Project: Analyzing a Market Test

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?
   This A/B test project is aimed to evaluate whether the new menu can increase enough sales. We need to decide whether launch a new menu or not.
2. What is the test period?
   The test period will be 2016-April-29 to 2016-July-21 (12 weeks.)
3. At what level (day, week, month, etc.) should the data be aggregated?
   Our data needs to be aggregated (grouped) by Store ID, Number of Weeks, Store Traffic, Gross Margin and Sales.

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.

The steps I followed:

- *Filter the data to the proper date range ([Invoice Date]>="2015-02-06" AND [Invoice Date]<"2016-07-22")*
- *Aggregate the data to get the weekly gross margin and weekly traffic count (count of unique invoices)*
- *Calculate Trend and Seasonality with the AB Trend Tool*
- *Label the data as treatment and control stores*
- *Calculate correlation between other numeric measures and the performance metric (gross margin)*
- *Match Treatment to control stores per region using the AB Controls Tool*
- *Calculate lift from control to treatment store with AB Analysis tool*

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.

StoreID	Week	Week_Start	Week_End	Count	Sum_Gross Margin	Sum_Sales
10018	1	2015-02-06	2015-02-12	308	2212.71	4741.48
10018	2	2015-02-13	2015-02-19	288	2164.01	4571.25
10018	3	2015-02-20	2015-02-26	204	1560.93	3348.25
10018	4	2015-02-27	2015-03-05	320	2342.98	5114.96

10018	5	2015-03-06	2015-03-12	284	2199.41	4799.48
10018	6	2015-03-13	2015-03-19	288	2103.14	4554.97
10018	7	2015-03-20	2015-03-26	194	1412.93	2999.55
10018	8	2015-03-27	2015-04-02	286	2124.37	4519.61
10018	9	2015-04-03	2015-04-09	274	2216.15	4727.05
10018	10	2015-04-10	2015-04-16	215	1686.25	3689.33

Apart from trend and seasonality...

1. What control variables should be considered? Note: Only consider variables in the RoundRoastersStore file.
   *Besides **trend** and **seasonality**, we should consider **Sq_Ft** and **Average Monthly Sales** for each store. Because they are numeric and seems more reasonable.*

2. What is the correlation between your each potential control variable and your performance metric?
   An association analysis is performed to find the correlation between the potential control variables and the performance metric (Gross Margin.) From the analysis, it is observed that AvgMonthSales is highly correlated to the performance metric where Sq_Ft is not.

Pearson Correlation Analysis

Focused Analysis on Field Sum_Gross.Margin

	Association Measure	p-value	
AvgMonthSales	0.788408	0.000000	***
Sq_Ft	-0.020474	0.039457	*

3. What control variables will you use to match treatment and control stores?
   *Eventually, we should choose **trend**, **seasonality** and **Average Monthly Sales** to match treatment and control stores. Because Sq_Ft doesn't have enough correlation with the target variable.*

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

Controls	Treatments	Distance
12536	2301	0.3769839
12586	2341	0.4598068
12686	2293	0.6895000
1863	1696	0.4131876
1964	1664	0.2946072
2014	1700	1.0075255
2214	1675	0.7034827
2568	2288	0.4072874
2572	2341	0.2622104
3185	2322	0.2791392

7037	1700	0.9185037
7162	1664	0.3461303
7284	1675	0.6638134
7334	1696	0.6614922
7434	1712	0.6120869
8162	1712	0.4878611
9081	2288	0.4469973
9238	2301	0.3832117
9388	2322	0.2313615
9639	2293	0.7303510

