

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

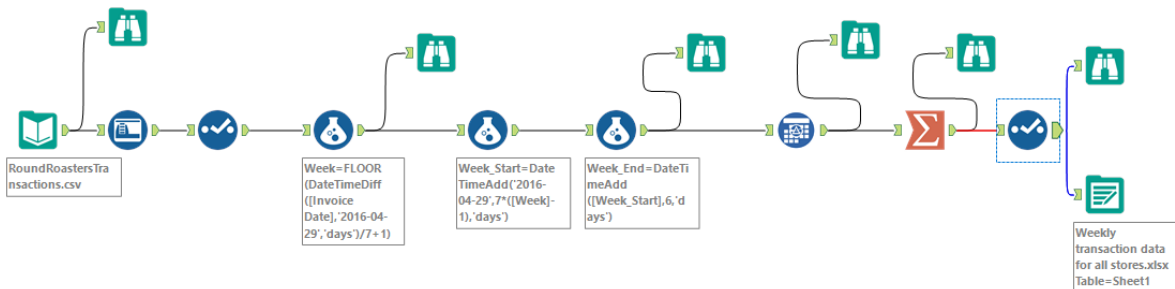
The metric to be used for the results evaluation is profit, and more specifically profit growth. This is represented by the gross_margin variable in our dataset. The test period for our experiment is from 2016-April-29 to 2016-July-21, a 12-week period. The data should be aggregated at a week level since the experiment time is given and measured in weeks.

Step 2: Clean Up Your Data

The data set was prepared for analysis, aggregating the transaction data for each store on a week level. Given that the experiment start date is the start date of Week 1, week numbers are ranging from -63 to 12, i.e. 52 + 12 weeks of data for the analysis as specified plus 12 weeks of experiment duration. These 76 weeks span from 06/02/2017 to 21/072016.

Fields: StoreID, Week, Week Start, Week End, No. of Invoices, Sum Sales, Sum Gross Margin

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.

In order to match treatment and control units, first we need to decide on which variables should be used in the analysis as the control variables. Given that we have transaction data for the whole 76 weeks, Trend and Seasonality should be calculated and used as control variables.

However, there are other candidate control variables in the RoundRoastersStore file. Our performance metric is Gross_Margin (numerical) and the only hence the Sq_Ft and AvgMonthSales variables should be investigated, in terms of correlation to the Gross_Margin. Using the Assosiation Analysis tool on Alteryx, we get the following correlation matrix.

Pearson Correlation Analysis

Focused Analysis on Field Sum_Sum_Gross.Margin

	Association Measure	p-value	
AvgMonthSales	0.990982	0.00000	***
Sq_Ft	-0.024255	0.78168	

