COMP 495 WEATHER ANALYSIS PROJECT

LITERATURE REVIEW

Weather analysis plays an important role in many industries for helping people make better decisions, and for smoother businesses operations. For example, in outdoor sports, it is essential to decide the venue which will host the game. The games are decided based on playing conditions, which is also impacted by weather conditions. If weather statistics are known in advance, then important decisions can be made beforehand and lots of time, money, and resources can be used efficiently.

A brief history of weather analysis:

There have been interesting ways in which weather patterns have been analyzed in history. In the past, weather patterns were analyzed using a combination of observational data, weather instruments, and statistical methods. Below, there are some interesting ways in which the analytics in weather began:

1. Early Observations: Before the development of sophisticated instruments, weather patterns were primarily analyzed through direct observations. People would make note of atmospheric conditions, such as cloud formations, wind direction, and precipitation, using simple tools like thermometers, barometers, and rain gauges. These observations were often recorded in weather diaries or journals.
2. Weather Instruments: As scientific knowledge advanced, specialized instruments were developed to measure various weather parameters accurately. In the 17th and 18th centuries, instruments like the anemometer (to measure wind speed), hygrometer (to measure humidity), and barometer (to measure atmospheric pressure) were invented. These instruments allowed for more precise measurements and helped gather consistent data for analysis.
3. Weather Networks: In the 19th century, weather networks started to emerge, connecting various observation stations to share weather data. One prominent example is the establishment of the British Meteorological Office in 1854, which encouraged the collection and sharing of weather observations. This network facilitated the analysis of weather patterns across larger regions.
4. Weather Maps: The development of weather maps revolutionized the analysis of weather patterns. In the mid-19th century, scientists began creating weather maps to visualize and interpret weather data. This approach allowed for the identification of weather systems, such as high and low-pressure areas, fronts, and atmospheric circulation patterns.
5. Statistical Analysis: In the early 20th century, statistical methods became integral to weather pattern analysis. Meteorologists started applying statistical techniques to identify patterns and relationships in weather data. For example, they used correlation analysis to examine the connections between atmospheric variables and determine the likelihood of certain weather conditions occurring together.
6. Modern Techniques: Today, weather pattern analysis combines historical data, remote sensing technologies (like satellites and radar), and advanced computer modeling. These tools allow for the collection of extensive data, monitoring of global weather patterns in real-time, and the development of more accurate forecasting models. Machine learning and artificial intelligence techniques are also being employed to improve weather prediction and identify patterns that might not be apparent to human analysts.

In this project, my goal is to explore different weather data sources to find out the best source that serves as a good reference point for obtaining the data, while fulfilling all project requirements.

Project Scope-

In the province of Alberta, where extreme weather conditions are common, being well-informed about the external weather conditions is crucial for personal safety and preparedness. This project aims to analyze various weather factors such as wind speed, rain, humidity, visibility, snow, and sky conditions. By studying historical trends and patterns, the project seeks to provide valuable insights to help individuals make informed decisions and be better prepared for weather-related situations.

Focus/Content:

1. Collect and Analyze Data: Gather historical weather data from reliable sources, including meteorological records and databases, for the selected region in Alberta. Focus on obtaining data related to wind speed, rain, humidity, visibility, snowfall, and sky conditions.
2. Identify Trends and Patterns: Utilize statistical analysis techniques to identify recurring weather patterns and trends in the collected data. Determine long-term averages, seasonal variations, and extremes for each weather factor of interest.
3. Correlation Analysis: Explore relationships and correlations between different weather factors. Investigate how changes in one factor may influence or coincide with changes in other factors. For instance, assess how wind speed relates to visibility or how humidity impacts rainfall.
4. Visualization: Develop visual representations of the analyzed data, such as charts, graphs, and maps, to provide a clear understanding of the weather patterns over time. Present the data in an easily interpretable manner to facilitate informed decision-making.
5. Preparedness Guidelines: Based on the analyzed data and identified weather patterns, provide guidelines and recommendations for individuals to enhance their preparedness for different weather conditions. For example, advise on appropriate clothing, safety precautions, and travel considerations based on specific weather scenarios.
6. Forecasting and Early Warning Systems: Explore the potential for developing simple forecasting models or early warning systems based on historical data. Investigate if certain weather patterns can be used as indicators to predict upcoming weather events, allowing individuals to take proactive measures.
7. User Interface Development: Design and develop a user-friendly interface to access the analyzed weather data, visualizations, and preparedness guidelines. Ensure the interface allows users to easily navigate and interpret the information to make informed decisions regarding their activities and safety measures.

