Project: Analyzing a Market Test

Plan Your Analysis

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

   Gross margin would be the performance metric use to evaluate the results of the test. 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.

2. **What is the test period?**

   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.

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

   The data should be aggregated at the week level.

Clean Up Your Data

Data Preparation Steps:

1. The data sets are added using *Input* tool.
2. Use *Select* tool to keep or remove fields.
   - Round-Roasters-Transactions: StoreID, Invoice Number, Invoice Date, Gross Margin and Sales
   - Round-Roaster-Stores: StoreID, Sq_Ft, AvgMonthSales and Region
3. Use *Join* tool to join Round Rosters transaction data to Round Rosters stores data so that the store transactions are matched with store information using **StoreID**.
4. Use *Filter* tool to limit the records to the 76 weeks of data from 6 Feb 2015 to 21 Jul 2016.
   - [Invoice Date]>="2015-02-06" AND [Invoice Date]<="2016-07-21"

To leverage the AB Analysis tool in Alteryx, at least 1 full year of data prior to the test, plus an additional 12 weeks is required to calculate trend and seasonality. Therefore, we need a total of 76 weeks of data in total.

Historical Data	52 weeks
Trends	12 weeks
Experiment	12 weeks
Total	**76 Weeks**

5. To assist with summarizing transaction data to store in week level, **Week** field, **Week_Start** field and **Week_End** field will be created.

Week	FLOOR(DateTimeDiff([Invoice Date], "2015-02-06", 'days') / 7+1)
Week_Start	DateTimeAdd('2015-02-06',7*([Week]-1),'days')
Week_End	DateTimeAdd([Week_Start]),6,'days')

6. Use *Summarize* tool to create a count of weeks for each store.
   - GroupBy **Store** and CountDistinct **Week**
7. Use the *Filter* tool to keep only those store that have 76 weeks of data.
8. Use a *Join* tool to add back in the transaction data but only for those stores that are identified as having 70 weeks of data. Join using the **Store** field and remove fields that are not needed.
9. Use *Summarize* tool to get the total gross margin and sales amount per invoice. GroupBy **region**, **storeID**, **invoice number**, **invoice date**, **week**, **week_start**, **week_end** and **AvgMonthSales**.
10. Use *Summarize* tool to get the number of invoices per store per week. GroupBy **storeID**, **week**, **week_start** and **week_end**.
11. Use *Output data* tool to export the data set that shows the invoices per store per week along with the data information.

Record #	StoreID	Week	Week_Start	Week_End	Count	Sum_Sum_Gross Margin
1	10018	1	2015-02-06	2015-02-12	308	2212.7105
2	10018	2	2015-02-13	2015-02-19	288	2164.007
3	10018	3	2015-02-20	2015-02-26	204	1560.929
4	10018	4	2015-02-27	2015-03-05	320	2342.984
5	10018	5	2015-03-06	2015-03-12	284	2199.4065
6	10018	6	2015-03-13	2015-03-19	288	2103.143
7	10018	7	2015-03-20	2015-03-26	194	1412.927
8	10018	8	2015-03-27	2015-04-02	286	2124.3715
9	10018	9	2015-04-03	2015-04-09	274	2216.152
10	10018	10	2015-04-10	2015-04-16	215	1686.246
11	10018	11	2015-04-17	2015-04-23	277	1938.1365
12	10018	12	2015-04-24	2015-04-30	251	1874.485
13	10018	13	2015-05-01	2015-05-07	201	1571.1225
14	10018	14	2015-05-08	2015-05-14	207	1586.3945

Figure 1 Weekly Store Traffic

12. Use *Summarize* tool group by **region**, **store** and **AvgMonthSales** to get a unique list of stores in each region.
13. Use *Import* tool to add treatment store's data and *Select* tool to keep **storeID** and **AvgMonthSales**.
14. Add a *Formula* tool to add a field to identify the treatment candidate stores. Output field as 'Test_Group' as a string with 'Treatment' as the expression.
15. With the *Join* tool to add the treatment store data to the main data.
16. Add a *Formula* tool to add a field to identify the control candidate stores. Output field as 'Test_Group' as a string with 'CC' as the expression.
17. Combine the two tables using *Union* tool.
18. Using *Output* tool to export a data set of a list of stores with the region and test group associated with each store.

