

Report on BlackFriday for Professor Yin Shen

Human Computer Interface, A210, Spring semester, 2018-2019

Submitted by Zijian Zhang

Report on BlackFriday for Professor Yin Shen

Executive Summary

Data Set and Objectives

Data Set

Objectives

Which city buys more?

Who are buying this?

Software

Analysis

Which city buys more?

Who are buying this?

Attachment

Executive Summary

Black Friday is a main way of promoting sales of many gaint companies like Amazon and Ebay. The data about their sales gives us great insight into consumer clusters and maybe finding out different clusters of consumers within the buyers of a specific product.

Data Set and Objectives

Data Set

The data set given by Prof. Shen contains 550 000 observations about the black Friday in a retail store, it contains different kinds of variables either numerical or categorical. The columns of this data set is displayed as folows.

- User_ID
- Product_ID
- Gender
- Age
- Occupation
- City_Catagory
- Stay_In_Current_City_Years
- Matital_Status
- Product_Catagory_1

- Product_Catagory_2
- Product_Catagory_3
- Purchase

Objectives

According to the features of this data set, I choose 2 objectives as my main goal of this lab.

Which city buys more?

As we all know that different cities have different ability of purchase. So I am trying to compare the purchase value of the same occupations in different cities.

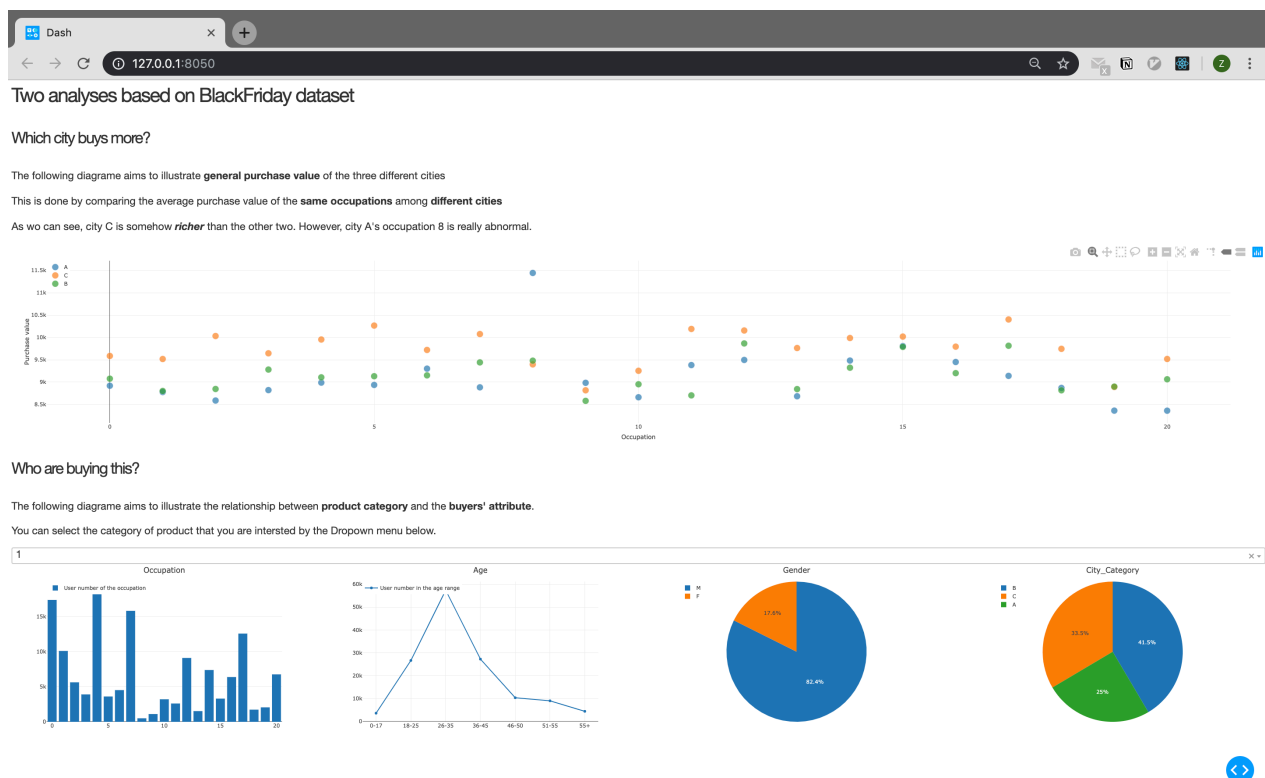
Who are buying this?

It is generally accepted that the buyer of different products are generally different. Thus I attempt to compare the statistiacs of differenent buyers to try to find some patterns.

Software

The analysis is accomplished with the help of **Pandas** and **Dash**.

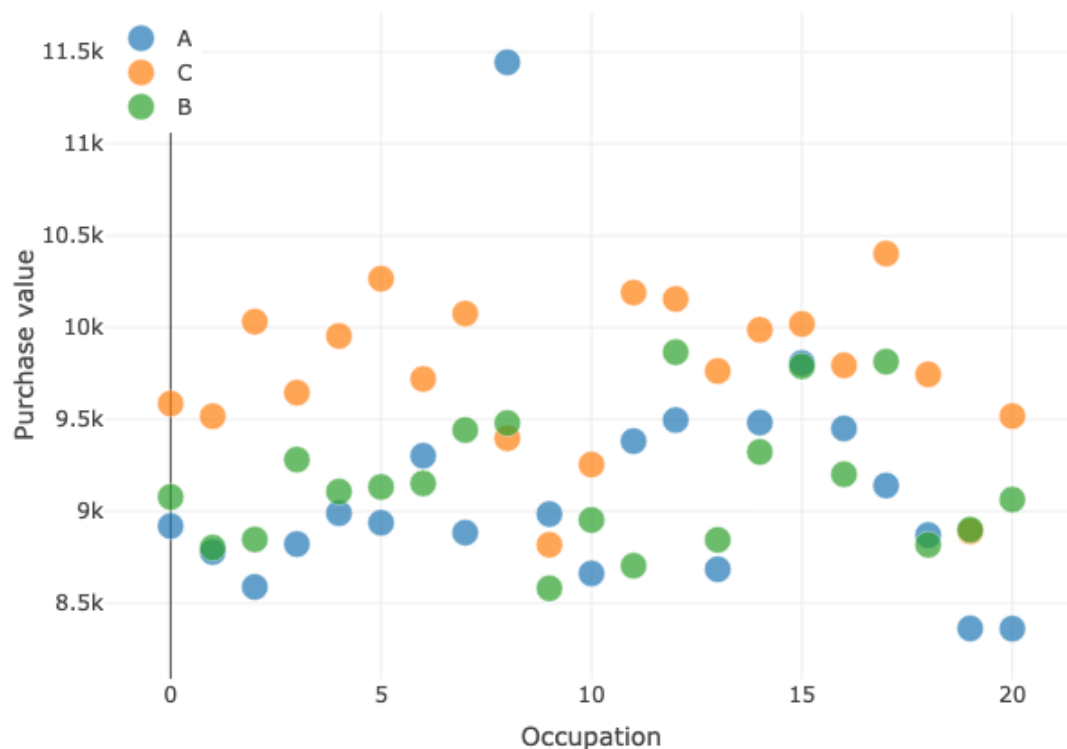
The dashboard is designed to be composed of two sections about the two goals respectively. The first section contains a scatter plot and the second part contains two bar charts as well as two pie charts. The overall strucure of the web page is illustrated below. In addition, the second part is interactive, which allows for clear demostration of various type of product category.



Analysis

Which city buys more?

I compare the purchase value of people from the same occupation but from different cities. The result diagram is shown blow.



As we can see, **city C's pruchase value is the highest, while city A's and B's purchase value is not very different.** However the purchase value of people in city A with occupation 8 is strikingly high. I examined the values of this category, and I think this is true in the given data set.

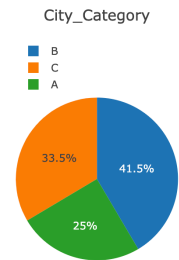
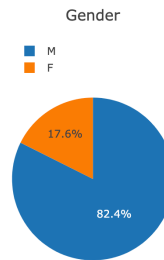
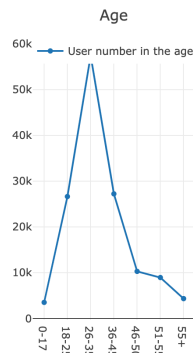
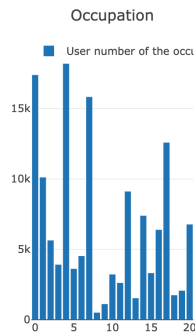
Who are buying this?

I examined the occupation, age, gender and city category of the buyer of a certain type.

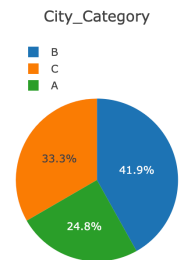
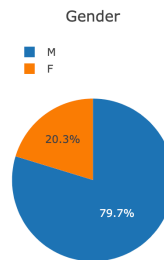
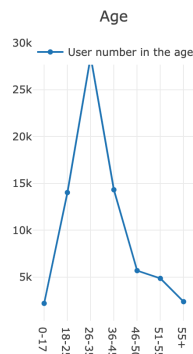
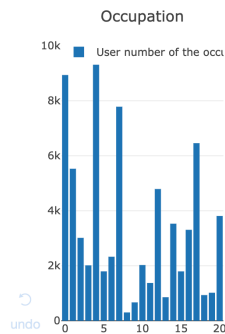
There are 18 types of products appear in product category. In case of inefficient space in my report, only product 1, 2 and 3 are shown.

Please note that the value below is about the total number of **User ID**, not about value directly.

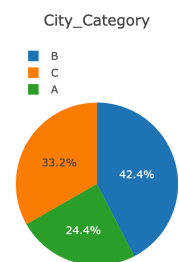
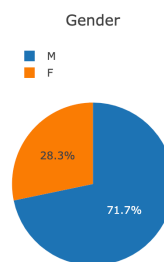
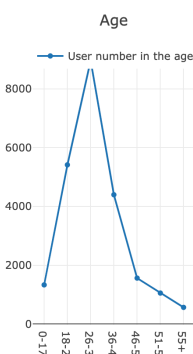
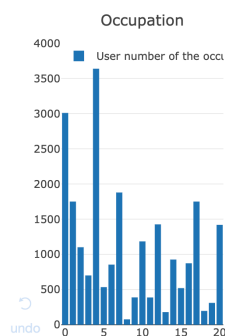
Product 1



Product 2



Product 3



Disappointingly, the result shows that similarities are everywhere.

- People of occupation 4 and occupation 0 buy more.
- Middle age people buy more.
- Male buy more.
- City **B** buys more, but surprisingly with less total value. I think that people in city **C** are buying more **types** of products, more scattered in other words.

Attachment

The source file is provided.