

# 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 sum of gross margin will be used to determine the results of the A/B test analysis. Based on the outcome of the analysis Round Roasters will decide on the introduction of a new menu selection across all stores.

### **2. What is the test period?**

The test ran for a period of 12 weeks (2016-April-29 to 2016-July-21)

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

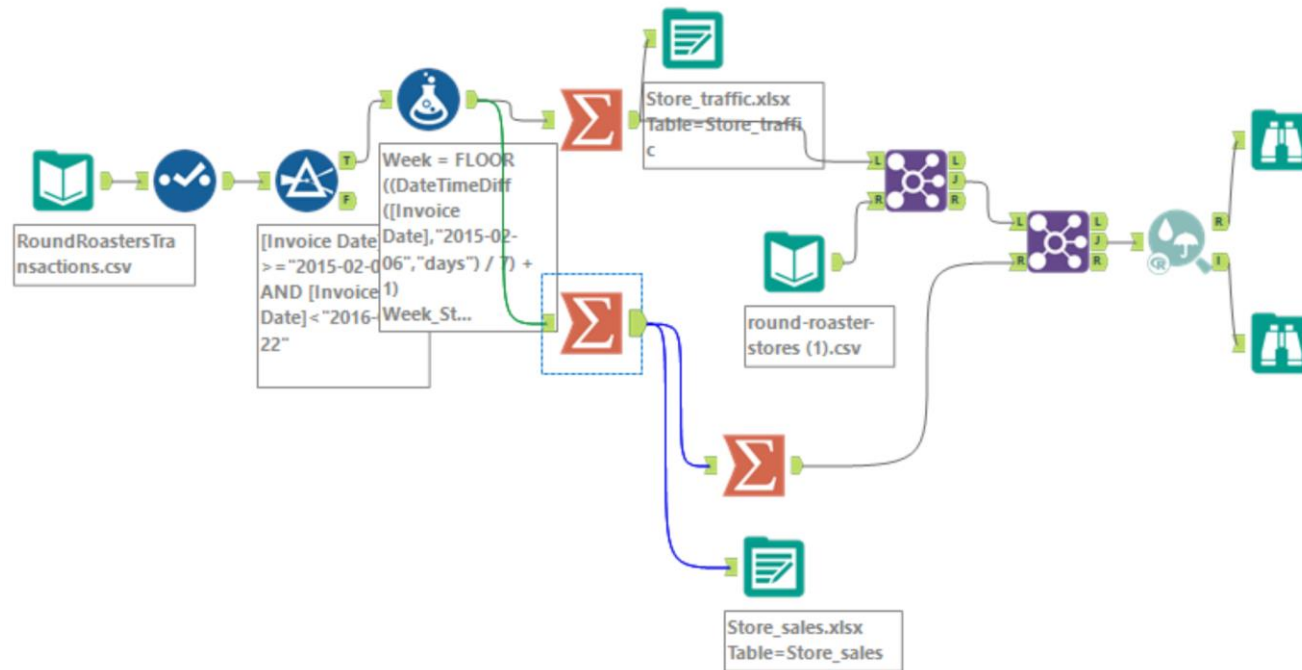
Weekly data is aggregated for this A/B test analysis.

## 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.*

There is no missing, incomplete, duplicate, or dirty data

| StoreID | Week | Week_Start | Week_End   | Count | StoreID | Week | Week_Start | Week_End   | Sum_Gross Margin |
|---------|------|------------|------------|-------|---------|------|------------|------------|------------------|
| 10018   | 1    | 2015-02-06 | 2015-02-12 | 308   | 10018   | 1    | 2015-02-06 | 2015-02-12 | 2212.7105        |
| 10018   | 2    | 2015-02-13 | 2015-02-19 | 288   | 10018   | 2    | 2015-02-13 | 2015-02-19 | 2164.007         |
| 10018   | 3    | 2015-02-20 | 2015-02-26 | 204   | 10018   | 3    | 2015-02-20 | 2015-02-26 | 1560.929         |



Workflow for data analysis and preparation

### 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 will Utilize the AvgMonthlySales and sq\_ft as a control Variable in our analysis.

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

The Correlation Matrix shows that we have a strong correlation between Sum\_Gross\_Margin and the AvgMonthSales 0.99 while Sq\_Ft has a poor correlation of -0.024. We will Omit Sq\_Ft in our analysis.

