

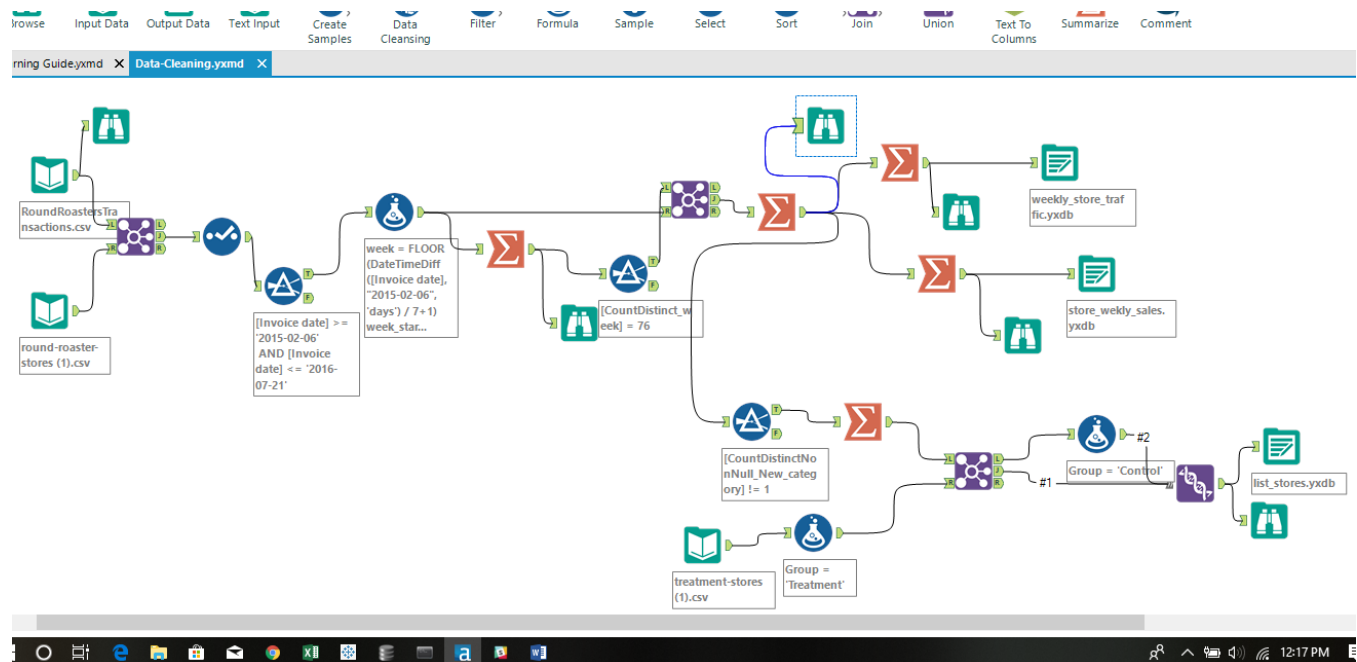
Project: Analyzing a Market Test

Step 1: Plan Your Analysis

1. What is the performance metric you'll use to evaluate the results of your test?
➔ To analyze, whether we should enroll the new menu or not, we will prefer comparing the profits of the two groups. So, here we will use gross margin as the metric to evaluate the results of my test.
2. What is the test period?
➔ Test period is of 12 weeks, i.e., from 29-Apr-2016 to 21-July-2016.
3. At what level (day, week, month, etc.) should the data be aggregated?
➔ The data should be aggregated to the week level, as a week will be enough to show one cycle of customers.

Step 2: Clean Up Your Data

- ➔ Data is extracted from the date 06-02-2015 to 21-07-2016.
- ➔ Week number, week start, week end dates are calculated.
- ➔ Data is verified using the count_distinct tool for week numbers and hence verified that the data is available for 76 weeks.
- ➔ Weekly_stores_transactions and store_weekly_sales database files are created.
- ➔ Data from sandwiches and wines is treated with a new_product flag and the flag is set to be 1.
- ➔ The data with flag 1 is treated as treatment and the rest is treated as control and the output of the join of these two is stored in a database store_lists.



Step 3: Match Treatment and Control Units

Apart from trend and seasonality...

1. What control variables should be considered?
➔ AvgMonthSales and sq.ft be considered as they can affect the gross margin.
2. What is the correlation between your each potential control variable and your performance metric?
➔ AvgMonthSales has a correlation of 0.988 and sq_ft has a correlation of -0.0203 with sum_gross_margin and with the table of p-values we can see that avgmonthsales is having 0 p-values which means it has a very strong relation with sum_gross_margin.

