

Capstone Data Analysis Project – Data Exploration
Biostatistics

Name: Emma O'Neil Score = 17.5 /25

GitHub repo: <https://github.com/eoneil214/ONeil-Capstone-Project>

Submitted on time? ☐ Y ☒ N

Project element	Value	Pts earned	Comments
Establish project <ul style="list-style-type: none"> 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	1	I had to go back and tell you to get the assignment set up with a code and data folder.
Set up data folder <ul style="list-style-type: none"> Add all csv files for project Add ≥ 1 metadata .txt file If > 1 csv, include README.txt to explain 	3	3	Please rename the data preparation metaFILE .txt to data_preparation_metadata.txt
Set up .qmd file in Code folder <ul style="list-style-type: none"> Check data for mistakes and outliers Change any var names or create new variables 	4	3	Please don't have 2 copies of the Data-exploration.qmd (get rid of the one in your project root directory and keep the one in your code folder). You did not check for mistakes in factor variables.
Exploratory Data Analysis- get to know your data <ul style="list-style-type: none"> summary statistics histograms boxplots group_by and summarise etc. 	6	4	<p>You should have a histogram for your quantitative variables as well as box plots as a means of looking at outliers.</p> <p>Your group_by and summarise should reflect one of the comparisons you wish to analyze and should include a plot of the results.</p>
Save cleaned dataset(s) <ul style="list-style-type: none"> Write code to save the cleaned, revised dataset in Data folder with clear name 	2	0	You didn't write a "cleaned" version of the data file in which you have checked/repared any issues with the data. I see that you didn't change anything, in which case you should have written in the .qmd that you didn't need to save a new copy.
Code is simple and clear and gives correct output <ul style="list-style-type: none"> Replaces as much human intervention as possible Provides correct summary values 	2	2	good

Capstone Data Analysis Project – Data Exploration
Biostatistics

Thought processes are well documented outside of code blocks, code is well commented, all steps prior to data analysis are finished	5	4	Please see data-exploration-feedback.qmd for more details.
Save and commit your changes and push to github <ul style="list-style-type: none">Send link to repo on Canvas when finished	1	0.5	I had to ask you to submit the link on canvas
<p>Additional feedback</p> <p>Please see data-exploration-feedback.qmd</p> <p>This assignment was 2.5 days late. I did not take off additional points, but should have. Please work harder to take initiative in communicating with me (though I know you haven't been feeling well) as I could have gotten you on the right track with re-setting git/github and your new laptop sooner.</p>			