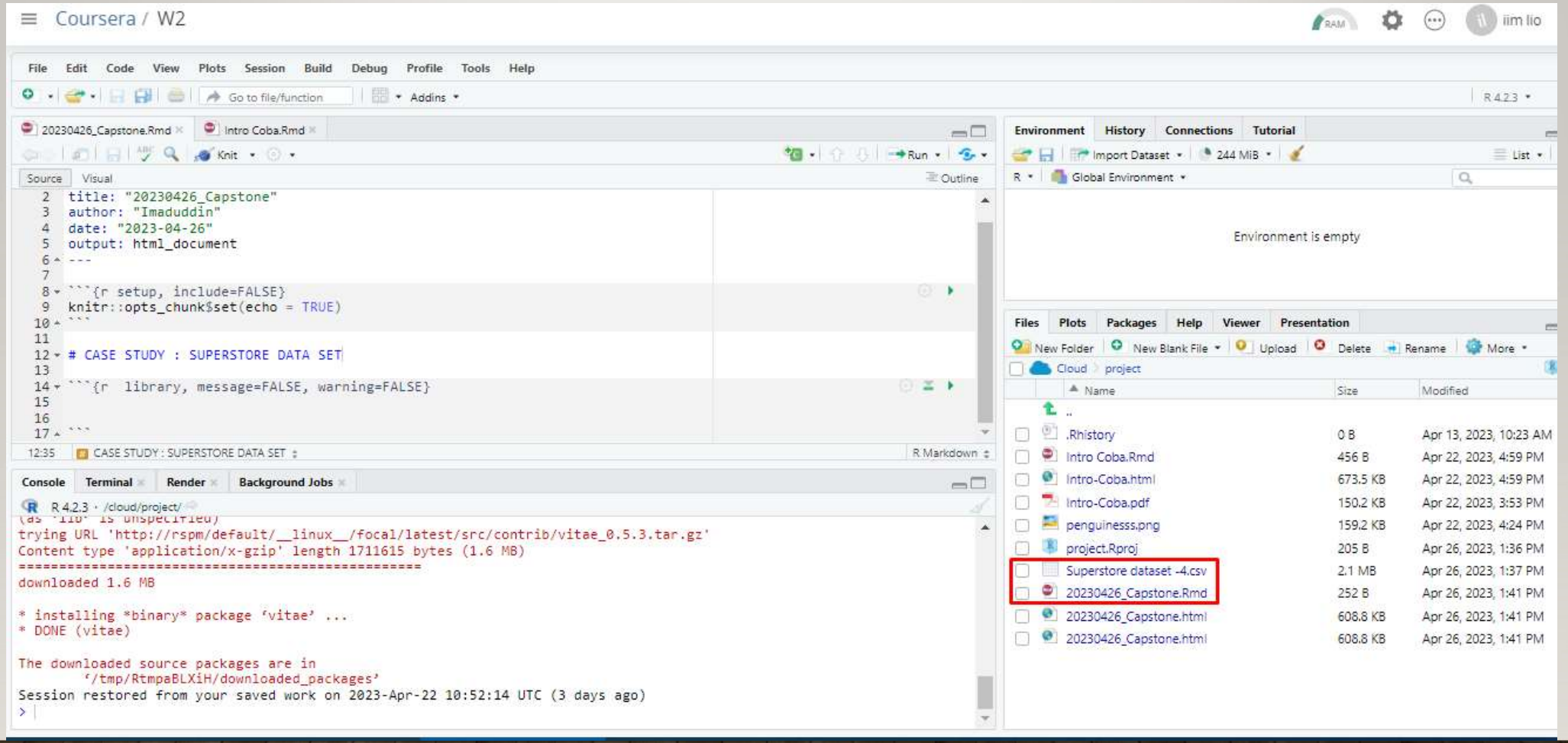
PROFIT & LOSS STORY

'IMADUDDIN / APRIL 2ND 2023

ASK – PREPARE – PROCESS – ANALYZE – ACT / SHARE

- Task : Finding the truth based on excessive discount on certain product and its effects on revenue
- Prepare : Superstore dataset uploaded via R.Studio, Creating Rmd file
- Process & Analyze : via R.Studio
- Act & Share : via Tableau

