# Multiple Linear Regression Results

1. Introduction:

   The goal of this analysis was to examine the business value of online search advertising strategies. Data was collected from the websiteoutlook.com website. More specifically, we wanted to examine whether online search strategies have impact on revenue and profits. The results of the analysis are presented below.

2. Exploratory Data Analysis

   As shown in the sample charts below, before transforming the data, outliers were noticeable, and the distributions were skewed:

Profit vs. Revenue
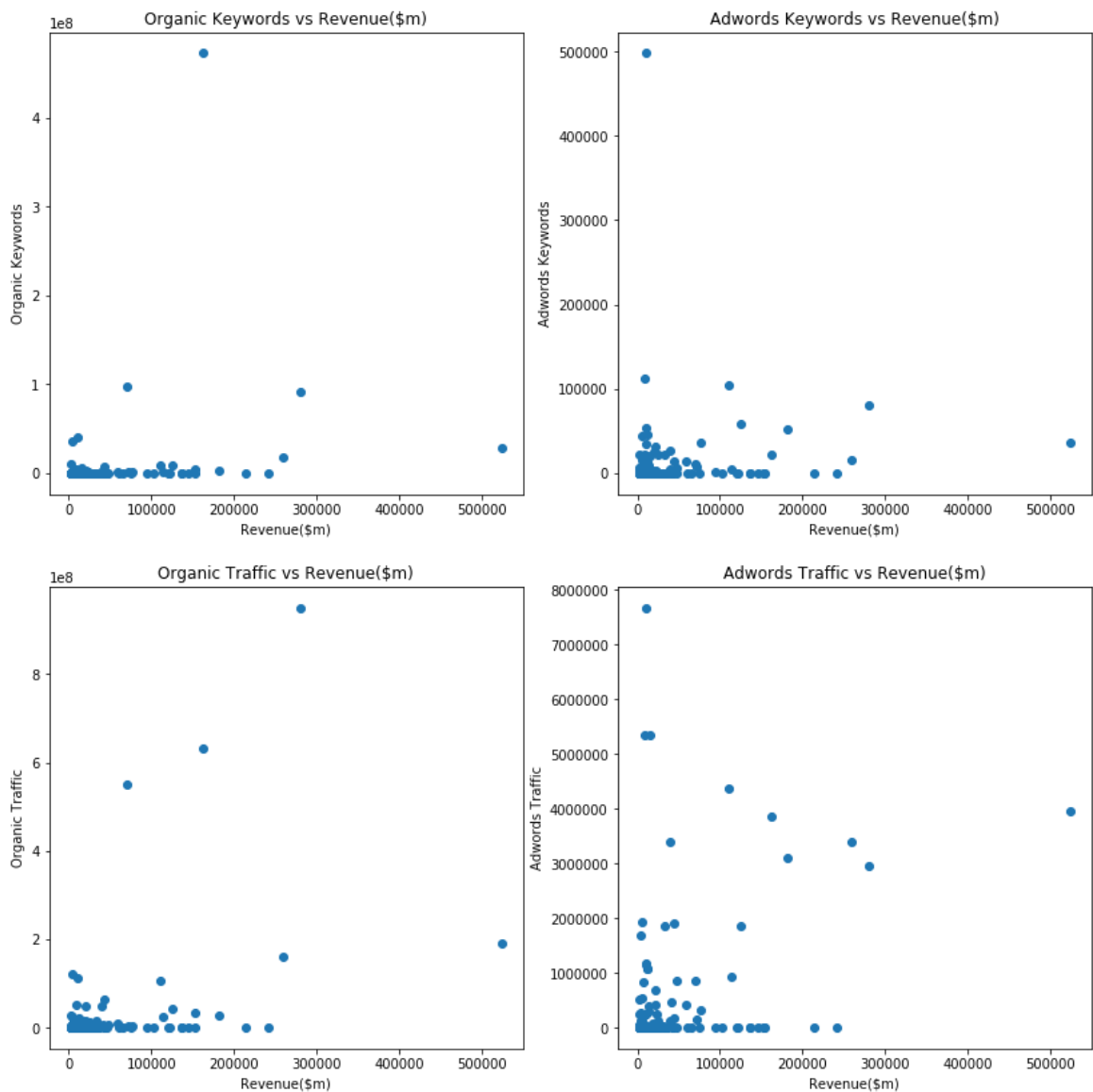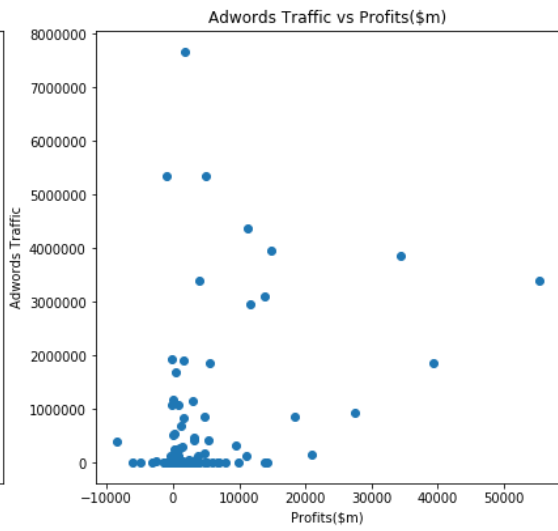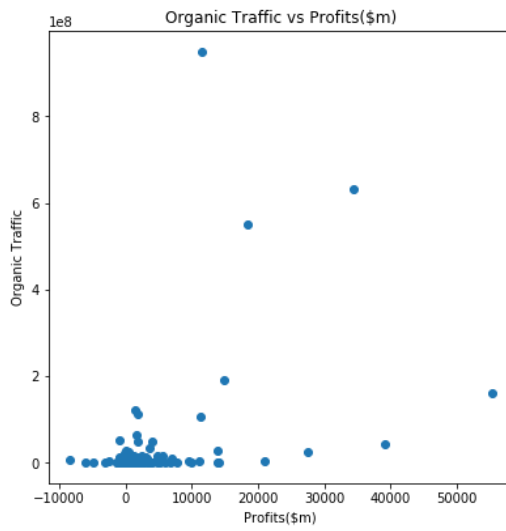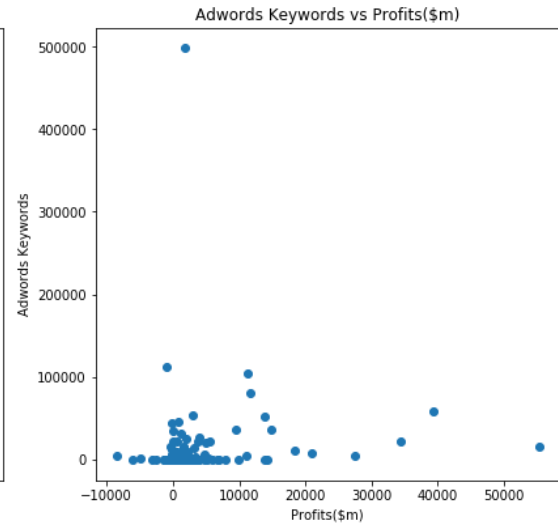