Project: Analyzing a Market Test

Project Details 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.**

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:

- What is the performance metric you'll use to evaluate the results of your test?

The performance metric I'll use to evaluate the results of my experiment is **gross margin**. We want to raise earnings by at least 18 percent, and the gross margin is the only indicator we have to track it.

- What is the test period?

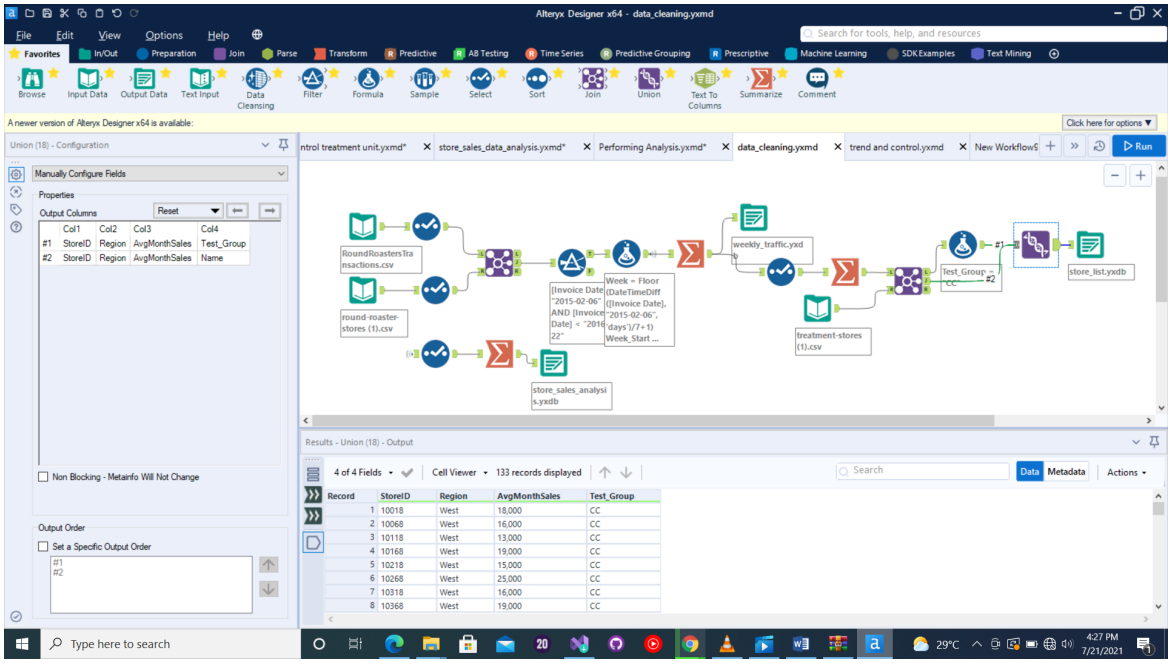
The test period is 12 weeks long (from **April 29th to July 21st, 2016**).

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

Data should be compiled on a **weekly** basis.

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. Note: Calculate the number of transactions per store per week to calculate trend and seasonality.

Apart from trends and seasonality…

- What control variables should be considered? Note: Only consider variables in the RoundRoastersStore file.

