

UNDERSTANDING THE DATASET

TEAM ID: PNT2022TMID44796

DATASETS:

Data Sets provide a structure to manage your uploaded data. A Data Set is a container that holds the data you upload to Analytics. Data Sets control how uploaded data gets joined with existing data.

CONTEXT:

The COVID-19 pandemic has forced people to limit their physical interactions, which has led to explosive growth in online grocery shopping and its currently the need of the hour. Let's see the two important aspects such as customer and product in online shopping.

CUSTOMER ANALYSIS:

Customer analysis is a way of analysing and using customer data while strategizing an organization's business plan. Customer analysis is a key element of customer acquisition strategies of marketing and advertising campaigns.

- Identifying who your customers are
- Making profitable segments of customers as per similar traits and customers behaviour
- A better understanding of purchasing decisions patterns of new customers
- Designing personalized content and marketing strategy on bestsuited marketing channels
- Extended decision making by suggesting significant modifications so product or service meets customer needs

- How many times will the customer buy the product?
- The frequency of the purchase and the time interval between two purchases?
- How much time did the customer take to decide what to buy?

PRODUCT ANALYSIS:

Product analysis is the process of gathering, defining, and analysing data about a product or service to make better decisions.

- How well your product meets customer needs
- Which country has the highest or lowest sales?
- What are the top ten products which makes more profit in each year?
- the overall cost of the goods
- How do sales affect the product's price
- Is there any increase in sales with the decrease in price at a day
- Based on location how the product delivery time varies

CONTENT:

After downloading the given dataset, we can see that the file contains the details of the online orders done by people across the globe within the time frame 1, January 2011 – 31, December 2014. Except the postal code column no other columns have missing values, which can be omitted.