CAP STONE PROJECT PROPOSALS

<https://www.kaggle.com/c/prudential-life-insurance-assessment/data>

In this dataset, you are provided over a hundred variables describing attributes of life insurance applicants. The task is to predict the "Response" variable for each Id in the test set. "Response" is an ordinal measure of risk that has 8 levels.

<https://www.kaggle.com/c/homesite-quote-conversion/data>

This dataset represents the activity of a large number of customers who are interested in buying policies from Homesite. Each QuoteNumber corresponds to a potential customer and the QuoteConversion\_Flag indicates whether the customer purchased a policy.

The provided features are anonymized and provide a rich representation of the prospective customer and policy. They include specific coverage information, sales information, personal information, property information, and geographic information. Your task is to predict QuoteConversion\_Flag for each QuoteNumber in the test set.

<https://www.kaggle.com/c/santander-productrecommendation/data?train_ver2.csv.zip>

In this competition, you are provided with 1.5 years of customers behavior data from Santander bank to predict what new products customers will purchase. The data starts at 2015-01-28 and has monthly records of products a customer has, such as "credit card", "savings account", etc. You will predict what **additional** products a customer will get in the last month, 2016-06-28, in addition to what they already have at 2016-05-28. These products are the columns named: ind\_(xyz)\_ult1, which are the columns #25 - #48 in the training data. You will predict what a customer will buy **in addition to what they already had at 2016-05-28**.

The test and train sets are split by time, and public and private leaderboard sets are split randomly.

<https://www.kaggle.com/c/customer-retention/data>

Two datasets will be provided (in the zip file rolled\_up\_data\_Kaggle\_V2.zip) during this competition:

* Dataset #1 (one record per policy term)
* Dataset #2 (multiple records per policy term.  The multiple records comes from having a new record everytime a change is made to the policy)

In all of the data, policies expire every 6 months.  A month before the customer's current policy ends, they will receive a letter in the mail informing them of their ability to renew their policy.  The customer’s renewal period is very important because this it the only time our company will actively change a customer’s price.  Of course, the customer can change their own price for an insurance policy by adding a car, changing coverage, etc.