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?
   In our analysis we used gross margin as the final performance measure.

2. What is the test period?
   The test period was from April 29, 2016 to July 21, 2016.

3. At what level (day, week, month, etc.) should the data be aggregated?
   The timeframe aggregation was performed at the week.

: Awesome: Excellent - all three questions were answered correctly.

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.

: : Suggestion: Using the Trend tool

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.

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.
   Trend, seasonality, square foot, and average monthly sales were used as variables to find control stores.

2. What is the correlation between your each potential control variable and your performance metric?
   Because gross margin is a sales minus the cost of goods and services, there is a strong correlation between the two.

3. What control variables will you use to match treatment and control stores?
   Trend, seasonality, square foot.

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

: Required: Please note that these besides trend and seasonality two variables need to be CONSIDERED at first as a control variable. To identify if they are good control variables we need to check their correlation with Gross Margin. I highly suggest including a correlation matrix heatmap for this. Then we need to justify why we chose to include or exclude either of them

: Required: Please include the correlation between AvgMonthlySales and Sq_ft with gross margin.

: Required: This variables is very low correlated with gross margin. Should we include it in this case?

Control	Treatment	Distance
6992	1664	0.278498966658067
7484	1664	0.282654466238954
1508	1675	0.319952331793155
7334	1675	0.456745403355615
7534	1696	0.245612639602244
7284	1696	0.275659003714054
1662	1700	0.37668445257404
7384	1700	0.652851195699666
7434	1712	0.144873144001897
7584	1712	0.267195888664056
11318	2288	0.775704885211254
8817	2288	0.950764607903188
11468	2293	0.412116480644685
8967	2293	0.653774396413601
2383	2301	0.536099237304478
11368	2301	0.693676802150037
11668	2322	0.481369837403396
3002	2322	0.658581665413576
2333	2341	0.230723446338339

: Awesome: For each treatment store there are 2 control stores.

9017 2341 0.438790005436986

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?
   Round Roasters should roll out the expanded menu of wine and sandwiches to more stores.

2. What is the lift from the new menu for West and Central regions (include statistical significance)?
   For the treatment stores, the lift was 39.5 %.

3. What is the lift from the new menu overall?
   In the test stores, within the central and west regions, we saw an over lift of 42.5%. With a 100% significance level.

: Awesome; The recommendation is correct - well done.

: Required: Please include the lifts for the Central and the West region along with the statistical significance.

: Awesome: The overall lift is in the accepted range. Well done!

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.