COVID-19 Deaths by Age, Sex, and State

Final Data Analysis Project

CDC Provisional COVID-19 Death Data

By: Caitlyn Nunez-Nole

Purpose & Audience

Purpose of the Analysis:
- Analyze how COVID-19 deaths differed by age, sex, time, and state

- Identify groups most impacted

Audience:
- Public health agencies, policymakers

Data & Methodology

Source: CDC provisional COVID-19 death data (2020–2023, CSV) via Kaggle

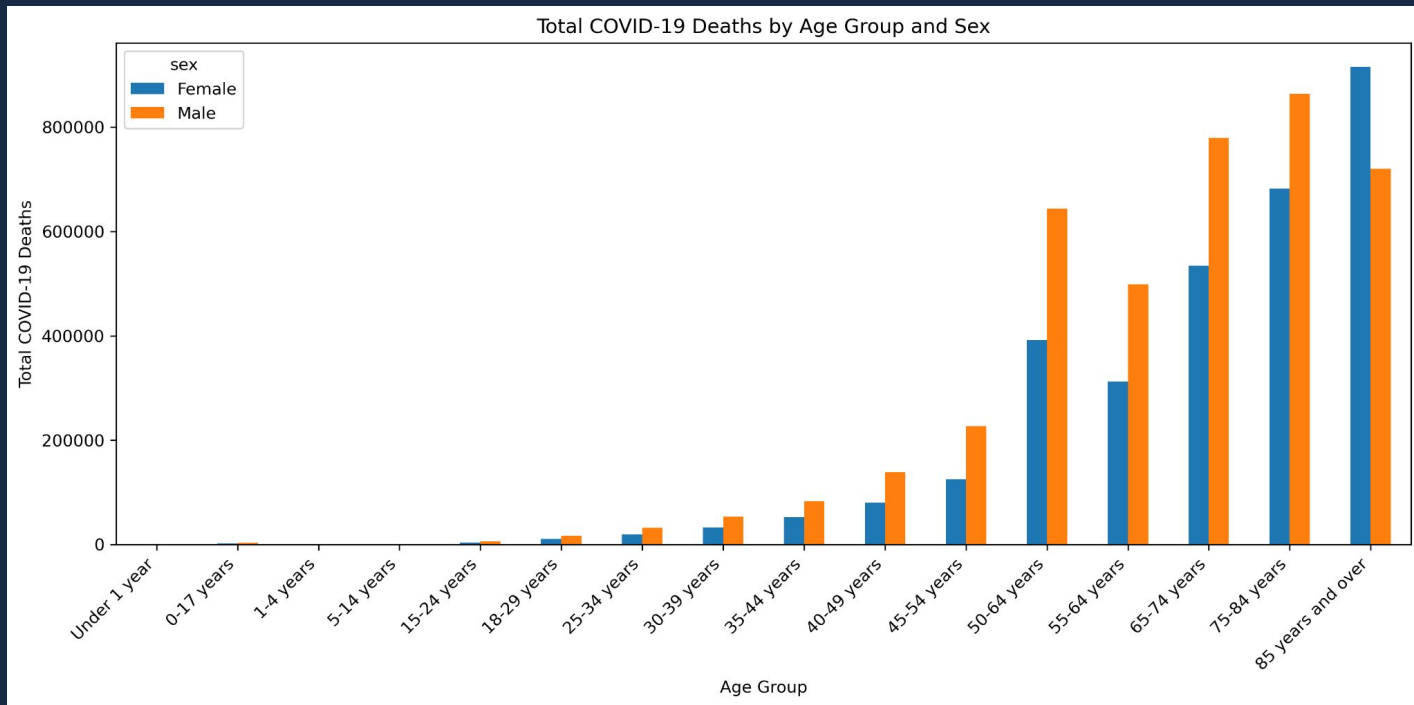Unit of analysis: death counts by age, sex, month, and state

Tools: Python, Pandas, Matplotlib, Jupyter Notebook

Question 1

How did provisional COVID-19 deaths differ between males and females across age groups?

