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. Its correlation with Sq_Ft is -0.0238, while its correlation with AvgMonthSales is 0.7867. So we shall only use AvgMonthSales.

Pearson Correlation Analysis

Full Correlation Matrix

	Trend	Seasonality	Sq_Ft	AvgMonthSales	Week	Sum_Gross.Margin
Trend	1.00000000	-0.83524919	0.19280405	-0.12182125	0.00021653	-0.06412232
Seasonality	-0.83524919	1.00000000	-0.23869123	0.13591684	-0.00024721	0.07442310
Sq_Ft	0.19280405	-0.23869123	1.00000000	-0.04037070	0.00048701	-0.02381035
AvgMonthSales	-0.12182125	0.13591684	-0.04037070	1.00000000	0.00050235	0.78666037
Week	0.00021653	-0.00024721	0.00048701	0.00050235	1.00000000	-0.08426519
Sum_Gross.Margin	-0.06412232	0.07442310	-0.02381035	0.78666037	-0.08426519	1.00000000
Sum_Sales	-0.06949458	0.08105287	-0.02767582	0.78828639	-0.08732209	0.99905199

	Sum_Sales
Trend	-0.06949458
Seasonality	0.08105287
Sq_Ft	-0.02767582
AvgMonthSales	0.78828639
Week	-0.08732209
Sum_Gross.Margin	0.99905199
Sum_Sales	1.00000000

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

Controls	Treatments	Distance	Region
7162	1664	0.339158	Central
8112	1664	0.788898	Central
1580	1675	0.561443	Central
1807	1675	0.83088	Central
1964	1696	0.351184	Central
1863	1696	0.525915	Central
2014	1700	0.799888	Central
1630	1700	0.920386	Central
8162	1712	0.818742	Central
7434	1712	0.965047	Central
9081	2288	0.265423	West
2568	2288	0.547031	West
12219	2293	0.296946	West
9524	2293	0.461749	West
3102	2301	0.219454	West
11668	2301	0.266407	West
2409	2322	0.232107	West
9388	2322	0.456012	West
12536	2341	0.323771	West
2572	2341	0.504903	West

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	11668
2322	2409	9388
2341	12536	2572

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: 38.1% lift, 99.3% significance

