INDIAN WEATHER DATA ANALYSIS

Exploring Real-Time Weather Data to Unveil Patterns and Trends Across Indian Cities.

Project Overview

Objective:

- Explore real-time weather data of Indian cities.
- Identify patterns, correlations, and visualize weather conditions.

Tools:

- Jupyter Notebook for analysis and visualization.

Libraries:

- **Pandas:** Data manipulation and exploration.
- **Seaborn and Matplotlib:** Data visualization.
- **Plotly:** Interactive charts and graphs.

Data Cleaning & Preprocessing

Column Overview:

- Checked for missing values.
- Handled nulls using appropriate strategies (mean, median, mode imputation).
- Converted data types where necessary.

Key Preprocessing Steps:

- Removed irrelevant or redundant columns.
- Standardized units and formatted data for consistency.

General Weather Statistics

Objective:

- Identify overall weather patterns across Indian cities.

Findings:

- Average Temperature: 29.70 °C
- Average Humidity: 59.96%
- Average Wind Speed: 12.55 km/h

```
Average Temperature: 29.70 °C
Average Humidity: 59.96%
Average Wind Speed: 12.55 km/h
```

```python
# Average temperature, humidity, and wind speed
avg_temp = weather_data['temperature_celsius'].mean()
avg_humidity = weather_data['humidity'].mean()
avg_wind_speed = weather_data['wind_kph'].mean()

print(f"Average Temperature: {avg_temp:.2f} °C")
print(f"Average Humidity: {avg_humidity:.2f}%")
print(f"Average Wind Speed: {avg_wind_speed:.2f} km/h")
```

Temperature Distribution by Region

Objective:

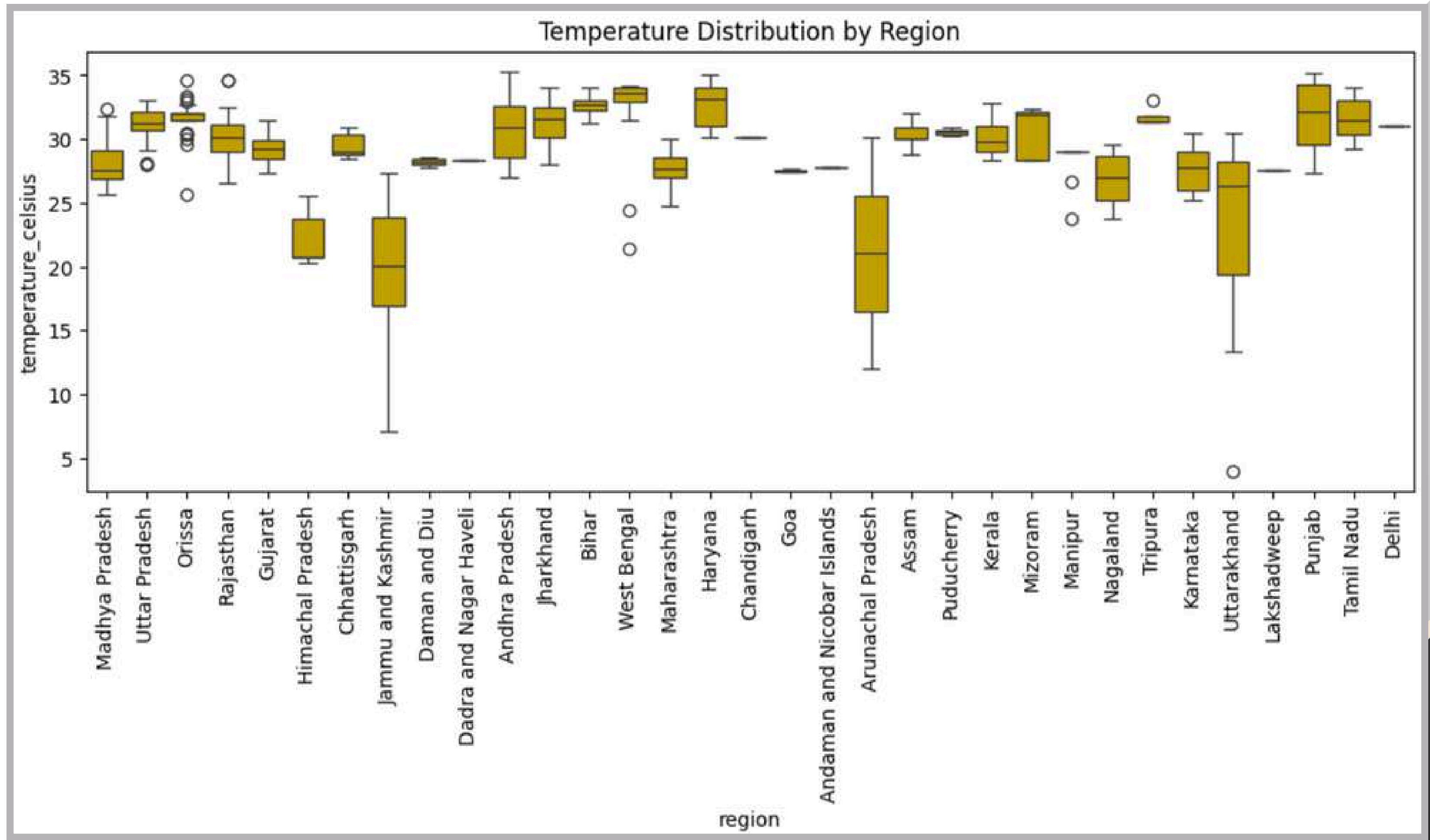
- Analyze how temperature varies by region.

