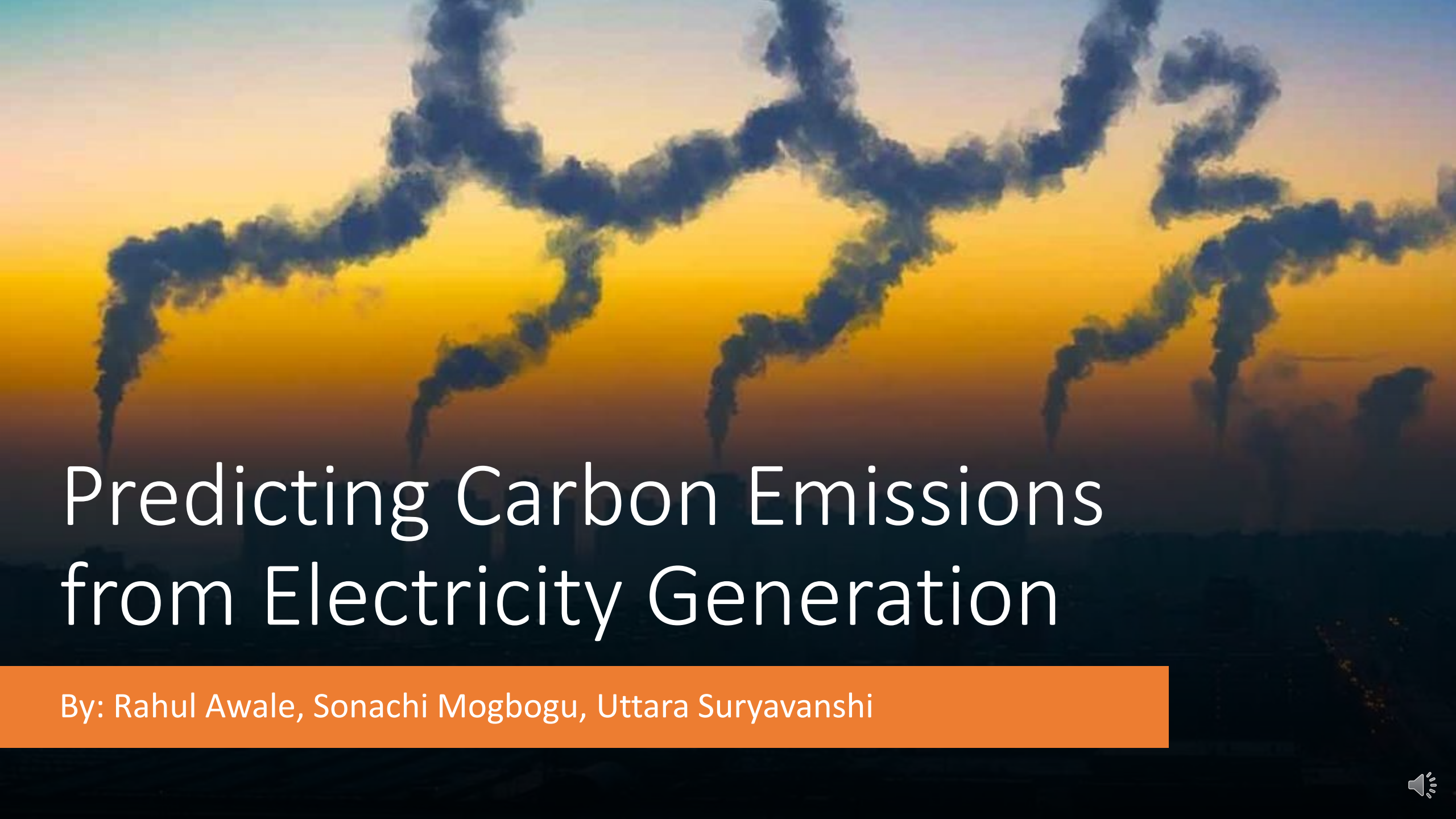


Predicting Carbon Emissions from Electricity Generation

By: Rahul Awale, Sonachi Mogbogu, Uttara Suryavanshi

Introduction

- Climate Change
- Increased CO_2 concentrations by 47%
- Increased by 90% since 1970
- Fossil Fuels and Industrial Process = 78% of CO_2 emission
- Using Time Series to Predict CO_2 emission

Motivation

- Apple Commits to be 100 percent carbon neutral by 2030
- Samsung to Offset Lifetime Carbon Footprint of All Washing Machines
- Reduce Carbon Emission
- What type of Carbon Emission is the top contributor
- Stop Global Warming

Research Question

Will Carbon Emission increase in the next 4 years?

Other Questions

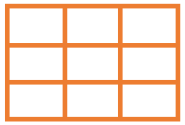
- Is There Any Correlation Between Type of Carbon Emission Source and Carbon Emission Value?
- What is the main source of Carbon Emission?
- What is the trend?

Data

U.S Energy Information Administration (EIA)

Monthly Records

Carbon Emission from 1973-2016



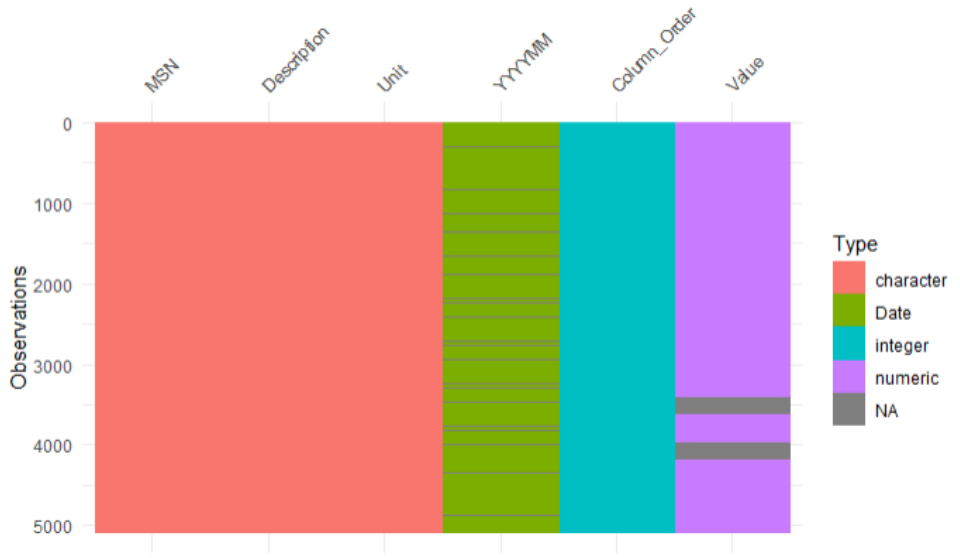
5094 x 6 Variables

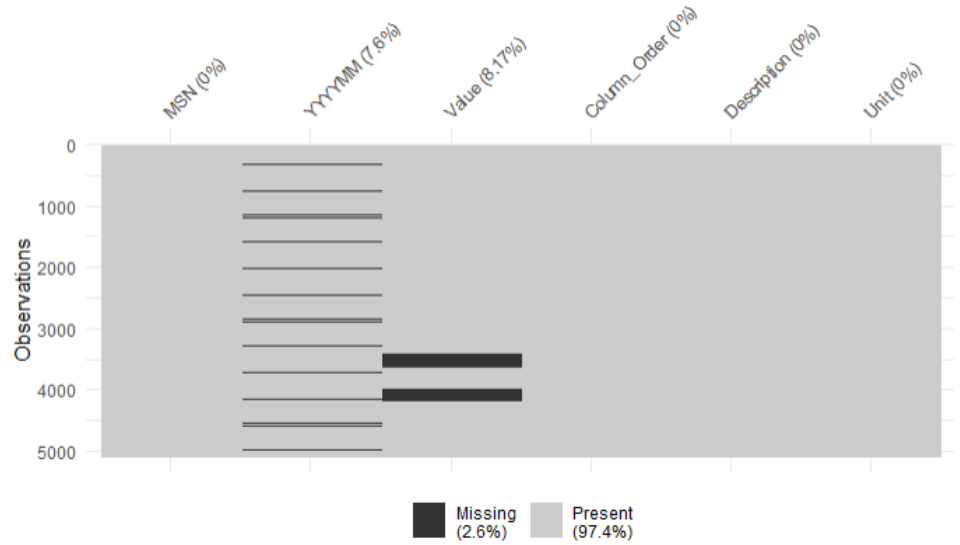
Variables

- MSN: Municipal Solid Waste – non-biomass component
- Description: Type
- Unit: Million Metric Tons of CO2
- YYYYMM: Date
- Column_Order: Column number based on type
- Value: Carbon Emission Value

