In [111…

```
import pandas as pd
from matplotlib import pyplot as plt
%matplotlib inline
import numpy as np
import seaborn as sns
from scipy import stats
import statsmodels.formula.api as smf
from sklearn.preprocessing import StandardScaler
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_absolute_error,mean_absolute_percentage_error,mean_squ

import warnings
warnings.filterwarnings('ignore')
```

In [112…

```
data_set = pd.read_csv('50_Startups.csv')
data_set
```

Out[112…

	R&D Spend	Administration	Marketing Spend	State	Profit
0	165349.20	136897.80	471784.10	New York	192261.83
1	162597.70	151377.59	443898.53	California	191792.06
2	153441.51	101145.55	407934.54	Florida	191050.39
3	144372.41	118671.85	383199.62	New York	182901.99
4	142107.34	91391.77	366168.42	Florida	166187.94
5	131876.90	99814.71	362861.36	New York	156991.12
6	134615.46	147198.87	127716.82	California	156122.51
7	130298.13	145530.06	323876.68	Florida	155752.60
8	120542.52	148718.95	311613.29	New York	152211.77
9	123334.88	108679.17	304981.62	California	149759.96
10	101913.08	110594.11	229160.95	Florida	146121.95
11	100671.96	91790.61	249744.55	California	144259.40
12	93863.75	127320.38	249839.44	Florida	141585.52
13	91992.39	135495.07	252664.93	California	134307.35
14	119943.24	156547.42	256512.92	Florida	132602.65
15	114523.61	122616.84	261776.23	New York	129917.04
16	78013.11	121597.55	264346.06	California	126992.93
17	94657.16	145077.58	282574.31	New York	125370.37
18	91749.16	114175.79	294919.57	Florida	124266.90
19	86419.70	153514.11	0.00	New York	122776.86
20	76253.86	113867.30	298664.47	California	118474.03
21	78389.47	153773.43	299737.29	New York	111313.02

	R&D Spend	Administration	Marketing Spend	State	Profit
22	73994.56	122782.75	303319.26	Florida	110352.25
23	67532.53	105751.03	304768.73	Florida	108733.99
24	77044.01	99281.34	140574.81	New York	108552.04
25	64664.71	139553.16	137962.62	California	107404.34
26	75328.87	144135.98	134050.07	Florida	105733.54
27	72107.60	127864.55	353183.81	New York	105008.31
28	66051.52	182645.56	118148.20	Florida	103282.38
29	65605.48	153032.06	107138.38	New York	101004.64
30	61994.48	115641.28	91131.24	Florida	99937.59
31	61136.38	152701.92	88218.23	New York	97483.56
32	63408.86	129219.61	46085.25	California	97427.84
33	55493.95	103057.49	214634.81	Florida	96778.92
34	46426.07	157693.92	210797.67	California	96712.80
35	46014.02	85047.44	205517.64	New York	96479.51
36	28663.76	127056.21	201126.82	Florida	90708.19
37	44069.95	51283.14	197029.42	California	89949.14
38	20229.59	65947.93	185265.10	New York	81229.06
39	38558.51	82982.09	174999.30	California	81005.76
40	28754.33	118546.05	172795.67	California	78239.91
41	27892.92	84710.77	164470.71	Florida	77798.83
42	23640.93	96189.63	148001.11	California	71498.49
43	15505.73	127382.30	35534.17	New York	69758.98
44	22177.74	154806.14	28334.72	California	65200.33
45	1000.23	124153.04	1903.93	New York	64926.08
46	1315.46	115816.21	297114.46	Florida	49490.75
47	0.00	135426.92	0.00	California	42559.73
48	542.05	51743.15	0.00	New York	35673.41
49	0.00	116983.80	45173.06	California	14681.40

In [113…

```
data_set.replace(0,np.nan,inplace=True)
data_set=data_set.dropna()
data_set
```

Out[113…

	R&D Spend	Administration	Marketing Spend	State	Profit
0	165349.20	136897.80	471784.10	New York	192261.83

	R&D Spend	Administration	Marketing Spend	State	Profit
1	162597.70	151377.59	443898.53	California	191792.06
2	153441.51	101145.55	407934.54	Florida	191050.39
3	144372.41	118671.85	383199.62	New York	182901.99
4	142107.34	91391.77	366168.42	Florida	166187.94
5	131876.90	99814.71	362861.36	New York	156991.12
6	134615.46	147198.87	127716.82	California	156122.51
7	130298.13	145530.06	323876.68	Florida	155752.60
8	120542.52	148718.95	311613.29	New York	152211.77
9	123334.88	108679.17	304981.62	California	149759.96
10	101913.08	110594.11	229160.95	Florida	146121.95
11	100671.96	91790.61	249744.55	California	144259.40
12	93863.75	127320.38	249839.44	Florida	141585.52
13	91992.39	135495.07	252664.93	California	134307.35
14	119943.24	156547.42	256512.92	Florida	132602.65
15	114523.61	122616.84	261776.23	New York	129917.04
16	78013.11	121597.55	264346.06	California	126992.93
17	94657.16	145077.58	282574.31	New York	125370.37
18	91749.16	114175.79	294919.57	Florida	124266.90
20	76253.86	113867.30	298664.47	California	118474.03
21	78389.47	153773.43	299737.29	New York	111313.02
22	73994.56	122782.75	303319.26	Florida	110352.25
23	67532.53	105751.03	304768.73	Florida	108733.99
24	77044.01	99281.34	140574.81	New York	108552.04
25	64664.71	139553.16	137962.62	California	107404.34
26	75328.87	144135.98	134050.07	Florida	105733.54
27	72107.60	127864.55	353183.81	New York	105008.31
28	66051.52	182645.56	118148.20	Florida	103282.38
29	65605.48	153032.06	107138.38	New York	101004.64
30	61994.48	115641.28	91131.24	Florida	99937.59
31	61136.38	152701.92	88218.23	New York	97483.56
32	63408.86	129219.61	46085.25	California	97427.84
33	55493.95	103057.49	214634.81	Florida	96778.92
34	46426.07	157693.92	210797.67	California	96712.80

	R&D Spend	Administration	Marketing Spend	State	Profit
35	46014.02	85047.44	205517.64	New York	96479.51
36	28663.76	127056.21	201126.82	Florida	90708.19
37	44069.95	51283.14	197029.42	California	89949.14
38	20229.59	65947.93	185265.10	New York	81229.06
39	38558.51	82982.09	174999.30	California	81005.76
40	28754.33	118546.05	172795.67	California	78239.91
41	27892.92	84710.77	164470.71	Florida	77798.83
42	23640.93	96189.63	148001.11	California	71498.49
43	15505.73	127382.30	35534.17	New York	69758.98
44	22177.74	154806.14	28334.72	California	65200.33
45	1000.23	124153.04	1903.93	New York	64926.08
46	1315.46	115816.21	297114.46	Florida	49490.75

