W. P. CAREY
SCHOOL of BUSINESS
ARIZONA STATE UNIVERSITY

**W. P. CAREY SCHOOL OF BUSINESS HONOR CODE MASTER'S PROGRAMS**

We, the students of the W.P. Carey Master's programs, are committed to maintaining the highest levels of honesty, integrity, and professional courtesy for ourselves and amongst our classmates. We believe that ethical behavior is not only the cornerstone of a good academic program, but of good business as well. As such, we are steadfast in our commitment to the following principles:

- Academic Integrity – We view the W.P. Carey Academic Integrity Policy as a living, breathing document, not just concepts on a page. As such, we are resolute in ensuring that we abide by the principles contained within, both in letter and in spirit. Furthermore, we will hold our classmates to the same high standards to which we hold ourselves.

- Consideration of Others – As W. P. Carey students, we believe that we should treat others as we wish to be treated. To this end, we will be considerate of classmates, professors, and recruiters in all aspects. Furthermore, we will constructively add to the collaborative environment by helping and encouraging others.

- Professionalism – As W. P. Carey students, we believe that the behaviors we practice today will become a habit in the future. To that end, we are committed to act in a professional manner in all school settings, including classes, networking events, and interviews. We will be cognizant of the fact that we are representing the program and our classmates at all times.

Full Name (printed): ______Sara Huntley________________

Signature___ ___________ Date ____4/18/2019_______

Full Name (printed): __Cameren Schaer_________________

Signature______ _________________ Date ____4/20/19_____

Full Name (printed): Adam Friio

Signature Date 4/21/19

Full Name (printed): Anuja Vaidya

Signature__ Date _____04/18/2019_________

Popeyes Louisiana Kitchen Data Analysis

Team S-01

Sara Huntley
Anuja Vaidya
Cameren Schaer
Adam Friio

## Table of Contents



## 1. Executive Summary

A fast food franchise owner, Paul Singh, recently opened a new Popeyes Louisiana Kitchen in Lynden, Washington. Paul is interested in applying data-driven analytical techniques in order to maximize the profitability of the new restaurants. Provided with a limited five month data set and minimal direction, our approach was to analyze each of the key profit margin drivers that Paul outlined, including net sales, customer count, location near the Canadian border, and labor hours and their associated costs. The data we received from Paul was a set of manually populated Excel spreadsheets that tracked daily net sales, customer counts, the breakdown of drive thru sales, dine in sales and several other attributes. In order to complete our analysis, we relied primarily on Excel and Tableau to identify correlations, project future sales, and prepare an optimized labor schedule.

The key results of our analysis focused on labor optimization and drivers of net sales. Using the data we were provided, we projected net sales for a typical week, and utilized those results to create an optimal labor schedule to bring labor hours to within a threshold that Paul desired. Similarly, we disproved an assumption Paul had about traffic crossing the Canadian border being a driver of net sales.

We were able to draw two main conclusions. The first conclusion was that Paul is scheduling too many employees from Monday to Thursday, and would be better off shifting those employees to the Friday and Saturday shifts. The second conclusion was that traffic to and from the border wasn't as impactful as Paul believed. Paul told us that border traffic was a significant source of his sales, however we found very little correlation between border traffic and sales. Our recommendation was to focus marketing efforts on local customers, as they are the drivers of his sales.

The limitations of our approach to this analysis were related to the data we were provided. Limited in volume, veracity, and variety, we could only scratch the surface of the analysis we could perform and validate. We found that the data was heavily skewed as a result of the increased sales the restaurant experienced during its first month of operations. While we feel our approach was sound, without a years worth of data, at a minimum, it is difficult to account for seasonality and adjust for the aforementioned skew.

## 2. Background

The fast food industry has a  profit margin of less than 3%, on average[1], and gross profit margins are heavily impacted by food shrinkage and excessive or too little labor cost. As a franchisee, Paul, would want to decrease immoderate spending in labor and raw food purchases and increase profits by tending to more customers. Focusing on key factors like these, the overall net profits would increase and improve the success of his restaurants. Most of the data for the restaurants is collected and managed by the franshier, which was not accessible by our team. Paul tracks some of his own sales data, manually, for personal use to determine the performance of his restaurants. He has a lot of experience in the fast food industry and has been able to make some assumptions and calculations to determine different key performance indicators (KPIs). Our team offered to take the data a step further, utilizing the analytical techniques we learned in our coursework to provide additional insight for the restaurant owner. While providing our own insights to the data, we also reviewed the sales drivers Paul had provided to determine if these were true KPIs.

## 3. Problem Statement

Paul has years of experience as a franchisee and has been able to keep up with his business solely on his own knowledge of the industry. As the owner of a few restaurants, Paul is able to have some say in how the business is run and can make decisions to optimize performance. He collects his own data but has never been able to take a deep dive into the analytics.

