Train Linear Regression model to Predict House Price

Analysis (or Training) Set of Data:

1. **a. Check simple Y*X plots for each X to see if the variables should be transformed to either the log or the square root scale in order to make the points spread out along a line.**

As can be seen from Figure 1 before transformation, apparently there is none of explanatory variable values versus dependent variable fall close to a line.
As can be seen from Figure 2 for log transformation, some explanatory variable values versus dependent variable fall close to a line, but some variables are hard to examine linearity because their data are squeezed on the right hand side
As can be seen from Figure 3 for square roots, all the explanatory variable values versus dependent variable fall close to a line.

Hence, we choose **square root** transformation.

b. **Once satisfied with the spread of the observations, find outliers in X, outliers in Y and influential points. You do not need to remove them just list a few of the ones found and indicate how you determine the points were outliers and/or influential.**

- Determining outliers of Y:
  (Studentized Y Values): Any value larger than t-table value with n-k- 1 degrees of freedom was an outlier for the Y value.
- Determining outliers of X:
  Any value above 2*(k-1)/n are considered outliers.
- Determining influential points:
  (Cook's Distance:) Any value above the 50th percentile of an F distribution k+1 and n-k- 1 degrees of freedom are considered influential.

Since k=10,n=300, using MS Excel, TINV(0.05,300-10-1) = 1.968206,2*(10-1)/300=0.06,FINV(0.05,10+1,300-10-1)= 1.821867.
So,

- Any values above 1.968206436 will be considered as an outlier of the Y value.

(Figure 4 gives a screenshot of a list of student_sqrt_SalePrice.)

student_sqrt_SalePrice
5.7744345
4.000663051
3.90058972
3.397488624
2.971691861
2.263473918
2.16545844

- Any values greater than 0.06 will be considered as an outlier of X, Figure 5 shows a list of h_sqrt_SalePrice.
- Any values above 1.821867 will be considered as influential point, Figure 6 shows a list of cookd_sqrt_SalePrice values (only 0.355927987).

cookd_sqrt_SalePrice

0.355927987

2. Check for assumption violations

a. if necessary use transformation to fix any assumptions that you note. Explain why you did or not find any assumption violations.

Residual Assumption check :

- Residual plots – If the residual plots show a random scattering of points, then we can conclude that we there is no violation assumption. If any patterns shows, then there will be possibly a violation against an assumption.
- Normality probability plots – If the normality probability plots show a straight line, then we can conclude that there is no violation assumption. If there is any other shape, then there might be possibly an assumption violation.

Before any transformation:
As can be seen from Figure 7, the points in QQ-plot of Residual for SalesPrices seems to deviate from the regression line , so there is a violation against normality. Some residual by Regressors for SalesPrice plots contain several separate cluster of points, do not show a random scattering of points, So we consider they reveal a residual plot violation.
So it is **necessary use transformation to fix the** assumptions.

b. check assumptions again after your transformation

After our **square root** transformation:
As can be seen from the Figure 8, the **square root** transformation helped to reduce the assumption violations against normality, the points in QQ-plot of Residual for sqrt_SalePrice seems more fall closer to a regression line than plot before transformation did . Although a violation of the normality assumption is possible, considering the big data sample size (n=300), we can accept the assumptions.
In Residual plots, Some chart in Regressors for SalesPrice plots seems to still contain several separate cluster of points, however, they seems be less extreme than the clusters from the original dataset.

c. comment on whether the assumption transformations helped or not. If they do not help, then revert back to the original variables.

The square root transformation help to reduce the violation the normality assumption and residual plot assumption, so we will accept the square root transformation the will not revert back to the original variables.

3. Using all possible regression, reduce the list of possible models down to three, explaining your reasoning for choosing those three models. Remember you should not always choose the model that has all the variables since you want a simple model that still adequately explains the variation in Y.

Models:

Model	Model Index	# in Model	R-Square	Variables in Model
1	176	4	0.796	sqrt Overall_Qual sqrt Year_Built sqrt Gr_Liv_Area sqrt Garage_Area
2	56	3	0.7692	sqrt Overall_Qual sqrt Year_Built sqrt Gr_Liv_Area
3	1	1	0.5738	sqrt Overall_Qual

Reason:

These models have the highest Square and least amount of variables in each models.

R-Square(as known as the Coefficient of Determination)is a number that indicates how well the data fits a statistical model.

An R-square of 1 indicates that the regression line perfectly fit the data, while a R-Square of 0 indicates that the line does not fit the data at all.

We chose model 1 with R-Square of 0.796 and 4 variables in the model as the model we would like to evaluate more deeply.

Figure 9 given show all the possible linear Regression, and selected three models.

4. Examine the three models in detail and choose one model. Explain why you choose that model and write down the least squares line for that model.

We chose model 1 with R-Square of 0.796 and 4 variables(sqrt Overall_Qual, sqrt Year_Built, sqrt Gr_Liv_Area, sqrt Garage_Area) in the model as the model we would like to evaluate more deeply.

Train Equation: predicted square root = -2629.84904 + 88.08128 (sqrt Overall_Qual) + 58.90220 (sqrt Year_Built) + 4.39973 (sqrt Gr_Liv_Area) + 2.00940 (sqrt Garage_Area)

5. Using the least squares equation found in 4 from your training set, predict the values of the dependent variable in your validation data set. Average the absolute differences between the actual value and the predicted value. Discuss the average error in prediction you would expect in the future using this model.

