

# Reverse Engineer the Formula for Air Quality Index in Calgary

Group

2024-11-25

```
air <- read.csv("/Users/shahzaibrahat/Downloads/Air.csv")

air$Season <- as.factor(air$Season)
air$Station.Name <- as.factor(air$Station.Name)

fullmodel <- lm(Air.Quality.Index ~ Station.Name + Season + Carbon.Monoxide + Methane + Nitric.Oxide + Nitrogen.Dioxide + PM2.5.Mass + Total.Hydrocarbons + Ozone + Non.methane.Hydrocarbons + Total.Oxides.Of.Nitrogen, data = air)

summary(fullmodel)

## 
## Call:
## lm(formula = Air.Quality.Index ~ Station.Name + Season + Carbon.Monoxide +
##     Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone + Non.methane.Hydrocarbons +
##     PM2.5.Mass + Total.Hydrocarbons + Total.Oxides.Of.Nitrogen,
##     data = air)
## 
## Residuals:
##      Min        1Q        Median        3Q        Max 
## -1.91256 -0.08855  0.00315  0.09043  2.00679 
## 
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)                0.7893546  0.0457078 17.270 < 2e-16 ***
## Station.NameCalgary Northwest -0.0407164  0.0080189 -5.078 4.03e-07 ***
## Station.NameCalgary Southeast -0.0647311  0.0094941 -6.818 1.09e-11 ***
## SeasonSpring                 0.0006750  0.0093561  0.072  0.94249  
## SeasonSummer                  0.0169526  0.0095345  1.778  0.07549  
## SeasonWinter                  0.0117499  0.0093987  1.250  0.21133  
## Carbon.Monoxide               0.5202057  0.0577702  9.005 < 2e-16 ***
## Methane                       8.4682907  2.7434515  3.087  0.00204 ** 
## Nitric.Oxide                  -8.9544223  2.9701025 -3.015  0.00259 ** 
## Nitrogen.Dioxide                7.0397357  2.7627719  2.548  0.01088 *  
## Ozone                          40.1656896  0.4176181 96.178 < 2e-16 ***
## Non.methane.Hydrocarbons       8.6813278  2.7450182  3.163  0.00158 ** 
## PM2.5.Mass                     0.0357418  0.0004993 71.586 < 2e-16 ***
## Total.Hydrocarbons              -8.3709886  2.7439713 -3.051  0.00230 ** 
## Total.Oxides.Of.Nitrogen        12.8073747  2.8791341  4.448 8.93e-06 *** 
## --- 
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 0.1677 on 3415 degrees of freedom
## Multiple R-squared:  0.8865, Adjusted R-squared:  0.886 
## F-statistic:  1905 on 14 and 3415 DF,  p-value: < 2.2e-16
```

```

str(air)

## 'data.frame': 3430 obs. of 14 variables:
## $ Date : chr "2015-04-07" "2015-04-08" "2015-04-09" "2015-04-10" ...
## $ Season : Factor w/ 4 levels "Fall","Spring",...: 2 2 2 2 2 2 2 2 ...
## $ Location : chr "(51.029944, -114.008111)" "(51.029944, -114.008111)" "(51.029944,
## $ Station.Name : Factor w/ 3 levels "Calgary Central-Inglewood",...: 1 1 1 1 1 1 1 1 1 1 1 1 1 1
## $ Air.Quality.Index : num 1.55 2.73 2.9 3.11 2.88 ...
## $ Carbon.Monoxide : num 0.183 0.179 0.188 0.221 0.133 ...
## $ Methane : num 1.93 1.94 1.92 1.93 1.89 ...
## $ Nitric.Oxide : num 0.00275 0.0045 0.003 0.0154 0.0016 0.0006 0.0263 0.0103 0.0013 0.0006
## $ Nitrogen.Dioxide : num 0.0107 0.0133 0.0189 0.0201 0.013 0.0087 0.0259 0.0194 0.0074 0.0064
## $ Non.methane.Hydrocarbons: num 0 0 0 0 0 0 0 0 0 0 0 0 0 ...
## $ Ozone : num 0.0303 0.0279 0.0297 0.0271 0.0329 0.0346 0.0231 0.0242 0.0358 0.0368
## $ PM2.5.Mass : num 6.38 6.38 8.88 10.71 6.71 ...
## $ Total.Hydrocarbons : num 1.93 1.94 1.92 1.93 1.89 ...
## $ Total.Oxides.Of.Nitrogen: num 0.0133 0.0178 0.0219 0.0355 0.0147 0.0093 0.0522 0.0297 0.0087 0.0087

fullmodel1 <- lm(
  Air.Quality.Index ~ Station.Name + Season + Carbon.Monoxide +
  Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone +
  Non.methane.Hydrocarbons + PM2.5.Mass +
  Total.Hydrocarbons + Total.Oxides.Of.Nitrogen,
  data = air
)
summary(fullmodel1)

##
## Call:
## lm(formula = Air.Quality.Index ~ Station.Name + Season + Carbon.Monoxide +
##     Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone + Non.methane.Hydrocarbons +
##     PM2.5.Mass + Total.Hydrocarbons + Total.Oxides.Of.Nitrogen,
##     data = air)
##
## Residuals:
##      Min        1Q        Median       3Q        Max 
## -1.91256 -0.08855  0.00315  0.09043  2.00679 
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)             0.7893546  0.0457078 17.270 < 2e-16 ***
## Station.NameCalgary Northwest -0.0407164  0.0080189 -5.078 4.03e-07 ***
## Station.NameCalgary Southeast -0.0647311  0.0094941 -6.818 1.09e-11 ***
## SeasonSpring            0.0006750  0.0093561  0.072  0.94249  
## SeasonSummer             0.0169526  0.0095345  1.778  0.07549 .  
## SeasonWinter             0.0117499  0.0093987  1.250  0.21133  
## Carbon.Monoxide          0.5202057  0.0577702  9.005 < 2e-16 ***
## Methane                  8.4682907  2.7434515  3.087  0.00204 ** 
## Nitric.Oxide              -8.9544223  2.9701025 -3.015  0.00259 ** 
## Nitrogen.Dioxide          7.0397357  2.7627719  2.548  0.01088 *  
## Ozone                     40.1656896  0.4176181  96.178 < 2e-16 ***
## Non.methane.Hydrocarbons 8.6813278  2.7450182  3.163  0.00158 ** 
## PM2.5.Mass                0.0357418  0.0004993  71.586 < 2e-16 ***

```

```

## Total.Hydrocarbons      -8.3709886  2.7439713  -3.051  0.00230 ***
## Total.Oxides.Of.Nitrogen 12.8073747  2.8791341   4.448  8.93e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1677 on 3415 degrees of freedom
## Multiple R-squared:  0.8865, Adjusted R-squared:  0.886
## F-statistic: 1905 on 14 and 3415 DF,  p-value: < 2.2e-16

```

## Finding the best variables using stepwise section procedure :

```

library(olsrr)

## 
## Attaching package: 'olsrr'

## The following object is masked from 'package:datasets':
## 
##      rivers

stepmod=ols_step_both_p(fullmodel1,p_enter = 0.1, p_remove = 0.3, details=TRUE)

## Stepwise Selection Method
## -----
## 
## Candidate Terms:
## 
## 1. Station.Name
## 2. Season
## 3. Carbon.Monoxide
## 4. Methane
## 5. Nitric.Oxide
## 6. Nitrogen.Dioxide
## 7. Ozone
## 8. Non.methane.Hydrocarbons
## 9. PM2.5.Mass
## 10. Total.Hydrocarbons
## 11. Total.Oxides.Of.Nitrogen
## 
## 
## Step    => 0
## Model   => Air.Quality.Index ~ 1
## R2      => 0
## 
## Initiating stepwise selection...
## 
## Step      => 1
## Selected  => PM2.5.Mass
## Model     => Air.Quality.Index ~ PM2.5.Mass
## R2        => 0.46

```

```

## Step      => 2
## Selected => Ozone
## Model     => Air.Quality.Index ~ PM2.5.Mass + Ozone
## R2        => 0.715
##
## Step      => 3
## Selected => Total.Oxides.Of.Nitrogen
## Model     => Air.Quality.Index ~ PM2.5.Mass + Ozone + Total.Oxides.Of.Nitrogen
## R2        => 0.871
##
## Step      => 4
## Selected => Nitric.Oxide
## Model     => Air.Quality.Index ~ PM2.5.Mass + Ozone + Total.Oxides.Of.Nitrogen + Nitric.Oxide
## R2        => 0.881
##
## Step      => 5
## Selected => Carbon.Monoxide
## Model     => Air.Quality.Index ~ PM2.5.Mass + Ozone + Total.Oxides.Of.Nitrogen + Nitric.Oxide + Carbon
## R2        => 0.883
##
## Step      => 6
## Selected => Methane
## Model     => Air.Quality.Index ~ PM2.5.Mass + Ozone + Total.Oxides.Of.Nitrogen + Nitric.Oxide + Carbon
## R2        => 0.884
##
## Step      => 7
## Selected => Station.Name
## Model     => Air.Quality.Index ~ PM2.5.Mass + Ozone + Total.Oxides.Of.Nitrogen + Nitric.Oxide + Carbon
## R2        => 0.885
##
## Step      => 8
## Selected => Non.methane.Hydrocarbons
## Model     => Air.Quality.Index ~ PM2.5.Mass + Ozone + Total.Oxides.Of.Nitrogen + Nitric.Oxide + Carbon
## R2        => 0.886
##
## Step      => 9
## Selected => Total.Hydrocarbons
## Model     => Air.Quality.Index ~ PM2.5.Mass + Ozone + Total.Oxides.Of.Nitrogen + Nitric.Oxide + Carbon
## R2        => 0.886
##
## Step      => 10
## Selected => Nitrogen.Dioxide
## Model     => Air.Quality.Index ~ PM2.5.Mass + Ozone + Total.Oxides.Of.Nitrogen + Nitric.Oxide + Carbon
## R2        => 0.886
##
## No more variables to be added or removed.

```

from this model i can see all my variables are good except for one variable , that is Season. which is very unusual.

## First reduced model:

```
red1 <- lm(  
  Air.Quality.Index ~ Station.Name + Carbon.Monoxide +  
  Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone +  
  Non.methane.Hydrocarbons + PM2.5.Mass +  
  Total.Hydrocarbons + Total.Oxides.Of.Nitrogen,  
  data = air  
)  
summary(red1)  
  
##  
## Call:  
## lm(formula = Air.Quality.Index ~ Station.Name + Carbon.Monoxide +  
##       Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone + Non.methane.Hydrocarbons +  
##       PM2.5.Mass + Total.Hydrocarbons + Total.Oxides.Of.Nitrogen,  
##       data = air)  
##  
## Residuals:  
##      Min        1Q     Median      3Q      Max  
## -1.91816 -0.08987  0.00283  0.09034  2.01554  
##  
## Coefficients:  
##                               Estimate Std. Error t value Pr(>|t|)  
## (Intercept)             0.7987853  0.0449516 17.770 < 2e-16 ***  
## Station.NameCalgary Northwest -0.0412573  0.0078878 -5.231 1.79e-07 ***  
## Station.NameCalgary Southeast -0.0657921  0.0094628 -6.953 4.27e-12 ***  
## Carbon.Monoxide          0.5241457  0.0576891  9.086 < 2e-16 ***  
## Methane                  8.5367609  2.7440588  3.111  0.00188 **  
## Nitric.Oxide             -8.2907247  2.9494466 -2.811  0.00497 **  
## Nitrogen.Dioxide          7.5727866  2.7480005  2.756  0.00589 **  
## Ozone                     40.1497183 0.3644937 110.152 < 2e-16 ***  
## Non.methane.Hydrocarbons 8.7704148  2.7456525  3.194  0.00141 **  
## PM2.5.Mass                0.0358271  0.0004866  73.621 < 2e-16 ***  
## Total.Hydrocarbons        -8.4403314  2.7446024 -3.075  0.00212 **  
## Total.Oxides.Of.Nitrogen  12.1444148  2.8641698  4.240 2.29e-05 ***  
## ---  
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 0.1678 on 3418 degrees of freedom  
## Multiple R-squared:  0.8863, Adjusted R-squared:  0.8859  
## F-statistic:  2422 on 11 and 3418 DF,  p-value: < 2.2e-16
```

all the variables are significant as seen in the summary.

## 2nd reduced model:

```
library(leaps)  
best.subsetExecSal=regsubsets(Air.Quality.Index ~ Station.Name + Season + Carbon.Monoxide +
```

```

Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone +
Non.methane.Hydrocarbons + PM2.5.Mass +
Total.Hydrocarbons + Total.Oxides.Of.Nitrogen, data= air, nv=14 )
summary(best.subsetExecSal)

## Subset selection object
## Call: regsubsets.formula(Air.Quality.Index ~ Station.Name + Season +
##   Carbon.Monoxide + Methane + Nitric.Oxide + Nitrogen.Dioxide +
##   Ozone + Non.methane.Hydrocarbons + PM2.5.Mass + Total.Hydrocarbons +
##   Total.Oxides.Of.Nitrogen, data = air, nv = 14)
## 14 Variables (and intercept)
##          Forced in Forced out
## Station.NameCalgary Northwest FALSE FALSE
## Station.NameCalgary Southeast FALSE FALSE
## SeasonSpring FALSE FALSE
## SeasonSummer FALSE FALSE
## SeasonWinter FALSE FALSE
## Carbon.Monoxide FALSE FALSE
## Methane FALSE FALSE
## Nitric.Oxide FALSE FALSE
## Nitrogen.Dioxide FALSE FALSE
## Ozone FALSE FALSE
## Non.methane.Hydrocarbons FALSE FALSE
## PM2.5.Mass FALSE FALSE
## Total.Hydrocarbons FALSE FALSE
## Total.Oxides.Of.Nitrogen FALSE FALSE
## 1 subsets of each size up to 14
## Selection Algorithm: exhaustive
##           Station.NameCalgary Northwest Station.NameCalgary Southeast
## 1  ( 1 )   " "
## 2  ( 1 )   " "
## 3  ( 1 )   " "
## 4  ( 1 )   " "
## 5  ( 1 )   " "
## 6  ( 1 )   " "
## 7  ( 1 )   "*"
## 8  ( 1 )   "*"
## 9  ( 1 )   "*"
## 10 ( 1 )   "*"
## 11 ( 1 )   "*"
## 12 ( 1 )   "*"
## 13 ( 1 )   "*"
## 14 ( 1 )   "*"
##           SeasonSpring SeasonSummer SeasonWinter Carbon.Monoxide Methane
## 1  ( 1 )   " "      " "      " "      " "      " "
## 2  ( 1 )   " "      " "      " "      " "      " "
## 3  ( 1 )   " "      " "      " "      " "      " "
## 4  ( 1 )   " "      " "      " "      " "      " "
## 5  ( 1 )   " "      " "      " "      "*"     " "
## 6  ( 1 )   " "      " "      " "      "*"     " "
## 7  ( 1 )   " "      " "      " "      "*"     " "
## 8  ( 1 )   " "      " "      " "      "*"     " "
## 9  ( 1 )   " "      "*      " "      "*"     " "

```

```

## 10  ( 1 ) " "      " "      " *"      " *"
## 11  ( 1 ) " "      " "      " *"      " *"
## 12  ( 1 ) " "      " *"      " "      " *"
## 13  ( 1 ) " "      " *"      " *"      " *"
## 14  ( 1 ) " *"      " *"      " *"      " *"
##
##          Nitric.Oxide Nitrogen.Dioxide Ozone Non.methane.Hydrocarbons
## 1  ( 1 ) " "      " "      " "      " "
## 2  ( 1 ) " "      " "      " *"      " "
## 3  ( 1 ) " "      " "      " *"      " "
## 4  ( 1 ) " *"      " "      " *"      " "
## 5  ( 1 ) " *"      " "      " *"      " "
## 6  ( 1 ) " "      " *"      " *"      " "
## 7  ( 1 ) " *"      " "      " *"      " "
## 8  ( 1 ) " *"      " "      " *"      " "
## 9  ( 1 ) " *"      " "      " *"      " "
## 10 ( 1 ) " *"      " "      " *"      " *"
## 11 ( 1 ) " *"      " *"      " *"      " *"
## 12 ( 1 ) " *"      " *"      " *"      " *"
## 13 ( 1 ) " *"      " *"      " *"      " *"
## 14 ( 1 ) " *"      " *"      " *"      " *"
##
##          PM2.5.Mass Total.Hydrocarbons Total.Oxides.Of.Nitrogen
## 1  ( 1 ) " *"      " "      " "
## 2  ( 1 ) " *"      " "      " "
## 3  ( 1 ) " *"      " "      " *"
## 4  ( 1 ) " *"      " "      " *"
## 5  ( 1 ) " *"      " "      " *"
## 6  ( 1 ) " *"      " *"      " *"
## 7  ( 1 ) " *"      " *"      " *"
## 8  ( 1 ) " *"      " *"      " *"
## 9  ( 1 ) " *"      " *"      " *"
## 10 ( 1 ) " *"      " *"      " *"
## 11 ( 1 ) " *"      " *"      " *"
## 12 ( 1 ) " *"      " *"      " *"
## 13 ( 1 ) " *"      " *"      " *"
## 14 ( 1 ) " *"      " *"      " *"

reg.summary=summary(best.subsetExecSal)
rsquare=c(reg.summary$rsq)
cp=c(reg.summary$cp)
AdjustedR=c(reg.summary$adjr2)
RSS=c(reg.summary$rss)
BIC=c(reg.summary$bic)
cbind(rsquare, cp, BIC, RSS, AdjustedR)

##           rsquare          cp          BIC          RSS AdjustedR
## [1,] 0.4599114 12823.52743 -2096.675 457.13025 0.4597539
## [2,] 0.7152136 5144.30464 -4283.714 241.04278 0.7150474
## [3,] 0.8711398 454.98890 -6995.602 109.06711 0.8710270
## [4,] 0.8812273 153.48841 -7267.063 100.52906 0.8810886
## [5,] 0.8832921 93.36438 -7319.077 98.78139 0.8831217
## [6,] 0.8840034 73.96549 -7331.903 98.17940 0.8838000
## [7,] 0.8848376 50.86528 -7348.521 97.47328 0.8846020
## [8,] 0.8856154 29.46498 -7363.624 96.81499 0.8853479
## [9,] 0.8857984 25.95767 -7360.977 96.66005 0.8854979

```

```

## [10,] 0.8860363 20.80228 -7359.987 96.45876 0.8857029
## [11,] 0.8862889 15.20104 -7359.459 96.24492 0.8859229
## [12,] 0.8864347 12.81434 -7355.720 96.12152 0.8860359
## [13,] 0.8864948 13.00520 -7349.396 96.07062 0.8860629
## [14,] 0.8864950 15.00000 -7341.261 96.07048 0.8860297

```

From the above matrix, we can see all models with 12 or more variables have very high adjusted R-squared. It does not matter which variables we choose above 12 because the only difference between the 12- and 14-variable models is the consideration of the Season variable as `SeasonSpring`, `SeasonSummer`, and `SeasonWinter`. So, we have to add Season as a variable in the model according to the above matrix.

OK, now I have two models that work: one without season as a variable and another with season in the model.

**ok now i have 2 models that works. one without season as a variable and another with seson in the model. # 2nd reduced model:**

```

red2 <- lm(
  Air.Quality.Index ~ Station.Name + Season + Carbon.Monoxide +
  Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone +
  Non.methane.Hydrocarbons + PM2.5.Mass +
  Total.Hydrocarbons + Total.Oxides.Of.Nitrogen,
  data = air
)
summary(red2)

##
## Call:
## lm(formula = Air.Quality.Index ~ Station.Name + Season + Carbon.Monoxide +
##     Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone + Non.methane.Hydrocarbons +
##     PM2.5.Mass + Total.Hydrocarbons + Total.Oxides.Of.Nitrogen,
##     data = air)
##
## Residuals:
##      Min        1Q        Median        3Q        Max 
## -1.91256 -0.08855  0.00315  0.09043  2.00679 
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)    
## (Intercept)                 0.7893546  0.0457078 17.270 < 2e-16 ***
## Station.NameCalgary Northwest -0.0407164  0.0080189 -5.078 4.03e-07 ***
## Station.NameCalgary Southeast -0.0647311  0.0094941 -6.818 1.09e-11 ***  
## SeasonSpring                  0.0006750  0.0093561  0.072  0.94249  
## SeasonSummer                   0.0169526  0.0095345  1.778  0.07549 .  
## SeasonWinter                   0.0117499  0.0093987  1.250  0.21133  
## Carbon.Monoxide                0.5202057  0.0577702  9.005 < 2e-16 ***
## Methane                         8.4682907  2.7434515  3.087  0.00204 ** 
## Nitric.Oxide                    -8.9544223  2.9701025 -3.015  0.00259 ** 
## Nitrogen.Dioxide                 7.0397357  2.7627719  2.548  0.01088 *  
## Ozone                            40.1656896  0.4176181  96.178 < 2e-16 ***  
## Non.methane.Hydrocarbons        8.6813278  2.7450182  3.163  0.00158 ** 
## PM2.5.Mass                      0.0357418  0.0004993  71.586 < 2e-16 *** 
## Total.Hydrocarbons              -8.3709886  2.7439713 -3.051  0.00230 ** 
## Total.Oxides.Of.Nitrogen         12.8073747  2.8791341  4.448 8.93e-06 *** 

```

```

## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1677 on 3415 degrees of freedom
## Multiple R-squared: 0.8865, Adjusted R-squared: 0.886
## F-statistic: 1905 on 14 and 3415 DF, p-value: < 2.2e-16

```

All of the season variables (SeasonSpring,SeasonSummer,SeasonWinter) is greater than 0.05. so i can conclude that i can drop season.

## Final additive model :

I am choosing the first reduced model as my final additive model, keeping all the variables except for the Season variable as proven above.

```

reduced_model<-lm(Air.Quality.Index ~ Station.Name + Carbon.Monoxide +
  Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone +
  Non.methane.Hydrocarbons + PM2.5.Mass +
  Total.Hydrocarbons + Total.Oxides.Of.Nitrogen,
  data = air)
summary(reduced_model)

##
## Call:
## lm(formula = Air.Quality.Index ~ Station.Name + Carbon.Monoxide +
##     Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone + Non.methane.Hydrocarbons +
##     PM2.5.Mass + Total.Hydrocarbons + Total.Oxides.Of.Nitrogen,
##     data = air)
##
## Residuals:
##       Min         1Q       Median        3Q       Max
## -1.91816 -0.08987  0.00283  0.09034  2.01554
##
## Coefficients:
##                               Estimate Std. Error t value Pr(>|t|)
## (Intercept)             0.7987853  0.0449516 17.770 < 2e-16 ***
## Station.NameCalgary Northwest -0.0412573  0.0078878 -5.231 1.79e-07 ***
## Station.NameCalgary Southeast -0.0657921  0.0094628 -6.953 4.27e-12 ***
## Carbon.Monoxide          0.5241457  0.0576891  9.086 < 2e-16 ***
## Methane                  8.5367609  2.7440588  3.111  0.00188 **
## Nitric.Oxide              -8.2907247  2.9494466 -2.811  0.00497 **
## Nitrogen.Dioxide          7.5727866  2.7480005  2.756  0.00589 **
## Ozone                     40.1497183  0.3644937 110.152 < 2e-16 ***
## Non.methane.Hydrocarbons 8.7704148  2.7456525  3.194  0.00141 **
## PM2.5.Mass                0.0358271  0.0004866  73.621 < 2e-16 ***
## Total.Hydrocarbons        -8.4403314  2.7446024 -3.075  0.00212 **
## Total.Oxides.Of.Nitrogen 12.1444148  2.8641698  4.240 2.29e-05 ***
##
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1678 on 3418 degrees of freedom
## Multiple R-squared: 0.8863, Adjusted R-squared: 0.8859

```

```
## F-statistic: 2422 on 11 and 3418 DF, p-value: < 2.2e-16
```

The final regression equation for predicting `Air.Quality.Index` is:

```
Air.Quality.Index = 0.7988 - 0.0413(Station.Name: Calgary Northwest) - 0.0658(Station.Name: Calgary Southeast)  
+ 0.5241(Carbon Monoxide) + 8.5368(Methane) - 8.2907(Nitric Oxide) + 7.5728(Nitrogen Dioxide) + 40.1497(Ozone)  
+ 8.7704(Non-methane Hydrocarbons) + 0.0358(PM2.5 Mass) - 8.4403(Total Hydrocarbons) + 12.1444(Total Oxides of Nitroge
```

#### Notes:

##### 1. Station Name Encoding:

- `Station.Name: Calgary Northwest` and `Station.Name: Calgary Southeast` are dummy variables for categorical encoding. The reference category is the baseline (likely another station not shown in the coefficients).

##### 2. Significance:

- Variables with \*\*\* ( $p < 0.001$ ) are highly significant in the model.

##### 3. Interpretation:

- Positive coefficients indicate a positive relationship with the Air Quality Index.
- Negative coefficients indicate a negative relationship.

---

This equation represents your final reduced regression model based on the data and significant predictors.

All the variables are significant with the p-value less than  $\alpha = 0.05$  level of significance.

