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?

    - ✓ The company is thinking of introducing a new menu to all stores based on a profit growth of at least 18%. In the data we have, profit is represented by gross margin so that's what we will use for evaluation.

2. What is the test period?

    - ✓ The test will run for a period of 12 weeks (2016-04-29 to 2016-07-21) where five stores in each of the test markets will be offered the updated menu along with television advertising.

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

    - ✓ Each period last for one week. For this reason, the data should be aggregated at the week 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.

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.

    - ✓ Based on the data, we see that both Sq_Ft and AvgMonthSales are good candidates for being control variables.

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

Pearson Correlation Analysis

Full Correlation Matrix

	Sum_Gross_Margin	Sq_Ft	AvgMonthSales
Sum_Gross_Margin	1.000000	-0.024255	0.990982
Sq_Ft	-0.024255	1.000000	-0.046967
AvgMonthSales	0.990982	-0.046967	1.000000

- ✓ The correlation between Sq_Ft area and the Sum_Gross_Margin (our target variable) is very low, it has barely any relationship between them, but looking at the sum of gross margin and the average monthly sales, we could see a strong relationship between them. Therefore, we eliminate the square footage area from our control variables list.

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

- ✓ *Trend*
- ✓ *Seasonality*
- ✓ Average monthly sales

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

Treatment Store	Control Store 1	Control Store 2
1664	7162	8112
1675	1580	1807
1696	1964	7334
1700	2014	1630
1712	8162	7434
2288	9081	2568
2293	12219	9589
2301	11668	9238
2322	2409	3235
2341	2572	3102

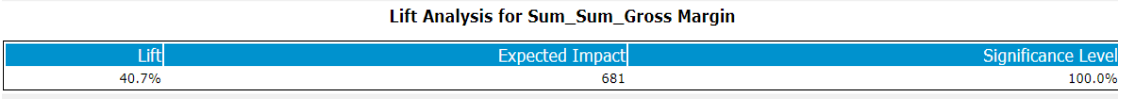
Step 4: Analysis and Writeup

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

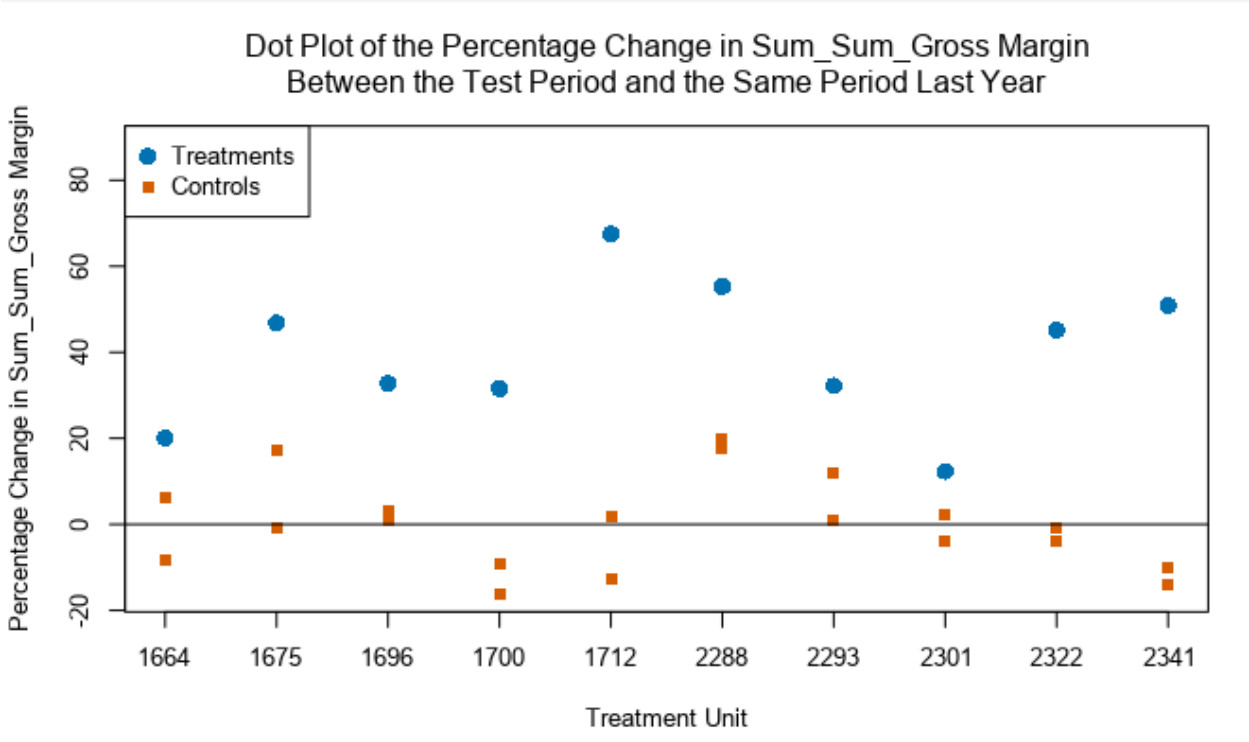
Answer these questions. Be sure to include visualizations from your analysis:

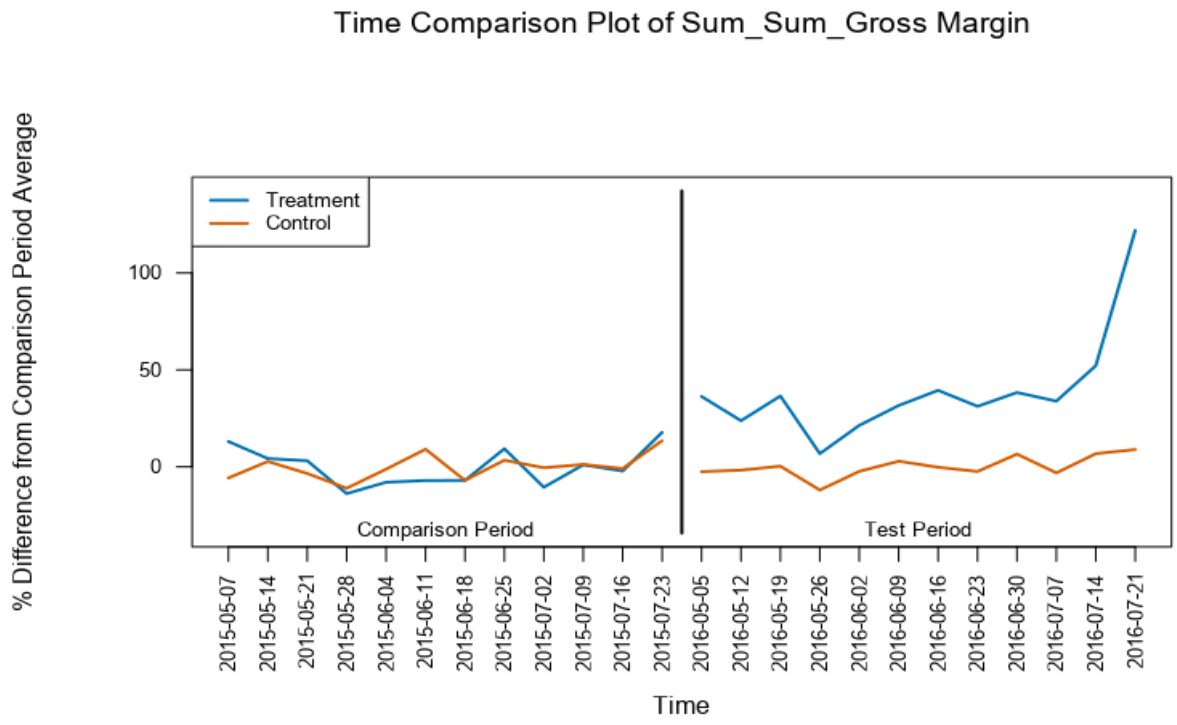
1. What is your recommendation - Should the company roll out the updated menu to all stores?

