Analyzing a Market Test with A/B Testing

Planned Analysis

The goal of the project is to determine if the introduction of a new menu will increase the amount of profit. This is accomplished through running a matched pairs A/B test of 10 locations located in two different regions of Central and West. The data will need to be prepared for analysis, the treatment stores matched to control stores, and then finally analysis will need to be run. If the increase in profit is greater than 18% we should roll out the new menu to all locations.

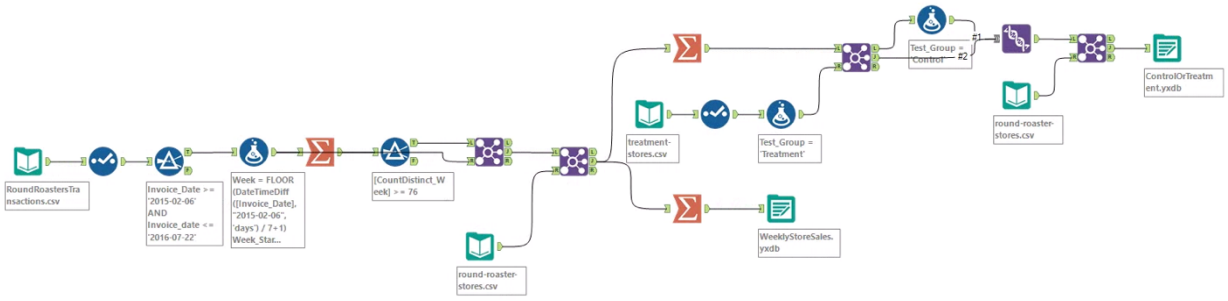
For this project our performance metric is gross margin as we want to see what yields the greatest profit. Our test period is between April 29th, 2016 – July 21st, 2016 and we will be using 12 periods of weekly data, therefore the data needs to be aggregated weekly. We will need 76 weeks of data with includes the minimum 52 weeks, plus 12 weeks for the historical data, and then an additional 12 weeks for the testing period.

We will be using Region as the level for analysis as at the state level there are too few of units in each category to draw accurate results. Typically having at least 15 units in each group will yield accurate results.

Record #	Region	State	Count
1	Central	IA	2
2	Central	IL	15
3	Central	KS	2
4	Central	MN	4
5	Central	MO	4
6	Central	TX	12
7	Central	WI	3
8	West	AZ	10
9	West	CA	42
10	West	CO	15
11	West	ID	2
12	West	NM	2
13	West	NV	2
14	West	OR	8
15	West	WA	10

Record #	Region	Count
1	Central	42
2	West	91

Data Clean Up

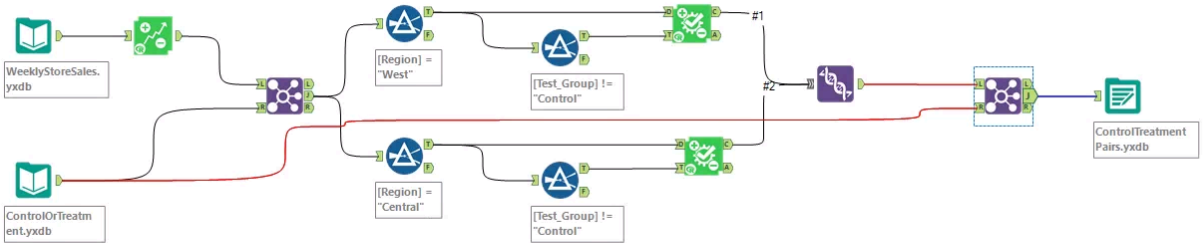


In order to perform this analysis, the data will need to be prepared by aggregating the data and retrieving the necessary data. For this analysis, we will need 2 sets of data before we can match the control and treatment pairs.

ControlOrTreatment – This file contains the field for each store and if it is in the control or treatment group. Also contained in this file is the additional control variable of AvgMonthSales.

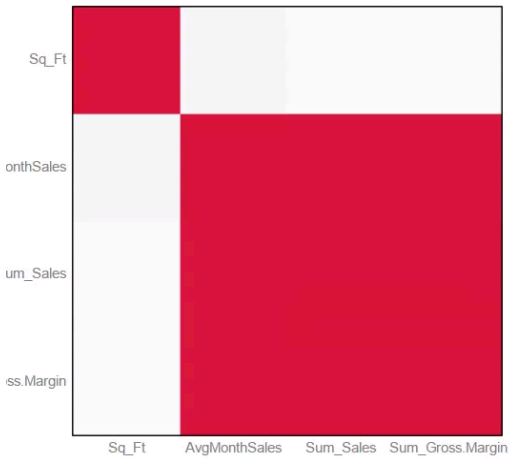
WeeklyStoreSales – This file contains transactional data on each store aggregated at the weekly level. The transactional data included is shown in number of transactions, gross margin, and total sales.

Match Treatment and Control Units



Here we matched treatment units to control units to form pairs, a number of factors needed to accounted for to ensure the highest probability of accuracy. We gain two of our control variables by using a AB trend analysis on the weekly store sales.

The first goal was to determine which control variables to account for. These variables once identified would need to be checked against the performance metric for correlation. In this project the potential control variables were both square feet (Sq_Ft) and average monthly sales (AvgMonthSales). The included images also show sales (Sum_Sales), but this variable is the same as our performance metric and therefore the correlation is expected and should not be included as a control variable.



