

```
In [437]: %matplotlib inline

import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

import statsmodels.api as sm
from scipy.stats import skew
from sklearn.cross_validation import train_test_split
from sklearn.cross_validation import KFold
```

```
In [438]: #Reading data
data = pd.read_csv('./data/train.csv')
```

## Data Exploration

### Basic Statistics & Observations

```
In [439]: #Read first 5 lines
data.head()
```

```
Out[439]:
```

	Id	MSSubClass	MSZoning	LotFrontage	LotArea	Street	Alley	LotShape	LandConto
0	1	60	RL	65.0	8450	Pave	NaN	Reg	Lvl
1	2	20	RL	80.0	9600	Pave	NaN	Reg	Lvl
2	3	60	RL	68.0	11250	Pave	NaN	IR1	Lvl
3	4	70	RL	60.0	9550	Pave	NaN	IR1	Lvl
4	5	60	RL	84.0	14260	Pave	NaN	IR1	Lvl

5 rows × 81 columns

```
In [440]: #general shape of the data
data.shape
```

```
Out[440]: (1460, 81)
```

```
In [441]: #Basic Statistics about all the *quantitative* columns
data.describe()
```

```
/Users/tiffanychien/anaconda/lib/python3.5/site-packages/numpy/lib/function_base.py:3834: RuntimeWarning: Invalid value encountered in percentile
RuntimeWarning)
```

Out[441]:

	Id	MSSubClass	LotFrontage	LotArea	OverallQual	OverallCond
<b>count</b>	1460.000000	1460.000000	1201.000000	1460.000000	1460.000000	1460.000000
<b>mean</b>	730.500000	56.897260	70.049958	10516.828082	6.099315	5.575342
<b>std</b>	421.610009	42.300571	24.284752	9981.264932	1.382997	1.112799
<b>min</b>	1.000000	20.000000	21.000000	1300.000000	1.000000	1.000000
<b>25%</b>	365.750000	20.000000	NaN	7553.500000	5.000000	5.000000
<b>50%</b>	730.500000	50.000000	NaN	9478.500000	6.000000	5.000000
<b>75%</b>	1095.250000	70.000000	NaN	11601.500000	7.000000	6.000000
<b>max</b>	1460.000000	190.000000	313.000000	215245.000000	10.000000	9.000000

8 rows × 38 columns

```
In [442]: #Basic Statistics about all the *categorical* columns
data.describe(include = ['O'])
```

Out[442]:

	MSZoning	Street	Alley	LotShape	LandContour	Utilities	LotConfig	LandSlope
<b>count</b>	1460	1460	91	1460	1460	1460	1460	1460
<b>unique</b>	5	2	2	4	4	2	5	3
<b>top</b>	RL	Pave	Grvl	Reg	Lvl	AllPub	Inside	Gtl
<b>freq</b>	1151	1454	50	925	1311	1459	1052	1382

4 rows × 43 columns

```
In [443]: #Correlations  
data.corr()
```

Out[443]:

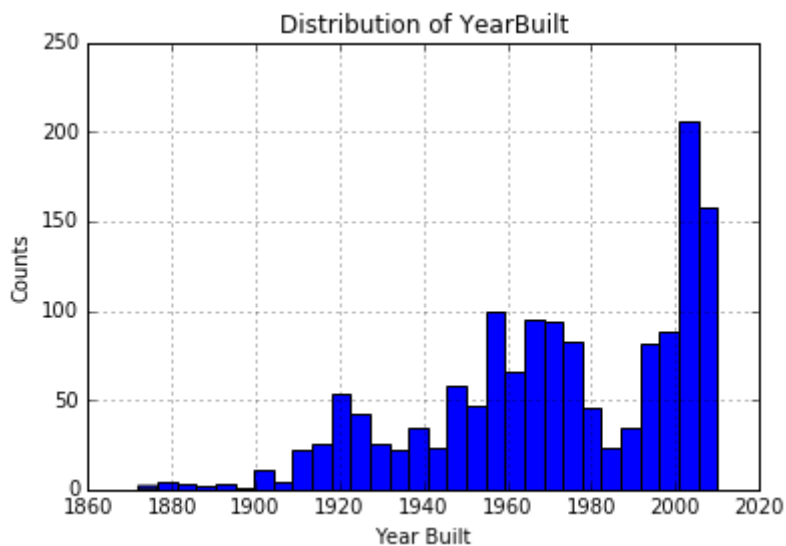
	<b>Id</b>	<b>MSSubClass</b>	<b>LotFrontage</b>	<b>LotArea</b>	<b>OverallQual</b>	<b>OverallCon</b>
<b>Id</b>	1.000000	0.011156	-0.010601	-0.033226	-0.028365	0.012609
<b>MSSubClass</b>	0.011156	1.000000	-0.386347	-0.139781	0.032628	-0.059316
<b>LotFrontage</b>	-0.010601	-0.386347	1.000000	0.426095	0.251646	-0.059213
<b>LotArea</b>	-0.033226	-0.139781	0.426095	1.000000	0.105806	-0.005636
<b>OverallQual</b>	-0.028365	0.032628	0.251646	0.105806	1.000000	-0.091932
<b>OverallCond</b>	0.012609	-0.059316	-0.059213	-0.005636	-0.091932	1.000000
<b>YearBuilt</b>	-0.012713	0.027850	0.123349	0.014228	0.572323	-0.375983
<b>YearRemodAdd</b>	-0.021998	0.040581	0.088866	0.013788	0.550684	0.073741
<b>MasVnrArea</b>	-0.050298	0.022936	0.193458	0.104160	0.411876	-0.128101
<b>BsmtFinSF1</b>	-0.005024	-0.069836	0.233633	0.214103	0.239666	-0.046231
<b>BsmtFinSF2</b>	-0.005968	-0.065649	0.049900	0.111170	-0.059119	0.040229
<b>BsmtUnfSF</b>	-0.007940	-0.140759	0.132644	-0.002618	0.308159	-0.136841
<b>TotalBsmtSF</b>	-0.015415	-0.238518	0.392075	0.260833	0.537808	-0.171098
<b>1stFlrSF</b>	0.010496	-0.251758	0.457181	0.299475	0.476224	-0.144203
<b>2ndFlrSF</b>	0.005590	0.307886	0.080177	0.050986	0.295493	0.028942
<b>LowQualFinSF</b>	-0.044230	0.046474	0.038469	0.004779	-0.030429	0.025494
<b>GrLivArea</b>	0.008273	0.074853	0.402797	0.263116	0.593007	-0.079686
<b>BsmtFullBath</b>	0.002289	0.003491	0.100949	0.158155	0.111098	-0.054942
<b>BsmtHalfBath</b>	-0.020155	-0.002333	-0.007234	0.048046	-0.040150	0.117821
<b>FullBath</b>	0.005587	0.131608	0.198769	0.126031	0.550600	-0.194149
<b>HalfBath</b>	0.006784	0.177354	0.053532	0.014259	0.273458	-0.060769
<b>BedroomAbvGr</b>	0.037719	-0.023438	0.263170	0.119690	0.101676	0.012980
<b>KitchenAbvGr</b>	0.002951	0.281721	-0.006069	-0.017784	-0.183882	-0.087001
<b>TotRmsAbvGrd</b>	0.027239	0.040380	0.352096	0.190015	0.427452	-0.057583
<b>Fireplaces</b>	-0.019772	-0.045569	0.266639	0.271364	0.396765	-0.023820
<b>GarageYrBlt</b>	0.000072	0.085072	0.070250	-0.024947	0.547766	-0.324297
<b>GarageCars</b>	0.016570	-0.040110	0.285691	0.154871	0.600671	-0.185758
<b>GarageArea</b>	0.017634	-0.098672	0.344997	0.180403	0.562022	-0.151521
<b>WoodDeckSF</b>	-0.029643	-0.012579	0.088521	0.171698	0.238923	-0.003334
<b>OpenPorchSF</b>	-0.000477	-0.006100	0.151972	0.084774	0.308819	-0.032589
<b>EnclosedPorch</b>	0.002889	-0.012037	0.010700	-0.018340	-0.113937	0.070356
<b>3SsnPorch</b>	-0.046635	-0.043825	0.070029	0.020423	0.030371	0.025504

	Id	MSSubClass	LotFrontage	LotArea	OverallQual	OverallCon
ScreenPorch	0.001330	-0.026030	0.041383	0.043160	0.064886	0.054811
PoolArea	0.057044	0.008283	0.206167	0.077672	0.065166	-0.001985
MiscVal	-0.006242	-0.007683	0.003368	0.038068	-0.031406	0.068777
MoSold	0.021172	-0.013585	0.011200	0.001205	0.070815	-0.003511
YrSold	0.000712	-0.021407	0.007450	-0.014261	-0.027347	0.043950
SalePrice	-0.021917	-0.084284	0.351799	0.263843	0.790982	-0.077856

38 rows × 38 columns

```
In [466]: # some graphs of different variables
data['YearBuilt'].hist(bins=30)
plt.title('Distribution of YearBuilt')
plt.xlabel('Year Built')
plt.ylabel('Counts')
```

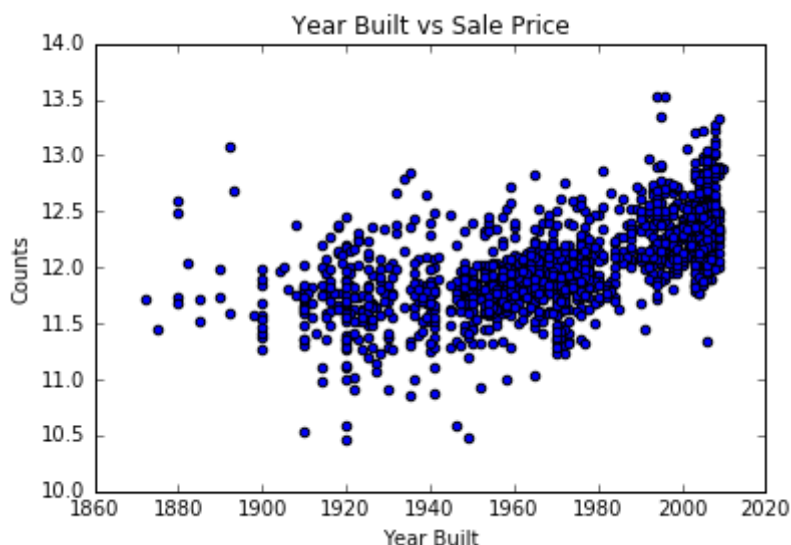
Out[466]: <matplotlib.text.Text at 0x11ba49710>



There are more new houses in the dataset than old houses, other than a dip around 1990.

```
In [468]: plt.scatter(data['YearBuilt'], data['SalePrice'])  
plt.title('Year Built vs Sale Price')  
plt.xlabel('Year Built')  
plt.ylabel('Sale Price')
```

```
Out[468]: <matplotlib.text.Text at 0x11eb84908>
```



There seems to be a somewhat weak positive correlation between the year the house was built and its price, with quite a few outliers.

