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 o-f your test?
   Gross margin was used as a performance metric to evaluate the results.

2. What is the test period?
   The data between April 29, 2016 and July 21, 2016 was used as the test period.

3. At what level (day, week, month, etc.) should the data be aggregated?
   The data was aggregated as weekly per requested in the requirements.

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.

- Dataset values for set properly and unnecessary columns were removed.
- Datasets were joined and new set was filtered based on the test period.
- Week information was identified and new week column was created accordingly.
- Based on the prepared data weekly, traffic and sale analysis was prepared and saved as new datasets.
- Treatment stores were identified in a new column and saved as a new dataset.

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 following two variables AvgMonthSales and Sq_Ft are considered as control variables.

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

From the Pearson Correlation Analysis below, it can be seen that *AvgMonthSales* variable has high correlation with the performance metric. However *Square Feet* (Sq_Ft) variable has a poor

correlation.

Pearson Correlation Analysis

Full Correlation Matrix

	AvgMonthSales	Sq_Ft	Sum_Sum_Gross.Margin
AvgMonthSales	1.000000	-0.046967	0.990982
Sq_Ft	-0.046967	1.000000	-0.024255
Sum_Sum_Gross.Margin	0.990982	-0.024255	1.000000

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

AvgMonthSales with Trend and Seasonality will be used to match treatment and control stores.

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

Treatment Store	Control Store 1	Control Store 2
1664	7162	7484
1675	2114	8162
1696	1863	7284
1700	1662	7384
1712	6992	7434
2288	2568	9081
2293	11768	9918
2301	12536	3185
2322	9238	9388
2341	11368	9488

Step 4: Analysis and Writeup

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

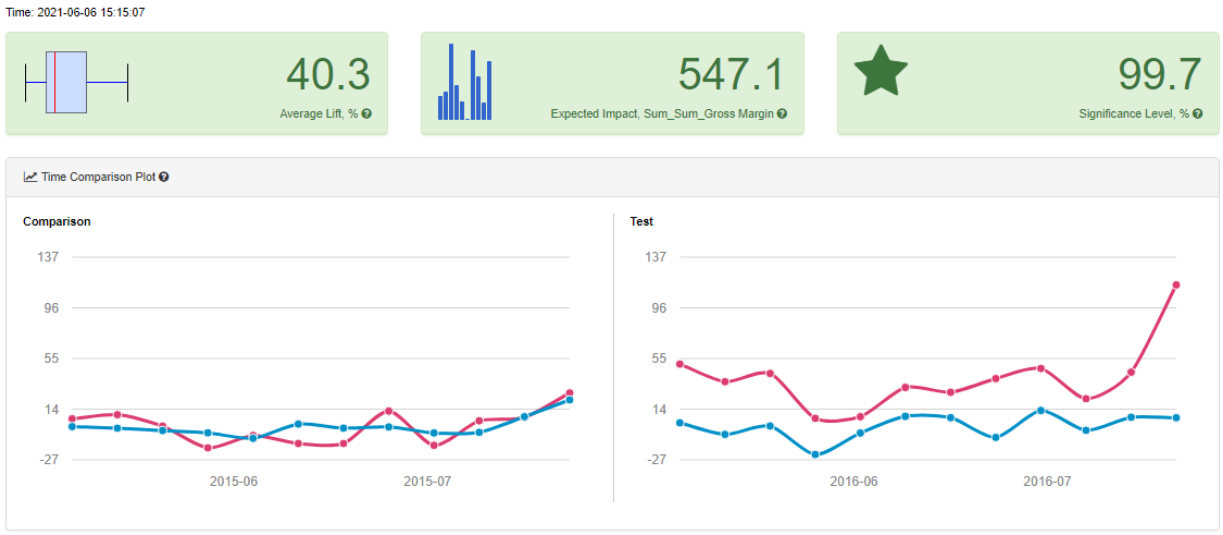
Answer these questions. Be sure to include visualizations from your analysis:

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

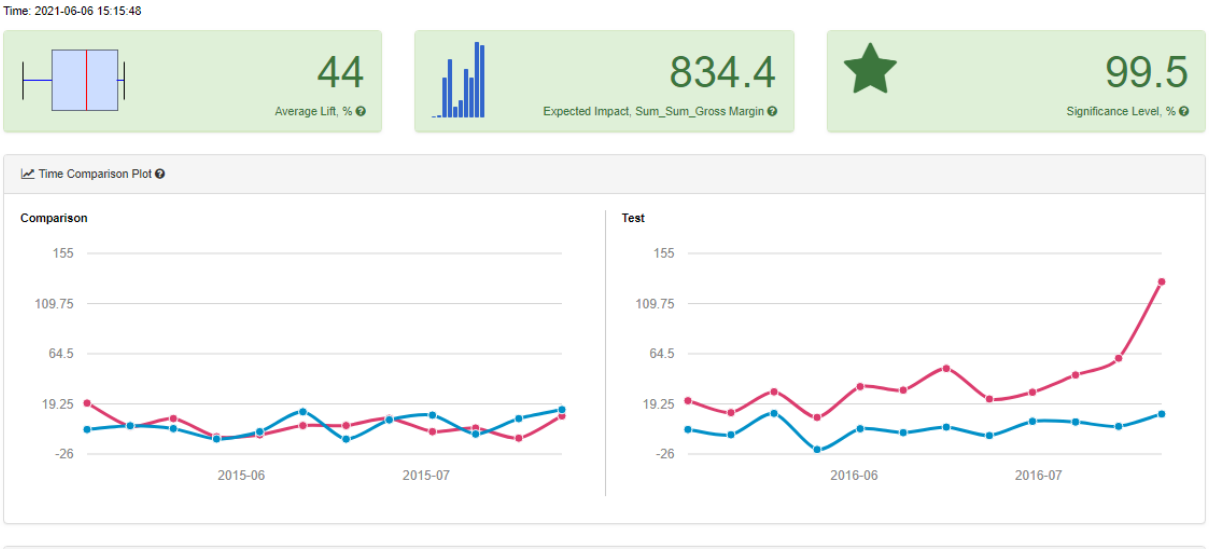
Based on the finding of this analysis, it is recommended that company roll out the updated menu to all stores. It is determined that both in West and Central region the profit exceeds 18% increase.

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

The figure below shows the report for West region. Average lift is 40.3%, the expected impact is $547.1 and significance level is 99.7%.



The figure belows illustrates the report for Central region. The average lift is 44%, the expected impact is $834.4, and the significance level is %99.5.

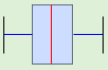


3. What is the lift from the new menu overall?

The overall lift details are shown below. The average lift is %42.2, the expected impact is $690.7 and significance is %100.

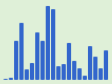
AB Test Analysis for Sum_Sum_Gross Margin