PROJECT DESIGN AND TECHNIAL STRATEGY

Introduction:

The project aims to conduct exploratory data analysis using Python to analyze weather data and gain valuable insights for improved preparedness. The design will focus on utilizing Python's core libraries such as NumPy, Pandas, Matplotlib, and Seaborn to import, manipulate, and visualize weather data.

Data Collection:

- Identify reliable sources for weather data specific to the selected region in Alberta.
- Explore meteorological databases, weather APIs, or government weather stations for comprehensive and historical data.
- Obtain weather data in a suitable format, such as CSV or Excel files, for easy import into Python.

Data Import and Preparation:

- Use Python's Pandas library to import the weather dataset into a Jupyter Notebook.

- Handle any missing or inconsistent data, ensuring data quality and integrity.

- Clean and preprocess the data, addressing outliers or anomalies as necessary.

- Structure the data into Pandas data frames for efficient manipulation and analysis.

Data Analysis Operations:

- Utilize Python's NumPy and Pandas libraries to perform exploratory data analysis on the weather dataset.

- Calculate descriptive statistics to understand the central tendencies, distributions, and variabilities of different weather factors like wind speed, rain, humidity, visibility, snowfall, and sky conditions.

- Apply filtering, grouping, and aggregation operations to extract relevant patterns and insights from the dataset.

- Explore correlation analysis to identify relationships and dependencies between different weather factors.

Data Visualization:

- Utilize Python's Matplotlib and Seaborn libraries to create visual representations of the weather data.

- Generate appropriate charts, graphs, and plots to illustrate the patterns, trends, and relationships discovered during the analysis.

- Customize visualizations with labels, legends, color schemes, and other elements to enhance interpretability and clarity.

Insights and Findings:

- Analyze the visualized data and derived statistical measures to extract meaningful insights and observations.

- Identify recurring weather patterns, trends, and anomalies that impact preparedness and decision-making.

- Present findings in a clear and concise manner, highlighting important weather factors and their implications for individuals in Alberta.

Project Report:

- Create comprehensive documentation outlining the entire project, including the data collection process, data preprocessing steps, analysis operations, and visualization choices.

- Provide clear explanations of the code implementation and logic behind each step.

- Include comments and annotations to enhance understanding and reproducibility of the project.

User Interface Development (PowerBI):

- Develop a user-friendly interface using visualization tool PowerBI which provides easy access to the analyzed weather data, and visualizations.

- Design an intuitive interface that allows users to interact with the data, explore visualizations, and access relevant information for better decision-making by using slicers and graphs.

The project design showcases how Python, along with key libraries, can be used for exploratory data analysis in weather analysis. By following this design, the project will provide valuable insights into weather patterns, correlations between weather factors, and trends in extreme weather conditions. The documentation and optional user interface will enhance the usability and accessibility of the project's findings, empowering individuals in Alberta to make informed decisions and be better prepared for various weather scenarios.

Project Look and Feel:

The project aims to leverage the capabilities of Python for data processing, analysis, and visualization, along with Power BI for creating impactful visualizations. The project's look and feel will focus on delivering a seamless and user-friendly experience throughout the data exploration and visualization process.

1. Data Processing and Analysis: Utilize Python's data processing and analysis libraries, such as Pandas and NumPy, to manipulate and analyze the weather dataset. Ensure that the code is well-structured, efficient, and follows best practices. Add comments and annotations to enhance code readability and maintainability.

2. Visualizations with Python: Leverage Python's visualization libraries, primarily Matplotlib and Seaborn, to create informative and visually appealing charts, graphs, and plots. Customize the visualizations to align with the project's overall look and feel, such as using color palettes consistent with the chosen visual theme.

3. Power BI Integration: After performing data processing, analysis, and visualization in Python, utilize Power BI to showcase the key findings and insights in a dynamic and interactive manner. Leverage Power BI's capabilities to create visually stunning dashboards and reports, incorporating the previously generated visualizations from Python.

4. User-Friendly Interface: Develop a user-friendly interface within Power BI that allows users to interact with the visualizations and explore the weather data effortlessly. Utilize Power BI's interactive features, such as drill-throughs, filters, and slicers, to enable users to delve into specific aspects of the data. Incorporate tooltips and explanatory text to provide additional context and guidance, enhancing the overall user experience.

8 WEEK MILESTONES WITH TIMELINE:

Week 1: Project Setup and Data Collection
*   Research and identify reliable sources for weather data specific to Alberta.
*   Obtain weather dataset in a suitable format (CSV or Excel) from the chosen sources.
*   Set up the development environment, including Python, Jupyter Notebook, and necessary libraries (NumPy, Pandas, Matplotlib, Seaborn)

Week 2: Data Import and Preparation
*   Import the weather dataset into a Jupyter Notebook using Pandas.
*   Clean and preprocess the data, handling missing values, outliers, and inconsistencies.
*   Organize the data into Pandas dataframes for efficient manipulation.

