Stats 195 Report

Cydney Vicentina

3/13/2020

During January of this year, I began an internship as a Brand Manager for the Chick-Fil-A in Westwood. The internship consisted of brainstorming creative marketing executions, bringing them to life, and then finally, doing some statistical analysis work.

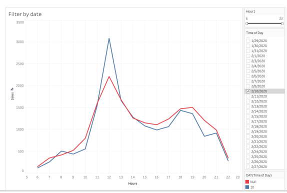
My supervisor, Eric Kim, had two questions that he hoped statistics could answer. 1. How are the marketing initiatives doing? Are they improving sales in some significant way? 2. Every 7th customer has the opportunity to fill out a customer survey. Do any of these categories (i.e. taste, temperature, ease of receiving order, etc.) correlate to sales?

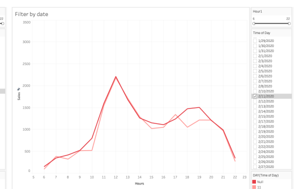
1. Efficacy of Marketing Initiatives

One of our first executions was a “Soft Reopening” for the week of Valentine’s Day. Each day, there were at least two promotions going on everyday, one in the morning/afternoon, and one at night. For example, on Tuesday, February 11th, if you went into the store from 2-5pm twinning with a friend, you could get a large shake and fries. Or from 5-10:30pm, if you were a Westwood employee, you could get a free sandwich with purchase. One nice aspect of Chick-Fil-A’s reporting page is that transaction information is easily accessible whether by day, by hour, or even by 15-minute intervals.

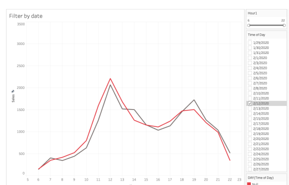
Graphs

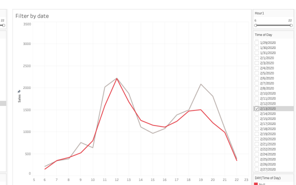
Using this information, sales data was graphed for the soft reopening week.





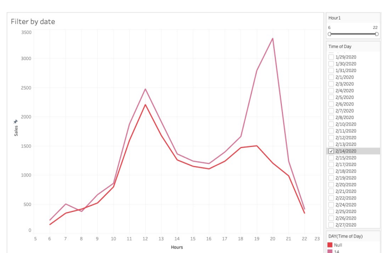
Monday Sales Tuesday Sales



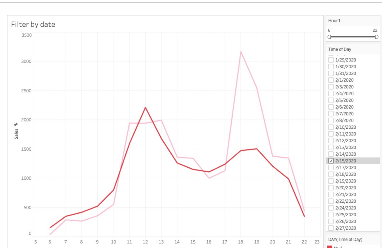


Wednesday Sales Thursday Sales

One new application that I learned for this project was Tableau, an app that easily creates graphs by using drag-and-drop technology. From the images, you can see the line graphs that I generated. The red line represents the “Null” category which is average sales data on the same weekday (i.e. Monday) across the open hours of the restaurant in January and the first week of February (usually 6:30am - 10:30pm). The other colored line indicates the sales for each hour during the soft reopening day.



Friday Sales

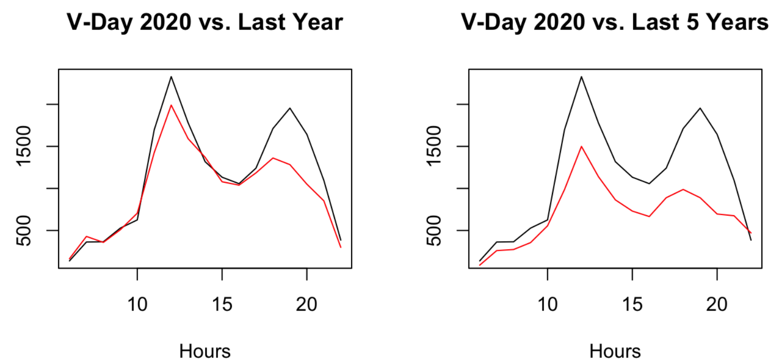


Saturday Sales

Two of the graphs stood out compared to the other days, Friday and Saturday. At a glance, our two best campaign times were Friday and Saturday evenings, both around 5pm-close. It’s unsure if these sales times were higher specifically because of our campaigns or because these were peak sales hours during a holiday weekend.