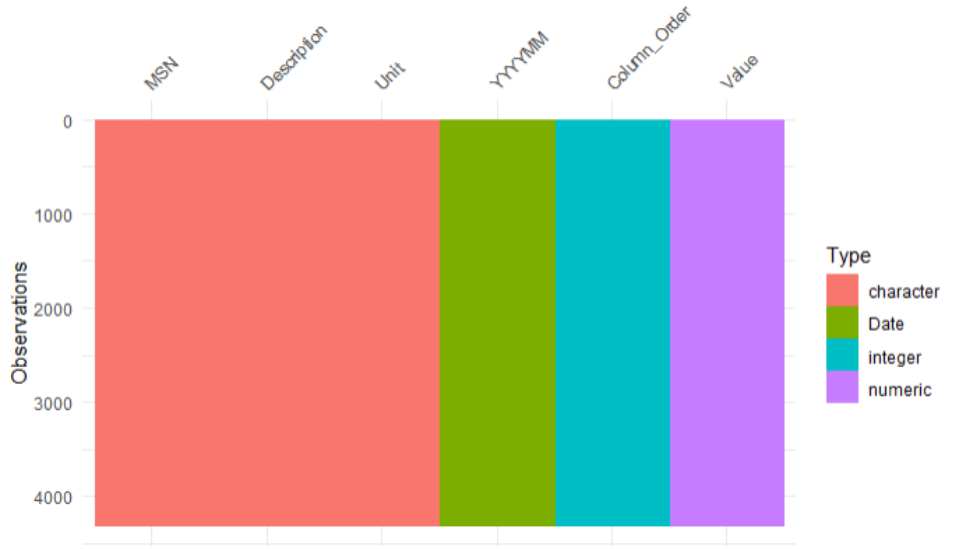
Raw Data

- 387 missing values(NA) for YYYYMM
- 416 missing values(NA) for Value
- Total 2.6% Missing Data





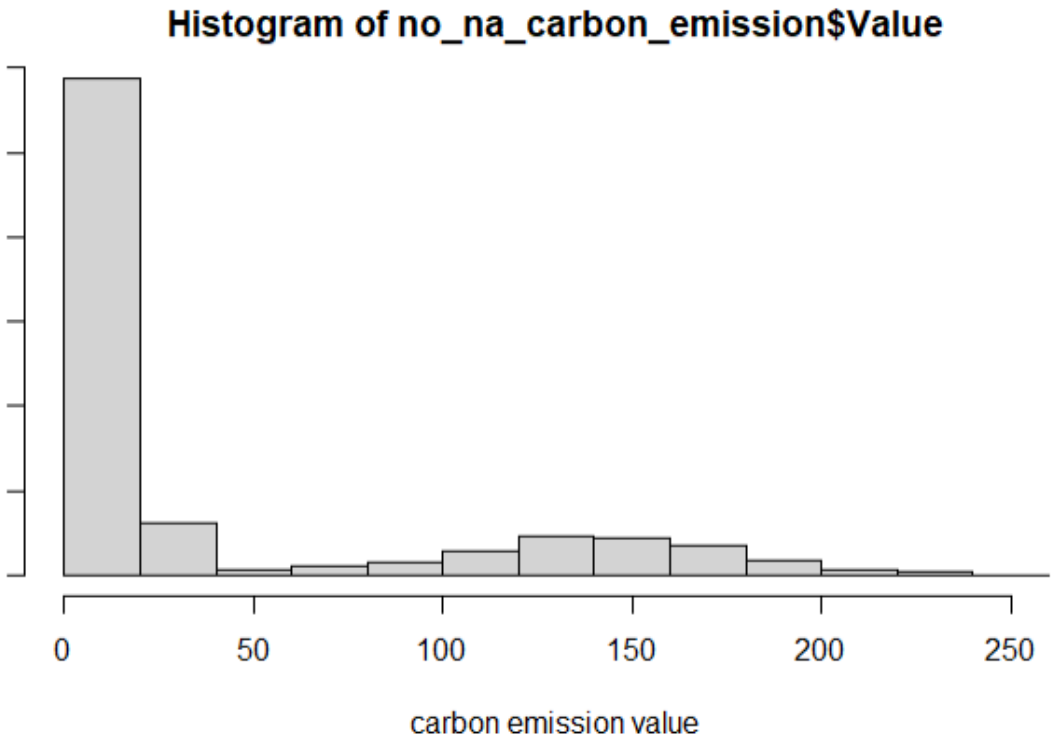
Pre-Processing



- Used mice function to remove NA
- New Data frame : no_na_carbon_emission
- 0 Missing Values

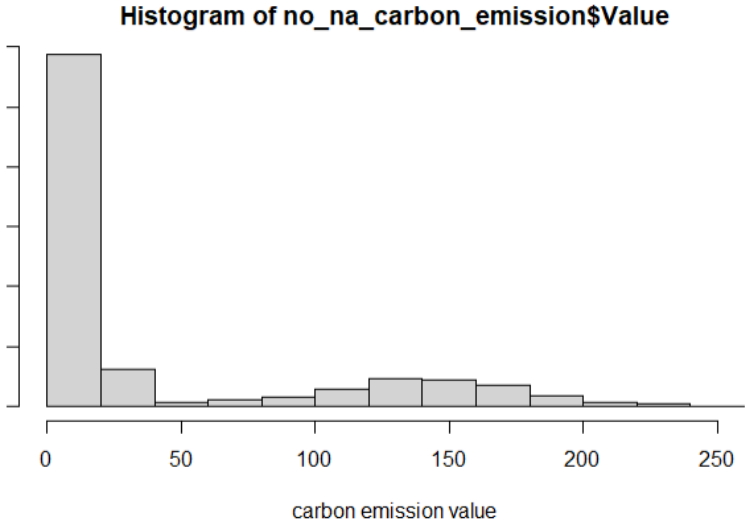
Normalization (Before)

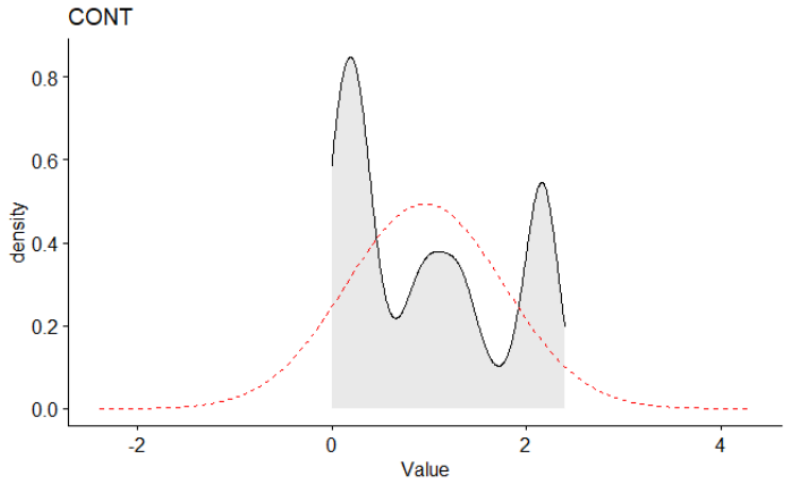
- Skewness= 1.43
- Kurtosis= 3.55
- W=0.67



Normalization (After)

