Weather_analysis_capstone_project

By Moumita Biswas

Contents

Introduction of data

This weather analysis dataset compiles a comprehensive range of meteorological parameters, including temperature, precipitation, wind speed, and atmospheric pressure, providing a thorough examination of weather patterns over time and space. With its detailed insights, it serves as a vital tool for researchers, policymakers, and industries seeking to understand and adapt to changing climate conditions. By leveraging this data, stakeholders can make informed decisions in areas such as agriculture, infrastructure planning, and emergency response, ultimately enhancing resilience and sustainability in the face of climate variability and extremes.

Insights from this data

Before diving into the specific questions, it's crucial to outline the potential insights we aim to derive from the weather data analysis. Firstly, we seek to understand global geographical patterns, examining how cities are distributed based on latitude and longitude coordinates. This analysis can shed light on urbanization trends and population density across different regions. Additionally, exploring climate variability across continents will provide insights into regional weather patterns. By analyzing temperature, humidity, wind speed, and air pressure trends over time, we aim to identify patterns, anomalies, and extreme weather events. Moreover, correlational analysis between weather variables can reveal their interplay and potential impacts. Spatial and temporal analyses will allow us to visualize variations and trends in weather conditions across different geographical locations and timeframes. Ultimately, through comparative analysis and impact assessment, we can understand the implications of weather conditions on various sectors and develop predictive models for future weather trends.

Overview of Weather_analysis dataset

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

Data cleaning

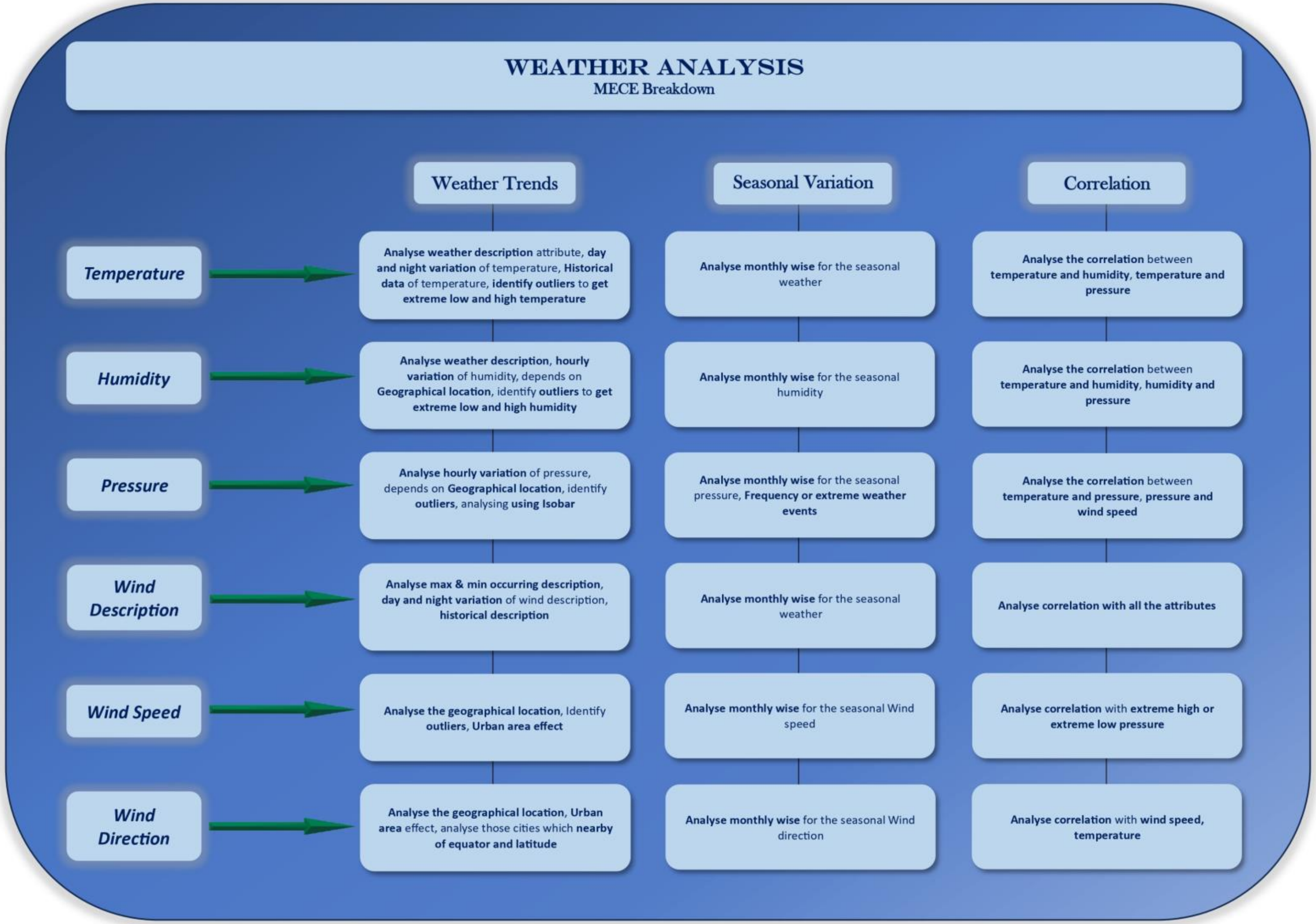
- There were many null values almost in each attributes, replaced those values with `N/A'
- After processing the raw data, made a new dataset called `Weather_analysis_transformed'
- Inside the dataset, made four lookup tables:
  - City_lookup
  - Country_lookup
  - Date_lookup
  - Time_lookup
- Made a fact table, called Final_fact, including 3 foreign keys

Overview of Weather_analysis_transformed dataset

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

MECE Breakdown

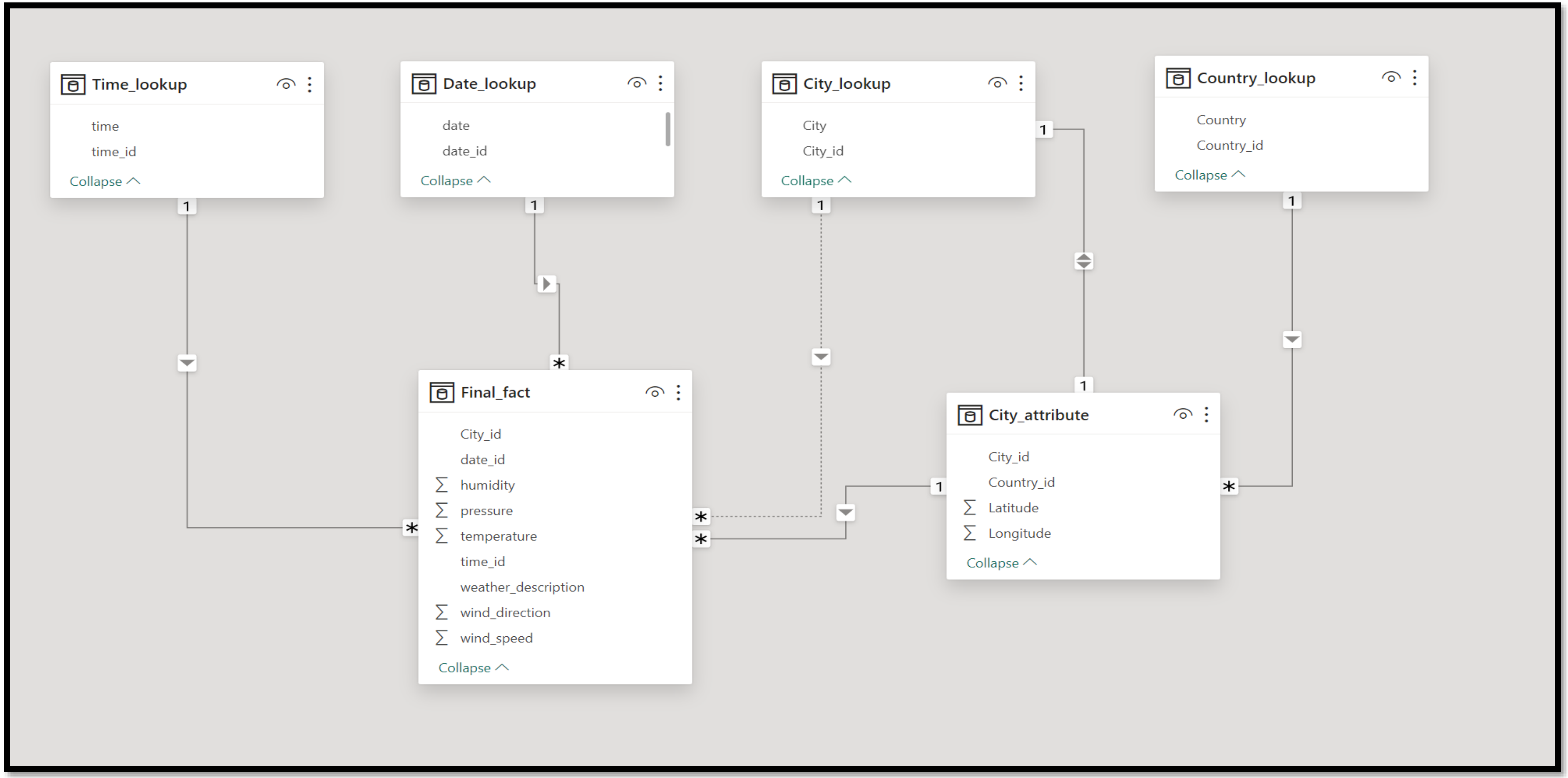
- *After understanding the data I've create a MECE breakdown for better understanding & to proceed further.*



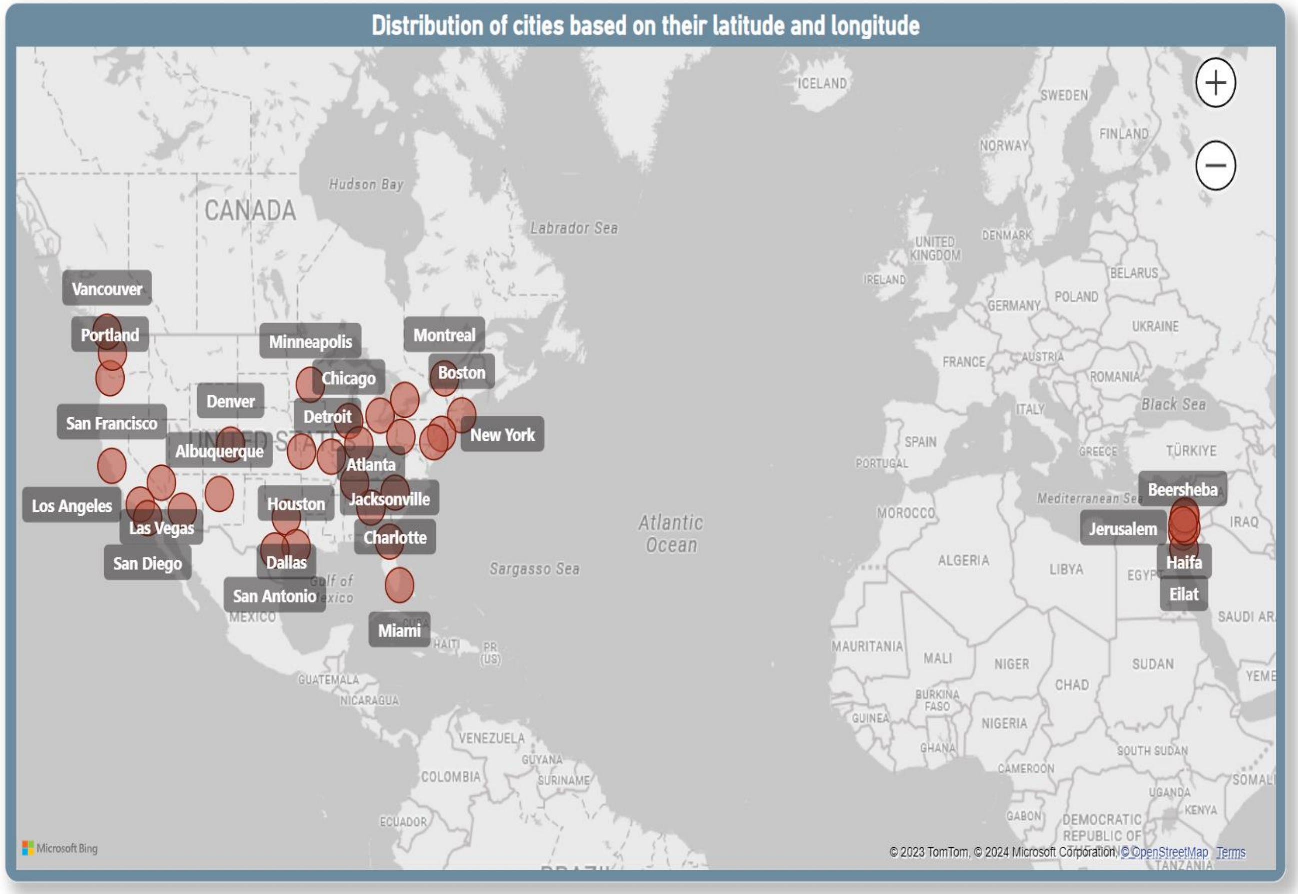
Understanding & solving PowerBI questions