For reference, the Friday evening promotion was a free ice cream cone giveaway and the Saturday evening promotion was free nuggets in return for a Yelp review. Both promotions went on from 5-10:30pm and required a purchase.

Valentine's Day t-tests



To help answer this question, I made some graphs comparing past Valentine's Weeks sales with this year. I found the sales data for Valentine's Day from 2015 to 2020, cleaned it, and calculated some averages.

```
t.test(vday20$Sales,vday19$Sales)
```

```
##
##  Welch Two Sample t-test
##
## data:  vday20$Sales and vday19$Sales
## t = 1.7475, df = 190.31, p-value = 0.08216
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -20.24936 334.84759
## sample estimates:
## mean of x mean of y
## 1139.3902  982.0911
```

```
t.test(vdayavg20,vdayavg3)
```

```
##
##  Welch Two Sample t-test
##
## data:  vdayavg20 and vdayavg3
## t = 7.181, df = 16, p-value = 2.186e-06
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##   803.0294 1475.7510
## sample estimates:
## mean of x mean of y
##   1139.39      0.00
```

Using t-tests, you can see that Valentine's Week sales for 2020 (`vday20$Sales`) were slightly different from last year's sales (`vday19$Sales`). You could technically considerate this difference significant if you use a p-value cutoff of 0.1. However, that decision must usually be made before looking at the results and 0.05, although somewhat arbitrary, is a more common cutoff point.

Furthermore, 2020 Valentine's Week sales were significantly different from the average Valentine's Week sales from the past three years (`vdayavg3`) and the past five years (`vdayavg` , p-value = 0.02484).

Note: Valentine's Day falls on different days of the week every year. So, it is also a possibility that the combination of Valentine's Day, it being a Friday, and having a promotion led to inflated sales. It is also possible that Chick-Fil-A sales have simply been on the incline this past year.

Conclusion

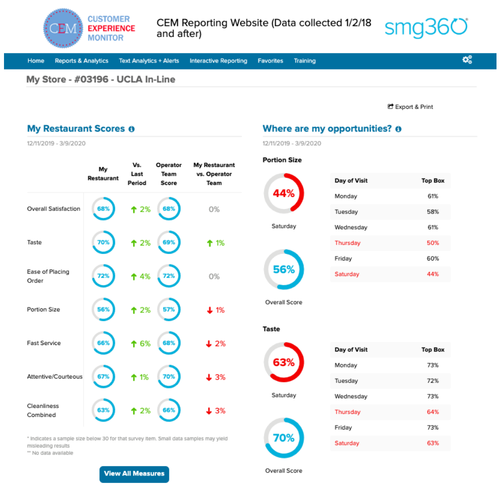
According to my supervisor, net sales were positive for the month of January (~5%) and positive for the month of February (~8%). While his goal is a 12% increase, it seems that our marketing team is on the right track.

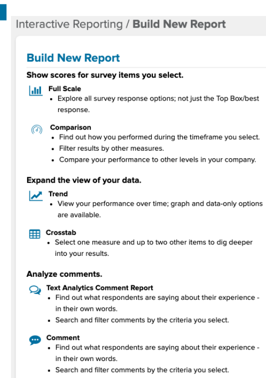
During our debrief of the soft reopening, some of our notes for future promotions include: expanding our follower base since social media was the only news outlet for the soft reopening, simpler text and giveaways because having different rules every day was tiresome to read through, and more preparation since many of us were gone for the weekend before and/or after Valentine's Day.

