

```
In [1]: import pandas as pd
import numpy as np
import statsmodels.api as sm
import matplotlib.pyplot as plt
import seaborn as sns

from sklearn.model_selection import train_test_split
from sklearn.preprocessing import OneHotEncoder
from sklearn.impute import SimpleImputer
from sklearn.compose import ColumnTransformer
from sklearn.pipeline import Pipeline
from sklearn.metrics import mean_absolute_error, mean_squared_error, r2_score
from statsmodels.stats.outliers_influence import variance_inflation_factor
```

```
In [11]: # Load Data
df = pd.read_excel(r"D:\internship\DS Internship Data.xlsx", sheet_name="Data")

df.drop(columns=["Store ID"], inplace=True)

datetime_cols = [col for col in df.columns if np.issubdtype(df[col].dtype, np.datetime64)]
df.drop(columns=datetime_cols, inplace=True)
y = df["Sales"]
X = df.drop(columns=["Sales"])
categorical_cols = X.select_dtypes(include="object").columns.tolist()
for col in categorical_cols:
    X[col] = X[col].astype(str)

numerical_cols = X.select_dtypes(include=np.number).columns.tolist()
```

```
In [12]: numeric_transformer = SimpleImputer(strategy="median")

categorical_transformer = Pipeline([
    ("imputer", SimpleImputer(strategy="most_frequent")),
    ("onehot", OneHotEncoder(drop="first", sparse_output=False, handle_unknown="ignore"))
])

preprocessor = ColumnTransformer([
    ("num", numeric_transformer, numerical_cols),
    ("cat", categorical_transformer, categorical_cols)
])

X_processed = preprocessor.fit_transform(X)

encoded_cat_cols = preprocessor.named_transformers_['cat'].named_steps['onehot'].get_feature_names_out()
feature_names = numerical_cols + list(encoded_cat_cols)

X_df = pd.DataFrame(X_processed, columns=feature_names)
```

```
In [13]: # Splitting of dataset
X_train, X_test, y_train, y_test = train_test_split(X_df, y, test_size=0.2, random_state=42)

X_train_sm = sm.add_constant(X_train)

model = sm.OLS(y_train, X_train_sm).fit()

print(model.summary())
```

OLS Regression Results

```

=====
Dep. Variable:          Sales      R-squared:          0.646
Model:                  OLS        Adj. R-squared:       0.485
Method:                 Least Squares  F-statistic:        4.013
Date:                  Thu, 29 May 2025  Prob (F-statistic):    1.25e-29
Time:                  11:12:53    Log-Likelihood:     -7707.1
No. Observations:      552        AIC:                1.576e+04
Df Residuals:          379        BIC:                1.651e+04
Df Model:              172
Covariance Type:       nonrobust
=====

```

```

=====
                                coef      std err          t      P>|t|      [0.02
5      0.975]
-----
const                        8.395e+05   4.89e+05     1.717     0.087   -1.22e+0
5      1.8e+06
Total Store Sq Ft           185.5647    82.498     2.249     0.025    23.35
4      347.776
Sales Floor Sq Ft          -77.8286    90.133    -0.863     0.388   -255.05
2      99.394
Girls Sq Ft                 18.4499    84.856     0.217     0.828   -148.39
8      185.298
Guys Sq Ft                 108.3261   116.800     0.927     0.354   -121.33
0      337.983
Accessories Sq Ft         -133.7613   113.231    -1.181     0.238   -356.40
1      88.878
Population Est CrYr        -2.747e+04  3.55e+04    -0.774     0.439   -9.73e+0
4      4.23e+04
Transient Pop Est           5.6451     2.700     2.091     0.037     0.33
7      10.953
Seasonal Pop Est           0.0227     3.300     0.007     0.995    -6.46
7      6.512
Sex Males CrYr             2.747e+04  3.55e+04     0.774     0.439   -4.23e+0
4      9.73e+04
Sex Females CrYr          2.747e+04  3.55e+04     0.774     0.439   -4.23e+0
4      9.73e+04
Age Average CrYr          1739.4376   1.07e+04     0.163     0.871   -1.92e+0
4      2.27e+04
Unemployment Rate         1991.1837   1.42e+04     0.141     0.888   -2.59e+0
4      2.98e+04
Average Household Income CrYr  10.2890     4.411     2.332     0.020     1.61
5      18.963
Householder under 25 years:  1.552e+04  1.56e+04     0.993     0.321   -1.52e+0
4      4.62e+04
Householder 25 to 44 years: -4428.9690  1.47e+04    -0.302     0.763   -3.32e+0
4      2.44e+04
Family Inc: Aggregate Average -2.265e-06  4.76e-05    -0.048     0.962   -9.58e-0
5      9.12e-05
Per Capita Income CrYr     -26.5985    11.232    -2.368     0.018   -48.68
4      -4.513
Wealth: Average Household   -0.0002     0.002    -0.120     0.905    -0.00
4      0.003
Poverty: In poverty        9.6906     14.199     0.683     0.495   -18.22
7      37.608
Male 14 to 35 CrYr         1.4836     15.877     0.093     0.926   -29.73
5      32.702
Female 14 to 35 CrYr       0.0347     16.197     0.002     0.998   -31.81

```

3	31.883					
Total Pop 14 to 35 CrYr		1.5079	2.188	0.689	0.491	-2.79
4	5.810					
Under 25 Low Income		-1.552e+04	1.56e+04	-0.993	0.321	-4.62e+0
4	1.52e+04					
Under 25 Med Income		-1.549e+04	1.56e+04	-0.991	0.322	-4.62e+0
4	1.52e+04					
Under 25 High Income		-1.564e+04	1.56e+04	-1.000	0.318	-4.64e+0
4	1.51e+04					
25 to 44 Low Income		4422.5932	1.47e+04	0.302	0.763	-2.44e+0
4	3.32e+04					
25 to 44 Med Income		4446.0178	1.47e+04	0.303	0.762	-2.44e+0
4	3.33e+04					
25 to 44 High Income		4435.7796	1.47e+04	0.303	0.762	-2.44e+0
4	3.32e+04					
Pop class_Suburban		-9.206e+04	8.01e+04	-1.150	0.251	-2.5e+0
5	6.54e+04					
Pop class_Urban		6.873e+04	7.17e+04	0.958	0.338	-7.23e+0
4	2.1e+05					
Centre Type_Outlet		2.067e+05	8.89e+04	2.326	0.021	3.2e+0
4	3.81e+05					
Centre Type_Strip		-1.064e+05	4.28e+04	-2.484	0.013	-1.91e+0
5	-2.22e+04					
Climate_Frigid		-1.997e+04	7.52e+04	-0.266	0.791	-1.68e+0
5	1.28e+05					
Climate_Hot		4.449e+04	6.61e+04	0.673	0.501	-8.54e+0
4	1.74e+05					
Climate_Moderate		4.571e+04	5e+04	0.914	0.361	-5.27e+0
4	1.44e+05					
Climate_Warm		3.424e+04	5.44e+04	0.629	0.530	-7.28e+0
4	1.41e+05					
Guys Segmentation_On Trend		-1.091e+05	2.02e+05	-0.539	0.590	-5.07e+0
5	2.89e+05					
Guys Segmentation_On trend		-1.701e+05	2.47e+05	-0.690	0.491	-6.55e+0
5	3.15e+05					
Girls Segmentation_On Trend		-9.524e+04	2.02e+05	-0.472	0.637	-4.92e+0
5	3.02e+05					
ChangeMade_no change		-6.539e+04	1.13e+05	-0.580	0.562	-2.87e+0
5	1.56e+05					
ChangeMade_relocation		3.826e+04	1.13e+05	0.340	0.734	-1.83e+0
5	2.6e+05					
ChangeMade_remodel		1.091e+05	2.86e+05	0.382	0.703	-4.53e+0
5	6.71e+05					
ChangeDate_2006-02-22 00:00:00		1.616e+05	4.07e+05	0.397	0.691	-6.38e+0
5	9.62e+05					
ChangeDate_2007-02-08 00:00:00		-17.6545	7.153	-2.468	0.014	-31.72
0	-3.589					
ChangeDate_2007-03-13 00:00:00		-3.449e+05	4.38e+05	-0.788	0.431	-1.21e+0
6	5.16e+05					
ChangeDate_2007-07-13 00:00:00		-5.944e+05	4.79e+05	-1.242	0.215	-1.54e+0
6	3.47e+05					
ChangeDate_2007-07-19 00:00:00		-5.215e+05	4.73e+05	-1.104	0.270	-1.45e+0
6	4.08e+05					
ChangeDate_2008-02-26 00:00:00		-7.79e+04	4.12e+05	-0.189	0.850	-8.88e+0
5	7.33e+05					
ChangeDate_2008-06-24 00:00:00		-6.411e+04	4.11e+05	-0.156	0.876	-8.72e+0
5	7.44e+05					
ChangeDate_2008-06-25 00:00:00		-3.55e+05	3.98e+05	-0.892	0.373	-1.14e+0
6	4.28e+05					
ChangeDate_2008-10-24 00:00:00		-3.052e+04	4.04e+05	-0.076	0.940	-8.24e+0

