

Making Data Visualizations

Rishi Valley 2018

Purpose

- Being able to create meaningful data visualizations is extremely important in order to effectively communicate information about large data sets.
- It's also important to be able to use visualizations to **simply “look”** at data that is too complex to make sense of by looking at the raw data alone.
- Anyone working with data should have some skills and facility with producing visualizations of the data to get a sense of what it contains. Visualizing the data allows you to see patterns, trends or relationships you might otherwise not.

Make a Quick Visualization

- On some survey 2,000 people were asked, "What do you do when you're bored?". Here are the most common responses by age group.

age	most common response	number	out of
18 and under	texting	157	500
19-64	watching TV	247	1200
65+	reading	54	300
all ages	talking with friends	451	2000

Make a Quick Visualization

- Take a few minutes by yourself and try to make a visual, graphical, explanation of this data. Try to communicate something through drawing while remaining true to the results of the data
- Take 3-5 minutes to draw
- Compare what you drew with an elbow partner and point out similarities and differences

- In this exercise what was challenging?
- What were the characteristics of the visualizations that effectively communicated this information visually?

Need a Computer?

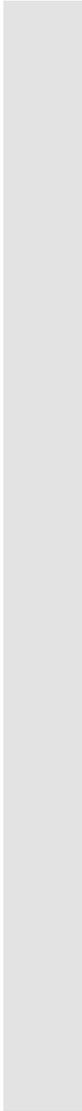
- Do you have to use a computer to create a data visualization?
- What are some reasons that you need to use a computer to manipulate data?

Using Visualizations

- **Do you have to use a computer to create a data visualization?**
- A large data set is too big to understand by looking at a table in a spreadsheet.
- Creating a data visualization with a computer is faster and more accurate than creating one by hand.
- Taking data from its raw state to the point where you can create a meaningful visualization involves several steps. Today we're going to use visualization in attempt to discover things in the data we might not otherwise see.
- It takes practice to create good visualizations. Let's get our feet wet by learning to create charts using Libre Calc.

Objectives

- Pick a single data set
- Create a **bar**, **line**, and **scatter chart** from a dataset using a computational tool.
- Use the settings of a data visualization tool to manipulate and refine the features of a data visualization.
- We are **not learning** to create the prettiest chart; it's about using charts to “tell the story” of what's really going on in the data.



