CAPSTONE PROJECT | DATA SCIENCE IMMERSIVE

PROJECT DELIVERABLES + TIMELINE

---

1. Pitch | Problem Statement - *Week 4*

- **Description**: Pitch us on potential ideas for a data-driven project. Think of topics you're passionate about, knowledge you're familiar with, or problems relevant to to industries you'd like to work with. What questions do you want to answer?
- **Requirements:** Lightning talk with 2-3 topics, including a problem statement, potential audience, goals, and success metrics, as well as possible data sources for each. Remember, if you can't find data, you can't do your project.
- **Format:** Slide deck

2. Dataset | Data Collection - *Week 7*

- **Description**: Use your newfound skills to source and collect the relevant data for your project. Data acquisition, transformation, and cleaning are typically the most time-consuming parts of data science projects, so don't procrastinate!
- **Requirements:** Source and format the data for your project. Perform preliminary data munging and cleaning of the data relevant to your project goals. Describe your data keeping the intended audience of your final report in mind.
- **Format:** Table, file, or database with relevant text file or notebook description.

3. EDA | Preliminary Analysis - *Week 8*

- **Description**: Begin quantitatively describing and visualizing your data. With rich datasets, EDA can go down an endless number of roads. Maintain perspective on your goals and scope your EDA accordingly. Managing your own time is a critical skill in analysis projects. Keep notes on your approach, results, setbacks, and findings.
- **Requirements:** Perform initial descriptive and visual analysis of your data. Identify outliers, summarize risks and limitations, and describe how your EDA will inform your modeling decisions.
- **Format:** Jupyter notebook with well-formatted code and written discussion.

4. Findings | Technical Report - *Week 11*

- **Description**: Share your technical findings with your fellow data scientists. Explain your goals, describe modeling choices, evaluate model performance, and discuss results. Data science reporting is technical, but don't forget that you should tell a compelling story about your data.
- **Requirements:** Summarize your goals and metrics for success, variables of interest, and removal of any outliers or data imputation. Your process description should be concise and relevant to your goals. Summarize statistical analysis, including model selection, implementation, evaluation, and inference. Be convincing – justify all important decisions! Clearly label plots and visualizations. Include an Executive Summary.
- **Format:** Jupyter Notebook

5. Presentation | Non-Technical Summary - *Week 12*

- **Description**: Take your findings and share a 10 minute presentation that delivers the most important insights from your project to a non-technical audience. Tell us the most interesting story about your data. Break down your process for a novice audience. Make sure to include compelling visuals. Time is short, so be sure to practice and include only the most relevant components of your project.
- **Requirements:** Convey your goals, limits/assumptions, methods and their justification, findings, and conclusions. Define technical terms. Include graphics and visualizations.
- **Format:** Interactive graphic presentation, website, or slide deck.
