

Exploratory Data Analysis Report

Report by dlookr package

2019-04-15

- 1 Introduction
 - 1.1 Information of Dataset
 - 1.2 Information of Variables
 - 1.3 About EDA Report
- 2 Univariate Analysis
 - 2.1 Descriptive Statistics
 - 2.2 Normality Test of Numerical Variables
 - 2.2.1 Statistics and Visualization of (Sample) Data
- 3 Relationship Between Variables
 - 3.1 Correlation Coefficient
 - 3.1.1 Correlation Coefficient by Variable Combination
 - 3.1.2 Correlation Plot of Numerical Variables
- 4 Target based Analysis
 - 4.1 Grouped Descriptive Statistics
 - 4.1.1 Grouped Numerical Variables
 - 4.1.2 Grouped Categorical Variables
 - 4.2 Grouped Relationship Between Variables
 - 4.2.1 Grouped Correlation Coefficient
 - 4.2.2 Grouped Correlation Plot of Numerical Variables

1 Introduction

The EDA Report provides exploratory data analysis information on objects that inherit data.frame and data.frame.

1.1 Information of Dataset

The dataset that generated the EDA Report is an '**data.frame**' object. It consists of **1,460 observations** and **62 variables**.

1.2 Information of Variables

The variable information of the data set that generated the EDA Report is shown in the following table.:

Information of Variables					
variables	types	missing_count	missing_percent	unique_count	unique_rate
MSZoning	factor	0	0	5	0.0034247
Street	factor	0	0	2	0.0013699
LotShape	factor	0	0	4	0.0027397
LandContour	factor	0	0	4	0.0027397
Utilities	factor	0	0	2	0.0013699
LotConfig	factor	0	0	5	0.0034247
LandSlope	factor	0	0	3	0.0020548
Neighborhood	factor	0	0	25	0.0171233
Condition1	factor	0	0	9	0.0061644
Condition2	factor	0	0	8	0.0054795
BldgType	factor	0	0	5	0.0034247
HouseStyle	factor	0	0	8	0.0054795
RoofStyle	factor	0	0	6	0.0041096
RoofMatl	factor	0	0	8	0.0054795
Exterior1st	factor	0	0	15	0.0102740
Exterior2nd	factor	0	0	16	0.0109589
ExterQual	factor	0	0	4	0.0027397
ExterCond	factor	0	0	5	0.0034247
Foundation	factor	0	0	6	0.0041096
Heating	factor	0	0	6	0.0041096
HeatingQC	factor	0	0	5	0.0034247
CentralAir	factor	0	0	2	0.0013699
KitchenQual	factor	0	0	4	0.0027397
Functional	factor	0	0	7	0.0047945
PavedDrive	factor	0	0	3	0.0020548
SaleType	factor	0	0	9	0.0061644
SaleCondition	factor	0	0	6	0.0041096
Id	integer	0	0	1460	1.0000000

MSSubClass	integer	0	0	15	0.0102740
LotArea	integer	0	0	1073	0.7349315
OverallQual	integer	0	0	10	0.0068493
OverallCond	integer	0	0	9	0.0061644
YearBuilt	integer	0	0	112	0.0767123
YearRemodAdd	integer	0	0	61	0.0417808
BsmtFinSF1	integer	0	0	637	0.4363014
BsmtFinSF2	integer	0	0	144	0.0986301
BsmtUnfSF	integer	0	0	780	0.5342466
TotalBsmtSF	integer	0	0	721	0.4938356
X1stFlrSF	integer	0	0	753	0.5157534
X2ndFlrSF	integer	0	0	417	0.2856164
LowQualFinSF	integer	0	0	24	0.0164384
GrLivArea	integer	0	0	861	0.5897260
BsmtFullBath	integer	0	0	4	0.0027397
BsmtHalfBath	integer	0	0	3	0.0020548
FullBath	integer	0	0	4	0.0027397
HalfBath	integer	0	0	3	0.0020548
BedroomAbvGr	integer	0	0	8	0.0054795
KitchenAbvGr	integer	0	0	4	0.0027397
TotRmsAbvGrd	integer	0	0	12	0.0082192
Fireplaces	integer	0	0	4	0.0027397
GarageCars	integer	0	0	5	0.0034247
GarageArea	integer	0	0	441	0.3020548
WoodDeckSF	integer	0	0	274	0.1876712
OpenPorchSF	integer	0	0	202	0.1383562
EnclosedPorch	integer	0	0	120	0.0821918
X3SsnPorch	integer	0	0	20	0.0136986
ScreenPorch	integer	0	0	76	0.0520548
PoolArea	integer	0	0	8	0.0054795
MiscVal	integer	0	0	21	0.0143836

MoSold	integer	0	0	12	0.0082192
YrSold	integer	0	0	5	0.0034247
SalePrice	integer	0	0	663	0.4541096

The target variable of the data is '**SalePrice**', and the data type of the variable is **integer**.

1.3 About EDA Report

EDA reports provide information and visualization results that support the EDA process. In particular, it provides a variety of information to understand the relationship between the target variable and the rest of the variables of interest.

2 Univariate Analysis

2.1 Descriptive Statistics

edaData

62 Variables 1460 Observations

MSZoning



n	missing	distinct
1460	0	5

Value	C (all)	FV	RH	RL	RM
Frequency	10	65	16	1151	218
Proportion	0.007	0.045	0.011	0.788	0.149

Street

n	missing	distinct
1460	0	2

Value	Grvl	Pave
Frequency	6	1454
Proportion	0.004	0.996

LotShape



n	missing	distinct

1460

0

4

Value	IR1	IR2	IR3	Reg
Frequency	484	41	10	925
Proportion	0.332	0.028	0.007	0.634

LandContour



n

missing

distinct

1460

0

4

Value	Bnk	HLS	Low	Lvl
Frequency	63	50	36	1311
Proportion	0.043	0.034	0.025	0.898

Utilities

n

missing

distinct

1460

0

2

Value	AllPub	NoSeWa
Frequency	1459	1
Proportion	0.999	0.001

LotConfig



n

missing

distinct

1460

0

5

Value	Corner	CulDSac	FR2	FR3	Inside
Frequency	263	94	47	4	1052
Proportion	0.180	0.064	0.032	0.003	0.721

LandSlope



n

missing

distinct

1460

0

3

Value	Gtl	Mod	Sev
Frequency	1382	65	13
Proportion	0.947	0.045	0.009

Neighborhood

n	missing	distinct
1460	0	25

lowest : Blmngtn Blueste BrDale BrkSide ClearCr , highest: Somerst StoneBr SWISU Timber Veenker

Condition1

n	missing	distinct
1460	0	9

Value	Artery	Feedr	Norm	PosA	PosN	RRAe	RRAn	RRNe	RRNn
Frequency	48	81	1260	8	19	11	26	2	5
Proportion	0.033	0.055	0.863	0.005	0.013	0.008	0.018	0.001	0.003

Condition2

n	missing	distinct
1460	0	8

Value	Artery	Feedr	Norm	PosA	PosN	RRAe	RRAn	RRNn
Frequency	2	6	1445	1	2	1	1	2
Proportion	0.001	0.004	0.990	0.001	0.001	0.001	0.001	0.001

BldgType

n	missing	distinct
1460	0	5

Value	1Fam	2fmCon	Duplex	Twnhs	TwnhsE
Frequency	1220	31	52	43	114
Proportion	0.836	0.021	0.036	0.029	0.078

HouseStyle

n	missing	distinct
1460	0	8

Value	1.5Fin	1.5Unf	1Story	2.5Fin	2.5Unf	2Story	SFoyer	SLvl
Frequency	1000	100	100	100	100	100	100	100
Proportion	0.689	0.068	0.068	0.068	0.068	0.068	0.068	0.068

Frequency	154	14	726	8	11	445	37	65
Proportion	0.105	0.010	0.497	0.005	0.008	0.305	0.025	0.045

RoofStyle

n	missing	distinct
1460	0	6

Value	Flat	Gable	Gambrel	Hip	Mansard	Shed
Frequency	13	1141	11	286	7	2
Proportion	0.009	0.782	0.008	0.196	0.005	0.001

RoofMatl

n	missing	distinct
1460	0	8

Value	ClyTile	CompShg	Membran	Metal	Roll	Tar&Grv	WdShake	WdShngl
Frequency	1	1434	1	1	1	11	5	6
Proportion	0.001	0.982	0.001	0.001	0.001	0.008	0.003	0.004

Exterior1st

n	missing	distinct
1460	0	15

Value	AsbShng	AsphShn	BrkComm	BrkFace	CBlock	CemntBd	HdBoard	ImStucc	MetalSd
Frequency	20	1	2	50	1	61	222	1	220
Proportion	0.014	0.001	0.001	0.034	0.001	0.042	0.152	0.001	0.151

Value	Plywood	Stone	Stucco	VinylSd	Wd	Sdng	WdShing
Frequency	108	2	25	515	206	26	
Proportion	0.074	0.001	0.017	0.353	0.141	0.018	

Exterior2nd

n	missing	distinct
1460	0	16

Value	AsbShng	AsphShn	Brk	Cmn	BrkFace	CBlock	CemntBd	HdBoard	ImStucc	MetalSd
Frequency	20	3	7	25	1	60	207	10	214	
Proportion	0.014	0.002	0.005	0.017	0.001	0.041	0.142	0.007	0.147	

Value	Other	Plywood	Stone	Stucco	VinylSd	Wd	Sdng	Wd	Shng
Frequency	1	142	5	26	504	197	38		
Proportion	0.001	0.097	0.003	0.018	0.345	0.135	0.026		

ExterQual

n	missing	distinct
1460	0	4

Value	Ex	Fa	Gd	TA
Frequency	52	14	488	906
Proportion	0.036	0.010	0.334	0.621

ExterCond

n	missing	distinct
1460	0	5

Value	Ex	Fa	Gd	Po	TA
Frequency	3	28	146	1	1282
Proportion	0.002	0.019	0.100	0.001	0.878

Foundation

n	missing	distinct
1460	0	6

Value	BrkTil	CBlock	PConc	Slab	Stone	Wood
Frequency	146	634	647	24	6	3
Proportion	0.100	0.434	0.443	0.016	0.004	0.002

Heating

n	missing	distinct
1460	0	6

Value	Floor	GasA	GasW	Grav	OthW	Wall
Frequency	1	1428	18	7	2	4
Proportion	0.001	0.978	0.012	0.005	0.001	0.003

HeatingQC

n	missing	distinct
1460	0	5

Value	Ex	Fa	Gd	Po	TA
Frequency	741	49	241	1	428
Proportion	0.508	0.034	0.165	0.001	0.293

CentralAir

n	missing	distinct
1460	0	2

Value	N	Y
Frequency	95	1365
Proportion	0.065	0.935

KitchenQual

n	missing	distinct
1460	0	4

Value	Ex	Fa	Gd	TA
Frequency	100	39	586	735
Proportion	0.068	0.027	0.401	0.503

Functional

n	missing	distinct
1460	0	7

Value	Maj1	Maj2	Min1	Min2	Mod	Sev	Typ
Frequency	14	5	31	34	15	1	1360
Proportion	0.010	0.003	0.021	0.023	0.010	0.001	0.932

PavedDrive

n	missing	distinct
1460	0	3

Value	N	P	Y
Frequency	90	30	1340
Proportion	0.062	0.021	0.918

SaleType

n	missing	distinct
1460	0	9

Value	COD	Con	ConLD	ConLI	ConLw	CWD	New	0th	WD
Frequency	43	2	9	5	5	4	122	3	1267
Proportion	0.029	0.001	0.006	0.003	0.003	0.003	0.084	0.002	0.868

SaleCondition

n	missing	distinct
1460	0	6

Value	Abnorml	AdjLand	Alloca	Family	Normal	Partial
Frequency	101	4	12	20	1198	125
Proportion	0.069	0.003	0.008	0.014	0.821	0.086

Id

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25
1460	0	1460	1	730.5	487	73.95	146.90	365.75
.50	.75	.90	.95					
730.50	1095.25	1314.10	1387.05					

lowest : 1 2 3 4 5 , highest: 1456 1457 1458 1459 1460

MSSubClass

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	15	0.94	56.9	43.19	20	20	20	50	70	120	160

Value	20	30	40	45	50	60	70	75	80	85	90	120
Frequency	536	69	4	12	144	299	60	16	58	20	52	87
Proportion	0.367	0.047	0.003	0.008	0.099	0.205	0.041	0.011	0.040	0.014	0.036	0.060

Value	160	180	190
Frequency	63	10	30
Proportion	0.043	0.007	0.021

LotArea

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	1073	1	10517	5718	3312	5000	7554	9478	11602	14382	17401

lowest : 1300 1477 1491 1526 1533 , highest: 70761 115149 159000 164660 215245

OverallQual

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	10	0.951	6.099	1.522	4	5	5	6	7	8	8

Value	1	2	3	4	5	6	7	8	9	10
Frequency	2	3	20	116	397	374	319	168	43	18
Proportion	0.001	0.002	0.014	0.079	0.272	0.256	0.218	0.115	0.029	0.012

OverallCond

n	missing	distinct	Info	Mean	Gmd
1460	0	9	0.814	5.575	1.111

Value	1	2	3	4	5	6	7	8	9
Frequency	1	5	25	57	821	252	205	72	22
Proportion	0.001	0.003	0.017	0.039	0.562	0.173	0.140	0.049	0.015

YearBuilt

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	112	1	1971	33.88	1916	1925	1954	1973	2000	2006	2007

lowest : 1872 1875 1880 1882 1885 , highest: 2006 2007 2008 2009 2010

YearRemodAdd

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	61	0.997	1985	23.05	1950	1950	1967	1994	2004	2006	2007

lowest : 1950 1951 1952 1953 1954 , highest: 2006 2007 2008 2009 2010

BsmtFinSF1

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	637	0.967	443.6	484.5	0.0	0.0	0.0	383.5	712.2	1065.5	1274.0

lowest : 0 2 16 20 24 , highest: 1904 2096 2188 2260 5644

BsmtFinSF2

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	144	0.305	46.55	86.58	0.0	0.0	0.0	0.0	0.0	117.2	396.2

lowest : 0 28 32 35 40 , highest: 1080 1085 1120 1127 1474

BsmtUnfSF

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	780	0.999	567.2	486.6	0.0	74.9	223.0	477.5	808.0	1232.0	1468.0

lowest : 0 14 15 23 26 , highest: 2042 2046 2121 2153 2336

TotalBsmtSF

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	721	1	1057	459.5	519.3	636.9	795.8	991.5	1298.2	1602.2	1753.0

lowest : 0 105 190 264 270 , highest: 3094 3138 3200 3206 6110

X1stFlrSF

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	753	1	1163	416.4	673.0	756.9	882.0	1087.0	1391.2	1680.0	1831.2

lowest : 334 372 438 480 483 , highest: 2633 2898 3138 3228 4692

X2ndFlrSF

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	417	0.817	347	450.2	0.0	0.0	0.0	0.0	728.0	954.2	1141.0

lowest : 0 110 167 192 208 , highest: 1611 1796 1818 1872 2065

LowQualFinSF

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	24	0.052	5.845	11.55	0	0	0	0	0	0	0

lowest : 0 53 80 120 144 , highest: 513 514 515 528 572

GrLivArea

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	861	1	1515	563.1	848	912	1130	1464	1777	2158	2466

lowest : 334 438 480 520 605 , highest: 3627 4316 4476 4676 5642

BsmtFullBath

n	missing	distinct	Info	Mean	Gmd																				
1460	0	4	0.733	0.4253	0.5085																				
<table> <tr> <td>Value</td> <td>0</td> <td>1</td> <td>2</td> <td>3</td> <td></td> </tr> <tr> <td>Frequency</td> <td>856</td> <td>588</td> <td>15</td> <td>1</td> <td></td> </tr> <tr> <td>Proportion</td> <td>0.586</td> <td>0.403</td> <td>0.010</td> <td>0.001</td> <td></td> </tr> </table>						Value	0	1	2	3		Frequency	856	588	15	1		Proportion	0.586	0.403	0.010	0.001			
Value	0	1	2	3																					
Frequency	856	588	15	1																					
Proportion	0.586	0.403	0.010	0.001																					
BsmtHalfBath																									
n	missing	distinct	Info	Mean	Gmd																				
1460	0	3	0.159	0.05753	0.1088																				
<table> <tr> <td>Value</td> <td>0</td> <td>1</td> <td>2</td> <td></td> <td></td> </tr> <tr> <td>Frequency</td> <td>1378</td> <td>80</td> <td>2</td> <td></td> <td></td> </tr> <tr> <td>Proportion</td> <td>0.944</td> <td>0.055</td> <td>0.001</td> <td></td> <td></td> </tr> </table>						Value	0	1	2			Frequency	1378	80	2			Proportion	0.944	0.055	0.001				
Value	0	1	2																						
Frequency	1378	80	2																						
Proportion	0.944	0.055	0.001																						
FullBath																									
n	missing	distinct	Info	Mean	Gmd																				
1460	0	4	0.766	1.565	0.5521																				
<table> <tr> <td>Value</td> <td>0</td> <td>1</td> <td>2</td> <td>3</td> <td></td> </tr> <tr> <td>Frequency</td> <td>9</td> <td>650</td> <td>768</td> <td>33</td> <td></td> </tr> <tr> <td>Proportion</td> <td>0.006</td> <td>0.445</td> <td>0.526</td> <td>0.023</td> <td></td> </tr> </table>						Value	0	1	2	3		Frequency	9	650	768	33		Proportion	0.006	0.445	0.526	0.023			
Value	0	1	2	3																					
Frequency	9	650	768	33																					
Proportion	0.006	0.445	0.526	0.023																					
HalfBath																									
n	missing	distinct	Info	Mean	Gmd																				
1460	0	3	0.706	0.3829	0.4852																				
<table> <tr> <td>Value</td> <td>0</td> <td>1</td> <td>2</td> <td></td> <td></td> </tr> <tr> <td>Frequency</td> <td>913</td> <td>535</td> <td>12</td> <td></td> <td></td> </tr> <tr> <td>Proportion</td> <td>0.625</td> <td>0.366</td> <td>0.008</td> <td></td> <td></td> </tr> </table>						Value	0	1	2			Frequency	913	535	12			Proportion	0.625	0.366	0.008				
Value	0	1	2																						
Frequency	913	535	12																						
Proportion	0.625	0.366	0.008																						
BedroomAbvGr																									
n	missing	distinct	Info	Mean	Gmd																				
1460	0	8	0.815	2.866	0.818																				
<table> <tr> <td>Value</td> <td>0</td> <td>1</td> <td>2</td> <td>3</td> <td>4</td> <td>5</td> <td>6</td> <td>8</td> <td></td> </tr> <tr> <td>Frequency</td> <td>6</td> <td>50</td> <td>358</td> <td>804</td> <td>213</td> <td>21</td> <td>7</td> <td>1</td> <td></td> </tr> </table>						Value	0	1	2	3	4	5	6	8		Frequency	6	50	358	804	213	21	7	1	
Value	0	1	2	3	4	5	6	8																	
Frequency	6	50	358	804	213	21	7	1																	

Proportion	0.004	0.034	0.245	0.551	0.146	0.014	0.005	0.001
------------	-------	-------	-------	-------	-------	-------	-------	-------

KitchenAbvGr

n	missing	distinct	Info	Mean	Gmd
1460	0	4	0.133	1.047	0.09174

Value	0	1	2	3
Frequency	1	1392	65	2
Proportion	0.001	0.953	0.045	0.001

TotRmsAbvGrd

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	12	0.958	6.518	1.762	4	5	5	6	7	9	10

Value	2	3	4	5	6	7	8	9	10	11	12	14
Frequency	1	17	97	275	402	329	187	75	47	18	11	1
Proportion	0.001	0.012	0.066	0.188	0.275	0.225	0.128	0.051	0.032	0.012	0.008	0.001

Fireplaces

n	missing	distinct	Info	Mean	Gmd
1460	0	4	0.806	0.613	0.6566

Value	0	1	2	3
Frequency	690	650	115	5
Proportion	0.473	0.445	0.079	0.003

GarageCars

n	missing	distinct	Info	Mean	Gmd
1460	0	5	0.802	1.767	0.7609

Value	0	1	2	3	4
Frequency	81	369	824	181	5
Proportion	0.055	0.253	0.564	0.124	0.003

GarageArea

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	441	1	473	234.9	0.0	240.0	334.5	480.0	576.0	757.1	850.1

lowest : 0 160 164 180 186 , highest: 1220 1248 1356 1390 1418

WoodDeckSF

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	274	0.858	94.24	125	0	0	0	0	168	262	335

lowest : 0 12 24 26 28 , highest: 668 670 728 736 857

OpenPorchSF

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	202	0.909	46.66	62.43	0	0	0	25	68	130	175

lowest : 0 4 8 10 11 , highest: 406 418 502 523 547

EnclosedPorch

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	120	0.369	21.95	39.39	0.0	0.0	0.0	0.0	0.0	112.0	180.1

lowest : 0 19 20 24 30 , highest: 301 318 330 386 552

X3SsnPorch

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	20	0.049	3.41	6.739	0	0	0	0	0	0	0

Value	0	23	96	130	140	144	153	162	168	180	182	196
Frequency	1436	1	1	1	1	2	1	1	3	2	1	1
Proportion	0.984	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.002	0.001	0.001	0.001
Value	216	238	245	290	304	320	407	508				
Frequency	2	1	1	1	1	1	1	1				
Proportion	0.001	0.001	0.001	0.001	0.001	0.001	0.001	0.001				

ScreenPorch

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	76	0.22	15.06	28.27	0	0	0	0	0	0	160

lowest : 0 40 53 60 63 , highest: 385 396 410 440 480

PoolArea

n	missing	distinct	Info	Mean	Gmd
1460	0	8	0.014	2.759	5.497

Value	0	480	512	519	555	576	648	738
Frequency	1453	1	1	1	1	1	1	1
Proportion	0.995	0.001	0.001	0.001	0.001	0.001	0.001	0.001

MiscVal

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	21	0.103	43.49	85.67	0	0	0	0	0	0	0

Value	0	50	350	400	450	500	550	600	700	800	1150	1200
Frequency	1408	1	1	11	4	10	1	5	5	1	1	2
Proportion	0.964	0.001	0.001	0.008	0.003	0.007	0.001	0.003	0.003	0.001	0.001	0.001

Value	1300	1400	2000	2500	3500	8300	15500
Frequency	1	1	4	1	1	1	1
Proportion	0.001	0.001	0.003	0.001	0.001	0.001	0.001

MoSold

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	12	0.985	6.322	3.041	2	3	5	6	8	10	11

Value	1	2	3	4	5	6	7	8	9	10	11	12
Frequency	58	52	106	141	204	253	234	122	63	89	79	59
Proportion	0.040	0.036	0.073	0.097	0.140	0.173	0.160	0.084	0.043	0.061	0.054	0.040

YrSold

n	missing	distinct	Info	Mean	Gmd
1460	0	5	0.955	2008	1.498

Value	2006	2007	2008	2009	2010
Frequency	314	329	304	338	175
Proportion	0.215	0.225	0.208	0.232	0.120

SalePrice

n	missing	distinct	Info	Mean	Gmd	.05	.10	.25	.50	.75	.90	.95
1460	0	663	1	180921	81086	88000	106475	129975	163000	214000	278000	326100

lowest : 34900 35311 37900 39300 40000 , highest: 582933 611657 625000 745000 755000

2.2 Normality Test of Numerical Variables

2.2.1 Statistics and Visualization of (Sample) Data

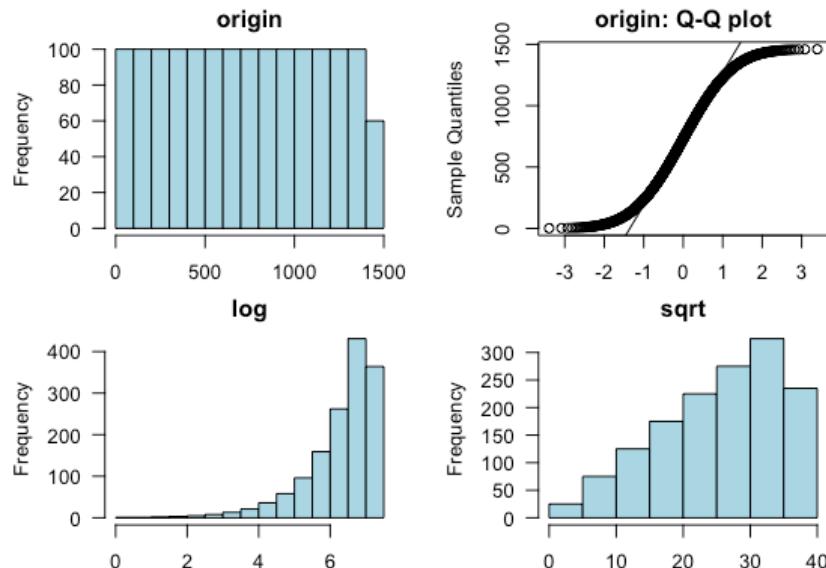
[Id]

normality test : Shapiro-Wilk normality test

```
statistic : 0.95484, p-value : 7.97201E-21
```

skewness and kurtosis

type	skewness	kurtosis
original	0.0000000	1.799999
log transformation	-1.8822380	7.762366
sqrt transformation	-0.5633059	2.393478



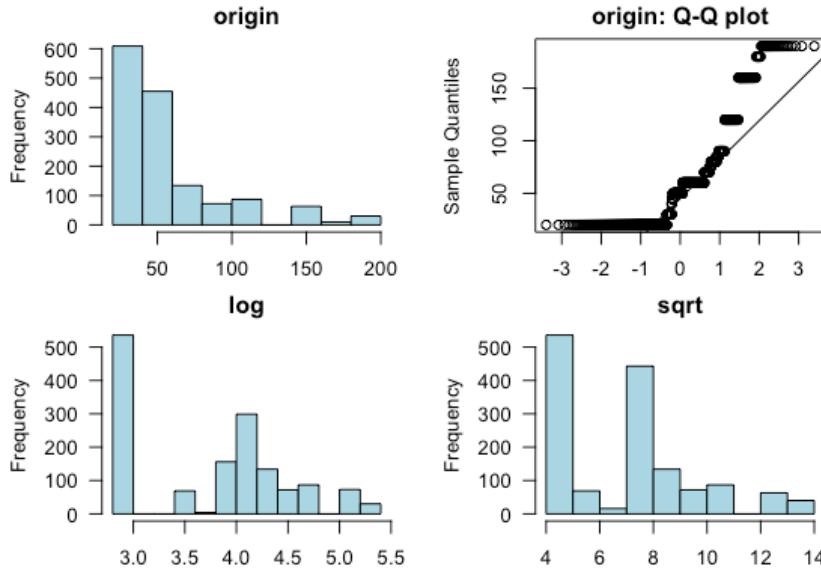
[MSSubClass]

normality test : Shapiro-Wilk normality test

```
statistic : 0.80457, p-value : 9.10819E-39
```

skewness and kurtosis

type	skewness	kurtosis
original	1.4062101	4.570675
log transformation	0.2322719	1.882137
sqrt transformation	0.7616640	2.864319



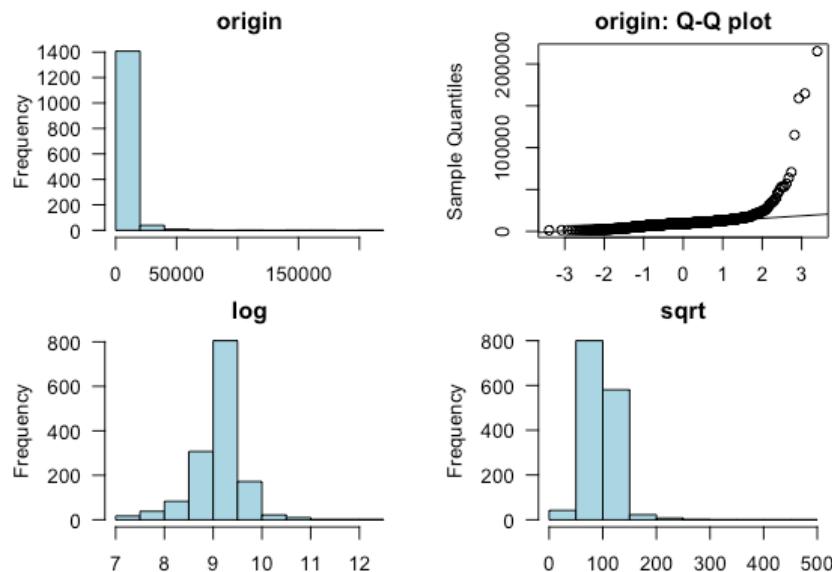
[LotArea]

normality test : Shapiro-Wilk normality test

```
statistic : 0.35106, p-value : 7.93365E-58
```

skewness and kurtosis

type	skewness	kurtosis
original	12.195142	205.543793
log transformation	-0.137852	7.693125
sqrt transformation	4.135459	43.383355



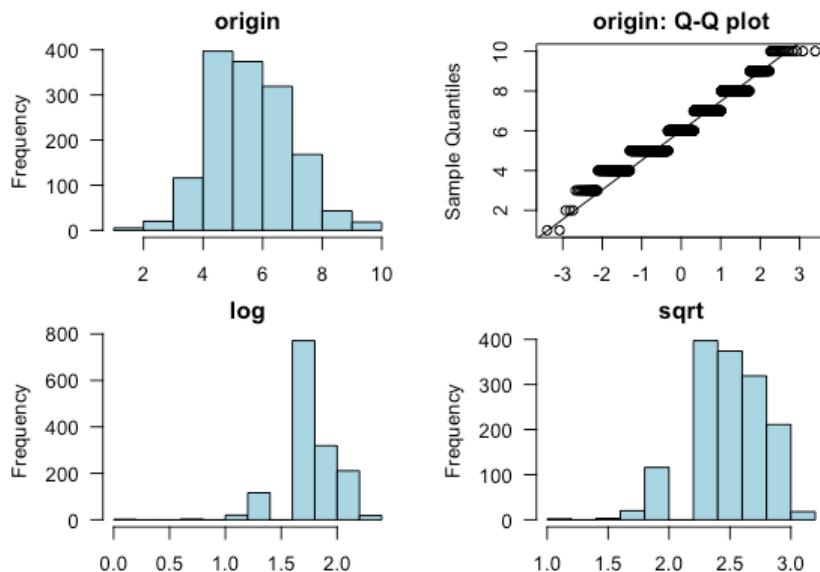
[OverallQual]

normality test : Shapiro-Wilk normality test

```
statistic : 0.94801, p-value : 2.68646E-22
```

skewness and kurtosis

type	skewness	kurtosis
original	0.2167210	3.091857
log transformation	-0.9288258	7.275053
sqrt transformation	-0.2180643	3.763105



