LEAD SCORE CASE STUDY

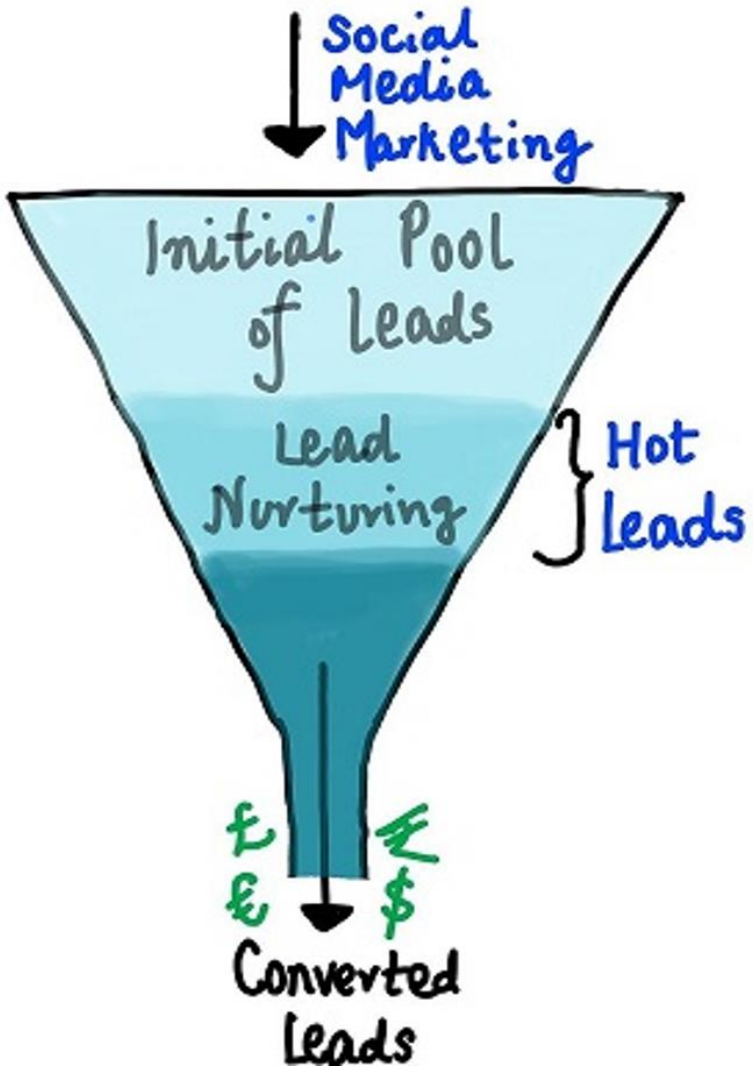
By-

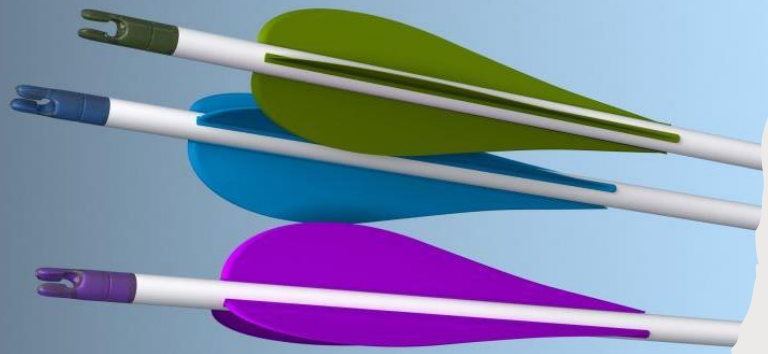
Pranjal Mehta, Ashish Arora,

Sarita Chand

PROBLEM STATEMENT

- An education company named X Education sells online courses to industry professionals. On any given day, many professionals who are interested in the courses land on their website and browse for courses.
- The company markets its courses on several websites and search engines like Google. Moreover, the company also gets leads through past referrals. Once these leads are acquired, employees from the sales team start making calls, writing emails, etc. Through this process, some of the leads get converted while most do not. The typical lead conversion rate at X education is around 30%.