  AvgMonthSales and Sq_Ft

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

Record 1 | Layout

Pearson Correlation Analysis

Full Correlation Matrix

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

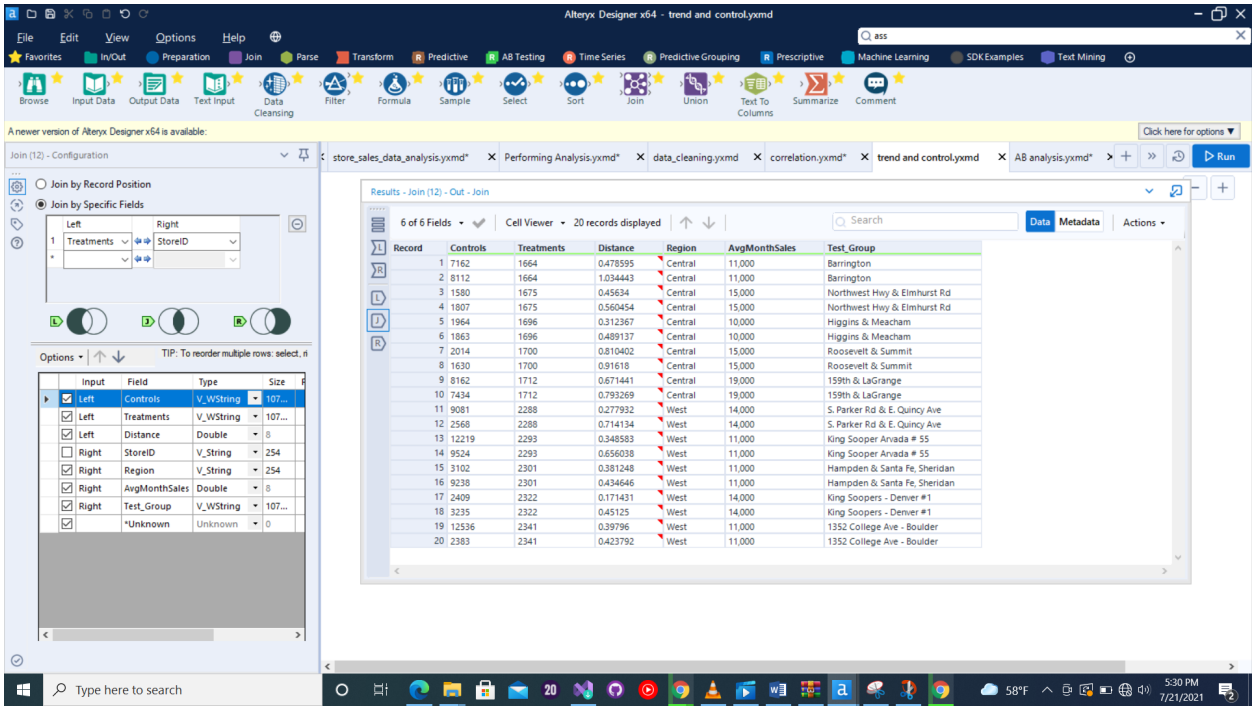
Using the Association Analysis tool, we discover that **Avg Month Sales** has a high correlation of **0.80** with the performance parameter, which is the **Sum of Gross Margin.** We can also note that **Square Feet**has a weak correlation of **-0.01**

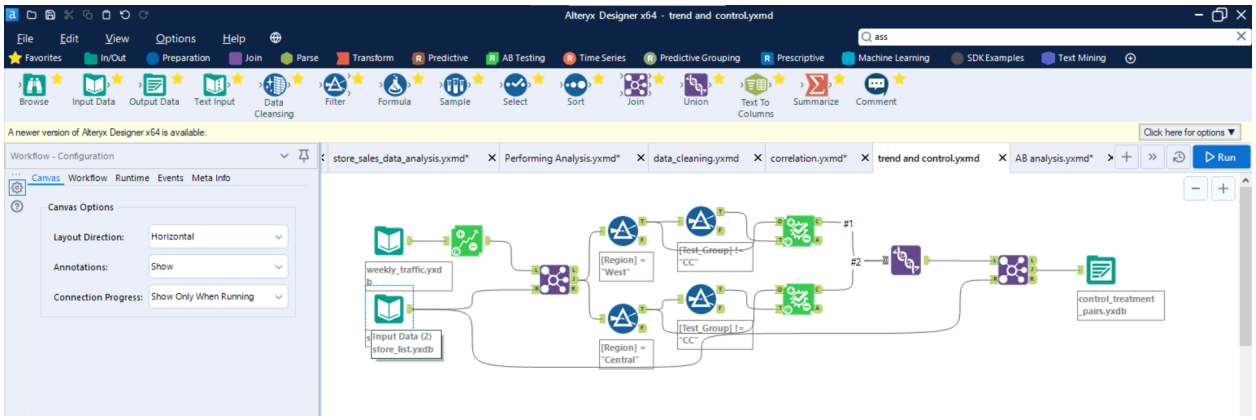
- What control variables will you use to match treatment and control stores?