In [114…

```
data_set.shape
```

Out[114…

```
(46, 5)
```

In [115…

```
data_set.isna().sum()
```

Out[115…

```
R&D Spend          0
Administration     0
Marketing Spend    0
State              0
Profit             0
dtype: int64
```

In [116…

```
data_set.dtypes
```

Out[116…

```
R&D Spend          float64
Administration     float64
Marketing Spend    float64
State               object
Profit             float64
dtype: object
```

In [117…

```
data_set.describe()
```

Out[117…

	R&D Spend	Administration	Marketing Spend	Profit
count	46.000000	46.000000	46.000000	46.000000
mean	78241.718043	121947.043478	228393.083261	117063.925217
std	43695.348335	26812.767940	111336.691801	35960.003971
min	1000.230000	51283.140000	1903.930000	49490.750000

	R&D Spend	Administration	Marketing Spend	Profit
25%	46117.032500	103730.875000	142431.385000	96537.832500
50%	74661.715000	122699.795000	239452.750000	108643.015000
75%	111370.977500	144842.180000	302423.767500	143590.930000
max	165349.200000	182645.560000	471784.100000	192261.830000

In [118…

```
data_set.drop(labels ='State',axis=1,inplace =True)
```

In [119…

```
data_set.head()
```

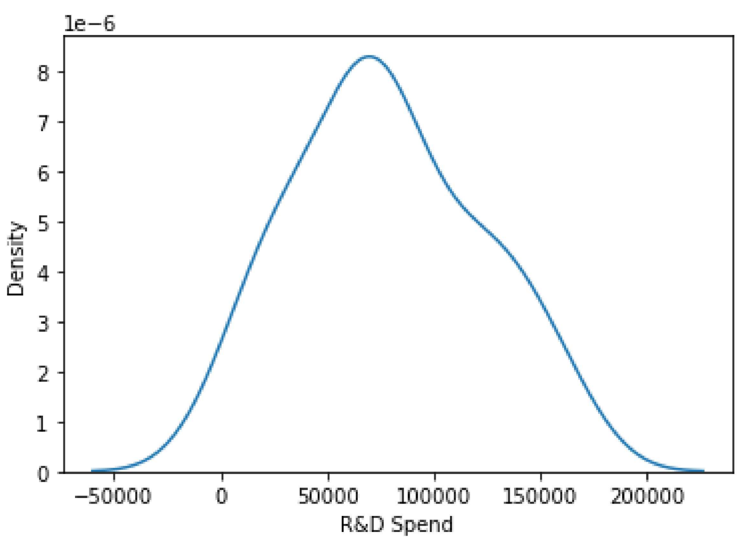
Out[119…

	R&D Spend	Administration	Marketing Spend	Profit
0	165349.20	136897.80	471784.10	192261.83
1	162597.70	151377.59	443898.53	191792.06
2	153441.51	101145.55	407934.54	191050.39
3	144372.41	118671.85	383199.62	182901.99
4	142107.34	91391.77	366168.42	166187.94

Data visualization

In [120…

```
sns.distplot(a = data_set['R&D Spend'],hist =False)
plt.show()
```



In [121…

```
data_set['R&D Spend'].skew()
```

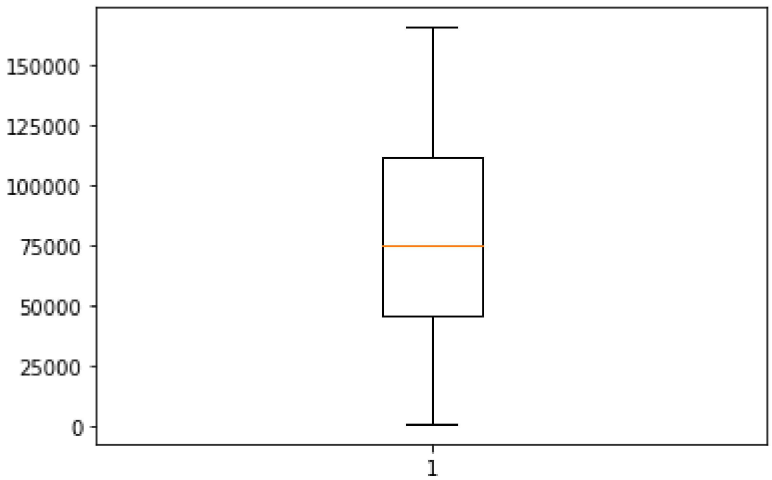
Out[121…

```
0.21386961761259485
```

In [122…

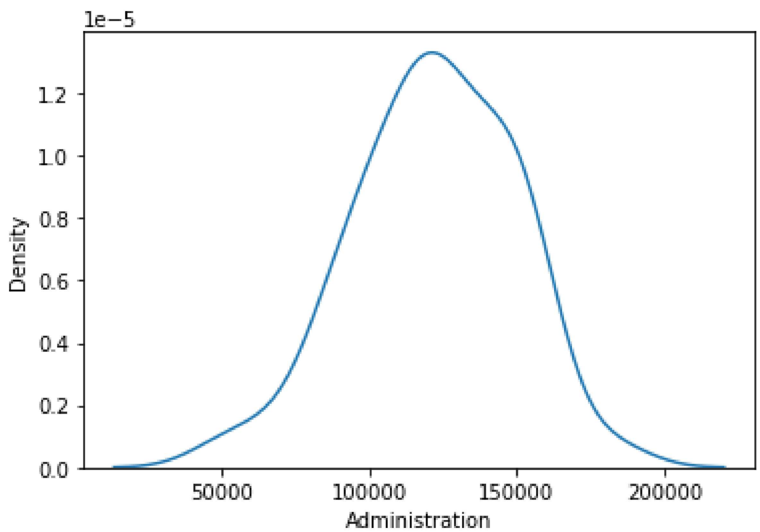
```
plt.boxplot(x = data_set['R&D Spend'])
```

```
plt.show()
```



In [123...

```
sns.distplot(a = data_set['Administration'],hist = False)
plt.show()
```



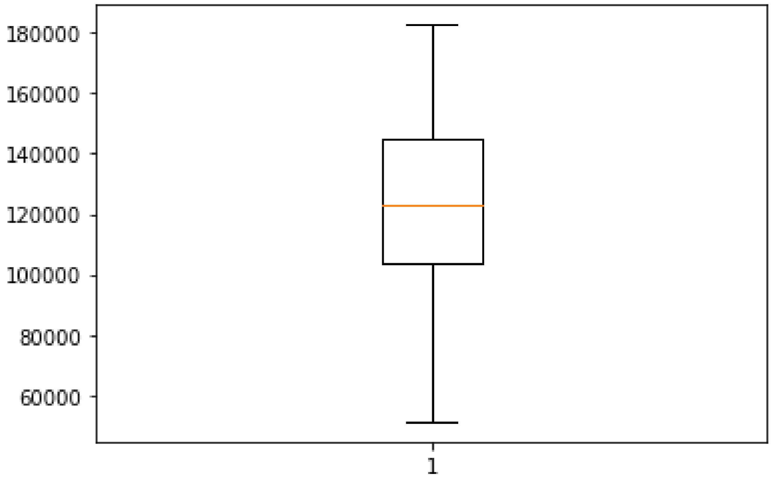
In [124...

```
data_set['Administration'].skew()
```