Making a Scatter Chart

Scatter Chart

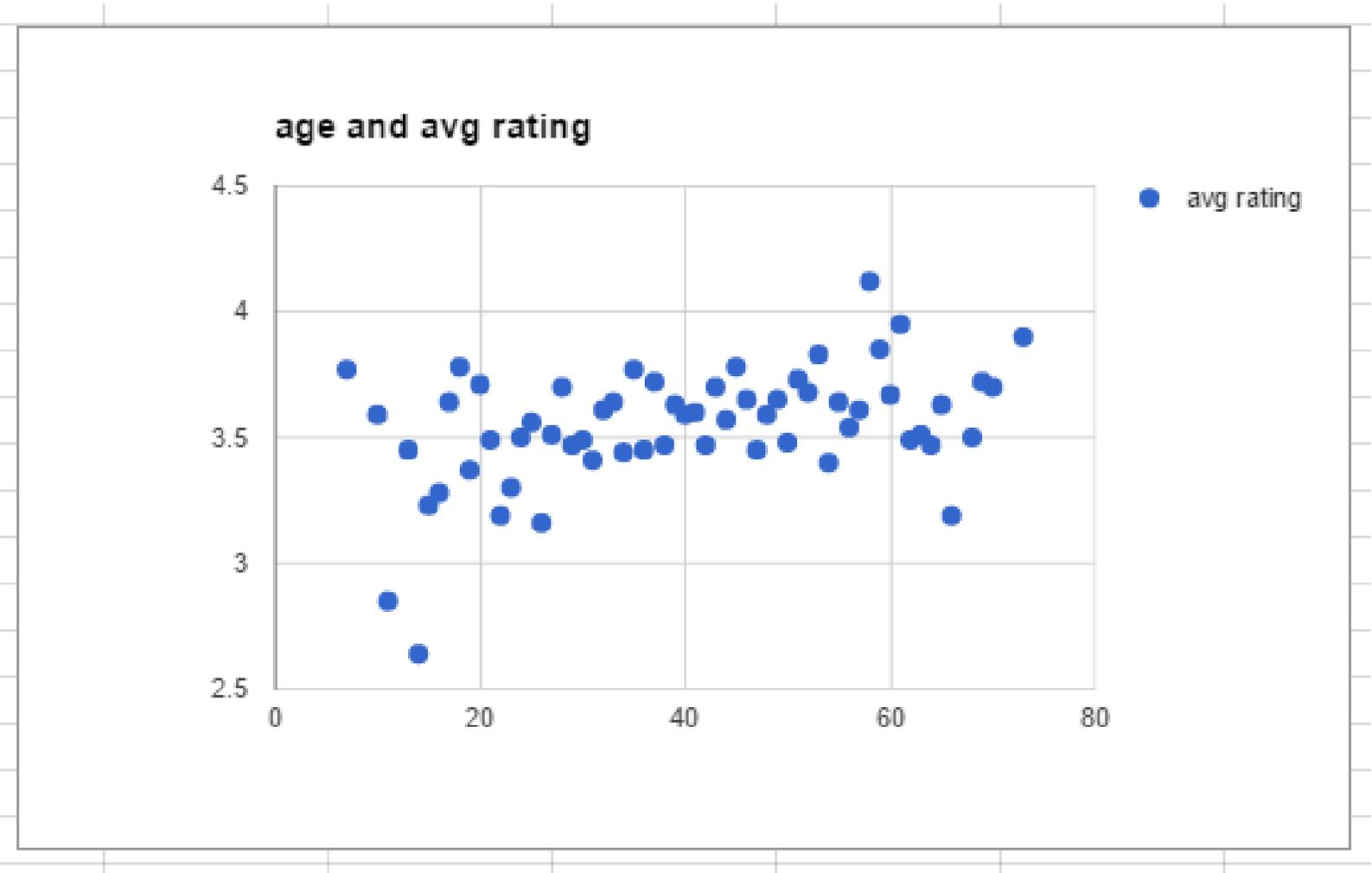
- The first type of chart you're going to build is a scatter chart.
- Scatter charts are useful for finding relationships between two types of data.
- In this exercise, you will build a scatter chart that shows the relationship between movie reviewer age and average movie rating.

Scatter Chart

- **Select Data**
- Select the "age" and "avg rating" columns
- **Insert Chart**
- Select "Scatter Chart" type
- Experiment with the options, and notice how the chart preview changes.
- "Use row 1 as headers." This makes sure your chart uses the column names in the first row of your spreadsheet.
- "Use column A as labels." This tells the chart that column A is your horizontal or "x" axis.
- Click "Insert" to add your chart to the spreadsheet.

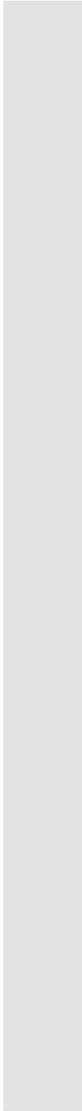
Do a visual
check

At this point, your chart should look something like this:



Debrief

- Can you understand what the chart is showing? Even though the title and axis labels are wrong or don't appear, you should be able to decode the basics of your chart.
- What does this chart help you notice about the data?
- Which age groups had the highest average ratings?
- Which age groups had the lowest?
- What other connections and trends can you see?



