

JANUARY 2024

Pakistan's E-Commerce Data Analysis Report

Optimizing the Launch and Operation of a New
E-commerce Business in Pakistan:
A Data-Driven Strategy.

PRESENTED TO
Business Stakeholders

PRESENTED BY
Ali Bin Kashif



Table of Contents

01 INTRODUCTION

02 PROBLEM STATEMENT

03 ABOUT THE DATASET

04 DATA PREPARATION AND CLEANING

05 EXPLORATORY DATA ANALYSIS

06 SUGGESTIONS

Introduction

In recent years, Pakistan's e-commerce scene has undergone a remarkable transformation, with notable figures showcasing its rapid growth. With over **100 million** internet users and nearly **40%** of the population accessing the internet via smartphones, the digital landscape has become a bustling marketplace. The **COVID-19** pandemic acted as a catalyst, pushing e-commerce to the forefront as online shopping surged, contributing to a **35%** increase in e-commerce sales in **2020** alone.

Companies such as **Daraz**, reported significant milestones, including a record-breaking "**11.11**" sale with over **3 million** products sold in a single day. The convenience of digital payments has played a pivotal role, with digital wallets like **EasyPaisa** and **JazzCash** gaining prominence.



Problem Statement

A group of entrepreneurs want to start a new E-Commerce business in Pakistan. They have gathered the data of **half a million e-commerce orders** in Pakistan from **March 2016** to **August 2018**. Now as a data analyst, I have to analyze and explore the data to find out useful insights, answer various analytical and research questions, test different hypothesis with data driven approach and make a **Market Research Report**, so that entrepreneurs can take informed decisions after looking what the data says.

Assumptions before starting

- The dataset is still valid as of 2024 and can be used to analyze in an efficient manner.
- Assuming that there are records of all months in 2016 and 2018 as the dataset only consists of records from March 2016 to August 2018.

About the Dataset

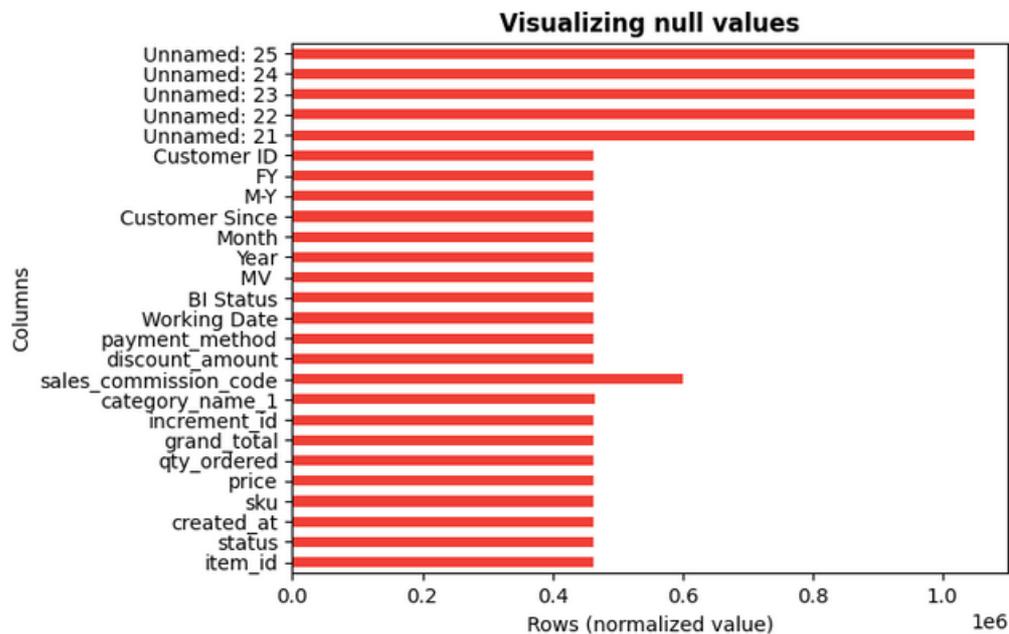
The dataset contains detailed information of half a million e-commerce orders in Pakistan from March 2016 to August 2018. It contains item details, shipping method, payment method like credit card, Easy-Paisa, Jazz-Cash, cash-on-delivery, product categories like fashion, mobile, electronics, appliance etc., date of order, SKU, price, quantity, total and customer ID. This is the most detailed dataset about e-commerce in Pakistan that you can find in the Public domain. The data was collected from various e-commerce merchants as part of a research study.

Key Variables:

- Item ID
- Order Status
- Date of Order
- SKU
- Price
- Quantity
- Grand Total
- Category
- Payment Method
- Customer ID.

Data Preparation & Cleaning

- Initially dataset had **1048575** records and **26** columns.
- Checked for null/irrelevant data.



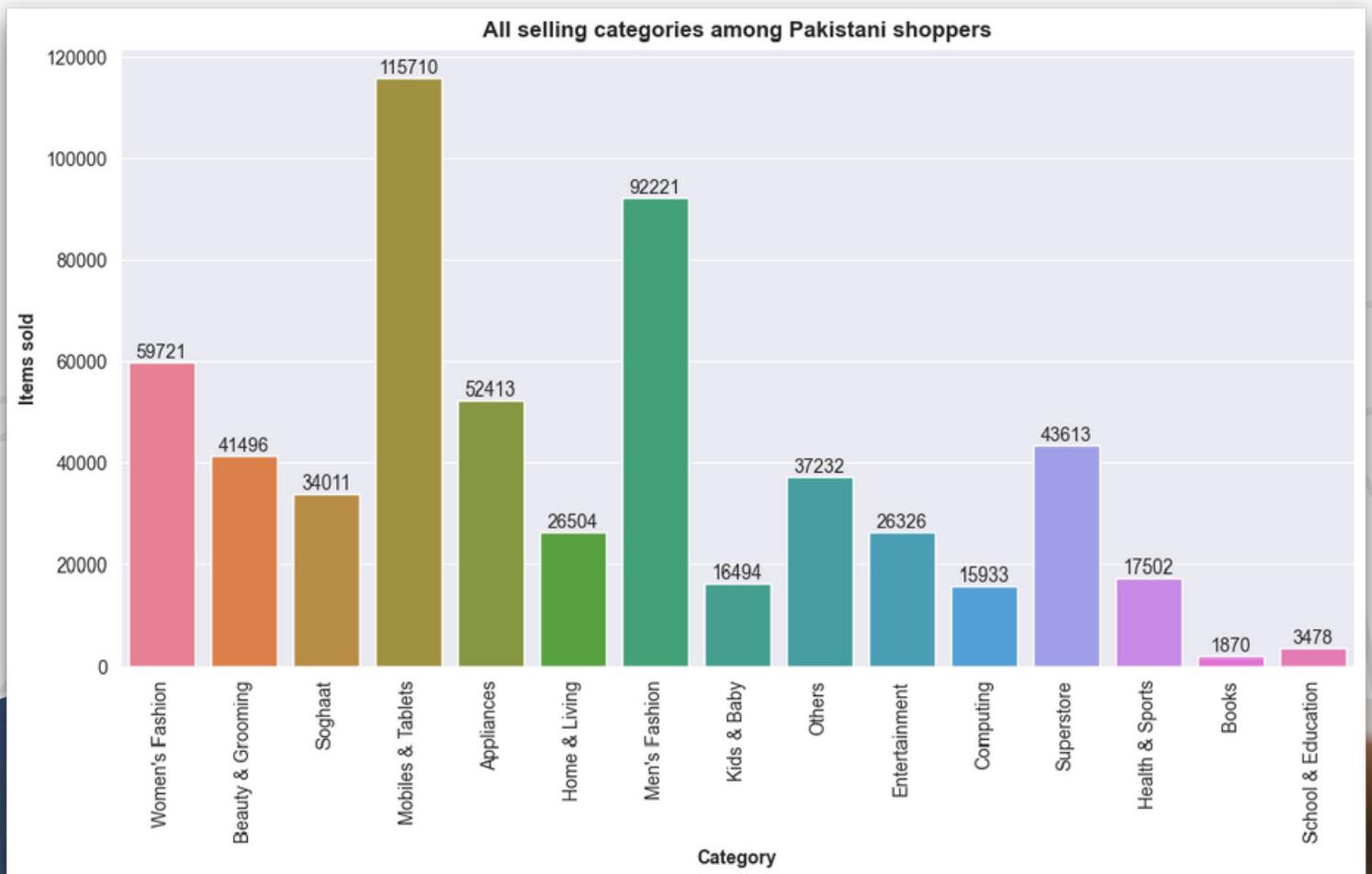
- As you can see 4 columns are 100% null and in other columns almost 50% rows are empty.
- All null rows and columns were removed, redundant columns were also dropped.
- Encoded date columns into date-time format.
- Identified many outliers, data was very spread out, as the dataset contains prices and products whose ranges varies alot. So no need to remove them as they are true outliers.
- After cleaning and preprocessing, dataset was reduced to **584524** records and **15** columns.

