

Creating a Histogram in Venngage

Part of a Series of Tutorials on using Google Sheets to work with data for making charts in Venngage



University of Colorado
Boulder



These materials are based upon work supported by the National Science Foundation under Grant Nos. IIS-1441561, IIS-1441471, & IIS-1441481. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

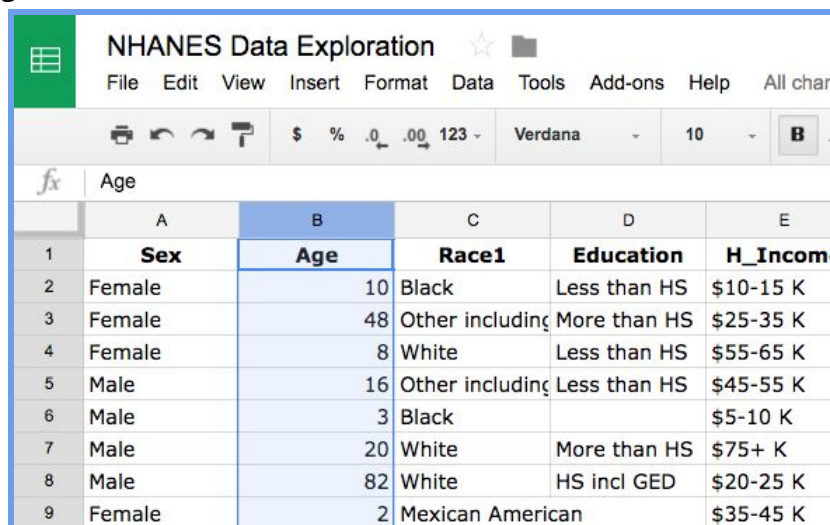
Start with a 2000-case sample collected from the EEPS NHANES data explorer (<http://www.eeps.com/zoo/nhanes/source/choose.php>) and placed in a Google Sheet. To see full instructions on how to do this, check out the mini lesson titled “Creating Graphs of Distributions in Venngage.” We will use these data to create a histogram in Venngage using Google Sheets features to prepare the data.

Setting up a Histogram in Google Sheets


Histograms are a way to graph continuous numerical data (i.e. 2.3, 4, 1.02, 16, etc) into bins that can be displayed on something like a bar chart. No matter what the size of the bins are, histograms allow viewers to easily grasp the distribution of a quantity and observe patterns that would be difficult to find otherwise.

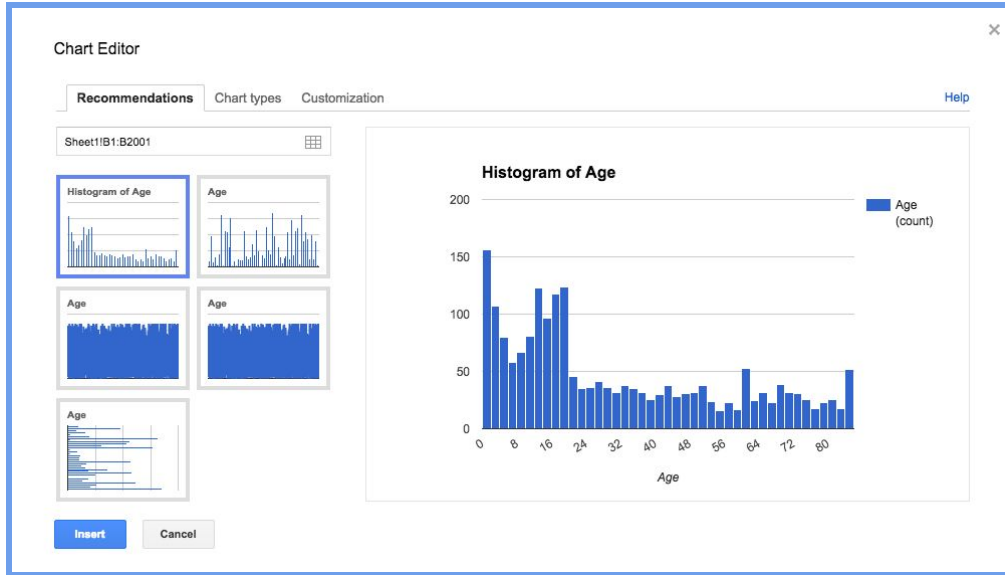
In the NHANES data, we might be interested in the distribution of ages in our sample. Are there a lot of children? Senior citizens? Mostly middle-aged people? A histogram will help us answer these questions.