Making a Line Chart

Line Chart

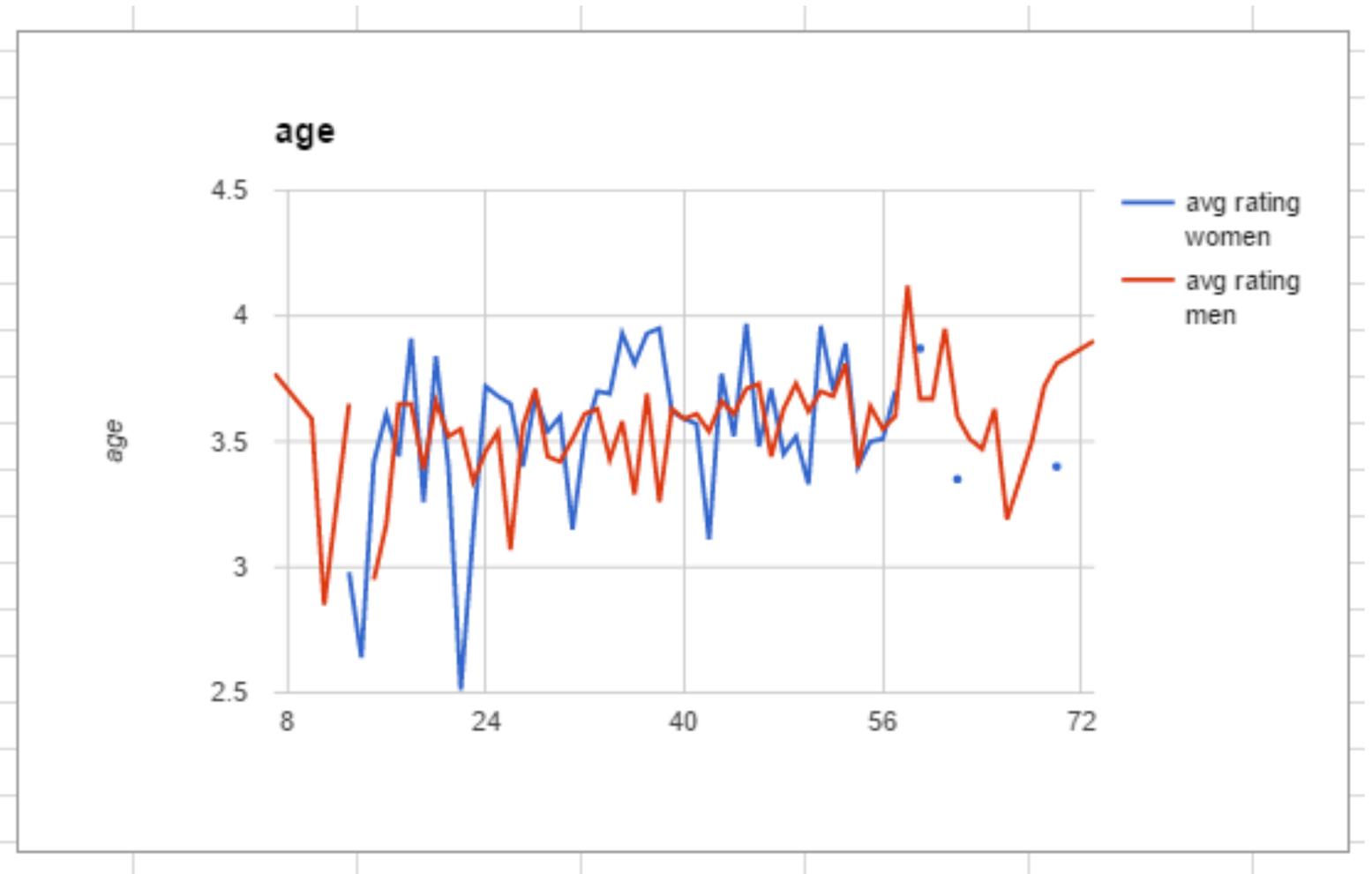
- You can use many different kinds of charts to look at the same data.
- You'll now investigate different columns of your data with a line chart.
- Line charts are helpful for showing the progression of values over time.
- In this case you will be showing how the average movie rating changes with the reviewer's age.