Analysis:
COVID-19 deaths increased with age for both males and females. Across nearly all age groups, males experienced higher death counts, especially in older age categories.
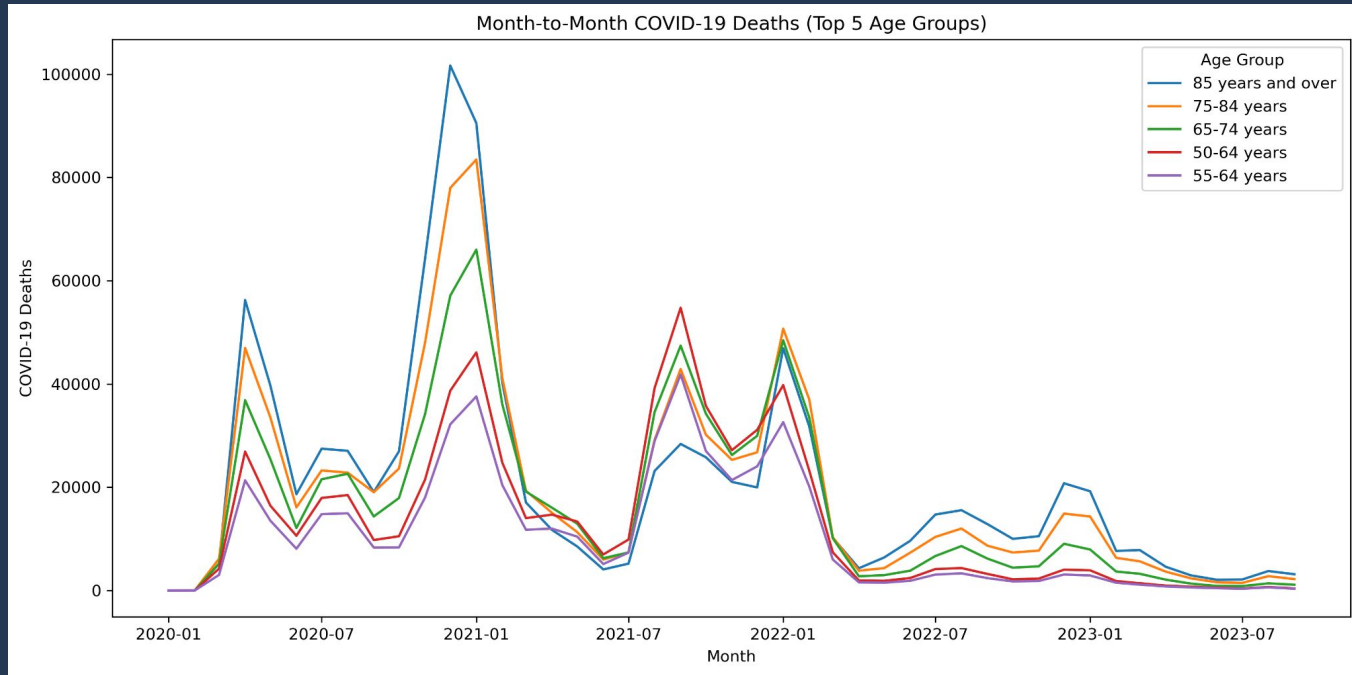

Total COVID-19 Deaths by Age Group and Sex

Question 2

How did COVID-19 deaths change month-to-month within age groups?

