# Problem Set 4

## Due 23 March

**Report on COVID-19, mask use, and vaccination**

You are hired to deliver a report to a client on the COVID-19 pandemic, mask use, and vaccinations in counties across the United States. Your data team *promised* to deliver coded data and completed visualizations. Let's just say they accomplished some of that. They send you the code to generate and clean the data along with the start of the visualizations and tables. It falls to you to pick up the pieces and complete the report.

Organize and number the body of the report as follows:

1. Summary: short overview of the report, including purpose and central findings
2. Data: describe data sources and summarize data for key variables
   - COVID deaths
   - Percentage "always" masking
   - Rates of vaccination
3. Analysis: Estimated impact of mask use and vaccination on 2022 COVID deaths

You'll need to use their baseline code (`PS4start.R`) as well as the CDC vaccination data (`cdc vax mar1.csv`). Use the same code to compile/clean the data. Then use the rest as a starting point for the code chunks in your markdown script. You may, but are not asked to, add graphs or estimates beyond what you find in the base code. Obviously, you will need to improve what's there.

Find further details on the county-level data sets for this report from the NYT Coronavirus Data in the United States (https://github.com/nytimes/covid-19-data) and the CDC COVID-19 Vaccinations in the United States, County reports (https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh/data).
Note that a 5-digit FIPS code uniquely identifies US counties.

## What to submit

Use R Markdown to compose the report described above. Write clearly, and format your tables and visuals with careful attention and professionalism. Knit to `.pdf`, and post all materials in a github repository by March 23. NOTE: *hide your code chunks in the main body.* Move all code to the end in an appendix.