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?

We want an 18% increase in profit when comparing the test results with the comparative period, in the data the profit is represented in the variable **Gross_margin**

2. What is the test period?

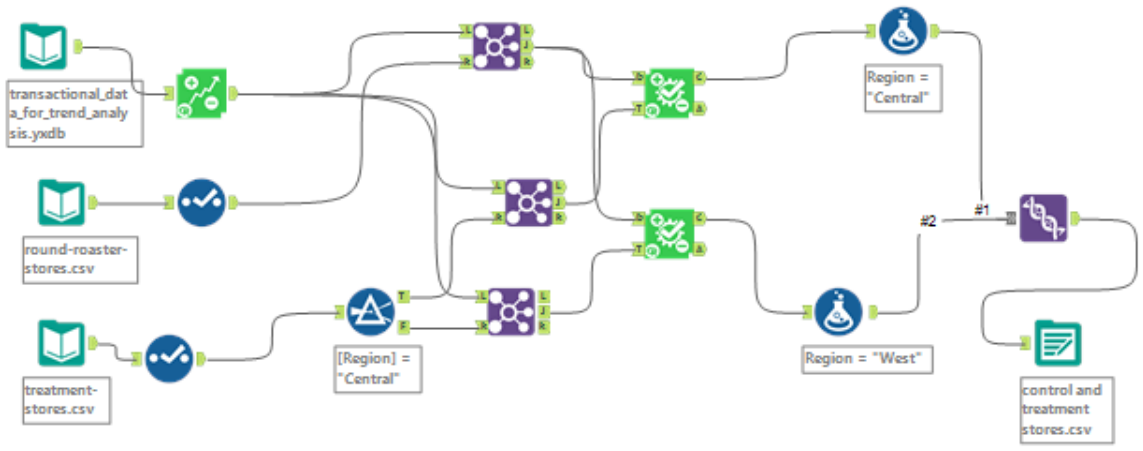
The menu test lasts 12 weeks (April 29 to July 21, 2016, for the analysis we must calculate the trend and seasonality to match the control and treatment units, in total the process requires 52 weeks of historical data plus another 6 periods of data in the upper part, since the test lasted a period of 12 weeks, each period was weekly, the weeks required for the analysis is as follows:

Historical Analysis: 52 weeks
Trend Analysis: 6 weeks
Testing: 12 weeks
Total: 70 weeks

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

   Data must be saved on a weekly period

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.

As unique identifier we will use Store Id, we must use The Sq_Ft, AvgMonthSales and probably state, region, time zone latitude and longitud

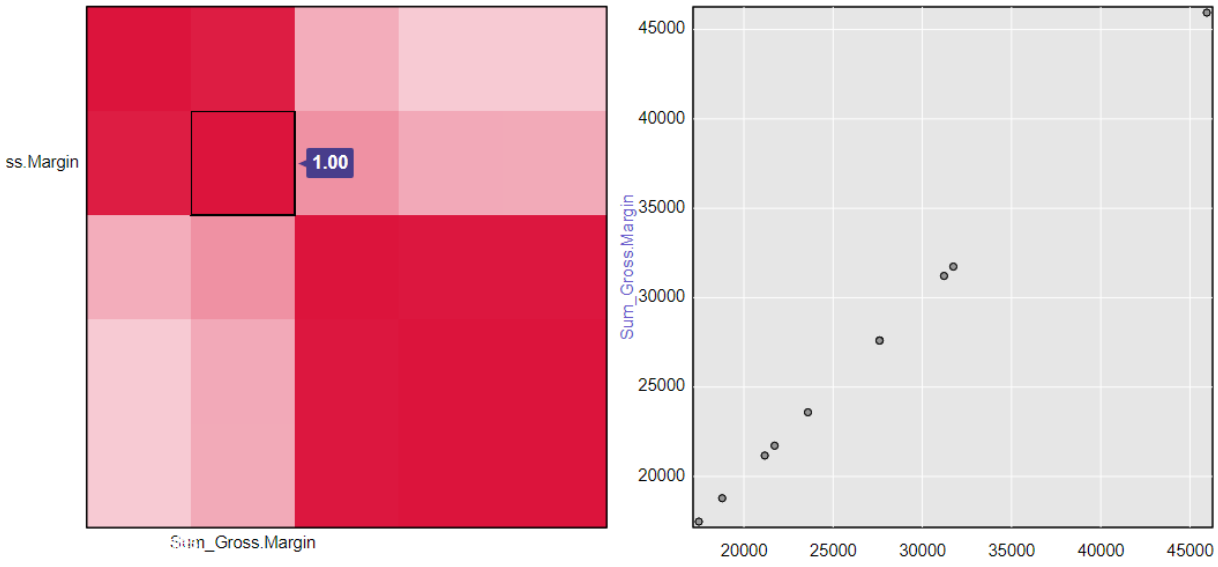
2. What is the correlation between your each potential control variable and your performance metric?

