QualityAlloys

Manish Grewal

26/07/2020

Table of Contents

[PART 1 - EXECUTIVE SUMMARY 3](#_Toc47060962)

[1. How many people visit the website, and how do they come to website 3](#_Toc47060963)

[Unique visits over time 3](#_Toc47060964)

[Summary of unique visits by period 3](#_Toc47060965)

[Mean unique visits by period 4](#_Toc47060966)

[How do people come to the website: 4](#_Toc47060967)

[2. Is the website generating interest, and does this interest yield actual sales? 4](#_Toc47060968)

[Summary of Lbs. Sold by Period: 4](#_Toc47060969)

[Mean Lbs. Sold by period 5](#_Toc47060970)

[Comparison of Unique Visits and Lbs. Sold after Z-score standardization: 5](#_Toc47060971)

[Scatterplot of Lbs. Sold and Unique Vists 6](#_Toc47060972)

[Coefficient of correlation of Lbs. Sold and Unique Visits 6](#_Toc47060973)

[3. Check whether the data variables (if applicable) follows the normality assumption. 7](#_Toc47060974)

[Normality assumption for Revenue 7](#_Toc47060975)

[Normality assumption for Profit 7](#_Toc47060976)

[Normality assumption for UniqueVisits 8](#_Toc47060977)

[Evaluation of the goals of the company website 8](#_Toc47060978)

[(a) drive new sales - NOT MET 8](#_Toc47060979)

[(b) make product and contact information available - MAY BE MET 8](#_Toc47060980)

[(c) give or add legitimacy to its brand. - MAY BE MET 8](#_Toc47060981)

[Recommendations from demographics: 9](#_Toc47060982)

[PART 2 - DATA ANALYSIS 10](#_Toc47060983)

[1. Column Charts 10](#_Toc47060984)

[a. Unique visits over time 10](#_Toc47060985)

[b. Revenue over time, 10](#_Toc47060986)

[c. Profit over time 11](#_Toc47060987)

[d. Pounds sold over time. 11](#_Toc47060988)

[2. Summary by period 11](#_Toc47060989)

[a. Summary table of Initial Period 11](#_Toc47060990)

[b. Summary table of Pre-promotion Period 12](#_Toc47060991)

[c. Summary table of Promotion Period 12](#_Toc47060992)

[d. Summary table of Post-promotion Period 12](#_Toc47060993)

[3. Column chart of the means over the four periods 13](#_Toc47060994)

[a. Mean Visits by Period 13](#_Toc47060995)

[b. Mean Unique Visits by Period 13](#_Toc47060996)

[c. Mean Revenue by Period 14](#_Toc47060997)

[d. Mean Profit by Period 14](#_Toc47060998)

[e. Mean Lbs. Sold by Period 15](#_Toc47060999)

[4. Summary of findings thus far 15](#_Toc47061000)

[5. Scatter plot of Revenue and Lbs. Sold 15](#_Toc47061001)

[b. Coefficient of Correlation 16](#_Toc47061002)

[6. Scatter plot of Visits and Revenue 16](#_Toc47061003)

[b. Coefficient of Correlation 16](#_Toc47061004)

[7. Summary of results 16](#_Toc47061005)

[8. Modeling data 16](#_Toc47061006)

[a. Summary of Lbs. Sold data per week from January 3, 2005, through the week of July 19, 2010. 16](#_Toc47061007)

[b. Histogram of the pounds of material sold data. 17](#_Toc47061008)

[c. Histogram description 17](#_Toc47061009)

[d. Empirical rule 17](#_Toc47061010)

[e. Refine empirical rule 17](#_Toc47061011)

[f. Is pound of material sold bell shaped? 18](#_Toc47061012)

[g. Skweness and Kurtosis 18](#_Toc47061013)

[9. Comparison of the distribution of the pounds sold data with daily visit data. 19](#_Toc47061014)

[10. Demographics 20](#_Toc47061015)

[a. All Traffic Sources 20](#_Toc47061016)

[b. Top Ten Referring Sites 20](#_Toc47061017)

[c. Top Ten Search Engine Sources of Visits 21](#_Toc47061018)

[d. Top Ten Geographic Sources by Sub Continent Region 22](#_Toc47061019)

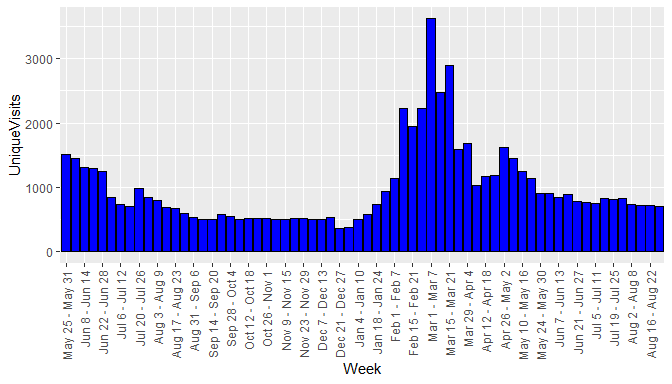
[e. Top Ten Browsers Used 22](#_Toc47061020)