5	7.63e+05						
ChangeDate_2008-11-17 00:00:00	2.602e+05	4.13e+05	0.630	0.529	-5.52e+0		
5	1.07e+06						
ChangeDate_2009-02-12 00:00:00	6.745e+04	3.96e+05	0.170	0.865	-7.11e+0		
5	8.46e+05						
ChangeDate_2009-03-12 00:00:00	-1.75e+05	3.96e+05	-0.442	0.659	-9.54e+0		
5	6.04e+05						
ChangeDate_2009-06-04 00:00:00	-6.178e+05	5e+05	-1.236	0.217	-1.6e+0		
6	3.65e+05						
ChangeDate_2009-09-17 00:00:00	-1.7086	5.520	-0.310	0.757	-12.56		
1	9.144						
ChangeDate_2009-09-18 00:00:00	-3.326e+05	3.99e+05	-0.833	0.406	-1.12e+0		
6	4.53e+05						
ChangeDate_2009-10-01 00:00:00	-1.188e+05	3.27e+05	-0.363	0.717	-7.62e+0		
5	5.24e+05						
ChangeDate_2010-05-06 00:00:00	-5.753e+05	3.97e+05	-1.451	0.148	-1.36e+0		
6	2.05e+05						
ChangeDate_2010-06-24 00:00:00	1.587e+05	4.91e+05	0.324	0.746	-8.06e+0		
5	1.12e+06						
ChangeDate_2010-07-22 00:00:00	11.6049	7.639	1.519	0.130	-3.41		
6	26.626						
ChangeDate_2010-09-02 00:00:00	5.3696	8.736	0.615	0.539	-11.80		
7	22.547						
ChangeDate_2010-10-07 00:00:00	-4.99e+05	3.41e+05	-1.463	0.144	-1.17e+0		
6	1.72e+05						
ChangeDate_2010-10-14 00:00:00	-4.536e+05	3.98e+05	-1.139	0.255	-1.24e+0		
6	3.29e+05						
ChangeDate_2011-02-10 00:00:00	-4.921e+05	3.97e+05	-1.239	0.216	-1.27e+0		
6	2.89e+05						
ChangeDate_2011-02-25 00:00:00	-2.322e+05	3.99e+05	-0.582	0.561	-1.02e+0		
6	5.53e+05						
ChangeDate_2011-03-31 00:00:00	-2.979e+05	3.95e+05	-0.754	0.451	-1.07e+0		
6	4.79e+05						
ChangeDate_2011-04-07 00:00:00	1.291e+05	4.14e+05	0.312	0.755	-6.85e+0		
5	9.43e+05						
ChangeDate_2011-04-29 00:00:00	-3.86e+05	3.98e+05	-0.970	0.333	-1.17e+0		
6	3.97e+05						
ChangeDate_2011-06-23 00:00:00	-2.51e+05	2.51e+05	-1.001	0.317	-7.44e+0		
5	2.42e+05						
ChangeDate_2011-06-30 00:00:00	1.2840	6.615	0.194	0.846	-11.72		
3	14.291						
ChangeDate_2011-09-29 00:00:00	-3.551e+05	3.92e+05	-0.905	0.366	-1.13e+0		
6	4.16e+05						
ChangeDate_2011-10-06 00:00:00	-1.278e+04	4.01e+05	-0.032	0.975	-8.02e+0		
5	7.77e+05						
ChangeDate_2011-10-13 00:00:00	14.6549	6.365	2.302	0.022	2.13		
9	27.170						
ChangeDate_2011-11-03 00:00:00	-5.563e+05	3.95e+05	-1.407	0.160	-1.33e+0		
6	2.21e+05						
ChangeDate_2012-03-29 00:00:00	5.143e+05	4.21e+05	1.222	0.222	-3.13e+0		
5	1.34e+06						
ChangeDate_2012-04-13 00:00:00	2.329e+05	3.13e+05	0.744	0.458	-3.83e+0		
5	8.49e+05						
ChangeDate_2012-05-17 00:00:00	-4.563e+05	3.96e+05	-1.151	0.250	-1.24e+0		
6	3.23e+05						
ChangeDate_2012-07-19 00:00:00	-3.405e+05	4e+05	-0.850	0.396	-1.13e+0		
6	4.47e+05						
ChangeDate_2012-08-30 00:00:00	-3.816e+05	4.08e+05	-0.936	0.350	-1.18e+0		
6	4.2e+05						
ChangeDate_2012-09-27 00:00:00	-4.128e+05	4.14e+05	-0.997	0.319	-1.23e+0		

6	4.01e+05						
ChangeDate_2012-10-19 00:00:00	2.629e+04	3.93e+05	0.067	0.947	-7.47e+0		
5	8e+05						
ChangeDate_2012-10-25 00:00:00	2.346e+04	4.19e+05	0.056	0.955	-8e+0		
5	8.47e+05						
ChangeDate_2012-11-08 00:00:00	-1.543e+05	3.17e+05	-0.487	0.626	-7.77e+0		
5	4.68e+05						
ChangeDate_2013-03-15 00:00:00	-4.713e+05	4.07e+05	-1.157	0.248	-1.27e+0		
6	3.3e+05						
ChangeDate_2013-05-09 00:00:00	-5.227e+05	4.01e+05	-1.303	0.193	-1.31e+0		
6	2.66e+05						
ChangeDate_2013-08-22 00:00:00	-16.2718	6.845	-2.377	0.018	-29.73		
2	-2.812						
ChangeDate_2013-09-26 00:00:00	-2.31e+05	4.23e+05	-0.546	0.586	-1.06e+0		
6	6.01e+05						
ChangeDate_2013-10-31 00:00:00	-4.225e+05	4.04e+05	-1.046	0.296	-1.22e+0		
6	3.71e+05						
ChangeDate_2013-11-07 00:00:00	-4.565e+05	3.03e+05	-1.505	0.133	-1.05e+0		
6	1.4e+05						
ChangeDate_2014-05-22 00:00:00	-5.041e+05	4.02e+05	-1.253	0.211	-1.3e+0		
6	2.87e+05						
ChangeDate_2014-06-05 00:00:00	-5.689e+05	3.96e+05	-1.437	0.152	-1.35e+0		
6	2.09e+05						
ChangeDate_2014-06-13 00:00:00	-2.296e+05	3.97e+05	-0.578	0.564	-1.01e+0		
6	5.52e+05						
ChangeDate_2014-07-03 00:00:00	-2.919e+05	3.37e+05	-0.867	0.387	-9.54e+0		
5	3.7e+05						
ChangeDate_2014-07-31 00:00:00	-7.9605	9.855	-0.808	0.420	-27.33		
7	11.416						
ChangeDate_2014-10-02 00:00:00	-9.7317	8.921	-1.091	0.276	-27.27		
3	7.810						
ChangeDate_2014-10-23 00:00:00	8.974e+04	3.95e+05	0.227	0.820	-6.87e+0		
5	8.66e+05						
ChangeDate_2014-10-30 00:00:00	1.826e+05	3.96e+05	0.461	0.645	-5.96e+0		
5	9.61e+05						
ChangeDate_2014-11-06 00:00:00	-3.474e+05	4.34e+05	-0.800	0.424	-1.2e+0		
6	5.06e+05						
ChangeDate_2014-11-13 00:00:00	3.678e+05	3.98e+05	0.923	0.356	-4.15e+0		
5	1.15e+06						
ChangeDate_2014-11-23 00:00:00	-1.3109	8.715	-0.150	0.881	-18.44		
6	15.824						
ChangeDate_2015-03-26 00:00:00	2.92e+05	4e+05	0.730	0.466	-4.94e+0		
5	1.08e+06						
ChangeDate_2015-04-16 00:00:00	-5.907e+05	4.19e+05	-1.409	0.160	-1.41e+0		
6	2.33e+05						
ChangeDate_2015-04-23 00:00:00	7.891e+05	4.18e+05	1.886	0.060	-3.34e+0		
4	1.61e+06						
ChangeDate_2015-04-29 00:00:00	6.915e+04	3.59e+05	0.193	0.847	-6.36e+0		
5	7.75e+05						
ChangeDate_2015-04-30 00:00:00	-2.722e+05	3.19e+05	-0.853	0.394	-9e+0		
5	3.55e+05						
ChangeDate_2015-05-20 00:00:00	3.57e+04	4.16e+05	0.086	0.932	-7.83e+0		
5	8.54e+05						
ChangeDate_2015-05-28 00:00:00	-2.299e+04	4.19e+05	-0.055	0.956	-8.47e+0		
5	8.01e+05						
ChangeDate_2015-05-29 00:00:00	1.9700	10.299	0.191	0.848	-18.28		
1	22.221						
ChangeDate_2015-07-23 00:00:00	-5.537e+05	3.99e+05	-1.389	0.166	-1.34e+0		
6	2.3e+05						
ChangeDate_2015-07-28 00:00:00	-5.575e+04	4.27e+05	-0.130	0.896	-8.96e+0		