In the fast food industry there are two main controllable costs that the franchisee has influence over. They are food loss (shrinkage, over produced) and labor costs. Without access to food costs and quantities, Paul tasked us with trying to optimize his labor cost without compromising his operational performance or level of customer service. He understands, and has experienced, the quality loss function without knowing it as such. He was looking for any insight we may provide to mitigate the impact he has experienced during his tenure as a franchisee. This includes a spike in sales during the first few months of opening, followed by a substantial correction until a normalization over the following months is achieved. With two new restaurants opened within 3-6 months of each other and a reference point of a more established location, Dairy Queen in Marysville as a sustainable control.

It is important for Paul to stay on top of the data and trends of his restaurant for the success of his business. Should he fall short in any way, the business could fail and he would be forced to close his restaurant. If Paul can determine where to improve in the process it can lead to a more optimized restaurant, better customer service, and increased net sales.

Our team has approached the problem that was presented to us with an open mind. We do not have much experience in the fast food industry, so it took some time for us to understand the process and what decisions Paul is able to make as a franchisee. Although there are some ways we were taught to test for increasing profits like coupons or sale items, we assumed that these were not something Paul could control, and therefore not an option for his franchises. We put our focus on what we could do with the data we were given and gave thought to outside data sources we could incorporate like weather data and border traffic.

[1] https://smallbusiness.chron.com/average-profit-margin-restaurant-13477.html

## 4. Methods

*<u>Why we took the approach we did</u>*

Due to the NDA restrictions, the only data that was made available to us is what Paul collected via Excel and pales in comparison to what the franchise business collects at the Point of Sale (POS) that are connected to the franchisor mainframe. The level of detail that Paul collects and calculated focus on the following:

- Number of customers based on the number of transactions at the till
- Weather
- Net sales
- Drive thru sales
- Labor percentage

Initially, our team started with an exploratory, descriptive data analysis. This included testing correlations between various weather metrics, comparing the population of Lyden and Marysville to see its impact on sales, the difference between drive thru sales and dine in sales, labor costs as a percentage of net sales, and employee discounts and its impact on net sales. The charts below (Figures 1.0 - 1.4) provide a sample of our preliminary findings:


*Figure 1.0 - Net sales of all restaurants, including the Popeyes locations in Marysville and Lynden and the DQ in Marysville.*

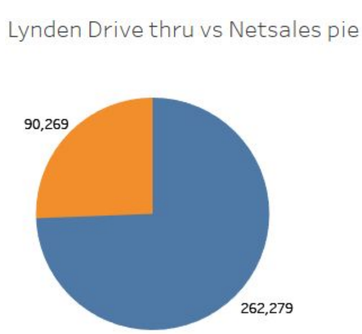



*Figure 1.1 - Drive Thru Sales vs. Dine in Sales*

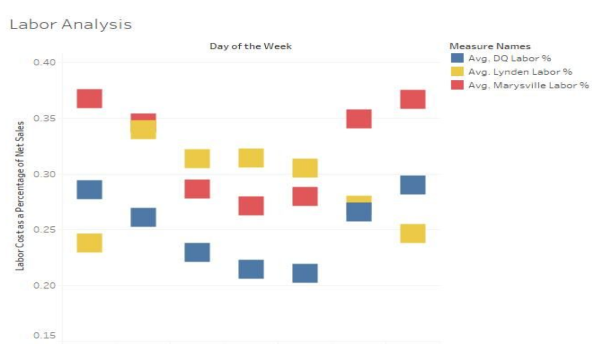



*Figure 1.2 - Labor Analysis as a % of Net Sales for all restaurants*


*Figure 1.3 - Lynden correlation between employee discount and net sales*

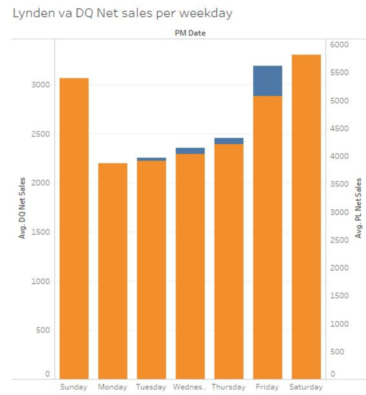



*Figure 1.4 - Lynden vs. DQ, net sales per weekday*

After encountering a number of dead ends, we met with Paul to discuss our preliminary findings. This meeting helped us to identify the key point of emphasis moving forward, namely labor costs and the impact on the location of the restaurant influence net sales. Paul explained to us that labor hours are sensitive in his restaurants, in that too few employees at any given time impacts the restaurant cleanliness, level of customer service and speed of service to the customer. Conversely, having too many employees at one time impacts his profit margin, as he has to cover the increased costs. He found his optimized labor cost to be around 25% of daily net sales. With the new Popeyes location in Lynden, he is experiencing swings between 24 - 34% and averaging about 29%, as seen in Figure 2.0 below.

