أكاديمية مسك
MISK ACADEMY

UDACITY

TECH LAB

Predictive Analytics for Business

Project #5 A/B Test a New Menu Launch

Name: Marwan Saeed Alsharabbi

Date: 29-Jan-2020

2020

INTRODUCTION

A/B Testing Summary

Summary: In an A/B test, a change is applied to a treatment group, and its performance is compared against a control group to estimate the impact of the change**.**

STEP 1: SELECT A PERFORMANCE METRIC

It's important to understand the metric used to evaluate the results of the test. Whether the goal is to increase sales, profit, conversion rate, etc., this should be specified at the upfront**.**

STEP 2: SELECT THE EXPERIMENT DESIGN

Matched pair - when the sample size is small and/or the data is difficult to collect, a matched pair experiment should be used.Randomized design - when the sample size is large and the data is easy to collect, then a randomized experiment should be used. Randomized experiments are very common for web-based AB tests.

STEP 3: SELECT TREATMENT AND CONTROL UNITS

Each individual in the test is considered a unit. The unit can be a person, store, etc. In a test, units are split into two groups, the treatment group and control group. Treatment and control units are compared against each other

Useful Alteryx Tool: Formula

STEP 4: SELECT EXPERIMENTAL AND CONTROL VARIABLES

Experimental variable - The experimental, or treatment, variable(s) is the variable that is different between treatment and control units. For example, if you are testing a new price point, the experimental variable would be price.

Control Variables - The control variables are the variables that should remain constant between test and control groups. These variables ensure that the treatment and control groups are representative of each other and that the results will apply to the population. Control variables are used to match each treatment unit to one or more control units.

Useful Alteryx Tool: AB Controls

STEP 5: DETERMINE TEST DURATION AND SAMPLE SIZE

These two go hand in hand and contribute most directly to statistical significance. You can improve statistical significance by either increasing the sample size or test duration. Generally, the duration of a test should be at least as long enough to capture a representative sample.

STEP 6: RUN THE TEST AND PREPARE THE DATA

Now it's time to run the test and collect the data. Preparing the data includes filtering for the dates of the test, ensuring there are no duplicate records, removing records with incomplete data, and removing outliers.

STEP 7: ANALYZE RESULTS

Lift - Compare the average performance between the two groups. It can also be useful to understand the distribution of the performance of the units.

Statistical Significance - Performing a t-test provides a p-value. P-values below 0.05, indicate statistically significant results. Paired t-test are used for matched pair experiments and unpaired t-test for randomized experiments.

Impact Estimation - In order to provide an expected impact of broad implementation of the treatment, apply the lift calculation to the entire population.

Useful Alteryx Tool: AB Analysis

Project Overview

You're a business analyst for Round Roasters, a coffee restaurant in the United States of America. The executive team conducted a market test with a new menu and needs to figure whether the new menu can drive enough sales to offset the cost of marketing the new menu. Your job is to analyze the A/B test and write up a recommendation to whether the Round Roasters chain should launch this new menu.

How Do I Complete this Project?

This project uses skills learned throughout the "A/B Testing" course. To complete this project:

Go through the course
Apply the skills learned in the course to solve the business problem given in the project details section.

Use our guidelines and rubric to help build your project.
When you're ready, submit it to us for review using the submission template found in the supporting materials section.

Skills Required

In order to complete this project, you must be able to:

Cleanup, format, and blend a wide range of data sources
Plan and analyze A/B tests

Steps to Success

Step 1: Plan Your Analysis

To perform the correct analysis, you will need to prepare a data set. Prior to rolling up your sleeves and preparing the data, it's a good idea to have a plan of what you need to do in order to prepare the correct data set. A good plan will help you with your analysis. Here are a few questions to get you started:

- What is the performance metric you'll use to evaluate the results of your test?
- What is the test period?
- At what level (day, week, month, etc.) should the data be aggregated?

Step 2: Clean Up Your Data

In this step, you should prepare the data for steps 3 and 4. You should aggregate the transaction data to the appropriate level and filter on the appropriate data ranges. You can assume that there is no missing, incomplete, duplicate, or dirty data. You're ready to move on to the next step when you have weekly transaction data for all stores.

Step 3: Match Treatment and Control Units

In this step, you should create the trend and seasonality variables, and use them along with you other control variable(s) to match two control units to each treatment unit. Treatment stores should be matched to control stores in the same region. *Note: Calculate the number of transactions per store per week and use 12 periods to calculate trend and seasonality.*

Apart from trend and seasonality...

- What control variables should be considered? Note: Only consider variables in the RoundRoastersStore file.
- What is the correlation between your each potential control variable and your performance metric? (Example of correlation matrix below)
- What control variables will you use to match treatment and control stores?

Step 4: Analysis and Writeup

Conduct your A/B analysis and create a short report outlining your results and recommendations.

In an AB Analysis we use the correlation matrix to find the most correlated variable to the performance metric to include in the AB controls tool to help find the best matches.

A/B Test Guide

The A/B test guide can help you map out the process you'll need to go through to complete an A/B test. This will be useful for the project, as well as any A/B tests you do in a professional setting.

Data

round-roaster-stores.csv - This file contains store information for each Round Roaster store in the USA.
treatment-stores.csv - This file contains store information for each store that offered the new menu items.
round-roaster-transactions.zip - This file contains transaction level information for all of Round Roaster's stores

Aggregate and Export

We recommend you save your aggregated transaction database as a separate file to further reduce your development time. You shouldn't need to re-aggregate the transaction database every time you want to test out a new workflow.

Creating a week variable

When calculating the week column, assign week 1 to the first week of the test period, so the week_begin variable for week 1 would be your test start date. This will cause weeks prior to the test period to be negative, which is perfectly fine. See here . If the first link doesn't work, please click here for an example of how to calculate the week variable using a formula tool in Alteryx.