Out[124... -0.3172677129155058

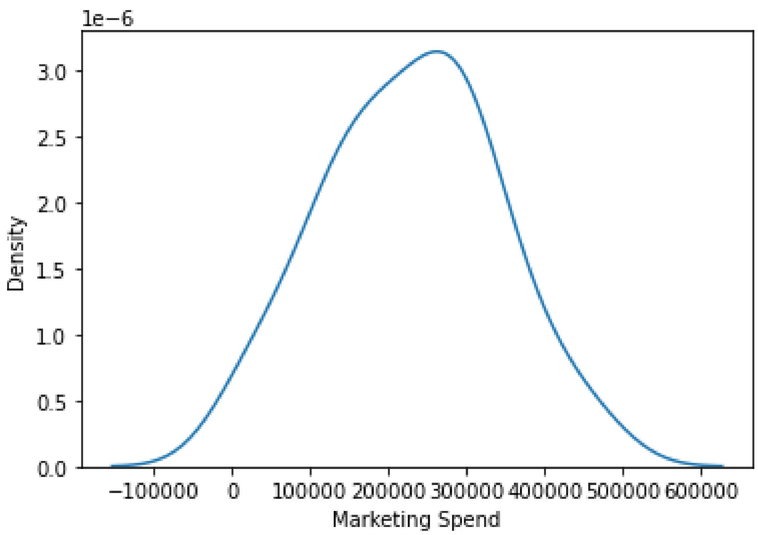
In [125...

```
plt.boxplot(data_set['Administration'])
plt.show()
```



In [126…

```
sns.distplot(a = data_set['Marketing Spend'],hist=False)
plt.show()
```



In [127…

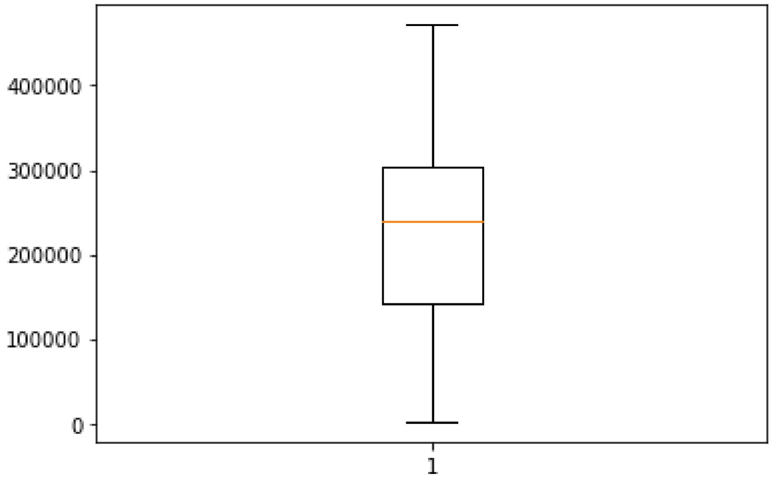
```
data_set['Marketing Spend'].skew()
```

Out[127…

```
-0.009246982105102828
```

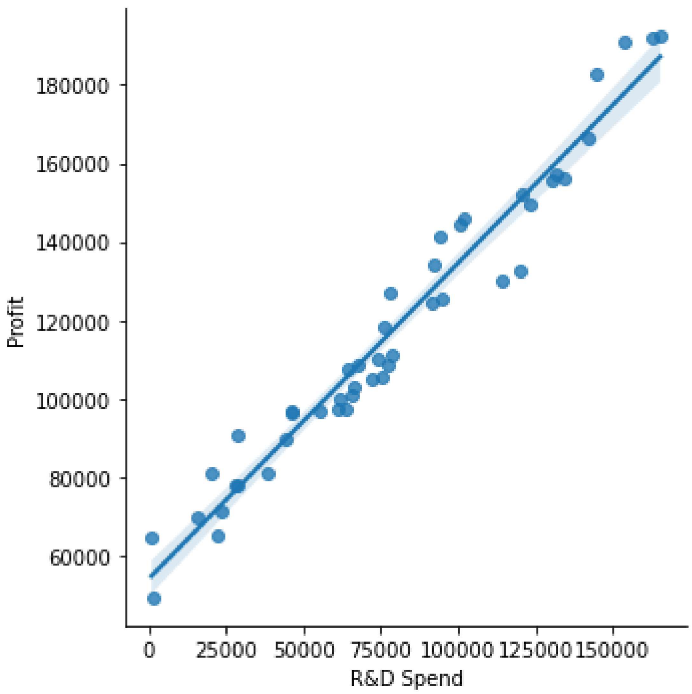
In [128…

```
plt.boxplot(data_set['Marketing Spend'])
plt.show()
```



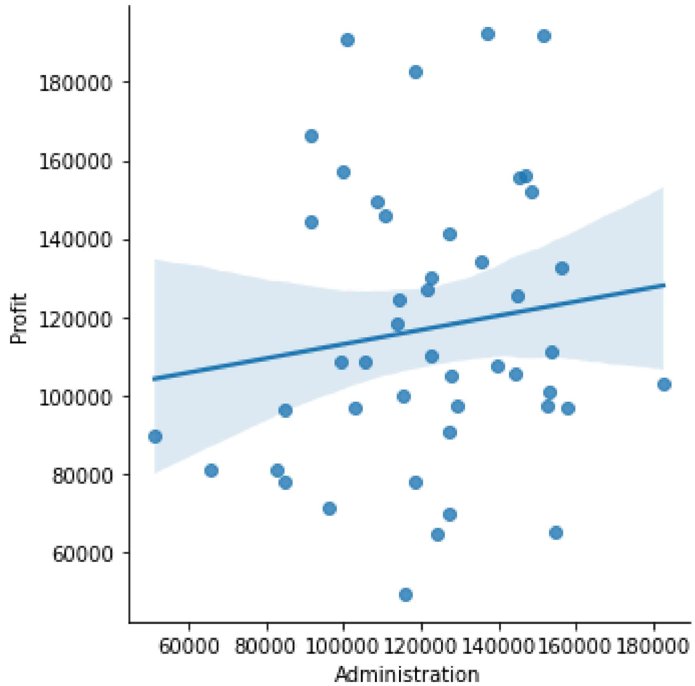
In [129…

```
sns.lmplot(x = 'R&D Spend',y ='Profit',data = data_set)
plt.show()
```



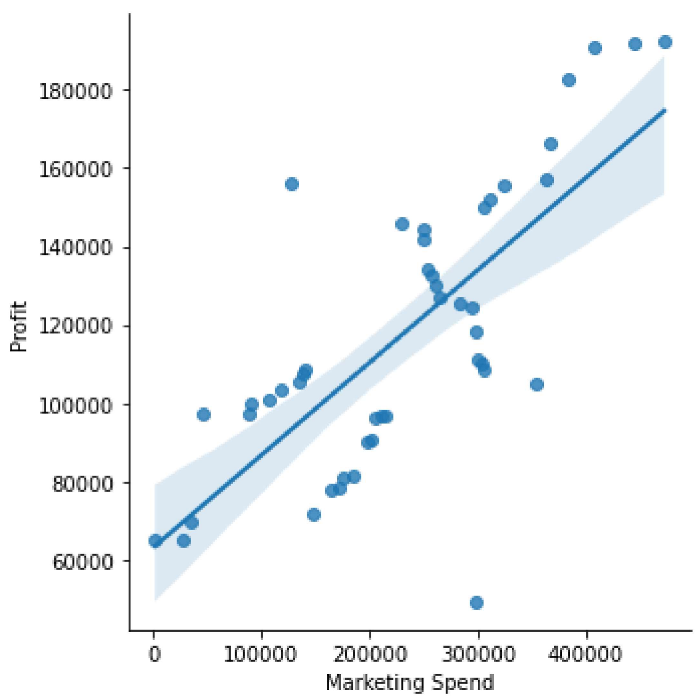
In [130…

```
sns.lmplot(x='Administration',y='Profit',data = data_set)
plt.show()
```



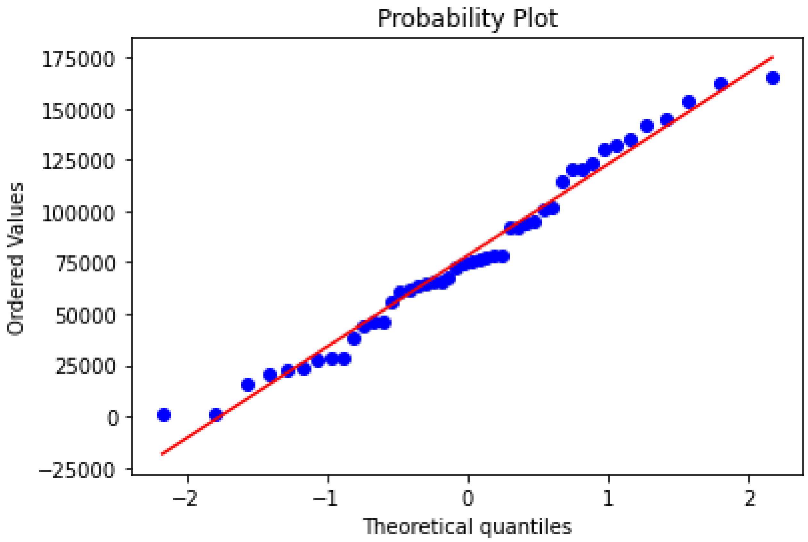
In [131…

```
sns.lmplot(x = 'Marketing Spend',y='Profit',data = data_set)
plt.show()
```



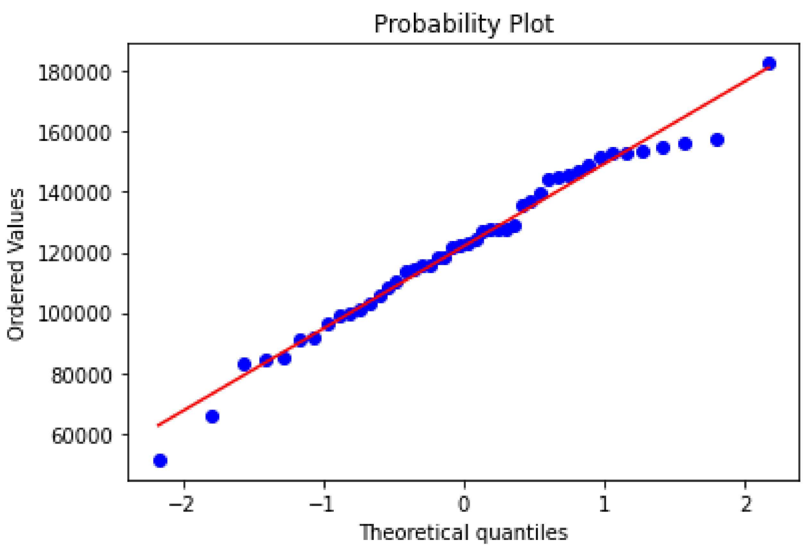
In [132…

```
stats.probplot(x = data_set['R&D Spend'],dist = 'norm',plot = plt)
plt.show()
```



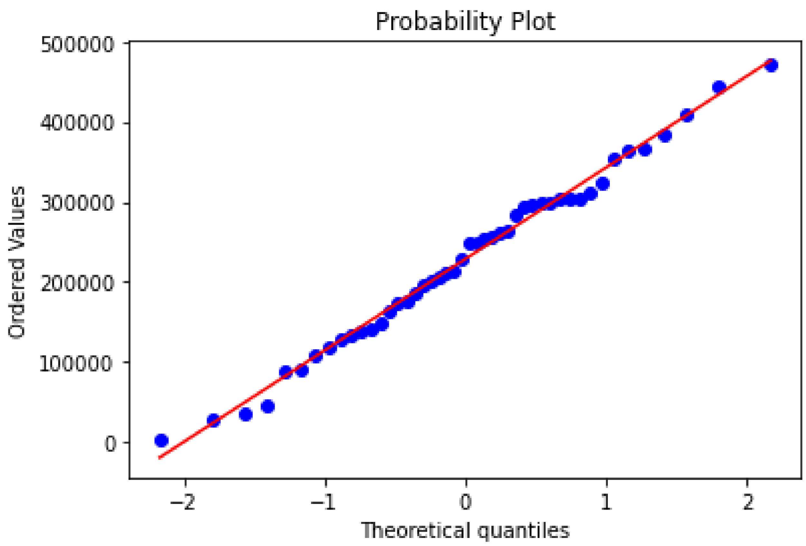
In [133…

```
stats.probplot(x = data_set['Administration'],dist = 'norm',plot = plt)
plt.show()
```



In [134…

```
stats.probplot(x = data_set['Marketing Spend'],dist = 'norm',plot = plt)
plt.show()
```



In [135…

```
data_set_corr_matrix = data_set.corr().round(4)
data_set_corr_matrix
```

