# 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 performance metric which interests us is the profit, as a result of the new menu implementation. Profit is represented by the **gross_margin** variable.

2. What is the test period?

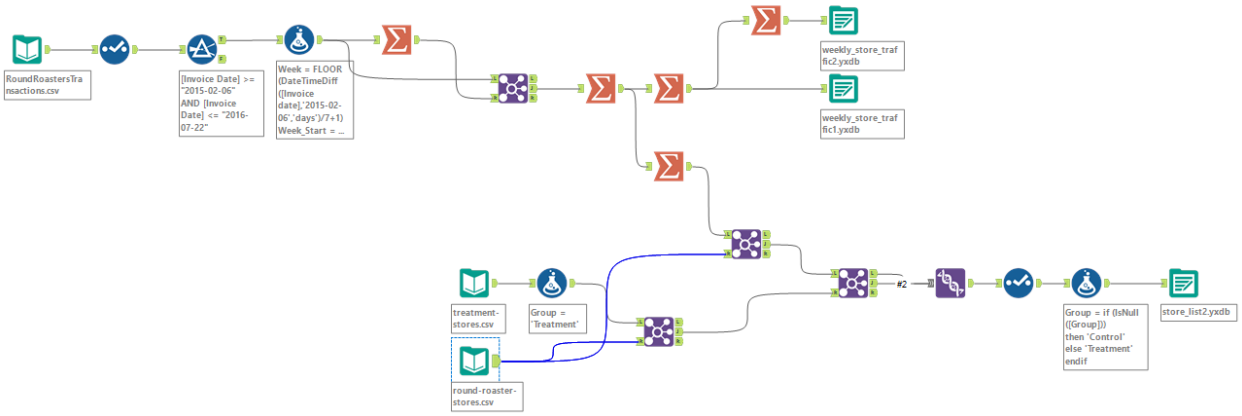
The test period consists of 12 weeks, between 2016-April-29 and 2016-July-21.

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

I aggregated it to weeks because the week is considered the time unit in which we will group our data and analysis.

## 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...*

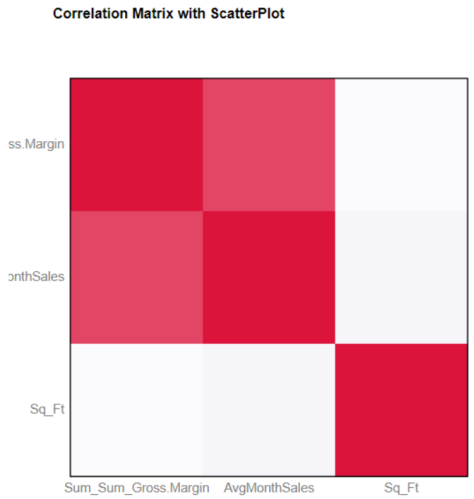
1. What control variables should be considered? Note: Only consider variables in the RoundRoastersStore file.

The AvgMonthlySales and Sq_Ft variables are considered our control variables; furthermore, they are the only numeric variables in RoundRoastersStore file. Another control variable which we use in our data filtering and analysis is Region.

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

We discovered by using Correlation Analysis Tool that:

1) There is clear connection between AvgMonthlySales and our performance metric (gross_margin), so we should use it as a control variable;
2) Sq_Ft variable should be excluded as the as the correlation value is negative.

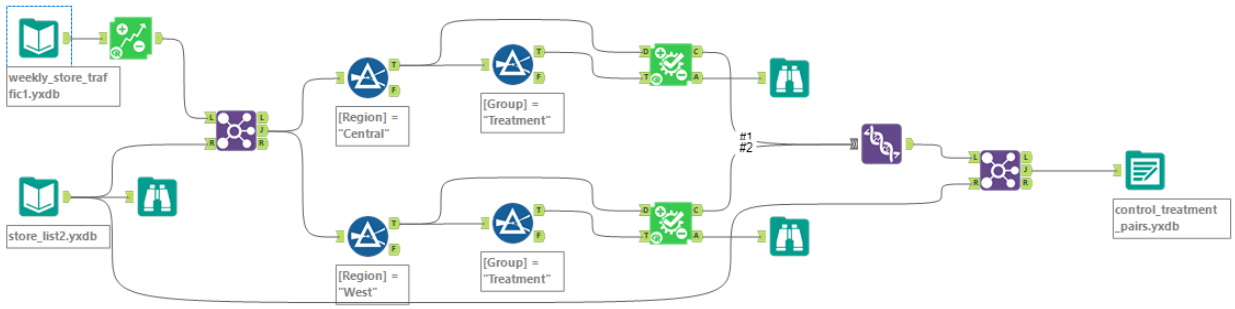



3. What control variables will you use to match treatment and control stores?

The control variables considered were trend and seasonality. Apart from those, I added also AvgMonthSales as it may have statistical significance in the overall match pairing selections.

