Project: Analyzing a Market Test

Step 1: Plan Your Analysis

1. What is the performance metric you'll use to evaluate the results of your test?

   For this problem, I would use profit - Gross Margin in the dataset as my performance metric.

2. What is the test period?

   The test would run for 12 weeks, from April 29, 2016 to July 21, 2016.

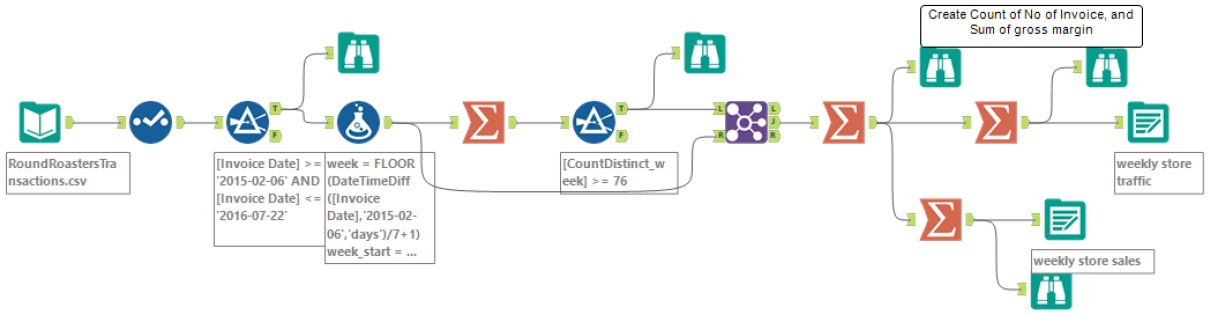
3. At what level (day, week, month, etc.) should the data be aggregated?

   Data should be aggregated at week level, since the test would run on week level.

Step 2: Clean Up Your Data

Summary: I implemented Alteryx and aggregated the transaction data to the week level and filter on the A/B test analysis data ranges. Then I created a store list including information of experiment groups for matching analysis.

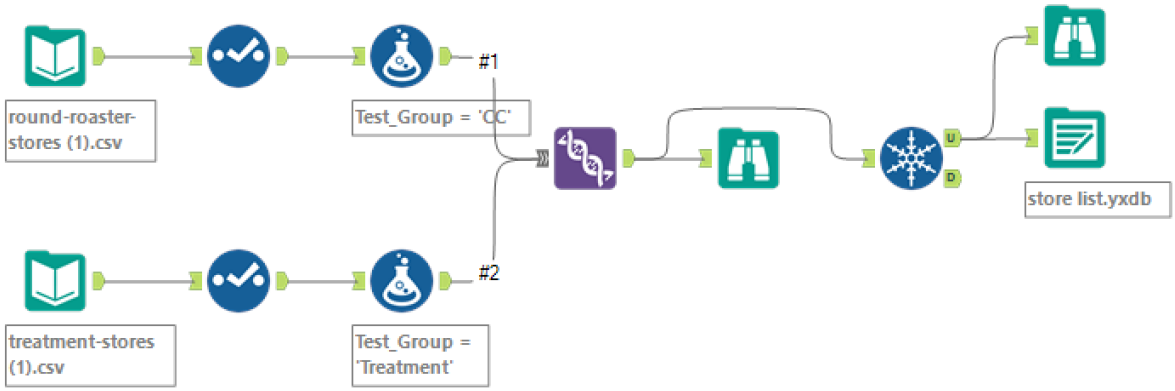
1. First I aggregate store transactions data, to get the sales information for each store and week at an aggregated level. Such as the sum of gross margin for StoreID 10018 at Week 1 is 2112.
   The workflow generate two datasets for use: Weekly store traffic dataset and Weekly store sales dataset.