Out[135…

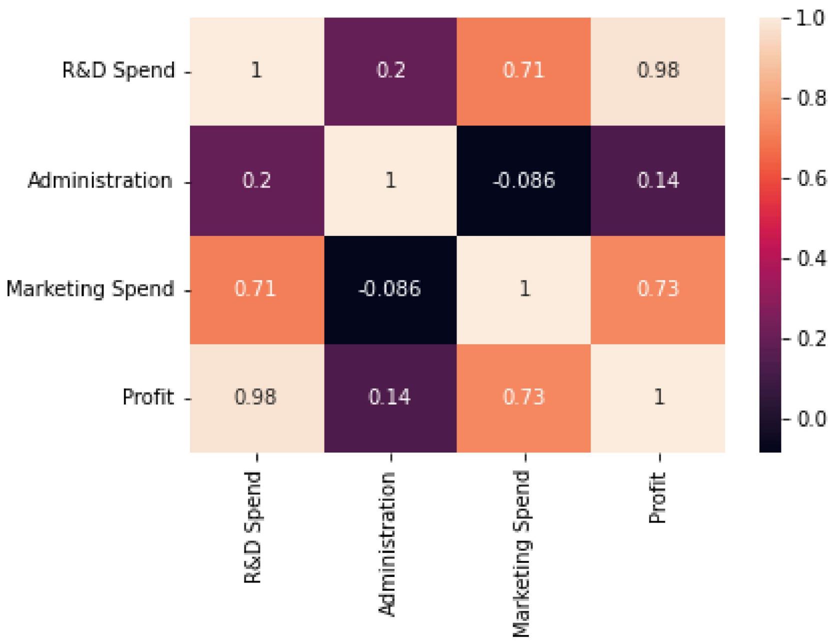
	R&D Spend	Administration	Marketing Spend	Profit
R&D Spend	1.0000	0.1961	0.7083	0.9777
Administration	0.1961	1.0000	-0.0858	0.1351
Marketing Spend	0.7083	-0.0858	1.0000	0.7323
Profit	0.9777	0.1351	0.7323	1.0000

In [136…

```
sns.heatmap(data_set_corr_matrix,annot=True)
plt.show
```

Out[136…

```
<function matplotlib.pyplot.show(close=None, block=None)>
```



In [137…

```
x = data_set.drop('Profit',axis = 1)
y = data_set[['Profit']]
```

In [138…

```
model = StandardScaler()
x_scaled = model.fit_transform(x)
x_scaled = pd.DataFrame(x_scaled,columns=['std_R&D Spend','std_Administration','std_Mar
x_scaled
```

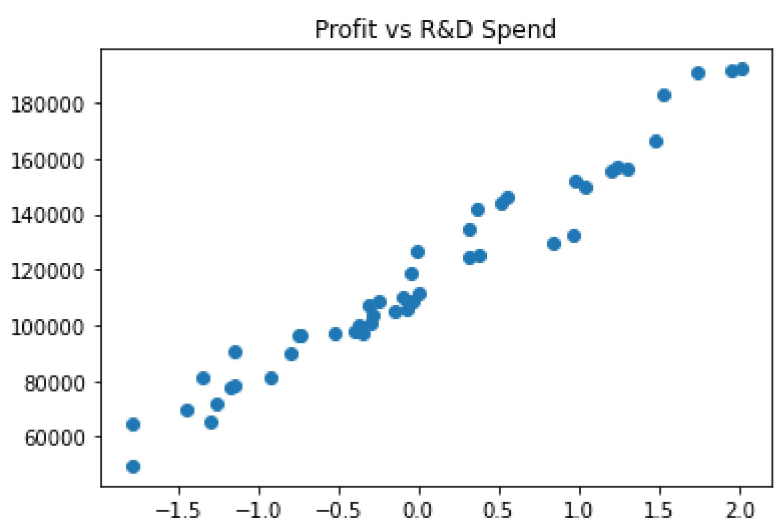
Out[138…

	std_R&D Spend	std_Administration	std_Marketing Spend
0	2.015547	0.563760	2.210237
1	1.951881	1.109761	1.957008
2	1.740019	-0.784378	1.630418
3	1.530173	-0.123500	1.405800
4	1.477762	-1.152172	1.251140
5	1.241044	-0.834561	1.221108
6	1.304411	0.952190	-0.914243
7	1.204514	0.889263	0.867088
8	0.978782	1.009509	0.755724
9	1.043394	-0.500302	0.695502
10	0.547723	-0.428094	0.006973
11	0.519005	-1.137132	0.193893
12	0.361472	0.202617	0.194755
13	0.318172	0.510866	0.220413
14	0.964916	1.304704	0.255357
15	0.839513	0.025257	0.303153
16	-0.005290	-0.013179	0.326489
17	0.379831	0.872201	0.492020
18	0.312544	-0.293037	0.604128
19	-0.045996	-0.304669	0.638135
20	0.003419	1.200103	0.647878
21	-0.098273	0.031513	0.680406
22	-0.247796	-0.610716	0.693568
23	-0.027713	-0.854673	-0.797479
24	-0.314153	0.663888	-0.821200
25	-0.067399	0.836696	-0.856730
26	-0.141935	0.223136	1.133226
27	-0.282064	2.288807	-1.001135

	std_R&D Spend	std_Administration	std_Marketing Spend
28	-0.292385	1.172147	-1.101116
29	-0.375939	-0.237776	-1.246477
30	-0.395794	1.159698	-1.272930
31	-0.343212	0.274232	-1.655540
32	-0.526352	-0.712283	-0.124939
33	-0.736170	1.347935	-0.159784
34	-0.745704	-1.391402	-0.207732
35	-1.147166	0.192655	-0.247605
36	-0.790688	-2.664579	-0.284814
37	-1.342321	-2.111602	-0.391646
38	-0.918215	-1.469282	-0.484870
39	-1.145070	-0.128244	-0.504881
40	-1.165002	-1.404097	-0.580480
41	-1.263387	-0.971255	-0.730041
42	-1.451624	0.204951	-1.751354
43	-1.297243	1.239044	-1.816733
44	-1.787261	0.083183	-2.056751
45	-1.779967	-0.231180	0.624060

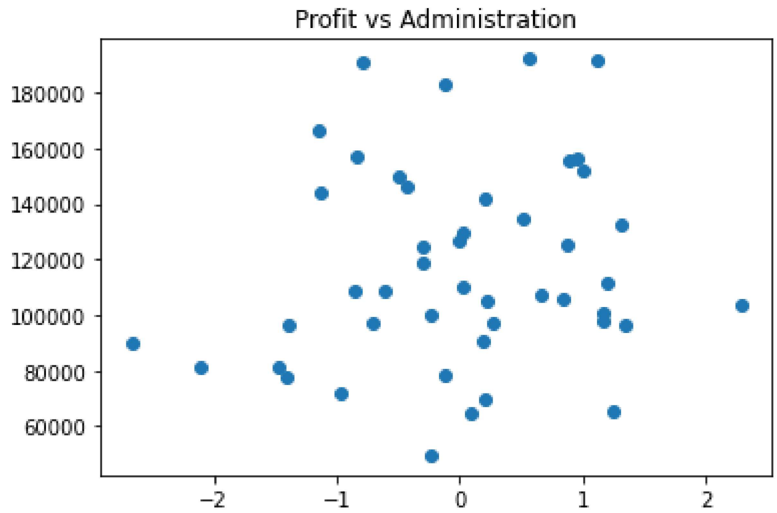
In [139…

```
plt.scatter(x=x_scaled['std_R&D Spend'],y=y)
plt.title('Profit vs R&D Spend')
plt.show()
```



