# Project: Analyzing a Market Test

Complete each section. When you are ready, save your file as a PDF document and submit it [here](https://classroom.udacity.com/nanodegrees/nd008/parts/11a7bf4c-2b69-47f3-9aec-108ce847f855/project).

## 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.

1. What is the test period?

The test period is 12 weeks between 2016-04-29 to 2016-07-21.

1. 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.*

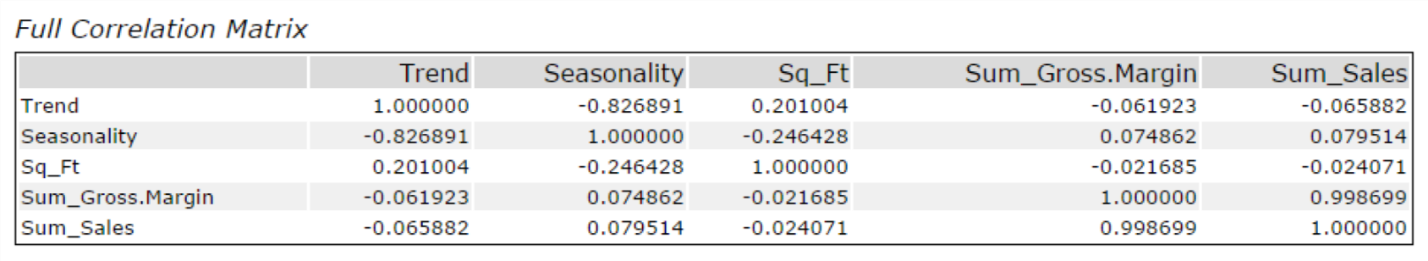
*Apart from trend and seasonality...*

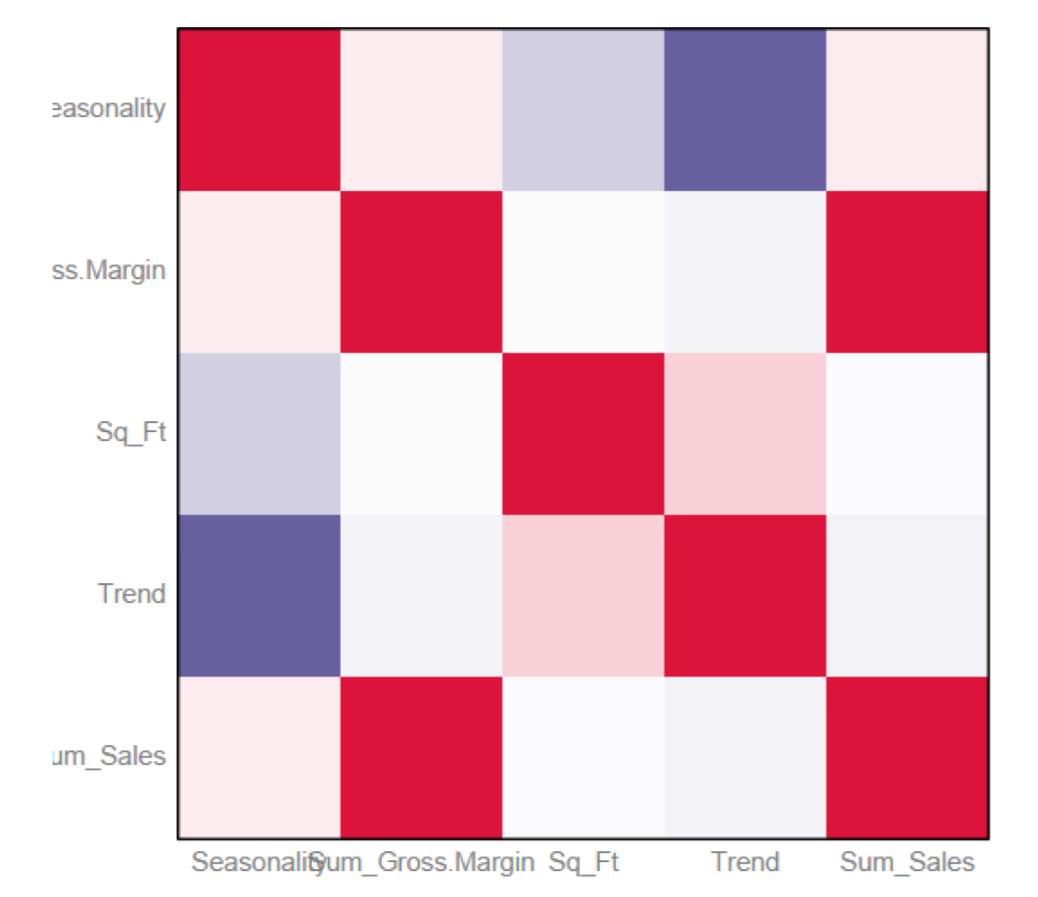
1. What control variables should be considered? Note: Only consider variables in the RoundRoastersStore file.