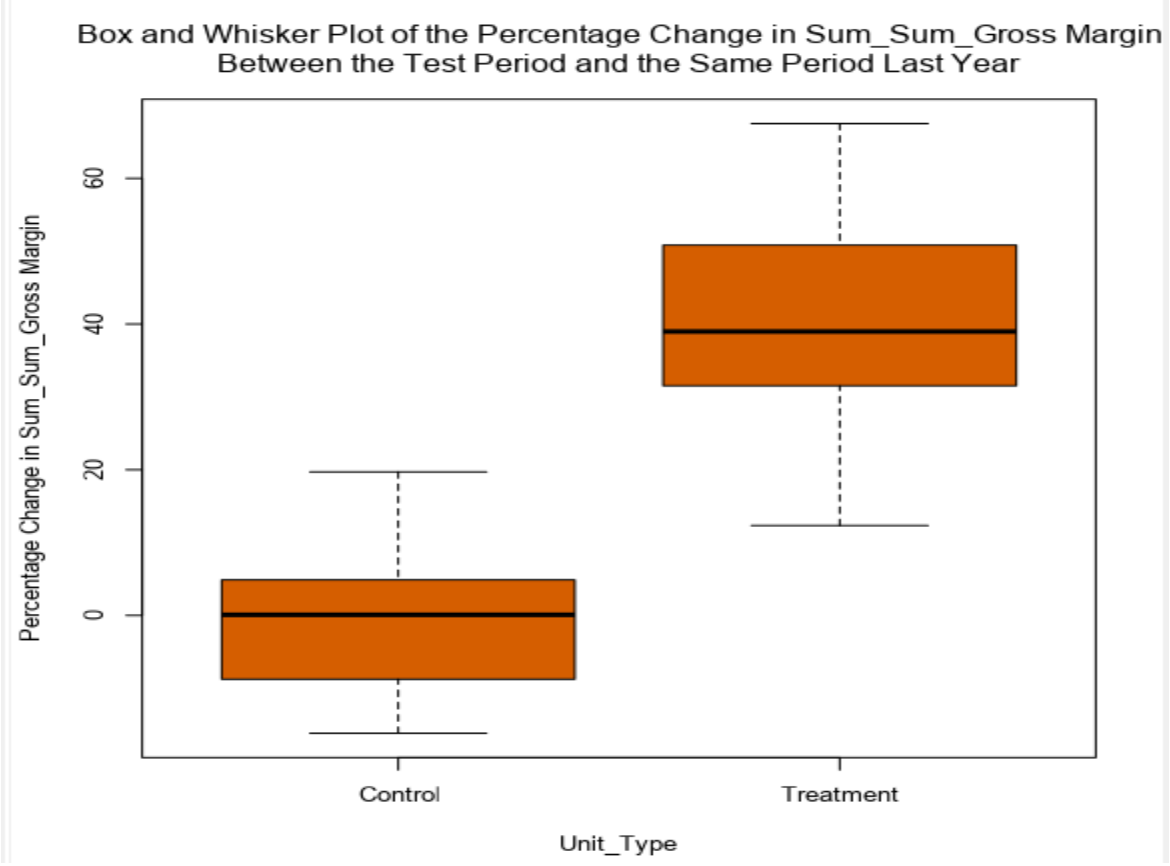
   - ✓ The A/B test results are conclusive: Round Roasters should roll out the updated menu to all stores. As shown below, the lift is greater than 18% and it is highly statistically significant.

