

Analyst

Angie aragon sogamoso

Goal

This project is based on fictitious data from a sports store. The store aims to ramp up sales and find new business opportunities by offering customers innovative sports experiences. The data includes age, gender, sports preferences, and purchasing behaviors, which will be analyzed to uncover trends and insights. Understanding customer needs will be key to offering personalized experiences. The data allows the store to analyze the details of each client, helping to improve the products offered and better manage stock investment.

This analysis will help the store tailor its offerings and marketing strategies to better meet customer preferences. By focusing on emerging trends and individual needs, the store can create a more engaging shopping experience that not only boosts sales but also builds long-term customer loyalty.

Methodology

The data covers 4 years after the COVID-19 pandemic. During this time, people around the world began to look for better ways to stay healthy. Many people now practice sports to improve their health and connect with others through shared hobbies and goals. This methodology looks at how these changes are shown in the data, focusing on trends in sports participation and community involvement. For this project, I used Python, Seaborn, and Matplotlib to analyze and visualize the data.

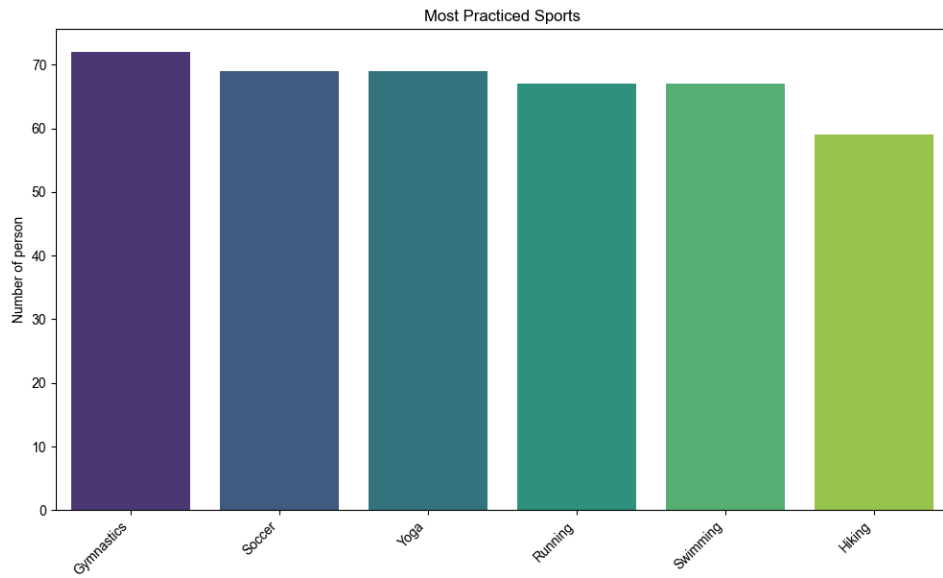
Data Analyzed

The analysis seeks to identify patterns in purchasing behavior to assist the store in making informed decisions. By examining data trends and customer preferences, the goal is to provide actionable insights that enhance decision-making processes and drive business growth.

