|  |
| --- |
|  |
| Project Report on Advanced Data Analytics |
| Rocket Fuel Inc. |
|  |
| **GOVIND KRISHNAN KA** |
| **8/14/2020** |

Table of Contents

[1.0 Executive Summary 2](#_Toc48332789)

[2.0 About Rocket Fuel Inc. 3](#_Toc48332790)

[3.0 Problem Statement 3](#_Toc48332791)

[4.0 Factors determining Control Group Size (4%) 4](#_Toc48332792)

[5.0 Methodology for conducting the ad-campaign 4](#_Toc48332793)

[6.0 Checking the randomness of user distribution in Test Group and Control Group 5](#_Toc48332794)

[7.0 Financials due to the new digital campaign 6](#_Toc48332795)

[8.0 Effectiveness of the advertising 7](#_Toc48332796)

[9.0 Recommendations based on Experiment conducted 12](#_Toc48332797)

# Executive Summary

Rocket Fuel Inc. is a leader in using data science for driving sales for companies operating in the consumer goods space. TaskaBella, an accessory company wants to increase its new handbags sales through the introduction of a new digital ad campaign. Rocket Fuel through its proprietary algorithms conducted a 6 month long experiment in ascertaining the optimum number of impressions required by an ad for a successful converted user. The experiment also focuses on the need of a control group who are shown a public service ad. The control group helps in understanding the lift achieved in converted users through the digital ad-campaign. The number of users in the test group i.e. users shown the actual digital ad and control group (users shown the PSA) have been allocated randomly.

The report highlights the economics of the project namely the cost, revenues, return on investment and opportunity cost for TaskaBella if any.

There is also an attempt to arrive at the optimum number of impressions that lead to successful conversion of a user. The time of the day and day of the week where prospective users are likely to be active are also highlighted in the later part of the report.

# About Rocket Fuel Inc.

Rocket Fuel Inc. was founded in March 2008 with a vision of transforming the digital advertising agency industry through big data and artificial intelligence. At a time when real time bidding (RTB) and ad exchanges were still nascent, Rocket Fuel quickly grew into the industry leader by bringing the precision and conclusiveness of science to the world of marketing to drive higher ROI. By 2016, Rocket Fuel had grown into a business with nearly half a billion dollars in annual revenue. Rocket Fuel’s full Programmatic Marketing Platform has been designed to leverage artificial intelligence (AI) to determine what marketing actions to take with a particular person in a particular moment of time. Using Machine Learning (ML) techniques the platform continuously updates the features of its underlying model to improve campaign effectiveness in real time.

# Problem Statement

TaskaBella, a handbags and related accessories firm was contemplating to reallocate a significant portion of its advertising spent on digital advertising. Before proceeding towards the major overhaul, TaskaBella approached Rocket Fuel to run a trial campaign for its newly released handbag model.

The ad-campaign planned by Rocket Fuel Inc. was aimed to target at least 0.5 million consumers online. The subsequent chapters will highlight the effectiveness of the ad-campaign undertaken by Rocket Fuel Inc.

**Input parameters considered while launching the Ad-campaign**

The table below indicates the various parameters considered for the ad-campaign:

|  |  |
| --- | --- |
| Parameters Considered | Values |
| Target number of consumers | 500,000 |
| Cost per thousand impressions | $9 |
| Average revenue from converted user | $40 |
| Size of control group\* | 4% |

*\*Control group is a group of randomly selected users who were shown ads related to public service announcements (PSA)*

# Factors determining Control Group Size (4%)

1. Conversion rates in display advertising are typically low compared to other experimental settings requiring a sizeable control group to detect lift
2. If advertising is effective, the lift might be small due to a very small size of control group considered
3. A large control group allows for a deeper analysis with potentially significant differences in different segments of the population

# Methodology for conducting the ad-campaign

The trial campaign ran from November 2015 to February 2016. The ads were first loaded into Rocket Fuel’s content delivery network by an operations associate, which enables the quick displaying of the ads anywhere in the world within milliseconds. When the user visits a publisher’s page (like CNN.com or NYtimes.com), the publisher partners with a supply-side vendor to put the impression opportunity out to auction. This all happens millions of times per second across the dozens of such exchanges that Rocket Fuel is integrated with. Rocket Fuel receives the bid request and needs to respond within 100 milliseconds in order for its bid to qualify.

Rocket Fuel receives a hashed cookie ID in the bid request, which allows it to anonymously identify the user in its system, along with many other parameters about the ad size and content. It then applies ML models to predict the probability that this user will take the required action or actions for each of its campaigns running at the time of the bid. For each campaign, users that have been assigned to the control group are not actually served a campaign ad, but a PSA.

