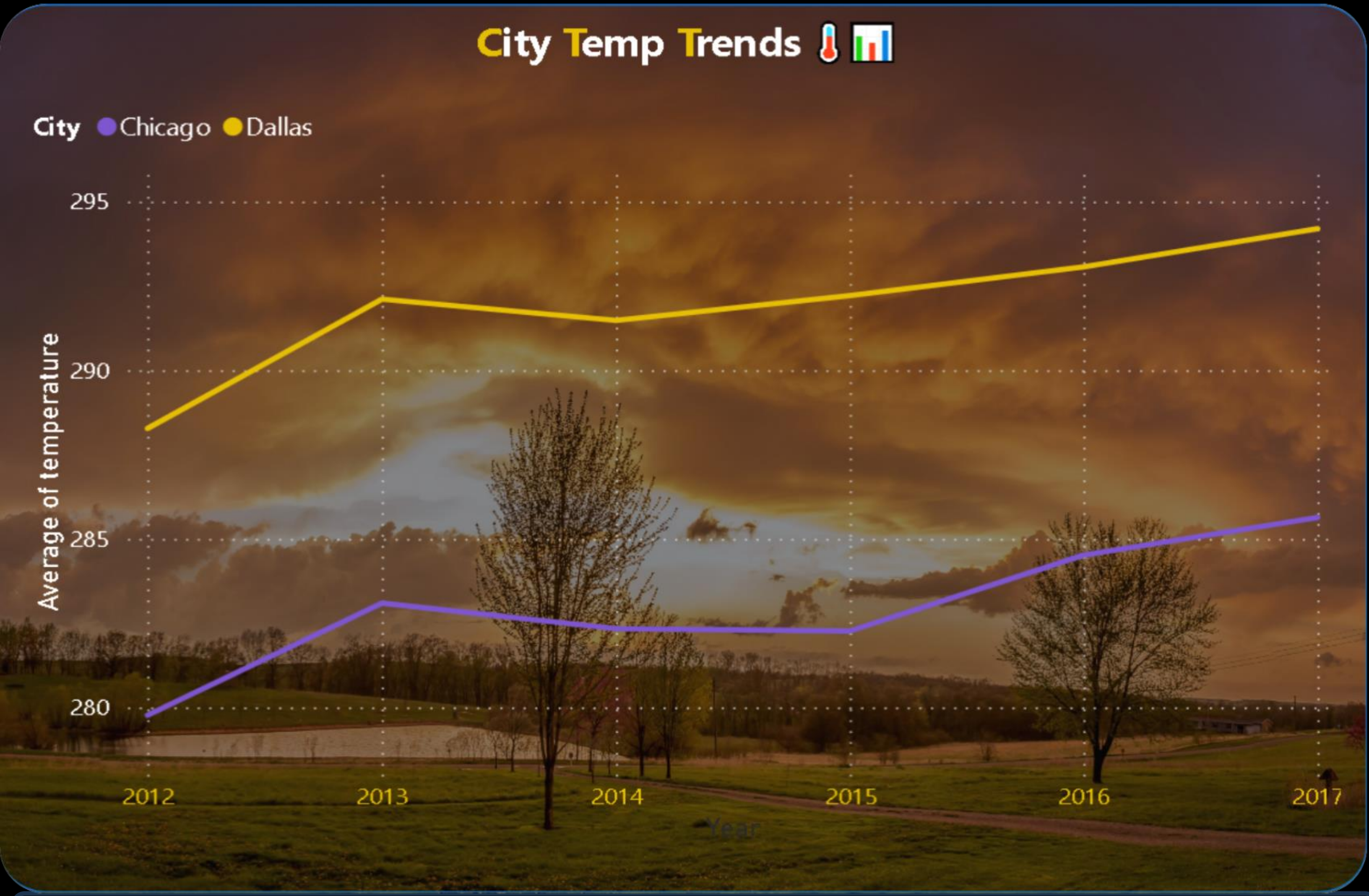
Create a line chart in Power BI to display the temperature trends over time for a selected city. Highlight extreme temperature events.

Extreme temperature events peaked in June and July, reaching 305.67°C and 305.85°C respectively. Lowest temperatures were in January and November at 268.73°C and 270.31°C. Overall, temperatures rose from January to July before declining slightly towards December.



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

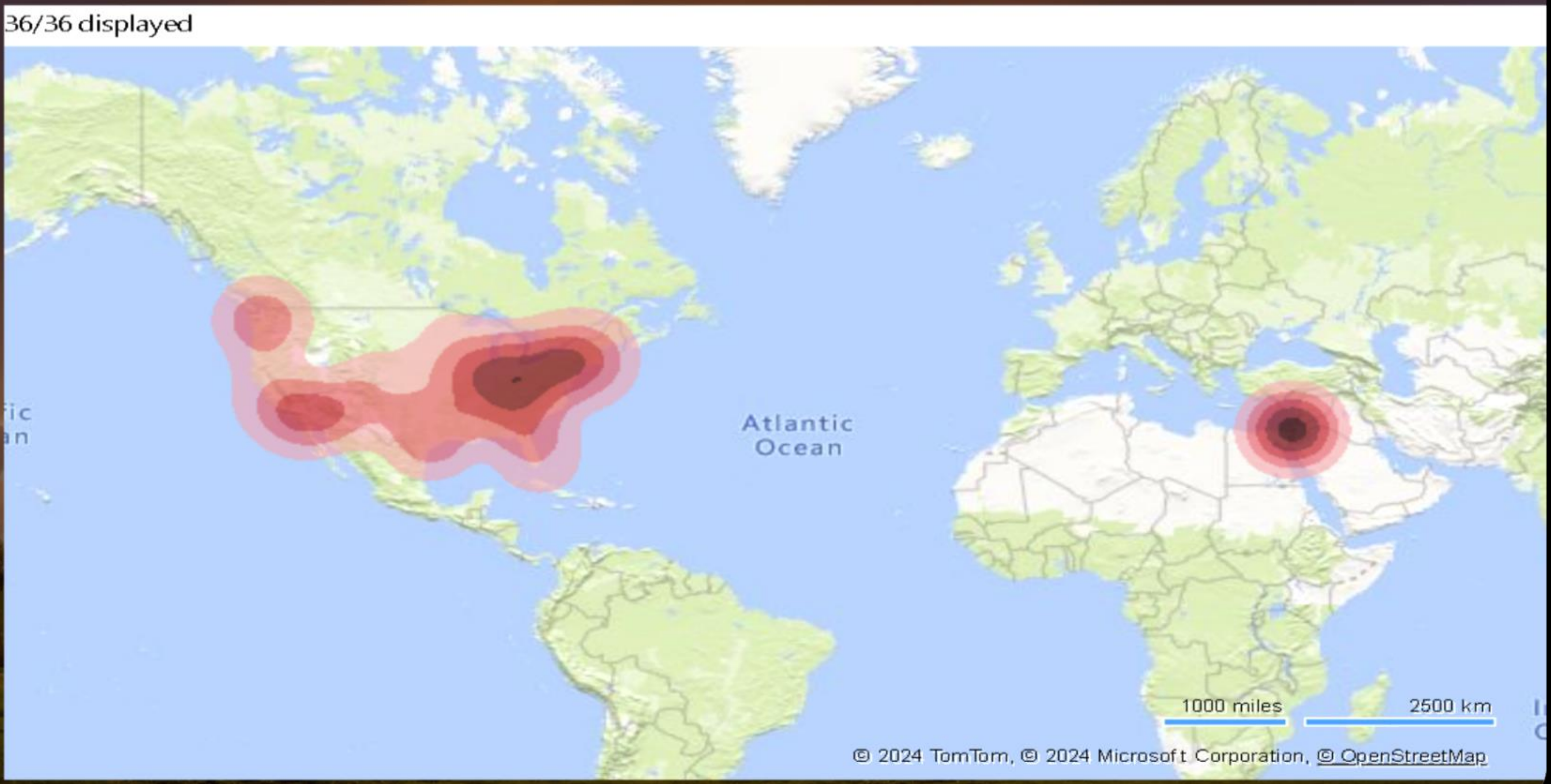
The average temperature steadily increased from 2012 to 2017, starting at 285.47°C in 2012 and reaching 290.15°C in 2017. Each year showed a slight increment, with the highest rise observed between 2016 and 2017.