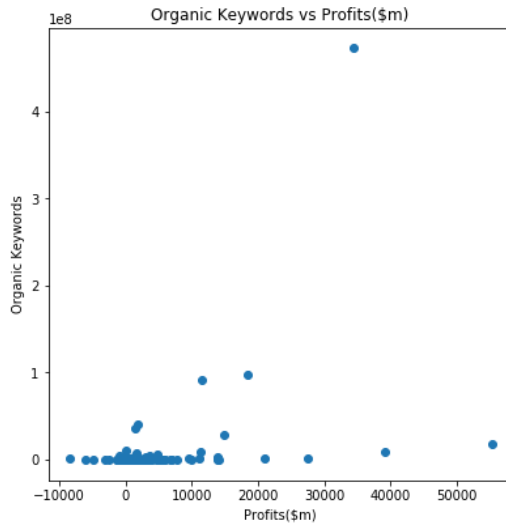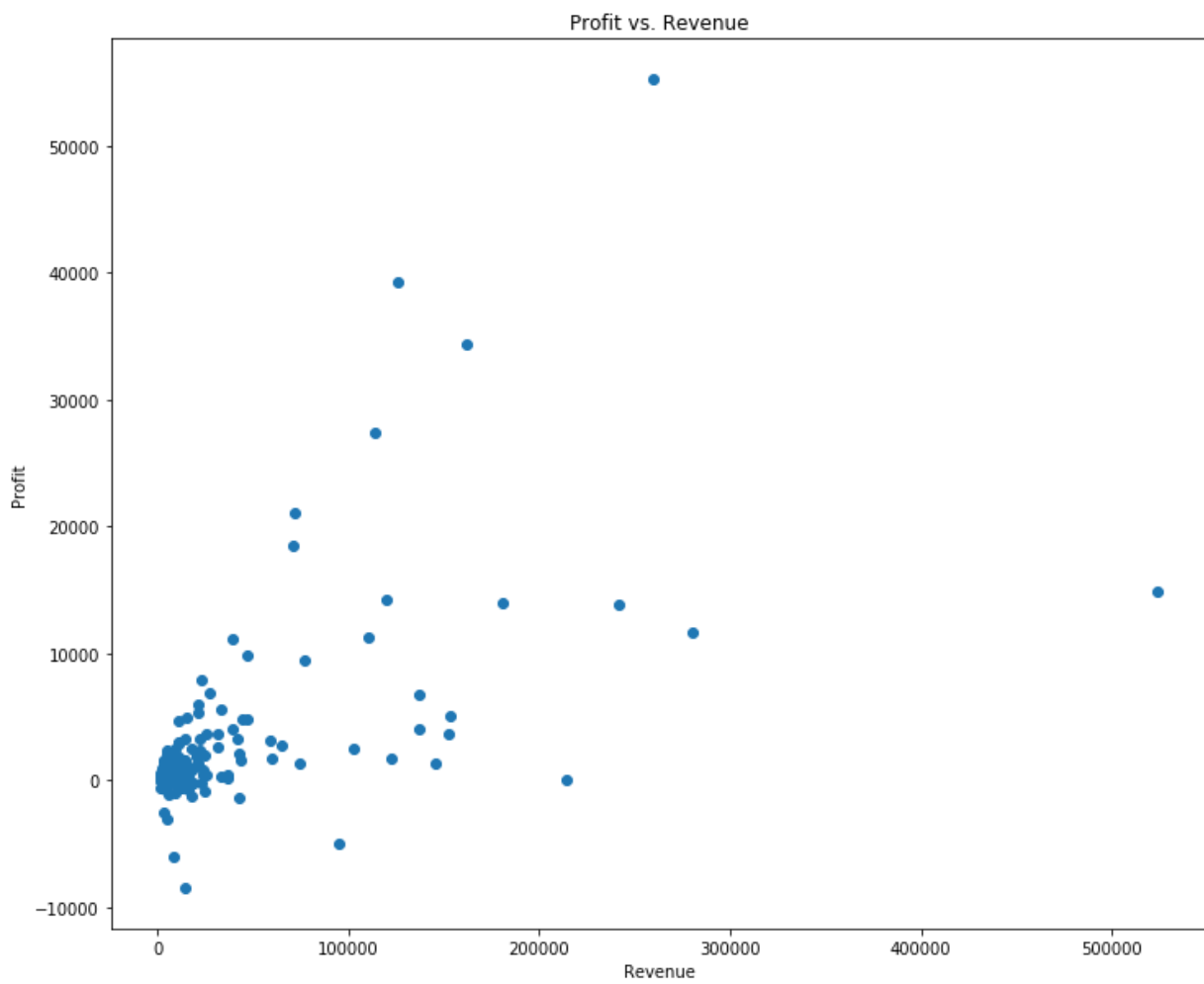
The seaborn heatmap below show variable correlations:
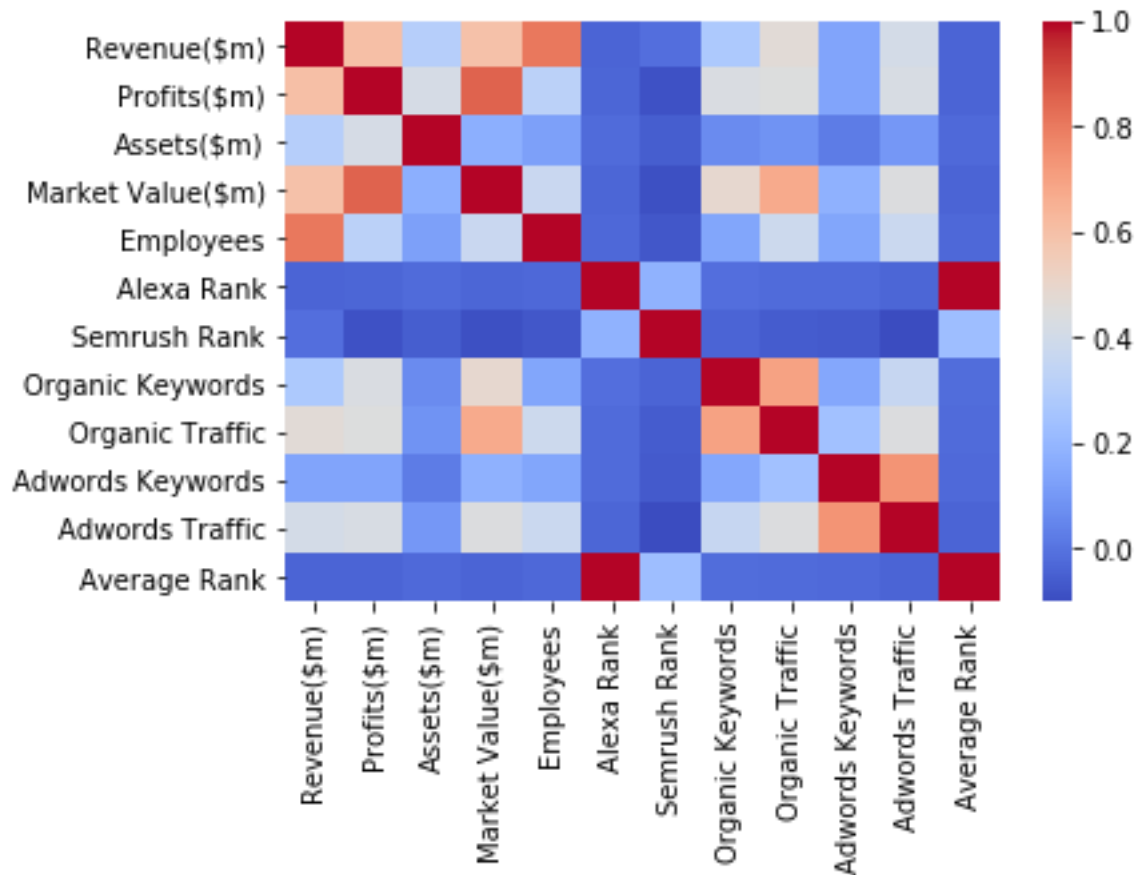
The dependent (outcome) variables in our analysis are Profit and Revenue.