5	7.84e+05							
ChangeDate_2015-07-30 00:00:00	1.889e+05	4.2e+05	0.450	0.653	-6.37e+0			
5	1.02e+06							
ChangeDate_2015-08-13 00:00:00	-1.433e+05	3.43e+05	-0.417	0.677	-8.18e+0			
5	5.32e+05							
ChangeDate_2015-08-19 00:00:00	4.162e+05	4.26e+05	0.976	0.329	-4.22e+0			
5	1.25e+06							
ChangeDate_2015-09-03 00:00:00	2.653e+05	3.18e+05	0.835	0.404	-3.6e+0			
5	8.9e+05							
ChangeDate_2015-09-17 00:00:00	-4.982e+05	3.27e+05	-1.523	0.129	-1.14e+0			
6	1.45e+05							
ChangeDate_2015-09-24 00:00:00	9556.4855	3.23e+05	0.030	0.976	-6.26e+0			
5	6.45e+05							
ChangeDate_2015-09-29 00:00:00	-6.5783	8.549	-0.769	0.442	-23.38			
8	10.232							
ChangeDate_2015-09-30 00:00:00	2.406e+05	4.18e+05	0.576	0.565	-5.81e+0			
5	1.06e+06							
ChangeDate_2015-10-01 00:00:00	-6.186e+04	4e+05	-0.155	0.877	-8.48e+0			
5	7.24e+05							
ChangeDate_2015-10-08 00:00:00	5.546e+05	3.21e+05	1.725	0.085	-7.75e+0			
4	1.19e+06							
ChangeDate_2015-10-15 00:00:00	1.518e+05	4.11e+05	0.370	0.712	-6.56e+0			
5	9.59e+05							
ChangeDate_2015-10-16 00:00:00	4.459e+05	4.16e+05	1.073	0.284	-3.71e+0			
5	1.26e+06							
ChangeDate_2015-10-22 00:00:00	1.306e+05	4.03e+05	0.324	0.746	-6.62e+0			
5	9.23e+05							
ChangeDate_2015-10-24 00:00:00	4.051e+05	4.14e+05	0.979	0.328	-4.09e+0			
5	1.22e+06							
ChangeDate_2015-10-29 00:00:00	1.234e+05	2.73e+05	0.452	0.652	-4.14e+0			
5	6.61e+05							
ChangeDate_2015-10-30 00:00:00	5.71e+05	3.97e+05	1.437	0.152	-2.1e+0			
5	1.35e+06							
ChangeDate_2015-11-04 00:00:00	1.174e+04	3.43e+05	0.034	0.973	-6.62e+0			
5	6.86e+05							
ChangeDate_2015-11-05 00:00:00	3.3048	10.605	0.312	0.755	-17.54			
8	24.157							
ChangeDate_2015-11-07 00:00:00	-2.638e+05	4.05e+05	-0.652	0.515	-1.06e+0			
6	5.32e+05							
ChangeDate_2015-11-11 00:00:00	-2.161e+05	3.07e+05	-0.703	0.482	-8.21e+0			
5	3.88e+05							
ChangeDate_2015-11-12 00:00:00	-1.742e+05	2.71e+05	-0.642	0.521	-7.08e+0			
5	3.6e+05							
ChangeDate_2015-11-18 00:00:00	-1.729e+05	4.17e+05	-0.414	0.679	-9.94e+0			
5	6.48e+05							
ChangeDate_2015-11-20 00:00:00	4.012e+05	4.35e+05	0.923	0.357	-4.54e+0			
5	1.26e+06							
ChangeDate_2015-12-03 00:00:00	-2.29e+05	4.02e+05	-0.569	0.570	-1.02e+0			
6	5.62e+05							
ChangeDate_2016-02-04 00:00:00	-2.622e+05	3.94e+05	-0.665	0.506	-1.04e+0			
6	5.13e+05							
ChangeDate_2016-02-11 00:00:00	5.157e+05	4.05e+05	1.272	0.204	-2.81e+0			
5	1.31e+06							
ChangeDate_2016-03-03 00:00:00	-5.135e+05	4.15e+05	-1.238	0.217	-1.33e+0			
6	3.02e+05							
ChangeDate_2016-03-10 00:00:00	-2.918e+05	2.95e+05	-0.990	0.323	-8.71e+0			
5	2.88e+05							
ChangeDate_2016-03-16 00:00:00	9.29e+04	4.21e+05	0.221	0.825	-7.34e+0			
5	9.2e+05							
ChangeDate_2016-03-21 00:00:00	-2.148e+05	4.17e+05	-0.515	0.607	-1.03e+0			