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 | 1863 | 1964 |
| 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?

Taking into consideration that the incremental lift overweighs significantly the expected 18% increase in the marketing budget, it justifies the implementation of the new menu offerings. Both regions benefit from this update.

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

Central region has a lift of 43.5%, while West region has 37.9%.
For Central region, the p-Value is 0.004491 < 0.05, which expresses a meaningful statistical significance between the variables.
West region displays a p-Value of 0.005225.

**Central:**

**Lift Analysis for Sum_Sum_Gross Margin**

| Lift | Expected Impact | Significance Level |
|---|---|---|
| 43.5% | 836 | 99.6% |

**Welch's Two Sample t-test(s) of Sum_Sum_Gross Margin by Test Group**

| Test | t-Statistic | Degrees of Freedom | p-Value |
|---|---|---|---|
| Control vs Treatment | -4.74884 | 5.25449 | 0.004491 |

**West:**

**Lift Analysis for Sum_Sum_Gross Margin**

| Lift | Expected Impact | Significance Level |
|---|---|---|
| 37.9% | 527 | 99.5% |

**Welch's Two Sample t-test(s) of Sum_Sum_Gross Margin by Test Group**

| Test | t-Statistic | Degrees of Freedom | p-Value |
|---|---|---|---|
| Control vs Treatment | -4.36941 | 5.74747 | 0.005225 |

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

According to the calculations, the overall lift is 40.7%.

**Lift Analysis for Sum_Sum_Gross Margin**

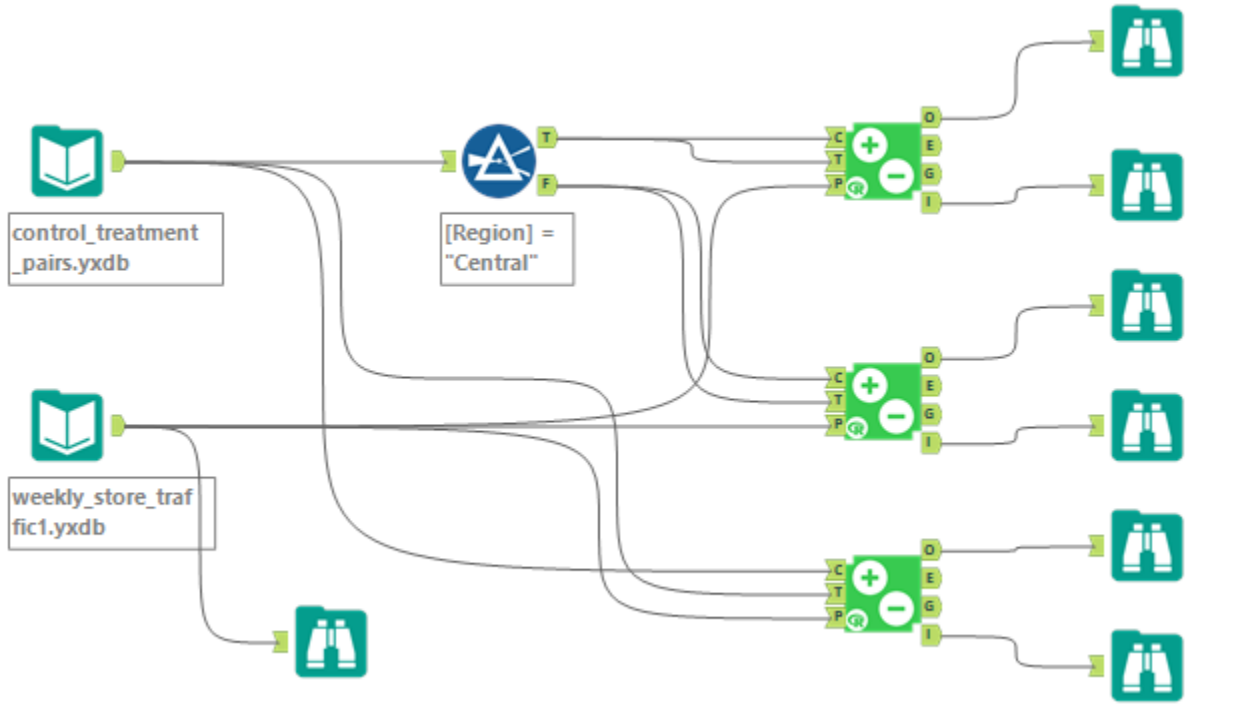
| Lift | Expected Impact | Significance Level |
|---|---|---|
| 40.7% | 681 | 100.0% |

