

Power BI

Final project

NTI – data analysis track

Agenda



Dataset



Understanding the
business problem



Data Cleaning and
Transformation



Dax Measures



Key Questions



Dashboard

1.Dataset

- **1.1 What are the data**

This dataset has the data of 1K+ Amazon Product's Ratings and Reviews

- **1.2 source of data:**

From Kaggle

- **1.3 columns:**

Product Id, product name , category , actual price, discount price, discount%, rating, rating count, about product, user Id, user name , review id, review title, review content

2. Understanding the Business Problem

2.1 what are we trying to solve?

We are trying to solve some questions related to the revenue and top products, as we will see ISA.

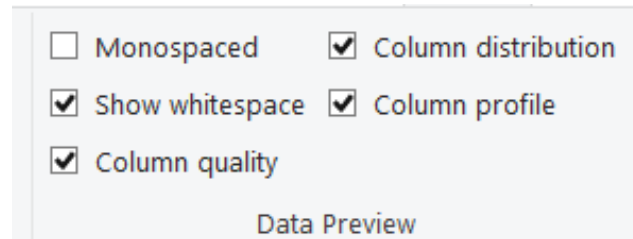
3. Data Cleaning and Transformation

- The process of data cleaning is almost one of the most important processes in the data analysis life cycle.
- Before Data cleaning there is an Important thing we must do this thing is data Exploration

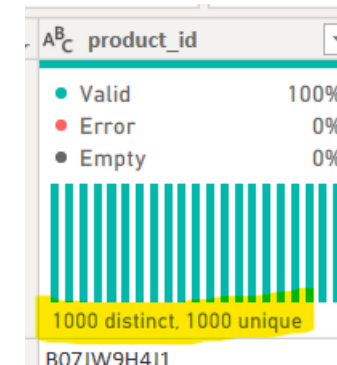
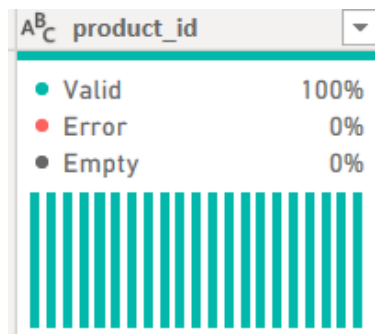


3.1 Data Exploration

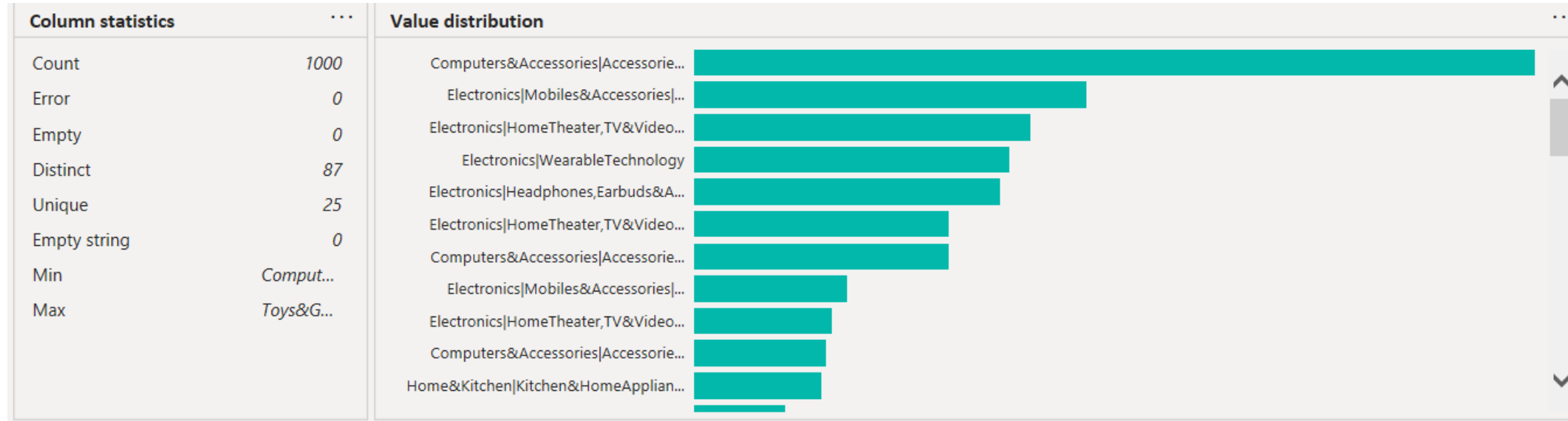
- After loading data into Power BI, I opened Power Query and saw the content of columns.
- Then there are a few buttons that can help you in the data exploration process.