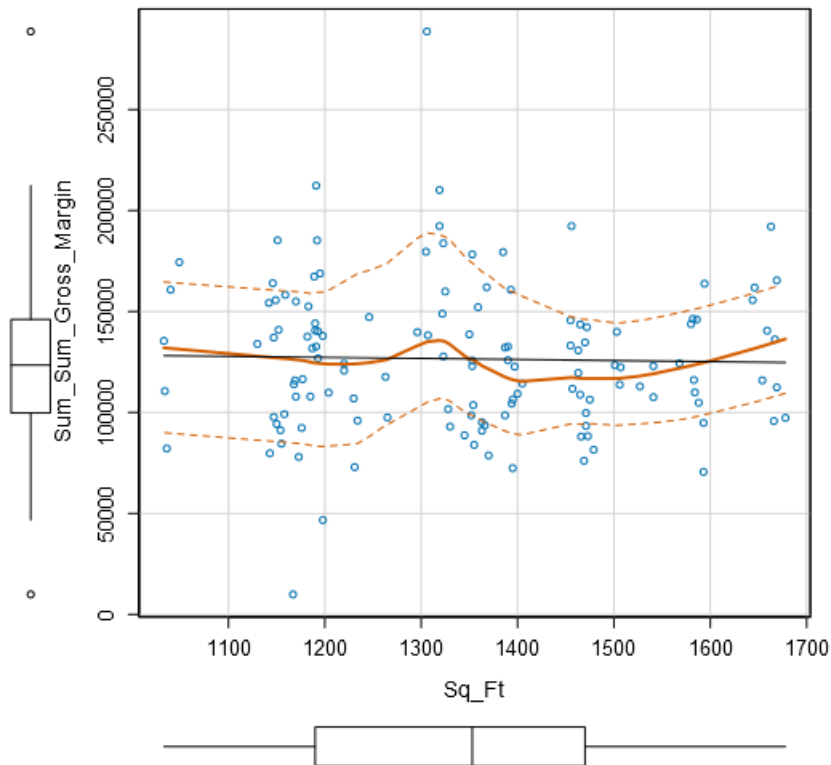
Full Correlation Matrix

	Sum_Sum_Gross.Margin	Sq_Ft	AvgMonthSales
Sum_Sum_Gross.Margin	1.000000	-0.024255	0.990982
Sq_Ft	-0.024255	1.000000	-0.046967
AvgMonthSales	0.990982	-0.046967	1.000000

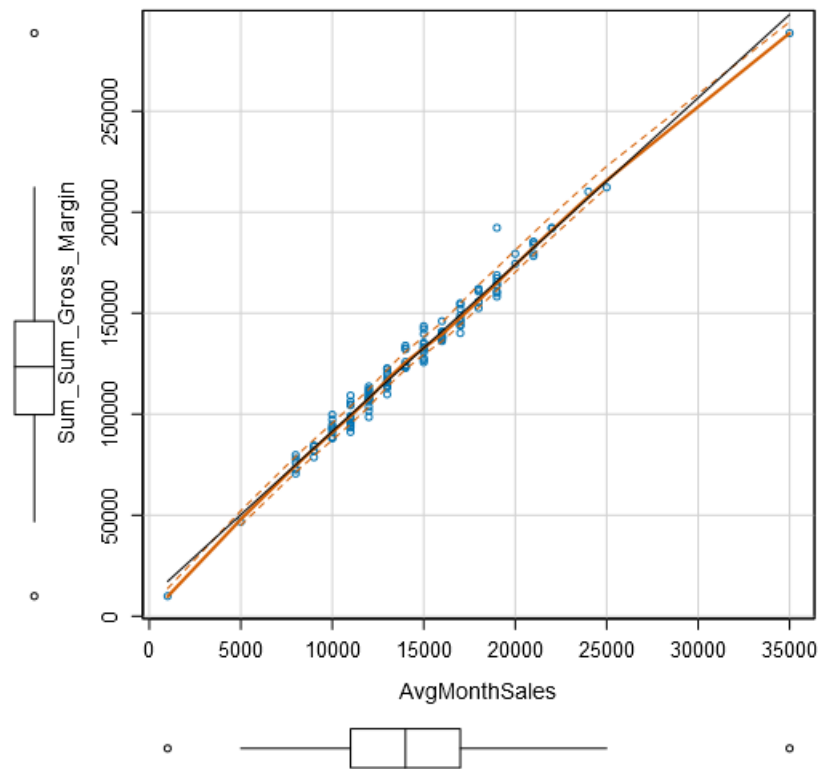
Matrix of Corresponding p-values

	Sum_Sum_Gross.Margin	Sq_Ft	AvgMonthSales
Sum_Sum_Gross.Margin		0.78168	0.00000
Sq_Ft	0.78168		0.59138
AvgMonthSales	0.00000	0.59138	

Scatterplot of Sq_Ft versus Sum_Sum_Gross_Margin



catterplot of AvgMonthSales versus Sum_Sum_Gross_M



It appears that only the AvgMonthSales variable has a strong correlation to our performance measure where as Sq_Ft does not correlate at all. More specifically, we see a 0.99 correlation and a 0.0000 p-value together with a *** statistical significance for the AvgMonthSales variable. This suggests a perfectly linear relationship between AvgMonthSales and Gross_Margin, as can be verified by looking at the scatterplot above. To sum up, we are keeping AvgMonth Sales and discarding Sq_Ft as control variables.

Calculation of Trend and Seasonality

Using the aggregated store data file and the AB Trend tool on Alteryx we can calculate our next two control variables. The configuration of the tool is shown below.

