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:

- What is the performance metric you'll use to evaluate the results of your test?
  - The performance metric I'll use to evaluate the results of my test is gross margin sales.
- What is the test period?
  - It is from 29/04/2016 to 21/07/2016. It's about 12 weeks
- At what level (day, week, month, etc.) should the data be aggregated?
  - It should the data be aggregated by 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.

I prepare the data for make training test. I aggregated the transaction data. After aggregation, I have 3 datasets. They are weekly_invoices.yxdb, weekly_sales.yxdb, and union1.yxdb

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...

- What control variables should be considered? Note: Only consider variables in the RoundRoastersStore file.
  - The control variables should be considered are Sq_Ft, AvgMonthSales, and Region
- What is the correlation between your each potential control variable and your performance metric?

Pearson Correlation Analysis

Full Correlation Matrix

	Sq_Ft	Sum_Gross.Margin
Sq_Ft	1.000000	-0.020353
Sum_Gross.Margin	-0.020353	1.000000

Pearson Correlation Analysis

Full Correlation Matrix

	AvgMonthSales	Sum_Gross.Margin
AvgMonthSales	1.00000	0.98822
Sum_Gross.Margin	0.98822	1.00000

As we see above, the correlation between Sq_Ft and Sum_Gross.Margin is -0.0203. It is negative weak correlation. On the other hand, there is a positive strong correlation between AvgMonthSales and Sum_Gross.Margin. It is 0.9882.

- What control variables will you use to match treatment and control stores?
  - Based on the correlations results, the control variable I'll use to match treatment and control stores is AvgMonthSales.
- 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	1863
1700	2014	1630
1712	8162	7434
2288	9081	2568
2293	12219	9524
2301	3102	9238
2322	2409	3235
2341	12536	2383

Step 4: Analysis and Writeup

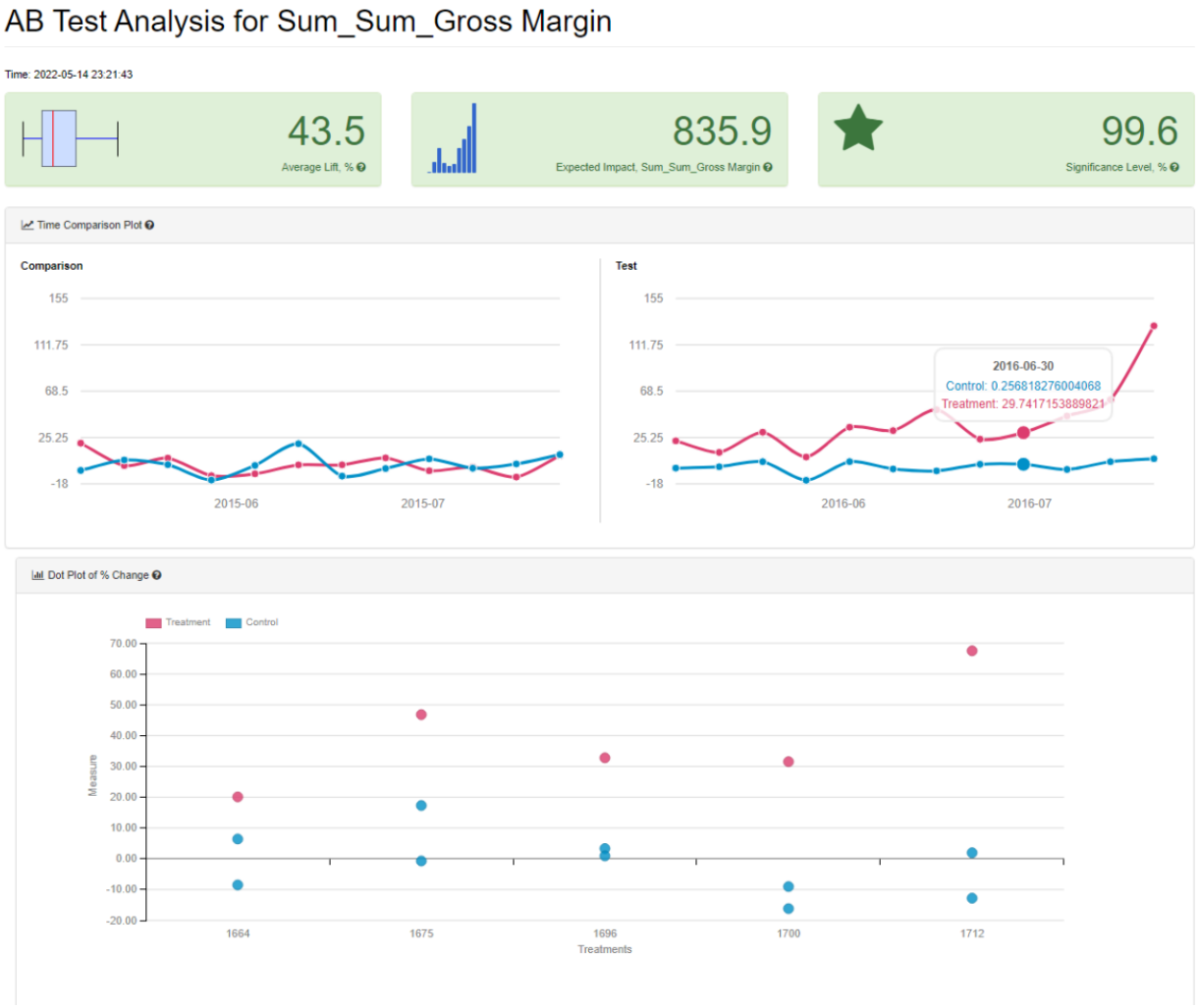
Conduct your A/B analysis and create a short report outlining your results and recommendations. (250 words limit)

