

Utkarsh Raj - Tried data preprocessing (using integer mapping and one-hot encoding - finally settled on one-hot encoding using `pd.get_dummies`) filling Null Values (numeric columns with the median & categorical column with mode), checked correlation, model selection( tried Random Forest, XGB, LBG, CatBoost Regressors - stacked), prediction (got log RMSE of 0.13405).

Richa - Worked on data preprocessing (using mean to fill NAN values) and trying out to fill categorical columns using both integer mapping and one hot encoding, model training of random forest regressor, got log RMSE 0.15.

Nidhish - Worked on preprocessing by filling NaN values with the median. Grouping data was attempted but did not improve RMSE. The team chose a stacked model of XGB, LightGBM, and CatBoost, with an RMSE of 0.13327. Hyperparameter optimization was attempted but was either too slow or ineffective.