Validation Equation:

predicted square root = -2426.29334 + 107.00913 (sqrt Overall_Qual) + 54.04595 (sqrt Year_Built) + 3.69960 (sqrt Gr_Liv_Area) + 1.65293 (sqrt Garage_Area)

House Sales Price (Dep Var):

House Sales Price (Dep Var): see Figure 10 for the House Sales Prices calculated using the Validation Equation.

Example is shown below.

fx =-2426.29334+107.00913*(N3)+54.04595*(P3)+3.6996*(S3)+1.65293*(V3)

	S	T	U	V	W	X	Y	Z	AA
Flr_SF	sqrt Gr_Liv_Area	sqrt Fireplaces	sqrt Garage_Cars	sqrt Garage_Area	sqrt Yr_Sold	sqrt SalePrice	Pred Sqrt (Sales Price)	Pred (Sales Price)	Abs Diff
29.33	41.47	-	1.41	23.77	44.79	443.28	$ 468.97	$ 219,936	$ 23,436
28.46	36.44	-	1.00	14.49	44.79	327.11	$ 317.75	$ 100,968	$ 6,032
28.35	41.01	-	1.41	22.87	44.83	447.77	$ 441.15	$ 194,617	$ 5,883

Absolute differences are shown in Column AA of the Figure 10(Validation).
The Average of the absolute differences are shown in Column AC of the Figure 10(Validation).

Calculation shown below.

ce)	Abs Diff
,936	$ 23,436
,968	$ 6,032
,617	$ 5,883
,308	$ 32,808
,557	$ 19,943

Average	$ 16,220
Std Dev	$ 15,533

=AVERAGE(AA:AA)

	S	T	U	V	W	X	Y	Z	AA	AB	AC	AD
_SF	sqrt Gr_Liv_Area	sqrt Fireplaces	sqrt Garage_Cars	sqrt Garage_Area	sqrt Yr_Sold	sqrt SalePrice	Pred Sqrt (Sales Price)	Pred (Sales Price)	Abs Diff		Average	=AVERAGE(A
.33	41.47	-	1.41	23.77	44.79	443.28	$ 468.97	$ 219,936	$ 23,436		Std Dev	$ 15,533
.46	36.44	-	1.00	14.49	44.79	327.11	$ 317.75	$ 100,968	$ 6,032			
.35	41.01	-	1.41	22.87	44.83	447.77	$ 441.15	$ 194,617	$ 5,883			
.67	45.37	1.00	1.41	19.90	44.81	359.86	$ 402.87	$ 162,308	$ 32,808			
.45	32.45	-	1.00	17.66	44.81	381.44	$ 354.34	$ 125,557	$ 19,943			
.64	42.39	-	1.41	25.28	44.80	462.06	$ 476.07	$ 226,643	$ 13,143			

Average Error in Prediction:
When you predict the price of a house, it could be ~15K (1 standard deviation) better than expected or ~15K worse than expected.

6. Using the validation data set, see if the model found in 4 is useful.

Train Equation:
predicted square root = -2629.84904 + 88.08128 (sqrt Overall_Qual) + 58.90220 (sqrt Year_Built) + 4.39973 (sqrt Gr_Liv_Area) + 2.00940 (sqrt Garage_Area)

Validation Equation:
predicted square root = -2426.29334 + 107.00913 (sqrt Overall_Qual) + 54.04595 (sqrt Year_Built) + 3.69960 (sqrt Gr_Liv_Area) + 1.65293 (sqrt Garage_Area)

Our validation equation is closely related to our train equation so we believe the model 1 is useful in predicting the future house prices.

Appendix:

Figure 1(plot before transformation):

Figure 2(log transformation plot):

