

## 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?

**Ans** – Round Roasters, a coffee restaurant in the United states of America, wants to conduct a market test to launch a new menu and needs to understand if the menu can drive enough sales to offset the cost of marketing the new menu. This project is to analyse the A/B test and recommend if the Round Roasters chain should launch the new menu or not.

The **Sum of Gross Margin** should be used as the performance metric to evaluate whether to introduce the new menu gourmet sandwiches and limited wine offering would increase the sales growth in Round Roasters.

2. What is the test period?

**Ans** – A period of 12 weeks from 29<sup>th</sup> April 2016 to 21<sup>st</sup> Jul 2016 is used run the test period. A total of 76 weeks of data would be needed till the end of the test period of 21<sup>st</sup> July 2016. The final period required for the data is 2015-02-06 till 2016-07-22

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

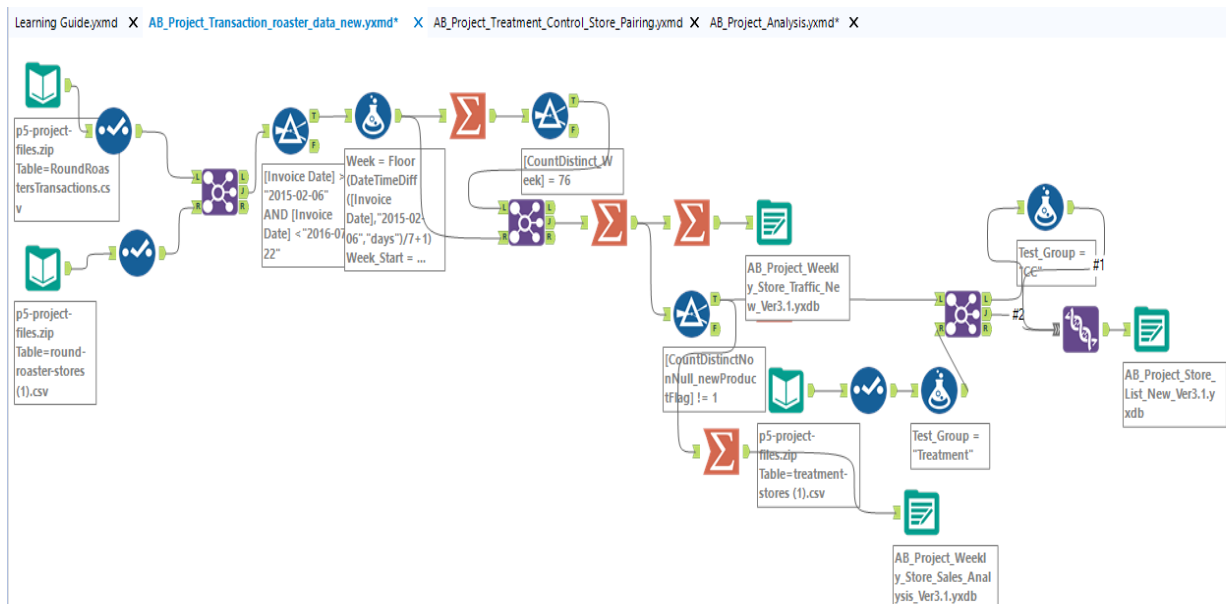
**Ans** – The data should be aggregated at a weekly level

## 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*

**Ans** – In this project, two files Round Roaster Transaction and Round Roaster stores datasets were used and first combined using the join tool. 76 weeks of data were used starting from 2015-02-06 to 2016-07-22 as A/B test requires 52 weeks of data in addition to a minimum of 12 weeks of data needed to calculate the seasonality. For the testing purpose another 12 weeks of data are added as the test period would last for 12 weeks.

Some new Week variables were created namely Week, Week\_start, Week\_end, newProduct flag to calculate the weekly traffic and sales for each store. Treatment\_store data set is then used to create a list of control and treatment stores.