6	6.05e+05						
ChangeDate_2016-03-23 00:00:00	-3.657e+05	4.22e+05	-0.866	0.387	-1.2e+0		
6	4.64e+05						
ChangeDate_2016-03-24 00:00:00	1.915e+05	2.85e+05	0.671	0.503	-3.69e+0		
5	7.52e+05						
ChangeDate_2016-04-07 00:00:00	-3e+05	4.01e+05	-0.748	0.455	-1.09e+0		
6	4.88e+05						
ChangeDate_2016-04-14 00:00:00	-1.744e+05	3.2e+05	-0.544	0.587	-8.04e+0		
5	4.56e+05						
ChangeDate_2016-04-21 00:00:00	-2.119e+05	3.39e+05	-0.624	0.533	-8.79e+0		
5	4.55e+05						
ChangeDate_2016-05-12 00:00:00	4.5565	6.541	0.697	0.486	-8.30		
4	17.417						
ChangeDate_2016-05-19 00:00:00	-1.677e+04	4.15e+05	-0.040	0.968	-8.33e+0		
5	8e+05						
ChangeDate_2016-05-25 00:00:00	2.673e+05	4.22e+05	0.634	0.527	-5.62e+0		
5	1.1e+06						
ChangeDate_2016-05-26 00:00:00	6.242e+04	4.17e+05	0.150	0.881	-7.57e+0		
5	8.82e+05						
ChangeDate_2016-06-08 00:00:00	-6.041e+05	4.01e+05	-1.506	0.133	-1.39e+0		
6	1.85e+05						
ChangeDate_2016-06-09 00:00:00	1.573e+05	2.92e+05	0.539	0.590	-4.17e+0		
5	7.31e+05						
ChangeDate_2016-06-16 00:00:00	6.537e+05	4.02e+05	1.624	0.105	-1.38e+0		
5	1.44e+06						
ChangeDate_2016-06-23 00:00:00	-1.728e+05	4.26e+05	-0.405	0.685	-1.01e+0		
6	6.65e+05						
ChangeDate_2016-06-30 00:00:00	-5.099e+05	3.18e+05	-1.601	0.110	-1.14e+0		
6	1.16e+05						
ChangeDate_2016-07-14 00:00:00	-4.211e+05	3.17e+05	-1.330	0.184	-1.04e+0		
6	2.01e+05						
ChangeDate_2016-07-21 00:00:00	-4.846e+05	3.97e+05	-1.220	0.223	-1.27e+0		
6	2.96e+05						
ChangeDate_2016-07-27 00:00:00	2.726e+05	4.16e+05	0.655	0.513	-5.46e+0		
5	1.09e+06						
ChangeDate_2016-07-28 00:00:00	1.395e+05	2.82e+05	0.495	0.621	-4.15e+0		
5	6.94e+05						
ChangeDate_2016-08-04 00:00:00	-2.053e+04	2.52e+05	-0.081	0.935	-5.16e+0		
5	4.75e+05						
ChangeDate_2016-08-11 00:00:00	-1.835e+05	3.98e+05	-0.461	0.645	-9.66e+0		
5	5.99e+05						
ChangeDate_2016-08-25 00:00:00	-1.939e+05	4.01e+05	-0.483	0.629	-9.83e+0		
5	5.95e+05						
ChangeDate_2016-09-01 00:00:00	11.6785	10.363	1.127	0.260	-8.69		
9	32.056						
ChangeDate_2016-09-22 00:00:00	6.869e+04	4.03e+05	0.171	0.865	-7.23e+0		
5	8.6e+05						
ChangeDate_2016-09-29 00:00:00	3.1603	10.933	0.289	0.773	-18.33		
7	24.658						
ChangeDate_2016-10-05 00:00:00	-7.1301	10.917	-0.653	0.514	-28.59		
6	14.336						
ChangeDate_2016-10-06 00:00:00	-1.839e+05	4.02e+05	-0.458	0.647	-9.74e+0		
5	6.06e+05						
ChangeDate_2016-10-13 00:00:00	-6.041e+04	3.45e+05	-0.175	0.861	-7.39e+0		
5	6.18e+05						
ChangeDate_2016-10-20 00:00:00	2.193e+05	4.11e+05	0.533	0.594	-5.89e+0		
5	1.03e+06						
ChangeDate_2016-10-27 00:00:00	-1.345e+05	3.15e+05	-0.427	0.669	-7.53e+0		
5	4.84e+05						
ChangeDate_2016-11-03 00:00:00	-8.3883	10.887	-0.771	0.441	-29.79		

```

4      13.017
ChangeDate_2016-11-05 00:00:00 -3.071e+05  4.09e+05  -0.751  0.453  -1.11e+0
6      4.97e+05
ChangeDate_2016-11-09 00:00:00 -3.312e+05  3.94e+05  -0.840  0.401  -1.11e+0
6      4.44e+05
ChangeDate_2016-11-10 00:00:00 -1.348e+05  2.48e+05  -0.544  0.587  -6.22e+0
5      3.53e+05
ChangeDate_2016-11-17 00:00:00  3.371e+05  3.97e+05  0.850  0.396  -4.43e+0
5      1.12e+06
ChangeDate_2016-11-18 00:00:00      3.9907    10.796  0.370  0.712  -17.23
6      25.218
ChangeDate_2016-12-08 00:00:00  6.682e+04  3.94e+05  0.169  0.866  -7.08e+0
5      8.42e+05
ChangeDate_2017-03-23 00:00:00  -8.7877    10.916  -0.805  0.421  -30.25
1      12.676
ChangeDate_2017-05-02 00:00:00 -5.029e+05  4.27e+05  -1.177  0.240  -1.34e+0
6      3.37e+05
ChangeDate_2017-05-03 00:00:00 -2.594e+05  4.88e+05  -0.531  0.595  -1.22e+0
6      7e+05
ChangeDate_2017-05-09 00:00:00 -1.704e+05  4.98e+05  -0.343  0.732  -1.15e+0
6      8.08e+05
ChangeDate_2017-05-11 00:00:00 -8.289e+04  4.02e+05  -0.206  0.837  -8.73e+0
5      7.08e+05
ChangeDate_2017-05-18 00:00:00 -5.942e+05  4.06e+05  -1.463  0.144  -1.39e+0
6      2.04e+05
ChangeDate_2017-05-25 00:00:00 -2.814e+05  3.93e+05  -0.716  0.474  -1.05e+0
6      4.91e+05
ChangeDate_2017-06-13 00:00:00 -4.589e+05  4.92e+05  -0.932  0.352  -1.43e+0
6      5.09e+05
ChangeDate_2017-11-24 00:00:00 -5.289e+05  5.12e+05  -1.034  0.302  -1.53e+0
6      4.77e+05
ChangeDate_2018-01-25 00:00:00  3.884e+05  3.09e+05  1.256  0.210  -2.2e+0
5      9.97e+05
ChangeDate_2018-02-03 00:00:00 -1.503e+05  4.06e+05  -0.370  0.712  -9.49e+0
5      6.48e+05
ChangeDate_2018-04-19 00:00:00  208.5277  4.09e+05  0.001  1.000  -8.04e+0
5      8.04e+05
ChangeDate_2018-05-10 00:00:00 -2.194e+05  4e+05  -0.549  0.584  -1.01e+0
6      5.67e+05
ChangeDate_2019-04-04 00:00:00  4.521e+05  3.94e+05  1.148  0.252  -3.22e+0
5      1.23e+06
ChangeDate_2019-09-12 00:00:00  9.476e+04  3.5e+05  0.270  0.787  -5.94e+0
5      7.84e+05
ChangeDate_2019-09-25 00:00:00 -9.481e-05  0.000  -0.875  0.382  -0.00
0      0.000
ChangeDate_2020-01-23 00:00:00 -2.516e+05  4.38e+05  -0.575  0.566  -1.11e+0
6      6.09e+05
ChangeDate_nan      -6.539e+04  1.13e+05  -0.580  0.562  -2.87e+0
5      1.56e+05

```

```

=====
Omnibus:      119.868  Durbin-Watson:      1.911
Prob(Omnibus): 0.000  Jarque-Bera (JB):      391.427
Skew:         1.000  Prob(JB):      1.01e-85
Kurtosis:     6.608  Cond. No.      3.05e+26
=====

```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The smallest eigenvalue is $1.35e-30$. This might indicate that there are strong multicollinearity problems or that the design matrix is singular.

```
In [14]: # Splitting of dataset
X_train, X_test, y_train, y_test = train_test_split(X_df, y, test_size=0.2, random_state=42)

X_train_sm = sm.add_constant(X_train)

model = sm.OLS(y_train, X_train_sm).fit()

print(model.summary())
```

OLS Regression Results

```

=====
Dep. Variable:      Sales      R-squared:      0.646
Model:              OLS       Adj. R-squared:  0.485
Method:              Least Squares  F-statistic:    4.013
Date:                Thu, 29 May 2025  Prob (F-statistic): 1.25e-29
Time:                11:13:14  Log-Likelihood: -7707.1
No. Observations:    552      AIC:              1.576e+04
Df Residuals:        379      BIC:              1.651e+04
Df Model:            172
Covariance Type:     nonrobust
=====

```

```

=====
                    coef      std err          t      P>|t|      [0.02
5      0.975]
-----
const                8.395e+05   4.89e+05     1.717     0.087   -1.22e+0
5      1.8e+06
Total Store Sq Ft      185.5647     82.498     2.249     0.025    23.35
4      347.776
Sales Floor Sq Ft     -77.8286     90.133    -0.863     0.388   -255.05
2      99.394
Girls Sq Ft           18.4499     84.856     0.217     0.828   -148.39
8      185.298
Guys Sq Ft           108.3261    116.800     0.927     0.354   -121.33
0      337.983
Accessories Sq Ft    -133.7613    113.231    -1.181     0.238   -356.40
1      88.878
Population Est CrYr   -2.747e+04   3.55e+04    -0.774     0.439   -9.73e+0
4      4.23e+04
Transient Pop Est      5.6451       2.700     2.091     0.037     0.33
7      10.953
Seasonal Pop Est      0.0227       3.300     0.007     0.995     -6.46
7      6.512
Sex Males CrYr        2.747e+04   3.55e+04     0.774     0.439   -4.23e+0
4      9.73e+04
Sex Females CrYr      2.747e+04   3.55e+04     0.774     0.439   -4.23e+0
4      9.73e+04
Age Average CrYr      1739.4376    1.07e+04     0.163     0.871   -1.92e+0
4      2.27e+04
Unemployment Rate     1991.1837    1.42e+04     0.141     0.888   -2.59e+0
4      2.98e+04
Average Household Income CrYr 10.2890     4.411     2.332     0.020     1.61
5      18.963
Householder under 25 years: 1.552e+04   1.56e+04     0.993     0.321   -1.52e+0
4      4.62e+04
Householder 25 to 44 years: -4428.9690   1.47e+04    -0.302     0.763   -3.32e+0
4      2.44e+04
Family Inc: Aggregate Average -2.265e-06   4.76e-05    -0.048     0.962   -9.58e-0
5      9.12e-05
Per Capita Income CrYr   -26.5985     11.232    -2.368     0.018   -48.68
4      -4.513
Wealth: Average Household -0.0002       0.002    -0.120     0.905     -0.00
4      0.003
Poverty: In poverty     9.6906     14.199     0.683     0.495   -18.22
7      37.608
Male 14 to 35 CrYr      1.4836     15.877     0.093     0.926   -29.73
5      32.702
Female 14 to 35 CrYr     0.0347     16.197     0.002     0.998   -31.81

