Story 1

Title	Goals, Steps, Choices	Available Data	Data Exploration	Geographical: Explanation	Geographical	Supply Chain Management

Sample Sales Data Set

Erin Howland

Story 1

Title	Goals, Steps, Choices	Available Data	Data Exploration	Geographical: Explanation	Geographical	Supply Chain Management

Steps, Goals, and Choices

General Steps

* Determine available data
* What vizualizations and analysis might be available and logical
* Create visuals
* What questions arise and do we have enough data to answer those questions?
* Create questions/insight statements to guide client where they might want to look a little deeper and/or provide additional data for more complete analysis

Goals

Without a specific target audience it is difficult to put together a sample of what I would produce for a client. Executives will want to know different things and be presented to differently than would a regional manager or a department head, for example. For this reason, I will try to include a variety of visuals and filtering options that get into four primary categories: supply chain management, operational standards, customers, and growth as a company.

Choices

Profit and Sales

In most cases, I have chosen to put sales and profit data side-by-side. Profit cannot occur without sales so these two are, by nature, inextricably linked. I have also chosen to do this so that a client can see which products might be a higher priority to optimize in order to get a higher profit margin. Depending on who this data is presented to, however, I may have made different choices in how segregated the sales and profit data is. If I was asked to give a presentation that was specifically geared toward sales managers, then marrying sales and profit together as I've chosen to do here would not necessarily be the most useful choice, though executives and operations managers would be more likely to find my chosen presentation more useful.

Time Slices: Order Date vs Ship Date

When thinking about time slices, I have used a mix of order date and ship date. I have tried to use ship date when it comes to presenting data that shows in financials since generally speaking, revenue, and therefore profits, are not realized until a good/service is turned over to the customer. Without detail on specific policy regarding whether the company uses Free on Board (FOB) shipping point or FOB destination and becauase we are lacking let shipping destination date data, I am making the assumption that the client is using FOB shipping policy, which means the customer takes possession at time of shipment rather than time of receipt.

I have used order date generally only when looking at data related to orders themselves. I debated on whether to use shipping or order date when considering whether a customer should be counted, and I determined to use ship date here, as well, in order to (1) maintain continuity of filters with other customer data such as customer sales and profit, and (2) to eliminate the possibility of inadvertently counting a new customer creating and order and then cancelling before shipment. As you'll see, this decision does cause a visually alarming chart when looking at unique customers by year should all years be shown (ship dates go through Q1 2018) but it can be explained and controlled for, and it's easy enough to provide this data by order date instead should that be their desire.

One place where I was unable to determine whether order date or ship date would be more appropriate is in relation to the People dimension table. Without knowing what role these people have, it's difficult to determine whether order or ship date makes more sense. If we assume these are sales managers, we could use order date if we want to consistently measure their performance against orders. This would make sense if there is an incentive for or desired measurement tool based on time of order rather than time of shipment. If there are commissions, having both order date and ship date might be more appropriate if we want transparency between a person's sales and their commissions; having both would allow for immediate transparency for any cancelled or returned orders that would therefore not be paid as commission.

Story 1

Title	Goals, Steps, Choices	Available Data	Data Exploration	Geographical: Explanation	Geographical	Supply Chain Management

The Data

* Data has been provided and is current as of Q1 2018
* The sales data has three tables: **Orders** (fact table), **People** (dimension table), and **Returns** (dimension table)
* I created relationships between the tables for analysis
* I did the initial data exploration by both looking at the data itself in its original format and through visualization to determine things like whether there was missing data, what outliers might exist, and general things to think about as I move forward. Some of visuals I used are shown on the next story point.

Orders

The **Orders** table is a fact table, so most of the useful data is here (customer, order details, sales, and profit). When thinking about what most executives and managers would be interested in, this is the table that will be a primary focus because of its direct tie to the financial well-being of the company.