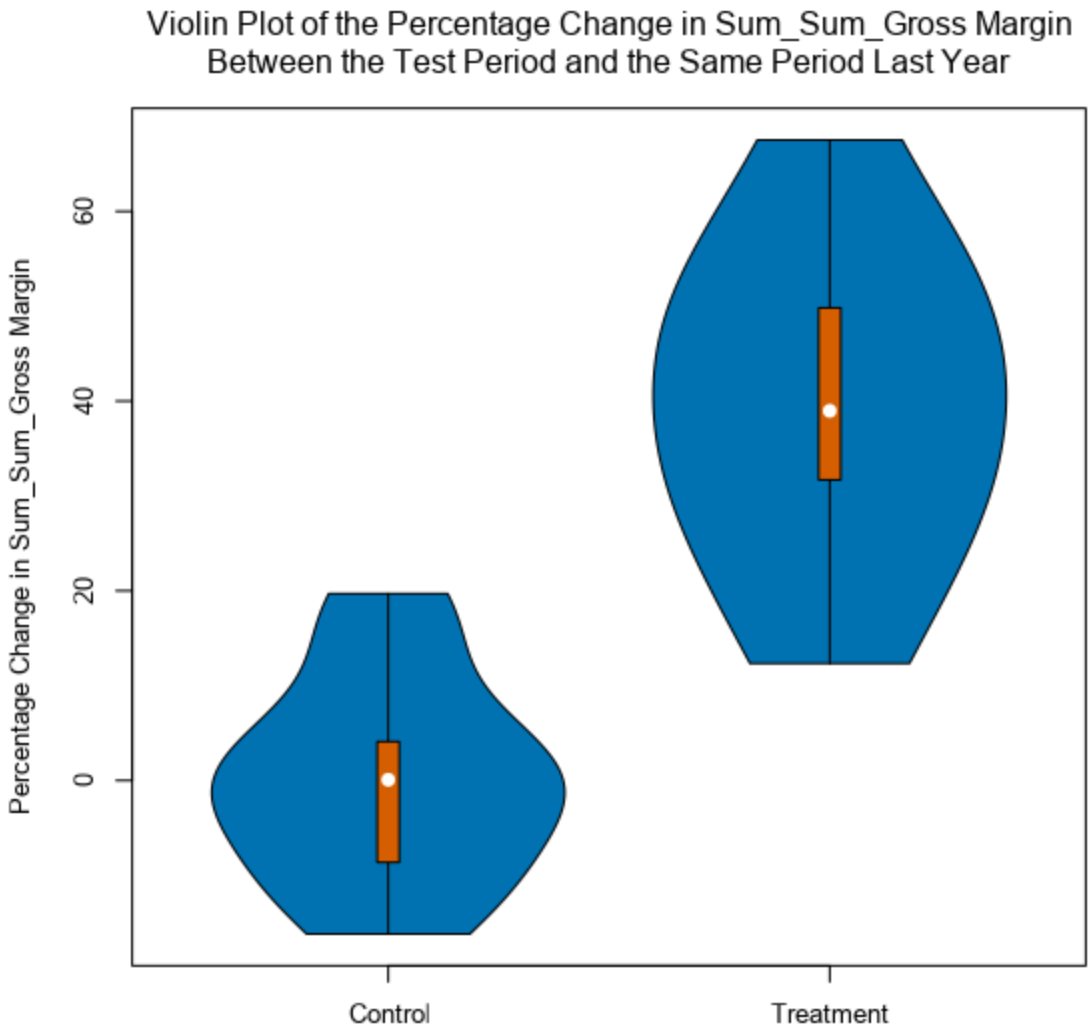


   - ✓ A comparison of the treatment-control pairs indicates an average lift in Sum_Gross Margin for the treatment units over the control units of 40.7%, which results in an expected impact of 681 on Sum_Gross Margin, with 100.0% of the treatment-control pairs exhibiting a positive lift for the treatment units.




Time Comparison Plot of Sum_Sum_Gross Margin
% Difference from Comparison Period Average
Treatment
Control
100
50
0
Comparison Period
Test Period
2015-05-07
2015-05-14
2015-05-21
2015-05-28
2015-06-04
2015-06-11
2015-06-18
2015-06-25
2015-07-02
2015-07-09
2015-07-16
2015-07-23
2016-05-05
2016-05-12
2016-05-19
2016-05-26
2016-06-02
2016-06-09
2016-06-16
2016-06-23
2016-06-30
2016-07-07
2016-07-14
2016-07-21
Time


Box and Whisker Plot of the Percentage Change in Sum_Sum_Gross Margin
Between the Test Period and the Same Period Last Year
Percentage Change in Sum_Sum_Gross Margin
0
20
40
60
Control
Treatment
Unit_Type