#### Pearson Correlation Analysis

*Focused Analysis on Field Sum\_Gross.Margin*

|               | Association Measure | p-value      |
|---------------|---------------------|--------------|
| AvgMonthSales | 0.990982            | 0.000000 *** |
| Sq_Ft         | -0.024255           | 0.014743 *   |

*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      |

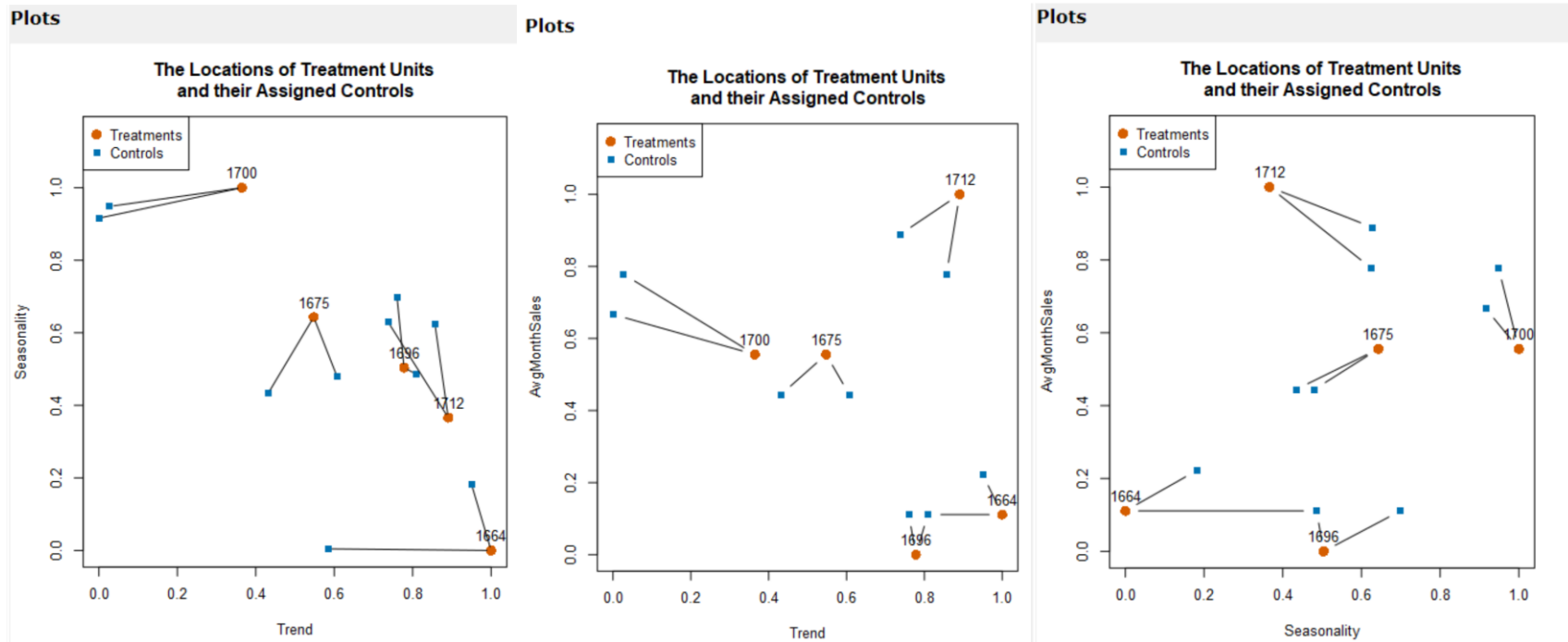
*Matrix of Corresponding p-values*

|                  | Sum_Gross.Margin | Sq_Ft      | AvgMonthSales |
|------------------|------------------|------------|---------------|
| Sum_Gross.Margin |                  | 1.4743e-02 | 0.0000e+00    |
| Sq_Ft            | 1.4743e-02       |            | 2.3119e-06    |
| AvgMonthSales    | 0.0000e+00       | 2.3119e-06 |               |

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

Trend, Seasonality and AvgMonthSales will be used to map out treatment and control stores.

