Extension Plan Report: Wildfire Smoke Impact on Healthcare in Boulder, Colorado

1. Motivation/Problem Statement

In recent years, Boulder, Colorado, has experienced a significant rise in the frequency and intensity of wildfires in its vicinity. This trend has led to elevated smoke levels in the community, posing severe health risks to its residents. Wildfire smoke contains pollutants like particulate matter (PM2.5), carbon monoxide, and other toxins that can lead to respiratory issues, cardiovascular problems, and long-term health impacts, especially for vulnerable populations such as children, the elderly, and those with preexisting health conditions. For these individuals, even short-term exposure to wildfire smoke can lead to severe health consequences, while longer-term exposure can contribute to chronic diseases.

In my initial analysis, I developed a time series model to predict future smoke exposure based on historical wildfire data. Extending this analysis to assess the health impacts on Boulder residents offers a valuable next step. Specifically, this extension aims to investigate correlations between increased smoke exposure and healthcare metrics, such as respiratory-related hospital admissions, emergency room visits, and mortality rates. By linking predicted smoke exposure levels to health outcomes, this analysis will offer Boulder's public health and emergency response teams data-driven insights to support planning and resource allocation.

From a scientific perspective, this analysis also contributes to the broader understanding of how environmental stressors like wildfire smoke affect urban populations' health. In the face of climate change, wildfires are expected to become more frequent and intense, making it essential to understand the full scope of their impacts on community health. By quantifying these impacts and understanding seasonal trends, the results of this analysis will help inform long-term policy decisions, allowing Boulder to build resilience against future wildfire events and protect the health of its residents.

2. Impact Focus

The primary focus of my extension is on healthcare impacts, with an emphasis on understanding how wildfire smoke exposure affects respiratory health in Boulder. Given that wildfire smoke contains pollutants like PM2.5, carbon monoxide, and other toxins, prolonged exposure can have a profound effect on public health. By examining the correlation between predicted smoke levels and respiratory-related illnesses, I aim to quantify the potential healthcare burden for Boulder during wildfire seasons. Specifically, this study will explore metrics such as emergency room visits, hospital admissions, and other indicators of respiratory distress and illness.

The insights gained from this analysis can serve as a vital resource for local healthcare providers, public health officials, and policymakers. By providing a clearer understanding of the relationship between wildfire smoke exposure and health outcomes, the results will support evidence-based decision-making around resource allocation, hospital preparedness, and targeted public health interventions. Additionally, this research will identify the most vulnerable populations—such as children, the elderly, and individuals with preexisting respiratory or cardiovascular conditions—who are likely to be disproportionately affected by increased smoke levels.

In practical terms, the findings could lead to the development of early-warning systems, public health advisories, and tailored healthcare protocols to mitigate the impact of wildfire smoke on Boulder's residents. By proactively addressing these health risks, the community can better prepare for and respond to future wildfire events, ultimately reducing the strain on healthcare resources and safeguarding public health. This healthcare-focused extension will also contribute valuable data to inform long-term policy decisions, enabling Boulder to build resilience against the anticipated increase in wildfire incidents due to climate change.

3. Data or Model to be Used

For this extended analysis, I plan to incorporate additional health-related datasets to assess the impact of wildfire smoke exposure on public health outcomes in Boulder. These datasets will provide information on hospital visits, admissions, respiratory illnesses, and other relevant health indicators. Specifically, I intend to utilize the following resources:

1. **Colorado Department of Public Health & Environment (CDPHE) Datasets**:
    - **Vision Portal**: This portal offers extensive health data, including hospital admissions, mortality rates, and incidences of specific respiratory diseases such as asthma and COPD. This resource is crucial for identifying trends in respiratory illnesses that may correlate with increased smoke exposure. The data can be accessed through the CDPHE Vision Portal.
    - **Vital Statistics Program**: This program provides comprehensive mortality data, along with detailed statistics on health conditions across Colorado, enabling a closer examination of the potential rise in respiratory-related mortality associated with smoke exposure. Further information is available through the CDPHE Vital Statistics Program.
    - **Data License**: To access these datasets, I have contacted CDPHE and submitted a data request form, as they require authorization for use in research. The data will serve as the foundation for analyzing the relationship between wildfire smoke exposure and hospital admissions for respiratory conditions in Boulder.
2. **Adapting the SARIMA Model**:
   Building on the existing SARIMA model developed for forecasting smoke exposure, I will integrate the CDPHE health datasets to correlate smoke estimates with public health

outcomes. Specifically, I aim to adapt the SARIMA model's output to investigate trends in health impacts, such as emergency room visits and hospital admissions, that align with smoke exposure peaks. This will involve:

- **Combining Smoke Predictions with Health Data**: By overlaying smoke estimates with hospital admission data, I can assess the extent to which increased smoke exposure impacts public health over time.
- **Analyzing Correlation and Causality**: This step will focus on identifying significant correlations between predicted smoke levels and health indicators, providing insights into the causal relationships that may exist between wildfire events and respiratory health conditions.
- **Assessing Vulnerability of Specific Demographics**: In addition to general health trends, the analysis will consider data by age group, socioeconomic status, and preexisting health conditions to highlight vulnerable populations.

Integrating the CDPHE datasets with the SARIMA predictions will allow for a comprehensive view of how smoke levels could affect public health, thus enabling Boulder's policymakers to anticipate and mitigate health risks associated with wildfire smoke. By aligning healthcare needs with smoke exposure forecasts, the model aims to inform proactive public health strategies to safeguard Boulder's residents.