- Skewness= 0.41
- Kurtosis= 1.69
- W=0.88

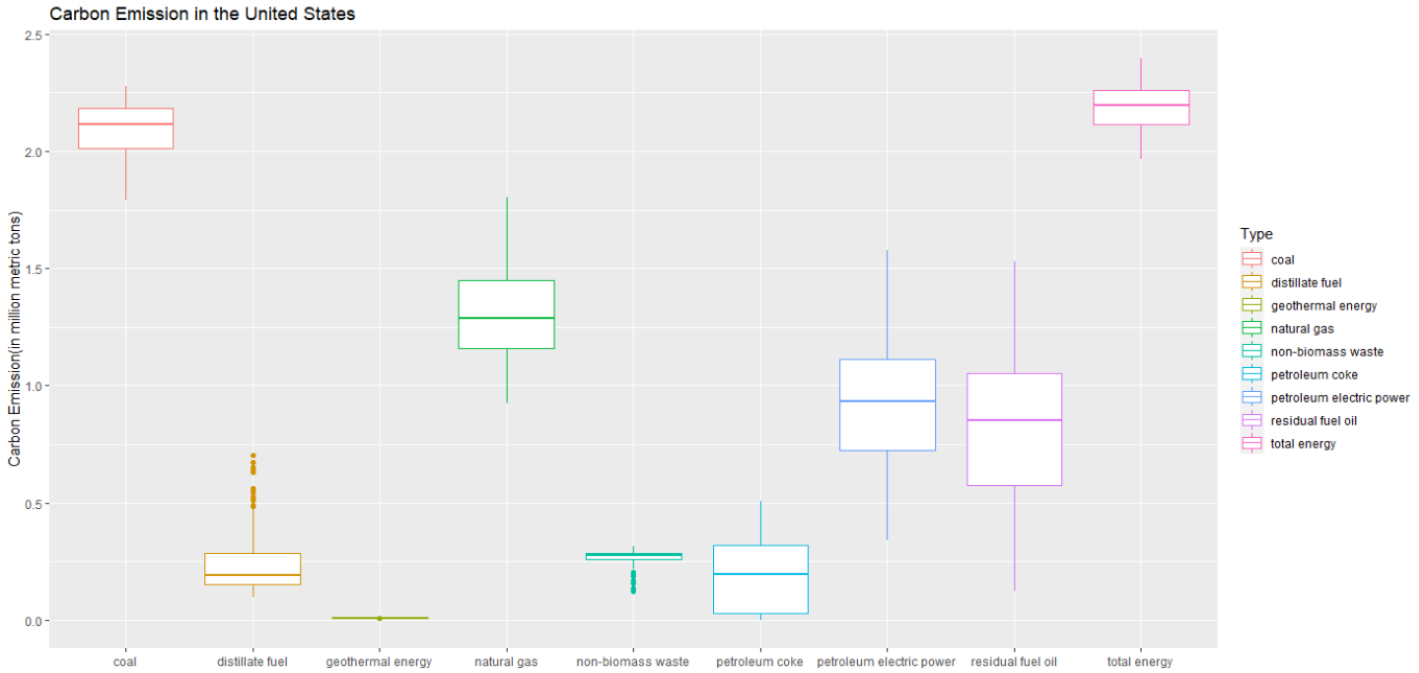




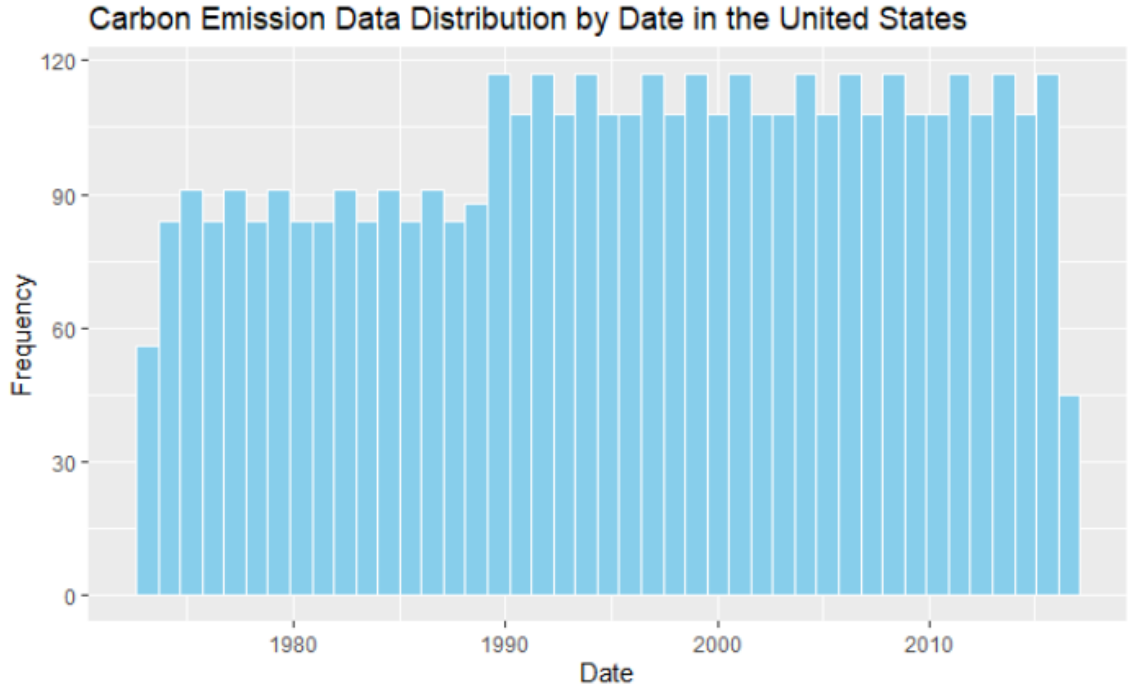
Exploratory Analysis

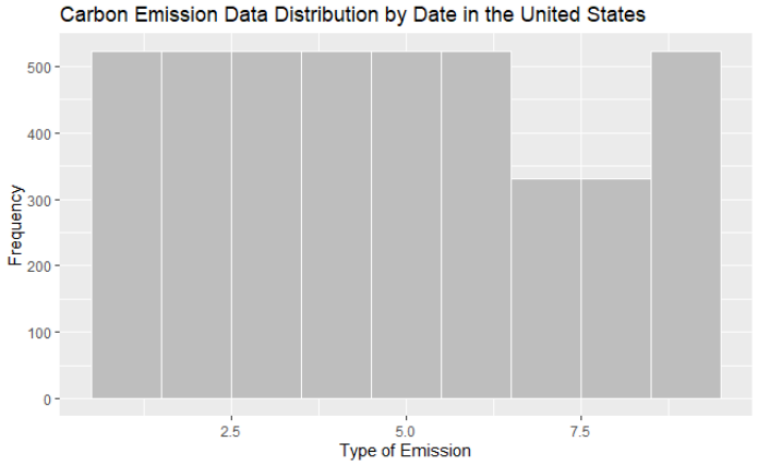
- Box Plot
- Histogram
- Correlation
- Bar Chart
- Line Chart
- Scatter Plot

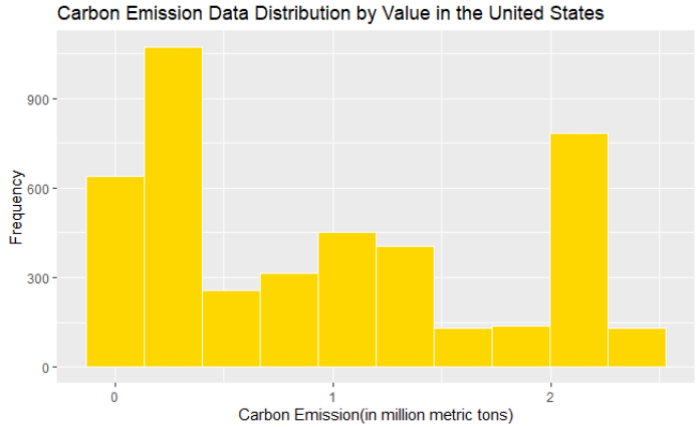
Boxplot


Carbon Emission in the United States
Carbon Emission(in million metric tons)
2.5
2.0
1.5
1.0
0.5
0.0
coal
distillate fuel
geothermal energy
natural gas
non-biomass waste
petroleum coke
petroleum electric power
residual fuel oil
total energy
Type
coal
distillate fuel
geothermal energy
natural gas
non-biomass waste
petroleum coke
petroleum electric power
residual fuel oil
total energy