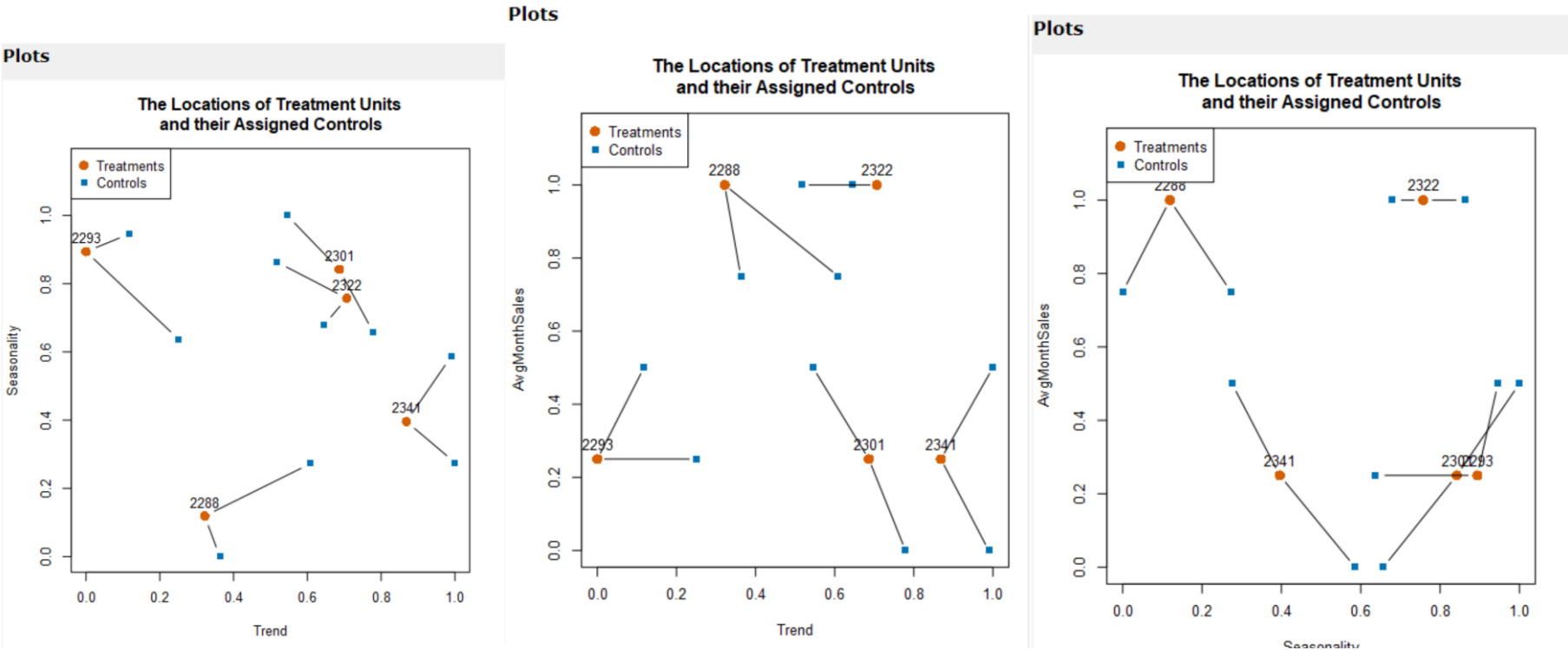
The Plots indicate the pairing of treatment and control units based on Seasonality, Trend and AvgMonthlySales for the Central Region.



Distance Between a Control Unit and its Assigned Treatment

| Control Unit | Assigned Treatment | Distance to Treatment |
|--------------|--------------------|-----------------------|
| 1964         | 1664               | 1.0249                |
| 7162         | 1664               | 0.4786                |
| 1580         | 1675               | 0.4563                |
| 1807         | 1675               | 0.5605                |
| 1863         | 1696               | 0.4891                |
| 1964         | 1696               | 0.3124                |
| 1630         | 1700               | 0.9162                |
| 2014         | 1700               | 0.8104                |
| 7434         | 1712               | 0.7933                |
| 8162         | 1712               | 0.6714                |

