Historical Weather Dynamics for the Years 2006 to 2016

Prepared by: Adaobi Cynthia Okonkwo

Tools Used: Excel, Data Visualization Techniques

Category: Environmental Data Analytics

1. Outline

This report follows a structured roadmap to present the analysis clearly and logically:

- Introduction
- Story of Data
- Data Splitting and Preprocessing
- Pre-Analysis
- In-Analysis
- Post-Analysis and Insights
- Data Visualizations & Charts
- Recommendations and Observations
- Conclusion
- References & Appendices

2. Introduction

This project analyzes historical weather dynamics from 2006 to 2016, focusing on understanding the key factors that shaped daily and seasonal weather conditions during this ten-year period. The goal is to identify trends across temperature, humidity, precipitation, wind behavior, visibility, and atmospheric pressure, while uncovering patterns that influence climate variability over time. The analysis specifically examines weather outcomes across different seasons, years, and climate variables. The core problem being addressed is understanding how major weather factors change over time and identifying patterns that can support informed decision-making in environmental studies, urban planning, and climate awareness.

3. Story of Data

- **Data Source:** The data originates from a structured historical weather dataset, compiled from real-world meteorological observations and made publicly available on Kaggle.com.

- **Data Collection:** The data was collected through automated weather-monitoring systems, including meteorological sensors and weather stations. Each record captures daily observations such as temperature, humidity, precipitation type, wind speed and direction, visibility, cloud cover, and atmospheric pressure.

- **Data Structure:**

- Rows: Each row represents individual daily weather observations.
- Columns: Variables include Formatted Date, Summary, Precip Type, Temperature (°C), Apparent Temperature (°C), Humidity, Wind Speed (km/h), Wind Bearing (degrees), Visibility (km), Loud Cover, Pressure (millibars), and Daily Summary.

- **Important Features and Their Significance:**
  - Temperature & Apparent Temperature: Important for identifying seasonal changes and understanding heat dynamics.
  - Humidity: Helps reveal moisture patterns that influence rainfall and comfort levels.
  - Precipitation Type: Indicates weather conditions such as rain or snow, essential for environmental and planning insights.
  - Wind Speed & Direction: Useful for understanding storm behavior and weather transitions.
  - Visibility: Reflects air clarity, foggy conditions, and potential hazards.
  - Pressure: Key for detecting changes in weather systems, such as approaching storms.
  - Cloud Cover: Helps explain sunlight variation and sky conditions.

Additional engineered features, such as Date, Year, Month, Day, Time (hrs), Timezone, Season, Wind Category, Wind Direction, Visibility Category, and Pressure Category, were added to enrich the pattern

- **Data Limitations or Biases:** Potential limitations include:
  - Time Zone Variations: Date formatting issues (e.g., "+0200") may require cleaning to ensure proper time alignment.
  - Potential Sensor Error: Automated systems may have slight inaccuracies or gaps in data collection.

These limitations were considered and addressed through data cleaning and validation to ensure reliable results.

4. Data Splitting and Preprocessing

- **Data Cleaning:** The following cleaning steps were applied:
  - Correcting Column Names: The column "Loud Cover" was corrected to "Cloud Cover".
  - Removing Duplicates: Duplicate rows were identified and removed to avoid skewed analysis.

    - Correcting Errors: Any inconsistencies in date formatting (e.g., different time zones) were standardized.
    - Consistency Checks: Temperature, wind speed, and other numeric values were checked for unrealistic entries (e.g., negative visibility or wind speed).

- **Handling Missing Values:** No null values are present.
- **Data Transformations:**
    - Date-Time Conversion: The "Formatted Date" column was converted into a standard datetime format for easy filtering and time-series analysis.
    - Derived Variables: New columns were created, such as Date, Year, Month, Day, Time (hrs), Time zone, Season from the formatted date, Wind Category, Wind Direction, from Wind Speed (km/h), Visibility Category from Visibility (km), Pressure Category from Pressure (millibars), to support seasonal trend analysis.
- **Data Splitting:** Variables were separated into:
    - Independent Variables: Formatted Date, Apparent Temperature (°C), Humidity, Wind Speed (km/h), Wind Bearing (degrees), Visibility (km), Cloud Cover, Pressure (millibars).
    - Dependent Variable: Temperature (°C), Daily Summary, Precip Type, Summary.