Week 3: Data Analysis Operations
*   Perform descriptive statistics on weather factors such as wind speed, rain, humidity, visibility, snowfall, and sky conditions.
*   Apply filtering, grouping, and aggregation operations to extract relevant patterns and insights from the dataset.

- Explore correlation analysis to identify relationships and dependencies between different weather factors.

Week 4: Data Visualization with Python
- Utilize Matplotlib and Seaborn libraries to create visual representations of the weather data.
- Generate appropriate charts, graphs, and plots to showcase patterns, trends, and relationships discovered during analysis.
- Customize visualizations with labels, legends, color schemes, and other elements for clarity and visual appeal.

Week 5: Insights and Findings
- Analyze the visualized data and derived statistical measures to extract meaningful insights and observations.
- Identify recurring weather patterns, trends, and anomalies impacting preparedness and decision-making.
- Document and summarize key findings, highlighting important weather factors and their implications.

Week 6: Project Documentation
- Create comprehensive documentation outlining the entire project, including data collection, preprocessing, analysis, and visualization.
- Provide clear explanations of the code implementation and logic behind each step.
- Include comments and annotations to enhance reproducibility and understanding of the project.

Week 7: Optional User Interface Development
- If applicable, develop a user-friendly interface using frameworks like Flask or Django.
- Implement features that allow users to interact with the analyzed weather data, explore visualizations, and access relevant information.
- Ensure the interface aligns with the project's look and feel, providing an intuitive and engaging user experience.
-

Week 8: Finalize and Review
- Conduct a thorough review of the entire project, including code, documentation, and optional user interface.
- Make necessary refinements and improvements based on feedback and observations.
- Prepare a final presentation or report summarizing the project's objectives, methodologies, findings, and potential applications.

PROJECT DELIVERABLES:

1. Weather Dataset: A curated weather dataset specific to Alberta, collected from reliable sources and provided in a suitable format (CSV or Excel) for analysis
2. Jupyter Notebook: A well-documented Jupyter Notebook containing Python code and step-by-step instructions for data import, preprocessing, analysis operations, and visualization using libraries such as NumPy, Pandas, Matplotlib, and Seaborn.
3. Data Analysis Results: A comprehensive analysis of weather factors, including descriptive statistics, patterns, trends, correlations, and significant insights extracted from the dataset.

4. Visualizations: Visual representations of the weather data using Matplotlib and Seaborn, showcasing patterns, trends, and relationships identified during the analysis. The visualizations should align with the project's look and feel, ensuring clarity and interpretability.
5. Documentation: A detailed project documentation outlining the entire process, including data collection, preprocessing, analysis, and visualization. It should provide clear explanations of the code implementation, logic behind each step, and the interpretation of results. The documentation should also include comments and annotations to facilitate understanding and reproducibility.
6. User Interface within PowerBI: A user-friendly interface developed using PowerBI, providing easy access to the analyzed weather data, interactive visualizations, and preparedness guidelines. The interface should have slicers, filters and graphs to visualize the data better and enhance user interactivity with the report/dashboard.

EVALUATION CRITERIA

The purpose of this evaluation proposal is to outline the criteria and process for evaluating the completeness and quality of the project deliverables for the Weather Analysis project. The evaluation will focus on assessing the Jupyter Notebook, PowerBI file, and project report submitted as part of the project.

Project Proposal (15%):

- Alignment with Objectives: Consistency between the proposed objectives, methodologies, and deliverables outlined in the project proposal and the actual implementation.
- Feasibility: Assessment of the project proposal's practicality, considering available resources and time constraints.

Python Code and Power BI File (50%):

- Code Quality: Organization, structure, readability, and adherence to best coding practices.
- Data Import and Preprocessing: Proper handling of data import, preprocessing, and cleaning techniques.
- Analysis Operations: Implementation of data analysis operations using NumPy and Pandas.
- Interactive Visualization: Effective use of Matplotlib and Seaborn to create informative and visually appealing visualizations. Utilization of Power BI's interactive features to enhance user experience and exploration of the weather data.

Since these are two pure implementations files- Python and PowerBI, these have comprehensive work behind them and have a word length of their own. Therefore, there is not a set 5000-7000 word limit, as these have equivalent or more work put in for the implementation of code and dashboard.

Project Report (35%):

- Structure and Organization: Logical flow and organization of the project report, including clear sections and subheadings.
- Clarity of Explanations: Clear and concise explanations of the project objectives, methodologies, and findings.
- Interpretation of Results: Insightful analysis and interpretation of the weather data, supported by appropriate visualizations.
- Completeness: Inclusion of all relevant information, including data sources, preprocessing steps, analysis techniques, and limitations.