ISyE 4031 T09 - Georgia Achievement Gaps in K-12 Schools with Regression

Ruwei Ma, Yichen Ma, Yang Yang

1. Introduction

Covid-19 had brought a big impact to the education system across US. National test results for 2022 reveal the pandemic's devastating effects on American schoolchildren, with the performance of 9-year-olds in math and reading dropping to the lowest levels from two decades ago [1]. This lagging effect from the pandemic applies to all races and income levels and sparks a collective decline in academics for the generation that experienced school closures, frequent reliance on virtual and remote learning, and other pandemic effects. The setbacks will occupy the low-performing students for up to 9 months to catch up with the average, prompting an urgent need for the underlying solution to the achievement gap [2]. This setback further adds to, and likely aggravates, the pre-pandemic disparity in student achievement outcomes for vulnerable and at-risk student populations, especially in Georgia. Based on some of my preliminary analysis of the 2021 achievement data across 2,180 schools in Georgia, we found that there are 2 prominent factors that affect achievement rate: the student's economic status and race. The achievement rate in 2021 of economically disadvantaged students is 46.11%, compared to 52.32% across all students. A similar gap can be observed in the difference in achievement rate between white and black students in Georgia, the former as high as 66.99%, compared to the 39.88% of the latter. The gap within the economically-disadvantaged students' group is vast and depends on the county or school they attend. Further analysis at the school level shows strong correlation between achievement rate and the school's other demographics.

2. Problem Goal

We aim to adopt regression modeling to identify gaps in national test achievement rates between different demographic groups in Georgia, and recommend robust strategies to address such disparities. Specifically, the objectives are: (1) visualize the disparities in school resources, such as teacher certifications and FTE (Full-time Equivalent), and quantify its correlation with the student's achievement outcomes, especially among marginalized minority groups (e.g., White, Black, vs. Hispanic students, economically disadvantaged vs. affluent students, and rural vs. Urban schools) (2) quantify the achievement gap at the county level across Georgia's 159 counties at the school level to identify factors that predict student achievement and highlight intervention or resource allocation strategies, and (3) evaluate the impact and predict the trajectory of the policies and strategies produced from step 2 with adjustments.

3. Executive Summary

4. Data Description

```r
# Input Dataset
library(readxl)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

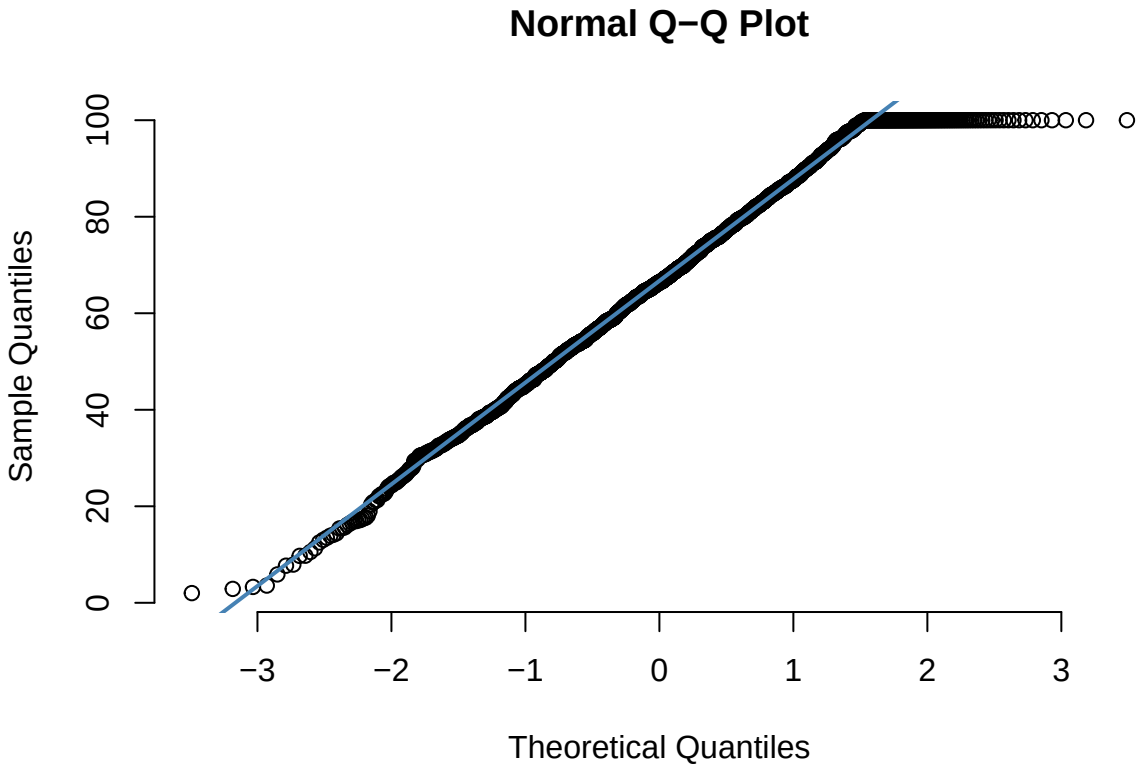
```r
Achievement_Rate = read.csv("2019 & 2021 Content Mastery Data.csv", header=TRUE)
Percentage = read.csv("Percentages & Certificates.csv", header=TRUE)
Salaries = read.csv("salaries.csv", header=TRUE)
Absent_Rate = read.csv("Absent Rate.csv", header=TRUE)
School_Expenditure = read_excel("2021_School-Level_PPE.xls")
```

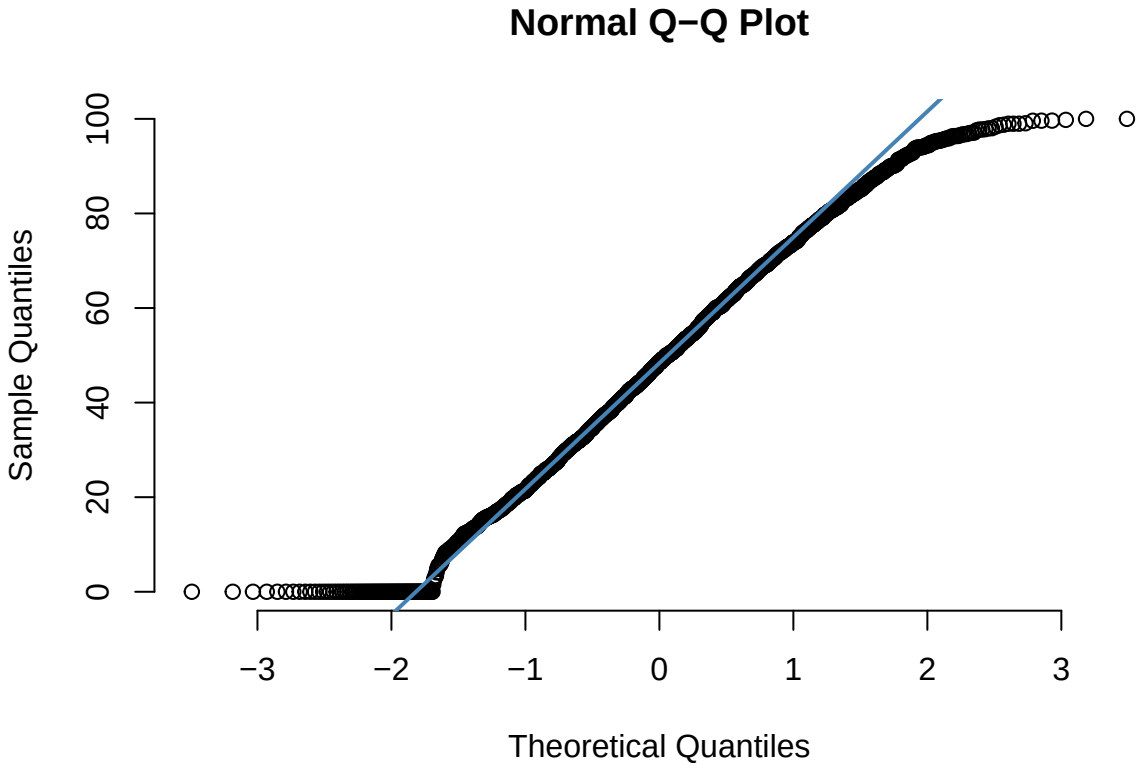
```
## Warning: Expecting numeric in Y2255 / R2255C25: got 'Non-Compliant'

## Warning: Expecting numeric in Z2255 / R2255C26: got 'Non-Compliant'

## Warning: Expecting numeric in AA2255 / R2255C27: got 'Non-Compliant'

## Warning: Expecting logical in AB2255 / R2255C28: got 'Note: This school did not
## report financial data for FY21.'

## New names:
## * `` -> `...28`
```

```r
School_Expenditure = select(School_Expenditure, schoolname, amount, school_ppe_21)
Poverty.Percentage = read_excel("2021_directly_certified_school.xls")
Poverty.Percentage = select(Poverty.Percentage, SCHOOL_NAME, direct_cert_perc)
Mobility = read_excel("2021_School_Mobility.xls")
Mobility = select(Mobility, school_name, mobility)
Enrollment = read.csv("Enrollment_by_Subgroups_Programs.csv", header=TRUE)
Enrollment = select(Enrollment, INSTN_NAME, ENROLL_PCT_GIFTED)

data = merge(Achievement_Rate, Percentage, by="School.Name")
data = merge(data, Salaries, by="School.Name")
data = merge(data, Absent_Rate, by="School.Name")

data = left_join(
          data %>% group_by(School.Name) %>% mutate(id = row_number()),
          School_Expenditure %>% group_by(schoolname) %>% mutate(id = row_number()),
          by = c("School.Name" = "schoolname", "id"))
```

```r
data = left_join(
          data %>% group_by(School.Name) %>% mutate(id = row_number()),
          Poverty.Percentage %>% group_by(SCHOOL_NAME) %>% mutate(id = row_number()),
          by = c("School.Name" = "SCHOOL_NAME", "id"))

data = left_join(
          data %>% group_by(School.Name) %>% mutate(id = row_number()),
          Mobility %>% group_by(school_name) %>% mutate(id = row_number()),
          by = c("School.Name" = "school_name", "id"))

data = left_join(
          data %>% group_by(School.Name) %>% mutate(id = row_number()),
          Enrollment %>% group_by(INSTN_NAME) %>% mutate(id = row_number()),
          by = c("School.Name" = "INSTN_NAME", "id"))
attach(data)

# Creating a Dummy Variable for Urban/Rural
data$u.r_dummy <- data$Urban.Rural
data$u.r_dummy <- as.character(data$u.r_dummy)
data$u.r_dummy[data$u.r_dummy == "Urban"] <- 1
data$u.r_dummy[data$u.r_dummy == "Rural"] <- 0
data$u.r_dummy <- as.numeric(data$u.r_dummy)
data$growth.rate.math <-data$X19.21.Difference.in.Math
```

a. Data Summary





```
##                2.5 %  97.5 %
## (Intercept) 65.23627 66.9811
```

```
##                2.5 %   97.5 %
## (Intercept) 46.99179 49.09119
```

We are 95% confident that the mean student achivement rate in Math in 2021 is higher than that in 2019.

```
##
## Attaching package: 'huxtable'
```

```
## The following object is masked from 'package:dplyr':
##
##     add_rownames
```

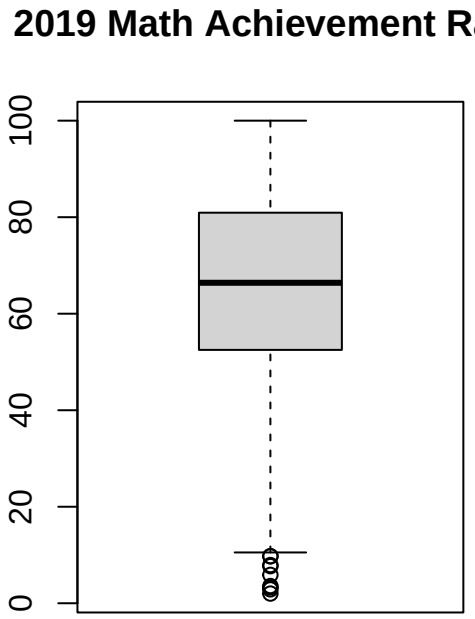
	2019	**2021**
Observations	2067.00	2067.00
Avg. Math achievement	66.1086840832124	48.0414900822448
Median Math achievement	66.41	48.47
Lower Bound of Math achievement	2.01	0
Upper Bound of Math achievement	100	100
Standard Deviation	20.2251098070928	24.3350230861084

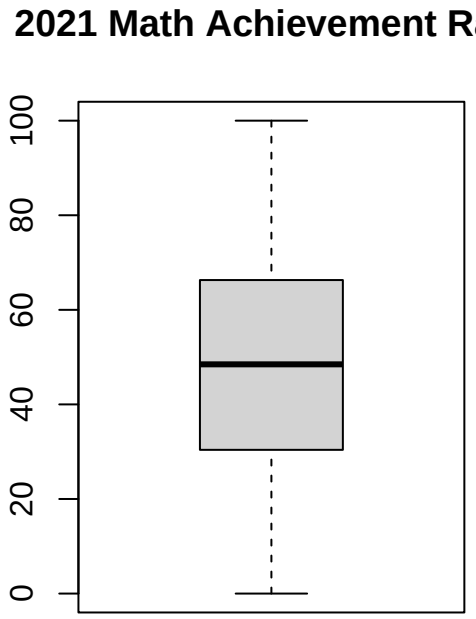
Mean and median Math test achievement rates are higher in 2019 than in 2021.