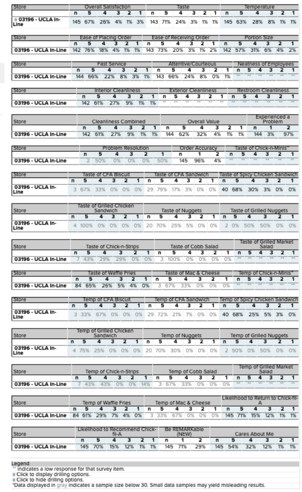
While these results can't be considered completely valid since an experiment was not conducted, this information gave my supervisor some insight into his best sales hours. It also inspired us to do more specific tracking for our next promotions (i.e. adding a "Student ID" button to the cash registers to track who comes in for our social media promotions aimed at students during Week 10). Hopefully, this will make statistical analysis more robust and clear in the future.

2. Survey Data

The second question Eric Kim asked me to look into was whether or not certain survey categories affected sales or transaction counts. Throughout the day, every 7th printed receipt receives a link at the bottom which takes the recipient to an online survey about their experience at Chick-Fil-A that day. This is mostly managed by an outside company, smg360, which compiles the data in a user-friendly “Customer Experience Monitor” website. Trying to build a new report mainly resulted in more compiled graphics as shown below.

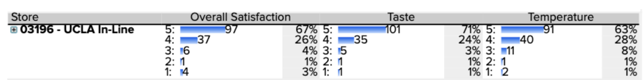






Problems with Data Collection

The first problem I ran into was collecting the data. Since this data is compiled into “helpful” charts and statistics by another company, my supervisor had no access to the raw data. The closest option to the raw data was “Full Scale” report building. However, there were more problems. Here’s an example of February’s data. The general categories in the survey include “Overall Satisfaction”, “Taste”, “Cleanliness”, etc. You are given an “n”, the number of people who took the survey which is not consistent. If you notice, “n” for “Overall Satisfaction” is 145 while the “n” for “Taste” is 143. This means that people are not obligated to fill out every section of the survey which can lead to NAs in the data. The other obvious problem is that in this “Data Only” report, you are given percentages of people who answered from 1-5.



Store	Overall Satisfaction		Taste		Temperature	
03196 - UCLA In-Line	5: 97	67%	5: 101	71%	5: 91	63%
	4: 37	26%	4: 35	24%	4: 40	28%
	3: 6	4%	3: 5	3%	3: 11	8%
	2: 1	1%	2: 1	1%	2: 1	1%
	1: 4	3%	1: 1	1%	1: 2	1%

In a different tab, you get numeric values. However, again, this data is compiled and not scrapable. Luckily, I talked to my supervisor and he was able to get the raw data with timestamps and comments, which I further cleaned.

Trying to Fit A Model

Using VLOOKUP in Excel Sheets, I was able to combine each row of survey data with the corresponding hourly sales data.

