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

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?

   Sum of the gross margin

2. What is the test period?

   12 Weeks, 4/29/16-7/21/16

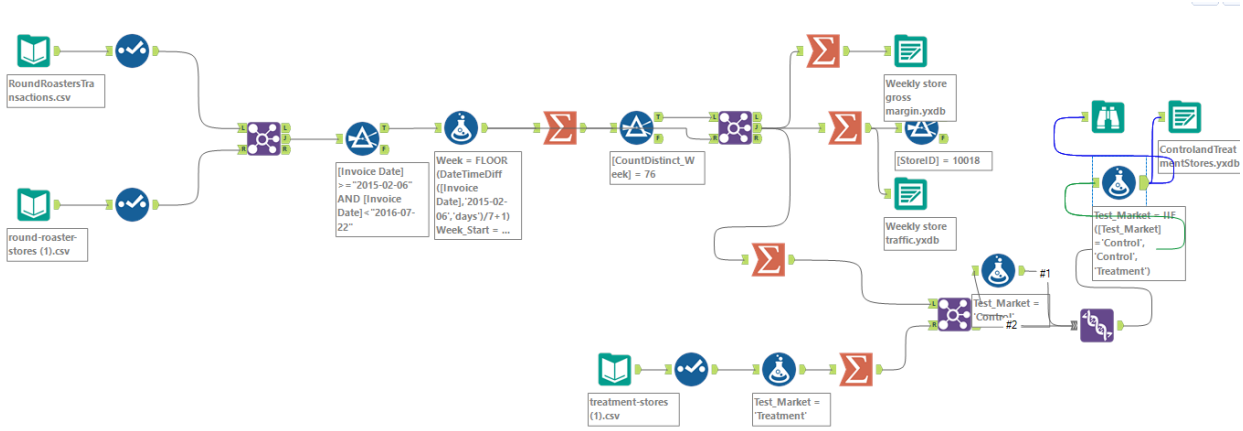
3. At what level (day, week, month, etc.) should the data be aggregated?

   Weekly

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.

Added 3 columns, Week, Week_Start and Week_End.
Generated Weekly store gross margin and Weekly store traffic files
Joined with Treatment stores input file and added column control or treatments to Control and Treatment Stores file