## Alteryx Workflow – Data Cleanup

Record #	StoreID	Week	Week_Start	Week_End	Invoive_Count	Sumtotal_Gross Margin
1	10018	1	2015-02-06	2015-02-12	308	2212.7105
2	10018	2	2015-02-13	2015-02-19	288	2164.007
3	10018	3	2015-02-20	2015-02-26	204	1560.929
4	10018	4	2015-02-27	2015-03-05	320	2342.984
5	10018	5	2015-03-06	2015-03-12	284	2199.4065
6	10018	6	2015-03-13	2015-03-19	288	2103.143

## Invoice count and Gross Margin Aggregation

### Step 3: Match Treatment and Control Units

*In this step, you should create the trend and seasonality variables, and use them along with your 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.

**Ans** – AvgMonthSales and Sq.Ft should be considered.

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

**Ans** – AvgMonthSales has a very high correlation with the performance metric sum of Gross Margin. Whereas Sq. Ft has a very poor correlation and needs to be ignored.

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

**Ans** – AvgMonthSales along with Trend and seasonality will be used when matching 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	12586	12736
1675	2114	11268
1696	1964	2301
1700	12119	10518
1712	7284	3102
2288	9081	8112
2293	11568	12219
2301	1964	10018
2322	1696	2409
2341	2333	6992

Record #	Controls	Treatments	Distance
1	9081	2288	0.171017
2	8112	2288	0.380848
3	11568	2293	0.150729
4	12219	2293	0.269406
5	1964	2301	0.10782
6	10018	2301	0.124762
7	1696	2322	0.156122
8	2409	2322	0.166552
9	2333	2341	0.114878
10	6992	2341	0.241309

Record #	Controls	Treatments	Distance
1	12586	1664	0.11718
2	12736	1664	0.17444
3	2114	1675	0.0405
4	11268	1675	0.156505
5	1964	1696	0.073437
6	2301	1696	0.121982
7	12119	1700	0.132016
8	10518	1700	0.163157
9	7284	1712	0.14585
10	3102	1712	0.148023

## 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?

**Ans** – Yes, the company should roll out the updated menu to all stores as the profitability has definitely increased by more than 18%.

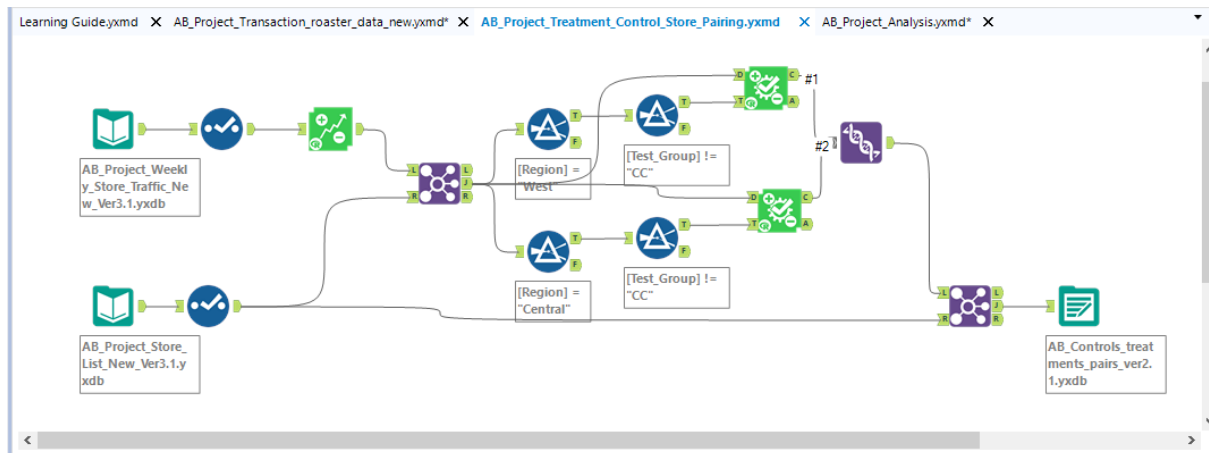
2. What is the lift from the new menu for West and Central regions (include statistical significance)?