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

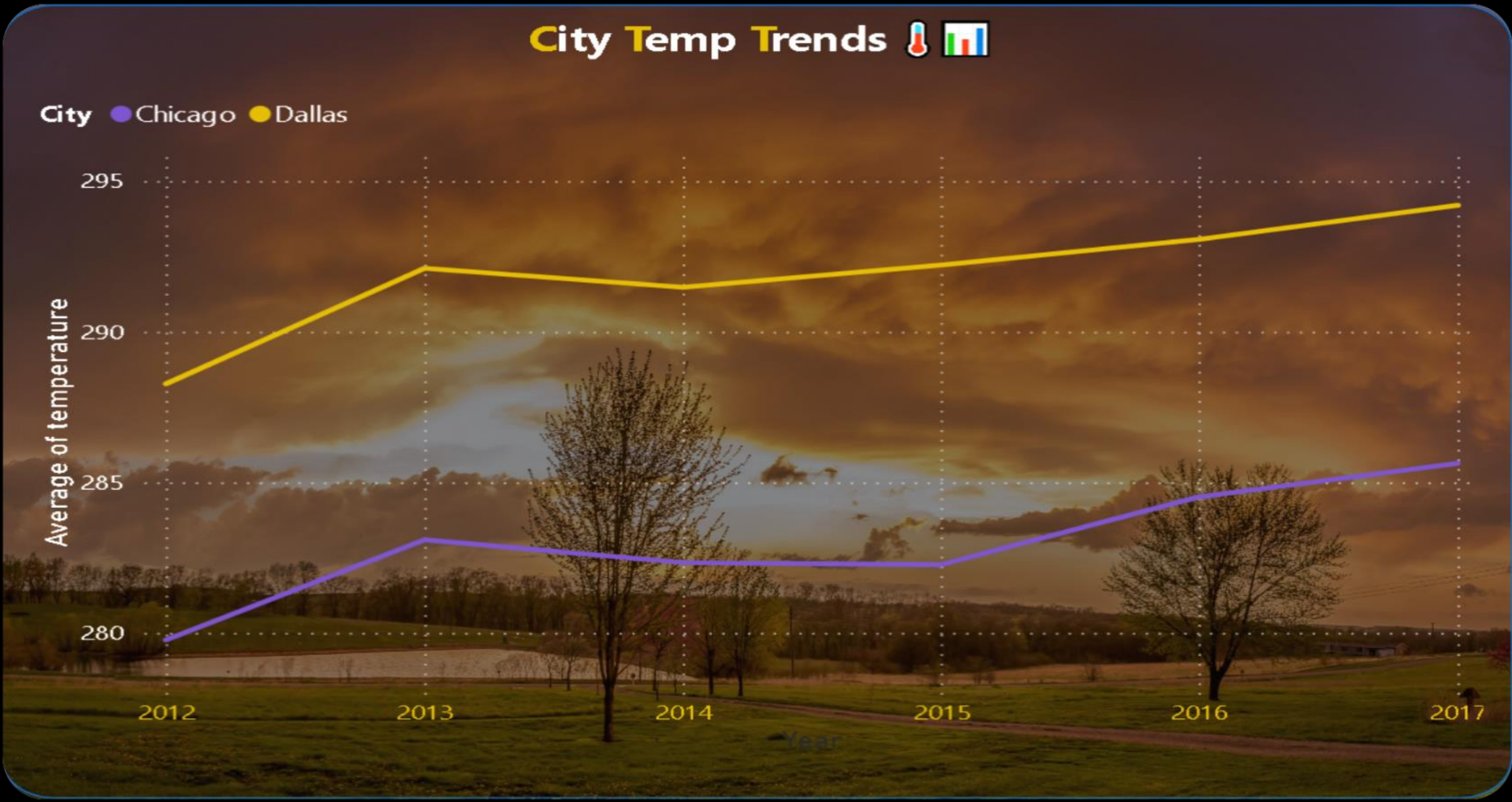
From 2012 to 2017, Atlanta experienced a gradual temperature rise, reaching 291.67°C in 2017. Boston also saw an upward trend, with temperatures climbing from 279.93°C in 2012 to 285.27°C in 2017. Atlanta consistently maintained higher temperatures compared to Boston throughout the period.

City Temperature Heatmap



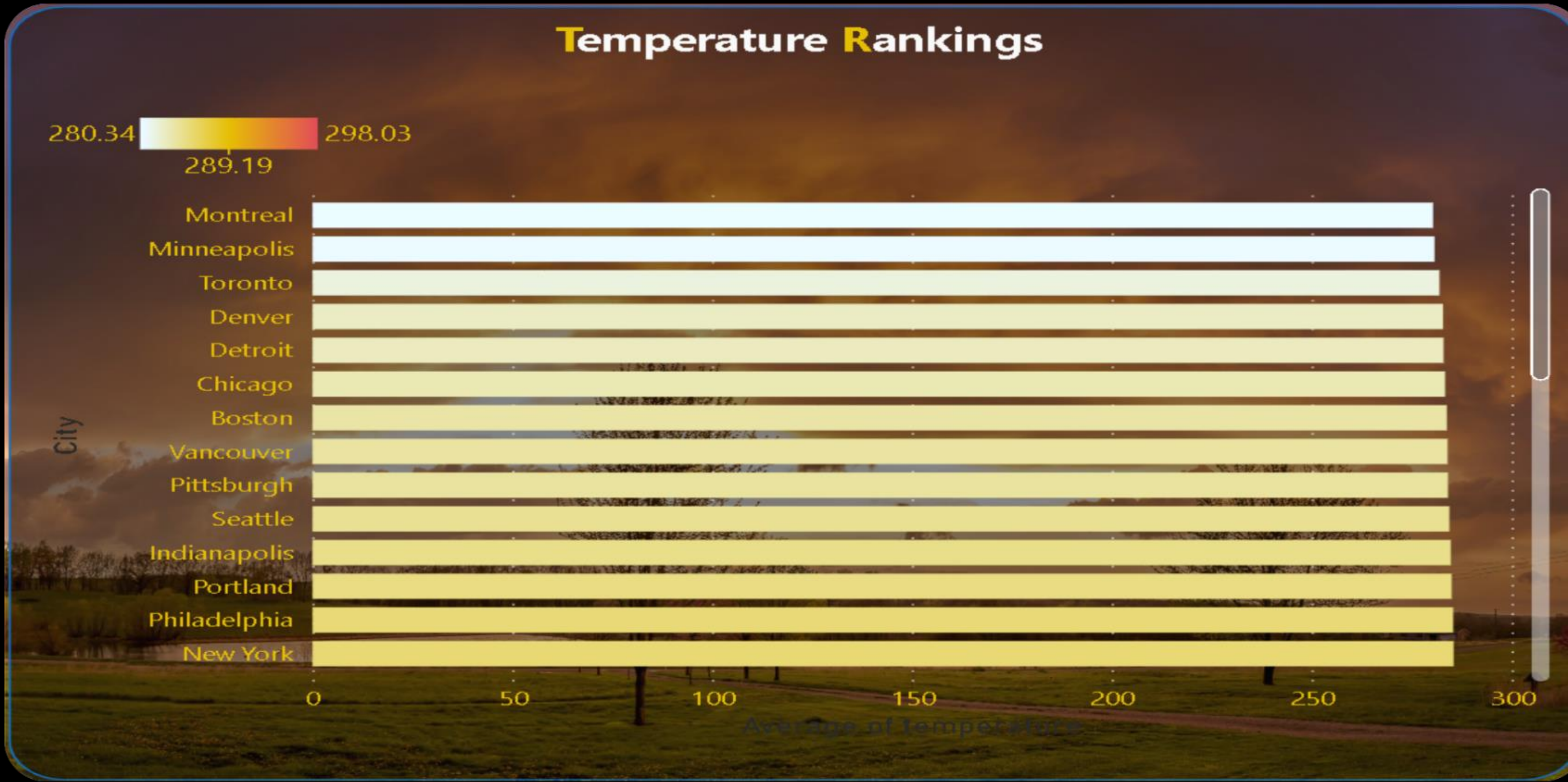
Can you build a heatmap in Power BI to show the temperature ranges for cities across different countries?

Cities exhibit diverse average temperatures: Miami, United States, records the highest at 298.03°C, while Montreal, Canada, experiences the lowest at 280.34°C. Tel Aviv District, Israel, sees a moderate 294.43°C. United States cities vary from 280.70°C in Minneapolis to 298.03°C in Miami.



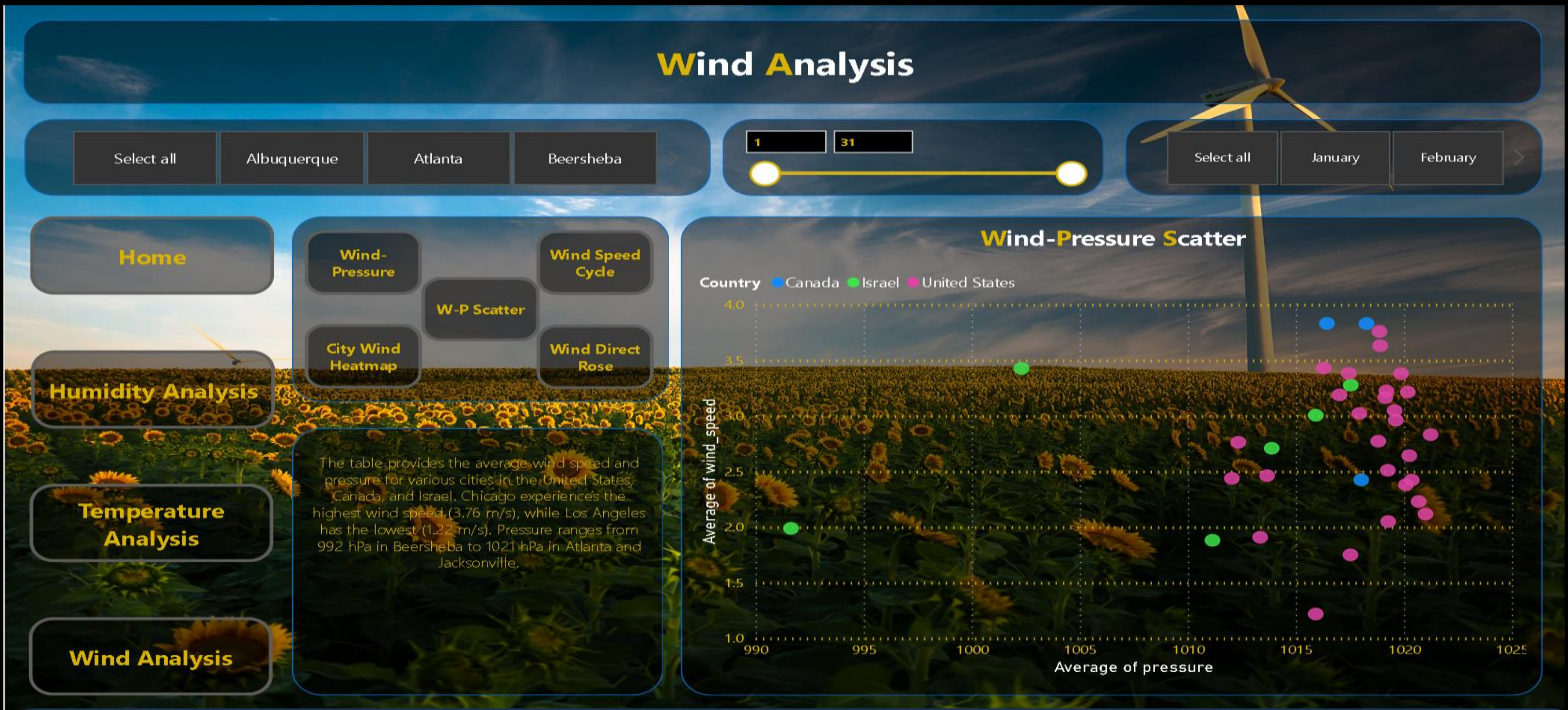
Can you build a heatmap in Power BI to show the temperature ranges for cities across different countries?

