Introduction to R for Applied Epidemiology

Code review checklist

Date revised: 2022-03-23



# Overview

In the final module of this course, you are asked to create an original R Markdown script to produce a situation report for the COVID-19 epidemic in Fulton County, Georgia, USA. After producing this report, you are asked to review the code of a partner.

* Provide **constructive feedback** on the content, organization, readability, output, and syntax.
* Write your comments below each bullet point, **in a different color (e.g. Red)**.

**THIS EXERCISE IS NOT A TEST** It is simply an opportunity to try applying your new skills.

# Output

* Are code chunks, warnings, and messages hidden from the output?
* Is there a title, subtitle, and date at the top of the report? (produced by the YAML)

# Script organization

* Is the first code chunk the setup chunk that sets echo = FALSE and other R Markdown settings?
* Is the first substantial R code chunk loading packages with {pacman}?

# Readability

* Are the purposes of the R packages int he {pacman} command well documented with comments?
* Are there # comments within the code chunks to describe the sections of code?
* Do # comments use spaces well, resulting in clear and readable comments?
* Is the R code generally “vertical” more than “horizontal” i.e. does the author use new lines and indentations successfully to improve readability?
* Is there sufficient white space between code so the script does not look “squished” or compressed?
* Is it easy to see where each section begins and ends?
* Are the author’s name and contact information documented somewhere near the top of the script?

# Inline code

* Does the “Summary” section at the top contain inline code that successfully produces a data date (June 23 2021), the total number of confirmed cases, the number of deaths (and proportion), and the number that required hospitalization?

# Syntax

* Are functions correctly written adjacent to their parentheses, like mutate() and not mutate ()?
* Do object names (names defined by the user) use short, descriptive names? are they lowercase and use underscores if necessary? e.g. linelist or joined\_data or city\_table.

# Age/Sex pyramid

* Is there a “Demographics” header?
* Does the pyramid show proportions? (not counts)
* Are the X, Y, and Legend labels adjusted to display with title case? (e.g. “Gender”, “Age group”, and “Percent of total”)
* Is there a title? Is there a caption?
* Are 10 year age groups used?
* Do the colors match the template? (green/purple)

# Epidemic curve

* Is there a header for the Weekly Epidemic Curve by City?
* Does the epicurve use date of report?
* Is the color palatte matching the template, or otherwise pleasant to see?
* Are the labels, title, and subtitle written well, to match the Template?
* Is there a caption that produces the current number of cases, and the maximum and minimum dates of report?
* Is the epicurve aligned at the bottom, flush with the axis ticks?
* Are the date labels on the X-axis efficient, informative, and pleasant to read?
* Are the y-axis break at appropriate intervals?

# Case incidence table

* Is the table an image (flextable) or editable?
* Is there one row for each City?
* Is there a column for total cases?
* Is there a column showing the number of cases in the last 14 days?
* Is there a column showing cumulative case incidence per 10,000 population?
* Are the rows ordered by cumulative incidence?
* Is there a total row at the bottom?

# Race and outcome table

* Is there one row per racial group?
* Is the table results stratified by outcome (Hospitalisation is Yes/No)?
* Is there a p-value?