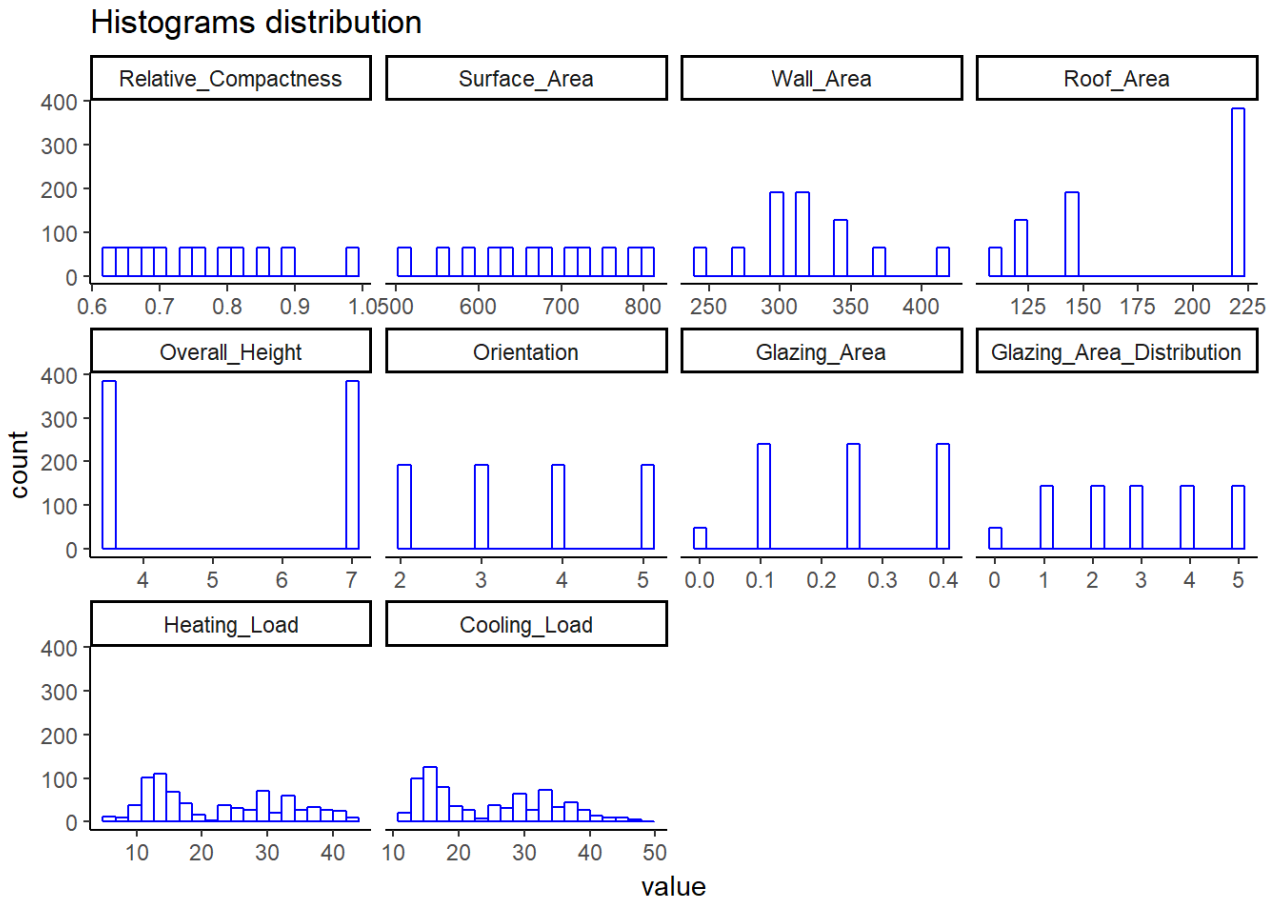
Analysis and Prediction of Energy Performance on Residential Buildings

