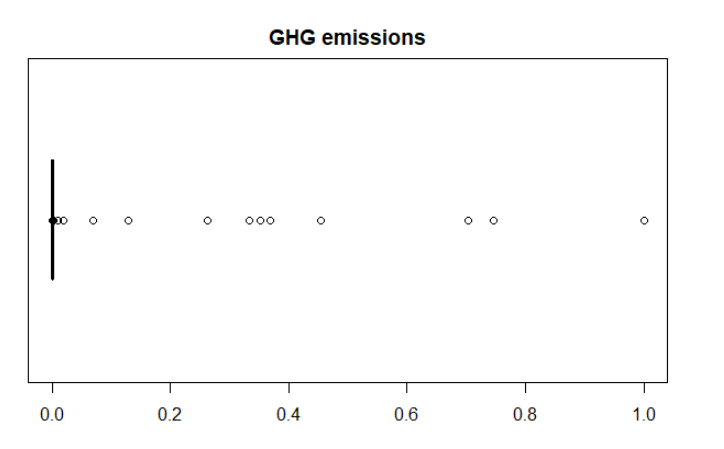
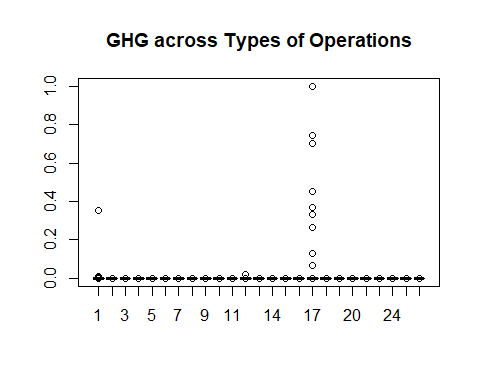
**Results**

The 4 sectors are Municipal, Post-Secondary Educational Institution, Public Hospital, School Board and the Operation types were reduced to

GHG emission varied from 0 to a very large number. Therefore normalized data was used for the study. See plot below.

**Plot 1: GHG emissions using normalized data**



**Plot2:** **GHG emissions using normalized data across operation types**

|  |  |
| --- | --- |
| **OperationType** | **TypeofOperation** |
| Administrative office | 1 |
| Library | 2 |
| Water treatment | 3 |
| Water pumping | 4 |
| Sewage treatment | 5 |
| Sewage pumping | 6 |
| Police station | 7 |
| Fire station | 8 |
| Storage | 9 |
| Community centres | 10 |
| Classrooms | 11 |
| Hospital | 12 |
| Ambulance station | 13 |
| Laboratories | 14 |
| Student residences | 15 |
| Recreational facilities | 16 |
| School | 17 |
| Parking | 18 |
| Indoor swimming pools | 19 |
| Indoor ice rinks | 20 |
| Multi-use | 21 |
| Art galleries | 22 |
| Performing arts facilities | 23 |
| Auditoriums | 24 |
| Other | 25 |
|  |  |

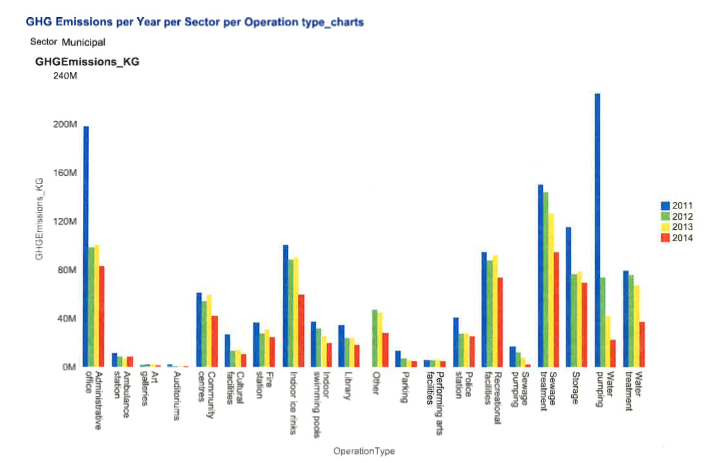
Sector School Board has the highest number of records in the dataset and also the highest GHG. Data below shows the total GHG using the normalized data.