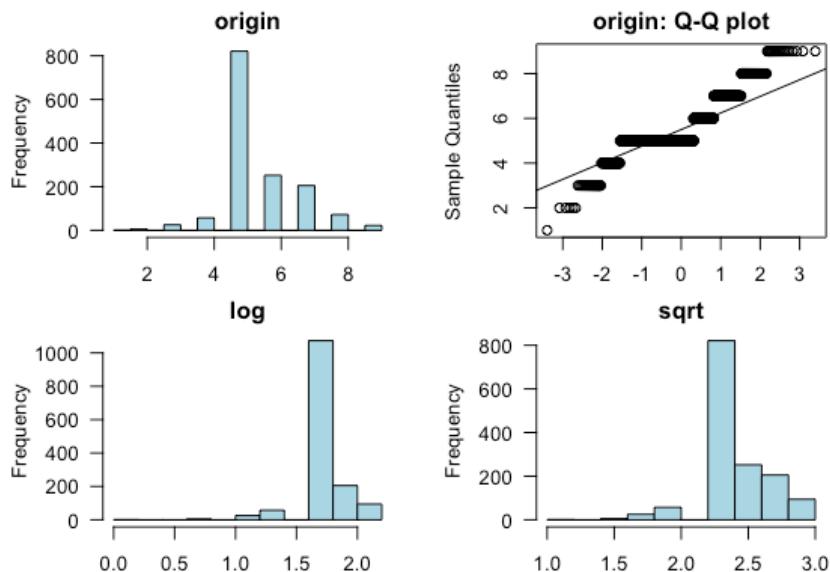
[OverallCond]

normality test : Shapiro-Wilk normality test

```
statistic : 0.82892, p-value : 6.77423E-37
```

skewness and kurtosis

type	skewness	kurtosis
original	0.6923552	4.098521
log transformation	-0.6360762	8.582312
sqrt transformation	0.1855279	4.861649



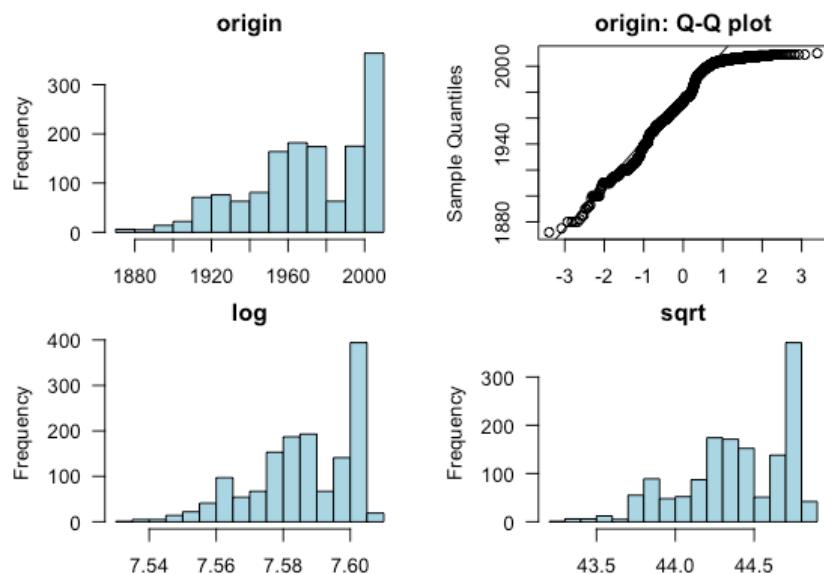
[YearBuilt]

normality test : Shapiro-Wilk normality test

```
statistic : 0.9256, p-value : 2.77022E-26
```

skewness and kurtosis

type	skewness	kurtosis
original	-0.6128307	2.557845
log transformation	-0.6404848	2.619312
sqrt transformation	-0.6265927	2.588173



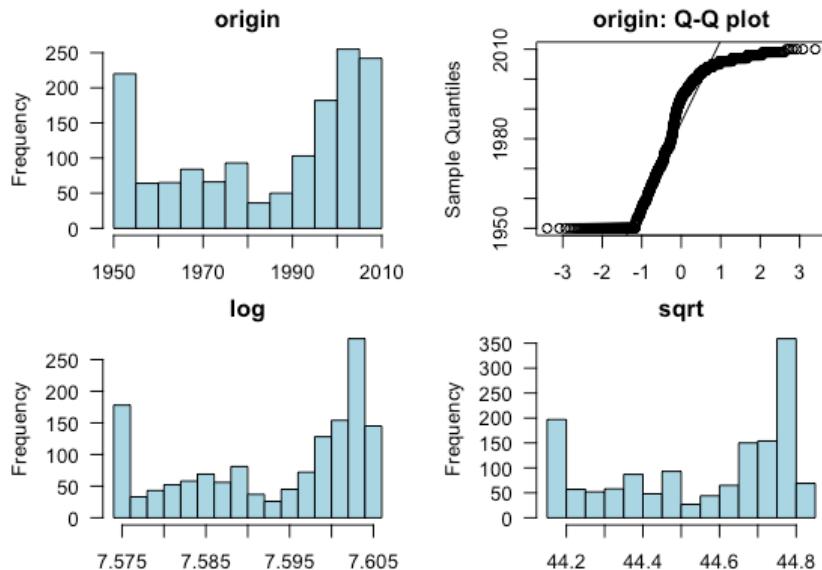
[YearRemodAdd]

normality test : Shapiro-Wilk normality test

```
statistic : 0.8628, p-value : 6.72028E-34
```

skewness and kurtosis

type	skewness	kurtosis
original	-0.5030445	1.728001
log transformation	-0.5104839	1.736856
sqrt transformation	-0.5067620	1.732411



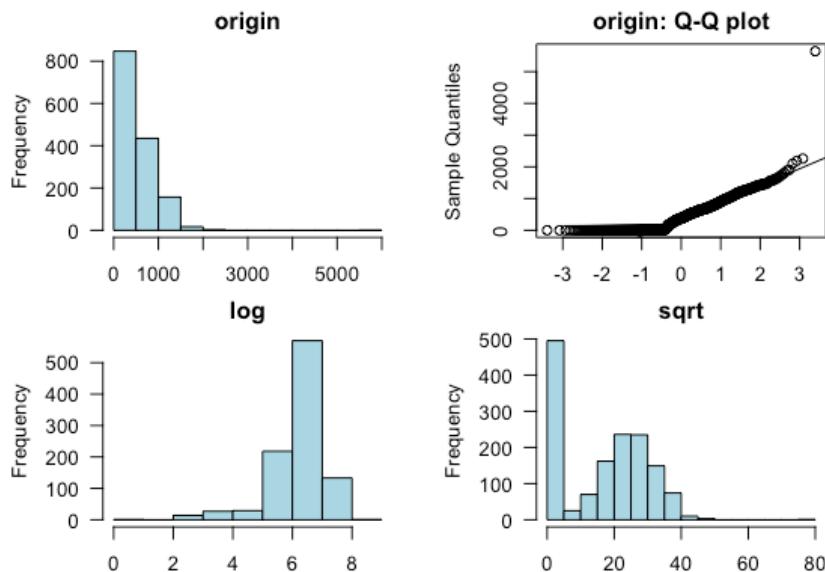
[BsmtFinSF1]

normality test : Shapiro-Wilk normality test

```
statistic : 0.84796, p-value : 2.81385E-35
```

skewness and kurtosis

type	skewness	kurtosis
original	1.6837709	14.076090
log transformation	NaN	NaN
sqrt transformation	-0.0299153	1.840678



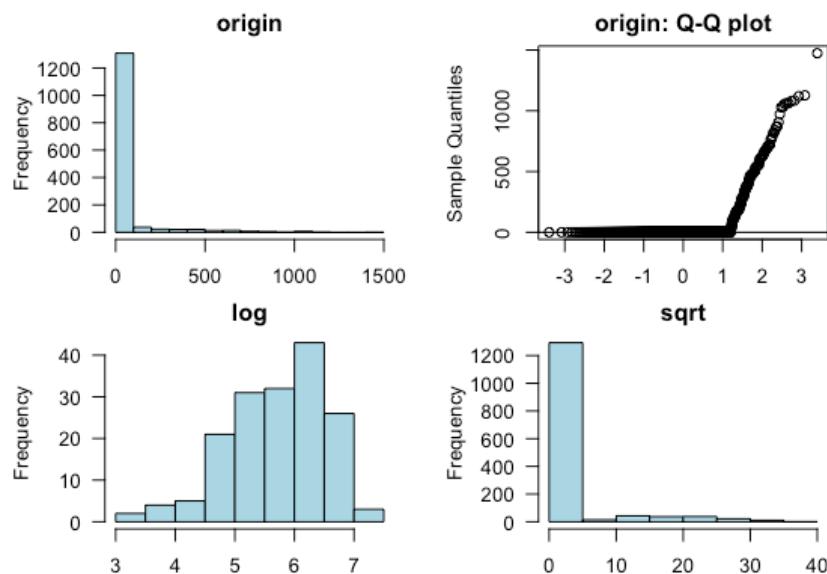
[BsmtFinSF2]

normality test : Shapiro-Wilk normality test

```
statistic : 0.32728, p-value : 1.85025E-58
```

skewness and kurtosis

type	skewness	kurtosis
original	4.250888	23.04042
log transformation	NaN	NaN
sqrt transformation	3.033689	11.33069



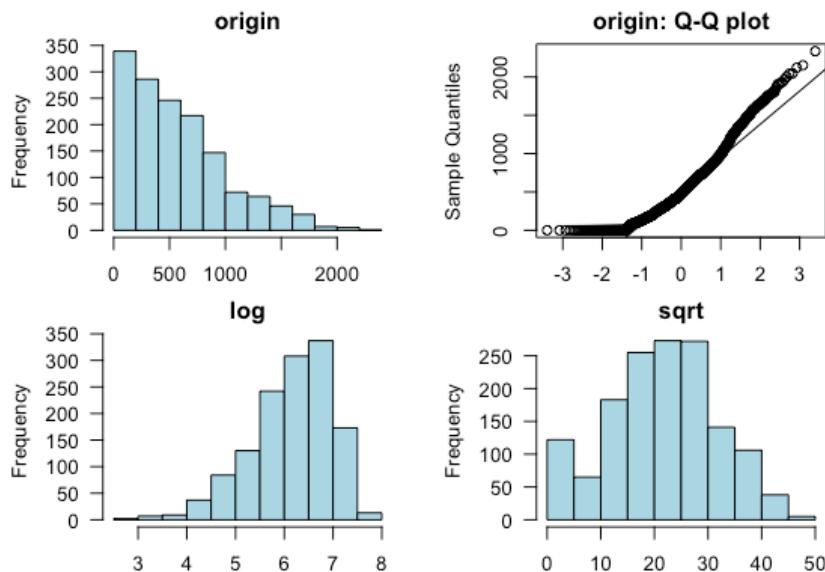
[BsmtUnfSF]

normality test : Shapiro-Wilk normality test

```
statistic : 0.93042, p-value : 1.63991E-25
```

skewness and kurtosis

type	skewness	kurtosis
original	0.9193227	3.469262
log transformation	NaN	NaN
sqrt transformation	-0.2462528	2.722352



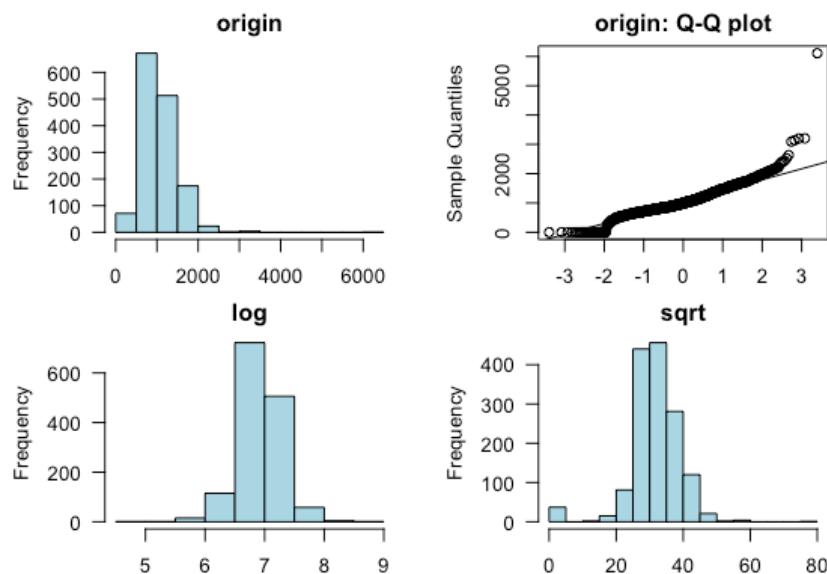
[TotalBsmtSF]

normality test : Shapiro-Wilk normality test

```
statistic : 0.91735, p-value : 1.61133E-27
```

skewness and kurtosis

type	skewness	kurtosis
original	1.522688	16.201042
log transformation	NaN	NaN
sqrt transformation	-1.273228	9.296421



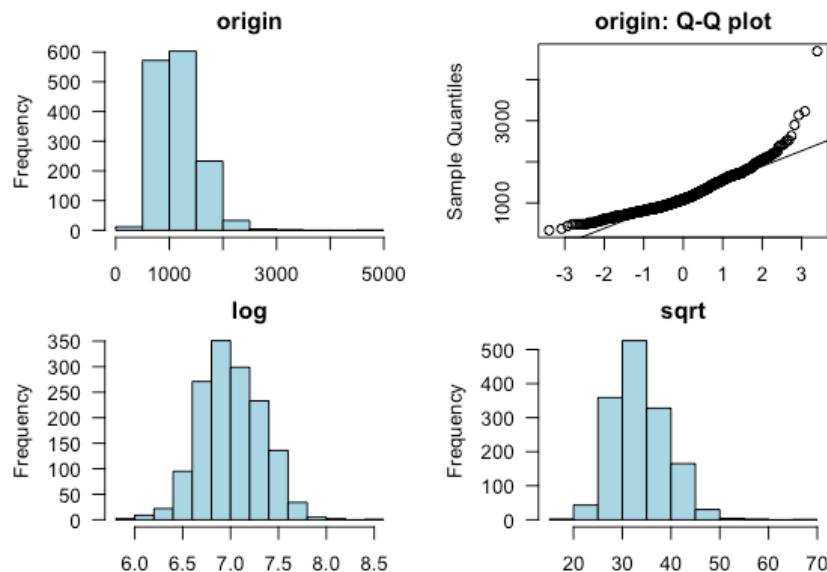
[X1stFlrSF]

normality test : Shapiro-Wilk normality test

```
statistic : 0.92695, p-value : 4.51322E-26
```

skewness and kurtosis

type	skewness	kurtosis
original	1.3753417	8.722076
log transformation	0.0790756	3.146185
sqrt transformation	0.6236212	4.197478



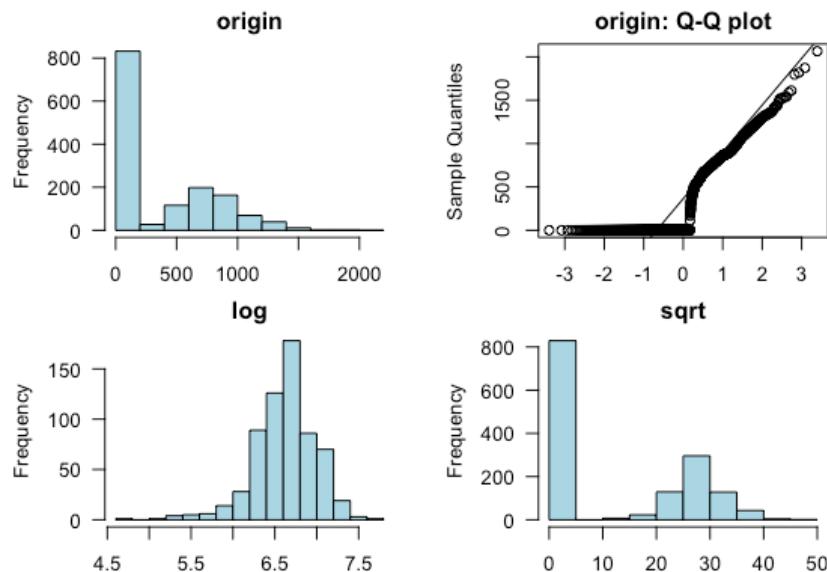
[X2ndFlrSF]

normality test : Shapiro-Wilk normality test

```
statistic : 0.7668, p-value : 2.51488E-41
```

skewness and kurtosis

type	skewness	kurtosis
original	0.8121943	2.444323
log transformation	NaN	NaN
sqrt transformation	0.4218984	1.368254



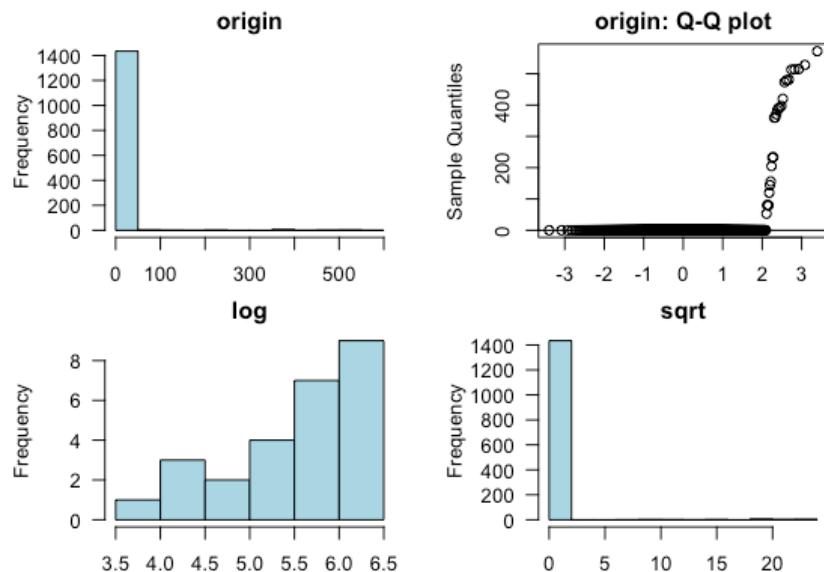
[LowQualFinSF]

normality test : Shapiro-Wilk normality test

```
statistic : 0.09799, p-value : 9.58925E-64
```

skewness and kurtosis

type	skewness	kurtosis
original	9.002080	85.94593
log transformation	NaN	NaN
sqrt transformation	8.057104	68.26267



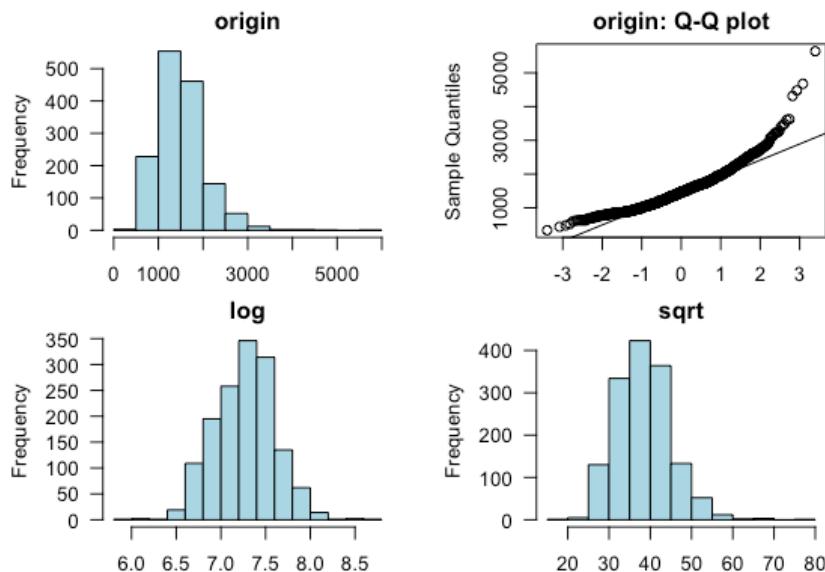
[GrLivArea]

normality test : Shapiro-Wilk normality test

```
statistic : 0.92798, p-value : 6.59761E-26
```

skewness and kurtosis

type	skewness	kurtosis
original	1.3651560	7.874266
log transformation	-0.0069880	3.277529
sqrt transformation	0.5936439	4.215916



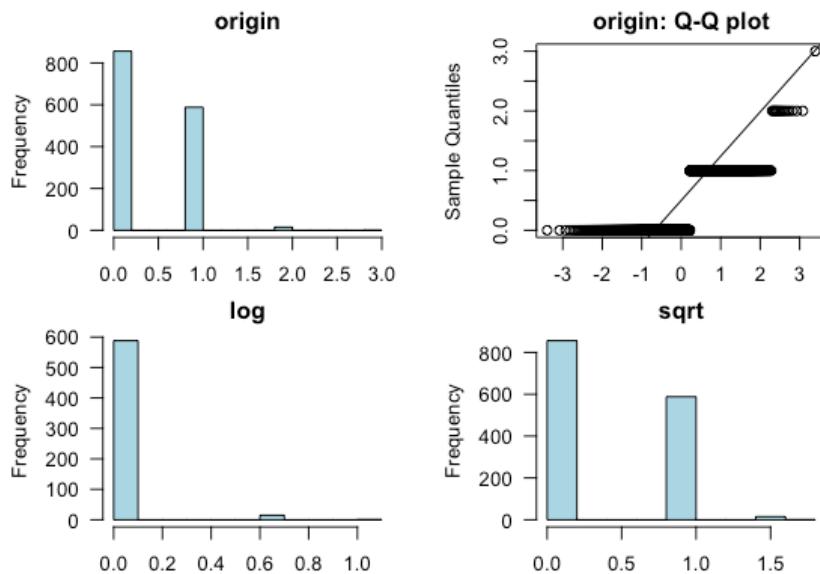
[BsmtFullBath]

normality test : Shapiro-Wilk normality test

```
statistic : 0.6583, p-value : 3.76067E-47
```

skewness and kurtosis

type	skewness	kurtosis
original	0.5954540	2.159666
log transformation	NaN	NaN
sqrt transformation	0.3830105	1.216578



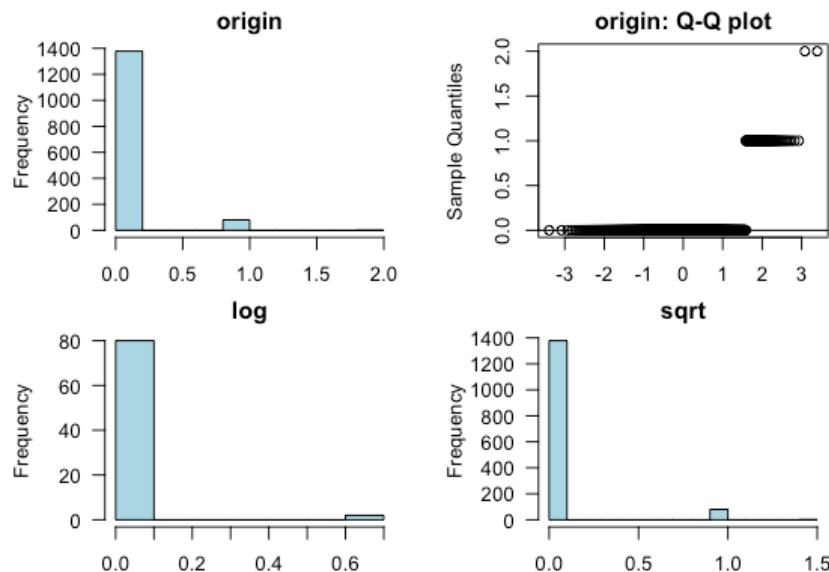
[BsmtHalfBath]

normality test : Shapiro-Wilk normality test

```
statistic : 0.24291, p-value : 1.46662E-60
```

skewness and kurtosis

type	skewness	kurtosis
original	4.099186	19.33644
log transformation	NaN	NaN
sqrt transformation	3.889992	16.28375



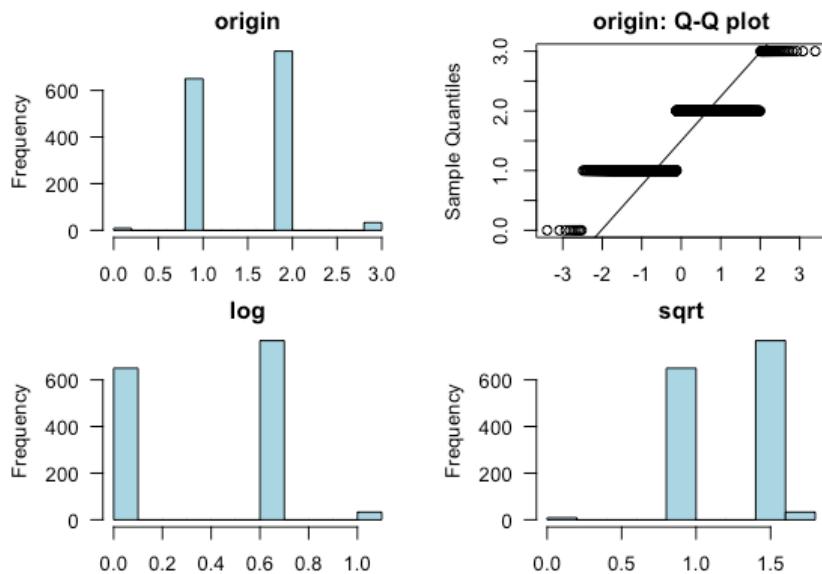
[FullBath]

normality test : Shapiro-Wilk normality test

```
statistic : 0.71936, p-value : 4.23149E-44
```

skewness and kurtosis

type	skewness	kurtosis
original	0.036524	2.141783
log transformation	NaN	NaN
sqrt transformation	-0.779461	5.441798



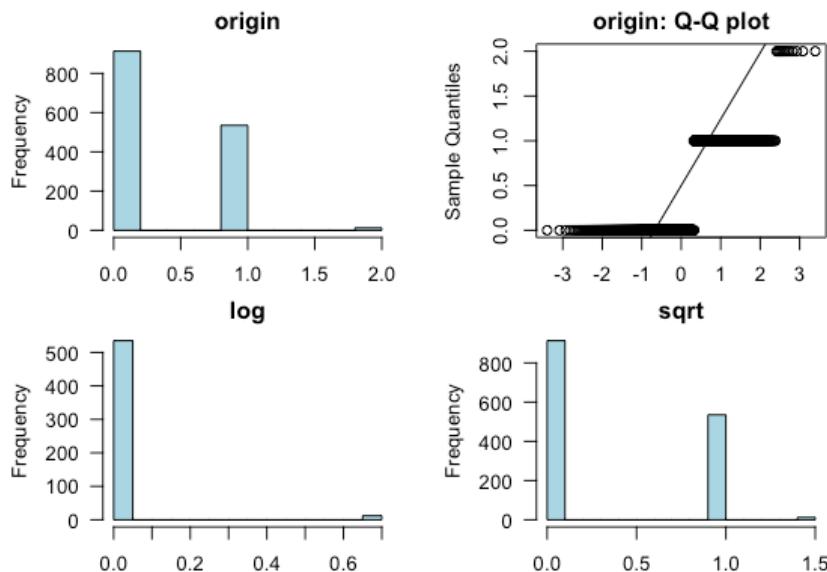
[HalfBath]

normality test : Shapiro-Wilk normality test

```
statistic : 0.638, p-value : 4.58158E-48
```

skewness and kurtosis

type	skewness	kurtosis
original	0.6752028	1.922651
log transformation	NaN	NaN
sqrt transformation	0.5403104	1.338718



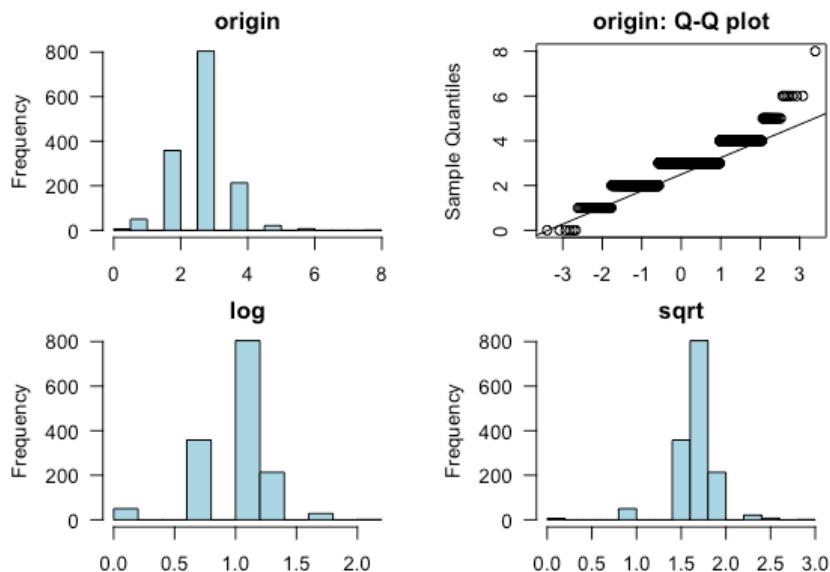
[BedroomAbvGr]

normality test : Shapiro-Wilk normality test

```
statistic : 0.8498, p-value : 4.11555E-35
```

skewness and kurtosis

type	skewness	kurtosis
original	0.2115724	5.219135
log transformation	NaN	NaN
sqrt transformation	-1.2457383	9.646511



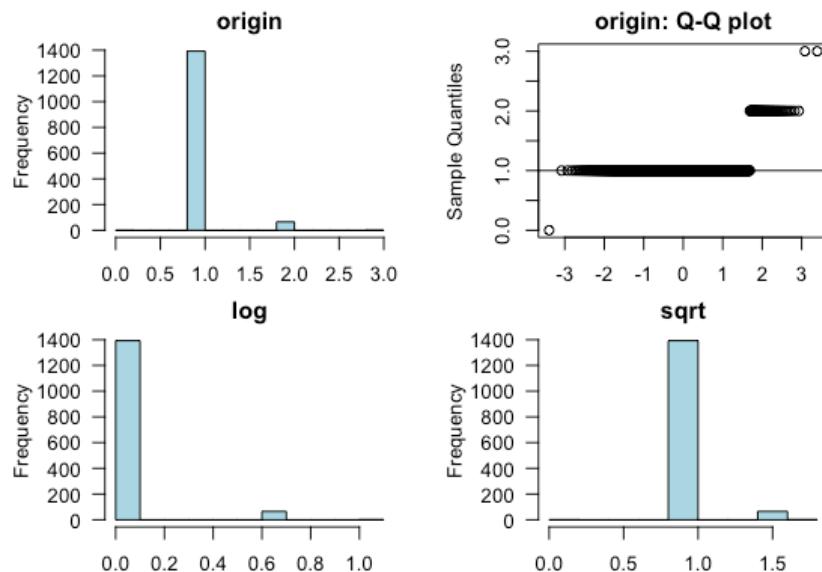
[KitchenAbvGr]

normality test : Shapiro-Wilk normality test

```
statistic : 0.2198, p-value : 4.2212E-61
```

skewness and kurtosis

type	skewness	kurtosis
original	4.483784	24.45463
log transformation	NaN	NaN
sqrt transformation	3.102307	28.78156



