University of St. Gallen

School of Management, Economics, Law, Social Sciences

International Affairs and Computer Science (HSG)

Chrono24 Luxury Watches Web Scraping Program

Group Project Documentation

Submitted by:

Boris Szelcsanyi (19-606-656)

Marco Beverari (22-620-157)

Noah Kreuzbichler (19-606-334)

Samir Trpeza (17-602-426)

Stefan Unterganschnigg (19-606-946)

Skills: Programming with Advanced Computer Languages (7,789 | 8,789)

University of St. Gallen

Mario Silic

22.12.2023

Contents

Table of Figures

1 What does the program do?

The Chrono24 Luxury Watches web scraping program is crafted on Python to perform a two-step process.

1. First, it engages in web scraping by extracting data from the initial page of Chrono24 using a specified reference number or model. This entails gathering information about watch listings associated with the provided reference number/ model.
2. Following the web scraping, the program then undertakes a statistical analysis of the collected data. This analysis encompasses:

    - Location Parameters: Calculating statistical measures related to the data's central tendency, such as minimum, maximum, and mean values.
    - Distribution Parameters: Examining the distribution of the data by assessing measures such as standard deviation (std) and quartiles. This analysis aids in understanding both the spread and central tendencies of the gathered data.

In essence, the program combines web scraping and statistical analysis, providing a comprehensive understanding of watchlisting-data from Chrono24, including descriptive features and distributional characteristics.

2 How does the Program work?

The Chrono24 Luxury Watches web scraping program uses several libraries and frameworks to automate the extraction of watch data from the Chrono24 website. Employed key technologies include Selenium WebDriver for automated web browsing, BeautifulSoup for HTML parsing, and Pandas for data manipulation and CSV file handling.

The program is structured around modular classes with specific tasks:

- **Header:** Manages HTTP headers to mimic real browser requests, reducing the risk of bot detection.
- **Driver:** Configures and controls the Selenium WebDriver with the ability to perform headless browsing.
- **Chrono:** Takes over from Driver and is tailored to interact specifically with the Chrono24 website. It handles search queries, pagination, data retrieval, and parsing.
- **Menu:** Provides a console-based user interface for entering search criteria, selecting data retrieval options, and deciding on data export.

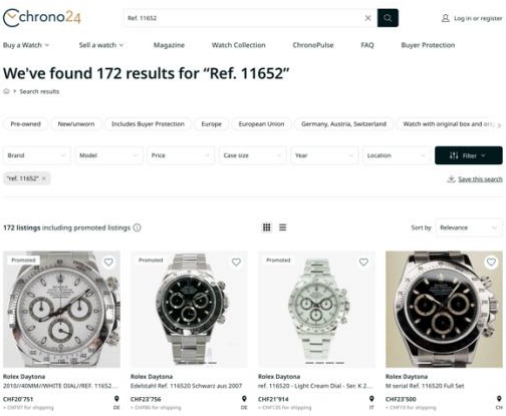

Figure 1: Chrono24 Reference Number Search

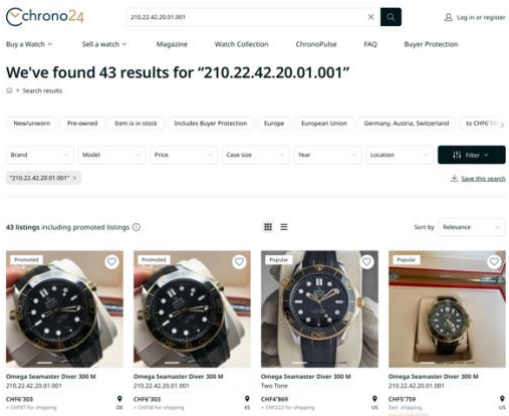

Figure 2: Chrono24 Serial Number Search

The main function acts as an entry point and organises the entire workflow. The **user is prompted for input** to enter search criteria, specifying either the model or reference number of watches. Based on the user input, the program starts the scraping process using the Chrono class to interact with the website, **retrieve watch data**, and parse the HTML content.

It also offers options to **save the collected data for export** in a CSV file using the Pandas library. The **process is repeated in a loop**, allowing the user to perform multiple searches until they choose to exit.

This tool simplifies the process of collecting detailed watch data from the Chrono24 online marketplace. It is user-friendly and versatile, utilising web scrapping and automation. Users can conveniently access and analyse data from the luxury watch market.

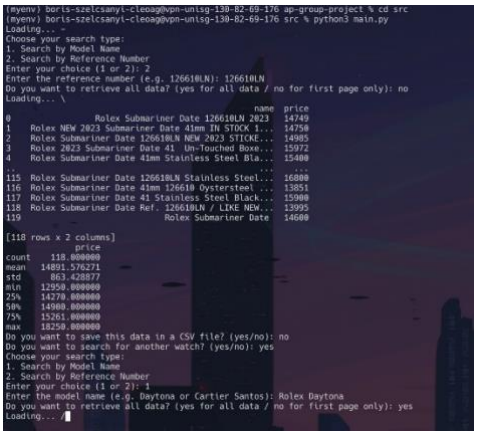

Figure 3: Screenshot of Output

Key Stats	
Model	**Rolex Daytona**
Count	5'298
Mean	47'369
Std	37'204
Min	14'850
1 Quartile	25'700
Median	40'000
3 Quartile	51'010
Max	450'000

Figure 4: Output Organized

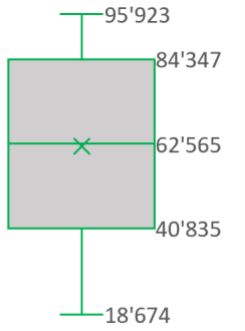

Figure 5: Example of Illustrated Output

3 What it can be used for

The Chrono24 Luxury Watches web scraping program can be used in multiple ways, offering users a spectrum of functionalities. Below are some key scenarios highlighting its practicability:

1. **Market Analysis and Price Trend Identification**

This program gives users a deep insight into what is happening in the luxury watch market. It gathers information from Chrono24 and shows patterns, ups and downs, and generally how prices have changed over time. This information is super important for experts who study the watch market, people who collect watches, and sellers who want to understand what buyers think and are willing to pay. It helps them figure out how much a watch is worth and make smart choices about pricing.

2. **Research and Decision Support for Buyers and Collectors**

This code is great for people thinking about buying a watch or who are just real watch enthusiasts. It helps them do thorough research before they decide to purchase. By putting together all this data, it shows what models are out there, how much they cost, and allows the user to develop then a deeper understanding of different features and conditions they have. This way, buyers can feel confident choosing a watch that matches their expectations and fits in their budget.

3. **Inventory and Investment Management for Sellers**

This program is a big help for sellers, as it helps them figure out what the market dynamics are. It allows them to manage their stock and helps them to spot chances of investing in new watches that could be popular in the future.

4. **Customized Analytics and Reporting**

The program's capability to extract, organize, and process data from the world's largest watch market site (Chrono24) allows users to create customized inputs for reports. These reports

can be customized based on the preferences of the user, offering insights into various metrics like brand popularity, model performance, regional market differences, and more.

5. Identified Potential: Real-time Monitoring and Alert Systems

To think this further, we thought that this program could potentially be expanded to including a real-time monitoring and alert system, which could notify users about significant market shifts/trends, changes in prices, or the availability of a specific watch model.

Declaration of Authorship

"I hereby declare

- that I have written this thesis without any help from others and without the use of documents and aids other than those stated above;
- that I have mentioned all the sources used and that I have cited them correctly according to established academic citation rules;
- that I have acquired any immaterial rights to materials I may have used such as images or graphs, or that I have produced such materials myself;
- that the topic or parts of it are not already the object of any work or examination of another course unless this has been explicitly agreed on with the faculty member in advance and is referred to in the thesis;
- that I will not pass on copies of this work to third parties or publish them without the University's written consent if a direct connection can be established with the University of St.Gallen or its faculty members;
- that I am aware that my work can be electronically checked for plagiarism and that I hereby grant the University of St.Gallen copyright in accordance with the Examination Regulations in so far as this is required for administrative action;
- that I am aware that the University will prosecute any infringement of this declaration of authorship and, in particular, the employment of a ghostwriter, and that any such infringement may result in disciplinary and criminal consequences which may result in my expulsion from the University or my being stripped of my degree."

.....

By submitting this academic term paper, I confirm through my conclusive action that I am submitting the Declaration of Authorship, that I have read and understood it, and that it is true.