Name: Likhith K Raj
LinkedIn: [LinkedIn Profile Link](#)
GitHub: [GitHub Profile Link](#)
Portfolio: [Portfolio Link](#)

Zeotap Data Science Intern Assignment

Task 1: Exploratory Data Analysis (EDA) and Business Insights

Google Colab Link:
https://colab.research.google.com/drive/1apsczs9x_YEUl8L6PcF6iFvw1jAy Kn7?usp=drive_link

Introduction:-

- There was 3 data csv files given, the first step was to clean data and merge them using common factor (Product Id, Customer Id).
- Checked for duplicates, null values unformatted values etc.., then processed with the EDA.
- Here is the data summary :

```
Customer Data Summary:
------------------------
- Total Customers: 200
- Most Popular Region: South America (with 59 orders)
- Most Active Customer: Customer C0109 (with 11 orders)
- First Signup Date: 2022-01-22 00:00:00
- New Signups in December 2024: 1

Product Data Summary:
-----------------------
- Total Products: 100
- Highest Priced Product: TechPro Textbook ($497.76)
- Most Ordered Product: SoundWave Cookbook (with 46 orders)
- Total Revenue: $332669.30
- Top Selling Category: Books

Transaction Data Summary:
--------------------------
- Total Transactions: 1000
- Busiest Month: 2024-01 (with 107 transactions)
- Highest Transaction Value: $995.52
- Average Transaction Value: $447.74
```
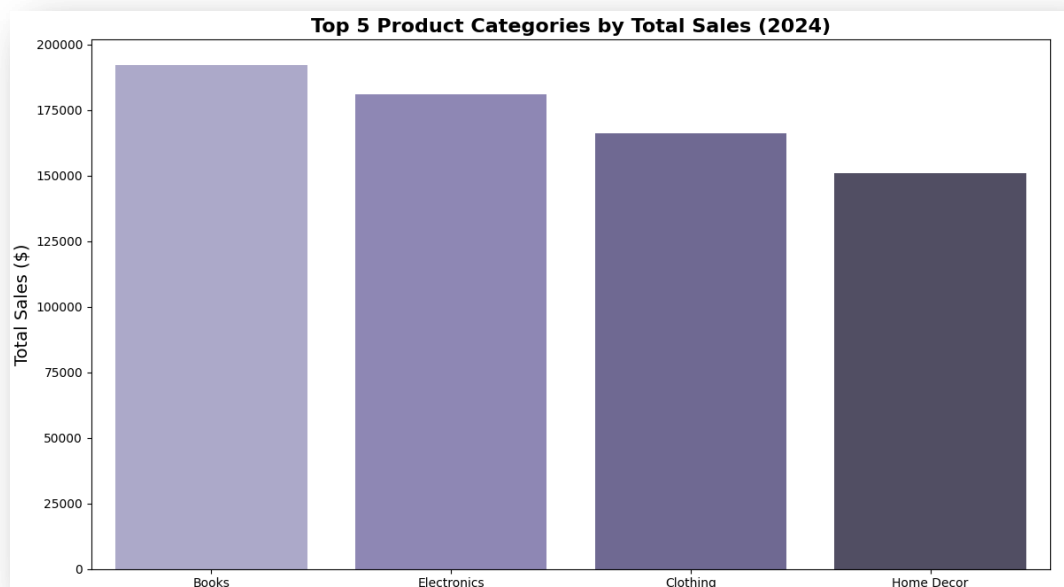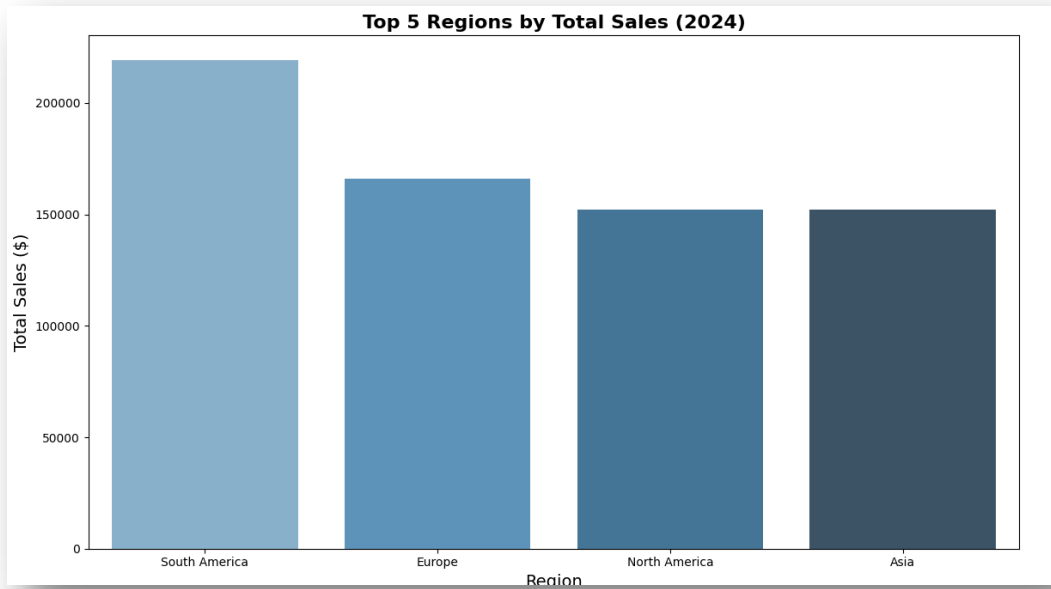
After this, we can move with some visualization.

