

Obs	address	info	z_address	neighborhood	usecode	bathrooms	bedrooms	finishedsqft	lastsolddate	lastsoldprice	latitude	longitude	totalrooms	yearbuilt	zestimate	zipcode
1	Address: 560 Missouri Street #B	San FranciscoSales price: 1495000Sales date: 02/17/2016Bedrooms: 3Square feet: 1425Lot size: 2496	560 Missouri St # B	Potero Hill	Condominium	4	3	1425	20501	1495000	37.7592	-122.397	6	2003	1708594	94107
2	Address: 350 Missouri Street	San FranciscoSales price: 2700000Sales date: 02/17/2016Bedrooms: 3Square feet: 1984Lot size: 2469	350 Missouri St	Potero Hill	SingleFamily	3	3	2231	20501	2700000	37.7619	-122.397	10	1927	2411236	94107
3	Address: 3658 Folsom Street	San FranciscoSales price: 1530000Sales date: 02/17/2016Bedrooms: 0Square feet: 924Lot size: 1750	3658 Folsom St	Bernal Heights	SingleFamily	3	3	1300	20501	1530000	37.7408	-122.413	4	1900	1918539	94110
4	Address: 156 Lowell Street	San FranciscoSales price: 4600000Sales date: 02/17/2016Bedrooms: 2Square feet: 1250Lot size: 3000	156 Lowell St	Crocker Amazon	SingleFamily	1	2	1250	20501	4600000	37.7102	-122.446	5	1924	703392	94112
5	Address: 735 Lakeview Avenue	San FranciscoSales price: 5320000Sales date: 02/17/2016Bedrooms: 3Square feet: 1032Lot size: 2495	735 Lakeview Ave	Oceanview	SingleFamily	1	3	1032	20501	532000	37.7180	-122.460	6	1939	741715	94112

#### Data after dropping variables using DATA step

Obs	usecode	bathrooms	bedrooms	finishedsqft	lastsolddate	lastsoldprice	totalrooms	yearbuilt	zestimate	zipcode
1	Condominium	4	3	1425	20501	1495000	6	2003	1708594	94107
2	SingleFamily	3	3	2231	20501	2700000	10	1927	2411236	94107
3	SingleFamily	3	3	1300	20501	1530000	4	1900	1918539	94110
4	SingleFamily	1	2	1250	20501	4600000	5	1924	703392	94112
5	SingleFamily	1	3	1032	20501	532000	6	1939	741715	94112

#### Data after dropping variables using PROC SQL

Obs	bathrooms	bedrooms	finishedsqft	lastsolddate	lastsoldprice	totalrooms	usecode	yearbuilt	zestimate	zipcode
1	4	3	1425	20501	1495000	6	Condominium	2003	1708594	94107
2	3	3	2231	20501	2700000	10	SingleFamily	1927	2411236	94107
3	3	3	1300	20501	1530000	4	SingleFamily	1900	1918539	94110
4	1	2	1250	20501	4600000	5	SingleFamily	1924	703392	94112
5	1	3	1032	20501	532000	6	SingleFamily	1939	741715	94112

#### Data after Adding a new variable using DATA step

Obs	usecode	bathrooms	bedrooms	finishedsqft	lastsolddate	lastsoldprice	totalrooms	yearbuilt	zestimate	zipcode	price_per_square_foot
1	Condominium	4	3	1425	20501	1495000	6	2003	1708594	94107	1049.12
2	SingleFamily	3	3	2231	20501	2700000	10	1927	2411236	94107	1210.22
3	SingleFamily	3	3	1300	20501	1530000	4	1900	1918539	94110	1176.92
4	SingleFamily	1	2	1250	20501	4600000	5	1924	703392	94112	368.00
5	SingleFamily	1	3	1032	20501	532000	6	1939	741715	94112	515.50

#### Data after Adding a new variable using PROC SQL

Obs	usecode	bathrooms	bedrooms	finishedsqft	lastsolddate	lastsoldprice	totalrooms	yearbuilt	zestimate	zipcode	price_per_square_foot
1	Condominium	4	3	1425	20501	1495000	6	2003	1708594	94107	1049.12
2	SingleFamily	3	3	2231	20501	2700000	10	1927	2411236	94107	1210.22
3	SingleFamily	3	3	1300	20501	1530000	4	1900	1918539	94110	1176.92
4	SingleFamily	1	2	1250	20501	4600000	5	1924	703392	94112	368.00
5	SingleFamily	1	3	1032	20501	532000	6	1939	741715	94112	515.50

#### Data after Adding a new variable using PROC SQL

The MEANS Procedure

Analysis Variable : lastsoldprice						
zipcode	N Obs	N	Mean	Std Dev	Minimum	Maximum
94102	318	318	835957.49	369222.84	12000.00	2599000.00
94103	359	359	1057131.44	743471.94	209888.00	7850000.00
94104	1	1	1300000.00	.	1300000.00	1300000.00
94105	340	340	1202028.24	879560.55	250000.00	8100000.00
94107	856	856	1115697.78	619521.75	40000.00	9500000.00
94108	114	114	1479360.53	1005902.05	600.0000000	4995000.00

Analysis Variable : lastsoldprice						
zipcode	N Obs	N	Mean	Std Dev	Minimum	Maximum
94109	653	653	1274315.54	1164511.97	74000.00	14000000.00
94110	935	935	1235720.02	632018.21	12920.00	8125000.00
94111	38	38	1078907.89	381671.68	535000.00	2100000.00
94112	877	877	763993.29	253747.09	535.0000000	1875000.00
94114	651	651	1685736.43	947387.93	15500.00	7000000.00
94115	534	534	1860896.42	2067537.08	35000.00	23889000.00
94116	655	655	1019365.35	393961.29	5700.00	3549000.00
94117	473	473	1454416.58	931391.69	18000.00	6500000.00
94118	466	466	1835655.89	1610192.04	70000.00	18000000.00
94121	527	527	1474686.53	1151000.33	299000.00	11000000.00
94122	601	601	1070692.68	405502.80	30500.00	3150000.00
94123	430	430	2449347.67	2002826.05	1000.00	16760000.00
94124	393	393	573805.64	189456.68	3000.00	1300000.00
94127	435	435	1365385.56	655196.27	5450.00	6285000.00
94131	687	687	1382432.57	758288.48	26000.00	5250000.00
94132	292	292	774159.25	320999.81	90000.00	1889000.00
94133	183	183	1718297.81	1350321.25	345000.00	9998000.00
94134	439	439	650733.71	177587.07	25500.00	1255000.00
94158	71	71	976035.21	312054.89	490000.00	2000000.00

#### Average Last Sold Price by Zipcode using DATA step

Obs	zipcode	_TYPE_	_FREQ_	avg_lastsoldprice
1	.	0	11328	1263970.37
2	94102	1	318	835957.49
3	94103	1	359	1057131.44
4	94104	1	1	1300000.00
5	94105	1	340	1202028.24
6	94107	1	856	1115697.78
7	94108	1	114	1479360.53
8	94109	1	653	1274315.54
9	94110	1	935	1235720.02
10	94111	1	38	1078907.89

#### Average Last Sold Price by Zipcode using PROC SQL

Obs	zipcode	avg_lastsoldprice
1	94102	835957.49
2	94103	1057131.44
3	94104	1300000.00
4	94105	1202028.24
5	94107	1115697.78
6	94108	1479360.53
7	94109	1274315.54
8	94110	1235720.02
9	94111	1078907.89
10	94112	763993.29

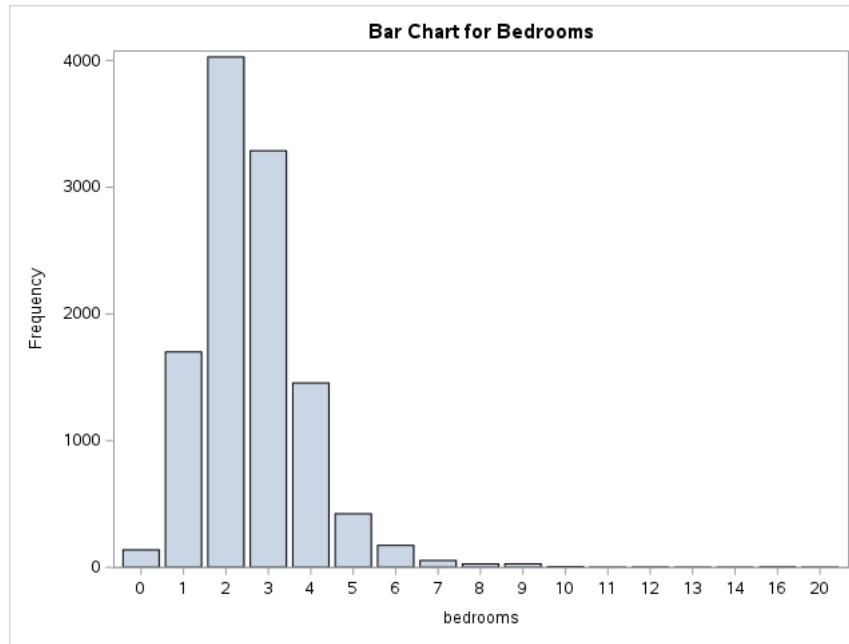
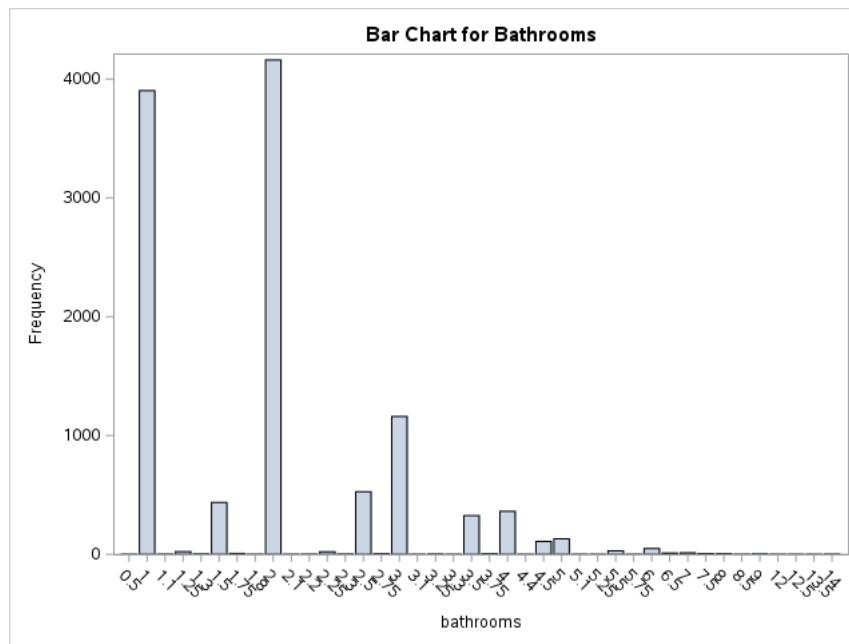
#### Average Last Sold Price by Use Code, Total Rooms, and Bedrooms using DATA step

