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. (250 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?
   I will use the sum of gross margin per store per week for evaluation.
2. What is the test period?
   The test period is 12 weeks between 2016-04-29 to 2016-07-21.
3. At what level (day, week, month, etc.) should the data be aggregated?
   To analyze the trend and seasonality, we will use weekly aggregated data.

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.

Step 3: Match Treatment and Control Units

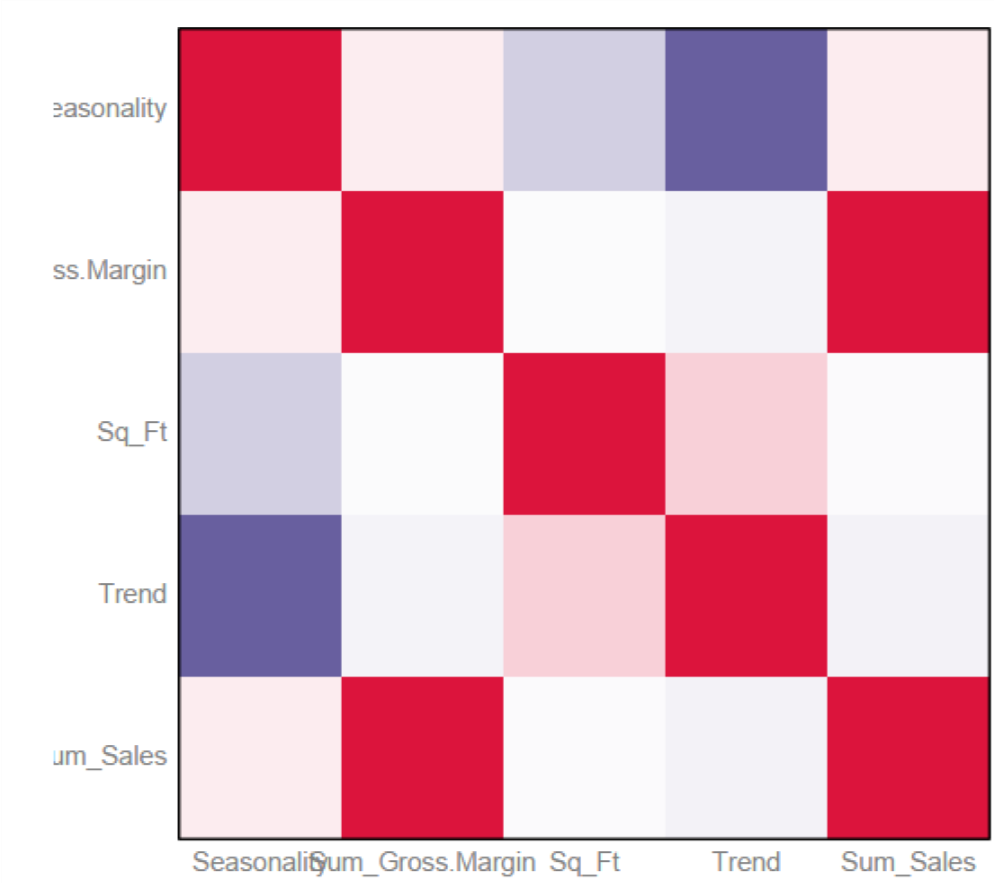
In this step, you should create the trend and seasonality variables, and use them along with you other control variable(s) to match two control units to each treatment unit. Note: Calculate the number of transactions per store per week to calculate trend and seasonality.

Apart from trend and seasonality...

1. What control variables should be considered? Note: Only consider variables in the RoundRoastersStore file.
   We shall only consider Sq_Ft and AvgMonthSales, all other variables are geographic, and we have already incorporated the region variable.
2. What is the correlation between your each potential control variable and your performance metric?
   Sum_Gross_Margin is our performance metric. It has correlation 1 with Sum_Sales, and very low correlations with other variables. So we will not use Sum_Sales. Also notice Trend and Seasonality have high negative correlation.