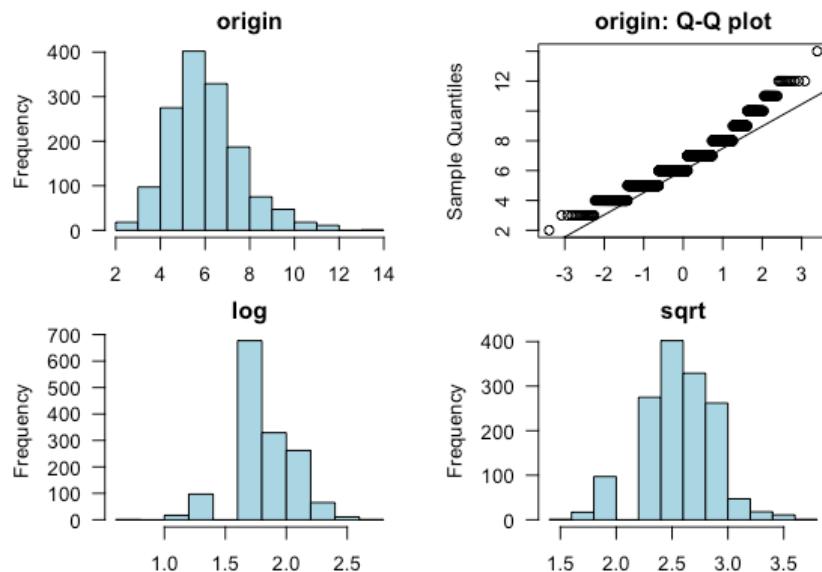
[TotRmsAbvGrd]

normality test : Shapiro-Wilk normality test

```
statistic : 0.94228, p-value : 2.00496E-23
```

skewness and kurtosis

type	skewness	kurtosis
original	0.6756458	3.873641
log transformation	-0.1905500	3.506792
sqrt transformation	0.2523529	3.375124



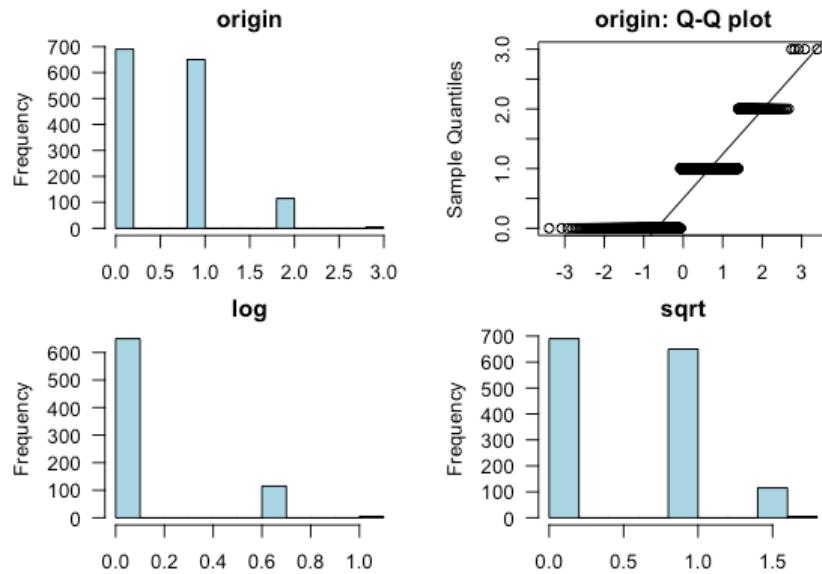
[Fireplaces]

normality test : Shapiro-Wilk normality test

```
statistic : 0.75523, p-value : 4.83098E-42
```

skewness and kurtosis

type	skewness	kurtosis
original	0.6488976	2.779399
log transformation	NaN	NaN
sqrt transformation	0.0451109	1.267465



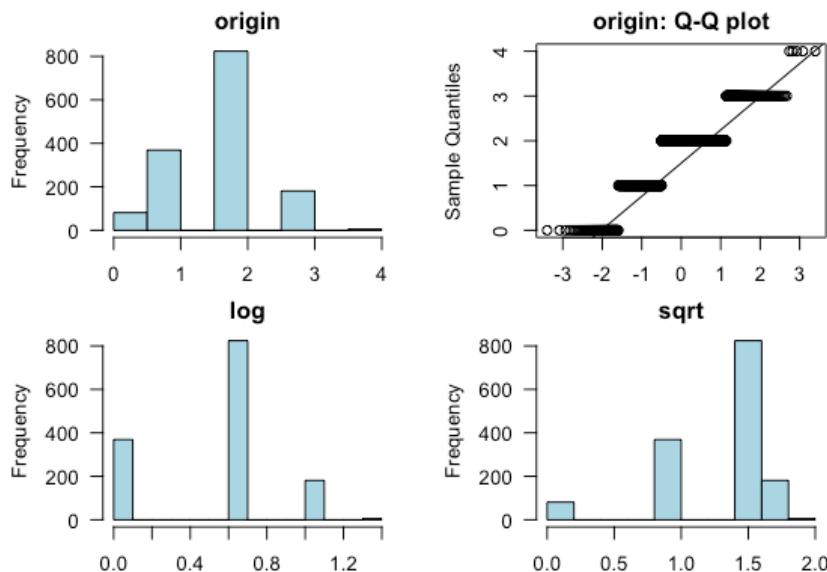
[GarageCars]

normality test : Shapiro-Wilk normality test

```
statistic : 0.83537, p-value : 2.30169E-36
```

skewness and kurtosis

type	skewness	kurtosis
original	-0.3421969	3.216135
log transformation	NaN	NaN
sqrt transformation	-1.8359830	7.017374



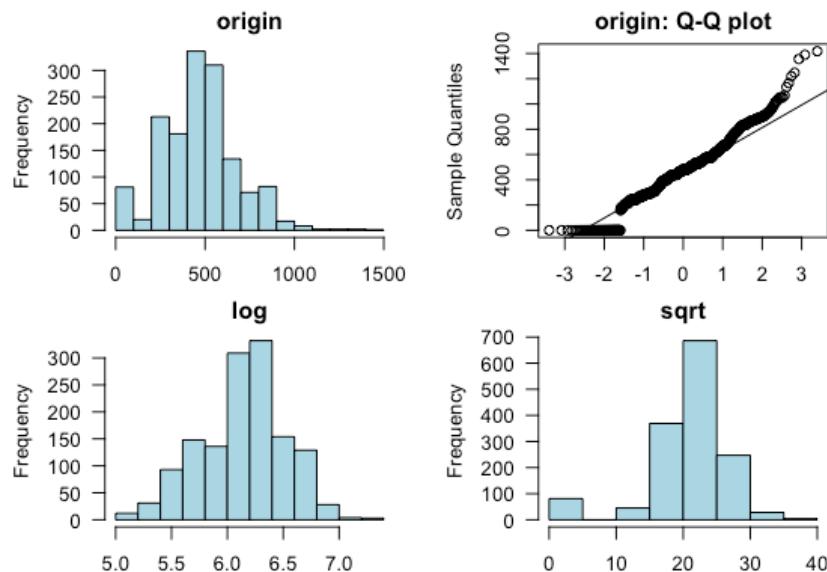
[GarageArea]

normality test : Shapiro-Wilk normality test

```
statistic : 0.97533, p-value : 4.01696E-15
```

skewness and kurtosis

type	skewness	kurtosis
original	0.1797959	3.909823
log transformation	NaN	NaN
sqrt transformation	-1.5983651	6.710425



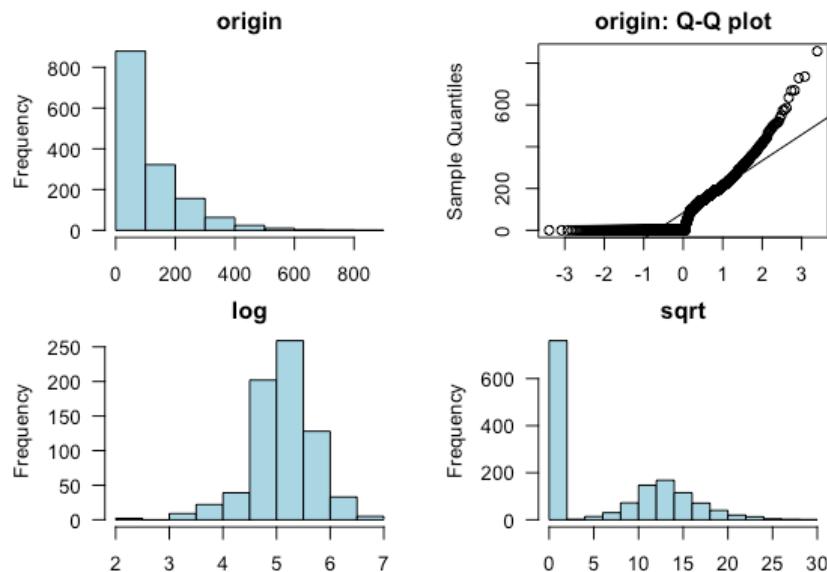
[WoodDeckSF]

normality test : Shapiro-Wilk normality test

```
statistic : 0.76852, p-value : 3.22799E-41
```

skewness and kurtosis

type	skewness	kurtosis
original	1.5397917	5.978604
log transformation	NaN	NaN
sqrt transformation	0.4860737	1.786097



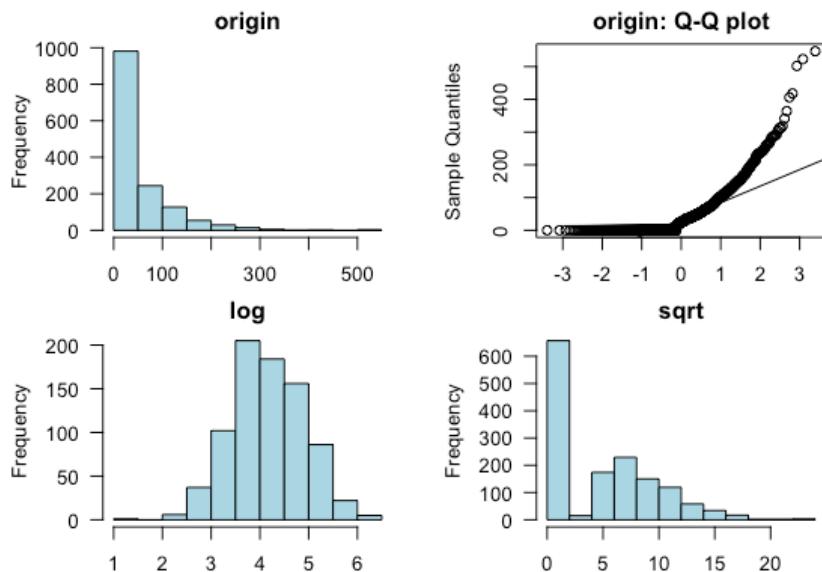
[OpenPorchSF]

normality test : Shapiro-Wilk normality test

```
statistic : 0.72717, p-value : 1.1359E-43
```

skewness and kurtosis

type	skewness	kurtosis
original	2.3619119	11.45718
log transformation	NaN	NaN
sqrt transformation	0.6002667	2.47021



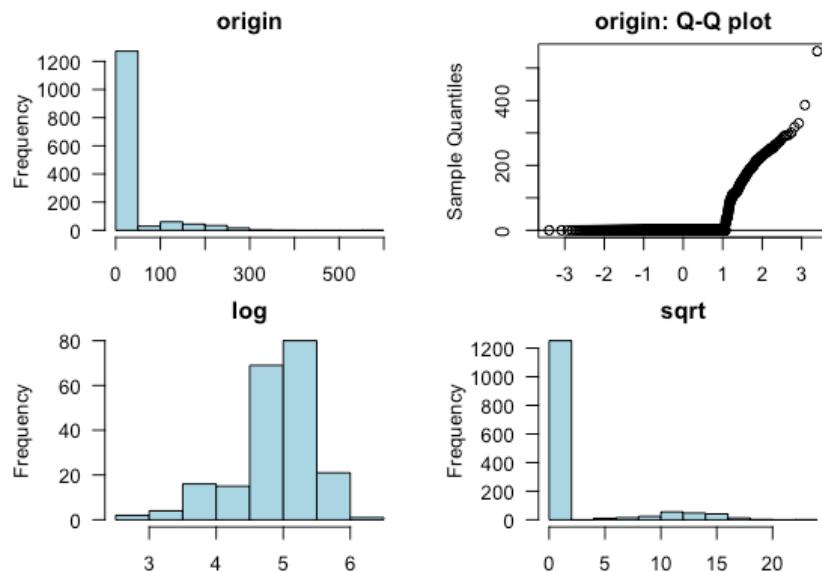
[EnclosedPorch]

normality test : Shapiro-Wilk normality test

```
statistic : 0.41444, p-value : 4.84948E-56
```

skewness and kurtosis

type	skewness	kurtosis
original	3.086696	13.390972
log transformation	NaN	NaN
sqrt transformation	2.345362	7.016066



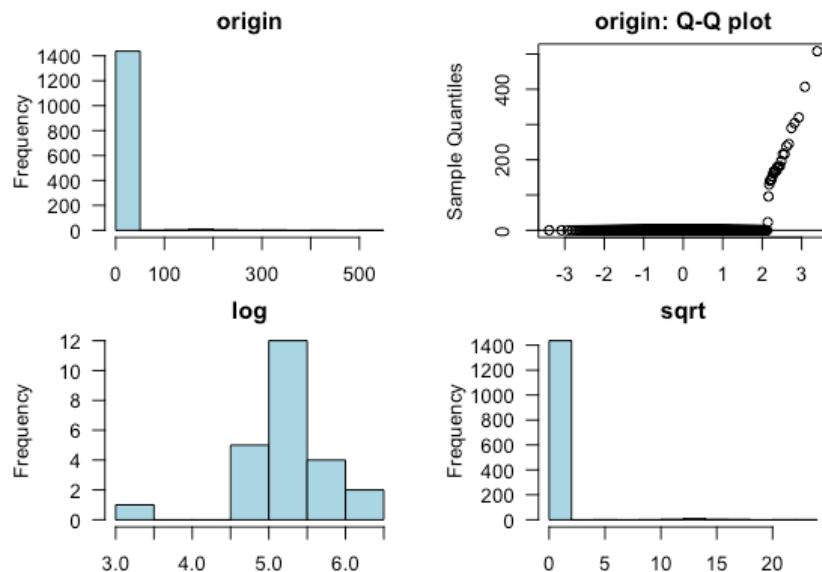
[X3SsnPorch]

normality test : Shapiro-Wilk normality test

```
statistic : 0.09493, p-value : 8.30727E-64
```

skewness and kurtosis

type	skewness	kurtosis
original	10.293752	126.23518
log transformation	NaN	NaN
sqrt transformation	8.268572	73.01529



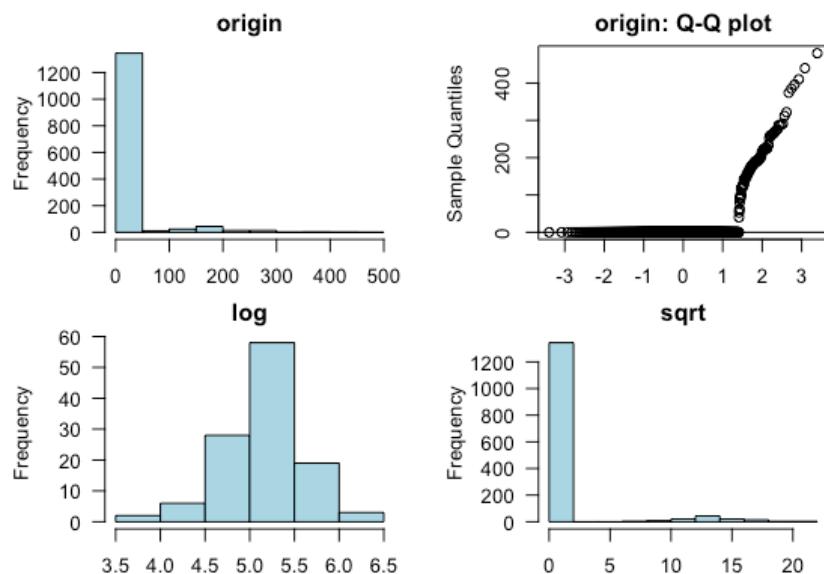
[ScreenPorch]

normality test : Shapiro-Wilk normality test

```
statistic : 0.29821, p-value : 3.30569E-59
```

skewness and kurtosis

type	skewness	kurtosis
original	4.117977	21.37187
log transformation	NaN	NaN
sqrt transformation	3.351332	12.86930



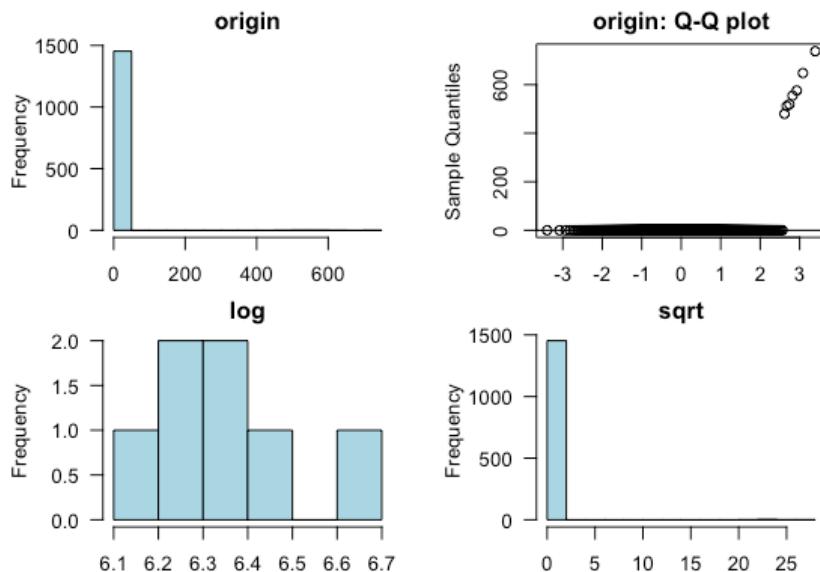
[PoolArea]

normality test : Shapiro-Wilk normality test

```
statistic : 0.0412, p-value : 7.11154E-65
```

skewness and kurtosis

type	skewness	kurtosis
original	14.81313	225.5005
log transformation	NaN	NaN
sqrt transformation	14.44994	210.9449



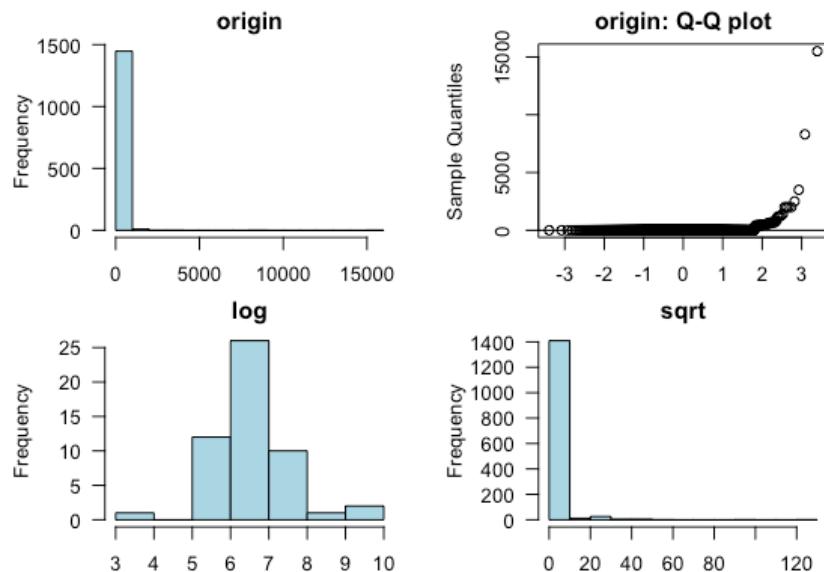
[MiscVal]

normality test : Shapiro-Wilk normality test

```
statistic : 0.05823, p-value : 1.52991E-64
```

skewness and kurtosis

type	skewness	kurtosis
original	24.451640	701.6008
log transformation	NaN	NaN
sqrt transformation	9.549306	131.6831



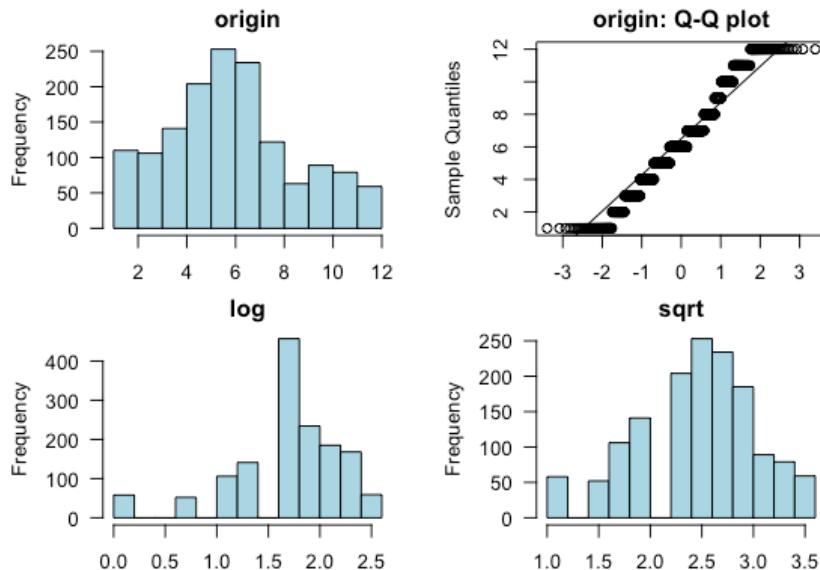
[MoSold]

normality test : Shapiro-Wilk normality test

```
statistic : 0.96878, p-value : 3.17897E-17
```

skewness and kurtosis

type	skewness	kurtosis
original	0.2118351	2.593167
log transformation	-1.3053445	5.069596
sqrt transformation	-0.4412232	3.081224



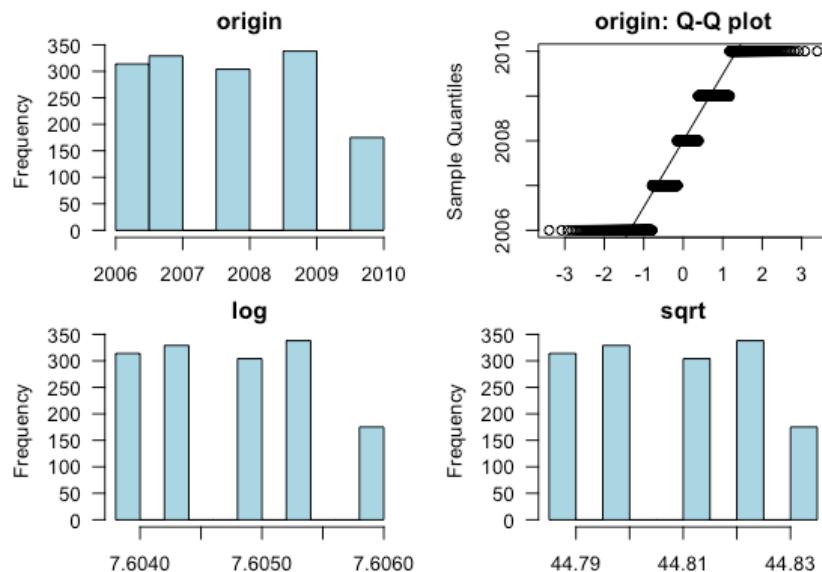
[YrSold]

normality test : Shapiro-Wilk normality test

```
statistic : 0.8971, p-value : 3.42019E-30
```

skewness and kurtosis

type	skewness	kurtosis
original	0.0961696	1.809366
log transformation	0.0953761	1.809028
sqrt transformation	0.0957728	1.809197



3 Relationship Between Variables

3.1 Correlation Coefficient

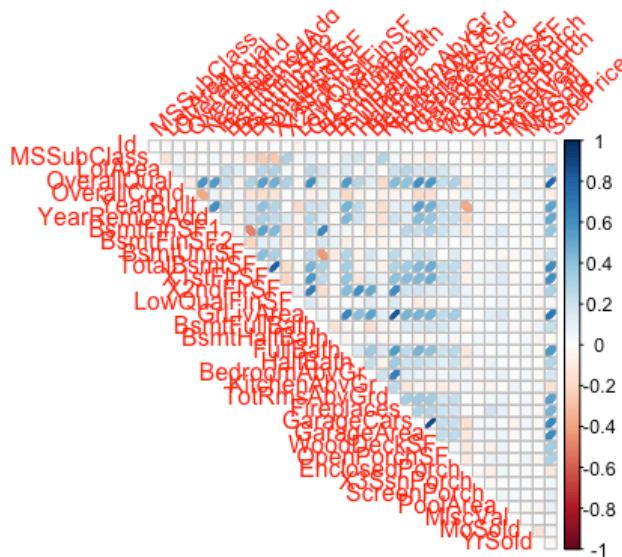
3.1.1 Correlation Coefficient by Variable Combination

Table of correlation coefficients (0.5 or more)

Variable1	Variable2	Correlation Coefficient
GarageArea	GarageCars	0.8824754
TotRmsAbvGrd	GrLivArea	0.8254894
X1stFlrSF	TotalBsmtSF	0.8195300
SalePrice	OverallQual	0.7909816
SalePrice	GrLivArea	0.7086245
GrLivArea	X2ndFlrSF	0.6875011
TotRmsAbvGrd	BedroomAbvGr	0.6766199
BsmtFullBath	BsmtFinSF1	0.6492118
SalePrice	GarageCars	0.6404092
FullBath	GrLivArea	0.6300116
SalePrice	GarageArea	0.6234314
TotRmsAbvGrd	X2ndFlrSF	0.6164226

SalePrice	TotalBsmtSF	0.6135806
HalfBath	X2ndFlrSF	0.6097073
SalePrice	X1stFlrSF	0.6058522
GarageCars	OverallQual	0.6006707
GrLivArea	OverallQual	0.5930074
YearRemodAdd	YearBuilt	0.5928550
YearBuilt	OverallQual	0.5723228
GrLivArea	X1stFlrSF	0.5660240
GarageArea	OverallQual	0.5620218
SalePrice	FullBath	0.5606638
TotRmsAbvGrd	FullBath	0.5547843
YearRemodAdd	OverallQual	0.5506839
FullBath	OverallQual	0.5505997
GarageCars	YearBuilt	0.5378501
TotalBsmtSF	OverallQual	0.5378085
SalePrice	TotRmsAbvGrd	0.5337232
SalePrice	YearBuilt	0.5228973
TotalBsmtSF	BsmtFinSF1	0.5223961
BedroomAbvGr	GrLivArea	0.5212695
SalePrice	YearRemodAdd	0.5071010
BedroomAbvGr	X2ndFlrSF	0.5029006

3.1.2 Correlation Plot of Numerical Variables



4 Target based Analysis

4.1 Grouped Descriptive Statistics

4.1.1 Grouped Numerical Variables

[Id]

Simple Linear Model Information

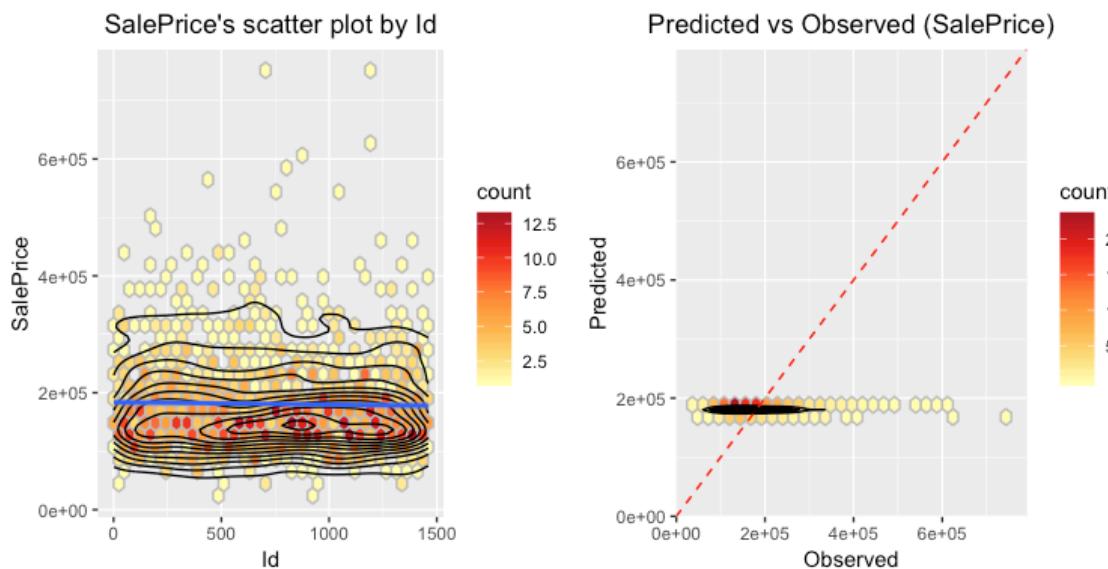
```
Residual standard error: 79451 on 1458 degrees of freedom
Multiple R-squared:  0.00048,  Adjusted R-squared: -0.00021
F-statistic: 1 on 1 and 1458 DF,  p-value: 0.402693799755773
```

Coefficients:

Simple Linear Model coefficients : Id

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	183937.94	4160.77	44.21	0.0
Id	-4.13	4.93	-0.84	0.4

Visualization:



[MSSubClass]

Simple Linear Model Information

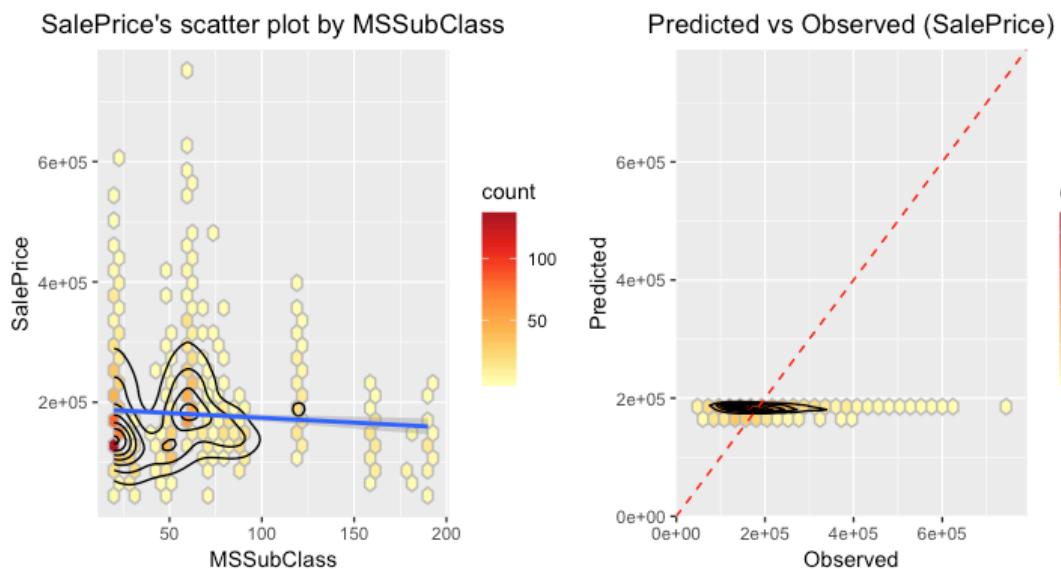
```
Residual standard error: 79187 on 1458 degrees of freedom
Multiple R-squared:  0.0071,   Adjusted R-squared:  0.00642
F-statistic: 10 on 1 and 1458 DF,  p-value: 0.00126647168474951
```

