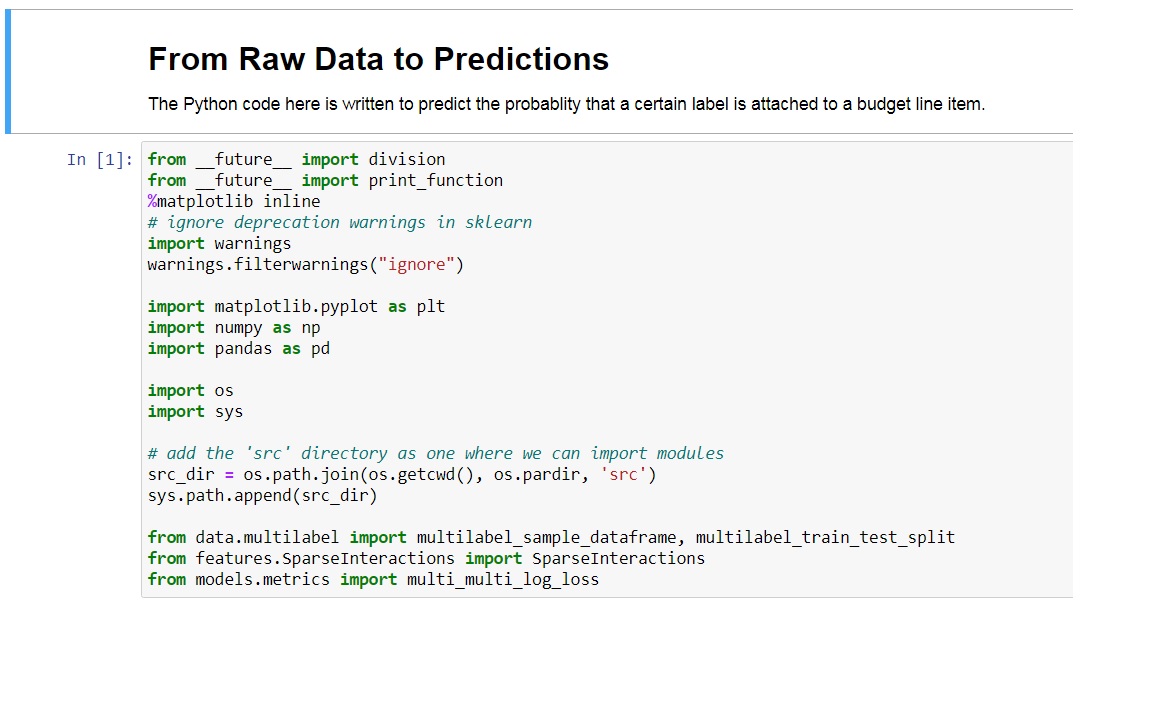
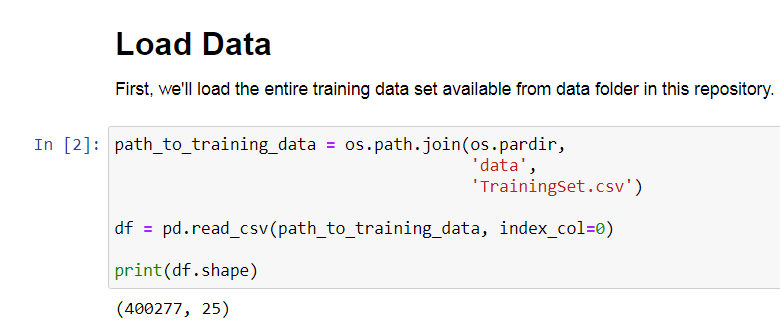
Following are the steps taken to solve the Challenge:

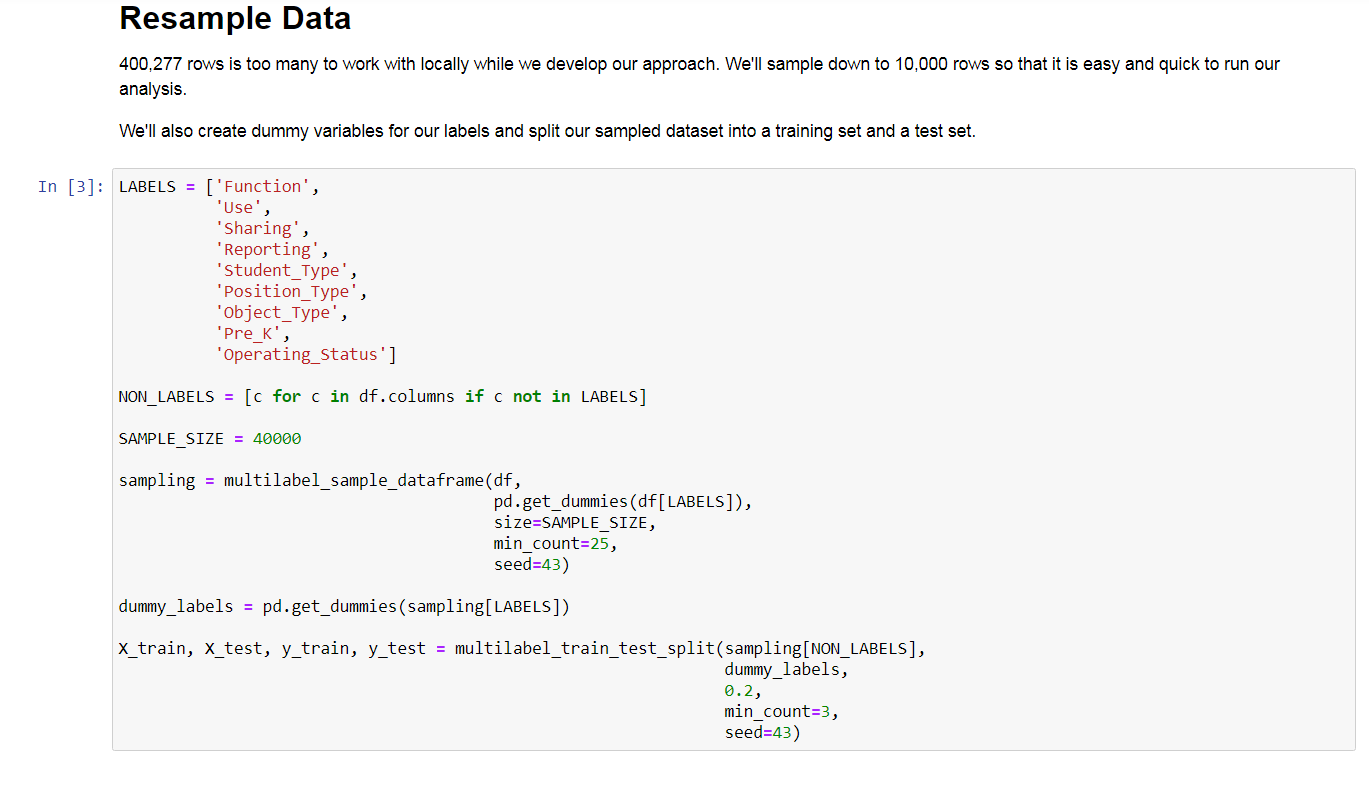
1. From Raw Data To Predictions



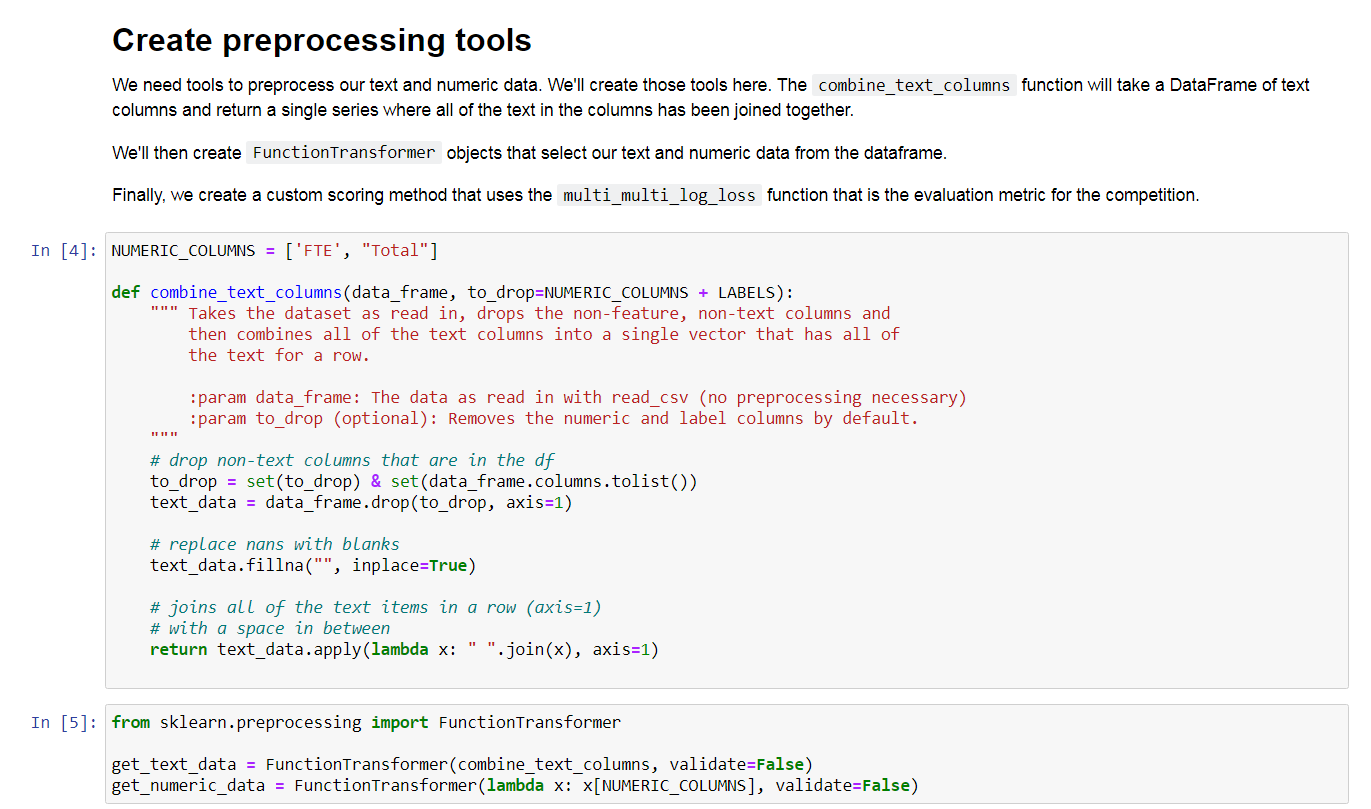
1. Loading the Data by reading the \*.csv file.