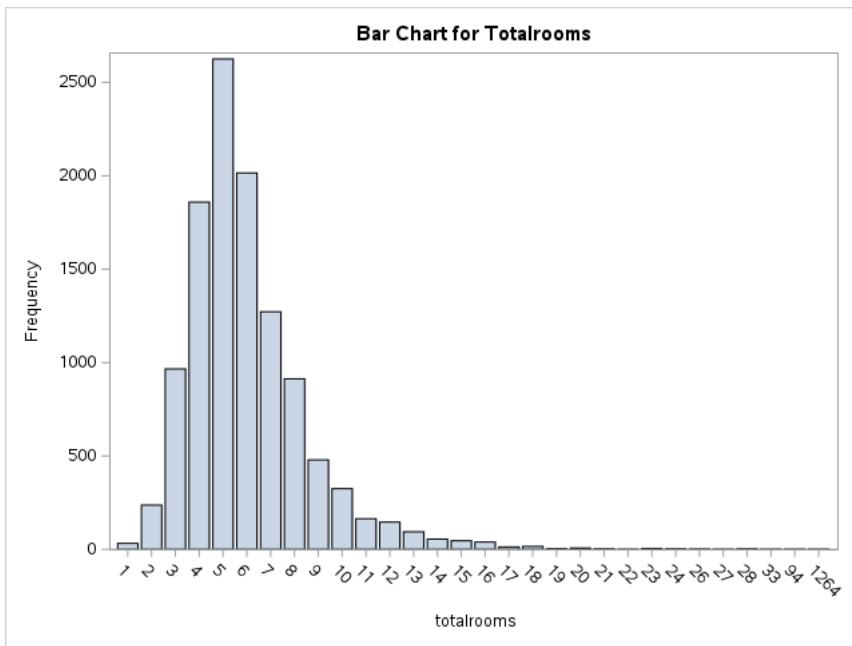
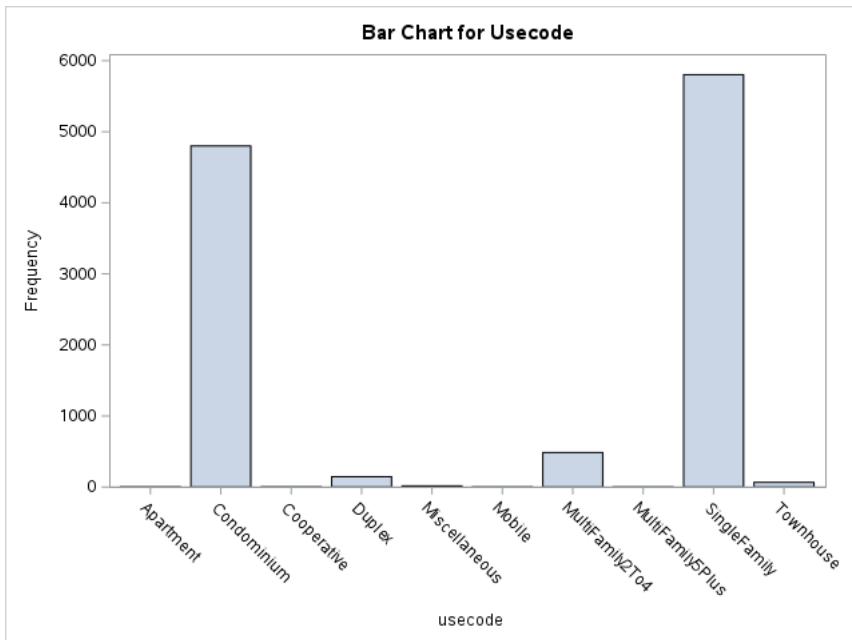
Obs	usecode	totalrooms	bedrooms	_TYPE_	_FREQ_	avg_lastsoldprice
1	.	.	.	0	11328	1263970.37
2	.	.	0	1	138	681605.07
3	.	.	1	1	1700	746408.89
4	.	.	2	1	4028	1050092.70
5	.	.	3	1	3288	1323012.77
6	.	.	4	1	1455	1780377.15
7	.	.	5	1	423	2318233.57
8	.	.	6	1	173	2796990.87
9	.	.	7	1	54	3336972.22
10	.	.	8	1	27	2202333.33

#### Average Last Sold Price by Use Code, Total Rooms, and Bedrooms using PROC SQL step

Obs	usecode	totalrooms	bedrooms	avg_lastsoldprice
1	Apartment	5	1	845000.00
2	Apartment	5	2	1850000.00
3	Apartment	7	2	1330000.00
4	Condominium	1	0	494366.67
5	Condominium	1	1	1141777.78

Obs	usecode	totalrooms	bedrooms	avg_lastsoldprice
6	Condominium	1	2	1007500.00
7	Condominium	2	0	474083.33
8	Condominium	2	1	660961.15
9	Condominium	2	2	1158444.44
10	Condominium	3	0	488590.91





**Histogram for Last Sold Price**

The UNIVARIATE Procedure  
Variable: bathrooms

Moments			
N	11328	Sum Weights	11328
Mean	1.98031427	Sum Observations	22433
Std Deviation	1.04740996	Variance	1.09706763
Skewness	1.97413883	Kurtosis	9.20546362
Uncorrected SS	56850.875	Corrected SS	12426.4851
Coeff Variation	52.8910983	Std Error Mean	0.00984102

Basic Statistical Measures			
Location		Variability	
Mean	1.980314	Std Deviation	1.04741
Median	2.000000	Variance	1.09707
Mode	2.000000	Range	13.50000
		Interquartile Range	1.00000

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	201.2306	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001

Tests for Location: Mu0=0			
Test	Statistic	p Value	
Signed Rank	S	32083728	Pr >=  S  <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	14.0
99%	5.5
95%	4.0
90%	3.0
75% Q3	2.0
50% Median	2.0
25% Q1	1.0
10%	1.0
5%	1.0
1%	1.0
0% Min	0.5

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0.5	4792	12.0	5504
1.0	11327	12.5	7817
1.0	11326	13.5	10848
1.0	11323	14.0	345
1.0	11319	14.0	4210

### Histogram for Last Sold Price

The UNIVARIATE Procedure  
Variable: bedrooms

Moments			
N	11328	Sum Weights	11328
Mean	2.61467161	Sum Observations	29619
Std Deviation	1.29947058	Variance	1.68862378
Skewness	1.66247966	Kurtosis	9.42389926
Uncorrected SS	96571	Corrected SS	19127.0416
Coeff Variation	49.6991887	Std Error Mean	0.01220927

Basic Statistical Measures			
Location		Variability	
Mean	2.614672	Std Deviation	1.29947
Median	2.000000	Variance	1.68862
Mode	2.000000	Range	20.00000
		Interquartile Range	1.00000

Tests for Location: Mu0=0			
Test	Statistic	p Value	
Student's t	t	214.1546	Pr >  t  <.0001
Sign	M	5595	Pr >=  M  <.0001
Signed Rank	S	31306823	Pr >=  S  <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	20
99%	7
95%	5
90%	4
75% Q3	3
50% Median	2
25% Q1	2
10%	1
5%	1
1%	0
0% Min	0

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0	11268	14	10286
0	11244	16	5688
0	11222	16	9200
0	11062	16	10726

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0	11053	20	6742

### Histogram for Last Sold Price

The UNIVARIATE Procedure  
Variable: finishedsqrt

Moments			
N	11328	Sum Weights	11328
Mean	1585.52904	Sum Observations	17960873
Std Deviation	922.023251	Variance	850126.875
Skewness	4.18246964	Kurtosis	62.838115
Uncorrected SS	3.81069E10	Corrected SS	9629387118
Coeff Variation	58.1524038	Std Error Mean	8.66293893

Basic Statistical Measures			
Location		Variability	
Mean	1585.529	Std Deviation	922.02325
Median	1362.000	Variance	850127
Mode	1200.000	Range	27274
		Interquartile Range	857.50000

Tests for Location: Mu0=0			
Test	Statistic		p Value
Student's t	t	183.0244	Pr >  t  <.0001
Sign	M	5664	Pr >=  M  <.0001
Signed Rank	S	32083728	Pr >=  S  <.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	27275.0
99%	4992.0
95%	3215.0
90%	2600.0
75% Q3	1876.5
50% Median	1362.0
25% Q1	1019.0
10%	780.0
5%	672.0
1%	493.0
0% Min	1.0

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
1	11135	9540	5703
1	10966	9699	9649
1	10651	10000	8086
1	7382	10236	2749
1	5102	27275	4996

### Histogram for Last Sold Price

The UNIVARIATE Procedure  
Variable: lastsolddate

Moments			
N	11328	Sum Weights	11328
Mean	19851.3646	Sum Observations	224876258
Std Deviation	527.668851	Variance	278434.416
Skewness	-8.5693936	Kurtosis	194.597322
Uncorrected SS	4.46725E12	Corrected SS	3153826634
Coeff Variation	2.65809864	Std Error Mean	4.95775245

Basic Statistical Measures			
Location		Variability	
Mean	19851.36	Std Deviation	527.66885
Median	19877.00	Variance	278434
Mode	19600.00	Range	16614
		Interquartile Range	627.00000

Note: The mode displayed is the smallest of 2 modes with a count of 46.

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	4004.106	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001
Signed Rank	S	32083728	Pr >=  S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	20541
99%	20501
95%	20440
90%	20384
75% Q3	20206
50% Median	19877
25% Q1	19579
10%	19318
5%	19250
1%	18914
0% Min	3927

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
3927	9926	20529	10771
4688	4736	20531	8130
7063	6907	20531	8929
10379	8434	20535	9141
11774	3974	20541	8676

### Histogram for Last Sold Price

The UNIVARIATE Procedure  
Variable: lastsoldprice

Moments			
N	11328	Sum Weights	11328
Mean	1263970.37	Sum Observations	1.43183E10
Std Deviation	1042159.8	Variance	1.0861E12
Skewness	5.30103726	Kurtosis	55.2651108
Uncorrected SS	3.04001E16	Corrected SS	1.23022E16
Coeff Variation	82.4512841	Std Error Mean	9791.69094

Basic Statistical Measures			
Location		Variability	
Mean	1263970	Std Deviation	1042160
Median	990000	Variance	1.0861E12
Mode	1200000	Range	23888465
		Interquartile Range	721000

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	129.086	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001
Signed Rank	S	32083728	Pr >=  S	<.0001

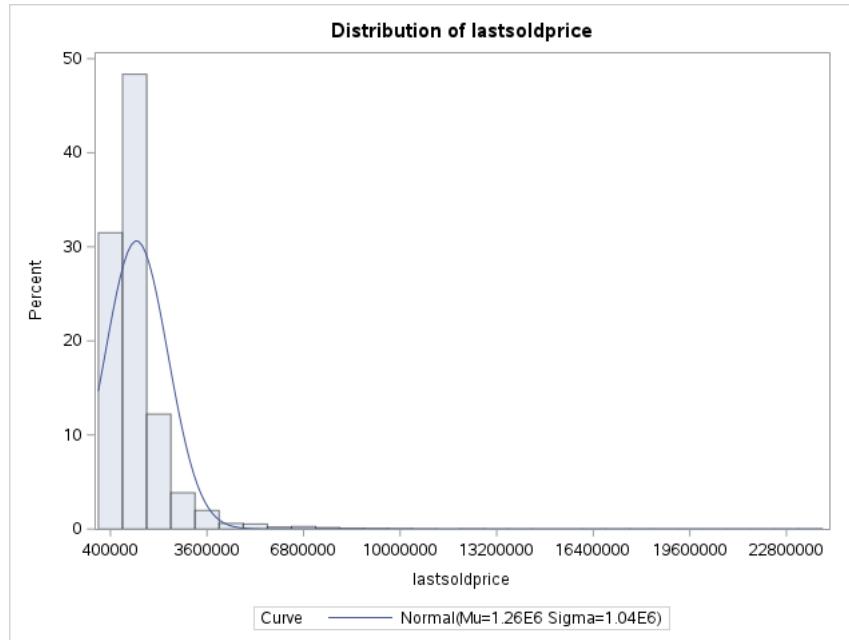
Quantiles (Definition 5)	
Level	Quantile
100% Max	23889000
99%	5500000
95%	2885000
90%	2150000
75% Q3	1450000
50% Median	990000
25% Q1	729000
10%	555000
5%	450000
1%	240000
0% Min	535

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
535	5520	14000000	3353
600	4541	16760000	1884
1000	6840	17100000	522
3000	1247	18000000	8230

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
5450	5064	23889000	2749

### Histogram for Last Sold Price

The UNIVARIATE Procedure



### Histogram for Last Sold Price

The UNIVARIATE Procedure  
Fitted Normal Distribution for lastsoldprice

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	1263970
Std Dev	Sigma	1042160

Goodness-of-Fit Tests for Normal Distribution			
Test	Statistic	p Value	
Kolmogorov-Smirnov	D	0.184425	Pr > D <0.010
Cramer-von Mises	W-Sq	167.309936	Pr > W-Sq <0.005
Anderson-Darling	A-Sq	922.016487	Pr > A-Sq <0.005