```
#average change in achievement rate
(52.23121-67.99686)/67.99686
```

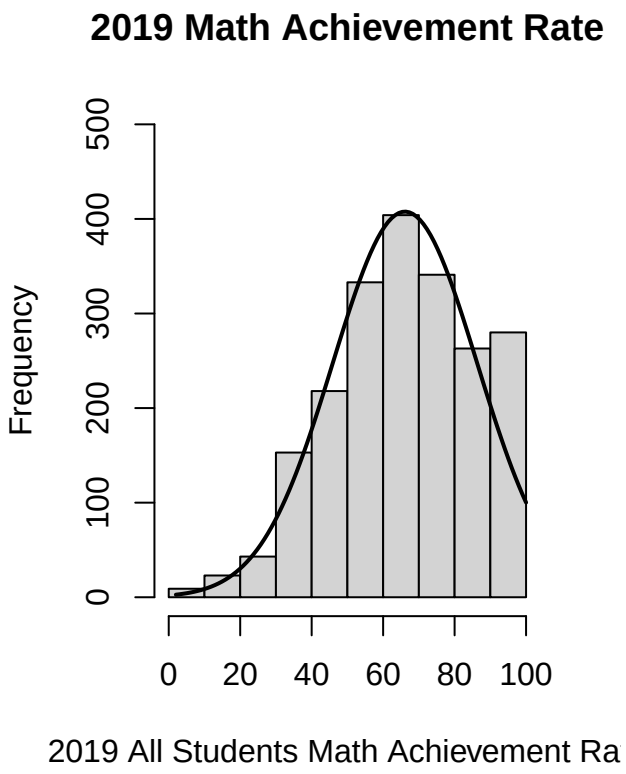
```
## [1] -0.2318585
```

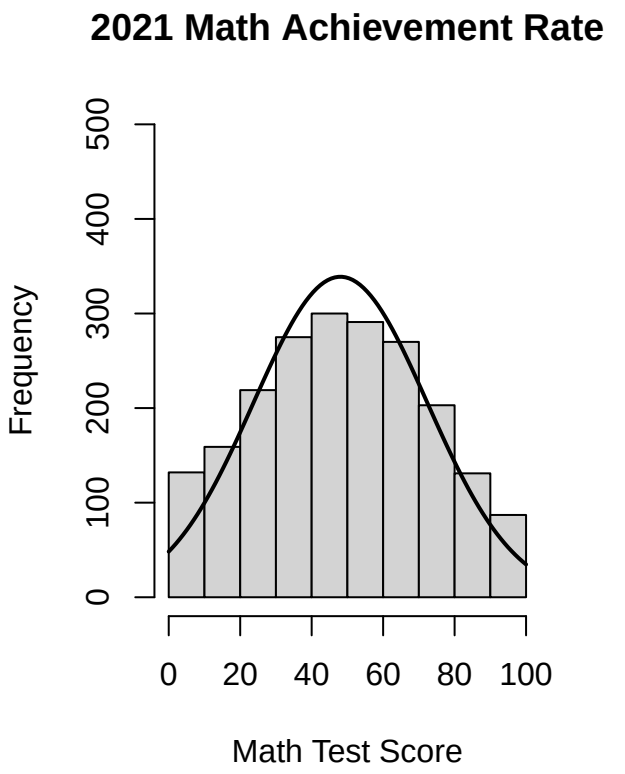
c. Data Visulization





The boxplot of both years' math achievement rate shows that in 2019, the data distribution is more compact, and all quartiles are significantly higher than those in 2021. A tremendous number of outliers are identified in both year's boxplots, suggesting many data points below the lower quartile by more than 1.5 interquartile range (IQR). Achievement rates are highly left skewed.





From both years' histogram, it can be confirmed that there is a very low frequency of math achievement rate between 0-30 for the 2019 data, as compared to the 2021 data. More outliers in the 2019 data could mean a higher . From plain sight, the 2019 data is better approximated by a normal distribution. The 2021 data seems skewed to the center.

d. Table of Variables

Variables	Description	Type
y1	2019 All Students Math Achievement Rate	Quantative
y2	2021 All Students Math Achievement Rate	Quantative
x1	Absent 0-5 Days Percentage	Quantative
x2	Absent 6-15 Days Percentage	Quantative
x3	Absent 15+ Days Percentage	Quantative
x4	Avg. Annual Salaries - Administrators	Quantative
x5	Avg. Annual Salaries - Teachers	Quantative
x6	Avg. Annual Salaries - Support.Personnel	Quantative
x7	Number of Teachers with a phd degree	Quantative
x8	Total Number of Certified Teachers	Quantative
x9	Post Grad Percentage	Quantative
x10	Total Students Enrolled	Quantative
x11	Teacher-Student Ratio	Quantative
x12	White Student Percentage	Quantative
x13	Black Student Percentage	Quantative
x14	Economically Disadvantaged Student Percentage	Quantative
x15	Directly Certified Students Percentage	Quantative
x16	Amount of Money Invested for Students	Quantative
x17	Per-Pupil Expenditure at School Level	Quantative
x18	Rate of Entries and Withdrawls to a School	Quantative
x19	Percentage of Gifted Students	Quantative
x20	Urban/Rural Area of the School	Qualitative

5. Regression Analysis

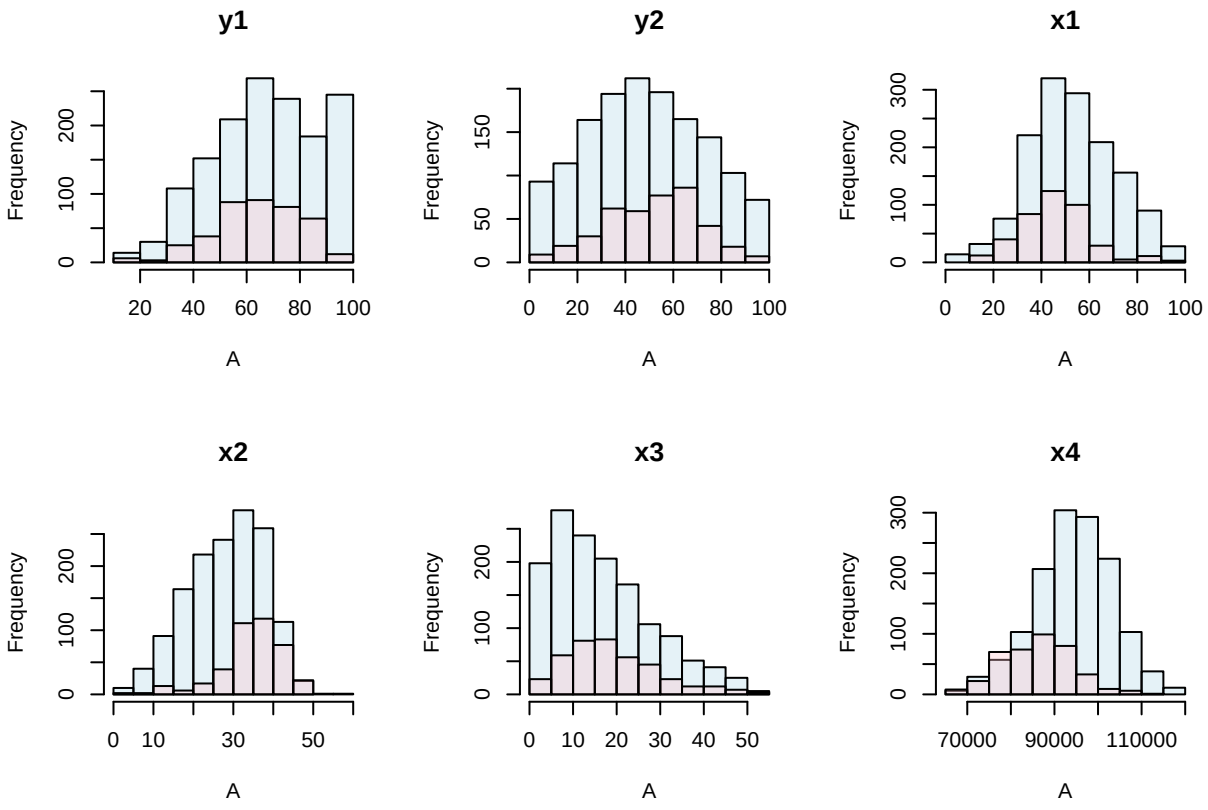
a. Iterations of the analysis process

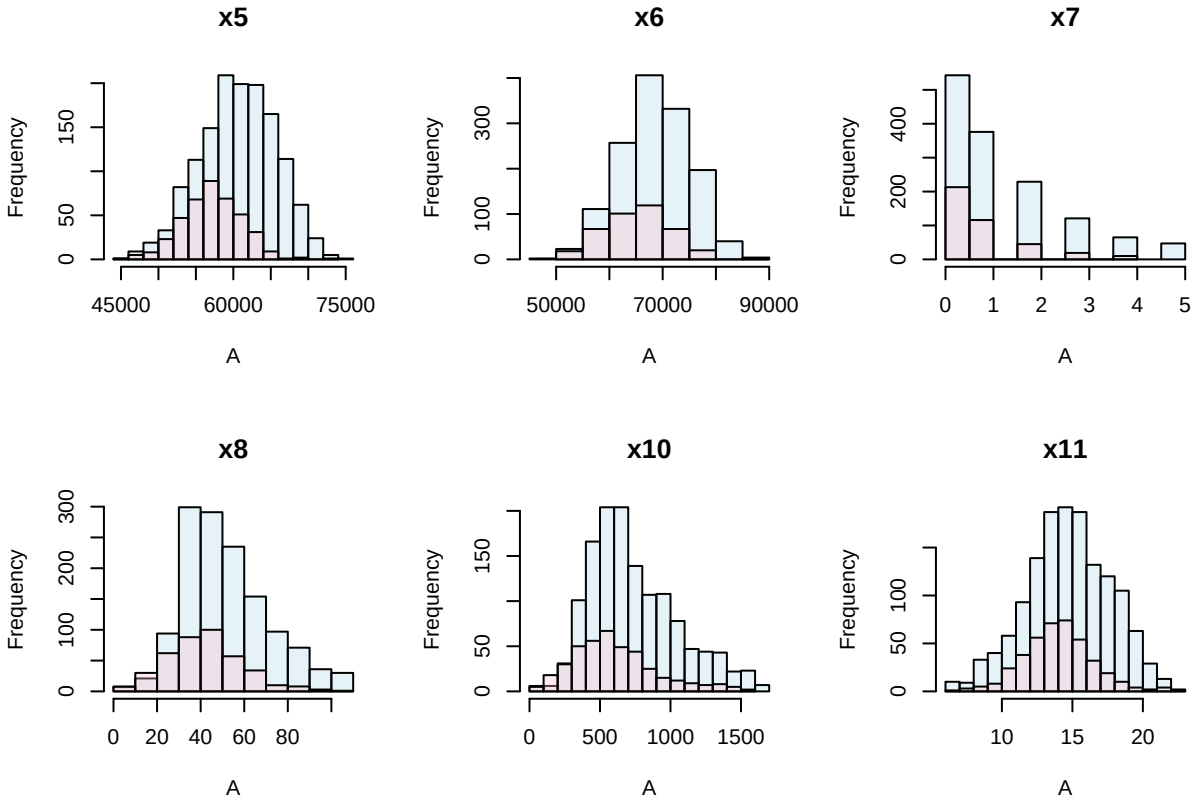
- paragraph description

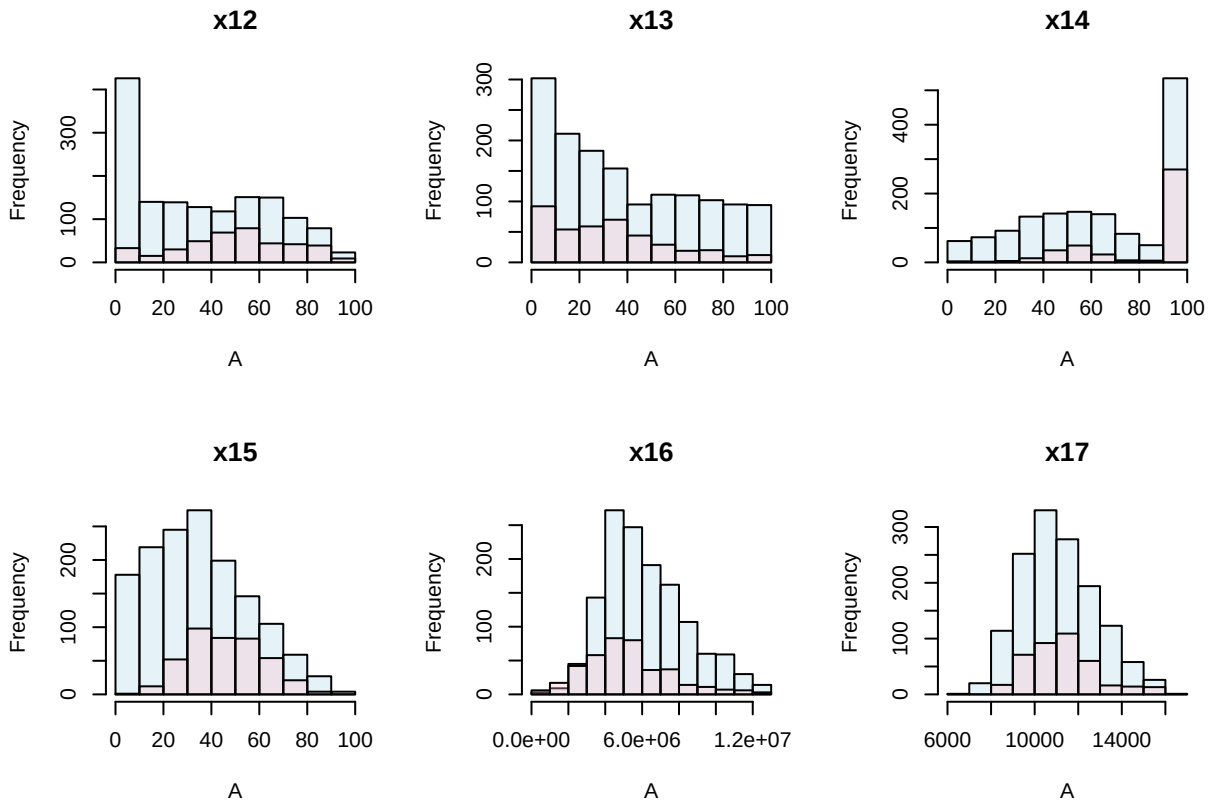
c. Plots of variables- Scatterplot

For the plots below, a light blue color indicates Urban Area and a light pink color indicates Rural Area.