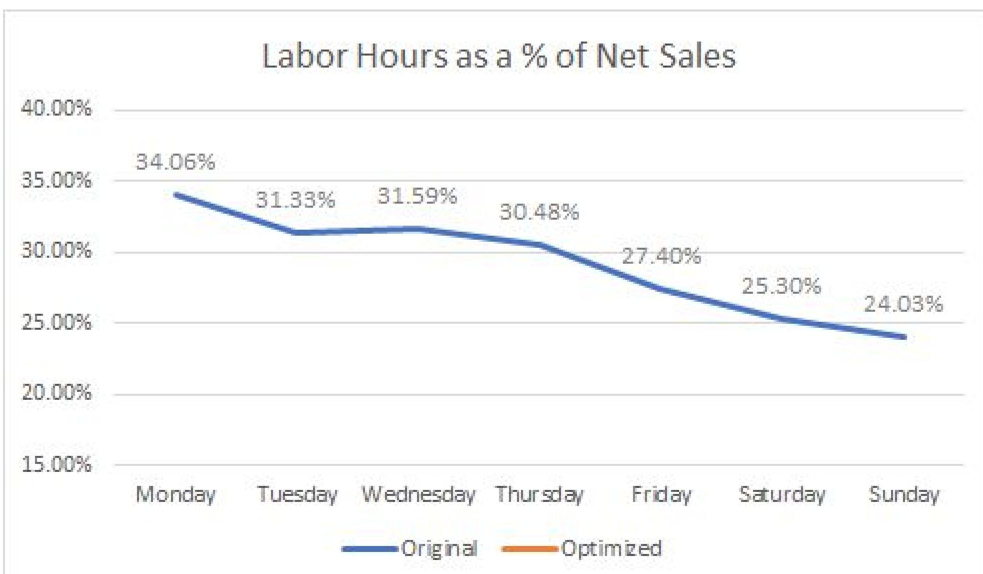



Figure 2.0 - Labor hours as a % of Average Net Daily Sales

With regard to the location impacting the sales at the restaurant, Paul was certain that the Lynden restaurant location near the Canadian border was a driver of net sales. This was a new perspective for us, so we gathered border crossing information that highlighted the number of cars that crossed the border each day, the average length of the line of cars crossing the border, and the length of time spent in the line[2].

[2] http://www.cascadegatewaydata.com/Crossing/136/2019/3?data=sum-vol%2Cavg-delay%2Cavg-qmtr

We utilized this data to test the null hypothesis that says there was no relationship between net sales and border traffic, or any of the other border metrics we found. As you can see in Figures 2.1 and 2.2 below, we found that the p-value across all of these metrics was greater than 0.05 which indicates it is not statistically significant, and we would not reject the null hypothesis.