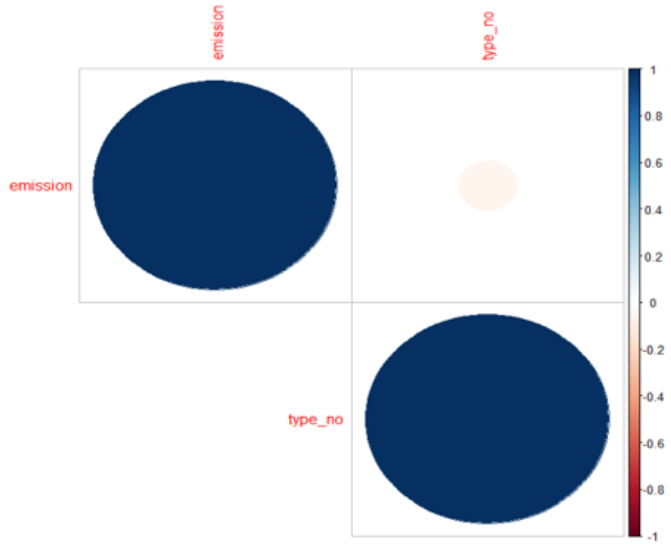
Histogram

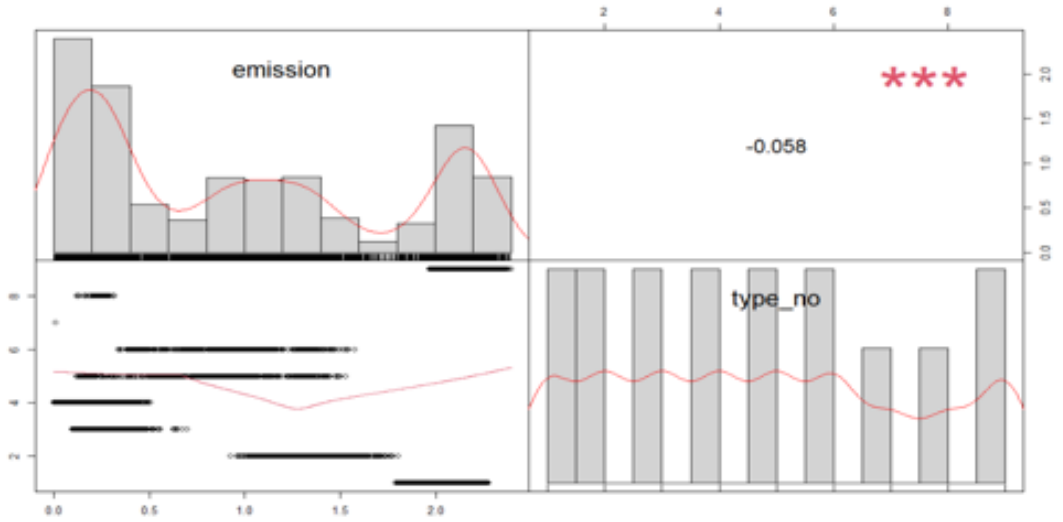

Carbon Emission Data Distribution by Date in the United States
120
90
60
30
0
Frequency
1980
1990
2000
2010
Date


Carbon Emission Data Distribution by Date in the United States
500
400
300
200
100
0
Frequency
2.5
5.0
7.5
Type of Emission

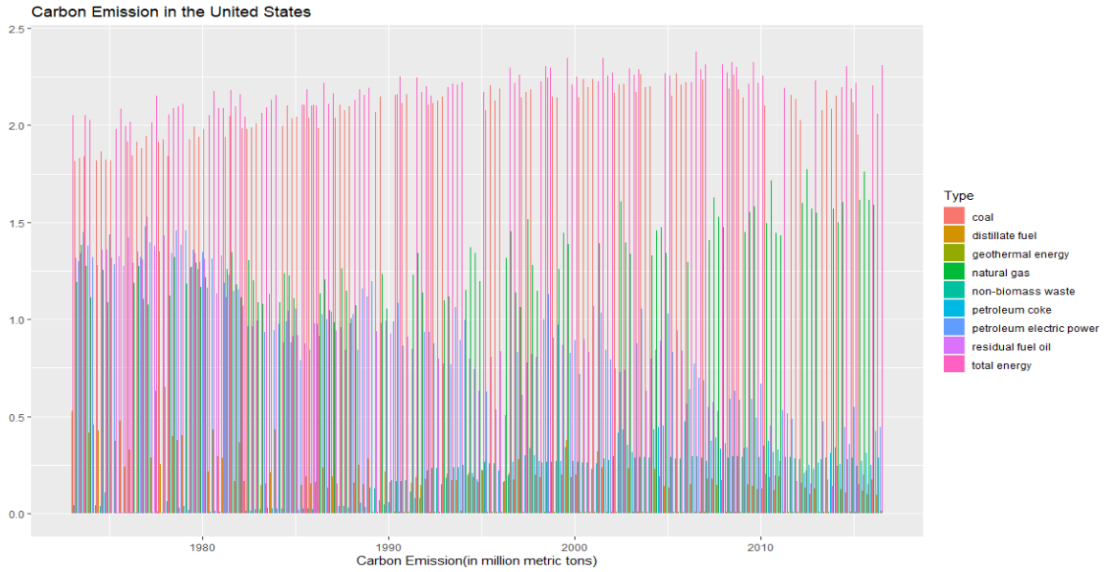

Carbon Emission Data Distribution by Value in the United States
900
600
300
0
Frequency
0
1
2
Carbon Emission(in million metric tons)

