Project: Analyzing a Market Test

Step 1: Plan Our Analysis

1. What is the performance metric we'll use to evaluate the results of your test?
   Gross margin is the best performance metric we should use since we are interested in profits rather than costs or revenues.
2. What is the test period?
   12 weeks start from 2016-April-29 to 2016-July-21.
3. At what level (day, week, month, etc.) should the data be aggregated?
   Week is a good period cycle to monitor all customers' behavior.

Step 2: Clean Up Our Data

Firstly, I've checked data with interesting fields only, there is no missing, incomplete, duplicate. I have resolved some issues that would help in getting data prepared before analysis process:

1- In 'RoundRoastersTransactions.csv', some fields like Invoice_Date, QTY, Gross Margin, and Sales are converted to their proper data type.

2- The store sales transactions have been filtered to keep only transactions with dates between 2015-February-06 and 2016-July-21 (76 weeks).

3- The store sales transactions have been aggregated by each week for each store.

4- Listing all stores with classifying whether a store is a treatment or not by using 'treatment-stores.csv'.

5- Joining store class list with 'round-roaster-stores.csv' in order to add AvgMonthSales field.

Step 3: Match Treatment and Control Units

1. What control variables should be considered?
   Average Monthly Sales and area square feets for each store.
2. What is the correlation between our potential control variables and our performance metric?
   Using the following correlation matrix, we found that area square feets doesn't affect the gross margin profit but Average Monthly Sales strongly do.

Full Correlation Matrix

	Sum_Gross.Margin	Sq_Ft	AvgMonthSales
Sum_Gross.Margin	1.000000	-0.019345	0.790358
Sq_Ft	-0.019345	1.000000	-0.046967
AvgMonthSales	0.790358	-0.046967	1.000000

3. What control variables will we use to match treatment and control stores?
   In addition to trend and seasonality we should use Average Monthly Sales too.
4. The table below represents our treatment and control stores pairs:

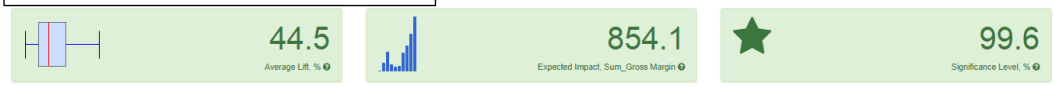
Treatment Store	Control Store 1	Control Store 2
1664	7162	8112
1675	1580	1807
1696	1964	1863
1700	2014	1630
1712	8162	7434
2288	9081	2568
2293	12219	9524
2301	3102	9238
2322	2409	3235
2341	12536	2383

Step 4: Analysis and Writeup

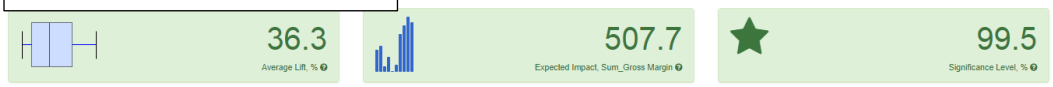
Answer these questions. Be sure to include visualizations from your analysis:

1. What is our recommendation - Should the company roll out the updated menu to all stores?
   The company should roll out the updated menu to all stores as a comparison of the treatment-control pairs indicates an average lift in Sum Gross Margin for the treatment units over the control units of 40.4%, which results in an expected impact of 680.9 with significance level 100%.
2. What is the lift from the new menu for West and Central regions (include statistical significance)?
   The lift for Central region is 44.5% while the lift for West region is 36.3% and both have a statistical significance of 99.6% and 99.5% respectively.

Central Region Gross Margin



West Region Gross Margin



3. What is the lift from the new menu overall?
   The overall lift is 40.4% with a significance level 100%.

Overall Gross Margin