An Econometric Analysis of Housing Demand

Motivation

A common concept within the field of economics is elasticity (i.e., the idea that price for goods/services will generally change in context of the quantity of an item available within a market) (Hayes, 2021). Housing is one of many commonly mentioned market goods in both macro and microeconomics, and housing is often subject to price fluctuations within the market. However, housing units are also vastly variable in quality (from studio apartments to mansions), so analysis of housing markets generally requires *ceteris paribus* conditions to analyze. Thankfully, the U.S. Department of Housing and Urban Development (HUD) has an Office of Policy Development & Research that provides publicly available data on the 50th percentile/median market rate for housing (specifically for studio and one-four-bedroom housing units) in every area within the United States, including territories such as Puerto Rico and Guam. (US Housing and Urban Development - Office of Policy Development & Research, 2021) As medians are centers of mean generally less prone to skews and bias, my intent is to see whether the basic concept of elasticity holds well for housing across the United States, with a secondary goal of using linear regression to establish a general idea of exactly *how* elastic (or inelastic) the demand for housing is, assuming median housing quality is present for the median market rate present in the 2020 housing data from HUD.

Data Description

As referenced in the above motivation statement, HUD has publicly available data for 50th percentile (median) rents for the entire United States as far back as 2001. The number of areas within the dataset is defined as n= 4766, accounting for several (if not all) counties, parishes, and municipalities within all states and territories in the United States. For the sake of using the most recent full calendar year, I am using the dataset for 2020 — as 2021 is not entirely over at the time of this project.

The dataset is an MS Excel spreadsheet with the following

- fips2010
- rent50_0
- rent50_1
- rent50_2
- rent50_3
- rent50_4
- state
- cbsasub21
- areaname21
- county
- cousub cntyname

- name
- pop2017
- hu2017
- state_alpha

For the purposes of SAS coding for this simple regression project, I have used “rent50_0”, “rent50_1”, “rent50_2”, “rent50_3”, “rent50_4” as potential predictor variables, as well as pop2017, hu2017, and transformations of pop2017 and hu2017 as potential predictor variables. I will specifically refer to “rent50_2” as my response variable in this written report from here on out, as two-bedroom housing units are often used as a benchmark for what constitutes affordable housing (Adamczyk, 2021), (Simone, 2021). Additionally, “rent50_2” serves as the median number of bedrooms within the range of rent prices, giving another good reason to use it as a metric.

The other variables outside of the assorted rent variables, population variable, and housing unit variable are all categorical variables that entail things such as metropolitan statistical areas, counties, states/territories, and other assorted government codes that further detail each region.

Data Exploration

I used the below SAS code snippet to import a slightly cleaned up version* of the HUD data referenced earlier for this paper:

```
FILENAME CSV "/home/u49665201/sasuser.v94/STA3064/Copy 2 of
FY2021_50_County.csv" TERMSTR=CRLF;

/** Import the CSV file.  **/

PROC IMPORT DATAFILE=CSV

            OUT=HUDdata

            DBMS=CSV

            REPLACE;

RUN;

/** Print the results. **/

PROC PRINT DATA=HUDdata;

RUN;
```

From this point, I created a secondary dataset I called "Housing" with a few additional changes.

```
data Housing;
 set HUDdata;
 PtoH = pop2017/hu2017;
 HtoP = hu2017/pop2017;
 transPop = pop2017**-1;
 transHu = hu2017**-1;
run;
```

In this instance, the most classic example of elasticity would be present in the model "rent50_2=hu2017" to mimic a classical "price=quantity" scenario. However, I made some additions for additional experimentation and exploration of this data.

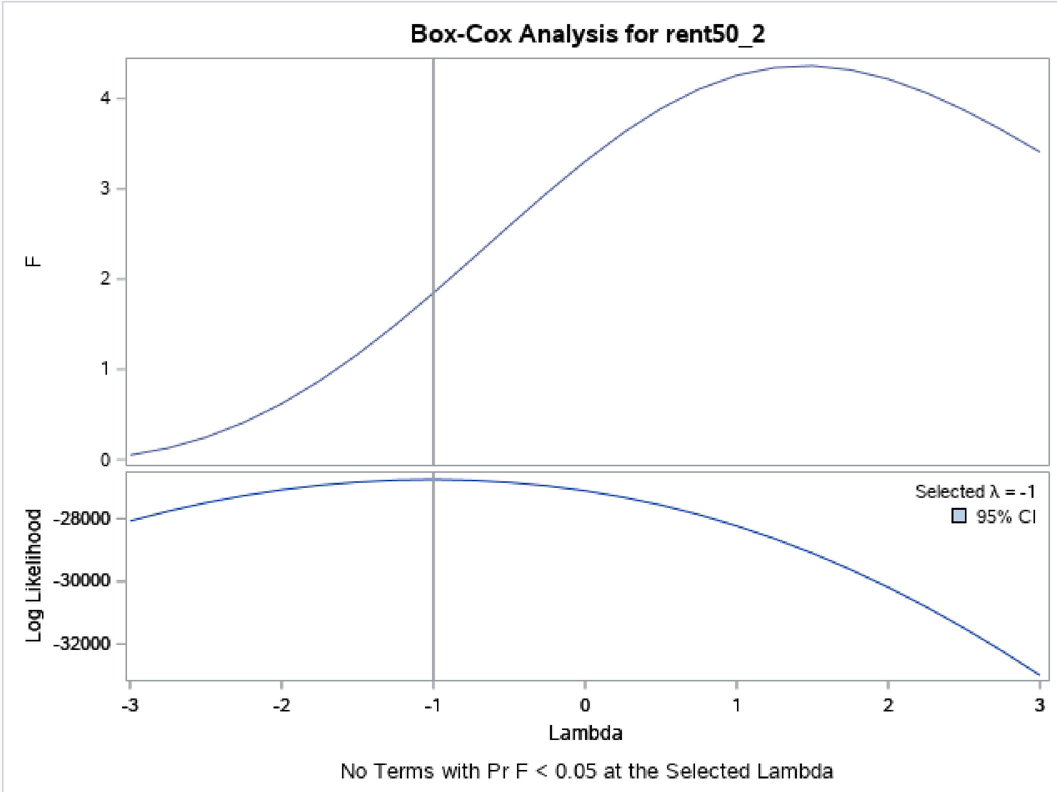
The first addition I made was creating population: housing (defined as 'PtoH') and housing: population (defined as 'HtoP') ratios from pop2017 and hu2017. While these ratios would not necessarily create the scenario for classic elasticity, my initial assumption was that they could serve as proxy variables for housing scarcity in lieu of being able to use multiple regression for this project.

The second addition I made was adding transformations to the variables pop2017 and hu2017, based on the recommended Box-Cox transformation recommended by the output generated by the below code:

```
proc transreg data=Housing;
model boxcox(rent50_2)=identity(HtoP);
run;
proc transreg data=Housing;
model boxcox(rent50_2)=identity(PtoH);
run;
proc transreg data=HUDdata;
model boxcox(rent50_2)=identity(pop2017);
run;
proc transreg data=HUDdata;
model boxcox(rent50_2)=identity(hu2017);
run;
```

STA 3064 PROJECT A

The TRANSREG Procedure

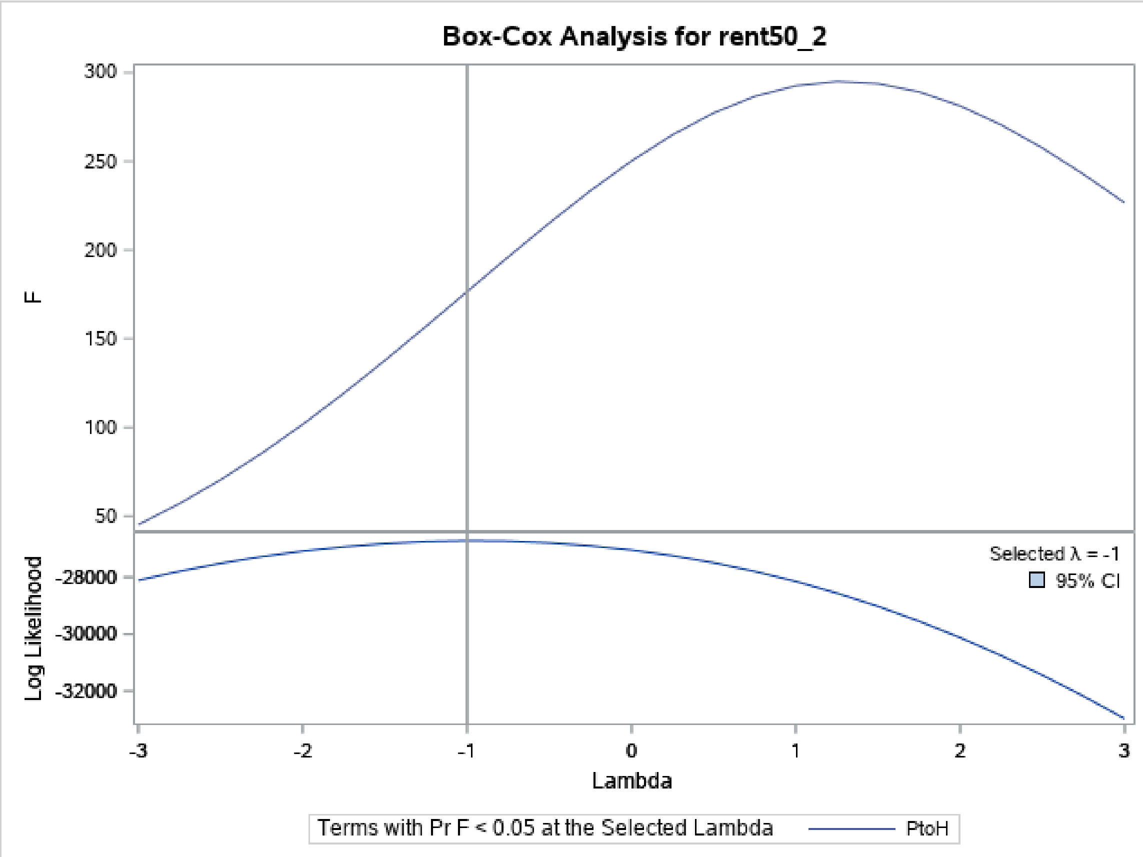


The TRANSREG Procedure

TRANSREG Univariate Algorithm Iteration History for BoxCox(rent50_2)					
Iteration Number	Average Change	Maximum Change	R-Square	Criterion Change	Note
1	0.00000	0.00000	0.00039		Converged

Algorithm converged.

The TRANSREG Procedure



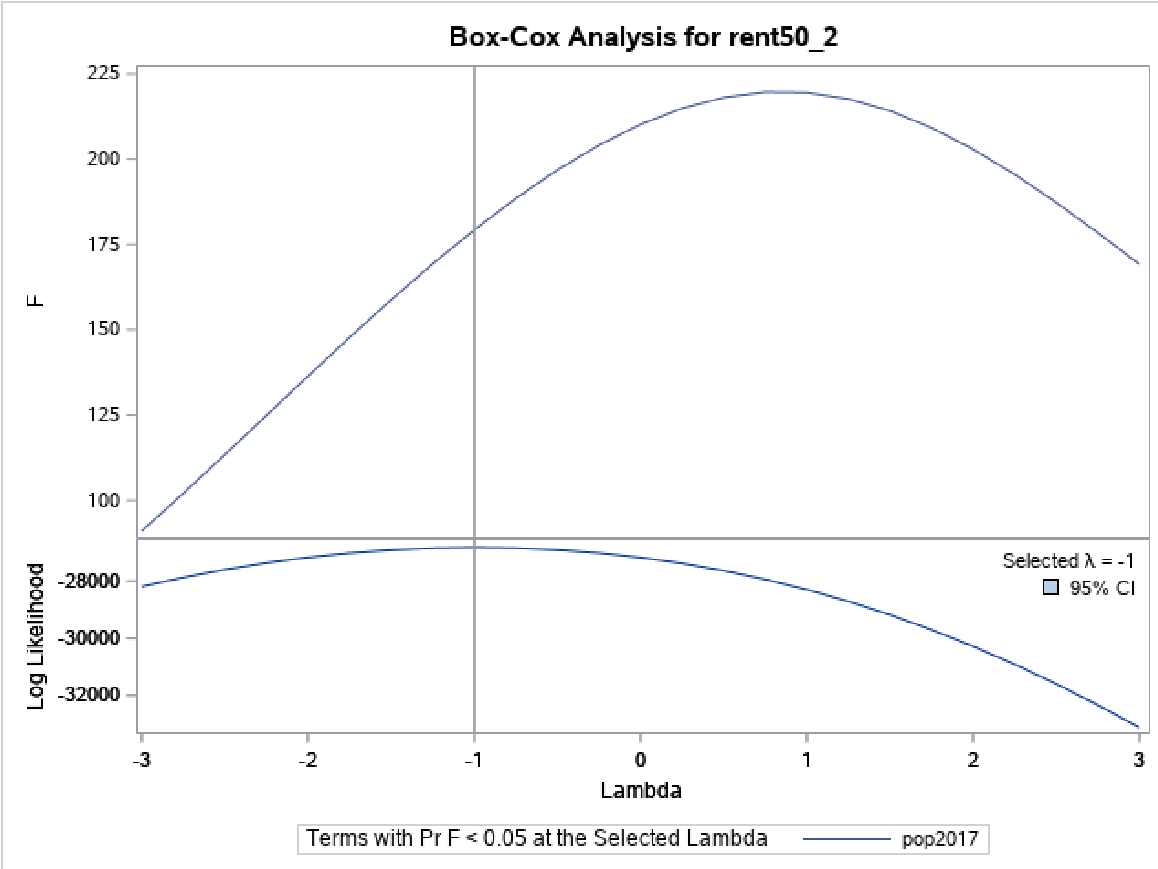
The TRANSREG Procedure

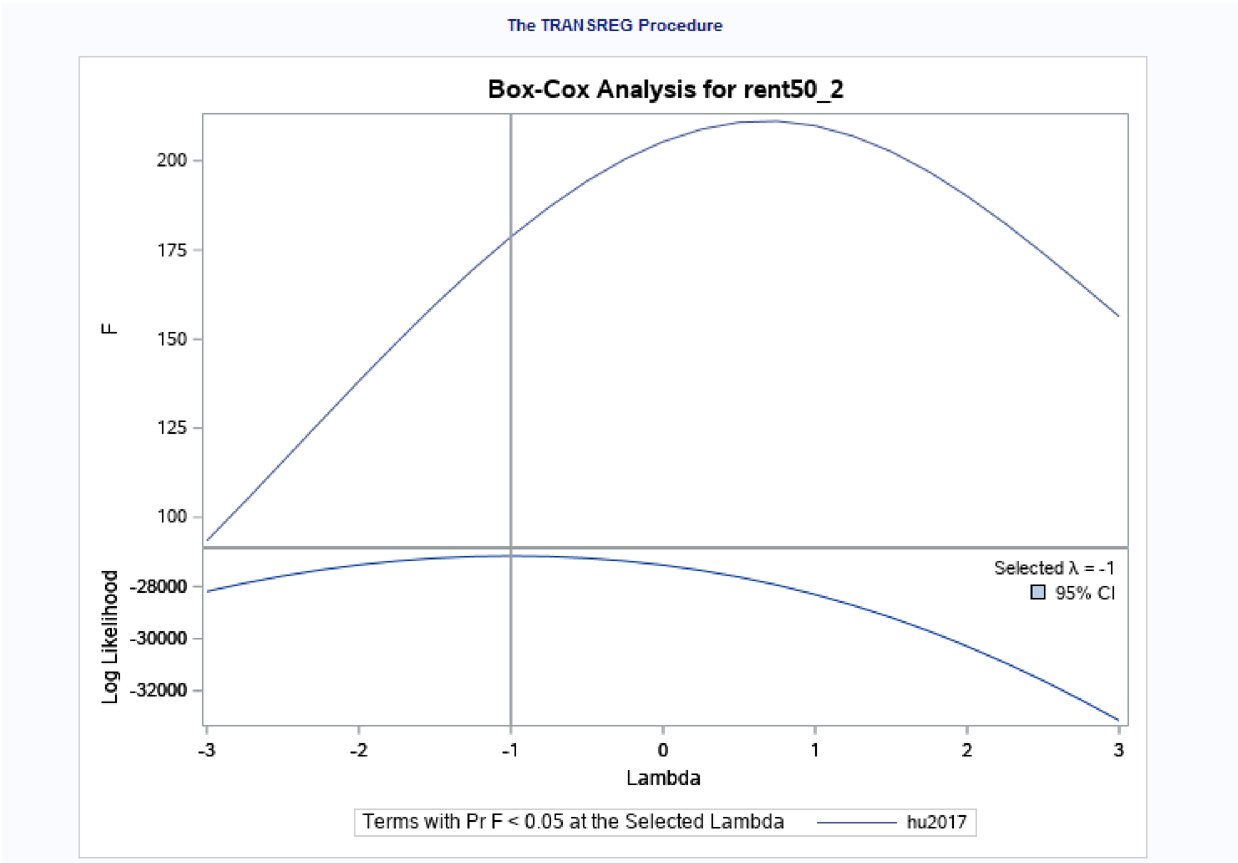
TRANSREG Univariate Algorithm Iteration History for BoxCox(rent50_2)					
Iteration Number	Average Change	Maximum Change	R-Square	Criterion Change	Note
1	0.00000	0.00000	0.03588		Converged

Algorithm converged.

STA 3064 PROJECT A

The TRANSREG Procedure





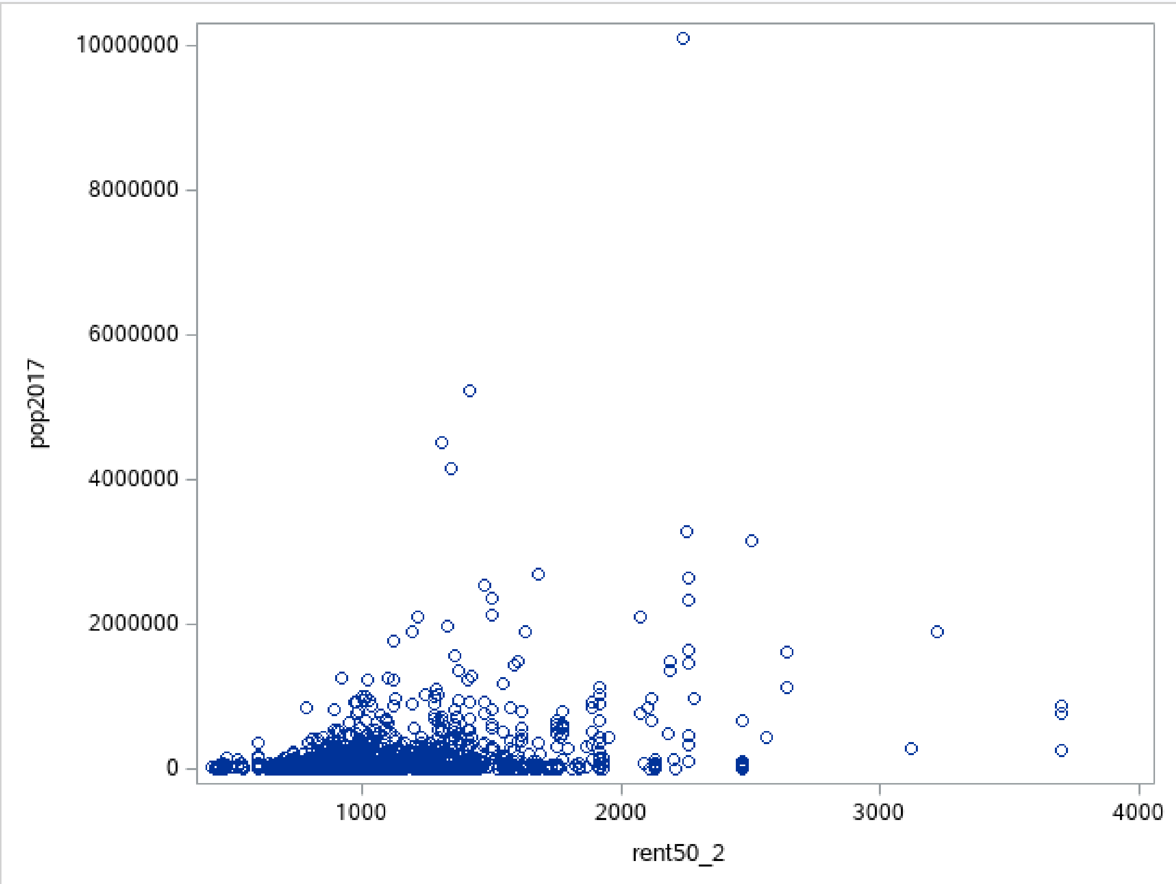
As shown above, both the ratio variables as well as pop2017 and hu2017 are recommended to be transformed to the -1^{st} power.