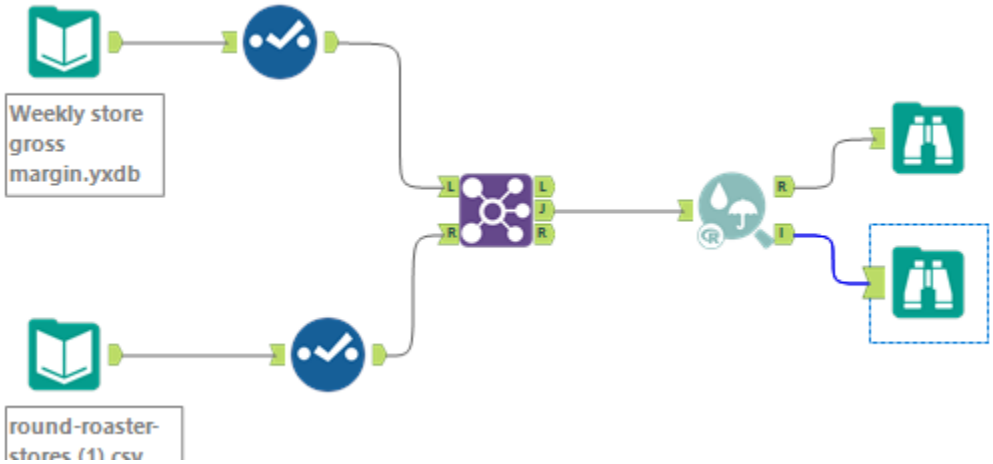


Step 3: Match Treatment and Control Units

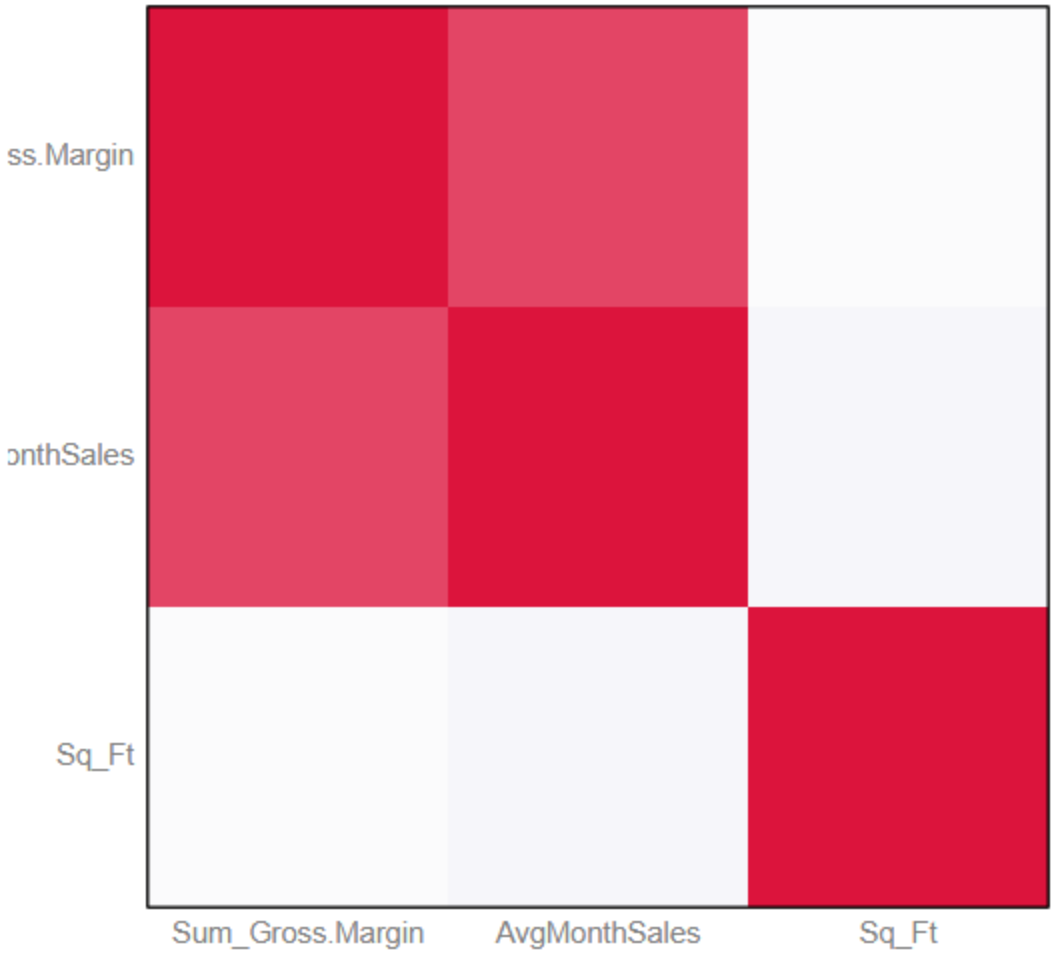
In this step, you should create the trend and seasonality variables, and use them along with you other control variable(s) to match two control units to each treatment unit. Note: Calculate the number of transactions per store per week to calculate trend and seasonality.

Apart from trend and seasonality...

1. What control variables should be considered? Note: Only consider variables in the RoundRoastersStore file.
   AvgMonthSales, Tren, Seasonality and Sq_Ft
2. What is the correlation between your each potential control variable and your performance metric?
   Using association analysis, AvgMonthlySales has high correlation and Sq_Ft is low.



Correlation Matrix with ScatterPlot



The left panel is an image of a correlation matrix, with blue = -1 and red = +1. Hover over pixels in the correlation matrix on the left to see the values; click to see the corresponding scatterplot on the right. The variables have been clustered based on degree of correlation, so that highly correlated variables appear adjacent to each other.

Pearson Correlation Analysis

Full Correlation Matrix

	Sum_Gross.Margin	Sq_Ft	AvgMonthSales
Sum_Gross.Margin	1.000000	-0.023163	0.789998
Sq_Ft	-0.023163	1.000000	-0.046967
AvgMonthSales	0.789998	-0.046967	1.000000

Matrix of Corresponding p-values

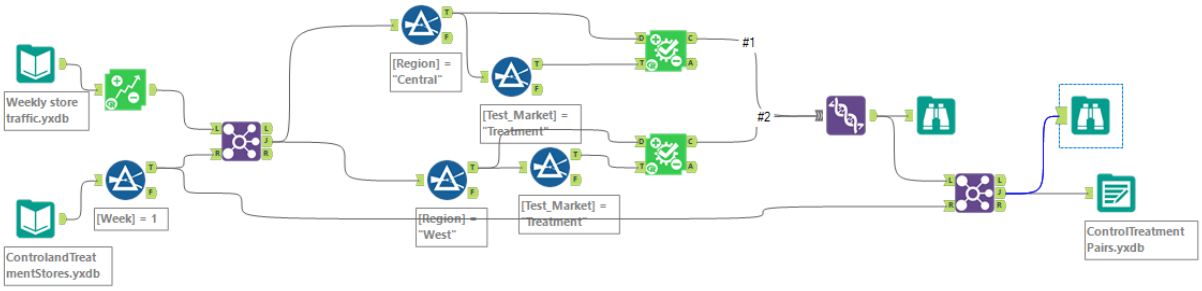
	Sum_Gross.Margin	Sq_Ft	AvgMonthSales
Sum_Gross.Margin		1.9871e-02	0.0000e+00
Sq_Ft	1.9871e-02		2.3119e-06
AvgMonthSales	0.0000e+00	2.3119e-06	

