**PREDICTION OF ENERGY USED BY APPLIANCES**

The contents of this are as follows:

* ***ANALYSIS AND VISUALIZATION***: In this various plots of target variable vs independent variable are analysed using ggplot2 package for better insights about data and various correlation plot to check correlation coefficients among them.
* ***LINEAR REGRESSION MODEL:***As the first approach for regression problems are straight away linear regression. But this model performs very poorly with adjusted R square as 0.1425 which is very low. Another insights is that this data is suffering from heteroskedasticity (funnel shape plot of residuals vs fitted values) which do not explain the variance between predicted vs actual values on large scale.
* ***LASSO REGRESSION MODEL*** : As our linear regression model failed we might get to lasso regression by improving regularization parameters L1 AND L2. Although this model gave a better model than linear regression but RMSE score is 108.36.
* ***RIDGE REGRESSION MODEL***: this is another model of regularization parameters for better accuracy .Although this model gave RMSE score of 106.705 with continuos cross-validation and improvement through iteration on alpha.
* ***XGBOOST MODEL***: Through xgboost model we have arrived at important gains of independent variables using inbuilt xgb.cv function and calculated best no of rounds for this model .

The data from kitchen ,living room and laundry were ranked highest in importance in energy prediction the prediction models with only the weather data ,selected the atmospheric pressure which is correlated to wind speed as the most relevant weather data variable in the prediction .

Data that measures temperature and Humidity increase prediction accuracy.

Pressure ,Air temperature and wind speed are important parameters in the prediction.

From correlation plots we can take inference about all variables on other variables ,we

Can see that T5,RH\_5,T6,RH6,T7,RH\_7,T8,RH\_8 which are less accessible rooms for members of family are of least importance which is intuitive.

Also plots for temperature vs Energy we can infer that most values are from 20 degree C to 23 degree Celsius and Energy is variated from 0 – 300 Wh.