Coefficients:

Simple Linear Model coefficients : MSSubClass

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	189927.44	3474.29	54.67	0
MSSubClass	-158.29	49.01	-3.23	0

Visualization:



[LotArea]

Simple Linear Model Information

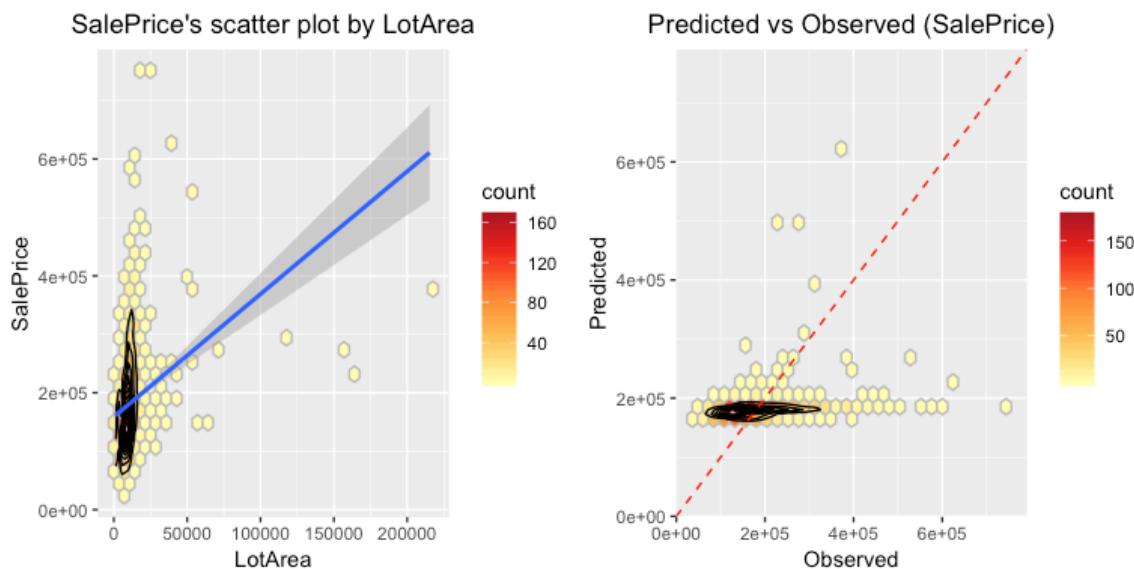
```
Residual standard error: 76654 on 1458 degrees of freedom
Multiple R-squared:  0.06961, Adjusted R-squared:  0.06898
F-statistic: 109 on 1 and 1458 DF,  p-value: 1.12313915491887e-24
```

Coefficients:

Simple Linear Model coefficients : LotArea

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	158836.1	2914.72	54.49	0
LotArea	2.1	0.20	10.44	0

Visualization:



[OverallQual]

Simple Linear Model Information

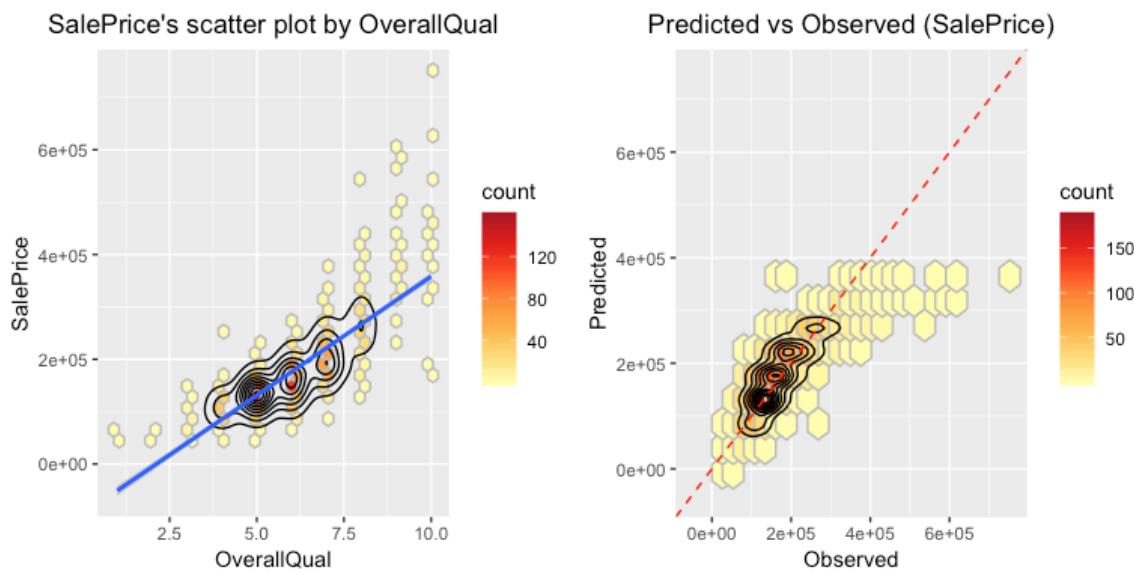
```
Residual standard error: 48623 on 1458 degrees of freedom
Multiple R-squared:  0.62565, Adjusted R-squared:  0.6254
F-statistic: 2437 on 1 and 1458 DF,  p-value: 2.18567526799377e-313
```

Coefficients:

Simple Linear Model coefficients : OverallQual

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-96206.08	5756.41	-16.71	0
OverallQual	45435.80	920.43	49.36	0

Visualization:



[OverallCond]

Simple Linear Model Information

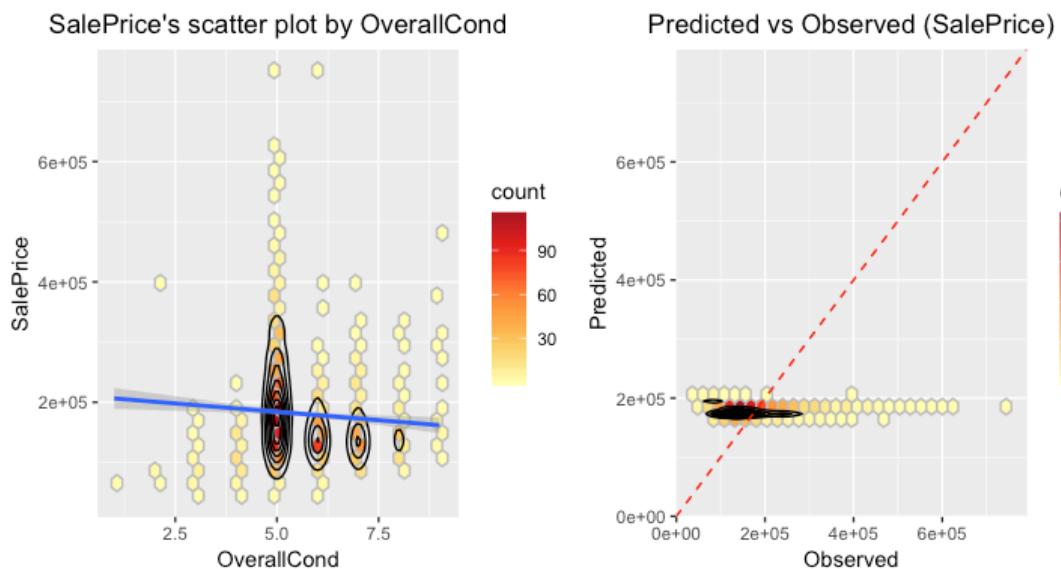
```
Residual standard error: 79229 on 1458 degrees of freedom
Multiple R-squared:  0.00606, Adjusted R-squared:  0.00538
F-statistic: 9 on 1 and 1458 DF,  p-value: 0.00291235126791591
```

Coefficients:

Simple Linear Model coefficients : OverallCond

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	211909.59	10597.07	20.00	0
OverallCond	-5558.12	1863.96	-2.98	0

Visualization:



[YearBuilt]

Simple Linear Model Information

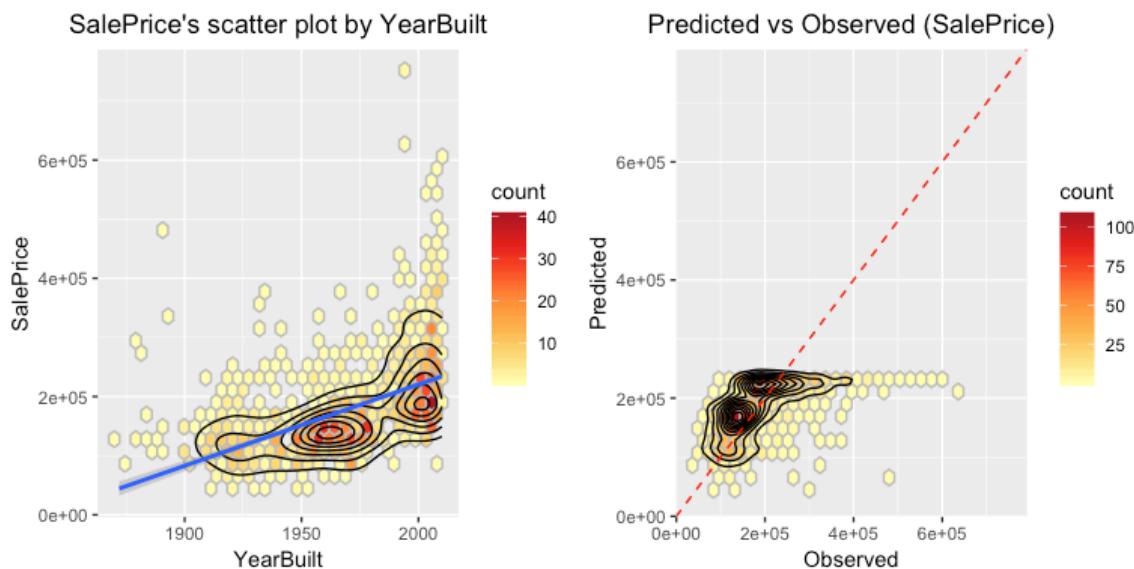
```
Residual standard error: 67740 on 1458 degrees of freedom
Multiple R-squared:  0.27342, Adjusted R-squared:  0.27292
F-statistic: 549 on 1 and 1458 DF,  p-value: 2.99022909901717e-103
```

Coefficients:

Simple Linear Model coefficients : YearBuilt

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-2530308.25	115761.26	-21.86	0
YearBuilt	1375.37	58.72	23.42	0

Visualization:



[YearRemodAdd]

Simple Linear Model Information

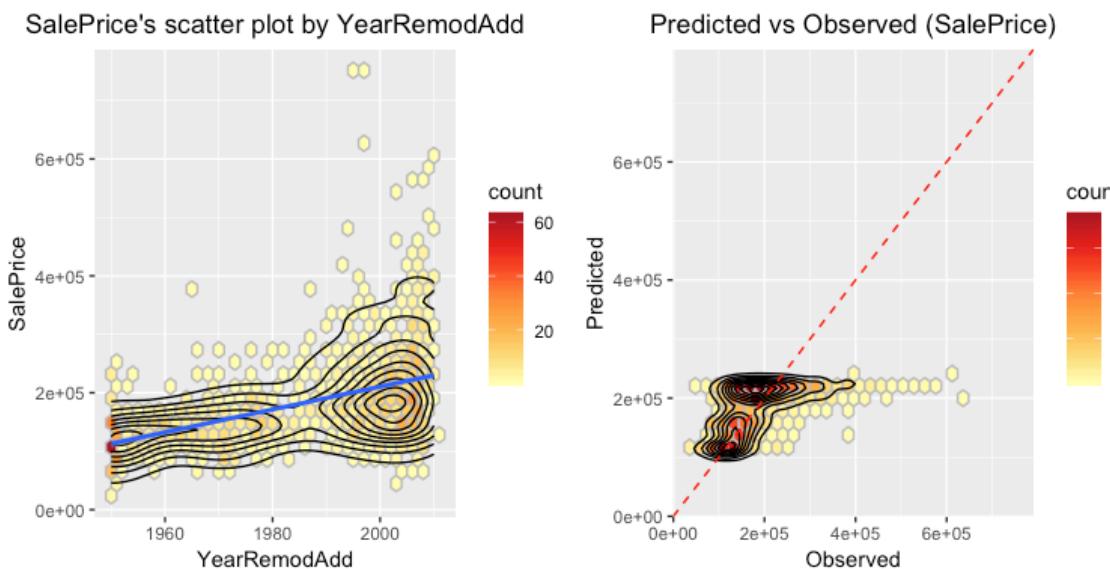
```
Residual standard error: 68494 on 1458 degrees of freedom
Multiple R-squared:  0.25715, Adjusted R-squared:  0.25664
F-statistic: 505 on 1 and 1458 DF,  p-value: 3.16494824191857e-96
```

Coefficients:

Simple Linear Model coefficients : YearRemodAdd

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-3692146.2	172407.23	-21.42	0
YearRemodAdd	1951.3	86.86	22.47	0

Visualization:



[BsmtFinSF1]

Simple Linear Model Information

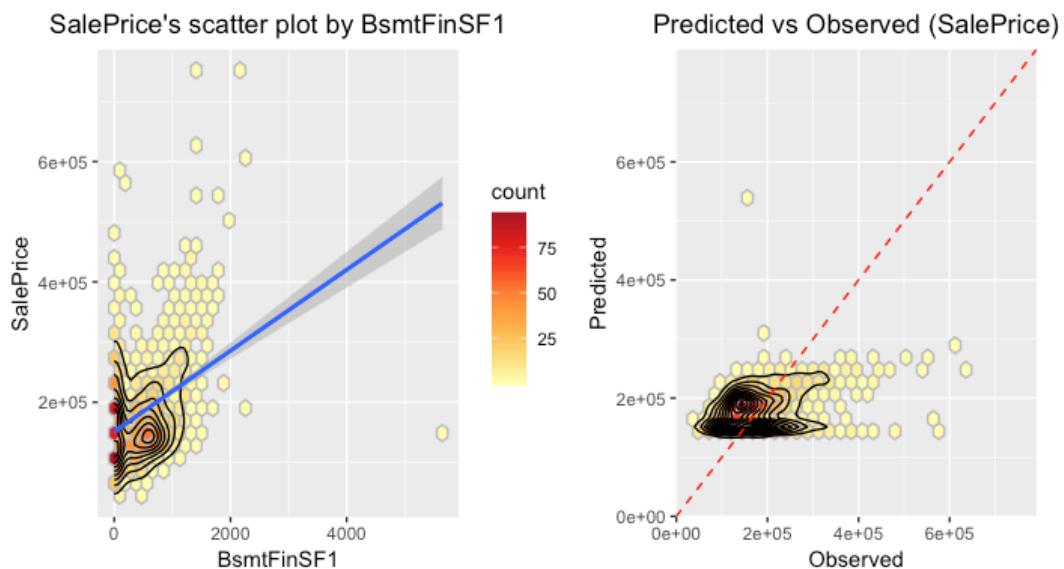
```
Residual standard error: 73297 on 1458 degrees of freedom
Multiple R-squared:  0.14932, Adjusted R-squared:  0.14874
F-statistic: 256 on 1 and 1458 DF,  p-value: 3.39410982928768e-53
```

Coefficients:

Simple Linear Model coefficients : BsmtFinSF1

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	151061.56	2676.49	56.44	0
BsmtFinSF1	67.31	4.21	16.00	0

Visualization:



[BsmtFinSF2]

Simple Linear Model Information

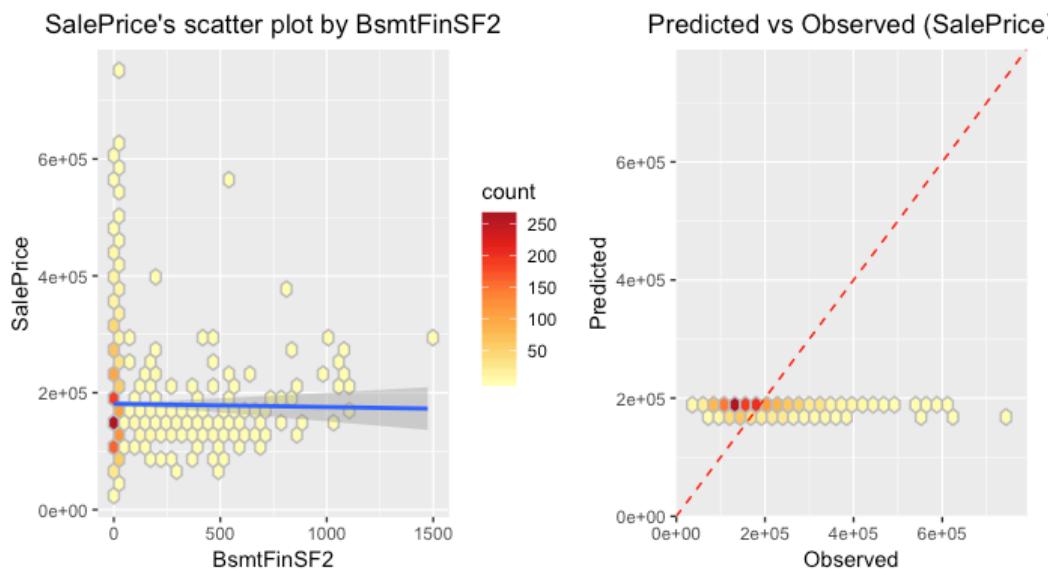
```
Residual standard error: 79465 on 1458 degrees of freedom
Multiple R-squared:  0.00013, Adjusted R-squared: -0.00056
F-statistic: 0 on 1 and 1458 DF,  p-value: 0.663998650834085
```

Coefficients:

Simple Linear Model coefficients : BsmtFinSF2

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	181182.0	2164.59	83.70	0.00
BsmtFinSF2	-5.6	12.90	-0.43	0.66

Visualization:



[BsmtUnfSF]

Simple Linear Model Information

```
Residual standard error: 77620 on 1458 degrees of freedom
Multiple R-squared:  0.046,   Adjusted R-squared:  0.04535
F-statistic: 70 on 1 and 1458 DF,  p-value: 1.18297579635946e-16
```

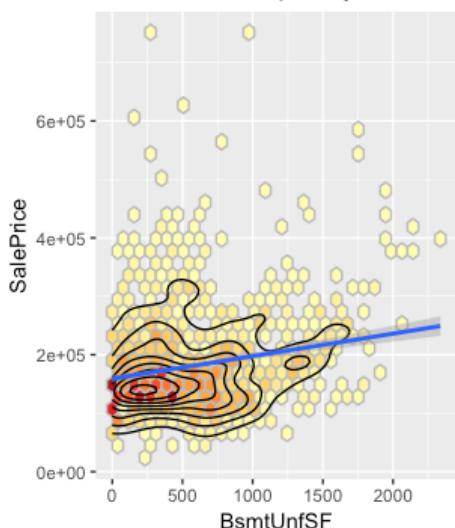
Coefficients:

Simple Linear Model coefficients : BsmtUnfSF

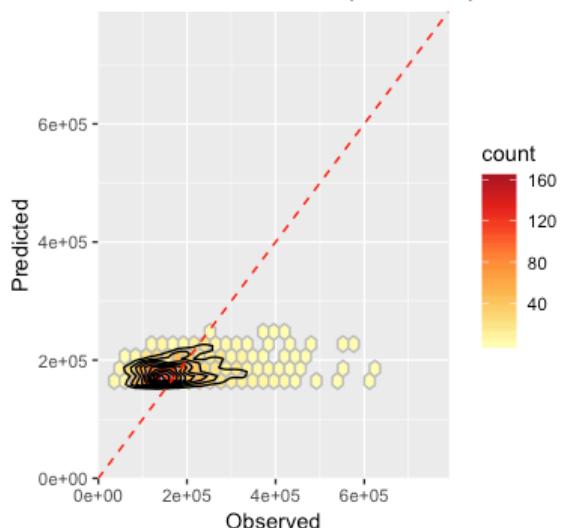
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	159047.93	3306.35	48.10	0
BsmtUnfSF	38.56	4.60	8.38	0

Visualization:

SalePrice's scatter plot by BsmtUnfSF



Predicted vs Observed (SalePrice)



[TotalBsmtSF]

Simple Linear Model Information

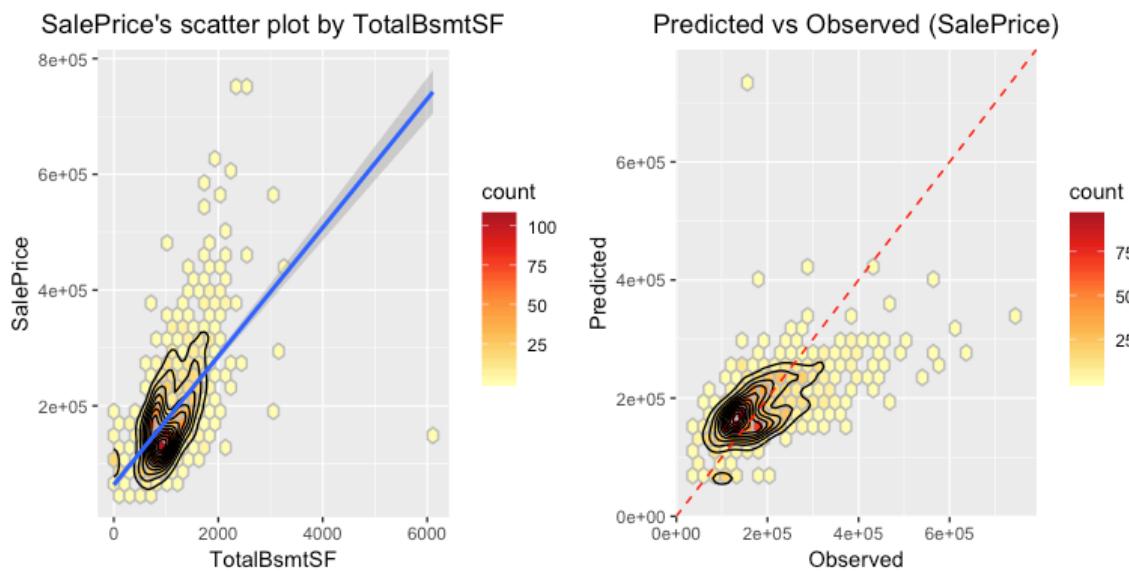
```
Residual standard error: 62752 on 1458 degrees of freedom
Multiple R-squared:  0.37648, Adjusted R-squared:  0.37605
F-statistic: 880 on 1 and 1458 DF,  p-value: 9.48422939149415e-152
```

Coefficients:

Simple Linear Model coefficients : TotalBsmtSF

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	63430.63	4286.89	14.80	0
TotalBsmtSF	111.11	3.74	29.67	0

Visualization:



[X1stFlrSF]

Simple Linear Model Information

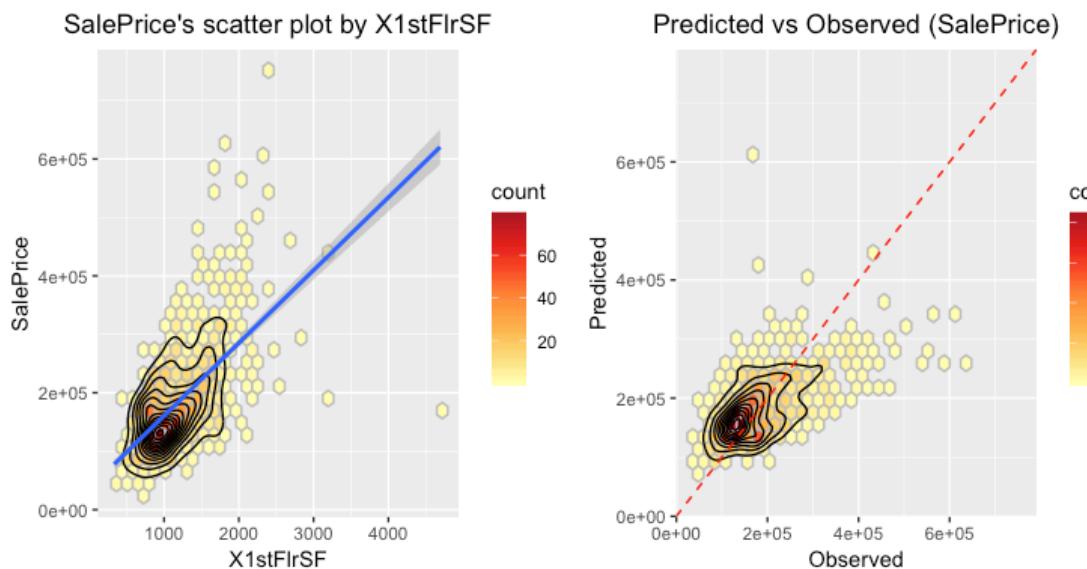
```
Residual standard error: 63224 on 1458 degrees of freedom
Multiple R-squared:  0.36706, Adjusted R-squared:  0.36662
F-statistic: 846 on 1 and 1458 DF,  p-value: 5.39471061897331e-147
```

Coefficients:

Simple Linear Model coefficients : X1stFlrSF

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	36173.45	5245.73	6.90	0
X1stFlrSF	124.50	4.28	29.08	0

Visualization:



[X2ndFlrSF]

Simple Linear Model Information

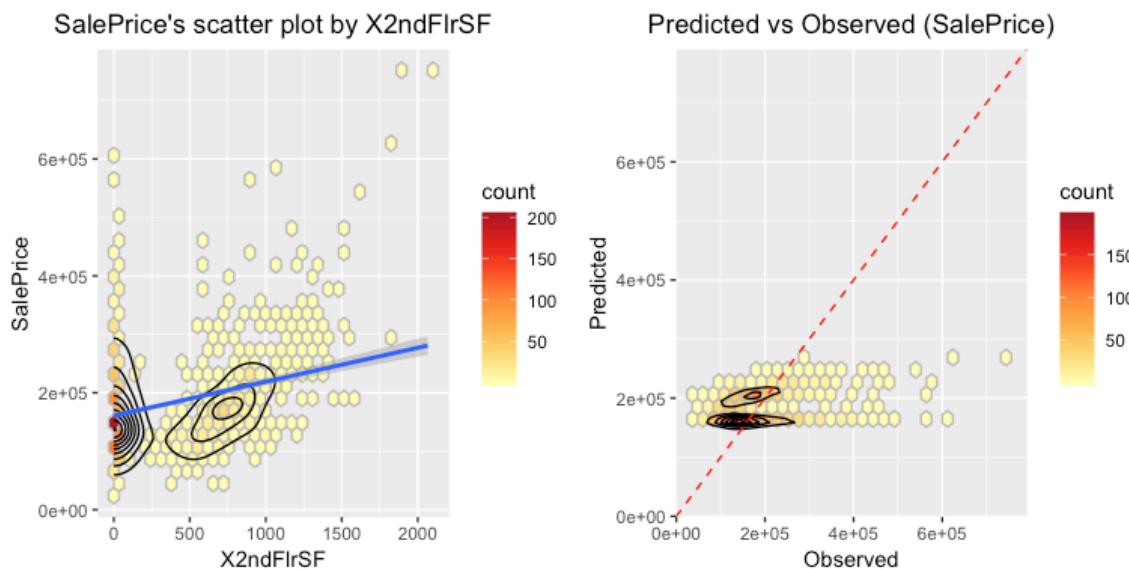
```
Residual standard error: 75309 on 1458 degrees of freedom
Multiple R-squared:  0.10197, Adjusted R-squared:  0.10136
F-statistic: 166 on 1 and 1458 DF,  p-value: 5.76433511918468e-36
```

Coefficients:

Simple Linear Model coefficients : X2ndFlrSF

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	160755.87	2518.07	63.84	0
X2ndFlrSF	58.11	4.52	12.87	0

Visualization:



[LowQualFinSF]

Simple Linear Model Information

```
Residual standard error: 79444 on 1458 degrees of freedom
Multiple R-squared:  0.00066, Adjusted R-squared: -3e-05
F-statistic: 1 on 1 and 1458 DF,  p-value: 0.328207309840526
```

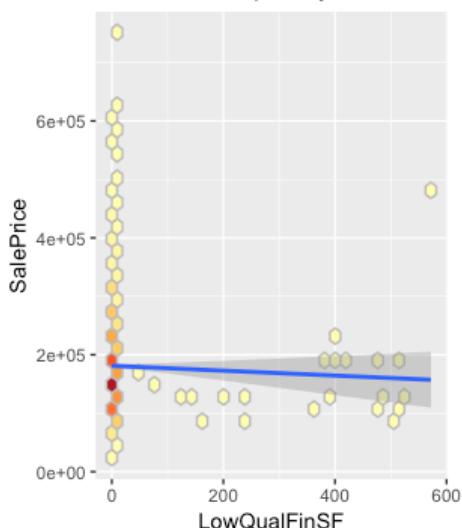
Coefficients:

Simple Linear Model coefficients : LowQualFinSF

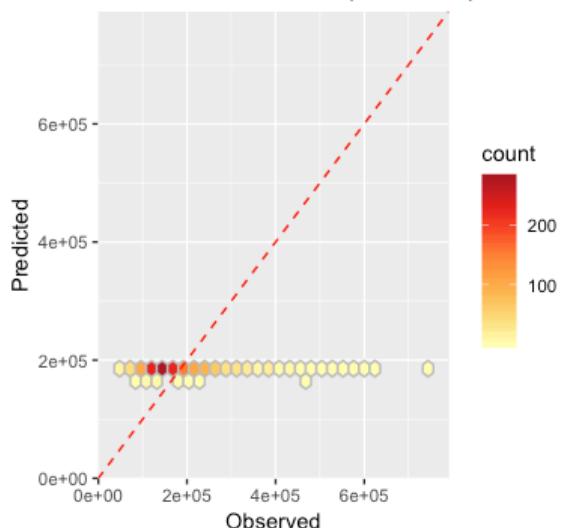
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	181165.71	2094.11	86.51	0.00
LowQualFinSF	-41.84	42.77	-0.98	0.33

Visualization:

SalePrice's scatter plot by LowQualFinSF



Predicted vs Observed (SalePrice)



[GrLivArea]

