# CMSE 830 Final Project

In your first project the focus was mainly on implementing data science methods to explore and analyze the dataset. In the second (final) project, you will focus on implementing modeling methods. The modeling methods are to be drawn from machine learning methods (linear regression, neural networks, support vector machines, gaussian process regression/classification, KNN, and any other regression-classification technique) as an exploratory or a predictive tool.
You can use the same dataset you used for your first project, if that makes sense, or start with a completely new dataset. For example, you can switch to a project that uses time series analysis; or, add a modeling component to your previous project, such as regression or classification.

Project Requirements:

- As before, the project will be in the form of a Streamlit app that is deployed at GitHub. You may wish to link to this from your CV, so make the app professional. Importantly, make the app realistic as if you are developing this for a company and not a class project. That is, avoid using phrases like “in this project” or titles like “violin plots”. Soon, you will either be applying for jobs or PhD programs and will want to show that you didn’t just solve a bunch of HW problems!
- Your project should have a goal. Think of this as answering a question or delivering a service - this is far beyond basic EDA.
- Narrative is a requirement for a successful presentation. You are trying to tell a story based on what you have learned. Narratives try to connect different pieces of your project and present those in a way that leads the

user to reach similar remarks, conclusions, or an understanding of the project as yours.
    - The narrative should not be a walkthrough for pandas or seaborn.
    - The narrative begins in the introduction, expands as the user progresses through the project, and ends with remarks and conclusions.
    - Without a proper narrative and context, a visualization is just a picture without a specific purpose or goal.
- Make the app organized. For example, don't put all of the content into one very long app that the user needs to scroll down through. Use a menu bar on the left or tabs to provide a convenient and intuitive layout. Examine the options and choose what works best for your purpose.
- Include an introduction, as it provides a brief overview of the project. Your introduction should inform the user of what the app does for them. You probably want a sentence like "This app will enable you to….". .
    - The introduction must include an explanation of the purpose for the potential user.
    - As discussed in class, think of the project as a showcase of your skills to achieve a purpose, and avoid using the term "project" within the app, as mentioned above.
    - Remember that the project should be attractive and appealing to the user, and the introduction as the first section of the project plays a vital part in keeping the user interested (remember the example of the first page of a newspaper).
    - For the final project, add a bio tab that gives us a little information about yourself, your skills, hobbies, and interests. The bio section adds a personal touch to the project.
    - Think of your project as a product of a company, not an assignment

- Interactivity is important. The interactivity can be achieved through streamlit functionality or the visualization itself. The streamlit functionality accounts for sidebar or input widgets, whereas the visualization interactivity is achieved through Plotly or Altair visualizations. Your final project must be interactive using the streamlit functionalities and visualizations.
    - While Streamlit provides several options to interact with users and ask for input, these options must be implemented appropriately. For example, while input within each tab affects that tab and the following materials, the sidebar inputs affect the entire web app.
    - Do not describe the functionality (defining sidebar, slider, ...) or visualization (violin plot is a plot that ...). We already know what these are called.
    - Make suggestions to users on how to interact with the project or the data, or what to look for. But, do not provide useless variations to the user for no reason; if you provide a capability (e.g., slider), suggest to the user why it is there (e.g., “By varying this quantity you can reveal that....”)
- As you have observed during the lectures, ICAs, and HWs, visualization is one of the tools that are particularly important in data exploration and analysis.
    - Even though visualizations are important, remember that the project is a showcase of how you solved a problem. Each visualization should serve a purpose and have a clear remark. (See related  comment above about interactivity.)
    - The project is not a Pandas or Seaborn walkthrough (no need to explore every type of visualization); make your choices professional and meaningful.

    - Review visualization libraries, choose the correct type, and follow the grammar of graphics and Tufte's rules for your visualizations.
    - Most of the time, 3D visualizations are not helpful since it is hard to comprehend or compare. It helps a 3D plot if you can rotate, as Plotly allows.
- While typos and errors are common within assignments and projects, the ultimate goal is to have a web app that works smoothly, without problems and errors.
    - If you are using a slider, ensure that the web app works for every possible input.
    - If you are asking the user for various inputs, instead of using text or number inputs, it would be better to use a selectbox or slider.
    - For a visualization that requires only two axes (otherwise two variables), the user should not be able to select more than two features.