HONR 490 – Homework 3

Justin A. Gould
gould29@purdue.edu

February 24, 2021

DUE DATE: 2021/09/12 23:59 EDT

Homework Instructions
To receive credit for the assignment, do the following:

1. Create an `.ipynb` file, and name it: `purduealias_honr490_homework_number.ipynb` (e.g., `gould29_honr490_homework_1.ipynb`)

2. Show all your work and follow the instructions below very carefully.

3. Submit a printout (e.g., as PDF) of your `.ipynb` file–and the file itself–to Brightspace by the due date.

4. You must show **all** your work and provide comments in your code explaining what you are doing.
(e.g., `#Adding title to map` or `#Load data`)

5. This homework assignment will have a non-code question (question 2). To receive credit for this question, please follow the question's instructions and submit a PDF of the requirements to Brightspace by the due date.

For grading this assignment, I will not leverage unit tests. I will look at your `.ipynb` file to grade question 1 and the PDF you submit for question 2. When in doubt, please show and comment all your work.

All files and data you need to complete this assignment are located on GitHub, under the Homework 3 folder. It contains the following:

- `UScounties`: A folder containing a shapefile (and supporting files) for the `POLYGON` of every U.S. county.
- `covid_us_counties.csv`: A CSV of county-level COVID-19-related data from January 2020 through February 2021.
- `co-est2019-alldata.csv`: A CSV of the population of every U.S. county, organized by FIPS code.

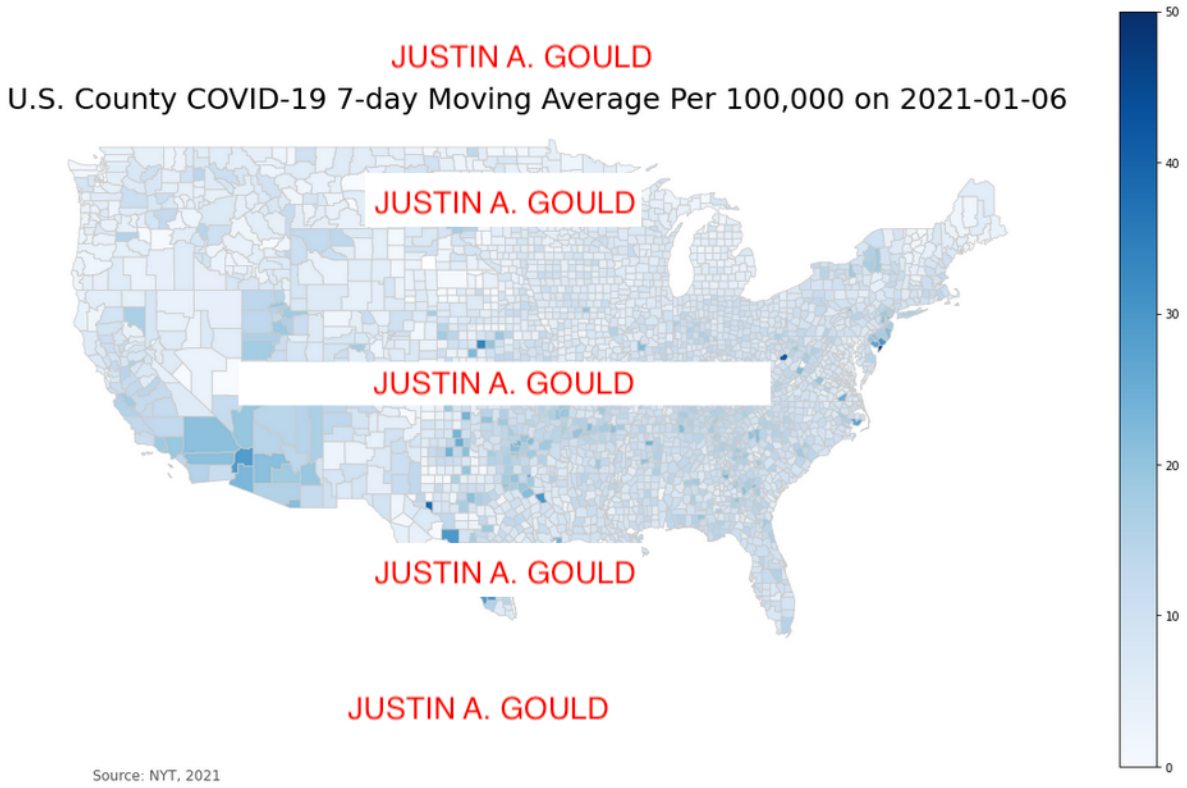

Figure 1: Example map output

Data Overview and Task Requirements
Your task is to create a choropleth map of COVID-19 case data for the domestic United States. Specifically, you're tasked with calculating and mapping the 7-day moving average of new cases per 100,000 residents on November 03, 2020.

HINT: I suggest working in Python first, as we did in class, to do the preprocessing, before moving to QGIS for the second question.

Your final map should look something like Figure 1.

Preprocessing Steps:

- Remove counties from Hawaii and Alaska from the shapefile.
- Calculate daily case increases at the FIPS code level.
- Calculate 7-day moving average of daily increase in cases at the FIPS code level...**HINT**: There is a built-in function in the `pandas` package you can use to do this!
- `JOIN` the following data into one Pandas DataFrame: FIPS geometry, NYT COVID-19 case data (your calculations), FIPS population data.
- Finally, be sure to filter the final DataFrame to show only the target date.

Problem 1

Visualizing a COVID Map via Python – 15 points
Create a choropleth map of 7-day average of new cases per 100,000 residents on November 03, 2020, via `pandas`, `geopandas`, and `matplotlib`.

Input: The data and files described on page 2.

Desired Output: Please submit your `.ipynb` and image of the map (either as `.png` or `.pdf`). Your notebook **must** show the `GeoDataFrame` used to generate the map.

Grading Criteria:

- You will receive 5 points for correctly generating the map.
- You will receive 10 points for showing the `GeoDataFrame` with the correct values (i.e., your preprocessing and moving average calculation were done correctly).

Problem 2

Visualizing a COVID Map via QGIS – 10 points
Create a choropleth map of 7-day average of new cases per 100,000 residents on November 03, 2020, via `QGIS`.

Input: The data and files described on page 2.

Desired Output: Please submit your QGIS project as a `.qgz` file and image of the map (either as `.png` or `.pdf`).