Blue Bell Ice-cream Recall Analysis

Rushan Shakya & Tahrima Mustafa

In Partial fulfillment of the requirements for the course project of

ISQS 6637- Scripting Languages

Supervisor

Prof. Dr. Jaeki Song

Rawls College of Business

October 10, 2017

# Executive Summary

Blue Bell is a very popular ice cream brand in the US. In 2015, the company’s ice creams faced a series of recalls due to Listeria concerns leading to the death of 3 people across the US. The company had to shut down its factories in Texas and Oklahoma to clean, repair, and replace machineries to stop the Listeria outbreak. We want to look at the sales volume of ice creams and other frozen products at the Ice cream section of the grocery chain around Texas Tech University due to this ice cream recall.

Blue Bell recall started on March 12, 2015. The company then recalled all its products from the market on April 20, 2015 and then ceased its operations for 5 months. The company resumed its ice cream production on August 31, 2015.

In this project, we have studied consumer behavior around Texas Tech University. The main objective of this project is to analyze if people around this area consumed other Ice creams or if people just stopped eating ice creams due to Blue Bell’s Listeria recalls. The problem statement we have worked on revolves around this incident. We want to know if this recall had a significant impact on ice-cream sales.

For our study, we have used grocery chain data from 3 stores that are within the 5-mile radius from Texas Tech University. We collected the data from September 2014 till August 2016.

We have broken down our analysis into 3 stages, namely: the pre-recall stage, the recall stage, and the post recall stage. We have looked at the over-all ice cream sales and the Blue Bell ice cream sales during these three stages for our analysis. Before starting our analysis, we expected to find Blue Bell sales to be dropping after the incident. We also assumed that many of Blue Bell’s customers must have shifted to other ice cream brands after the Listeria incident. We also expected people’s ice cream buying behavior to be lower during the recall. Similarly, we expect the buying behavior to come back to normal after certain interval of time.

With all these goals in mind we have done several graphical interpretations on our data. We have used scatter plot, mosaic plot, box plot, pie chart, line graphs and bubble charts for our analysis. We will discuss about those in the coming sections.

During the analysis phase of our work, we have come to the realization that most of our assumptions were true on the Blue Bell product sales. For example, Blue Bell sales dropped drastically when they returned to market after almost 5 months of recall period. Also after Blue Bell returned to market, some products’ sales continued to go higher even though they had poor sales before the Blue Bell listeria incident. However, we did not find any change in the ice cream buying behavior of people after the deadly listeria incident. Ice cream selling rate continued to grow over the whole period.

After the analysis, we have reached to the conclusion that the listeria incident has a significant effect on the Blue Bell brand. While they were out of market for the recall process, other brands became successful in attracting consumers. After Blue Bell came back to market, they lost almost half of their market shares to other brands.

# Table of Contents

[Executive Summary i](#_Toc495411234)

[1. Introduction 1](#_Toc495411236)

[1.1. Goals 2](#_Toc495411237)

[1.2. Assumptions and Limitations 2](#_Toc495411238)

[1.3. Sources of Data 3](#_Toc495411239)

[2. Data Cleansing Process 3](#_Toc495411240)

[3. Graphical Analysis for the Recall 6](#_Toc495411241)

[3.1. Analysis Based on Recall Periods 9](#_Toc495411242)

[1.1.1. Pre-recall Period 9](#_Toc495411243)

[1.1.2. Recall Period 10](#_Toc495411244)

[3.2. Bubble Chart 12](#_Toc495411245)

[4. Conclusion 14](#_Toc495411246)

[4.1. Next Steps 14](#_Toc495411247)

[References 15](#_Toc495411248)

[Appendix 16](#_Toc495411249)

# Introduction

Blue Bell is one of the most famous ice cream brands in the US. The company sells ice creams and other ice cream novelties (like ice cream bars, popsicles, etc.) under its Blue Bell Brand. It was founded in 1907 in Brenham, Texas. Blue Bell is the most recognized ice cream brand in the US south and the third largest ice cream maker inside the US. According to Forbes (2015), the Blue Bell brand was valued at $1.66 billion dollars.

In 2015, Blue Bell had to recall its ice creams due to the listeriosis concerns which is thought to have started in its Broken Arrow, Oklahoma factory. It was reported that a total of 10 people were infected, out of which 3 people lost their lives. Listeriosis was reported from 4 states: Texas, Oklahoma, Kansas, and Arizona. (CDC, 2015)

The listeriosis outbreak was first linked to its Texas and Oklahoma facilities. Then, after a multistate inspection, the company voluntarily recalled all of its ice-cream products. Blue Bell recalls started on March 12, 2015. The company then recalled all its products from the market on April 20, 2015 and then ceased its operations for 4 months. The company resumed its ice cream production on August 31, 2015. (Ramzy, 2015)

In this report, we are going to analyze the effect of listeria incident in the sales of Blue Bell products over the 24 months period. We will also analyze the sales of other ice cream brands competing with Blue Bell in the same period.

## Goals

For this project, it was interesting to analyze, how listeria outbreak affected Blue Bell. It was once considered to take up to about 20% market share of the ice cream market and huge loyal customer base. The problem statement for our study revolves around finding out the following:

1. What happened to Blue Bell sales after recall?
2. Did other ice cream brands take Blue Bell’s market share?
3. Did Blue Bell recall affect the sales of other ice creams?

