
Analysis of the Success of a Promotional Offer

Arvind Ramkumar

Email – aramkum4@asu.edu | arvindramkumar1996@gmail.com

Phone – 480(859)-2393

Table of Contents

| | | |
|----------------------------|--------------|-----------|
| Problem Statement | | 3 |
| Assumption | | 3 |
| Data Preprocessing | | 3 |
| Dashboard | | 4 |
| Data Analysis | | 6 |
| Findings and Result | | 16 |
| Future Work | | 16 |

Problem Statement

- This is a detailed analysis of whether or not this Promotional offer was actually profitable for the company, because sometimes promotional offer has the tendency to show a profitability when it is actually not.
- Also, suggestions are given on how to improve the benefits of the Promotional offer

Assumption

- As the Move Out Date is missing for most of the data, we assume that the date when the data was collected (06/11/2014) to be the move out dates
- Here, we have compared between With Promotion and Without Promotion to understand the success of the Promotional Offer

Data Preprocessing

Basics

1. Import Libraries
2. Import Data Set
3. Explaining the Variables used in Analysis

Quality Check

1. The Data was checked for Number of Null and Not Null Values, column wise
2. The Data Frame's Shape, Dtypes are checked to have a better understanding of the Model
3. Check the type of Unique Values (In Categorical Variable) Present in the Data

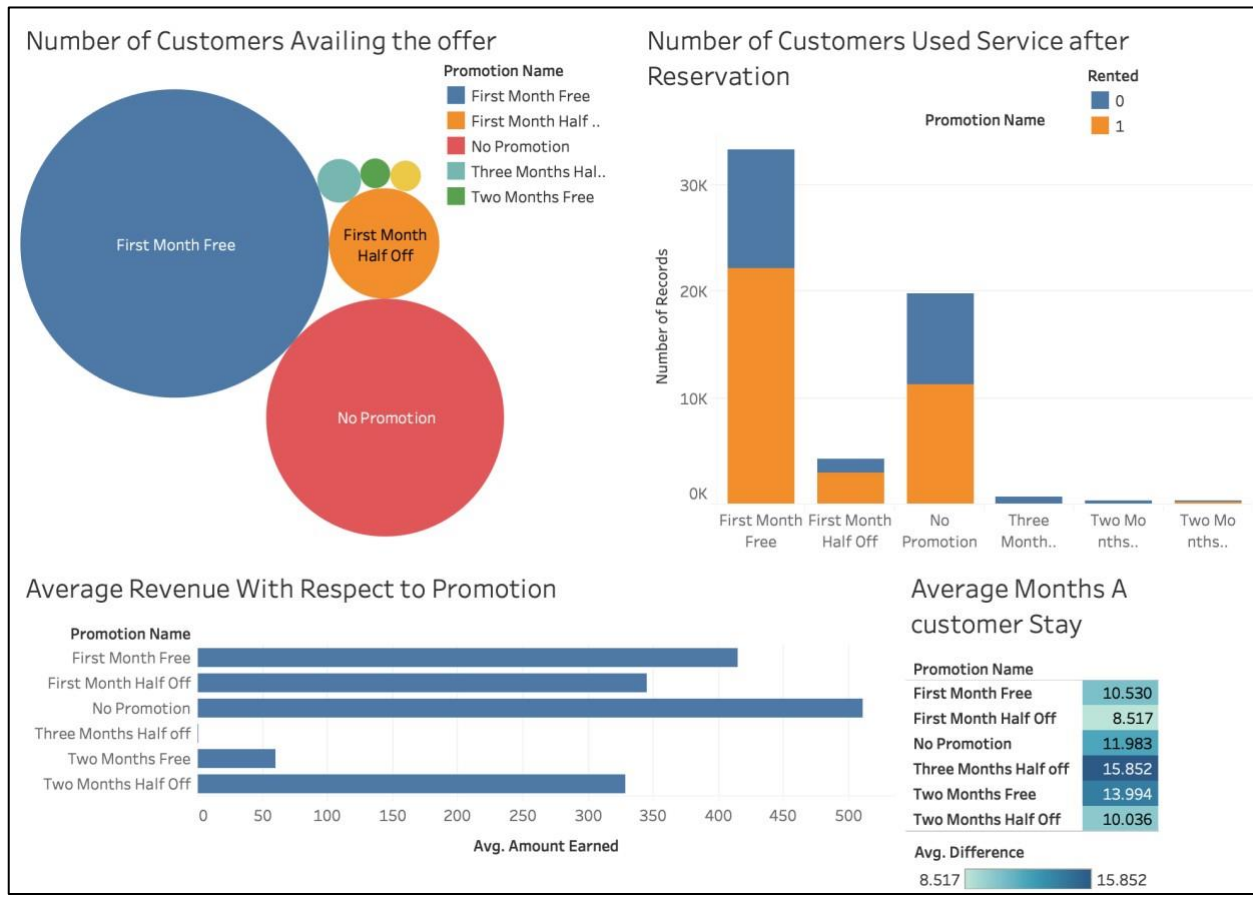
Data Munging

1. Remove all the Spaces in the Column Names
2. Convert the Rent from Boolean to 1 or 0 - 1 for True, 0 for False
3. Filling in all Null Values in Promotion with a new category - No Promotion
4. Remove \$ from RentedRate and StreetRate and Convert it to Float - Also remove % From Concession
5. Rename the column names by removing ?
6. Convert the Reserve Date and Move In Date from Object to Date Format
7. Assume all the Move Out Date which is not given as (6/11/2014)
8. Obtaining the length of stay of each customer
9. Obtaining the Revenue of the Company

Dashboard

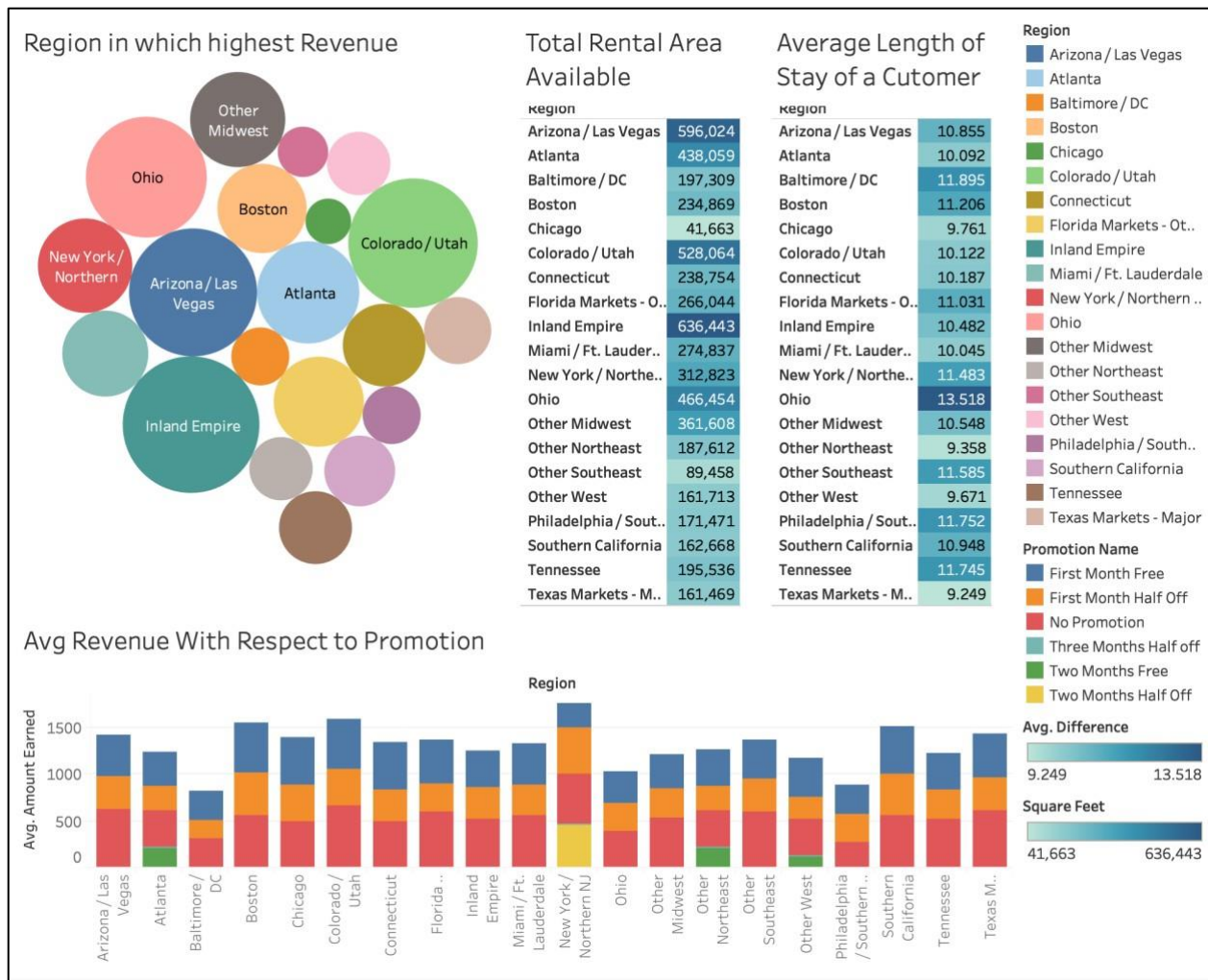
Dashboards are a great way to visually tell a story from a data, and here we have attempted to understand the data using two different dashboards created using Tableau.

1. Visualization how Promotion affects the revenue and number of customers
2. Visualization on the effect of the Region Type on the Revenue and Number of Customers



Visualization of How Promotion affects the revenue and the number of customers

1. Number of Customers Availing the offer – We see that First Month Free were offered more to the customers to an extent Customers without promotion is less than customers with promotion
2. Number of Customers Used Service After Reservation – The Grouped bar chart will help us to visualize the Number of Customers who did use the service and who did not after the Reservation was made grouped on the type of Promotion Basis
3. Average Revenue with Respect to Promotion – We can see that Average Revenue by a customer without promotion is higher than a customer with promotion
4. Average Months A customer stays – This gives us the average period of stay of a customer grouped on the type of promotion basis. Three Months Half off has high average period of stay because of the minimal customers availed that offer.



Visualization on the effect of the Region Type on the Revenue and Number of Customers

1. Region in which highest Revenue – We can see that Arizona/ Las Vegas, NYC, Ohio, Colorado, Inland Empire has generated a higher revenue than any other place. So, we can try to understand on why the revenue is high and make an effort to use that strategy to increase the revenue in other Regions as well
2. Total Rental Area Available – Inland Empire has the highest area available to rent. We see that NYC has lower area but was among the top Revenue Generators. So, we might want to consider that as well.
3. Average Length of Stay of a customer – Ohio is the region that the customer stays the longest.
4. Average Revenue with Respect to promotion – Average Revenue Earned in every region grouped by the type of promotion

Strange – We see that NYC has highest Revenue, but it has less rental area available, and the average stay of a customer is pretty much same as others. But It has one of the highest revenue generators. And also, NYC has offered the highest Two Months Half off. So, we might want to consider and study how two months half off has affected the revenue in NYC. It seems at this point that Two months half off seems to me more reliable than any other promotion

Data Analysis

Question 1

Free month promotions are effectively a discount for the customer, though hopefully generate some ultimate benefit to revenue. Using the dataset, can you assess how much incremental revenue is generated (or loss) from the free month promotions?

Number of Customers Necessary to break even

Step 1: Obtain the Lost Revenue = Revenue Possible without Promotion – Actual Revenue obtained after the Promotion

Step 2: To Compensate this loss, obtain the value of number of customers are needed without promotion, taking the average revenue by customer without promotion

Step 3: Now obtain the Number of reservations needed to have the number of customers renting (From Step 2) taking the ratio of Number of people Who Used and didn't use the Rental after reservation.

The Total Number of Reservations Required to Break Even No Promotion is 0.0

The Total Number of Reservations Required to Break Even Two Months Half Off is 32.02585077593871

The Total Number of Reservations Required to Break Even First Month Half Off is 272.6769835696182

The Total Number of Reservations Required to Break Even First Month Free is 3636.3758838818885

The Total Number of Reservations Required to Break Even Three Months Half off is 0.0

The Total Number of Reservations Required to Break Even Two Months Free is 13.662832188201225

The Total Number of Reservations Required to Break Even the Promotional Loss Incurred is 3954.7415504156465

When will the promotion break Even in Revenue Perspective?

Difference in Average Revenue with and Without Promotion for Every Quarter

Quarter 1 - 60.27413798547764

Quarter 2 - 49.348163488445834

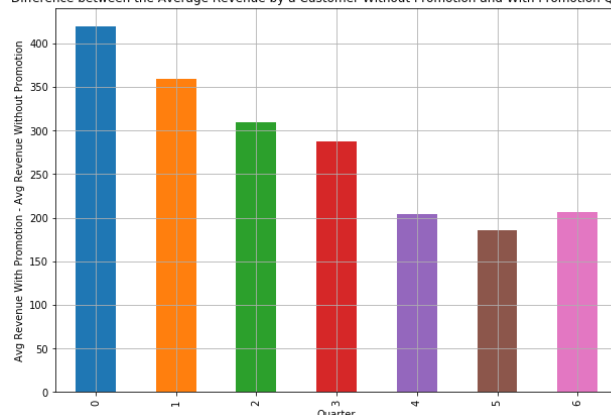
Quarter 3 - 22.06865925402394

Quarter 4 - 83.29524563348019

Quarter 5 - 19.23278083848595

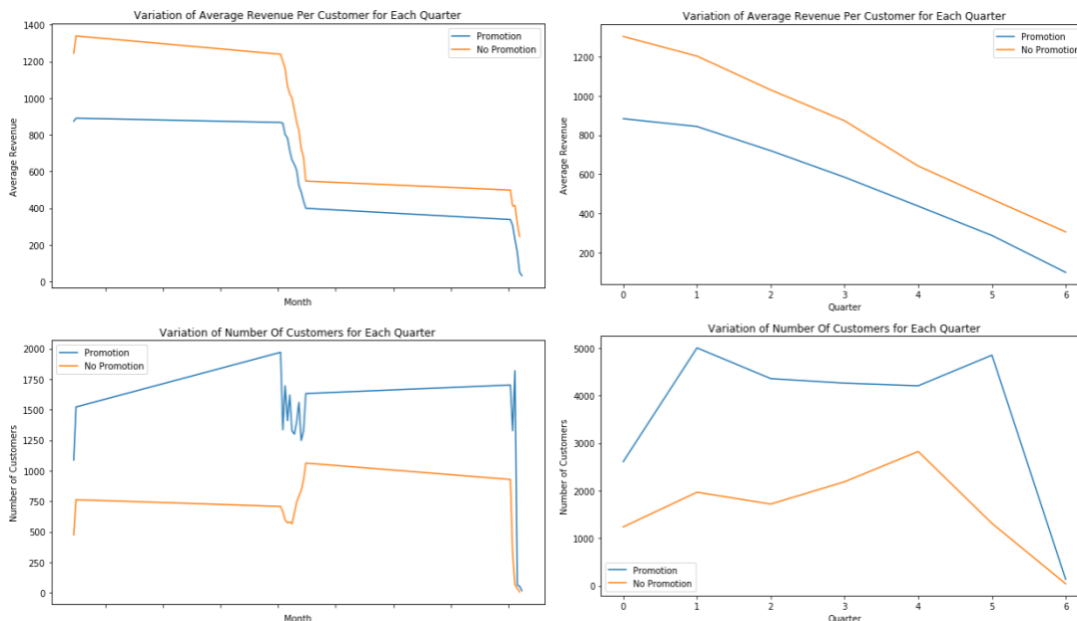
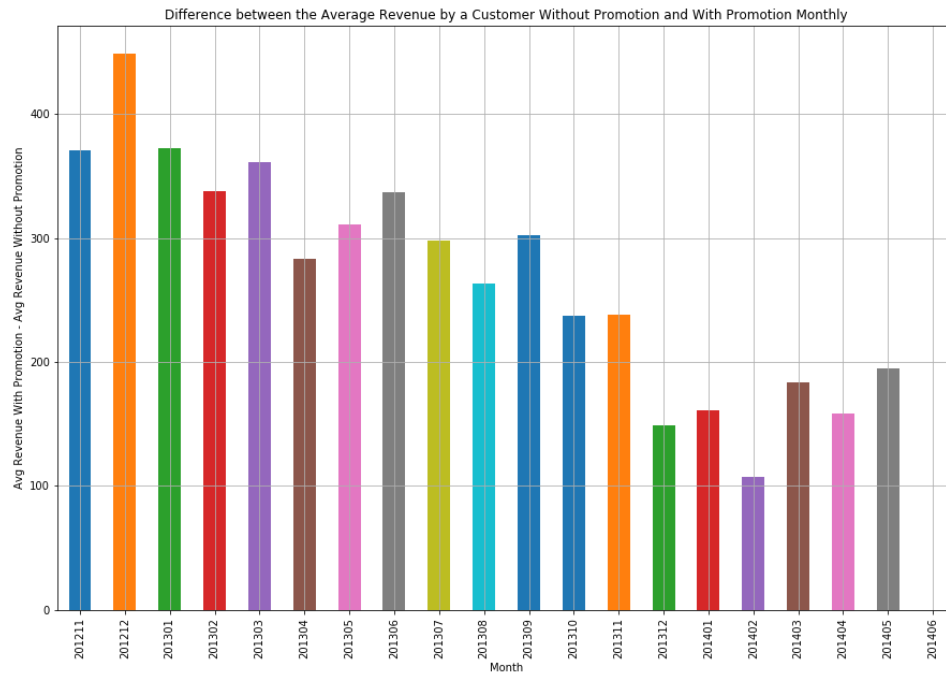
Quarter 6 - 21.187078558100296

Difference between the Average Revenue by a Customer Without Promotion and With Promotion Quarterly



Average Reduction in (Amount Earned W/O Promotion - W/ Promotion For Every Quarter)) is 35.50531810696888
 Number of Quarters Required to Make the Promotion Profitable Following the trend 5.814168714043318

Average Revenue with promotion – Average Revenue Without Promotion – Monthly



The Left two curves show how Average Revenue and Number of Customers with and without promotion on a monthly basis and on the right, it is on a quarterly basis

We see that The Promotions, bought in an increase in the number of customers, but the Average Revenue by a customer without promotion is high than an Average Revenue by a with promotion - So, the promotion is not a success in revenue perspective of view.

Also, with the following trend, we can see that with the next 6 quarters, there is a possibility that the Average Revenue with Promotion will be higher than Average Revenue without promotion - So there is a high chance that after 6 quarters ~ 1.5 Years, we might see a success in giving promotion

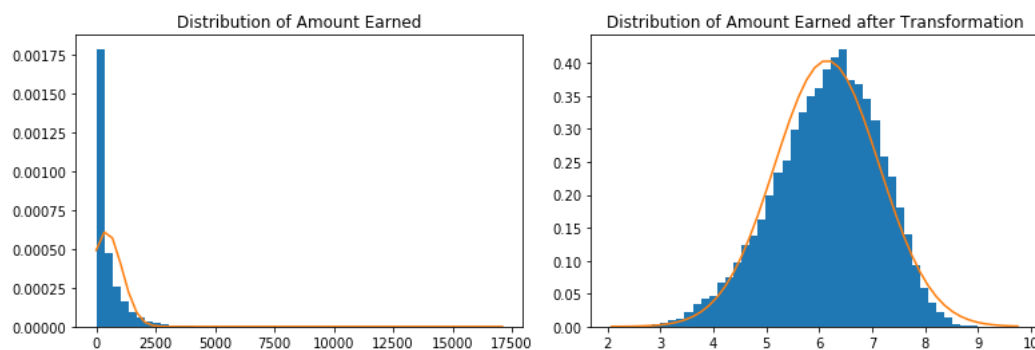
Also, the total number of reservations required to breakeven the current loss incurred by promotion, taking the number of people who reserve and don't use is ~ 3956 Customers

Question 2

Free month promotions are effectively a discount for the customer, though hopefully generate some ultimate benefit to revenue. Using the dataset, can you assess how much incremental revenue is generated (or loss) from the free month promotions?

The Only measure of Criteria that we have is to compare between Promotion and No Promotion. So, if we are able to say that the Average Revenue of Customer with Promotion should at least be equal or greater than Average Revenue of the customer with promotion, for a promotion to be successful.

So, we compare the mean of Average Revenue with Promotion and Average Revenue without Promotion and it could be statistically proved by the T test. For a t test, the basic assumption is the normality, but the Amount Earned is not normally distributed. So, we use Log transform and then we can proceed, as the log transformation made the data to be normally distributed. The graph of distribution before and after transformation is given below



T test to compare the Means of Promotions and No Promotions

- Now, our motive is to understand whether Promotions was more significant than No Promotions. For this we consider only the cases where the customer actually did rent (Rented? = True).
- For understanding the success, we compare the average revenue bought by a single customer without promotions against the average revenue bought by a customer with Promotions. The Average Revenue bought by the customer with promotions should be at least equal or greater than the Average Revenue bought by the customer without promotion.
- Since, the data is normally distributed (After the transformation) and Independent, we can perform t- test to compare the means of these two.

Average Revenue with Promotions – μ_1 | Average Revenue without Promotions – μ_2

Null Hypothesis – $H_0: \mu_1 = \mu_2$
Alternate Hypothesis – $H_1: \mu_1 > \mu_2$

- Here, as far as this problem is concerned, it is preferable to do Welch's T test, because there is high possibility that the Variance of these samples might not be equal.
- So, for the Rejection Criteria for the Null Hypothesis, the test statistic $< -t(\alpha, V)$. Since the Degree of Freedom is Really high, we choose the comparing statistic to be 1.96 (From t table).
- Also, if p-value is less than 0.5, it also rejects the null hypothesis.
- Reject Null Hypothesis, mean of No Promotion is greater than mean of Promotions
- The Difference between Average Revenue Per Customer W/O Promotions and W/ Promotions is 289.08627012480076

We Compare, the Revenue from Customer with Promotion and Revenue from Customer without Promotion, and the mean of them is compared statistically using T-Test, and we see that Average Without Promotion $<$ Average With Promotion. This means that the Promotion wasn't a success.

Also, the incremental loss was given by (Average without Promotion - Average with Promotion) = 289.0862 Dollars Loss

Question 3

To what degree does the length of customers' stays impact your responses above and how should it be considered? What is the most appropriate way to consider customer length of stay, given the number of customers still renting?

- The Customer's Length of Stay is the most crucial part in the analysis. The Revenue Generated is Entirely Dependent on the length of the customer stay.
- Here, we have made an assumption that, the customer did move out on the day we collected the data (06/11/2014). But this is a very vague approximation.
- There is a sophisticated way to obtain the length of the customer's stay.
- We are sure about the Region in which the customer rent and also the Purpose and the Reason for the Storage (Available for most of the Rented = 1).
- With these three approximations, it is possible to have a better estimate of the length of stay of the customer.
- Say, there are a few Customer - Renting in Arizona | Purpose of Business | Storage of business Inventory. There could be 10 such customers with the same set of parameters, and out of which 6 might have moved out and 4 didn't. We average the length of stay of the 6 customers and use that Average as the length of stay for other 4 customers who didn't move out and proceed with analysis.
- But sometimes it is impossible to make it work for all possible combinations of three variables. If that is the case, we can use purpose and Region, and even for a very few cases these two combinations might not work and that case, we can use the average stay of a customer based on region and this average could be used for customers who didn't move.

Question 4

How do customer type (based on survey responses), geographic/regional variations, reservation channel, and price concessions impact consideration of (1) and (2) above?

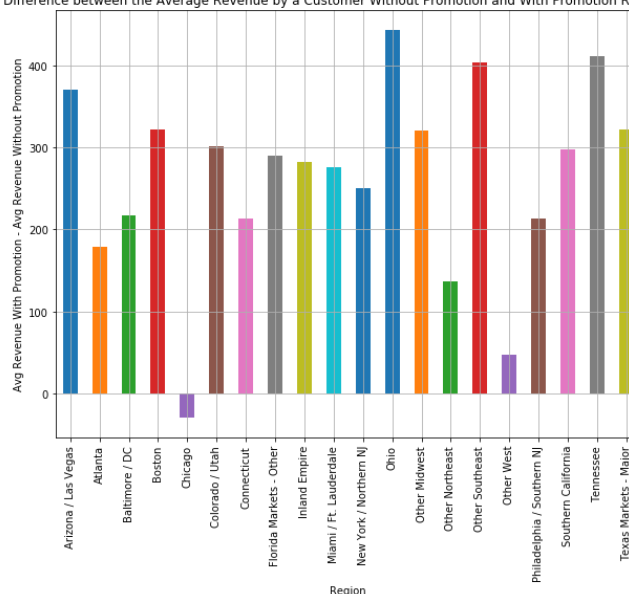
To understand how these parameters affect the Promotion, we compare the Average Revenue Per Customer with Promotion and Without promotion Region wise, Reservation channel wise and Concession wise.

Region

T test for Individual Regions

Inland Empire - Reject Null Hypothesis
Colorado / Utah - Reject Null Hypothesis
Florida Markets - Other - Reject Null Hypothesis
Other Southeast - Reject Null Hypothesis
Baltimore / DC - Reject Null Hypothesis
New York / Northern NJ - Reject Null Hypothesis
Ohio - Reject Null Hypothesis
Southern California - Reject Null Hypothesis
Other Midwest - Reject Null Hypothesis
Boston - Reject Null Hypothesis
Other Northeast - Reject Null Hypothesis
Arizona / Las Vegas - Reject Null Hypothesis
Other West - Accept Null Hypothesis
Atlanta - Reject Null Hypothesis
Texas Markets - Major - Reject Null Hypothesis
Connecticut - Reject Null Hypothesis
Miami / Ft. Lauderdale - Reject Null Hypothesis
Tennessee - Reject Null Hypothesis
Chicago - Accept Null Hypothesis
Philadelphia / Southern NJ - Reject Null Hypothesis

Difference between the Average Revenue by a Customer Without Promotion and With Promotion Regionwise



The Region wise actually will vary the entire concept of promotion. We see that only Chicago and Other West had a significant positive impact on Promotion and we might consider stopping giving the promotion in Regions like Ohio, Arizona, Tennessee.

This Inference is also backed up by the T Test performed between Average Revenue with Promotion and Average revenue without promotion for individual regions

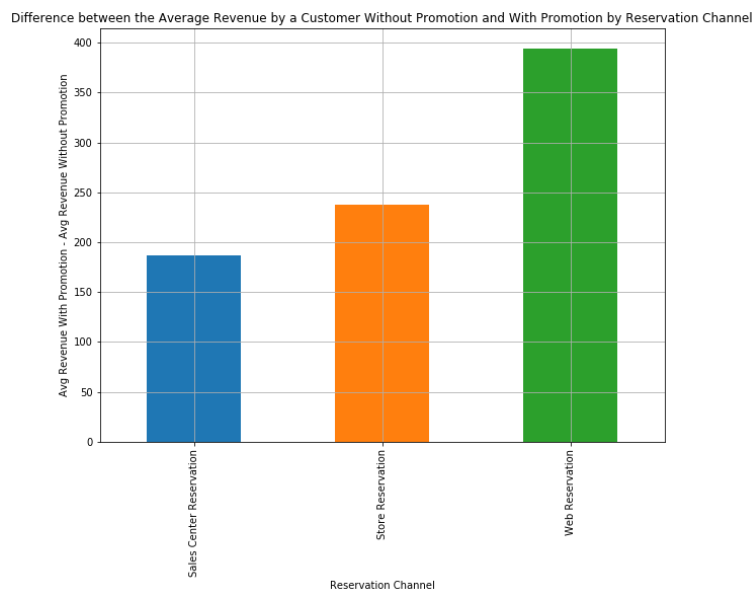
Reservation Channel

T test for Individual Reservation Channel

Web Reservation - Reject Null Hypothesis

Sales Center Reservation - Reject Null Hypothesis

Store Reservation - Reject Null Hypothesis



None of the reservation channel did change the effect of promotions. In Every of the Reservation Channel, we see that Average Revenue with Promotion < Average Revenue Without Promotion, and it is also supported by the T Test performed above

Concession

Since the concession had a very wide range, I just had it as whether concession was offered or not.

T test for concession

No - Reject Null Hypothesis

Yes - Reject Null Hypothesis

The Concession is approximated into whether Concession was given or Not. And with Concession, Yes or No, the result still stays the same. Average Revenue with Promotion < Average Revenue Without Promotion for Both Concession Type

But, may be if we categorize the Concession for 20 bins, between 0 - 5% and so on, may be certain percentage of concession might actually will give a better result, for which the revenue was better with Promotion

Question 5

What type of restrictions and/or creative adjustments to the structure of the free month promotions would you recommend in order to increase the profitability of the promotion? Can these recommendations be justified by the data provided?

There are two major problems that we encounter

1. Because of Monthly Lease, we are unsure of the length of the stay
2. The Promotion offer is given out at random and that is not how it has to be done. There should be a criterion for the promotion offer

Problems:

1. There are few Customers, who will get the promotional offer (One Month Free) and they Move Out Next Month. So, the effective revenue is zero. There is difference between a customer offered Promotional Offer and then stay for just one month and Customer offered promotional offer and stayed for a prolonged period
2. There are few customers who are staying for a prolonged period and they don't get any promotional offer, that need to be considered, in the customer retention point of view

Solution

1. The Promotional Offer should be given to the customers who are disclosing their length of stay prior in hand. Or if they are ready to make a commitment for at least say 5 - 6 Months they could be eligible for the promotional offer. Depending on the number of months they are willing to sign the lease, we could modify the customers Promotional Offer.
2. If they are sure that they are going to stay for less than two months or just one month, instead of full month free, we can give half month free promotion.
3. Another advantage of knowing the customer length of stay is, Say, a customer plans to rent for 10 Months and the rent per month is 10 Dollars. If we offer him the first month promotion, and can increase the rent by just 1 Dollar, we can still break even our actual expectation.
4. Or as I said above, dealing with the Length of the stay of the customer, if we are able to develop an algorithm that can predict the average stay of the customer, based on this average length, we get the input from customers for their purpose and Storage Reason, we can predict what their average stay would be and then

offer them promotion based on that. This could be more accurate than forcing the customers to disclose their length of stay to be eligible for promotion offer

Question 6

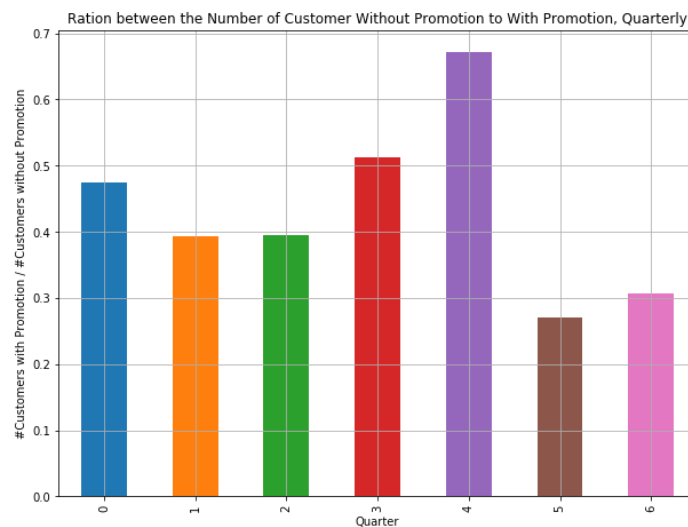
What additional information would be helpful to have to better determine the profitability of the First Month Free promotion?

1. If we had known the revenue or number of customers for any period (Without Promotion) before or after the given time period, it will be possible to know the success
2. If we know the target Revenue, or the target number of customers, it could help us in understanding the success

Question 7

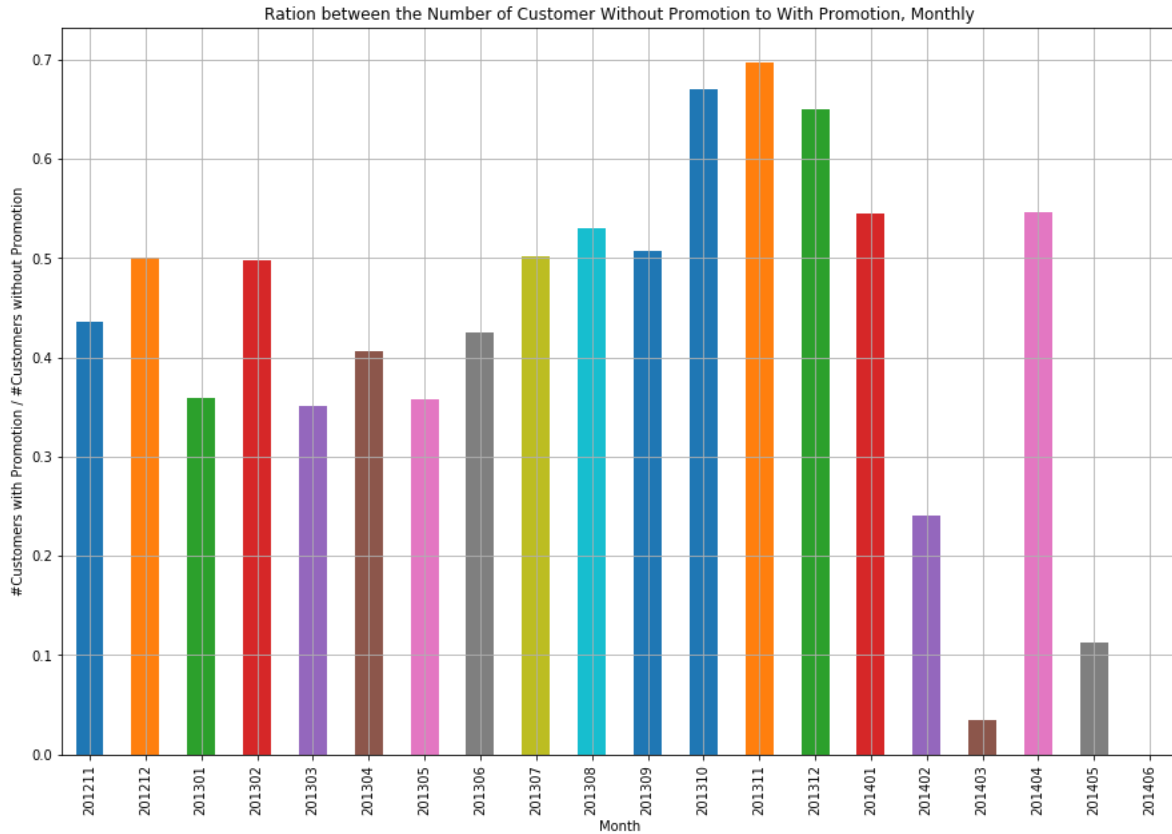
Estimate probabilistically the likelihood that CubeSmart received more rentals as of result of offering promotions

The Promotion is a success in increasing the number of customers perspective. We see these two graphs below (Quarterly and Monthly) and see that Number of Customers with Promotion > Number of Customers without Promotion



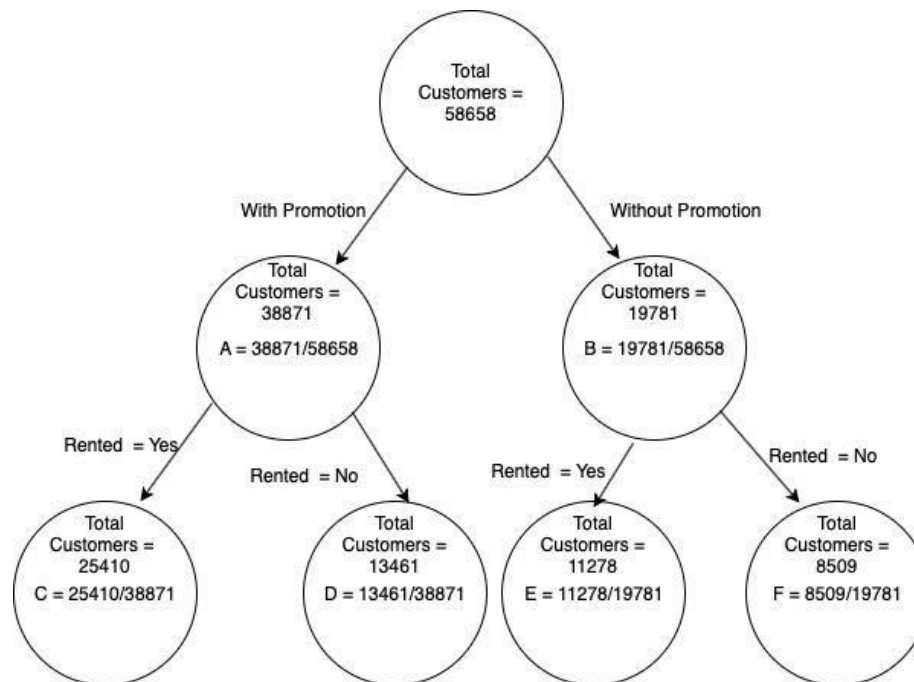
Quarter

| | |
|---|----------|
| 0 | 0.473745 |
| 1 | 0.393121 |
| 2 | 0.394259 |
| 3 | 0.513031 |
| 4 | 0.671265 |
| 5 | 0.270421 |
| 6 | 0.305970 |



The above two graphs represent the Ratio of Number of Customers with Promotion To Number of Customers without Promotion, Quarterly and Monthly Basis

| | | AmountEarned | | |
|--------------|--------|--------------|-------|------------|
| | | sum | count | mean |
| Promo | Rented | | | |
| No Promotion | 0 | 0.00 | 8509 | 0.000000 |
| | 1 | 10098662.44 | 11278 | 895.430257 |
| Promotion | 0 | 0.00 | 13461 | 0.000000 |
| | 1 | 15407200.71 | 25410 | 606.343987 |



LEGEND

A - Probability (Number of Customers with Promotion)
 B - Probability (Number of Customers with No Promotion)
 C - Probability (Number Of Customers Renting I Promotion)
 D - Probability (Number Of Customers Not Renting I Promotion)
 E - Probability (Number Of Customers Renting I No Promotion)
 F - Probability (Number Of Customers Not Renting I No Promotion)

Flow chart division based on the Rental and Promotion

- $P(\text{Number of Rentals} \mid \text{Promotion}) = 25410 / (13461 + 25410)$
- $P(\text{Number of Rentals} \mid \text{No Promo}) = 11278 / (8509 + 11278)$
- Likelihood that we receive more rentals due to promotion = $P(\text{Number of Rentals} \mid \text{Promotion}) / P(\text{Number of Rentals} \mid \text{No Promotion})$
- Probabilistically, the likelihood of Receiving more rentals due to promotion is 1.1469033336466181

Probabilistically, because of promotion, we are 1.14 times more likely to have more rentals than if we don't have promotion

Question 8

Estimate probabilistically the likelihood CubeSmart increased revenue when offering these promotions

We know the Revenue for the company with promotion and without promotion. So, to obtain the likelihood for the increased revenue when offering these promotions is given by $[(\text{Revenue with Promotion}) - (\text{Revenue without Promotion})] / [(\text{Revenue with Promotion}) + (\text{Revenue without Promotion})]$

Probabilistically, the likelihood of CubeSmart receiving Increased Revenue when offering the promotion is 0.21

Findings and Results

1. To Summarize the entire work, we were able to understand that although promotional offer was successful in the number of customers perspective, it failed in the revenue perspective. The Average Revenue from a customer with promotion < Average Revenue Without Promotion
2. The likelihood of having more rentals because of Promotion is 1.14
3. So, to break even, we need around ~4000 reservations to break in, and with the current trend it is also possible to observe that after 6 quarters, the Average Revenue with Promotion will exceed the Average Revenue Without Promotion.
4. The Incremental loss of the Promotion is given by $[\text{Average Revenue (Without Promotion)} - \text{Average Revenue (With Promotion)}] = 289.08622$ Dollars
5. Considering the Promotion and No promotion with respect to region, reservation channel and Concession didn't alter our findings to a greater extent
6. The other possible works that could be done for the better Revenue of the company and the better profitability of the promotion is detailed below

Future Work

1. As mentioned above, regarding a way to work with the Length of stay of the customer, I would like to predict the customer length based on the aggregation of the three other variables (Region, purpose and Reason)
2. We have assumed that the variance of Amount Earned With promotion and without promotion to be unequal and we used Welch T Test for all comparison of mean. But we can do Chi squared test to know whether the Variance are equal between the two samples and decide whether to use pooled T test or Welch T Test to compare the mean of the two samples for better significance.
3. Also, another way to deal with the Length of stay is to consider all the instances that have the actual length of stay and build a model based on this (Probably without the Move Out Date), and once we are able to develop a prediction model, then it is possible to predict the length of stay for all the unknown instances.
4. For Concession, I have approximated into two different categories, Whether Concession was offered, Yes or No. But it would be more relevant to make the concession into categories of 20 bins (0 - 5% in one bin, 6 - 10% in one bin), and see if there was any range in which promotion might have made a positive impact
5. We can develop a geographical plot (Geographical heat map), to understand whether any pattern is there on the Amount Earned or the Number of customers. Say there could be more concentration customers around California and AZ and the entire west coast than the east coast. If we are able to understand these geographical impacts, then it will be possible to make better understanding about the promotion

6. We have made an estimate that in the next 6 quarters, there is a possibility that the Promotion does actually impact the revenue positively. So, we can try to see whether this is true using the Time Series Analysis. We aggregate the average salary monthly from 11/2012 - 05/2016 and then we use the Seasonal ARIMA to develop Model for the Amount Earned Vs Time, and using this model, we predict the future amount earned for both with promotion and Without Promotion. Now we can see whether the Promotion was making an impact positively to the revenue
7. To study how NYC is one of the highest revenue generators despite its less rental area. There could be a possibility that the Two Months half off promotion worked in NYC. Because it is the only placed that has offered this more and it has generated high revenue. Need to study on how Two months half off did affect the revenue on NYC.