## Data Cleaning & Feature Engineering

```
In [444]: #log-transforming response variables  
data['SalePrice'] = np.log1p(data['SalePrice'])
```

```
In [445]: #Splitting categorical variables into dummies  
data = pd.get_dummies(data, drop_first=True)  
  
data.shape
```

```
Out[445]: (1460, 247)
```

```
In [1]: #Filling NAs with column means
data = data.fillna(data.mean())

#And checking if any NAs exist after the code running
data.isnull().values.any()
```

```
-----
-----
NameError                                Traceback (most recent call 1
ast)
<ipython-input-1-756ca97d35b1> in <module>()
      1 #Filling NAs with column means
----> 2 data = data.fillna(data.mean())
      3
      4 #And checking if any NAs exist after the code running
      5 data.isnull().values.any()

NameError: name 'data' is not defined
```

## Model Building

```
In [447]: #Splitting datasets into train & test
local_train, local_test = train_test_split(data, test_size = 0.2, random
_state=123)
```

```
In [448]: #Removing SalesPrices (is the predicting y-variable)

local_train_y = local_train['SalePrice']
local_train_x = local_train.drop('SalePrice', axis = 1)

local_test_y = local_test['SalePrice']
local_test_x = local_test.drop('SalePrice', axis = 1)
```

```
In [449]: #General statistics to make sure everything is right
local_train.shape[0] + local_test.shape[0] == data.shape[0]
```

Out[449]: True

```
In [450]: #General statistics to make sure everything is right
local_train.shape[1] == local_test.shape[1] == data.shape[1]
```

Out[450]: True

```
In [451]: #Fitting model  
clf = sm.OLS(local_train_y,local_train_x)  
result = clf.fit()  
result.summary()
```



Out[451]: OLS Regression Results

<b>Dep. Variable:</b>	SalePrice	<b>R-squared:</b>	1.000
<b>Model:</b>	OLS	<b>Adj. R-squared:</b>	1.000
<b>Method:</b>	Least Squares	<b>F-statistic:</b>	6.203e+04
<b>Date:</b>	Sun, 02 Oct 2016	<b>Prob (F-statistic):</b>	0.00
<b>Time:</b>	23:10:46	<b>Log-Likelihood:</b>	1089.3
<b>No. Observations:</b>	1168	<b>AIC:</b>	-1701.
<b>Df Residuals:</b>	929	<b>BIC:</b>	-490.6
<b>Df Model:</b>	239		
<b>Covariance Type:</b>	nonrobust		

	coef	std err	t	P> t	[95.0% Conf. Int.]
<b>Id</b>	-6.255e-06	8.34e-06	-0.750	0.454	-2.26e-05 1.01e-05
<b>MSSubClass</b>	-0.0003	0.000	-0.662	0.508	-0.001 0.001
<b>LotFrontage</b>	0.0005	0.000	2.046	0.041	2.16e-05 0.001
<b>LotArea</b>	2.351e-06	6.09e-07	3.858	0.000	1.16e-06 3.55e-06
<b>OverallQual</b>	0.0373	0.006	6.760	0.000	0.026 0.048
<b>OverallCond</b>	0.0368	0.005	7.909	0.000	0.028 0.046
<b>YearBuilt</b>	0.0017	0.000	4.227	0.000	0.001 0.002
<b>YearRemodAdd</b>	0.0007	0.000	2.502	0.013	0.000 0.001
<b>MasVnrArea</b>	1.891e-06	3.21e-05	0.059	0.953	-6.12e-05 6.5e-05
<b>BsmtFinSF1</b>	5.557e-05	1.5e-05	3.699	0.000	2.61e-05 8.5e-05
<b>BsmtFinSF2</b>	4.926e-05	2.96e-05	1.662	0.097	-8.91e-06 0.000
<b>BsmtUnfSF</b>	-9.627e-06	1.45e-05	-0.665	0.506	-3.8e-05 1.88e-05
<b>TotalBsmtSF</b>	9.52e-05	1.92e-05	4.958	0.000	5.75e-05 0.000
<b>1stFlrSF</b>	7.979e-05	3.37e-05	2.370	0.018	1.37e-05 0.000
<b>2ndFlrSF</b>	6.321e-05	2.93e-05	2.158	0.031	5.71e-06 0.000
<b>LowQualFinSF</b>	3.398e-06	7.5e-05	0.045	0.964	-0.000 0.000
<b>GrLivArea</b>	0.0001	3.09e-05	4.735	0.000	8.57e-05 0.000
<b>BsmtFullBath</b>	0.0263	0.010	2.532	0.012	0.006 0.047
<b>BsmtHalfBath</b>	0.0029	0.017	0.173	0.862	-0.030 0.036
<b>FullBath</b>	0.0113	0.012	0.973	0.331	-0.011 0.034
<b>HalfBath</b>	0.0150	0.011	1.340	0.181	-0.007 0.037

<b>BedroomAbvGr</b>	0.0053	0.007	0.722	0.471	-0.009 0.020
<b>KitchenAbvGr</b>	-0.0577	0.037	-1.553	0.121	-0.131 0.015
<b>TotRmsAbvGrd</b>	0.0088	0.005	1.717	0.086	-0.001 0.019
<b>Fireplaces</b>	0.0176	0.013	1.383	0.167	-0.007 0.042
<b>GarageYrBlt</b>	-0.0004	0.000	-1.081	0.280	-0.001 0.000
<b>GarageCars</b>	0.0231	0.012	1.870	0.062	-0.001 0.047
<b>GarageArea</b>	0.0001	4.29e-05	2.656	0.008	2.98e-05 0.000
<b>WoodDeckSF</b>	0.0001	3.06e-05	3.509	0.000	4.73e-05 0.000
<b>OpenPorchSF</b>	6.613e-05	6.39e-05	1.035	0.301	-5.93e-05 0.000
<b>EnclosedPorch</b>	2.88e-05	6.85e-05	0.420	0.674	-0.000 0.000
<b>3SsnPorch</b>	0.0001	0.000	1.285	0.199	-7.68e-05 0.000
<b>ScreenPorch</b>	0.0003	6.77e-05	3.787	0.000	0.000 0.000
<b>PoolArea</b>	0.0001	0.000	0.690	0.490	-0.000 0.000
<b>MiscVal</b>	3.64e-06	7.13e-06	0.511	0.610	-1.03e-05 1.76e-05
<b>MoSold</b>	0.0008	0.001	0.623	0.533	-0.002 0.003
<b>YrSold</b>	0.0015	0.001	3.011	0.003	0.001 0.003
<b>MSZoning_FV</b>	0.4914	0.061	8.069	0.000	0.372 0.611
<b>MSZoning_RH</b>	0.4877	0.060	8.155	0.000	0.370 0.605
<b>MSZoning_RL</b>	0.4702	0.051	9.191	0.000	0.370 0.571
<b>MSZoning_RM</b>	0.4413	0.048	9.227	0.000	0.347 0.535
<b>Street_Pave</b>	0.1054	0.062	1.695	0.090	-0.017 0.228
<b>Alley_Pave</b>	0.0472	0.025	1.910	0.056	-0.001 0.096
<b>LotShape_IR2</b>	0.0302	0.023	1.305	0.192	-0.015 0.076
<b>LotShape_IR3</b>	0.0128	0.050	0.255	0.798	-0.085 0.111
<b>LotShape_Reg</b>	0.0036	0.009	0.415	0.678	-0.014 0.021
<b>LandContour_HLS</b>	0.0439	0.028	1.585	0.113	-0.010 0.098
<b>LandContour_Low</b>	-0.0321	0.034	-0.947	0.344	-0.099 0.034
<b>LandContour_Lvl</b>	0.0237	0.020	1.189	0.235	-0.015 0.063
<b>Utilities_NoSeWa</b>	-0.2010	0.128	-1.571	0.117	-0.452 0.050
<b>LotConfig_CulDSac</b>	0.0350	0.018	1.923	0.055	-0.001 0.071
<b>LotConfig_FR2</b>	-0.0297	0.021	-1.385	0.166	-0.072 0.012
<b>LotConfig_FR3</b>	-0.0636	0.091	-0.702	0.483	-0.241 0.114
<b>LotConfig_Inside</b>	-0.0097	0.010	-1.005	0.315	-0.029 0.009

<b>LandSlope_Mod</b>	0.0643	0.023	2.844	0.005	0.020 0.109
<b>LandSlope_Sev</b>	-0.1578	0.066	-2.400	0.017	-0.287 -0.029
<b>Neighborhood_Blueste</b>	-0.1314	0.117	-1.121	0.263	-0.361 0.099
<b>Neighborhood_BrDale</b>	-0.0952	0.058	-1.648	0.100	-0.208 0.018
<b>Neighborhood_BrkSide</b>	-0.0300	0.051	-0.585	0.558	-0.131 0.071
<b>Neighborhood_ClearCr</b>	-0.0094	0.050	-0.186	0.852	-0.108 0.090
<b>Neighborhood_CollgCr</b>	-0.0509	0.039	-1.300	0.194	-0.128 0.026
<b>Neighborhood_Crawfor</b>	0.0653	0.046	1.417	0.157	-0.025 0.156
<b>Neighborhood_Edwards</b>	-0.1064	0.043	-2.458	0.014	-0.191 -0.021
<b>Neighborhood_Gilbert</b>	-0.0558	0.041	-1.347	0.178	-0.137 0.026
<b>Neighborhood_IDOTRR</b>	-0.0489	0.057	-0.853	0.394	-0.162 0.064
<b>Neighborhood_MeadowV</b>	-0.2280	0.059	-3.895	0.000	-0.343 -0.113
<b>Neighborhood_Mitchel</b>	-0.0991	0.044	-2.227	0.026	-0.186 -0.012
<b>Neighborhood_NAmes</b>	-0.0668	0.042	-1.574	0.116	-0.150 0.016
<b>Neighborhood_NPkVill</b>	-0.0332	0.076	-0.435	0.664	-0.183 0.117
<b>Neighborhood_NWAmes</b>	-0.0718	0.043	-1.651	0.099	-0.157 0.014
<b>Neighborhood_NoRidge</b>	-0.0088	0.046	-0.193	0.847	-0.099 0.081
<b>Neighborhood_NridgHt</b>	0.0747	0.041	1.826	0.068	-0.006 0.155
<b>Neighborhood_OldTown</b>	-0.0702	0.051	-1.362	0.173	-0.171 0.031
<b>Neighborhood_SWISU</b>	-0.0284	0.051	-0.561	0.575	-0.128 0.071
<b>Neighborhood_Sawyer</b>	-0.0726	0.044	-1.646	0.100	-0.159 0.014
<b>Neighborhood_SawyerW</b>	-0.0481	0.042	-1.148	0.251	-0.130 0.034
<b>Neighborhood_Somerst</b>	-0.0057	0.048	-0.119	0.905	-0.100 0.089
<b>Neighborhood_StoneBr</b>	0.1035	0.044	2.345	0.019	0.017 0.190
<b>Neighborhood_Timber</b>	-0.0283	0.043	-0.656	0.512	-0.113 0.056
<b>Neighborhood_Veenker</b>	0.0177	0.054	0.330	0.742	-0.088 0.123
<b>Condition1_Feedr</b>	0.0687	0.028	2.457	0.014	0.014 0.124
<b>Condition1_Norm</b>	0.1108	0.023	4.752	0.000	0.065 0.157
<b>Condition1_PosA</b>	0.0708	0.052	1.350	0.177	-0.032 0.174
<b>Condition1_PosN</b>	0.1054	0.041	2.579	0.010	0.025 0.186
<b>Condition1_RRAe</b>	-0.0006	0.047	-0.012	0.991	-0.093 0.092
<b>Condition1_RRAn</b>	0.0839	0.036	2.306	0.021	0.012 0.155
<b>Condition1_RRNe</b>	0.0476	0.084	0.565	0.572	-0.118 0.213