Tajudeen Abdulazeez and Greg Miller

SYRACUSE UNIVERSITY
School of Information Studies

Data Collected

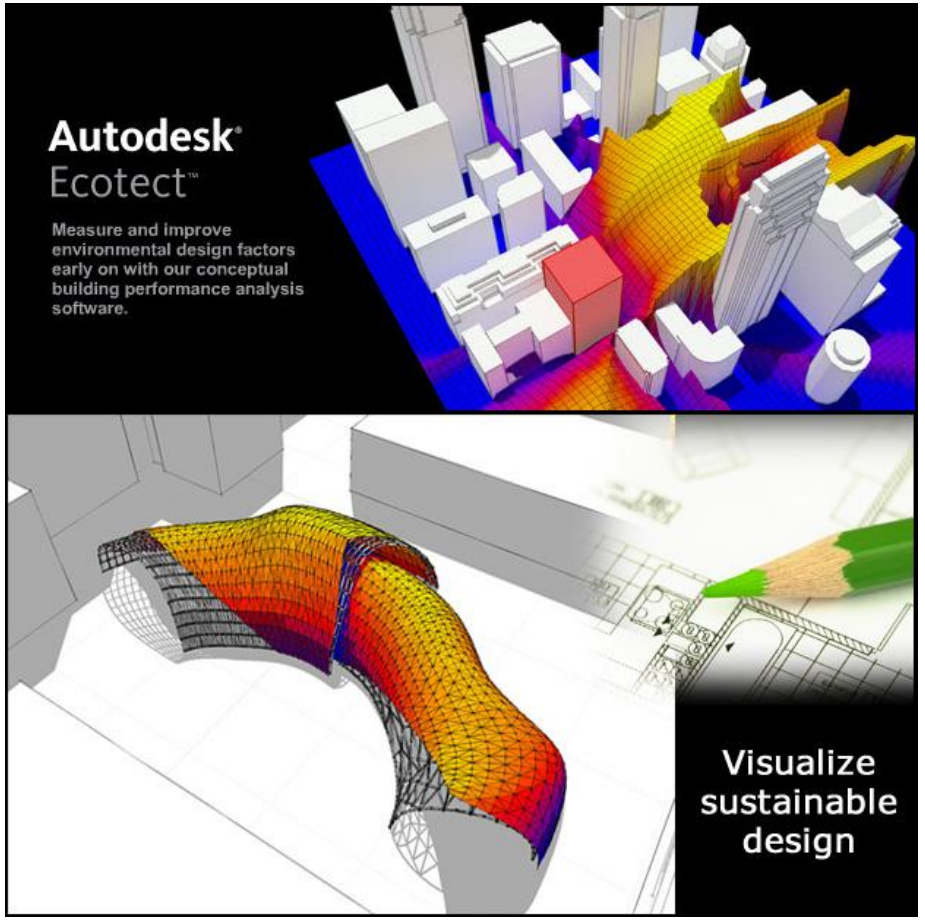


Input/Output Variable	Possible Values
Relative Compactness	12
Surface Area	12
Wall Area	7
Roof Area	4
Overall Height	2
Orientation	4
Glazing Area	4
Glazing Area Distribution	6
Heating Load	586
Cooling	636

