Project: Analyzing a Market Test

Step 1: Plan Your Analysis

Predicting how well the new menu offerings performs is based on the new television advertising test in two cities (Denver and Chicago), serving has the test markets each consisting of five stores. These test markets were selected based on their similarity to all stores across the entire chain of stores.
To evaluate the tests results of this analysis, the **Weekly Gross Margin** will be the performance metric. The weekly gross margin represents the profit which is needed to justify the need for an increase in marketing budget; at least 18% increase in profit growth.

For the analysis the test will run for a period of 12 weeks (2016-April-29 to 2016-July-21) and in addition with the 52 weeks and 12 weeks of data. The **total test period is 76 weeks** (2015-January-21 to 2016-July-21).

Since the test period is in weeks, the data will be **aggregated weekly.**

Step 2: Clean Up Your Data

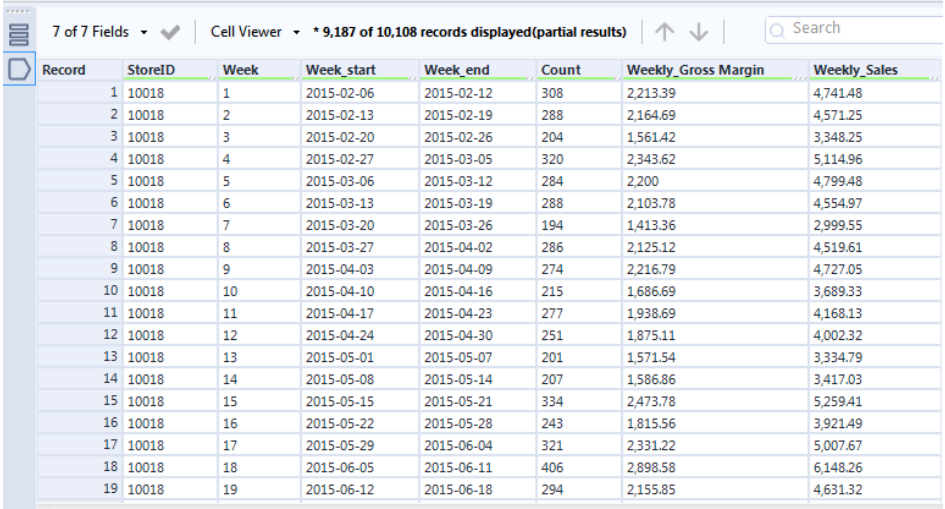
The dataset available for the analysis are:

Round-roaster-stores.csv – Consists of store information for each Round Roaster store in the USA.
Treatment-stores.csv - This is store information for each store that offered the new menu items.
Round-roaster-transactions.csv - Contains transaction level information for all of Round Roaster's stores.

These raw data files were cleaned and prepared to generate the necessary files to be used for the AB Test Analysis:

Weekly_store _traffic: This data consists of store id, the weeks of the test period, week_start, week_end, count, weekly_gross_margin and weekly_sales. These results were achieved by modifying some of the data types, filtering the test period of the data, calculating, aggregating and grouping the data variables in the round-roaster-transactions.csv. Table 1 shows the results below:

Table 1. Weekly Store Traffic Data.



7 of 7 Fields | Cell Viewer | * 9,187 of 10,108 records displayed(partial results) | Search

