Predictive Analytics Nanodegree
Project: A/B Test a New Menu Launch

Step 1: Plan your Analysis

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

   The company wants to know whether there will be a significantly high positive impact on the profitability by new menu additions to justify increasing the marketing budget on the menu launch. Since the focus is on profitability, we choose the total gross margin as the performance metric to evaluate the results

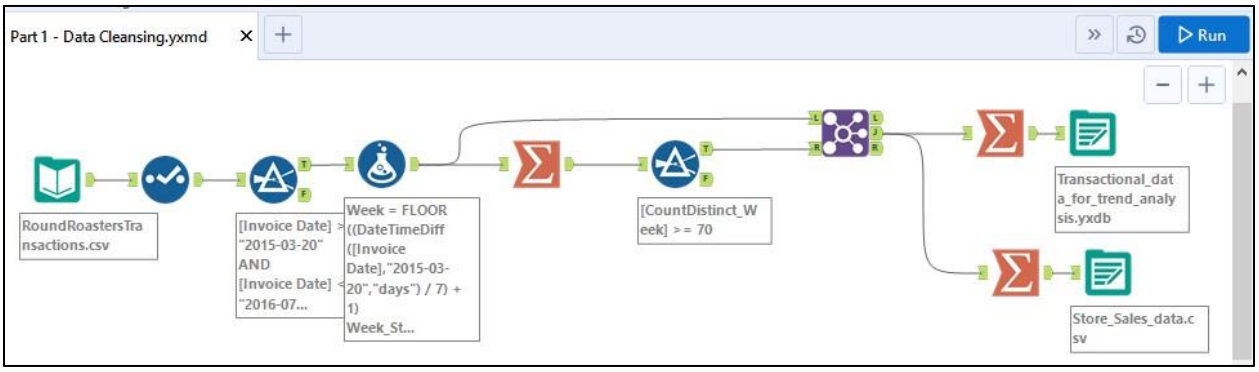
2. What is the test period?

   The test period is 12 weeks long ranging from '2016-29-04' to '2016-07-21'.

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

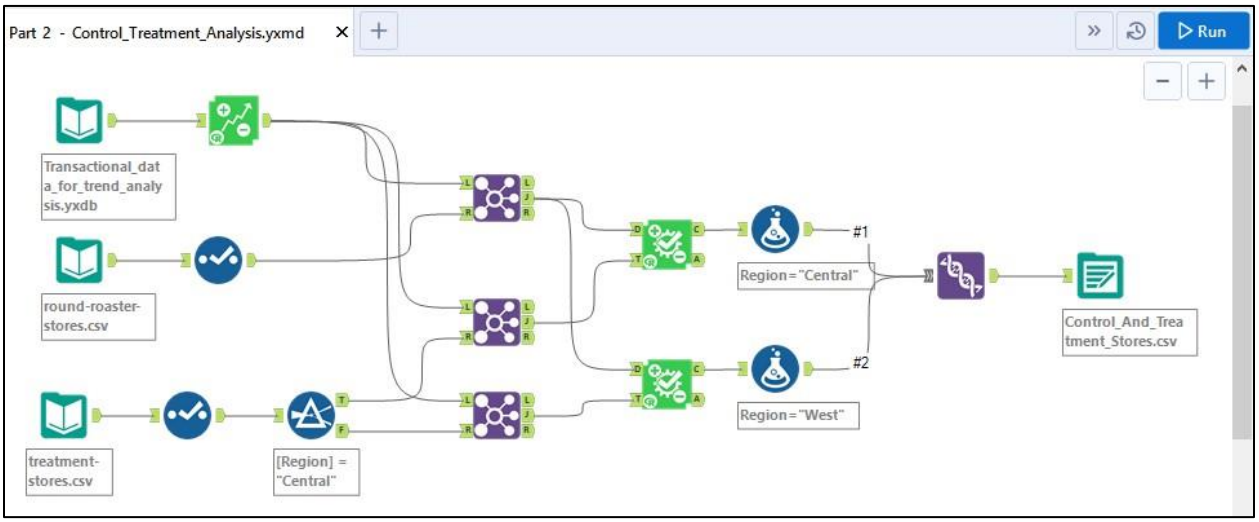
   Because the test is being run for 12 weeks, it would make sense to do the aggregation on a weekly basis per store. Therefore, the aggregation is done using Store ID and Week.

Step 2: Clean Up Your Data



- Set the correct datatypes and select the required data from 2015-03-20 to 2016-07-21 (the 12-week test dates and the 58-week window as required by Alteryx for the A/B Testing
- Created the fields – Week, Week_Start and Week_end – for aggregating weekly data
- Aggregated the data by Store ID and Gross Margin.
- Filtered out stores with data less than 70 weeks.
- Joined with original data to get the full information.
- Aggregated the data get weekly invoice data and total gross margin data.
- Saved the output files - weekly sales data as Store_Sales_data.csv and the total gross margin data as Transactional_data_for_trend_analysis.yxdb files.