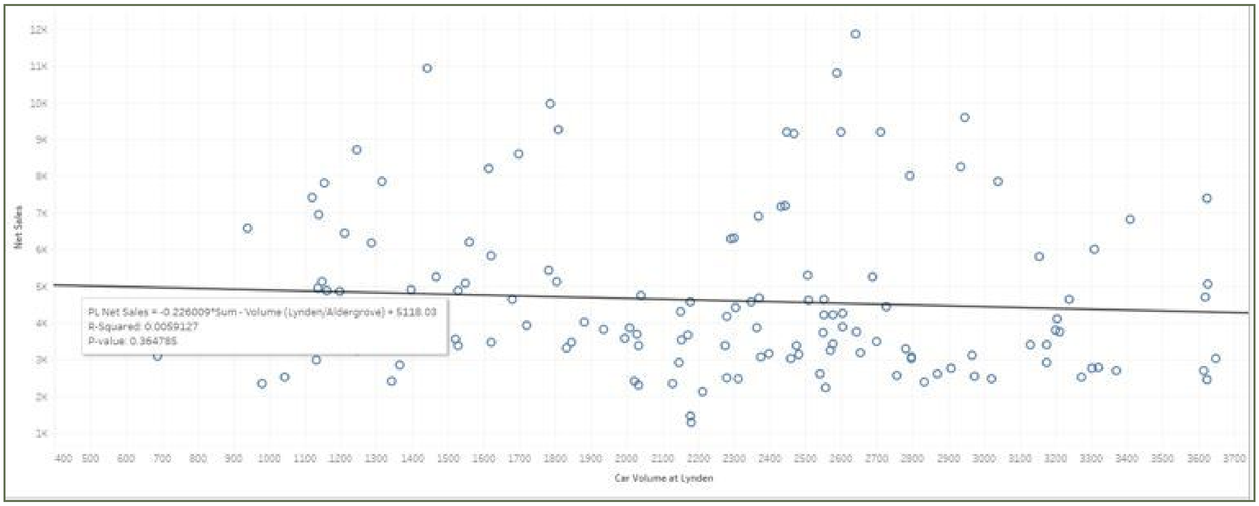



*Figure 2.1 - Net Sales vs. Car Volume through the Canadian border crossing*

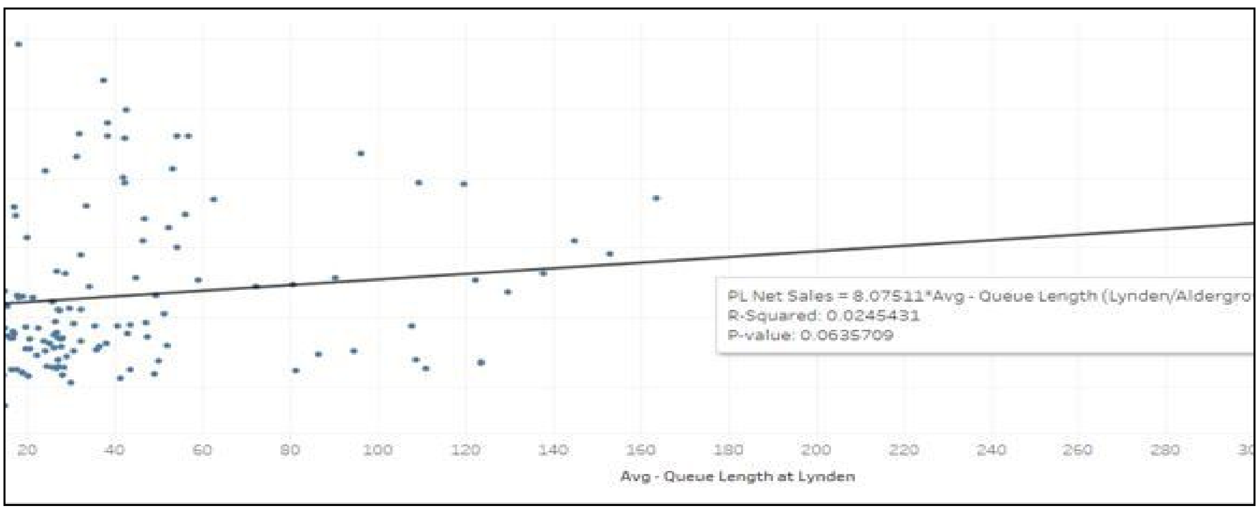



*Figure 2.2 - Net Sales vs. Average Queue Length at Lynden Border*

## *The datasets we used*

Paul provided manually updated Excel data files that included daily sales data for several of his restaurants. For Popeyes, we had data from October 2018, through February 2019 for the Lynden and Marysville locations. We were also provided Dairy Queen sales data from January 2018 through February 2019 to compare a sustainable location.

Since the data was manually stored, the team had to clean the file to make it a normalized, uniform dataset. We had to remove color coding, rename columns, and remove any special text formatting. We changed date formats to make the fields uniform. There were one-off calculated fields that needed to be removed at the bottom of the Excel sheets. We added in more detailed weather data from a website called Wunderground.com. For any missing data, we took the average for each month and filled in the missing values for each location. For data that we could not take an average calculation, we left as an empty value. There was not a significant amount of empty data points.

We also used a dataset that tracked border crossing data from the ‘Cascade Gateway Border Data Warehouse’ to test if there was a correlation between Net Sales and Border Crossing information. The data was retrieved from (http://www.cascadegatewaydata.com/Crossing/136/2019/3?data=sum-vol%2Cavg-delay%2Cavg-qmtr). Exportable as individual months, we downloaded the data for October 2018 through February 2019 as individual csv files. The data was complete without any missing values, so it did not require any cleaning, or significant transformation. We simply combined the csv files to run our analysis.

*The analytical models we used*

**Excel**

In order to address Paul’s desire to keep his daily labor costs around 25% of net sales, we created a solver model that outlines the average number of employees required for each hour the restaurant is open. Solver was the most logical choice for this model as we had a number of constraints, such as staffing per time of day, and labor hours as a percentage of daily sales, and we wanted to use a straightforward model that Paul could use moving forward.

We focused our solver model off of our sales projections for the week of February 24th, 2019 and broke them out into the three four hour shifts that Paul uses to schedule his staff, the morning shift (10 am to 2 pm), afternoon shift (2 pm to 6 pm), and evening shift (6 pm to 10 pm). We developed our model using two key constraints:

- Labor costs for each day could not exceed 25% of projected daily sales.
- Labor costs for the morning shift could not exceed 30% of projected weekly sales, while the afternoon and evening shifts could not exceed 35% of projected weekly sales, respectively.

The second constraint was in response to a comment Paul made in our second meeting with him, wherein he stated that the morning shifts were the slowest, and required less staff overall. We felt that this constraint was the most straightforward, in that it accounted for the restaurants being 5% slower in the morning than the other shifts, but wasn’t dramatically lower and could lead to skewed results.

To fine tune our model, we added additional parameters that restricted the lower limits of the values of the variable results that were being generated. This step was necessary, as our initial simulations were giving unrealistic values, like one labor hour worked for the morning shift on Tuesdays. To combat this, we added two additional constraints as daily labor costs and labor costs for each of the three shifts. For the daily labor costs, we constrained the total for each day, so it could not dip below 23.75% of daily sales. We felt that this value was realistic as a lower limit, as anything lower could result in the restaurant being understaffed. For the labor costs for each shift, we constrained the lower limit to be 90% of total weekly sales for each respective shift. We found that this 10% variance to be the most appropriate, as it gave the model flexibility, without restricting the solutions it could generate, and provided results within the 27% to 30% range for the morning shift and 31.5% and 35% for the afternoon and evening shifts. This was a level we agreed was acceptable for real life scheduling.

