KirigoY /
dsc-phase-1-project

<> **Code**   ⊗ Pull requests   ▶ Actions   ▦ Projects   ▢ Wiki   ⊘ Security   ⬚ Insights   ⚙ Settings

⚖ View license

☆ **0** stars   ⑂ **466** forks   ⊙ **0** watching   ⑂ **1** Branch   ⬚ **0** Tags   ⋀ Activity

🌐 Public repository · Forked from learn-co-curriculum/dsc-phase-1-project

⑂ master ▾   ⑂ 1 Branch   ⬚ 0 Tags   ⑂   ⬚   🔍 Go to file   t   Go to file   +   Add file ▾   Code   ⋯

This branch is up to date with `learn-co-curriculum/dsc-phase-1-project:master` .   ⑂ **Contribute** ▾   ⟳ Sync fork ▾

Cheffrey2000 Update .canvas ⋯   3 years ago ⋯ 🕓

📁 zippedData	Initial commit	4 years ago
📄 .canvas	Update .canvas	3 years ago
📄 .gitignore	Initial commit	4 years ago
📄 CONTRIBUTING.md	Initial commit	4 years ago
📄 LICENSE.md	Initial commit	4 years ago
📄 README.md	Update README.md	4 years ago
📄 awesome.gif	Initial commit	4 years ago
📄 student.ipynb	Initial commit	4 years ago

📖 **README**   ⚖ License   ✎ ☰

Phase 1 Project

You've made it all the way through the first phase of this course - take a minute to celebrate your awesomeness!

Now you will put your new skills to use with a large end-of-Phase project! This project should take 20 to 30 hours to complete.

Project Overview

For this project, you will use exploratory data analysis to generate insights for a business stakeholder.

Business Problem

Microsoft sees all the big companies creating original video content and they want to get in on the fun. They have decided to create a new movie studio, but they don't know anything about creating movies. You are charged with exploring what types of films are currently doing the best at the box office. You must then translate those findings into actionable insights that the head of Microsoft's new movie studio can use to help decide what type of films to create.

The Data

In the folder `zippedData` are movie datasets from:

- [Box Office Mojo](#)
- [IMDB](#)
- [Rotten Tomatoes](#)
- [TheMovieDB](#)
- [The Numbers](#)

It is up to you to decide what data from this to use and how to use it. If you want to make this more challenging, you can scrape websites or make API calls to get additional data. If you are feeling overwhelmed or behind (e.g. struggled with the Phase 1 Code Challenge), we recommend you use only the following data files:

- imdb.title.basics
- imdb.title.ratings
- bom.movie_gross

Deliverables

There are three deliverables for this project:

- A **GitHub repository**
- A **Jupyter Notebook**
- A **non-technical presentation**

Review the "Project Submission & Review" page in the "Milestones Instructions" topic for instructions on creating and submitting your deliverables. Refer to the rubric associated with this assignment for specifications describing high-quality deliverables.

Key Points

- **Your analysis should yield three concrete business recommendations.** The ultimate purpose of exploratory analysis is not just to learn about the data, but to help an organization perform better. Explicitly relate your findings to business needs by recommending actions that you think the business (Microsoft) should take.

- **Communicating about your work well is extremely important.** Your ability to provide value to an organization - or to land a job there - is directly reliant on your ability to communicate with them about what you have done and why it is valuable. Create a storyline your audience (the head of Microsoft's new movie studio) can follow by walking them through the steps of your process, highlighting the most important points and skipping over the rest.

- **Use plenty of visualizations.** Visualizations are invaluable for exploring your data and making your findings accessible to a non-technical audience. Spotlight visuals in your presentation, but only ones that relate directly to your recommendations. Simple visuals are usually best (e.g. bar charts and line graphs), and don't forget to format them well (e.g. labels, titles).

Getting Started

Please start by reviewing this assignment, the rubric at the bottom of it, and the "Project Submission & Review" page. If you have any questions, please ask your instructor ASAP.

Next, we recommend you check out the Phase 1 Project Templates and Examples repo and use the MVP template for your project.

Alternatively, you can fork the Phase 1 Project Repository, clone it locally, and work in the `student.ipynb` file. Make sure to also add and commit a PDF of your presentation to your repository with a file name of `presentation.pdf`.

Project Submission and Review

Review the "Project Submission & Review" page in the "Milestones Instructions" topic to learn how to submit your project and how it will be reviewed. Your project must pass review for you to progress to the next Phase.

Summary

This project will give you a valuable opportunity to develop your data science skills using real-world data. The end-of-phase projects are a critical part of the program because they give you a chance to bring together all the skills you've learned, apply them to realistic projects for a business stakeholder, practice communication skills, and get feedback to help you improve. You've got this!

Releases

No releases published
Create a new release

Packages

No packages published
Publish your first package

Languages

- **Jupyter Notebook** 100.0%