<b>Condition1_RRn</b>	0.1036	0.064	1.612	0.107	-0.023 0.230
<b>Condition2_Feetr</b>	0.0887	0.116	0.763	0.446	-0.139 0.317
<b>Condition2_Norm</b>	0.0464	0.101	0.460	0.646	-0.152 0.244
<b>Condition2_PosA</b>	0.2207	0.187	1.178	0.239	-0.147 0.588
<b>Condition2_PosN</b>	-0.8195	0.137	-5.985	0.000	-1.088 -0.551
<b>Condition2_RRAe</b>	-0.5692	0.240	-2.374	0.018	-1.040 -0.099
<b>Condition2_RRAn</b>	-0.0424	0.153	-0.277	0.782	-0.343 0.258
<b>Condition2_RRn</b>	-0.0119	0.134	-0.089	0.929	-0.275 0.251
<b>BldgType_2fmCon</b>	0.0275	0.068	0.405	0.686	-0.106 0.161
<b>BldgType_Duplex</b>	0.0042	0.045	0.093	0.926	-0.084 0.092
<b>BldgType_Twnhs</b>	-0.0489	0.054	-0.905	0.366	-0.155 0.057
<b>BldgType_TwnhsE</b>	-0.0283	0.049	-0.578	0.563	-0.125 0.068
<b>HouseStyle_1.5Unf</b>	-0.0026	0.046	-0.056	0.955	-0.094 0.089
<b>HouseStyle_1Story</b>	-0.0395	0.024	-1.653	0.099	-0.086 0.007
<b>HouseStyle_2.5Fin</b>	-0.1062	0.064	-1.668	0.096	-0.231 0.019
<b>HouseStyle_2.5Unf</b>	0.1057	0.061	1.722	0.085	-0.015 0.226
<b>HouseStyle_2Story</b>	-0.0204	0.019	-1.091	0.275	-0.057 0.016
<b>HouseStyle_SFoyer</b>	-0.0237	0.035	-0.686	0.493	-0.091 0.044
<b>HouseStyle_SLvl</b>	-0.0118	0.030	-0.395	0.693	-0.070 0.047
<b>RoofStyle_Gable</b>	-0.0140	0.094	-0.150	0.881	-0.198 0.170
<b>RoofStyle_Gambrel</b>	-0.0089	0.102	-0.088	0.930	-0.209 0.191
<b>RoofStyle_Hip</b>	-0.0112	0.094	-0.120	0.905	-0.196 0.173
<b>RoofStyle_Mansard</b>	0.0620	0.110	0.563	0.573	-0.154 0.278
<b>RoofStyle_Shed</b>	0.4625	0.185	2.496	0.013	0.099 0.826
<b>RoofMatl_CompShg</b>	2.7998	0.195	14.367	0.000	2.417 3.182
<b>RoofMatl_Membran</b>	3.1696	0.262	12.103	0.000	2.656 3.684
<b>RoofMatl_Metal</b>	3.0197	0.258	11.726	0.000	2.514 3.525
<b>RoofMatl_Roll</b>	2.8164	0.229	12.314	0.000	2.368 3.265
<b>RoofMatl_Tar&amp;Grv</b>	2.8128	0.215	13.096	0.000	2.391 3.234
<b>RoofMatl_WdShake</b>	2.6856	0.223	12.032	0.000	2.248 3.124
<b>RoofMatl_WdShngl</b>	2.8669	0.199	14.399	0.000	2.476 3.258
<b>Exterior1st_AsphShn</b>	2.016e-15	8.98e-16	2.245	0.025	2.54e-16 3.78e-15
<b>Exterior1st_BrkComm</b>	-0.1985	0.146	-1.362	0.174	-0.485 0.088

<b>Exterior1st_BrkFace</b>	0.1115	0.080	1.399	0.162	-0.045 0.268
<b>Exterior1st_CBlock</b>	-0.0092	0.068	-0.134	0.894	-0.143 0.125
<b>Exterior1st_CemntBd</b>	-0.1175	0.118	-0.994	0.321	-0.349 0.115
<b>Exterior1st_HdBoard</b>	-0.0005	0.078	-0.006	0.995	-0.154 0.153
<b>Exterior1st_ImStucc</b>	6.596e-15	7.22e-16	9.133	0.000	5.18e-15 8.01e-15
<b>Exterior1st_MetalSd</b>	0.0762	0.086	0.886	0.376	-0.093 0.245
<b>Exterior1st_Plywood</b>	0.0098	0.077	0.127	0.899	-0.142 0.162
<b>Exterior1st_Stone</b>	0.0983	0.176	0.557	0.578	-0.248 0.445
<b>Exterior1st_Stucco</b>	0.0082	0.089	0.093	0.926	-0.166 0.182
<b>Exterior1st_VinylSd</b>	0.0303	0.079	0.381	0.703	-0.126 0.186
<b>Exterior1st_Wd Sdng</b>	-0.0202	0.076	-0.264	0.792	-0.170 0.130
<b>Exterior1st_WdShing</b>	0.0095	0.081	0.116	0.907	-0.150 0.169
<b>Exterior2nd_AsphShn</b>	-0.0312	0.117	-0.266	0.790	-0.261 0.199
<b>Exterior2nd_Brk Cmn</b>	-0.0165	0.110	-0.150	0.881	-0.232 0.199
<b>Exterior2nd_BrkFace</b>	-0.0890	0.080	-1.118	0.264	-0.245 0.067
<b>Exterior2nd_CBlock</b>	-0.0092	0.068	-0.134	0.894	-0.143 0.125
<b>Exterior2nd_CmentBd</b>	0.1360	0.116	1.172	0.242	-0.092 0.364
<b>Exterior2nd_HdBoard</b>	-0.0311	0.075	-0.414	0.679	-0.179 0.116
<b>Exterior2nd_ImStucc</b>	0.0148	0.085	0.173	0.862	-0.152 0.182
<b>Exterior2nd_MetalSd</b>	-0.0783	0.083	-0.937	0.349	-0.242 0.086
<b>Exterior2nd_Other</b>	-0.1486	0.136	-1.091	0.275	-0.416 0.119
<b>Exterior2nd_Plywood</b>	-0.0398	0.073	-0.543	0.587	-0.184 0.104
<b>Exterior2nd_Stone</b>	-0.0042	0.105	-0.040	0.968	-0.210 0.202
<b>Exterior2nd_Stucco</b>	-0.0236	0.086	-0.275	0.784	-0.192 0.145
<b>Exterior2nd_VinylSd</b>	-0.0373	0.076	-0.494	0.621	-0.186 0.111
<b>Exterior2nd_Wd Sdng</b>	-0.0101	0.073	-0.137	0.891	-0.154 0.134
<b>Exterior2nd_Wd Shng</b>	-0.0250	0.076	-0.327	0.744	-0.175 0.125
<b>MasVnrType_BrkFace</b>	0.0280	0.029	0.961	0.337	-0.029 0.085
<b>MasVnrType_None</b>	0.0251	0.029	0.856	0.392	-0.032 0.083
<b>MasVnrType_Stone</b>	0.0319	0.032	1.010	0.313	-0.030 0.094
<b>ExterQual_Fa</b>	-0.0466	0.063	-0.741	0.459	-0.170 0.077
<b>ExterQual_Gd</b>	0.0165	0.027	0.617	0.537	-0.036 0.069
<b>ExterQual_TA</b>	0.0200	0.030	0.678	0.498	-0.038 0.078

<b>ExterCond_Fa</b>	-0.1134	0.088	-1.287	0.199	-0.286 0.060
<b>ExterCond_Gd</b>	-0.0753	0.083	-0.912	0.362	-0.237 0.087
<b>ExterCond_Po</b>	-0.1086	0.156	-0.695	0.487	-0.415 0.198
<b>ExterCond_TA</b>	-0.0471	0.082	-0.573	0.567	-0.209 0.114
<b>Foundation_CBlock</b>	0.0239	0.017	1.419	0.156	-0.009 0.057
<b>Foundation_PConc</b>	0.0406	0.018	2.236	0.026	0.005 0.076
<b>Foundation_Slab</b>	-0.0387	0.049	-0.788	0.431	-0.135 0.058
<b>Foundation_Stone</b>	0.0938	0.055	1.721	0.086	-0.013 0.201
<b>Foundation_Wood</b>	-0.1449	0.072	-2.022	0.043	-0.286 -0.004
<b>BsmtQual_Fa</b>	-0.0149	0.031	-0.479	0.632	-0.076 0.046
<b>BsmtQual_Gd</b>	-0.0269	0.017	-1.594	0.111	-0.060 0.006
<b>BsmtQual_TA</b>	-0.0394	0.021	-1.909	0.057	-0.080 0.001
<b>BsmtCond_Gd</b>	0.0441	0.028	1.573	0.116	-0.011 0.099
<b>BsmtCond_Po</b>	0.3252	0.151	2.147	0.032	0.028 0.622
<b>BsmtCond_TA</b>	0.0304	0.022	1.353	0.176	-0.014 0.074
<b>BsmtExposure_Gd</b>	0.0234	0.016	1.437	0.151	-0.009 0.055
<b>BsmtExposure_Mn</b>	-0.0033	0.016	-0.204	0.838	-0.036 0.029
<b>BsmtExposure_No</b>	-0.0133	0.012	-1.109	0.268	-0.037 0.010
<b>BsmtFinType1_BLQ</b>	-0.0097	0.015	-0.657	0.511	-0.039 0.019
<b>BsmtFinType1_GLQ</b>	0.0153	0.013	1.166	0.244	-0.010 0.041
<b>BsmtFinType1_LwQ</b>	-0.0212	0.020	-1.083	0.279	-0.060 0.017
<b>BsmtFinType1_Rec</b>	-0.0093	0.016	-0.574	0.566	-0.041 0.022
<b>BsmtFinType1_Unf</b>	-0.0038	0.015	-0.251	0.802	-0.033 0.026
<b>BsmtFinType2_BLQ</b>	-0.0665	0.038	-1.742	0.082	-0.141 0.008
<b>BsmtFinType2_GLQ</b>	0.0058	0.049	0.118	0.906	-0.090 0.102
<b>BsmtFinType2_LwQ</b>	-0.0301	0.037	-0.806	0.420	-0.103 0.043
<b>BsmtFinType2_Rec</b>	-0.0225	0.034	-0.656	0.512	-0.090 0.045
<b>BsmtFinType2_Unf</b>	-0.0082	0.034	-0.241	0.809	-0.075 0.059
<b>Heating_GasA</b>	0.1424	0.124	1.153	0.249	-0.100 0.385
<b>Heating_GasW</b>	0.2225	0.131	1.693	0.091	-0.035 0.480
<b>Heating_Grav</b>	0.1932	0.160	1.209	0.227	-0.120 0.507
<b>Heating_OthW</b>	0.1105	0.154	0.717	0.474	-0.192 0.413
<b>Heating_Wall</b>	0.1678	0.150	1.119	0.263	-0.126 0.462

HeatingQC_Fa	-0.0165	0.025	-0.666	0.505	-0.065 0.032
HeatingQC_Gd	-0.0200	0.011	-1.813	0.070	-0.042 0.002
HeatingQC_Po	-0.0873	0.131	-0.666	0.505	-0.344 0.170
HeatingQC_TA	-0.0324	0.011	-2.980	0.003	-0.054 -0.011
CentralAir_Y	0.0605	0.021	2.879	0.004	0.019 0.102
Electrical_FuseF	-0.0284	0.033	-0.858	0.391	-0.093 0.037
Electrical_FuseP	-0.0688	0.096	-0.715	0.475	-0.258 0.120
Electrical_Mix	-0.0942	0.226	-0.417	0.677	-0.538 0.350
Electrical_SBrkr	-0.0024	0.016	-0.153	0.878	-0.033 0.029
KitchenQual_Fa	-0.0073	0.036	-0.200	0.841	-0.078 0.064
KitchenQual_Gd	-0.0600	0.019	-3.081	0.002	-0.098 -0.022
KitchenQual_TA	-0.0551	0.022	-2.527	0.012	-0.098 -0.012
Functional_Maj2	-0.2985	0.082	-3.654	0.000	-0.459 -0.138
Functional_Min1	0.0624	0.047	1.326	0.185	-0.030 0.155
Functional_Min2	0.0240	0.047	0.516	0.606	-0.067 0.115
Functional_Mod	-0.0563	0.059	-0.957	0.339	-0.172 0.059
Functional_Sev	6.608e-16	3.09e-16	2.139	0.033	5.45e-17 1.27e-15
Functional_Typ	0.0756	0.042	1.818	0.069	-0.006 0.157
FireplaceQu_Fa	-0.0046	0.027	-0.173	0.863	-0.057 0.048
FireplaceQu_Gd	0.0048	0.016	0.303	0.762	-0.027 0.036
FireplaceQu_Po	-0.0036	0.035	-0.100	0.920	-0.073 0.066
FireplaceQu_TA	0.0113	0.017	0.650	0.516	-0.023 0.046
GarageType_Attchd	0.1118	0.055	2.023	0.043	0.003 0.220
GarageType_Basment	0.1019	0.064	1.588	0.113	-0.024 0.228
GarageType_BuiltIn	0.0975	0.058	1.676	0.094	-0.017 0.212
GarageType_CarPort	0.1229	0.072	1.696	0.090	-0.019 0.265
GarageType_Detchd	0.0986	0.055	1.789	0.074	-0.010 0.207
GarageFinish_RFn	-0.0030	0.010	-0.287	0.774	-0.024 0.018
GarageFinish_Unf	-0.0124	0.013	-0.970	0.332	-0.037 0.013
GarageQual_Fa	-0.0681	0.041	-1.663	0.097	-0.148 0.012
GarageQual_Gd	0.0165	0.050	0.331	0.741	-0.081 0.114
GarageQual_Po	-0.1694	0.111	-1.530	0.126	-0.387 0.048
GarageQual_TA	-0.0074	0.040	-0.185	0.853	-0.086 0.071