- 1-Column Quality (Valid, Empty, or Error values) 2- column distribution (Unique & distinct values).



- 3-Column Profile (Statistics like min, max, and mean)



3.2 Data cleaning

- I removed unnecessary columns that were not relevant for data visualization and updated the data types



3.2 Data Transformation

- I added a new column for revenue by subtracting the discount price from the actual price to accurately reflect earnings



3.2 Data Transformation

Calculate Sub category from category column:

- The dataset does not include a separate subcategory field, but the category column contains hierarchical information in the following format:
Computers&Accessories | Accessories&Peripherals | Cables&Accessories | Cables | USB Cables
- Since the subcategory name appears at the end of the category string, I split the column to extract the subcategory

Split Column by Delimiter

Specify the delimiter used to split the text column.

Select or enter delimiter

--Custom--

|

Split at

☐ Left-most delimiter

☒ Right-most delimiter

☐ Each occurrence of the delimiter


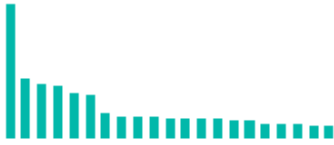
▸ Advanced options

Quote Character

"

☐ Split using special characters

Insert special character

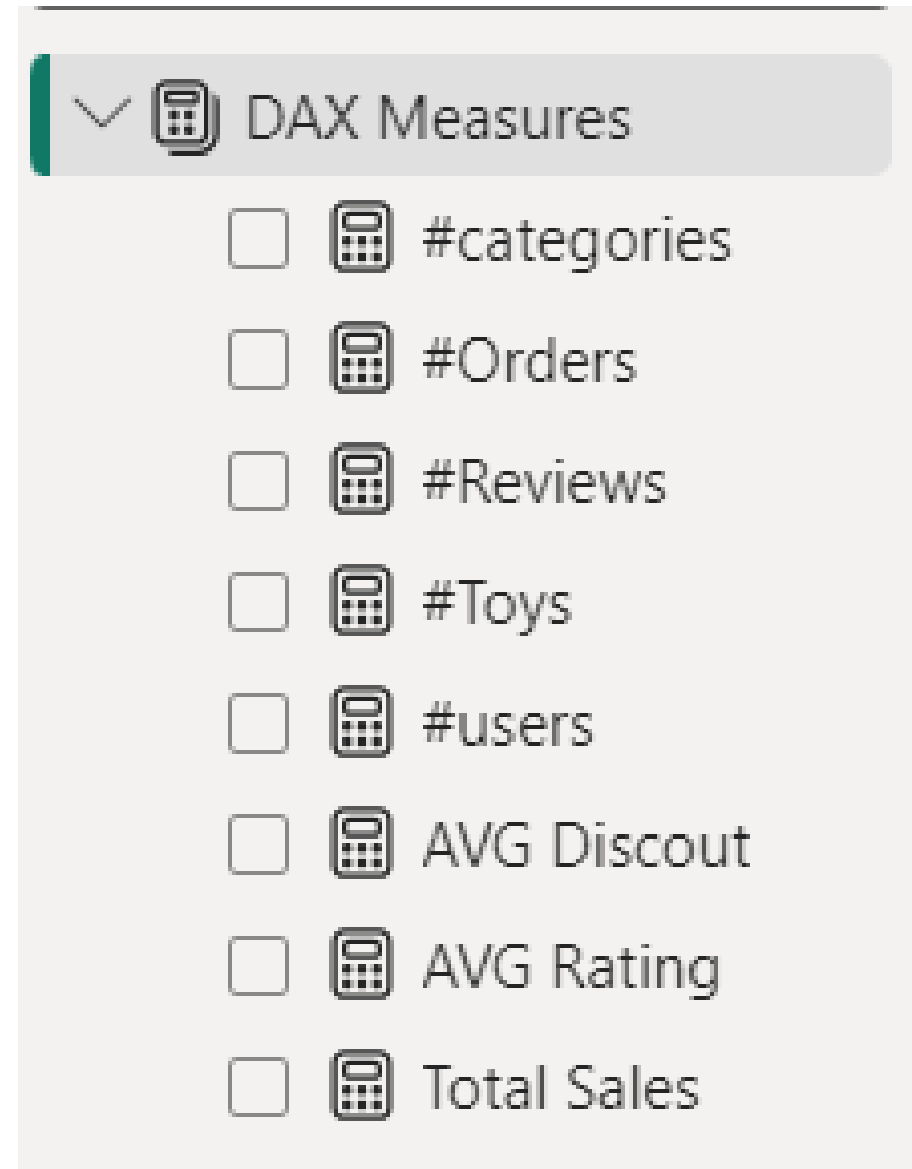
ABC category	ABC Sub_Category
<ul style="list-style-type: none"> Valid 100% Error 0% Empty 0%  <p>103 distinct, 31 unique</p>	<ul style="list-style-type: none"> Valid 100% Error 0% Empty 0%  <p>207 distinct, 72 unique</p>
Computers&Accessories Accessories&Peripherals Cables&Accessorie...	USBCables
Computers&Accessories Accessories&Peripherals Cables&Accessorie...	USBCables
Computers&Accessories Accessories&Peripherals Cables&Accessorie...	USBCables
Computers&Accessories Accessories&Peripherals Cables&Accessorie...	USBCables
Computers&Accessories Accessories&Peripherals Cables&Accessorie...	USBCables
Computers&Accessories Accessories&Peripherals Cables&Accessorie...	USBCables
Computers&Accessories Accessories&Peripherals Cables&Accessorie...	USBCables
Computers&Accessories Accessories&Peripherals Cables&Accessorie...	USBCables
Computers&Accessories NetworkingDevices NetworkAdapters	WirelessUSBAdapters
Computers&Accessories Accessories&Peripherals Cables&Accessorie...	USBCables
Computers&Accessories Accessories&Peripherals Cables&Accessorie...	USBCables
Computers&Accessories Accessories&Peripherals Cables&Accessorie...	USBCables
Electronics HomeTheater,TV&Video Accessories Cables	HDMICables
Computers&Accessories Accessories&Peripherals Cables&Accessorie...	USBCables
Computers&Accessories Accessories&Peripherals Cables&Accessorie...	USBCables
Computers&Accessories Accessories&Peripherals Cables&Accessorie...	USBCables