Cities exhibit diverse average temperatures: Miami, United States, records the highest at 298.03°C, while Montreal, Canada, experiences the lowest at 280.34°C. Tel Aviv District, Israel, sees a moderate 294.43°C. United States cities vary from 280.70°C in Minneapolis to 298.03°C in Miami.

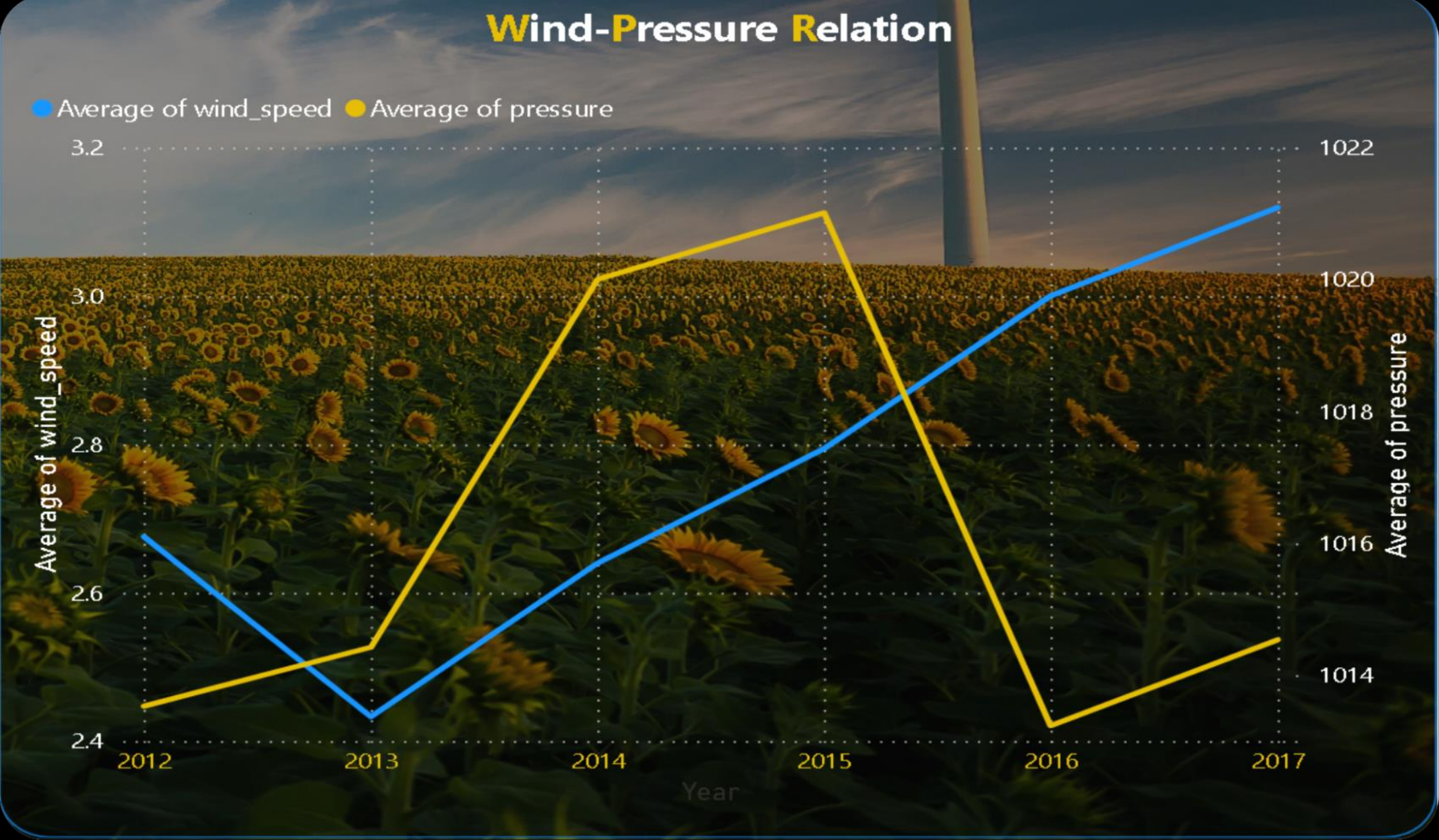


Create a bar chart in Power BI to highlight cities with the highest and lowest average temperatures in the dataset.

Cities exhibit varied average temperatures: Miami, USA, records the highest at 298.03°C, while Montreal, Canada, experiences the lowest at 280.34°C. Tel Aviv District, Israel, sees 294.43°C, while major US cities range from 280.70°C in Minneapolis to 298.03°C in Miami.



The Wind Analysis Page on our Power BI Weather Dashboard delivers a thorough exploration of wind dynamics across multiple cities. Users can examine wind speed and direction trends, facilitating informed decision-making in various industries reliant on weather data. Interactive visualizations empower users to compare wind patterns over time and across different geographical regions effortlessly. With intuitive navigation functionalities and customizable filters, users can refine their analysis based on specific criteria, ensuring precision in their assessments. This page serves as a pivotal tool for understanding wind behavior and its potential impact, aiding stakeholders in optimizing operations and resource management strategies.

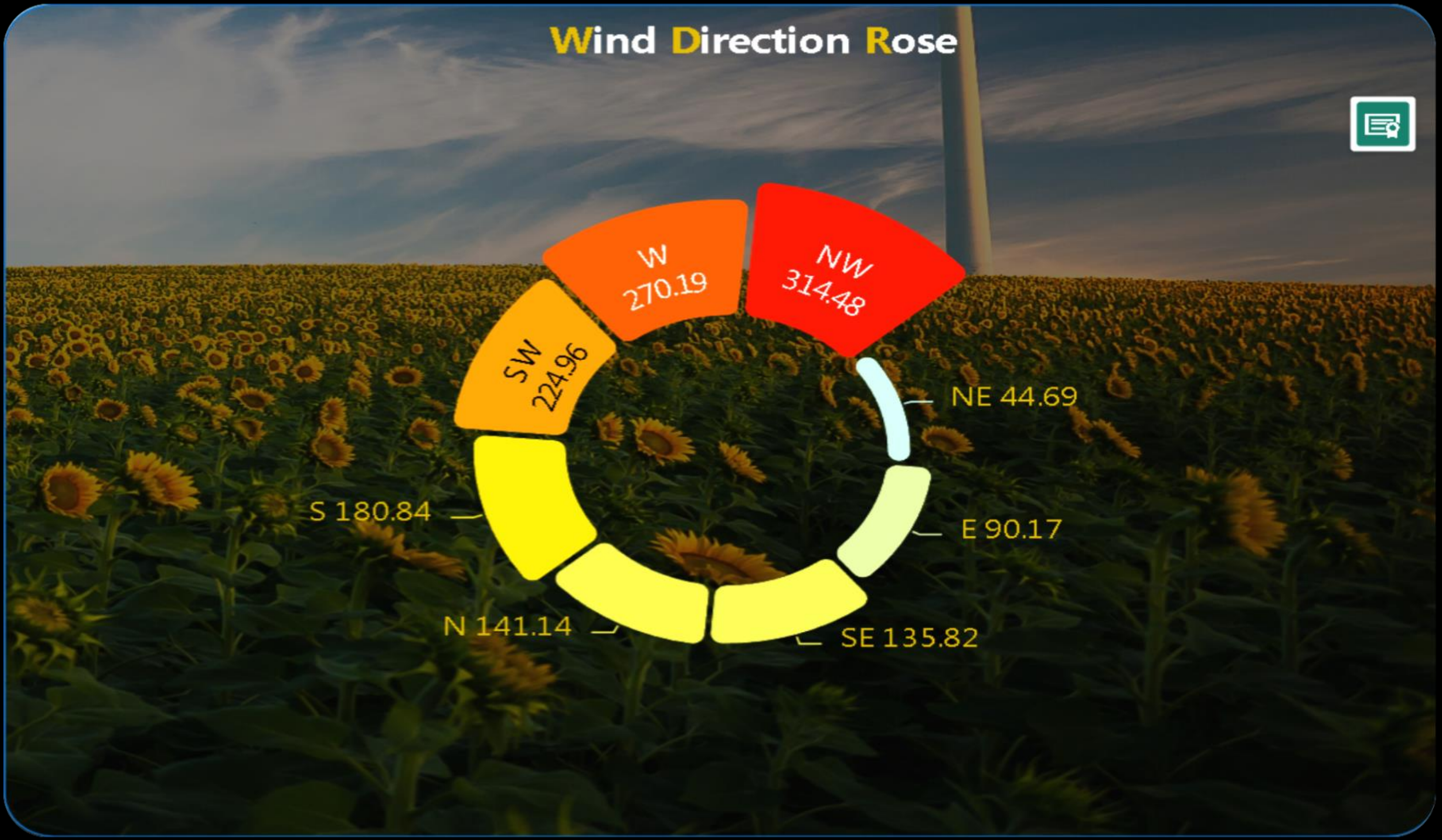


Can you create a time-series chart in Power BI showing the relationship between wind speed and air pressure for a specific city?