To validate our results we compared them to the average cost of labor as a percentage of daily sales for each day of the week from the data Paul provided. As you can see in Figure 3.0 below our optimized model lowered the range of the labor percentage from 24.03% - 34.06% to 23.83% - 24.00%.

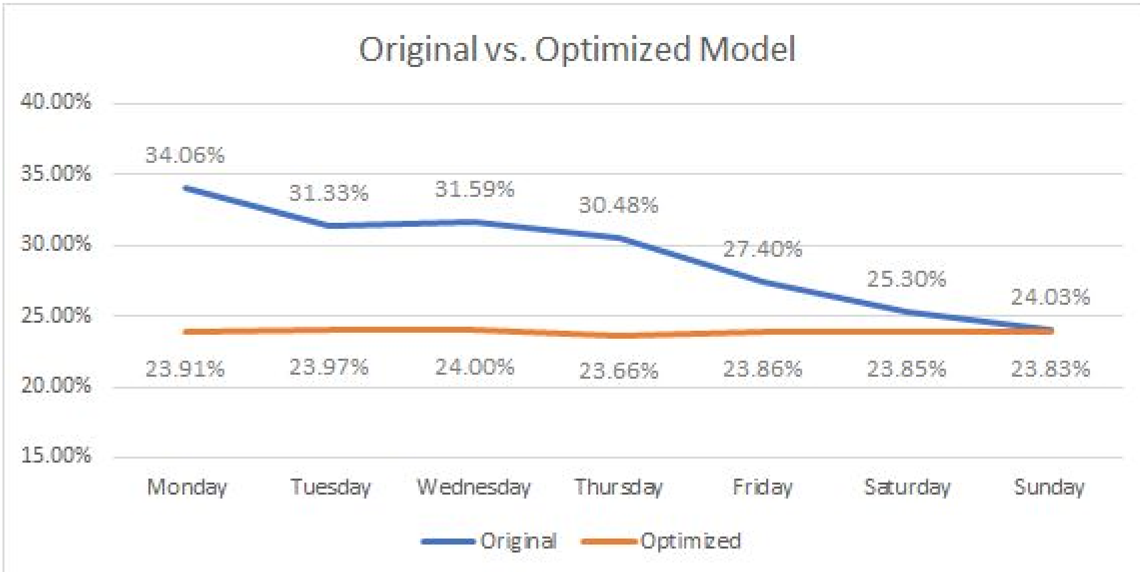



*Figure 3.0 - Average Daily Labor Costs as a % of Net Sales vs. Optimized Labor Costs as a % of Net Sales*

The validation shows that our optimized schedule brings the labor hours below Pauls desired 25% threshold and gives him a baseline to work off of when creating the schedule each week. Similarly, the weekly labor cost from our model was $5,424.00, compared to the average weekly labor cost of $8,070.00, providing Paul with a projected savings of $2,646.00 for the week and lowered total labor hours worked from 672.50 to 452. A reduction of 220.50 labor hours.

**Tableau**

Paul wanted to estimate the customer count and net sales for the upcoming week. We performed an initial exploratory analysis on the past daily net sales and customer count data for the past few months. Based on our analysis, we observed that the new restaurant experienced a sales peak upon its first two months of opening. Thereafter, the sales stabilized over the next three months. We observed that historically Mondays and Tuesdays where the slowest periods of the week, increasing toward the weekends where the peak sales would be experienced week after week. We tried several forecasting methods like simple moving average, however, this resulted in taking into account past averages without considering the seasonality. Then we tried the ARIMA (0,0,7) model, however, we could not find a strong correlation between the independent variables and dependent variables like net sales and customer count. After several attempts to find the model fitness with the available data, we found the exponential smoothing model to closely replicate the historical trends experienced.

- Smoothing is a statistical term used to describe an average of the past demand used to predict the demand for the future business. In smoothing we remove the random variation from our historical demand by the process of averaging out based on the recent history.
- Weighted moving average is used in exponential smoothing by adapting our moving average calculation and applying various weights (usually in the form of percentage) to specific past periods, which aggregate to 100%.

For example, for the past 4 months, we apply the weight of 35% to the most recent, then 25% to the second most recent and then 25% and 15% to the remaining months respectively to add up to 100%.

- The rationale behind using this model was to take daily variation, like Mon./ Tues. into account as well as associate higher weights to the most recent data. We could not completely ignore the trends observed in the weeks following the restaurants opening, however, those weeks did not portray the true picture of the sustained net sales and the customer count. Hence upon further analysis, the results of the most recent weeks provided closer and more accurate picture in predicting the net sales and the customer count.
- Another important advantage of exponential smoothing is upon addition of new period's demand, we need not recalculate against each previous period, instead we reuse the output of the exponential smoothing calculation from the previous period to represent all of the previous periods.
- The formulation for exponential smoothing was done based on the calculation as follows:
  The net sales for the most recent period multiplied by the smoothing factor.
  PLUS
  The most recent period's forecast multiplied by (one minus the smoothing factor).