Since Venngage does not have a histogram chart option, we will use the Venngage column chart structure and take advantage of Google Sheets’ histogram ability to get the numbers to put into Venngage. Start off in Google Sheets with the NHANES data and select the column “Age” by clicking on the column letter above it.

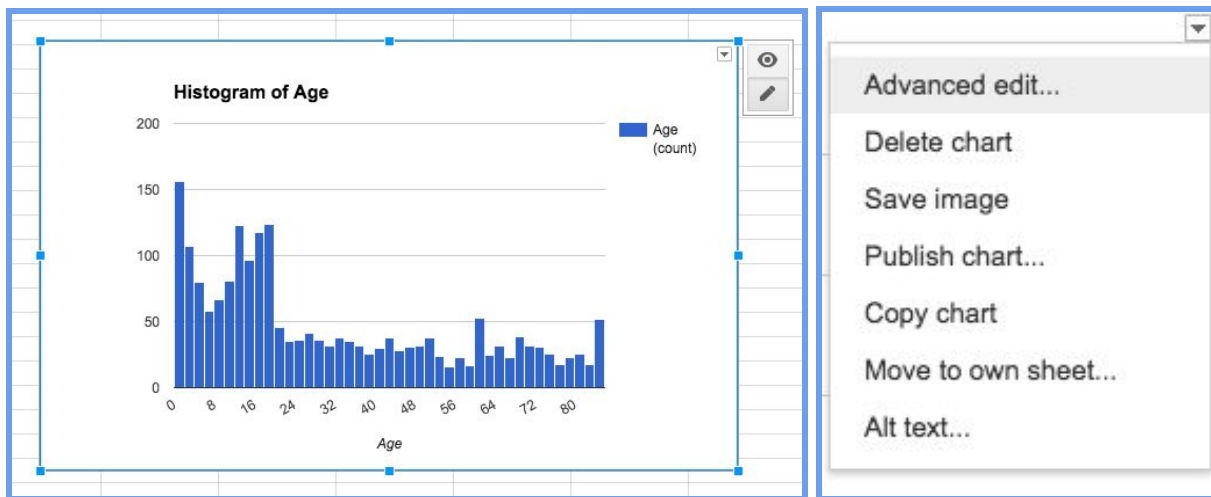


	A	B	C	D	E
1	Sex	Age	Race1	Education	H_Income
2	Female	10	Black	Less than HS	\$10-15 K
3	Female	48	Other including	More than HS	\$25-35 K
4	Female	8	White	Less than HS	\$55-65 K
5	Male	16	Other including	Less than HS	\$45-55 K
6	Male	3	Black		\$5-10 K
7	Male	20	White	More than HS	\$75+ K
8	Male	82	White	HS incl GED	\$20-25 K
9	Female	2	Mexican American		\$35-45 K

