

E-Commerce Database

Background/General information

E-commerce, also known as electronic commerce or internet commerce, refers to buying and selling goods or services using the internet and transferring money and data to execute these transactions. A few of the largest E-commerce companies include Amazon, Shopify, and Alibaba. All these companies have to deal with large amounts of data related to customers, products, merchants, inventories, and more. A well-structured database powers e-commerce and manages all the interactions within the system. They can use their database to keep track of sales and inventory and derive meaningful information to improve profit margins.

Questions:

1. When were the largest orders made?
2. Which customers made an order with the most no of items?
3. What is the most common item bought within a specific price range(Will decide range later)
4. Display the order contents of a particular customer
5. Which company is doing the most business
6. How much is the average order cost?
7. Display all the products that have been bought by a particular customer or by all customer
8. Which company makes the most number of products
9. How many orders are being made by each customer state/city
10. How many orders were made each month of a year

Data Source

1. The Data in the below dataset is going to be used in the Customer and Order tables
<https://www.kaggle.com/datasets/benroshan/ecommerce-data>
2. The below dataset is going to be used in the products table
<https://www.kaggle.com/datasets/mukuldeshantri/ecommerce-fashion-dataset>
3. The below link was used to create the manufacturer table
<https://www.kaggle.com/code/kerneler/starter-manufacturer-sql-dataset-e70be8e3-2/data?select=sample-data.sql>

Tables

The tables that I am planning to use for this project are:

Customer(cust_id,name,city,state)

products(prod_id,Category,prod_title,brand,mrp,price,offer,stock_availability)

order(order_id,date,cust_id)

order_summary(order_id,prod_id,quantity)

brands(id,companyName,ContactName,city,country,phone)

Data cleaning and pre-processing

Tools used - Google sheets and vscode

The following data pre-processing and cleaning were done/going to be done

- [Dataset 1](#) was loaded into google sheets to preview. I split this dataset's data into 2 tables, Customer and orders tables. Duplicate customers were deleted in order to get a list of all unique customers which along with city and state is being used in the customer table. Unique cust_id was assigned to each row.
- [Dataset 2](#) was loaded into a google Sheets to preview. Uniq column was renamed to prod_id and columns that were not going to be used in the project were deleted (Crawl Timestamp, Site Name, Product Description, Pack Size Or Quantity, Combo Offers, Image URLs). The offer column was in a percentage format, it was reformatted to a floating point to make future calculations easier. Null values as well as rows with special characters were removed.
- [Dataset 3](#) was initially cleaned using vscode tools to remove extra characters and symbols. The data was then formatted into csv format and then duplicates and null values were removed using sheets.
- All the above-cleaned CSV files will then be loaded into their respective table via the /copy command that was used in week 1 of classes

Screenshots of the Customer table

cust_id	name	city	state
CA01	Bharat	Gujarat	Ahmedabad
CA02	Pearl	Maharashtra	Pune
CA03	Jahan	Madhya Pradesh	Bhopal
CA04	Divsha	Rajasthan	Jaipur
CA05	Kasheen	West Bengal	Kolkata
CA06	Hazel	Karnataka	Bangalore
CA07	Sonakshi	Jammu and Kashmir	Kashmir
CA08	Aarushi	Tamil Nadu	Chennai
CA09	Jitesh	Uttar Pradesh	Lucknow
CA10	Yogesh	Bihar	Patna
CA11	Anita	Kerala	Thiruvananthapuram
CA12	Shrichand	Punjab	Chandigarh
CA13	Mukesh	Haryana	Chandigarh
CA14	Vandana	Himachal Pradesh	Simla
CA15	Bhavna	Sikkim	Gangtok
CA16	Kanak	Goa	Goa
CA17	Sagar	Nagaland	Kohima
CA18	Manju	Andhra Pradesh	Hyderabad
CA19	Ramesh	Gujarat	Ahmedabad
CA20	Sarita	Maharashtra	Pune
CA21	Deepak	Madhya Pradesh	Bhopal

Screenshot of the product table

prod_id	category	brand	mrp	price	offer	stock_availability	prod_title
f0b68a06deb9253226d6a029b065f38	Make up	Homely 100pcs/bag Japan Rough Triangle Metal Circle Stud Rivet Nail Art Decoration 3D Nail Accessories DIY Nail Tools series 11 : 7					
3100ab469207d37b288e1e2da1bb001	Make up	Homely Fashlady™ F819: 2 Folhas De Lavanda Flor Etiqueta Do Prego Decalques De Á Gua Roxo Blooming Slides Diy Projeto Nail Art Decalques De Transferência	1817	678	0.33	YES	
c1a		Fashlady™	1949	1299	0.33	NO	
b8f4d3a59d5c36d3ccaa94740887d78	Make up	Fashlady™ 1 Sheet Nailmad Diamond Heart Nail Art Sticker Artascope 3D Nail Sticker Triangle Nail Sticker	1949	1299	0.33	YES	
2e0dfcb33b248a154693a1353d716f5	Bath & Shower	Fa Coconut Water Shower Gel - 250ml					
1f41d00d574e35e0436798d24051005c	Make up	Fa Veena OTwo0 Brush Set Soft And Easy To Wear Professional Make Up Brush Set Eyebrow Eyeshadow Eyeliner Lip Make Up Tools Kit 12 Pcs9962F01	595	495	0.17	NO	
bc8645d634457484d7886ce7d88646f2	Make up	Veena Homely 100pcs/bag Silver/Gold Lucky-Palm-Tree Mushroom non-adhesive Soft Metal Sticker Nail Art Decoration : MS347-2 Gold	1586.4	1065	0.29	NO	
fcc1a66a4bdf1f6d12f2dc94ba1f2f4f	Skin Care	Honey Philosophy Fresh cream Lip Polishing S...	713	713	0.00	YES	
01362502f4e6d328944399ad4a60af29	Make up	Philosophy Fashlady™ 04: Nail Art Stamping Placa Rendas Selo Do Prego De AÇ O InoxidÁ Vel Manicure Ferramenta De Impressora Polaco Padrô Es Selo Carimba	5989	5989	0.00	YES	
d8b47403d489ebfd3f839dbfb29a682c	Skin Care	Fashlady™ 04: Nail Art Stamping Placa Rendas Selo Do Prego De AÇ O InoxidÁ Vel Manicure Ferramenta De Impressora Polaco Padrô Es Selo Carimba	2949	1299	0.33	YES	
640174ff6d5a03a843320ce4d0c52625	Fragrance	BuyChoice Acne Scars Cream Acne Treatment Face Care Anti Acne Removal Gel Whitening Moisturizing Cream	1197	798	0.33	YES	
62806229bc505f63a4e0257b79b8c65	Skin Care	Parag Fragrances Cool Woman Spray Perfume 50ml (Long Lasting / 100% Real Fragrance/Geniune Product) Best Long Lasting Perfume For Men and Women	2800	699	0.50	NO	
bf3a93a9db223f54e4d393d32a6ceb	Skin Care	[the SAEM] Snail Essential EX Wrinkle Solution Emulsion 150ml	6680	5840	0.13	YES	
		The Saem Xplus Bath Loofah(Pack of 3)	450	400	0.11	NO	