Predictive Analytics for Business Nanodegree

Project: Analyzing a Market Test

Latifa M.Alyaeesh

Misk & Udacity

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.

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?**

We need to make decision in order to create a new launch a new menu or not. In addition, we need to evaluate whether the new menu can increase enough sales. 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 represented in the *gross_margin* variable.

2. **What is the test period?**

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

Hint: we need total 76 weeks of data which includes 12 weeks of test period+12 weeks of the compared period+52 weeks \ 1 year of data prior the test.

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

The data should aggregated by Store ID , Number of weeks , Store Traffic , Gross Margin and Sales 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.

By using raw data, I created three data files:

- Weekly Store Traffic
- Store Sales Analysis Data
- Store list

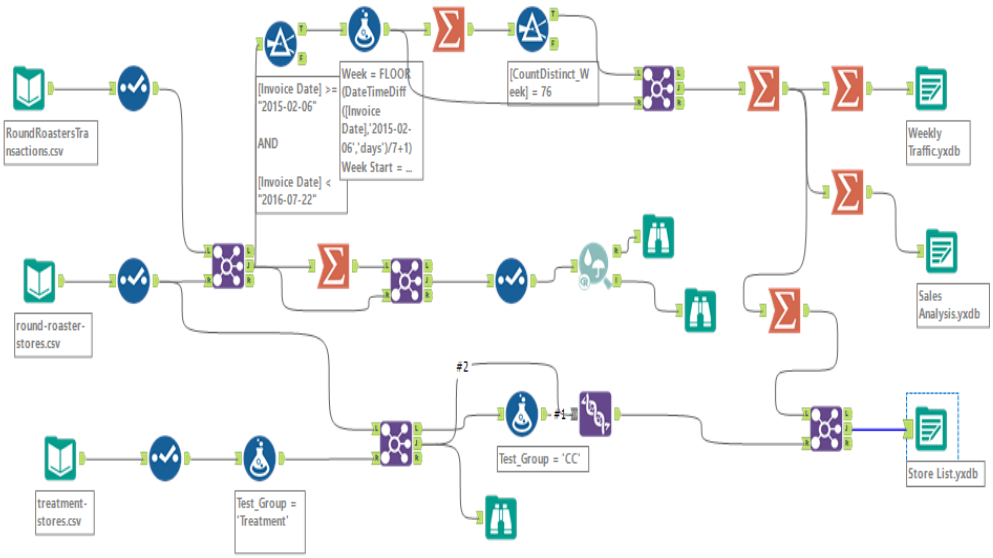

Figure 4: Clean up data workflow

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 square footage of the store and the average monthly sales for the store considered as the control variables in our test.

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

The correlation of AvgMonthSales & Ft on gross margin is AvgMonthSales=0.85 and Sq_Ft =-0.086

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

Based on the correlation coefficient – AvgMonthSales will be used to match treatment and control store, in addition, AvgMonthSales is statistically significant because the P-value <0.05

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	12219	9524
2301	3102	9238
2322	3235	2409
2341	12536	2383

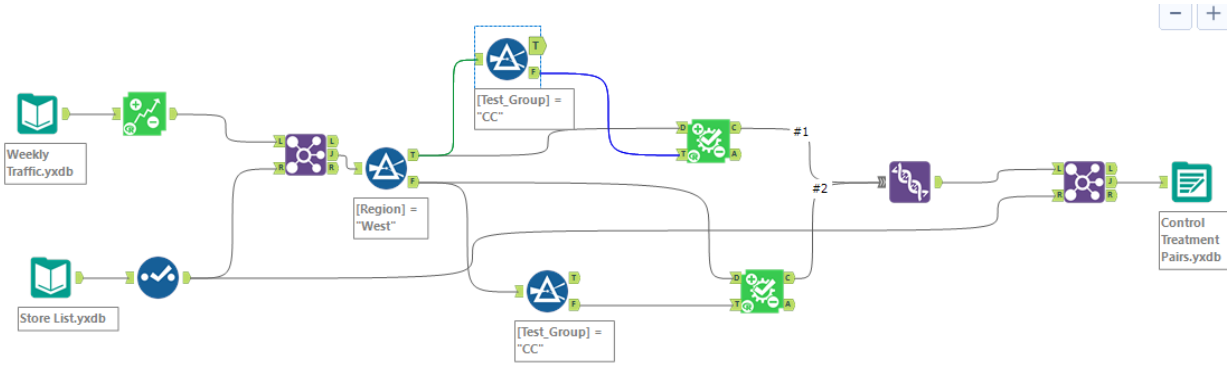

Figure 5: Control Treatment Pairs

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, the company should roll out update to all the stores as the average lift for both the regions was well above 18% that justifies the increase in the marketing budget.

