RocketFuel

Manish Grewal

17/07/2020

## 1. Was the advertising campaign effective? Did additional consumers convert as a result of the ad campaign?

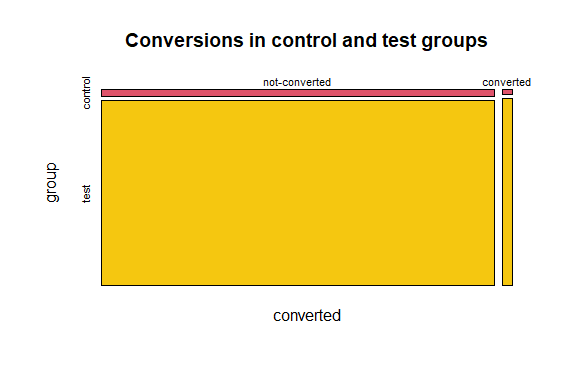
#### Number of users:

|  |  |  |
| --- | --- | --- |
|  | control | test |
| not-converted | 23104 | 550154 |
| converted | 420 | 14423 |

#### Proportion of users as percent:

|  |  |  |
| --- | --- | --- |
|  | control | test |
| not-converted | 98.214589 | 97.445344 |
| converted | 1.785411 | 2.554656 |

#### Mosaicplot



The conversion rate of exposed group 2.555% is greater than the conversion rate for control group 1.785%.

1.785% of users in the test group would have converted even without the ad impressions.

2.555% - 1.785% = 0.769% additional users in the test group have converted.

However, to be able to say that additional users converted as a result of the ad campaign, we need to check if the gain in conversion rate was statistically significant.

### t-test to check bias:

t.test(dat$tot\_impr ~dat$ftest)

##   
## Welch Two Sample t-test  
##   
## data: dat$tot\_impr by dat$ftest  
## t = -0.218, df = 25608, p-value = 0.8274  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -0.6217286 0.4972735  
## sample estimates:  
## mean in group control mean in group test   
## 24.76114 24.82337

As the mean of total impressions for the control and test groups are almost the same, and the t-value is reported as insignificant, we do not detect a bias in the total impressions for the control group and test group.

Yes, additional consumers converted as a result of the ad campaign.

## 2. Was the campaign profitable?

### a. How much more money did TaskaBella make by running the campaign (excluding advertising costs)?

Number of users in exposed group = 564577  
Profit per conversion = $40

**Additional profit**  
= lift \* (Number of users in exposed group) \* (Profit per conversion) / 100  
= lift \* num\_exposed \* 40 / 100  
= 0.7692453 \* 564577 \* 40 / 100  
= 173719.2858357

**Additional profit** = $173719.29

## b. What was the cost of the campaign?

Cost per 1000 (CPM) = 9

**Cost of campaign**  
= tot\_impr \* CPM / 1000  
= 14597182 \* 9 / 1000  
= 131374.638

**Cost of campaign** = $131374.64

## c. Calculate the ROI of the campaign. Was the campaign profitable?

**ROI**  
= (addl\_profit - cost\_campaign) / cost\_campaign \* 100  
= 173719.2858357 - 131374.638 / 131374.638 \* 100  
= 32.2319806

**ROI** = 32.23%

Yes, The campaign was profitable as ROI was good.

## d. What was the opportunity cost of including a control group; how much more could have TaskaBella made with a smaller control group or not having a control group at all?

**Opportunity Cost**  
= lift \* (number of users in control group) \* (Profit per conversion) / 100  
= lift \* num\_control \* 40 / 100  
= 0.7692453 \* 23524 \* 40 / 100  
= $7238.2907557

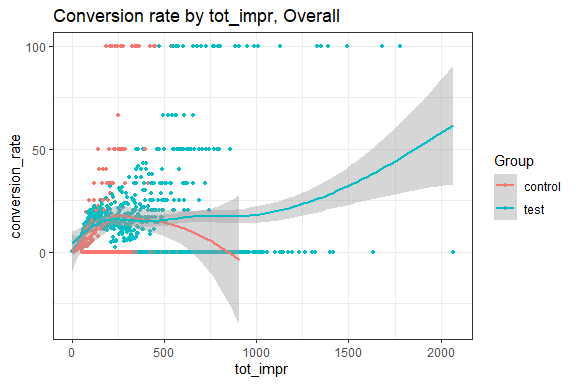
**Opportunity Cost** = $7238.29

If the control group had been shown ads for the product instead of PSA, an additional profit of $7238.2907557 could have been obtained.

## 3. How did the number of impressions seen by each user influence the effectiveness of advertising?

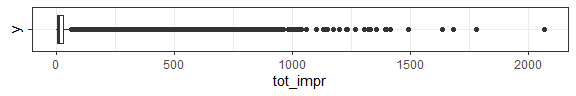
### a.1 Create a chart of conversion rates as a function of the number of ads displayed to users.

#### Conversion rates by total\_impr - Overall



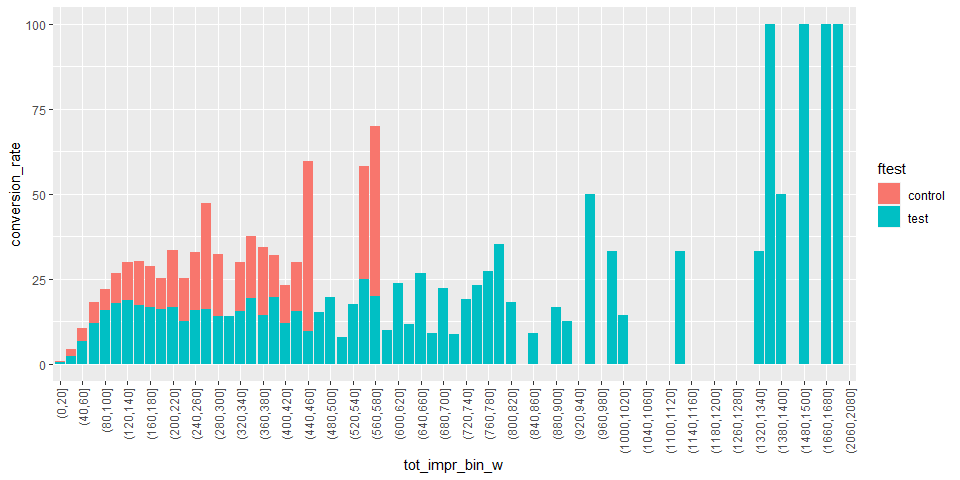
The scatter plot does not show a clear pattern. The number of points decrease as the total impressions increase, suggesting a presence of outliers. We bin the data using equal width and equal frequency binning to see if a patter emerges.

#### Boxplot total impressions



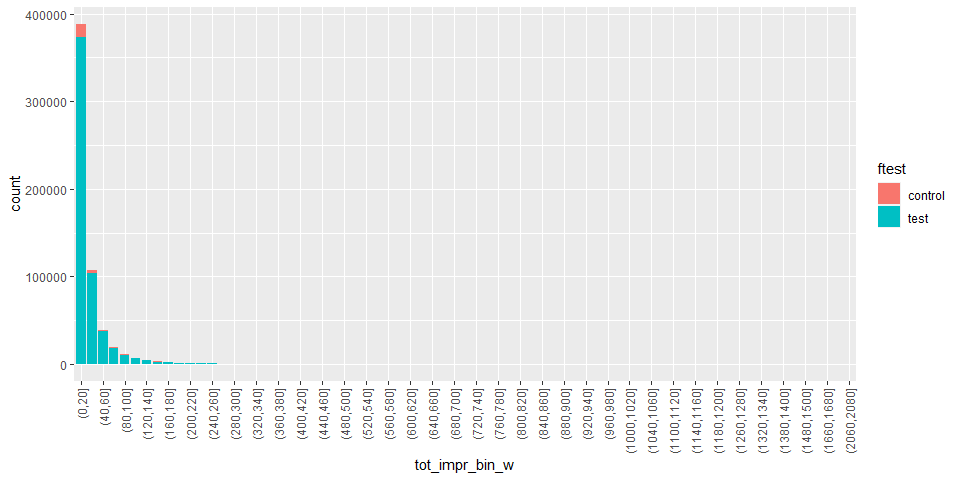
It appears that the values of total impressions above 60 are outliers.

#### Conversion rates by total\_impr - Overall - Equal width binning



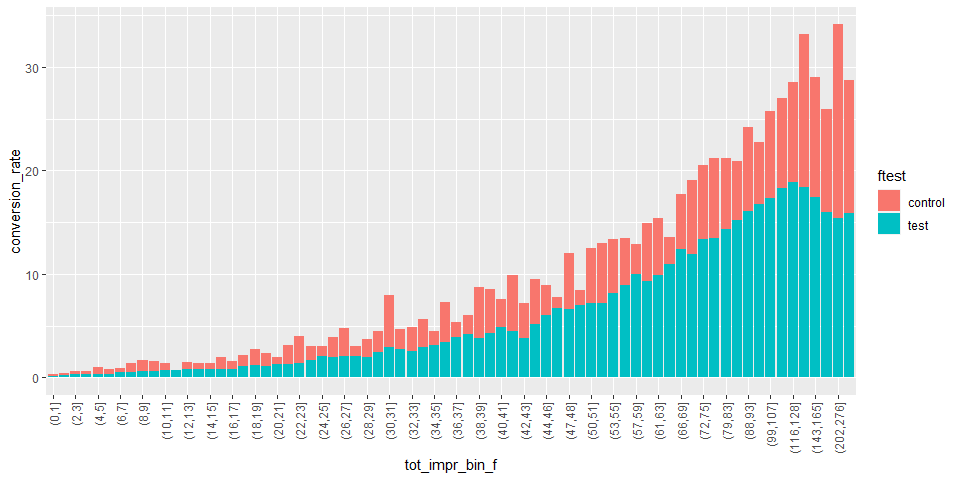
We observe a steady increase in conversion rate till around (120, 140] total impressions. Further increase in total impressions slightly reduces the conversion rate till (320,340] total impressions. Further increase in total impressions leads to an erratic conversion rate which sometimes falls to 0 for certain values of total impressions or increases dramatically.

#### Number of users by total impressions - Overall - Equal width binning



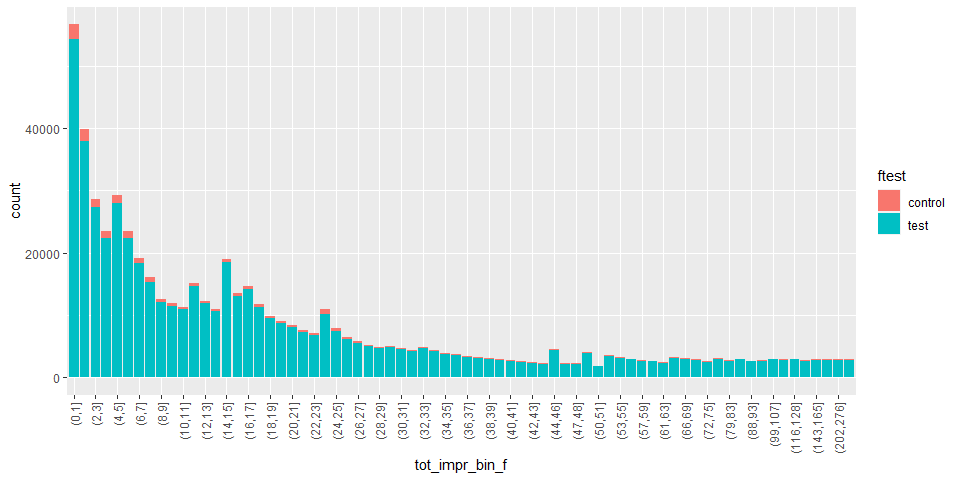
We see more than 60% users in the first bin (0,20]. We need to keep in mind the number of users under each bin is sharply decreasing with increasing total impressions and falls to insignificant numbers above 300 total impressions. Hence, equal width binning does not give a good representaion of the data.

#### Conversion rates by total\_impr - Overall - Equal frequency binning



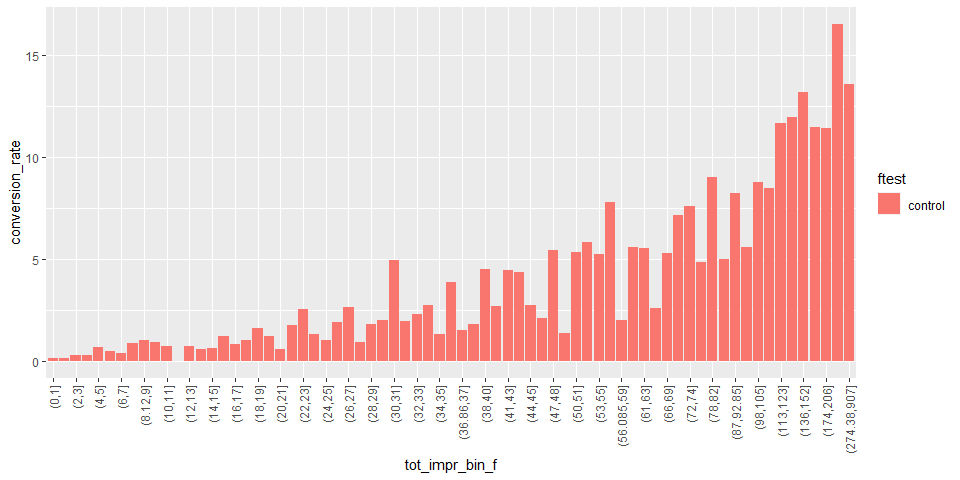
We can see the conversion rate steadily increases and reaches a peak at (116,128] total impressions. Further increase in total impressions leads to a declining conversion rate.

#### Number of users by total impressions - Overall - Equal frequency binning



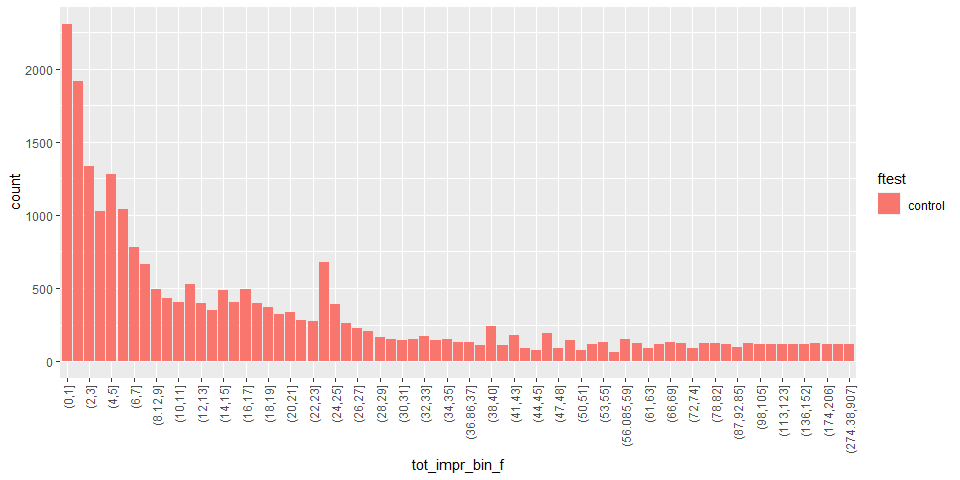
### 3.a.2 Plot conversion rates for those who were in the control group and for those who were exposed to the ad.

#### Conversion rates by total\_impr - Control Group - Equal frequency binning

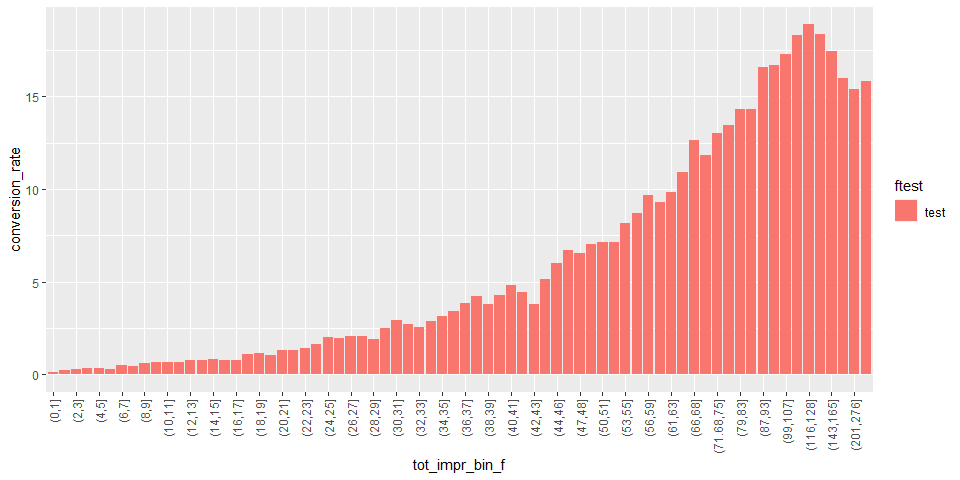


We can see the conversion rate steadily increases and reaches a peak at (116,128] total impressions. Further increase in total impressions leads to a declining conversion rate.

#### Number of users by total impressions - Control Group - Equal frequency binning

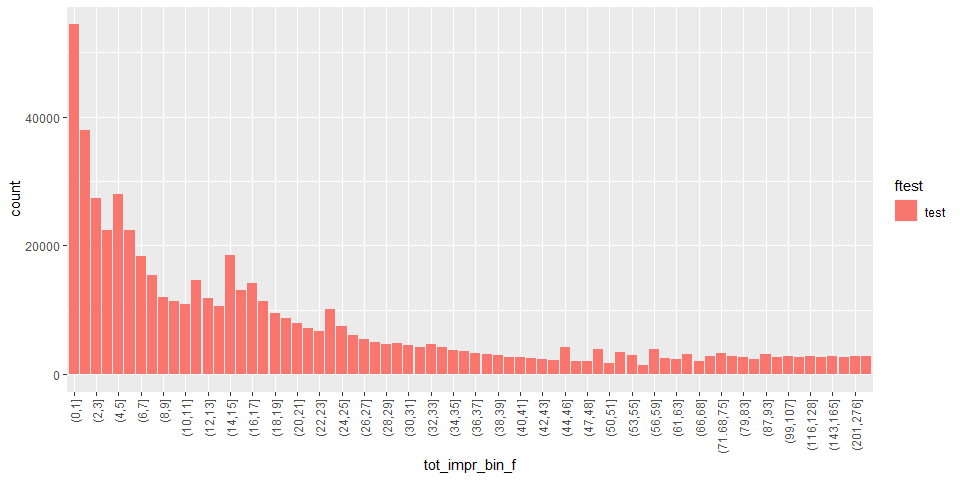


#### Conversion rates by total\_impr - Test Group - Equal frequency binning



We can see the conversion rate steadily increases and reaches a peak at (116,128] total impressions. Further increase in total impressions leads to a declining conversion rate.

#### Number of users by total impressions - Control Group - Equal frequency binning



### 3.a.3 Group together number of impressions as necessary to obtain a meaningful plot.

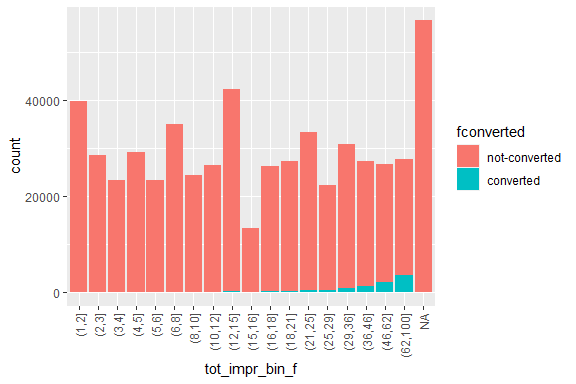
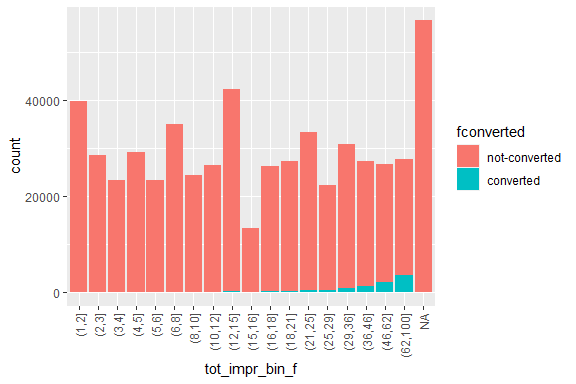
#### Boxplot

From looking at the quantiles, we see that 99.5% of that records have tot\_impr less than or equal to 276. So we will split the dataset into four parts for ease of analysis.

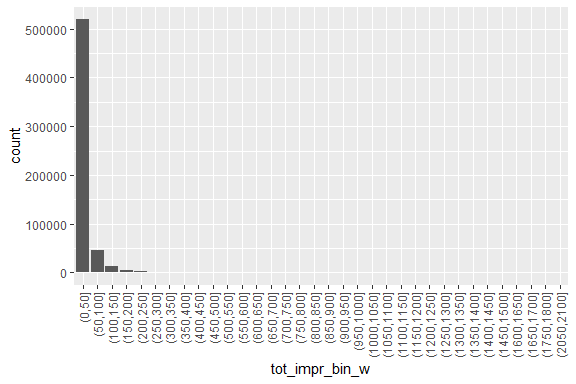
## 5% 10% 15% 20% 25% 30% 35% 40% 45% 50% 55% 60% 65% 70% 75% 80%   
## 1 1 2 3 4 5 6 8 10 12 15 16 18 21 25 29   
## 85% 90% 95% 100%   
## 36 46 62 100

## (1,2] (2,3] (3,4] (4,5] (5,6] (6,8] (8,10] (10,12]   
## 39827 28661 23426 29303 23409 35132 24411 26477   
## (12,15] (15,16] (16,18] (18,21] (21,25] (25,29] (29,36] (36,46]   
## 42250 13467 26360 27313 33380 22350 30833 27341   
## (46,62] (62,100] NA's   
## 26704 27787 56606

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 1.00 4.00 12.00 18.33 25.00 100.00

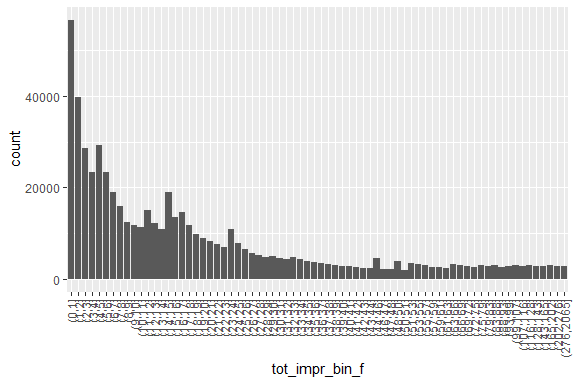


#### Equal width binning



#### Equal frequency binning

## 0% 20% 40% 60% 80% 100%   
## 1 3 8 17 33 2065



## b. What can you infer from the charts? In what region is advertising most effective?

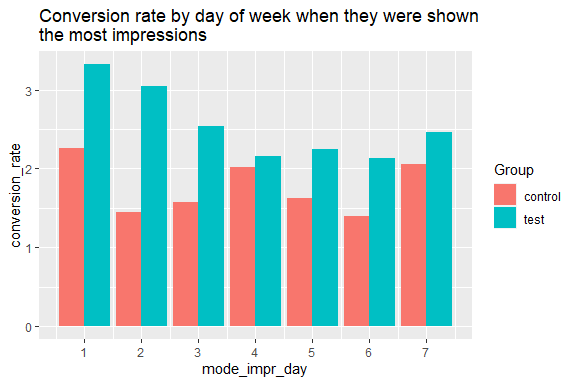
The conversion rate increases as the number of impressions is increased. However, as the number of impressions is increased beyond a limit (appears to be 250), we see a flattening of the smoothing curve and also a concentration of non conversions. Need to zoom in to the chart to get a better picture.

## c. What do the above figures imply for the design of the next campaign assuming that consumer response would be similar?

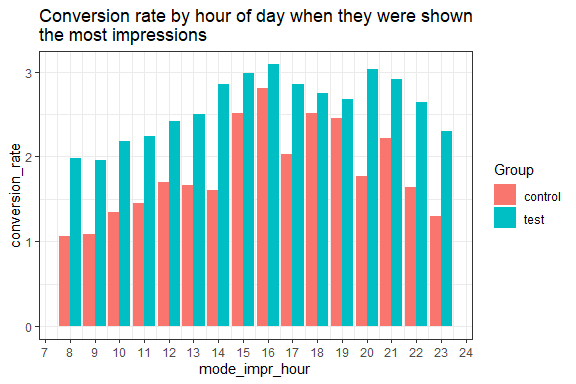
Total impressions can be limited to up to 500 to ensure cost savings. At higher number of impression, conversion rate can dip unexpectedly.

# 4. How does consumer response to advertising vary on different days of the week and at different times of the day?

## a. Create a chart with the conversion rates for the control group and the exposed group as a function of the day of week when they were shown the most impressions.



## b. Create the same chart for hours within a day (excluding the period between midnight and 8 a.m.).



## c. What days/hours is advertising most/least effective?

#### From the graph, we can read the highest and lowest values for the exposed group. We can also cross check by arranging summ\_day and summ\_hour variables by conversion rate

### Most effective day

|  |  |  |
| --- | --- | --- |
| mode\_impr\_day | ftest | conversion\_rate |
| 1 | test | 3.32412 |

#### => Most effective on Mondays.

### Least effective day

|  |  |  |
| --- | --- | --- |
| mode\_impr\_day | ftest | conversion\_rate |
| 6 | test | 2.130657 |

#### => Least effective on Saturdays

### Most effective hour

|  |  |  |
| --- | --- | --- |
| mode\_impr\_hour | ftest | conversion\_rate |
| 16 | test | 3.089286 |

#### => Most effective between 4 to 5 pm

### Least effective hour

|  |  |  |
| --- | --- | --- |
| mode\_impr\_hour | ftest | conversion\_rate |
| 8 | control | 1.062215 |

#### => Least effective between 9 to 10 am