

# Project: Analyzing a Market Test

## 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 performance metric to evaluate the A/B Test will be weekly gross margin per store, that is the benefit of each store per week.

2. What is the test period?

The test period is 12 weeks, from 2016-April-29 to 2016-July-21.

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

The data should be aggregated at the week level, considering the stores containing data 76 weeks back from the end day of the test, to have enough data to create trend and seasonality variables to use as control variables. The data should be grouped by store and region.

## 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.*

Figure 1 shows the workflow to aggregate the data to week level and filter the necessary data ranges.

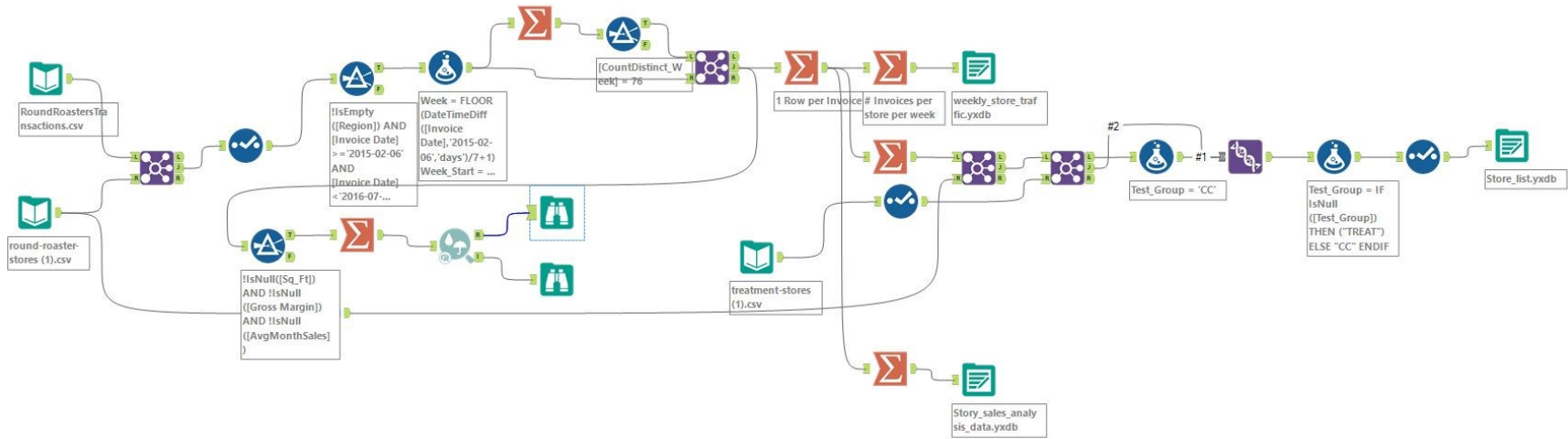


Figure 1. Workflow to aggregate data by week and select data from 76 weeks before the test is performed and create two databases with a list of stores and information about the sales.

## 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 control variable that should be considered are AvgMonthSales, the average sales per month and Square feet (Sq\_Ft). Trend and seasonality will also be considered.

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

Figure 2 shows the full correlation matrix and the related p-values. There is a high correlation between AvgMonthSales and gross margin (0.7904) with a p-value of 0. On the contrary, Sq\_Ft is non-significant (p-value > 0.05) and with an almost zero correlation with weekly gross margin (rho = -0.019).

### Pearson Correlation Analysis

*Focused Analysis on Field weekly\_Gross.Margin*

	Association Measure	p-value
AvgMonthSales	0.790358	0.000000 ***
Sq_Ft	-0.019345	0.051796 .

*Full Correlation Matrix*

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

*Matrix of Corresponding p-values*

	weekly_Gross.Margin	Sq_Ft	AvgMonthSales
weekly_Gross.Margin		5.1796e-02	0.0000e+00
Sq_Ft	5.1796e-02		2.3119e-06
AvgMonthSales	0.0000e+00	2.3119e-06	

Figure 2. Pearson full correlation matrix with the corresponding p-values, for the variables weekly gross margin, square feet and average month sales.

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

In order to match control units to treatment units the following numeric variables are used: AvgMonthSales (significant correlation with weekly gross margin), trend and seasonality. Trend and seasonality are obtained with the AB Trend tool in the comparison time (one year before the test started).

4. Please fill out the table below with your 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

*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?  
Yes, the company should offer the updated menu to all stores because it will provide more benefit.
2. What is the lift from the new menu for West and Central regions (include statistical significance)?

For the Central regions, the lift is 43.5% with a statistical significance of 99.6%. The expected impact in weekly gross margin per store is \$835.9, as shown in Figure 3.

## AB Test Analysis for weekly\_gross\_margin\_store

Time: 2018-12-29 17:29:30



Figure 3. A/B test results in Central regions in a test performed for 12 months including sandwiches and wine in treatment coffee shops.

For the Western regions, the lift is lower than in the Central regions, 37.9% with a statistical significance of 99.5%. The expected impact in weekly gross margin per store is \$526.5 (Figure 4).

## AB Test Analysis for weekly\_gross\_margin\_store

Time: 2018-12-29 17:29:56

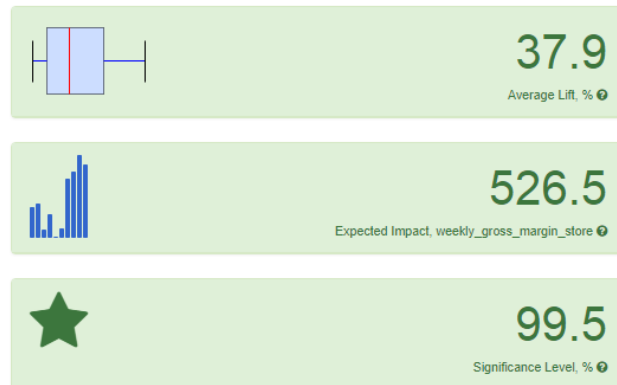


Figure 4. A/B test results in Western regions in a test performed for 12 months including sandwiches and wine in treatment coffee shops.

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

Overall, the lift is 40.7% with a statistical significance of 100%. The expected impact in weekly gross margin per store is \$681.2.

## AB Test Analysis for weekly\_gross\_margin\_store

Time: 2018-12-29 16:10:41

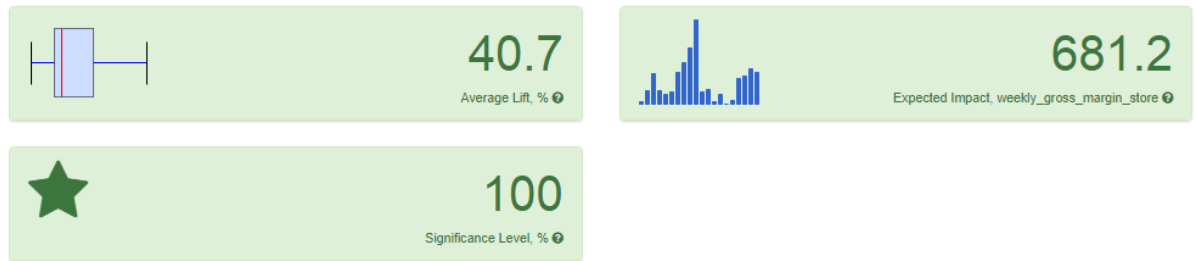


Figure 5. A/B test results in Central and Western regions in a test offering sandwiches and wine in treatment coffee shops.

Figure 6 contains the workflow to pair a treatment store with two control stores and Figure 7 shows the workflow for the A/B analysis.