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

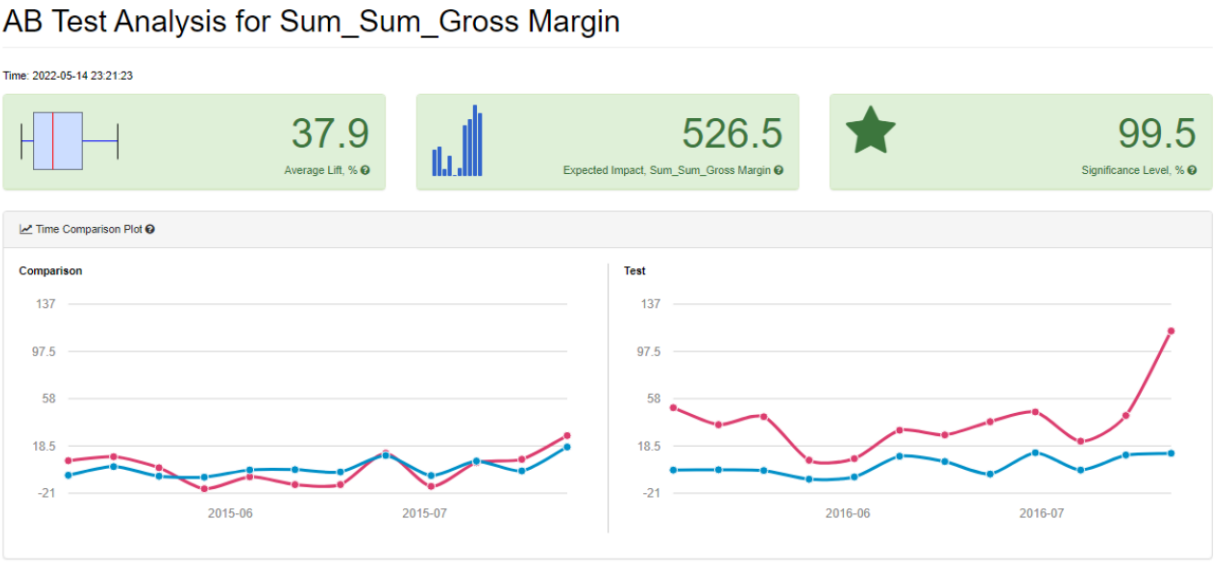
1. What is your recommendation - Should the company roll out the updated menu to all stores?

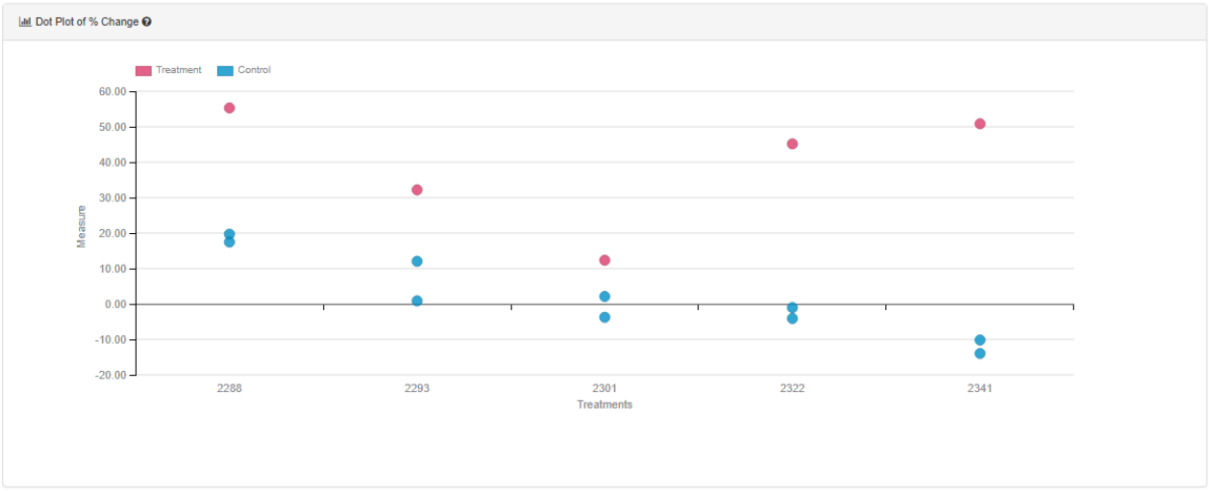
As we see from predictive results, the gross margin was increased by using the new menu. So, I recommend the company should roll out the updated menu to all stores.