Number of Weeks

The trend tool is used to create trend and seasonality variables to use as control variables. To do this, you need at least 52 weeks of data, plus the number of weeks you select in the tool to calculate trend, before the beginning of the test start date. In lesson 4, you used 6 weeks to calculate the trend, so you needed 58 weeks prior to the test start date. For the project, you are asked to use 12 weeks to calculate trend, so you'll need 64 weeks of data prior to the test start date. Since the test lasts for 12 weeks, this means you'll need a total 76 weeks of data.
Your filter at the beginning of your workflow should go back 76 weeks from the end of the test period 2016-July-21. In Alteryx the expression shuold look like [Invoice Date]>="2015-02-06" AND [Invoice Date]<"2016-07-22"
All stores should have 76 weeks of Data.

AB Trend Tool

The Test start date is 2016-April-29

The performance metric for this tool is the invoice count per week which represent weekly foot traffic. You had to create this variable with a summarize tool.

AB Controls Tool

You should use 3 numeric measures to match treatment and control stores.

1. Trend
2. Seasonality
3. AvgMonthSales (This should be determined by looking at the correlation between the appropriate numeric variables in the round roasters stores file AvgMonthSales and Sq_ft with the performance metric gross margin.) - This variable is in the round-roaster-stores.csv file

AB Analysis Tool

Make sure to use weekly gross margin per store in all lift calculations and not total sales. Your data has gross margin in it but you will have to use a summarize tool to get weekly gross margin per store.

Steps

1. Filter the data to the proper date range
2. Aggregate the data to get the weekly gross margin and weekly traffic count (count of unique invoices)
3. Calculate Trend and Seasonality with the AB Trend Tool
4. Label the data as treatment and control stores
5. Calculate correlation between other numeric measures and the performance metric (gross margin)
6. Match Treatment to control stores per region using the AB Controls Tool
7. Calculate lift from control to treatment store with AB Analysis tool

If you run into errors in Alteryx or unexpected results from a tool, we have a guide to help you figure out what is going on.

Alteryx Debugging Guide: See the Resources tab in the left most panel of your classroom for a downloadable PDF of the Alteryx Debugging Guide

Step 1: Plan Your Analysis

To perform the correct analysis, you will need to prepare a data set. (500 word limit)
Answer the following questions to help you plan out your analysis:

The Business Problem

Round Roasters is an upscale coffee chain with locations in the western United States of America. The past few years have resulted in stagnant growth at the coffee chain, and a new management team was put in place to reignite growth at their stores.

The first major growth initiative is to introduce gourmet sandwiches to the menu, along with limited wine offerings. The new management team believes that a television advertising campaign is crucial to drive people into the stores with these new offerings.

However, the television campaign will require a significant boost in the company's marketing budget, with an unknown return on investment (ROI). Additionally, there is concern that current customers will not buy into the new menu offerings.

To minimize risk, the management team decides to test the changes in two cities with new television advertising. Denver and Chicago cities were chosen to participate in this test because the stores in these two cities (or markets) perform similarly to all stores across the entire chain of stores; performance in these two markets would be a good proxy to predict how well the updated menu performs.

The test ran for a period of 12 weeks (2016-April-29 to 2016-July-21) where five stores in each of the test markets offered the updated menu along with television advertising.

The comparative period is the test period, but for last year (2015-April-29 to 2015-July-21).

You've been asked to analyze the results of the experiment to determine whether the menu changes should be applied to all stores. The predicted impact to profitability should be enough to justify the increased marketing budget: at least 18% increase in profit growth compared to the comparative period while compared to the control stores; otherwise known as incremental lift. In the data, profit is represented in the gross_margin variable.

You have been able to gather three data files to use for your analysis:

- Transaction data for all stores from 2015-January-21 to 2016-August-18
- A listing of all-Round Roasters stores
- A listing of the 10 stores (5 in each market) that were used as test markets.

1-What is the performance metric you'll use to evaluate the results of your test?

analyze the results of the experiment to determine whether the menu changes should be applied to all stores. The predicted impact to profitability should be enough to justify the increased marketing budget: at least 18% increase in profit growth compared to the comparative period while compared to the control stores; otherwise known as incremental lift. In the data, profit is represented in the gross_margin variable. the increased marketing budget necessary for the addition of gourmet sandwiches and limited wine offerings to Round Roaster's menu. And decide on whether the TV ad campaign will have an impact on sales.

2-What is the test period?

The test period is between 2016-April-29 to 2016-July-21, which is about 12 weeks.

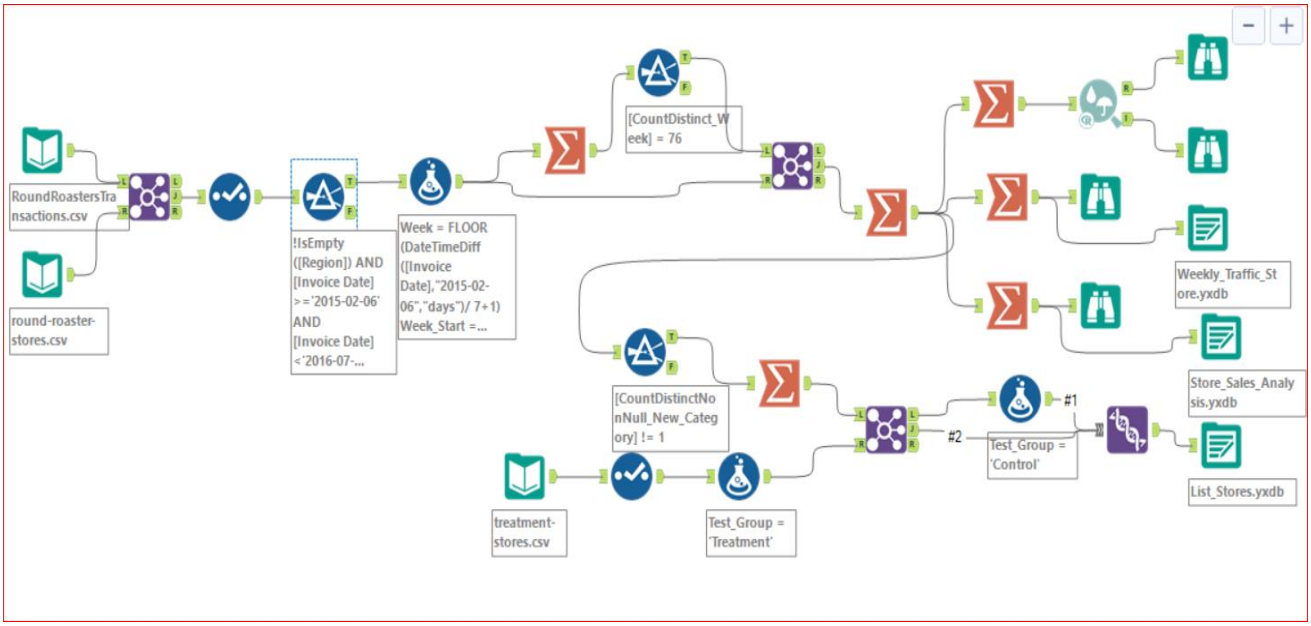
3-At what level (day, week, month, etc.) should the data be aggregated?