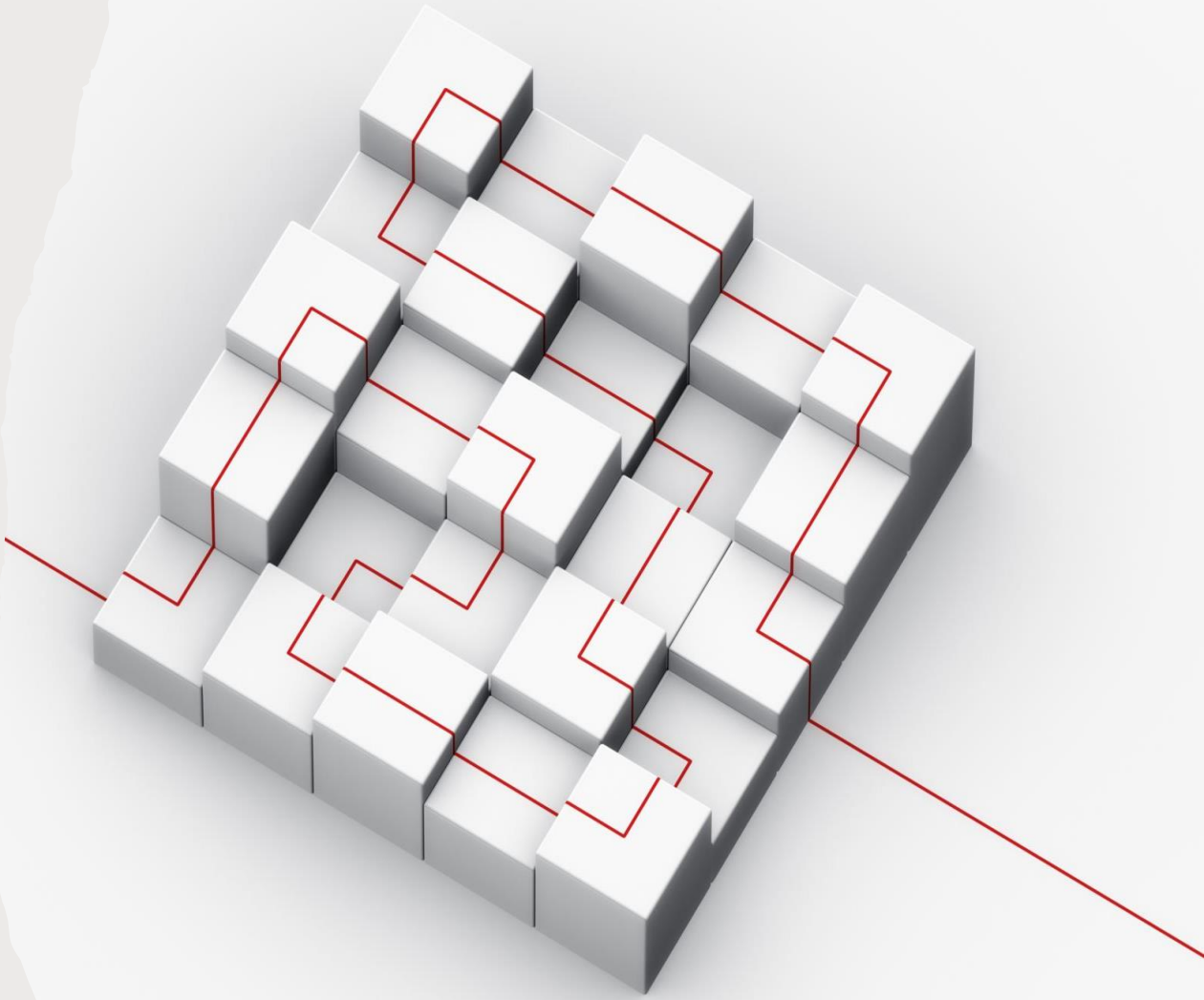
BUSINESS GOAL

- To select the most promising leads, i.e. the leads that are most likely to convert into paying customers by building a model wherein we need to assign a lead score to each of the leads such that the customers with a higher lead score have a higher conversion chance and the customers with a lower lead score have a lower conversion chance.

