Project: Analyzing a Market Test

Step 1: Plan Your Analysis

To perform the correct analysis, you will need to prepare a data set. (250 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?
   I'll use "gross margin" as performance metric.
2. What is the test period?
   Test period is 12 weeks, from April 29th, 2016 to July 21st, 2016.
3. At what level (day, week, month, etc.) should the data be aggregated?
   Data should be aggregated in weeks.

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.

   These variables should be considered
   - Average sale,
   - Size of store,
   - Location of store (City, State, Region)
   - Time zone

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

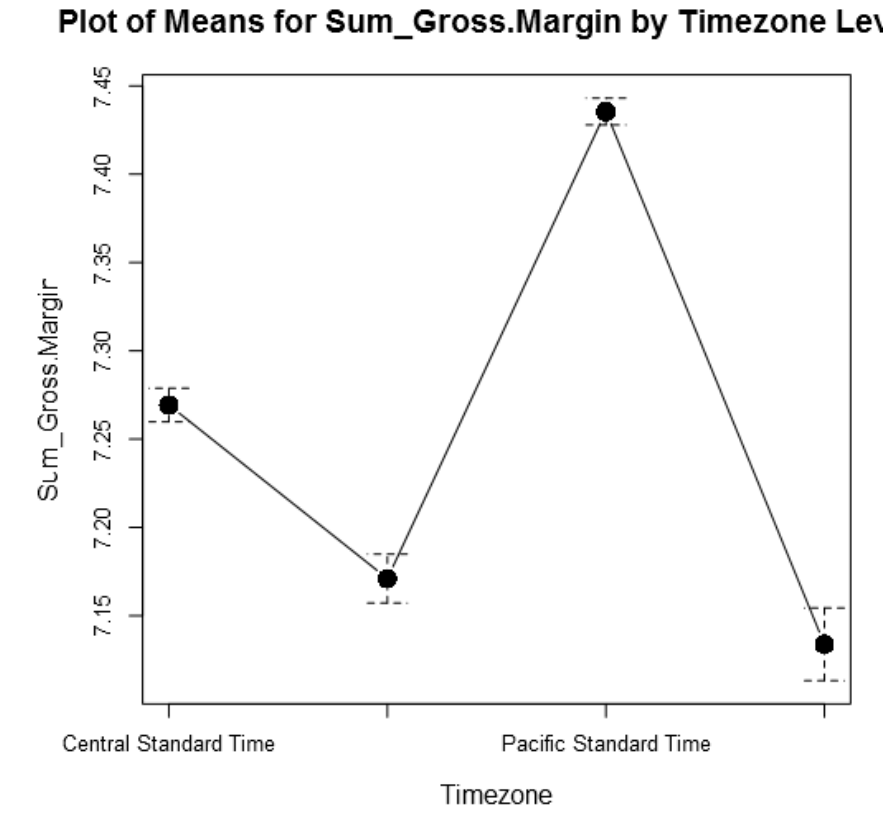
   Sq_Ft showed no correlation with my performance metric.
   Time zone didn't have any correlation with my performance metric.
   I counted total numbers of store by aggregating with region, state, city, and postal code variables. The total numbers of stores exceed 15 only when it is aggregated by region.
   There is strong correlation between average monthly sales and my performance metric.

   So I decided to exclude city, state, time zone and size of store (Sq_Ft) from my list of control variables and, also decided to include average monthly sales and region along with trend and seasonality in my list of control variables.

Pearson Correlation Analysis

Full Correlation Matrix

	Sq_Ft	AvgMonthSales	Sum_Gross.Margin
Sq_Ft	1.000000	-0.046967	-0.020353
AvgMonthSales	-0.046967	1.000000	0.988219
Sum_Gross.Margin	-0.020353	0.988219	1.000000



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

I will use
- Trend,
- Seasonality,
- AvgMonthSales and
- Region as control variables to match treatment and control stores.