Quantiles for Normal Distribution		
Percent	Quantile	
	Observed	Estimated
1.0	240000	-1160455.9
5.0	450000	-450230.0
10.0	555000	-71611.2
25.0	729000	561044.3
50.0	990000	1263970.4
75.0	1450000	1966896.5
90.0	2150000	2599551.9
95.0	2885000	2978170.7
99.0	5500000	3688396.6

### Histogram for Last Sold Price

The UNIVARIATE Procedure  
Variable: totalrooms

Moments			
N	11328	Sum Weights	11328
Mean	6.11202331	Sum Observations	69237
Std Deviation	12.1268378	Variance	147.060196
Skewness	98.5980606	Kurtosis	10222.9567
Uncorrected SS	2088929	Corrected SS	1665750.84
Coeff Variation	198.409549	Std Error Mean	0.11393862

Basic Statistical Measures				
Location		Variability		
Mean	6.112023	Std Deviation		12.12684
Median	5.000000	Variance		147.06020
Mode	5.000000	Range		1263
		Interquartile Range		3.00000

Tests for Location: Mu0=0				
Test	Statistic	p Value		
Student's t	t	53.64312	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001
Signed Rank	S	32083728	Pr >=  S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	1264
99%	15
95%	11
90%	9
75% Q3	7
50% Median	5
25% Q1	4
10%	3
5%	3
1%	2
0% Min	1

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
1	11222	28	2749
1	10987	28	10726
1	10594	33	11174
1	10432	94	8299
1	9858	1264	8966

### Histogram for Last Sold Price

The UNIVARIATE Procedure  
Variable: yearbuilt

Moments			
N	11328	Sum Weights	11328
Mean	1948.48808	Sum Observations	22072473
Std Deviation	37.9069693	Variance	1436.93832
Skewness	0.34358094	Kurtosis	-1.2386919
Uncorrected SS	4.30242E10	Corrected SS	16276200.4
Coeff Variation	1.94545554	Std Error Mean	0.35615779

Basic Statistical Measures			
Location		Variability	
Mean	1948.488	Std Deviation	37.90697
Median	1940.000	Variance	1437
Mode	1900.000	Range	156.00000
		Interquartile Range	70.50000

Tests for Location: Mu0=0				
Test	Statistic	p Value		
Student's t	t	5470.856	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001
Signed Rank	S	32083728	Pr >=  S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	2016.0
99%	2014.0
95%	2009.0
90%	2006.0
75% Q3	1986.0
50% Median	1940.0
25% Q1	1915.5
10%	1904.0
5%	1900.0
1%	1890.0
0% Min	1860.0

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
1860	6635	2015	3402
1863	9649	2015	4210
1866	7735	2015	6878
1866	3733	2015	10416
1870	5267	2016	8513

### Histogram for Last Sold Price

The UNIVARIATE Procedure  
Variable: zestimate

Moments			
N	11328	Sum Weights	11328
Mean	1565795.91	Sum Observations	1.77373E10
Std Deviation	1229500.14	Variance	1.51167E12
Skewness	4.13692999	Kurtosis	25.6680924
Uncorrected SS	4.48957E16	Corrected SS	1.71227E16
Coeff Variation	78.5223752	Std Error Mean	11551.8612

Basic Statistical Measures			
Location		Variability	
Mean	1565796	Std Deviation	1229500
Median	1231054	Variance	1.51167E12
Mode	1101226	Range	15100868
		Interquartile Range	826533

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	135.5449	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001
Signed Rank	S	32083728	Pr >=  S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	15533253
99%	7045051
95%	3699838
90%	2639675
75% Q3	1731778
50% Median	1231054
25% Q1	905245
10%	737918
5%	641875
1%	536665
0% Min	432385

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
432385	1733	14272621	2749
433999	2498	14658165	8230
458909	1218	15048610	426
464318	2708	15463813	8086
466307	1112	15533253	522

### Histogram for Last Sold Price

The UNIVARIATE Procedure  
Variable: zipcode

Moments			
N	11328	Sum Weights	11328
Mean	94116.9141	Sum Observations	1066156403
Std Deviation	9.40033672	Variance	88.3663305
Skewness	0.77787596	Kurtosis	0.97142959
Uncorrected SS	1.00343E14	Corrected SS	1000925.43
Coeff Variation	0.00998794	Std Error Mean	0.08832157

Basic Statistical Measures			
Location		Variability	
Mean	94116.91	Std Deviation	9.40034
Median	94115.00	Variance	88.36633
Mode	94110.00	Range	56.00000

Basic Statistical Measures				
Location		Variability		
		Interquartile Range		
		13.00000		

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	1065616	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001
Signed Rank	S	32083728	Pr >=  S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	94158
99%	94134
95%	94133
90%	94131
75% Q3	94123
50% Median	94115
25% Q1	94110
10%	94107
5%	94103
1%	94102
0% Min	94102

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
94102	11295	94158	10367
94102	11268	94158	10686
94102	11245	94158	10973
94102	11238	94158	10979
94102	11231	94158	11263

### Histogram for Last Sold Price

The UNIVARIATE Procedure  
Variable: price\_per\_square\_foot

Moments			
N	11328	Sum Weights	11328
Mean	1760.90617	Sum Observations	19947545.1
Std Deviation	33864.2358	Variance	1146786467
Skewness	40.4361909	Kurtosis	1791.28167
Uncorrected SS	1.30248E13	Corrected SS	1.29897E13
Coeff Variation	1923.11415	Std Error Mean	318.173979

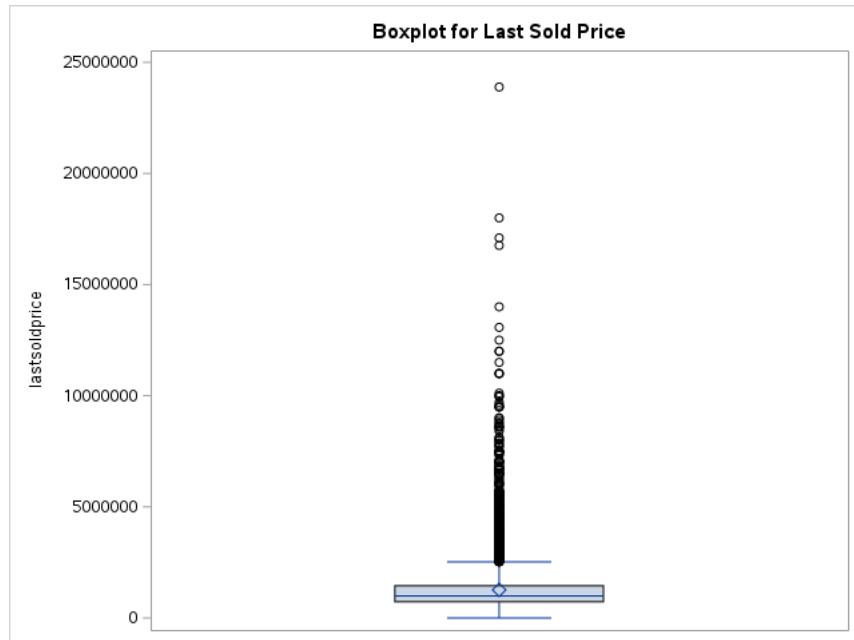
Basic Statistical Measures				
Location		Variability		
Mean	1760.906	Std Deviation	33864	
Median	781.866	Variance	1146786467	
Mode	1000.000	Range	1930000	
		Interquartile Range	417.41596	

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	5.534413	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001
Signed Rank	S	32083728	Pr >=  S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	1.93000E+06
99%	2.02203E+03
95%	1.38053E+03
90%	1.21385E+03
75% Q3	9.99683E+02
50% Median	7.81866E+02
25% Q1	5.82267E+02
10%	4.29630E+02
5%	3.37621E+02
1%	1.86296E+02
0% Min	4.67249E-01

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0.467249	5520	1200000	953
0.648088	6840	1225000	1464
1.016949	4541	1231000	915
2.375000	4038	1485000	1617
3.000000	1247	1930000	3404



**Histogram for Zestimate**

The UNIVARIATE Procedure  
Variable: bathrooms

Moments			
<b>N</b>	11328	<b>Sum Weights</b>	11328
<b>Mean</b>	1.98031427	<b>Sum Observations</b>	22433
<b>Std Deviation</b>	1.04740996	<b>Variance</b>	1.09706763
<b>Skewness</b>	1.97413883	<b>Kurtosis</b>	9.20546362
<b>Uncorrected SS</b>	56850.875	<b>Corrected SS</b>	12426.4851
<b>Coeff Variation</b>	52.8910983	<b>Std Error Mean</b>	0.00984102

Basic Statistical Measures			
Location		Variability	
<b>Mean</b>	1.980314	<b>Std Deviation</b>	1.04741
<b>Median</b>	2.000000	<b>Variance</b>	1.09707
<b>Mode</b>	2.000000	<b>Range</b>	13.50000
		<b>Interquartile Range</b>	1.00000

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	201.2306	<b>Pr &gt;  t </b>	<.0001
Sign	M	5664	<b>Pr &gt;=  M </b>	<.0001
Signed Rank	S	32083728	<b>Pr &gt;=  S </b>	<.0001

Quantiles (Definition 5)	
Level	Quantile
<b>100% Max</b>	14.0
<b>99%</b>	5.5
<b>95%</b>	4.0
<b>90%</b>	3.0
<b>75% Q3</b>	2.0
<b>50% Median</b>	2.0
<b>25% Q1</b>	1.0
<b>10%</b>	1.0
<b>5%</b>	1.0
<b>1%</b>	1.0
<b>0% Min</b>	0.5

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0.5	4792	12.0	5504
1.0	11327	12.5	7817
1.0	11326	13.5	10848
1.0	11323	14.0	345
1.0	11319	14.0	4210

### Histogram for Zestimate

The UNIVARIATE Procedure  
Variable: bedrooms

Moments			
N	11328	Sum Weights	11328
Mean	2.61467161	Sum Observations	29619
Std Deviation	1.29947058	Variance	1.68862378
Skewness	1.66247966	Kurtosis	9.42389926
Uncorrected SS	96571	Corrected SS	19127.0416
Coeff Variation	49.6991887	Std Error Mean	0.01220927

Basic Statistical Measures			
Location		Variability	
Mean	2.614672	Std Deviation	1.29947
Median	2.000000	Variance	1.68862
Mode	2.000000	Range	20.00000
		Interquartile Range	1.00000

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	214.1546	Pr >  t	<.0001
Sign	M	5595	Pr >=  M	<.0001
Signed Rank	S	31306823	Pr >=  S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	20
99%	7
95%	5
90%	4
75% Q3	3
50% Median	2
25% Q1	2
10%	1
5%	1
1%	0
0% Min	0

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0	11268	14	10286
0	11244	16	5688
0	11222	16	9200
0	11062	16	10726
0	11053	20	6742

### Histogram for Zestimate

The UNIVARIATE Procedure  
Variable: finishedsoft

Moments			
N	11328	Sum Weights	11328
Mean	1585.52904	Sum Observations	17960873
Std Deviation	922.023251	Variance	850126.875
Skewness	4.18246964	Kurtosis	62.838115
Uncorrected SS	3.81069E10	Corrected SS	9629387118
Coeff Variation	58.1524038	Std Error Mean	8.66293893

Basic Statistical Measures			
Location		Variability	
Mean	1585.529	Std Deviation	922.02325
Median	1362.000	Variance	850127
Mode	1200.000	Range	27274

Basic Statistical Measures				
Location		Variability		
		Interquartile Range		
		857.50000		

Tests for Location: Mu0=0				
Test	Statistic	p Value		
Student's t	t	183.0244	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001
Signed Rank	S	32083728	Pr >=  S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	27275.0
99%	4992.0
95%	3215.0
90%	2600.0
75% Q3	1876.5
50% Median	1362.0
25% Q1	1019.0
10%	780.0
5%	672.0
1%	493.0
0% Min	1.0

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
1	11135	9540	5703
1	10966	9699	9649
1	10651	10000	8086
1	7382	10236	2749
1	5102	27275	4996

### Histogram for Zestimate

The UNIVARIATE Procedure  
Variable: lastsolddate

Moments			
N	11328	Sum Weights	11328
Mean	19851.3646	Sum Observations	224876258
Std Deviation	527.668851	Variance	278434.416
Skewness	-8.5693936	Kurtosis	194.597322
Uncorrected SS	4.46725E12	Corrected SS	3153826634
Coeff Variation	2.65809864	Std Error Mean	4.95775245

Basic Statistical Measures			
Location		Variability	
Mean	19851.36	Std Deviation	527.66885
Median	19877.00	Variance	278434
Mode	19600.00	Range	16614
		Interquartile Range	627.00000

Note: The mode displayed is the smallest of 2 modes with a count of 46.