Findings:

- Northern cities (Delhi, Jaipur) have wider temperature fluctuations.
- Southern cities (Chennai, Bengaluru) have moderate temperature variations.

Visualization:

- Box plot comparing temperature distributions by region.




```
plt.figure(figsize=(12, 4))
sns.boxplot(x='region', y='temperature_celsius', data=weather_data)
plt.xticks(rotation=90)
plt.title('Temperature Distribution by Region')
plt.show()
```

Correlation Between Weather Variables

Objective:

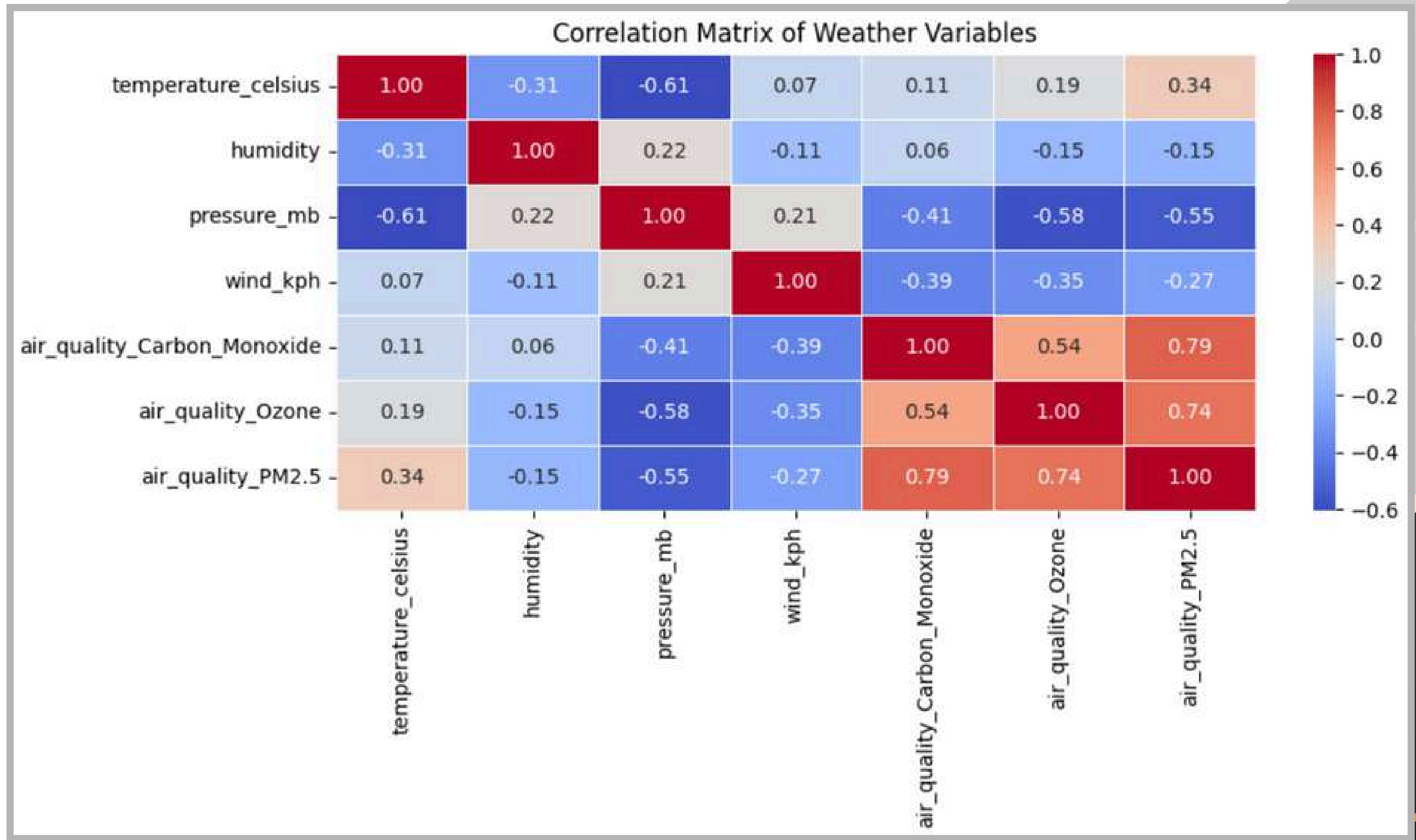
- Identify correlations between different weather parameters.

Findings:

- Temperature ↔ Humidity: Negative correlation.
- Wind Speed ↔ Temperature: Slight negative correlation.
- Cloud Cover ↔ Humidity: Positive correlation.

Visualization:

- Heatmap showing correlation matrix.




```python
# Select relevant numerical columns
corr_cols = ['temperature_celsius', 'humidity', 'pressure_mb', 'wind_kph',
             'air_quality_Carbon_Monoxide', 'air_quality_Ozone', 'air_quality_PM2.5']

# Correlation matrix
corr_matrix = weather_data[corr_cols].corr()

# Heatmap
plt.figure(figsize=(10, 4))
sns.heatmap(corr_matrix, annot=True, cmap='coolwarm', fmt=".2f", linewidths=0.5)
plt.title('Correlation Matrix of Weather Variables')
plt.show()
```

Location-wise Temperature Heatmap

Objective:

- Analyze how temperature varies by region.

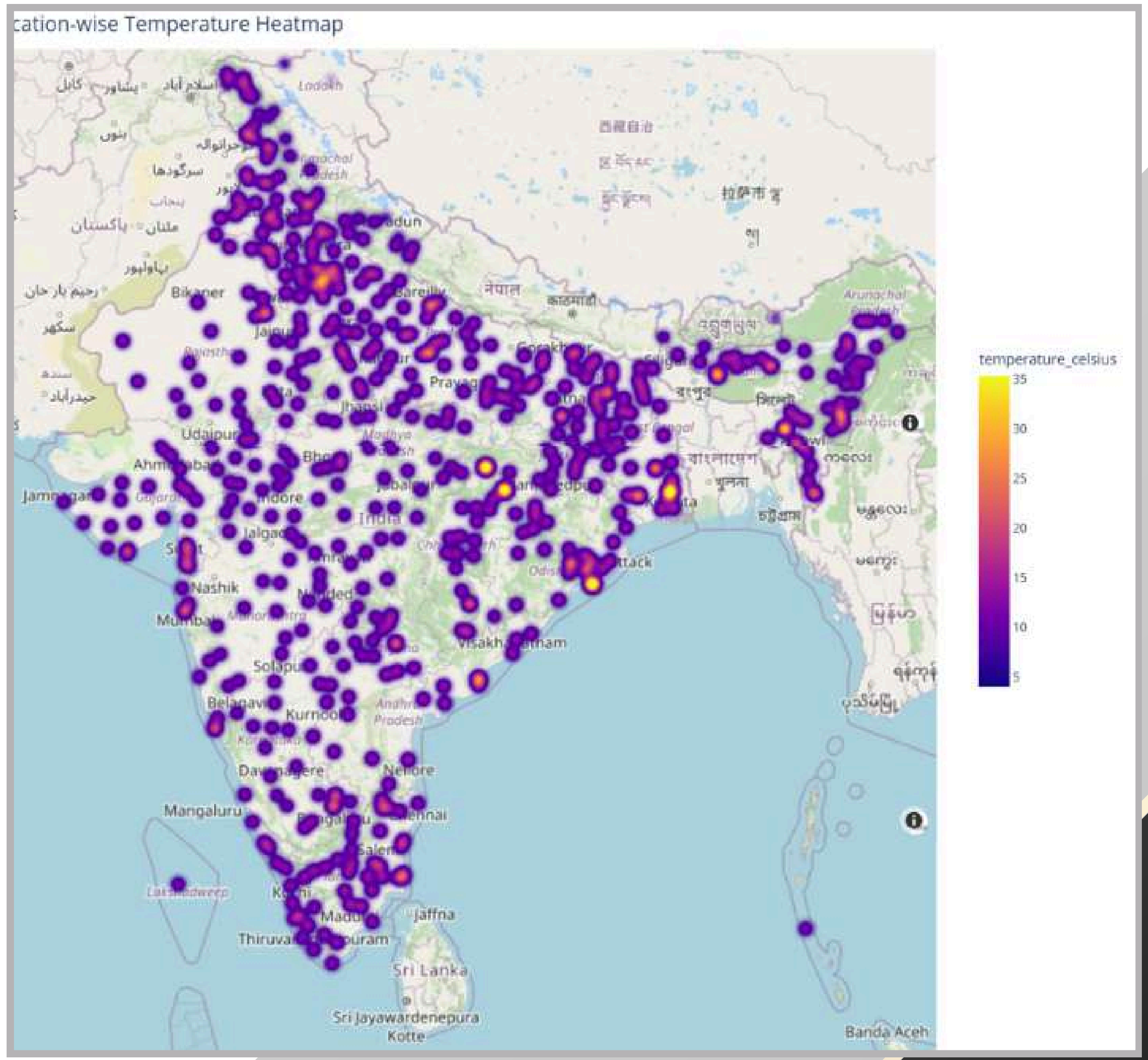
Findings:

- Northern cities (Delhi, Jaipur) have wider temperature fluctuations.
- Southern cities (Chennai, Bengaluru) have moderate temperature variations.

Visualization:

- Mapbox comparing temperature distributions by region.

```
fig10 = px.density_mapbox(weather_data, lat="latitude", lon="longitude", z="temperature_celsius", radius=11,
                          title="Location-wise Temperature Heatmap")
fig10.update_layout(mapbox_style="open-street-map")
fig10.show()
```