Time: 2021-06-06 15:16:29



42.2

Average Lift, %



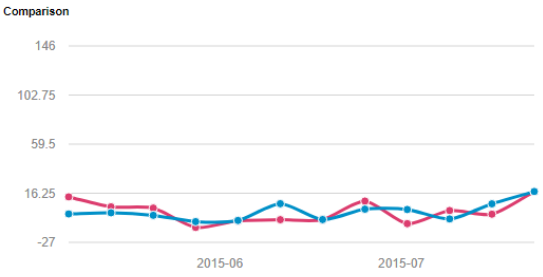
690.7

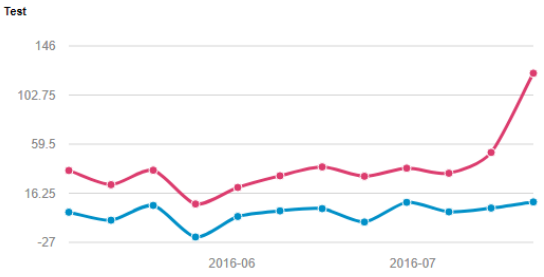
Expected Impact, Sum_Sum_Gross Margin

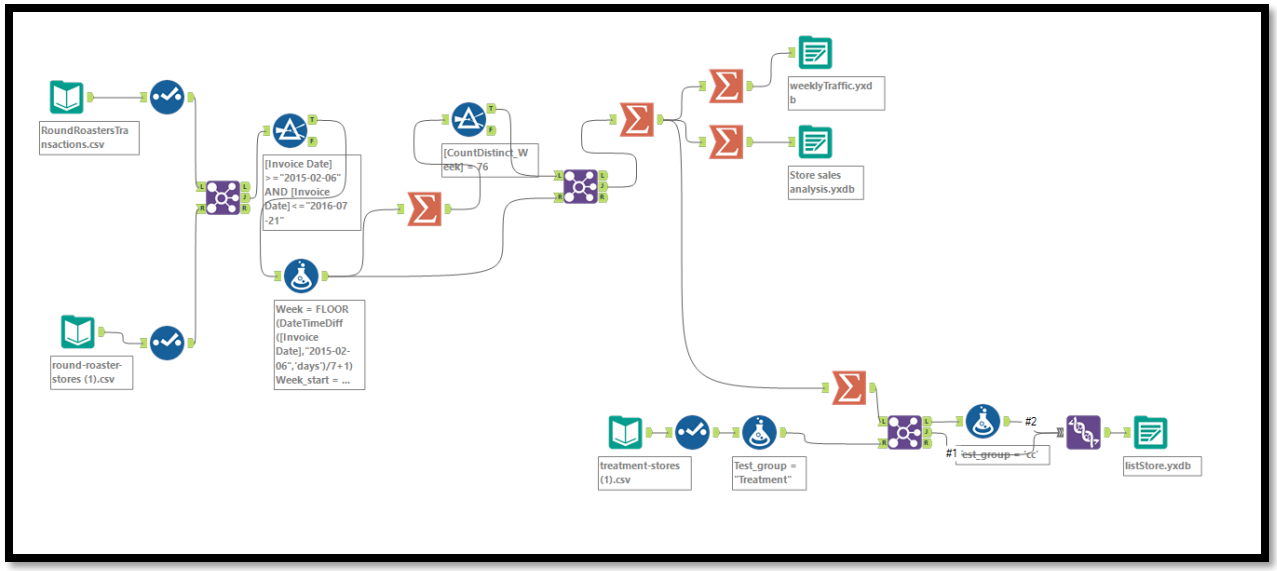
100

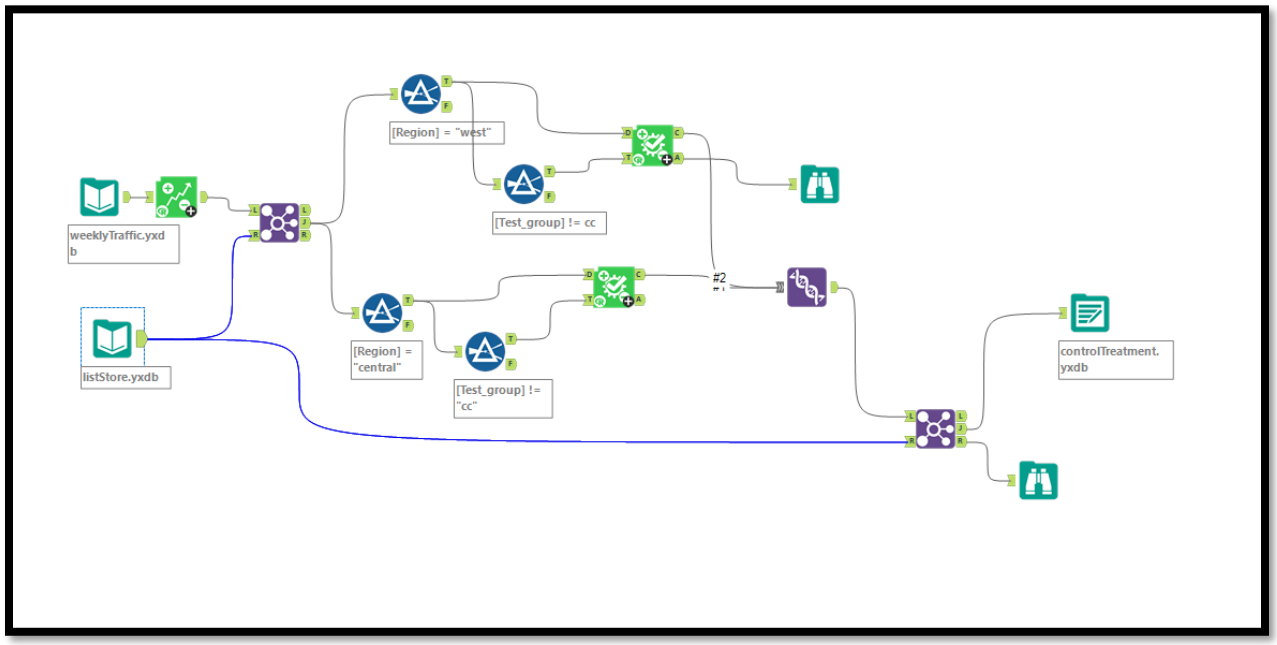
Significance Level, %

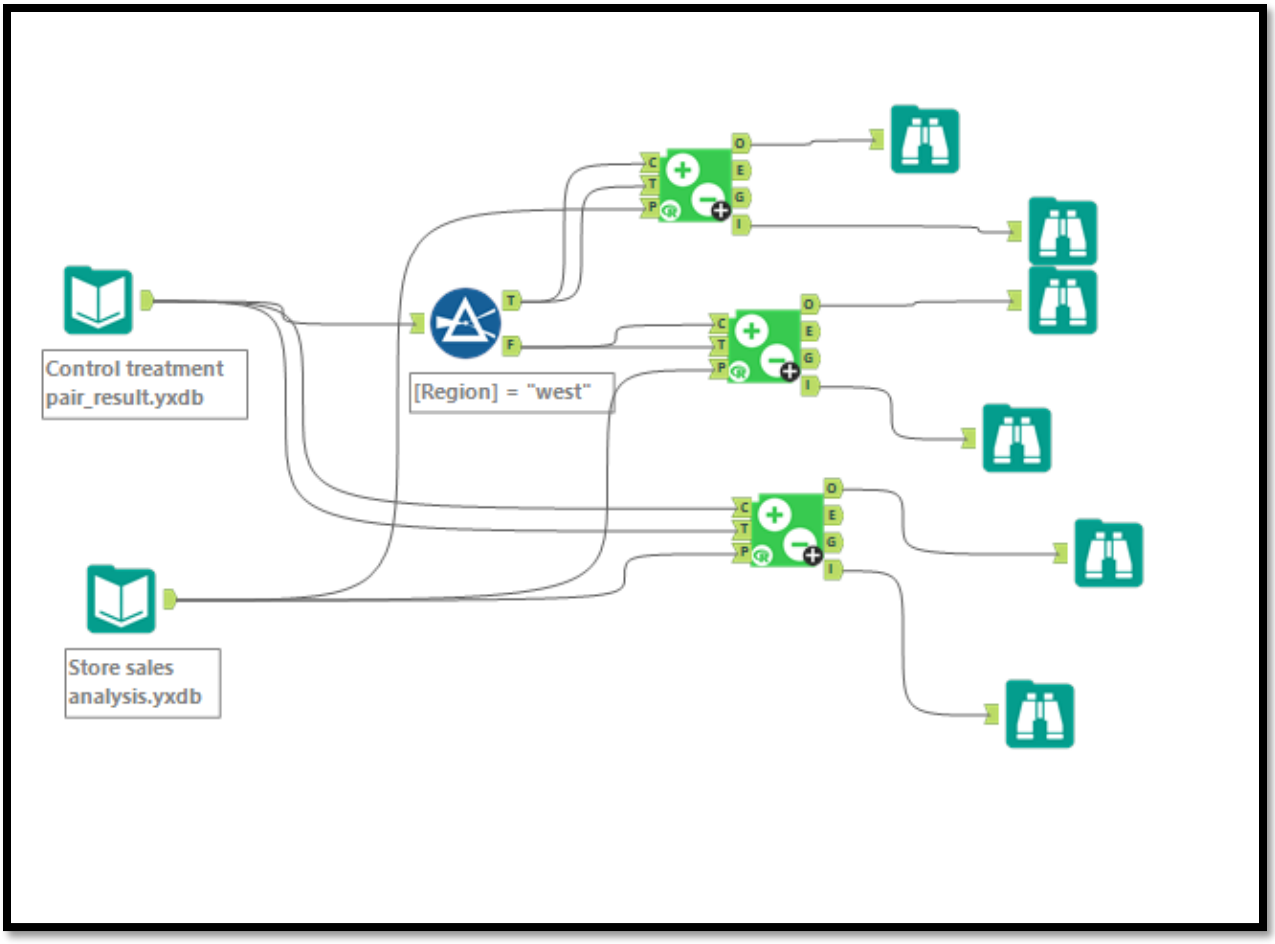
Time Comparison Plot






RoundRoastersTransactions.csv
[Invoice Date] >= "2015-02-06" AND [Invoice_Date] <= "2016-07-21"
[CountDistinct_Week] = 76
weeklyTraffic.yxdb
Store sales analysis.yxdb
round-roaster-stores (1).csv
Week = FLOOR (DateTimeDiff ([Invoice Date],"2015-02-06",'days')/7+1) Week_start = ...
treatment-stores (1).csv
Test_group = "Treatment"
#1 test_group = "cc"
#2
listStore.yxdb


[Region] = "west"
[Test_group] != cc
weeklyTraffic.yxdb
listStore.yxdb
[Region] = "central"
[Test_group] != "cc"
#2
controlTreatment.yxdb



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.