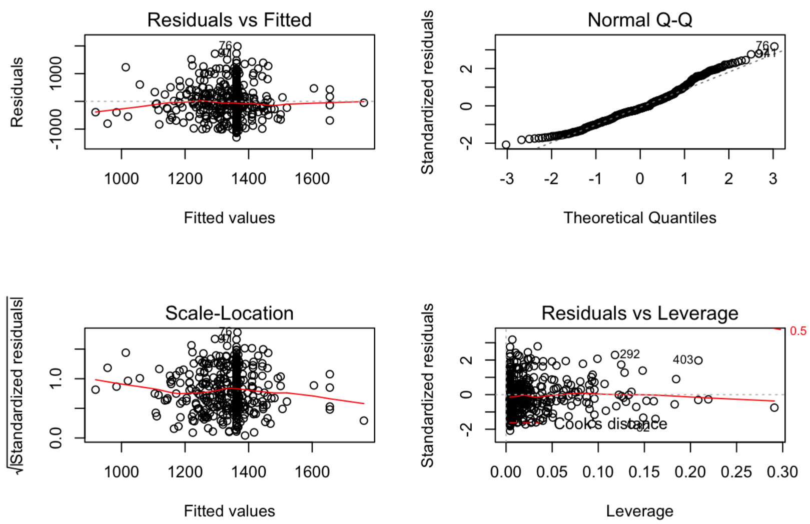
Note: I decided to remove columns like “Taste of Chicken Sandwich” or any categories that referenced individual menu items or other categories like “Problem Resolution” which consistently resulted in NAs and had no data (my guess would be that these columns indicated short answer responses which are also compiled on the front page but are not given with the numeric data).

```
lmodel3 <- lm(Sales ~., data = survey_new)
summary(lmodel3)
```

```
##
## Call:
## lm(formula = Sales ~ ., data = survey_new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1301.06  -430.01   -87.83   339.91  1984.36
##
## Coefficients:
##                        Estimate Std. Error t value Pr(>|t|)
## (Intercept)             596.293    395.478   1.508    0.132
## Attentive.and.Courteous  32.393     56.688   0.571    0.568
## Fast.Service            -81.193     63.910  -1.270    0.205
## Taste                    -6.484     52.959  -0.122    0.903
## Interior.Cleanliness     29.263     56.096   0.522    0.602
## Portion                  -3.005     42.765  -0.070    0.944
## Ease.of.Placing.Order   -18.660     65.593  -0.284    0.776
## Ease.of.Receiving.Order  90.220     79.007   1.142    0.254
## Accuracy                293.704    192.519   1.526    0.128
## Likelihood.to.Return    101.341     62.678   1.617    0.107
## Likelihood.to.Recommend -48.838     60.070  -0.813    0.417
## Experienced.a.Problem    -1.723    134.662  -0.013    0.990
##
## Residual standard error: 626 on 394 degrees of freedom
## Multiple R-squared:  0.02143,    Adjusted R-squared:  -0.005886
## F-statistic: 0.7846 on 11 and 394 DF,  p-value: 0.6556
```

To begin, I fit a simple linear model on the data. However, notice that the R-squared value is pretty small so diagnostic plots are needed.

```
par(mfrow = c(2, 2))
plot(lmodel3)
```



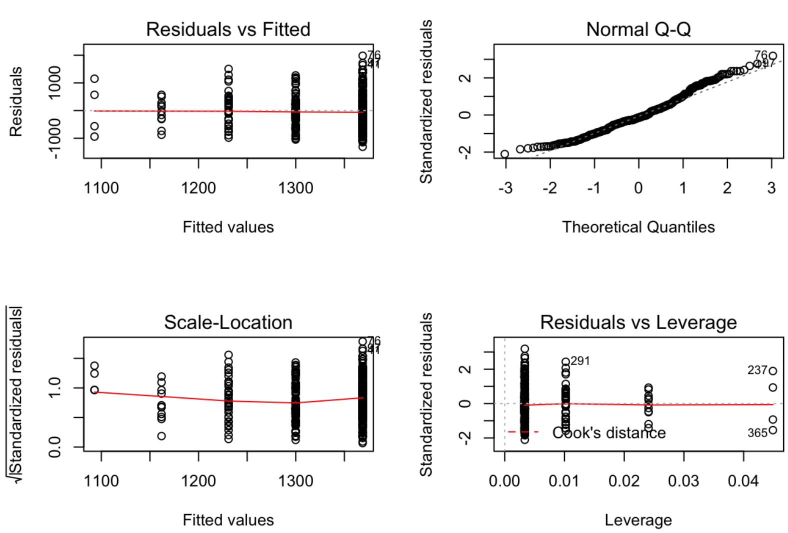
The diagnostic plots look okay so the assumptions of the model are upheld and it's possible that none of the variables are significant.

