U6614: Assignment 2: COVID-19 Country Case Data

Your Name (your-uni)

2021-01-18

Please submit your knitted .pdf file along with the corresponding R markdown (.rmd) via Courseworks by 11:59pm on Monday, January 25th.

Before knitting your rmd file as a pdf, you will need to install TinyTex for Latex distribution by running the following code:

```
tinytex::install_tinytex()
```

Please visit this link for more information on TinyTex installation.

If you run intro problems knitting as a pdf, please test by first knitting as an html file. If your rmd file knits as an html file but not as a pdf, then it is likely an RStudio/TinyTex installation issue. A quick fix is to open your knit html file in Chrome or your preferred browser, print as a pdf and submit that file to Courseworks.

If you are unable to knit as an html file, then the issue lies with your code and not your RStudio setup.

Introduction

Load packages:

```
#load packages here
```

1 Load and prep the data

Load the COVID-19 data from class and only keep observations without NA values for cases (not deaths). Data source: https://www.ecdc.europa.eu/en/publications-data/data-national-14-day-notification-rate-covid-19

[FOR EACH QUESTION, ADD YOUR WRITE-UP (IF APPLICABLE) UNDER THE QUESTIONS/HEADERS FOR ORGANIZATION]

```
#include any code you use to arrive at your answers as code chunks
#but don't include code you don't need to arrive at your answers!
#remember to use comments liberally to explain/organize your code
```

In your written answers, use inline code chunks to refer to any numbers from your analysis rather than hard-coding numbers. See Lecture2.2, Section 5.3 for an overview of in-line code chunks.

2 Describe the data

Provide the following, along with any other information you think might be useful for the reader to know about the data.

- *unit of observation*
- *date range observed in the data*
- *number of countries (or administrative entities reporting data)*

3 Latest global 14-day case rate (per 100,000 population)

a. Create a new data frame that only includes observations for the most recent week.

Note: don't hard-code a date to filter on, find the latest week, store as a data object, and then refer back to (the element in) that object (see Lecture2-inclass.r for guidance)

b. What was the max 14-day rate for the most recent week observed in the data?

c. List the top 10 countries by 14-day rate for the most recent week observed in the data?

d. How many countries had zero reported cases for the most recent week?

4 Panama 14-day case rates

a. Create a new data frame for the 14-day rate for Panama only. Sort in descending data order.

b. Find the weekly mean, min, and max 14-day rate for Panama over all included data and name each column appropriately.

c. What was the average 14-day rate in Panama over the *last* 10 weeks of reported data?

[HINT: See Lecture2.1 -> Section 4.2 for examples of subsetting syntax that can help you refer to the first 10 rows of sorted data. If you're having trouble, you can also try using the `row_number()` *function]*

d. What was the average 14-day rate in Panama over the *first* 10 weeks of reported data?