Record #	Region	StoreID	AvgMonthSales	Test_Group
1	West	10018	18000	CC
2	West	10068	16000	CC
3	West	10118	13000	CC
4	West	10168	19000	CC
5	West	10218	15000	CC
6	West	10268	25000	CC
7	West	10318	16000	CC
8	West	10368	19000	CC
9	West	10418	19000	CC
10	West	10468	21000	CC
11	West	10518	11000	CC
12	West	10568	21000	CC
13	West	10618	15000	CC
14	West	11268	12000	CC

Figure 2 Store List

19. Connect a *Summarize* tool, groupby **region**, **storeID**, **week**, **week_start** and **week_end**. Sum gross margin and sales information.
20. Add an *Output* tool to store the data as store sales analysis data.

Record #	Region	StoreID	Week	Week_Start	Week_End	Sum_Sum_Gross Margin	Sum_Sum_Sales
1	Central	1508	1	2015-02-06	2015-02-12	1429.439	3124.61
2	Central	1508	2	2015-02-13	2015-02-19	2031.258	4360.54
3	Central	1508	3	2015-02-20	2015-02-26	1022.043	2217.27
4	Central	1508	4	2015-02-27	2015-03-05	1372.2495	2996.95
5	Central	1508	5	2015-03-06	2015-03-12	1698.859	3704.73
6	Central	1508	6	2015-03-13	2015-03-19	1744.468	3749.92
7	Central	1508	7	2015-03-20	2015-03-26	815.2905	1744.58
8	Central	1508	8	2015-03-27	2015-04-02	1439.5665	3089.46
9	Central	1508	9	2015-04-03	2015-04-09	1575.9425	3384.61
10	Central	1508	10	2015-04-10	2015-04-16	1154.9785	2464.87
11	Central	1508	11	2015-04-17	2015-04-23	1485.977	3188.36
12	Central	1508	12	2015-04-24	2015-04-30	1946.1405	4143.61

Figure 3 Store Sales Analysis Data

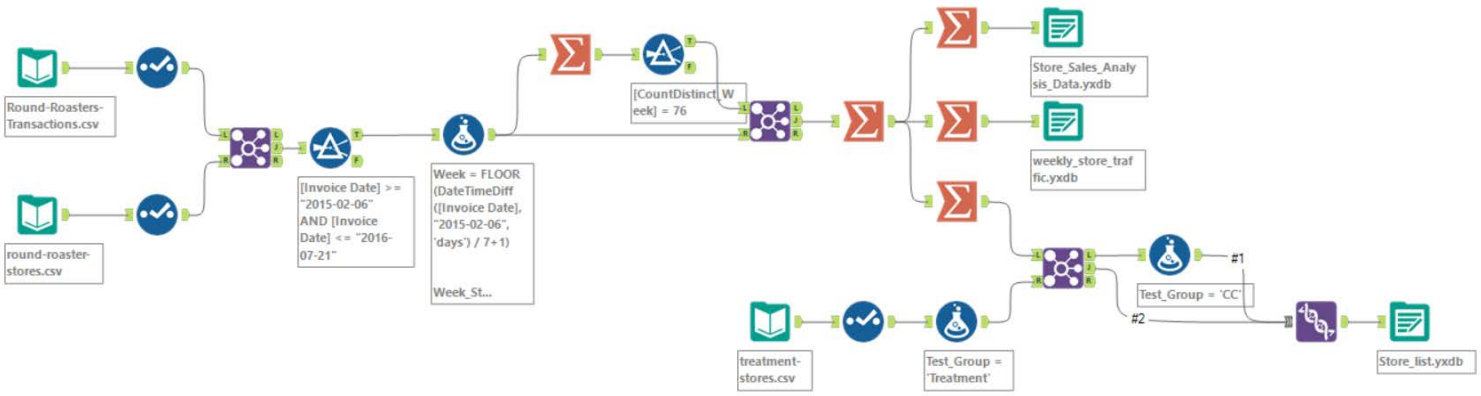

Figure 4 Alteryx Flow (Clean Data)

Match Treatment and Control Units

1. **What control variables should be considered?**

AvgMonthSales and **Sq_Ft** are both numeric variables which should be considered as constant variables.