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

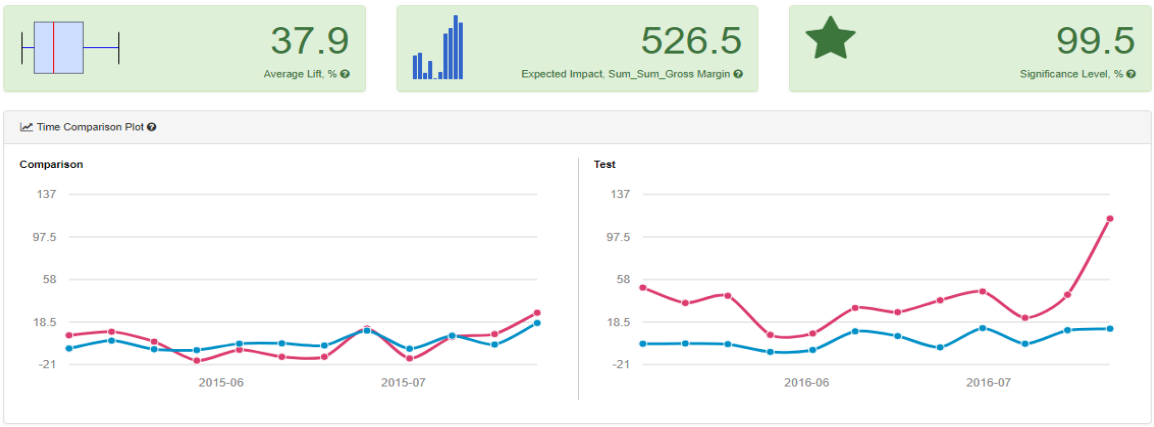

Figure 6: AB test Analysis for West Region

The report shows that after rolling out the new menu for West Region, showed 37.9% improvement at a significance of 99.5% over the existing menu. The average lift as a result from changing the new menu would be 37.9% per store per week, or approximately $526.5 per store per week.

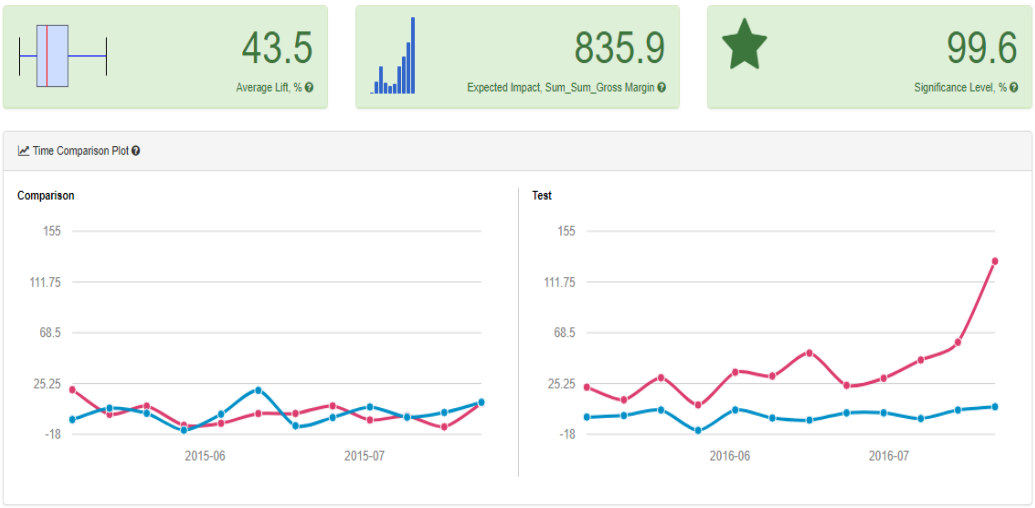

Figure 7: AB test Analysis for Central Region

The report shows that after rolling out the new menu for West Region, showed 43.5% improvement at a significance of 99.6% over the existing menu. The average lift as a result from changing the new menu would be 43.5% per store per week, or approximately $835.9 per store per week.

