Capstone Project Executive Summary

OVERVIEW

The following executive summary gives a detailed explanation of the thought process and steps in which we took to solve the questions asked of us. This document is intended to guide you through the actions we took in order to reach our conclusions in the most digestible and intuitive way possible.

PROBLEM

Pilot Flying J has been running the "Thungry" promotional campaign for food and beverage items at its gas stations in different divisions (geographical areas) during the 2017 and 2018 fiscal year. Consumers have the option of receiving discounts on select items or combinations of items during the specified months. Pilot Flying J would like to know the affects their promotion has had on sales.

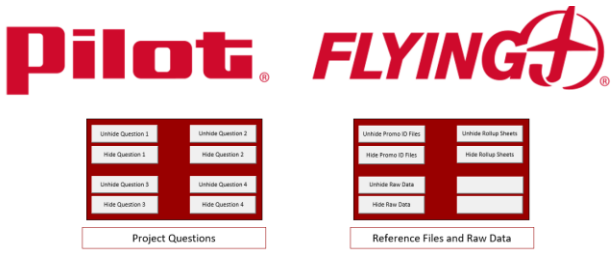
KEY TAKEAWAYS

- **The featured beverage has a very small, if any impact on promotional sales**
- **Divisions spend more on certain featured beverages regardless of promotion**
- **There is a strong, positive correlation between the number of items in a promotion and the sales it generates**
- **Certain combinations of brands are highly impacted by promotions**

WORKBOOK ARCHITECTURE

Dashboard

The Dashboard is a path to all Capstone questions and related sheets that helped us reach our conclusions. You can reach all the sheets that were used in finding our answers/insights from the dashboard.



Question buttons

The question buttons unhide said question sheets. Each question button will display one main question sheet that we used to formulate our answers. There will also be a couple of other sheets that were directly used in formatting and transforming data from the *Rollup* sheets into digestible charts and tables used on the main question sheet.

Main Question Sheet

Rollup Sheets and Raw Data

These macro buttons will unhide the *Master Rollup* sheets and sheets that were used in creating the *Master Rollup* sheets.

Promo ID Sheets

- These sheets contain promotion sales data by month and division. They list the total sales of the brands offered in the promotion for said month.

Raw Data

- The *Raw Data* sheets include *UPC Description, GroceryCooler_Sales_Month, GroceryCooler_Sales_crosstab,* and *2017-2018 Thungry Schedule*.
- Note: An =INDEX(MATCH()) was used in the *GroceryCooler_Sales_crosstab* in order to pull brand names from the *UPC Description* sheet. Also, *GroceryCooler_Sales_Month* is the same as *GroceryCooler_Sales_crosstab* but broken out by month instead of fiscal week. This was completed with a =SUM() formula.

Rollup Sheets

- These files contain summary data from the sheets hidden under the *Raw Data* and the *Promo ID Sheets Dashboard* macro buttons. We used the *Rollup* sheets to pivot and visualize data in order to reach our conclusions.

PROCESS

To begin, we were first given a zip file containing CSV and EXCEL files which contained sales and item description data. The first step we did was save all the CSV files as excel files to reduce the chances of data loss, and we combined all the workbooks into one workbook. Having all the worksheets in one place would make the entire process much more efficient.

Next, we examined the questions. Before we could work with data, we needed to understand our end goal. Being future-oriented greatly helped us work through this project. After

examining the questions, we broke each one down into key words that would help guide our data cleaning process:

1. What is the effect of running Coke or Pepsi pictures in the advertisement for each promotion? Does picturing Coke or Pepsi impact sales of either product in the promotion? What is the difference in sales when picturing one over the other? Please consider this overall and by division.

- Promotion Sales by month, featured beverage by month, promotional division sales by month

2. Are different divisions affected differently by different promotions? If so, what are the implications? Which products sell more effectively in which divisions? Are these sales directly related to promotions or are these sales regularly higher in these divisions?