In addition to producing scatterplots for the classical "rent50_2=hu2017" model, I also generated five additional scatterplots using the pop2017 variable, as well as the two transformed versions of those variables and the ratios of those variables using the following code:

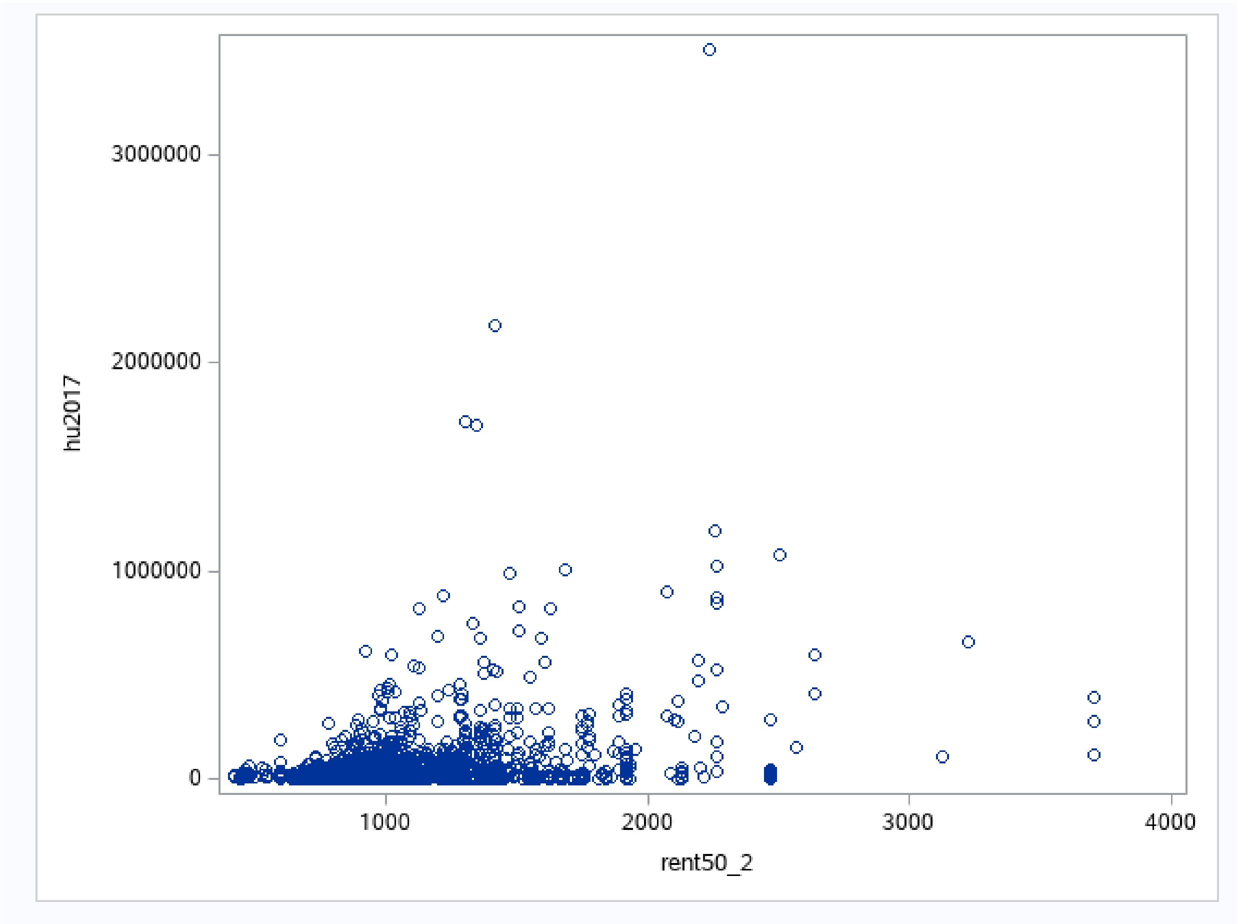
```
proc sgplot data=Housing;
    scatter X=rent50_2 Y=pop2017;
run;
proc sgplot data=Housing;
    scatter X=rent50_2 Y=hu2017;
run;
proc sgplot data=Housing;
    scatter X=rent50_2 Y=HtoP;
```

```
run;
proc sgplot data=Housing;
	scatter X=rent50_2 Y=PtoH;
run;
proc sgplot data=Housing;
	scatter X=rent50_2 Y=transHu;
run;
proc sgplot data=Housing;
	scatter X=rent50_2 Y=transPop;
run;
```

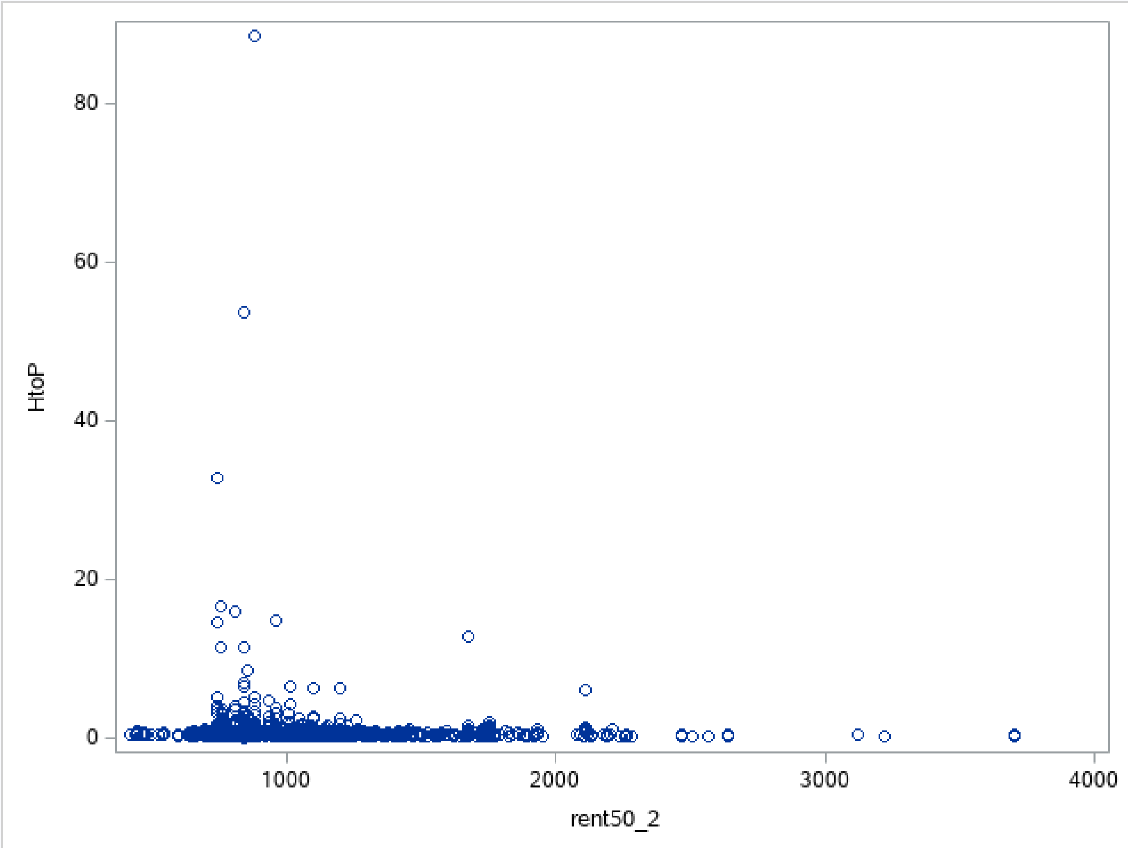
STA 3064 PROJECT A



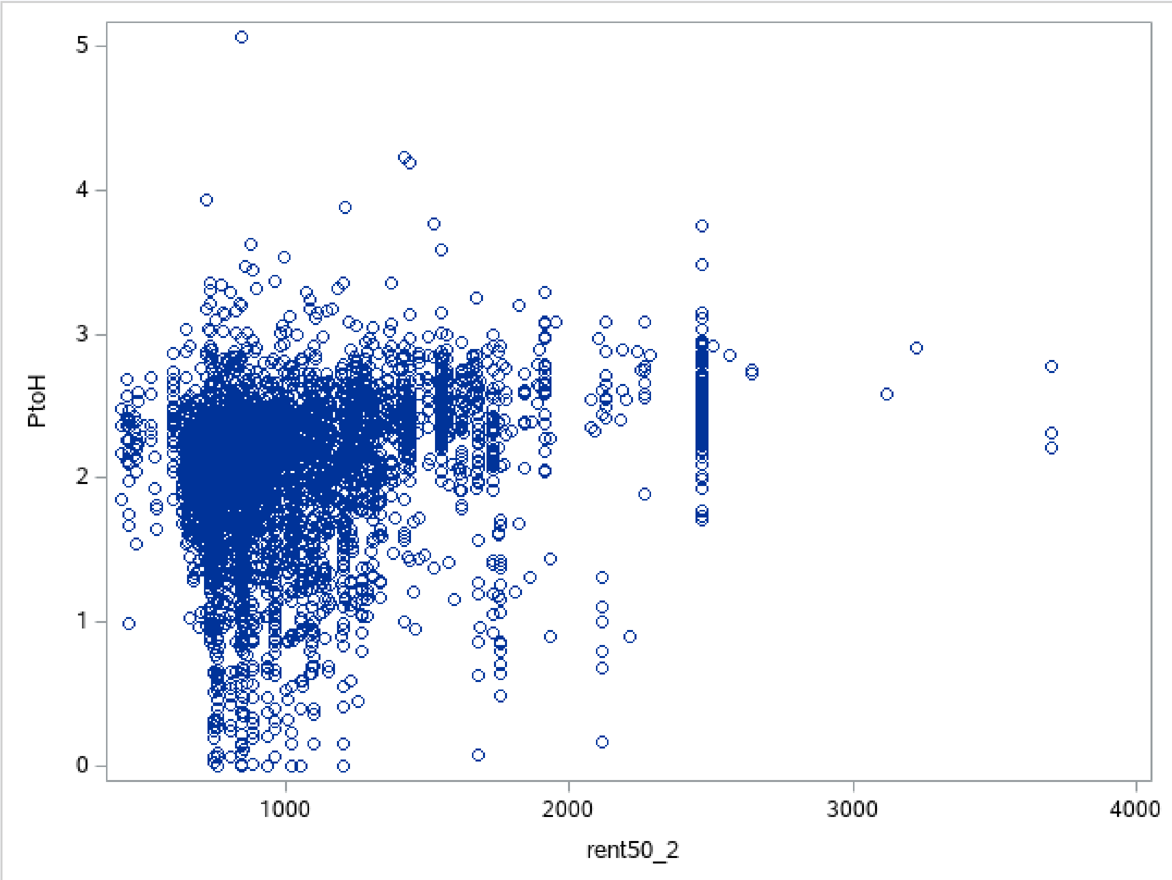
STA 3064 PROJECT A



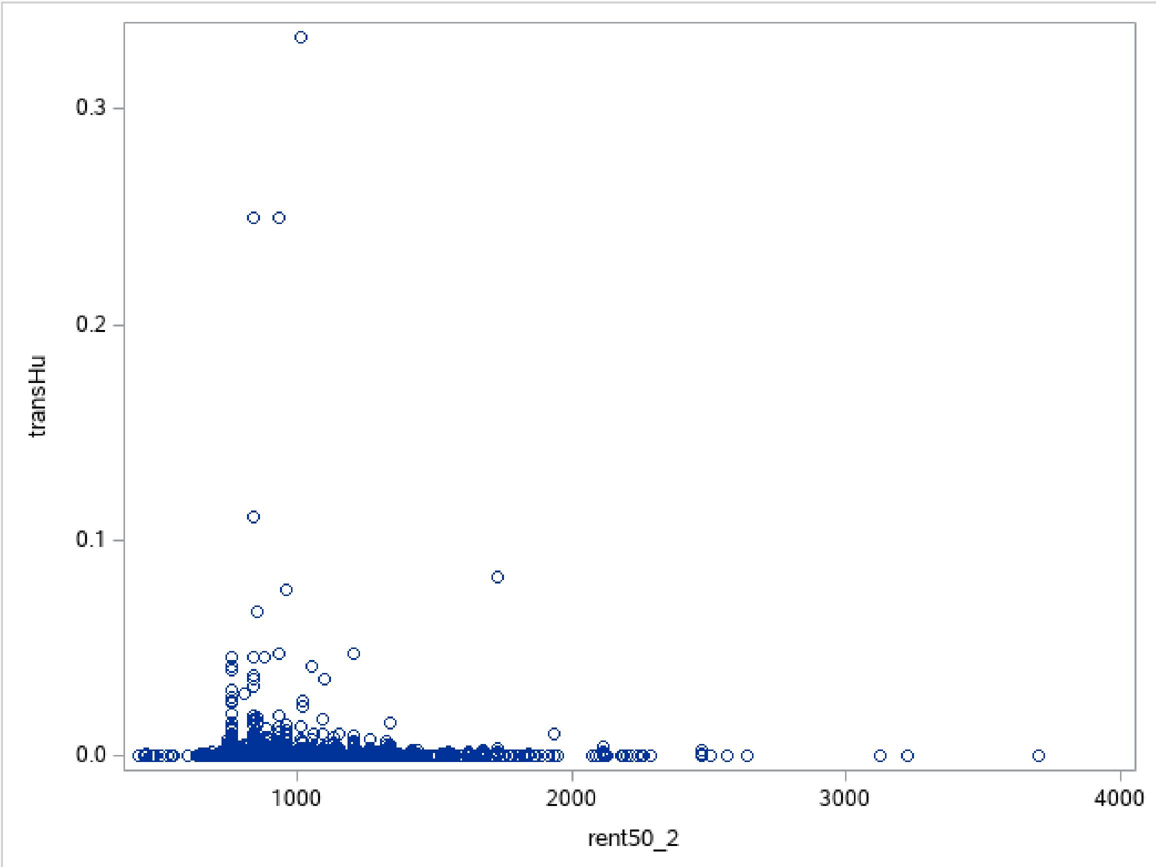
STA 3064 PROJECT A



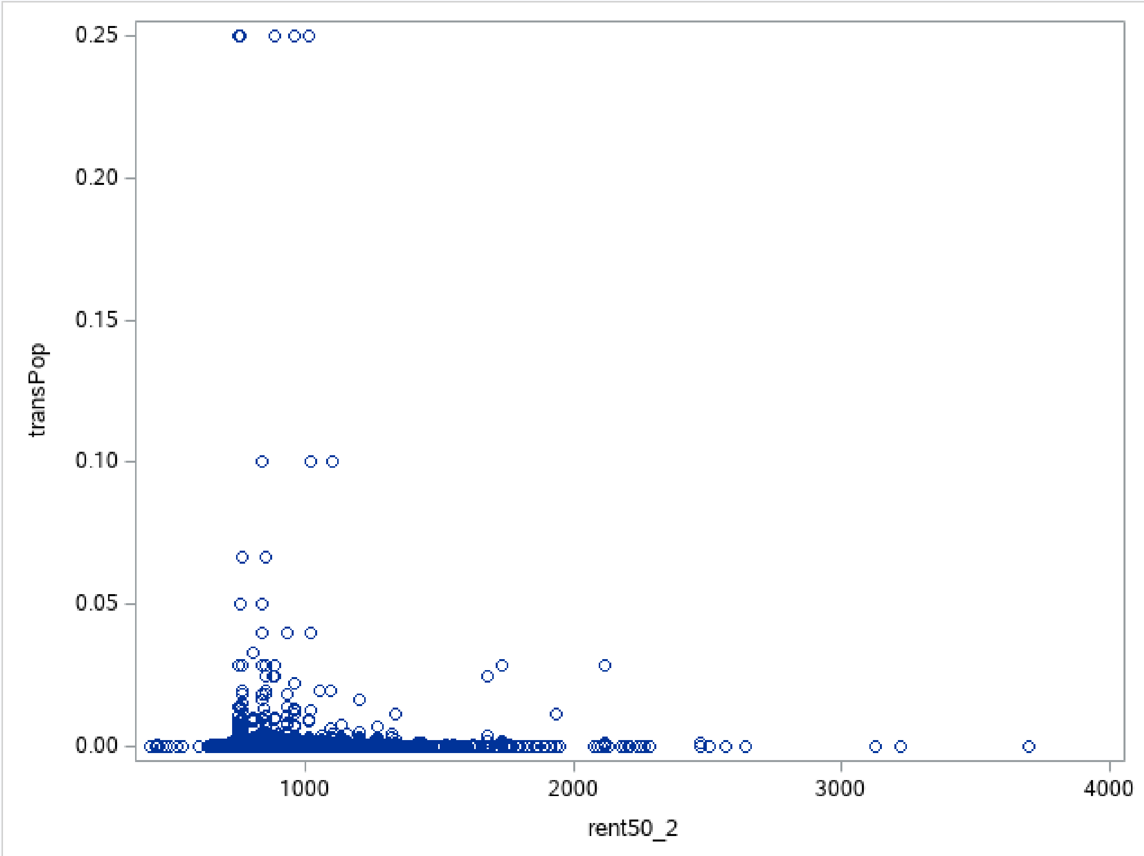
STA 3064 PROJECT A



STA 3064 PROJECT A



STA 3064 PROJECT A



STA 3064 PROJECT A

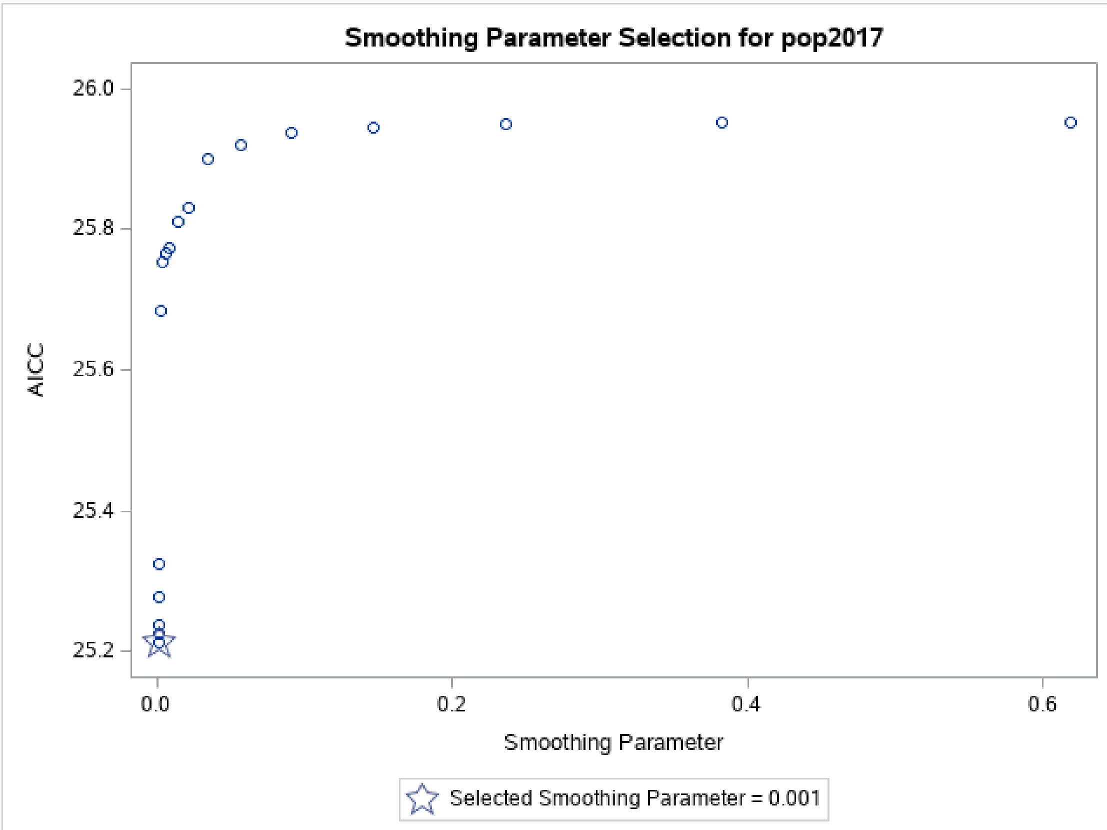
In addition to the previous code and output for proc sgplot, I also used the following code to generate LOESS curves:

```
proc loess data=Housing;
    model pop2017=rent50_2;
run;
proc loess data=Housing;
    model pop2017=rent50_2;
run;
proc loess data=Housing;
    model pop2017=HtoP;
run;
proc loess data=Housing;
    model pop2017=PtoH;
run;
proc loess data=Housing;
    model transPop=rent50_2;
run;
proc loess data=Housing;
    model transHu=rent50_2;
run;
```

The LOESS Procedure

Independent Variable Scaling	
Scaling applied: None	
Statistic	rent50_2
Minimum Value	421
Maximum Value	3700

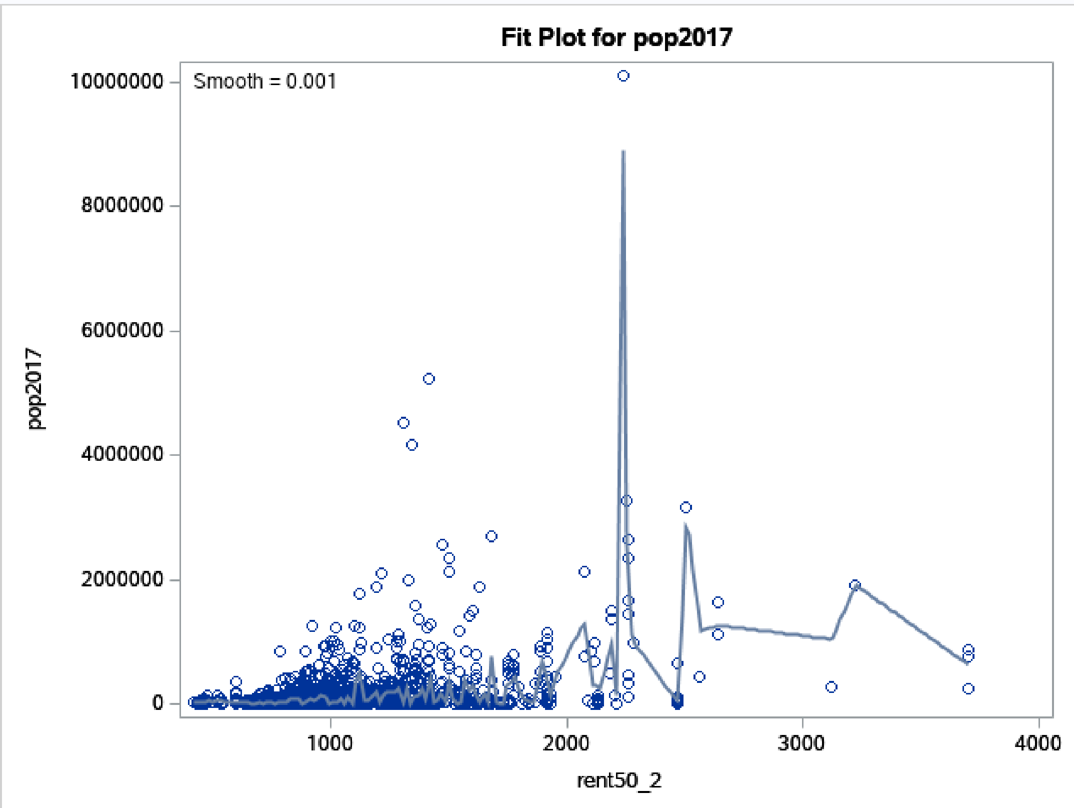
The LOESS Procedure
Dependent Variable: pop2017