Tests for Location: Mu0=0				
Test	Statistic	p Value		
Student's t	t	4004.106	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001
Signed Rank	S	32083728	Pr >=  S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	20541
99%	20501
95%	20440
90%	20384
75% Q3	20206
50% Median	19877
25% Q1	19579
10%	19318
5%	19250
1%	18914
0% Min	3927

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
3927	9926	20529	10771
4688	4736	20531	8130
7063	6907	20531	8929
10379	8434	20535	9141
11774	3974	20541	8676

### Histogram for Zestimate

The UNIVARIATE Procedure  
Variable: lastsoldprice

Moments			
N	11328	Sum Weights	11328
Mean	1263970.37	Sum Observations	1.43183E10
Std Deviation	1042159.8	Variance	1.0861E12
Skewness	5.30103726	Kurtosis	55.2651108
Uncorrected SS	3.04001E16	Corrected SS	1.23022E16
Coeff Variation	82.4512841	Std Error Mean	9791.69094

Basic Statistical Measures			
Location		Variability	
Mean	1263970	Std Deviation	1042160
Median	990000	Variance	1.0861E12
Mode	1200000	Range	23888465
		Interquartile Range	721000

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	129.086	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001
Signed Rank	S	32083728	Pr >=  S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	23889000
99%	5500000
95%	2885000
90%	2150000
75% Q3	1450000
50% Median	990000
25% Q1	729000
10%	555000
5%	450000
1%	240000
0% Min	535

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
535	5520	14000000	3353
600	4541	16760000	1884
1000	6840	17100000	522
3000	1247	18000000	8230
5450	5064	23889000	2749

### Histogram for Zestimate

The UNIVARIATE Procedure  
Variable: totalrooms

Moments			
N	11328	Sum Weights	11328
Mean	6.11202331	Sum Observations	69237
Std Deviation	12.1268378	Variance	147.060196
Skewness	98.5980606	Kurtosis	10222.9567
Uncorrected SS	2088929	Corrected SS	1665750.84
Coeff Variation	198.409549	Std Error Mean	0.11393862

Basic Statistical Measures			
Location		Variability	
Mean	6.112023	Std Deviation	12.12684
Median	5.000000	Variance	147.06020
Mode	5.000000	Range	1263

Basic Statistical Measures			
Location		Variability	
		Interquartile Range	
		3.00000	

Tests for Location: Mu0=0				
Test	Statistic	p Value		
Student's t	t	53.64312	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001
Signed Rank	S	32083728	Pr >=  S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	1264
99%	15
95%	11
90%	9
75% Q3	7
50% Median	5
25% Q1	4
10%	3
5%	3
1%	2
0% Min	1

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
1	11222	28	2749
1	10987	28	10726
1	10594	33	11174
1	10432	94	8299
1	9858	1264	8966

### Histogram for Zestimate

The UNIVARIATE Procedure  
Variable: yearbuilt

Moments			
N	11328	Sum Weights	11328
Mean	1948.48808	Sum Observations	22072473
Std Deviation	37.9069693	Variance	1436.93832
Skewness	0.34358094	Kurtosis	-1.2386919
Uncorrected SS	4.30242E10	Corrected SS	16276200.4
Coeff Variation	1.94545554	Std Error Mean	0.35615779

Basic Statistical Measures			
Location		Variability	
Mean	1948.488	Std Deviation	37.90697
Median	1940.000	Variance	1437
Mode	1900.000	Range	156.00000
		Interquartile Range	70.50000

Tests for Location: Mu0=0				
Test	Statistic	p Value		
Student's t	t	5470.856	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001
Signed Rank	S	32083728	Pr >=  S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	2016.0
99%	2014.0
95%	2009.0
90%	2006.0
75% Q3	1986.0
50% Median	1940.0
25% Q1	1915.5
10%	1904.0
5%	1900.0
1%	1890.0
0% Min	1860.0

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
1860	6635	2015	3402
1863	9649	2015	4210
1866	7735	2015	6878
1866	3733	2015	10416
1870	5267	2016	8513

### Histogram for Zestimate

The UNIVARIATE Procedure  
Variable: zestimate

Moments			
N	11328	Sum Weights	11328
Mean	1565795.91	Sum Observations	1.77373E10
Std Deviation	1229500.14	Variance	1.51167E12
Skewness	4.13692999	Kurtosis	25.6680924
Uncorrected SS	4.48957E16	Corrected SS	1.71227E16
Coeff Variation	78.5223752	Std Error Mean	11551.8612

Basic Statistical Measures			
Location		Variability	
Mean	1565796	Std Deviation	1229500
Median	1231054	Variance	1.51167E12
Mode	1101226	Range	15100868
		Interquartile Range	826533

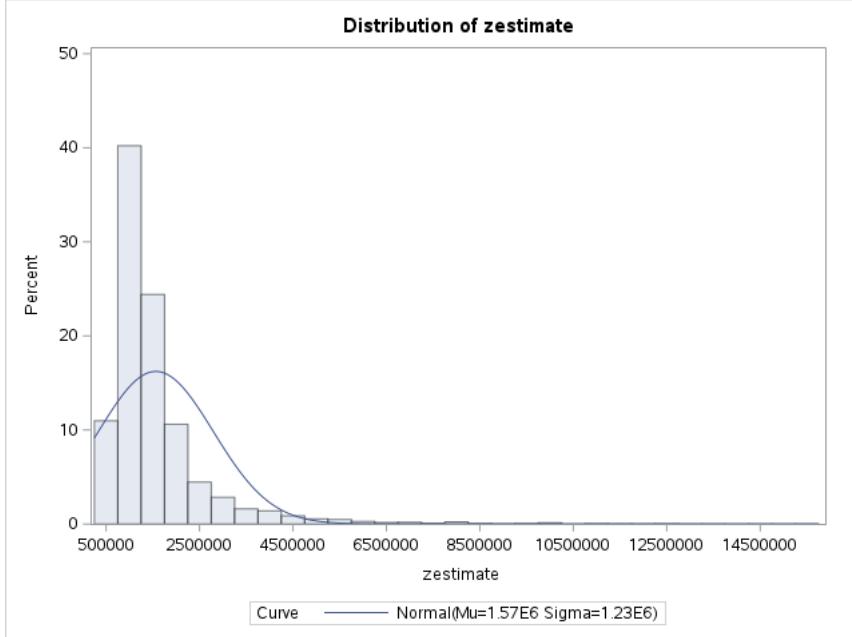
Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	135.5449	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001
Signed Rank	S	32083728	Pr >=  S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	15533253
99%	7045051
95%	3699838
90%	2639675
75% Q3	1731778
50% Median	1231054
25% Q1	905245
10%	737918
5%	641875
1%	536665
0% Min	432385

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
432385	1733	14272621	2749
433999	2498	14658165	8230
458909	1218	15048610	426
464318	2708	15463813	8086
466307	1112	15533253	522

### Histogram for Zestimate

The UNIVARIATE Procedure



### Histogram for Zestimate

The UNIVARIATE Procedure  
Fitted Normal Distribution for zestimate

Parameters for Normal Distribution		
Parameter	Symbol	Estimate
Mean	Mu	1565796
Std Dev	Sigma	1229500

Goodness-of-Fit Tests for Normal Distribution				
Test	Statistic		p Value	
Kolmogorov-Smirnov	D	0.19733	Pr > D	<0.010
Cramer-von Mises	W-Sq	189.66562	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	1032.65403	Pr > A-Sq	<0.005

Quantiles for Normal Distribution		
Percent	Quantile	
	Observed	Estimated
1.0	536665	-1294449.13
5.0	641875	-456551.85
10.0	737918	-9871.92
25.0	905245	736510.67
50.0	1231054	1565795.91
75.0	1731778	2395081.15
90.0	2639675	3141463.74
95.0	3699838	3588143.68
99.0	7045051	4426040.95

### Histogram for Zestimate

The UNIVARIATE Procedure  
Variable: zipcode

Moments			
N	11328	Sum Weights	11328
Mean	94116.9141	Sum Observations	1066156403
Std Deviation	9.40033672	Variance	88.3663305
Skewness	0.77787596	Kurtosis	0.97142959
Uncorrected SS	1.00343E14	Corrected SS	1000925.43
Coeff Variation	0.00998794	Std Error Mean	0.08832157

Basic Statistical Measures			
Location		Variability	
Mean	94116.91	Std Deviation	9.40034
Median	94115.00	Variance	88.36633
Mode	94110.00	Range	56.00000
		Interquartile Range	13.00000

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	1065616	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001
Signed Rank	S	32083728	Pr >=  S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	94158
99%	94134
95%	94133
90%	94131
75% Q3	94123
50% Median	94115
25% Q1	94110
10%	94107
5%	94103
1%	94102
0% Min	94102

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
94102	11295	94158	10367
94102	11268	94158	10686
94102	11245	94158	10973
94102	11238	94158	10979
94102	11231	94158	11263

### Histogram for Zestimate

The UNIVARIATE Procedure  
Variable: price\_per\_square\_foot

Moments			
N	11328	Sum Weights	11328
Mean	1760.90617	Sum Observations	19947545.1
Std Deviation	33864.2358	Variance	1146786467
Skewness	40.4361909	Kurtosis	1791.28167
Uncorrected SS	1.30248E13	Corrected SS	1.29897E13
Coeff Variation	1923.11415	Std Error Mean	318.173979

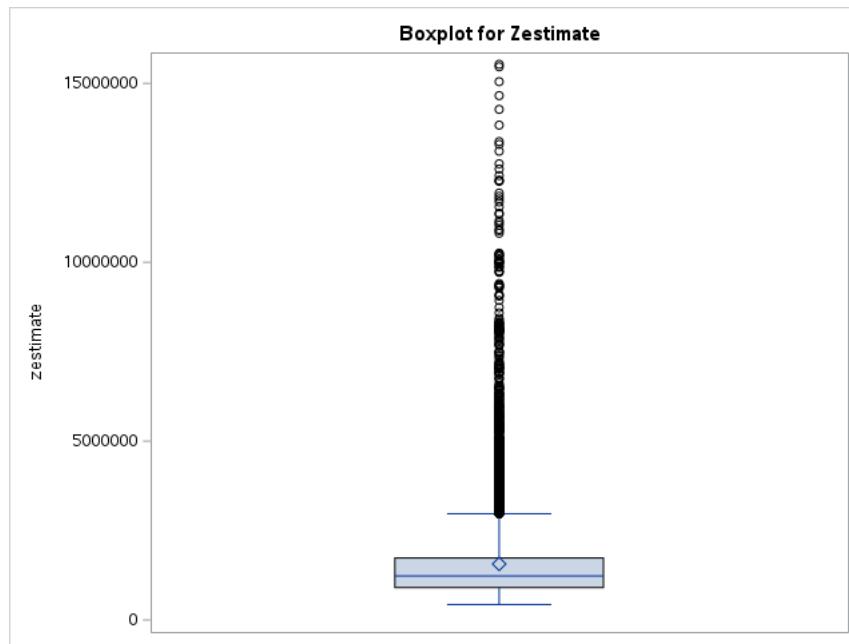
Basic Statistical Measures			
Location		Variability	
Mean	1760.906	Std Deviation	33864
Median	781.866	Variance	1146786467
Mode	1000.000	Range	1930000
		Interquartile Range	417.41596