Analysis:
COVID-19 deaths went up and down over time, with noticeable spikes during major parts of the pandemic. Older age groups consistently had higher death counts than younger groups.
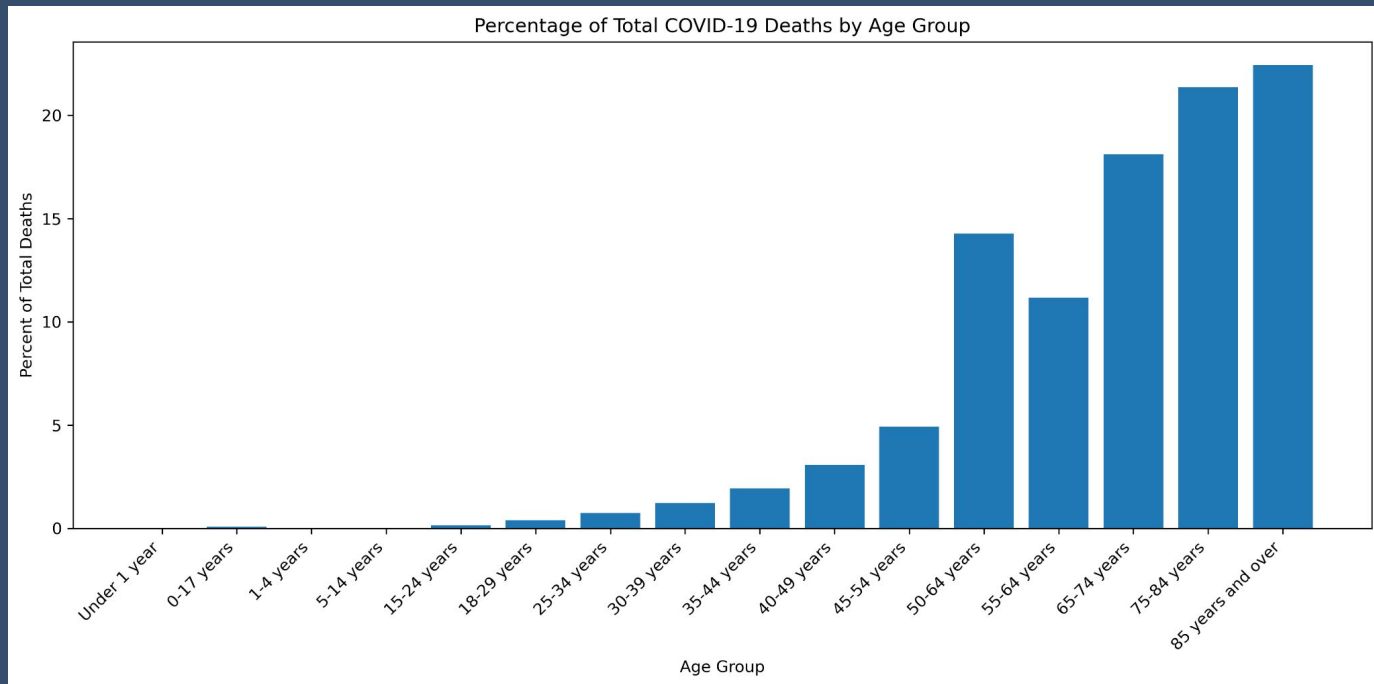

Month-to-Month COVID-19 Deaths (Top 5 Age Groups)

Question 3

Which age groups account for the largest share of total COVID-19 deaths?