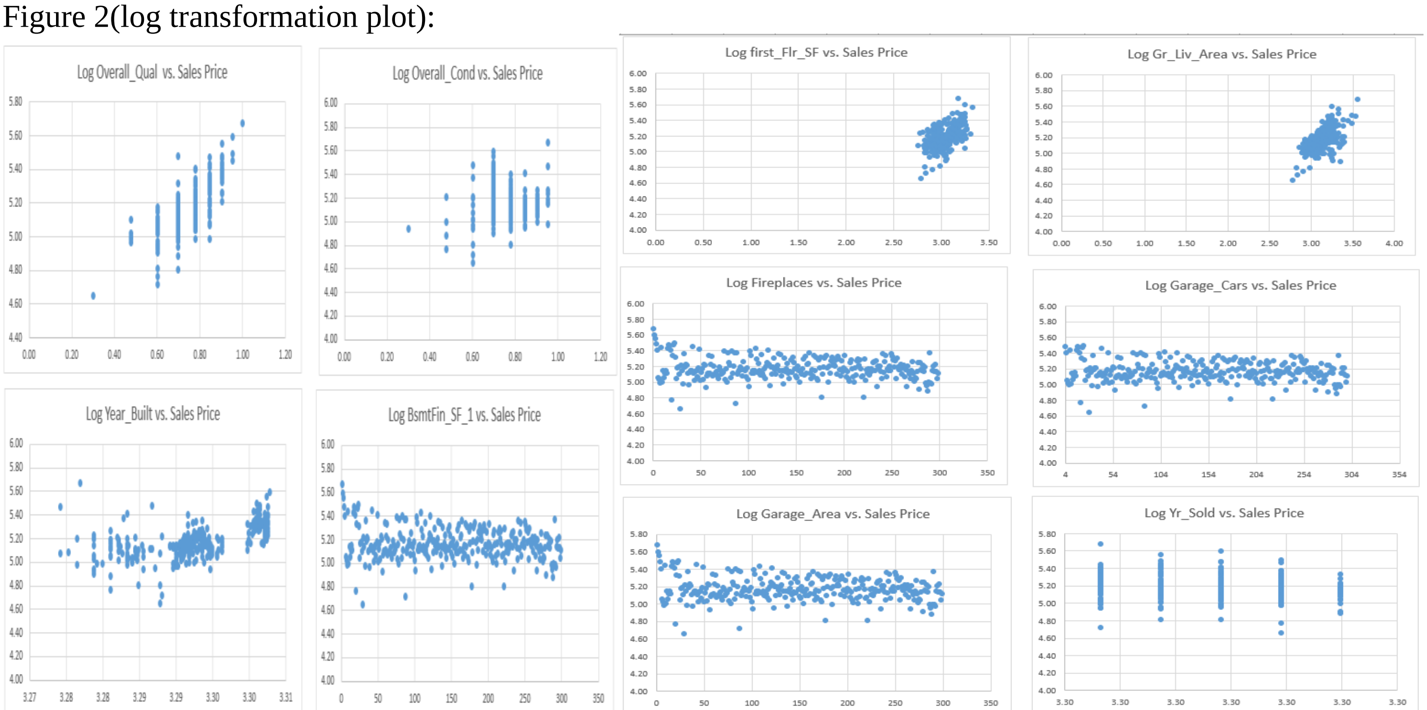


Figure 3(square roots transformation plot):

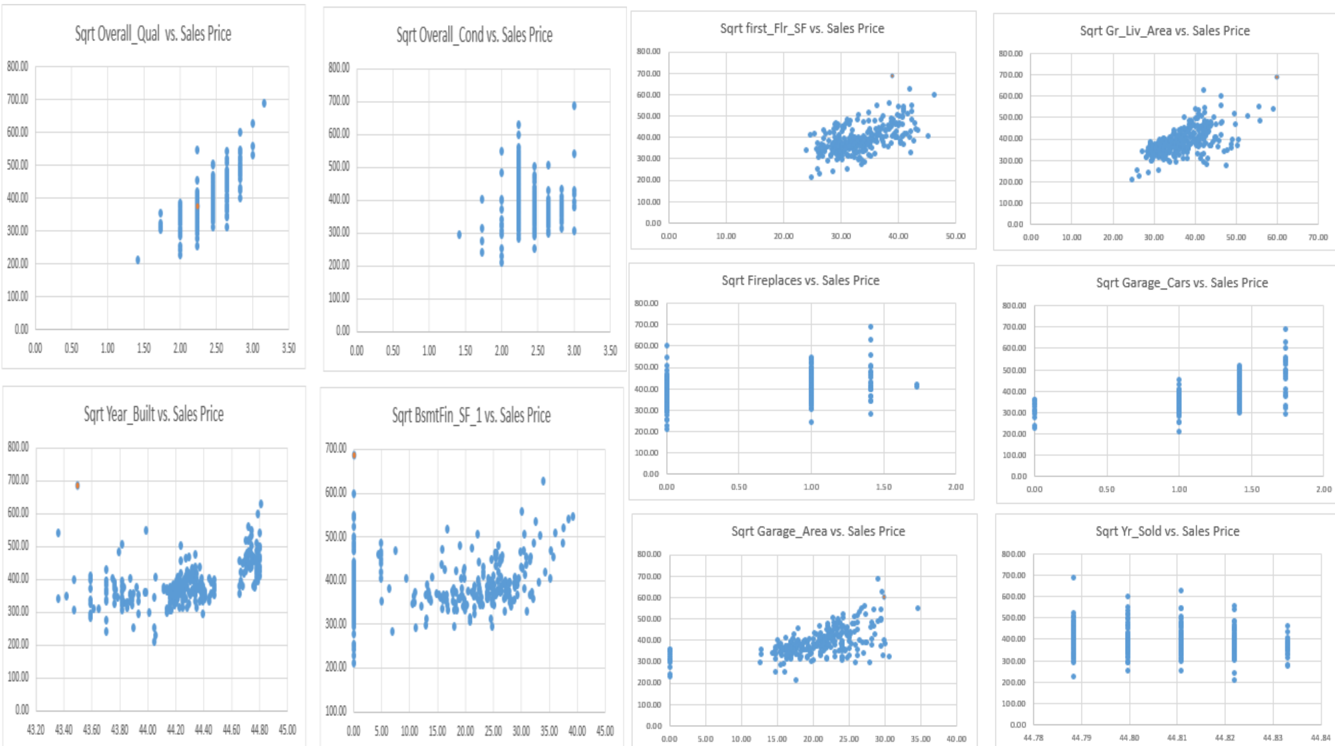


Figure 4 (outliers of Y, yellow highlighted student_sqrt_SalePrice column)

sqrt_Bsmt	sqrt_first_	sqrt_Gr_Li	sqrt_Firep	sqrt_Gara	sqrt_Gara	sqrt_Yr_S	sqrt_SalePri	student_sqrt_SalePric	cookd_sq	h_sqrt_SalePrice		
0	38.96152	60.06663	1.414214	1.732051	28.98275	44.78839	689.202438	5.7744345	0.355928	0.105080012		
0	36.16628	55.55178	0	1.732051	34.64102	44.79955	549.181209	4.000663051	0.162031	0.10020099	n=300,k=10	
33.95585	42.0238	42.0238	1.414214	1.732051	29.56349	44.81071	628.52128	3.90058972	0.057853	0.040148183	t(0.05,300-10-1)	1.968206
0	46.26013	46.26013	0	1.732051	29.89983	44.79955	600	3.397488624	0.068089	0.060933028		
33.31666	37.2827	37.2827	1.414214	1.414214	23.45208	44.78839	501.996016	2.971691861	0.024134	0.029184969		
0	27.12932	38.10512	0	0	0	44.82187	336.154726	2.263473918	0.031153	0.062694192		
0	26.34388	35.97221	0	0	0	44.81071	325.96012	2.16545844	0.025126	0.05565924		