```

3	31.883					
Total Pop 14 to 35 CrYr	1.5079	2.188	0.689	0.491	-2.79	
4	5.810					
Under 25 Low Income	-1.552e+04	1.56e+04	-0.993	0.321	-4.62e+0	
4	1.52e+04					
Under 25 Med Income	-1.549e+04	1.56e+04	-0.991	0.322	-4.62e+0	
4	1.52e+04					
Under 25 High Income	-1.564e+04	1.56e+04	-1.000	0.318	-4.64e+0	
4	1.51e+04					
25 to 44 Low Income	4422.5932	1.47e+04	0.302	0.763	-2.44e+0	
4	3.32e+04					
25 to 44 Med Income	4446.0178	1.47e+04	0.303	0.762	-2.44e+0	
4	3.33e+04					
25 to 44 High Income	4435.7796	1.47e+04	0.303	0.762	-2.44e+0	
4	3.32e+04					
Pop class_Suburban	-9.206e+04	8.01e+04	-1.150	0.251	-2.5e+0	
5	6.54e+04					
Pop class_Urban	6.873e+04	7.17e+04	0.958	0.338	-7.23e+0	
4	2.1e+05					
Centre Type_Outlet	2.067e+05	8.89e+04	2.326	0.021	3.2e+0	
4	3.81e+05					
Centre Type_Strip	-1.064e+05	4.28e+04	-2.484	0.013	-1.91e+0	
5	-2.22e+04					
Climate_Frigid	-1.997e+04	7.52e+04	-0.266	0.791	-1.68e+0	
5	1.28e+05					
Climate_Hot	4.449e+04	6.61e+04	0.673	0.501	-8.54e+0	
4	1.74e+05					
Climate_Moderate	4.571e+04	5e+04	0.914	0.361	-5.27e+0	
4	1.44e+05					
Climate_Warm	3.424e+04	5.44e+04	0.629	0.530	-7.28e+0	
4	1.41e+05					
Guys Segmentation_On Trend	-1.091e+05	2.02e+05	-0.539	0.590	-5.07e+0	
5	2.89e+05					
Guys Segmentation_On trend	-1.701e+05	2.47e+05	-0.690	0.491	-6.55e+0	
5	3.15e+05					
Girls Segmentation_On Trend	-9.524e+04	2.02e+05	-0.472	0.637	-4.92e+0	
5	3.02e+05					
ChangeMade_no change	-6.539e+04	1.13e+05	-0.580	0.562	-2.87e+0	
5	1.56e+05					
ChangeMade_relocation	3.826e+04	1.13e+05	0.340	0.734	-1.83e+0	
5	2.6e+05					
ChangeMade_remodel	1.091e+05	2.86e+05	0.382	0.703	-4.53e+0	
5	6.71e+05					
ChangeDate_2006-02-22 00:00:00	1.616e+05	4.07e+05	0.397	0.691	-6.38e+0	
5	9.62e+05					
ChangeDate_2007-02-08 00:00:00	-17.6545	7.153	-2.468	0.014	-31.72	
0	-3.589					
ChangeDate_2007-03-13 00:00:00	-3.449e+05	4.38e+05	-0.788	0.431	-1.21e+0	
6	5.16e+05					
ChangeDate_2007-07-13 00:00:00	-5.944e+05	4.79e+05	-1.242	0.215	-1.54e+0	
6	3.47e+05					
ChangeDate_2007-07-19 00:00:00	-5.215e+05	4.73e+05	-1.104	0.270	-1.45e+0	
6	4.08e+05					
ChangeDate_2008-02-26 00:00:00	-7.79e+04	4.12e+05	-0.189	0.850	-8.88e+0	
5	7.33e+05					
ChangeDate_2008-06-24 00:00:00	-6.411e+04	4.11e+05	-0.156	0.876	-8.72e+0	
5	7.44e+05					
ChangeDate_2008-06-25 00:00:00	-3.55e+05	3.98e+05	-0.892	0.373	-1.14e+0	
6	4.28e+05					
ChangeDate_2008-10-24 00:00:00	-3.052e+04	4.04e+05	-0.076	0.940	-8.24e+0	

5	7.63e+05						
ChangeDate_2008-11-17 00:00:00	2.602e+05	4.13e+05	0.630	0.529	-5.52e+0		
5	1.07e+06						
ChangeDate_2009-02-12 00:00:00	6.745e+04	3.96e+05	0.170	0.865	-7.11e+0		
5	8.46e+05						
ChangeDate_2009-03-12 00:00:00	-1.75e+05	3.96e+05	-0.442	0.659	-9.54e+0		
5	6.04e+05						
ChangeDate_2009-06-04 00:00:00	-6.178e+05	5e+05	-1.236	0.217	-1.6e+0		
6	3.65e+05						
ChangeDate_2009-09-17 00:00:00	-1.7086	5.520	-0.310	0.757	-12.56		
1	9.144						
ChangeDate_2009-09-18 00:00:00	-3.326e+05	3.99e+05	-0.833	0.406	-1.12e+0		
6	4.53e+05						
ChangeDate_2009-10-01 00:00:00	-1.188e+05	3.27e+05	-0.363	0.717	-7.62e+0		
5	5.24e+05						
ChangeDate_2010-05-06 00:00:00	-5.753e+05	3.97e+05	-1.451	0.148	-1.36e+0		
6	2.05e+05						
ChangeDate_2010-06-24 00:00:00	1.587e+05	4.91e+05	0.324	0.746	-8.06e+0		
5	1.12e+06						
ChangeDate_2010-07-22 00:00:00	11.6049	7.639	1.519	0.130	-3.41		
6	26.626						
ChangeDate_2010-09-02 00:00:00	5.3696	8.736	0.615	0.539	-11.80		
7	22.547						
ChangeDate_2010-10-07 00:00:00	-4.99e+05	3.41e+05	-1.463	0.144	-1.17e+0		
6	1.72e+05						
ChangeDate_2010-10-14 00:00:00	-4.536e+05	3.98e+05	-1.139	0.255	-1.24e+0		
6	3.29e+05						
ChangeDate_2011-02-10 00:00:00	-4.921e+05	3.97e+05	-1.239	0.216	-1.27e+0		
6	2.89e+05						
ChangeDate_2011-02-25 00:00:00	-2.322e+05	3.99e+05	-0.582	0.561	-1.02e+0		
6	5.53e+05						
ChangeDate_2011-03-31 00:00:00	-2.979e+05	3.95e+05	-0.754	0.451	-1.07e+0		
6	4.79e+05						
ChangeDate_2011-04-07 00:00:00	1.291e+05	4.14e+05	0.312	0.755	-6.85e+0		
5	9.43e+05						
ChangeDate_2011-04-29 00:00:00	-3.86e+05	3.98e+05	-0.970	0.333	-1.17e+0		
6	3.97e+05						
ChangeDate_2011-06-23 00:00:00	-2.51e+05	2.51e+05	-1.001	0.317	-7.44e+0		
5	2.42e+05						
ChangeDate_2011-06-30 00:00:00	1.2840	6.615	0.194	0.846	-11.72		
3	14.291						
ChangeDate_2011-09-29 00:00:00	-3.551e+05	3.92e+05	-0.905	0.366	-1.13e+0		
6	4.16e+05						
ChangeDate_2011-10-06 00:00:00	-1.278e+04	4.01e+05	-0.032	0.975	-8.02e+0		
5	7.77e+05						
ChangeDate_2011-10-13 00:00:00	14.6549	6.365	2.302	0.022	2.13		
9	27.170						
ChangeDate_2011-11-03 00:00:00	-5.563e+05	3.95e+05	-1.407	0.160	-1.33e+0		
6	2.21e+05						
ChangeDate_2012-03-29 00:00:00	5.143e+05	4.21e+05	1.222	0.222	-3.13e+0		
5	1.34e+06						
ChangeDate_2012-04-13 00:00:00	2.329e+05	3.13e+05	0.744	0.458	-3.83e+0		
5	8.49e+05						
ChangeDate_2012-05-17 00:00:00	-4.563e+05	3.96e+05	-1.151	0.250	-1.24e+0		
6	3.23e+05						
ChangeDate_2012-07-19 00:00:00	-3.405e+05	4e+05	-0.850	0.396	-1.13e+0		
6	4.47e+05						
ChangeDate_2012-08-30 00:00:00	-3.816e+05	4.08e+05	-0.936	0.350	-1.18e+0		
6	4.2e+05						
ChangeDate_2012-09-27 00:00:00	-4.128e+05	4.14e+05	-0.997	0.319	-1.23e+0		