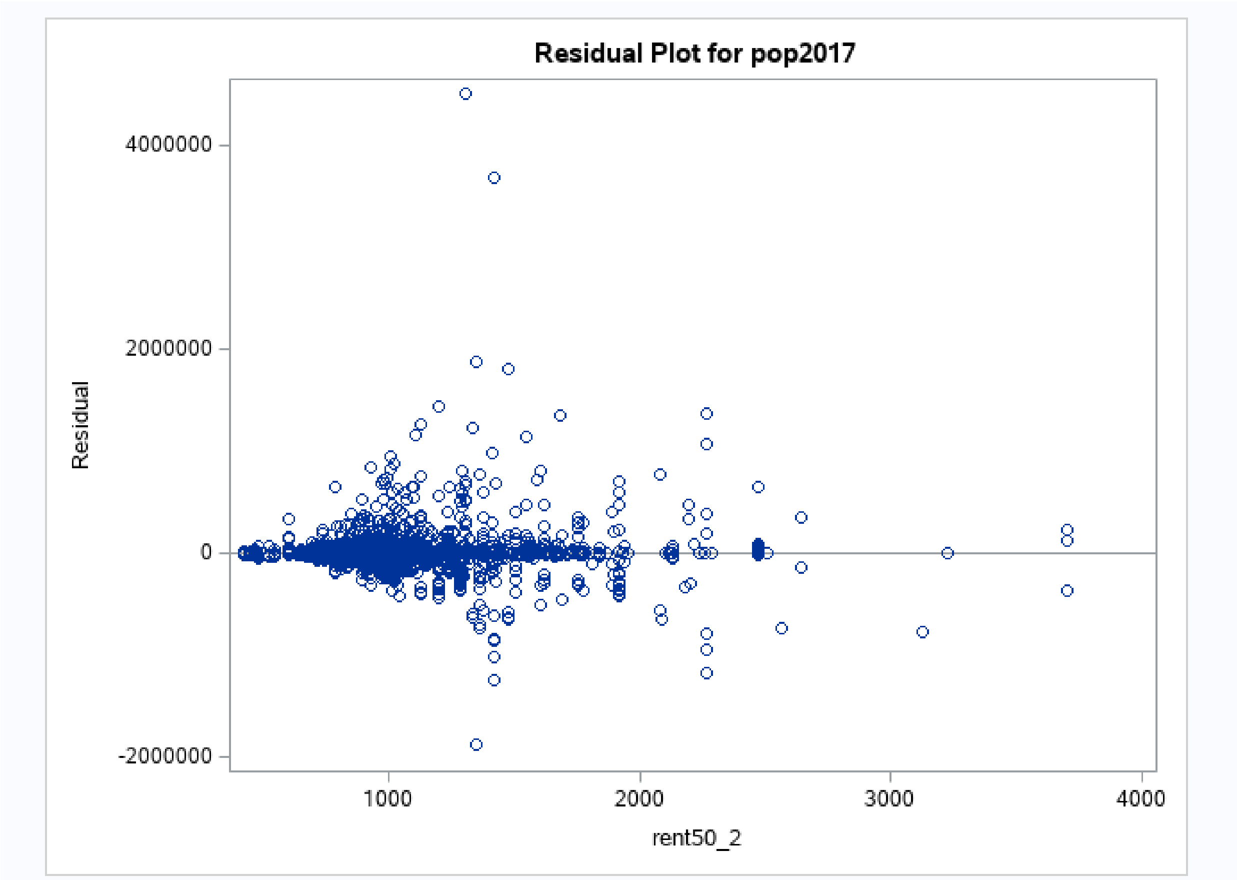


Optimal Smoothing Criterion	
AICC	Smoothing Parameter
25.21332	0.00115

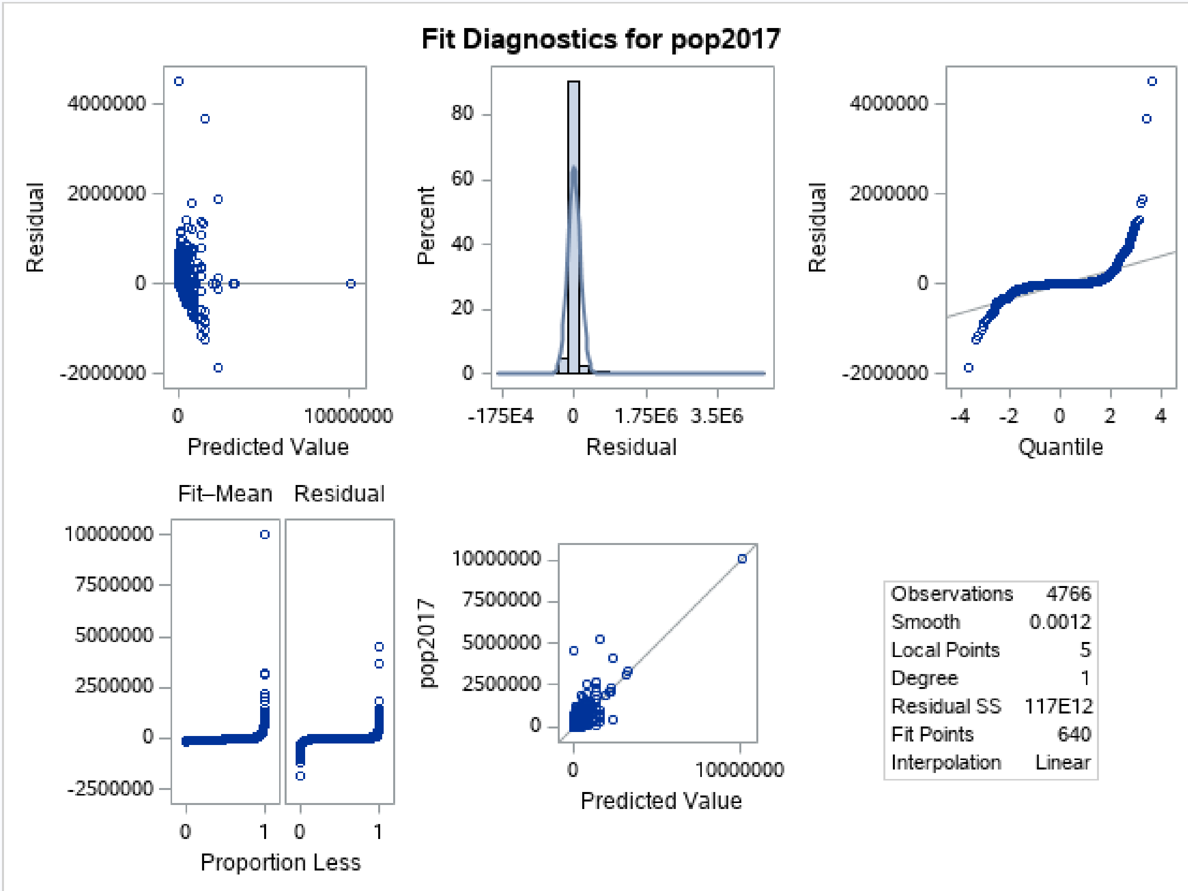
The LOESS Procedure
Selected Smoothing Parameter: 0.001
Dependent Variable: pop2017

Fit Summary	
Fit Method	kd Tree
Blending	Linear
Number of Observations	4766
Number of Fitting Points	640
kd Tree Bucket Size	1
Degree of Local Polynomials	1
Smoothing Parameter	0.00115
Points in Local Neighborhood	5
Residual Sum of Squares	1.165671E14
Trace[L]	608.05040
GCV	6742449
AICC	25.21332




Residual Plot for pop2017
Residual
4000000
2000000
0
-2000000
1000
2000
3000
4000
rent50_2

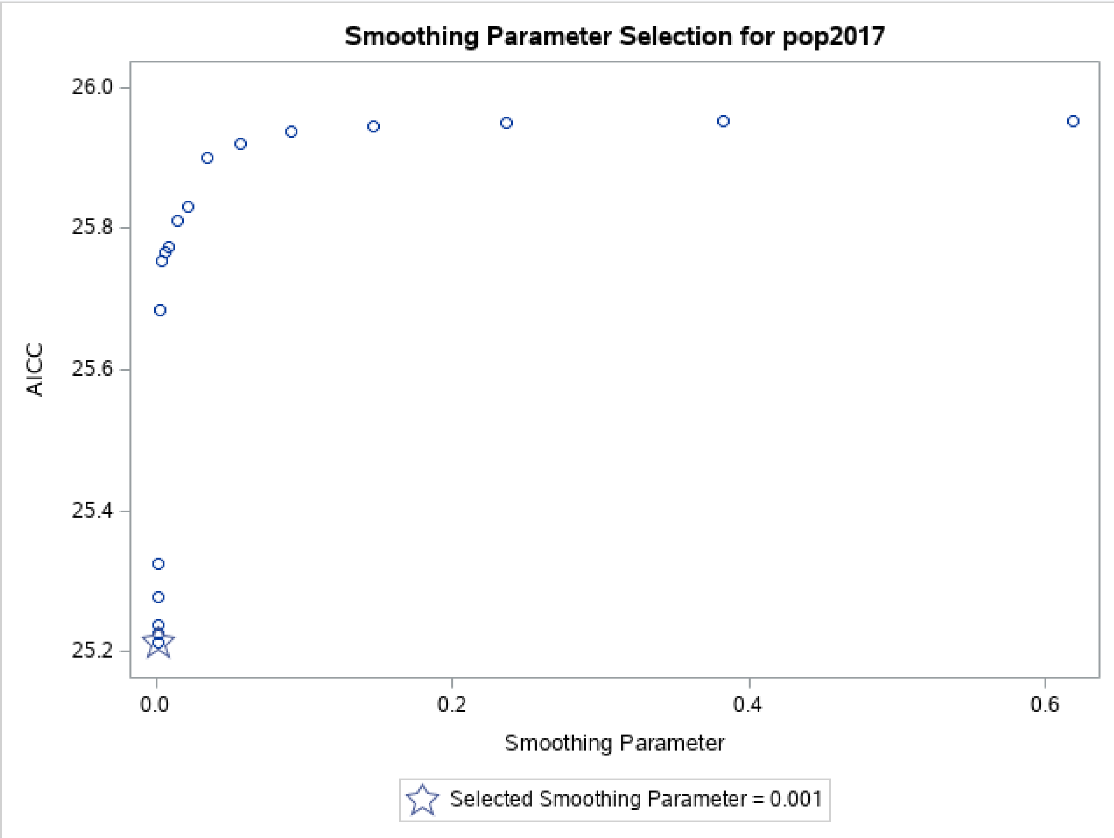
STA 3064 PROJECT A



The LOESS Procedure

Independent Variable Scaling	
Scaling applied: None	
Statistic	rent50_2
Minimum Value	421
Maximum Value	3700

The LOESS Procedure
Dependent Variable: pop2017

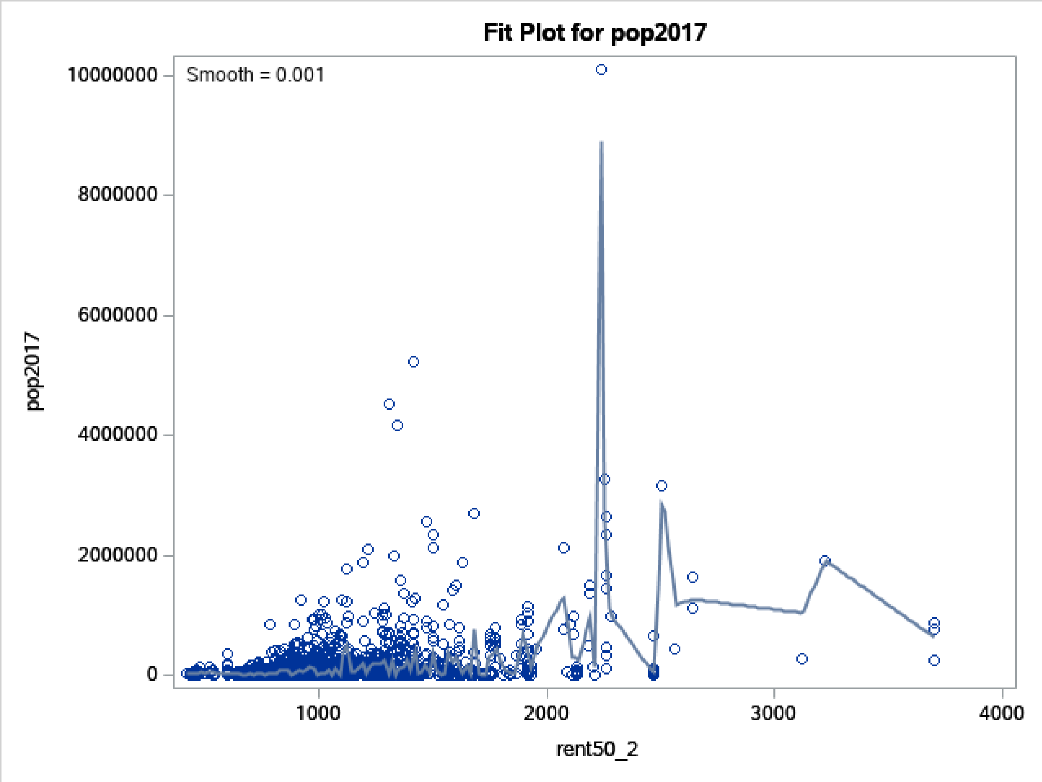


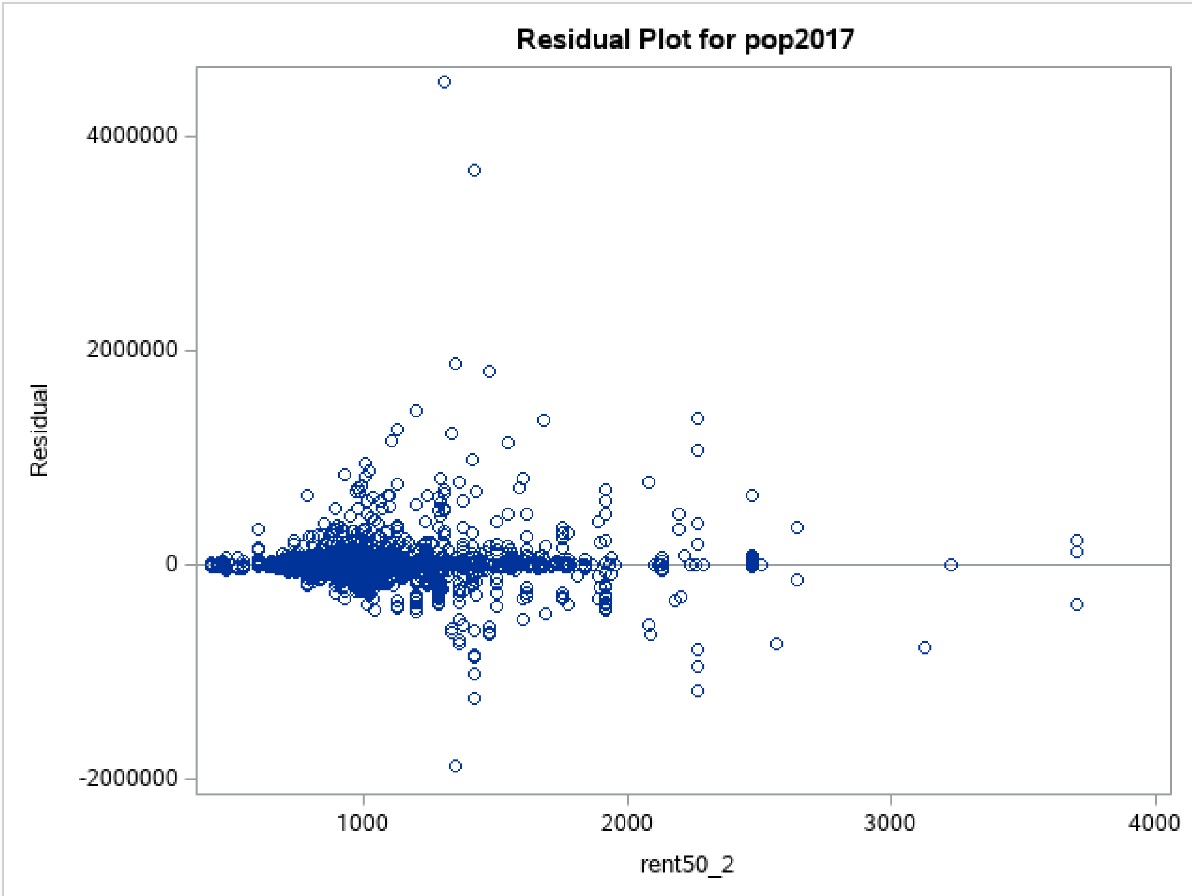
Optimal Smoothing Criterion	
AICC	Smoothing Parameter
25.21332	0.00115

STA 3064 PROJECT A

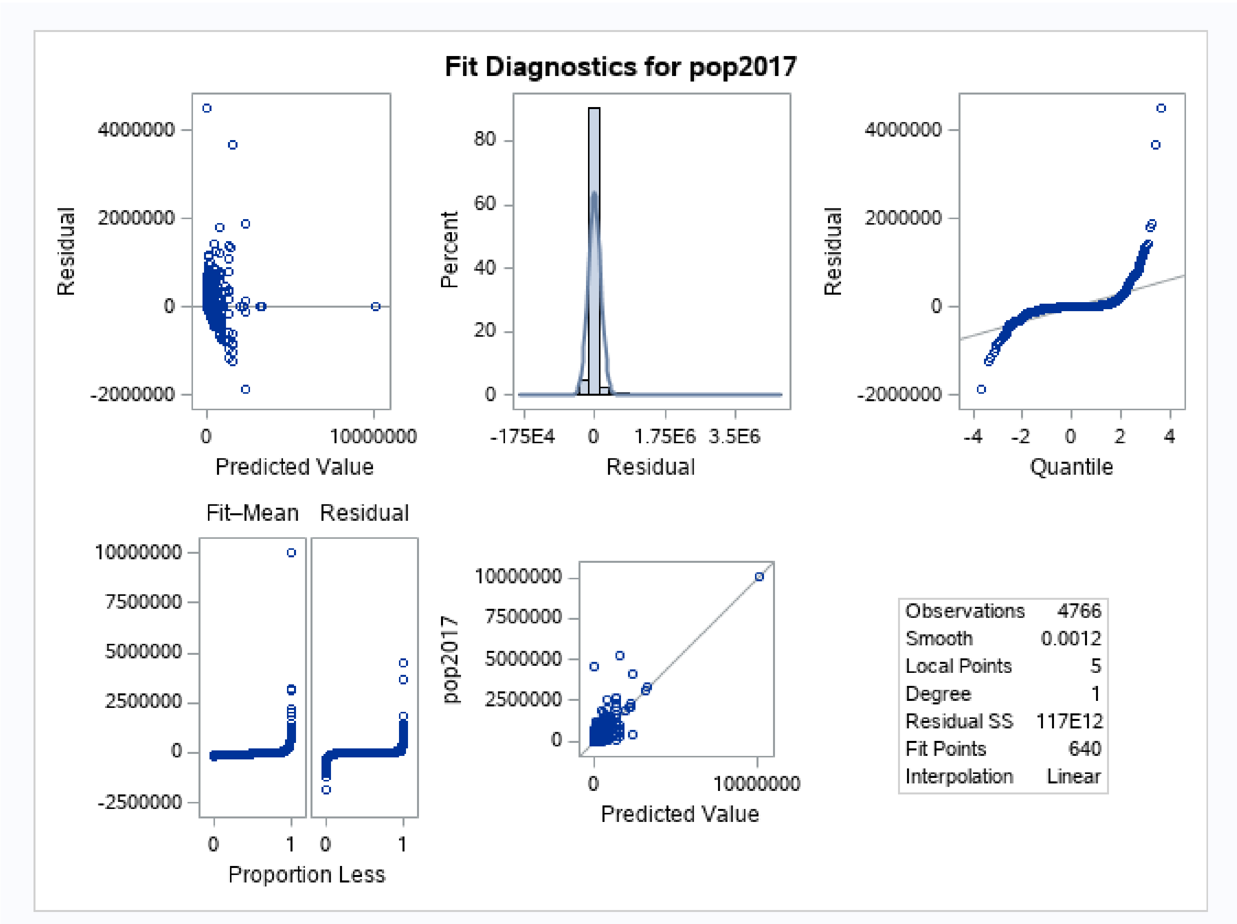
The LOESS Procedure
Selected Smoothing Parameter: 0.001
Dependent Variable: pop2017

Fit Summary	
Fit Method	kd Tree
Blending	Linear
Number of Observations	4766
Number of Fitting Points	640
kd Tree Bucket Size	1
Degree of Local Polynomials	1
Smoothing Parameter	0.00115
Points in Local Neighborhood	5
Residual Sum of Squares	1.165671E14
Trace[L]	608.05040
GCV	6742449
AICC	25.21332




Residual Plot for pop2017
Residual
4000000
2000000
0
-2000000
1000
2000
3000
4000
rent50_2

STA 3064 PROJECT A

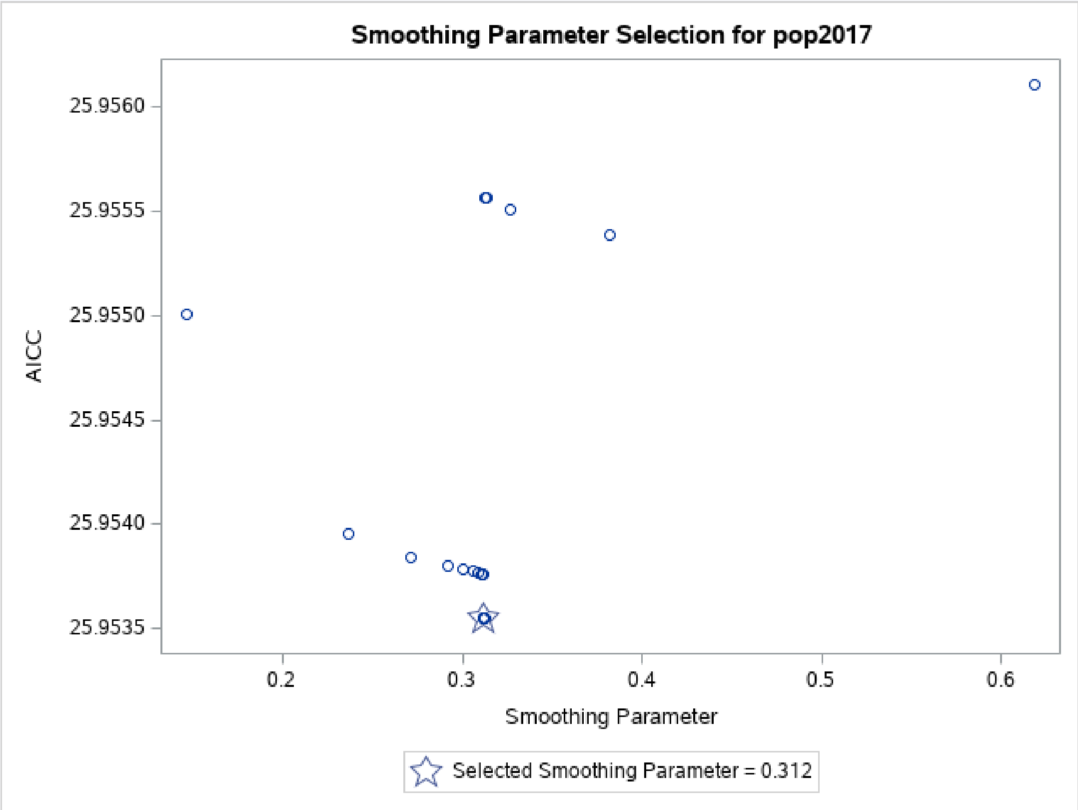


STA 3064 PROJECT A

The LOESS Procedure

Independent Variable Scaling	
Scaling applied: None	
Statistic	HtoP
Minimum Value	0
Maximum Value	88.50000