Simple Linear Model Information

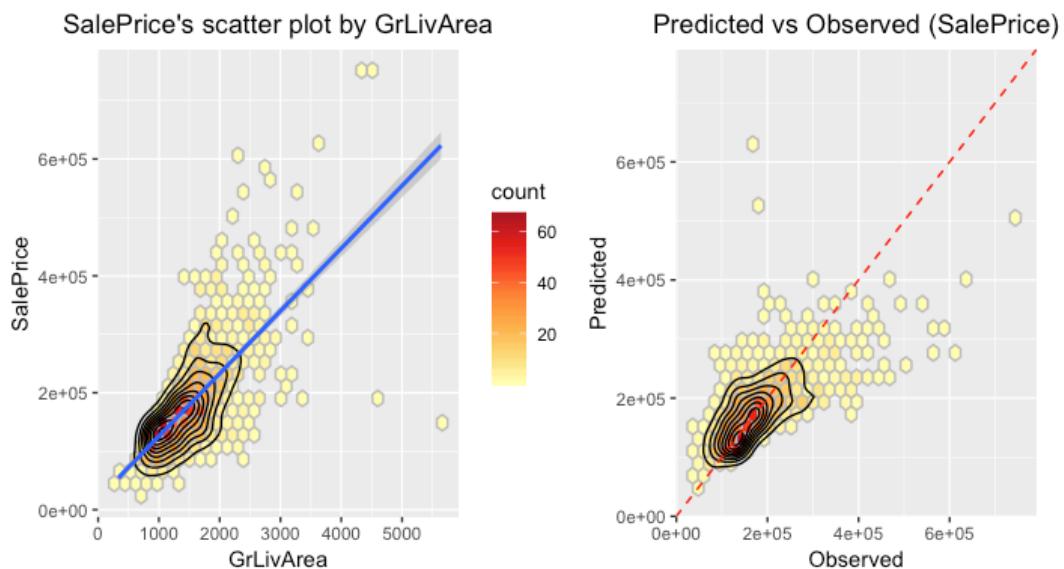
```
Residual standard error: 56073 on 1458 degrees of freedom
Multiple R-squared:  0.50215, Adjusted R-squared:  0.50181
F-statistic: 1471 on 1 and 1458 DF,  p-value: 4.51803364677319e-223
```

Coefficients:

Simple Linear Model coefficients : GrLivArea

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	18569.03	4480.75	4.14	0
GrLivArea	107.13	2.79	38.35	0

Visualization:



[BsmtFullBath]

Simple Linear Model Information

```
Residual standard error: 77393 on 1458 degrees of freedom
Multiple R-squared:  0.05158, Adjusted R-squared:  0.05093
F-statistic: 79 on 1 and 1458 DF,  p-value: 1.55034413721318e-18
```

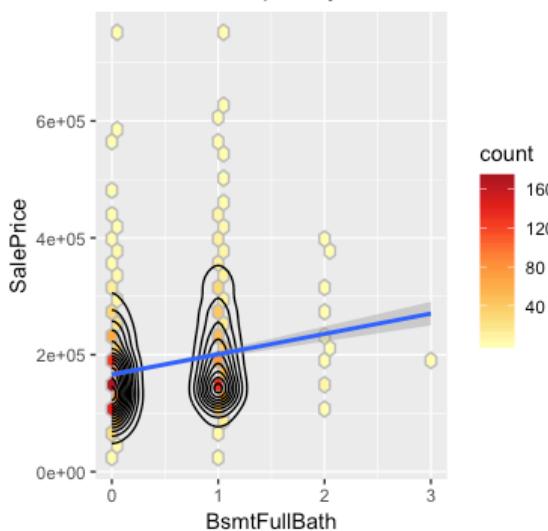
Coefficients:

Simple Linear Model coefficients : BsmtFullBath

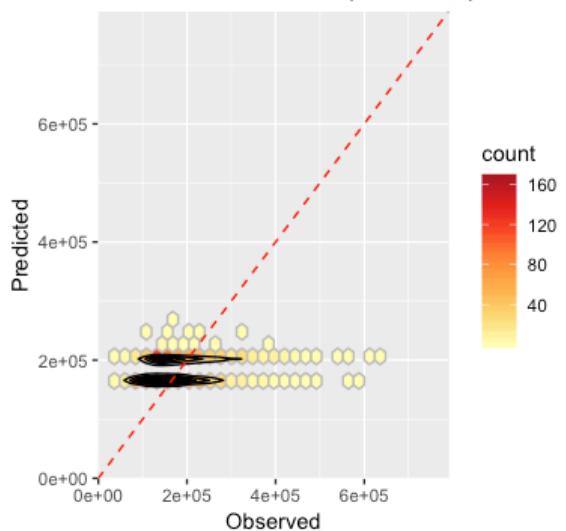
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	166131.52	2619.31	63.43	0
BsmtFullBath	34771.23	3904.64	8.91	0

Visualization:

SalePrice's scatter plot by BsmtFullBath



Predicted vs Observed (SalePrice)



[BsmtHalfBath]

Simple Linear Model Information

```
Residual standard error: 79458 on 1458 degrees of freedom
Multiple R-squared:  0.00028, Adjusted R-squared: -4e-04
F-statistic: 0 on 1 and 1458 DF,  p-value: 0.520153635727761
```

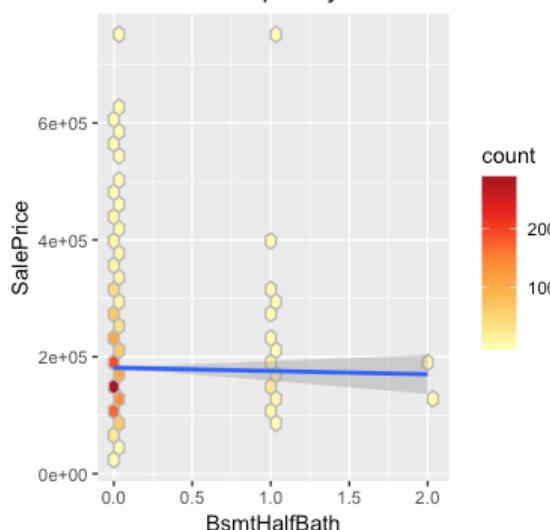
Coefficients:

Simple Linear Model coefficients : BsmtHalfBath

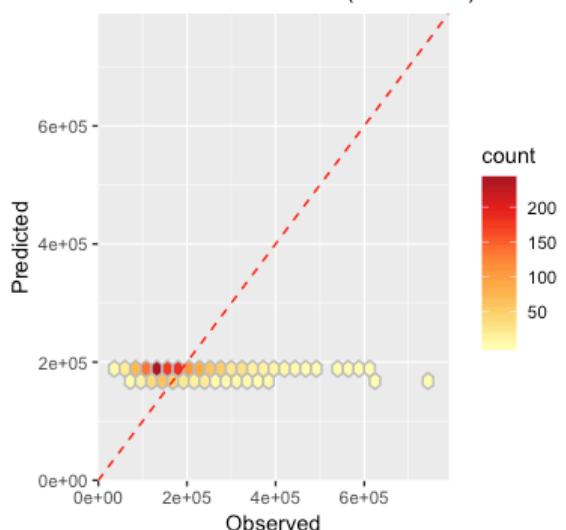
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	181243.66	2139.09	84.73	0.00
BsmtHalfBath	-5604.72	8712.93	-0.64	0.52

Visualization:

SalePrice's scatter plot by BsmtHalfBath



Predicted vs Observed (SalePrice)



[FullBath]

Simple Linear Model Information

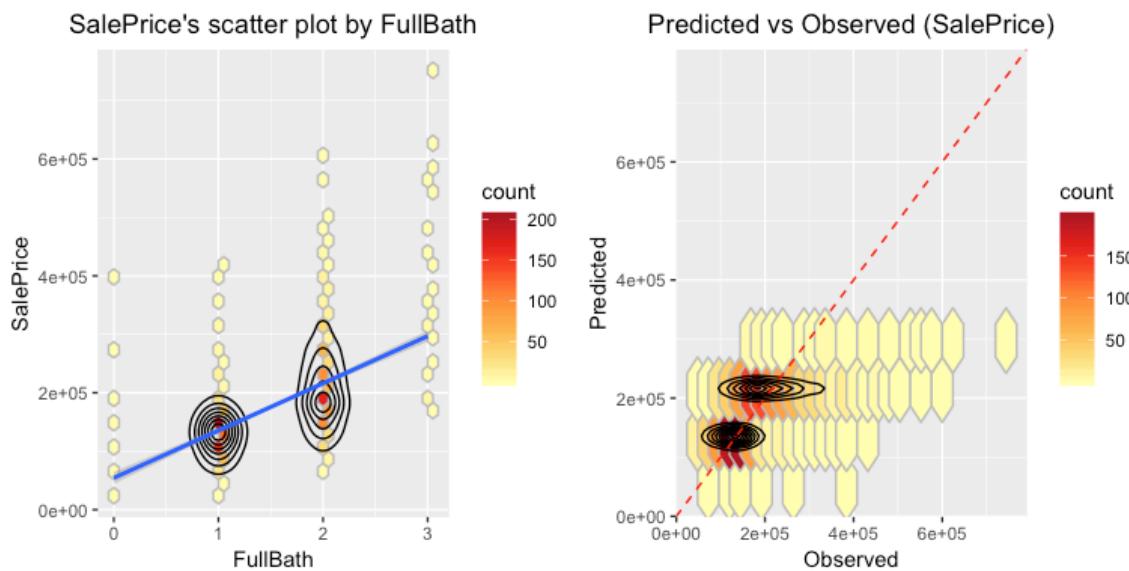
```
Residual standard error: 65804 on 1458 degrees of freedom
Multiple R-squared:  0.31434, Adjusted R-squared:  0.31387
F-statistic: 668 on 1 and 1458 DF,  p-value: 1.236470066764e-121
```

Coefficients:

Simple Linear Model coefficients : FullBath

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	54388.28	5188.3	10.48	0
FullBath	80848.17	3127.1	25.85	0

Visualization:



[HalfBath]

Simple Linear Model Information

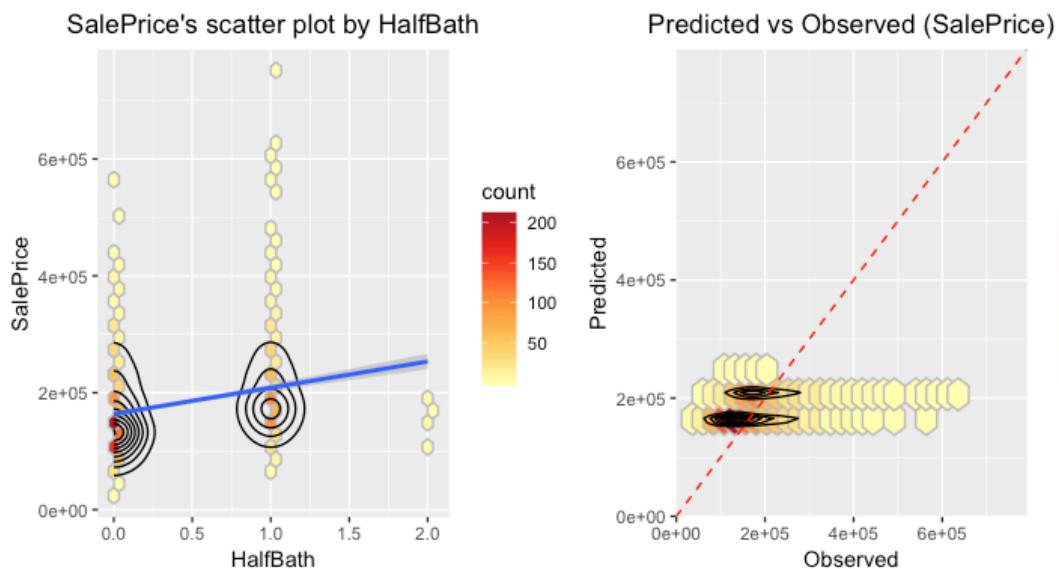
```
Residual standard error: 76195 on 1458 degrees of freedom
Multiple R-squared:  0.08072, Adjusted R-squared:  0.08009
F-statistic: 128 on 1 and 1458 DF,  p-value: 1.65047339557196e-28
```

Coefficients:

Simple Linear Model coefficients : HalfBath

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	163737.13	2506.61	65.32	0
HalfBath	44881.45	3966.70	11.31	0

Visualization:



[BedroomAbvGr]

Simple Linear Model Information

```
Residual standard error: 78337 on 1458 degrees of freedom
Multiple R-squared:  0.0283,   Adjusted R-squared:  0.02763
F-statistic: 42 on 1 and 1458 DF,  p-value: 9.92749732617604e-11
```

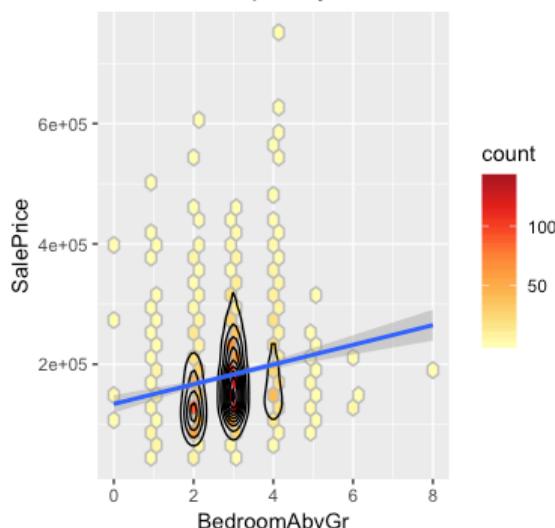
Coefficients:

Simple Linear Model coefficients : BedroomAbvGr

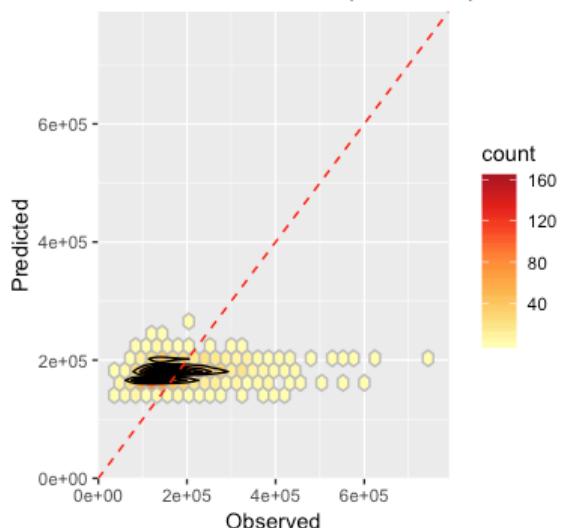
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	133966.02	7492.25	17.88	0
BedroomAbvGr	16381.02	2514.02	6.52	0

Visualization:

SalePrice's scatter plot by BedroomAbvGr



Predicted vs Observed (SalePrice)



[KitchenAbvGr]

Simple Linear Model Information

```
Residual standard error: 78732 on 1458 degrees of freedom
Multiple R-squared:  0.01847, Adjusted R-squared:  0.0178
F-statistic: 27 on 1 and 1458 DF,  p-value: 1.8604260320753e-07
```

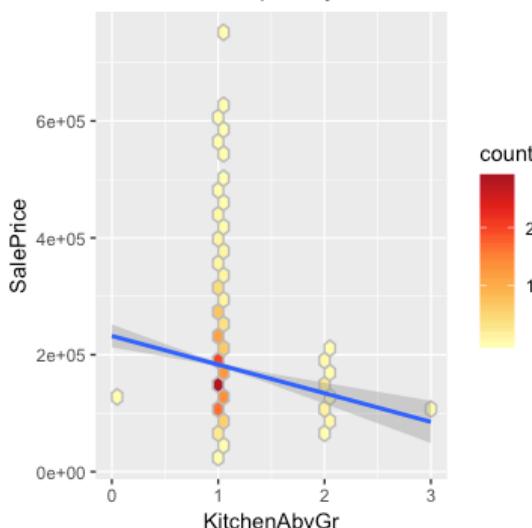
Coefficients:

Simple Linear Model coefficients : KitchenAbvGr

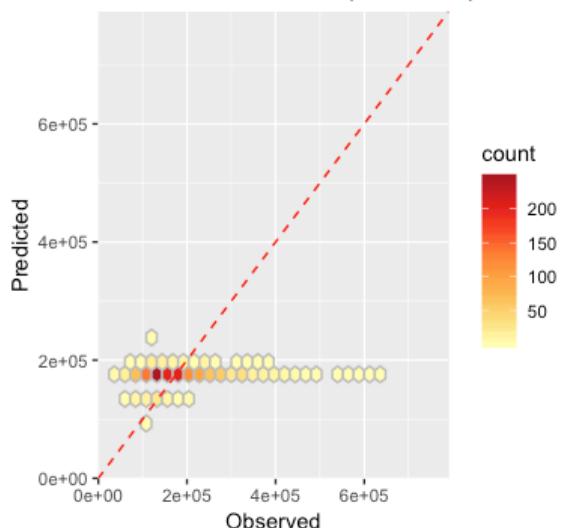
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	232204.58	10005.02	23.21	0
KitchenAbvGr	-49001.13	9354.83	-5.24	0

Visualization:

SalePrice's scatter plot by KitchenAbvGr



Predicted vs Observed (SalePrice)



[TotRmsAbvGrd]

Simple Linear Model Information

```
Residual standard error: 67204 on 1458 degrees of freedom
Multiple R-squared:  0.28486, Adjusted R-squared:  0.28437
F-statistic: 581 on 1 and 1458 DF,  p-value: 2.77228093295957e-108
```

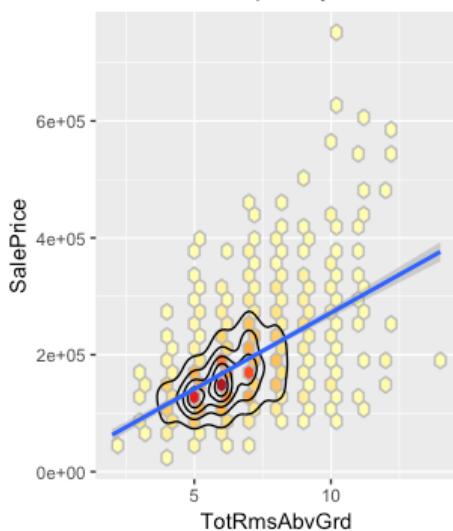
Coefficients:

Simple Linear Model coefficients : TotRmsAbvGrd

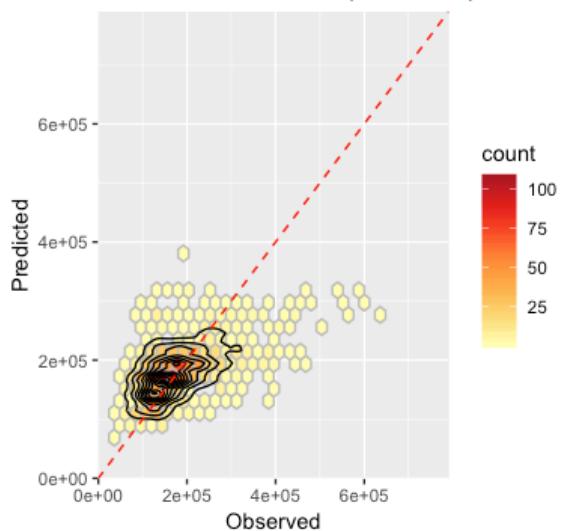
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	10896.47	7271.18	1.5	0.13
TotRmsAbvGrd	26086.18	1082.46	24.1	0.00

Visualization:

SalePrice's scatter plot by TotRmsAbvGrd



Predicted vs Observed (SalePrice)



[Fireplaces]

Simple Linear Model Information

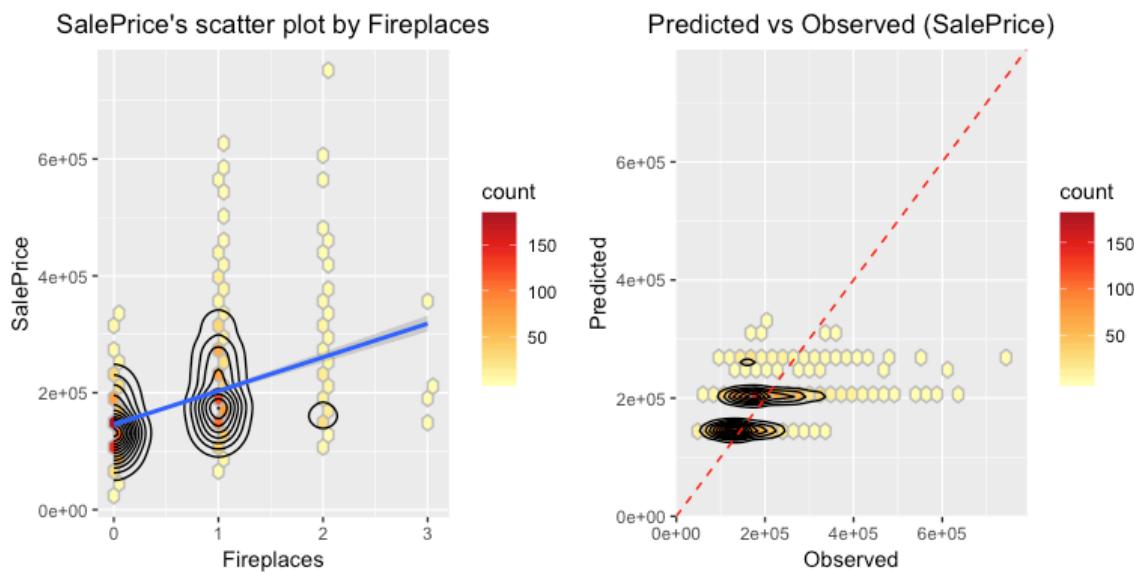
```
Residual standard error: 70275 on 1458 degrees of freedom
Multiple R-squared:  0.21802, Adjusted R-squared:  0.21749
F-statistic: 407 on 1 and 1458 DF,  p-value: 6.14148746111316e-80
```

Coefficients:

Simple Linear Model coefficients : Fireplaces

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	145648.49	2538.35	57.38	0
Fireplaces	57539.83	2853.88	20.16	0

Visualization:



[GarageCars]

Simple Linear Model Information

```
Residual standard error: 61035 on 1458 degrees of freedom
Multiple R-squared:  0.41012, Adjusted R-squared:  0.40972
F-statistic: 1014 on 1 and 1458 DF,  p-value: 2.49864416716425e-169
```

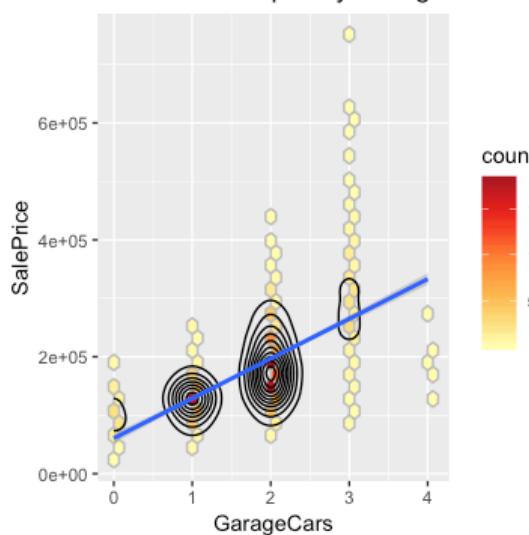
Coefficients:

Simple Linear Model coefficients : GarageCars

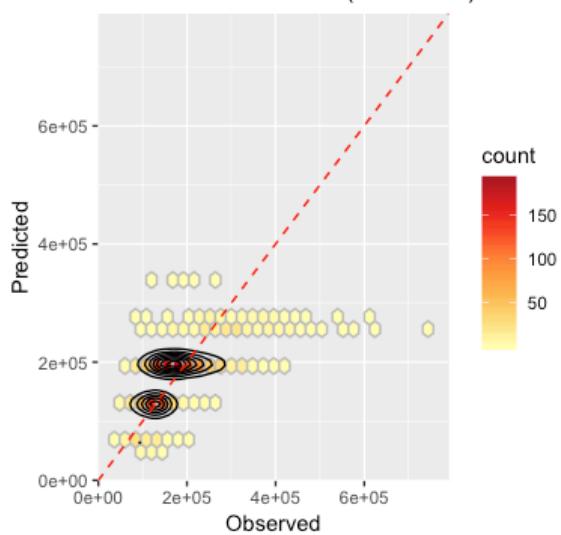
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	60618.98	4102.26	14.78	0
GarageCars	68078.00	2138.21	31.84	0

Visualization:

SalePrice's scatter plot by GarageCars



Predicted vs Observed (SalePrice)



[GarageArea]

Simple Linear Model Information

```
Residual standard error: 62136 on 1458 degrees of freedom
Multiple R-squared:  0.38867, Adjusted R-squared:  0.38825
F-statistic: 927 on 1 and 1458 DF,  p-value: 5.26503816797725e-158
```

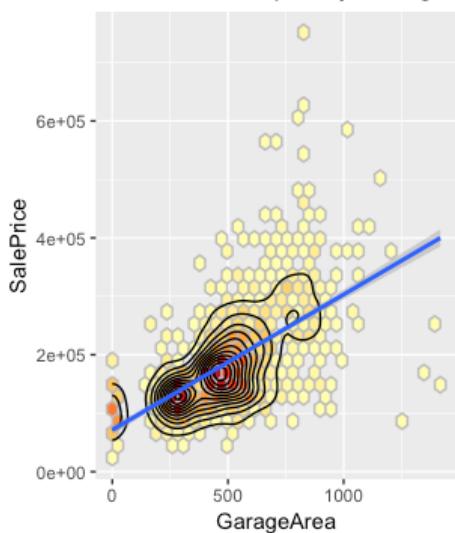
Coefficients:

Simple Linear Model coefficients : GarageArea

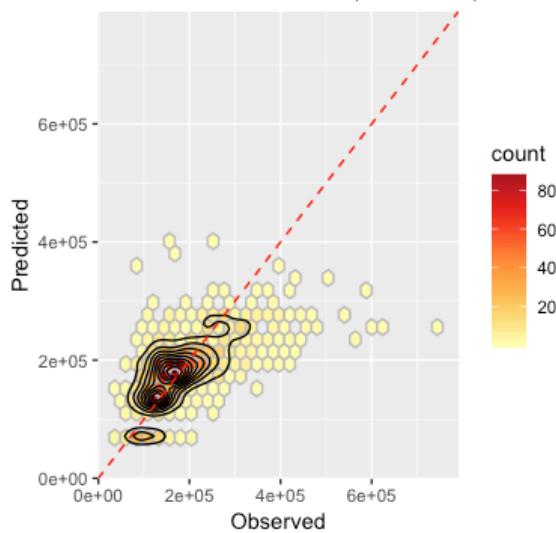
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	71357.42	3949.00	18.07	0
GarageArea	231.65	7.61	30.45	0

Visualization:

SalePrice's scatter plot by GarageArea



Predicted vs Observed (SalePrice)



[WoodDeckSF]

Simple Linear Model Information

```
Residual standard error: 75172 on 1458 degrees of freedom
Multiple R-squared:  0.10524, Adjusted R-squared:  0.10463
F-statistic: 171 on 1 and 1458 DF,  p-value: 3.97221653562683e-37
```

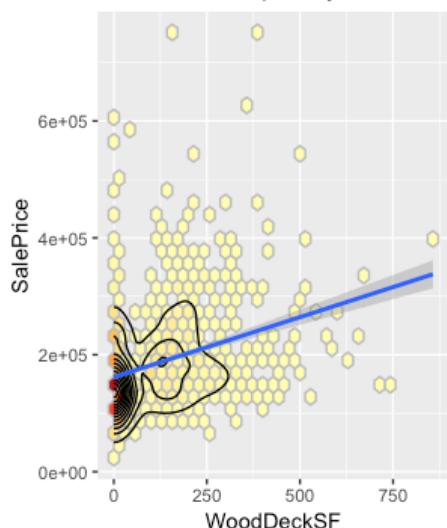
Coefficients:

Simple Linear Model coefficients : WoodDeckSF

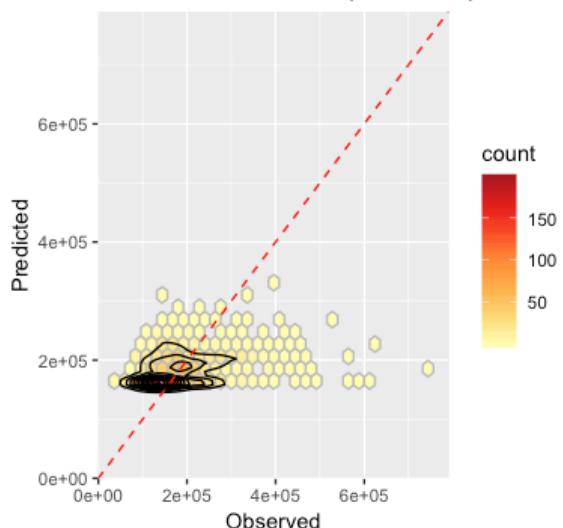
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	161542.60	2461.74	65.62	0
WoodDeckSF	205.62	15.70	13.10	0

Visualization:

SalePrice's scatter plot by WoodDeckSF



Predicted vs Observed (SalePrice)



[OpenPorchSF]

Simple Linear Model Information

```
Residual standard error: 75401 on 1458 degrees of freedom
Multiple R-squared:  0.09977, Adjusted R-squared:  0.09915
F-statistic: 162 on 1 and 1458 DF,  p-value: 3.49337362309578e-35
```

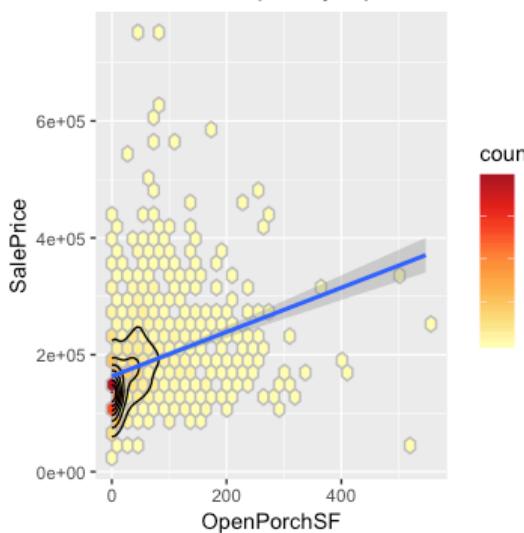
Coefficients:

Simple Linear Model coefficients : OpenPorchSF

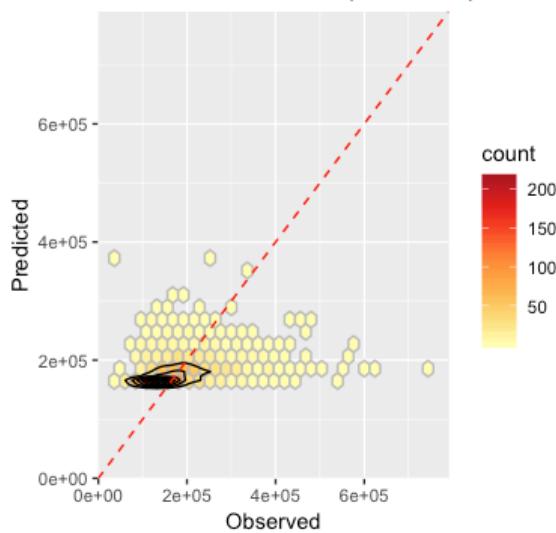
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	163250.07	2413.86	67.63	0
OpenPorchSF	378.72	29.79	12.71	0