[f. Top Ten Operating Systems Used 23](#_Toc47061021)

# PART 1 - EXECUTIVE SUMMARY

## 1. How many people visit the website, and how do they come to website

### Unique visits over time

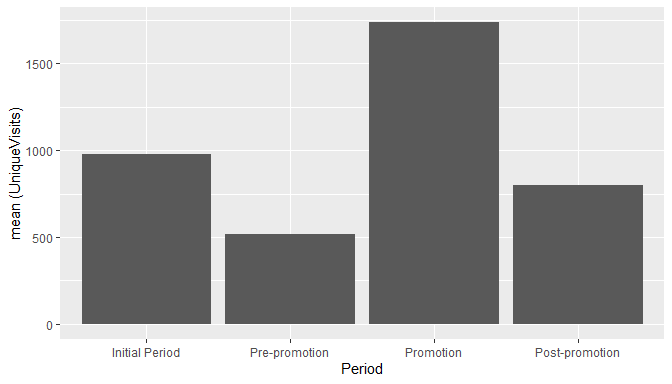


### Summary of unique visits by period

UNIQUE VISTS BY PERIOD

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Period | mean | median | sd | minimum | maximum |
| Initial Period | 975.9286 | 845.5 | 319.59721 | 594 | 1509 |
| Pre-promotion | 516.8095 | 510.0 | 70.93914 | 366 | 734 |
| Promotion | 1738.8235 | 1585.0 | 743.02534 | 930 | 3617 |
| Post-promotion | 800.7857 | 800.0 | 72.35615 | 709 | 912 |

### Mean unique visits by period



### How do people come to the website:

* Most of the people come through referring sites, search engines and direct traffic in that order.
* Google and sedoparking.com are the predominant sources of referral traffic.
* Google and yahoo are the top referring search engines.
* Bulk of the traffic is from the American continents.
* Bulk of the users use Internet Explore or Firefox on Windows.

##### Note: Refer to Q10(a) to (f) for details of how do people come to the website.

## 2. Is the website generating interest, and does this interest yield actual sales?

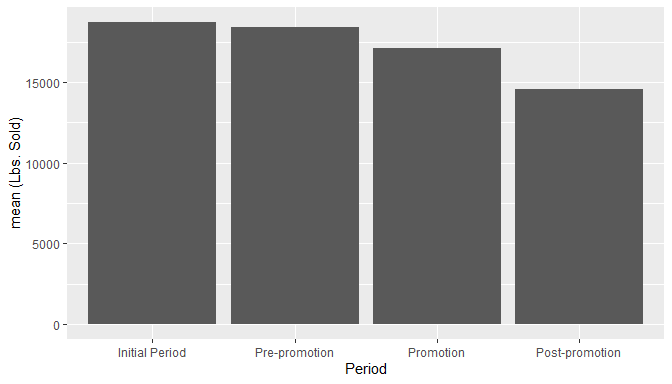
The website generated interest in the Promotion period which can be seen above by the increased number of Unique Visits over the promotion period. However, the interest did not yield actual sales as we see a decreasing Lbs. Sold over the four periods.

### Summary of Lbs. Sold by Period:

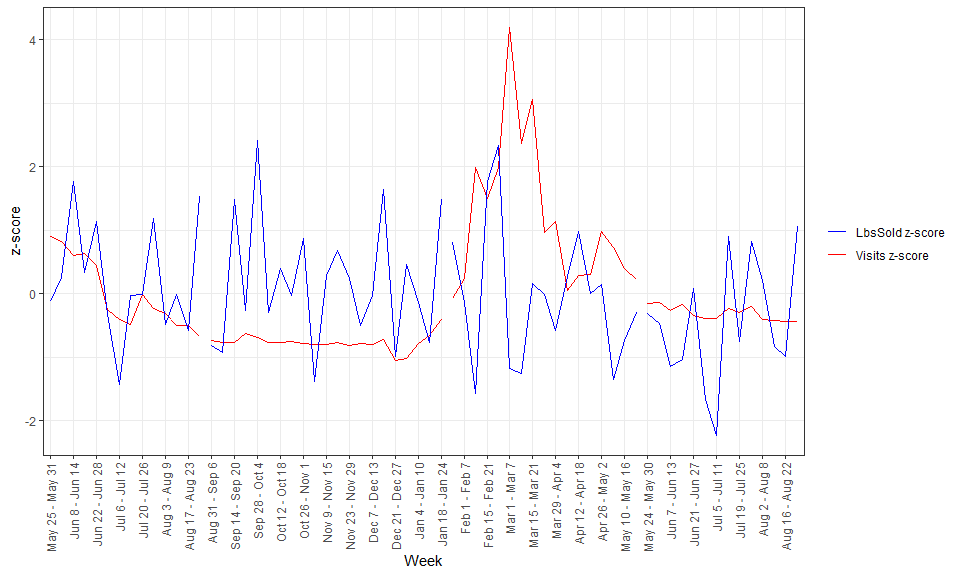
LBS. SOLD BY PERIOD

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Period | mean | median | sd | minimum | maximum |
| Initial Period | 18736.73 | 17269.70 | 5427.393 | 8633.059 | 28052.92 |
| Pre-promotion | 18440.77 | 17215.12 | 5965.630 | 8992.422 | 31968.98 |
| Promotion | 17112.92 | 17299.12 | 6519.071 | 7814.050 | 31496.26 |
| Post-promotion | 14577.79 | 13646.89 | 5941.556 | 3825.748 | 23761.61 |