  Seasonality, Trend, Avg month sales

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

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


Alteryx Designer x64 - trend and control.yxmd
File Edit View Options Help
Browse Input Data Output Data Text Input Data Cleansing Filter Formula Sample Select Sort Join Union Text To Columns Summarize Comment
A newer version of Alteryx Designer x64 is available:
Join (12) - Configuration
Join by Record Position
Join by Specific Fields
Left Right
Treatments StoreID
Options
TIP: To reorder multiple rows: select, ri
Input Field Type Size
Left Controls V_WString 107...
Left Treatments V_WString 107...
Left Distance Double 8
Right StoreID V_String 254
Right Region V_String 254
Right AvgMonthSales Double 8
Right Test_Group V_WString 107...
*Unknown Unknown 0
store_sales_data_analysis.yxmd*
Performing Analysis.yxmd*
data_cleaning.yxmd
correlation.yxmd*
trend and control.yxmd
AB analysis.yxmd*
Run
Results - Join (12) - Out - Join
6 of 6 Fields
Cell Viewer
20 records displayed
Search
Data Metadata
Actions
Record Controls Treatments Distance Region AvgMonthSales Test_Group
1 7162 1664 0.478595 Central 11,000 Barrington
2 8112 1664 1.034443 Central 11,000 Barrington
3 1580 1675 0.45634 Central 15,000 Northwest Hwy & Elmhurst Rd
4 1807 1675 0.560454 Central 15,000 Northwest Hwy & Elmhurst Rd
5 1964 1696 0.312367 Central 10,000 Higgins & Meacham
6 1863 1696 0.489137 Central 10,000 Higgins & Meacham
7 2014 1700 0.810402 Central 15,000 Roosevelt & Summit
8 1630 1700 0.91618 Central 15,000 Roosevelt & Summit
9 8162 1712 0.671441 Central 19,000 159th & LaGrange
10 7434 1712 0.793269 Central 19,000 159th & LaGrange
11 9081 2288 0.277932 West 14,000 S. Parker Rd & E. Quincy Ave
12 2568 2288 0.714134 West 14,000 S. Parker Rd & E. Quincy Ave
13 12219 2293 0.348583 West 11,000 King Sooper Arvada # 55
14 9524 2293 0.656038 West 11,000 King Sooper Arvada # 55
15 3102 2301 0.381248 West 11,000 Hampden & Santa Fe, Sheridan
16 9238 2301 0.434646 West 11,000 Hampden & Santa Fe, Sheridan
17 2409 2322 0.171431 West 14,000 King Soopers - Denver #1
18 3235 2322 0.45125 West 14,000 King Soopers - Denver #1
19 12536 2341 0.39796 West 11,000 1352 College Ave - Boulder
20 2383 2341 0.423792 West 11,000 1352 College Ave - Boulder
Type here to search
58°F
5:30 PM
7/21/2021


Alteryx Designer x64 - trend and control.yxmd
File Edit View Options Help
Browse Input Data Output Data Text Input Data Cleansing Filter Formula Sample Select Sort Join Union Text To Columns Summarize Comment
A newer version of Alteryx Designer x64 is available:
Workflow - Configuration
Canvas Workflow Runtime Events Meta Info
Canvas Options
Layout Direction: Horizontal
Annotations: Show
Connection Progress: Show Only When Running
store_sales_data_analysis.yxmd*
Performing Analysis.yxmd*
data_cleaning.yxmd
correlation.yxmd*
trend and control.yxmd
AB analysis.yxmd*
Run
weekly_traffic.yxdb
Input Data (2)
store_list.yxdb
[Region] = "West"
[Region] = "Central"
[Test_Group] != "CC"
[Test_Group] != "CC"
#1
#2
control_treatment_pairs.yxdb