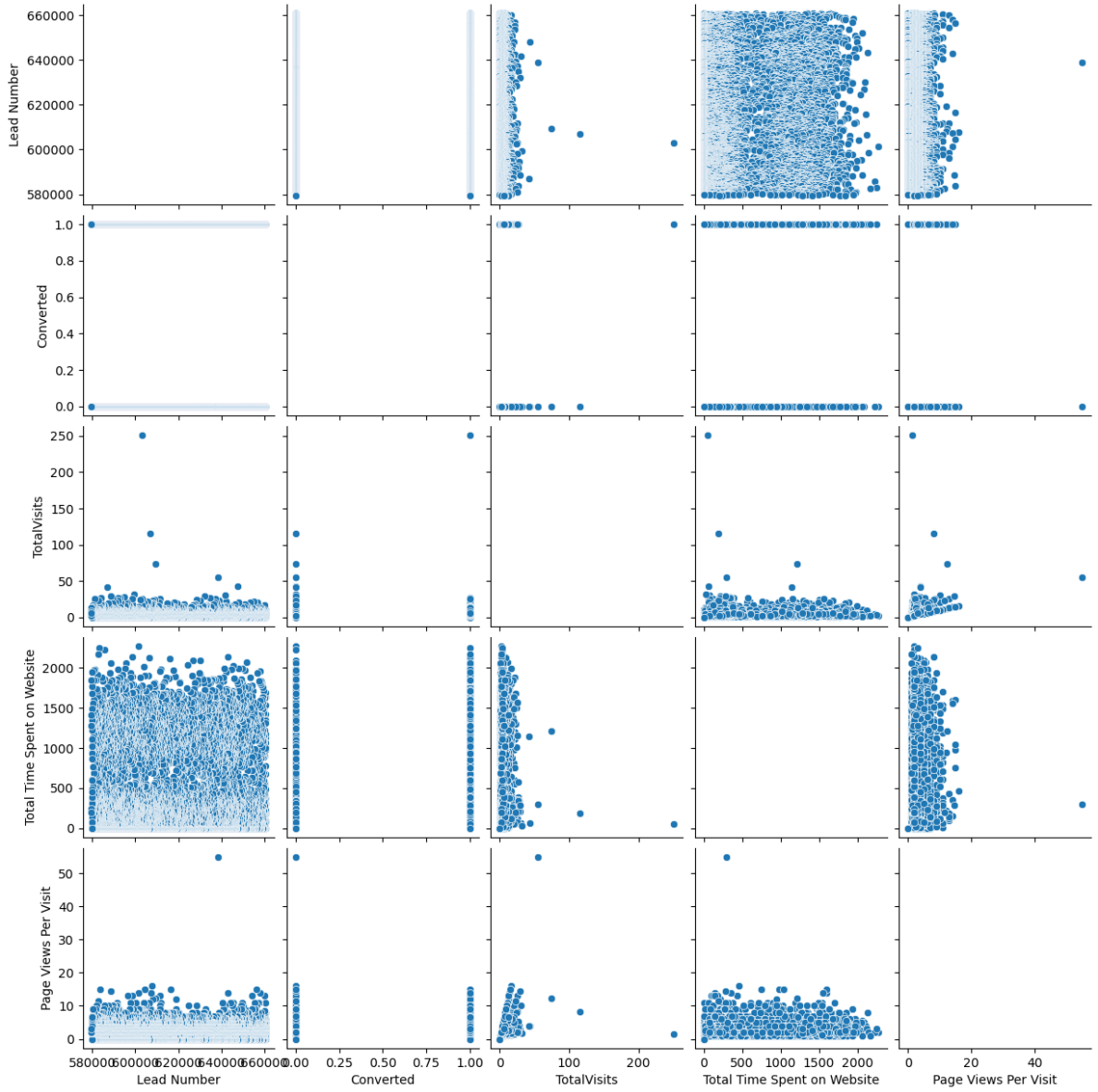
- The model to be built in lead conversion rate should be around 80% or more.

