**MARKETING AND RETAIL ANALYTICS**

1. **Data cleaning and exploration:**
   1. Orders sheet:

* Checked for blank order\_id by using the formula ‘COUNTBLANK’ and also checked for duplicate order\_id by conditional formatting.
* Filtered out order\_status that were not ‘delivered’ and deleted the rows.
* Checked for null values in all columns, found some null cells in ‘order\_approved\_at’ and ‘order\_delivered\_timestamp’
* Replaced the null values for ‘order\_approved\_at’ column same as ‘order\_purchase\_timestamp’ as the approved and purchase dates are somewhere same for every order.
* Replaced the null values for ‘order\_delivered\_timestamp column same as ‘order\_estimated\_delivery\_date’ as the delivered dates are somewhat same for every order.
  1. Order\_items sheet:
* Checked for null values in every column, no null values found.
  1. Customers sheet:
* Checked for null values in every column, no null values found.
* Checked for duplicate customer\_id. 3345 found and deleted.
  1. Payments sheet:
* Checked for null values in every column, no null values found.
  1. Products sheet:
* Checked for null values, found them in columns: product\_category\_name, product\_weight\_g, product\_length\_cm, product\_height\_cm, product\_width\_cm.
* Replaced rows with ‘toys’ where product\_category\_name was null as the most ordered product is toys.
* Replaced with average using the formula ‘AVERAGE’ for the columns: product\_weight\_g, product\_length\_cm, product\_height\_cm, product\_width\_cm.

1. **Outlier Detection:**

Some outliers were detected for product weight, length, height and width but can be ignored as we have different categories including furniture, toys, etc.

1. **Market Basket data creation:**

* Created a new sheet for market basket analysis which consisted only order\_id and product\_name using VLOOKUP from order\_items and product sheets.
* Later removed the order\_id that were present only once using conditional formatting as we have to perform analysis on the items that were bought together in one single order.