```
## Warning in hist.default(A, breaks = 12, plot = FALSE, na.rm = TRUE): argument
## '...' is not made use of

## Warning in hist.default(B, breaks = 12, plot = FALSE, na.rm = TRUE): argument
## '...' is not made use of

## Warning in hist.default(A, breaks = 12, plot = FALSE, na.rm = TRUE): argument
## '...' is not made use of

## Warning in hist.default(B, breaks = 12, plot = FALSE, na.rm = TRUE): argument
## '...' is not made use of

## Warning in hist.default(A, breaks = 12, plot = FALSE, na.rm = TRUE): argument
## '...' is not made use of

## Warning in hist.default(B, breaks = 12, plot = FALSE, na.rm = TRUE): argument
## '...' is not made use of

## Warning in hist.default(A, breaks = 12, plot = FALSE, na.rm = TRUE): argument
## '...' is not made use of

## Warning in hist.default(B, breaks = 12, plot = FALSE, na.rm = TRUE): argument
## '...' is not made use of

## Warning in hist.default(A, breaks = 12, plot = FALSE, na.rm = TRUE): argument
## '...' is not made use of

## Warning in hist.default(B, breaks = 12, plot = FALSE, na.rm = TRUE): argument
## '...' is not made use of

## Warning in hist.default(A, breaks = 12, plot = FALSE, na.rm = TRUE): argument
## '...' is not made use of

## Warning in hist.default(B, breaks = 12, plot = FALSE, na.rm = TRUE): argument
## '...' is not made use of
```

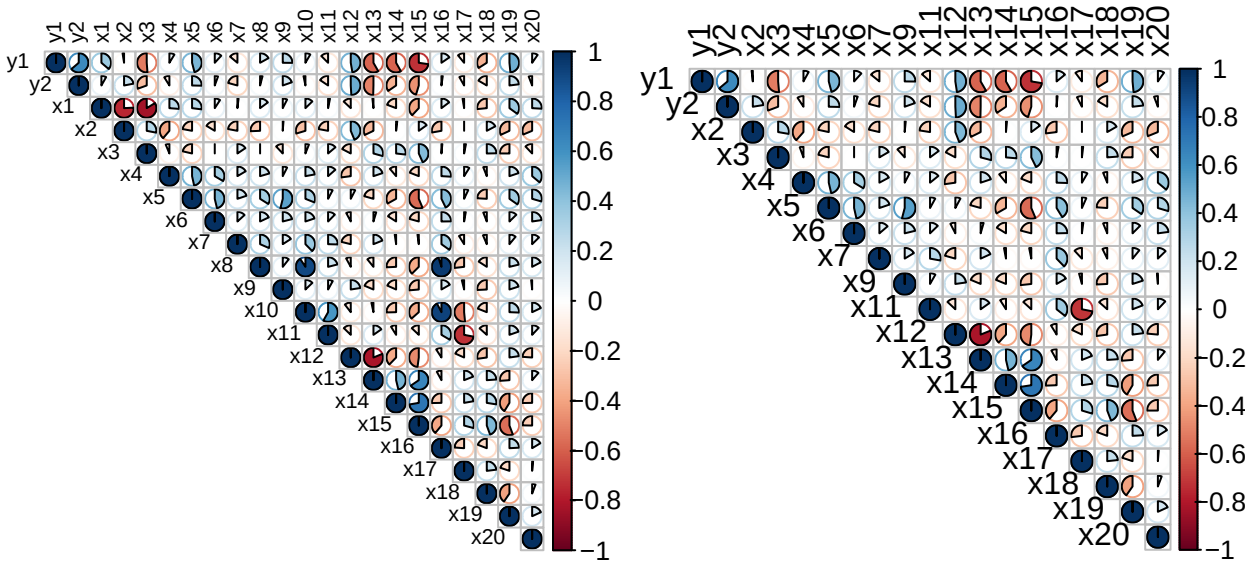

y1
Frequency
0
100
200
20
40
60
80
100
A
y2
Frequency
0
50
150
0
20
40
60
80
100
A
x1
Frequency
0
100
200
300
0
20
40
60
80
100
A
x2
Frequency
0
100
200
0
10
30
50
A
x3
Frequency
0
100
200
0
10
20
30
40
50
A
x4
Frequency
0
100
200
300
70000
90000
110000
A


x5
Frequency
0
50
150
45000
60000
75000
A
x6
Frequency
0
100
300
50000
70000
90000
A
x7
Frequency
0
200
400
0
1
2
3
4
5
A
x8
Frequency
0
100
200
300
0
20
40
60
80
A
x10
Frequency
0
50
150
0
500
1000
1500
A
x11
Frequency
0
50
100
10
15
20
A



b. Multicollinearity

```
## corrplot 0.92 loaded
```



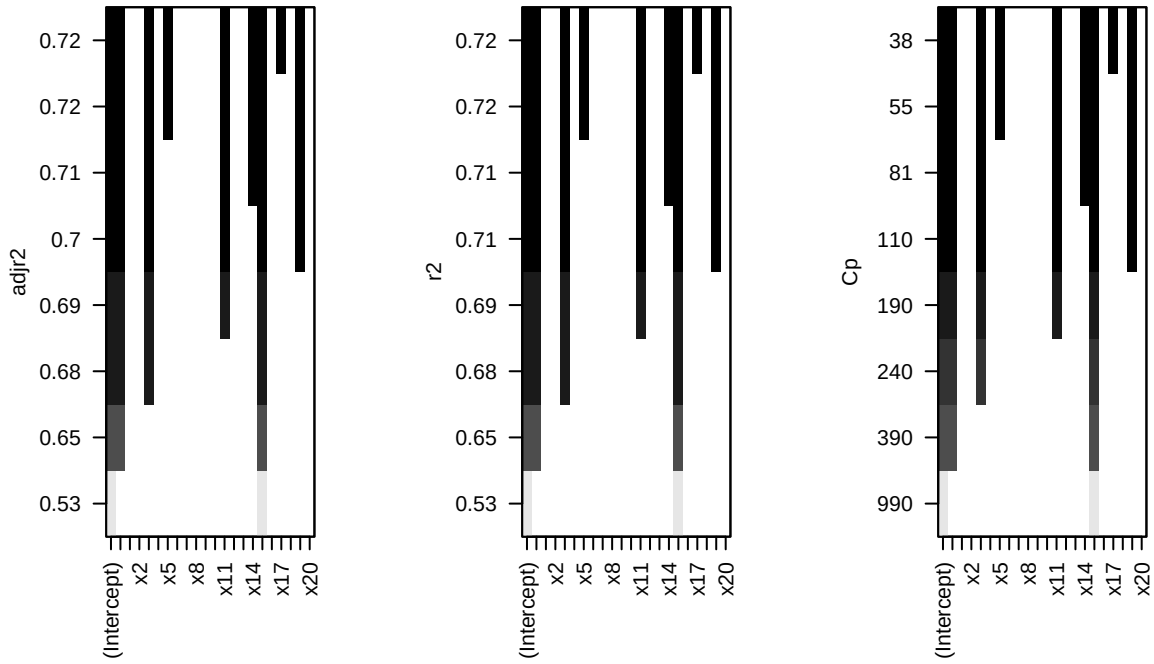

```
## Loading required package: lattice

## Loading required package: survival

## Loading required package: Formula

## Loading required package: ggplot2

##
## Attaching package: 'ggplot2'

## The following object is masked from 'package:huxtable':
##
##     theme_grey

##
## Attaching package: 'Hmisc'

## The following objects are masked from 'package:huxtable':
##
##     contents, label, label<-

## The following objects are masked from 'package:dplyr':
##
##     src, summarize
```

```
## The following objects are masked from 'package:base':
##
##     format.pval, units
```

Before doing the model selection process, a Multicollinearity check produces high correlation of (x1:x3), (x8:x10,x16), (x10:x16), (x12: x13), (x11: x17), (x15: x13, x14). And another set of variables that have a high correlation is y1 and y2, since we are modeling them separately as response variables, we do not need to drop any of them. The renewed plot is on the right.

d. Model Selection

2019 Model Selection




```
##   (Intercept)    y2    x1    x2    x3    x4    x5    x6    x7    x8    x9   x10
## 1        TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## 2        TRUE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## 3        TRUE  TRUE FALSE FALSE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## 4        TRUE  TRUE FALSE FALSE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## 5        TRUE  TRUE FALSE FALSE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## 6        TRUE  TRUE FALSE FALSE  TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
## 7        TRUE  TRUE FALSE FALSE  TRUE FALSE  TRUE FALSE FALSE FALSE FALSE FALSE
## 8        TRUE  TRUE FALSE FALSE  TRUE FALSE  TRUE FALSE FALSE FALSE FALSE FALSE
##     x11   x12   x13   x14  x15   x16   x17   x18   x19   x20
## 1 FALSE FALSE FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE
## 2 FALSE FALSE FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE
```

```
## 3 FALSE FALSE FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE
## 4  TRUE FALSE FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE
## 5  TRUE FALSE FALSE FALSE TRUE FALSE FALSE FALSE  TRUE FALSE
## 6  TRUE FALSE FALSE  TRUE TRUE FALSE FALSE FALSE  TRUE FALSE
## 7  TRUE FALSE FALSE  TRUE TRUE FALSE FALSE FALSE  TRUE FALSE
## 8  TRUE FALSE FALSE  TRUE TRUE FALSE  TRUE FALSE  TRUE FALSE

##
## Attaching package: 'olsrr'

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

## Stepwise Selection Method
## --------------------------
##
## Candidate Terms:
##
## 1. y2
## 2. x1
## 3. x2
## 4. x3
## 5. x4
## 6. x5
## 7. x6
## 8. x7
## 9. x8
## 10. x9
## 11. x10
## 12. x11
## 13. x12
## 14. x13
## 15. x14
## 16. x15
## 17. x16
## 18. x17
## 19. x18
## 20. x19
## 21. x20
##
## We are selecting variables based on p value...
##
##
## Stepwise Selection: Step 1
##
## - x15 added
##
##                         Model Summary
## ----------------------------------------------------------------
## R                       0.702       RMSE                14.159
## R-Squared               0.492       Coef. Var           21.333
## Adj. R-Squared          0.492       MSE                200.489
```

