Web Scraping British Airways Reviews

What is Web Scraping?

Web scraping is the automated process of extracting data from websites. It is commonly used to collect customer reviews, pricing information, and market trends. This project focuses on scraping British Airways reviews from AirlineQuality.com, a well-known airline review platform.

Steps Taken in This Project:

1. ### Importing Necessary Libraries
    - **requests** – Fetches the HTML content of the AirlineQuality.com review page.
    - **BeautifulSoup** – Parses and extracts specific elements from the HTML.
    - **pandas** – Structures and stores the extracted data for further analysis.

2. ### Identifying the Target Web Page

    The project targets the British Airways review section of AirlineQuality.com:
    🔗 **URL Scraped:**
    https://www.airlinequality.com/airline-reviews/british-airways/page/2/?sortby=post_date%3ADesc&pagesize=1000

    This page contains **1,000 customer reviews per request**, making it ideal for large-scale data extraction.

3. ### Sending a Request and Parsing the HTML

    The script:

    ✅ Uses requests.get() to fetch the webpage content.

    ✅ Parses the response using BeautifulSoup.

    ✅ Locates review-specific elements within the page's structure

4. Extracting Key Review Information

- Passenger Information
  - **Passenger_Name – The name of the reviewer.**
  - **Nationality – The country of origin of the reviewer.**

- Review Metadata
  - **date_of_review – The date when the review was posted.**
  - **Title – The headline/title of the review.**
  - **Review – The full text of the review.**
  - **Verification – Whether the review is verified or not.**

- Flight Details
  - **Aircraft – The aircraft model used for the flight (if mentioned).**
  - **Type Of Traveller – The category of traveler (e.g., solo, couple, family).**
  - **Seat Type – The class of the seat (e.g., Economy, Business, First Class).**
  - **Route – The flight route taken by the reviewer.**
  - **Date Flown – The date when the flight took place.**

- Service & Experience Ratings

  (These are rated individually, usually **on a scale of 1-5**)

  - **Seat Comfort – How comfortable the seating was.**
  - **Cabin Staff Service – The service quality of the flight attendants.**
  - **Food & Beverages – The quality of in-flight meals and drinks.**
  - **Inflight Entertainment – The availability and quality of movies, music, etc.**

- **Ground Service – The quality of service at the airport (check-in, boarding).**
- **Wifi & Connectivity – The quality of onboard internet services.**
- **Value For Money – Whether the passenger felt the service was worth the cost.**

- Overall Ratings & Recommendation
  - **Recommended – Whether the passenger recommends British Airways (Yes/No).**
  - **OverAll_Rating – The overall rating given to the airline (out of 10).**

5. Saving the Data for Further Analysis

Once extracted, the reviews are organized in a Pandas DataFrame and The final dataset is saved as:

📁 **Filename:** `british_airways_reviews.csv`

This allows for:

- Sentiment analysis (e.g., positive vs. negative reviews).
- Customer experience trend tracking.
- Service improvement insights for British Airways.