**Final Project:**

Research Question: Are COVID-19 vaccination rates associated with fewer cases and deaths (excess mortality)?

Two data sources (minimum):

1. **Our World in Data:** <https://ourworldindata.org/covid-vaccinations>
2. **World Bank Data:** <https://data.worldbank.org/>

**Basic Data Preparation Instructions:**

1) Use the covid vaccine data from OneDrive (provided by the professor): owid\_covid\_data\_final\_project\_Nov42021.xlsx

2) Before importing to SAS, rename any variables over 32 characters in length (Using Excel)

3) Drop any record that has missing, zero, or negative values for people\_fully\_vaccinated\_per\_hundred or new\_cases\_per\_million. Note: Ignore missing values for excess mortality for now (we will merge in new data later).

4) Create a month and year variable using month=month(date) and year=year(date). Next, keep only data from September and October of 2021.

5) For key predictor and dependent variables, create an average for September and October of 2021.

6) Print out the average for all variables to be used in our analyses (check for missing, errors etc.)

7) Run a proc corr for all variables that will be used in regressions later.

8) Merge the 5 control-variable files that were chosen by the professor into the main covid\_vaccine data. Note: These data include an update of the excess mortality data that were uploaded by the professor. You should drop the original excess mortality data from the main file (too many missing).

9) Run regressions to predict new\_cases\_per\_million using people\_fully\_vaccinated\_per\_hundred as the only predictor. Repeat this step to predict excess\_mortality.

**Reporting of Results:**

Overview: You will create 3 tables and write a paragraph about the meanings of each table.

1. Create table of “Summary or Descriptive Statistics” (e.g. Mean, SD, or Frequency N) for all variables utilized.
2. Create a table for both regressions (explaining new cases, excess mortality)
3. Write a paragraph that explains the meanings of each table

**Examples of the Reported Results (with example tables) below:**

**Title: An Examination of the Relationships between Vaccination, COVID-19 Case Rates and Excess Mortality**

Research Question: Are COVID-19 vaccination rates associated with fewer cases and deaths (excess mortality)? Using data that were obtained from Our World in Data and the Worldbank International Database, the purpose of this study was to determine whether vaccination rates significant reduced new COVID-19 cases or excess mortality in approximately 165 nations during fall, 2021. Two linear regressions (Tables 2 and 3 below) were conducted while including controls for national wealth (gross domestic product), medical expenditures, etc. etc.

Table 1. Descriptive Statistics.

Variable Mean Standard Deviation N

New Case Rate (per hundred) 61.1 22.2 165

Fully Vaccinated Rate (per hundred) 55.5 12.4 165

Gross Domestic Product 1,234,555.12 222.3 165

Excess Mortality Rate 123.33 44.1 77

Include all other relevant variables

The descriptive statistics in Table 1 indicate that the average case rate across all nations during September and October of 2021 was 61.1. The average full vaccination rate was 55.5, indicating that the majority of adults among the 165 nations were fully vaccinated by the end of October, 2021. The average gross domestic product was xyz etc. etc. [Note: Were any of these descriptive statistics unusual or higher/lower than expected? If so, you should mention this].

Table 2. Regression: New Case Rate as Predicted by Full Vaccination Rate During Fall, 2021.

Variable Coef. S.E. Sig.

Fully vaccinated rate (per hundred) 1.99 0.09 .001

Gross Domestic Product -1.22 0.003 .02

Amount Spent on Medical -3.44 1.10 .65

Include all other relevant variables 11.11 1.11 .999

N=165 R2 = 0.05

Contrary to expectations, the regression coefficients from Table 2 indicate that that countries with higher vaccination rates experienced higher overall new cases of COVID-19. [Interpret the vaccination rate coefficient 1.99 here]. Additionally, nations with greater overall wealth (measured via GDP) had lower new case rates [Interpret the GDP coefficient -1.22 here]. Similarly, nations that invested more in medical care also experienced lower new COVID case rates [Interpret -3.44 here]. Explain all control variable coefficients etc. etc. Nevertheless, these cross-sectional regressions cannot scientifically demonstrate cause-effect relationships. Consequently, the results in Table 2 could also be interpreted as indicating nations with high COVID-19 exposure were more likely to encourage vaccination.

- Continued -

Table 3. Regression: Excess Mortality Rate as Predicted by Full Vaccination Rate During Fall, 2021.

Variable Coef. S.E. Sig.

Fully vaccinated rate (per hundred) -0.99 0.09 .001

Gross Domestic Product -1.22 0.003 .02

Amount Spent on Medical -3.44 1.10 .65

Include all other relevant variables 11.11 1.11 .999

N=77 R2 = 0.05

Regression coefficients from Table 3 indicate that nations with higher overall vaccination rates experienced fewer excess deaths. [Interpret the vaccine coefficient effect on excess mortality here]. Explain the coefficients for all remaining control variables here.

Conclusions: Write and briefly explain your answers to the two research questions here.

Did higher vaccination rates reduce COVID-19 cases? (yes, no, unclear).

Did higher vaccination rates reduce excess mortality? (yes, no, unclear).

Submitting your work to blackboard:

1. Submit a sample of your final fully merge data (about 10 to 20 observations) in the first link under the Final Project folder on Blackboard.
2. Next, submit all of the SAS code that you used to prepare, merge, and analyze the final files (including your final regression statements).