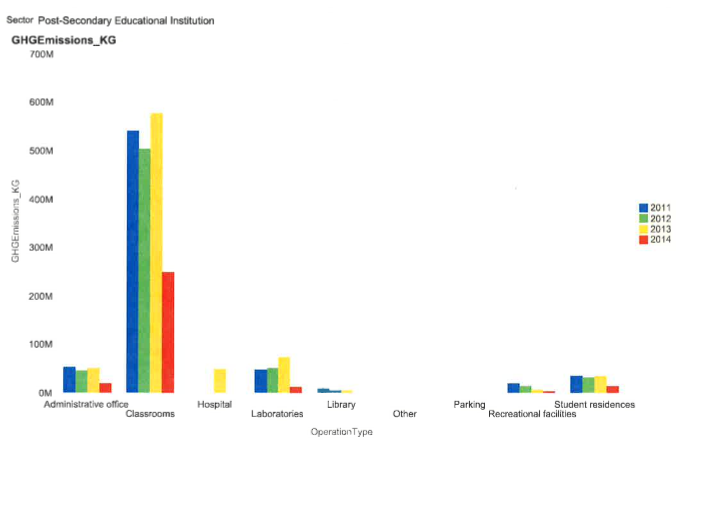
Sector NumofRows totalGHG  
  
1 School Board 25016 4.46   
2 Public Hospital 2262 0.0330  
3 Municipal 55951 0.0159  
4 Post-Secondary Educational Institution 4545 0.0123

Chart 1 and Chart 2 were obtained using Oracle Business Intelligence Enterprise Edition. They show the variation in GHG across operation types through the years 2011-2014.

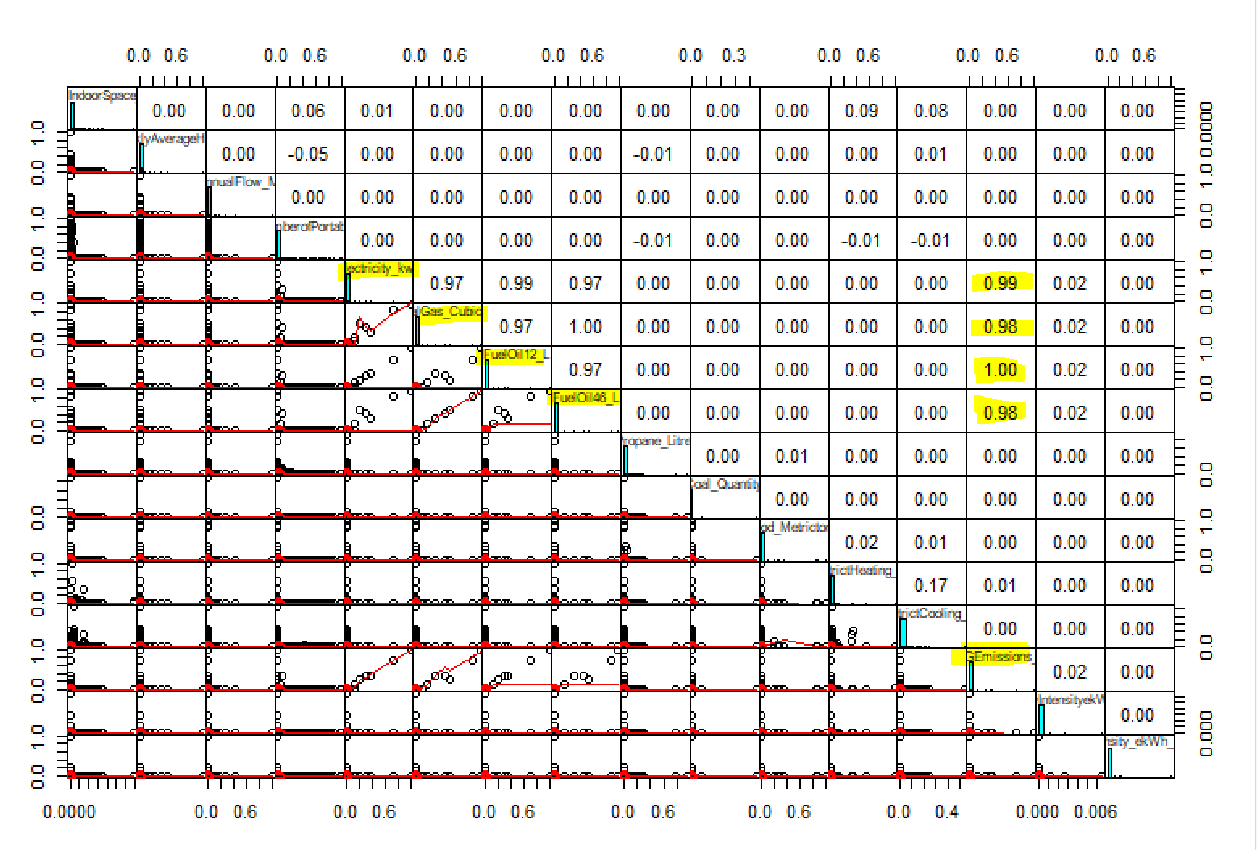
Principal Component Analysis was not an efficient method to reduce the number of dimension as it showed almost the same number of attributes at the conclusion of the analysis. Dimensionality reduction using Forward selection worked well. It showed that GHG can be determined from the attributes Electricity\_kwh, NaturalGas\_Cubicmeter, FuelOil12\_L , FuelOil46\_L, DistrictHeating\_GJ, DistrictCooling\_GJ and TypeofSector.

