

# Data Mining on e-Commerce Behavior Data from Multi-category Store

Fajrur Rahman Suprapto  
[fajrur.rs@gmail.com](mailto:fajrur.rs@gmail.com)

# ABOUT DATASET

This dataset contain behavior data on October 2019 from a large multi-category online store. There are 9 columns and 42,448,764 rows in this dataset. The dataset can be found on [Kaggle](#).

Property	Description
event_time	Time when event happened at (in UTC).
event_type	Only one kind of event: purchase.
product_id	ID of a product
category_id	Product's category ID
category_code	Product's category taxonomy (code name) if it was possible to make it. Usually present for meaningful categories and skipped for different kinds of accessories.
brand	Downcased string of brand name. Can be missed.
price	Float price of a product. Present.
user_id	Permanent user ID.
** user_session**	Temporary user's session ID. Same for each user's session. Is changed every time user come back to online store from a long pause.

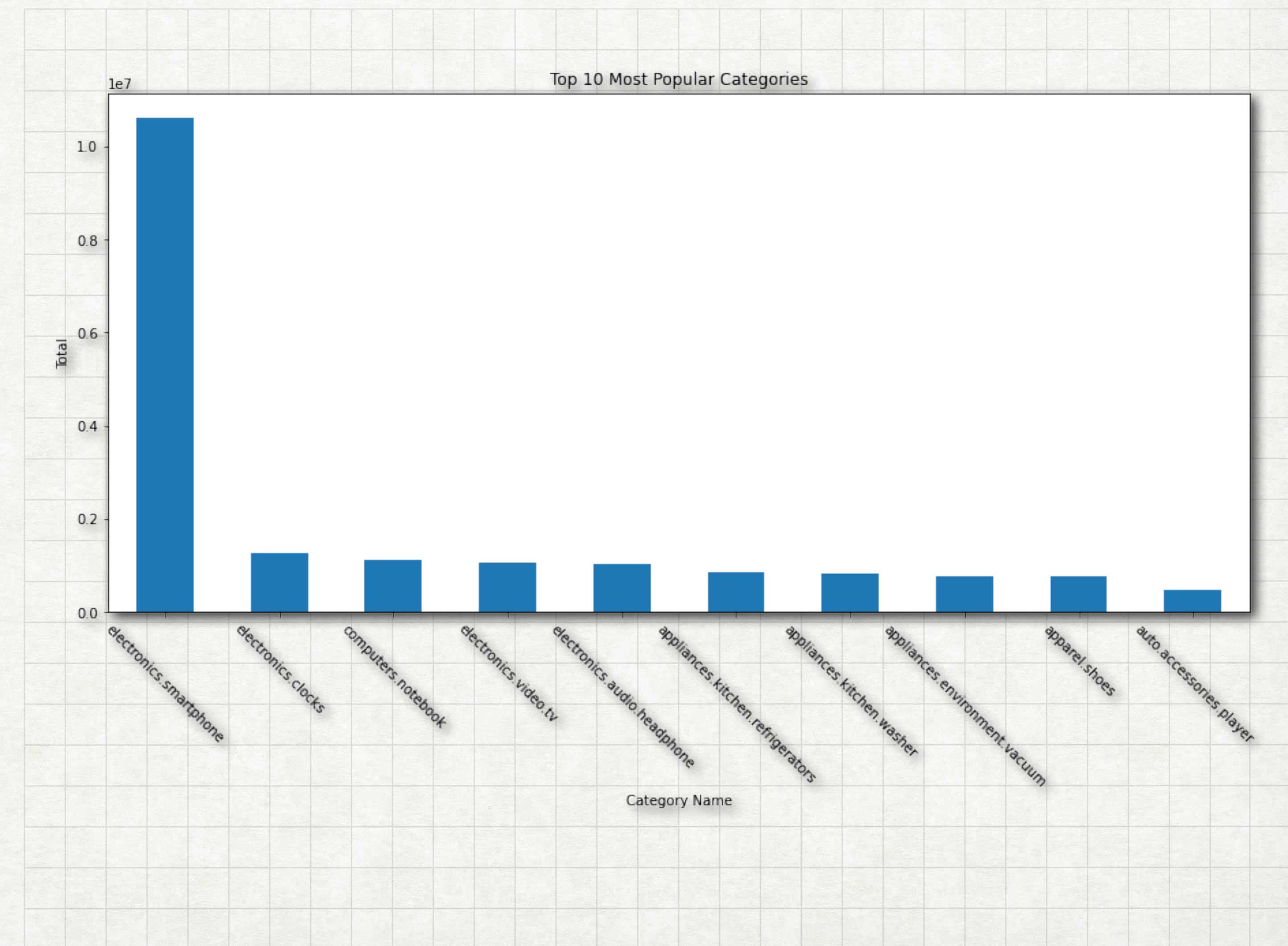
# BUSINESS QUESTIONS

This project focus on 5 questions below:

- What are the most popular categories and brands?
- What are the most purchase categories and brands?
- Who are the most customers with the highest number of purchase?
- How about daily visitor of this month?
- How about costumer activity?