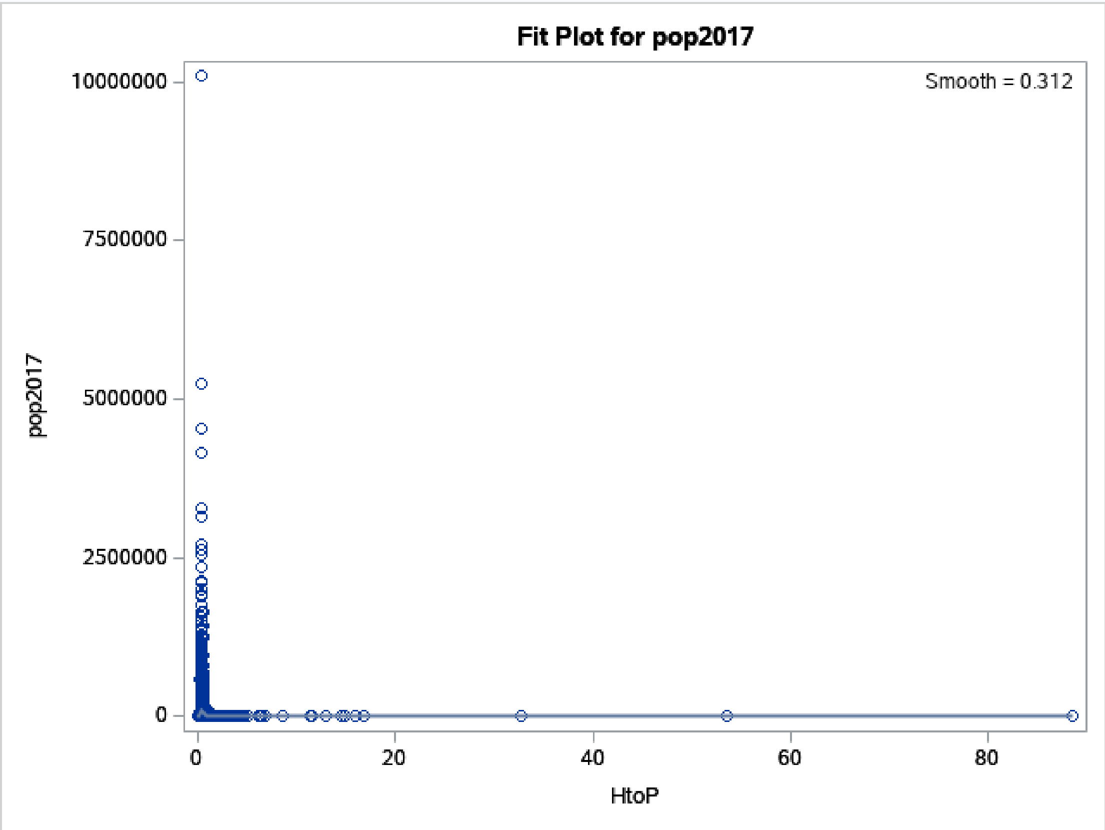
The LOESS Procedure
Dependent Variable: pop2017

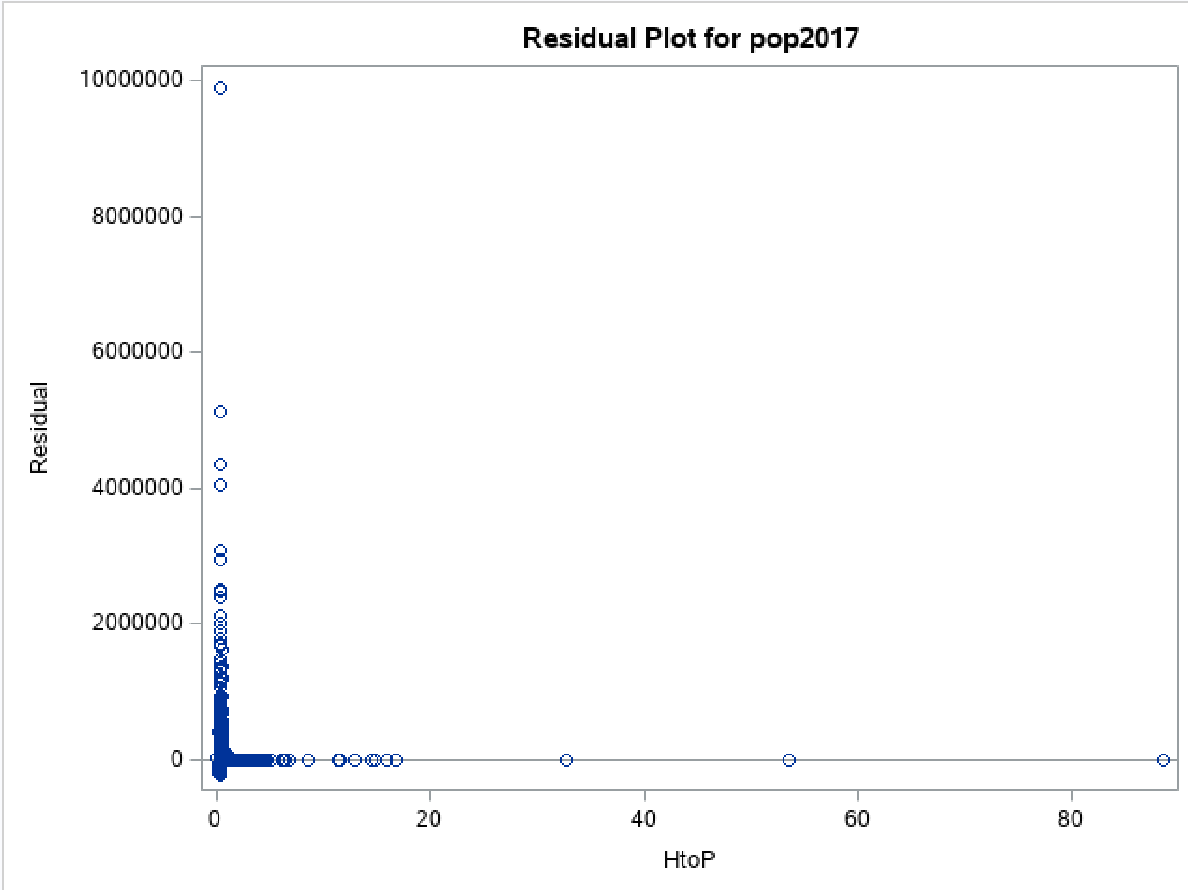


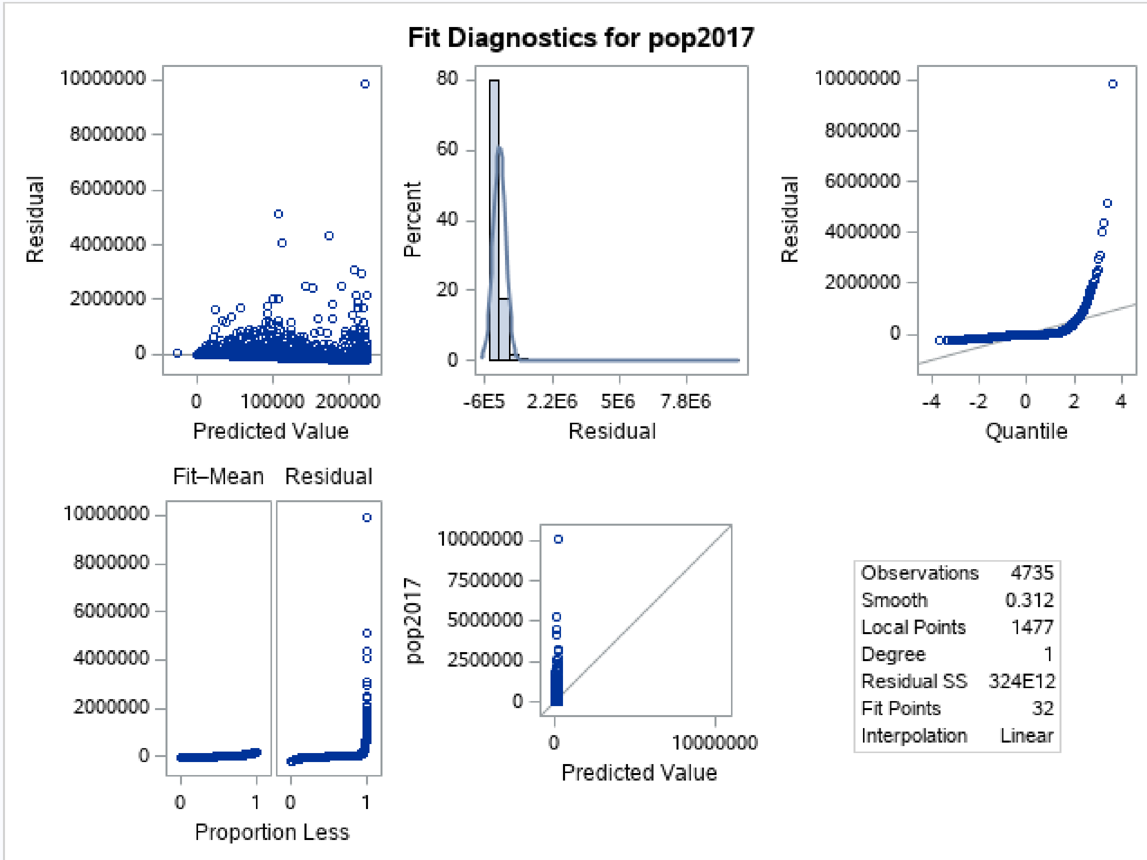
Optimal Smoothing Criterion	
AICC	Smoothing Parameter
25.95355	0.31204

The LOESS Procedure
Selected Smoothing Parameter: 0.312
Dependent Variable: pop2017

Fit Summary	
Fit Method	kd Tree
Blending	Linear
Number of Observations	4735
Number of Fitting Points	32
kd Tree Bucket Size	295
Degree of Local Polynomials	1
Smoothing Parameter	0.31204
Points in Local Neighborhood	1477
Residual Sum of Squares	3.241161E14
Trace[L]	8.83183
GCV	14510497
AICC	25.95355




Residual Plot for pop2017
10000000
8000000
6000000
4000000
2000000
0
Residual
0
20
40
60
80
HtoP

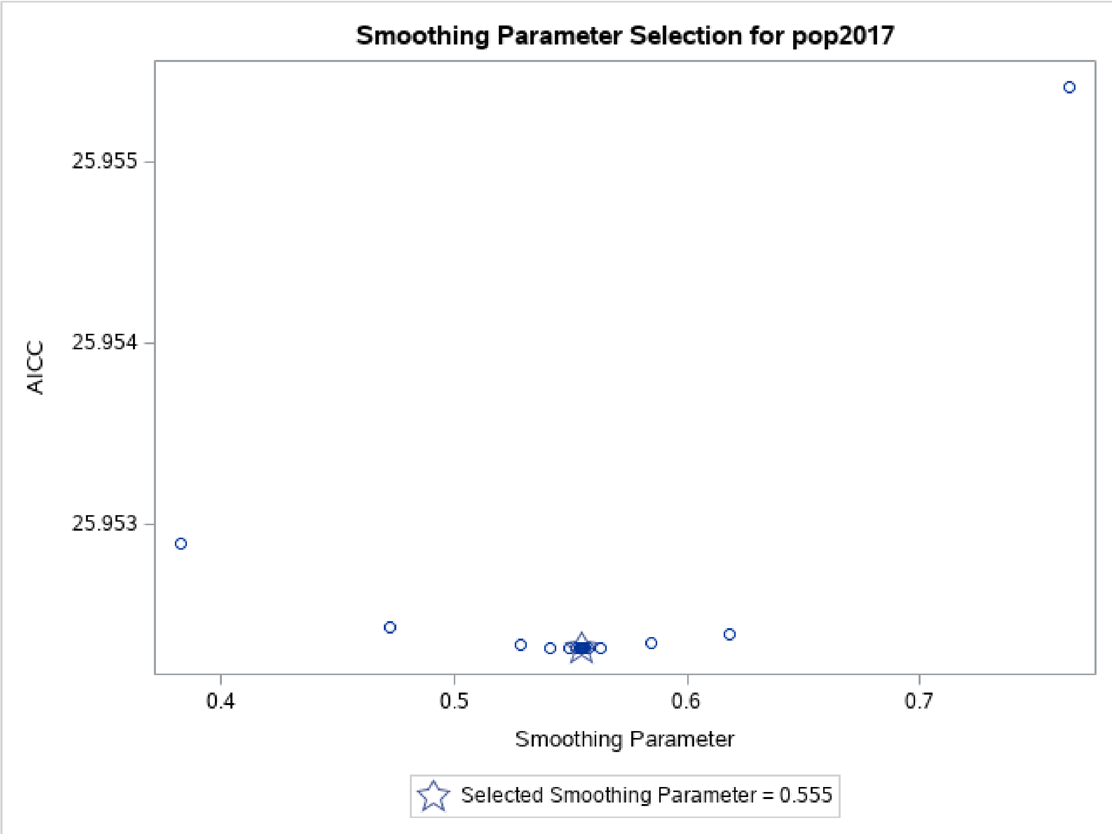

Fit Diagnostics for pop2017
Residual
Predicted Value
Percent
Residual
Quantile
Fit–Mean
Residual
Proportion Less
pop2017
Predicted Value
Observations 4735
Smooth 0.312
Local Points 1477
Degree 1
Residual SS 324E12
Fit Points 32
Interpolation Linear

STA 3064 PROJECT A

The LOESS Procedure

Independent Variable Scaling	
Scaling applied: None	
Statistic	PtoH
Minimum Value	0
Maximum Value	5.06422

The LOESS Procedure
Dependent Variable: pop2017

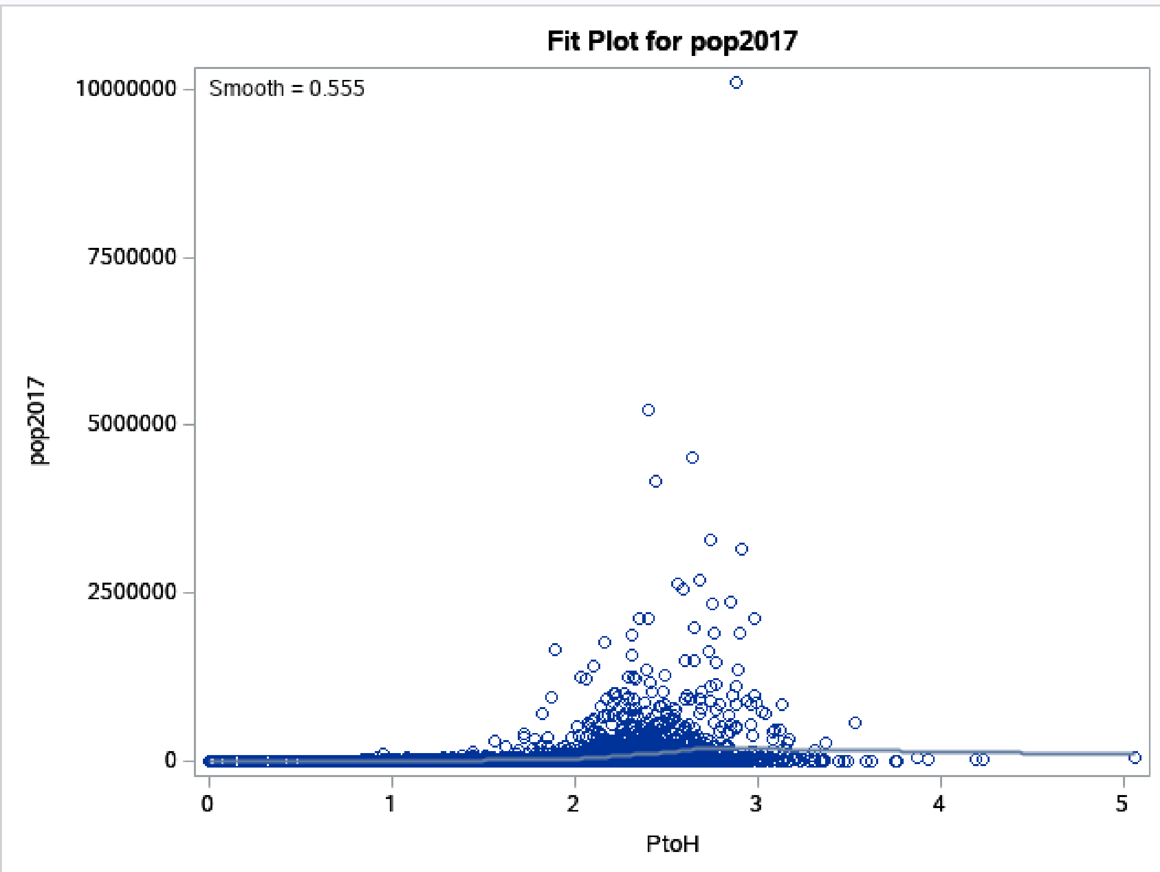


Optimal Smoothing Criterion	
AICC	Smoothing Parameter
25.95231	0.55489

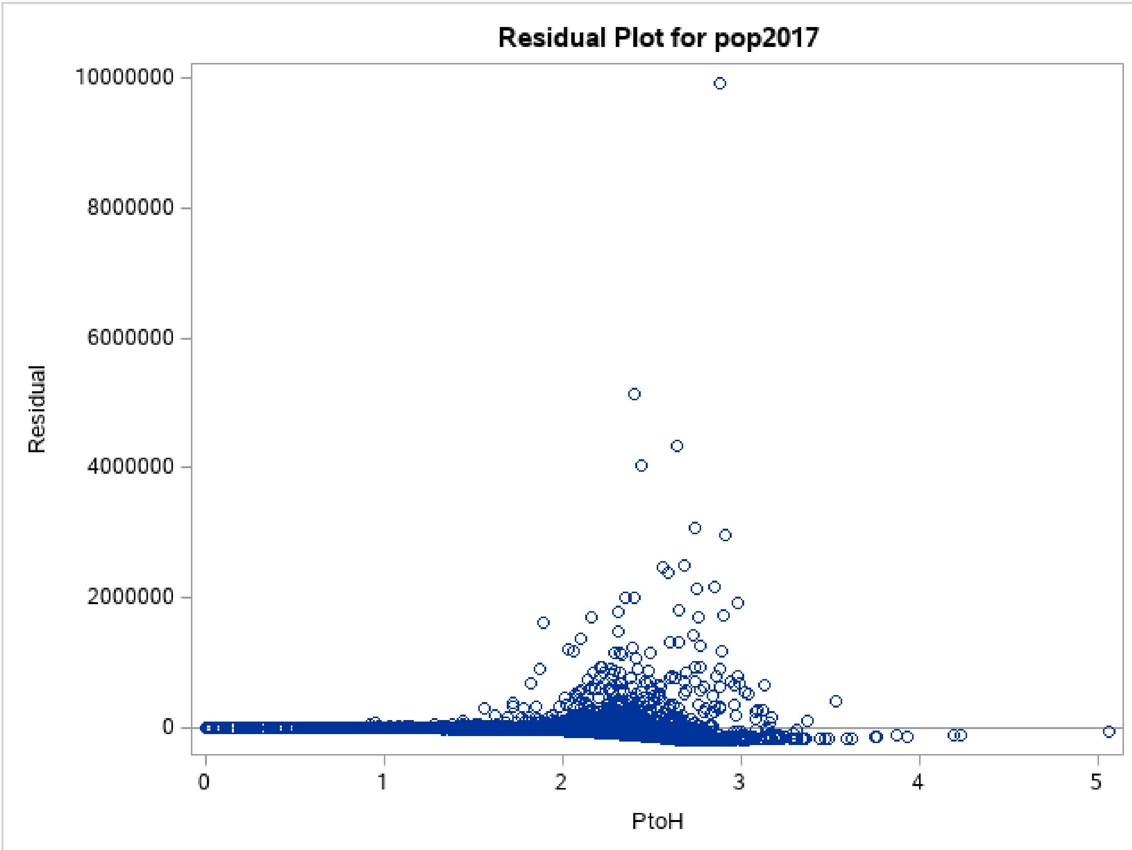
STA 3064 PROJECT A

The LOESS Procedure
Selected Smoothing Parameter: 0.555
Dependent Variable: pop2017

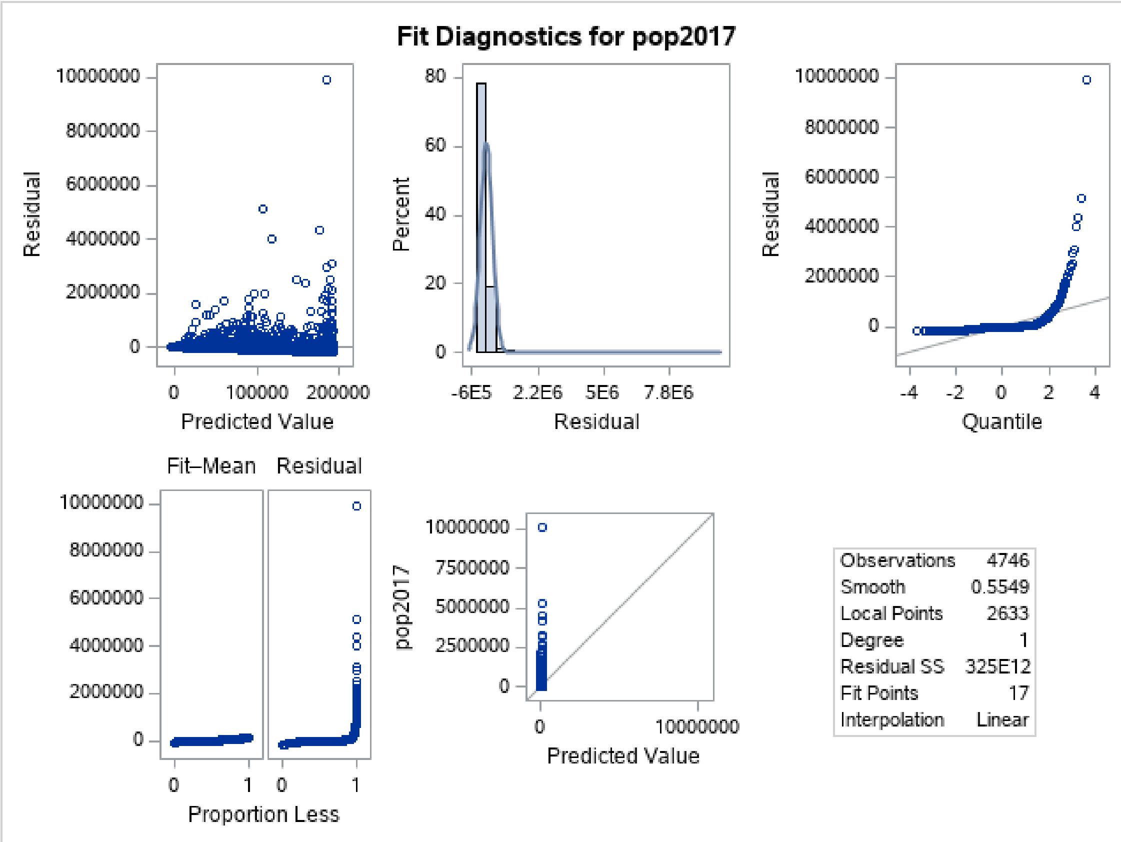
Fit Summary	
Fit Method	kd Tree
Blending	Linear
Number of Observations	4746
Number of Fitting Points	17
kd Tree Bucket Size	526
Degree of Local Polynomials	1
Smoothing Parameter	0.55489
Points in Local Neighborhood	2633
Residual Sum of Squares	3.249007E14
Trace[L]	5.69377
GCV	14458988
AICC	25.95231



