Our lesson is on linear regression. We will begin this assignment by introducing the OLS linear regression problem and showing students how to get a closed form solution analytically. Students will be using matplotlib to visualize the data and their results throughout the assignment. They will then implement the analytical solution for simple and multiple linear regression problems. After that, students will be introduced to Scikit Learn’s Linear Regression implementation for modelling and doing prediction. At the same time, we will introduce one-hot encoding and have students do linear regression on one-hot encoded vectors. We will then provide a tougher regression problem that is nonlinear and introduce featurization as a way to do nonlinear regression. Students will featurize their data and perform linear regression to get high accuracy on this tougher regression problem. Finally, we will show students that linear regression can be used for not only regression tasks, but also classification tasks. We will have the students use linear regression to classify a real-world dataset and running regression on it. This will be for both two classes and multiple classes. To this end, we will also help students understand the downsides of OLS to motivate and hint towards regularization.

TODO: Talk about how to navigate folder