

Data201 Diary

Author: Zhuoheng Chen

Team member: Zhuoheng Chen

04/10/2023:

Start finding database from the Internet and download it while trying to import the database into Jupyter lab.

05/10/2023:

Start doing some preliminary wrangling on the data. Such as weeding out useless data and trying to minimize databases.

07/10/2023:

I found that I could not log in mathvlinux22001p on the math portal. I tried to send an email to the IT department of the school, but there was no response because it was a weekend. I couldn't log in successfully until Tuesday (10/10/2023) of the following week since I went to the IT office. I used R-Studio to work in this period.

11/10/2023:

I found that the original database could not achieve the desired effect. After some searching and learning on the Internet, I decided to analyze from the perspective of survival rate, so I changed another database and start working on that one. Continued working on this until the 15th.

15/10/2023:

Finished all the coding work and started to make the slides for the presentation. Import code generated images and some key data into the slides, add some simple comments, write a simple script for the speech.

17/10/2023:

Finished off the group project presentation.

19/10/2023:

Do the final collating work, but I didn't clarify the logic of relational database, because I felt that it was more used for the connection of multiple databases, and I only have one database, so I am not sure how to proceed, but give it a goal at the end.