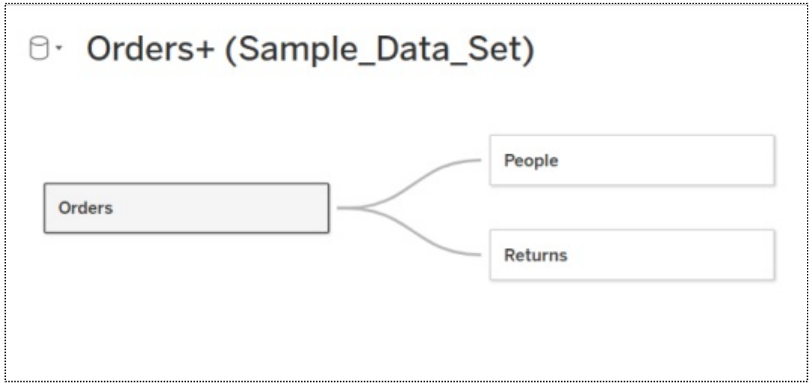
People

The **People** dimension table appears to show regional or district managers. This would be helpful if/when we find insights about a specific state/region so the client knows where to direct further questions. We would need more context as to the role of these people before we can dig further for analysis - for example:

* Are they directly responsible for being able to offer discounts?
* Are they running a warehouse and creating budgets and therefore have direct control over operational expenses (OpEx)?

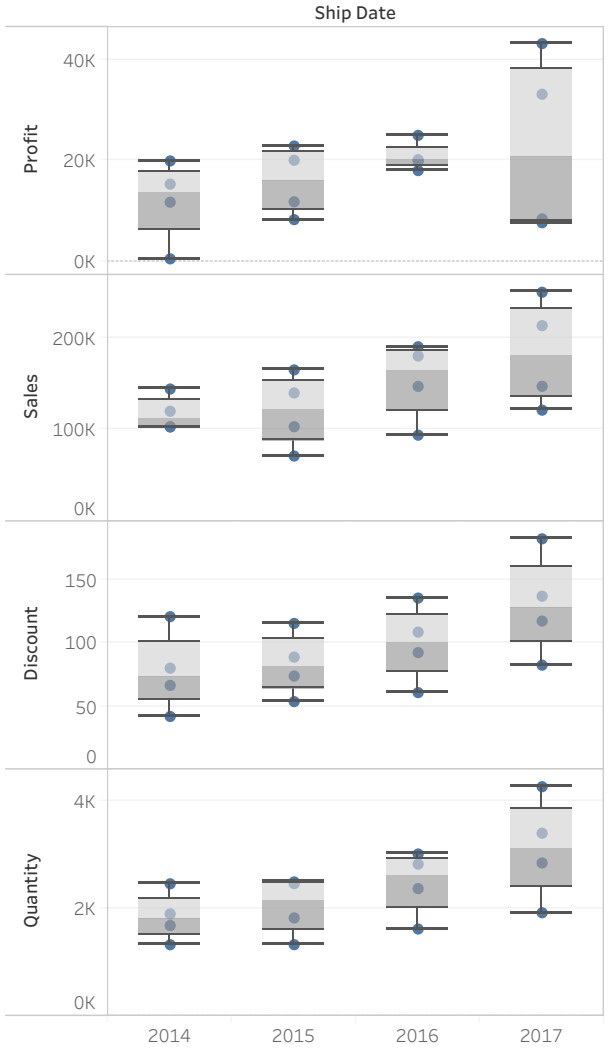
Returns

The **Returns** dimension table contains *Order ID* and a *Return Status*. The granularity for this table is at the *Order ID* level, so it's impossible to tell whether the entire order was returned or if a subset of the items orders was returned. Because of this, we cannot reliably assign any dollar value to the return despite being able to match to the **Orders** table on *Order ID*. We can, however, use this to find customer patterns regarding returns (e.g. serial returners or segments with higher than normal return rates). Because of the mismatch of granularity between the **Orders** and **Returns** tables, I did not factor for returns in *Sales* and *Profit* analysis. Were this to be a real-life situation, I would work with data engineering to see if I could get the required granularity to conduct and more thorough analysis.

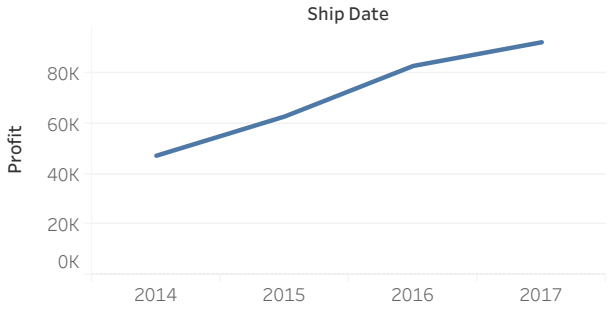


Story 1

Avg Profit, Sales, Discount, and Quantity by Year



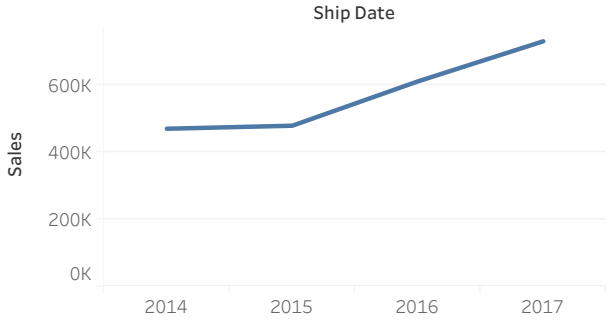
Annual Profit



Region
- ✔ Central
- ✔ East
- ✔ South
- ✔ West

Year of Ship Date
- ✔ 2014
- ✔ 2015
- ✔ 2016
- ✔ 2017
- ☐ 2018

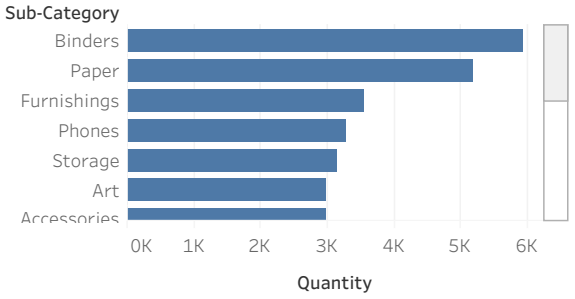
Annual Sales



Quarter of Ship Date
- ✔ Q1
- ✔ Q2
- ✔ Q3
- ✔ Q4

Segment
- ✔ Consumer
- ✔ Corporate
- ✔ Home Office

SubCategories by Sales Quantity



Story 1

Goals, Steps, Choices	Available Data	Data Exploration	Geographical: Explanation	Geographical	Supply Chain Management	Customers: Explanation

Geography

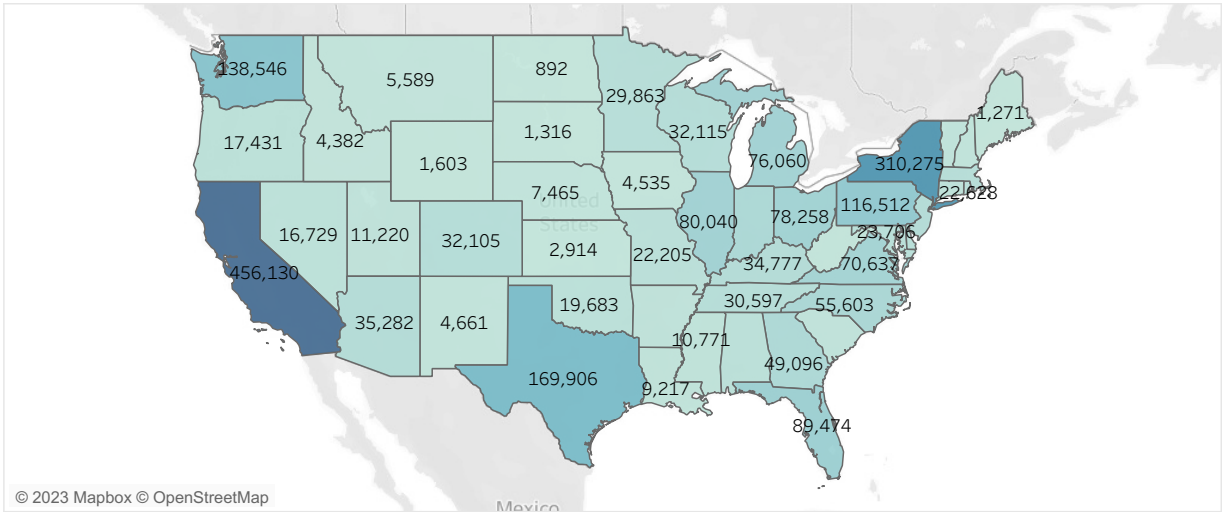
The purpose of this is to see at a high level where things are happening - this focuses solely on state and regional information regarding profit and sales.

Insights

We can see that the company is operating within the contiguous US. Domestically, Alaska, Hawaii, and US territories are open for growth as are Canada and Mexico if the client is open to nearby international expansion. We might want to consider why this client isn't operating in those areas yet. Some potential topics we may want to raise on this issue include competitors, cost (of shipping, legal fees tied to new markets, operations), capacity, and goals (e.g. is the goal to better penetrate areas that are easier to ship to before expanding to routes that may be more logistically complex?).

Story 1

Sales By State



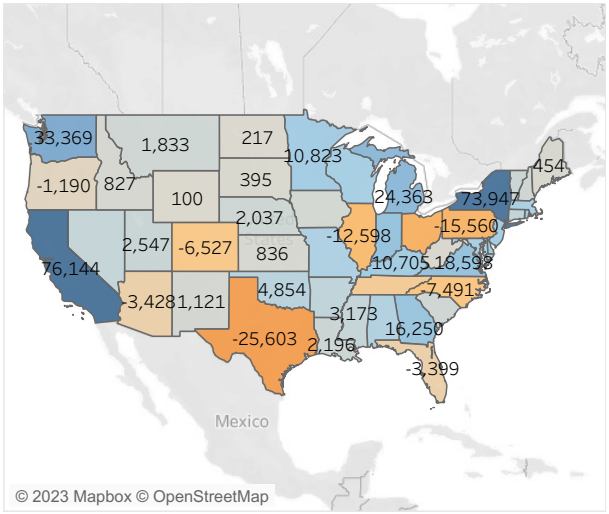
Region
- ✔ Central
- ✔ East
- ✔ South
- ✔ West

Profit
-25,729.3563 to 76,381...

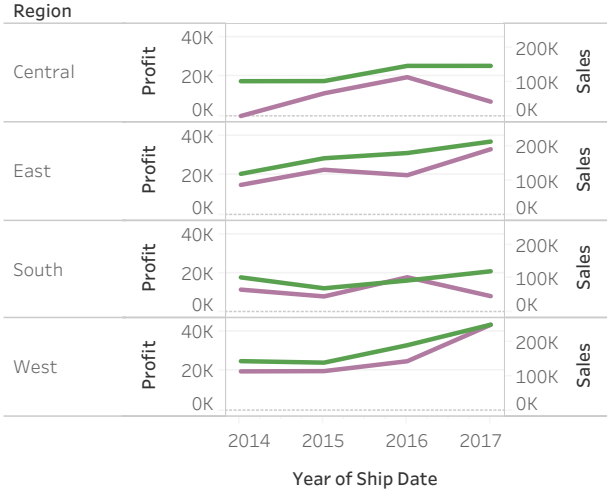
Year of Ship Date
- ✔ 2014
- ✔ 2015
- ✔ 2016
- ✔ 2017
- ☐ 2018

Quarter of Ship Date
- ✔ Q1
- ✔ Q2
- ✔ Q3
- ✔ Q4

Profit by State



Profit and Sales by Region



Segment
- ✔ Consumer
- ✔ Corporate
- ✔ Home Office

Sales
892 — 456,130

Profit
-25,603 — 76,144

Measure Names
- Profit
- Sales

Story 1

Supply Chain Management: Shipping

Turnaround from order date to shipping date is a way to evaluate supply chain management by showing whether there are issues with:

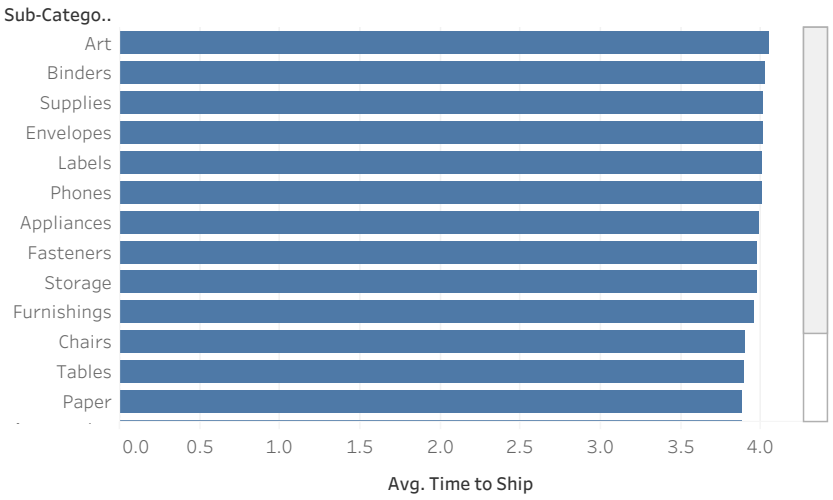
* System - is the order system is working appropriately, both for upstream and downstream orders?
* Staffing - is there is proper staffing to handle order volume in a time frame the company has determined appropriate?
* Supply - are there known or potential issues upstream that would cause downstream problems they may need to alert customers to in order to manage expectations?

Supply chain issues could also ultimately have an effect on profitability of an order if, for example, steep discounts were given to offset longer than normal wait time.

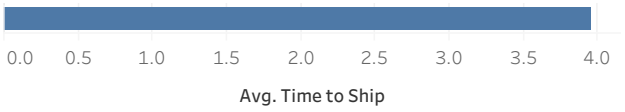
Insights

Nothing in particular stands out here. Average days between order date and ship date are fairly consistent. I don't think there's anything I would point out to this client here unless I learned of something that would be of value for them to address unless they provided more information that would give context to these figures. For example, is their average shipping over industry standard or do they want to implement a new policy to decrease the turnaround to 2 days but they're also considering layoffs in that department?

Avg Days From Order to Ship by SubCategory



Days From Order to Ship: Company Average



Year of Order Date

- [x] 2014
- [x] 2015
- [x] 2016
- [x] 2017

Quarter of Order Date

- [x] Q1
- [x] Q2
- [x] Q3
- [x] Q4

Story 1

Customers

An important aspect of any organization is to understand its customers. Some of the most basic questions that we can start with regardless of organization type include:

* WHO are the customers (names, demographics, client type, etc.)?
* WHAT do they want/need?
* WHEN do they want/need those products/services (e.g. is it constant, cyclical, dependent on disposable income, etc.) and WHEN do they engage with us?
* WHERE are they located?
* WHY are they selecting us?
* HOW do they engage with us (e.g. social media, mailers, text, phone, etc.) and HOW are they evolving over time?

I have included the top 10% of customers by both *Profit* and *Sales*. Because *Profit* and *Sales* do not always align in a predictable fashion, I have added color coding to the top customers by *Sales* visual to also point out the associated *Profit*, and vice versa for the *Profit* visual. This may be of use should a client choose to investigate why a particular customer may be less profitable than expected given their sales data.

One thing I would point out to this client is that the total number of distinct customers they have in the top right line graph does not match the total number of distinct customers in the map. This is because of the way the disctinct count is done. The line graph shows the true number of distinct customers the client has. The map shows distinct customers by state, and many customers are associated with multiple states. Whether this is because they have moved or because they might be ordering for a multi-state company we don't know, but this is something to be aware of when looking at the data and it should be clearly disclosed to the client.

To reduce the number of visuals shown in the story - there is a balance between level of detail presented for decision-making and succinct presentations - I have incuded customer segments as a filter rather than as a study in and of itself.

Insights

Positive

* Customers are increasing over time. This may be a result of marketing efforts or it could just be continued presense and therefore brand awareness or anything in between. We'd want to have a decent idea of cause to determine how to move forward; if we had more data, we'd be able to provide this analysis.
* There are customers in every state, though the customers are fairly concentrated by zip code (chart not shown in presentation). There may be room for increased market penetration but we would want to conduct this study in relation to population density, income, and other demographic data (e.g. census data). For example, we could look at sales by post code in relation to population density to look for cold spots.

Negative

There are several customers who have high sales but represent a disproportionate loss in terms of profit. Why is this occuring and how can we stop it from happening? Without doing a full root cause analysis, it's likely that this may have something to do with discounts and/or products. Discounts reduce profits and if a product has a narrow margin, a discount could make that particular product unprofitable.

Story 1

Top 10% Customers by Profit

Customer Name	
Tamara Chand	8,981
Raymond Buch	6,976
Sanjit Chand	5,757
Hunter Lopez	5,622
Adrian Barton	5,445
Tom Ashbrook	4,704
Christopher Martinez	3,900
Keith Dawkins	3,039
Andy Reiter	2,885
Daniel Raglin	2,869
Tom Boeckenhauer	2,798
Nathan Mautz	2,752
Sanjit Engle	2,651
Bill Shonely	2,616
Harry Marie	2,438
Todd Sumrall	2,372
Brian Moss	2,199
Christopher Conant	2,177
Jane Waco	2,174
Helen Wasserman	2,164
Greg Tran	2,163
Laura Armstrong	2,059
Adam Bellavance	2,055
Fred Hopkins	2,050
Pete Kriz	2,038
Steven Roelle	1,990
Shirley Daniels	1,985
Clay Ludtke	1,934
Robert Marley	1,903

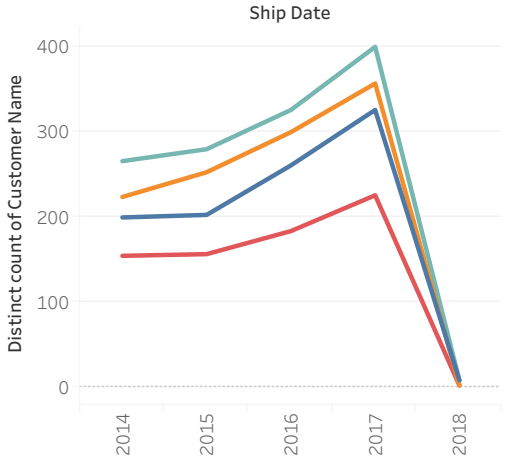
Top 10% Customers by Sales

Customer Name	
Sean Miller	25,043
Tamara Chand	19,052
Raymond Buch	15,117
Tom Ashbrook	14,596
Adrian Barton	14,474
Ken Lonsdale	14,175
Sanjit Chand	14,142
Hunter Lopez	12,873
Sanjit Engle	12,209
Christopher Conant	12,129
Todd Sumrall	11,892
Greg Tran	11,820
Becky Martin	11,790
Seth Vernon	11,471
Caroline Jumper	11,165
Clay Ludtke	10,881
Maria Etezadi	10,664
Karen Ferguson	10,604
Bill Shonely	10,502
Edward Hooks	10,311
John Lee	9,800
Grant Thornton	9,351
Helen Wasserman	9,300
Tom Boeckenhauer	9,134
Peter Fuller	9,063
Christopher Martinez	8,954
Justin Deggeller	8,828
Joe Elijah	8,698
Laura Armstrong	8,673

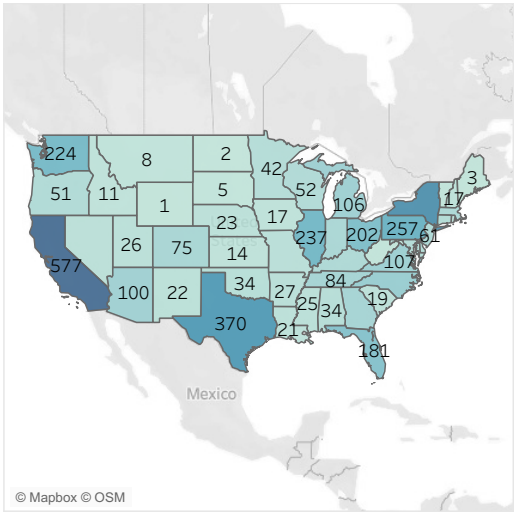
Sales
3,419 — 19,052

Profit
-4,109 — 8,981

Count Distinct Customers Over Time



Distinct Customers by Location



Region
- Central
- East
- South
- West

Segment
- Consumer
- Corporate
- Home Office

Region
- Central
- East
- South
- West

Year of Ship Date
- 2014
- 2015
- 2016
- 2017
- 2018

State
- Alabama
- Arizona
- Arkansas
- California
- Colorado
- Connecticut
- Delaware
- District of Columbia
- Florida
- Georgia

Distinct count of Custom..
1 — 577

Story 1

Products

Categories

There are three categories of goods this client provides: furniture, office supplies, and technology.

SubCategories

I had wanted to show profit and sales by subcategory using the same kind of visual for continuity, but I noticed that when I did that, three subcategories were missing from the visual: bookcases, tables, and supplies. These three subcategories are not profitable (supplies was in 2014 and 2015 but by 2016 it was no longer profitable) and therefore do not show on the visual I had initially selected. While the visual I had originally selected is great for showing the profitable subcategories, I thought it was important to also draw attention to the ones that were not profitable so a client could look into those and figure out why they aren't profitable (we are unable to provide potential answers for them with the given data) and/or evaluate whether those categories are ones they might want to drop.

Products

There are too many products to examine them all. A better approach is to look at them through their subcategories and categories except if we notice specific things related to specific products that may want to address.

Insights

Positive

Overall, the company is in the black for profits.

Negative

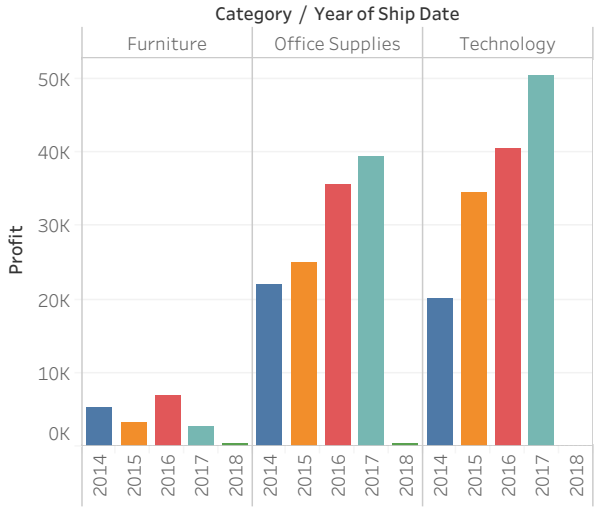
There are three subcategories that are, cumulatively, not profitable: supplies, bookcases, and tables. When we look at individual years, notice that tables have always been unprofitable while bookcases and supplies has been inconsistent. New for 2017 was the addition of machines to the list of unprofitable subcategories. There are a few things to think about here from a business perspective:

1. How do these (sub)categories relate to sales?
2. Are there specific products within these (sub)categories that are causing them to be unprofitable?
3. What is the associated OpEx involved that might be making these subcategories (or individual products, as it may be) unprofitable and is there a way forward to ultimately have all granularities of products in the black?

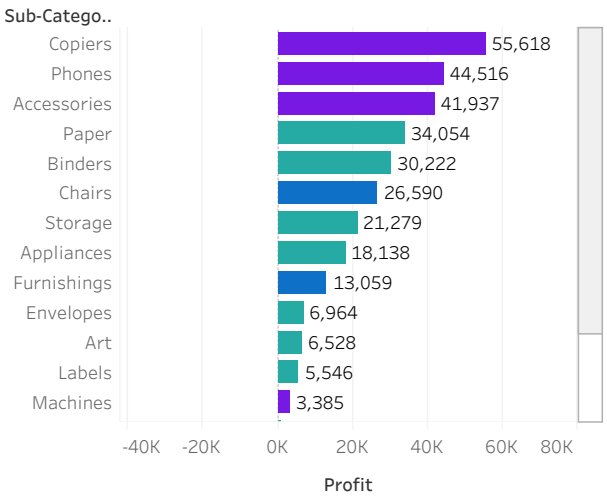
Story 1

Geographical: Explanation | Geographical | Supply Chain Management | Customers: Explanation | Customers | Products: Explanation | **Products**

Profit by Category



Profit by SubCategory



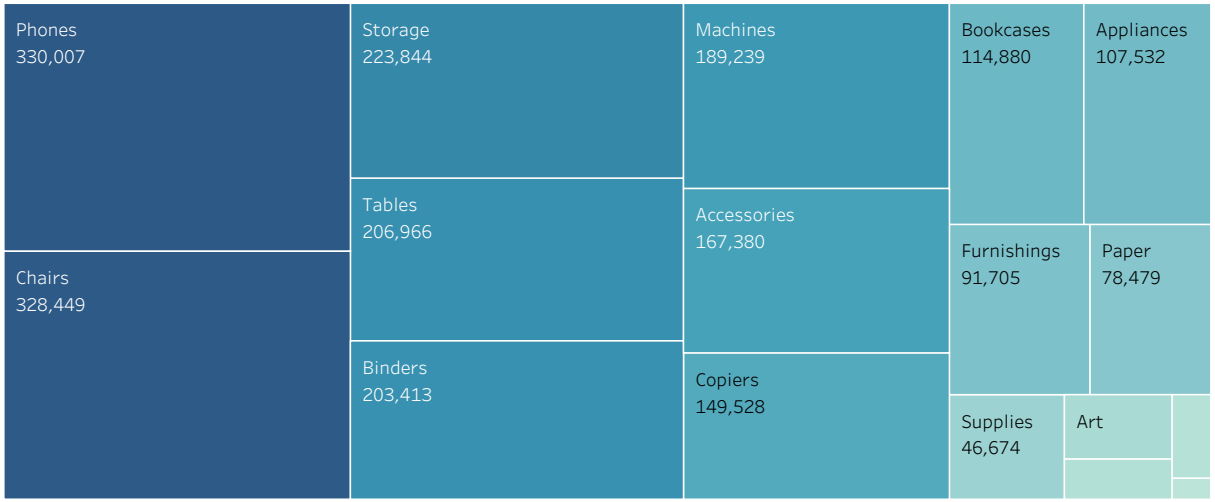
Category
- Furniture
- Office Supplies
- Technology

Category
- ✔ Furniture
- ✔ Office Supplies
- ✔ Technology

Year of Ship Date
- ✔ 2014
- ✔ 2015
- ✔ 2016
- ✔ 2017
- ✔ 2018

Quarter of Ship Date
- ✔ Q1
- ✔ Q2
- ✔ Q3
- ✔ Q4

Sales by SubCategory



Segment
- ✔ Consumer
- ✔ Corporate
- ✔ Home Office

Region
- ✔ Central
- ✔ East
- ✔ South
- ✔ West

Sales

3,024 — 330,007