This is a tutorial that is taken from the Tableau training site.
The url of the site is https://www.tableau.com/beginners-data-visualization.

The Beginner's Guide to Data Visualization

Data Visualization Training

- On Demand Training
- Live Online Training

## Data Visualization Whitepapers

---

- Visual Analysis for Everyone
- Designing Great Visualizations
- Improve Your Vision & Expand Your Mind With Visual Analytics

---

Data Visualization Training for Everyone

Data visualization refers to the process of converting data into a graphical or visual representation. Because our brains process visual information so efficiently, visualizing data graphically greatly speeds up the process of data analysis. To quickly illustrate the difficulty of understanding data in numerical form, time yourself as you search for the outlier in the following data set:

	A	B
1	Year	Sales
2	1981	1.4622
3	1982	1.47004
4	1983	1.49253
5	1984	1.49118
6	1985	1.49722
7	1986	1.50138
8	1987	1.50008
9	1988	1.51493
10	1989	1.50781
11	1990	1.50899
12	1991	1.53037
13	1992	1.58137
14	1993	1.54299
15	1994	1.53307
16	1995	1.55845
17	1996	1.56213
18	1997	1.54488
19	1998	1.56927
20	1999	1.55305
21	2000	1.5571
22	2001	1.56235
23	2002	1.58847
24	2003	1.59309
25	2004	1.58303
26	2005	1.5947

Now time yourself as you search for it in the same data set, but this time with the data presented visually:
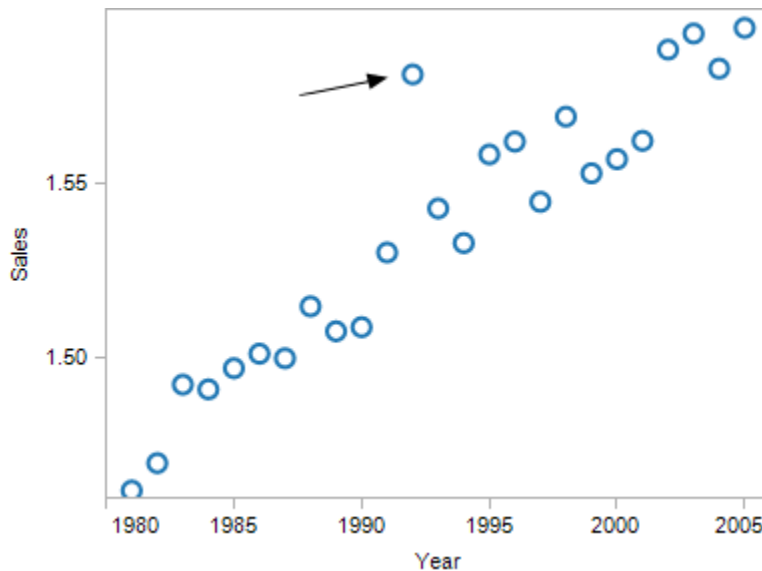
Much faster, isn't it? This was a simplistic example, but illustrates nicely the time saved by using visualization as an analytic tool, rather than simply generating nice charts and graphs for presentations. As the data sets become more complex, the potential time savings become even greater.

The General Social Survey

In this example, we'll be using Tableau's data visualization software to explore some data from the General Social Survey. The General Social Survey is an NSF-funded survey, interviewing more than 50,000 Americans over nearly 3 decades. Survey questions cover topics ranging from political, to economic, to educational, and more. But to keep things interesting, we'll be talking about premarital sex.

Acceptance of premarital sex has been steadily increasing in the US. In the 2008 General Social Survey, for the first time ever, a majority of respondents said they believed that premarital sex was "not wrong at all". In this tutorial, we'll be exploring historical data from the GSS, using data visualization to uncover hidden trends.

Connecting to Your Data

You should already have a copy of Tableau Desktop installed on your computer. Download the CSV file with the GSS data to your computer, then simply drag and drop the CSV file onto the Tableau Desktop icon or into an open Tableau window.

Tableau automatically knows the settings for a Text File Connection, so just click OK. You might get a dialog box afterwards asking if you'd like to create an extract or remember this data connection – for now, just say no.
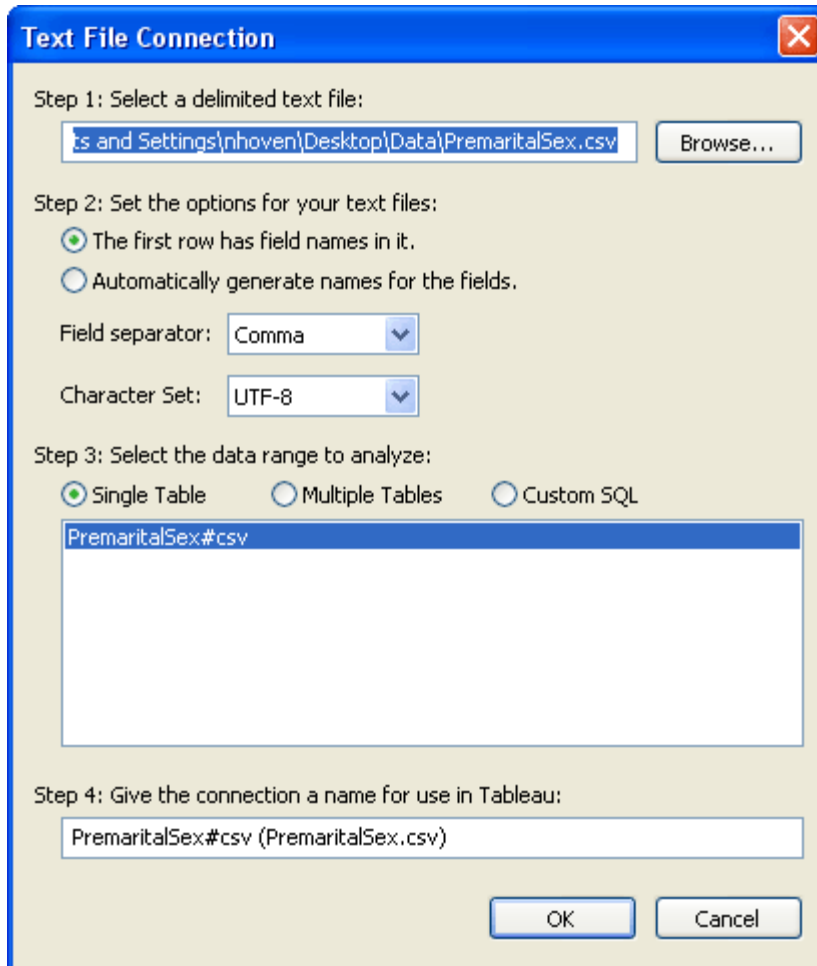
Tableau automatically reads the field names from the data and populates the dimensions and measures areas. Measures contain numerical data like AGE, while dimensions contain categorical data, like marital status.

Measures and Dimensions

Tableau did a pretty good job of figuring out what goes where, but notice that Sex is included under measures when it should really be a category. That's because our data file used "1" and "2" to represent male and women, so Tableau thought it was a numerical field. Let's fix that. Simply drag SEX from the measures area and drop it in the dimensions area.
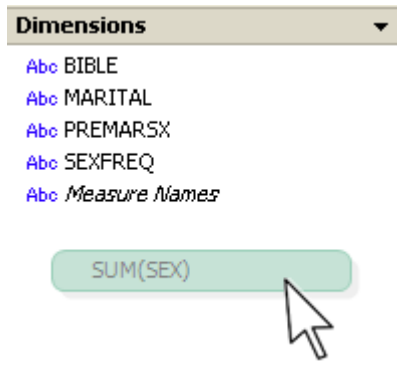
Editing Aliases and Data Labels

While we're at it, right click SEX, choose "Edit Mark Properties >> Aliases", change the value "1" to "Men" and the value "2" to "Women". Then right click SEX again, select "Rename" and change SEX to GENDER. That will make things less confusing when we're talking about premarital sex and frequency of sex.

