EDA Report: women's clothing e-commerce reviews

Dataset: Women's Clothing E-Commerce Reviews Dataset

Data Overview:

\* The dataset contains information about women's clothing e-commerce reviews.

\* It has 23486 rows × 13 columns

\* The columns include information such as customer age, division, department, product name, rating, and recommendation status.

Data Cleaning and Preprocessing:

\* Dropped unnecessary columns and rows.

\* Handled missing values by dropping records with null values.

\* Removed duplicate records.

\* Renamed columns for clarity.

Data Exploration:

Customer Age Distribution:

\* Most customers are between 20 and 40 years old.

\* The distribution is relatively uniform across this age range.

Correlation Matrix:

\* There is a moderate positive correlation between age and rating.

\* There is a weak positive correlation between age and price.

Division and Department Analysis:

\* Tops and Dresses are the most popular divisions.

\* Intimate and Sleep are the least popular divisions.

\* Juniors and Intimates are the most popular departments.

Product Name Analysis:

\* Pants are the most popular product name.

\* Dresses and Sweaters are also popular product names.

Average Ratings for Each Product:

\* The average rating for most products is around 4.0.

\* Some products, such as Dresses and Pants, have slightly higher average ratings.

Percentage of Recommended Products:

\* Approximately 80% of the products are recommended.

Relationship between Age and Product Preferences:

\* There is no clear relationship between age and product preferences.

\* Customers of all ages purchase a variety of products.

Conclusion:

\* The dataset contains a diverse range of women's clothing reviews.

\* There are some interesting trends and patterns in the data.

\* The data can be used to gain insights into customer preferences and to improve product recommendations.