Homework-4-2

March 31, 2017

1 Scraping Hotel Ratings on Tripadvisor

To be completed INDIVIDUALLY and due on April 7 at 3pm.

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)**

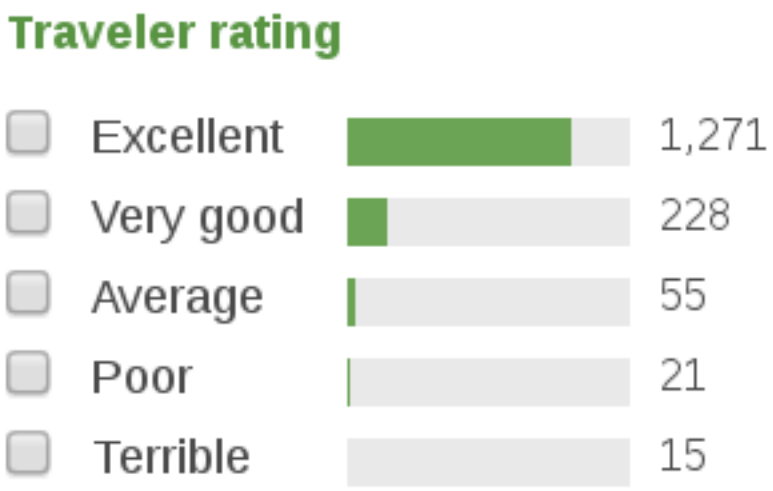

Figure 1: Information to be scraped

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 [ ]:

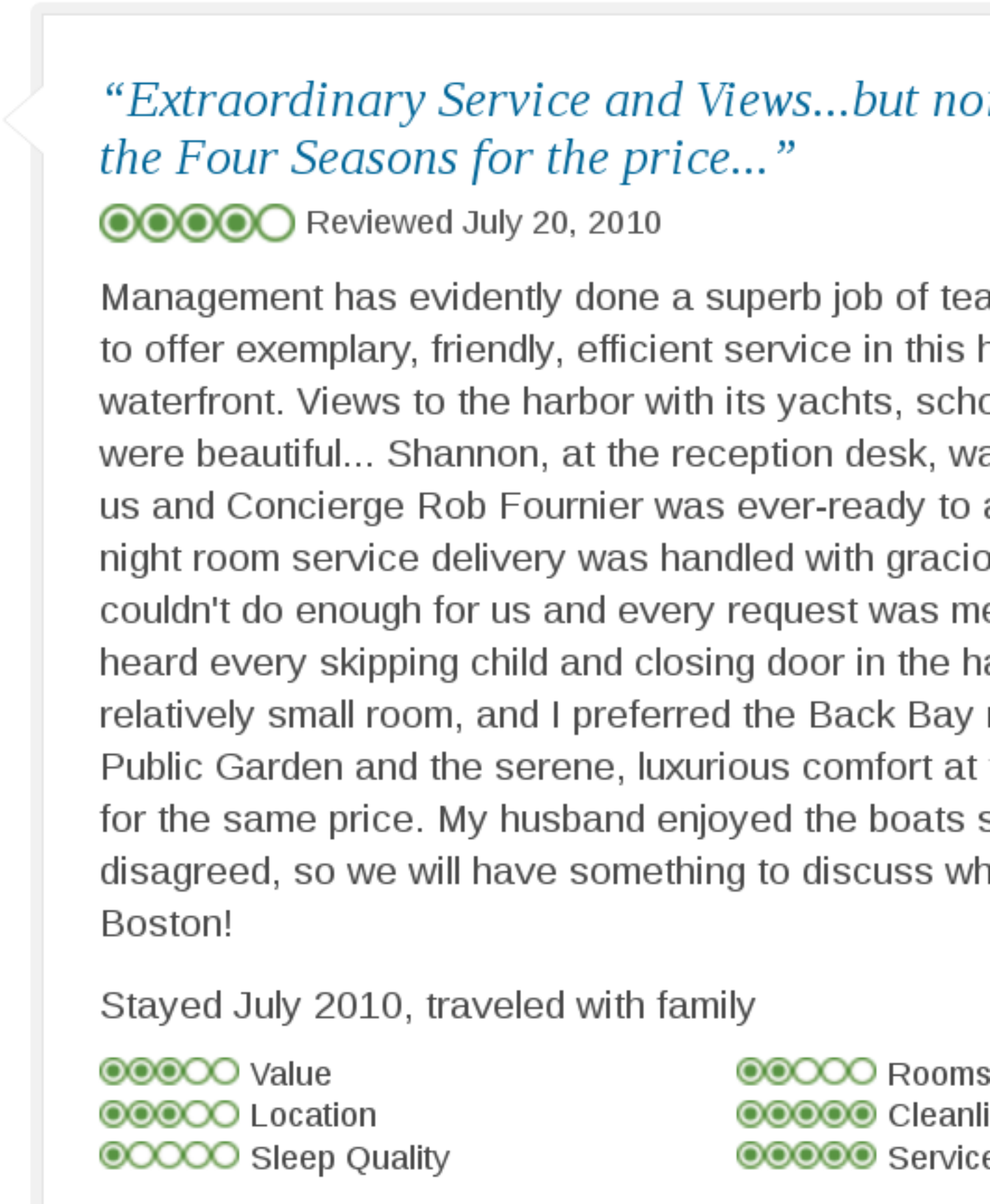

Figure 2: Information to be scraped