

SEND FEEDBACK

### Project Overview

You're a business analyst for Round Roasters, a coffee restaurant in the United States of America. The executive team conducted a market test with a new menu and needs to figure whether the new menu can drive enough sales to offset the cost of marketing the new menu. Your job is to analyze the A/B test and write up a recommendation to whether the Round Roasters chain should launch this new menu.

### How Do I Complete this Project?

This project uses skills learned throughout the "A/B Testing " course. To complete this project:

- Go through the course
- Apply the skills learned in the course to solve the business problem given in the project details section.
- Use our guidelines and rubric to help build your project.
- When you're ready, submit it to us for review using the submission template found in the supporting materials section.

### Skills Required

In order to complete this project, you must be able to:

- Cleanup, format, and blend a wide range of data sources
- Plan and analyze A/B tests

NEXT

SEND FEEDBACK

### The Business Problem

Round Roasters is an upscale coffee chain with locations in the western United States of America. The past few years have resulted in stagnant growth at the coffee chain, and a new management team was put in place to reignite growth at their stores.

The first major growth initiative is to introduce gourmet sandwiches to the menu, along with limited wine offerings. The new management team believes that a television advertising campaign is crucial to drive people into the stores with these new offerings.

However, the television campaign will require a significant boost in the company's marketing budget, with an unknown return on investment ([ROI](#)). Additionally, there is concern that current customers will not buy into the new menu offerings.

To minimize risk, the management team decides to test the changes in two cities with new television advertising. Denver and Chicago cities were chosen to participate in this test because the stores in these two cities (or markets) perform similarly to all stores across the entire chain of stores; performance in these two markets would be a good proxy to predict how well the updated menu performs.

The test ran for a period of 12 weeks (2016-April-29 to 2016-July-21) where five stores in each of the test markets offered the updated menu along with television advertising.

The comparative period is the test period, but for last year (2015-April-29 to 2015-July-21).

You've been asked to analyze the results of the experiment to determine whether the menu changes should be applied to all stores. The predicted impact to profitability should be enough to justify the increased marketing budget: at least 18% increase in profit growth compared to the comparative period while compared to the control stores; otherwise known as *incremental lift*. In the data, profit is represented in the *gross\_margin* variable.

You have been able to gather three data files to use for your analysis:

- Transaction data for all stores from 2015-January-21 to 2016-August-18
- A listing of all Round Roasters stores
- A listing of the 10 stores (5 in each market) that were used as test markets.

## Steps to Success

### Step 1: Plan Your Analysis

To perform the correct analysis, you will need to prepare a data set. Prior to rolling up your sleeves and preparing the data, it's a good idea to have a plan of what you need to do in order to prepare the correct data set. A good plan will help you with your analysis. Here are a few questions to get you started:

- What is the performance metric you'll use to evaluate the results of your test?
- What is the test period?
- At what level (day, week, month, etc.) should the data be aggregated?

### 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 your other control variable(s) to match two control units to each treatment unit.

Treatment stores should be matched to control stores in the same region. *Note:*

*Calculate the number of transactions per store per week and use 12 periods to calculate trend and seasonality.*

Apart from trend and seasonality...

- What control variables should be considered? Note: Only consider variables in the RoundRoastersStore file.
- What is the correlation between your each potential control variable and your performance metric? (Example of correlation matrix below)
- What control variables will you use to match treatment and control stores?

### Step 4: Analysis and Writeup

Conduct your A/B analysis and create a short report outlining your results and recommendations.

## Pearson Correlation Analysis

### Full Correlation Matrix

	random	fake	alternative
random	1.0000000	-0.0023871	-0.0262367
fake	-0.0023871	1.0000000	-0.0331166
alternative	-0.0262367	-0.0331166	1.0000000

In an AB Analysis we use the correlation matrix to find the most correlated variable to the performance metric to include in the AB controls tool to help find the best matches.

NEXT

### ; Review

Use the [project rubric](#) to review your project. If you are happy with your submission, then you're ready to submit your project. If you see room for improvement, keep working to improve your project.

### A/B Test Guide

The A/B test guide can help you map out the process you'll need to go through to complete an A/B test. This will be useful for the project, as well as any A/B tests you do in a professional setting.

### Submission Template

Use the submission template at the bottom of this section to submit your project. After filling it out, save it as a PDF and submit the PDF in the next section. You may also include your Alteryx workflow if you'd like. If your submission does not meet specifications, having the workflow may help the review identify mistakes.