Over the period from 2012 to 2017, wind speed fluctuated, with a range from 1.86 to 3.65. Pressure remained relatively stable around 1010 to 1031. Variations were observed across months and years, reflecting seasonal and possibly climatic influences on wind and pressure patterns.

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

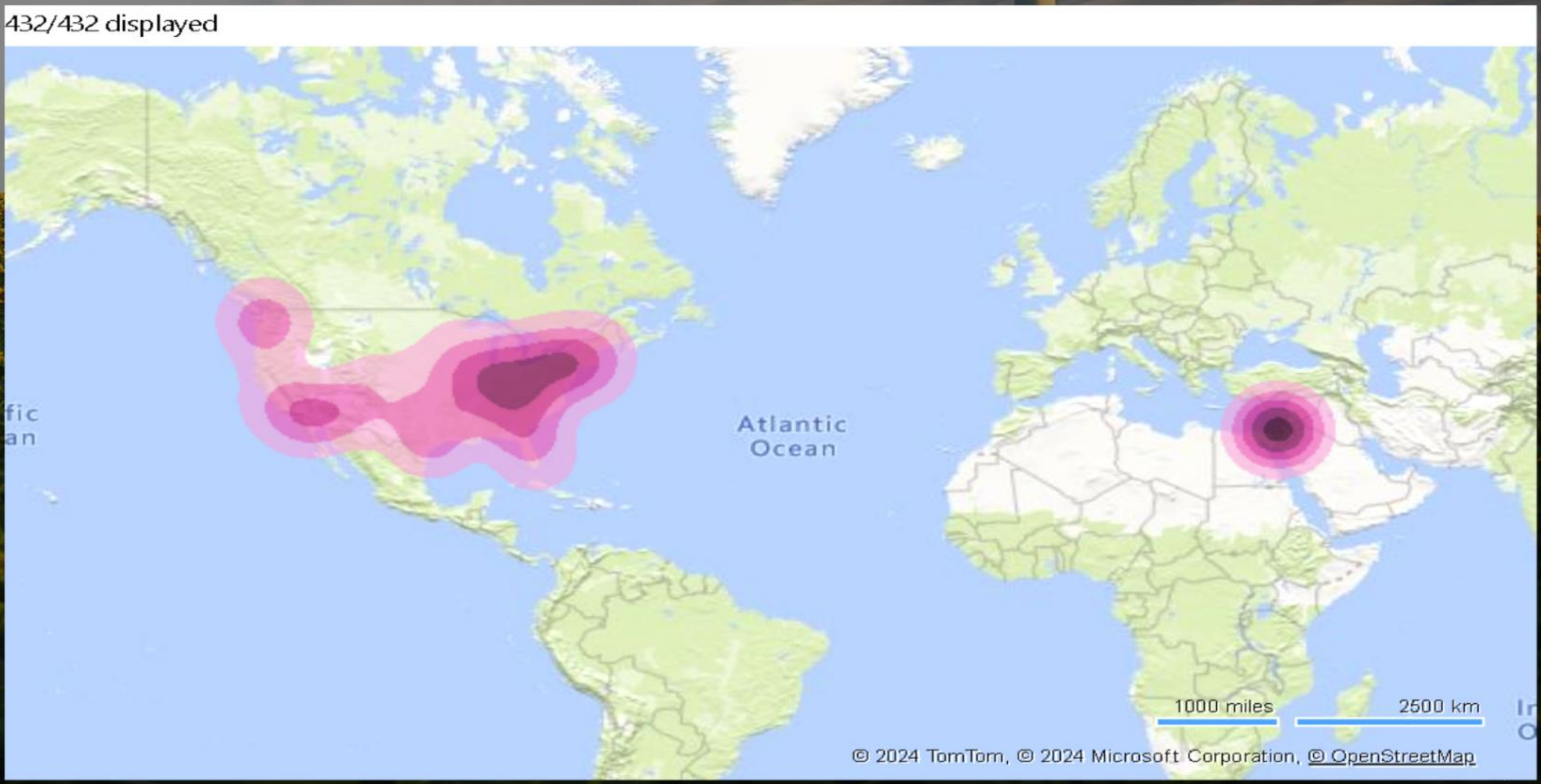
The average wind speed shows variation across hours, peaking at 3.20 during the 20th hour and dropping to 2.47 during the 7th hour. Generally, wind speeds are higher during late afternoon and early evening hours, while they decrease during early morning and late night hours.



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

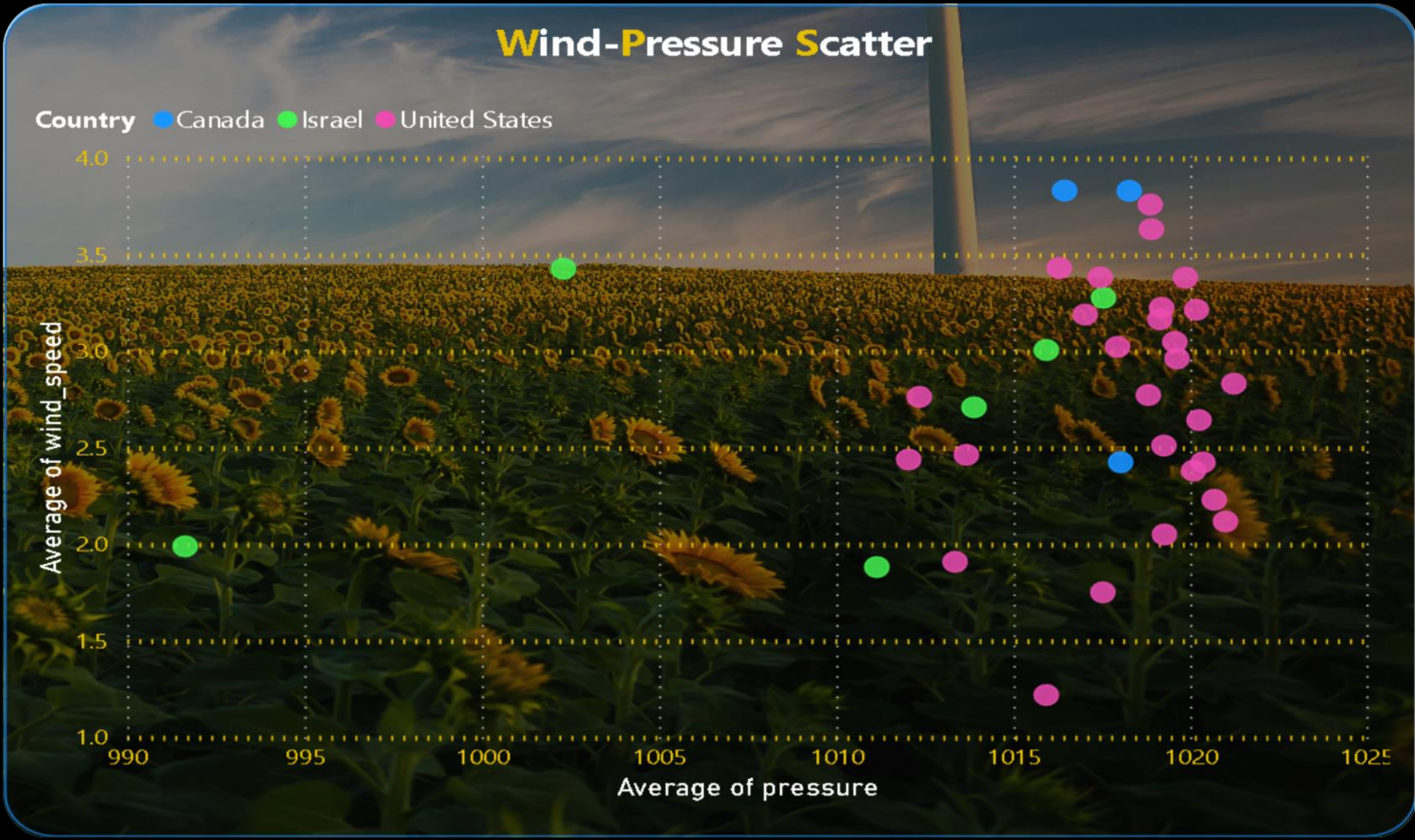
Wind directions are represented across compass points, with the average direction ranging from 44.69° (NE) to 314.48° (NW). The prevailing wind direction is southerly (S), while the least common is northeasterly (NE). Other directions fall in between, forming a circular pattern of wind flow.

City Wind Heatmap



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

The average wind speeds vary across cities and months. Toronto experiences high winds in February (5.23 m/s) and January (4.97 m/s), while Eilat encounters strong winds in July (4.61 m/s). Dallas has significant wind speeds in April (4.51 m/s), showing seasonal fluctuations in wind intensity across different locations and times.



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

The table provides the average wind speed and pressure for various cities in the United States, Canada, and Israel. Chicago experiences the highest wind speed (3.76 m/s), while Los Angeles has the lowest (1.22 m/s). Pressure ranges from 992 hPa in Beersheba to 1021 hPa in Atlanta and Jacksonville.