- Promotion Sales by month, featured beverage by month, division sales by month, product sales by division, product sales by month

3. Does the promotion work better when Pilot Flying J offers a broad category of options in their promotion or when specific brands are offered as part of the promotion? (Is more beef jerky sold under a jerky promotion or under a specific brand promotion?)

- Number of unique items offered by promotion, total promotional sales by month

4. Are there any products that sell as well (or almost as well) whether they are on promotion or not? (This would mean that the client does not need to run them on special promotion).

- Brands in each promotion, total sales by brand by month

After breaking apart the questions, we began to clean the data. Once we started examining the worksheets, we noticed that the data we needed to answer the questions would have to come from many different sheets. So, we decided to build three *Master Rollup* sheets. The idea was that we did not need to have every individual item sold from the *GroceryCooler_Sales_crosstab* or every item sold from the Promo ID sheets. Instead, we decided to build these summary sheets that we named *Master_Rollup_Division, Master_Rollup_Month,* and *Promo_Rollup_Brand*.

Building the *Master Rollup* sheets

The first thing we wanted in the *Rollup* sheets was promotional sales by month and division. To do this, we created subtotals in the Promo ID tabs. We added sum subtotals to sales at each change in *OPER_DIVISION_ID* (division).

OPER_DIVISION_	PROMOTION_	FISCAL_WEEK_	MAIN_UPC	Brand	SALES
2 Total					$197,469.90
4 Total					$200,738.98
5 Total					$202,816.60
9 Total					$184,465.82
11 Total					$144,274.80
13 Total					$227,391.74
78 Total					$2,343.23
Total Total					$2,316,658.91

This made it very easy to copy and paste the data from the Promo ID tabs into the shell of the soon to be *Total Promo Sales table*. This process was repeated for every Promo ID sheet until the entire *Promo Sales by Division* table in the *Master_Rollup_Month* sheet was populated.

After the *Total Promo Sales* table was complete, we built three more tables, one for Coke, Pepsi, and Mountain Dew. This process of building these three tables was slightly different from the *Total Promo Sales table*.

First, we built the *GrocceryCooler_Sales_byMonth* sheet. This sheet was created from *GroceryCooler_Sales_crosstab* to summarize purchases of every product by month. We thought the data would be easier to digest and to insert into the *Master Rollup* sheet if it was summarized. To keep our data consistent, we decided to use the last purchase by fiscal week code in the Promo ID sheets to determine when to create our month total columns in the *GrocceryCooler_Sales_byMonth* sheet. For example, the last fiscal week code in the *13798 Jan-Feb 2017* sheet was "20170308." So, in the *GrocceryCooler_Sales_byMonth* sheet, we would insert a new column after the same fiscal code and title the column *Jan/Feb 2017*. Then we used a =SUM() formula to calculate the total sales for each UPC for that month. Then we used the fill handle to copy the formula down to the bottom of the table. We continued this process until we ran out of columns.

Next, we built shells for the tables and began to input total sales by division and month. To input the data, we used a =SUMIFS() formula.

COKE							
DIVISION 2	DIVISION 3	DIVISION 4	DIVISION 5	DIVISION 6	DIVISION 7	DIVISION 8	DIVISION 9
$ 633,356.91	$ -	$ 429,231.41	$ 498,011.01	$ 30,128.50	$ -	$ -	$ 347,224.33

=SUMIFS(GroceryCooler_Sales_Month!$AF$5:$AF$51805,GroceryCooler_Sales_Month!$A$5:$A$51805,D$45,GroceryCooler_Sales_Month!$C$5:$C$51805,"*COKE*")

This formula is set to sum any value in *Mar/Apr 17* column of the *GrocceryCooler_Sales_byMonth* if the value's associated divisions number equaled "DIVISION 2" (D$45 in the COKE table) and if the brand name included the word "COKE". Because there was a "COKE" and "DIET COKE" brand name we used a wildcard in the formula to look for any brands that included "COKE." The same formula was used for the Pepsi and Mountain Dew tables too, but we changed the sum criteria to "*Mountain Dew*" and "*Pepsi*".