## Checking for multicollinearity

```
library(ggplot2)  
library(lmtest)  
  
## Loading required package: zoo  
  
##  
## Attaching package: 'zoo'  
  
## The following objects are masked from 'package:base':  
##  
##     as.Date, as.Date.numeric  
  
library(mctest)  
library(car)  
  
## Loading required package: carData
```

```

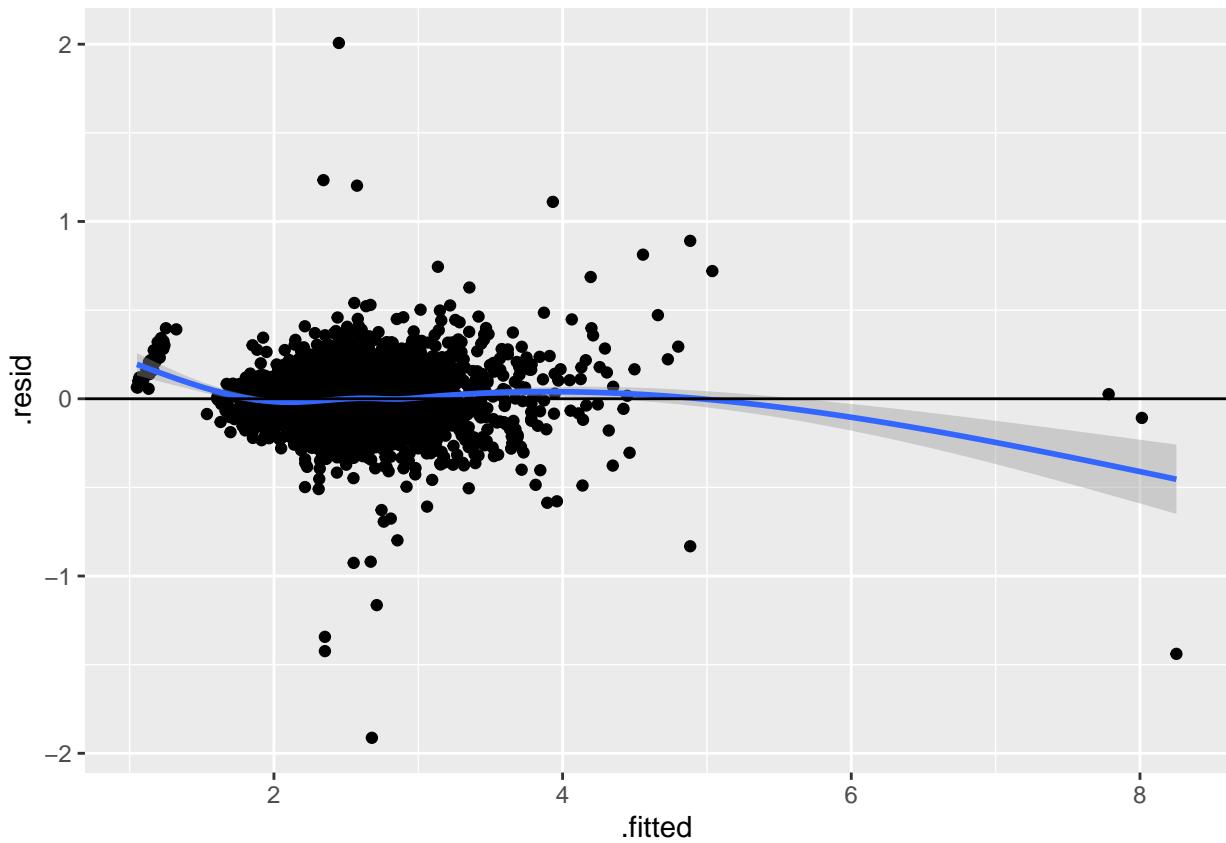
library(GGally)

## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2

ggplot(fullmodel, aes(x=.fitted, y=.resid)) +geom_point() +geom_smooth() +geom_hline(yintercept = 0)

## `geom_smooth()` using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'

```



```
imcdiag(fullmodel, method="VIF")
```

```

##
## Call:
## imcdiag(mod = fullmodel, method = "VIF")
##
##
## VIF Multicollinearity Diagnostics
##
##                                     VIF detection
## Station.NameCalgary Northwest      1.7496      0
## Station.NameCalgary Southeast       2.4557      0
## SeasonSpring                        2.2018      0

```

```

## SeasonSummer           2.0297    0
## SeasonWinter          1.9008    0
## Carbon.Monoxide       4.2949    0
## Methane                30901.7792   1
## Nitric.Oxide          161.6668   1
## Nitrogen.Dioxide      72.8051   1
## Ozone                  2.2808    0
## Non.methane.Hydrocarbons 2471.6600   1
## PM2.5.Mass            1.7993    0
## Total.Hydrocarbons    40564.0524   1
## Total.Oxides.Of.Nitrogen 400.3398   1
##
## Multicollinearity may be due to Methane Nitric.Oxide Nitrogen.Dioxide Non.methane.Hydrocarbons Total
##
## 1 --> COLLINEARITY is detected by the test
## 0 --> COLLINEARITY is not detected by the test
##
## =====

```

The variables Methane, Nitric.Oxide, Nitrogen.Dioxide, Non.methane.Hydrocarbons, Total.Hydrocarbons, and Total.Oxides.Of.Nitrogen show high collinearity .

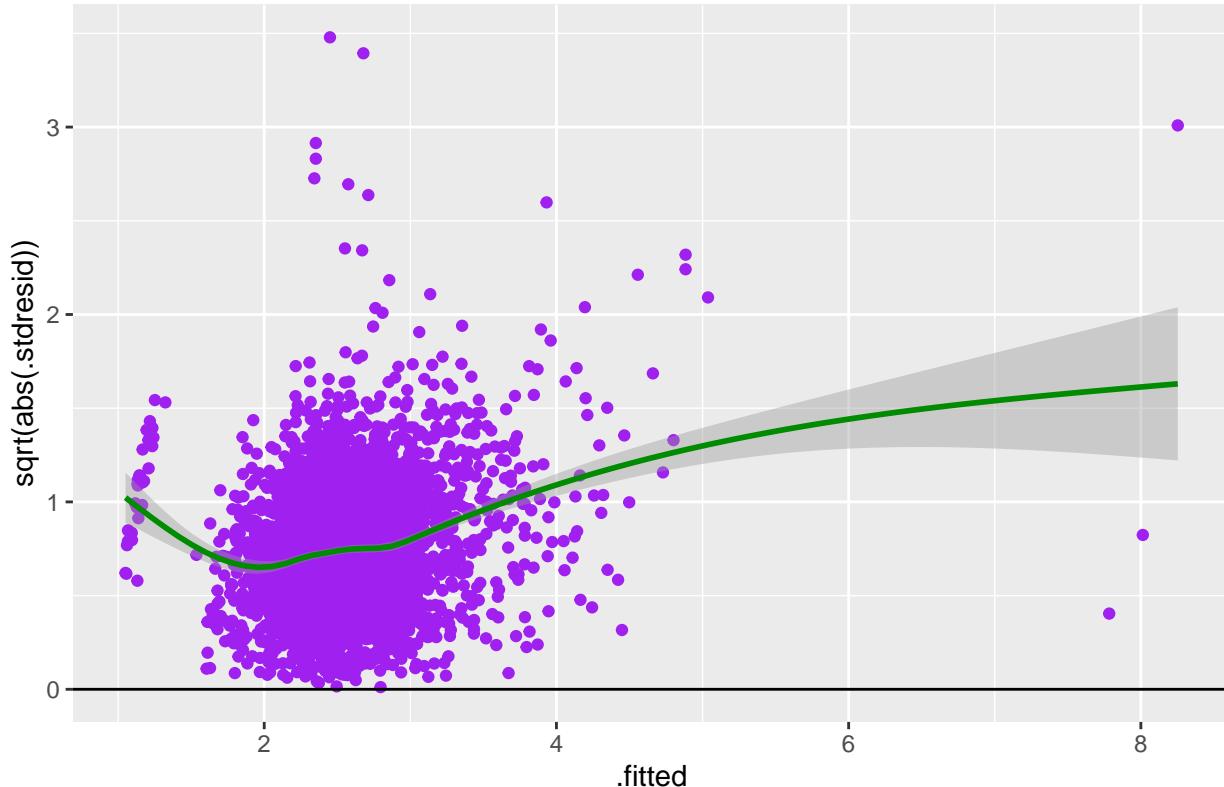
```

ggplot(fullmodel, aes(x=.fitted, y=sqrt(abs(.stdresid)))) +
  geom_point(colour = "purple") +
  geom_hline(yintercept = 0) +
  geom_smooth( colour = "green4")+
  ggtitle("Scale-Location plot : Standardized Residual vs Fitted values")

```

```
## `geom_smooth()` using method = 'gam' and formula = 'y ~ s(x, bs = "cs")'
```

Scale–Location plot : Standardized Residual vs Fitted values

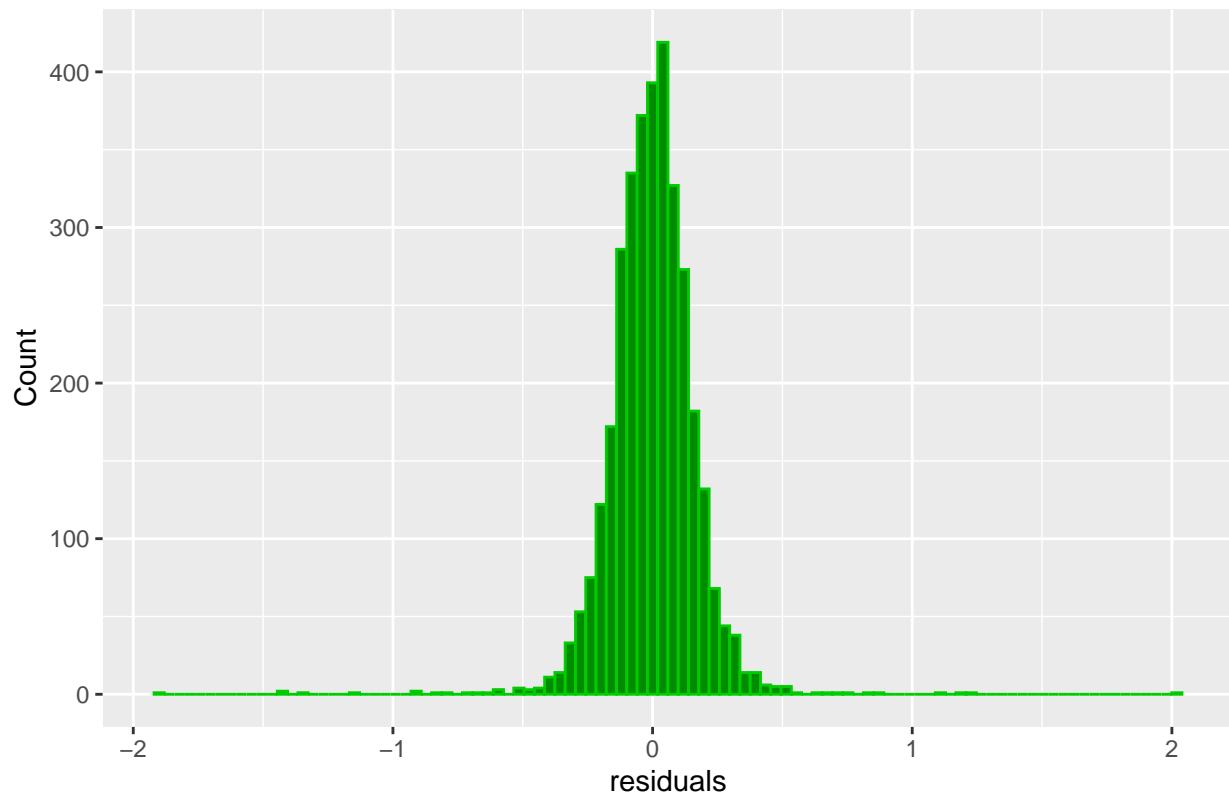


## Checking for the assumptions

The distribution of model residuals should be approximately normal.

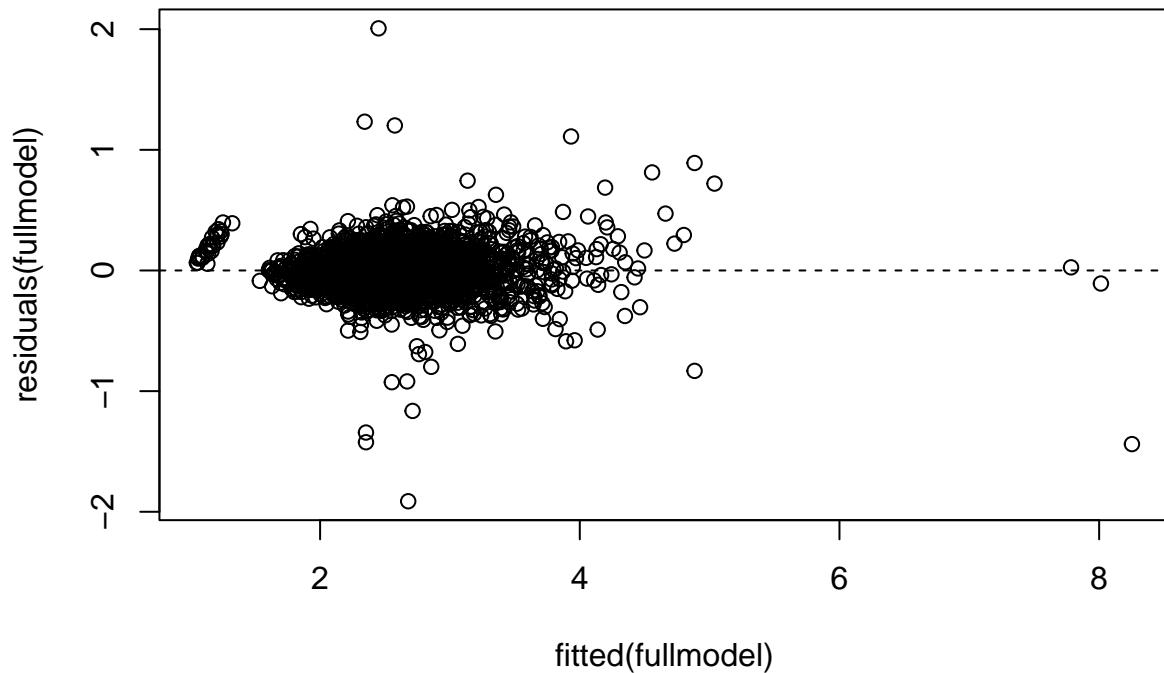
```
ggplot(data=fullmodel, aes(residuals(fullmodel))) +  
  geom_histogram(col="green3", fill="green4", bins = 100) +  
  labs(title="Histogram for residuals") +  
  labs(x="residuals", y="Count")
```

Histogram for residuals



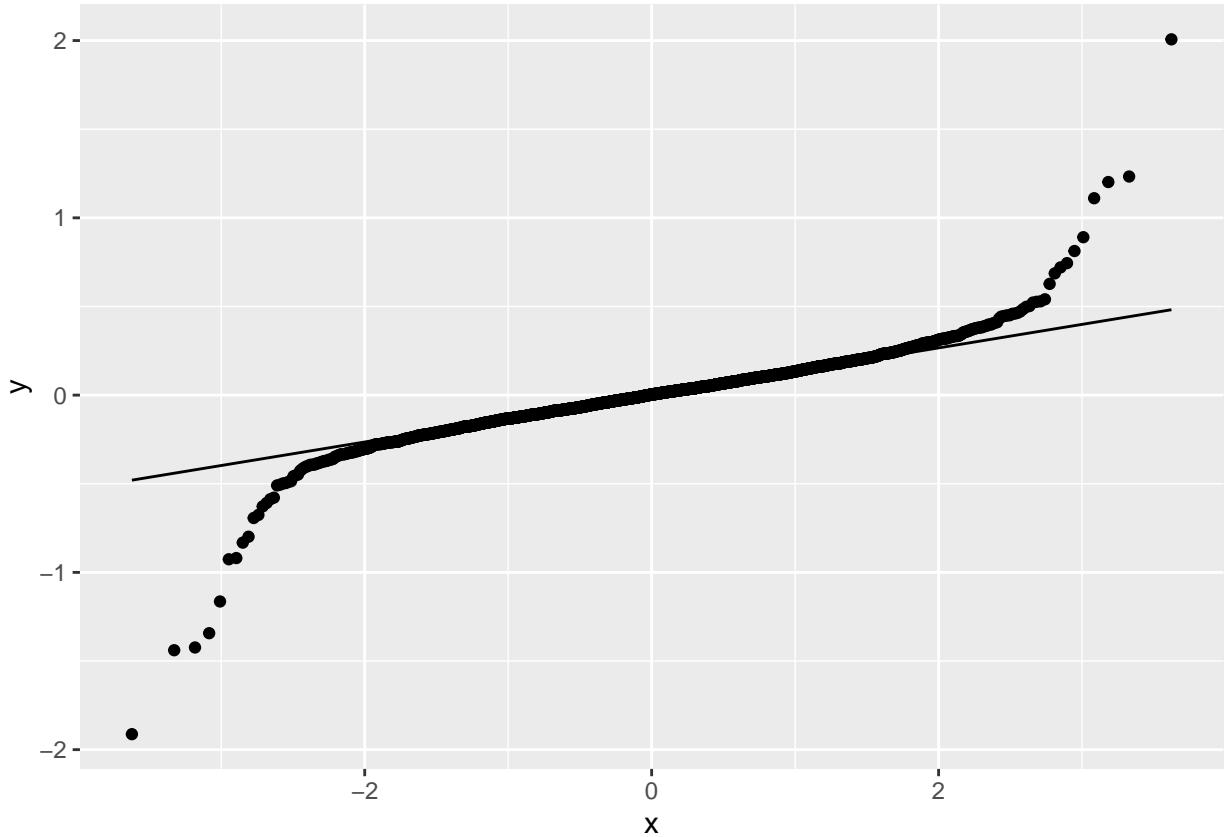
The variance of the residuals should be consistent for all observations.

```
plot(fitted(fullmodel), residuals(fullmodel))
abline(h = 0, lty = 2)
```



QQ plot

```
ggplot(fullmodel, aes(sample=fullmodel$residuals)) +stat_qq() +stat_qq_line()
```



### Breusch-Pagan Test

$H_0$  : the variance of the residuals is constant ( homoscedasticity). \  $H_a$  : the variance of the residuals is not constant .

```
bptest(fullmodel)
```

```
##
## studentized Breusch-Pagan test
##
## data: fullmodel
## BP = 272.6, df = 14, p-value < 2.2e-16
```

The variance of the residuals is not significant and we reject the null hypothesis.

### Shapiro Test

$H_0$  : We have normality \  $H_a$  : We don't have the normality .

```
shapiro.test(residuals(fullmodel))
```

```
##
```

```

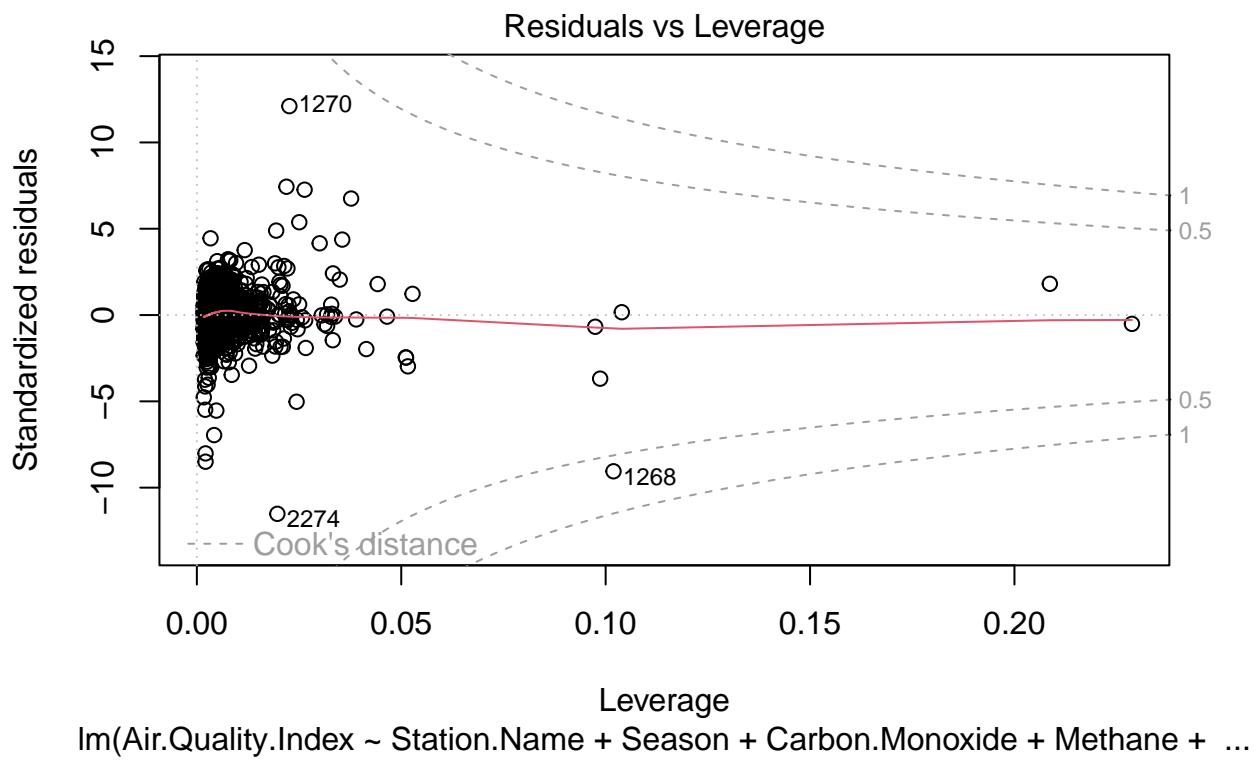
## Shapiro-Wilk normality test
##
## data: residuals(fullmodel)
## W = 0.88906, p-value < 2.2e-16

```

This suggests that the full model is not normally distributed and we reject the null hypothesis.