```
## Pred R-Squared          0.492       MAE                  10.974
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F          Sig.
## --------------------------------------------------------------------------
## Regression    400015.583           1     400015.583    1995.198    0.0000
## Residual      412205.640        2056        200.489
## Total         812221.223        2057
## --------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------------
## (Intercept)    91.915         0.651                141.087    0.000    90.637    93.193
##         x15    -0.671         0.015       -0.702    -44.668    0.000    -0.701    -0.642
## ----------------------------------------------------------------------------------------------
##
##
##
## Stepwise Selection: Step 2
##
## - y2 added
##
##                           Model Summary
## ----------------------------------------------------------------
## R                       0.804       RMSE                11.827
## R-Squared               0.646       Coef. Var           17.819
## Adj. R-Squared          0.646       MSE                139.870
## Pred R-Squared          0.644       MAE                  8.564
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F          Sig.
## --------------------------------------------------------------------------
## Regression    524787.972           2     262393.986    1875.982    0.0000
## Residual      287433.251        2055        139.870
## Total         812221.223        2057
## --------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
```

```
## ---------------------------------------------------------------------------------------------
## (Intercept)    66.569         1.008                66.033    0.000    64.592    68.546
##         x15    -0.468         0.014       -0.490   -32.832    0.000    -0.496    -0.440
##          y2     0.366         0.012        0.446    29.867    0.000     0.342     0.390
## ---------------------------------------------------------------------------------------------
##
##
##
##                           Model Summary
## ------------------------------------------------------------------
## R                       0.804       RMSE                 11.827
## R-Squared               0.646       Coef. Var            17.819
## Adj. R-Squared          0.646       MSE                 139.870
## Pred R-Squared          0.644       MAE                   8.564
## ------------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                  ANOVA
## --------------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F          Sig.
## --------------------------------------------------------------------------------
## Regression    524787.972           2     262393.986    1875.982    0.0000
## Residual      287433.251        2055        139.870
## Total         812221.223        2057
## --------------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ---------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ---------------------------------------------------------------------------------------------
## (Intercept)    66.569         1.008                66.033    0.000    64.592    68.546
##         x15    -0.468         0.014       -0.490   -32.832    0.000    -0.496    -0.440
##          y2     0.366         0.012        0.446    29.867    0.000     0.342     0.390
## ---------------------------------------------------------------------------------------------
##
##
##
## Stepwise Selection: Step 3
##
## - x3 added
##
##                           Model Summary
## ------------------------------------------------------------------
## R                       0.812       RMSE                 11.294
## R-Squared               0.659       Coef. Var            16.751
## Adj. R-Squared          0.658       MSE                 127.548
## Pred R-Squared          0.656       MAE                   8.318
## ------------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
```

```
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F          Sig.
## --------------------------------------------------------------------------
## Regression    484915.179           3     161638.393    1267.273    0.0000
## Residual      251269.917        1970        127.548
## Total         736185.095        1973
## --------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------------
## (Intercept)    71.865         1.056                 68.045    0.000    69.794    73.936
##         x15    -0.424         0.015       -0.441    -28.509    0.000    -0.454    -0.395
##          y2     0.331         0.012        0.409     27.184    0.000     0.307     0.354
##          x3    -0.286         0.024       -0.170    -11.683    0.000    -0.334    -0.238
## ----------------------------------------------------------------------------------------------
##
##
##
##                           Model Summary
## ----------------------------------------------------------------
## R                       0.812       RMSE                11.294
## R-Squared               0.659       Coef. Var           16.751
## Adj. R-Squared          0.658       MSE                127.548
## Pred R-Squared          0.656       MAE                  8.318
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F          Sig.
## --------------------------------------------------------------------------
## Regression    484915.179           3     161638.393    1267.273    0.0000
## Residual      251269.917        1970        127.548
## Total         736185.095        1973
## --------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------------
## (Intercept)    71.865         1.056                 68.045    0.000    69.794    73.936
##         x15    -0.424         0.015       -0.441    -28.509    0.000    -0.454    -0.395
##          y2     0.331         0.012        0.409     27.184    0.000     0.307     0.354
##          x3    -0.286         0.024       -0.170    -11.683    0.000    -0.334    -0.238
## ----------------------------------------------------------------------------------------------
##
```

```
##
##
## Stepwise Selection: Step 4
##
## - x19 added
##
##                          Model Summary
## ----------------------------------------------------------------
## R                       0.819       RMSE                10.717
## R-Squared               0.671       Coef. Var           15.669
## Adj. R-Squared          0.671       MSE                114.860
## Pred R-Squared          0.669       MAE                  7.989
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F          Sig.
## --------------------------------------------------------------------------
## Regression    442933.711           4     110733.428    964.074    0.0000
## Residual      216740.546        1887        114.860
## Total         659674.257        1891
## --------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------------------
## (Intercept)    67.335         1.247                 53.985    0.000    64.889    69.782
##         x15    -0.357         0.017       -0.373    -20.736    0.000    -0.390    -0.323
##          y2     0.313         0.012        0.393     26.410    0.000     0.289     0.336
##          x3    -0.295         0.024       -0.180    -12.303    0.000    -0.342    -0.248
##         x19     0.316         0.035        0.144      8.939    0.000     0.247     0.386
## ----------------------------------------------------------------------------------------------------
##
##
##
##                          Model Summary
## ----------------------------------------------------------------
## R                       0.819       RMSE                10.717
## R-Squared               0.671       Coef. Var           15.669
## Adj. R-Squared          0.671       MSE                114.860
## Pred R-Squared          0.669       MAE                  7.989
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
```

```
##                   Squares          DF    Mean Square       F         Sig.
## ------------------------------------------------------------------------
## Regression    442933.711           4     110733.428    964.074    0.0000
## Residual      216740.546        1887        114.860
## Total         659674.257        1891
## ------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------
## (Intercept)    67.335         1.247                53.985    0.000    64.889    69.782
##         x15    -0.357         0.017       -0.373   -20.736    0.000    -0.390    -0.323
##          y2     0.313         0.012        0.393    26.410    0.000     0.289     0.336
##          x3    -0.295         0.024       -0.180   -12.303    0.000    -0.342    -0.248
##         x19     0.316         0.035        0.144     8.939    0.000     0.247     0.386
## ----------------------------------------------------------------------------------------
##
##
##
## Stepwise Selection: Step 5
##
## - x11 added
##
##                          Model Summary
## ----------------------------------------------------------------
## R                       0.829       RMSE                10.416
## R-Squared               0.688       Coef. Var           15.192
## Adj. R-Squared          0.687       MSE                108.497
## Pred R-Squared          0.684       MAE                  7.843
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                  ANOVA
## ------------------------------------------------------------------------
##                   Sum of
##                   Squares          DF    Mean Square       F         Sig.
## ------------------------------------------------------------------------
## Regression    440535.448           5      88107.090    812.068    0.0000
## Residual      200177.339        1845        108.497
## Total         640712.787        1850
## ------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------
## (Intercept)    81.760         1.897                43.108    0.000    78.040    85.479
##         x15    -0.405         0.017       -0.424   -23.175    0.000    -0.440    -0.371
##          y2     0.283         0.012        0.356    23.932    0.000     0.260     0.307
##          x3    -0.256         0.024       -0.157   -10.649    0.000    -0.303    -0.209
##         x19     0.343         0.035        0.158     9.840    0.000     0.275     0.412
```

```
##         x11    -0.830         0.091       -0.127     -9.153    0.000    -1.007    -0.652
## ----------------------------------------------------------------------------------------
##
##
##
##                           Model Summary
## ----------------------------------------------------------------
## R                        0.829       RMSE                10.416
## R-Squared                0.688       Coef. Var           15.192
## Adj. R-Squared           0.687       MSE                108.497
## Pred R-Squared           0.684       MAE                  7.843
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                  ANOVA
## ---------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F         Sig.
## ---------------------------------------------------------------------------
## Regression    440535.448           5      88107.090    812.068    0.0000
## Residual      200177.339        1845        108.497
## Total         640712.787        1850
## ---------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------
## (Intercept)    81.760         1.897                 43.108    0.000    78.040    85.479
##         x15    -0.405         0.017       -0.424    -23.175    0.000    -0.440    -0.371
##          y2     0.283         0.012        0.356     23.932    0.000     0.260     0.307
##          x3    -0.256         0.024       -0.157    -10.649    0.000    -0.303    -0.209
##         x19     0.343         0.035        0.158      9.840    0.000     0.275     0.412
##         x11    -0.830         0.091       -0.127     -9.153    0.000    -1.007    -0.652
## ----------------------------------------------------------------------------------------
##
##
##
## Stepwise Selection: Step 6
##
## - x14 added
##
##                           Model Summary
## ----------------------------------------------------------------
## R                        0.835       RMSE                10.247
## R-Squared                0.698       Coef. Var           14.945
## Adj. R-Squared           0.697       MSE                105.002
## Pred R-Squared           0.694       MAE                  7.774
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
```

```
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F         Sig.
## --------------------------------------------------------------------------
## Regression    447089.897           6      74514.983    709.656    0.0000
## Residual      193622.890        1844        105.002
## Total         640712.787        1850
## --------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------------
## (Intercept)    84.269         1.893                44.524    0.000    80.557    87.981
##         x15    -0.313         0.021       -0.328   -15.061    0.000    -0.354    -0.272
##          y2     0.279         0.012        0.350    23.867    0.000     0.256     0.301
##          x3    -0.264         0.024       -0.162   -11.155    0.000    -0.310    -0.218
##         x19     0.333         0.034        0.153     9.683    0.000     0.265     0.400
##         x11    -0.804         0.089       -0.123    -9.005    0.000    -0.979    -0.629
##         x14    -0.084         0.011       -0.142    -7.901    0.000    -0.105    -0.063
## ----------------------------------------------------------------------------------------------
##
##
##
##                           Model Summary
## ------------------------------------------------------------------
## R                       0.835       RMSE                10.247
## R-Squared               0.698       Coef. Var           14.945
## Adj. R-Squared          0.697       MSE                105.002
## Pred R-Squared          0.694       MAE                  7.774
## ------------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F         Sig.
## --------------------------------------------------------------------------
## Regression    447089.897           6      74514.983    709.656    0.0000
## Residual      193622.890        1844        105.002
## Total         640712.787        1850
## --------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------------
## (Intercept)    84.269         1.893                44.524    0.000    80.557    87.981
##         x15    -0.313         0.021       -0.328   -15.061    0.000    -0.354    -0.272
##          y2     0.279         0.012        0.350    23.867    0.000     0.256     0.301
```

```
##          x3    -0.264         0.024       -0.162    -11.155    0.000    -0.310    -0.218
##         x19     0.333         0.034        0.153      9.683    0.000     0.265     0.400
##         x11    -0.804         0.089       -0.123     -9.005    0.000    -0.979    -0.629
##         x14    -0.084         0.011       -0.142     -7.901    0.000    -0.105    -0.063
## ----------------------------------------------------------------------------------------
##
##
##
## Stepwise Selection: Step 7
##
## - x5 added
##
##                         Model Summary
## ----------------------------------------------------------------
## R                       0.845       RMSE                 9.952
## R-Squared               0.715       Coef. Var           14.507
## Adj. R-Squared          0.713       MSE                 99.043
## Pred R-Squared          0.711       MAE                  7.564
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F         Sig.
## --------------------------------------------------------------------------
## Regression    444444.576          7      63492.082    641.058    0.0000
## Residual      177583.362        1793         99.043
## Total         622027.938        1800
## --------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------
## (Intercept)    62.229         4.147                 15.006    0.000    54.096    70.363
##         x15    -0.290         0.024       -0.301    -12.310    0.000    -0.336    -0.244
##          y2     0.264         0.012        0.332     22.880    0.000     0.242     0.287
##          x3    -0.271         0.024       -0.166    -11.439    0.000    -0.317    -0.224
##         x19     0.310         0.034        0.143      9.188    0.000     0.244     0.377
##         x11    -0.952         0.093       -0.140    -10.234    0.000    -1.135    -0.770
##         x14    -0.081         0.011       -0.136     -7.416    0.000    -0.103    -0.060
##          x5     0.000         0.000        0.109      7.166    0.000     0.000     0.001
## ----------------------------------------------------------------------------------------
##
##
##
##                         Model Summary
## ----------------------------------------------------------------
## R                       0.845       RMSE                 9.952
## R-Squared               0.715       Coef. Var           14.507
## Adj. R-Squared          0.713       MSE                 99.043
```

