Justin Wang and Austin Lee

jwang592, alee235

CS 183 - Ali Davinian

31 October 2021

CS183 Project Proposal

Problem Description

Oftentimes, we come across large amounts of data online and want a faster way to process and reformat the data. One example would be finding the right hotel among the many hotel listings on booking.com. The web interface might work fine for the average user, but having to perform multiple manual searches to find the right fit ends up being very inefficient and time-wasting. The problem becomes even more apparent as the number of people searching increases, adding to the total time wasted.

Proposed Solution

Our solution will use a web scraping script written in bash, which can automatically scrape the data on booking.com and pipe it into a python application. The python application would be able to store the data into database format using the sqlite3 library, from which searching would be faster and more efficient. The application would also output the data in tabular format using the pandas library. If time permits, we also plan to implement analytics features and display them using a visualization library such as Seaborn.

This project serves as a practical introduction to automation using bash scripts, which will be very useful for future system administration automation tasks. The visualization portion is also relevant to system administration as many unix systems revolve around graphing trends and finding useful insights in massive amounts of data.

Expected Deliverables

- bash script that will scrape booking.com
- python application
    - stores piped data into a database
    - outputs database in tabular format
    - displays analytics visualization

[Booking.com : Hotels in Rishīkesh . Book your hotel now!](#)

[Scraping Hotel Listings From Booking.com With Python and BeautifulSoup | by Mohan Ganesan | The Startup | Medium](#)