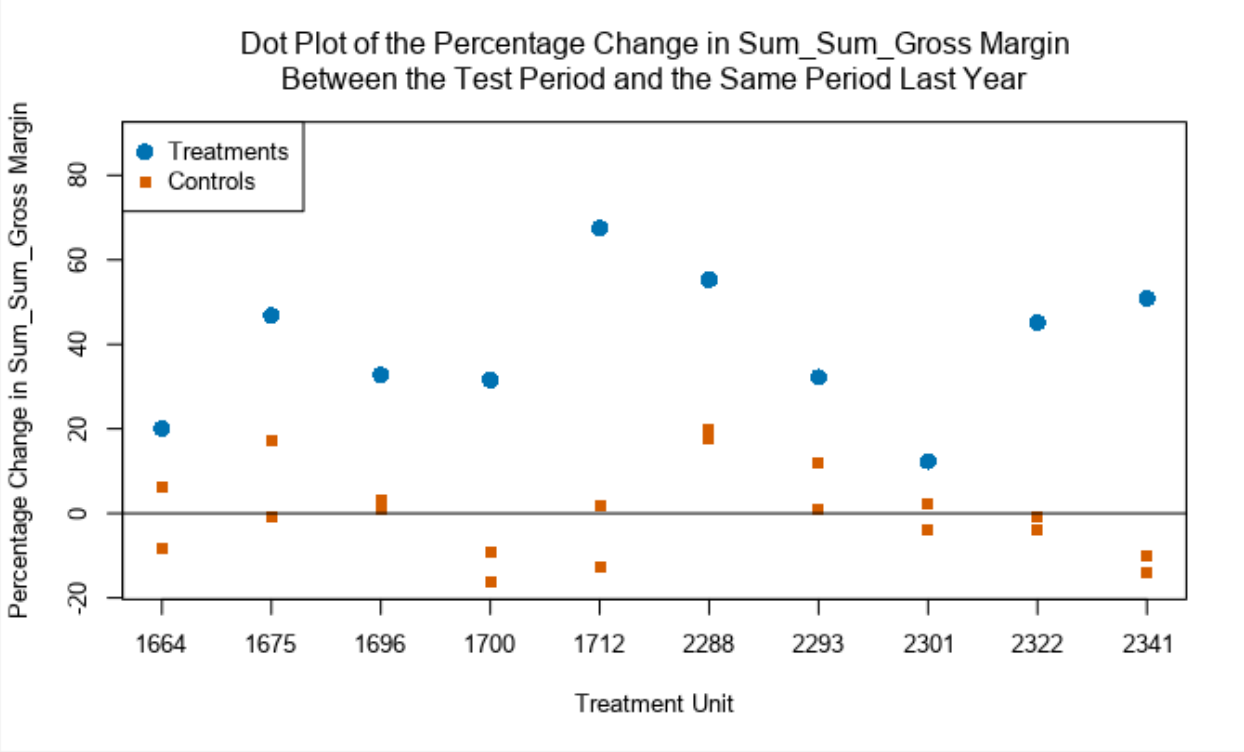
**Summary Statistics for Sum_Sum_Gross Margin by Test Group**