### Data

*round-roaster-stores.csv* - This file contains store information for each Round Roaster store in the USA.

*treatment-stores.csv* - This file contains store information for each store that offered the new menu items.

*round-roaster-transactions.zip* - This file contains transaction level information for all of Round Roaster's stores

### Project Checklist

This is here to make it easier for you to assess if your project is close to passing so you can try debugging your workflow and project before submission. If you do not feel confident in the topics addressed please review the course material or reach out on Knowledge or Study Groups!

#### Task List

- ☐

The lift in my project for each Central and West is between 30-50%

- ☐

The significance in my project for each Central and West is above 85%

- ☐

Showed the correlation between each potential control variable and the performance metric

- ☐

I checked the rubric <https://review.udacity.com/#!/rubrics/287/view>

- ☐

I have gone to the Knowledge or Study Groups before submitting

If your Lift or Significance levels are not within the ranges stated above, please review your workflow. There may be a small detail missed somewhere causing the error.

- Forums: <https://knowledge.udacity.com>
- Study Groups

**To Download Files below please right click on the link and select "Save Link As"**

### Supporting Materials

- [Round Roaster Stores](#)
- [Treatment Stores](#)
- [Round Roaster Transactions Database](#)
- [AB Test Guide](#)
- [P5 Submission Template](#)
- [All Project Files](#)

NEXT

;

SEND FEEDBACK

### Aggregate and Export

We recommend you save your aggregated transaction database as a separate file to further reduce your development time. You shouldn't need to re-aggregate the transaction database every time you want to test out a new workflow.

### Creating a week variable

When calculating the week column, assign week 1 to the first week of the test period, so the week\_begin variable for week 1 would be your test start date. This will cause weeks prior to the test period to be negative, which is perfectly fine. See [here](#) . If the first link doesn't work, please click [here](#) for an example of how to calculate the week variable using a formula tool in Alteryx.

### Number of Weeks

The trend tool is used to create trend and seasonality variables to use as control variables. To do this, you need at least 52 weeks of data, plus the number of weeks you select in the tool to calculate trend, before the beginning of the test start date. In lesson

4, you used 6 weeks to calculate the trend, so you needed 58 weeks prior to the test start date. For the project, you are asked to use 12 weeks to calculate trend, so you'll need 64 weeks of data prior to the test start date. Since the test lasts for 12 weeks, this means you'll need a total 76 weeks of data.

Your filter at the beginning of your workflow should go back 76 weeks from the end of the test period 2016-July-21. In Alteryx the expression should look like `[Invoice Date]>="2015-02-06" AND [Invoice Date]<"2016-07-22"`

All stores should have 76 weeks of Data.

## **AB Trend Tool**

The Test start date is 2016-April-29

The performance metric for this tool is the invoice count per week which represent weekly foot traffic. You had to create this variable with a summarize tool.

## **AB Controls Tool**

You should use 3 numeric measures to match treatment and control stores.

1. Trend
2. Seasonality
3. AvgMonthSales (This should be determined by looking at the correlation between the appropriate numeric variables in the round roasters stores file AvgMonthSales and Sq\_ft with the performance metric gross margin.) - This variable is in the round-roaster-stores.csv file

## **AB Analysis Tool**

Make sure to use `weekly gross margin` per store in all lift calculations and not total sales. Your data has gross margin in it but you will have to use a summarize tool to get weekly gross margin per store.

## **Steps**

1. Filter the data to the proper date range
2. Aggregate the data to get the weekly gross margin and weekly traffic count (count of unique invoices)
3. Calculate Trend and Seasonality with the AB Trend Tool
4. Label the data as treatment and control stores
5. Calculate correlation between other numeric measures and the performance metric (gross margin)
6. Match Treatment to control stores per region using the AB Controls Tool
7. Calculate lift from control to treatment store with AB Analysis tool

If you run into errors in Alteryx or unexpected results from a tool we have a guide to help you figure out what is going on.

**Alteryx Debugging Guide:** See the Resources tab in the left most panel of your classroom for a downloadable PDF of the Alteryx Debugging Guide

[Google Doc](#) Please note: If you are accessing Udacity via a corporate internet network or have certain settings on your computer, you may not be able to see the Google Doc, please see the resources tab instead.

Forum: <https://knowledge.udacity.com/>

## Supporting Materials

- [Debugging Alteryx](#)

NEXT

.

,