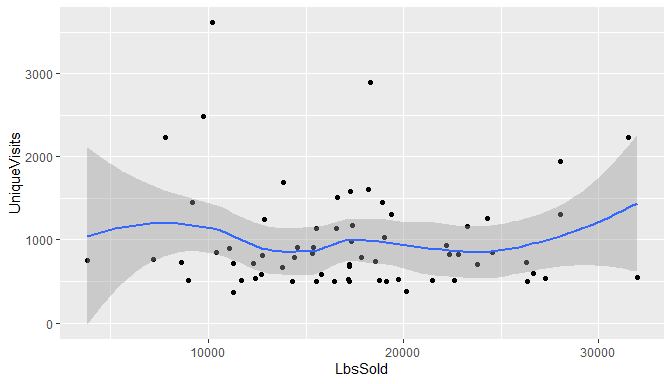
### Mean Lbs. Sold by period



### Comparison of Unique Visits and Lbs. Sold after Z-score standardization:



### Scatterplot of Lbs. Sold and Unique Vists



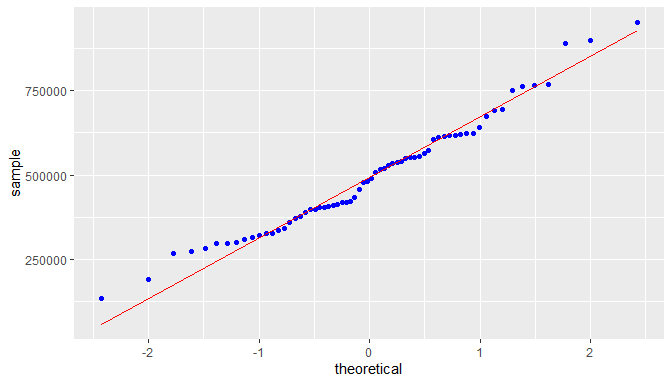
### Coefficient of correlation of Lbs. Sold and Unique Visits

## [1] -0.05622245

Small negative value shows that Lbs. Sold does not increase or decrease with Unique Visits.

## 3. Check whether the data variables (if applicable) follows the normality assumption.

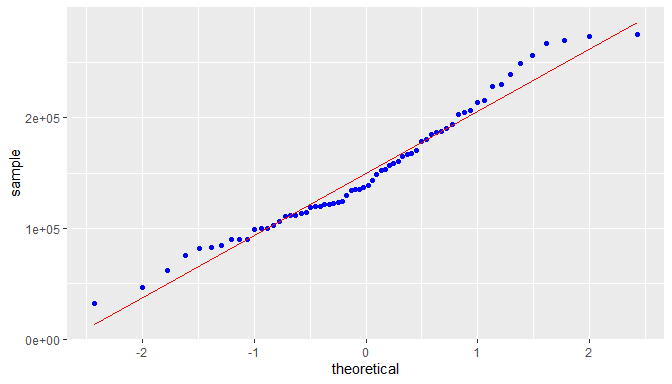
### Normality assumption for Revenue



Revenue is approximately normal but slightly right skewed:

## [1] 0.1850349

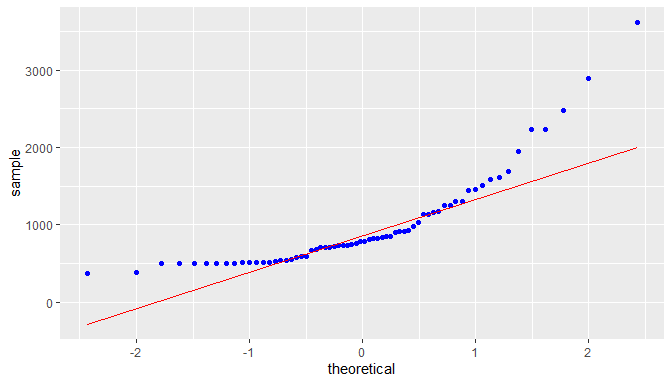
### Normality assumption for Profit



Profit is approximately normal but significantly right skewed. Skewness:

## [1] 0.6745705

### Normality assumption for UniqueVisits



UniqueVisits shows a high degree of right skewness:

## [1] 0.9622686

##### Note: Refer to Q8(f) for normality assumption check of Lbs. Sold and Q9. for normality assumption check of Dailty Visits

## Evaluation of the goals of the company website

### (a) drive new sales - NOT MET

The data analysis presented in PART B of this document clearly shows that revenue and quantity of material sold have been reducing over the period of analysis. Thus, we can say that the website has not been able to drive new sales even during the promotion period.

### (b) make product and contact information available - MAY BE MET

This goal can be considered as met if we assume that the contact information has been presented well on the website, .

