



CUSTOMER PURCHASE DATA

Presented by: Kolade Adediji



OVERVIEW

01

Objective

02

Dataset

03

Data Cleaning

04

Visualizations

05

Insights

06

Machine Learning
model

07

Conclusion





GOALS AND OBJECTIVES

Analysis

This project focuses on drawing valuable insights from customer's data using important data visualization tools and python using jupyter notebooks functions

Machine Learning Model

Designing a model to predict customer data i.e whether a customer would purchase a good or not considering various features like their Age and sex and others and check the model with the highest accuracy and precision



DATASET

The dataset used in this analysis was sourced from kaggle data catalogue, accessible at the link:

<https://www.kaggle.com/datasets/rabieelkharoua/predict-customer-purchase-behavior-dataset>





DATASET

The Dataset contains 9 columns namely: Age, Gender, Annual Income, Number of Purchases, Product Category, Time Spent On Website, Loyalty Website, Discount Availed, Purchases Status. It also contains 1500 rows, contains no missing data and no duplicate

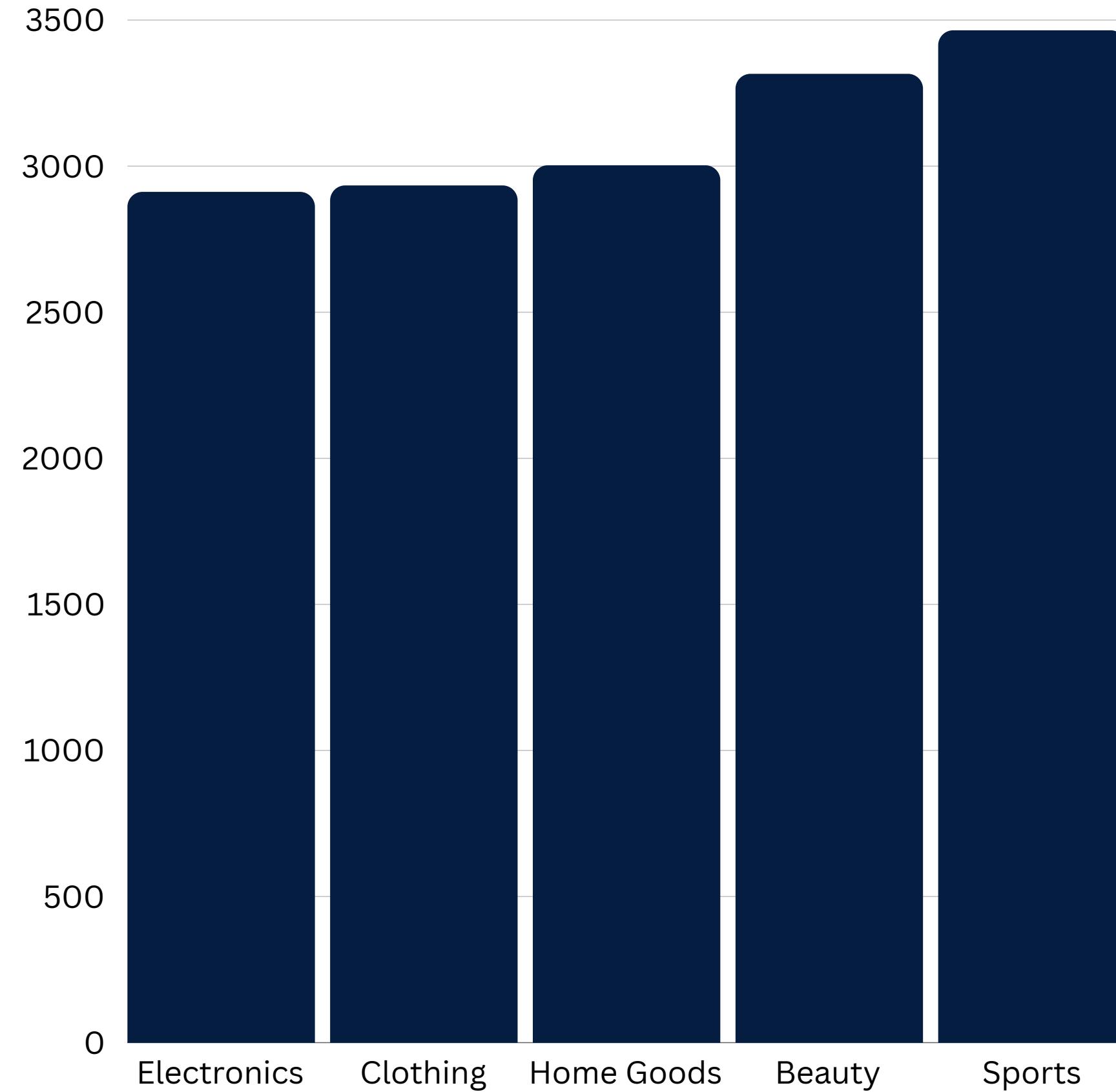




VISUALIZATIONS

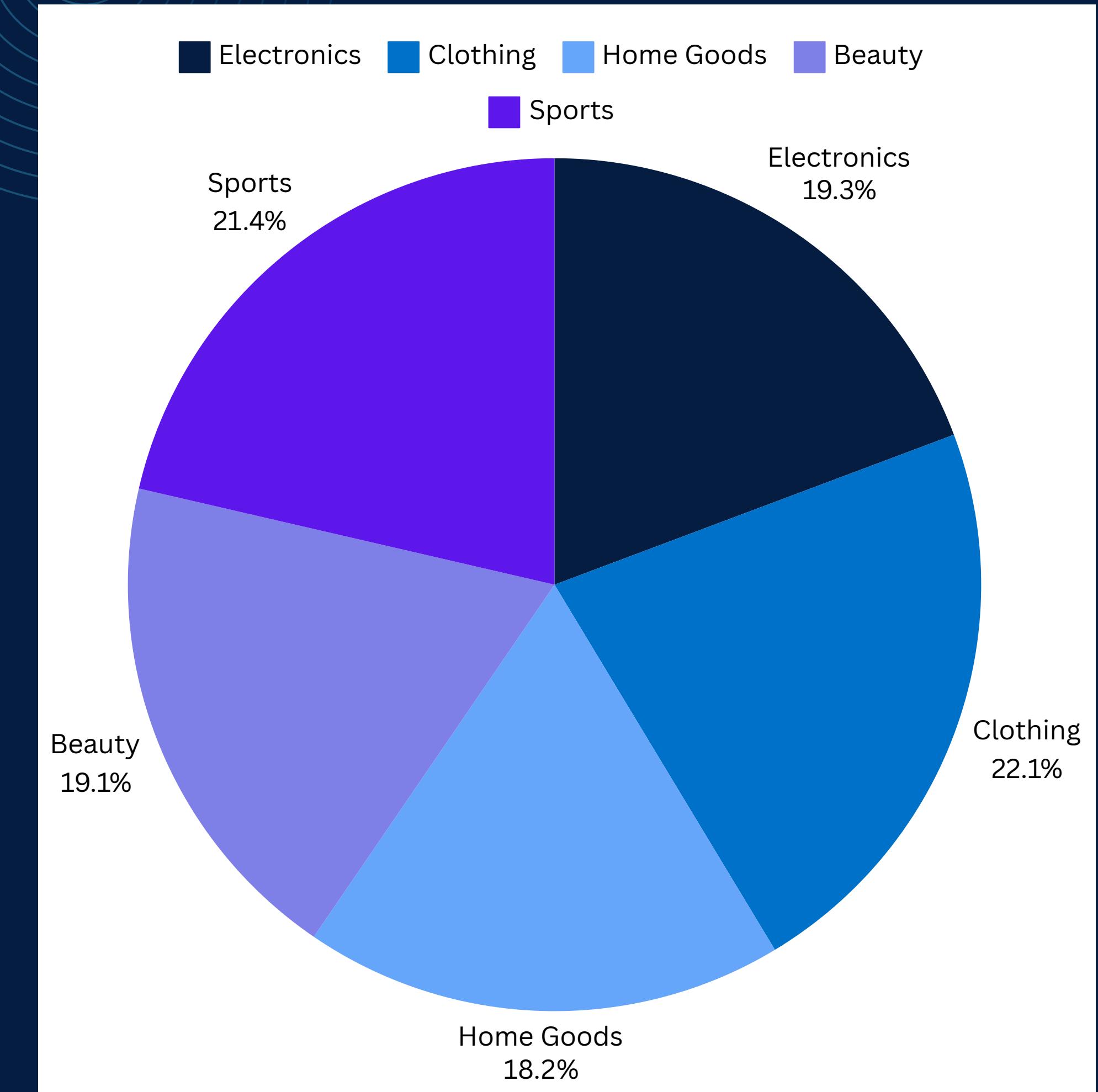
Number Of Purchases by Product Category

- The Sports section made more sales with Electronics making the least sales



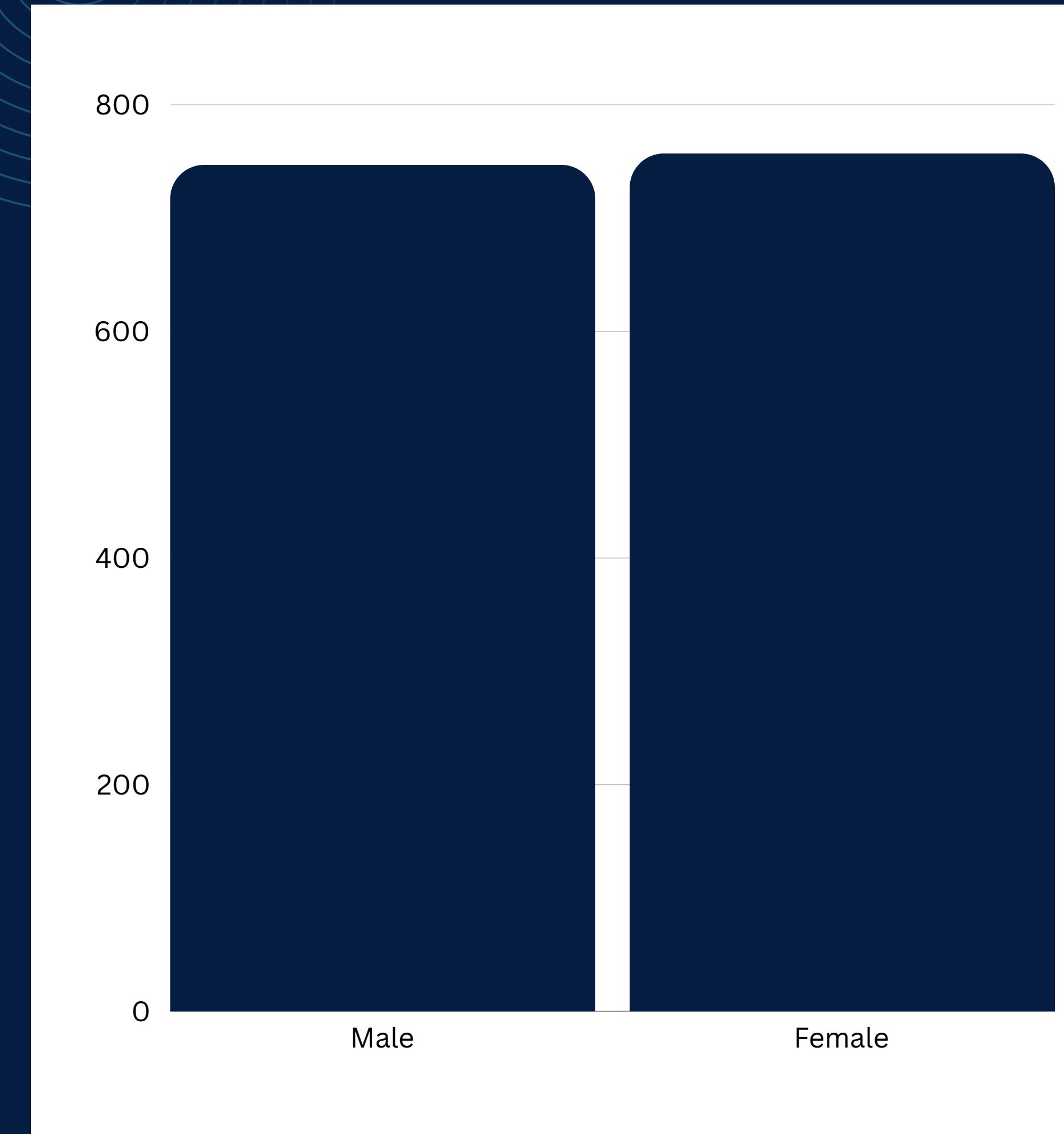
Number Of Purchases by Product Category

- The Sports section made more sales with Electronics making the least sales



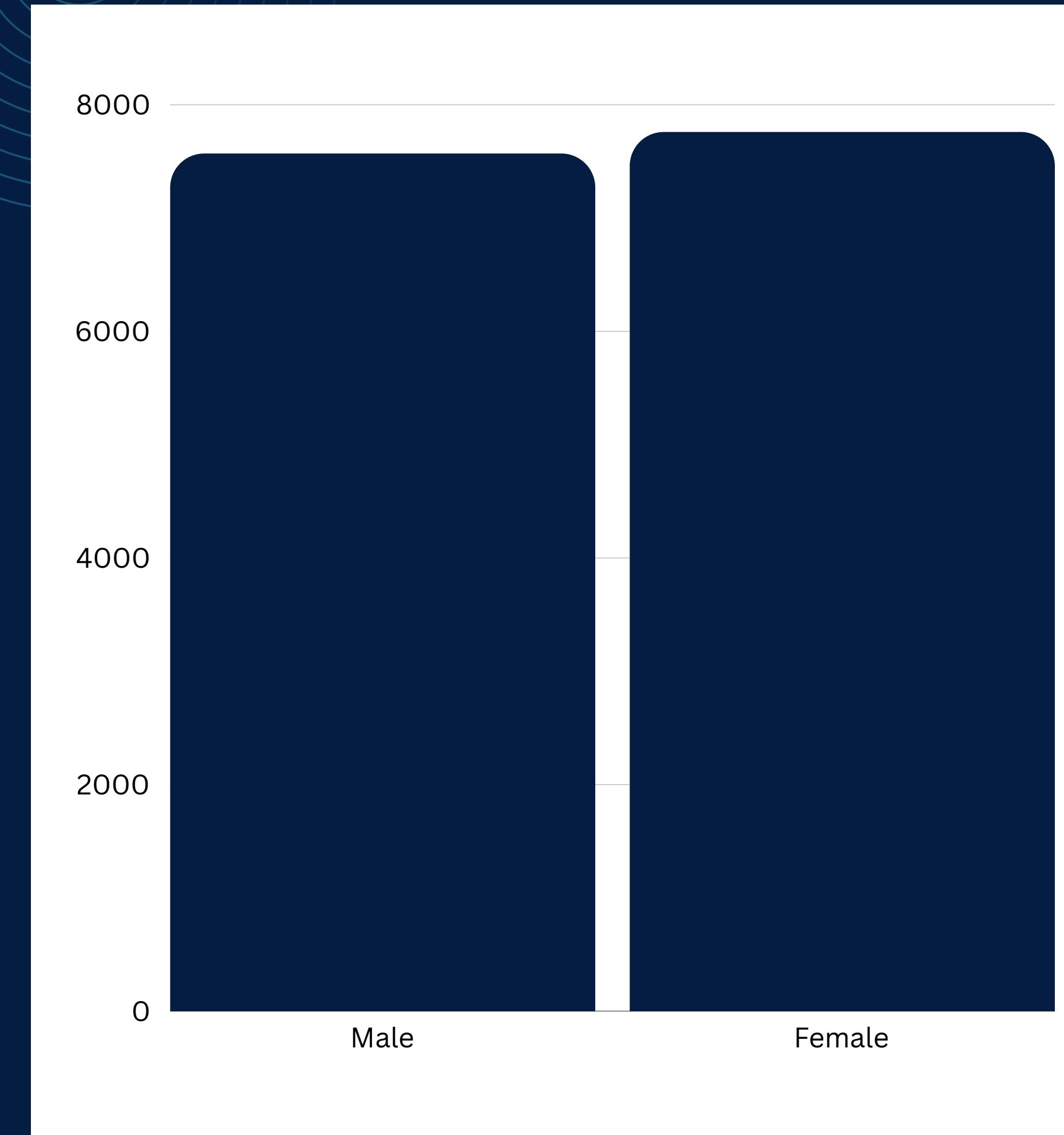
Distribution of Gender

- The dataset contains more females than male



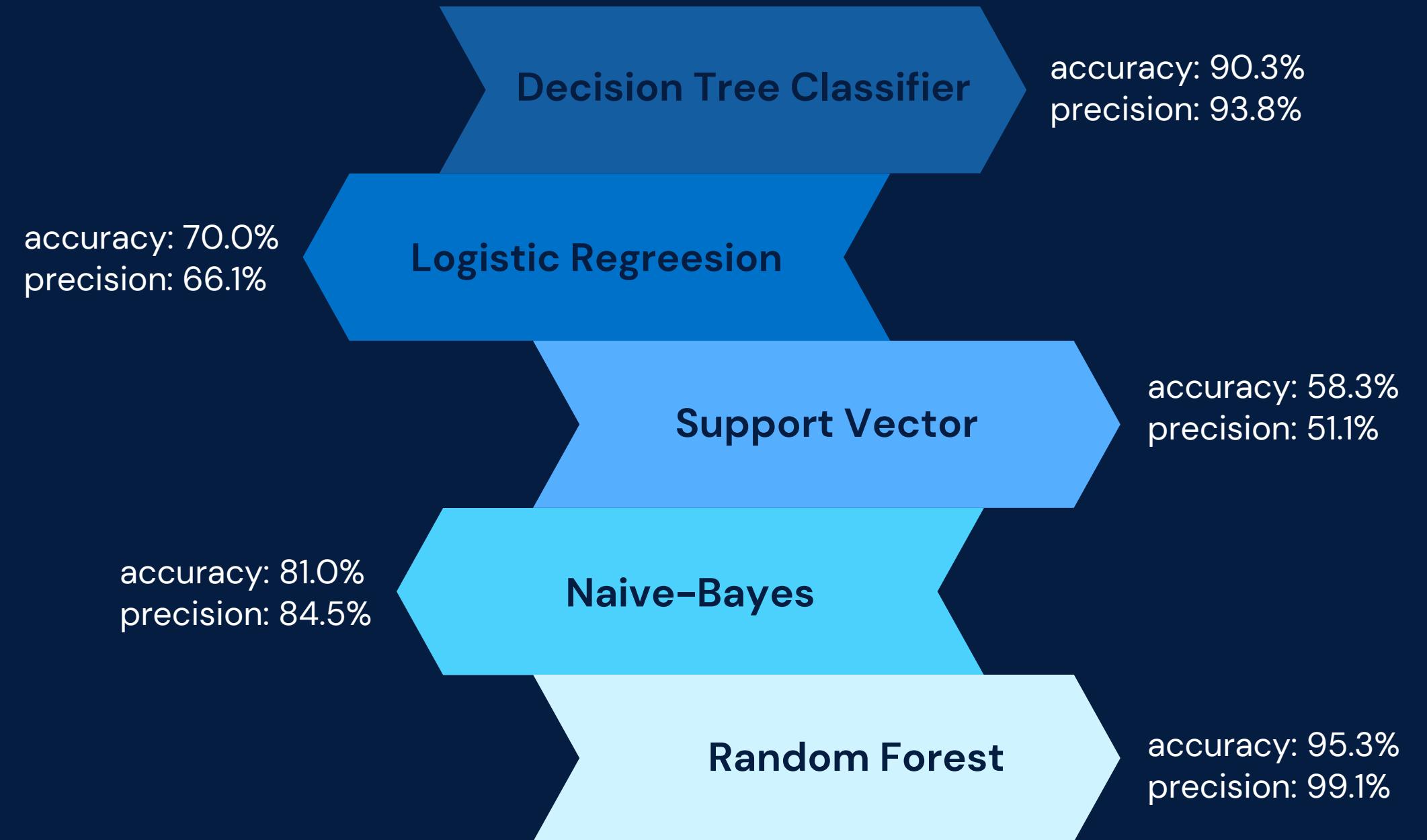
Distribution of Gender

- The females also made more purchases than the males





MODELLING



STATISTICS

(RANDOM FOREST CLASSIFIER)

95%

Accuracy

The Random Forest Classifier Model gave the highest accuracy of all the model used with 95%

99%

Precision

It also gave the highest precision of all the model with 99%



CONCLUSION

To get more Insights and more views about the project please check :

https://github.com/Koladeadediji/Customer_Purchase_Data