### cooks distance (checking for outliers)

```
plot(fullmodel,which=5)
```



## ANOVA MODEL

Null hypothesis : none of the predictors are significant alterantive : atleast one of the predictor is signifacnt

```

model1<-lm(Air.Quality.Index ~ 1, data=air)
print(anova(model1,fullmodel))

```

```

## Analysis of Variance Table
##

```

```

## Model 1: Air.Quality.Index ~ 1
## Model 2: Air.Quality.Index ~ Station.Name + Season + Carbon.Monoxide +
##      Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone + Non.methane.Hydrocarbons +
##      PM2.5.Mass + Total.Hydrocarbons + Total.Oxides.Of.Nitrogen
##   Res.Df     RSS Df Sum of Sq    F    Pr(>F)
## 1    3429  846.40
## 2    3415  96.07 14    750.33 1905.1 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

## The interaction model

```

full_interaction_model<-lm(Air.Quality.Index ~ (Station.Name+Carbon.Monoxide+Methane+Nitric.Oxide
+Nitrogen.Dioxide+Ozone
+Non.methane.Hydrocarbons+PM2.5.Mass
+Total.Hydrocarbons+Total.Oxides.Of.Nitrogen)^2,
data=air)
summary(full_interaction_model)

##
## Call:
## lm(formula = Air.Quality.Index ~ (Station.Name + Carbon.Monoxide +
##      Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone + Non.methane.Hydrocarbons +
##      PM2.5.Mass + Total.Hydrocarbons + Total.Oxides.Of.Nitrogen)^2,
##      data = air)
##
## Residuals:
##      Min        1Q        Median        3Q        Max 
## -1.77860 -0.07634  0.00332  0.07583  1.66233 
##
## Coefficients:
##                               Estimate Std. Error
## (Intercept)                -1.550e-01  3.089e-01
## Station.NameCalgary Northwest          4.596e-01  1.710e-01
## Station.NameCalgary Southeast          -2.854e-01  1.588e-01
## Carbon.Monoxide              7.378e-01  8.532e-01
## Methane                     -3.611e+01  1.586e+01
## Nitric.Oxide                 -2.591e+02  8.687e+01
## Nitrogen.Dioxide              -2.370e+02  8.581e+01
## Ozone                        7.628e+01  5.706e+00
## Non.methane.Hydrocarbons      -3.582e+01  1.589e+01
## PM2.5.Mass                   4.518e-02  7.487e-03
## Total.Hydrocarbons            3.663e+01  1.586e+01
## Total.Oxides.Of.Nitrogen      2.724e+02  8.638e+01
## Station.NameCalgary Northwest:Carbon.Monoxide 4.807e-01  1.526e-01
## Station.NameCalgary Southeast:Carbon.Monoxide 2.776e-01  2.248e-01
## Station.NameCalgary Northwest:Methane           1.022e+01  9.027e+00
## Station.NameCalgary Southeast:Methane           3.059e+00  1.039e+01
## Station.NameCalgary Northwest:Nitric.Oxide      2.959e+02  1.058e+02
## Station.NameCalgary Southeast:Nitric.Oxide       2.151e+02  7.963e+01
## Station.NameCalgary Northwest:Nitrogen.Dioxide  2.853e+02  1.058e+02

```

## Station.Name	Calgary	Southeast:Nitrogen.Dioxide	2.032e+02	7.952e+01
## Station.Name	Calgary	Northwest:Ozone	-2.527e+00	9.484e-01
## Station.Name	Calgary	Southeast:Ozone	9.758e-01	1.145e+00
## Station.Name	Calgary	Northwest:Non.methane.Hydrocarbons	1.126e+01	9.023e+00
## Station.Name	Calgary	Southeast:Non.methane.Hydrocarbons	2.688e+00	1.039e+01
## Station.Name	Calgary	Northwest:PM2.5.Mass	-5.996e-03	1.278e-03
## Station.Name	Calgary	Southeast:PM2.5.Mass	-2.686e-04	1.613e-03
## Station.Name	Calgary	Northwest:Total.Hydrocarbons	-1.041e+01	9.023e+00
## Station.Name	Calgary	Southeast:Total.Hydrocarbons	-2.934e+00	1.039e+01
## Station.Name	Calgary	Northwest:Total.Oxides.Of.Nitrogen	-2.971e+02	1.058e+02
## Station.Name	Calgary	Southeast:Total.Oxides.Of.Nitrogen	-2.119e+02	7.958e+01
## Carbon.Monoxide:Methane			7.876e+01	7.295e+01
## Carbon.Monoxide:Nitric.Oxide			8.368e+00	6.240e+01
## Carbon.Monoxide:Nitrogen.Dioxide			4.416e+01	6.202e+01
## Carbon.Monoxide:Ozone			-5.262e+01	6.836e+00
## Carbon.Monoxide:Non.methane.Hydrocarbons			7.543e+01	7.293e+01
## Carbon.Monoxide:PM2.5.Mass			1.793e-03	1.508e-03
## Carbon.Monoxide:Total.Hydrocarbons			-7.827e+01	7.295e+01
## Carbon.Monoxide:Total.Oxides.Of.Nitrogen			-3.414e+01	6.211e+01
## Methane:Nitric.Oxide			6.342e+03	5.191e+03
## Methane:Nitrogen.Dioxide			4.545e+03	5.089e+03
## Methane:Ozone			1.063e+03	3.934e+02
## Methane:Non.methane.Hydrocarbons			-2.522e-01	1.020e+00
## Methane:PM2.5.Mass			1.123e+00	7.756e-01
## Methane:Total.Hydrocarbons			-5.752e-02	4.970e-02
## Methane:Total.Oxides.Of.Nitrogen			-5.488e+03	5.131e+03
## Nitric.Oxide:Nitrogen.Dioxide			-9.942e+02	1.538e+02
## Nitric.Oxide:Ozone			2.149e+02	3.848e+02
## Nitric.Oxide:Non.methane.Hydrocarbons			6.302e+03	5.194e+03
## Nitric.Oxide:PM2.5.Mass			2.587e-01	4.973e-01
## Nitric.Oxide:Total.Hydrocarbons			-6.320e+03	5.191e+03
## Nitric.Oxide:Total.Oxides.Of.Nitrogen			1.284e+02	2.751e+01
## Nitrogen.Dioxide:Ozone			4.914e+01	3.661e+02
## Nitrogen.Dioxide:Non.methane.Hydrocarbons			4.557e+03	5.091e+03
## Nitrogen.Dioxide:PM2.5.Mass			4.387e-01	4.863e-01
## Nitrogen.Dioxide:Total.Hydrocarbons			-4.536e+03	5.088e+03
## Nitrogen.Dioxide:Total.Oxides.Of.Nitrogen			5.577e+02	7.495e+01
## Ozone:Non.methane.Hydrocarbons			1.052e+03	3.938e+02
## Ozone:PM2.5.Mass			-5.091e-01	6.690e-02
## Ozone:Total.Hydrocarbons			-1.074e+03	3.935e+02
## Ozone:Total.Oxides.Of.Nitrogen			-6.419e+01	3.734e+02
## Non.methane.Hydrocarbons:PM2.5.Mass			1.157e+00	7.760e-01
## Non.methane.Hydrocarbons:Total.Hydrocarbons			2.054e-01	8.018e-01
## Non.methane.Hydrocarbons:Total.Oxides.Of.Nitrogen			-5.454e+03	5.134e+03
## PM2.5.Mass:Total.Hydrocarbons			-1.118e+00	7.757e-01
## PM2.5.Mass:Total.Oxides.Of.Nitrogen			-5.641e-01	4.918e-01
## Total.Hydrocarbons:Total.Oxides.Of.Nitrogen			5.465e+03	5.131e+03
##		t value	Pr(> t )	
## (Intercept)		-0.502	0.61585	
## Station.Name	Calgary	Northwest	2.688	0.00723 **
## Station.Name	Calgary	Southeast	-1.798	0.07233 .
## Carbon.Monoxide			0.865	0.38725
## Methane			-2.276	0.02288 *
## Nitric.Oxide			-2.983	0.00288 **

## Nitrogen.Dioxide	-2.762	0.00578	**
## Ozone	13.367	< 2e-16	***
## Non.methane.Hydrocarbons	-2.255	0.02422	*
## PM2.5.Mass	6.035	1.76e-09	***
## Total.Hydrocarbons	2.309	0.02099	*
## Total.Oxides.Of.Nitrogen	3.154	0.00163	**
## Station.NameCalgary Northwest:Carbon.Monoxide	3.150	0.00165	**
## Station.NameCalgary Southeast:Carbon.Monoxide	1.235	0.21705	
## Station.NameCalgary Northwest:Methane	1.132	0.25768	
## Station.NameCalgary Southeast:Methane	0.294	0.76840	
## Station.NameCalgary Northwest:Nitric.Oxide	2.797	0.00519	**
## Station.NameCalgary Southeast:Nitric.Oxide	2.702	0.00693	**
## Station.NameCalgary Northwest:Nitrogen.Dioxide	2.697	0.00702	**
## Station.NameCalgary Southeast:Nitrogen.Dioxide	2.555	0.01067	*
## Station.NameCalgary Northwest:Ozone	-2.664	0.00775	**
## Station.NameCalgary Southeast:Ozone	0.853	0.39397	
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons	1.248	0.21196	
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons	0.259	0.79599	
## Station.NameCalgary Northwest:PM2.5.Mass	-4.690	2.84e-06	***
## Station.NameCalgary Southeast:PM2.5.Mass	-0.166	0.86777	
## Station.NameCalgary Northwest:Total.Hydrocarbons	-1.154	0.24860	
## Station.NameCalgary Southeast:Total.Hydrocarbons	-0.283	0.77757	
## Station.NameCalgary Northwest:Total.Oxides.Of.Nitrogen	-2.809	0.00500	**
## Station.NameCalgary Southeast:Total.Oxides.Of.Nitrogen	-2.663	0.00779	**
## Carbon.Monoxide:Methane	1.080	0.28036	
## Carbon.Monoxide:Nitric.Oxide	0.134	0.89333	
## Carbon.Monoxide:Nitrogen.Dioxide	0.712	0.47648	
## Carbon.Monoxide:Ozone	-7.698	1.80e-14	***
## Carbon.Monoxide:Non.methane.Hydrocarbons	1.034	0.30111	
## Carbon.Monoxide:PM2.5.Mass	1.188	0.23476	
## Carbon.Monoxide:Total.Hydrocarbons	-1.073	0.28334	
## Carbon.Monoxide:Total.Oxides.Of.Nitrogen	-0.550	0.58259	
## Methane:Nitric.Oxide	1.222	0.22191	
## Methane:Nitrogen.Dioxide	0.893	0.37186	
## Methane:Ozone	2.702	0.00693	**
## Methane:Non.methane.Hydrocarbons	-0.247	0.80466	
## Methane:PM2.5.Mass	1.448	0.14779	
## Methane:Total.Hydrocarbons	-1.157	0.24719	
## Methane:Total.Oxides.Of.Nitrogen	-1.070	0.28491	
## Nitric.Oxide:Nitrogen.Dioxide	-6.466	1.15e-10	***
## Nitric.Oxide:Ozone	0.559	0.57648	
## Nitric.Oxide:Non.methane.Hydrocarbons	1.213	0.22506	
## Nitric.Oxide:PM2.5.Mass	0.520	0.60296	
## Nitric.Oxide:Total.Hydrocarbons	-1.218	0.22348	
## Nitric.Oxide:Total.Oxides.Of.Nitrogen	4.668	3.16e-06	***
## Nitrogen.Dioxide:Ozone	0.134	0.89324	
## Nitrogen.Dioxide:Non.methane.Hydrocarbons	0.895	0.37080	
## Nitrogen.Dioxide:PM2.5.Mass	0.902	0.36707	
## Nitrogen.Dioxide:Total.Hydrocarbons	-0.891	0.37273	
## Nitrogen.Dioxide:Total.Oxides.Of.Nitrogen	7.440	1.27e-13	***
## Ozone:Non.methane.Hydrocarbons	2.672	0.00757	**
## Ozone:PM2.5.Mass	-7.609	3.57e-14	***
## Ozone:Total.Hydrocarbons	-2.730	0.00637	**
## Ozone:Total.Oxides.Of.Nitrogen	-0.172	0.86351	

```

## Non.methane.Hydrocarbons:PM2.5.Mass           1.491  0.13612
## Non.methane.Hydrocarbons:Total.Hydrocarbons    0.256  0.79781
## Non.methane.Hydrocarbons:Total.Oxides.Of.Nitrogen -1.062  0.28809
## PM2.5.Mass:Total.Hydrocarbons                 -1.442  0.14946
## PM2.5.Mass:Total.Oxides.Of.Nitrogen            -1.147  0.25142
## Total.Hydrocarbons:Total.Oxides.Of.Nitrogen      1.065  0.28689
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1462 on 3364 degrees of freedom
## Multiple R-squared:  0.915, Adjusted R-squared:  0.9134
## F-statistic: 557.3 on 65 and 3364 DF,  p-value: < 2.2e-16

```

Some of the interactions terms are not significant at 0.05 level of significance, so we will drop those to get the best fit first order interaction model.

```
airinter1 <- lm(Air.Quality.Index ~ factor(Station.Name)+Carbon.Monoxide+Methane+Nitric.Oxide+Nitrogen.Dioxide, data = air)
summary(airinter1)
```

```

##
## Call:
## lm(formula = Air.Quality.Index ~ factor(Station.Name) + Carbon.Monoxide +
##     Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone + Non.methane.Hydrocarbons +
##     PM2.5.Mass + Total.Hydrocarbons + Total.Oxides.Of.Nitrogen +
##     factor(Station.Name) * Carbon.Monoxide + factor(Station.Name) * *
##     Nitric.Oxide + factor(Station.Name) * Nitrogen.Dioxide +
##     factor(Station.Name) * Ozone + factor(Station.Name) * PM2.5.Mass +
##     factor(Station.Name) * Total.Oxides.Of.Nitrogen + Carbon.Monoxide *
##     Ozone + Methane * Ozone + Nitric.Oxide * Nitrogen.Dioxide +
##     Nitric.Oxide * Total.Oxides.Of.Nitrogen + Nitrogen.Dioxide *
##     Total.Oxides.Of.Nitrogen + Ozone * Non.methane.Hydrocarbons +
##     Ozone * PM2.5.Mass + Ozone * Total.Hydrocarbons, data = air)
##
## Residuals:
##       Min     1Q   Median     3Q    Max 
## -1.83850 -0.07865  0.00468  0.07990  1.76976
##
## Coefficients:
## (Intercept)                               Estimate
## factor(Station.Name)Calgary Northwest          5.659e-01
## factor(Station.Name)Calgary Southeast          6.666e-02
## Carbon.Monoxide                            3.901e-03
## Methane                                    1.067e+00
## Nitric.Oxide                                -5.133e+00
## Nitrogen.Dioxide                            -2.174e+02
## Ozone                                       -2.176e+02
## Non.methane.Hydrocarbons                   5.965e+01
## PM2.5.Mass                                 -5.002e+00
## Total.Hydrocarbons                          4.828e-02
## Total.Oxides.Of.Nitrogen                  5.213e+00
## factor(Station.Name)Calgary Northwest:Carbon.Monoxide 2.294e+02
## factor(Station.Name)Calgary Southeast:Carbon.Monoxide 7.226e-01
## 
```

## factor(Station.Name)Calgary Northwest:Nitric.Oxide	3.199e+02
## factor(Station.Name)Calgary Southeast:Nitric.Oxide	2.334e+02
## factor(Station.Name)Calgary Northwest:Nitrogen.Dioxide	3.095e+02
## factor(Station.Name)Calgary Southeast:Nitrogen.Dioxide	2.200e+02
## factor(Station.Name)Calgary Northwest:Ozone	-2.622e+00
## factor(Station.Name)Calgary Southeast:Ozone	-1.200e+00
## factor(Station.Name)Calgary Northwest:PM2.5.Mass	-4.151e-03
## factor(Station.Name)Calgary Southeast:PM2.5.Mass	2.619e-03
## factor(Station.Name)Calgary Northwest:Total.Oxides.Of.Nitrogen	-3.218e+02
## factor(Station.Name)Calgary Southeast:Total.Oxides.Of.Nitrogen	-2.307e+02
## Carbon.Monoxide:Ozone	-4.046e+01
## Methane:Ozone	6.771e+02
## Nitric.Oxide:Nitrogen.Dioxide	-1.086e+03
## Nitric.Oxide:Total.Oxides.Of.Nitrogen	2.199e+01
## Nitrogen.Dioxide:Total.Oxides.Of.Nitrogen	5.703e+02
## Ozone:Non.methane.Hydrocarbons	6.702e+02
## Ozone:PM2.5.Mass	-3.915e-01
## Ozone:Total.Hydrocarbons	-6.808e+02
##	Std. Error
## (Intercept)	7.521e-02
## factor(Station.Name)Calgary Northwest	3.165e-02
## factor(Station.Name)Calgary Southeast	3.512e-02
## Carbon.Monoxide	1.239e-01
## Methane	4.681e+00
## Nitric.Oxide	8.166e+01
## Nitrogen.Dioxide	8.155e+01
## Ozone	3.312e+00
## Non.methane.Hydrocarbons	4.688e+00
## PM2.5.Mass	1.390e-03
## Total.Hydrocarbons	4.683e+00
## Total.Oxides.Of.Nitrogen	8.159e+01
## factor(Station.Name)Calgary Northwest:Carbon.Monoxide	1.274e-01
## factor(Station.Name)Calgary Southeast:Carbon.Monoxide	1.593e-01
## factor(Station.Name)Calgary Northwest:Nitric.Oxide	1.087e+02
## factor(Station.Name)Calgary Southeast:Nitric.Oxide	8.163e+01
## factor(Station.Name)Calgary Northwest:Nitrogen.Dioxide	1.086e+02
## factor(Station.Name)Calgary Southeast:Nitrogen.Dioxide	8.159e+01
## factor(Station.Name)Calgary Northwest:Ozone	8.232e-01
## factor(Station.Name)Calgary Southeast:Ozone	9.378e-01
## factor(Station.Name)Calgary Northwest:PM2.5.Mass	1.106e-03
## factor(Station.Name)Calgary Southeast:PM2.5.Mass	1.161e-03
## factor(Station.Name)Calgary Northwest:Total.Oxides.Of.Nitrogen	1.086e+02
## factor(Station.Name)Calgary Southeast:Total.Oxides.Of.Nitrogen	8.161e+01
## Carbon.Monoxide:Ozone	4.292e+00
## Methane:Ozone	2.330e+02
## Nitric.Oxide:Nitrogen.Dioxide	1.476e+02
## Nitric.Oxide:Total.Oxides.Of.Nitrogen	1.668e+01
## Nitrogen.Dioxide:Total.Oxides.Of.Nitrogen	6.669e+01
## Ozone:Non.methane.Hydrocarbons	2.332e+02
## Ozone:PM2.5.Mass	4.875e-02
## Ozone:Total.Hydrocarbons	2.331e+02
##	t value Pr(> t )
## (Intercept)	7.524 6.77e-14
## factor(Station.Name)Calgary Northwest	2.106 0.035294

## factor(Station.Name)Calgary Southeast	0.111 0.911562
## Carbon.Monoxide	8.610 < 2e-16
## Methane	-1.097 0.272848
## Nitric.Oxide	-2.662 0.007806
## Nitrogen.Dioxide	-2.668 0.007658
## Ozone	18.007 < 2e-16
## Non.methane.Hydrocarbons	-1.067 0.286119
## PM2.5.Mass	34.727 < 2e-16
## Total.Hydrocarbons	1.113 0.265710
## Total.Oxides.Of.Nitrogen	2.812 0.004959
## factor(Station.Name)Calgary Northwest:Carbon.Monoxide	5.670 1.55e-08
## factor(Station.Name)Calgary Southeast:Carbon.Monoxide	2.751 0.005974
## factor(Station.Name)Calgary Northwest:Nitric.Oxide	2.944 0.003263
## factor(Station.Name)Calgary Southeast:Nitric.Oxide	2.859 0.004269
## factor(Station.Name)Calgary Northwest:Nitrogen.Dioxide	2.850 0.004397
## factor(Station.Name)Calgary Southeast:Nitrogen.Dioxide	2.697 0.007031
## factor(Station.Name)Calgary Northwest:Ozone	-3.185 0.001459
## factor(Station.Name)Calgary Southeast:Ozone	-1.279 0.200909
## factor(Station.Name)Calgary Northwest:PM2.5.Mass	-3.753 0.000177
## factor(Station.Name)Calgary Southeast:PM2.5.Mass	2.256 0.024149
## factor(Station.Name)Calgary Northwest:Total.Oxides.Of.Nitrogen	-2.963 0.003067
## factor(Station.Name)Calgary Southeast:Total.Oxides.Of.Nitrogen	-2.827 0.004721
## Carbon.Monoxide:Ozone	-9.426 < 2e-16
## Methane:Ozone	2.906 0.003685
## Nitric.Oxide:Nitrogen.Dioxide	-7.356 2.36e-13
## Nitric.Oxide:Total.Oxides.Of.Nitrogen	1.318 0.187601
## Nitrogen.Dioxide:Total.Oxides.Of.Nitrogen	8.552 < 2e-16
## Ozone:Non.methane.Hydrocarbons	2.874 0.004080
## Ozone:PM2.5.Mass	-8.031 1.33e-15
## Ozone:Total.Hydrocarbons	-2.921 0.003508
##	
## (Intercept)	***
## factor(Station.Name)Calgary Northwest	*
## factor(Station.Name)Calgary Southeast	
## Carbon.Monoxide	***
## Methane	
## Nitric.Oxide	**
## Nitrogen.Dioxide	**
## Ozone	***
## Non.methane.Hydrocarbons	
## PM2.5.Mass	***
## Total.Hydrocarbons	
## Total.Oxides.Of.Nitrogen	**
## factor(Station.Name)Calgary Northwest:Carbon.Monoxide	***
## factor(Station.Name)Calgary Southeast:Carbon.Monoxide	**
## factor(Station.Name)Calgary Northwest:Nitric.Oxide	**
## factor(Station.Name)Calgary Southeast:Nitric.Oxide	**
## factor(Station.Name)Calgary Northwest:Nitrogen.Dioxide	**
## factor(Station.Name)Calgary Southeast:Nitrogen.Dioxide	**
## factor(Station.Name)Calgary Northwest:Ozone	**
## factor(Station.Name)Calgary Southeast:Ozone	
## factor(Station.Name)Calgary Northwest:PM2.5.Mass	***
## factor(Station.Name)Calgary Southeast:PM2.5.Mass	*
## factor(Station.Name)Calgary Northwest:Total.Oxides.Of.Nitrogen	**

```

## factor(Station.Name)Calgary Southeast:Total.Oxides.Of.Nitrogen ***
## Carbon.Monoxide:Ozone ***                                 ***
## Methane:Ozone **                                     **
## Nitric.Oxide:Nitrogen.Dioxide ***                   ***
## Nitric.Oxide:Total.Oxides.Of.Nitrogen               ***
## Nitrogen.Dioxide:Total.Oxides.Of.Nitrogen           ***
## Ozone:Non.methane.Hydrocarbons **                  **
## Ozone:PM2.5.Mass ***                                ***
## Ozone:Total.Hydrocarbons **                         **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1508 on 3398 degrees of freedom
## Multiple R-squared:  0.9087, Adjusted R-squared:  0.9078
## F-statistic:  1090 on 31 and 3398 DF,  p-value: < 2.2e-16

airinter2 <- lm(Air.Quality.Index ~ factor(Station.Name)+Carbon.Monoxide+Methane+Nitric.Oxide+Nitrogen.Dioxide+Ozone+Non.methane.Hydrocarbons+PM2.5.Mass+Total.Hydrocarbons+Total.Oxides.Of.Nitrogen+factor(Station.Name)*Carbon.Monoxide+factor(Station.Name)*Nitric.Oxide+factor(Station.Name)*Nitrogen.Dioxide+factor(Station.Name)*Ozone+factor(Station.Name)*PM2.5.Mass+factor(Station.Name)*Total.Oxides.Of.Nitrogen+Carbon.Monoxide*Ozone+Methane*Ozone+Nitric.Oxide*Nitrogen.Dioxide+Nitrogen.Dioxide*Total.Oxides.Of.Nitrogen+Ozone*Non.methane.Hydrocarbons+Ozone*PM2.5.Mass+Ozone*Total.Hydrocarbons, data = air)
##
## Residuals:
##      Min      1Q      Median      3Q      Max
## -1.83942 -0.07893  0.00454  0.07970  1.77458
##
## Coefficients:
##                               Estimate
## (Intercept)                5.739e-01
## factor(Station.Name)Calgary Northwest        6.599e-02
## factor(Station.Name)Calgary Southeast         1.900e-03
## Carbon.Monoxide             1.057e+00
## Methane                     -5.033e+00
## Nitric.Oxide                -2.195e+02
## Nitrogen.Dioxide             -2.181e+02
## Ozone                        5.924e+01
## Non.methane.Hydrocarbons    -4.893e+00
## PM2.5.Mass                  4.836e-02
## Total.Hydrocarbons          5.108e+00
## Total.Oxides.Of.Nitrogen    2.308e+02
## factor(Station.Name)Calgary Northwest:Carbon.Monoxide 7.167e-01
## factor(Station.Name)Calgary Southeast:Carbon.Monoxide 4.367e-01
## factor(Station.Name)Calgary Northwest:Nitric.Oxide     3.205e+02
## factor(Station.Name)Calgary Southeast:Nitric.Oxide     2.336e+02

```

## factor(Station.Name)Calgary Northwest:Nitrogen.Dioxide	3.106e+02
## factor(Station.Name)Calgary Southeast:Nitrogen.Dioxide	2.206e+02
## factor(Station.Name)Calgary Northwest:Ozone	-2.642e+00
## factor(Station.Name)Calgary Southeast:Ozone	-1.190e+00
## factor(Station.Name)Calgary Northwest:PM2.5.Mass	-4.105e-03
## factor(Station.Name)Calgary Southeast:PM2.5.Mass	2.617e-03
## factor(Station.Name)Calgary Northwest:Total.Oxides.Of.Nitrogen	-3.226e+02
## factor(Station.Name)Calgary Southeast:Total.Oxides.Of.Nitrogen	-2.310e+02
## Carbon.Monoxide:Ozone	-4.006e+01
## Methane:Ozone	6.733e+02
## Nitric.Oxide:Nitrogen.Dioxide	-9.163e+02
## Nitrogen.Dioxide:Total.Oxides.Of.Nitrogen	5.080e+02
## Ozone:Non.methane.Hydrocarbons	6.660e+02
## Ozone:PM2.5.Mass	-3.946e-01
## Ozone:Total.Hydrocarbons	-6.769e+02
##	Std. Error
## (Intercept)	7.497e-02
## factor(Station.Name)Calgary Northwest	3.165e-02
## factor(Station.Name)Calgary Southeast	3.510e-02
## Carbon.Monoxide	1.237e-01
## Methane	4.681e+00
## Nitric.Oxide	8.165e+01
## Nitrogen.Dioxide	8.156e+01
## Ozone	3.298e+00
## Non.methane.Hydrocarbons	4.688e+00
## PM2.5.Mass	1.389e-03
## Total.Hydrocarbons	4.683e+00
## Total.Oxides.Of.Nitrogen	8.159e+01
## factor(Station.Name)Calgary Northwest:Carbon.Monoxide	1.274e-01
## factor(Station.Name)Calgary Southeast:Carbon.Monoxide	1.593e-01
## factor(Station.Name)Calgary Northwest:Nitric.Oxide	1.087e+02
## factor(Station.Name)Calgary Southeast:Nitric.Oxide	8.163e+01
## factor(Station.Name)Calgary Northwest:Nitrogen.Dioxide	1.086e+02
## factor(Station.Name)Calgary Southeast:Nitrogen.Dioxide	8.159e+01
## factor(Station.Name)Calgary Northwest:Ozone	8.232e-01
## factor(Station.Name)Calgary Southeast:Ozone	9.379e-01
## factor(Station.Name)Calgary Northwest:PM2.5.Mass	1.106e-03
## factor(Station.Name)Calgary Southeast:PM2.5.Mass	1.161e-03
## factor(Station.Name)Calgary Northwest:Total.Oxides.Of.Nitrogen	1.086e+02
## factor(Station.Name)Calgary Southeast:Total.Oxides.Of.Nitrogen	8.162e+01
## Carbon.Monoxide:Ozone	4.282e+00
## Methane:Ozone	2.330e+02
## Nitric.Oxide:Nitrogen.Dioxide	7.247e+01
## Nitrogen.Dioxide:Total.Oxides.Of.Nitrogen	4.703e+01
## Ozone:Non.methane.Hydrocarbons	2.332e+02
## Ozone:PM2.5.Mass	4.870e-02
## Ozone:Total.Hydrocarbons	2.331e+02
##	t value Pr(> t )
## (Intercept)	7.655 2.51e-14
## factor(Station.Name)Calgary Northwest	2.085 0.037157
## factor(Station.Name)Calgary Southeast	0.054 0.956822
## Carbon.Monoxide	8.546 < 2e-16
## Methane	-1.075 0.282343
## Nitric.Oxide	-2.688 0.007214

## Nitrogen.Dioxide	-2.674	0.007537
## Ozone	17.960	< 2e-16
## Non.methane.Hydrocarbons	-1.044	0.296728
## PM2.5.Mass	34.824	< 2e-16
## Total.Hydrocarbons	1.091	0.275446
## Total.Oxides.Of.Nitrogen	2.828	0.004705
## factor(Station.Name)Calgary Northwest:Carbon.Monoxide	5.627	1.99e-08
## factor(Station.Name)Calgary Southeast:Carbon.Monoxide	2.741	0.006154
## factor(Station.Name)Calgary Northwest:Nitric.Oxide	2.949	0.003206
## factor(Station.Name)Calgary Southeast:Nitric.Oxide	2.862	0.004237
## factor(Station.Name)Calgary Northwest:Nitrogen.Dioxide	2.860	0.004267
## factor(Station.Name)Calgary Southeast:Nitrogen.Dioxide	2.703	0.006901
## factor(Station.Name)Calgary Northwest:Ozone	-3.210	0.001339
## factor(Station.Name)Calgary Southeast:Ozone	-1.269	0.204468
## factor(Station.Name)Calgary Northwest:PM2.5.Mass	-3.713	0.000208
## factor(Station.Name)Calgary Southeast:PM2.5.Mass	2.254	0.024230
## factor(Station.Name)Calgary Northwest:Total.Oxides.Of.Nitrogen	-2.970	0.003000
## factor(Station.Name)Calgary Southeast:Total.Oxides.Of.Nitrogen	-2.830	0.004678
## Carbon.Monoxide:Ozone	-9.356	< 2e-16
## Methane:Ozone	2.890	0.003880
## Nitric.Oxide:Nitrogen.Dioxide	-12.644	< 2e-16
## Nitrogen.Dioxide:Total.Oxides.Of.Nitrogen	10.801	< 2e-16
## Ozone:Non.methane.Hydrocarbons	2.856	0.004319
## Ozone:PM2.5.Mass	-8.103	7.41e-16
## Ozone:Total.Hydrocarbons	-2.904	0.003703
##		
## (Intercept)		***
## factor(Station.Name)Calgary Northwest		*
## factor(Station.Name)Calgary Southeast		
## Carbon.Monoxide		***
## Methane		
## Nitric.Oxide		**
## Nitrogen.Dioxide		**
## Ozone		***
## Non.methane.Hydrocarbons		
## PM2.5.Mass		***
## Total.Hydrocarbons		
## Total.Oxides.Of.Nitrogen		**
## factor(Station.Name)Calgary Northwest:Carbon.Monoxide		***
## factor(Station.Name)Calgary Southeast:Carbon.Monoxide		**
## factor(Station.Name)Calgary Northwest:Nitric.Oxide		**
## factor(Station.Name)Calgary Southeast:Nitric.Oxide		**
## factor(Station.Name)Calgary Northwest:Nitrogen.Dioxide		**
## factor(Station.Name)Calgary Southeast:Nitrogen.Dioxide		**
## factor(Station.Name)Calgary Northwest:Ozone		**
## factor(Station.Name)Calgary Southeast:Ozone		
## factor(Station.Name)Calgary Northwest:PM2.5.Mass		***
## factor(Station.Name)Calgary Southeast:PM2.5.Mass		*
## factor(Station.Name)Calgary Northwest:Total.Oxides.Of.Nitrogen		**
## factor(Station.Name)Calgary Southeast:Total.Oxides.Of.Nitrogen		**
## Carbon.Monoxide:Ozone		***
## Methane:Ozone		**
## Nitric.Oxide:Nitrogen.Dioxide		***
## Nitrogen.Dioxide:Total.Oxides.Of.Nitrogen		***