```
## Pred R-Squared          0.711       MAE                  7.564
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F         Sig.
## --------------------------------------------------------------------------
## Regression    444444.576           7      63492.082    641.058    0.0000
## Residual      177583.362        1793         99.043
## Total         622027.938        1800
## --------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------------
## (Intercept)    62.229         4.147                  15.006    0.000    54.096    70.363
##         x15    -0.290         0.024       -0.301    -12.310    0.000    -0.336    -0.244
##          y2     0.264         0.012        0.332     22.880    0.000     0.242     0.287
##          x3    -0.271         0.024       -0.166    -11.439    0.000    -0.317    -0.224
##         x19     0.310         0.034        0.143      9.188    0.000     0.244     0.377
##         x11    -0.952         0.093       -0.140    -10.234    0.000    -1.135    -0.770
##         x14    -0.081         0.011       -0.136     -7.416    0.000    -0.103    -0.060
##          x5     0.000         0.000        0.109      7.166    0.000     0.000     0.001
## ----------------------------------------------------------------------------------------------
##
##
##
## Stepwise Selection: Step 8
##
## - x13 added
##
##                          Model Summary
## ----------------------------------------------------------------
## R                        0.848       RMSE                 9.868
## R-Squared                0.719       Coef. Var           14.385
## Adj. R-Squared           0.718       MSE                 97.376
## Pred R-Squared           0.716       MAE                  7.546
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F         Sig.
## --------------------------------------------------------------------------
## Regression    447530.042           8      55941.255    574.487    0.0000
## Residual      174497.896        1792         97.376
```

```
## Total         622027.938        1800
## ----------------------------------------------------------------------
##
##                                 Parameter Estimates
## ------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ------------------------------------------------------------------------------------------
## (Intercept)    55.710         4.272                13.041    0.000    47.331    64.088
##         x15    -0.196         0.029       -0.203    -6.824    0.000    -0.252    -0.140
##          y2     0.246         0.012        0.310    20.686    0.000     0.223     0.270
##          x3    -0.280         0.024       -0.172   -11.895    0.000    -0.326    -0.234
##         x19     0.332         0.034        0.153     9.847    0.000     0.266     0.398
##         x11    -0.818         0.095       -0.120    -8.578    0.000    -1.005    -0.631
##         x14    -0.089         0.011       -0.149    -8.092    0.000    -0.110    -0.067
##          x5     0.000         0.000        0.132     8.439    0.000     0.000     0.001
##         x13    -0.074         0.013       -0.109    -5.629    0.000    -0.099    -0.048
## ------------------------------------------------------------------------------------------
##
##
##
##                         Model Summary
## ----------------------------------------------------------------
## R                       0.848       RMSE                 9.868
## R-Squared               0.719       Coef. Var           14.385
## Adj. R-Squared          0.718       MSE                 97.376
## Pred R-Squared          0.716       MAE                  7.546
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                   ANOVA
## ----------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F          Sig.
## ----------------------------------------------------------------------
## Regression    447530.042          8      55941.255    574.487    0.0000
## Residual      174497.896        1792         97.376
## Total         622027.938        1800
## ----------------------------------------------------------------------
##
##                                 Parameter Estimates
## ------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ------------------------------------------------------------------------------------------
## (Intercept)    55.710         4.272                13.041    0.000    47.331    64.088
##         x15    -0.196         0.029       -0.203    -6.824    0.000    -0.252    -0.140
##          y2     0.246         0.012        0.310    20.686    0.000     0.223     0.270
##          x3    -0.280         0.024       -0.172   -11.895    0.000    -0.326    -0.234
##         x19     0.332         0.034        0.153     9.847    0.000     0.266     0.398
##         x11    -0.818         0.095       -0.120    -8.578    0.000    -1.005    -0.631
##         x14    -0.089         0.011       -0.149    -8.092    0.000    -0.110    -0.067
##          x5     0.000         0.000        0.132     8.439    0.000     0.000     0.001
##         x13    -0.074         0.013       -0.109    -5.629    0.000    -0.099    -0.048
```

```
## ---------------------------------------------------------------------------------------
##
##
##
## Stepwise Selection: Step 9
##
## - x7 added
##
##                          Model Summary
## ----------------------------------------------------------------
## R                        0.849       RMSE                 9.830
## R-Squared                0.720       Coef. Var           14.336
## Adj. R-Squared           0.719       MSE                 96.621
## Pred R-Squared           0.716       MAE                  7.526
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## ---------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F         Sig.
## ---------------------------------------------------------------------------
## Regression    427996.762          9      47555.196    492.181    0.0000
## Residual      166285.265        1721         96.621
## Total         594282.027        1730
## ---------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------
## (Intercept)    50.169         4.419                 11.353    0.000    41.502    58.836
##         x15    -0.198         0.029       -0.206    -6.783    0.000    -0.255    -0.141
##          y2     0.233         0.012        0.295     19.088    0.000     0.209     0.257
##          x3    -0.259         0.024       -0.159    -10.679    0.000    -0.307    -0.212
##         x19     0.338         0.036        0.150      9.462    0.000     0.268     0.408
##         x11    -0.707         0.099       -0.102     -7.153    0.000    -0.901    -0.513
##         x14    -0.092         0.011       -0.155     -8.184    0.000    -0.114    -0.070
##          x5     0.001         0.000        0.155      9.636    0.000     0.000     0.001
##         x13    -0.066         0.013       -0.098     -4.905    0.000    -0.092    -0.039
##          x7    -1.026         0.209       -0.071     -4.902    0.000    -1.436    -0.615
## ----------------------------------------------------------------------------------------
##
##
##
##                          Model Summary
## ----------------------------------------------------------------
## R                        0.849       RMSE                 9.830
## R-Squared                0.720       Coef. Var           14.336
## Adj. R-Squared           0.719       MSE                 96.621
## Pred R-Squared           0.716       MAE                  7.526
## ----------------------------------------------------------------
```

```
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F         Sig.
## --------------------------------------------------------------------------
## Regression    427996.762           9      47555.196    492.181    0.0000
## Residual      166285.265        1721         96.621
## Total         594282.027        1730
## --------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------------
## (Intercept)    50.169         4.419                 11.353    0.000    41.502    58.836
##         x15    -0.198         0.029       -0.206     -6.783    0.000    -0.255    -0.141
##          y2     0.233         0.012        0.295     19.088    0.000     0.209     0.257
##          x3    -0.259         0.024       -0.159    -10.679    0.000    -0.307    -0.212
##         x19     0.338         0.036        0.150      9.462    0.000     0.268     0.408
##         x11    -0.707         0.099       -0.102     -7.153    0.000    -0.901    -0.513
##         x14    -0.092         0.011       -0.155     -8.184    0.000    -0.114    -0.070
##          x5     0.001         0.000        0.155      9.636    0.000     0.000     0.001
##         x13    -0.066         0.013       -0.098     -4.905    0.000    -0.092    -0.039
##          x7    -1.026         0.209       -0.071     -4.902    0.000    -1.436    -0.615
## ----------------------------------------------------------------------------------------------
##
##
##
## Stepwise Selection: Step 10
##
## - x17 added
##
##                         Model Summary
## ----------------------------------------------------------------
## R                        0.850       RMSE                 9.739
## R-Squared                0.722       Coef. Var           14.130
## Adj. R-Squared           0.720       MSE                 94.841
## Pred R-Squared           0.717       MAE                  7.462
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F         Sig.
## --------------------------------------------------------------------------
## Regression    399261.115          10      39926.112    420.981    0.0000
## Residual      153926.516        1623         94.841
```

```
## Total         553187.632        1633
## ----------------------------------------------------------------------
##
##                                 Parameter Estimates
## ------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ------------------------------------------------------------------------------------------
## (Intercept)    64.853         5.452                11.895    0.000    54.159    75.547
##         x15    -0.190         0.030       -0.194    -6.228    0.000    -0.249    -0.130
##          y2     0.224         0.013        0.283    17.694    0.000     0.199     0.249
##          x3    -0.258         0.025       -0.154   -10.257    0.000    -0.307    -0.209
##         x19     0.353         0.036        0.159     9.754    0.000     0.282     0.424
##         x11    -1.245         0.163       -0.172    -7.659    0.000    -1.564    -0.926
##         x14    -0.088         0.012       -0.148    -7.556    0.000    -0.111    -0.065
##          x5     0.001         0.000        0.175     9.604    0.000     0.001     0.001
##         x13    -0.058         0.014       -0.084    -4.113    0.000    -0.085    -0.030
##          x7    -0.914         0.214       -0.063    -4.263    0.000    -1.334    -0.493
##         x17    -0.001         0.000       -0.094    -4.276    0.000    -0.002    -0.001
## ------------------------------------------------------------------------------------------
##
##
##
##                         Model Summary
## ----------------------------------------------------------------
## R                       0.850       RMSE                 9.739
## R-Squared               0.722       Coef. Var           14.130
## Adj. R-Squared          0.720       MSE                 94.841
## Pred R-Squared          0.717       MAE                  7.462
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                  ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F          Sig.
## --------------------------------------------------------------------------
## Regression    399261.115          10      39926.112    420.981    0.0000
## Residual      153926.516        1623         94.841
## Total         553187.632        1633
## --------------------------------------------------------------------------
##
##                                 Parameter Estimates
## ------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ------------------------------------------------------------------------------------------
## (Intercept)    64.853         5.452                11.895    0.000    54.159    75.547
##         x15    -0.190         0.030       -0.194    -6.228    0.000    -0.249    -0.130
##          y2     0.224         0.013        0.283    17.694    0.000     0.199     0.249
##          x3    -0.258         0.025       -0.154   -10.257    0.000    -0.307    -0.209
##         x19     0.353         0.036        0.159     9.754    0.000     0.282     0.424
##         x11    -1.245         0.163       -0.172    -7.659    0.000    -1.564    -0.926
##         x14    -0.088         0.012       -0.148    -7.556    0.000    -0.111    -0.065
```

```
##          x5     0.001         0.000        0.175      9.604    0.000     0.001     0.001
##         x13    -0.058         0.014       -0.084     -4.113    0.000    -0.085    -0.030
##          x7    -0.914         0.214       -0.063     -4.263    0.000    -1.334    -0.493
##         x17    -0.001         0.000       -0.094     -4.276    0.000    -0.002    -0.001
## ----------------------------------------------------------------------------------------
##
##
##
## Stepwise Selection: Step 11
##
## - x20 added
##
##                         Model Summary
## ----------------------------------------------------------------
## R                       0.856       RMSE                 9.596
## R-Squared               0.732       Coef. Var           13.931
## Adj. R-Squared          0.730       MSE                 92.086
## Pred R-Squared          0.727       MAE                  7.334
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F          Sig.
## --------------------------------------------------------------------------
## Regression    385782.260          11      35071.115    380.853    0.0000
## Residual      141259.613        1534         92.086
## Total         527041.874        1545
## --------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------
## (Intercept)    68.403         5.728                 11.943    0.000    57.168    79.637
##         x15    -0.207         0.032       -0.211     -6.574    0.000    -0.269    -0.145
##          y2     0.230         0.013        0.289     17.665    0.000     0.204     0.256
##          x3    -0.275         0.026       -0.165    -10.774    0.000    -0.325    -0.225
##         x19     0.361         0.037        0.161      9.803    0.000     0.289     0.434
##         x11    -1.304         0.169       -0.179     -7.727    0.000    -1.635    -0.973
##         x14    -0.092         0.012       -0.152     -7.666    0.000    -0.115    -0.068
##          x5     0.001         0.000        0.164      8.799    0.000     0.000     0.001
##         x13    -0.044         0.015       -0.063     -2.927    0.003    -0.073    -0.014
##          x7    -0.881         0.218       -0.060     -4.037    0.000    -1.308    -0.453
##         x17    -0.001         0.000       -0.090     -3.946    0.000    -0.002    -0.001
##         x20    -1.612         0.660       -0.037     -2.441    0.015    -2.908    -0.317
## ----------------------------------------------------------------------------------------
##
##
##
##                         Model Summary
```

```
## ----------------------------------------------------------------
## R                       0.856       RMSE                 9.596
## R-Squared               0.732       Coef. Var           13.931
## Adj. R-Squared          0.730       MSE                 92.086
## Pred R-Squared          0.727       MAE                  7.334
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                ANOVA
## ---------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F          Sig.
## ---------------------------------------------------------------------------
## Regression    385782.260          11      35071.115    380.853    0.0000
## Residual      141259.613        1534         92.086
## Total         527041.874        1545
## ---------------------------------------------------------------------------
##
##                                 Parameter Estimates
## ----------------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------------------
## (Intercept)    68.403         5.728                 11.943    0.000    57.168    79.637
##         x15    -0.207         0.032       -0.211     -6.574    0.000    -0.269    -0.145
##          y2     0.230         0.013        0.289     17.665    0.000     0.204     0.256
##          x3    -0.275         0.026       -0.165    -10.774    0.000    -0.325    -0.225
##         x19     0.361         0.037        0.161      9.803    0.000     0.289     0.434
##         x11    -1.304         0.169       -0.179     -7.727    0.000    -1.635    -0.973
##         x14    -0.092         0.012       -0.152     -7.666    0.000    -0.115    -0.068
##          x5     0.001         0.000        0.164      8.799    0.000     0.000     0.001
##         x13    -0.044         0.015       -0.063     -2.927    0.003    -0.073    -0.014
##          x7    -0.881         0.218       -0.060     -4.037    0.000    -1.308    -0.453
##         x17    -0.001         0.000       -0.090     -3.946    0.000    -0.002    -0.001
##         x20    -1.612         0.660       -0.037     -2.441    0.015    -2.908    -0.317
## ----------------------------------------------------------------------------------------------------
##
##
##
## Stepwise Selection: Step 12
##
## - x10 added
##
##                          Model Summary
## ----------------------------------------------------------------
## R                       0.855       RMSE                 9.610
## R-Squared               0.731       Coef. Var           14.000
## Adj. R-Squared          0.729       MSE                 92.361
## Pred R-Squared          0.725       MAE                  7.330
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
```