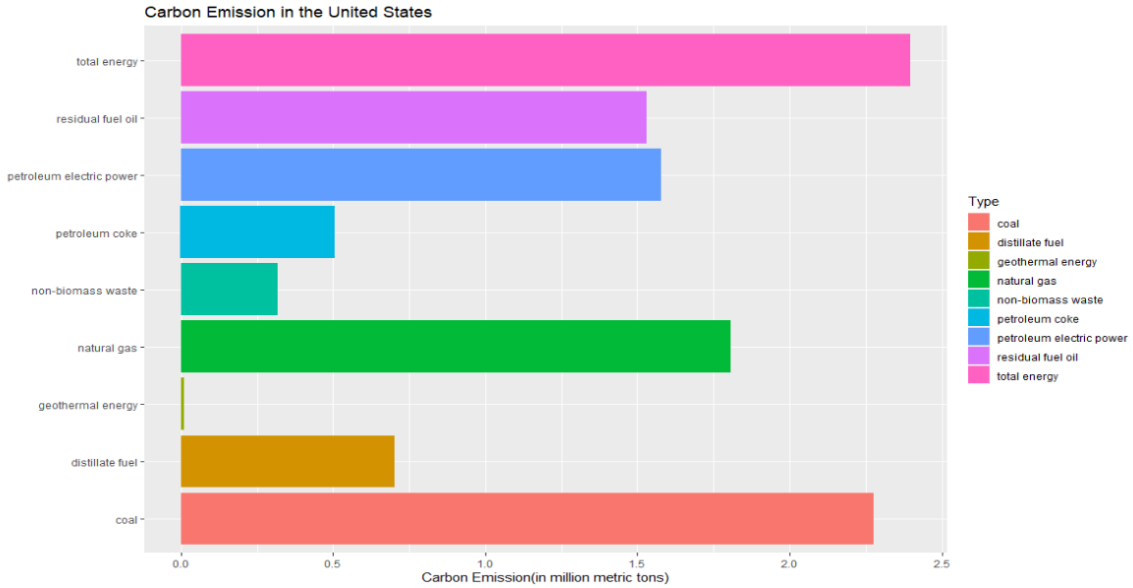
Correlation



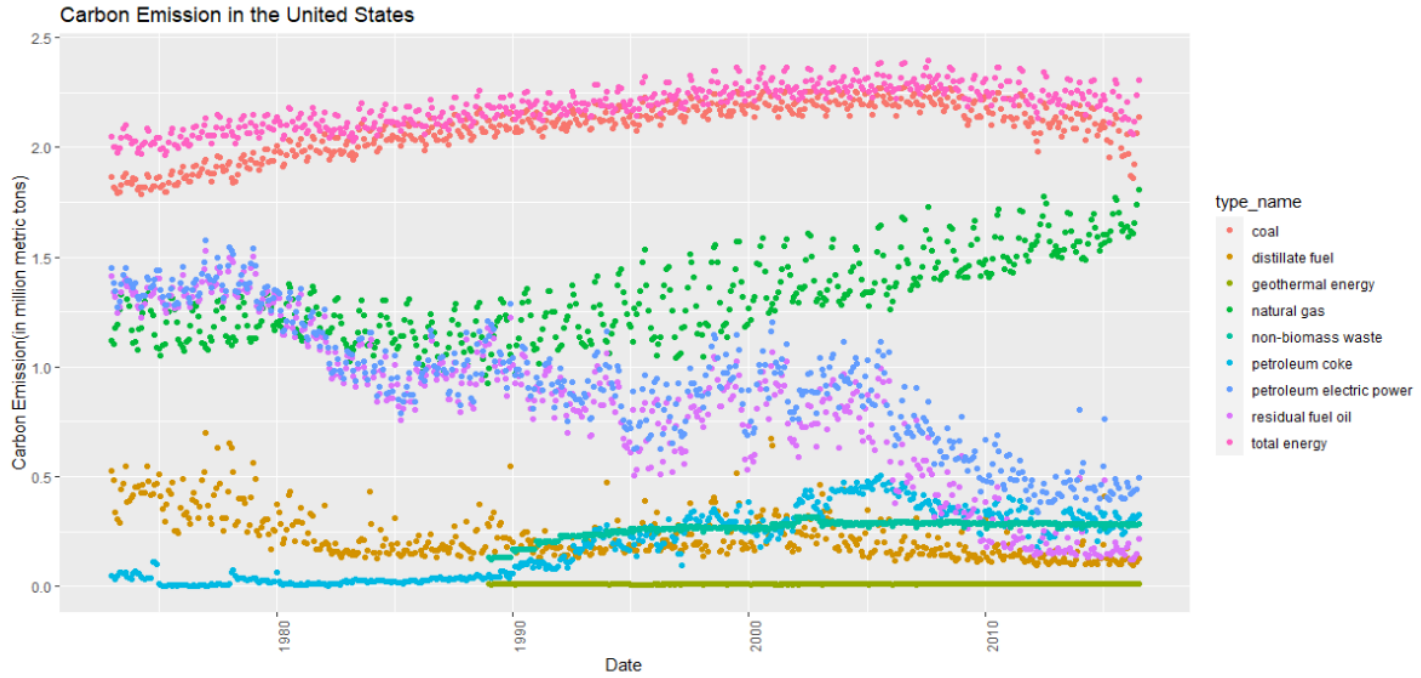


Bar Chart

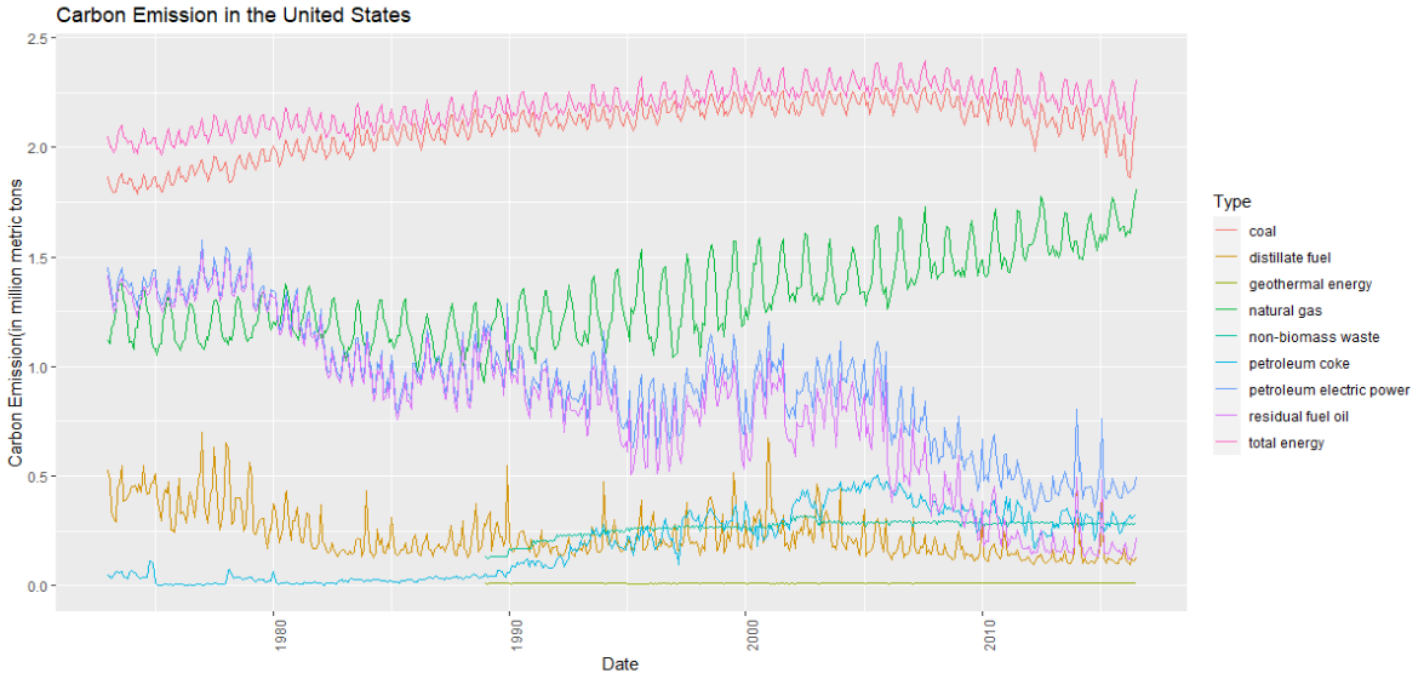

Carbon Emission in the United States
2.5
2.0
1.5
1.0
0.5
0.0
1980
1990
2000
2010
Carbon Emission(in million metric tons)
Type
coal
distillate fuel
geothermal energy
natural gas
non-biomass waste
petroleum coke
petroleum electric power
residual fuel oil
total energy


Carbon Emission in the United States
total energy
residual fuel oil
petroleum electric power
petroleum coke
non-biomass waste
natural gas
geothermal energy
distillate fuel
coal
0.0
0.5
1.0
1.5
2.0
2.5
Carbon Emission(in million metric tons)
Type
coal
distillate fuel
geothermal energy
natural gas
non-biomass waste
petroleum coke
petroleum electric power
residual fuel oil
total energy

Scatter Plot



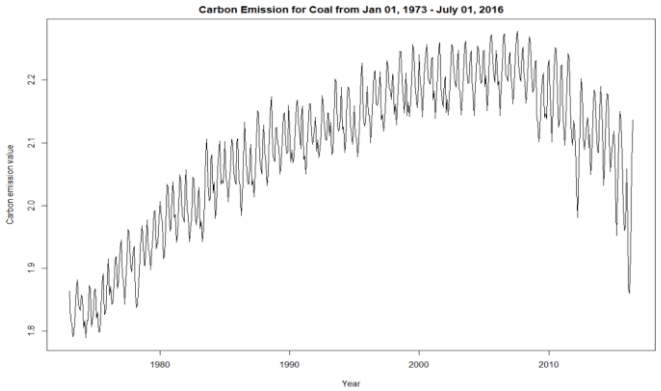
Line Chart

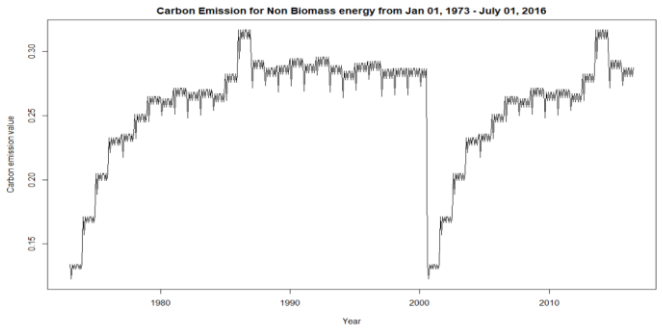

Carbon Emission in the United States
Carbon Emission(in million metric tons
2.5
2.0
1.5
1.0
0.5
0.0
1980
1990
2000
2010
Date
Type
coal
distillate fuel
geothermal energy
natural gas
non-biomass waste
petroleum coke
petroleum electric power
residual fuel oil
total energy

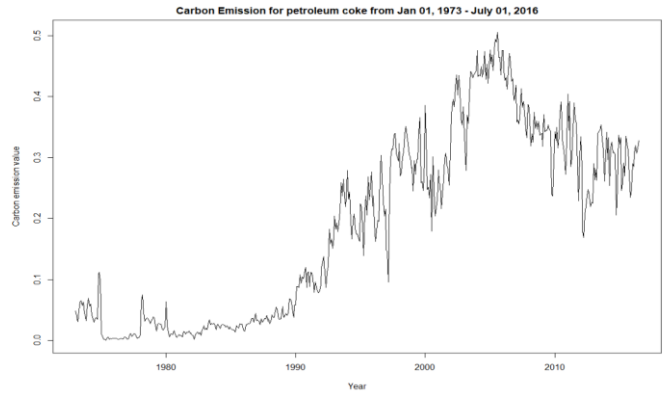
Methods

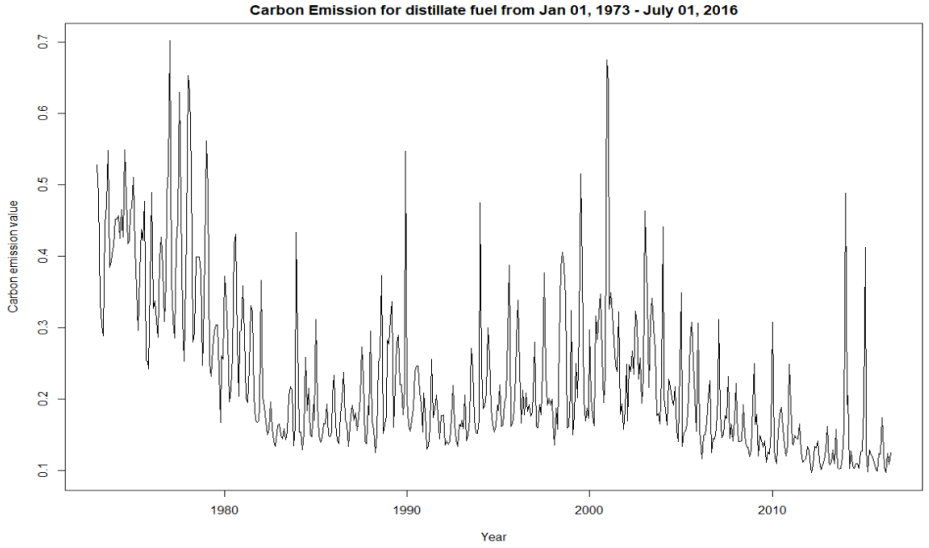
- Time Series
  - Cleaning Data
  - Understanding Data
  - Forecasting Data