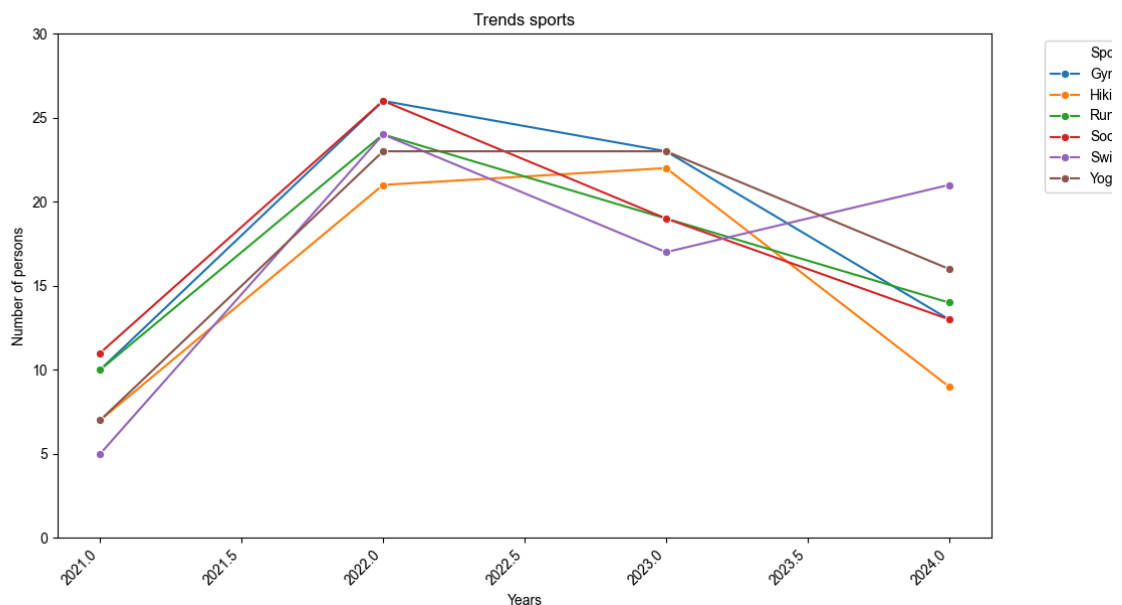
1. Customer Behavior

The store collects data on each customer, and to gain deeper insights into our clients, we decided to analyze these details.

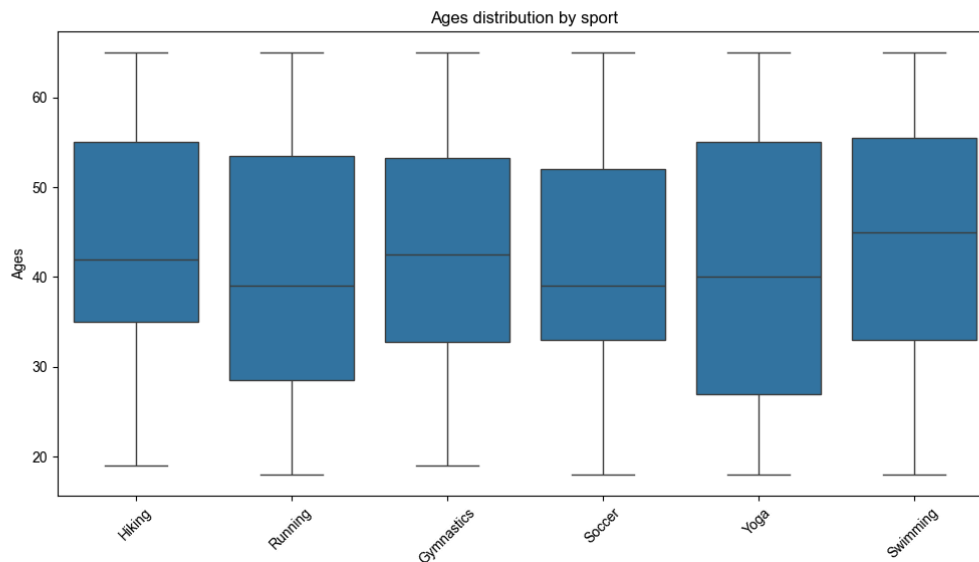
- **Most practiced sport by number of people:** In the bar chart, we can see that the 7 sports show no significant differences among them; however, the sport with the most participants is Gymnastics. This may be due to trends influenced by the Coronavirus pandemic, which encouraged people to improve their health.