This separation allowed for trend analysis, correlation studies, and visualization of how independent weather factors influence daily conditions.

- **Industry Context:** The data belongs to the meteorology and climate analytics industry, which supports weather prediction, environmental research, urban planning, agriculture, and disaster preparedness.
- **Stakeholders:** Key stakeholders include meteorologists, climate scientists, urban planners, agriculture managers, and energy sector analysts, all of whom can use the insights to make informed decisions based on historical weather trends.
- **Value to the Industry:** Insights from this analysis can help meteorological organizations, environmental agencies, and related industries improve forecasting accuracy, plan resources effectively, mitigate risks from extreme weather events, and optimize operations impacted by weather conditions.

5. Pre-Analysis

The pre-analysis phase provides an initial understanding of the dataset and uncovers early trends before deeper statistical evaluation.

- **Identify Key Trends:** Early exploration of the 2006 - 2016 weather data revealed several notable patterns:
    - Seasonal Temperature Patterns: Temperatures tend to be higher during summer months (June–August) and lower during winter months (December–February), showing a clear seasonal cycle.
    - Humidity Trends: Higher humidity levels are observed during rainy months, while drier periods correlate with lower precipitation.
    - Precipitation Patterns: Rainfall is concentrated in specific months, indicating seasonal spikes, while snow occurs mainly in winter months in colder regions.
    - Wind Speed and Direction: Stronger winds appear more frequently during transitional seasons (spring and autumn).
    - Visibility: Lower visibility is often observed during periods of rain, snow, or fog, consistent with precipitation patterns.
- **Potential Correlations:** Initial observations suggest possible relationships between key variables:
    - Temperature vs. Apparent Temperature: Apparent Temperature closely follows Temperature, but is slightly lower when humidity and wind effects are strong.
    - Humidity vs. Precipitation Type: Higher humidity is often associated with rainy days, suggesting a direct relationship between moisture levels and precipitation occurrence.
    - Pressure vs. Weather Conditions: Drops in atmospheric pressure often precede stormy or rainy weather, indicating a potential predictive correlation.
    - Wind Speed vs. Temperature/Pressure: Wind patterns may influence apparent temperature, with stronger winds making colder days feel even colder.
- **Initial Insights:**
    - Seasonal variations are very pronounced across multiple variables, suggesting the importance of considering month or season in deeper analysis.
    - Rain and snow events appear concentrated in specific periods, which could help in forecasting or understanding extreme weather trends.
    - Certain variables, like Temperature, Humidity, and Pressure, appear strongly correlated, which may simplify analysis when exploring patterns or building models.
    - Early observations suggest the dataset is suitable for trend visualization, seasonal analysis, and correlation studies.

These emerging patterns highlight key areas, such as seasonal temperature changes, precipitation trends, and wind and humidity behavior, that warrant further investigation.

6. In-Analysis

This section represents the core of the analysis, where the weather data from 2006 to 2016 is examined in detail to uncover meaningful insights, relationships among variables, and potential implications. The focus is on understanding how different weather factors interact and how these interactions shape overall weather dynamics.