Once the tables were complete, we built smaller tables out to the right that summed total sales by month and featured beverage. We also used conditional formatting on the total sales and sorted the sales from largest to smallest. Building the tables with the month and featured

beverage and then sorting by largest to smallest allowed us to see exactly which months had the highest sales per beverage and which beverage was featured in the promotion.

Once we completed the *Master_Rollup_Month* sheet we created the *Master_Rollup_Division* sheet to display the same data but with more emphasis on the divisions. Creating the *Master_Rollup_Division tables* involved copying and pasting (transposing and keep number format) the values from *Master_Rollup_Month* into the *Master_Rollup_Division.* The smaller summary tables out to the right were created the same way the ones in the *Master_Rollup_Month* were created.

It is important to note that in the *Master_Rollup_Division* sheet there are less divisions than in the *Master_Rollup_Month* sheet. We removed all division that did not have any Coke, Pepsi, Mountain Dew, or promotional sales. Also, if there was an instance where there were sales in division X but no promotion sales in division X, we removed it. We believed that a division that didn't sell any of the featured beverages or didn't partake in the promotion should not be included in our results.

In short, the *Master_Rollup_Division* sheet is a refined version of the *Master_Rollup_Month* sheet with transposed table headings. We used the *Master_Rollup_Division* for most of our charts and calculations because there were less holes in the data.

Creating the *Promo_Rollup_Brand* Sheet

Because questions 3 and 4 required us to have information on the brands within each promotion, and the number of brands in each promotion, we created the *Promo_Rollup_Brand* sheet.

To create this sheet, we used the *Promo ID* files. These files contained the brands associated with each promotion. In order to get each unique brand name, we copy and pasted the brand column outside of the table in the sheet. Then, we removed duplicates in the new brand column we had pasted. This allowed us to see the exact brands in each promotion. To count the amount of items in each promotion we used a =COUNTA() formula. Then, we copy and pasted this data from this *Promo ID* sheets into the *Promo_Rollup_Brand* sheet.

We used these Rollup sheets for all the questions.

QUESTIONS

Once we had completed the master rollup sheets, we felt we had cleaned and formatted the data to where it was ready to be pivoted and manipulated to answer the questions.

1. What is the effect of running Coke or Pepsi pictures in the advertisement for each promotion? Does picturing Coke or Pepsi impact sales of either product in the promotion? What is the difference in sales when picturing one over the other? Please consider this overall and by division.

To answer this question, we pivoted the *Promo Sales* and the *Total Promo Sales* tables in our *Master_Rollup_Division* sheet and created two charts. One shows the sum of all sales for all divisions by month. The other is the same graph but has a slicer that is filtering the data by division. There are also two charts that were created from the *Promo Sales by Month* table in the *Master_Rollup_Month* sheet which shows the total sales by beverage for 2017 and 2018.

We also included the smaller summary tables from the *Rollup* sheets.

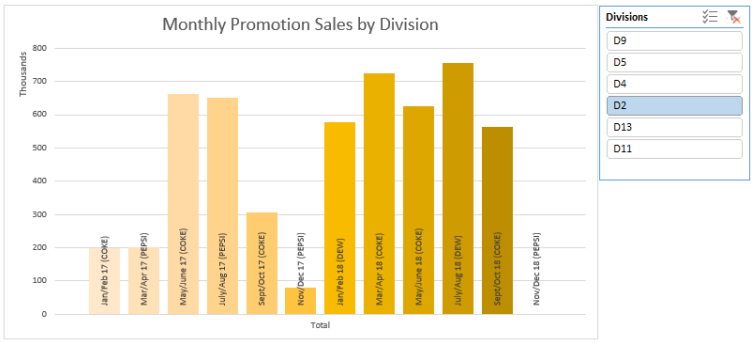
From the data, it seems there is a lack of evidence to conclude that certain signage has an impact on "Thungry" promotional sales.