Line Chart

- Select the "age," "avg rating women," and "avg rating men" columns.
- Select Insert -> Chart to open the Chart Editor.
- Set the chart type to "Line chart."
- "Use row 1 as headers" and "Use column A as labels."
- Click "Insert" to add your chart to the spreadsheet.

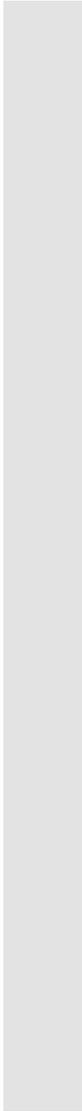
Visual Check

- At this point, your chart should look something like this:



Debrief

- Once your chart looks close to the one above, take a moment to look at your visualization.
- What does this chart help you notice about the data?
- For which ages were the average ratings similar between men and women?
- For which ages were they different?
- What other connections and trends can you see from this chart?



Making a Bar Chart

Bar Chart

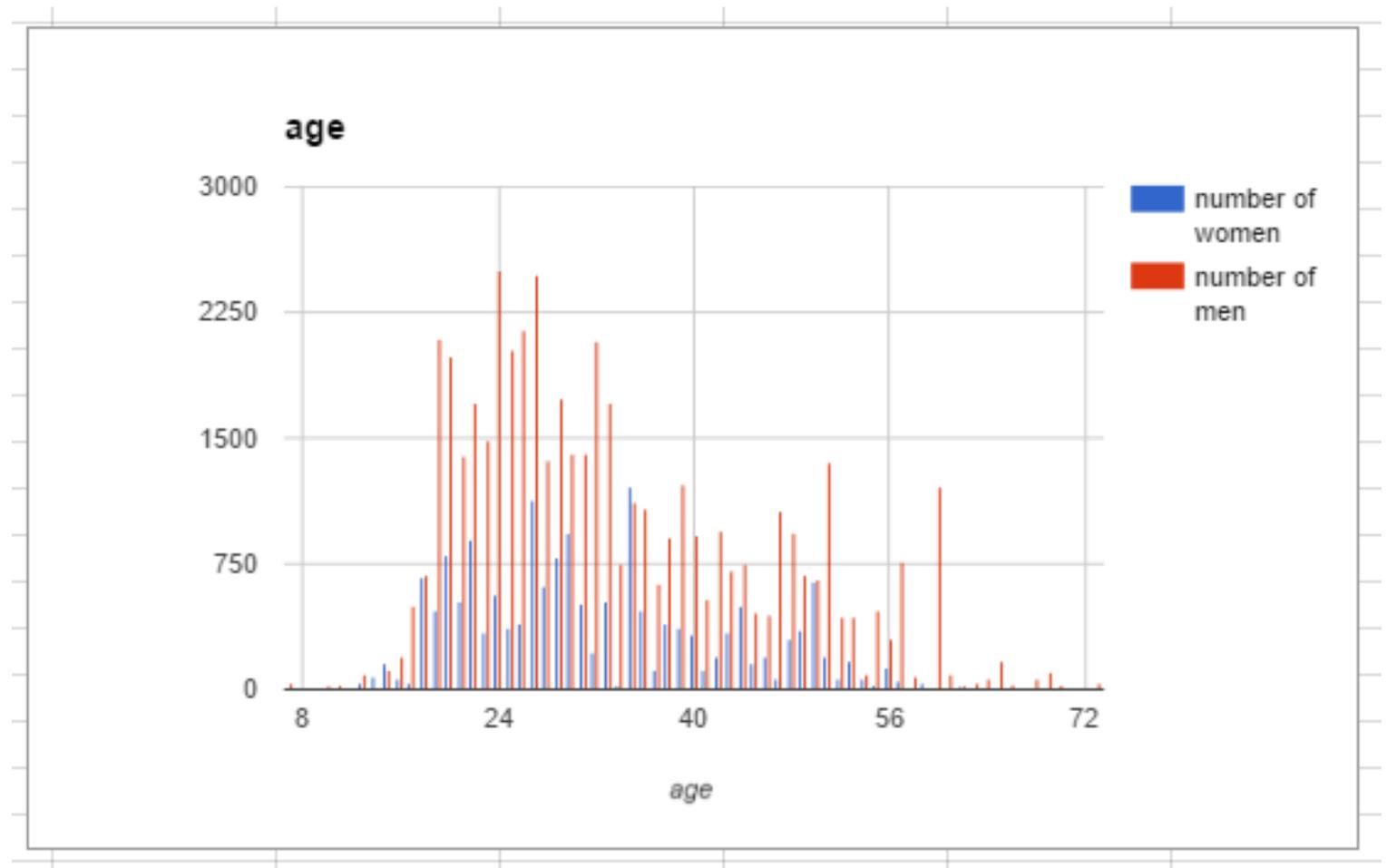
- In this exercise, you will use some different columns in the spreadsheet to create a bar chart.
- Bar charts are useful for viewing data grouped by different categories.