https://archive.ics.uci.edu/ml/datasets/Energy+efficiency

Scope

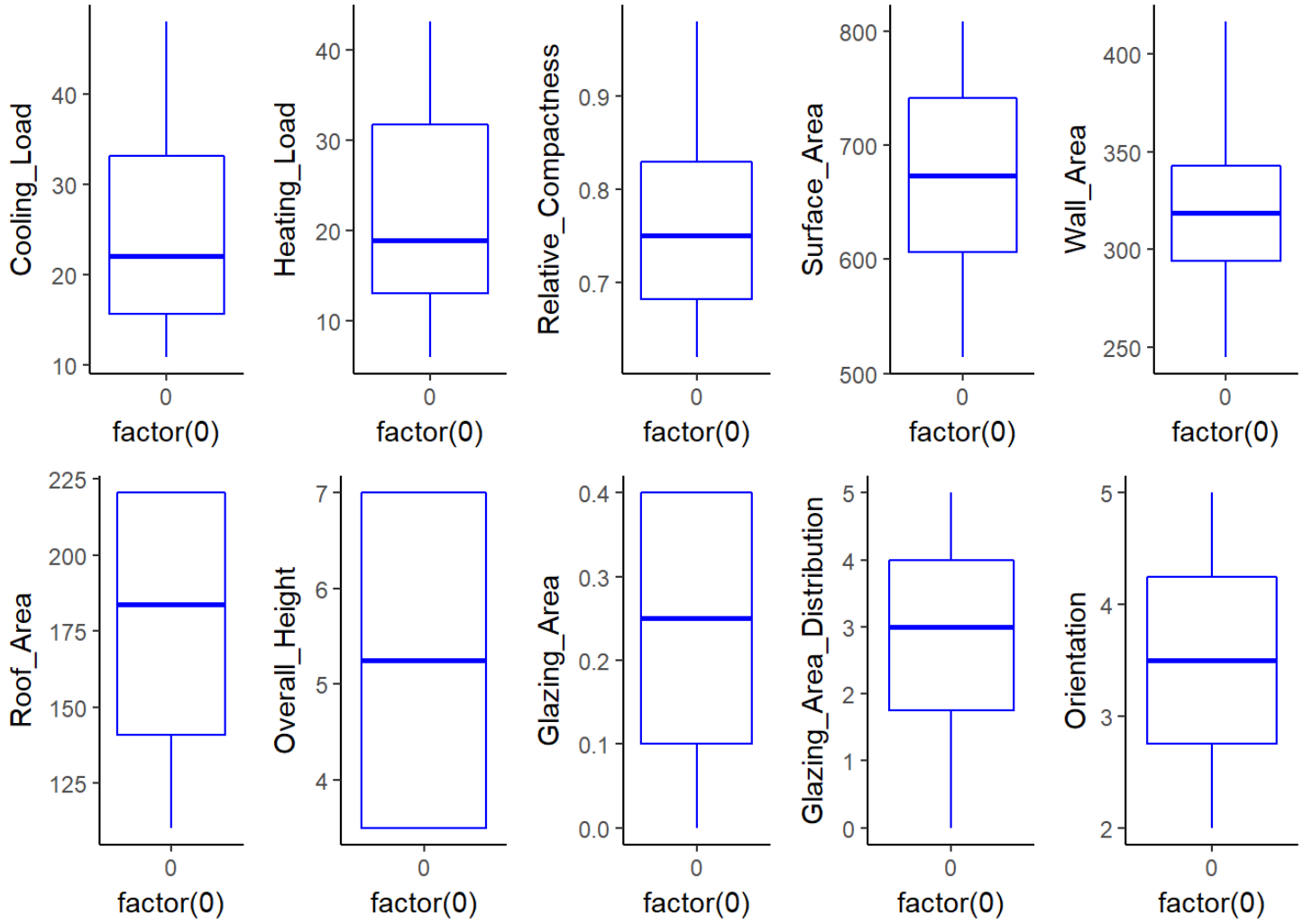
- Ecotect is a simulation tool for designing buildings
- Heating and Cooling Loads indicate the energy efficiency of the building
- Ecotect simulates these values through a difficult interface
- Create a model to better understand input variables and attempt a model that would predict the Heating and Cooling Loads