Configuration - AB Trend

Input Data

Select the unit identifier
StoreID

Select the field with the reporting period dates
Week_End

Select the performance measure to use
CountDistinct_Invoice Number

Date Values

Report Period Type
Weekly

Number of periods to calculate the trend.
12

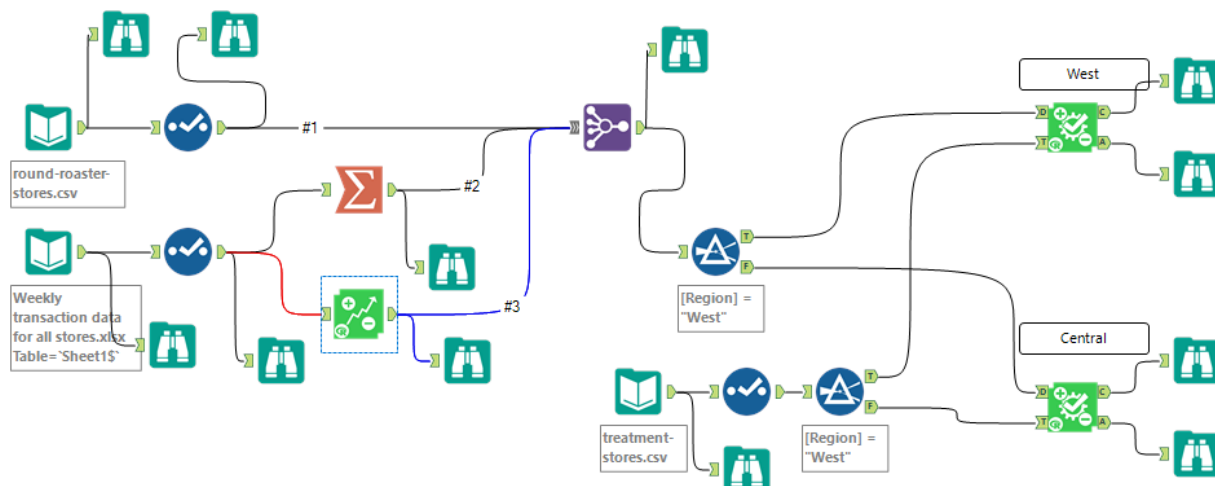
Test Start Date

April 2016						
Mon	Tue	Wed	Thu	Fri	Sat	Sun
28	29	30	31	1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	1
2	3	4	5	6	7	8

Today: 09/05/2017

Note: The input data must have data for at least a year plus the number of periods prior to this date, for example for weekly data with 2 periods selected, you need 54 prior records in your data stream.

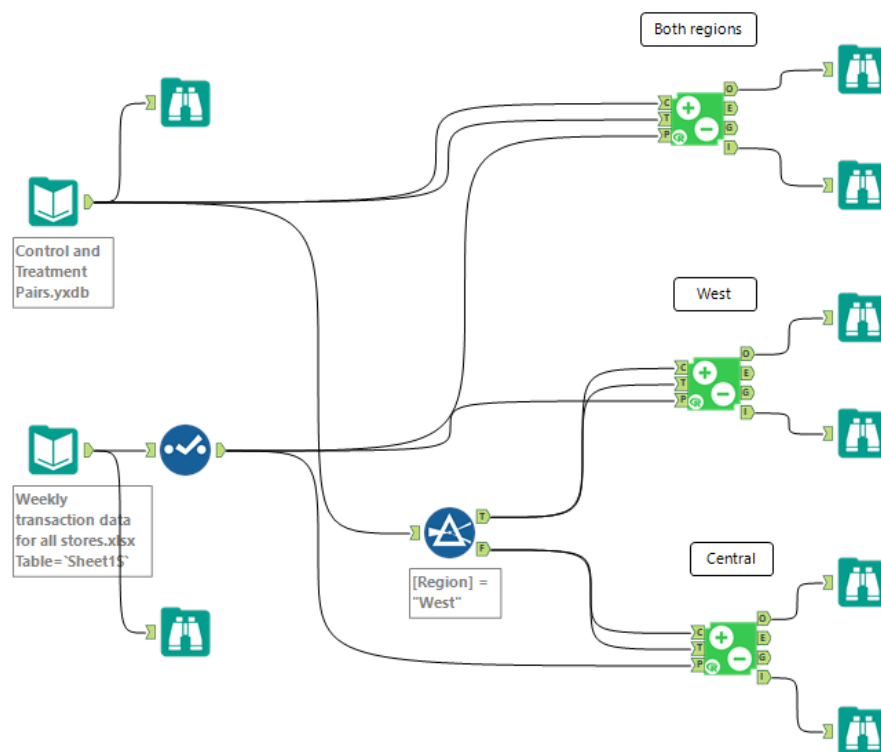
Having chosen our three control variables (AbgMonthSales, Trend, Seasonality), we are ready to match each treatment unit with two control units. This will be a separate workflow for each region, i.e. West and Central. The workflow to get to the matchings table is shown below:



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

Step 4: Analysis and Writeup

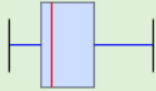
The business requires that three separate AB analyses are done, i.e. for both regions, for the West region, and for the Central region. Using the AB Analysis tool on Alteryx :



Results for Both Regions

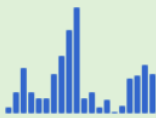
AB Test Analysis for Sum_Gross Margin

Time: 2017-05-09 16:43:25



41.9

Average Lift, %



693.8

Expected Impact, Sum_Gross Margin

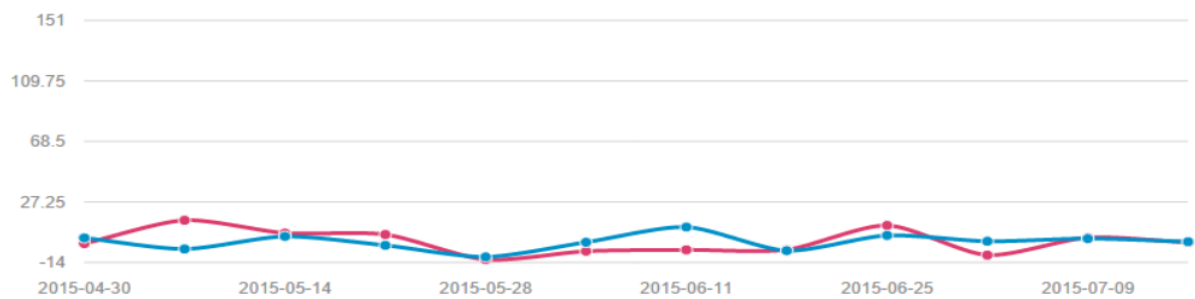


100

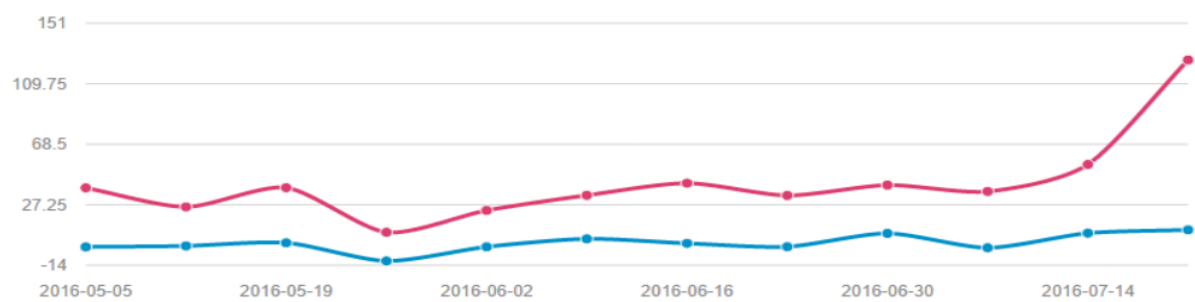
Significance Level, %

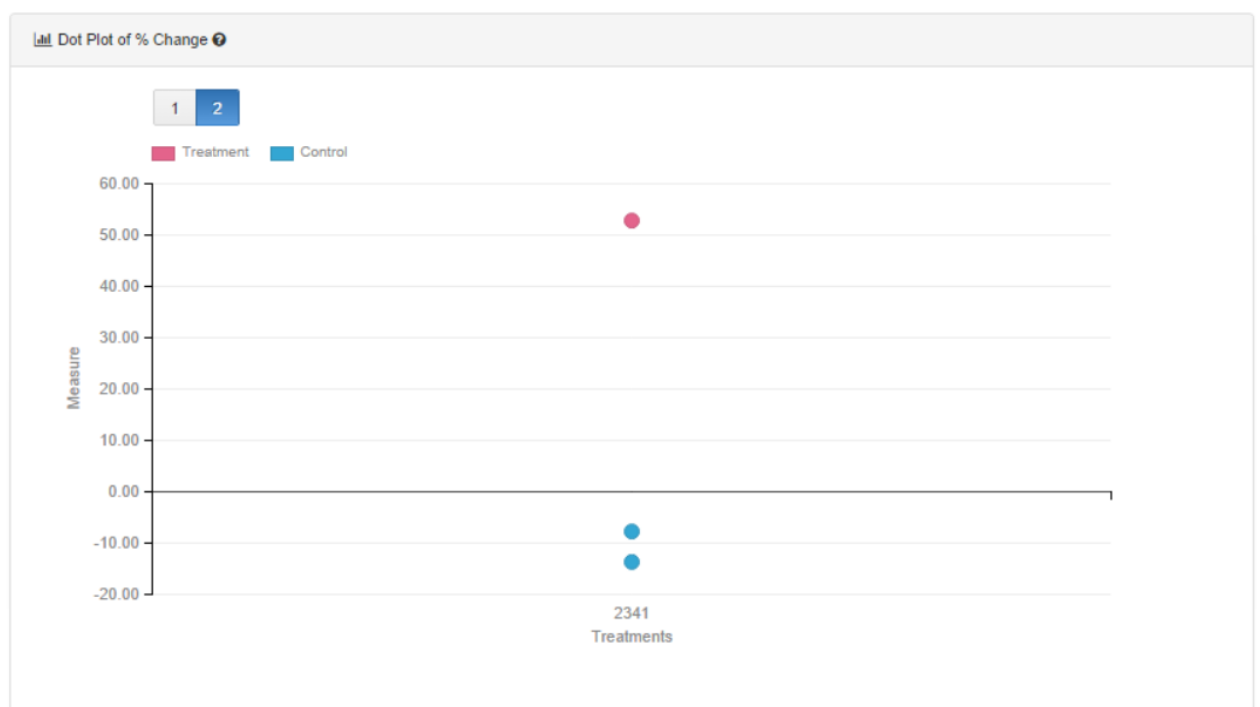
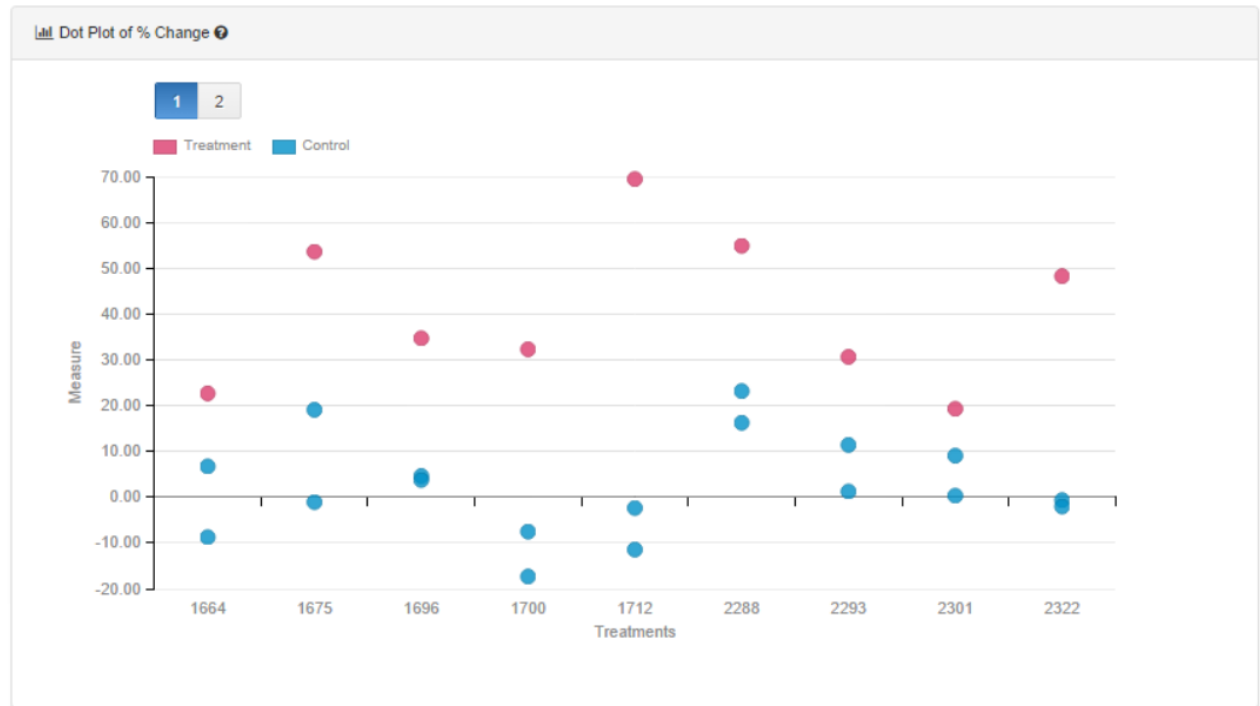
Time Comparison Plot