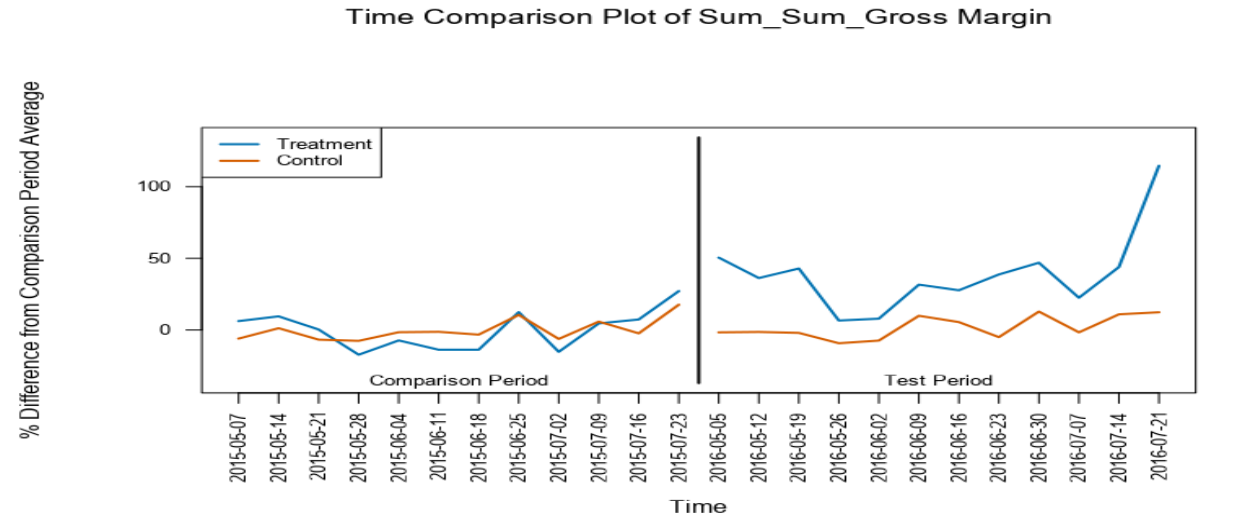

Violin Plot of the Percentage Change in Sum_Sum_Gross Margin
Between the Test Period and the Same Period Last Year
Percentage Change in Sum_Sum_Gross Margin
0
20
40
60
Control
Treatment

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

   ✓ For the West region, I got the following results:

Lift Analysis for Sum_Sum_Gross Margin

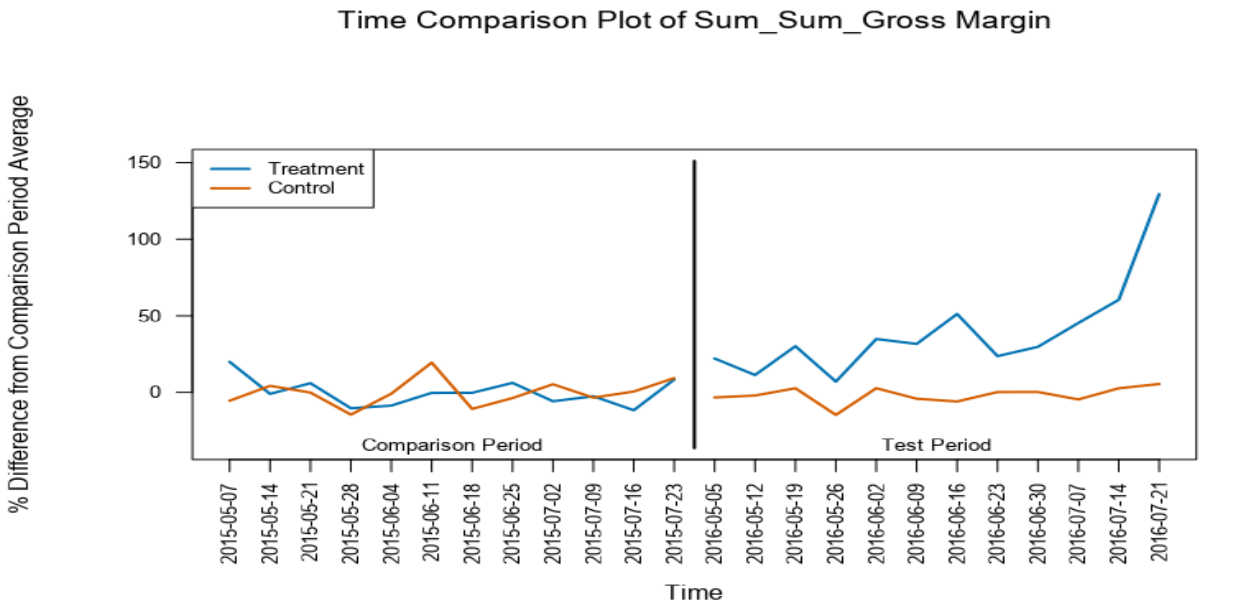
Lift	Expected Impact	Significance Level
37.9%	527	99.5%



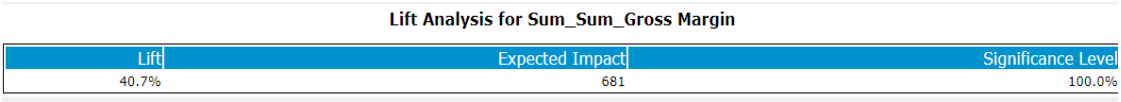
   ✓ For the Central region, I got the following results:

Lift Analysis for Sum_Sum_Gross Margin

Lift	Expected Impact	Significance Level
43.5%	836	99.6%



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



Lift Analysis for Sum_Sum_Gross Margin

Lift	Expected Impact	Significance Level
40.7%	681	100.0%

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.