Record	StoreID	Week	Week_start	Week_end	Count	Weekly_Gross Margin	Weekly_Sales
1	10018	1	2015-02-06	2015-02-12	308	2,213.39	4,741.48
2	10018	2	2015-02-13	2015-02-19	288	2,164.69	4,571.25
3	10018	3	2015-02-20	2015-02-26	204	1,561.42	3,348.25
4	10018	4	2015-02-27	2015-03-05	320	2,343.62	5,114.96
5	10018	5	2015-03-06	2015-03-12	284	2,200	4,799.48
6	10018	6	2015-03-13	2015-03-19	288	2,103.78	4,554.97
7	10018	7	2015-03-20	2015-03-26	194	1,413.36	2,999.55
8	10018	8	2015-03-27	2015-04-02	286	2,125.12	4,519.61
9	10018	9	2015-04-03	2015-04-09	274	2,216.79	4,727.05
10	10018	10	2015-04-10	2015-04-16	215	1,686.69	3,689.33
11	10018	11	2015-04-17	2015-04-23	277	1,938.69	4,168.13
12	10018	12	2015-04-24	2015-04-30	251	1,875.11	4,002.32
13	10018	13	2015-05-01	2015-05-07	201	1,571.54	3,334.79
14	10018	14	2015-05-08	2015-05-14	207	1,586.86	3,417.03
15	10018	15	2015-05-15	2015-05-21	334	2,473.78	5,259.41
16	10018	16	2015-05-22	2015-05-28	243	1,815.56	3,921.49
17	10018	17	2015-05-29	2015-06-04	321	2,331.22	5,007.67
18	10018	18	2015-06-05	2015-06-11	406	2,898.58	6,148.26
19	10018	19	2015-06-12	2015-06-18	294	2,155.85	4,631.32

Store_sales_analysis_data: Compiles the store_id, week, week_start, week_end, count, weekly_gross_margin, sales, and region; which were used along with the control and treatment group data to analyze the lift and statistical significance of the overall new. Table 2 shows the results below:

Table 2. Store Sales Analysis Data

8 of 8 Fields * 8,173 of 10,108 records displayed(partial results) Search Data Metadata Action

Record	StoreID	Region	Week	Week_start	Week_end	Count	Weekly_Gross Margin	Weekly_Sales
1	10018	West	1	2015-02-06	2015-02-12	308	2,213.39	4,741.48
2	10018	West	2	2015-02-13	2015-02-19	288	2,164.69	4,571.25
3	10018	West	3	2015-02-20	2015-02-26	204	1,561.42	3,348.25
4	10018	West	4	2015-02-27	2015-03-05	320	2,343.62	5,114.96
5	10018	West	5	2015-03-06	2015-03-12	284	2,200	4,799.48
6	10018	West	6	2015-03-13	2015-03-19	288	2,103.78	4,554.97
7	10018	West	7	2015-03-20	2015-03-26	194	1,413.36	2,999.55
8	10018	West	8	2015-03-27	2015-04-02	286	2,125.12	4,519.61
9	10018	West	9	2015-04-03	2015-04-09	274	2,216.79	4,727.05
10	10018	West	10	2015-04-10	2015-04-16	215	1,686.69	3,689.33
11	10018	West	11	2015-04-17	2015-04-23	277	1,938.69	4,168.13
12	10018	West	12	2015-04-24	2015-04-30	251	1,875.11	4,002.32
13	10018	West	13	2015-05-01	2015-05-07	201	1,571.54	3,334.79
14	10018	West	14	2015-05-08	2015-05-14	207	1,586.86	3,417.03
15	10018	West	15	2015-05-15	2015-05-21	334	2,473.78	5,259.41
16	10018	West	16	2015-05-22	2015-05-28	243	1,815.56	3,921.49
17	10018	West	17	2015-05-29	2015-06-04	321	2,331.22	5,007.67
18	10018	West	18	2015-06-05	2015-06-11	406	2,898.58	6,148.26

Store list: Adds the treatment test group to the controls and treatment data set. The round-roaster-stores.csv and treatment-stores.csv was prepared to create the file. Table 3 shows the results below:

Table 3. Store List Data.

