



[Course](#) > [Capstone Project: I...](#) > [Project Submission:...](#) > Choose Your Own ...

Choose Your Own Project Submission

CHOOSE YOUR OWN SUBMISSION

This assignment has several steps. In the first step, you'll provide a response to the prompt. The other steps appear below the **Your Response** field.

▼ **Your Response** due Jan 10, 2020 18:59 EST (in 4 months, 3 weeks) IN PROGRESS

Enter your response to the prompt. You can save your progress and return to complete your response at any time before the due date (Friday, Jan 10, 2020 18:59 EST). **After you submit your response, you cannot edit it.**

The prompt for this section

Your submission for this project is three files:

1. Your report in PDF format
2. Your report in Rmd format
3. A script in R format that performs a supervised machine learning task

To upload and submit your files press the "Choose Files" button, select three files at once (using the control key on a Windows machine or command key on a Mac) and press "Choose," type a description for each, and then press the "Upload files" button.

You must also provide access to your dataset (either via automatic download or inclusion in a GitHub repository). We recommend providing a link to a GitHub repository containing your dataset and the three files above.

Note that when downloading files for peer assessments, R and Rmd files will be downloaded as txt files by default.

Your response (optional)

Enter your response to the prompt above.

Preview in LaTeX

CLICK TO PREVIEW YOUR SUBMISSION IN LATEX.

Save your progress

THIS RESPONSE HAS NOT BEEN SAVED.

Choose Files No file chosen

Upload files

You may continue to work on your response until you submit it.

Submit your response and move to the next step

Assess Peers due Jan 17, 2020 18:59 EST (in 4 months, 4 weeks) NOT AVAILABLE

Staff Grade NOT AVAILABLE

▸ **Your Grade:** Not Started

Grading Rubric

Files (5 points possible)

The appropriate files are submitted in the correct formats: a report in both PDF and Rmd format and an R script in R format.

- 0 points: No files provided
- 3 points: At least one file is missing and/or not in the correct format

- 5 points: All 3 files were submitted in the requested formats

Report (25 points possible)

The report documents the analysis and presents the findings, along with supporting statistics and figures. The report must be written in English and uploaded. The report must include at least the following sections:

1. an introduction/overview/executive summary section that describes the dataset and summarizes the goal of the project and key steps that were performed
 2. a methods/analysis section that explains the process and techniques used, such as data cleaning, data exploration and visualization, any insights gained, and your modeling approach
 3. a results section that presents the modeling results and discusses the model performance
 4. a conclusion section that gives a brief summary of the report, its limitations, and future work (the last two are recommended but not necessary)
- 0 points: The report is either not uploaded or contains very minimal information OR the report appears to violate the terms of the edX Honor Code.
 - 5 points: One or more required sections of the report are missing.
 - 10 points: The report includes all required sections, but the report is significantly difficult to follow or missing significant supporting detail in multiple sections.
 - 15 points: The report includes all required sections, but the report is difficult to follow or missing supporting detail in one section (or has minor flaws in multiple sections).
 - 20 points: The report includes all required sections and is easy to follow, but with minor flaws in one section.
 - 25 points: The report includes all required sections, is easy to follow with good supporting detail throughout, and is insightful and innovative.

Code (20 points)

The code in the R script should run without errors and should be well-commented and easy to follow. It should also use relative file paths and automatically install missing packages. The dataset you use should either be automatically be downloaded by your code or provided in your GitHub repo.

- 0 points: Code does not run and produces many errors OR code appears to violate the terms of the edX Honor Code.
- 5 points: Code runs but does not produce output consistent with what is presented in the report OR there is overtraining (the test set is used for training steps).
- 10 points: Code runs but is difficult to follow and/or may not produce output entirely consistent with what is presented in the report.
- 15 points: Code runs, can be followed, is at least mostly consistent with the report, but is lacking (sufficient) comments and explanation OR uses absolute paths instead of relative paths OR does not automatically install missing packages OR does not provide easy access to the dataset (either via automatic download or inclusion in a GitHub repository).
- 20 points: Code runs easily, is easy to follow, is consistent with the report, and is well-commented. All file paths are relative and missing packages are automatically installed with `if(!require)` statements.

Have a question about the choose your own project? Want to bounce some ideas for an analysis to do or a dataset to pick off someone else? Need some feedback on the best approach to take or some troubleshooting for a snippet of your code? You can ask your questions here!

You are encouraged to discuss **general approaches** to your project. It is okay to post **small snippets** of code if you're having trouble getting a particular piece of code to run. However, you **may not post your entire R script** for the project.

Discussion: Choose Your Own Project

Hide Discussion

Topic: Choose Your Own Project / Choose Your Own Project[Add a Post](#)[◀ All Posts](#)

Choose your own project

question posted 26 days ago by [Mukhan169](#)

Just want to make sure that we I will be doing same analysis including RMSE as I did on the movie dataset. Just the data would be different. Sorry may seem like a silly question but no where in instruction it says that we need to create RMSE. And if thats the case what will be the minimum value of RMSE acceptable?

This post is visible to everyone.

wonyoungcheong (Staff)25 days ago - marked as answer 25 days ago by [wonyoungcheong](#) (Staff)

There is no minimum score. For the Choose Your Own Project, I encourage you to be creative and tackle a problem that interests you by finding (or creating) and analyzing data. Depending on your project, RMSE may not even be the metric used!

There is a lot of flexibility so that you can tailor the project to your preferences and needs. Please be sure to review the grading rubric. I would like to see thorough exploration (EDA, visualizations) and analysis (comparison of machine learning algorithms), and a comprehensive report documenting the process and your findings.

Preview

Submit