Step 3: Match Treatment and Control Units



Apart from trend and seasonality...

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

   Trend, Seasonality, Sq_Ft, and AvgMonthSales.

2. What is the correlation between your each potential control variable and your performance metric? (Example of correlation matrix below)

   Full Correlation Matrix

	Sum_Gross.Margin	Trend	Seasonality	Sq_Ft	AvgMonthSales
Sum_Gross.Margin	1.000000	-0.066564	0.088777	-0.020453	0.988207
Trend	-0.066564	1.000000	-0.772280	0.197444	-0.132040
Seasonality	0.088777	-0.772280	1.000000	-0.229980	0.153059
Sq_Ft	-0.020453	0.197444	-0.229980	1.000000	-0.046967
AvgMonthSales	0.988207	-0.132040	0.153059	-0.046967	1.000000

   With respect to Sum_Gross_Margin, the highest correlation was with AvgMonthSales (0.988207) and the lowest with Sq_Ft (- 0.020453).

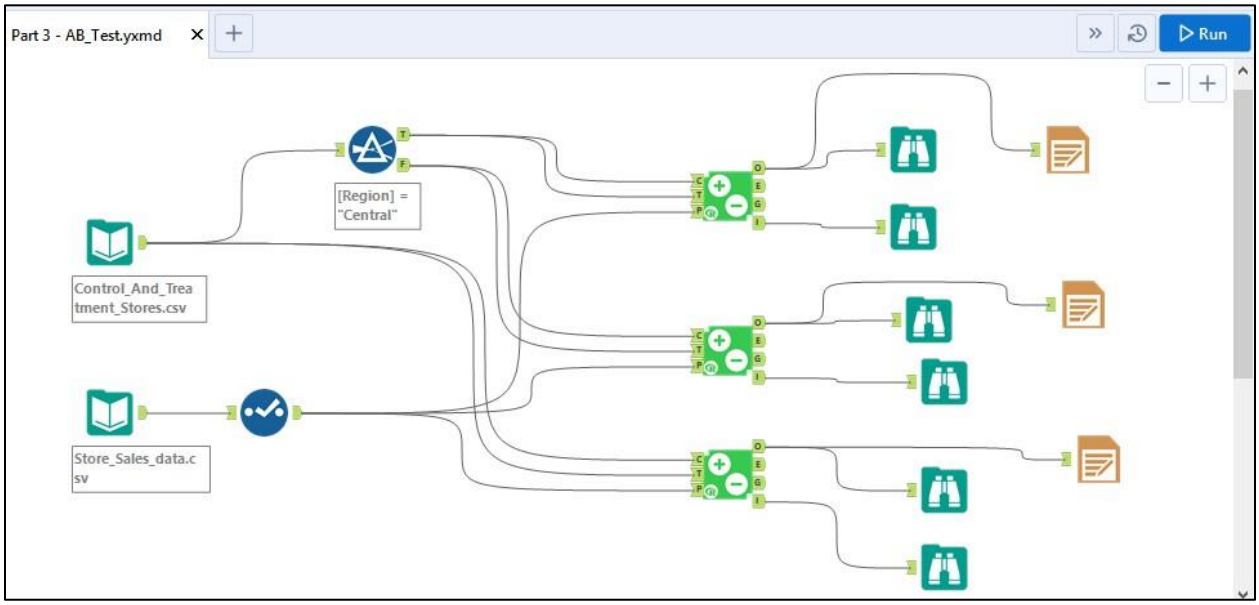
3. What control variables will you use to match treatment and control stores?

   The control variables that make most sense are AvgMonthSales, Trend and Seasonality. Sq_Ft has very low correlation and therefore, will not be considered.

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

Treatment Store	Control Store 1	Control Store 2
1664	11868	12019
1675	2409	3235
1696	2383	7334
1700	1508	8717
1712	7434	9017
2288	1857	7484
2293	7770	7811
2301	1863	11268
2322	1807	7284
2341	7584	8817

Step 4: Analysis and Writeup



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

   The company needs to show at least 18% increase in profit growth compared to the comparative period while compared to the control stores (*incremental lift*) to justify the menu launch. From the A/B test, we see an overall lift of 40.2%. Therefore, I recommend that the company roll out the updated menu to all the stores.

2. What is the lift from the new menu for West and Central regions?

   The lift for West region is 33.3% with a statistical significance level of 99.2% while the lift for Central region is 47.1% with a statistical significance of 99.6% respectively.

   Lift Analysis for West Region

   Lift Analysis for Sum_Gross Margin

Lift	Expected Impact	Significance Level
33.3%	466	99.2%

   Lift Analysis for Central Region

   Lift Analysis for Sum_Gross Margin

Lift	Expected Impact	Significance Level
47.1%	855	99.6%

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

   The overall lift for the new menu for all regions is 40.2% with a statistical significance of 100%.

   Lift Analysis for Sum_Gross Margin

Lift	Expected Impact	Significance Level
40.2%	660	100.0%