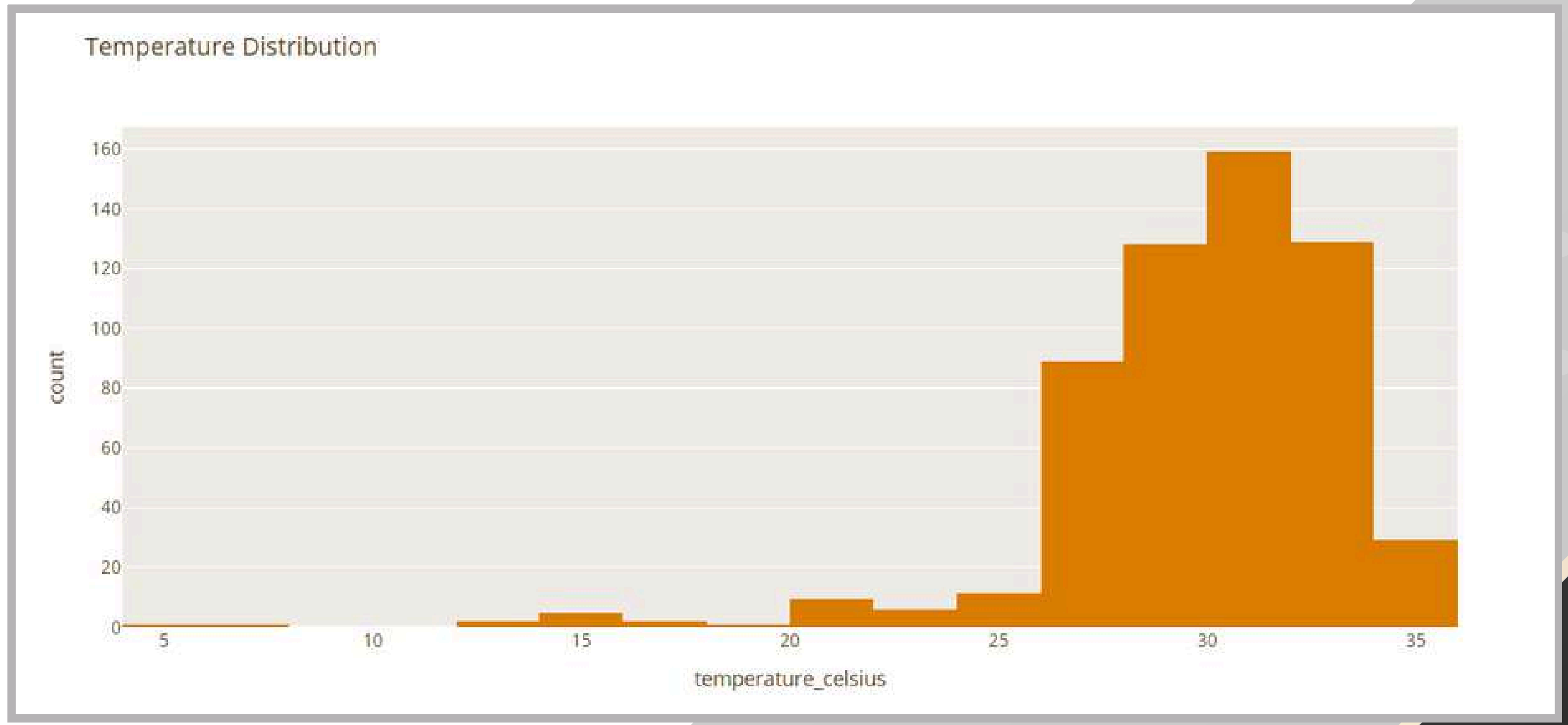
Temperature Distribution

Objective:

- Analyze how temperature is distributed across different regions.

Findings:

- The histogram shows that most temperatures range between 25°C and 35°C, indicating warm to hot weather conditions.
- Peak frequency occurs around 30-32°C, suggesting this is the most common temperature range across the dataset.
- Few occurrences of lower temperatures (below 20°C), indicating that colder weather is less frequent.



Visualization:

- The image displays a temperature distribution histogram.

```
fig1 = px.histogram(weather_data, x="temperature_celsius", nbins=20, title="Temperature Distribution")
fig1.show()
```

