

<u>Course</u> > <u>Capstone Project: I...</u> > <u>Project Submission:</u>... > Choose Your Own ...

Choose Your Own Project Submission

CHOOSE YOUR OWN SUBMISSION

Status

You have completed this assignment. Review your grade and your assessment details.

- ► Your Response due Jan 10, 2020 18:59 EST (in 0 minutes) ✓ COMPLETE
- ► Assess Peers due Jan 17, 2020 18:59 EST (in 4 days, 10 hours) ✓ 5 COMPLETED

Staff Grade ✓ COMPLETE

Your Grade: 40 out of 50

The question 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

See GitHub link:

https://github.com/jbalzani/edx-Economic-Growth-or-Contraction-Project

I appreciate your feedback!

Your Upload

script (Economic Growth or Contraction R Script.R)

rmd file (Economic-Growth-or-Contraction.Rmd)

Final Report (Economic-Growth-or-Contraction.pdf)

Assessments of Your Response

5 / 5 POINTS

▼ Files

STAFF GRADE - 5 POINTS

All correct 1

STAFF COMMENTS

Great job including the proper files and link to a github repository!

• PEER MEDIAN GRADE

All correct 1

PEER 1 - ALL CORRECT
All files are there. OK.
PEER 2 - ALL CORRECT
No issues
PEER 3 - ALL CORRECT
Great Job

▼ Report

20 / 25 POINTS

STAFF GRADE - 20 POINTS

Minor flaws **1**

STAFF COMMENTS

I enjoyed your report overall. Good job working to optimize each of your modeling techniques. I think you could have improved the report by providing more description in certain sections. The section where you perform the hypothesis tests in preparation for generating your models could have included more explanation of what each test was accomplishing or an interpretation of the results for each. Great job discussing potential areas for future work in your conclusion and discussing the limitations of the current models.

PEER MEDIAN GRADE

Excellent 1

PEER 1 - EXCELLENT

Throughout report, perfect.

PEER 2 - MINOR FLAWS

Tough to follow. Lots of code and light on explanation. Needs more project discussion.

PEER 3 - EXCELLENT

Great Job

▼ Code

STAFF GRADE - 15 POINTS

Good **1**

STAFF COMMENTS

Good job overall with your code. I felt that it was easy to follow and was appropriately commented, though some comments could have included a higher level of detail. You could possibly include some comments describing your results as well. I would recommend using different object names when loading in the datasets, as they are relatively unintuitive names to start. This is improved when you combine the datasets, but they are called enough before then that it may have been helpful to rename them.

• PEER MEDIAN GRADE

Excellent **1**

PEER 1 - EXCELLENT

Clean code as well.

PEER 2 - EXCELLENT

No issues running the code. Well commented

PEER 3 - EXCELLENT

Great Job

▼ Additional comments on your response

STAFF COMMENTS

It seems that a large limitation of this analysis is that there were so few countries which did not experience growth, as demonstrated by the sensitivity ratings of 0. The random forest model and ensemble model in particular predicted every observation to be a 1, which seems to be influenced by the very small sample size and even smaller sample of values which were actually 0. To strengthen this analysis, you should try to repeat your work on a larger dataset, as it seems this was the biggest limitation in your work.

• PEER 1

Interesting project and results. Here is some white feedback: - When dealing with time series (from what I remember in the ts courses I took long time ago) you can not claim stationarity in a plot like that, there was no trend but there could have been seasonality. The difference in scale also didn't help. I appreciated that you ran a specialized significance test for stationarity. Good job! - You didn't preprocess or scale (standarize / z-score) your data, I believe this would have helped you reach better accuracy. - I would have appreciated seeing a correlogram as part of your data visualization, from my experience dealing with ts data it has helped fitting the best p or the AR(p) model to find out up to which p would help lagging your data with and then using all the p lagging levels as predictor variables. - Another idea, is to split your ts into components: trend (which doesn't seem to be there), seasonal and random. Then use each separately as a predictor variable.