We shall only consider Sq\_Ft and AvgMonthSales, all other variables are geographic, and we have already incorporated the region variable.

1. 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.

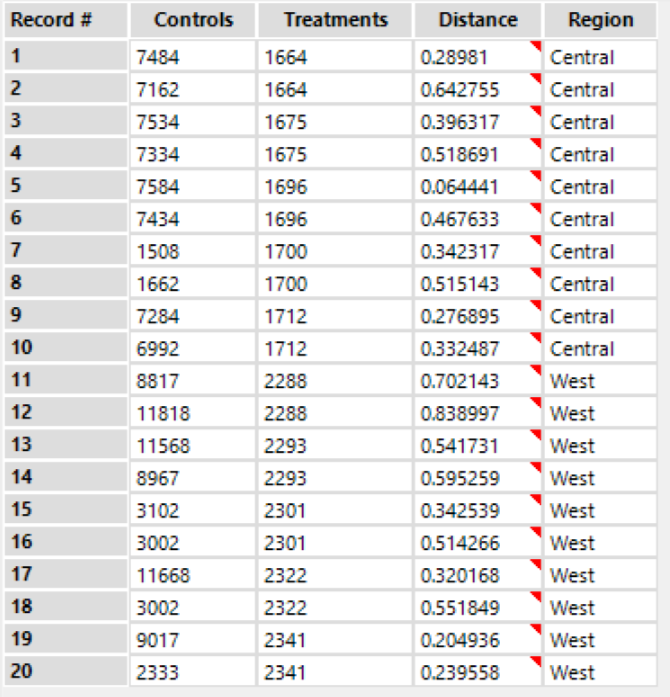




1. What control variables will you use to match treatment and control stores?

Finally, we choose Trend, Seasonality and Sq\_Ft as control variables.

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



|  |  |  |
| --- | --- | --- |
| 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)*

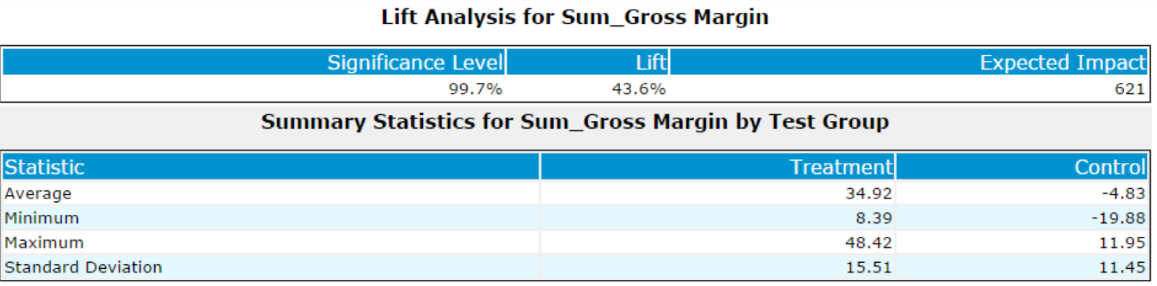
*Answer these questions. Be sure to include visualizations from your analysis:*

