Exploratory Data Analysis

of

Black Friday Sales

Black Friday Sales

Black Friday originated in the 1960s in the United States as a term to describe heavy traffic and chaos the day after Thanksgiving. By the 1980s, retailers began using it to promote special sales and discounts. Over time, it evolved into a major shopping event, with stores opening early and offering significant discounts to kick off the holiday shopping season. The rise of online shopping further amplified Black Friday, with both brick-and-mortar and e-commerce retailers participating in the sales frenzy. Despite controversies and occasional chaos, Black Friday remains a crucial part of the holiday shopping tradition.

Table of Contents

Introduction

- Exploratory Data Analysis (EDA) of Black Friday Sales unveils a comprehensive insight into consumer behavior and retail trends during this annual shopping extravaganza. As a pivotal event in the retail calendar, Black Friday attracts millions of shoppers globally, shaping the landscape of both traditional and online commerce.

- This analysis delves into the intricate patterns of purchasing preferences, examining the impact of discounts, product categories, and demographic factors on sales. By scrutinizing vast datasets, we aim to unravel hidden correlations, identify popular products, and discern regional variations in spending habits.

- EDA allows us to uncover unique market dynamics, shedding light on the evolving nature of consumerism and providing valuable perspectives for businesses seeking to optimize their strategies during this high-stakes retail phenomenon.

Exploratory Data Analysis

❖ **What is the distribution of purchases in the dataset?**

Distribution of Purchases
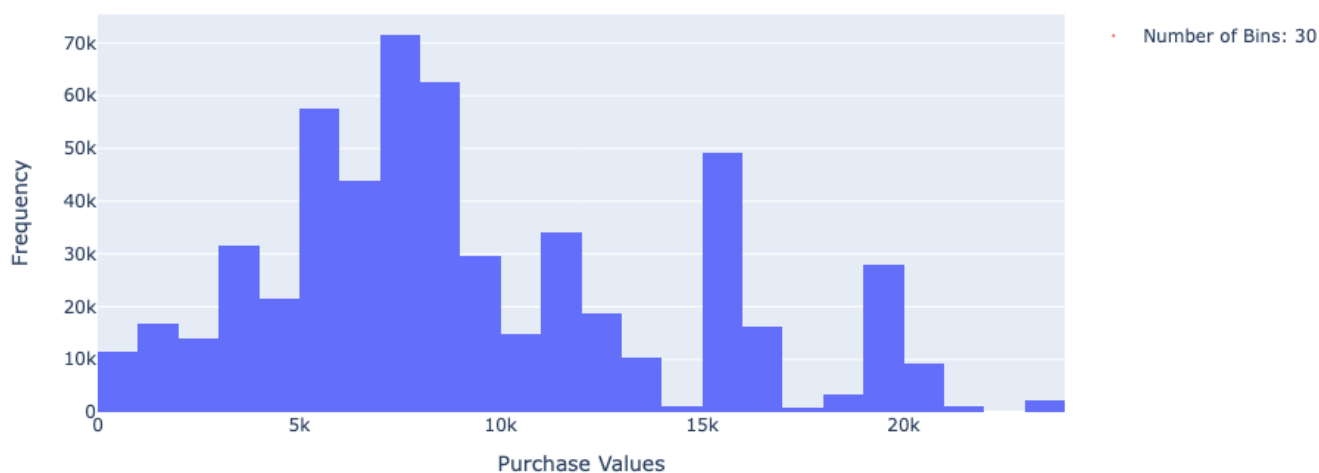
- **Highest Count Range (7000 to 8000):**

The histogram exhibits a prominent peak in the purchase values, with the highest count reaching 7100 occurrences.

The concentrated distribution in the range of 7000 to 8000 suggests a popular price range for purchases.
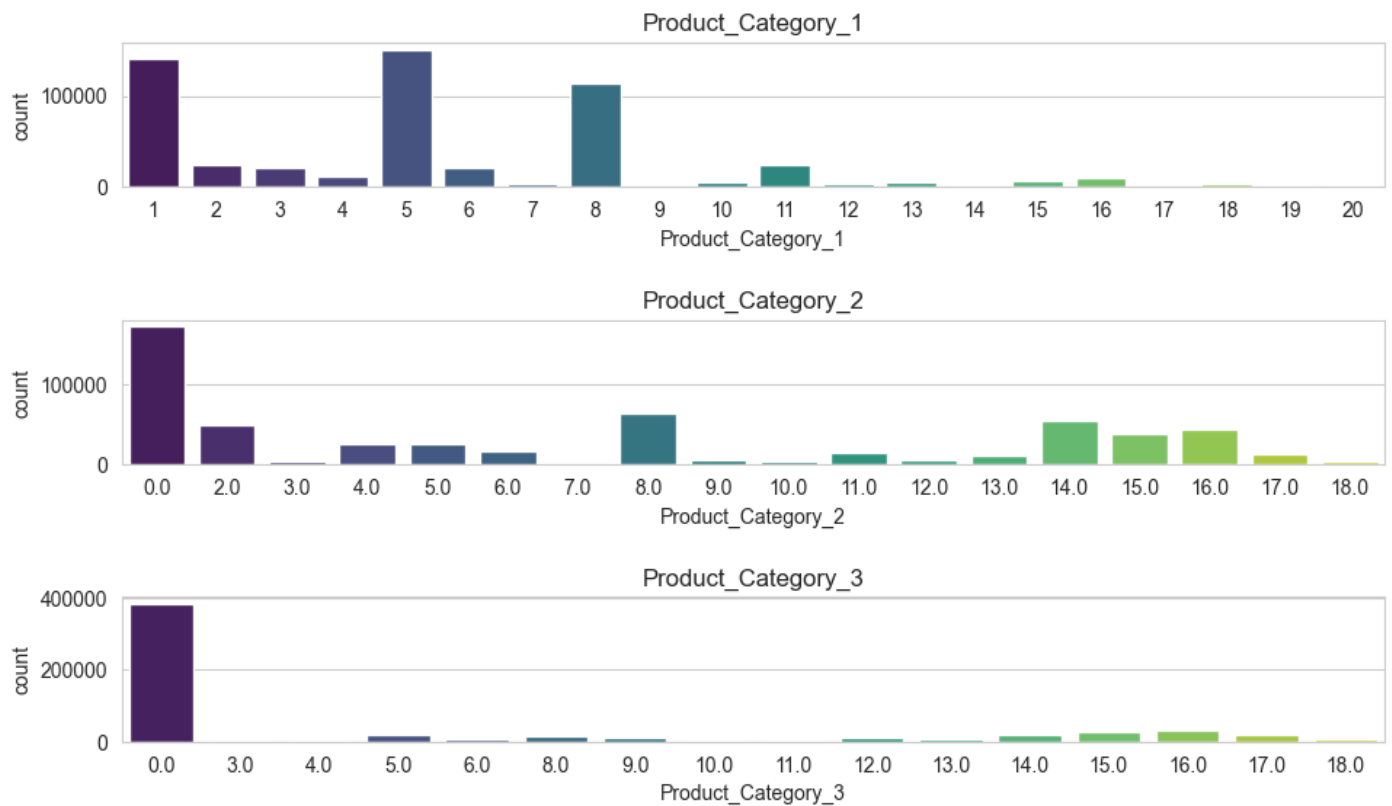
- **Lowest Count Range (2300 to 2400):**

There is a distinct dip in the histogram, indicating a lower count of occurrences.

The purchase values in the range of 2300 to 2400 exhibit the lowest count, with only 35 occurrences.

❖ **How is the purchase amount distributed across different product categories?**

- Analyzing product category preferences during Black Friday indicates a notable attraction towards specific categories: 1, 5, and 8 in Category_1, 0 and 8 in Category_2, and 0 in Category_3.

- This observation suggests that these particular categories hold a heightened appeal among Black Friday shoppers. The popularity of these categories signifies potential trends or preferences guiding consumer choices during the sales event. Retailers can leverage this insight to tailor their product offerings and promotional strategies, ensuring alignment with the preferences and interests of Black Friday shoppers and optimizing overall sales performance.

- Understanding the appeal of these specific product categories is instrumental for businesses aiming to enhance their competitiveness during the annual shopping event.

❖ **what is the distribution of purchases among different age groups and genders?**

Purchase Count by Age and Gender
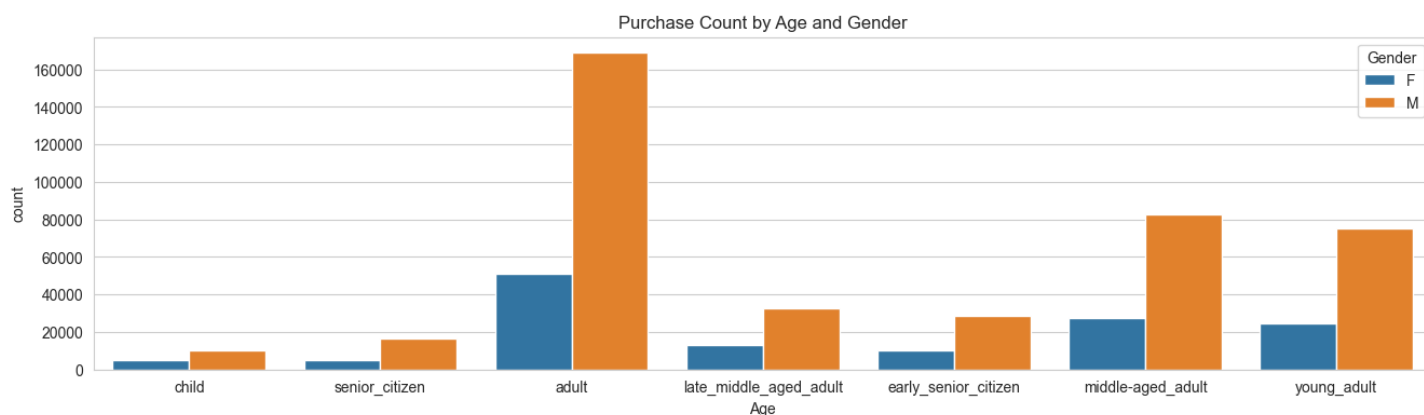
- The examination of age distribution during Black Friday reveals a pronounced prevalence of adults, middle-aged adults, and young adults, encompassing both males and females. This signifies that individuals falling within these age categories are the primary contributors to the Black Friday shopping phenomenon.
- The dominance of these age groups implies a collective inclination towards participation in the event, potentially driven by specific preferences or purchasing behaviors associated with these demographics.

- Understanding the prominence of adults, middle-aged adults, and young adults is pivotal for retailers, enabling them to tailor their Black Friday marketing strategies to better resonate with the preferences and shopping habits of these key age segments.

❖ **What is the distribution of purchases among different genders?**

Purchase Count by Gender
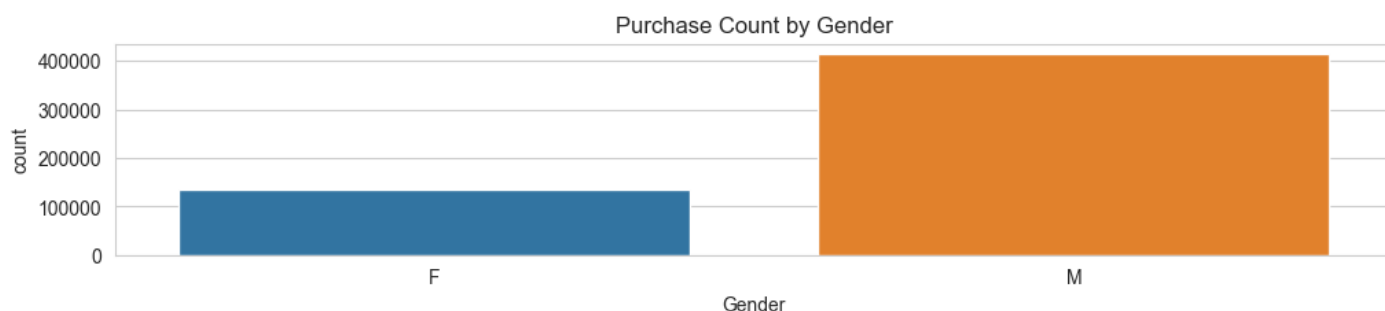
- The visual analysis of gender distribution in Black Friday sales reveals a notable imbalance, with a higher representation of males compared to females. This observation suggests that the event tends to attract a predominantly male audience.

- The higher ratio of male participants underscores a potential trend in consumer behavior, influencing retailers to tailor their marketing strategies and product offerings to cater to the preferences and interests of this significant demographic. Understanding and leveraging this gender-based trend can be crucial for businesses seeking to optimize their Black Friday sales strategies and enhance overall customer satisfaction.

❖ **How do occupation relate to the purchase amount?**


Purchase Amount by Occupation
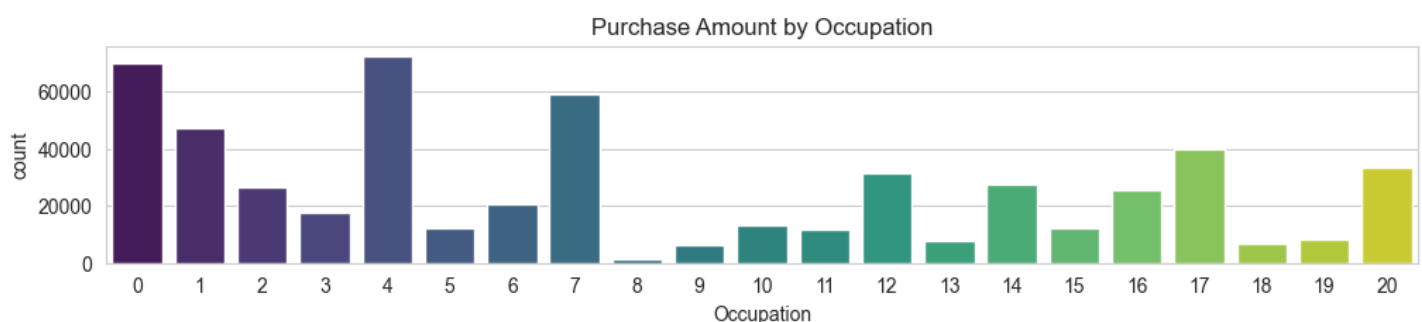
- The data analysis highlights a notable concentration of customers with occupations 4, 0, and 7 during Black Friday sales. This observation suggests a likelihood that individuals holding these specific occupations are more inclined to participate in the event.

- The prevalence of customers within these occupational categories could be indicative of certain job-related factors or preferences that drive increased engagement with Black Friday, offering valuable insights for retailers aiming to tailor their strategies to the preferences of these specific occupational groups.

- Understanding and acknowledging this concentration can be instrumental for businesses in optimizing their Black Friday sales approaches to effectively reach and engage this targeted demographic.

❖ How do marital status relate to the purchase amount?

Purchase Amount by Marital_Status
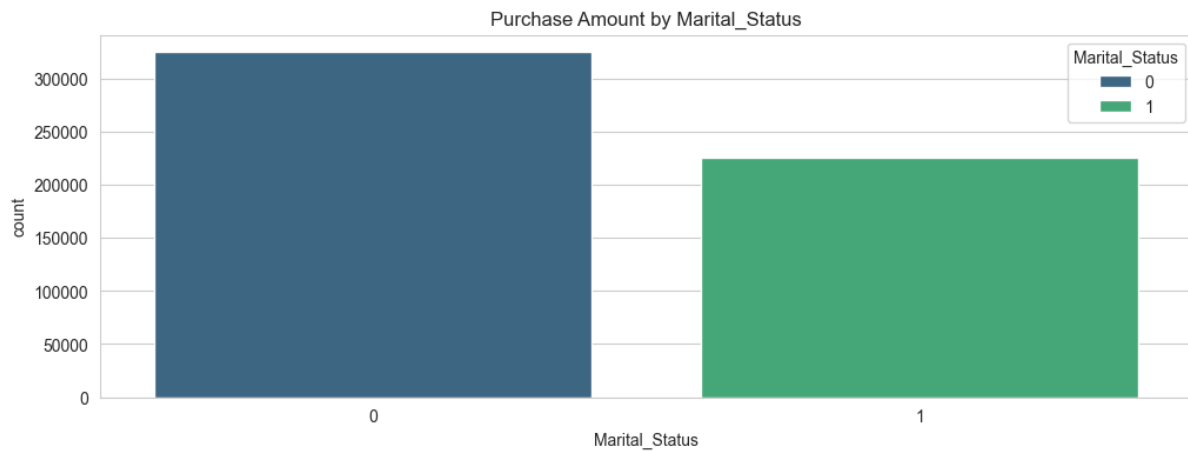
- Examining the Marital Status distribution during Black Friday indicates a prevalent majority of married customers, considering 0 as yes and 1 as no. This observation suggests that the demographic makeup of Black Friday shoppers leans significantly towards married individuals.

- The higher representation of married customers may reflect specific consumer behaviors or household dynamics that influence their participation in the sales event. Recognizing this marital status predominance is essential for retailers, guiding them to tailor their Black Friday strategies to the preferences and needs of the married consumer segment. Understanding and catering to this demographic can enhance overall customer satisfaction and engagement.

❖ How do city category relate to the purchase amount?

Purchase Amount by City_Category

- The analysis of City Category distribution during Black Friday reveals a predominant presence of customers from City Category B. This observation indicates that City Category B holds a substantial representation within the Black Friday sales demographic.

- The dominance of this city category suggests potential regional preferences or economic factors that contribute to the higher participation of residents from City Category B in the annual shopping event. Recognizing and understanding this dominance is crucial for retailers, enabling them to tailor marketing strategies and product offerings to align with the specific preferences and characteristics of customers residing in City Category B.

❖ What is the overall distribution of gender, age, and marital status in the dataset?

Gender Distribution | Age Distribution | Marital Status Distribution
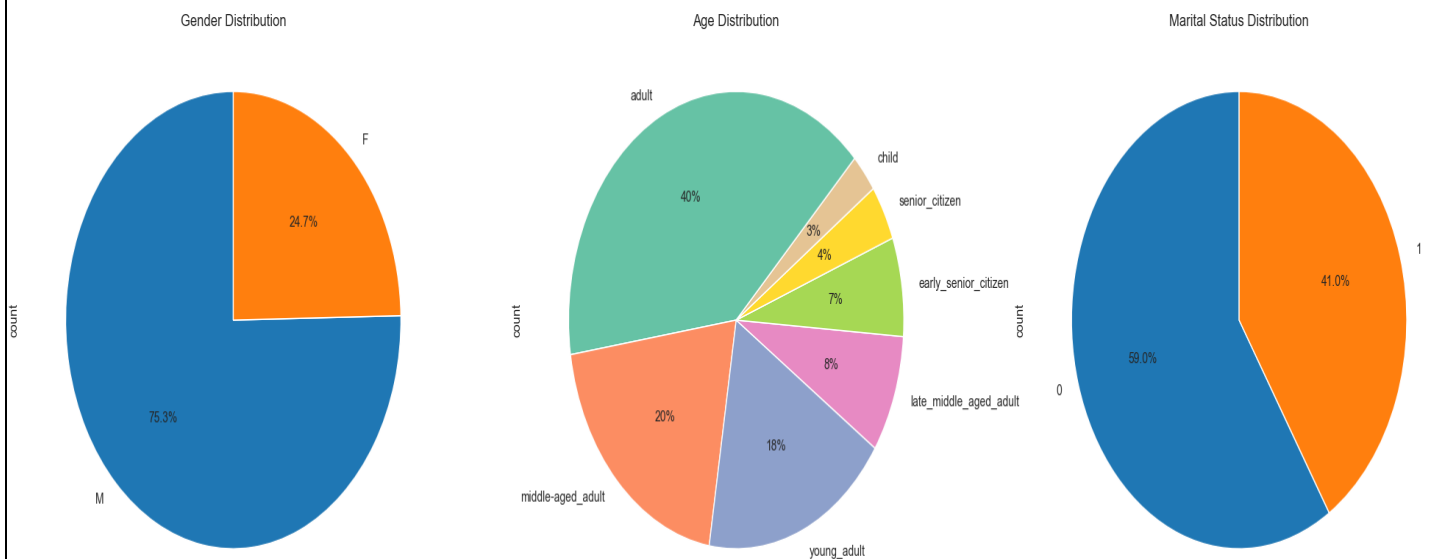
- **Gender Distribution**

The gender distribution analysis reveals a notable disparity, with males constituting the majority at 75.3%, while females account for 24.7%. This observation underscores a significant gender imbalance, indicating a higher participation of males in Black Friday sales. The substantial majority of male participants suggests a prevailing trend in consumer demographics during this shopping event.

- **Marital Straus Distribution**

The marital status distribution analysis, considering 0 as "married" and 1 as "not married," reveals a notable pattern with 59% of participants categorized as married and 41% as not married. This observation underscores a prevalent representation of married individuals among Black Friday shoppers. The majority of participants being married may reflect specific consumer behaviors, household dynamics, or economic factors influencing their engagement with Black Friday sales.

- **Age Distribution**

The age distribution analysis for Black Friday participation reveals a diverse demographic landscape. While children constitute a minor 3%, and senior citizens contribute a modest 4%, there is a more substantial engagement from early senior citizens (7%) and late middle-aged adults (8%). Notably, young adults (18%) and middle-aged adults (20%) make up a significant portion of participants, emphasizing the event's broad appeal across these age groups. Adults, as a collective category, dominate the landscape with 40%, underscoring Black Friday's primary resonance with the adult demographic. This comprehensive breakdown provides retailers with valuable insights for targeted marketing strategies, recognizing the varying degrees of interest and participation among different age segments during the Black Friday sales event.

❖ Correlation of all numeric variables

Correlation Between Numeric Columns
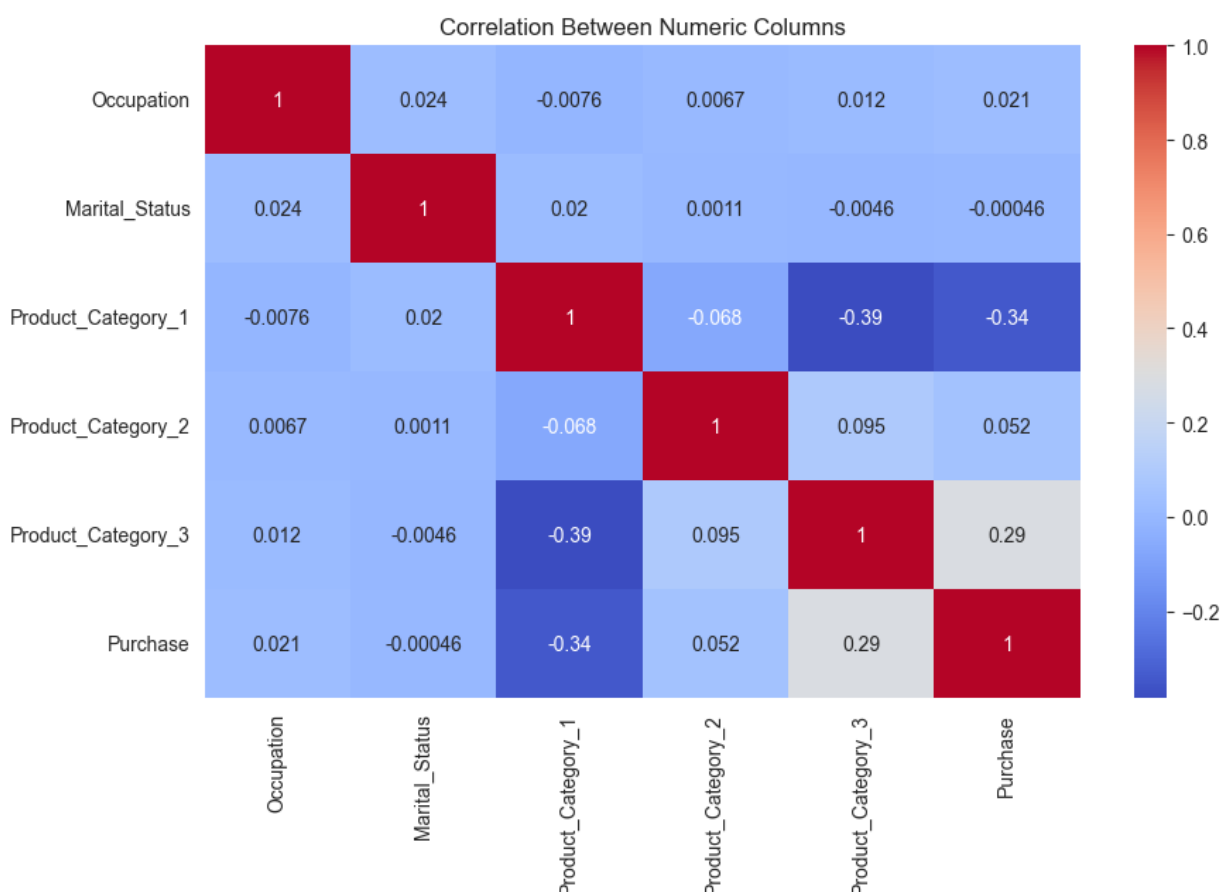
- Occupation and Marital Status: A weak positive correlation (0.024) suggests a minor association between occupation and marital status.

- Occupation and Product Categories: Weak correlations (close to zero) indicate limited influence of occupation on product category preferences.

- Occupation and Purchase: A weak positive correlation (0.021) suggests a minor impact of occupation on overall purchase amounts.

- Marital Status and Product Categories: Weak correlations (around zero) imply that marital status has minimal influence on product category preferences.

- Marital Status and Purchase: An extremely weak correlation (-0.00046) suggests that marital status does not significantly impact overall purchase amounts.