In [140…

```
plt.scatter(x=x_scaled['std_Administration'],y=y)
plt.title('Profit vs Administration')
plt.show()
```



In [141...

```
plt.scatter(x=x_scaled['std_Marketing Spend'],y=y)
plt.title('Profit vs Marketing Spend')
plt.show()
```

In [142...

```
linear_model = LinearRegression()
linear_model.fit(x_scaled,y)
```

Out[142...

```
LinearRegression()
```

In [143...

```
linear_model.coef_
```

Out[143...

```
array([[33585.5015273 , -1598.79094766,  2120.62207726]])
```

In [144...

```
linear_model.intercept_
```

Out[144...

```
array([117063.92521739])
```

In [145...

```
y_pred = linear_model.predict(x_scaled)
```

```
y_pred
```

Out[145…

```
array([[188542.8230628 ],
       [184994.62678374],
       [180214.90791276],
       [171634.1734915 ],
       [171190.59125946],
       [162668.81166146],
       [157412.09273868],
       [157935.13591697],
       [149925.42607178],
       [154381.59493502],
       [136158.68492535],
       [136724.17370809],
       [129293.21181254],
       [127400.52166174],
       [147926.66860343],
       [145861.87995931],
       [117599.69951904],
       [129469.65085564],
       [129310.4892184 ],
       [117359.46320106],
       [116633.9373571 ],
       [115155.86574365],
       [111188.78105914],
       [115808.45242181],
       [103710.05841339],
       [111645.78404533],
       [114343.36306294],
       [101808.29936342],
       [103034.95633614],
       [102174.68520376],
       [ 99217.46993946],
       [101587.7677164 ],
       [100259.97404795],
       [ 89845.37125579],
       [ 93803.10685529],
       [ 77702.7026391 ],
       [ 94164.40969467],
       [ 74526.89333457],
       [ 87546.07447869],
       [ 77740.55319307],
       [ 78950.63823043],
       [ 74637.13647202],
       [ 64268.76335085],
       [ 67661.79016718],
       [ 52543.26892927],
       [ 58975.82938982]])
```

In [146…

```
error = y-y_pred
error
```

Out[146…

	Profit
0	3719.006937
1	6797.433216
2	10835.482087

	Profit
3	11267.816509
4	-5002.651259
5	-5677.691661
6	-1289.582739
7	-2182.535917
8	2286.343928
9	-4621.634935
10	9963.265075
11	7535.226292
12	12292.308187
13	6906.828338
14	-15324.018603
15	-15944.839959
16	9393.230481
17	-4099.280856
18	-5043.589218
20	1114.566799
21	-5320.917357
22	-4803.615744
23	-2454.791059
24	-7256.412422
25	3694.281587
26	-5912.244045
27	-9335.053063
28	1474.080637
29	-2030.316336
30	-2237.095204
31	-1733.909939
32	-4159.927716
33	-3481.054048
34	6867.428744
35	2676.403145
36	13005.487361

	Profit
37	-4215.269695
38	6702.166665
39	-6540.314479
40	499.356807
41	-1151.808230
42	-3138.646472
43	5490.216649
44	-2461.460167
45	12382.811071
46	-9485.079390

In [147…

```
mean_absolute_error(y,y_pred)
```

Out[147…

```
5865.380022385375
```

In [148…

```
mean_absolute_percentage_error(y,y_pred)
```

Out[148…

```
0.054140744295296826
```

In [149…

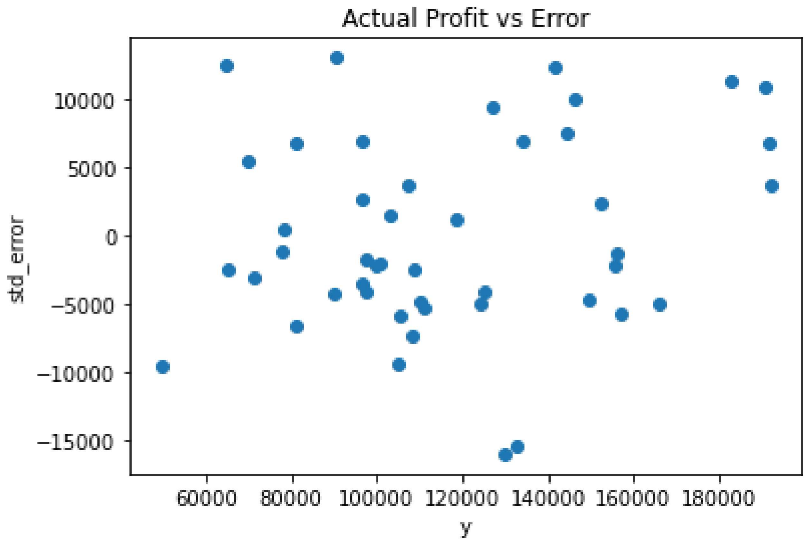
```
mean_squared_error(y,y_pred)
```

Out[149…

```
49559086.4165616
```

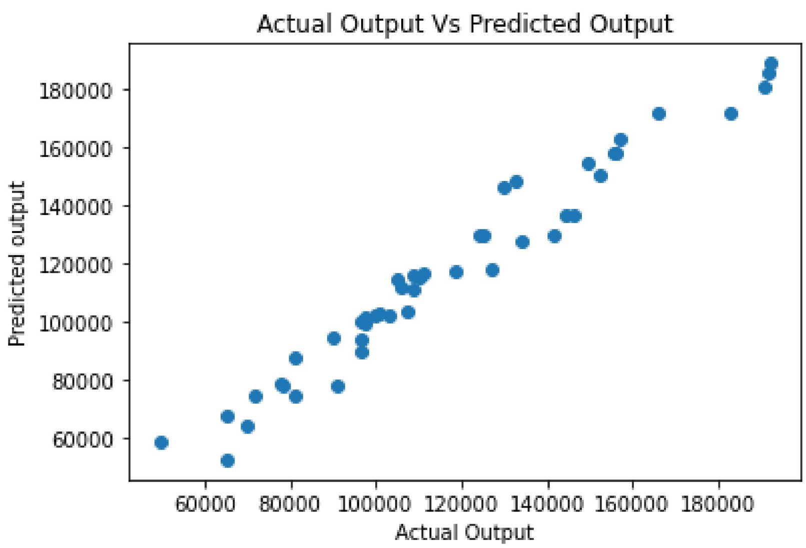
In [150…

```
plt.scatter(y,error)
plt.xlabel('y')
plt.ylabel('std_error')
plt.title('Actual Profit vs Error')
plt.show()
```



In [151…

```
plt.scatter(y,y_pred)
plt.title('Actual Output Vs Predicted Output')
plt.xlabel('Actual Output')
plt.ylabel('Predicted output')
plt.show()
```