Tests for Location: Mu0=0				
Test	Statistic		p Value	
Student's t	t	5.534413	Pr >  t	<.0001
Sign	M	5664	Pr >=  M	<.0001
Signed Rank	S	32083728	Pr >=  S	<.0001

Quantiles (Definition 5)	
Level	Quantile
100% Max	1.93000E+06
99%	2.02203E+03
95%	1.38053E+03
90%	1.21385E+03
75% Q3	9.99683E+02
50% Median	7.81866E+02
25% Q1	5.82267E+02
10%	4.29630E+02
5%	3.37621E+02
1%	1.86296E+02
0% Min	4.67249E-01

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
0.467249	5520	1200000	953
0.648088	6840	1225000	1464
1.016949	4541	1231000	915
2.375000	4038	1485000	1617

Extreme Observations			
Lowest		Highest	
Value	Obs	Value	Obs
3.000000	1247	1930000	3404



**Boxplot for Zestimate**

Obs	median_lastsoldprice
1	990000

**Boxplot for Zestimate**

Obs	median_zestimate
1	1231054

**T-Test for Average Zestimate by Zip Code**

ZipCode	AverageZestimate
94107	1362007
94109	1580086

**T-Test for Average Zestimate by Zip Code**

The TTEST Procedure

Variable: zestimate

zipcode	Method	N	Mean	Std Dev	Std Err	Minimum	Maximum
94107		856	1362007	619937	21189.0	548210	5044710
94109		653	1580086	1311224	51312.1	495193	11933093
Diff (1-2)	Pooled		-218078	980765	50958.4		
Diff (1-2)	Satterthwaite		-218078		55515.0		

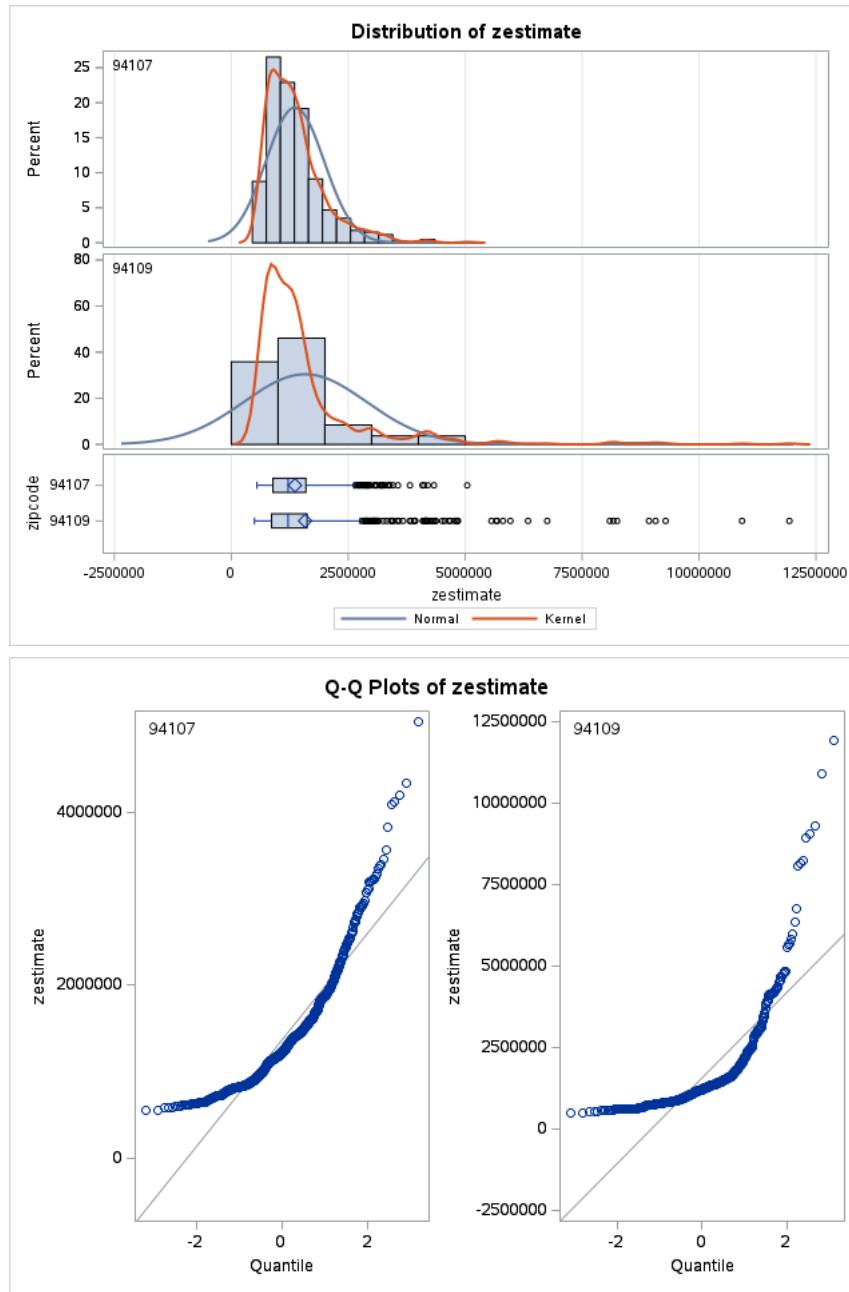
zipcode	Method	Mean	95% CL Mean	Std Dev	95% CL Std Dev
94107		1362007	1320419	1403596	619937
94109		1580086	1479329	1680843	1311224
Diff (1-2)	Pooled	-218078	-318035	-118121	980765
Diff (1-2)	Satterthwaite	-218078	-327037	-109120	946970
					1017080

Method	Variances	DF	t Value	Pr >  t
Pooled	Equal	1507	-4.28	<.0001
Satterthwaite	Unequal	873.94	-3.93	<.0001

Equality of Variances

Method	Num DF	Den DF	F Value	Pr > F
--------	--------	--------	---------	--------

Equality of Variances				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	652	855	4.47	<.0001



#### T-Test for Average Zestimate by Zip Code

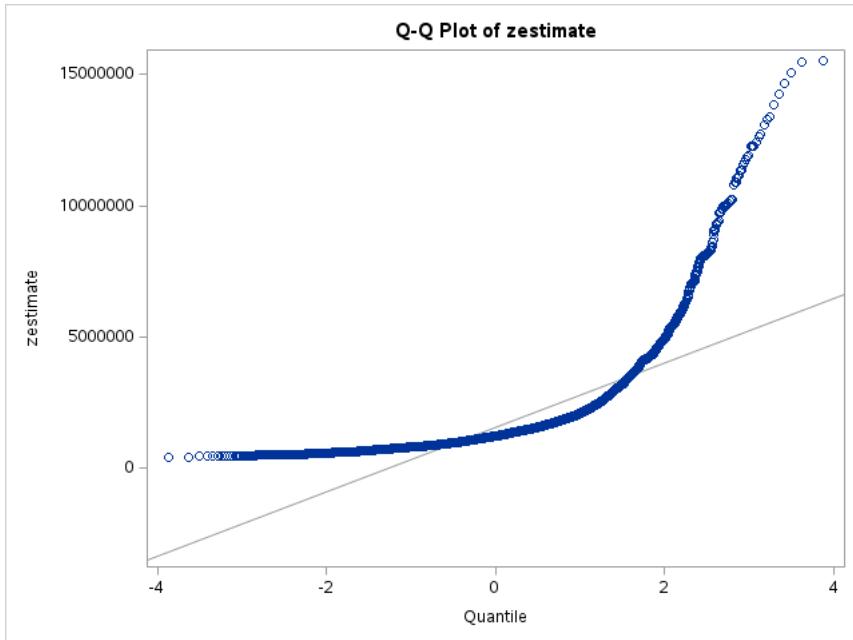
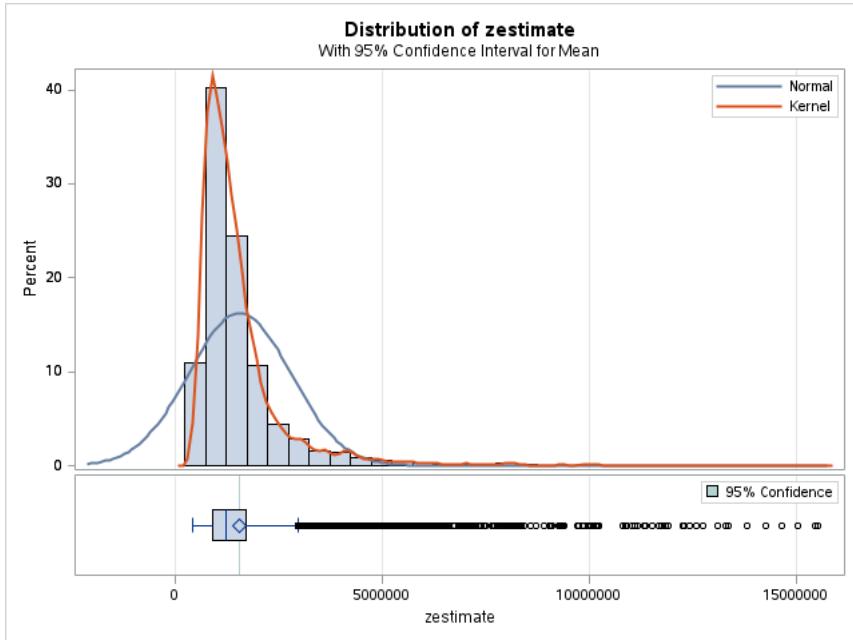
The TTEST Procedure

Variable: zestimate

N	Mean	Std Dev	Std Err	Minimum	Maximum
11328	1565796	1229500	11551.9	432385	15533253

Mean	95% CL Mean	Std Dev	95% CL Std Dev
1565796	1543152	1588440	1229500

DF	t Value	Pr >  t
11327	135.54	<.0001

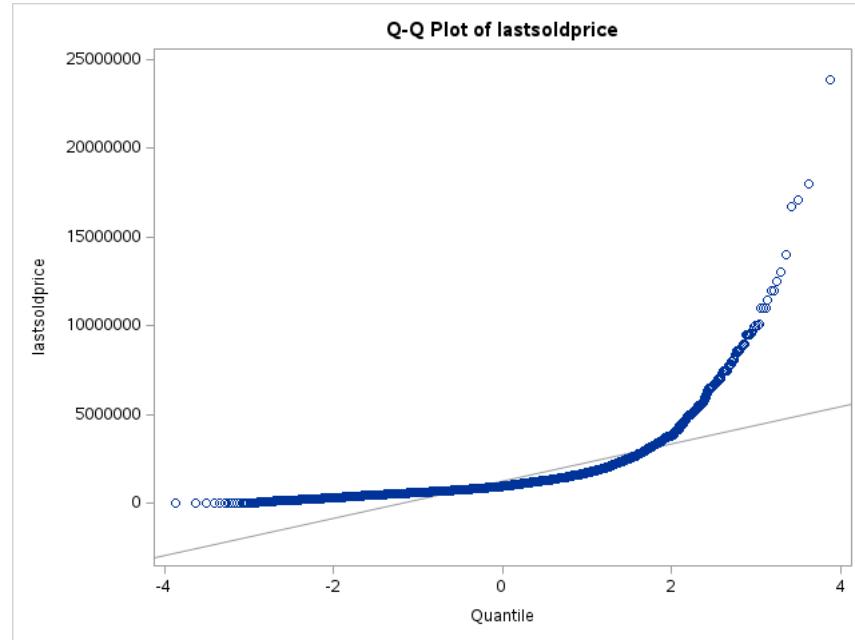
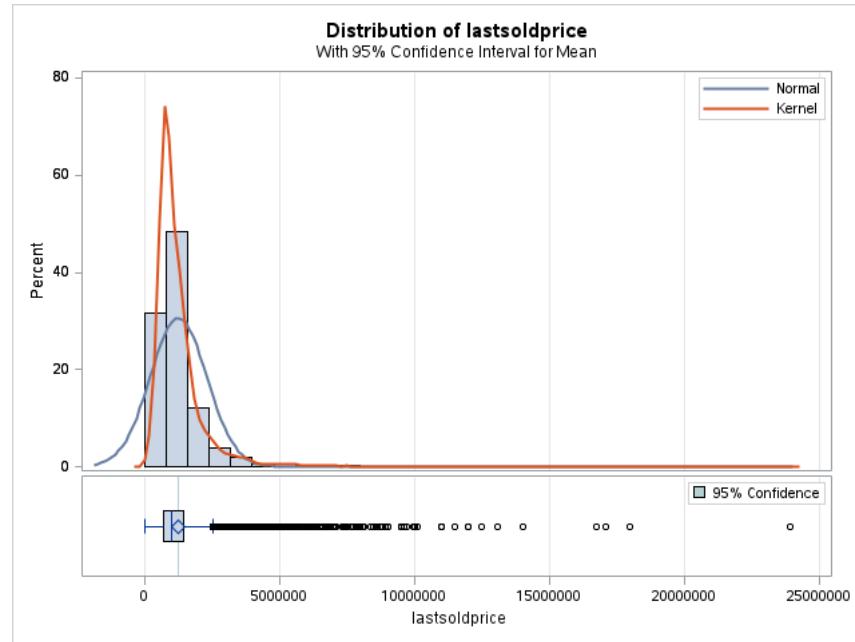