- **Unconfirmed Insights (Hypotheses Under Evaluation):** During initial exploration, several patterns appeared promising but require further validation with statistical tools:
  - Temperature and Humidity: Periods of higher temperatures often appear to coincide with lower humidity levels, suggesting an inverse relationship that may vary by season.
  - Pressure and Precipitation: Drops in atmospheric pressure frequently align with rainy or stormy conditions, indicating a possible predictive link between pressure changes and precipitation events.
  - Wind Speed and Apparent Temperature: Higher wind speeds seem to reduce apparent temperature, particularly during colder months, supporting the hypothesis that wind intensifies the feeling of cold.
  - Seasonal Variability: Certain years appear to experience more extreme temperature fluctuations than others, which could point to underlying climatic variability.
- **Recommendations (Preliminary and Based on Emerging Patterns):** Although the analysis is ongoing, early patterns allow for initial recommendations:
  - Seasonal Segmentation: Future analyses should segment data by season or month to improve the accuracy of trend interpretation.
  - Weather Preparedness: Stakeholders such as urban planners and environmental agencies can use pressure and precipitation patterns to improve early warning systems for adverse weather.
  - Enhanced Monitoring: Increased focus on wind and humidity patterns can support a better understanding of comfort levels and weather-related risks.

- Further Statistical Analysis: Applying correlation coefficients or regression analysis in future work would help validate the observed relationships.

These recommendations will be refined once hypotheses are fully validated.

- **Analysis Techniques Used in Excel:** Excel served as the primary tool for exploring and validating relationships in the dataset. Key functionalities included:

  - Pivot Tables and Pivot Charts: Used to summarize weather variables across years, months, and seasons.
  - Sorting and Filtering: Applied to isolate specific weather conditions such as rainy days, extreme temperatures, or low-pressure events.
  - Descriptive Statistics: Functions such as AVERAGE, MIN, MAX, and COUNT were used to understand data distribution.
  - Calculated Columns: Created to extract year, month, and seasonal indicators from the date field.
  - Charts and Graphs: Line charts, bar charts, and trend visuals were used to illustrate temporal patterns and variable relationships.

7. Post-Analysis and Insights

This section summarizes the conclusions drawn from the analysis of historical weather data from 2006 to 2013 and evaluates whether the findings supported the initial assumptions and observed patterns. The study confirms several expected weather behaviors while also revealing subtle variations across years and seasons.

- **Key Findings:** Based on the completed analysis, several significant insights emerged:

  - There is a clear and consistent disparity between warm and cold seasons, with summer months recording significantly higher temperatures compared to winter months, highlighting a strong seasonal influence on weather patterns.
  - Apparent temperature closely follows actual temperature trends but is consistently lower during periods of high wind speed, emphasizing the impact of wind on perceived weather conditions.

- Precipitation type is closely linked to temperature, with rainfall occurring mainly during warmer conditions, snowfall limited to cold, high-humidity conditions, rain occurs in warmer, lower-pressure environments, and non-precipitation days align with stable, high-pressure and mild weather.
- Rain dominates across all seasons, with snow occurring mainly in winter and non-precipitation days being rare.
- Wind patterns are dominated by north–south flows, with northwest and southeast directions also common, while easterly winds occur least frequently, indicating limited weather movement from the east.
- Most days experience low to moderate wind speeds, while high-wind conditions are rare, indicating generally calm to moderate atmospheric conditions over the period analyzed.
- Visibility is highest during rainy conditions, lowest during snow, and moderately reduced on days without precipitation, indicating that snow has the strongest impact on reducing visibility.
- Humidity and Precipitation Relationship: Periods with rainfall are associated with notably higher humidity levels, confirming moisture as a key driver of precipitation events.
- While overall seasonal patterns remain consistent, certain years exhibit more pronounced temperature and pressure fluctuations, indicating periods of increased climatic variability.
- Cloud cover remains consistently at zero, indicating clear skies throughout the dataset.

- **Comparison with Initial Findings**

Comparing the validated results to initial expectations provides both alignment and unexpected discoveries:

- Validated Insights: Seasonal temperature cycles identified during pre-analysis were strongly confirmed. The anticipated relationship between humidity and precipitation was validated, as was the inverse relationship between wind speed and apparent temperature during colder periods.
- Surprising or Counter-Intuitive Results: Some periods with moderate temperatures still recorded low visibility due to cloud cover rather than precipitation, which was not initially

expected. Additionally, certain high-pressure periods experienced brief rainfall events, suggesting that pressure alone is not always a definitive predictor of weather stability.