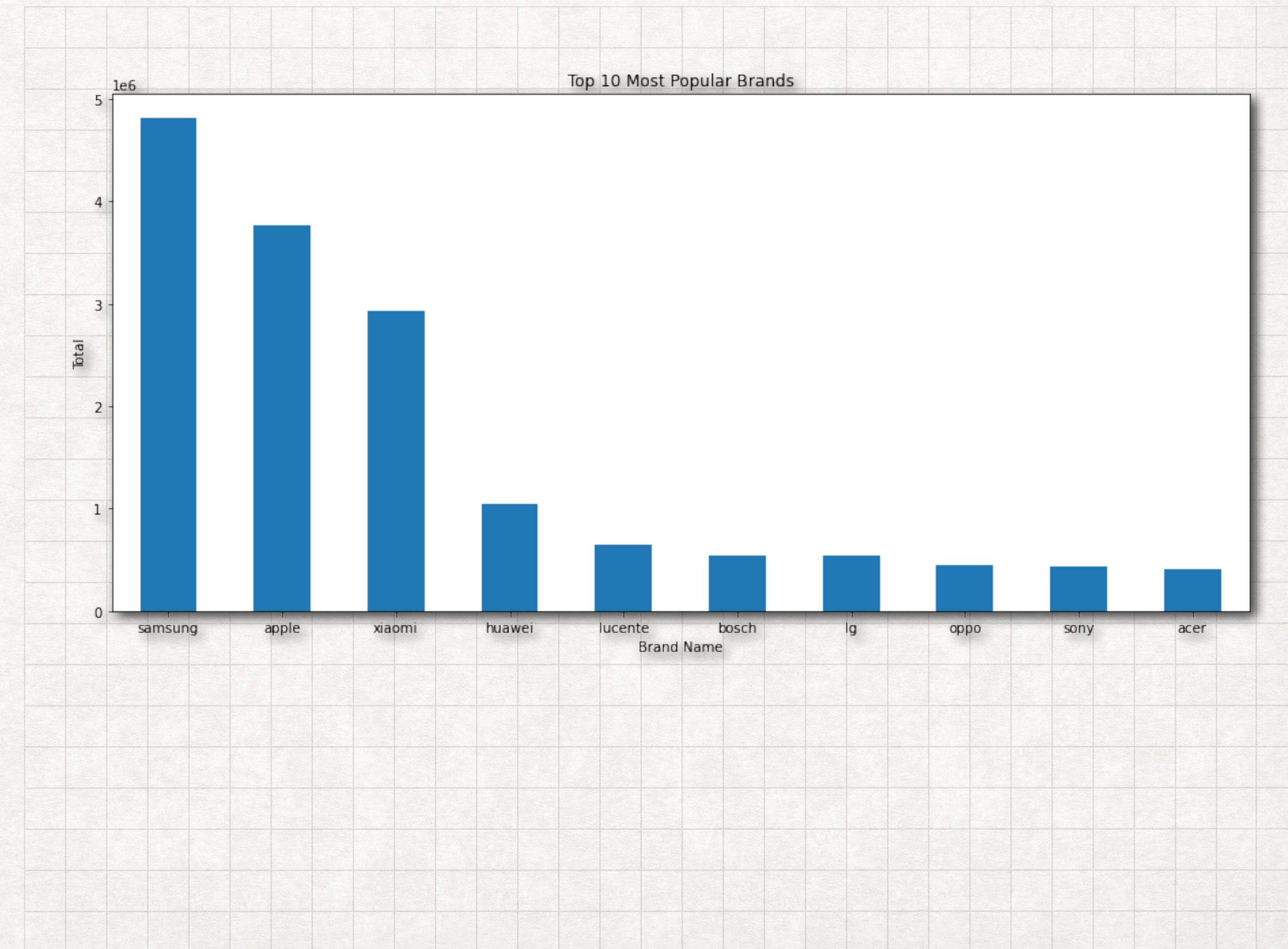
# MOST POPULAR CATEGORIES AND BRANDS

Smartphone became the most popular category in October 2019. This category has been viewed more than 110,000,000 times.



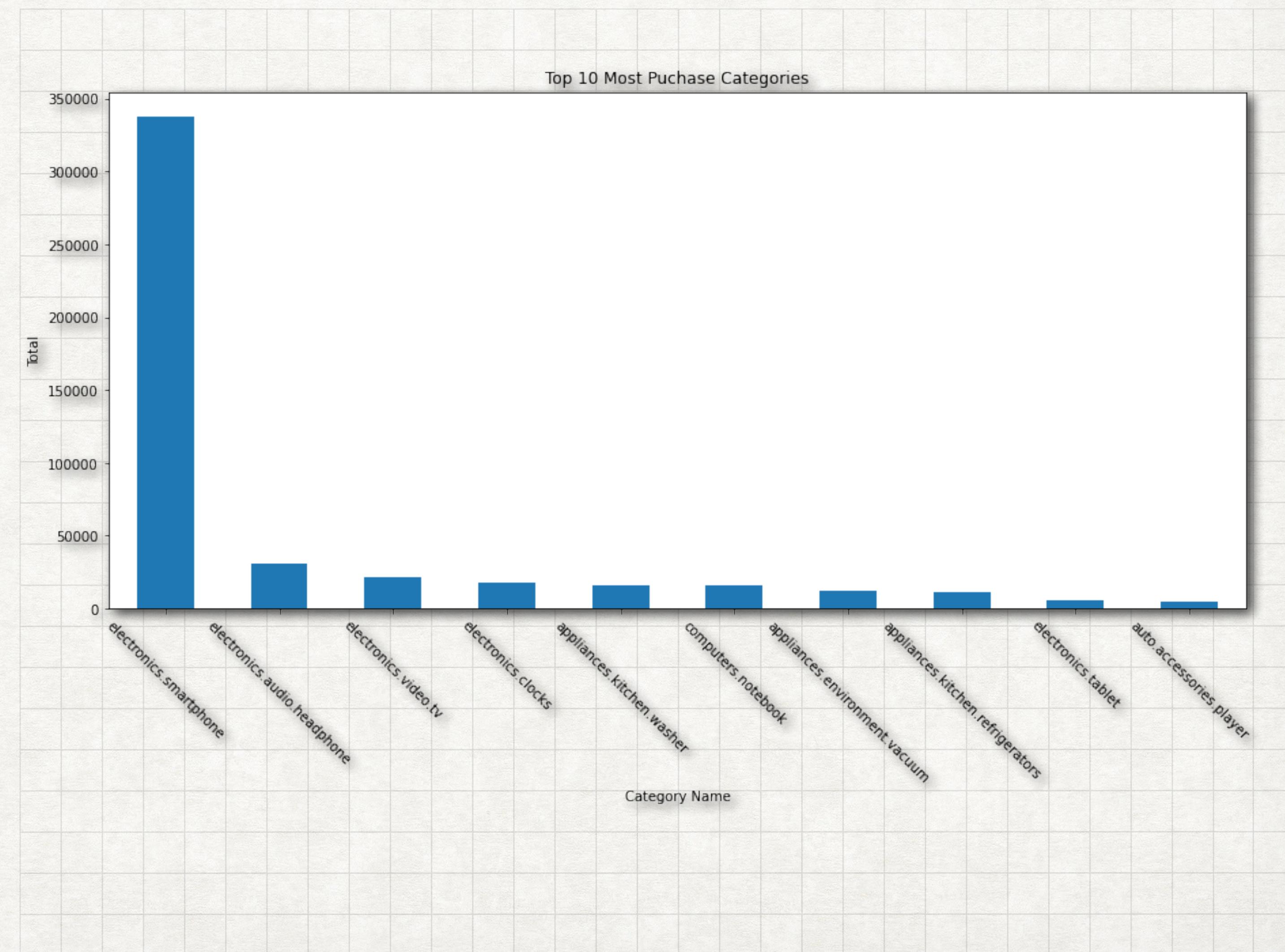
# MOST POPULAR CATEGORIES AND BRANDS

Samsung became the most popular brand in October 2019, followed by Apple and Xiaomi in second and third place. Samsung's product has been viewed almost 5,000,000 times.



# MOST PURCHASE CATEGORIES AND BRANDS

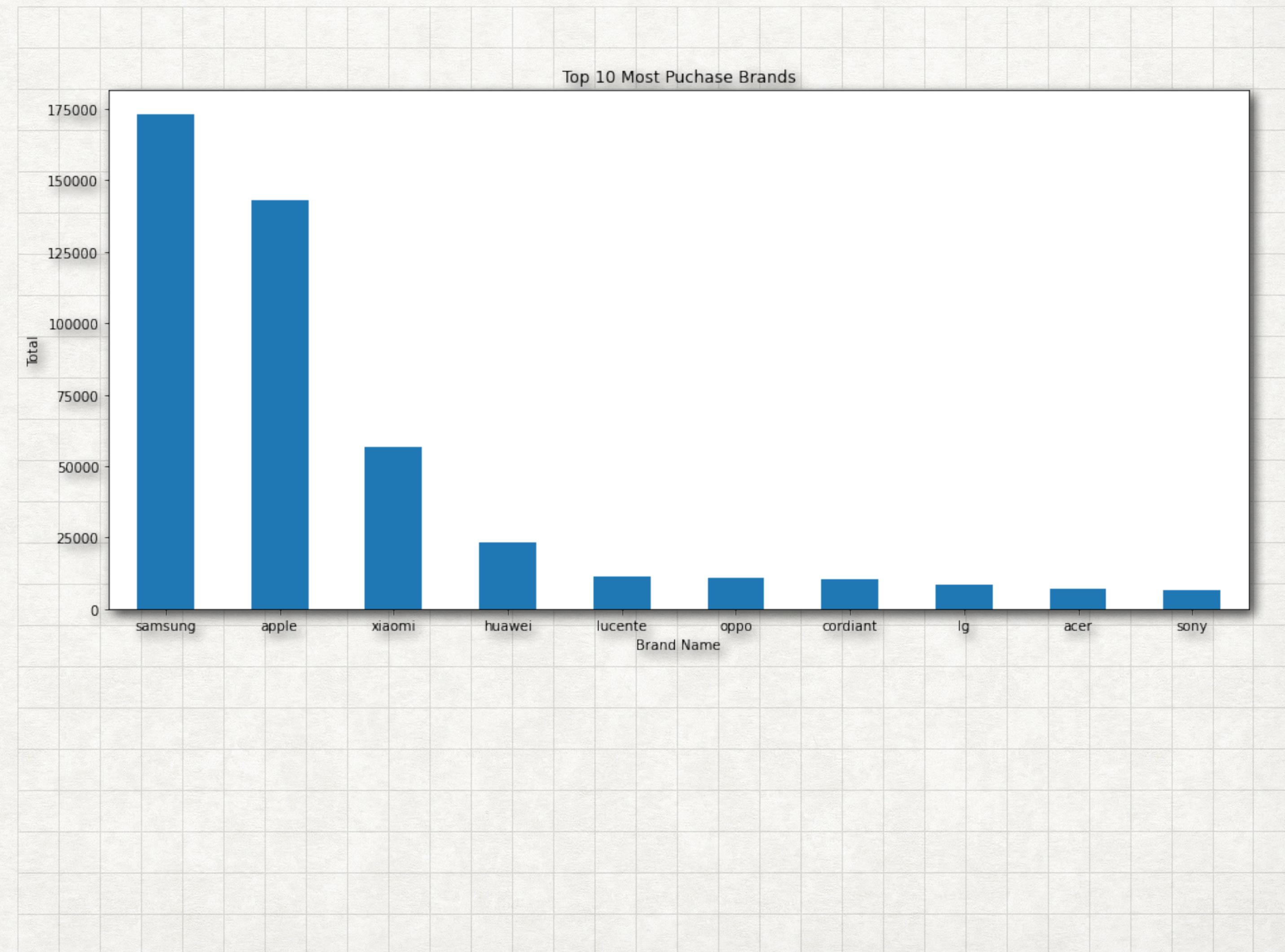
Smartphone became the most purchase category in October 2019. This category has been purchased almost 350,000 times. Headphone and TV are in the second and third place.



# MOST PURCHASE CATEGORIES AND BRANDS

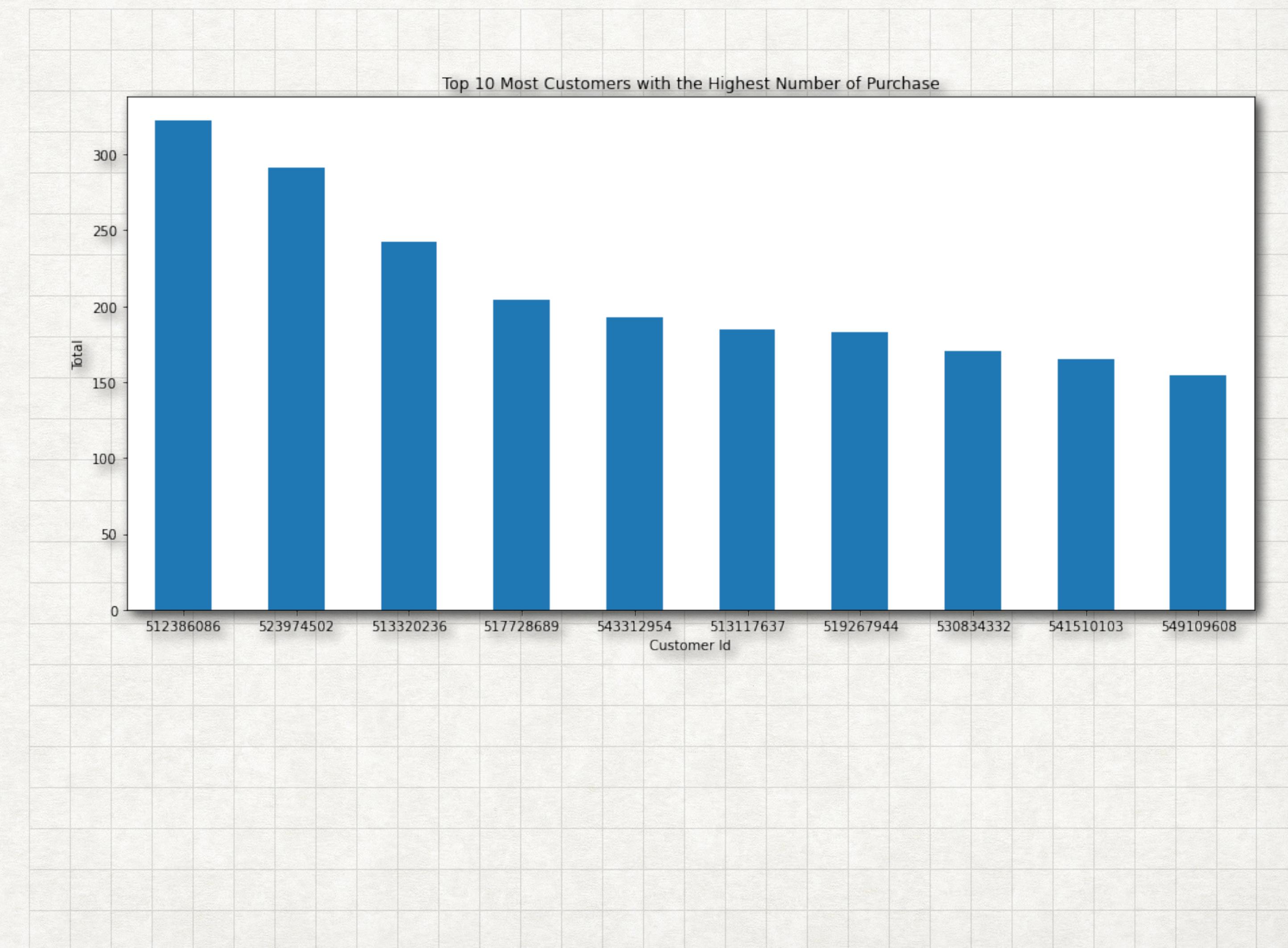
Samsung became the most purchase brand in October 2019, followed by Apple and Xiaomi in second and third place. Samsung's product has been purchased almost 175,000 times.

The position of the top 10 most purchase brands and the top 10 most popular brands are nearly identical.



# MOST CUSTOMER WITH THE HIGHEST NUMBER OF PURCHASE

The customer with User ID 512386086 became the most active customer for making the highest number of purchases. This customer has made more than 325 purchases.



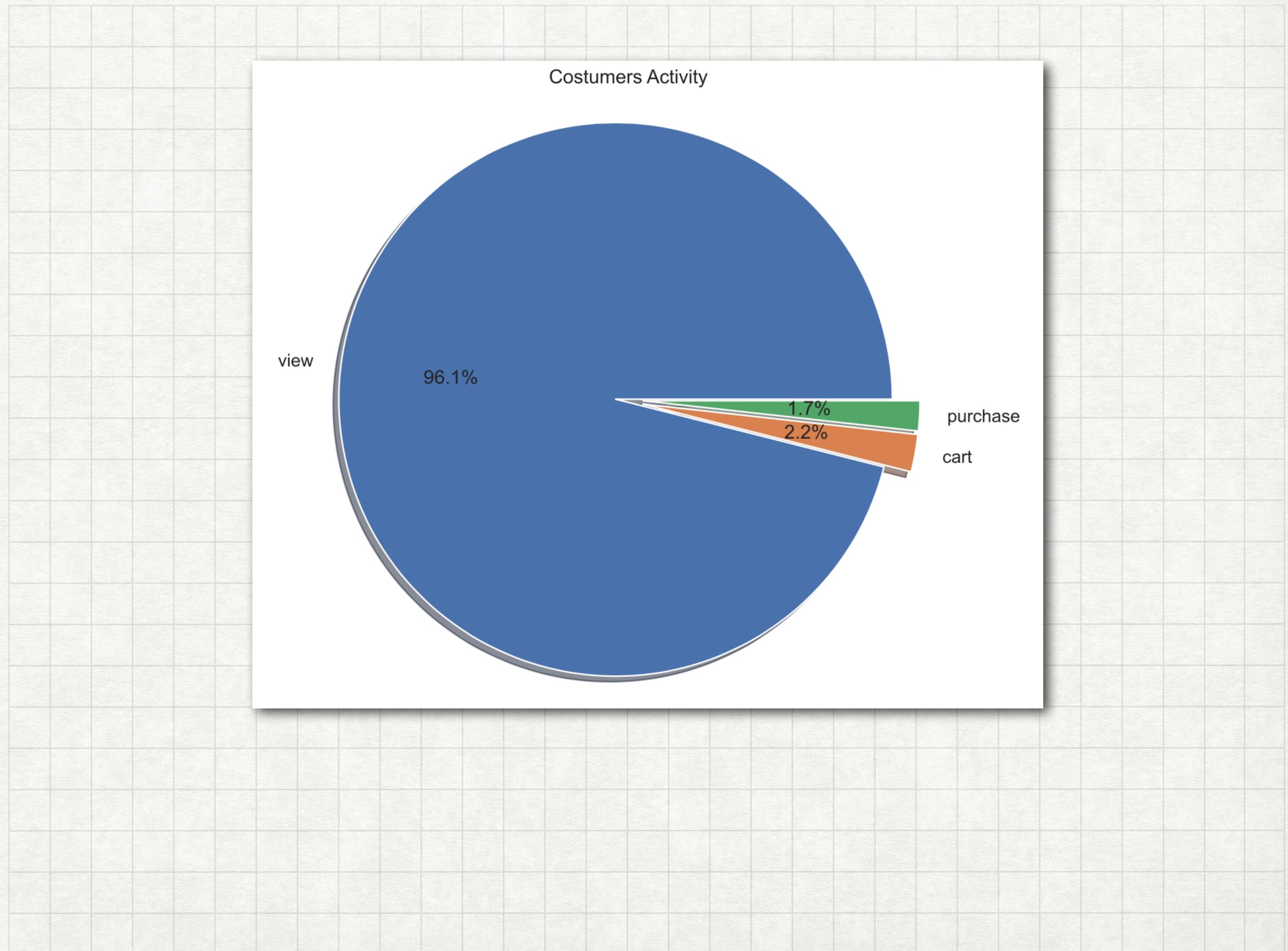
# DAILY VISITOR

In October 2019, the daily number of visitor fluctuated a lot. The visitors tends to increase from the beginning of the month until the middle of the month. However, it tends to decrease after the middle of the month.



# CONSUMERS ACTIVITY

Most of the customers activity are view the products (96.1%); only a small percentage of customers add items to their shopping cart or complete the transactions. There are 1.7% or 742,849 customers who made a transaction.



# CONCLUSION AND RECOMMENDATION

Based on the data mining activities that have been carried above:

- Provide discounts on products that aren't selling well, so its can be sold right away.
- Give a promotions at the beginning and end of the month to attract or increase website visitors.
- Because many customers just view the products, rewards can be given to customers who make frequent transactions in order to increase the number of transactions. This will encourage more customers to make transactions.