Software Engineering for Data Scientists (SEDS), Homework 5

Due Wed. 3/15 @ 5PM

Jim and I really enjoyed your project ideas that you shared with us on Wed. Homework 5 will help you focus your project goals and check your understanding of use cases, component specifications, etc. Don't wait to start as answering these questions will help you with your presentation on Monday, 3/13.

Tasks
1. Create a new repository for the project. Add all the project participants as collaborators. You can choose one member of the team to host the project in their GitHub account or you can create an organization that will host the repository. Try to have fun thinking of the name for the software. We all know software isn't real until it has a funny name. (1pt)
2. Make sure there is a LICENSE file and a README.md that describes the directory structure, a sentence about the choice of license and a short, descriptive title for the project. You will keep this README up to date over the project with things like dependencies and usage instructions so keep it clean and make it polished. (1pt)
3. Create a Markdown document called Design_Considerations.md . This is where you will place all of your content regarding the big picture design of your software. For this week, layout the use cases (you should have one to three) and for each one describe the component specification. Feel free to use figures, interaction diagrams (e.g. https://www.websequencediagrams.com), etc. to communicate your design. (1pt)
4. Make another document that describes the data sets, their size, dimensionality, data model (what specific columns, etc. are available). Verify that the data are available. Is it possible to create a subset of the data for unit tests? Address this in your document. (1 pt)
5. Make sure everyone on the team does at least one commit! (1 pt)