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



As we see above, the average lift from the new menu for Central region is 43.5% and the significance level is 99.6%.

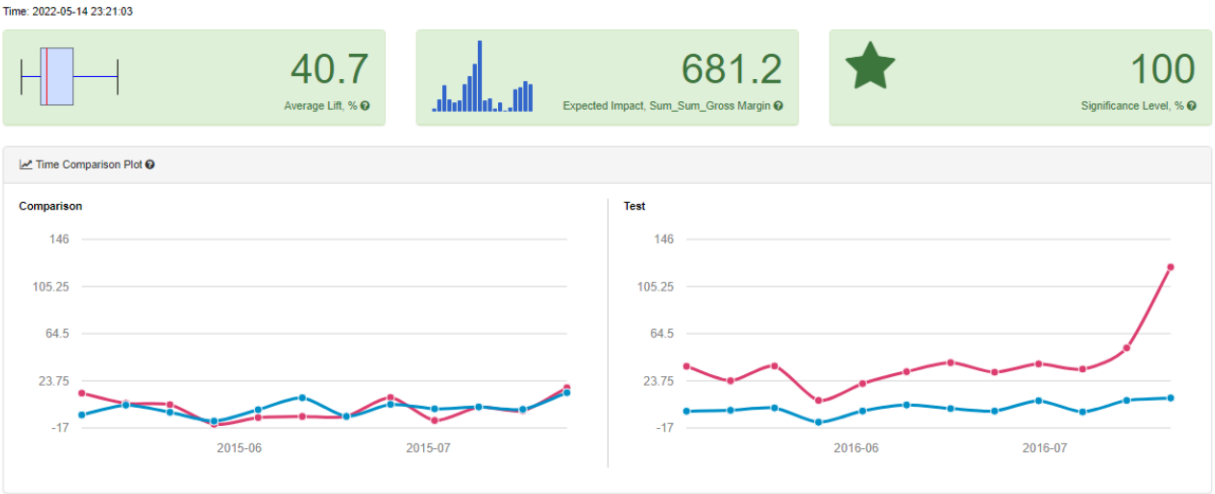


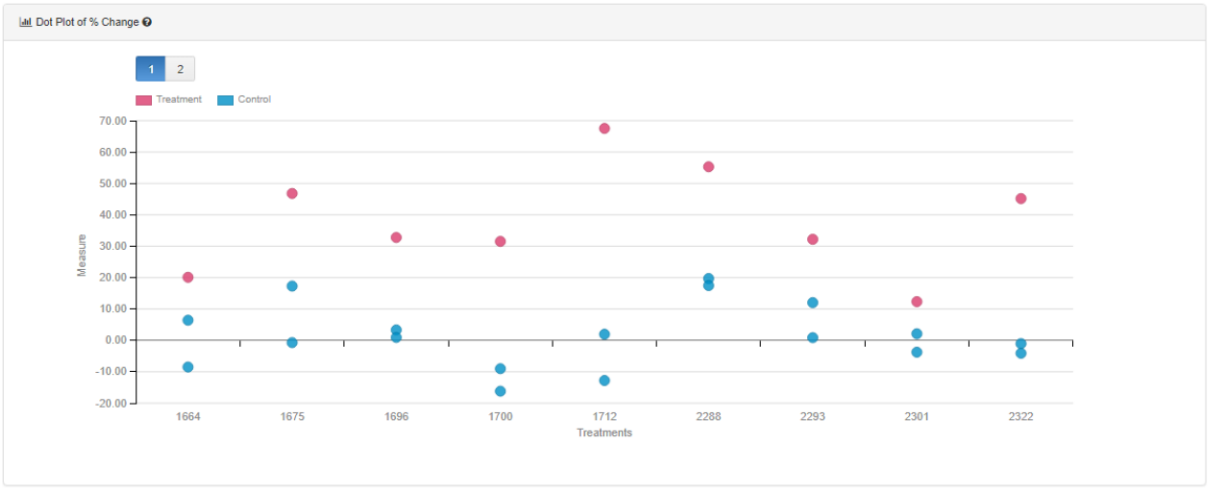


As we see above, the average lift from the new menu for West region is 37.9% and the significance level is 99.5%

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

AB Test Analysis for Sum_Sum_Gross Margin





As we see above, the average lift from the new menu for overall is 40.7% and the significance level is 100%

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.