Zillow Housing Cost Estimator User Guide

TFP Group 16

Aiden MacKay, Bailin Kirk

Computing II Final Project

Contents

Overview

This Vaadin application allows users to sort through Zillow listings in a timely manner. This tool was designed with realtors and investors in mind. The program can streamline the process of finding listings within a client's certain requirements. Our project uses Zillow API scrapers and Geocoding APIs to determine the exact users' inputs and output the users' possible choices as well as the location.

Requirements

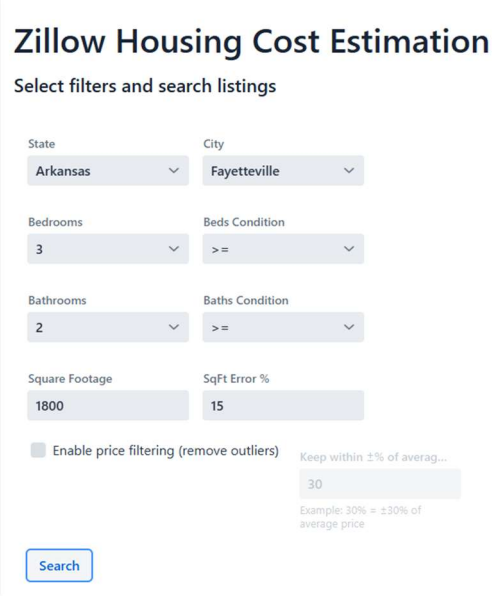
- An application that can run Java 17 or later
- An active internet connection
- Node.js installation (Expanded on within the document)

Using the User Interface

Page 1: Zillow User Inputs

When the user first launches the code, they will be greeted with a webpage. This first page they will see will be named *Zillow User Inputs*. This page has a place for the user to input the State, City, number of bathrooms, number of bedrooms, and square footage desired. There are also conditional boxes for the bedroom and bathroom. These conditional boxes allow the user to define If they want exactly the number of bedrooms/bathrooms inputted or if they want to also include listings with more than inputted. This allows the user to set a minimum requirement for their house of choice. The square footage input also has a square footage error box next to it. This box allows the users square footage to be an average value they would like to have. The error input allows the user to see listings within a set percentage of the given value. Each of these boxes allow the user to type or to choose from a dropdown menu to increase the user experience. Under the dropdown menus, there is a checkbox labeled *Enable Price Filtering*. This

checkbox allows the outliers to be excluded from the user's search. Sometimes Zillow listings may lack information or may have incorrect information. This checkmark allows for the user to filter the "bad" price data out of their search. The user can also define what range percentage they would like to see for the listings. After the user is done inputting their desired values, they can press the search button located at the bottom of the boxes. This will begin the process of searching and gathering all the information from Zillow that meets the users' specifications. The user also does not have to input values for every box as well. If there is flexibility in the clients' specifications, then the user may opt to leave a box blank. The only required information is the City and State name. If the rest are blank, the code will run to pull everything listed within the given location. Once the user presses the submit button, the information will populate underneath with the listings address, price, number of beds, number of baths, square footage, and the URL for easy lookup. The program will also output the average cost for the houses meeting the users' inputs. If there is no output, simply relax the square footage error box by 5% to make sure that there are listings available that may not meet your requirements. If there are no listings populating, keep in mind that there may not be any available within your requested City.



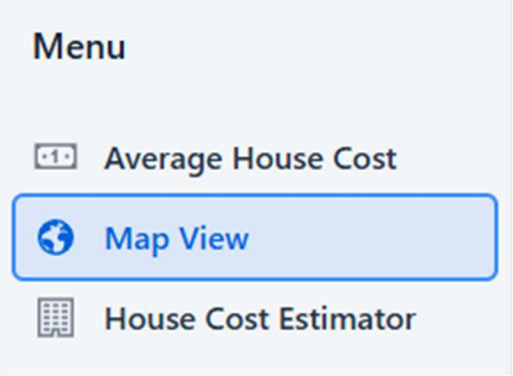
User input boxes as described in the writing.