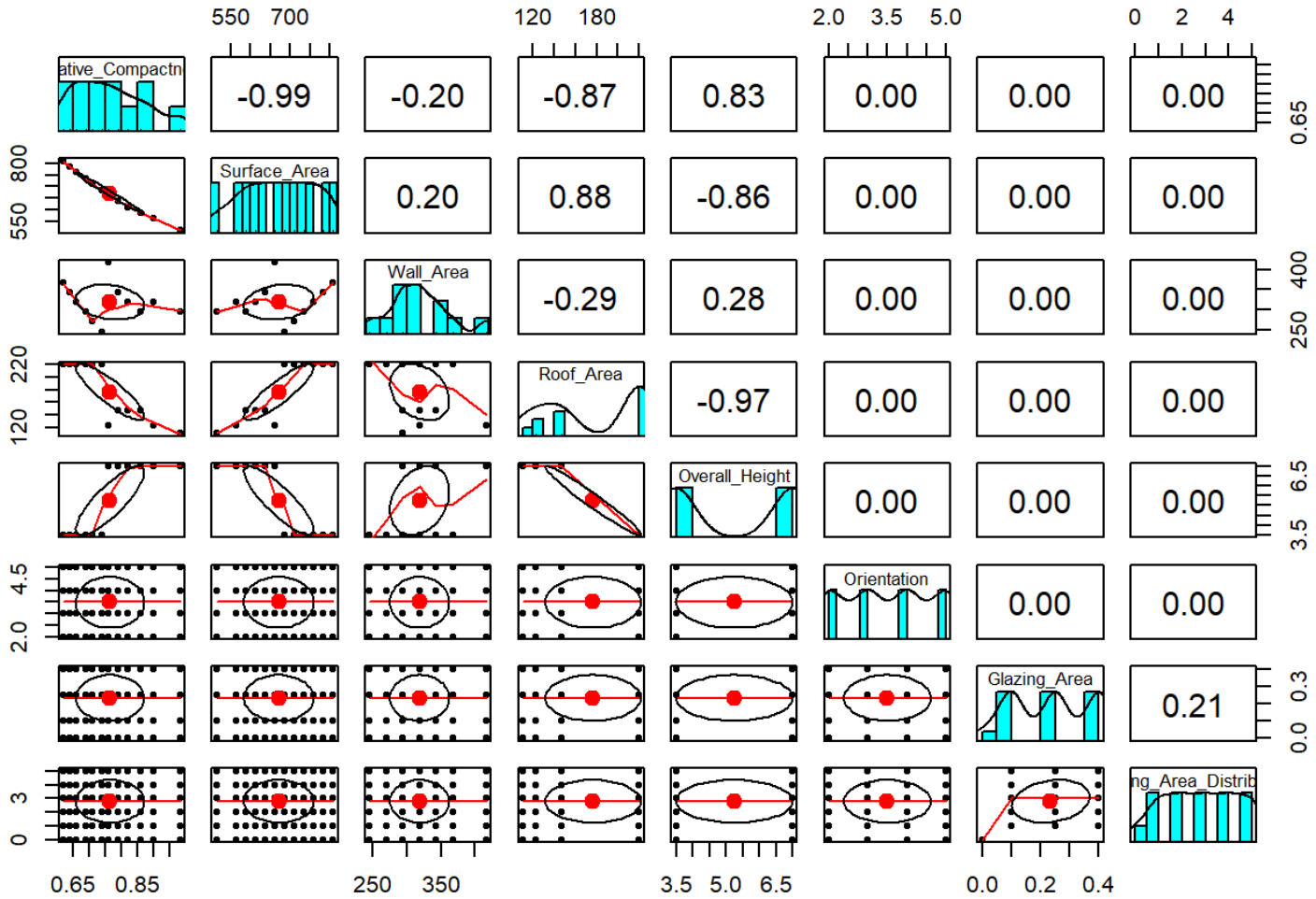
Summary

```
##  Relative_Compactness  Surface_Area     Wall_Area       Roof_Area
##  Min.   :0.6200       Min.   :514.5   Min.   :245.0   Min.   :110.2
##  1st Qu.:0.6825       1st Qu.:606.4   1st Qu.:294.0   1st Qu.:140.9
##  Median :0.7500       Median :673.8   Median :318.5   Median :183.8
##  Mean   :0.7642       Mean   :671.7   Mean   :318.5   Mean   :176.6
##  3rd Qu.:0.8300       3rd Qu.:741.1   3rd Qu.:343.0   3rd Qu.:220.5
##  Max.   :0.9800       Max.   :808.5   Max.   :416.5   Max.   :220.5
##  Overall_Height  Orientation    Glazing_Area    Glazing_Area_Distribution
##  Min.   :3.50   Min.   :2.00   Min.   :0.0000   Min.   :0.000
##  1st Qu.:3.50   1st Qu.:2.75   1st Qu.:0.1000   1st Qu.:1.750
##  Median :5.25   Median :3.50   Median :0.2500   Median :3.000
##  Mean   :5.25   Mean   :3.50   Mean   :0.2344   Mean   :2.812
##  3rd Qu.:7.00   3rd Qu.:4.25   3rd Qu.:0.4000   3rd Qu.:4.000
##  Max.   :7.00   Max.   :5.00   Max.   :0.4000   Max.   :5.000
##   Heating_Load    Cooling_Load
##  Min.   : 6.01   Min.   :10.90
##  1st Qu.:12.99   1st Qu.:15.62
##  Median :18.95   Median :22.08
##  Mean   :22.31   Mean   :24.59
##  3rd Qu.:31.67   3rd Qu.:33.13
##  Max.   :43.10   Max.   :48.03
```