4. Unknowns and Dependencies

There are several factors that may influence the feasibility and accuracy of this extended analysis:

1. **Data Access and Permissions**:
   The analysis heavily depends on access to detailed health data from the Colorado Department of Public Health & Environment (CDPHE). Although I have submitted a formal data request, any delays in receiving access, limitations on data granularity, or outright denial could significantly impact my ability to draw meaningful correlations between smoke exposure and health outcomes. If access is delayed or denied, I may need to find alternate data sources or adjust the scope of my analysis.
2. **Data Compatibility and Time Series Alignment**:
   Combining time series data from different sources can be challenging, especially when each dataset may use different temporal resolutions or time frames. For example, my smoke impact estimates are on an annual basis, while health data may be collected daily, monthly, or annually. Harmonizing these time scales for consistent analysis will require preprocessing, and in some cases, data aggregation. Additional preprocessing will be needed to ensure the data aligns accurately, preserving the integrity of the trends without introducing temporal biases.
3. **Model Adaptation and Validation**:
   Modifying the SARIMA model to include health data will require careful adjustments, as health-related time series data might exhibit patterns, trends, and seasonality different from environmental smoke data. Adjusting the model's parameters to accurately capture these variations will involve additional model tuning and validation. Testing multiple

model configurations may be necessary to ensure that the health impacts are captured effectively within the seasonal and trend-based framework of SARIMA. Each adjustment will require time and might involve iterative testing.

4. **Interpretation of Results and External Variables**:
   While the analysis aims to establish a relationship between smoke exposure and health impacts, it's crucial to recognize that many external factors—such as weather, air quality, socioeconomic factors, and public health interventions—also influence health outcomes. Differentiating the impact of wildfire smoke from these other variables will be challenging. The analysis will focus on identifying correlations rather than establishing causality. Careful interpretation of the results is essential to avoid overgeneralization, as the observed correlations may not imply a direct cause-and-effect relationship. Addressing these limitations transparently in the analysis will be necessary to provide a realistic context for the findings.

5. Timeline to Completion

To complete this extended analysis on time, I have developed a detailed timeline that outlines key milestones, significant tasks, and deadlines. Each phase is designed to build upon the last, ensuring a structured and systematic approach to the integration of health impact data with smoke estimates.

1. **By November 10: Data Access and Preprocessing**
   - *Goal*: Secure access to health datasets from the Colorado Department of Public Health & Environment (CDPHE).
   - *Tasks*:
     - Confirm access to requested health data, specifically related to hospital admissions, respiratory illnesses, and other relevant metrics.
     - Clean and preprocess the datasets to prepare them for analysis. This will involve handling missing values, aligning temporal scales, and ensuring consistency in variable names and formats.
2. **By November 17: Data Integration and Initial Trend Analysis**
   - *Goal*: Combine health data with smoke estimates to gain preliminary insights into trends.
   - *Tasks*:
     - Merge health outcome data with existing smoke impact estimates for Boulder to create a combined dataset.
     - Conduct exploratory data analysis (EDA) to understand trends over time, identify potential correlations, and observe any seasonal or cyclical patterns in health data that might align with wildfire smoke exposure.
     - Document initial observations that might guide further analysis.
3. **By November 20: Model Adaptation and Validation**
   - *Goal*: Adapt the SARIMA model to incorporate health data for correlation analysis with smoke exposure.
   - *Tasks*:

        - Modify the SARIMA model to integrate health outcome data, accounting for both the seasonal and trend components unique to healthcare metrics.
        - Use cross-validation techniques to validate the model's predictive performance. Testing on smaller subsets will help ensure robustness and prevent overfitting.
        - Fine-tune the model parameters to improve the accuracy of predictions and correlations with smoke impact.
4. **By November 23: Visualization Development**
    - *Goal*: Create visualizations to effectively communicate findings and correlations between smoke exposure and health outcomes.
    - *Tasks*:
        - Develop a series of visualizations, such as line plots to show year-over-year changes in smoke levels and associated health outcomes, and heatmaps to illustrate high-correlation periods.
        - Finalize charts that clearly show the trends and potential correlations, making them accessible and understandable for stakeholders in Boulder.
        - Ensure visualizations highlight key findings while providing transparency on uncertainty and confidence intervals where applicable.
5. **By November 27: Presentation Preparation and Delivery**
    - *Goal*: Prepare and deliver a presentation summarizing key findings and methods.
    - *Tasks*:
        - Create a concise presentation focusing on the methodology, data integration, and the observed relationship between smoke exposure and health outcomes.
        - Emphasize insights valuable for city planning, particularly around public health preparedness and resource allocation.
        - Practice the presentation to ensure clarity and address anticipated questions about model assumptions and data limitations.
6. **By December 1: Final Analysis and Refinement**
    - *Goal*: Refine analysis based on any feedback from the presentation and complete the extended analysis.
    - *Tasks*:
        - Integrate feedback from the presentation to clarify any points that may need additional detail or explanation.
        - Finalize the analysis, double-check data handling steps, and ensure that all steps are reproducible for transparency.
        - Begin drafting the final report, ensuring that all findings are well-supported by the data and presented in a clear, organized manner.
7. **By December 4: Final Report and Repository Submission**
    - *Goal*: Submit the completed project, including all code, datasets, and the project report, with a focus on providing recommendations.
    - *Tasks*:
        - Complete the written report, covering the extended analysis and specific recommendations for Boulder's public health planning.

    - Finalize the project repository with clearly documented code, organized datasets, and instructions for replicating the analysis.
    - Ensure that the repository README provides an overview, instructions, and acknowledgments, guiding users through the analysis steps.

6. Links to Data and Resources

- **Colorado Department of Public Health & Environment (CDPHE) - Vision Portal**: CDPHE Vision Portal
- **CDPHE Vital Statistics Program**: CDPHE Vital Statistics
- **Wildfire Data**: ScienceBase Wildfire Data
- **EPA Air Quality Data**: EPA AQS Data