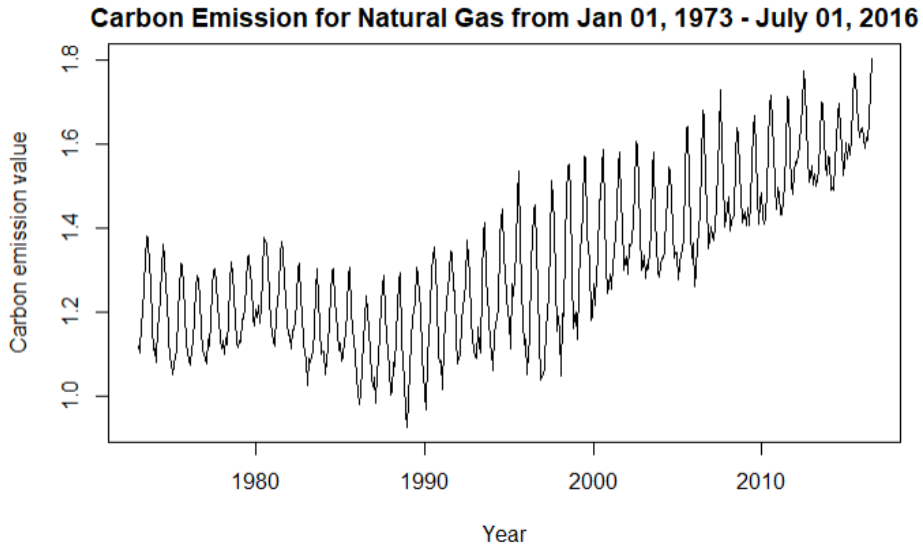
Visualization of Time Series Plots of Each Carbon Emission Over Time

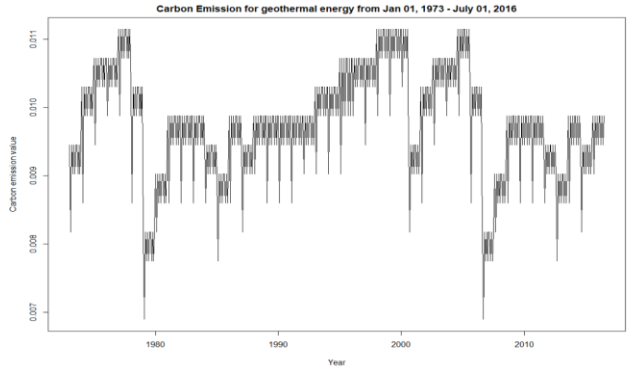


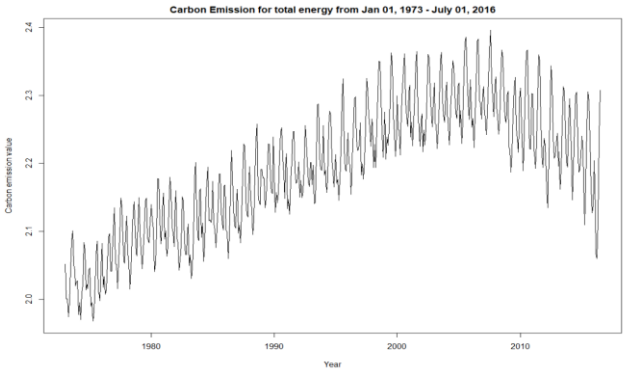


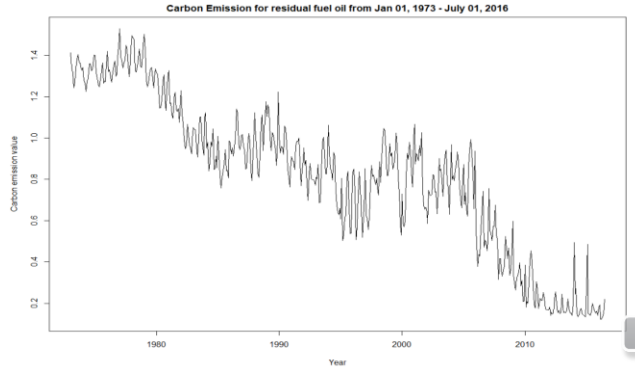










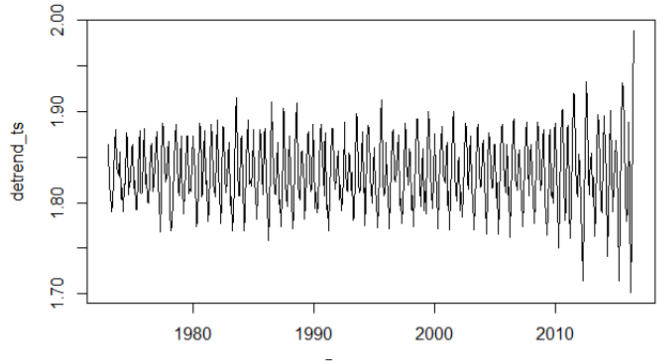


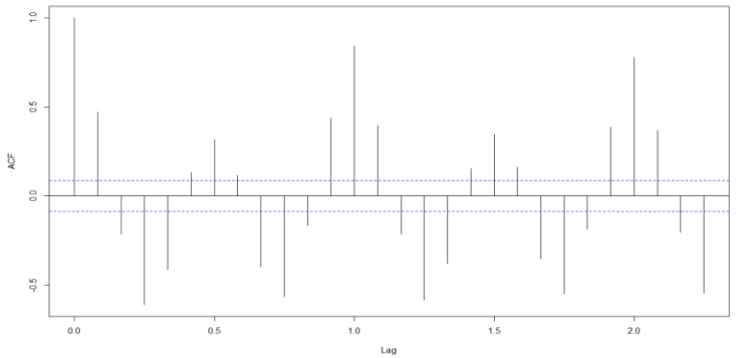
Augmented Dickey-Fuller Test

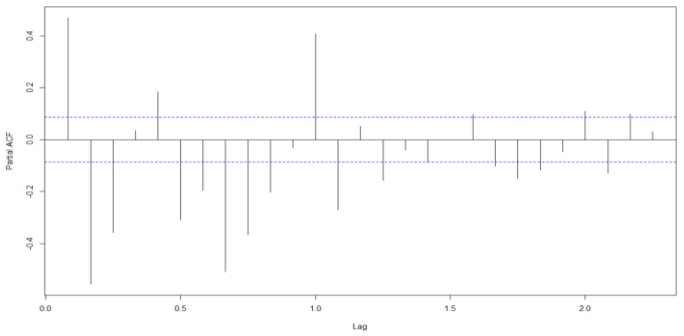
- Data: ts_allEmission
- -0.47868
- Lag Order = 8
- P-Value= 0.9828
- Stationary

Detrending

- Remove trend component
- Make it Stationary
- Dickey-Fuller- -25.335
- Lag order= 8
- P-value= 0.01
- Stationary



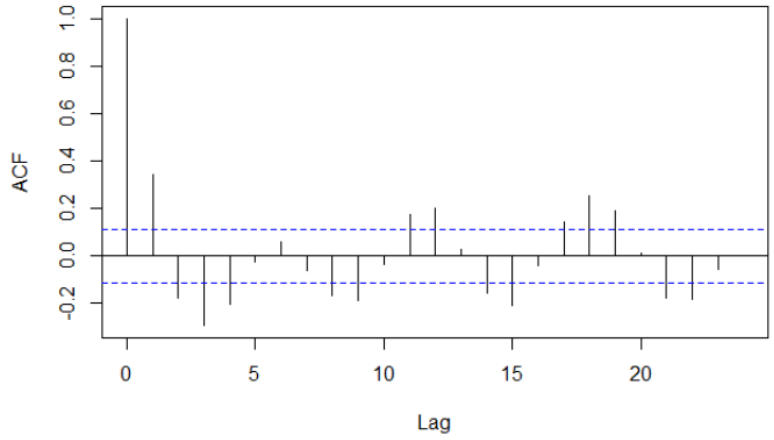


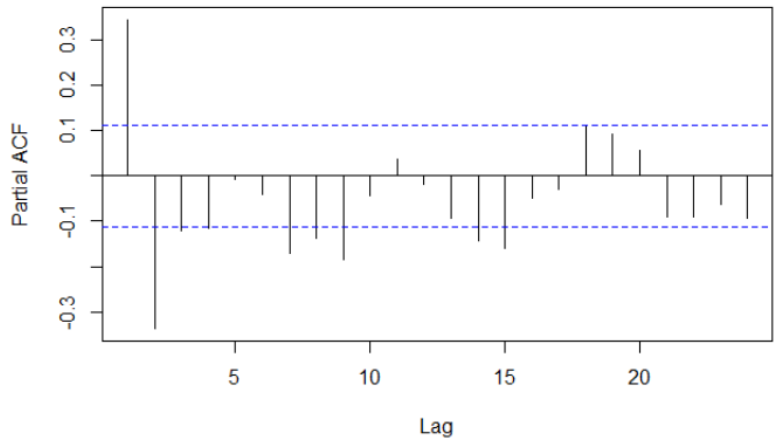


Train/Test


Train= (1990,1)-(2015,7)
Test= (2015,8)-(2016,7)
Time= time(carbonEmission_train)
Seas= factor(cycle(carbonEmission_train))
carbonEmissiong.lm<- carbonEmisssion_train~0+Time+Seas

