Kacper Ksieski

Project: Analyzing a Market Test

Complete each section. When you are ready, save your file as a PDF document and submit it here.

Step 1: Plan Your Analysis

To perform the correct analysis, you will need to prepare a data set. (250 word limit)
Answer the following questions to help you plan out your analysis:

1. What is the performance metric you'll use to evaluate the results of your test?
   Round Roasters will decide to roll out the new menu to all stores based on a profit growth of at least 18%. In the data, profit is represented by gross margin.
2. What is the test period?
   The test ran from 2016-04-29 to 2016-07-21. During this time, select stores featured the additional menu items alongside television advertisements.
3. At what level (day, week, month, etc.) should the data be aggregated?
   Each period lasts for one week. Therefore, the data should be aggregated at the week level.

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. 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.
   The most likely control variables are those which deal with the geography or size/sales volume of the store. The following are the candidates:
   - Sq_Ft
   - AvgMonthSales
   - Postal Code
   - Latitude
   - Longitude
   - Current Timezone Offset
2. What is the correlation between your each potential control variable and your performance metric?
   The correlations between the candidate control variables and the gross profit metric are as in shown in the screenshot below.

Record #	FieldName	Sq_Ft	AvgMonthSales	Postal Code	Latitude	Longitude	Current Timezone Offset	Sum_Sum_Gross Margin
1	Sq_Ft	1	-0.049482	-0.039312	0.999999	0.041569	0.011375	-0.019933
2	AvgMonthSales	-0.049482	1	0.156999	-0.049434	-0.267602	-0.305861	0.758595
3	Postal Code	-0.039312	0.156999	1	-0.039176	-0.391528	-0.385508	0.086861
4	Latitude	0.999999	-0.049434	-0.039176	1	0.04133	0.011149	-0.019927
5	Longitude	0.041569	-0.267602	-0.391528	0.04133	1	0.963611	-0.173813
6	Current Timezone Offset	0.011375	-0.305861	-0.385508	0.011149	0.963611	1	-0.205647
7	Sum_Sum_Gross Margin	-0.019933	0.758595	0.086861	-0.019927	-0.173813	-0.205647	1

3. What control variables will you use to match treatment and control stores?
   As indicated in the correlation matrix above, the only candidate control variable with a high correlation to the gross profit is the AvgMonthSales variable with an approximate Pearson R value of .76. This shows that AvgMonthSales strongly moves in the same direction as the gross profit and can be used as a control variable.
4. Please fill out the table below with your treatment and control stores pairs:

Treatment Store	Control Store 1	Control Store 2
1664	7162	8112
1675	1580	1807
1696	1964	7334
1700	2014	1630
1712	8162	7434
2288	9081	2568
2293	12219	9589
2301	11668	9238
2322	2409	3235
2341	2572	3102

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:

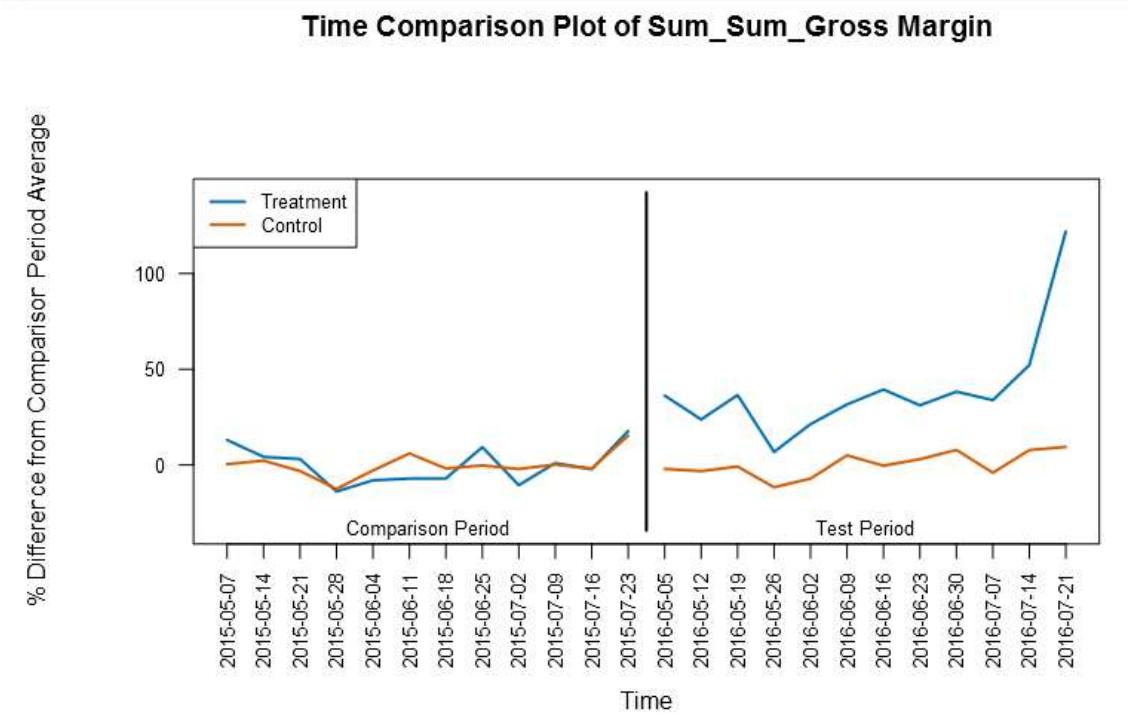
1. What is your recommendation - Should the company roll out the updated menu to all stores?
   The A/B test results are conclusive: Round Roasters should roll out the updated menu to all stores. As shown below, the lift is greater than 18% and it is highly statistically significant.

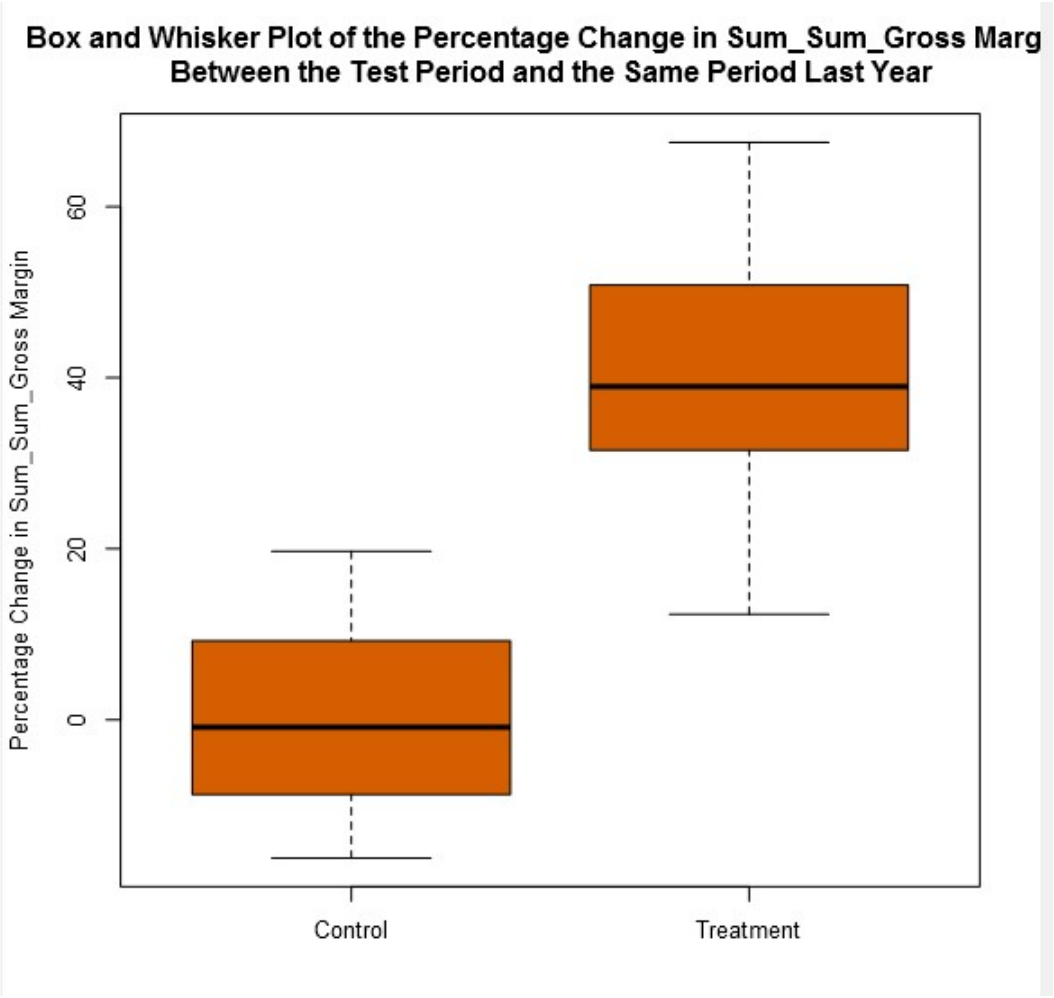
Lift Analysis for Sum_Sum_Gross Margin

Significance Level	Lift	Expected Impact
100.0%	40.0%	672

I have also added two charts below which highlight the stark difference between performance of the controls and the treatments.

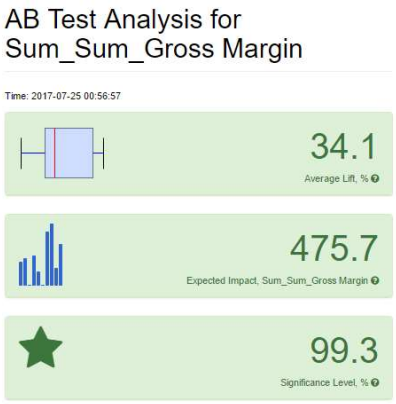
Kacper Ksieski



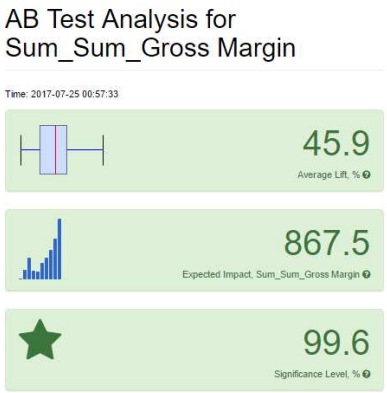


Kacper Ksieski

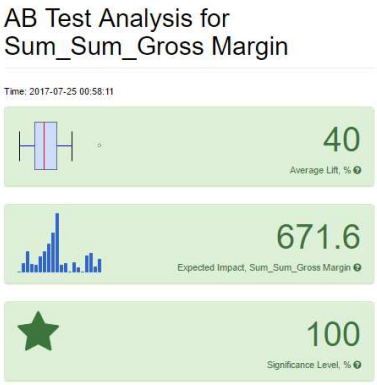
2. What is the lift from the new menu for West and Central regions (include statistical significance)?
   The list from the new menu for the West region is 34.1% with a statistical significance of 99.3%. Please reference screenshot below.



   The list from the new menu for the Central region is 45.9% with a statistical significance of 99.6%. Please reference screenshot below.



3. What is the lift from the new menu overall?
   The lift from the new menu overall is 40% with a statistical significance of 100%. This justifies the rollout of the new menu to all Round Roasters stores nationwide. Please see screenshot below as a reference.



Kacper Ksieski

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.