**Week of February 11st - February 17th**

**Blog 3: February 15th , 2019**

In the weekly meeting, our group has a chat with Mariem. We get familiar with Power BI, a data visualization tool which could be install in our CreditSuisse VMs.

We also know more background about our CSV data files, which are APACAMER.csv , Atlas opportunity file Audit history.csv and so on.

There are some variables that are important for our analysis (APACAMER.csv) .

* Account ID: Always a number. / If its the same account, itll be the same ID
* When is a deal supposed to close?
* Lot of fields are null (check for data cleaning before to replace)
* Client ID
* Stage
* Probability: note there are few fields added due to system migration (they will be labeled with an underscore) sth. Starts or ends with amd
* Eg a business process with start with a B
* Probabiliies will be distinct values- not continuous- could see them as categorical
* Amount could be different currencies: default would be USD
* Status and probabilities
* Opportunity for an account or a product;
* It could be a coverage 1 person or coverage 2 person : If people are trying to analyse- what matters the most is coverage 1(main coverage)- person you will go to for the opportunity basically - there are people rarely for coverage 2- more for backup! GO TO IS COVERAGE 1 (maybe coverage 2 could be interesting) Coverage 2 can be ignored / EDA may need it
* Modified by: system or someone else edit it
* Products; hierarchy: product group level structure (PGL)=> PGL A,PGL B, PGL C >> Product (product is more detailed), A is a parent company (Alphabet would be A and Google might be product)

Also for the excel Atlas opportunity file Audit history.csv.

* More about opportunity
* Equivalent
* This is one of the archive files
* Nice and Clearer
* Amount yield
* mapping : opportunity ID: you can join different files through this
* NOT complete
* Tie back an opportunity : winning or losing
* Status/stage: is something 0 as opposed to 1: which of the features affected it

During this week, we have done some descriptive analysis and showed our notebook to Mariem. In the next step, after figuring out our target variable and meaning of some variable, we will try to build a regression model to analysis the realtionship between the dependent and independent variables.