6	4.01e+05						
ChangeDate_2012-10-19 00:00:00	2.629e+04	3.93e+05	0.067	0.947	-7.47e+0		
5	8e+05						
ChangeDate_2012-10-25 00:00:00	2.346e+04	4.19e+05	0.056	0.955	-8e+0		
5	8.47e+05						
ChangeDate_2012-11-08 00:00:00	-1.543e+05	3.17e+05	-0.487	0.626	-7.77e+0		
5	4.68e+05						
ChangeDate_2013-03-15 00:00:00	-4.713e+05	4.07e+05	-1.157	0.248	-1.27e+0		
6	3.3e+05						
ChangeDate_2013-05-09 00:00:00	-5.227e+05	4.01e+05	-1.303	0.193	-1.31e+0		
6	2.66e+05						
ChangeDate_2013-08-22 00:00:00	-16.2718	6.845	-2.377	0.018	-29.73		
2	-2.812						
ChangeDate_2013-09-26 00:00:00	-2.31e+05	4.23e+05	-0.546	0.586	-1.06e+0		
6	6.01e+05						
ChangeDate_2013-10-31 00:00:00	-4.225e+05	4.04e+05	-1.046	0.296	-1.22e+0		
6	3.71e+05						
ChangeDate_2013-11-07 00:00:00	-4.565e+05	3.03e+05	-1.505	0.133	-1.05e+0		
6	1.4e+05						
ChangeDate_2014-05-22 00:00:00	-5.041e+05	4.02e+05	-1.253	0.211	-1.3e+0		
6	2.87e+05						
ChangeDate_2014-06-05 00:00:00	-5.689e+05	3.96e+05	-1.437	0.152	-1.35e+0		
6	2.09e+05						
ChangeDate_2014-06-13 00:00:00	-2.296e+05	3.97e+05	-0.578	0.564	-1.01e+0		
6	5.52e+05						
ChangeDate_2014-07-03 00:00:00	-2.919e+05	3.37e+05	-0.867	0.387	-9.54e+0		
5	3.7e+05						
ChangeDate_2014-07-31 00:00:00	-7.9605	9.855	-0.808	0.420	-27.33		
7	11.416						
ChangeDate_2014-10-02 00:00:00	-9.7317	8.921	-1.091	0.276	-27.27		
3	7.810						
ChangeDate_2014-10-23 00:00:00	8.974e+04	3.95e+05	0.227	0.820	-6.87e+0		
5	8.66e+05						
ChangeDate_2014-10-30 00:00:00	1.826e+05	3.96e+05	0.461	0.645	-5.96e+0		
5	9.61e+05						
ChangeDate_2014-11-06 00:00:00	-3.474e+05	4.34e+05	-0.800	0.424	-1.2e+0		
6	5.06e+05						
ChangeDate_2014-11-13 00:00:00	3.678e+05	3.98e+05	0.923	0.356	-4.15e+0		
5	1.15e+06						
ChangeDate_2014-11-23 00:00:00	-1.3109	8.715	-0.150	0.881	-18.44		
6	15.824						
ChangeDate_2015-03-26 00:00:00	2.92e+05	4e+05	0.730	0.466	-4.94e+0		
5	1.08e+06						
ChangeDate_2015-04-16 00:00:00	-5.907e+05	4.19e+05	-1.409	0.160	-1.41e+0		
6	2.33e+05						
ChangeDate_2015-04-23 00:00:00	7.891e+05	4.18e+05	1.886	0.060	-3.34e+0		
4	1.61e+06						
ChangeDate_2015-04-29 00:00:00	6.915e+04	3.59e+05	0.193	0.847	-6.36e+0		
5	7.75e+05						
ChangeDate_2015-04-30 00:00:00	-2.722e+05	3.19e+05	-0.853	0.394	-9e+0		
5	3.55e+05						
ChangeDate_2015-05-20 00:00:00	3.57e+04	4.16e+05	0.086	0.932	-7.83e+0		
5	8.54e+05						
ChangeDate_2015-05-28 00:00:00	-2.299e+04	4.19e+05	-0.055	0.956	-8.47e+0		
5	8.01e+05						
ChangeDate_2015-05-29 00:00:00	1.9700	10.299	0.191	0.848	-18.28		
1	22.221						
ChangeDate_2015-07-23 00:00:00	-5.537e+05	3.99e+05	-1.389	0.166	-1.34e+0		
6	2.3e+05						
ChangeDate_2015-07-28 00:00:00	-5.575e+04	4.27e+05	-0.130	0.896	-8.96e+0		

5	7.84e+05							
ChangeDate_2015-07-30 00:00:00	1.889e+05	4.2e+05	0.450	0.653	-6.37e+0			
5	1.02e+06							
ChangeDate_2015-08-13 00:00:00	-1.433e+05	3.43e+05	-0.417	0.677	-8.18e+0			
5	5.32e+05							
ChangeDate_2015-08-19 00:00:00	4.162e+05	4.26e+05	0.976	0.329	-4.22e+0			
5	1.25e+06							
ChangeDate_2015-09-03 00:00:00	2.653e+05	3.18e+05	0.835	0.404	-3.6e+0			
5	8.9e+05							
ChangeDate_2015-09-17 00:00:00	-4.982e+05	3.27e+05	-1.523	0.129	-1.14e+0			
6	1.45e+05							
ChangeDate_2015-09-24 00:00:00	9556.4855	3.23e+05	0.030	0.976	-6.26e+0			
5	6.45e+05							
ChangeDate_2015-09-29 00:00:00	-6.5783	8.549	-0.769	0.442	-23.38			
8	10.232							
ChangeDate_2015-09-30 00:00:00	2.406e+05	4.18e+05	0.576	0.565	-5.81e+0			
5	1.06e+06							
ChangeDate_2015-10-01 00:00:00	-6.186e+04	4e+05	-0.155	0.877	-8.48e+0			
5	7.24e+05							
ChangeDate_2015-10-08 00:00:00	5.546e+05	3.21e+05	1.725	0.085	-7.75e+0			
4	1.19e+06							
ChangeDate_2015-10-15 00:00:00	1.518e+05	4.11e+05	0.370	0.712	-6.56e+0			
5	9.59e+05							
ChangeDate_2015-10-16 00:00:00	4.459e+05	4.16e+05	1.073	0.284	-3.71e+0			
5	1.26e+06							
ChangeDate_2015-10-22 00:00:00	1.306e+05	4.03e+05	0.324	0.746	-6.62e+0			
5	9.23e+05							
ChangeDate_2015-10-24 00:00:00	4.051e+05	4.14e+05	0.979	0.328	-4.09e+0			
5	1.22e+06							
ChangeDate_2015-10-29 00:00:00	1.234e+05	2.73e+05	0.452	0.652	-4.14e+0			
5	6.61e+05							
ChangeDate_2015-10-30 00:00:00	5.71e+05	3.97e+05	1.437	0.152	-2.1e+0			
5	1.35e+06							
ChangeDate_2015-11-04 00:00:00	1.174e+04	3.43e+05	0.034	0.973	-6.62e+0			
5	6.86e+05							
ChangeDate_2015-11-05 00:00:00	3.3048	10.605	0.312	0.755	-17.54			
8	24.157							
ChangeDate_2015-11-07 00:00:00	-2.638e+05	4.05e+05	-0.652	0.515	-1.06e+0			
6	5.32e+05							
ChangeDate_2015-11-11 00:00:00	-2.161e+05	3.07e+05	-0.703	0.482	-8.21e+0			
5	3.88e+05							
ChangeDate_2015-11-12 00:00:00	-1.742e+05	2.71e+05	-0.642	0.521	-7.08e+0			
5	3.6e+05							
ChangeDate_2015-11-18 00:00:00	-1.729e+05	4.17e+05	-0.414	0.679	-9.94e+0			
5	6.48e+05							
ChangeDate_2015-11-20 00:00:00	4.012e+05	4.35e+05	0.923	0.357	-4.54e+0			
5	1.26e+06							
ChangeDate_2015-12-03 00:00:00	-2.29e+05	4.02e+05	-0.569	0.570	-1.02e+0			
6	5.62e+05							
ChangeDate_2016-02-04 00:00:00	-2.622e+05	3.94e+05	-0.665	0.506	-1.04e+0			
6	5.13e+05							
ChangeDate_2016-02-11 00:00:00	5.157e+05	4.05e+05	1.272	0.204	-2.81e+0			
5	1.31e+06							
ChangeDate_2016-03-03 00:00:00	-5.135e+05	4.15e+05	-1.238	0.217	-1.33e+0			
6	3.02e+05							
ChangeDate_2016-03-10 00:00:00	-2.918e+05	2.95e+05	-0.990	0.323	-8.71e+0			
5	2.88e+05							
ChangeDate_2016-03-16 00:00:00	9.29e+04	4.21e+05	0.221	0.825	-7.34e+0			
5	9.2e+05							
ChangeDate_2016-03-21 00:00:00	-2.148e+05	4.17e+05	-0.515	0.607	-1.03e+0			