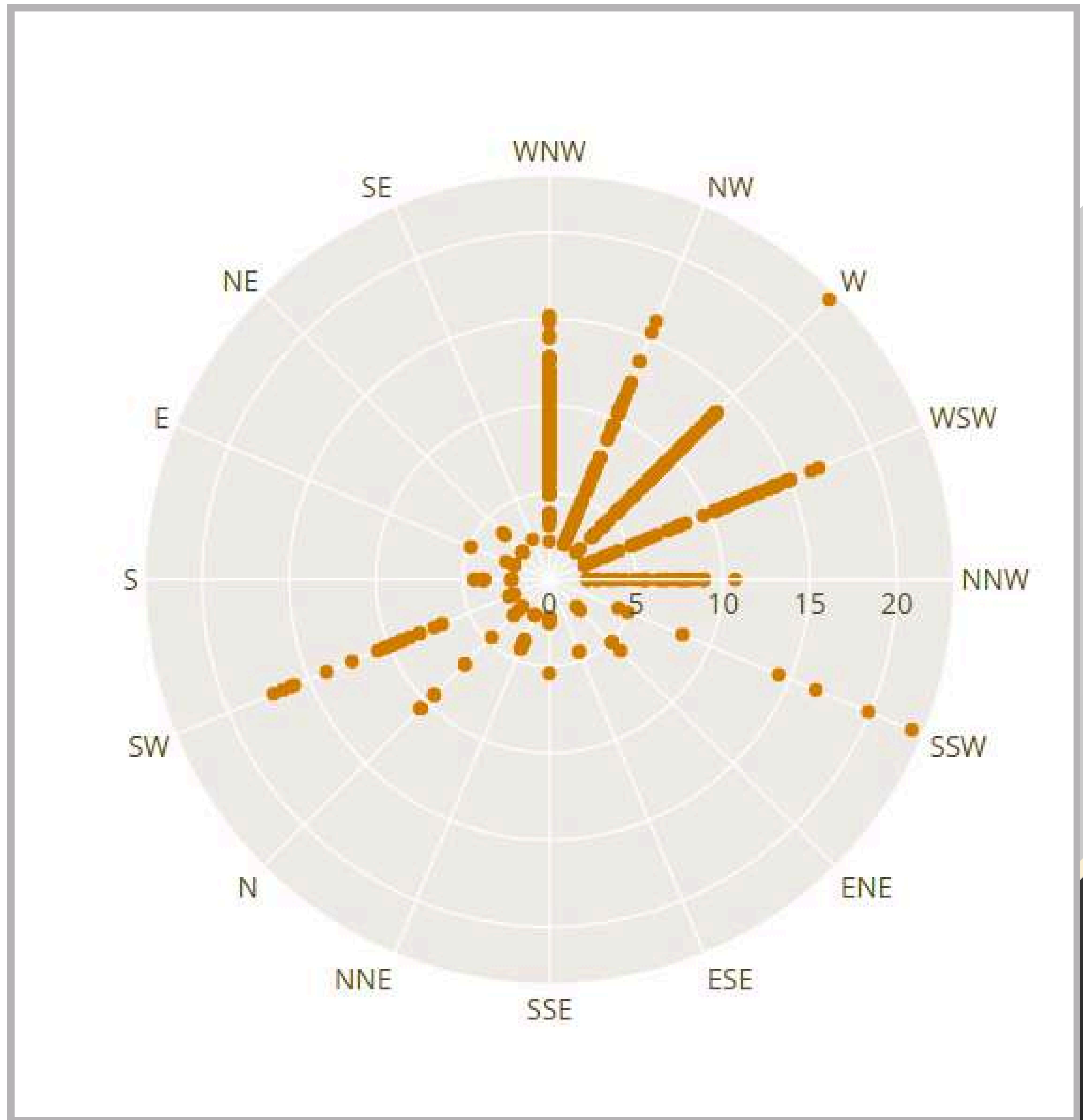
Wind Speed vs Wind Direction

Graph Description:

- The polar plot illustrates wind speed against wind direction. The length of the points indicates the wind speed, while their position reflects the direction.

Findings:

- Most of the wind speeds are concentrated in the W (West), NW (Northwest), and WNW (West-Northwest) directions.
- Higher wind speeds tend to occur towards the SSW (South-Southwest) direction.
- The wind speed is generally below 20 km/h, with only a few occurrences exceeding this value.




```
fig2 = px.scatter_polar(weather_data, r="wind_mph", theta="wind_direction", title="Wind Speed vs Wind Direction")
fig2.show()
```