<b>GarageCond_Fa</b>	-0.1160	0.045	-2.555	0.011	-0.205 -0.027
<b>GarageCond_Gd</b>	-0.0605	0.056	-1.078	0.281	-0.171 0.050
<b>GarageCond_Po</b>	0.0170	0.086	0.198	0.843	-0.152 0.186
<b>GarageCond_TA</b>	-0.0690	0.041	-1.696	0.090	-0.149 0.011
<b>PavedDrive_P</b>	-0.0022	0.030	-0.073	0.942	-0.061 0.056
<b>PavedDrive_Y</b>	0.0141	0.018	0.766	0.444	-0.022 0.050
<b>PoolQC_Fa</b>	0.0108	0.166	0.065	0.948	-0.315 0.336
<b>PoolQC_Gd</b>	0.1406	0.148	0.948	0.343	-0.150 0.432
<b>Fence_GdWo</b>	-0.0324	0.019	-1.734	0.083	-0.069 0.004
<b>Fence_MnPrv</b>	-0.0112	0.012	-0.917	0.359	-0.035 0.013
<b>Fence_MnWw</b>	-0.0471	0.041	-1.143	0.254	-0.128 0.034
<b>MiscFeature_Othr</b>	-0.1506	0.149	-1.013	0.312	-0.442 0.141
<b>MiscFeature_Shed</b>	-0.0071	0.021	-0.344	0.731	-0.048 0.033
<b>MiscFeature_TenC</b>	-0.2110	0.169	-1.249	0.212	-0.542 0.120
<b>SaleType_CWD</b>	0.0862	0.062	1.385	0.166	-0.036 0.208
<b>SaleType_Con</b>	0.0979	0.084	1.163	0.245	-0.067 0.263
<b>SaleType_ConLD</b>	0.2808	0.059	4.787	0.000	0.166 0.396
<b>SaleType_ConLI</b>	-0.0379	0.062	-0.609	0.543	-0.160 0.084
<b>SaleType_ConLw</b>	-0.0452	0.069	-0.656	0.512	-0.180 0.090
<b>SaleType_New</b>	0.1404	0.076	1.845	0.065	-0.009 0.290
<b>SaleType_Oth</b>	0.0559	0.070	0.804	0.422	-0.081 0.192
<b>SaleType_WD</b>	-0.0137	0.021	-0.637	0.524	-0.056 0.028
<b>SaleCondition_AdjLand</b>	0.1236	0.080	1.550	0.121	-0.033 0.280
<b>SaleCondition_Alloca</b>	0.0511	0.047	1.093	0.275	-0.041 0.143
<b>SaleCondition_Family</b>	-0.0018	0.034	-0.054	0.957	-0.068 0.064
<b>SaleCondition_Normal</b>	0.0588	0.016	3.781	0.000	0.028 0.089
<b>SaleCondition_Partial</b>	-0.0480	0.073	-0.656	0.512	-0.192 0.096

<b>Omnibus:</b>	313.878	<b>Durbin-Watson:</b>	2.041
<b>Prob(Omnibus):</b>	0.000	<b>Jarque-Bera (JB):</b>	4103.148
<b>Skew:</b>	-0.858	<b>Prob(JB):</b>	0.00
<b>Kurtosis:</b>	12.020	<b>Cond. No.</b>	1.39e+16



```
In [452]: local_train_x.shape
```

```
Out[452]: (1168, 246)
```

```
In [453]: local_train_x.drop(['MasVnrArea', 'LowQualFinSF', 'BsmtHalfBath', 'Condition1_RRAe', 'Condition2_RRNn', 'BldgType_Duplex', 'HouseStyle_1.5Unf', 'RoofStyle_Gable', 'RoofStyle_Gambrel', 'RoofStyle_Hip', 'Exterior1st_CBlock', 'Exterior1st_HdBoard', 'Exterior1st_Plywood', 'Exterior1st_Stucco', 'Exterior1st_WdShing', 'Exterior2nd_Brk Cmn', 'Exterior2nd_CBlock', 'Exterior2nd_ImStucc', 'Exterior2nd_Stone', 'Exterior2nd_Wd Sdng', 'BsmtExposure_Mn', 'BsmtFinType1_Unf', 'BsmtFinType2_GLQ', 'BsmtFinType2_Unf', 'Electrical_SBrkr', 'KitchenQual_Fa', 'FireplaceQu_Fa', 'FireplaceQu_Po', 'GarageQual_TA', 'GarageCond_Po', 'PavedDrive_P', 'PoolQC_Fa', 'SaleCondition_Family'], axis=1, inplace=True)
local_test_x.drop(['MasVnrArea', 'LowQualFinSF', 'BsmtHalfBath', 'Condition1_RRAe', 'Condition2_RRNn', 'BldgType_Duplex', 'HouseStyle_1.5Unf', 'RoofStyle_Gable', 'RoofStyle_Gambrel', 'RoofStyle_Hip', 'Exterior1st_CBlock', 'Exterior1st_HdBoard', 'Exterior1st_Plywood', 'Exterior1st_Stucco', 'Exterior1st_WdShing', 'Exterior2nd_Brk Cmn', 'Exterior2nd_CBlock', 'Exterior2nd_ImStucc', 'Exterior2nd_Stone', 'Exterior2nd_Wd Sdng', 'BsmtExposure_Mn', 'BsmtFinType1_Unf', 'BsmtFinType2_GLQ', 'BsmtFinType2_Unf', 'Electrical_SBrkr', 'KitchenQual_Fa', 'FireplaceQu_Fa', 'FireplaceQu_Po', 'GarageQual_TA', 'GarageCond_Po', 'PavedDrive_P', 'PoolQC_Fa', 'SaleCondition_Family'], axis=1, inplace=True)
local_train_x.shape
```

```
Out[453]: (1168, 213)
```

```
In [454]: #Refitting model  
clf = sm.OLS(local_train_y,local_train_x)  
result = clf.fit()  
result.summary()
```

Out[454]: OLS Regression Results

<b>Dep. Variable:</b>	SalePrice	<b>R-squared:</b>	1.000
<b>Model:</b>	OLS	<b>Adj. R-squared:</b>	1.000
<b>Method:</b>	Least Squares	<b>F-statistic:</b>	7.314e+04
<b>Date:</b>	Sun, 02 Oct 2016	<b>Prob (F-statistic):</b>	0.00
<b>Time:</b>	23:10:47	<b>Log-Likelihood:</b>	1088.7
<b>No. Observations:</b>	1168	<b>AIC:</b>	-1759.
<b>Df Residuals:</b>	959	<b>BIC:</b>	-701.3
<b>Df Model:</b>	209		
<b>Covariance Type:</b>	nonrobust		

	coef	std err	t	P> t	[95.0% Conf. Int.]
<b>Id</b>	-6.633e-06	8.03e-06	-0.826	0.409	-2.24e-05 9.13e-06
<b>MSSubClass</b>	-0.0003	0.000	-0.753	0.452	-0.001 0.000
<b>LotFrontage</b>	0.0005	0.000	2.079	0.038	2.89e-05 0.001
<b>LotArea</b>	2.359e-06	5.89e-07	4.005	0.000	1.2e-06 3.51e-06
<b>OverallQual</b>	0.0372	0.005	7.023	0.000	0.027 0.048
<b>OverallCond</b>	0.0369	0.005	8.195	0.000	0.028 0.046
<b>YearBuilt</b>	0.0017	0.000	4.495	0.000	0.001 0.002
<b>YearRemodAdd</b>	0.0007	0.000	2.593	0.010	0.000 0.001
<b>BsmtFinSF1</b>	5.466e-05	1.14e-05	4.797	0.000	3.23e-05 7.7e-05
<b>BsmtFinSF2</b>	5.584e-05	2.24e-05	2.497	0.013	1.2e-05 9.97e-05
<b>BsmtUnfSF</b>	-1.286e-05	1.19e-05	-1.079	0.281	-3.63e-05 1.05e-05
<b>TotalBsmtSF</b>	9.763e-05	1.8e-05	5.428	0.000	6.23e-05 0.000
<b>1stFlrSF</b>	7.892e-05	9.88e-05	0.799	0.425	-0.000 0.000
<b>2ndFlrSF</b>	6.108e-05	9.4e-05	0.650	0.516	-0.000 0.000
<b>GrLivArea</b>	0.0001	9.75e-05	1.518	0.129	-4.33e-05 0.000
<b>BsmtFullBath</b>	0.0260	0.010	2.710	0.007	0.007 0.045
<b>FullBath</b>	0.0109	0.011	0.969	0.333	-0.011 0.033
<b>HalfBath</b>	0.0140	0.011	1.296	0.195	-0.007 0.035
<b>BedroomAbvGr</b>	0.0050	0.007	0.716	0.474	-0.009 0.019
<b>KitchenAbvGr</b>	-0.0548	0.027	-2.042	0.041	-0.107 -0.002
<b>TotRmsAbvGrd</b>	0.0092	0.005	1.853	0.064	-0.001 0.019

