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

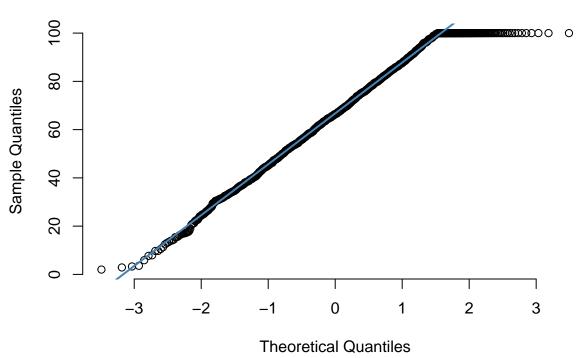
4. Data Description

```
# 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
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'
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"))
```

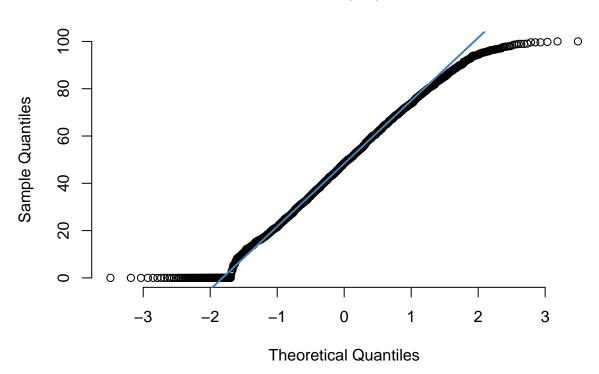
```
data = left_join(
          data %>% group_by(School.Name) %>% mutate(id = row_number()),
          Poverty.Percentage %% group_by(SCHOOL_NAME) %>% mutate(id = row_number()),
          by = c("School.Name" = "SCHOOL_NAME", "id"))
data = left_join(
          data %>% group_by(School.Name) %>% mutate(id = 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</pre>
data$u.r_dummy <- as.character(data$u.r_dummy)</pre>
data$u.r_dummy[data$u.r_dummy == "Urban"] <- 1</pre>
data$u.r_dummy[data$u.r_dummy == "Rural"] <- 0</pre>
data$u.r_dummy <- as.numeric(data$u.r_dummy)</pre>
data$growth.rate.math <-data$X19.21.Difference.in.Math</pre>
```

a. Data Summary





Normal Q-Q Plot



```
## 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
```

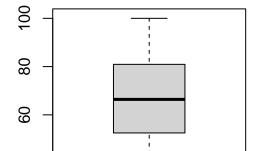
[1] -0.2318585

c. Data Visulization

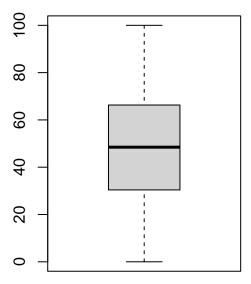
20

0

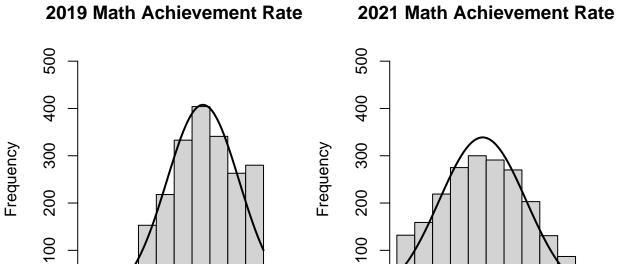
2019 Math Achievement Rate



2021 Math Achievement Rate



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.



2019 All Students Math Achievement Ra

Math Test Score

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
х3	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

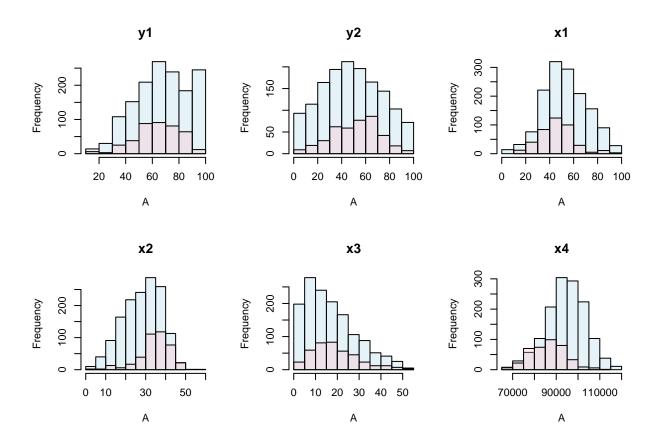
a. Iterations of the analysis process

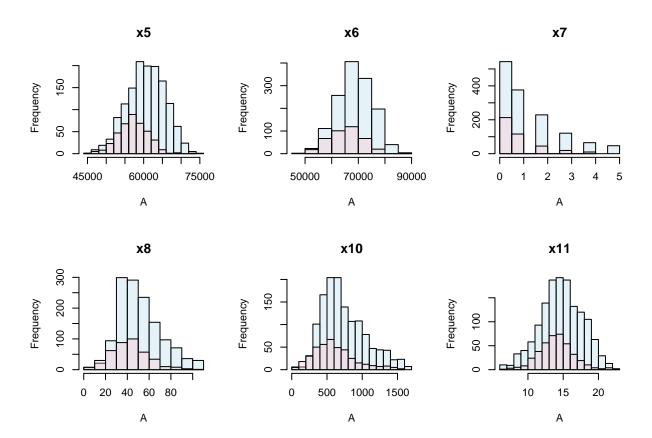
• paragraph description

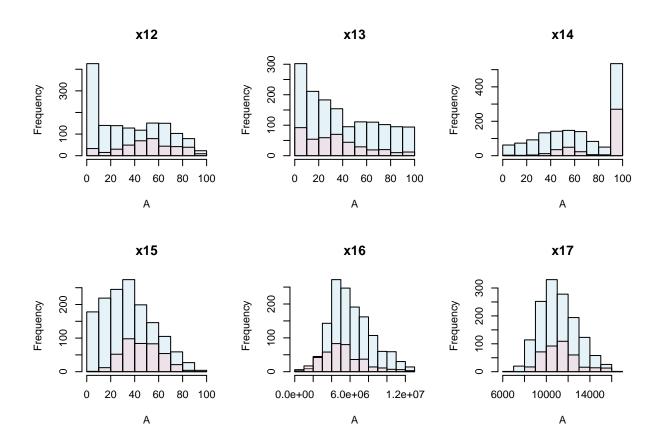
c. Plots of variables- Scatterplot

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
```







b. Multicollinearity

corrplot 0.92 loaded

