

Electronics Data set

The electronic product dataset reveals fluctuating sales for an online store, with certain products experiencing suboptimal performance. The store owner is seeking strategies to stabilize sales throughout the day and night, and to identify key cities contributing significantly to overall sales. The provided dataset is currently unclean, prompting the need for data cleaning procedures.

Order ID:- Order id of the purchase

Product:- name of product

Quantity Ordered: number of quantity

Price Each :price

Order Date: order date

Purchase Address: address of the customer.

Strategy

Data cleaning, Data moulding, Extract particular information from data set, Geographical, time and product analysis.

Proceed by solving below questions.

QUESTION 1 Load the dataset into pandas and provide a summary of columns in the loaded dataset without any modification.

Question 2 : Prepare the data frame for further analysis. Clean it, Mold it in any manner you see fit

Step 1 check for null values, drop it and cross check

Step 2 check for duplicate values and drop it

Step 3 rectify column name

Step 4 Extract date from order date

Step 5 extract month from order date

Step 6 extract house_no ,city and pin code from address

Step 7 Check for month column for bad data entry and remove it

Step 8 correct data types

Step 9 make Sales Column (hint $\text{Quantity_ordered} * \text{Price_each}$)

Question 3 : Sales Value by Each Month

Question 4 Which city had the highest sales overall?

Question 5 Find out the distribution of sales over different times of the day (Morning, Afternoon, Evening, Night) and answer the following for each city.

Question 6 Highest Product Sold by Amount and Numbers

Conclusion : Recommend your solution to the clients