   We can deduce that there is a high correlation between the gross margin and AvgMonthSales, Sq Ft and with the variables that represent the region

Full Correlation Matrix

	Sq_Ft	AvgMonthSales	Latitude	Longitude	Sum_Gross.Margin
Sq_Ft	1.00000	0.22629	0.99999	0.98632	0.36553
AvgMonthSales	0.22629	1.00000	0.22428	0.34759	0.96221
Latitude	0.99999	0.22428	1.00000	0.98623	0.36320
Longitude	0.98632	0.34759	0.98623	1.00000	0.47003
Sum_Gross.Margin	0.36553	0.96221	0.36320	0.47003	1.00000

Correlation Matrix with ScatterPlot



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

The control variables we will use are AvgMonthSales and Region.

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

Treatment Store	Control Store 1	Control Store 2
1664	11868	12019
1675	3235	2409
1696	2383	7334
1700	8717	1508
1712	9017	7434
2288	7484	1857
2293	7811	7770
2301	1863	11268
2322	1807	7284
2341	8817	7584

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 expected impact on profits should be sufficient justification for the increase in marketing (greater than 18% compared to the comparative period vs control stores.
If we look at the results of the average increase, we can say that the company has to implement the updated menu

Lift Analysis for Sum_Gross Margin

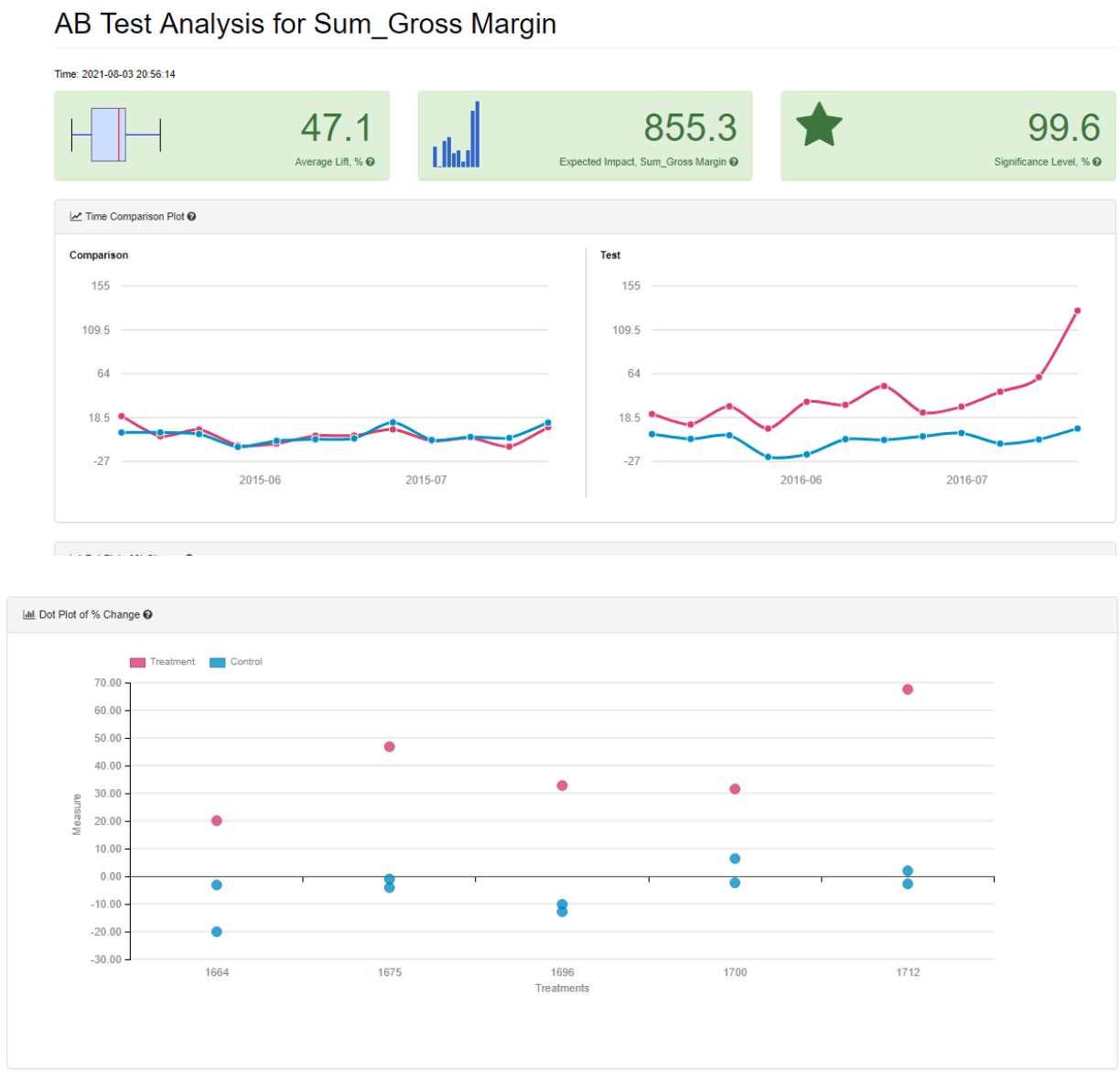
Lift	Expected Impact	Significance Level
40.2%	660	100.0%