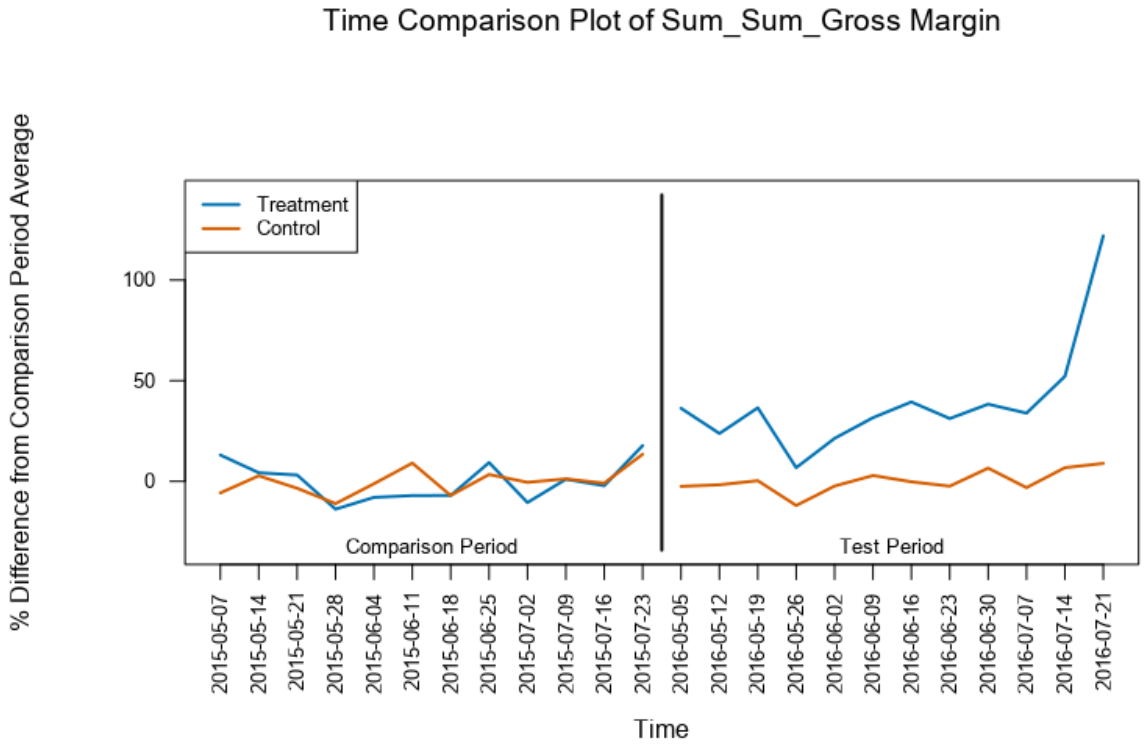
| Statistic | Treatment | Control |
|---|---|---|
| Average | 39.45 | 0.09 |
| Minimum | 12.34 | -16.18 |
| Maximum | 67.52 | 19.70 |
| Standard Deviation | 16.30 | 10.54 |

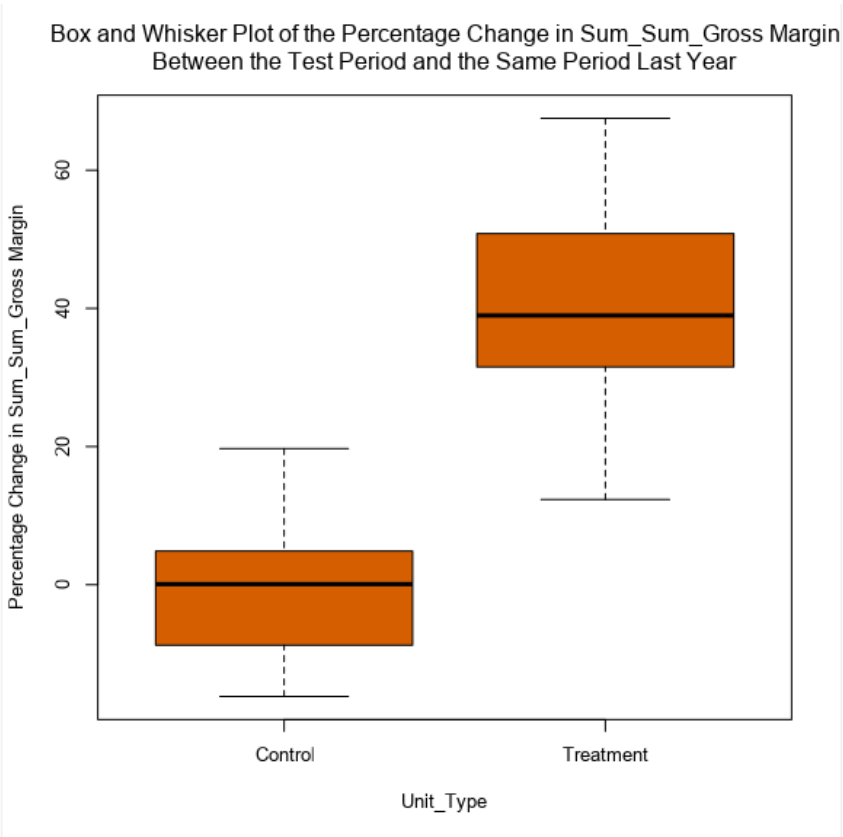
**Welch's Two Sample t-test(s) of Sum_Sum_Gross Margin by Test Group**