Prediction Using Linear Regression Model





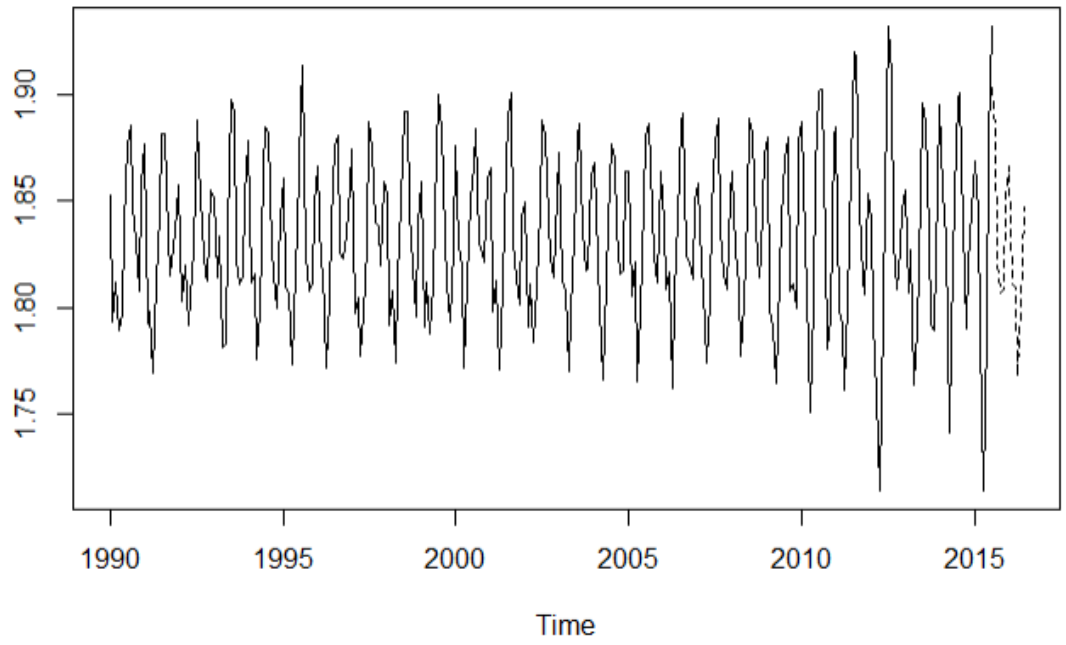

```
##        1        2        3        4        5        6        7        8
## 1.891336 1.889355 1.834269 1.814647 1.811901 1.855129 1.865905 1.810784
##        9       10       11       12
## 1.809450 1.767552 1.801533 1.847200
```

Mean absolute error: 2.4856 percent

Analysis

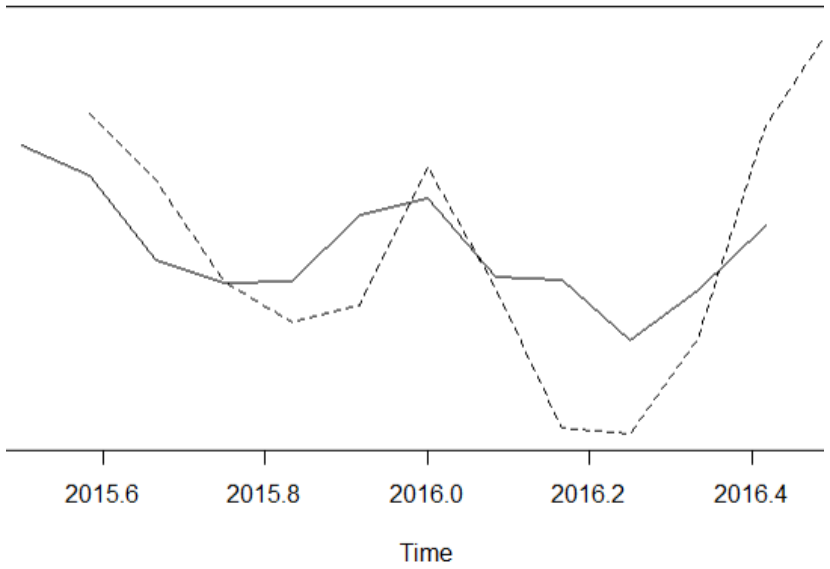
- Predicting The Next 12 Months
- Naive Forecasting
- Simple Exponential Smoothing
- Holt’s Trend
- ARIMA
- TBATS

Predicting The Next Twelve Months



Prediction vs Test

- Means Absolute Error= 2.485639%



Naive Forecasting

- MAPE= 6.28

```
Forecast method: Naive method

Model Information:
Call: naive(y = carbonEmission_train, h = 12)

Residual sd: 0.039

Error measures:
                      ME       RMSE        MAE          MPE      MA
Training set 0.0002587966 0.03901394 0.03258609 -0.009051039 1.7816
                    ACF1
Training set 0.1594616

Forecasts:
         Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
Aug 2015       1.931701 1.881703 1.981700 1.855235 2.008167
Sep 2015       1.931701 1.860993 2.002410 1.823562 2.039841
Oct 2015       1.931701 1.845102 2.018301 1.799259 2.064144
Nov 2015       1.931701 1.831705 2.031698 1.778770 2.084633
Dec 2015       1.931701 1.819902 2.043501 1.760718 2.102684
Jan 2016       1.931701 1.809231 2.054172 1.744399 2.119004
Feb 2016       1.931701 1.799418 2.063985 1.729392 2.134011
Mar 2016       1.931701 1.790285 2.073118 1.715423 2.147980
Apr 2016       1.931701 1.781706 2.081697 1.702304 2.161099
May 2016       1.931701 1.773593 2.089810 1.689895 2.173508
Jun 2016       1.931701 1.765876 2.097527 1.678093 2.185310
Jul 2016       1.931701 1.758502 2.104901 1.666816 2.196587
```

Simple Exponential Smoothing

- MAPE= 4.12

```
## Simple exponential smoothing
##
## Call:
##  ses(y = carbonEmission_train, h = 12)
##
##   Smoothing parameters:
##     alpha = 1e-04
##
##   Initial states:
##     l = 1.8325
##
##   sigma:  0.0385
##
##       AIC      AICc       BIC
## -237.1452 -237.0660 -225.9647
##
## Error measures:
##                         ME       RMSE        MAE         MPE     MAPE     MASE
## Training set 2.783062e-06 0.03841044 0.03155226 -0.04378553 1.721291 2.522318
##                   ACF1
## Training set 0.474498
##
## Forecasts:
##          Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## Aug 2015       1.832539 1.783153 1.881925 1.757009 1.908068
## Sep 2015       1.832539 1.783153 1.881925 1.757009 1.908068
## Oct 2015       1.832539 1.783153 1.881925 1.757009 1.908068
## Nov 2015       1.832539 1.783153 1.881925 1.757009 1.908068
## Dec 2015       1.832539 1.783153 1.881925 1.757009 1.908068
## Jan 2016       1.832539 1.783153 1.881925 1.757009 1.908068
## Feb 2016       1.832539 1.783153 1.881925 1.757009 1.908068
## Mar 2016       1.832539 1.783153 1.881925 1.757009 1.908068
## Apr 2016       1.832539 1.783153 1.881925 1.757009 1.908068
## May 2016       1.832539 1.783153 1.881925 1.757009 1.908068
## Jun 2016       1.832539 1.783153 1.881925 1.757009 1.908068
## Jul 2016       1.832539 1.783153 1.881925 1.757009 1.908068
```

