



OList Marketing and Retail Analytics: Capstone Project

AKSHATA GURAYYA ANNIGERI & PAPANA
RAVI TEJA

Agenda

1. Objectives
2. Background
3. Visualizations
4. Insights
5. Recommendations
6. Appendix –
 - Data Sources
 - Data Methodology
 - Data Assumptions



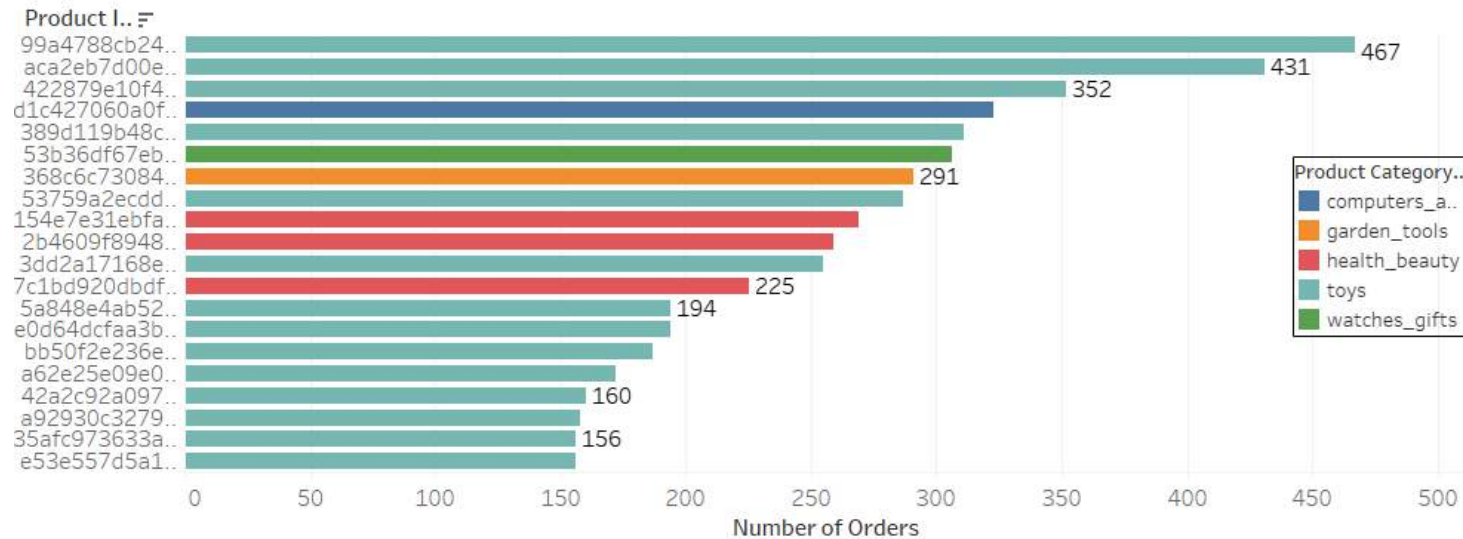
Objective

- Identify revenue-driving products and top product category through Pareto Analysis.
- Enhance insights with market basket analysis to understand customer purchase behavior.
- Determine individual and complementary product preferences to optimize sales strategy.

Background

OList, an e-commerce firm, aims to mitigate recent losses by optimizing inventory management to minimize unnecessary costs. To meet customer demands, efficient warehousing is crucial. Strategic inventory planning is essential to streamline operations and control expenses associated with storage.

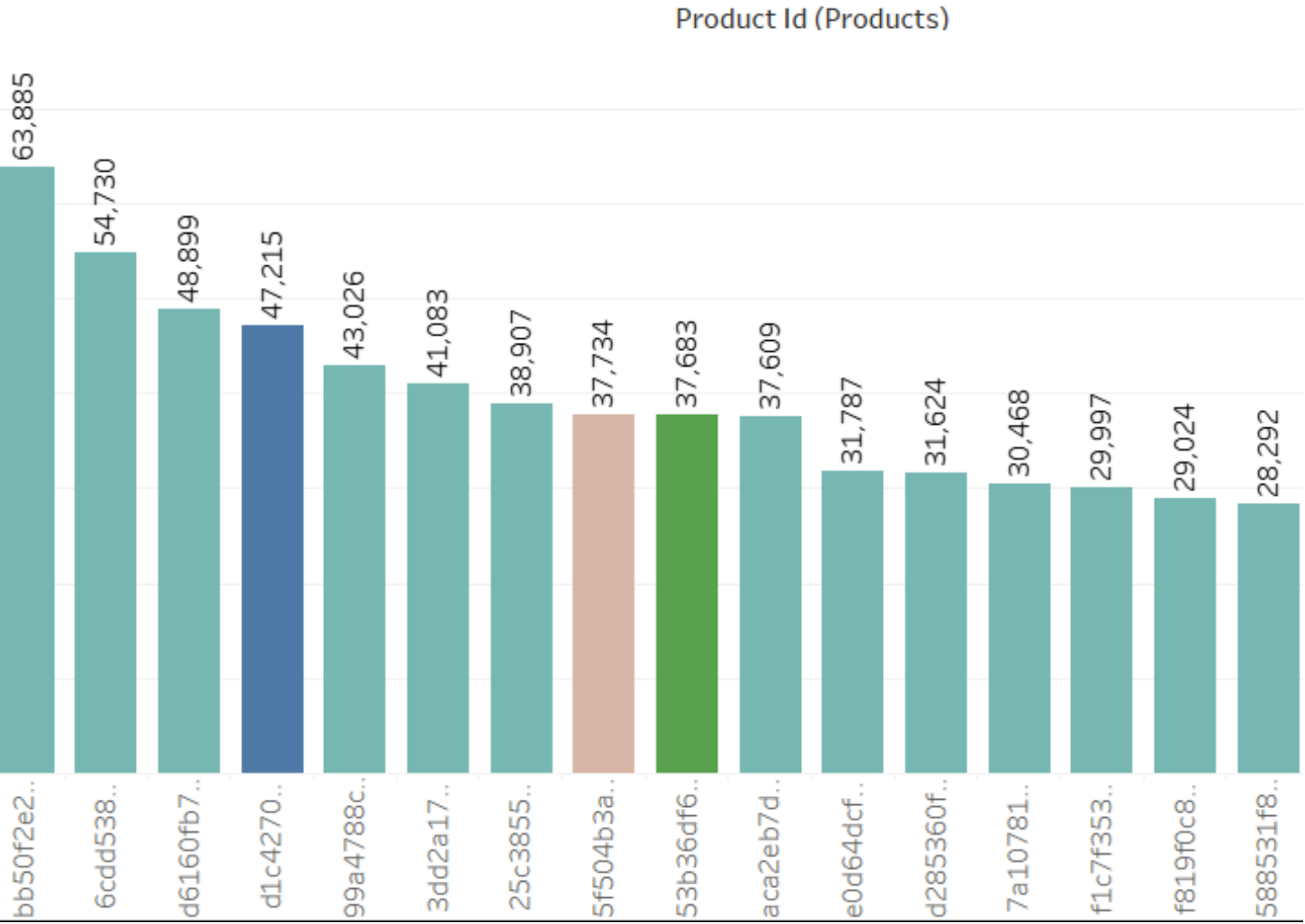
Top 20 Ordered Products by Quantity



Top 20 Ordered Products by Quantity

- The highest ordered product is from the Toys category and has been ordered 467 times.
- Most of the products in the Top 20 that are frequently ordered belong to the Toys category.

Ordered Products by Revenue



Top 20 Ordered Products by Revenue

- The highest revenue generation is 63, 885 which belongs to the Toys Category.
- Most of the products in the Top 20 list generating high revenue belong to the Toys category.

Percent Running Totals

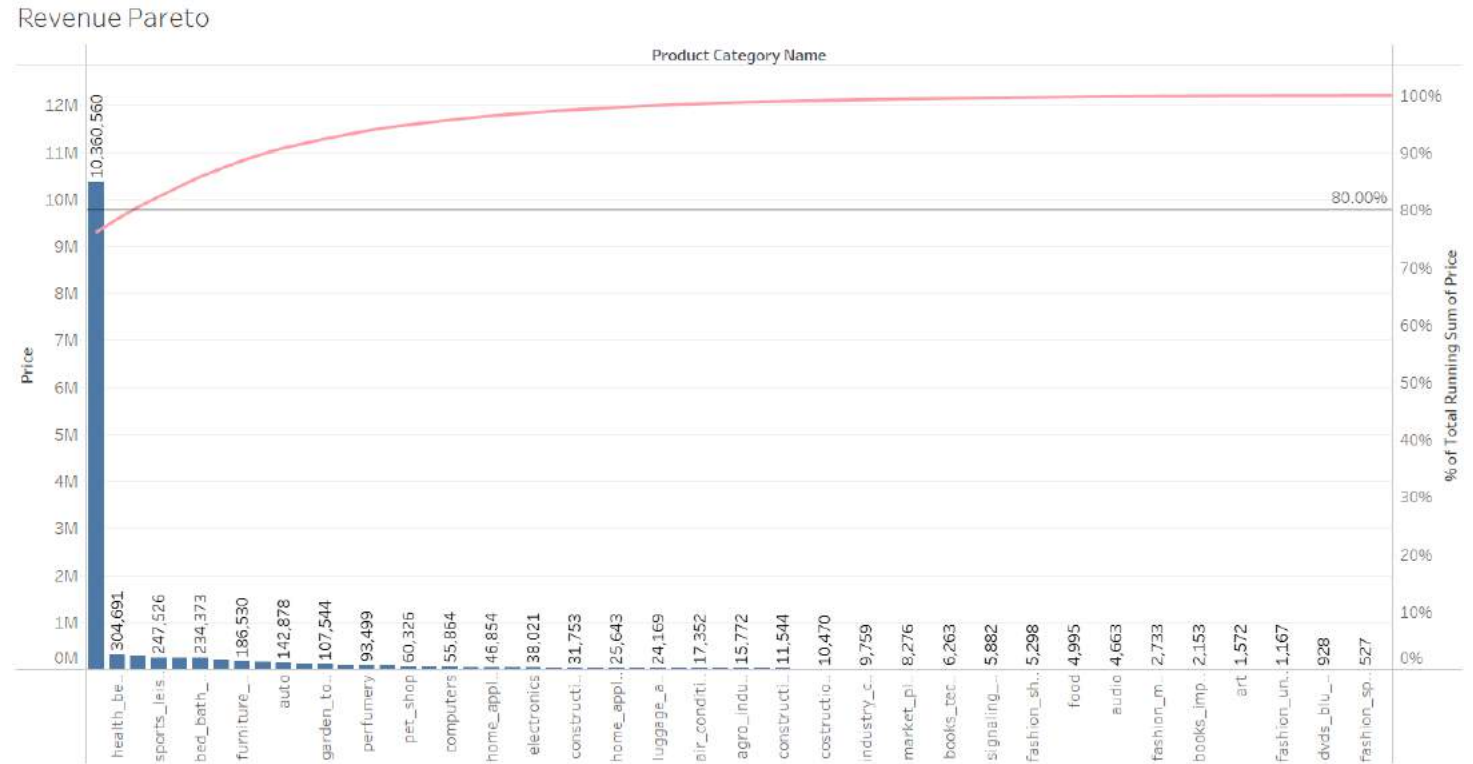
- The Percentage of Total Running Revenue and Quantity Ordered has been broken down by Product Id.
- The contribution of each product towards the total revenue can be identified.

Percentage Running Totals by Revenue and Orders

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

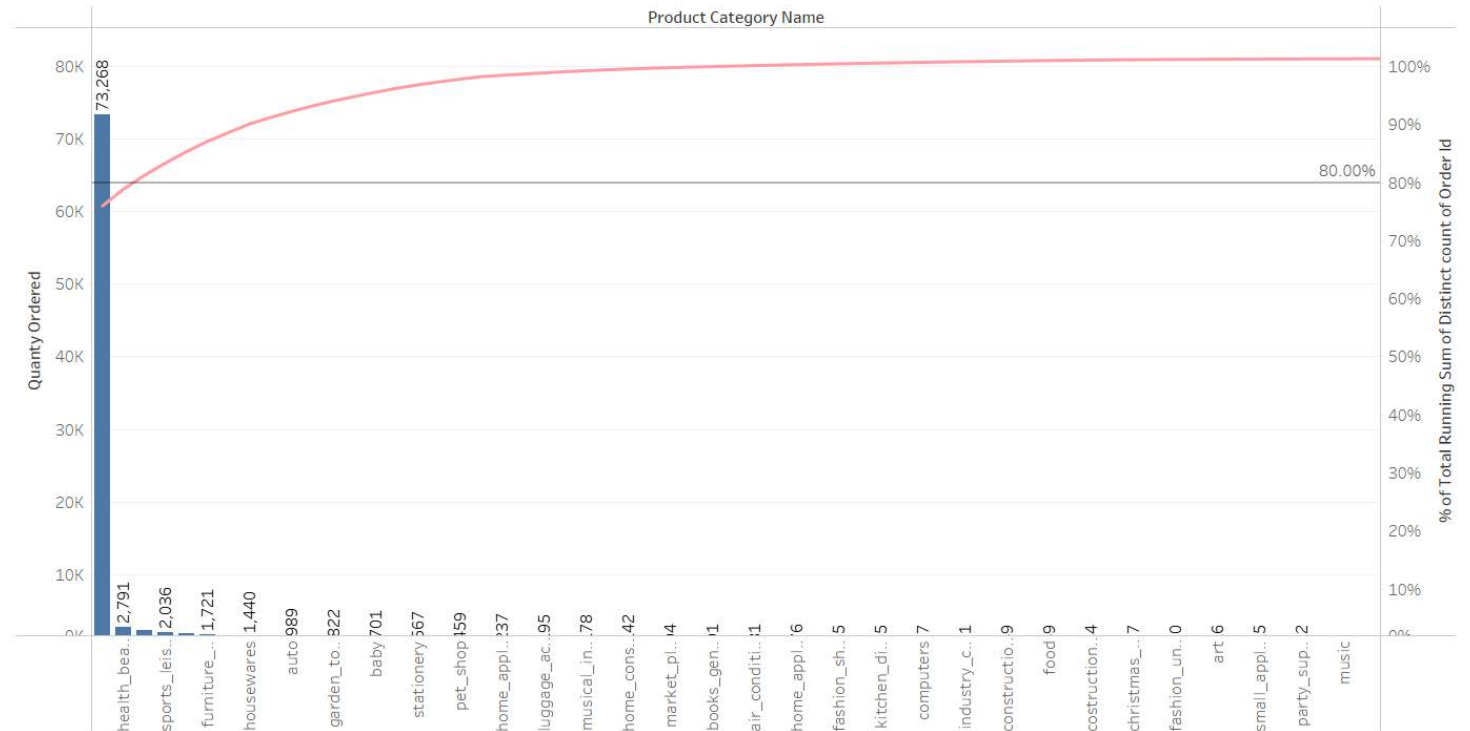
- Toys, health_beauty and watches_gift combine generate 80.56% of the revenue.
- Toys alone generates 76.23% of the revenue.
- The rest of the 70+ product categories generates 19.44% of the revenue.



Quantity Pareto

- 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.

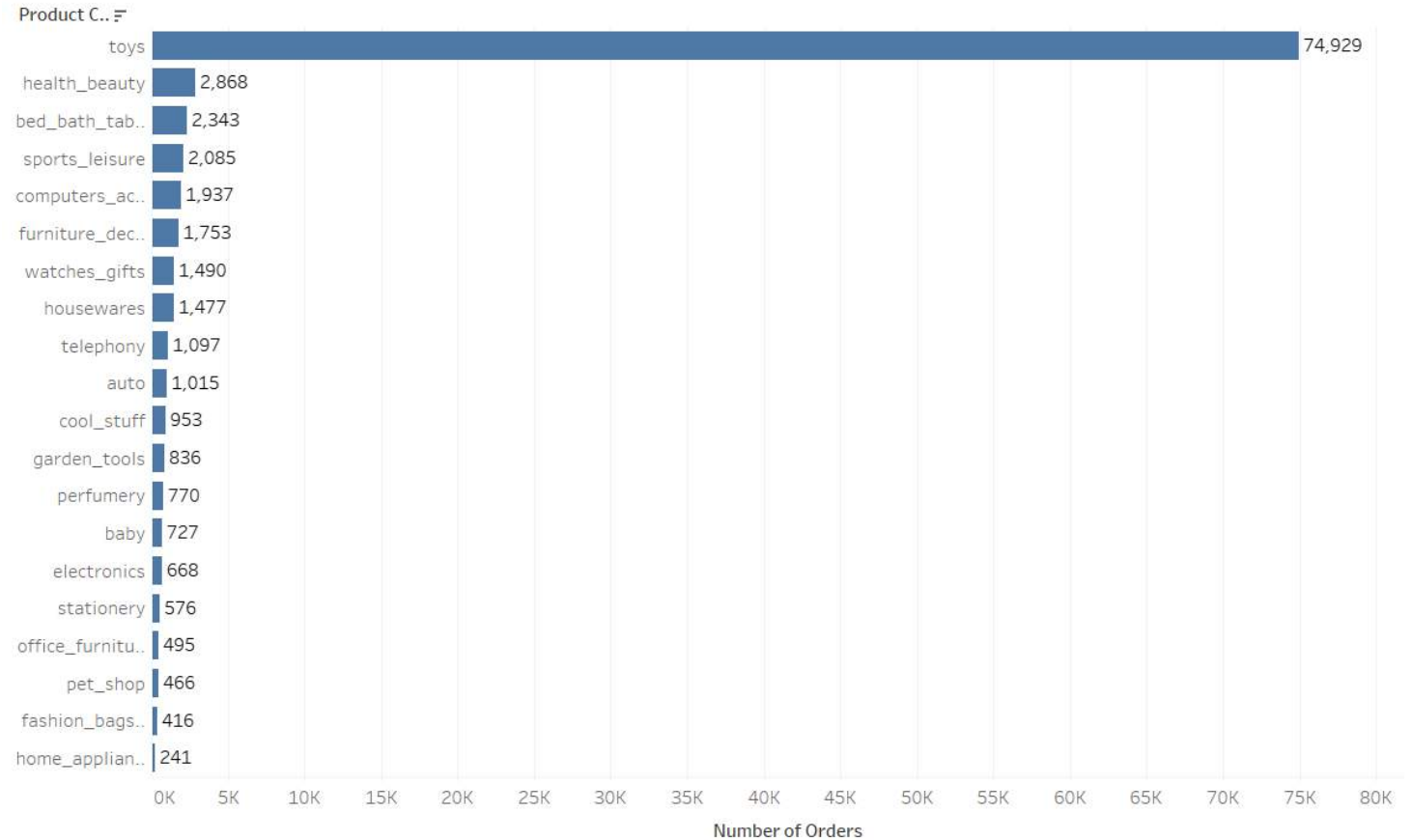
Quantity Pareto



Product Category Ordered >5 Times

- 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



Market Basket Analysis

- Market Basket Analysis is performed to identify the frequently ordered category association.
- Toys are the most ordered category along with the categories of bed_bath_table, furniture_decor, garden_tools, housewares, computers_accessories and health_beauty.

Market Basket Analysis

Product Category Name (Order!Prod1)	Product Category Name (Order!Prod2)								
	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_accessories	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_accessories	42								
auto	40								
perfumery	32								
telephony	32								
stationery	31								
cool_stuff	25							2	
pet_shop	23								
baby	21								
office_furniture	18								

Insights

- ❖ **Revenue Focus:** Toys, comprising 20% of products, generate a significant 80% of revenue.
- ❖ **Price Resilience:** Despite higher prices, specific products are consistently popular among customers.
- ❖ **Top Ordered Categories:** Bed_bath_table, furniture_decor, computers_accessories, and health_beauty are frequently ordered.
- ❖ **Basket Trends:** Customers commonly purchase products from Toys and other noted categories, revealing popular combinations.
- ❖ **High-Value Products:** Price doesn't deter frequent purchases, highlighting a willingness to invest in certain items.

Recommendations

- 1. Optimize Stock Levels:** Prioritize categories generating over 80% of revenue, ensuring consistent availability.
- 2. Targeted Marketing:** Direct marketing efforts towards customers inclined to purchase toys, the most frequently ordered category, to boost overall sales.
- 3. Promotional Strategy:** Introduce promo codes or discounts on frequently ordered category associations, fostering cross-selling and increasing average transaction value.
- 4. Streamlined Inventory:** Consider reducing low-performing subcategories to optimize resources and focus on high-impact products.

-

Appendix - Data Sources

- Here is a snapshot of our data dictionary:
 - Order details such as order id, order status, order purchased timestamp, etc.
 - Order Items detail such as order item id, seller id, price, shipping charges, etc.
 - Customer details such as customer is, customer city, customer state, etc.
 - Payment details such as payment type, payment value, etc.
 - Product details such as product id, product category name, product dimensions, etc.
- The following data sources were used:
 - OList retail dataset containing order-related information.
 - The data consisted for the year 2016 to 2018.

Appendix - Data Methodology

A thorough analysis of the OList Retail Dataset was conducted. The process included:

- The dataset was cleaned and transformed using the python libraries of Pandas and Numpy in the Jupyter Notebook.
- The missing values for the various columns were replaced with the best values.
- The redundant and duplicate records were discarded and only first occurrence is kept.
- Exploratory data analysis was done using the python libraries of Matplotlib and Seaborn in the Jupyter Notebook.
- A new dataset consisting of order id and product category name was created for Market Basket Analysis.
- Various visualizations and Market Basket Analysis was conducted in Tableau.

Appendix - Data Assumptions

- Only the cases having order status as 'delivered' are considered.
- We assumed that the data provided was achieving the desired revenue.
- We assumed that the company does not want to expand to new warehouses.
- The company's strategies are decided considering there is constant growth in sales.

Thank you