Visualization:

SalePrice's scatter plot by OpenPorchSF



Predicted vs Observed (SalePrice)



[EnclosedPorch]

Simple Linear Model Information

```
Residual standard error: 78810 on 1458 degrees of freedom
Multiple R-squared:  0.01653, Adjusted R-squared:  0.01586
F-statistic: 25 on 1 and 1458 DF,  p-value: 8.2557704751123e-07
```

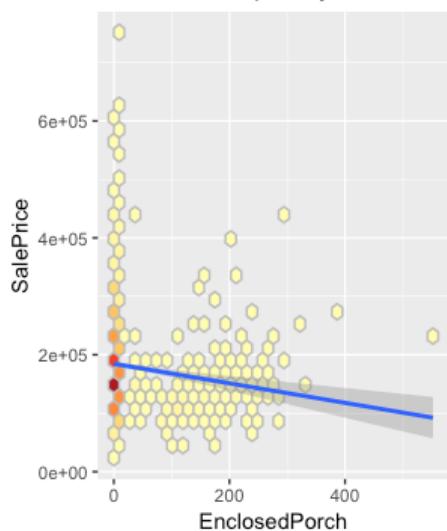
Coefficients:

Simple Linear Model coefficients : EnclosedPorch

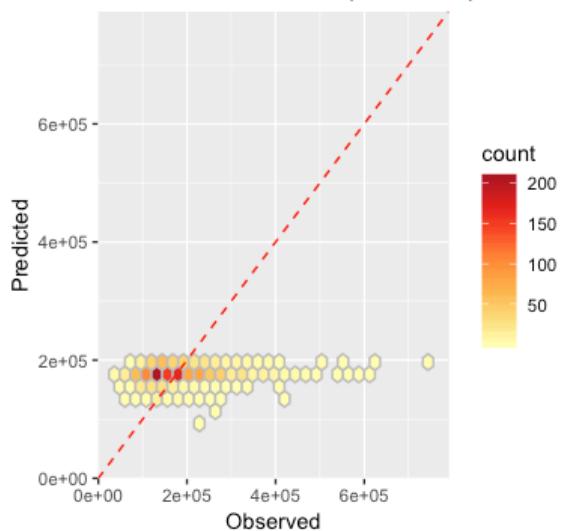
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	184590.28	2191.67	84.22	0
EnclosedPorch	-167.13	33.76	-4.95	0

Visualization:

SalePrice's scatter plot by EnclosedPorch



Predicted vs Observed (SalePrice)



[X3SsnPorch]

Simple Linear Model Information

```
Residual standard error: 79391 on 1458 degrees of freedom
Multiple R-squared:  0.00199, Adjusted R-squared:  0.0013
F-statistic: 3 on 1 and 1458 DF,  p-value: 0.0885817035805621
```

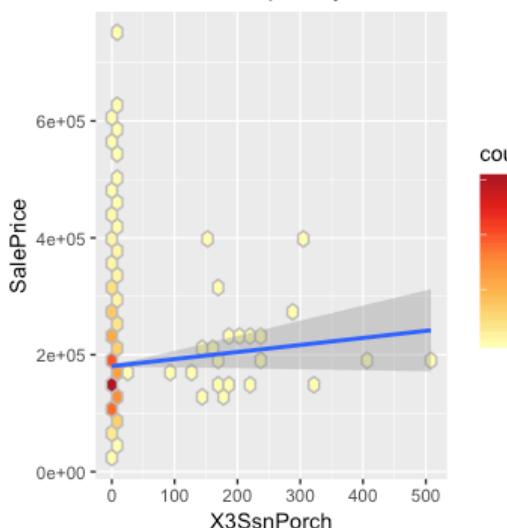
Coefficients:

Simple Linear Model coefficients : X3SsnPorch

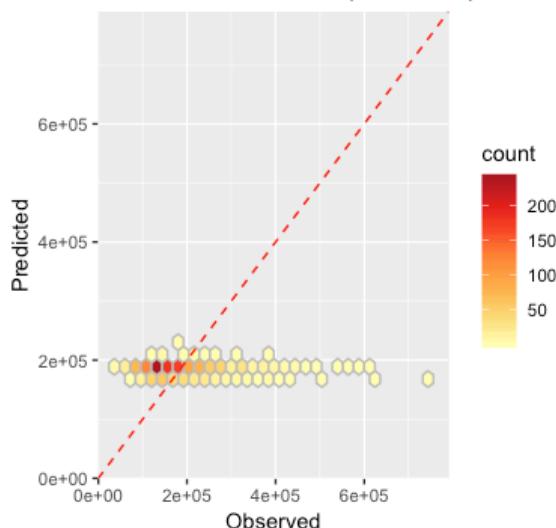
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	180509.28	2091.76	86.3	0.00
X3SsnPorch	120.81	70.90	1.7	0.09

Visualization:

SalePrice's scatter plot by X3SsnPorch



Predicted vs Observed (SalePrice)



[ScreenPorch]

Simple Linear Model Information

```
Residual standard error: 78975 on 1458 degrees of freedom
Multiple R-squared:  0.01242, Adjusted R-squared:  0.01174
F-statistic: 18 on 1 and 1458 DF,  p-value: 1.97214001946914e-05
```

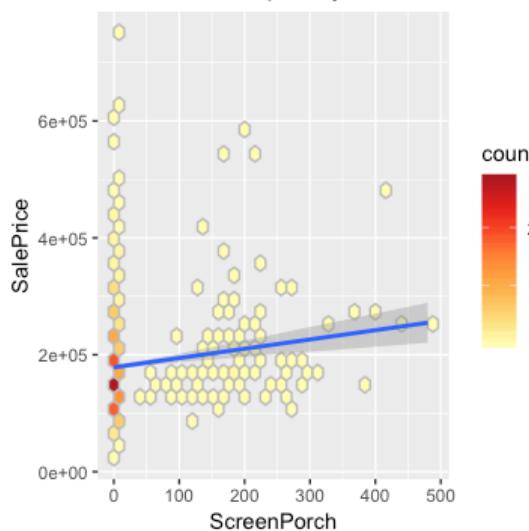
Coefficients:

Simple Linear Model coefficients : ScreenPorch

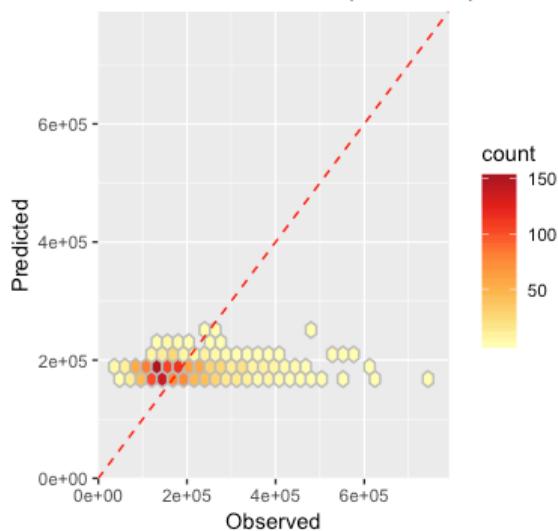
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	178529.70	2140.99	83.39	0
ScreenPorch	158.79	37.08	4.28	0

Visualization:

SalePrice's scatter plot by ScreenPorch



Predicted vs Observed (SalePrice)



[PoolArea]

Simple Linear Model Information

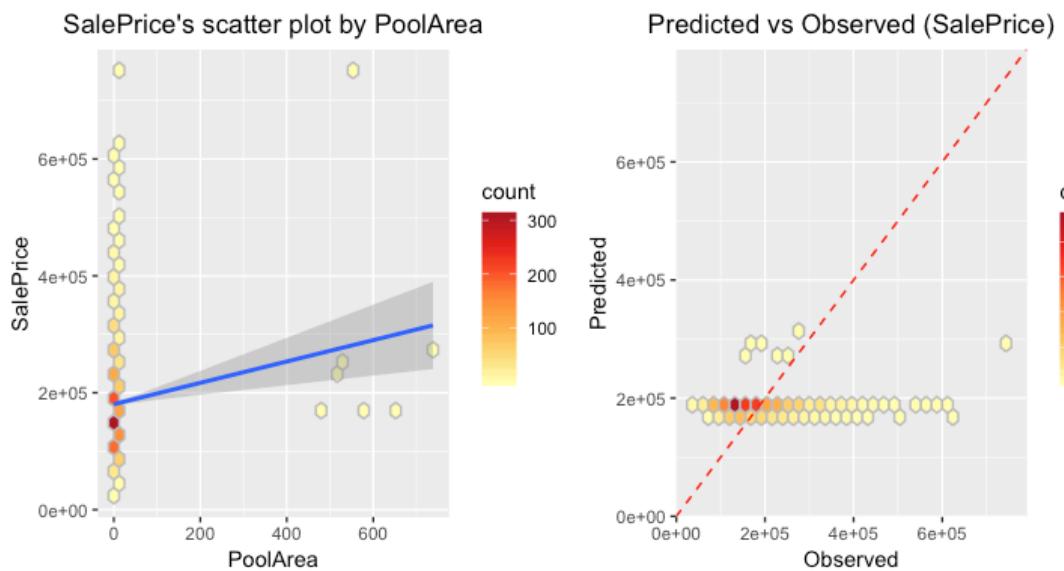
```
Residual standard error: 79130 on 1458 degrees of freedom
Multiple R-squared:  0.00854, Adjusted R-squared:  0.00786
F-statistic: 13 on 1 and 1458 DF,  p-value: 0.000407348960119601
```

Coefficients:

Simple Linear Model coefficients : PoolArea

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	180417.12	2075.80	86.91	0
PoolArea	182.71	51.56	3.54	0

Visualization:



[MiscVal]

Simple Linear Model Information

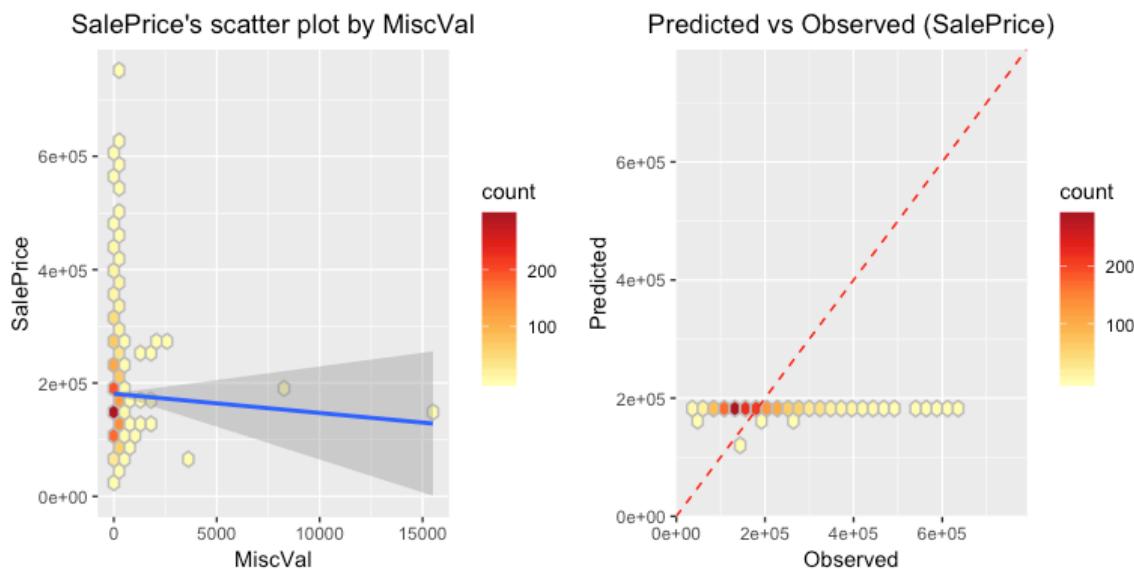
```
Residual standard error: 79452 on 1458 degrees of freedom
Multiple R-squared:  0.00045, Adjusted R-squared: -0.00024
F-statistic: 1 on 1 and 1458 DF,  p-value: 0.418486349408032
```

Coefficients:

Simple Linear Model coefficients : MiscVal

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	181068.75	2087.33	86.75	0.00
MiscVal	-3.39	4.19	-0.81	0.42

Visualization:



[MoSold]

Simple Linear Model Information

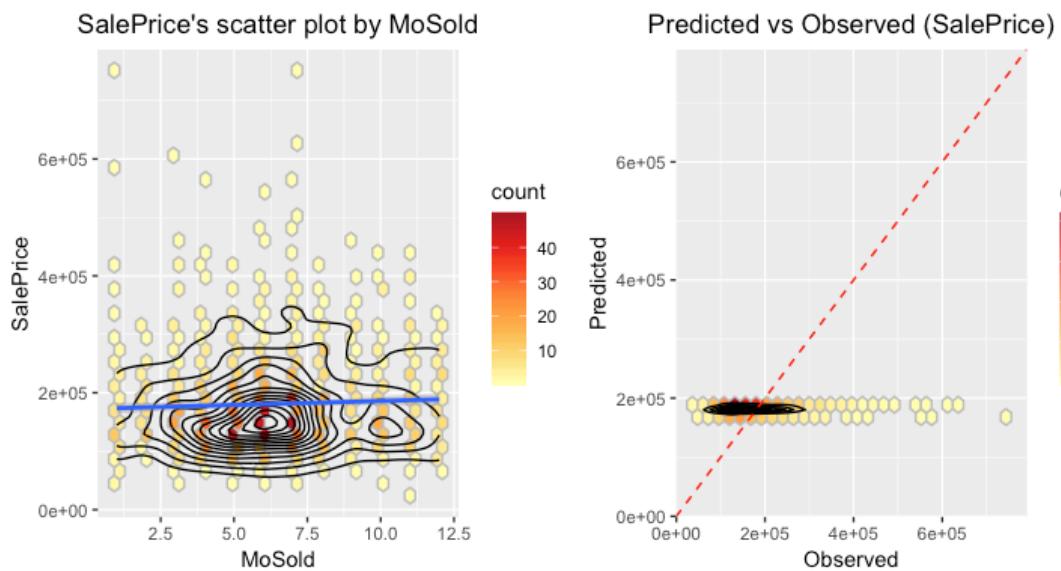
```
Residual standard error: 79384 on 1458 degrees of freedom
Multiple R-squared:  0.00216, Adjusted R-squared:  0.00147
F-statistic: 3 on 1 and 1458 DF,  p-value: 0.0761275785060439
```

Coefficients:

Simple Linear Model coefficients : MoSold

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	172295.88	5285.15	32.60	0.00
MoSold	1364.35	768.70	1.77	0.08

Visualization:



[YrSold]

Simple Linear Model Information

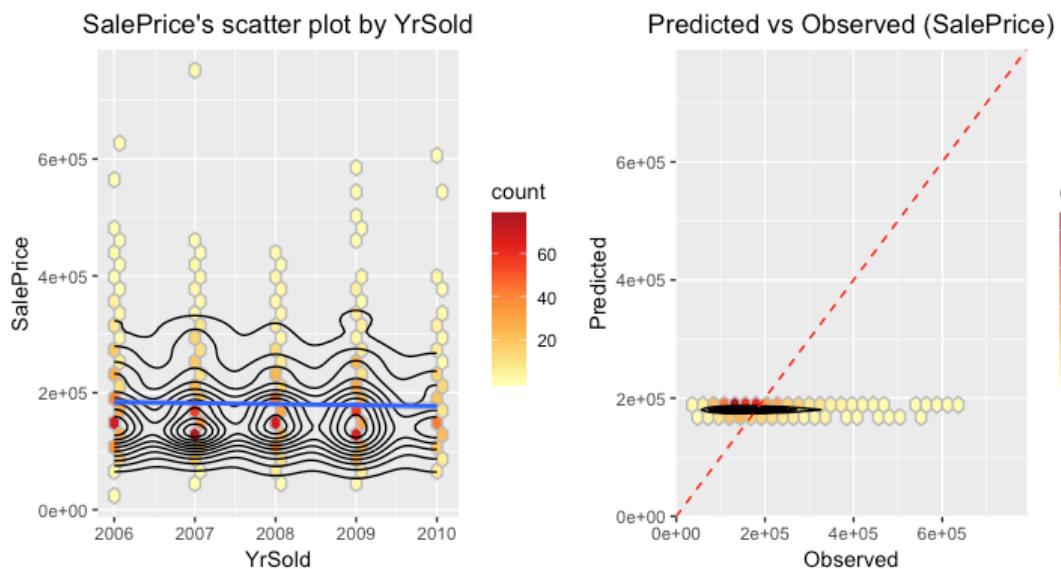
```
Residual standard error: 79436 on 1458 degrees of freedom
Multiple R-squared:  0.00084, Adjusted R-squared:  0.00015
F-statistic: 1 on 1 and 1458 DF,  p-value: 0.269413193280875
```

Coefficients:

Simple Linear Model coefficients : YrSold

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	3654560.37	3144033.8	1.16	0.25
YrSold	-1730.06	1565.9	-1.10	0.27

Visualization:



4.1.2 Grouped Categorical Variables

[MSZoning]

1. Analysis of Variance

Analysis of Variance Table : MSZoning

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
MSZoning	4	9.904000e+11	247600007088	43.84	0
Residuals	1455	8.217511e+12	5647774094	NA	NA

2. Simple Linear Model Information

```
Residual standard error: 75152 on 1455 degrees of freedom
Multiple R-squared:  0.10756, Adjusted R-squared:  0.10511
F-statistic: 44 on 4 and 1455 DF,  p-value: 5.46617371994783e-08
```

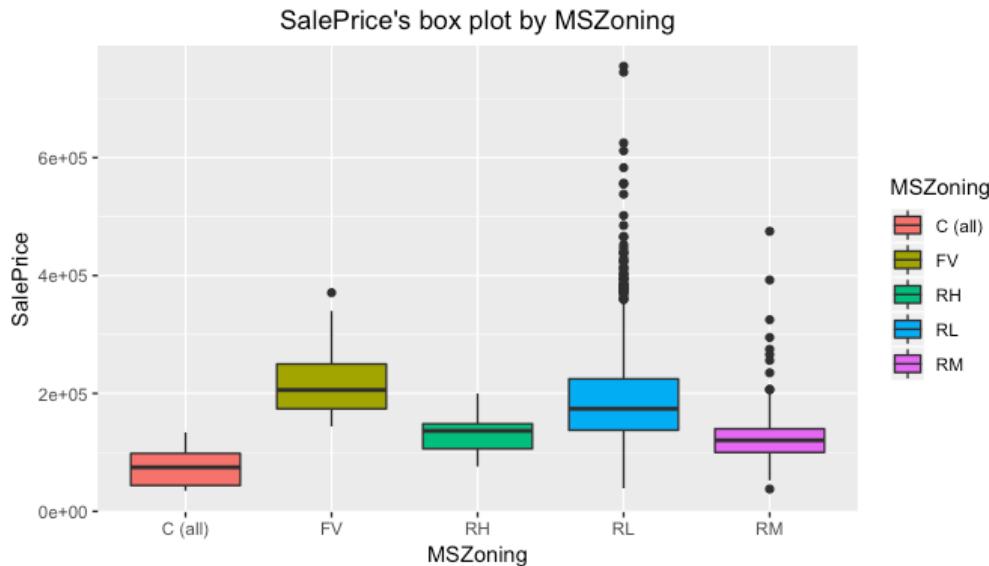
Coefficients:

Simple Linear Model coefficients : MSZoning

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	74528.00	23765.05	3.14	0.00
MSZoningFV	139486.06	25527.75	5.46	0.00

MSZoningRH	57030.37	30294.61	1.88	0.06
MSZoningRL	116476.99	23868.06	4.88	0.00
MSZoningRM	51788.83	24304.00	2.13	0.03

3. Visualization:



[Street]

1. Analysis of Variance

Analysis of Variance Table : Street

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
Street	1	1.550534e+10	15505341615	2.46	0.12
Residuals	1458	9.192406e+12	6304805208	NA	NA

2. Simple Linear Model Information

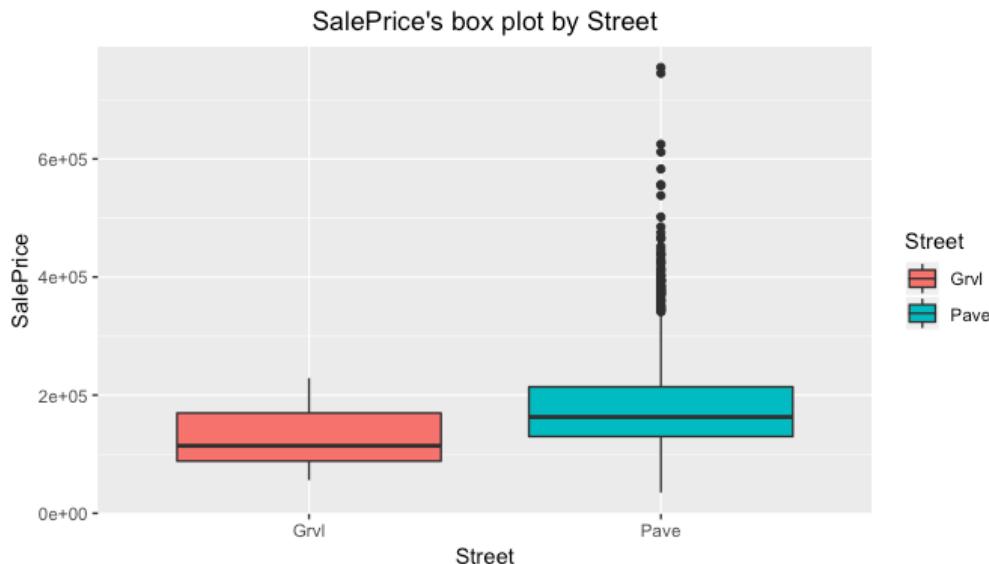
```
Residual standard error: 79403 on 1458 degrees of freedom
Multiple R-squared:  0.00168, Adjusted R-squared:  0.001
F-statistic: 2 on 1 and 1458 DF,  p-value: 0.117048604067807
```

Coefficients:

Simple Linear Model coefficients : Street

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	130190.50	32416.06	4.02	0.00
StreetPave	50940.04	32482.87	1.57	0.12

3. Visualization:



[LotShape]

1. Analysis of Variance

Analysis of Variance Table : LotShape

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
LotShape	3	7.032607e+11	234420247686	40.13	0
Residuals	1456	8.504651e+12	5841106176	NA	NA

2. Simple Linear Model Information

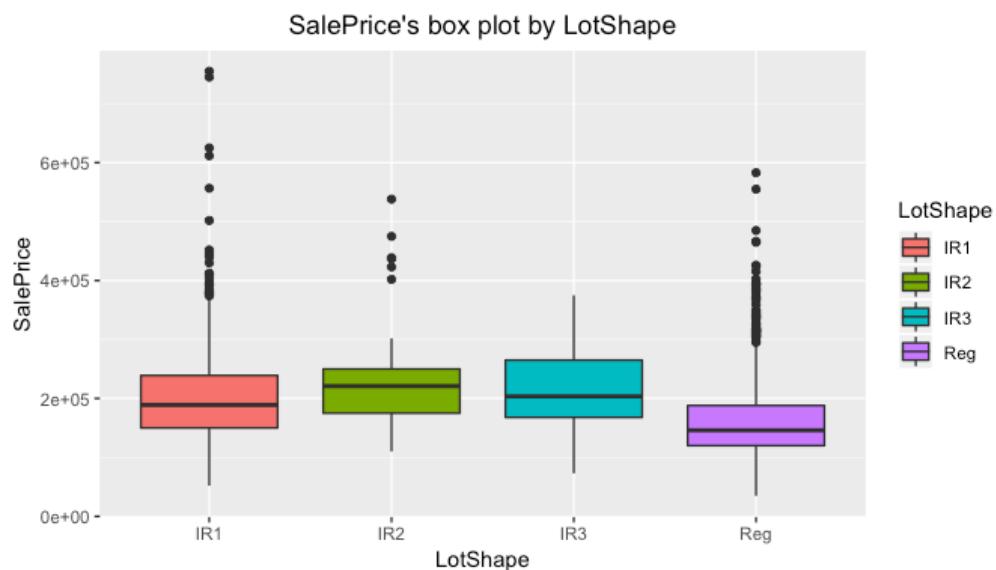
```
Residual standard error: 76427 on 1456 degrees of freedom
Multiple R-squared:  0.07638, Adjusted R-squared:  0.07447
F-statistic: 40 on 3 and 1456 DF,  p-value: 0.00673669009335628
```

Coefficients:

Simple Linear Model coefficients : LotShape

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	206101.67	3473.96	59.33	0.00
LotShapeIR2	33731.70	12431.19	2.71	0.01
LotShapeIR3	9934.83	24416.78	0.41	0.68
LotShapeReg	-41346.85	4287.55	-9.64	0.00

3. Visualization:



[LandContour]

1. Analysis of Variance

Analysis of Variance Table : LandContour

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
LandContour	3	2.375097e+11	79169883962	12.85	0
Residuals	1456	8.970402e+12	6160990167	NA	NA

2. Simple Linear Model Information

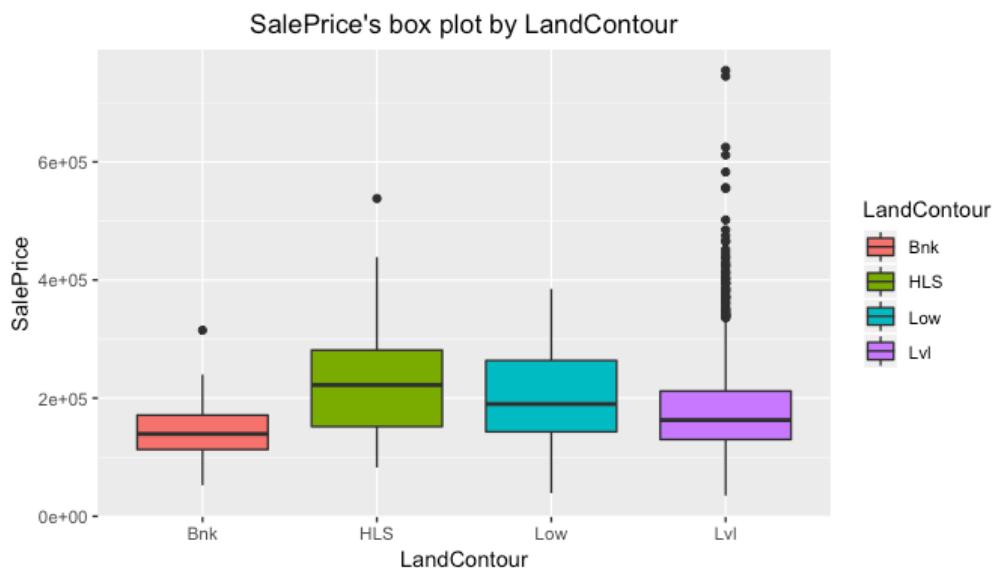
Residual standard error: 78492 on 1456 degrees of freedom
 Multiple R-squared: 0.02579, Adjusted R-squared: 0.02379
 F-statistic: 13 on 3 and 1456 DF, p-value: 3.38850951608263e-09

Coefficients:

Simple Linear Model coefficients : LandContour

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	143104.08	9889.06	14.47	0
LandContourHLS	88429.86	14866.52	5.95	0
LandContourLow	60557.03	16399.15	3.69	0
LandContourLvl	37079.67	10123.88	3.66	0

3. Visualization:



[Utilities]

1. Analysis of Variance

Analysis of Variance Table : Utilities

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
Utilities	1	1.886693e+09	1886692508	0.3	0.58
Residuals	1458	9.206025e+12	6314145845	NA	NA

2. Simple Linear Model Information

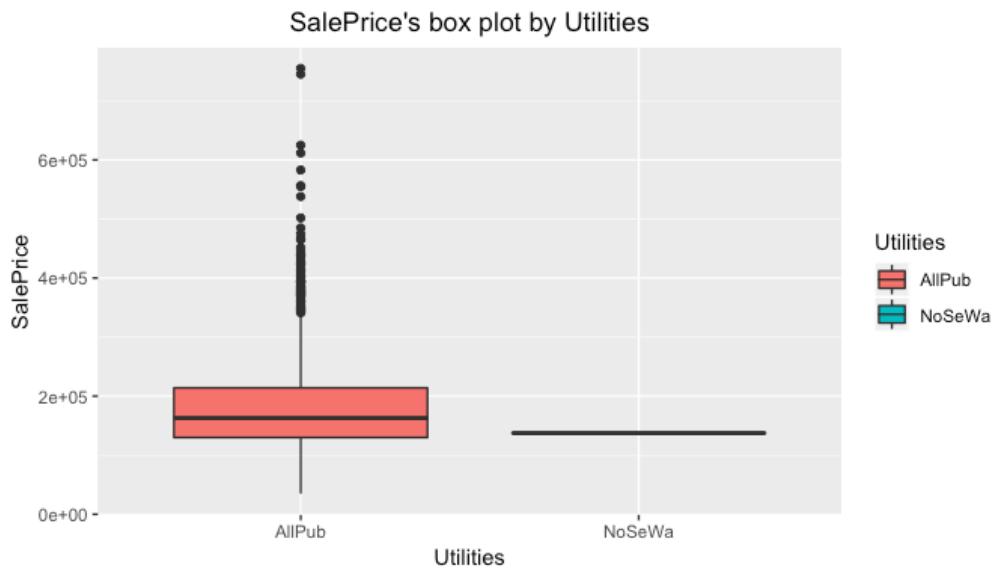
```
Residual standard error: 79462 on 1458 degrees of freedom
Multiple R-squared:  2e-04,   Adjusted R-squared: -0.00048
F-statistic: 0 on 1 and 1458 DF,  p-value: 0.584716773968918
```

Coefficients:

Simple Linear Model coefficients : Utilities

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	180950.96	2080.32	86.98	0.00
UtilitiesNoSeWa	-43450.96	79488.83	-0.55	0.58

3. Visualization:



[LotConfig]

1. Analysis of Variance

Analysis of Variance Table : LotConfig

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
LotConfig	4	1.935444e+11	48386111244	7.81	0
Residuals	1455	9.014367e+12	6195441161	NA	NA