In [154…

```
x_scaled = x_scaled.rename(columns={'std_R&D Spend':'RDS','std_Administration':'ADM','s
```

In [155…

```
x_scaled['Profit'] = y
x_scaled.head()
```

Out[155…

	RDS	ADM	MRKS	Profit
0	2.015547	0.563760	2.210237	192261.83
1	1.951881	1.109761	1.957008	191792.06
2	1.740019	-0.784378	1.630418	191050.39
3	1.530173	-0.123500	1.405800	182901.99

	RDS	ADM	MRKS	Profit
4	1.477762	-1.152172	1.251140	166187.94

In [156…

```
model_1 = smf.ols(formula = 'Profit ~ RDS',data = x_scaled).fit()
print('R-Square                    :',round(model_1.rsquared,4))
print('Adjusted R-square           :',round(model_1.rsquared_adj,4))
print('AIC                         :',round(model_1.aic,4))
print('BIc                         :',round(model_1.bic,4))
```

```
R-Square                    : 0.9487
Adjusted R-square           : 0.9475
AIC                         : 938.4225
BIc                         : 942.0359
```

In [160…

```
model_2 = smf.ols(formula = 'Profit ~ RDS+ADM',data = x_scaled).fit()
print('R-Square                    :',round(model_2.rsquared,4))
print('Adjusted R-square           :',round(model_2.rsquared_adj,4))
print('AIC                         :',round(model_2.aic,4))
print('BIc                         :',round(model_2.bic,4))
```

```
R-Square                    : 0.9517
Adjusted R-square           : 0.9494
AIC                         : 937.7134
BIc                         : 943.1334
```

In [162…

```
model_3 = smf.ols(formula = 'Profit ~ RDS+ADM+MRKS',data = x_scaled).fit()
print('R-Square                    :',round(model_3.rsquared,4))
print('Adjusted R-square           :',round(model_3.rsquared_adj,4))
print('AIC                         :',round(model_3.aic,4))
print('BIc                         :',round(model_3.bic,4))
```

```
R-Square                    : 0.9572
Adjusted R-square           : 0.9541
AIC                         : 934.2043
BIc                         : 941.4309
```

In [166…

```
data_set = data_set.rename(columns={'R&D Spend':'RDS','Administration':'ADM','Marketing
```

In [167…

```
log_x = np.log(data_set.drop('Profit',axis = 1))
log_y = data_set[['Profit']]
```

In [168…

```
log_x.head()
```

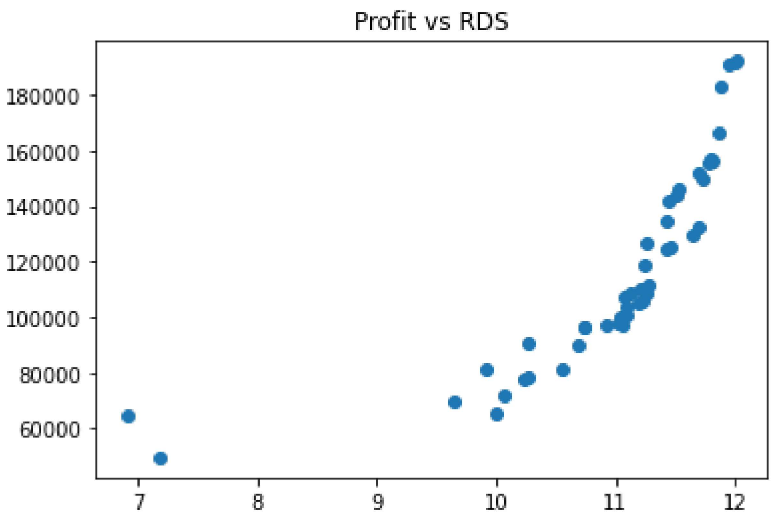
Out[168…

	RDS	ADM	MRKS
0	12.015815	11.826990	13.064277
1	11.999034	11.927533	13.003351
2	11.941075	11.524316	12.918862
3	11.880151	11.684117	12.856311

	RDS	ADM	MRKS
4	11.864338	11.422911	12.810849

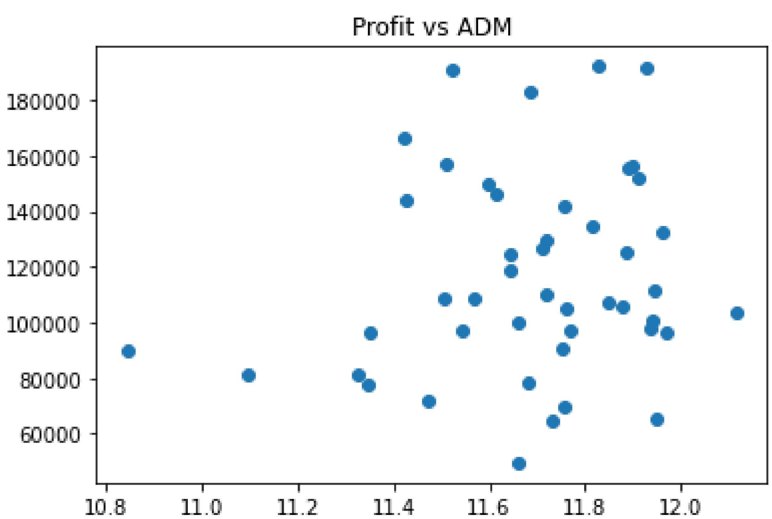
In [169…

```
plt.scatter(x=log_x['RDS'],y=log_y)
plt.title('Profit vs RDS')
plt.show()
```



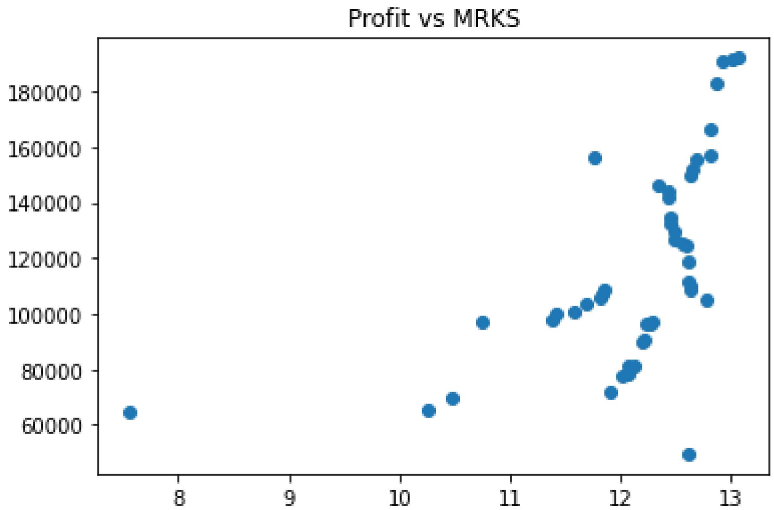
In [170…

```
plt.scatter(x=log_x['ADM'],y=log_y)
plt.title('Profit vs ADM')
plt.show()
```