When looking at total promotion sales, independent of month, featuring Coke results in the most amount of sale. But when looking at sales by month the data tells a different story.

In most cases, featuring Coke does result in more sales. Although, there are other times when sales are higher when featuring Mountain Dew or Pepsi. All featured beverages seem to have a random impact on total sales. For that reason, we do not believe the sales are directly correlated to featured beverages. This finding is also consistent with the sales by division data found in the *Monthly Promotional Sales by Division* graph.

The only real trend that we can see is that sales tend to depend upon the time of year. The five top selling months from the total promotional sales are Mar/Apr 2018, May/June 2017, May/June 2018, July/Aug 2018, and July/Aug2017. Also, while the top three months featured coke, the other two featured Mountain Dew and Pepsi, respectively. It seems that the promotion is more affective during warmer parts of the year.

Something else to note, the highest sales month for all divisions was Mar/Apr 2018, except for Division 2. Division 2's highest sales month was July/Aug 2018, when Mountain Dew was the featured beverage. The sales were not exponentially higher than other months, but they were still higher. Which could mean Division 2 prefers Mountain Dew.



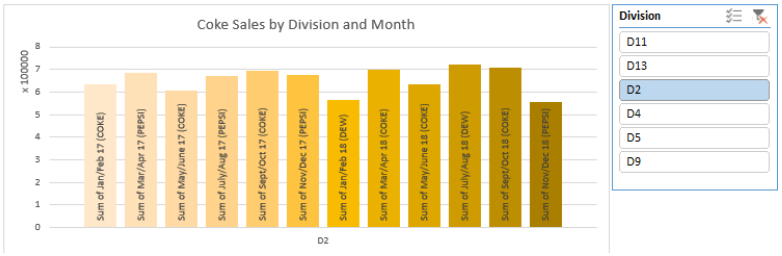
Sheet: Question_1

2. Are different divisions affected differently by different promotions? If so, what are the implications? Which products sell more effectively in which divisions? Are these sales directly related to promotions or are these sales regularly higher in these divisions?

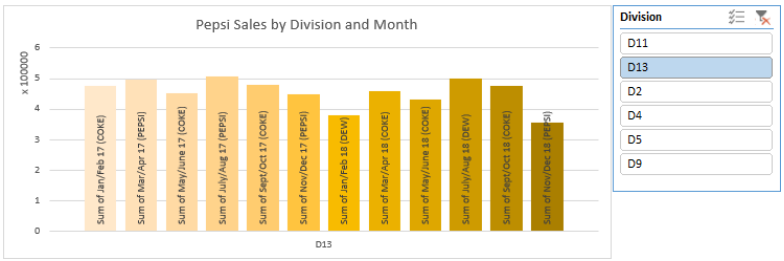
To answer this question, we used the *Promo Sales* table from the *Master_Rollup_Month* sheet and the *Coke, Pepsi,* and *Mountain Dew* tables from the *Master_Rollup_Division* sheet. This data would allow us to see promotional sales by division and featured beverage sale by division. First, we created a *Monthly Sales by Division* graph like the one for question 2. Then we pivoted the *Coke, Pepsi,* and *Mountain Dew* tables from the *Master_Rollup_Division* sheet in order to create charts that we could insert slicers into. We also used the smaller, summary tables, *Coke Sales by Division, Pepsi Sales by Division,* and *Mountain Dew Sales by Division, to* create three charts that showed total featured beverage sales by division.

From our data, we believe that total sales for certain featured beverages are indeed higher in certain divisions: Division 2 prefers Coke more than other divisions, Division 13 prefers Pepsi more than other divisions, and Division 11 prefers Mountain Dew more than other divisions. Although, we do not believe that the promotion had an impact on sales of featured beverages by that much, if at all. We met this conclusion by filtering the *Coke, Pepsi, and Mountain Dew Sales by Division and Month* graphs by the division that accumulated the most sales of that product.