## Assumptions and Limitations

With human deaths reported, it was a lot of negative publicity for the Blue Bell brand. So, we expected to see the post-recall sales to be highly lower than what it was before the recall. Also, we expected the Blue Bell brand to lose most of it market share to its competitors like the Blue-Ribbon ice creams and others.

We expect our data to have some limitations to what happened in terms of ice-cream sales. Our data has been aggregated to months and then to just two brands (i.e. we have divided the brands to Blue Bell and others), we might have lost important insights into how Blue Bell’s competitors gained on customer during the recall period and how they continued to hold on market share after the recall. Also, the post recall period for our study seemed inadequate, and at least a year worth of post-recall data would have made our analysis much more accurate.

## Sources of Data

For our analysis, we used data from three grocery chains inside Lubbock. We selected three stores within the 5-mile radius from the Rawls College of Business as we wanted to understand the ice cream sales’ dynamics at Lubbock during the recall.

We collected the daily transaction data from three of these stores from September 2014 till August 2016 for our analysis. We had about 1.5 million ice cream transaction records spread across the entire two years period for our analysis.

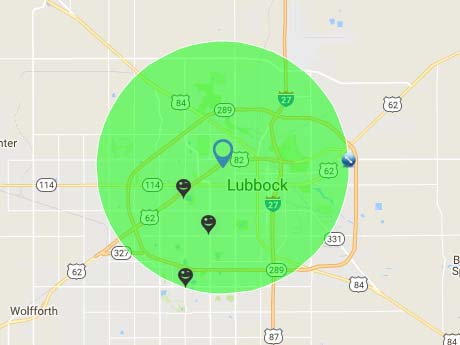


Figure 1: Three grocery chains selected for analysis

# Data Cleansing Process

For our data cleansing process, we mostly used R. We used R-packages like dplyr, and sqldf to clean our initial raw file. Our raw data had about 1,448,437 ice cream transaction observations with 14 variables.

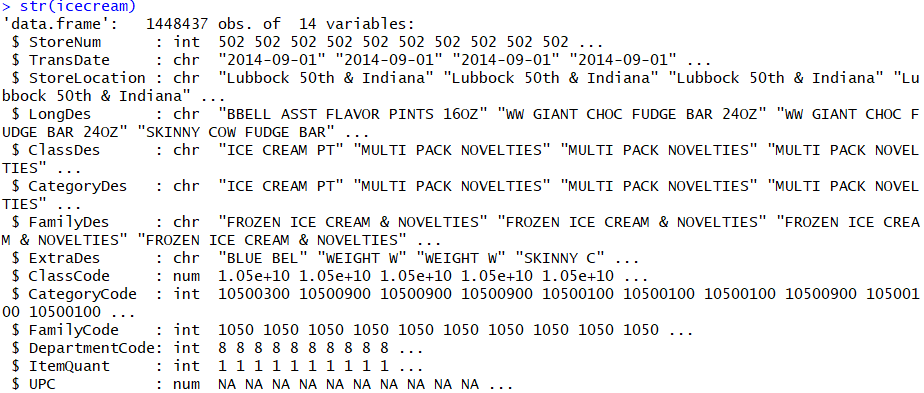
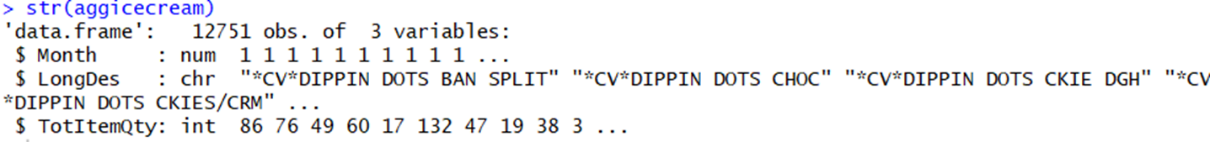


Figure 1 R-output to show the structure for icecream.csv file

Once we had our raw data imported into R, we then added the Month variable to our initial data file. We added the Month variable based on the transaction date of each ice cream sale transaction. The following logic was used to determine the months for each of the transaction dates in our data set:

*Figure 2 R-output to show the structure for the aggregated sales quantity grouped by Month and LongDes variables*

After the months were added to our original data set, we then aggregated the total sales quantity of the data set according to the Months and the Name/Long Description of our transactions. So, our aggregated data set looked like the following:

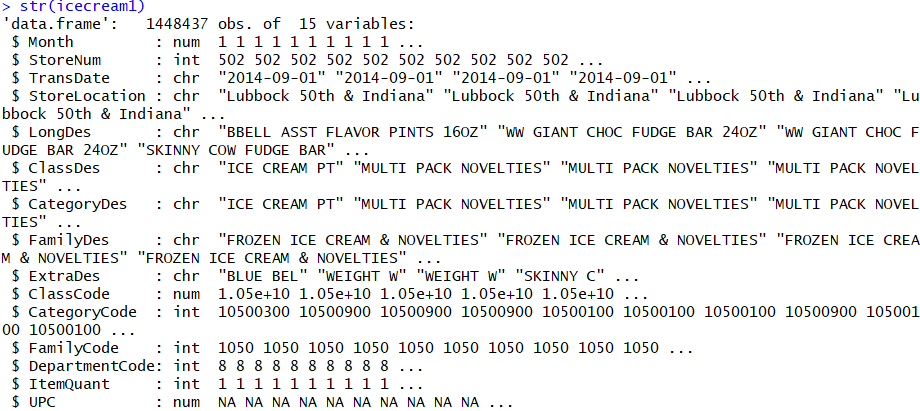


Figure 3 R-output for the structure of the data set after adding Month variable

After, the data was aggregated according to the Sales Quantity, we further aggregated our data according to the Brands (i.e. Blue Bell and Other). The structure of our data set after we further aggregated it is shown below:

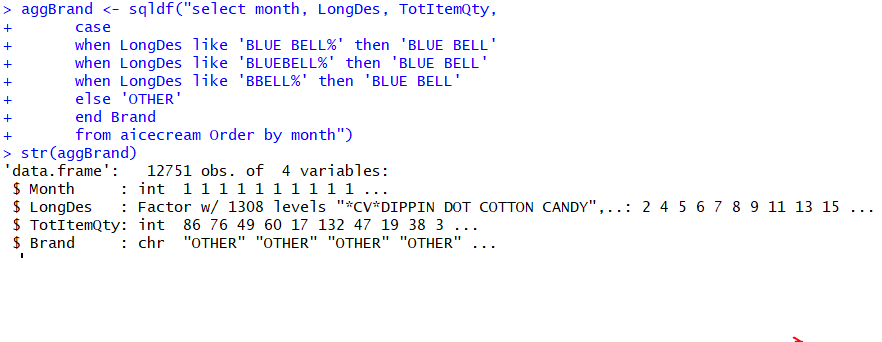
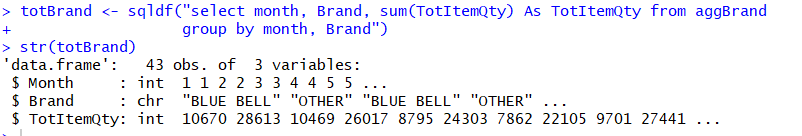


Figure 4. R-output to show the structure for the aggregated sales quantity with added Brand variable

Figure 5. R-output to show the structure for the aggregated sales quantity grouped by Brand and Month variables

Last, we were done with the data aggregation, we further divided our data set into three different time periods:

Table 1. Dividing the data set into 3-time periods

|  |  |
| --- | --- |
| **Period Description** | **Period Range** |
| Pre-recall Period | Month 1 to Month 7 |
| Recall Period | Month 8 to Month 15 |
| Post-recall period | Month 16 to Month 24 |

# Graphical Analysis for the Recall

In this part, we are going to show graphical interpretations of Blue Bell and all other brands in 3 stores over 2-year period. We analyzed our data using scatter plot, mosaic plot, box plot, pie chart and bubble charts using R. Also, we did our line graphs using Python.

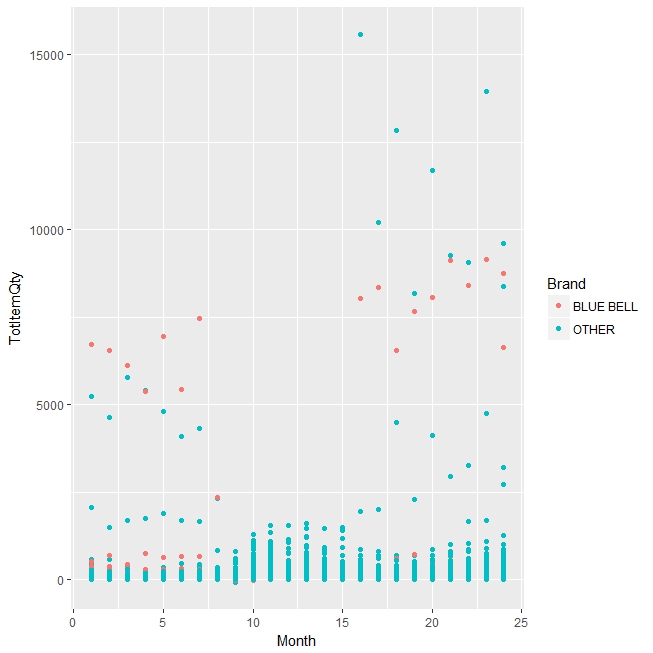


Figure 6. Scatter plot to graphically represent sales data for all the brands and Blue Bell

For our analysis, we first, used a scatter plot to show the sales of ice-cream products. Here we have plotted the months in the X-axis and total item quantity (TotItemQty) in the Y-axis. In the scatter plot, the red dots are individual Blue Bell products and the blue dots are other products. Here we can see Blue Bell has a high selling rate in the pre-recall period compared to other brands. Then from 9th to 15th month we can see no Blue Bell products were sold. That is the recall period when Blue Bell was out of market.

One interesting thing we observed from the scatter plot was that, in the recall period, except for Blue Bell, there were no other brands having sales quantity greater than 2500. This may be due to the outbreak of death news from listeria and people were scared to eat ice-cream overall.

Then in the post recall period we saw Blue Bell releasing fewer products than usual and getting higher sales compared to others.