Bar Chart

- Select the "age," "number of women," and "number of men" columns
- Select Insert -> Chart from the main toolbar to open the Chart Editor.
- Set the chart type to "Column chart."
- "Use row 1 as headers" and "Use column A as labels."
- Click "Insert" to add your chart to the spreadsheet.

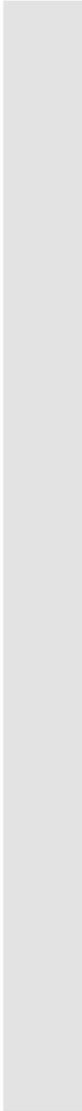
Visual Check

- At this point, your chart should look something like this:



Debrief

- Once your chart looks close to the one above, take a moment to look at your visualization.
- What does this chart help you notice about the data?
- For which ages were the number of ratings similar between men and women?
- For which ages were they different?
- What other connections and trends can you see from this chart?



Give your charts a
makeover

A makeover!

- You've now successfully learned to create charts!
- Now learn to further customize the appearance of your chart, which will make your visualization easier to read and understand.

Chart Title

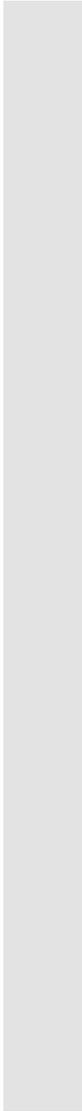
- A good chart title should effectively summarize the data story in the chart.
- Do This: Change the title of your line chart to "Average Rating by Age of Reviewer".

Axis Labels

- Your chart should include labels that indicate what the axes represent. Include measurement units, if applicable.
- Do This: Rename your horizontal axis to "Age (years)" and your left vertical axis to "Movie Rating".

Legend

- You may notice that the labels in your legend are not very official-looking. To change the text that appears on the chart, you have to change the text in the column headers themselves.
- Do This: Change the legend labels to "Avg. Rating: Women" and "Avg. Rating: Men".
- Note: If you can't read your legend labels because the text is overlapping, try adjusting the style (bold, italics, etc.) of your legend.



Now its your turn

Now it's your turn

- Now that you know the basics of how to create charts, try making some visualizations of your own!
- You can continue exploring the spreadsheet from the previous exercises, or download this additional data set:
- `MovieRating_disneyMovies.csv`
- Remember, this exercise is all about exploring trends and discovering connections in the data. Don't stress about trying to create the perfect chart. Experiment with different chart types, try using different combinations of columns, and see what else you can learn about this data!

Discuss

- What was the most interesting visualization you were able to create with the data set provided?
- What did it help you discover about the data?
 - What type of chart it was
 - What specific data it plotted
 - What it helped you discover and/or why it was the most interesting to you