Data were aggregated weekly level transaction data from 06-02-2015 to 22-07-2016 76 week all stores

Step 2: Clean Up Your Data

In this step, you should prepare the data for steps 3 and 4. You should aggregate the transaction data to the appropriate level and filter on the appropriate data ranges. You can assume that there is no missing, incomplete, duplicate, or dirty data. You're ready to move on to the next step when you have weekly transaction data for all stores.

In this step, I collected data from the two tables *round-roaster-transactions.csv* and *round-roaster-stores.csv* because there was no Region and I chose the appropriate fields for analysis as I learned from lessons I prepared the data by aggregating the transaction data to the weekly level with the appropriate dates. We were given a file for transaction data for all stores from 2015-January-21 to 2016-August-18 but we only needed 76 weeks of data, that is between 2015-Feb-06 and 2016-Jul-22.
I resumed all this by creating a Weekly_Traffic_Store file ,Store_Sales_Analysis file. And List_Stores Then I created a store-list file that group all the stores with a treatment or control flag for each store. See the Alteryx Workflow next page



Step 3: Match Treatment and Control Units

In this step, you should create the trend and seasonality variables, and use them along with you other control variable(s) to match two control units to each treatment unit. Note: Calculate the number of transactions per store per week to calculate trend and seasonality.

Apart from trend and seasonality...

1-What control variables should be considered? Note: Only consider variables in the RoundRoastersStore file.

I am select variables from two table that should be considered (StoreID, Invoice Number, Invoice Date, Category, Product, Gross Margin, Sales, Sq_Ft, AvgMonthSales, Region) These variables needed in the implementation of the project

2-What is the correlation between your each potential control variable and your performance metric?

We can see that AvgMonthSales has a high correlation of 0.99 with the Sum of Gros Margin. But, sq_Ft has a poor correlation of -0.04.

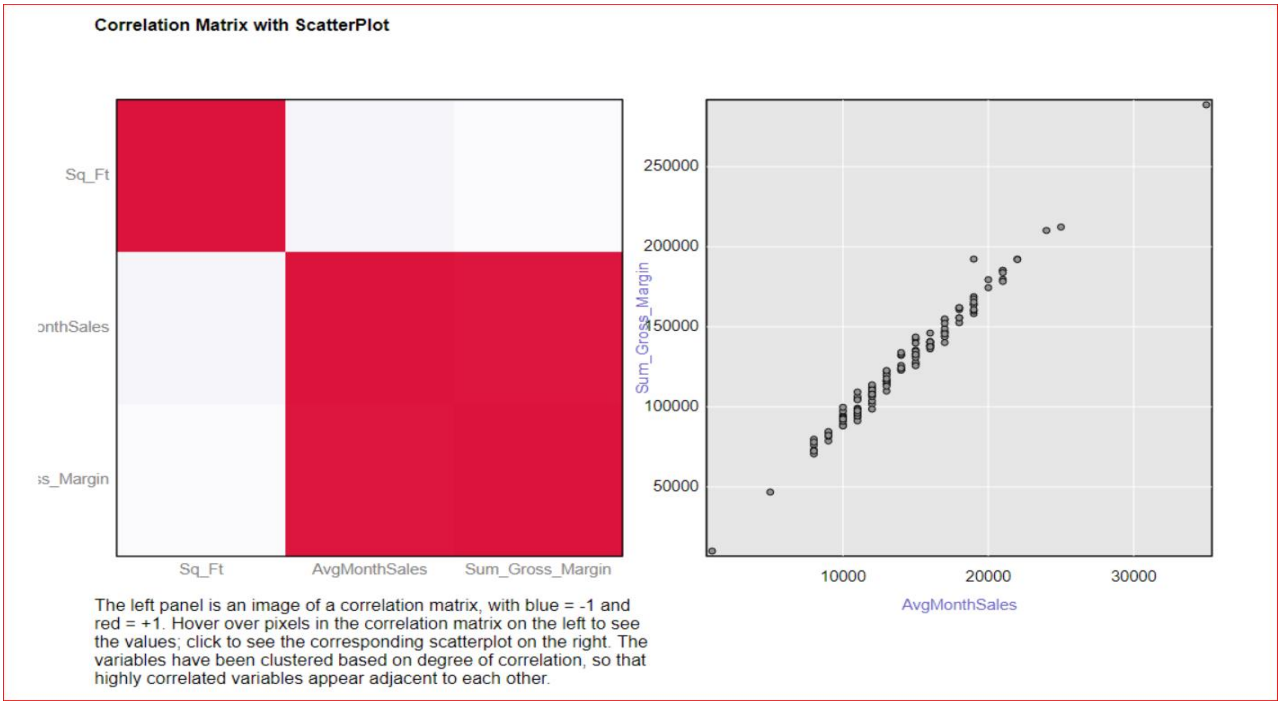
Pearson Correlation Analysis

Full Correlation Matrix

	Sum_Gross_Margin	AvgMonthSales	Sq_Ft
Sum_Gross_Margin	1.000000	0.990982	-0.024255
AvgMonthSales	0.990982	1.000000	-0.046967
Sq_Ft	-0.024255	-0.046967	1.000000