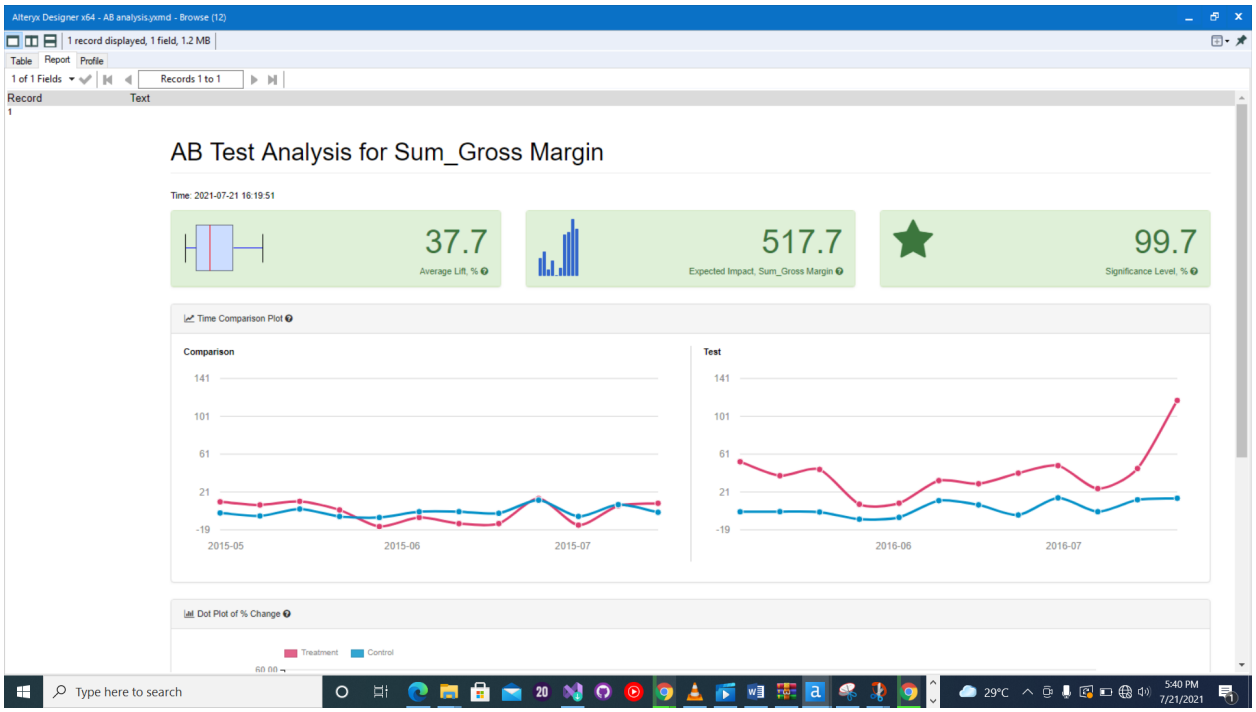
Step 4: Analysis and Writeup

- What is your recommendation - Should the company roll out the updated menu to all stores?