Correlation analysis showed a strong positive linear relationship between GHG and Electricity\_kwh, NaturalGas\_Cubicmeter, FuelOil12\_L, FuelOil46\_L (Chart 3) with correlations of 0.99, 0.98, 1.00, 0.98 respectively.

**Chart 1: Municipal Sector – Variation of GHG across operation types for 2011-2014**

**Chart 2: Post-Secondary Educational Sector – Variation of GHG across operation types for 2011-2014**.

**Chart 3: Correlation between Attributes**



The simple linear regression model algorithm treated TypeofSector as numerical although it is a factor and therefore this attribute was excluded in models 2 and 3.

train.set$GHGEmissions\_KG ~ train.set$FuelOil12\_L + train.set$NaturalGas\_Cubicmeter + train.set$DistrictHeating\_GJ + train.set$FuelOil46\_L + train.set$DistrictCooling\_GJ + train.set$Electricity\_kwh

In the glm model Akaike Information Criteria AIC is -1034690 which is low (good).

Use of simple linear regression and general linear regression gave the same root mean square error rmse 0.013633. RMSE is the square root of the variance of the residuals. It shows how closely the predicted value matches the actual value. It is a good measure of the fit of the model.

This low value for rmse shows that the model has high accuracy.

Both lm and glm gave the same intercept and coefficients. It can be concluded that GHG can be calculated using the formula equation

GHGEmissions\_KG = -4.054311e-07 + 7.441562e-01 (FuelOil12\_L) + 2.968370e-02(NaturalGas\_Cubicmeter) + 1.041156e-02 (DistrictHeating\_GJ) + 1.779203e-01 (FuelOil46\_L) + -3.223940e-03 (DistrictCooling\_GJ) + 5.522074e-02(Electricity\_kwh)

# Conclusions

Schools in the province as a whole emit the largest amount of GHG. Inspection of the raw data showed a big spread in GHG emission across organizations of the same operation type. Floor area and volume of water do not directly influence the GHG produced.

This study shows that the quantity of GHG emitted is ultimately dependent on the amount of energy consumed. The use of coal, wood and renewables are insignificant in quantity and therefore are not major contributors to GHG. Significant contributors to GHG are the quantity of electricity, gas and fuel oil used for heating and cooling.

The following equation can be used to predict the GHG for any sector and operation type if we insert the values for the independent variables.

GHGEmissions\_KG = -4.054311e-07 + 7.441562e-01 (FuelOil12\_L) + 2.968370e-02(NaturalGas\_Cubicmeter) + 1.041156e-02 (DistrictHeating\_GJ) + 1.779203e-01 (FuelOil46\_L) + -3.223940e-03 (DistrictCooling\_GJ) + 5.522074e-02(Electricity\_kwh)