APPROACH

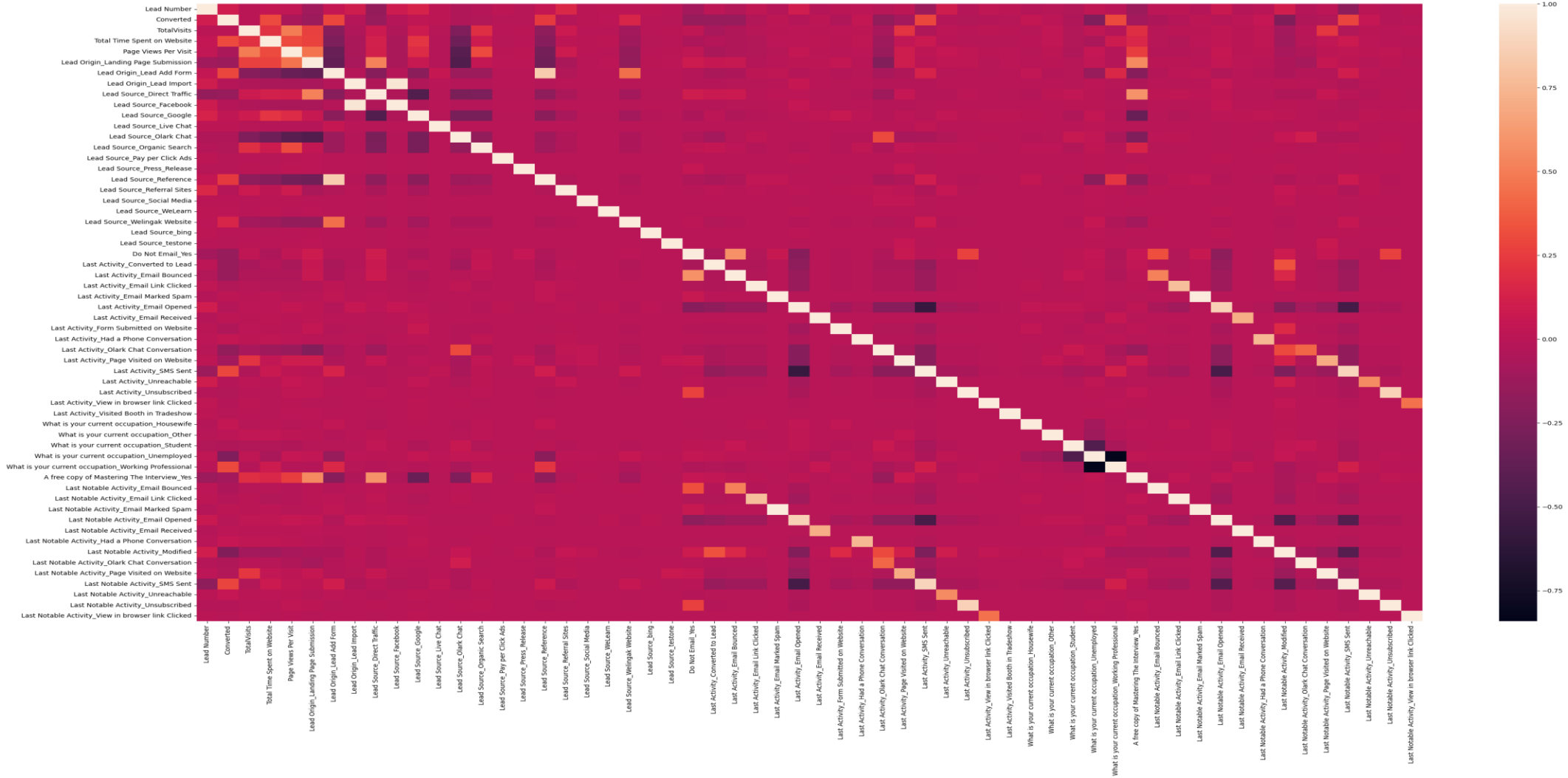
- Import data.
- Clean and prepare the data for further analysis.
- EDA for most helpful attributes for conversion.
- Scaling features
- Data preparation for model building
- Assigning lead score for each lead
- Model test on train set
- Model evaluation
- Model accuracy measurement



