Hotel_Data_Scrapping - Booking.com

Project Overview

This project is a web scraper that extracts hotel information from Booking.com for New Delhi. The script collects comprehensive hotel data including pricing, ratings, reviews, and location details, then saves it in a clean, formatted CSV file.

Features

· **Data Extraction:** Scrapes hotel names, locations, prices, ratings, scores, and review counts
· **Price Formatting:** Converts currency to standardized "RS" format with clean numerical values
· **Review Processing:** Extracts and cleans review counts to show only numerical data
· **Data Cleaning:** Removes extra spaces and normalizes text formatting across all fields
· **CSV Export:** Saves structured data in a well-formatted CSV file

Technical Stack

· Python with BeautifulSoup for web scraping
· Requests for HTTP handling
· CSV module for data export
· Regex for data cleaning and formatting

Output Format

The scraper generates a CSV file with the following columns:

· Hotel Name
· Locality
· Price (formatted as "RS X,XXX")
· Rating
· Score
· Reviews (numerical count only)
· Link

Use Cases

· Market research for hotel pricing in New Delhi
· Competitive analysis for hospitality businesses
· Data collection for travel applications
· Academic research in tourism and pricing trends