2. Simple Linear Model Information

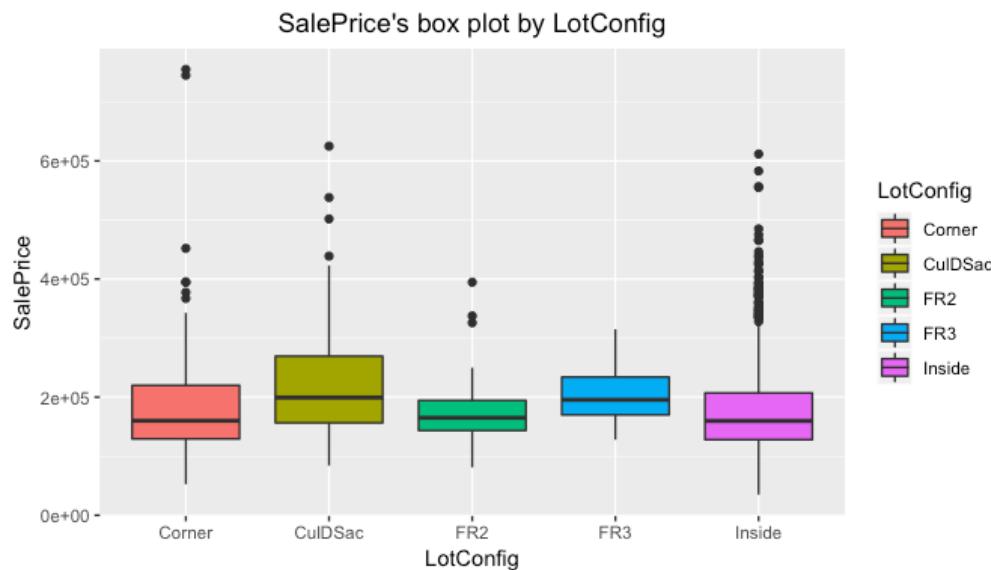
```
Residual standard error: 78711 on 1455 degrees of freedom
Multiple R-squared:  0.02102, Adjusted R-squared:  0.01833
F-statistic: 8 on 4 and 1455 DF,  p-value: 8.63169509087829e-06
```

Coefficients:

Simple Linear Model coefficients : LotConfig

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	181623.43	4853.54	37.42	0.00
LotConfigCulDSac	42231.19	9458.63	4.46	0.00
LotConfigFR2	-3688.85	12464.94	-0.30	0.77
LotConfigFR3	26851.57	39653.71	0.68	0.50
LotConfigInside	-4685.38	5426.42	-0.86	0.39

3. Visualization:



[LandSlope]

1. Analysis of Variance

Analysis of Variance Table : LandSlope

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
LandSlope	2	2.469217e+10	12346085214	1.96	0.14

Residuals	1457	9.183219e+12	6302827155	NA	NA
-----------	------	--------------	------------	----	----

2. Simple Linear Model Information

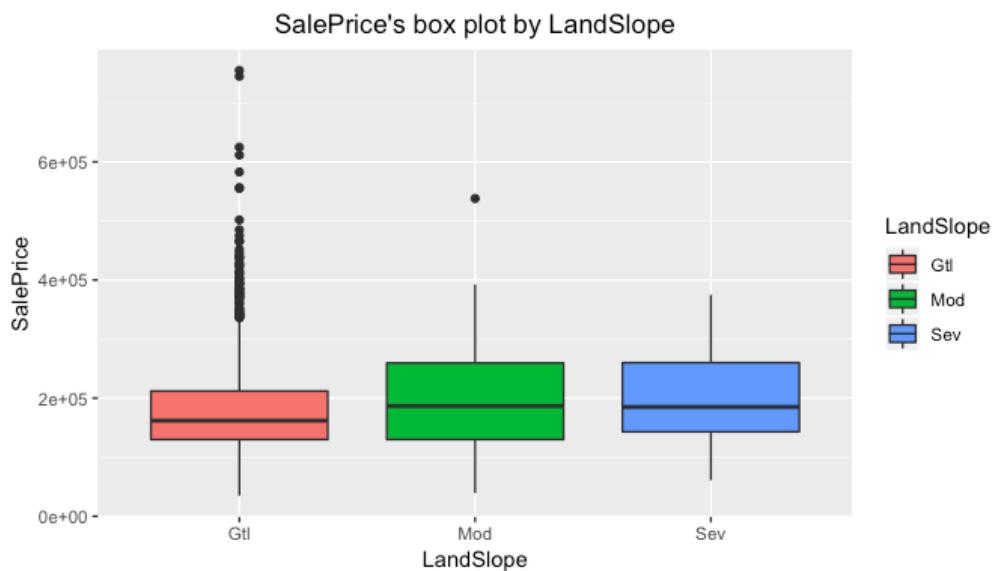
Residual standard error: 79390 on 1457 degrees of freedom
 Multiple R-squared: 0.00268, Adjusted R-squared: 0.00131
 F-statistic: 2 on 2 and 1457 DF, p-value: 0.0961143064472277

Coefficients:

Simple Linear Model coefficients : LandSlope

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	179956.80	2135.57	84.27	0.00
LandSlopeMod	16777.34	10076.07	1.67	0.10
LandSlopeSev	24422.43	22122.24	1.10	0.27

3. Visualization:



[Neighborhood]

1. Analysis of Variance

Analysis of Variance Table : Neighborhood

Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
----	--------	---------	---------	--------

Neighborhood	24	5.023606e+12	209316922573	71.78	0
Residuals	1435	4.184305e+12	2915892120	NA	NA

2. Simple Linear Model Information

Residual standard error: 53999 on 1435 degrees of freedom
 Multiple R-squared: 0.54557, Adjusted R-squared: 0.53797
 F-statistic: 72 on 24 and 1435 DF, p-value: 0.155463281670235

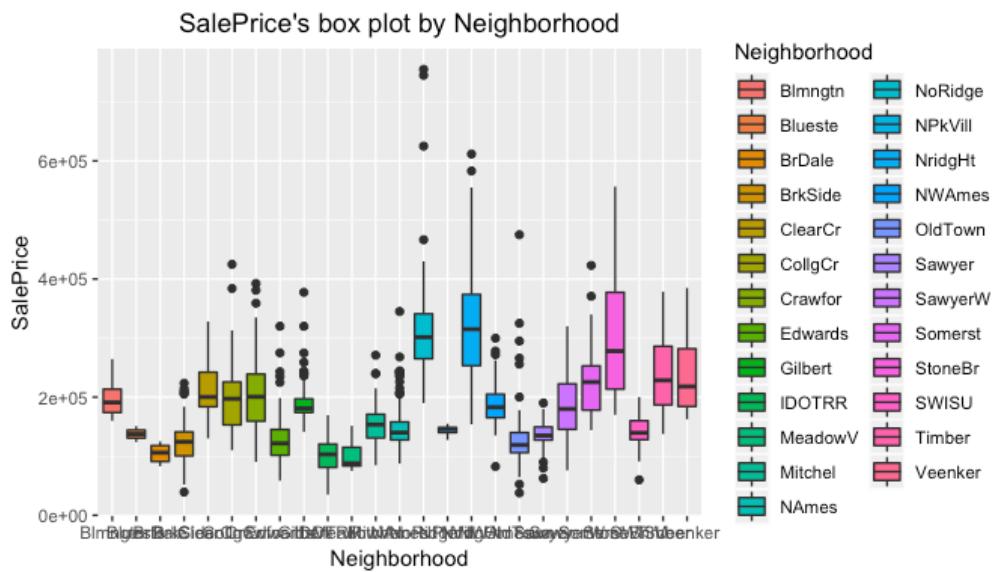
Coefficients:

Simple Linear Model coefficients : Neighborhood

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	194870.88	13096.68	14.88	0.00
NeighborhoodBlueste	-57370.88	40366.68	-1.42	0.16
NeighborhoodBrDale	-90377.13	18808.68	-4.81	0.00
NeighborhoodBrkSide	-70036.83	14892.85	-4.70	0.00
NeighborhoodClearCr	17694.55	16603.07	1.07	0.29
NeighborhoodCollgCr	3094.89	13818.91	0.22	0.82
NeighborhoodCrawfor	15753.84	15122.75	1.04	0.30
NeighborhoodEdwards	-66651.18	14166.23	-4.70	0.00
NeighborhoodGilbert	-2016.38	14437.21	-0.14	0.89
NeighborhoodIDOTRR	-94747.10	15821.85	-5.99	0.00
NeighborhoodMeadowV	-96294.41	18521.50	-5.20	0.00
NeighborhoodMitchel	-38600.76	15199.71	-2.54	0.01
NeighborhoodNAmes	-49023.80	13582.44	-3.61	0.00
NeighborhoodNoRidge	140424.43	15576.98	9.01	0.00
NeighborhoodNPkVill	-52176.44	22260.08	-2.34	0.02
NeighborhoodNridgHt	121399.74	14470.38	8.39	0.00
NeighborhoodNWAmes	-5820.81	14541.90	-0.40	0.69
NeighborhoodOldTown	-66645.58	14047.33	-4.74	0.00
NeighborhoodSawyer	-58077.75	14523.33	-4.00	0.00
NeighborhoodSawyerW	-8315.09	14864.22	-0.56	0.58

NeighborhoodSomerst	30508.95	14332.79	2.13	0.03
NeighborhoodStoneBr	115628.12	16975.24	6.81	0.00
NeighborhoodSWISU	-52279.52	16975.24	-3.08	0.00
NeighborhoodTimber	47376.57	15756.18	3.01	0.00
NeighborhoodVeenker	43901.84	20895.08	2.10	0.04

3. Visualization:



[Condition1]

1. Analysis of Variance

Analysis of Variance Table : Condition1

	Df	Sum.Sq	Mean.Sq	F.value	Pr.F.
Condition1	8	3.004600e+11	37557504942	6.12	0
Residuals	1451	8.907451e+12	6138836179	NA	NA

2. Simple Linear Model Information

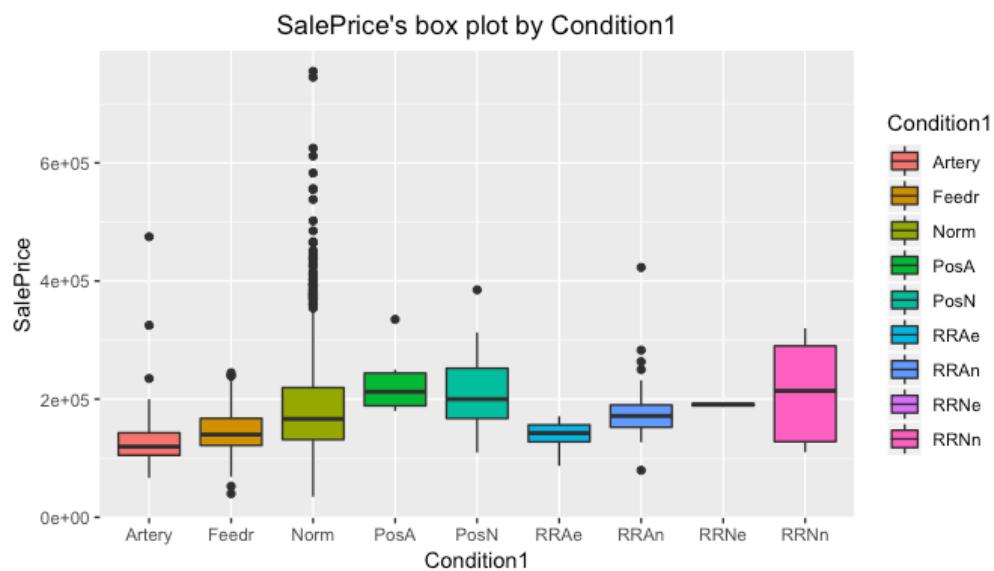
Residual standard error: 78351 on 1451 degrees of freedom
 Multiple R-squared: 0.03263, Adjusted R-squared: 0.0273
 F-statistic: 6 on 8 and 1451 DF, p-value: 0.604972762340044

Coefficients:

Simple Linear Model coefficients : Condition1

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	135091.67	11308.95	11.95	0.00
Condition1Feedr	7383.81	14271.67	0.52	0.60
Condition1Norm	49403.83	11522.35	4.29	0.00
Condition1PosA	90783.33	29920.68	3.03	0.00
Condition1PosN	80092.54	21236.50	3.77	0.00
Condition1RRAe	3308.33	26191.00	0.13	0.90
Condition1RRAn	49304.95	19078.82	2.58	0.01
Condition1RRNe	55658.33	56544.77	0.98	0.33
Condition1RRNn	77308.33	36819.28	2.10	0.04

3. Visualization:



[Condition2]

1. Analysis of Variance

Analysis of Variance Table : Condition2

Df	Sum.Sq	Mean.Sq	F.value	Pr..F.

Condition2	7	9.115059e+10	13021513452	2.07	0.04
Residuals	1452	9.116761e+12	6278760841	NA	NA

2. Simple Linear Model Information

Residual standard error: 79239 on 1452 degrees of freedom
 Multiple R-squared: 0.0099, Adjusted R-squared: 0.00513
 F-statistic: 2 on 7 and 1452 DF, p-value: 0.820693580038708

Coefficients:

Simple Linear Model coefficients : Condition2

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	106500.00	56030.17	1.90	0.06
Condition2Feedr	14666.67	64698.07	0.23	0.82
Condition2Norm	74669.41	56068.94	1.33	0.18
Condition2PosA	218500.00	97047.11	2.25	0.02
Condition2PosN	178375.00	79238.63	2.25	0.02
Condition2RRAe	83500.00	97047.11	0.86	0.39
Condition2RRAn	30405.00	97047.11	0.31	0.75
Condition2RRNn	-9750.00	79238.63	-0.12	0.90

3. Visualization:



[BldgType]

1. Analysis of Variance

Analysis of Variance Table : BldgType

	Df	Sum.Sq	Mean.Sq	F.value	Pr.F.
BldgType	4	3.179863e+11	79496564405	13.01	0
Residuals	1455	8.889925e+12	6109914142	NA	NA

2. Simple Linear Model Information

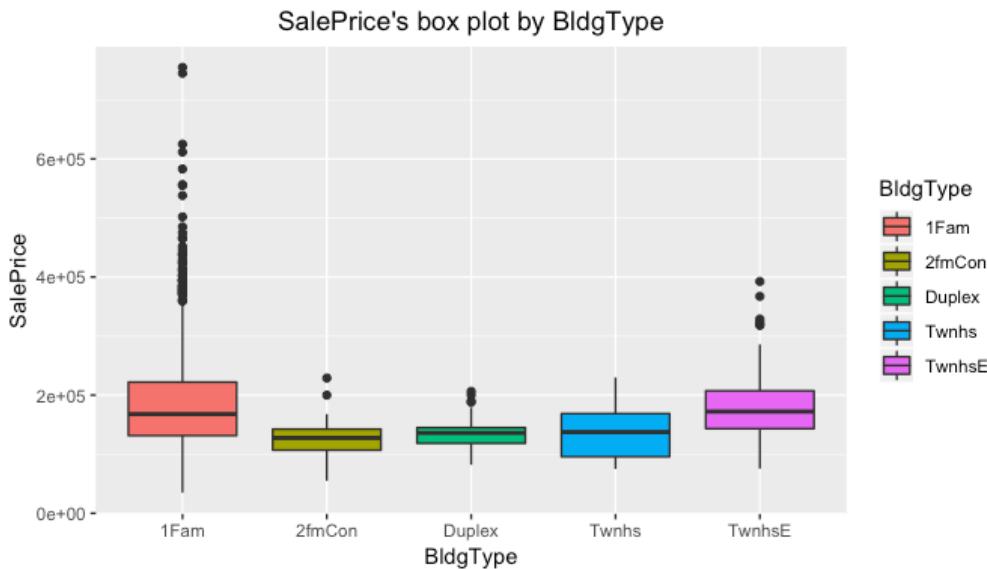
Residual standard error: 78166 on 1455 degrees of freedom
 Multiple R-squared: 0.03453, Adjusted R-squared: 0.03188
 F-statistic: 13 on 4 and 1455 DF, p-value: 5.79707083982456e-05

Coefficients:

Simple Linear Model coefficients : BldgType

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	185763.81	2237.88	83.01	0.00
BldgType2fmCon	-57331.55	14216.26	-4.03	0.00
BldgTypeDuplex	-52222.73	11068.26	-4.72	0.00
BldgTypeTwnhs	-49852.18	12128.44	-4.11	0.00
BldgTypeTwnhsE	-3804.47	7655.32	-0.50	0.62

3. Visualization:



[HouseStyle]

1. Analysis of Variance

Analysis of Variance Table : HouseStyle

	Df	Sum.Sq	Mean.Sq	F.value	Pr.F.
HouseStyle	7	7.947590e+11	113537002754	19.6	0
Residuals	1452	8.413152e+12	5794182035	NA	NA

2. Simple Linear Model Information

```
Residual standard error: 76120 on 1452 degrees of freedom
Multiple R-squared:  0.08631, Adjusted R-squared:  0.08191
F-statistic: 20 on 7 and 1452 DF,  p-value: 0.121001628491992
```

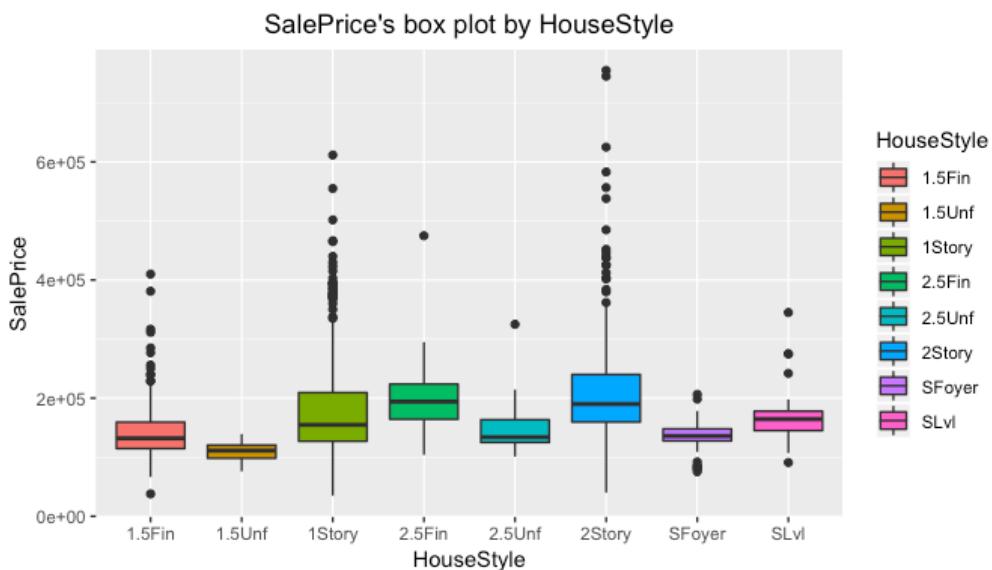
Coefficients:

Simple Linear Model coefficients : HouseStyle

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	143116.74	6133.89	23.33	0.00

HouseStyle1.5Unf	-32966.74	21248.40	-1.55	0.12
HouseStyle1Story	32868.74	6753.19	4.87	0.00
HouseStyle2.5Fin	76883.26	27602.49	2.79	0.01
HouseStyle2.5Unf	14237.81	23756.44	0.60	0.55
HouseStyle2Story	66935.02	7116.54	9.41	0.00
HouseStyleSoyer	-8042.25	13936.43	-0.58	0.56
HouseStyleSLvl	23586.64	11259.03	2.09	0.04

3. Visualization:



[RoofStyle]

1. Analysis of Variance

Analysis of Variance Table : RoofStyle

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
RoofStyle	5	5.312655e+11	106253090905	17.81	0
Residuals	1454	8.676646e+12	5967431829	NA	NA

2. Simple Linear Model Information

Residual standard error: 77249 on 1454 degrees of freedom

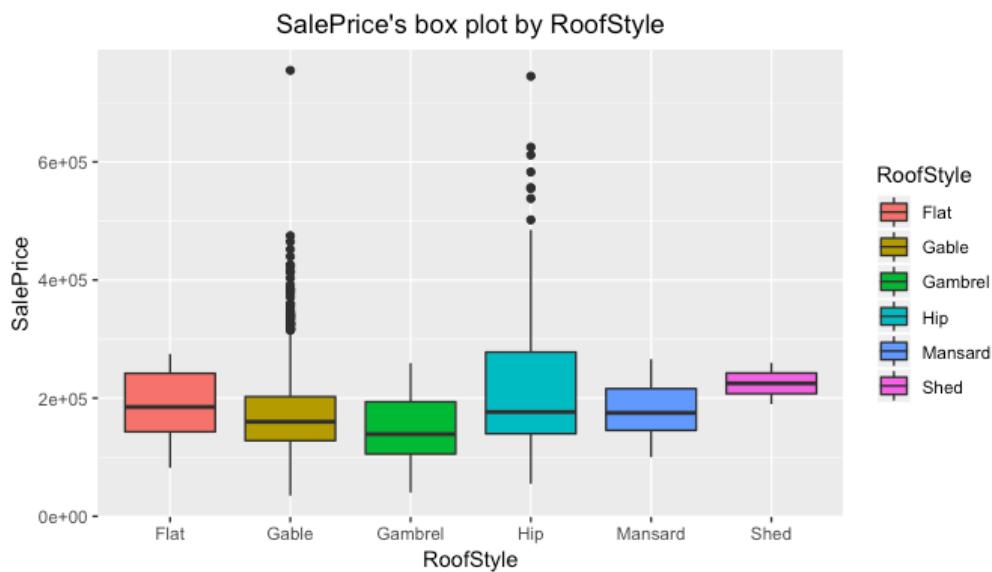
Multiple R-squared: 0.0577, Adjusted R-squared: 0.05446
 F-statistic: 18 on 5 and 1454 DF, p-value: 0.281655305106236

Coefficients:

Simple Linear Model coefficients : RoofStyle

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	194690.00	21425.06	9.09	0.00
RoofStyleGable	-23206.04	21546.77	-1.08	0.28
RoofStyleGambrel	-45780.91	31646.91	-1.45	0.15
RoofStyleHip	24186.93	21906.58	1.10	0.27
RoofStyleMansard	-14121.57	36214.96	-0.39	0.70
RoofStyleShed	30310.00	58674.94	0.52	0.61

3. Visualization:



[RoofMatl]

1. Analysis of Variance

Analysis of Variance Table : RoofMatl

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
RoofMatl	7	2.892493e+11	41321321676	6.73	0

Residuals	1452	8.918662e+12	6142329258	NA	NA
-----------	------	--------------	------------	----	----

2. Simple Linear Model Information

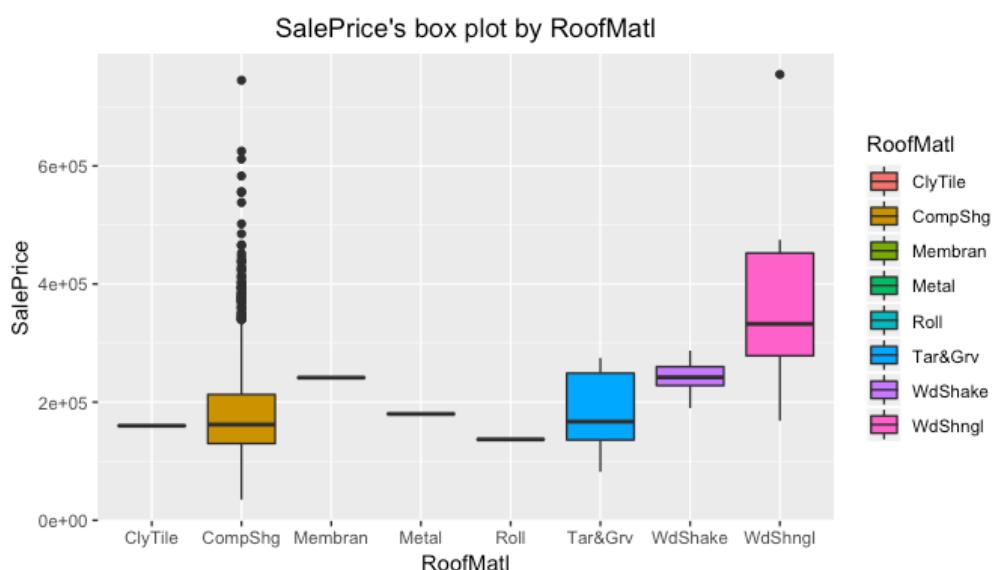
Residual standard error: 78373 on 1452 degrees of freedom
 Multiple R-squared: 0.03141, Adjusted R-squared: 0.02674
 F-statistic: 7 on 7 and 1452 DF, p-value: 0.800615503126109

Coefficients:

Simple Linear Model coefficients : RoofMatl

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	160000.00	78373.01	2.04	0.04
RoofMatlCompShg	19803.68	78400.34	0.25	0.80
RoofMatlMembran	81500.00	110836.18	0.74	0.46
RoofMatlMetal	20000.00	110836.18	0.18	0.86
RoofMatlRoll	-23000.00	110836.18	-0.21	0.84
RoofMatlTar&Grv	25406.36	81857.94	0.31	0.76
RoofMatlWdShake	81400.00	85853.33	0.95	0.34
RoofMatlWdShngl	230250.00	84652.53	2.72	0.01

3. Visualization:



[Exterior1st]

1. Analysis of Variance

Analysis of Variance Table : Exterior1st

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
Exterior1st	14	1.406721e+12	100480098015	18.61	0
Residuals	1445	7.801190e+12	5398747379	NA	NA

2. Simple Linear Model Information

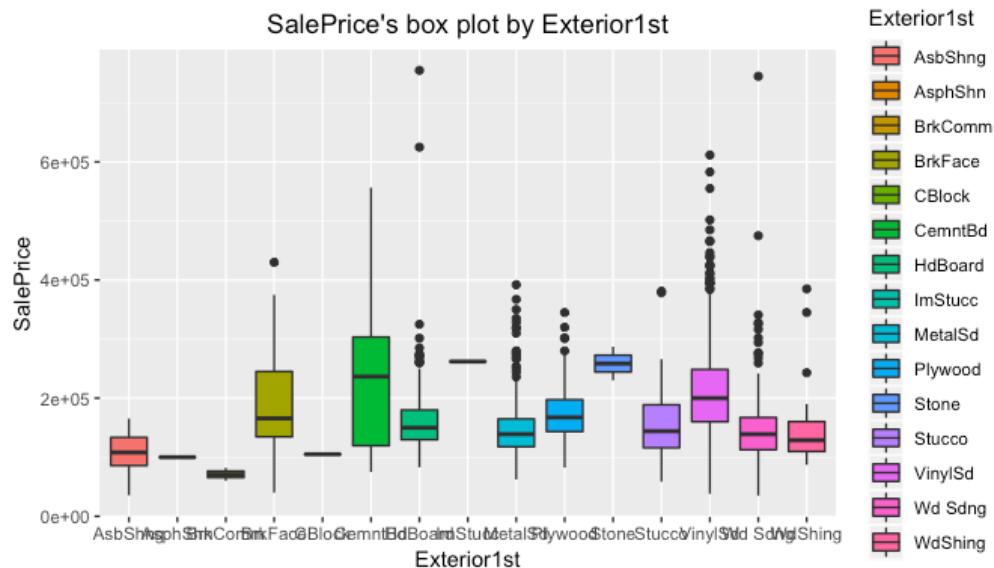
Residual standard error: 73476 on 1445 degrees of freedom
 Multiple R-squared: 0.15277, Adjusted R-squared: 0.14456
 F-statistic: 19 on 14 and 1445 DF, p-value: 0.921871391717156

Coefficients:

Simple Linear Model coefficients : Exterior1st

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	107385.55	16429.77	6.54	0.00
Exterior1stAsphShn	-7385.55	75290.67	-0.10	0.92
Exterior1stBrkComm	-36385.55	54491.39	-0.67	0.50
Exterior1stBrkFace	87187.45	19439.97	4.48	0.00
Exterior1stCBlock	-2385.55	75290.67	-0.03	0.97
Exterior1stCemntBd	124305.11	18932.55	6.57	0.00
Exterior1stHdBoard	55691.90	17153.89	3.25	0.00
Exterior1stImStucc	154614.45	75290.67	2.05	0.04
Exterior1stMetalSd	42036.63	17160.34	2.45	0.01
Exterior1stPlywood	68556.83	17886.47	3.83	0.00
Exterior1stStone	151114.45	54491.39	2.77	0.01
Exterior1stStucco	55604.45	22042.85	2.52	0.01
Exterior1stVinylSd	106347.35	16745.76	6.35	0.00
Exterior1stWd Sdng	42456.10	17208.86	2.47	0.01
Exterior1stWdShing	43269.53	21853.64	1.98	0.05

3. Visualization:



[Exterior2nd]

1. Analysis of Variance

Analysis of Variance Table : Exterior2nd

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
Exterior2nd	15	1.416452e+12	94430114176	17.5	0
Residuals	1444	7.791460e+12	5395747661	NA	NA

2. Simple Linear Model Information

Residual standard error: 73456 on 1444 degrees of freedom
 Multiple R-squared: 0.15383, Adjusted R-squared: 0.14504
 F-statistic: 18 on 15 and 1444 DF, p-value: 0.598704474111426

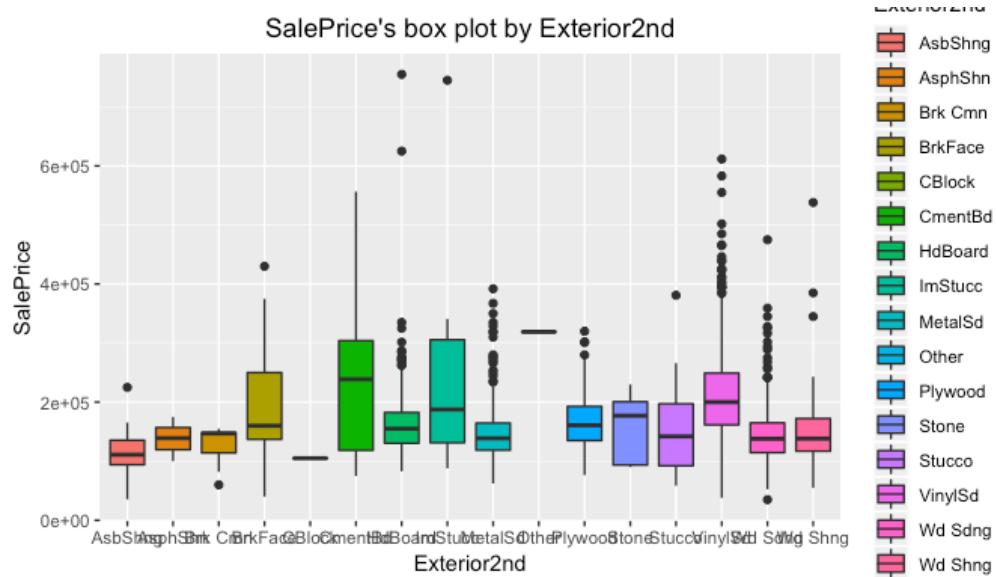
Coefficients:

Simple Linear Model coefficients : Exterior2nd

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	114060.55	16425.21	6.94	0.00
Exterior2ndAsphShn	23939.45	45479.34	0.53	0.60

Exterior2ndBrk Cmn	12653.74	32258.46	0.39	0.69
Exterior2ndBrkFace	81757.45	22036.73	3.71	0.00
Exterior2ndCBlock	-9060.55	75269.75	-0.12	0.90
Exterior2ndCmentBd	116033.28	18966.19	6.12	0.00
Exterior2ndHdBoard	53601.02	17200.40	3.12	0.00
Exterior2ndImStucc	138009.45	28449.29	4.85	0.00
Exterior2ndMetalSd	35742.62	17175.60	2.08	0.04
Exterior2ndOther	204939.45	75269.75	2.72	0.01
Exterior2ndPlywood	54051.84	17543.82	3.08	0.00
Exterior2ndStone	44164.25	36727.88	1.20	0.23
Exterior2ndStucco	41844.60	21847.57	1.92	0.06
Exterior2ndVinylSd	100371.91	16747.93	5.99	0.00
Exterior2ndWd Sdng	34325.52	17238.82	1.99	0.05
Exterior2ndWd Shng	47268.40	20292.38	2.33	0.02