5 of 5 Fields Cell Viewer 133 records displayed Search

Record	StoreID	Avg_Sq_Ft	Avg_AvgMonthSales	Test Group	Region
1	10018	1,183	18,000	CC	West
2	10068	1,198	16,000	CC	West
3	10118	1,204	13,000	CC	West
4	10168	1,195	19,000	CC	West
5	10218	1,193	15,000	CC	West
6	10268	1,191	25,000	CC	West
7	10318	1,147	16,000	CC	West
8	10368	1,159	19,000	CC	West
9	10418	1,146	19,000	CC	West
10	10468	1,151	21,000	CC	West
11	10518	1,147	11,000	CC	West
12	10568	1,192	21,000	CC	West
13	10618	1,187	15,000	CC	West
14	11268	1,328	12,000	CC	West
15	11318	1,319	24,000	CC	West
16	11368	1,319	22,000	CC	West
17	11418	1,323	15,000	CC	West
18	11468	1,325	19,000	CC	West

Step 3: Match Treatment and Control Units

The control variables in the round roaster store file to be considered in matching the treatment and control units are apart from trend and seasonability are **Sq_ft** and **Average_month_sales**

The correlation between the performance metric and the two control variable was tested using association analysis and Avg_AvgMonthSales **highly correlated (0.99)** with the performance metric 'Weekly Gross Margin' while Avg_Sq_ft **poorly correlated (-0.02).** Table 4 below shows the correlation results:

Table 4. Correlation Analysis of Control Variable and Performance Metric.

Pearson Correlation Analysis

Full Correlation Matrix

	Sum_Gross.Margin	Avg_Sq_Ft	Avg_AvgMonthSales
Sum_Gross.Margin	1.000000	-0.024224	0.990978
Avg_Sq_Ft	-0.024224	1.000000	-0.046967
Avg_AvgMonthSales	0.990978	-0.046967	1.000000

Matrix of Corresponding p-values

	Sum_Gross.Margin	Avg_Sq_Ft	Avg_AvgMonthSales
Sum_Gross.Margin		0.78196	0.00000
Avg_Sq_Ft	0.78196		0.59138
Avg_AvgMonthSales	0.00000	0.59138	

From the correlation results above, the control variables used to match the treatment and control stores are **Trend, Seasonality, and Avg_AvgMonthly Sales.** Avg_sq_ft was not included because it had a low correlation with performance metric 'Weekly Gross Margin'**.**

Table 5. The treatment and control stores pairs:

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

Step 4: Analysis and Write-up

Recommendation:
From this analysis I recommend the company to roll out their updated menu to all stores. As the results of the overall lift in weekly gross margin was 40.7% which is over the profit needed to justify the need for an increase in marketing budget; (greater than 18%).

Lift for Central Region:
The average percentage change in the Weekly_Gross Margin was 39.7% for the treatment units in the test period relative to the comparison period. The Weekly_Gross Margin lift for the central region is 43.5%, which is highly statistically significant (significance level 99.5%) The statistical significance can be found in table 6 and figure 1-3 below:

Table 6. Lift_Analysis_for_Central_Region.