Year is also included as a measure, but it should really be a category as well. We typically segment data by year, we don't add or average years together. Drag YEAR to the category box, too.

Creating Calculated Fields

In Tableau, it's easy to get more information about a particular field. Right click SEXFREQ and choose "Describe". If you click "Load", you'll see a list of all the existing values in this field. The field contains each interviewee's response to the question "How many times did you have sex last year?" Their responses are grouped into 7 bins (plus a null value for those interviewees who gave no response or weren't asked), but it would be more interesting if we had a numerical value for each person's yearly sexual frequency.

Fortunately, we can approximate that. Close the description dialog box, then right click in the measures area and choose "Create calculated field". Name it "SexPerYear" and paste the following code into the formula box:

```
Case [SEXFREQ]
When "Not at all" Then 0
When "Once or twice" Then 1.5
When "Once a month" Then 12
When "2-3 times a month" Then 30
When "Weekly" Then 52
When "2-3 times per week" Then 130
When "4+ times per week" Then 250
Else Null
End
```
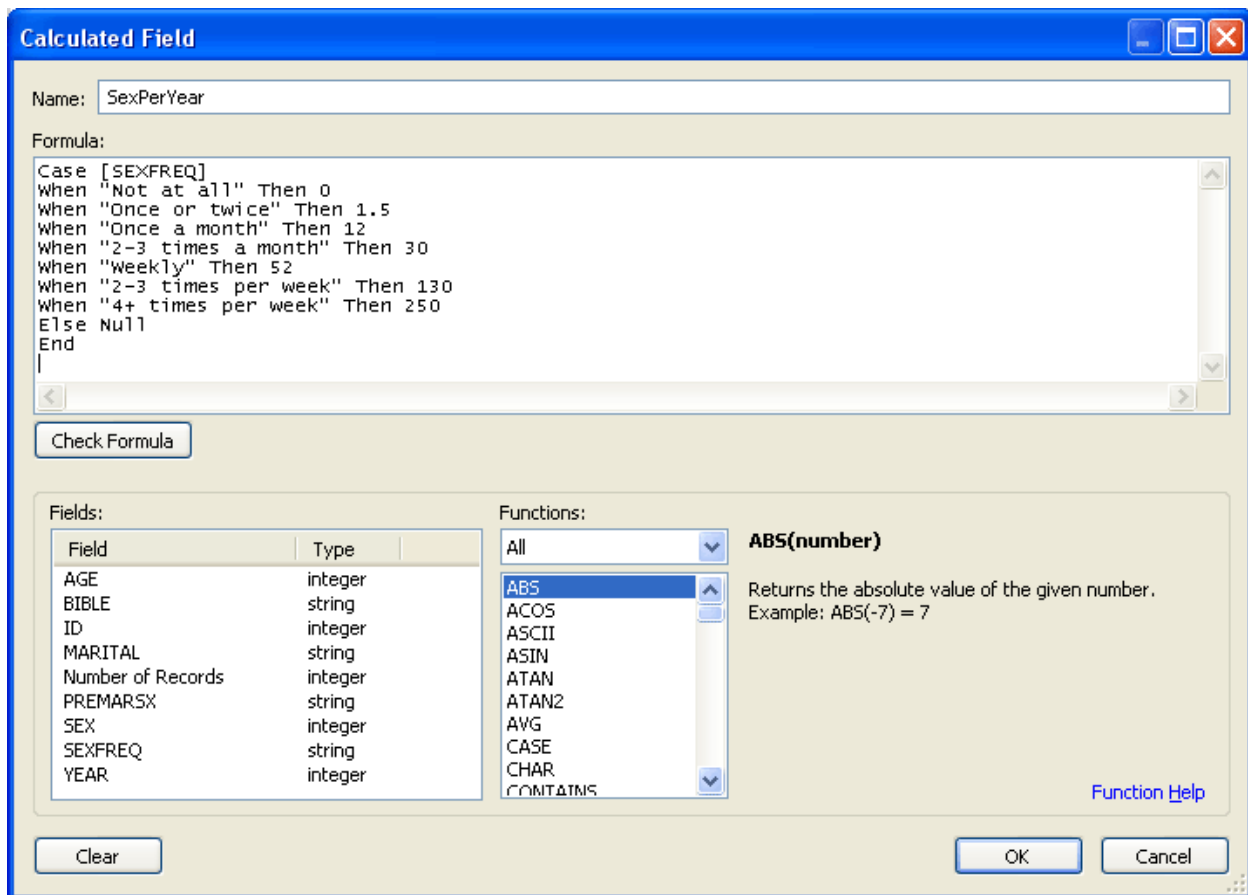
This code creates a numerical field for each respondent. The field contains the approximate number of times they had sex in the past year, based on their answer to the survey question.

