

# Business Intelligence

## Assignment 2 - OLAP Cube

### Group Members

- Hamna Inam Abro (27113)
- Zara Masood (26928)
- Zuha Aqib (26106)

### Table of Contents

Group Members .....	1
Task 1.....	2
Task 2.....	2
Task 5.....	3
Facts .....	3
Dimensions.....	3
Useless Variables .....	3
Task 6.....	4
Business Problem #1: Hamna Inam Abro .....	4
Business Problem #2: Zara Masood.....	9
Business Problem #3: Zuha Aqib.....	13
Task 7.....	16

## Task 1

Select a dataset of your choice from the provided list at the end of this document.

<https://www.kaggle.com/prachil3/customer-analytics>

#	A	B	C	D	E	F	G	H	I	J	K	L
1	ID	Warehouse_block	Mode_of_Shipment	Customer_care_calls	Customer_rating	Cost_of_the_Product	Prior_purchases	Product_importance	Gender	Discount_offered	Weight_in_gms	Reached.on.Time_Y.N
2	1	D	Flight	4	2	177	3	low	F	44	1233	1
3	2	F	Flight	4	5	216	2	low	M	59	3088	1
4	3	A	Flight	2	2	183	4	low	M	48	3374	1
5	4	B	Flight	3	3	176	4	medium	M	10	1177	1
6	5	C	Flight	2	2	184	3	medium	F	46	2484	1
7	6	F	Flight	3	1	162	3	medium	F	12	1417	1
8	7	D	Flight	3	4	250	3	low	F	3	2371	1
9	8	F	Flight	4	1	233	2	low	F	48	2804	1
10	9	A	Flight	3	4	150	3	low	F	11	1861	1
11	10	B	Flight	3	2	164	3	medium	F	29	1187	1
12	11	C	Flight	3	4	189	2	medium	M	12	2888	1
13	12	F	Flight	4	5	232	3	medium	F	32	3253	1
14	13	D	Flight	3	5	198	3	medium	F	1	3667	1
15	14	F	Flight	4	4	275	3	high	M	29	2602	1
16	15	A	Flight	4	3	152	3	low	M	43	1009	1
17	16	B	Flight	4	3	227	3	low	F	45	2707	1
18	17	C	Flight	3	4	143	2	medium	F	6	1194	1
19	18	F	Ship	5	5	227	3	medium	M	36	3952	1
20	19	D	Ship	5	5	239	3	high	M	18	2495	1
21	20	F	Ship	4	5	145	3	medium	M	45	1059	1
22	21	A	Ship	3	3	161	2	medium	F	38	1521	1
23	22	B	Ship	3	1	232	4	medium	F	51	2899	1
24	23	C	Ship	2	5	156	2	low	M	2	1750	1
25	24	F	Ship	4	3	211	3	high	M	12	3922	1
26	25	D	Ship	4	5	251	2	medium	F	28	3561	1

## Task 2

Study the columns thoroughly in the selected dataset. You may do your research to understand the data better.

- **ID:** ID Number of Customers.
- **Warehouse\_block:** The Company have big Warehouse, which is divided in to block such as A, B, C, D, E.
- **Mode\_of\_Shipment:** The Company Ships the products in multiple way such as Ship, Flight and Road.
- **Customer\_care\_calls:** The number of calls made from enquiry for enquiry of the shipment.
- **Customer\_rating:** The company has rated from every customer. 1 is the lowest (Worst), 5 is the highest (Best).
- **Cost\_of\_the\_Product:** Cost of the Product in US Dollars.
- **Prior\_purchases:** The Number of Prior Purchase.
- **Product\_importance:** The company has categorized the product in the various parameters such as low, medium, high.
- **Gender:** Male and Female.
- **Discount\_offered:** Discount offered on that specific product.
- **Weight\_in\_gms:** It is the weight in grams.
- **Reached.on.Time\_Y.N:** It is the target variable, where 1 Indicates that the product has NOT reached on time and 0 indicates it has reached on time.

## Task 5

**Identify the facts, dimensions, and useless variables.**

### Facts

- Cost\_of\_the\_Product
- Discount\_offered
- Weight\_in\_gms
- Customer\_care\_calls
- Customer\_rating
- Prior\_purchases
- Reached.on.Time\_Y.N

### Dimensions

- Warehouse\_block
- Mode\_of\_Shipment
- Product\_importance
- Gender

### Useless Variables

- ID (Simply identifies each row uniquely, adds nothing to the analysis)

## Task 6

For each problem, create a process diagram which explains your drill-down process of solving this problem (NB: drill-down is an essential step in BI to solve a problem). You may make use of MS Word's SmartArt Feature or any tool of your choice. Since this diagram is the path to the solution, it carries most marks.

### Business Problem #1: Hamna Inam Abro

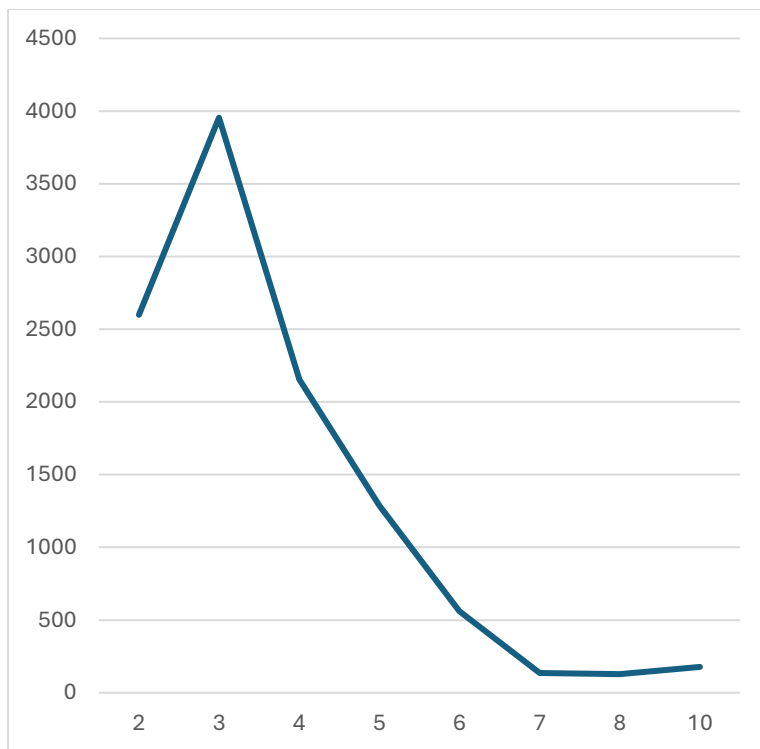
#### What is the business problem?

The business problem I am focusing on is understanding customer behavior after multiple purchases and identifying factors that impact long-term engagement and retention. I aim to determine whether discounts, delivery performance, or customer support influence customer disengagement and what strategies can help improve retention

#### How did you execute the drill down?

I started by analyzing prior purchases, customer ratings, and the count of customers to understand general purchasing trends. I noticed that the number of customers decreases as prior purchases increase, which led me to investigate further.

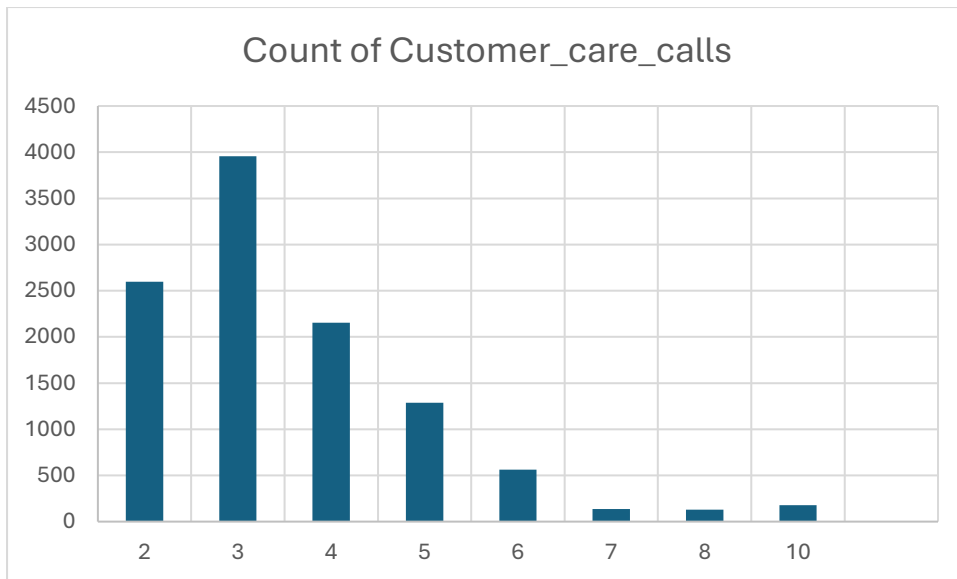
Count of ID	Customer ratings					Grand Total
	1	2	3	4	5	
Prior purchases						
2	563	515	498	520	503	2599
3	775	791	829	787	773	3955
4	436	392	450	427	450	2155
5	275	265	259	236	252	1287
6	104	116	113	123	105	561
7	26	32	32	26	20	136
8	24	25	26	28	25	128
10	32	29	32	42	43	178
Grand Total	2235	2165	2239	2189	2171	10999



The graph clearly shows that after 3 prior purchases, engagement drops significantly and continues decreasing. By 6+ prior purchases, very few customers remain engaged.

To investigate this decline in customer engagement, I did the following:

1. Examined customer care interactions and discovered that customer care calls decrease as prior purchases increase, indicating that dissatisfied customers may not report their issues before leaving.



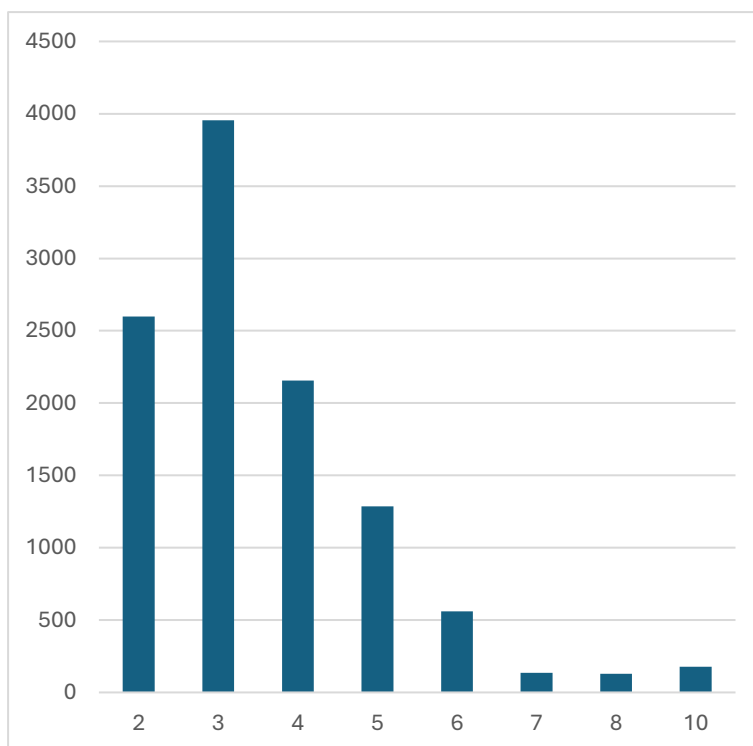
Prior purchases	Count of ID	Count of Customer_care_calls
2	2599	2599
3	3955	3955
4	2155	2155
5	1287	1287
6	561	561
7	136	136
8	128	128
10	178	178
<b>Grand Total</b>	<b>10999</b>	<b>10999</b>

2. Looked at discount patterns and found that:

- Customers with 7+ prior purchases received higher discounts (15-65%) compared to newer customers.
- Despite these discounts, engagement still declined, meaning discounts may not fully prevent churn.

3. Analyzed delivery performance for these customers and found that many of their deliveries were late, suggesting a possible reason for disengagement.

Prior purchases	Count of ID
<b>7</b>	<b>136</b>
0	44
1	92
<b>8</b>	<b>128</b>
0	45
1	83
<b>10</b>	<b>178</b>
0	76
1	102
<b>Grand Total</b>	<b>442</b>



## What did you identify about the business?

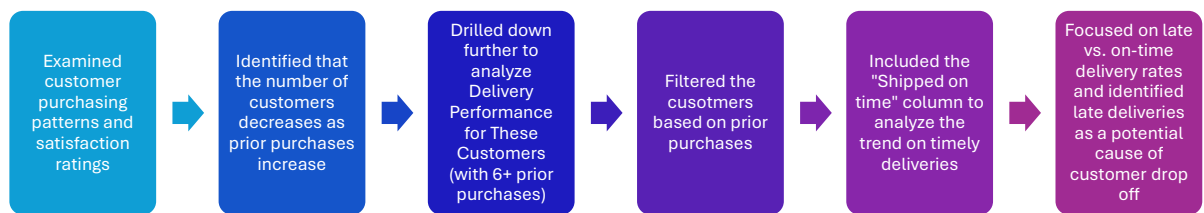
From my analysis, I identified several key insights:

1. Late deliveries could be a major factor in customer disengagement after six purchases.
2. Customer care calls decrease over time, meaning customers may be silently leaving without expressing dissatisfaction.
3. Discounts help retain some customers, but they are not a long-term solution—they delay disengagement rather than preventing it.

## What Impacts Do Your Identifications Imply?

1. If delivery performance does not improve, customers will continue to disengage despite discounts.
2. Since fewer customers report their dissatisfaction, the business lacks visibility into churn reasons, making it harder to proactively address concerns.
3. Discounts alone are not enough—other retention strategies like loyalty programs and improved service are needed to sustain engagement.

### Drill down diagram:





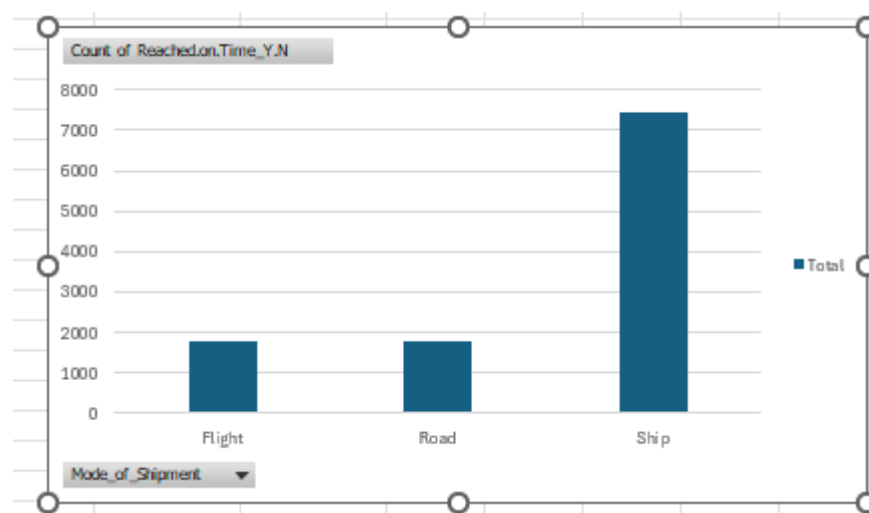
## Business Problem #2: Zara Masood

### What is the business problem?

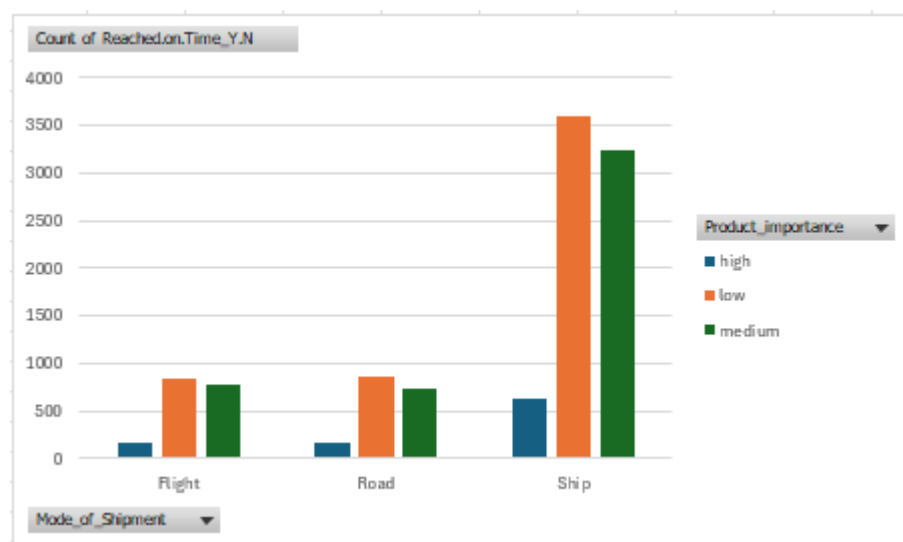
The business problem I am focusing on is identifying the best performing warehouse block, along with the mode of shipment and optimize the shipping and delivery process to ensure that the products reach the customers on time, improving customer satisfaction. Including discount offered, will help us understand how discounts impact on-time delivery performance, and determine if certain warehouse blocks are more effective in leveraging discounts to improve on-time delivery.

### How did you execute the drill down?

I started with the lowest, base-level, retrieving the count of on-time products based on each of the Mode of Shipments (*Flight*, *Road*, *Ship*). Clearly, Shipping the product turned out to be the most timely shipment method.

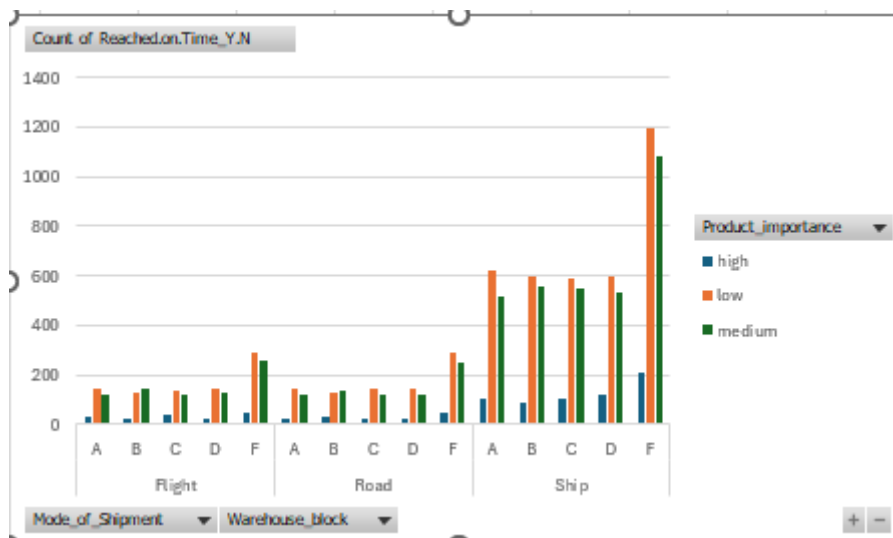


Next, I added the "Product Importance", at the legend (column), to understand the distribution of on-time products, based on product importance, grouped by the 3 modes of shipment.

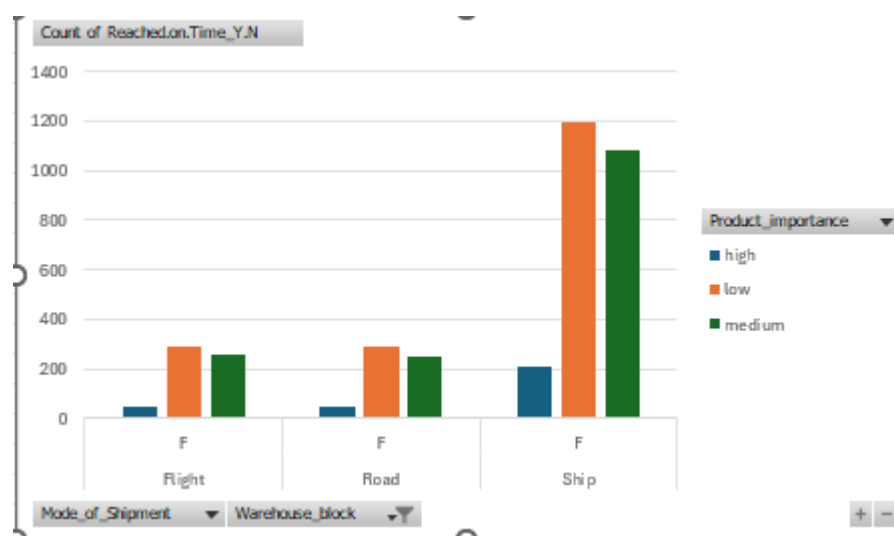


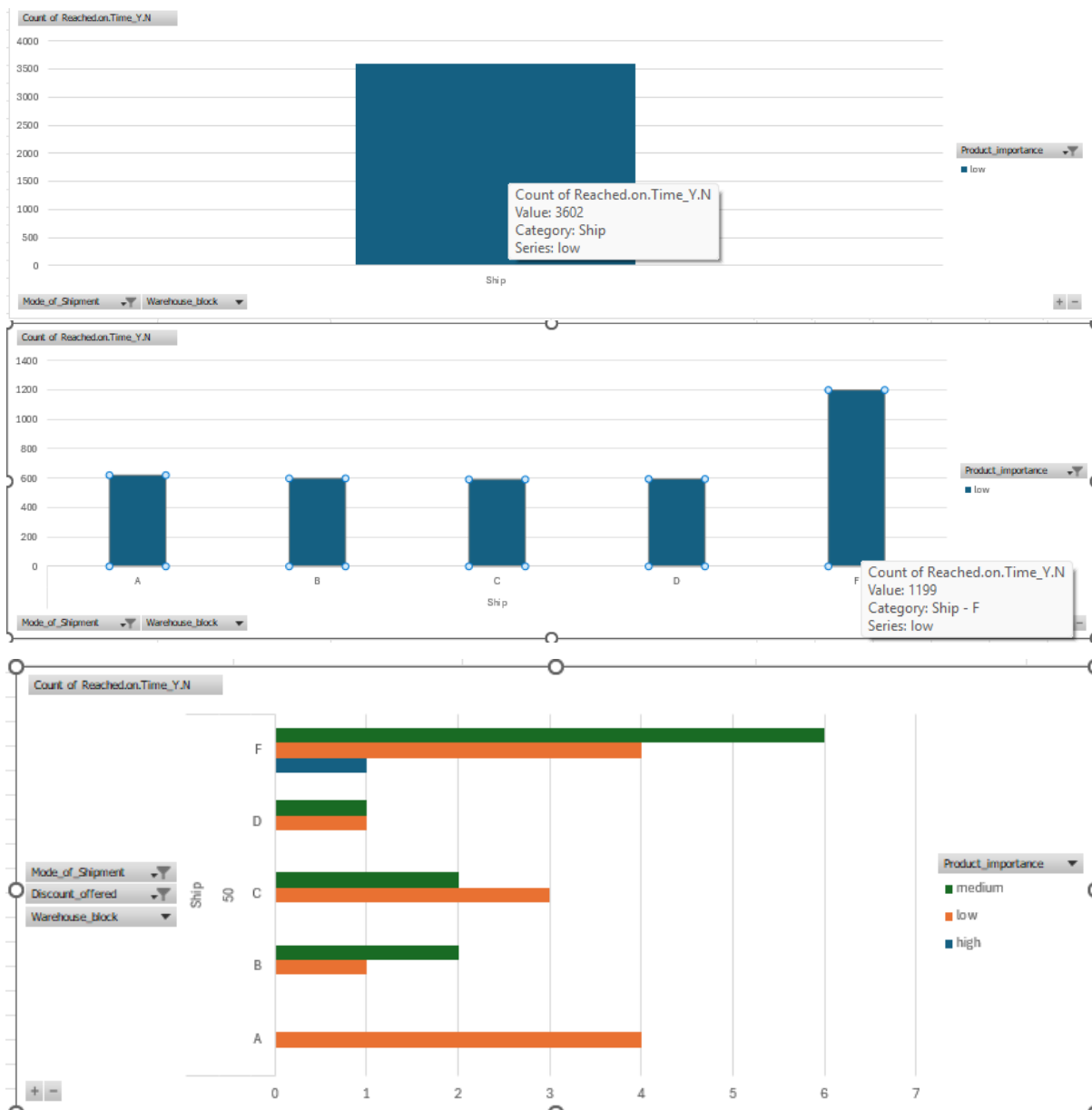
The chart showed that Shipping dominated for the 3 categories of product importance (high, low, medium), that reached on time. Especially the low importance products.

I added the warehouse blocks as the 2<sup>nd</sup> axis of the pivot table, which gave me the following chart:



Via Shipping, Warehouse Block F, produced the most products that reached on time. Importance-wise, the low-importance products, shipped through shipping were the highest in number. This is shown by filtering the chart to just display for Warehouse Block F. I have attached the following 3 figures to explain how I drilled down to find this insight.





To analyse further, I added Discount offered as one of the Axes. Particularly filtering out at Value = 50, this showed that Warehouse F, offered discount on all the 3 product importance categories, while the highest count of products reached on time, this time was dominated by "medium", particularly via Shipping. All the Warehouses offered discounts on the low importance categories.

### What did you identify about the business?

- The best performing shipping method was Ship along with the warehouse block F that delivered the most products on time. These were majorly products with lower product importance.
- Ship is the most efficient shipping method for delivering products on time, outperforming Flight and Road.

- Warehouse Block F offered discounts across all three product importance categories (high, medium, low), and it still maintained the highest count of on-time deliveries.
- When filtering for discount = 50, medium-importance products shipped via Ship had the highest count of on-time deliveries.

### What Impacts Do Your Identifications Imply?

- Ship should be prioritized for time-sensitive deliveries, especially for low-importance products.
- The low-importance products are easier to manage and deliver on time, possibly due to lower handling complexity or fewer customer expectation thus are delivered on time mostly out of the 3 categories.
- Discounts may incentivize customers to accept slightly longer delivery times or reduce the pressure on warehouses to expedite shipments, leading to better on-time delivery performance.
- Medium-importance products are more responsive to discounts in terms of delivery performance, possibly because they are neither too critical (high importance) nor too low-priority (low importance).

### Drill Down Diagram:



## Business Problem #3: Zuha Aqib

### What is the business problem?

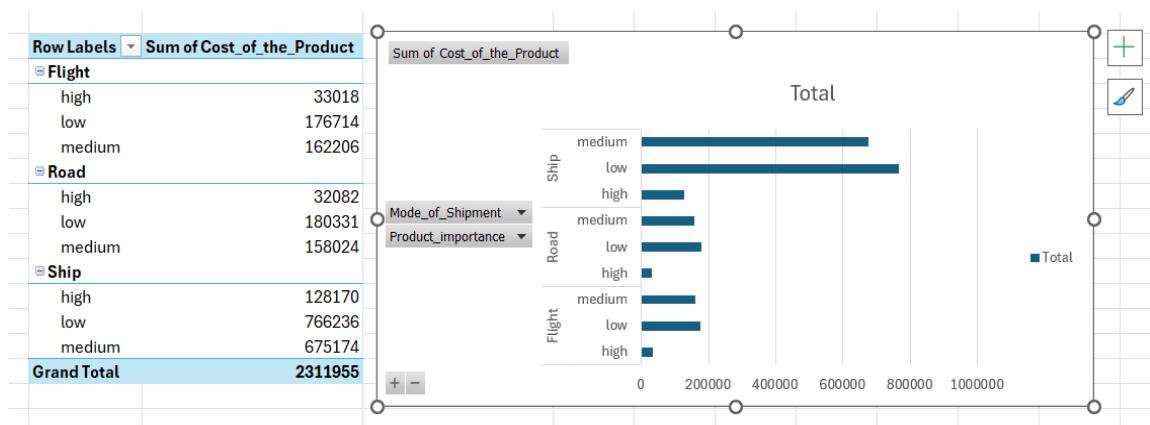
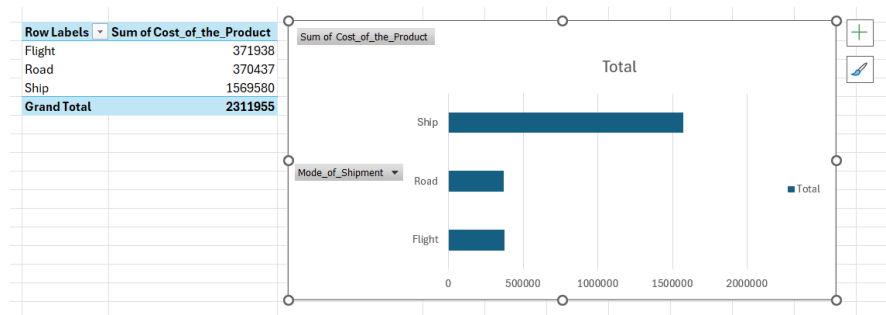
This is an analysis of transport. This dataset identifies 3 types of shipping mode,

- Flight
- Road
- Ship

By identifying which mode is the most used, and at which warehouse, we can then identify which products are mostly shipped using that mode and whether they are finally delivered on time. This will help us identify whether this most used shipping mode is effective or not - is the product being delivered on time or not? Why was it delayed? Is this mode the only one being delayed or are they all delayed? Let's dive into it.

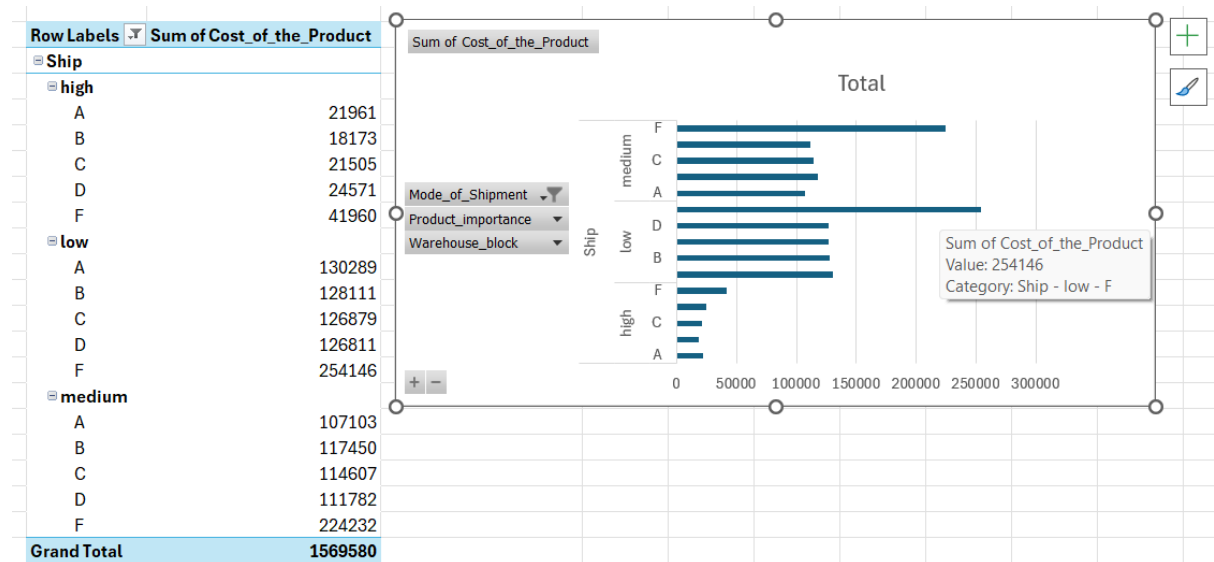
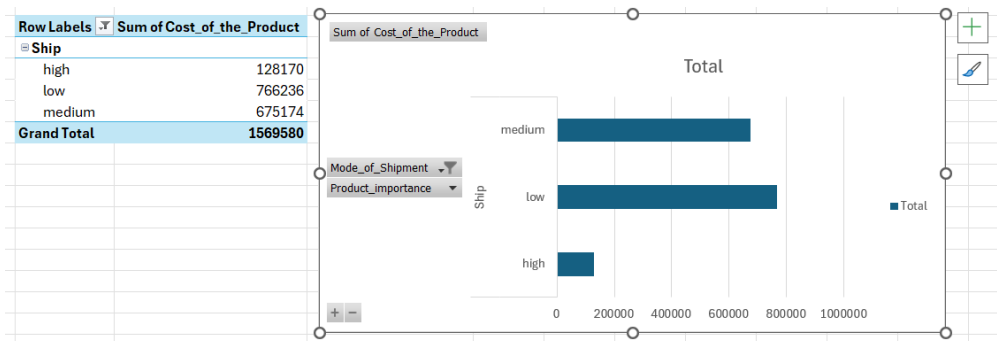
### How did you execute the drill down?

So first we saw that according to our sales, the "Ship" mode was the most used shipping mode. It was nearly thrice as much as road and flight.

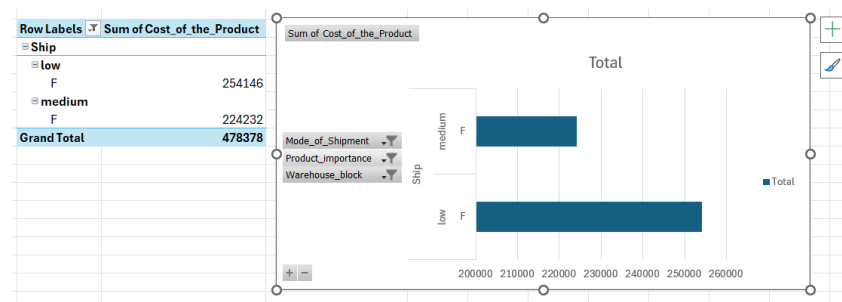


Then we saw that mostly "medium" priority or "low" priority goods were transported using the "Ship" mode of shipping - nearly three or four times more.

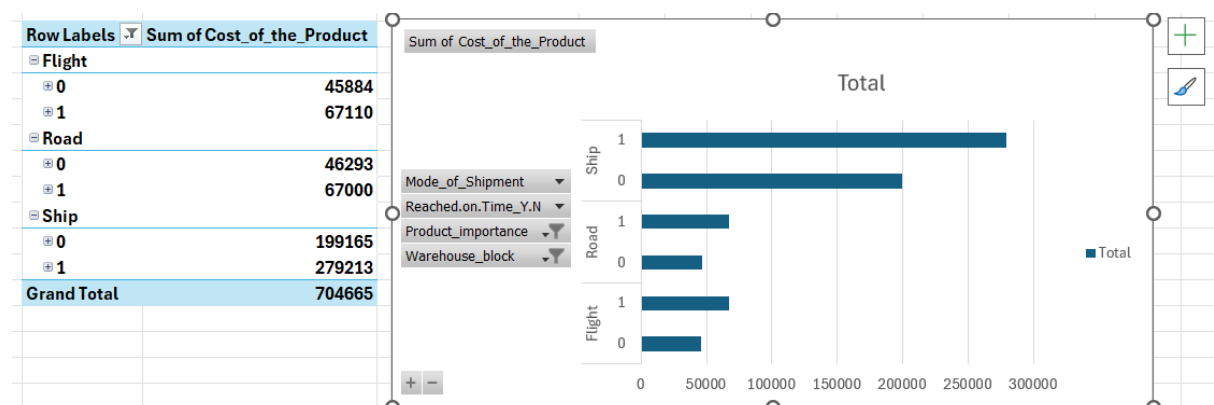
Then we filtered out the "Ship" in the shipment mode



So we see that "warehouse F" is the most selling than others (approximately double). In all shipping modes, "warehouse F" is the most selling.



So then we filtered out the "low", "medium" priority products and then "warehouse F"



Then we brought back all the shipping modes and added the target variable of "reached\_on\_time" - but we moved it up and drilled down to see each mode with its "reached\_on\_time". We saw that while "Road" and "Flight" mode had negligible difference between the reached-on-time and not-reached-on-time, however "Ship" has a large difference, nearly 60% is not reaching-on-time, which is **bad**, it deteriorates our customer service and reputation.

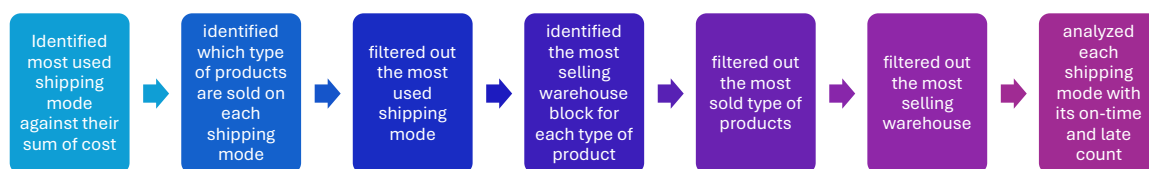
### What did you identify about the business?

The business has a good sale in "Warehouse F", and "Ship" is the most used mode of Shipment. Mostly "low" and "medium" priority products are transported using the "Ship" mode. However, this most used mode comes with its cons. Nearly 60% of shipments sent using the "Ship" mode is late, and not on time.

### What impacts do your identifications imply?

- As the "Ship" mode has products shipping majority late, customers prefer to send their "low" or "medium" priority products using that method. It could be that the "Ship" mode is the cheapest, but we cannot say this without proper data.
- The late shipping can have a negative impact on the business, as it ruins the image of not meeting deadlines.

### Drill Down Diagram



This is the drill down diagram, which is also reflected in the drill down process (with pictures) above.

## Task 7

**Additionally, think in terms of business recommendations and make a note about it. That note should be directed to the key decision makers of the company who will get a summarized view of the problem and recommended solution. (You may be required to do some research on your end to understand the business dynamics)**

1. Introduce proactive customer engagement, such as post-purchase check-ins or personalized outreach when engagement drops.
2. Redesign the discount strategy to focus on loyalty rewards, exclusive perks, or free shipping instead of generic percentage-based discounts.
3. Track silent disengagement patterns using data analytics to identify customers at risk of leaving and intervene before they churn.
4. Since Ship is the most efficient method for on-time delivery, allocate more low- and medium-importance products to this shipping method. While reserve Flight for only high-importance products that require urgent delivery.
5. Analyse the operational processes of Warehouse Block F to identify best practices (for e.g., efficient inventory management, optimized logistics). Implement these practices in other warehouse blocks (A, B, C, D) to improve their on-time delivery performance.
6. Investigate why other warehouse blocks (A, B, C, D) have lower on-time delivery rates. Possible issues could include:
  - a. Longer distances to customers.
  - b. Inefficient inventory management.
  - c. Lack of coordination with shipping carriers.
7. Since medium-importance products showed the highest on-time delivery performance when discounts were offered, consider targeting discounts primarily on these products.
8. We can use discounts as a tool to improve customer satisfaction and loyalty for medium-importance products
9. Because the "Ship" mode has nearly three times more sales than the other two modes, it should also be allocated three times more time, money and effort to improve it and make it more profitable.
10. The business should analyze the number of days it takes to "Ship" a product and quote an average of those days, perhaps maybe a day or two more than that, so that in the customer eyes, the business is "on-time".
11. The business should also investigate possible reasons for the delay in shipment - they could notify customers in



delays in case of weather issues beforehand or find a better, more reliable shipping company.

12. The business could also invest more into the other two shipping modes: "Road" and "Flight". Why do people not use them? Are they too costly when compared to the price of the product? Is there any way they can be cheaper? Our graphs show that even in both these modes the product is delivered late more than on time. Why is this a common trait for all shipping modes? Is it an internal issue that orders are processed late and they all reach (majority) late?