New datasets example:

StoreID	week	week_start	week_end	Count
10018	1	2015-02-06	2015-02-12	308
10018	2	2015-02-13	2015-02-19	288
10018	3	2015-02-20	2015-02-26	204
10018	4	2015-02-27	2015-03-05	320
10018	5	2015-03-06	2015-03-12	284
10018	6	2015-03-13	2015-03-19	288
10018	7	2015-03-20	2015-03-26	194
10018	8	2015-03-27	2015-04-02	286

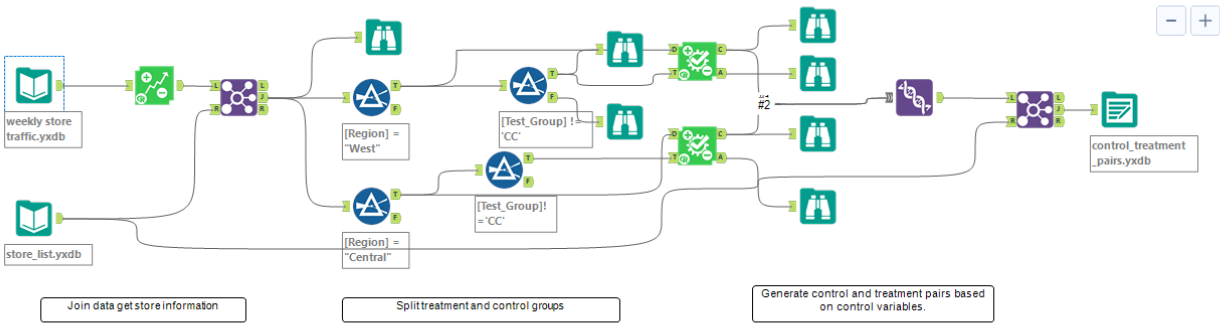
2. To categorize stores into two groups: treatment group and control group for the next matching step.
   This workflow generated one dataset: store_list.



StoreID	AvgMonthSales	Region	Test_Group
10018	18,000	West	CC
10068	16,000	West	CC
10118	13,000	West	CC
10168	19,000	West	CC
10218	15,000	West	CC
10268	25,000	West	CC

Step 3: Match Treatment and Control Units

Alteryx workflow:



1. What control variables should be considered? Note: Only consider variables in the RoundRoastersStore file.

   AvgMonthSales and **Sq_Ft**. From the RoundRoastersStore file, we can see these two variables have the same data type as our performance metric. Besides, the business rationale is very clear, high monthly sales may be correlated with high gross margin, and larger stores with more square feet area may attract more customers, thus correlated with gross margin.

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

   From the correlation analysis below, we can see AvgMonthSales variable is highly correlated with our performance metric, Gross Margin. And the Sq_Ft did not exhibit a relationship with Gross Margin. Therefore, I would choose AvgMonthSales as my control variable in the A/B test.

   Pearson Correlation Matrix Result:

FieldName	Sum_Gross Margin	Sum_Sales	Sq_Ft	AvgMonthSales
Sum_Gross Margin	1.000	0.999	-0.020	0.788
Sum_Sales	0.999	1.000	-0.023	0.786
Sq_Ft	-0.020	-0.023	1.000	-0.047
AvgMonthSales	0.788	0.786	-0.047	1.000

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

   From the above result, we can see **AvgMonthSales** would be the best control variables to match treatment and control stores.

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

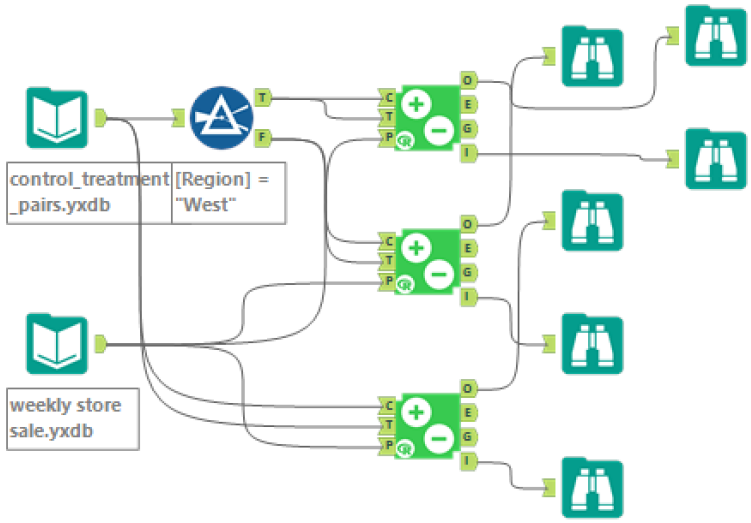
Treatment Store	Control Store 1	Control Store 2
1664	7162	8112

1675	1580	1807
1696	1964	1863
1700	2014	1630
1712	8162	7434
2288	9081	2568
2293	12219	9524
2301	3102	9238
2322	2409	3235
2341	12536	2383