Matrix of Corresponding p-values

	Sum_Gross_Margin	AvgMonthSales	Sq_Ft
Sum_Gross_Margin		0.00000	0.78168
AvgMonthSales	0.00000		0.59138
Sq_Ft	0.78168	0.59138	



3-What control variables will you use to match treatment and control stores?

will use AvgMonthSales with Trend and Seasonality when matching treatment and control stores.

4-Please fill out the table below with your treatment and control stores pairs:

Treatment Store	Control Store 1	Control Store 2
1664	7182	8112
1675	1580	1807
1696	1964	1863
1700	2014	1630
1712	8162	7434
2288	9081	2568
2293	12219	9524
2301	3102	9238
2322	2409	3235
2341	12536	2383

Step 4: Analysis and Writeup

Conduct your A/B analysis and create a short report outlining your results and recommendations. (250 words limit)

Answer these questions. Be sure to include visualizations from your analysis:

1-What is your recommendation - Should the company roll out the updated menu to all stores?

Based on the A/B analysis, Round Roasters should roll out the updated menu to all stores because the predicted impact to profitability justifies the increased marketing budget as the lift from the new menu offerings exceeds the 18% threshold.
For the central region, the lift was 43.5% with a 99.6% significance level.
For the west region, the lift was 37.9% with a 99.5% significance level.
The overall lift for the new menu was 40.7% with a 100% significance level.

2-What is the lift from the new menu for West and Central regions (include statistical significance)?

West regions:

Analysis of the Test on the Measure Weekly_Gross_Margin

Test Start Date: 2016-04-29
Test End Date: 2016-07-21
Additional Information:

Test Summary

The average percentage change in Weekly_Gross_Margin was 39.2% for the treatment units in the test period relative to the comparison period. This same measure was 1.9% for the control units, with the difference between the treatment and control units being 37.2%, which is highly statistically significant. More detailed summary statistics for the treatment and control groups are contained in the first table (which immediately follows), while the details of the hypothesis test of a significant difference in the mean average percentage change in Weekly_Gross_Margin is contained in a table at the end of this report.

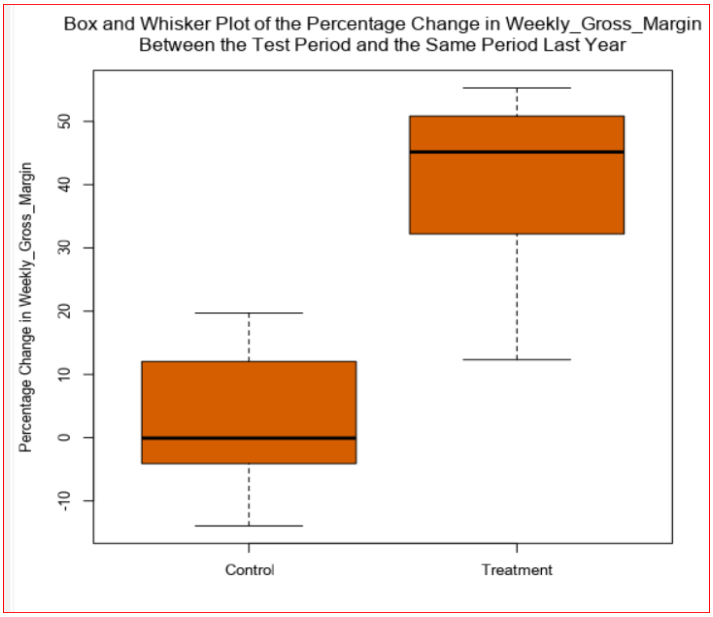
A comparison of the treatment-control pairs indicates an average lift in Weekly_Gross_Margin for the treatment units over the control units of 37.9%, which results in an expected impact of 527 on Weekly_Gross_Margin, with 100.0% of the treatment-control pairs exhibiting a positive lift for the treatment units.

Lift Analysis for Weekly_Gross_Margin

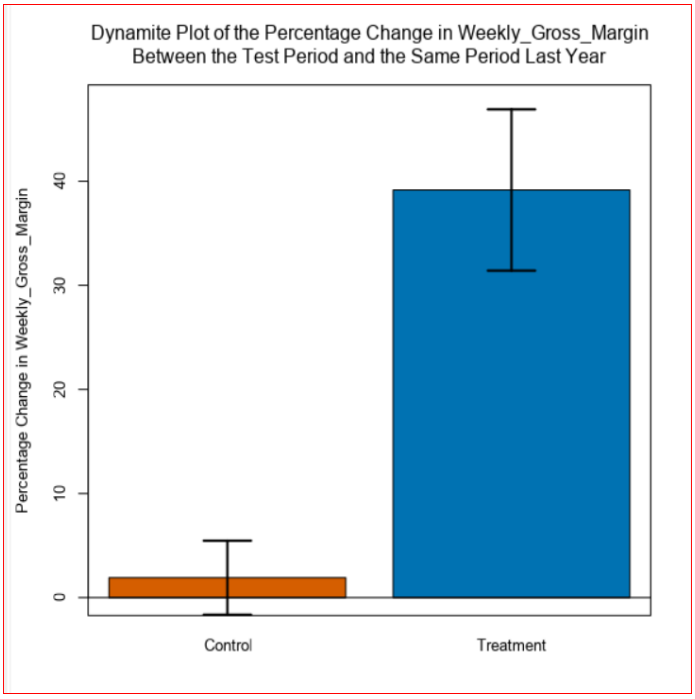
Lift	Expected Impact	Significance Level
37.9%	527	99.5%

Summary Statistics for Weekly_Gross_Margin by Test Group

Statistic	Treatment	Control
Average	39.17	1.92
Minimum	12.34	-13.96
Maximum	55.30	19.70
Standard Deviation	16.34	11.24



The box and whisker plot provides information about the distribution of the percentage change in Weekly_Gross_Margin between the test period and the same time period one year earlier for both the treatment and control units. In the plots, the bottom and top of the orange box are at the first and third quartiles, and the line inside the box is at the median for each test group. The vertical bars that extend above and below the box indicate the range of the data that is within 1.5 times the interquartile range (where the interquartile range is the difference between the values of the third and first quartiles).



In the dynamite chart the blue bar gives the mean value of the percentage change in Weekly_Gross_Margin between the test period and the same time period one year earlier for both the treatment and control groups, while the black error bars provide a one standard error range around the mean values. If there is a statistically significant effect associated with the treatment, then the error bars for the treatment and control groups should not overlap.

Central Regions:

Analysis of the Test on the Measure Weekly_Gross_Margin

Test Start Date: 2016-04-29
Test End Date: 2016-07-21
Additional Information:

Test Summary

The average percentage change in Weekly_Gross_Margin was 39.7% for the treatment units in the test period relative to the comparison period. This same measure was -1.7% for the control units, with the difference between the treatment and control units being 41.5%, which is highly statistically significant. More detailed summary statistics for the treatment and control groups are contained in the first table (which immediately follows), while the details of the hypothesis test of a significant difference in the mean average percentage change in Weekly_Gross_Margin is contained in a table at the end of this report.

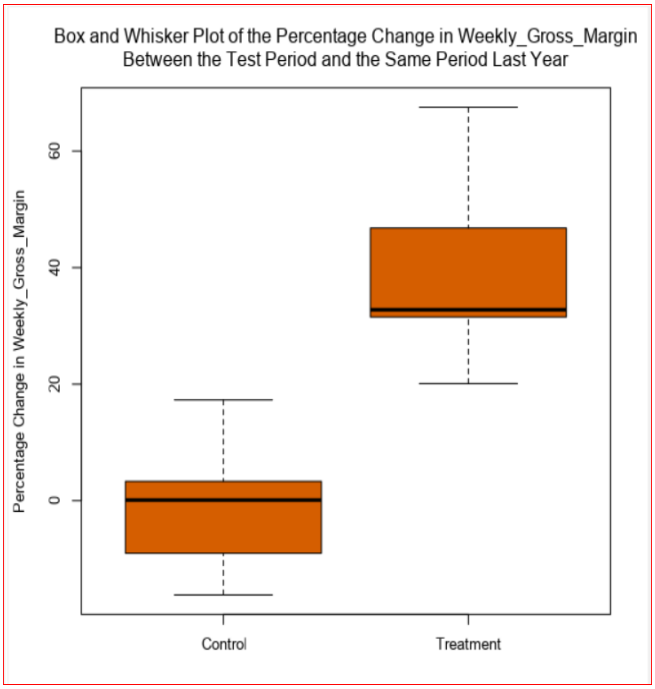
A comparison of the treatment-control pairs indicates an average lift in Weekly_Gross_Margin for the treatment units over the control units of 43.5%, which results in an expected impact of 836 on Weekly_Gross_Margin, with 100.0% of the treatment-control pairs exhibiting a positive lift for the treatment units.

Lift Analysis for Weekly_Gross_Margin

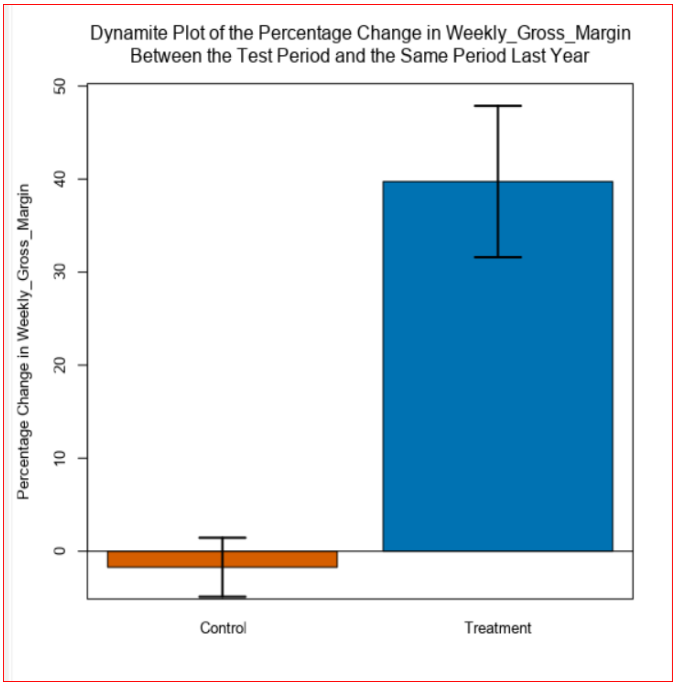
Lift	Expected Impact	Significance Level
43.5%	836	99.6%

Summary Statistics for Weekly_Gross_Margin by Test Group

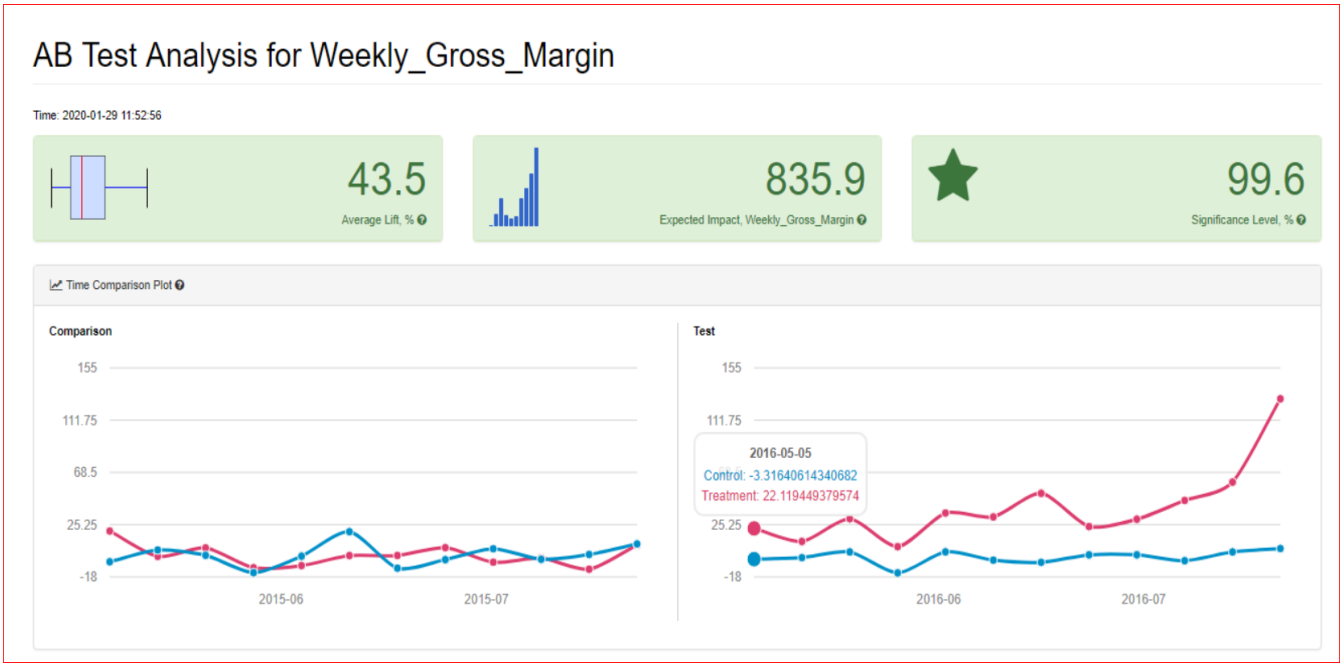
Statistic	Treatment	Control
Average	39.74	-1.73
Minimum	20.09	-16.18
Maximum	67.52	17.29
Standard Deviation	17.15	10.03



The box and whisker plot provides information about the distribution of the percentage change in Weekly_Gross_Margin between the test period and the same time period one year earlier for both the treatment and control units. In the plots, the bottom and top of the orange box are at the first and third quartiles, and the line inside the box is at the median for each test group. The vertical bars that extend above and below the box indicate the range of the data that is within 1.5 times the interquartile range (where the interquartile range is the difference between the values of the third and first quartiles).



In the dynamite chart the blue bar gives the mean value of the percentage change in Weekly_Gross_Margin between the test period and the same time period one year earlier for both the treatment and control groups, while the black error bars provide a one standard error range around the mean values. If there is a statistically significant effect associated with the treatment, then the error bars for the treatment and control groups should not overlap.



3-What is the lift from the new menu overall?

overall Regions:

Analysis of the Test on the Measure Weekly_Gross_Margin

Test Start Date: 2016-04-29
Test End Date: 2016-07-21
Additional Information:

Test Summary

The average percentage change in Weekly_Gross_Margin was 39.5% for the treatment units in the test period relative to the comparison period. This same measure was 0.1% for the control units, with the difference between the treatment and control units being 39.4%, which is highly statistically significant. More detailed summary statistics for the treatment and control groups are contained in the first table (which immediately follows), while the details of the hypothesis test of a significant difference in the mean average percentage change in Weekly_Gross_Margin is contained in a table at the end of this report.

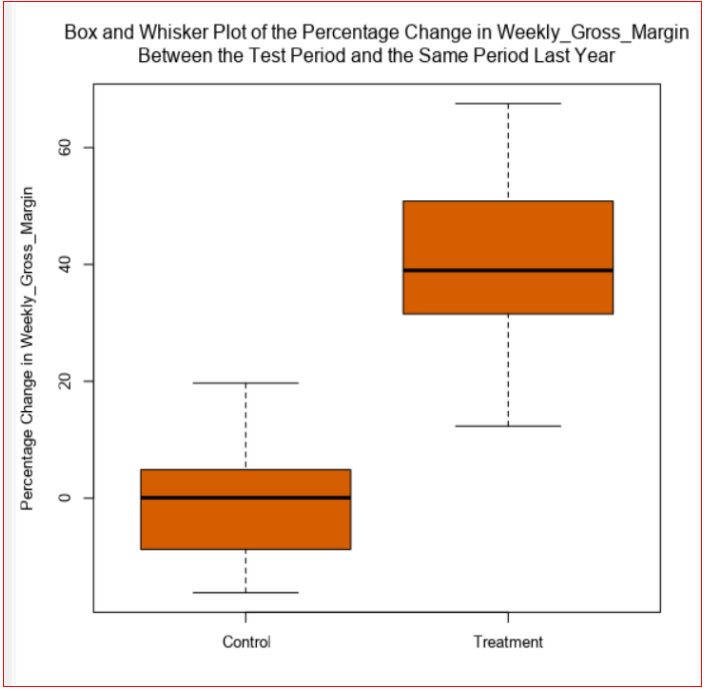
A comparison of the treatment-control pairs indicates an average lift in Weekly_Gross_Margin for the treatment units over the control units of 40.7%, which results in an expected impact of 681 on Weekly_Gross_Margin, with 100.0% of the treatment-control pairs exhibiting a positive lift for the treatment units.

Lift Analysis for Weekly_Gross_Margin

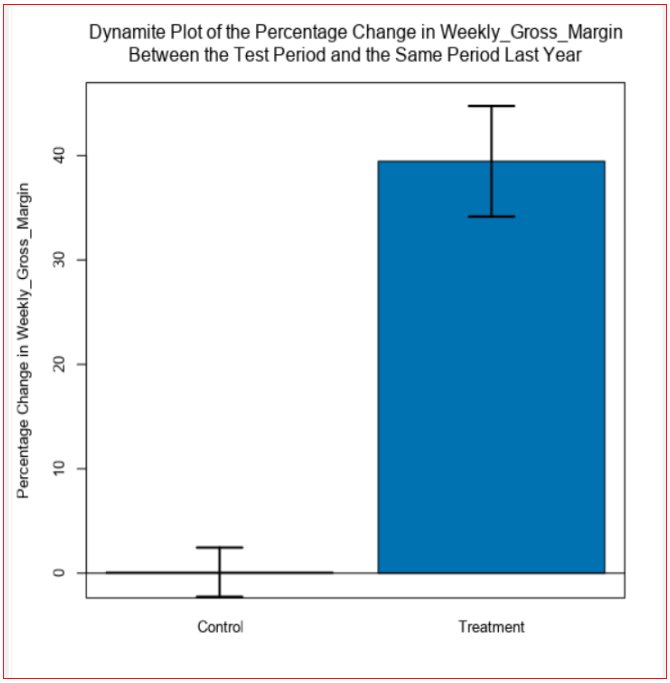
Lift	Expected Impact	Significance Level
40.7%	681	100.0%

Summary Statistics for Weekly_Gross_Margin by Test Group

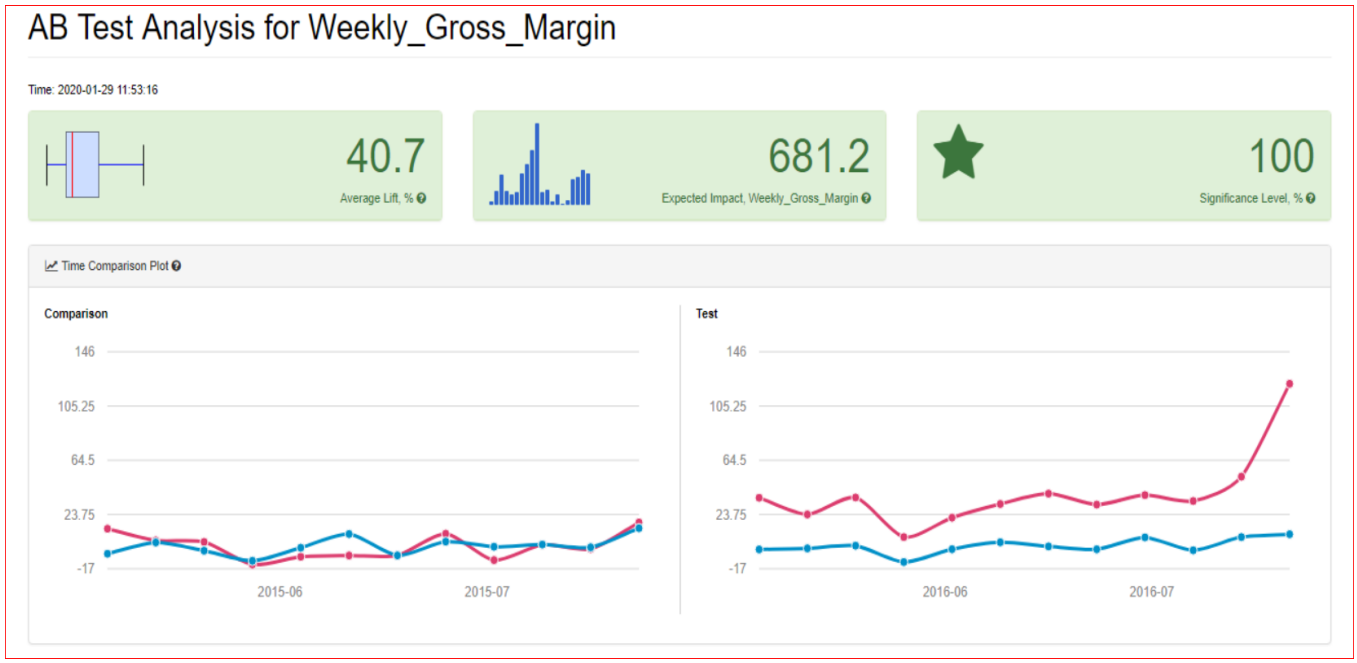
Statistic	Treatment	Control
Average	39.45	0.09
Minimum	12.34	-16.18
Maximum	67.52	19.70
Standard Deviation	16.30	10.54



The box and whisker plot provides information about the distribution of the percentage change in Weekly_Gross_Margin between the test period and the same time period one year earlier for both the treatment and control units. In the plots, the bottom and top of the orange box are at the first and third quartiles, and the line inside the box is at the median for each test group. The vertical bars that extend above and below the box indicate the range of the data that is within 1.5 times the interquartile range (where the interquartile range is the difference between the values of the third and first quartiles).