Holt’s Trend Method

- MAPE= 4.16

```
##
## Model Information:
## Holt's method
##
## Call:
##  holt(y = carbonEmission_train, h = 12)
##
##   Smoothing parameters:
##     alpha = 0.0115
##     beta  = 0.0115
##
##   Initial states:
##     l = 1.83
##     b = 3e-04
##
##   sigma:  0.0397
##
##       AIC      AICc       BIC
## -216.5203 -216.3210 -197.8861
##
## Error measures:
##                          ME       RMSE        MAE         MPE     MAPE     MASE
## Training set -1.184666e-05 0.03946467 0.03234292 -0.04574211 1.764425 2.585525
##                    ACF1
## Training set 0.4854102
##
## Forecasts:
##          Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## Aug 2015       1.834710 1.783801 1.885618 1.756851 1.912568
## Sep 2015       1.835003 1.784081 1.885926 1.757125 1.912882
## Oct 2015       1.835297 1.784345 1.886250 1.757372 1.913222
## Nov 2015       1.835591 1.784585 1.886598 1.757584 1.913599
## Dec 2015       1.835885 1.784795 1.886975 1.757749 1.914021
## Jan 2016       1.836179 1.784968 1.887390 1.757858 1.914500
## Feb 2016       1.836473 1.785098 1.887848 1.757901 1.915044
## Mar 2016       1.836767 1.785178 1.888355 1.757869 1.915664
## Apr 2016       1.837060 1.785203 1.888918 1.757752 1.916369
## May 2016       1.837354 1.785167 1.889541 1.757541 1.917167
```

ARIMA

- MAPE= 2.31

```
## Series: carbonEmission_train
## ARIMA(0,0,1)(1,1,2)[12]
##
## Coefficients:
##          ma1     sar1     sma1     sma2
##       0.4248  -0.3761  -0.3518  -0.3371
## s.e.  0.0523   0.4080   0.3995   0.3026
##
## sigma^2 estimated as 0.0001537:  log likelihood=873.29
## AIC=-1736.58   AICc=-1736.37   BIC=-1718.14
##
## Training set error measures:
##                        ME       RMSE         MAE          MPE      MAPE
## Training set 0.0002402486 0.01206913 0.009260892 0.007434987 0.5053679
##                    MASE         ACF1
## Training set 0.7403246 -0.000975859
```

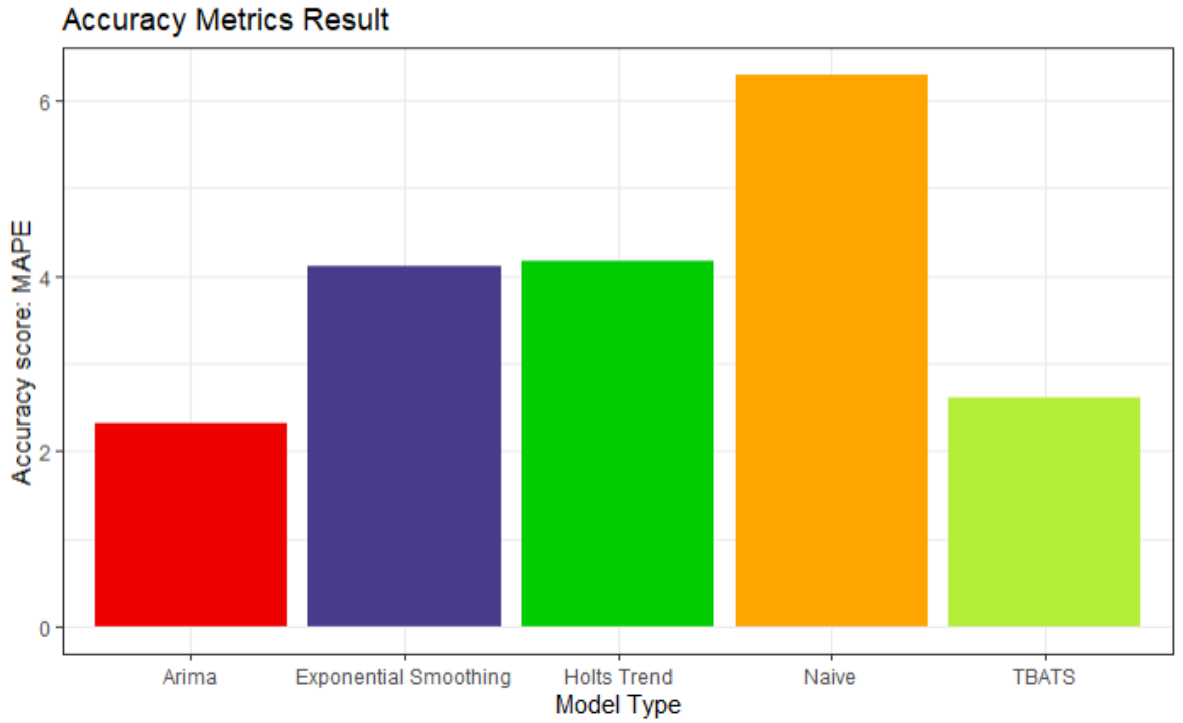
TBATS

- MAPE=2.61

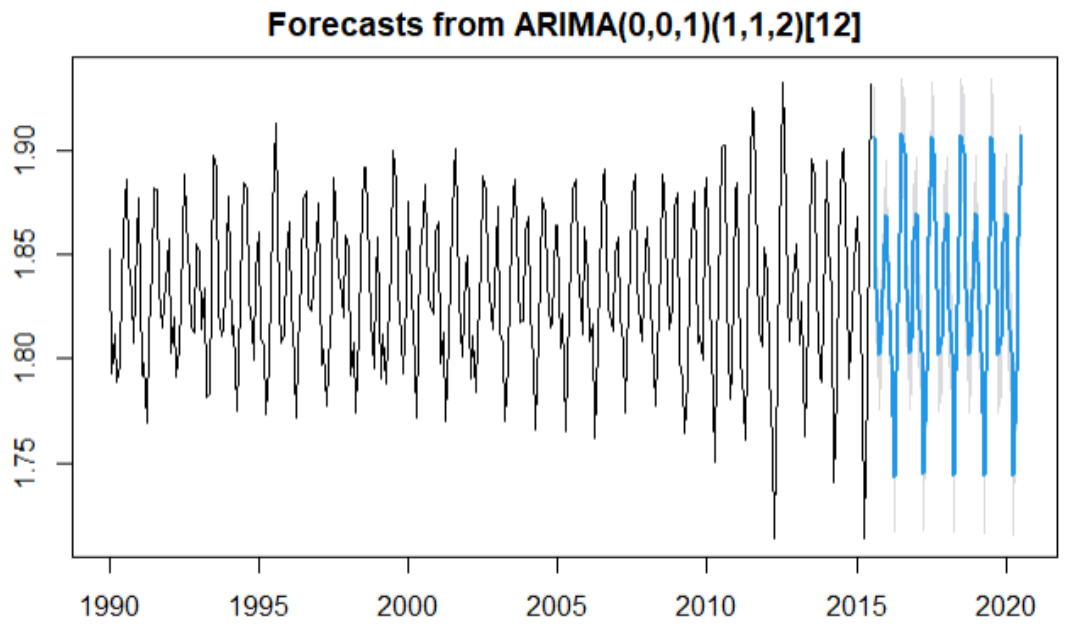
```
##          Point Forecast    Lo 80    Hi 80    Lo 95    Hi 95
## Aug 2015       1.898832 1.883603 1.914061 1.875541 1.922123
## Sep 2015       1.822015 1.806299 1.837731 1.797980 1.846050
## Oct 2015       1.805088 1.789073 1.821102 1.780596 1.829579
## Nov 2015       1.802837 1.785853 1.819821 1.776862 1.828812
## Dec 2015       1.854034 1.836801 1.871267 1.827679 1.880389
## Jan 2016       1.861378 1.844080 1.878675 1.834923 1.887832
## Feb 2016       1.811907 1.794593 1.829221 1.785428 1.838386
## Mar 2016       1.806102 1.788785 1.823420 1.779618 1.832587
## Apr 2016       1.769276 1.751958 1.786594 1.742791 1.795762
## May 2016       1.798493 1.781175 1.815811 1.772007 1.824979
## Jun 2016       1.849080 1.831762 1.866398 1.822594 1.875565
## Jul 2016       1.888402 1.871084 1.905721 1.861917 1.914888
```

RESULTS

- Naive =6.2849
- Exponential Smoothing =4.1167
- Holts Trend= 4.1617
- Arima = 2.3084 ***
- Tbats = 2.6093



Forecasting Next Four Years



Conclusion

- No Correlation was found between the Source type of Carbon Emission and the Carbon Emission Value
- Coal is the main source of Carbon Emission followed by Natural Gas
- All Source Type except Natural Gas has a downtrend starting 2010
- Overall Emission by Electricity Generation has a downtrend = Lower Carbon Emission
- Linear Regression model had a MAPE = 2.49%
- Best Model to Predict Carbon Emission: ARIMA =2.31%
- Percent Change in Carbon Emission between 07/2016 to 07/2020= 5.70%
- After Forecasting, Carbon Emission is predicted to increase in the next five years