3. Visualization:



[ExterQual]

1. Analysis of Variance

Analysis of variance table : ExterQual

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
ExterQual	3	4.395744e+12	1.465248e+12	443.33	0
Residuals	1456	4.812167e+12	3.305060e+09	NA	NA

2. Simple Linear Model Information

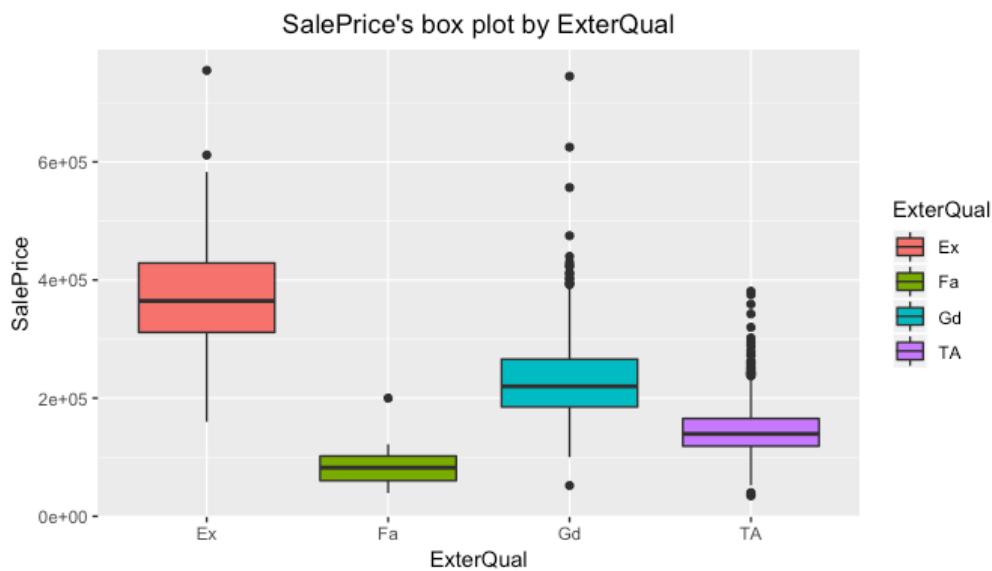
Residual standard error: 57490 on 1456 degrees of freedom
 Multiple R-squared: 0.47739, Adjusted R-squared: 0.47631
 F-statistic: 443 on 3 and 1456 DF, p-value: 4.91251702337871e-54

Coefficients:

Simple Linear Model coefficients : ExterQual

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	367361.0	7972.38	46.08	0
ExterQualFa	-279375.8	17309.95	-16.14	0
ExterQualGd	-135727.5	8386.39	-16.18	0
ExterQualTA	-223019.6	8197.98	-27.20	0

3. Visualization:



[ExterCond]

1. Analysis of Variance

Analysis of Variance Table : ExterCond

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
ExterCond	4	2.174690e+11	54367239271	8.8	0
Residuals	1455	8.990442e+12	6178998198	NA	NA

2. Simple Linear Model Information

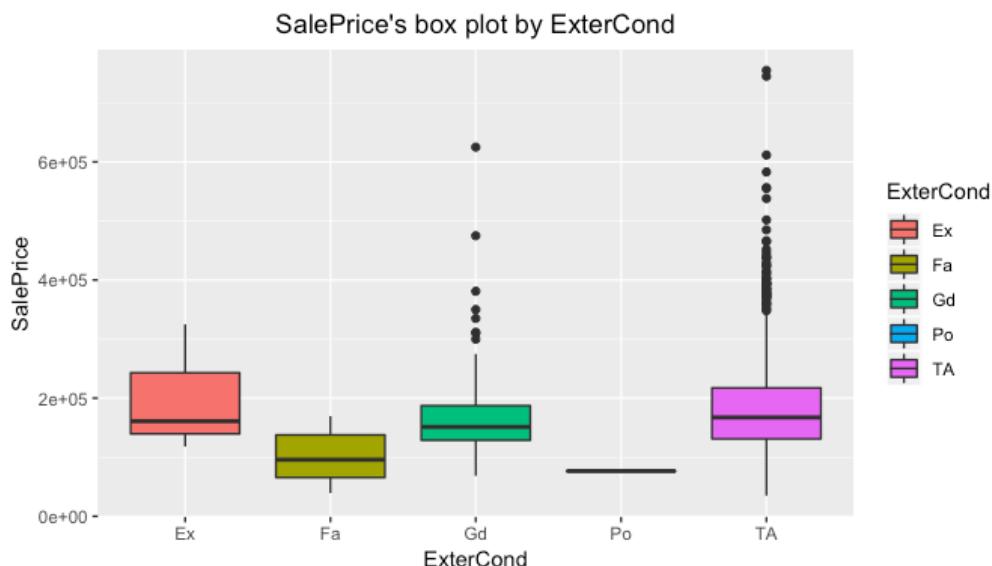
Residual standard error: 78607 on 1455 degrees of freedom
 Multiple R-squared: 0.02362, Adjusted R-squared: 0.02093
 F-statistic: 9 on 4 and 1455 DF, p-value: 0.0388458359298857

Coefficients:

Simple Linear Model coefficients : ExterCond

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	201333.33	45383.54	4.44	0.00
ExterCondFa	-98738.19	47752.95	-2.07	0.04
ExterCondGd	-32435.76	45847.44	-0.71	0.48
ExterCondPo	-124833.33	90767.09	-1.38	0.17
ExterCondTA	-17298.44	45436.61	-0.38	0.70

3. Visualization:



[Foundation]

1. Analysis of Variance

Analysis of Variance Table : Foundation

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
Foundation	5	2.360618e+12	472123502058	100.25	0
Residuals	1454	6.847294e+12	4709280484	NA	NA

2. Simple Linear Model Information

Residual standard error: 68624 on 1454 degrees of freedom
 Multiple R-squared: 0.25637, Adjusted R-squared: 0.25381
 F-statistic: 100 on 5 and 1454 DF, p-value: 0.00550023467582817

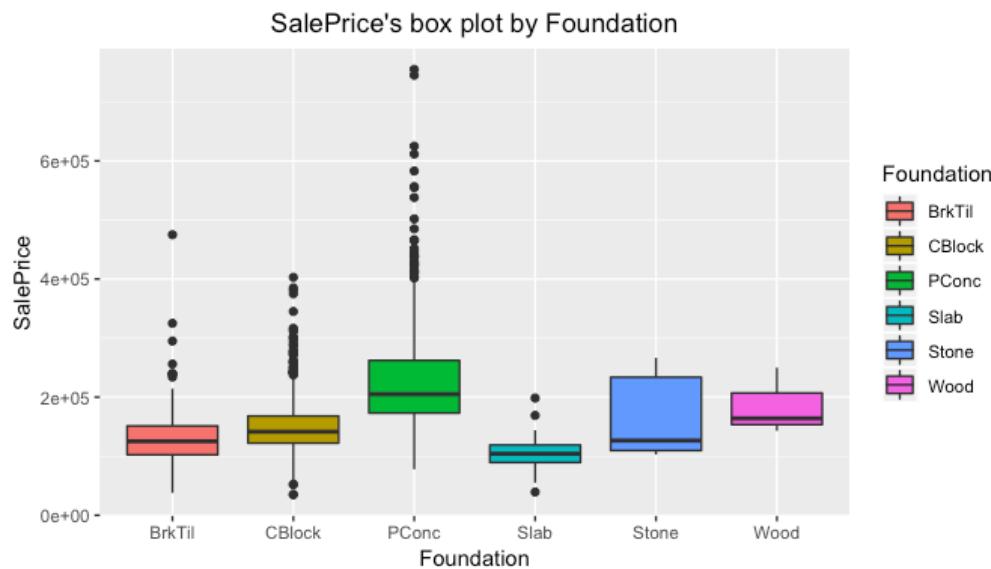
Coefficients:

Simple Linear Model coefficients : Foundation

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	132291.08	5679.38	23.29	0.00
FoundationCBlock	17514.64	6299.46	2.78	0.01
FoundationPConc	92939.37	6287.61	14.78	0.00
FoundationSlab	-24925.45	15115.40	-1.65	0.10
FoundationStone	33668.09	28585.58	1.18	0.24

FoundationWood	53375.59	40025.19	1.33	0.18
----------------	----------	----------	------	------

3. Visualization:



[Heating]

1. Analysis of Variance

Analysis of Variance Table : Heating

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
Heating	5	1.329359e+11	26587172586	4.26	0
Residuals	1454	9.074975e+12	6241386157	NA	NA

2. Simple Linear Model Information

Residual standard error: 79002 on 1454 degrees of freedom
 Multiple R-squared: 0.01444, Adjusted R-squared: 0.01105
 F-statistic: 4 on 5 and 1454 DF, p-value: 0.166015538248806

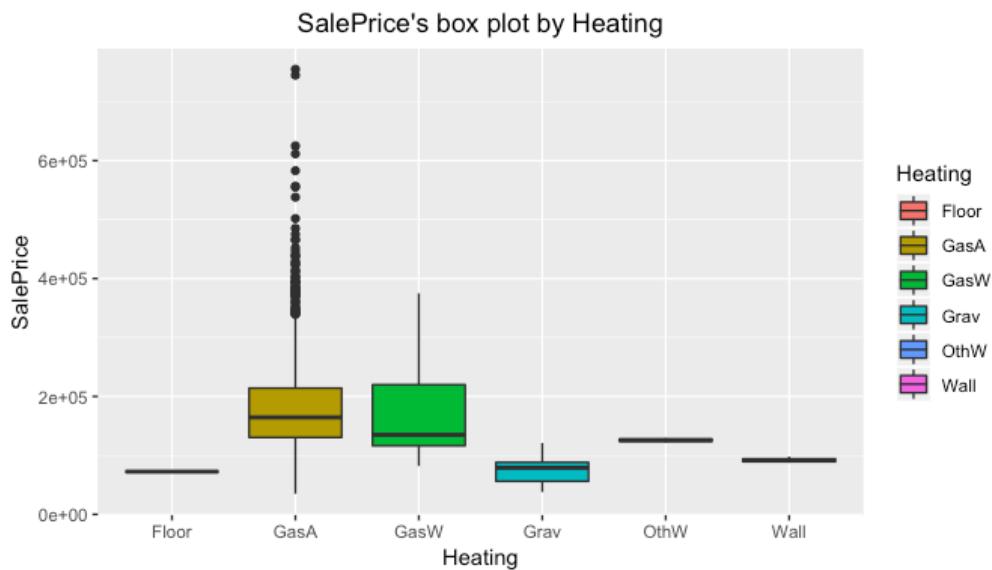
Coefficients:

Simple Linear Model coefficients : Heating

Estimate	Std. Error	t value	Pr(> t)
----------	------------	---------	----------

(Intercept)	72500.00	79002.44	0.92	0.36
HeatingGasA	109521.20	79030.10	1.39	0.17
HeatingGasW	94132.17	81167.30	1.16	0.25
HeatingGrav	2771.43	84457.17	0.03	0.97
HeatingOthW	53250.00	96757.84	0.55	0.58
HeatingWall	19600.00	88327.42	0.22	0.82

3. Visualization:



[HeatingQC]

1. Analysis of Variance

Analysis of Variance Table : HeatingQC

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
HeatingQC	4	1.800151e+12	450037784872	88.39	0
Residuals	1455	7.407760e+12	5091244120	NA	NA

2. Simple Linear Model Information

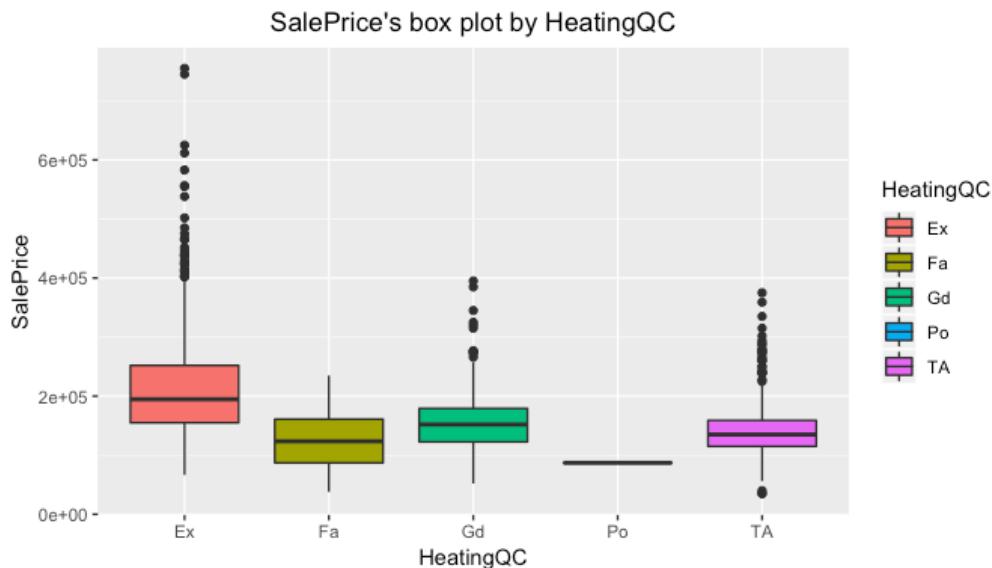
Residual standard error: 71353 on 1455 degrees of freedom
 Multiple R-squared: 0.1955, Adjusted R-squared: 0.19329
 F-statistic: 88 on 4 and 1455 DF, p-value: 1.38753554741728e-17

Coefficients:

Simple Linear Model coefficients : HeatingQC

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	214914.43	2621.22	81.99	0.00
HeatingQCFa	-90994.94	10524.91	-8.65	0.00
HeatingQCGd	-58055.56	5291.15	-10.97	0.00
HeatingQCPo	-127914.43	71401.08	-1.79	0.07
HeatingQCTA	-72551.55	4332.00	-16.75	0.00

3. Visualization:



[CentralAir]

1. Analysis of Variance

Analysis of Variance Table : CentralAir

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
CentralAir	1	5.816255e+11	581625508784	98.31	0
Residuals	1458	8.626286e+12	5916519771	NA	NA

2. Simple Linear Model Information

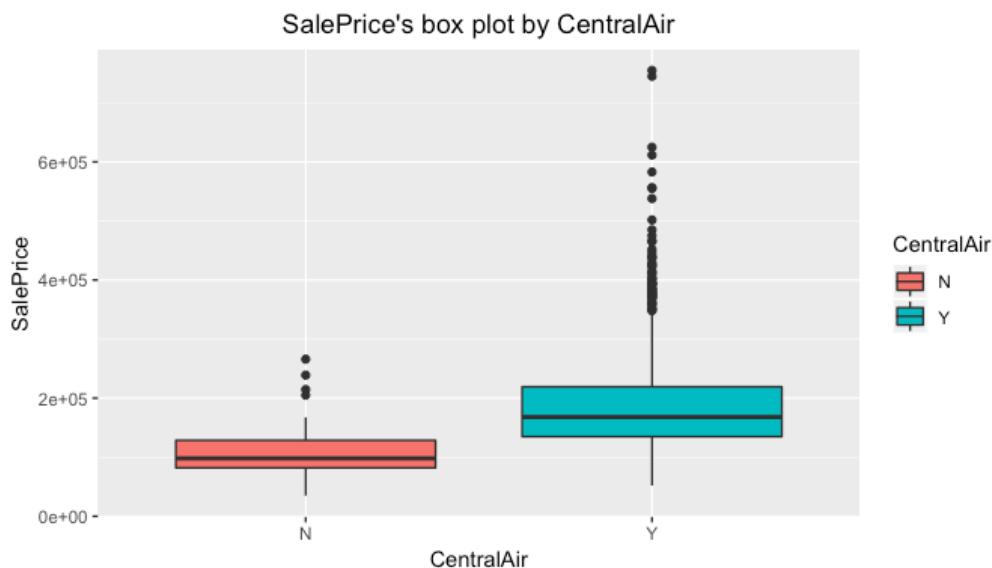
Residual standard error: 76919 on 1458 degrees of freedom
 Multiple R-squared: 0.06317, Adjusted R-squared: 0.06252
 F-statistic: 98 on 1 and 1458 DF, p-value: 1.80950615592689e-22

Coefficients:

Simple Linear Model coefficients : CentralAir

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	105264.07	7891.71	13.34	0
CentralAirY	80922.64	8161.72	9.91	0

3. Visualization:



[KitchenQual]

1. Analysis of Variance

Analysis of Variance Table : KitchenQual

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
KitchenQual	3	4.204320e+12	1.401440e+12	407.81	0
Residuals	1456	5.003592e+12	3.436533e+09	NA	NA

2. Simple Linear Model Information

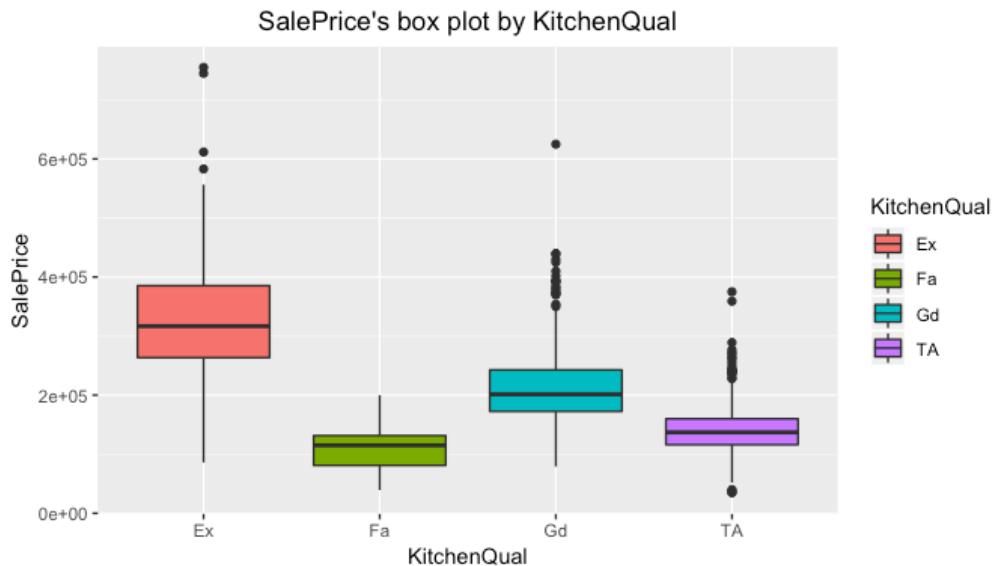
Residual standard error: 58622 on 1456 degrees of freedom
 Multiple R-squared: 0.4566, Adjusted R-squared: 0.45548
 F-statistic: 408 on 3 and 1456 DF, p-value: 7.77473123170266e-80

Coefficients:

Simple Linear Model coefficients : KitchenQual

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	328554.7	5862.19	56.05	0
KitchenQualFa	-222989.5	11067.14	-20.15	0
KitchenQualGd	-116438.6	6342.69	-18.36	0
KitchenQualTA	-188592.2	6248.27	-30.18	0

3. Visualization:



[Functional]

1. Analysis of Variance

Analysis of Variance Table : Functional

Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
1	1234567890	1234567890	1234567890	1.23456789e-10

Functional	6	1.517499e+11	25291655080	4.06	0
Residuals	1453	9.056161e+12	6232733244	NA	NA

2. Simple Linear Model Information

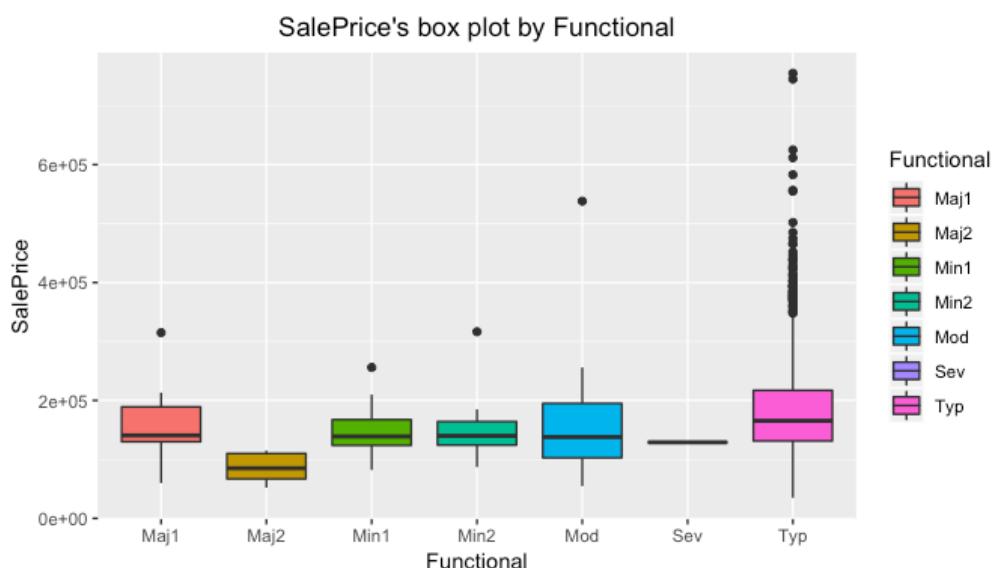
Residual standard error: 78948 on 1453 degrees of freedom
 Multiple R-squared: 0.01648, Adjusted R-squared: 0.01242
 F-statistic: 4 on 6 and 1453 DF, p-value: 0.0977627802903312

Coefficients:

Simple Linear Model coefficients : Functional

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	153948.14	21099.65	7.30	0.00
FunctionalMaj2	-68148.14	41130.79	-1.66	0.10
FunctionalMin1	-7562.66	25421.47	-0.30	0.77
FunctionalMin2	-9707.50	25070.12	-0.39	0.70
FunctionalMod	14445.19	29337.87	0.49	0.62
FunctionalSev	-24948.14	81718.59	-0.31	0.76
FunctionalTyp	29481.00	21207.97	1.39	0.16

3. Visualization:



[PavedDrive]

1. Analysis of Variance

Analysis of Variance Table : PavedDrive

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
PavedDrive	2	5.021970e+11	251098491341	42.02	0
Residuals	1457	8.705714e+12	5975095643	NA	NA

2. Simple Linear Model Information

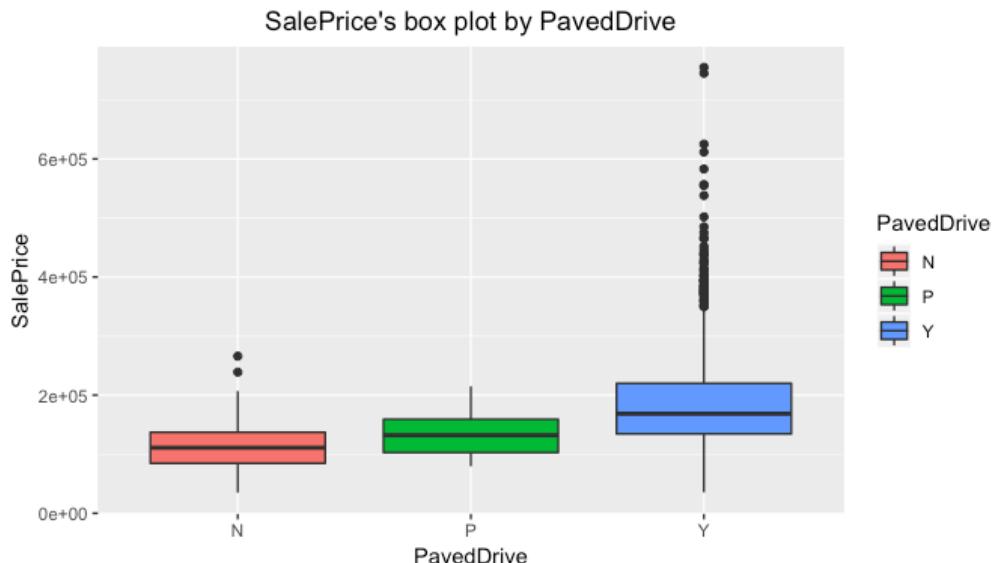
```
Residual standard error: 77299 on 1457 degrees of freedom
Multiple R-squared:  0.05454, Adjusted R-squared:  0.05324
F-statistic: 42 on 2 and 1457 DF,  p-value: 0.288843004907257
```

Coefficients:

Simple Linear Model coefficients : PavedDrive

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	115039.12	8148.00	14.12	0.00
PavedDriveP	17290.88	16296.01	1.06	0.29
PavedDriveY	71394.85	8417.18	8.48	0.00

3. Visualization:



[SaleType]

1. Analysis of Variance

Analysis of Variance Table : SaleType

	Df	Sum.Sq	Mean.Sq	F.value	Pr.F.
SaleType	8	1.264131e+12	158016375931	28.86	0
Residuals	1451	7.943780e+12	5474693540	NA	NA

2. Simple Linear Model Information

Residual standard error: 73991 on 1451 degrees of freedom
 Multiple R-squared: 0.13729, Adjusted R-squared: 0.13253
 F-statistic: 29 on 8 and 1451 DF, p-value: 0.0190502552472954

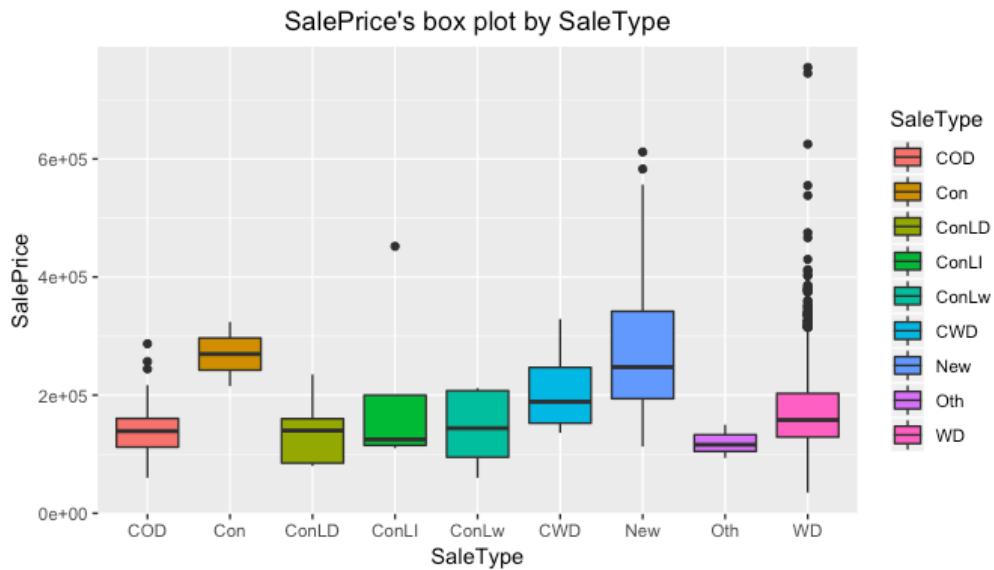
Coefficients:

Simple Linear Model coefficients : SaleType

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	143973.26	11283.55	12.76	0.00
SaleTypeCon	125626.74	53522.57	2.35	0.02
SaleTypeConLD	-5192.37	27122.27	-0.19	0.85
SaleTypeConLI	56416.74	34960.79	1.61	0.11
SaleTypeConLw	-273.26	34960.79	-0.01	0.99

SaleTypeCWD	66626.74	38678.05	1.72	0.09
SaleTypeNew	130972.16	13122.23	9.98	0.00
SaleTypeOth	-24123.26	44183.89	-0.55	0.59
SaleTypeWD	29428.58	11473.42	2.56	0.01

3. Visualization:



[SaleCondition]

1. Analysis of Variance

Analysis of Variance Table : SaleCondition

	Df	Sum.Sq	Mean.Sq	F.value	Pr..F.
SaleCondition	5	1.247649e+12	249529750618	45.58	0
Residuals	1454	7.960263e+12	5474733550	NA	NA

2. Simple Linear Model Information

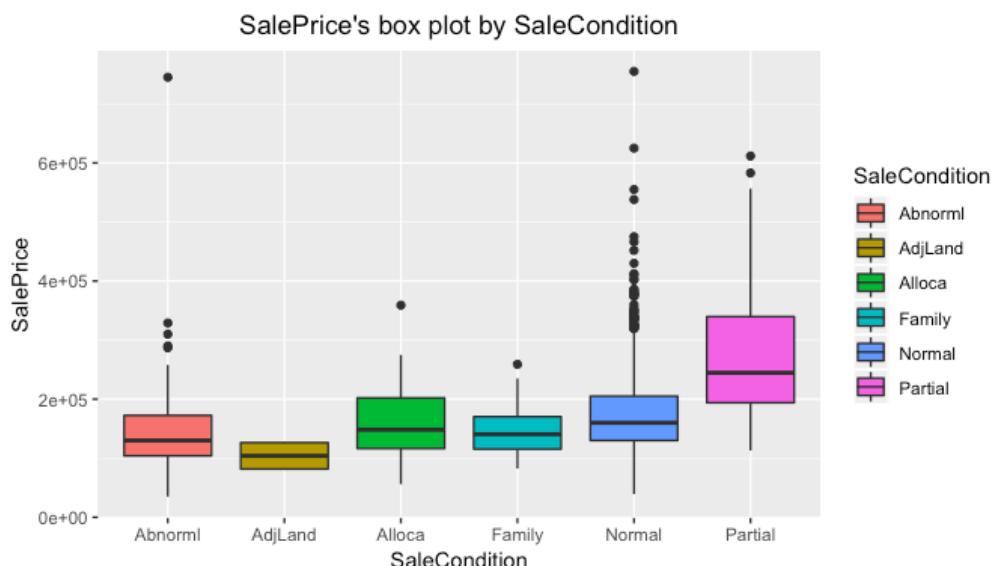
Residual standard error: 73991 on 1454 degrees of freedom
 Multiple R-squared: 0.1355, Adjusted R-squared: 0.13252
 F-statistic: 46 on 5 and 1454 DF, p-value: 0.261164845241487

Coefficients:

Simple Linear Model coefficients : SaleCondition

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	146526.62	7362.42	19.90	0.00
SaleConditionAdjLand	-42401.62	37721.20	-1.12	0.26
SaleConditionAllocac	20850.79	22592.77	0.92	0.36
SaleConditionFamily	3073.38	18109.17	0.17	0.87
SaleConditionNormal	28675.60	7666.50	3.74	0.00
SaleConditionPartial	125765.13	9899.65	12.70	0.00

3. Visualization:



4.2 Grouped Relationship Between Variables

4.2.1 Grouped Correlation Coefficient

Numerical target variables are not supported.

4.2.2 Grouped Correlation Plot of Numerical Variables

Numerical target variables are not supported.