```

## Ozone:Non.methane.Hydrocarbons          **
## Ozone:PM2.5.Mass                      ***
## Ozone:Total.Hydrocarbons                **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ',' 1
##
## Residual standard error: 0.1509 on 3399 degrees of freedom
## Multiple R-squared:  0.9086, Adjusted R-squared:  0.9078
## F-statistic:  1126 on 30 and 3399 DF,  p-value: < 2.2e-16

```

90.78 % of the variation is explained by the interaction model.

## HYPOTHESIS

$H_0 : \setminus H_a :$

```
anova(airinter2,airinter1)
```

```

## Analysis of Variance Table
##
## Model 1: Air.Quality.Index ~ factor(Station.Name) + Carbon.Monoxide +
##           Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone + Non.methane.Hydrocarbons +
##           PM2.5.Mass + Total.Hydrocarbons + Total.Oxides.Of.Nitrogen +
##           factor(Station.Name) * Carbon.Monoxide + factor(Station.Name) *
##           Nitric.Oxide + factor(Station.Name) * Nitrogen.Dioxide +
##           factor(Station.Name) * Ozone + factor(Station.Name) * PM2.5.Mass +
##           factor(Station.Name) * Total.Oxides.Of.Nitrogen + Carbon.Monoxide *
##           Ozone + Methane * Ozone + Nitric.Oxide * Nitrogen.Dioxide +
##           Nitrogen.Dioxide * Total.Oxides.Of.Nitrogen + Ozone * Non.methane.Hydrocarbons +
##           Ozone * PM2.5.Mass + Ozone * Total.Hydrocarbons
## Model 2: Air.Quality.Index ~ factor(Station.Name) + Carbon.Monoxide +
##           Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone + Non.methane.Hydrocarbons +
##           PM2.5.Mass + Total.Hydrocarbons + Total.Oxides.Of.Nitrogen +
##           factor(Station.Name) * Carbon.Monoxide + factor(Station.Name) *
##           Nitric.Oxide + factor(Station.Name) * Nitrogen.Dioxide +
##           factor(Station.Name) * Ozone + factor(Station.Name) * PM2.5.Mass +
##           factor(Station.Name) * Total.Oxides.Of.Nitrogen + Carbon.Monoxide *
##           Ozone + Methane * Ozone + Nitric.Oxide * Nitrogen.Dioxide +
##           Nitric.Oxide * Total.Oxides.Of.Nitrogen + Nitrogen.Dioxide *
##           Total.Oxides.Of.Nitrogen + Ozone * Non.methane.Hydrocarbons +
##           Ozone * PM2.5.Mass + Ozone * Total.Hydrocarbons
##   Res.Df   RSS Df Sum of Sq    F Pr(>F)
## 1     3399 77.35
## 2     3398 77.31  1  0.039521 1.7371 0.1876

```

The p-vale for the F-statistic is greater than 0.05 level of significance and we can say that the reduced interaction model 'airinter2' is better than the 'airinter1' in the estimation of the air quality index.

## finding the best interaction variables using stepwise section procedure :

we already have found the best interaction terms in the previous steps but just to reconfirm, i will also use stepwise section procedure action to find the interaction terms that best fits in the interaction model.

```
full_interaction_model<-lm(Air.Quality.Index ~ (Station.Name+Carbon.Monoxide+Methane+Nitric.Oxide
+Nitrogen.Dioxide+Ozone
+Non.methane.Hydrocarbons+PM2.5.Mass
+Total.Hydrocarbons+Total.Oxides.Of.Nitrogen)^2,
data=air)
airinter3=ols_step_both_p(full_interaction_model,p_enter = 0.1, p_remove = 0.3, details=FALSE)
summary(airinter3$model)

##
## Call:
## lm(formula = paste(response, "~", paste(preds, collapse = " + ")),
##     data = 1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -1.84084 -0.07885  0.00409  0.07778  1.69286
##
## Coefficients:
##                               Estimate Std. Error
## (Intercept)                -3.225e-01  2.684e-01
## Total.Oxides.Of.Nitrogen    7.476e-01  3.152e+00
## Ozone                      7.885e+01  4.851e+00
## Nitrogen.Dioxide            2.594e+01  5.298e+00
## Station.NameCalgary Northwest 4.090e-01  1.630e-01
## Station.NameCalgary Southeast -3.681e-01  1.330e-01
## Methane                     2.086e+00  2.969e+00
## Nitric.Oxide                9.777e+00  3.942e+00
## Total.Hydrocarbons          -1.408e+00  2.979e+00
## Non.methane.Hydrocarbons    1.996e+00  2.989e+00
## Carbon.Monoxide              1.967e+00  1.851e-01
## PM2.5.Mass                  3.108e-02  5.782e-03
## Ozone:PM2.5.Mass            -4.519e-01  5.961e-02
## PM2.5.Mass:Total.Hydrocarbons 1.030e-02  2.570e-03
## PM2.5.Mass:Nitric.Oxide     -3.548e-01  5.024e-02
## Ozone:Carbon.Monoxide       -5.694e+01  5.755e+00
## Nitric.Oxide:Carbon.Monoxide -2.126e+01  3.930e+00
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide 5.550e+02  6.593e+01
## Nitrogen.Dioxide:Nitric.Oxide   -9.117e+02  1.475e+02
## Total.Oxides.Of.Nitrogen:Nitric.Oxide 1.028e+02  2.058e+01
## Nitrogen.Dioxide:Station.NameCalgary Northwest -1.094e+01  1.548e+00
## Nitrogen.Dioxide:Station.NameCalgary Southeast -7.672e+00  1.804e+00
## PM2.5.Mass:Station.NameCalgary Northwest -4.405e-03  1.129e-03
## PM2.5.Mass:Station.NameCalgary Southeast 1.235e-03  1.175e-03
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons 1.672e+01  5.653e+00
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons 4.389e+00  9.048e+00
## Station.NameCalgary Northwest:Total.Hydrocarbons     -1.578e+01  5.652e+00
## Station.NameCalgary Southeast:Total.Hydrocarbons     -4.386e+00  9.037e+00
```

## Nitrogen.Dioxide:Methane	-3.559e+01	9.606e+00
## Ozone:Total.Hydrocarbons	-1.260e+01	2.541e+00
## Ozone:Station.NameCalgary Northwest	-2.228e+00	8.715e-01
## Ozone:Station.NameCalgary Southeast	6.508e-01	9.681e-01
## Ozone:Nitric.Oxide	1.676e+02	6.148e+01
## Station.NameCalgary Northwest:Methane	1.560e+01	5.652e+00
## Station.NameCalgary Southeast:Methane	4.541e+00	9.035e+00
## Methane:Total.Hydrocarbons	-9.573e-02	4.566e-02
## Non.methane.Hydrocarbons:Carbon.Monoxide	-2.349e+00	6.912e-01
## Nitrogen.Dioxide:Total.Hydrocarbons	2.550e+01	9.170e+00
## Station.NameCalgary Northwest:Carbon.Monoxide	3.812e-01	1.320e-01
## Station.NameCalgary Southeast:Carbon.Monoxide	3.849e-01	1.831e-01
##		
## (Intercept)	t value	Pr(> t )
## Total.Oxides.Of.Nitrogen	-1.201	0.229699
## Ozone	0.237	0.812533
## Nitrogen.Dioxide	16.254	< 2e-16 ***
## Station.NameCalgary Northwest	4.897	1.02e-06 ***
## Station.NameCalgary Southeast	2.510	0.012135 *
## Methane	-2.766	0.005700 **
## Nitric.Oxide	0.703	0.482213
## Total.Hydrocarbons	2.481	0.013167 *
## Non.methane.Hydrocarbons	-0.473	0.636474
## Carbon.Monoxide	0.668	0.504157
## PM2.5.Mass	10.626	< 2e-16 ***
## Ozone:PM2.5.Mass	5.375	8.16e-08 ***
## PM2.5.Mass:Total.Hydrocarbons	-7.582	4.37e-14 ***
## PM2.5.Mass:Nitric.Oxide	4.007	6.29e-05 ***
## Ozone:Carbon.Monoxide	-7.063	1.97e-12 ***
## Nitric.Oxide:Carbon.Monoxide	-9.894	< 2e-16 ***
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide	-5.410	6.73e-08 ***
## Nitrogen.Dioxide:Nitric.Oxide	8.418	< 2e-16 ***
## Total.Oxides.Of.Nitrogen:Nitric.Oxide	-6.179	7.21e-10 ***
## Nitrogen.Dioxide:Station.NameCalgary Northwest	4.997	6.12e-07 ***
## Nitrogen.Dioxide:Station.NameCalgary Southeast	-7.070	1.87e-12 ***
## PM2.5.Mass:Station.NameCalgary Northwest	-4.254	2.16e-05 ***
## PM2.5.Mass:Station.NameCalgary Southeast	-3.902	9.73e-05 ***
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons	1.051	0.293499
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons	2.958	0.003114 **
## Station.NameCalgary Northwest:Total.Hydrocarbons	0.485	0.627629
## Station.NameCalgary Southeast:Total.Hydrocarbons	-2.792	0.005273 **
## Nitrogen.Dioxide:Methane	-0.485	0.627470
## Ozone:Total.Hydrocarbons	-3.705	0.000215 ***
## Ozone:Station.NameCalgary Northwest	-4.956	7.54e-07 ***
## Ozone:Station.NameCalgary Southeast	-2.557	0.010612 *
## Ozone:Nitric.Oxide	0.672	0.501468
## Station.NameCalgary Northwest:Methane	2.726	0.006435 **
## Station.NameCalgary Southeast:Methane	2.761	0.005798 **
## Methane:Total.Hydrocarbons	0.503	0.615264
## Non.methane.Hydrocarbons:Carbon.Monoxide	-2.097	0.036100 *
## Nitrogen.Dioxide:Total.Hydrocarbons	-3.399	0.000684 ***
## Station.NameCalgary Northwest:Carbon.Monoxide	2.781	0.005448 **
## Station.NameCalgary Southeast:Carbon.Monoxide	2.888	0.003908 **
## ---	2.102	0.035646 *

```

## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1468 on 3390 degrees of freedom
## Multiple R-squared:  0.9136, Adjusted R-squared:  0.9126
## F-statistic: 919.6 on 39 and 3390 DF,  p-value: < 2.2e-16

```

## Comparing both interaction models :

### Introduction

This document compares two interaction models for a regression analysis:

1. The **Second Model**, created using the function `ols_step_both_p` from the `olsrr` library.
2. The **First Model**, built by manually removing interaction terms in groups based on relevance and performance.

We will evaluate and document the differences in their variables, and provide insights to determine the best interaction model.

#### Second Model Variables with function `ols_step_both_p`:

"Intercept", "factor(Station.Name)Calgary Northwest", "factor(Station.Name)Calgary Southeast", "Carbon.Monoxide", "Methane", "Nitric.Oxide", "Nitrogen.Dioxide", "Ozone", "Non.methane.Hydrocarbons", "PM2.5.Mass", "Total.Hydrocarbons", "Total.Oxides.Of.Nitrogen", "factor(Station.Name)Calgary Northwest:Carbon.Monoxide", "factor(Station.Name)Calgary Southeast:Carbon.Monoxide", "factor(Station.Name)Calgary Northwest:Nitric.Oxide", "factor(Station.Name)Calgary Southeast:Nitric.Oxide", "factor(Station.Name)Calgary Northwest:Nitrogen.Dioxide", "factor(Station.Name)Calgary Southeast:Nitrogen.Dioxide", "factor(Station.Name)Calgary Northwest:Ozone", "factor(Station.Name)Calgary Southeast:Ozone", "factor(Station.Name)Calgary Northwest:PM2.5.Mass", "factor(Station.Name)Calgary Southeast:PM2.5.Mass", "factor(Station.Name)Calgary Northwest:Total.Oxides.Of.Nitrogen", "factor(Station.Name)Calgary Southeast:Total.Oxides.Of.Nitrogen", "Carbon.Monoxide:Ozone", "Methane:Ozone", "Nitric.Oxide:Nitrogen.Dioxide", "Nitrogen.Dioxide:Total.Oxides.Of.Nitrogen", "Ozone:Non.methane.Hydrocarbons", "Ozone:PM2.5.Mass", "Ozone:Total.Hydrocarbons"

#### First Model Variables :

"Intercept", "Total.Oxides.Of.Nitrogen", "Ozone", "Nitrogen.Dioxide", "Station.NameCalgary Northwest", "Station.NameCalgary Southeast", "Methane", "Nitric.Oxide", "Total.Hydrocarbons", "Non.methane.Hydrocarbons", "Carbon.Monoxide", "PM2.5.Mass", "Ozone:PM2.5.Mass", "PM2.5.Mass:Total.Hydrocarbons", "PM2.5.Mass:Nitric.Oxide", "Ozone:Carbon.Monoxide", "Nitric.Oxide:Carbon.Monoxide", "Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide", "Nitrogen.Dioxide:Nitric.Oxide", "Total.Oxides.Of.Nitrogen:Nitric.Oxide", "Nitrogen.Dioxide:Station.NameCalgary Northwest", "Nitrogen.Dioxide:Station.NameCalgary Southeast", "PM2.5.Mass:Station.NameCalgary Northwest", "PM2.5.Mass:Station.NameCalgary Southeast", "Station.NameCalgary Northwest:Non.methane.Hydrocarbons", "Station.NameCalgary Southeast:Non.methane.Hydrocarbons", "Station.NameCalgary Northwest:Total.Hydrocarbons", "Nitrogen.Dioxide:Methane", "Ozone:Total.Hydrocarbons", "Ozone:Station.NameCalgary Northwest", "Ozone:Station.NameCalgary Southeast", "Ozone:Nitric.Oxide", "Station.NameCalgary Northwest:Methane", "Station.NameCalgary Southeast:Methane", "Methane:Total.Hydrocarbons", "Non.methane.Hydrocarbons:Carbon.Monoxide", "Nitrogen.Dioxide:Total.Hydrocarbons", "Station.NameCalgary Northwest:Carbon.Monoxide", "Station.NameCalgary Southeast:Carbon.Monoxide"

I will compare both of these interaction model by calculating AdjustedR, Mallows' Cp, and Akaike Information Criterion (AIC)

```

library(olsrr)

# Adjusted R
adj_r2_model1 <- summary(airinter2)$adj.r.squared
print(adj_r2_model1)

## [1] 0.9078063

# Mallows' Cp (assuming full_model is defined)
cp_model1 <- ols_mallows_cp(airinter2, full_interaction_model)
print(cp_model1)

## [1] 249.8147

# AIC
aic_model1 <- AIC(airinter2)
print(aic_model1)

## [1] -3208.56

# Get the airinter3 (from ols_step_both_p)
final_model <- airinter3$model

# Residual sum of squares (RSS)
RSS <- sum(residuals(final_model)^2)

# Estimated variance of the residuals (sigma^2)
n <- length(residuals(final_model)) # Number of observations
p <- length(coef(final_model)) # Number of parameters (including the intercept)

# Estimated variance (mean squared error)
sigma_squared <- RSS / (n - p)

# Mallows' Cp
cp_value <- (RSS / sigma_squared) - (n - 2 * p)
print(paste("Mallows' Cp:", cp_value))

## [1] "Mallows' Cp: 40"

# Adjusted R-squared for airinter3 model
adj_r_squared_final <- summary(final_model)$adj.r.squared
print(paste("Adjusted R-squared for final model:", adj_r_squared_final))

## [1] "Adjusted R-squared for final model: 0.912648218076459"

# AIC for airinter3 model
aic_value_final <- AIC(final_model)
print(paste("AIC for final model:", aic_value_final))

## [1] "AIC for final model: -3384.69622315295"

```

Based on all the above results we can clearly see that model `airinter3` has high adj R-square, lower AIC value and CP closer to k+1 (where k in number of variable terms in the model). **i can conclude that airinter3 is the most fit interaction model.**

**Important note:** There are many terms that are not significant in the model but should be consider in the model due to Hierarchy principal.

## Final Interaction Model Equation

The equation for the final interaction model is:

$$\begin{aligned}
 \hat{y} = & -0.3225 + 0.7476 \times \text{Total.Oxides.Of.Nitrogen} + 78.85 \times \text{Ozone} + 25.94 \times \text{Nitrogen.Dioxide} \\
 & + 0.409 \times \text{Station.NameCalgary Northwest} - 0.368 \times \text{Station.NameCalgary Southeast} + 2.086 \times \text{Methane} + 9.777 \times \text{Nitric.Oxide} \\
 & - 1.408 \times \text{Total.Hydrocarbons} + 1.996 \times \text{Non.methane.Hydrocarbons} + 1.967 \times \text{Carbon.Monoxide} + 0.03108 \times \text{PM2.5.Mass} \\
 & - 0.4519 \times (\text{Ozone} \times \text{PM2.5.Mass}) + 0.0103 \times (\text{PM2.5.Mass} \times \text{Total.Hydrocarbons}) - 0.3548 \times (\text{PM2.5.Mass} \times \text{Nitric.Oxide}) \\
 & - 56.94 \times (\text{Ozone} \times \text{Carbon.Monoxide}) - 21.26 \times (\text{Nitric.Oxide} \times \text{Carbon.Monoxide}) + 555.0 \times (\text{Total.Oxides.Of.Nitrogen} \times \text{Nitrogen.Dioxide}) \\
 & - 911.7 \times (\text{Nitrogen.Dioxide} \times \text{Nitric.Oxide}) + 102.8 \times (\text{Total.Oxides.Of.Nitrogen} \times \text{Nitric.Oxide}) - 10.94 \times (\text{Nitrogen.Dioxide} \times \text{Station.NameCalgary Northwest}) \\
 & - 7.672 \times (\text{Nitrogen.Dioxide} \times \text{Station.NameCalgary Southeast}) + 0.0044 \times (\text{PM2.5.Mass} \times \text{Station.NameCalgary Northwest}) \\
 & + 0.00124 \times (\text{PM2.5.Mass} \times \text{Station.NameCalgary Southeast}) + 16.72 \times (\text{Station.NameCalgary Northwest} \times \text{Non.methane.Hydrocarbons}) \\
 & + 4.389 \times (\text{Station.NameCalgary Southeast} \times \text{Non.methane.Hydrocarbons}) - 15.78 \times (\text{Station.NameCalgary Northwest} \times \text{Total.Hydrocarbons}) \\
 & - 4.386 \times (\text{Station.NameCalgary Southeast} \times \text{Total.Hydrocarbons}) - 35.59 \times (\text{Nitrogen.Dioxide} \times \text{Methane}) \\
 & - 12.60 \times (\text{Ozone} \times \text{Total.Hydrocarbons}) - 2.228 \times (\text{Ozone} \times \text{Station.NameCalgary Northwest}) + 0.6508 \times (\text{Ozone} \times \text{Station.NameCalgary Southeast}) \\
 & + 167.6 \times (\text{Ozone} \times \text{Nitric.Oxide}) + 15.60 \times (\text{Station.NameCalgary Northwest} \times \text{Methane}) + 4.541 \times (\text{Station.NameCalgary Southeast} \times \text{Nitric.Oxide})
 \end{aligned}$$

$$-0.0957 \times (\text{Methane} \times \text{Total.Hydrocarbons}) - 2.349 \times (\text{Non.methane.Hydrocarbons} \times \text{Carbon.Monoxide})$$

$$+ 25.50 \times (\text{Nitrogen.Dioxide} \times \text{Total.Hydrocarbons}) + 0.3812 \times (\text{Station.NameCalgary Northwest} \times \text{Carbon.Monoxide})$$

$$+ 0.3849 \times (\text{Station.NameCalgary Southeast} \times \text{Carbon.Monoxide})$$

It is a long equation.