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

ER-Diagram for weather_analysis_transformed dataset



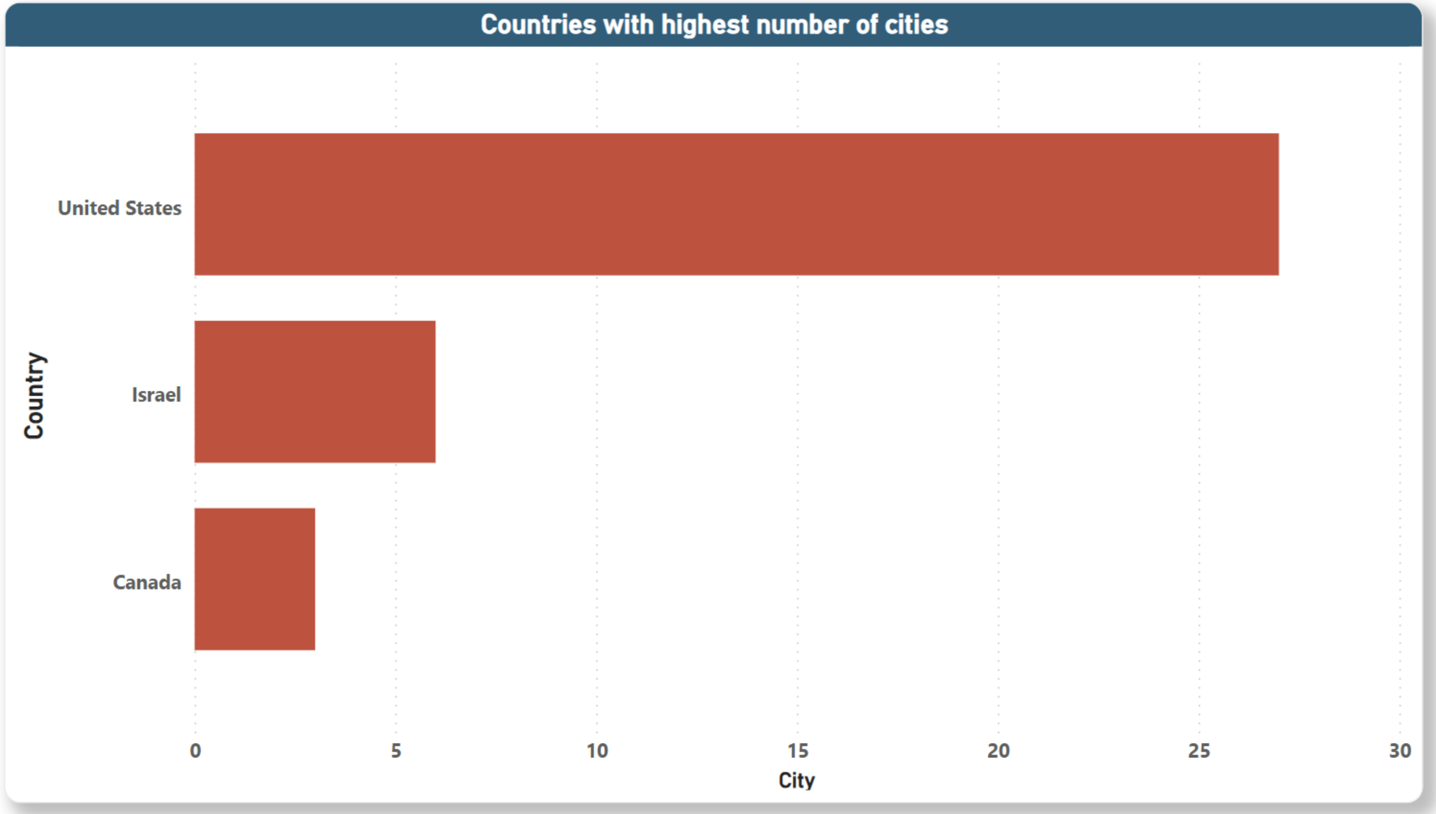
1. Can you create a geographical map in Power BI showing the distribution of cities in the dataset based on their latitude and longitude?



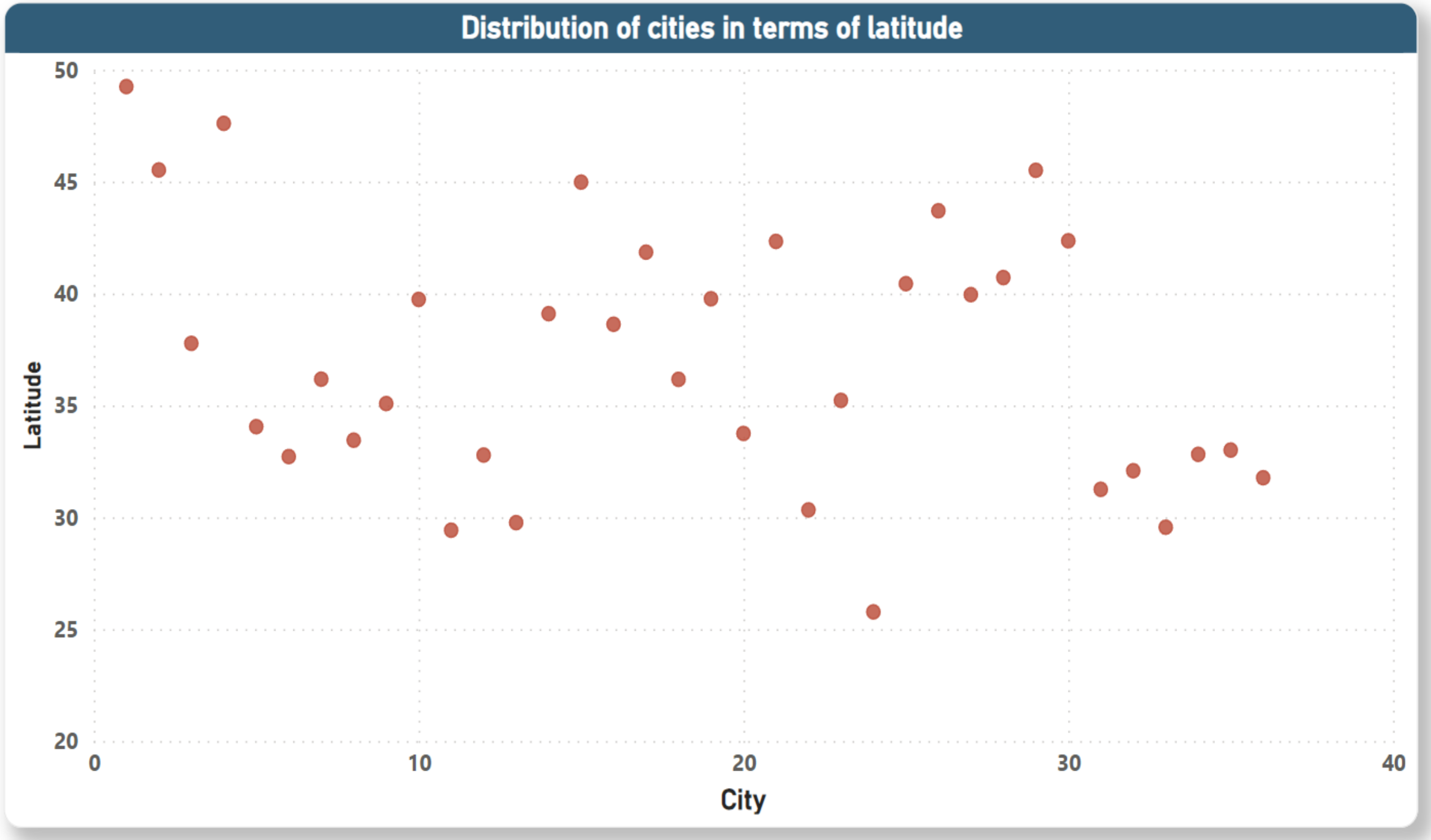
Created a geographical map to show the distribution of cities base on their latitude and longitude.

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

Unfortunately, we have only 3 countries consisting 36 cities. So, here I created a bar chart to show the countries in descending order to get the country with highest number of cities.

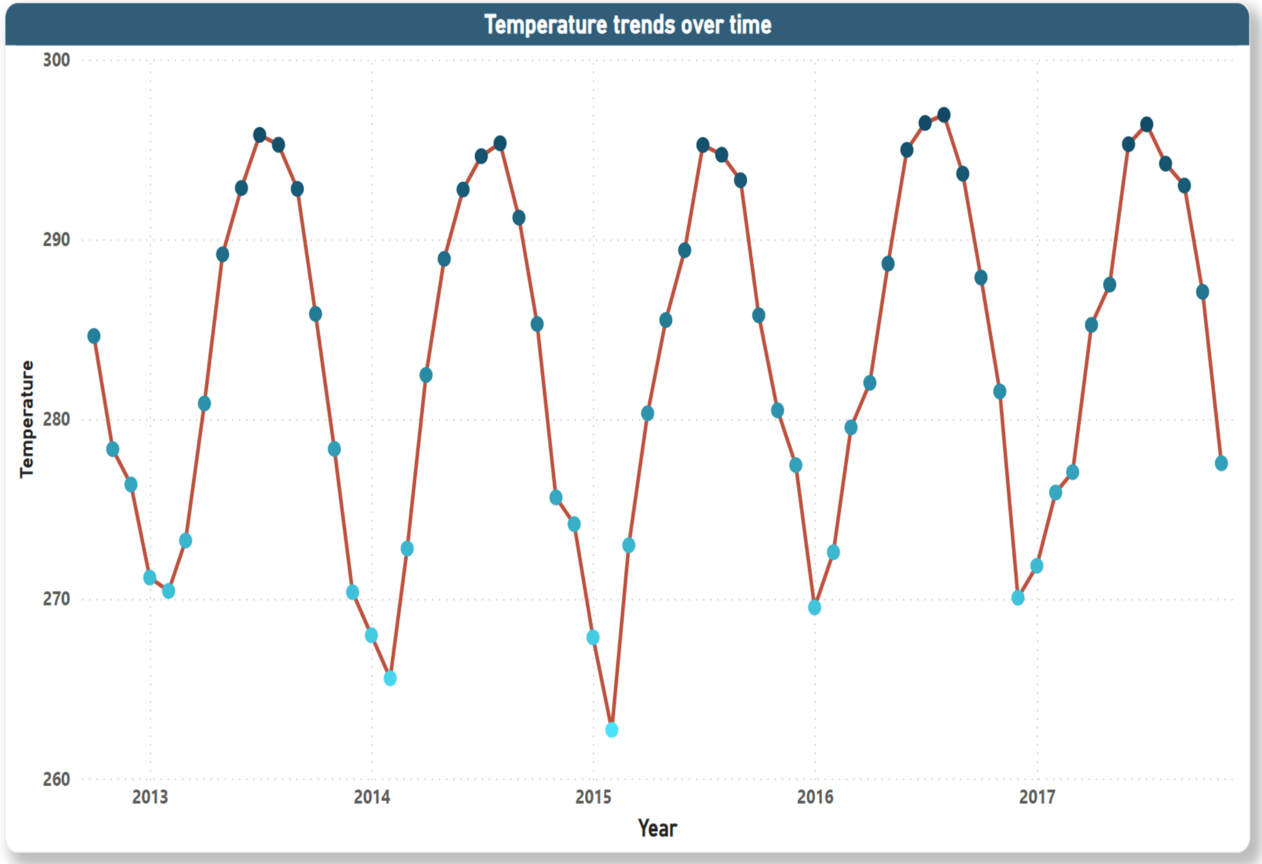


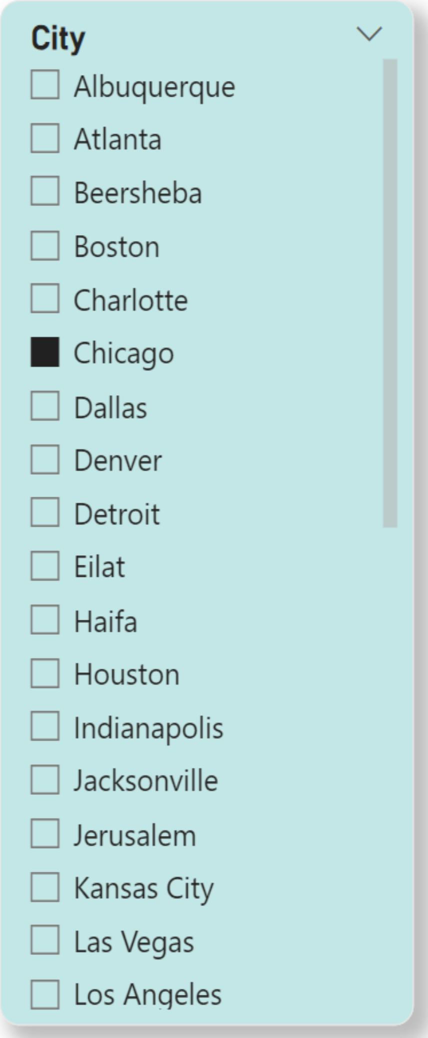
3. How does the distribution of cities in terms of latitude vary across different continents? Create a scatter plot in Power BI to illustrate this.



Created a scatter plot to show the distribution of cities in terms of latitude.

