Capstone Data Analysis Project – Data Exploration
Biostatistics

Name: Margaret Dougher Score = 22 /25

GitHub repo: https://github.com/margaretdougher14/Dougher-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	2.5	good job detailing what is in your data folder. Note that the file names in your folder are not exactly the same as what are listed in your README and Data Set Preparation.txt files.
Set up .qmd file in Code folder • Check data for mistakes and outliers • Change any var names or create new variables	4	4	your code is good and you worked through some good analysis - you didn't tell me what you learned, though (see below).
Exploratory Data Analysis- get to know your data • summary statistics • histograms • boxplots • group_by and summarise • etc.	6	5	I don't see many histograms at all of your data.
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	3.5	You reallly need to work to record your scientific thinking in your .qmd file, both before a code chunk to tell your plan and why you are doing it and after code chunks to tell yourself (and me) what you learned.
Save and commit your changes and push to github • Send link to repo on Canvas when finished	1	1	good.
Additional feedback			