Variable: lastsoldprice

N	Mean	Std Dev	Std Err	Minimum	Maximum
11328	1263970	1042160	9791.7	535.0	23889000

Mean	95% CL Mean	Std Dev	95% CL Std Dev
1263970	1244777	1283164	1042160

DF	t Value	Pr >  t
11327	129.09	<.0001



#### T-Test for Average Zestimate by Zip Code

The FREQ Procedure

bedrooms	Table of bedrooms by usecode											
	Apartment	Condominium	Cooperative	Duplex	Miscellaneous	Mobile	MultiFamily2To4	MultiFamily5Plus	SingleFamily	Townhouse	Total	
0	0 0.00 0.00 0.00	100 0.88 72.46 2.08	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	0 0.00 0.00 0.00	22 0.19 15.94 4.53	0 0.00 0.00 0.00	16 0.14 11.59 0.28	0 0.00 0.00 0.00	138 1.22 0.00 0.00	
1	1 0.01 0.06 33.33	1462 12.91 86.00 30.46	0 0.00 0.00 0.00	2 0.02 0.12 1.37	1 0.01 0.06 5.88	0 0.00 0.00 0.00	73 0.64 4.29 15.02	0 0.00 0.00 0.00	153 1.35 9.00 2.64	8 0.07 0.47 12.12	1700 15.01	
2	2 0.02 0.05 66.67	2178 19.23 54.07 45.38	1 0.01 0.02 33.33	11 0.10 0.27 7.53	7 0.06 0.17 41.18	1 0.01 0.02 50.00	105 0.93 2.61 21.60	2 0.02 0.05 100.00	1684 14.87 41.81 29.02	37 0.33 0.92 56.06	4028 35.56	
3	0 0.00 0.00 0.00	906 8.00 27.55 18.88	1 0.01 0.03 33.33	17 0.15 0.52 11.64	4 0.04 0.12 23.53	1 0.01 0.03 50.00	44 0.39 1.34 9.05	0 0.00 0.00 0.00	2299 20.29 69.92 39.62	16 0.14 0.49 24.24	3288 29.03	
4	0 0.00 0.00 0.00	143 1.26 9.83 2.98	0 0.00 0.00 0.00	47 0.41 0.23 32.19	2 0.02 0.14 11.76	0 0.00 0.00 0.00	77 0.68 5.29 15.84	0 0.00 0.00 0.00	1183 10.44 81.31 20.39	3 0.03 0.21 4.55	1455 12.84	
5	0 0.00 0.00 0.00	7 0.06 1.65 0.15	0 0.00 0.00 0.00	16 0.14 3.78 10.96	1 0.01 0.24 5.88	0 0.00 0.00 0.00	62 0.55 14.66 12.76	0 0.00 0.00 0.00	337 2.97 79.67 5.81	0 0.00 0.00 0.00	423 3.73	

Table of bedrooms by usecode													
bedrooms	usecode												
	Apartment	Condominium	Cooperative	Duplex	Miscellaneous	Mobile	MultiFamily2To4	MultiFamily5Plus	SingleFamily	Townhouse	Total		
6	0	1	0	26	0	0	48	0	96	2	173		
	0.00	0.01	0.00	0.23	0.00	0.00	0.42	0.00	0.85	0.02	1.53		
	0.00	0.58	0.00	15.03	0.00	0.00	27.75	0.00	55.49	1.16			
	0.00	0.02	0.00	17.81	0.00	0.00	9.88	0.00	1.65	3.03			
7	0	0	0	16	1	0	17	0	20	0	54		
	0.00	0.00	0.00	0.14	0.01	0.00	0.15	0.00	0.18	0.00	0.48		
	0.00	0.00	0.00	29.63	1.85	0.00	31.48	0.00	37.04	0.00			
	0.00	0.00	0.00	10.96	5.88	0.00	3.50	0.00	0.34	0.00			
8	0	0	0	8	0	0	13	0	6	0	27		
	0.00	0.00	0.00	0.07	0.00	0.00	0.11	0.00	0.05	0.00	0.24		
	0.00	0.00	0.00	29.63	0.00	0.00	48.15	0.00	22.22	0.00			
	0.00	0.00	0.00	5.48	0.00	0.00	2.67	0.00	0.10	0.00			
9	0	2	0	3	0	0	15	0	8	0	28		
	0.00	0.02	0.00	0.03	0.00	0.00	0.13	0.00	0.07	0.00	0.25		
	0.00	7.14	0.00	10.71	0.00	0.00	53.57	0.00	28.57	0.00			
	0.00	0.04	0.00	2.05	0.00	0.00	3.09	0.00	0.14	0.00			
10	0	0	0	0	1	0	4	0	0	0	5		
	0.00	0.00	0.00	0.00	0.01	0.00	0.04	0.00	0.00	0.00	0.04		
	0.00	0.00	0.00	20.00	0.00	0.00	80.00	0.00	0.00	0.00			
	0.00	0.00	0.00	5.88	0.00	0.00	0.82	0.00	0.00	0.00			
11	0	0	0	0	0	0	0	0	1	0	1		
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01		
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00			
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00			
12	0	0	0	0	0	0	2	0	0	0	2		
	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.02		
	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00		
	0.00	0.00	0.00	0.00	0.00	0.00	0.41	0.00	0.00	0.00	0.00		
13	0	0	0	0	0	0	1	0	0	0	1		
	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01		
	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00		
	0.00	0.00	0.00	0.00	0.00	0.00	0.21	0.00	0.00	0.00	0.00		
14	0	0	1	0	0	0	0	0	0	0	1		
	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01		
	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
	0.00	33.33	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
16	0	0	0	0	0	0	3	0	0	0	3		
	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.03		
	0.00	0.00	0.00	0.00	0.00	0.00	100.00	0.00	0.00	0.00	0.00		
	0.00	0.00	0.00	0.00	0.00	0.00	0.62	0.00	0.00	0.00	0.00		
20	0	1	0	0	0	0	0	0	0	0	0		
	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
<b>Total</b>	3	4800	3	146	17	2	486	2	5803	66	11328		
	0.03	42.37	0.03	1.29	0.15	0.02	4.29	0.02	51.23	0.58	100.00		

Statistics for Table of bedrooms by usecode

Statistic	DF	Value	Prob
Chi-Square	144	8899.5646	<.0001
Likelihood Ratio Chi-Square	144	4403.9478	<.0001
Mantel-Haenszel Chi-Square	1	1947.5340	<.0001
Phi Coefficient		0.8864	
Contingency Coefficient		0.6633	
Cramer's V		0.2955	
WARNING: 78% of the cells have expected counts less than 5. Chi-Square may not be a valid test.			

Sample Size = 11328

## T-Test for Average Estimate by Zip Code

The FREQ Procedure

Frequency Percent Row Pct Col Pct	Table of bedrooms by bathrooms																	bathrooms							
bedrooms	0.5	1	1.1	1.25	1.3	1.5	1.75	1.8	2	2.1	2.2	2.25	2.3	2.5	2.75	3	3.1	3.25	3.3	3.5	3.75	4	4.4	4.5	
0	1	120	0	0	0	1	0	0	10	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	
	0.01	1.06	0.00	0.00	0.00	0.01	0.00	0.00	0.09	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	
	0.72	86.96	0.00	0.00	0.00	0.72	0.00	0.00	7.25	0.00	0.00	0.00	0.00	0.00	0.00	2.17	0.00	0.00	0.00	0.00	0.00	1.45	0.00	0.00	0.00
	100.00	3.08	0.00	0.00	0.00	0.23	0.00	0.00	0.24	0.00	0.00	0.00	0.00	0.00	0.00	0.26	0.00	0.00	0.00	0.00	0.00	0.55	0.00	0.00	0.00
1	0	1464	0	3	1	93	1	0	134	0	0	0	0	0	0	0	3	0	0	0	0	0	1	0	0
	0.00	12.92	0.00	0.03	0.01	0.82	0.01	0.00	1.18	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
	0.00	86.12	0.00	0.18	0.06	5.47	0.06	0.00	7.88	0.00	0.00	0.00	0.00	0.00	0.00	0.18	0.00	0.00	0.00	0.00	0.00	0.06	0.00	0.00	0.00
	0.00	37.52	0.00	13.04	25.00	21.28	12.50	0.00	3.22	0.00	0.00	0.00	0.00	0.00	0.00	0.26	0.00	0.00	0.00	0.00	0.00	0.28	0.00	0.00	0.00
2	0	1789	1	16	2	164	0	1	1809	0	2	3	0	142	0	88	0	0	0	0	2	0	7	0	1
	0.00	15.79	0.01	0.14	0.02	1.45	0.00	0.01	15.97	0.00	0.02	0.03	0.00	1.25	0.00	0.78	0.00	0.00	0.00	0.02	0.00	0.06	0.00	0.01	0.00
	0.00	44.41	0.02	0.40	0.05	4.07	0.00	0.02	44.91	0.00	0.05	0.07	0.00	3.53	0.00	2.18	0.00	0.00	0.00	0.05	0.00	0.17	0.00	0.02	0.00
	0.00	45.85	50.00	69.57	50.00	37.53	0.00	100.00	43.49	0.00	100.00	13.64	0.00	26.94	0.00	7.59	0.00	0.00	0.00	0.61	0.00	1.93	0.00	0.92	0.00
3	0	484	1	4	1	163	6	0	1709	0	0	12	2	269	4	469	1	1	0	106	0	42	0	6	
	0.00	4.27	0.01	0.04	0.01	1.44	0.05	0.00	15.09	0.00	0.00	0.11	0.02	2.37	0.04	4.14	0.01	0.01	0.00	0.94	0.00	0.37	0.00	0.05	0.00
	0.00	14.72	0.03	0.12	0.03	4.96	0.18	0.00	51.98	0.00	0.00	0.36	0.06	8.18	0.12	14.26	0.03	0.03	0.00	3.22	0.00	1.28	0.00	0.18	0.00
	0.00	12.40	50.00	17.39	25.00	37.30	75.00	0.00	41.08	0.00	0.00	54.55	66.67	51.04	0.00	66.67	0.00	100.00	33.33	0.00	32.52	0.00	11.60	0.00	5.50
4	0	41	0	0	0																				