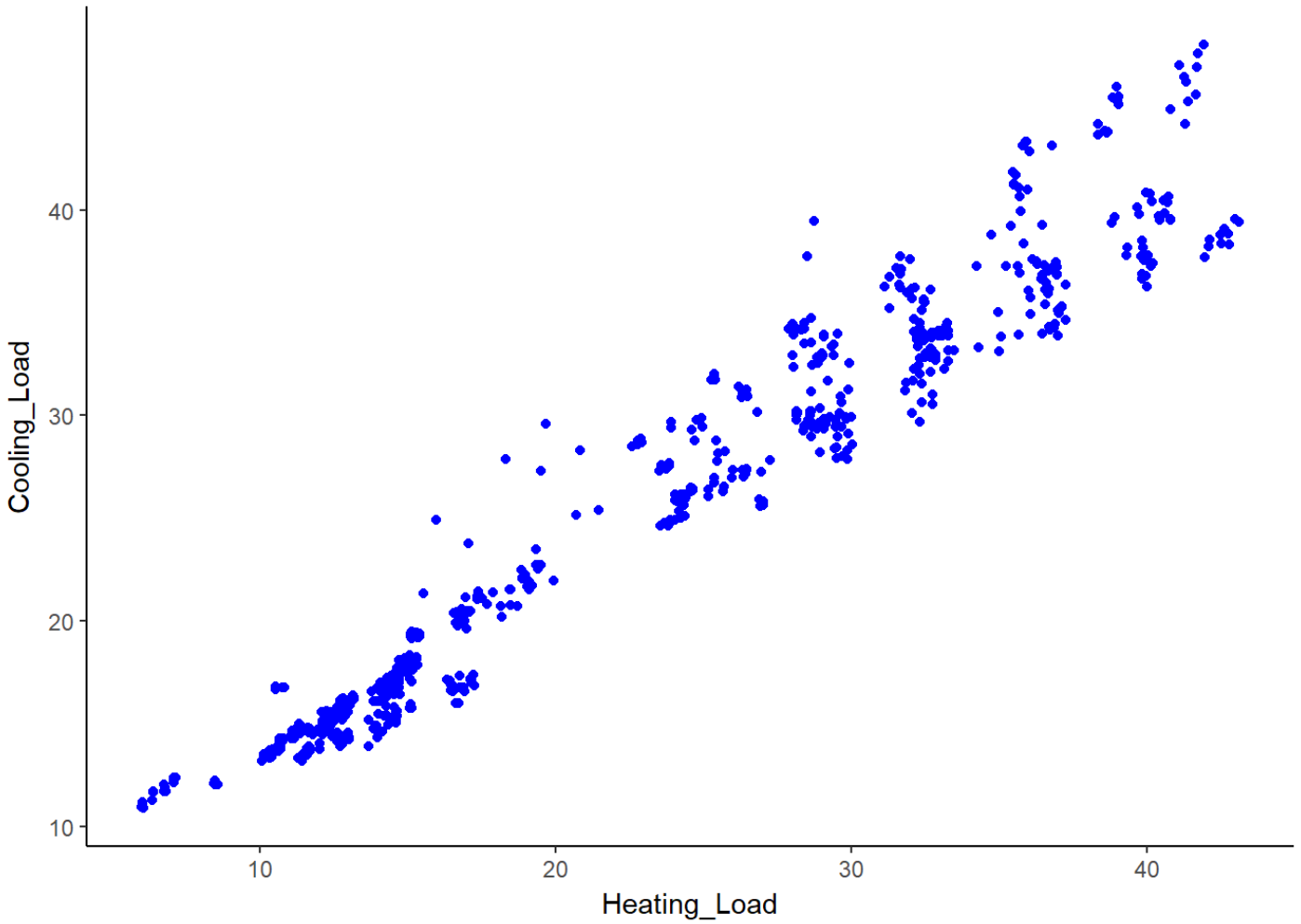
Data Distribution



Input Correlation



Output Correlation



Linear Model

```
##
## Call:
## lm(formula = Heating_Load ~ ., data = Mh)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -10.3862  -1.3667  -0.0142   1.3162   7.5555
##
## Coefficients: (1 not defined because of singularities)
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)            84.386471  19.111765   4.415 1.15e-05 ***
## Relative_Compactness -64.773432  10.333611  -6.268 6.11e-10 ***
## Surface_Area          -0.087289   0.017149  -5.090 4.51e-07 ***
## Wall_Area              0.060813   0.006676   9.109  < 2e-16 ***
## Roof_Area                   NA         NA      NA       NA
## Overall_Height         4.169954   0.339441  12.285  < 2e-16 ***
## Glazing_Area          20.437968   0.798727  25.588  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.947 on 762 degrees of freedom
## Multiple R-squared:  0.9153, Adjusted R-squared:  0.9147
## F-statistic:  1646 on 5 and 762 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = Cooling_Load ~ ., data = Mc)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -8.7240 -1.6017 -0.2631  1.3417 11.3251
##
## Coefficients: (1 not defined because of singularities)
##                          Estimate Std. Error t value Pr(>|t|)
## (Intercept)            97.761848  20.756339   4.710 2.94e-06 ***
## Relative_Compactness -70.787707  11.222822  -6.307 4.80e-10 ***
## Surface_Area          -0.088245   0.018624  -4.738 2.57e-06 ***
## Wall_Area              0.044682   0.007251   6.162 1.16e-09 ***
## Roof_Area                   NA         NA      NA       NA
## Overall_Height         4.283843   0.368650  11.620  < 2e-16 ***
## Glazing_Area          14.817971   0.867458  17.082  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.2 on 762 degrees of freedom
## Multiple R-squared:  0.8876, Adjusted R-squared:  0.8868
## F-statistic:  1203 on 5 and 762 DF,  p-value: < 2.2e-16
```

Heating Load Random Forest

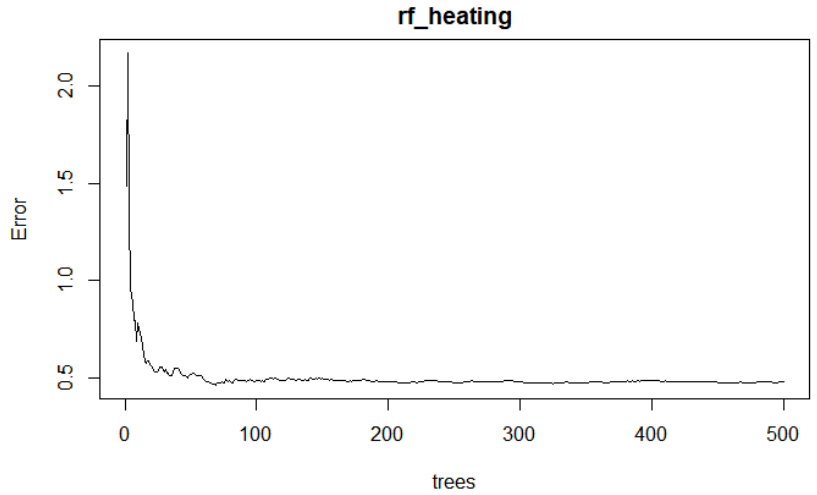
```
#heating load

rf_heating <- randomForest(Heating_Load ~ ., data = Ml, mtry=3, importance = TRUE, na.action = na.omit)

rf_heating
```