<b>Fireplaces</b>	0.0161	0.010	1.597	0.111	-0.004 0.036
<b>GarageYrBlt</b>	-0.0004	0.000	-1.193	0.233	-0.001 0.000
<b>GarageCars</b>	0.0238	0.012	1.991	0.047	0.000 0.047
<b>GarageArea</b>	0.0001	4.15e-05	2.745	0.006	3.25e-05 0.000
<b>WoodDeckSF</b>	0.0001	2.97e-05	3.596	0.000	4.86e-05 0.000
<b>OpenPorchSF</b>	6.734e-05	6.18e-05	1.089	0.276	-5.4e-05 0.000
<b>EnclosedPorch</b>	3.093e-05	6.58e-05	0.470	0.638	-9.82e-05 0.000
<b>3SsnPorch</b>	0.0001	0.000	1.325	0.185	-7.06e-05 0.000
<b>ScreenPorch</b>	0.0003	6.58e-05	3.917	0.000	0.000 0.000
<b>PoolArea</b>	0.0002	0.000	1.232	0.218	-8.92e-05 0.000
<b>MiscVal</b>	3.682e-06	6.96e-06	0.529	0.597	-9.98e-06 1.73e-05
<b>MoSold</b>	0.0008	0.001	0.629	0.529	-0.002 0.003
<b>YrSold</b>	0.0015	0.000	3.140	0.002	0.001 0.002
<b>MSZoning_FV</b>	0.4906	0.059	8.350	0.000	0.375 0.606
<b>MSZoning_RH</b>	0.4883	0.058	8.463	0.000	0.375 0.602
<b>MSZoning_RL</b>	0.4700	0.049	9.565	0.000	0.374 0.566
<b>MSZoning_RM</b>	0.4399	0.046	9.547	0.000	0.349 0.530
<b>Street_Pave</b>	0.1029	0.059	1.749	0.081	-0.013 0.218
<b>Alley_Pave</b>	0.0469	0.024	1.977	0.048	0.000 0.093
<b>LotShape_IR2</b>	0.0318	0.022	1.426	0.154	-0.012 0.075
<b>LotShape_IR3</b>	0.0133	0.048	0.274	0.784	-0.082 0.108
<b>LotShape_Reg</b>	0.0039	0.008	0.464	0.643	-0.013 0.020
<b>LandContour_HLS</b>	0.0444	0.027	1.671	0.095	-0.008 0.096
<b>LandContour_Low</b>	-0.0329	0.033	-1.006	0.315	-0.097 0.031
<b>LandContour_Lvl</b>	0.0247	0.019	1.301	0.194	-0.013 0.062
<b>Utilities_NoSeWa</b>	-0.1962	0.121	-1.617	0.106	-0.434 0.042
<b>LotConfig_CulDSac</b>	0.0342	0.018	1.942	0.052	-0.000 0.069
<b>LotConfig_FR2</b>	-0.0295	0.021	-1.408	0.159	-0.071 0.012
<b>LotConfig_FR3</b>	-0.0664	0.088	-0.754	0.451	-0.239 0.106
<b>LotConfig_Inside</b>	-0.0092	0.009	-0.994	0.320	-0.027 0.009
<b>LandSlope_Mod</b>	0.0648	0.022	2.959	0.003	0.022 0.108
<b>LandSlope_Sev</b>	-0.1556	0.064	-2.445	0.015	-0.281 -0.031
<b>Neighborhood_Blueste</b>	-0.1292	0.115	-1.128	0.260	-0.354 0.096

<b>Neighborhood_BrDale</b>	-0.0934	0.055	-1.711	0.087	-0.201 0.014
<b>Neighborhood_BrkSide</b>	-0.0275	0.049	-0.560	0.576	-0.124 0.069
<b>Neighborhood_ClearCr</b>	-0.0042	0.048	-0.087	0.931	-0.098 0.090
<b>Neighborhood_CollgCr</b>	-0.0500	0.038	-1.324	0.186	-0.124 0.024
<b>Neighborhood_Crawfor</b>	0.0670	0.044	1.516	0.130	-0.020 0.154
<b>Neighborhood_Edwards</b>	-0.1052	0.042	-2.524	0.012	-0.187 -0.023
<b>Neighborhood_Gilbert</b>	-0.0551	0.040	-1.379	0.168	-0.133 0.023
<b>Neighborhood_IDOTRR</b>	-0.0464	0.055	-0.845	0.399	-0.154 0.061
<b>Neighborhood_MeadowV</b>	-0.2256	0.056	-4.006	0.000	-0.336 -0.115
<b>Neighborhood_Mitchel</b>	-0.0984	0.043	-2.301	0.022	-0.182 -0.014
<b>Neighborhood_NAmes</b>	-0.0655	0.041	-1.605	0.109	-0.146 0.015
<b>Neighborhood_NPkVill</b>	-0.0356	0.056	-0.640	0.522	-0.145 0.074
<b>Neighborhood_NWAmes</b>	-0.0696	0.042	-1.669	0.095	-0.151 0.012
<b>Neighborhood_NoRidge</b>	-0.0055	0.043	-0.128	0.898	-0.089 0.078
<b>Neighborhood_NridgHt</b>	0.0747	0.039	1.915	0.056	-0.002 0.151
<b>Neighborhood_OldTown</b>	-0.0675	0.049	-1.371	0.171	-0.164 0.029
<b>Neighborhood_SWISU</b>	-0.0278	0.049	-0.568	0.570	-0.124 0.068
<b>Neighborhood_Sawyer</b>	-0.0723	0.042	-1.708	0.088	-0.155 0.011
<b>Neighborhood_SawyerW</b>	-0.0464	0.040	-1.158	0.247	-0.125 0.032
<b>Neighborhood_Somerst</b>	-0.0046	0.046	-0.100	0.920	-0.095 0.086
<b>Neighborhood_StoneBr</b>	0.1035	0.042	2.452	0.014	0.021 0.186
<b>Neighborhood_Timber</b>	-0.0274	0.042	-0.655	0.513	-0.109 0.055
<b>Neighborhood_Veenker</b>	0.0199	0.052	0.384	0.701	-0.082 0.121
<b>Condition1_Feedr</b>	0.0692	0.024	2.920	0.004	0.023 0.116
<b>Condition1_Norm</b>	0.1111	0.019	5.824	0.000	0.074 0.148
<b>Condition1_PosA</b>	0.0722	0.050	1.455	0.146	-0.025 0.170
<b>Condition1_PosN</b>	0.1063	0.038	2.810	0.005	0.032 0.180
<b>Condition1_RRAn</b>	0.0840	0.033	2.548	0.011	0.019 0.149
<b>Condition1_RRNe</b>	0.0478	0.081	0.588	0.557	-0.112 0.207
<b>Condition1_RRNn</b>	0.1093	0.060	1.813	0.070	-0.009 0.228
<b>Condition2_Feedr</b>	0.0975	0.083	1.178	0.239	-0.065 0.260
<b>Condition2_Norm</b>	0.0576	0.062	0.925	0.355	-0.065 0.180
<b>Condition2_PosA</b>	0.2246	0.169	1.333	0.183	-0.106 0.555

<b>Condition2_PosN</b>	-0.8095	0.110	-7.386	0.000	-1.025 -0.594
<b>Condition2_RRAe</b>	-0.5561	0.219	-2.542	0.011	-0.985 -0.127
<b>Condition2_RRAn</b>	-0.0331	0.128	-0.259	0.795	-0.283 0.217
<b>BldgType_2fmCon</b>	0.0262	0.052	0.507	0.612	-0.075 0.128
<b>BldgType_Twnhs</b>	-0.0492	0.048	-1.024	0.306	-0.144 0.045
<b>BldgType_TwnhsE</b>	-0.0278	0.043	-0.643	0.521	-0.113 0.057
<b>HouseStyle_1Story</b>	-0.0392	0.021	-1.864	0.063	-0.081 0.002
<b>HouseStyle_2.5Fin</b>	-0.1049	0.061	-1.718	0.086	-0.225 0.015
<b>HouseStyle_2.5Unf</b>	0.1054	0.059	1.780	0.075	-0.011 0.222
<b>HouseStyle_2Story</b>	-0.0198	0.018	-1.122	0.262	-0.054 0.015
<b>HouseStyle_SFoyer</b>	-0.0227	0.032	-0.701	0.484	-0.086 0.041
<b>HouseStyle_SLvl</b>	-0.0100	0.027	-0.367	0.714	-0.064 0.044
<b>RoofStyle_Mansard</b>	0.0788	0.056	1.397	0.163	-0.032 0.189
<b>RoofStyle_Shed</b>	0.4802	0.164	2.934	0.003	0.159 0.801
<b>RoofMatl_CompShg</b>	2.8039	0.181	15.477	0.000	2.448 3.159
<b>RoofMatl_Membran</b>	3.1916	0.229	13.915	0.000	2.741 3.642
<b>RoofMatl_Metal</b>	3.0358	0.226	13.455	0.000	2.593 3.479
<b>RoofMatl_Roll</b>	2.8252	0.212	13.300	0.000	2.408 3.242
<b>RoofMatl_Tar&amp;Grv</b>	2.8268	0.184	15.335	0.000	2.465 3.189
<b>RoofMatl_WdShake</b>	2.6833	0.210	12.780	0.000	2.271 3.095
<b>RoofMatl_WdShngl</b>	2.8706	0.185	15.488	0.000	2.507 3.234
<b>Exterior1st_AsphShn</b>	3.02e-15	4.82e-16	6.263	0.000	2.07e-15 3.97e-15
<b>Exterior1st_BrkComm</b>	-0.2126	0.093	-2.278	0.023	-0.396 -0.029
<b>Exterior1st_BrkFace</b>	0.1049	0.028	3.781	0.000	0.050 0.159
<b>Exterior1st_CemntBd</b>	-0.1236	0.085	-1.453	0.147	-0.291 0.043
<b>Exterior1st_ImStucc</b>	-4.526e-16	2.97e-16	-1.525	0.128	-1.04e-15 1.3e-16
<b>Exterior1st_MetalSd</b>	0.0711	0.041	1.738	0.083	-0.009 0.151
<b>Exterior1st_Stone</b>	0.1054	0.154	0.683	0.495	-0.198 0.409
<b>Exterior1st_VinylSd</b>	0.0238	0.039	0.611	0.542	-0.053 0.100
<b>Exterior1st_Wd Sdng</b>	-0.0288	0.020	-1.459	0.145	-0.068 0.010
<b>Exterior2nd_AsphShn</b>	-0.0230	0.086	-0.266	0.790	-0.193 0.147
<b>Exterior2nd_BrkFace</b>	-0.0800	0.034	-2.355	0.019	-0.147 -0.013
<b>Exterior2nd_CmentBd</b>	0.1421	0.086	1.651	0.099	-0.027 0.311

<b>Exterior2nd_HdBoard</b>	-0.0305	0.021	-1.460	0.145	-0.071 0.010
<b>Exterior2nd_MetalSd</b>	-0.0731	0.042	-1.727	0.085	-0.156 0.010
<b>Exterior2nd_Other</b>	-0.1430	0.119	-1.204	0.229	-0.376 0.090
<b>Exterior2nd_Plywood</b>	-0.0322	0.021	-1.558	0.120	-0.073 0.008
<b>Exterior2nd_Stucco</b>	-0.0130	0.034	-0.389	0.698	-0.079 0.053
<b>Exterior2nd_VinylSd</b>	-0.0304	0.039	-0.773	0.439	-0.107 0.047
<b>Exterior2nd_Wd Shng</b>	-0.0164	0.025	-0.649	0.517	-0.066 0.033
<b>MasVnrType_BrkFace</b>	0.0283	0.029	0.994	0.321	-0.028 0.084
<b>MasVnrType_None</b>	0.0249	0.028	0.891	0.373	-0.030 0.080
<b>MasVnrType_Stone</b>	0.0321	0.031	1.037	0.300	-0.029 0.093
<b>ExterQual_Fa</b>	-0.0533	0.054	-0.997	0.319	-0.158 0.052
<b>ExterQual_Gd</b>	0.0135	0.025	0.538	0.591	-0.036 0.063
<b>ExterQual_TA</b>	0.0168	0.028	0.601	0.548	-0.038 0.072
<b>ExterCond_Fa</b>	-0.1184	0.084	-1.404	0.161	-0.284 0.047
<b>ExterCond_Gd</b>	-0.0800	0.078	-1.021	0.307	-0.234 0.074
<b>ExterCond_Po</b>	-0.1198	0.149	-0.802	0.423	-0.413 0.173
<b>ExterCond_TA</b>	-0.0519	0.078	-0.665	0.506	-0.205 0.101
<b>Foundation_CBlock</b>	0.0250	0.016	1.557	0.120	-0.007 0.057
<b>Foundation_PConc</b>	0.0408	0.017	2.349	0.019	0.007 0.075
<b>Foundation_Slab</b>	-0.0317	0.043	-0.744	0.457	-0.115 0.052
<b>Foundation_Stone</b>	0.0917	0.053	1.742	0.082	-0.012 0.195
<b>Foundation_Wood</b>	-0.1447	0.070	-2.066	0.039	-0.282 -0.007
<b>BsmtQual_Fa</b>	-0.0185	0.029	-0.628	0.530	-0.076 0.039
<b>BsmtQual_Gd</b>	-0.0292	0.016	-1.807	0.071	-0.061 0.003
<b>BsmtQual_TA</b>	-0.0418	0.020	-2.141	0.032	-0.080 -0.003
<b>BsmtCond_Gd</b>	0.0432	0.027	1.609	0.108	-0.009 0.096
<b>BsmtCond_Po</b>	0.3258	0.146	2.236	0.026	0.040 0.612
<b>BsmtCond_TA</b>	0.0288	0.021	1.376	0.169	-0.012 0.070
<b>BsmtExposure_Gd</b>	0.0240	0.015	1.609	0.108	-0.005 0.053
<b>BsmtExposure_No</b>	-0.0128	0.009	-1.362	0.174	-0.031 0.006
<b>BsmtFinType1_BLQ</b>	-0.0075	0.013	-0.573	0.567	-0.033 0.018
<b>BsmtFinType1_GLQ</b>	0.0166	0.011	1.485	0.138	-0.005 0.039
<b>BsmtFinType1_LwQ</b>	-0.0191	0.018	-1.090	0.276	-0.053 0.015

<b>BsmtFinType1_Rec</b>	-0.0075	0.014	-0.529	0.597	-0.035 0.020
<b>BsmtFinType2_BLQ</b>	-0.0620	0.026	-2.396	0.017	-0.113 -0.011
<b>BsmtFinType2_LwQ</b>	-0.0254	0.024	-1.055	0.292	-0.073 0.022
<b>BsmtFinType2_Rec</b>	-0.0176	0.021	-0.841	0.400	-0.059 0.023
<b>Heating_GasA</b>	0.1327	0.118	1.122	0.262	-0.099 0.365
<b>Heating_GasW</b>	0.2099	0.125	1.677	0.094	-0.036 0.456
<b>Heating_Grav</b>	0.1801	0.146	1.232	0.218	-0.107 0.467
<b>Heating_OthW</b>	0.1028	0.148	0.696	0.487	-0.187 0.393
<b>Heating_Wall</b>	0.1601	0.140	1.141	0.254	-0.115 0.435
<b>HeatingQC_Fa</b>	-0.0175	0.024	-0.741	0.459	-0.064 0.029
<b>HeatingQC_Gd</b>	-0.0203	0.011	-1.899	0.058	-0.041 0.001
<b>HeatingQC_Po</b>	-0.0924	0.116	-0.793	0.428	-0.321 0.136
<b>HeatingQC_TA</b>	-0.0330	0.011	-3.134	0.002	-0.054 -0.012
<b>CentralAir_Y</b>	0.0602	0.020	3.031	0.002	0.021 0.099
<b>Electrical_FuseF</b>	-0.0258	0.029	-0.898	0.370	-0.082 0.031
<b>Electrical_FuseP</b>	-0.0702	0.086	-0.820	0.412	-0.238 0.098
<b>Electrical_Mix</b>	-0.0907	0.217	-0.418	0.676	-0.517 0.336
<b>KitchenQual_Gd</b>	-0.0585	0.016	-3.590	0.000	-0.090 -0.027
<b>KitchenQual_TA</b>	-0.0531	0.017	-3.109	0.002	-0.087 -0.020
<b>Functional_Maj2</b>	-0.2975	0.079	-3.787	0.000	-0.452 -0.143
<b>Functional_Min1</b>	0.0631	0.046	1.387	0.166	-0.026 0.153
<b>Functional_Min2</b>	0.0262	0.045	0.585	0.559	-0.062 0.114
<b>Functional_Mod</b>	-0.0569	0.057	-0.992	0.321	-0.169 0.056
<b>Functional_Sev</b>	-9.474e-16	1.67e-16	-5.671	0.000	-1.28e-15 -6.2e-16
<b>Functional_Typ</b>	0.0759	0.040	1.893	0.059	-0.003 0.155
<b>FireplaceQu_Gd</b>	0.0061	0.013	0.457	0.648	-0.020 0.032
<b>FireplaceQu_TA</b>	0.0131	0.014	0.916	0.360	-0.015 0.041
<b>GarageType_Attchd</b>	0.1133	0.048	2.361	0.018	0.019 0.208
<b>GarageType_Basment</b>	0.1036	0.058	1.800	0.072	-0.009 0.217
<b>GarageType_BuiltIn</b>	0.0991	0.051	1.944	0.052	-0.001 0.199
<b>GarageType_CarPort</b>	0.1280	0.064	1.994	0.046	0.002 0.254
<b>GarageType_Detchd</b>	0.1005	0.048	2.107	0.035	0.007 0.194
<b>GarageFinish_RFn</b>	-0.0033	0.010	-0.327	0.744	-0.023 0.016



<b>GarageFinish_Unf</b>	-0.0125	0.012	-1.019	0.308	-0.037 0.012
<b>GarageQual_Fa</b>	-0.0613	0.023	-2.657	0.008	-0.107 -0.016
<b>GarageQual_Gd</b>	0.0247	0.035	0.705	0.481	-0.044 0.094
<b>GarageQual_Po</b>	-0.1570	0.111	-1.408	0.159	-0.376 0.062
<b>GarageCond_Fa</b>	-0.1272	0.054	-2.358	0.019	-0.233 -0.021
<b>GarageCond_Gd</b>	-0.0717	0.066	-1.079	0.281	-0.202 0.059
<b>GarageCond_TA</b>	-0.0790	0.048	-1.639	0.102	-0.174 0.016
<b>PavedDrive_Y</b>	0.0149	0.015	0.980	0.327	-0.015 0.045
<b>PoolQC_Gd</b>	0.1282	0.117	1.095	0.274	-0.102 0.358
<b>Fence_GdWo</b>	-0.0321	0.018	-1.780	0.075	-0.068 0.003
<b>Fence_MnPrv</b>	-0.0109	0.012	-0.931	0.352	-0.034 0.012
<b>Fence_MnWw</b>	-0.0467	0.040	-1.170	0.242	-0.125 0.032
<b>MiscFeature_Othr</b>	-0.1512	0.141	-1.069	0.285	-0.429 0.126
<b>MiscFeature_Shed</b>	-0.0073	0.020	-0.362	0.718	-0.047 0.032
<b>MiscFeature_TenC</b>	-0.2123	0.132	-1.614	0.107	-0.470 0.046
<b>SaleType_CWD</b>	0.0863	0.060	1.433	0.152	-0.032 0.204
<b>SaleType_Con</b>	0.0958	0.082	1.162	0.246	-0.066 0.258
<b>SaleType_ConLD</b>	0.2811	0.057	4.970	0.000	0.170 0.392
<b>SaleType_ConLI</b>	-0.0406	0.060	-0.681	0.496	-0.158 0.076
<b>SaleType_ConLw</b>	-0.0433	0.066	-0.658	0.511	-0.173 0.086
<b>SaleType_New</b>	0.1395	0.074	1.884	0.060	-0.006 0.285
<b>SaleType_Oth</b>	0.0544	0.068	0.799	0.424	-0.079 0.188
<b>SaleType_WD</b>	-0.0137	0.021	-0.667	0.505	-0.054 0.027
<b>SaleCondition_AdjLand</b>	0.1270	0.075	1.690	0.091	-0.020 0.274
<b>SaleCondition_Alloca</b>	0.0509	0.044	1.149	0.251	-0.036 0.138
<b>SaleCondition_Normal</b>	0.0588	0.013	4.387	0.000	0.033 0.085
<b>SaleCondition_Partial</b>	-0.0477	0.071	-0.670	0.503	-0.187 0.092

<b>Omnibus:</b>	314.318	<b>Durbin-Watson:</b>	2.040
<b>Prob(Omnibus):</b>	0.000	<b>Jarque-Bera (JB):</b>	4130.026
<b>Skew:</b>	-0.858	<b>Prob(JB):</b>	0.00
<b>Kurtosis:</b>	12.051	<b>Cond. No.</b>	9.27e+16

```
In [455]: local_train_x.drop(['2ndFlrSF', 'EnclosedPorch', 'MiscVal', 'MoSold', 'LotShape_IR3', 'LotShape_Reg', 'Neighborhood_BrkSide', 'Neighborhood_ClearCr', 'Neighborhood_NPkvill', 'Neighborhood_NoRidge', 'Neighborhood_SWISU', 'Neighborhood_Somerst', 'Neighborhood_Timber', 'Neighborhood_Veenker', 'Condition1_RRNe', 'Condition2_RRAn', 'BldgType_2fmCon', 'BldgType_TwnhsE', 'HouseStyle_SLvl', 'Exterior1st_VinylSd', 'Exterior2nd_AsphShn', 'Exterior2nd_Stucco', 'Exterior2nd_Wd Shng', 'ExterQual_Gd', 'ExterQual_TA', 'ExterCond_TA', 'BsmtQual_Fa', 'BsmtFinType1_BLQ', 'BsmtFinType1_Rec', 'Electrical_Mix', 'Functional_Min2', 'FireplaceQu_Gd', 'GarageFinish_RFn', 'MiscFeature_Shed', 'SaleType_ConLw', 'SaleType_WD', 'SaleCondition_Partial'], axis=1, inplace=True)
local_test_x.drop(['2ndFlrSF', 'EnclosedPorch', 'MiscVal', 'MoSold', 'LotShape_IR3', 'LotShape_Reg', 'Neighborhood_BrkSide', 'Neighborhood_ClearCr', 'Neighborhood_NPkvill', 'Neighborhood_NoRidge', 'Neighborhood_SWISU', 'Neighborhood_Somerst', 'Neighborhood_Timber', 'Neighborhood_Veenker', 'Condition1_RRNe', 'Condition2_RRAn', 'BldgType_2fmCon', 'BldgType_TwnhsE', 'HouseStyle_SLvl', 'Exterior1st_VinylSd', 'Exterior2nd_AsphShn', 'Exterior2nd_Stucco', 'Exterior2nd_Wd Shng', 'ExterQual_Gd', 'ExterQual_TA', 'ExterCond_TA', 'BsmtQual_Fa', 'BsmtFinType1_BLQ', 'BsmtFinType1_Rec', 'Electrical_Mix', 'Functional_Min2', 'FireplaceQu_Gd', 'GarageFinish_RFn', 'MiscFeature_Shed', 'SaleType_ConLw', 'SaleType_WD', 'SaleCondition_Partial'], axis=1, inplace=True)
local_train_x.shape
```

```
Out[455]: (1168, 176)
```

```
In [456]: #Refitting model  
clf = sm.OLS(local_train_y,local_train_x)  
result = clf.fit()  
result.summary()
```