```
0.8
                                                               0.8
0.6
                                                               0.6
0.4
                                                               ₩.4
0.2
                                                              Ю.2
                                                                0
-0.2
                                                               -0.2
-0.4
                                                                -0.4
-0.6
                                                                0.6
-0.8
                                                               -0.8
```

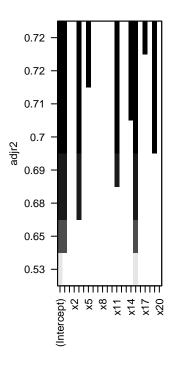
```
## Loading required package: lattice
## Loading required package: survival
## Loading required package: Formula
## Loading required package: ggplot2
##
## Attaching package: 'ggplot2'
   The following object is masked from 'package:huxtable':
##
##
       theme_grey
##
## Attaching package: 'Hmisc'
   The following objects are masked from 'package:huxtable':
##
##
       contents, label, label <-
## The following objects are masked from 'package:dplyr':
##
##
       src, summarize
```

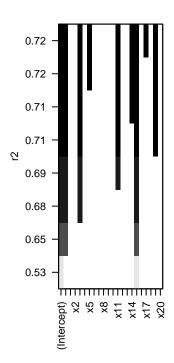
```
## The following objects are masked from 'package:base':
##
## format.pval, units
```

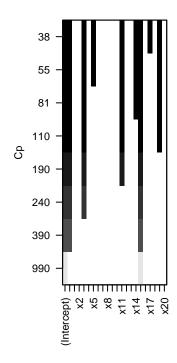
Before doing the model selection process, a Multicollinearity check produces high correlation of (x1:x3), (x8:x10,x16), (x10:x16), (x12: x13), (x11: x17), (x15: x13, x14). And another set of variables that have a high correlation is y1 and y2, since we are modeling them separately as response variables, 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
                                    xЗ
                                          x4
                                                x5
                                                      x6
                                                            x7
                                                                 8x
                                                                       x9
                                                                            x10
## 1
           TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
  2
##
           TRUE
                 TRUE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE
##
  3
           TRUE
                 TRUE FALSE FALSE
                                  TRUE FALSE FALSE FALSE FALSE FALSE FALSE
##
  4
           TRUE
                 TRUE FALSE FALSE
                                  TRUE FALSE FALSE FALSE FALSE FALSE FALSE
                 TRUE FALSE FALSE
                                  TRUE FALSE FALSE FALSE FALSE FALSE FALSE
##
  5
           TRUE
  6
           TRUE
                 TRUE FALSE FALSE
                                  TRUE FALSE FALSE FALSE FALSE FALSE FALSE
##
  7
                                              TRUE FALSE FALSE FALSE FALSE
##
           TRUE
                 TRUE FALSE FALSE
                                  TRUE FALSE
##
  8
           TRUE
                 TRUE FALSE FALSE
                                  TRUE FALSE
                                              TRUE FALSE FALSE FALSE FALSE
##
      x11
            x12
                  x13
                        x14
                            x15
                                  x16
                                        x17
                                              x18
                                                    x19
                                                          x20
## 1 FALSE FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE
## 2 FALSE FALSE FALSE TRUE FALSE FALSE FALSE FALSE
```

```
## 3 FALSE FALSE FALSE TRUE FALSE FALSE FALSE FALSE FALSE
## 4 TRUE FALSE FALSE FALSE TRUE FALSE FALSE FALSE FALSE
## 5 TRUE FALSE FALSE TRUE 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
## 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
                     0.702 RMSE
0.492 Coef. Var
0.492 MSE
                                                        14.159
## R-Squared
                                                        21.333
## Adj. R-Squared
                                                        200.489
```

	ed			AE 	10.974	_		
RMSE: Root MSE: Mean S	Mean Square	Error						
			ANOV					
			DF	Mean Square		S:	ig.	
Regression Residual Total	400015.583 412205.640 812221.223	i 1	1 2056 2057	400015.583 200.489	1995.198			
			Pai	rameter Estima	tes			
model	Beta	Std.	Error	Std. Beta	t	Sig	lower	
(Intercept)	91.915		0.651		141.087	0.000	90.637	
x15				-0.702	-44.668 	0.000	-0.701	-0.6
-	ection: Step.	2						
- y2 added	ection: Step		L Summary	y				
- y2 added		Model				_		
- y2 added		Model 0.804	Ri		11.827	_		
- y2 addedR R-Squared Adj. R-Squar		Model 0.804 0.646 0.646	RI Co MS	MSE pef. Var SE	11.827 17.819 139.870	_		
- y2 addedR R-Squared Adj. R-Squar	red red	Model 0.804 0.646 0.646 0.644	RI Co MS	MSE pef. Var	11.827 17.819	-		
- y2 added R R-Squared Adj. R-Squar Pred R-Squar RMSE: Root MSE: Mean S MAE: Mean A	ed ed Mean Square	Model 0.804 0.646 0.644 Error	RI Co MS	MSE pef. Var SE	11.827 17.819 139.870	_		
- y2 added R R-Squared Adj. R-Squar Pred R-Squar RMSE: Root MSE: Mean S MAE: Mean A	red red Mean Square square Error	Model 0.804 0.646 0.644 Error	RI Co MS	MSE pef. Var SE AE	11.827 17.819 139.870	_		
- y2 added R R-Squared Adj. R-Squar Pred R-Squar RMSE: Root MSE: Mean S MAE: Mean A	red red Mean Square square Error	Model 0.804 0.646 0.644 Error	RI Co MS MA	MSE pef. Var SE AE	11.827 17.819 139.870	_		
- y2 added R R-Squared Adj. R-Squar Pred R-Squar RMSE: Root MSE: Mean S MAE: Mean A	red red Mean Square Square Error bsolute Erro Sum of Squares	Model 0.804 0.646 0.644 Error	RI Co MS MA	MSE Def. Var SE AE WA Mean Square	11.827 17.819 139.870 8.564	-	 ig.	
- y2 added R R-Squared Adj. R-Squar Pred R-Squar RMSE: Root MSE: Mean S MAE: Mean A	red red Mean Square Square Error bsolute Erro Sum of Squares 524787.972	Model 0.804 0.646 0.644 	ANOV	MSE Def. Var SE AE Mean Square 262393.986	11.827 17.819 139.870 8.564 	-		
- y2 added R R-Squared Adj. R-Squar Pred R-Squar RMSE: Root MSE: Mean S MAE: Mean A	red red Mean Square Square Error bsolute Erro Sum of Squares 524787.972	Model 0.804 0.646 0.644 Error	RI Co MS MA ANOV	MSE Def. Var SE AE Mean Square 262393.986	11.827 17.819 139.870 8.564 	- 		
- y2 added R R-Squared Adj. R-Squar Pred R-Squar RMSE: Root MSE: Mean A Regression Residual Total	Mean Square Square Error Sum of Squares 524787.972 287433.251	Model 0.804 0.646 0.644 Error	ANOV DF 2 2055 2057	MSE Def. Var SE AE Mean Square 262393.986	11.827 17.819 139.870 8.564 	- 		
- y2 added R R-Squared Adj. R-Squar Pred R-Squar MSE: Root MSE: Mean S MAE: Mean A	Sum of Squares 524787.972 287433.251 812221.223	Model 0.804 0.646 0.644 Error	ANOV DF 2 2055 2057	MSE Def. Var SE AE Mean Square 262393.986	11.827 17.819 139.870 8.564 F 1875.98	- 		

	66.569 -0.468 0.366		1.008 0.014 0.012	-	0.490 0.446	66.033 -32.832 29.867	0.000	64.592 -0.496 0.342	68.54 -0.44 0.39
		Modo	l Cummo						
			l Summa	т у			_		
R		0.804		RMSE		11.827			
R-Squared		0.646		Coef. V	ar	17.819			
Adj. R-Squared		0.646		MSE		139.870			
Pred R-Squared	i 	0.644 		MAE		8.564	_		
RMSE: Root Me	ean Square	Error							
MSE: Mean Squ	uare Error								
MAE: Mean Abs	solute Erro	r							
			AN	AVO					
	Sum of Squares		DE	' Mos	n Sauare	F		Sig.	
	=		DI	nec					
Regression	524787.972		2	26	2393.986	1875.98	2 0.0	0000	
Residual					139.870				
	812221.223		2057						
			P	aramete	r Estima	tes			
model	Beta	Std.	Error	Std.	Beta	t	Sig	lower	uppe
(Intercept)	66.569		1.008			66.033	0.000	64.592	68.54
x15	-0.468		0.014	-	0.490	-32.832	0.000	-0.496	-0.44
у2	0.366		0.012		0.446	29.867	0.000	0.342	0.39
Stepwise Selec	ction: Step	3							
	ction: Step	3							
- x3 added		Mode	l Summa	-			_		
- x3 added		Mode:		-		 11.294	-		
- x3 added		Mode: 		RMSE			-		
- x3 added		Mode 0.812 0.659		RMSE		11.294	-		

MAE: Mean Absolute Error

## ##				AN	OVA				
## ##		 Sum of							
##		Squares			Mean Square	F	Ç	Sig.	
## ##	Regression Residual	484915.179		3 1970	161638.393 127.548		3 0.0	0000	
## ##									
##					arameter Estimat	tes			
##		Beta	Std.	Error	Std. Beta	t	•	lower	upper
	(Intercept)	71.865		1.056		68.045			
##					-0.441				
##					0.409 -0.170				
## ##		-∪.∠ŏō 		U.UZ4 	-0.170	-11.083		-0.334 	-0.238
##									
##									
##									
##			Mode	l Summa	ry				
## ##	D		010		RMSE	11.294	_		
	R-Squared				Coef. Var				
	Adj. R-Squared					127.548			
	Pred R-Squared		0.656		MAE	8.318			
##							_		
##	RMSE: Root Me	_	Error						
## ##	MSE: Mean Squ MAE: Mean Abs		-						
## ##	MAE: Mean AD	solute Ello	Ľ						
##				AN	OVA				
##									
##		Sum of							
##		Squares		DF	Mean Square	F		Sig.	
	Regression				161638 303	1267.273		2000	
	Residual						5 0.0	3000	
		736185.095		1973					
##									
##									
##				P	arameter Estimat	tes			
## ##					Std. Beta				upper
	moder							TOMET	
##	(Intercept)	71.865		1.056		68.045	0.000	69.794	73.936
##					-0.441				
##	•				0.409				
##	x3	-0.286		0.024	-0.170	-11.683	0.000	-0.334	-0.238

##

```
##
##
## Stepwise Selection: Step 4
## - x19 added
##
                   Model Summary
## -----
                       RMSE
Coef. Var
## R
                  0.819
                                        10.717
## R-Squared
                  0.671
                                       15.669
## Adj. R-Squared
                  0.671
                         MSE
                                       114.860
                  0.669
## Pred R-Squared
                         MAE
                                         7.989
## -----
## RMSE: Root Mean Square Error
## MSE: Mean Square Error
## MAE: Mean Absolute Error
##
##
                        ANOVA
## ------
              Sum of
##
             Squares
                        DF Mean Square F
## ------
                       4
                                       964.074 0.0000
## Regression
                              110733.428
           442933.711
## Residual
                      1887
           216740.546
                              114.860
## Total
           659674.257
                       1891
##
                         Parameter Estimates
          Beta Std. Error Std. Beta
     model
## (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
      у2
##
            0.313
                      0.012
                              0.393
                                     26.410 0.000
                                                   0.289
                                                          0.336
          -0.295
                                    -12.303 0.000 -0.342
##
       x3
                      0.024
                              -0.180
                                                          -0.248
      x19
##
                      0.035
                               0.144
                                      8.939 0.000
            0.316
                                                   0.247
                                                          0.386
##
##
##
                  Model Summary
##
  ______
                  0.819 RMSE
0.671 Coef. Var
0.671 MSE
0.669 MAE
                                        10.717
## R-Squared
                                        15.669
## Adj. R-Squared
                                       114.860
## Pred R-Squared
                                         7.989
## -----
 RMSE: Root Mean Square Error
## MSE: Mean Square Error
## MAE: Mean Absolute Error
##
##
                         ANOVA
## -----
##
              Sum of
```

##		Squares						F	Sig.		
## ##	Regression Residual Total	442933.711		4 1887		1107		964.074	0.0000	_	
##										_	
## ##				р	arame	ter	Estimat	tes			
##											
##	model	Beta 	Std.	Error	St	d. B	eta 	t	Sig	lower	upper
	(Intercept)							53.985			69.782
##	_					-0.	373	-20.736	0.000	-0.390	-0.323
##	у2	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 Sele	ction: Step	5								
##											
##	- x11 added										
##											
##				l Summa	ry						
									-		
##								10.416			
	-					Var					
	Adj. R-Square				MSE			108.497			
	Pred R-Square							7.843			
									-		
	RMSE: Root M	-	Error								
##	1										
##	MAE: Mean Ab	solute Erro	r								
##				A M	OVA						
										_	
##		Sum of									
##		Squares		DF	М	ean	Square	F	Sig.		
##								- 		_	
##	Regression	440535.448		5		881	07.090	812.068	0.0000	ı	
	Residual	200177.339		1845			08.497				
	Total	640712.787		1850							
##										_	
##											
##				P	arame	ter	Estimat	tes			
##											
##	model	Beta	Std.	Error	St	d. B	eta	t	Sig	lower	upper
##											
##	(Intercept)	81.760		1.897				43.108		78.040	85.479
##	x15	-0.405		0.017		-0.		-23.175		-0.440	-0.371
##	у2	0.283		0.012			356	23.932	0.000	0.260	0.307
##	x3	-0.256		0.024		-0.		-10.649		-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
                           RMSE
Coef. Var
## R
                     0.829
                                               10.416
## R-Squared
                     0.688
                                               15.192
## Adj. R-Squared
                             MSE
                     0.687
                                              108.497
                             MAE
## Pred R-Squared
                    0.684
                                               7.843
   RMSE: Root Mean Square Error
## MSE: Mean Square Error
## MAE: Mean Absolute Error
##
##
                             ANOVA
##
                Sum of
##
              Squares
                            DF Mean Square
                                                F
                                                        Sig.
  _____
                          5
## Regression 440535.448
                                  88107.090
                                              812.068 0.0000
                        1845
1850
## Residual 200177.339
                                    108.497
## Total
             640712.787
##
                              Parameter Estimates
       model
              Beta Std. Error
                                 Std. Beta
                                                     0.000
              81.760
                                                            78.040
## (Intercept)
                          1.897
                                            43.108
                                                                    85.479
                                   -0.424 -23.175 0.000 -0.440 -0.371
##
   x15
             -0.405
                          0.017
       y2 0.283
x3 -0.256
##
              0.283
                         0.012
                                   0.356 23.932 0.000 0.260 0.307
##
                         0.024
                                   -0.157 -10.649 0.000 -0.303
                                                                    -0.209
       x19
              0.343
                                                    0.000
                                                            0.275
##
                          0.035
                                    0.158
                                            9.840
                                                                    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
##
                     0.835 RMSE
## R
                                               10.247
## R-Squared
                     0.698
                             Coef. Var
                                               14.945
                             MSE
## Adj. R-Squared
                     0.697
                                              105.002
                     0.694
## Pred R-Squared
                               MAE
                                               7.774
## RMSE: Root Mean Square Error
## MSE: Mean Square Error
```

21

MAE: Mean Absolute Error

#									
:# :#				ANO 	VA 			_	
:# :# :#		Sum of Squares		DF	Mean Square	F	Sig.	_	
# #	Regression Residual Total	447089.897 193622.890 640712.787		6 1844 1850	74514.983 105.002		0.0000	_	
:# :#				Pa	rameter Estima [.]	tes			
#									
:# :#	model	Beta 	Std.	Error	Std. Beta	t 	Sig	lower	upper
	(Intercept) x15	84.269 -0.313		1.893 0.021	-0.328	44.524 -15.061		80.557 -0.354	
#	у2	0.279		0.012	0.350		0.000	0.256	0.301
#	x3	-0.264		0.024	-0.162				
#		0.333		0.034		9.683	0.000		
#		-0.804		0.089	-0.123				
#	x14	-0.084		0.011	-0.142	-7.901	0.000	-0.105	-0.063
# # # # #	R R-Squared		Mode: 0.835		y MSE oef. Var	 10.247 14.945	-		
	Adj. R-Square		0.698		oei. var SE	105.002			
	Pred R-Square	d (0.694	M	AE	7.774			
######	RMSE: Root M MSE: Mean Sq MAE: Mean Ab	uare Error	Error	ANO	VA			_	
#		Sum of Squares		DF	Mean Square	F	Sig.		
# #	Regression Residual Total			6 1844 1850	74514.983 105.002	709.656	0.0000	_	
# # # #					rameter Estima				
#			Std.	Error	Std. Beta	t	Sig	lower	upper
# # #	(Intercept)			1.893 0.021		44.524 -15.061		80.557 -0.354	87.981 -0.272
#	у2	0.279		0.012	0.350	23.867		0.256	0.301

```
xЗ
           -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
                               -0.142 -7.901 0.000 -0.105 -0.063
                       0.011
##
##
##
##
## Stepwise Selection: Step 7
##
## - x5 added
##
##
                  Model Summary
## -----
                  0.845 RMSE
0.715 Coef.
                                         9.952
## R-Squared
                  0.715
                           Coef. Var
                                        14.507
## Adj. R-Squared
                          MSE
                 0.713
                                        99.043
                          MAE
## Pred R-Squared
                  0.711
                                        7.564
## -----
## RMSE: Root Mean Square Error
## MSE: Mean Square Error
## MAE: Mean Absolute Error
##
                         ANOVA
  ______
              Sum of
##
             Squares
                         DF
                             Mean Square
                                         F
                                                 Sig.
                       7 63492.082
## Regression 444444.576
                                        641.058 0.0000
## Residual 177583.362
## Total 622027 938
                       1793
                                99.043
                     1800
## Total
            622027.938
##
##
                          Parameter Estimates
      model
            Beta Std. Error
                             Std. Beta
                                               Sig
                                                     lower upper
 ______
## (Intercept)
            62.229
                                              0.000
                                                           70.363
                       4.147
                                      15.006
                                                    54.096
                              -0.301 -12.310
                                            0.000 -0.336 -0.244
##
  x15
            -0.290
                      0.024
      у2
                     0.012
##
            0.264
                              0.332 22.880 0.000 0.242 0.287
                     0.024
##
       x3 -0.271
                              -0.166 -11.439 0.000 -0.317 -0.224
                     0.034
0.093
                                       9.188 0.000
                                                    0.244
##
      x19
            0.310
                               0.143
                                                           0.377
                               -0.140 -10.234 0.000 -1.135
##
      x11 -0.952
                                                           -0.770
##
                                       -7.416 0.000 -0.103
      x14 -0.081
                      0.011
                               -0.136
                                                           -0.060
       x5 0.000
                     0.000
                              0.109
                                       7.166 0.000 0.000 0.001
##
##
##
                  Model Summary
## --
                 0.845 RMSE
0.715 Coef. Var
0.713 MSE
## R
                                        9.952
## R-Squared
                                       14.507
## Adj. R-Squared
                                        99.043
```

Pred R-Squ	ared	(0.711	M	IAE	7.564			
RMSE: Roc MSE: Mear MAE: Mear	Square	Error					-		
				ANC	DVA				
		 Sum of						-	
	S	quares		DF	Mean Squar	re F	Sig.		
Regression	4444	44.576		7		82 641.058	0.0000	_	
Residual				1793	99.04	13			
Total	6220	27.938 		1800 				_	
					rameter Estim	nates			
mode	el B	eta	Std.	Error	Std. Beta	t	Sig	lower	upp
(Intercept	.) 62			4.147		15.006	0.000	 54.096	70.3
	.5 -0.				-0.301				
У				0.012		22.880			
3				0.024	-0.166	-11.439		-0.317	
x1	.9 0.	310		0.034					
x1	.1 -0.	952		0.093		-10.234		-1.135	
	.4 -0.				-0.136				
X	:5 0.	000 		0.000	0.109	7.166 	0.000	0.000	0.0
Stepwise S	Selection	: Step	8						
200P220 Z	.02000201	. 200р							
- x13 adde	ed								
		I	Model	Summary	7				
 R			 n 8/18	F	 RMSE	9.868	-		
R-Squared					Coef. Var				
Adj. R-Squ	ared				ISE	97.376			
Pred R-Squ	ared	(0.716	M	1AE	7.546			
RMSE: Roc MSE: Mear MAE: Mear	ot Mean S Square	quare l Error	Error				-		
				ANC	DVA				
		Sum of						=	
	S	quares		DF	Mean Squar	re F	Sig.		
Regression Residual	4475	30.042		8	55941.25 97.37	55 574.487 66	7 0.0000	_	

Total		622027.93	3 	1800) 					
				P	aramet	ter Estima	tes			
mode	 el	Beta	Std.	Error	Sto	d. Beta	t	Sig	lower	upper
(Intercept	 t)	55.710		4.272			13.041	0.000	47.331	64.088
x1		-0.196		0.029		-0.203	-6.824	0.000	-0.252	-0.140
7	y2	0.246		0.012		0.310	20.686	0.000	0.223	0.27
X	x3	-0.280		0.024		-0.172	-11.895	0.000	-0.326	-0.23
x1	19	0.332		0.034		0.153	9.847	0.000	0.266	0.398
x1	11	-0.818		0.095		-0.120	-8.578	0.000	-1.005	-0.63
x1	14	-0.089		0.011		-0.149	-8.092	0.000	-0.110	-0.06
×	x5	0.000		0.000		0.132	8.439	0.000	0.000	0.00
x1	13	-0.074		0.013		-0.109	-5.629	0.000	-0.099	-0.04
			Model	Summar	.у					
R			0.848		RMSE		9.868			
R-Squared			0.719		Coef.	Var	14.385			
					MOD		97.376			
Adj. R-Squ			0.718		MSE					
Pred R-Squ	uared		0.718		MAE		7.546			
Pred R-Squ	uared		0.716							
Pred R-Squ RMSE: Roc	uared ot Me	an Square	0.716							
Pred R-Squ 	uared ot Me n Squ	an Square	0.716 Error							
Pred R-Squ 	uared ot Me n Squ	an Square	0.716 Error							
Pred R-Squ 	uared ot Me n Squ	an Square	0.716 Error		MAE 					
Pred R-Squ 	uared ot Me n Squ	an Square	0.716 Error							
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean	uared ot Me n Squ	an Square	0.716 Error	AN	MAE 					
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean	uared ot Me n Squ	an Square are Error olute Err	0.716 Error	AN	MAE	ean Square	7.546	Si	 g. 	
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean	uared	an Square are Error olute Erro Sum of Square 447530.04	0.716 Error	AN DF	MAE OVA Me	55941.255	7.546 F 574.487			
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean Regression Residual	uared	an Square are Error olute Erro Sum of Square 447530.04	0.716 Error or f s	AN DF 8 1792	MAE OVA Me		7.546 F 574.487			
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean MAE: Mean Regression Residual Total	uared	an Square are Error olute Error Sum of Square 447530.04	0.716 Error or f s	AN DF 8 1792 1800	MAE OVA Me	55941.255	7.546 F 574.487			
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean Regression Residual Total	uared	an Square are Error olute Error Sum of Square 447530.04	0.716 Error	AN DF 8 1792 1800	MAE OVA Me OVA	55941.255 97.376	7.546 			
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean Regression Residual Total	uared	an Square are Error olute Error Sum of Square 447530.04	0.716 Error	AN DF 8 1792 1800	MAE OVA Me OVA	55941.255 97.376	7.546 			
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean Regression Residual Total mode	uared ot Me n Squ n Abs	an Square are Error olute Erro Sum of Square 447530.04 174497.89 622027.93	0.716 Error Or f s 	AN DF 1792 1800 F Error	MAE OVA Me Caramet	55941.255 97.376 er Estima	7.546 F 574.487	0.00	 00 	uppe:
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean MAE: Mean Total mode	uared ot Me n Squ n Abs n	an Square are Error olute Erro Sum of Square 447530.04 174497.89 622027.93	0.716 Error or f s S S Std.	AN DF 1792 1800 F Error	MAE OVA Me Paramet	55941.255 97.376 er Estima	7.546 F 574.487	0.00	 00 lower	
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean MAE: Mean Total mode (Intercept	uared ot Me n Squ n Abs n	an Square are Error olute Erro Sum of Square 447530.04 174497.89 622027.93 Beta 55.710	0.716 Error or f s S S Std.	AN 	MAE TOVA Model Paramet	55941.255 97.376 eer Estima	7.546 F 574.487 tes table 13.041	0.000 Sig	 00 lower 47.331	64.08
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean MAE: Mean Regression Residual Total mode (Intercept	uared ot Me n Squ n Abs n	an Square are Error olute Erro Sum of Square 447530.04 174497.89 622027.93 Beta 55.710 -0.196	0.716 Error or f s 3 3 	AN DF 1792 1800 F Error 4.272 0.029	MAE OVA Me Paramet	55941.255 97.376 eer Estima	7.546 F 574.487 tes t 13.041 -6.824	0.000 Sig 0.000 0.000	lower 47.331 -0.252	64.08 -0.14
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean MAE: Mean Regression Residual Total mode (Intercept x1	uared ot Me n Squ n Abs n el t)	an Square are Error olute Error sum of Square 447530.04174497.89622027.93888888888888888888888888888888888888	0.716 Error	AN DF 1792 1800 F Error 4.272 0.029 0.012	MAE OVA Me Caramet	55941.255 97.376 cer Estima 1. Beta -0.203 0.310	7.546 F 574.487 tes 13.041 -6.824 20.686	0.000 Sig 0.000 0.000 0.000	lower	64.08 -0.14 0.27
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean MAE: Mean Regression Residual Total mode (Intercept x1	uared ot Me n Squ n Abs n t) 15 y2 x3	an Square are Error olute Error olute Error Sum of Square 447530.04174497.89622027.936	0.716 Error	AND DF 8 1792 1800 F Error 4.272 0.029 0.012 0.024	MAE OVA Me Caramet	55941.255 97.376 er Estima d. Beta -0.203 0.310 -0.172	7.546 F 574.487 tes 13.041 -6.824 20.686 -11.895	0.000 Sig 0.000 0.000 0.000 0.000	lower	64.08 -0.14 0.27 -0.23
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean MAE: Mean Regression Residual Total mode (Intercept	uared ot Me n Squ n Abs n el t) 15 y2 x3 19	an Square are Error olute Error olute Error Sum of Square 447530.04 174497.89 622027.93 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	0.716 Error	AN DF 1792 1800 F Error 4.272 0.029 0.012 0.024 0.034	MAE OVA Me Paramet	55941.255 97.376 er Estima -0.203 0.310 -0.172 0.153	7.546 F 574.487 tes 13.041 -6.824 20.686 -11.895 9.847	0.000 Sig 0.000 0.000 0.000 0.000 0.000 0.000	lower	64.088 -0.140 0.270 -0.234 0.398
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean MAE: Mean Regression Residual Total mode (Intercept x1 yx x1 x1	uared pot Me n Squ n Abs n t) 15 y2 x3 19 11	an Square are Error olute Error olute Error sum of Square 447530.04 174497.89 622027.93 622027.7	0.716 Error	AN DF 1792 1800 F Error 4.272 0.029 0.012 0.024 0.034 0.095	MAE OVA Me Paramet	55941.255 97.376 Ser Estima -0.203 0.310 -0.172 0.153 -0.120	7.546 F 574.487 tes 13.041 -6.824 20.686 -11.895 9.847 -8.578	0.000 Sig 0.000 0.000 0.000 0.000 0.000 0.000 0.000	lower 47.331 -0.252 0.223 -0.326 0.266 -1.005	64.083 -0.144 0.276 -0.23- 0.393 -0.63
Pred R-Squ RMSE: Roo MSE: Mean MAE: Mean MAE: Mean Regression Residual Total mode (Intercept x1 x1 x1	uared ot Me n Squ n Abs n t) 15 y2 x3 19 11 14	an Square are Error olute Error olute Error Sum of Square 447530.04 174497.89 622027.93 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8 8	0.716 Error	AN DF 1792 1800 F Error 4.272 0.029 0.012 0.024 0.034	MAE OVA Me Carameter Sto	55941.255 97.376 er Estima -0.203 0.310 -0.172 0.153	7.546 F 574.487 tes t 13.041 -6.824 20.686 -11.895 9.847 -8.578 -8.092	0.000 Sig 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000	lower	-0.140

```
##
##
##
## Stepwise Selection: Step 9
## - x7 added
##
                               Model Summary
## ----
                                           RMSE
## R
                               0.849
                                                                     9.830
## R-Squared
                               0.720
                                             Coef. Var
                                                                   14.336
## Adj. R-Squared
                               0.719
                                             MSE
                                                                    96.621
                               0.716
## Pred R-Squared
                                                                     7.526
                                             MAE
    RMSE: Root Mean Square Error
## MSE: Mean Square Error
   MAE: Mean Absolute Error
##
##
                                           ANOVA
##
                        Sum of
##
                                    DF Mean Square F Sig.
                       Squares
                                     -----
## -----
                                     9
## Regression 427996.762
                                                   47555.196 492.181 0.0000
                                       1721
## Residual 166285.265
                                                      96.621
                 594282.027
## Total
                                       1730
##
                                            Parameter Estimates
                                                 Std. Beta
                                                                               Sig
         model
                       Beta
                                Std. Error
                                                                   t
                                                                                          lower
                    50.169
                                                                 11.353
                                                                              0.000
                                                                                        41.502
                                                                                                     58.836
## (Intercept)
                                       4.419

      0.029
      -0.206
      -6.783
      0.000
      -0.255
      -0.141

      0.012
      0.295
      19.088
      0.000
      0.209
      0.257

      0.024
      -0.159
      -10.679
      0.000
      -0.307
      -0.212

      0.036
      0.150
      9.462
      0.000
      0.268
      0.408

      0.099
      -0.102
      -7.153
      0.000
      -0.901
      -0.513

      0.011
      -0.155
      -8.184
      0.000
      -0.114
      -0.070

##
    x15
                  -0.198
          у2
##
                    0.233
##
            x3 -0.259
##
         x19
                    0.338
          x11 -0.707
##
          x14 -0.092
##
##
                    0.001
                                    0.000
                                                    0.155
                                                                 9.636 0.000
                                                                                        0.000
            x5
                                                                                                    0.001
           x13 -0.066
##
                                     0.013
                                                    -0.098
                                                                 -4.905 0.000 -0.092
                                                                                                    -0.039
             x7
                                                                                                    -0.615
                    -1.026
                                     0.209
                                                     -0.071
                                                                 -4.902 0.000 -1.436
##
##
##
                               Model Summary
                      0.849 RMSE
0.720 Coef.
0.719 MSE
0.716 MAE
## R
                                                                     9.830
## R-Squared
                                             Coef. Var
                                                                    14.336
## Adj. R-Squared
                                                                    96.621
## Pred R-Squared
                                                                    7.526
```

RMSE: Root Mean Square Error ## MSE: Mean Square Error ## MAE: Mean Absolute Error ## ANOVA ## -----Sum of ## Squares DF Mean Square ______ 9 47555.196 492.181 0.0000 ## Regression 427996.762 ## Residual 166285.265 1721 96.621 594282.027 1730 ## Total ## Parameter Estimates Std. Beta ${\tt model}$ Beta Std. Error t Sig lower 11.353 0.000 41.502 ## (Intercept) 4.419 58.836 50.169 0.029 0.012 -0.206 -6.783 0.000 -0.255 x15 -0.198 -0.141 0.295 19.088 0.000 0.209 0.257 ## у2 0.233
 0.024
 -0.159
 -10.679
 0.000
 -0.307
 -0.212

 0.036
 0.150
 9.462
 0.000
 0.268
 0.408

 0.099
 -0.102
 -7.153
 0.000
 -0.901
 -0.513

 0.011
 -0.155
 -8.184
 0.000
 -0.114
 -0.070
 ## x3 -0.259 0.338 ## x19 x11 -0.707 ## ## x14 -0.092 0.001 x5 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 ## -----RMSE Coef. Var ## R 0.850 9.739 ## R-Squared 0.722 14.130 ## Adj. R-Squared 0.720 MSE 94.841 ## Pred R-Squared 0.717 MAE7.462 ## -----## RMSE: Root Mean Square Error ## MSE: Mean Square Error ## MAE: Mean Absolute Error ## ## ANOVA ## Sum of ## DF Mean Square F Squares ## -----10 39926.112 420.981 0.0000 ## Regression 399261.115 1623 ## Residual 153926.516 94.841

##	Total	553187.632	2	1633					
##								_	
## ##				P	arameter Estimat	ces			
## ##	model	Beta	 Std.	Error	Std. Beta	 t	 Sig	lower	upper
##									
##	(Intercept)	64.853		5.452	0.404	11.895		54.159	75.547
##	x15	-0.190		0.030	-0.194	-6.228		-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		-0.307	-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		-1.564	-0.926
##	x14	-0.088		0.012	-0.148	-7.556		-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		-0.085	-0.030
##	x7	-0.914		0.214	-0.063	-4.263		-1.334	-0.493
##	x17	-0.001		0.000	-0.094	-4.276	0.000	-0.002	-0.001
##									
##									
##									
##									
##			Model	Summar	У				
##									
##	R		0.850		RMSE	9.739			
	R-Squared		0.722		Coef. Var	14.130			
	Adj. R-Square		0.720		MSE	94.841			
	Pred R-Square		0.717]	MAE	7.462			
## ##	RMSE: Root M		Error						
##	MSE: Mean Sq	_	EIIOI						
##	MAE: Mean Ab		or.						
##	TIAL: TICALI AD	BOILUC LII	71						
##				AN	DVA				
##								_	
##		Sum of		DE	Mana Carra	F	G:		
		Squares		DF	Mean Square	F 	Sig.	_	
			5	10	39926.112	420.981	0.0000	1	
	Residual								
11 H	Iotal	553187.632	2	1000					
	Total	553187.632						_	
## ##								_	
## ## ##				P	arameter Estimat	ces			
## ## ## ##	model	 Beta	Std.	P:	arameter Estimat		Sig	- lower	 upper
## ## ## ## ##	model	Beta	Std.	P. Error	arameter Estimat	t			
## ## ## ## ##	model (Intercept)	Beta 64.853	Std.	Pro	arameter Estimat	t 11.895	0.000	 54.159	75.547
## ## ## ## ## ##	model (Intercept) x15	Beta 64.853 -0.190	Std.	Error 5.452 0.030	arameter Estimates Std. Beta	t 11.895 -6.228	0.000 0.000	 54.159 -0.249	75.547 -0.130
## ## ## ## ## ##	model (Intercept) x15 y2	Beta 64.853 -0.190 0.224	Std.	Error 5.452 0.030 0.013	Std. Beta -0.194 0.283	t 11.895 -6.228 17.694	0.000 0.000 0.000	54.159 -0.249 0.199	75.547 -0.130 0.249
## ## ## ## ## ## ##	model (Intercept) x15 y2 x3	Beta64.853 -0.190 0.224 -0.258	Std.	Error 5.452 0.030 0.013 0.025	Std. Beta -0.194 0.283 -0.154	t 11.895 -6.228 17.694 -10.257	0.000 0.000 0.000 0.000	54.159 -0.249 0.199 -0.307	75.547 -0.130 0.249 -0.209
## ## ## ## ## ## ##	model (Intercept) x15 y2 x3 x19	Beta 64.853 -0.190 0.224 -0.258 0.353	Std.	Error 5.452 0.030 0.013 0.025 0.036	Std. Beta -0.194 0.283 -0.154 0.159	t 11.895 -6.228 17.694 -10.257 9.754	0.000 0.000 0.000 0.000 0.000	54.159 -0.249 0.199 -0.307 0.282	75.547 -0.130 0.249 -0.209 0.424
## ## ## ## ## ## ##	model (Intercept) x15 y2 x3	Beta64.853 -0.190 0.224 -0.258	Std.	Error 5.452 0.030 0.013 0.025	-0.194 0.283 -0.154 0.159 -0.172	t 11.895 -6.228 17.694 -10.257 9.754 -7.659	0.000 0.000 0.000 0.000 0.000 0.000	54.159 -0.249 0.199 -0.307	75.547 -0.130 0.249 -0.209 0.424 -0.926

```
0.175
                                               0.000
##
       x5
             0.001
                    0.000
                                        9.604
                                                       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
## -----
                   0.856 RMSE
0.732 Coef.
                                           9.596
## R-Squared
                    0.732
                            Coef. Var
                                           13.931
## Adj. R-Squared
                           MSE
                   0.730
                                          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.
                                 35071.115
## Regression 385782.260
                        11
                                           380.853 0.0000
## Residual 141259.613
                        1534
                                  92.086
                      1545
## Total
            527041.874
##
##
                            Parameter Estimates
      model
             Beta
                    Std. Error
                               Std. Beta
                                                 Sig
                                                        lower upper
## (Intercept)
                                                 0.000
                                                               79.637
             68.403
                        5.728
                                         11.943
                                                       57.168
                                       -6.574
                                                      -0.269
##
  x15
             -0.207
                       0.032
                                 -0.211
                                                0.000
                                                              -0.145
##
             0.230
                                0.289
                                        17.665 0.000 0.204
       у2
                       0.013
                                                             0.256
##
            -0.275
                       0.026
                                -0.165
                                       -10.774 0.000 -0.325
                                                              -0.225
       xЗ
##
       x19
             0.361
                       0.037
                                0.161
                                         9.803 0.000
                                                       0.289
                                                               0.434
##
             -1.304
                                         -7.727 0.000 -1.635
       x11
                       0.169
                                 -0.179
                                                              -0.973
##
                                -0.152 -7.666 0.000 -0.115
       x14
            -0.092
                       0.012
                                                              -0.068
##
                       0.000
                                        8.799 0.000
                                                       0.000
       x5
             0.001
                                0.164
                                                              0.001
                       0.015
                                -0.063
                                         -2.927
                                                      -0.073
##
        x13
             -0.044
                                                0.003
                                                              -0.014
            -0.881
                                                              -0.453
##
       x7
                       0.218
                                 -0.060
                                        -4.037 0.000 -1.308
##
                                -0.090
                                         -3.946
       x17
            -0.001
                       0.000
                                                0.000 -0.002
                                                              -0.001
##
        x20 -1.612
                        0.660
                                 -0.037
                                         -2.441
                                                0.015
                                                       -2.908
                                                              -0.317
##
##
##
##
##
                    Model Summary
```

29