Net Sales Forecast

| Day of PL Date | |
|---|---|
| March 2, 2019 | 4,176 |
| March 1, 2019 | 3,772 |
| February 28, 2019 | 3,043 |
| February 27, 2019 | 2,950 |
| February 26, 2019 | 2,703 |
| February 25, 2019 | 2,509 |
| February 24, 2019 | 3,576 |
| February 23, 2019 | 4,176 |
| February 22, 2019 | 3,772 |
| February 21, 2019 | 2,926 |
| February 20, 2019 | 2,781 |
| February 19, 2019 | 3,162 |
| February 18, 2019 | 3,088 |
| February 17, 2019 | 3,925 |
| February 16, 2019 | 3,895 |
| February 15, 2019 | 3,369 |
| February 14, 2019 | 2,296 |
| February 13, 2019 | 2,412 |
| February 12, 2019 | 1,287 |
| February 11, 2019 | 1,454 |
| February 10, 2019 | 3,388 |
| February 9, 2019 | 4,030 |
| February 8, 2019 | 3,648 |

*Figure 3.1 - Net Sales Forecast*

In the above example, we obtained the data through February 22nd, 2019. We predict the net sales for the upcoming week until March 2nd, 2019.

For February 23rd, we used the weighted average of the past week in the decreasing percentages and then summed the values to predict the net sales for February 23rd, 2019.

For February 24th, we assigned higher weightage to the predicted value of February 23rd compared to February 22nd and then weightages in the decreasing order.

Similarly, we calculated the predictions for the upcoming week through March 2nd, 2019.

Using our method, we observe similar trends for the predicted days within the week compared to the past data. The net sales value for March 23rd, 2019 is same as April 2nd, 2019.

## 5. Results and Conclusions

*Key Results*

The key results of our findings were focused on two areas, current levels of labor and which customers are driving sales for the new restaurant. Our most notable finding was disproving Paul's belief that border traffic near Lynden was a key driver of net sales. Based on our analysis, there is no notable correlation between border traffic and sales, which to us means that local customers make up more of the customer base than Paul expected. Using this information, Paul could focus future marketing efforts on the local communities that surround his restaurants in an effort to drive sales. The other finding was that current labor costs are to high and cutting into Paul's margins. A lot of the variance in labor hours can be attributed to the relative infancy of the restaurant, however it is telling that our model saw an over 10% improvement in labor costs as a percentage of net sales. This tells us that, at the very least, the labor schedule needs to be reevaluated to bring labor costs to a more sustainable level.

The input to the solver was the predictions on the weekly net sales and customer count. To help Paul with accurate estimations, we used exponential smoothing method to forecast these metrics. These metrics would help Paul with stock keeping and inventory management. Labour schedule and overall planning can also help Paul increase profitability by saving big bucks.

*Why are they important?*

From the words of Peter Drucker - "*you can't manage what you can't measure*". [3] From a business standpoint, how well or how poorly the business is performing is unknown until it is tracked or measured to some consistent manner. Without tracking key areas, like labor percentage, Paul would not know if any individual day, week, month or year was a profitable one. Granted this is just one of many KPI for his business, it is one of his controllable costs in his operations. By measuring and tracking various situations of daily business, it helps to distinguish between special cause and common caused operation variations that occur. If special causes are identified they can either be adopted as positive or eliminated as negative impacts to the business performance. From the words of Edward Deming - "*Management of a system requires knowledge of the interrelationships between all of the components within the system and of everybody that works in it*." [4]

Our other key evaluation was about the customer base of Popeyes Lynden. It is extremely critical for every business to know it customer base. Without that understanding and focus area how are you expected to provide a service? "*Profit in business comes from repeat customers, customers that boast about your product and service, and that bring*

[3] https://guavabox.com/if-you-cant-measure-it-you-cant-improve-it/

[4] https://quotes.deming.org/authors/W._Edwards_Deming?page=8

*friends with them*"[5] - Deming. We as a team hope and expect that it is the repeat customers, as we suspect and Deming has also stated, that sustain the Lyden location. However being near the border to Canada does propose an uncommon variance to his customer base. Without performing some level of a survey of his actual customers, it will be hard to have a definitive answer to Pauls prediction. Inversely relying on our data obtained only and our analysis, we found that the border traffic did not have a statistical correlation with his profit margin. This alone expresses the point of how critical it is to know your customer base and how to wow them to the point that they freely boast your service to their friends. And to Paul's point he does not compromise his labor throughout the day, because he understands the impact it can have directly on his customer base with the lack or speed of service, cleanliness, etc.

*Who can use these solutions and how?*

We feel that our analysis and solutions can be replicated for any fast food restaurant. We also believe our staffing solver model could be a great starting point for most simple businesses knowing some form of adjustments would be needed to closely replicate their business structure.