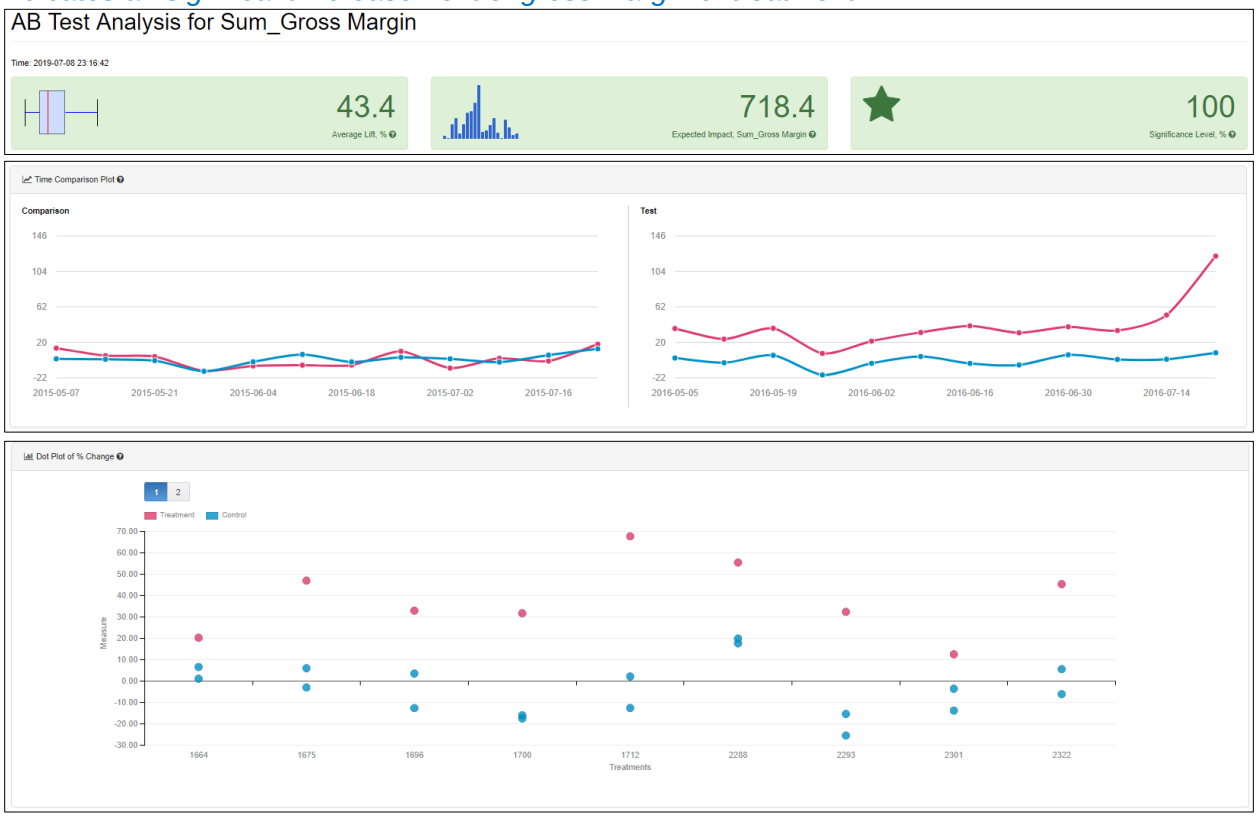
Step 4: Analysis and Writeup

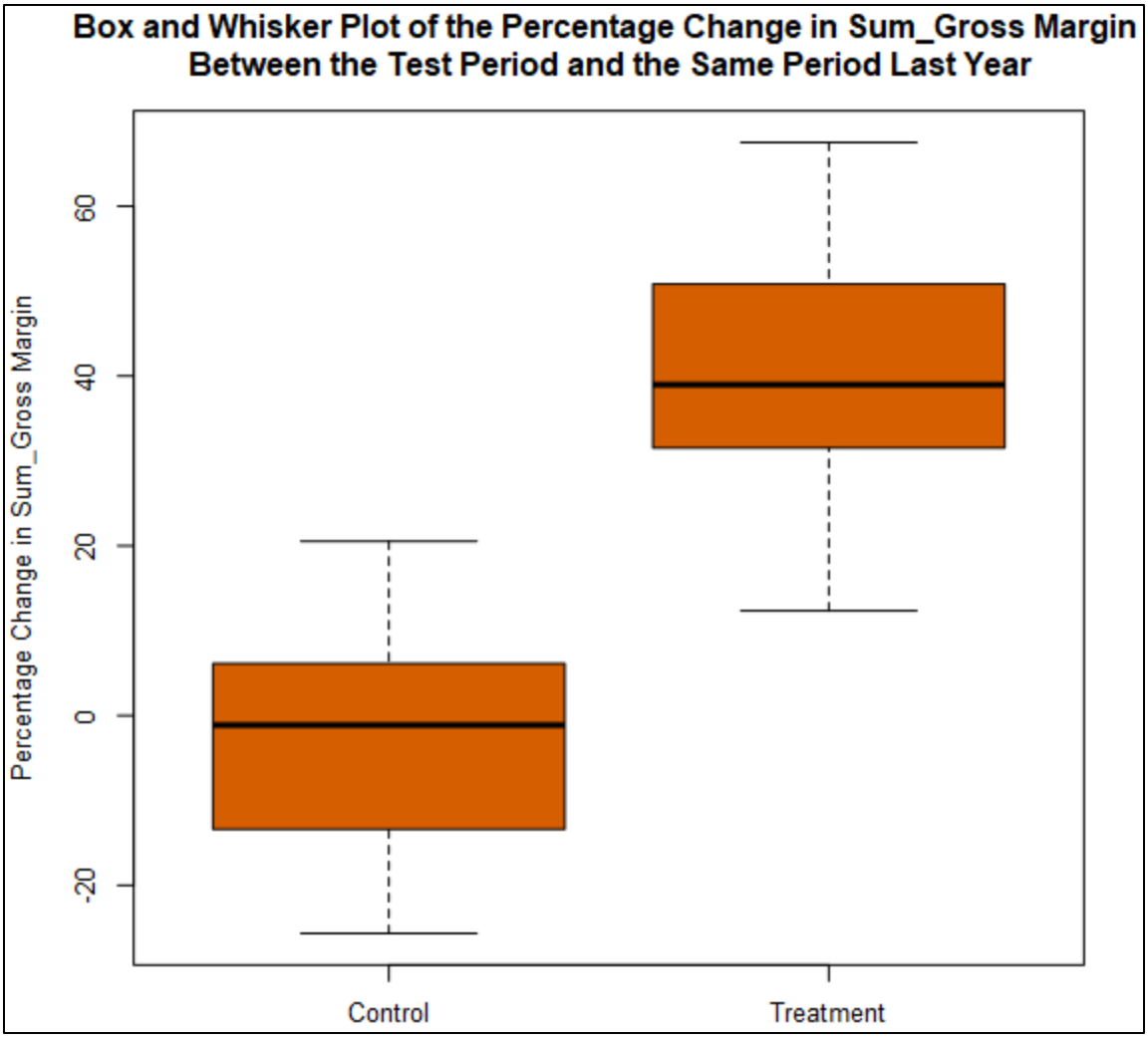
Conduct your A/B analysis and create a short report outlining your results and recommendations. (250 words limit)

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?

   Yes. From the analysis, we should launch a new menu to all stores. The below figures indicates an significant increasement of gross margin of treatment.





