

**Data Analysis Case:**  
**Retail Sales Forecasting**

Anna Avetissian, David Gukasyan, Caleb Harding,  
Nicole Joseph, Alissa Lara, Shayne Yamazaki

Team 2

BUS 312: Data Literacy for Business (20656 - SP2024)

Professor Wayne Smith

April 10th, 2024

# MEMORANDUM

**Date:** April 10th, 2024

**To:** Professor Wayne Smith

**From:** Team 2

**Subject:** Brazilian Retailer (Firm Name Undisclosed)

---

Our team consists of individuals who have different backgrounds and skills in business and data analysis. This allowed accurate data analysis to the end of the primary goal, which was to accurately forecast demand for 2016Q4. The team approached the analysis by reviewing the historical sales data provided by the firm to forecast demand for 2016Q4. This demand forecast should assist the firm in managing inventory levels for the upcoming quarter.

It has been a delightful experience for the team to analyze your sales data and produce insights for management. By forecasting demand, it should be easier to give assistance to supply chain practitioners and have the appropriate measures carried out. If there are any further questions or concerns, please feel free to contact us. We look forward to working together in the future.

# Table of Contents

Section I - Executive Summary.....	1
Section II - Introduction .....	2
Team Members .....	2
Case Choice: Overview .....	3
Operations in Retail .....	3
Section III - Getting an Understanding .....	4
Section IV - Intermediate Forecast for Current Period (2016Q3) .....	6
Section V - Forecast for Next Period (2016Q4) .....	11
Key Findings .....	13
Section VI - Ethical Data-Handling .....	14
The Importance of Ethical Data-Handling .....	14
U.N. Global Compact Classification & Data Privacy .....	15
Classical Ethical Theories & Data-Handling .....	16
Section VII - Conclusion .....	17
Limitations and Future Analysis .....	17
Recommendations .....	20
Works Cited .....	21
Appendix .....	22

# EXECUTIVE SUMMARY

This case follows an undisclosed Brazilian retailer and daily business data for one of its products, including selling price, quantity sold, quantity in stock. Data is recorded for all days between January 1, 2014 and July 31st, 2016, leaving two months of 2016Q3 and all of 2016Q4 yet to be recorded. The ensuing analysis centers around the production of an intermediate demand forecast for the remaining portion of 2016Q3 and an ensemble demand forecast for 2016Q4 using a trend-adjusted moving average and linear regression to facilitate management of inventory levels for the last quarter of 2016.

The key recommendation for this Brazilian retailer is to manage inventory levels such that on October 1, 2016, inventory level for the product in question is at a minimum of 13,000 units plus the minimum allowable inventory level, which will function as a cushion in the event of an under-forecast. Additionally, the firm should adjust the sale price of this product to match market equilibrium as of the 2016Q4 upon reaching October to ensure that they do not incur negative economic profit as a result of mismatched supply and selling price, as the demand forecast has been performed in a manner that accounts for projected changes in underlying economic conditions.

## Introduction

### Team Members

**Anna Avetissian** is a third-year pursuing a degree in Business Management. Her skills include statistics, Excel (education), problem solving, and time management (experience). She optimizes Excel usage by efficiently managing time for spreadsheet creation, data analysis, and report generation, while applying problem-solving techniques to resolve intricate business issues through financial analysis, or identifying trends.

**David Gukasyan** is a third-year Business Analytics Major with a minor in Quality Management & Assurance. His skills include statistics, Excel, SQL, data visualization (education), and Python (experience). These skills are essential to data-driven decision-making in business because they collectively cover all components of analysis, including model development, deployment, and evaluation.

**Caleb Harding** a third year Pre-Business Law Major. His skills include Excel, Statistics (Education), and general Finance (Education). These skills show the foundation for Pre-Law, and to understand the basis of law,finance, and computing skills.

**Nicole Joseph** is a third-year pursuing a degree in Financial Analysis. Skills she has fostered include Excel, analyzing sales data (experience) and moving average forecasting (education). She applies these skills to promote accurate business decisions supported by experience and data. She was drawn to retail sales forecast data because it is something many contribute to.

**Alissa Lara** is a second-year student who is pursuing a degree in Marketing. The skills that she has acquired include analyzing various types of data and product trends (education), information

visualization (experience), and Excel. These skills are valuable when making business decisions that are driven by data because they allow her to properly analyze, interpret and communicate the significance of the data that is given, which then effectively allows her to make an informed business decision based on said data.

**Shayne Yamazaki** is a third-year Pre-Accountancy Major. His skills include Excel (education), financial analysis (education), and ethical awareness (education). These skills allow him to understand the data and make business decisions based on his analysis of the data.

### Case Choice: Overview

The team chose the Retail Sales Forecasting dataset as the primary option because, given the diverse majors and academic background within the group, it was apparent that opting for a more straightforward dataset would streamline the analysis process, particularly in contrast to potential datasets featuring a complexity of up to twenty variables. Another case choice considered was NYC Airbnb Open Data, which could offer insights into the dynamics of the sharing economy and the behavior of hosts and travelers, and could offer valuable information about accommodation trends and preferences in New York. But in the end, the simpler dataset was favored.