Data Visualization:

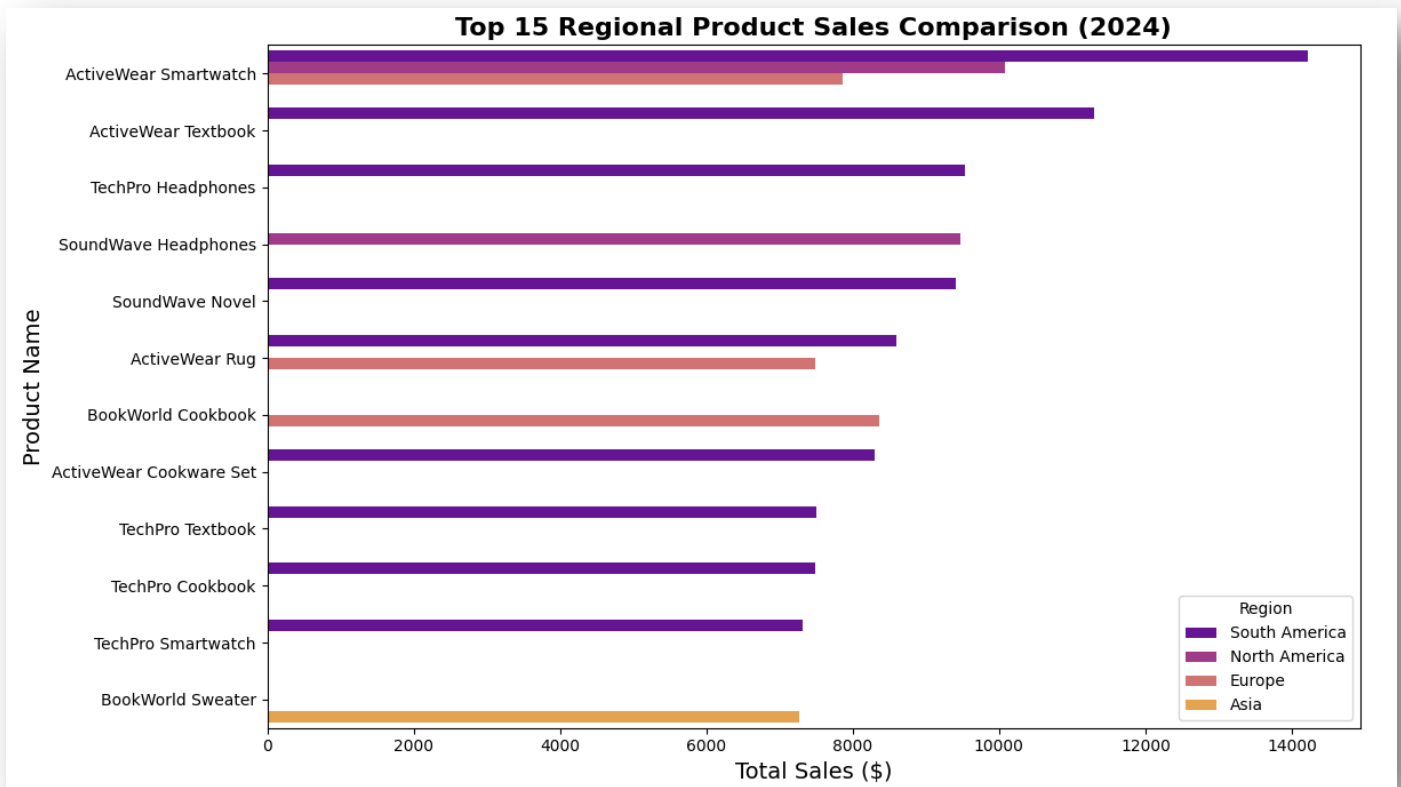In visualization part, I have included 6 graphs which showcase to business insights.

