**Final Project Proposal**

**Coal Energy and Mortality: An Analysis of Towns with and without Coal Energy Plants**

Trevor D. Snedden and Aaron J. S. Roberts

College of Applied Science and Technology, University of Arizona

APCV 361 Data Analysis and Visualization

Dr. Li Xu

March 26, 2023

**Introduction**

Coal energy has been a prevalent source of electricity in the USA for decades. According to the U.S. Energy Information Administration (2023), it accounts for 19% of energy production in the U.S. However, its environmental and health impacts are a cause for concern. “When coal is burned it releases a number of airborne toxins and pollutants. They include mercury, lead, sulfur dioxide, nitrogen oxides, particulates, and various other heavy metals. Health impacts can range from asthma and breathing difficulties, to brain damage, heart problems, cancer, neurological disorders, and premature death.” (Union of Concerned Scientists, 2017). The question is then raised if communities located close to coal energy plants pose a greater mortality risk to those residents than those living in counties with non-coal energy plants.

**Project Objectives**

This research project aims to analyze the mortality rate of towns with and without coal energy plants and determine whether there is a significant difference between the two. Specifically, this project aims to answer the following research questions:

1. What is the mortality rate in towns with coal energy plants compared to counties without non-coal energy plants?
2. Is there a significant difference in mortality rates between these two types of counties?

**The Background and Significance of the Project**

Although the popularity of coal as an energy has dramatically reduced in the U.S. over the past few decades, it is still a prevalent source of energy in the U.S. If our project can demonstrate a significant difference between mortality rates in counties with coal energy power plants that could indicate an issue that policy makers should address regarding the placement of coal plants near residential areas. This project could highlight the risks associated with coal as an energy source. Furthermore, this could be used to advocate for more the necessity for alternative energy sources. Ultimately, This project could impact both public health and policy.

**Project Design or Development**

Data Collection & Cleaning:

To address the research questions, we will collect data on the mortality rates of counties with and without non-coal-powered energy plants. Non-coal plants consist of Oil, Hydro, Gas, Wind, and Biomass. The few plants that burned another fuel alongside coal are placed into the coal plant category. The primary data source utilized in this study will be obtained from the EPA for all energy-producing power plants in the United States and the University of Wisconsin Population Health Institute’s county health rankings & roadmaps documentation for the mortality rates per county within the United States. We will clean the data by removing duplicates, missing values, and outliers. We will also standardize the data to ensure that all data points are comparable. This will be accomplished by filtering matching counties between the EPA’s data set with the county mortality data set retrieved from the County Health Rankings & Roadmap’s Analytic Data. All unmatched counties will be categorized as outliers and not accounted for in this study. We will do this using Jupyter Notebooks and the Pandas library.

Data Analysis / Visualization:

We will use statistical analysis methods, such as t-tests and regression analysis, to analyze the mortality rates of towns and counties with coal energy plants and without non-coal energy plants. We will compare the mortality rates of the two groups and determine if there is a statistically significant difference between them. We will use data visualization tools, like graphs and charts to present our findings clearly and concisely. This will help us to better communicate our results to policymakers and the public. We use Python’s MATPLOTLIB library to create graphs and charts.

Interpretation and Conclusion:

We will interpret our findings and draw conclusions based on our analysis. Suppose we find a statistically significant difference in mortality rates between towns and counties with and without coal energy plants. In that case, we will discuss the policy implications of our results and recommend alternative energy sources to coal. If we do not find a significant difference, we will discuss the limitations of our study and guide further research to better understand the health impacts of coal energy.

**Project Timeline**

* Project log will be updated on a weekly basis

Week 1 (3/27-4/2):

* Formulate specific research questions and define the scope of the project
* Start gathering and cleaning data
* Begin working on the project log report

Week 2 (4/3-4/9):

* Continue data cleaning and pre-processing
* Begin exploratory data analysis
* Refine methodology and algorithms for data analysis

Week 3 (4/10-4/16):

* Complete exploratory data analysis
* Begin developing data visualization techniques
* Test and refine algorithms for data analysis

Week 4 (4/17-4/23):

* Finalize data visualization techniques
* Continue testing and refining algorithms for data analysis
* Begin working on the final project notebook

Week 5 (4/24-4/30):

* Complete the final project notebook
* Review and refine project results
* Begin drafting the peer review document

Week 6 (5/1-5/7):

* Finalize the peer review document and the project log report
* Review and edit the final project package
* Submit the final project package on or before the due date (May 8, 2023)

**Project References**

Union of Concerned Scientists. (2017, November 15). *Coal Power Impacts*. Union of Concerned Scientists. Retrieved March 22, 2023, from https://www.ucsusa.org/resources/coal-power-impacts

U.S. Energy Information Administration. (2023, March 7). *Short-Term Energy Outlook: Coal*. Short-Term Energy Outlook - U.S. Energy Information Administration (EIA). Retrieved March 22, 2023, from https://www.eia.gov/outlooks/steo/report/coal.php

United States Environmental Protection Agency. (2023, February 16). Power Plants and Neighboring Communities. Retrieved March 22, 2023, from https://www.epa.gov/airmarkets/power-plants-and-neighboring-communities#additional

University of Wisconsin Population Health Institute. (2022). *Explore health rankings: Rankings Data & Documentation*. County Health Rankings & Roadmaps. Retrieved March 22, 2023, from https://www.countyhealthrankings.org/explore-health-rankings/rankings-data-documentation