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

Full Correlation Matrix

	Sum_Sum_Gross.Margin	Sq_Ft	AvgMonthSales
Sum_Sum_Gross.Margin	1.000000	-0.016284	0.786797
Sq_Ft	-0.016284	1.000000	-0.046967
AvgMonthSales	0.786797	-0.046967	1.000000

Figure 5 Full Correlation Matrix

Based on the full correlation matrix, **AvgMonthSales** has high correlation of 0.79 with the performance metric sum of Gross margin. On the other hand, **Sq_Ft** has a poor correlation of -0.02.

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

AvgMonthSales will be used together with **Trend** and **Seasonality** when matching treatment and control stores.

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

Treatment Store	Control Store 1	Control Store 2
1644	1964	7162
1675	7284	2214
1696	1863	7334
1700	7037	2014
1712	8162	7434
2288	2568	9081
2293	12686	9639
2301	12536	9238
2322	9388	3185
2341	2572	12586

Record #	Controls	Treatments	Distance	Region	AvgMonthSales	Test_Group
1	1964	1664	0.294607	Central	11000	Treatment
2	7162	1664	0.34613	Central	11000	Treatment
3	7284	1675	0.663813	Central	15000	Treatment
4	2214	1675	0.703483	Central	15000	Treatment
5	1863	1696	0.413188	Central	10000	Treatment
6	7334	1696	0.661492	Central	10000	Treatment
7	7037	1700	0.918504	Central	15000	Treatment
8	2014	1700	1.007526	Central	15000	Treatment
9	8162	1712	0.487861	Central	19000	Treatment
10	7434	1712	0.612087	Central	19000	Treatment
11	2568	2288	0.407287	West	14000	Treatment
12	9081	2288	0.446997	West	14000	Treatment
13	12686	2293	0.6895	West	11000	Treatment
14	9639	2293	0.730351	West	11000	Treatment
15	12536	2301	0.376984	West	11000	Treatment
16	9238	2301	0.383212	West	11000	Treatment
17	9388	2322	0.231362	West	14000	Treatment
18	3185	2322	0.279139	West	14000	Treatment
19	2572	2341	0.26221	West	11000	Treatment
20	12586	2341	0.459807	West	11000	Treatment

Figure 6 Control Treatment Pairs

Alteryx Configuration:

1. For *AB Trend* tool, select **StoreID** as unit identifier, **Week_End** as reporting period dates, **Sum_Gross Margin** as performance measure. Report period type as Weekly, 12 number of period and 29 Apr 2016 as test start date.
2. For AB Controls tool, select **storeID** as unit identifier and data measures to match treatments to controls. **Trend**, **Seasonality** and **AvgMonthSales** as numeric measures.

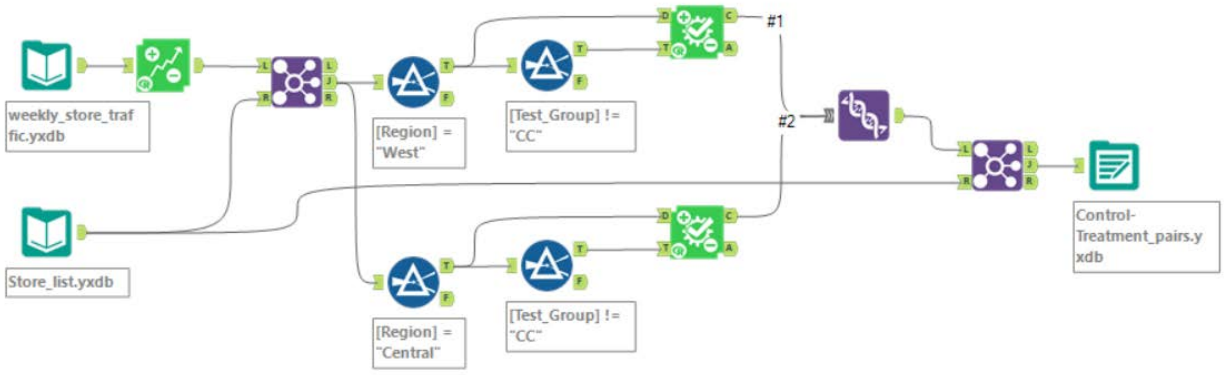

Figure 7 Alteryx Flow (Match Treatment and Control Units)

Analysis and Writeup

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

The company should roll out the updated menu to all stores as based on the A/B test analysis result, the profit growth are more than 18%. Which the predicted impact to profitability would 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.