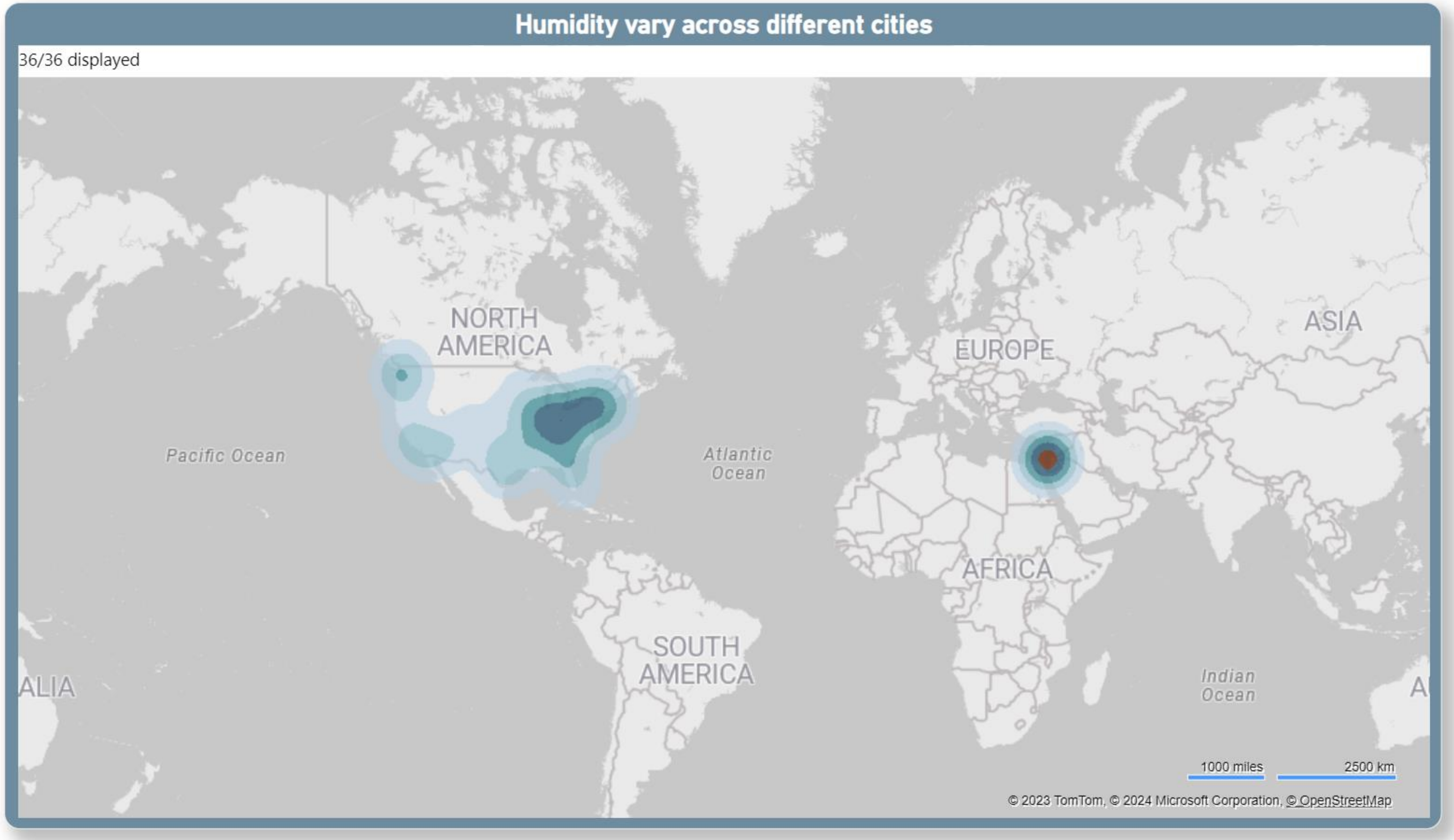
4. Create a line chart in Power BI to display the temperature trends over time for a selected city. Highlight extreme temperature events.





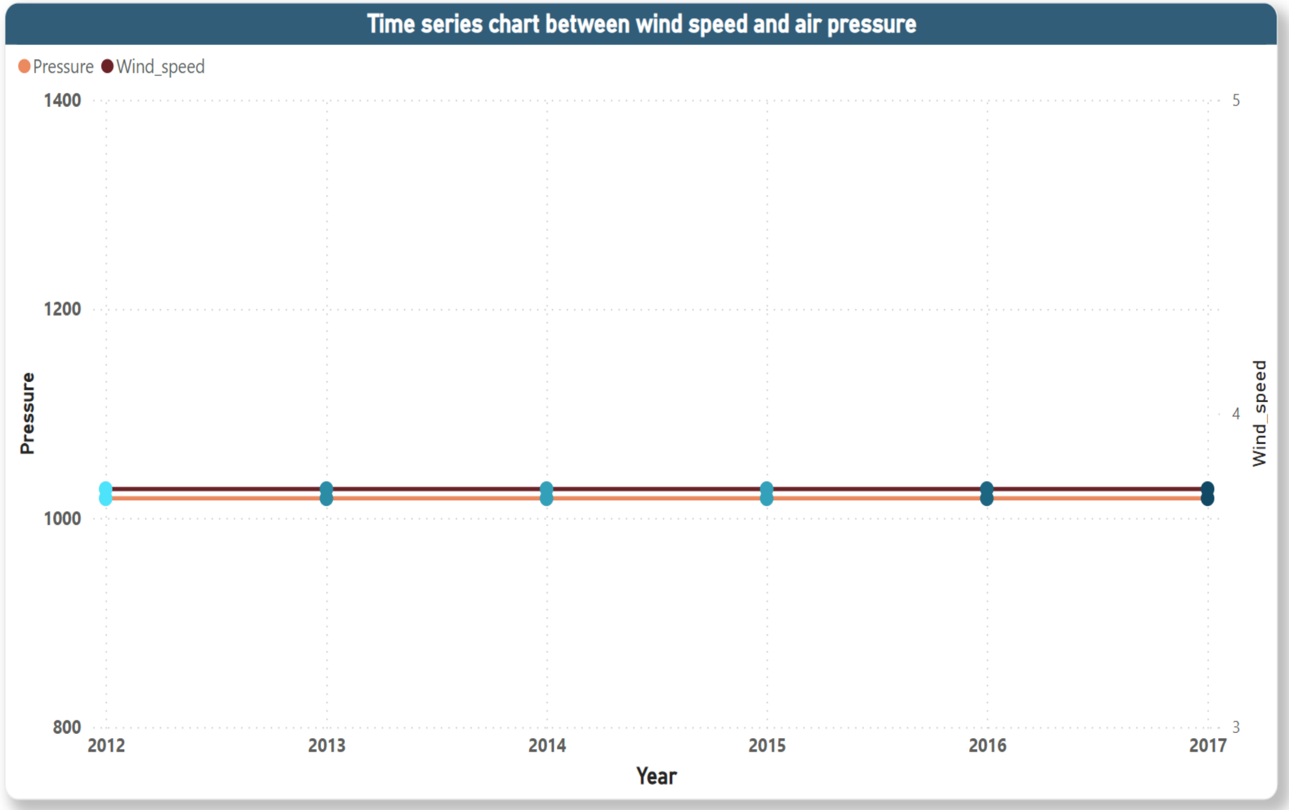
Created a line chart to show the temperature trends over time, highlighted highest to lowest points in the line as the temperature going. to select a specific city created a city slicer.

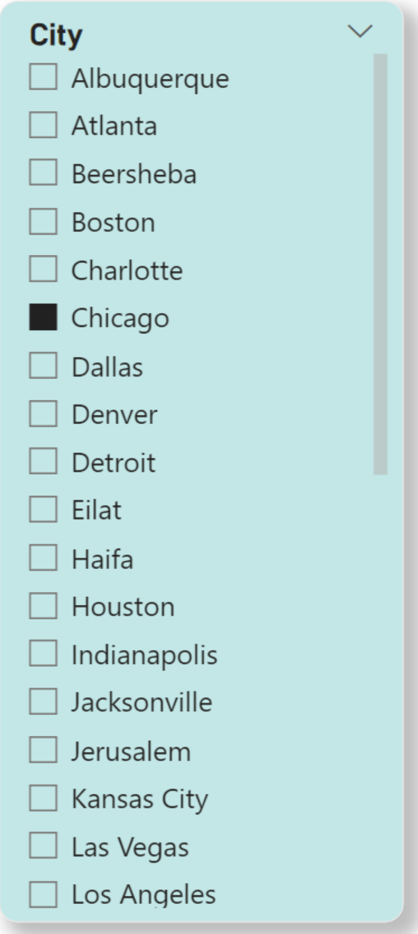
5. How does humidity vary across different cities? Generate a heatmap in Power BI to visualize this variation.



Created a heatmap to show the clusters of humidity across different countries. The deepest brown color showing the max humidity level, and light green to dark green showing here the lowest to 2nd highest humidity level.

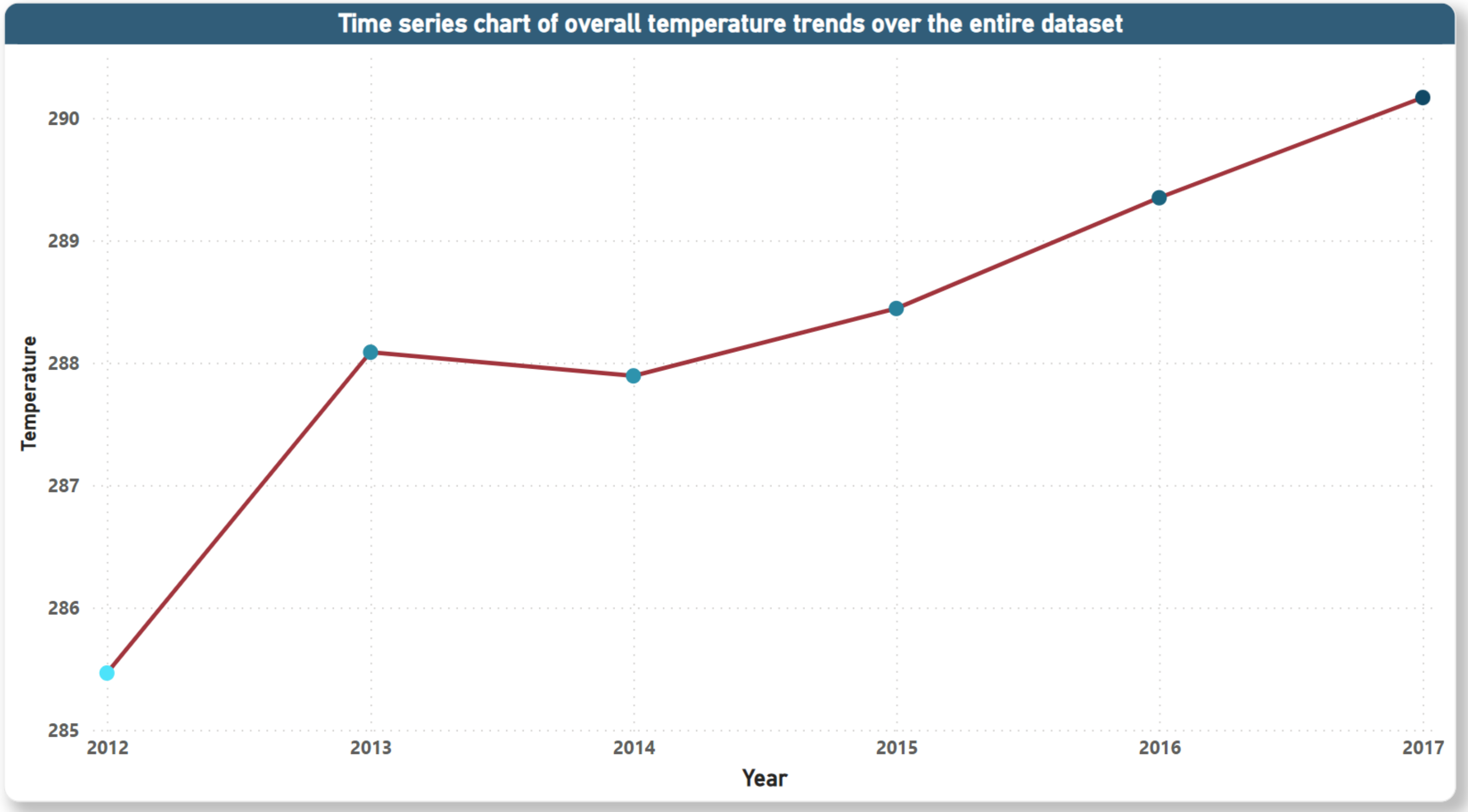
6. Can you create a time-series chart in Power BI showing the relationship between wind speed and air pressure for a specific city?





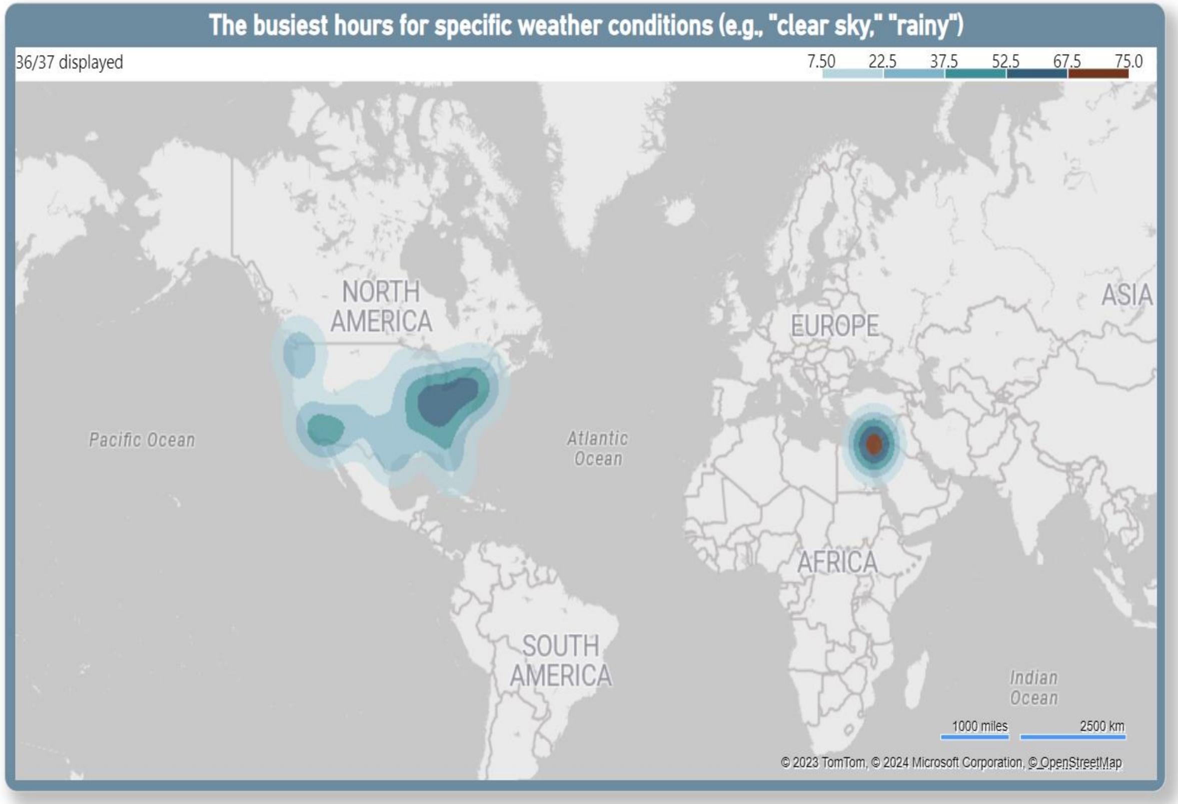
Created a time series line chart to show the relationship between wind speed and air pressure, to select a specific city created a city slicer.

