U6614: Assignment 2: COVID-19 Country Case Data

Your Name (your-uni)

2020-09-15

Please submit your knitted .pdf file along with the corresponding R markdown (.rmd) via Courseworks by 11:59pm on Monday, September 21st.

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.

Load and prep the data

Load the coronavirus.rda data from class and only keep confirmed cases. Data source: https://github.com/RamiKrispin/coronavirus/tree/master/data

[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
#remember to use comments liberally to explain/organize your code
```

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)

Latest global case counts

a. Create a new data frame that only includes observations for the most recent day only. *Note: don't hard-code a date to filter on, find the last day, store as a data object, and then refer back to (the element in) that object (see Lecture2-inclass.r for guidance)*

b. What was max case count for the most recent day observed in the data?

c. List the top 5 countries (or administrative entities) by case count for the most recent day observed in the data?

d. How many countries (or administrative entities) had zero confirmed cases for the most recent day?

Oman case counts

a. Create a new data frame for daily confirmed case counts for Oman only. Sort in descending data order.

b. Find the daily mean, min, and max case counts for Oman over the duration of the andemic and name each column appropriately.

c. What was the average daily case count in Oman over *last* 30 days of reported data? *[HINT: See Lecture2.1 -> Section 4.2 for examples of subsetting syntax that can help you refer to the first 30 rows of sorted data. If you're having trouble, you can also try using the* `row_number()` *function]*

d. What was the average daily case count in Oman over the *first* 30 days of reported data?