**Debt issuance cohorts / clusters**

Many large companies make use of debt markets to fund their operations and expansion.  Companies issuing new debt represents an opportunities for banks such as ANZ, so understanding what drives companies to issue new debt and the ability to predict when a company is likely to complete a new debt issuance would be a major driver of value for the bank.

We have a number of hypotheses relating to this that we would like to test:

Within the debt markets, some companies act as "leaders" within the debt markets, spotting new opportunities before other companies, while other companies act as "followers".  An example might be that lending conditions are more favourable in Australia than other markets.  These "leaders" will respond to this opportunity by issuing debt in this market.  Other companies will then respond to the leaders (and market conditions) and then issue debt later.

If this was the case, then when we see "leaders" behaving in a particular way, ANZ can then market the "followers" with solutions similar to those that the "leaders" have implemented.

For leaders and followers to make sense, we need to cluster companies into groups that exhibit leadership / following behaviour.  E.g. we need to find "like" companies who enter into the same markets at staggered intervals, and thus have some predictive power.

This is a mammoth task, and one that years of work could be poured into, so we don't expect a complete solution (but if you get one, that would be great!).

To help get you started, we would like to propose the following tasks:

1. Use "debt\_issuance" to filter the other files to remove data not from companies with debt issuances.

2. Using debt\_balance and debt\_issuance, find all the companies issuing debt in "new" markets (in currencies they haven't issued debt in the last 3 years), and the dates in which they issued debt. ("market\_entries")

3. Using "market\_entries", rank companies based on the timing of the entrance into that market (this rank will be your leadership score)

4. Cluster companies based on their financial metrics, so that there is a spread of leaders and followers

5. Given the clusters, create some explanation on why these companies should be grouped together that a business user (non-ml, non stats) person could understand.

6. Assess the predictive power for predicting a debt issuance

To enable this analysis, we have provided you with the following CSV files (tab de-limited):

debt\_balance.tsv

For each quarterly reporting period, list all the debt products a company has on their books, including the rates, and outstanding balances. This is the place to look for new debt issuance's, as a company must declare its debt products in its reports.

metrics.tsv

Translation from the IQ\_ codes into more meaningful names

balance\_sheet.tsv

For each quarterly reporting period, list the balance sheet metrics for the company.  The metric codes are defined in metrics.csv.

income\_statement.tsv

For each quarterly reporting period, list the income statement metrics for the company. The metric codes are defined in metrics.csv.

cashflow.tsv

For each quarterly reporting period, list the cash flow metrics for the company. The metric codes are defined in metrics.csv.

------------------------

The data is / will be available:

[https://www.dropbox.com/s/jyy5dwxtr8aqdc0/anz\_debt\_](https://www.dropbox.com/s/jyy5dwxtr8aqdc0/anz_debt_issuance.zip?dl=0)