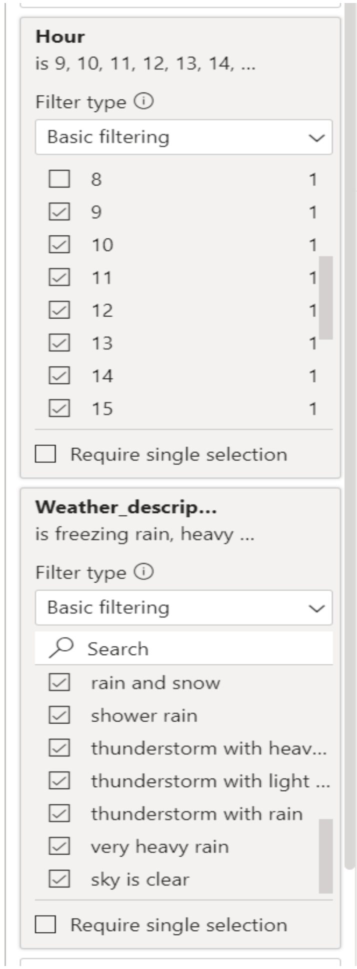
7. Create a time-series line chart in Power BI to show the overall temperature trends over the entire dataset.



Created a time-series line chart to show the overall temperature trends over the entire dataset, means it showing the trends for overall cities.

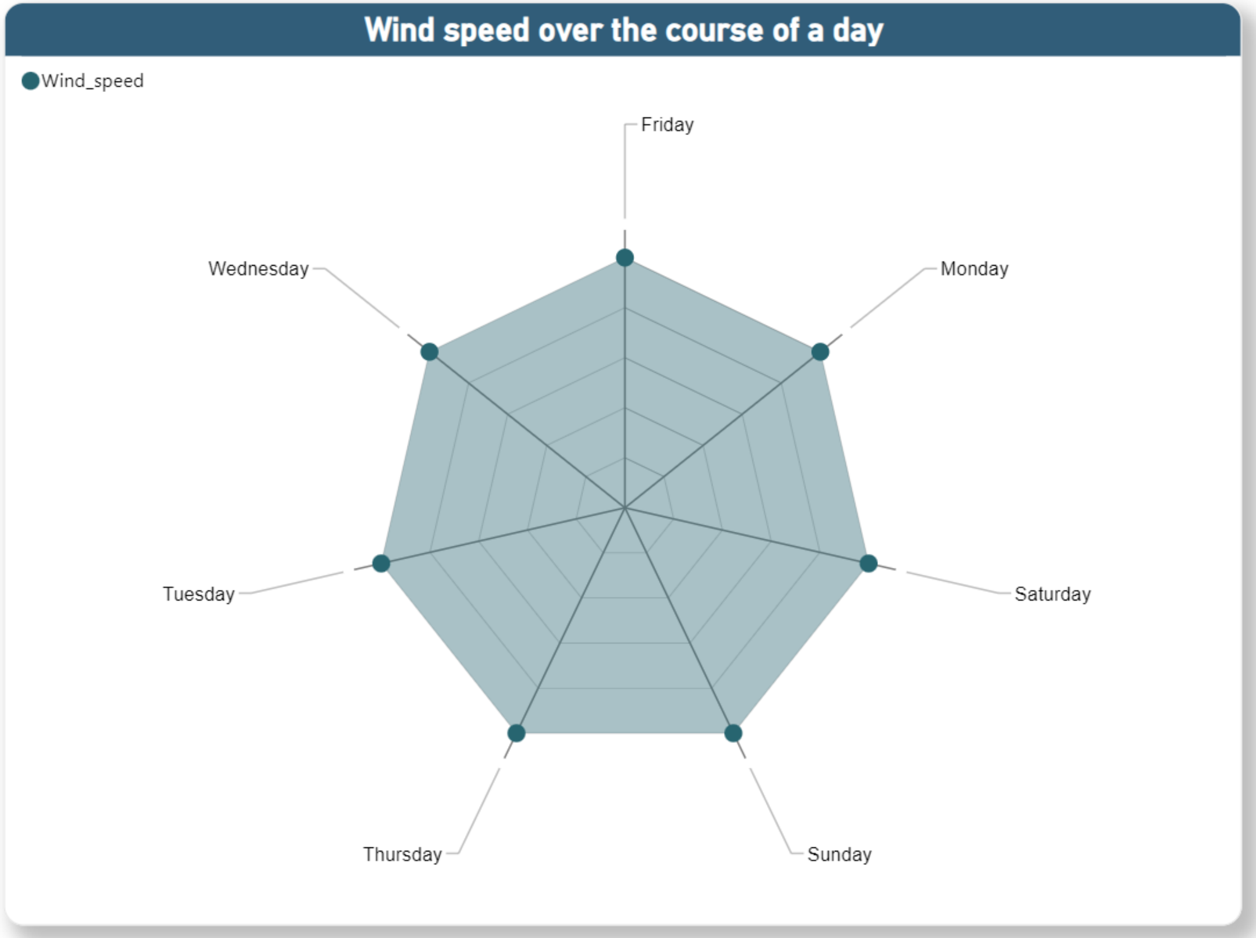
8. Can you create a heatmap in Power BI to visualize the busiest hours for specific weather conditions (e.g., "clear sky," "rainy")?





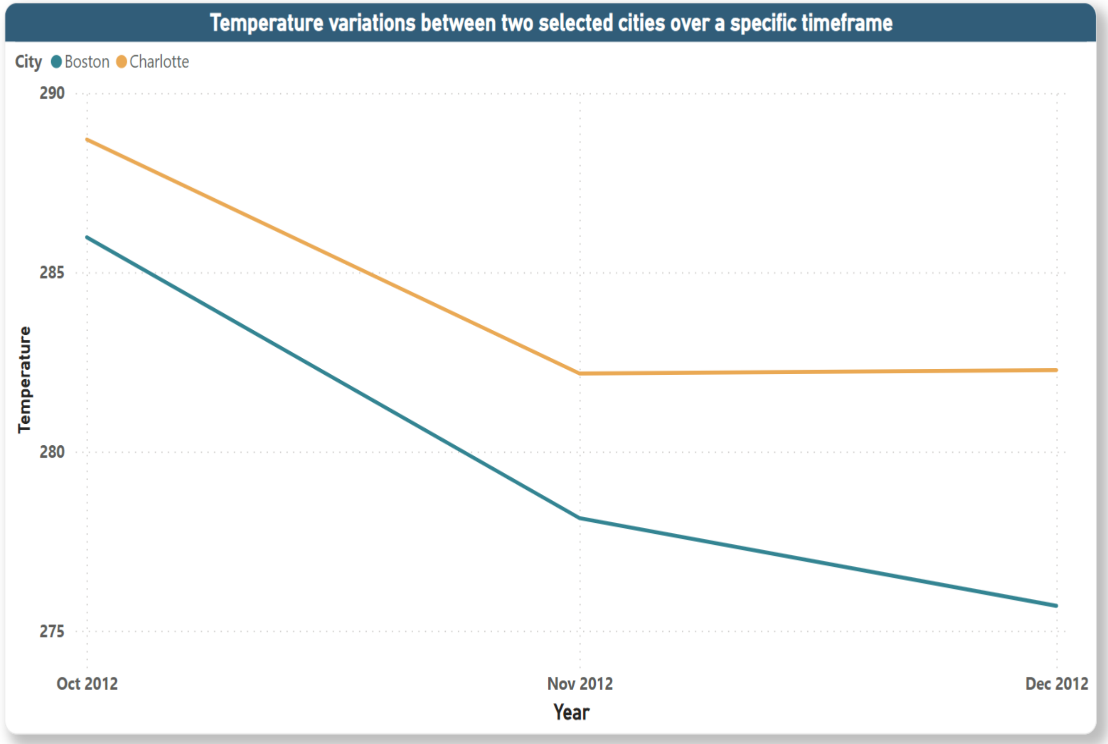
Created a heatmap to visualize the busiest hours (office hours 9 am-5 pm) for specific weather conditions (e.g., "clear sky," "rainy"). Filtered out the hours and those specific weather for a perfect match.

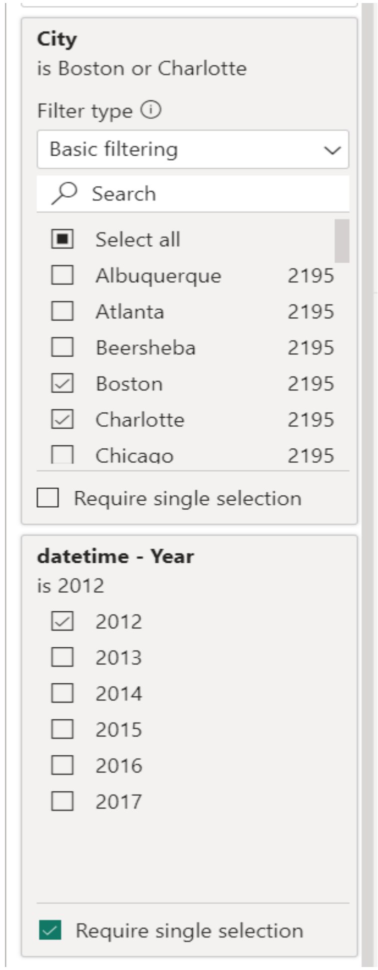
9. How does the wind speed change over the course of a day? Create a radial chart in Power BI to represent this.



Created a radial chart to represent the wind speed change over the course of a day, here I've taken the day name to showcase the course of the day.

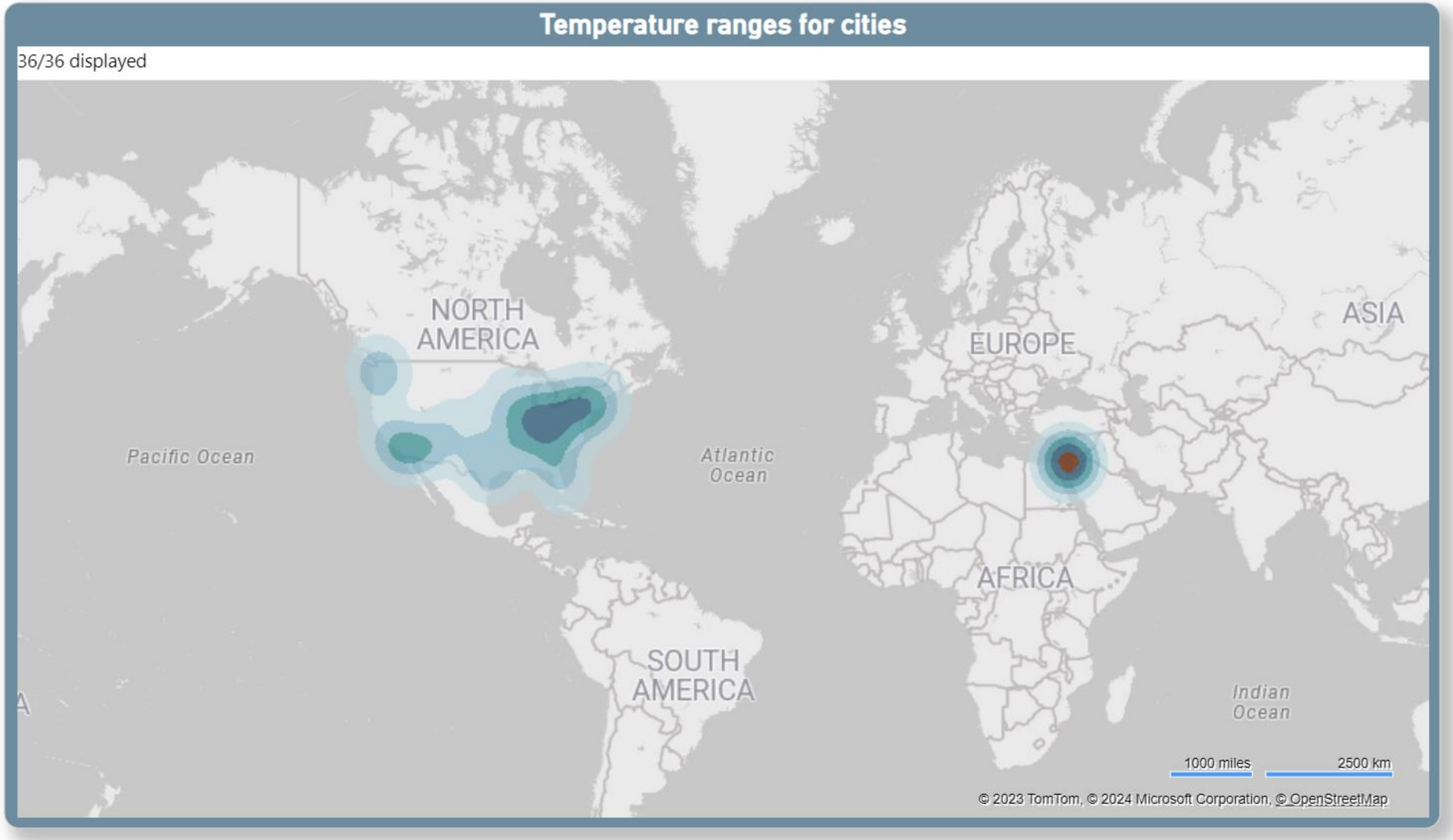
10. Create a Power BI chart comparing the temperature variations between two selected cities over a specific timeframe.