The table indicates the results that were obtained based on the ad-campaign run by Rocket Fuel Inc.

|  |  |
| --- | --- |
| Parameters Considered | Values |
| No. of impressions served | 14,50,000 |
| No. of users identified | 590,000 |
| Users buying the new handbag | 15,000 |

The table below indicates the number of users in the control and test group. The values corresponding to converted represent the number of users who bought the hand bag and values corresponding to not-converted represent the number of users who did not buy the hand bag.

|  |  |  |
| --- | --- | --- |
| # | Test Group | Control Group |
| Not-Converted | 550,154 | 23,104 |
| Converted | 14,423 | 420 |
| Total | 564,577 | 23,524 |

The above table represents randomly selected users who were shown the digital ad for TaskaBella and PSA. The numbers of users engaged in the activity were 588,101. As stated previously, the size of control group was restricted to **4%** of the total group size

# Checking the randomness of user distribution in Test Group and Control Group

From the data obtained, it was evident that the conversion rate of users was dependent on the number of impressions. In order to validate the randomness of the distribution of the users in the two groups namely test and control, a **“t”** test was carried out. It was found that the mean number of impressions in test and control group were nearly equal.

Mean Impressions in control group: **24.76**

Mean Impressions in test group: **24.82**

The table below represents the percentage of converted users and non-converted users in test and control group:

|  |  |  |
| --- | --- | --- |
| # | Test Group | Control Group |
| Not-Converted | 97.44% | 98.21% |
| Converted | 2.56% | 1.79% |
| Total | 100.00% | 100.00% |

*The lift observed due to introduction of this ad-campaign is* ***0.77%***

**Hence, the ad-campaign introduced was effective**

# Financials due to the new digital campaign

**Additional revenues for TaskaBella**

|  |  |
| --- | --- |
| **#** | **Values** |
| Revenue from one converted user | $40 |
| Total number of users in test group | 564,577 |
| Lift observed | 0.77% |
| Additional revenues due to campaign | $173,709.1 |

**Cost of Ad-Campaign**

|  |  |
| --- | --- |
| **#** | **Values** |
| Total number of impressions | 14,597,182 |
| Cost per 1000 impressions | $9 |
| Total cost of Ad-Campaign | $131,374.6 |

**Return on Investment Calculation**

|  |  |
| --- | --- |
| **#** | **Values** |
| Additional revenues due to campaign | $173,709.1 |
| Total cost of Ad-Campaign | $131,374.6 |
| Return on Investment | 32.22% |

**Opportunity Cost due to control group**

|  |  |
| --- | --- |
| **#** | **Values** |
| Revenue from converted user | $40 |
| Total number of users in control group | 23,524 |
| Lost lift | 0.77% |
| Opportunity cost | $7,235.98 |

# Effectiveness of the advertising

The table below indicates the distribution of the total number of impressions:

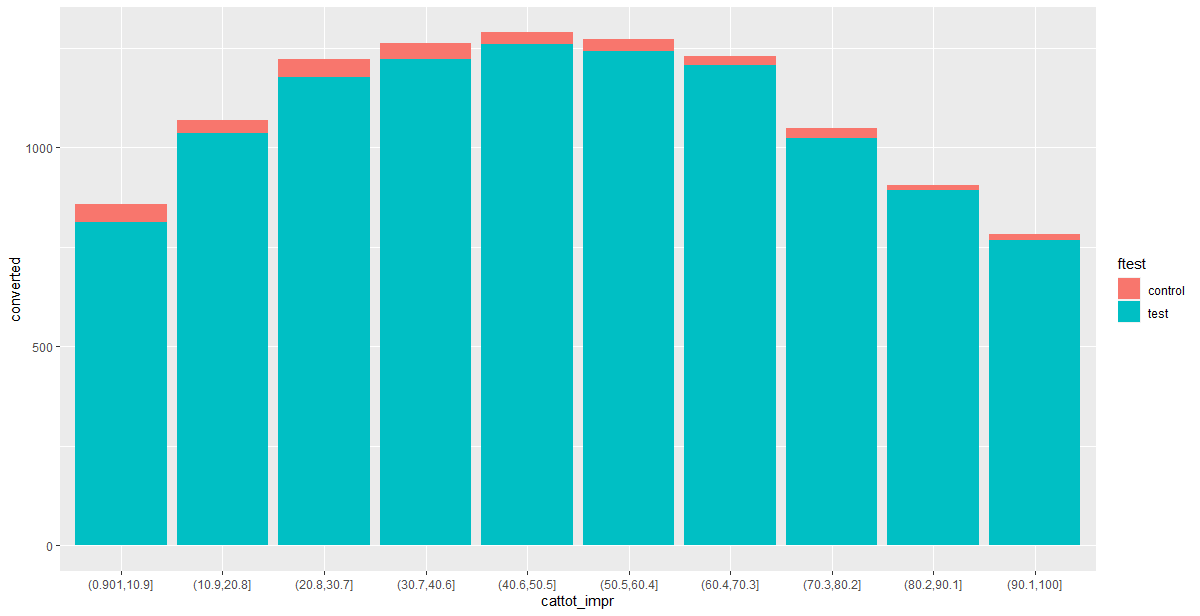
|  |  |
| --- | --- |
| # | Values |
| Minimum | 1.00 |
| 1st Quartile | 4.00 |
| Median | 13.00 |
| Mean | 24.82 |
| 3rd Quartile | 27.00 |
| Maximum | 2065.00 |

The above table clearly indicates that the total numbers of impressions amongst the users were highly skewed. Hence, the users were divided based on the number of impressions. The categories based on the number of impressions are:

1. Users with impressions less than or equal to 100
2. Users with impressions greater than 100 but less or equal to 200
3. Users with impressions greater than 200

For a more in-depth analysis, the group of users having impressions less than 100 was further divided into 10 equal groups. The groups were people having impressions between 0-10, 11-20 and so on…

Based on this division, the number of converted users was found out. The graph depicted below clearly indicates that the number of converted users increased from 0 to 50 impressions and then slowly tapered off. Maximum converted users with impressions between (40-50)



The table below indicates the percentage of converted users in test and control groups having less than or equal to **100** impressions.

|  |  |  |
| --- | --- | --- |
| # | Control | Test |
| Not-converted | 98.67% | 98.04% |
| Converted | 1.33% | 1.96% |

The total number of users having impressions between 0 and 100 are **565,037.** The lift observed was **0.63%.** Additional revenue earned by TaskaBella was **$136,715.8**

The table below indicates the percentage of converted users in test and control groups having greater than **100** impressions but less than 200 impressions.

|  |  |  |
| --- | --- | --- |
| # | Control | Test |
| Not-converted | 89.36% | 82.32% |
| Converted | 10.64% | 17.68% |

The total number of users having impressions between 100 and 200 are **17,112.** The lift observed was **7.04%.** Additional revenue earned by TaskaBella was **$46,069.76**

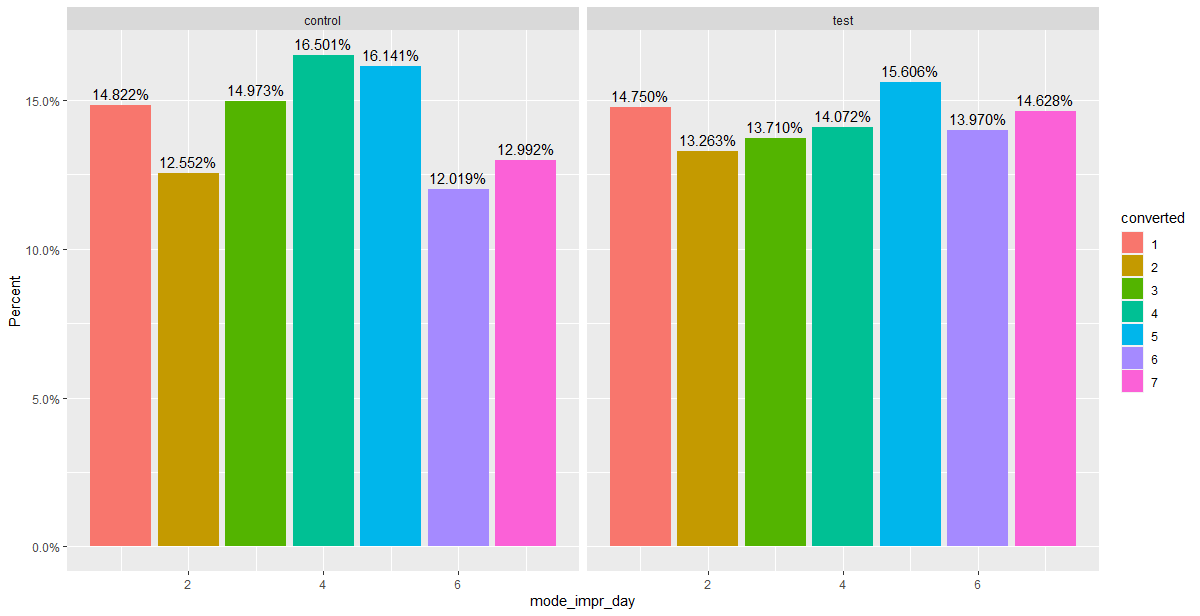
The table below indicates the percentage of converted users in test and control groups having greater than **200 i**mpressions:

|  |  |  |
| --- | --- | --- |
| # | Control | Test |
| Not-converted | 84.5% | 84.4% |
| Converted | 15.5% | 15.6% |

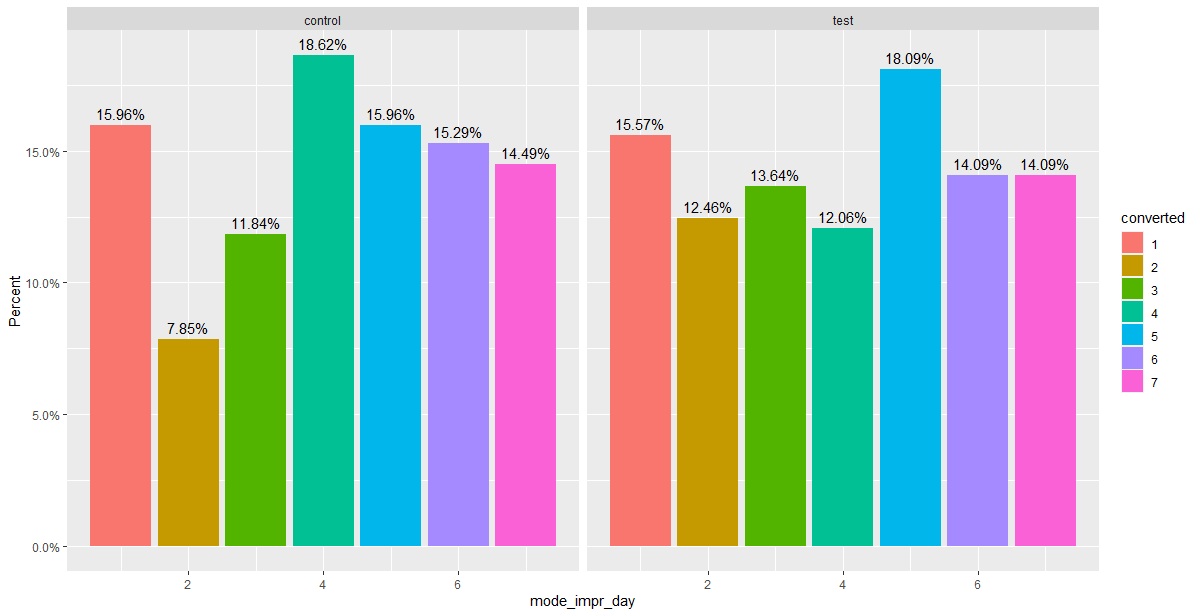
The total number of users having impressions between 100 and 200 are **5,952.** The lift observed was **0%. Negligible revenue** was obtained from users having **greater than 200 impressions**

*Based on the experiment, it was observed that chances of conversion were higher for users having impressions between* ***100 and 200***

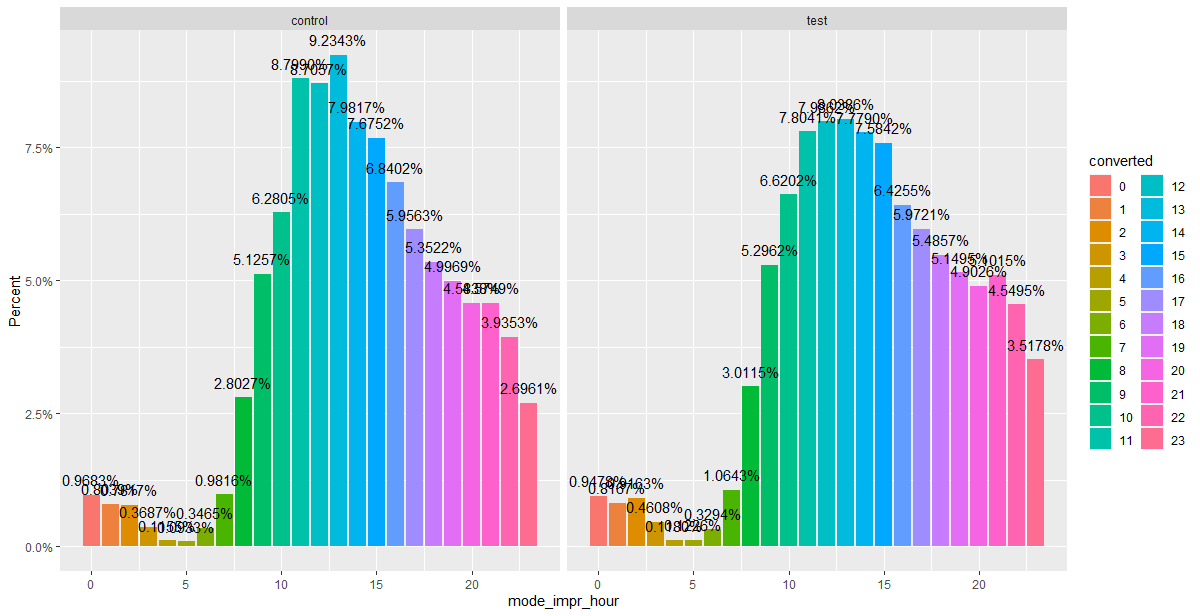
**Consumer response to ads based on Days of the Week and Hours of the Day**



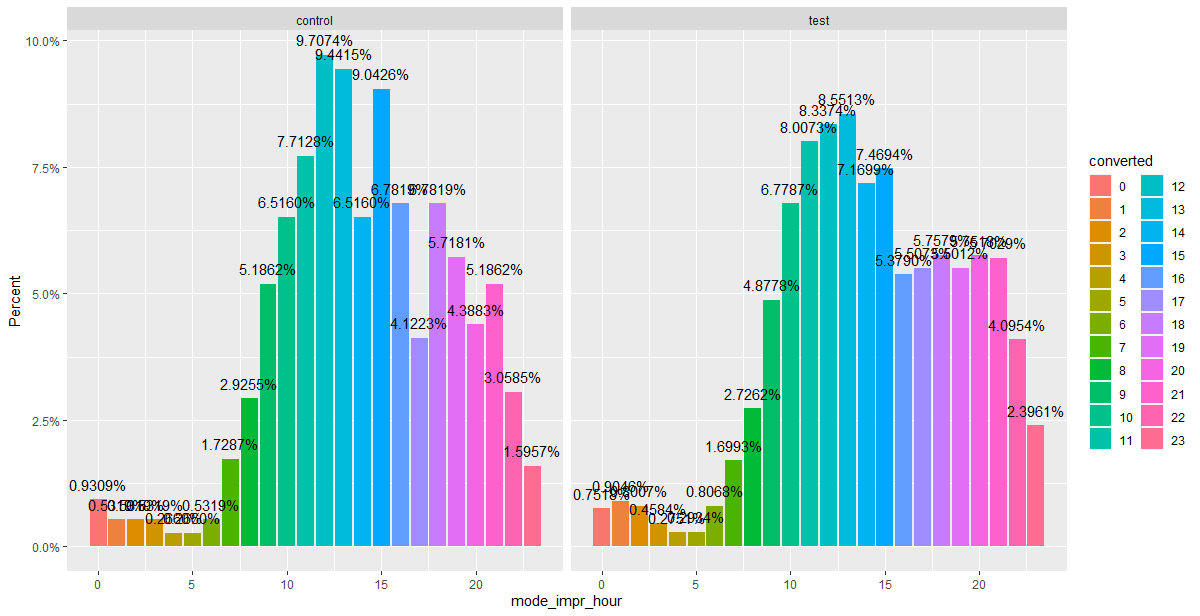
Based on the data collected, the 5th day of the week (**Friday**) was most effective for test group users (with less than or equal to 100 impressions). Conversion percentage was **15.6%. Tuesday** was the least productive day with conversion percentage of **13.3%**



Based on the data collected, the 5th day of the week (**Friday**) was most effective for test group users (with greater than100 impressions but less than 200 impressions). Conversion percentage was **18.1%. Thursday** was the least productive day with conversion percentage of **12.1%**



Based on the data collected, the 13th hour (**1 PM**) seems to be the the hour of the day where user conversion is maximum (almost 8%) for users with impressions lesser than equal to 100. The user conversion is extremely during wee hours



Based on the data collected, the 13th hour (**1 PM**) seems to be the the hour of the day where user conversion is maximum (almost **8.55%)** for users having impressions between 100 and 200. The user conversion is extremely during wee hours

# Recommendations based on Experiment conducted

1. The experiment clearly reiterates the fact that the user conversion is a function of number of impressions. However, repeated targeting of a user through digital ads might not necessarily lead to conversion. The optimal number of impressions for a prospective converted user shall be limited between 100 and 200
2. It is important to have a control group, of sufficient size to detect any significant lift due to the digital ad. The experiment clearly shows a lift of 0.77% when the control group size was limited to 4% of total users
3. Hour of the day and day of the week also play important rules. Hence, ads should be regulated as per the traffic observed on the site