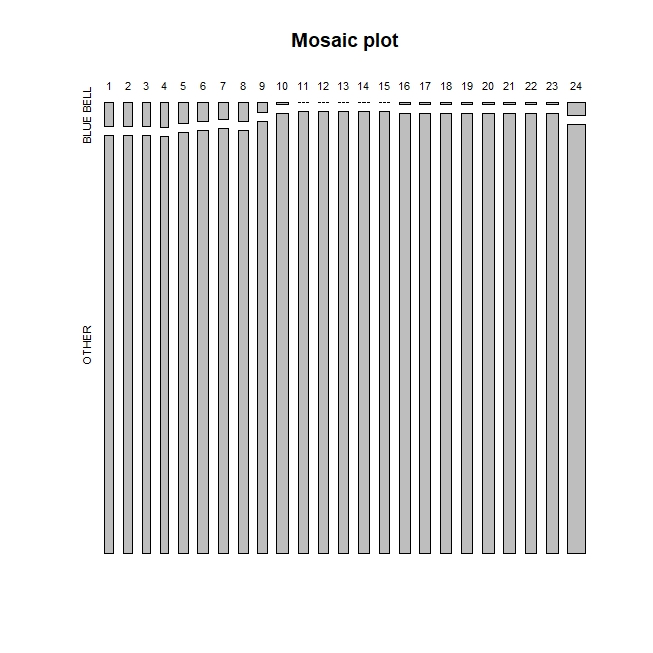


Figure 7. Mosaic plot to graphically represent sales data for all the brands and Blue Bell

Next, we show the aggregated item quantity for Blue Bell and all the other brands for the 3 stores over the 24 months period in a mosaic plot. Blue Bell products dominated the market in the pre-recall period in terms of total display items at the stores and then suddenly going out from the market during the recall period. After coming back, their product quantity was fewer compared to the pre-recall period. Though we can see that the manufacturer is gradually increasing the number of Blue Bell Products after the recall.

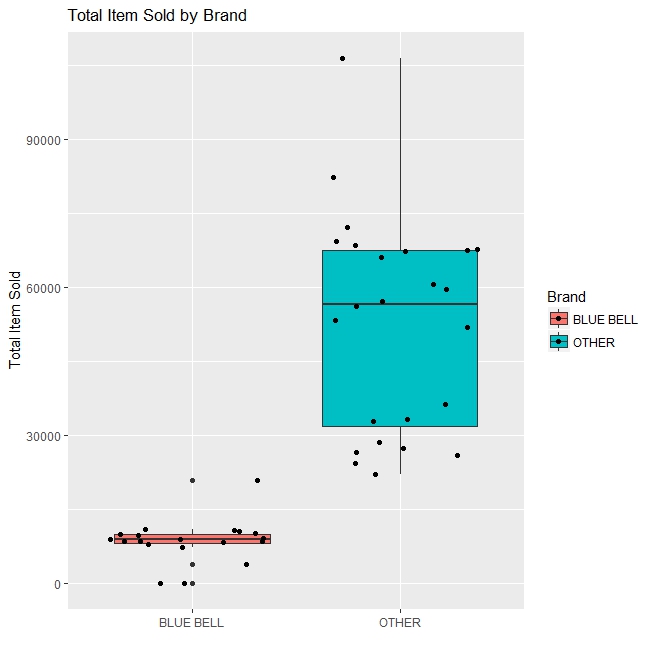


Figure 8. Box plot to graphically represent sales data for all the brands and Blue Bell

Next, we used boxplot. Box plots are used to find possible outliers in the data. It also shows skewness of the data. From this plot, we have observed that Blue Bell has consistent sales over the time since their maximum, minimum and median values are very close to each other. On the contrary, other brands have a lot of variations in their data, which is predictable since we have plotted all the brands’ sales quantity together.

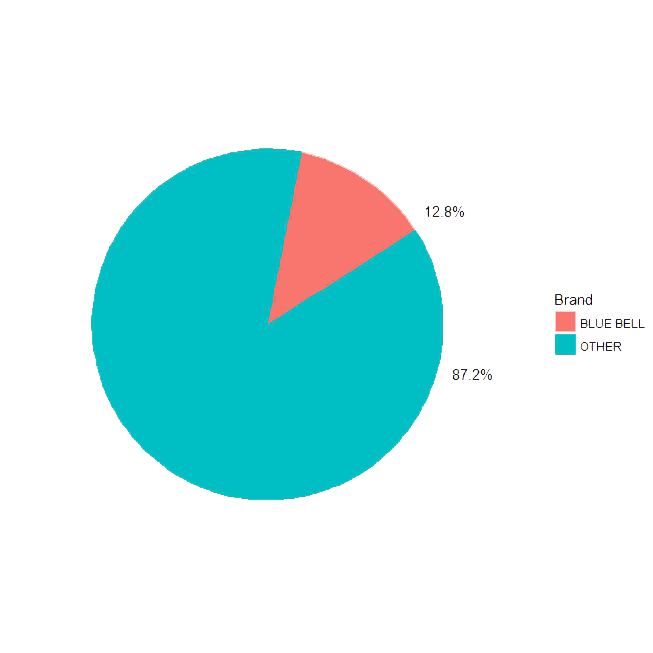
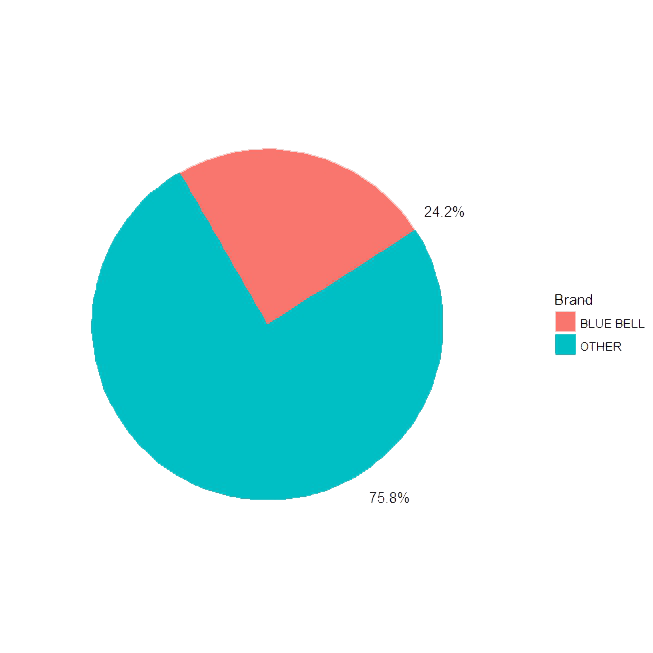