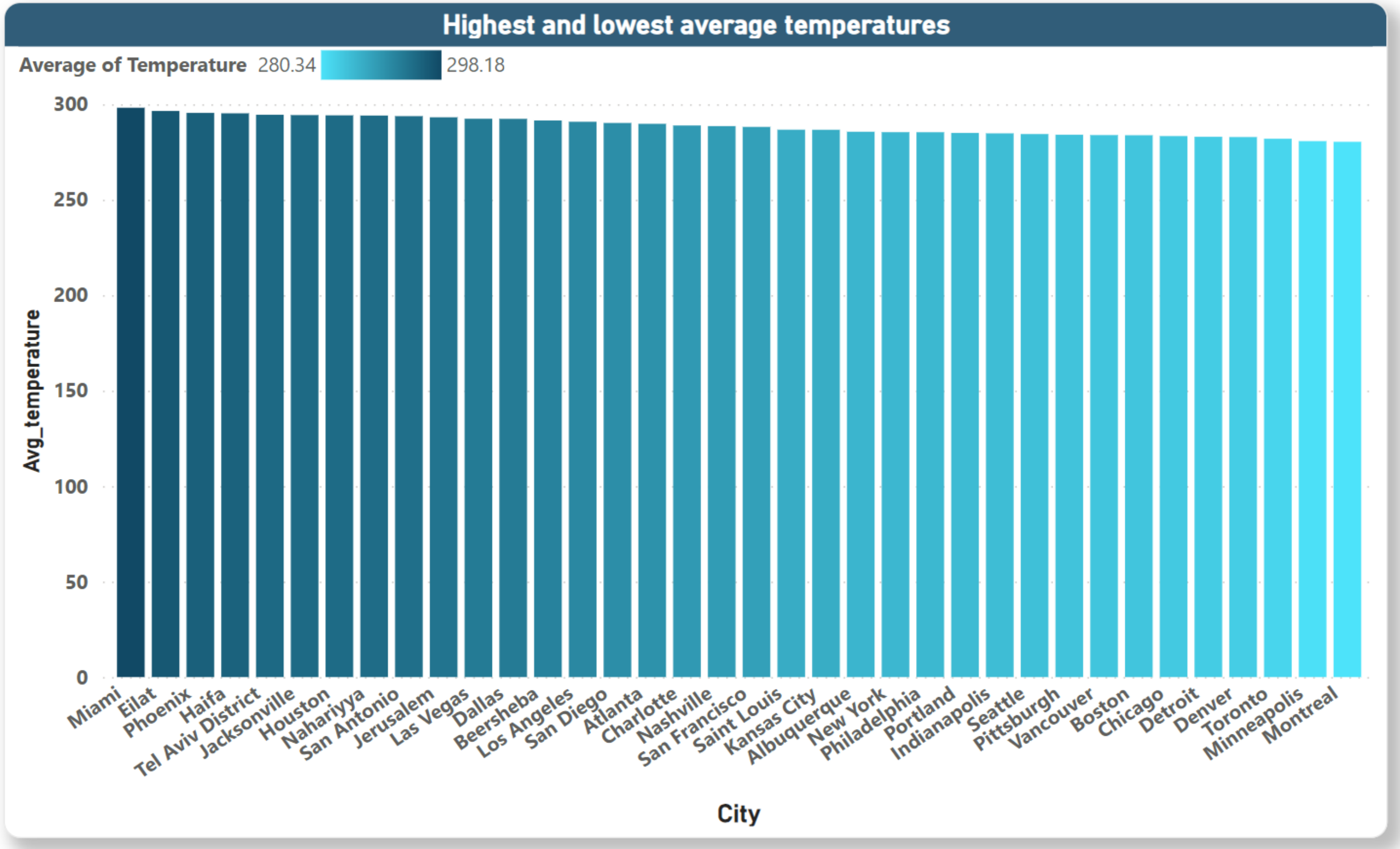
Created a line chart to show the temperature variations between 2 selected cities (Boston & Charlotte) over a specific timeframe (2012). Filtered out the cities and year to showcase the perfect match.

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



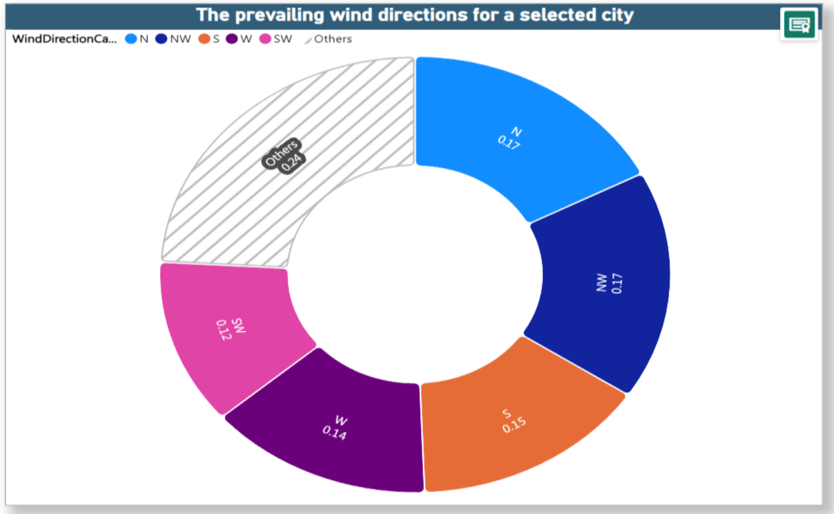
Created a heatmap to show the temperature ranges for cities across different countries. The deepest brown color showing the maximum temperature, and light green to dark green showing here the lowest to 2nd highest temperature.

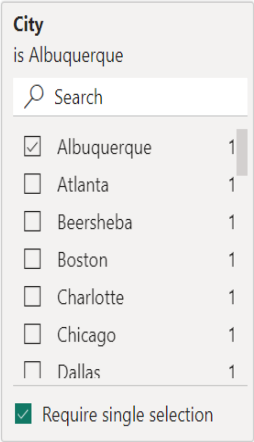
12. Create a bar chart in Power BI to highlight cities with the highest and lowest average temperatures in the dataset.

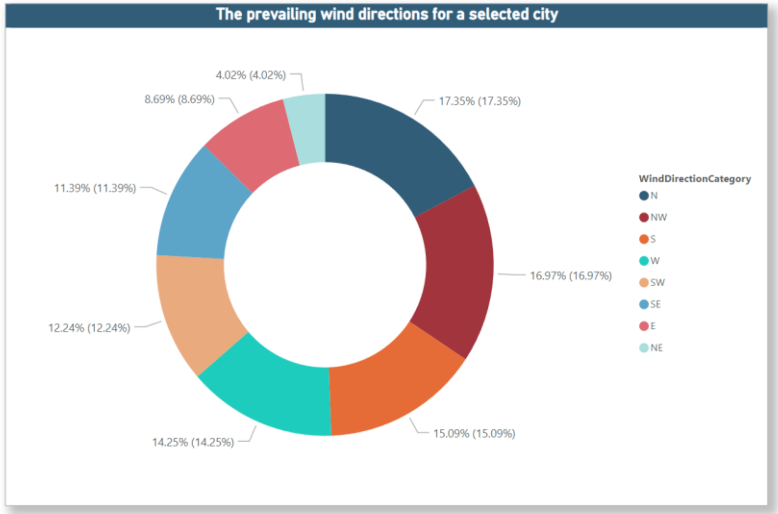


Created a bar chart to highlight the cities with the highest and lowest temperature in the dataset. The darkest green color referring the highest temperature & the lightest green color referring the lowest temperature.

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



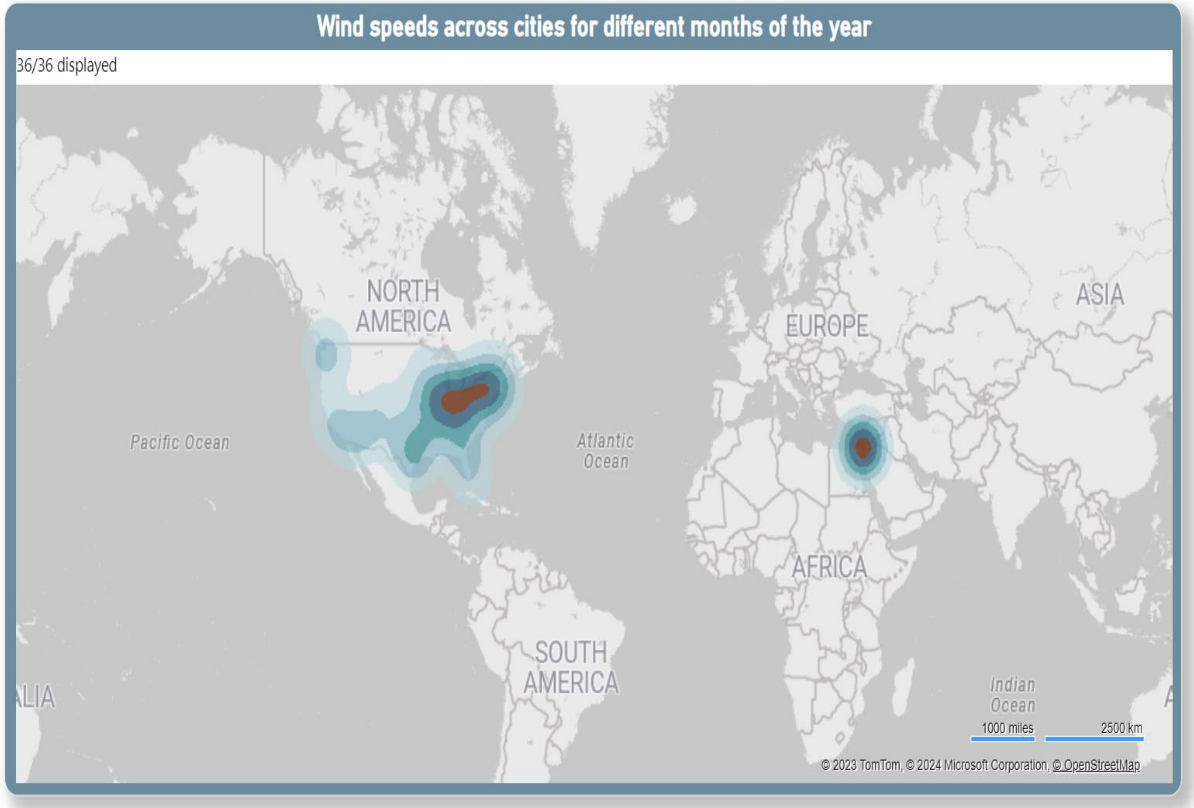




Unfortunately, wind rose chart was not available, so created a rose donut chart (left), also a donut chart (right) to visualize the prevailing wind directions for a selected city, Albuquerque. Filtered out the city in filter section.

Note: we can use any of this chart to visualize this condition.

