This is my first Github Project. This project was inspired from the Data School video series on SciKit-learn. From that ten-part series I recreate and later build off of what I learned using my own data on 401k participation.

The goal of this project is to recreate what I learned from that video series. The steps of the project include three different methods used in the video series:

1. Linear Regression
   1. Model Training - Train and test
   2. linear regression
   3. Interpretting the model
   4. making predictions
   5. Evaluate metrics
   6. Cross-validation - feature selection
2. KNN
   1. Model Training
   2. Model Evaluation
   3. Cross-validation
      1. K fold Cross validation: parameter tuning using GridSearchCV
   4. Model Evaluation
   5. Compare Train/test Split w/ kfold
   6. Evaluate results
3. Logit regression
   1. Model Training
   2. Model Evaluation
   3. Cross-validation
      1. K fold Cross validation: parameter tuning using GridSearchCV
   4. Model Evaluation
   5. Compare Train/test Split w/ kfold
   6. Evaluate results
4. Cross-Validation: model selection

This is to build confidence and showcase my ability to use python for data science. All operations were done in a Jupyter Notebook.