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

   AvgMonthSales, Trend and Seasonality

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

Treatment Store	Control Store 1	Control Store 2
1164	2214	1857
1675	2114	7770
1696	7434	1964
1700	1508	2014
1712	8212	6992
2288	2568	12536
2293	11568	10618
2301	10018	3002
2322	10468	11318
2341	8817	12286



6 of 6 Fields | Cell Viewer

Record #	Controls	Treatments	Distance	StoreID	Region	Test_Market
1	2214	1664	0.197301	1664	Central	Treatment
2	1857	1664	0.294089	1664	Central	Treatment
3	2114	1675	0.132899	1675	Central	Treatment
4	7770	1675	0.241155	1675	Central	Treatment
5	7434	1696	0.289172	1696	Central	Treatment
6	1964	1696	0.300062	1696	Central	Treatment
7	1508	1700	0.286046	1700	Central	Treatment
8	2014	1700	0.514526	1700	Central	Treatment
9	8212	1712	0.205904	1712	Central	Treatment
10	6992	1712	0.223096	1712	Central	Treatment
11	2568	2288	0.308527	2288	West	Treatment
12	12536	2288	0.349303	2288	West	Treatment
13	11568	2293	0.262358	2293	West	Treatment
14	10618	2293	0.318339	2293	West	Treatment
15	10018	2301	0.081238	2301	West	Treatment
16	3002	2301	0.153201	2301	West	Treatment
17	10468	2322	0.147195	2322	West	Treatment
18	11318	2322	0.238718	2322	West	Treatment
19	8817	2341	0.159338	2341	West	Treatment
20	12286	2341	0.200322	2341	West	Treatment

Step 4: Analysis and Writeup

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

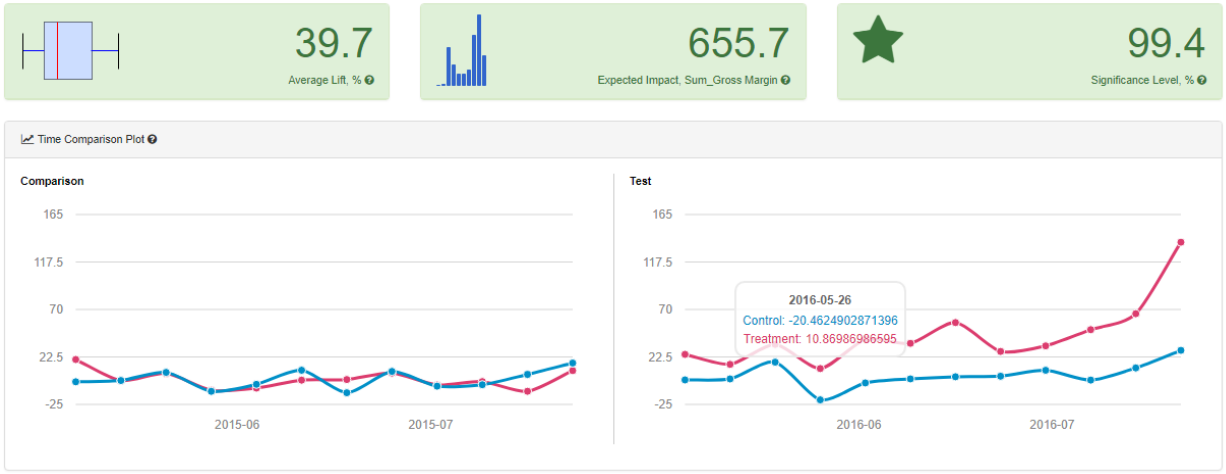
Answer these questions. Be sure to include visualizations from your analysis:

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

   Yes, New menus should be rolled out to all the stores due to a 43.8% overall lift, $623.9 increase in overall gross margin and 100% significance level.

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

The lift for the West region is 39.7% @ 99.4% statistical level.