**Ans** – The lift for West Region is 34 % and the lift for Central region is 45% and both the regions have statistical significance level of 99.3% and 99.6 % respectively.

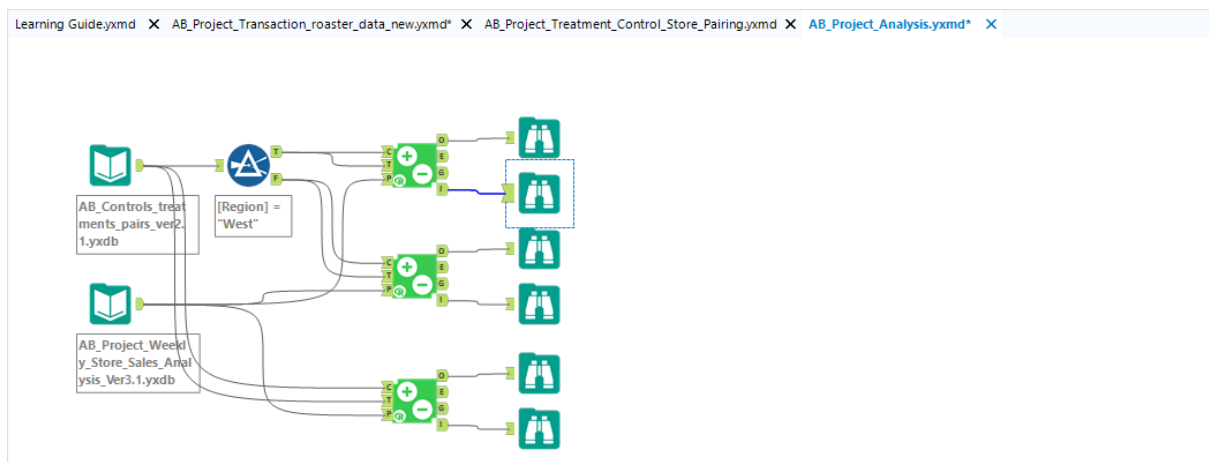
3. What is the lift from the new menu overall?

**Ans** – The lift from the new menu over all is 39.9 % with a statistical significance level of 100%.

### **Alteryx Workflow – Treatment-Control-Store Pairing**



### **Alteryx Workflow - Analysis**



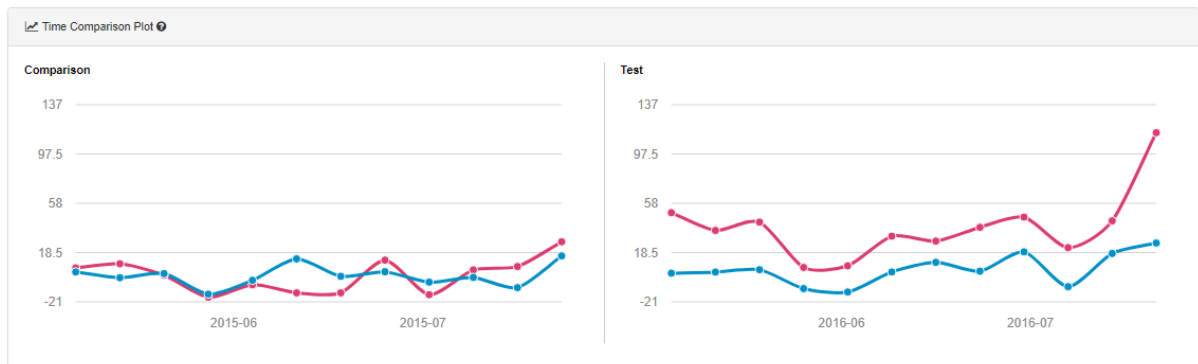
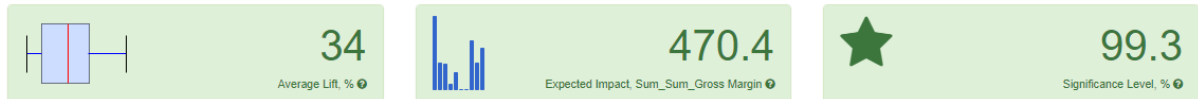
## West Region