Figure 5 (outliers of X, yellow highlighted h_sqrt_SalePrice column)

sqrt_Bsmt	sqrt_first_	sqrt_Gr_Li	sqrt_Firep	sqrt_Gara	sqrt_Gara	sqrt_Yr_S	sqrt_SalePri	student_sqrt_SalePric	cookd_sqrt_SalePrice	h_sqrt_SalePrice
0	33.09078	42.11888	0	1.732051	27.91057	44.78839	294.957624	-0.67903364	0.006351874	0.131593658
0	33.58571	47.49737	0	0	0	44.83302	276.586334	-1.781064981	0.037934902	0.116252031
27.34959	33.82307	33.82307	1	1	28.21347	44.78839	378.153408	-0.733172018	0.006269266	0.113704156
0	28.56571	28.56571	1	0	0	44.82187	241.867732	1.012974159	0.011001847	0.105497704
0	38.96152	60.06663	1.414214	1.732051	28.98275	44.78839	689.202438	5.7744345	0.355927987	0.105080012
32.18695	36.87818	55.7853	1	1.414214	28.19574	44.78839	484.767986	-0.581409735	0.003597598	0.104799905
0	36.16628	55.55178	0	1.732051	34.64102	44.79955	549.181209	4.000663051	0.162030871	0.10020099
29.32576	36.13862	36.13862	0	1	27.49545	44.82187	391.152144	-0.639146161	0.00401366	0.097535571
13	31.27299	43.17407	0	0	0	44.82187	312.2499	-1.55329293	0.023404002	0.096415097
0	32.04684	32.04684	1	0	0	44.81071	346.410162	1.072442978	0.010940142	0.09472167
0	40.92676	59.10161	1	1.732051	29.49576	44.81071	543.139025	1.934448724	0.033767332	0.090297245
0	26.34388	34.8425	0	1.732051	30.59412	44.78839	324.037035	1.474811374	0.018885976	0.087185028
16	34.10279	34.10279	0	1	14.8324	44.78839	355.211205	0.800542906	0.005006753	0.079136164
23.85372	35	35	0	0	0	44.83302	359.096087	1.730530977	0.022804536	0.077289523
0	24.73863	24.73863	0	1	17.54993	44.82187	212.132034	0.182924963	0.000253944	0.077048258
0	26.32489	26.32489	0	0	0	44.78839	229.128785	-0.238091407	0.000427848	0.076657973
17.94436	29.29164	29.29164	0	1	26.15339	44.79955	331.662479	-0.543149685	0.002170401	0.074868176
0	30.14963	42.09513	0	1.414214	28.56571	44.79955	430	-0.223987602	0.000364356	0.073976342
0	33.18132	41.24318	0	0	0	44.82187	307.489837	-0.061392602	2.72739E-05	0.073730195
24.77902	29.29164	29.29164	0	0	0	44.78839	296.647939	0.33671956	0.000807614	0.072660487
0	34.322	47.01064	0	1	14.31782	44.79955	342.78273	-0.34749362	0.000845757	0.071533745
24.97999	28.84441	38.49675	0	1.414214	18	44.82187	378.153408	-1.03536279	0.007154484	0.068394028
25.7682	42.0238	49.13247	1	1.414214	22.80351	44.78839	400	-0.217987598	0.000315867	0.068137385
0	45.15529	45.15529	1.732051	1.732051	28.12472	44.78839	408.656335	-1.862718889	0.022628573	0.066936992
0	27.64055	38.26225	0	0	0	44.82187	359.1657	1.834302552	0.021868878	0.066724755
0	26.15339	32.86335	0	0	0	44.81071	308.2207	1.09474231	0.007766122	0.066538001
24.81935	32.24903	32.24903	0	0	0	44.82187	352.136337	0.786207216	0.003991504	0.066321249
0	39.69887	50.53712	0	1.414214	26.533	44.79955	399.374511	-1.940644152	0.023189842	0.063436046
0	34.78505	34.78505	0	1	24.81935	44.82187	320.780299	-1.476265786	0.013357908	0.06316345
0	24.65766	37.36308	1	1.414214	20.97618	44.82187	412.310563	0.299153488	0.000546511	0.062945961
0	27.12932	38.10512	0	0	0	44.82187	336.154726	2.263473918	0.031153318	0.062694192
0	35.29873	47.85394	0	1.414214	23.66432	44.83302	350	-0.223998859	0.000304226	0.062525386
0	32.24903	32.24903	0	1.414214	20	44.82187	304.795013	0.795334379	0.003755659	0.061306046
0	46.26013	46.26013	0	1.732051	29.89983	44.79955	600	3.397488624	0.068089404	0.060933028
19.74842	29.66479	42.04759	1.414214	1.414214	17.88854	44.78839	403.732585	-0.268825045	0.000424262	0.060660951

