1. **Credit cards- classification**
2. *What is the problem you want to solve?*

Credit is a risk encompassing business where the probability of default of a credit is important for the lending companies. A portion of customers might default on payment and it is imperative for the financial lending companies to predict the probability of default. This problem estimates the real probability of default.

1. *Who is your client and why do they care about this problem? In other words, what will your client DO or DECIDE based on your analysis that they wouldn’t have otherwise?*

Companies in the global financial and lending space and global payments like VISA, MASTERCARD, AMERICAN EXPRESS as well as financial institutions which lend money would be interested in the problem

These companies in the financial and payments space would be able to classify customers based on their payment history and take decisions on lending based on this analysis.

1. *What data are you going to use for this? How will you acquire this data?*

The data available in UCI machine learning repository will be used for this project. This data is available as a public dataset. <https://archive.ics.uci.edu/ml/datasets/default+of+credit+card+clients>

1. *In brief, outline your approach to solving this problem (knowing that this might change later).*
   * Data Wrangling methods is used to cleaned for any missing values
   * Exploratory Data Analysis is done to understand the patterns in data
   * Applicable methodology is selected to build the model. The model will be built to predict the probability of default using test data
   * The model is tested and fine tuned
   * Report is created with exploratory data analysis charts, findings and recommendations.
2. *What are your deliverables? Typically, this would include code, along with a paper and/or a slide deck.*

* R Code used for the analysis in github.
* Project report outlining the problem, approach and findings.
* Slide deck for Springboard’s technical blog.