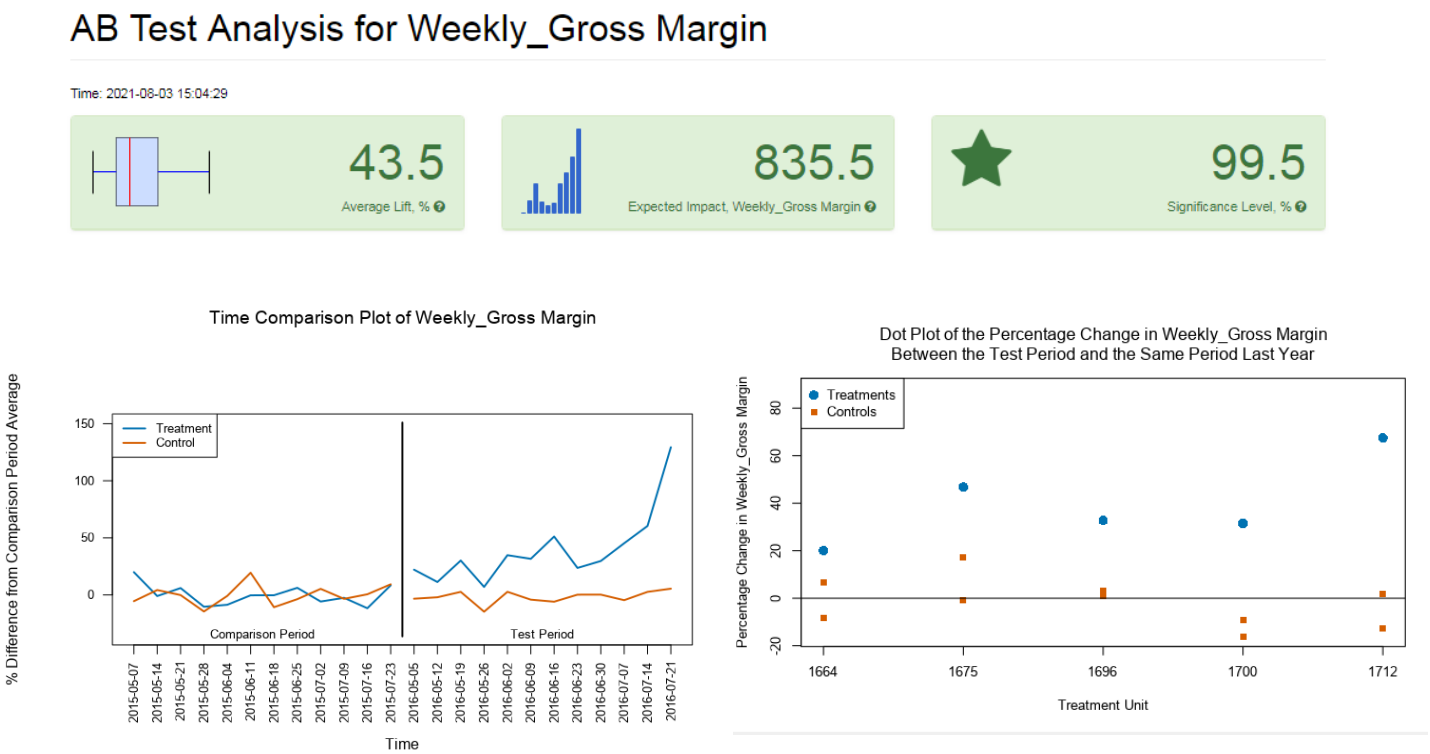
Lift Analysis for Weekly_Gross Margin

Lift	Expected Impact	Significance Level
43.5%	835	99.5%

Summary Statistics for Weekly_Gross Margin by Test Group

Statistic	Treatment	Control
Average	39.71	-1.72
Minimum	20.07	-16.18
Maximum	67.48	17.30
Standard Deviation	17.15	10.03

Figure 1-3. Visualizations_for_the_Lift_in_Central_Region



Lift for West Region:

The Weekly_Gross Margin lift for the west region was 37.9%, which is highly statistically significant (significance level 99.5%) The average percentage change in the Weekly_Gross Margin was 39.1% for the treatment units in the test period relative to the comparison period. The statistical significance can be found in table 7 and figure 4-6 below:

Table 7. Lift_Analysis_for_West_Region.

