

Visualizing Relationships Between Variables

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



University of Colorado
Boulder

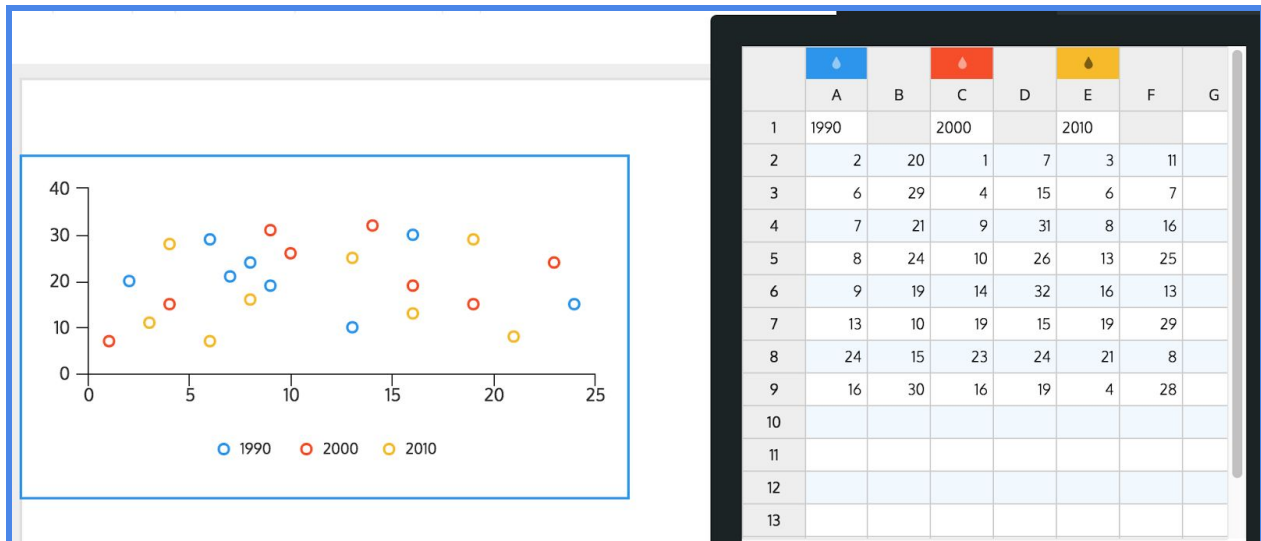


These materials are based upon work supported by the National Science Foundation under Grant Nos. IIS-1441561, IIS-1441471, & IIS-1441481. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF.

One way to investigate whether two variables are related to one another is to create a scatterplot, which plots one variable on the X-axis and one on the Y-axis.

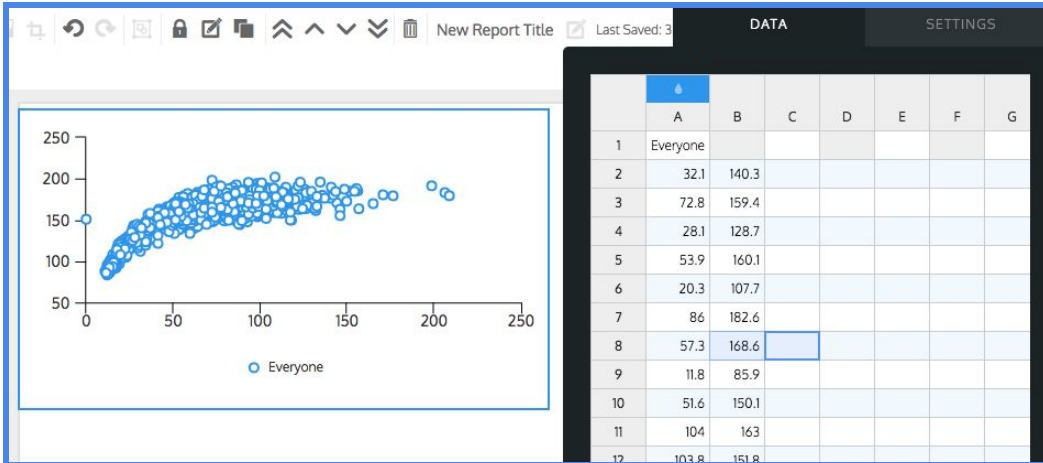
We will be using the same Google Sheet file of NHANES data from the previous exercise (Creating Graphs of Distribution in Venngage). If you do not already have the dataset in Google Sheets, please refer to the previous exercise for instructions on getting the data.

In the NHANES dataset, two variables that we might imagine have a relationship to one another are height and weight. We will create a scatterplot in Venngage to explore this relationship.



In Venngage, drag a scatter plot onto the canvas and double-click it to edit. The default data table that comes with the scatterplot is set up for three different “series” of data. Row 1 contains the labels for each of these series. Then the columns alternate between x-axis and y-axis for each series. To create our own scatterplot, we first have to clear out the data table by selecting all the data in the data table and deleting it by hitting the Delete or Backspace key.

Our x-axis will be Weight, and our y-axis will be Height. In Google Sheets, select all the data in the column for Weight, not including the label on the first row. Paste this into the data table in Venngage, starting at cell A2, leaving row 1 for the label. Copy and paste the Height data into column B, starting at cell B2. We’ll label this series “Everyone” in cell A1. Our scatter plot will look like this:



It's hard to interpret this graph, though, without knowing what each axis represents, so we need to label the axes. In the Settings tab, click on "SHOW X-AXIS" to reveal the details of the axis and enter the x-axis Title, "Weight in kg". Notice that it's not just "Weight" because that doesn't tell people what the units are.

CANCEL SAVE

DATA SETTINGS

SHOW TITLE

SHOW X-AXIS

AXIS TITLE

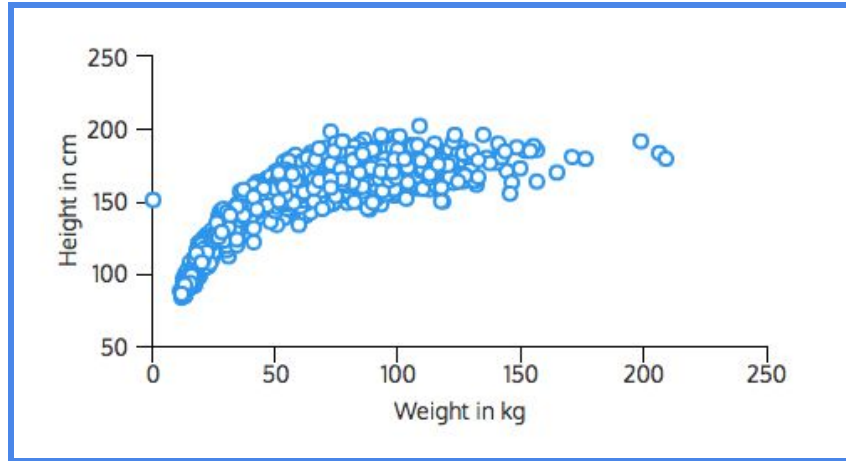
Enter a title

FONT STYLE

Oxygen 14 B I U

Do the same with the y-axis and label it "Height in cm". The kg and cm come from the EEPS NHANES data exploration tool web page.

You can use the On/Off slider in Settings to hide the Legend, which isn't very illuminating in this case. Our scatter plot now looks like this:



What does this scatter plot reveal about the relationship between height and weight in this sample?