```
# Fit the linear regression model with interaction terms
interaction_model <- lm(Air.Quality.Index ~
  Total.Oxides.Of.Nitrogen + Ozone + Nitrogen.Dioxide + Station.Name + Methane +
  Nitric.Oxide + Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
  Ozone * PM2.5.Mass +
  PM2.5.Mass * Total.Hydrocarbons +
  PM2.5.Mass * Nitric.Oxide +
  Ozone * Carbon.Monoxide +
  Nitric.Oxide * Carbon.Monoxide +
  Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide +
  Nitrogen.Dioxide * Nitric.Oxide +
  Total.Oxides.Of.Nitrogen * Nitric.Oxide +
  Nitrogen.Dioxide * Station.Name +
  PM2.5.Mass * Station.Name +
  Station.Name * Non.methane.Hydrocarbons +
  Station.Name * Total.Hydrocarbons +
  Nitrogen.Dioxide * Methane +
  Ozone * Total.Hydrocarbons +
  Ozone * Station.Name +
  Ozone * Nitric.Oxide +
  Station.Name * Methane +
  Methane * Total.Hydrocarbons +
  Non.methane.Hydrocarbons * Carbon.Monoxide +
  Nitrogen.Dioxide * Total.Hydrocarbons +
  Station.Name * Carbon.Monoxide,
  data = air)

# Display the model summary
summary(interaction_model)
```

```
##
## Call:
## lm(formula = Air.Quality.Index ~ Total.Oxides.Of.Nitrogen + Ozone +
##     Nitrogen.Dioxide + Station.Name + Methane + Nitric.Oxide +
##     Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
##     PM2.5.Mass + Ozone * PM2.5.Mass + PM2.5.Mass * Total.Hydrocarbons +
##     PM2.5.Mass * Nitric.Oxide + Ozone * Carbon.Monoxide + Nitric.Oxide *
##     Carbon.Monoxide + Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide +
##     Nitrogen.Dioxide * Nitric.Oxide + Total.Oxides.Of.Nitrogen *
##     Nitric.Oxide + Nitrogen.Dioxide * Station.Name + PM2.5.Mass *
```

```

## Station.Name + Station.Name * Non.methane.Hydrocarbons +
## Station.Name * Total.Hydrocarbons + Nitrogen.Dioxide * Methane +
## Ozone * Total.Hydrocarbons + Ozone * Station.Name + Ozone *
## Nitric.Oxide + Station.Name * Methane + Methane * Total.Hydrocarbons +
## Non.methane.Hydrocarbons * Carbon.Monoxide + Nitrogen.Dioxide *
## Total.Hydrocarbons + Station.Name * Carbon.Monoxide, data = air)
##
## Residuals:
##      Min     1Q   Median     3Q    Max
## -1.84084 -0.07885  0.00409  0.07778  1.69286
##
## Coefficients:
## (Intercept)                               Estimate Std. Error
## Total.Oxides.Of.Nitrogen                  -3.225e-01 2.684e-01
## Ozone                                     7.476e-01 3.152e+00
## Nitrogen.Dioxide                          7.885e+01 4.851e+00
## Station.NameCalgary Northwest             2.594e+01 5.298e+00
## Station.NameCalgary Southeast              4.090e-01 1.630e-01
## Methane                                    -3.681e-01 1.330e-01
## Nitric.Oxide                             2.086e+00 2.969e+00
## Total.Hydrocarbons                      9.777e+00 3.942e+00
## Non.methane.Hydrocarbons                 -1.408e+00 2.979e+00
## Carbon.Monoxide                           1.996e+00 2.989e+00
## PM2.5.Mass                                1.967e+00 1.851e-01
## Ozone:PM2.5.Mass                         3.108e-02 5.782e-03
## Total.Hydrocarbons:PM2.5.Mass            -4.519e-01 5.961e-02
## Nitric.Oxide:PM2.5.Mass                  1.030e-02 2.570e-03
## Ozone:Carbon.Monoxide                    -3.548e-01 5.024e-02
## Nitric.Oxide:Carbon.Monoxide             -5.694e+01 5.755e+00
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide 2.126e+01 3.930e+00
## Nitrogen.Dioxide:Nitric.Oxide            5.550e+02 6.593e+01
## Total.Oxides.Of.Nitrogen:Nitric.Oxide    -9.117e+02 1.475e+02
## Nitrogen.Dioxide:Station.NameCalgary Northwest -1.028e+02 2.058e+01
## Nitrogen.Dioxide:Station.NameCalgary Southeast -1.094e+01 1.548e+00
## Station.NameCalgary Northwest:PM2.5.Mass   -7.672e+00 1.804e+00
## Station.NameCalgary Southeast:PM2.5.Mass    -4.405e-03 1.129e-03
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons 1.235e-03 1.175e-03
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons 1.672e+01 5.653e+00
## Station.NameCalgary Northwest:Total.Hydrocarbons   4.389e+00 9.048e+00
## Station.NameCalgary Southeast:Total.Hydrocarbons   -1.578e+01 5.652e+00
## Nitrogen.Dioxide:Methane                  -4.386e+00 9.037e+00
## Ozone:Total.Hydrocarbons                 -3.559e+01 9.606e+00
## Ozone:Station.NameCalgary Northwest       -1.260e+01 2.541e+00
## Ozone:Station.NameCalgary Southeast        -2.228e+00 8.715e-01
## Ozone:Nitric.Oxide                      6.508e-01 9.681e-01
## Station.NameCalgary Northwest:Methane      1.676e+02 6.148e+01
## Station.NameCalgary Southeast:Methane       1.560e+01 5.652e+00
## Methane:Total.Hydrocarbons                4.541e+00 9.035e+00
## Non.methane.Hydrocarbons:Carbon.Monoxide -9.573e-02 4.566e-02
## Nitrogen.Dioxide:Total.Hydrocarbons       -2.349e+00 6.912e-01
## Station.NameCalgary Northwest:Carbon.Monoxide 2.550e+01 9.170e+00
## Station.NameCalgary Southeast:Carbon.Monoxide 3.812e-01 1.320e-01
## Station.NameCalgary Southeast:Carbon.Monoxide 3.849e-01 1.831e-01
## t value Pr(>|t|)
```

```

## (Intercept) -1.201 0.229699
## Total.Oxides.Of.Nitrogen 0.237 0.812533
## Ozone 16.254 < 2e-16 ***
## Nitrogen.Dioxide 4.897 1.02e-06 ***
## Station.NameCalgary Northwest 2.510 0.012135 *
## Station.NameCalgary Southeast -2.766 0.005700 **
## Methane 0.703 0.482213
## Nitric.Oxide 2.481 0.013167 *
## Total.Hydrocarbons -0.473 0.636474
## Non.methane.Hydrocarbons 0.668 0.504157
## Carbon.Monoxide 10.626 < 2e-16 ***
## PM2.5.Mass 5.375 8.16e-08 ***
## Ozone:PM2.5.Mass -7.582 4.37e-14 ***
## Total.Hydrocarbons:PM2.5.Mass 4.007 6.29e-05 ***
## Nitric.Oxide:PM2.5.Mass -7.063 1.97e-12 ***
## Ozone:Carbon.Monoxide -9.894 < 2e-16 ***
## Nitric.Oxide:Carbon.Monoxide -5.410 6.73e-08 ***
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide 8.418 < 2e-16 ***
## Nitrogen.Dioxide:Nitric.Oxide -6.179 7.21e-10 ***
## Total.Oxides.Of.Nitrogen:Nitric.Oxide 4.997 6.12e-07 ***
## Nitrogen.Dioxide:Station.NameCalgary Northwest -7.070 1.87e-12 ***
## Nitrogen.Dioxide:Station.NameCalgary Southeast -4.254 2.16e-05 ***
## Station.NameCalgary Northwest:PM2.5.Mass -3.902 9.73e-05 ***
## Station.NameCalgary Southeast:PM2.5.Mass 1.051 0.293499
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons 2.958 0.003114 **
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons 0.485 0.627629
## Station.NameCalgary Northwest:Total.Hydrocarbons -2.792 0.005273 **
## Station.NameCalgary Southeast:Total.Hydrocarbons -0.485 0.627470
## Nitrogen.Dioxide:Methane -3.705 0.000215 ***
## Ozone:Total.Hydrocarbons -4.956 7.54e-07 ***
## Ozone:Station.NameCalgary Northwest -2.557 0.010612 *
## Ozone:Station.NameCalgary Southeast 0.672 0.501468
## Ozone:Nitric.Oxide 2.726 0.006435 **
## Station.NameCalgary Northwest:Methane 2.761 0.005798 **
## Station.NameCalgary Southeast:Methane 0.503 0.615264
## Methane:Total.Hydrocarbons -2.097 0.036100 *
## Non.methane.Hydrocarbons:Carbon.Monoxide -3.399 0.000684 ***
## Nitrogen.Dioxide:Total.Hydrocarbons 2.781 0.005448 **
## Station.NameCalgary Northwest:Carbon.Monoxide 2.888 0.003908 **
## Station.NameCalgary Southeast:Carbon.Monoxide 2.102 0.035646 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ',' 1
##
## Residual standard error: 0.1468 on 3390 degrees of freedom
## Multiple R-squared: 0.9136, Adjusted R-squared: 0.9126
## F-statistic: 919.6 on 39 and 3390 DF, p-value: < 2.2e-16

```

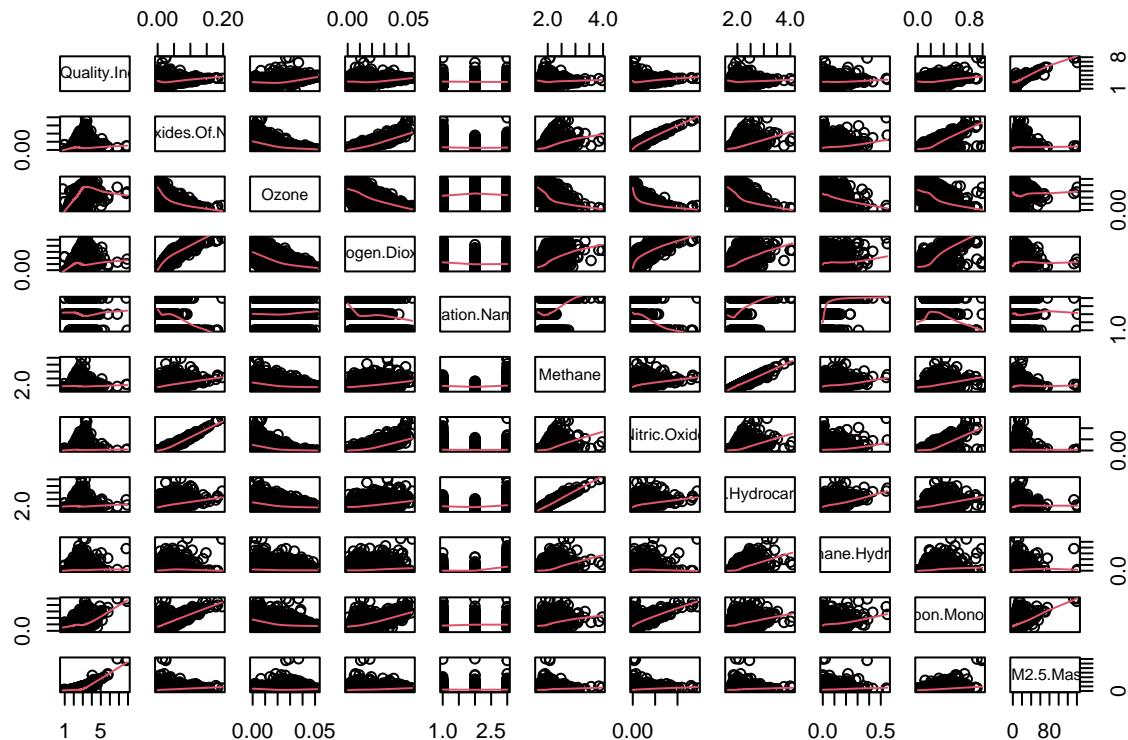
## HIGHER ORDER INTERACTIONS

Now that i have the interaction model i can look for only higher order terms from our interaction model.

```

library(GGally)
CreditImportant = data.frame( air$Total.Oxides.Of.Nitrogen , air$Ozone , air$Nitrogen.Dioxide , air$Station.Name , air$Methane , air$Nitric.Oxide , air$Total.Hydrocarbons , air$Non.methane.Hydrocarbons , air$Carbon.Monoxide , air$PM2.5.Mass )
pairs(~Air.Quality.Index+ Total.Oxides.Of.Nitrogen + Ozone + Nitrogen.Dioxide + Station.Name + Methane + Nitric.Oxide + Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide + PM2.5.Mass)

```



By zooming in on the above ggplot , it can be observed in the first column of line graphs that Total.Oxides.Of.Nitrogen,Ozone, Nitrogen.Dioxide , Carbon.Monoxide and PM2.5.Mass are all the variables that have non liner lines in thier graphs with Air.Quality.Index.

My strategy is deal with one vraible at a time . i will continue to increase the degree's until the highest degree term becomes insignificant.

Total.Oxides.Of.Nitrogen

```

model1 <- lm(Air.Quality.Index ~
  Total.Oxides.Of.Nitrogen+I(Total.Oxides.Of.Nitrogen^2) + Ozone + Nitrogen.Dioxide +
  Nitric.Oxide + Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
  Ozone * PM2.5.Mass +
  PM2.5.Mass * Total.Hydrocarbons +
  PM2.5.Mass * Nitric.Oxide +
  Ozone * Carbon.Monoxide +
  Nitric.Oxide * Carbon.Monoxide +
  Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide +
  Nitrogen.Dioxide * Nitric.Oxide +
  Total.Oxides.Of.Nitrogen * Nitric.Oxide +
  PM2.5.Mass)

```

```

Nitrogen.Dioxide * Station.Name +
PM2.5.Mass * Station.Name +
Station.Name * Non.methane.Hydrocarbons +
Station.Name * Total.Hydrocarbons +
Nitrogen.Dioxide * Methane +
Ozone * Total.Hydrocarbons +
Ozone * Station.Name +
Ozone * Nitric.Oxide +
Station.Name * Methane +
Methane * Total.Hydrocarbons +
Non.methane.Hydrocarbons * Carbon.Monoxide +
Nitrogen.Dioxide * Total.Hydrocarbons +
Station.Name * Carbon.Monoxide,
data = air)

# Display the model summary
summary(model1)

```

```

##
## Call:
## lm(formula = Air.Quality.Index ~ Total.Oxides.Of.Nitrogen + I(Total.Oxides.Of.Nitrogen^2) +
##     Ozone + Nitrogen.Dioxide + Station.Name + Methane + Nitric.Oxide +
##     Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
##     PM2.5.Mass + Ozone * PM2.5.Mass + PM2.5.Mass * Total.Hydrocarbons +
##     PM2.5.Mass * Nitric.Oxide + Ozone * Carbon.Monoxide + Nitric.Oxide *
##     Carbon.Monoxide + Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide +
##     Nitrogen.Dioxide * Nitric.Oxide + Total.Oxides.Of.Nitrogen *
##     Nitric.Oxide + Nitrogen.Dioxide * Station.Name + PM2.5.Mass *
##     Station.Name + Station.Name * Non.methane.Hydrocarbons +
##     Station.Name * Total.Hydrocarbons + Nitrogen.Dioxide * Methane +
##     Ozone * Total.Hydrocarbons + Ozone * Station.Name + Ozone *
##     Nitric.Oxide + Station.Name * Methane + Methane * Total.Hydrocarbons +
##     Non.methane.Hydrocarbons * Carbon.Monoxide + Nitrogen.Dioxide *
##     Total.Hydrocarbons + Station.Name * Carbon.Monoxide, data = air)
##
## Residuals:
##      Min        1Q    Median        3Q       Max
## -1.84112 -0.07906  0.00375  0.07833  1.69516
##
## Coefficients:
##                               Estimate Std. Error
## (Intercept)                -4.120e-01  2.700e-01
## Total.Oxides.Of.Nitrogen      1.353e+01  5.493e+00
## I(Total.Oxides.Of.Nitrogen^2) -5.356e+02  1.886e+02
## Ozone                      8.012e+01  4.867e+00
## Nitrogen.Dioxide             1.473e+01  6.605e+00
## Station.NameCalgary Northwest 4.207e-01  1.629e-01
## Station.NameCalgary Southeast -3.956e-01  1.333e-01
## Methane                     2.132e+00  2.966e+00
## Nitric.Oxide                 -2.266e+00 5.787e+00
## Total.Hydrocarbons           -1.392e+00 2.976e+00
## Non.methane.Hydrocarbons      1.891e+00  2.986e+00
## Carbon.Monoxide              1.959e+00  1.849e-01

```

## PM2.5.Mass	3.069e-02	5.778e-03
## Ozone:PM2.5.Mass	-4.529e-01	5.955e-02
## Total.Hydrocarbons:PM2.5.Mass	1.050e-02	2.568e-03
## Nitric.Oxide:PM2.5.Mass	-3.552e-01	5.019e-02
## Ozone:Carbon.Monoxide	-5.665e+01	5.750e+00
## Nitric.Oxide:Carbon.Monoxide	-2.101e+01	3.926e+00
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide	1.108e+03	2.056e+02
## Nitrogen.Dioxide:Nitric.Oxide	-9.732e+02	1.490e+02
## Total.Oxides.Of.Nitrogen:Nitric.Oxide	6.450e+02	1.920e+02
## Nitrogen.Dioxide:Station.NameCalgary Northwest	-1.119e+01	1.549e+00
## Nitrogen.Dioxide:Station.NameCalgary Southeast	-7.642e+00	1.802e+00
## Station.NameCalgary Northwest:PM2.5.Mass	-4.359e-03	1.128e-03
## Station.NameCalgary Southeast:PM2.5.Mass	1.253e-03	1.174e-03
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons	1.682e+01	5.648e+00
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons	3.824e+00	9.041e+00
## Station.NameCalgary Northwest:Total.Hydrocarbons	-1.584e+01	5.646e+00
## Station.NameCalgary Southeast:Total.Hydrocarbons	-3.811e+00	9.030e+00
## Nitrogen.Dioxide:Methane	-3.842e+01	9.648e+00
## Ozone:Total.Hydrocarbons	-1.320e+01	2.548e+00
## Ozone:Station.NameCalgary Northwest	-2.361e+00	8.719e-01
## Ozone:Station.NameCalgary Southeast	5.562e-01	9.676e-01
## Ozone:Nitric.Oxide	1.643e+02	6.143e+01
## Station.NameCalgary Northwest:Methane	1.566e+01	5.646e+00
## Station.NameCalgary Southeast:Methane	3.981e+00	9.028e+00
## Methane:Total.Hydrocarbons	-1.047e-01	4.572e-02
## Non.methane.Hydrocarbons:Carbon.Monoxide	-2.303e+00	6.907e-01
## Nitrogen.Dioxide:Total.Hydrocarbons	2.744e+01	9.186e+00
## Station.NameCalgary Northwest:Carbon.Monoxide	3.810e-01	1.319e-01
## Station.NameCalgary Southeast:Carbon.Monoxide	3.903e-01	1.830e-01
##		
## (Intercept)	t	value Pr(> t )
## Total.Oxides.Of.Nitrogen	-1.526	0.127076
## I(Total.Oxides.Of.Nitrogen^2)	2.463	0.013840 *
## Ozone	-2.839	0.004547 **
## Nitrogen.Dioxide	16.463	< 2e-16 ***
## Station.NameCalgary Northwest	2.230	0.025829 *
## Station.NameCalgary Southeast	2.583	0.009845 **
## Methane	-2.968	0.003014 **
## Nitric.Oxide	0.719	0.472196
## Total.Hydrocarbons	-0.392	0.695448
## Non.methane.Hydrocarbons	-0.468	0.639882
## Carbon.Monoxide	0.633	0.526563
## PM2.5.Mass	10.592	< 2e-16 ***
## Ozone:PM2.5.Mass	5.312	1.16e-07 ***
## Total.Hydrocarbons:PM2.5.Mass	-7.606	3.63e-14 ***
## Nitric.Oxide:PM2.5.Mass	4.088	4.45e-05 ***
## Ozone:Carbon.Monoxide	-7.077	1.78e-12 ***
## Nitric.Oxide:Carbon.Monoxide	-9.852	< 2e-16 ***
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide	-5.352	9.30e-08 ***
## Nitrogen.Dioxide:Nitric.Oxide	5.390	7.54e-08 ***
## Total.Oxides.Of.Nitrogen:Nitric.Oxide	-6.533	7.42e-11 ***
## Nitrogen.Dioxide:Station.NameCalgary Northwest	3.359	0.000792 ***
## Nitrogen.Dioxide:Station.NameCalgary Southeast	-7.226	6.11e-13 ***
## Station.NameCalgary Northwest:PM2.5.Mass	-4.242	2.28e-05 ***
	-3.865	0.000113 ***

```

## Station.NameCalgary Southeast:PM2.5.Mass           1.067 0.286015
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons 2.978 0.002917 **
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons 0.423 0.672312
## Station.NameCalgary Northwest:Total.Hydrocarbons      -2.805 0.005065 **
## Station.NameCalgary Southeast:Total.Hydrocarbons      -0.422 0.673061
## Nitrogen.Dioxide:Methane                          -3.982 6.98e-05 ***
## Ozone:Total.Hydrocarbons                         -5.181 2.34e-07 ***
## Ozone:Station.NameCalgary Northwest            -2.708 0.006803 **
## Ozone:Station.NameCalgary Southeast             0.575 0.565425
## Ozone:Nitric.Oxide                            2.675 0.007500 **
## Station.NameCalgary Northwest:Methane          2.773 0.005580 **
## Station.NameCalgary Southeast:Methane          0.441 0.659313
## Methane:Total.Hydrocarbons                     -2.290 0.022103 *
## Non.methane.Hydrocarbons:Carbon.Monoxide     -3.334 0.000864 ***
## Nitrogen.Dioxide:Total.Hydrocarbons          2.987 0.002834 **
## Station.NameCalgary Northwest:Carbon.Monoxide 2.889 0.003893 **
## Station.NameCalgary Southeast:Carbon.Monoxide 2.133 0.032974 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ',' 1
##
## Residual standard error: 0.1467 on 3389 degrees of freedom
## Multiple R-squared:  0.9138, Adjusted R-squared:  0.9128
## F-statistic: 898.7 on 40 and 3389 DF,  p-value: < 2.2e-16

```

```

library(olsrr)
library(ggplot2)
library(lmtest)
library(mctest)
library(car)
library(GGally)
air$Season <- as.factor(air$Season)
air$Station.Name <- as.factor(air$Station.Name)

fullmodel <- lm(Air.Quality.Index ~ Station.Name + Season + Carbon.Monoxide + Methane + Nitric.Oxide + Nitrogen.Dioxide, data = air)
summary(fullmodel)

```

```

##
## Call:
## lm(formula = Air.Quality.Index ~ Station.Name + Season + Carbon.Monoxide +
##     Methane + Nitric.Oxide + Nitrogen.Dioxide + Ozone + Non.methane.Hydrocarbons +
##     PM2.5.Mass + Total.Hydrocarbons + Total.Oxides.Of.Nitrogen,
##     data = air)
##
## Residuals:
##       Min     1Q   Median     3Q    Max 
## -1.91256 -0.08855  0.00315  0.09043  2.00679
##
## Coefficients:
## (Intercept)          Estimate Std. Error t value Pr(>|t|)    
## (Intercept)          0.7893546  0.0457078 17.270 < 2e-16 ***
## Station.NameCalgary Northwest -0.0407164  0.0080189 -5.078 4.03e-07 ***
## Station.NameCalgary Southeast -0.0647311  0.0094941 -6.818 1.09e-11 ***
## SeasonSpring          0.0006750  0.0093561  0.072  0.94249  
## SeasonSummer           0.0169526  0.0095345  1.778  0.07549 .

```

```

## SeasonWinter          0.0117499  0.0093987  1.250  0.21133
## Carbon.Monoxide      0.5202057  0.0577702  9.005 < 2e-16 ***
## Methane               8.4682907  2.7434515  3.087  0.00204 **
## Nitric.Oxide          -8.9544223  2.9701025 -3.015  0.00259 **
## Nitrogen.Dioxide       7.0397357  2.7627719  2.548  0.01088 *
## Ozone                  40.1656896  0.4176181  96.178 < 2e-16 ***
## Non.methane.Hydrocarbons 8.6813278  2.7450182  3.163  0.00158 **
## PM2.5.Mass             0.0357418  0.0004993  71.586 < 2e-16 ***
## Total.Hydrocarbons     -8.3709886  2.7439713 -3.051  0.00230 **
## Total.Oxides.Of.Nitrogen 12.8073747  2.8791341  4.448  8.93e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1677 on 3415 degrees of freedom
## Multiple R-squared:  0.8865, Adjusted R-squared:  0.886
## F-statistic:  1905 on 14 and 3415 DF,  p-value: < 2.2e-16

```

```
str(air)
```

```

## 'data.frame': 3430 obs. of 14 variables:
## $ Date : chr "2015-04-07" "2015-04-08" "2015-04-09" "2015-04-10" ...
## $ Season : Factor w/ 4 levels "Fall","Spring",...: 2 2 2 2 2 2 2 2 2 ...
## $ Location : chr "(51.029944, -114.008111)" "(51.029944, -114.008111)" "(51.029944,
## $ Station.Name : Factor w/ 3 levels "Calgary Central-Inglewood",...: 1 1 1 1 1 1 1 1 1 1
## $ Air.Quality.Index : num 1.55 2.73 2.9 3.11 2.88 ...
## $ Carbon.Monoxide : num 0.183 0.179 0.188 0.221 0.133 ...
## $ Methane : num 1.93 1.94 1.92 1.93 1.89 ...
## $ Nitric.Oxide : num 0.00275 0.0045 0.003 0.0154 0.0016 0.0006 0.0263 0.0103 0.0013 0.00
## $ Nitrogen.Dioxide : num 0.0107 0.0133 0.0189 0.0201 0.013 0.0087 0.0259 0.0194 0.0074 0.00
## $ Non.methane.Hydrocarbons: num 0 0 0 0 0 0 0 0 0 ...
## $ Ozone : num 0.0303 0.0279 0.0297 0.0271 0.0329 0.0346 0.0231 0.0242 0.0358 0.00
## $ PM2.5.Mass : num 6.38 6.38 8.88 10.71 6.71 ...
## $ Total.Hydrocarbons : num 1.93 1.94 1.92 1.93 1.89 ...
## $ Total.Oxides.Of.Nitrogen: num 0.0133 0.0178 0.0219 0.0355 0.0147 0.0093 0.0522 0.0297 0.0087 0.00

```

considering **Total.Oxides.Of.Nitrogen** higher drgress.