Figure 9. Blue Bell market share pre-recall (left) and post recall (right) period

Here we have plotted Blue Bell market share with the shares of all other brands in a pie chart. The left chart shows that Blue Bell had almost 24% market shares among all other brands in the pre-recall period. But their shares decreased to 12.8% after the recall.

## 3.1. Analysis Based on Recall Periods

### Pre-recall Period

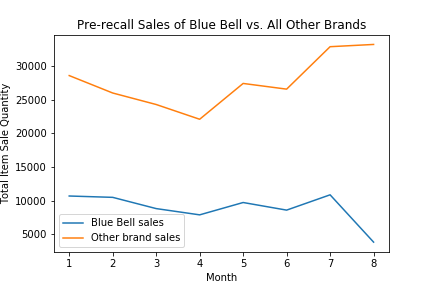


Figure 10. Line graph showing sales quantity of Blue Bell and other brands in the pre-recall period.

For the graphical analysis on pre-recall period, we have first divided the existing preprocessed data into 3 parts: Pre-recall, Recall & Post recall period. This work was done using Python.

First, we have analyzed the data for pre-recall period and plotted line graphs using Pyplot of Matplotlib library.

From the line graph we can deduce that, Blue Bell had a huge number of share in the market comparing all other ice-cream brands. But after 7th month of the pre-recall period, a downward spike is visible for Blue Bell while the other brands’ sales are spiked up. Around this period, news spread about the listeria incident and we can guess that people started to avoid Blue Bell products.

### Recall Period

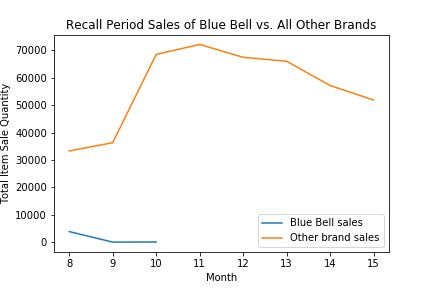


Figure 11. Line graph showing sales quantity of Blue Bell and other brands in the recall period.

As we have discussed before, in the recall period, all the Blue Bell products were called back from the market. We can see this from the line graph that the sales of Blue Bell went to negative first and then vanished completely from the graph. For the next few months, Blue Bell stopped selling their products in the market. In the meantime, other products started having higher sales than usual. From this visualization, we can deduce that people did not stop eating ice-cream completely after the listeria incident. Rather they just switched to different brands.

### Post-recall Period

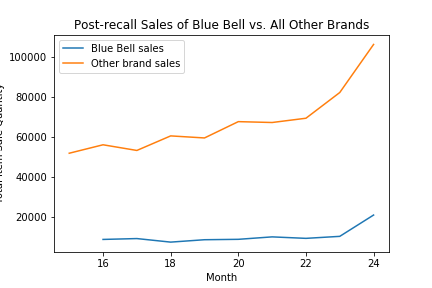


Figure 12. Line graph showing sales quantity of Blue Bell and other brands in the post-recall period.

Last, we have analyzed the post-recall period. After 5 months of not being in the market, Blue Bell returned and started selling their products. From the line graph, we can say that they have a consistent rate of sale which is not as high as pre-recall period. This can be due to the fact that they were releasing less products in the beginning of the post-recall phase. Though during the 23rd month, they had a spike on their sale that people were buying their products and they started to bring more items in the stores.

Also, there can be 2 possibilities of Blue Bell’s low sale in the post recall period: while they were out of market, some other brands came and grab Blue Bell’s consumers. Another reason might be the fact that people were trusting other brands because they are too scared to eat Blue Bell products again. Though after the listeria event, Blue Bell is still getting some sales proves that they have some loyal customers who remained with Blue Bell in the post recall period.

## Bubble Chart

The last part of our graphical analysis is done with Bubble chart using GoogleVis library in R. Bubble charts are very useful tool for data visualization because these charts can represent more variables in one place at a time. In all our bubble charts, we have plotted the months in X-axis and total sale in Y-axis. Also, another new feature for these charts is that the size of the circles (or bubbles) are showing the sales quantity. The bigger the bubble, the higher the sales for that item. Also, when we hover mouse on a circle, it shows the details of each product.