Out[456]: OLS Regression Results

<b>Dep. Variable:</b>	SalePrice	<b>R-squared:</b>	1.000
<b>Model:</b>	OLS	<b>Adj. R-squared:</b>	1.000
<b>Method:</b>	Least Squares	<b>F-statistic:</b>	9.121e+04
<b>Date:</b>	Sun, 02 Oct 2016	<b>Prob (F-statistic):</b>	0.00
<b>Time:</b>	23:10:47	<b>Log-Likelihood:</b>	1081.7
<b>No. Observations:</b>	1168	<b>AIC:</b>	-1819.
<b>Df Residuals:</b>	996	<b>BIC:</b>	-948.6
<b>Df Model:</b>	172		
<b>Covariance Type:</b>	nonrobust		

	coef	std err	t	P> t	[95.0% Conf. Int.]
<b>Id</b>	-7.964e-06	7.77e-06	-1.025	0.306	-2.32e-05 7.29e-06
<b>MSSubClass</b>	-0.0004	0.000	-2.967	0.003	-0.001 -0.000
<b>LotFrontage</b>	0.0006	0.000	2.591	0.010	0.000 0.001
<b>LotArea</b>	2.444e-06	5.27e-07	4.642	0.000	1.41e-06 3.48e-06
<b>OverallQual</b>	0.0367	0.005	7.427	0.000	0.027 0.046
<b>OverallCond</b>	0.0378	0.004	8.905	0.000	0.029 0.046
<b>YearBuilt</b>	0.0017	0.000	5.035	0.000	0.001 0.002
<b>YearRemodAdd</b>	0.0008	0.000	2.803	0.005	0.000 0.001
<b>BsmtFinSF1</b>	5.359e-05	1.08e-05	4.951	0.000	3.24e-05 7.48e-05
<b>BsmtFinSF2</b>	5.41e-05	2.13e-05	2.544	0.011	1.24e-05 9.58e-05
<b>BsmtUnfSF</b>	-1.185e-05	1.11e-05	-1.067	0.286	-3.37e-05 9.95e-06
<b>TotalBsmtSF</b>	9.583e-05	1.69e-05	5.668	0.000	6.27e-05 0.000
<b>1stFlrSF</b>	1.757e-05	3.16e-05	0.556	0.578	-4.44e-05 7.96e-05
<b>GrLivArea</b>	0.0002	2.43e-05	8.739	0.000	0.000 0.000
<b>BsmtFullBath</b>	0.0272	0.009	2.942	0.003	0.009 0.045
<b>FullBath</b>	0.0080	0.011	0.752	0.452	-0.013 0.029
<b>HalfBath</b>	0.0137	0.010	1.312	0.190	-0.007 0.034
<b>BedroomAbvGr</b>	0.0067	0.007	0.997	0.319	-0.006 0.020
<b>KitchenAbvGr</b>	-0.0383	0.023	-1.640	0.101	-0.084 0.008
<b>TotRmsAbvGrd</b>	0.0085	0.005	1.835	0.067	-0.001 0.018
<b>Fireplaces</b>	0.0190	0.007	2.596	0.010	0.005 0.033

<b>GarageYrBlt</b>	-0.0003	0.000	-1.137	0.256	-0.001 0.000
<b>GarageCars</b>	0.0245	0.012	2.133	0.033	0.002 0.047
<b>GarageArea</b>	0.0001	3.97e-05	2.926	0.004	3.83e-05 0.000
<b>WoodDeckSF</b>	0.0001	2.88e-05	3.842	0.000	5.4e-05 0.000
<b>OpenPorchSF</b>	6.825e-05	5.92e-05	1.153	0.249	-4.79e-05 0.000
<b>3SsnPorch</b>	0.0001	0.000	1.314	0.189	-6.92e-05 0.000
<b>ScreenPorch</b>	0.0003	6.34e-05	4.154	0.000	0.000 0.000
<b>PoolArea</b>	0.0001	0.000	1.158	0.247	-9.39e-05 0.000
<b>YrSold</b>	0.0014	0.000	3.323	0.001	0.001 0.002
<b>MSZoning_FV</b>	0.4987	0.051	9.742	0.000	0.398 0.599
<b>MSZoning_RH</b>	0.4890	0.056	8.791	0.000	0.380 0.598
<b>MSZoning_RL</b>	0.4716	0.048	9.923	0.000	0.378 0.565
<b>MSZoning_RM</b>	0.4338	0.045	9.733	0.000	0.346 0.521
<b>Street_Pave</b>	0.1060	0.056	1.905	0.057	-0.003 0.215
<b>Alley_Pave</b>	0.0445	0.023	1.964	0.050	4.32e-05 0.089
<b>LotShape_IR2</b>	0.0278	0.021	1.303	0.193	-0.014 0.070
<b>LandContour_HLS</b>	0.0440	0.025	1.742	0.082	-0.006 0.094
<b>LandContour_Low</b>	-0.0269	0.031	-0.867	0.386	-0.088 0.034
<b>LandContour_Lvl</b>	0.0278	0.018	1.557	0.120	-0.007 0.063
<b>Utilities_NoSeWa</b>	-0.2121	0.114	-1.853	0.064	-0.437 0.012
<b>LotConfig_CulDSac</b>	0.0377	0.017	2.272	0.023	0.005 0.070
<b>LotConfig_FR2</b>	-0.0255	0.020	-1.284	0.199	-0.065 0.013
<b>LotConfig_FR3</b>	-0.0683	0.086	-0.796	0.426	-0.237 0.100
<b>LotConfig_Inside</b>	-0.0075	0.009	-0.844	0.399	-0.025 0.010
<b>LandSlope_Mod</b>	0.0646	0.021	3.072	0.002	0.023 0.106
<b>LandSlope_Sev</b>	-0.1609	0.059	-2.749	0.006	-0.276 -0.046
<b>Neighborhood_Blueste</b>	-0.1133	0.107	-1.055	0.292	-0.324 0.097
<b>Neighborhood_BrDale</b>	-0.0691	0.039	-1.769	0.077	-0.146 0.008
<b>Neighborhood_CollgCr</b>	-0.0378	0.015	-2.565	0.010	-0.067 -0.009
<b>Neighborhood_Crawfor</b>	0.0836	0.021	3.943	0.000	0.042 0.125
<b>Neighborhood_Edwards</b>	-0.0880	0.017	-5.100	0.000	-0.122 -0.054
<b>Neighborhood_Gilbert</b>	-0.0413	0.019	-2.150	0.032	-0.079 -0.004
<b>Neighborhood_IDOTRR</b>	-0.0200	0.030	-0.665	0.506	-0.079 0.039

<b>Neighborhood_MeadowV</b>	-0.2116	0.041	-5.181	0.000	-0.292 -0.131
<b>Neighborhood_Mitchel</b>	-0.0860	0.022	-3.981	0.000	-0.128 -0.044
<b>Neighborhood_NAmes</b>	-0.0508	0.015	-3.325	0.001	-0.081 -0.021
<b>Neighborhood_NWAmes</b>	-0.0545	0.019	-2.873	0.004	-0.092 -0.017
<b>Neighborhood_NridgHt</b>	0.0823	0.020	4.145	0.000	0.043 0.121
<b>Neighborhood_OldTown</b>	-0.0395	0.021	-1.908	0.057	-0.080 0.001
<b>Neighborhood_Sawyer</b>	-0.0587	0.019	-3.048	0.002	-0.096 -0.021
<b>Neighborhood_SawyerW</b>	-0.0326	0.019	-1.753	0.080	-0.069 0.004
<b>Neighborhood_StoneBr</b>	0.1114	0.027	4.066	0.000	0.058 0.165
<b>Condition1_Feedr</b>	0.0660	0.023	2.914	0.004	0.022 0.110
<b>Condition1_Norm</b>	0.1103	0.018	6.148	0.000	0.075 0.145
<b>Condition1_PosA</b>	0.0706	0.048	1.460	0.145	-0.024 0.165
<b>Condition1_PosN</b>	0.1017	0.036	2.821	0.005	0.031 0.172
<b>Condition1_RRAAn</b>	0.0875	0.031	2.804	0.005	0.026 0.149
<b>Condition1_RRNn</b>	0.1176	0.057	2.074	0.038	0.006 0.229
<b>Condition2_Feedr</b>	0.0817	0.075	1.087	0.277	-0.066 0.229
<b>Condition2_Norm</b>	0.0643	0.053	1.207	0.228	-0.040 0.169
<b>Condition2_PosA</b>	0.2672	0.141	1.897	0.058	-0.009 0.544
<b>Condition2_PosN</b>	-0.8036	0.102	-7.865	0.000	-1.004 -0.603
<b>Condition2_RRAe</b>	-0.5166	0.199	-2.600	0.009	-0.906 -0.127
<b>BldgType_Twnhs</b>	-0.0306	0.024	-1.253	0.210	-0.079 0.017
<b>HouseStyle_1Story</b>	-0.0403	0.015	-2.701	0.007	-0.070 -0.011
<b>HouseStyle_2.5Fin</b>	-0.1303	0.050	-2.612	0.009	-0.228 -0.032
<b>HouseStyle_2.5Unf</b>	0.1052	0.055	1.906	0.057	-0.003 0.214
<b>HouseStyle_2Story</b>	-0.0187	0.016	-1.205	0.229	-0.049 0.012
<b>HouseStyle_SFoyer</b>	-0.0165	0.026	-0.625	0.532	-0.069 0.035
<b>RoofStyle_Mansard</b>	0.0653	0.052	1.250	0.212	-0.037 0.168
<b>RoofStyle_Shed</b>	0.4748	0.155	3.063	0.002	0.171 0.779
<b>RoofMatl_CompShg</b>	2.8404	0.170	16.740	0.000	2.507 3.173
<b>RoofMatl_Membran</b>	3.2300	0.215	15.003	0.000	2.808 3.653
<b>RoofMatl_Metal</b>	3.0784	0.211	14.602	0.000	2.665 3.492
<b>RoofMatl_Roll</b>	2.8548	0.200	14.277	0.000	2.462 3.247
<b>RoofMatl_Tar&amp;Grv</b>	2.8568	0.172	16.601	0.000	2.519 3.194

<b>RoofMatl_WdShake</b>	2.7366	0.197	13.891	0.000	2.350 3.123
<b>RoofMatl_WdShngl</b>	2.9120	0.173	16.859	0.000	2.573 3.251
<b>Exterior1st_AsphShn</b>	-5.808e-16	4.55e-16	-1.277	0.202	-1.47e-15 3.12e-16
<b>Exterior1st_BrkComm</b>	-0.1955	0.090	-2.168	0.030	-0.373 -0.019
<b>Exterior1st_BrkFace</b>	0.1078	0.026	4.219	0.000	0.058 0.158
<b>Exterior1st_CemntBd</b>	-0.1215	0.082	-1.478	0.140	-0.283 0.040
<b>Exterior1st_ImStucc</b>	-2.46e-15	3.77e-16	-6.525	0.000	-3.2e-15 -1.72e-15
<b>Exterior1st_MetalSd</b>	0.0672	0.039	1.732	0.084	-0.009 0.143
<b>Exterior1st_Stone</b>	0.0997	0.147	0.676	0.499	-0.190 0.389
<b>Exterior1st_Wd Sdng</b>	-0.0238	0.015	-1.552	0.121	-0.054 0.006
<b>Exterior2nd_BrkFace</b>	-0.0744	0.033	-2.276	0.023	-0.138 -0.010
<b>Exterior2nd_CmentBd</b>	0.1439	0.082	1.762	0.078	-0.016 0.304
<b>Exterior2nd_HdBoard</b>	-0.0242	0.016	-1.475	0.141	-0.056 0.008
<b>Exterior2nd_MetalSd</b>	-0.0629	0.039	-1.600	0.110	-0.140 0.014
<b>Exterior2nd_Other</b>	-0.1107	0.111	-0.999	0.318	-0.328 0.107
<b>Exterior2nd_Plywood</b>	-0.0268	0.017	-1.547	0.122	-0.061 0.007
<b>Exterior2nd_VinylSd</b>	-0.0017	0.016	-0.104	0.917	-0.033 0.030
<b>MasVnrType_BrkFace</b>	0.0310	0.027	1.135	0.257	-0.023 0.085
<b>MasVnrType_None</b>	0.0283	0.027	1.054	0.292	-0.024 0.081
<b>MasVnrType_Stone</b>	0.0365	0.030	1.229	0.219	-0.022 0.095
<b>ExterQual_Fa</b>	-0.0668	0.043	-1.555	0.120	-0.151 0.018
<b>ExterCond_Fa</b>	-0.0699	0.031	-2.275	0.023	-0.130 -0.010
<b>ExterCond_Gd</b>	-0.0285	0.012	-2.434	0.015	-0.052 -0.006
<b>ExterCond_Po</b>	-0.0825	0.121	-0.684	0.494	-0.319 0.154
<b>Foundation_CBlock</b>	0.0219	0.016	1.414	0.158	-0.009 0.052
<b>Foundation_PConc</b>	0.0412	0.017	2.432	0.015	0.008 0.074
<b>Foundation_Slab</b>	-0.0325	0.041	-0.801	0.423	-0.112 0.047
<b>Foundation_Stone</b>	0.0850	0.051	1.677	0.094	-0.014 0.184
<b>Foundation_Wood</b>	-0.1528	0.068	-2.257	0.024	-0.286 -0.020
<b>BsmtQual_Gd</b>	-0.0230	0.013	-1.734	0.083	-0.049 0.003
<b>BsmtQual_TA</b>	-0.0347	0.015	-2.349	0.019	-0.064 -0.006
<b>BsmtCond_Gd</b>	0.0388	0.026	1.503	0.133	-0.012 0.089
<b>BsmtCond_Po</b>	0.2532	0.106	2.381	0.017	0.045 0.462

