Capstone Data Analysis Project – Data Exploration
Biostatistics

Name: August Kotula   Score = 22   /25

GitHub repo: https://github.com/akotu21/Kotula-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
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	good but I'd like to see a link to Peter Menzies on Kaggle...
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	4.5	I didn't see you check levels for some of your important factor variables. Also, I'd like you to think harder about how you could examine the number of observations/date data - think more creatively about some exploratory plots you could view. Finally, I'd like to see histograms of any continuous variables.
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	4.5	This is much better - keep working on providing a better break down of scientific plan (what you are about to do) -> code -> what you learn from code.
Save and commit your changes and push to github • Send link to repo on Canvas when finished	1	1	good

Additional feedback