

Title Page

DEA 111: Visual Analysis for Business Intelligence (BI) with  
Tableau  
(Final Project Report)

-----

Profitability Intelligence and Customer Segmentation

Safak Atakan Celik  
CT1009470

## INTRODUCTION

This report follows the Final Report Structure outlined in the project guidelines. There are no separate Phase 2 and Phase 3 uploads. Phase 1 has already been uploaded before. Here are phase 2 and phase 3 deliverables linked to the sections in this report:

- .twbx file: uploaded separately
- Data preparation: has a section called Data Preparation
- Filtering, sorting, grouping: worksheets allow filtering and sorting, and some of them are already filtered and sorted. There's manual grouping of views, but there's no grouping as in creating a group in Tableau.
- LOD calculations and KPIs: in the section called Key Calculations and KPIs
- Interactive charts: The .twbx file contains many interactive charts.
- Dashboards: in the .twbx file and explained in the section called Dashboard Design
- Stories: in the worksheets and explained in the section called Storytelling Flow
- Final report: entirety of this report

The objective of this project is to analyze a store's data to gain information into their profitability across different dimensions and measures, and provide actionable insights to help improve the business. The analysis starts from the customer segment level, and drills down to categories, sub-categories, and also provides insight on individual customers' profitability by using the measures in the dataset and joining a separate dataset segmenting customers further.

## DATA PREPARATION

The dataset used is the Superstore dataset. I prepared it with Tableau Prep combining the components of it, and created a newer version with a Python notebook.

The screenshot shows the Tableau Data Source interface. In the top left, there are two connections: "orders\_oct27\_assigned\_unique\_productIDs" (Text file) and "customers\_df\_clustered(purchasefreq-lifetime-against\_churn\_rate)" (Text file). Below the connections, under "Files", there are two files listed: "customers\_df...rn.rate.csv" and "orders\_oct27...oductIDs.csv". A "New Union" button is also present. On the right, a preview of the "orders\_oct27\_assigned\_unique\_productIDs.csv" file is shown, indicating it has 35 fields and 9952 rows. The preview table includes columns for Name, Type, Field Name, Physical Table, and Remarks. The first few rows show data for field F1.

Name	Type	Field Name	Physical Table	Remarks
#	F1	orders_oct27_assigned_unique_produc	orders_oct27_assigned_unique_produc	Is Duplicate Row?
1				Days to Ship
2				
3				
4				

<https://github.com/safakatakancelik/portfolio-public/tree/main/notebooks/data/datasets/Superstore>

files:

- orders\_oct27\_assigned\_unique\_productIDs.csv
- customers\_df\_clustered(purchasefreq-lifetime-against\_churn\_rate).csv

More detailed steps about the entire dataset are outlined in the README.md in the repository. Because I only used orders, and the clustered customers dataframe I created, I'm only outlining them in this document:

Orders Central	<ul style="list-style-type: none"> <li>- Created the “Region” calculated field</li> <li>- Created the “Order Date” and “Ship Date” calculated fields and removed partial date fields(order year, order month, order day, ...)</li> <li>- Replaced None values with 0 in “Discounts” field, and changed its data type to Number (decimal)</li> </ul>
Orders West	<ul style="list-style-type: none"> <li>- Removed the duplicate fields with the prefix “Right_” (20 fields in total)</li> <li>- Grouped values of states with their abbreviations.</li> </ul>
Orders East	<ul style="list-style-type: none"> <li>- Removed letters “USD” from Sales numbers, and changed its data type to Number (decimal)</li> </ul>
Orders South	<ul style="list-style-type: none"> <li>- None, they were following the same format and unioned directly.</li> </ul>

	- Unioned all
	- Removed duplicate records (16302 to 9952)

The prepared data is the same as the one I prepared in the MT project, however, there's an additional step of assigning new unique product IDs to each product ID - product name pair due to the conflict I found during the MT project. Earlier and current versions of the dataset along with the reasoning and the actual processes can be found in the repository.

## KEY CALCULATIONS and KPIs

Calculation Name	Calculation Formula
Fixed Segment Avg Profit	{ FIXED [Segment]: AVG([Profit])}
Fixed Cat Avg Profit	{ FIXED [Category]: AVG([Profit])}
Fixed SubCat Avg Profit	{ FIXED [Sub-Category]: AVG([Profit])}
Segment Avg Profit Deviation	AVG([Profit]) - AVG([Fixed Segment Avg Profit])
Cat Avg Profit Deviation	AVG([Profit]) - AVG([Fixed Cat Avg Profit])
SubCat Avg Profit Deviation	AVG([Profit]) - AVG([Fixed SubCat Avg Profit])
Profit Margin	SUM([Profit]) / SUM([Sales])

There are 3 main calculations.

- Profit Margin is an overall calculation that shows different values based on views and filters that are in effect, and provides information on the portion of sales that yield to profits.
- Fixed Avg Profit calculations provide average profits fixed to the related levels.
- Avg Profit Deviation calculations provide level of detail (LOD) analysis values. Based on filters applied on the view to select a sample, the values show the sample's deviation from the population(company wide average profits) mean.

KPIs in this analysis are these calculated fields along with Sales and Profits. These are used to assess profitability of different dimensions to provide insights.

## DASHBOARD DESIGN

Within this analysis I've applied a simpler approach to focus better. Instead of many different visuals with custom filters and parameters, the focus is on a few views that are repeated across different dimensions. However, there's also some variety as well.

There are 5 dashboards used in 5 stories that are the product of 15 worksheets.

Dashboards:

KPI Summary Over 4 Years	<ul style="list-style-type: none"><li>- Line chart showing actual profit, sales, and profit margins over 4 years</li><li>- Line chart showing YoY growth rate for profit, sales, and profit margins over 4 years.</li></ul>
Segment Profitability View	<ul style="list-style-type: none"><li>- Line chart showing actual profit margins of segments</li><li>- Bar chart showing actual sales and profits along with percentages over their respective measures (e.g. consumer segment's profits over all segments' profits)</li><li>- Bar chart showing the deviations of average profits based on the applied filters.</li></ul>
Category Profitability View	<ul style="list-style-type: none"><li>- Same as Segment Profitability view but applied on the Category dimension</li></ul>
Sub-Category Profitability View	<ul style="list-style-type: none"><li>- Same as Segment Profitability view but applied on the Sub-Category dimension</li></ul>
Customer Profitability View and Segmentation	<ul style="list-style-type: none"><li>- Scatter plot of all customers with dimensions Profit and Sales with a hue of positive/negative profit ratio.</li><li>- Scatter plot of analyzed customers with dimensions Customer Lifetime (days) and Purchase Frequency with a hue of assigned clusters.</li><li>- Bar chart showing the distribution of customers within each cluster with a hue of churned boolean dimension.</li></ul>

The main filter provided and applied is a filter of date (aggregated into year), because this project is focused on understanding the decline in the last year by analyzing all the data and changes over the years.

The last dashboard Customer Profitability View and Segmentation is a product of ongoing work I'm doing for the class DEA 112. There I used the same dataset. While it does not provide the main takeaway of this analysis, I've included it here because it is related and can be useful. Within this dashboard I allow filtering with clusters and the churned dimension.

Color coding: blue: sales, green: profits, red: loss, any other color: a unique dimension

## STORYTELLING FLOW

I created 5 stories using the 5 dashboards I created, and aimed to tell one comprehensive story that contains all 5 stories using the STAR (Situation, Task, Action, Result) framework.

Story Name	Link to the STAR framework
4 Year Overview	Situation + Task
Segment Profitability Analysis	Action + Result
Category Profitability Analysis	Action + Result
Sub-Category Profitability Analysis	Action + Result
Customer Profitability and Segmentation Analysis	Action + Result

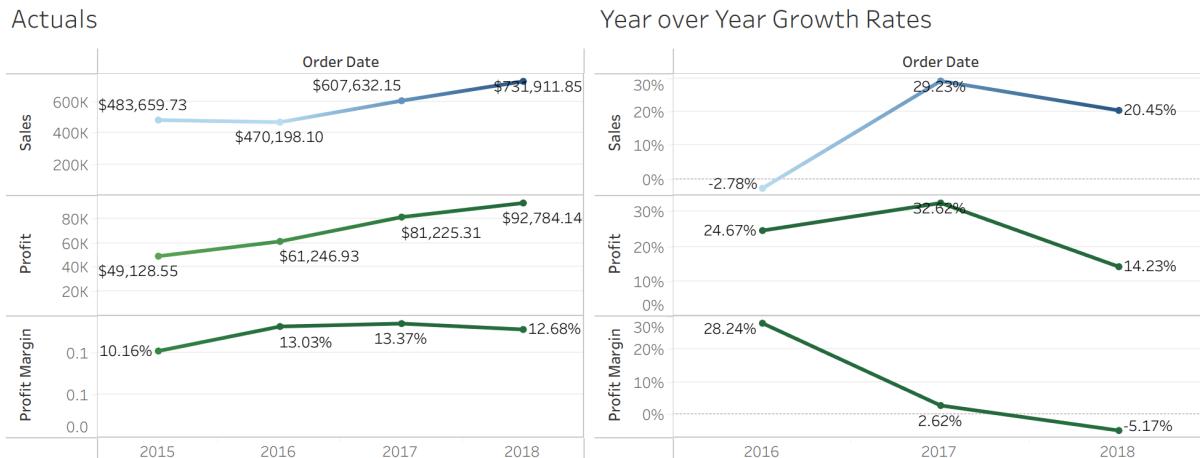
The 4 Year Overview story assesses the situation the business is in currently, and points to a task that needs to be done in order to help the business. Each of the remaining stories essentially works on this task in action and provides insights as a result. The conclusion section of this report will serve as the main result combining the key insights and recommendations for the business.

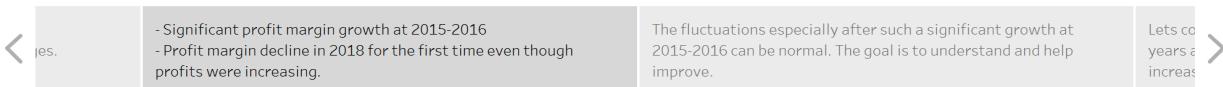
### Story: 4 Year Overview

#### 4 Year Overview



### KPI Summary Over 4 Years



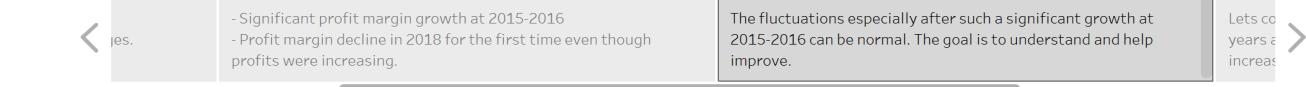


## KPI Summary Over 4 Years

Actuals

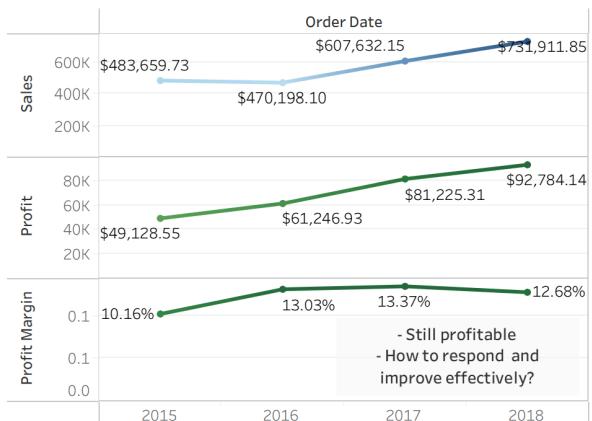


Year over Year Growth Rates

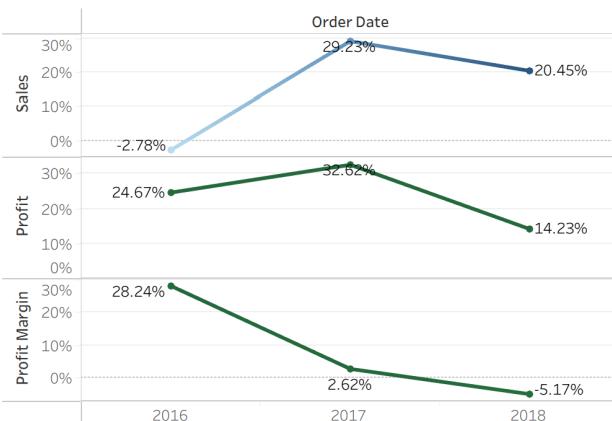


## KPI Summary Over 4 Years

Actuals



Year over Year Growth Rates

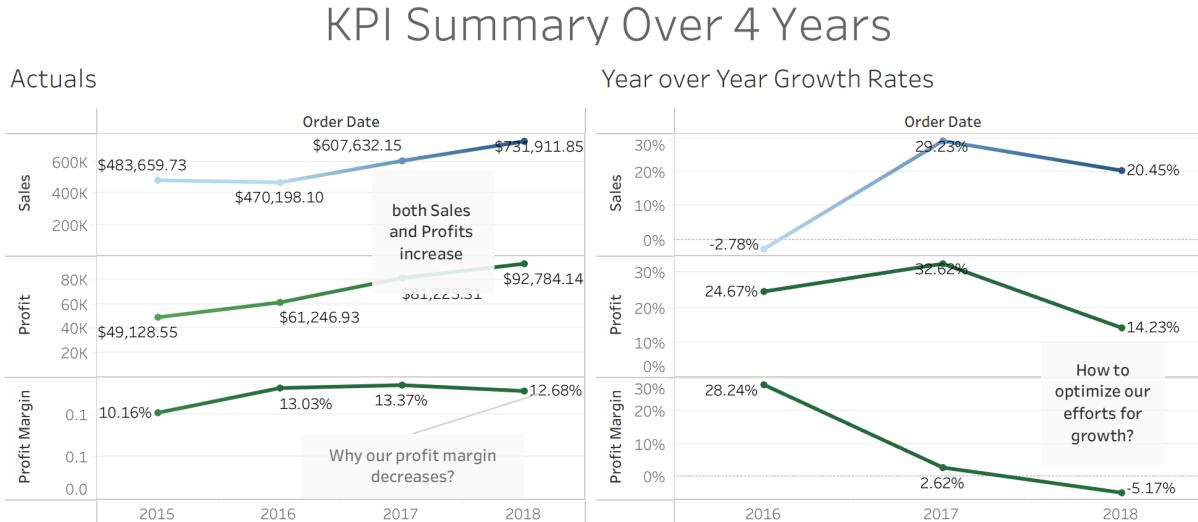


- Still profitable  
- How to respond and improve effectively?

16  
ne even though

The fluctuations especially after such a significant growth at 2015-2016 can be normal. The goal is to understand and help improve.

Lets compare various measures across dimensions over these 4 years and learn: How to recover and improve?

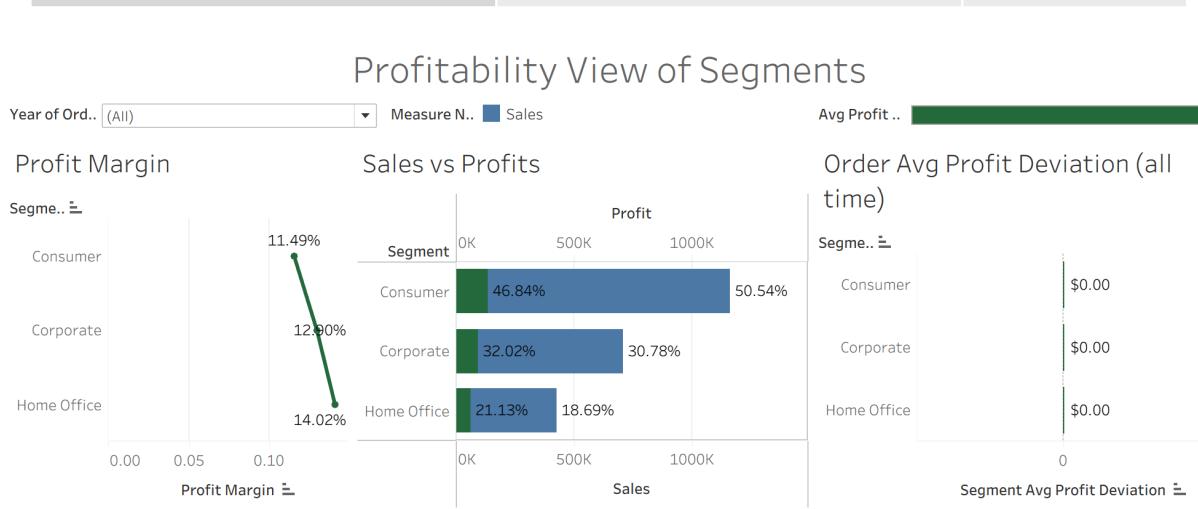


### Story: Segment Profitability Analysis

Overall our profits margins are higher in segments we have less sales, and lower in segments we have higher sales.

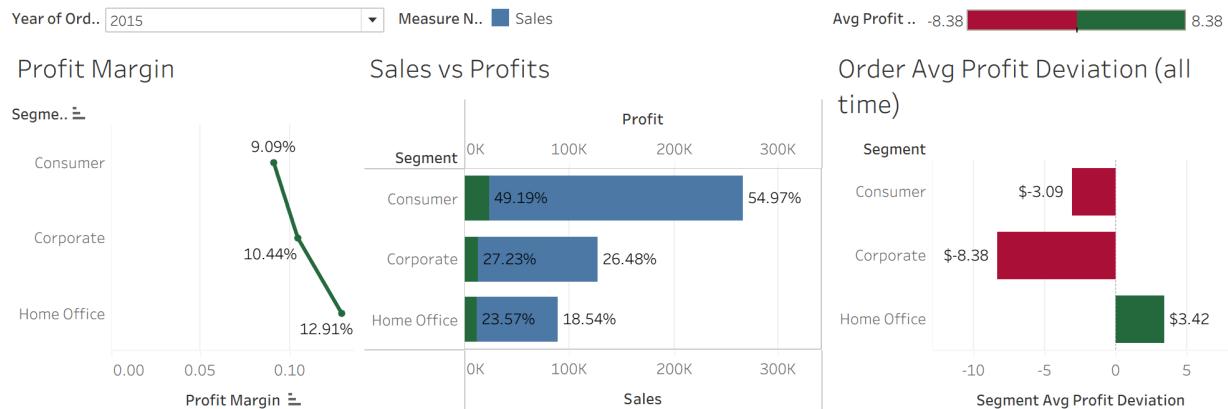
2015: Corporate avg profits are \$-8.38 lower than all years. Lowest profit margins, yet follows a similar pattern across segments over years.

2016: Corporate avg profits are \$-8.38 lower than all years. Lowest profit margins, yet follows a similar pattern across segments over years.

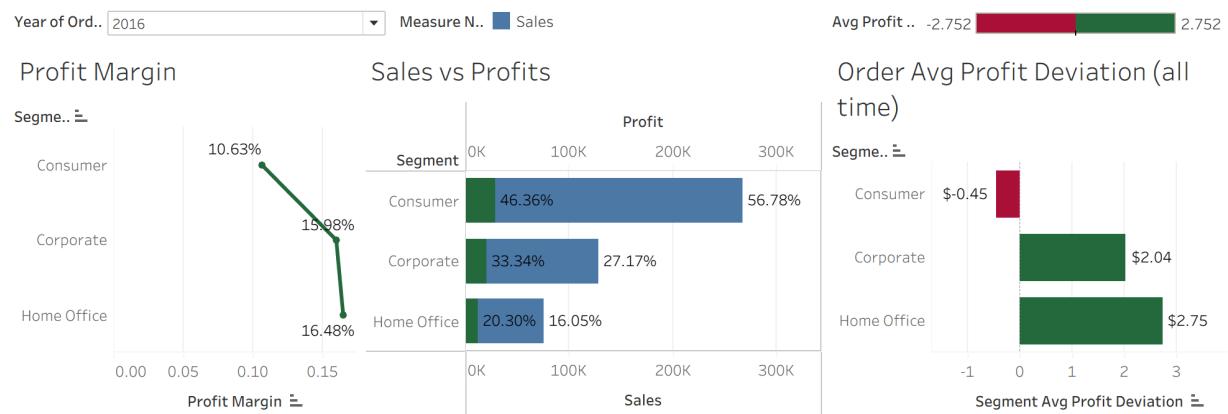


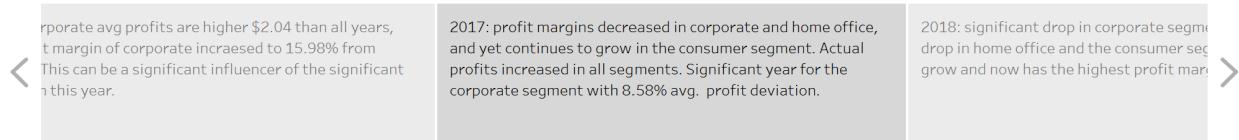


## Profitability View of Segments

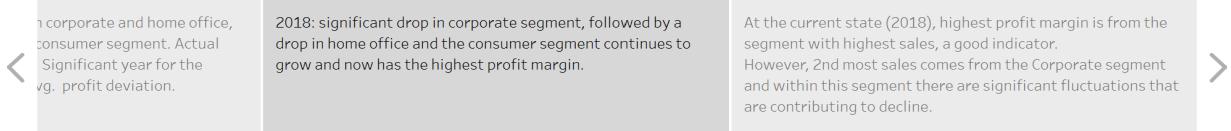
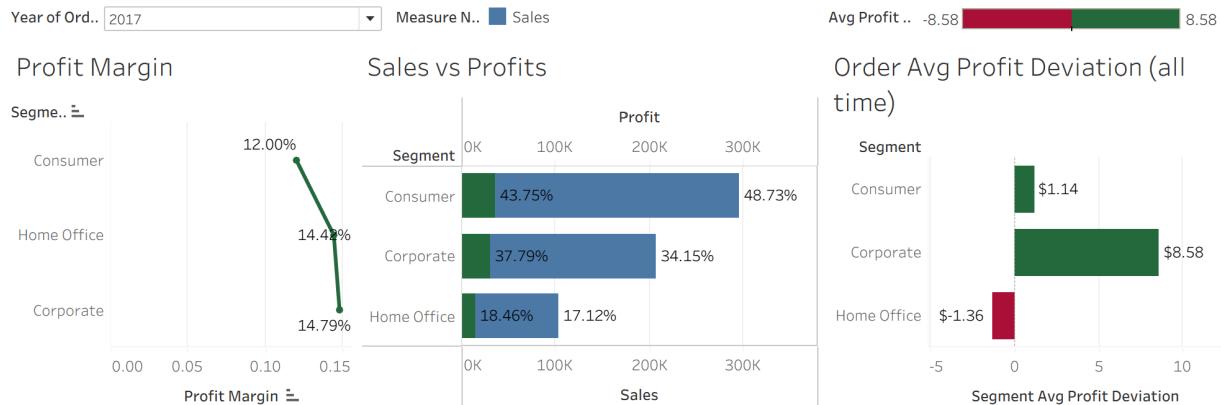


## Profitability View of Segments

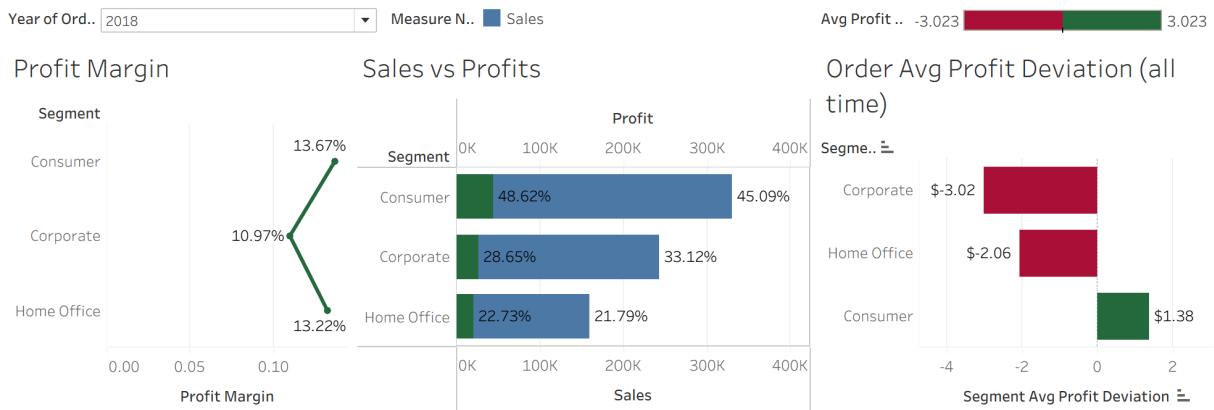


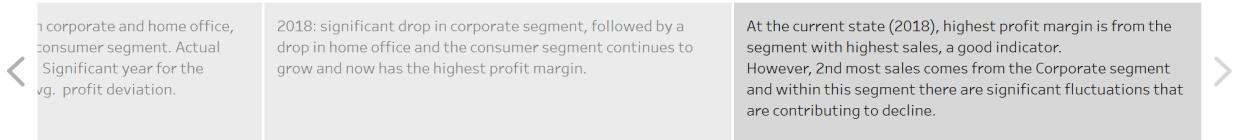


## Profitability View of Segments

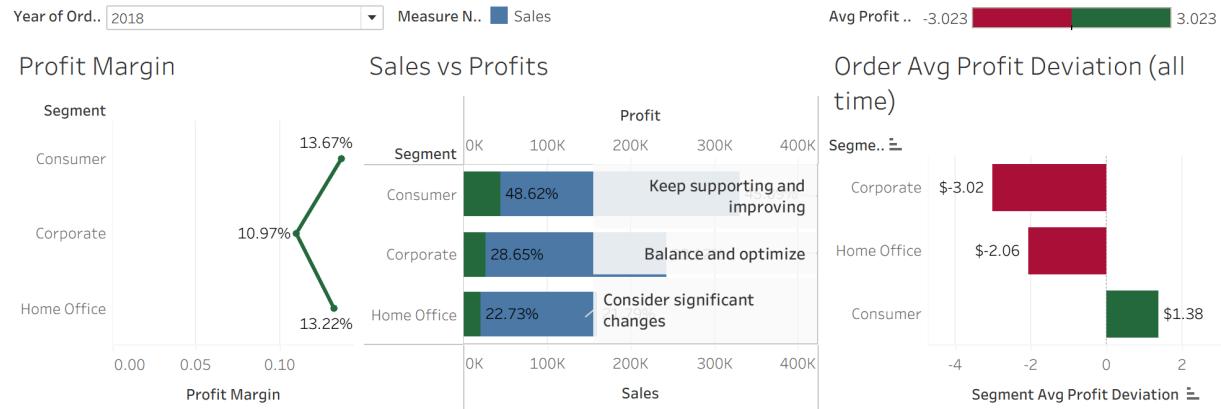


## Profitability View of Segments





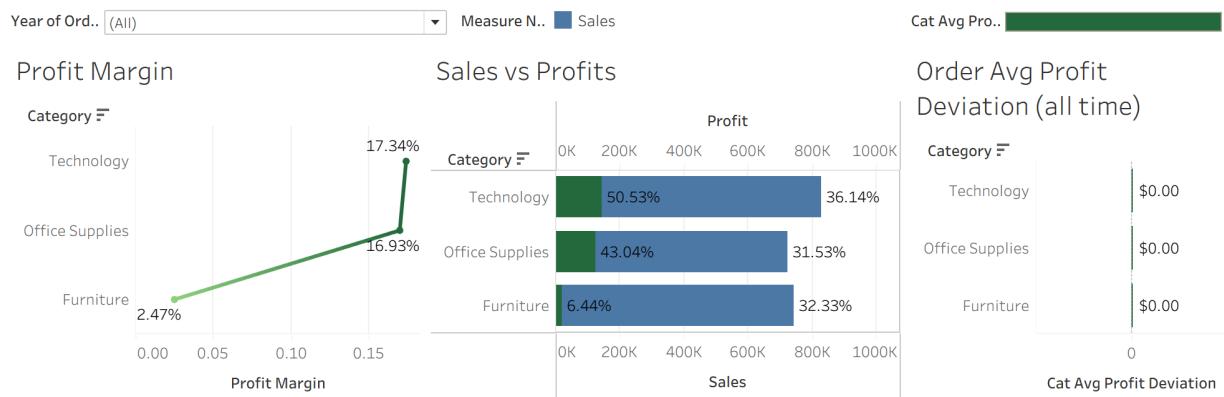
## Profitability View of Segments



## Story: Category Profitability Analysis



## Profitability View of Categories



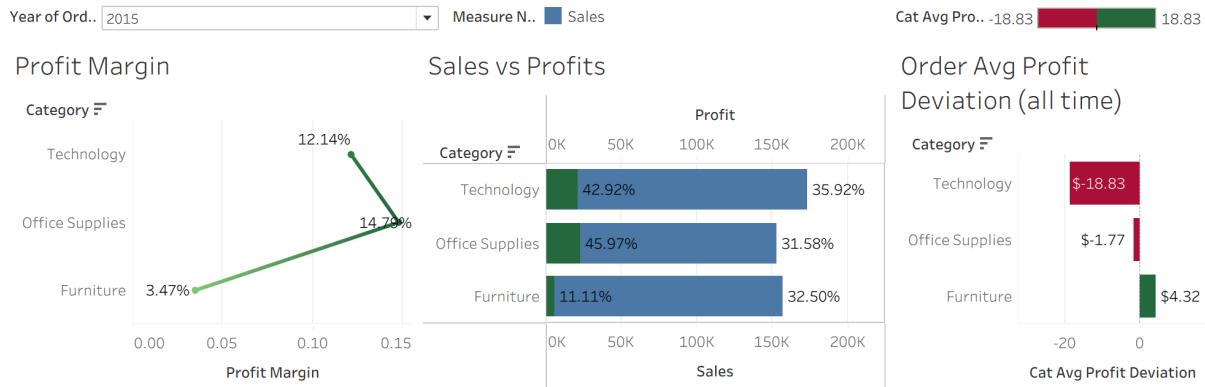
t year to year and aligns with the overall view.

2015 is the worst performant year in these 4 years. In comparison the only upside of the year 2015 is its performance in furniture category, however, this is our worst performing category.

Furniture category has 32.33% 6.44% of our profits. Additionally profit margins. We can consider



## Profitability View of Categories



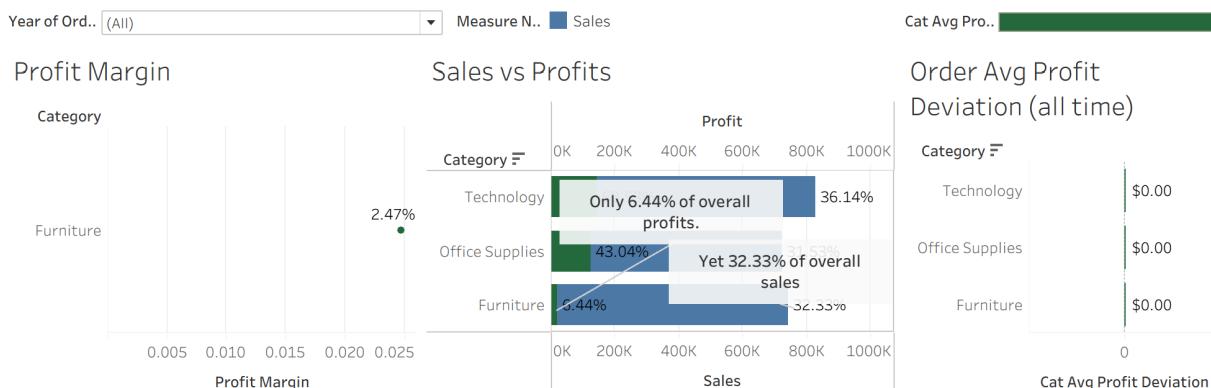
2015 is the worst performant year in these 4 years. In comparison the only upside of the year 2015 is its performance in furniture category, however, this is our worst performing category.

Furniture category has 32.33% of our sales while the category holds the 6.44% of our profits. Additionally, the category has significantly lower profit margins. (a problem to investigate)

2017 h terms

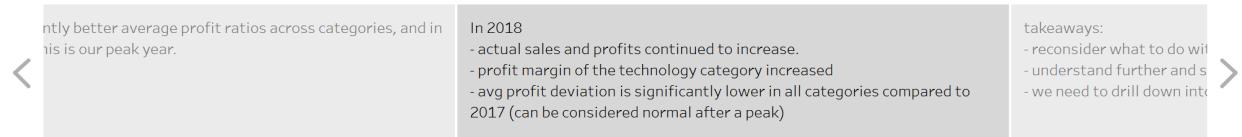
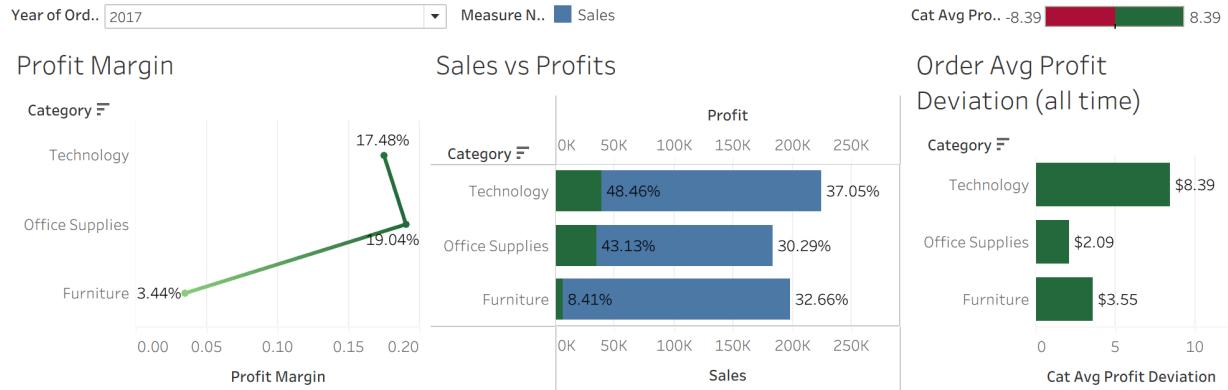


## Profitability View of Categories

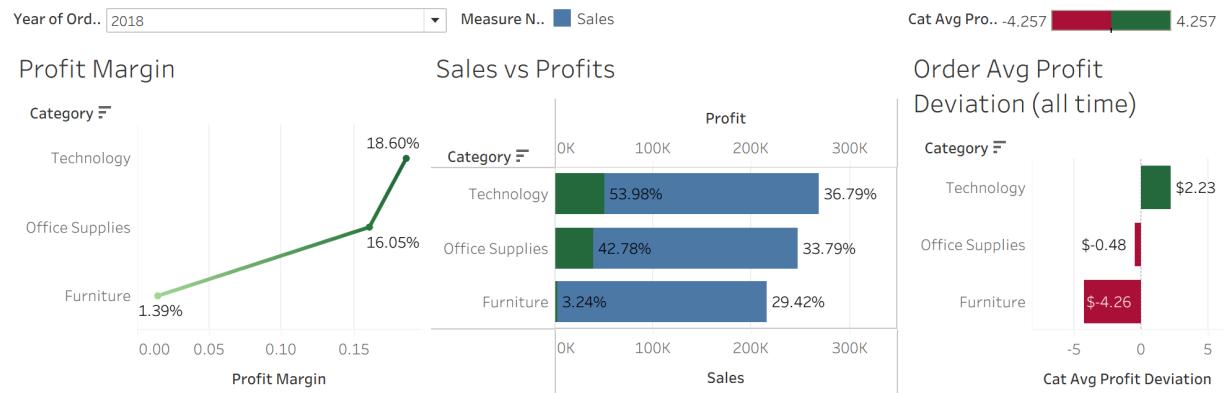




## Profitability View of Categories

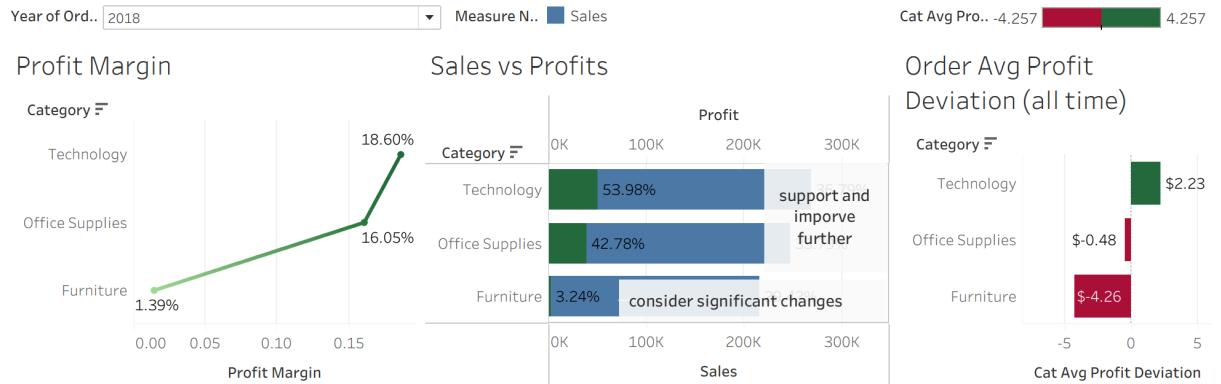


## Profitability View of Categories





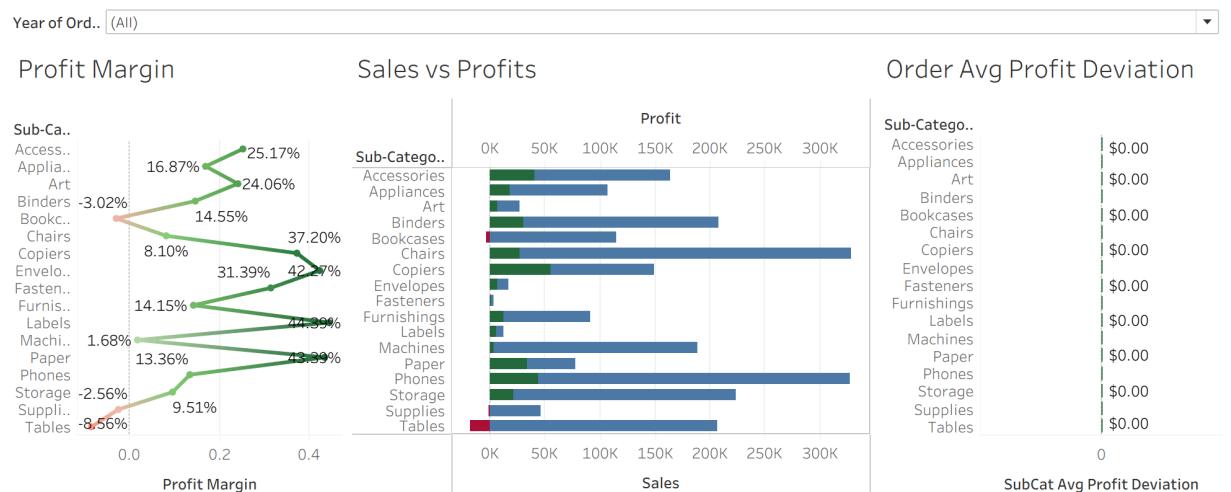
## Profitability View of Categories

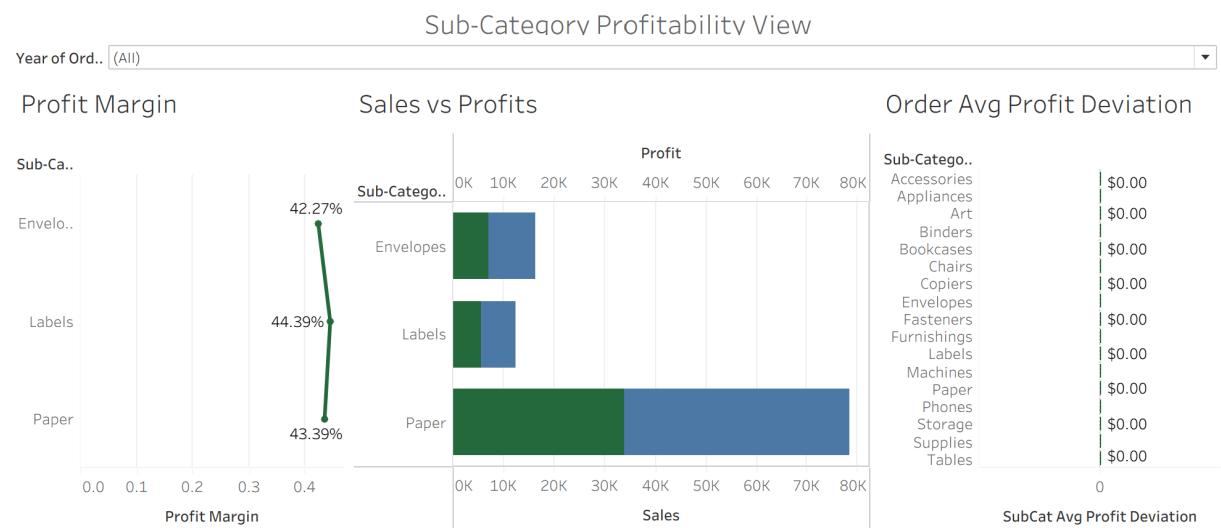
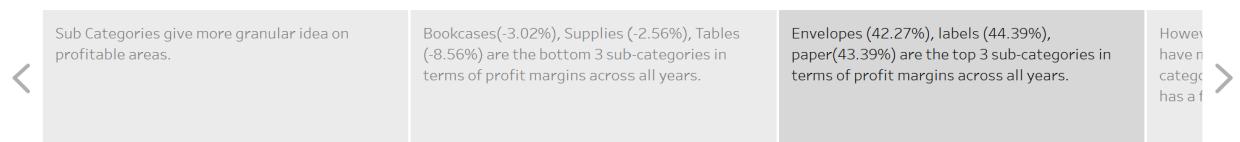
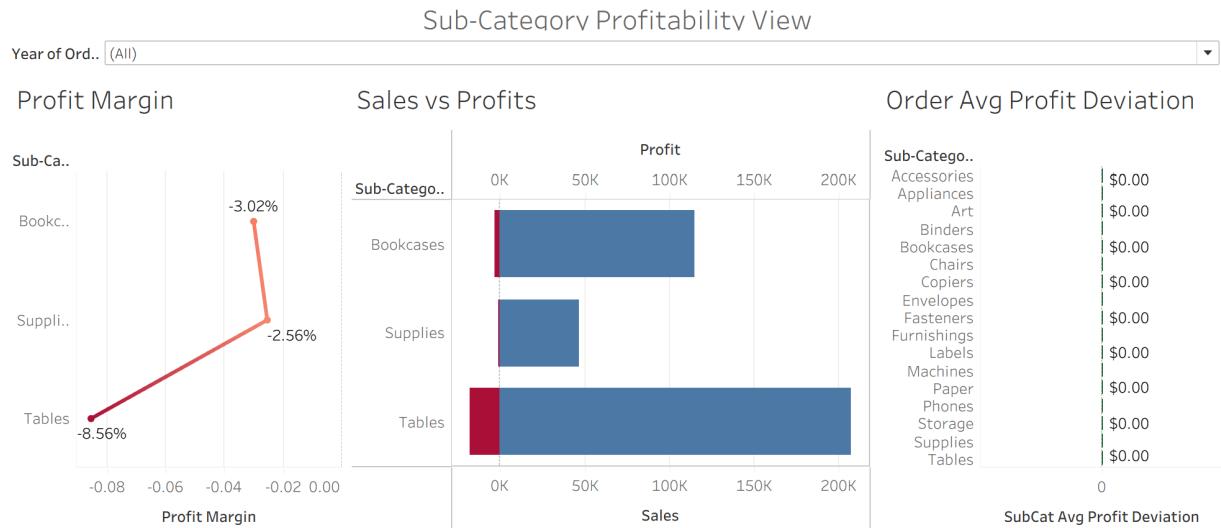
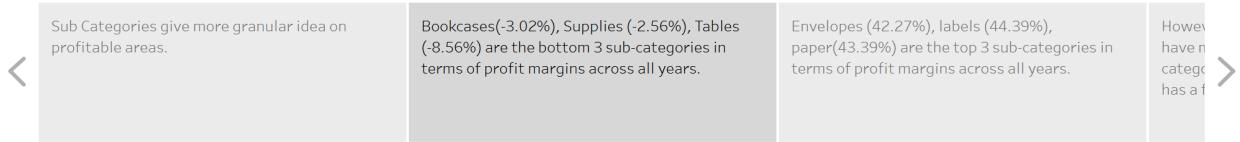


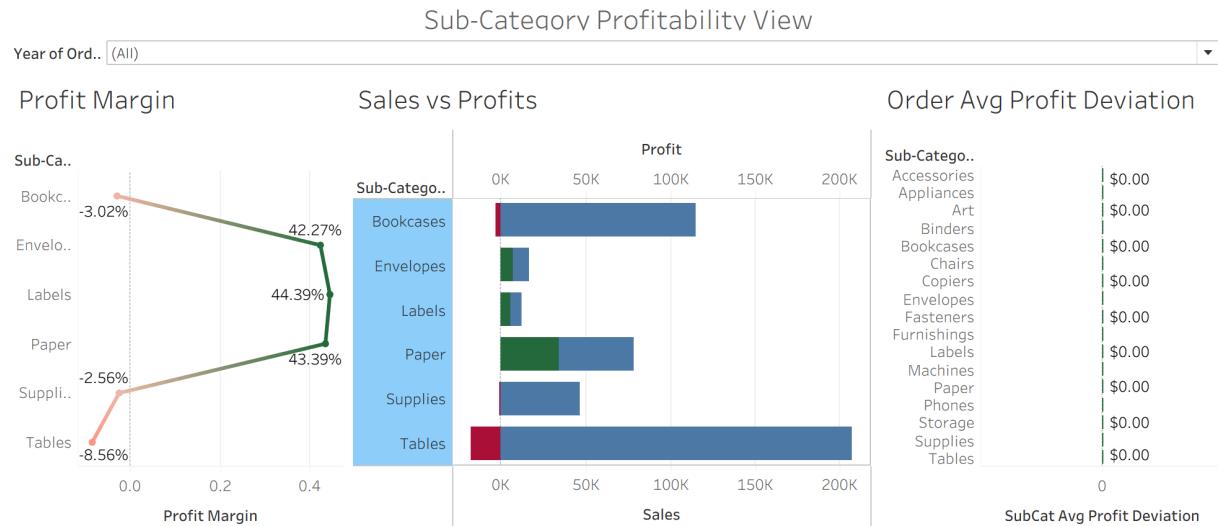
## Story: Sub-Category Profitability Analysis



## Sub-Category Profitability View





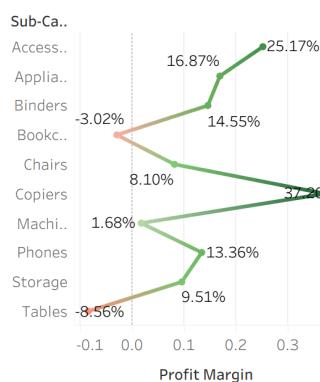




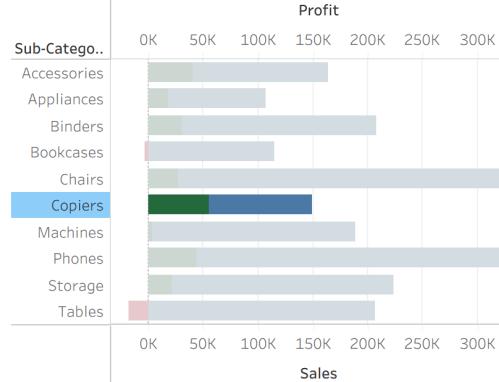
### Sub-Category Profitability View

Year of Ord.. (All)

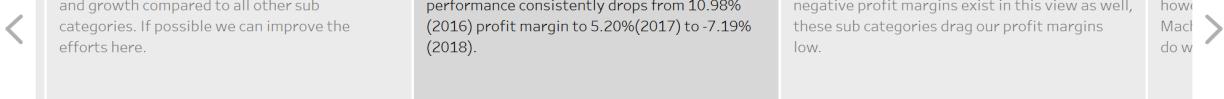
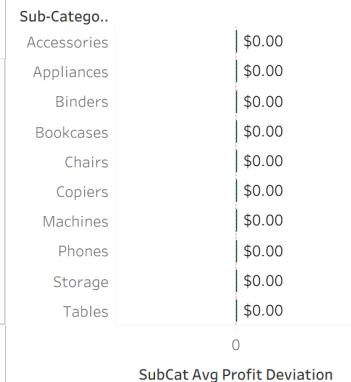
Profit Margin



Sales vs Profits



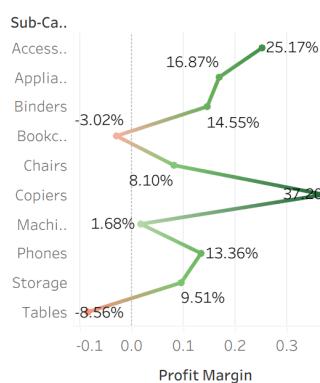
Order Avg Profit Deviation



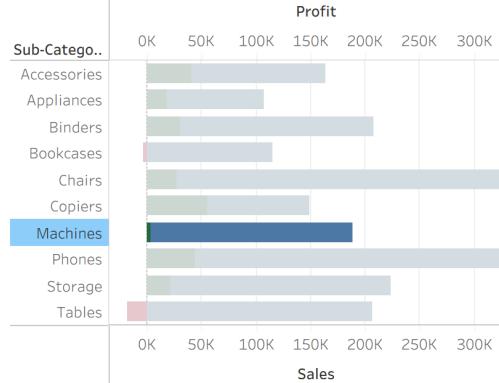
### Sub-Category Profitability View

Year of Ord.. (All)

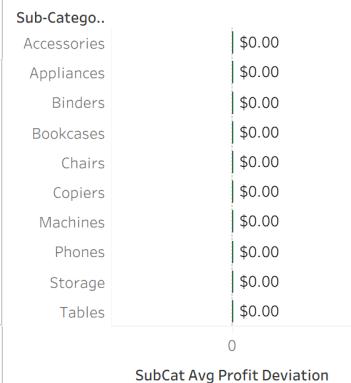
Profit Margin



Sales vs Profits



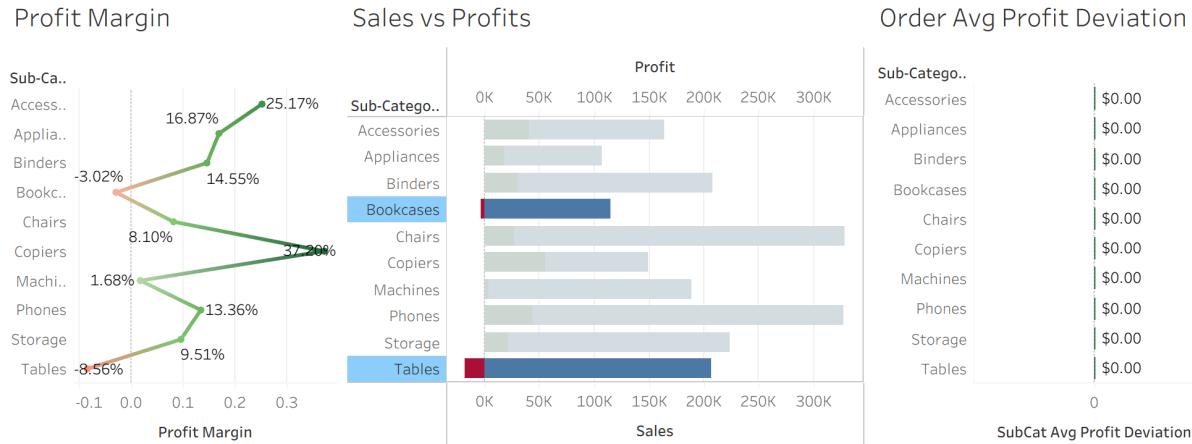
Order Avg Profit Deviation



Margins	Machines have significant sales, yet its performance consistently drops from 10.98% (2016) profit margin to 5.20%(2017) to -7.19% (2018).	Bookcases (-3.02%) and Tables(-8.56%) with negative profit margins exist in this view as well, these sub categories drag our profit margins low.	Last year (2018), shows similar patterns, however, also points out the problems in Machines clearly. We need to consider what to do with the worst performant categories.
---------	---	--	---

### Sub-Category Profitability View

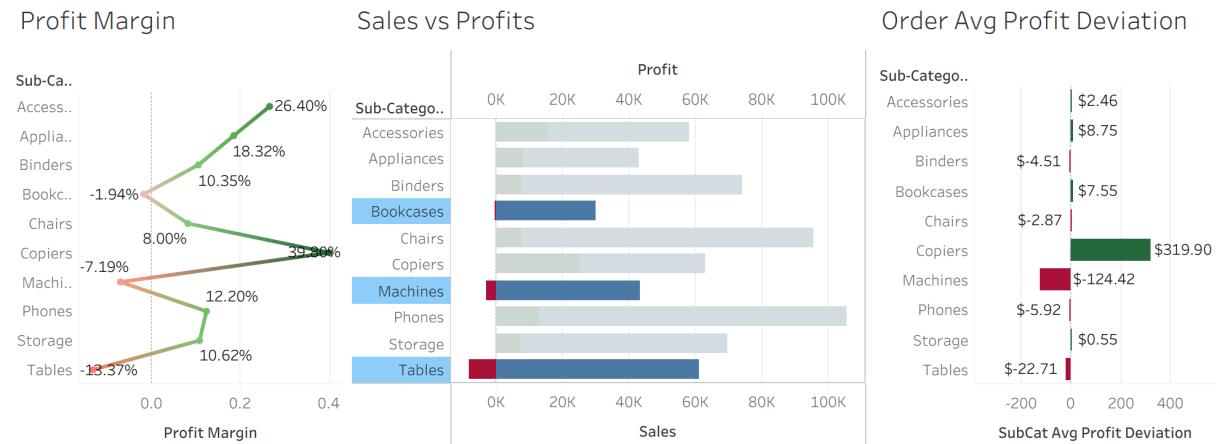
Year of Ord.. (All) ▼



Margins	Machines have significant sales, yet its performance consistently drops from 10.98% (2016) profit margin to 5.20%(2017) to -7.19% (2018).	Bookcases (-3.02%) and Tables(-8.56%) with negative profit margins exist in this view as well, these sub categories drag our profit margins low.	Last year (2018), shows similar patterns, however, also points out the problems in Machines clearly. We need to consider what to do with the worst performant categories.
---------	---	--	---

### Sub-Category Profitability View

Year of Ord.. 2018 ▼



## Story: Customers Profitability and Segmentation Analysis

Not all customers are profitable. Although some of those might be tolerated depending on other factors, and if they are in the relationship builder segment. However, if a customer is not churned and has a long tenure we should expect them to be profitable. This ongoing work will help us to have deeper and better understanding and overall opportunity.



Not all customers are profitable. Although some of those might be tolerated depending on other factors, and if they are in the relationship builder segment. However, if a customer is not churned and has a long tenure we should expect them to be profitable. This ongoing work will help us to have deeper and better understanding and overall opportunity.



Not all customers are profitable.

Although some of those might be tolerated depending on other factors, and if they are in the relationship builder segment.

However, if a customer is not churned and has a long tenure we should expect them to be profitable.

This ongoing work can be continued to dive deeper and help tailor marketing campaigns and overall optimization efforts if seems fit.



Alough some of those might be tolerated depending on other factors, and if they are in the relationship builder segment.

However, if a customer is not churned and has a long tenure we should expect them to be profitable.

This ongoing work can be continued to dive deeper and help tailor marketing campaigns and overall optimization efforts if seems fit.



## CONCLUSION

This analysis has been conducted to understand the store's profitability across different dimensions over 4 years(2015-2018), especially what happened between 2017-2018. It was noticed that even though the profits and sales were increasing, the profit margin of the store has declined from 13.37% in 2017 to 12.68% in 2018. The operations are still profitable, and the company has a good stance. The overall interpretation suggestion is that our operations peaked in 2017 and we saw a decline after it. The main objective is to understand the reasons behind this negative fluctuation, and suggest actionable insights on how to recover/improve:

- Segments
  - Consumer segment's profit margins consistently improved between 2015-2018, from 9.09% to 10.63% to 12% to 13.67%, and the segment has the most sales.
    - Keep supporting and improving the segment
  - Overall our profit margins are lower in segments with higher sales.
  - Corporate has the 2nd most sales while home office has the least. Both have fluctuated inconsistently over 4 years, and declined between 2017-2018. The main reason for the decline might be here.
    - In 2018, the corporate segment had 10.97% profit margin and the home office segment had 13.22% profit margin.
    - Balance and optimize the corporate segment, consider significant changes in the home office segment.
- Categories
  - Furniture category has 32.33% of our sales while the category holds 6.44% of our profits.
    - Evaluate the related efforts and reconsider actions
  - Technology with profit margin 18.60% and Office Supplies (16.05%) have the most sales.
    - Efforts can be focused more on these categories.
- Sub-Categories
  - Bookcases(-3.02%), Supplies (-2.56%), Tables (-8.56%) are the bottom 3 sub-categories in terms of profit margins across all years.
  - Envelopes (42.27%), labels (44.39%), paper(43.39%) are the top 3 sub-categories in terms of profit margins across all years.
  - Machines have significant sales, yet its performance consistently drops from 10.98%(2016) profit margin to 5.20%(2017) to -7.19%(2018).
  - Worse performant sub categories have relatively high sales and we can limit or optimize them, and shift our focus to better performing sub categories.
- Customers
  - Scatter plots reveal individual customers who are profitable to us with all of their orders combined.
  - Clustering puts them into 3 bins: long tenure - high purchase frequency, long tenure - low purchase frequency, short tenure - variant purchase frequency.
    - We can target these clusters separately based on business needs.

after discussing with the stakeholders, we can dive deeper and focus into any of these areas.