6	6.05e+05						
ChangeDate_2016-03-23 00:00:00	-3.657e+05	4.22e+05	-0.866	0.387	-1.2e+0		
6	4.64e+05						
ChangeDate_2016-03-24 00:00:00	1.915e+05	2.85e+05	0.671	0.503	-3.69e+0		
5	7.52e+05						
ChangeDate_2016-04-07 00:00:00	-3e+05	4.01e+05	-0.748	0.455	-1.09e+0		
6	4.88e+05						
ChangeDate_2016-04-14 00:00:00	-1.744e+05	3.2e+05	-0.544	0.587	-8.04e+0		
5	4.56e+05						
ChangeDate_2016-04-21 00:00:00	-2.119e+05	3.39e+05	-0.624	0.533	-8.79e+0		
5	4.55e+05						
ChangeDate_2016-05-12 00:00:00	4.5565	6.541	0.697	0.486	-8.30		
4	17.417						
ChangeDate_2016-05-19 00:00:00	-1.677e+04	4.15e+05	-0.040	0.968	-8.33e+0		
5	8e+05						
ChangeDate_2016-05-25 00:00:00	2.673e+05	4.22e+05	0.634	0.527	-5.62e+0		
5	1.1e+06						
ChangeDate_2016-05-26 00:00:00	6.242e+04	4.17e+05	0.150	0.881	-7.57e+0		
5	8.82e+05						
ChangeDate_2016-06-08 00:00:00	-6.041e+05	4.01e+05	-1.506	0.133	-1.39e+0		
6	1.85e+05						
ChangeDate_2016-06-09 00:00:00	1.573e+05	2.92e+05	0.539	0.590	-4.17e+0		
5	7.31e+05						
ChangeDate_2016-06-16 00:00:00	6.537e+05	4.02e+05	1.624	0.105	-1.38e+0		
5	1.44e+06						
ChangeDate_2016-06-23 00:00:00	-1.728e+05	4.26e+05	-0.405	0.685	-1.01e+0		
6	6.65e+05						
ChangeDate_2016-06-30 00:00:00	-5.099e+05	3.18e+05	-1.601	0.110	-1.14e+0		
6	1.16e+05						
ChangeDate_2016-07-14 00:00:00	-4.211e+05	3.17e+05	-1.330	0.184	-1.04e+0		
6	2.01e+05						
ChangeDate_2016-07-21 00:00:00	-4.846e+05	3.97e+05	-1.220	0.223	-1.27e+0		
6	2.96e+05						
ChangeDate_2016-07-27 00:00:00	2.726e+05	4.16e+05	0.655	0.513	-5.46e+0		
5	1.09e+06						
ChangeDate_2016-07-28 00:00:00	1.395e+05	2.82e+05	0.495	0.621	-4.15e+0		
5	6.94e+05						
ChangeDate_2016-08-04 00:00:00	-2.053e+04	2.52e+05	-0.081	0.935	-5.16e+0		
5	4.75e+05						
ChangeDate_2016-08-11 00:00:00	-1.835e+05	3.98e+05	-0.461	0.645	-9.66e+0		
5	5.99e+05						
ChangeDate_2016-08-25 00:00:00	-1.939e+05	4.01e+05	-0.483	0.629	-9.83e+0		
5	5.95e+05						
ChangeDate_2016-09-01 00:00:00	11.6785	10.363	1.127	0.260	-8.69		
9	32.056						
ChangeDate_2016-09-22 00:00:00	6.869e+04	4.03e+05	0.171	0.865	-7.23e+0		
5	8.6e+05						
ChangeDate_2016-09-29 00:00:00	3.1603	10.933	0.289	0.773	-18.33		
7	24.658						
ChangeDate_2016-10-05 00:00:00	-7.1301	10.917	-0.653	0.514	-28.59		
6	14.336						
ChangeDate_2016-10-06 00:00:00	-1.839e+05	4.02e+05	-0.458	0.647	-9.74e+0		
5	6.06e+05						
ChangeDate_2016-10-13 00:00:00	-6.041e+04	3.45e+05	-0.175	0.861	-7.39e+0		
5	6.18e+05						
ChangeDate_2016-10-20 00:00:00	2.193e+05	4.11e+05	0.533	0.594	-5.89e+0		
5	1.03e+06						
ChangeDate_2016-10-27 00:00:00	-1.345e+05	3.15e+05	-0.427	0.669	-7.53e+0		
5	4.84e+05						
ChangeDate_2016-11-03 00:00:00	-8.3883	10.887	-0.771	0.441	-29.79		

```

4      13.017
ChangeDate_2016-11-05 00:00:00 -3.071e+05  4.09e+05  -0.751  0.453  -1.11e+0
6      4.97e+05
ChangeDate_2016-11-09 00:00:00 -3.312e+05  3.94e+05  -0.840  0.401  -1.11e+0
6      4.44e+05
ChangeDate_2016-11-10 00:00:00 -1.348e+05  2.48e+05  -0.544  0.587  -6.22e+0
5      3.53e+05
ChangeDate_2016-11-17 00:00:00  3.371e+05  3.97e+05  0.850  0.396  -4.43e+0
5      1.12e+06
ChangeDate_2016-11-18 00:00:00      3.9907    10.796  0.370  0.712  -17.23
6      25.218
ChangeDate_2016-12-08 00:00:00  6.682e+04  3.94e+05  0.169  0.866  -7.08e+0
5      8.42e+05
ChangeDate_2017-03-23 00:00:00  -8.7877    10.916  -0.805  0.421  -30.25
1      12.676
ChangeDate_2017-05-02 00:00:00 -5.029e+05  4.27e+05  -1.177  0.240  -1.34e+0
6      3.37e+05
ChangeDate_2017-05-03 00:00:00 -2.594e+05  4.88e+05  -0.531  0.595  -1.22e+0
6      7e+05
ChangeDate_2017-05-09 00:00:00 -1.704e+05  4.98e+05  -0.343  0.732  -1.15e+0
6      8.08e+05
ChangeDate_2017-05-11 00:00:00 -8.289e+04  4.02e+05  -0.206  0.837  -8.73e+0
5      7.08e+05
ChangeDate_2017-05-18 00:00:00 -5.942e+05  4.06e+05  -1.463  0.144  -1.39e+0
6      2.04e+05
ChangeDate_2017-05-25 00:00:00 -2.814e+05  3.93e+05  -0.716  0.474  -1.05e+0
6      4.91e+05
ChangeDate_2017-06-13 00:00:00 -4.589e+05  4.92e+05  -0.932  0.352  -1.43e+0
6      5.09e+05
ChangeDate_2017-11-24 00:00:00 -5.289e+05  5.12e+05  -1.034  0.302  -1.53e+0
6      4.77e+05
ChangeDate_2018-01-25 00:00:00  3.884e+05  3.09e+05  1.256  0.210  -2.2e+0
5      9.97e+05
ChangeDate_2018-02-03 00:00:00 -1.503e+05  4.06e+05  -0.370  0.712  -9.49e+0
5      6.48e+05
ChangeDate_2018-04-19 00:00:00  208.5277  4.09e+05  0.001  1.000  -8.04e+0
5      8.04e+05
ChangeDate_2018-05-10 00:00:00 -2.194e+05  4e+05  -0.549  0.584  -1.01e+0
6      5.67e+05
ChangeDate_2019-04-04 00:00:00  4.521e+05  3.94e+05  1.148  0.252  -3.22e+0
5      1.23e+06
ChangeDate_2019-09-12 00:00:00  9.476e+04  3.5e+05  0.270  0.787  -5.94e+0
5      7.84e+05
ChangeDate_2019-09-25 00:00:00 -9.481e-05  0.000  -0.875  0.382  -0.00
0      0.000
ChangeDate_2020-01-23 00:00:00 -2.516e+05  4.38e+05  -0.575  0.566  -1.11e+0
6      6.09e+05
ChangeDate_nan      -6.539e+04  1.13e+05  -0.580  0.562  -2.87e+0
5      1.56e+05

```

```

=====
Omnibus:      119.868  Durbin-Watson:      1.911
Prob(Omnibus):      0.000  Jarque-Bera (JB):      391.427
Skew:      1.000  Prob(JB):      1.01e-85
Kurtosis:      6.608  Cond. No.      3.05e+26
=====

```

Notes:

[1] Standard Errors assume that the covariance matrix of the errors is correctly specified.