- Insights: The final analysis confirms that a combination of temperature, humidity, wind, pressure, cloud cover, and seasonal timing shapes weather behavior. Understanding these interactions supports better forecasting, improved preparedness for adverse weather conditions, and more informed planning in climate-sensitive sectors such as agriculture, energy, and urban development.

8. Data Visualizations & Charts

Visual representations were created in Microsoft Excel to simplify complex data relationships and highlight key patterns related to the weather patterns. A summary of the visualizations and their interpretations is presented below.

- **Visuals**

  - **Actual vs Perceived Temperature Trend Over Time (Line Chart)**

This chart was used to show the trends and relationship between temperature and apparent temperature over time.

Insight: Temperatures peak in summer (July–August) with minimal difference between actual and apparent temperature, while winter months (December–January) are coldest, with apparent temperature lower due to wind chill.

  - **Overall Precipitation Type Distribution (Pie Chart)**

This chart visualizes the distribution of precipitation types (rain, snow, and days without precipitation) across the dataset.

Insight: Rain is the dominant precipitation type, while snow is occasional and clear or unclassified conditions are rare, indicating a primarily rain-prone climate.

  - **Average Temperature by Precipitation Type (Column Chart)**

This visualization shows how average temperature varies across different precipitation types—rain, snow, and no precipitation.

Insight: Precipitation type is temperature-dependent: rainfall occurs in mild to warm conditions, snowfall happens in cold, sub-zero weather, and days without precipitation are moderately cool.

- **Overall Seasonal Distribution of Precipitation Types (Stacked Bar Chart)**

This visualizes the proportion of rain, snow, and no-precipitation days across different seasons, highlighting seasonal patterns in weather conditions.

Insight: Rain dominates all seasons, with summer entirely rainy, spring and autumn mostly rainy, and winter the only season with substantial snow; days without precipitation are very rare.

- **Overall Climate Profile by Precipitation Type (Clustered Column Chart)**

This visualizes average temperature, humidity, and pressure for each precipitation type (rain, snow, and no precipitation), highlighting how climate conditions vary with different weather events.

Insight: Precipitation type is linked to atmospheric conditions: snow occurs in cold, low-pressure, humid conditions; rain occurs in warmer, lower-pressure, moderately humid conditions; and no precipitation happens in stable, high-pressure, mild-temperature environments.

- **Overall Distribution of Wind Conditions (Column Chart)**

This visualization shows the frequency of low, moderate, and high wind speed days, highlighting how often different wind conditions occur in the dataset.

Insight: Most days experience low to moderate wind speeds, while high-wind conditions are rare, indicating generally calm to moderate wind patterns in the region.

- **Overall Distribution of Wind Direction Frequency (Radar Chart)**

This visualization shows the frequency of different wind directions, highlighting which wind directions are most and least common throughout the dataset.

Insight: North and south winds are the most frequent, northwest and southeast winds are also common, while east winds are the least frequent, indicating prevailing wind patterns are primarily north–south and diagonal.

- **Overall Average Visibility Under Different Weather Conditions (Donut Chart)**

This visualization shows how average visibility varies across different weather conditions (rain, snow, and no precipitation), highlighting the impact of each condition on visibility.

Insight: Visibility is highest during rain, lowest during snow, and moderately reduced on days without precipitation, indicating that snow has the greatest impact on reducing visibility.

- **Historical Weather Dynamics for the Year 2006 to 2016 (Excel Dashboard)**

A consolidated dashboard was developed using slicers, pivot charts, and KPI indicators to allow interactive exploration of: Actual vs Perceived Temperature Trend Over Time, Overall Precipitation Type Distribution, Average Temperature by Precipitation Type, Overall Climate Profile by Precipitation Type, Overall Average Visibility Under Different Weather Conditions, and Overall Distribution of Wind Direction Frequency.

