[Project Submission Link:](https://nam10.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdocs.google.com%2Fforms%2Fd%2Fe%2F1FAIpQLSdzJ1gItVy7FvH1XFxEeyA5rFFb2kIDN39NjksB_mn2TxDQNw%2Fviewform&data=05%7C02%7Cjjo5541%40psu.edu%7C0f513659ffb543fdbb3d08dcd84198a6%7C7cf48d453ddb4389a9c1c115526eb52e%7C0%7C0%7C638623023443006477%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C0%7C%7C%7C&sdata=hGu5%2BJTjM8V1VX59AphXkO3EituVZMCoK03vicA%2FOEk%3D&reserved=0)

Use Scikit learn for programing white box models

Machine Learning – training a computer to read data and respond to it

Linear Regression – predict a variable based on the value of another variable

Logistic Regression – is used to generate a “yes” or “no” type of response based on independent variables

SVM (support vector machine) – separates data into two parts

Used for classification, supper fast, good with high dimensionality

Decision Trees – algorithm that uses a tree structure

K-Means Clustering – cluster the groups together based on classification