Finally, I referenced my 101C notes on all the models we learned and attempted to fit some models using other methods.

```
#CPS and BICS (1 significant variable)
summary(lm(Sales ~ `Likelihood.to.Return` , data = survey_new))
```

```
##
## Call:
## lm(formula = Sales ~ Likelihood.to.Return, data = survey_new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1308.11  -427.22   -84.45   345.33  1980.32
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)           1024.03     167.73   6.105 2.41e-09 ***
## Likelihood.to.Return    68.95      36.70   1.879    0.061 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 622.2 on 404 degrees of freedom
## Multiple R-squared:  0.008664,   Adjusted R-squared:  0.00621
## F-statistic: 3.531 on 1 and 404 DF,  p-value: 0.06096
```

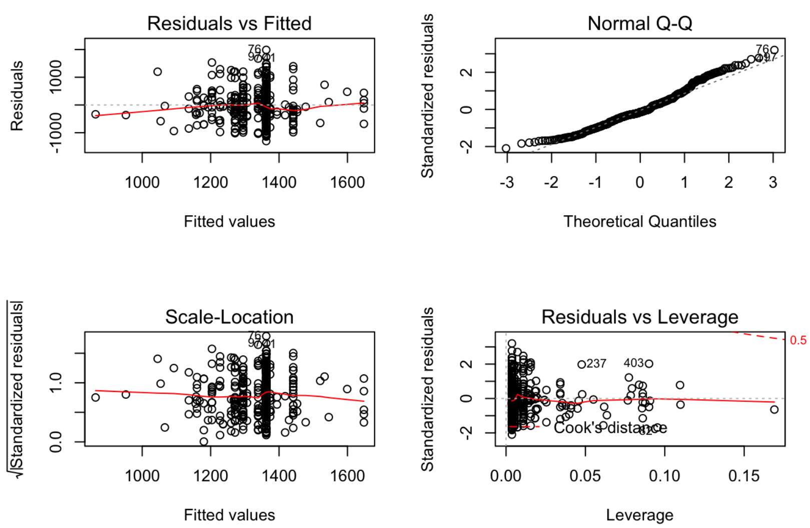
```
par(mfrow = c(2, 2))
plot(lm(Sales ~ `Likelihood.to.Return` , data = survey_new))
```



```
#AR (4 significant variables)
summary(lm(Sales ~ Fast.Service + Ease.of.Receiving.Order + Accuracy + Likelihood.to.Return , data = survey_new))
```

```
##
## Call:
## lm(formula = Sales ~ Fast.Service + Ease.of.Receiving.Order +
##     Accuracy + Likelihood.to.Return, data = survey_new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1301.92  -413.06   -83.99   349.88  1986.51
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)               621.09     326.01   1.905   0.0575 .
## Fast.Service              -78.88      61.34  -1.286   0.1992
## Ease.of.Receiving.Order   102.59      63.14   1.625   0.1050
## Accuracy                  285.40     180.29   1.583   0.1142
## Likelihood.to.Return       67.51      40.49   1.667   0.0962 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 621.4 on 401 degrees of freedom
## Multiple R-squared:  0.0186, Adjusted R-squared:  0.00881
## F-statistic:   1.9 on 4 and 401 DF,  p-value: 0.1096
```

```
par(mfrow = c(2, 2))
plot(lm(Sales ~ Fast.Service + Ease.of.Receiving.Order + Accuracy + Likelihood.to.Return , data = survey_new))
```



Using regsubsets had poor results again. I used the “exhaustive” approach because neither forward or backward did any better. As you can see, the R-squared values are still small. With the diagnostic plots, you can see that they are valid models and still, none of the variables are significant.

Here, I tried using LASSO for variable selection.

```
## 12 x 1 sparse Matrix of class "dgCMatrix"
##                                    1
## (Intercept)             1329.7277975
## Attentive.and.Courteous    .
## Fast.Service               .
## Taste                      .
## Interior.Cleanliness       .
## Portion                    .
## Ease.of.Placing.Order      .
## Ease.of.Receiving.Order    .
## Accuracy                   .
## Likelihood.to.Return       0.9068663
## Likelihood.to.Recommend    .
## Experienced.a.Problem      .
```

It was interesting that both Cps, Bics, and Lasso all narrowed the important variables down to “Likelihood to Return” even though when inputting this information into a model, it becomes insignificant.