PREPARE


Coursera / W2
File Edit Code View Plots Session Build Debug Profile Tools Help
20230426_Capstone.Rmd
Intro Coba.Rmd
Knit
Run
Source Visual
Outline
2 title: "20230426_Capstone"
3 author: "Imaduddin"
4 date: "2023-04-26"
5 output: html_document
6 ---
8 ```{r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = TRUE)
10 ```
12 # CASE STUDY : SUPERSTORE DATA SET
14 ```{r library, message=FALSE, warning=FALSE}
17 ```
12:35 CASE STUDY : SUPERSTORE DATA SET
R Markdown
Console Terminal Render Background Jobs
R 4.2.3 · /cloud/project/
trying URL 'http://rspm/default/__linux__/focal/latest/src/contrib/vitae_0.5.3.tar.gz'
Content type 'application/x-gzip' length 1711615 bytes (1.6 MB)
downloaded 1.6 MB
* installing *binary* package 'vitae' ...
* DONE (vitae)
The downloaded source packages are in
'/tmp/RtmpaBLXiH/downloaded_packages'
Session restored from your saved work on 2023-Apr-22 10:52:14 UTC (3 days ago)
Environment History Connections Tutorial
Import Dataset
Global Environment
Environment is empty
Files Plots Packages Help Viewer Presentation
New Folder
New Blank File
Upload
Delete
Rename
More
Cloud project
Name Size Modified
.Rhistory 0 B Apr 13, 2023, 10:23 AM
Intro Coba.Rmd 456 B Apr 22, 2023, 4:59 PM
Intro-Coba.html 673.5 KB Apr 22, 2023, 4:59 PM
Intro-Coba.pdf 150.2 KB Apr 22, 2023, 3:53 PM
penguinesss.png 159.2 KB Apr 22, 2023, 4:24 PM
project.Rproj 205 B Apr 26, 2023, 1:36 PM
Superstore dataset -4.csv 2.1 MB Apr 26, 2023, 1:37 PM
20230426_Capstone.Rmd 252 B Apr 26, 2023, 1:41 PM
20230426_Capstone.html 608.8 KB Apr 26, 2023, 1:41 PM
20230426_Capstone.html 608.8 KB Apr 26, 2023, 1:41 PM

PROCESS & ANALYZE

HTTPS://2AD61D15183E4D2B9D044A32A52658DF.APP.POSIT.CLOUD/?VIEW=RMARKDOWN

```
## DATA INSPECTION & RECAP
```{r check str summary, message=FALSE, warning=FALSE}
str(superstore)
summary(superstore)
```


CLEANING
```{r}
superstore_clean <- mutate(superstore,"Postal Code"=as.character("Postal Code"))
superstore_clean$"Order Date"<- strptime(as.character(superstore_clean$"Order Date"), "%d/%m/%Y")
superstore_clean$"Ship Date"<- strptime(as.character(superstore_clean$"Ship Date"), "%d/%m/%Y")
superstore_clean$Discount <- as.numeric(gsub('\\$|,', '', superstore_clean$Discount))
superstore_clean$"Discount to Sales Ratio" <- as.numeric(gsub('\\%|,', '', superstore_clean$"Discount to Sales Ratio"))
superstore_clean$Profit <- as.numeric(gsub('\\$|,', '', superstore_clean$Profit))
superstore_clean$"Profit Ratio" <- as.numeric(gsub('\\%|,', '', superstore_clean$"Profit Ratio"))
superstore_clean$Sales <- as.numeric(gsub('\\$|,', '', superstore_clean$Sales))

#rename column
superstore_clean <- superstore_clean %>%
  rename(postal_code="Postal Code", profit_ratio="Profit Ratio",discount_to_sales_ratio="Discount to Sales Ratio",order_date="Order Date",ship_date="Ship Date")

colnames(superstore_clean)
str(superstore_clean)
summary(superstore_clean)
skim(superstore_clean)
```

```
   ship_date                     Ship Mode            State
 Min.   :2011-01-02 00:00:00.00   Length:9994        Length:9994
 1st Qu.:2012-04-11 00:00:00.00   Class :character   Class :character
 Median :2013-05-11 00:00:00.00   Mode  :character   Mode  :character
 Mean   :2013-04-05 12:37:52.50
 3rd Qu.:2014-04-10 18:00:00.00
 Max.   :2015-06-01 00:00:00.00
 NA's   :6116
 Sub-Category          Discount       discount_to_sales_ratio Number of Records
 Length:9994        Min.   :0.00000   Min.   : 0.000          Min.   :1
 Class :character   1st Qu.:0.00000   1st Qu.: 0.000          1st Qu.:1
 Mode  :character   Median :0.00000   Median : 0.020          Median :1
                    Mean   :0.09226   Mean   : 1.359          Mean   :1
                    3rd Qu.:0.00000   3rd Qu.: 0.470          3rd Qu.:1
                    Max.   :1.00000   Max.   :80.000          Max.   :1
                                      NA's   :1
     Profit         profit_ratio        Quantity         Sales
 Min.   :-6599.98   Min.   :-275.00   Min.   : 1.00   Min.   :    0.0
 1st Qu.:    1.73   1st Qu.:   7.00   1st Qu.: 2.00   1st Qu.:   17.0
 Median :    8.67   Median :  27.00   Median : 3.00   Median :   54.5
 Mean   :   28.66   Mean   :  12.05   Mean   : 3.79   Mean   :  229.9
 3rd Qu.:   29.36   3rd Qu.:  36.00   3rd Qu.: 5.00   3rd Qu.:  210.0
 Max.   : 8399.98   Max.   :  50.00   Max.   :14.00   Max.   :22638.0
 NA's   :1865
```

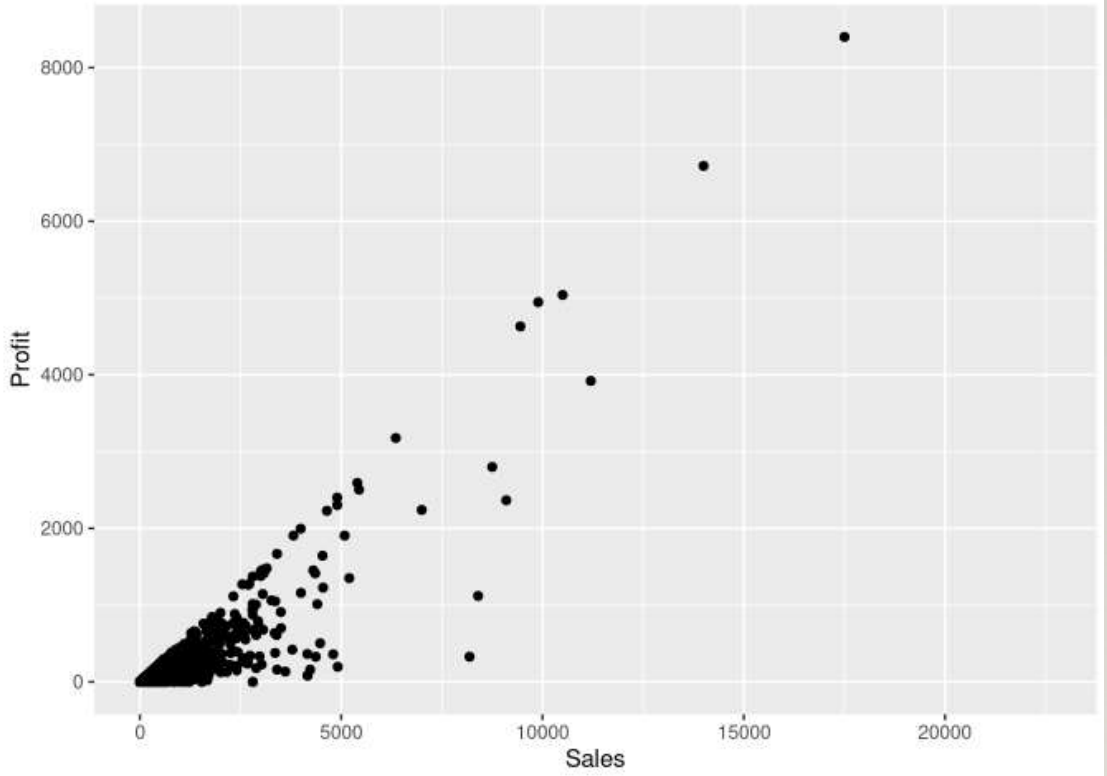
Variable type: numeric

skim_variable	n_missing	complete_rate	mean	sd	p0	p25	p50	p75	p100	hist
Discount	0	1.00	0.09	0.29	0	0	0.00	0.00	1	▇▁▁▁▁
discount_to_sales_ratio	1	1.00	1.36	5.66	0	0	0.02	0.47	80	▇▁▁▁▁
Number of Records	0	1.00	1.00	0.00	1	1	1.00	1.00	1	▁▁▇▁▁
Profit	1865	0.81	54.43	212.83	0	5	13.00	40.00	8400	▇▁▁▁▁
profit_ratio	0	1.00	12.05	46.69	-275	7	27.00	36.00	50	▁▁▁▁▇
Quantity	0	1.00	3.79	2.23	1	2	3.00	5.00	14	▇▇▃▁▁
Sales	0	1.00	229.87	623.25	0	17	54.50	210.00	22638	▇▁▁▁▁
```