```
## R
               0.856 RMSE
0.732 Coef. Var
                                   9.596
## R-Squared
                                   13.931
## Adj. R-Squared
                0.730
                      MSE
                                   92.086
               0.727
## Pred R-Squared
                        MAE
## -----
## RMSE: Root Mean Square Error
## MSE: Mean Square Error
## MAE: Mean Absolute Error
##
##
                      ANOVA
## -----
            Sum of
                    DF Mean Square F Sig.
            Squares
```

385782.260

11 35071.115 380.853 0.0000 1534 92.086 1545 ## Residual 141259.613 ## Total 527041.874

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
##	у2	0.230	0.013	0.289	17.665	0.000	0.204	0.256
##	хЗ	-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

Regression

##

##		Model Summa	ry	
##				
##	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

##											
##				Al	AVOI						
##										_	
##		Sum of Squares		DH	7 M2	ean Squar	e F		Sig.		
##										_	
##	Regression	367021.341		12	2	30585.11	2 331.1	48	0.0000		
	-	135216.130		1464	1	92.36					
##	Total	502237.471		1476	3						
##										_	
##				_							
##					aramet	ter Estim	ates 				
##	model	Beta	Std.	Error	Sto	i. Beta	t	S	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			-0.285	-0.156
##	у2	0.227		0.013		0.287				0.201	0.253
##	х3	-0.279		0.027		-0.167				-0.332	-0.227
##	x19	0.342		0.038		0.150	9.001			0.267	0.416
##	x11	-1.247		0.177		-0.164	-7.061			-1.593	-0.901
##	x14	-0.090		0.012		-0.147				-0.114	-0.066
##	x5	0.001		0.000			8.690			0.001	0.001
##	x13	-0.037		0.015		-0.055	-2.461			-0.067	-0.008
##	x7 x17	-0.813		0.246		-0.051	-3.302 -4.193			-1.296	-0.330
##		-0.001 -1.564		0.000		-0.099 -0.036				-0.002 -2.871	-0.001 -0.258
##	x10	-0.002		0.001		-0.037				-0.005	0.000
##											
##											
##											
##											
##			Model	Summan	ĵу						
##											
##			0.855		RMSE	17	9.61				
	R-Squared		0.731		Coef.	var	14.00				
	Adj. R-Square Pred R-Square		0.729 0.725		MSE MAE		92.36 7.33				
##							۰۰.۵۰				
##	RMSE: Root M										
##		_									
##	MAE: Mean Ab		r								
##											
##					AVOI						
										_	
##		Sum of		זמ	г м.	oon Sauor	o E		Sia		
		Squares		וע			e F 		nıR.	_	
	Regression										
	Residual										
		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
##	-	-0.220		0.033	-0.223	-6.726		-0.285	-0.156
##		0.227		0.013	0.287	17.065		0.201	0.253
 ‡#	x3	-0.279		0.027	-0.167	-10.505		-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		-1.593	-0.901
#	x14	-0.090		0.012	-0.147	-7.334		-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		-0.067	-0.008
#	x7	-0.813		0.246	-0.051	-3.302		-1.296	-0.330
#	x17	-0.001		0.000	-0.099	-4.183		-0.002	-0.001
#	x20	-1.564		0.666	-0.036			-2.871	-0.258
#	x10	-0.002		0.000	-0.037			-0.005	0.230
#	X10	0.002		0.001	0.037	1.071	0.002	0.005	0.000
#									
:#									
#									
	C+		- 12						
	Stepwise Sele	ection: Ste	p 13						
#	0 11 1								
	- x9 added								
#			Model	Summary					
#	R		0.856	 R	MSE	9.585			
#	R-Squared		0.733	C	oef. Var	13.948			
	Adj. R-Square		0.730	M	SE	91.879			
ш	Pred R-Square	•	0 700		. —	7 000			
	_		0.726	М	AE	7.308			
# # # #	RMSE: Root M	 Mean Square Nuare Error	Error	 	AE 	7.308			
# # # # #	RMSE: Root M	 Mean Square Nuare Error	Error	 ANO		7.308			
#######	RMSE: Root M	 Mean Square Nuare Error	Error			7.308 		-	
######	RMSE: Root M	 Mean Square Nuare Error	Error			7.308		_	
########	RMSE: Root M	Mean Square quare Error osolute Err	Error or				 Sig.	-	
########	RMSE: Root M MSE: Mean So MAE: Mean Ab	Mean Square quare Error psolute Err Sum o	Error or f	ANO	VA 		Sig.	-	
##########	RMSE: Root M MSE: Mean So MAE: Mean Ab	Mean Square quare Error psolute Err Sum o Square	Error or f	ANO	VA 	F		-	
##########	RMSE: Root M MSE: Mean So MAE: Mean Ab	Mean Square quare Error osolute Err Sum o Square	Error or f s8	ANO DF	VA Mean Square	F		-	
###########	RMSE: Root M MSE: Mean So MAE: Mean Ab	Mean Square quare Error psolute Err Sum o Square 364539.72	Error or f s 8	ANO DF 13	VA Mean Square 	F		-	