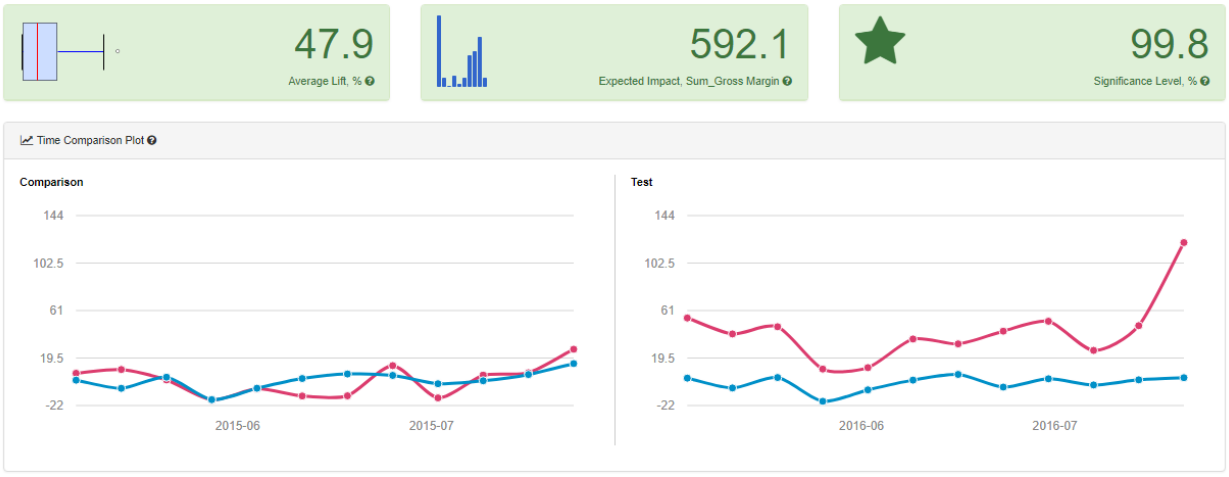
Lift Analysis for Sum_Gross Margin

Lift	Expected Impact	Significance Level
39.7%	656	99.4%

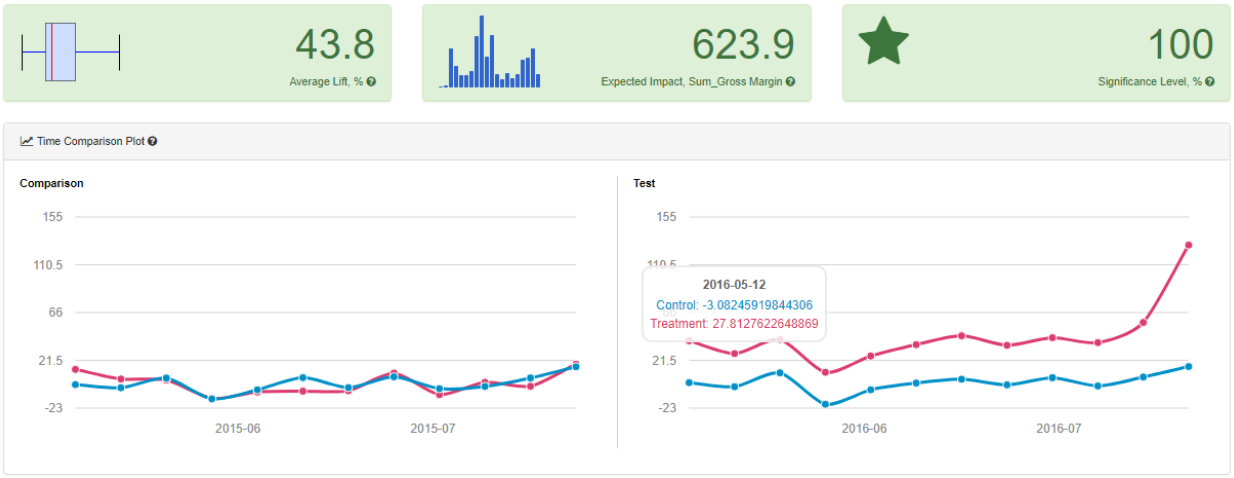
Summary Statistics for Sum_Gross Margin by Test Group

Statistic	Treatment	Control
Average	44.46	4.08
Minimum	23.56	-16.60
Maximum	73.40	27.12
Standard Deviation	17.94	11.07

The lift for the Central region is 47.9% @ 99.8% statistical level.



3. What is the lift from the new menu overall?
   The overall lift is 43.8% @ 100% statistical level.



Lift Analysis for Sum_Gross Margin

Lift	Expected Impact	Significance Level
43.8%	624	100.0%

Summary Statistics for Sum_Gross Margin by Test Group

Statistic	Treatment	Control
Average	43.89	0.81
Minimum	16.06	-16.60
Maximum	73.40	27.12
Standard Deviation	16.85	11.68

Before you Submit

Please check your answers against the requirements of the project dictated by the rubric here. Reviewers will use this rubric to grade your project.