Franchise restaurants like Popeyes and Dairy Queen typically run in a similar fashion. The quick service restaurants include all the same variables like labor costs, food costs, and customer service that affect net sales. Our team focused on the variables shared by most franchises and all restaurants may find benefit in the results we have produced.

*What are the limitations of the solution and how can this work be extended in the future?*

We had two main limiting factors during the project. First, our team did not have a strong background in the fast food industry. Although we were able to come up with general ideas about attributes to analyze, we may have been limited to key factors we could have reviewed. For example, it was not until Paul had mentioned border traffic affecting sales that we started to look into this notion. The second limiting factor was access to more pertinent data. Fortunately Paul had been collecting his own data over a period of time that we were able to derive analysis from. The inability to establish a NDA between us and the Franchisor really limited the depth of the data, that we would have preferred. . This greatly limited the type of analysis we could provide to Paul.

In order to extend our work moving forward, access to all relevant data will be key, obviously, but not just the aforementioned point of sale and purchase data. We believe that even if Paul takes small steps to expand the variety of data he maintains, it can provide significant insight moving forward. If Paul could track simple data, such as transactions and sales by shift, when coupons were sent to customers, and how many orders come via food delivery apps such as GrubHub, we could better project net sales each day, provide a more thorough labor optimization schedule, and see which promotions, and which medium they are delivered through, are the most impactful. Similarly, if Paul started tracking the types of items customers are ordering, beverages, food, etc., like he does in his DQ location, further insight in the seasonality and demand of certain menu items could limit food spoilage. Ultimately, with the full data set that Popeye's keeps, the possibilities are endless to the types of analysis you one perform to drive efficiencies in this restaurant.

[5] https://quotes.deming.org/authors/W._Edwards_Deming?page=9

## 6. References

"Abbotsford, Canada History." *Weather Underground*, www.wunderground.com/history/daily/us/wa/lynden/CYXX/date/2019-4-28?cm_ven=localwx_history.

"Lynden/Aldergrove Border Crossing 2018 - 2019." *Cascade Gateway Border Data Warehouse,* *http://www.cascadegatewaydata.com/Crossing/136/2019?data=sum-vol*

Piasecki, Dave. "Exponential Smoothing Explained." *Exponential Smoothing Explained*, www.inventoryops.com/articles/exponential_smoothing.htm.

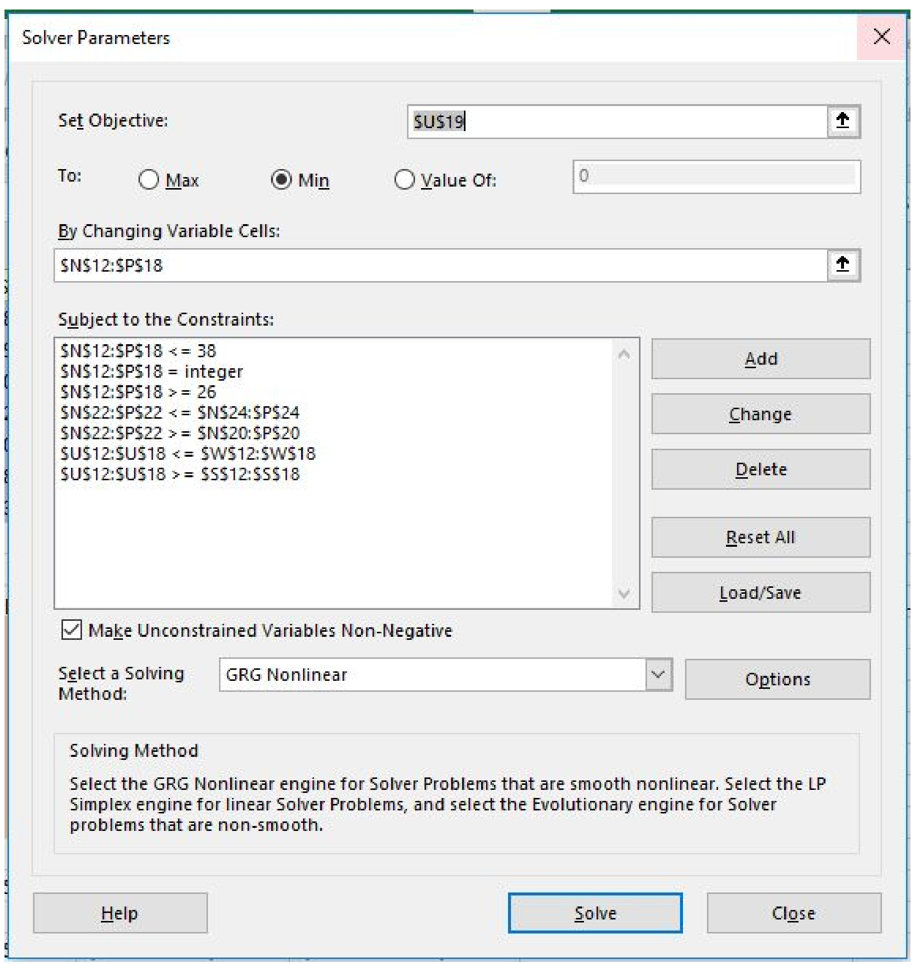
## 7. Appendix – Reproduction of Results