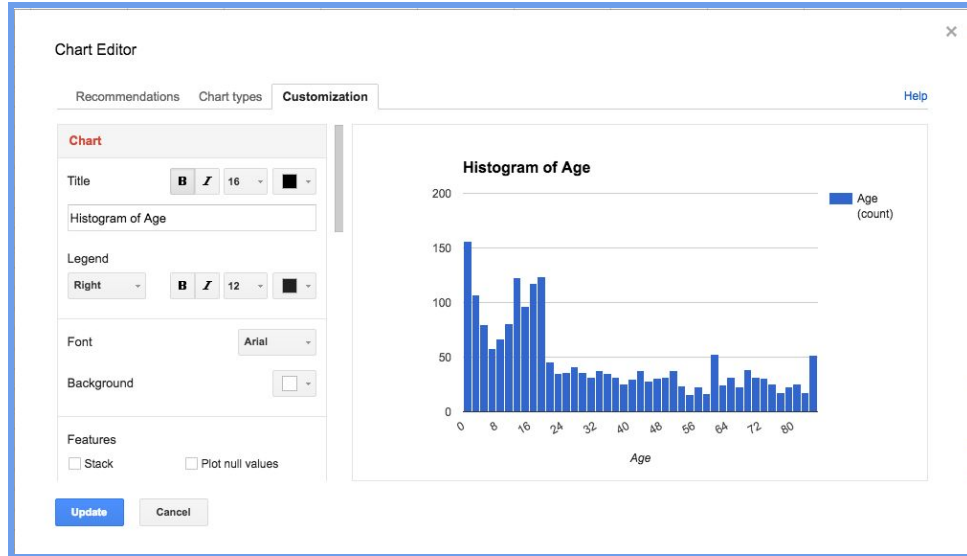
Next, click on the chart icon  on the menu. A Chart Editor dialog box will appear. Since the column for Age contains continuous numerical data, Google Sheets will automatically recommend a histogram as the first option under Recommendations.



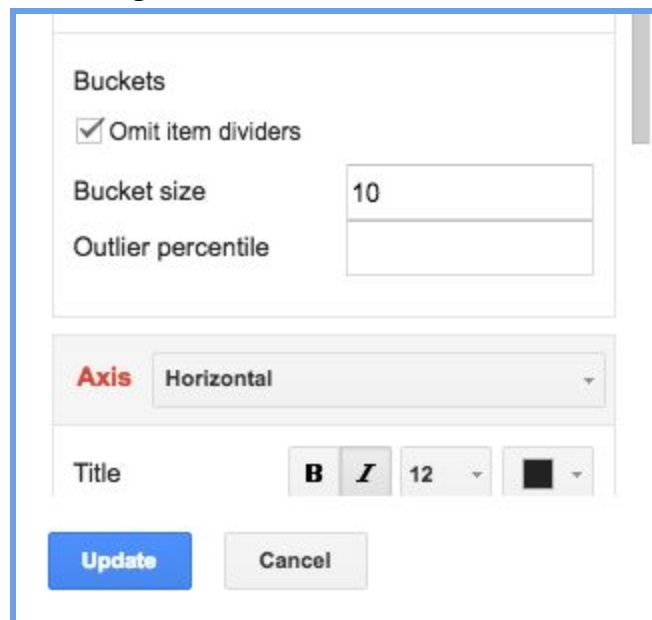
Click on the Insert button on the bottom left of the dialog box to insert the histogram into your Google sheet so that you can edit it. You'll first want to decide how wide you'd like the bins to be. Look at the x-axis. In this case, it appears that the bin width has been automatically set to 2, since there are four bars before the number "8" on the x-axis. We're going to change it to bins of 10 to mark the different decades. On the inserted histogram, you will see a small downward-pointing triangle on the top right corner of the histogram. Click on that, and then choose the first option, Advanced Edit.



A Chart Editor Customization dialog box will appear.

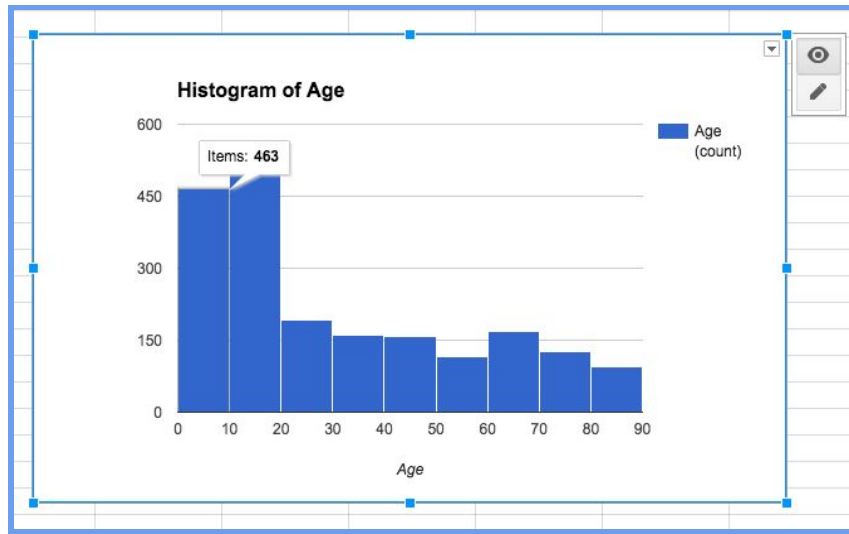


On the left, scroll down to where you see a section for “Buckets” (Google Sheets’ term for bins) and set the Bucket size to “10”, and then click the Update button on the bottom. Your histogram should now have larger bin sizes of 10.



Now you can use the histogram to get numbers to enter into Vennage. Click on the histogram and make sure that you are in “View Mode”. It’s the eye icon in the upper right corner of the histogram. With “View Mode” on, you can hover your mouse cursor over each bin, and it will show you how many items are in each bin. Here, I have 463 people whose ages are between 0 to 10. (Your number will probably be different from mine since we don’t have exactly the same NHANES data sample subset. The NHANES data explorer provides a random sample each time data is requested.) You can get the number of items in each bin

in a similar way; it is a good idea to write them down so that you will have them ready to enter into Venngage.



Creating a Histogram in Venngage

In a different browser tab, open up Venngage and insert a column chart into your infographic canvas. Double-click on the chart to change the default data.

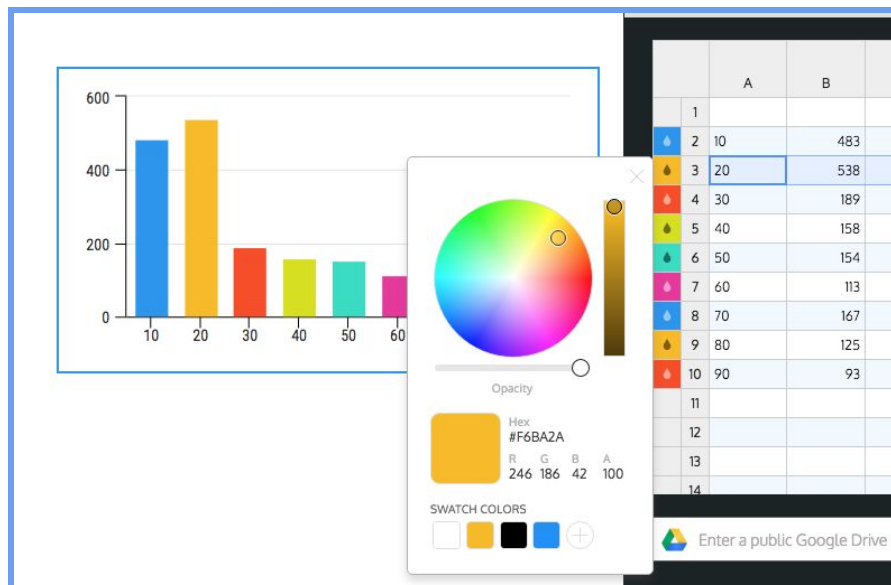
	A	B	C	D	E
1		2015			
2	May	18			
3	June	12			
4	July	6			
5					
6					

First, select and delete the default data. Starting from row 2, enter in the end bin number in column A (for example, if the bin is from “0 to 10”, enter “10”) and then the number of items in the bin in column B. Row 1 in Venngage data tables are reserved for labels, regardless of whether it appears in the chart. Keep referring back to the Google Sheets histogram (or your written list of bin contents) to fill in the table, adding a bin “label” and the number of entries

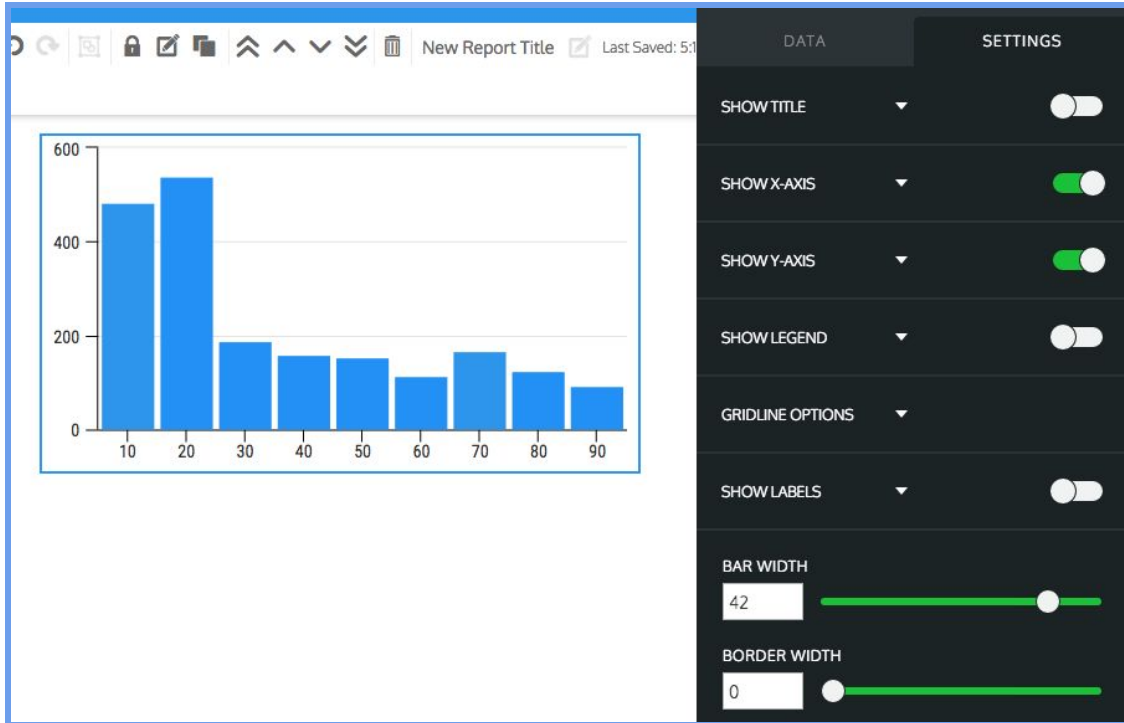
in each bin in each row. Once you've filled in the table, you have something like a histogram, but with a few important differences. We'll try to minimize those differences in the next few steps.

Making the Column Chart Look like a Histogram

Vennage automatically makes each bar of a column chart a different color, because column charts are designed to show the relative size of different categories of objects (e.g. males and females). But all of our values in this graph are ages, so the convention is that the bars should all be the same color. In order to make all the bars the same color, double-click on the colored square with a water droplet icon inside. A color picker pop-up box will appear and from there you can change each of them to the same color of your choosing. To save a color to be used multiple times, first select the color you want using the circle for hue and brightness, vertical bar for darkness, and then opacity/translucency, if you like. Then under the section for swatch colors, you'll see a circle with a "+" inside. Click on that and it will be added to the swatch colors. As you change the other colors for the different columns, you can just click on the swatch color you created and have the exact same color.



Another difference between the column chart and a histogram is the space between bars. In a column chart, there is space between the bars because they represent different categories, but in a histogram, there are no values between adjacent bars, so the spaces are somewhat misleading. You can get rid of most of the space between the bars of your column chart by making the bars wider. In the Settings tab of the chart, adjust the bar width option until the bars look like they're very close to each other but do not quite touch. Make sure you label your x-axis with "Age". Then Save. You're done!



Extra tip: Another difference between what you have and a histogram is the way the bins are labeled. If you want your graph to look more like a histogram, you can add a white rectangle to cover up the x-axis numbers and the tick marks, and add in a single text box with the numbers spaced out on top of the white box.

Final result:

