**Project 2: COVID-19 Data Analysis**

* Analyze main conditions related to death from COVID-19.
* Track the time line of COVID-19-related deaths in the US
* Calculate the net deaths from COVID-19

*The data used for this step might not be very accurate, so the result might be disputable. But we can use the example to learn the methodology. The experimental design here is similar to A/B testing, which is commonly used in marketing research.*

**data source:**

1) COVID-19-related deaths

<https://data.cdc.gov/NCHS/Conditions-Contributing-to-COVID-19-Deaths-by-Stat/hk9y-quqm>

Click on Export and then click on

AH Monthly Provisional Counts of Deaths by Age Group and HHS region for Select Causes of Death, 2019-2021

2) Deaths from all causes

<https://data.cdc.gov/NCHS/AH-Monthly-Provisional-Counts-of-Deaths-by-Age-Gro/ezfr-g6hf>

HHS: Department of Health and Human Services

3) US population by age group

<https://en.wikipedia.org/wiki/Demographics_of_the_United_States>

Table

Description automatically generated

3) Shape files for geo-mapping

<https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html>

Graphical user interface, text, application

Description automatically generated