All of our data and process steps have been stored on our teamFile Exchange. Outlined below are the overviews from our meetings, data sets, and steps taken to reproduce our results:

1. The notes from our initial meeting and follow up questions for Paul are saved as 'Meeting Minutes - Meeting w/ Paul 3.2.19' and 'Follow up Questions to Ask Paul', respectively.
2. The original datasets we received from Paul are saved as follows:
   a. Lynden Popeyes Location
      i. Lynden - Original Sales - October
      ii. Lynden - Original Sales - Nov & Dec
      iii. Lynden - Original Data - Jan & Feb
   b. Marysville Popeyes Location
      i. Marysville - 2018
      ii. Marysville - 2019
   c. Marysville Dairy Queen Location
      i. Marysville DQ - 2018
      ii. Marysville DQ - 2019
3. The cleaned datasets are saved as 'Marysville Popeyes Data', 'Marysville DQ Data', and 'Lynden Popeyes Data' respectively. The data we cleaned and used for the basis of our analysis can be found in the 'Lynden Cleaned Data' file. The description of the steps we took to clean the data can be found in the 'Steps to Clean Data' file.
4. Our Tableau analysis, as well as a brief description of the analysis completed in each file is as follows:
   a. *Labor Analysis - Interim Presentation Analysis* - A comparison of labor costs as a percentage of net sales for all three restaurants. The results of this analysis are captured in Figure 1.2 above.
   *b. Analysis of data on canadian border* - An overall analysis of the border crossing data and its impact on net sales. The results of this analysis are captured in Figure 2.2 above.
   c. *Net Sales vs Avg Delay* - Net Sales compared to the average delay at the border crossing for each day
   *d. Net Sales vs Car volume* - Net Sales compared to the volume of cars that crossed the border that day. The results of this analysis are captured in Figure 2.1 above.

   e. *Net Sales vs Drive Thru* - Net Dine-in Sales compared to Drive Thru sales. The results of this analysis are captured in Figure 1.1 above.
   f. *Final Project* - The sales projections for the upcoming week. The projection outputs are captured in Figure 3.1 above.

5. After meeting with Paul we focused on developing the solver model to optimize labor hours. We have saved the solver model as ‘Team 1 - Lynden Labor Analysis’ in the File Exchange. The initial results and optimized results of this analysis can be found in Figures 2.0 and 3.0, respectively.
   a. There are two tabs in the file:
      i. Historic - where we used average daily sales since the restaurant opened for the optimized schedule
      ii. Projected - where we used our sales projections to optimize the schedule
   b. Both tabs isolate the relevant information from the more robust Lynden data we received in Columns A through J.
   c. In Columns L through AA is where we have developed our model and performed any supporting calculations.
   d. In cells N20:P24 are the constraints related to the distribution of labor hours.
   e. In cells S12:W18 are the constraints related to labor hours.
   f. The constraints and parameters of the solver model are as follows:



6. Also following the meeting with Paul, we sourced border crossing data from the ‘Cascade Gateway Border Data Warehouse’ to see if there was a correlation between Net Sales and Border Crossing information. The data was retrieved from (http://www.cascadegatewaydata.com/Crossing/136/2019/3?data=sum-vol%2Cavg-delay%2Cavg-qmtr), and is saved as ‘Border Crossing Data - Combined’ in the File Exchange.
    1) We analyzed three metrics from the cross border data against the daily net sales.
    2) The Tableau file-Net Sales vs Car Volume contains analysis for each metric with the net sales.
    3) This file has been shared on the file exchange and include the below analysis.
    4) Each weekday has an average net sales plotted against average vehicles in queue, drive thru
        1) Net Sales vs Avg Delay
        2) Net Sales vs Avg vehicles in Queue
        3) Net Sales vs Car Volume
    5) The relationship of each of these metrics with net sales is non-linear over a given week.
    6) This relationship helped us prove that there is no correlation between the sales at Paul’s restaurant and the Canadian border traffic.
7. One outcome from the follow up meeting with Paul also indicated the need for forecasting the customer count and net sales for better stock keeping, inventory planning and labour schedule.
   We used the past net sales and customer count data from the daily sales sheets after the files were cleaned for calculating the future projections.
    1) We tried different methods to best predict these metrics.
    2) However, exponential smoothing method helped us predict these metrics really well since the data had seasonality.
    3) The file name Final Projects is a Tableau file which shows the calculated predictions for these metrics.
    4) This file has been uploaded on the file exchange.
    5) The weighted average have been applied to the past data with the most recent from the past data having more comparatively weightage i.e. percentage.
    6) Since, we have weekly data of the past 5 months, we are able to derive weekly predictions for net sales and customer count.
    7) With time, Paul will be able to derive monthly predictions.