## Table of bedrooms by bathrooms

	bathrooms																							
bedrooms	0.5	1	1.1	1.25	1.3	1.5	1.75	1.8	2	2.1	2.2	2.25	2.3	2.5	2.75	3	3.1	3.25	3.3	3.5	3.75	4	4.4	4.5
6	0	1	0	0	0	0	0	0	8	0	0	0	1	0	1	44	0	0	0	11	0	41	1	9
	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.01	0.00	0.01	0.39	0.00	0.00	0.00	0.10	0.00	0.36	0.01	0.08
	0.00	0.58	0.00	0.00	0.00	0.00	0.00	0.00	4.62	0.00	0.00	0.00	0.58	0.00	0.58	25.43	0.00	0.00	0.00	6.36	0.00	23.70	0.58	5.20
	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.19	0.00	0.00	0.00	33.33	0.00	16.67	3.79	0.00	0.00	0.00	3.37	0.00	11.33	100.00	8.26
7	0	0	0	0	0	0	0	0	6	0	0	0	0	0	0	8	0	0	0	0	0	8	0	1
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.00	0.00	0.00	0.00	0.07	0.00	0.01
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	11.11	0.00	0.00	0.00	0.00	0.00	0.00	14.81	0.00	0.00	0.00	0.00	0.00	14.81	0.00	1.85
8	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	4	0	0	0	0	0	5	0	0
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.04	0.00	0.00
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.70	0.00	0.00	0.00	0.00	0.00	0.00	14.81	0.00	0.00	0.00	0.00	0.00	18.52	0.00	0.00
9	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	3	0	0	0	1	0	1	0	0
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	3.57	0.00	0.00	0.00	0.00	0.00	0.00	10.71	0.00	0.00	0.00	0.00	0.00	3.57	0.00	0.00
10	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	3	0	0	0	1	0	1	0	0
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.26	0.00	0.00	0.00	0.00	0.00	0.31	0.00	0.00
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
11	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
12	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	50.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
13	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
20	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	0.00	100.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total	1	3902	2	23	4	437	8	1	4160	1	2	22	3	527	6	1160	1	3	1	326	6	362	1	109
	0.01	34.45	0.02	0.20	0.04	3.86	0.07	0.01	36.72	0.01	0.02	0.19	0.03	4.65	0.05	10.24	0.01	0.03	0.01	2.88	0.05	3.20	0.01	0.96

## Statistics for Table of bedrooms by bathrooms

Statistic	DF	Value	Prob
Chi-Square	624	17999.9644	<.0001
Likelihood Ratio Chi-Square	624	8662.2665	<.0001
Mantel-Haenszel Chi-Square	1	5778.5307	<.0001
Phi Coefficient		1.2605	
Contingency Coefficient		0.7834	
Cramer's V		0.3151	

**WARNING:** 90% of the cells have expected counts less than 5. Chi-Square may not be a valid test.

**Sample Size = 11328**

## Correlation of Numerical Variables with Zestimate

## The CORR Procedure

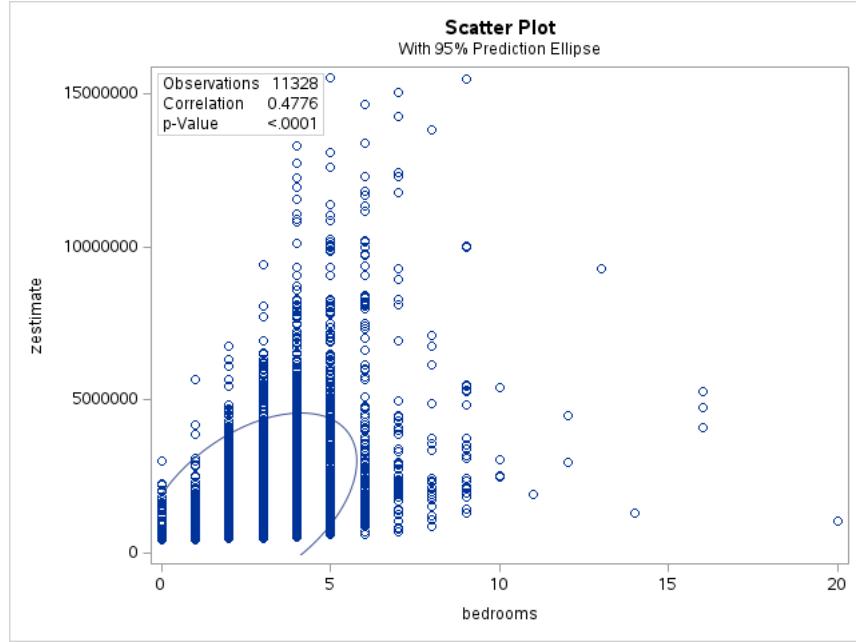
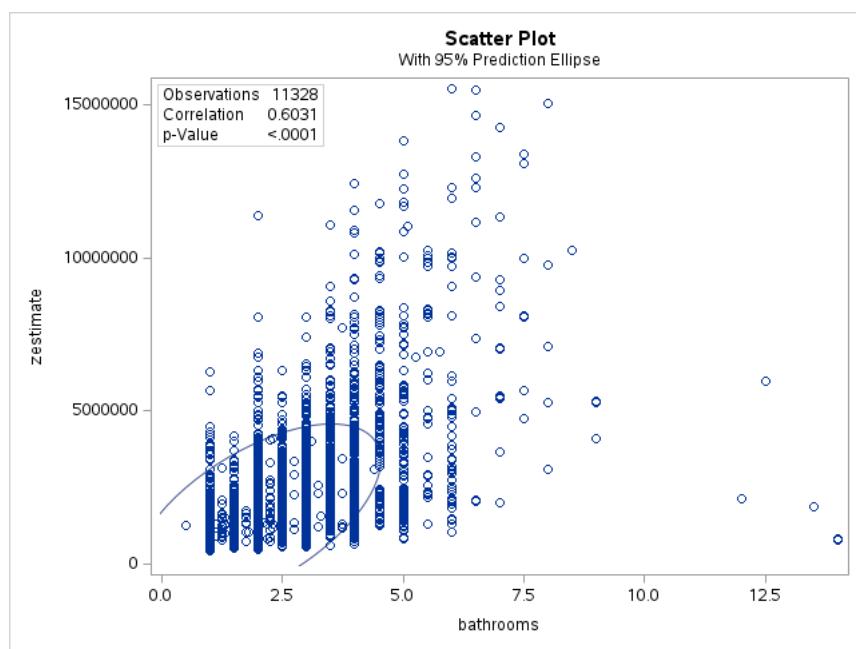
<b>1 With Variables:</b>	zestimate
<b>10 Variables:</b>	bathrooms bedrooms finishedsquarefeet lastsolddate lastsoldprice totalrooms yearbuilt zestimate zipcode price_per_square_foot

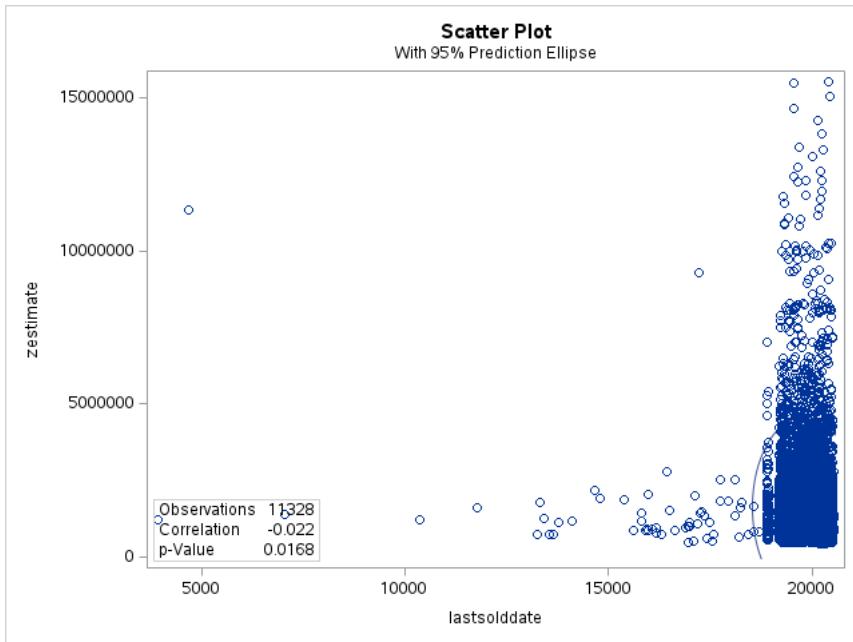
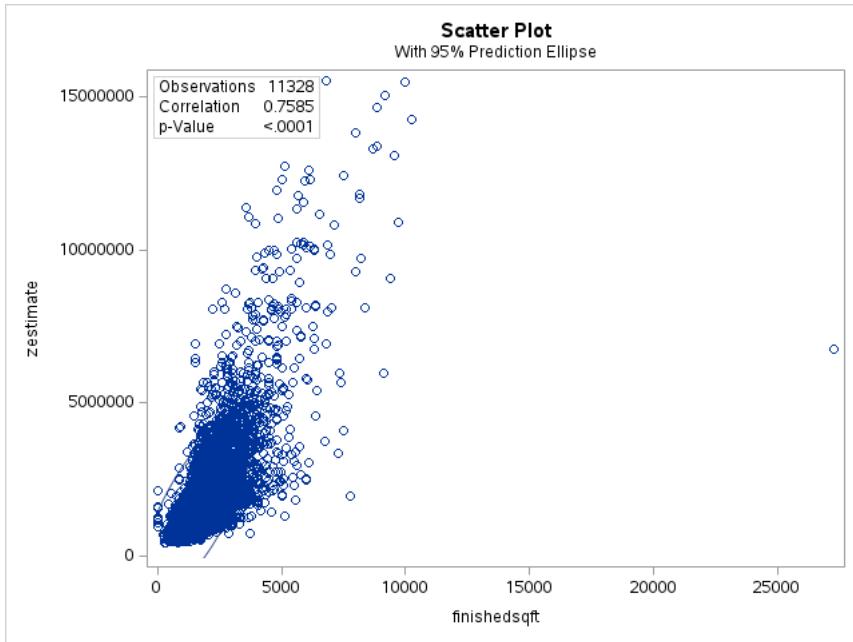
Simple Statistics						
Variable	N	Mean	Std Dev	Sum	Minimum	Maximum
zestimate	11328	1565796	1229500	1.77373E10	432385	15533253
bathrooms	11328	1.98031	1.04741	22433	0.50000	14.00000
bedrooms	11328	2.61467	1.29947	29619	0	20.00000
finishedsqft	11328	1586	922.02325	17960873	1.00000	27275
lastsolddate	11328	19851	527.66885	224876258	3927	20541
lastsoldprice	11328	1263970	1042160	1.43183E10	535.00000	23889000
totalrooms	11328	6.11202	12.12684	69237	1.00000	1264
yearbuilt	11328	1948	37.90697	22072473	1860	2016
zipcode	11328	94117	9.40034	1066156403	94102	94158
price_per_square_foot	11328	1761	33864	19947545	0.46725	1930000

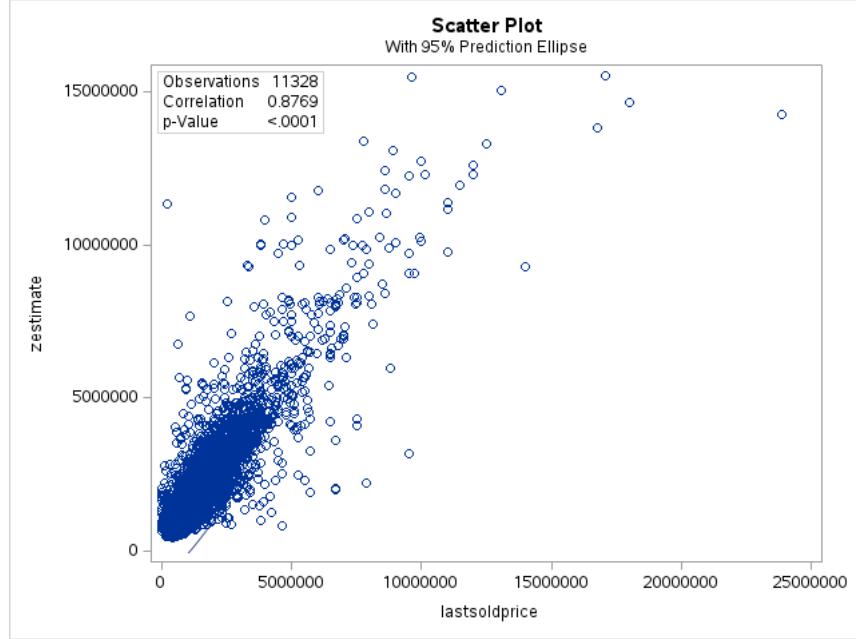
Pearson Correlation Coefficients, N = 11328											
	Prob >  r  under H0: Rho=0										
	bathrooms	bedrooms	finishedsqft	lastsolddate	lastsoldprice	totalrooms	yearbuilt	zestimate	zipcode	price_per_square_foot	
zestimate	0.60314 <.0001	0.47764 <.0001	0.75848 <.0001	-0.02247 0.0168	0.87694 <.0001	0.11197 <.0001	-0.21249 <.0001	1.00000	0.03444 0.0002	0.00093 0.9212	

Correlation of Numerical Variables with Zestimate

The CORR Procedure







question 13) regression model first three most correlated variables.

