A CASE STUDY TO IDENTIFY OPPORTUNITIES TO OPTIMISE INVENTORY COST

By Sagnik Chakravarty and Hashim Ali

AGENDA OR OBJECTIVE

- In the recent past, e-commerce companies have emerged and flourished in the industry. They offer the convenience to order from a wide variety of options from the comfort of one's home. But how do they offer these "wide variety of options or products"? To be able to meet the demands of the customers, any e-commerce company would obviously need to store tons and tons of products in warehouses. Now, some of these warehoused products might be fast-moving products which sell very quickly and some others might be slow-moving. Each of the products being stored incurs a cost to the company in terms of space and maintenance. Since storing these products obviously add to the costs that the company incurs, it is absolutely necessary for the organisations to plan their inventory well.
- OList is also facing a similar problem as it has incurred losses recently due to unoptimal management of it's inventory. It's obviously very important to manage their inventory very well so as to reduce any unnecessary costs that they might be bearing.
- So in this case study, we aim to identify the top products that contribute to the revenue and also use market basket analysis to study 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.

CONTENTS

- ***Exploratory Data Analysis***

- ***Insights from Market Basket Analysis***
  - *Toys are dominating in terms of contribution to total revenue*
  - *Toys are clearly the biggest contributor to total orders*
  - *Top 10 product categories in terms of average weight: Office furniture, Home Appliances and Computers are the heaviest products on average*
  - Frequently bought Itemsets.

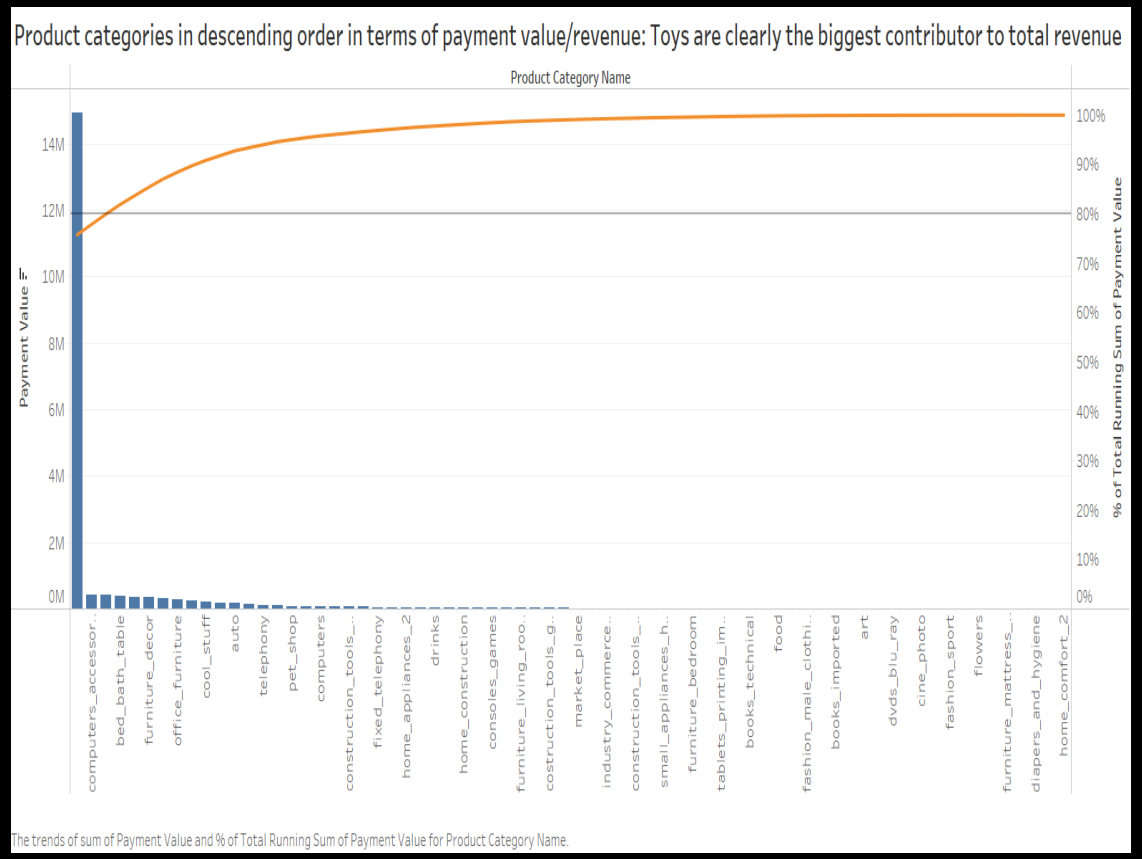
- Conclusion

EXPLORATORY DATA ANALYSIS

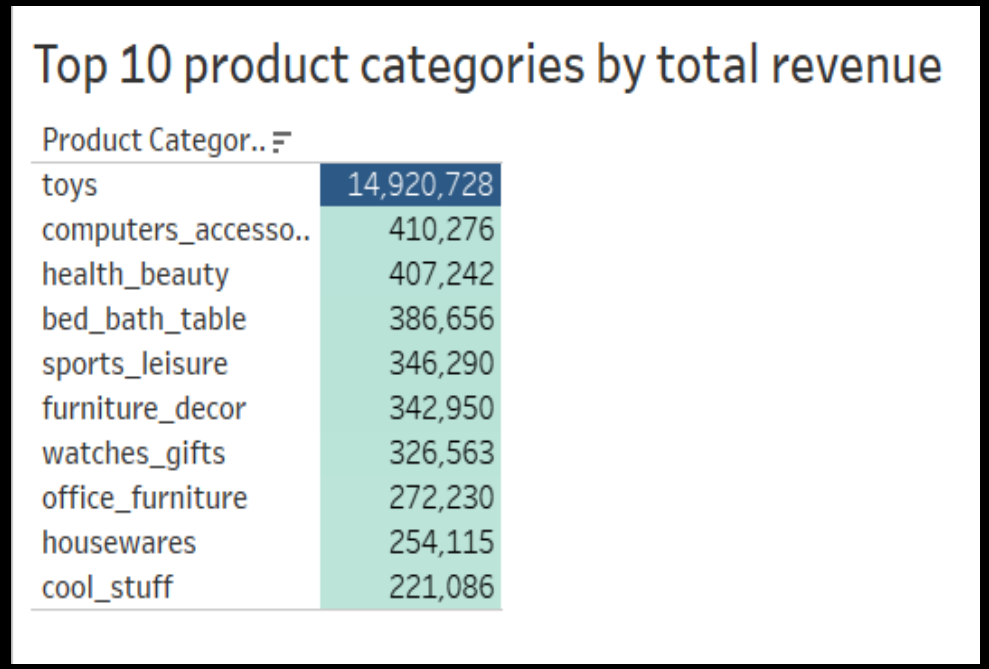
- Importing Libraries - Pandas, numpy, matpltlib amd seaborn
- Importing Dataset from different sheets
- Checking for duplicates - Duplicates only in customer table and we have removed duplicates due to small number.
- Checking for Null values - Removing null values in Orders and products table as it is only small portion.
- Outlier analysis - Using boxplot we checked for outliers. wherever there was a big difference between mean and median, we capped the outliers .
- Merged different datasets into a single dataset.

TOYS ARE DOMINATING IN TERMS OF CONTRIBUTION TO TOTAL REVENUE

- We plotted a pareto chart to compare the payment value or revenue generated by each product categories and toys are clearly dominating by a very wide margin.
- We can see that toys, computer accessories and health beauty product categories alone are contributing almost 80% of the total revenue.

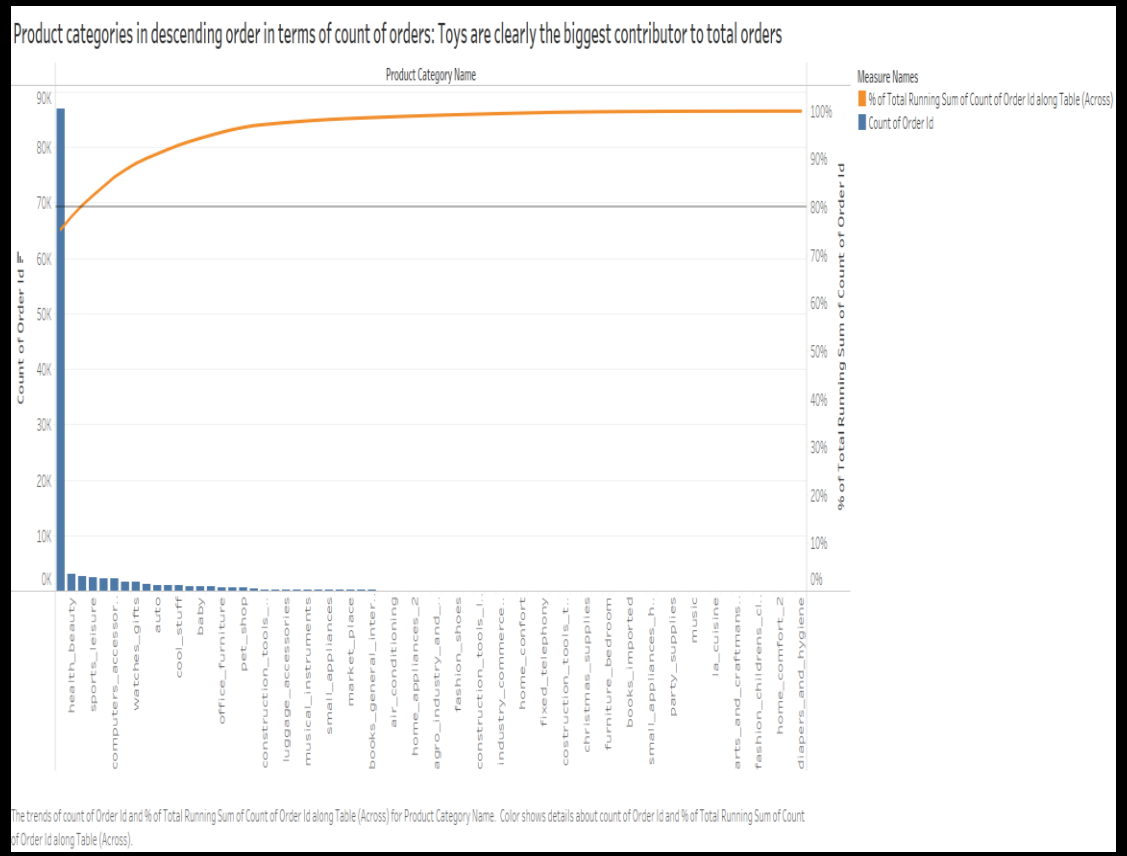


- Here showing the top 10 product categories by total revenue.

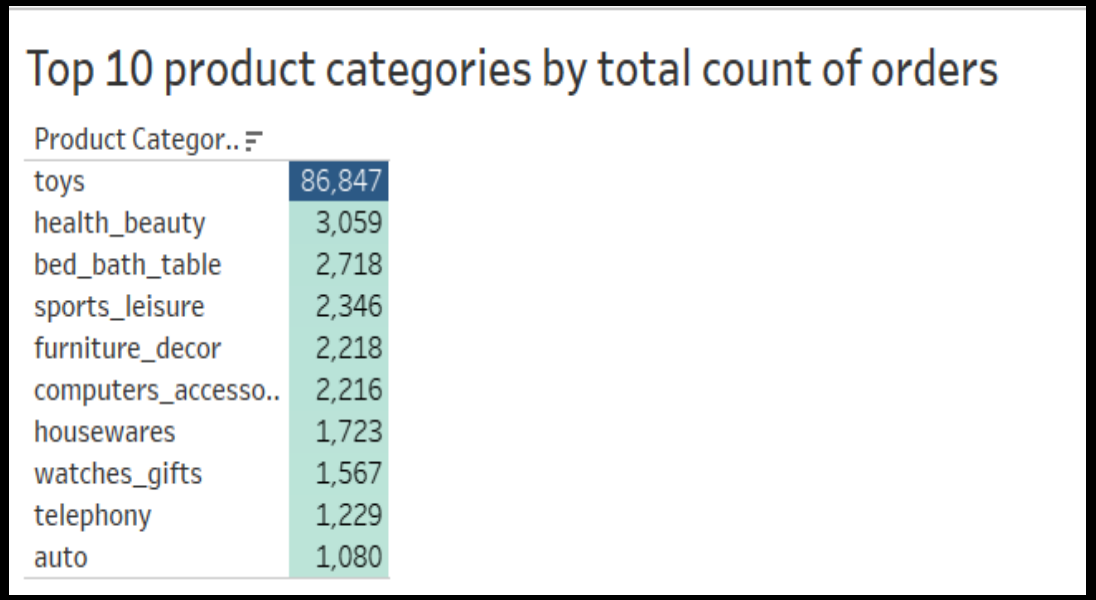


TOYS ARE CLEARLY THE BIGGEST CONTRIBUTOR TO TOTAL ORDERS

- We plotted another pareto chart to compare the product categories by their count of orders and toys are dominating again.
- We can see that toys, health beauty products and bed bath table alone are contributing 80% to the total count of orders.

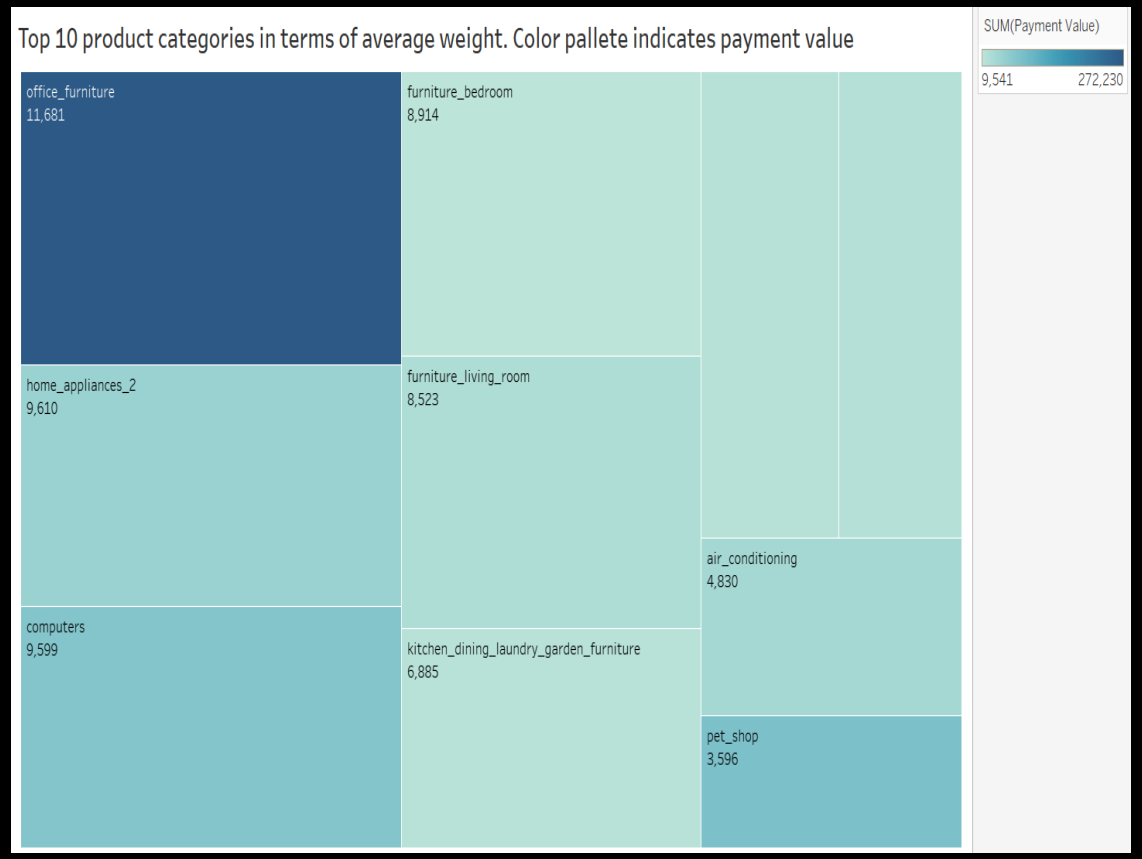


- Here showing the top 10 product categories by their count of orders.

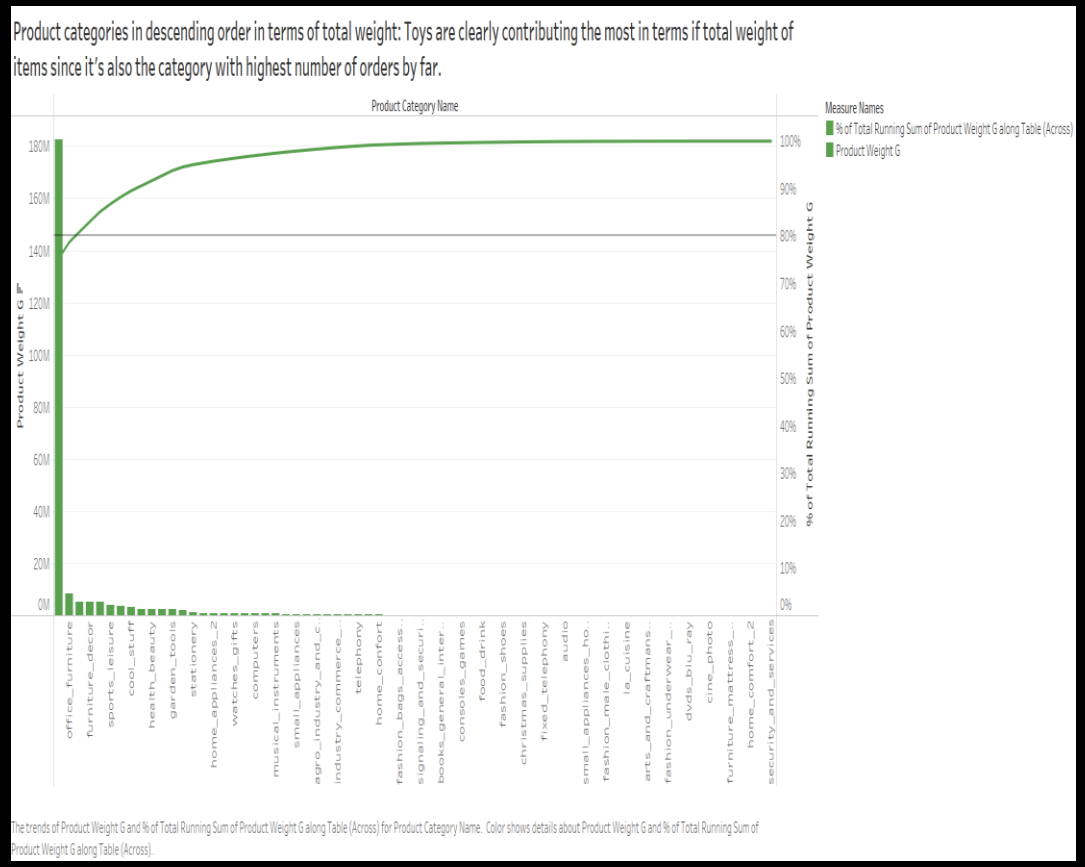


TOP 10 PRODUCT CATEGORIES IN TERMS OF AVERAGE WEIGHT: OFFICE FURNITURE, HOME APPLIANCES AND COMPUTERS ARE THE HEAVIEST PRODUCTS ON AVERAGE

- Heavy items take more than usual effort when it comes to storage so we created a heatmap to check which categories are the heaviest on average.
- It turns out that office furniture, home appliances and computers are the heaviest on average.



- While we checked the product categories by average weight, usually reality can't be measured by averages alone. So we created a pareto chart to see which product categories contributed to the total weight of products the most.
- It turns out that toys are again at the top contributing most to the total weight by far even though it's average weight is very low. This can be attributed to it's very large amount of orders. Hence, toys are kept in stock in large quantities to meet the demand.
- Toys are followed by office furniture and housewares. Together these 3 categories contribute 80% to the total weight of items.



FREQUENT ITEMSETS

- We used apriori algorithm to find frequent item sets in groups of two as well as three. These items are the product combinations that are most likely to be bought together by the customers.
- We set the minimum support at 0.05 or 5% and lift threshold to 0.8.

The top 5 frequent itemsets in groups of two were as follows:

a. 97% of customers buying toys are likely to but bed bath table

b. 97% of customers buying toys are likely to buy watches gifts.

c. 97% of customers buying toys are likely to buy computer accessories.

d. 97% of customers buying toys are likely to buy health beauty.

e. 97% of customers buying toys are likely to buy furniture decorations.

```
In [61]: rules = association_rules(frequent_itemsets, metric='lift',min_threshold = 0.8).sort_values('lift', ascending=False).reset_inde
         rules
```

```
C:\Users\user\anaconda3\lib\site-packages\ipykernel\ipkernel.py:287: DeprecationWarning: `should_run_async` will not call `tr
ansform_cell` automatically in the future. Please pass the result to `transformed_cell` argument and any exception that happe
n during thetransform in `preprocessing_exc_tuple` in IPython 7.17 and above.
  and should_run_async(code)
```

Out[61]:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(toys)	(bed_bath_table)	0.971139	0.230889	0.226989	0.233735	1.012325	0.002764	1.003714
1	(bed_bath_table)	(toys)	0.230889	0.971139	0.226989	0.983108	1.012325	0.002764	1.708580
2	(toys)	(watches_gifts)	0.971139	0.058502	0.056942	0.058635	1.002260	0.000128	1.000140
3	(watches_gifts)	(toys)	0.058502	0.971139	0.056942	0.973333	1.002260	0.000128	1.082293
4	(toys)	(computers_accessories)	0.971139	0.084243	0.080343	0.082731	0.982047	-0.001469	0.998351
5	(computers_accessories)	(toys)	0.084243	0.971139	0.080343	0.953704	0.982047	-0.001469	0.623401
6	(toys)	(health_beauty)	0.971139	0.065523	0.062402	0.064257	0.980685	-0.001229	0.998647
7	(health_beauty)	(toys)	0.065523	0.971139	0.062402	0.952381	0.980685	-0.001229	0.606084
8	(toys)	(furniture_decor)	0.971139	0.127145	0.119345	0.122892	0.966546	-0.004131	0.995151

The top 3 frequent itemsets in groups of three were as follows:

a. telephony, toys and cine photo

b. home construction, toys and computer accessories

c. furniture decorations, toys and electronics

Out[71]:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(telephony, toys)	(cine_photo)	0.1	0.1	0.1	1.00	10.0	0.09	inf
1	(cine_photo)	(telephony)	0.1	0.1	0.1	1.00	10.0	0.09	inf
2	(toys, cine_photo)	(telephony)	0.1	0.1	0.1	1.00	10.0	0.09	inf
3	(telephony)	(toys, cine_photo)	0.1	0.1	0.1	1.00	10.0	0.09	inf
4	(cine_photo)	(telephony, toys)	0.1	0.1	0.1	1.00	10.0	0.09	inf
5	(telephony)	(cine_photo)	0.1	0.1	0.1	1.00	10.0	0.09	inf
6	(home_construction)	(toys, computers_accessories)	0.1	0.2	0.1	1.00	5.0	0.08	inf
7	(computers_accessories)	(home_construction)	0.2	0.1	0.1	0.50	5.0	0.08	1.8
8	(toys, computers_accessories)	(home_construction)	0.2	0.1	0.1	0.50	5.0	0.08	1.8
9	(toys, home_construction)	(computers_accessories)	0.1	0.2	0.1	1.00	5.0	0.08	inf
10	(home_construction)	(computers_accessories)	0.1	0.2	0.1	1.00	5.0	0.08	inf
11	(computers_accessories)	(toys, home_construction)	0.2	0.1	0.1	0.50	5.0	0.08	1.8
12	(furniture_decor)	(toys, electronics)	0.4	0.1	0.1	0.25	2.5	0.06	1.2
13	(toys, electronics)	(furniture_decor)	0.1	0.4	0.1	1.00	2.5	0.06	inf
14	(garden_tools)	(toys, computers_accessories)	0.2	0.2	0.1	0.50	2.5	0.06	1.6
15	(electronics)	(toys, furniture_decor)	0.1	0.4	0.1	1.00	2.5	0.06	inf
16	(electronics)	(furniture_decor)	0.1	0.4	0.1	1.00	2.5	0.06	inf
17	(garden_tools)	(computers_accessories)	0.2	0.2	0.1	0.50	2.5	0.06	1.6
18	(toys, furniture_decor)	(electronics)	0.4	0.1	0.1	0.25	2.5	0.06	1.2
19	(furniture_decor)	(electronics)	0.4	0.1	0.1	0.25	2.5	0.06	1.2

CONCLUSION

- *Toys are dominating in both terms of contribution to total revenue or payment value as well as count of orders. So it's without a doubt the most valuable product category.*
- *We can see that toys, computer accessories and health beauty product categories alone are contributing almost 80% of the total revenue.*
- *We can also see that toys, health beauty products and bed bath table alone are contributing 80% to the total count of orders.*
- *Office furniture, home appliances and computers are the heaviest on average but toys are again at the top contributing most to the total weight by far even though it's average weight is very low. This can be attributed to it's very large amount of orders. Hence, toys are kept in stock in large quantities to meet the demand.*

The top 5 frequent itemsets in groups of two were as follows:

a. 97% of customers buying toys are likely to but bed bath table

b. 97% of customers buying toys are likely to buy watches gifts.

c. 97% of customers buying toys are likely to buy computer accessories.

d. 97% of customers buying toys are likely to buy health beauty.

e. 97% of customers buying toys are likely to buy furniture decorations.

The top 3 frequent itemsets in groups of three were as follows:

a. telephony, toys and cine photo

b. home construction, toys and computer accessories

c. furniture decorations, toys and electronics

- Inorder to optimize the inventory we can remove product categories that are contributing heavily to the total weight of items or are heavy on average but are not part of any frequent itemset combination with toys and not part of the top 10 product categories towards total revenue contribution for e.g. home appliances, computers, furniture bedroom, etc.