### Lift Analysis for Sum\_Sum\_Gross Margin

Lift	Expected Impact	Significance Level
34.0%	470	99.3%
Summary Statistics for Sum_Sum_Gross Margin by Test Group		
Statistic	Treatment	Control
Average	39.17	4.79
Minimum	12.34	-12.16
Maximum	55.30	32.77
Standard Deviation	16.34	13.54

### AB Test Analysis for Sum\_Sum\_Gross Margin

Time: 2018-11-20 17:44:29



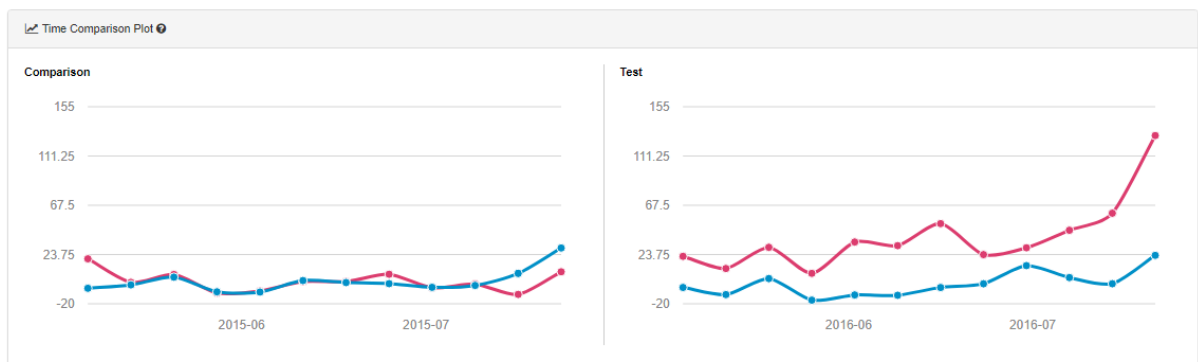
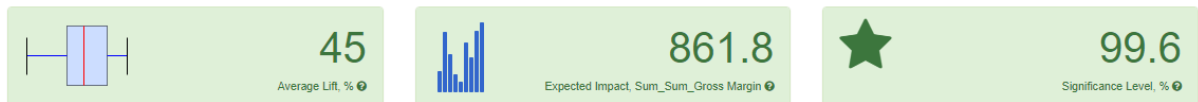
## Central Region

### Lift Analysis for Sum\_Sum\_Gross Margin

Lift	Expected Impact	Significance Level
45.0%	862	99.6%
Summary Statistics for Sum_Sum_Gross Margin by Test Group		
Statistic	Treatment	Control
Average	39.74	-2.17
Minimum	20.09	-18.50
Maximum	67.52	20.53
Standard Deviation	17.15	12.16

### AB Test Analysis for Sum\_Sum\_Gross Margin

Time: 2018-11-20 17:44:55



Overall

Lift Analysis for Sum_Sum_Gross Margin			
Lift	Expected Impact		Significance Level
39.5%	666		100.0%
Summary Statistics for Sum_Sum_Gross Margin by Test Group			
Statistic	Treatment		Control
Average	39.45		1.31
Minimum	12.34		-18.50
Maximum	67.52		32.77
Standard Deviation	16.30		13.02

AB Test Analysis for Sum\_Sum\_Gross Margin

Time: 2018-11-20 17:45:20