Analysis:
Older age groups made up a much larger share of total COVID-19 deaths, while younger age groups accounted for a smaller percentage. This shows that COVID-19 had a greater impact on older people.
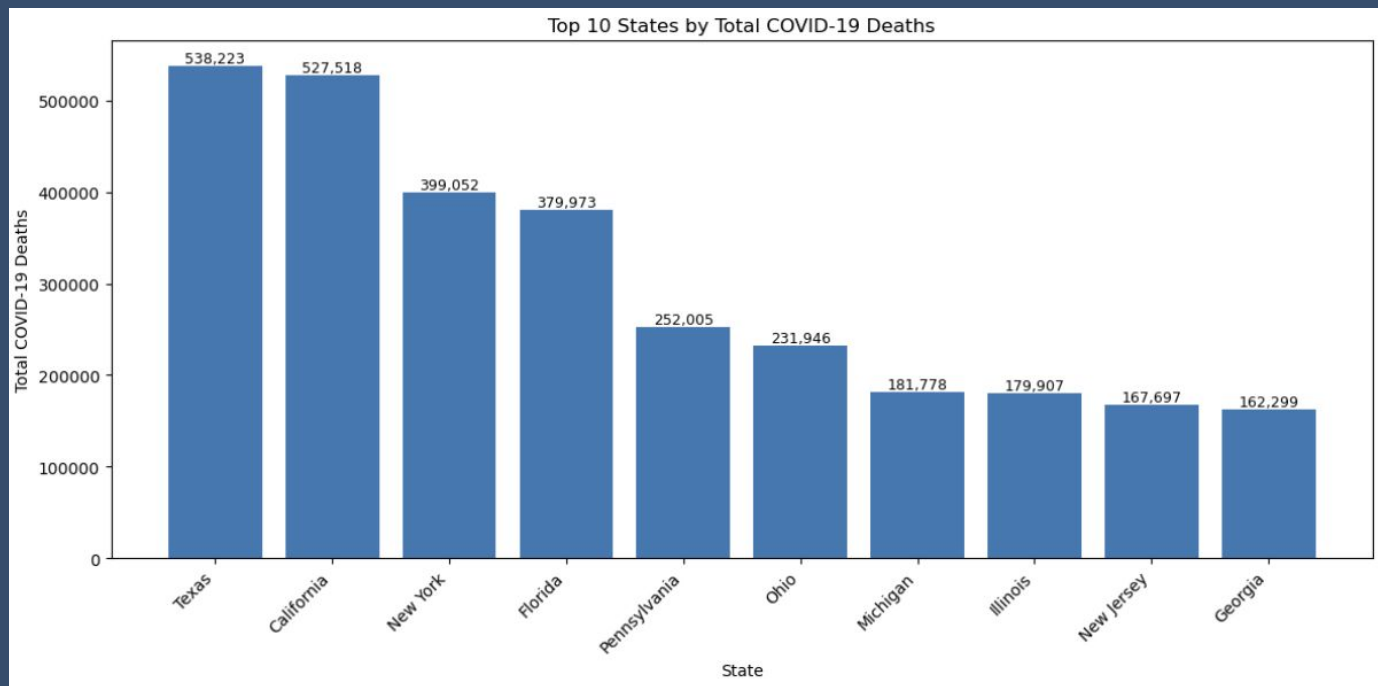

Percentage of Total COVID-19 Deaths by Age Group

Question 4

Which states experienced the highest total number of COVID-19 deaths?

Analysis:
The data shows that a small number of states had the highest total COVID-19 death counts. This suggests that COVID-19 affected some states more heavily than others.

Top 10 States by Total COVID-19 Deaths

Future Work

What new questions could be explored?:

- How do COVID-19 death rates compare when adjusted for population size?
- How did vaccination rollout affect death trends over time?

What still needs to be answered?:

- Why were some states impacted more than others?
- How did access to healthcare influence outcomes?