4. DAX measures

- I used DAX measures to calculate key metrics such as Total Sales and the Number of Categories.
- DAX (Data Analysis Expressions) enables dynamic aggregations in Power BI, allowing for calculations like SUM, COUNT, and other essential functions to analyze data efficiently



-
- So I created a new table and added DAX Measures



-
- One of the challenges I faced was calculating the number of occurrences for each category, given that the category column contained multiple hierarchical levels in the following format:

Toys&Games | Arts&Crafts | Drawing&PaintingSupplies

- To address this, I wrote the following DAX code:

```
1 Toys = CALCULATE(  
2     COUNT(amazon[category]),  
3     CONTAINSSTRING(amazon[category], "Toys"))
```


5. Key Questions

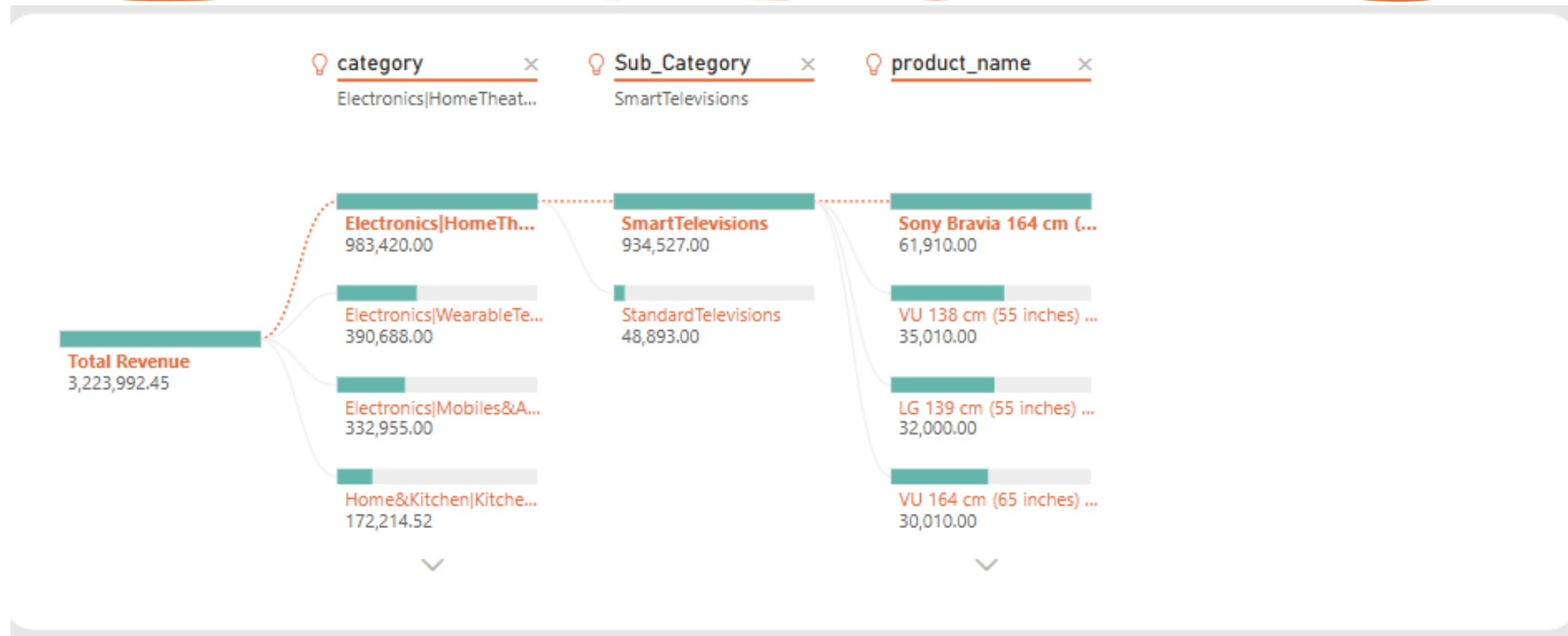
- 1- Which Category captured the market by their price or order quantity?

Which Category captured the market by their price ?

Electronics|HomeTheater,TV&Video|Televisions

- I determined this by applying a 'Top 1' filter on total revenue, identifying the highest-performing category

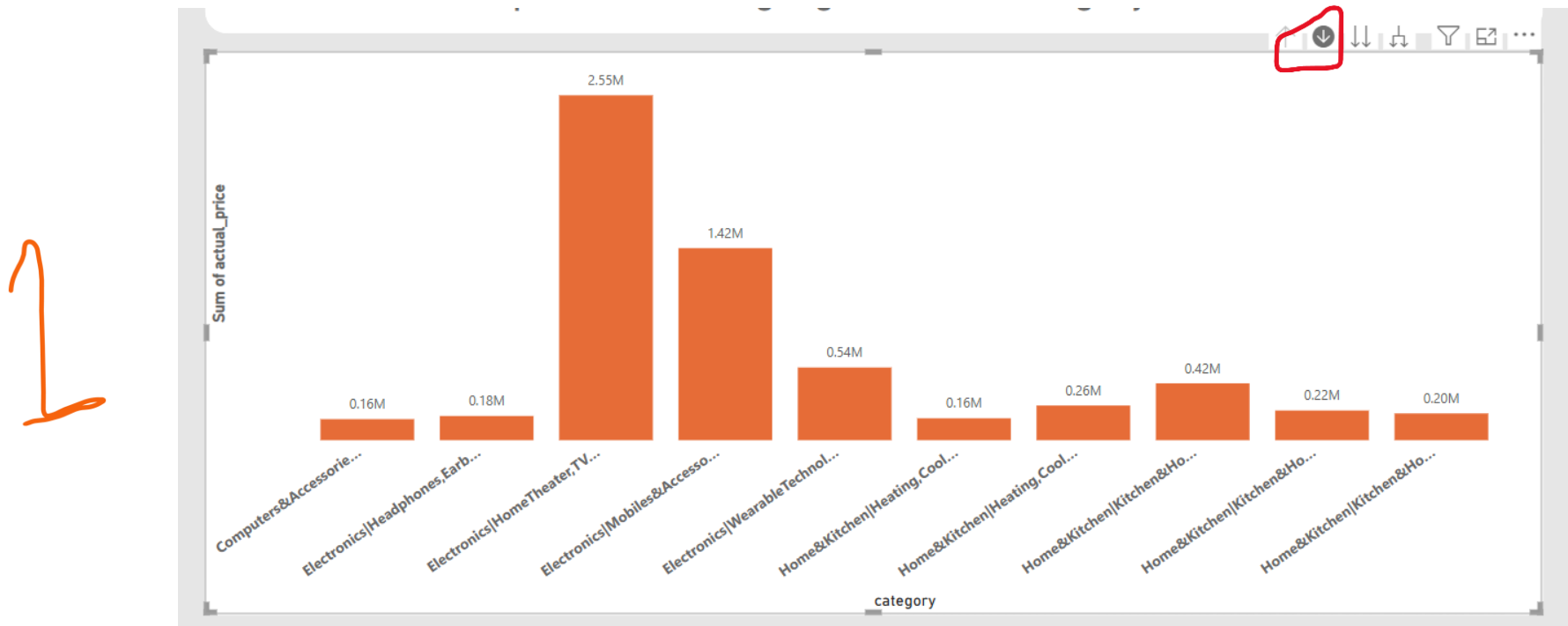
- 2-What is the total revenue they generate from the individual category and sub-category?



-
- The previous visual, the Decomposition Tree, is an excellent tool for analyzing revenue across categories, subcategories, and individual products.
 - It allows you to drill down into each category, revealing revenue distribution and identifying the top-performing products within each segment

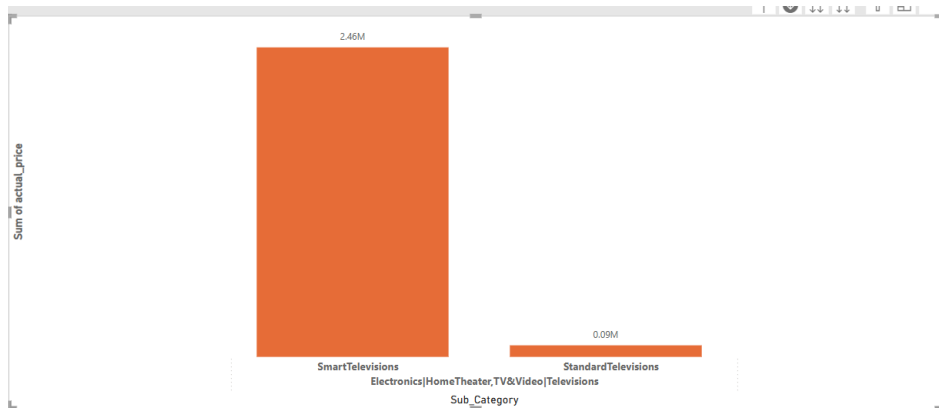
3-Price of the products having high demand category wise in market?

- To address this question, we first need to identify the top 10 categories and then analyze the pricing of products within those categories. The following steps, shown in the images, illustrate this process

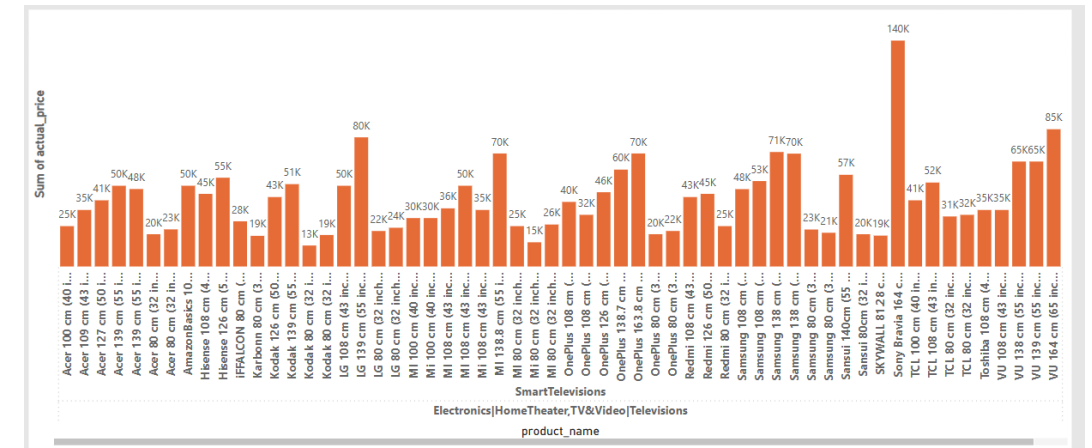


Click the dropdown button and select any category of interest.
The corresponding subcategories will then appear—click on any subcategory to view the related products.
Hover over a product to display its price using the tooltip feature

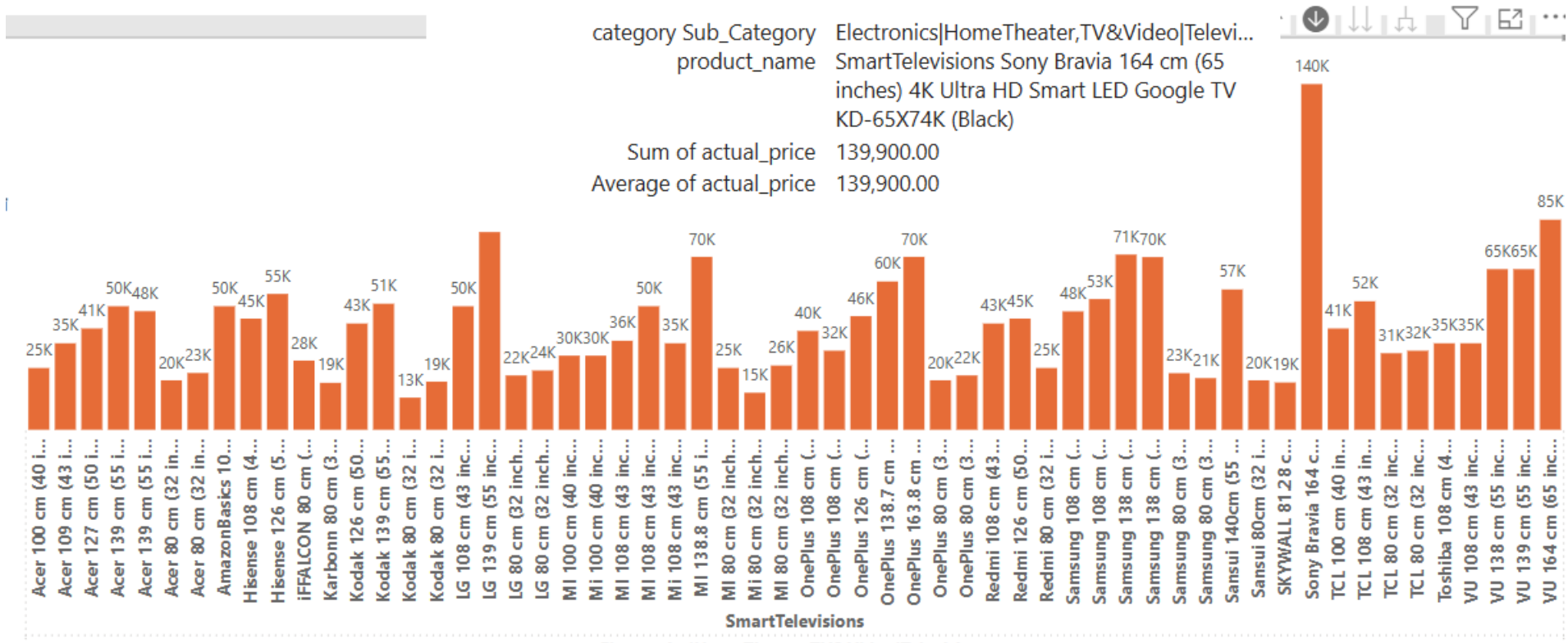
2



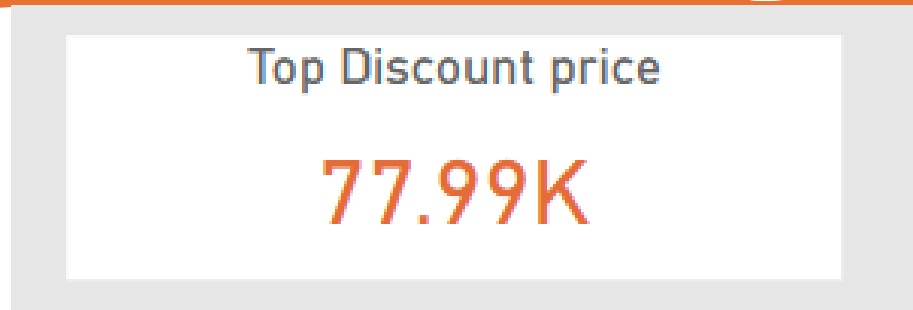
3



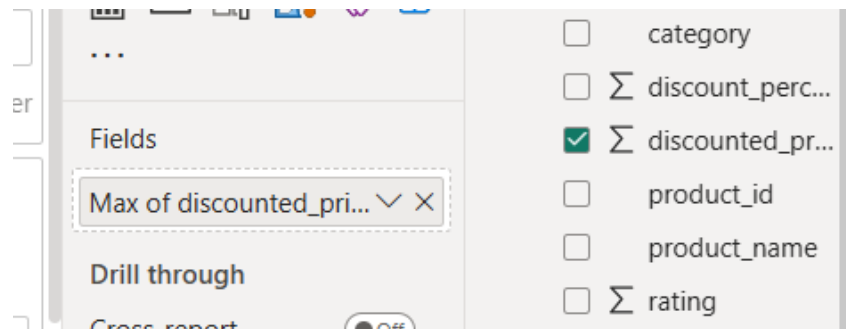
4



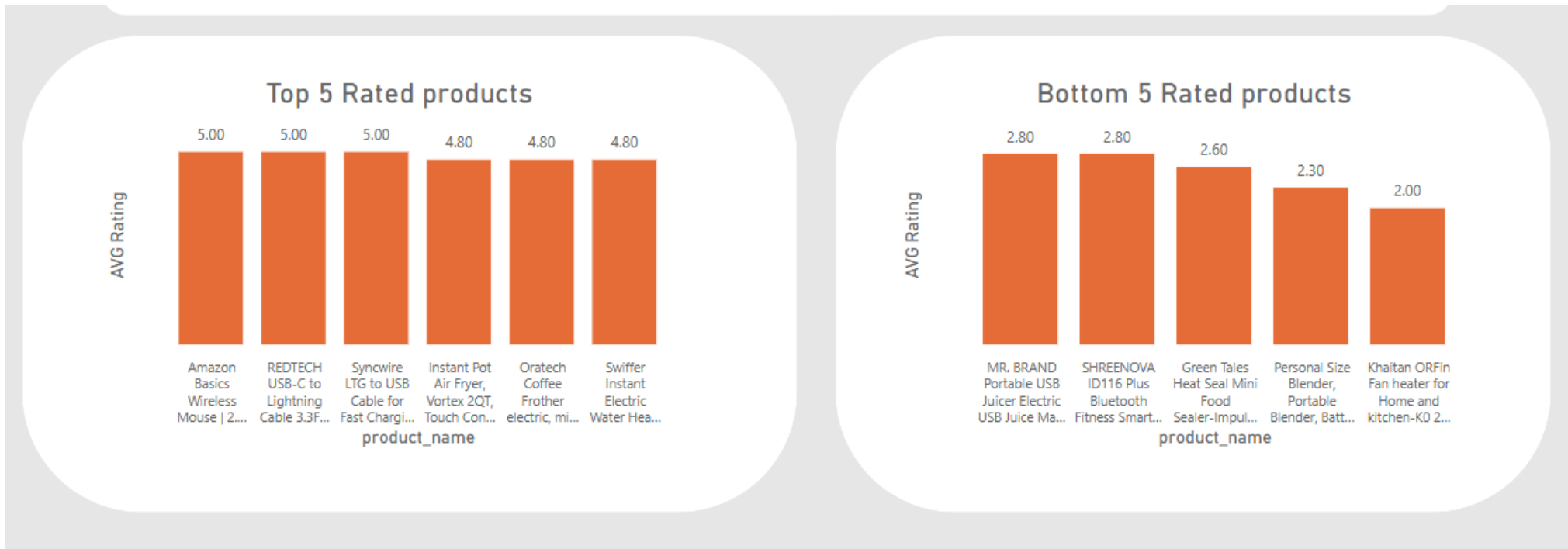
- 4- Top discount price?



- I calculated it by using the MAX function to determine the highest discount price, applying aggregation functions



- 5-What are the top and low-rated products, so we can see what needs to improve to increase the market's profit?



6. Whole Dashboard





\$8M

Total Sales

4.09

AVG Rating

3.30K

AVG Discout

1186

#users

103

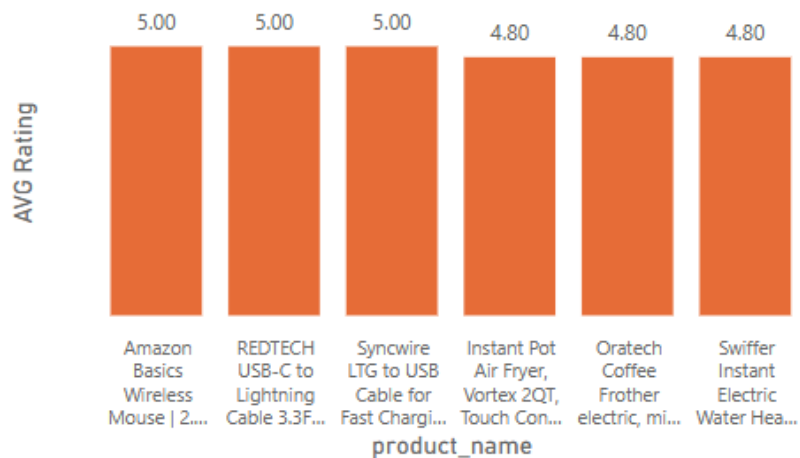
#categories

1351

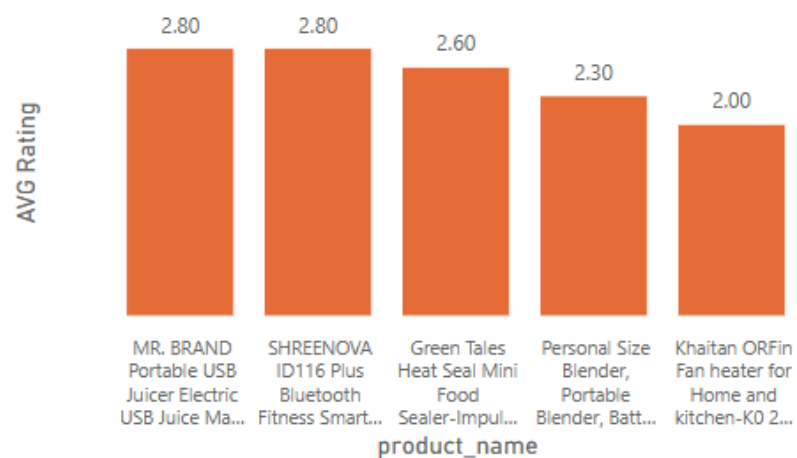
#Orders

top and low-rated products, so we can see what needs to improve to increase the market's profit.

Top 5 Rated products



Bottom 5 Rated products



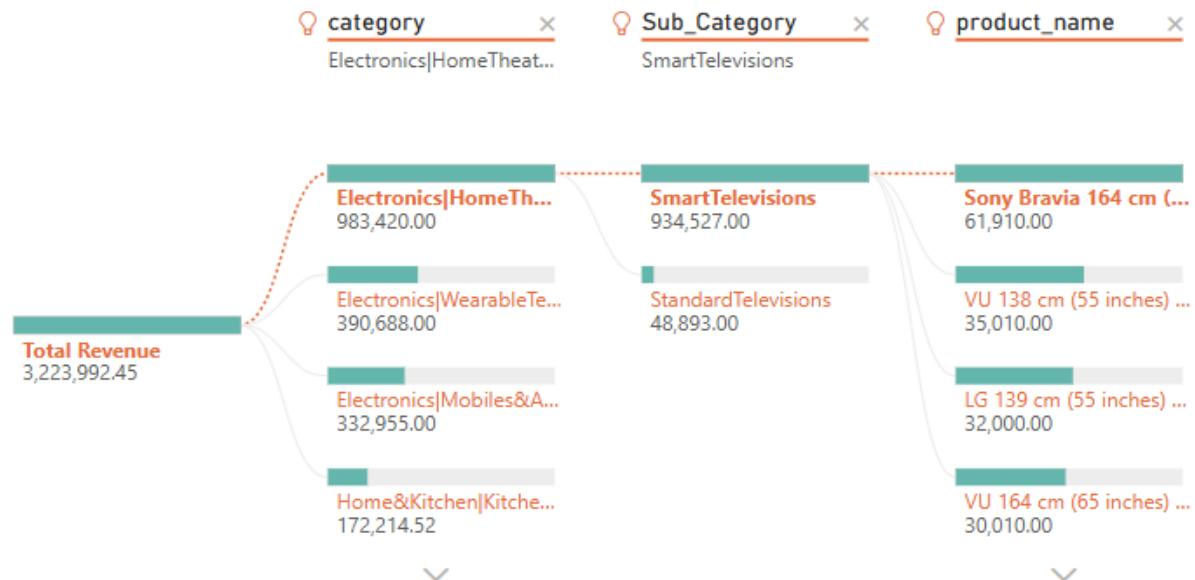
Which Category captured the market by their price ?

Electronics|HomeTheater,TV&Video|Televisions

Top Discount price

77.99K

total revenue they generate from the individual category , sub-category and products



Price of the products having high demand category wise in market.