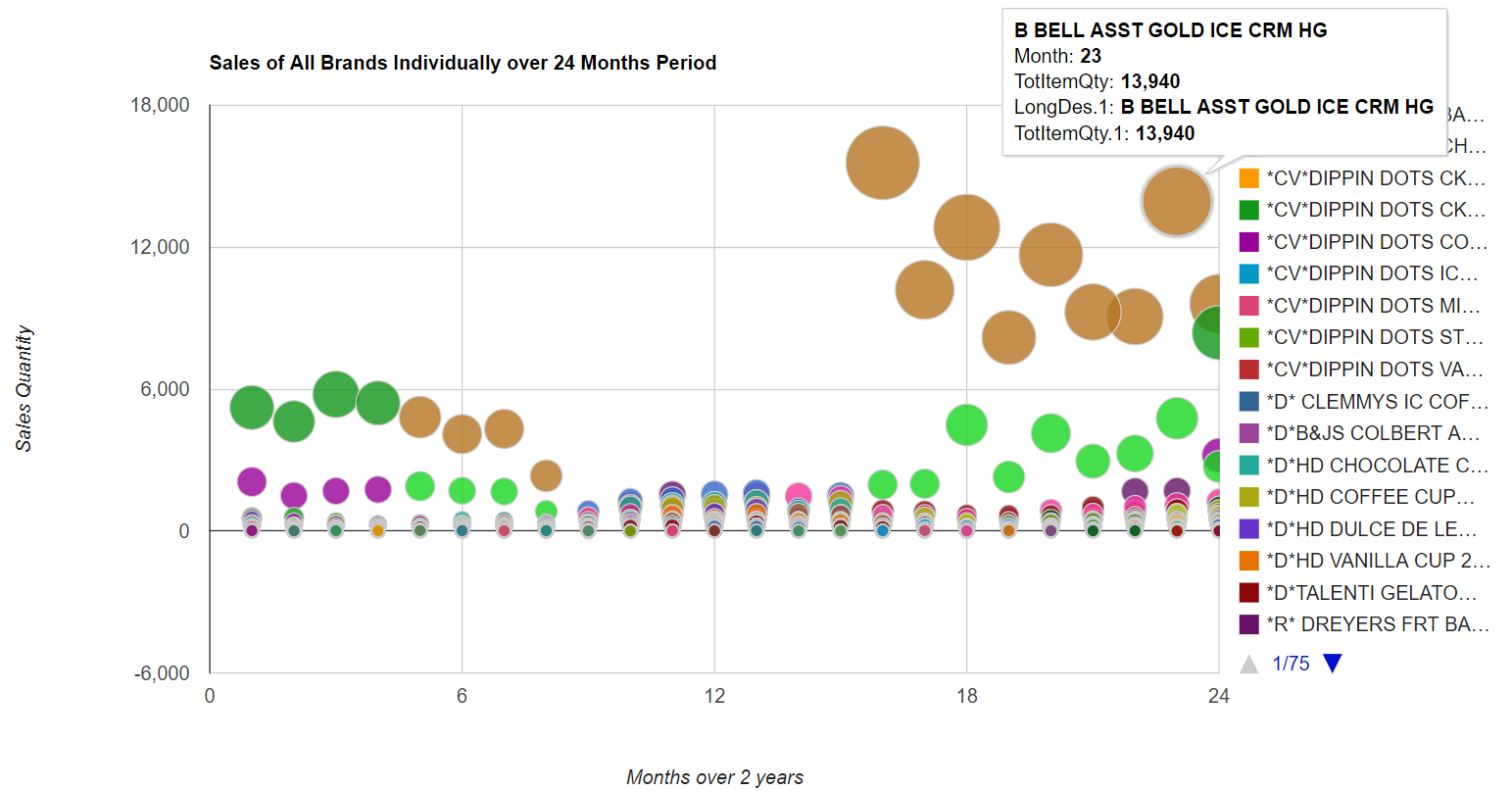


Figure 13. Bubble chart representing sales of all brands including Blue Bell brand products over 24 months.

In the chart, we can see the sales of all the ice cream brands in terms of quantity. Also, we can interpret the sales quantity from bubble size. The recall period is visible here in the chart as well. One reason behind the chart showing Blue Bell circles bigger in the post recall period is that we are showing individual brands here. While the aggregated sales of all other brands are higher than Blue Bell, individually they can’t possibly compete Blue Bell.