Comparison



Test



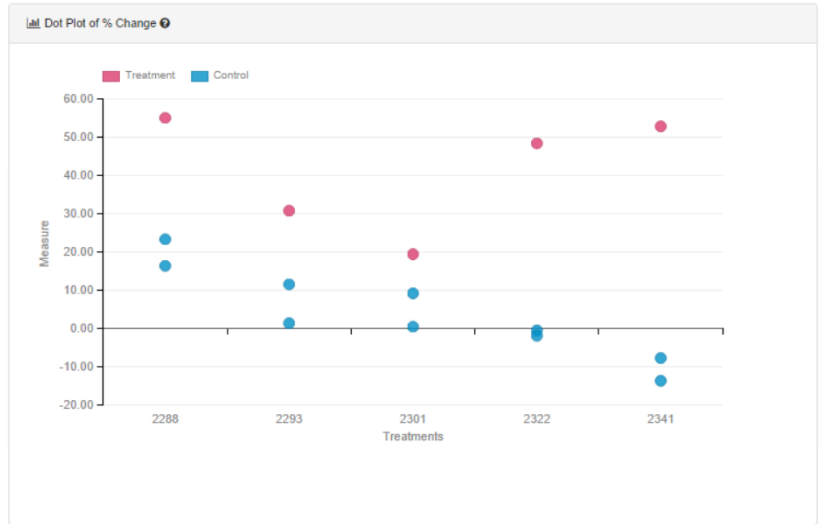
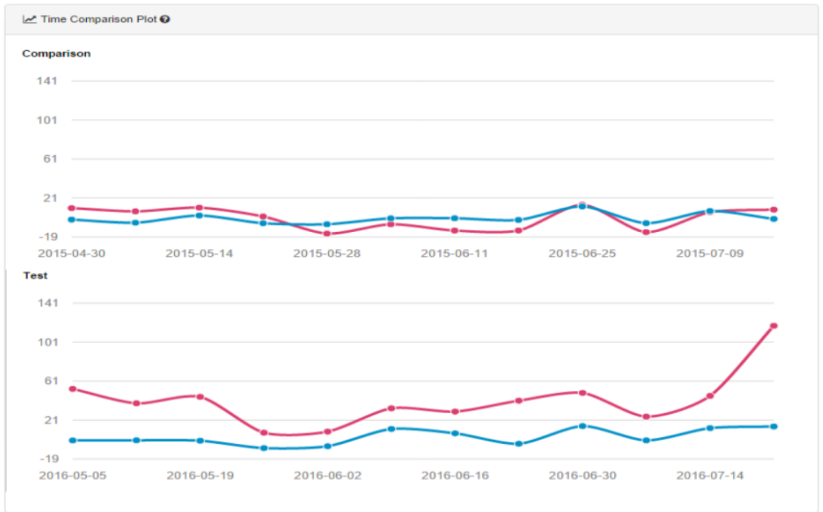
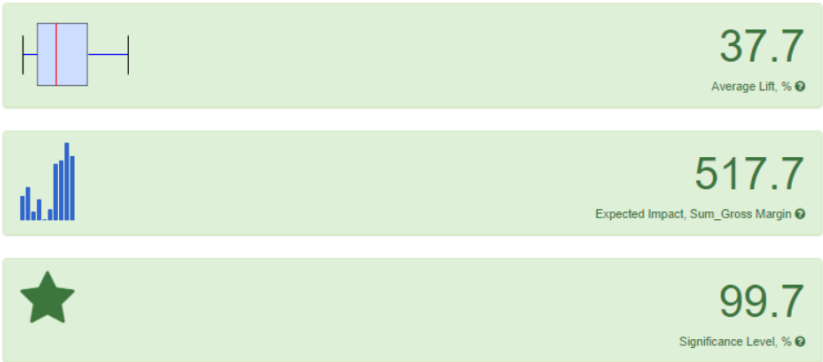


The analysis for both regions suggests that the introduction of the new items in the menu will have a positive impact on the company profits (Gross_Margin). More specifically, we note a 41.9% lift on average, which means that it would cause customers to spend \$693.8 per week more on average. The t-test returns a significance level of 100% which is strongly supporting the move to introduce the new items to the menu. Note: all treatment units are performing better than the control units.