<b>BsmtCond_TA</b>	0.0240	0.020	1.188	0.235	-0.016 0.064
<b>BsmtExposure_Gd</b>	0.0237	0.014	1.653	0.099	-0.004 0.052
<b>BsmtExposure_No</b>	-0.0144	0.009	-1.595	0.111	-0.032 0.003
<b>BsmtFinType1_GLQ</b>	0.0163	0.010	1.558	0.120	-0.004 0.037
<b>BsmtFinType1_LwQ</b>	-0.0166	0.016	-1.021	0.308	-0.048 0.015
<b>BsmtFinType2_BLQ</b>	-0.0636	0.025	-2.573	0.010	-0.112 -0.015
<b>BsmtFinType2_LwQ</b>	-0.0278	0.023	-1.195	0.233	-0.073 0.018
<b>BsmtFinType2_Rec</b>	-0.0166	0.020	-0.826	0.409	-0.056 0.023
<b>Heating_GasA</b>	0.1447	0.113	1.277	0.202	-0.078 0.367
<b>Heating_GasW</b>	0.2227	0.120	1.852	0.064	-0.013 0.459
<b>Heating_Grav</b>	0.1645	0.138	1.194	0.233	-0.106 0.435
<b>Heating_OthW</b>	0.1085	0.141	0.768	0.443	-0.169 0.386
<b>Heating_Wall</b>	0.1650	0.134	1.229	0.219	-0.099 0.429
<b>HeatingQC_Fa</b>	-0.0178	0.023	-0.787	0.432	-0.062 0.027
<b>HeatingQC_Gd</b>	-0.0203	0.010	-1.960	0.050	-0.041 2.53e-05
<b>HeatingQC_Po</b>	-0.0869	0.113	-0.769	0.442	-0.309 0.135
<b>HeatingQC_TA</b>	-0.0334	0.010	-3.283	0.001	-0.053 -0.013
<b>CentralAir_Y</b>	0.0543	0.019	2.861	0.004	0.017 0.092
<b>Electrical_FuseF</b>	-0.0282	0.028	-1.010	0.313	-0.083 0.027
<b>Electrical_FuseP</b>	-0.0673	0.078	-0.863	0.388	-0.220 0.086
<b>KitchenQual_Gd</b>	-0.0559	0.015	-3.778	0.000	-0.085 -0.027
<b>KitchenQual_TA</b>	-0.0493	0.016	-3.107	0.002	-0.080 -0.018
<b>Functional_Maj2</b>	-0.3334	0.065	-5.140	0.000	-0.461 -0.206
<b>Functional_Min1</b>	0.0419	0.029	1.443	0.149	-0.015 0.099
<b>Functional_Mod</b>	-0.0781	0.045	-1.745	0.081	-0.166 0.010
<b>Functional_Sev</b>	1.096e-16	6.77e-17	1.619	0.106	-2.33e-17 2.43e-16
<b>Functional_Typ</b>	0.0559	0.019	2.871	0.004	0.018 0.094
<b>FireplaceQu_TA</b>	0.0089	0.010	0.857	0.391	-0.012 0.029
<b>GarageType_Attchd</b>	0.1252	0.046	2.730	0.006	0.035 0.215
<b>GarageType_Basment</b>	0.1132	0.055	2.059	0.040	0.005 0.221
<b>GarageType_BuiltIn</b>	0.1111	0.049	2.287	0.022	0.016 0.206
<b>GarageType_CarPort</b>	0.1350	0.062	2.190	0.029	0.014 0.256
<b>GarageType_Detchd</b>	0.1086	0.046	2.380	0.017	0.019 0.198



GarageFinish_Unf	-0.0105	0.010	-1.039	0.299	-0.030 0.009
GarageQual_Fa	-0.0635	0.022	-2.831	0.005	-0.107 -0.019
GarageQual_Gd	0.0277	0.034	0.810	0.418	-0.039 0.095
GarageQual_Po	-0.1541	0.096	-1.598	0.110	-0.343 0.035
GarageCond_Fa	-0.1381	0.052	-2.670	0.008	-0.240 -0.037
GarageCond_Gd	-0.1003	0.063	-1.580	0.114	-0.225 0.024
GarageCond_TA	-0.0933	0.046	-2.026	0.043	-0.184 -0.003
PavedDrive_Y	0.0141	0.015	0.961	0.337	-0.015 0.043
PoolQC_Gd	0.1410	0.114	1.240	0.215	-0.082 0.364
Fence_GdWo	-0.0345	0.018	-1.967	0.049	-0.069 -8.8e-05
Fence_MnPrv	-0.0107	0.011	-0.942	0.346	-0.033 0.012
Fence_MnWw	-0.0504	0.039	-1.290	0.197	-0.127 0.026
MiscFeature_Othr	-0.0999	0.129	-0.776	0.438	-0.352 0.153
MiscFeature_TenC	-0.2068	0.128	-1.621	0.105	-0.457 0.044
SaleType_CWD	0.0925	0.056	1.661	0.097	-0.017 0.202
SaleType_Con	0.1174	0.077	1.534	0.125	-0.033 0.268
SaleType_ConLD	0.2813	0.050	5.615	0.000	0.183 0.380
SaleType_ConLI	-0.0307	0.055	-0.561	0.575	-0.138 0.077
SaleType_New	0.1045	0.018	5.706	0.000	0.069 0.140
SaleType_Oth	0.0600	0.064	0.934	0.351	-0.066 0.186
SaleCondition_AdjLand	0.1109	0.072	1.549	0.122	-0.030 0.251
SaleCondition_Alloca	0.0448	0.043	1.044	0.297	-0.039 0.129
SaleCondition_Normal	0.0569	0.012	4.600	0.000	0.033 0.081

<b>Omnibus:</b>	300.168	<b>Durbin-Watson:</b>	2.043
<b>Prob(Omnibus):</b>	0.000	<b>Jarque-Bera (JB):</b>	3855.185
<b>Skew:</b>	-0.809	<b>Prob(JB):</b>	0.00
<b>Kurtosis:</b>	11.752	<b>Cond. No.</b>	1.02e+17

```
In [457]: #Predictions based off our model
preds = result.predict(local_test_x)
preds[0:5]
```

```
Out[457]: array([ 12.26705006,  11.3662339 ,  11.86414711,  12.45139966,  11.7648
1408])
```

```
In [458]: #RMSE
rmse = (sum((local_test_y - preds) ** 2) / preds.shape[0]) ** 0.5
rmse
```

```
Out[458]: 0.11376542632433698
```

## Actual Kaggle Dataset

```
In [459]: #Import the test dataset (doesn't contain the test SalePrice values)
actual_test = pd.read_csv('./data/test.csv')
actual_test.shape
```

```
Out[459]: (1459, 80)
```

```
In [460]: #Use the whole dataset to re-train
#Removing SalesPrices (is the predicting y-variable)
train_y = data['SalePrice']
train_x = data.drop('SalePrice', axis = 1)
```

```
In [461]: #Feature engineer test dataset predictor columns
actual_test = pd.get_dummies(actual_test, drop_first=True)
actual_test = actual_test.fillna(actual_test.mean())
actual_test.drop(['MasVnrArea', 'LowQualFinSF', 'BsmtHalfBath', 'Condition1_RRAe', 'BldgType_Duplex', 'HouseStyle_1.5Unf', 'RoofStyle_Gable', 'RoofStyle_Gambrel', 'RoofStyle_Hip', 'Exterior1st_CBlock', 'Exterior1st_HdBoard', 'Exterior1st_Plywood', 'Exterior1st_Stucco', 'Exterior1st_WdShng', 'Exterior2nd_BrkCmn', 'Exterior2nd_CBlock', 'Exterior2nd_ImStucc', 'Exterior2nd_Stone', 'Exterior2nd_WdSdng', 'BsmtExposure_Mn', 'BsmtFinType1_Unf', 'BsmtFinType2_GLQ', 'BsmtFinType2_Unf', 'Electrical_SBrkr', 'KitchenQual_Fa', 'FireplaceQu_Fa', 'FireplaceQu_Po', 'GarageQual_TA', 'GarageCond_Po', 'PavedDrive_P', 'SaleCondition_Family'], axis=1, inplace=True)
actual_test.drop(['2ndFlrSF', 'EnclosedPorch', 'MiscVal', 'MoSold', 'LotShape_IR3', 'LotShape_Reg', 'Neighborhood_BrkSide', 'Neighborhood_ClearCr', 'Neighborhood_NPkVill', 'Neighborhood_NoRidge', 'Neighborhood_SWISU', 'Neighborhood_Somerst', 'Neighborhood_Timber', 'Neighborhood_Veenker', 'Condition1_RRNe', 'BldgType_2fmCon', 'BldgType_TwnhsE', 'HouseStyle_SLvl', 'Exterior1st_VinylSd', 'Exterior2nd_AsphShn', 'Exterior2nd_Stucco', 'Exterior2nd_WdShng', 'ExterQual_Gd', 'ExterQual_TA', 'ExterCond_TA', 'BsmtQual_Fa', 'BsmtFinType1_BLQ', 'BsmtFinType1_Rec', 'Functional_Min2', 'FireplaceQu_Gd', 'GarageFinish_RFn', 'MiscFeature_Shed', 'SaleType_ConLw', 'SaleType_WD', 'SaleCondition_Partial'], axis=1, inplace=True)

#Also remove columns not in actual_test from train_x
train_x = train_x.iloc[:, train_x.columns.isin(actual_test.columns)]
train_x.shape
```

```
Out[461]: (1460, 162)
```

```
In [462]: #Fitting model
actual_clf = sm.OLS(train_y, train_x)
actual_result = actual_clf.fit()
actual_result.summary()

actual_preds = actual_result.predict(actual_test)
actual_preds[0:5]
```

```
Out[462]: array([ 11.66476646,  11.9162187 ,  12.08137814,  12.19944302,  12.1846
 3637])
```

```
In [463]: # Converting predictions back to actual housing values
actual_preds = np.exp(actual_preds) - 1
actual_preds
```

```
Out[463]: array([ 116396.51463577,  149673.57267838,  176552.3083296 , ...,
 161159.65985258,  110287.76978393,  222581.85988514])
```

```
In [464]: #Exporting to CSV
submission = pd.concat([pd.Series(actual_test['Id']), pd.Series(actual_p
reds, name = 'SalePrice')], axis = 1)
submission.to_csv('submission.csv', index = False)
```

## Model Interpretation

From the results of the linear regression significance tests, the most important features ar predicting house sale price include features related to the size of the house (Lot Area, SF of living area), Overall Quality and Condition, its newness (Year built and remodeled), and location (MSZoning, some neighborhoods).

Those are mostly expected features to be important, but one surprising feature that the model indicates to be very important is roof material (P=0.000 for all dummies of RoofMatl)

In [ ]: