

CDACL-006 Market Analysis

Database credentials:

DB Name: **project_orders**

Host: 18.136.157.135

Domain Name: projects.datamites.com

Username: **dm_team20**

Password:)KiC1mip045X

Task:

1. What are the top 10 aisles with the highest number of products?
2. How many unique departments are there in the dataset?
3. What is the distribution of products across departments?
4. What are the top 10 products with the highest reorder rates?
5. How many unique users have placed orders in the dataset?
6. What is the average number of days between orders for each user?
7. What are the peak hours of order placement during the day?
8. How does order volume vary by day of the week?
9. What are the top 10 most ordered products?
10. How many users have placed orders in each department?
11. What is the average number of products per order?
12. What are the most reordered products in each department?
13. How many products have been reordered more than once?
14. What is the average number of products added to the cart per order?
15. How does the number of orders vary by hour of the day?
16. What is the distribution of order sizes (number of products per order)?
17. What is the average reorder rate for products in each aisle?
18. How does the average order size vary by day of the week?
19. What are the top 10 users with the highest number of orders?

20. How many products belong to each aisle and department?

Also, analyze customer purchasing behavior and product performance to optimize marketing strategies and improve customer satisfaction.

Attributes information:

1. Aisles Dataset:

- aisle_id: Unique identifier for the aisle. (**Primary Keys**)
- aisle: Name of the aisle.

2. Departments Dataset:

- department_id: Unique identifier for the department. (**Primary Keys**)
- department: Name of the department.

3. Order Products Prior Dataset:

- order_id: Unique identifier for the order. (**Primary Keys**)
- product_id: Unique identifier for the product.
- add_to_cart_order: Order in which the product was added to the cart.
- reordered: Indicates whether the product was reordered in this order (1 for reordered, 0 for not reordered).

4 Orders Dataset:

- order_id: Unique identifier for the order. (**Primary Keys**)
- user_id: Unique identifier for the user.
- eval_set: Evaluation set (prior, train, test).
- order_number: Order sequence number for the user.
- order_dow: Day of the week the order was placed.
- order_hour_of_day: Hour of the day the order was placed.
- days_since_prior_order: Number of days since the last order.

5. Products Dataset:

- product_id: Unique identifier for the product.

- product_name: Name of the product.
- aisle_id: Identifier for the aisle the product belongs to. (**Primary Keys**)
- department_id: Identifier for the department the product belongs to.

Note: For Merging the table, you have to use columns where specified as Primary Keys.

Note: You'll need to perform data manipulation, aggregation, merging, and analysis across multiple datasets, such as joining orders with products, aisles, and departments to get all the required information.

You can provide your inputs/solution as a PPT presentation and you can explain your project, record it and send it along with the PPT file.