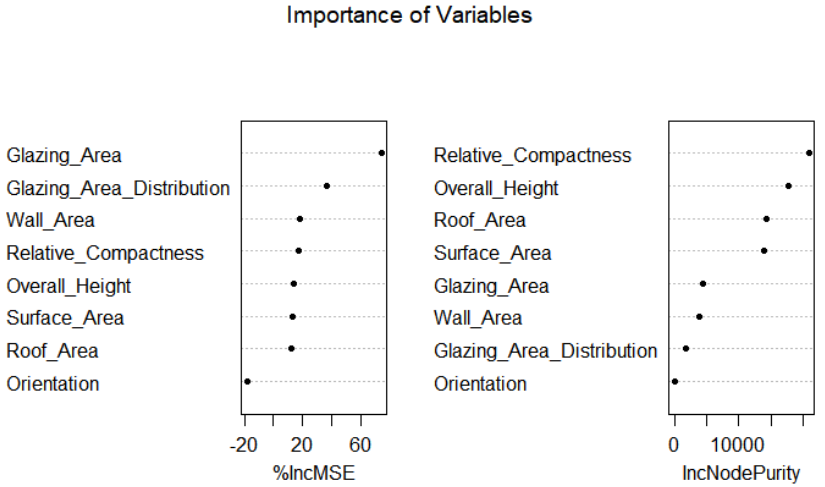
```
##
## Call:
##  randomForest(formula = Heating_Load ~ ., data = Ml, mtry = 3,      importance = TRUE, na.action = na.omit)
##                Type of random forest: regression
##                      Number of trees: 500
## No. of variables tried at each split: 3
##
##           Mean of squared residuals: 0.4447002
##                     % Var explained: 99.56
```

Feature importatnce

```
round( importance( rf_heating ), 2 )
```

```
##                           %IncMSE IncNodePurity
## Relative_Compactness        17.04      20747.03
## Surface_Area                13.48      13741.94
## Wall_Area                   16.62       3631.61
## Roof_Area                   13.26      16200.41
## Overall_Height              13.32      16616.68
## Orientation                -18.72         56.31
## Glazing_Area                83.13       4440.37
## Glazing_Area_Distribution   37.85       1840.84
```





Cooling Load Random Forest

```
rf_cooling <- randomForest(Cooling_Load ~ ., data = m2, mtry=3, importance = TRUE, na.action = na.omit)
rf_cooling
```

```
##
## Call:
##  randomForest(formula = Cooling_Load ~ ., data = m2, mtry = 3,      importance = TRUE, na.action = na.omit)
##                Type of random forest: regression
##                      Number of trees: 500
## No. of variables tried at each split: 3
##
##           Mean of squared residuals: 3.046427
##                     % Var explained: 96.63
```

```
round( importance( rf_cooling ) ,2 )
```

```
##                           %IncMSE IncNodePurity
## Relative_Compactness        15.40      15751.47
## Surface_Area                16.22      17650.40
## Wall_Area                   19.68       3318.23
## Roof_Area                   13.18      14323.40
## Overall_Height              12.06      12810.72
## Orientation                  2.76        251.59
## Glazing_Area                84.32       2456.18
## Glazing_Area_Distribution   25.02        826.77
```