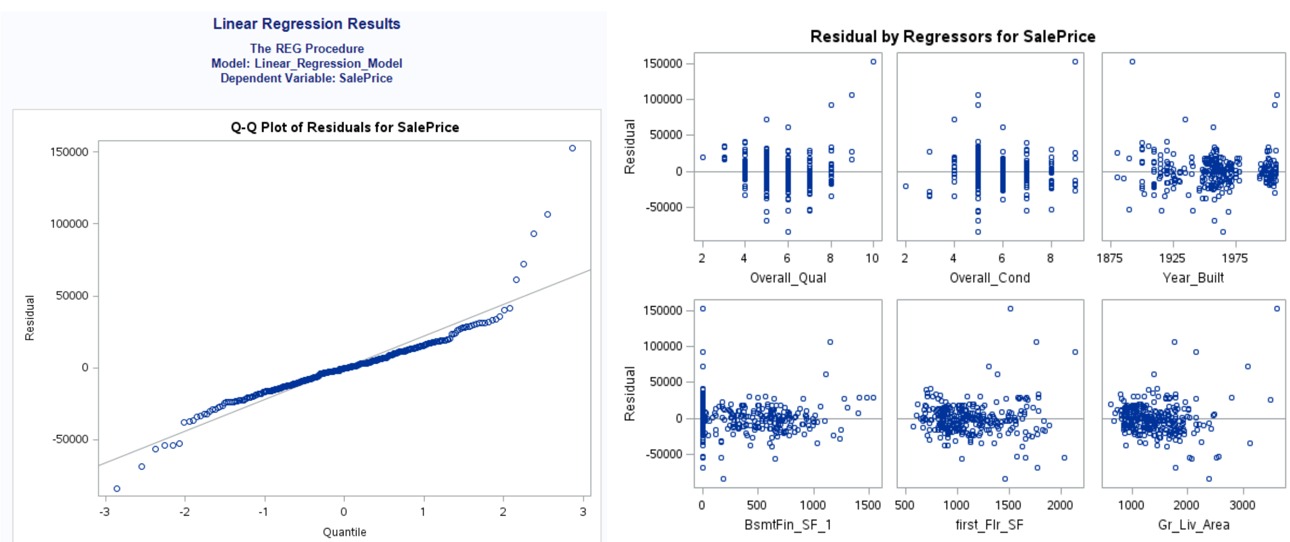
n=300,k=10	
2*(10-1)/300	0.06

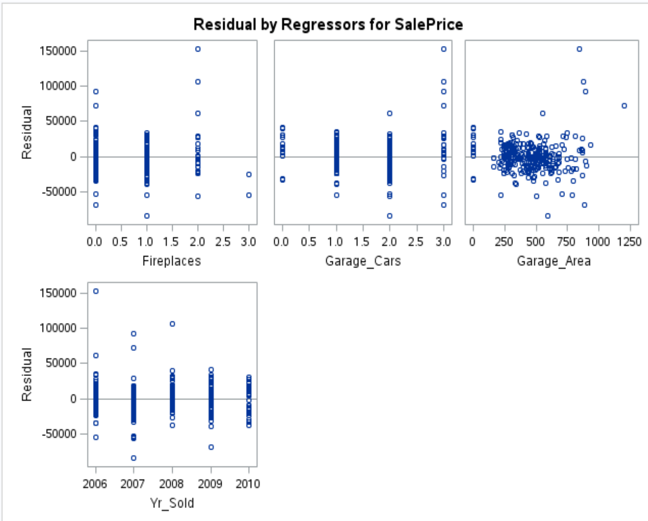
Figure 6 (influential point, yellow highlighted cookd_sqrt_SalePrice column)

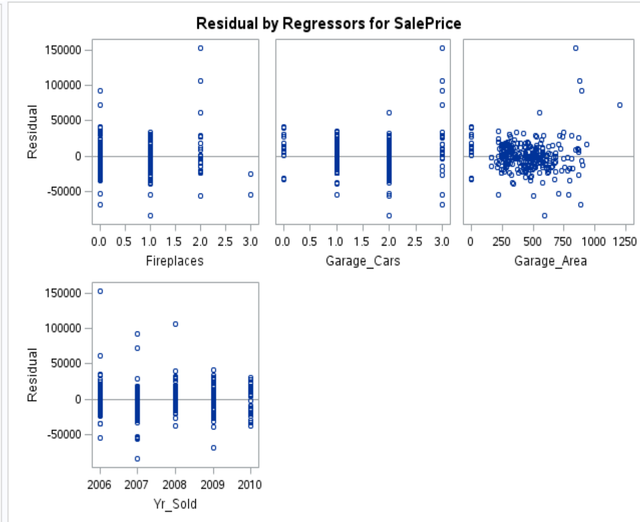
Overall_Q	sqrt_Over	sqrt_Year_	sqrt_Bsmt	sqrt_first_	sqrt_Gr_Li	sqrt_Firep	sqrt_Gara	sqrt_Gara	sqrt_Yr_S	sqrt_SalePri	student_sqrt_SalePric	cookd_sqrt_SalePrice	h_sqrt_SalePrice
3.162278	2.44949	44.23799	0	38.96152	60.06663	1.414214	1.732051	28.98275	44.78839	689.202438	5.7744345	0.355927987	0.105080012

n=300,k=10	
F(0.05,10+1,300-10-1)	1.821867

Figure 7(residual plots before transformation of variables)







Linear Regression Results

The REG Procedure
Model: Linear_Regression_Model
Dependent Variable: SalePrice

Number of Observations Read	300
Number of Observations Used	299
Number of Observations with Missing Values	1

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	10	7.003954E11	70039541407	139.47	<.0001
Error	288	1.44631E11	502191091		
Corrected Total	298	8.450264E11			

Root MSE	22410	R-Square	0.8288
Dependent Mean	155339	Adj R-Sq	0.8229
Coeff Var	14.42627		