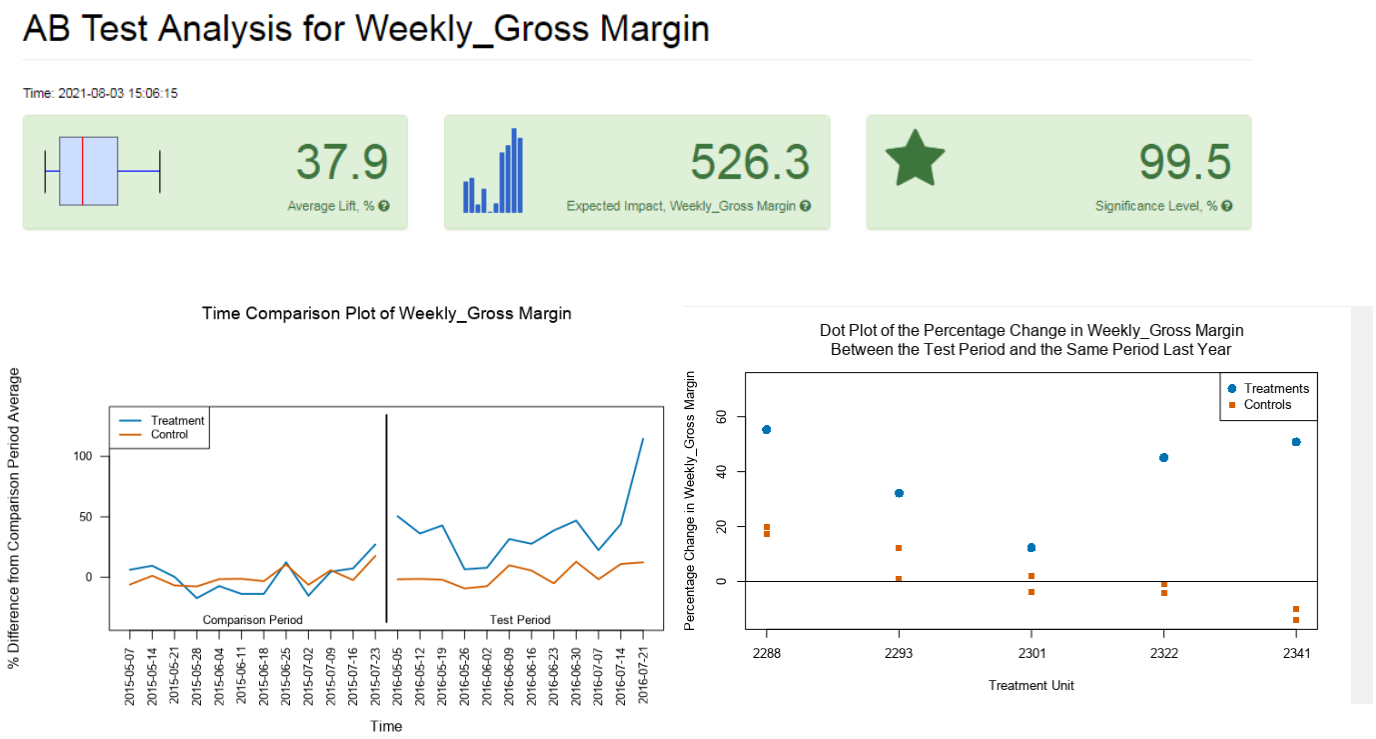
Lift Analysis for Weekly_Gross Margin

Lift	Expected Impact	Significance Level
37.9%	526	99.5%

Summary Statistics for Weekly_Gross Margin by Test Group

Statistic	Treatment	Control
Average	39.14	1.92
Minimum	12.31	-13.96
Maximum	55.27	19.70
Standard Deviation	16.33	11.24

Figure 4-6. Visualizations_for_the_Lift_in_West_Region.



Lift for New Menu Overall:

The lift for the new menu overall (Central and West Region) was 40.7%. Average percentage change in Weekly_Gross Margin was 39.4% for the treatment units in the test period relative to the comparison period. The lift was highly statistically significant at a 100%. The table 8 and fig 7-9 below shows this:

Table 8. Overall_Lift_for_New_Menu.

Lift Analysis for Weekly_Gross Margin

Lift	Expected Impact	Significance Level
40.7%	681	100.0%

Summary Statistics for Weekly_Gross Margin by Test Group

Statistic	Treatment	Control
Average	39.42	0.10
Minimum	12.31	-16.18
Maximum	67.48	19.70
Standard Deviation	16.30	10.54

Figure 7-9. Visualizations_for_the_Lift_Overall_Menu.