```

model2 <- lm(Air.Quality.Index ~
              Total.Oxides.Of.Nitrogen + I(Total.Oxides.Of.Nitrogen^2) + Ozone + Nitrogen.Dioxide +
              Nitric.Oxide + Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
              Ozone * PM2.5.Mass +
              PM2.5.Mass * Total.Hydrocarbons +
              PM2.5.Mass * Nitric.Oxide +
              Ozone * Carbon.Monoxide +
              Nitric.Oxide * Carbon.Monoxide +
              Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide +
              Nitrogen.Dioxide * Nitric.Oxide +
              Total.Oxides.Of.Nitrogen * Nitric.Oxide +
              Nitrogen.Dioxide * Station.Name +
              PM2.5.Mass * Station.Name +
              Station.Name * Non.methane.Hydrocarbons +

```

```

Station.Name * Total.Hydrocarbons +
Nitrogen.Dioxide * Methane +
Ozone * Total.Hydrocarbons +
Ozone * Station.Name +
Ozone * Nitric.Oxide +
Station.Name * Methane +
Methane * Total.Hydrocarbons +
Non.methane.Hydrocarbons * Carbon.Monoxide +
Nitrogen.Dioxide * Total.Hydrocarbons +
Station.Name * Carbon.Monoxide,
data = air)

# Display the model summary
summary(model2)

## 
## Call:
## lm(formula = Air.Quality.Index ~ Total.Oxides.Of.Nitrogen + I(Total.Oxides.Of.Nitrogen^2) +
##     Ozone + Nitrogen.Dioxide + Station.Name + Methane + Nitric.Oxide +
##     Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
##     PM2.5.Mass + Ozone * PM2.5.Mass + PM2.5.Mass * Total.Hydrocarbons +
##     PM2.5.Mass * Nitric.Oxide + Ozone * Carbon.Monoxide + Nitric.Oxide *
##     Carbon.Monoxide + Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide +
##     Nitrogen.Dioxide * Nitric.Oxide + Total.Oxides.Of.Nitrogen *
##     Nitric.Oxide + Nitrogen.Dioxide * Station.Name + PM2.5.Mass *
##     Station.Name + Station.Name * Non.methane.Hydrocarbons +
##     Station.Name * Total.Hydrocarbons + Nitrogen.Dioxide * Methane +
##     Ozone * Total.Hydrocarbons + Ozone * Station.Name + Ozone *
##     Nitric.Oxide + Station.Name * Methane + Methane * Total.Hydrocarbons +
##     Non.methane.Hydrocarbons * Carbon.Monoxide + Nitrogen.Dioxide *
##     Total.Hydrocarbons + Station.Name * Carbon.Monoxide, data = air)
##
## Residuals:
##      Min    1Q   Median    3Q   Max
## -1.84112 -0.07906  0.00375  0.07833  1.69516
##
## 
## Coefficients:
##                               Estimate Std. Error
## (Intercept)                 -4.120e-01  2.700e-01
## Total.Oxides.Of.Nitrogen        1.353e+01  5.493e+00
## I(Total.Oxides.Of.Nitrogen^2) -5.356e+02  1.886e+02
## Ozone                         8.012e+01  4.867e+00
## Nitrogen.Dioxide                1.473e+01  6.605e+00
## Station.NameCalgary Northwest  4.207e-01  1.629e-01
## Station.NameCalgary Southeast -3.956e-01  1.333e-01
## Methane                        2.132e+00  2.966e+00
## Nitric.Oxide                  -2.266e+00  5.787e+00
## Total.Hydrocarbons             -1.392e+00  2.976e+00
## Non.methane.Hydrocarbons       1.891e+00  2.986e+00
## Carbon.Monoxide                 1.959e+00  1.849e-01
## PM2.5.Mass                      3.069e-02  5.778e-03
## Ozone:PM2.5.Mass                -4.529e-01  5.955e-02
## Total.Hydrocarbons:PM2.5.Mass  1.050e-02  2.568e-03

```

## Nitric.Oxide:PM2.5.Mass	-3.552e-01	5.019e-02
## Ozone:Carbon.Monoxide	-5.665e+01	5.750e+00
## Nitric.Oxide:Carbon.Monoxide	-2.101e+01	3.926e+00
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide	1.108e+03	2.056e+02
## Nitrogen.Dioxide:Nitric.Oxide	-9.732e+02	1.490e+02
## Total.Oxides.Of.Nitrogen:Nitric.Oxide	6.450e+02	1.920e+02
## Nitrogen.Dioxide:Station.NameCalgary Northwest	-1.119e+01	1.549e+00
## Nitrogen.Dioxide:Station.NameCalgary Southeast	-7.642e+00	1.802e+00
## Station.NameCalgary Northwest:PM2.5.Mass	-4.359e-03	1.128e-03
## Station.NameCalgary Southeast:PM2.5.Mass	1.253e-03	1.174e-03
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons	1.682e+01	5.648e+00
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons	3.824e+00	9.041e+00
## Station.NameCalgary Northwest:Total.Hydrocarbons	-1.584e+01	5.646e+00
## Station.NameCalgary Southeast:Total.Hydrocarbons	-3.811e+00	9.030e+00
## Nitrogen.Dioxide:Methane	-3.842e+01	9.648e+00
## Ozone:Total.Hydrocarbons	-1.320e+01	2.548e+00
## Ozone:Station.NameCalgary Northwest	-2.361e+00	8.719e-01
## Ozone:Station.NameCalgary Southeast	5.562e-01	9.676e-01
## Ozone:Nitric.Oxide	1.643e+02	6.143e+01
## Station.NameCalgary Northwest:Methane	1.566e+01	5.646e+00
## Station.NameCalgary Southeast:Methane	3.981e+00	9.028e+00
## Methane:Total.Hydrocarbons	-1.047e-01	4.572e-02
## Non.methane.Hydrocarbons:Carbon.Monoxide	-2.303e+00	6.907e-01
## Nitrogen.Dioxide:Total.Hydrocarbons	2.744e+01	9.186e+00
## Station.NameCalgary Northwest:Carbon.Monoxide	3.810e-01	1.319e-01
## Station.NameCalgary Southeast:Carbon.Monoxide	3.903e-01	1.830e-01
##		
## (Intercept)	-1.526	0.127076
## Total.Oxides.Of.Nitrogen	2.463	0.013840 *
## I(Total.Oxides.Of.Nitrogen^2)	-2.839	0.004547 **
## Ozone	16.463	< 2e-16 ***
## Nitrogen.Dioxide	2.230	0.025829 *
## Station.NameCalgary Northwest	2.583	0.009845 **
## Station.NameCalgary Southeast	-2.968	0.003014 **
## Methane	0.719	0.472196
## Nitric.Oxide	-0.392	0.695448
## Total.Hydrocarbons	-0.468	0.639882
## Non.methane.Hydrocarbons	0.633	0.526563
## Carbon.Monoxide	10.592	< 2e-16 ***
## PM2.5.Mass	5.312	1.16e-07 ***
## Ozone:PM2.5.Mass	-7.606	3.63e-14 ***
## Total.Hydrocarbons:PM2.5.Mass	4.088	4.45e-05 ***
## Nitric.Oxide:PM2.5.Mass	-7.077	1.78e-12 ***
## Ozone:Carbon.Monoxide	-9.852	< 2e-16 ***
## Nitric.Oxide:Carbon.Monoxide	-5.352	9.30e-08 ***
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide	5.390	7.54e-08 ***
## Nitrogen.Dioxide:Nitric.Oxide	-6.533	7.42e-11 ***
## Total.Oxides.Of.Nitrogen:Nitric.Oxide	3.359	0.000792 ***
## Nitrogen.Dioxide:Station.NameCalgary Northwest	-7.226	6.11e-13 ***
## Nitrogen.Dioxide:Station.NameCalgary Southeast	-4.242	2.28e-05 ***
## Station.NameCalgary Northwest:PM2.5.Mass	-3.865	0.000113 ***
## Station.NameCalgary Southeast:PM2.5.Mass	1.067	0.286015
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons	2.978	0.002917 **
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons	0.423	0.672312
t value	Pr(> t )	

```

## Station.NameCalgary Northwest:Total.Hydrocarbons      -2.805 0.005065 ***
## Station.NameCalgary Southeast:Total.Hydrocarbons     -0.422 0.673061
## Nitrogen.Dioxide:Methane                          -3.982 6.98e-05 ***
## Ozone:Total.Hydrocarbons                         -5.181 2.34e-07 ***
## Ozone:Station.NameCalgary Northwest            -2.708 0.006803 **
## Ozone:Station.NameCalgary Southeast              0.575 0.565425
## Ozone:Nitric.Oxide                            2.675 0.007500 **
## Station.NameCalgary Northwest:Methane           2.773 0.005580 **
## Station.NameCalgary Southeast:Methane            0.441 0.659313
## Methane:Total.Hydrocarbons                      -2.290 0.022103 *
## Non.methane.Hydrocarbons:Carbon.Monoxide       -3.334 0.000864 ***
## Nitrogen.Dioxide:Total.Hydrocarbons             2.987 0.002834 **
## Station.NameCalgary Northwest:Carbon.Monoxide   2.889 0.003893 **
## Station.NameCalgary Southeast:Carbon.Monoxide    2.133 0.032974 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1467 on 3389 degrees of freedom
## Multiple R-squared:  0.9138, Adjusted R-squared:  0.9128
## F-statistic: 898.7 on 40 and 3389 DF,  p-value: < 2.2e-16

```

All the terms including the interaction terms are significant till now. Now degree 3, let's test that

```

model3 <- lm(Air.Quality.Index ~
               Total.Oxides.Of.Nitrogen + I(Total.Oxides.Of.Nitrogen^2) + I(Total.Oxides.Of.Nitrogen +
               Nitric.Oxide + Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
               Ozone * PM2.5.Mass +
               PM2.5.Mass * Total.Hydrocarbons +
               PM2.5.Mass * Nitric.Oxide +
               Ozone * Carbon.Monoxide +
               Nitric.Oxide * Carbon.Monoxide +
               Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide +
               Nitrogen.Dioxide * Nitric.Oxide +
               Total.Oxides.Of.Nitrogen * Nitric.Oxide +
               Nitrogen.Dioxide * Station.Name +
               PM2.5.Mass * Station.Name +
               Station.Name * Non.methane.Hydrocarbons +
               Station.Name * Total.Hydrocarbons +
               Nitrogen.Dioxide * Methane +
               Ozone * Total.Hydrocarbons +
               Ozone * Station.Name +
               Ozone * Nitric.Oxide +
               Station.Name * Methane +
               Methane * Total.Hydrocarbons +
               Non.methane.Hydrocarbons * Carbon.Monoxide +
               Nitrogen.Dioxide * Total.Hydrocarbons +
               Station.Name * Carbon.Monoxide,
               data = air)

# Display the model summary
summary(model3)

```

##

```

## Call:
## lm(formula = Air.Quality.Index ~ Total.Oxides.Of.Nitrogen + I(Total.Oxides.Of.Nitrogen^2) +
##      I(Total.Oxides.Of.Nitrogen^3) + Ozone + Nitrogen.Dioxide +
##      Station.Name + Methane + Nitric.Oxide + Total.Hydrocarbons +
##      Non.methane.Hydrocarbons + Carbon.Monoxide + PM2.5.Mass +
##      Ozone * PM2.5.Mass + PM2.5.Mass * Total.Hydrocarbons + PM2.5.Mass *
##      Nitric.Oxide + Ozone * Carbon.Monoxide + Nitric.Oxide * Carbon.Monoxide +
##      Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide + Nitrogen.Dioxide *
##      Nitric.Oxide + Total.Oxides.Of.Nitrogen * Nitric.Oxide +
##      Nitrogen.Dioxide * Station.Name + PM2.5.Mass * Station.Name +
##      Station.Name * Non.methane.Hydrocarbons + Station.Name *
##      Total.Hydrocarbons + Nitrogen.Dioxide * Methane + Ozone *
##      Total.Hydrocarbons + Ozone * Station.Name + Ozone * Nitric.Oxide +
##      Station.Name * Methane + Methane * Total.Hydrocarbons + Non.methane.Hydrocarbons *
##      Carbon.Monoxide + Nitrogen.Dioxide * Total.Hydrocarbons +
##      Station.Name * Carbon.Monoxide, data = air)
##
## Residuals:
##       Min     1Q   Median     3Q    Max
## -1.84218 -0.07686  0.00297  0.07685  1.72384
##
## Coefficients:
##                               Estimate Std. Error
## (Intercept)                -2.997e-01 2.705e-01
## Total.Oxides.Of.Nitrogen        1.259e+01 5.482e+00
## I(Total.Oxides.Of.Nitrogen^2) -5.008e+02 1.883e+02
## I(Total.Oxides.Of.Nitrogen^3)  4.813e+02 1.091e+02
## Ozone                         7.758e+01 4.888e+00
## Nitrogen.Dioxide              1.356e+01 6.592e+00
## Station.NameCalgary Northwest 4.425e-01 1.625e-01
## Station.NameCalgary Southeast -3.602e-01 1.331e-01
## Methane                        1.713e+00 2.959e+00
## Nitric.Oxide                  6.707e+00 6.119e+00
## Total.Hydrocarbons            -1.064e+00 2.969e+00
## Non.methane.Hydrocarbons      1.647e+00 2.978e+00
## Carbon.Monoxide               1.951e+00 1.844e-01
## PM2.5.Mass                     3.052e-02 5.762e-03
## Ozone:PM2.5.Mass              -4.538e-01 5.938e-02
## Total.Hydrocarbons:PM2.5.Mass  1.032e-02 2.562e-03
## Nitric.Oxide:PM2.5.Mass       -3.329e-01 5.031e-02
## Ozone:Carbon.Monoxide         -5.415e+01 5.762e+00
## Nitric.Oxide:Carbon.Monoxide -1.982e+01 3.925e+00
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide 1.033e+03 2.057e+02
## Nitrogen.Dioxide:Nitric.Oxide -8.491e+02 1.512e+02
## Total.Oxides.Of.Nitrogen:Nitric.Oxide 4.289e+02 1.977e+02
## Nitrogen.Dioxide:Station.NameCalgary Northwest -1.117e+01 1.545e+00
## Nitrogen.Dioxide:Station.NameCalgary Southeast -6.999e+00 1.803e+00
## Station.NameCalgary Northwest:PM2.5.Mass -3.827e-03 1.131e-03
## Station.NameCalgary Southeast:PM2.5.Mass  1.879e-03 1.180e-03
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons 1.692e+01 5.632e+00
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons 3.666e+00 9.017e+00
## Station.NameCalgary Northwest:Total.Hydrocarbons     -1.603e+01 5.631e+00
## Station.NameCalgary Southeast:Total.Hydrocarbons     -3.791e+00 9.006e+00
## Nitrogen.Dioxide:Methane        -3.456e+01 9.661e+00

```

```

## Ozone:Total.Hydrocarbons           -1.165e+01  2.565e+00
## Ozone:Station.NameCalgary Northwest -2.787e+00  8.748e-01
## Ozone:Station.NameCalgary Southeast  2.027e-01  9.683e-01
## Ozone:Nitric.Oxide                4.404e+01  6.706e+01
## Station.NameCalgary Northwest:Methane 1.585e+01  5.631e+00
## Station.NameCalgary Southeast:Methane 3.952e+00  9.004e+00
## Methane:Total.Hydrocarbons          -9.348e-02  4.566e-02
## Non.methane.Hydrocarbons:Carbon.Monoxide -2.207e+00  6.891e-01
## Nitrogen.Dioxide:Total.Hydrocarbons   2.444e+01  9.187e+00
## Station.NameCalgary Northwest:Carbon.Monoxide 3.109e-01  1.325e-01
## Station.NameCalgary Southeast:Carbon.Monoxide 2.799e-01  1.842e-01
## t value Pr(>|t|)
## (Intercept)                      -1.108  0.267898
## Total.Oxides.Of.Nitrogen          2.296  0.021756 *
## I(Total.Oxides.Of.Nitrogen^2)     -2.660  0.007856 **
## Ozone                           4.412  1.06e-05 ***
## Nitrogen.Dioxide                 15.873 < 2e-16 ***
## Station.NameCalgary Northwest    2.057  0.039750 *
## Station.NameCalgary Southeast     2.723  0.006501 **
## Methane                          -2.705  0.006864 **
## Nitric.Oxide                     0.579  0.562647
## Total.Hydrocarbons               1.096  0.273160
## Non.methane.Hydrocarbons        -0.358  0.720073
## Carbon.Monoxide                  0.553  0.580277
## PM2.5.Mass                      10.579 < 2e-16 ***
## Ozone:PM2.5.Mass                5.296  1.26e-07 ***
## Total.Hydrocarbons:PM2.5.Mass   -7.642  2.77e-14 ***
## Nitric.Oxide:PM2.5.Mass         4.030  5.69e-05 ***
## Ozone:Carbon.Monoxide            -6.617  4.23e-11 ***
## Nitric.Oxide:Carbon.Monoxide    -9.397 < 2e-16 ***
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide  -5.050  4.65e-07 ***
## Nitrogen.Dioxide:Nitric.Oxide    -5.616  2.12e-08 ***
## Total.Oxides.Of.Nitrogen:Nitric.Oxide  5.021  5.41e-07 ***
## Nitrogen.Dioxide:Station.NameCalgary Northwest -7.231  5.90e-13 ***
## Nitrogen.Dioxide:Station.NameCalgary Southeast  -3.882  0.000105 ***
## Station.NameCalgary Northwest:PM2.5.Mass      -3.382  0.000727 ***
## Station.NameCalgary Southeast:PM2.5.Mass       1.593  0.111298
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons 3.005  0.002677 **
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons 0.407  0.684369
## Station.NameCalgary Northwest:Total.Hydrocarbons   -2.847  0.004439 **
## Station.NameCalgary Southeast:Total.Hydrocarbons   -0.421  0.673838
## Nitrogen.Dioxide:Methane          -3.577  0.000353 ***
## Ozone:Total.Hydrocarbons          -4.544  5.72e-06 ***
## Ozone:Station.NameCalgary Northwest -3.185  0.001458 **
## Ozone:Station.NameCalgary Southeast  0.209  0.834174
## Ozone:Nitric.Oxide                0.657  0.511421
## Station.NameCalgary Northwest:Methane 2.815  0.004900 **
## Station.NameCalgary Southeast:Methane 0.439  0.660768
## Methane:Total.Hydrocarbons        -2.047  0.040722 *
## Non.methane.Hydrocarbons:Carbon.Monoxide -3.202  0.001375 **
## Nitrogen.Dioxide:Total.Hydrocarbons   2.661  0.007831 **
## Station.NameCalgary Northwest:Carbon.Monoxide 2.347  0.019008 *
## Station.NameCalgary Southeast:Carbon.Monoxide 1.520  0.128623

```

```
## ---  
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Residual standard error: 0.1463 on 3388 degrees of freedom  
## Multiple R-squared: 0.9143, Adjusted R-squared: 0.9133  
## F-statistic: 882 on 41 and 3388 DF, p-value: < 2.2e-16
```

I have one interaction Ozone:Nitric.Oxide that becomes insignificant (0.51142). Because of this reason i cannot consider Total.Oxides.Of.Nitrogen of degree 3 in the model. so the highest degree of Total.Oxides.Of.Nitrogen can be 2 without losing the model fit.

considering Ozone higher drgress.

```

## Nitric.Oxide + Nitrogen.Dioxide * Station.Name + PM2.5.Mass *
## Station.Name + Station.Name * Non.methane.Hydrocarbons +
## Station.Name * Total.Hydrocarbons + Nitrogen.Dioxide * Methane +
## Ozone * Total.Hydrocarbons + Ozone * Station.Name + Ozone *
## Nitric.Oxide + Station.Name * Methane + Methane * Total.Hydrocarbons +
## Non.methane.Hydrocarbons * Carbon.Monoxide + Nitrogen.Dioxide *
## Total.Hydrocarbons + Station.Name * Carbon.Monoxide, data = air)
##
## Residuals:
##      Min     1Q   Median     3Q     Max
## -1.83383 -0.07401  0.00247  0.07614  1.68186
##
## Coefficients:
##                               Estimate Std. Error
## (Intercept)                6.398e-01  2.774e-01
## Total.Oxides.Of.Nitrogen    1.426e+01  5.373e+00
## I(Total.Oxides.Of.Nitrogen^2) -4.457e+02  1.846e+02
## Ozone                      4.414e+01  5.573e+00
## I(Ozone^2)                  3.402e+02  2.741e+01
## Nitrogen.Dioxide            9.263e+00  6.475e+00
## Station.NameCalgary Northwest 3.927e-01  1.593e-01
## Station.NameCalgary Southeast -3.966e-01  1.304e-01
## Methane                     2.155e+00  2.901e+00
## Nitric.Oxide                 -9.335e+00  5.689e+00
## Total.Hydrocarbons           -1.984e+00  2.911e+00
## Non.methane.Hydrocarbons    2.128e+00  2.920e+00
## Carbon.Monoxide              1.869e+00  1.810e-01
## PM2.5.Mass                   3.712e-02  5.675e-03
## Ozone:PM2.5.Mass             -5.504e-01  5.877e-02
## Total.Hydrocarbons:PM2.5.Mass 8.489e-03  2.517e-03
## Nitric.Oxide:PM2.5.Mass      -3.622e-01  4.909e-02
## Ozone:Carbon.Monoxide        -5.262e+01  5.634e+00
## Nitric.Oxide:Carbon.Monoxide -1.974e+01  3.842e+00
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide 8.484e+02  2.022e+02
## Nitrogen.Dioxide:Nitric.Oxide -6.535e+02  1.480e+02
## Total.Oxides.Of.Nitrogen:Nitric.Oxide 5.448e+02  1.880e+02
## Nitrogen.Dioxide:Station.NameCalgary Northwest -8.453e+00  1.531e+00
## Nitrogen.Dioxide:Station.NameCalgary Southeast -7.542e+00  1.762e+00
## Station.NameCalgary Northwest:PM2.5.Mass -4.630e-03  1.104e-03
## Station.NameCalgary Southeast:PM2.5.Mass 9.081e-04  1.149e-03
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons 1.672e+01  5.524e+00
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons 8.298e-01  8.847e+00
## Station.NameCalgary Northwest:Total.Hydrocarbons -1.579e+01  5.523e+00
## Station.NameCalgary Southeast:Total.Hydrocarbons -8.214e-01  8.836e+00
## Nitrogen.Dioxide:Methane      -2.894e+01  9.468e+00
## Ozone:Total.Hydrocarbons     -4.162e+00  2.596e+00
## Ozone:Station.NameCalgary Northwest -5.823e-01  8.648e-01
## Ozone:Station.NameCalgary Southeast -2.374e-01  9.486e-01
## Ozone:Nitric.Oxide            3.656e+02  6.224e+01
## Station.NameCalgary Northwest:Methane 1.558e+01  5.523e+00
## Station.NameCalgary Southeast:Methane 1.006e+00  8.834e+00
## Methane:Total.Hydrocarbons    -3.692e-02  4.505e-02
## Non.methane.Hydrocarbons:Carbon.Monoxide -1.755e+00  6.770e-01
## Nitrogen.Dioxide:Total.Hydrocarbons 2.254e+01  8.994e+00

```

```

## Station.NameCalgary Northwest:Carbon.Monoxide      3.771e-01  1.290e-01
## Station.NameCalgary Southeast:Carbon.Monoxide    3.483e-01  1.790e-01
##
## t value Pr(>|t|)
## (Intercept)                                2.307 0.021124 *
## Total.Oxides.Of.Nitrogen                   2.654 0.007993 **
## I(Total.Oxides.Of.Nitrogen^2)              -2.414 0.015827 *
## Ozone                                      7.920 3.20e-15 ***
## I(Ozone^2)                                 12.415 < 2e-16 ***
## Nitrogen.Dioxide                           1.431 0.152660
## Station.NameCalgary Northwest               2.465 0.013755 *
## Station.NameCalgary Southeast                -3.043 0.002362 **
## Methane                                     0.743 0.457529
## Nitric.Oxide                               -1.641 0.100932
## Total.Hydrocarbons                         -0.681 0.495627
## Non.methane.Hydrocarbons                  0.729 0.466328
## Carbon.Monoxide                            10.326 < 2e-16 ***
## PM2.5.Mass                                6.540 7.06e-11 ***
## Ozone:PM2.5.Mass                          -9.365 < 2e-16 ***
## Total.Hydrocarbons:PM2.5.Mass            3.372 0.000754 ***
## Nitric.Oxide:PM2.5.Mass                 -7.377 2.02e-13 ***
## Ozone:Carbon.Monoxide                     -9.340 < 2e-16 ***
## Nitric.Oxide:Carbon.Monoxide             -5.137 2.94e-07 ***
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide 4.197 2.77e-05 ***
## Nitrogen.Dioxide:Nitric.Oxide            -4.417 1.03e-05 ***
## Total.Oxides.Of.Nitrogen:Nitric.Oxide   2.898 0.003784 **
## Nitrogen.Dioxide:Station.NameCalgary Northwest -5.521 3.61e-08 ***
## Nitrogen.Dioxide:Station.NameCalgary Southeast -4.279 1.93e-05 ***
## Station.NameCalgary Northwest:PM2.5.Mass   -4.195 2.80e-05 ***
## Station.NameCalgary Southeast:PM2.5.Mass    0.790 0.429336
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons 3.026 0.002493 **
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons 0.094 0.925277
## Station.NameCalgary Northwest:Total.Hydrocarbons     -2.859 0.004275 **
## Station.NameCalgary Southeast:Total.Hydrocarbons     -0.093 0.925938
## Nitrogen.Dioxide:Methane                  -3.057 0.002255 **
## Ozone:Total.Hydrocarbons                  -1.603 0.108996
## Ozone:Station.NameCalgary Northwest       -0.673 0.500749
## Ozone:Station.NameCalgary Southeast        -0.250 0.802387
## Ozone:Nitric.Oxide                      5.875 4.63e-09 ***
## Station.NameCalgary Northwest:Methane      2.822 0.004803 **
## Station.NameCalgary Southeast:Methane       0.114 0.909367
## Methane:Total.Hydrocarbons                -0.819 0.412595
## Non.methane.Hydrocarbons:Carbon.Monoxide -2.592 0.009587 **
## Nitrogen.Dioxide:Total.Hydrocarbons        2.506 0.012253 *
## Station.NameCalgary Northwest:Carbon.Monoxide 2.924 0.003482 **
## Station.NameCalgary Southeast:Carbon.Monoxide 1.946 0.051752 .
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ',' 1
##
## Residual standard error: 0.1435 on 3388 degrees of freedom
## Multiple R-squared:  0.9176, Adjusted R-squared:  0.9166
## F-statistic: 920.2 on 41 and 3388 DF,  p-value: < 2.2e-16

```

in the ozone 2nd degree model we found interaction terms that are insignificant like `Ozone:Total.Hydrocarbons` with p-value= 0.108. So, we are not considering any higher order terms for `Ozone`.

## considering Nitrogen.Dioxide higher drgress.