Parameter Estimates						
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	1	898416	2033846	0.44	0.6590	0
Overall_Qual	1	13378	1552.37410	8.62	<.0001	2.08508
Overall_Cond	1	5800.40303	1231.27256	4.71	<.0001	1.32746
Year_Built	1	511.68947	60.69867	8.43	<.0001	2.18829
BsmtFin_SF_1	1	20.96230	3.91975	5.35	<.0001	1.25025
first_Flr_SF	1	13.47759	5.45752	2.47	0.0141	1.66992
Gr_Liv_Area	1	45.40134	4.20077	10.81	<.0001	2.30856
Fireplaces	1	3458.37940	2249.42476	1.54	0.1253	1.33653
Garage_Cars	1	2453.28600	4036.45972	0.61	0.5438	4.71835
Garage_Area	1	42.90510	13.75576	3.12	0.0020	4.49853
Yr_Sold	1	-979.45341	1008.98625	-0.97	0.3325	1.03364

Figure 8 :

Linear Regression Results

The REG Procedure
Model: Linear_Regression_Model
Dependent Variable: sqrt SalePrice

Number of Observations Read	300
Number of Observations Used	299
Number of Observations with Missing Values	1

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	10	990736	99074	146.36	<.0001
Error	288	194950	676.91088		
Corrected Total	298	1185686			

Root MSE	26.01751	R-Square	0.8356
Dependent Mean	389.06741	Adj R-Sq	0.8299
Coeff Var	6.68715		

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Error	288	194950	676.91088		
Corrected Total	298	1185686			

Root MSE	26.01751	R-Square	0.8356
Dependent Mean	389.06741	Adj R-Sq	0.8299
Coeff Var	6.68715		

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	1	2565.98903	4721.12019	0.54	0.5872	0
sqrt Overall_Qual	1	74.10503	8.52516	8.69	<.0001	2.04899
sqrt Overall_Cond	1	37.13510	6.94167	5.35	<.0001	1.32621
sqrt Year_Built	1	60.75603	6.18811	9.82	<.0001	2.15786
sqrt BsmtFin_SF_1	1	0.61980	0.13749	4.51	<.0001	1.23553
sqrt first_Flr_SF	1	1.30531	0.43202	3.02	0.0027	1.64278
sqrt Gr_Liv_Area	1	4.48087	0.37956	11.81	<.0001	2.25783
sqrt Fireplaces	1	4.32935	3.15717	1.37	0.1714	1.31367
sqrt Garage_Cars	1	-3.16244	11.85668	-0.27	0.7899	7.63463
sqrt Garage_Area	1	1.87705	0.69820	2.69	0.0076	7.46700
sqrt Yr_Sold	1	-120.17452	105.04847	-1.14	0.2536	1.03500

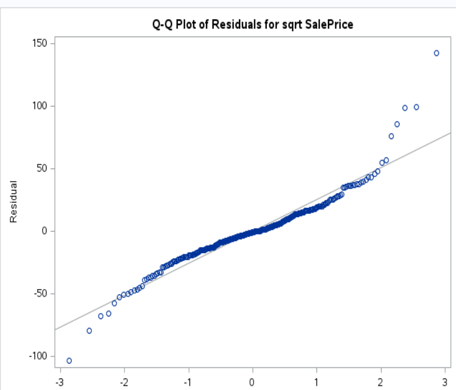
Generated by the SAS System ('SASApp', Linux) on April 23, 2018 at 5:07:04 PM

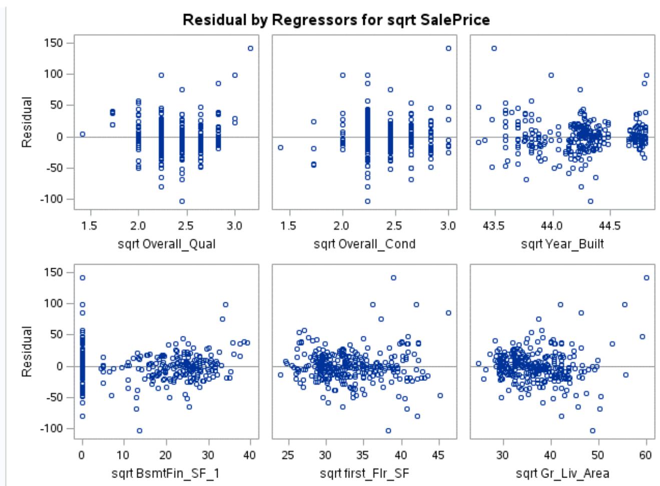
Page Break

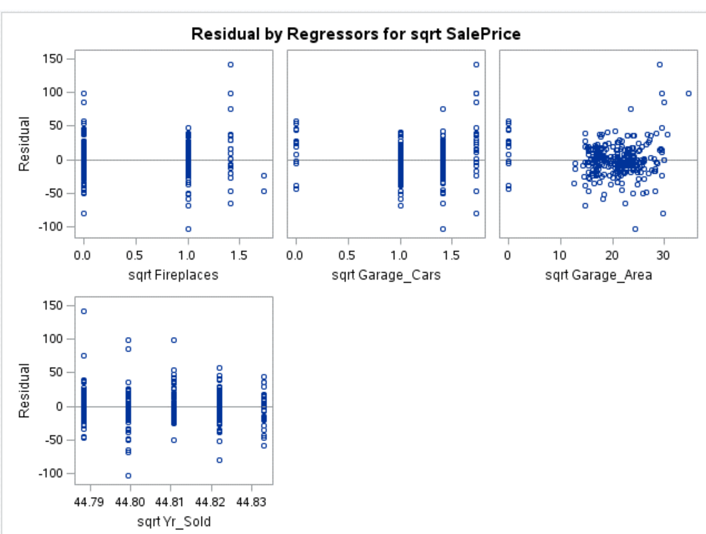
Linear Regression Results

Linear Regression Results

The REG Procedure
Model: Linear_Regression_Model
Dependent Variable: sqrt SalePrice







