**Web Scraping and Data Encryption – Report**

**1. Description of the Website Scraped**

The website used for scraping is Books to Scrape, a publicly available site designed for practicing web scraping. It contains multiple pages of book listings with structured HTML, making it ideal for extracting metadata such as titles, prices, ratings, and availability.

**2. Fields Extracted**

From each book listing, the following fields were extracted:

* Title: The name of the book
* Price: The listed price in GBP (£)
* Rating: Star rating (e.g., One, Two, Three)
* Availability: Stock status (e.g., In stock)

The data was stored in a pandas DataFrame and saved as a CSV file (books\_data.csv).

**3. Security Approach Used**

Encryption using Fernet (cryptography library):

* The scraped CSV file was encrypted using symmetric encryption.
* A secure key was generated and stored in encryption.key.
* The encrypted output was saved as books\_data\_encrypted.csv.
* A separate script (decryption.py) was created to decrypt the file when needed.

**4. Challenges Faced**

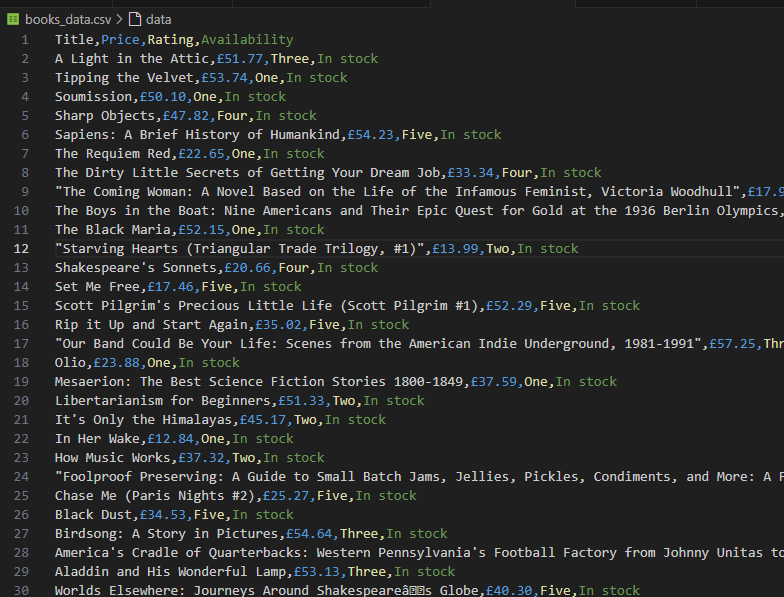
* Encoding issues: The currency symbol (£) appeared incorrectly as “Â£” due to encoding mismatch. This was resolved using string replacement.
* Python setup: Initial issues with pip not being recognized were resolved by installing Python correctly and using python -m pip.
* Data cleaning: Minor formatting adjustments were made to ensure readability (e.g., capitalizing ratings, removing extra whitespace).

**5. Ethical Considerations**

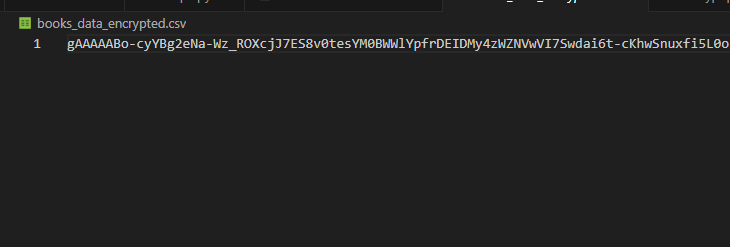
* Scraped only 3 pages to avoid excessive requests
* Used a User-Agent header to mimic browser behavior
* Added a delay (time.sleep(1)) between requests
* Verified that the site allows scraping and does not block bots

**6. Screenshots**

**1.Book data**

****

**2. book data decrypted**

****