

Objective:

We will analyze the Google Merchandise Store (also known as GStore, where Google swag is sold) customer dataset to predict revenue per customer.

Date Source:

<https://www.kaggle.com/c/ga-customer-revenue-prediction/data>

Work distribution:

The project will go through standard 6 phased Data Science life cycle :

1. Business Understanding
2. Data Cleaning
3. Data Exploration
4. Feature Engineering
5. Predictive Modelling
6. Data Visualizations

Every phase will have two leaders responsible for dividing the tasks for the phase and task distribution. Once the leads have decided a plan for the phase, all team members will work toward the phase deliverable. Selecting the leads and task allocation is and will be carried out based on individual interests, time constraints, and team capability.

Activity	Leads
1. Business Understanding	Sandeep, Nisha
2. Data Cleaning	Nisha, Prashant
3. Data Exploration	Prashant, Prathima
4. Feature Engineering	Prathima, Sandeep
5. Predictive Modelling	Prashant, Sandeep
6. Data Visualizations	Nisha, Prathima

Midterm Deliverable:

We will submit cleaned dataset, along with our understanding of the data generating distribution.
Stretch goal: Finish data exploration and get an understanding of the data.

Analyzing the outcome:

The dataset is a part of an ongoing kaggle competition. We will use the error values as an indicator of the model's success. We will then compare our error values with the ones from Kaggle competition to evaluate our project's success.

Team members: Nisha Dayananda, Prashant Rathod, Prathima Devanath, Sandeep Raju