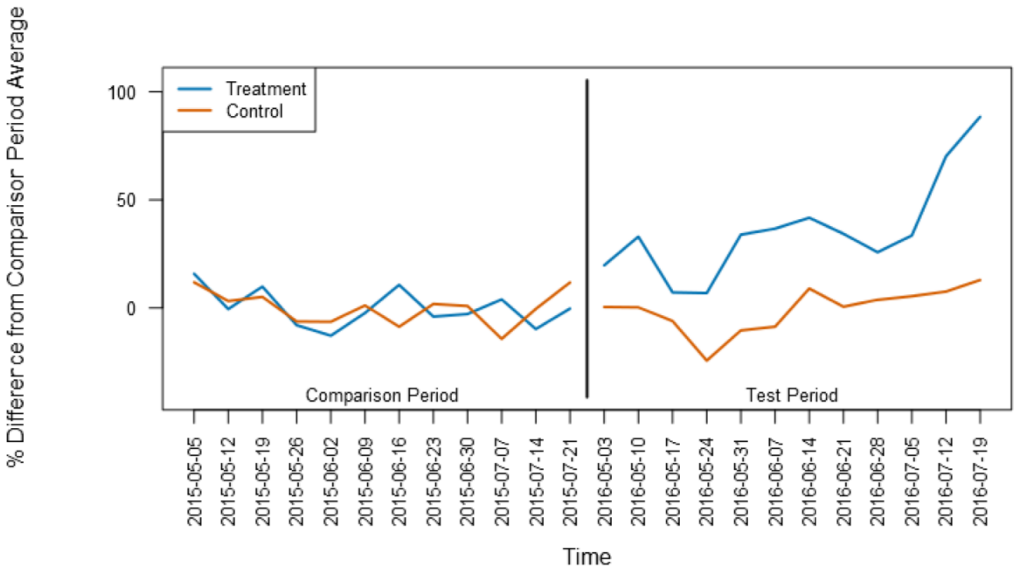
Lift Analysis for Sum_Gross Margin

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

Time Comparison Plot of Sum_Gross Margin



   CENTRAL:43.6% lift, 99.7% significance

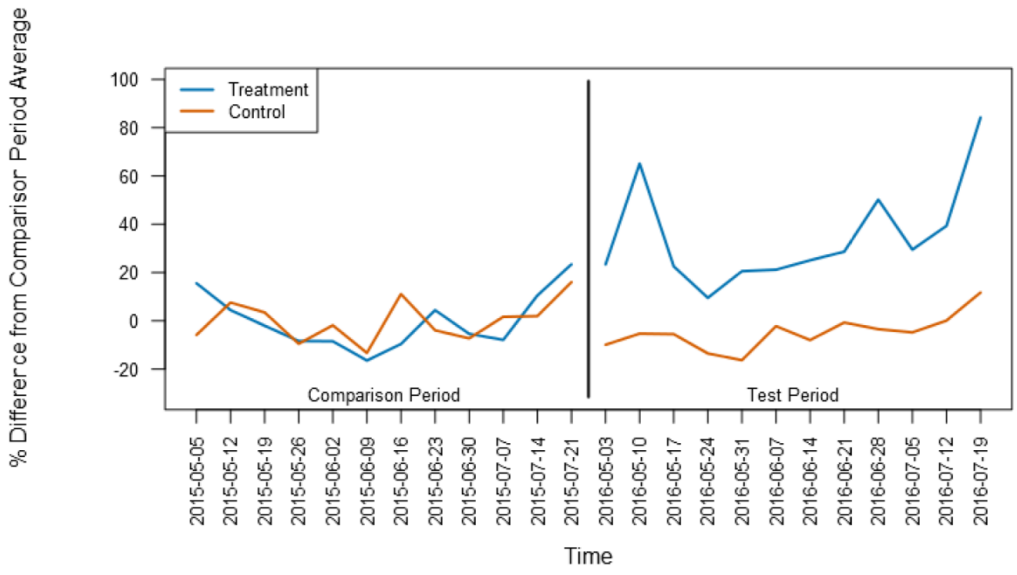
Lift Analysis for Sum_Gross Margin

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

Time Comparison Plot of Sum_Gross Margin



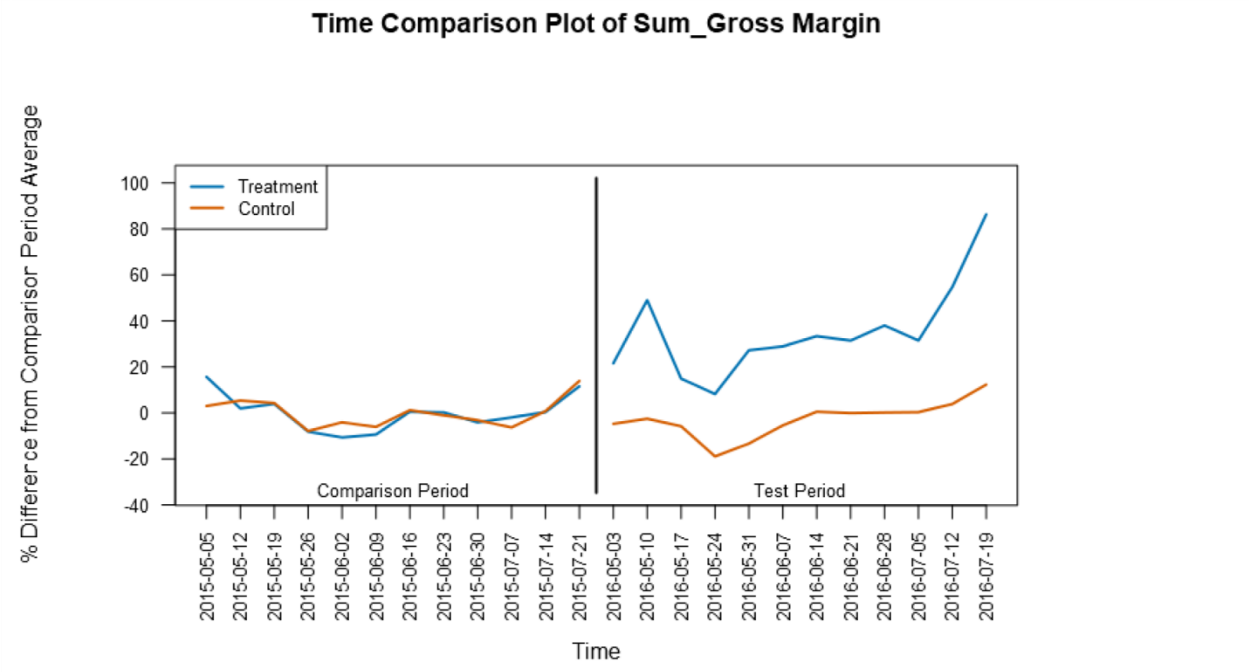
3. What is the lift from the new menu overall?
   Overall Lift: 40.9%. Overall significance: 100.0%.

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



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.