**PYTHON DATA VISUALISATION PROJECT SUMMARY**

ROLL NO:221BCADA26

**DATASET**: Customer Shopping Dataset - Retail Sales Data

**LINK**: https://www.kaggle.com/datasets/mehmettahiraslan/customer-shopping-dataset/

**DESCRIPTION:**

This dataset contains shopping information from 10 different shopping malls in the city of Istanbul between 2021 and 2023. We have gathered data from various age groups and genders to provide a comprehensive view of shopping habits in Istanbul. The dataset includes essential information such as invoice numbers, customer IDs, age, gender, payment methods, product categories, quantity, price, order dates, and shopping mall locations. We hope that this dataset will serve as a valuable resource for researchers, data analysts, and machine learning enthusiasts who want to gain insights into shopping trends and patterns in Istanbul.

**DATA FIELDS:**

* **invoice\_no:** Invoice number. Nominal. A combination of the letter 'I' and a 6-digit integer uniquely assigned to each operation.
* **customer\_id:** Customer number. Nominal. A combination of the letter 'C' and a 6-digit integer uniquely assigned to each operation.
* **gender:** String variable of the customer's gender.
* **age:** Positive Integer variable of the customers age.
* **category:** String variable of the category of the purchased product.
* **quantity:** The quantities of each product (item) per transaction. Numeric.
* **price:** Unit price. Numeric. Product price per unit in Turkish Liras (TL).
* **payment\_method:** String variable of the payment method (cash, credit card or debit card) used for the transaction.
* **invoice\_date:** Invoice date. The day when a transaction was generated.

**OBJECTIVES:**

To find out the interest and shopping patterns of the public for example

1. How does the category of the products differs while shopping in the case of both male and female
2. How are most of the transactions done (Cash, Credit, Debit)
3. Accordingly, what products has highest and lowest selling prices
4. Which is the most visited mall by the public

**FUTURE SCOPE:**

Further analysis in the dataset can also help us to create a model to predict the prices based on the types of products people buy.

Supermarkets can tailor marketing and promotions to individual customers, improving customer engagement

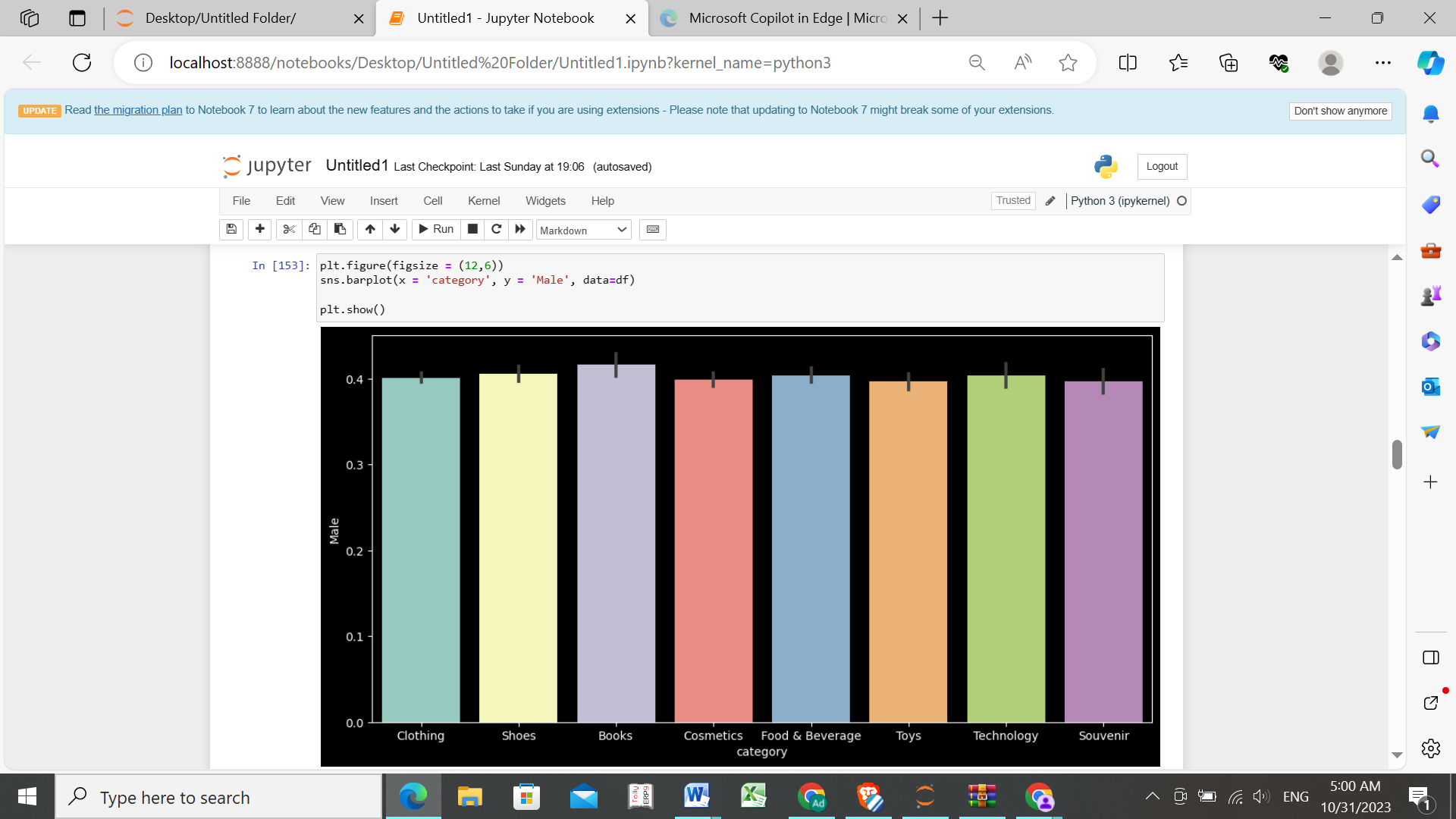
Identifying which products are commonly purchased together allows supermarkets to cross-sell and bundle products effectively.

**PROBLEM STATEMENT FOR THE DATASET**

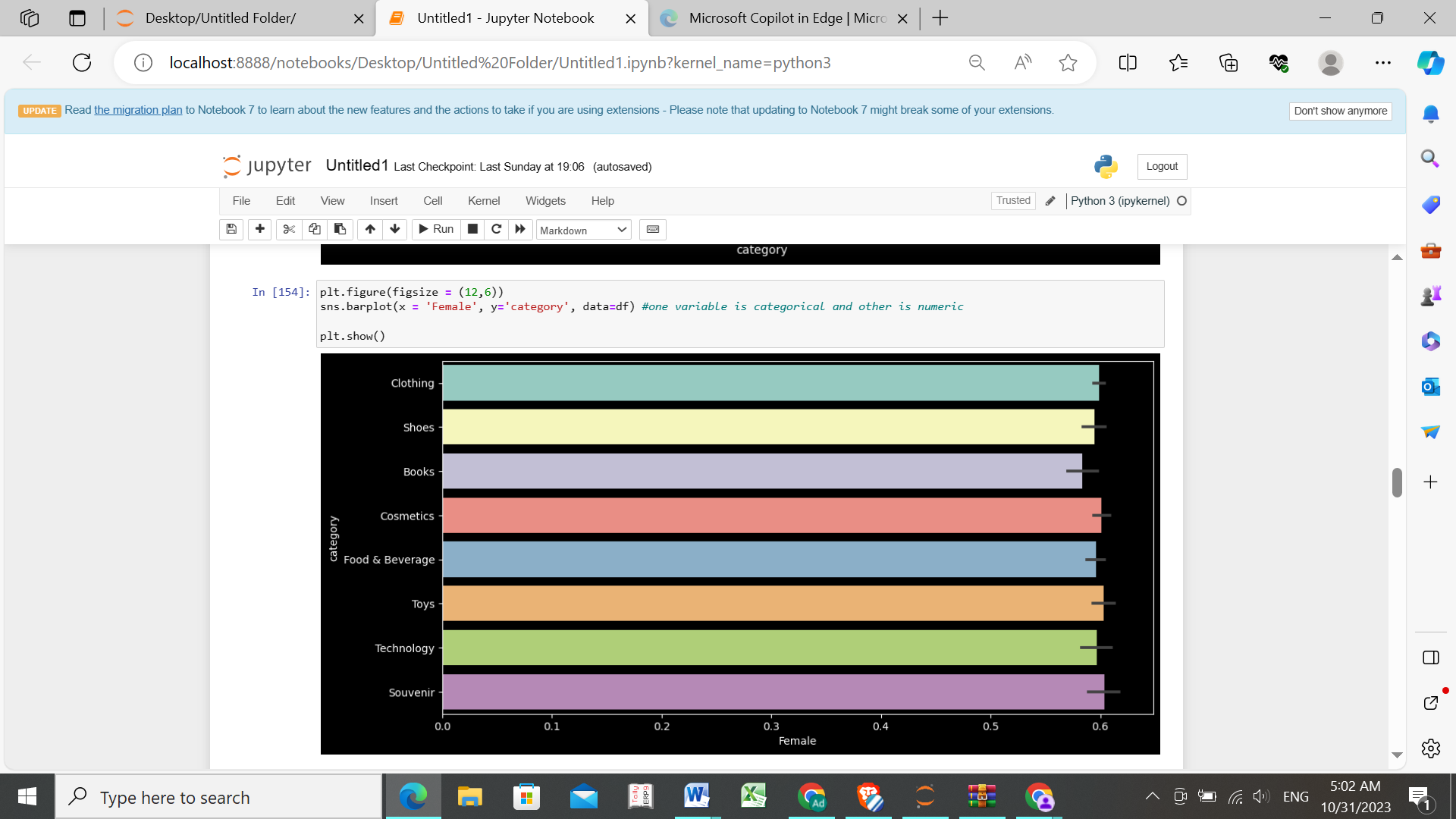
If an individual intends to establish a retail outlet within a shopping centre in Istanbul, which shopping mall is considered the most favoured choice, what products would bring more profit and what are the underlying reasons for this preference?

**IMPLEMENTATION**

How do the products differs while shopping in the case of both male and female

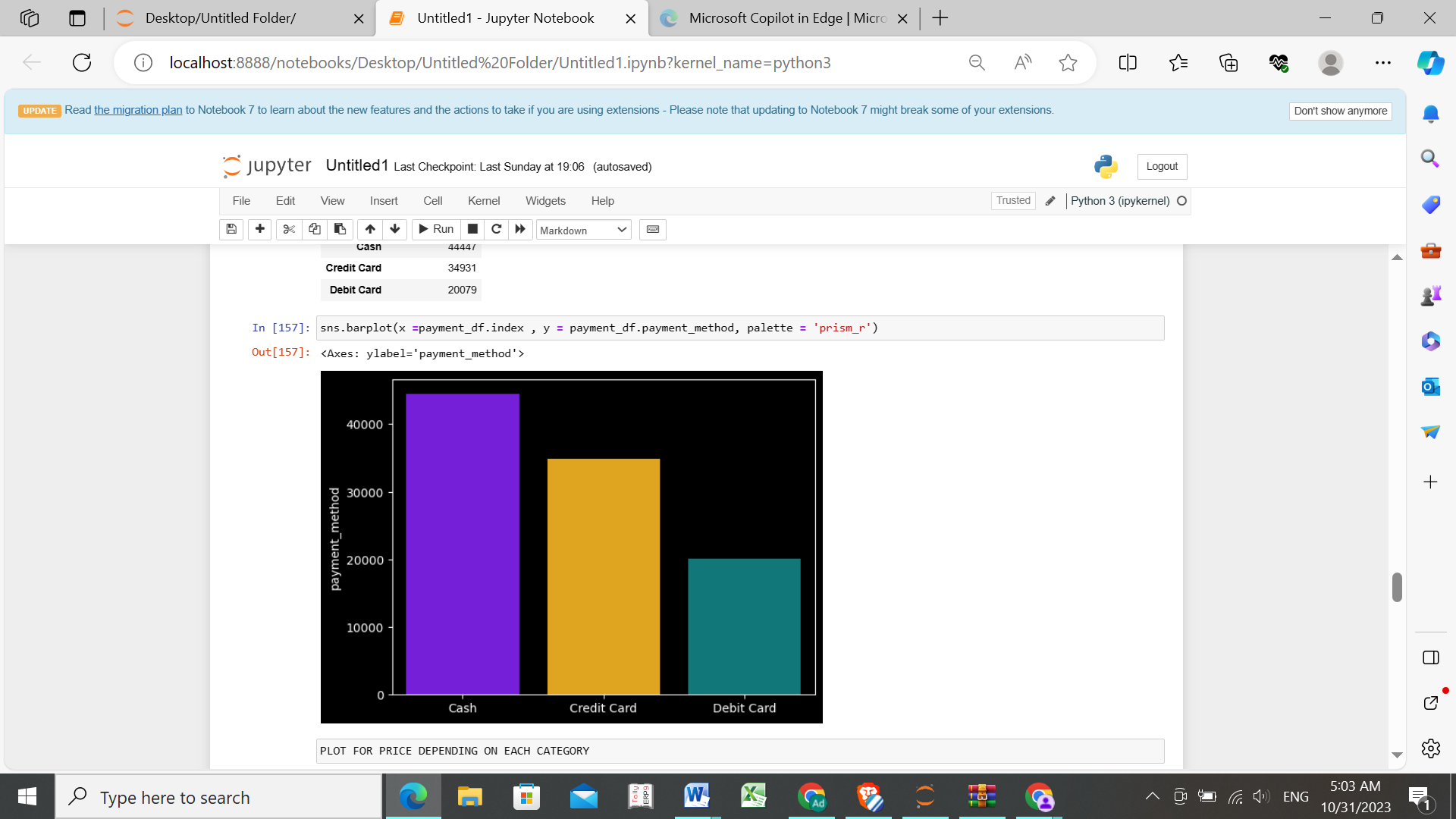


Most Male customers have interest in all the categories equally but the purchase of Books is maximum



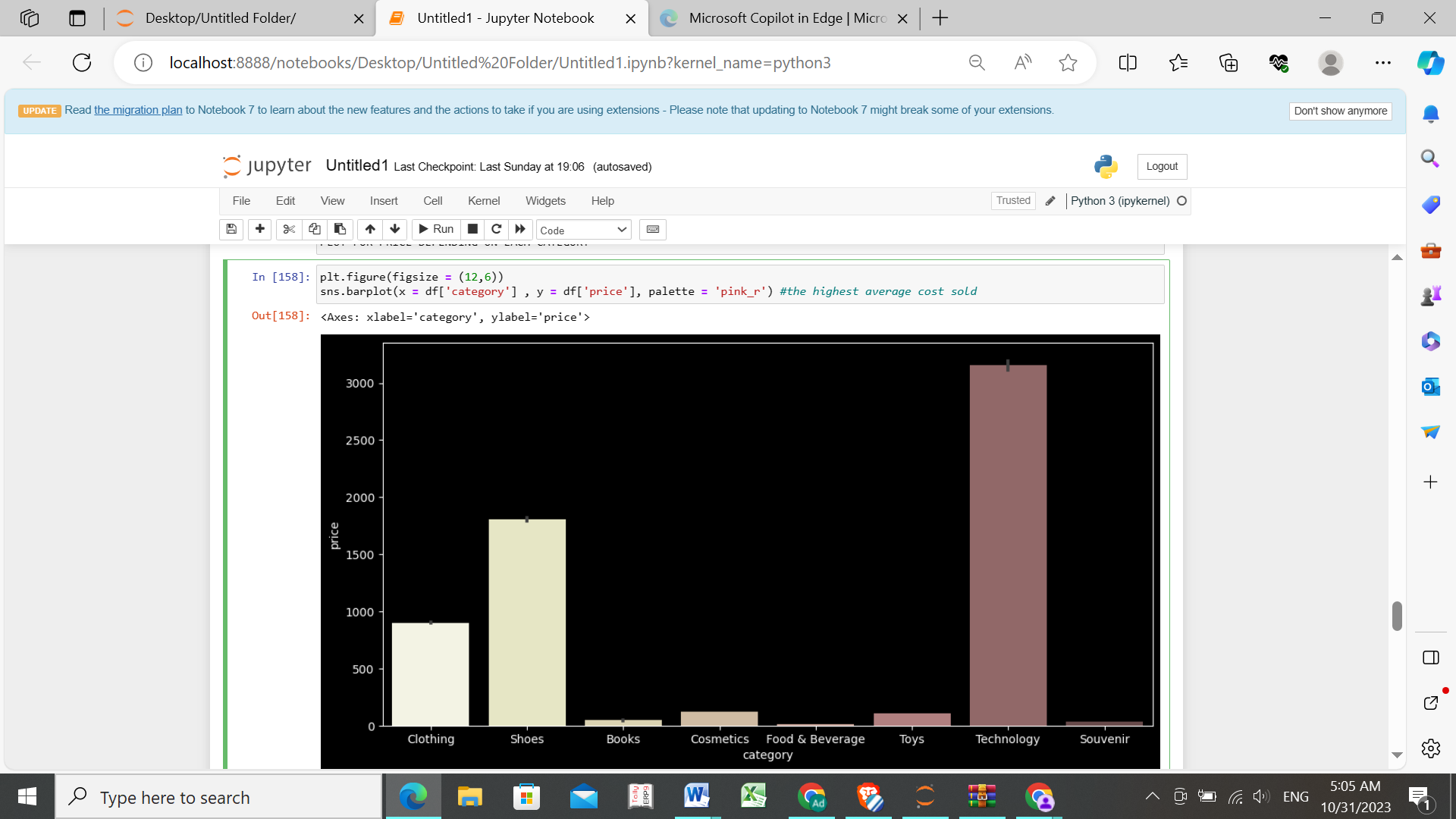
Most Female customers have interest in all the categories equally but the purchase of Toys and Souvenir is maximum

How are most of the transactions done (Cash, Credit, Debit)



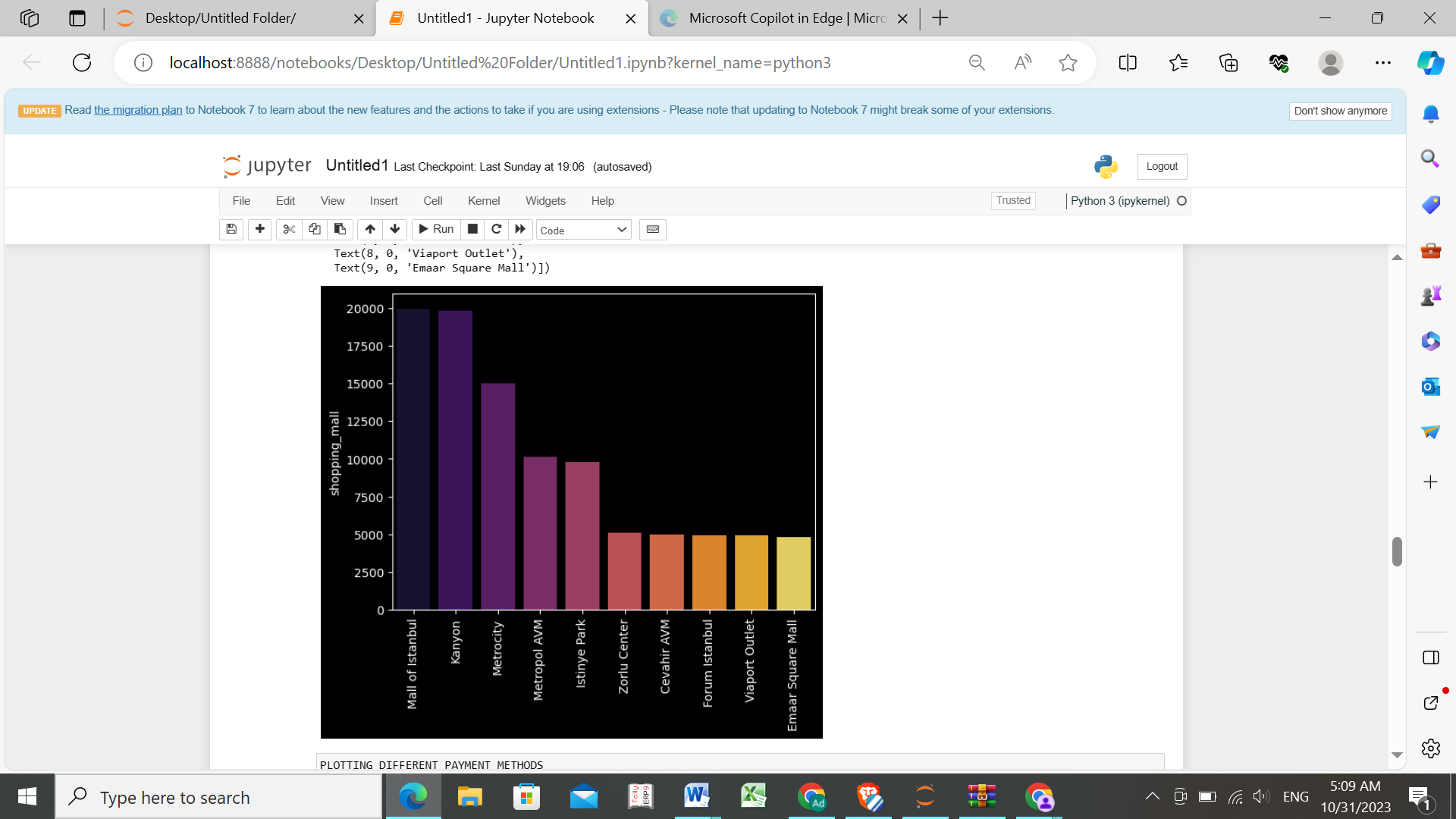
Most of the transactions are done in cash

Accordingly, what products has highest and lowest selling prices



Highest: Technology  
Lowest: Souvenir/Food&Beverages

Which is the most visited mall by the public



Mall of Istanbul is the most visited mall

PROBLEM STATEMENT FOR THE DATASET

If an individual intends to establish a retail outlet within a shopping centre in Istanbul, which shopping mall is considered the most favoured choice, what products would bring more profit and what are the underlying reasons for this preference?

Conclusion:

A person can establish a retail outlet at the *Mall of Istanbul* mall as it has the least number of customers. Since there is more number of Females all over, the products that would be popular are *toys* and *souvenir* based on our analysis above and therefore the person can invest in similar products.