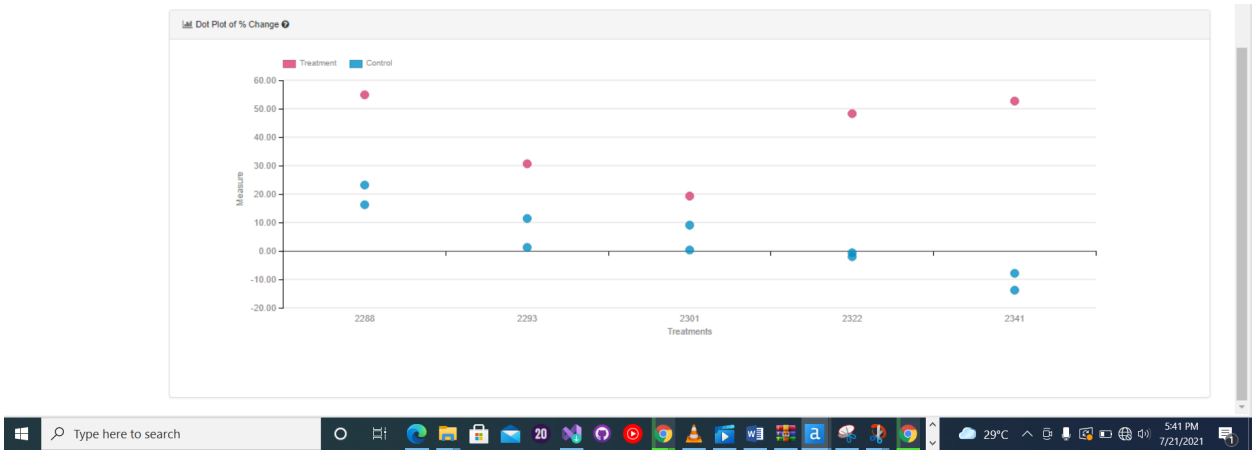
Round Roasters should make the improved menu available at all of its locations. The predicted increase in gross margin exceeds the 18 percent limit, with lifts of **46 percent and 37.7 percent** in the **Central and West regions**, respectively.