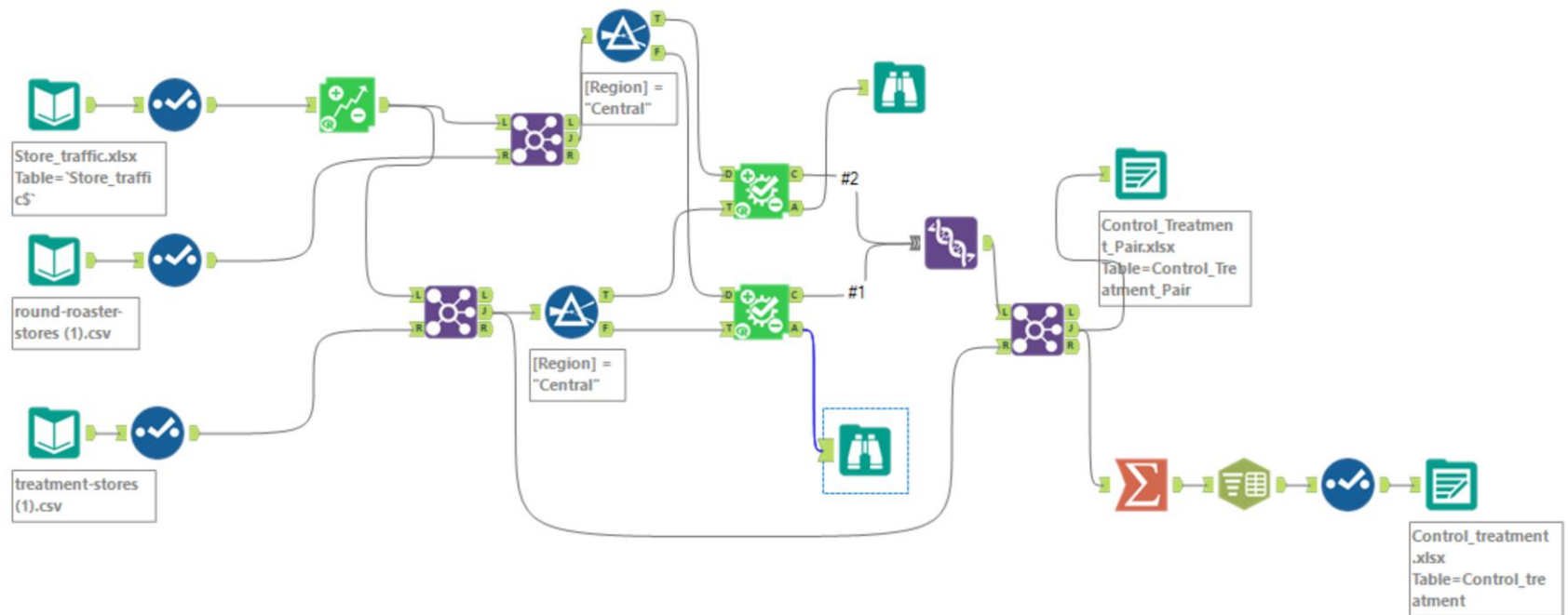
The Plots indicate the pairing of treatment and control units based on Seasonality, Trend and AvgMonthlySales for the Western Region.



Distance Between a Control Unit and its Assigned Treatment

| Control Unit | Assigned Treatment | Distance to Treatment |
|--------------|--------------------|-----------------------|
| 2568         | 2288               | 0.7141                |
| 9081         | 2288               | 0.2779                |
| 12219        | 2293               | 0.3486                |
| 9524         | 2293               | 0.6560                |
| 3102         | 2301               | 0.3812                |
| 9238         | 2301               | 0.4346                |
| 2409         | 2322               | 0.1714                |
| 3235         | 2322               | 0.4513                |
| 12536        | 2341               | 0.3980                |
| 2383         | 2341               | 0.4238                |