[2] The smallest eigenvalue is 1.35e-30. This might indicate that there are strong multicollinearity problems or that the design matrix is singular.

```
In [6]: vif_df = pd.DataFrame()
vif_df["Feature"] = X_train_sm.columns
vif_df["VIF"] = [variance_inflation_factor(X_train_sm.values, i) for i in range(X_train_sm.shape[1])]
print("\nVariance Inflation Factor (VIF):")
print(vif_df.sort_values("VIF", ascending=False))
```

```
C:\Users\mahima\anaconda3\Lib\site-packages\statsmodels\stats\outliers_influence.py:198: RuntimeWarning: divide by zero encountered in scalar divide
  vif = 1. / (1. - r_squared_i)
C:\Users\mahima\anaconda3\Lib\site-packages\statsmodels\regression\linear_model.py:1781: RuntimeWarning: invalid value encountered in scalar divide
  return 1 - self.ssr/self.centered_tss
```

```
Variance Inflation Factor (VIF):
```

	Feature	VIF
20	Male 14 to 35 CrYr	inf
22	Total Pop 14 to 35 CrYr	inf
21	Female 14 to 35 CrYr	inf
6	Population Est CrYr	2.151744e+12
195	ChangeDate_nan	1.705263e+12
..
166	ChangeDate_2016-10-05 00:00:00	NaN
171	ChangeDate_2016-11-03 00:00:00	NaN
176	ChangeDate_2016-11-18 00:00:00	NaN
178	ChangeDate_2017-03-23 00:00:00	NaN
193	ChangeDate_2019-09-25 00:00:00	NaN

[196 rows x 2 columns]

```
In [15]: # Prepare test set for prediction
X_test_sm = sm.add_constant(X_test)
y_pred = model.predict(X_test_sm)

def mean_absolute_percentage_error(y_true, y_pred):
    return np.mean(np.abs((y_true - y_pred) / y_true)) * 100

mape = mean_absolute_percentage_error(y_test, y_pred)
mae = mean_absolute_error(y_test, y_pred)
rmse = np.sqrt(mean_squared_error(y_test, y_pred))
r2 = r2_score(y_test, y_pred)

print(f"\nModel Performance on Test Data:")
print(f"MAPE: {mape:.2f}%")
print(f"MAE: {mae:.2f}")
print(f"RMSE: {rmse:.2f}")
print(f"R²: {r2:.4f}")
```

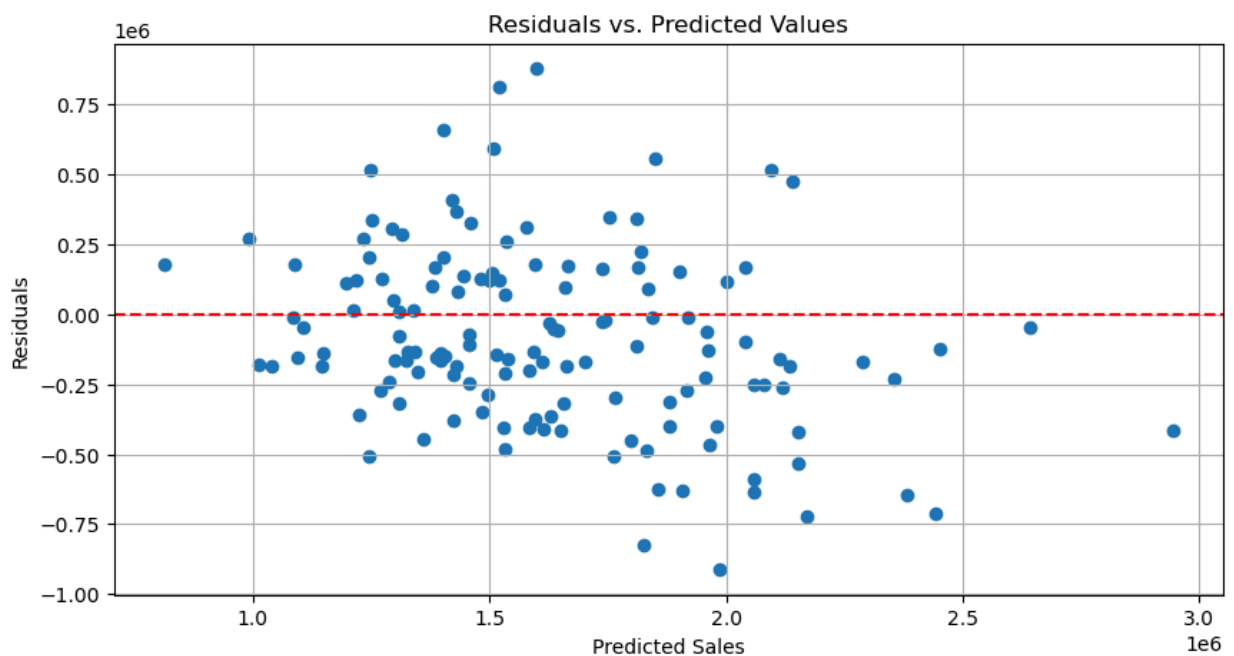
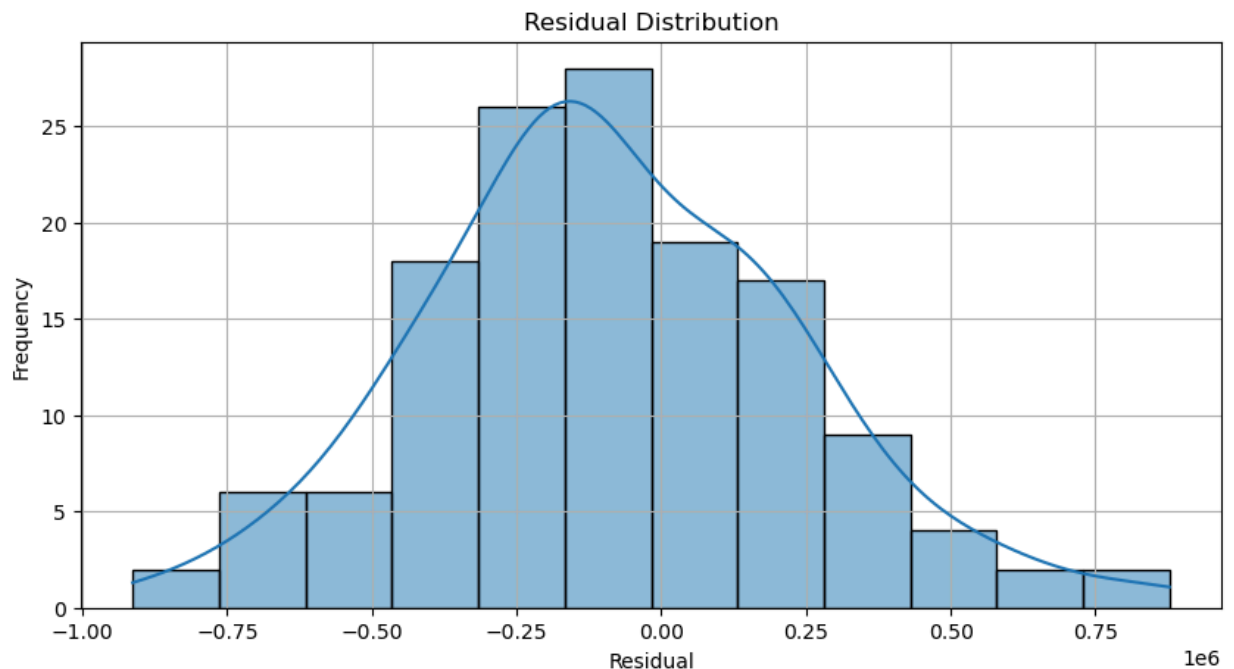
```
Model Performance on Test Data:
MAPE: 18.74%
MAE: 268855.94
RMSE: 333030.93
R²: 0.3197
```

```
In [8]: residuals = y_test - y_pred

plt.figure(figsize=(10, 5))
sns.histplot(residuals, kde=True)
plt.title("Residual Distribution")
```

```
plt.xlabel("Residual")
plt.ylabel("Frequency")
plt.grid(True)
plt.show()

plt.figure(figsize=(10, 5))
plt.scatter(y_pred, residuals)
plt.axhline(0, color='red', linestyle='--')
plt.title("Residuals vs. Predicted Values")
plt.xlabel("Predicted Sales")
plt.ylabel("Residuals")
plt.grid(True)
plt.show()
```



```
In [16]: summary_df = pd.DataFrame({
    "Feature": model.params.index,
    "Coefficient": model.params.values,
```

```
        "P-value": model.pvalues.values
    })
summary_df.to_excel(r"D:\internship\model_summary.xlsx", index=False)
vif_df.to_excel(r"D:\internship\vif_analysis.xlsx", index=False)

print("\n Results exported to 'model_summary.xlsx' and 'vif_analysis.xlsx' in D:\\inte

Results exported to 'model_summary.xlsx' and 'vif_analysis.xlsx' in D:\internship
```

In []: