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. (500 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?
   Gross Margin
2. What is the test period?
   From April 29, 2016 to July 21, 2016 for total 12 weeks.
3. At what level (day, week, month, etc.) should the data be aggregated?
   Week

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.
   Region, Sq_Ft, and AvgMonthSales
2. What is the correlation between your each potential control variable and your performance metric?
   0.174, -0.024, and 0.991 (see table below)

Pearson Correlation Analysis

Focused Analysis on Field Region.num

	Association Measure	p-value	
AvgMonthSales	0.197643	0.022586	*
Sum_Gross.Margin	0.173728	0.045517	*
Sq_Ft	-0.056715	0.516713	

Full Correlation Matrix

	Region.num	Sq_Ft	AvgMonthSales	Sum_Gross.Margin
Region.num	1.000000	-0.056715	0.197643	0.173728
Sq_Ft	-0.056715	1.000000	-0.046967	-0.024255
AvgMonthSales	0.197643	-0.046967	1.000000	0.990982
Sum_Gross.Margin	0.173728	-0.024255	0.990982	1.000000

Matrix of Corresponding p-values

	Region.num	Sq_Ft	AvgMonthSales	Sum_Gross.Margin
Region.num		0.516713	0.022586	0.045517
Sq_Ft	0.516713		0.591380	0.781685
AvgMonthSales	0.022586	0.591380		0.000000
Sum_Gross.Margin	0.045517	0.781685	0.000000	

3. What control variables will you use to match treatment and control stores?
   Trend, Seasonality and AvgMonthSales.
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	1863
1700	2014	1630
1712	8162	7434
2288	9081	2568
2293	9524	12219
2301	3102	9238
2322	2409	3235
2341	2383	12536

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?
   Yes, they should.
2. What is the lift from the new menu for West and Central regions (include statistical significance)?
   43.5% for Central region and 37.9% for West region.
   Central's region report

Report

Analysis of the Test on the Measure Sum_Sum_Gross Margin

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

Test Summary

The average percentage change in Sum_Sum_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 Sum_Sum_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 Sum_Sum_Gross Margin for the treatment units over the control units of 43.5%, which results in an expected impact of 836 on Sum_Sum_Gross Margin, with 100.0% of the treatment-control pairs exhibiting a positive lift for the treatment units.

Lift Analysis for Sum_Sum_Gross Margin

Lift	Expected Impact	Significance Level
43.5%	836	99.6%

West's region report

Report

Analysis of the Test on the Measure Sum_Sum_Gross Margin

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

Test Summary

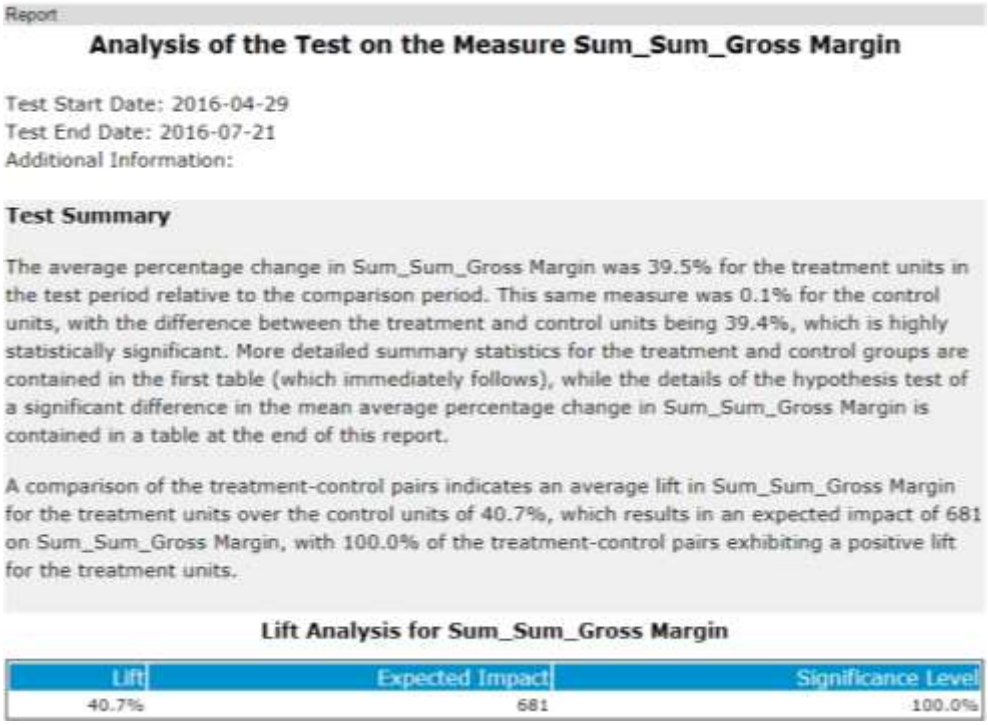
The average percentage change in Sum_Sum_Gross Margin was 39.2% for the treatment units in the test period relative to the comparison period. This same measure was 1.9% for the control units, with the difference between the treatment and control units being 37.2%, which is highly statistically significant. 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 Sum_Sum_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 Sum_Sum_Gross Margin for the treatment units over the control units of 37.9%, which results in an expected impact of 527 on Sum_Sum_Gross Margin, with 100.0% of the treatment-control pairs exhibiting a positive lift for the treatment units.

Lift Analysis for Sum_Sum_Gross Margin

Lift	Expected Impact	Significance Level
37.9%	527	99.5%

3. What is the lift from the new menu overall?
   40.7% based on report below:



Report

Analysis of the Test on the Measure Sum_Sum_Gross Margin

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

Test Summary

The average percentage change in Sum_Sum_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 Sum_Sum_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 Sum_Sum_Gross Margin for the treatment units over the control units of 40.7%, which results in an expected impact of 681 on Sum_Sum_Gross Margin, with 100.0% of the treatment-control pairs exhibiting a positive lift for the treatment units.

Lift Analysis for Sum_Sum_Gross Margin

Lift	Expected Impact	Significance Level
40.7%	681	100.0%

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.