Black Friday Sales Data Analysis

-**Prerequisites for this project:**Pandas, Numpy, Seaborn.

Black Friday is a shopping holiday  that takes place on the day after Thanksgiving. It is known for its deep discounts and special deals on a wide range of products, including electronics, home goods, clothing, and more. Many retailers offer special doorbuster deals and extended hours on Black Friday, and it is traditionally one of the busiest shopping days of the year.

The dataset that we are having contains 537578 rows and 12 columns.

Our main goal will be to do various kind of analysis and get inferences which will be able to help to provide valuable insights on the market and hence will be beneficial for the company.

We are going to divide the project into 7 parts:

1. Dataset Walkthrough
2. Analyzing Columns
3. Analyzing Gender
4. Analysing Age & Marital Status
5. Multi Column Analysis
6. Occupation and Products Analysis
7. Combining Gender & Marital Status

**Dataset Walkthrough:**

Here we are basically trying to check out the basic things regarding our dataset such as the column names, the data type of the columns, if there are any null values present in our data, etc. After some analysis we were able to detect a lot of null values in the column Product\_Category\_2 and Product\_Category\_3. So now if we try to use the function df.dopna() it would remove all the rows and might create a data loss problem later. Therefore we decide to delete the whole column instead and we get rid of all the null values present in our dataset.

**Analyzing Columns:**

In this section we have mainly used the unique() and nunique() functions to know the number of unique values present in every individual columns. For example when we have used the nunique() function with the columns User\_ID and Product\_ID we are able to get the actual total number of customers and total number of products present in our dataset. Similarly we have used the unique() function with the columns Gender and Age to get the different unique values present in those columns. Hence analyzing these individual columns like this tells us about the different kind of unique values present in our dataset.

**Analyzing Gender**

In this section we are going to specifically analyze the Gender Column. So in the gender column we basically have Male and Female values. After doing some basic analysis we have noticed that the Male data is much more than the Female data in our dataset. We have even used the groupby function to see which Gender is purchasing more in our dataset. With the help of pie chart and bar plot we have presented them in a very nice manner.

**Analysing Age & Marital Status**

In this section we are going to analyze the age and Martial Status columns. So we can again use the groupby function to see which age group is making the maximum number of order in our dataset. We can also see which age group has the highest purchasing amount in our dataset and can plot all the inferences with the help of pie chart and bar plots. After doing some basic analysis on Marital Status we can see that 60% of the people in our dataset is unmarried and 40% of them are married. We have even displayed them in pie chart forms for better understanding.

**Multi Column Analysis**

So all this time we were just talking about single column analysis but now let's get started with multi column analysis. We will use the Seaborn library to visualise them and present them in a more beautiful manner. We have done analysis with the help of Age and Gender columns and presented them with the help of the Seaborn library. We have also added legends which is basically the hue parameter within the seaborn function. So we can basically do analysis for any number of columns based on our choice with the help of the seaborn library.

**Occupation and Products Analysis**

We are going to perform the same data visualizations which we did till now for the Occupation ,Product\_ID and Product\_Category\_1 columns. Using the countplot we can get various insights about which kind of data is present the most in our dataset. We have also plotted normal bar plots with the help of groupby function to get information such as which Product\_ID has got the maximum purchasing amount in our dataset.

**Combining Gender & Marital Status**

So this the last subtopic in this project and here we have combined the Gender and Martial Status. Hence using seaborn we are able to perform various kinds of data visualizations and get meaningful insights of our data when Gender and Martial Status is combined with the help of mainly the countplot.

**#Link for Dataset**

<https://raw.githubusercontent.com/AshishJangra27/Data-Analysis-with-Python-GFG/main/26.%20Black%20Friday%20-%20Combining%20Age%20%26%20Marital%20Status/BlackFriday.csv>

**#Link for Code  
Reference**

1. Dataset Walkthrough :

<https://github.com/AshishJangra27/Data-Analysis-with-Python-GFG/blob/main/20.%20Black%20Friday%20-%20Walkthrough/Black%20Friday%20-%20Walkthrough.ipynb>

2. Analyzing Columns :

<https://github.com/AshishJangra27/Data-Analysis-with-Python-GFG/blob/main/21.%20Black%20Friday%20-%20Analysing%20Columns/Black%20Friday%20-%20Analysing%20Columns.ipynb>

3. Analyzing Gender :

https://github.com/AshishJangra27/Data-Analysis-with-Python-GFG/blob/main/22.%20Black%20Friday%20-%20Analysing%20Gender/Black%20Friday%20-%20Analysing%20Gender.ipynb

4. Analysing Age & Marital Status :

<https://github.com/AshishJangra27/Data-Analysis-with-Python-GFG/blob/main/23.%20Black%20Friday%20-%20Analysing%20Age%20%26%20Marital%20Status/Black%20Friday%20-%20Analysing%20Age%20%26%20Marital%20Status.ipynb>

5. Multi Column Analysis :

https://github.com/AshishJangra27/Data-Analysis-with-Python-GFG/blob/main/24.%20Black%20Friday%20-%20MultiColumn%20Analysis/Black%20Friday%20-%20MultiColumn%20Analysis.ipynb

6. Occupation and Products Analysis :

<https://github.com/AshishJangra27/Data-Analysis-with-Python-GFG/blob/main/25.%20Black%20Friday%20-%20Occupation%20and%20Products%20Analysis/Black%20Friday%20-%20Occupation%20and%20Products%20Analysis.ipynb>

7. Combining Gender & Marital Status :

https://github.com/AshishJangra27/Data-Analysis-with-Python-GFG/blob/main/26.%20Black%20Friday%20-%20Combining%20Age%20%26%20Marital%20Status/Black%20Friday%20-%20Combining%20Age%20%26%20Marital%20Status.ipynb

Mark as Read

Report An Issue