### (c) give or add legitimacy to its brand. - MAY BE MET

Just the existence of an official company website does add legitimacy to the branch. However, a survey given to the right audience, customers, potential customers and general public, may be able to shed more light on this. The available data is not sufficient to answer whether this goal has been met.

## Recommendations from demographics:

There is significant Direct Traffic, which could be a result of the website link socialized in the promotion. Traffic from search engines could be increased with the help of SEO (Search engine optimization).

The paid listings on globalspec.com are generating more referrals than thomasnet.com, however, both of them together are not generating much referrals compared to google ads.

The traffic from sedoparking.com may need to be investigated more as it could indicate users mistyping the QA website address and landing on a parked domain with a similar name.

As expected, google is the top search engine source of visits. SEO should be targeted at google search engine.

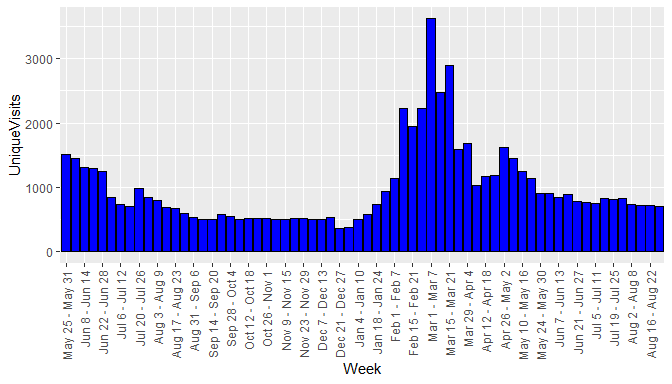
As the traffic is coming from various geographies, it may be a good idea to support multiple languages on the website (internationalization).

Website development / testing should focus on Internet Explorer and Firefox compatibility / support on Windows platform, followed by Mac and Linux as the majority of users are using this stack.

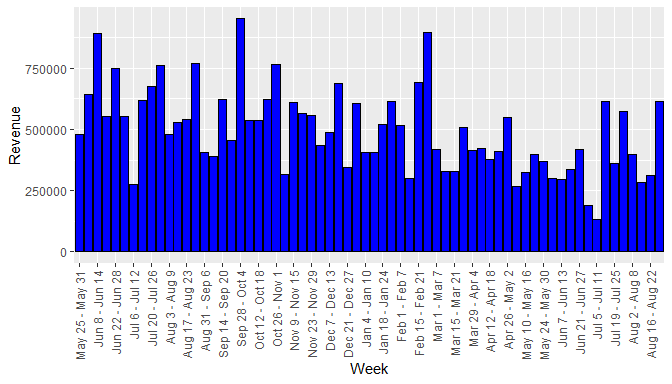
# PART 2 - DATA ANALYSIS

## 1. Column Charts

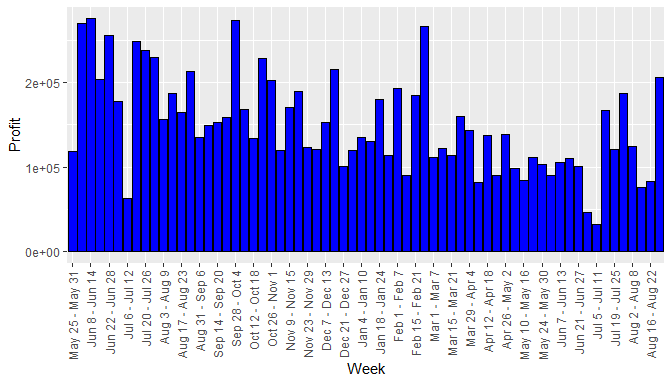
### a. Unique visits over time



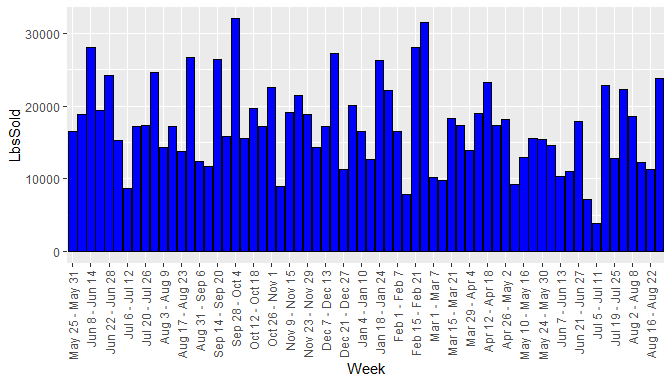
### b. Revenue over time,



### c. Profit over time



### d. Pounds sold over time.



## 2. Summary by period

### a. Summary table of Initial Period

VISIT AND FINANCIAL SUMMARY MEASURES - INITIAL PERIOD

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Visits | UniqueVisits | Revenue | Profit | LbsSold |
| mean | 1055.2143 | 975.9286 | 608250.1 | 200233.41 | 18736.729 |
| median | 899.0000 | 845.5000 | 586169.7 | 208913.00 | 17269.696 |
| std. dev. | 355.0333 | 319.5972 | 155930.4 | 60691.55 | 5427.393 |
| minimum | 626.0000 | 594.0000 | 274567.6 | 62580.40 | 8633.059 |
| maximum | 1632.0000 | 1509.0000 | 890076.7 | 275218.10 | 28052.924 |