The REG Procedure  
Model: MODEL1  
Dependent Variable: zestimate

Number of Observations Read	11328
Number of Observations Used	11328

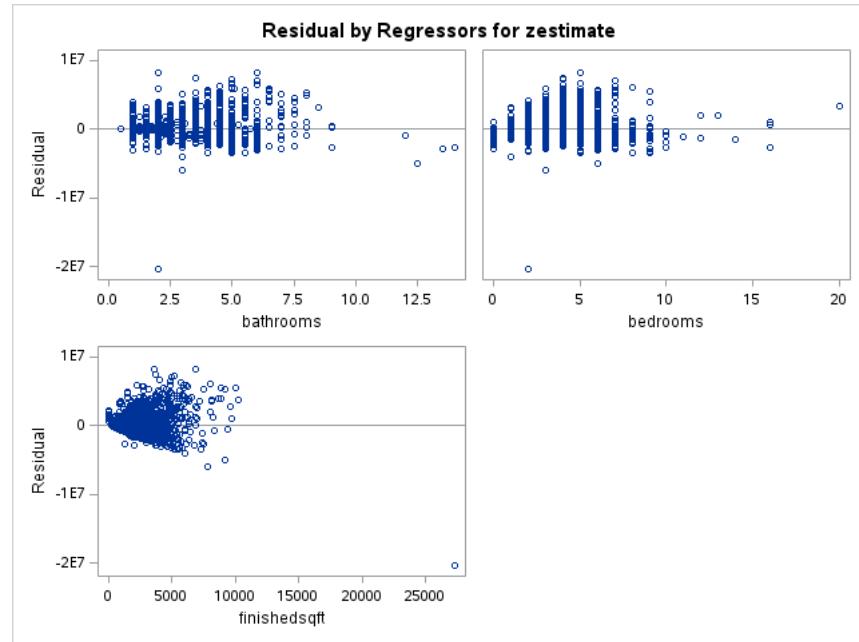
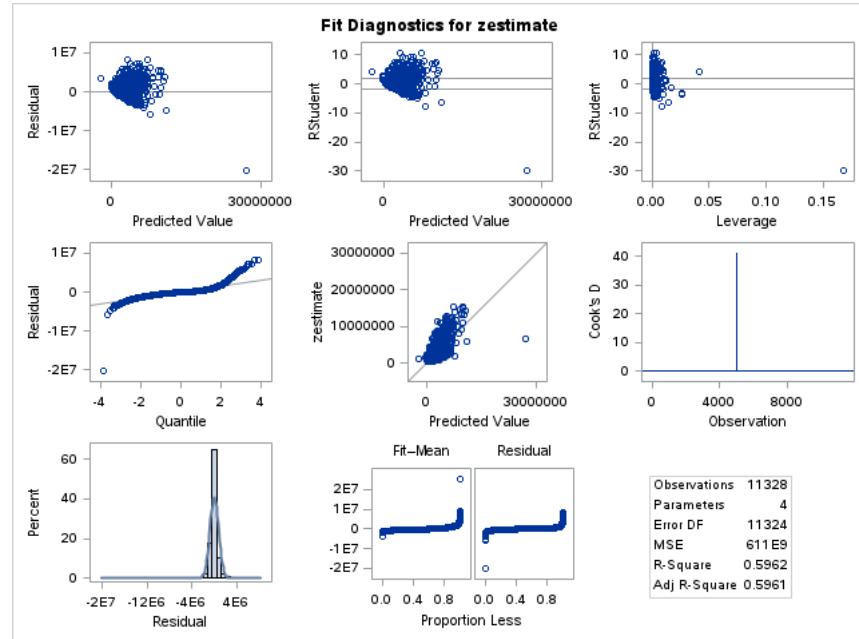
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	1.020917E16	3.403056E15	5574.03	<.0001
Error	11324	6.913526E15	6.105197E11		
Corrected Total	11327	1.712269E16			

Root MSE	781358	R-Square	0.5962
Dependent Mean	1565796	Adj R-Sq	0.5961
Coeff Var	49.90163		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr >  t
Intercept	1	-16025	17215	-0.93	0.3520
bathrooms	1	234880	11156	21.05	<.0001
bedrooms	1	-173019	8719.04238	-19.84	<.0001
finishedsqft	1	989.62021	12.37630	79.96	<.0001

question 13) regression model first three most correlated variables.

The REG Procedure  
Model: MODEL1  
Dependent Variable: zestimate



question 14) regression model first five most correlated variables.

The REG Procedure  
Model: MODEL1  
Dependent Variable: zestimate

Number of Observations Read	11328
Number of Observations Used	11328

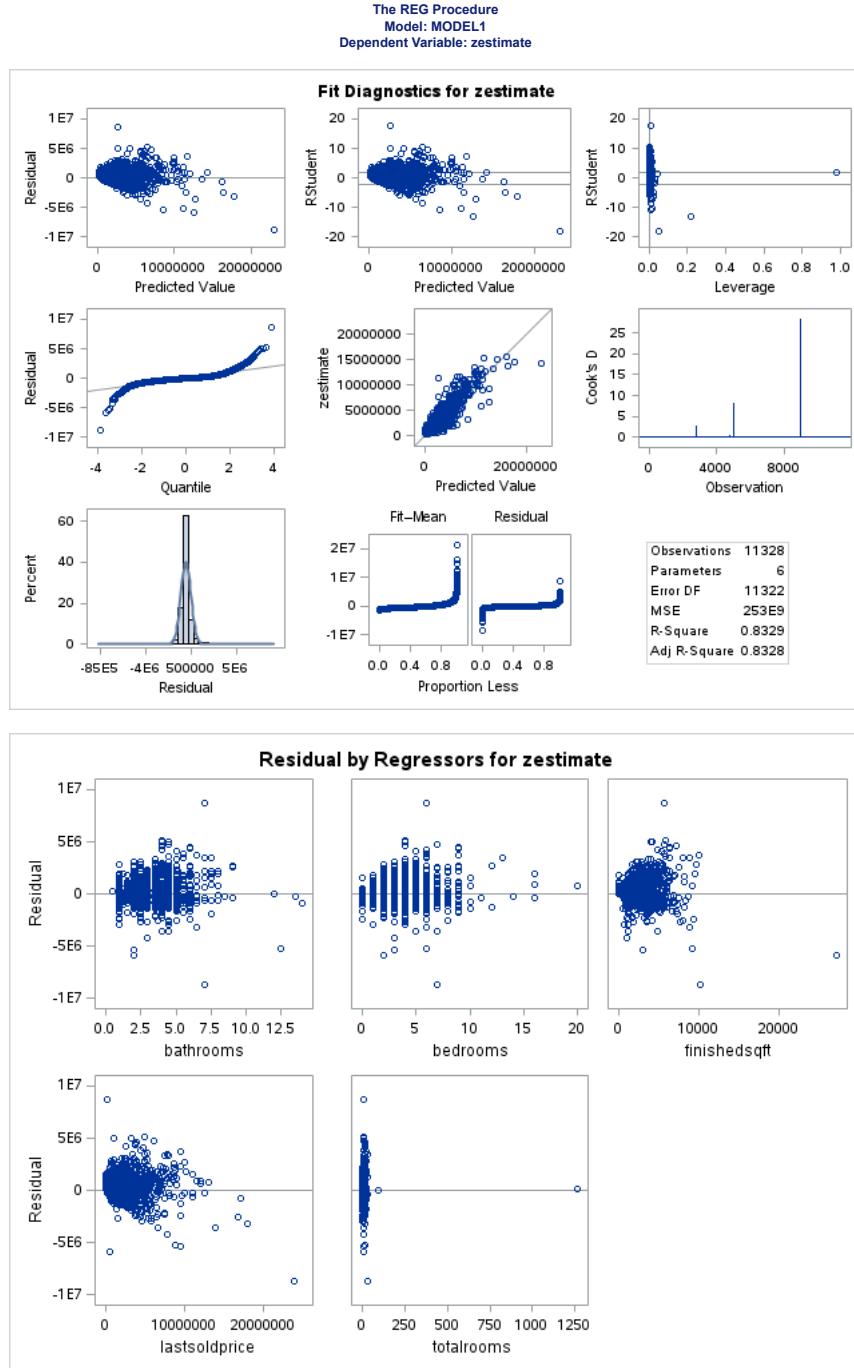
Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	5	1.426071E16	2.852141E15	11283.1	<.0001
Error	11322	2.861987E15	2.52781E11		
Corrected Total	11327	1.712269E16			

Root MSE	502773	R-Square	0.8329
Dependent Mean	1565796	Adj R-Sq	0.8328
Coeff Var	32.10976		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr >  t
Intercept	1	-106275	11127	-9.55	<.0001

Parameter Estimates				
Variable	DF	Parameter Estimate	Standard Error	t Value
bathrooms	1	50735	7324.21751	6.93
bedrooms	1	-44051	5710.51918	-7.71
finishedsqft	1	448.53948	9.05190	49.55
lastsoldprice	1	0.77216	0.00610	126.60
totalrooms	1	-62.77636	395.28076	-0.16

question 14) regression model first five most correlated variables.



question 14) regression model first five most correlated variables.

Obs	predicted_zestimate
1	2148807.90
2	1596614.38
3	2855494.04

Obs	predicted_zestimate
4	1788838.05

---

#### Average price\_per\_square\_foot by zipcode

Obs	zipcode	_TYPE_	_FREQ_	mean_price
1	.	0	11328	1760.91
2	94102	1	318	928.85
3	94103	1	359	3023.18
4	94104	1	1	866.67
5	94105	1	340	2856.13
6	94107	1	856	1660.76
7	94108	1	114	1001.95
8	94109	1	653	919.28
9	94110	1	935	3474.67
10	94111	1	38	823.54
11	94112	1	877	573.12
12	94114	1	651	1012.22
13	94115	1	534	940.64
14	94116	1	655	650.60
15	94117	1	473	2691.56
16	94118	1	466	834.23
17	94121	1	527	695.39
18	94122	1	601	642.40
19	94123	1	430	1188.12
20	94124	1	393	431.21
21	94127	1	435	710.00
22	94131	1	687	3703.84
23	94132	1	292	3190.92
24	94133	1	183	7832.94
25	94134	1	439	499.77
26	94158	1	71	21914.18

---

#### Average zestimate by totalrooms

Obs	totalrooms	_TYPE_	_FREQ_	mean_price
1	.	0	11328	1565795.91
2	1	1	32	864987.34
3	2	1	237	790535.91
4	3	1	966	872278.78
5	4	1	1859	1096460.69
6	5	1	2625	1269967.61
7	6	1	2015	1530213.50
8	7	1	1272	1738700.67
9	8	1	913	1927580.83
10	9	1	479	2472959.43
11	10	1	325	2840516.66
12	11	1	164	3267640.75
13	12	1	146	3047366.25
14	13	1	94	3692985.06
15	14	1	55	3530035.11
16	15	1	46	4273266.89
17	16	1	39	4120780.28
18	17	1	13	3178413.92
19	18	1	16	4178196.19
20	19	1	4	3739224.25
21	20	1	7	6404786.86
22	21	1	3	2012261.00
23	22	1	1	5684941.00
24	23	1	5	3106397.40
25	24	1	3	3662498.33
26	26	1	2	8765272.00
27	27	1	1	5430463.00
28	28	1	3	8381745.00
29	33	1	1	9304469.00
30	94	1	1	502285.00
31	1264	1	1	1591608.00

Procedure

Procedure