```
model2 <- lm(Air.Quality.Index ~
              Total.Oxides.Of.Nitrogen + I(Total.Oxides.Of.Nitrogen^2) + Ozone + Nitrogen.Dioxide +
              Nitric.Oxide + Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
              Ozone * PM2.5.Mass +
              PM2.5.Mass * Total.Hydrocarbons +
              PM2.5.Mass * Nitric.Oxide +
              Ozone * Carbon.Monoxide +
              Nitric.Oxide * Carbon.Monoxide +
              Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide +
              Nitrogen.Dioxide * Nitric.Oxide +
              Total.Oxides.Of.Nitrogen * Nitric.Oxide +
              Nitrogen.Dioxide * Station.Name +
              PM2.5.Mass * Station.Name +
              Station.Name * Non.methane.Hydrocarbons +
              Station.Name * Total.Hydrocarbons +
              Nitrogen.Dioxide * Methane +
              Ozone * Total.Hydrocarbons +
              Ozone * Station.Name +
              Ozone * Nitric.Oxide +
              Station.Name * Methane +
              Methane * Total.Hydrocarbons +
              Non.methane.Hydrocarbons * Carbon.Monoxide +
              Nitrogen.Dioxide * Total.Hydrocarbons +
              Station.Name * Carbon.Monoxide,
              data = air)

# Display the model summary
summary(model2)
```

```
##
## Call:
## lm(formula = Air.Quality.Index ~ Total.Oxides.Of.Nitrogen + I(Total.Oxides.Of.Nitrogen^2) +
##      Ozone + Nitrogen.Dioxide + I(Nitrogen.Dioxide^2) + Station.Name +
##      Methane + Nitric.Oxide + Total.Hydrocarbons + Non.methane.Hydrocarbons +
##      Carbon.Monoxide + PM2.5.Mass + Ozone * PM2.5.Mass + PM2.5.Mass *
##      Total.Hydrocarbons + PM2.5.Mass * Nitric.Oxide + Ozone *
##      Carbon.Monoxide + Nitric.Oxide * Carbon.Monoxide + Total.Oxides.Of.Nitrogen *
##      Nitrogen.Dioxide + Nitrogen.Dioxide * Nitric.Oxide + Total.Oxides.Of.Nitrogen *
##      Nitric.Oxide + Nitrogen.Dioxide * Station.Name + PM2.5.Mass *
##      Station.Name + Station.Name * Non.methane.Hydrocarbons +
##      Station.Name * Total.Hydrocarbons + Nitrogen.Dioxide * Methane +
##      Ozone * Total.Hydrocarbons + Ozone * Station.Name + Ozone *
##      Nitric.Oxide + Station.Name * Methane + Methane * Total.Hydrocarbons +
##      Non.methane.Hydrocarbons * Carbon.Monoxide + Nitrogen.Dioxide *
##      Total.Hydrocarbons + Station.Name * Carbon.Monoxide, data = air)
##
## Residuals:
##       Min     1Q   Median     3Q    Max
## -1.84008 -0.07945  0.00394  0.07828  1.68420
##
```

	Estimate	Std. Error
## Coefficients:		
##	-4.569e-01	2.705e-01
## (Intercept)	2.882e+01	8.526e+00
## Total.Oxides.Of.Nitrogen	3.898e+02	4.376e+02
## I(Total.Oxides.Of.Nitrogen^2)	8.041e+01	4.865e+00
## Ozone	6.345e-01	8.929e+00
## Nitrogen.Dioxide	5.040e+03	2.150e+03
## I(Nitrogen.Dioxide^2)	4.330e-01	1.629e-01
## Station.NameCalgary Northwest	-4.020e-01	1.332e-01
## Station.NameCalgary Southeast	2.064e+00	2.964e+00
## Methane	-1.813e+01	8.905e+00
## Nitric.Oxide	-1.287e+00	2.974e+00
## Total.Hydrocarbons	1.683e+00	2.985e+00
## Non.methane.Hydrocarbons	1.992e+00	1.854e-01
## Carbon.Monoxide	2.867e-02	5.838e-03
## PM2.5.Mass	-4.550e-01	5.951e-02
## Ozone:PM2.5.Mass	1.149e-02	2.601e-03
## Total.Hydrocarbons:PM2.5.Mass	-3.613e-01	5.022e-02
## Nitric.Oxide:PM2.5.Mass	-5.732e+01	5.753e+00
## Ozone:Carbon.Monoxide	-2.170e+01	3.935e+00
## Nitric.Oxide:Carbon.Monoxide	-4.888e+03	2.567e+03
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide	4.137e+03	2.185e+03
## Nitrogen.Dioxide:Nitric.Oxide	-2.839e+02	4.404e+02
## Total.Oxides.Of.Nitrogen:Nitric.Oxide	-1.118e+01	1.548e+00
## Nitrogen.Dioxide:Station.NameCalgary Northwest	-7.108e+00	1.815e+00
## Nitrogen.Dioxide:Station.NameCalgary Southeast	-4.120e-03	1.132e-03
## Station.NameCalgary Northwest:PM2.5.Mass	1.493e-03	1.178e-03
## Station.NameCalgary Southeast:PM2.5.Mass	1.695e+01	5.644e+00
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons	4.488e+00	9.040e+00
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons	-1.590e+01	5.643e+00
## Station.NameCalgary Northwest:Total.Hydrocarbons	-4.459e+00	9.028e+00
## Station.NameCalgary Southeast:Total.Hydrocarbons	-3.676e+01	9.667e+00
## Nitrogen.Dioxide:Methane	-1.331e+01	2.546e+00
## Ozone:Total.Hydrocarbons	-2.348e+00	8.713e-01
## Ozone:Station.NameCalgary Northwest	6.190e-01	9.674e-01
## Ozone:Station.NameCalgary Southeast	1.681e+02	6.141e+01
## Ozone:Nitric.Oxide	1.572e+01	5.642e+00
## Station.NameCalgary Northwest:Methane	4.631e+00	9.026e+00
## Station.NameCalgary Southeast:Methane	-1.134e-01	4.584e-02
## Methane:Total.Hydrocarbons	-1.977e+00	7.040e-01
## Non.methane.Hydrocarbons:Carbon.Monoxide	2.555e+01	9.216e+00
## Nitrogen.Dioxide:Total.Hydrocarbons	3.616e-01	1.320e-01
## Station.NameCalgary Northwest:Carbon.Monoxide	3.587e-01	1.833e-01
## Station.NameCalgary Southeast:Carbon.Monoxide	t value	Pr(> t )
##	-1.689	0.091319 .
## (Intercept)	3.380	0.000734 ***
## Total.Oxides.Of.Nitrogen	0.891	0.373036
## I(Total.Oxides.Of.Nitrogen^2)	16.527	< 2e-16 ***
## Ozone	0.071	0.943348
## Nitrogen.Dioxide	2.344	0.019160 *
## I(Nitrogen.Dioxide^2)	2.659	0.007878 **
## Station.NameCalgary Northwest	-3.018	0.002567 **
## Station.NameCalgary Southeast	0.696	0.486184
## Methane		

```

## Nitric.Oxide           -2.036 0.041787 *
## Total.Hydrocarbons    -0.433 0.665174
## Non.methane.Hydrocarbons 0.564 0.573007
## Carbon.Monoxide        10.748 < 2e-16 ***
## PM2.5.Mass              4.912 9.44e-07 ***
## Ozone:PM2.5.Mass        -7.645 2.71e-14 ***
## Total.Hydrocarbons:PM2.5.Mass 4.416 1.04e-05 ***
## Nitric.Oxide:PM2.5.Mass -7.194 7.73e-13 ***
## Ozone:Carbon.Monoxide   -9.964 < 2e-16 ***
## Nitric.Oxide:Carbon.Monoxide -5.516 3.73e-08 ***
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide -1.904 0.056943 .
## Nitrogen.Dioxide:Nitric.Oxide      1.893 0.058472 .
## Total.Oxides.Of.Nitrogen:Nitric.Oxide -0.645 0.519131
## Nitrogen.Dioxide:Station.NameCalgary Northwest -7.221 6.36e-13 ***
## Nitrogen.Dioxide:Station.NameCalgary Southeast -3.916 9.17e-05 ***
## Station.NameCalgary Northwest:PM2.5.Mass -3.640 0.000277 ***
## Station.NameCalgary Southeast:PM2.5.Mass      1.268 0.204961
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons 3.003 0.002694 **
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons 0.497 0.619564
## Station.NameCalgary Northwest:Total.Hydrocarbons     -2.819 0.004852 **
## Station.NameCalgary Southeast:Total.Hydrocarbons      -0.494 0.621428
## Nitrogen.Dioxide:Methane          -3.803 0.000146 ***
## Ozone:Total.Hydrocarbons         -5.228 1.82e-07 ***
## Ozone:Station.NameCalgary Northwest -2.695 0.007082 **
## Ozone:Station.NameCalgary Southeast      0.640 0.522323
## Ozone:Nitric.Oxide             2.737 0.006238 **
## Station.NameCalgary Northwest:Methane      2.786 0.005364 **
## Station.NameCalgary Southeast:Methane       0.513 0.607932
## Methane:Total.Hydrocarbons      -2.474 0.013397 *
## Non.methane.Hydrocarbons:Carbon.Monoxide -2.809 0.005003 **
## Nitrogen.Dioxide:Total.Hydrocarbons      2.772 0.005602 **
## Station.NameCalgary Northwest:Carbon.Monoxide 2.738 0.006209 **
## Station.NameCalgary Southeast:Carbon.Monoxide 1.956 0.050497 .

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1466 on 3388 degrees of freedom
## Multiple R-squared:  0.914, Adjusted R-squared:  0.9129
## F-statistic: 878.1 on 41 and 3388 DF, p-value: < 2.2e-16

```

For 2nd degree of Nitrogen.Dioxide we have one interaction term that has very large p-value.that is Total.Oxides.Of.Nitrogen:Nitric.Oxide with p-value 0.519131. And Nitrogen.Dioxide 2nd order term itself is not has p-value greater than 0.05. so no higher order terms for Nitrogen.Dioxide.

**considering Carbon.Monoxide higher drgress.**

```

model2 <- lm(Air.Quality.Index ~
               Total.Oxides.Of.Nitrogen + I(Total.Oxides.Of.Nitrogen^2) + Ozone + Nitrogen.Dioxide +
               Nitric.Oxide + Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
               Ozone * PM2.5.Mass +
               PM2.5.Mass * Total.Hydrocarbons +

```

```

PM2.5.Mass * Nitric.Oxide +
Ozone * Carbon.Monoxide +
Nitric.Oxide * Carbon.Monoxide +
Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide +
Nitrogen.Dioxide * Nitric.Oxide +
Total.Oxides.Of.Nitrogen * Nitric.Oxide +
Nitrogen.Dioxide * Station.Name +
PM2.5.Mass * Station.Name +
Station.Name * Non.methane.Hydrocarbons +
Station.Name * Total.Hydrocarbons +
Nitrogen.Dioxide * Methane +
Ozone * Total.Hydrocarbons +
Ozone * Station.Name +
Ozone * Nitric.Oxide +
Station.Name * Methane +
Methane * Total.Hydrocarbons +
Non.methane.Hydrocarbons * Carbon.Monoxide +
Nitrogen.Dioxide * Total.Hydrocarbons +
Station.Name * Carbon.Monoxide,
data = air)

# Display the model summary
summary(model2)

```

```

##
## Call:
## lm(formula = Air.Quality.Index ~ Total.Oxides.Of.Nitrogen + I(Total.Oxides.Of.Nitrogen^2) +
##      Ozone + Nitrogen.Dioxide + Station.Name + Methane + Nitric.Oxide +
##      Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
##      I(Carbon.Monoxide^2) + PM2.5.Mass + Ozone * PM2.5.Mass +
##      PM2.5.Mass * Total.Hydrocarbons + PM2.5.Mass * Nitric.Oxide +
##      Ozone * Carbon.Monoxide + Nitric.Oxide * Carbon.Monoxide +
##      Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide + Nitrogen.Dioxide *
##      Nitric.Oxide + Total.Oxides.Of.Nitrogen * Nitric.Oxide +
##      Nitrogen.Dioxide * Station.Name + PM2.5.Mass * Station.Name +
##      Station.Name * Non.methane.Hydrocarbons + Station.Name *
##      Total.Hydrocarbons + Nitrogen.Dioxide * Methane + Ozone *
##      Total.Hydrocarbons + Ozone * Station.Name + Ozone * Nitric.Oxide +
##      Station.Name * Methane + Methane * Total.Hydrocarbons + Non.methane.Hydrocarbons *
##      Carbon.Monoxide + Nitrogen.Dioxide * Total.Hydrocarbons +
##      Station.Name * Carbon.Monoxide, data = air)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.83636 -0.07740  0.00302  0.07594  1.56582
##
## Coefficients:
##                               Estimate Std. Error
## (Intercept)                 -4.493e-01  2.687e-01
## Total.Oxides.Of.Nitrogen      1.466e+01  5.468e+00
## I(Total.Oxides.Of.Nitrogen^2) -5.698e+02  1.877e+02
## Ozone                         8.129e+01  4.845e+00
## Nitrogen.Dioxide                1.659e+01  6.578e+00

```

## Station.NameCalgary Northwest	4.287e-01	1.620e-01
## Station.NameCalgary Southeast	-3.435e-01	1.329e-01
## Methane	2.639e+00	2.951e+00
## Nitric.Oxide	-1.017e+00	5.761e+00
## Total.Hydrocarbons	-1.896e+00	2.962e+00
## Non.methane.Hydrocarbons	2.705e+00	2.973e+00
## Carbon.Monoxide	1.362e+00	2.088e-01
## I(Carbon.Monoxide^2)	1.095e+00	1.812e-01
## PM2.5.Mass	3.730e-02	5.851e-03
## Ozone:PM2.5.Mass	-5.240e-01	6.039e-02
## Total.Hydrocarbons:PM2.5.Mass	7.533e-03	2.602e-03
## Nitric.Oxide:PM2.5.Mass	-3.458e-01	4.995e-02
## Ozone:Carbon.Monoxide	-5.247e+01	5.762e+00
## Nitric.Oxide:Carbon.Monoxide	-2.894e+01	4.121e+00
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide	1.133e+03	2.045e+02
## Nitrogen.Dioxide:Nitric.Oxide	-1.001e+03	1.483e+02
## Total.Oxides.Of.Nitrogen:Nitric.Oxide	6.987e+02	1.912e+02
## Nitrogen.Dioxide:Station.NameCalgary Northwest	-1.215e+01	1.549e+00
## Nitrogen.Dioxide:Station.NameCalgary Southeast	-9.503e+00	1.819e+00
## Station.NameCalgary Northwest:PM2.5.Mass	-5.354e-03	1.134e-03
## Station.NameCalgary Southeast:PM2.5.Mass	2.093e-03	1.176e-03
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons	1.667e+01	5.618e+00
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons	2.714e+00	8.996e+00
## Station.NameCalgary Northwest:Total.Hydrocarbons	-1.575e+01	5.617e+00
## Station.NameCalgary Southeast:Total.Hydrocarbons	-2.836e+00	8.985e+00
## Nitrogen.Dioxide:Methane	-5.531e+01	9.997e+00
## Ozone:Total.Hydrocarbons	-1.364e+01	2.535e+00
## Ozone:Station.NameCalgary Northwest	-2.794e+00	8.703e-01
## Ozone:Station.NameCalgary Southeast	2.007e-01	9.644e-01
## Ozone:Nitric.Oxide	1.373e+02	6.127e+01
## Station.NameCalgary Northwest:Methane	1.558e+01	5.617e+00
## Station.NameCalgary Southeast:Methane	2.975e+00	8.983e+00
## Methane:Total.Hydrocarbons	-8.840e-02	4.556e-02
## Non.methane.Hydrocarbons:Carbon.Monoxide	-3.902e+00	7.363e-01
## Nitrogen.Dioxide:Total.Hydrocarbons	4.343e+01	9.514e+00
## Station.NameCalgary Northwest:Carbon.Monoxide	4.114e-01	1.313e-01
## Station.NameCalgary Southeast:Carbon.Monoxide	5.718e-01	1.845e-01
##	t value	Pr(> t )
## (Intercept)	-1.672	0.094563 .
## Total.Oxides.Of.Nitrogen	2.682	0.007361 **
## I(Total.Oxides.Of.Nitrogen^2)	-3.035	0.002422 **
## Ozone	16.777	< 2e-16 ***
## Nitrogen.Dioxide	2.522	0.011699 *
## Station.NameCalgary Northwest	2.646	0.008190 **
## Station.NameCalgary Southeast	-2.586	0.009765 **
## Methane	0.894	0.371305
## Nitric.Oxide	-0.177	0.859841
## Total.Hydrocarbons	-0.640	0.522104
## Non.methane.Hydrocarbons	0.910	0.363004
## Carbon.Monoxide	6.525	7.79e-11 ***
## I(Carbon.Monoxide^2)	6.040	1.71e-09 ***
## PM2.5.Mass	6.375	2.07e-10 ***
## Ozone:PM2.5.Mass	-8.677	< 2e-16 ***
## Total.Hydrocarbons:PM2.5.Mass	2.895	0.003811 **

```

## Nitric.Oxide:PM2.5.Mass           -6.924 5.24e-12 ***
## Ozone:Carbon.Monoxide            -9.107 < 2e-16 ***
## Nitric.Oxide:Carbon.Monoxide      -7.023 2.61e-12 ***
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide 5.542 3.22e-08 ***
## Nitrogen.Dioxide:Nitric.Oxide      -6.751 1.72e-11 ***
## Total.Oxides.Of.Nitrogen:Nitric.Oxide 3.653 0.000263 ***
## Nitrogen.Dioxide:Station.NameCalgary Northwest -7.841 5.93e-15 ***
## Nitrogen.Dioxide:Station.NameCalgary Southeast -5.225 1.85e-07 ***
## Station.NameCalgary Northwest:PM2.5.Mass       -4.721 2.44e-06 ***
## Station.NameCalgary Southeast:PM2.5.Mass        1.779 0.075289 .
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons 2.968 0.003022 **
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons 0.302 0.762875
## Station.NameCalgary Northwest:Total.Hydrocarbons    -2.804 0.005081 **
## Station.NameCalgary Southeast:Total.Hydrocarbons    -0.316 0.752328
## Nitrogen.Dioxide:Methane          -5.533 3.39e-08 ***
## Ozone:Total.Hydrocarbons          -5.378 8.03e-08 ***
## Ozone:Station.NameCalgary Northwest -3.210 0.001339 **
## Ozone:Station.NameCalgary Southeast 0.208 0.835157
## Ozone:Nitric.Oxide                2.241 0.025096 *
## Station.NameCalgary Northwest:Methane 2.774 0.005571 **
## Station.NameCalgary Southeast:Methane 0.331 0.740510
## Methane:Total.Hydrocarbons        -1.940 0.052424 .
## Non.methane.Hydrocarbons:Carbon.Monoxide -5.299 1.24e-07 ***
## Nitrogen.Dioxide:Total.Hydrocarbons     4.564 5.19e-06 ***
## Station.NameCalgary Northwest:Carbon.Monoxide 3.134 0.001741 **
## Station.NameCalgary Southeast:Carbon.Monoxide 3.100 0.001953 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1459 on 3388 degrees of freedom
## Multiple R-squared: 0.9148, Adjusted R-squared: 0.9137
## F-statistic: 886.8 on 41 and 3388 DF, p-value: < 2.2e-16

```

everything looks good. lets try , 3rd order now

```

model3 <- lm(Air.Quality.Index ~
              Total.Oxides.Of.Nitrogen + I(Total.Oxides.Of.Nitrogen^2) + Ozone + Nitrogen.Dioxide +
              Nitric.Oxide + Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
              Ozone * PM2.5.Mass +
              PM2.5.Mass * Total.Hydrocarbons +
              PM2.5.Mass * Nitric.Oxide +
              Ozone * Carbon.Monoxide +
              Nitric.Oxide * Carbon.Monoxide +
              Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide +
              Nitrogen.Dioxide * Nitric.Oxide +
              Total.Oxides.Of.Nitrogen * Nitric.Oxide +
              Nitrogen.Dioxide * Station.Name +
              PM2.5.Mass * Station.Name +
              Station.Name * Non.methane.Hydrocarbons +
              Station.Name * Total.Hydrocarbons +
              Nitrogen.Dioxide * Methane +
              Ozone * Total.Hydrocarbons +
              Ozone * Station.Name +
              Ozone * Nitric.Oxide +

```

```

Station.Name * Methane +
Methane * Total.Hydrocarbons +
Non.methane.Hydrocarbons * Carbon.Monoxide +
Nitrogen.Dioxide * Total.Hydrocarbons +
Station.Name * Carbon.Monoxide,
data = air)

