Muzaffer Estelik

Prediction of Ratings
Based on Women's E-Commerce Reviews

Data Science Career Track Capstone-2 Project

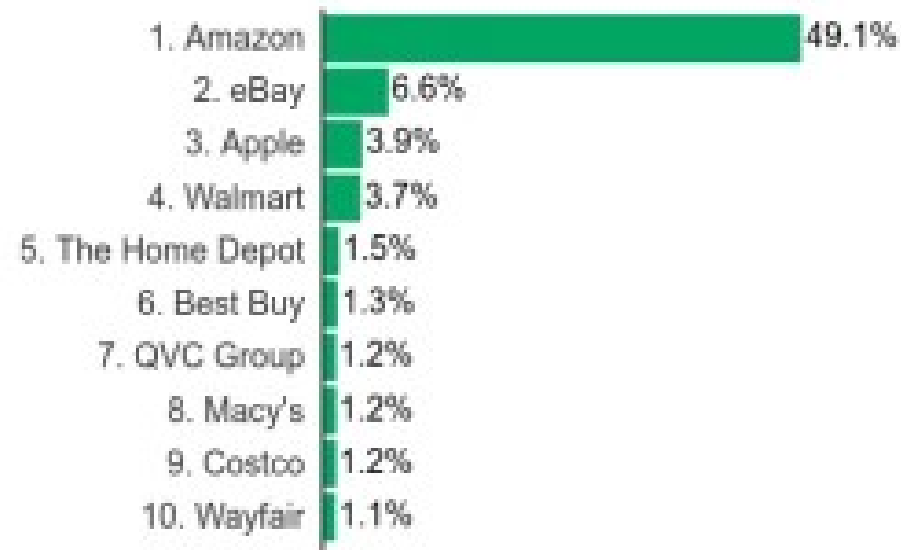**Oct 29th, 2018**

Table of Contents

*   **Problem**

*   **Data Set**

*   **Exploratory Data Analysis (EDA)**

*   **Machine Learning**

*   **Conclusion**

*   **Appendix**

Problem

Top 10 U.S. companies based on % of e-commerce sales

1. Amazon — 49.1%
2. eBay — 6.6%
3. Apple — 3.9%
4. Walmart — 3.7%
5. The Home Depot — 1.5%
6. Best Buy — 1.3%
7. QVC Group — 1.2%
8. Macy's — 1.2%
9. Costco — 1.2%
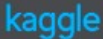10. Wayfair — 1.1%

Source: eMarketer, July 2018

E-commerce is becoming more important day by day, together with all the aspects of it.

Problem

- ✔ Promotion and discounts

- ✔ Content management

- ✔ Shopping cart and checkout process

- ✔ Search engine optimized code and layout

- ✔ Reporting tools

- ✔ Email marketing

Data Set

https://www.kaggle.com/nicapotato/womens-ecommerce-clothing-reviews

Exploratory Data Analysis (EDA)

Number of reviews per Age

25-50 is the most reviewing age group

Exploratory Data Analysis (EDA)

Rating Distribution per Age
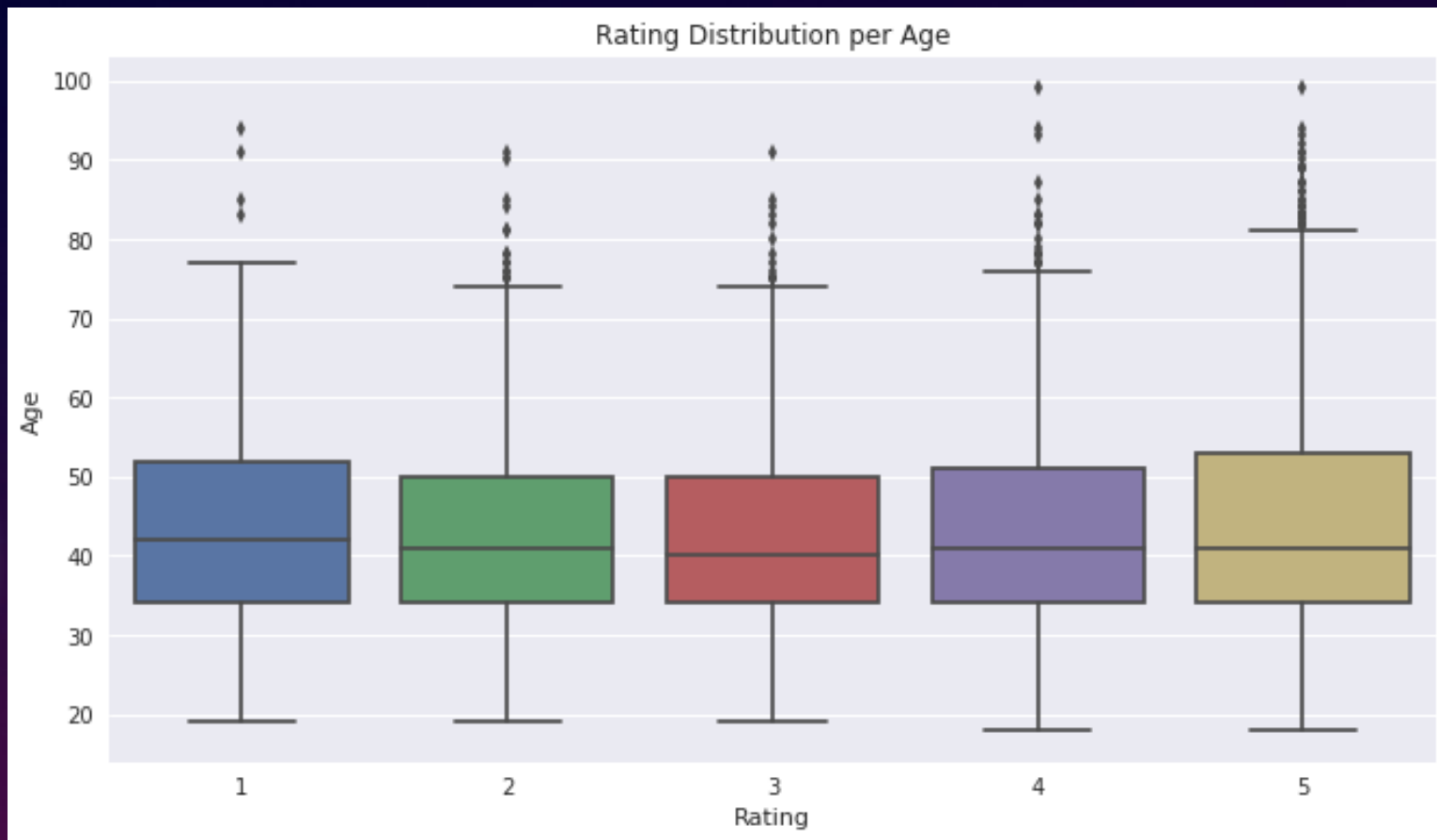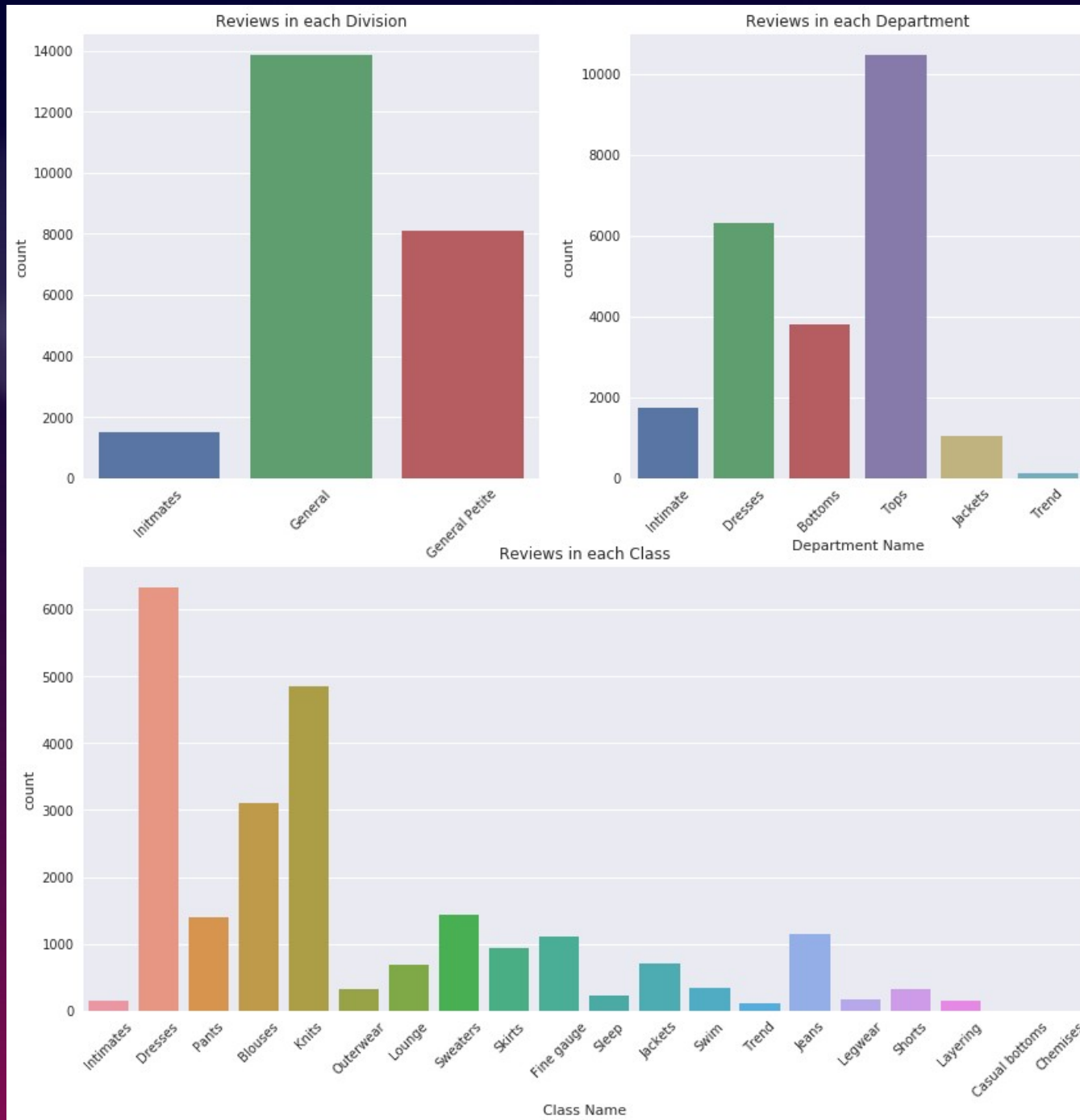
age is not a factor on the ratings of the reviews

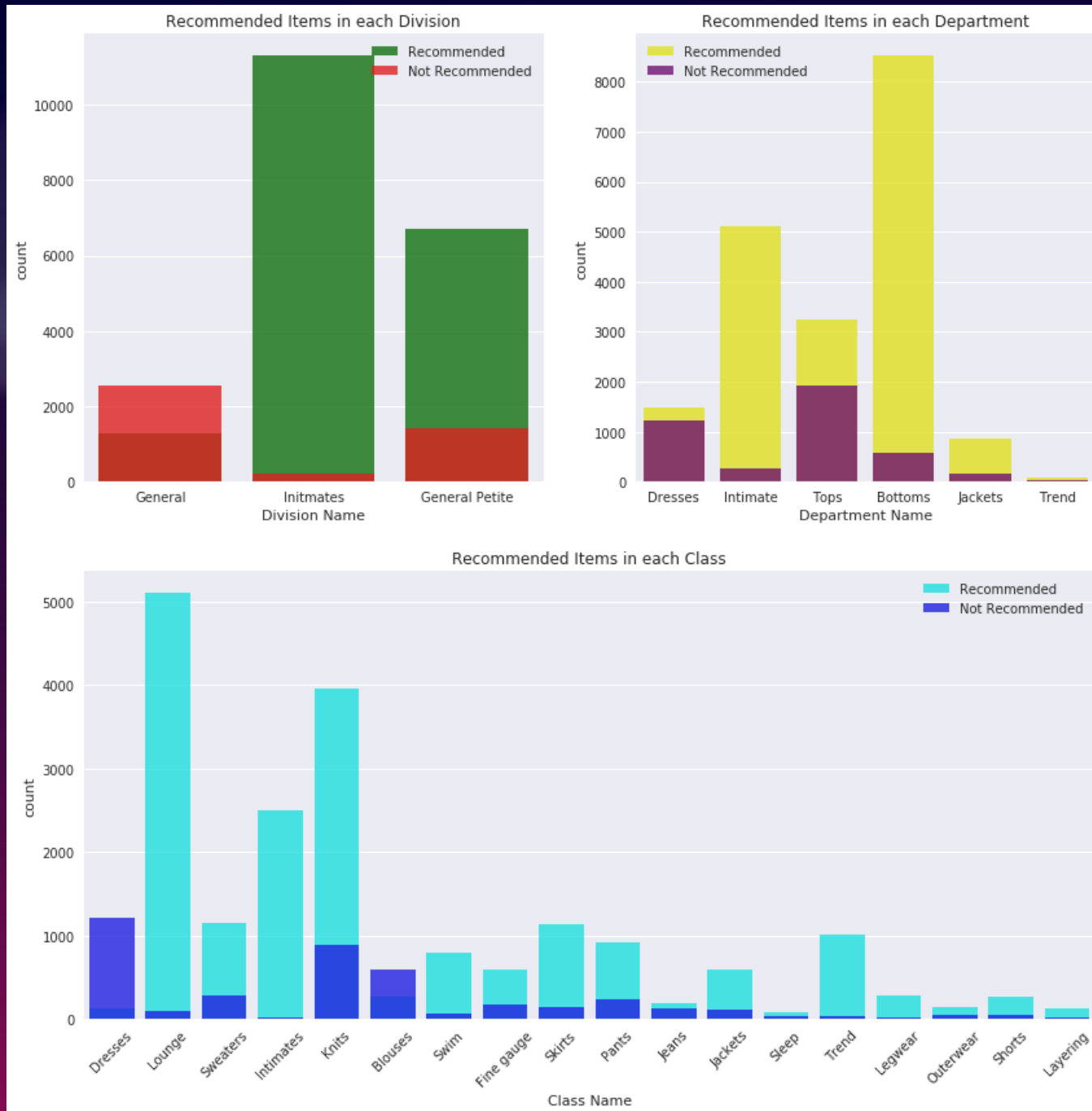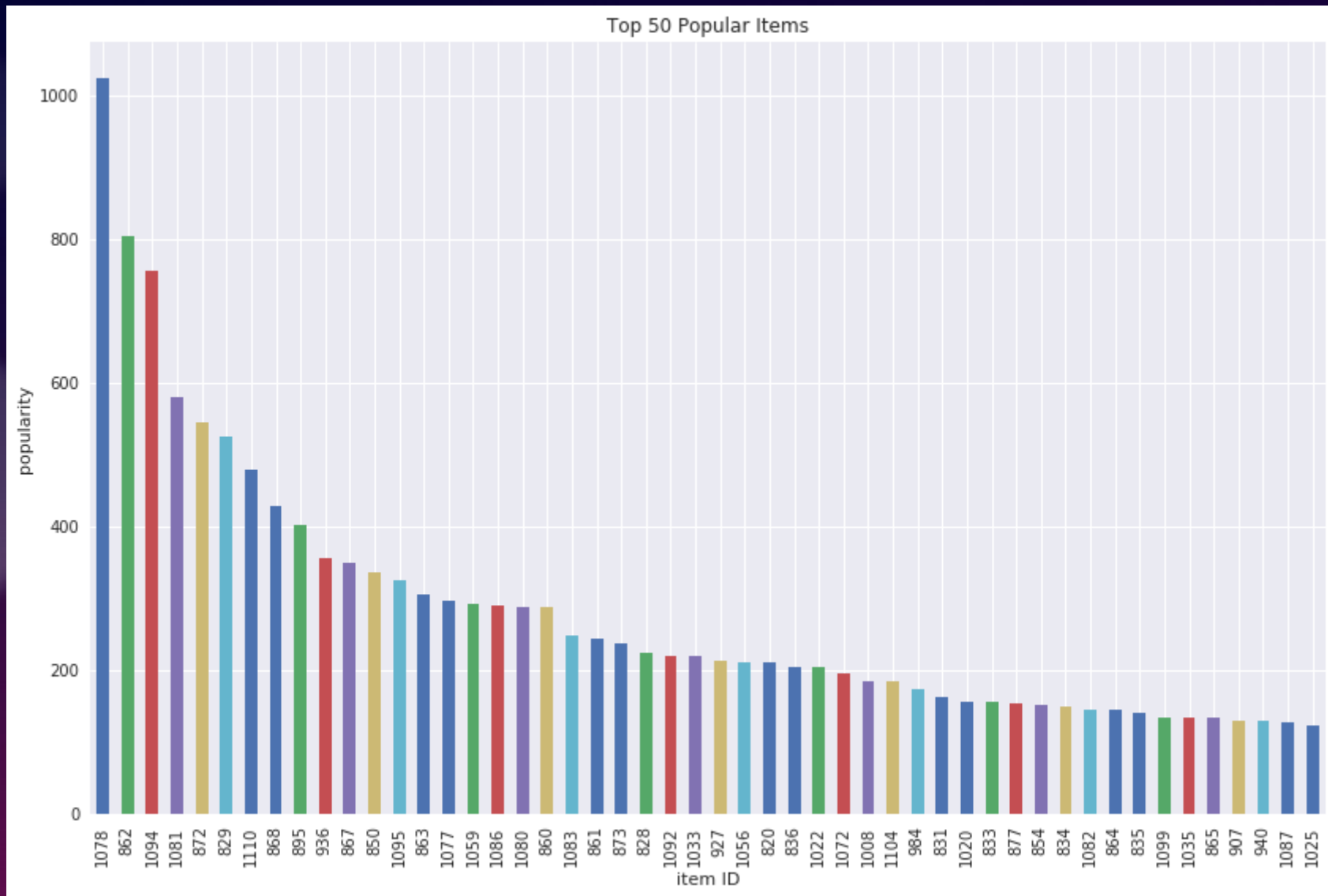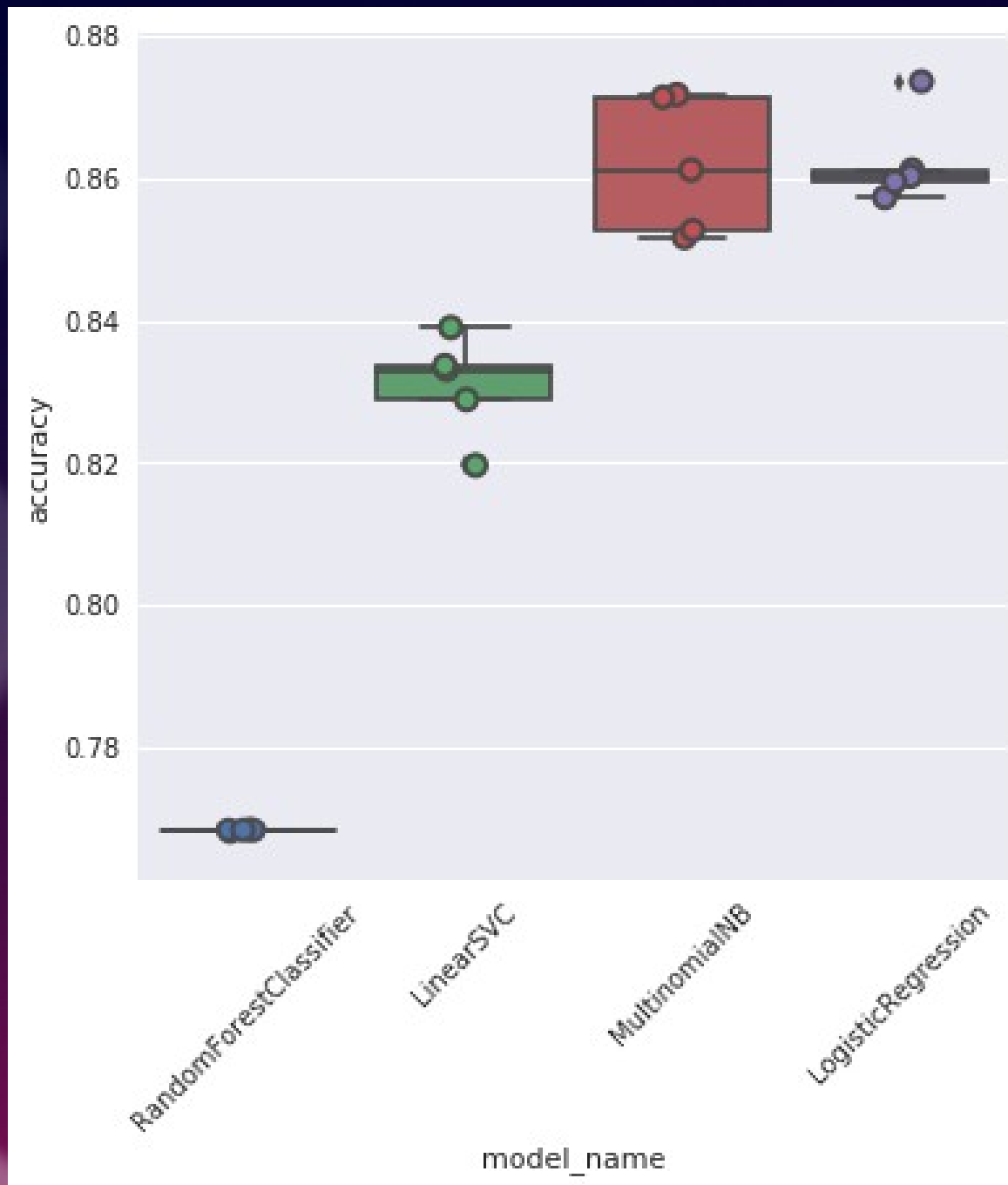# Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA)

Top 50 Popular Items

Machine Learning

Conclusion

Model Name	Train Accuracy Score	Test Accuracy Score
Linear SVC	0.830866	0.832948
Naive Bayes	0.861688	0.859608
Logistic Regression	0.862371	0.845893
Random Forest	0.768390	0.774387