```
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F         Sig.
## --------------------------------------------------------------------------
## Regression    367021.341          12      30585.112    331.148    0.0000
## Residual      135216.130        1464         92.361
## Total         502237.471        1476
## --------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ----------------------------------------------------------------------------------------------
## (Intercept)    70.017         5.888                 11.891    0.000    58.467    81.566
##         x15    -0.220         0.033       -0.223     -6.726    0.000    -0.285    -0.156
##          y2     0.227         0.013        0.287     17.065    0.000     0.201     0.253
##          x3    -0.279         0.027       -0.167    -10.505    0.000    -0.332    -0.227
##         x19     0.342         0.038        0.150      9.001    0.000     0.267     0.416
##         x11    -1.247         0.177       -0.164     -7.061    0.000    -1.593    -0.901
##         x14    -0.090         0.012       -0.147     -7.334    0.000    -0.114    -0.066
##          x5     0.001         0.000        0.167      8.690    0.000     0.001     0.001
##         x13    -0.037         0.015       -0.055     -2.461    0.014    -0.067    -0.008
##          x7    -0.813         0.246       -0.051     -3.302    0.001    -1.296    -0.330
##         x17    -0.001         0.000       -0.099     -4.183    0.000    -0.002    -0.001
##         x20    -1.564         0.666       -0.036     -2.348    0.019    -2.871    -0.258
##         x10    -0.002         0.001       -0.037     -1.871    0.062    -0.005     0.000
## ----------------------------------------------------------------------------------------------
##
##
##
##                          Model Summary
## ----------------------------------------------------------------
## R                        0.855       RMSE                 9.610
## R-Squared                0.731       Coef. Var           14.000
## Adj. R-Squared           0.729       MSE                 92.361
## Pred R-Squared           0.725       MAE                  7.330
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F         Sig.
## --------------------------------------------------------------------------
## Regression    367021.341          12      30585.112    331.148    0.0000
## Residual      135216.130        1464         92.361
## Total         502237.471        1476
## --------------------------------------------------------------------------
##
##                                  Parameter Estimates
```

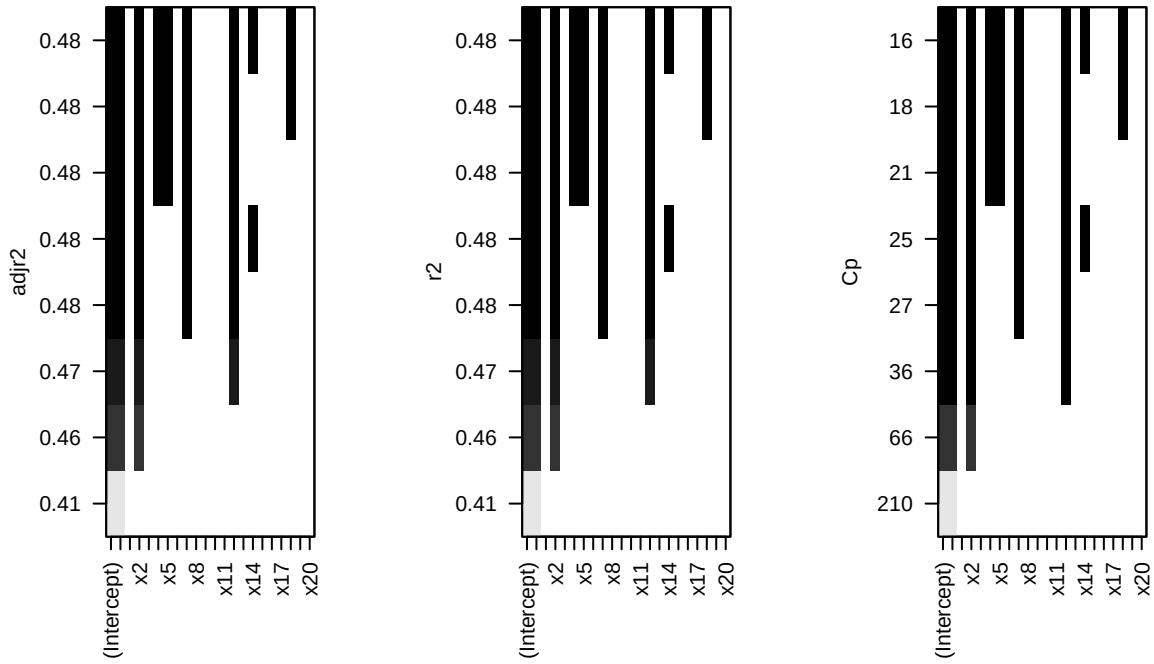
```
## ---------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig      lower     upper
## ---------------------------------------------------------------------------------------
## (Intercept)    70.017         5.888                11.891    0.000    58.467    81.566
##         x15    -0.220         0.033       -0.223    -6.726    0.000    -0.285    -0.156
##          y2     0.227         0.013        0.287    17.065    0.000     0.201     0.253
##          x3    -0.279         0.027       -0.167   -10.505    0.000    -0.332    -0.227
##         x19     0.342         0.038        0.150     9.001    0.000     0.267     0.416
##         x11    -1.247         0.177       -0.164    -7.061    0.000    -1.593    -0.901
##         x14    -0.090         0.012       -0.147    -7.334    0.000    -0.114    -0.066
##          x5     0.001         0.000        0.167     8.690    0.000     0.001     0.001
##         x13    -0.037         0.015       -0.055    -2.461    0.014    -0.067    -0.008
##          x7    -0.813         0.246       -0.051    -3.302    0.001    -1.296    -0.330
##         x17    -0.001         0.000       -0.099    -4.183    0.000    -0.002    -0.001
##         x20    -1.564         0.666       -0.036    -2.348    0.019    -2.871    -0.258
##         x10    -0.002         0.001       -0.037    -1.871    0.062    -0.005     0.000
## ---------------------------------------------------------------------------------------
##
##
##
## Stepwise Selection: Step 13
##
## - x9 added
##
##                         Model Summary
## ----------------------------------------------------------------
## R                       0.856       RMSE                 9.585
## R-Squared               0.733       Coef. Var           13.948
## Adj. R-Squared          0.730       MSE                 91.879
## Pred R-Squared          0.726       MAE                  7.308
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                  ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F         Sig.
## --------------------------------------------------------------------------
## Regression    364539.728          13      28041.518    305.201    0.0000
## Residual      133040.481        1448         91.879
## Total         497580.208        1461
## --------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ---------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig       lower     upper
## ---------------------------------------------------------------------------------------
## (Intercept)    68.504         5.933                11.547    0.000     56.866    80.141
##         x15    -0.204         0.034       -0.206    -6.066    0.000     -0.269    -0.138
##          y2     0.231         0.013        0.291    17.184    0.000      0.204     0.257
##          x3    -0.285         0.027       -0.171   -10.703    0.000     -0.338    -0.233
##         x19     0.351         0.038        0.154     9.217    0.000      0.276     0.425
```

```
##         x11    -1.221         0.178       -0.160     -6.843    0.000     -1.571    -0.871
##         x14    -0.092         0.013       -0.148     -7.262    0.000     -0.116    -0.067
##          x5     0.001         0.000        0.189      8.420    0.000      0.001     0.001
##         x13    -0.042         0.015       -0.061     -2.697    0.007     -0.072    -0.011
##          x7    -0.746         0.248       -0.047     -3.013    0.003     -1.232    -0.260
##         x17    -0.001         0.000       -0.104     -4.383    0.000     -0.002    -0.001
##         x20    -1.731         0.680       -0.040     -2.545    0.011     -3.065    -0.397
##         x10    -0.002         0.001       -0.040     -2.035    0.042     -0.005     0.000
##          x9    -6.263         3.318       -0.032     -1.888    0.059    -12.772     0.246
## ----------------------------------------------------------------------------------------
##
##
##
##                         Model Summary
## ----------------------------------------------------------------
## R                       0.856       RMSE                 9.585
## R-Squared               0.733       Coef. Var           13.948
## Adj. R-Squared          0.730       MSE                 91.879
## Pred R-Squared          0.726       MAE                  7.308
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F         Sig.
## --------------------------------------------------------------------------
## Regression    364539.728          13      28041.518    305.201    0.0000
## Residual      133040.481        1448         91.879
## Total         497580.208        1461
## --------------------------------------------------------------------------
##
##                                 Parameter Estimates
## ----------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig       lower     upper
## ----------------------------------------------------------------------------------------
## (Intercept)    68.504         5.933                 11.547    0.000     56.866    80.141
##         x15    -0.204         0.034       -0.206     -6.066    0.000     -0.269    -0.138
##          y2     0.231         0.013        0.291     17.184    0.000      0.204     0.257
##          x3    -0.285         0.027       -0.171    -10.703    0.000     -0.338    -0.233
##         x19     0.351         0.038        0.154      9.217    0.000      0.276     0.425
##         x11    -1.221         0.178       -0.160     -6.843    0.000     -1.571    -0.871
##         x14    -0.092         0.013       -0.148     -7.262    0.000     -0.116    -0.067
##          x5     0.001         0.000        0.189      8.420    0.000      0.001     0.001
##         x13    -0.042         0.015       -0.061     -2.697    0.007     -0.072    -0.011
##          x7    -0.746         0.248       -0.047     -3.013    0.003     -1.232    -0.260
##         x17    -0.001         0.000       -0.104     -4.383    0.000     -0.002    -0.001
##         x20    -1.731         0.680       -0.040     -2.545    0.011     -3.065    -0.397
##         x10    -0.002         0.001       -0.040     -2.035    0.042     -0.005     0.000
##          x9    -6.263         3.318       -0.032     -1.888    0.059    -12.772     0.246
## ----------------------------------------------------------------------------------------
##
```

```
##
##
## No more variables to be added/removed.
##
##
## Final Model Output
## ------------------
##
##                         Model Summary
## ----------------------------------------------------------------
## R                       0.856       RMSE                 9.585
## R-Squared               0.733       Coef. Var           13.948
## Adj. R-Squared          0.730       MSE                 91.879
## Pred R-Squared          0.726       MAE                  7.308
## ----------------------------------------------------------------
##  RMSE: Root Mean Square Error
##  MSE: Mean Square Error
##  MAE: Mean Absolute Error
##
##                                 ANOVA
## --------------------------------------------------------------------------
##                   Sum of
##                  Squares          DF    Mean Square       F         Sig.
## --------------------------------------------------------------------------
## Regression    364539.728          13      28041.518    305.201    0.0000
## Residual      133040.481        1448         91.879
## Total         497580.208        1461
## --------------------------------------------------------------------------
##
##                                  Parameter Estimates
## ----------------------------------------------------------------------------------------
##       model      Beta    Std. Error    Std. Beta       t        Sig       lower     upper
## ----------------------------------------------------------------------------------------
## (Intercept)    68.504         5.933                 11.547    0.000     56.866    80.141
##         x15    -0.204         0.034       -0.206    -6.066    0.000     -0.269    -0.138
##          y2     0.231         0.013        0.291    17.184    0.000      0.204     0.257
##          x3    -0.285         0.027       -0.171   -10.703    0.000     -0.338    -0.233
##         x19     0.351         0.038        0.154     9.217    0.000      0.276     0.425
##         x11    -1.221         0.178       -0.160    -6.843    0.000     -1.571    -0.871
##         x14    -0.092         0.013       -0.148    -7.262    0.000     -0.116    -0.067
##          x5     0.001         0.000        0.189     8.420    0.000      0.001     0.001
##         x13    -0.042         0.015       -0.061    -2.697    0.007     -0.072    -0.011
##          x7    -0.746         0.248       -0.047    -3.013    0.003     -1.232    -0.260
##         x17    -0.001         0.000       -0.104    -4.383    0.000     -0.002    -0.001
##         x20    -1.731         0.680       -0.040    -2.545    0.011     -3.065    -0.397
##         x10    -0.002         0.001       -0.040    -2.035    0.042     -0.005     0.000
##          x9    -6.263         3.318       -0.032    -1.888    0.059    -12.772     0.246
## ----------------------------------------------------------------------------------------

##
##                                  Stepwise Selection Summary
## ----------------------------------------------------------------------------------------
##                      Added/                   Adj.
## Step    Variable     Removed     R-Square    R-Square      C(p)          AIC         RMSE
```