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

West Region

Figure 8 AB Test Analysis (West Region)

The report shows that after rolling out the new menu for West region, showed 39.1% improvement at a significance of 99.6% over the existing menu. The average lift as a result from changing the new menu would be 39.1% per store per week, or approximately $530.5 per store per week.

Central Region

Figure 9 AB Test Analysis (Central Region)

The report shows that after rolling out the new menu for Central region, showed 47.6% improvement at a significance of 99.6% over the existing menu. The average lift as a result from changing the new menu would be 47.6% per store per week, or approximately $906.3 per store per week.

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

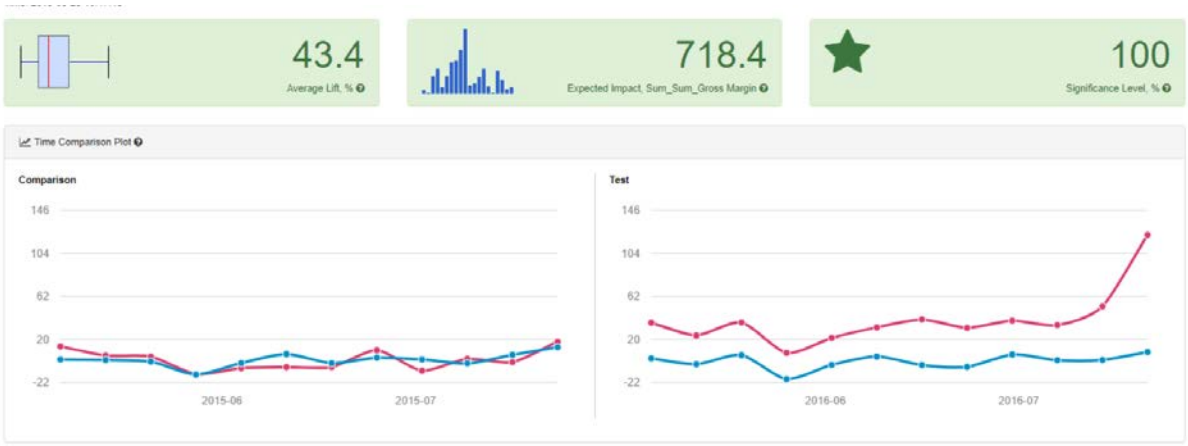

Figure 10 AB Test Analysis (Overall)

The report shows that after rolling out the new menu, showed 43.4% improvement at a significance of 100% over the existing menu. The average lift as a result from changing the new menu would be 43.4% per store per week, or approximately $718.4 per store per week.

Rolling out the new menu would improve gross margin; therefore the change should be rolled out across all stores.

Alteryx Configuration:

1. For *AB Analysis* tool, select **Controls** as control unit identifier, **Treatments** as treatment unit mapping field and treatment unit identifier. **StoreID** as performance data unit identifier, **Week_End** as reporting period data information and **Sum_Gross Margin** as performance measure.

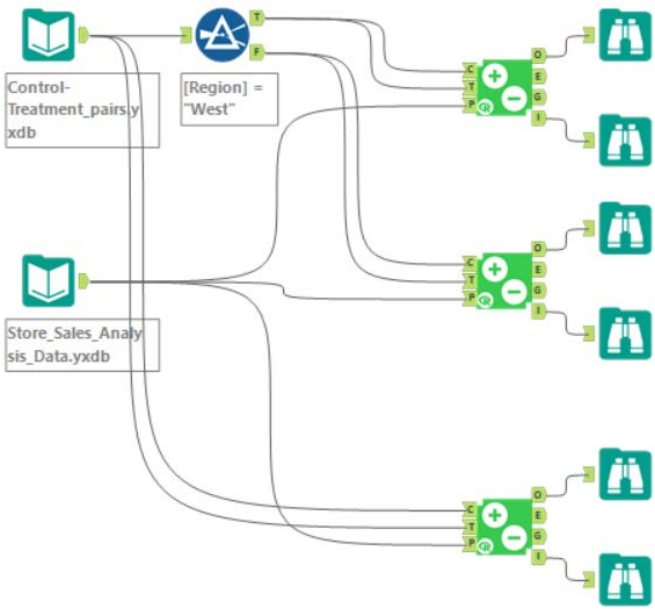

Figure 11 Alteryx Flow (AB Test Analysis)