



# Lead Score Case Study

**Members:**

Amarjeet Singh

Anand

Ankit



# Problem Statement

- X Education sells online courses to industry professionals.
- X Education gets a lot of leads, its lead conversion rate is very poor. For example, if, say, they acquire 100 leads in a day, only about 30 of them are converted.
- To make this process more efficient, the company wishes to identify the most potential leads, also known as 'Hot Leads'.
- Once this set of leads is identified, the lead conversion rate should go up as the sales team will now be focusing more on communicating with the potential leads rather than making calls to everyone.

**Business Objective:** X education wants to know most promising leads. For that they want to build a Model which identifies the hot leads. Deployment of the model for the future use.



# Solution Approach

- Data cleaning and data manipulation. 1.
  - a. Check and handle duplicate data.
  - b. Check and handle NA values and missing values.
  - c. Drop columns, if it contains large amount of missing values and not useful for the analysis.
  - d. Imputation of the values, if necessary.
  - e. Check and handle outliers in data.
- EDA
  - a. Univariate data analysis: value count, distribution of variable etc.
  - b. Bivariate data analysis: correlation coefficients and pattern between the variables etc.



## Solutions Contd ...

- Feature Scaling & Dummy Variables and encoding of the data.
- Classification technique: logistic regression used for the model making and prediction.
- Validation of the model.
- Model presentation.
- Conclusions and recommendations.

# EDA