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

Central: Lift=47.6%, Significance Level=99.60%

Lift Analysis for Sum_Gross Margin

Lift	Expected Impact	Significance Level
47.6%	906	99.6%

West: Lift=39.10%, Significance Level=99.60%

Lift Analysis for Sum_Gross Margin

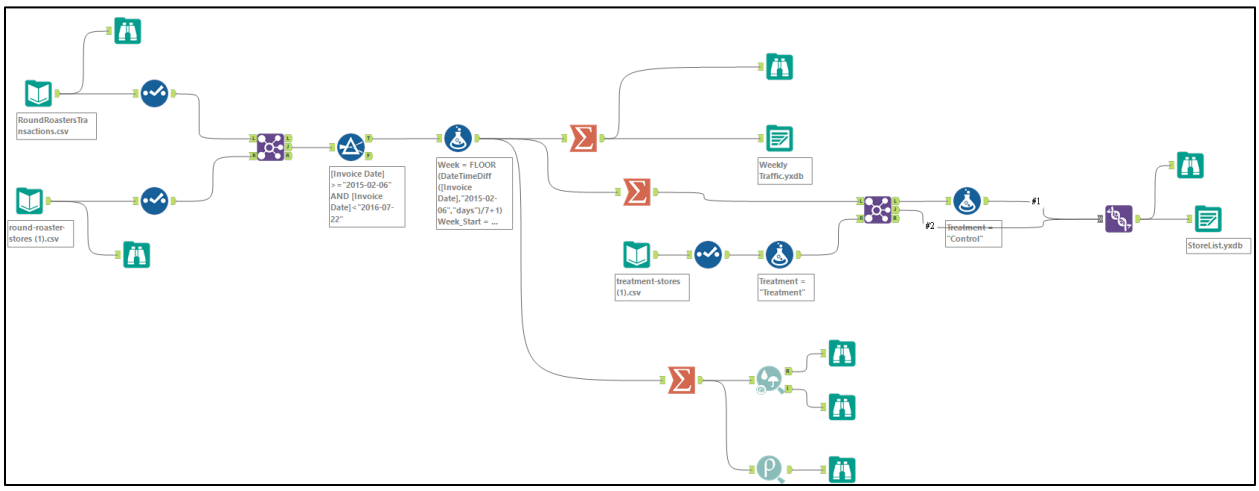
Lift	Expected Impact	Significance Level
39.1%	530	99.6%

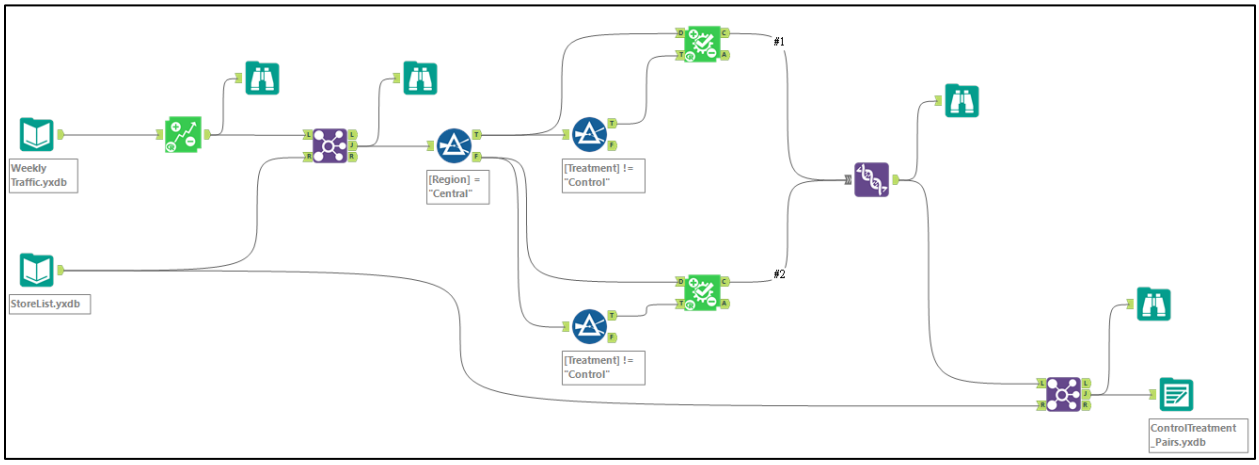
3. What is the lift from the new menu overall?