Pearson Correlation Analysis

Focused Analysis on Field Sum_Gross.Margin

	Association Measure	p-value
AvgMonthSales	0.988219	0.00000 ***
Sq_Ft	-0.020353	0.81612

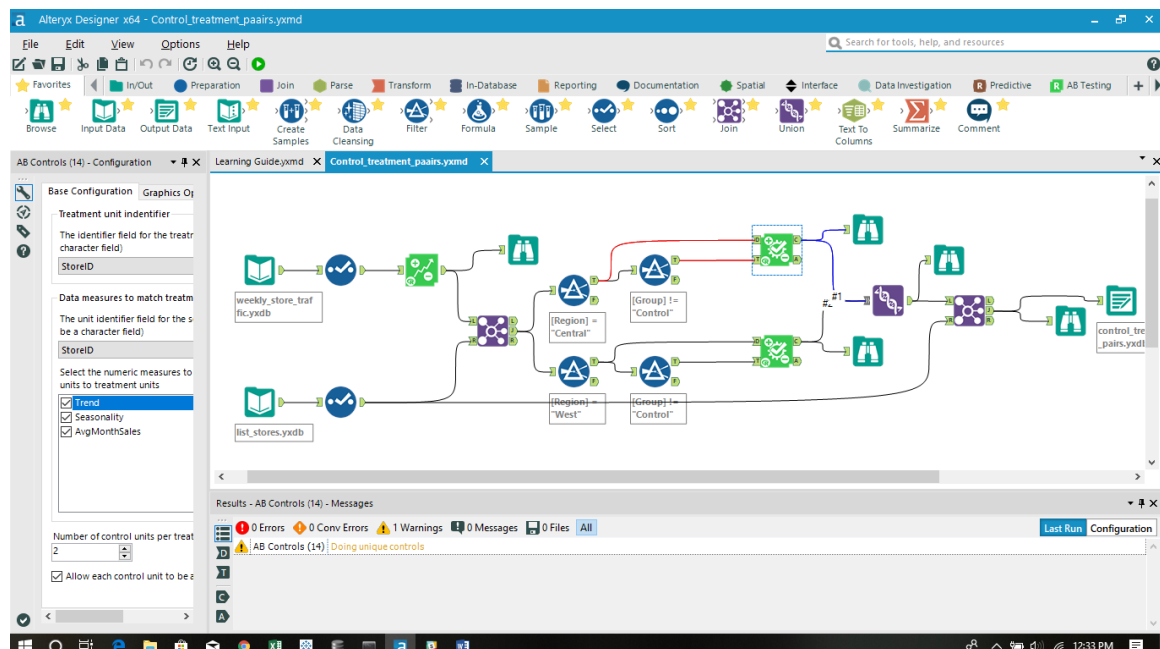
Full Correlation Matrix

	Sum_Gross.Margin	Sq_Ft	AvgMonthSales
Sum_Gross.Margin	1.000000	-0.020353	0.988219
Sq_Ft	-0.020353	1.000000	-0.046967
AvgMonthSales	0.988219	-0.046967	1.000000

Matrix of Corresponding p-values

	Sum_Gross.Margin	Sq_Ft	AvgMonthSales
Sum_Gross.Margin		0.81612	0.00000
Sq_Ft	0.81612		0.59138
AvgMonthSales	0.00000	0.59138	

3. What control variables will you use to match treatment and control stores?
➔ After conclusions from the above table, I will be using only the AvgMonthSales as the control variable to match the 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