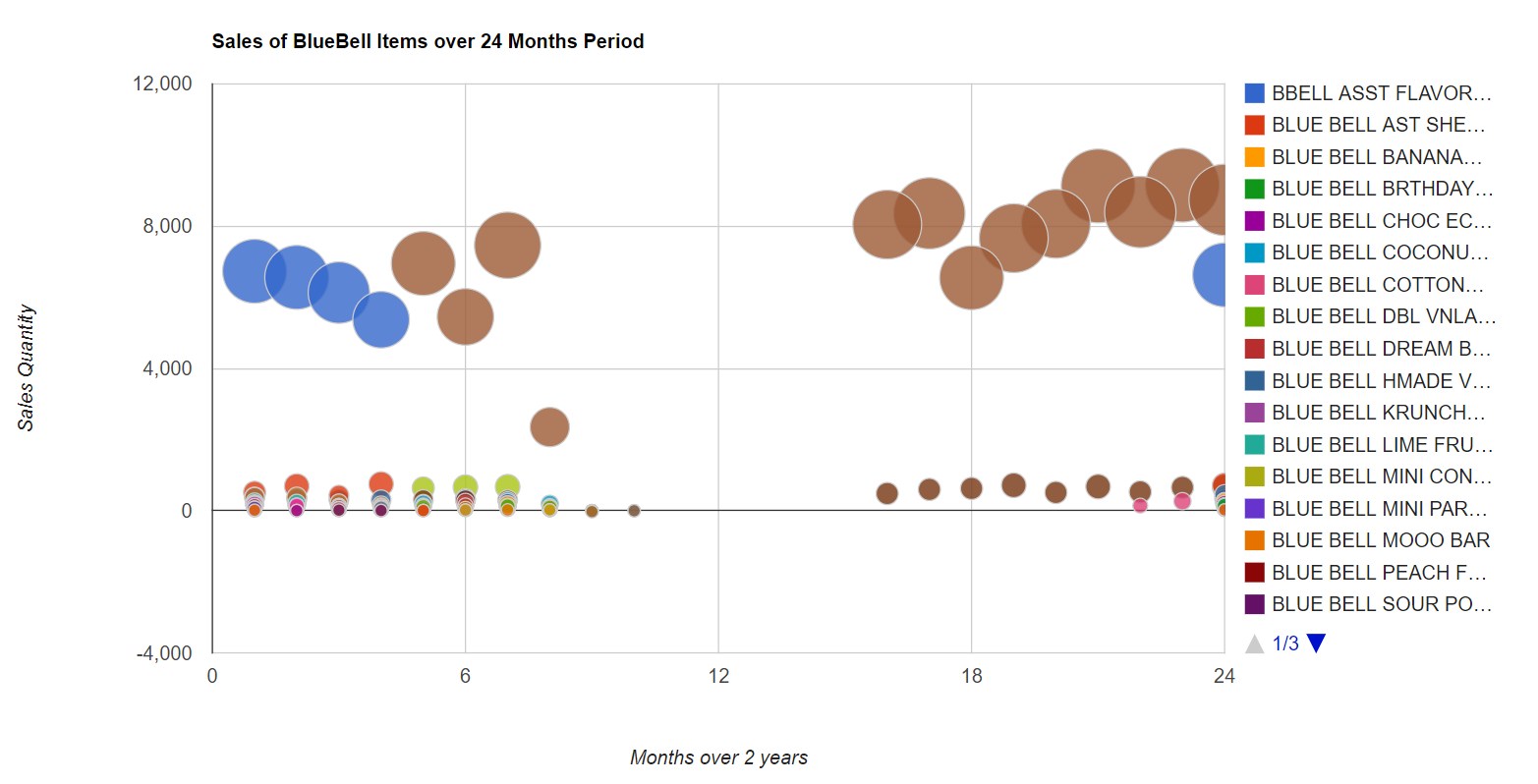


Figure 14. Bubble chart representing sales of all products of Blue Bell brand over 24 months.

The next chart is showing only the Blue Bell product sales in the 24 months period. From this chart, we can deduce that there are some Blue Bell products which became popular post recall period. One reason behind this statistic may be that because this brand was releasing less items in the market, Blue Bell fans were buying whichever item they can get in the stores.

