



Collecting Zillow Data Into a Spreadsheet

A Step-by-Step guide on
how to harvest 'sold' data to analyze the value of your property

The **RE**ssentials session on *Collecting Zillow Data into a Spreadsheet*, is a companion module to the **RE**ssentials session on *The Palace Method*.

It is designed to help those who don't have access to Multiple Listing Service, MLS, data in their area. In this presentation we will guide you through harvesting the sold properties information available on Zillow.com into easy to analyze lists showing the range of \$/SqFt values of the collected data.

This data information we are going to collect is available on other sites than Zillow.com, such as Realtor.com or Homes.com, and the process outlined in this session can easily be adapted to those sites if you are partial to them.

So, let's get right to it and see how you can harvest Zillow data into a valuable spreadsheet showing the range of sold property data in your area...

Zillow Navigation

1. The first thing to do is go to Zillow.com in your browser and either search on your actual property address or the general location description. For our purposes here, I'll use a town name.
2. Zillow will bring up a map of the defined area showing current properties for sale with red dots. Using the **+** & **-** zoom icons in the upper right of the map, and clicking on and moving the map, we can zoom to the sub area we want to analyze.
3. Once we have a closer view of the target area, click on a property for sale, and then click on the **EXPAND** menu option in the upper right.
4. Scroll down past the Home Value section, the Price and Tax History section, etc. until you get to the Neighborhood section of the page. There you should see another map of properties in the area (possibly in aerial photo view).



5. But now you will see not only properties on the market as red dots, but recent sold properties as well, which are the yellow dotted properties.
6. Zoom out until you can see the broader area in the map window, which hopefully will have quite a few yellow dots, meaning sold homes, to investigate.
7. When you click on a yellow dot, you will see all the data we need in a popup, including date sold. Ideally we are looking for properties sold within the previous 12 months. If we have to go older than that, so be it, however we should be aware of the average annual area appreciation, which Zillow may have stated above the map. With properties sold over 12 months ago, you might increase their sale price by the appreciation for your data analysis.

Copy and Paste to a Text Editor

1. Next we want to open a text editor. I'm on a windows pc, and Windows Notepad is what I use. On MAC you can use TextEdit. It is probably easiest to reduce the Zillow window enough to have the text editor appearing next to it.
2. On the data popups of the properties we select as comparable sales we are going to highlight and copy the information. And then we paste the data into the text editor.
3. Repeat this process for as many comps as you can find. This may take 30 minutes or more, but in terms of understanding your market, this is some of the most valuable time you can spend. Also the data we are saving can be used later during the sales and negotiation phases, so this is extremely important data to be saving. In essence, it makes you as smart about your market as any agent you will talk to. BTW ... if you click into the property by mistake and lose your dotted map view, don't worry. You can get back to the map simply by using the browser's back arrow. You will have to zoom out again with a few clicks, but you should be back where you were before clicking into the property by mistake... Let's fast forward through the rest of these...
4. We are going to do some more editing in this text file, but before we do, let's make sure we have it saved. I'm going to save it to my desktop and call it MyComps.txt

If you don't want to build a spreadsheet...

If you are not familiar with how spreadsheets work and don't want to know, you can just use this file, as-is, to analyze your area.



You need to get out a calculator and for each property divide the sale price by the number of SqFt to get each property's \$\$SqFt value. Then identify the highest and lowest \$\$SqFt property values and use them for ***The Palace Method*** Market Value Range.

Preparing the text data for importing to a spreadsheet

Each property's data we have copied into the text file is on multiple lines, and our goal is to get all of each property's data onto one line, and each value, like address, price, sale date and SqFt separated by commas.

The reason we are doing this is to create what is called a comma separated format that can easily be imported into a spreadsheet.

That means we also want to remove commas from the price and SqFt numbers, so they won't get divided up by mistake as separate fields. Also, we can get rid of some unneeded text.

First, we will go through manually editing the lines of a property's data to understand the process...

1. So, the first thing I'm going to do is get rid of the state and zip since we don't need it for our purposes (we know where we are!), and I'm also going to remove the words "SOLD ON" so that we just have the date remaining. When I delete that selection, the date moves up to the first line.
2. Next, I'm going to replace the colon after the date with a comma since we are wanting to have the data fields separated with commas, and then pressing delete brings the price up to the first line. Also, we need to remember to remove the comma from the price, so that the price number doesn't get broken up into two fields by mistake.
3. Then we want to have another comma after price, and pressing delete after that brings the rest of the data up to the first line.
4. Finally we want to remove the comma from SqFt field like we did in the price field, and then we don't want the "SqFt" description so we will get rid of that.
5. So now what we see is a neatly arrayed line of data with commas between each data field for the property.



Now let's do essentially the same process, but use the Edit Menu's Replace option to accomplish a lot of the edits globally across the whole file...

1. We can do this to the rest of the properties a bit faster using the Edit Replace tool.
2. First, we want to remove the state and zip, so I'm copying it into the Find what box and we want to replace it with nothing, so the Replace with box stays empty and clicking the Replace All button removes the words on each property.
3. We will do the same for the words SOLD ON.
4. We will replace the colon after the date with a comma.
5. Next, we want to get that comma out of the price number and placed after the price. So what I'm going to do is replace the ,000 with 000,
6. Finally, we want to remove the words SqFt, so we will leave, once again the Replace with box stays empty.
7. At this point we can close the Edit box, and we are going to start using the delete key at the end of each row to reduce the multiple rows down to one. And as we do this there is one last comma to get rid of in each property's SqFt field.
8. Once we have the data edited neatly into one line for each property with each data field separated by a comma, we can resave the file.

A final note here... depending on the type of property you are researching, and the price values associated with it, etc., the data formatting you copy and paste may not be exactly as it has appeared here in our demonstration. So be prepared to follow a slightly different editing procedure in the text file to reduce all the properties into single rows of comma separated fields of data. The key thing is to understand the process of editing out unneeded words and cleanly separating each data point with commas, making sure to remove the commas from the price and SqFt fields.

Working the data into a spreadsheet

Now we are ready to import it into spreadsheet. If you are an Excel user, you can open Excel (either the online or Office 365 version is fine) and then open our saved file.

This same process works for Google sheets as well if you prefer that app suite.

1. When opening/importing the text file, you will be asked how to delimit the data fields: select 'commas'.



2. Once the file is opened into the spreadsheet, we will see one row for each property with the data fields separated into columns. What we want to do is create a last column that will be the calculated $\$SqFt$ for each property.
3. So that we can be clear about what fields are what, I'm going to create a header row with a few labels for price and SqFt which we have and a new column for $\$SqFt$ which we want to create.
4. Click into the empty cell after the first property's SqFt. To get the $\$SqFt$ we need to divide the price cell by the SqFt cell, so we type the equals sign and click on the price cell, then type the division sign which is the forward slash on the question mark key on your keyboard, then click on the SqFt cell. Pressing the enter key will complete that calculation.
5. You will see the result in dollars and cents. For our purposes we just need to know the whole dollar amount, so let's get rid of the cents by using the decimal collapse icon on the Home toolbar.
6. To replicate this for each property click back on the new $\$SqFt$ cell we just calculated, and the cell will have a border around it with a little square in the bottom right corner. Just click on the little square and drag straight down the rest of the property rows in the same column. This will copy the formula for each property and show the results when you stop the drag process.
7. Finally, we are ready to finish getting what we have been after this whole time: the Highest and Lowest $\$SqFt$ value. To do this we just need to leave the $\$SqFt$ data column values highlighted and use the sort tool on the Home menu bar. I would sort from the highest to lowest value, since it's always nice to see the highest value at the top. When you start this sort, Excel will show a Sort Warning dialogue and you want to select expand the selection to make sure all the row data is kept together during the sort.

And there we have it: the complete list of comparable properties showing the highest $\$SqFt$ property on the first row and the lowest $\$SqFt$ property on the last row.

Conclusion

For some of you with properties in densely populated areas that have a lot of annual sales data to collect, this presentation is easy to use.

For others who have properties in less dense area who don't have many recent sales to choose from, the process we have outlined can be more challenging.



In those cases, you may have to cast a wider geographic net, searching for similar properties to yours in geographically dispersed but comparable areas, perhaps within the same distance of a city or town hub.

Or if you are looking at rural property comparisons, you may even have to scan over entire counties or states to find other meaningful comparable sales.

Regardless, the process we have outlined should help you create a valuable list of easy to analyze data points that will be useful throughout the sales cycle for your property, from initial pricing through the negotiation phase coming to a final sale price with a buyer.



The PaLaCe[®] Method Worksheet

$$P = \frac{L_{(\$sqft)} + C_{(\$sqft)}}{2}$$

	Location	Condition	Market Value Range \$\$sqft	
Best				High
Worst				Low