

Final Grade Reflection

Michelle Kairy

Through my work in this class so far, I believe I have earned an A-.

This quarter, I turned in most all of my assignments complete and on time, and all of the assignments I have turned in demonstrate my proficiency in the learning targets.

When it comes to working with data, I showed proficiency in reading in, selecting, filtering (with simple filters and anti/semi_joins), mutating, and joining data. In my `data_import_practice` document I read a variety of data formats (CSV, xlsx, and txt) using the `here` function. I even used optional arguments for `read.csv` and `read_delim` to read the data cleanly (WD-1 and R-1). I demonstrated proficiency in selecting columns in my Lab 3 assignment in question 10, where I selected several columns simultaneously (WD-2). My submission for parts 9 and 10 of my Lab 3 assignment best reflect my ability to filter factor and numeric type data. In problem 9, I filtered by numeric variables using the `>`, `==` and `!=` comparison operators, and used the `%in%` operator in part 10. For date types, I filtered by date in Lab 5 question 3.1 (WD-3). I showed that I could modify data for various data types and purposes in Lab 3 question 4, where I used the `mutate` function to convert character types to numeric and factor types. In question 7 of Lab 3, I used the `ifelse` function within the `mutate` function to create binary variables (WD-4). Finally, I showed that I could filter and join data frames in the "Data Cleaning" portion of my Challenge 4 assignment. I used a `semi_join` to filter and then used multiple `full_joins` to combine four data sets into one (WD-5 and WD-6).

In addition to my skills in working with data, I've been able to create reproducible and readable data. I showed that I could use the `here` function to load data into a quarto document in question 1 of my Lab 2 assignment, which worked because I completed it within my R project (R-1). My code is also tidy, as throughout Lab 5, I use proper white space, line breaks, and consistent operators (R-2). Finally, my code is reproducible because I don't hardcode or use relative references that make changing inputs more difficult. I show this skill in Lab 3 question 7, where I use the names of columns, rather than their position (R-3).

Throughout this course, I've created visualizations for categorical and continuous variables. In Lab 2 question 4, I used a scatter plot to represent the relationship between two continuous variables. In question 8, I used boxplots, and in Lab 3 question 8.a, I used a bar chart to show

the counts of two categorical variables (DVS-1). In creating these plots I've had to modify my axis labels, titles, and legends to make them easier for the viewer to understand. In Lab 4, I used the guides and `n.dodge` function and argument to use rows in my axis labels so they wouldn't run into each other and used `guide_legend` to change the legend title (DVS-2). I showed creativity in my plots by choosing different plot themes using `ggthemr` or other packages, as in in Lab 4, where I used the `dust` theme in `ggthemr`, and in Challenge 2, where I used the `viridis` package (DVS-3). I showed that I could create numerical summaries, even using the `across` function, in Lab 3 problem 7 (DVS-4 and DVS-5).

Lastly, I've proven my code is efficient throughout all of my assignments. Part 9.c of Lab 3 shows how I use the `filter` function to filter multiple variables at the same time, rather than calling it multiple times (PE-1). Using the `across` function, as I did in Lab 4 part 7, I reduce repetition by having the function perform the same operation on multiple columns simultaneously (PE-3). The functions I use are also modern, as I have been able to adapt to the `full_join` and `pivot_longer` functions as I did in Challenge 4 in the "Data Cleaning" portion (PE-4).

I've only been able to meet all of these learning requirements because I have taken every opportunity to revise my work and learn from it. In Lab 3 question 1, I got a "G" on my description of the data set, where I left out much of the important information about how and where the data was collected. After revising this question, I revised my answer for question 1 in Lab 4 to include as much relevant information about the data as possible. I have also showed revised thinking by including line breaks to keep my code from wrapping in all of my labs and challenges since the first lab.

Beyond revising the work I've already done, I always find opportunities to try out new functions, techniques, or edits to my HTML. I used different fonts and background colors in all of my labs starting with Lab 2. I added in-line R code to my Lab 5 and Lab 4 assignments, and I used `stringr` functions before we learned them officially. I also put extra time in my challenge problems, as I completed all of the challenges in Challenge 2. Perhaps what I am most proud of, though, is my Challenge 4 submission, where I used data from FRED that required me to use just about every data cleaning technique we have learned so far to create a unified, clean data set.

Aside from making sure *my* assignments are quality work, I help my group members do the same in each class. I am always present in class, and I come prepared, having done the preview activities, videos, and reading. Because I come with some knowledge of the problems we will be asked, I have been able to answer most of my group members' questions, along with helping to troubleshoot any errors or weird output. I also make myself readily available over text, answering questions outside of class when anyone in the group needs help.

Outside of my group, I help other students by putting time into my peer reviews. Even when one of my peers' lab looks almost perfect, I try to find some areas that could make their code easier to read or just offer a creative way to solve a problem.