Address	Price	Beds	Baths	SqFt	URL
675 S Neon Ave Fay...	370000	4	3	1892	https://www.zillow.c...
4435 W Canyon Ru...	372000	4	2	1788	https://www.zillow.c...
2225 N Mockingbir...	385000	3	2	1648	https://www.zillow.c...
1000 E Oaks Manor ...	475000	4	2	2112	https://www.zillow.c...
4287 W Divide Dr F...	365000	3	2	1615	https://www.zillow.c...
2641 Villa Blvd Faye...	340000	3	2	1442	https://www.zillow.c...
2172 W Million Ln F...	343500	3	3	2020	https://www.zillow.c...
581 N Sabine Pass R...	442900	4	3	2020	https://www.zillow.c...
2180 W Million Ln F...	342500	3	3	2012	https://www.zillow.c...

Average Price: $364,172.73

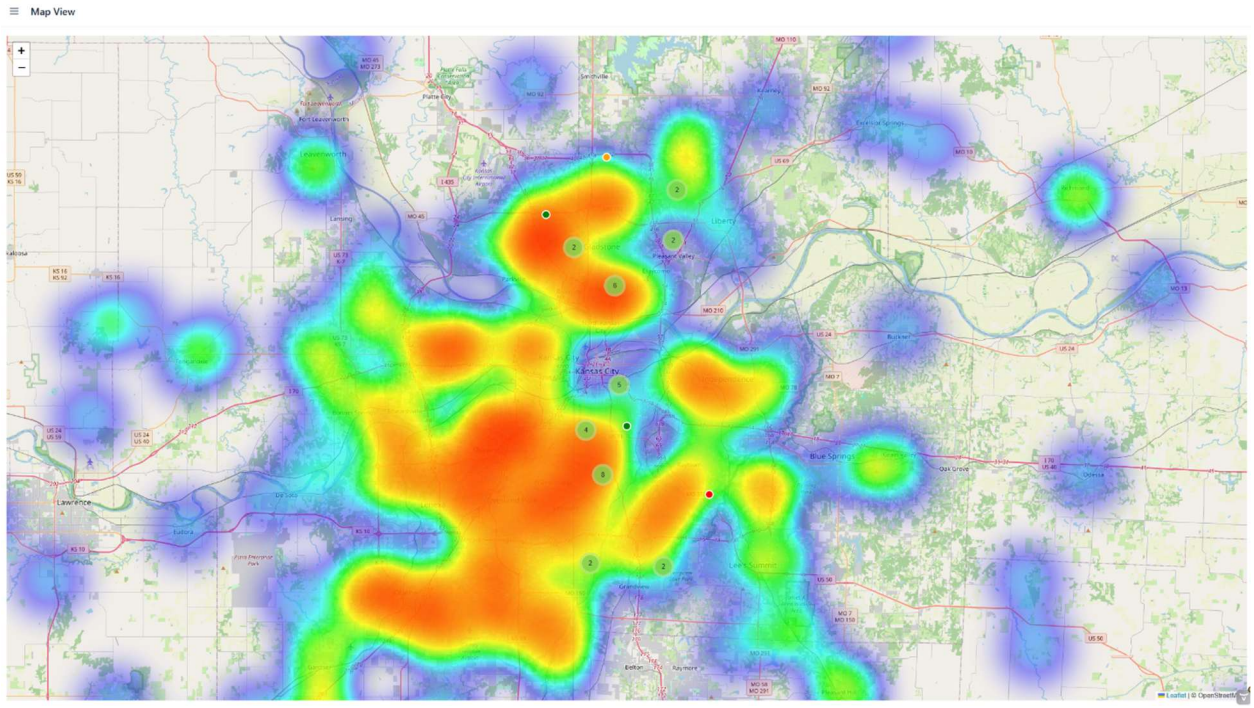
Example output for Fayetteville, AR. 3 Bedrooms or greater. 2 Bathrooms or greater. Square footage of 1800 with 15% error

Page 2: Map View



Screen capture of the labels within the menu.

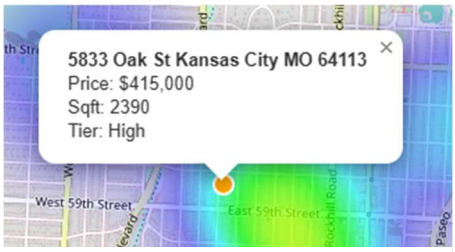
The Map View is our way of creating a more unified and user-friendly experience. It may be nice to have all the information listed out on a grid, but our map view takes the visual aspect to the next level. After the user is done on the Housing Cost Estimation page, they will navigate to the Map View via the menu in the upper left part of the screen. From here, they will click Map View which will take them directly to the map. When the user inputs anything on page one and presses the search button, the map is clustering the data and providing the user with a gradient heatmap controlled by the price per square foot. The code will run all of the houses through and place all of them on



User inputs Kansas City, MO on the Average House Cost page. Once they open the map, this is the image they are greeted with. The gradient heatmap is the average price per square foot. If red, it is greater than the national average. If blue, it is less than the national average.

the map according to their location listed on Zillow.com. The map will make the points either green, yellow, orange, or red. This signals Low tier, Middle tier, High tier, and Luxury tier respectively based on the price percentiles(30th,70th,85th) calculated from the data. This ranks the properties of all listings rather than fixed prices. When the user clicks on the dot, there will be a textbox displaying the listings Address, Price, Square footage (Sqft) and

the tier as well. This can be extremely useful for a realtor showing listings from around the city. It can be used in multiple different ways such as showing the wealthy areas with the heatmap, showing the distance to the actual center of the town. City services surrounding the listing such as fire stations, police stations, and hospitals. It can also show the user train tracks and airports which can be useful for people trying to avoid them. If there are no dots or clusters within the map, keep in mind that it must have an output from the Average House Cost page. If there is no output from that, there will be nothing to map. If there seems to be less listings on the map than what were labeled on the grid from Average House Cost, this is because the Geocoding failed. Geocoding is where the code takes an address and outputs a latitude and longitude for the house. This needs to be done for our locations to be properly mapped. The code was programmed to remove these listings that fail from both the map and the grip on Average House Cost. There may be a time when there will be a wrong listing on the grid. It may populate with it, but the geocoding will fail so it most definitely will not be listed on the map.



User clicks on the location of the dot. All of the information is broken down for the user.

Page 3: Housing Cost Predictor

The third page on the website is meant for people looking to build a house. This tool can be used to estimate the listing price within a city to show them if it would be worth the cost of investing. The user will navigate here via the menu again just like previously. The user will click on House Cost Estimator which will take them to a very simple page. This page includes dropdown boxes for State, City, Bed number, bath number and a textbox for square footage. The user must enter everything within these boxes. If a box is left empty, the code will not run as it does not have the sufficient information to provide the user with an accurate result. The information for the estimator is programmed to be run through a random forest model. This means that the information may be different every time that it is run. The website will output a user's estimated price as well as a range. This range is defined

Sample input of Fayetteville, Arkansas, 3 beds, 2 baths, and a square footage of 1650.

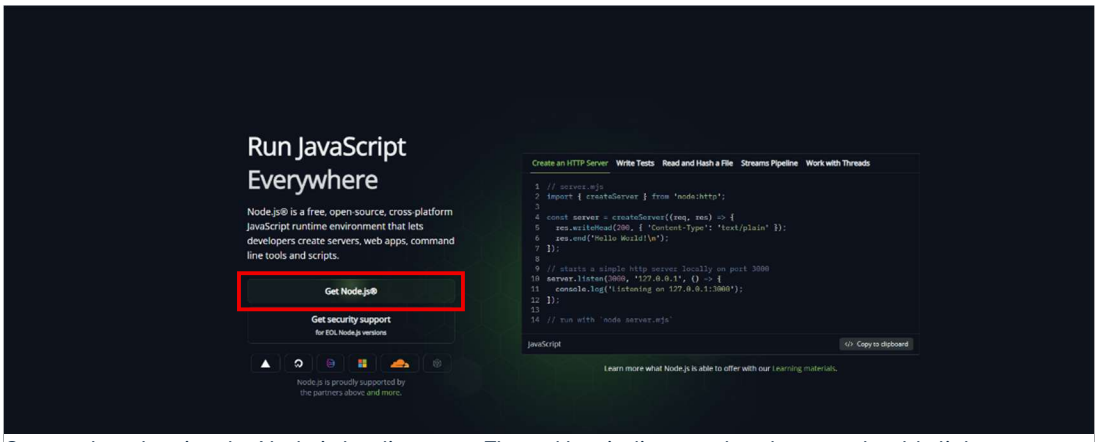
as 90% and 110% of the output estimated price. Some cities may not provide a fully accurate result to the estimator. This is not the codes or the data's fault. Some cities, such as New York City, only have so many listings. There are very few houses ever sold within the city limits of New York. This may create an unrealistic output for the user's input. This works better for cities which have an abundance of data such as Kansas City or Columbus. These medium/large cities provide the user with the best experience.

Installing node.js

Node.js is an integral part of our map view. This will be needed in order to get any output on the map. Node.js was required because it is based off the Leaflet heatmap plugin. This was the bypass to get this working the way we had previously intended.

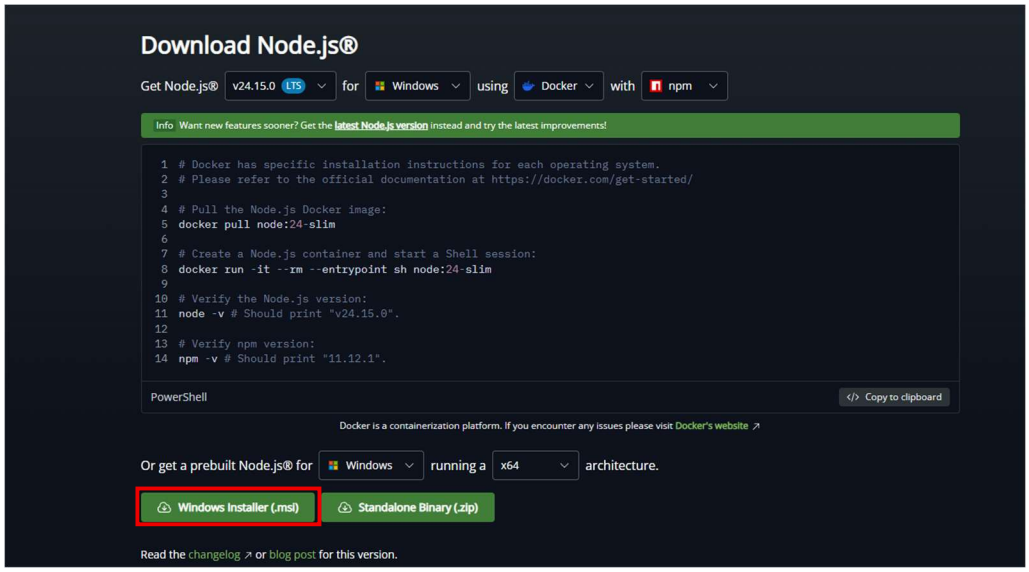
Installment steps

Step 1) Access https://nodejs.org/en/download.



Screenshot showing the Node.js landing page. The red box indicates what the user should click to access the download page.

Step 2) After clicking “Get Node.js”, Click Windows Installer (.msi)



The user will be prompted to download node.js in two separate ways. The user must select Windows Installer which is labeled by the red box.

Step 3) Run the downloaded .msi file. Follow all the steps. When prompted to install required dependencies, select the checkbox next to it. This may open Windows PowerShell. Allow this to happen. This is how it creates the files.

Step 4) Open IntelliJ or another Java IDE. Access the built-in terminal. Type *node -v* followed by the enter key. Then type *npm -v* followed by the enter key. Then type *npm install* followed by the enter key again. This should start a spinning wheel to indicate the process has been executed.

Step 5) Close the IDE and reopen it to verify it has been installed properly.

After this, the maps heatmap and clustering feature will work properly.

Frequently Asked Questions

Q: Why is the map not loading?

A: The map may not load if Node.js is not installed or if npm install has not been run in the project terminal. Ensure you have an active internet connection.

Q: Why are no houses showing up in my search?

A: This may happen if your search criteria are too specific. Try removing some filters or increasing the square footage or percent error.

Q: Why is the application running slowly?

A: The Zillow scraper may take extra time when searching through a large number of listings or when internet speed is slow. The Geocoding takes a long time to process as the API is rate limited. This means that 1 second has to pass before another one can be geocoded. If you choose a location with a lot of houses, this may take a while to process. Make sure you only press search once as this could also increase your wait times.

Q: Do I need to fill in every search box?

A: No. Only City and State are required. All other fields are optional.

Q: Why is my price estimator not giving an output?

A: Please make sure you have all these boxes filled in on the price estimator page. This is required on this page to ensure the most accurate result.