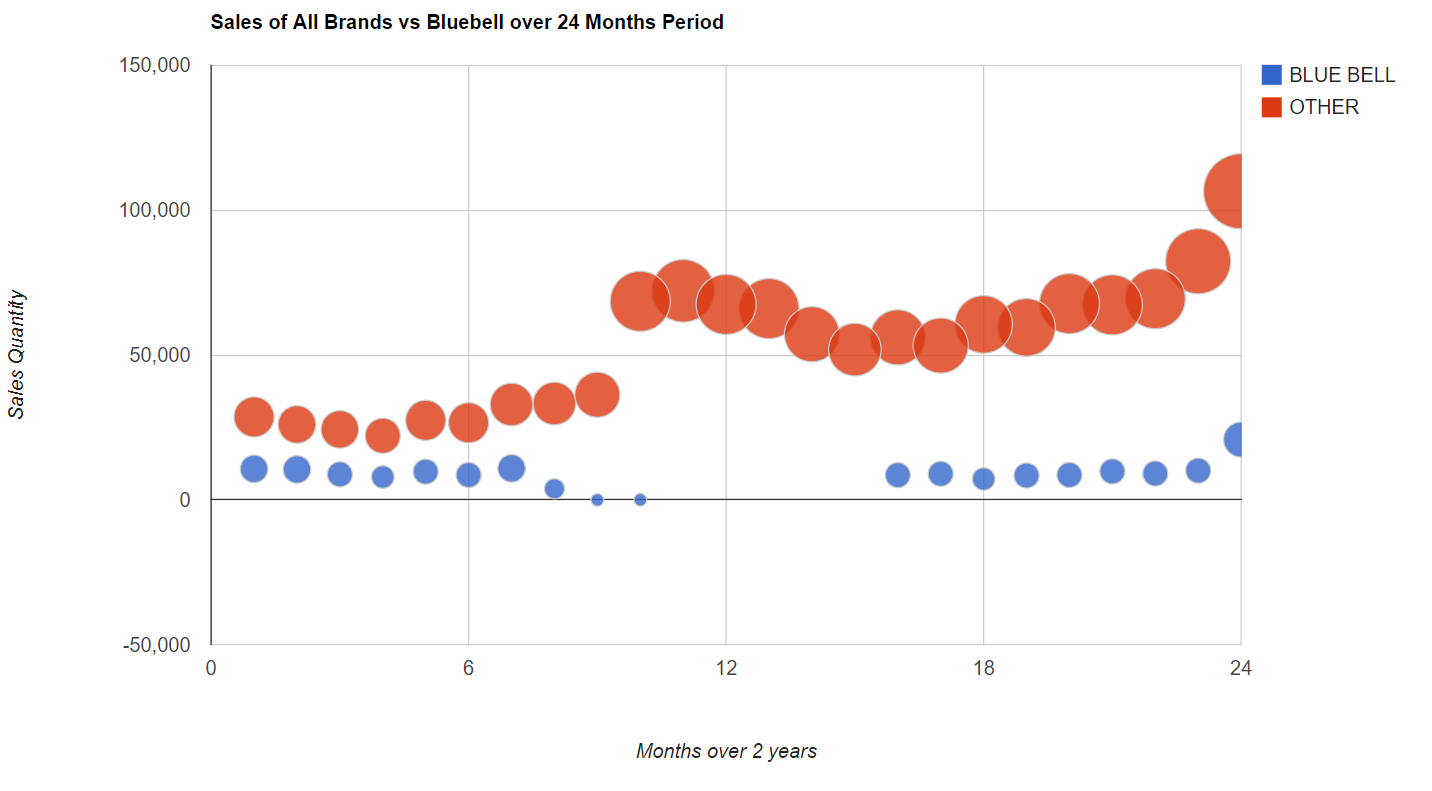


Figure 15. Bubble chart representing sales of all brands with Blue Bell over 24 months.

In the last graph, we are interpreting the overall sales statistics for Blue Bell and other brands. Here we can see that the sales of other products became very high during the recall and post recall period.

# Conclusion

The Blue Bell Recall had a huge impact on the Blue Bell post-recall sales. It seems that the other direct competitors of Blue Bell had a direct advantage over Blue Bell, over all ice-cream and other ice cream novelty sales over the recall period and then after. Our dataset lacked enough post-recall data to fully analyze the post-recall effect of Blue Bell. The company’s post-recall re-entry into the ice cream market was divided within several months and 5 different phases. We only saw the reintroduction of few of its products during the post-recall period we took into consideration for our study.

### Next Steps

The next step for our project is going to be incorporating a year worth of post-recall data to fully analyze the level of sales and determining the level of loyalty Blue Bell customers showed to this brand. Being the biggest market giant among ice cream products, Blue Bell is going to be an interesting customer loyalty project to tell if it’s loyal customers will return to it’s ice-creams.

We should also incorporate other ice brands that compete with Blue Bell for further analysis of what had happened when Blue Bell was completely out from the market.

## References

CDC. (2015, June 10). *Multistate Outbreak of Listeriosis Linked to Blue Bell Creameries Products (Final Update).* Retrieved from CDC: https://www.cdc.gov/listeria/outbreaks/ice-cream-03-15/

Ramzy, A. (2015, April 21). *Blue Bell Recalls All Products After Listeria Outbreak.* Retrieved from The New York Times: https://www.nytimes.com/2015/04/22/business/blue-bell-recalls-all-products-after-listeria-outbreak.html?\_r=0

# 

# Appendix

Table 2: Logic behind Month labeling

|  |  |  |
| --- | --- | --- |
| **Month** | **Transaction Start Date (YYYY-MM-DD)** | **Transaction End Date**  **(YYYY-MM-DD)** |
| 1 | 2014-09-01 | 2014-09-31 |
| 2 | 2014-10-01 | 2014-10-31 |
| 3 | 2014-11-01 | 2014-11-31 |
| 4 | 2014-12-01 | 2014-12-31 |
| 5 | 2015-01-01 | 2015-01-31 |
| 6 | 2015-02-01 | 2015-02-31 |
| 7 | 2015-03-01 | 2015-03-31 |
| 8 | 2015-04-01 | 2015-04-31 |
| 9 | 2015-05-01 | 2015-05-31 |
| 10 | 2015-06-01 | 2015-06-31 |
| 11 | 2015-07-01 | 2015-07-31 |
| 12 | 2015-08-01 | 2015-08-31 |
| 13 | 2015-09-01 | 2015-09-31 |
| 14 | 2015-10-01 | 2015-10-31 |
| 15 | 2015-11-01 | 2015-11-31 |
| 16 | 2015-12-01 | 2015-12-31 |
| 17 | 2016-01-01 | 2016-01-31 |
| 18 | 2016-02-01 | 2016-02-31 |
| 19 | 2016-03-01 | 2016-03-31 |
| 20 | 2016-04-01 | 2016-04-31 |
| 21 | 2016-05-01 | 2016-05-31 |
| 22 | 2016-06-01 | 2016-06-31 |
| 23 | 2016-07-01 | 2016-07-31 |
| 24 | 2016-08-01 | 2016-08-31 |