### b. Summary table of Pre-promotion Period

VISIT AND FINANCIAL SUMMARY MEASURES - PRE PROMOTION PERIOD

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Visits | UniqueVisits | Revenue | Profit | LbsSold |
| mean | 562.95238 | 516.80952 | 534313.5 | 159932.03 | 18440.774 |
| median | 558.00000 | 510.00000 | 534541.7 | 152476.40 | 17215.118 |
| std. dev. | 80.87118 | 70.93914 | 150502.8 | 42682.68 | 5965.630 |
| minimum | 383.00000 | 366.00000 | 315647.1 | 100388.40 | 8992.422 |
| maximum | 795.00000 | 734.00000 | 951216.2 | 273174.70 | 31968.976 |

### c. Summary table of Promotion Period

VISIT AND FINANCIAL SUMMARY MEASURES - PROMOTION PERIOD

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Visits | UniqueVisits | Revenue | Profit | LbsSold |
| mean | 1814.3529 | 1738.8235 | 456398.8 | 131929.90 | 17112.923 |
| median | 1663.0000 | 1585.0000 | 413937.1 | 114328.40 | 17299.115 |
| std. dev. | 758.1017 | 743.0253 | 161741.1 | 47776.85 | 6519.071 |
| minimum | 1000.0000 | 930.0000 | 268159.5 | 81841.40 | 7814.050 |
| maximum | 3726.0000 | 3617.0000 | 897163.7 | 266476.70 | 31496.257 |

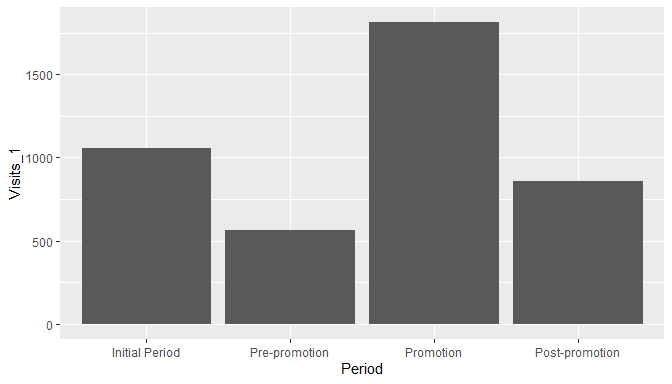
### d. Summary table of Post-promotion Period

VISIT AND FINANCIAL SUMMARY MEASURES - POST PROMOTION PERIOD

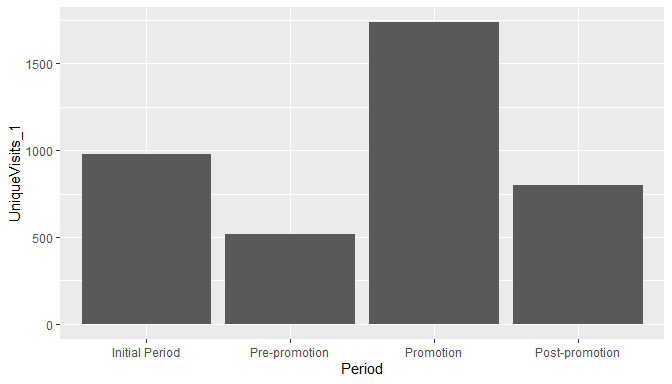
|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Visits | UniqueVisits | Revenue | Profit | LbsSold |
| mean | 856.57143 | 800.78571 | 371728.0 | 111045.82 | 14577.794 |
| median | 847.50000 | 800.00000 | 348397.1 | 104530.45 | 13646.895 |
| std. dev. | 70.88855 | 72.35615 | 145728.3 | 49065.27 | 5941.556 |
| minimum | 772.00000 | 709.00000 | 133966.9 | 32825.30 | 3825.748 |
| maximum | 963.00000 | 912.00000 | 615950.2 | 206441.20 | 23761.614 |