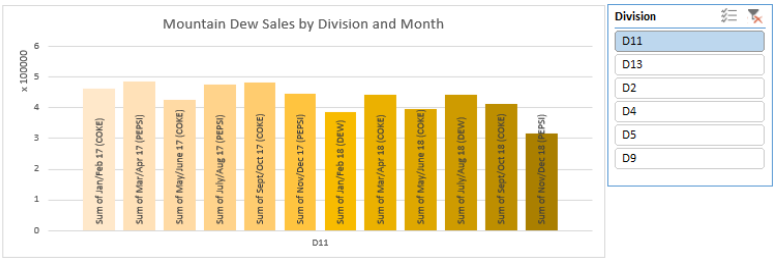
Coke Sales by Division	
D2	$ 7,856,033.86
D13	$ 7,218,822.72
D11	$ 6,389,910.03
D5	$ 6,318,180.43
D4	$ 6,035,899.20
D9	$ 4,607,754.49



Pepsi Sales by Division	
D13	$ 5,464,163.37
D5	$ 4,171,184.01
D2	$ 3,740,511.87
D9	$ 3,506,371.33
D4	$ 3,139,511.28
D11	$ 2,425,065.22



Mountain Dew Sales by Division	
D11	$10,895,459.12
D2	$ 9,199,816.96
D13	$ 8,590,156.37
D9	$ 7,277,956.43
D4	$ 5,172,609.37
D5	$ 4,788,404.61



Sheet: Question_2

From the graphs, sales of featured beverages seem to be independent of the promotion. For instance, Division 2, which spent the most on Coke products, had some of their highest coke spending months during the promotion of Pepsi and Mountain Dew.

It is also important to note that certain divisions are in fact faced with heavier spending during certain promotions. Like what was said in question 1, Division 2 was the only division to have their highest spending month not be Mar/Apr 2018. Their highest selling month was July/Aug 2018, when Mountain Dew was advertised. Although Division 2 did not spend the most on Mountain Dew products, it could still prefer them over Coke and Pepsi. We came to this conclusion because the population and number of Pilot gas stations was not listed, which could make this data misleading.

The key takeaway from this is that certain divisions do spend more on certain featured beverages. But this spending seems to remain constant throughout the year and is not impacted by promotion. Although, Promotional sales do spike during certain times of the year, and drastically spike during certain months. This could be because of the brands that are sold within each promotion.

3. Does the promotion work better when Pilot Flying J offers a broad category of options in their promotion or when specific brands are offered as part of the promotion? (Is more beef jerky sold under a jerky promotion or under a specific brand promotion?)

To start question three, we went back to our initial question key points and combed the Rollup sheets for said data. The data needed to answer this question was in the *Master_Rollup_Month* sheet and the *Promo_Rollup_Brand* sheet. We took the total sales by month from the *Total Promo Sales* table in the *Master_Rollup_Month* sheet and the number of items in each promotion form the *Promo_Rollup_Brand* sheet. This data was then combined in the *Sales_by_Item* sheet under the *Question 3* macro button.