###########	RMSE: Root M MSE: Mean Ab MAE: Mean Ab	Mean Square quare Error psolute Err Sum o Square 364539.72 133040.48	Error or f s 8	ANO DF 13 1448	VA Mean Square 	F		-	
#############	RMSE: Root M MSE: Mean Ab MAE: Mean Ab	Mean Square quare Error psolute Err Sum o Square 364539.72 133040.48	Error or f s 8	ANO DF 13 1448	VA Mean Square 	F		-	
##############	RMSE: Root M MSE: Mean Ab MAE: Mean Ab	Mean Square quare Error psolute Err Sum o Square 364539.72 133040.48	Error or f s 8	ANO DF 13 1448 1461	VA Mean Square 	F 305.201		-	
##############	RMSE: Root M MSE: Mean Ab MAE: Mean Ab	Mean Square quare Error psolute Err Sum o Square 364539.72 133040.48	Error or f s 8	ANO DF 13 1448 1461	VA Mean Square 	F 305.201		-	
################	RMSE: Root M MSE: Mean Ab MAE: Mean Ab	Mean Square quare Error psolute Err Sum o Square 364539.72 133040.48	Error or f s 8 1 8	ANO DF 13 1448 1461	VA Mean Square 	F 305.201		- lower	uppe:
##################	RMSE: Root N MSE: Mean Ab MAE: Mean Ab Regression Residual Total	Mean Square quare Error psolute Err Sum o Square 364539.72 133040.48 497580.20	Error or f s 8 1 8	ANO DF 13 1448 1461 Pa	Mean Square 28041.518 91.879 rameter Estima	F 305.201	0.0000	lower	
##################	RMSE: Root M MSE: Mean So MAE: Mean Ab Regression Residual Total model	Mean Square quare Error psolute Err Sum o Square 364539.72 133040.48 497580.20	Error or f s 8 1 8	ANO DF 13 1448 1461 Pa Error	Mean Square 28041.518 91.879 rameter Estima	F 305.201	0.0000 Sig		80.14
###################	RMSE: Root M MSE: Mean So MAE: Mean Ab Regression Residual Total model (Intercept)	Mean Square quare Error psolute Err Sum o Square 364539.72 133040.48 497580.20 Beta 68.504	Error or f s 8 1 8	ANO DF 13 1448 1461 Pa Error 5.933	Mean Square 28041.518 91.879 rameter Estimates Std. Beta	F 305.201	0.0000 Sig	56.866	80.14: -0.138
############	RMSE: Root M MSE: Mean Ab MAE: Mean Ab Regression Residual Total model (Intercept) x15	Sum o Square 364539.72 133040.48 497580.20 Beta 68.504 -0.204	Error or f s 8 1 8	DF 13 1448 1461 Pa Error 5.933 0.034	Mean Square 28041.518 91.879 rameter Estima Std. Beta -0.206	F 305.201 tes t	0.0000 Sig 0.000 0.000	56.866 -0.269	upper 80.141 -0.138 0.257 -0.233

##	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 # -----

ar 13.948	
10.010	
91.879	
7.308	
	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

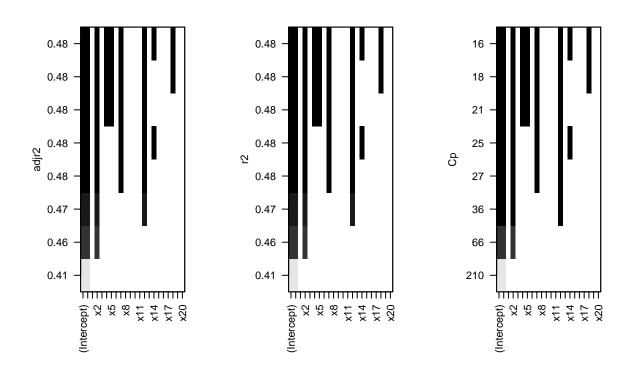
##								
##	model	Beta	Std. Error	Std. Beta	t	Sig	lower	upper
##								
##	(Intercept)	68.504	5.933		11.547	0.000	56.866	80.141
##	x15	-0.204	0.034	-0.206	-6.066	0.000	-0.269	-0.138
##	у2	0.231	0.013	0.291	17.184	0.000	0.204	0.257
##	x3	-0.285	0.027	-0.171	-10.703	0.000	-0.338	-0.233
##	x19	0.351	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
##	х9	-6.263	3.318	-0.032	-1.888	0.059	-12.772	0.246

##

##									
##	No more varia	ahlas to ha :	added.	/removed					
##	NO MOTE VALLE	ibles to be	added	removed.					
##									
	Final Model (-							
## ##]	Model	Summary					
##				RMS		9.585			
	R-Squared				ef. Var				
	Adj. R-Square					91.879			
##			J. 126	MA: 	E 	1.300			
	RMSE: Root M								
	MSE: Mean So	_							
##	MAE: Mean Ab	solute Erro	r						
##									
##				ANOV	A 				
##		Sum of						-	
##				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					
								-	
##				Dar	ameter Estimat	- 05			
##									
##	model	Beta	Std.	Error	Std. Beta	t	Sig	lower	upper
	(Intercept)	68.504		5.933	-0.206 0.291 -0.171	11.547	0.000	56.866	80.141
##	X12	-0.204		0.034	-0.206	-6.066 17 10/	0.000	0.269	-0.138
##	y2 x3	0.231 -0.285		0.013	-0.291	-10 703	0.000	-0.338	-0.237
##	x19	0.351		0.038	0.171	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
11									
##									
##				Stepwi	se Selection S	Summary			
##									
##		Add			Adj.				
##	Step Varia	able Remo	ved	R-Squa:	re R-Square	c(p)		AIC	RMSE

##								
##	1	x15	addition	0.492	0.492	2435.3480	16753.3141	14.1594
##	2	у2	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)

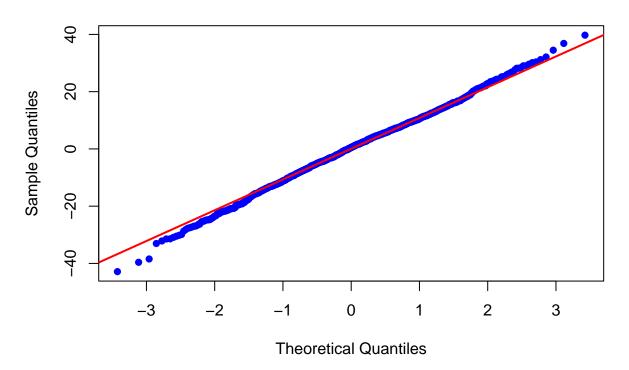
2019

2021

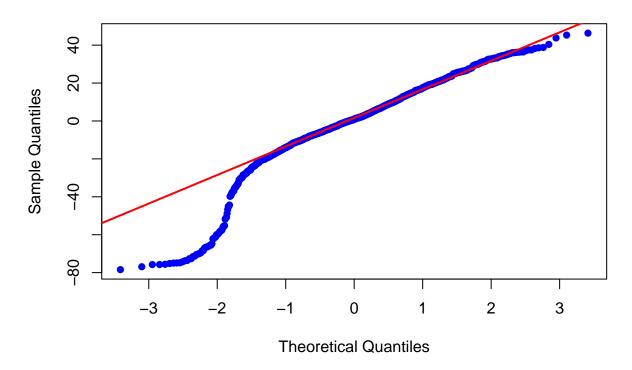
f. Normality Check