```
## ------------------------------------------------------------------------------------------
##    1      x15       addition       0.492       0.492    2435.3480    16753.3141    14.1594
##    2       y2       addition       0.646       0.646    1078.4470    16013.3399    11.8267
##    3       x3       addition       0.659       0.658     770.5910    15178.8928    11.2937
##    4      x19       addition       0.671       0.671     478.5300    14351.3613    10.7173
##    5      x11       addition       0.688       0.687     341.1390    13936.0276    10.4162
##    6      x14       addition       0.698       0.697     271.7550    13876.4054    10.2470
##    7       x5       addition       0.715       0.713     149.0680    13397.5842     9.9520
##    8      x13       addition       0.719       0.718     117.4640    13368.0173     9.8679
##    9       x7       addition       0.720       0.719     100.0190    12836.3897     9.8296
##   10      x17       addition       0.722       0.720      64.4200    12088.3472     9.7386
##   11      x20       addition       0.732       0.730      16.4640    11393.4374     9.5961
##   12      x10       addition       0.731       0.729      21.6440    10890.9488     9.6104
##   13       x9       addition       0.733       0.730      14.9490    10773.8362     9.5853
## ------------------------------------------------------------------------------------------
```

2021 Model Selection



d. Best Model

```
##
## Attaching package: 'modelsummary'

## The following object is masked from 'package:Hmisc':
```

```
##
##     Mean

##
## Attaching package: 'kableExtra'

## The following object is masked from 'package:huxtable':
##
##     add_footnote

## The following object is masked from 'package:dplyr':
##
##     group_rows
```

Based on the model selection, the best model for the 2019 Math Achievement Rate consists of independent variables of 'Absent 0-5 Days Percentage', 'Avg. Annual Salaries for Teachers', 'Number of Teachers with a phd degree', 'White Student Percentage', 'Black Student Percentage', 'Economically Disadvantaged Student Percentage', 'Percentage of Gifted Students', and 'Urban/Rural Area of the School'. The best model for the 2021 Math Achievement Rate consists of independent variables of 'Absent 0-5 Days Percentage', 'Avg. Annual Salaries for Teachers', 'Number of Teachers with a phd degree', 'White Student Percentage', 'Economically Disadvantaged Student Percentage', 'Amount of Money Invested for Students', 'Per-Pupil Expenditure at School Level', and 'Urban/Rural Area of the School'.

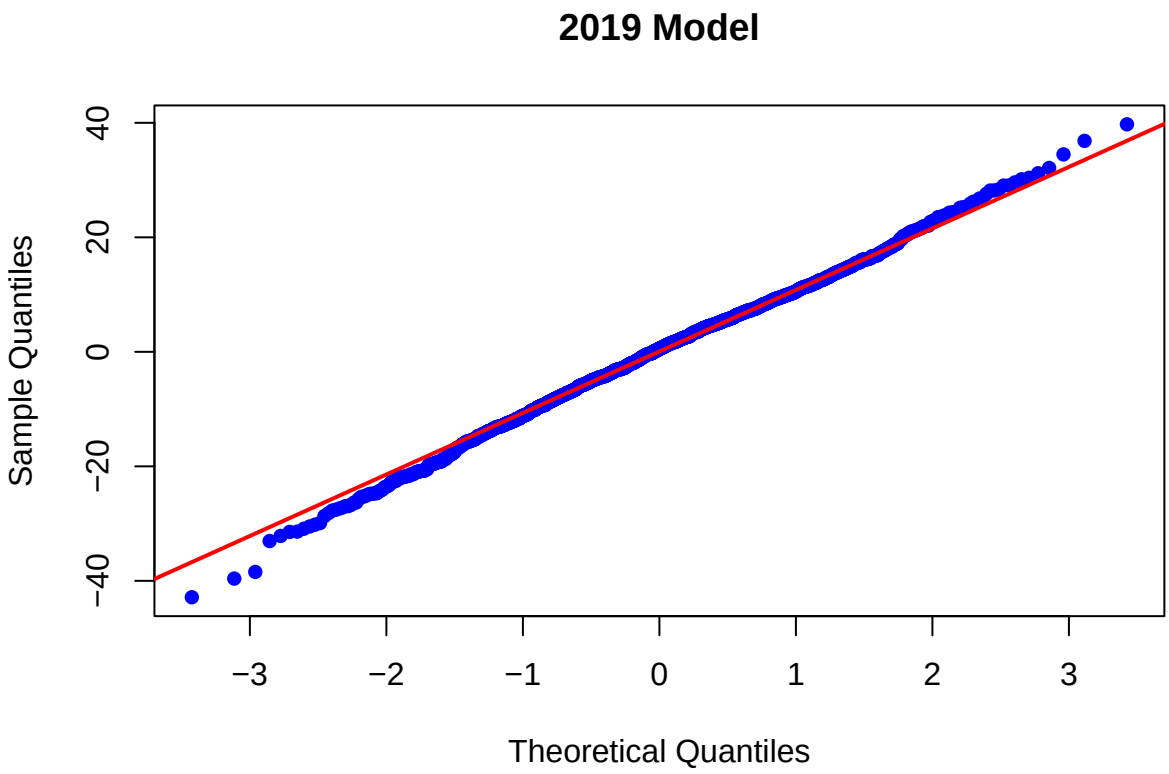
e. Best Model (Outlier Excluded)

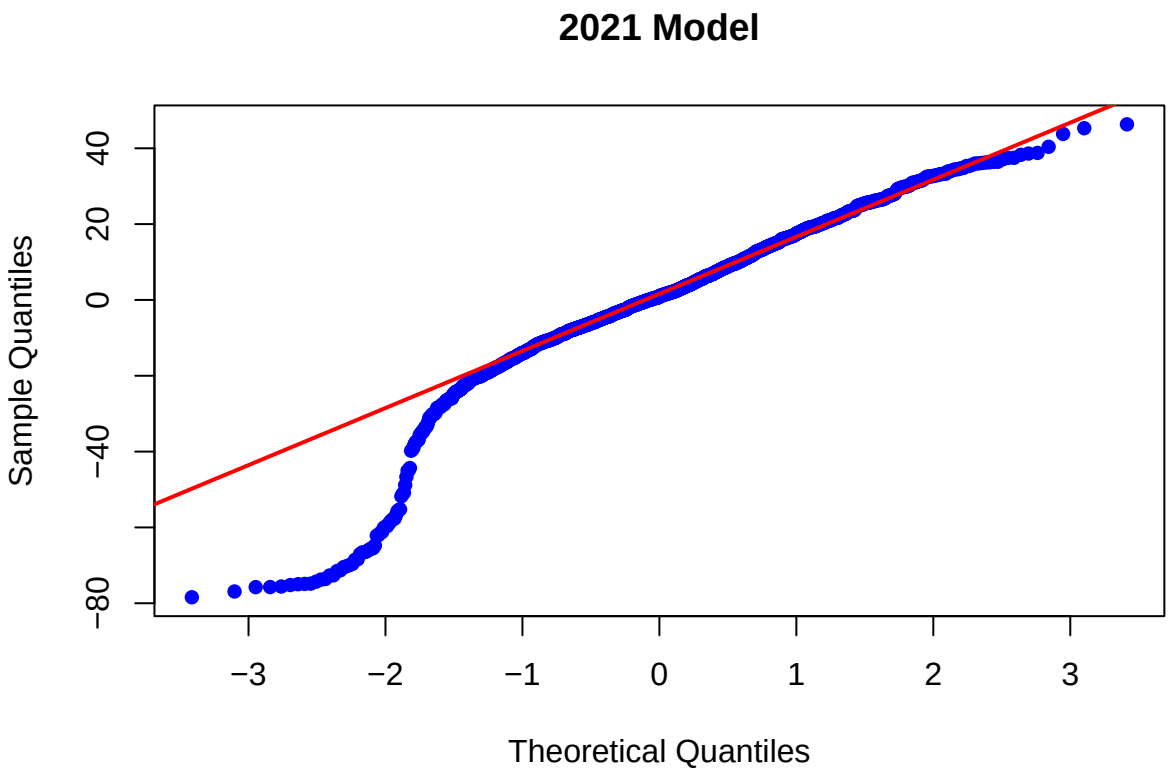
2019

2021

f. Normality Check

```
##
##  Anderson-Darling normality test
##
## data:  resid(best_model_2019)
## A = 1.2077, p-value = 0.003783

##
##  Anderson-Darling normality test
##
## data:  resid(best_model_2021)
## A = 20.141, p-value < 2.2e-16
```


2019 Model
Sample Quantiles
40
20
0
−20
−40
−3
−2
−1
0
1
2
3
Theoretical Quantiles



g. Transformation

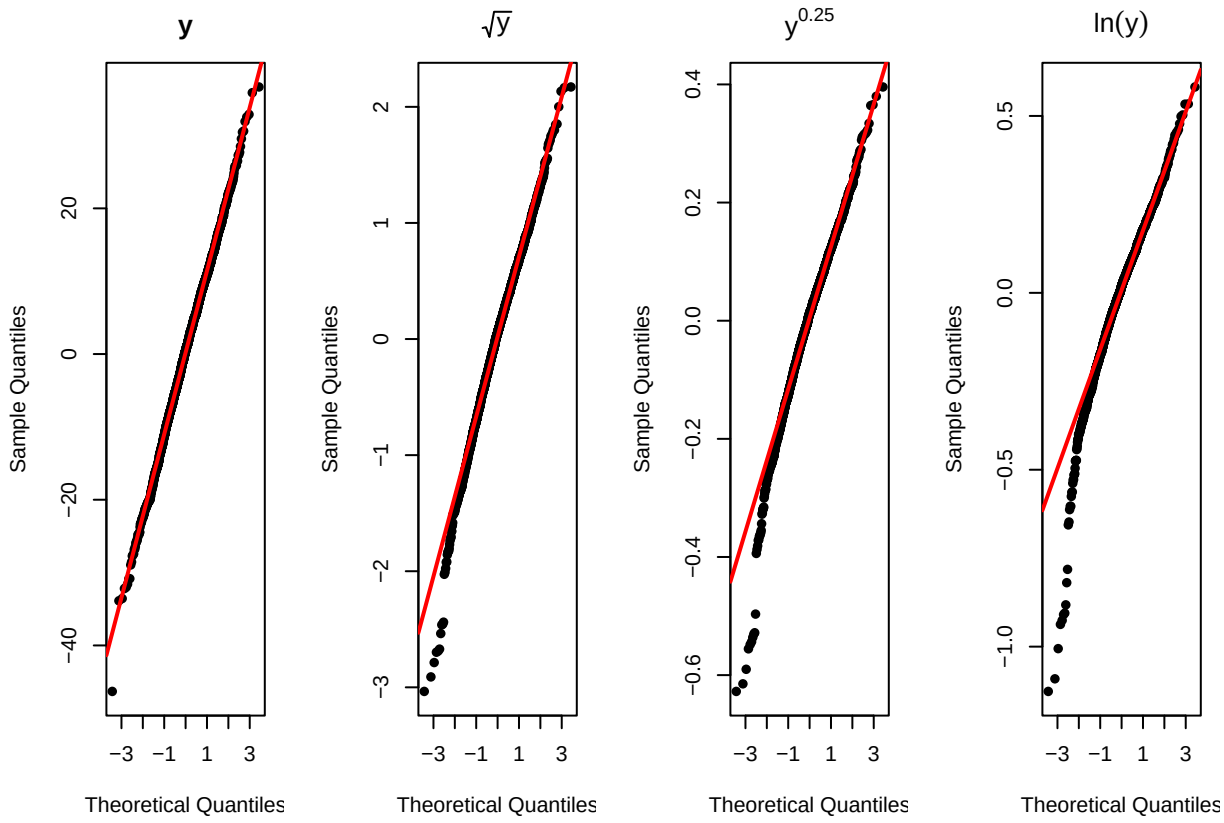
2019

```
##
## Call:
## lm(formula = y1 ~ x1 + x2 + x5 + x7 + x12 + x14 + x17 + x19,
##     data = data_numeric)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -46.313  -7.365   0.463   7.765  36.663
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -3.112e+01  4.865e+00  -6.397 2.06e-10 ***
## x1           4.374e-01  2.554e-02  17.125  < 2e-16 ***
## x2           5.872e-01  5.295e-02  11.088  < 2e-16 ***
## x5           1.004e-03  6.577e-05  15.261  < 2e-16 ***
## x7          -1.630e+00  2.429e-01  -6.710 2.66e-11 ***
## x12          1.403e-01  1.362e-02  10.302  < 2e-16 ***
## x14         -1.561e-01  1.108e-02 -14.087  < 2e-16 ***
## x17          5.903e-05  1.846e-04   0.320    0.749
## x19          5.076e-01  4.392e-02  11.559  < 2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.22 on 1647 degrees of freedom
##   (411 observations deleted due to missingness)
## Multiple R-squared:  0.6354, Adjusted R-squared:  0.6336
## F-statistic: 358.7 on 8 and 1647 DF,  p-value: < 2.2e-16