Appendix

- NYC is reported separately to CDC

- NYC has its own health department so its data is tracked independently

Checking values

```
df["state"].unique()
```

```
array(['United States', 'Alabama', 'Alaska', 'Arizona', 'Arkansas',
       'California', 'Colorado', 'Connecticut', 'Delaware',
       'District of Columbia', 'Florida', 'Georgia', 'Hawaii', 'Idaho',
       'Illinois', 'Indiana', 'Iowa', 'Kansas', 'Kentucky', 'Louisiana',
       'Maine', 'Maryland', 'Massachusetts', 'Michigan', 'Minnesota',
       'Mississippi', 'Missouri', 'Montana', 'Nebraska', 'Nevada',
       'New Hampshire', 'New Jersey', 'New Mexico', 'New York',
       'New York City', 'North Carolina', 'North Dakota', 'Ohio',
       'Oklahoma', 'Oregon', 'Pennsylvania', 'Rhode Island',
       'South Carolina', 'South Dakota', 'Tennessee', 'Texas', 'Utah',
       'Vermont', 'Virginia', 'Washington', 'West Virginia', 'Wisconsin',
       'Wyoming', 'Puerto Rico'], dtype=object)
```
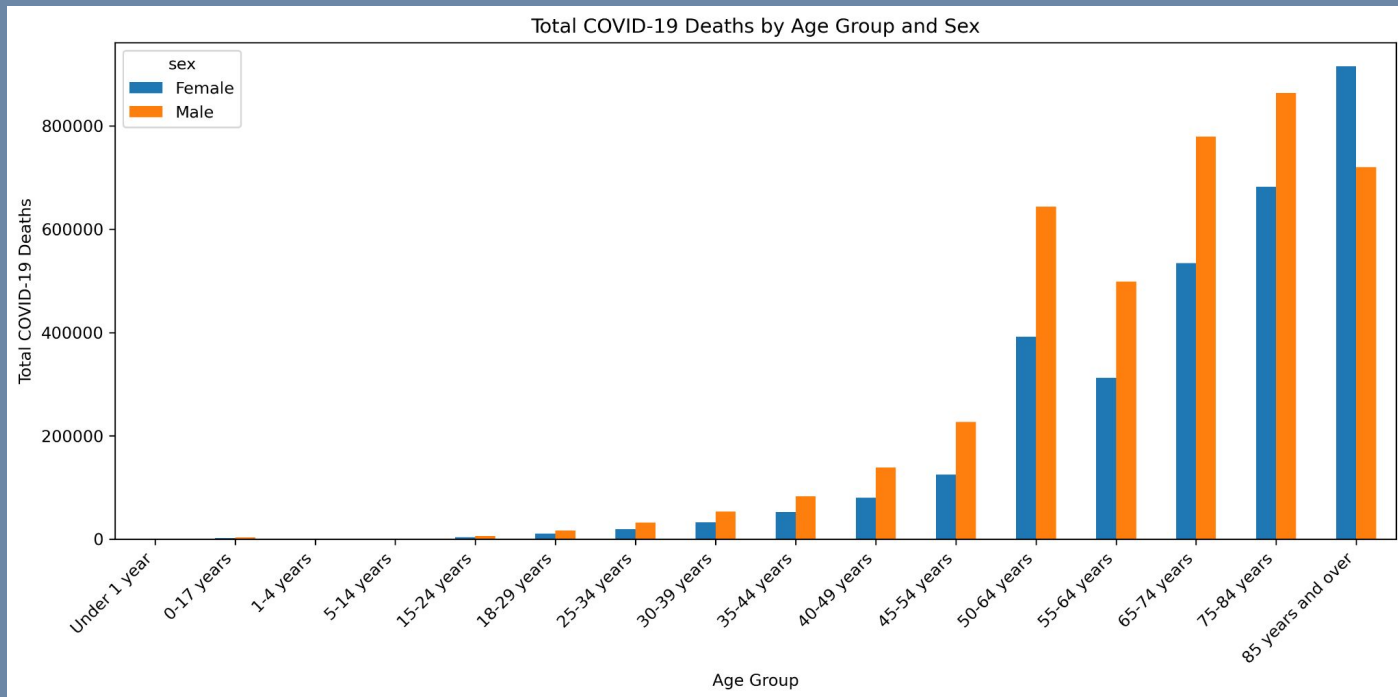
Merging New York City into New York State

```
df["state"] = df["state"].replace("New York City", "New York")
```

Question 5

Which age and sex groups account for the highest number of COVID-19 deaths?

Analysis:
When age and sex are looked at together, older males had the highest number of COVID-19 deaths. This shows that risk increases when multiple factors are combined.

Total COVID-19 Deaths by Age Group and Sex

Questions ?

Thank you!