Summary Statistics for Sum_Gross Margin by Test Group

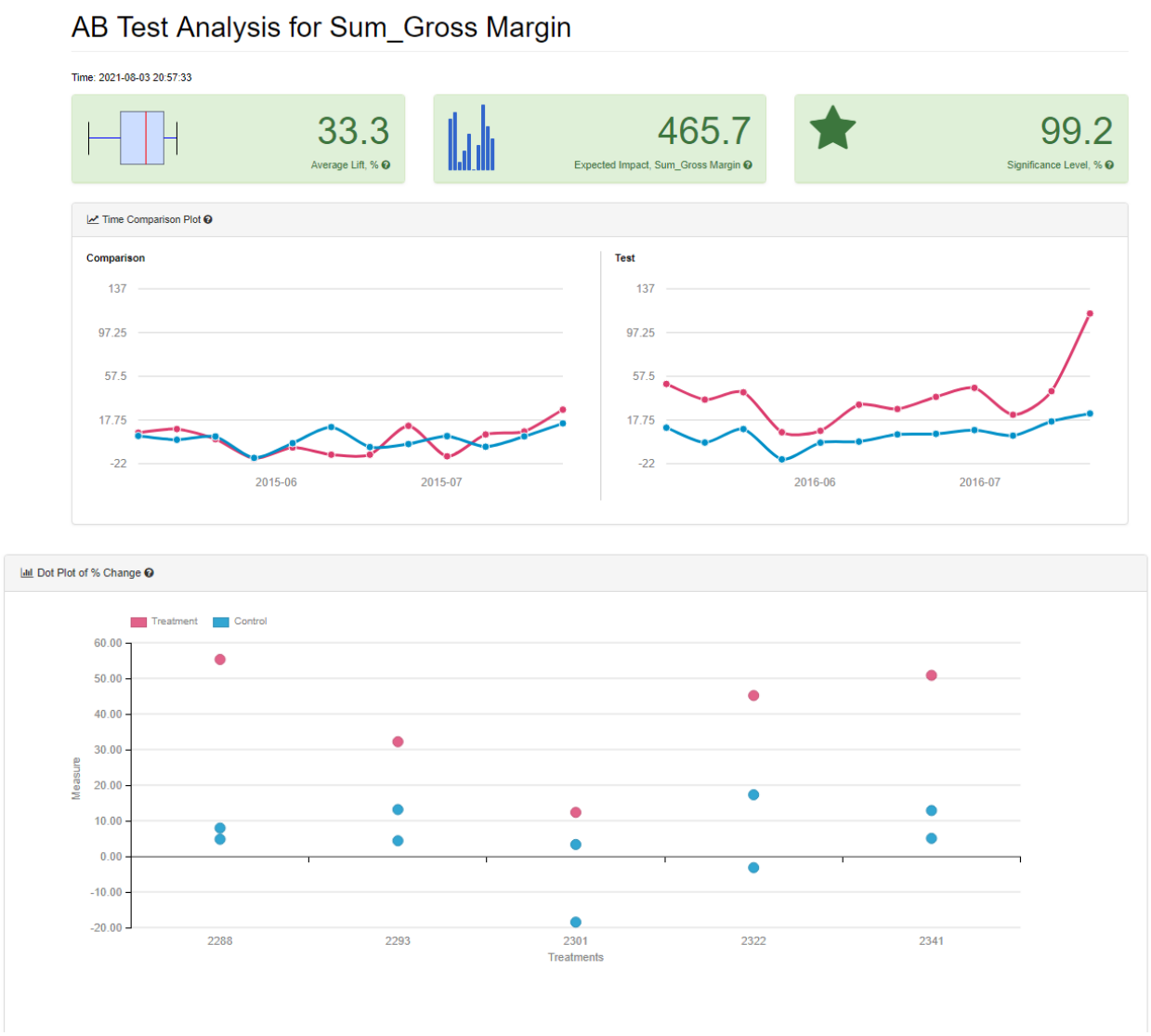
Statistic	Treatment	Control
Average	39.45	-0.05
Minimum	12.34	-20.05
Maximum	67.52	17.29
Standard Deviation	16.30	9.97

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

For the CENTRAL region, the expected increase is 47.1%, p-value of 0.003938. with a 99.6% level of significance, statistically significant.



For the **WEST REGION**, the expected increase is 33.3% and a p-value of 0.007863, with the significance level of 99.2%



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

The increase in the new menu is 40.2%.

Lift Analysis for Sum_Gross Margin

Lift	Expected Impact	Significance Level
40.2%	660	100.0%

Summary Statistics for Sum_Gross Margin by Test Group

Statistic	Treatment	Control
Average	39.45	-0.05
Minimum	12.34	-20.05
Maximum	67.52	17.29
Standard Deviation	16.30	9.97

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.