**Project Title: Book Price Comparison Tool**

**Project Description**

The **Book Price Comparison Tool** is a web scraping project designed to help users to find the best prices for books across multiple online bookstores. By inputting a book's **ISBN** or **title**, the tool scrapes real-time price data from several websites and presents a comparative view, allowing users to make informed purchasing decisions.

**Target Websites for Scraping**

The tool will extract data from the following e-commerce platforms:

1. **Book Depository (**[**https://www.bookdepository.com**](https://www.bookdepository.com)**)**
2. **AbeBooks (**[**https://www.abebooks.com**](https://www.abebooks.com)**)**
3. **Barnes & Noble** **(** [**https://www.barnesandnoble.com**](https://www.barnesandnoble.com)**)**

**Types of Data to Be Scraped**

For each book search (by title or ISBN), the scraper will extract:

Comprehensive metadata for books listed on the website. This encompasses the book title, listed price, customer rating , stock availability status and the associated category.

**Overall Goal of the Project**

The primary goal of this project is to develop an intelligent and user-friendly system that:

* Aggregates book prices from various sources in real-time.
* Helps users **save money** by finding the best deals across stores.
* Provides a **unified and clean interface** for searching books by ISBN or title.

**Project Plan :**

Data Extraction:

Utilize Python libraries such as requests and BeautifulSoup to automate the scraping of book-related data from multiple pages of the target websites.

Data Cleaning & Regular Expressions:

Perform data cleaning to eliminate unwanted characters (e.g., currency symbols) from price values. Leverage Regular Expressions (Regex) to convert qualitative ratings into structured numerical format, and to standardize category labels and availability status for consistency.

Data Analysis:

Conduct descriptive statistical analysis to compute metrics such as average price and average rating per book category. Investigate potential correlations between variables, such as pricing and customer ratings, and explore availability trends.

Data Visualization:

Develop clear and informative visualizations-including bar charts and heatmaps-to highlight key insights, allowing for intuitive interpretation of trends and relationships within the dataset.

Data Storage:

Store the final processed dataset in a relational database (Microsoft SQL Server Studio), enabling efficient querying and potential integration with data-driven applications or dashboards.

Bonus Task:

Build an interactive web app using Streamlit to explore book statistics and search within categories and to enhance the user experience.