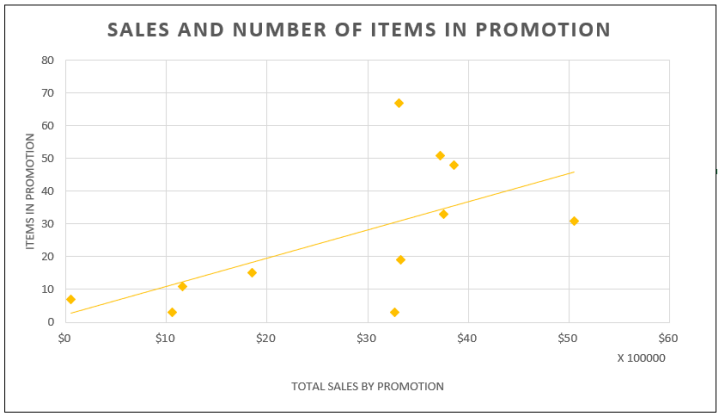
We built a table and used conditional formatting and sorting to find out the total promotional sales dependent on the number of items in the promotion. Once formatted, we created a scatter chart to show the correlation between the two variables. We also used the formula =CORREL() to get the exact correlation between the two variable. Having both numerical and visual data seemed like the most beneficial way to solidify our findings.

We also created another chart that did not include, what we believed to be, an outlier. This point (outlier) was the only point that really stood out and did not fit in with the rest of the data.

Based on the correlation we saw between the total sales by promotion and number of items in a promotion, we believe that promotions works better when more items are offered.

The first chart, which includes all data points, has a moderate correlation of 42%. Then, once the outlier was removed, we saw a correlation of 61%, which is considered strong.

Because of this, we believe that the more items Pilot offers in its promotions the more sales it will generate.



CORRELATION: 61%

Sheet: Question_3

4. Are there any products that sell as well (or almost as well) whether they are on promotion or not? (This would mean that the client does not need to run them on special promotion).

To start, we retrieved the data necessary to answer this question: brands by promotion, and total sales by brand by month. The brands by promotion was retrieved from the *Promo_Rollup_Brand* sheet. Total sales by brand by month was taken from the *GroceryCooler_Sales_Month.* In order to get the results we wanted, we thought we should use advanced filtering and macros to change the filter.

First, we created a copy of the *GroceryCooler_Sales_Month* sheet for the advanced filter in the sheet *Q_4 Chart and Table*. We created a copy so the macros could work without the *Raw Data* sheets present. Next, we created a table using =SUM() at bottom of the *Q_4 Chart and Table* to show total sales by month. Then we copy and pasted the brands by promotion data from the *Promo_Rollup_Brand* sheet at the bottom of the *Question_4* sheet (again, in order to run the macro with the *Raw Data* sheets hidden). Then, we created macros that would run an advanced filter on that data table based on the brands in each promotion. The table of brands that were used in the advanced filter are located at the bottom of *Question_4.*

With this, we could see how well the bundle of goods within each promotion sold whether there was a promotion or not. This would allow us to see if the promotion had an impact on monthly sales or if said bundle of goods would sell as well independent of a promotion.

Also, On *Question_4*, because you cannot filter our chart unless the previous filter was cleared we included a =IF(AND()) formula that would let the user know if the chart was "Ready" or "Needs Reset." The formula would display "Ready" if the total sales per month matched the *Reset Criteria* table(copy of total sales per month not affected by the filter) at the bottom of the *Q_4 Chart and Table* sheet and if cell *A51805* contained the value "Division 13." This was the cell value when the table was not filtered. If these two criteria were met, then the table had to have cleared all preexisting filters. And, if these criteria were not met the table must still contain a filter, thus resulting in the formula to display "Needs Reset."

We also used conditional formatting on the table below the chart which would highlight the highest *Month Sales* and change as the filter changed the monthly sales.

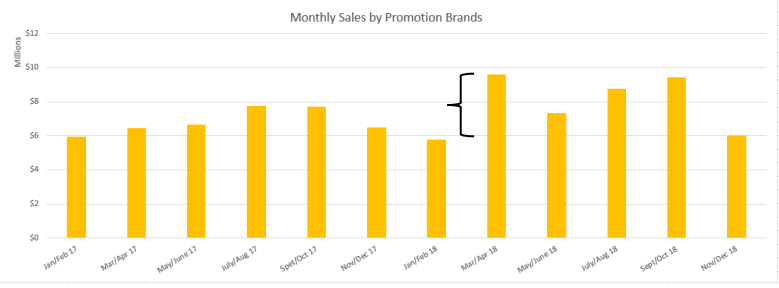
Once complete, this chart would let us see how each bundle of goods did by month. This would allow us to see if the promotion had an impact on the sales of the brands it offered. So, if there was not a spike in sales during the month that the promotion was offered then we would know the promotion did not have an effect.

