**Deliverable 4 – Descriptive Analytics Report**

**Introduction**

This report aims to provide descriptive and exploratory data analysis on the CDC dataset titled "Provisional COVID-19 Deaths by Sex and Age." The dataset contains information about COVID-19-related deaths categorized by sex and age, providing a valuable resource for understanding the demographic patterns of COVID-19 fatalities.

**Descriptive Analysis**

**Data Characteristics**

The dataset comprises several key attributes, including age group, sex, number of deaths, and week ending date. It covers a wide range of age groups and includes both male and female categories.

**Descriptive Statistics**

* **Frequencies and Distributions:** The dataset offers a comprehensive count of COVID-19 deaths categorized by age group and sex, allowing us to discern patterns in mortality rates across different demographics.
* **Central Tendency:** Calculating the mean, median, and mode of deaths provides insights into the typical number of fatalities within each demographic category.
* **Spread:** Assessing the range, quartiles, variance, or standard deviation helps understand the variability in death counts within age and sex groups.

**Visualizations**

To effectively present the descriptive analytics results, the following visualizations are utilized:

* **Bar Charts:** Bar charts illustrate the distribution of COVID-19 deaths across different age groups and sexes, providing a clear visual representation of mortality patterns.
* **Box Plots:** Box plots help visualize the central tendency and spread of death counts within each demographic category, enabling comparison between groups.
* **Scatter Plots:** Scatter plots may be employed to explore potential correlations between age, sex, and COVID-19 mortality.

**Prioritization of Attributes**

Based on the problem context, attributes such as age group and sex are prioritized due to their direct relevance to understanding demographic patterns of COVID-19 deaths. Further prioritization may be informed by relevant literature on factors influencing COVID-19 mortality, such as underlying health conditions and socioeconomic status.

**Data Exploration**

**Project Goals**

Project goals formulated in relevance to the problem context may include:

* Identifying demographic factors associated with higher COVID-19 mortality rates.
* Exploring temporal trends in COVID-19 fatalities.
* Investigating disparities in mortality outcomes across different population groups.

**Promising Attributes**

Attributes such as age group and sex are deemed promising for further analysis, as they are directly linked to demographic patterns of COVID-19 deaths. Additionally, attributes related to geographic location or comorbidities could be explored for deeper insights.

**Insights from Exploratory Graphics**

Exploratory graphics, such as histograms and scatter plots, provide valuable insights into the distribution and relationships between key attributes. For example, scatter plots may reveal any correlations between age, sex, and COVID-19 mortality, while histograms can showcase the distribution of deaths across age groups.

**New Characteristics and Interactions**

Exploratory data analysis may uncover new characteristics of the data, such as emerging trends or disparities in mortality rates across demographics. Additionally, interactions between data attributes, such as the impact of age and sex on COVID-19 mortality outcomes, may be identified and explored.

**Subset Identification**

Subsets of data may be identified for in-depth analysis based on specific criteria, such as age groups with disproportionately high mortality rates or regions experiencing surges in COVID-19 cases. These subsets enable focused investigations into particular aspects of the data.

**Changes in Project Goals**

Exploratory data analysis can influence project goals by revealing new insights or emphasizing certain aspects of the problem. For instance, if exploratory analysis identifies age as a significant predictor of COVID-19 mortality, project goals may shift to prioritize understanding the age-related factors contributing to mortality outcomes.

**Conclusion**

Descriptive and exploratory data analysis provide valuable insights into the demographic patterns and trends of COVID-19 deaths. By leveraging descriptive statistics and visualizations, as well as conducting thorough exploratory analysis, researchers can gain a comprehensive understanding of the factors influencing COVID-19 mortality rates and inform targeted interventions and public health strategies.