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

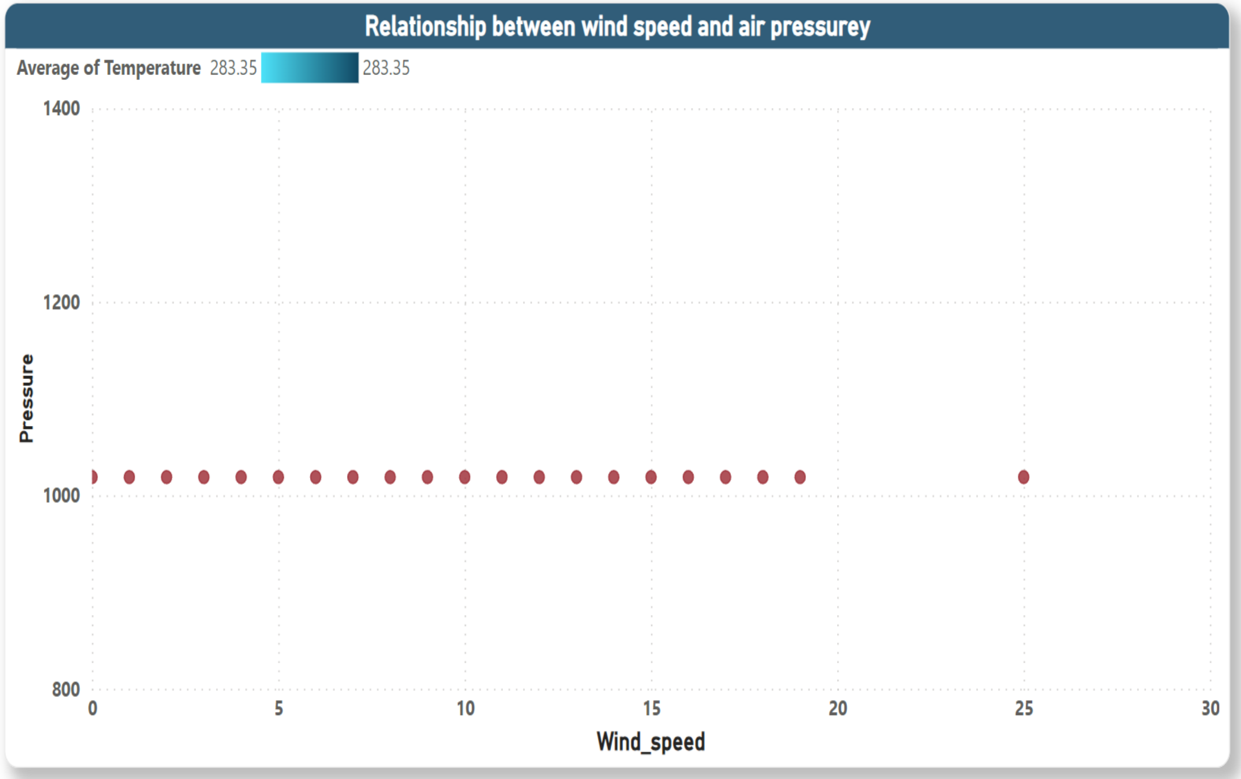


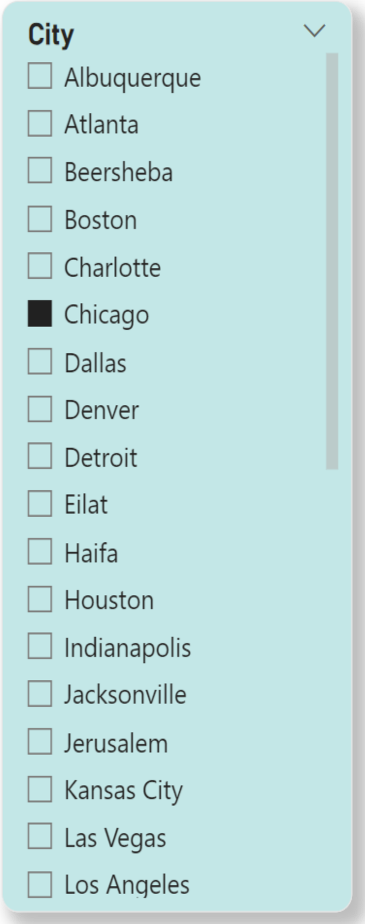
Month Name

- [ ] April
- [ ] August
- [ ] December
- [ ] February
- [x] January
- [ ] July
- [ ] June
- [ ] March
- [ ] May
- [ ] November
- [ ] October
- [ ] September

Created a heatmap to illustrating the average wind speed across cities for different months of the year. Created a slicer of Month name to show the single month (January). The deepest brown color showing the maximum wind speed, and light green to dark green showing here the lowest to 2nd highest wind speed.

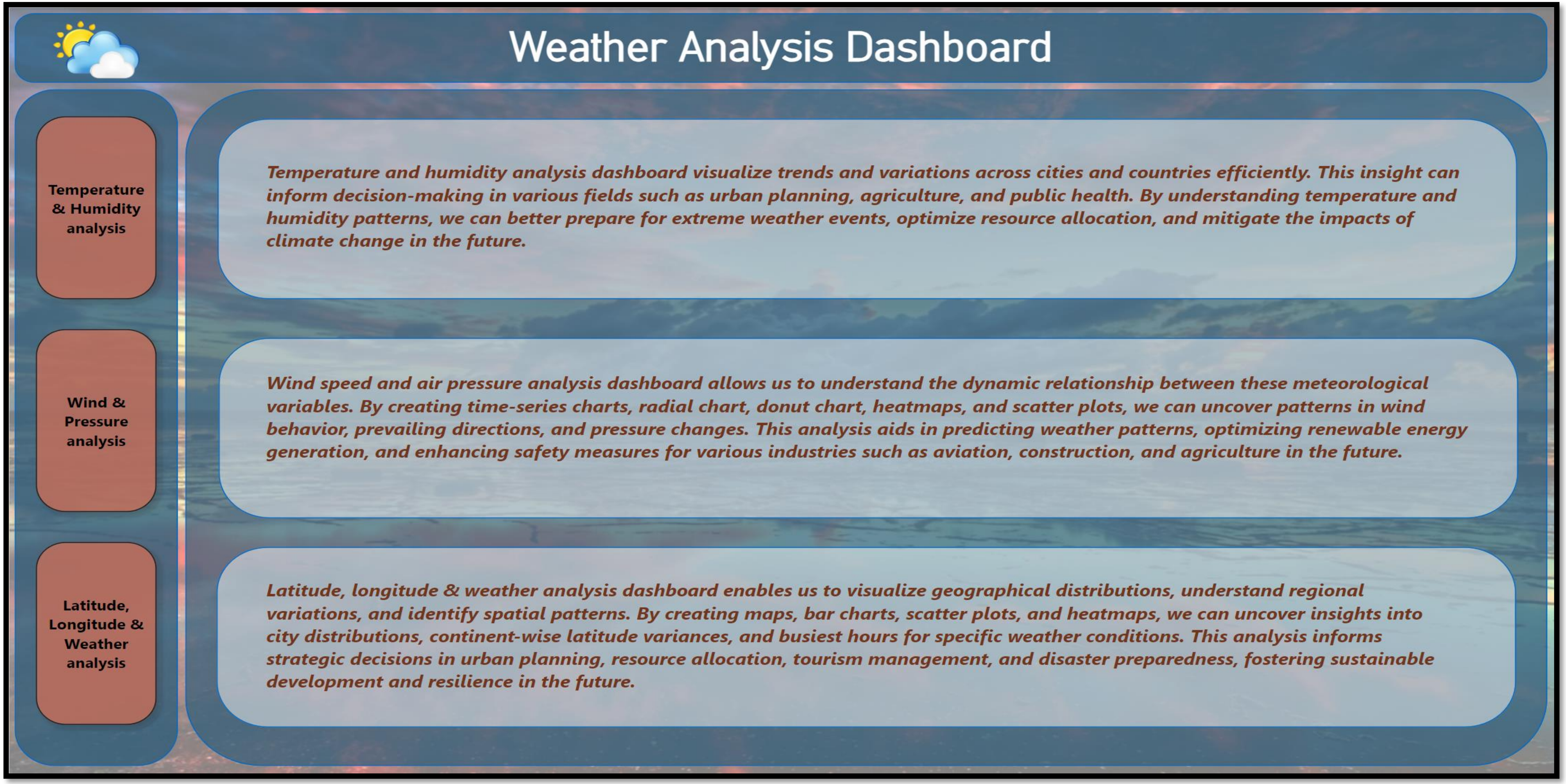
15. Create a Power BI scatter plot to show the relationship between wind speed and air pressure for a specific city.



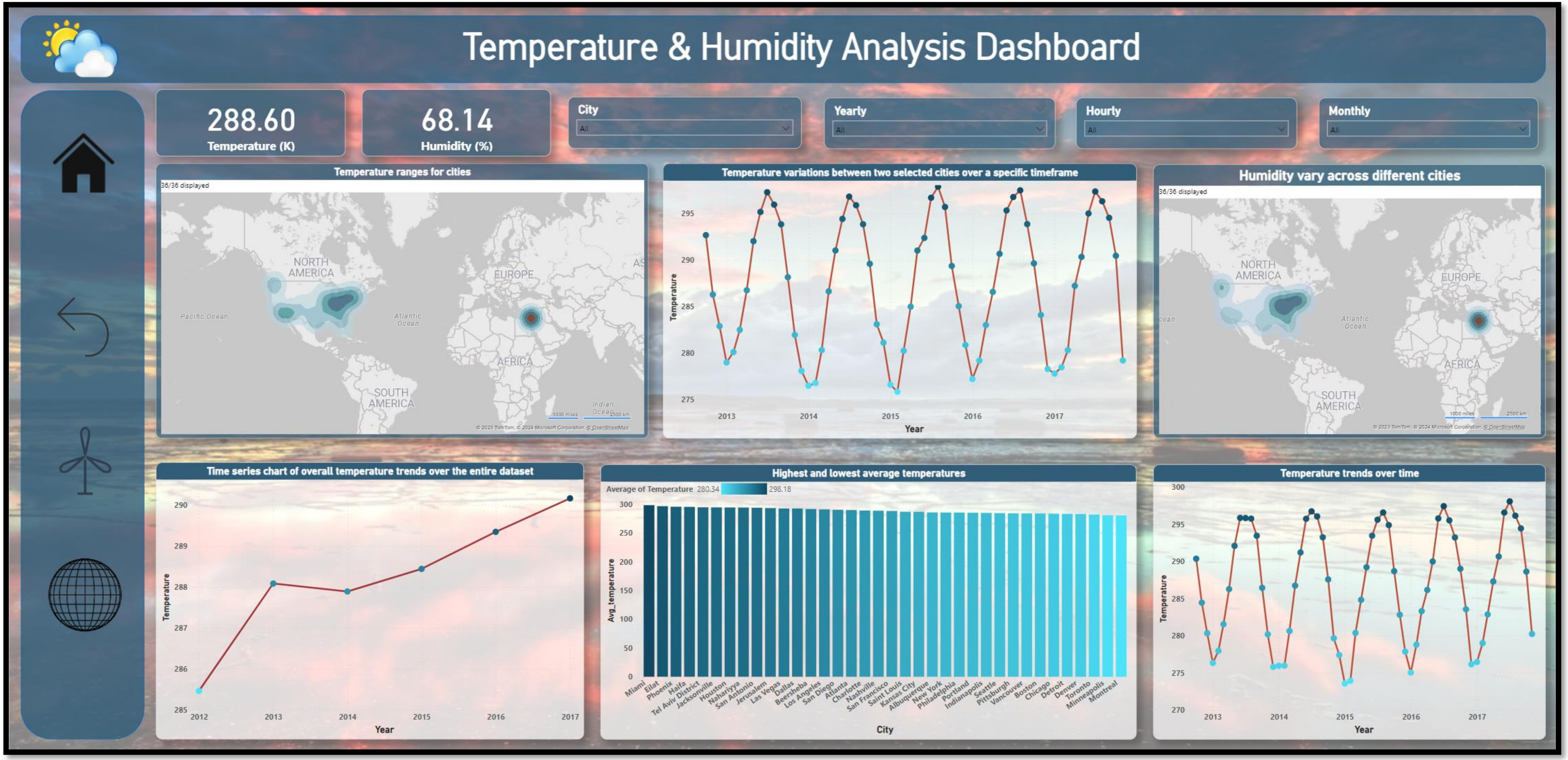


Created a scatter plot to show the relationship between wind speed and air pressure for a pressure for a specific city (Chicago). Here I've taken the air pressure and not summarized wind speed. Created a slicer to show a specific city.