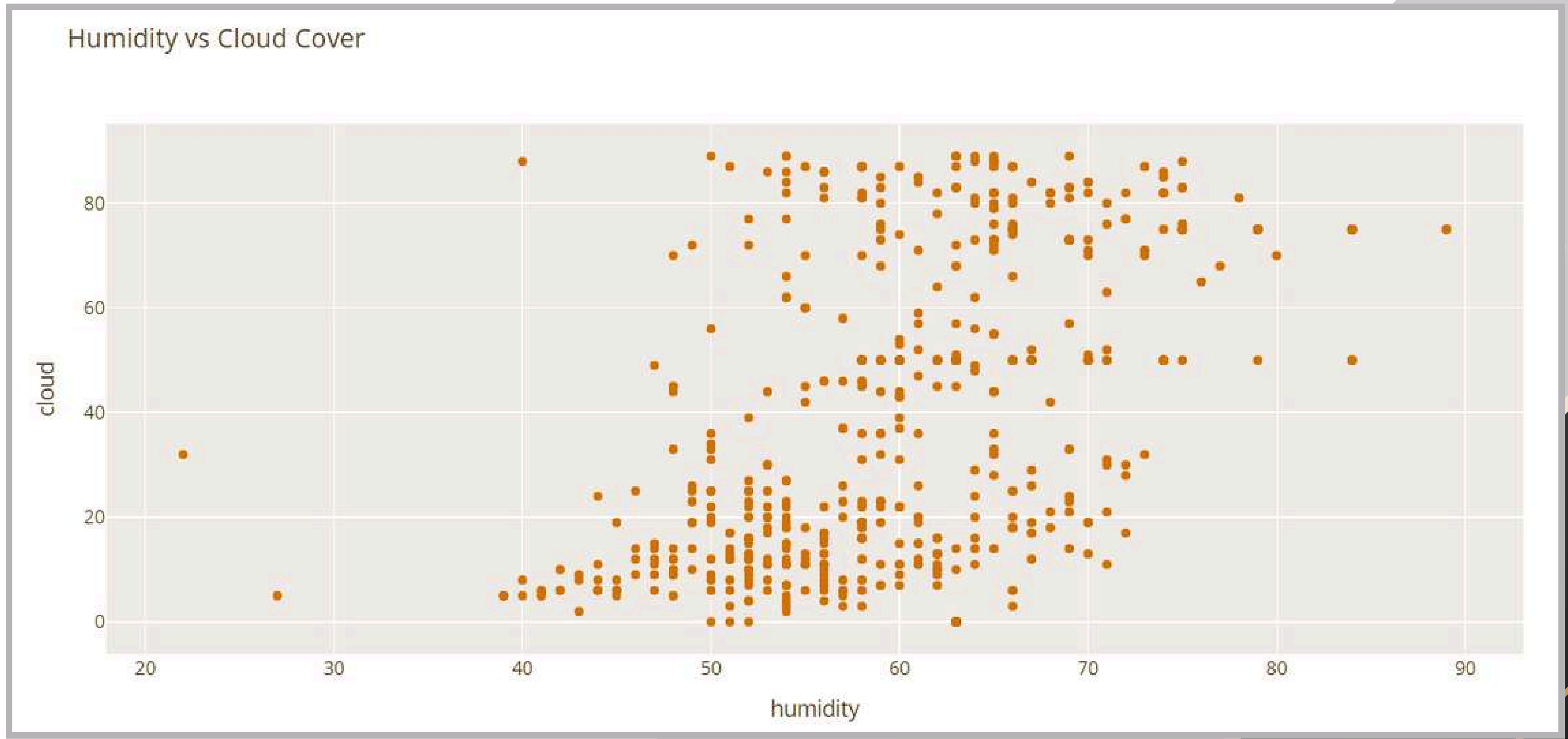
Humidity vs Cloud Cover

Graph Description:

- This scatter plot shows the relationship between humidity and cloud cover.

Findings:

- There is a positive correlation between humidity and cloud cover. As humidity increases, cloud cover also tends to increase.
- The cloud cover varies significantly when humidity is between 40% and 70%, indicating inconsistent cloud formation.
- At higher humidity levels, cloud cover tends to stabilize, forming denser cloud masses.




```
fig4 = px.scatter(weather_data, x="humidity", y="cloud", title="Humidity vs Cloud Cover")
fig4.show()
```

Air Quality Index Comparison

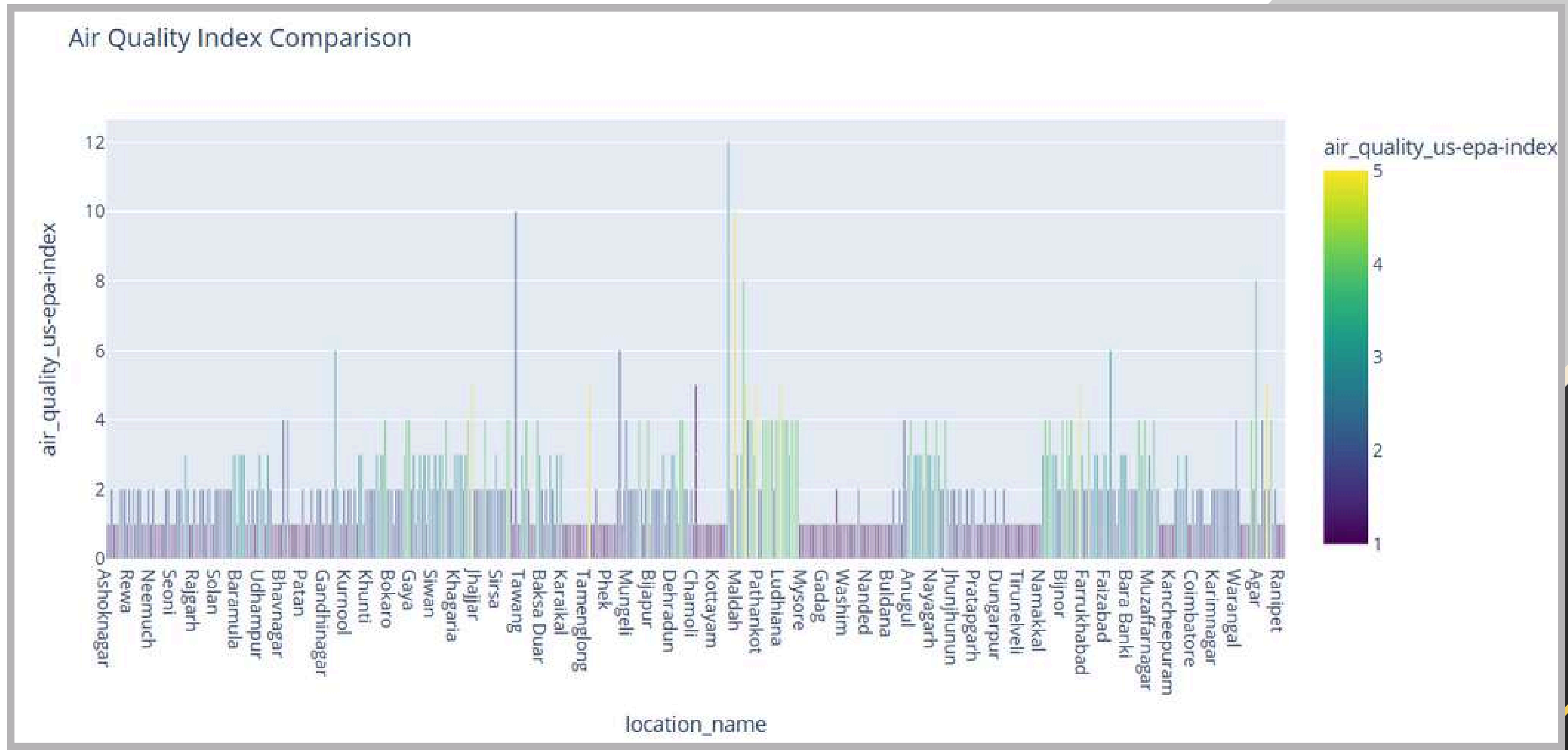
Graph Description:

- This bar chart compares the Air Quality Index (AQI) across multiple locations.