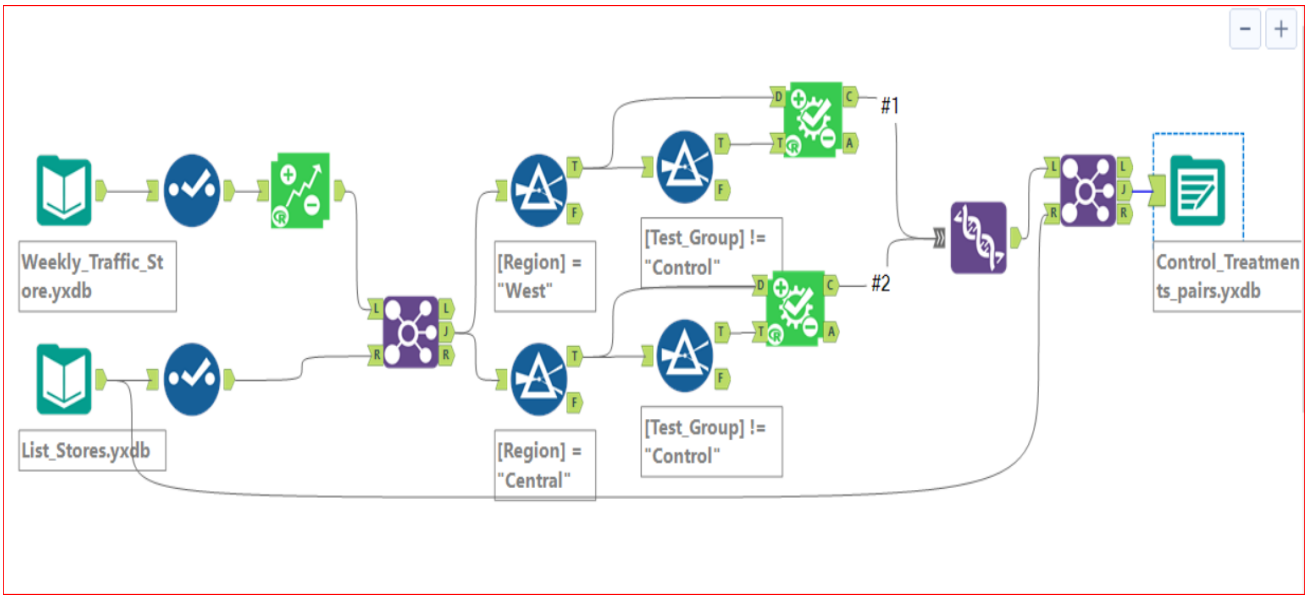


In the dynamite chart the blue bar gives the mean value of the percentage change in Weekly_Gross_Margin between the test period and the same time period one year earlier for both the treatment and control groups, while the black error bars provide a one standard error range around the mean values. If there is a statistically significant effect associated with the treatment, then the error bars for the treatment and control groups should not overlap.

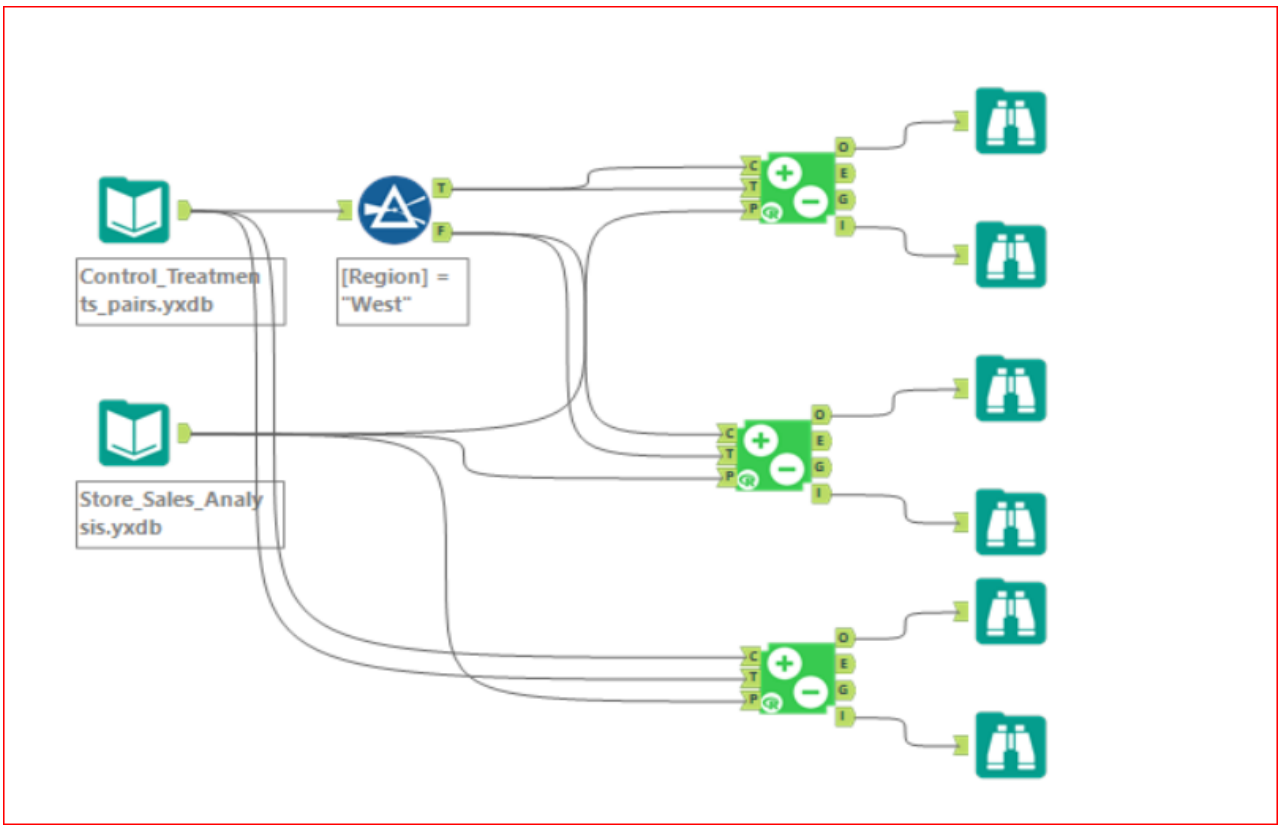


Alteryx Workflow

A/B Trends/Controls:



A/B Analysis:



Before you Submit

Please check your answers against the requirements of the project dictated by the rubric here. Reviewers will use this rubric to grade your project.

I hope to be home to the project requirements despite the valuable information that we learned from the lessons and also the project, but we do not know the exact correct result

Help resources Forums :https://knowledge.udacity.com

I wish success to all.

M*arwan Saeed Alsharabbi*