

## Dataset Source.

Search Code

ARPIT MISHRA · UPDATED 2 YEARS AGO 59 Code Download ⋮

# AmazonSalesReport

In this project I am doing basic EDA on Amazon Sales .



Data Card Code (20) Discussion (1) Suggestions (0)

## About Dataset

### Amazon Sales Report

#### Overview:

This dataset provides detailed sales data from Amazon, offering a comprehensive look at various product categories and their performance over time. It includes information on sales figures, order details, product categories, and customer demographics.

#### Features:

- 1. Order ID**  
Description: A unique identifier for each order placed on Amazon. This field helps to track individual orders and link related records.
- 2. Dates**  
Description: The date when the order was placed. This field is crucial for analyzing sales trends over time and identifying seasonal patterns.
- 3. Status**  
Description: The current status of the order (e.g., Shipped, Delivered, Pending). This field provides insight into the order fulfillment process and helps monitor order processing efficiency.
- 4. Fulfillment**

Usability: 10.00  
License: Attribution 4.0 International (CC BY)  
Expected update frequency: Never  
Tags: Business, Tabular, Transportation, Beginner, Social Science

Search Code

AmazonSalesReport 59 Code Download ⋮

Data Card Code (20) Discussion (1) Suggestions (0)

over time. It includes information on sales figures, order details, product categories, and customer demographics.

#### Features:

- 1. Order ID**  
Description: A unique identifier for each order placed on Amazon. This field helps to track individual orders and link related records.
- 2. Dates**  
Description: The date when the order was placed. This field is crucial for analyzing sales trends over time and identifying seasonal patterns.
- 3. Status**  
Description: The current status of the order (e.g., Shipped, Delivered, Pending). This field provides insight into the order fulfillment process and helps monitor order processing efficiency.
- 4. Fulfillment**

Never  
Tags: Business, Tabular, Transportation, Beginner, Social Science, E-Commerce Services, pandas, Matplotlib

View more

Promoted first row as headers and changed the type of columns.

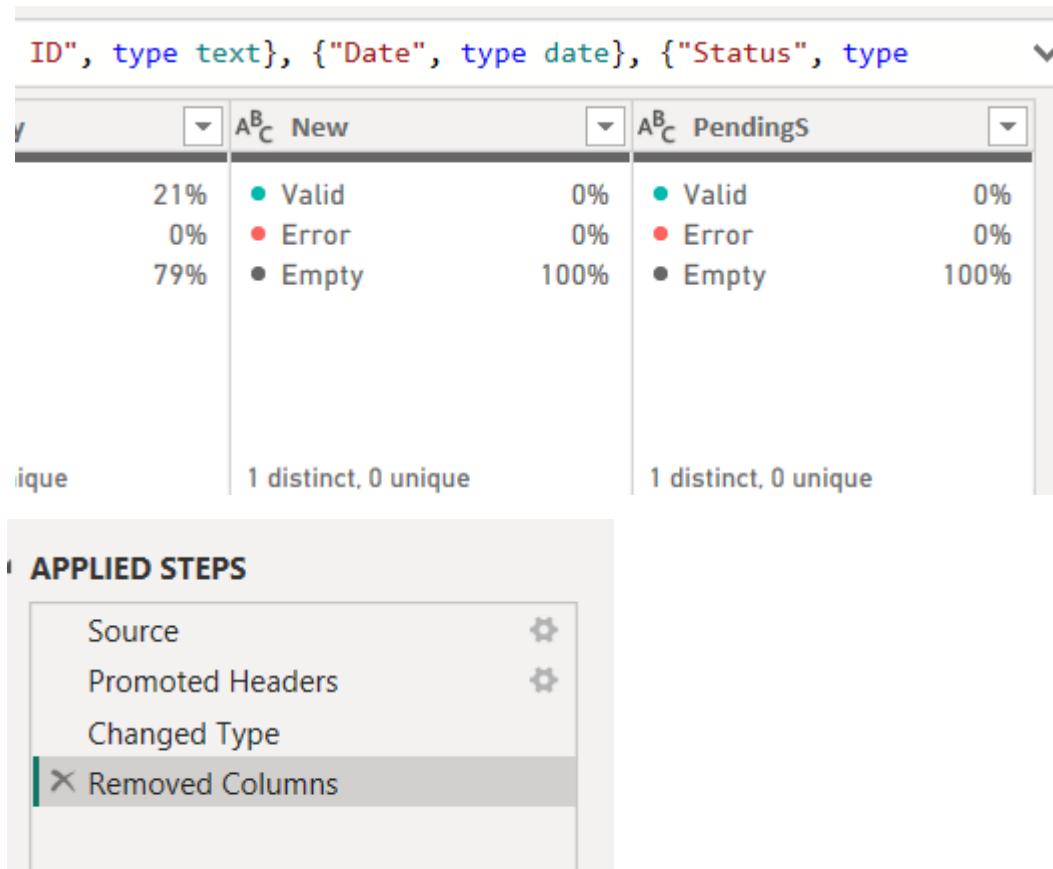
Table Any Column Date Column Number Column Date & Time Column Scripts

Queries [1] Amazon Sale Report Table.TransformColumnTypes("#Promoted Headers", [{"index": 1, "type": "text"}, {"index": 2, "type": "date"}, {"index": 3, "type": "text"}])

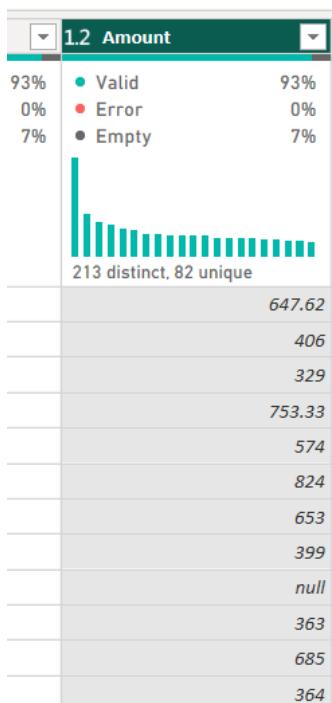
Properties Name: Amazon Sale Report  
Applied Steps: Promoted Headers, Changed Type

index	Order ID	Date	Status	Fulfillment	Sales Channel	ship-service
1	405-8078784-5731345	4/30/2022	Cancelled	Merchant	Amazon.in	Standard
2	171-9198151-1101146	4/30/2022	Shipped - Delivered to Buyer	Merchant	Amazon.in	Standard
3	204-0687676-7273146	4/30/2022	Shipped	Amazon	Amazon.in	Expedited
4	3403-9615377-8133951	4/30/2022	Cancelled	Merchant	Amazon.in	Standard
5	407-1089790-7240320	4/30/2022	Shipped	Amazon	Amazon.in	Expedited

Removed empty columns



Replaced null values in amount to 0 as they correlate to canceled orders

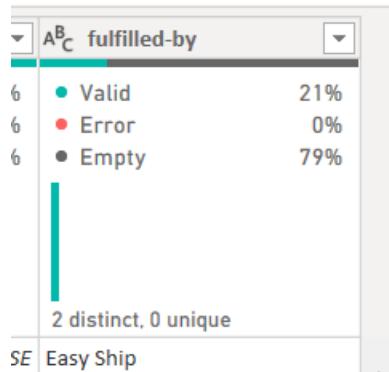


All Properties

APPLIED STEPS

- Source
- Promoted Headers
- Changed Type
- Removed Columns
- Replaced Value

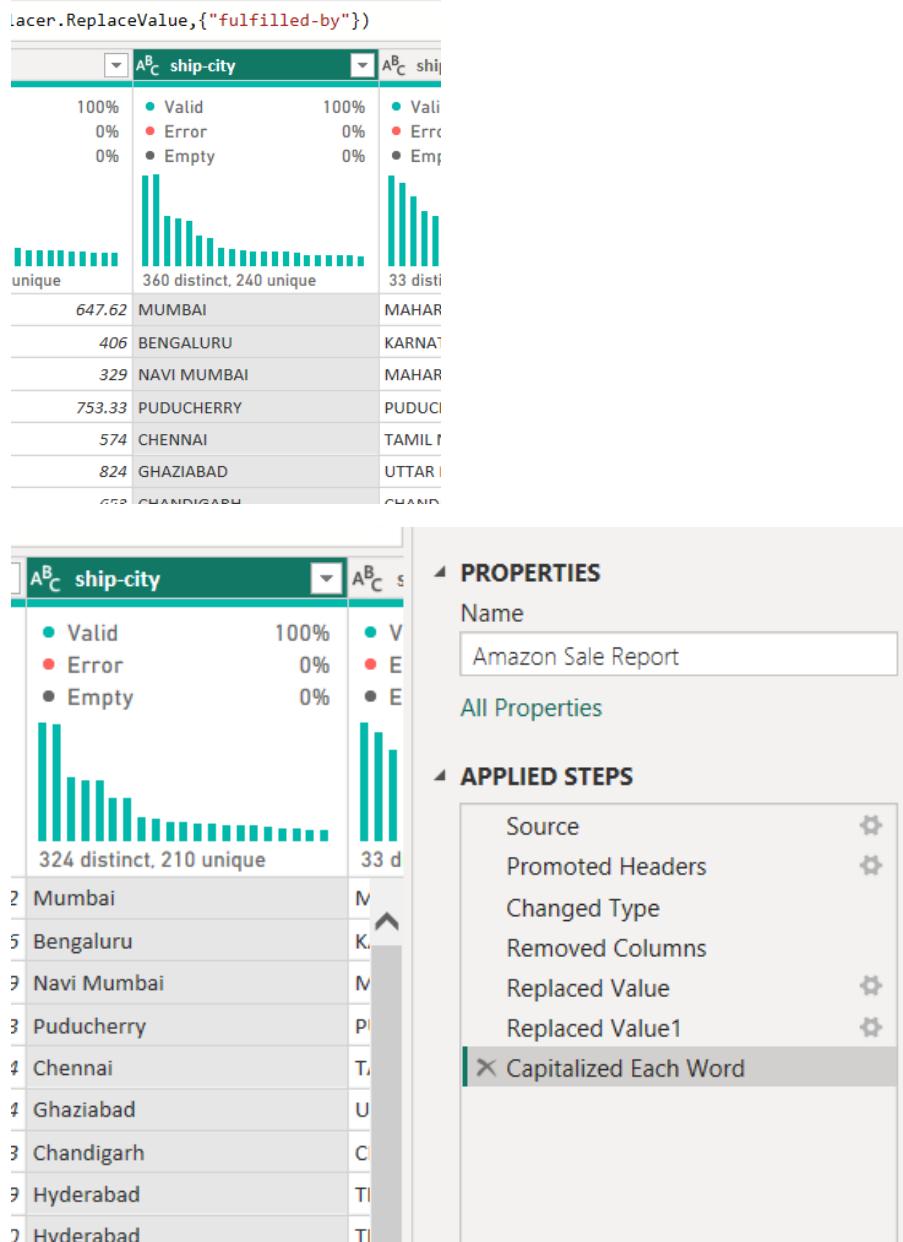
Changing nulls in the fulfilled by column with Amazon



APPLIED STEPS

- Source
- Promoted Headers
- Changed Type
- Removed Columns
- Replaced Value
- Replaced Value1

Capitalized each word in the ship city column



Inserted year and month name columns

All Properties

#### APPLIED STEPS

Source	⚙️
Promoted Headers	⚙️
Changed Type	⚙️
Removed Columns	⚙️
Replaced Value	⚙️
Replaced Value1	⚙️
Capitalized Each Word	⚙️
Inserted Year	⚙️
Inserted Month Name	⚙️

Inserting conditional column

Add Conditional Column

Add a conditional column that is computed from the other columns or values.

New column name: Order\_Value\_Tier

Column Name: If	Operator: is greater than	Value: ABC 123	Output: High Value	...
Amount	is greater than	1000	ABC 123	

Add Clause

Else: ABC 123

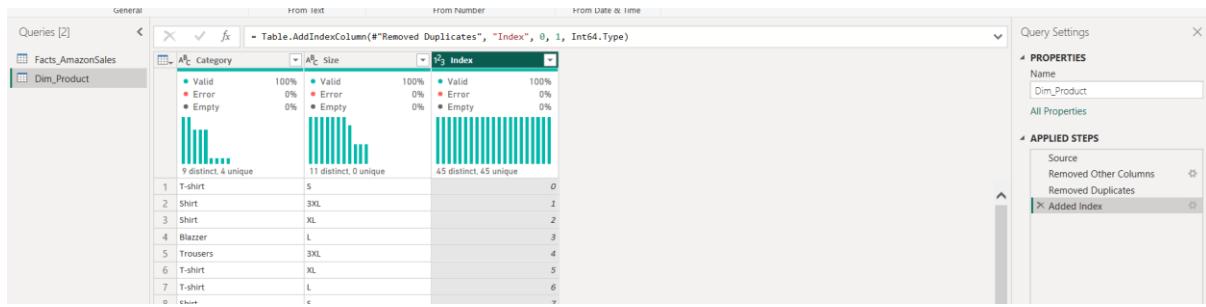
OK Cancel

The screenshot shows the Power BI Data Editor interface. On the left, there's a preview pane with a table titled "Order\_Value\_Tier". The table has three columns: "Order\_Value" (with values "Valid", "Error", and "Empty"), "Tier" (with values "100%", "0%", and "0%"), and a third column. Below this is a list of rows, all of which have the value "Low" in the first column. On the right, there are two sections: "PROPERTIES" and "APPLIED STEPS". In the "PROPERTIES" section, the "Name" is set to "Amazon Sale Report". In the "APPLIED STEPS" section, a list of steps is shown, with the last step, "Added Conditional Column", highlighted.

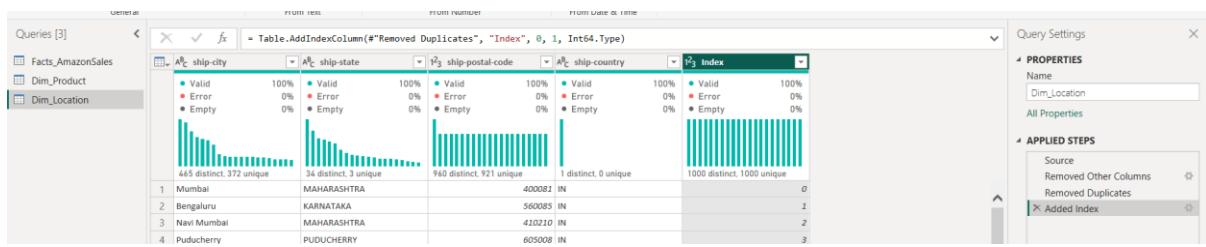
Renamed the query

The screenshot shows the Power BI Query Editor. The title bar says "Queries [1]". Below it, there is a list containing one item: "Facts\_AmazonSales".

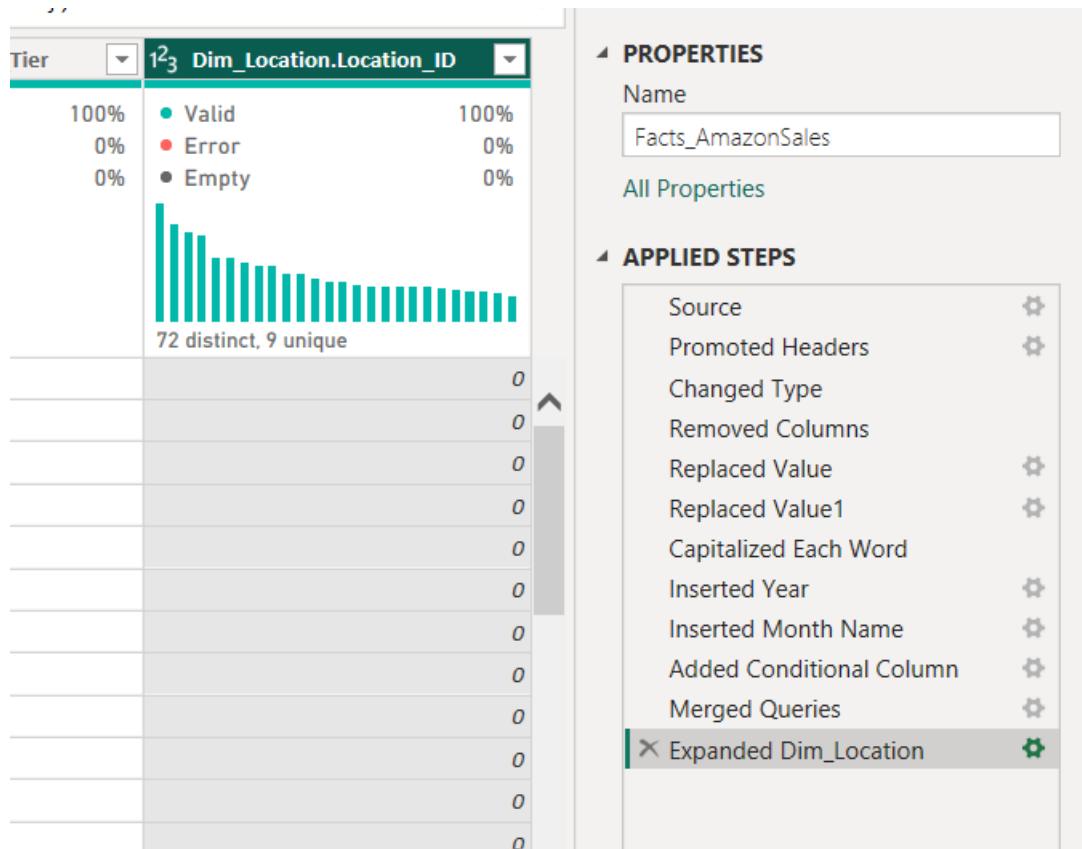
Created dim\_product table by referencing facts\_amazon\_table and removed all column except category and size. And then removed duplicates and then added and index column



Created dim\_location table by referencing facts\_amazon\_table and removed all column except ship city, ship state, ship postal code, and ship country. And then removed duplicates and then added and index column



Merged the dim\_location table to main table and expanded to keep location id for relationship in data modeling.



Merged the dim\_product table to main table and expanded to keep product id for relationship in data modeling.

**1<sup>2</sup> Dim\_Product.Product\_ID**

Category	Percentage
Valid	100%
Error	0%
Empty	0%

33 distinct, 2 unique

Value	Count
0	5
5	5
1	1
1	1
1	1
13	13
13	13
13	13

**PROPERTIES**

Name: Facts\_AmazonSales

[All Properties](#)

**APPLIED STEPS**

- Source
- Promoted Headers
- Changed Type
- Removed Columns
- Replaced Value
- Replaced Value1
- Capitalized Each Word
- Inserted Year
- Inserted Month Name
- Added Conditional Column
- Merged Queries
- Expanded Dim\_Location
- Merged Queries1
- Expanded Dim\_Product**

Now that I have the id columns I removed the redundant columns.

## Choose Columns

Choose the columns to keep

<input checked="" type="checkbox"/> (Select All Columns)
<input checked="" type="checkbox"/> index
<input checked="" type="checkbox"/> Order ID
<input checked="" type="checkbox"/> Date
<input checked="" type="checkbox"/> Status
<input checked="" type="checkbox"/> Fulfilment
<input checked="" type="checkbox"/> Sales Channel
<input checked="" type="checkbox"/> ship-service-level
<input checked="" type="checkbox"/> Category
<input checked="" type="checkbox"/> Size
<input checked="" type="checkbox"/> Courier Status
<input checked="" type="checkbox"/> Qty
<input checked="" type="checkbox"/> currency
<input checked="" type="checkbox"/> Amount
<input checked="" type="checkbox"/> ship-city
<input checked="" type="checkbox"/> ship-state
<input checked="" type="checkbox"/> ship-postal-code
<input checked="" type="checkbox"/> ship-country
<input checked="" type="checkbox"/> B2B
<input checked="" type="checkbox"/> fulfilled-by

Search Columns A  
Z

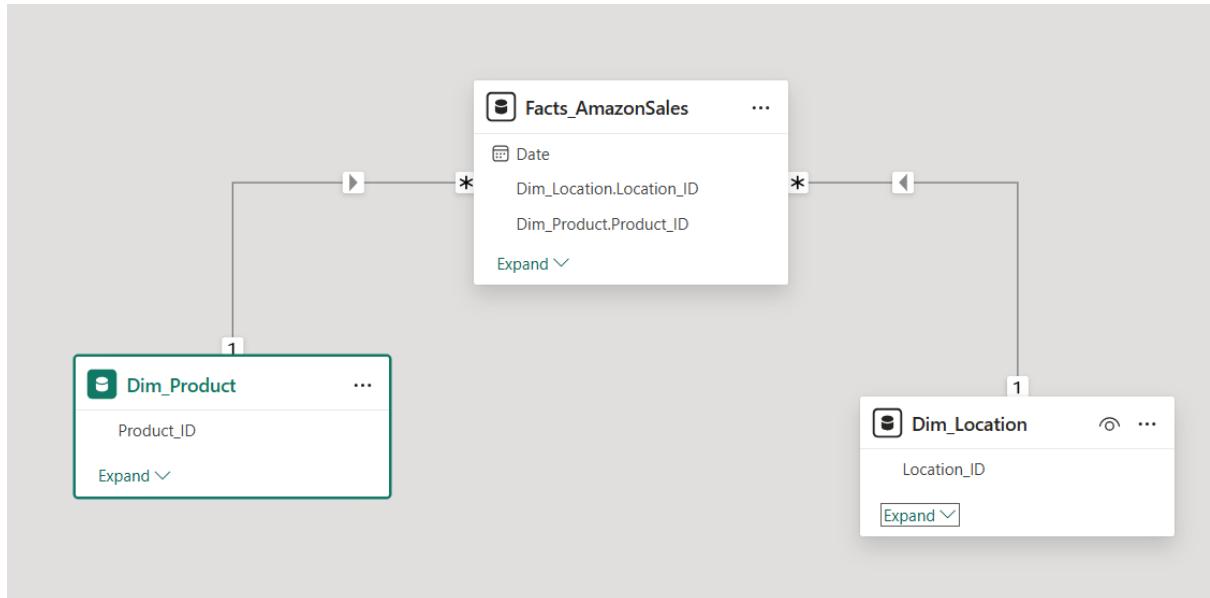
- Order ID
- Date
- Status
- Fulfilment
- Sales Channel
- ship-service-level
- Category
- Size
- Courier Status
- Qty
- currency
- Amount
- ship-city
- ship-state
- ship-postal-code
- ship-country
- B2B
- fulfilled-by
- Year
- Month Name

**APPLIED STEPS**

- Source
- Promoted Headers
- Changed Type
- Removed Columns
- Replaced Value
- Replaced Value1
- Capitalized Each Word
- Inserted Year
- Inserted Month Name
- Added Conditional Column
- Merged Queries
- Expanded Dim\_Location
- Merged Queries1
- Expanded Dim\_Product
- Removed Other Columns**

Data modeling.

Model view



Relationship between facts table and Dim\_product.

### Edit relationship

Select tables and columns that are related.

From table

Facts\_AmazonSales

Amount	B2B	Courier Status	currency	Date	Dim_Location...	Dim_Prod
399	false	Shipped	INR	Wednesday, ...	2	14
399	false	Shipped	INR	Friday, April 2...	15	14
399	false	Shipped	INR	Friday, April 2...	19	14

To table

Dim\_Product

Category	Product_ID	Size
T-shirt	0	S
Shirt	1	3XL
Shirt	2	XL

Cardinality

Many to one (\*:1)

Cross-filter direction

Single

Make this relationship active

Apply security filter in both directions

Assume referential integrity

**Save** **Cancel**

Relationship between facts table and din location

## Edit relationship



Select tables and columns that are related.

### From table

Facts\_AmazonSales



Amount	B2B	Courier Status	currency	Date	Dim_Location	Dim_Prod
399	false	Shipped	INR	Wednesday, ...	2	14
399	false	Shipped	INR	Friday, April 2...	15	14
399	false	Shipped	INR	Friday, April 2...	19	14

### To table

Dim\_Location



Location_ID	ship-city	ship-country	ship-postal-c...	ship-state	
766	Ernakulam	IN	682017	KERALA	
776	Ernakulam	IN	686664	KERALA	
1087	Ernakulam	IN	682039	KERALA	

### Cardinality

Many to one (\*:1)

### Cross-filter direction

Single



Make this relationship active

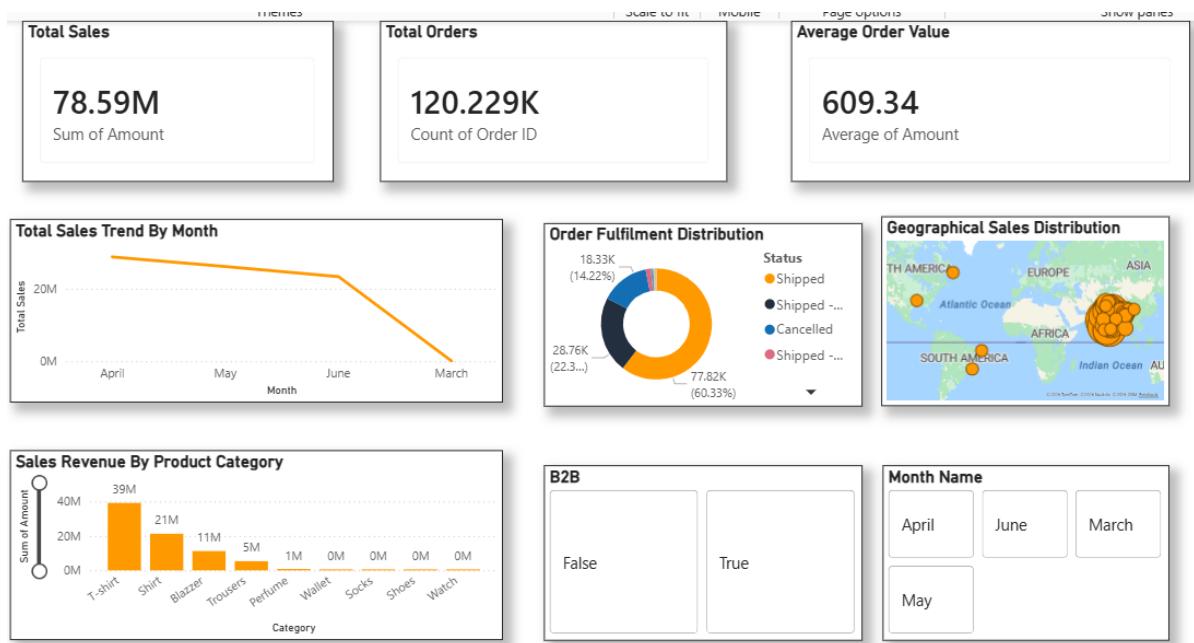
Apply security filter in both directions

Assume referential integrity

Save

Cancel

## Dashboard screenshot



<https://app.powerbi.com/view?r=eyJrljoiNWY1NDYxNTUtYjc5Ny00YTU3LWEyMjgtZTNlYWNlYzZkMTE4IiwidCI6IjE2ZDgzZWU2LTI1NGEtNDY5ZC1hNmNjLTU0ZTJjYTlzMjMTNlNyIsImMiOjh9>