For ridge regression, I took out variables that had coefficients that would round to 0. This still produced non-interesting results.

```
## 12 x 1 sparse Matrix of class "dgCMatrix"
##                                  1
## (Intercept)            1270.2957347
## Attentive.and.Courteous   1.2792566
## Fast.Service              0.3827141
## Taste                     1.1999420
## Interior.Cleanliness      1.7393029
## Portion                   0.9556624
## Ease.of.Placing.Order     1.2971722
## Ease.of.Receiving.Order   1.6648241
## Accuracy                  5.6747304
## Likelihood.to.Return      2.2953104
## Likelihood.to.Recommend   1.0004616
## Experienced.a.Problem     2.0802520
```

```
##
## Call:
## lm(formula = Sales ~ . - Fast.Service, data = survey_new)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1307.50  -441.11   -88.15   342.13  1972.90
##
## Coefficients:
##                         Estimate Std. Error t value Pr(>|t|)
## (Intercept)              596.128    395.786   1.506    0.133
## Attentive.and.Courteous   19.645     55.836   0.352    0.725
## Taste                     -1.380     52.847  -0.026    0.979
## Interior.Cleanliness      25.868     56.076   0.461    0.645
## Portion                   -8.034     42.615  -0.189    0.851
## Ease.of.Placing.Order    -20.375     65.630  -0.310    0.756
## Ease.of.Receiving.Order   42.631     69.616   0.612    0.541
## Accuracy                 257.024    190.489   1.349    0.178
## Likelihood.to.Return     101.447     62.727   1.617    0.107
## Likelihood.to.Recommend  -54.926     59.925  -0.917    0.360
## Experienced.a.Problem     -4.678    134.747  -0.035    0.972
##
## Residual standard error: 626.4 on 395 degrees of freedom
## Multiple R-squared:  0.01743,    Adjusted R-squared:  -0.007449
## F-statistic: 0.7005 on 10 and 395 DF,  p-value: 0.7241
```

Again, notice that none of these variables are significant and the R-squared value is still small even though the diagnostic plots for this model were also good.

Conclusion

The fact that this survey data is managed by another company made data analysis difficult. Since the target audience for the survey company is business owners, the focus is heavy on graphs, easy-to-read charts and percentages that are clearly color-coded and nicely presented. While nice for the average manager, this made data collection, cleaning, and analysis tiresome. I did ask my supervisor to reach out to the company to obtain the raw data which they provided. However, they were unable to offer names or specific transactions which might have made the data more accurate (since a specific sale amount and time could have been analyzed instead of generalizing to an entire hour of sales).

I would further argue that connecting survey data to sales in this context is somewhat silly. Since customers are given the survey after they've ordered and paid, this means that the sale has already been made. Pouring resources into 'Taste', for example, won't necessarily translate into more sales. The participant has already gone into the store and made a purchase to evaluate 'Taste'. To further this point, imagine a customer scored 'Taste' a 5 but scored 'Overall Satisfaction' a 1 and never returned to the store. In theory, increasing the taste of the food won't bring the customer back. Therefore, this data structure is not ideal for predicting sales (plus, none were significant anyway) and can only really be used to gauge a customer's feedback after their experience in the store. This helps explain why none of the variables ended up being significant.

Discussion

Overall, I had a lot of fun working on this project. I really enjoy my internship as part of the marketing team at Chick-Fil-A. Getting to see and work on some of their data and hopefully help out my supervisor by giving him some insights was really rewarding!

This project also taught me how tough real-life statistics work really is. From data collection to cleaning to analysis, there were often roadblocks at each step. Plus, I really felt the necessity of constantly practicing coding or fitting models because it can be easy to forget or to lose a certain skill.