

# TABLEAU PROJECT I: SUPERSTORE SALES

AN EXPLORATORY ANALYSIS ON THE EFFECT OF DISCOUNTS OF SALES PROFIT

**Chinh X. Mai**

*Last update: March 3, 2022, v0.4*

## CONTENTS

I.	Case description and Personas establishment .....	2
I.1.	Case description .....	2
I.2.	Personas .....	2
2.	Storytelling design checklist .....	3
2.1.	Background of the case .....	3
2.2.	Audience personas .....	3
2.3.	Checklist .....	3
3.	Variable choices and KPI construction .....	4
4.	Analysis details .....	4
5.	Story construction .....	4
6.	Design choices .....	4
7.	Handouts and Final visualization .....	5
8.	Final visualization analysis .....	5
9.	Related Information and Data .....	5
9.1.	Dataset .....	5
9.2.	Change log .....	5
9.3.	What to add .....	5
Appendix and References .....		5
Appendix .....		6
References .....		6

## I. CASE DESCRIPTION AND PERSONAS ESTABLISHMENT

### I.1. CASE DESCRIPTION

Imagine you are an analyst working for the Super Store Corporation. Your boss Sylvia is the VP of Sales and she is conducting a review of the company discount policy. As research for her review, she has asked you to perform an exploratory analysis of the Super Store discount data. Specifically, Sylvia has asked you to investigate if there is a relationship between discounted sales and profit, and how much the company is profiting or losing based on discounted sales.

She also wants to see if there are opportunities for efficiencies with respect to how discounts are handled. Perhaps certain customers are receiving more discounts than other customers. Perhaps certain product categories or product segments are receiving more discounts than others. Sylvia has given you wide latitude to explore this question however you see fit, and will ultimately make a recommendation to the board of directors based on your findings.

Sylvia will have only 5-7 minutes to present to board members, and she knows that her board members don't understand the minutiae of the business and prefer quickly hearing only the top-level facts. She also knows the room the presentation will be held in is large, meaning the visualization will be viewed from a great distance across the room. Please consult the Persona document included in this project for specific requirements of this audience.

### I.2. PERSONAS

#### Persona 1: Sylvia



*"ACME Superstores wants to make the most of sales data for better, more nimble decisions. I want to make sure that process works well."*

Sylvia needs to turn the many findings by the retail analysts into a sales strategy she can sell to the executive team.

**Role:** VP of Sales

**Organization:** ACME Superstore (your company)

**Goals:**

- Use sales data to help make more nimble inventory and distribution decisions
- Anticipate next moves in sales based on trends
- Be able to make a case for her sales strategy to her colleagues

**Challenges and Needs**

- Limited face time with executives.
- She's a good sales person but not great with numbers and stats
- She has to choose the most meaningful findings in the sales data to create a strategy but it's sometimes hard to tell what is most meaningful in the data.

#### Persona 2: Terrence (Super Store board member)



***I am interested in any idea that will make this company more efficient***

Terrence is a conservative, highly connected and financially-oriented board member. He feels a responsibility for ensuring the financial stability of the company. Is borderline obsessed with finding "efficiency" in business operations, where ever it may exist.

**Role:** Active Board Member of Super Store Corporation

**Age:** 56

**Gender:** Male

**Education:** MBA

**Goals:**

- Wants quick clear and credible updates
- Wants to ensure financial stability and make company more efficient
- Wishes to end fraud and abuse of company policy
- Prefers data-driven decision making
- Needs large fonts (poor eyesight)

**Challenges and Needs:**

Terrence has heard anecdotal reports of the company discount policy being abused. He is interested in learning if discounted sales are resulting in more or less profitable sales, and how much the company is profiting or losing based on discounted sales.

**Context:**

Monthly board meeting. There are 15 people in the audience looking at a big screen across a large room and also have a printed summary (handouts).

## 2. STORYTELLING DESIGN CHECKLIST

### 2.1. BACKGROUND OF THE CASE

#### Main tasks:

- Perform an exploratory analysis of the Super Store discount data
- Investigate if there is a relationship between discounted sales and profits
- Investigate by how much the company is profiting or losing based on discounted sales
- Investigate the efficiencies of how discounts are handled, (1) by individuals, and (2) by product segments

#### Primary requirements:

- Short presentation of 5-7 minutes in length
- Focus on top-level facts
- Visualization is viewed from a great distance across the room

### 2.2. AUDIENCE PERSONAS

#### VP of Sales **Sylvia:**

- **Needs:** (1) make more effective inventory and distribution decisions with data, (2) anticipate next moves in sales based on trends, and (3) communicate with her colleagues about sales strategy
- **Characteristics:** (1) limited face time with executives, (2) not great with numbers and stats, (3) finds it hard to find what the data conveys

#### Active Board Member **Terrence:**

- **Needs:** (1) learn whether discounted sales make more or less profitable sales, (2) by how much
- **Characteristics:** (1) efficiency oriented, (2) want to end abuse of company's policies, (3) prefers data-driven decision making, and (4) poor eyesight

### 2.3. CHECKLIST

#### WHO

- **Stakeholders:** Sylvia is not good with stats but needs clear insights from data
- **Audience:** Terrence needs insights from data for data-driven decision making presented in big fonts
- **Subject-Matter Experts:** Not available with the provided personas

#### WHAT

- **Data Sources:** Provided dataset, assumed to be from a reliable source
- **Data Quality:** Huge dataset with almost 210K rows and 21 columns, seems to be good (no missing values) and contains enough data for accessing the questions being asked
- **Data Timeliness:** Unidentifiable, but the data dates back to the year 2014, so it might be timely at the point of making the analysis

#### WHY

- **Business Case/Other Goals:** Exploratory analysis on the impact and magnitude of discounts on sales
- **Intended Outcome:** Negative impacts backed up by inefficient distribution of discounts among product segments and individual customers

#### HOW

- **Format(s):** Dashboard with a big font and illustrations for the board members
- **Presentation Vehicle:** Large projection board and handouts

### 3. VARIABLE CHOICES AND KPI CONSTRUCTION

As the goals of the case focus on the exploratory analysis of the impact and magnitude of discounts on sales, the main variables are **Profit**, **Discount**, **Sales**, **Region**, **Category**, and **Sub-category**. They can be divided into two groups, the first group is regarded as Metrics, which includes variables used for calculation, and the second group is regarded as Aggregation, which includes variables used for aggregating data.

However, it is not so intuitive to work directly with the raw metrics as profit is related to sales, so they should be analyzed together. This requires some transformations for variables in the calculation group. These transformations include

- Calculating **Profit Ratio**, defined as total profit divided by total sales aggregated by aggregation variables
- Forming **Profit Ratio KPI** as follows: High PR when Profit Ratio > 33%, Low PR when 33% > Profit Ratio > 0%, and Negative when Profit Ratio < 0%.

The variable **Discount** can be aggregated in two ways: either by counting or averaging, respectively represent for the amount of discounting coupons received and the average discount rate. These distinguished aggregations generate two different variables, whose effects to Profit Ratio might be different.

### 4. ANALYSIS DETAILS

### 5. STORY CONSTRUCTION

The story is simple:

*"There are many abnormal sales whose profit is negative despite their large magnitude, which calls for an exploratory analysis to figure out the cause. The analysis results show that discounts are distributed unevenly among customers and product categories, but more discounts do not necessarily lead to higher profit. Higher discount rates worsen the profit ratio while the discount coupons only work on customers with negative profit. The suggestions are that we should give less discount coupons and lower the discount rate to customers with positive profit while giving more discount coupons with lower discount rate to the customers with loss."*

The key messages here are:

- Discount coupons do not work with customers whose profit is positive
- Higher discount rates worsen profit in all cases

Hence, the story is constructed surrounding these two messages. The story being told follows a three-part structure: Problem, Investigation, and Solution, so it will have around 5 story points to convey all the findings, which is appropriate for a short presentation of 5 ~ 7 minutes.

### 6. DESIGN CHOICES

The consideration focuses on many factors:

- **Font size:** as large as necessary as the audience needs to grasp the information quickly
- **Color:** Main colors include Green, Yellow, and Red, which are also the colors assigned for the three categories of the implemented KPI. These colors are chosen due as they are intuitive for the audience to understand the magnitude and also attract their attention to what's important in visualization
- **Alignment:** Important information is aligned to the top left of all visualizations as the audience reads from left to right and top to bottom
- **Pre-attentive attributes and Gestalt principles:** Color, Enclosure, Proximity, and Similarity
- **Types of visualization:** Simple text is used to convey message to the audience, barchart is used for aggregating discounts among sub-categories, treemap is used to visualize the magnitude of discounts (counts and average rates) among customers, and scatter plot is used to present the correlation between Profit Ratio and Discount aggregations. The choices of barchart

and scatter plot are intuitively conventional, and treemap is suitable as it presents the distribution of discounts on ALL customers while showing the relative magnitude of discount among them

**Highlights:** key messages are highlighted to reduce the clutter and processing time of the audience

## 7. HANDOUTS AND FINAL VISUALIZATION

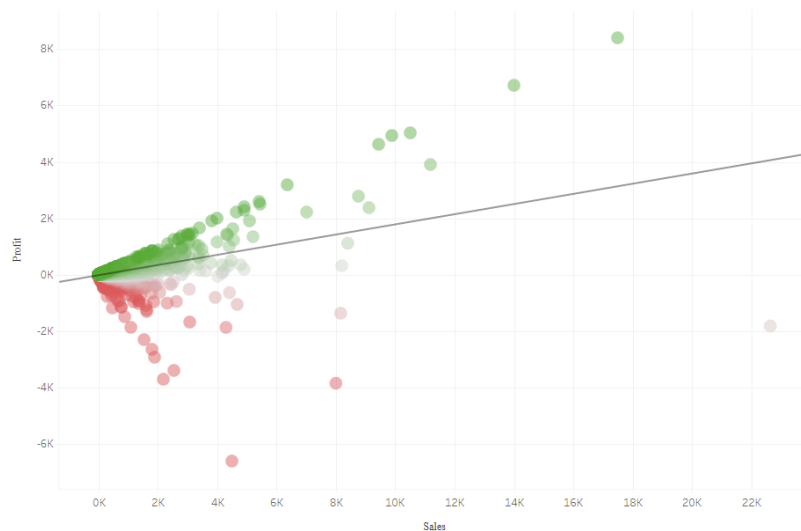
The one-page handout for the audience can be downloaded from the link below:

[Handout Page](#)

The full story can be found on Tableau public by following the link below

[An introductory analysis to the impact of discount on profit](#)

Profit and Sales Correlation



In the next **six minutes...**

### OUR GOAL:

**1** Understand **how discounts are distributed among customers and product categories**

**2** Use this knowledge to **investigate the effect of discounts on sales profit**

We will end with a **specific recommendation**

## 8. FINAL VISUALIZATION ANALYSIS

## 9. RELATED INFORMATION AND DATA

### 9.1. DATASET

[Superstore Dataset](#)

### 9.2. CHANGE LOG

- **02/03/2022:** v0.1 main contents established
- **02/03/2022:** v0.2 correct some typos
- **03/03/2022:** v0.3 correct some typos, add more details to the explanation, and adjust the structure of the document
- **03/03/2022:** v0.4 structure update

### 9.3. WHAT TO ADD

- Story point design discussion
- Analysis process
- Appendix

## APPENDIX AND REFERENCES

APPENDIX

REFERENCES