Exploratory Data Analysis

Explored the cleaned dataset to extract insights and findings according to the research/analytical questions of the stakeholders.

Some key insights and findings

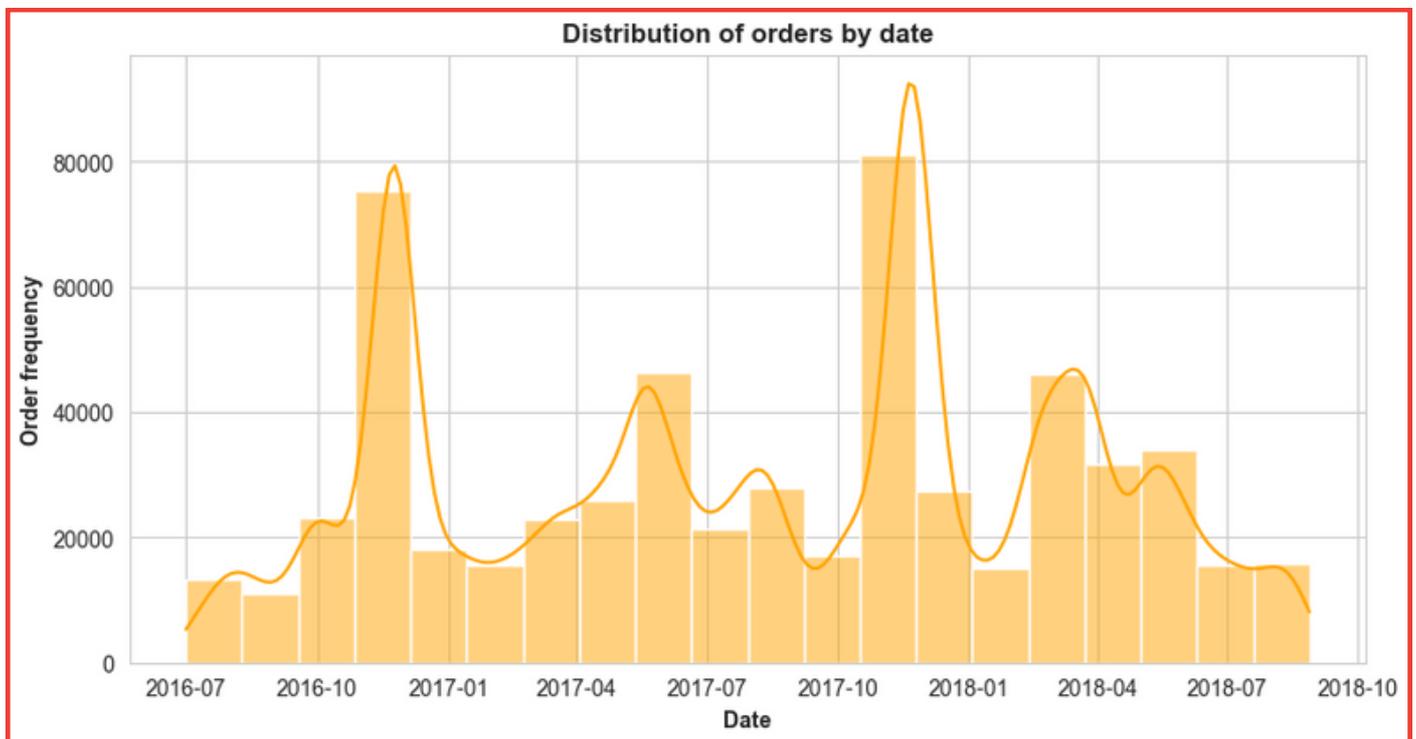
Best-selling product categories among Pakistani online shoppers.



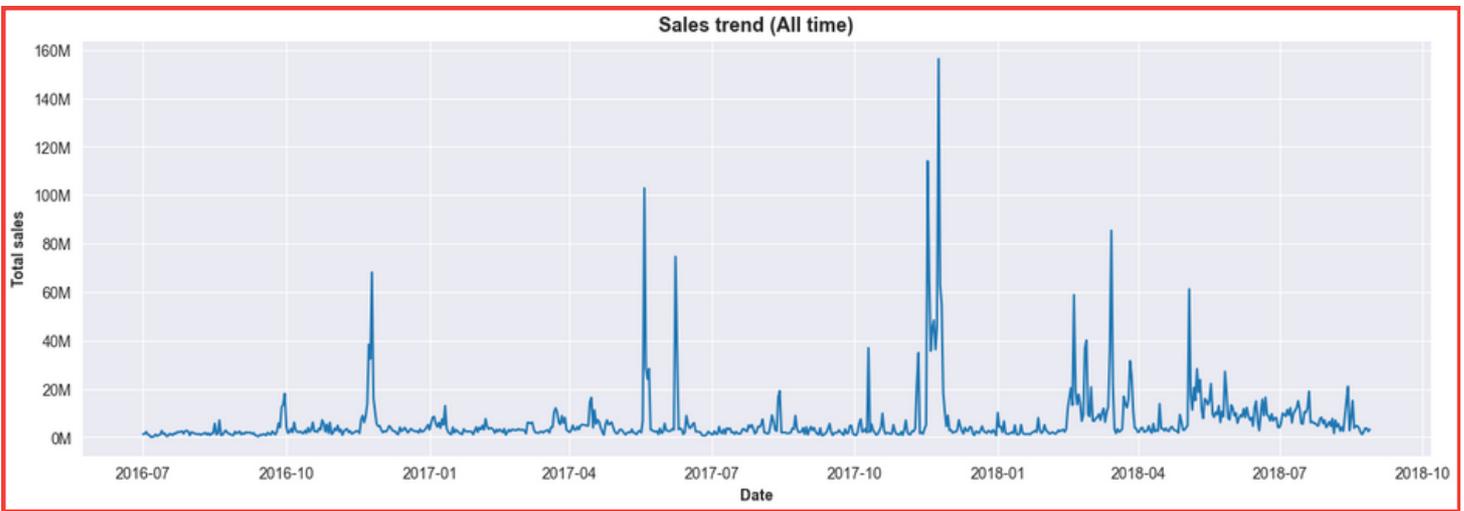
We can see that:

- Mobiles and Tablets is the best selling category, along with Men and Women's Fashion and appliances.
- Books and education stuffs are the least sold.