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

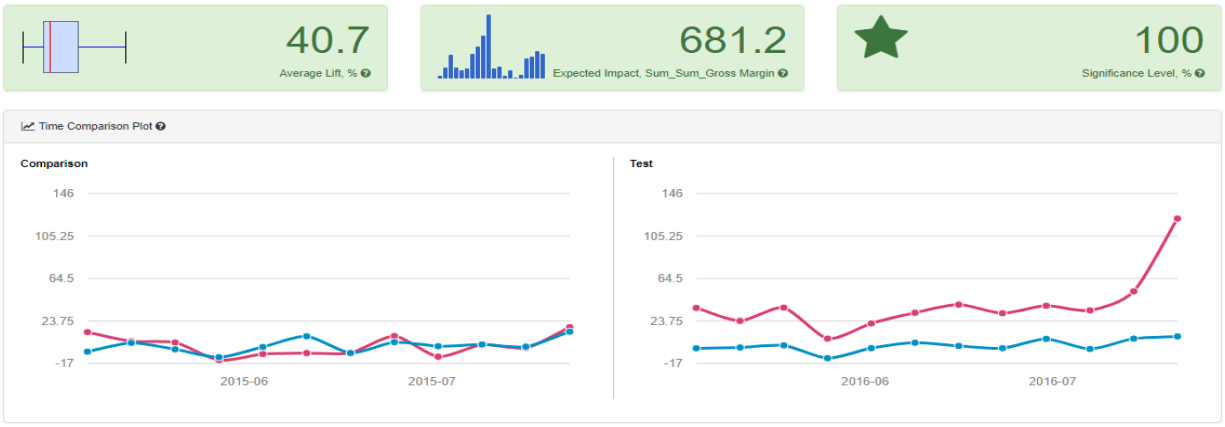

Figure 8: AB test Analysis for overall

The report shows that after rolling out the new menu, showed 40% improvement at a significance of 100% over the existing menu. The average lift as a result from changing the new menu would be 40% per store per week, or approximately $681.2 per store per week. Rolling out the new menu would improve gross margin, therefore the change should rolled out across all stores.

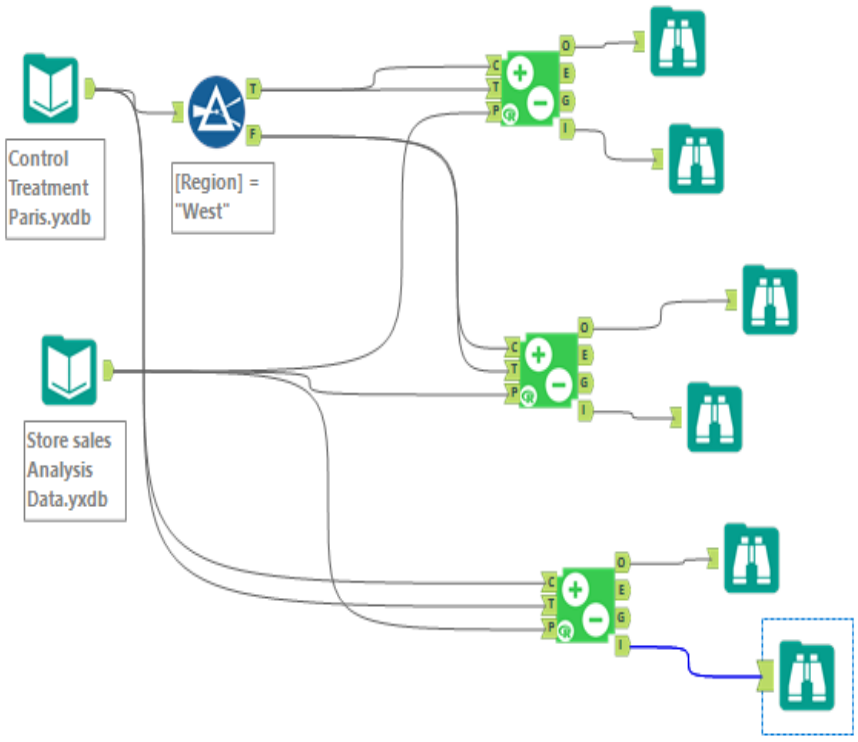

Figure 9: AB Test Analysis workflow