Dashboard Insights: The dashboard provides a clear, consolidated view of how temperature, precipitation, wind, and other weather factors interact over time. It allows stakeholders to quickly identify seasonal patterns, extreme conditions, and key trends influencing the region's overall climate behavior.

9. Recommendations and Observations

This section translates the findings from the weather analysis into actionable recommendations and practical observations for stakeholders such as meteorologists, urban planners, agriculture managers, and energy sector analysts.

- **Actionable Insights**

  - Since snow occurs primarily in winter and rain dominates other seasons, agencies should prioritize snow management and winter weather alerts during cold months and focus on flood or rain mitigation strategies in warmer months.

  - Snow significantly reduces visibility, while rain has minimal impact. Transportation and aviation sectors should strengthen monitoring and safety measures during snowfall periods.

  - Low to moderate wind conditions are common, but occasional high winds, though rare, can affect safety. Forecasts should highlight these periods for risk mitigation.

- Outdoor activities, energy usage, and crop planning should consider apparent temperature trends, as colder months feel even colder due to wind chill, affecting human comfort and operational efficiency.

- **Optimizations or Business Decisions**
  - Resource Allocation: Deploy personnel and equipment based on seasonal weather patterns, e.g., snowplows and de-icing resources in winter.
  - Infrastructure Planning: Urban planning, transport systems, and energy grids can optimize scheduling and capacity according to expected precipitation and temperature patterns.
  - Forecast Improvement: Combining temperature, pressure, humidity, and wind trends allows more accurate weather predictions, supporting better decision-making for climate-sensitive sectors.
- **Unexpected Outcomes**
  - Occasional rainfall occurred during high-pressure periods, suggesting that pressure alone is not always a definitive predictor of precipitation.
  - Some moderate temperature days had reduced visibility despite no precipitation, likely due to fog, haze, or mist, highlighting the need for multi-variable monitoring rather than relying solely on precipitation data.

10. Conclusion

This section summarizes the key findings, acknowledges limitations, and outlines directions for future analysis, providing closure to the report while highlighting actionable insights.

- **Key Learnings**
  - Rain is the dominant precipitation type across all seasons, while snow is largely restricted to winter.
  - Snowfall occurs at sub-zero temperatures and significantly reduces visibility.
  - Apparent temperature closely matches actual temperature in warm months but drops lower in colder periods due to wind chill.
  - Rain typically occurs under mild to warm temperatures with relatively good visibility.

- Weather behavior is influenced by the combined interaction of multiple atmospheric variables rather than a single factor.

- **Limitations**
    - Analysis was limited to the variables available in the dataset.
    - Cloud cover values were constant, preventing meaningful cloud-related analysis.
    - "Null" precipitation entries may represent missing or unclassified data.
    - Geographic differences and extreme weather events were not captured.
    - The dataset only covers 2006 - 2016, so long-term climate change analysis outside this range is limited.

- **Future Research**
    - Include additional variables such as variable cloud cover, solar radiation, or geographic location.
    - Apply time-series or predictive models to forecast temperature and precipitation trends.
    - Expand the dataset to cover longer periods and extreme weather conditions for deeper climate insights.

11. References

This section provides supporting materials and resources for readers who want to explore the analysis in more detail. It ensures transparency, reproducibility, and gives credit to sources and tools used.

- Dataset Source: Weather Dataset
- Tools Used: Microsoft Excel (Pivot Tables, Charts, Dashboard).
- External Resources:
    - Research articles on weather patterns, climate variability, and meteorological data analysis.
- **Appendices**

- Additional Charts and Tables: Detailed visualizations showing precipitation type distribution and seasonal patterns.
- Raw Data Samples: Subsets of the dataset used for analysis.
- Formulas and Functions**:** Step-by-step explanations of key Excel formulas used, including:

  Pivot tables were used to aggregate weather metrics by season, precipitation type, wind condition, and month.
- Dashboard Screenshots: Illustrations of the consolidated dashboard displaying KPIs and critical metrics for stakeholders.