Marketing and Retail Analytics: Capstone Project

Content:

- Background on project
- Objectives
- Task-1: Preparation of Data
- Task-2:Creation of Tableau Dashboard
- Dashboard Snapshots and insights
- Market Basket Analysis
- Inferences
- Recommendations

BACKGROUND:

OList is an e-commerce company that has faced some losses recently and they want to manage their inventory very well so as to reduce any unnecessary costs that they might be bearing.

- need to manage the inventory cost of this e-commerce company.

- need to identify top products that contribute to the revenue and also use market basket to analyse the purchase behaviour of individual customers to estimate with relative   certainty

- What items are more likely to be purchased individually or in combination with some   other products.

- need to help OList to identify the product categories which they can get rid of without significantly impacting business

OBJECTIVES:

TASK-1 PYTHON NOTEBOOK

- Missing values are correctly identified and treated accordingly, i.e., no missing value is present in the dataset.

- Duplicates are correctly identified and removed from the dataset, i.e., no duplicates are present in the dataset.

- The records are filtered to include only orders with order status as 'delivered'.

- The required tables are identified and joined in the correct way during data import.

TASK-2 TABLEAU

- The top 20 ordered products by quantity are identified and visualized.

- The percentage running totals by revenue and number of orders to be depicted correctly for each product.

- The product categories which are ordered more than 5 times to be identified.

- Combinations of product categories which are frequently ordered together are identified and visualised appropriately

TASK-1 PREPARING DATA USING PYTHON

• Read Data for each sheet provided in Excel sheet Retail_dataset

• Check Null values for each sheet provided in Excel sheet Retail_dataset

1. Orders sheet

```
In [2]:   # reading the orders sheet
          orders = pd.read_excel("Retail_dataset.xlsx", sheet_name = "orders")
```

```
In [3]:   #inspecting data
          print(orders.shape)

          (99441, 7)
```

```
In [8]:   #checking null values for each column
          (orders.isnull().sum()/len(orders)*100).sort_values(ascending = False).head(70)

Out[8]:   order_approved_at                0.014511
          order_delivered_timestamp        0.008292
          order_id                         0.000000
          customer_id                      0.000000
          order_status                     0.000000
          order_purchase_timestamp         0.000000
          order_estimated_delivery_date    0.000000
          dtype: float64
```

TASK-1 PREPARING DATA CONTD....

• Now we will export the cleaned dataset to start our analysis on Tableau

• A new dataset consisting of order id and product category name was created for Market Basket Analysis.

Now we will export the cleaned dataset to start our analysis on Tableau

```python
[39]:  # create a Pandas Excel writer using XlsxWriter as the engine
       Market = pd.ExcelWriter("Retail_dataset_new.xlsx", engine = 'xlsxwriter')

[40]:  orders.to_excel(Market, sheet_name = "order", index = False)
       order_items.to_excel(Market, sheet_name = "order_items", index = False)
       customers.to_excel(Market, sheet_name = "customers", index = False)
       payments.to_excel(Market, sheet_name = "payments", index = False)
       products.to_excel(Market, sheet_name = "products", index = False)

[41]:  Market.save()
       print("Files exported successfully.")

       Files exported successfully.

[42]:  # check if the export was successful
       import os
       os.getcwd()

Out[42]:  'C:\\Users\\ASUS'
```
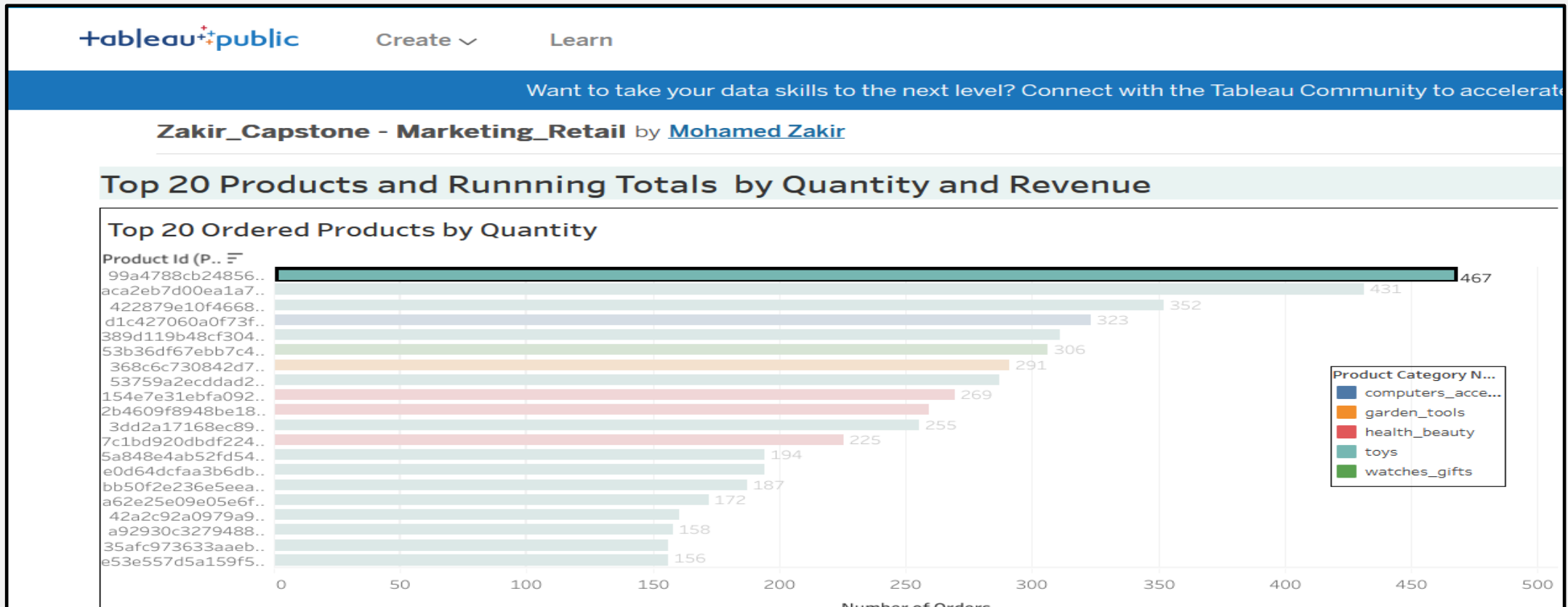
TASK-2 PREPARING TABLEAU DASHBOARD

- Form the connection between tables through clean retail dataset and marketing basket excel sheet details extracted using python notebook

The top 20 ordered products by quantity are identified and visualized.

• The highest ordered product is from the Toys category and has been ordered 467 times.

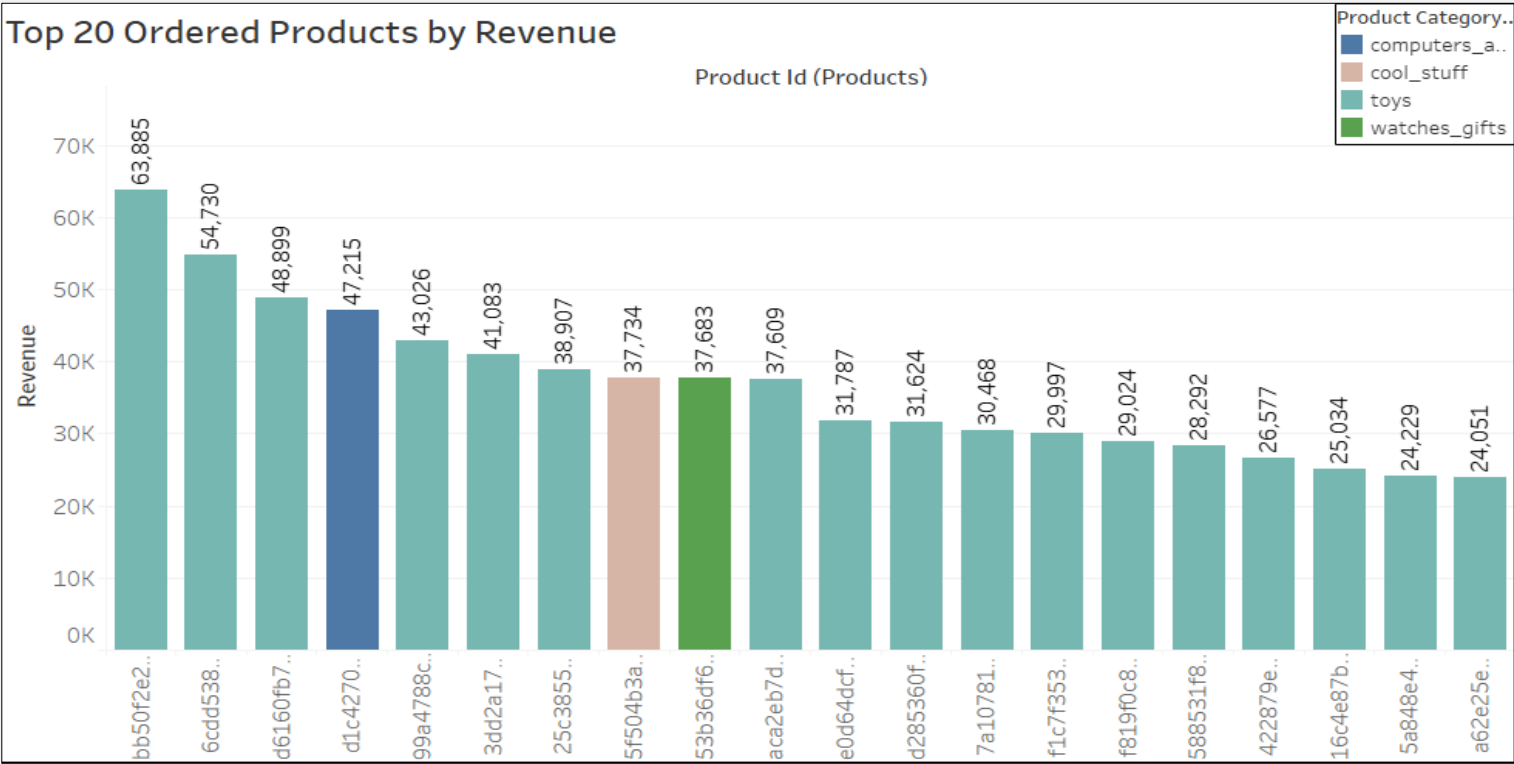• From Dashboard we can infer that Toys are ordered more under top 20 as green highlights more

Top 20 Ordered Products by Revenue

Product Category..
- computers_a..
- cool_stuff
- toys
- watches_gifts

• The highest revenue generation is 63,885 which belongs to the Toys Category.

The Percentage of Total Running Revenue and Quantity Ordered has been broken down by Product Id.

Product Id	Revenue	% of Total Running Revenue	Total Quantity Ordered	% of Total Running Quantity Ordered
bb50f2e236e5..	63,885	0.47%	186	0.19%
6cdd5384349..	54,730	0.87%	148	0.35%
d6160fb7873f..	48,899	1.23%	33	0.38%
d1c427060a0f..	47,215	1.58%	313	0.70%
99a4788cb248..	43,026	1.90%	456	1.18%
3dd2a17168ec..	41,083	2.20%	253	1.44%
25c38557cf79..	38,907	2.48%	38	1.48%
5f504b3a1c75..	37,734	2.76%	63	1.54%
53b36df67ebb..	37,683	3.04%	304	1.86%
aca2eb7d00ea..	37,609	3.32%	425	2.30%
e0d64dcfaa3b..	31,787	3.55%	193	2.50%
d285360f29ac..	31,624	3.78%	118	2.62%
7a107816372..	30,468	4.01%	140	2.77%
f1c7f353075c..	29,997	4.23%	149	2.92%
f819f0c84a64f..	29,024	4.44%	44	2.97%
588531f8ec37..	28,292	4.65%	19	2.99%
422879e10f46..	26,577	4.85%	352	3.35%
16c4e87b98a9..	25,034	5.03%	13	3.37%
5a848e4ab52f..	24,229	5.21%	187	3.56%
a62e25e09e05..	24,051	5.38%	170	3.74%
2b4609f8948b..	22,717	5.55%	254	4.00%
fd0065af7f09..	22,000	5.71%	10	4.01%
a5215a7a9f46..	21,740	5.87%	16	4.03%
bc4cd4da98dd..	21,500	6.03%	17	4.04%

Revenue pareto analysis is performed to understand ideal category depth for each category.

- Toys, health_beauty and watches_gift combine generate 80.56% of the revenue.

- Toys alone generates 76.23% of the revenue.

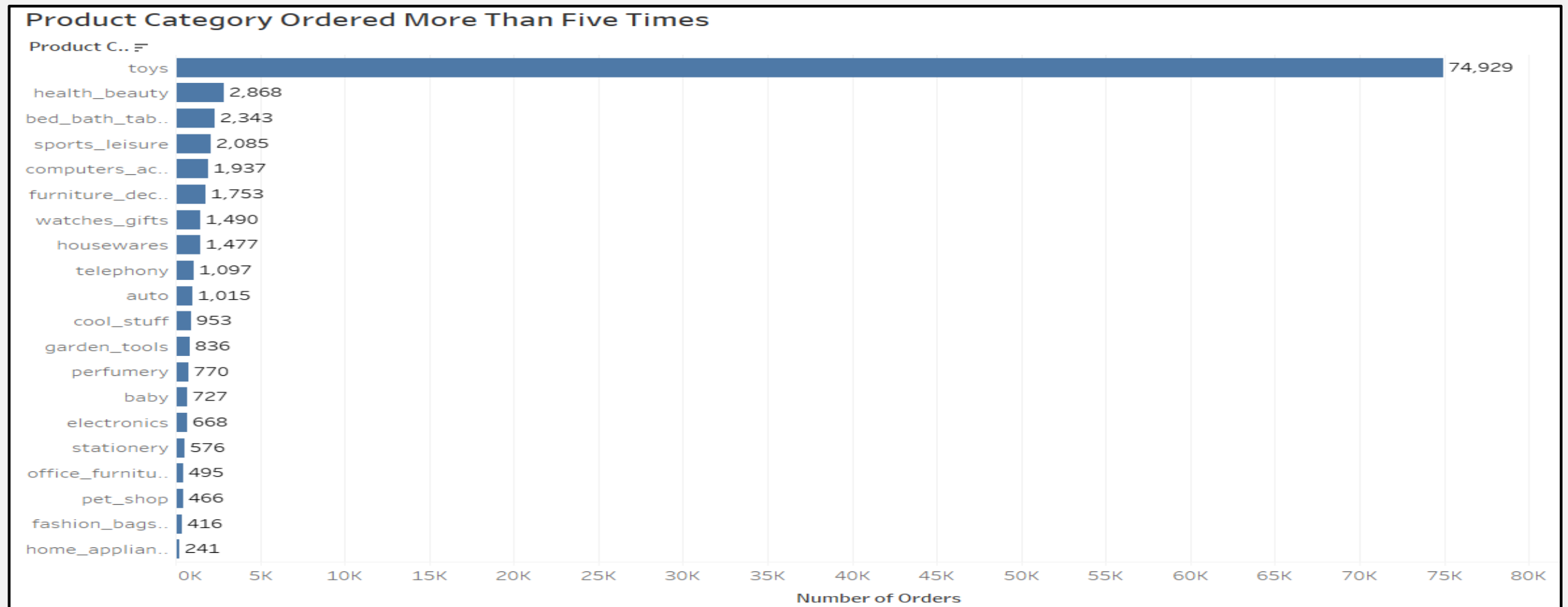- The rest of the 70+ product categories generates 19.44% of the revenue.
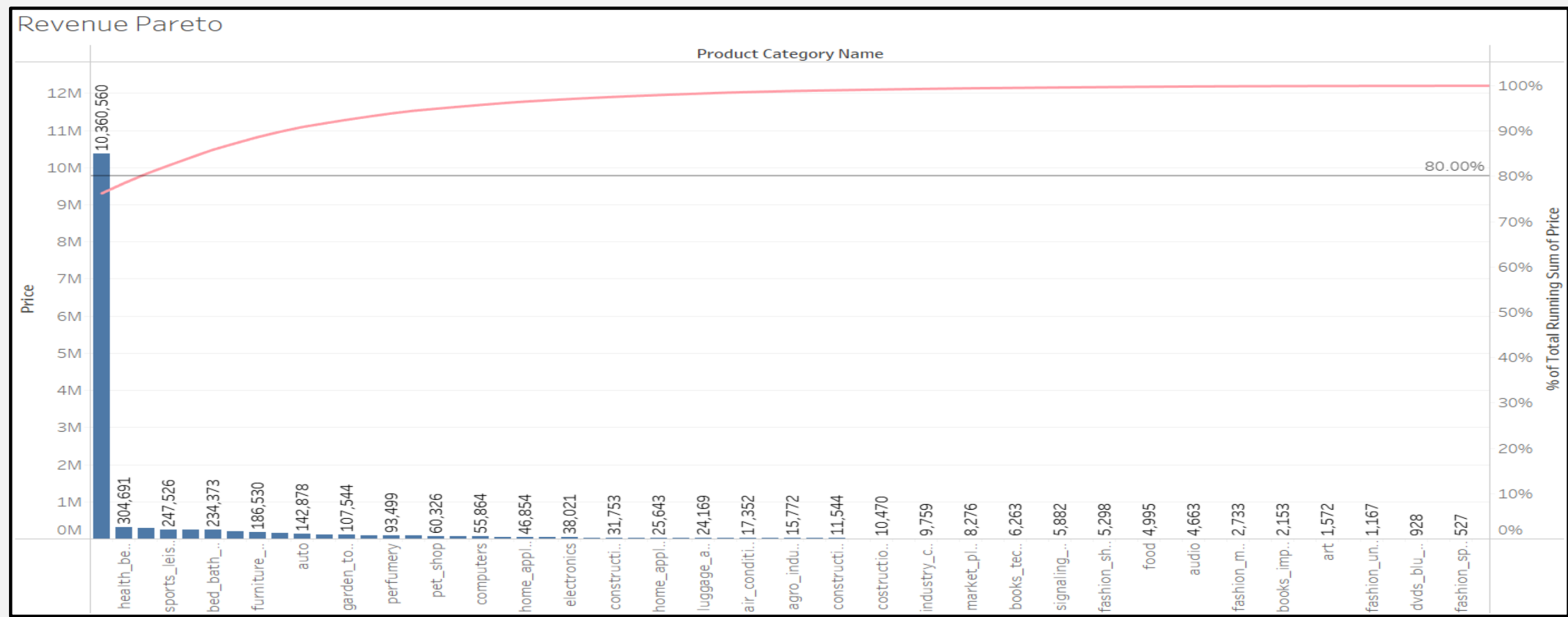
- Toys, health_beauty and bed_bath_table make up 80.38% of the total orders.

- Toys alone has 75.94% of the total orders.

- The rest of the 70+ product categories generate 19.62% of the total orders.
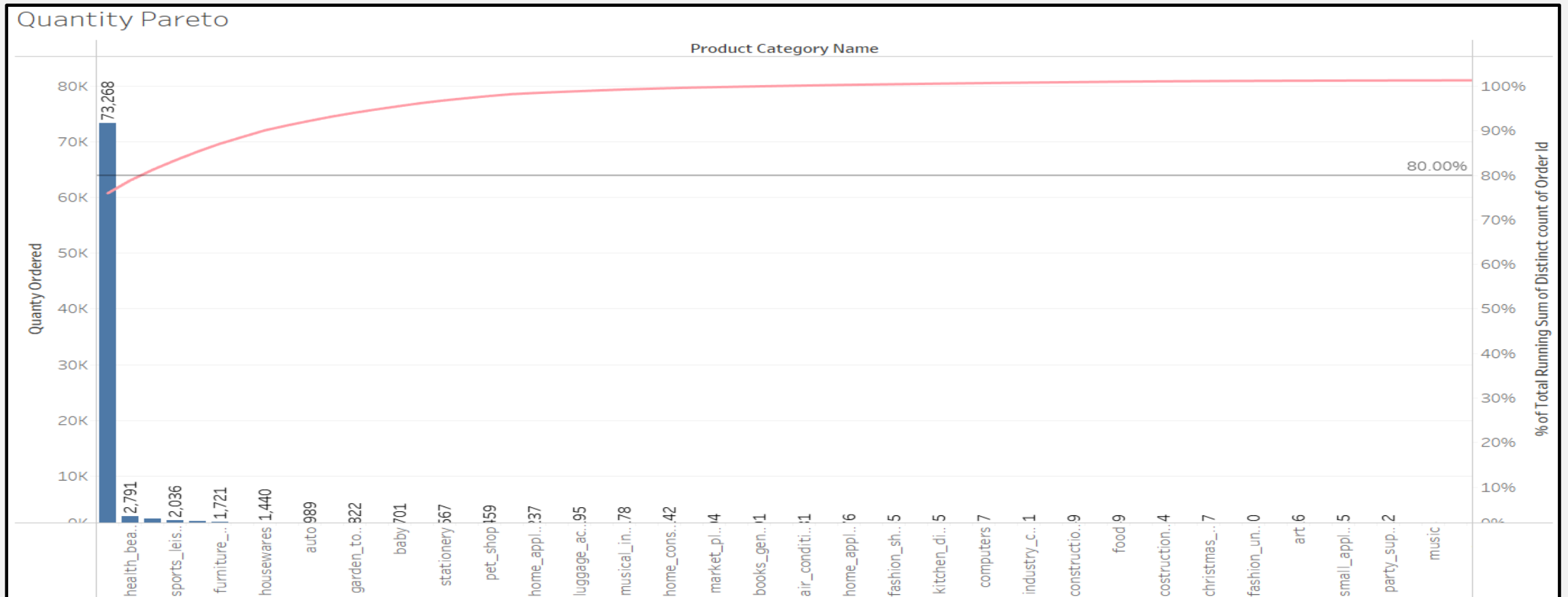
The product categories which are ordered more than 5 times to be identified.

• Toys category is the most ordered category with a total of 74,929 orders.

• Health_beauty, bed_bath_table and sports_leisure are the next most ordered category.

Product Category Ordered More Than Five Times

Product C.. =

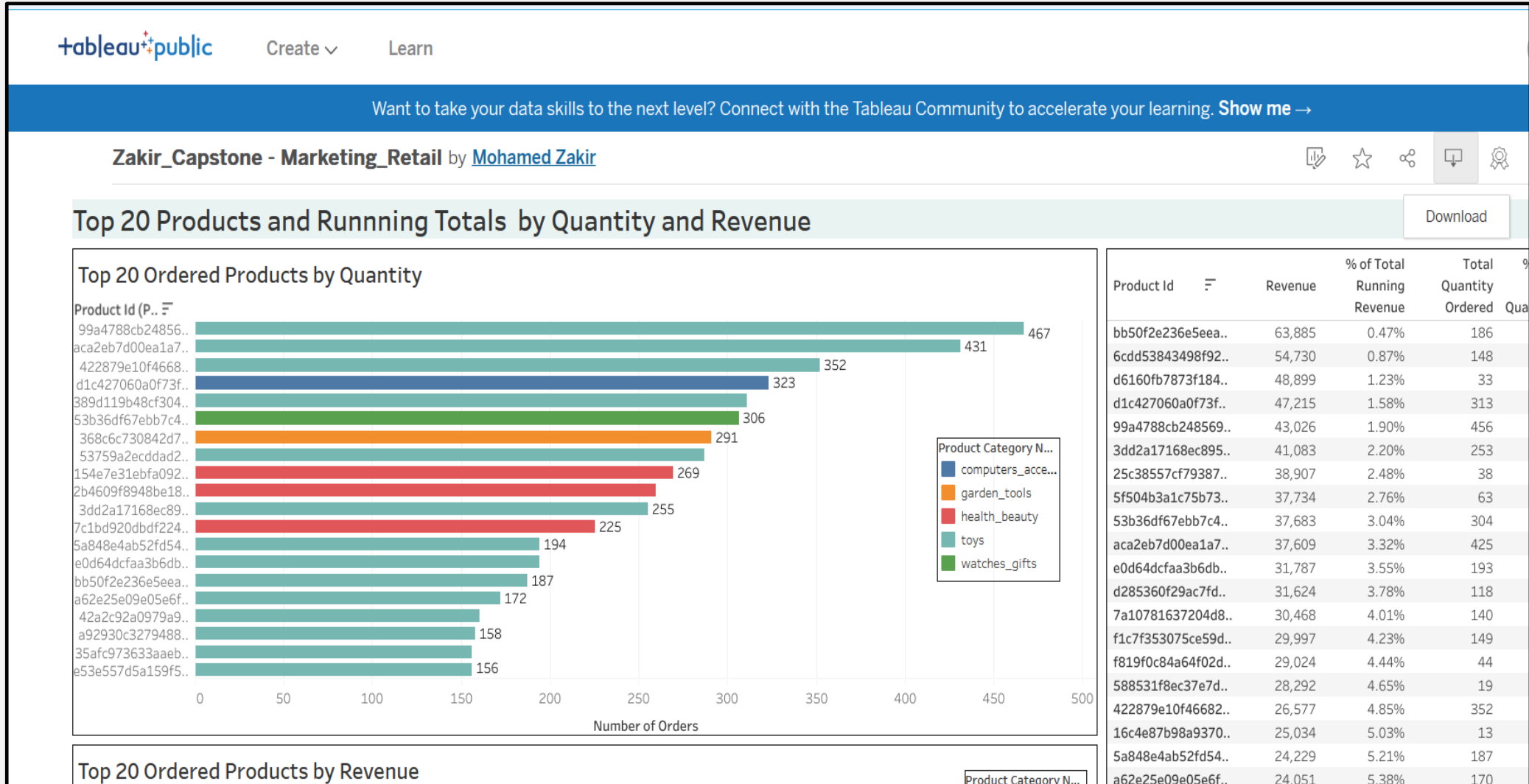| Category	Number of Orders
toys	74,929
health_beauty	2,868
bed_bath_tab..	2,343
sports_leisure	2,085
computers_ac..	1,937
furniture_dec..	1,753
watches_gifts	1,490
housewares	1,477
telephony	1,097
auto	1,015
cool_stuff	953
garden_tools	836
perfumery	770
baby	727
electronics	668
stationery	576
office_furnitu..	495
pet_shop	466
fashion_bags..	416
home_applian..	241

Number of Orders

Market Basket Analysis

❑Market basket analysis is performed to identify the frequently ordered category association.

❑Toys are often purchased with various other categories as shown in this Market Basket Analysis.

Market Basket Analysis

Product Category Name (Order!Prod2)

Product Category Name (Order!Prod1)	toys	watches_gifts	sports_leisure	furniture_decor	garden_tools	housewares	home_construction	cool_stuff	health_beauty
toys	1,193	73		3	2				
bed_bath_table	294			6		2			
furniture_decor	156			8	3				
computers_accessor..	104				2	2			
health_beauty	84		2						2
watches_gifts		75							
sports_leisure	62		9						
housewares	63					6			
garden_tools	45		2		7				
fashion_bags_acces..	42								
auto	40								
perfumery	32								
telephony	32								
stationery	31								
cool_stuff	25							2	
pet_shop	23								
baby	21								
office_furniture	18								

- Save it online so data connection remains and work with easily

- Download Dashboard file for submission

Inferences:

- The category 'toys' is the most ordered category as it is ordered 74929 times (76% of the total number of orders)

- Apart from 'toys', the categories 'health_beauty','bed_bath_table','sports_leisure', 'computer_accessories' and 'furniture_decor' are the most frequently ordered categories.The above categories with 'toys' or/and with each other are most frequent in customers' basket.

- It is observed that despite of the high price, some products are frequently purchased by the customers.

- Only the cases having order status as 'delivered' are considered.

- We assumed that the data provided was achieving the desired revenue.

Recommendations:

- Target customers who have children to boost up sales as they are most likely to purchase 'toys' which is the most ordered category.

- Offer Promo-codes or discounts on frequently ordered category associations and the most ordered products to attract more customers.

- Consider the ideal category depth to minimize the inventory cost by getting rid of the products which are seldom ordered and/or do not have a significant contribution to the total revenue under each product category.