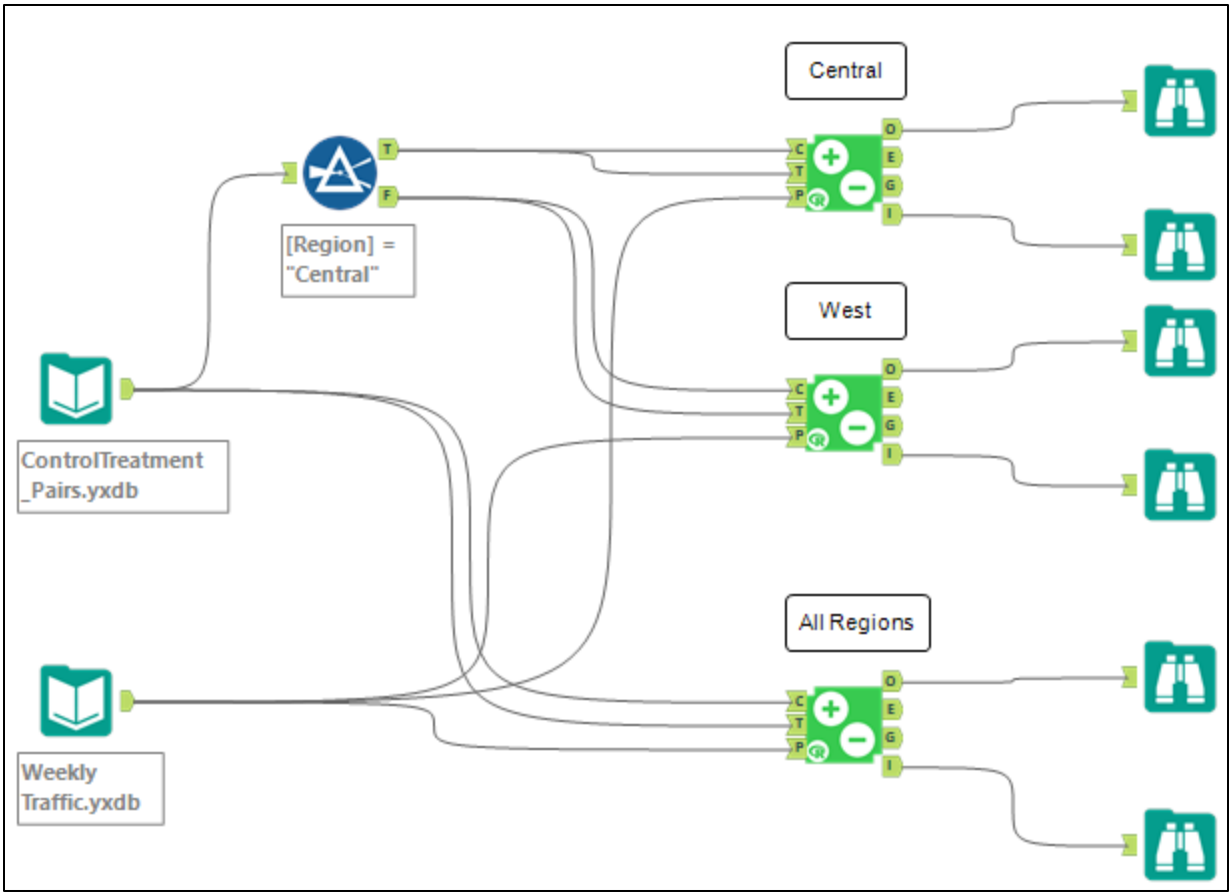
The overall lift is 43.4%.

Lift Analysis for Sum_Gross Margin		
Lift	Expected Impact	Significance Level
43.4%	718	100.0%

Appendix: Alteryx Workflows






Central
West
All Regions
[Region] =
"Central"
ControlTreatment
_Pairs.yxdb
Weekly
Traffic.yxdb
T
F
C
T
P
O
E
G
I