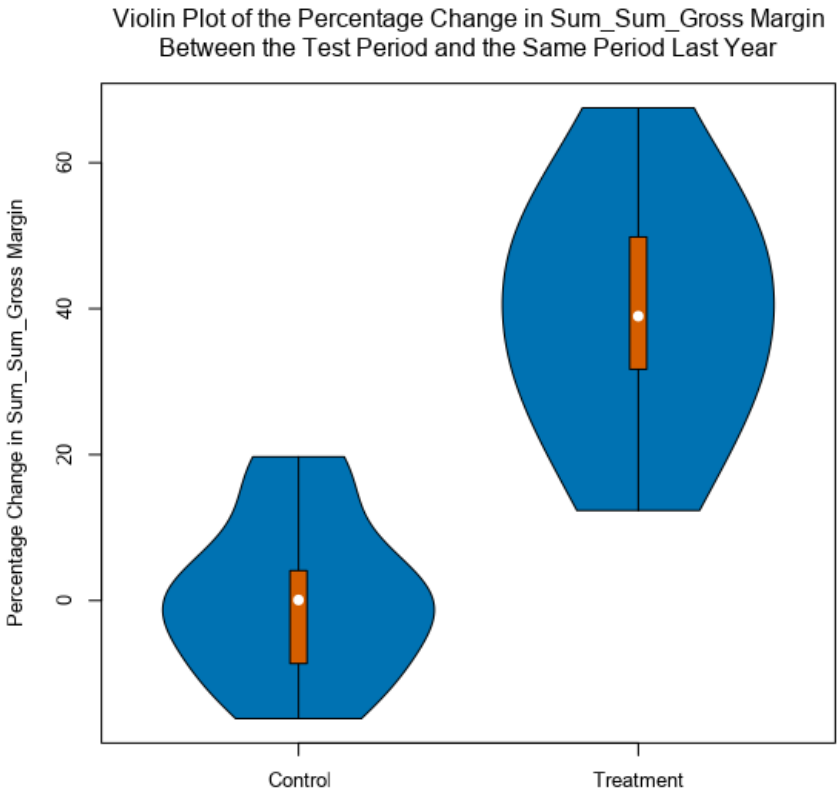
| Test | t-Statistic | Degrees of Freedom | p-Value |
|---|---|---|---|
| Control vs Treatment | -6.78894 | 12.67716 | 0.000015 |

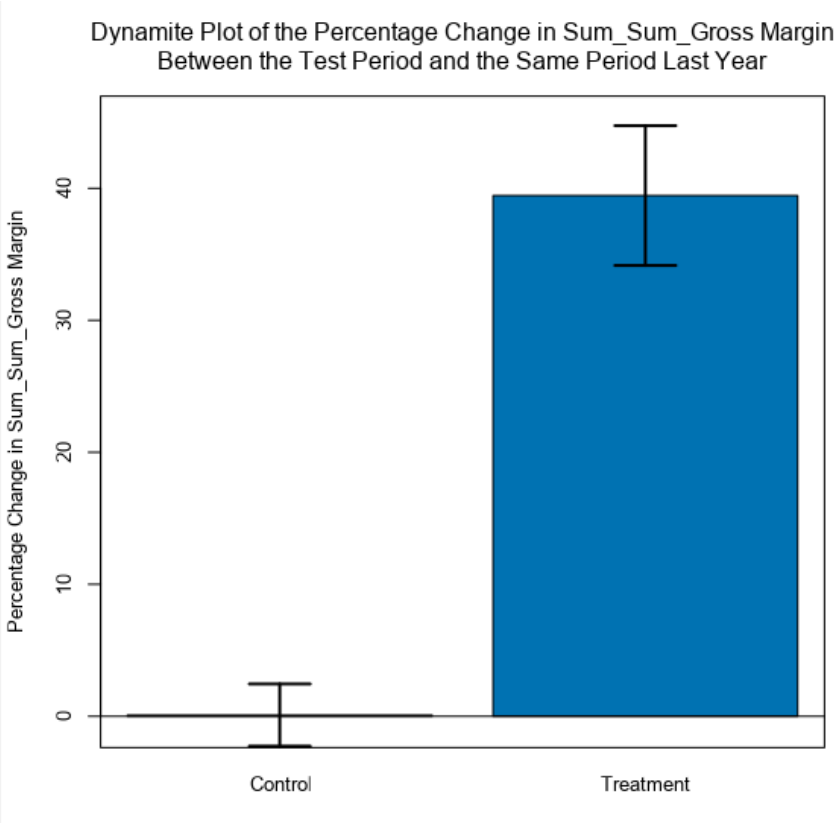



Dot Plot of the Percentage Change in Sum_Sum_Gross Margin
Between the Test Period and the Same Period Last Year
Treatments
Controls
Percentage Change in Sum_Sum_Gross Margin
-20
0
20
40
60
80
1664
1675
1696
1700
1712
2288
2293
2301
2322
2341
Treatment Unit


Time Comparison Plot of Sum_Sum_Gross Margin
Treatment
Control
% Difference from Comparison Period Average
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




## 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.