Results for the West Region

AB Test Analysis for Sum_Gross Margin

Time: 2017-05-09 16:40:28

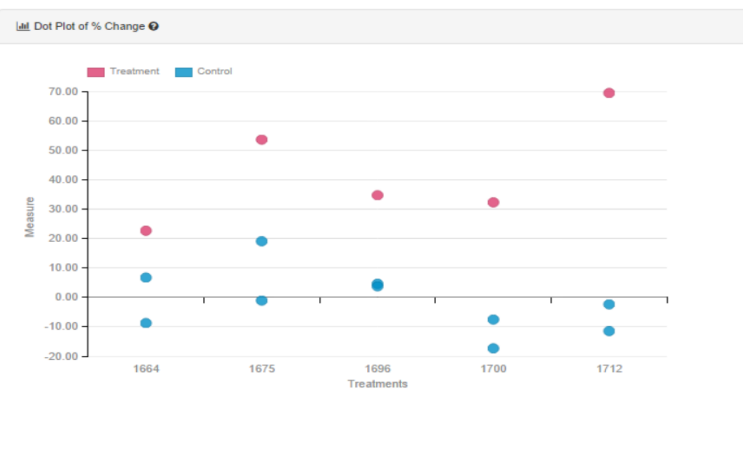
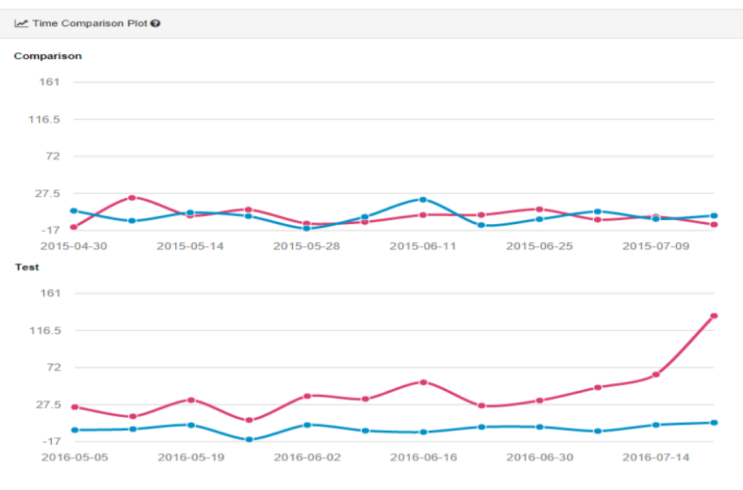
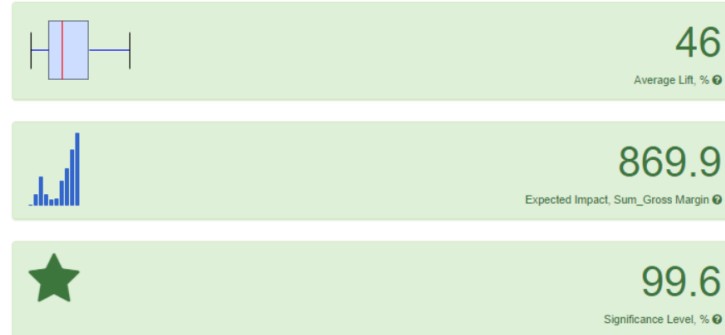


The analysis for the West region also suggests that the new items should be introduced to the menu. There is a 37.7% lift on average, i.e. customers would spend \$517.7 per week more on average. This comes at a 99.7% significance level and along with the fact that all treatment units are performing better than control units, the analysis supports the decision for introducing the new menu in the West region.

Results for the Central region

AB Test Analysis for Sum_Gross Margin

Time: 2017-05-09 16:34:03



Finally, the results are also encouraging for the Central region. This means that:

- We see a 46% lift on average
- Customers would spend \$869.9 more per week on average
- T-test returns a 99.6% significance level
- All treatment units are performing better than control units

Based on the above analyses, it is our recommendation that the company goes forward with launching the new menu in both regions and all stores as it would significantly increase profits.