Findings:

- The majority of locations have an AQI value below 4, indicating moderate to good air quality.
- A few locations experience spikes above 6, suggesting areas with poorer air quality.
- The color gradient highlights variations in AQI, with darker shades indicating better air quality and lighter shades representing poorer conditions.

```python
fig5 = px.bar(weather_data, x="location_name", y="air_quality_us-epa-index",
              title="Air Quality Index Comparison",
              color="air_quality_us-epa-index",
              color_continuous_scale="Viridis"
             )
fig5.show()
```



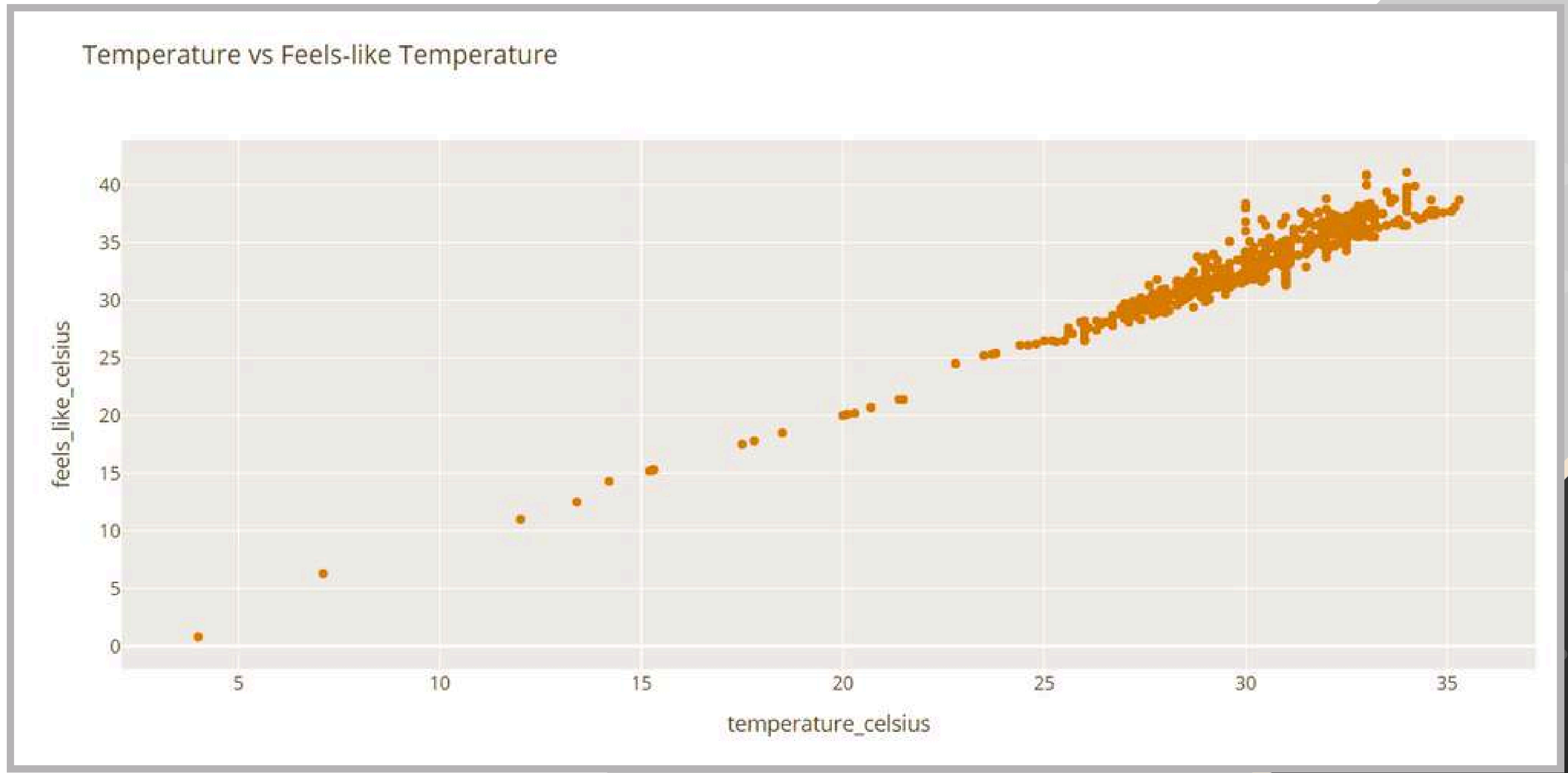
Temperature vs Feels-like Temperature

Graph Description:

- This scatter plot shows the relationship between actual temperature and perceived (feels-like) temperature.