STA 3064 PROJECT A



STA 3064 PROJECT A

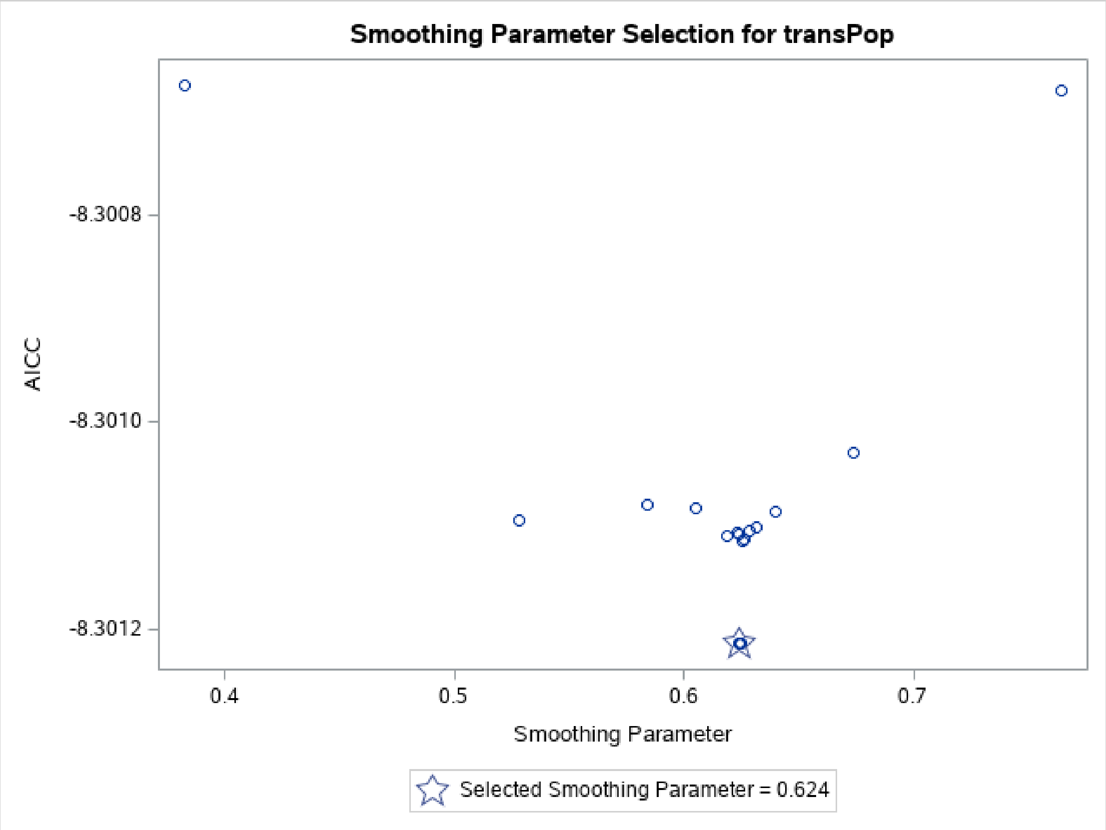


STA 3064 PROJECT A

The LOESS Procedure

Independent Variable Scaling	
Scaling applied: None	
Statistic	rent50_2
Minimum Value	421
Maximum Value	3700

The LOESS Procedure
Dependent Variable: transPop

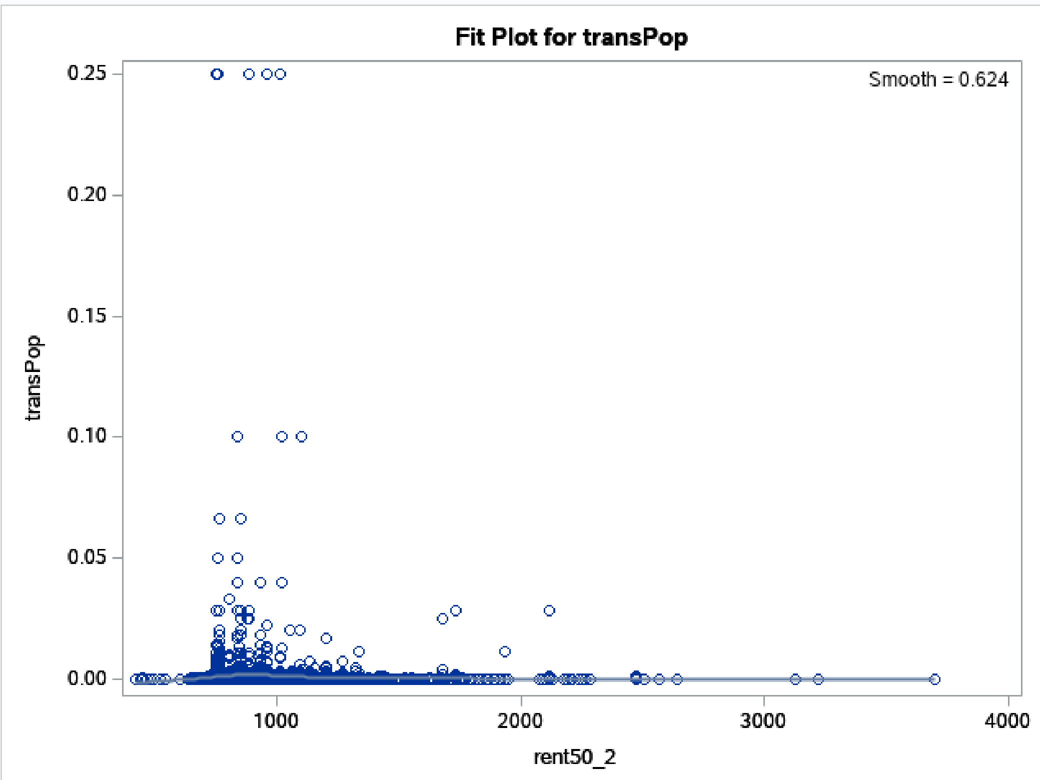


Optimal Smoothing Criterion	
AICC	Smoothing Parameter
-8.30121	0.62418

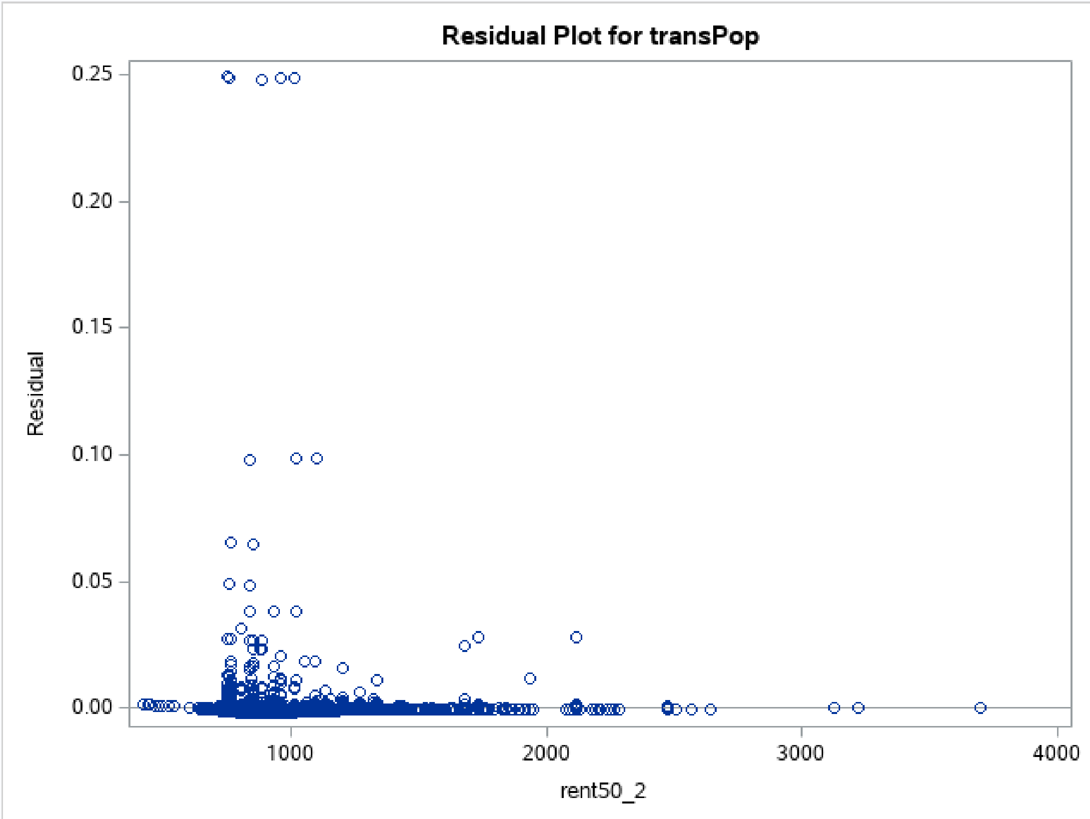
STA 3064 PROJECT A

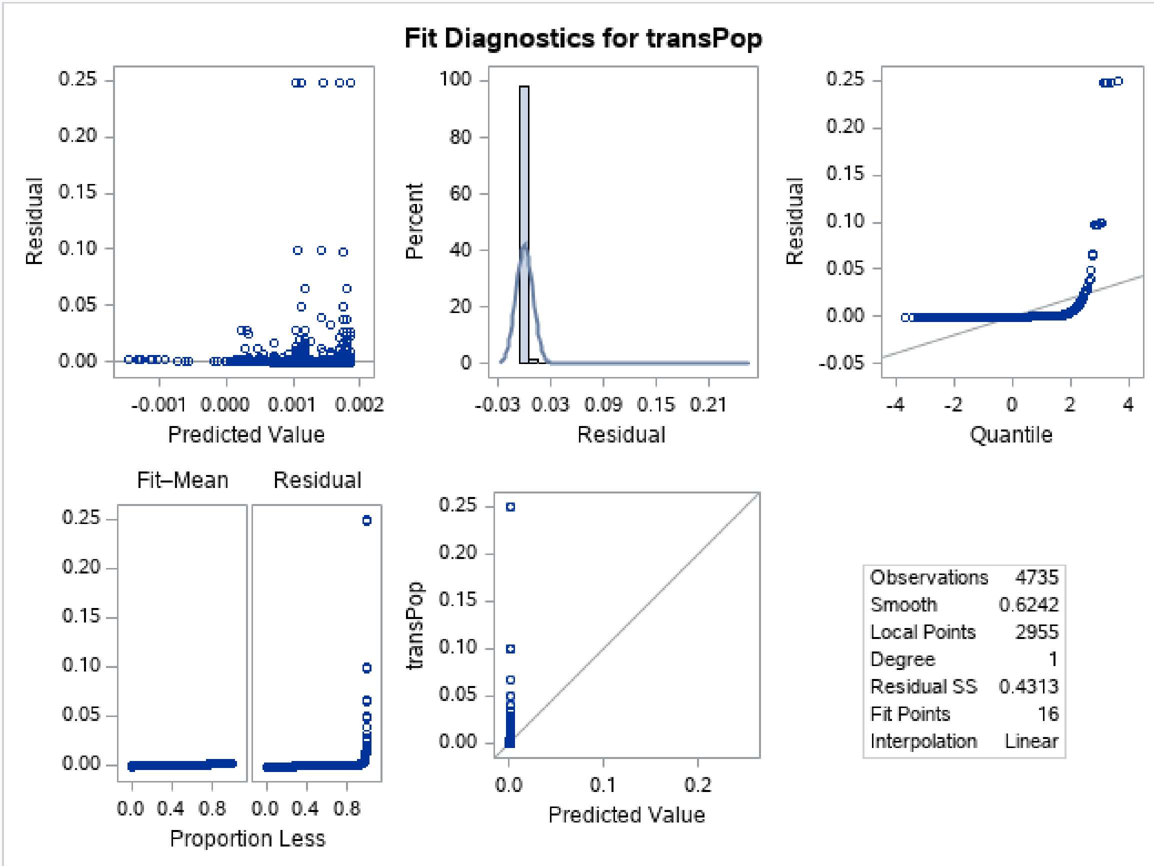
The LOESS Procedure
Selected Smoothing Parameter: 0.624
Dependent Variable: transPop

Fit Summary	
Fit Method	kd Tree
Blending	Linear
Number of Observations	4735
Number of Fitting Points	16
kd Tree Bucket Size	591
Degree of Local Polynomials	1
Smoothing Parameter	0.62418
Points in Local Neighborhood	2955
Residual Sum of Squares	0.43132
Trace[L]	4.75369
GCV	1.927657E-8
AICC	-8.30121



STA 3064 PROJECT A



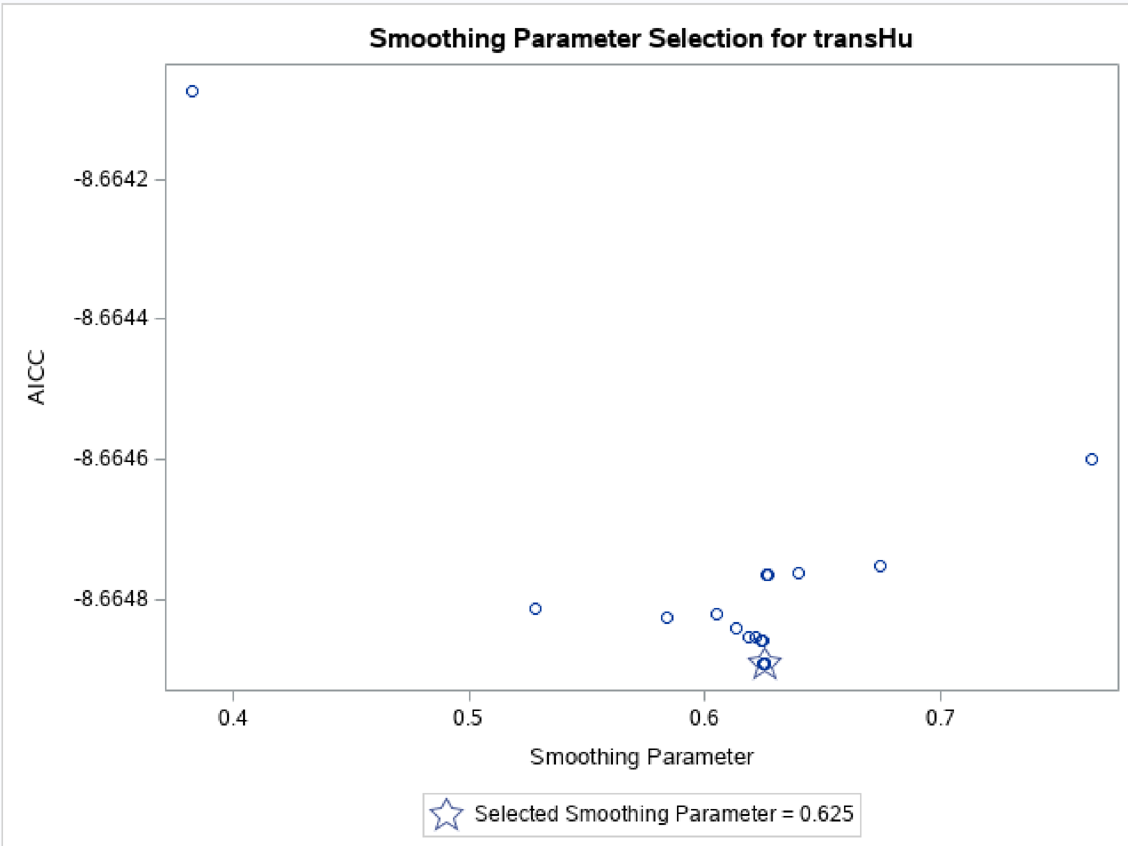

Fit Diagnostics for transPop
Residual
Predicted Value
Percent
Residual
Quantile
Fit–Mean
Residual
Proportion Less
transPop
Predicted Value
Observations 4735
Smooth 0.6242
Local Points 2955
Degree 1
Residual SS 0.4313
Fit Points 16
Interpolation Linear

STA 3064 PROJECT A

The LOESS Procedure

Independent Variable Scaling	
Scaling applied: None	
Statistic	rent50_2
Minimum Value	421
Maximum Value	3700

The LOESS Procedure
Dependent Variable: transHu

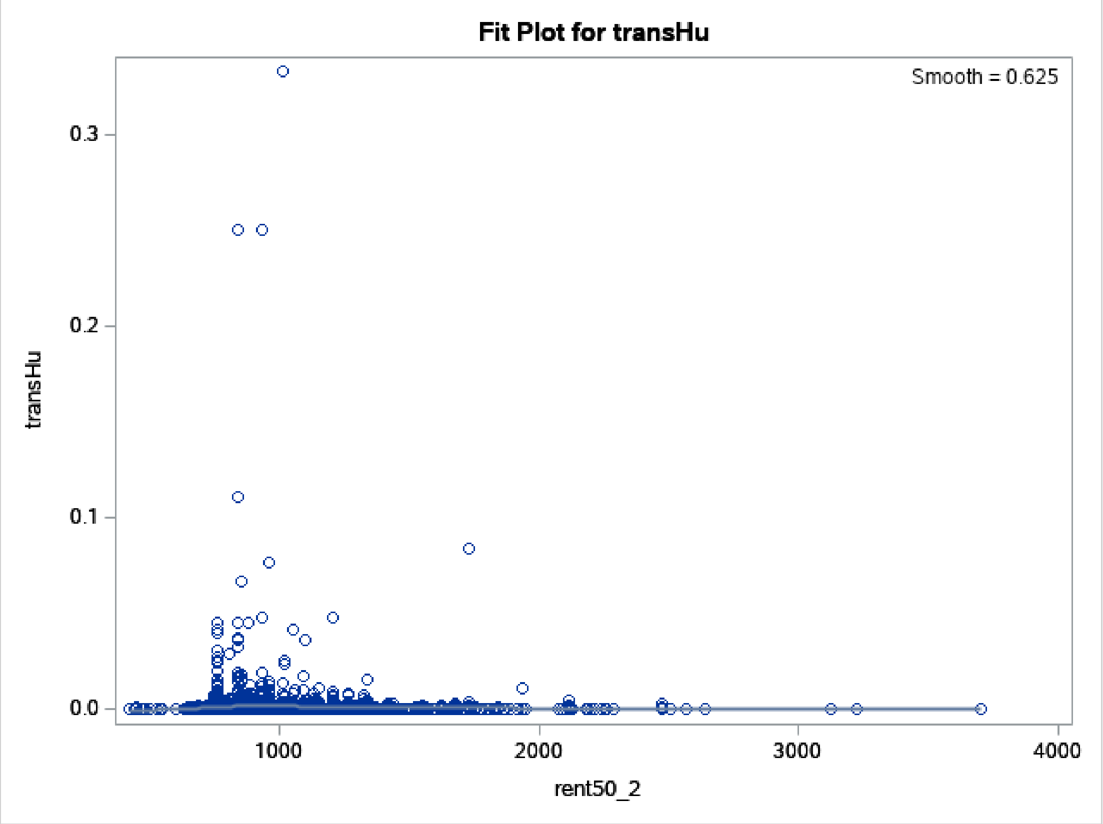


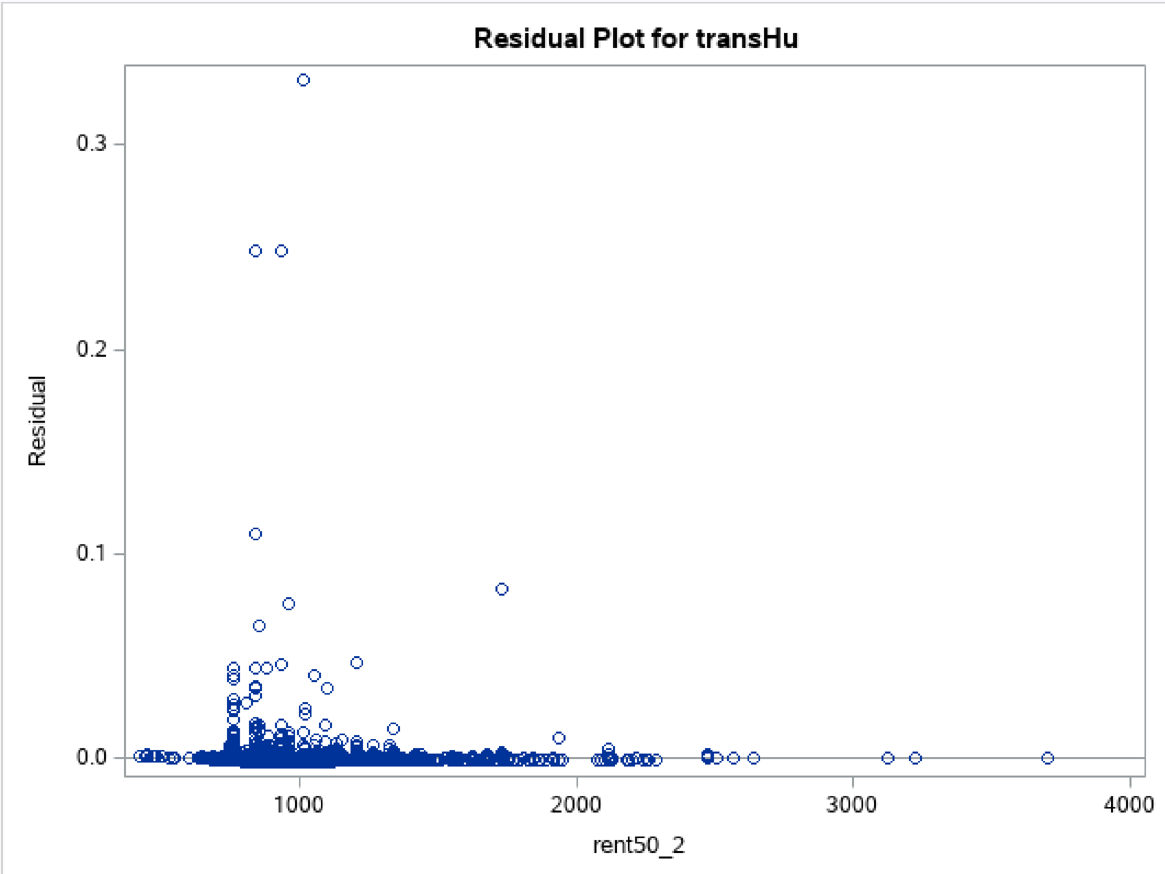
Optimal Smoothing Criterion	
AICC	Smoothing Parameter
-8.66489	0.62505