PEER 2

Nice work! Add more project discussion to the report. It's a bit hard to follow and this topic needs more explanation for non-finance people

PEER 3

Great Job

▼ Provide feedback on peer assessments

Course staff will be able to see any feedback that you provide here when they review course records.

Select the statements below that best reflect your experience with peer assessments.
✓ These assessments were useful.
☐ These assessments were not useful.
☐ I disagree with one or more of the peer assessments of my response.
Some comments I received were inappropriate.
Provide feedback on the grade or comments that you received from your peers.
Very useful feedback. One peer in particular had taken a course in working with time series, so was able to tell me what I could have done differently.
Submit feedback on peer assessments

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 <u>edX Honor Code</u>.
- 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 <u>automatically install missing packages</u> 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 <u>packages are automatically installed</u> 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!

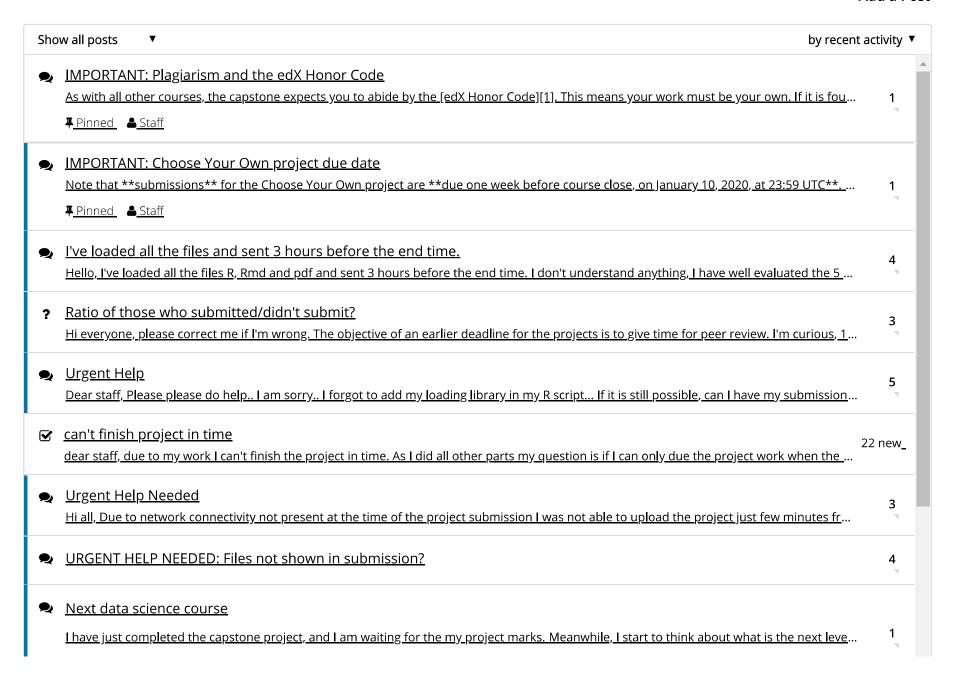
You are encouraged to discuss **general approaches** to your project. It is okay to post **small snippets** of code if you're having trouble getting a particular piece of code to run. However, you **may not post your entire R script** for the project.

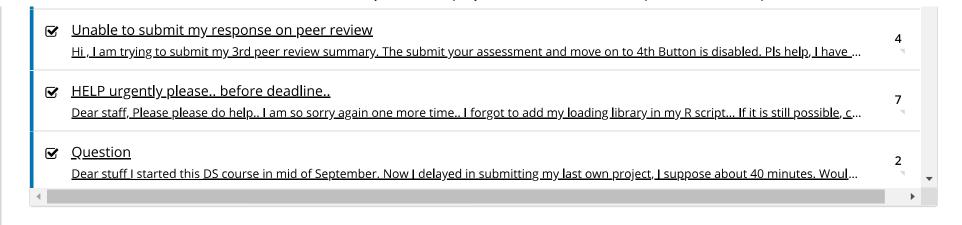
Discussion: Choose Your Own Project

Hide Discussion

Topic: Choose Your Own Project / Choose Your Own Project

Add a Post





© All Rights Reserved