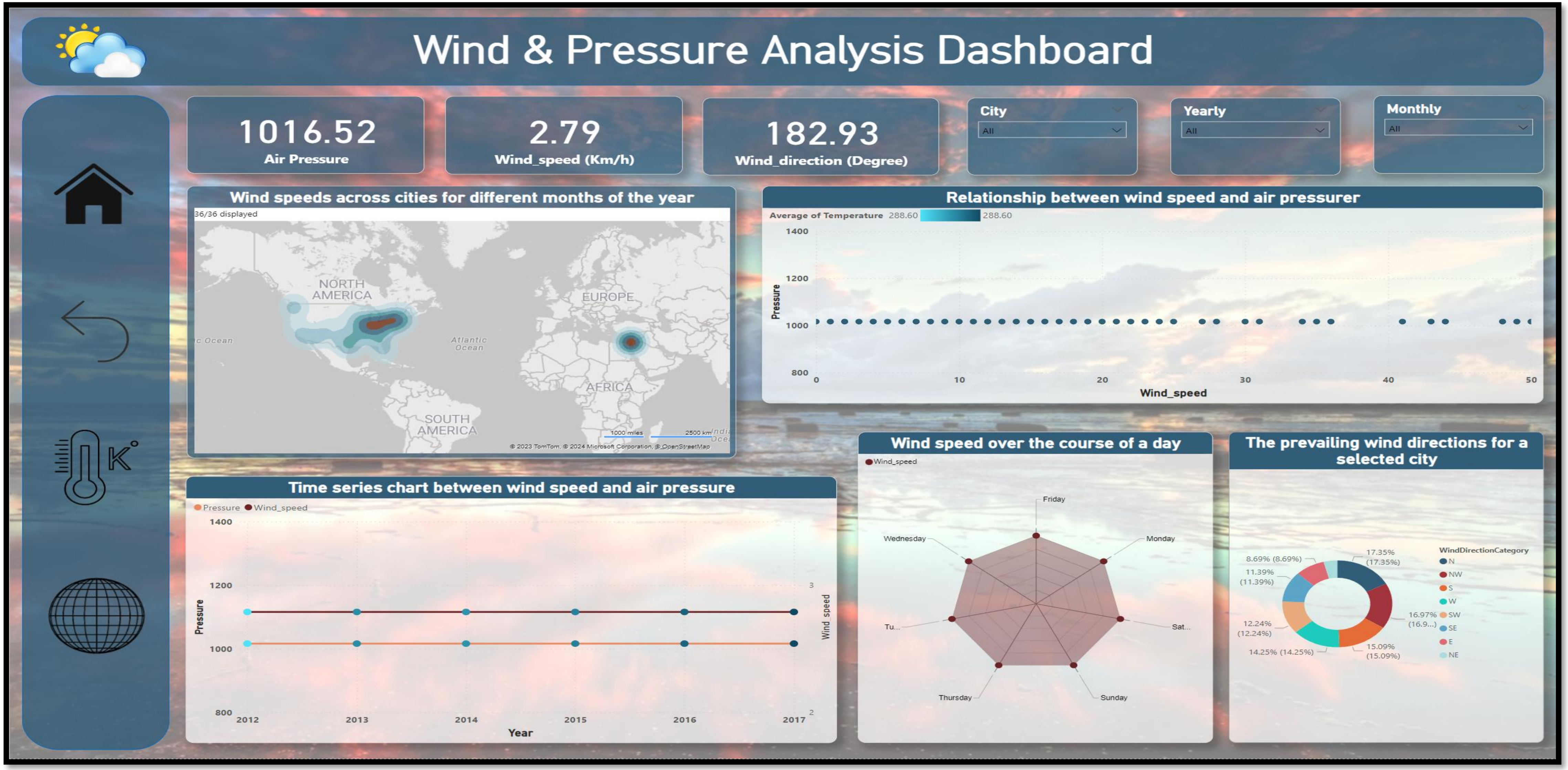
Weather Analysis Dashboard



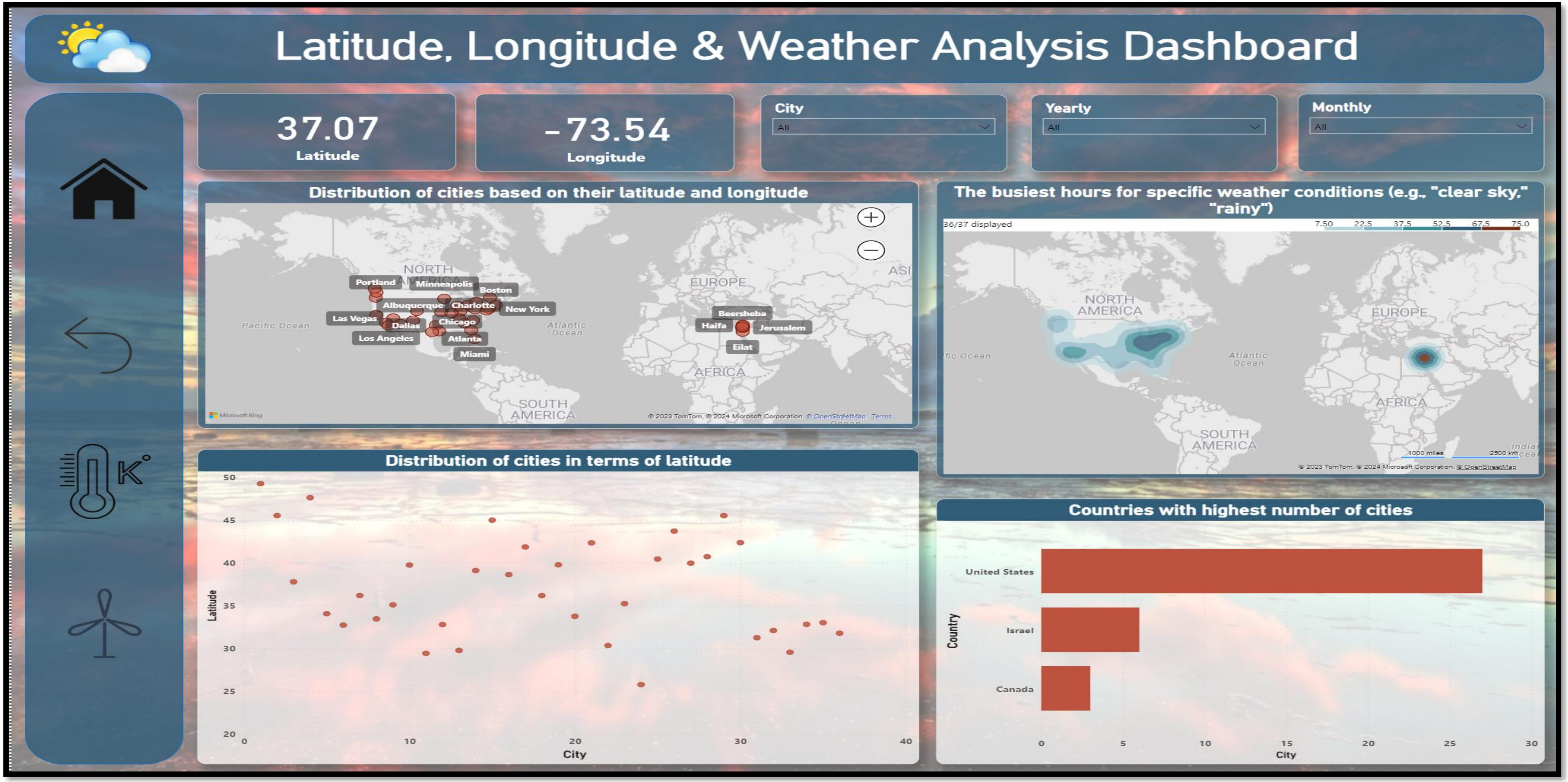
Temperature & Humidity Analysis Dashboard



Wind & Pressure Analysis Dashboard



Latitude, Longitude & Weather Analysis Dashboard



Understanding & Solving EDA questions

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

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

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

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

- ❑ A latitude above 60 degrees North is often considered to be an extreme northern latitude or Extreme maximum latitude. Some cities like, Iqaluit, Canada (63.7°N), Troms, Norway (69.6°N), Yakutsk, Russia (62.0°N), Fairbanks, Alaska (64.8° N) situated in extreme North hemisphere, These cities witnessing a polar night, extreme cold snaps, mostly snowfall, wind speeds can reach more than 40 km/h, more than avg -40 degree K in winter.
- ❑ A latitude below -60 degrees South is considered an extreme southern latitude or Extreme minimum latitude. Cities like, Ushuaia, Argentina (54.8°S), McMurdo Station, Antarctica (77.8°S) situated in extreme South hemisphere. Witness of wind speed 100 km/h, avg -28 degree c temperature through the year.

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

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

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

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

These cities has Latitude and Longitude within 1 degree of distance. Distance between 1-2 degree can be considered approximately similar Latitude and Longitude. So, we can say these cities has similar Latitude and Longitude values. To fetch these cities we can use any of these above SQL queries.

Noticeable facts –

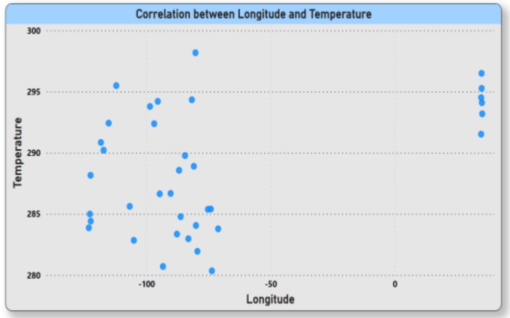
- Cities at similar latitudes often experience similar climate patterns, which can influence their development and infrastructure.
- Similar economic activities, such as agriculture, industry, or technology, may attract cities to specific regions.
- Proximity to major transportation routes, such as highways, railways, or airports, can influence the development and connectivity of cities.
- Cultural and political factors can play a role in the development and clustering of cities.

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

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

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

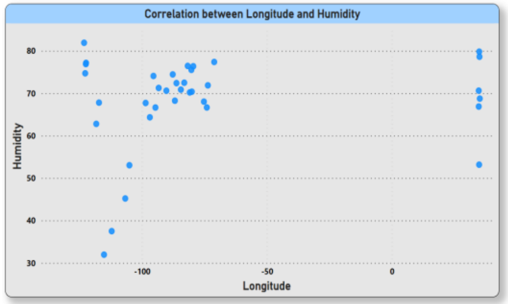
Correl_bw_longitude_temperature
0.20575749576366825




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

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

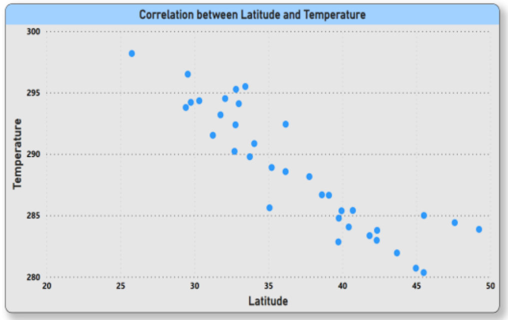
Correl_bw_longitude_humidity
0.06713517937490078




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

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

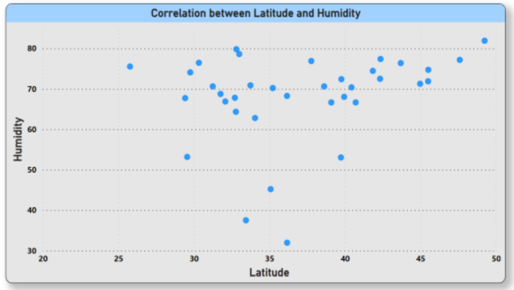
Correl_bw_latitude_temperature
-0.4301449468982176




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

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

Correl_bw_latitude_humidity
0.11692809006383399



Correlation b/w Latitude & Temperature	-0.4301449
Correlation b/w Longitude & Temperature	0.20575749
Correlation b/w Latitude & Humidity	0.11692809
Correlation b/w Longitude & Humidity	0.067135179

- (-0.4301449) indicates a moderate negative correlation b/w Latitude & Temperature.
- (0.20575749) indicates a weak positive correlation b/w Longitude & Temperature.
- (0.11692809) indicates a very weak positive correlation b/w Latitude & Humidity.
- (0.067135179) also indicates a very weak positive correlation b/w Longitude & Humidity.

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

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

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

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

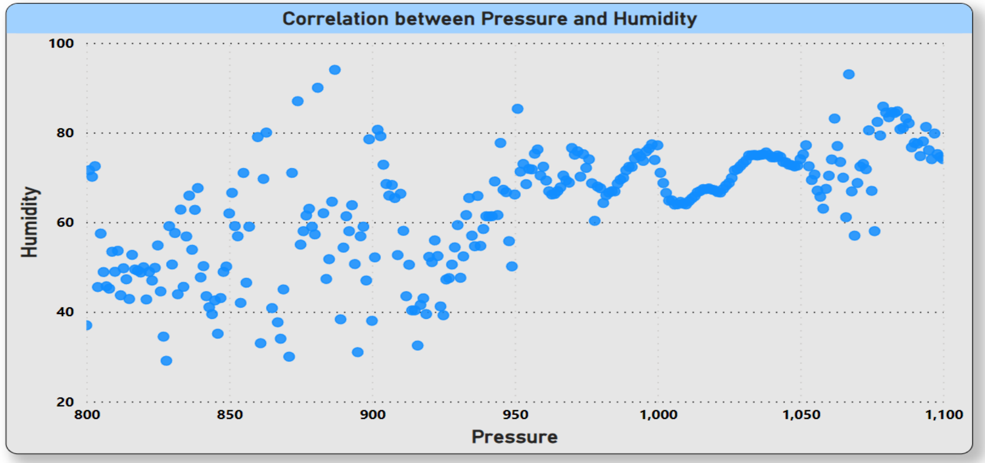
```
with cte as
( select cl.City, ff.weather_description, count(*) rainy_days
from city_lookup cl
join final_fact ff
on cl.City_id = ff.City_id
where ff.weather_description like "%rain%"
group by 1,2)
select cte.City, cte.weather_description from cte
order by rainy_days desc
```

Toronto, Montreal & Vancouver are the top 3 cities with the most frequent occurrence of rainy weather based on weather description.

- Spring (March-May) and Fall (September-November) average around 75-85mm of rain per month in Toronto.
- Spring (March-May) sees the most rain, averaging around 75-90mm per month in Montreal.
- Winter (December-February) brings the most rain, averaging around 150-200mm per month in Vancouver.

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

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




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

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

Correl_bw_pressure_humidity
0.0771615214995301 6

A correlation of 0.077 between humidity levels and air pressure is a relatively weak positive relationship.

- Weak correlation means humidity level and air pressure do not strongly drive each other. Other factors have a bigger direct impact on weather.
- With higher pressure, subsiding air warms and loses relative humidity. Lower pressure allows rising, cooling air to retain more moisture.
- Humidity and pressure may correlate more strongly on short time scales, but other dynamics dominate over long periods.
- Higher pressure typically brings more settled, stable weather. Lower pressure indicates storm systems are more likely.

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

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

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

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

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

Onshore winds (blowing from the ocean towards the land):

- Typically bring cooler air over land, especially during summer. This is because ocean water generally has a lower temperature than the land during the warm season.
- Can bring warmer air during winter if the ocean water is warmer than the land.

Offshore winds (blowing from the land towards the ocean):

- Typically bring warmer air over land, especially during summer. This is because the land heats up faster than the ocean during the warm season.
- Can bring cooler air during winter if the land is colder than the ocean water.

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

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

City	month	avg_temp	max_temp	max_temp_fluctuation
Minneapolis	January	264.1889103	295.3295214	31.14061107
Minneapolis	February	265.1646816	295.3295214	30.16483977
Montreal	February	264.8355044	294.3005872	29.46508282
Montreal	January	265.6189056	294.3005872	28.68168168
Minneapolis	December	267.2912517	295.3295214	28.03826963
Detroit	February	268.8679292	295.6821602	26.81423106
Las Vegas	December	279.7797912	306.2685698	26.48877857
Toronto	February	268.1386153	294.5796928	26.44107754

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

City	month	avg_temp	min_temp	min_temp_fluctuation
Minneapolis	July	295.3295214	264.1889103	31.14061107
Minneapolis	August	293.9916383	264.1889103	29.80272804
Montreal	July	294.3005872	264.8355044	29.46508282
Minneapolis	June	293.3294662	264.1889103	29.14055591
Montreal	August	293.0499331	264.8355044	28.21442863
Detroit	July	295.6821602	268.8679292	26.81423106
Minneapolis	September	290.9390717	264.1889103	26.75016145
Las Vegas	July	306.2685698	279.7797912	26.48877857

From this query, I'm getting the average temperature of per month in last 5 years, and also what is the maximum average temperature and the minimum average temperature. By this, we can understand what can be the temperature fluctuations in each month. Cities can experience significant temperature fluctuations due to various factors, including:

- **Seasonal changes**
- **Continental vs. maritime climates**
- **Urban heat island effect**
- **Synoptic weather patterns**
- **Local factors**

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

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

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

Top Periods	Count of weather_description
2017	3238
2016	2382
2015	2033
2014	1087
2013	1975
2012	153

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

As we can see the period of 2017 was the year of extreme weather events, such as storms or heatwaves, by analyzing the time-based data , also in table format I've shown the weather patterns emerge in those periods.

