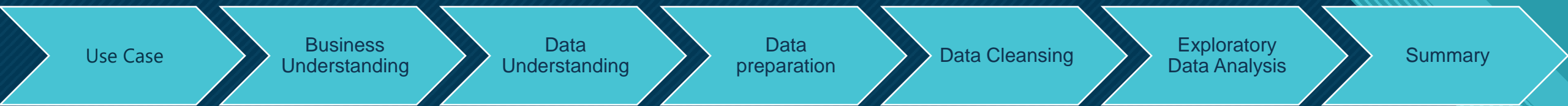




E-Commerce Behavior Data

Data Mining Project By Suwarti

Mind Map E-Commerce Behavior Data



Use Case: E-Commerce Behavior Data

Periode: October 2019

Use Case Summary

Objective Statement:

- Get business insight about what is the best product to sell in the specific time of the day.
- Get business insight about what is the best Event to predict that User most likely to buy a product.
- Get business insight about how is the behaviour of our repeat user vs new user.

Source Data:

Ecommerce Data Behaviour

Challenges:

- Large size of data, can not maintain by excel spreadsheet.
- Demography data have a lot missing values and typo.

Methodology / Analytic Technique:

- Descriptive analysis
- Graph analysis

Success Criteria:

Explore insigh the behaviour of ecommerce business.

Expected Outcome

- The best product to sell in the specific time of the day.
- The best Event to predict that User most likely to buy a product.
- The behaviour of our repeat user vs new user.

Business Understanding

- E-Commerce is a business model that lets firms and individuals buy and sell things over the internet.
- This case has some business question using the data:
 - What is the best product to sell in the specific time of the day?
 - What is the best Event to predict that User most likely to buy a product?
 - How is the behaviour of our repeat user vs new user?

Data Understanding

- Behavior data for October 2019 from a large multi-category online store. Each row in the file represents an event. All events are related to products and users. Each event is like many-to-many relation between products and users.
- **Source Data:** [Ecommerce Data Behaviour](#)
- Data Dictionary:

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.

Data preparation

Code Used:

- Python Version: 3.7.6
- Packages: Pandas, Numpy and Matplotlib.

Data Cleansing

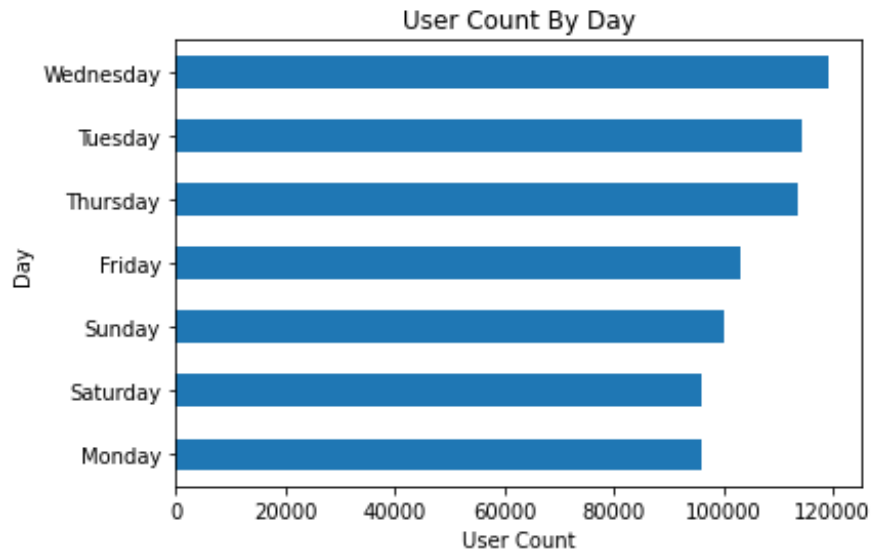
- Check for each column and find if any column is redundant or useless.
- Check for missing values.
- Check for outliers.
- Check if the data format is already suitable for algorithm.
- Check for value that are not consistent with general common sense.

Exploratory Data Analysis

Looked at the distributions of the data and the value counts for the various categorical variables.

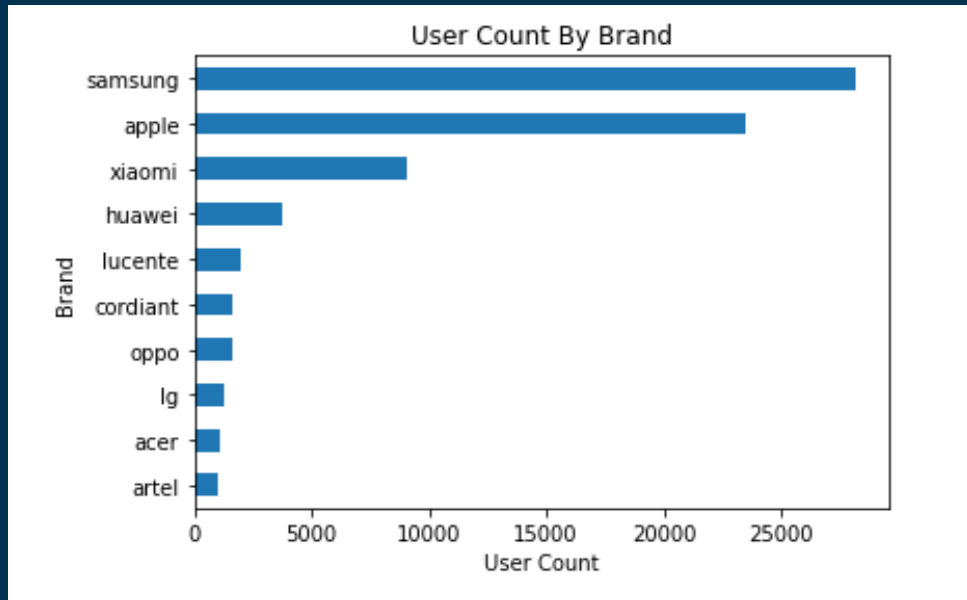
Explore the data to answer the business question that created before.

- **The best product to sell in the specific time of the day is Wednesday.**



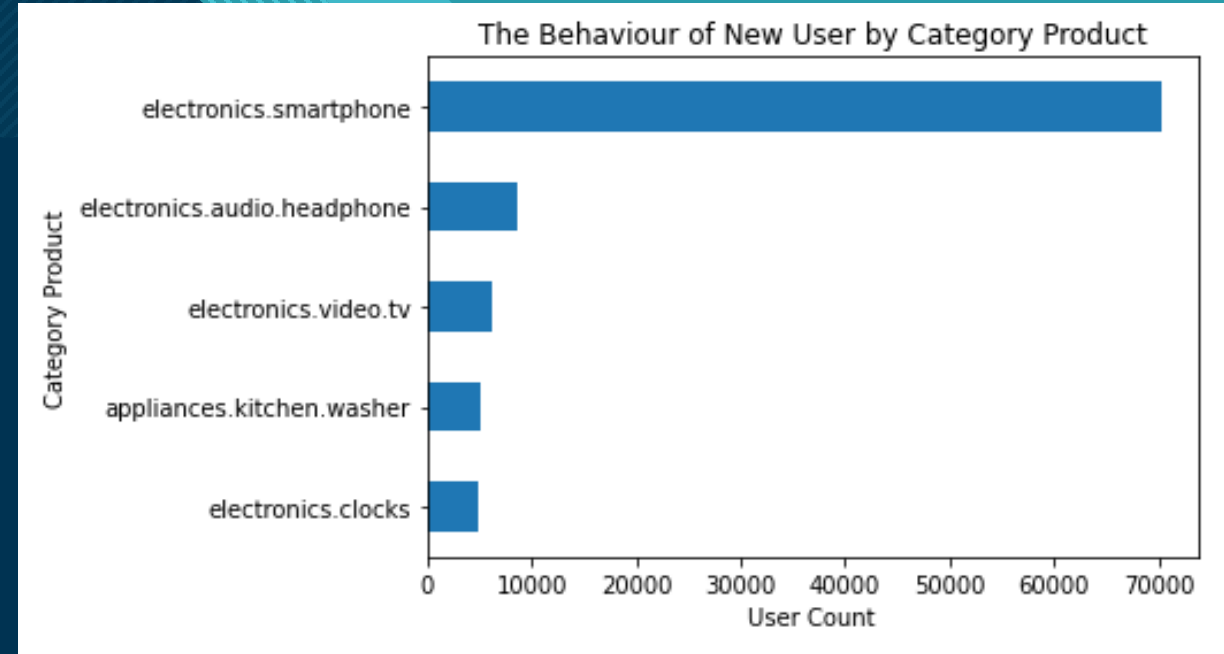
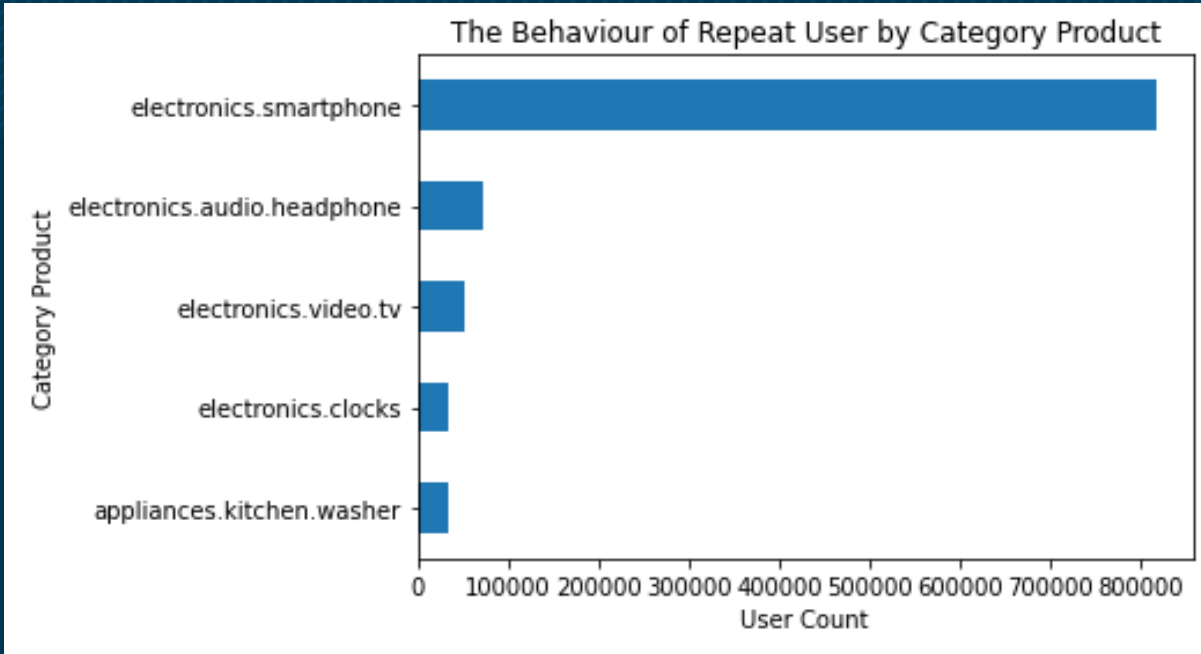
Most consumers shop at e-commerce on Wednesday. Actually, consumers have seen the product since Sunday. After Sunday consumers are still considering whether to buy or not. Beside that consumers are still comparing one product to another. The final decision to buy is mostly made on Wednesday. The e-commerce business team can provide a campaign to convince customers. The campaign can be intensified on Wednesday.

- The best Event to predict that User most likely to buy a product is purchase of brand Samsung.



Many consumers choose brand Samsung because it already has a strong and positive image for the smartphone segment. Samsung also has a simple and trend-following model. In addition, Samsung is diligent in innovating and releasing the latest technology. Seeing Samsung's very large consumer market, the e-commerce business team can make bundling package promos to expand the market.

The behaviour of our repeat user vs new user by category product



The behavior of repeat user by category product:

Most repeat users buy smartphone electronics products. Meanwhile product electronics audio headphones have a very large gap. The e-commerce business team can make product bundling between electronics smartphone products and electronics audio headphones to increase profits from product categories that are still low in orders.

The behavior of new user by category product:

Most new users buy smartphone electronics products. The e-commerce business team can make promo of this product category to attract more new market users.

Summary

- Most consumers shop at e-commerce on Wednesday. The e-commerce business team can provide a campaign to convince customers. The campaign can be intensified on Wednesday.
- Many consumers choose brand Samsung because it already has a strong and positive image for the smartphone segment. The e-commerce business team can make bundling package promos to expand the market.
- Most repeat users buy smartphone electronics products. Meanwhile product electronics audio headphones have a very large gap. The e-commerce business team can make product bundling between electronics smartphone products and electronics audio headphones to increase profits from product categories that are still low in orders.
- Most new users buy smartphone electronics products. The e-commerce business team can make promo of this product category to attract more new market users.

Thank You



suwarti.shubuh@gmail.com



Suwarti Suwarti



github.com/suwarti