### Operations in Retail

An article concerning retail forecasting discusses the decline in German retail sales by 1.9% in February 2024, failing to meet the anticipated 0.5% increase, highlighting persistent challenges within the sector amidst subdued consumer confidence. This marks the fourth consecutive month of declining sales, with significant drops in food retail and online sales.

Weak consumer confidence persists, hindering purchasing willingness, but optimism remains for potential improvements in household spending.

Another article details Swedish retailer Hennes & Mauritz (H&M) exceeding earnings expectations in February 2024, driven by strong sales of spring collections. Despite challenges like lower consumer spending and competition from online rivals, H&M aims to achieve a 10% operating margin in 2024 through cost-saving measures and strategic initiatives. The company is closely monitoring supply chain disruptions, while accelerating store upgrades and expanding local sourcing to mitigate risks and improve performance.

The dataset selected for analysis in this project comprises comprehensive historical sales data from a leading Brazilian retailer, covering a wide range of SKUs and stores. Transformed to protect the retailer's identity, this dataset serves as a valuable resource for exploring inventory management challenges in the retail and consumer goods industries. With the perpetual dilemma of balancing inventory levels to avoid excess costs or lost sales, accurate demand forecasting is crucial. The dataset's span offers an opportunity to develop short-term forecasting models, aiding in the optimization of inventory levels for the next quarter. By leveraging advanced analytical techniques and collaborative efforts, actionable recommendations were able to be made.

## **Getting an Understanding**

### *Descriptive Analytics*

In an effort to better understand the dataset, several visualizations were constructed and exploratory analyses were performed. It appears that the distributions of **sales** and **revenue**

(each recorded daily in the dataset, since **revenue** was derived from multiplying **sales** and **price**) are both right-skewed (Figure 1, 2). A majority of daily unit sales are 0-100 units, and a majority of daily revenue is between 0-200 Brazilian Reals (R\$).

With daily distributions of **sales** and **revenue** in mind, it was considered that there could be time-based trends that result in the observed skews. As shown, both total unit sales per quarter and total revenue per quarter are trending upwards (Figure 3, 4). It should be noted that the sharp decline in values for these variables in the most recent period, 2016Q3, are not accurate, as the dataset only records values for these variables up to July 31st of 2016, which left only one month of data to accumulate within this quarter as opposed to the typical three. Similarly, the same upwards trend and sharp decline in the most recent quarter is observed in quarterly stock purchases (Figure 6). This coinciding upwards trend with **sales** implies that more stock of the product in question is being purchased in order to satisfy increasing demand.

Viewing the distributions of **sales** within each quarter by day, it is observable that the most common bin for **sales** is 1-100 (Figure 6). Nearly every single quarter follows this pattern, and they all appear to be skewed-right, as there is more data to the right of the mode than to the left. This implies that the overall right-skewness of **sales** holds within each quarter as well, and is not simply a product of the concatenation of quarterly data into population data.

In addition to **sales**, **price** also appears to be skewed-right (Figure 7). This is corroborated by the fact that the mode is less than the median and that, in turn, is less than the mean. Dispersion wise, it appears that the standard deviations of **sales** and **price** are 80.7 sales and R\$0.53, respectively. This means that the typical amount by which a day's **sales** deviates from the



mean of 90.5 is 80.7 sales, and the typical amount by which a day's **price** deviates from the mean of R\$1.59 is R\$0.53.

The count of 937 for both variables indicates that data for 937 days is recorded, as each row corresponds to one day of data (Figure 7). This, however, does not correspond to a full three years, despite containing data from 2014-2016. The difference is accounted for by the fact that data for 2016 is incomplete, as data collection ended with the first month of 2016Q3, leaving a full five months of 2016 that have yet to be recorded. From an operations perspective, this raises a most critical question about what demand for these remaining five months might be, prompting a forecast analysis for the rest of 2016Q3 and the entirety of 2016Q4 to help the firm meet demand efficiently.

### **Intermediate Forecast for Current Period (2016Q3)**

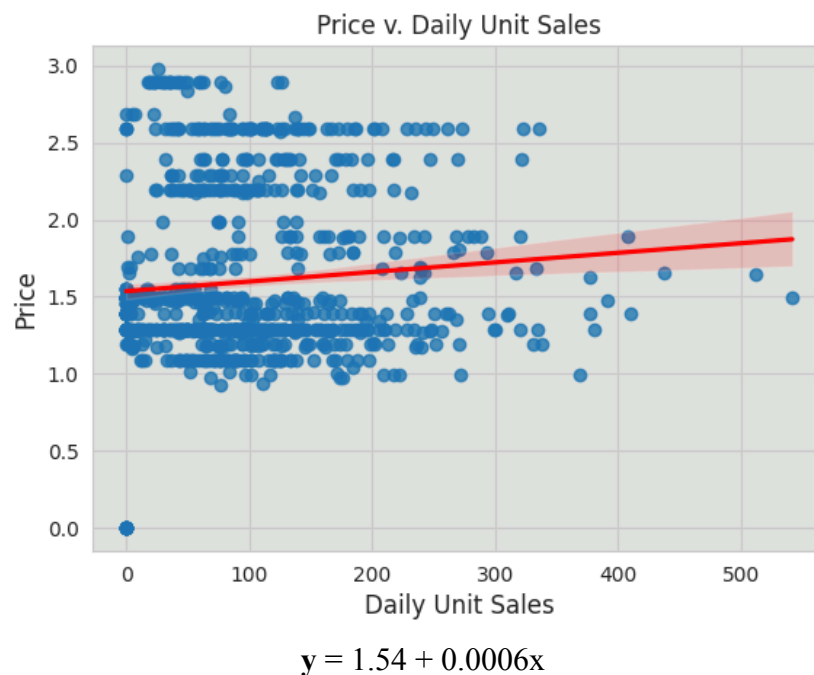
#### *Diagnostic-Level Analytics*

With the objective of accurately forecasting demand for 2016Q4 in mind, it is imperative that data from 2016Q3 is included in the forecast calculation. However, because this dataset only contains records up through July 31st, the first month of 2016Q3, a problem is encountered because reliable time series forecasting methods assume that data is collected in equal intervals. On a quarter-by-quarter basis, this means that each record on quarterly sales must contain sales data for an entire quarter, as opposed to the final record containing data at an interval of one month from the previous record.

From a business perspective, it is unwise to wait until all data for 2016Q3 is collected before making a forecast for 2016Q4, as all data will be collected only when Q4 has already been

reached, leaving no time to implement forecast-based recommendations before they are demanded of the firm. Therefore, an intermediate forecast, one that forecasts the total demand for 2016Q3, is required to accurately produce a Q4 forecast.

Quantity demanded, however, is dependent upon the price that a product is sold at. Because the firm regularly adjusts its sale price, one might consider that an approach that relies solely upon values for **sales**, recording quantity sold, is inappropriate, as the corresponding price would confound any forecast made. Therefore, one might assume that an approach that considers both **sales** and **price** would be most appropriate. However, this line of reasoning assumes that the firm's price changes take effect under equivalent economic conditions, and therefore all price changes and subsequent quantities demanded belong to the same demand curve. This core assumption of a **sales/price**-based approach is demonstrated invalid by the relationship between **sales** and **price**:



If it is assumed that the product being sold is a normal good (that is, that quantity demanded tends to be lower at higher prices), the weak relationship between **sales** and **price** (slope = 0,  $R^2 = 0.008$ ) implies that price changes are a response to changing economic conditions, and therefore, quantity demanded is consistently in equilibrium when sales occur, with price changes being reflective of changes to the supply/demand curves themselves rather than taking place along the same demand curve. With this knowledge, it can be concluded that an approach that considers only **sales** is appropriate, as the sale price itself can simply be adjusted to be sold at the equilibrium price dictated by the market's underlying conditions in 2016Q4.

The approach to make this intermediate forecast for 2016Q3 will rely on using patterns in past data to infer the remaining sales yet to be made for August and September of 2016 (the remaining months of 2016Q3), which will be added to the actual sales made in July. For this, it must be determined whether the daily sales made in July of 2016Q3 are statistically significant relative to the daily sales of the entire population of days, as selecting the appropriate forecasting method relies on the similarity of recent data to historical data. With t-test conditions being met as best as possible for time series data (random sample selection being less relevant since it is desired to deliberately check the most recent values against historical values for significance, with independence and normality met outright since each day's sales are not affected by previous days and  $n > 30$ ), a two-sided t-test was used where the hypothesized mean is the mean **sales** of all days for which data is available and the sample mean is the mean of **sales** for all days in July 2016 (the current available data for 2016Q3):

#### Diagnosing Daily Sales Relative to All Quarters

Hypothesized Mean: 90.53  
Sample Mean: 166.26  
Sample Std: 77.48  
alpha: 0.05  
t\*: 5.44  
p-value: 6.699427143782265e-06

Conclusion:  
Reject the null hypothesis

With such a low p-value, it can be concluded that July 2016's **sales** are statistically significant, with the positive t-statistic indicating that they are statistically significant in the positive direction. In other words, July 2016 **sales** are significantly greater than what would be expected based on the entire population of daily sales, meaning that as of July 31st, 2016Q3 is strongly outperforming what would be expected based on the population mean.

Two things, however, should also be considered: 1) Such a hypothesis test does not account for the margin of error for July 2016 **sales**, and 2) Apparent statistically-significant results may be the result of cyclicalities. To account for margin of error, a one-sample t-interval (c-level = 0.95) was constructed for July 2016, with the following results:

Confidence interval: (137.84, 194.68)

Even when re-running the hypothesis test with the least favorable sample mean for July 2016 (137, rounded down from the lower boundary), the resulting p-value ( $p = 0.002$ ) is still well into the 5% significance threshold, indicating that, for all intents and purposes, it is safe to assume that July 2016 is significantly outperforming the population.

Cyclicalities not being accounted for limits the ability to draw inferences from t-test results. For example, Q4, for most retailers, consistently outperforms the other quarters in sales,

as Christmas shopping places upwards pressure on sales. This pattern is cyclical, so it is expected to occur every year. To confirm whether or not the apparently significant results observed for 2016Q3 are the result of cyclicalities as opposed to a genuine outperformance of the population, 2016Q3 **sales** must also be compared independently to all Julys and Q3 sales for the available years (2014-2015). For this, two-sided t-tests were used, with the first using hypothesized mean of mean **sales** for all days in Q3 of 2014-2015 and the second using the mean of **sales** for all days in Julys of 2014-2015 (data from 2016 is excluded from the population, since the low availability of years would potentially cause 2016Q3 to skew the data). Both tests use a sample mean of the mean of **sales** for all days in July 2016:

#### Diagnosing Daily Sales Relative to Previous 3rd-Quarters

Hypothesized Mean: 68.24  
Sample Mean: 166.26  
Sample Std: 77.48  
alpha: 0.05  
t\*: 7.04  
p-value: 7.889322097873018e-08

Conclusion:  
Reject the null hypothesis

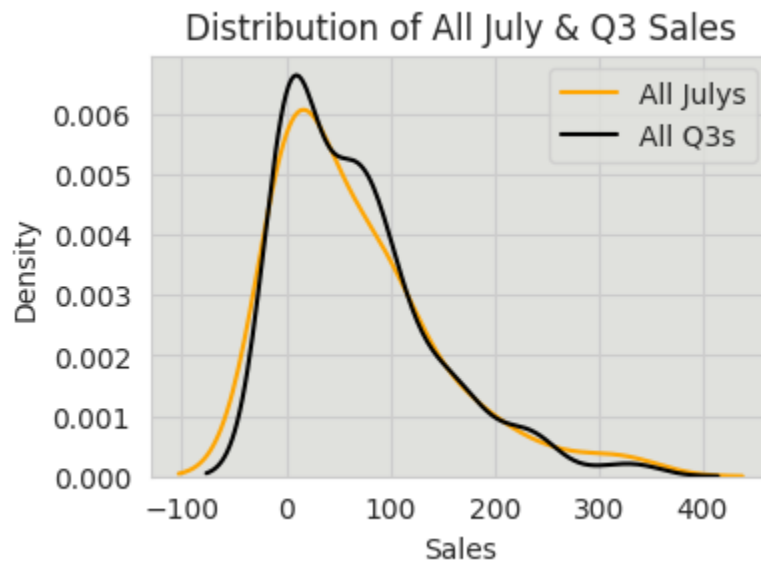
#### Diagnosing Daily Sales Relative to Previous Julys

Hypothesized Mean: 68.31  
Sample Mean: 166.26  
Sample Std: 77.48  
alpha: 0.05  
t\*: 7.04  
p-value: 7.991132800230547e-08

Conclusion:  
Reject the null hypothesis

Again, the low p-values and the positive t-statistics indicate that July 2016's **sales** are statistically significant in the positive direction. Therefore, it is reasonable to assume that the remaining months of 2016Q3 will likely be statistically significant as well, since July 2016 is

significantly outperforming other Q3 months from previous years as well. Further, the identical t-statistics reveal that there is not much variation in **sales** between July and the other Q3 months in a given year, as also shown by this visualization:



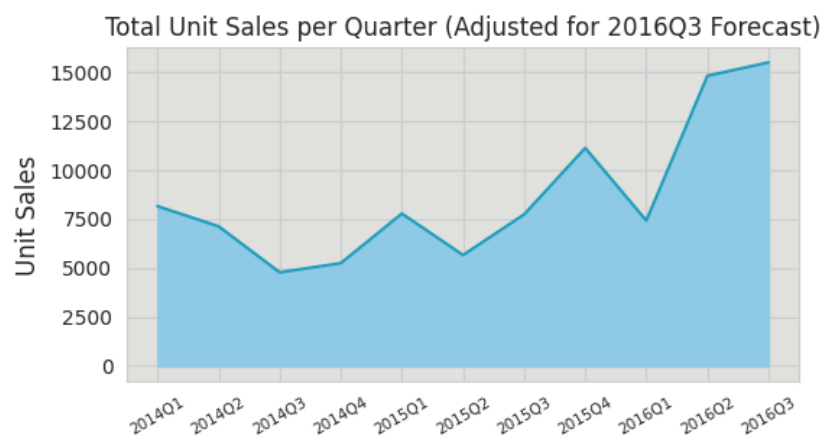
In making an intermediate forecast for 2016Q3, it should be considered that the aforementioned results indicate that the forecasted **sales** should be greater than the actual **sales** for previous Q3s. Extrapolating the relationship between July sales and overall Q3 sales (near 1:1), it can be inferred with some certainty that an appropriate forecast for each of the remaining two months in 2016Q3 is equivalent to the actual total value of **sales** for the first month of the quarter, July. Therefore, since all three months are projected to have equivalent **sales**, an appropriate intermediate forecast for 2016Q3 is three times the July 2016 **sales** of 5,154, for a final forecast of 15,462 unit sales.

### **Forecast for Next Period (2016Q4)**

*Predictive-Level Analytics*

With a forecast of 15,462 unit sales for 2016Q3, 10,308 of which are yet to be realized, a demand forecast for 2016Q4 is preparable. While an exponential smoothing model would be preferred, the lack of actual values for 2016Q3's total sales makes it such that exponential smoothing will always produce a forecast of 15,462 sales no matter the smoothing constant when operating under the assumption that 2016Q3's forecast is 100% accurate. Any other assumption for the actual value would be unscientific and even more prone to error. For this reason, a trend-adjusted moving average will be used to forecast 2016Q4.

Since 2014Q3, sales per quarter have demonstrated an upwards trend:



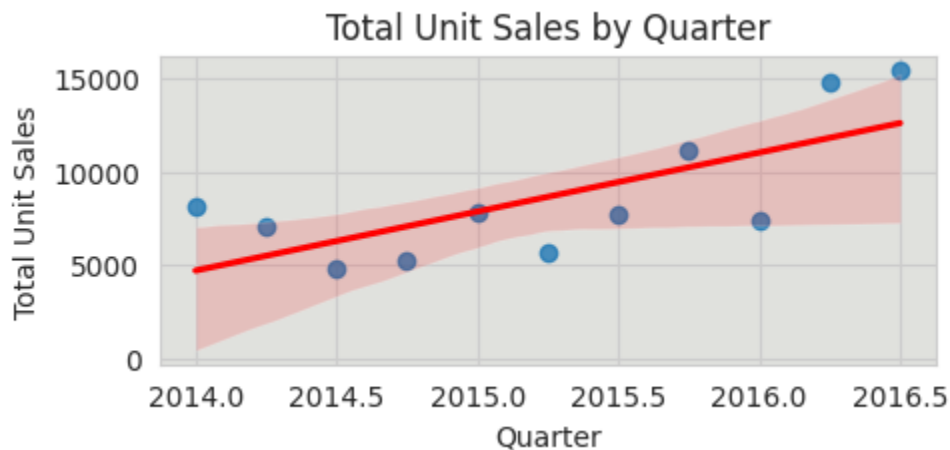
Assuming that this trend will continue, it would be reasonable to infer that the most relevant values for calculating a next-period forecast are the most recent values. Incorporating a trend-adjusted moving average using the three most recent values (2016Q1, Q2, and Q3), where the trend-adjustment comes in the form of adding the mean change in sales of the previous three data points to the moving average produces the following forecast:

	quarter	current_mean_change_in_sales	change_in_sales	sales
0	2015Q4	1453.333333	NaN	11102.0
1	2016Q1	1453.333333	-3699.0	7403.0
2	2016Q2	1453.333333	7379.0	14782.0
3	2016Q3	1453.333333	680.0	15462.0
4	2016Q4	1453.333333	NaN	NaN

	quarter	current_mean_change_in_sales	change_in_sales	sales
0	2015Q4	725.0	NaN	11102.0
1	2016Q1	725.0	-3699.0	7403.0
2	2016Q2	725.0	7379.0	14782.0
3	2016Q3	725.0	680.0	15462.0
4	2016Q4	725.0	-1460.0	14002.0

With the previous three data points having a mean **sales** of 12,549 and a mean **change\_in\_sales** of 1,453, the forecast for 2016Q4 **sales** sums to 14,002 units.

Comparing this approach to simple linear regression, it is found the trend-adjusted moving average forecast produces similar results:



$$y = -6358285.3 + 3159.38x \mid r = 0.72 \mid R^2 = 0.52$$

Since quarters are denoted by decimal point (2014.0 indicating 2014Q1, 2014.25 indicating 2015Q2, etc.), a forecast for 2016Q4 can be derived by plugging in “2016.75” for “x”. Upon



doing so, the forecasted total unit sales for 2016Q4 comes out to 13,384 units, a more modest forecast relative to the trend-adjusted moving average approach, but within the same ballpark. With a positive slope and correlation coefficient, **quarter** (time) and **sales** have a positive relationship, indicating that later periods in time tend to have higher **sales**. Given the decently high R<sup>2</sup>-value (0.52), it appears that this model fits the data somewhat well, as 52% of the variation in total unit **sales** can be explained by variation in time, as measured by quarter.

### Key Findings

All this leads to the conclusion that an appropriate forecast for 2016Q4 **sales** is in the 13,000-14,000 units range, and therefore, the firm should take steps to ensure an inventory stock of at least 13,000 units plus the minimum inventory threshold potentially implemented by the firm to act as a cushion. Accordingly, the firm should place an order for the appropriate quantity of this particular item to meet this forecasted demand.

### **Ethical Data-Handling**

#### *Firm-level Ethical Analysis*

### The Importance of Ethical Data-Handling

Upon review of the dataset, the team has concluded that ethical issues do not arise from this dataset because the firm took active steps to protect against unethical use. Because the identity of the retailer was not disclosed and the dataset was transformed to protect the identity of the retailer, there was limited room for unethical conduct outside of the fallout of a potential leak of the actual data or identity of the retailer. Therefore, instead of an ethical analysis on the dataset itself, a search was conducted for instances of retail data being stolen or used unethically

in a way that could have been avoided had the data been kept more private as this Brazilian retailer has done.

An article from Forbes titled, “Target Data Breach Spilled On As Many As 70 Million Customers,” states that “information stolen included customer names, credit or debit card number, the card’s expiration date, and CVV... Target said an additional 70 million people were affected, and the stolen customer information includes names, mailing addresses, phone numbers, and email addresses,” (McGrath, 2014). This caused millions of consumers emotional distress due to having to change valuable information and fear of identity theft. While this article is about consumer information breach as opposed to business information breach, it still exhibits the impacts of negligent/unethical handling of data for firms. A more comprehensive list of the impacts of unethical handling of data for firms include damaging of reputation, financial loss, sensitive data loss, legal implications and actions, and operational downtime. Because firms often struggle to deal with the consequences of negligent handling of data, it is wise to take steps, as this Brazilian retailer did, to protect business data. Without intentional steps to protect business data, competitors would be able to use the firm’s data to come up with strategies against them, an outcome functionally equivalent to that of corporate espionage.

Another issue is sensitive data loss. A data breach of the Brazilian retailer could cause significant harm that would breach the firm’s moral responsibility to its owners and employees, which is a particularly important matter in a country like Brazil that has a high unemployment rate of 24.01% as of 2016 (Statista Research Department, 2024). A breach of employee data, such as annual tax documents, could cause employees to worry about identity theft just as Target customers did.

### U.N. Global Compact Classification & Data Privacy

A data breach of firm Team 2 would fall under the U.N. Global Compact Classification of Principle 1 of Human Rights. It falls under Principle 1 since the firm would not be “supporting and respecting the protection of internationally proclaimed human rights” (Principle 1 / UN Global Compact, 2024). The data breach would violate the right of privacy and right of security, as employees’ personal information such as name, birthday, financial information, address, etc. would be exposed to potential bad actors. Causing employees to worry for their security could have adverse business consequences as well, as employees may be less trusting of the firm and less motivated to work, even to the extent of causing financial harm to the firm. This right of security would be violated if sensitive employee information is accessed without authorization. This Brazilian firm, however, avoids such problems by not disclosing their identity and transforming their business data prior to sharing online, exemplifying careful and ethical data-handling.

### Classical Ethical Theories & Data-Handling

The Brazilian retailer who provided this retail forecasting data has done a successful job in being able to acquire data ethically. The data has been transformed, protecting not only the retailer but the customers who participated in sharing their data as well, as the publicly-available data excludes any and all consumer information. Based on the information that was provided on how the data was acquired, the ethical approach that is felt most appropriate is deontology. Deontology is the ethical decision-making theory in which the morality of decisions are based on the intentions of actions rather than their outcome.

Recently it is very common for data to be stolen, used or accessed unethically. An article written by Min-Seok Pang and Anthony Vance discusses how the number of data breaches has significantly increased and how individuals are affected by data breaches. The article states, “In 2022, there were 1,802 reported incidents of data being compromised in the U.S., just short of the record 1,862 in 2021, according to the Identity Theft Resource Center.” (Pang & Vance). With the number of breaches rapidly increasing it is important for companies to develop further security measures that protect the data of their customers and for their company as a whole as well, it is important that if there was ever a breach that no crucial information about the customers or company is out in the open to be taken and misused. By making business decisions using deontological ethics, the possibility of acting ethically by happenstance if a firm “gets away with it” (as in, the leak somehow does not result in significant negative consequences for others) provided by utilitarianism is bypassed, resulting in less problems relating to negligent data-handling by firms.

Although deontology is what is believed best fits the case, another ethical theory that could be justified to use is right-based ethics. Rights-based ethics is the theory that every individual is entitled rights such as privacy, equal treatment, security, and freedom to achieve self-actualization. In this instance, the retailer upheld customer privacy by keeping their information private. However, it has been seen that some datasets are not as protected as this one, and in those instances, a data breach results in consumer information being exposed, which violates their rights to privacy and security. Consequently, a good principle when making business decisions is to consider the rights of every individual before making data-related decisions.

## Conclusions

### Limitations & Future Analysis

While evaluating the dataset, the team remained aware of cognitive biases that may distort or influence the decision making process. In an effort to overcome and avoid these biases, the team discussed. One common bias, hindsight bias, is the idea that future events are perceived as easier to forecast or foresee after the fact than during the actual forecasting process. In an effort to avoid this bias, the team has committed to affirming the inherently unpredictable nature of forecasts in any way dependent upon economic conditions. Furthermore, the previously-acquired knowledge of each team member has been acknowledged amongst the team in order to hold one another accountable for judgements that may be overtly influenced by such biases.

Certain limitations of the dataset make extrapolation more difficult. According to the dataset's description on Kaggle.com, this data has been transformed to protect the identity of the retailer. The actual value of the provided variables is unknown, as, for example, **sales** could have been divided by 10 without disclosure. Perhaps the original data was transformed by dividing all continuous variables by 10. If this were the case, the relationships among the variables would be preserved, allowing a simple proportional adjustment to make the team's findings applicable. It is also a possibility, however, that the transformations made do not preserve the relationships between the variables. For example, a transformation that adds 0.2 to **price** but multiplies **sales** and **stock** by 5 would not preserve the relationships, as such a transformation would alter relative proportions. Because there is no way of knowing the exact

transformations performed, this analysis' findings are not guaranteed to be applicable in a real world scenario.

Fluctuations in currency exchange rates can significantly impact sales figures, particularly for businesses that operate across borders. Tracking currency exchange rates over time and correlating them with sales data could reveal insights into how changes in exchange rates affect consumer behavior and purchasing power. The issue, however, is that since no data dictionary was provided, there is no way to know what currency **price** is measured in. In conjunction with the potential transformation to **price** and the lack of information about the product being sold, there is no way to certainly infer the currency used to measure **price**. Given these circumstances, **price** was assumed to be measured in Brazilian Reals (R\$) since the firm in question is a Brazilian retailer.

Third, an issue involving **stock\_purchases\_from\_previous**, which was engineered by subtracting the previous day's **stock** from the current day's **stock**, and then adding the current day's **sales**, is noted. This engineering approach assumes that **stock**, or inventory, is recorded at the end of the day, which appears to be a reasonable assumption given that it results in most days having a value of 0 for **stock\_purchases\_from\_previous**, indicating no inventory purchases. When this variable is negative for a given day, it indicates that there was some sort of shoplifting or damage to inventory that did not result in a sale. However, when this variable is non-negative, it is unknown whether shoplifting or damage occurred, as potential inventory purchases could conceal the inventory-reducing effect of shoplifting or damage.

Future analysis should focus on model evaluation and adjustment. Additional data on actual stock purchases would assist in expanding the analysis to produce insights on shoplifting prevention and proper product handling. Lastly, incorporation of economics and operations experts may assist in producing additional insights that could facilitate proper inventory management.

### Recommendations

All in all, this information would be valuable for making strategic decisions regarding pricing, inventory management, and expansion into new markets. With the intermediate demand forecast of a total of 15,492 units for 2016Q3 and a demand forecast of 13,000-14,000 units for 2016Q4, the firm has the necessary information to plan operations for the remainder of 2016. An order should be placed to bring **stock** up to at least 13,000 units plus the minimum allowable inventory in order to meet demand through the end of 2016Q4. Additionally, only **price** should be adjusted upon reaching 2016Q4 in order to maximize economic profit, as the 2016Q4 demand forecast accounts for changes in underlying economic conditions, while **price** must be manually adjusted to equilibrium.

## Work Cited

Pang, Min-Seok, and Anthony Vance. "Your Personal Data Has Been Stolen. Will It Really Cost You Anything? - WSJ." *The Wall Street Journal*, 6 June 2023, [www.wsj.com/articles/personal-data-stolen-what-cost-you-f0c8e6a1](https://www.wsj.com/articles/personal-data-stolen-what-cost-you-f0c8e6a1).

Chopping, Dominic. "Hennes & Mauritz Net Profit Beats Forecasts as Sales Pick ..." *The Wall Street Journal*, 27 Mar. 2024, [www.wsj.com/business/earnings/hennes-mauritz-net-profit-beats-forecasts-as-sales-pick-up-82eee846](https://www.wsj.com/business/earnings/hennes-mauritz-net-profit-beats-forecasts-as-sales-pick-up-82eee846).

Frankl, Ed. "In Surprise, German Retail Sales Drop - WSJ." *The Wall Street Journal*, 28 Mar. 2024, [www.wsj.com/articles/SB10001424052748703859204575525370047124374](https://www.wsj.com/articles/SB10001424052748703859204575525370047124374).

McGrath, Maggie. "Target Data Breach Spilled Info on as Many as 70 Million Customers." *Forbes*, Forbes Magazine, 28 June 2021, [www.forbes.com/sites/maggiemcgrath/2014/01/10/target-data-breach-spilled-info-on-as-many-as-70-million-customers/amp/](https://www.forbes.com/sites/maggiemcgrath/2014/01/10/target-data-breach-spilled-info-on-as-many-as-70-million-customers/amp/).

"Principle 1: UN Global Compact." *Principle 1 | UN Global Compact*, [unglobalcompact.org/what-is-gc/mission/principles/principle-1](https://unglobalcompact.org/what-is-gc/mission/principles/principle-1).

Statista Research Department, and Mar 7. "Brazil: Unemployment Rate by Gender 2022." *Statista*, 7 Mar. 2024, [www.statista.com/statistics/1343569/unemployment-rate-by-gender-brazil/](https://www.statista.com/statistics/1343569/unemployment-rate-by-gender-brazil/).



## Appendix

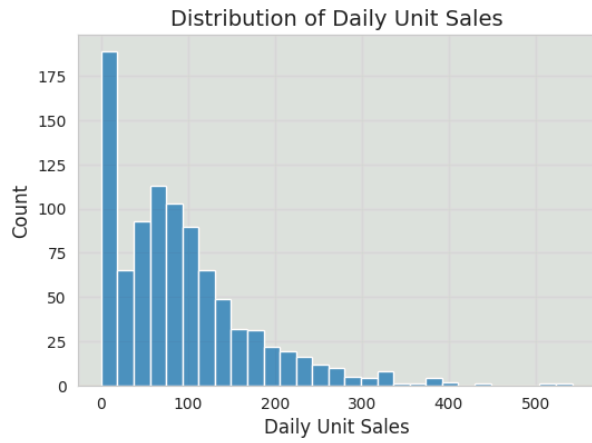


Figure 1

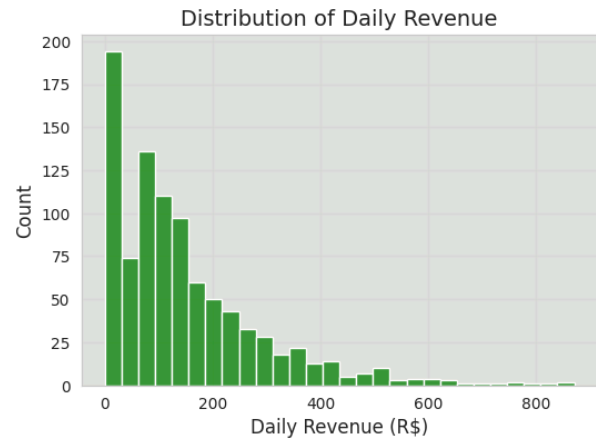


Figure 2

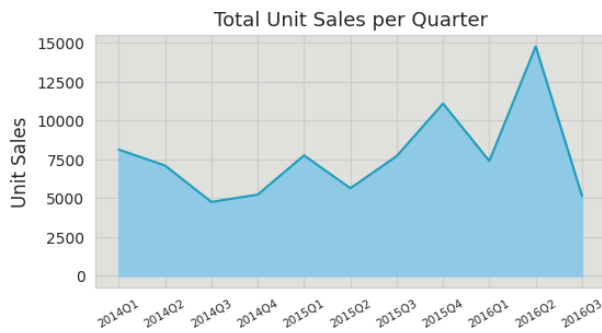


Figure 3

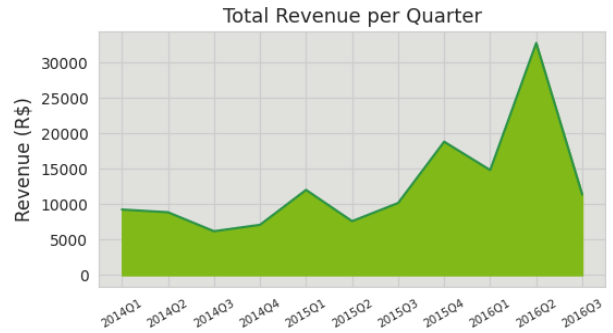


Figure 4

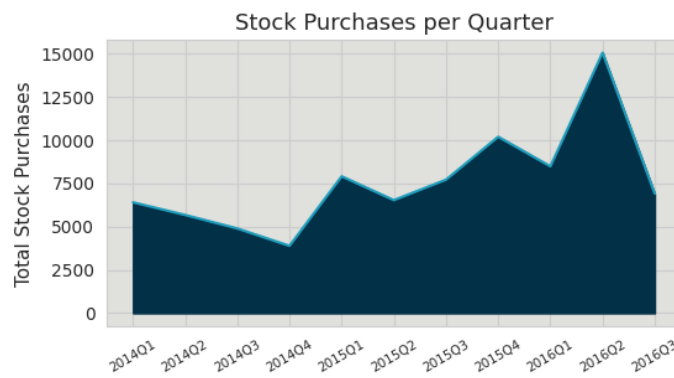


Figure 5

Count of sales Column Labels <span>▼</span>								
Row Labels <span>▼</span>	-99-0	1-100	101-200	201-300	301-400	401-500	501-600	Grand Total
2014Q1	1	56	22	4	2			85
2014Q2	5	62	20	3	1			91
2014Q3	34	41	11	3	2			91
2014Q4	27	49	14	1			1	92
2015Q1	3	62	19	5	1			90
2015Q2	25	40	22	4				91
2015Q3	1	61	23	7				92
2015Q4	4	43	32	9	3	1		92
2016Q1	8	56	22	5				91
2016Q2	4	19	40	19	6	2	1	91
2016Q3		8	14	7	2			31
<b>Grand Total</b>	<b>112</b>	<b>497</b>	<b>239</b>	<b>67</b>	<b>17</b>	<b>3</b>	<b>2</b>	<b>937</b>

Figure 6

<i>sales</i>		<i>price</i>	
Mean	90.53361793	Mean	1.592572038
Standard Error	2.635768974	Standard Error	0.017298062
Median	76	Median	1.39
Mode	0	Mode	1.29
Standard Deviation	80.68208948	Standard Deviation	0.529501561
Sample Variance	6509.599563	Sample Variance	0.280371903
Kurtosis	3.028263869	Kurtosis	0.474218903
Skewness	1.42699613	Skewness	0.771903909
Range	542	Range	2.98
Minimum	0	Minimum	0
Maximum	542	Maximum	2.98
Sum	84830	Sum	1492.24
Count	937	Count	937

Figure 7