Lab4-Report

1. Introduction

This is 2052844 Cao Xiaoci's user interaction technical job 4 Data processing.

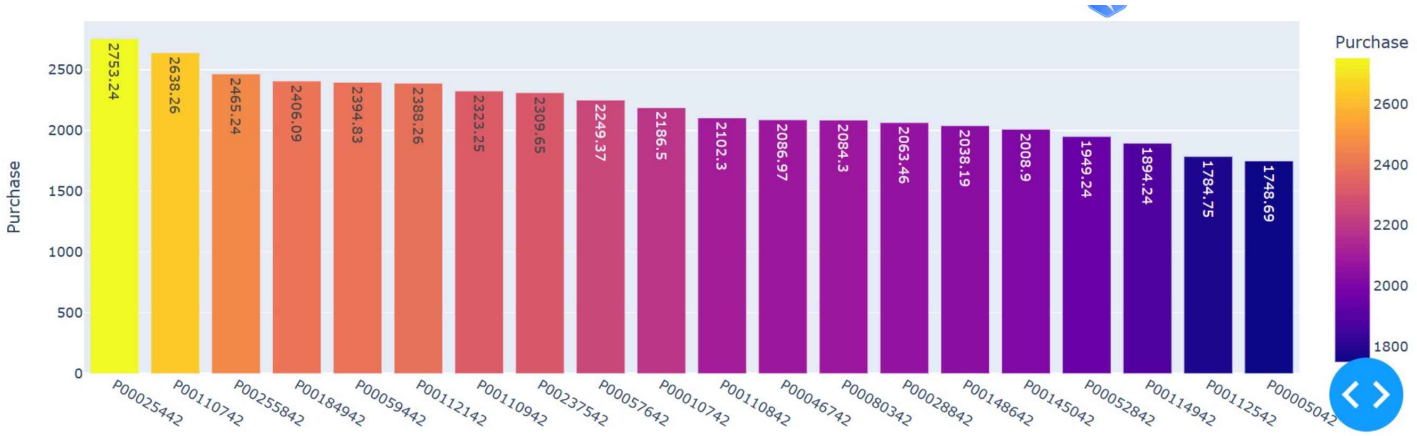
The data used by the job is the Black Friday dataset.

Displayed using **python dash**, plotly.express drawing.

In this section, you can see the impact of gender, age, occupation, and length of residence on sales and sales, which is displayed on a pie chart.

You can also view total sales pie charts, sales box charts, and column charts for the ratio of male and female sales in different product categories.

左图展示了不同因素对消费额与消费数量的对比图（以性别为例），右图为不同商品类别销售金额与男女间比例图



该部分展示的是销售额最高的20件商品销售额对比图

2. Data analysis task

The background of the dataset relies on Black Friday, which is a major procurement in Western countries, similar to China's "Double Eleven" event. On Black Friday, the mall will launch a large number of discounts and promotions as the last large-scale promotion before Christmas, aiming to maximize profits by the end of the year.

The Black Friday dataset is a dataset containing approximately 540000 black Friday transaction samples from retail stores, including 12 fields. Among them, product category_ 1 cannot be empty, product category_ 2. Product categories_ 3 can be empty. There are no missing values in the dataset.

Black Friday is the second day after Thanksgiving in the United States and also the first day of procurement for the National People's Congress. The mall will launch a large number of discounts and promotional activities. This dataset contains demographic information of users (age, gender, marital status, city category, and residence time), detailed information of products (product ID and product category), and total purchase amount. Analyzing these datasets can help businesses better understand user purchasing behavior and provide relevant intelligence and reference for next Black Friday.

The analysis objects of the Black Friday dataset are mainly users and products, and the main characteristics are as follows:

In terms of products, the sales revenue and popularity of products can be analyzed from dimensions such as product ID and product category.

For example, product category 1 has the highest sales volume, and product ID P00265242 ranks second in sales,.

In terms of users, their purchasing behavior and preferences can be analyzed from dimensions such as gender, age, occupation, marital status, city category, and length of residence.

For example, the consumption quantity and amount of male users are much higher than that of female users, with a ratio of approximately 3:1. In addition, the age group of 26 to 35 is the main consumer, and users with a career number of 4 shop the most.

3. Layout

The overall layout of the page is relatively simple, mainly composed of three images, located in the upper left, upper right and just below the three images. The relevant code is as follows:

```
app.layout = html.Div(
    children=[
        html.H1(
            children='BlackFriday Dash',
            style={
                'textAlign': 'center'
```

```
            }
        ),
        html.Div([
            html.Div([
                dcc.Markdown(
                    }
                ),
                dcc.Dropdown(
                    id="dropdown",
                    options=[
                        {"label": "Gender", "value": "Gender"},
                        {"label": "Age", "value": "Age"},
                        {"label": "Occupation", "value": "Occupation"},
                        {"label": "Stay_In_Current_City_Years", "value": "Stay_I
                    ],
                    value="Gender",
                    clearable=False
                ),
                dcc.Graph(
                    id='example-graph-2',
                    figure=fig
                ),
            ], className="six columns"),  # 使用six columns类来指定宽度为一半
            html.Div([
                dcc.Markdown(
                    }
                ),
                dcc.Dropdown(
                ),
                dcc.Graph(
                    id='myfig-2',
                    figure=fig5  # 饼图
                ),
            ], className="six columns"),  # 使用six columns类来指定宽度为一半
        ], className="row"),  # 使用row类来指定布局为水平排
        html.Div([
            dcc.Markdown(
            ),

            dcc.Graph(
                id='fig-6',
                figure=fig6  # top20
            ),
        ], className="row"),

    ], className="row twelve columns")
```