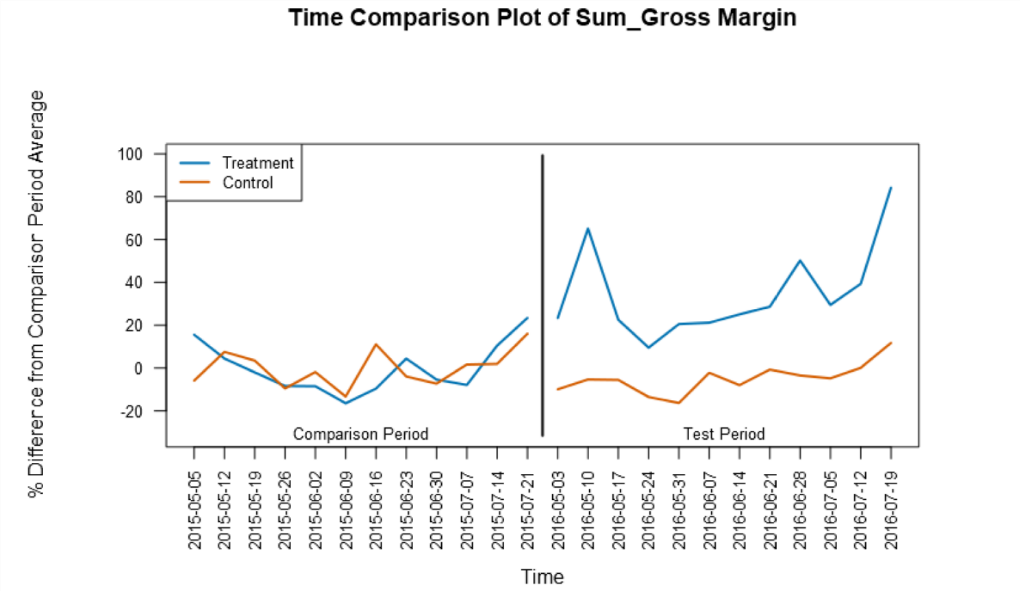
1. What is your recommendation - Should the company roll out the updated menu to all stores?

The company should roll out the updated menu to all stores.

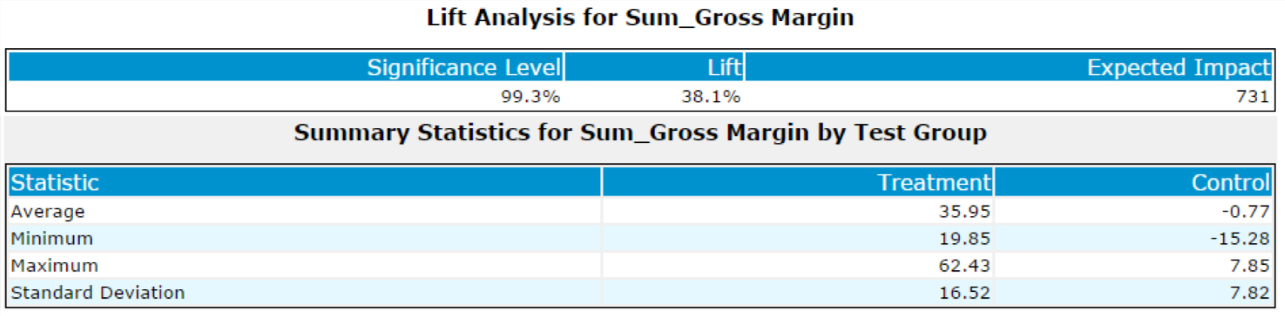
1. What is the lift from the new menu for West and Central regions (include statistical significance)?