# Display the model summary
summary(model3)

## 
## Call:
## lm(formula = Air.Quality.Index ~ Total.Oxides.Of.Nitrogen + I(Total.Oxides.Of.Nitrogen^2) +
##     Ozone + Nitrogen.Dioxide + Station.Name + Methane + Nitric.Oxide +
##     Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
##     I(Carbon.Monoxide^2) + I(Carbon.Monoxide^3) + PM2.5.Mass +
##     Ozone * PM2.5.Mass + PM2.5.Mass * Total.Hydrocarbons + PM2.5.Mass *
##     Nitric.Oxide + Ozone * Carbon.Monoxide + Nitric.Oxide * Carbon.Monoxide +
##     Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide + Nitrogen.Dioxide *
##     Nitric.Oxide + Total.Oxides.Of.Nitrogen * Nitric.Oxide +
##     Nitrogen.Dioxide * Station.Name + PM2.5.Mass * Station.Name +
##     Station.Name * Non.methane.Hydrocarbons + Station.Name *
##     Total.Hydrocarbons + Nitrogen.Dioxide * Methane + Ozone *
##     Total.Hydrocarbons + Ozone * Station.Name + Ozone * Nitric.Oxide +
##     Station.Name * Methane + Methane * Total.Hydrocarbons + Non.methane.Hydrocarbons *
##     Carbon.Monoxide + Nitrogen.Dioxide * Total.Hydrocarbons +
##     Station.Name * Carbon.Monoxide, data = air)
## 
## Residuals:
##      Min       1Q   Median      3Q      Max
## -1.82763 -0.07690  0.00331  0.07632  1.53521
## 
## Coefficients:
##                               Estimate Std. Error
## (Intercept)                -4.031e-01  2.679e-01
## Total.Oxides.Of.Nitrogen        1.682e+01  5.468e+00
## I(Total.Oxides.Of.Nitrogen^2) -5.966e+02  1.872e+02
## Ozone                      8.122e+01  4.829e+00
## Nitrogen.Dioxide             1.630e+01  6.556e+00
## Station.NameCalgary Northwest 3.592e-01  1.621e-01
## Station.NameCalgary Southeast -3.595e-01  1.325e-01
## Methane                     2.440e+00  2.942e+00
## Nitric.Oxide                 -2.862e+00  5.754e+00
## Total.Hydrocarbons           -1.640e+00  2.952e+00
## Non.methane.Hydrocarbons      2.503e+00  2.964e+00
## Carbon.Monoxide              2.714e-01  3.064e-01
## I(Carbon.Monoxide^2)          3.820e+00  5.901e-01
## I(Carbon.Monoxide^3)          -2.344e+00  4.831e-01
## PM2.5.Mass                  3.284e-02  5.904e-03
## Ozone:PM2.5.Mass             -5.159e-01  6.022e-02
## Total.Hydrocarbons:PM2.5.Mass 9.890e-03  2.638e-03
## Nitric.Oxide:PM2.5.Mass      -3.933e-01  5.074e-02
## Ozone:Carbon.Monoxide        -4.262e+01  6.091e+00

```

## Nitric.Oxide:Carbon.Monoxide	-2.576e+01	4.159e+00
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide	1.142e+03	2.039e+02
## Nitrogen.Dioxide:Nitric.Oxide	-1.021e+03	1.478e+02
## Total.Oxides.Of.Nitrogen:Nitric.Oxide	7.312e+02	1.907e+02
## Nitrogen.Dioxide:Station.NameCalgary Northwest	-1.315e+01	1.558e+00
## Nitrogen.Dioxide:Station.NameCalgary Southeast	-9.208e+00	1.814e+00
## Station.NameCalgary Northwest:PM2.5.Mass	-5.632e-03	1.132e-03
## Station.NameCalgary Southeast:PM2.5.Mass	1.618e-03	1.177e-03
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons	1.722e+01	5.601e+00
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons	2.826e+00	8.966e+00
## Station.NameCalgary Northwest:Total.Hydrocarbons	-1.650e+01	5.601e+00
## Station.NameCalgary Southeast:Total.Hydrocarbons	-3.030e+00	8.955e+00
## Nitrogen.Dioxide:Methane	-4.995e+01	1.002e+01
## Ozone:Total.Hydrocarbons	-1.440e+01	2.532e+00
## Ozone:Station.NameCalgary Northwest	-3.042e+00	8.689e-01
## Ozone:Station.NameCalgary Southeast	1.792e-01	9.612e-01
## Ozone:Nitric.Oxide	9.111e+01	6.181e+01
## Station.NameCalgary Northwest:Methane	1.637e+01	5.601e+00
## Station.NameCalgary Southeast:Methane	3.188e+00	8.953e+00
## Methane:Total.Hydrocarbons	-1.058e-01	4.555e-02
## Non.methane.Hydrocarbons:Carbon.Monoxide	-3.408e+00	7.409e-01
## Nitrogen.Dioxide:Total.Hydrocarbons	3.808e+01	9.546e+00
## Station.NameCalgary Northwest:Carbon.Monoxide	4.829e-01	1.317e-01
## Station.NameCalgary Southeast:Carbon.Monoxide	5.054e-01	1.844e-01
##		
## (Intercept)	-1.505	0.132528
## Total.Oxides.Of.Nitrogen	3.076	0.002112 **
## I(Total.Oxides.Of.Nitrogen^2)	-3.187	0.001450 **
## Ozone	16.819	< 2e-16 ***
## Nitrogen.Dioxide	2.486	0.012961 *
## Station.NameCalgary Northwest	2.216	0.026779 *
## Station.NameCalgary Southeast	-2.714	0.006677 **
## Methane	0.829	0.406942
## Nitric.Oxide	-0.497	0.619022
## Total.Hydrocarbons	-0.555	0.578638
## Non.methane.Hydrocarbons	0.845	0.398333
## Carbon.Monoxide	0.886	0.375764
## I(Carbon.Monoxide^2)	6.474	1.09e-10 ***
## I(Carbon.Monoxide^3)	-4.852	1.28e-06 ***
## PM2.5.Mass	5.563	2.86e-08 ***
## Ozone:PM2.5.Mass	-8.568	< 2e-16 ***
## Total.Hydrocarbons:PM2.5.Mass	3.749	0.000181 ***
## Nitric.Oxide:PM2.5.Mass	-7.752	1.19e-14 ***
## Ozone:Carbon.Monoxide	-6.996	3.16e-12 ***
## Nitric.Oxide:Carbon.Monoxide	-6.195	6.52e-10 ***
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide	5.602	2.29e-08 ***
## Nitrogen.Dioxide:Nitric.Oxide	-6.908	5.85e-12 ***
## Total.Oxides.Of.Nitrogen:Nitric.Oxide	3.834	0.000128 ***
## Nitrogen.Dioxide:Station.NameCalgary Northwest	-8.442	< 2e-16 ***
## Nitrogen.Dioxide:Station.NameCalgary Southeast	-5.077	4.03e-07 ***
## Station.NameCalgary Northwest:PM2.5.Mass	-4.976	6.83e-07 ***
## Station.NameCalgary Southeast:PM2.5.Mass	1.375	0.169173
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons	3.075	0.002124 **
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons	0.315	0.752665
t value	Pr(> t )	

```

## Station.NameCalgary Northwest:Total.Hydrocarbons      -2.945 0.003247 **
## Station.NameCalgary Southeast:Total.Hydrocarbons     -0.338 0.735128
## Nitrogen.Dioxide:Methane                          -4.983 6.58e-07 ***
## Ozone:Total.Hydrocarbons                         -5.687 1.40e-08 ***
## Ozone:Station.NameCalgary Northwest            -3.501 0.000470 ***
## Ozone:Station.NameCalgary Southeast             0.186 0.852126
## Ozone:Nitric.Oxide                            1.474 0.140539
## Station.NameCalgary Northwest:Methane           2.923 0.003494 **
## Station.NameCalgary Southeast:Methane            0.356 0.721810
## Methane:Total.Hydrocarbons                      -2.324 0.020210 *
## Non.methane.Hydrocarbons:Carbon.Monoxide       -4.599 4.39e-06 ***
## Nitrogen.Dioxide:Total.Hydrocarbons              3.989 6.77e-05 ***
## Station.NameCalgary Northwest:Carbon.Monoxide   3.667 0.000249 ***
## Station.NameCalgary Southeast:Carbon.Monoxide    2.741 0.006156 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1454 on 3387 degrees of freedom
## Multiple R-squared:  0.9154, Adjusted R-squared:  0.9143
## F-statistic: 872.1 on 42 and 3387 DF,  p-value: < 2.2e-16

```

Because of this interaction term `Ozone:Nitric.Oxide` having p-value =0.140539 , I cannot consider 3rd order term of `Carbon.Monoxide` in our model . Only considering 2nd order term for `Carbon.Monoxide`.

## considering PM2.5.Mass higher drgress.

2nd order

```

model2 <- lm(Air.Quality.Index ~
              Total.Oxides.Of.Nitrogen + I(Total.Oxides.Of.Nitrogen^2) + Ozone + Nitrogen.Dioxide +
              Nitric.Oxide + Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
              Ozone * PM2.5.Mass +
              PM2.5.Mass * Total.Hydrocarbons +
              PM2.5.Mass * Nitric.Oxide +
              Ozone * Carbon.Monoxide +
              Nitric.Oxide * Carbon.Monoxide +
              Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide +
              Nitrogen.Dioxide * Nitric.Oxide +
              Total.Oxides.Of.Nitrogen * Nitric.Oxide +
              Nitrogen.Dioxide * Station.Name +
              PM2.5.Mass * Station.Name +
              Station.Name * Non.methane.Hydrocarbons +
              Station.Name * Total.Hydrocarbons +
              Nitrogen.Dioxide * Methane +
              Ozone * Total.Hydrocarbons +
              Ozone * Station.Name +
              Ozone * Nitric.Oxide +
              Station.Name * Methane +
              Methane * Total.Hydrocarbons +
              Non.methane.Hydrocarbons * Carbon.Monoxide +
              Nitrogen.Dioxide * Total.Hydrocarbons +
              Station.Name * Carbon.Monoxide,

```

```

    data = air)

# Display the model summary
summary(model2)

## 
## Call:
## lm(formula = Air.Quality.Index ~ Total.Oxides.Of.Nitrogen + I(Total.Oxides.Of.Nitrogen^2) +
##     Ozone + Nitrogen.Dioxide + Station.Name + Methane + Nitric.Oxide +
##     Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
##     I(Carbon.Monoxide^2) + PM2.5.Mass + I(PM2.5.Mass^2) + Ozone *
##     PM2.5.Mass + PM2.5.Mass * Total.Hydrocarbons + PM2.5.Mass *
##     Nitric.Oxide + Ozone * Carbon.Monoxide + Nitric.Oxide * Carbon.Monoxide +
##     Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide + Nitrogen.Dioxide *
##     Nitric.Oxide + Total.Oxides.Of.Nitrogen * Nitric.Oxide +
##     Nitrogen.Dioxide * Station.Name + PM2.5.Mass * Station.Name +
##     Station.Name * Non.methane.Hydrocarbons + Station.Name *
##     Total.Hydrocarbons + Nitrogen.Dioxide * Methane + Ozone *
##     Total.Hydrocarbons + Ozone * Station.Name + Ozone * Nitric.Oxide +
##     Station.Name * Methane + Methane * Total.Hydrocarbons + Non.methane.Hydrocarbons *
##     Carbon.Monoxide + Nitrogen.Dioxide * Total.Hydrocarbons +
##     Station.Name * Carbon.Monoxide, data = air)
##
## Residuals:
##      Min       1Q   Median      3Q      Max
## -1.83346 -0.07787  0.00329  0.07660  1.49399
##
## 
## Coefficients:
##                               Estimate Std. Error
## (Intercept)                -4.058e-01 2.688e-01
## Total.Oxides.Of.Nitrogen        1.540e+01 5.468e+00
## I(Total.Oxides.Of.Nitrogen^2) -5.884e+02 1.876e+02
## Ozone                      8.119e+01 4.840e+00
## Nitrogen.Dioxide             1.620e+01 6.572e+00
## Station.NameCalgary Northwest 3.522e-01 1.640e-01
## Station.NameCalgary Southeast -3.892e-01 1.336e-01
## Methane                     2.827e+00 2.949e+00
## Nitric.Oxide                 -1.955e-01 5.762e+00
## Total.Hydrocarbons            -2.072e+00 2.959e+00
## Non.methane.Hydrocarbons      2.826e+00 2.970e+00
## Carbon.Monoxide               1.078e+00 2.304e-01
## I(Carbon.Monoxide^2)          1.666e+00 2.674e-01
## PM2.5.Mass                   3.420e-02 5.942e-03
## I(PM2.5.Mass^2)              -3.486e-05 1.202e-05
## Ozone:PM2.5.Mass              -4.937e-01 6.123e-02
## Total.Hydrocarbons:PM2.5.Mass 9.481e-03 2.684e-03
## Nitric.Oxide:PM2.5.Mass       -3.489e-01 4.991e-02
## Ozone:Carbon.Monoxide         -4.912e+01 5.870e+00
## Nitric.Oxide:Carbon.Monoxide -3.424e+01 4.503e+00
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide 1.150e+03 2.044e+02
## Nitrogen.Dioxide:Nitric.Oxide -1.011e+03 1.481e+02
## Total.Oxides.Of.Nitrogen:Nitric.Oxide 7.291e+02 1.913e+02
## Nitrogen.Dioxide:Station.NameCalgary Northwest -1.204e+01 1.548e+00

```

## Nitrogen.Dioxide:Station.NameCalgary Southeast	-9.253e+00	1.819e+00
## Station.NameCalgary Northwest:PM2.5.Mass	-5.782e-03	1.143e-03
## Station.NameCalgary Southeast:PM2.5.Mass	1.407e-03	1.199e-03
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons	1.665e+01	5.612e+00
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons	2.518e+00	8.986e+00
## Station.NameCalgary Northwest:Total.Hydrocarbons	-1.572e+01	5.611e+00
## Station.NameCalgary Southeast:Total.Hydrocarbons	-2.578e+00	8.975e+00
## Nitrogen.Dioxide:Methane	-5.533e+01	9.986e+00
## Ozone:Total.Hydrocarbons	-1.397e+01	2.535e+00
## Ozone:Station.NameCalgary Northwest	-2.824e+00	8.694e-01
## Ozone:Station.NameCalgary Southeast	2.851e-01	9.638e-01
## Ozone:Nitric.Oxide	1.118e+02	6.183e+01
## Station.NameCalgary Northwest:Methane	1.559e+01	5.611e+00
## Station.NameCalgary Southeast:Methane	2.744e+00	8.973e+00
## Methane:Total.Hydrocarbons	-9.875e-02	4.565e-02
## Non.methane.Hydrocarbons:Carbon.Monoxide	-3.920e+00	7.355e-01
## Nitrogen.Dioxide:Total.Hydrocarbons	4.331e+01	9.504e+00
## Station.NameCalgary Northwest:Carbon.Monoxide	3.783e-01	1.316e-01
## Station.NameCalgary Southeast:Carbon.Monoxide	5.512e-01	1.844e-01
##		
## (Intercept)	-1.510	0.131176
## Total.Oxides.Of.Nitrogen	2.816	0.004892 **
## I(Total.Oxides.Of.Nitrogen^2)	-3.136	0.001727 **
## Ozone	16.775	< 2e-16 ***
## Nitrogen.Dioxide	2.465	0.013742 *
## Station.NameCalgary Northwest	2.147	0.031836 *
## Station.NameCalgary Southeast	-2.912	0.003611 **
## Methane	0.958	0.337886
## Nitric.Oxide	-0.034	0.972934
## Total.Hydrocarbons	-0.700	0.483783
## Non.methane.Hydrocarbons	0.951	0.341465
## Carbon.Monoxide	4.681	2.97e-06 ***
## I(Carbon.Monoxide^2)	6.228	5.31e-10 ***
## PM2.5.Mass	5.756	9.37e-09 ***
## I(PM2.5.Mass^2)	-2.900	0.003750 **
## Ozone:PM2.5.Mass	-8.062	1.03e-15 ***
## Total.Hydrocarbons:PM2.5.Mass	3.532	0.000418 ***
## Nitric.Oxide:PM2.5.Mass	-6.992	3.25e-12 ***
## Ozone:Carbon.Monoxide	-8.367	< 2e-16 ***
## Nitric.Oxide:Carbon.Monoxide	-7.603	3.73e-14 ***
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide	5.629	1.97e-08 ***
## Nitrogen.Dioxide:Nitric.Oxide	-6.826	1.03e-11 ***
## Total.Oxides.Of.Nitrogen:Nitric.Oxide	3.811	0.000141 ***
## Nitrogen.Dioxide:Station.NameCalgary Northwest	-7.779	9.63e-15 ***
## Nitrogen.Dioxide:Station.NameCalgary Southeast	-5.087	3.83e-07 ***
## Station.NameCalgary Northwest:PM2.5.Mass	-5.061	4.39e-07 ***
## Station.NameCalgary Southeast:PM2.5.Mass	1.173	0.240689
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons	2.967	0.003028 **
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons	0.280	0.779303
## Station.NameCalgary Northwest:Total.Hydrocarbons	-2.801	0.005118 **
## Station.NameCalgary Southeast:Total.Hydrocarbons	-0.287	0.773950
## Nitrogen.Dioxide:Methane	-5.541	3.23e-08 ***
## Ozone:Total.Hydrocarbons	-5.509	3.88e-08 ***
## Ozone:Station.NameCalgary Northwest	-3.248	0.001175 **
t value	Pr(> t )	

```

## Ozone:Station.NameCalgary Southeast          0.296 0.767381
## Ozone:Nitric.Oxide                      1.809 0.070581 .
## Station.NameCalgary Northwest:Methane      2.779 0.005477 **
## Station.NameCalgary Southeast:Methane       0.306 0.759779
## Methane:Total.Hydrocarbons                -2.163 0.030595 *
## Non.methane.Hydrocarbons:Carbon.Monoxide -5.329 1.05e-07 ***
## Nitrogen.Dioxide:Total.Hydrocarbons        4.558 5.36e-06 ***
## Station.NameCalgary Northwest:Carbon.Monoxide 2.874 0.004080 **
## Station.NameCalgary Southeast:Carbon.Monoxide 2.989 0.002819 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1458 on 3387 degrees of freedom
## Multiple R-squared:  0.915, Adjusted R-squared:  0.9139
## F-statistic: 867.8 on 42 and 3387 DF,  p-value: < 2.2e-16

```

- There is one interaction term `Ozone:Nitric.Oxide` that has p-value greater than 0.05. So I cannot consider higher order terms for the `PM2.5.Mass`.

## Our final model is :

```

final_inter_model <- lm(Air.Quality.Index ~
                        Total.Oxides.Of.Nitrogen + I(Total.Oxides.Of.Nitrogen^2) + Ozone + Nitrogen.Dioxide +
                        Nitric.Oxide + Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
                        Ozone * PM2.5.Mass +
                        PM2.5.Mass * Total.Hydrocarbons +
                        PM2.5.Mass * Nitric.Oxide +
                        Ozone * Carbon.Monoxide +
                        Nitric.Oxide * Carbon.Monoxide +
                        Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide +
                        Nitrogen.Dioxide * Nitric.Oxide +
                        Total.Oxides.Of.Nitrogen * Nitric.Oxide +
                        Nitrogen.Dioxide * Station.Name +
                        PM2.5.Mass * Station.Name +
                        Station.Name * Non.methane.Hydrocarbons +
                        Station.Name * Total.Hydrocarbons +
                        Nitrogen.Dioxide * Methane +
                        Ozone * Total.Hydrocarbons +
                        Ozone * Station.Name +
                        Ozone * Nitric.Oxide +
                        Station.Name * Methane +
                        Methane * Total.Hydrocarbons +
                        Non.methane.Hydrocarbons * Carbon.Monoxide +
                        Nitrogen.Dioxide * Total.Hydrocarbons +
                        Station.Name * Carbon.Monoxide,
                        data = air)

# Display the model summary
summary(final_inter_model)

```

##

```

## Call:
## lm(formula = Air.Quality.Index ~ Total.Oxides.Of.Nitrogen + I(Total.Oxides.Of.Nitrogen^2) +
##     Ozone + Nitrogen.Dioxide + Station.Name + Methane + Nitric.Oxide +
##     Total.Hydrocarbons + Non.methane.Hydrocarbons + Carbon.Monoxide +
##     I(Carbon.Monoxide^2) + PM2.5.Mass + Ozone * PM2.5.Mass +
##     PM2.5.Mass * Total.Hydrocarbons + PM2.5.Mass * Nitric.Oxide +
##     Ozone * Carbon.Monoxide + Nitric.Oxide * Carbon.Monoxide +
##     Total.Oxides.Of.Nitrogen * Nitrogen.Dioxide + Nitrogen.Dioxide *
##     Nitric.Oxide + Total.Oxides.Of.Nitrogen * Nitric.Oxide +
##     Nitrogen.Dioxide * Station.Name + PM2.5.Mass * Station.Name +
##     Station.Name * Non.methane.Hydrocarbons + Station.Name *
##     Total.Hydrocarbons + Nitrogen.Dioxide * Methane + Ozone *
##     Total.Hydrocarbons + Ozone * Station.Name + Ozone * Nitric.Oxide +
##     Station.Name * Methane + Methane * Total.Hydrocarbons + Non.methane.Hydrocarbons *
##     Carbon.Monoxide + Nitrogen.Dioxide * Total.Hydrocarbons +
##     Station.Name * Carbon.Monoxide, data = air)
##
## Residuals:
##      Min       1Q    Median      3Q      Max
## -1.83636 -0.07740  0.00302  0.07594  1.56582
##
## Coefficients:
##                               Estimate Std. Error
## (Intercept)                -4.493e-01  2.687e-01
## Total.Oxides.Of.Nitrogen        1.466e+01  5.468e+00
## I(Total.Oxides.Of.Nitrogen^2) -5.698e+02  1.877e+02
## Ozone                      8.129e+01  4.845e+00
## Nitrogen.Dioxide              1.659e+01  6.578e+00
## Station.NameCalgary Northwest 4.287e-01  1.620e-01
## Station.NameCalgary Southeast -3.435e-01  1.329e-01
## Methane                     2.639e+00  2.951e+00
## Nitric.Oxide                 -1.017e+00  5.761e+00
## Total.Hydrocarbons            -1.896e+00  2.962e+00
## Non.methane.Hydrocarbons      2.705e+00  2.973e+00
## Carbon.Monoxide               1.362e+00  2.088e-01
## I(Carbon.Monoxide^2)          1.095e+00  1.812e-01
## PM2.5.Mass                   3.730e-02  5.851e-03
## Ozone:PM2.5.Mass              -5.240e-01  6.039e-02
## Total.Hydrocarbons:PM2.5.Mass 7.533e-03  2.602e-03
## Nitric.Oxide:PM2.5.Mass       -3.458e-01  4.995e-02
## Ozone:Carbon.Monoxide         -5.247e+01  5.762e+00
## Nitric.Oxide:Carbon.Monoxide -2.894e+01  4.121e+00
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide 1.133e+03  2.045e+02
## Nitrogen.Dioxide:Nitric.Oxide -1.001e+03  1.483e+02
## Total.Oxides.Of.Nitrogen:Nitric.Oxide 6.987e+02  1.912e+02
## Nitrogen.Dioxide:Station.NameCalgary Northwest -1.215e+01  1.549e+00
## Nitrogen.Dioxide:Station.NameCalgary Southeast -9.503e+00  1.819e+00
## Station.NameCalgary Northwest:PM2.5.Mass -5.354e-03  1.134e-03
## Station.NameCalgary Southeast:PM2.5.Mass 2.093e-03  1.176e-03
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons 1.667e+01  5.618e+00
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons 2.714e+00  8.996e+00
## Station.NameCalgary Northwest:Total.Hydrocarbons -1.575e+01  5.617e+00
## Station.NameCalgary Southeast:Total.Hydrocarbons -2.836e+00  8.985e+00
## Nitrogen.Dioxide:Methane      -5.531e+01  9.997e+00

```

```

## Ozone:Total.Hydrocarbons          -1.364e+01  2.535e+00
## Ozone:Station.NameCalgary Northwest -2.794e+00  8.703e-01
## Ozone:Station.NameCalgary Southeast 2.007e-01  9.644e-01
## Ozone:Nitric.Oxide             1.373e+02  6.127e+01
## Station.NameCalgary Northwest:Methane 1.558e+01  5.617e+00
## Station.NameCalgary Southeast:Methane 2.975e+00  8.983e+00
## Methane:Total.Hydrocarbons        -8.840e-02  4.556e-02
## Non.methane.Hydrocarbons:Carbon.Monoxide -3.902e+00  7.363e-01
## Nitrogen.Dioxide:Total.Hydrocarbons 4.343e+01  9.514e+00
## Station.NameCalgary Northwest:Carbon.Monoxide 4.114e-01  1.313e-01
## Station.NameCalgary Southeast:Carbon.Monoxide 5.718e-01  1.845e-01
## t value Pr(>|t|)
## (Intercept) -1.672 0.094563 .
## Total.Oxides.Of.Nitrogen 2.682 0.007361 **
## I(Total.Oxides.Of.Nitrogen^2) -3.035 0.002422 **
## Ozone 16.777 < 2e-16 ***
## Nitrogen.Dioxide 2.522 0.011699 *
## Station.NameCalgary Northwest 2.646 0.008190 **
## Station.NameCalgary Southeast -2.586 0.009765 **
## Methane 0.894 0.371305
## Nitric.Oxide -0.177 0.859841
## Total.Hydrocarbons -0.640 0.522104
## Non.methane.Hydrocarbons 0.910 0.363004
## Carbon.Monoxide 6.525 7.79e-11 ***
## I(Carbon.Monoxide^2) 6.040 1.71e-09 ***
## PM2.5.Mass 6.375 2.07e-10 ***
## Ozone:PM2.5.Mass -8.677 < 2e-16 ***
## Total.Hydrocarbons:PM2.5.Mass 2.895 0.003811 **
## Nitric.Oxide:PM2.5.Mass -6.924 5.24e-12 ***
## Ozone:Carbon.Monoxide -9.107 < 2e-16 ***
## Nitric.Oxide:Carbon.Monoxide -7.023 2.61e-12 ***
## Total.Oxides.Of.Nitrogen:Nitrogen.Dioxide 5.542 3.22e-08 ***
## Nitrogen.Dioxide:Nitric.Oxide -6.751 1.72e-11 ***
## Total.Oxides.Of.Nitrogen:Nitric.Oxide 3.653 0.000263 ***
## Nitrogen.Dioxide:Station.NameCalgary Northwest -7.841 5.93e-15 ***
## Nitrogen.Dioxide:Station.NameCalgary Southeast -5.225 1.85e-07 ***
## Station.NameCalgary Northwest:PM2.5.Mass -4.721 2.44e-06 ***
## Station.NameCalgary Southeast:PM2.5.Mass 1.779 0.075289 .
## Station.NameCalgary Northwest:Non.methane.Hydrocarbons 2.968 0.003022 **
## Station.NameCalgary Southeast:Non.methane.Hydrocarbons 0.302 0.762875
## Station.NameCalgary Northwest:Total.Hydrocarbons -2.804 0.005081 **
## Station.NameCalgary Southeast:Total.Hydrocarbons -0.316 0.752328
## Nitrogen.Dioxide:Methane -5.533 3.39e-08 ***
## Ozone:Total.Hydrocarbons -5.378 8.03e-08 ***
## Ozone:Station.NameCalgary Northwest -3.210 0.001339 **
## Ozone:Station.NameCalgary Southeast 0.208 0.835157
## Ozone:Nitric.Oxide 2.241 0.025096 *
## Station.NameCalgary Northwest:Methane 2.774 0.005571 **
## Station.NameCalgary Southeast:Methane 0.331 0.740510
## Methane:Total.Hydrocarbons -1.940 0.052424 .
## Non.methane.Hydrocarbons:Carbon.Monoxide -5.299 1.24e-07 ***
## Nitrogen.Dioxide:Total.Hydrocarbons 4.564 5.19e-06 ***
## Station.NameCalgary Northwest:Carbon.Monoxide 3.134 0.001741 **
## Station.NameCalgary Southeast:Carbon.Monoxide 3.100 0.001953 **

```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1459 on 3388 degrees of freedom
## Multiple R-squared:  0.9148, Adjusted R-squared:  0.9137
## F-statistic: 886.8 on 41 and 3388 DF,  p-value: < 2.2e-16
```