Capstone Data Analysis Project – Data Exploration
Biostatistics

Name: __Kate Gardner__   Score = __22.5___ /25

GitHub repo: __https://github.com/klgard/Gardner-Capstone-Project__

Submitted on time?   ⦿ Y   ◯ N

Project element	Value	Pts earned	Comments
Establish project • Create GitHub repository for project called "Lastname-Capstone-Project" • Create new RStudio project tied to git hub repo • Set up project with Code, Data folder	2	2	good. Consider deleting the README.md in the root folder of project since it has nothing in it.
Set up data folder • Add all csv files for project • Add ≥ 1 metadata .txt file • If > 1 csv, include README.txt to explain	3	2.5	good but I'd like to see something more clear in the metadata file indicating the names of the .csv files that relate to this project so it is clear what is original data vs. data that might be manipulated and saved with write.csv() later.
Set up .qmd file in Code folder • Check data for mistakes and outliers • Change any var names or create new variables	4	4	good
Exploratory Data Analysis- get to know your data • summary statistics • histograms • boxplots • group_by and summarise • etc.	6	5	I'd like to see you plot some histograms of your data as part of the process of getting a feel for the data. I think it might also be good to plot these values by date (not month, but actual date) to get a feel for the data, even if that isn't your actual analysis.
Save cleaned dataset(s) • Write code to save the cleaned, revised dataset in Data folder with clear name	2	2	good
Code is simple and clear and gives correct output • Replaces as much human intervention as possible • Provides correct summary values	2	2	good

Thought processes are well documented outside of code blocks, code is well commented, all steps prior to data analysis are finished	5	4	See notes in data-exploration-feedback.qmd
Save and commit your changes and push to github • Send link to repo on Canvas when finished	1	1	good

Additional feedback