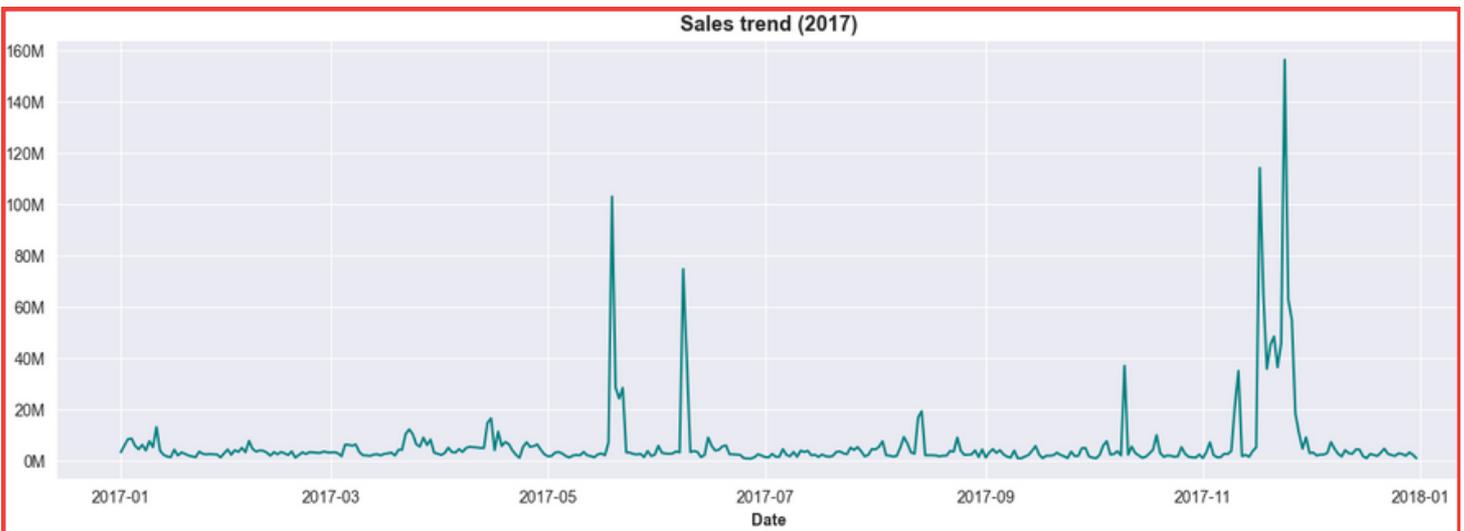
Sale trends of previous years and months with the most sales.

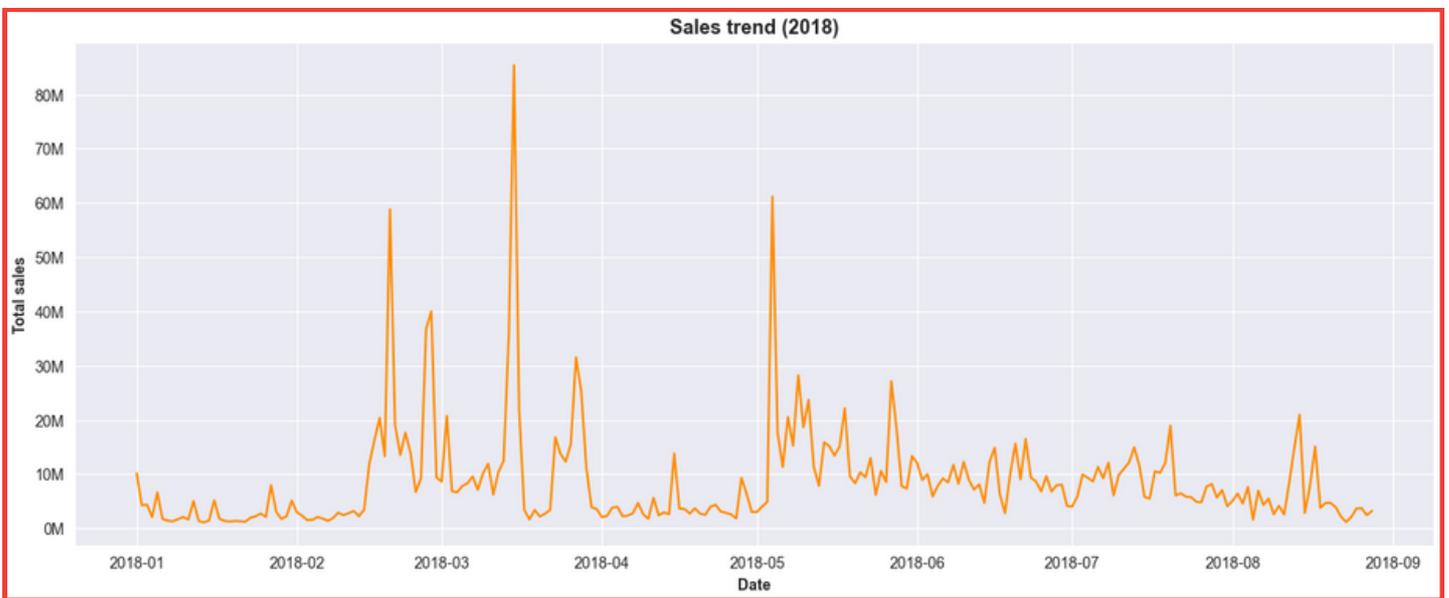


- It is a multimodal distribution.
- You can see different peaks which represents there are large number of orders in that specific time period.
- There are high peaks are at the end of each year.
- Now we will see the sale trends.

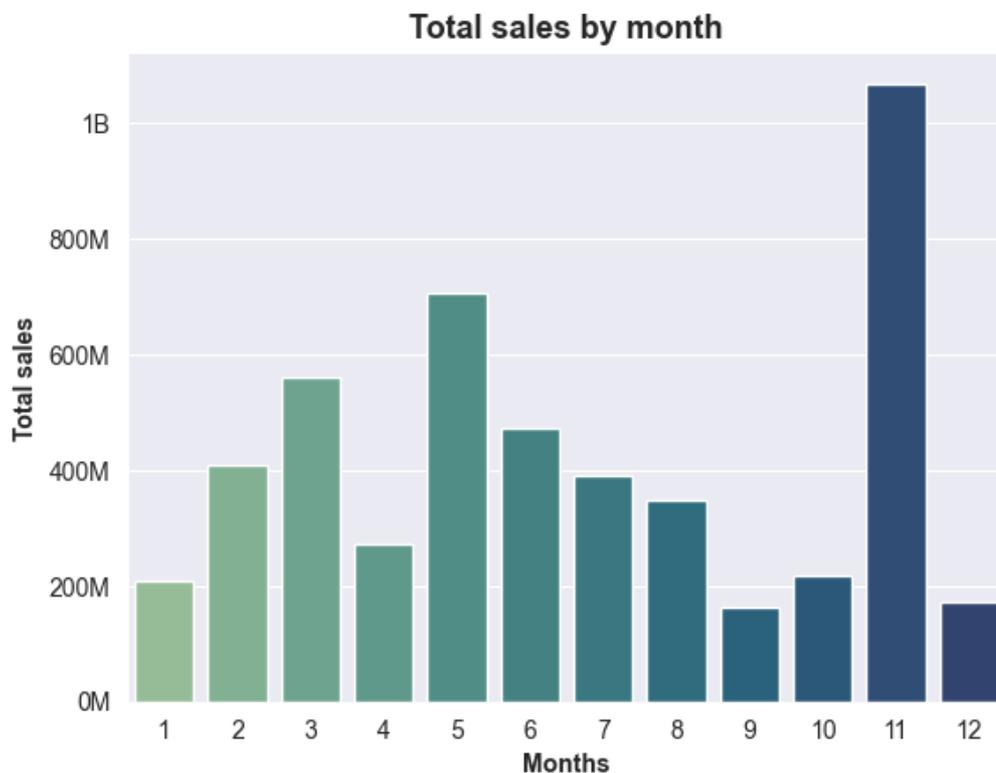


- There are hikes especially at the end of years. Hike indicates a lot of sale, these are due to promotional and sales campaigns of different E-Commerce companies and special events. Let's have a closer look at each year.