## 3. Column chart of the means over the four periods

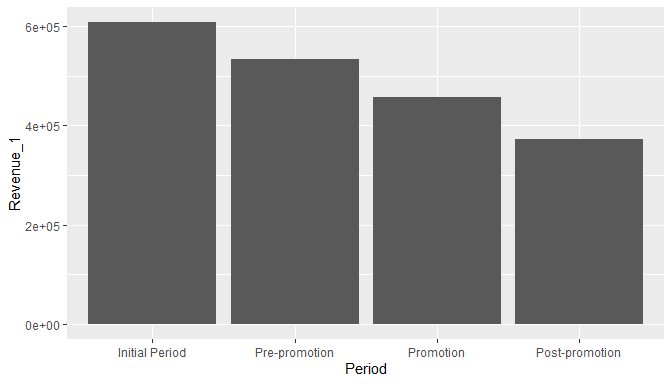
### a. Mean Visits by Period



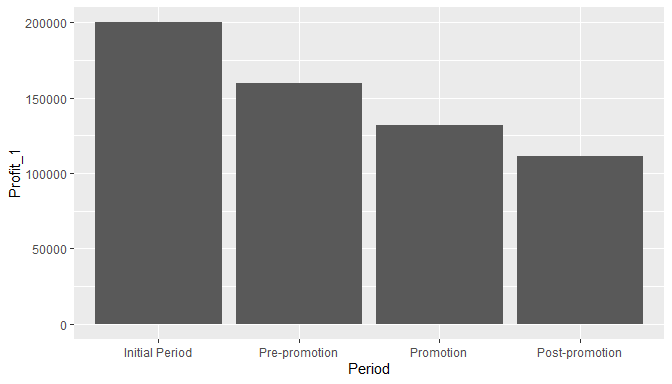
### b. Mean Unique Visits by Period



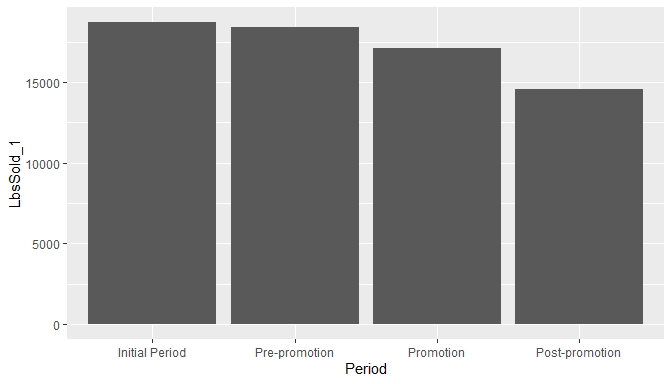
### c. Mean Revenue by Period



### d. Mean Profit by Period



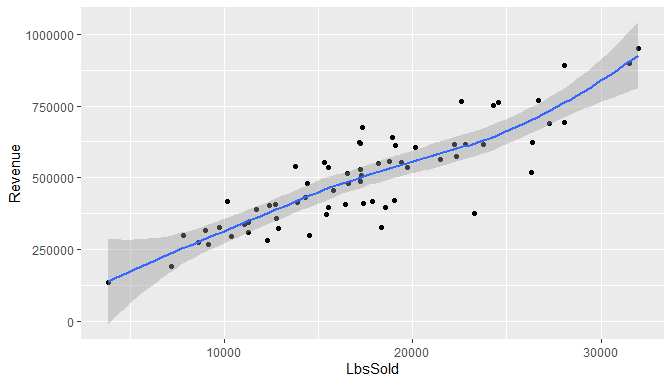
### e. Mean Lbs. Sold by Period



## 4. Summary of findings thus far

Average visits and average unique visits have shown a marked increase in promotion period. However, profit, revenue and lbs. sold show a steady declining trend over the four periods. This suggests that the website promotion did not have a positive effect on financials.

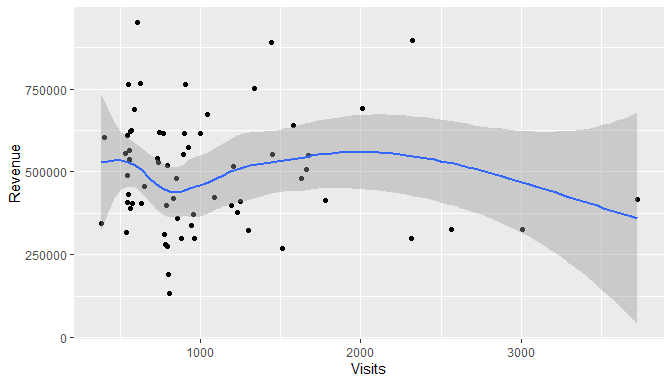
## 5. Scatter plot of Revenue and Lbs. Sold



### b. Coefficient of Correlation

## [1] 0.8689297

## 6. Scatter plot of Visits and Revenue



### b. Coefficient of Correlation

## [1] -0.05939183

## 7. Summary of results

The scatter plot of Revenue and Lbs. Sold shows a high degree of correlation.

The scatter plot of Revenue and Visits mirrors the earlier observations from the column chart - revenue does not increase or decrease with visits. Thus the website did not meet the goal of driving sales.

## 8. Modeling data

### a. Summary of Lbs. Sold data per week from January 3, 2005, through the week of July 19, 2010.

## [1] "Mean: 18681.5551724138"

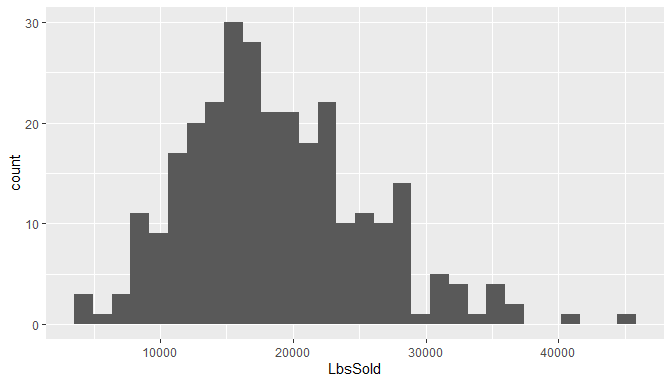
## [1] "Median: 17673"

## [1] "Std. Dev.: 6840.50793952086"

## [1] "Minimum: 3826"

## [1] "Maximum: 44740"

### b. Histogram of the pounds of material sold data.



### c. Histogram description

Histogram appears bell shaped with a few outliers on the right side. Also, mean is greater than median, so the data is right skewed.