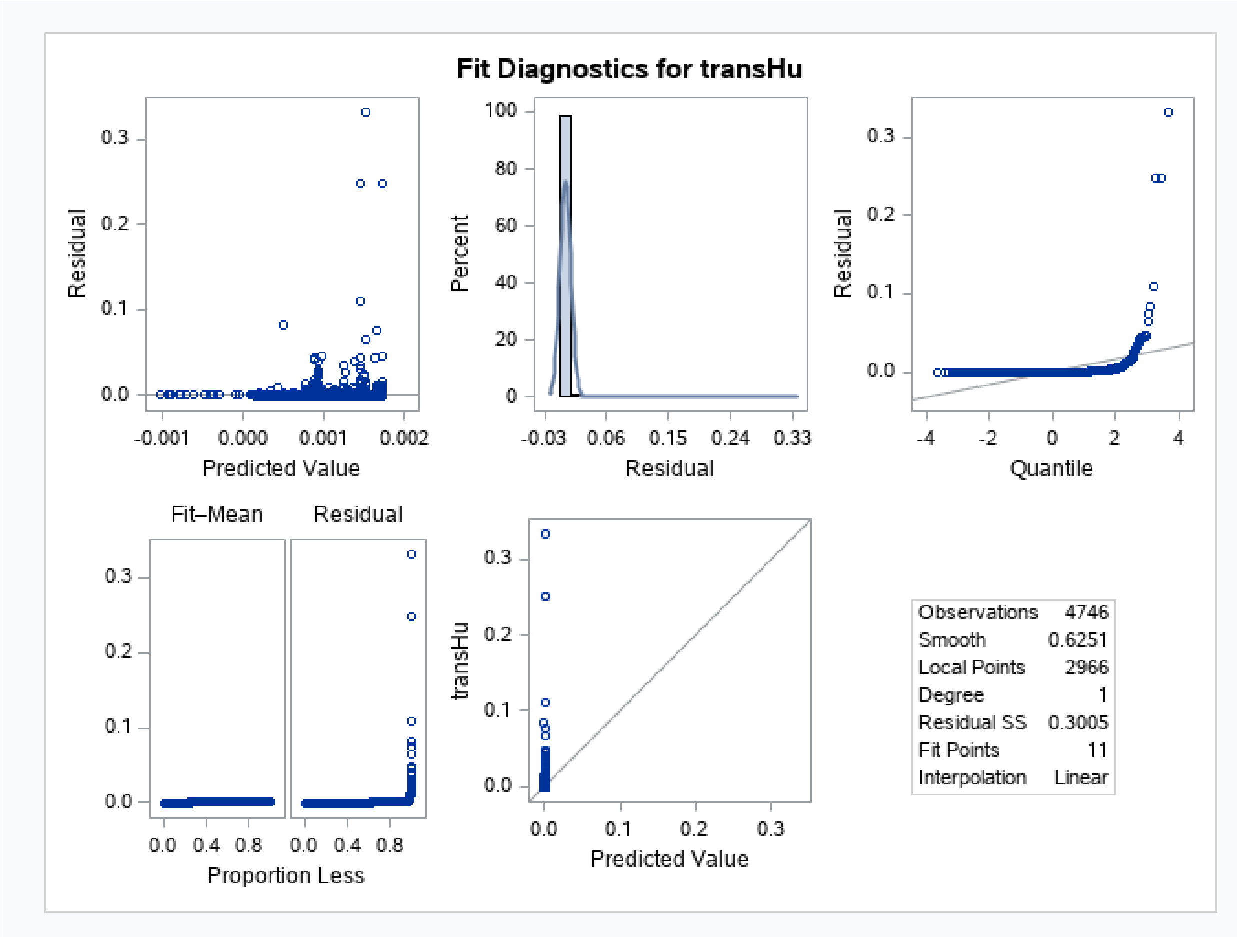
STA 3064 PROJECT A

The LOESS Procedure
Selected Smoothing Parameter: 0.625
Dependent Variable: transHu

Fit Summary	
Fit Method	kd Tree
Blending	Linear
Number of Observations	4746
Number of Fitting Points	11
kd Tree Bucket Size	593
Degree of Local Polynomials	1
Smoothing Parameter	0.62505
Points in Local Neighborhood	2966
Residual Sum of Squares	0.30052
Trace[L]	4.69079
GCV	1.336839E-8
AICC	-8.66489




Residual Plot for transHu
Residual
0.3
0.2
0.1
0.0
1000
2000
3000
4000
rent50_2


Fit Diagnostics for transHu
Residual
Predicted Value
Percent
Residual
Residual
Quantile
Fit–Mean
Residual
Proportion Less
transHu
Predicted Value
Observations 4746
Smooth 0.6251
Local Points 2966
Degree 1
Residual SS 0.3005
Fit Points 11
Interpolation Linear

Model Fitting and Analysis

While I generated ***several*** potential models, I would like to start with the analysis of the initially assumed relationship of "rent50_2" being the response variable, and "hu2017" being the predictor variable. Using the below code with the original HUD dataset;

proc reg data=HUDdata;

model rent50_2=hu2017;

run;

I generated the following output:

The REG Procedure
Model: MODEL1
Dependent Variable: rent50_2

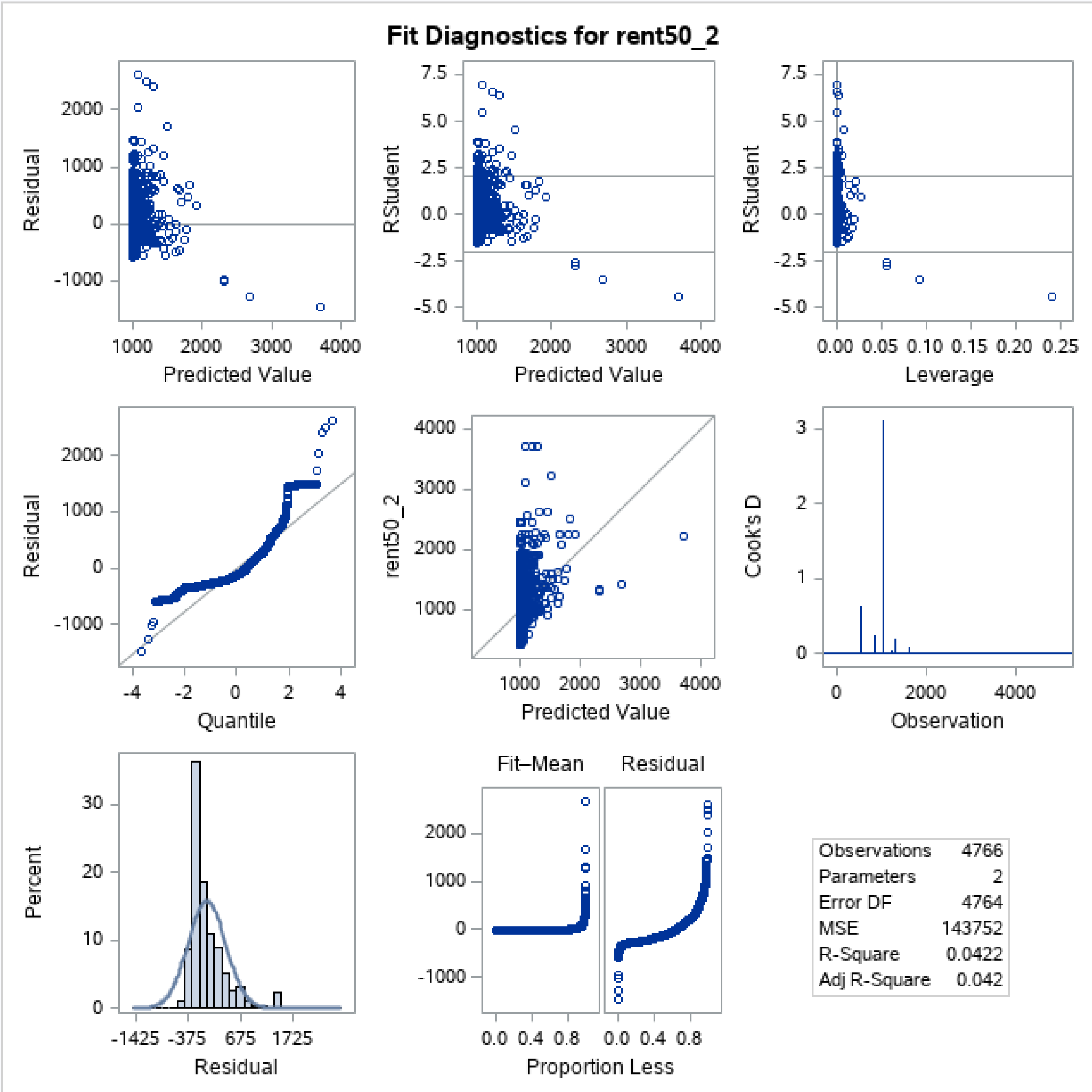
Number of Observations Read	4766
Number of Observations Used	4766

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	30166295	30166295	209.85	<.0001
Error	4764	684835845	143752		
Corrected Total	4765	715002139			

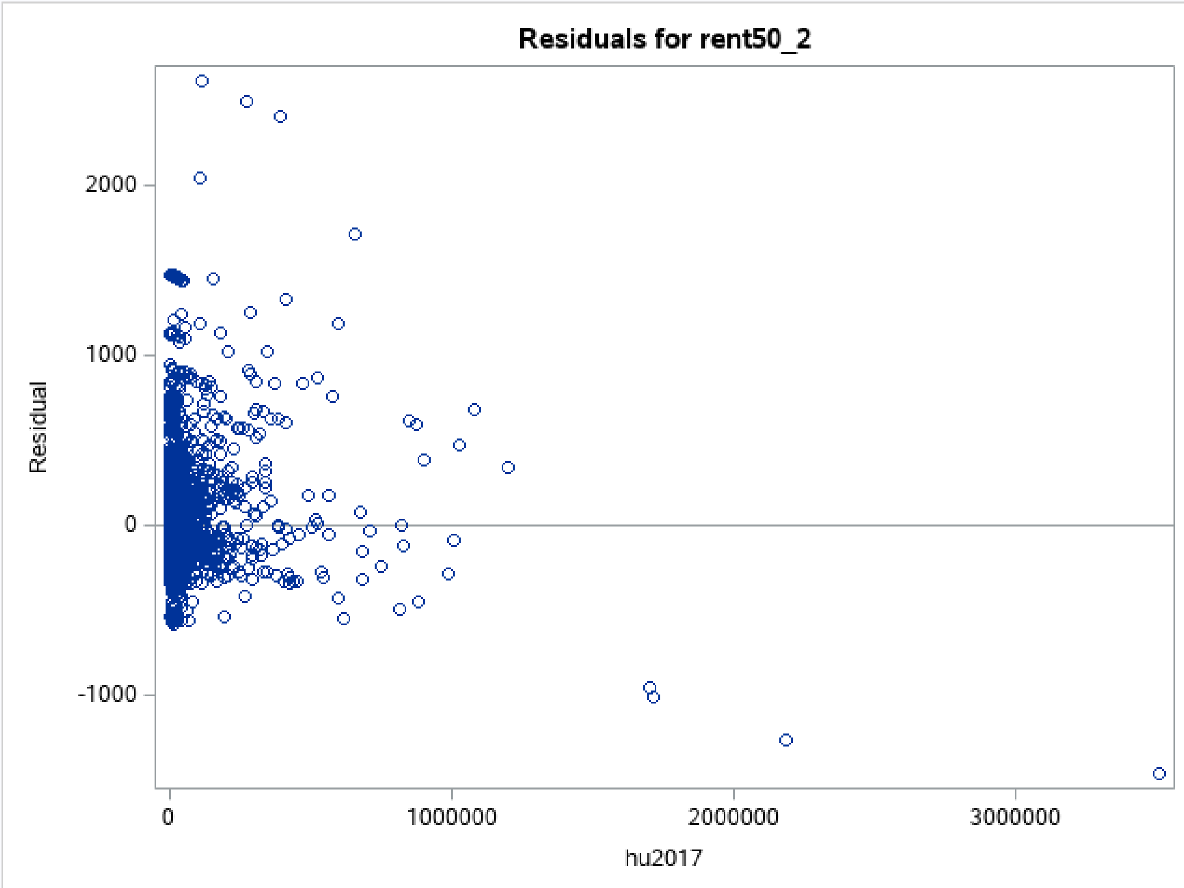
Root MSE	379.14677	R-Square	0.0422
Dependent Mean	1014.65065	Adj R-Sq	0.0420
Coeff Var	37.36722		

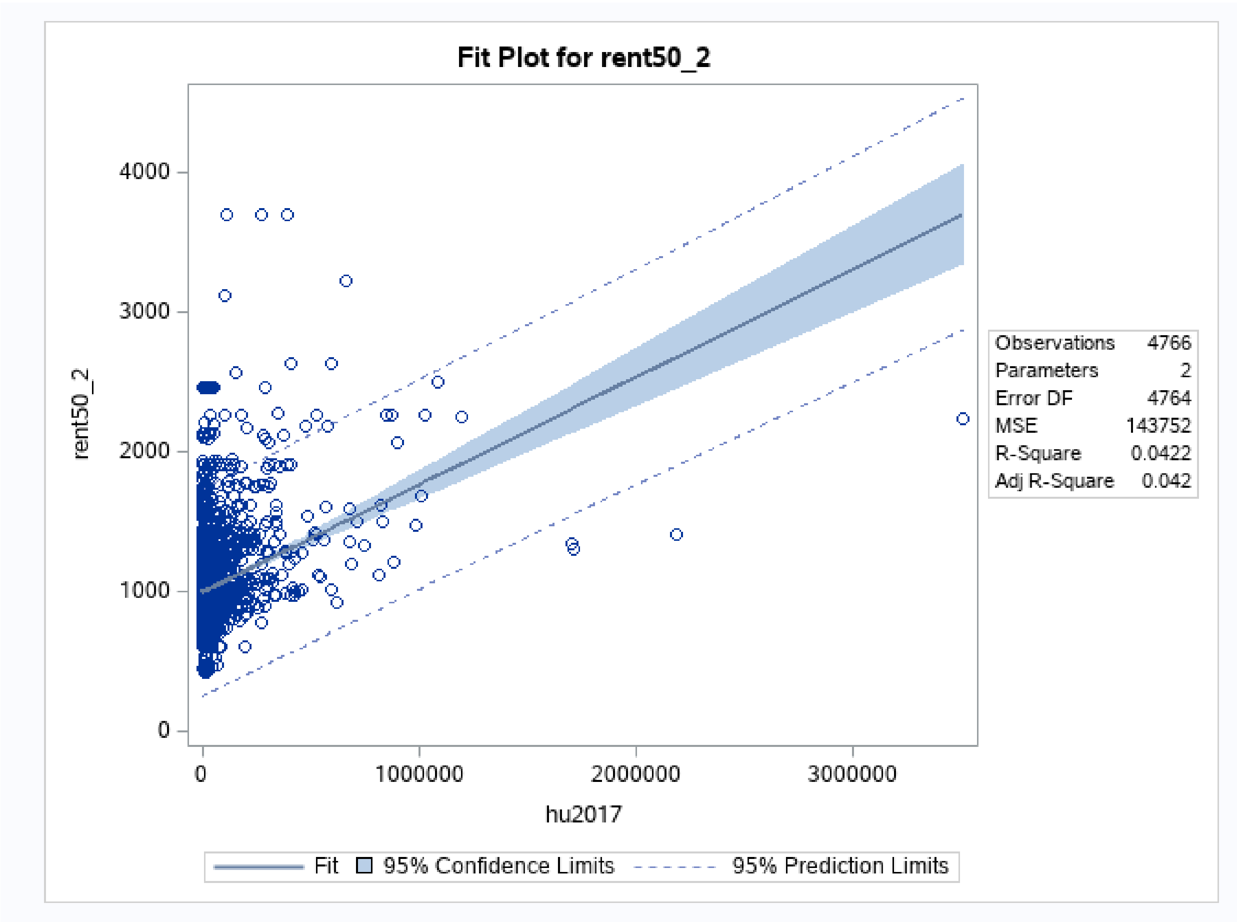
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	992.40167	5.70271	174.02	<.0001
hu2017	1	0.00077344	0.00005339	14.49	<.0001

The REG Procedure
Model: MODEL1
Dependent Variable: rent50_2



STA 3064 PROJECT A





Using the above information, we can note that a model estimate of

$$rent50\_2 = 992.40167 + 0.00077344X + \varepsilon$$

is not a very practical regression model. With an R-Square value of 0.0422, we can also note that the relationship between these two single variables is close to non-existent, despite what would make economic sense. It is, however, interesting to note that the 95% confidence intervals get wider with a larger number of housing units available.

Looking at the residual graphs, this model violates *several* assumptions; predominantly on constant variance and normality, as the observations curve away from a normal distribution, and seem to clump together on top of each other in a line. I would hesitate to say that the relationship is not linear…although it specifically lines up in a vertical pattern that is almost close to having an undefined slope. In economics, this data would still be useful, as it shows housing has extremely inelastic demand that veers close to perfectly inelastic demand (Minnesota State University, 2017), although this model is effectively useless for statistical prediction, and does not meet the statistical definition for linearity, either.

As noted from the previous section with Box-Cox transformations, I used the modified housing data with a transformation of -1 and coded this proc reg statement

```
proc reg data=Housing;
model rent50_2=transHu;
run;
```

to generate the following output:

The REG Procedure
Model: MODEL1
Dependent Variable: rent50_2

Number of Observations Read	4766
Number of Observations Used	4746
Number of Observations with Missing Values	20

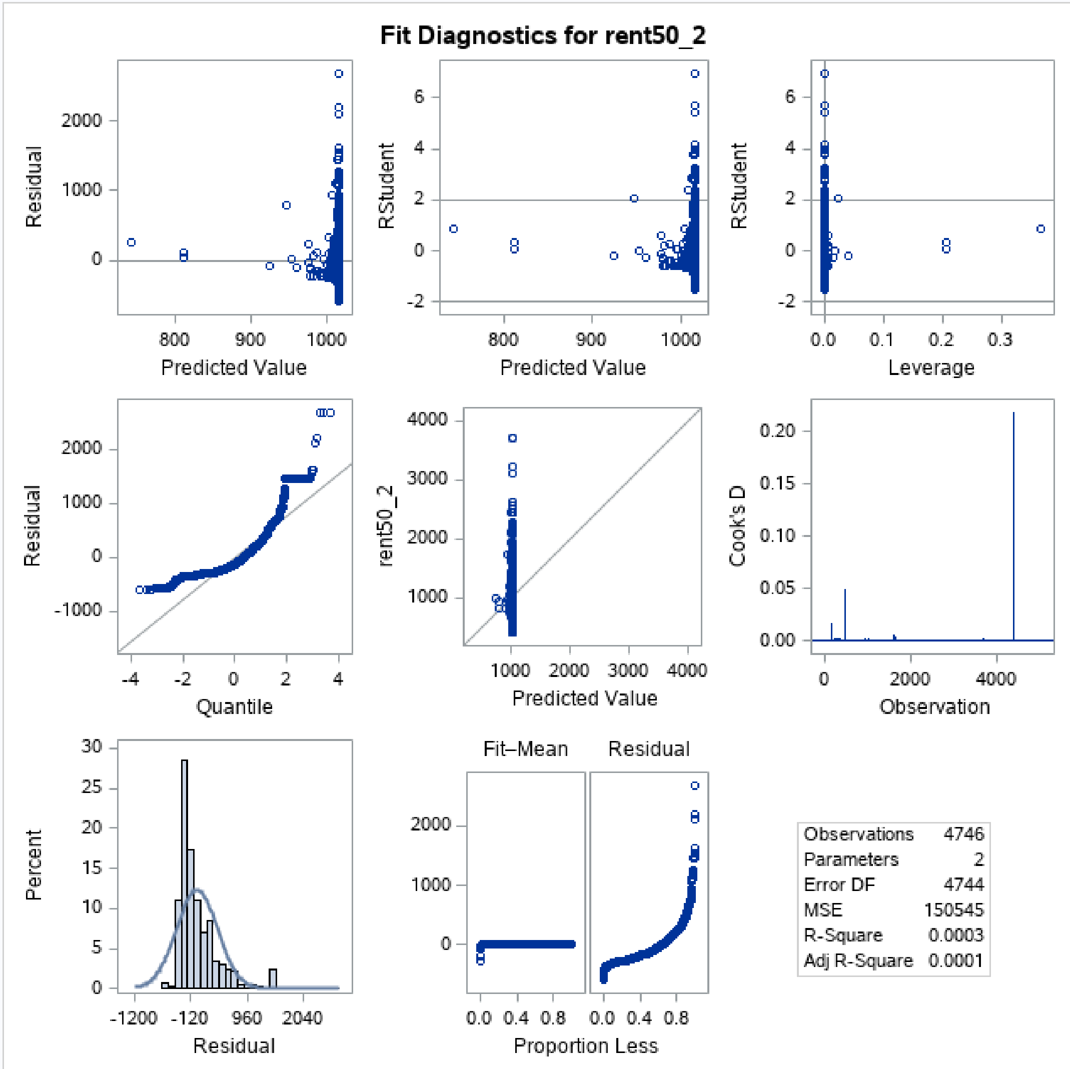
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	203901	203901	1.35	0.2446
Error	4744	714187171	150545		
Corrected Total	4745	714391072			

