Homework-4-2

March 31, 2017

1   Scraping Hotel Ratings on Tripadvisor

To be completed INDIVIDUALLY and due on April 10 at 7pm.

In this homework we will practice web scraping. Let's get some basic information for each hotel in Boston. On each hotel page, scrape the Traverler ratings. **(15 pts)**
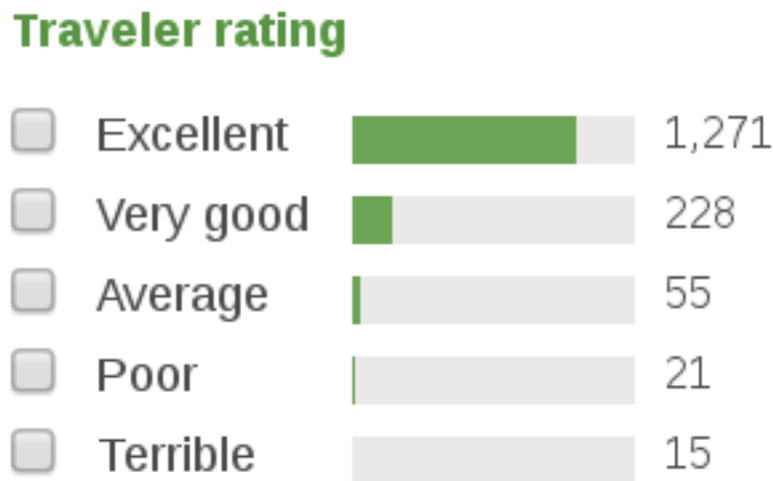

Figure 1: Information to be scraped

Save the data in "traveler_ratings.csv" in the following format:

hotel_name, rating, count

To receive credit, you must commit traveler_ratings.csv to Github.

In [ ]:

Next, scrape all the reviews of each hotel for the star ratings of the following attributes: Value, Location, Sleep Quality, Rooms, Cleanliness, Service. Note that some reviews may not have attribute ratings and some may only have some of the attributes. **(25 pts)**

Save the data in "attribute_ratings.csv" in the following format:

hotel_name, review_id, attribute, star_value

To receive credit, you must commit attribute_ratings.csv to Github.

In [ ]:

"Extraordinary Service and Views...but no[t] the Four Seasons for the price..."

◉◉◉◉○ Reviewed July 20, 2010

Management has evidently done a superb job of tea[m...] to offer exemplary, friendly, efficient service in this h[...] waterfront. Views to the harbor with its yachts, sch[...] were beautiful... Shannon, at the reception desk, wa[...] us and Concierge Rob Fournier was ever-ready to a[...] night room service delivery was handled with gracio[...] couldn't do enough for us and every request was me[...] heard every skipping child and closing door in the ha[...] relatively small room, and I preferred the Back Bay [...] Public Garden and the serene, luxurious comfort at [...] for the same price. My husband enjoyed the boats s[...] disagreed, so we will have something to discuss wh[...] Boston!

Stayed July 2010, traveled with family

◉◉◉○○ Value   ◉◉○○○ Rooms
◉◉◉○○ Location   ◉◉◉◉◉ Cleanli[ness]
◉◉○○○ Sleep Quality   ◉◉◉◉◉ Service

Figure 2: Information to be scraped