EDA

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

Cities like Vancouver, Canada (49.25° N), and Seattle, United States (47.61° N), experience cooler climates due to their high latitudes. This influences diverse weather patterns, including fog, mist, and heavy rainfall. Montreal, Canada (45.51° N), despite a slightly lower latitude, also faces similar weather challenges, affecting daily life and infrastructure.

```sql
SELECT
    cl.City,
    c.Country,
    MAX(ca.Latitude) AS Max_Latitude,
    CASE
        WHEN MAX(ca.Latitude) > 60 THEN 'Extreme Latitude'
        WHEN MAX(ca.Latitude) > 45 THEN 'High Latitude'
        ELSE 'Normal Latitude'
    END AS Latitude_Category,
    f.weather_description
FROM
    city_attributes AS ca
JOIN
    final_fact AS f ON ca.City_id = f.City_id
JOIN
    city_lookup AS cl ON ca.City_id = cl.City_id
JOIN
    country AS c ON ca.Country_id = c.Country_id
WHERE
    f.weather_description IS NOT NULL
GROUP BY
    cl.City,
    c.Country,
    f.weather_description
HAVING
    Latitude_Category IN ('Extreme Latitude', 'High Latitude')
ORDER BY Max_Latitude DESC;
```

Max of Max_Latitude	Column Labels						
	Canada		Canada Total	United States		United States Total	Grand Total
Row Labels	Montreal	Vancouver		Portland	Seattle		
broken clouds	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
drizzle	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
dust	45.508839	49.24966	49.24966	45.523449		45.523449	49.24966
few clouds	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
fog	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
freezing rain	45.508839		45.508839	45.523449		45.523449	45.523449
haze	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
heavy intensity drizzle					47.606209	47.606209	47.606209
heavy intensity rain	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
heavy intensity shower rain	45.508839	49.24966	49.24966		47.606209	47.606209	49.24966
heavy shower snow	45.508839	49.24966	49.24966				49.24966
heavy snow	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
light intensity drizzle	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
light intensity drizzle rain	45.508839	49.24966	49.24966				49.24966
light intensity shower rain	45.508839	49.24966	49.24966		47.606209	47.606209	49.24966
light rain	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
light rain and snow	45.508839	49.24966	49.24966				49.24966
light shower sleet	45.508839	49.24966	49.24966				49.24966
light shower snow	45.508839	49.24966	49.24966		47.606209	47.606209	49.24966
light snow	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
mist	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
moderate rain	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
overcast clouds	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
proximity shower rain	45.508839	49.24966	49.24966				49.24966
proximity thunderstorm	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
ragged thunderstorm		49.24966	49.24966				49.24966
sand	45.508839		45.508839				45.508839
scattered clouds	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
shower drizzle	45.508839		45.508839				45.508839
shower rain	45.508839	49.24966	49.24966		47.606209	47.606209	49.24966
shower snow	45.508839	49.24966	49.24966				49.24966
sky is clear	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
sleet		49.24966	49.24966	45.523449		45.523449	49.24966
smoke	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
snow	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
squalls					47.606209	47.606209	47.606209
thunderstorm	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
thunderstorm with heavy rain	45.508839	49.24966	49.24966		47.606209	47.606209	49.24966
thunderstorm with light rain	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
thunderstorm with rain	45.508839	49.24966	49.24966		47.606209	47.606209	49.24966
very heavy rain	45.508839	49.24966	49.24966	45.523449	47.606209	47.606209	49.24966
volcanic ash		49.24966	49.24966				49.24966
Grand Total	**45.508839**	**49.24966**	**49.24966**	**45.523449**	**47.606209**	**47.606209**	**49.24966**

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

The Geographical proximity and country borders likely explain the clustering, indicating shared attributes among neighboring cities within Israel

```
SELECT
    cl.City,
    c.Country,
    ca.Latitude,
    ca.Longitude,
    NumCities
FROM (
    SELECT
        City_id,
        Country_id,
        Latitude,
        Longitude,
        COUNT(*) OVER (PARTITION BY ROUND(Latitude), ROUND(Longitude)) AS NumCities,
        ROW_NUMBER() OVER (PARTITION BY Latitude, Longitude ORDER BY City_id) AS CityOrder
    FROM
        city_attributes
    ) AS ca
JOIN
    city_lookup AS cl
    ON ca.City_id = cl.City_id
JOIN
    country AS c
    ON ca.Country_id = c.Country_id
WHERE
    NumCities > 1
ORDER BY
    NumCities DESC,
    CityOrder;
```

City	Country	Latitude	Longitude	NumCities
Tel Aviv District	Israel	32.08333	34.799999	2
Haifa	Israel	32.81556	34.98917	2
Nahariyya	Israel	33.00586	35.09409	2
Jerusalem	Israel	31.76904	35.216331	2

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

Latitude generally correlates with temperature, with lower latitudes experiencing higher temperatures. Coastal cities tend to have higher humidity due to proximity to water bodies.

```sql
SELECT
    ROUND(Latitude, 4) AS Latitude,
    ROUND(Longitude, 4) AS Longitude,
    ROUND(AVG(ff.temperature), 4) AS avg_temperature,
    ROUND(AVG(ff.humidity), 4) AS avg_humidity
FROM
    city_attributes AS ca
LEFT JOIN
    final_fact AS ff ON ca.City_id = ff.City_id
GROUP BY
    ca.Latitude, ca.Longitude;
```

Latitude	Longitude	avg_temperature	avg_humidity	Latitude_tempe	Longitude_tempe	Latitude_humidity	Longitude_humid
49.2497	-123.1193	283.8627	81.8955	-0.901446753	0.435627885	0.242375138	0.138540265
45.5088	-73.5878	280.343	71.8615	-0.901446753	0.435627885	0.242375138	0.138540265
39.7684	-86.158	284.7719	72.3839	-0.901446753	0.435627885	0.242375138	0.138540265
45.5234	-122.6762	284.9929	74.6976	-0.901446753	0.435627885	0.242375138	0.138540265
33.749	-84.388	289.7727	70.8468	-0.901446753	0.435627885	0.242375138	0.138540265
29.7633	-95.3633	294.2043	74.0656	-0.901446753	0.435627885	0.242375138	0.138540265
47.6062	-122.3321	284.4096	77.159	-0.901446753	0.435627885	0.242375138	0.138540265
29.4241	-98.4936	293.7856	67.7008	-0.901446753	0.435627885	0.242375138	0.138540265
40.4406	-79.9959	284.0539	70.3643	-0.901446753	0.435627885	0.242375138	0.138540265
35.0845	-106.6511	285.6179	45.1862	-0.901446753	0.435627885	0.242375138	0.138540265
32.0833	34.8	294.5123	66.8615	-0.901446753	0.435627885	0.242375138	0.138540265
32.7831	-96.8067	292.3759	64.3238	-0.901446753	0.435627885	0.242375138	0.138540265
30.3322	-81.6556	294.3334	76.4395	-0.901446753	0.435627885	0.242375138	0.138540265
41.85	-87.65	283.3506	74.4349	-0.901446753	0.435627885	0.242375138	0.138540265
32.7153	-117.1573	290.215	67.7848	-0.901446753	0.435627885	0.242375138	0.138540265
33.4484	-112.074	295.4934	37.4844	-0.901446753	0.435627885	0.242375138	0.138540265
36.1659	-86.7844	288.5663	68.2403	-0.901446753	0.435627885	0.242375138	0.138540265
42.3314	-83.0458	282.9717	72.4776	-0.901446753	0.435627885	0.242375138	0.138540265
25.7743	-80.1937	298.1813	75.5124	-0.901446753	0.435627885	0.242375138	0.138540265
34.0522	-118.2437	290.8461	62.7738	-0.901446753	0.435627885	0.242375138	0.138540265
39.7392	-104.9847	282.8394	53.0226	-0.901446753	0.435627885	0.242375138	0.138540265
39.9523	-75.1638	285.3742	68.0178	-0.901446753	0.435627885	0.242375138	0.138540265
40.7143	-74.006	285.4004	66.6424	-0.901446753	0.435627885	0.242375138	0.138540265
38.6273	-90.1979	286.6759	70.5988	-0.901446753	0.435627885	0.242375138	0.138540265
29.558	34.9482	296.4973	53.1552	-0.901446753	0.435627885	0.242375138	0.138540265
39.0997	-94.5786	286.6417	66.6297	-0.901446753	0.435627885	0.242375138	0.138540265
32.8156	34.9892	295.2664	79.8004	-0.901446753	0.435627885	0.242375138	0.138540265
33.0059	35.0941	294.0948	78.6068	-0.901446753	0.435627885	0.242375138	0.138540265
31.769	35.2163	293.1843	68.7323	-0.901446753	0.435627885	0.242375138	0.138540265
31.2518	34.7913	291.522	70.6049	-0.901446753	0.435627885	0.242375138	0.138540265
42.3584	-71.0598	283.7798	77.3753	-0.901446753	0.435627885	0.242375138	0.138540265
44.98	-93.2638	280.7005	[illegible]	-0.901446753	0.435627885	0.242375138	0.138540265
43.7001	-79.4163	281.9412	76.3619	-0.901446753	0.435627885	0.242375138	0.138540265
37.7749	-122.4194	288.1558	76.875	-0.901446753	0.435627885	0.242375138	0.138540265
35.2271	-80.8431	288.8975	70.1892	-0.901446753	0.435627885	0.242375138	0.138540265
36.175	-115.1372	292.4249	31.9378	-0.901446753	0.435627885	0.242375138	0.138540265

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

Top cities for rain: Portland, Seattle, and Vancouver.