Drag and Drop Data Visualization

We can examine this new variable visually in just a few clicks. Drag YEAR to the columns shelf and drag SexPerYear to the rows shelf. The default aggregation is sum, so we're looking at the total number of times all survey respondents had sex each year. Since there were different numbers of people surveyed each year, this chart isn't particularly informative. Right click SUM(SexPerYear) on the Rows shelf and select "Measure (SUM) >> Average". Now we have the average number of times survey respondents had sex every year.
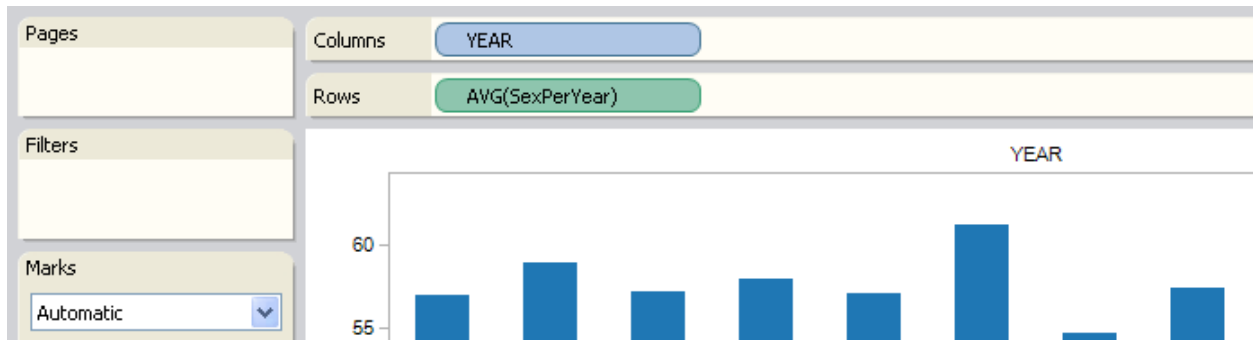
Answering Questions with Visual Analysis

Now let's start asking questions. What effect do you think an interviewee's opinion on premarital sex has on the number of times they have sex each year? Drag PREMARSX to the color shelf to find out.

I'm curious about trends over time, so change marks to Line. The line for NULL (people who weren't asked this survey question) doesn't tell us anything, so right click NULL in the color legend and choose "Exclude".

Grouping Data Set Members

Let's say we've also decided that we don't really need to differentiate between people who think premarital sex is "always wrong" and those who think it's "almost always wrong". Instead, we'll group them together. By holding down the control key, select "almost always wrong" and "always wrong", right click, and choose "group".

We have an interesting result already. Over the past 20 years, those opposed to premarital sex have been having more and more sex on average each year, while those who think it is not wrong at all have been having less and less. Why do you think this is?
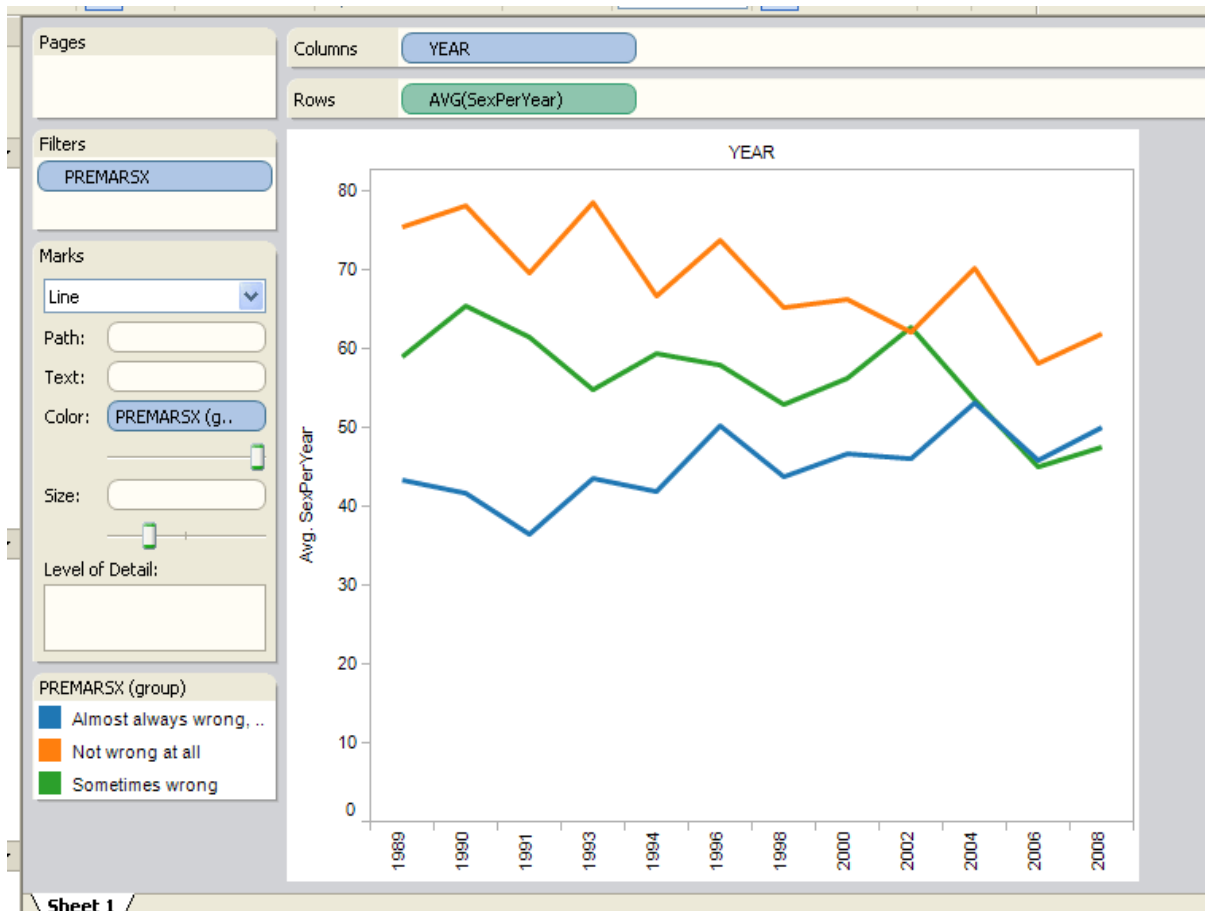
Editing Marks and Color Legends

We're going to attempt to answer this question, but to make our graph even easier to understand, let's change the color legend. Right click inside the color legend and select "Edit Colors". Change the Color Palette dropdown from Automatic to Traffic Light. Click on the "Almost always wrong, Always wrong" group and then select a nice shade of red. Click "Not wrong at all" and select a shade of green, and then select "Sometimes wrong" and select a shade of yellow. Close and then in the color legend, drag "Not wrong at all" to the bottom so the opinions are in order.
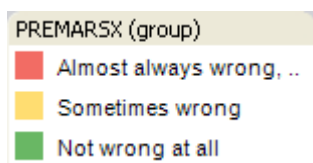
We don't want to confuse correlation with causality, though. It's probable that there's a confounding variable at work here, and one of the most likely culprits is age. Let's look at how the average age of these three groups has changed over the years.

Editing an Existing Visualization

Right click on the "Sheet 1" tab at the bottom of the Tableau window and choose "Duplicate sheet". On the new sheet, "Sheet 2", drag Age from the Measures box and drop it directly on top of the "AVG(SexPerYear)" field on the Rows shelf. "AVG(SexPerYear)" will disappear and be replaced with "SUM(AGE)". Right click "SUM(AGE)" and select "Measure (SUM) >> Average". We see now that the average age of those opposing premarital sex has been decreasing, while the average age of those supporting it has been increasing.
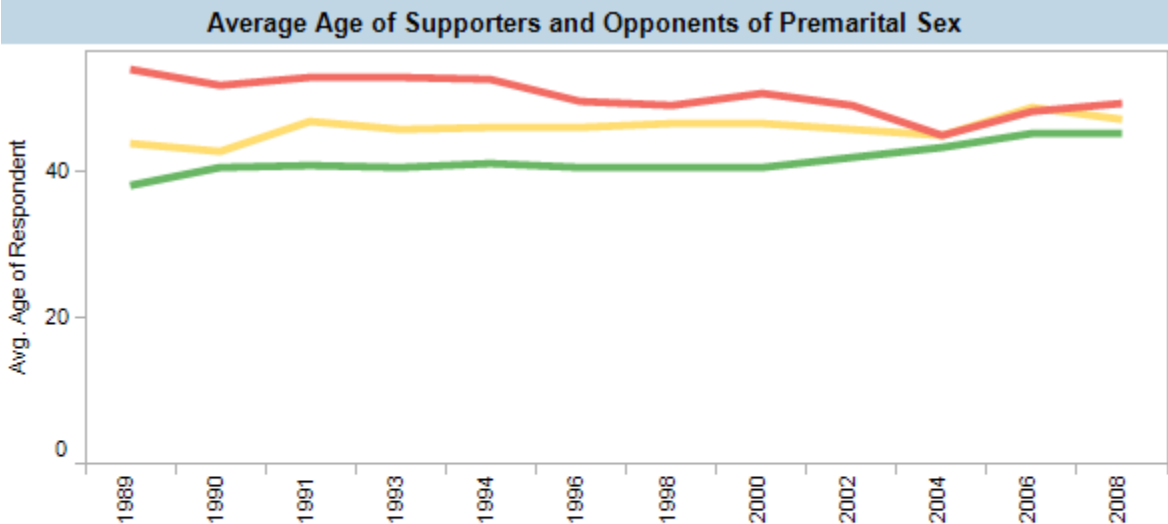
The average age of premarital sex supporters has been increasing every year, while the average age of premarital sex opponents has been steadily decreasing.

Respondent's opinion on premarital sex

- Always wrong
- Sometimes wrong, sometimes right
- Not wrong at all

Since younger people have sex much more frequently than older people (create a new worksheet plotting SexPerYear against AGE if you want to see for yourself), we can see that the puzzling trends in sexual activity are likely linked to the changing average age of these difference groups.

The question remains, however, why is the demographic makeup of these groups changing? Why would the average premarital sex opponent be getting younger and younger? Does it have something to do with recent trends in abstinence education? That seems unlikely, as the shift in age has been going on for twenty years.

Let's examine more closely the changing opinions among the young and the old. Create a new calculated field, this time in the dimensions box. For name, put "Binned Age (<40)" and paste the following into the formula:

```
IF [AGE]<40 THEN '<40'
ELSE '40+'
END
```

Switching Between Different Chart Types

Duplicate Sheet 2, and on the new sheet, switch the Marks dropdown menu from "Line" to "Bar". Drag "Number of Records" from the Measures box to the rows shelf and use it to replace the "AVG(AGE)" field.

We can now see how many respondents each year had each opinion on premarital sex. However, the total number of respondents each year vary by quite a bit, so looking at percentages will make the graph much more understandable.

Calculating Percentages

From the Analysis menu, choose "Percentage of >> Cell". It appears that acceptance of premarital sex is slowly growing among the general population. However, we were particularly interested in how opinions were changing within age brackets. So drag the "Binned Age (<40)" dimension to the column shelf and drop it to the left of YEAR.
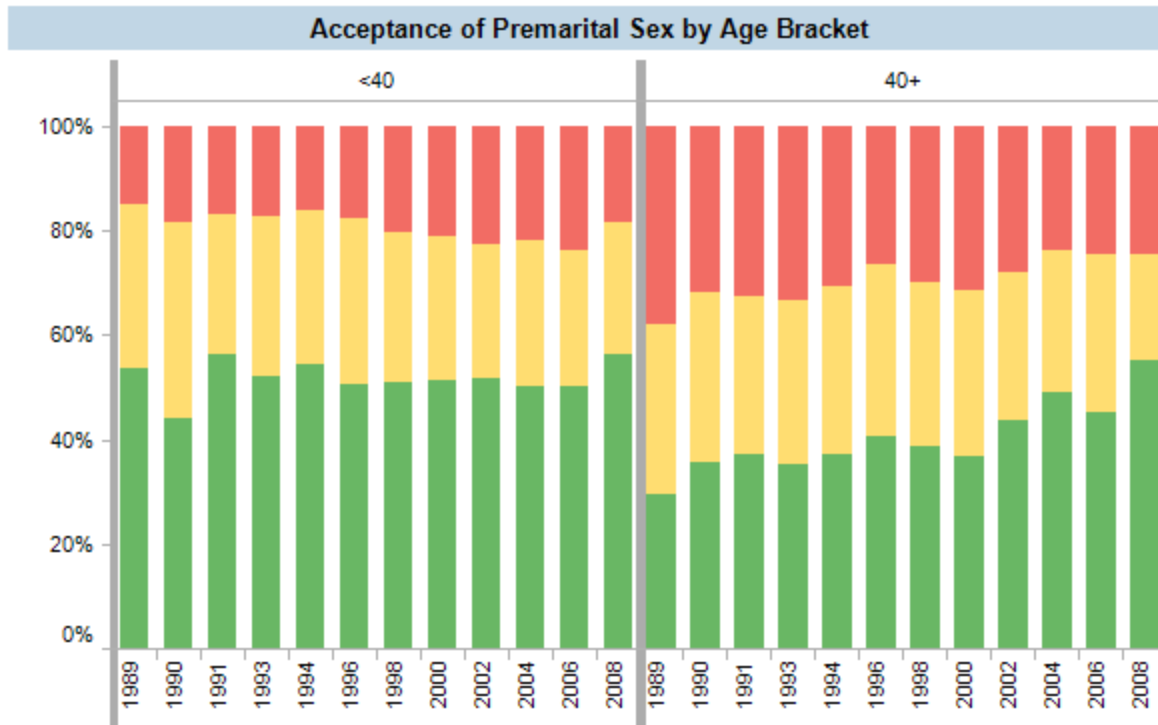
Acceptance of Premarital Sex by Age Bracket

In the under-40 age bracket, beliefs have remained remarkably consistent over the past twenty years. In the over-40 age bracket, on the other hand, there has been a clear shift towards increased acceptance of premarital sex.

Respondent's opinion on premarital sex

- Always wrong
- Sometimes wrong, sometimes right
- Not wrong at all

The answer appears when we segment the population into under-40 and over-40 age brackets. In the under-40 age bracket, beliefs have remained remarkably consistent over the past twenty years. In the over-40 age bracket, on the other hand, there has been a clear shift towards increased acceptance of premarital sex. Do you think this shift represents an actual shift in belief within individuals, or a shift in the population of individuals who make up the older age bracket?

Though on-the-fly data visualization, we were able to rapidly pull new insights out of our date. Rather than simply creating a chart to report results we already, we used visual analysis to explore data and discover new findings. To do the same with your own data, simply open a new workbook and click "Connect to Data". We'll be here for you if you want more instruction, with online training available 24-7 and customer support just a click away.