Browse (19) - Control_treatment_pairs.xml

20 records displayed, 7 fields, 2457 bytes

Table Profile

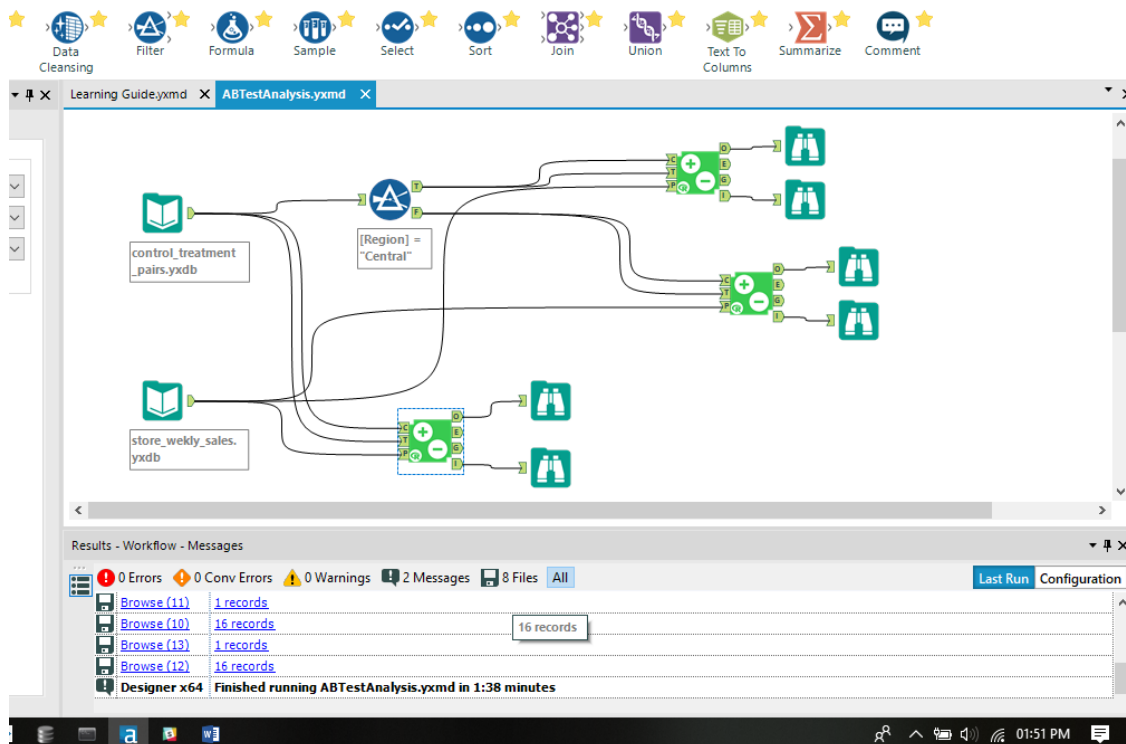
7 of 7 Fields Cell Viewer

Record #	Controls	Treatments	Distance	StoreID	AvgMonthSales	Region	Group
1	7162	1664	0.478595	1664	11000	Central	Treatment
2	8112	1664	1.034443	1664	11000	Central	Treatment
3	1580	1675	0.45634	1675	15000	Central	Treatment
4	1807	1675	0.560454	1675	15000	Central	Treatment
5	1964	1696	0.312367	1696	10000	Central	Treatment
6	1863	1696	0.489137	1696	10000	Central	Treatment
7	2014	1700	0.810402	1700	15000	Central	Treatment
8	1630	1700	0.91618	1700	15000	Central	Treatment
9	8162	1712	0.671441	1712	19000	Central	Treatment
10	7434	1712	0.793269	1712	19000	Central	Treatment
11	9081	2288	0.277932	2288	14000	West	Treatment
12	2568	2288	0.714134	2288	14000	West	Treatment
13	12219	2293	0.348583	2293	11000	West	Treatment
14	9524	2293	0.656038	2293	11000	West	Treatment
15	3102	2301	0.381248	2301	11000	West	Treatment
16	9238	2301	0.434646	2301	11000	West	Treatment
17	2409	2322	0.171431	2322	14000	West	Treatment
18	3235	2322	0.45125	2322	14000	West	Treatment
19	12536	2341	0.39796	2341	11000	West	Treatment
20	2383	2341	0.423792	2341	11000	West	Treatment

Step 4: Analysis and Writeup

Answer these questions. Be sure to include visualizations from your analysis:

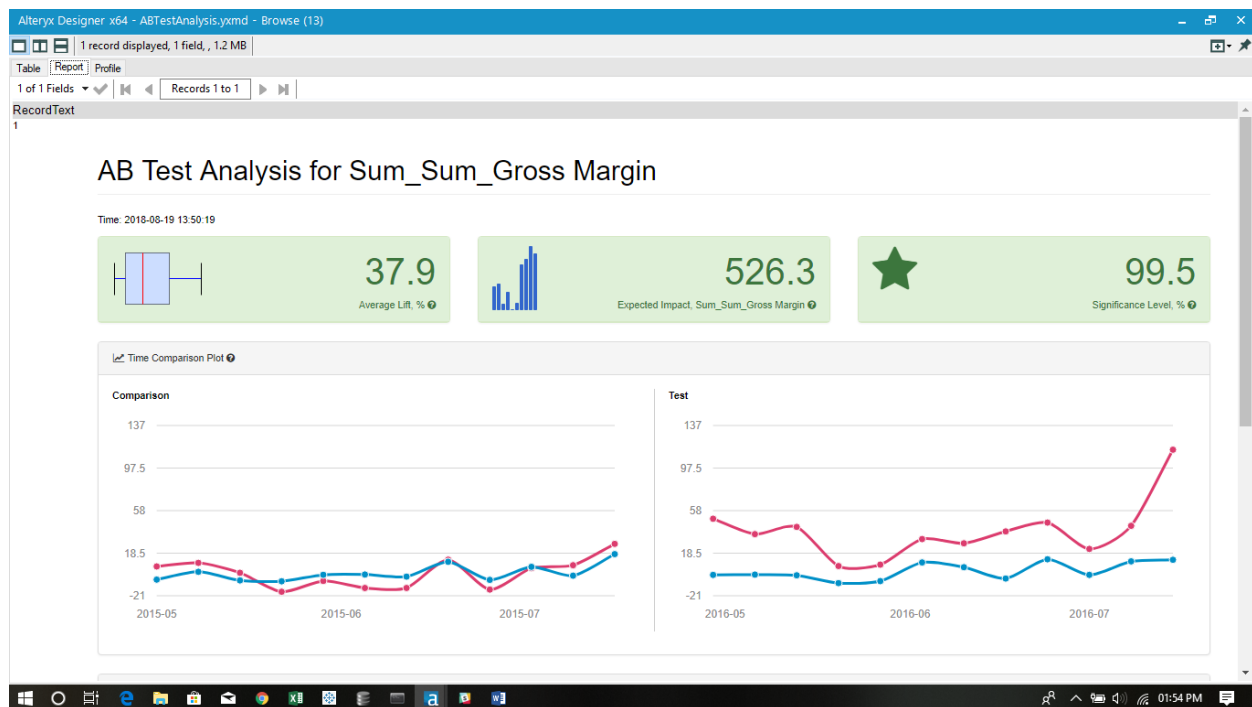
- What is your recommendation - Should the company roll out the updated menu to all stores?
 - ➔ After analyzing the AB-tests, I strongly recommend to enroll out the new menu as It will definitely help the company to increase its profit.



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

→ **For West Region :-**

Average lift is 37.9 % and significance level is 99.5 % .



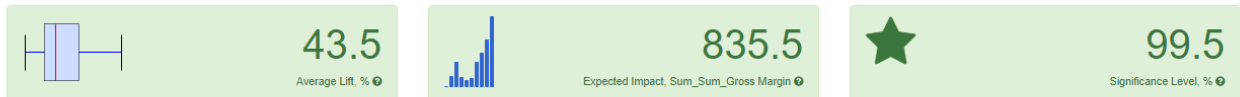
➔ **For Central Region :-**

Average lift is 43.5 % and significance level is 99.5% . Lift % is better than the west region.

2X1

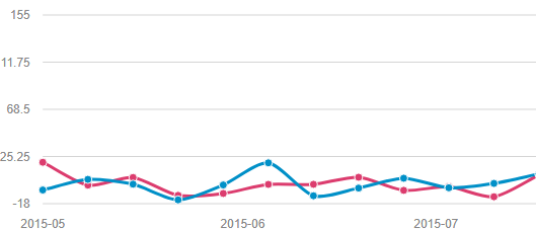
AB Test Analysis for Sum_Sum_Gross Margin

Time: 2018-08-19 13:49:47

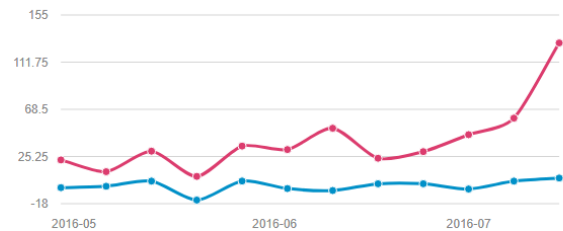


Time Comparison Plot

Comparison



Test

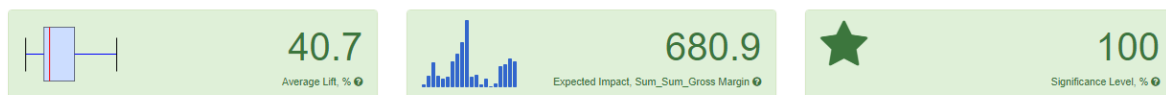


3. What is the lift from the new menu overall?

➔ **Overall Menu :-** Overall lift is 40.7 % and significance level is 100 %.

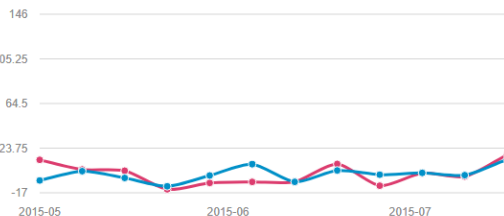
AB Test Analysis for Sum_Sum_Gross Margin

Time: 2018-08-19 13:49:14



Time Comparison Plot

Comparison



Test