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

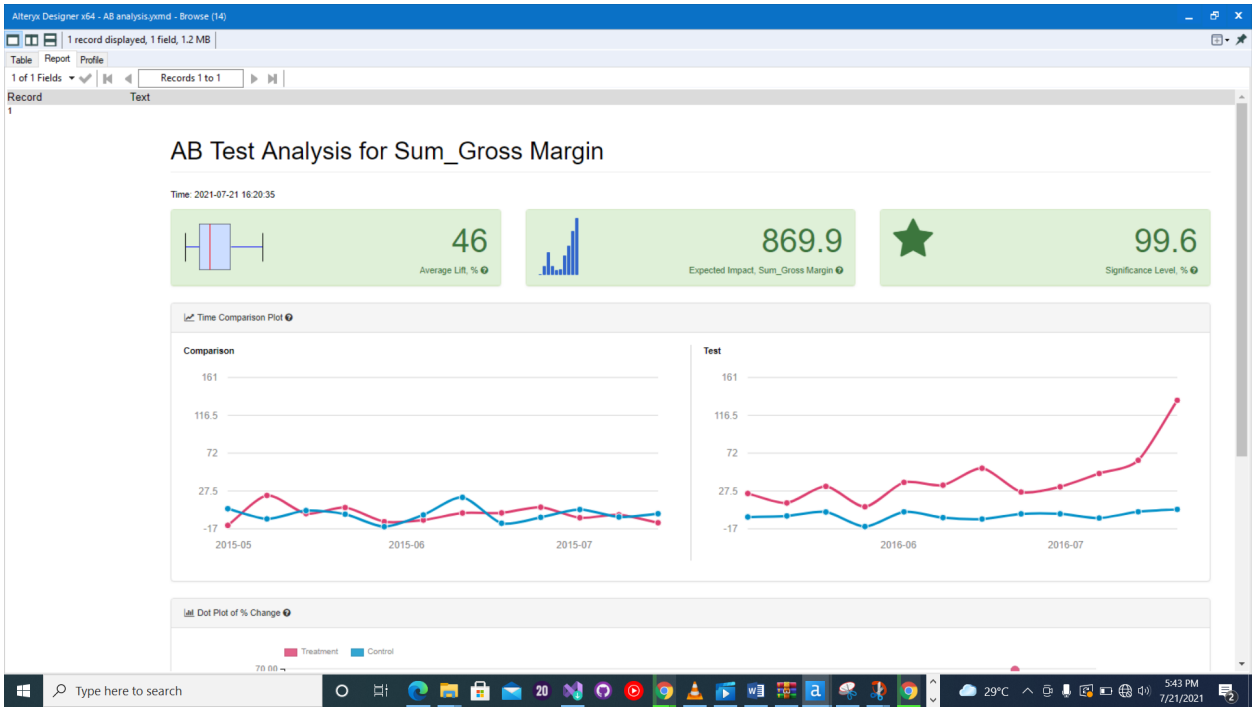
West Region



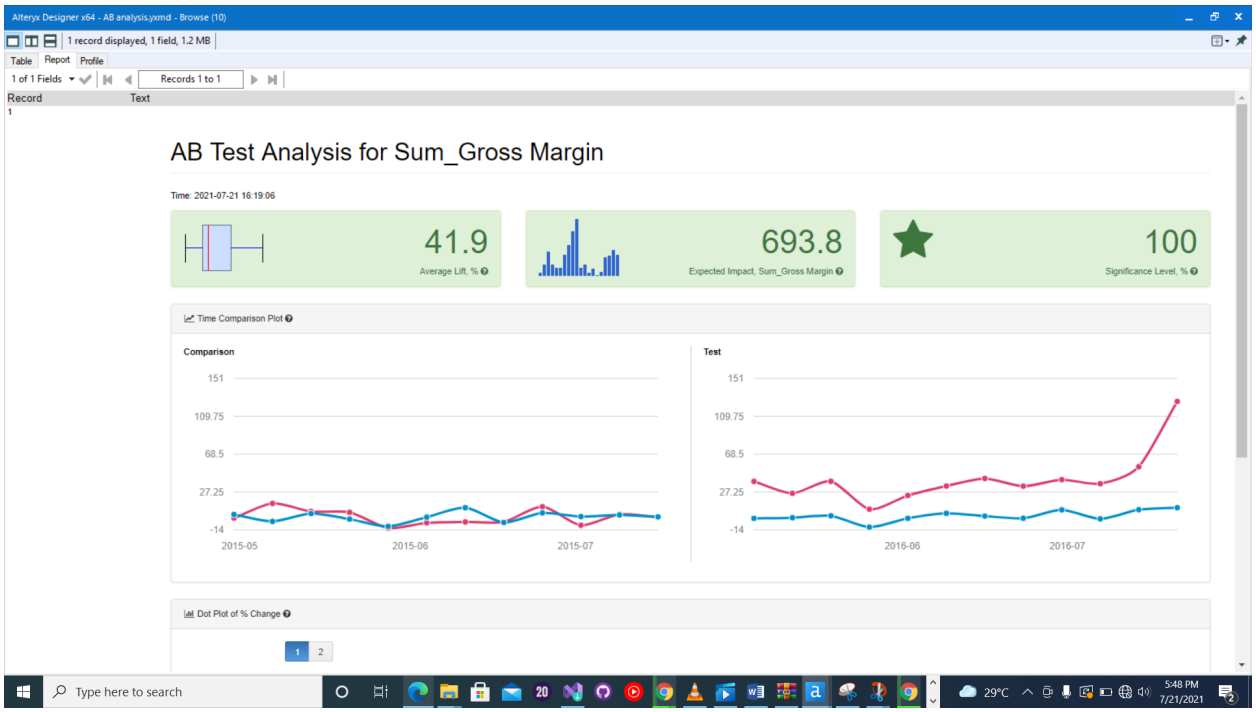
	Lift	Statistical Significance
West	37.7%	99.7%
Central	46%	99.6%