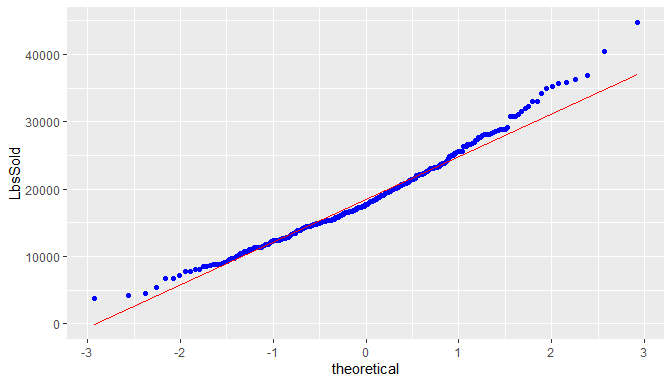
### d. Empirical rule

|  |  |  |  |
| --- | --- | --- | --- |
| Interval | TheoreticalPctOfData | TheoreticalNoOfObs | ActualNoOfObs |
| mean B1 1 std. dev. | 68 | 197.2 | 201 |
| mean B1 2 std. dev. | 95 | 275.5 | 276 |
| mean B1 3 std. dev. 99% | 99 | 287.1 | 288 |

### e. Refine empirical rule

|  |  |  |  |
| --- | --- | --- | --- |
| Interval | TheoreticalPctOfData | TheoreticalNoOfObs | ActualNoOfObs |
| mean + 1 std. dev. | 34.101 | 99 | 84 |
| mean - 1 std. dev. | 34.065 | 99 | 117 |
| 1 std. dev. to 2 std. dev. | 13.669 | 40 | 35 |
| -1 std. dev. to -2 std. dev. | 13.631 | 40 | 40 |
| 2 std. dev. to 3 std. dev. | 2.139 | 6 | 9 |
| -2 std. dev. to -3 std. dev. | 2.131 | 6 | 3 |

### f. Is pound of material sold bell shaped?



From the normal probability plot, we can see the LbsSold data approximates the bell shape but not exactly.

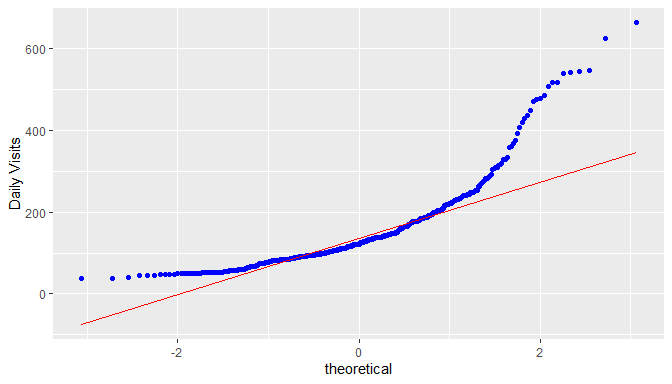
### g. Skweness and Kurtosis

## [1] "Skewness using formula 3\*(mean-median)/sd: 0.442315913378401"

## [1] "Skewness using library: 0.629032218484559"

## [1] "Kurtosis: 3.53337138531912"

## 9. Comparison of the distribution of the pounds sold data with daily visit data.



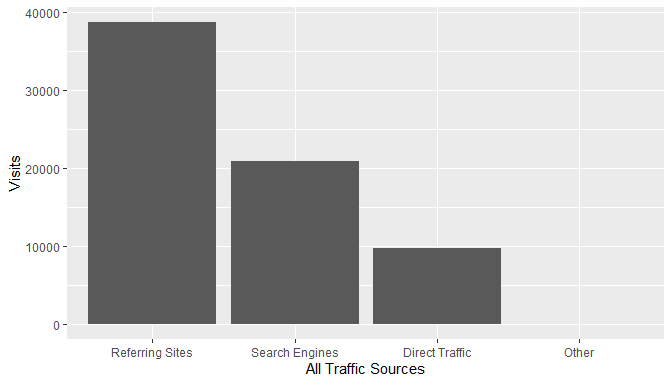
Daily visits data is less bell shaped compared to the Lbs sold data.

The empirical rule shows the actual number of observations matches the theoretical number of observations within 2nd and 3rd sd, but there is huge difference in the first sd - 314 (68%) theoretical vs 392 (85%) actual).

The mean is greater than median, so the data is right skewed. Skewness and Kurtosis are also high compared to Lbs Sold data.

