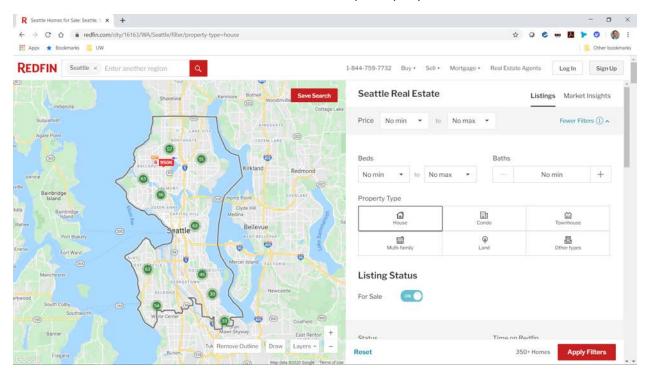
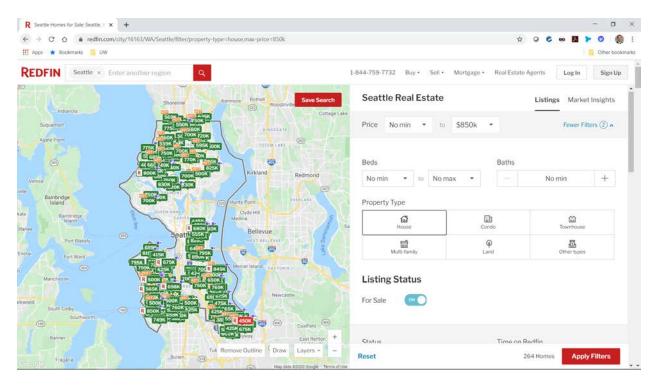
Redfin Data Download Guide

Corey Christopherson DATA515 – Final Project

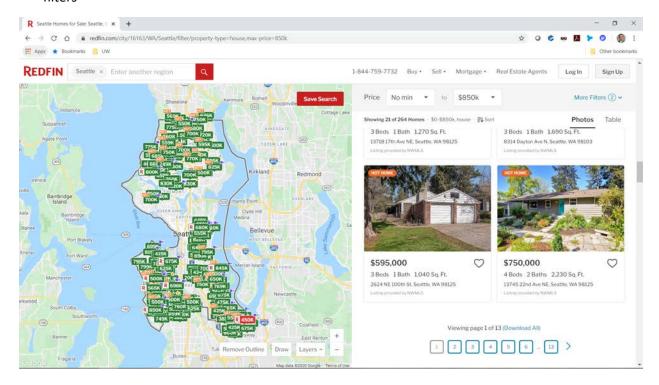
- 1. Go to www.redfin.com and enter Seattle into the search box in the upper left. Note that you can also search by zip, school, neighborhood, and many other divisions
- 2. Click on the More Filters text on the right side (shown as Fewer Filters in the image below) and select the House for the Property Type option.
- 3. Note in the bottom right it displays the number of results as 350+. The current maximum output is 350 rows, so this is how we know we need to break up the query.



4. Price seems to be the best way to break up the query into smaller chunks without the risk of losing data. In the image below I've set a max price of \$850k which drops our total number of results to 264. Note that the URL for each page is a simple comma separated list of filters. This will make it very easy to parameterize in our code.



- 5. Click on the red Apply Filters button to return to the main page. Note that all of the matches we generated are now displayed in a pane view on the right side.
- 6. Scroll down to the bottom of the listing pictures to see the page navigation options to show the next group of listings. Click on the Download All link to download all of the data generated from our filters



I was exploring this final page a bit and noticed that if you copy the link when hovering over the button you extract the exact API call used to generate the CSV file as shown below. Now we don't need to worry about using curl or beautifulsoup to scrape the data and can just go straight to the open API.

https://www.redfin.com/stingray/api/gis-csv?al=1&market=seattle&max_price=850000&min_stories=1&num_homes=350&ord=redfin-recommended-asc&page_number=1®ion_id=16163®ion_type=6&sf=1,2,3,5,6,7&status=1&uipt=1&v=8

I took out the price filter and changed the max rows to 1000 (shown below) and was able to get back 500 rows and a row limit warning. This means we'll still need to break up the queries but won't need as many as we originally thought. Small wins.

https://www.redfin.com/stingray/api/gis-csv?al=1&market=seattle&min_stories=1&num_homes=1000&ord=redfin-recommended-asc&page_number=1®ion_id=16163®ion_type=6&sf=1,2,3,5,6,7&status=1&uipt=1&v=8

Finally, there isn't any terms of use or license information on this main page, but if you go to the Redfin Data Center (https://www.redfin.com/blog/data-center/) you can find the following list of guidelines.

Guidelines for Using this Data

You are welcome to use this data for your own purposes, we just ask that you cite the source. Please include proper citation to Redfin for the first reference on a page, post, or article.

For example:

- According to Redfin, a national real estate brokerage, ...
- Data provided by Redfin, a national real estate brokerage.

Please note the proper way to spell, capitalize and describe Redfin. Thank you and please contact us if you have any questions regarding the use of our data or if you would like to speak to one of our housing market experts.

This should be sufficient for our not-for-profit academic purposes.