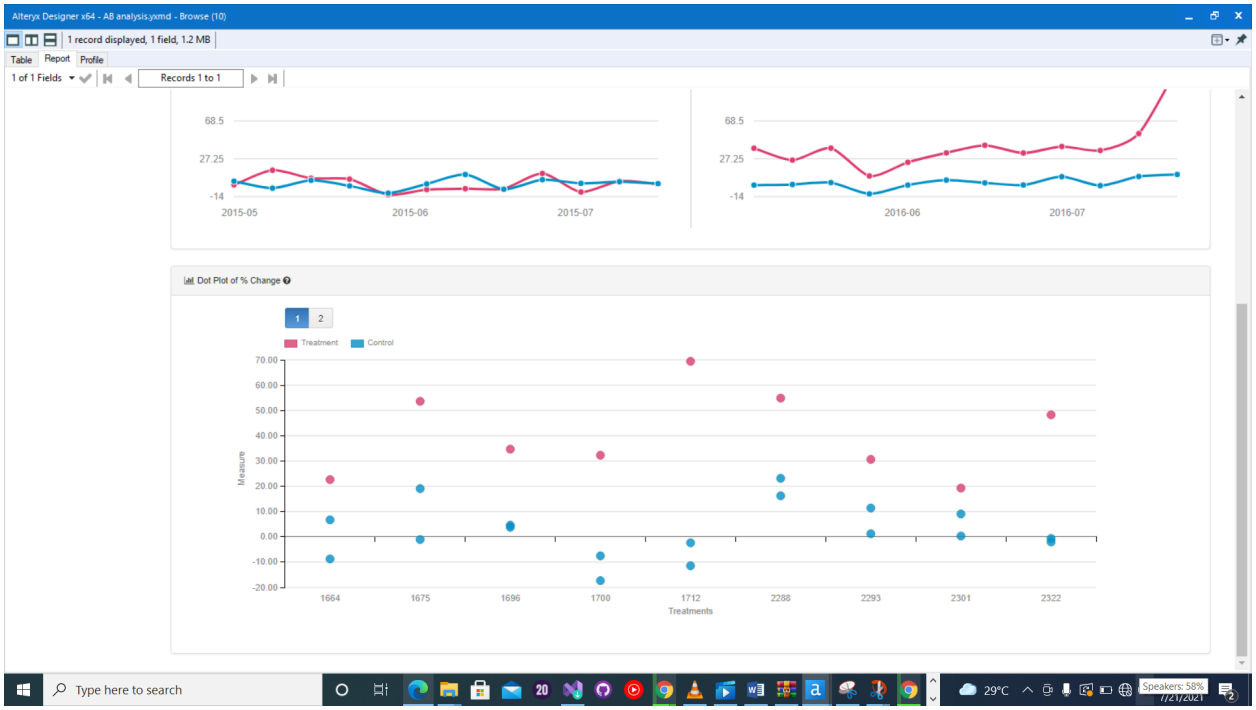


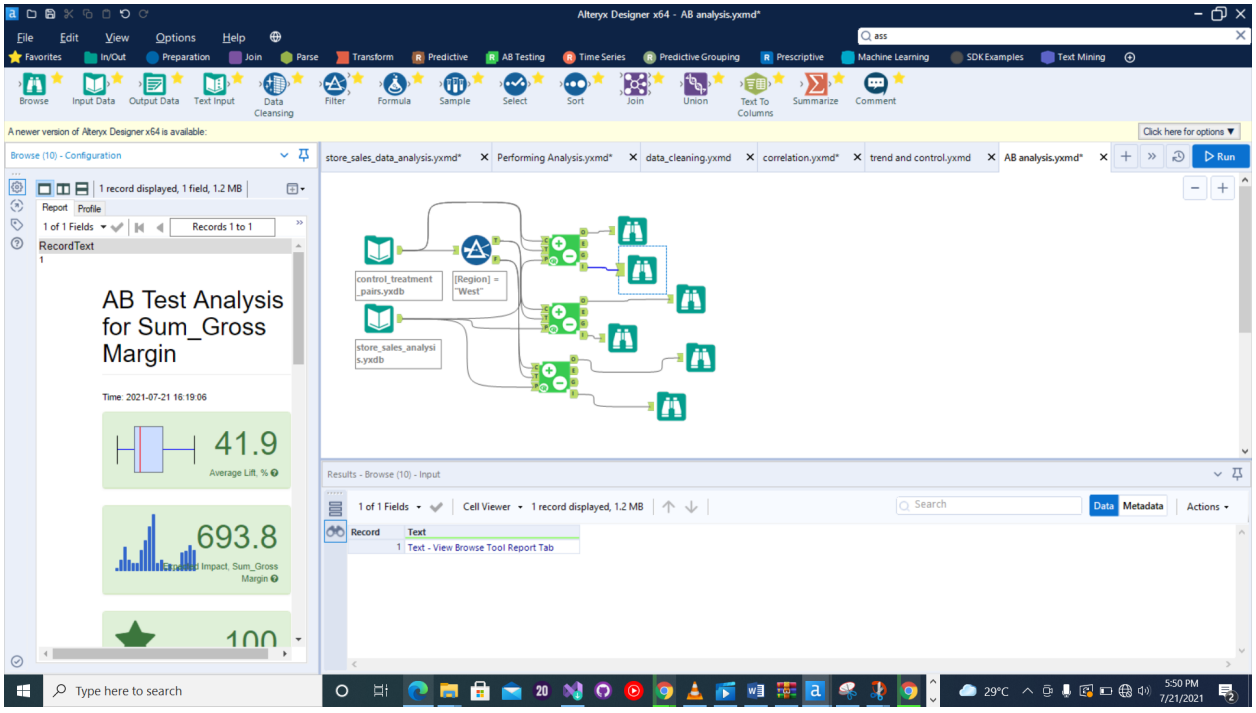
Central Region



- What is the lift from the new menu overall? **41.9% and Statistical Signicance of 100%**




Alteryx Designer x64 - AB analysis.yxmd - Browse (10)
1 record displayed, 1 field, 1.2 MB
Table
Report
Profile
1 of 1 Fields
Records 1 to 1
68.5
27.25
-14
2015-05
2015-06
2015-07
2016-06
2016-07
Dot Plot of % Change
Treatment
Control
Measure
70.00
60.00
50.00
40.00
30.00
20.00
10.00
0.00
-10.00
-20.00
1664
1675
1696
1700
1712
2288
2293
2301
2322
Treatments
Type here to search
29°C
Speakers: 58%


Alteryx Designer x64 - AB analysis.yxmd*
File
Edit
View
Options
Help
Favorites
In/Out
Preparation
Join
Parse
Transform
Predictive
AB Testing
Time Series
Predictive Grouping
Prescriptive
Machine Learning
SDK Examples
Text Mining
Browse
Input Data
Output Data
Text Input
Data Cleansing
Filter
Formula
Sample
Select
Sort
Join
Union
Text To Columns
Summarize
Comment
A newer version of Alteryx Designer x64 is available:
Click here for options
Browse (10) - Configuration
store_sales_data_analysis.yxmd*
Performing Analysis.yxmd*
data_cleaning.yxmd
correlation.yxmd*
trend and control.yxmd
AB analysis.yxmd*
Run
1 record displayed, 1 field, 1.2 MB
Report
Profile
1 of 1 Fields
Records 1 to 1
RecordText
1
AB Test Analysis for Sum_Gross Margin
Time: 2021-07-21 16:19:06
41.9
Average Lift, %
693.8
Impact, Sum_Gross Margin
control_treatment_pairs.yxdb
[Region] = "West"
store_sales_analysis.yxdb
Results - Browse (10) - Input
1 of 1 Fields
Cell Viewer
1 record displayed, 1.2 MB
Search
Data
Metadata
Actions
Record
Text
1
Text - View Browse Tool Report Tab
Type here to search
29°C
5:50 PM
7/21/2021