**Blog 4: March 1st, 2019 (Week of Feb 18th - Feb 28th)**

During this week, our team focused on the initial data exploration and processing.

We first do some statistics on the target variable: “statuscodename” in APAC file. It is a categorical variable with categories: Initialize, Due Diligence, finalize, Lost(Dead), Lost(Declined), Won(Funded), Won(On Recommendation List), Won(Unfunded). It is really confusing. Why we have so many different categories? What’s the meaning of won-unfunded? And why the distribution is so imbalanced? We found that in our dataset, about 80% samples’ status is either won or lose. We assumed that maybe transactions with initialize/finalize/duediligence status are still on-going. We also found that salesperson change the close date multiple times, and we don’t know why. So we communicated with Mariem, and put forward our questions.

According to Mariem, the won-unfounded is a subcategory of winning a deal, it means that the salesperson won the deal but still waited to receive money. And also, we can treat initialize/finalize/duediligence to be outliers. And she also mentioned that, it is interesting that salesperson change the close date multiple times and it worth digging into.

We also got constructive and helpful feedback on our project draft from Credit Suisse. We eliminated some of content that our clients were not interested in, and added some suggestions according to the feedback. The feedback is very useful and helps us effectively adjust our project goal and research direction and prioritize our strategies.

However, we encountered some problems. One problem is how to construct time series model. Time series model is often used to forecast future values based on previously observed values, but the problem here is that there are too many different transactions from different clients, different companies, it is unlikely to construct a dataset to use some companies’ transactions to predict other companies’ transactions. We need further discuss it with Professor Jason.