**C5T 2 – Lessons learnt for future projects (One Page)**

**Lessons Learnt**

* Did you learn anything of potential business value from this analysis?

There could be immense value for business for decision support for loan approvals. Management can make policy decisions for acceptable risk levels and tie the probabilities to financial risks and rewards for the organization. Organization can evaluate the customer profiles deeply to further augment the decisions making process. In this specific case, it is valuable for management to know that our default rates are 22% which can be compared with competitor firms.

* What are the main lessons you've learned from this experience?
* Business Insights
  + High default rates profiles
    - Females
    - Young population
    - University Graduates
    - Customers with high spending levels
    - Customers with revolving credits (Highest Default Rates)
  + Default rates for married and singles are similar
  + No significant impact in other categories (more analysis may be valuable)
* Data/Analytics Insights
  + Bill payments data (by month) for is skewed. It has many data points with lower or zero amounts
  + Using log function makes the distribution more normal
  + Removal of zero values that make the data skewed reduces the data points
  + Data for August seems suspect quality
* What recommendations would you give to the Data Science team regarding your findings?
* Remove outliers and analyze them separately
* Be careful about ‘restriction of range’
* Try to analyze the data that skews the plot
* Check for correlations by segmenting data further across features and see if we can find better predictability.
* Compare results from multiple algorithms (not all predict equally well in all situations)
* Consider binning for limit balances and age ranges