

#### **Abstract**

The goal of the Women's E-Commerce Dataset project is to help the marketing department better target their markets and how they can use their customers' feedback in order to improve their business and products. The unique goal of the project is to know the products with the highest number and rating and the number of products for each type to make a successful marketing plan to help attract their customers.

# Design

Statistical analysis of the project on e-commerce reviews is designed based on the number of reviews for the reactions of each product, the effectiveness of each product, the satisfaction and attraction of customers to buy it again, and the number of times each product is evaluated.

#### **Data**

The project's statistical analysis data on e-commerce reviews contains a large number of product names that customers recommend to buy and high ratings for each product. The data model I chose is about 20 products and how each product affects each product's average life, rating, and customer recommendation for each product. The dataset contains 23,486 rows and 10 feature variables.

# **Algorithms**

- 1. Load Dataset
- 2. Explore Dataset
  - Work to discover data through some methods such as :
  - df.head() , df.info() .

#### 3. Cleaning Dataset

- **Clean** the data, identify the rows and columns that contain an missing value, and delete the rows that we do not need.
- Rename some rows to have meaning in the dataset.

## 4. Methodology of exploring the data:

After cleaning the data, identifying the rows that contain a missing value, deleting them, renaming some columns, and then analyzing the data. Then, calculate the average of the most rated products for the products and calculate the average ages for each product's recommendation. Calculate the number of product values for each category

## **Tools**

- Python and Jupyter Notebook
- Pandas for data manipulation
- Matplotlib and Seaborn for plotting visuialization

## **Communication**





