🔒 ga-students / **DSI_LDN_1_LESSON_NOTES** Private

⊙ Watch ▾ 49   ★ Star 0   ⑂ Fork 26

<> Code | ⊘ Issues **0** | ⑂ Pull requests **2** | ▦ Projects **0** | 📖 Wiki | ⚡ Pulse | 📊 Graphs

Branch: master ▾   **DSI_LDN_1_LESSON_NOTES** / projects / capstone / **part-01** /

Create new file | Upload files | Find file | History

👤 **juanginzo** added capstone PT1   Latest commit a45401d 12 days ago

..

| 📄 part-01-rubric.md | added capstone PT1 | 12 days ago |
| 📄 readme.md | added capstone PT1 | 12 days ago |

📖 **readme.md**

🅖🅐 Part 1: Capstone Topic + Dataset Validation

Overview

In the field of data science, projects are **practical**. A good project is manageable and relates to your working domain; however, it can be hard to scope good ideas when you're new to a field. That's where Part 1 of your Capstone Project comes in.

One of the best ways to test expectations and get feedback is to *share your ideas with others*. For part 1 of your Capstone, you'll come up with a problem you want to solve with data, then present it to the class. You must include at least one potential dataset that could be used to solve your problem, as well as at least one potential audience for your findings.

This project will provide you with immediate feedback and guidance to help you select an awesome topic and build a polished Capstone portfolio. You will need to identify your problem statement, goals & criteria for success, potential audience, and 1-2 potential datasets. You will present this information in a slide deck, and should be prepared to answer questions and defend your data selection(s). Presentations should take no more than 3-5 minutes.

Goal: A lightning talk presentation describing your proposed project, criteria, audience, and potential dataset(s)

Requirements

Your work must:

- Define your problem statement, goals, and success criteria.
- Describe 1-2 potential datasets that address your problem statement
- Identify a potential audience of stakeholders who may be interested in your findings
- Create a 3-5 minute presentation slide deck

- Be prepared to discuss and defend your topic & dataset(s)

- *Bonus:*

  - Identify the source and the format of your dataset(s)
  - Write the Python code needed to read in your dataset(s) (**Hint**: numpy method getfromtxt)
  - Create a blog post of at least 500 words (and 1-2 graphics!) that describes your project idea, data, and audience. Link to it in your presentation appendix.

Necessary Deliverables / Submission

- Presentation Slide Deck of 3-5 minutes, including:
    - Description, problem statement, audience, goals, & success metrics.
    - 1-2 Potential Datasets related to topic
- Presentation should be prepared by the end of Week 6

Suggested Ways to Get Started

- Read the docs for whatever technologies you use. Most of the time, there is a tutorial that you can follow, but not always, and learning to read documentation is crucial to your success!
- Write pseudocode before you write actual code. Thinking through the logic of something helps.
- Document **everything**.

Useful Resources

- How to find the data you need
- How to give a good lightning talk

Project Feedback + Evaluation

Attached here is a complete rubric for this project.

Your instructors will score each of your technical requirements using the scale below:

Score	Expectations
0	*Incomplete.*
1	*Does not meet expectations.*
2	*Meets expectations, good job!*
3	*Exceeds expectations, you wonderful creature, you!*

This will serve as a helpful overall gauge of whether you met the project goals, but **the more important scores are the individual ones** above, which can help you identify where to focus your efforts for the next project!