Findings:

- There is a strong positive correlation between temperature and feels-like temperature.
- At higher temperatures (above 30°C), the feels-like temperature is slightly higher, indicating the impact of humidity and heat index.
- The difference between actual and feels-like temperature is more noticeable in the higher temperature range.




```
fig8 = px.scatter(weather_data, x="temperature_celsius", y="feels_like_celsius", title="Temperature vs Feels-like Temperature")
fig8.show()
```

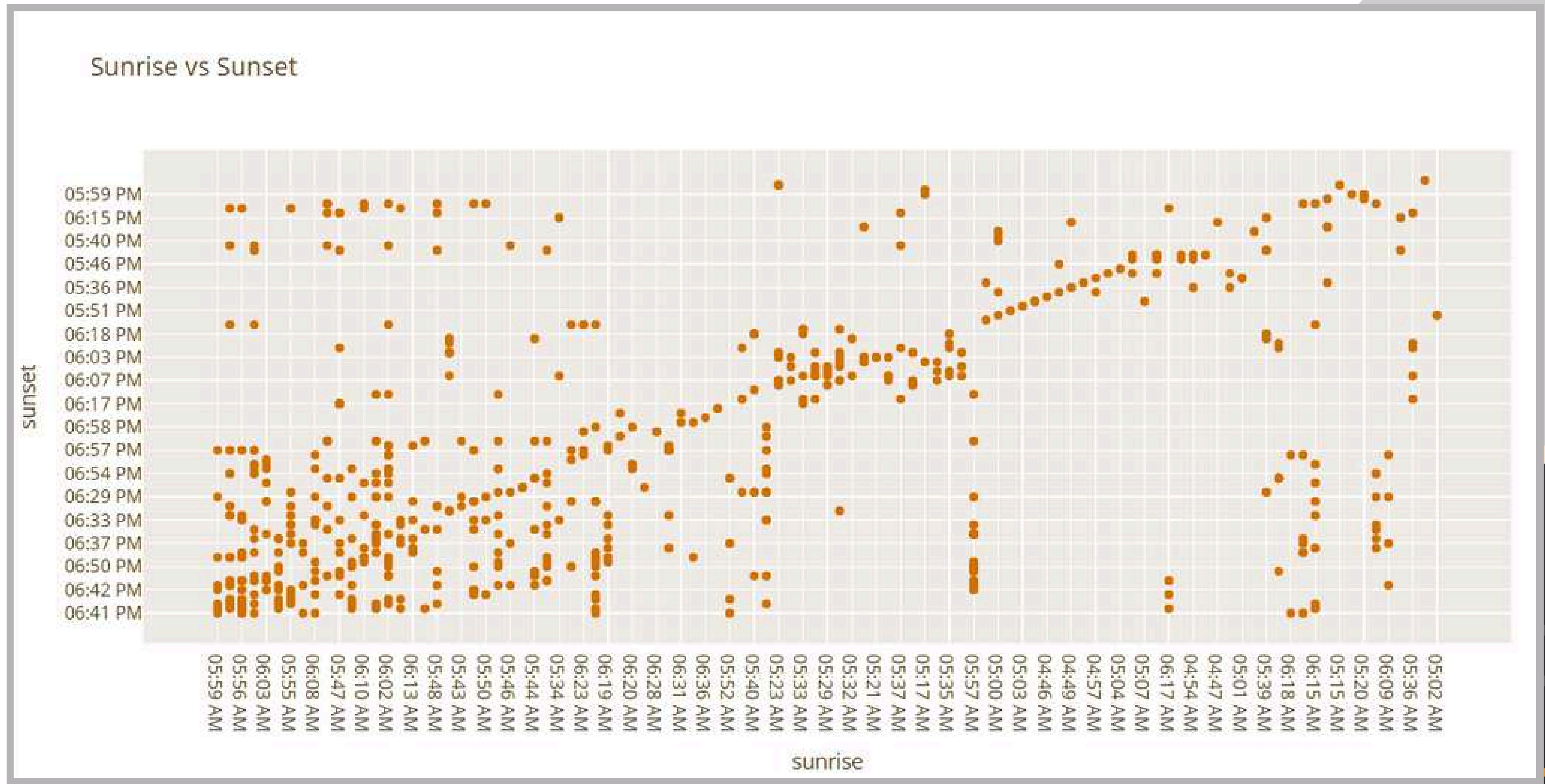
Sunrise vs Sunset

Graph Description:

- The scatter plot depicts the relationship between sunrise and sunset times.

Findings:

- There is a clear upward trend, indicating that as sunrise times occur later, sunset times also tend to be later.
- The sunrise times range from around 4:45 AM to 6:00 AM, while sunset times vary between 5:30 PM and 7:00 PM.
- The plot shows some outliers, where sunset occurs unusually early or late compared to the general trend.




```
fig9 = px.scatter(weather_data, x="sunrise", y="sunset", title="Sunrise vs Sunset")
fig9.show()
```

Key Insights and Outcomes

- **Insights:**
    - Northern regions experience greater temperature variability.
    - Southern regions have more consistent weather patterns.
    - Clear correlation between humidity and cloud cover.
    - Wind direction varies significantly by region.

- **Outcomes:**
    - Real-time weather data helps predict patterns.
    - Correlation analysis aids in understanding interdependencies.

Conclusion & Future Enhancements

- **Conclusion:**
    - The analysis provides insights into weather patterns across Indian cities.
    - Visualizations highlight key trends and relationships.
- **Future Enhancements:**
    - Include more weather parameters (e.g., precipitation, UV index).
    - Expand dataset with longer timeframes for seasonal analysis.
    - Use machine learning models to predict future weather trends.

THANK YOU