

---

# Project Proposal - Capstone 2

02<sup>nd</sup> July 2018

## Problem and Client

Santander Group wants to provide their current customers and new customers services not found in other banking providers. The first step in achieving this level of service, Santander Group wants to **determine the amount or value of a customer's transaction**. This will enable them to personalize their services to current and new customers. Customers today are more likely to do business with a company if they provide personalized service; this applies to all sectors of business.

## Data

The data for this project has been provided by the Santander Group in the form a Kaggle competition (<https://www.kaggle.com/c/santander-value-prediction-challenge>). Santander Group has provided 3 csv files; 1) a sample submission file, 2) a training set file and, 3) a test set file. The training and test files are an anonymized dataset containing numeric feature variables, the numeric `target` column, and a string `ID` column.

## Approach and Deliverables

The approach will be to use the training set to develop a model to predict a value for the target column using the test set. The deliverables for this project will be all the code, a report detailing the work completed and a slide deck providing the story and answers.