ACT & SHARE

SCATTER PLOT

```
ggplot(data = superstore_clean) +
  geom_point(mapping = aes(x = Sales, y = Profit))
```



SHARE VIA TABLEAU

- https://bit.ly/story_20230401_DW4
- 3 specific choice :
  - The overview : customer all over the region/state with each sales, profit, discount on product categories
  - Examining the geographical factor sales that clearing the ambiguity and avoid false narrative
  - Finding the truth based on excessive discount on certain product
- My recommendation are specifics (last page)

Profit and Loss Story

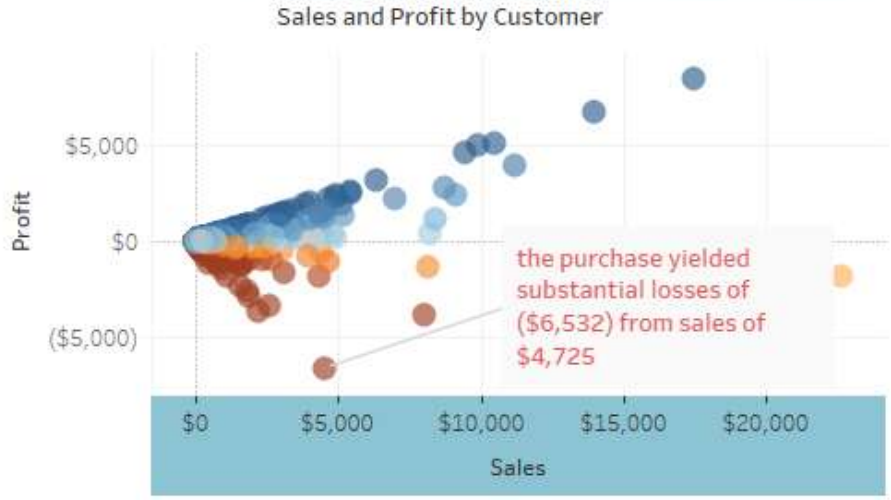
Introduction | **All Customers** | Overview, Selected Customer, Profit and Discount | Customer by State | Discount & Profit Ratio Dashboard

This section examines all of the customer. Go to next section for details why customers' purchase lead to a loss while most seem to make a profit

Year (All) Segment (All) Category (All)

Region	Sales per Cust	Sales	Quantity	Discount	Profit	Profit Ratio
Central				$558.34		
East	$1,007	$678,781	10,618		$91,523	13.48%
South				$350.20		
West	$1,058	$725,458	12,266		$108,418	14.94%

Profit Ratio -50.00% 50.00%



Profit and Loss Story

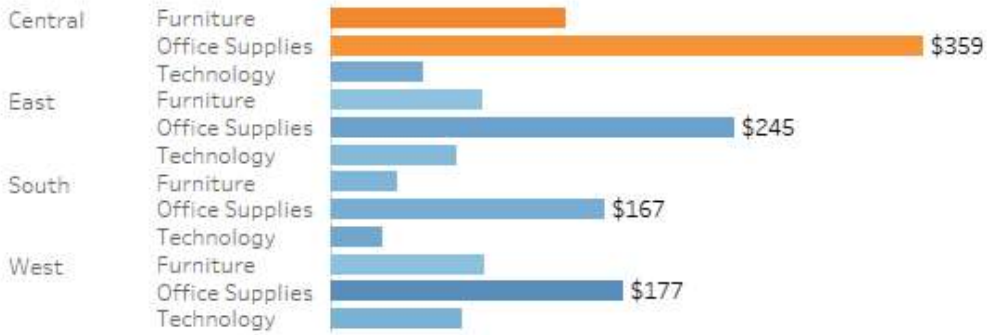
Introduction | All Customers | Overview, Selected Customer, Profit and Discount | Customer by State | Discount & Profit Ratio Dashboard

Year (A... Region (All) Segment (All) Category (All) Sub-Catego.. (All) Avg. Profit .. -50.00% 50.00%

Avg. Profit .. -50.00% 50.00%

Profit Ratio by State

Discount per Region and Category



Discount by Segment and Category

Profit and Loss Story

Introduction | All Customers | Overview, Selected Customer, Profit and Discount | **Customer by State** | Discount & Profit Ratio Dashboard

Year of Ord.. (All) Segment (All) Category (All) Customer .. (Multiple values) Sub-Catego.. (All)

There is no obvious evidence on state-level difference in profit between who results loss and profit

Profit Ratio -50.00% 50.00%

State	Category	Seth Vernon	Ted Trevino	Michelle Moray	Caroline Jumper	Tom Prescott
		Customer Name				
California	Furniture					
	Office Supplies			28.00%		
Colorado	Furniture			-170.00%		
	Office Supplies			-74.42%		
Delaware	Furniture					
Florida	Furniture					
	Office Supplies					
Illinois	Furniture			-50.80%		
	Office Supplies			9.38%		
Louisiana	Furniture					
Montana	Office Supplies			28.00%		
	Technology			31.13%		
New York	Furniture			13.33%		
	Office Supplies			43.73%		
	Technology			21.00%		

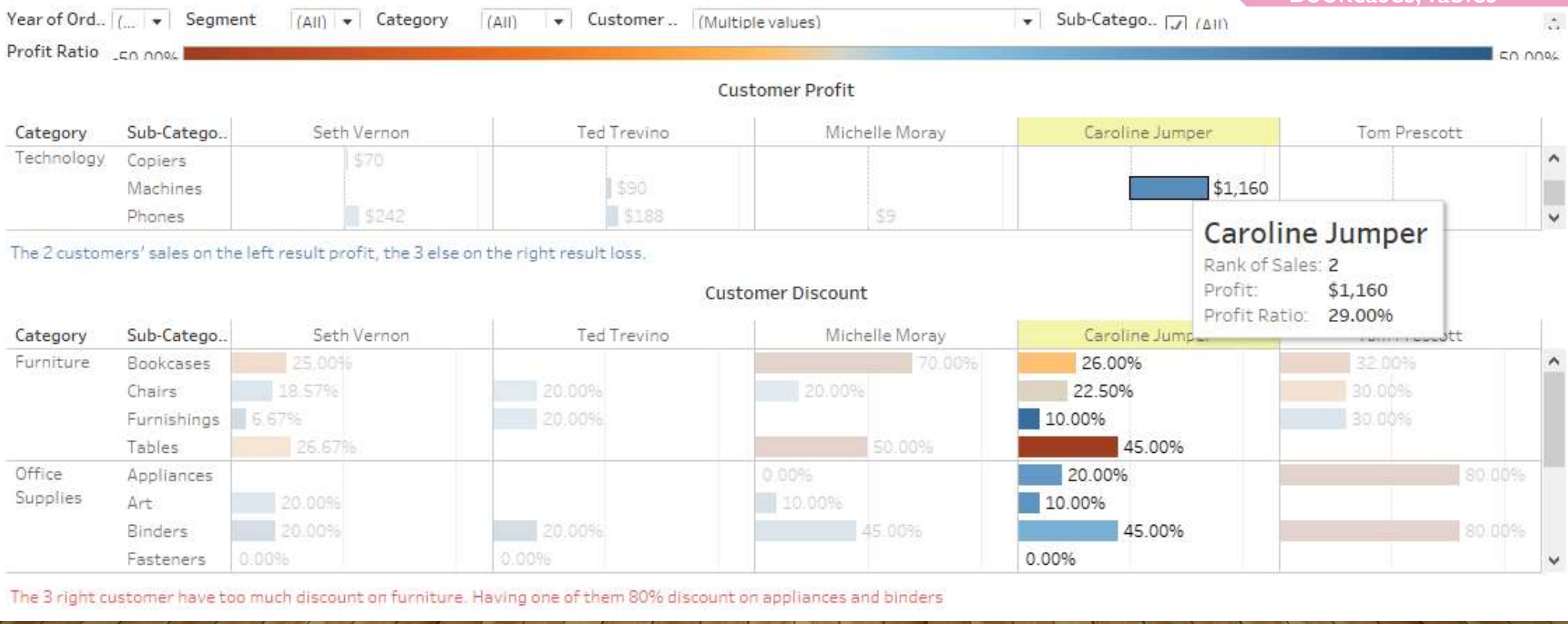
Michelle Moray
Rank of Sales:
Profit: ($248)
Sales: $146
Profit Ratio: -170.00%
Avg Discount: 70.00%

Yet, it's not significant on geography scale (State-Region). So it's not the problem

Profit and Loss Story

Introduction | All Customers | Overview, Selected Customer, Profit and Discount | Customer by State | Discount & Profit Ratio Dashboard

The pain point is the discount on certain product sub-category : Appliances, Binders, Bookcases, Tables



RECOMMENDATION

- State-Region is not the problem based on the pattern even there're significant outliers
- Sales on Central and West Region need to be boosted
- Discount amount on Furnitures (esp. Tables & Bookcases) and also Office Supplies (esp. Binders & Appliances) need to be evaluated to maintain profit