PREPARE THE DATA FOR MODELLING

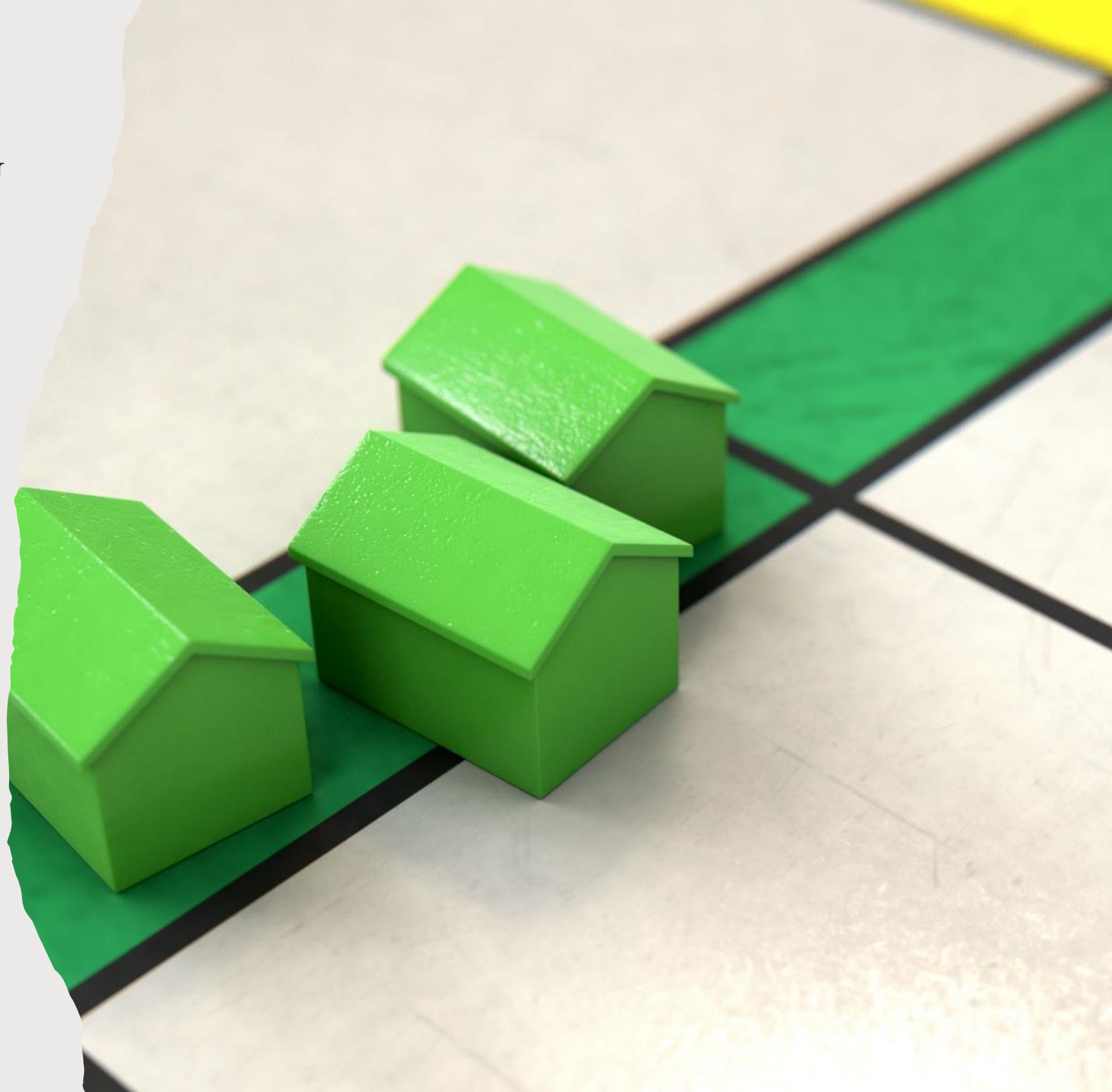


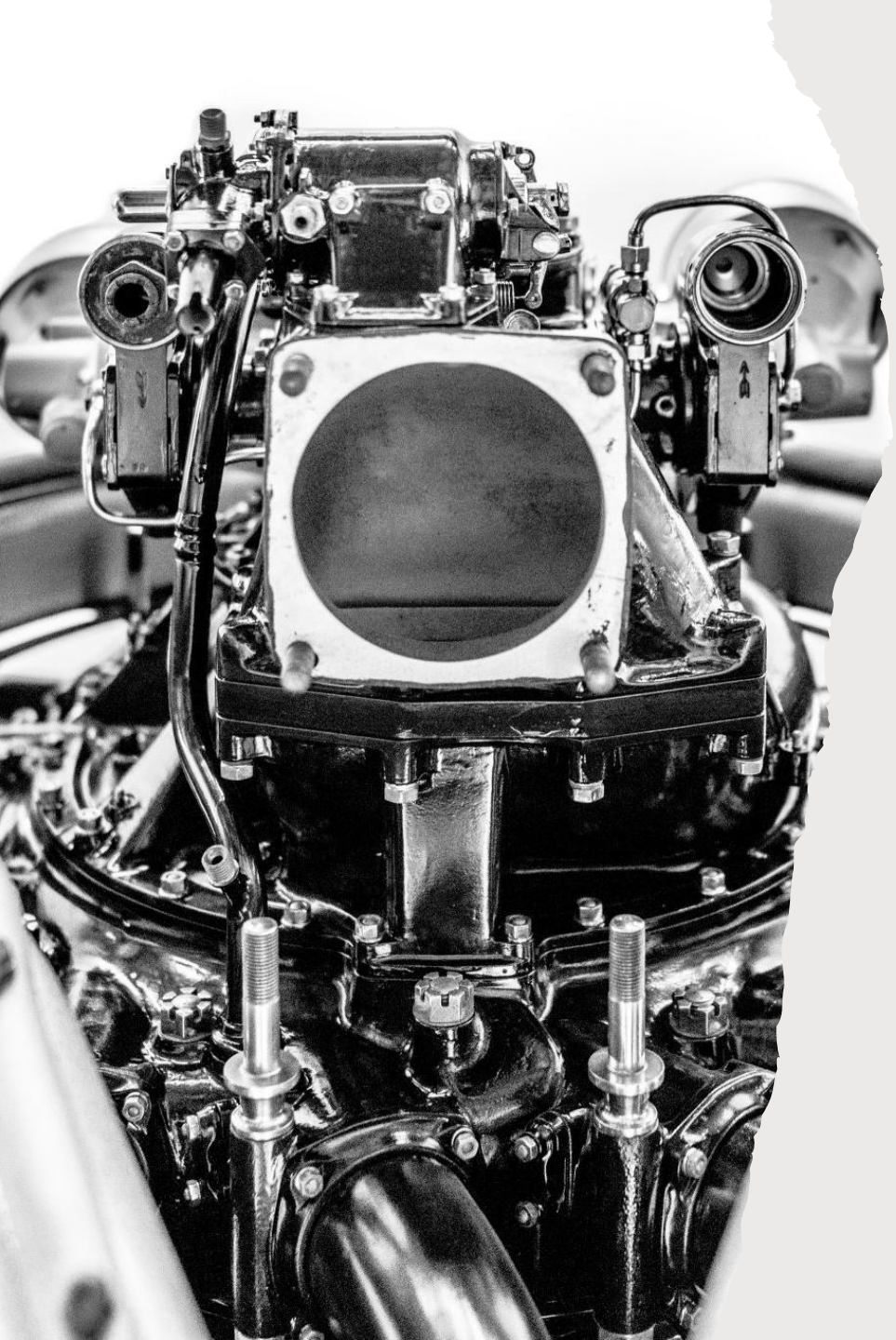
LOOKING AT THE CORRELATIONS



MODEL BUILDING

There are a lot of variables present in the dataset which we cannot deal with. So the best way to approach this is to select a small set of features from this pool of variables using RFE.





FINAL OBSERVATIONS

VIFs for all features are very low.

There is hardly any multicollinearity present.