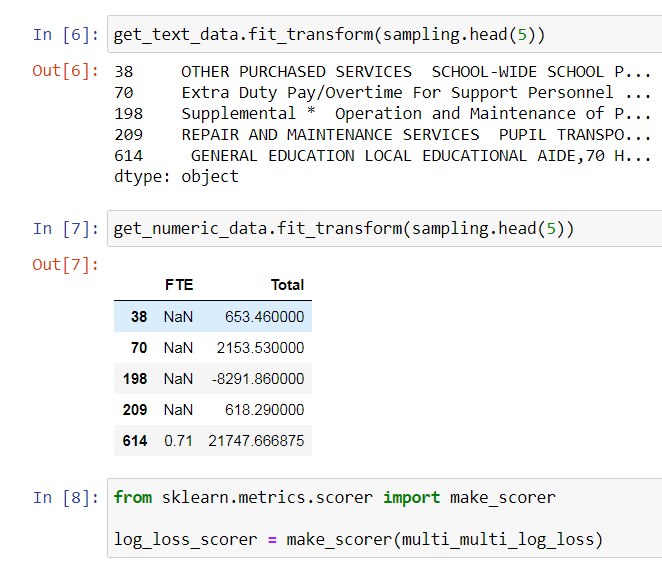
1. Resample the Data



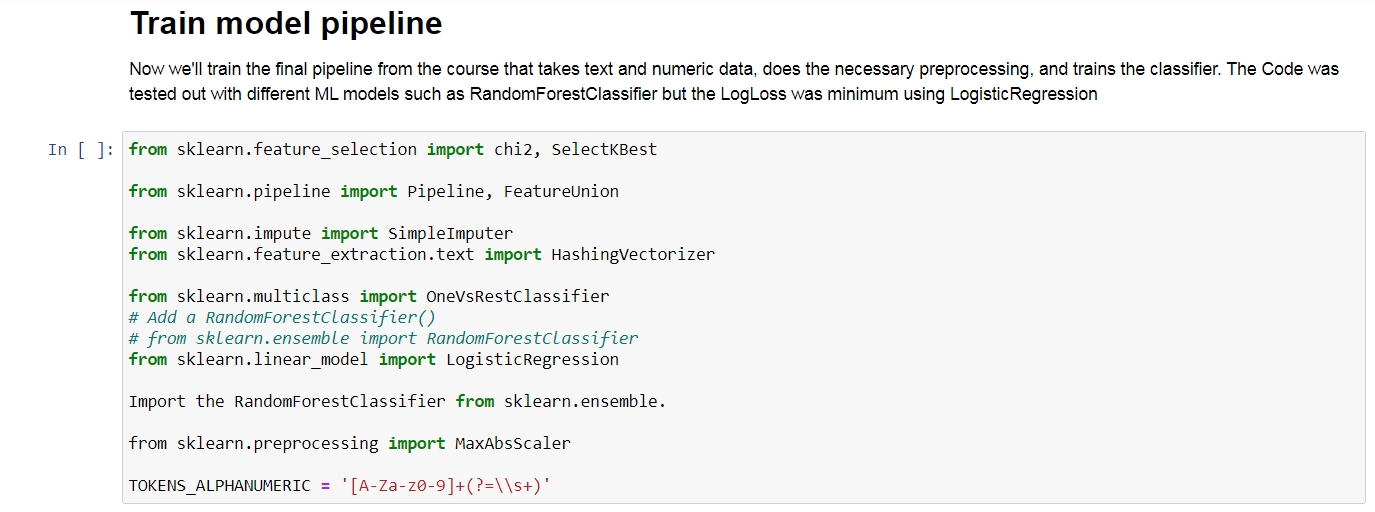
1. Pre-processing the data. Writing the function to do it.



1. Check the first five rows after fitting and transformation



1. Training the Model using the pipeline from scikit learn and calculating the LogLoss





**Note: Lesser the Logloss, better is the model.**

***The model was evaluated on***

***LogisticRegression***

***RandomForest Classifier***

***k-nearest neighbors algorithm (k-NN)***

The Log Loss score for each of them was calculated and it was found Logistic Regression was best suited.

Finally, the Training Model is applied on the Test Data to calculate the prediction.