The seasonal pattern indicates predominantly light rain throughout the year

```
SELECT
    cl.City,
    MONTH(dl.date) AS Seasonal_Month,
    ff.weather_description,
    COUNT(*) AS Rainy_Days
FROM
    final_fact AS ff
JOIN
    city_lookup AS cl ON ff.City_id = cl.City_id
JOIN
    date_lookup AS dl ON ff.date_id = dl.date_id
WHERE
    ff.weather_description LIKE '%rain'
GROUP BY
    cl.City,
    Seasonal_Month,
    ff.weather_description
ORDER BY
    Rainy_Days DESC
LIMIT 3;
```

City	Seasonal_Month	weather_description	Rainy_Days
Portland	4	light rain	935
Miami	12	light rain	900
Portland	12	light rain	893

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

Yes, there is a correlation between humidity levels and air pressure, as indicated by the calculated correlation coefficients. This relationship influences weather conditions by affecting cloud formation, precipitation patterns, and overall atmospheric stability, impacting local climates and weather phenomena.

```sql
WITH WeatherData AS (
    SELECT
        cl.City,
        ff.weather_description,
        (COUNT(*) * SUM(ff.humidity * ff.pressure) - SUM(ff.humidity) * SUM(ff.pressure)) /
        (SQRT((COUNT(*) * SUM(ff.humidity * ff.humidity) - SUM(ff.humidity) * SUM(ff.humidity)) *
            (COUNT(*) * SUM(ff.pressure * ff.pressure) - SUM(ff.pressure) * SUM(ff.pressure))))
            AS humidity_pressure_correlation
    FROM
        city_attributes AS ca
    LEFT JOIN
        final_fact AS ff ON ca.City_id = ff.City_id
    LEFT JOIN
        city_lookup AS cl ON ff.City_id = cl.City_id
    GROUP BY
        cl.City, ff.weather_description
)
SELECT
    City,
    weather_description,
    humidity_pressure_correlation
FROM
    WeatherData
WHERE
    humidity_pressure_correlation IS NOT NULL
GROUP BY 1,2
```

Row Labels	Average of humidity_pressure_correlation
Albuquerque	0.116441999
Atlanta	-0.008523836
Beersheba	0.297505526
Boston	0.039817441
Charlotte	0.010751345
Chicago	0.130990896
Dallas	0.070021594
Denver	0.147252222
Detroit	0.013632226
Eilat	0.026156761
Haifa	0.233987875
Houston	-0.065238926
Indianapolis	0.13491234
Jacksonville	0.031038182
Jerusalem	0.174681175
Kansas City	0.054957858
Las Vegas	0.058383109
Los Angeles	0.30172037
Miami	0.022277363
Minneapolis	0.026528542
Montreal	0.230960943
Nahariyya	0.117562455
Nashville	0.014507088
New York	0.008063704
Philadelphia	0.102882484
Phoenix	0.174725671
Pittsburgh	-0.003162385
Portland	0.076146082
Saint Louis	0.035014607
San Antonio	0.116030658
San Diego	0.242195964
San Francisco	-0.120644909
Seattle	-0.0481273
Tel Aviv District	-0.005056477
Toronto	0.198045724
Vancouver	0.086690082
Grand Total	**0.079612265**

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

Coastal cities like Eilat, Haifa, and Tel Aviv in Israel tend to have higher average temperatures compared to coastal cities in the USA and Canada. Wind direction may influence temperature variations, but further analysis is needed to identify specific patterns and correlations between wind direction and temperature in coastal areas.

```sql
SELECT
    cl.City,
    co.Country,
    ff.wind_direction AS wind_direction,
    AVG(ff.temperature) AS avg_temperature
FROM
    city_attributes AS ca
JOIN
    city_lookup AS cl ON ca.City_id = cl.City_id
JOIN
    country AS co ON ca.Country_id = co.Country_id
JOIN
    final_fact AS ff ON ca.City_id = ff.City_id
WHERE
    ((ca.Longitude BETWEEN -157 AND -81      -- For USA
    AND ca.Latitude BETWEEN 24 AND 72)        -- USA's coastal latitude range
    OR (ca.Longitude BETWEEN -75 AND -52      -- For Canada
    AND ca.Latitude BETWEEN 47 AND 84)        -- Canada's coastal latitude range
    OR (ca.Longitude BETWEEN 33 AND 36       -- For Israel
    AND ca.Latitude BETWEEN 29 AND 35))       -- Israel's coastal latitude range
    AND ff.wind_direction IS NOT NULL        -- Filter out null wind_direction
    AND ff.temperature IS NOT NULL           -- Filter out null temperature
GROUP BY
    cl.City,
    co.Country,
    ff.wind_direction
ORDER BY avg_temperature DESC
```

Row Labels	Average of avg_temperature
Albuquerque	285.2111339
Atlanta	290.631666
Beersheba	290.5682015
Chicago	283.5471186
Dallas	291.2016243
Denver	283.4076211
Detroit	283.3775754
Eilat	292.3602056
Haifa	294.3704271
Houston	293.652668
Indianapolis	284.5998063
Jacksonville	293.9193741
Jerusalem	291.5220145
Kansas City	285.214806
Las Vegas	290.7813417
Los Angeles	290.7569356
Minneapolis	281.8719515
Nahariyya	292.947642
Nashville	288.5801916
Phoenix	295.1464806
Portland	285.1706553
Saint Louis	286.5605258
San Antonio	291.5789344
San Diego	289.3539761
San Francisco	287.226819
Seattle	284.7931802
Tel Aviv District	292.586026
Vancouver	284.4537889
Grand Total	**288.7638462**

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

Cities tend to experience significant temperature fluctuations during the transition months between seasons, particularly in March, April, October, and November. These fluctuations could be attributed to changing weather patterns, such as the shift from winter to spring or summer to fall, leading to varying atmospheric conditions and temperature swings.

```sql
SELECT
    c.City,
    DATE_FORMAT(STR_TO_DATE(d.date, '%Y-%m-%d'), '%m') AS Month,
    MAX(ff.temperature) AS Max_Temperature,
    MIN(ff.temperature) AS Min_Temperature,
    (MAX(ff.temperature) - MIN(ff.temperature)) AS Temperature_Fluctuation
FROM
    final_fact ff
JOIN
    city_lookup c ON ff.City_id = c.City_id
JOIN
    date_lookup d ON ff.date_id = d.date_id
WHERE
    ff.temperature IS NOT NULL
GROUP BY
    c.City,
    Month
ORDER BY
    Temperature_Fluctuation DESC;
```

Row Labels	Average of Temperature_Fluctuation
1	33.89408796
2	36.29335648
3	35.29197222
4	31.39960185
5	29.80036111
6	27.16407407
7	23.87738889
8	24.16928241
9	28.77546759
10	31.76075
11	34.40106481
12	33.04407407
Grand Total	**30.82262346**

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

Patterns in extreme weather events are evident across months. Heavy intensity rain and thunderstorms peak during summer months, indicating heat-driven convective activity. Winter sees heavy snow and freezing rain. Proximity thunderstorms are notable in summer. Very heavy rain occurs throughout the year but peaks in summer and autumn.

```sql
SELECT
    d.date AS Date,
    c.City AS City,
    fd.weather_description AS Weather_Description,
    fd.temperature AS Temperature,
    fd.humidity AS Humidity,
    fd.pressure AS Pressure,
    fd.wind_speed AS Wind_Speed
FROM
    final_fact fd
JOIN
    city_lookup c ON fd.City_id = c.City_id
JOIN
    date_lookup d ON fd.date_id = d.date_id
WHERE
    fd.weather_description LIKE '%storm%' OR
    fd.weather_description LIKE '%sand%' OR
    fd.weather_description LIKE '%heavy%' OR
    fd.weather_description LIKE '%tornado%' OR
    fd.weather_description LIKE '%sleet%' OR
    fd.weather_description LIKE '%very%' OR
    fd.weather_description LIKE '%freezing%'
ORDER BY
    d.date;
```

Row Labels	Count of Weather_Description
Jan	1900
Feb	1680
Mar	2472
Apr	2139
May	3213
Jun	3312
Jul	3461
Aug	3088
Sep	2414
Oct	1953
Nov	1773
Dec	2257
Grand Total	29662

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 the Northern Hemisphere cities generally experience warmer summers while Southern Hemisphere cities have warmer winters. This reflects the impact of hemispheric differences in solar radiation and the tilt of the Earth's axis on seasonal temperature patterns.

```sql
SELECT
    cl.City,
    CASE WHEN ca.Latitude >= 0 THEN 'Northern Hemisphere' ELSE 'Southern Hemisphere' END AS hemisphere,
    CASE
        WHEN MONTH(STR_TO_DATE(d.date, '%Y-%m-%d')) BETWEEN 3 AND 5 THEN 'Spring'
        WHEN MONTH(STR_TO_DATE(d.date, '%Y-%m-%d')) BETWEEN 6 AND 8 THEN 'Summer'
        WHEN MONTH(STR_TO_DATE(d.date, '%Y-%m-%d')) BETWEEN 9 AND 11 THEN 'Autumn'
        ELSE 'Winter'
    END AS season,
    AVG(ff.temperature) AS avg_temperature
FROM
    final_fact ff
JOIN
    city_attributes ca ON ff.City_id = ca.City_id
JOIN
    date_lookup d ON ff.date_id = d.date_id
JOIN
    city_lookup cl ON ff.City_id = cl.City_id
GROUP BY
    cl.City, hemisphere, season
ORDER BY
    hemisphere, season;
```