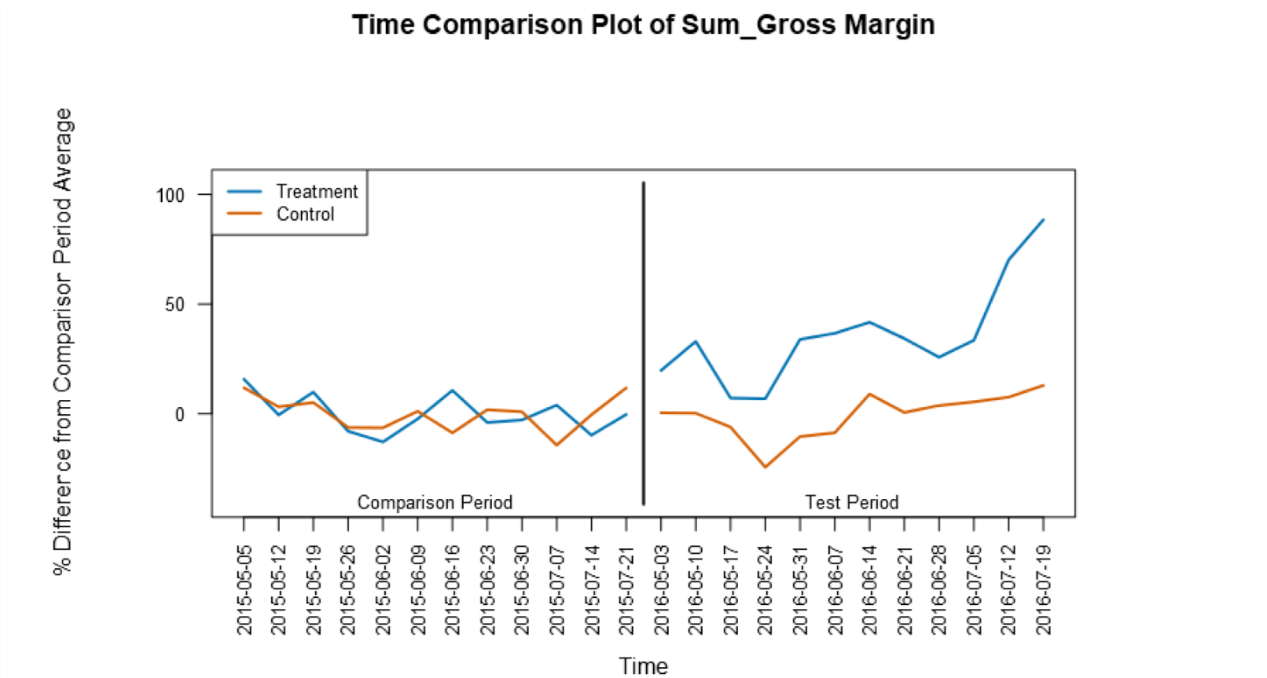
WEST: 43.6% lift, 99.7% significance





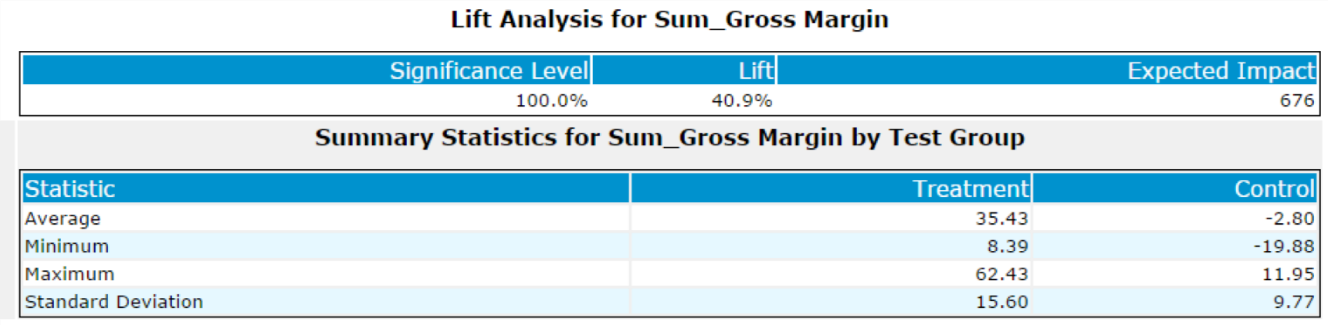
CENTRAL:38.1% lift, 99.3% significance

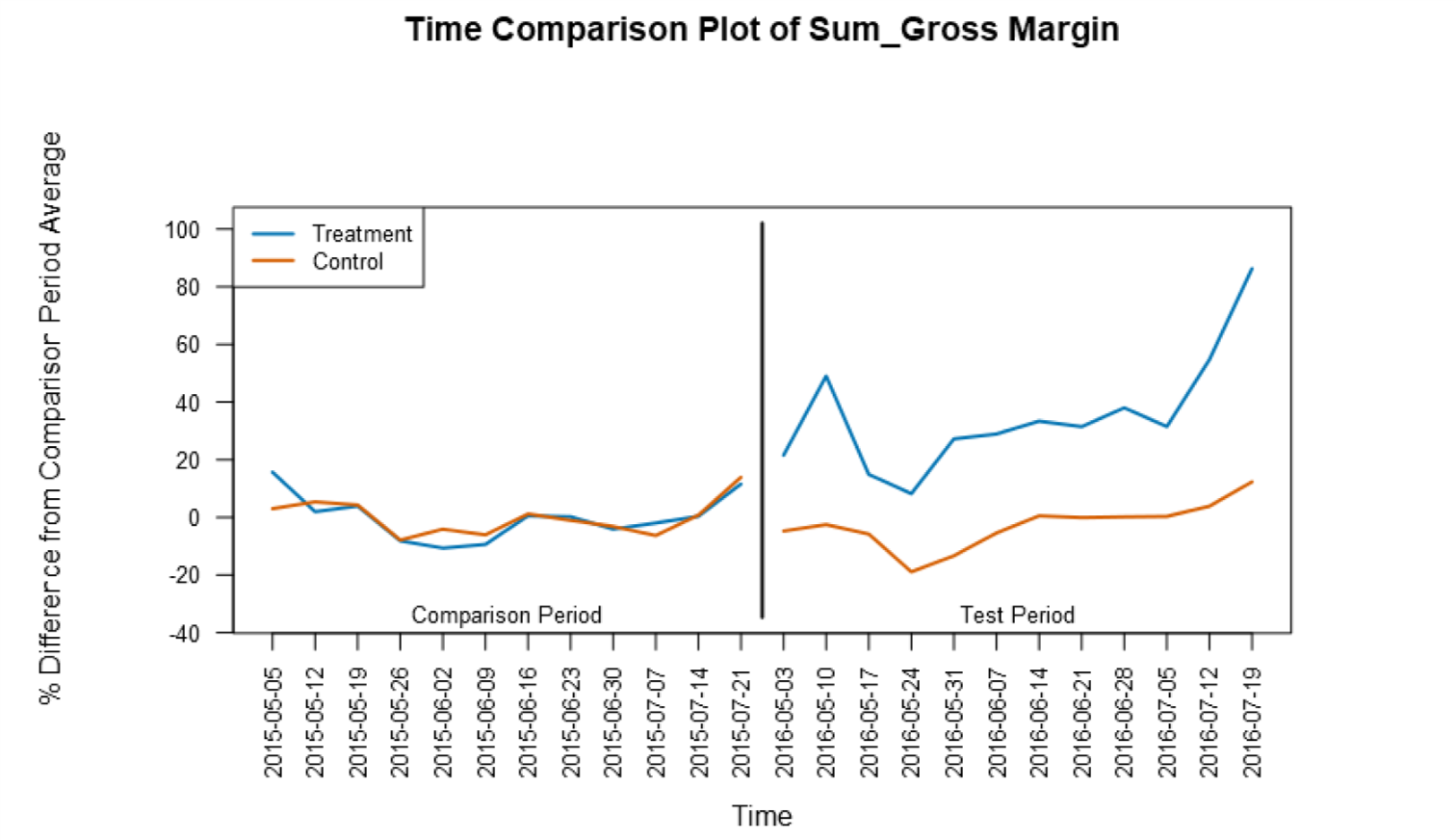




1. What is the lift from the new menu overall?

Overall Lift: 40.9%. Overall significance: 100.0%.





## Before you Submit

Please check your answers against the requirements of the project dictated by the [rubric](https://review.udacity.com/#!/rubrics/287/view) here. Reviewers will use this rubric to grade your project.