**Main progress**

* Time Series Research for probability
* Get New Dataset
* Data Exploration – Audit history

**1. Time Series Research for probability**

Professor Jason gave us some hint about using time series. After some thinking, I think I can use the model I built in the last blog to build a time series model to see the trend of the probability change. For example, saying that I have a probability set that have 5 probabilities choice for each row record. I can calculate the real probabilities for each probabilities in each given quarter. Then I could see the trend that each probabilities change with the time in their own range. I could use this to forecast the accurate probabilities for the next quarter to make better financial predictions.

The problem is that we need the history audit data with probabilities. Unfortunately, with the dataset I have in hand, we can’t make it happen.

**2. Get New Dataset**

We get a new data set from Credit Suisse. This dataset is a short cut of the large audit history dataset. In this dataset, it just has the opportunity id, changed field name, old value, new value, and modify date. It doesn’t include the information that if the deal is won or lost.

I think I can use this data set to build a model to see if any change would have a significant influence on winning a deal.

**3. Data Exploration – Audit History**

The first step is to merge APAC opportunity and new audit history dataset. I left join new history and APAC on opportunity id and previous opportunity id respectively. I do this because the audit history doesn’t have the information if the deal is won or not. After some data preparation, I group the win and lose numbers by changed field name and I do a data visualization for three graphs:

1. Total row record number for each field name

2. Total win row record number for each field name

3. Win Ration for each field name