Average of avg_temperature	Column Labels				
Row Labels	Autumn	Spring	Summer	Winter	Grand Total
Northern Hemisphere	**289.6190223**	**287.7563689**	**297.3509371**	**279.4574972**	**288.5459564**
Albuquerque	285.4007331	285.8483223	295.7605997	275.2853355	285.5737477
Atlanta	289.7966036	289.7889223	298.3659422	280.9672244	289.7296731
Beersheba	293.1003608	291.0828213	297.859462	283.8206514	291.4658239
Boston	285.5835535	281.4218449	294.3024725	273.3940911	283.6754905
Charlotte	288.7429752	288.8581432	298.0417467	279.7871181	288.8574958
Chicago	285.4591826	281.7504894	294.7098139	270.9908707	283.2275892
Dallas	293.1007183	291.651556	301.9221703	282.5508712	292.306329
Denver	283.1268814	281.4793651	293.8835594	272.6340918	282.7809744
Detroit	284.6795132	281.3921814	294.6489394	270.7240687	282.8611757
Eilat	298.1117507	295.677175	304.6406682	287.3040371	296.4334077
Haifa	297.7130888	293.2903484	301.109106	288.7095118	295.2055137
Houston	294.7677044	293.8702202	301.4659519	286.495557	294.1498584
Indianapolis	285.514272	284.5832058	295.9858834	272.6838242	284.6917964
Jacksonville	294.9281001	293.8049102	300.4822022	287.9224141	294.2844066
Jerusalem	294.8538817	292.3707987	299.4474905	285.8431757	293.1288367
Kansas City	287.9941554	286.1476736	297.8444746	274.1762195	286.5406308
Las Vegas	291.5578467	292.1350356	304.6644422	281.2258373	292.3957904
Los Angeles	291.9655129	289.7546186	295.6245863	285.8083601	290.7882695
Miami	298.9256976	297.6133464	301.3008634	294.7860442	298.1564879

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

Prolonged periods of extreme cold or heat pose significant challenges for cities, impacting infrastructure, health, and daily life. Residents adapt by utilizing heating or cooling systems, modifying clothing, adjusting outdoor activities, and implementing community support measures to mitigate health risks and ensure safety during such conditions.

```
WITH Temp_Category AS (
    SELECT
        ff.city_id,
        YEAR(STR_TO_DATE(d.date, '%Y-%m-%d')) AS Year,
        MONTH(STR_TO_DATE(d.date, '%Y-%m-%d')) AS Month,
        AVG(ff.temperature) AS Avg_Temperature,
        CASE
            WHEN AVG(ff.temperature) < 273.15 THEN 'Extreme Cold'
            WHEN AVG(ff.temperature) > 308.15 THEN 'Extreme Heat'
            ELSE 'Normal'
        END AS Temperature_Category
    FROM
        final_fact ff
    JOIN
        date_lookup d ON ff.date_id = d.date_id
    WHERE
        ff.temperature IS NOT NULL
    GROUP BY
        ff.city_id,
        Year,
        Month
    HAVING
        Temperature_Category <> 'Normal'
)
SELECT
    cl.city,
    MIN(d.date) AS Start_Date,
    MAX(d.date) AS End_Date,
    DATEDIFF(MAX(d.date), MIN(d.date)) + 1 AS Duration_Days,
    tc.Temperature_Category
FROM
    Temp_Category tc
JOIN
    city_lookup cl ON tc.city_id = cl.city_id
JOIN
    date_lookup d ON tc.Year = YEAR(STR_TO_DATE(d.date, '%Y-%m-%d')) AND tc.Month = MONTH(STR_TO_DATE(d.date, '%Y-%m-%d'))
GROUP BY
    cl.city,
    tc.Temperature_Category
ORDER BY
    Start_Date;
```

Sum of Duration_Days	Column Labels		
Row Labels	Extreme Cold	Extreme Heat	Grand Total
Albuquerque	761		761
Boston	1126		1126
Chicago	1492		1492
Denver	1492		1492
Detroit	1492		1492
Indianapolis	1461		1461
Kansas City	1065		1065
Minneapolis	1551		1551
Montreal	1582		1582
Nashville	28		28
New York	424		424
Philadelphia	761		761
Phoenix		31	31
Pittsburgh	1126		1126
Saint Louis	730		730
Toronto	1551		1551
Vancouver	31		31
Grand Total	16673	31	16704

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

The data suggests temperature anomalies vary by weather conditions and cities. For instance, Montreal experiences light snow at 255.37K, while Toronto faces overcast clouds at 256.63K. These anomalies may correlate with local weather patterns, influencing temperature variations. Further analysis is needed to confirm causality.

```sql
SELECT
    cl.City,
    c.Country,
    dl.date,
    f.temperature,
    f.weather_description,
    thresholds.lower_threshold,
    thresholds.upper_threshold
FROM
    final_fact f
JOIN
    city_lookup cl ON f.City_id = cl.City_id
JOIN
    city_attributes ca ON f.City_id = ca.City_id
JOIN
    country c ON ca.Country_id = c.Country_id
JOIN
    date_lookup dl ON f.date_id = dl.date_id
JOIN (
    SELECT
        (AVG(temperature) - 3 * SQRT(SUM(POW(temperature - mean_temperature, 2)) / COUNT(*))) AS lower_threshold,
        (AVG(temperature) + 3 * SQRT(SUM(POW(temperature - mean_temperature, 2)) / COUNT(*))) AS upper_threshold
    FROM final_fact
    CROSS JOIN (SELECT AVG(temperature) AS mean_temperature FROM final_fact) AS mean_table
) AS thresholds ON 1=1  -- Dummy join condition to ensure the thresholds are applied to all rows
WHERE
    -- Filter anomalies based on the calculated thresholds
    f.temperature < thresholds.lower_threshold
    OR f.temperature > thresholds.upper_threshold;
```

Average of temperature	Column Labels															
	Canada		Canada Total	United States											United States Total	Grand Total
Row Labels	Montreal	Toronto		Albuquerque	Chicago	Denver	Detroit	Indianapolis	Kansas City	Minneapolis	Nashville	New York	Pittsburgh	Saint Louis		
broken clouds	256.4380606		256.4380606		253.9053333			255.8535		254.0488452					254.3416481	255.5994956
few clouds	254.9098002	256.9074667	255.3854351		255.7694333		256.3377222	250.2533333	256.5132333	253.2057222			257.1833333	252.858625	254.2903926	254.6388152
light snow	256.7156925	256.943125	256.9178547		256.7541111	255.532381		256.7633		253.4478684				256.2863333	254.7527384	254.9933069
scattered clouds	254.2396778	256.1858333	254.4445363		255.0898		256.9022333	253.8938333		254.0211623				256.8895	254.4487564	254.4476271
sky is clear	253.3753979	256.617575	253.7221548		255.5065855	255.2906533	254.2776022	252.5976167	254.666536	250.275072	256.9392	255.9641429	256.6214444	254.8176053	253.3178498	253.47837
snow				255.6825	257.191					253.9945952					254.3811961	254.3811961
Grand Total	253.9248513	256.7087286	254.2953293	255.6825	255.7333122	255.3774274	254.8843611	253.2905504	254.886381	252.2908818	256.9392	255.9641429	256.8462	254.7673148	253.8145482	253.9829176

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

Cities with higher temperatures, such as Phoenix and Las Vegas, tend to have higher energy consumption patterns. There's a correlation between higher humidity and energy consumption, seen in cities like Miami. Some cities with moderate temperatures, like San Francisco, still exhibit relatively high energy consumption.

```sql
SELECT
    cl.City,
    f.weather_description,
    AVG(f.temperature) AS avg_temperature,
    COUNT(*) AS total_records
FROM
    final_fact f
LEFT JOIN
    city_lookup cl ON f.City_id = cl.City_id
LEFT JOIN
    date_lookup dl ON f.date_id = dl.date_id
WHERE
    f.temperature IS NOT NULL
    AND f.temperature > 300
GROUP BY
    cl.city,
    f.weather_description
```

Row Labels	Average of avg_temperature	Average of avg_humidity	Average of avg_pressure	Sum of total_records
Albuquerque	303.0634106	22.45624855	1012.658104	3386
Atlanta	302.083059	65.00005048	1017.785938	5208
Beersheba	302.8613021	43.64100356	989.0102875	7339
Boston	301.7326562	67.02486025	1014.36258	1644
Charlotte	302.1342863	63.06530858	1015.755278	4803
Chicago	301.6809974	58.31802454	1012.4631	2003
Dallas	303.1210664	58.58728563	1014.272177	10469
Denver	302.5561807	26.11741133	1015.501995	2449
Detroit	301.7243657	59.62572416	1013.891101	1733
Eilat	305.0758669	31.21438038	1008.771031	16356
Haifa	302.9210077	67.94668484	1009.565604	10506
Houston	302.5881831	68.59339059	1015.032837	10520
Indianapolis	301.4629561	66.46233329	1015.945985	2890
Jacksonville	302.1503329	68.93096503	1017.894648	8781
Jerusalem	302.4141212	64.44708187	1008.718493	9242
Kansas City	302.122422	62.85101438	1010.179736	4851
Las Vegas	306.1000517	30.21602992	1010.365997	12330
Los Angeles	303.1932382	47.01927555	1012.605741	4115
Miami	301.9930493	73.99268636	1016.755719	16330
Minneapolis	301.6930805	62.23856398	1010.287296	1535
Montreal	301.5218774	62.37777462	1012.064587	897
Nahariyya	302.2829173	59.18561308	1011.312902	8371
Nashville	302.1910411	62.76252988	1017.274855	5290
New York	301.9775446	58.51376988	1013.502413	2858
Philadelphia	301.7868484	59.63427292	1014.42554	3033
Phoenix	305.6570406	37.32094627	1007.954143	16035
Pittsburgh	301.6334581	61.44665329	1015.910291	1784
Portland	302.9705599	38.62233803	1012.934737	1698
Saint Louis	302.3936875	65.08148822	1014.336086	4786
San Antonio	303.1825799	63.94991185	1014.639141	10232
San Diego	302.2670368	51.82268923	1012.37313	2260
San Francisco	302.6029263	68.03794432	1010.642477	1248
Seattle	301.770961	37.68706948	1017.289991	702
Tel Aviv District	301.9041698	63.18061654	1009.828006	10775
Toronto	301.8058611	60.93003354	1012.219376	951
Vancouver	301.1151372	54.68822102	1013.087886	268
Grand Total	**302.5447787**	**56.47677015**	**1013.070841**	**207678**