Full Correlation Matrix

	Trend	Seasonality	Sq_Ft	Sum_Gross.Margin	Sum_Sales
Trend	1.000000	-0.826891	0.201004	-0.061923	-0.065882
Seasonality	-0.826891	1.000000	-0.246428	0.074862	0.079514
Sq_Ft	0.201004	-0.246428	1.000000	-0.021685	-0.024071
Sum_Gross.Margin	-0.061923	0.074862	-0.021685	1.000000	0.998699
Sum_Sales	-0.065882	0.079514	-0.024071	0.998699	1.000000



3. What control variables will you use to match treatment and control stores?
   Finally, we choose Trend, Seasonality and Sq_Ft as control variables.
4. Please fill out the table below with your treatment and control stores pairs:

Record #	Controls	Treatments	Distance	Region
1	7484	1664	0.28981	Central
2	7162	1664	0.642755	Central
3	7534	1675	0.396317	Central
4	7334	1675	0.518691	Central
5	7584	1696	0.064441	Central
6	7434	1696	0.467633	Central
7	1508	1700	0.342317	Central
8	1662	1700	0.515143	Central
9	7284	1712	0.276895	Central
10	6992	1712	0.332487	Central
11	8817	2288	0.702143	West
12	11818	2288	0.838997	West
13	11568	2293	0.541731	West
14	8967	2293	0.595259	West
15	3102	2301	0.342539	West
16	3002	2301	0.514266	West
17	11668	2322	0.320168	West
18	3002	2322	0.551849	West
19	9017	2341	0.204936	West
20	2333	2341	0.239558	West

Treatment Store	Control Store 1	Control Store 2
1664	7484	7162
1675	7534	7334
1696	7584	7434
1700	1662	7284
1712	7284	6992
2288	8817	11818
2293	11568	8967
2301	3102	3002
2322	11668	3002
2341	9017	2333

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?
   The company should roll out the updated menu to all stores.
2. What is the lift from the new menu for West and Central regions (include statistical significance)?

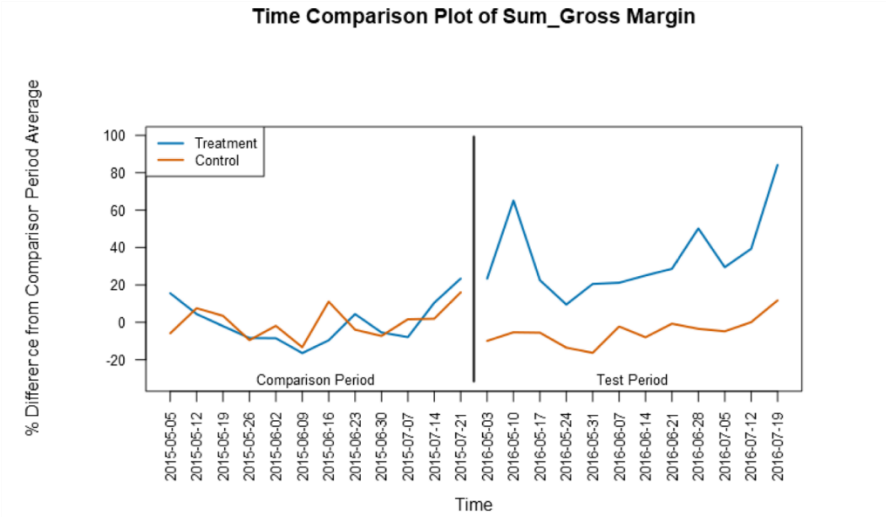
   WEST: 43.6% lift, 99.7% significance

Lift Analysis for Sum_Gross Margin

Significance Level	Lift	Expected Impact
99.7%	43.6%	621

Summary Statistics for Sum_Gross Margin by Test Group

Statistic	Treatment	Control
Average	34.92	-4.83
Minimum	8.39	-19.88
Maximum	48.42	11.95
Standard Deviation	15.51	11.45