Generated by the SAS System ('SASApp', Linux) on April 23, 2018 at 5:07:04 PM

Page Break

Figure 9:

Chosen:

Model	Model Index	# in Model	R-Square	Variables in Model
1	176	4	0.796	sqrt Overall_Qual sqrt Year_Built sqrt Gr_Liv_Area sqrt Garage_Area
2	56	3	0.7692	sqrt Overall_Qual sqrt Year_Built sqrt Gr_Liv_Area
3	1	1	0.5738	sqrt Overall_Qual

From List:

Linear Regression Results

The REG Procedure
Model: Linear_Regression_Model
Dependent Variable: sqrt SalePrice

R-Square Selection Method

Number of Observations Read	300
Number of Observations Used	299
Number of Observations with Missing Values	1

Max R-Square
0.8356

Min R-Square
0.0044

Model Index	Number in Model	R-Square	Variables in Model
1	1	0.5738	sqrt Overall_Qual
2	1	0.3775	sqrt Gr_Liv_Area
3	1	0.3263	sqrt Garage_Cars
4	1	0.3213	sqrt first_Flr_SF
5	1	0.3139	sqrt Garage_Area
6	1	0.251	sqrt Year_Built
7	1	0.1592	sqrt Fireplaces
8	1	0.0505	sqrt BsmtFin_SF_1
9	1	0.0258	sqrt Yr_Sold
10	1	0.0044	sqrt Overall_Cond
11	2	0.6922	sqrt Year_Built sqrt Gr_Liv_Area
12	2	0.6667	sqrt Overall_Qual sqrt Garage_Area
13	2	0.6661	sqrt Overall_Qual sqrt first_Flr_SF
14	2	0.6543	sqrt Overall_Qual sqrt Garage_Cars
15	2	0.6525	sqrt Overall_Qual sqrt Gr_Liv_Area
16	2	0.6216	sqrt Overall_Qual sqrt Year_Built
17	2	0.6027	sqrt Overall_Qual sqrt BsmtFin_SF_1
18	2	0.5934	sqrt Overall_Qual sqrt Fireplaces
19	2	0.5822	sqrt Overall_Qual sqrt Yr_Sold
20	2	0.5786	sqrt Overall_Qual sqrt Overall_Cond
21	2	0.5343	sqrt Gr_Liv_Area sqrt Garage_Cars

Figure 10(Validation):

