Final Project Summary: COVID-19 Death Analysis (2020–2023)

The purpose of this project was to analyze COVID-19 death data to better understand how the pandemic affected different groups of people in the United States. In the dataset of 2020-2023 COVID-19 death data from the CDC, each row represented total COVID-19 deaths grouped by age range, gender, month, year, and state. This analysis was completed in Python using a Jupyter Notebook, with pandas used for cleaning and organizing the data and matplotlib used to create graphs. Rows with missing death counts were removed to keep results accurate, while missing background details were left alone.

I created 5 different questions to lead the analysis. These questions included whether COVID-19 deaths differed between men and women across age groups, how death counts changed over time, which age groups made up the largest share of deaths, and which states were impacted the most. I displayed the results as plots, which gave insights that helped to answer each question. One of the findings was that COVID-19 deaths increased significantly with age, with older groups experiencing far more deaths than younger ones. Additionally, across most age groups, males had more deaths than females. Deaths also jumped up during large outbreak periods, and specific states accounted for a large portion of total deaths nationwide.

If I were to present this information to a client, I would recommend focusing on older populations and states that were more heavily impacted by COVID-19. If I continued analyzing this data, I would want to compare death counts by population size and explore how factors like vaccination rates affected outcomes.