In [171…

```
plt.scatter(x=log_x['MRKS'],y=log_y)
plt.title('Profit vs MRKS')
plt.show()
```



In [172…

```
linear_model_2 = LinearRegression()
```

In [173…

```
linear_model_2.fit(log_x,log_y)
```

Out[173… LinearRegression()

In [174…

```
linear_model_2.coef_
```

Out[174…

```
array([[23876.95032092, 10674.92455481,  4809.28567067]])
```

In [175…

```
linear_model_2.intercept_
```

Out[175…

```
array([-328028.6101341])
```

In [176…

```
log_y_pred = linear_model_2.predict(log_x)
log_y_pred
```

Out[176…

```
array([[147954.46910188],
       [148334.07798969],
       [142239.53811076],
       [142189.92075091],
       [138805.33904455],
       [137918.87221202],
       [137534.61868678],
       [141109.85247259],
       [139297.43446899],
       [136392.50178307],
       [130649.01867171],
       [128780.77562625],
       [130603.46456391],
       [130840.99121316],
       [138790.25481553],
       [135176.09638436],
       [125967.50950406],
       [132790.31714992],
       [129693.97274194],
```

```
       [125308.77122517],
       [129192.77861182],
       [125469.7190631 ],
       [121716.6433655 ],
       [120467.45030752],
       [119829.61038521],
       [123680.93868257],
       [126017.83415117],
       [122463.26140815],
       [119942.64433956],
       [114822.00833762],
       [117300.53970092],
       [113266.75656827],
       [115067.1249283 ],
       [115261.19564638],
       [108335.13830577],
       [101215.1484265 ],
       [101701.68800387],
       [ 85499.03143477],
       [103078.8974305 ],
       [ 99820.22832407],
       [ 95269.12834683],
       [ 92169.15511562],
       [ 78235.32742594],
       [ 87772.94758489],
       [ -1559.74961912],
       [ 28527.31720701]])
```

In [177…

```
lg_error = log_y - log_y_pred
lg_error
```

Out[177…

	Profit
0	44307.360898
1	43457.982010
2	48810.851889
3	40712.069249
4	27382.600955
5	19072.247788
6	18587.891313
7	14642.747527
8	12914.335531
9	13367.458217
10	15472.931328
11	15478.624374
12	10982.055436
13	3466.358787
14	-6187.604816

	Profit
15	-5259.056384
16	1025.420496
17	-7419.947150
18	-5427.072742
20	-6834.741225
21	-17879.758612
22	-15117.469063
23	-12982.653365
24	-11915.410308
25	-12425.270385
26	-17947.398683
27	-21009.524151
28	-19180.881408
29	-18938.004340
30	-14884.418338
31	-19816.979701
32	-15838.916568
33	-18288.204928
34	-18548.395646
35	-11855.628306
36	-10506.958427
37	-11752.548004
38	-4269.971435
39	-22073.137431
40	-21580.318324
41	-17470.298347
42	-20670.665116
43	-8476.347426
44	-22572.617585
45	66485.829619
46	20963.432793

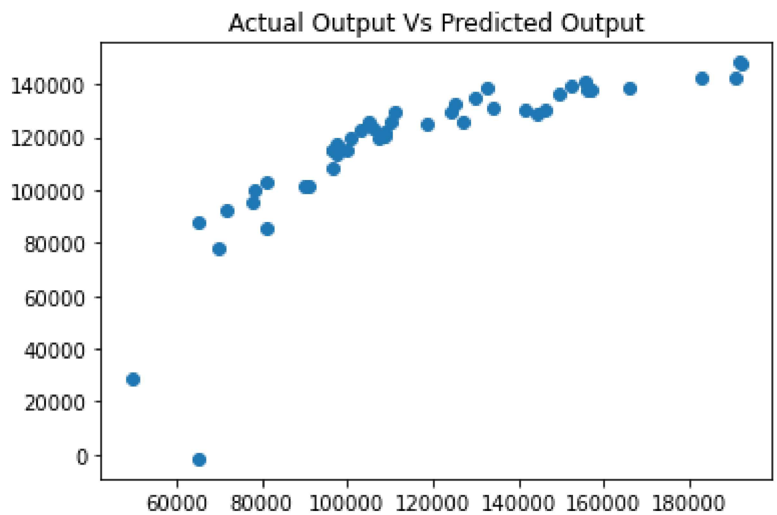
In [178…

```
mean_absolute_percentage_error(log_y,log_y_pred)
```

Out[178… 0.1708947103355232

In [180…
```
plt.scatter(log_y,lg_error)
plt.title('Actual Price vs Error')
plt.show()
```

In [181…
```
plt.scatter(log_y,log_y_pred)
plt.title('Actual Output Vs Predicted Output')
plt.show()
```



In [182…
```
log_x['Profit'] = log_y
log_x.head()
```

Out[182…

	RDS	ADM	MRKS	Profit
0	12.015815	11.826990	13.064277	192261.83
1	11.999034	11.927533	13.003351	191792.06
2	11.941075	11.524316	12.918862	191050.39
3	11.880151	11.684117	12.856311	182901.99

	RDS	ADM	MRKS	Profit
4	11.864338	11.422911	12.810849	166187.94

In [183…

```
lg_model = smf.ols(formula = 'Profit ~ RDS',data = log_x).fit()
print('R-Square                          :',round(lg_model.rsquared,4))
print('Adjusted R-square                 :',round(lg_model.rsquared_adj,4))
print('AIC                               :',round(lg_model.aic,4))
print('BIc                               :',round(lg_model.bic,4))
```

```
R-Square                          : 0.6066
Adjusted R-square                 : 0.5977
AIC                               : 1055.71
BIc                               : 1059.3673
```

In [184…

```
lg_model_2 = smf.ols(formula = 'Profit ~ RDS+ADM',data = log_x).fit()
print('R-Square                          :',round(lg_model_2.rsquared,4))
print('Adjusted R-square                 :',round(lg_model_2.rsquared_adj,4))
print('AIC                               :',round(lg_model_2.aic,4))
print('BIc                               :',round(lg_model_2.bic,4))
```

```
R-Square                          : 0.6087
Adjusted R-square                 : 0.5905
AIC                               : 1057.4694
BIc                               : 1062.9553
```

In [185…

```
lg_model_3 = smf.ols(formula = 'Profit ~ RDS+ADM+MRKS',data = log_x).fit()
print('R-Square                          :',round(lg_model_3.rsquared,4))
print('Adjusted R-square                 :',round(lg_model_3.rsquared_adj,4))
print('AIC                               :',round(lg_model_3.aic,4))
print('BIc                               :',round(lg_model_3.bic,4))
```

```
R-Square                          : 0.6168
Adjusted R-square                 : 0.5894
AIC                               : 1058.5087
BIc                               : 1065.8232
```

In [ ]: