# User Churn Prediction for E-commerce Platform

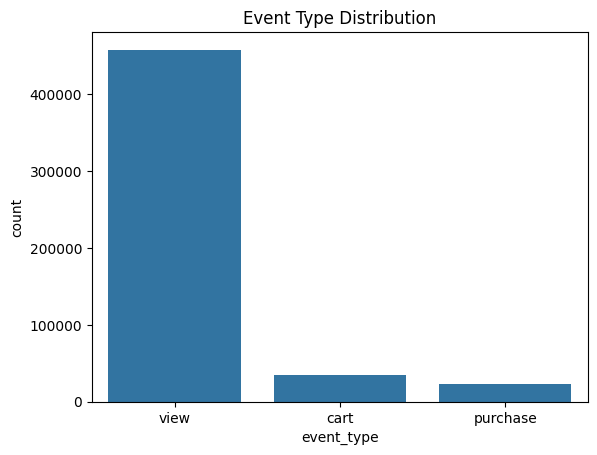
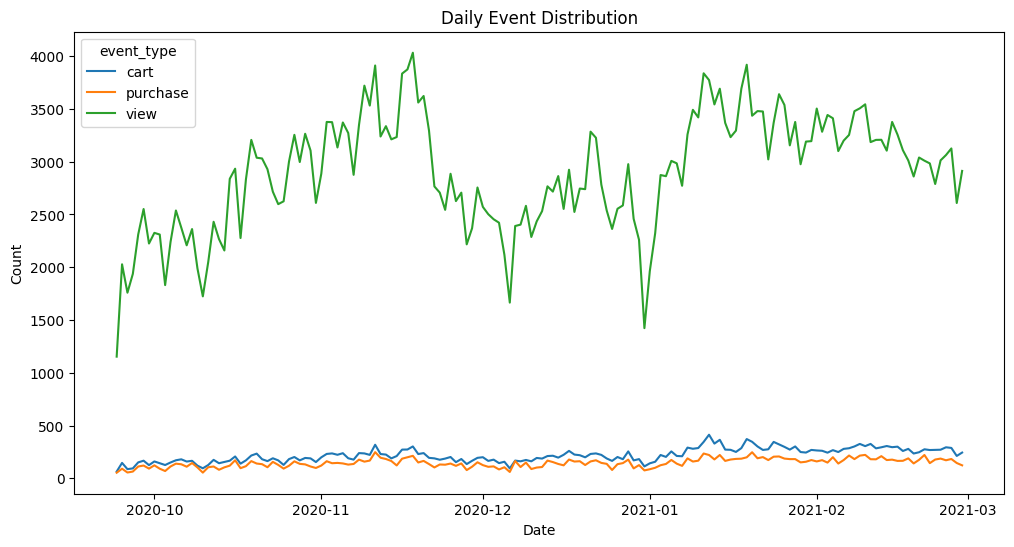
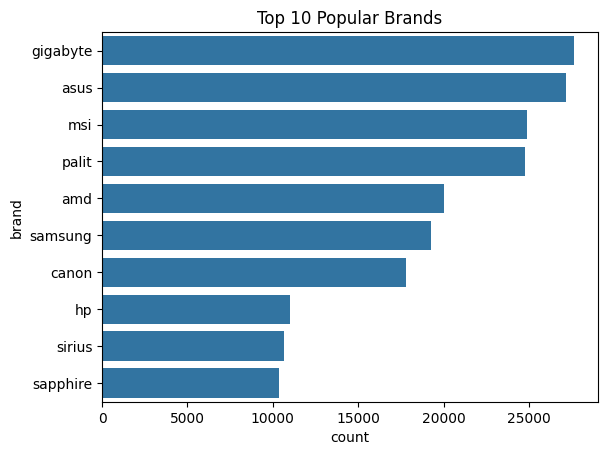
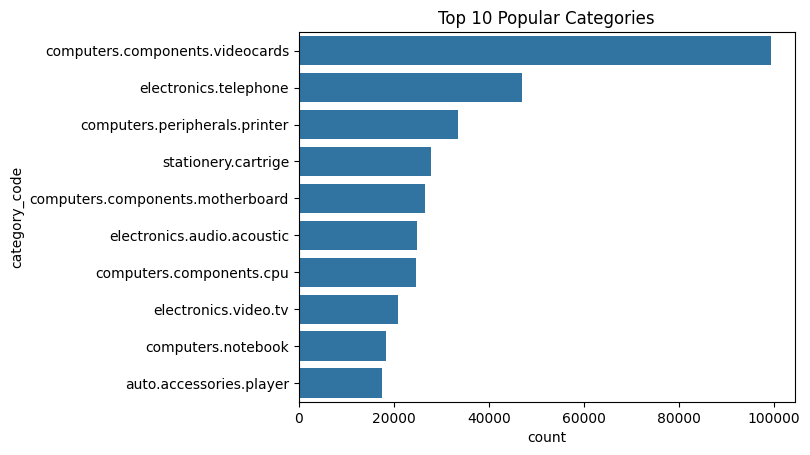
## Problem Definition

The objective is to predict user churn on an e-commerce platform by analyzing user activity data and providing actionable business insights. Churn is defined as users who stop engaging with the platform (e.g., no purchases within the last 30 days).

## Data Loading and Cleaning

The dataset was loaded from Google Drive using the path '/content/drive/MyDrive/Colab Notebooks/events.csv'.  
  
Key steps included handling missing values, removing duplicate rows, and converting the 'event\_time' column to datetime format. After cleaning, the dataset contained no missing or duplicate values, and data types were confirmed for analysis.

## Exploratory Data Analysis (EDA)

EDA provided insights into user behavior, including:  
1. Event distributions ('view', 'cart', 'purchase').  
   
  
2. Daily activity trends.   
  
3. Popular brands and categories.  
  


## Defining Churn

Churn was defined as users with no purchases in the last 30 days. The last purchase date for each user was calculated, and users with no activity within this threshold were flagged as churned. A total of 10,096 churned users were identified.

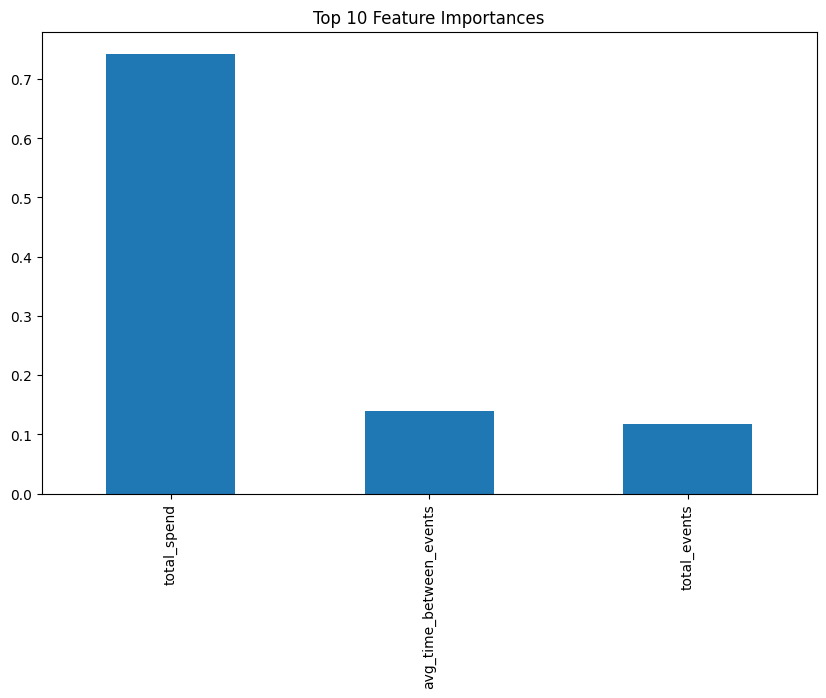
## Feature Engineering

Features created included:  
- \*\*Total Events per User\*\*: Count of all user activities.  
- \*\*Total Spend per User\*\*: Sum of prices for all purchases.  
- \*\*Average Time Between Events\*\*: Mean time difference between consecutive events.  
  
\*(Attach a bar graph showing feature importance here)\*

## Predictive Modeling

A Random Forest Classifier was used to predict churn. The data was split into training (80%) and testing (20%) sets. Evaluation metrics included a confusion matrix and a classification report. The model achieved an accuracy of 98% with the following key metrics:  
  
Confusion Matrix:  
[[42477 527]  
 [ 166 1825]]  
  
Classification Report:  
- Precision: 78%  
- Recall: 92%  
- F1-score: 84%

## Interpretation & Insights

Feature importance analysis revealed that total events, average time between events, and total spend were the key predictors of churn. Users with infrequent purchases or long session gaps are at a higher risk of churn.  
  


## Business Recommendations

1. Target users with high activity but no recent purchases using personalized offers.  
2. Re-engage users with long session gaps through email or app notifications.  
3. Optimize inventory and marketing strategies by focusing on popular brands and categories.