Linear Regression Pipeline:

1. Importing the standard libraries
2. Importing the dataset
   1. Using sklearn
3. Viewing the dataset
   1. “the number of features”
   2. “the features names:”
   3. “the number of examples”
4. Transfer to polynomial feature
   1. Using api
   2. Print the shape of new freature
5. Implementing the function that return the close form w for linear regression
   1. Return w values
6. Implementing the function that return the close form w for logistic regression
   1. Return w values
7. Implementing the evaluation function
   1. Return the train\_error and test\_error values
8. Finish writting the k\_fold\_cross\_validation function.
   1. Returns the average training error and average test error from the k-fold cross validation
   2. use Sklearns K-Folds cross-validator: <https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.KFold.html>
   3. centering the data so we do not need the intercept term (we could have also chose w\_0=average y value)
   4. scaling the data matrix
   5. determine the training error and the test error
9. Predict testing set using w.

Reference

Closed-form solution: <https://sebastianraschka.com/faq/docs/closed-form-vs-gd.html>

K-fold cross validation: <https://machinelearningmastery.com/k-fold-cross-validation/>

Ridge regression: <https://towardsdatascience.com/ridge-and-lasso-regression-a-complete-guide-with-python-scikit-learn-e20e34bcbf0b>