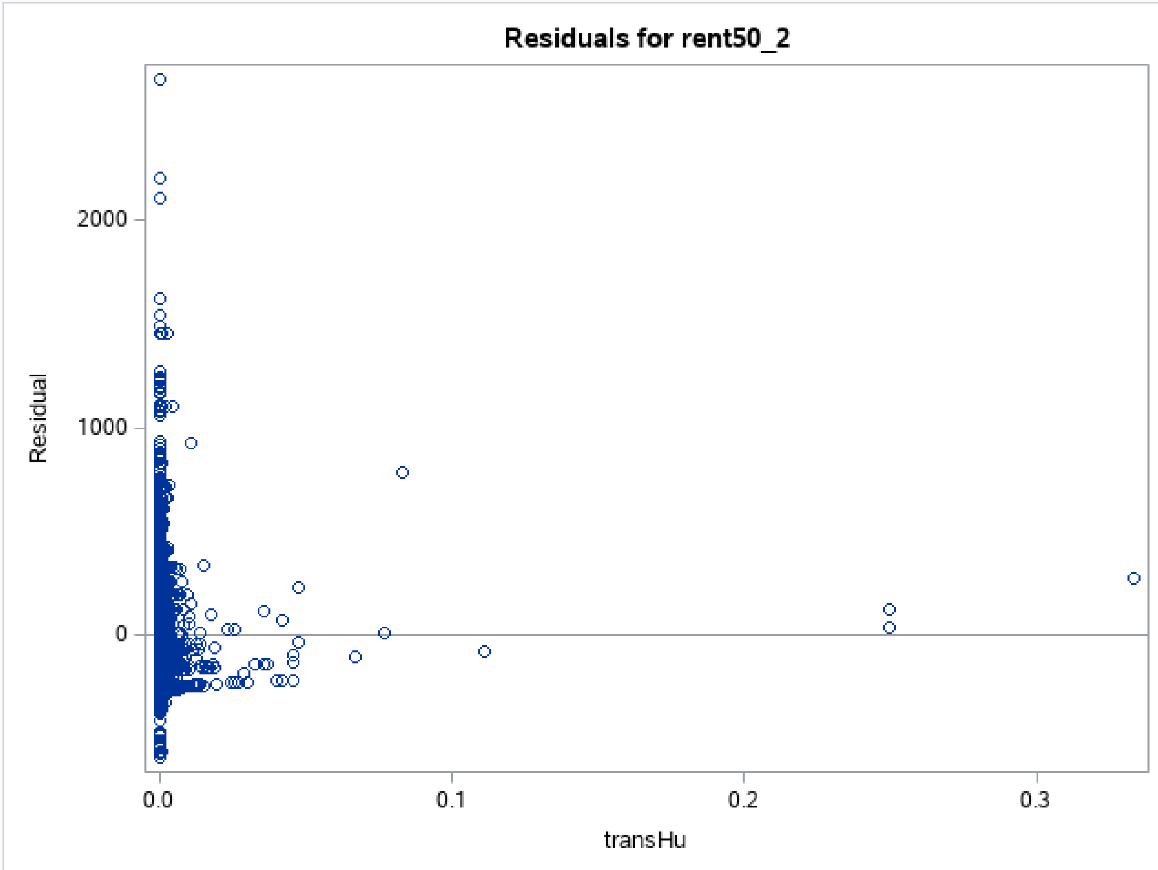
Root MSE	388.00175	R-Square	0.0003
Dependent Mean	1015.28845	Adj R-Sq	0.0001
Coeff Var	38.21591		

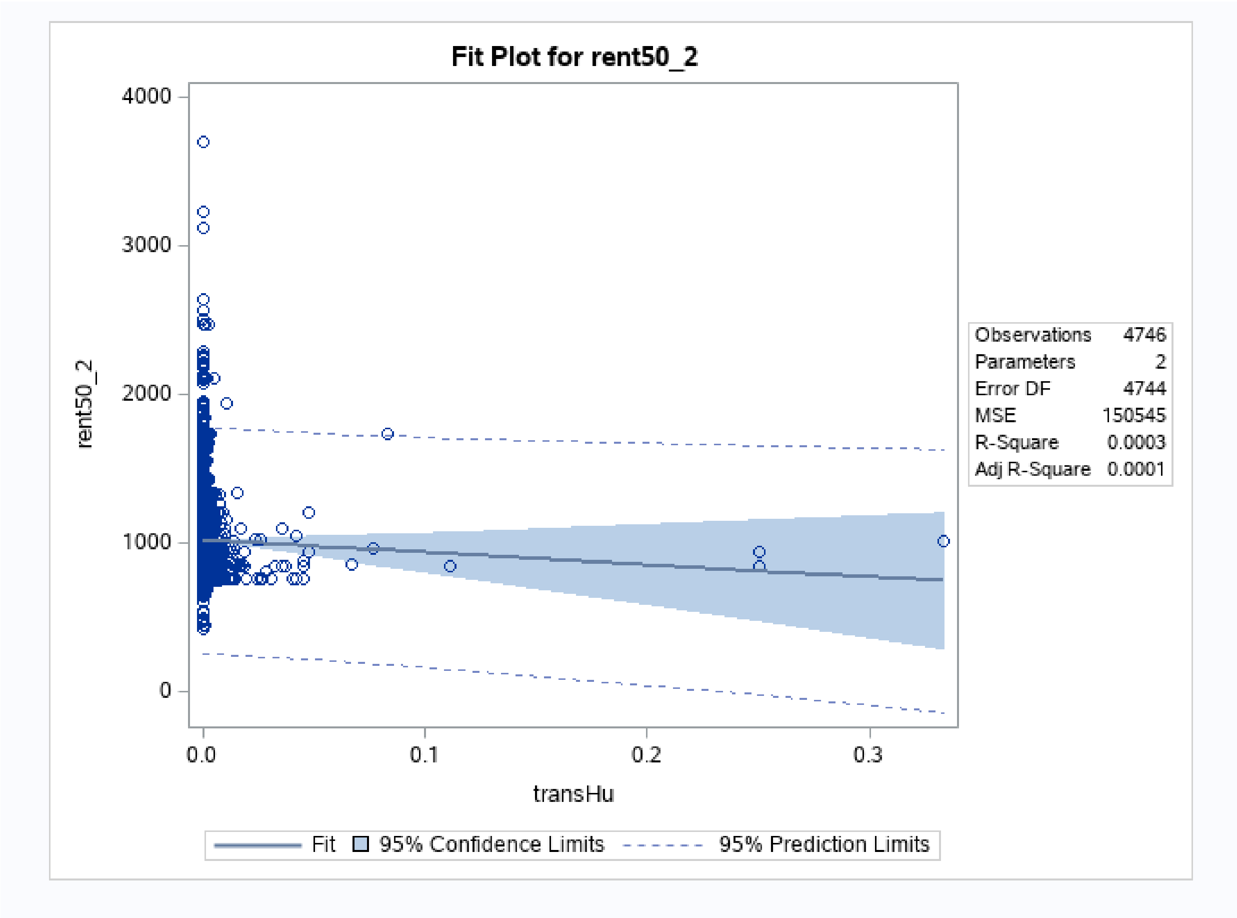
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	1016.11525	5.67672	179.00	<.0001
transHu	1	-821.67499	706.03177	-1.16	0.2446

STA 3064 PROJECT A

The REG Procedure
Model: MODEL1
Dependent Variable: rent50_2




Residuals for rent50_2
Residual
2000
1000
0
0.0
0.1
0.2
0.3
transHu



Surprisingly, the transformation recommended by the Box-Cox output from proc transreg made this model even less useful. The above output leads to the following equation of

rent50_2 = 1016.11525 -821.67499X +ε

that provides practically no use, as it has an R-Square value of 0.0003. I would almost say that there is almost *no relationship* between these two variables when it is transformed like this.

There is *one* model that produced slightly better results than the original "rent50_2=hu2017" model. Namely, using the ratio of population to housing units as a proxy for housing scarcity produced marginally better results, although not by much. By using this code;

proc reg data=Housing;

model rent50_2=PtoH;

run;

I was able to produce the below output:

The REG Procedure
Model: MODEL1
Dependent Variable: rent50_2

Number of Observations Read	4766
Number of Observations Used	4746
Number of Observations with Missing Values	20

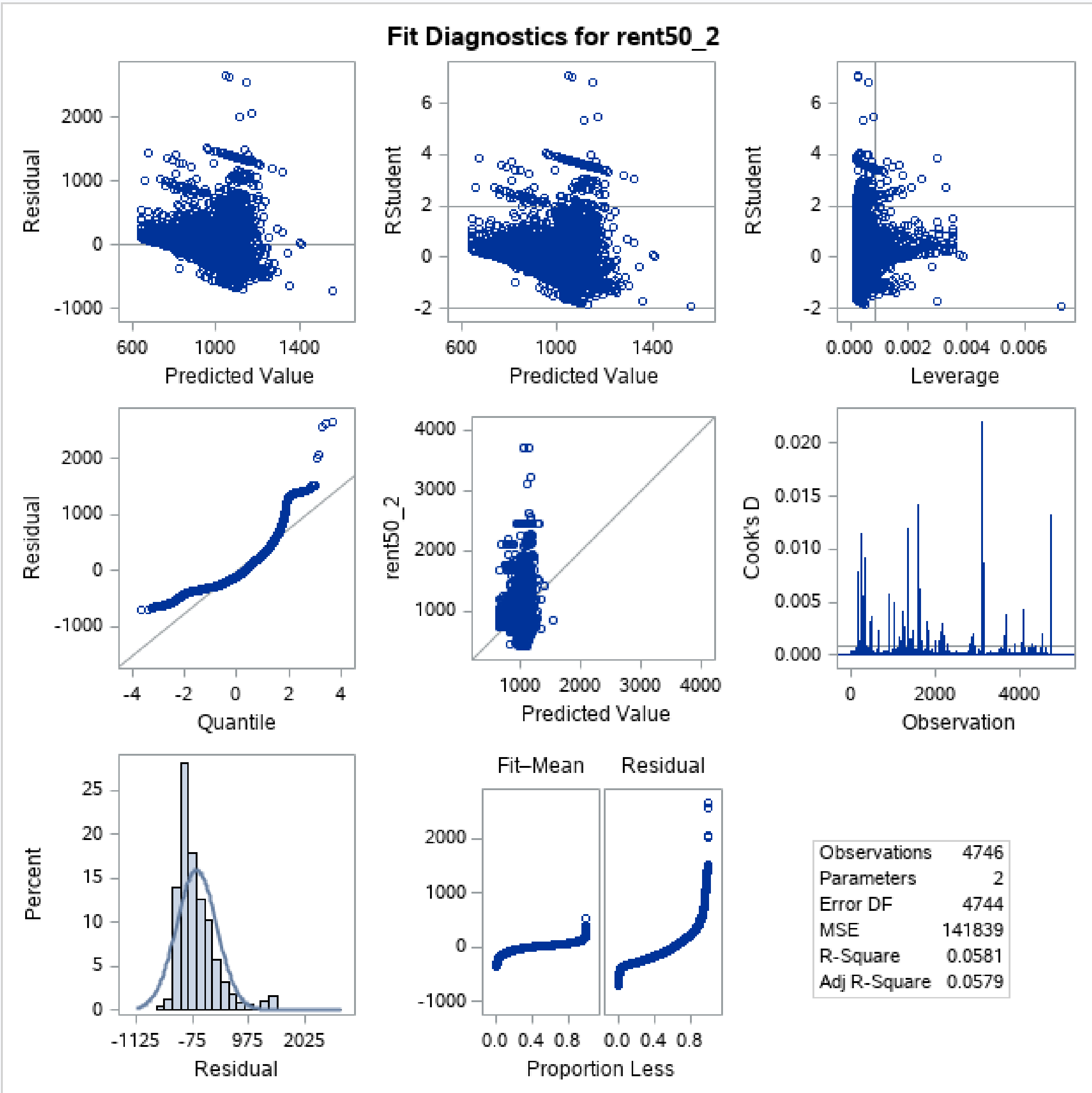
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	41505258	41505258	292.62	<.0001
Error	4744	672885814	141839		
Corrected Total	4745	714391072			

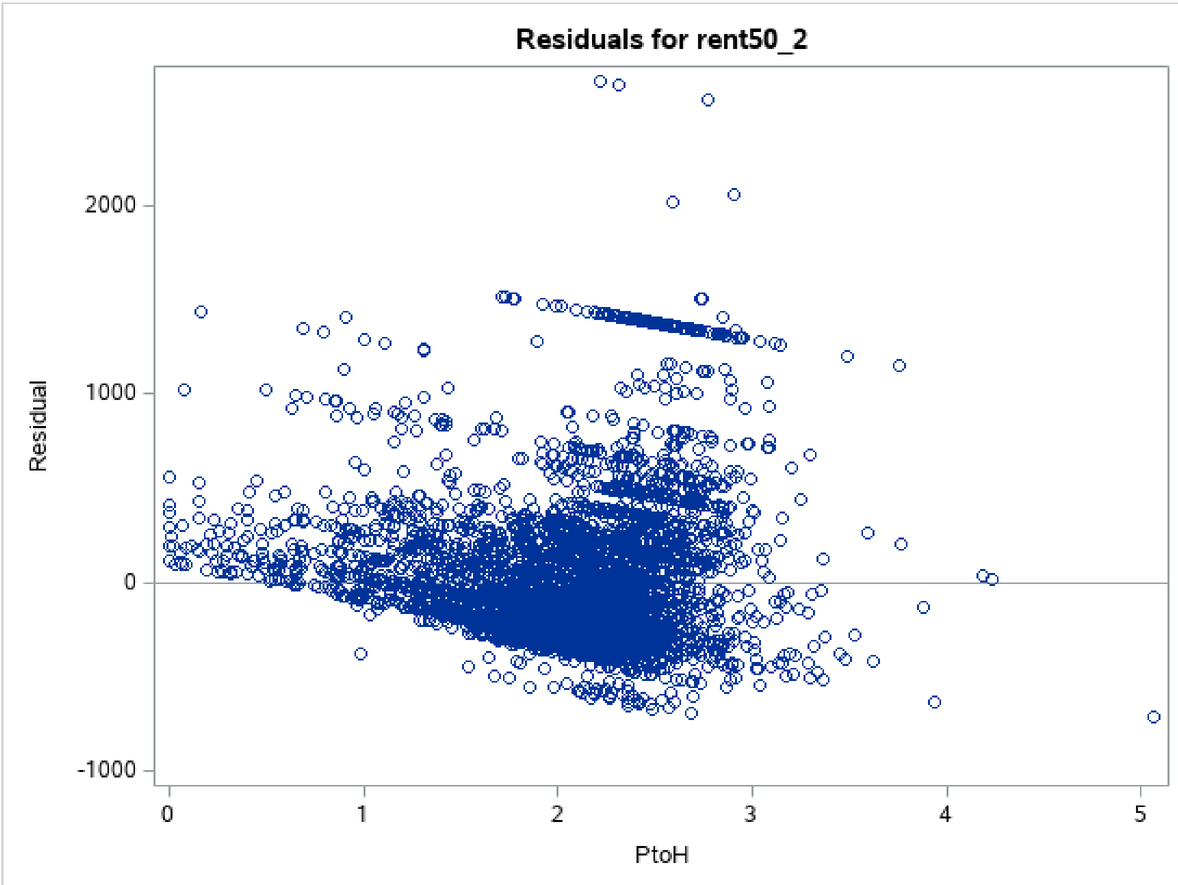
Root MSE	376.61564	R-Square	0.0581
Dependent Mean	1015.28845	Adj R-Sq	0.0579
Coeff Var	37.09445		

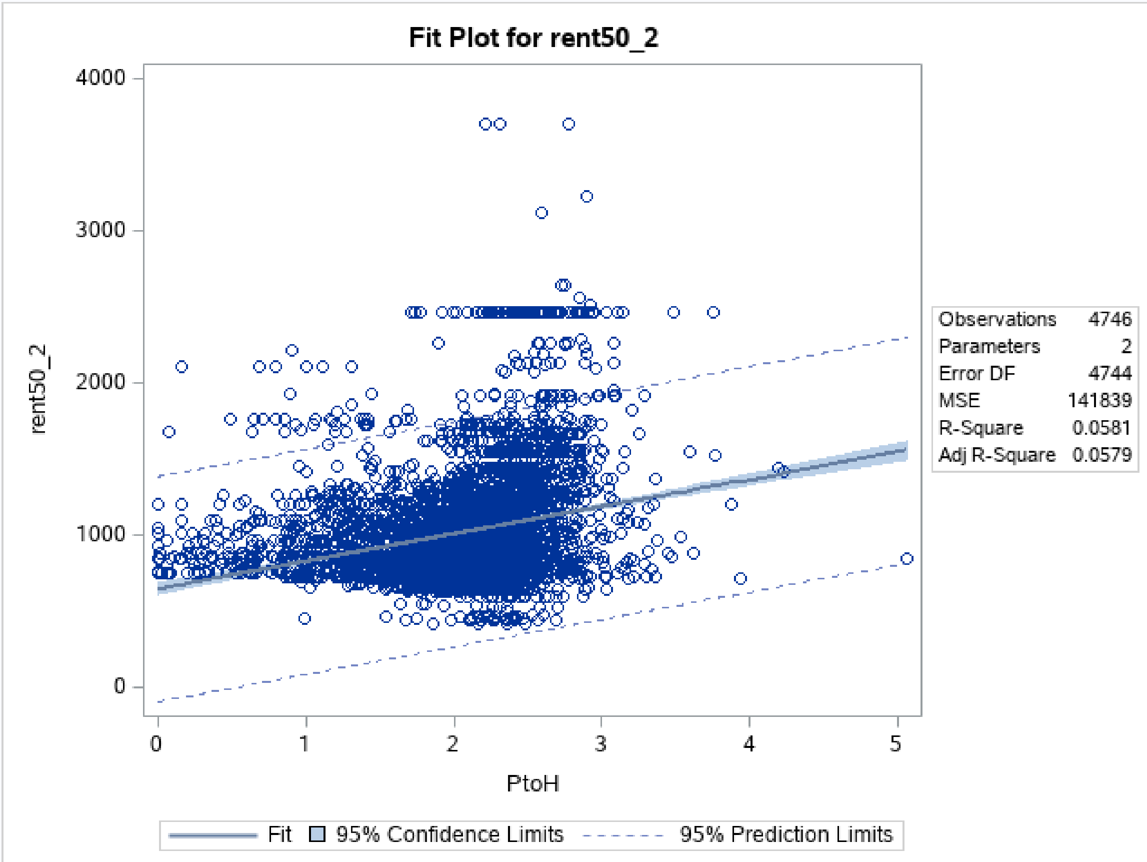
Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	645.80518	22.28048	28.99	<.0001
PtoH	1	179.79938	10.51078	17.11	<.0001

STA 3064 PROJECT A

The REG Procedure
Model: MODEL1
Dependent Variable: rent50_2




Residuals for rent50_2
Residual
2000
1000
0
-1000
0
1
2
3
4
5
PtoH



The above output shows a regression model that follows the below line:

rent50_2 = 645.80518 + 179.79938X +ε

The above output also presents with an ever so slightly more promising R-square value of 0.0581. Additionally, it can be noted that the residuals still show a linear, clumped pattern, but they *are* somewhat more spread out than the other models I was able to generate and fit somewhat closer to a normal distribution than evaluating the quantity of housing units without the context of population sizes…which would model situations where multiple people might be applying for one unit simultaneously, but only one person can get that housing unit.

Lastly, I ran a proc ANOVA data step for both the classic “Price=Quantity” demand curve that would be reflected with “rent50_2=hu2017”, as well as my modified “rent50_2=PtoH” model with the following code:

```
proc ANOVA data=Housing;

class hu2017;

model rent50_2=hu2017;

run;

proc ANOVA data=Housing;

class PtoH;

model rent50_2=PtoH;

run;
```

The above code generated the following output:

The ANOVA Procedure

Class Level Information

Class	Levels	Values
hu2017	4253	0 3 4 9 12 13 15 21 22 24 25 27 28 31 33 35 37 39 41 43 51 54 55 58 59 62 65 66 67 69 71 72 73 75 76 79 81 82 86 87 94 95 97 98 99 100 101 107 109 111 113 114 118 119 120 121 122 125 127 129 130 131 133 134 135 139 140 141 142 143 144 145 148 151 152 155 157 158 159 160 161 163 164 165 178 180 181 182 183 184 186 189 193 195 196 197 199 202 205 209 210 212 213 214 222 224 228 229 230 231 233 234 235 238 241 243 244 246 249 253 256 259 261 262 263 264 265 266 270 274 275 277 278 279 280 282 283 284 285 287 288 289 295 297 298 301 302 303 307 313 314 316 319 320 322 323 327 330 332 334 340 342 343 345 346 348 349 351 353 354 355 356 358 363 365 366 368 370 372 373 374 375 378 379 381 386 389 391 392 394 395 399 400 402 403 405 406 407 409 414 416 418 419 420 421 422 423 426 427 428 429 433 434 435 438 439 440 441 447 448 449 450 452 454 455 456 457 458 459 460 461 462 469 470 473 477 478 479 481 483 484 487 489 490 494 495 497 499 501 506 508 513 514 516 523 524 525 527 528 530 532 533 534 535 536 537 538 539 542 544 546 549 550 551 552 553 555 558 559 563 564 567 569 570 571 572 573 575 576 578 581 582 583 586 587 588 589 592 593 594 596 597 599 600 603 604 605 606 609 611 613 614 616 618 619 620 625 626 627 628 631 632 633 634 636 638 639 641 643 647 651 652 654 658 659 660 662 663 664 666 673 674 676 679 680 683 686 687 688 689 691 693 694 695 699 703 705 707 708 711 713 714 718 719 722 724 725 726 727 728 729 731 732 733 734 736 739 740 743 747 749 750 755 760 761 764 766 767 768 769 770 771 772 778 780 781 784 786 787 788 789 790 791 792 793 794 796 797 799 800 804 806 807 808 810 811 812 814 816 818 822 826 829 833 834 836 837 839 841 842 843 846 847 849 850 851 852 854 855 856 859 860 862 863 867 868 871 872 877 878 879 880 884 889 892 894 897 898 899 901 903 907 908 909 915 916 919 921 922 924 926 927 930 931 935 938 939 940 941 942 943 944 945 946 950 951 958 959 960 962 965 967 968 971 972 976 980 982 983 984 985 987 988 989 990 991 992 993 994 996 997 999 1002 1003 1004 1006 1007 1008 1010 1011 1012 1013 1015 1016 1018 1019 1020 1022 1024 1028 1030 1036 1037 1042 1046 1048 1049 1051 1055 1056 1057 1058 1060 1063 1064 1066 1067 1068 1070 1071 1073 1078 1082 1083 1087 1088 1091 1094 1096 1102 1105 1106 1108 1110 1111 1113 1115 1122 1125 1126 1131 1134 1136 1137 1139 1141 1142 1143 1145 1146 1147 1150 1151 1153 1154 1156 1161 1164 1165 1167 1174 1176 1179 1184 1185 1186 1189 1193 1194 1200 1201 1203 1206 1208 1214 1215 1217 1219 1221 1224 1228 1232 1235 1237 1238 1239 1240 1243 1244 1247 1249 1251 1254 1255 1256 1258 1263 1265 1270 1273 1275 1279 1281 1284 1285 1286 1290 1297 1300 1301 1303 1310 1312 1315 1316 1323 1324 1326 1328 1331 1334 1336 1338 1340 1343 1346 1353 1354 1356 1357 1359 1361 1369 1372 1373 1375 1376 1377 1380 1381 1382 1384 1385 1388 1391 1392 1393 1395 1398 1403 1406 1407 1422 1425 1428 1433 1437 1438 1439 1441 1442 1446 1447 1449 1450 1459 1463 1464 1466 1467 1469 1476 1479 1482 1483 1486 1488 1490 1492 1494 1500 1501 1503 1510 1512 1514 1516 1518 1521 1531 1532 1544 1547 1550 1552 1554 1559 1563 1566 1567 1568 1574 1576 1578 1580 1584 1585 1587 1596 1601 1605 1611 1613 1614 1616 1619 1622 1623 1629 1632 1636 1640 1644 1646 1648 1649 1652 1656 1660 1666 1667 1669 1670 1672 1675 1676 1677 1679 1680 1685 1687 1688 1694 1695 1703 1704 1706 1712 1724 1725 1726 1727 1729 1730 1731 1732 1733 1737 1739 1741 1742 1746 1750 1752 1754 1756 1758 1759 1760 1761 1764 1766 1770 1772 1775 1776 1777 1778 1783 1785 1789 1797 1800 1803 1806 1811 1813 1816 1823 1826 1827 1828 1833 1835 1836 1842 1845 1849 1853 1858 1866 1867 1870 1871 1876 1885 1895 1897 1899 1903 1906 1907 1909 1910 1922 1925 1926 1927 1930 1931 1932 1938 1940 1941 1942 1944 1946 1947 1948 1954 1958 1959 1965 1967 1968 1969 1975 1979 1980 1983 1987 1993 1994 1998 2000 2002 2006 2010 2015 2020 2027 2030 2033 2035 2036 2037 2039 2040 2046 2051 2056 2060 2062 2065 2069 2073 2076 2081 2082 2086 2088 2096 2098 2101 2104 2108 2112 2120 2127 2129 2133 2139 2142 2145 2150 2152 2156 2160 2163 2168 2173 2178 2185 2186 2190 2195 2196 2198 2205 2207 2216 2220 2221 2222 2225 2228 2235 2237 2239 2240 2242 2244 2248 2250 2252 2256 2257 2262 2263 2268 2271 2272 2273 2276 2282 2286 2287 2305 2308 2312 2313 2314 2317 2320 2327 2332 2334 2338 2340 2341 2342 2343 2349 2351 2352 2353 2355 2358 2362 2367 2373 2382 2388 2391 2393 2397 2398 2402 2404 2405 2406 2415 2424 2435 2438 2440 2444 2445 2447 2451 2453 2454 2455 2467 2469 2473 2483 2494 2498 2503 2518 2523 2528 2534 2535 2539 2544 2546 2547 2549 2550 2553 2554 2555 2568 2574 2579 2583 2589 2591 2594 2598 2606 2612 2613 2614 2617 2619 2623 2632 2640 2643 2657 2661 2662 2664 2668 2670 2672 2673 2674 2675 2678 2681 2682 2686 2688 2690 2692 2693 2696 2712 2713 2715 2717 2718 2719 2726 2727 2734 2735 2737 2739 2742 2743 2744 2748 2754 2762 2763 2765 2768 2770 2773 2775 2776 2777 2778 2781 2784 2788 2789 2792 2803 2812 2820 2825 2827 2835 2837 2838 2839 2848 2863 2865 2866 2867 2870 2883 2888 2892 2895 2906 2911 2912 2913 2916 2918 2923 2927 2930 2937 2938 2941 2943 2944 2949 2952 2953 2954 2956 2965 2973 2975 2978 2983 2985 2987 2990 2992 2999 3003 3004 3006 3012 3018 3020 3021 3023 3025 3026 3032 3033 3043 3047 3051 3072 3076 3077 3080 3082 3084 3085 3087 3088 3093 3094 3096 3103 3104 3111 3121 3123 3125 3129 3132 3134 3136 3145 3151 3153 3155 3167 3169 3170 3176 3179 3182 3183 3184 3185 3188 3189 3196 3197 3206 3208 3210 3215 3223 3226 3228 3231 3232 3239 3247 3248 3255 3260 3261 3267 3268 3270 3276 3283 3287 3290 3291 3292 3299 3302 3305 3311 3318 3323 3324 3325 3340 3342 3344 3346 3349 3353 3360 3372 3384 3388 3394 3395 3396 3398 3405 3407 3408 3411 3418 3424 3428 3430 3443 3449 3453 3454 3457 3463 3468 3472 3473 3475 3477 3481 3483 3484 3488 3495 3502 3508 3509 3510 3511 3513 3520 3522 3533 3539 3540 3545 3548 3557 3558 3560 3561 3564 3565 3569 3575 3588 3593 3595 3599 3602 3605 3616 3618 3621 3623 3627 3633 3637 3641 3643 3644 3648 3665 3668 3671 3675 3685 3690 3694 3704 3712 3722 3730 3734 3735 3743 3744 3750 3754 3758 3763 3778 3782 3790 3793 3799 3809 3811 3816 3818 3834 3836 3845 3851 3852 3857 3858 3859 3860 3861 3863 3869 3870 3884 3885 3888 3891 3892 3893 3897 3901 3907 3911 3912 3913 3915 3916 3919 3930 3945 3948 3953 3956 3958 3959 3970 3973 3984 3994 4000 4014 4017 4020 4027 4034 4035 4040 4041 4045 4046 4056 4060 4061 4066 4067 4069 4073 4077 4078 4082 4084 4108 4112 4116 4120 4124 4130 4136 4137 4138 4143 4145 4149 4153 4156 4158 4164 4169 4170 4172 4175 4177 4178 4180 4185 4186 4192 4193 4195 4197 4204 4213 4215 4216 4219 4222 4224 4227 4229 4239 4240 4241 4243 4245 4254 4270 4272 4274 4298 4300 4302 4323 4324 4326 4336 4347 4350 4351 4355 4356 4360 4366 4369 4371 4375 4381 4384 4389 4394 4395 4400 4416 4422 4425 4426 4431 4435 4440 4445 4447 4450 4461 4473 4480 4488 4489 4490 4492 4494 4498 4503 4504 4505 4512 4516 4518 4528 4531 4533 4536 4537 4544 4554 4555 4558 4560 4565 4566 4572 4575 4578 4580 4582 4584 4589 4592 4593 4598 4617 4622 4625 4629 4630 4637 4645 4648 4649 4657 4658 4665 4684 4686 4694 4706 4711 4717 4718 4719 4723 4726 4727 4735 4737 4742 4748 4751 4752 4755 4758 4764 4767 4771 4775 4784 4790 4791 4797 4798 4799 4806 4807 4809 4811 4828 4830 4831 4837 4843 4848 4849 4850 4856 4857 4859 4866 4869 4870 4878 4880 4881 4890 4892 4895 4903 4904 4907 4914 4921 4930 4935 4936 4948 4952 4968 4970 4976 4983 4984 4985 4988 4990 4994 4995 4997 5014 5018 5026 5031 5032 5036 5039 5047 5051 5056 5070 5081 5095 5101 5104 5106 5113 5125 5130 5132 5138 5141 5157 5158 5168 5169 5171 5181 5182 5186 5187 5188 5190 5196 5199 5202 5210 5212 5214 5217 5224 5228 5231 5233 5234 5240 5242 5244 5250 5267 5268 5278 5292 5299 5305 5308 5310 5311 5314 5315 5318 5320 5330 5332 5337 5339 5340 5341 5346 5349 5358 5359 5360 5363 5372 5373 5382 5399 5405 5407 5412 5414 5426 5429 5433 5441 5442 5445 5450 5451 5454 5456 5459 5462 5477 5482 5483 5486 5498 5505 5512 5519 5531 5541 5543 5548 5550 5558 5560 5564 5565 5571 5579 5586 5587 5598 5602 5610 5624 5626 5642 5645 5655 5659 5678 5679 5683 5693 5698 5702 5709 5711 5717 5721 5732 5736 5753 5756 5760 5762 5766 5775 5776 5782 5786 5794 5795 5805 5818 5821 5823 5825 5826 5829 5831 5833 5835 5836 5843 5871 5878 5880 5882 5886 5892 5894 5898 5906 5907 5920 5927 5936 5937 5938 5939 5950 5953 5954 5957 5964 5968 5980 5981 5983 5984 5990 5995 6001 6007 6008 6027 6032 6035 6043 6044 6049 6050 6054 6065 6084 6096 6100 6104 6107 6109 6111 6113 6123 6126 6140 6143 6144 6153 6154 6164 6169 6176 6177 6179 6189 6210 6221 6236 6246 6260 6261 6263 6267 6273 6286 6301 6304 6307 6309 6310 6317 6321 6329 6330 6336 6344 6347 6356 6362 6368 6376 6379 6393 6401 6404 6405 6412 6413 6415 6433 6434 6441 6457 6461 6466 6471 6474 6479 6493 6494 6499 6504 6507 6519 6527 6531 6538 6541 6543 6546 6558 6562 6565 6571 6573 6587 6596 6600 6618 6619 6623 6634 6635 6637 6652 6661 6662 6672 6680 6682 6689 6702 6714 6724 6725 6731 6732 6736 6741 6743 6749 6754 6758 6760 6763 6770 6772 6775 6781 6784 6797 6808 6811 6813 6823 6827 6833 6839 6841 6843 6846 6854 6867 6875 6878 6891 6903 6913 6922 6958 6959 6960 6968 6970 6985 6999 7004 7011 7045 7047 7049 7050 7053 7057 7058 7062 7063 7084 7086 7095 7096 7105 7125 7134 7135 7139 7149 7172 7184 7185 7188 7189 7201 7206 7208 7211 7214 7224 7232 7238 7252 7254 7263 7271 7279 7281 7285 7292 7293 7294 7299 7301 7303 7306 7310 7311 7312 7318 7319 7329 7332 7333 7335 7338 7347 7354 7357 7358 7363 7365 7369 7384 7389 7390 7407 7421 7428 7434 7436 7444 7448 7455 7474 7484 7491 7496 7497 7504 7505 7506 7508 7512 7537 7539 7541 7542 7550 7553 7554 7558 7565 7577 7581 7586 7587 7589 7591 7592 7605 7606 7617 7621 7627 7630 7632 7659 7669 7671 7675 7687 7688 7691 7694 7700 7705 7712 7716 7718 7741 7748 7753 7754 7760 7767 7771 7778 7787 7798 7800 7802 7805 7807 7823 7834 7853 7855 7880 7900 7901 7903 7918 7926 7934 7937 7942 7947 7948 7957 7964 7968 7971 7973 7981 7989 7990 7993 7994 8003 8016 8020 8035 8047 8055 8058 8065 8071 8074 8080 8087 8091 8092 8110 8112 8128 8143 8146 8152 8155 8161 8163 8171 8177 8178 8179 8199 8202 8203 8217 8224 8229 8233 8246 8260 8264 8275 8278 8289 8290 8308 8310 8320 8324 8328 8334 8346 8370 8397 8420 8429 8431 8434 8435 8453 8456 8458 8481 8490 8503 8508 8528 8530 8532 8536 8547 8558 8566 8571 8575 8577 8583 8584 8591 8599 8600 8616 8618 8619 8621 8627 8637 8645 8653 8655 8658 8671 8684 8693 8695 8713 8726 8727 8734 8739 8750 8757 8759 8800 8808 8813 8814 8817 8823 8843 8846 8853 8870 8872 8882 8885 8886 8895 8903 8908 8909 8913 8917 8924 8930 8931 8941 8945 8952 8956 8958 8965 8971 8973 8975 8998 9000 9016 9024 9025 9033 9036 9040 9049 9055 9061 9065 9067 9074 9078 9097 9105 9112 9121 9128 9131 9141 9152 9156 9165 9173 9184 9185 9201 9203 9212 9213 9218 9221 9230 9232 9234 9235 9250 9254 9257 9263 9282 9297 9303 9305 9316 9317 9318 9319 9324 9325 9326 9331 9344 9347 9350 9357 9369 9372 9377 9383 9385 9403 9414 9417 9422 9447 9452 9454 9460 9465 9466 9489 9490 9492 9502 9513 9519 9524 9525 9526 9531 9538 9565 9582 9583 9593 9601 9629 9645 9651 9668 421152 425716 427292 439403 453948 456265 471031 487519 506468 517631 524488 525186 535380 546040 563638 564349 574342 596504 596898 619305 658409 678673 679876 683786 711900 748333 815472 818382 826704 850422 872645 877617 902107 985738 1008908 1028383 1081701 1195868 1699628 1714340 2183987 3506903

Number of Observations Read	4766
Number of Observations Used	4766

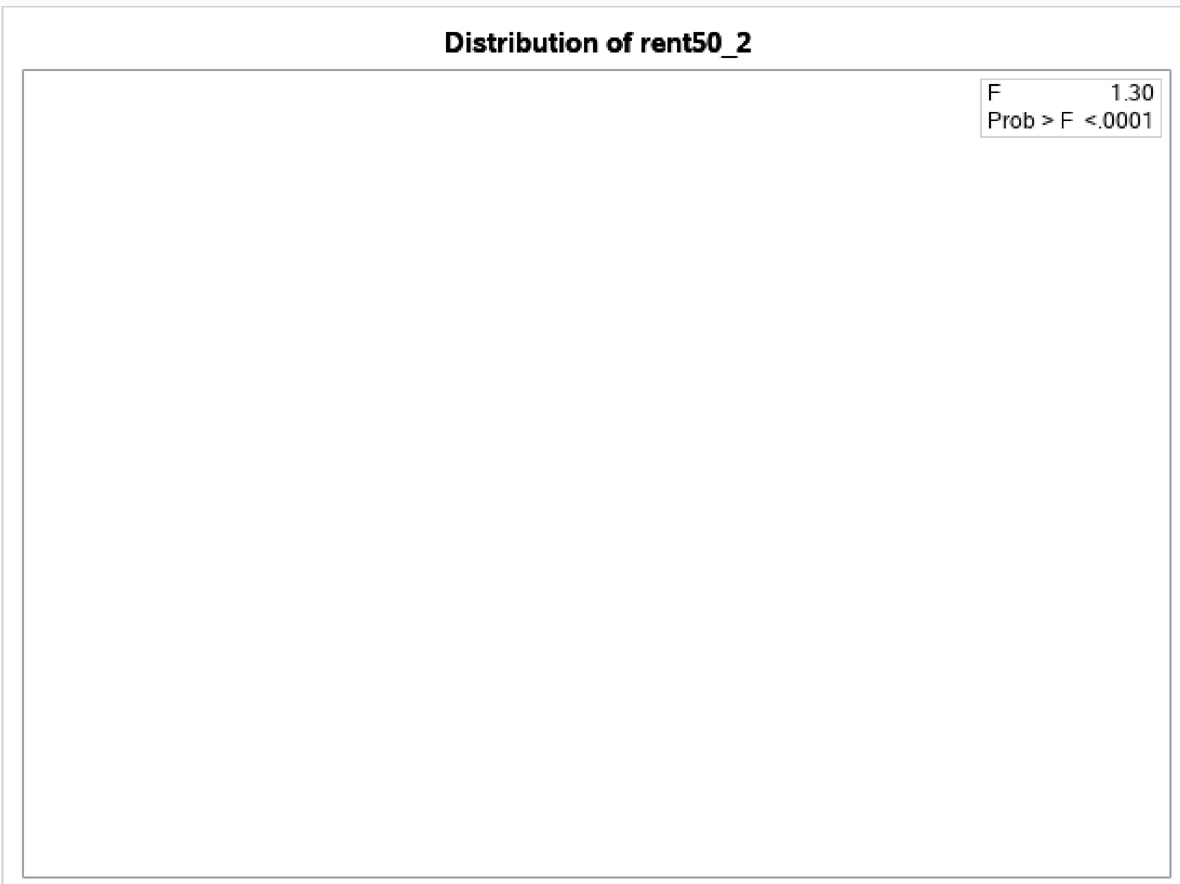
The ANOVA Procedure

Dependent Variable: rent50_2

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4252	654451619.9	153916.2	1.30	<.0001
Error	513	60550519.5	118032.2		
Corrected Total	4765	715002139.3			

R-Square	Coeff Var	Root MSE	rent50_2 Mean
0.915314	33.85975	343.5581	1014.651

Source	DF	Anova SS	Mean Square	F Value	Pr > F
hu2017	4252	654451619.9	153916.2	1.30	<.0001



The ANOVA Procedure

Class Level Information	
Class	Levels
PtoH	4704

Number of Observations Read	4766
Number of Observations Used	4746

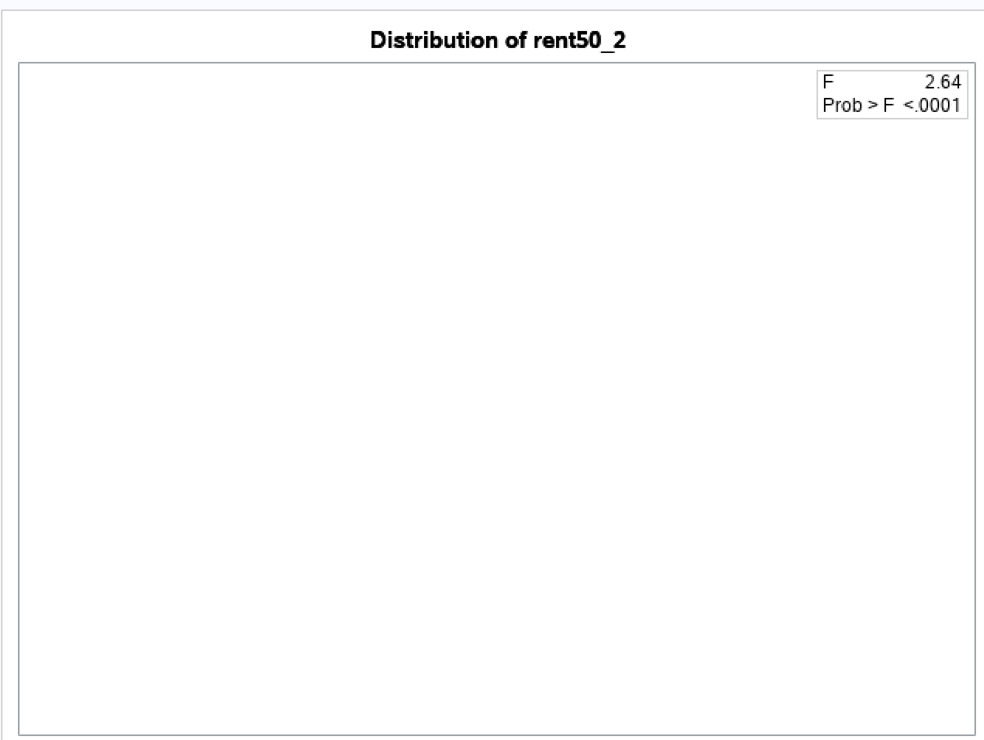
The ANOVA Procedure

Dependent Variable: rent50_2

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	4703	711982304.2	151389.0	2.64	<.0001
Error	42	2408767.9	57351.6		
Corrected Total	4745	714391072.1			

R-Square	Coeff Var	Root MSE	rent50_2 Mean
0.996628	23.58758	239.4820	1015.288

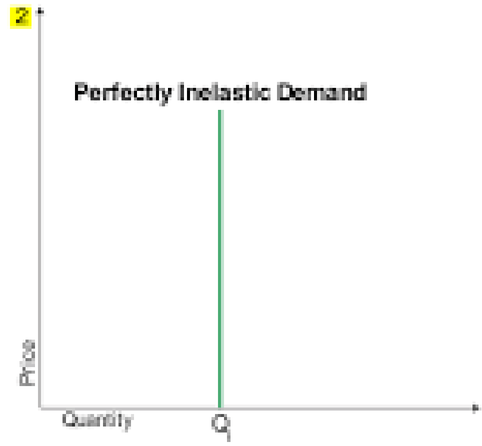
Source	DF	Anova SS	Mean Square	F Value	Pr > F
PtoH	4703	711982304.2	151389.0	2.64	<.0001



Due to the *contradictory* ANOVA output, I am honestly unsure as to whether I should reject the null hypothesis or not. The F-Test statistics are 1.30 and 2.64 for the first and second model, which would imply to keep the null hypothesis that there is no relationship for now (which would make sense, given the other metrics). However, the p-values for the ANOVA tests are both <.0001, which would imply to reject the null hypothesis in favor of the alternative hypothesis that there is a statistically meaningful relationship of rent to housing quantity.

Conclusion

While the data applied to a simple regression model does *not* support the traditionally understood economic idea that prices are caused by quantity alone, it does yield other interesting insights. Namely, the distribution of the observations and residuals seems to imply that the demand for housing is close to perfectly inelastic.



In layman's terms, without other factors to consider such as house quality (which is wildly variable, considering that everything from studio apartments to mansions exist), region, proximity to work, etc; people will pay nearly anything for shelter. This would still make economic sense, considering that necessities such as food and healthcare are typically inelastic, while goods such as video games, luxury goods, and other non-essentials typically have more elastic demand. As shelter is one of the core things necessary for human survival and a basic standard of living, it would make sense that it also has an extremely inelastic demand akin to food and healthcare. While this observation is neutral, it can be interpreted in a vast number of ways on the political spectrum, which can and does impact economic policy regarding affordable housing.

Additional Notes

*The original dataset from HUD was slightly modified to account for SAS being unable to read 'ñ', 'ó', 'í ', 'á', or 'ü' in context of several observations of Puerto Rican regions without additional coding— the following characters were changed to 'n', 'o', 'I', 'a', and 'u' in the dataset for the sake of SAS being able to read the observations. (Although I would like to fix this for the sake of language integrity, the documentation that detailed fixing this would probably take more time to learn than one week — especially given how complex the dataset is for analysis purposes as is. Thus, I deemed it more time efficient to change the special characters in MS Excel.)

References

Adamczyk, A. (2021, July 14). *CNBC*. Retrieved from cnbc.com: https://www.cnbc.com/2021/07/14/full-time-minimum-wage-workers-cant-afford-rent-anywhere-in-the-us.html

Hayes, A. (2021, August 01). *Investopedia*. Retrieved from investopedia.com: https://www.investopedia.com/terms/e/elasticity.asp

Minnesota State University. (2017). *Minnesota State University*. Retrieved from http://mnstate.edu: http://web.mnstate.edu/stutes/Econ202/Econ202/Fall%202017/study3.htm

Simone, J. (2021, July 20). *Coalition for the Homeless*. Retrieved from https://www.coalitionforthehomeless.org/: https://www.coalitionforthehomeless.org/todays-read-out-of-reach-2021/

US Housing and Urban Development - Office of Policy Development & Research. (2021, January 01). *HUDUser*. Retrieved from HUDUser.gov: https://www.huduser.gov/portal/datasets/50per.html#2022