##
## Call:
## lm(formula = trans_y1 ~ x1 + x2 + x5 + x7 + x12 + x14 + x17 +
##     x19, data = data_numeric)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.03527 -0.44353  0.05277  0.48535  2.17113
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.750e+00  3.109e-01   5.629 2.13e-08 ***
## x1           2.840e-02  1.632e-03  17.395  < 2e-16 ***
## x2           4.048e-02  3.384e-03  11.963  < 2e-16 ***
## x5           6.324e-05  4.203e-06  15.044  < 2e-16 ***
## x7          -1.006e-01  1.553e-02  -6.482 1.20e-10 ***
## x12          9.002e-03  8.704e-04  10.342  < 2e-16 ***
## x14         -9.254e-03  7.083e-04 -13.065  < 2e-16 ***
## x17          2.116e-06  1.179e-05   0.179    0.858
## x19          3.064e-02  2.807e-03  10.915  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.7173 on 1647 degrees of freedom
##   (411 observations deleted due to missingness)
## Multiple R-squared:  0.6237, Adjusted R-squared:  0.6219
## F-statistic: 341.3 on 8 and 1647 DF,  p-value: < 2.2e-16

##
## Call:
## lm(formula = trans_y2 ~ x1 + x2 + x5 + x7 + x12 + x14 + x17 +
##     x19, data = data_numeric)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.62758 -0.07664  0.01175  0.08596  0.39569
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.685e+00  5.664e-02  29.758  < 2e-16 ***
## x1           5.158e-03  2.974e-04  17.343  < 2e-16 ***
## x2           7.575e-03  6.165e-04  12.286  < 2e-16 ***
## x5           1.132e-05  7.658e-07  14.783  < 2e-16 ***
## x7          -1.777e-02  2.829e-03  -6.283 4.24e-10 ***
## x12          1.621e-03  1.586e-04  10.220  < 2e-16 ***
## x14         -1.601e-03  1.290e-04 -12.405  < 2e-16 ***
## x17          2.480e-07  2.149e-06   0.115    0.908
```

```
## x19          5.373e-03  5.113e-04  10.508  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1307 on 1647 degrees of freedom
##   (411 observations deleted due to missingness)
## Multiple R-squared:  0.6128, Adjusted R-squared:  0.6109
## F-statistic: 325.8 on 8 and 1647 DF,  p-value: < 2.2e-16
```

```
##
## Call:
## lm(formula = trans_y3 ~ x1 + x2 + x5 + x7 + x12 + x14 + x17 +
##     x19, data = data_numeric)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.12661 -0.10550  0.01824  0.12123  0.58191
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2.476e+00  8.377e-02  29.558  < 2e-16 ***
## x1           7.538e-03  4.399e-04  17.137  < 2e-16 ***
## x2           1.140e-02  9.119e-04  12.503  < 2e-16 ***
## x5           1.632e-05  1.133e-06  14.406  < 2e-16 ***
## x7          -2.519e-02  4.184e-03  -6.022 2.12e-09 ***
## x12          2.343e-03  2.345e-04   9.988  < 2e-16 ***
## x14         -2.223e-03  1.909e-04 -11.646  < 2e-16 ***
## x17          1.825e-07  3.178e-06   0.057    0.954
## x19          7.592e-03  7.563e-04  10.039  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1933 on 1647 degrees of freedom
##   (411 observations deleted due to missingness)
## Multiple R-squared:  0.5974, Adjusted R-squared:  0.5955
## F-statistic: 305.5 on 8 and 1647 DF,  p-value: < 2.2e-16
```



2021

h. Influential Points

```
## named numeric(0)
```

```
## named numeric(0)
```

VIF

```
## Loading required package: carData
```

```
##
## Attaching package: 'carData'
```

```
## The following object is masked _by_ '.GlobalEnv':
##
##     Salaries
```

```
##
## Attaching package: 'car'
```

```
## The following object is masked from 'package:dplyr':
##
##     recode
```

```
##       x1       x5       x7      x12      x13      x14      x19      x20
## 1.251926 1.349700 1.160524 3.499350 3.446830 1.613986 1.470418 1.320757
```

```
##       x1       x5       x7      x12      x14      x16      x17      x20
## 1.141084 1.541745 1.234421 1.564701 1.592821 1.575792 1.269641 1.297179
```

7. Residual Plot

```
# Residual Plot
# plot(data_numeric$y2, resid(best_model_2021), pch=16, col="blue")
# abline(0, 0, col = "red", lwd = 3)
# plot(fitted(best_model_2021), resid(best_model_2021), pch=16, col="blue", ylab=bquote(paste("e")))
# abline(0, 0, col = "red", lwd = 3)
```

Category

1. Urban & Rural

```
urban = data[data$Urban.Rural == "Urban", ]
rural = data[data$Urban.Rural == "Rural", ]
```

Testing if mean of Urban and Rural Math Achievement Rates are equal

$$H_0 : \mu_{Urban} - \mu_{Rural} = 0$$
$$H_0 : \mu_{Urban} - \mu_{Rural} > 1$$
$$p - value = 0.006737 < \alpha = 0.05 \rightarrow Reject\ H_0$$

```
mean(urban$All.Students.Math.Achievement)
```

```
## [1] 67.43092
```

```
mean(rural$All.Students.Math.Achievement)
```

```
## [1] 63.96621
```

```
t.test(urban$All.Students.Math.Achievement, rural$All.Students.Math.Achievement,
       mu=1, alternative='greater')
```

```
##
##  Welch Two Sample t-test
##
## data:  urban$All.Students.Math.Achievement and rural$All.Students.Math.Achievement
```

```
## t = 2.5042, df = 809.73, p-value = 0.006234
## alternative hypothesis: true difference in means is greater than 1
## 95 percent confidence interval:
##  1.843955      Inf
## sample estimates:
## mean of x mean of y
##  67.43092  63.96621
```

2. Race

Testing if the difference in mean of White and Black Math Achievement Rates is greater than 13

$$H_0 : \mu_{White} - \mu_{Black} = 0$$
$$H_0 : \mu_{White} - \mu_{Black} > 13$$
$$p - value = 0.004886 < \alpha = 0.05 \rightarrow Reject\ H_0$$

```
mean(data$White.Math.Achievement)
```

```
## [1] 63.89831
```

```
mean(data$Black.Math.Achievement)
```

```
## [1] 48.41171
```

```
t.test(data$White.Math.Achievement, data$Black.Math.Achievement,
       mu=13, alternative='greater')
```

```
##
##  Welch Two Sample t-test
##
## data:  data$White.Math.Achievement and data$Black.Math.Achievement
## t = 2.6881, df = 3705.6, p-value = 0.003609
## alternative hypothesis: true difference in means is greater than 13
## 95 percent confidence interval:
##  13.96467      Inf
## sample estimates:
## mean of x mean of y
##  63.89831  48.41171
```

```
mean(urban$White.Percentage)
```

```
## [1] 35.32944
```

```
mean(rural$White.Percentage)
```

```
## [1] 50.33252
```

```r
mean(urban$Black.Percentage)
```

```
## [1] 38.83198
```

```r
mean(rural$Black.Percentage)
```

```
## [1] 32.6066
```

3. Economy

```r
# 100% Econ Disadv Percentage
Econ_Dia_100 = data[data$Econ.Disadvantaged.Percentage == '100', ]
Econ_Dia_100_urban = Econ_Dia_100[Econ_Dia_100$Urban.Rural == "Urban",]
Econ_Dia_100_rural = Econ_Dia_100[Econ_Dia_100$Urban.Rural == "Rural",]
# 2019
c(mean(Econ_Dia_100_urban$All.Students.Math.Achievement),
  mean(Econ_Dia_100_rural$All.Students.Math.Achievement))
```

```
## [1] 54.89206 58.82841
```

```r
# 2021
c(mean(Econ_Dia_100_urban$X2021.All.Students.Math.Achievement),
  mean(Econ_Dia_100_rural$X2021.All.Students.Math.Achievement))
```

```
## [1] 36.54889 47.08022
```

$$H_0 : \mu_{Rural\ EconDis} - \mu_{Urban\ EconDis} = 0$$
$$H_0 : \mu_{Rural\ EconDis} - \mu_{Urban\ EconDis} > 15$$
$$p{-}value = 0.04061 < \alpha = 0.05 \rightarrow Reject\ H_0$$

```r
mean(urban$Econ.Disadvantaged.Percentage)
```

```
## [1] 65.87646
```

```r
mean(rural$Econ.Disadvantaged.Percentage)
```

```
## [1] 83.43863
```

```r
t.test(rural$Econ.Disadvantaged.Percentage, urban$Econ.Disadvantaged.Percentage ,
       mu=15, alternative='greater')
```

```
##
##  Welch Two Sample t-test
##
## data:  rural$Econ.Disadvantaged.Percentage and urban$Econ.Disadvantaged.Percentage
```

```
## t = 1.7462, df = 809.96, p-value = 0.04058
## alternative hypothesis: true difference in means is greater than 15
## 95 percent confidence interval:
##  15.14594      Inf
## sample estimates:
## mean of x mean of y
##  83.43863  65.87646
```

4. Teacher Certificates

$$H_0 : \mu_{Urban\ Certificates} - \mu_{Rural\ Certificates} = 0$$
$$H_0 : \mu_{Urban\ Certificates} - \mu_{Rural\ Certificates} > 10$$
$$p{-}value = 0.001039 < \alpha = 0.05 \rightarrow Reject\ H_0$$

```
# Number of total certificates at school level
mean(urban$Total)
```

```
## [1] 59.34454
```

```
mean(rural$Total)
```

```
## [1] 44.75061
```

```
t.test(urban$Total, rural$Total,
       mu=10, alternative='greater')
```

```
##
##  Welch Two Sample t-test
##
## data:  urban$Total and rural$Total
## t = 3.1267, df = 806.62, p-value = 0.0009156
## alternative hypothesis: true difference in means is greater than 10
## 95 percent confidence interval:
##  12.17444      Inf
## sample estimates:
## mean of x mean of y
##  59.34454  44.75061
```

Reference

[1] Mervosh, Sarah. "The Pandemic Erased Two Decades of Progress in Math and Reading." The New York Times, The New York Times, 1 Sept. 2022, https://www.nytimes.com/2022/09/01/us/national-test-scores-math-reading- pandemic.html?smid=nytcore-ios-share&referringSource=articleShare.

[2] Stern, Paul. "The Pandemic Worsened Racial Achievement Gaps. Making up the Difference Won't Be Easy." CT Mirror, 23 May 2022, https://ctmirror.org/2022/05/22/the-pandemic-worsened-racial-achievement-gaps-making-up-the-difference-wont-be-easy/.

[3] Georgia Department of Education. CCRPI Reports. Retrieved from https://www.gadoe.org/CCRPI/

Pages/default.aspx

[4] The Governor's Office of Student Achievement. Downloadable Dataset. Retrieved from https://gosa.georgia.gov/dashboards-data-report-card/downloadable-data

row	column	cor	p
x10	x16	0.896	0
x8	x16	0.894	0
x8	x10	0.844	0
x1	x3	-0.833	0
x12	x13	-0.815	0
y1	x15	-0.702	0
x14	x15	0.699	0
x13	x15	0.688	0
y1	y2	0.679	0
x11	x17	-0.64	0
y1	x13	-0.639	0
x1	x2	-0.592	0
y1	x14	-0.588	0
x15	x19	-0.583	0
x5	x9	0.575	0
y2	x13	-0.553	0
y1	x12	0.55	0
y2	x12	0.541	0
x12	x15	-0.532	0
x5	x15	-0.504	0
x10	x11	0.503	0
y1	x19	0.501	0
x5	x6	0.494	0
x13	x14	0.489	0
x4	x5	0.488	0
y1	x3	-0.478	0
x10	x17	-0.472	0
y2	x15	-0.472	0
x12	x14	-0.431	0
x15	x18	0.43	0
x14	x19	-0.43	0
x18	x19	-0.425	0
x2	x12	-0.424	0

	2019 Best Model	2021 Best Model
(Intercept)	16.436	−12.483
	(4.505)	(7.417)
x1	0.221	0.106
	(0.019)	(0.031)
x5	0.0009	0.001
	(0.000 07)	(0.0001)
x7	−2.141	−3.135
	(0.237)	(0.458)
x12	0.058	0.349
	(0.019)	(0.022)
x13	−0.197	
	(0.019)	
x14	−0.150	−0.091
	(0.012)	(0.020)
x19	0.409	
	(0.043)	
x20	−2.143	−2.880
	(0.764)	(1.283)
x16		0.000 000 3
		(0.000 000 3)
x17		−0.0004
		(0.0003)
Num.Obs.	1628	1561
R2	0.638	0.355
R2 Adj.	0.636	0.352
AIC	12 515.8	13 646.9
BIC	12 569.7	13 700.4
Log.Lik.	−6247.892	−6813.438
F		106.702
RMSE	11.23	19.03