Full Correlation Matrix

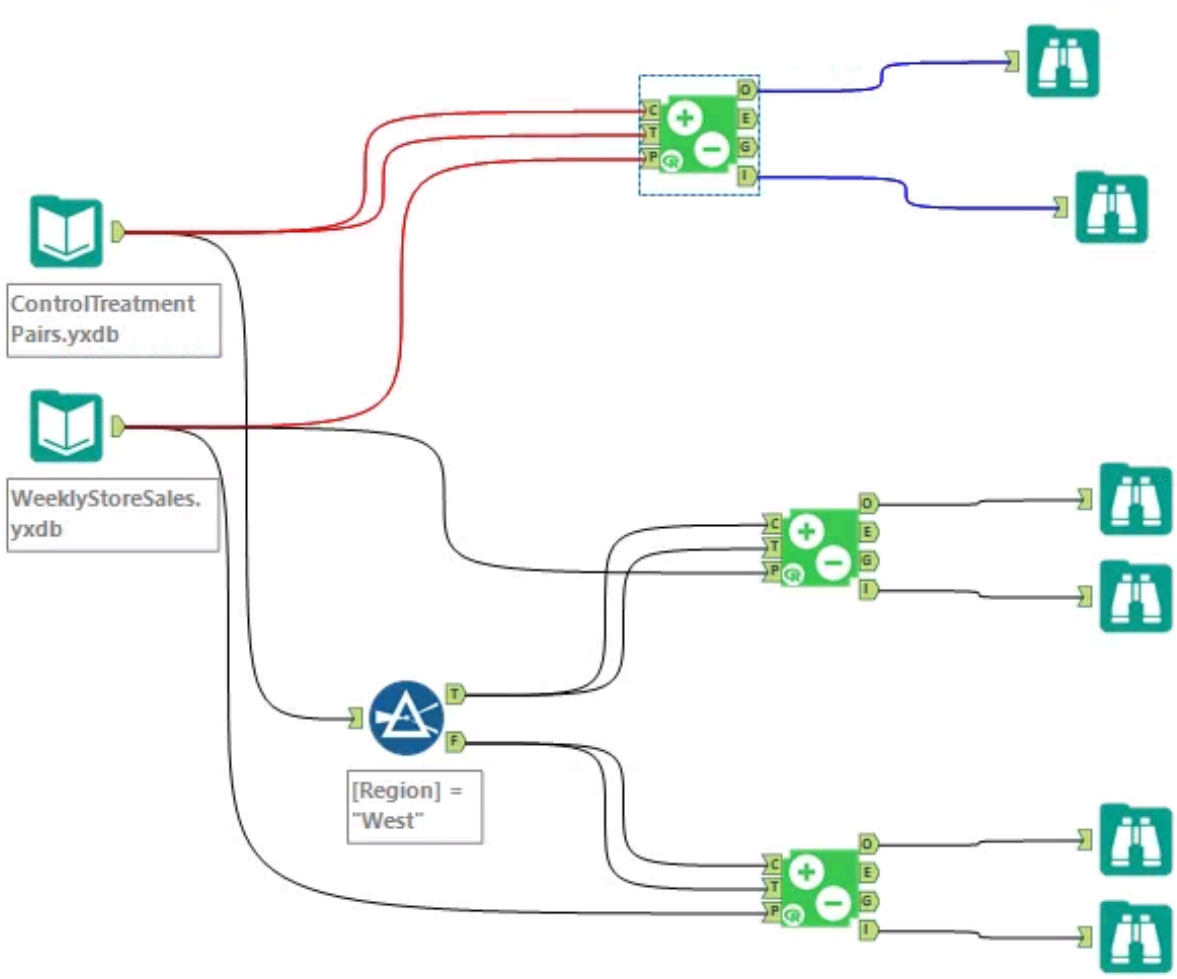
	Sq_Ft	AvgMonthSales	Sum_Gross.Margin	Sum_Sales
Sq_Ft	1.000000	-0.046967	-0.020353	-0.023430
AvgMonthSales	-0.046967	1.000000	0.988219	0.987602
Sum_Gross.Margin	-0.020353	0.988219	1.000000	0.999339
Sum_Sales	-0.023430	0.987602	0.999339	1.000000

After running an association analysis between our potential control variables and the performance metric, we will now include average monthly sales (AvgMonthSales) as a control variable and remove square feet (Sq_Ft) as a control variable. Any variable with a correlation above .7 is considered “high” and should be included as a control variable. Average monthly sales (AvgMonthSales) has a correlation of 0.99 and will be included as a control variable and square feet (Sq_Ft) has a correlation of -0.02 and therefore will not be included as a control variable.

After adjusting for region, by ensuring the treatment units were paired with same region control units, we the use our three control variables (trend, seasonality, AvgMonthSales) to determine which control units should be paired with the treatment units. After running an AB Control analysis, the following pairs were matched:

Treatment Store	Control Store 1	Control Store 2
1664	1964	7162
1675	2214	7284
1696	1863	7334
1700	2014	7037
1712	7434	8162
2288	2568	9081
2293	9639	12686
2301	9238	12536
2322	3185	9388
2341	2572	12586

Analysis and Writeup



Three instances of the analysis are performed, one on each region (West and Central) and one on all units.

Our results for West show a lift of 39.1% at a significance of 99.6%

Our results for Central show a lift of 47.6% at a significance of 99.6%

Our results overall show a lift of 43.4% at a significance of 100%

Based on these results the company should roll out the new menu across all of their stores.