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



**Workflow for Control and Treatment store pairing**

| 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            |

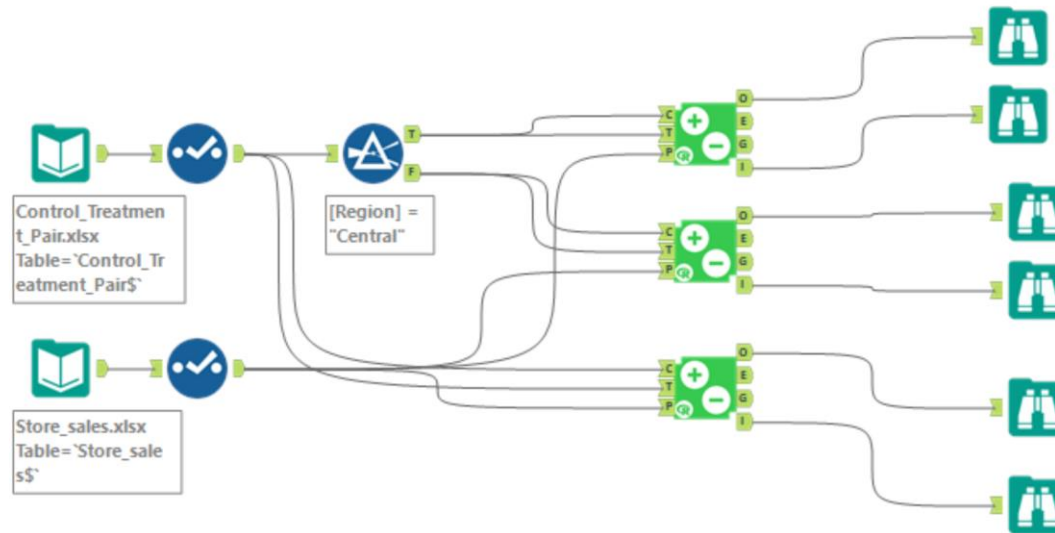
## Step 4: Analysis and Writeup

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?

The Overall Lift for the combined regions is at 40.70% >18% (Incremental Lift ) which justifies that menu can be rolled out across all markets.



**Workflow for A/B test analysis**

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

Central Region Analysis Results: Lift 43.5%

Report

### Analysis of the Test on the Measure Sum\_Gross Margin

Test Start Date: 2016-04-29

Test End Date: 2016-07-21

Additional Information:

#### Test Summary

The average percentage change in Sum\_Gross Margin was 39.7% for the treatment units in the test period relative to the comparison period. This same measure was -1.7% for the control units, with the difference between the treatment and control units being 41.5%, which is highly statistically significant. More detailed summary statistics for the treatment and control groups are contained in the first table (which immediately follows), while the details of the hypothesis test of a significant difference in the mean average percentage change in Sum\_Gross Margin is contained in a table at the end of this report.

A comparison of the treatment-control pairs indicates an average lift in Sum\_Gross Margin for the treatment units over the control units of 43.5%, which results in an expected impact of 836 on Sum\_Gross Margin, with 100.0% of the treatment-control pairs exhibiting a positive lift for the treatment units.

#### Lift Analysis for Sum\_Gross Margin

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

Western Region Analysis Results: Lift 37.9%

Report

### Analysis of the Test on the Measure Sum\_Gross Margin

Test Start Date: 2016-04-29

Test End Date: 2016-07-21

Additional Information:

#### Test Summary

The average percentage change in Sum\_Gross Margin was 39.2% for the treatment units in the test period relative to the comparison period. This same measure was 1.9% for the control units, with the difference between the treatment and control units being 37.2%, which is highly statistically significant. More detailed summary statistics for the treatment and control groups are contained in the first table (which immediately follows), while the details of the hypothesis test of a significant difference in the mean average percentage change in Sum\_Gross Margin is contained in a table at the end of this report.

A comparison of the treatment-control pairs indicates an average lift in Sum\_Gross Margin for the treatment units over the control units of 37.9%, which results in an expected impact of 527 on Sum\_Gross Margin, with 100.0% of the treatment-control pairs exhibiting a positive lift for the treatment units.

#### Lift Analysis for Sum\_Gross Margin

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



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

Combined Regions Analysis Results: Lift 40.70%

Report

### Analysis of the Test on the Measure Sum\_Gross Margin

Test Start Date: 2016-04-29

Test End Date: 2016-07-21

Additional Information:

#### Test Summary

The average percentage change in Sum\_Gross Margin was 39.5% for the treatment units in the test period relative to the comparison period. This same measure was 0.1% for the control units, with the difference between the treatment and control units being 39.4%, which is highly statistically significant. More detailed summary statistics for the treatment and control groups are contained in the first table (which immediately follows), while the details of the hypothesis test of a significant difference in the mean average percentage change in Sum\_Gross Margin is contained in a table at the end of this report.

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.

#### Lift Analysis for 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.