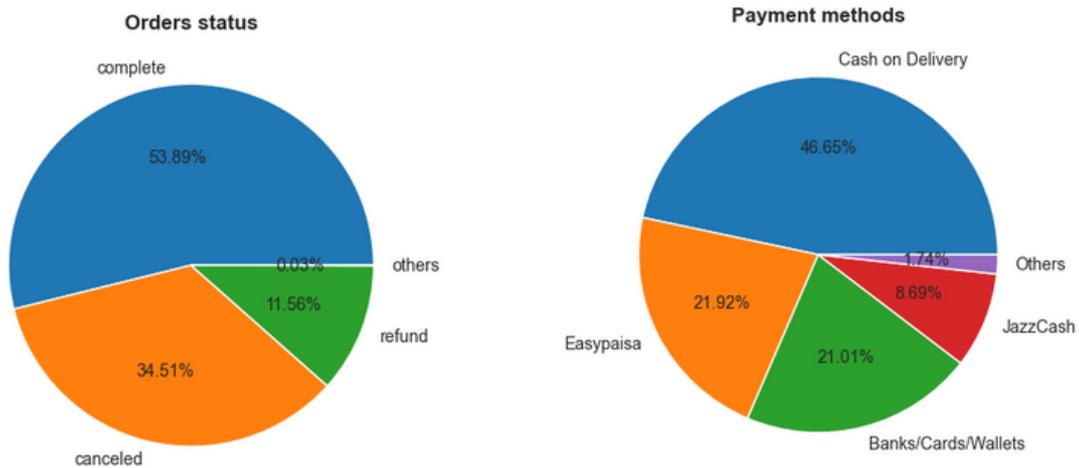
- These were the sales trends of each year, the highest peak recorded was in **November 2017** that almost touch **160 million PKR**. Let's see total amount of sales by each month.



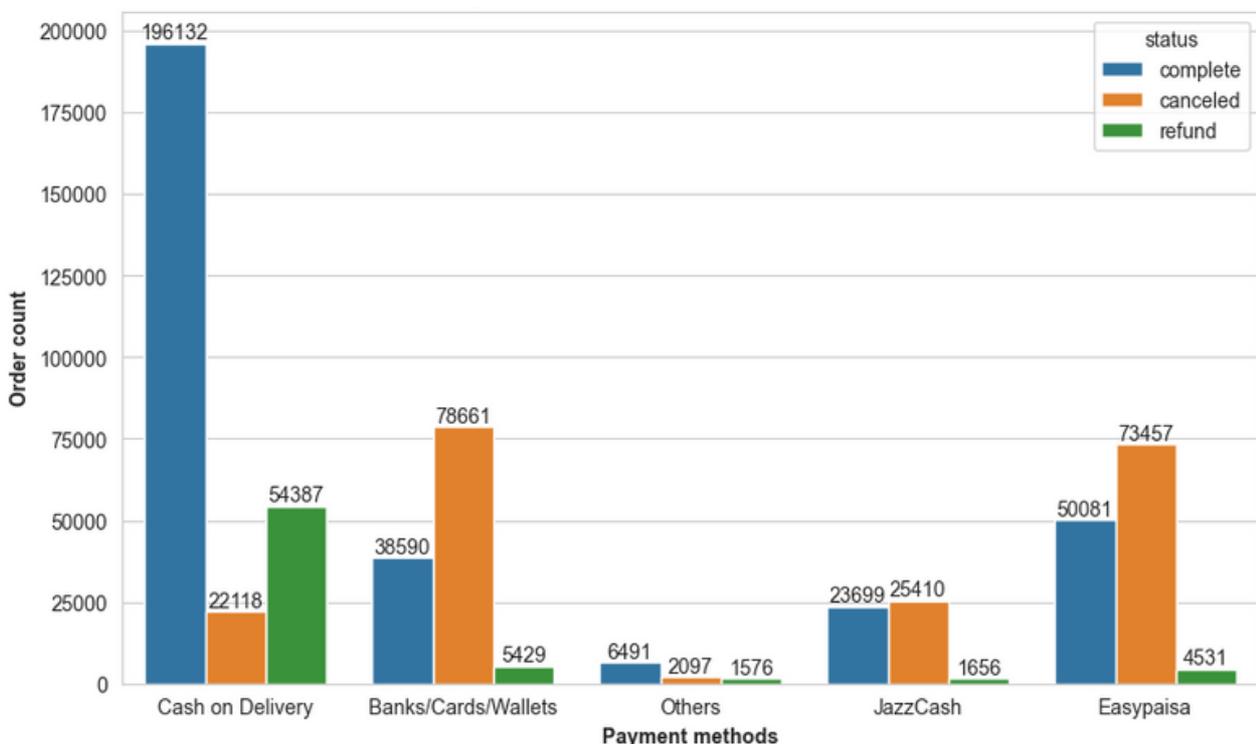
- In November there are the most sales due to different sales and campaigns e.g: **11-11**
- Sales are also high in May and June due to events like Ramadan and Eid.

How does the choice of payment method (credit card, Easy-Paisa, Jazz-Cash, cash-on-delivery) influence the completion and satisfaction of e-commerce orders?

Percentage count of orders by their status and payment methods



No. of orders for different payment methods with status

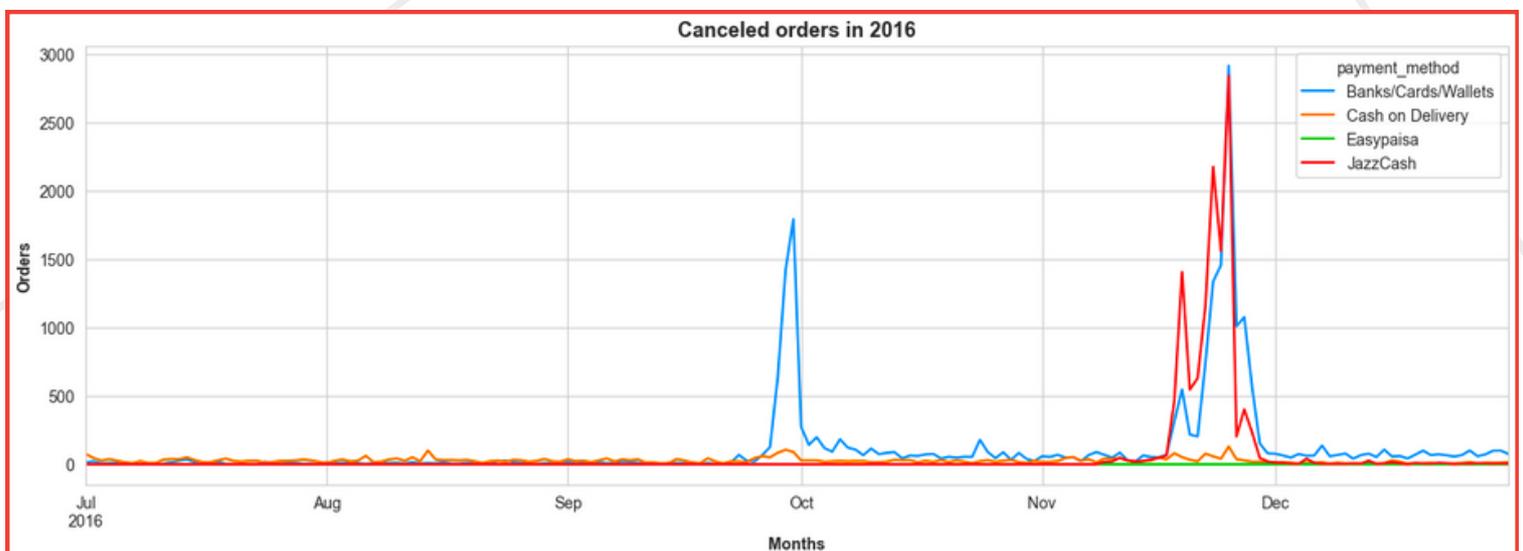
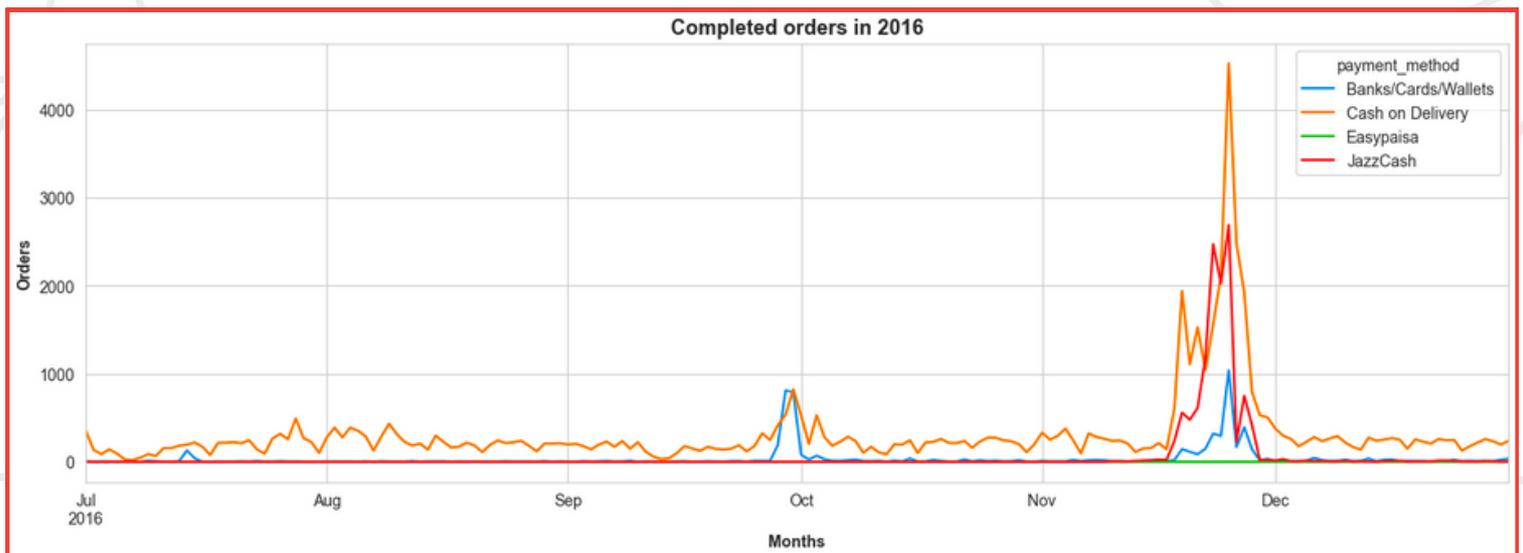


- There is a **high cancellation** rate in online payments.
- **Easypaisa** and **Banks/Cards/Wallets** has the highest rate of cancellation.
- **Cash on Delivery** orders are the most successful with most them marked as completed.

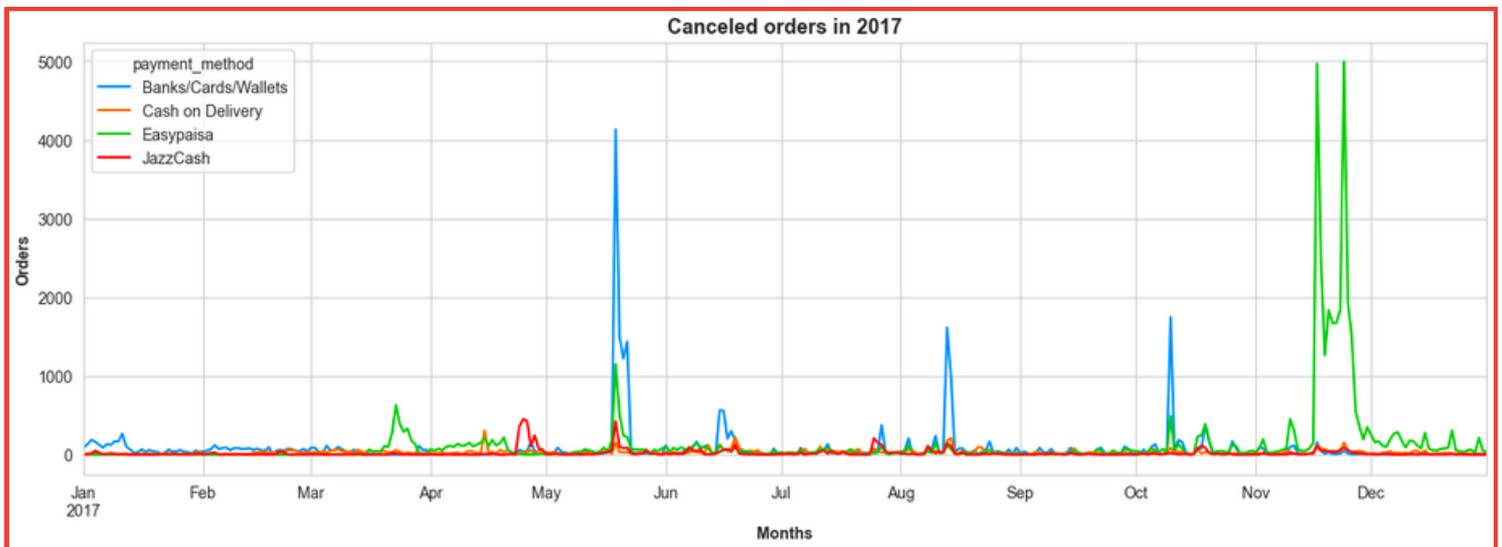
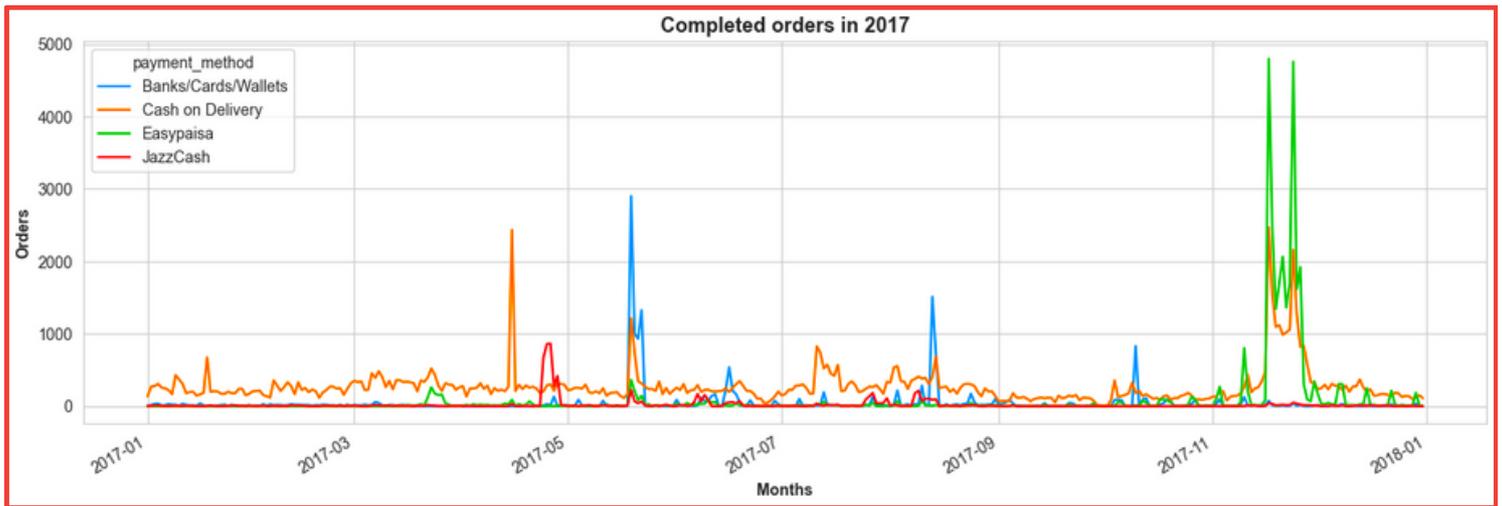
Let's look how market is performing over the years with different payment methods.

- Now I have categorized orders into **completed** and **cancelled** orders.
- Each line is of different color and represent payment methods as shown in the legend.
- We have date on the Y-axis and no. of orders on the X-axis.

In 2016



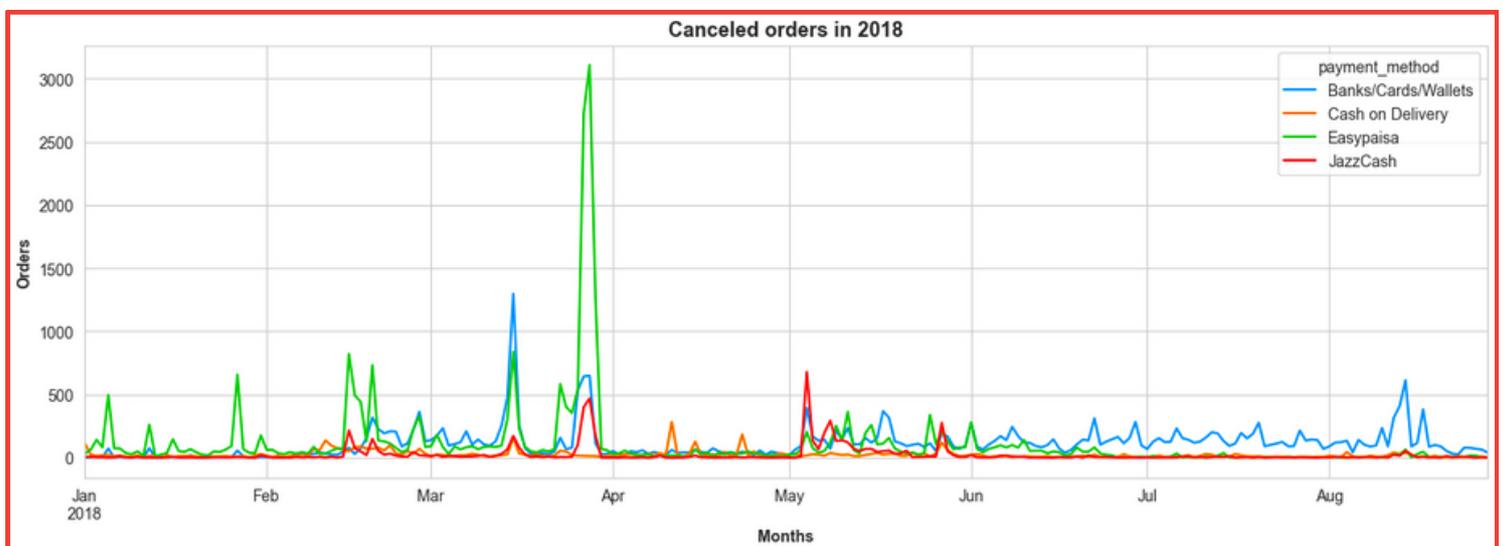
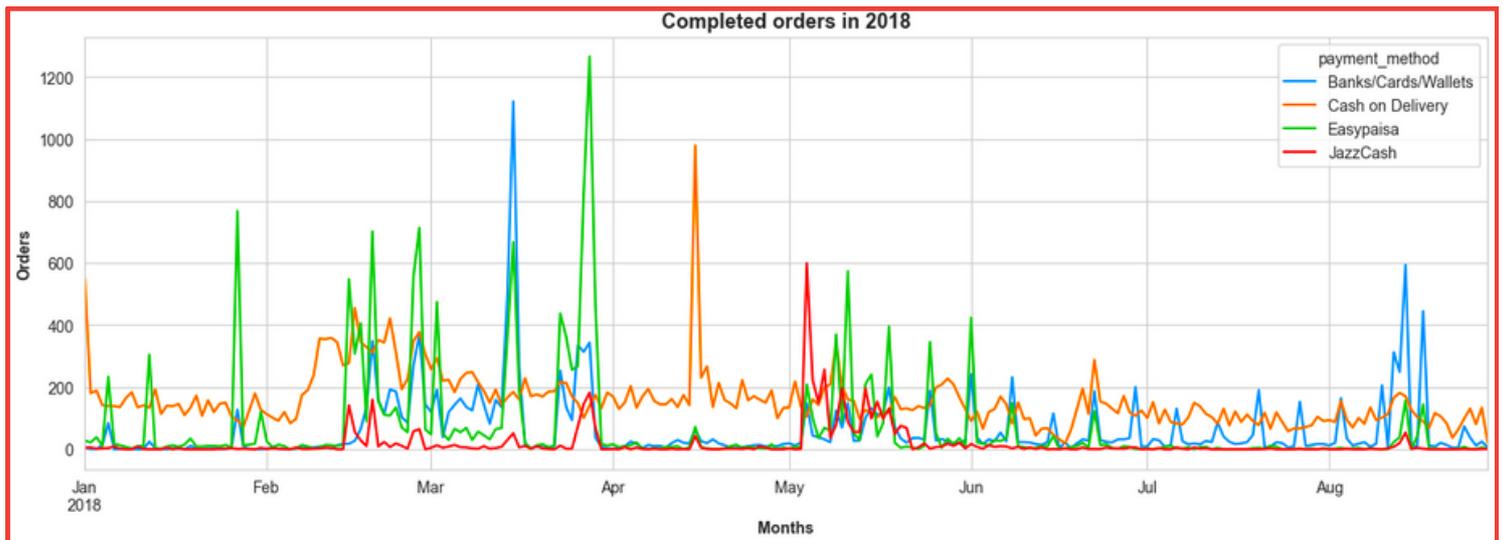
In 2017



- Orders with **Easypaisa** as their payment method hiked in 2017.



In 2018

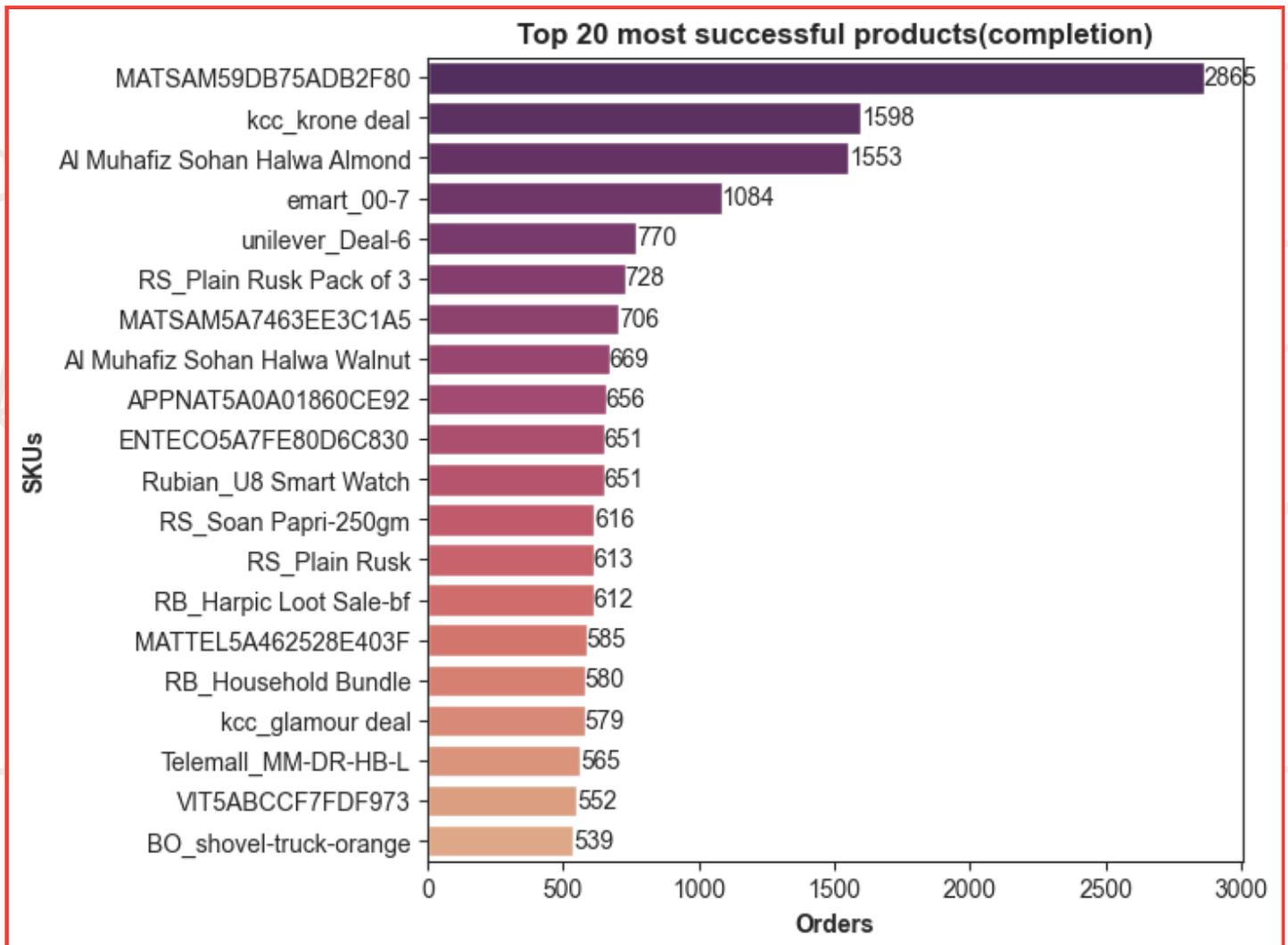


Findings:

- Customers find it more convenient to pay **cash on delivery** for the order, so this method is quite stable and also have very **less order cancellations**.
- There are **more cancellations** in the **online payment methods**.
- People prefer Cash on delivery more, but you can see **hikes** in **online payments**, it may be due to **discount vouchers, sales** and different **campaigns** which are eligible for **mobile wallets** and **internet banking**.

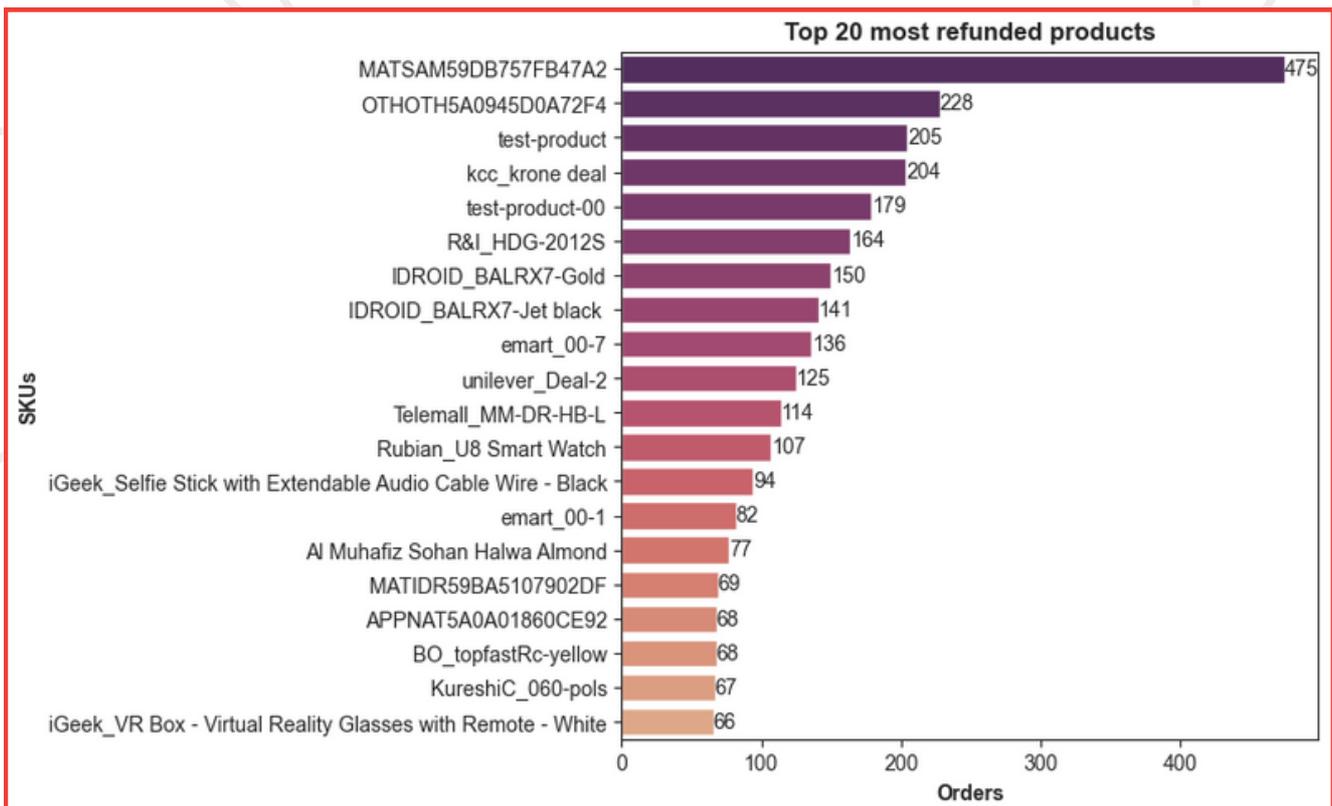
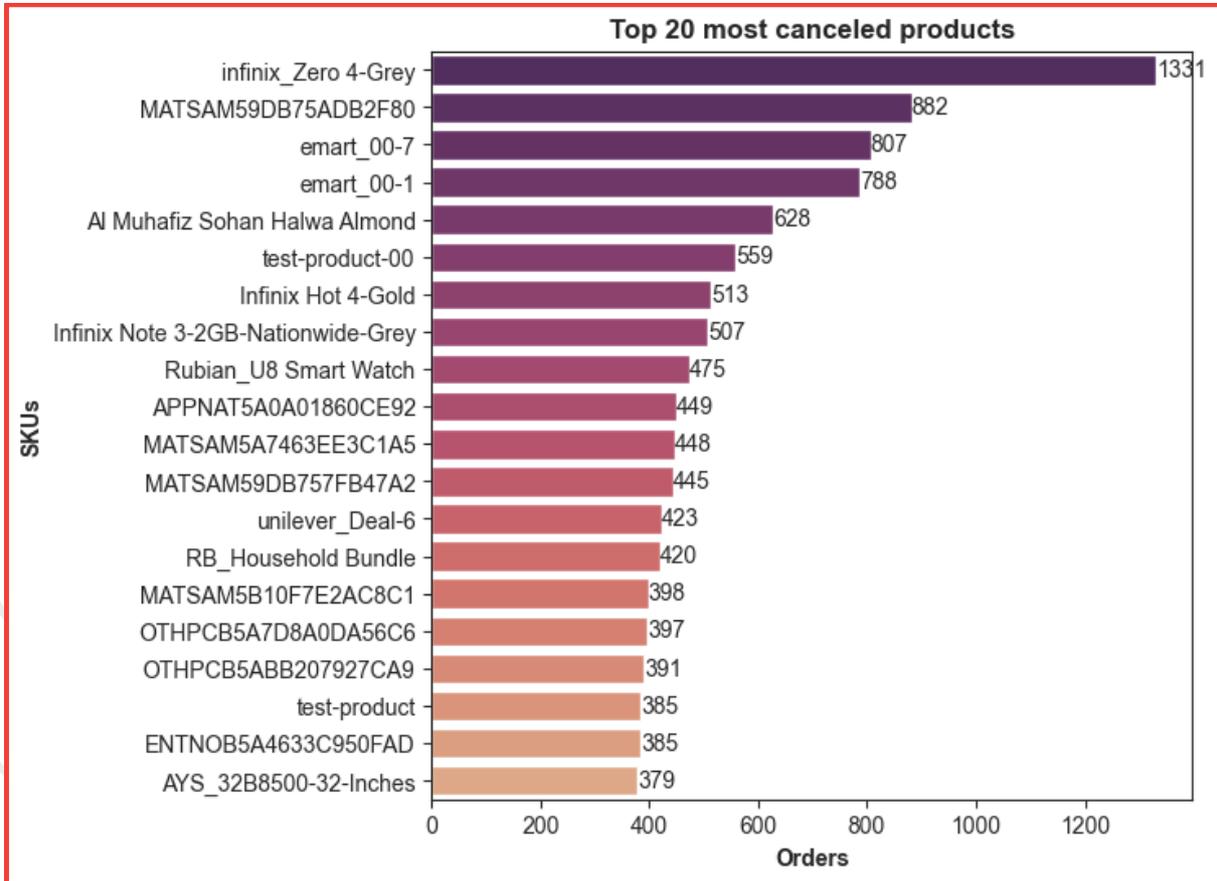
Performance of different categories and their products.

- Now let's look at some of the best selling products along with most cancelled, returned and refunded products also.

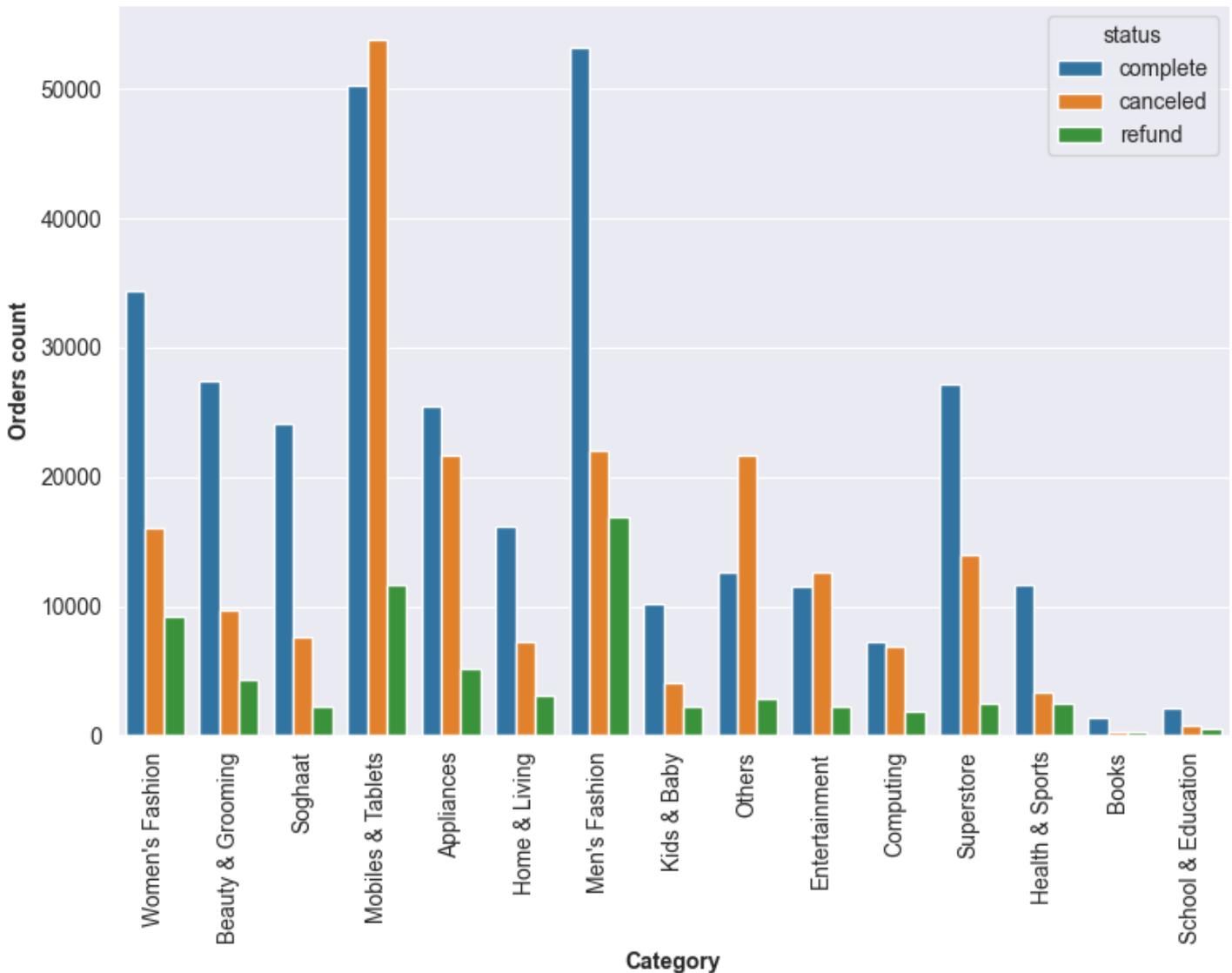


- In e-commerce, a cancellation is when a customer cancels an order for products or services. This can include orders that are in the process of being purchased, as well as orders that have already been delivered.

- A refund can only be performed on transactions where the funds have already been transferred to the bank account. A cancellation or deletion can be done before a payment has been fully processed.



Product categories with their order count and status.



- **Men's and Women's fashion** are both successful categories as they have more **completions** as compared to cancellations and refund.
- **Mobiles & Tablet** is also a good category but also has a high risk. As you can see the cancellation rate is more than the completed rate.

Suggestions

Optimize Online Payment Process

Given the high cancellation rate in online payments, it's essential to evaluate and optimize the online payment process. This may involve improving the user interface, providing clearer instructions, and addressing any technical issues that could contribute to cancellations.

Encourage Digital Payment Adoption

While Cash on Delivery orders are successful, consider incentivizing digital payment methods to shift towards a more digitalized payment landscape. This could involve offering discounts, exclusive deals, or loyalty rewards for customers using online payment options.

Mitigate Risks in Mobiles & Tablets Category:

Recognizing the high-risk nature of the Mobiles & Tablets category, implement measures to mitigate risks such as offering detailed product information, clear return policies, and perhaps additional customer support for this category.

Strategic Marketing for November and Festive Seasons:

Leverage the high sales periods in November and during Ramadan/Eid by implementing targeted marketing strategies. Plan special promotions, discounts, and exclusive deals to capitalize on the increased consumer activity during these times.