Step 4: Analysis and Writeup

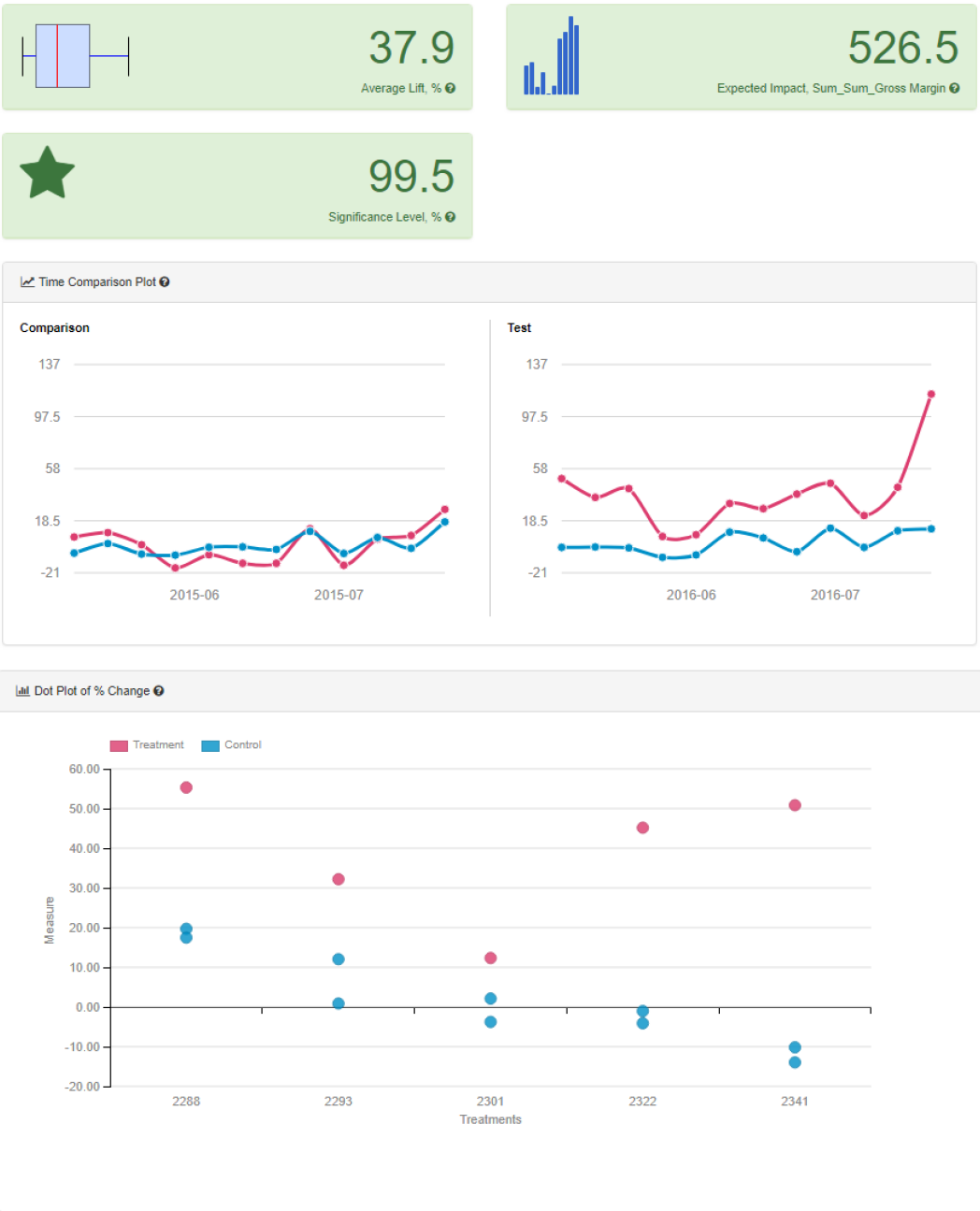
A/B test Result:

- Alteryx workflow

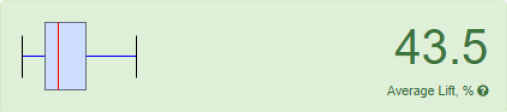


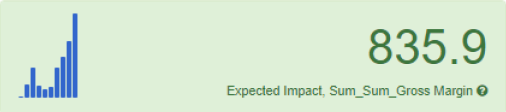
- A/B test analysis:

West Region:

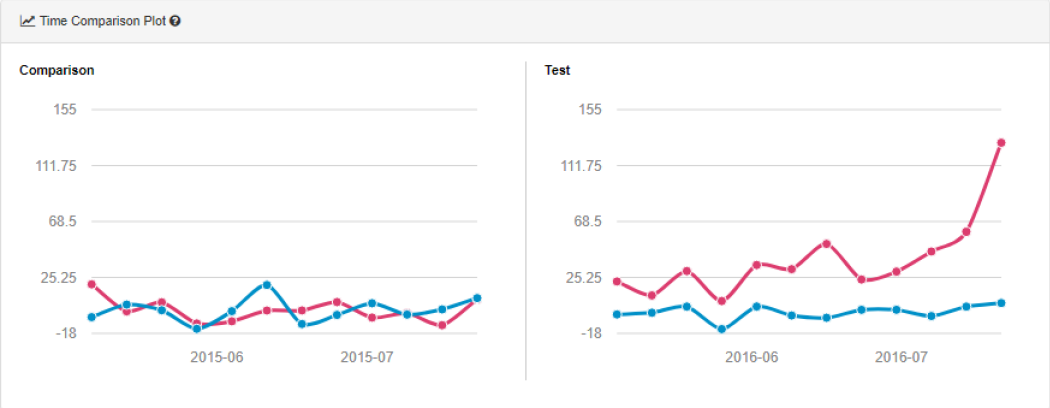


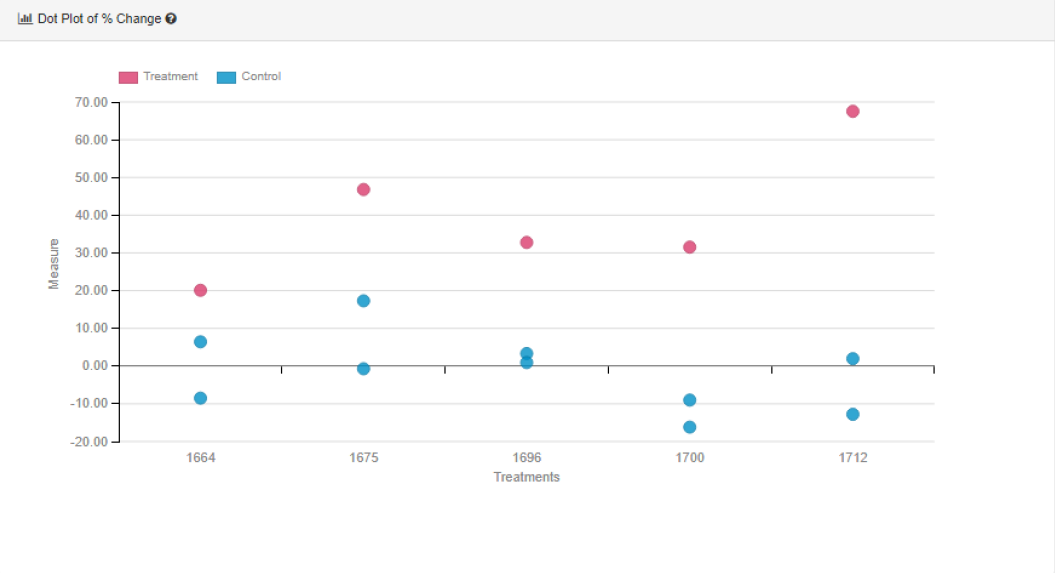
Central Region:


43.5
Average Lift, %

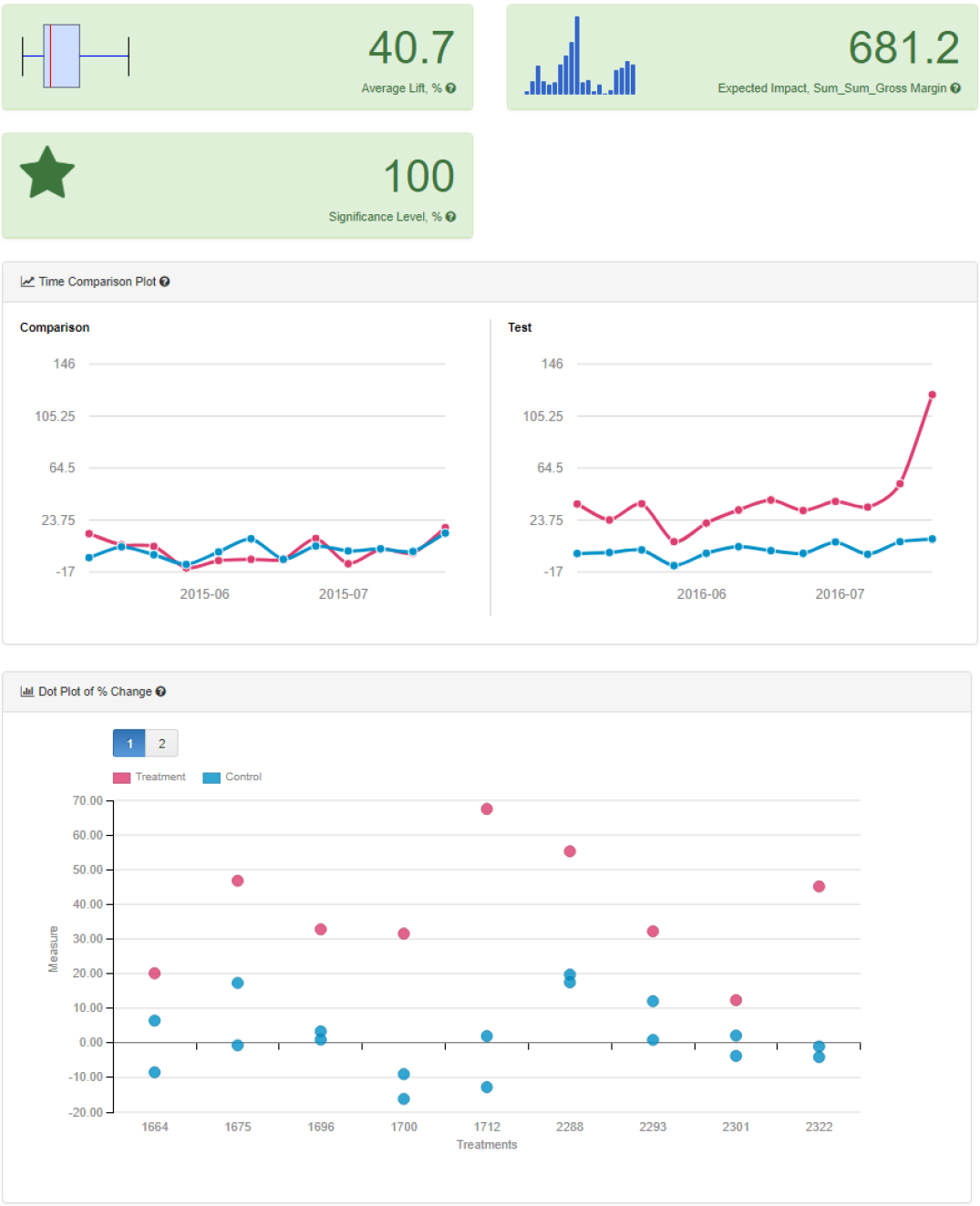

835.9
Expected Impact, Sum_Sum_Gross Margin

99.6
Significance Level, %


Time Comparison Plot
Comparison
155
111.75
68.5
25.25
-18
2015-06
2015-07
Test
155
111.75
68.5
25.25
-18
2016-06
2016-07


Dot Plot of % Change
Treatment
Control
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
-10.00
-20.00
Measure
1664
1675
1696
1700
1712
Treatments

Overall:



1. <u>What is your recommendation - Should the company roll out the updated menu to all stores?</u>

   I would recommend the company roll out the updated menu to all stores. Because in all regions the profit growth is larger than 18%. For the West region, the lift is 37.9%, for the

Central region, the lift is 43.5%, and for Overall region, the lift is 40.7%. And the test results are significant.

2. What is the lift from the new menu for West and Central regions (include statistical significance)?
   For the West region, the lift is 37.9% with significance level 99.5%, for the Central region, the lift is 43.5% with significance level 99.6%.

3. What is the lift from the new menu overall?
   For the Overall, the lift is 40.7% with significance level 100.0%.