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

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

But, there are many significant differences in temperature trends between northern and southern hemisphere cities over the year, due to the Earth's tilt on its axis and its revolution around the sun.

Northern Hemisphere:

Summer	June-August
Autumn	September-November
Winter	December-February
Spring	March-May

Longer days and more direct sunlight lead to warmer temperatures. The higher the latitude, the shorter the days and cooler the temperatures become.

Southern Hemisphere:

Summer	December-February
Autumn	March-May
Winter	June-August
Spring	September-November

Opposite to the north, this is the hottest time with longer days and more direct sunlight. Higher latitudes are cooler.

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

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

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

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

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

Consequences of Extreme Cold:

- Extreme cold poses health risks including hypothermia, frostbite, and heart attacks, while also causing infrastructure challenges like frozen pipes and transportation disruptions. Additionally, it leads to economic losses, energy consumption spikes, and exacerbates social and mental health issues.

Adaptation Strategies to Extreme Cold:

- Residents prepare with layered clothing and snowmobiles, but prolonged cold strains infrastructure and health. Adequate insulation, snow removal, and emergency kits are vital for safety and comfort. Cities prioritize vulnerable populations' access to shelter and healthcare during extreme events.

Consequences of Extreme Heat:

- Extreme heat poses health risks like heat-related illnesses and strains energy resources, leading to power outages and infrastructure damage. Prolonged heatwaves affect crops, water supply, and urban areas, impacting productivity and increasing costs.

Adaptation Strategies to Extreme Heat:

- Cities establish cooling centers for relief from high temperatures and promote water conservation during heatwaves. Increasing shade and green spaces in urban areas help cool environments, while public awareness campaigns educate on heat risks and safety measures. Additionally, ensuring vulnerable populations access shelter, cooling centers, and healthcare during extreme events is prioritized.

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

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

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

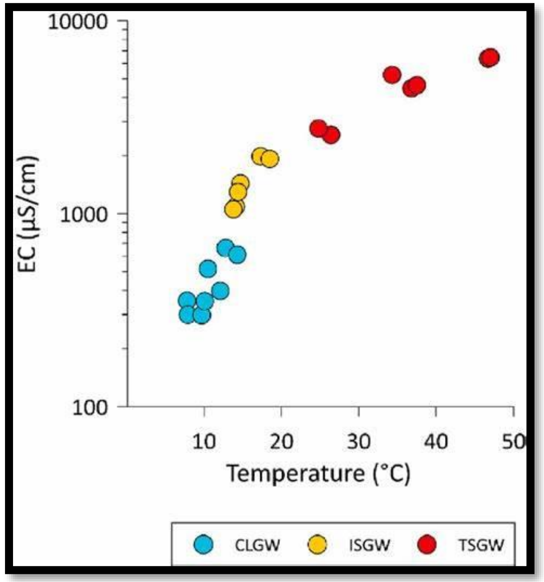
Weather_events
broken clouds
few clouds
fog
haze
heavy shower snow
heavy snow
light intensity drizzle
light rain
light shower snow
light snow
mist
moderate rain
overcast clouds
proximity shower rain
scattered clouds
sky is clear
snow
very heavy rain

City	Average of Unusual_temp	Count of weather_description
Tel Aviv District	320.93	1
Haifa	320.375	2
Phoenix	320.3535714	14
Eilat	320.15	2
Nashville	256.8724722	12
Charlotte	256.5072	5
Albuquerque	256.163	4
Saint Louis	255.9150594	101
Philadelphia	255.601952	59
Chicago	255.5147911	596
Kansas City	255.4126315	168
Boston	255.3898604	61
Pittsburgh	255.353999	200
Toronto	255.2935548	441
Denver	255.2381346	389
New York	255.2303855	72
Indianapolis	254.7733258	317
Detroit	254.4457814	443
Montreal	254.1798454	1500
Minneapolis	253.2845196	2260
Vancouver	249.8021122	4

The comfortable or ideal temperature for a city around the year between 275-295 degree K. In this 12 cities from Weather_analysis dataset has the average anomalies temperature, which is less than 275K and more than 295K.
And, these are the noticeable certain weather events occur in these temperature

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

Unfortunately, I don't have any real time data about the impact of temperature on energy consumption patterns in cities.
Though, Let's try to analyze the impact of temperature on energy consumption patterns in cities.



Temperature & City Energy Use:

•**Summer:** Higher temps -> More AC -> Higher electricity use (warmer climates, denser cities hit harder).
•**Winter:** Colder temps -> More heating -> Higher consumption of gas/electricity/fuel (varies by infrastructure).
•**Mild seasons:** Less heating/cooling -> Lower overall energy use (potentially).
•**Building efficiency:** Better insulation/systems = less impact from temp changes.
•**Urban design:** Walkable, mixed-use cities encourage walking/cycling -> less transport energy use.
•**Socioeconomic factors:** Low-income communities more vulnerable to energy price fluctuations.
•**Renewable energy:** More renewables = less susceptibility to price hikes during peak demand.
•**Trends:** Strong correlation between temp & electricity use, expected to increase with climate change.
•**Mitigation strategies:** Cities promoting energy efficiency, renewable energy, & smart grids.

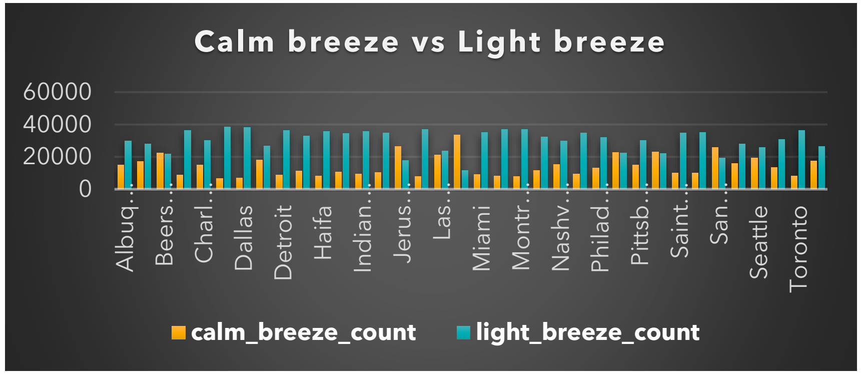
13. How do specific wind patterns impact air quality and pollution dispersion in urban areas? Analyze wind direction data for insights.

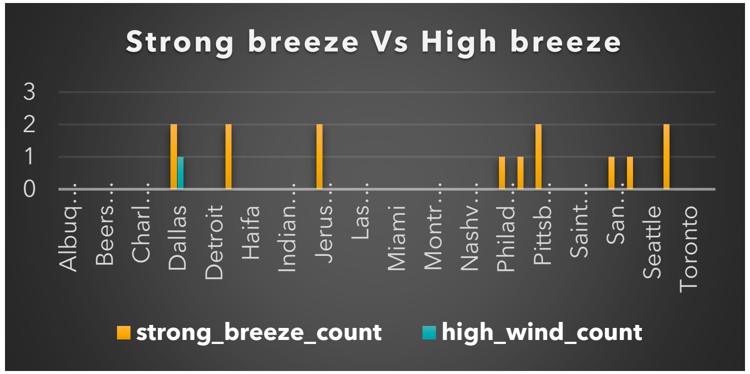
a) Calm breeze: 0-1.8 km/h (or 0-1 knots)
b) Light breeze: 1.9-11.2 km/h (or 2-6 knots)
c) Moderate breeze: 11.3-28.8 km/h (or 7-16 knots)
d) Strong breeze: 28.9-49.9 km/h (or 17-27 knots)
e) High wind: 50 km/h or higher (or 28 knots or higher)

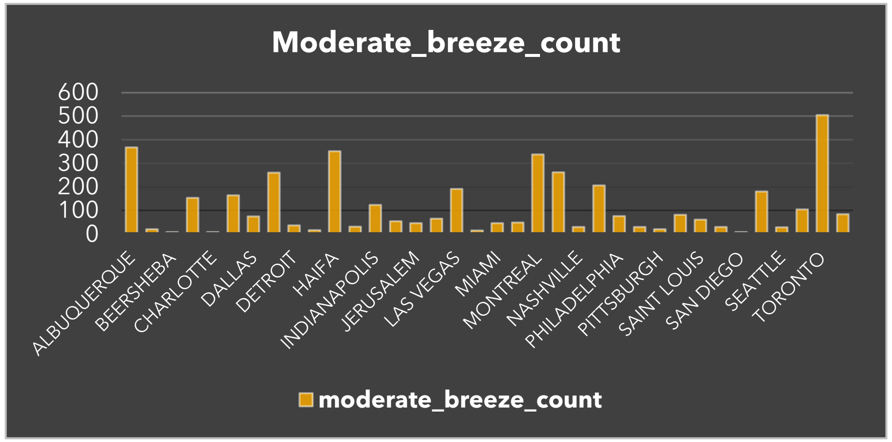
Note: *Well, I will prefer to analyze the wind_speed to get the insights for air quality & pollution instead of Wind_direction*

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

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







- The higher wind speed go, result in very low pollution levels and excellent air quality. Exposure to harmful pollutants is minimal, leading to improved respiratory health and reduced health risks for the population.
- The lighter wind speed go, it helps disperse pollutants and improve air quality by enhancing ventilation and mixing of the atmosphere. This can reduce local pollution concentrations and mitigate the effects of emissions from traffic, industry, and other sources.

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

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

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

Natural disaster
heavy intensity drizzle
heavy intensity rain
heavy intensity shower rain
heavy shower snow
heavy snow
heavy thunderstorm
proximity sand/dust whirls
proximity thunderstorm
proximity thunderstorm with drizzle
proximity thunderstorm with rain
ragged thunderstorm
rain and snow
sand
sand/dust whirls
shower snow
thunderstorm
thunderstorm with drizzle
thunderstorm with heavy drizzle
thunderstorm with heavy rain
thunderstorm with rain
tornado
very heavy rain
volcanic ash

In all 36 cities are prone area. For these natural disaster, cities are facing so many life challenges and transportation issues.

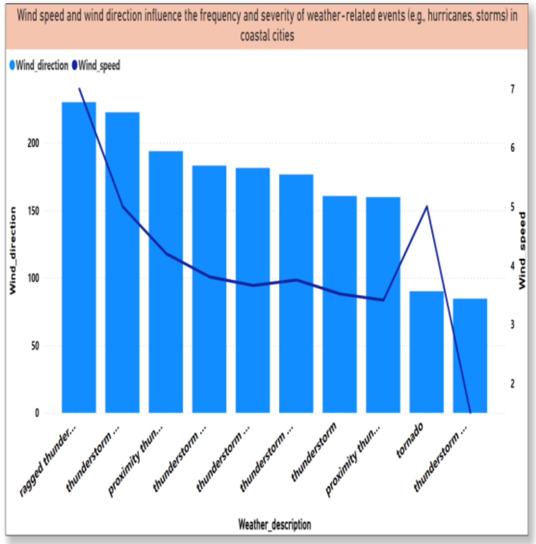
- Strong winds can cause structural damage to buildings, bridges, power lines, and other infrastructure. This can lead to power outages, disruptions in communication, and even building collapses.
- Strong winds can increase the severity of tropical cyclones like hurricanes for coastal cities in their path. This brings high winds, storm surge and flooding.
- Blizzards and whiteout conditions become more likely in areas prone to cold, icy winds.
- Air travel can be disrupted by takeoff and landing restrictions due to crosswinds. Flight delays and cancellations may increase. Also, Takeoff and landing of small aircraft may be prohibited during strong winds due to safety issues.

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

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

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

City id	City	Wind_direction	Wind_speed	weather_description	Month_Name
1	Vancouver	80	7	thunderstorm with rain	August
1	Vancouver	180	4	thunderstorm with light rain	April
1	Vancouver	80	4	proximity thunderstorm	July
1	Vancouver	0	1	proximity thunderstorm	September
1	Vancouver	170	8	thunderstorm with heavy rain	September
1	Vancouver	330	1	thunderstorm	September
1	Vancouver	173	1	thunderstorm with light rain	April



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

- Miami has maximum variation in the change of climate and Eilat has minimum variation in the change of climate.
- The mean Wind direction of these cities lying between 149-218 degree and mean Windspeed lying between 3.14-4.40 km/h
- The Standard_deviation of Wind_direction lying between 80.92-143.93 degree and the Standard_deviation of Wind_speed lying between 1.80-3.01km/h
- We get the maximum and minimum wind_direction and wind_speed of each coastal cities.

Conclusion my understanding about this data

- The weather analysis conducted provides a comprehensive overview of historical climate data, enabling stakeholders to gain valuable insights into past trends and patterns.
- Through meticulous examination, key metrics such as temperature variations, precipitation levels, and occurrences of extreme weather events have been identified and analyzed.
- The identification of seasonal fluctuations and temperature anomalies offers critical information for forecasting future weather conditions and anticipating potential risks.
- Advanced analytical techniques have been employed to uncover correlations between different meteorological parameters, enhancing our understanding of weather dynamics.
- This analysis serves as a valuable resource for decision-makers across various sectors, allowing them to make informed choices and develop strategies to adapt to changing environmental circumstances.
- Moving forward, continuous refinement of methodologies and the utilization of cutting-edge data analysis tools will be essential for further enhancing our understanding of weather variability and improving resilience in the face of uncertainty.