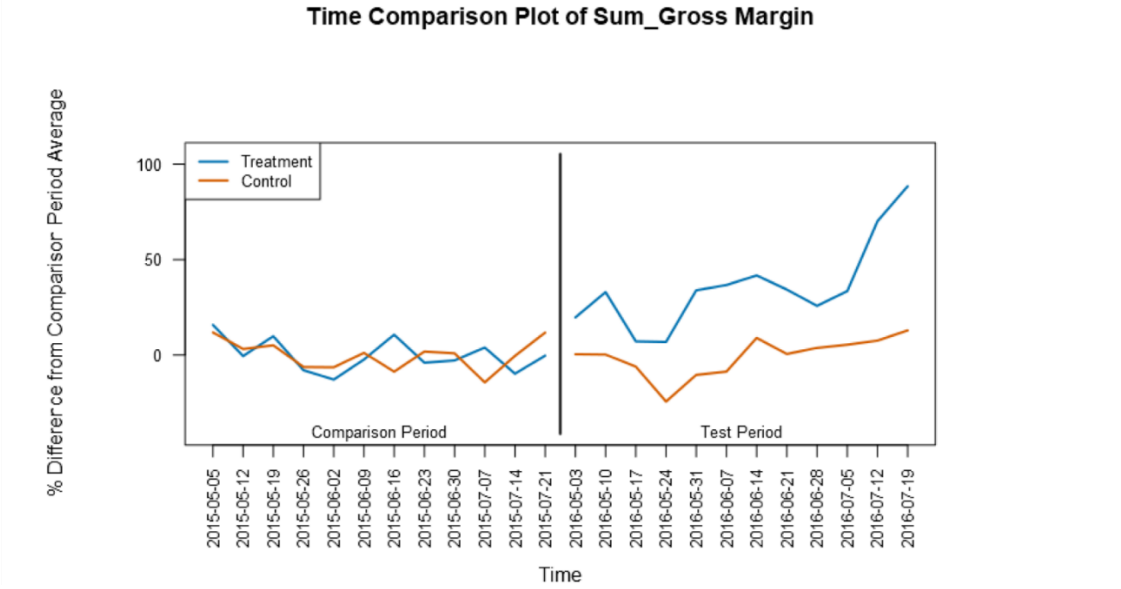
CENTRAL:38.1% lift, 99.3% significance

Lift Analysis for Sum_Gross Margin

Significance Level	Lift	Expected Impact
99.3%	38.1%	731

Summary Statistics for Sum_Gross Margin by Test Group

Statistic	Treatment	Control
Average	35.95	-0.77
Minimum	19.85	-15.28
Maximum	62.43	7.85
Standard Deviation	16.52	7.82



3. What is the lift from the new menu overall?
   Overall Lift: 40.9%. Overall significance: 100.0%.

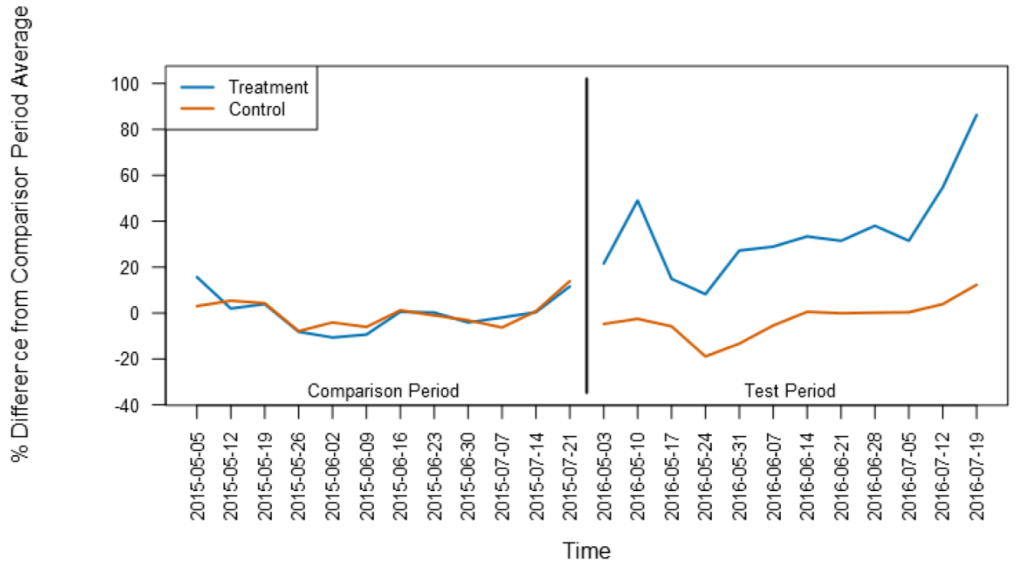
Lift Analysis for Sum_Gross Margin

Significance Level	Lift	Expected Impact
100.0%	40.9%	676

Summary Statistics for Sum_Gross Margin by Test Group

Statistic	Treatment	Control
Average	35.43	-2.80
Minimum	8.39	-19.88
Maximum	62.43	11.95
Standard Deviation	15.60	9.77

Time Comparison Plot of Sum_Gross Margin



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.