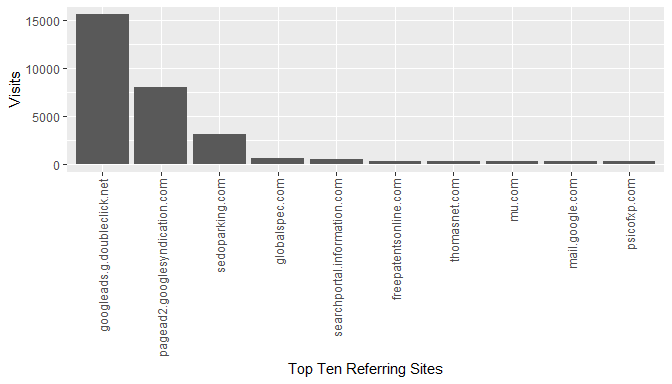
## 10. Demographics

## a. All Traffic Sources



There is significant Direct Traffic, which could be a result of the website link socialized in the promotion. Traffic from search engines could be increased with the help of SEO (Search engine optimization).

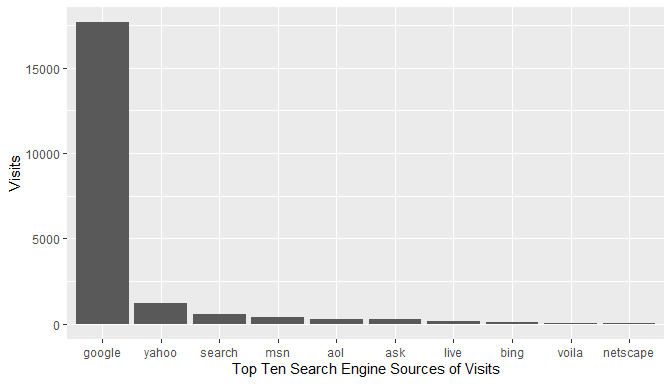
## b. Top Ten Referring Sites



The paid listings on globalspec are generating more referrals than thomasnet, however, both of them together are not generating much referrals compared to google ads.

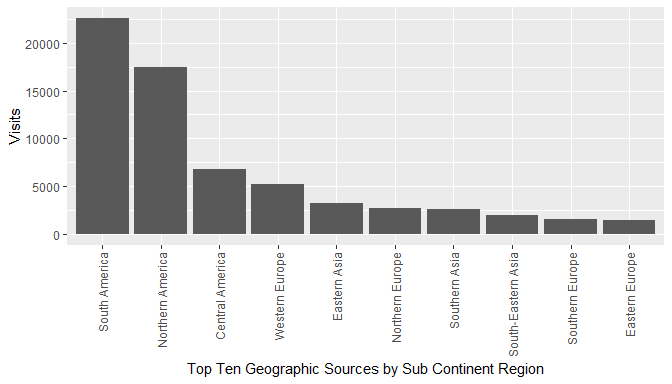
The traffic from sedoparking.com may need to be investigated more as it could indicate users mistyping the QA website address and landing on a parked domain with a similar name.

## c. Top Ten Search Engine Sources of Visits



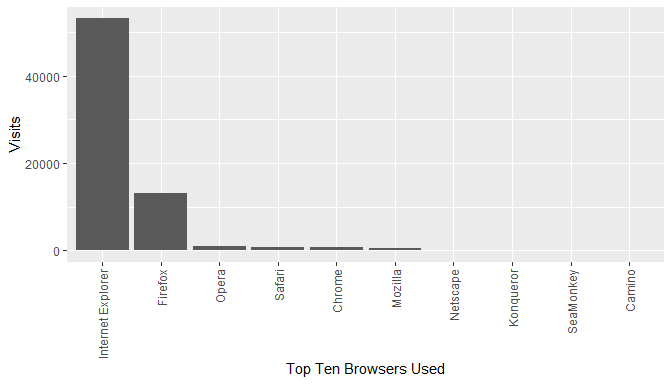
As expected, google is the top search engine source of visits. SEO should be targeted at google search engine.

## d. Top Ten Geographic Sources by Sub Continent Region



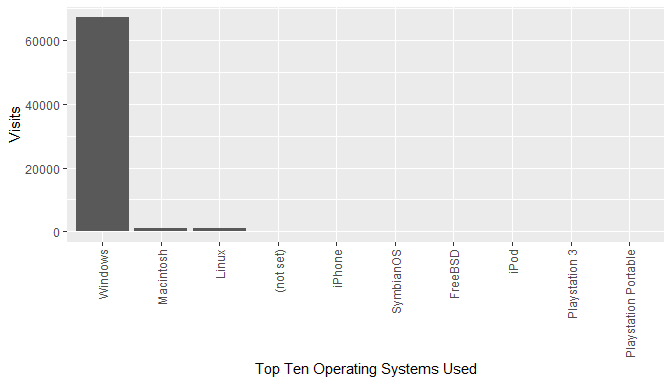
As the traffic is coming from various geographies, it may be a good idea to support multiple languages on the website (internationalization).

## e. Top Ten Browsers Used



Website development / testing should focus on Internet Explorer and Firefox compatibility / support.

## f. Top Ten Operating Systems Used



Website development / testing should focus on Windows platform, followed by Mac and Linux.