From the chart and *Month Sales* column, we can conclude that the following promotions had an impact on the bundle of goods it they were promoting:

"Thungry Buy Drink GetChxChzCmb B1G1F" (Mar/Apr 2017)

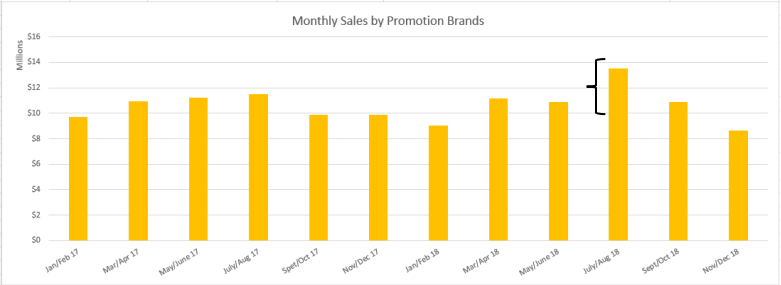
Month Sales
$ 763,695.58
$ 1,607,102.00
$ 899,141.38
$ 1,049,134.37
$ 874,690.78
$ 950,276.86
$ 735,401.51
$ 1,016,822.28
$ 1,279,072.98
$ 1,230,054.29
$ 917,174.75
$ 732,899.04

"Thungry BOGO Jerky Bags" (Mar/Apr 2018)



Month Sales
$ 5,952,773.57
$ 6,438,826.45
$ 6,654,357.36
$ 7,752,125.20
$ 7,715,841.78
$ 6,476,275.54
$ 5,788,206.85
$ 9,609,806.61
$ 7,329,387.47
$ 8,745,229.02
$ 9,421,212.76
$ 6,022,089.62

"Thungry BOGO Candy Peg Bags and Pouches" (July/Aug 2018)



Month Sales	
$	9,702,052.43
$	10,922,037.53
$	11,233,786.81
$	11,489,308.17
$	9,864,997.11
$	9,862,027.21
$	9,033,759.75
$	11,187,013.83
$	10,887,815.24
$	13,498,122.36
$	10,906,346.39
$	8,628,721.36

"Thungry BOGO SJ, Legend, Cattlemans Cut" (Sept/Oct 2018)

Month Sales	
$	2,420,536.13
$	2,316,705.02
$	2,369,549.12
$	2,764,311.35
$	3,539,418.57
$	2,363,341.43
$	2,116,822.90
$	3,781,040.91
$	2,484,702.60
$	3,265,740.08
$	5,178,924.14
$	2,170,266.53

Sheets: Question_4

In conclusion, the products that were offered in the above promotions sell better when they are a part of a promotion.

All other bundles of items seem to generate the same amount of sales whether under a promotion or not.

INSIGHTS

From our findings, we would recommend Pilot Flying J change the following with the "Thungry" promotion:

- Increase the average number of goods provided in each promotion
- Put an emphasis on prompting only brands that need promoting
- Advertise the snack items offered in each promotion – not the beverage

Also, because of the constant low sales during winter months, it may be beneficial to invest in hotter, more winter-oriented drinks. Possibly, having coffee as a featured beverage during colder months could have an impact on sales. We would imagine that an ice-cold Coke probably doesn't sound that enticing to a trucker who's been driving in 30-degree weather for 7 hours.

CONTRIBUTIONS

Alyssa Marie Menavich: Q1

James Graham Chalfant: Q2

Mary Catherine Whitfield: Q3

James Graham Chalfant: Q4

Georgia Louise Cotton: Executive Summary