Semester Recap Report

Paul Larison

University of Arkansas Little Rock

BINS 5351: Data Analysis and Reporting

December 4, 2022

Table of Contents

Executive Summary

Adventureworks has a profitability problem. While the company is making money as a whole, the bottom line is being held back by losses in specific wholesale product subcategories. The biggest losses come from bicycle sales, with additional losses in components. Smaller losses are occurring in the component and clothing categories, with only the tiny accessories category avoiding negative sales.

I compared product performance between internet retail sales and traditional wholesale sales and discovered that retail is doing quite well, but wholesale is operating at a net loss. Most of the losses are occurring in wholesale bicycle sales. Mountain bikes were doing fairly well until profitable models were discontinued, but touring and road bike categories, while profitable when sold on the website, are bleeding money in wholesale. Closer analysis determined that these losses are consistent across regions and sales teams and not tied to any special discounts beyond the normal 40% wholesale discount.

All of the products in the Road and Touring bike categories, as well as the Mountain-200 model, are priced too low relative to their current production costs when the wholesale discount is applied. If costs cannot be reduced, we should increase the prices and/or lower the wholesale discount. This will likely reduce sales volume, but the remaining sales will make money instead of losing it. E-commerce prices can be maintained by offering discounts on the new, higher official prices. I also recommend bringing back the Mountain-100 and 500 bikes, which were popular and profitable before their discontinuation.

Business Analysis

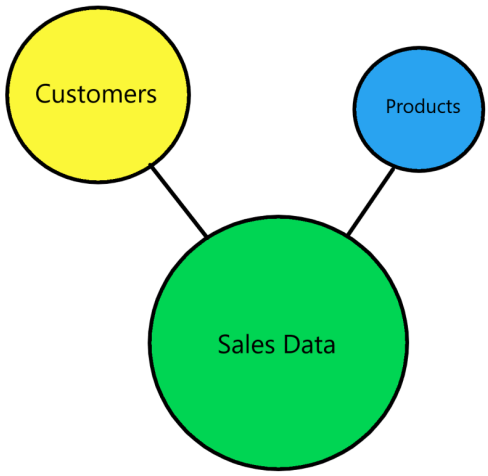
Company Overview:

AdventureWorks is a bicycle manufacturing company that builds and distributes mountain, road, and touring bikes, as well as cycling clothing and accessories. The company employs about 500 workers in its Bothell, WA location, where its final products are assembled and stored. Adventureworks recently purchased one of its major suppliers and now owns a component factory in Mexico, but still purchases many other components from outside suppliers.

AdventureWorks sells its products in North America, Europe, Asia and Australia. The bulk of wholesale sales are in North America, with five sales regions in the United States and another in Canada. Overseas sales regions are Australia, France, Germany and the United Kingdom. Each of these regions is assigned to a sales team responsible for managing relations with resellers. AdventureWorks also has a direct-to-consumer web store.

Data Structure Overview

We have a great deal of data, but today we will focus on our sales data, which is formed by the interaction of customers and products. We arrange this data to mirror the real world, and maintain a list of all our products. This list includes properties such as cost and list price, and the items are sorted into a hierarchy of groups for organization. We also maintain a much longer list of customers that has properties of its own. While we have many more customers than products, they are only grouped into two categories.



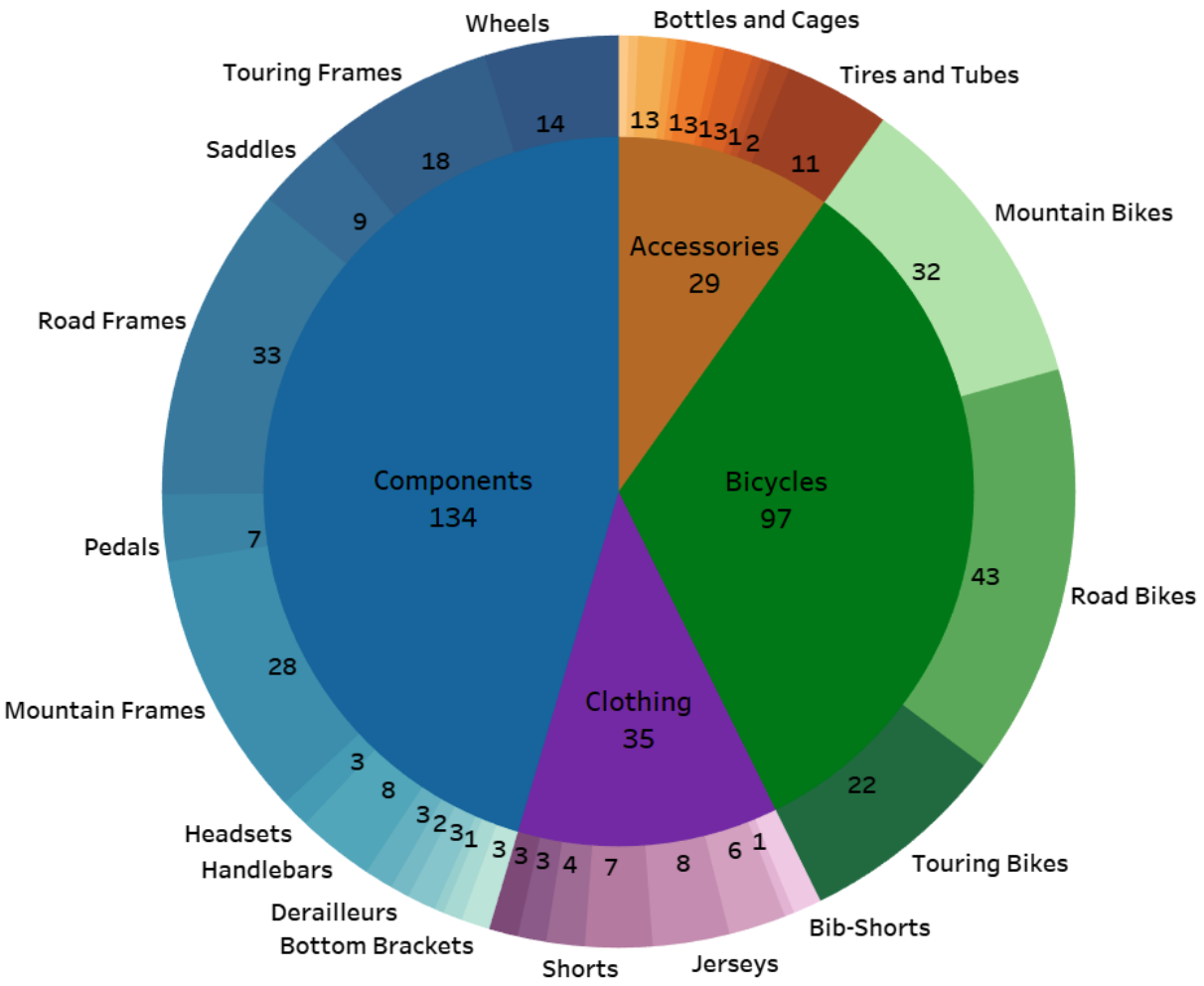
We are analyzing the sales data in search of missed profits and outright losses. To do this we must first understand our products and our customers.

Products

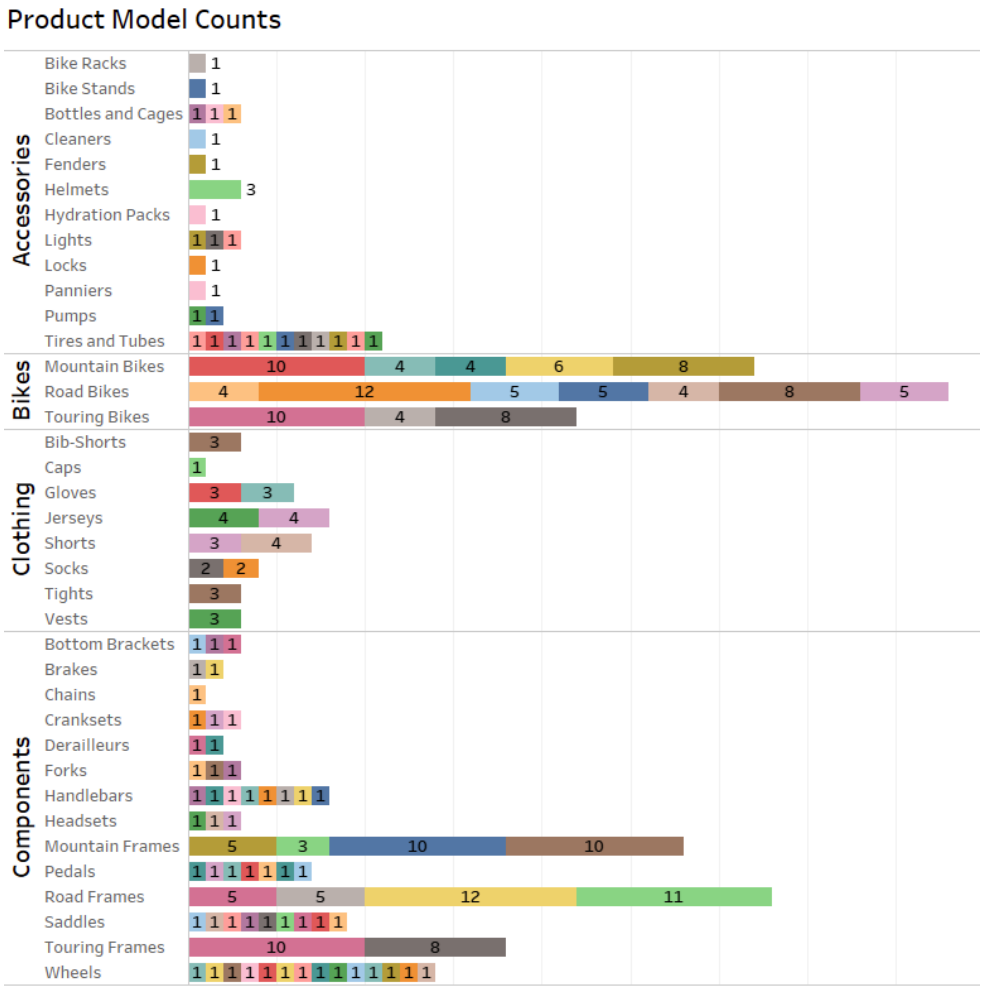
Our products are broken down into various categories and subcategories. We will ignore the 209 unfinished products in this analysis as they do not directly generate profit or loss. Finished products are grouped into four categories: Accessories, Bikes, Clothing and Components. The category with the highest product count is Components, which is not surprising as products in the second largest category, Bikes, are made up of multiple components. Unlike the other three categories, Components are only sold wholesale and are not available on the web store. The Accessories and Clothing categories are smaller, containing supplemental products meant to promote brand awareness and customer loyalty. These categories are further divided into 37 subcategories, as shown below:

Number of Categories	4
Subcategories	37
Product Models	119
Saleable Products	295
All Products	504

Categories and Subcategories by Number of Products

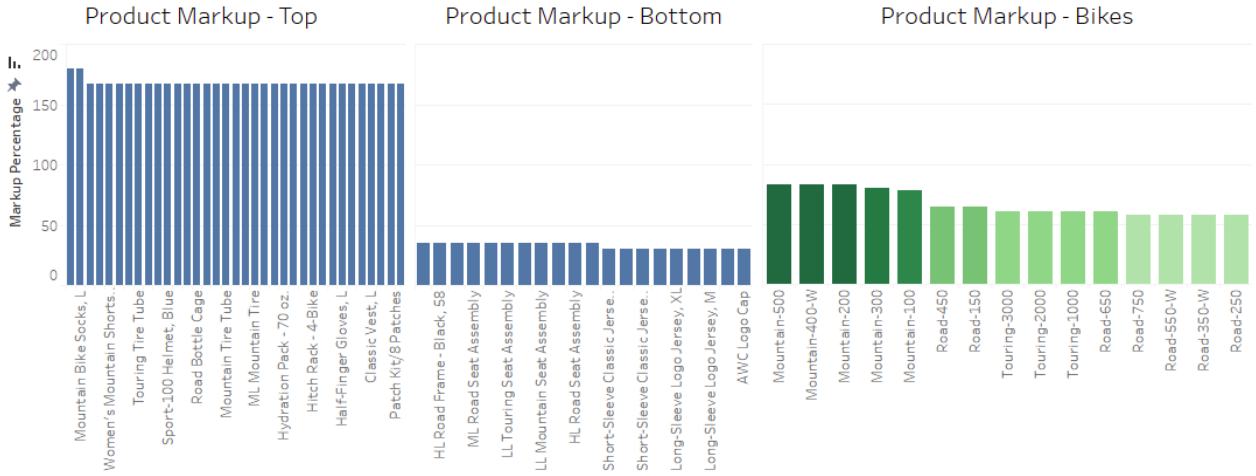


Like products are grouped into Product Models. For example, the Mountain-500 model contains 4 sizes and two colors of otherwise identical bicycles. Many models only contain one product.

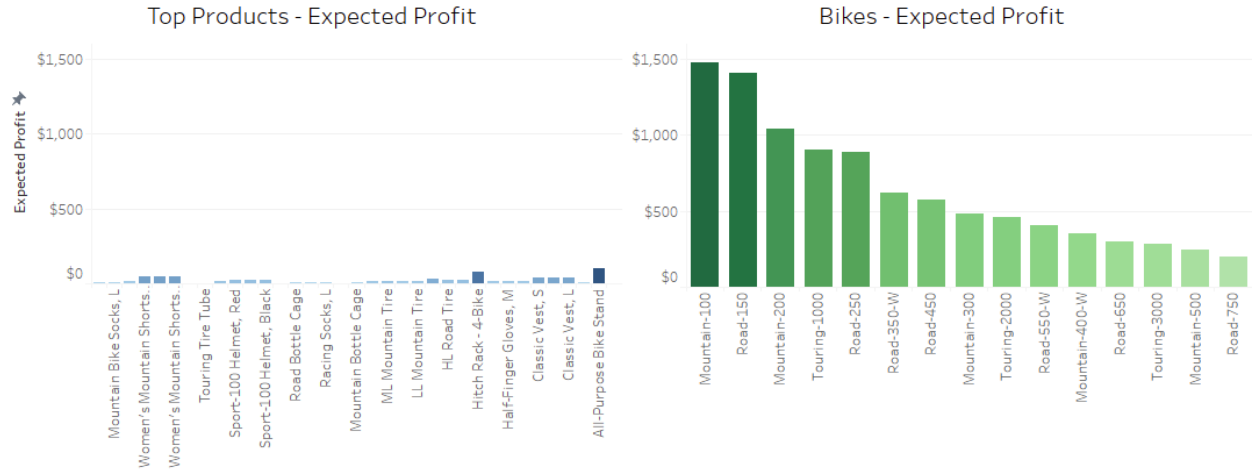


Expected Profitability

Expected profitability is calculated based on standard costs and list prices. So if a product that costs $100 to create is listed at $150, its expected profitability would be $50. The ratio of this difference is the Markup Percentage. The highest markups are on clothing and accessory products, while the smallest markups are on components and on outerwear capable of prominently displaying AdventureWorks branding.



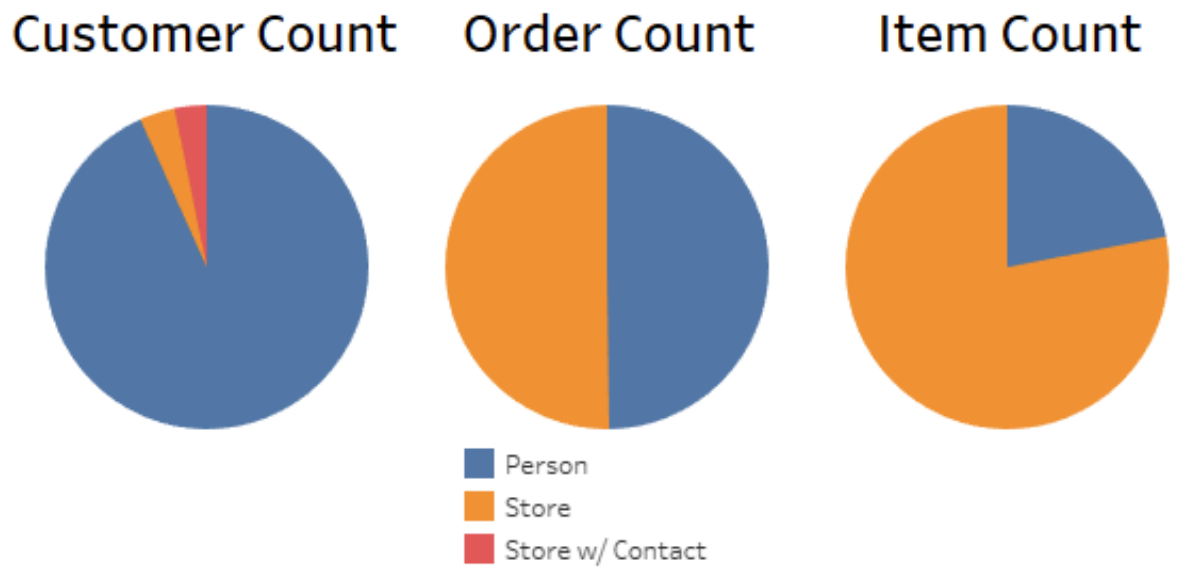
It should be noted that high markups do not always equal high profits, as higher value products will generate larger sums from smaller percentages. That means that the highest expected profitability is associated with products in the Bikes category, and we would need to sell a very large volume of high-markup socks to generate the profit of a single bike.



Customers

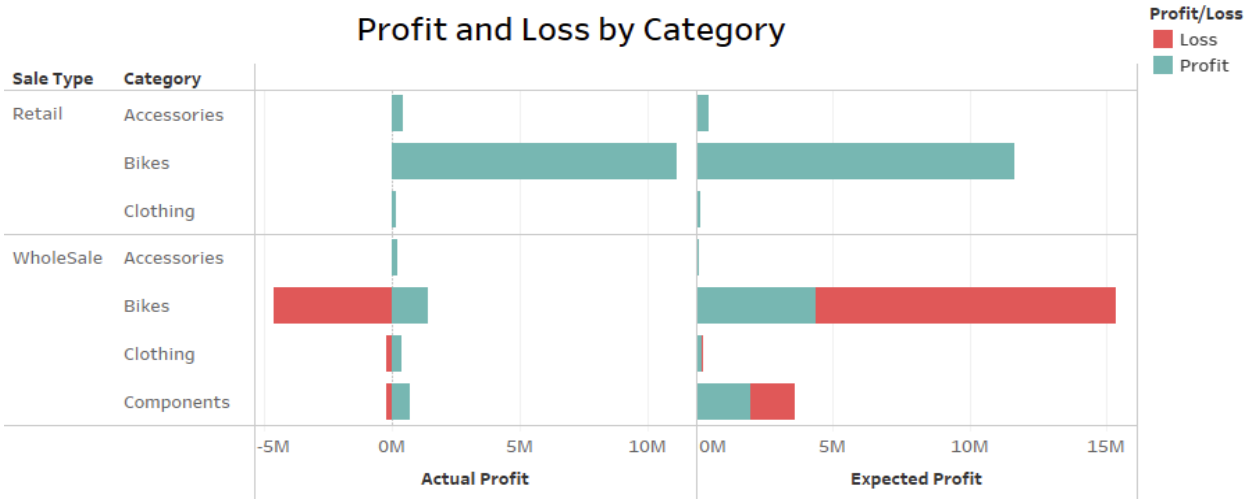
We have two kinds of customer: Retail customers, who purchase their products from our web store, and wholesale customers, to whom we make the bulk of our sales. Wholesalers are further divided into businesses that do or don't have a single point of contact.

Most of our customers are individuals, but businesses account for most of our sales, as visualized below. (Store types were combined when calculating order and item counts.) This illustrates that, while the recently added direct to consumer business is going well, the core of our customer base is still wholesale.

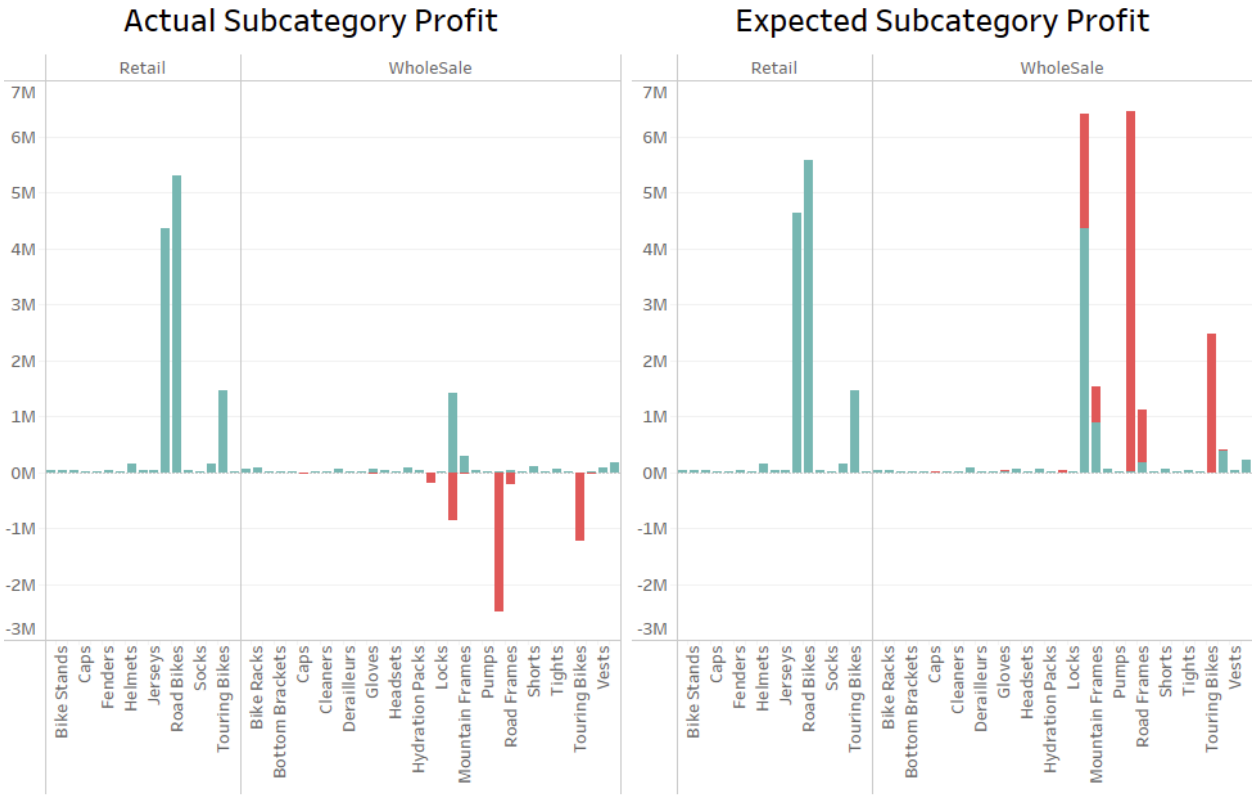


Sales: Retail and Wholesale

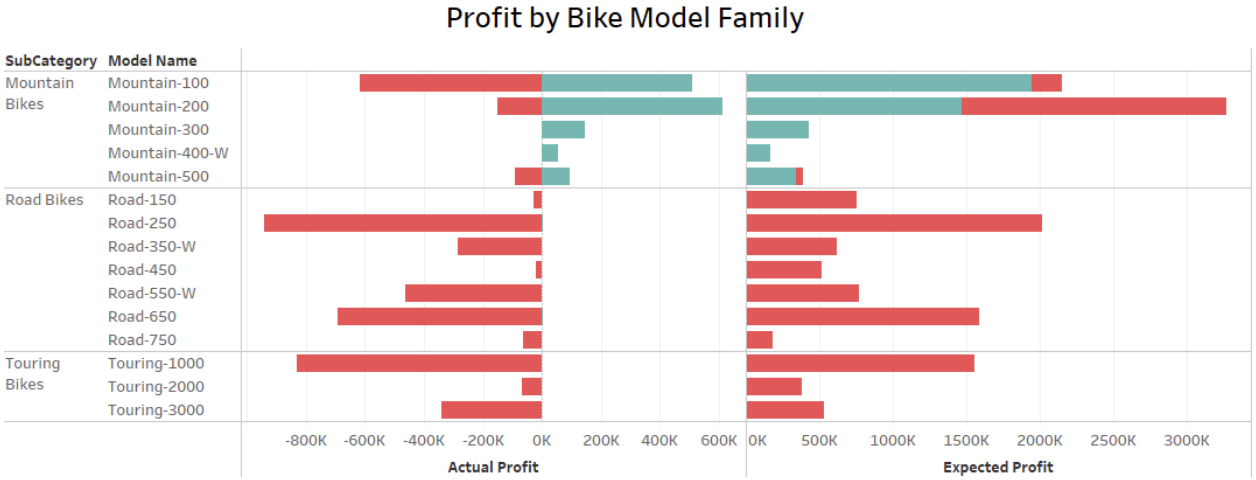
With these factors in mind, we will examine real profits vs expected profits separately for retail and wholesale customers. An overview chart at the category level shows that retail profits are doing just fine, but wholesale is vastly underperforming. Despite accounting for 78% of products sold, wholesale is losing money, primarily in the Bikes category.

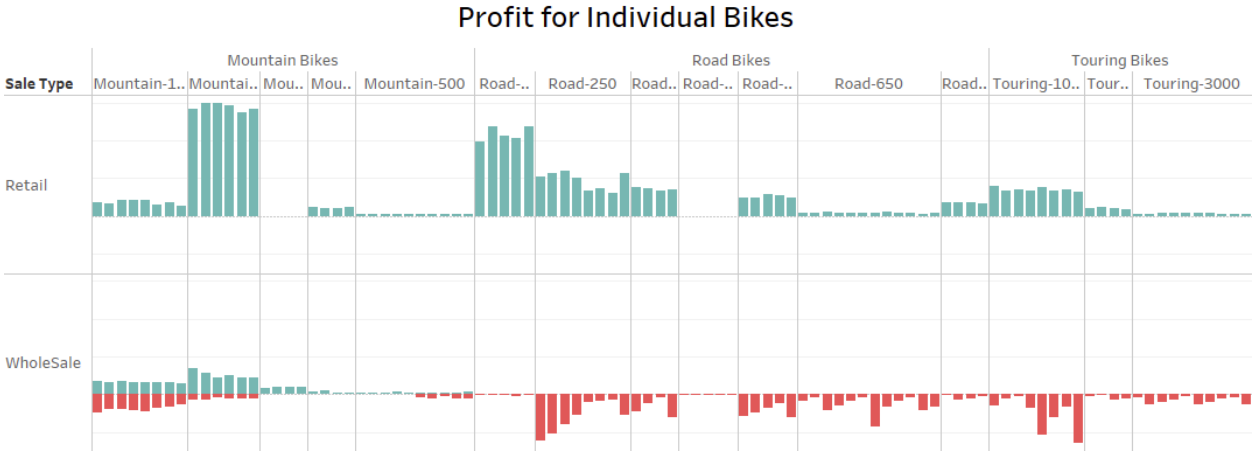


Zooming into the Subcategory level tells much the same story. Wholesale sales of Mountain and Road bikes were expected to result in the highest profits of all segments. Mountain bikes instead produced mixed results, while Road bikes, along with the also-sizeable category of Touring bikes, produced a loss during the measured timeframe.



Now we'll set the retail results aside, since that section seems to be doing fine, and drill down into the model and product levels of wholesale bike sales. It quickly becomes apparent that Mountain bikes are our best performers simply by not being disasters. All products in the Road and Touring subcategories are losing money any time they are not sold through the web store.





Loss Analysis

To determine where these losses are coming from, we will first chart model losses by sales reason. This chart shows us a reason for the odd performance of the Mountain-100, which seems to have been discontinued despite being our most profitable wholesale bike by a large margin. A similar fate befell the lower volume, but still profitable, Mountain-500. The Mountain-200 seems to still be in production and did post a profit, but an examination of the "No Discount" losses shows that nearly 1,700 products sold for small losses, cutting the expected profits for this model in half. Note the "No Discount" does not include the 40% off that is automatically applied to all wholesale sales.

Model Name	Discontinued Product	Excess Inventory	New Product	No Discount	Volume Discount
Cycling Cap				-7,439	-3,392
Full-Finger Gloves					-50
HL Mountain Frame				-20,638	
HL Road Frame				-87,300	
HL Touring Frame					-6,526
LL Road Frame				-46,264	-761
LL Touring Frame					-601
Long-Sleeve Logo Jersey				-89,851	-23,967
ML Road Frame-W				-75,212	-1,449
Mountain-100	-617,514				
Mountain-200				-129,180	-23,525
Mountain-500	-91,840				
Road-150				-26,790	
Road-250				-910,872	-31,426
Road-350-W				-188,596	-96,633
Road-450				-21,008	-1,400
Road-550-W				-446,212	-16,275
Road-650		-97,973		-514,992	-78,734
Road-750				-54,977	-9,549
Short-Sleeve Classic Jer..				-45,974	-29,873
Touring-1000			-547,905	-199,856	-81,883
Touring-2000				-56,287	-12,965
Touring-3000			-277,071	-56,042	-5,587

The next chart shows the totals of all profits and losses for each bike model, broken down by sales team. Of note:

- The Mountain-100 bikes were doing quite well until the steep Discontinued Product discounts kicked in.
- Mountain-300 and 400 bikes are lower volume, but neither had any losses.
- The Road-250 model resulted in a profitable sale exactly once.
- Only 2 teams managed to turn a (small) profit on the Road-650 model.
- The rest of the Road bike, along with the entire Touring line, produced only losses with no offsetting profits.

Wholesale Profit/Loss by Model

	Mountain-100		Mountain-200		Mountain-300	Mountain-400-W	Mountain-500		Road-150	Road-250		Road-350-W	Road-450	Road-550-W	Road-650		Road-750	Touring-1000	Touring-2000	Touring-3000
	Loss	Profit	Loss	Profit	Profit	Profit	Loss	Profit	Loss	Loss	Profit	Loss	Loss	Loss	Loss	Profit	Loss	Loss	Loss	Loss
285				250														-15,517	-918	-8,388
286				9,665			-962	858		-178								-84,601	-5,852	-26,502
274	-74,541	1,274	-45	3,768		126		733							-17			-2,005	-157	-1,145
278	-24,354	22,784	-4,214	11,920	4,163	838	-2,404	2,738	-1,484	-31,131		-18,994	-2,403	-21,227	-59,275		-3,851	-29,153	-2,755	-15,839
282	-32,552	50,801							-3,820				-4,114		-60,241					
289			-25,294	101,763	22,847	8,257	-10,338	15,176		-144,526		-33,633		-66,680	-38,641		-8,753	-66,167	-3,254	-25,213
287			-452	250	595					-335		-248					-79	-3,033	-210	-1,096
290			-6,572	44,238	5,005	1,635	-3,606	5,245		-45,249		-31,550		-22,690	-21,543		-5,171	-85,669	-8,744	-36,635
287														-122		77	-39	-10,491	-682	-5,061
288				14,458		880	-3,366	2,176		-4,002	721	-3,220		-2,202		3,390	-1,002	-79,335	-5,356	-38,414
282			-9,743	40,746	8,475	3,521	-9,857	6,403		-42,141		-18,614		-18,711	-14,034		-4,748	-64,394	-6,160	-29,504
287			-815	5,665	694	251		1,001		-7,200		-495		-3,012	-1,082		-197	-5,142	-499	-2,626
274		3,186	-1,223	5,524	1,982	754	-3,847	746	-438	-14,436		-3,220	-119	-9,212	-9,247		-1,396	-4,069	-289	-305
275	-36,561	28,013	-16,456	77,748	15,504	4,108	-14,906	11,611	-3,455	-162,761		-34,947	-4,330	-76,271	-117,290		-8,401	-30,639	-2,204	-7,043
276	-123,189	92,817	-25,494	102,110	25,747	12,838	-22,599	17,130	-2,117	-113,179		-47,027	-1,864	-53,897	-71,341		-5,327	-71,451	-6,953	-26,488
277	-29,802	76,353	-16,954	43,629	19,576	3,269	-5,770	8,320	-5,791	-142,047		-45,315	-3,847	-70,158	-123,541		-10,860	-24,937	-2,965	-16,616
279	-85,348	96,235	-11,980	40,478	11,944	5,072	-2,404	7,641	-3,236	-74,627		-15,231	-2,241	-36,478	-66,210		-4,717	-39,782	-3,569	-18,572
280		44,571	-7,767	38,377	6,959	3,353	-4,568	3,630	-2,458	-7,803			-1,279	-4,626	-23,576			-27,300	-3,621	-18,034
281	-180,004	48,347	-13,734	24,738	9,713	4,317	-962	4,161	-2,774	-86,619		-16,841	-1,596	-41,692	-59,903		-5,484	-86,641	-4,790	-29,740
283	-31,162	45,958	-9,858	35,734	7,830	2,683	-3,125	5,358	-1,217	-27,425		-6,359	-615	-15,486	-19,906		-1,160	-47,402	-4,906	-13,364
284			-2,102	11,799	2,627	1,132	-3,125	1,620		-38,637		-9,535		-20,024	-5,852		-3,341	-51,915	-5,368	-18,116

It is possible to drill down further, but we have found the most important information: Once the 40% wholesale discount is applied, non-Mountain bikes always sell for less than their cost.

Recommendations

First we should adjust our definition of Expected Profitability to account for possible discounts, wholesale and otherwise. This will allow us to set our prices and projections more accurately. The formulation of a proper, statistics based Expected Profitability formula is beyond the scope of this ~~class~~ paper.

The optimum solution to this problem would be to lower the cost of bike manufacture enough to generate a profit at current prices. This would keep both kinds of customer happy, but external realities could render this goal impossible.

If costs cannot be lowered, prices must be raised, and/or the wholesale discount must be reduced. In this case, I recommend lowering the effective wholesale discount by raising the list price and offering discounts on the web store that keep the true retail prices approximately where they are now. This may result in fewer wholesale sales, but that's still an improvement from the current situation.

I also recommend bringing back the popular and profitable Mountain-100 and 500 bikes. This will likely incur costs associated with restarting production, but it will be worth it to get two of our best products back.

Problem Solving Skills

The first step in solving a problem is to clearly define the problem and the goal, as it is important to make sure that I truly understand what the problem is. Proper problem framing results in proper communication with the client, which in turn produces proper results. It's not enough to know what I'm looking for, I need to know why I'm looking for it. Knowing why allows me to identify and respond to a greater variety of available inputs. I'll also need a general overview of the business processes that are related to the problem. At that point, it's time to get to know the data, starting with the data dictionary and ERDs of any section of the database that looks possibly relevant.

Data exploration also helps with the next part of the process, which is to understand the context and the circumstances of the problem. The structure and contents of the database will reveal additional details about the business and the problem. It is likely that I will have follow up questions during this process. I will then search out any other information that seems relevant to the problem. Once I have all the data, it's time to start hunting for the problem, starting with the most zoomed-out view. I will then drill down a step at a time, searching for patterns and similarities at each stage and evaluating each instance in turn. This will continue until the root cause of the problem is found.

After that it's time to formulate and evaluate possible solutions. I will start by listing ways that the data inputs could be adjusted to fix the problem. These ideas will then be evaluated for realism by listing the pros and cons of each and a recommendation will be formed.

Once a solution has been chosen and implemented, it's time to evaluate. The new data will be compared with the old data, and the results compared to the expected results. A new analysis may be performed at this time in order to determine if the problem persists or if new problems have occurred.

Technical Skills

ERD and Data Dictionaries
This semester began with the data dictionary, a documentation research that will, hopefully, if properly maintained, list every property and detail for each table, view, procedure, function, etc in the database. We then created an ERD of the Products table and all of its related tables, giving us a visual representation of the portion of the database that we were working with.

Join and Where Clauses
On week 3 we started into the SQL. We started with another ERD before learning how to format output, which is important because pretty numbers are also easier to read. We then used WHERE and JOIN functionality to familiarize ourselves with the products and their groupings. The most important thing I learned all semester was the processing order for SQL queries.

Math and Aggregation
We built on the previous week's SQL query by using multiple joins to retrieve all of our desired information at once. We then used Math functions to create calculated fields, and aggregation functions such as COUNT. We then sorted the output based on various criteria. Importantly, we learned how to sort based on calculated fields, which can be tricky due to processing order.

Reporting
We learned how to use SSRS in Visual Studio to generate reports on SQL Server databases. This included basic functionality like connecting to a data source and generating a query with and without the wizard. The overlap between building reports in SSRS and MS Access appears to be considerable.

Case and Concatenate Clauses
Case statements are just If/Then statements for SQL, and concatenation is a familiar function as well, so using those was just a matter of learning the syntax. We also got some more practice with the COUNT function.

Complex Math and Aggregation
This was similar to the previous Math and Aggregation lesson, but included more complex concepts like using calculations that rely on other calculations. We also go more practice using ORDER BY with calculations and aggregations.

Excel Integration and Visualization
We started by creating an ERD relating to sales orders and examining each table. We then created a query to retrieve or calculate thirteen fields. The results were then pasted into Excel. We then made a half-dozen pivot tables and used these to create bar charts.

Data Warehouse Implementation
This lesson detoured into an overview of data warehousing. The primary focus was on the differences between DW and OLTP data structures.

Subqueries and Row to Column Pivot
The bulk of this assignment was creating the SQL query that was ultimately used to create the business analysis section of this paper. A set of subquery was included in order to retrieve related data without multiplying the number of results. The data was then moved to Excel where we explored the relationship between customer type and the total profit of categories, subcategories and products.

Summary
SQL skills: Queries, subqueries, calculated fields, data formatting, query formatting, aggregation, concatenation, functions, case statements, nested case statements, math algorithms, aliases, joins, filtering, aggregation, output sorting.

Data Warehousing: Data structure and use cases

Documentation: Entity-relationship diagrams and data dictionaries

SSRS: Data source connection, data formatting, report generation, report formatting

Excel: Data query import, calculated fields, pivot tables, pivot charts

Tableau: Pretty pictures