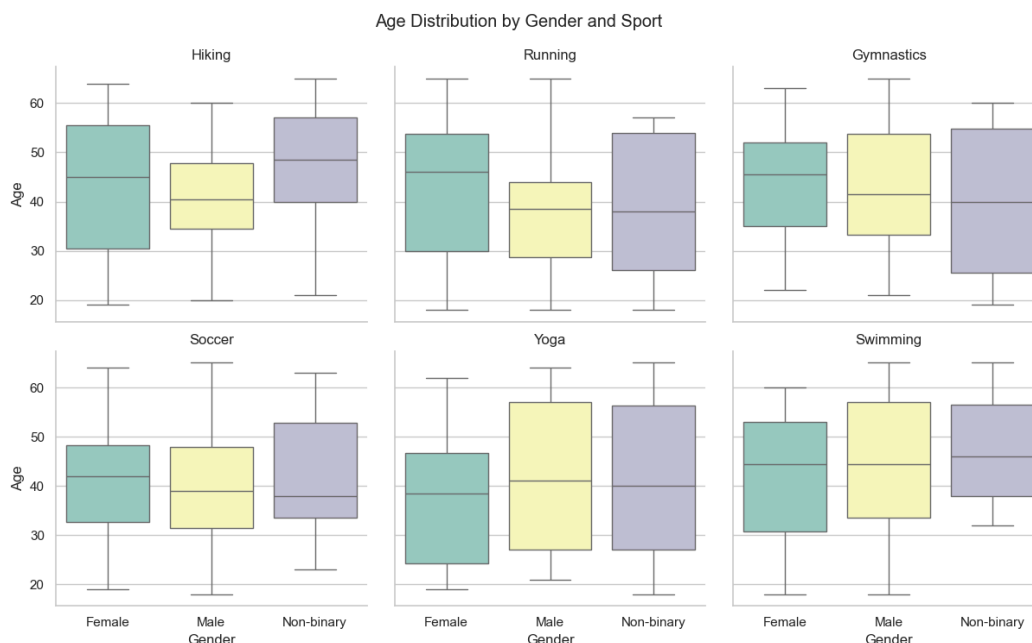
- **Sports trends by age and year:** Seven sports were analyzed across different years 2021, 2022, 2023, and 2024 to see how many people practiced each sport during this time period. We can observe that sports such as Gymnastics, Hiking, Running, Soccer, Swimming, and Yoga increased from 2021 to 2022. However, Hiking and Yoga remained stable until 2023 and then dropped in 2024. Soccer and Running began to decline in 2022, while Gymnastics and Yoga slowly decreased from 2022 to 2023. During the period leading up to 2024, the number of participants in these sports continued to decrease. In contrast, Swimming saw a drop in participants in 2023 but showed an upward trend from 2023 to 2024.



- **Age distribution by sport:** A wide range of people aged between 25 and 55 practice Yoga, while people aged between 28 and 54 practice Running. This indicates that these sports appeal to a broad variety of generations and preferences. In contrast, sports such as Hiking, Soccer, Swimming, and Gymnastics show age distributions that remain within the range of 30 to 50, attracting primarily young adults and adults.



- **Age distribution by gender and sport:** In Hiking, the age distribution varies across the three genders. Females predominantly practice this sport between the ages of 30 and 55, with the highest concentration around age 45. Participation declines towards age 30 and picks up again between ages 55 and 65. Males show slower participation between the ages of 35 and 47, with the highest concentration around ages 40 to 47. Non-binary individuals participate between the ages of 40 and 67, with a peak concentration around age 48.



Insights

The store has several ways to increase sales. For example:

- The trends can help us offer better products that match what customers need based on the demographic data we have.
- Each sport is different and depends on who is practicing it, including their physical condition, age, and whether they are a beginner, amateur, or professional.

RECOMMENDATIONS

1. Sports trends changed before the coronavirus pandemic, with less conventional sports like swimming gaining popularity due to their health benefits. The age of people who practice swimming ranges from 18 to 65, with a higher concentration between the ages of 34 and 55, and peak participation around age 45. Keep the age ranges in mind, as there are economic differences between young adults and adults, as well as different goals in sports.

For people between 18 and 35: Increase stock of beginner swimmer products, such as swimming poles, flotation devices, support equipment, swimming caps, and swimming goggles. Consider offering customization options, like personalized designs, to motivate and engage customers.

People between the ages of 34 and 45

This age range can be considered more advanced, so the stock should focus on self-care products such as chlorine-removal shampoos, sunscreen, and moisturizers. Additionally, offering natural nutritional supplements can support better nutrition. Swimmers at this level may require personalized attention to help them achieve their athletic goals

Master Swimmers Aged 45 to 55

In this category, athletes also require self-care products such as polarized sunglasses, muscle recovery creams, and nutritional supplements. Offering low-impact training programs focused on improving technique and strength would also be a valuable addition for the store.

Analysis of Sales Products

This report focuses on the sales performance of products in the sports store. Its goal is to analyze purchasing patterns and identify opportunities to improve stock management, sales, and marketing strategies.

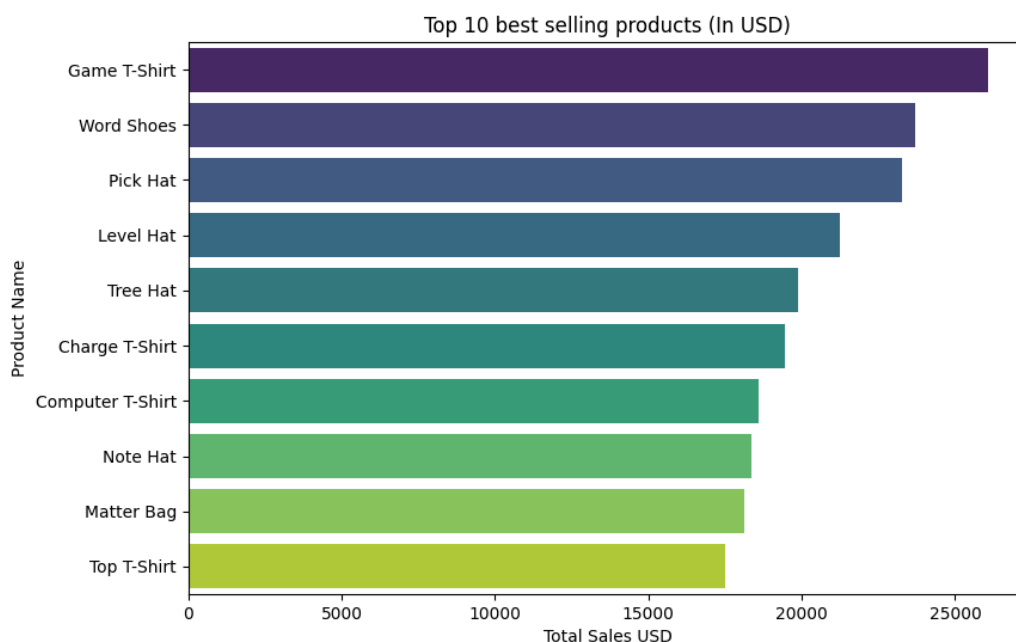
Methodology

In this report, we analyzed data from after the Coronavirus pandemic, as customer behavior has changed significantly. Online shopping has become the norm, and current trends have greatly influenced customer preferences. We used Python, Seaborn, and Matplotlib for data visualization, and applied ETL techniques to clean and process the data. Some statistical methods were used to better understand sales patterns during this period.

Data analyzed

The stock was analyzed by sports categories, subcategories, and their respective sales to understand the overall average behavior.

1. Top ten best selling product

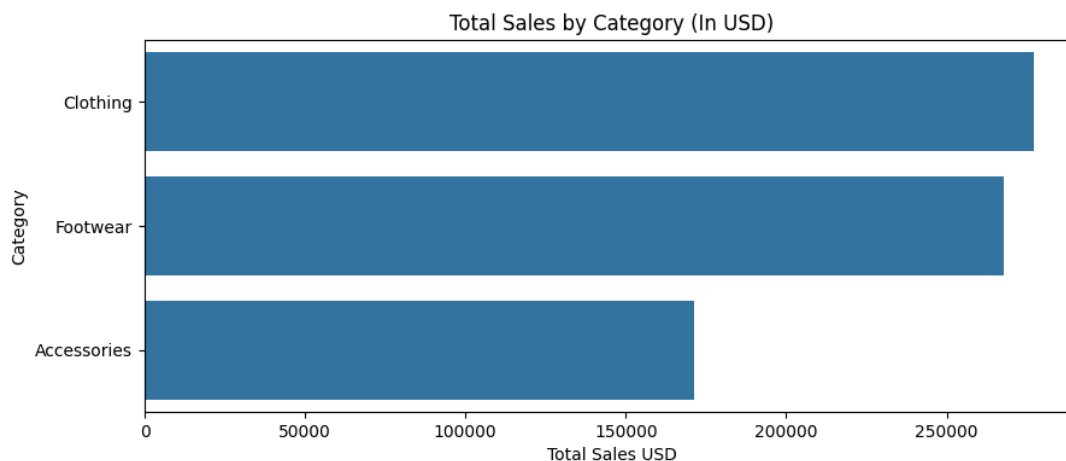


The chart displays the top 10 best-selling products based on total sales in USD. The Game T-shirt stands out as the highest-selling product, significantly outperforming other items. Among the products, there are four types of T-shirts: **Game T-shirt**, **Charge T-shirt**, **Computer T-shirt**, and **Top T-shirt**. These variations could be targeted at different activities or customer groups, with the **Game T-shirt** clearly being the most popular.

While we see a strong preference for the Game T-shirt, other products like the **Word Shoes** and **Pick Hat** also perform well. The presence of different hats suggests a range of popular accessories. The variety in products may indicate diverse customer interests, potentially targeting different sports or demographics, although the exact sports categories are not specified in this data.

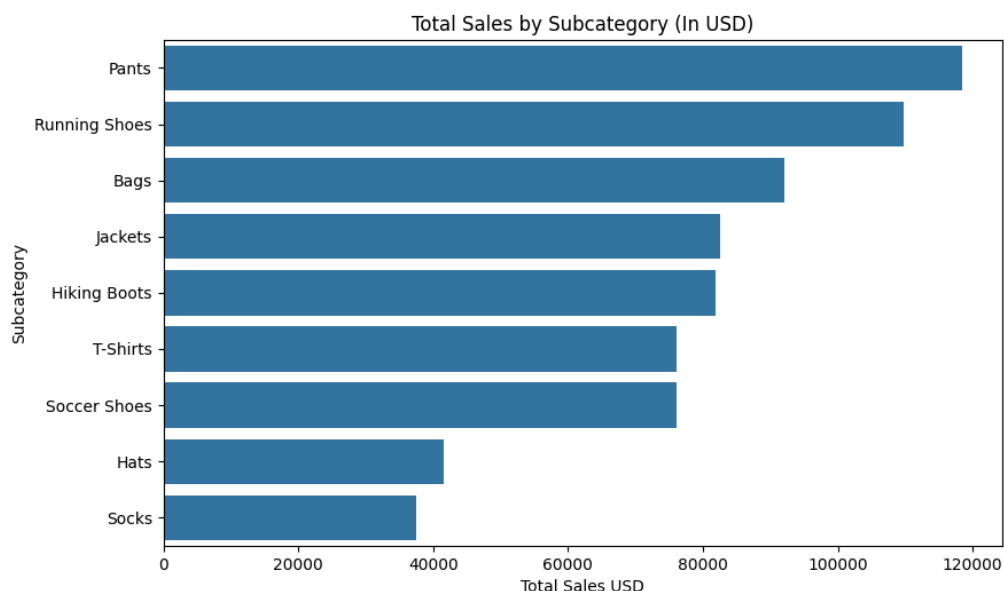
In conclusion, T-shirts and hats dominate the top-selling products, with potential differentiation across sports or casual wear.

2. Categories of products



In this category, the highest sales were in Clothing, likely due to different preferences for each sport, whether for casual wear or sportswear, with sales exceeding 250,000 USD. The next category is Footwear, with sales around 255,000 USD, but not as high as expected. This could be because people often use the same footwear for different sports. The last category is Accessories, with sales around 150,000 USD. While every sport requires accessories, it seems that customers may not be fully aware of this.

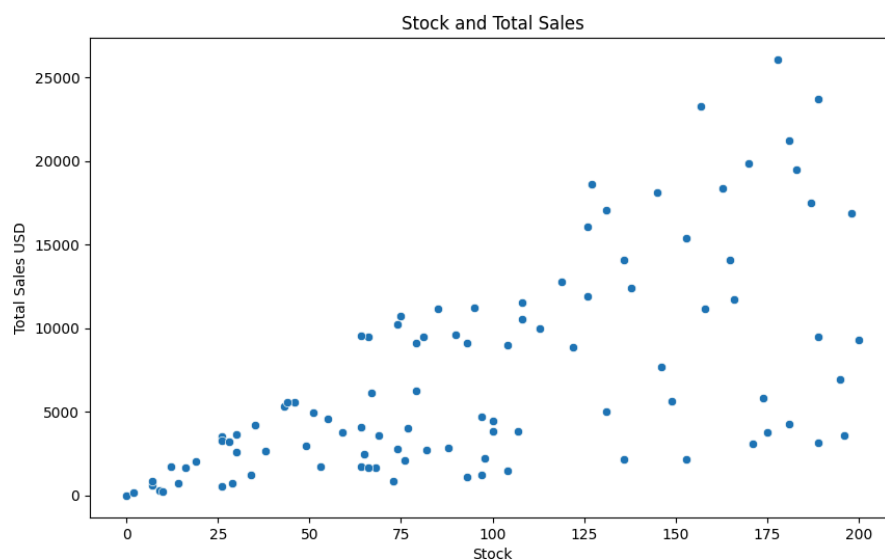
3. Sales by category



In the general analysis of sales by subcategory, we see that the category with the most sales is Pants, with 120,000 USD, followed by Running Shoes, which have more than 100,000 USD. Other subcategories like Bags, Jackets, and Hiking Boots have sales below 100,000 USD.

Even though in the chart of the top 10 best-selling products, the Game T-Shirt leads with 25,000 USD, this only shows the success of individual products in the t-shirt subcategory. However, this does not mean that the T-Shirts subcategory has more sales than other categories, like Pants. In fact, in the first chart, the "T-Shirts" subcategory has much lower sales compared to Pants.

4. Stock and Total sales



The relationship between stock and sales is weak, as sales don't always increase proportionally when stock increases. The data points are spread out, showing that having more stock doesn't always lead to more sales. After about 125 units in stock, sales tend to stay the same. To understand this better, you can check the correlation charts by category and subcategory in the Product Data folder. This suggests that more stock doesn't always mean more sales.

RECOMMENDATIONS

1. Sales change depending on the number of units, categories, and subcategories, which means we need a deeper cost analysis to understand profits, losses, and investments, and to optimize stock. This cost analysis should include the acquisition costs of each product, the storage costs incurred while holding inventory, and the opportunity costs of unsold items. Additionally, it is important to consider potential losses from products that may not sell or need to be discounted. By comparing sales revenue with the total costs of maintaining stock, we can identify areas to improve profitability and better manage inventory levels.
2. As we saw in the analysis, product sales go up or down based on how customers use them. Some people prefer comfortable clothes for working, studying, being at home, or doing sports. A demographic study would help marketing to segment by these specific uses.
3. For accessories, even though sales are lower, it would be useful to create promotions or sports combos that match what users need, whether for sports or daily use. Marketing could also run campaigns to teach amateur athletes about the importance of having the right accessories for each activity.

Conclusion

The analysis of customer behavior and sales data from the sports store shows several key opportunities to improve both inventory management and marketing strategies. In particular, we observe that:

- **Product category preferences:** Clothing and footwear lead in sales, which suggests that optimizing inventory in these categories should be a priority. However, it's important to also focus on accessories, where promotions and bundles could help increase sales.
- **Age and sport segmentation:** Customer preferences change significantly based on their age and the sport they practice. This gives the store a chance to personalize product offers and create more attractive experiences, which could improve customer loyalty and increase sales.
- **Relationship between stock and sales:** The analysis shows that increasing stock doesn't always result in higher sales. This highlights the need for a detailed cost analysis to make better decisions on inventory investment.

