Report Stage: Qualitative Market Analysis and Evaluation of the Marketing Strategy to be Applied to a Potential E-Commerce, through the Construction of KPIs, in the Roller Sports Sector

1. Objectives

The purpose of this work, carried out between February and May 2025, was to conduct an initial qualitative market analysis through competitor observation and to develop consistent evaluation metrics to support a marketing strategy, with a view to the potential creation of an e-commerce activity in the roller sports sector. The reference point for the definition and categorization of roller sports, as well as for the related equipment, was the official website of the Italian Roller Sports Federation (fisr.it).

The evaluation metrics for this qualitative market analysis in the roller sports field were based on KPIs (Key Performance Indicators), themselves of a qualitative/descriptive nature, since the reference information for building these metrics was derived from existing online sales websites, with data extracted through web scraping.

The choice of sources was therefore shaped according to the KPIs to be developed and, given that the available data only concerned information published on the respective websites (hence no sales volume data was accessible to create predictive KPIs), the starting generic metrics mainly focused on capturing price variations, discount variations, product ratings (in score format), number of reviews, and most sought-after brands.
On this basis, the sources deemed suitable—i.e., those providing at least product-related data on price, score, reviews, and brand name, along with a structure not overly complex or heavily protected by anti-scraping measures, and a product catalog sufficiently broad both in overall volume and in product categories—turned out to be two: one Italian and one English, namely Decathlon Italy and RollerSnakes UK**.** Both sources clearly offered the same key information to ensure continuity and consistency in the collected data.

The choice of Decathlon was further supported by the accessible nature of its brand reputation and by the fact that it has always presented itself as “The sports store,” a reference point for anyone wishing to approach a specific sport, thanks also to its vast range of sports categories, catering to beginners as well as experts and enthusiasts. RollerSnakes UK, instead, was chosen as the foreign reference, since among the main international roller sites, it proved the most suitable given the premises mentioned above and the available computational resources.

2. Development

The first step of this work was to analyze the structure of the most suitable websites for the project’s purpose, using developer tools, some already integrated in browsers such as Safari and Chrome. These tools allow the inspection of site elements in their HTML form, as well as the observation of network calls made by the site, which can be useful for inferring possible API calls and thereby retrieving information more directly. In this way, the HTML tags necessary to develop the respective scripts for Decathlon and RollerSnakes were retrieved, such as “div.page-container” and

“div.collection__main,” which identify the product-container elements for each page of the site, respectively for the two websites.

The method employed for both was a combination of **Selenium**, which enabled the execution of a webdriver navigating the site, together with **BeautifulSoup** for HTML parsing of the pages and the actual data extraction, in a local working environment, specifically VS Code.
Both scripts follow the same logic: first, access to the starting page is carried out, then the product links from all pages of each product category are collected into a list—as in the case of Decathlon, which had 5 pages, or directly from the starting page, as for RollerSnakes. Once all pages were exhausted and the product links collected, the actual data extraction began, according to the HTML tags of the respective websites; finally, the extracted data was saved into a JSON file, within a folder stored locally by the script itself.

After the scraping period—which was carried out weekly (to avoid possible IP blocking by the websites involved) between the beginning of April and mid-May 2025—the JSON files were uploaded to **MongoDB**, which was particularly suitable for the project’s goals and the subsequent development of queries for KPI construction. In addition, **MongoDB Compass** was used as the database interface, allowing an aggregated visualization of the data, since locally they were separate files, and enabling the verification of queries later developed in a Python notebook on VS Code, through the import of the files already aggregated on MongoDB Compass with the **PyMongo** library.

For KPI development, however, Decathlon’s data turned out—after an initial parsing of the extracted data—to be more significant both qualitatively and quantitatively. For this reason, the final notebook created to conclude the stage contains observations based on Decathlon’s data. This choice was also supported by the fact that the KPIs maintained a general and flexible character, so that they could be applied in the future to other datasets, and therefore, from a purely conceptual standpoint, this did not hinder the success of the work.

Thus, the final project pipeline was:

- **Step 1**: analysis of website structures suitable for the purpose and available resources
- **Step 2**: development of scripts and fine-tuning of scrapers
- **Step 3**: weekly data extraction and saving in JSON file format
- **Step 4**: file storage on MongoDB and aggregation through MongoDB Compass
- **Step 5**: creation of queries for significant KPIs and graphical representations

3. Results / Conclusions

Before presenting the most important KPIs developed for this internship project, it must be noted that the data collection period—about two months (ending one week before the end of the internship to allow KPI development)—was not long enough to capture more significant and appreciable variations related to prices, discounts, and other indicators.
For this reason, in August, an additional data extraction was carried out to try to observe variations in the main variables considered, namely price, score, and reviews.

This time, however, it was not possible to collect data from Decathlon due to a technical issue caused by a change in the website’s structure and a subsequent unsuccessful attempt to retrieve data, probably hindered by too many requests to the site within a few hours.

For this reason, **RollerSnakes UK** data was used instead, since its scraper still worked perfectly, and it proved useful for developing the marketing strategy in line with the project's objectives.

The most important KPIs developed are as follows:

- **Average Price per Month and Category**
- **Average Price Trend per Month and Category**
- **Average Weekly Discount Trend (also shown by category)**
- **Average Discount (%) by Category (over the period)**
- **Top Brand by Reviews**
- **Average Reviews by Discount Range**
- **Most Stable Products by Score**
- **Product Availability by Brand**
- **Correlation between Discount and Number of Reviews**

Although RollerSnakes data did not include discount information (unlike Decathlon), it nonetheless provided, thanks to the August extraction, useful insights to support a sales strategy. This was made evident by the KPI **"Average Price per Month and Category,"** which, when represented graphically, shows a clear seasonality of prices that, for some product categories, proved to be lower on average.

Despite the limited number of observations to obtain more appreciable and strategically useful insights, it is believed that—at least from a conceptual and interpretative standpoint—the developed KPIs can still be useful and potentially employed for an initial qualitative evaluation of this sports sector's market. In the context of this project, they could become more meaningful over a longer observation period and therefore with a larger data volume than the one collected.