There are extremely high correlations between Assets and Revenue, Revenue and Average rank, Profit and Market Value, Profit/Revenue and Adwords Keywords, Revenue and Adwords Keywords.

3. **Results of the Explanation Model**

Overall Model Assessment (F-statistic and r-squared)

Profits:-

From the results the F-statistic is 20.53 and p-value <0.01. Therefore the model is statistically significant. The adjusted R-Squared is 0.250. This implies that the model variables explain only 25% of the variation in Profit.

Revenue:-

From the results the F-statistic is 33.03 and p-value <0.01. Therefore the model is statistically significant. The adjusted R-Squared is 0.353. This implies that the model variables explain only 35.3% of the variation in Profit.

Individual Coefficients (t ratios)

Profits:-

For the individual coefficients, estimated coefficients for Organic Traffic (b=9237.25, p>0.05) is not statistically as significant. The estimate coefficients for Assets (b=2.199e+04, p<0.01) and Adwords Traffic (b=1858e+04, p<0.01) are statistically significant.

```
                        OLS Regression Results
==============================================================================
Dep. Variable:          Profits($m)   R-squared:                       0.263
Model:                          OLS   Adj. R-squared:                  0.250
Method:               Least Squares   F-statistic:                     20.53
Date:              Mon, 15 Jun 2020   Prob (F-statistic):           1.99e-11
Time:                      17:44:21   Log-Likelihood:                 -1758.0
No. Observations:               177   AIC:                             3524.
Df Residuals:                   173   BIC:                             3537.
Df Model:                         3
Covariance Type:            nonrobust
==============================================================================
                   coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const            778.6016    400.703      1.943      0.054     -12.294    1569.497
Assets($m)       2.199e+04   4938.084     4.454      0.000    1.22e+04    3.17e+04
Organic Traffic  9237.2518   5368.988     1.720      0.087   -1359.903    1.98e+04
Adwords Traffic  1.858e+04   3983.256     4.664      0.000    1.07e+04    2.64e+04
==============================================================================
Omnibus:                      224.462   Durbin-Watson:                   1.881
Prob(Omnibus):                  0.000   Jarque-Bera (JB):            11936.707
Skew:                           5.111   Prob(JB):                         0.00
Kurtosis:                      41.911   Cond. No.                         15.3
==============================================================================
```

Revenue:-

The estimate coefficients for Organic traffic(b= 2.803e+05, p<0.01), Assets (b= 1.643e+05, p<0.01) and Adwords Traffic (b= 1.465e+05, p<0.01) are statistically significant.

```
                        OLS Regression Results
==============================================================================
Dep. Variable:            Revenue($m)   R-squared:                       0.364
Model:                           OLS   Adj. R-squared:                  0.353
Method:                Least Squares   F-statistic:                     33.03
Date:               Mon, 15 Jun 2020   Prob (F-statistic):           6.23e-17
Time:                       17:44:22   Log-Likelihood:                 -2151.7
No. Observations:                177   AIC:                             4311.
Df Residuals:                    173   BIC:                             4324.
Df Model:                          3
Covariance Type:           nonrobust
==============================================================================
                    coef    std err          t      P>|t|      [0.025      0.975]
------------------------------------------------------------------------------
const            1.257e+04   3704.794      3.393      0.001    5258.231    1.99e+04
Assets($m)       1.643e+05   4.57e+04      3.598      0.000    7.42e+04    2.54e+05
Organic Traffic  2.803e+05   4.96e+04      5.646      0.000    1.82e+05    3.78e+05
Adwords Traffic  1.465e+05   3.68e+04      3.979      0.000    7.38e+04    2.19e+05
==============================================================================
Omnibus:                     196.495   Durbin-Watson:                   2.023
Prob(Omnibus):                 0.000   Jarque-Bera (JB):             6010.945
Skew:                          4.314   Prob(JB):                         0.00
Kurtosis:                     30.214   Cond. No.                         15.3
==============================================================================
```

Managerial Implications

Based on the analysis, management should focus more on Organic traffic, and Adwords traffic, majorly on Adwords traffic. Improvement to regulate this traffic will help generate more profit and revenue. Assets directly impact profit and revenue, maintaining those will help keep a stability in profits and revenue.

4. **Prediction Results:**

   After explanatory, prediction models were developed (OLS, Ridge and Lasso). Following are RMSE stats:

Profits

Method	Alpha	RMSE
OLS	-	3034.631708958245
Ridge		3434.855601080441
Lasso		3034.63894324319

Revenue

Method	Alpha	RMSE
OLS	-	33456.601265177684
Ridge		28756.969314745525
Lasso		33456.59250530985

The make out-of-sample predictions, we used the Lasso model to predict the expected Profits and Revenue:

Input	Asset ($m)	Organic Traffic	Adwords Traffic
Transform ed value	0.25	0.15	0.21
Actual value	875830	142562520	1611306

The predicted value for profit was **11563.56046845** and for revenue was **126447.49003641**