1. <u>Total Number of sales and region</u>

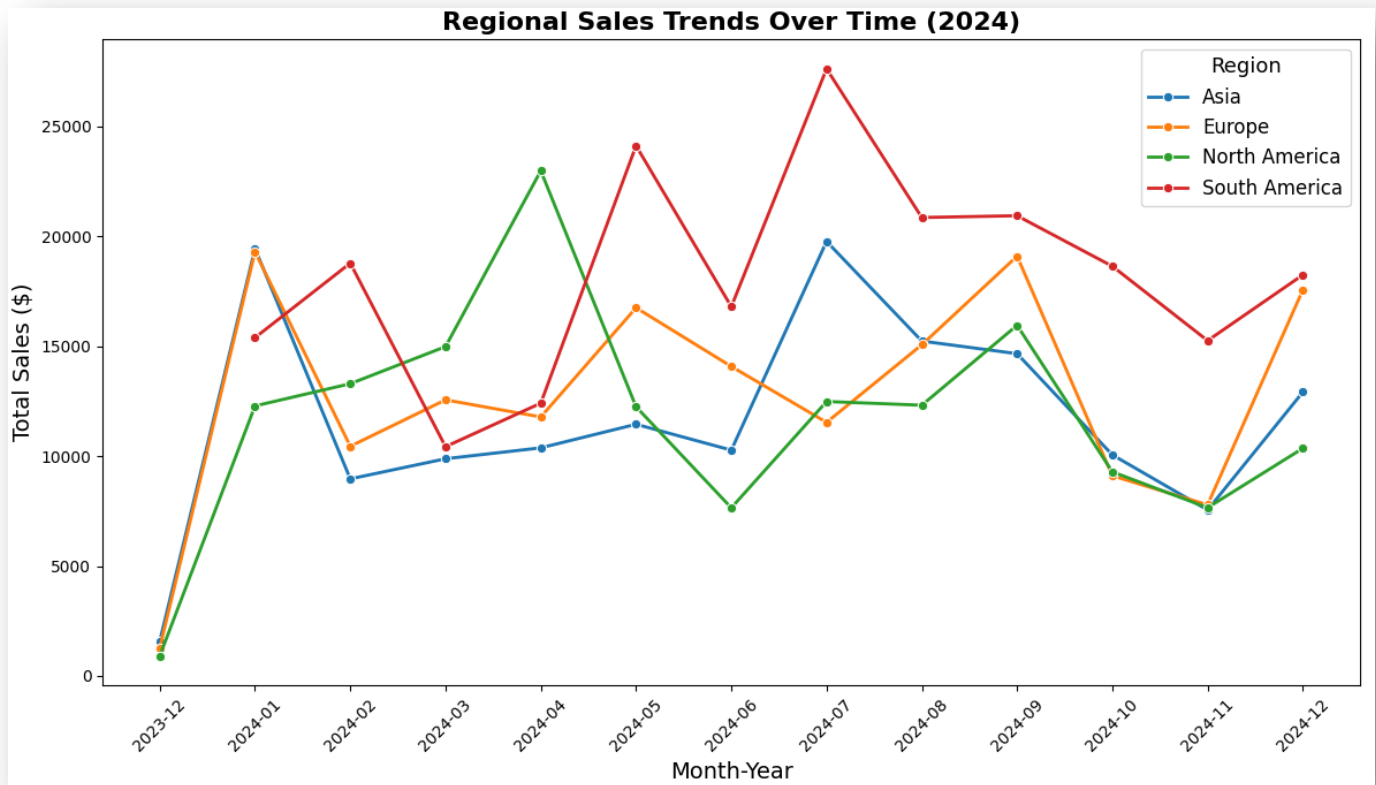According to the **data timed 2024, South Africa** has the major contributions in total sales, and in **product categories books** are on the top.

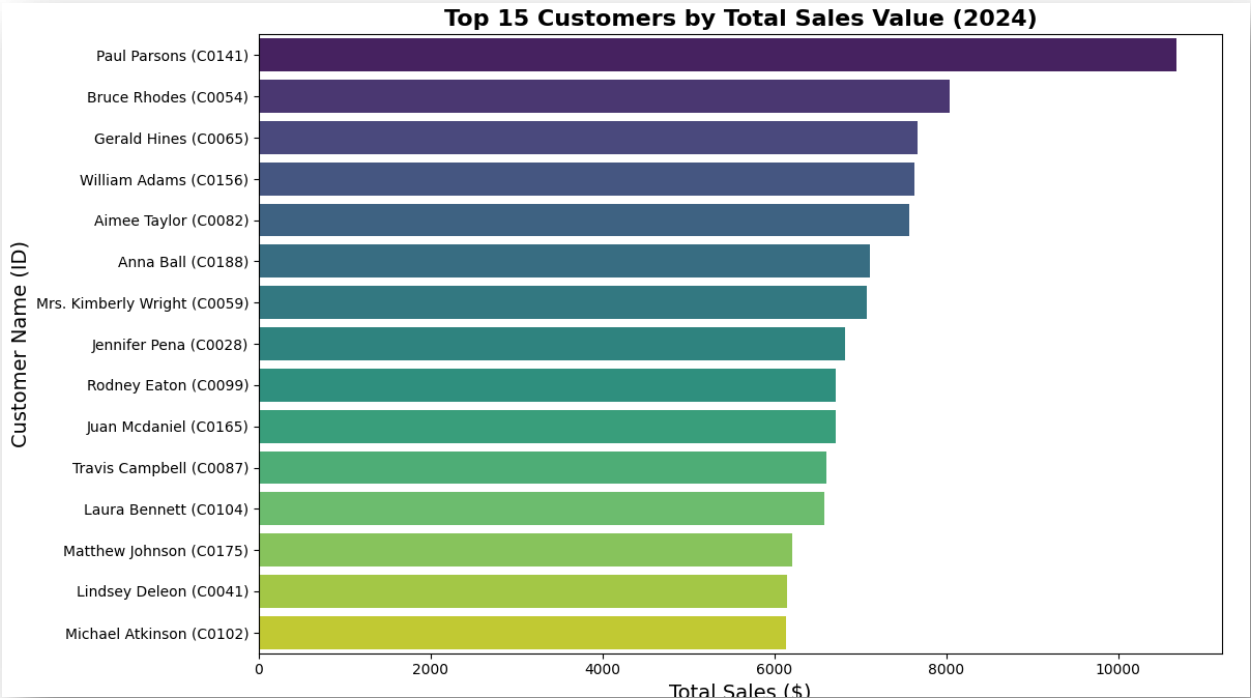## 2. Top 15 Regional product which was sold in 2024

Top 15 Regional Product Sales Comparison (2024)

"Active Wear Smart watch" was the most sold regional product with **$14,000**. And most of the products were sold in **South Africa (2024)**

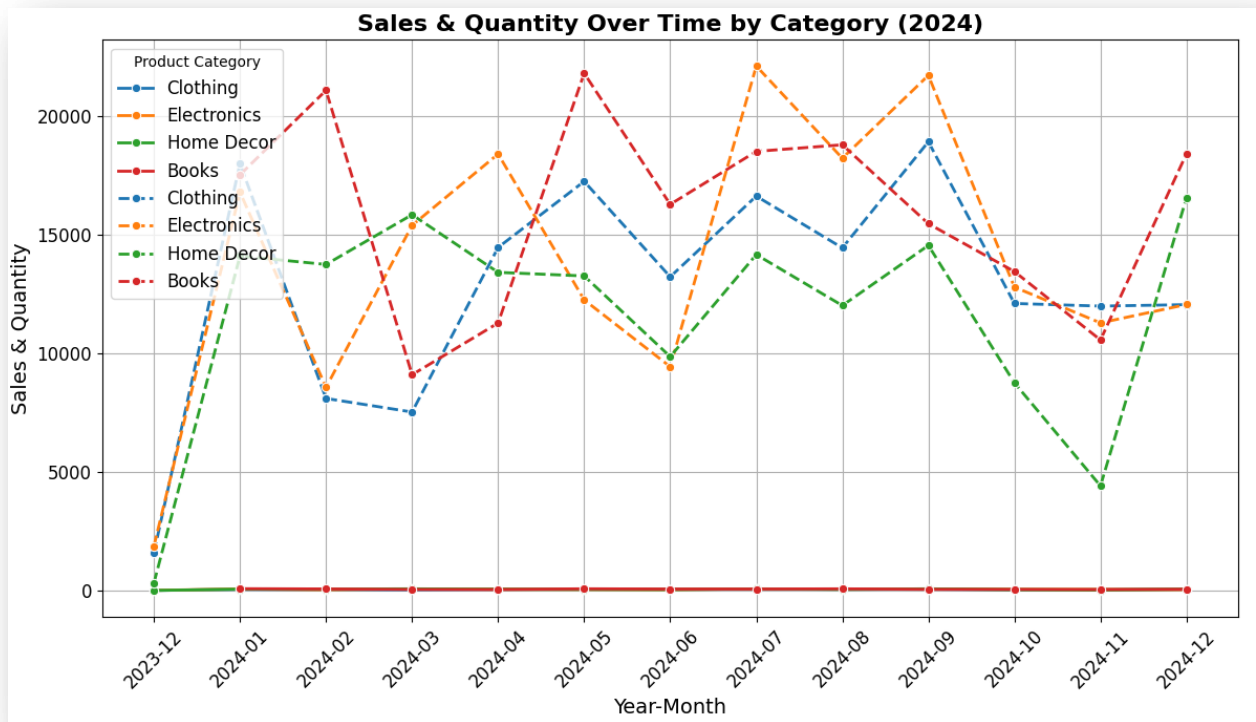### 3. <u>**Regional trends of sales overtime (2024):**</u>

The majority of the sales was generated on the month of **August (2024)** and **South Africa scores highest again.**

4. Top 15 Customer contribution to sales on 2024

Top 15 Customers by Total Sales Value (2024)

Customer ID (C0141) has ordered the most and **contributed more than $10,000** in 2024, belong to the **region South Africa**

5. Sales and Quantity distribution over the time (2024)

Sales & Quantity Over Time by Category (2024)

The **books** has the high sales **on May (2024**), and **Electronics** has high sales on the **month of July (2024).**