Steps Followed in the Hackathon:

Day 1:

1. Understanding the Problem Statement
2. Understanding the Data and Hypothesis Generation
3. Exploratory Data Analysis : (Univariate and multivariate analysis)
   1. Numerical Data
   2. Categorical Data
   3. Textual Data
   4. Date Time Data
4. Basic Rule Based Benchmark Models

Day 2:

1. Build ML model
2. Feature engineering ideas, Do feature selection to check performance
3. Build multiple models, Do grid search to find the best set of hyperparameters
4. Ensemble model to improve performance

Steps to make a baseline regression model:

1. Concatenate/ Join the train and test datasets. All kinds of preprocessing must be done on both.
2. Convert all categorical values to numbers – one hot encoding ( since linear model) else Label Encoding.
3. Fill null values.
4. Data Processing/ Feature Engineering steps
5. Split dataset back into train and test
6. Split train into train and validation set
7. Fit regression model on train and check its results on validation
8. Predict for test set
9. Save Submission file