```
##
## 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</pre>
```

2019 Model



2021 Model



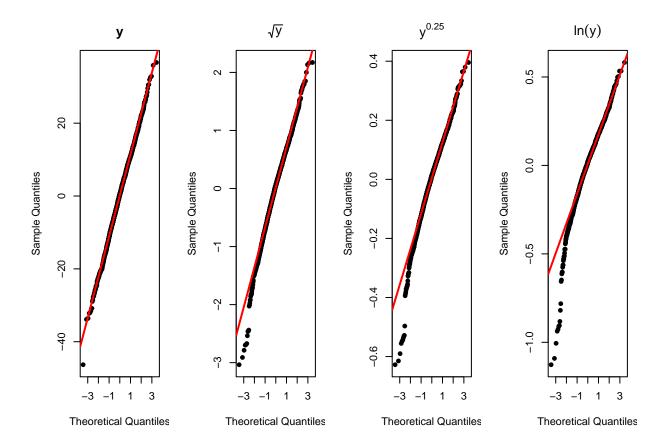
g. Transformation

2019

```
##
## Call:
## lm(formula = y1 ~ x1 + x2 + x5 + x7 + x12 + x14 + x17 + x19,
       data = data_numeric)
##
##
## Residuals:
##
      Min
                1Q Median
                                ЗQ
                                       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 ***
                4.374e-01 2.554e-02 17.125 < 2e-16 ***
## x1
## 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 ***
               1.403e-01 1.362e-02 10.302
## x12
                                             < 2e-16 ***
## x14
               -1.561e-01 1.108e-02 -14.087
                                              < 2e-16 ***
                                       0.320
               5.903e-05 1.846e-04
                                                0.749
## x17
## x19
               5.076e-01 4.392e-02 11.559 < 2e-16 ***
## ---
```

```
## Signif. codes: 0 '***' 0.001 '**' 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 ***
               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 ***
               9.002e-03 8.704e-04 10.342 < 2e-16 ***
## x12
              -9.254e-03 7.083e-04 -13.065 < 2e-16 ***
## x14
## 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:
                 1Q
                     Median
                                           Max
       Min
                                   30
## -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
                                    14.783 < 2e-16 ***
               1.132e-05 7.658e-07
## x7
              -1.777e-02 2.829e-03 -6.283 4.24e-10 ***
## x12
               1.621e-03 1.586e-04 10.220
                                            < 2e-16 ***
              -1.601e-03 1.290e-04 -12.405 < 2e-16 ***
## x14
              2.480e-07 2.149e-06
## x17
                                     0.115
                                               0.908
```

```
5.373e-03 5.113e-04 10.508 < 2e-16 ***
## x19
## ---
## Signif. codes: 0 '***' 0.001 '**' 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 ***
              -2.223e-03 1.909e-04 -11.646 < 2e-16 ***
## x14
## 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
                            x7
                                    x12
                                              x14
                                                       x16
                                                                x17
                                                                          x20
         x1
                  x5
## 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

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

2. Race

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

```
\begin{aligned} 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 \end{aligned}
```

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

[1] 63.89831

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

[1] 48.41171

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

[1] 35.32944

mean(rural\$White.Percentage)

[1] 50.33252

```
mean(urban$Black.Percentage)
## [1] 38.83198
mean(rural$Black.Percentage)
## [1] 32.6066
3. Economy
# 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$All.Students.Math.Achievement))
## [1] 54.89206 58.82841
# 2021
c(mean(Econ_Dia_100_urban$X2021.All.Students.Math.Achievement),
  mean(Econ_Dia_100_rural$X2021.All.Students.Math.Achievement))
## [1] 36.54889 47.08022
                           H_0: \mu_{Rural\ EconDis} - \mu_{Urban\ EconDis} = 0
                           H_0: \mu_{Rural\ EconDis} - \mu_{Urban\ EconDis} > 15
                           p-value = 0.04061 < \alpha = 0.05 \rightarrow Reject H_0
mean(urban$Econ.Disadvantaged.Percentage)
## [1] 65.87646
mean(rural$Econ.Disadvantaged.Percentage)
## [1] 83.43863
t.test(rural$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
```

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

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	-04.43	0
x18	x19	-0.425	0
0	19	0.494	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
	(0.00007)	(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		0.0000003
		(0.0000003)
x17		-0.0004
		(0.0003)
Num.Obs.	1628	1561
R2	0.638	0.355
R2 Adj.	0.636	0.352
AIC	12515.8	13646.9
BIC	12569.7	13700.4
Log.Lik.	-6247.892	-6813.438
F		106.702
RMSE	11.23	19.03