2	Garage_Cars	Garage_Area	Yr_Sold	SalePrice	id	sqrt Overall_Qual	sqrt Overall_Cond	sqrt Year_Built	sqrt BsmtFin_SF_1	sqrt first_Flr_SF	sqrt Gr_Liv_Area	sqrt Fireplaces	sqrt Garage_Cars	sqrt Garage_Area	sqrt Yr_Sold	sqrt SalePrice	Pred Sqrt (Sales Price)	Pred (Sales Price)	Abs Diff
3	2	565	2006	196500	538	2.65	2.24	44.77	-	29.33	41.47	-	1.41	23.77	44.79	443.28	$ 468.97	$ 219,936	$ 23,436
4	1	210	2006	107000	866	2.00	2.00	43.87	-	28.46	36.44	-	1.00	14.49	44.79	327.11	$ 317.75	$ 100,968	$ 6,032
5	2	523	2010	200500	544	2.45	2.24	44.70	22.47	28.35	41.01	-	1.41	22.87	44.83	447.77	$ 441.15	$ 194,617	$ 5,883
6	2	396	2008	129500	907	2.45	2.24	43.78	17.86	33.67	45.37	1.00	1.41	19.90	44.81	359.86	$ 402.87	$ 162,308	$ 32,808
7	1	312	2008	145500	742	2.24	2.65	44.26	28.39	32.45	32.45	-	1.00	17.66	44.81	381.44	$ 354.34	$ 125,557	$ 19,943
8	2	639	2007	213500	560	2.65	2.24	44.79	-	30.64	42.39	-	1.41	25.28	44.80	462.06	$ 476.07	$ 226,643	$ 13,143
9	1	392	2010	107500	624	2.24	2.65	44.22	-	28.76	28.76	1.00	1.00	19.80	44.83	327.87	$ 341.77	$ 116,804	$ 9,304
10	2	506	2006	119000	644	2.24	2.24	44.35	-	39.80	39.80	-	1.41	22.49	44.79	344.96	$ 394.40	$ 155,548	$ 36,548
11	2	400	2006	93500	613	1.73	2.00	44.16	-	32.25	32.25	-	1.41	20.00	44.79	305.78	$ 298.02	$ 88,818	$ 4,682
12	2	495	2008	186000	814	2.65	2.65	44.25	25.24	36.22	36.22	1.00	1.41	22.25	44.81	431.28	$ 419.10	$ 175,646	$ 10,354
13	1	288	2008	114000	880	2.24	2.83	43.87	-	27.33	34.04	-	1.00	16.97	44.81	337.64	$ 338.24	$ 114,409	$ 409
14	1	184	2007	89500	848	2.24	2.24	43.84	27.96	31.97	31.97	1.00	1.00	13.56	44.80	299.17	$ 323.09	$ 104,385	$ 14,885
15	1	216	2009	179900	943	2.65	2.83	43.76	-	35.67	44.09	-	1.00	14.70	44.82	424.15	$ 409.33	$ 167,549	$ 12,351
16	2	512	2007	138500	708	2.24	2.65	44.05	11.40	30.13	30.13	-	1.41	22.63	44.80	372.16	$ 342.35	$ 117,200	$ 21,300
17	1	200	2008	78000	839	2.24	2.45	43.70	-	29.98	29.98	-	1.00	14.14	44.81	279.28	$ 309.29	$ 95,659	$ 17,659
18	2	588	2008	178400	807	2.65	3.00	44.28	29.44	31.61	31.61	-	1.41	24.25	44.81	422.37	$ 407.17	$ 165,785	$ 12,615
19	1	312	2010	108538	625	2.24	2.24	44.26	30.59	34.73	34.73	-	1.00	17.66	44.83	329.45	$ 362.77	$ 131,600	$ 23,062
20	3	754	2006	294900	602	2.83	2.24	44.78	33.50	39.87	39.87	-	1.73	27.46	44.79	543.05	$ 489.31	$ 239,425	$ 55,475
21	2	544	2006	200000	543	2.45	2.24	44.75	18.44	41.23	41.23	-	1.41	23.32	44.79	447.21	$ 445.74	$ 198,680	$ 1,320
22	2	342	2007	137000	915	2.00	2.83	43.59	-	34.00	42.40	-	1.41	18.49	44.80	370.14	$ 330.97	$ 109,544	$ 27,456
23	2	440	2006	116000	884	2.45	2.83	43.82	-	27.09	33.44	-	1.41	20.98	44.79	340.59	$ 362.37	$ 131,314	$ 15,314
24	3	725	2008	274000	597	3.00	2.24	44.73	34.55	36.96	44.12	1.00	1.73	26.93	44.81	523.45	$ 520.10	$ 270,501	$ 3,499
25	2	484	2008	210000	554	2.65	2.24	44.77	30.76	37.89	37.89	-	1.41	22.00	44.81	458.26	$ 452.81	$ 205,037	$ 4,963
26	2	576	2006	217000	564	2.45	2.24	44.69	34.64	38.65	38.65	-	1.41	24.00	44.79	465.83	$ 433.69	$ 188,085	$ 28,915
27	2	430	2006	167000	794	2.45	2.45	44.32	16.12	38.57	38.57	1.00	1.41	20.74	44.79	408.66	$ 407.97	$ 166,437	$ 563
28	1	384	2008	140000	716	2.45	2.45	44.33	29.87	33.70	33.70	1.00	1.00	19.60	44.81	374.17	$ 388.67	$ 151,068	$ 11,068
29	1	286	2009	150000	753	2.24	2.65	44.23	29.26	33.82	33.82	-	1.00	16.91	44.82	387.30	$ 356.35	$ 126,982	$ 23,018
30	2	558	2007	226000	573	2.65	2.24	44.75	31.83	38.46	38.46	-	1.41	23.62	44.80	475.39	$ 456.97	$ 208,822	$ 17,178
31	2	434	2006	195000	536	2.65	2.24	44.77	-	29.10	44.14	1.00	1.41	20.83	44.79	441.59	$ 473.97	$ 224,649	$ 29,649
32	4	1488	2008	139000	917	2.24	2.24	43.84	-	35.41	37.89	1.00	2.00	38.57	44.81	372.83	$ 386.35	$ 149,267	$ 10,267
33	1	730	2010	105000	621	2.24	2.45	44.28	21.63	29.93	29.93	-	1.00	27.02	44.83	324.04	$ 361.71	$ 130,837	$ 25,837
34	1	509	2009	193500	818	2.45	2.45	44.36	21.17	34.38	46.28	1.00	1.00	22.56	44.82	439.89	$ 441.93	$ 195,306	$ 1,806
35	1	308	2008	97000	854	2.24	2.24	44.23	-	29.70	29.70	-	1.00	17.55	44.81	311.45	$ 342.14	$ 117,061	$ 20,061
36	2	484	2010	130000	908	2.24	2.65	43.92	15.30	30.43	36.28	-	1.41	22.00	44.83	360.56	$ 357.28	$ 127,648	$ 2,352
37	1	336	2007	150000	754	2.45	2.45	44.24	27.87	35.27	35.27	1.41	1.00	18.33	44.80	387.30	$ 387.49	$ 150,151	$ 151
38	1	180	2009	135000	912	2.24	2.65	43.77	-	28.43	40.69	-	1.00	13.42	44.82	367.42	$ 351.42	$ 123,497	$ 11,503
39	2	420	2009	95000	852	1.73	2.24	43.59	20.98	31.84	41.22	-	1.41	20.49	44.82	308.22	$ 301.23	$ 90,739	$ 4,261
40	2	400	2006	155000	768	2.45	2.45	44.17	-	39.69	39.69	1.00	1.41	20.00	44.79	393.70	$ 402.92	$ 162,347	$ 7,347
41	1	308	2010	150000	503	2.24	2.83	44.46	14.07	32.03	32.03	-	1.00	17.55	44.83	387.30	$ 363.57	$ 132,182	$ 17,818
42	3	1052	2007	424870	608	2.83	2.24	44.79	33.91	38.29	53.18	1.00	1.73	32.43	44.80	651.82	$ 547.36	$ 299,601	$ 125,269
43	2	484	2007	168675	512	2.65	2.24	44.79	-	27.17	38.61	1.00	1.41	22.00	44.80	410.70	$ 456.68	$ 208,553	$ 39,878
44	3	904	2009	239900	587	2.83	2.24	44.81	-	38.31	38.31	1.00	1.73	30.07	44.82	489.80	$ 489.66	$ 239,765	$ 135
45	2	360	2010	115000	882	2.24	2.45	43.90	-	32.06	32.06	1.00	1.41	18.97	44.83	339.12	$ 335.45	$ 112,530	$ 2,470

Average	$ 16,220
Std Dev	$ 15,533