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

Treatment Store	Control Store 1	Control Store 2
1664	1964	7162
1675	1807	7584
1696	1863	7334
1700	7037	2014
1712	8162	7434
2288	2568	9081
2293	12686	9639
2301	12019	9238
2322	3235	9388
2341	2572	3102

Step 4: Analysis and Writeup

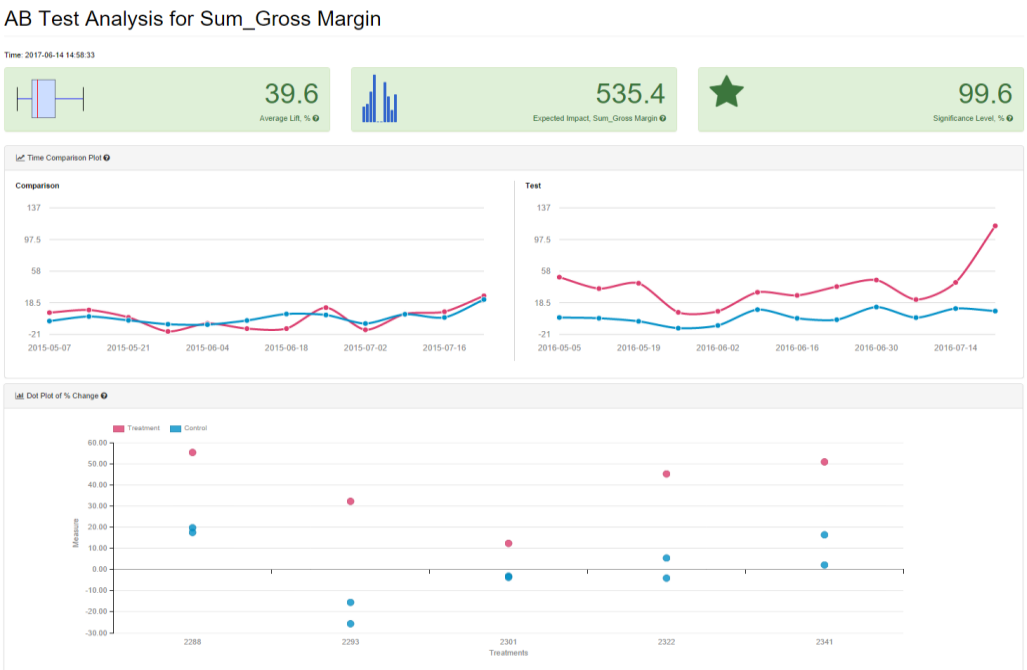
Conduct your A/B analysis and create a short report outlining your results and recommendations. (250 words limit)

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 management expect at least 18% average lift to applied menu updates. My AB analysis shows overall average lift of 42.3 % with statistical significance value of 100%. So my recommendation is – the company should update menu in all stores.

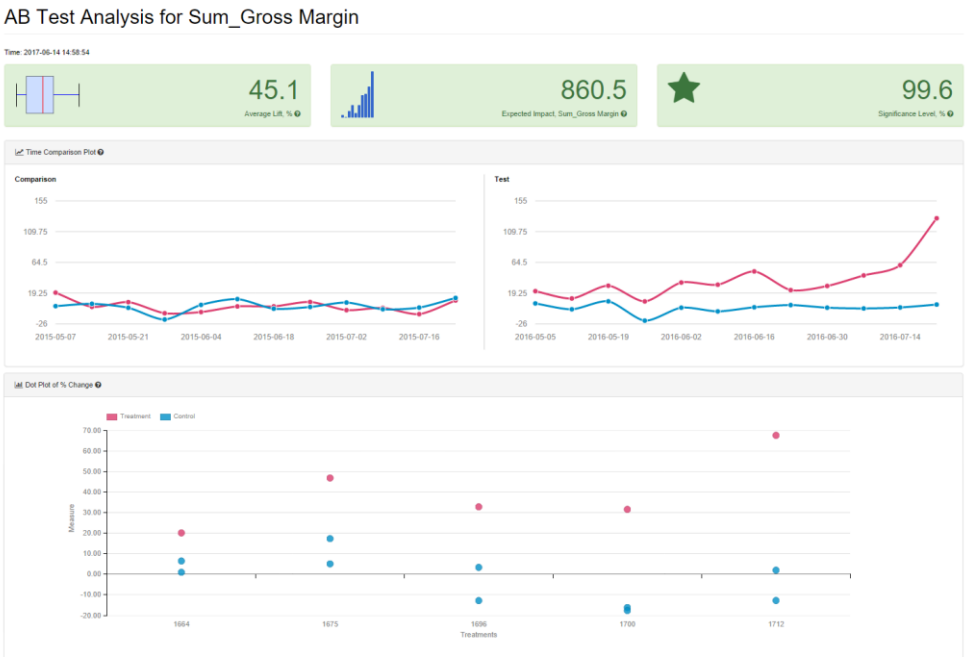
2. What is the lift from the new menu for West and Central regions (include statistical significance)?

   For West Regions
   - average lift is 39.6 % with significance level of 99.6 %.



For Central Regions

- average lift is 45.1 % with significance level of 99.6 %



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

Overall average lift is 42.3%