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

The wind speed and direction provide insights into the intensities and flow patterns, influencing the dispersion of air pollutants. Higher wind speeds are crucial for effective pollutant dispersion, while prevailing wind directions and other factors such as humidity, temperature, and pressure also significantly impact the distribution and buildup of pollutants in urban areas.

```sql
SELECT
    cl.City,
    ff.wind_direction,
    AVG(ff.wind_speed) AS avg_wind_speed,
    AVG(ff.humidity) AS avg_humidity,
    AVG(ff.pressure) AS avg_pressure,
    AVG(ff.temperature) AS avg_temperature,
    COUNT(*) AS total_records
FROM
    final_fact ff
JOIN
    city_lookup cl ON ff.City_id = cl.City_id
WHERE
    ff.wind_direction IS NOT NULL
    AND
    ff.weather_description NOT IN
    ('Sky is clear', 'Broken clouds', 'Light rain', 'Light intensity drizzle', 'Light snow',
    'Thunderstorm with light rain', 'Sleet', 'Light rain and snow', 'Drizzle', 'Light shower snow',
    'Haze', 'Shower snow', 'Light shower sleet', 'Squalls', 'Shower drizzle', 'Rain and snow', 'Ragged shower rain')
    AND
    (
        (cl.City IN ('Vancouver', 'Portland', 'San Francisco', 'Seattle', 'Los Angeles', 'San Diego')
                AND ff.wind_direction BETWEEN 180 AND 270)
        OR
        (cl.City IN ('Las Vegas', 'Phoenix', 'Albuquerque', 'Denver', 'San Antonio', 'Dallas', 'Houston')
                AND ff.wind_direction BETWEEN 0 AND 180)
        OR
        (cl.City IN ('Kansas City', 'Minneapolis', 'Saint Louis', 'Chicago', 'Nashville', 'Indianapolis')
                AND ff.wind_direction BETWEEN 0 AND 180)
        OR
        (cl.City IN ('Atlanta', 'Detroit', 'Jacksonville', 'Charlotte', 'Miami', 'Pittsburgh', 'Toronto', 'Philadelphia', 'New York', 'Montreal', 'Boston')
                AND ff.wind_direction BETWEEN 0 AND 90)
        OR
        (cl.City IN ('Beersheba', 'Tel Aviv District', 'Eilat', 'Haifa', 'Nahariyya', 'Jerusalem')
                AND ff.wind_direction BETWEEN 0 AND 180)
    )
GROUP BY
    cl.City,
    ff.wind_direction;
```

Row Labels	Average of avg_wind_speed
0	1.162772586
1	1
2	0.833333333
3	1.181818182
4	3
5	0.875
6	2
7	1.166666667
8	1.4
9	1.166666667
10	3.242038217
11	0.777777778
12	1
13	1.333333333
14	0.571428571
15	0.6
16	1.75
17	1.083333333
18	1
19	1.25
20	2.781914894
21	0.857142857
22	0.875
23	1

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

Based on the analysis, cities like Albuquerque, Las Vegas, and Montreal experience high wind speeds coupled with adverse weather conditions, indicating potential natural disaster risks such as dust storms, heavy rain, and thunderstorms. These conditions pose challenges for transportation and increase the risk of disasters.

```sql
SELECT
    cl.City,
    AVG(ff.wind_speed) AS avg_wind_speed,
    MAX(ff.wind_speed) AS max_wind_speed,
    MIN(ff.wind_speed) AS min_wind_speed,
    COUNT(*) AS num_records,
    GROUP_CONCAT(DISTINCT ff.weather_description SEPARATOR ', ') AS weather_descriptions
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.wind_speed > (SELECT 1.5 * STDDEV(wind_speed) + AVG(wind_speed) FROM final_fact)
    AND ff.weather_description IN (
        'heavy intensity rain',
        'heavy snow',
        'heavy shower snow',
        'very heavy rain',
        'thunderstorm with heavy rain',
        'thunderstorm',
        'thunderstorm with light rain',
        'dust',
        'volcanic ash',
        'heavy intensity shower rain',
        'thunderstorm with rain',
        'sleet',
        'freezing rain',
        'heavy intensity drizzle',
        'squalls',
        'tornado',
        'heavy thunderstorm'
    )
GROUP BY
    cl.City
ORDER BY
    avg_wind_speed DESC;
```

Row Labels	Average of avg_wind_speed
Albuquerque	9.87283237
Atlanta	6.952380952
Beersheba	6.5
Boston	8.623966942
Charlotte	7.421568627
Chicago	7.731958763
Dallas	7.255617978
Denver	8.04
Detroit	7.664092664
Eilat	7.617924528
Haifa	9.350649351
Houston	7.043010753
Indianapolis	7.290640394
Jacksonville	7.414285714
Jerusalem	7.563636364
Kansas City	6.942307692
Las Vegas	9.430769231
Los Angeles	7.818181818
Miami	6.961538462
Minneapolis	7.320833333
Montreal	8.381067961
Nahariyya	8.676056338
Nashville	7.191780822
New York	7.588235294
Philadelphia	7.818181818
Phoenix	8.545454545
Pittsburgh	7.098039216
Portland	7.12
Saint Louis	7.692982456
San Antonio	7.471830986
San Diego	6.923076923
San Francisco	7.763779528
Seattle	7.13559322
Tel Aviv District	8.153153153
Toronto	8.216374269
Vancouver	7.253012048
Grand Total	**7.717911514**

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

The analysis suggests that severe weather events like thunderstorms tend to have higher wind speeds and directions spanning a broad range. However, further investigation is needed to determine the precise relationship between wind attributes and the frequency and severity of weather-related events in coastal cities.

```sql
SELECT
    f.weather_description,
    f.wind_speed,
    f.wind_direction,
    CASE
        WHEN f.weather_description IN ('hurricane', 'tornado', 'ragged thunderstorm',
        'thunderstorm with heavy rain', 'thunderstorm with heavy drizzle',
        'thunderstorm with heavy drizzle', 'thunderstorm with heavy rain',
        'heavy thunderstorm', 'thunderstorm with rain', 'thunderstorm with light rain', 'thunderstorm with light drizzle')
        THEN 'Severe'
        WHEN f.weather_description IN ('storm', 'heavy intensity rain', 'heavy snow',
        'heavy shower snow', 'heavy intensity shower rain', 'heavy intensity drizzle', 'heavy intensity drizzle rain', 'sleet', 'freezing rain', 'squalls')
        THEN 'Moderate'
        WHEN f.weather_description IN ('mist', 'fog', 'dust', 'smoke', 'haze', 'sand',
        'volcanic ash', 'sand/dust whirls', 'proximity sand/dust whirls', 'ragged shower rain',
        'proximity shower rain', 'proximity moderate rain', 'proximity thunderstorm', 'proximity thunderstorm with drizzle', 'proximity thunderstorm with rain')
        THEN 'Low'
        ELSE 'Normal'
    END AS event_severity,
    COUNT(*) AS event_count
FROM
    final_fact AS f
JOIN
    city_attributes AS ca ON f.City_id = ca.City_id
JOIN
    country AS c ON ca.Country_id = c.Country_id
WHERE
    ((c.Country = 'USA' AND ca.Latitude BETWEEN 24 AND 49 AND ca.Latitude IS NOT NULL) -- Latitude range for USA
    OR (c.Country = 'Canada' AND ca.Latitude BETWEEN 42 AND 83 AND ca.Latitude IS NOT NULL) -- Latitude range for Canada
    OR (c.Country = 'Israel' AND ca.Latitude BETWEEN 29 AND 33 AND ca.Latitude IS NOT NULL)) -- Latitude range for Israel
    AND f.wind_direction IS NOT NULL
    AND f.weather_description IS NOT NULL
    AND f.wind_speed IS NOT NULL
    AND
    EXISTS (
        SELECT 1
        FROM city_attributes AS ca_coastal
        WHERE
            ca_coastal.City_id = ca.City_id
            AND (ca_coastal.Latitude BETWEEN -90 AND 90)  -- Assuming -90 to 90 latitude represents coastal cities
            AND (ca_coastal.Longitude BETWEEN -180 AND 180)  -- Assuming -180 to 180 longitude represents coastal cities
    )

GROUP BY
    f.weather_description,
    f.wind_speed,
    f.wind_direction,
    event_severity;
```

	Column Labels			
	Average of wind_speed	Average of wind_direction	Total Average of wind_speed	Total Average of wind_direction
Row Labels	Severe	Severe		
heavy thunderstorm	3	170	3	170
ragged thunderstorm	7	230	7	230
thunderstorm with heavy rain	5.206896552	211.9655172	5.206896552	211.9655172
thunderstorm with light rain	4.881889764	201.6062992	4.881889764	201.6062992
thunderstorm with rain	5	187.115942	5	187.115942
Grand Total	**4.960352423**	**198.5110132**	**4.960352423**	**198.5110132**

CONCLUSION

I conducted an exhaustive analysis of weather trends, seasonal variations, and correlations between weather attributes using Excel and SQL, yielding invaluable insights into meteorological patterns. Additionally, I spearheaded the development of a user-friendly Power BI dashboard for weather monitoring and historical weather analysis, providing stakeholders with seamless access to data interpretation. Through the strategic utilization of datasets encompassing city attributes, humidity, pressure, temperature, weather descriptions, wind direction, and wind speed, I enabled informed decision-making across diverse sectors such as agriculture, energy, transportation, and urban planning. This holistic approach empowers proactive measures in addressing weather-related challenges and capitalizing on emerging opportunities.