Data Visualization: Bar Graphs and Histograms in Excel, Tableau, and R

Kylie Bemis
Kylie Bemis Kuiper

Data Visualization: Bar Graphs and Histograms in Excel, Tableau, and R

Slides are available online here and here. Source code, data, and PDF slides downloaded and modified.

Bar Graphs and bar charts

This ggplot2 visualization from the Thyroiscopy Trial uses a 100% stacked column chart to compare the proportion of BRAF V600E mutations across three different tumor groups. The series is either stacked "Soft Tissue" or unstacked "Thyroid Tissue", and the x-axis, "Tumors", is ordered by increasing risk.

Bar or Column Charts

Bar or column charts (related pie/donut charts from 100% series) are the most well-known, basic charts, meaning that a simple Google Image Search finds many visualizations of this type. Bar charts are used to compare discrete or categorical data, and are distinctive from binning visualizations (with no white or whitish space between bars).

Clearly-labeled bar charts offer the displaying discrete categorical data, for comparing values across categories, and for illustrating change over time (with time on the x-axis).

Bar or Column Charts

Sex
Male
Female
Transgender
Gender
Male
Female
Transgender

5. Next, drag the Bins to columns and calculated field to rows!

The easier way…

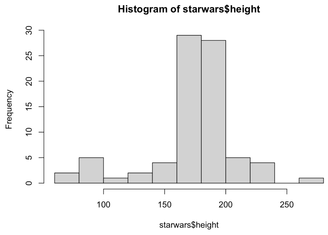
1. Click on your calculated field count of heights.
2. Click on the drop down menu Show me (right upper corner) and select Histogram.
3. Voila!
4. You can adjust colors, titles, axis labels, etc. the same way as with the bar charts!!

In R Studio:

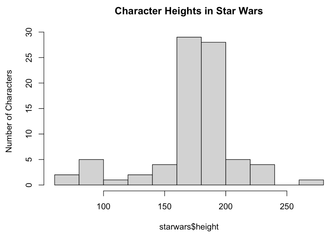
The code chunks below assumes you have already uploaded the dataset and correct libraries (as shown above for bar graphs).

The first way we'll make histograms uses the base R graphics system:

```
#using base r
hist(starwars$height)
```



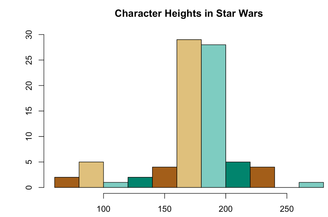

```
#add in the titles
hist(starwars$height, main = "Character Heights in Star Wars",  ylab = "Number of Characters")
```




```
#How would we add in the x axis title?
```

Try out different colors!!

```
#adjust the colors
hist(starwars$height, main = "Character Heights in Star Wars",  ylab = "", xlab = "", col = brewertheme)
```

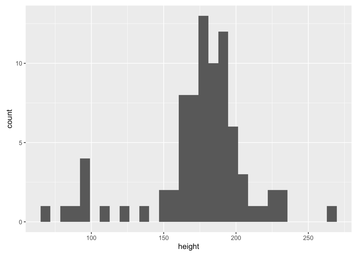


Now, we will make it with ggplot. Ggplot works by layering the data and is very customizable in this way!

```
#now make it with ggplot!!
ggplot(starwars, aes(x=height)) + geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

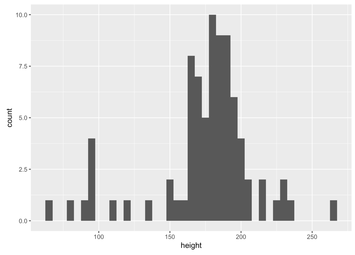
```
## Warning: Removed 6 rows containing non-finite values (stat_bin).
```



To adjust things, including the width of the bins

```
#lets adjust a few things
#change the width of the bins
ggplot(starwars, aes(x=height)) + geom_histogram(binwidth = 5)
```

```
## Warning: Removed 6 rows containing non-finite values (stat_bin).
```

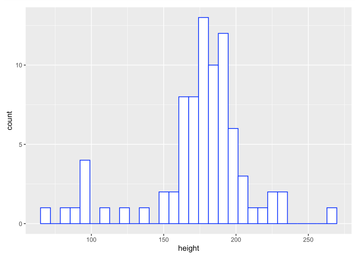


To change colors, feel free to edit the below code or test out other options below!

```
ggplot(starwars, aes(x=height)) + geom_histogram(color="blue", fill="white")
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

```
## Warning: Removed 6 rows containing non-finite values (stat_bin).
```



Looking ahead!
We will learn how to make scatterplots in Excel, Tableau, and R on November 11th, same time same place, and on the 18th of November we will learn how to make area graphs in Excel, Tableau, and R, also via zoom 4pm est.

Please reach out with any questions, comments, or feedback at katherine.kuiper25@uga.edu!

Works Cited

Cash, Ivan. https://i.imgur.com/DNxvj.png.
Chang, Agnes. 2020. Introduction to Data Visualization: Course Syllabus. Columbiaviz.github.io
Ciesliak, Jack. 13 Powerful ways to visualize your data. Sisense.com
Durcevel, Sandra. Designing Charts and Graphs: How to Choose the Right Data visualization Types. May 2nd, 2019. datapine. https://www.smithsonianmag.com/history/surprising-history-infographic-180959563/
Durcevel, Sandra. Designing Charts and Graphs: How to Choose the Right Data visualization Types. May 2nd, 2019. datapine.
ETA+. 2020. Unsplash.
Friendly, Michael. Historical Milestones. http://www.datavis.ca/
Friendly, Michael. 2008, "The Golden Age of Statistical Graphics,"Statistical Science 8(4), 502—535. https://ggplot2.tidyverse.org
Gupta, Sujata. 2020. Florence Nightingale Understood the Power of Visualizing Science. Science News. https://www.sciencenews.org/article/florence-nightingale-birthday-power-visualizing-science
Harper, Douglas. 2001. Bar: Online Etymology Dictionary. https://www.etymonline.com/word/bar?ref=etymonline_crossreference#etymonline_v_5227
Hill, Asha. 2018. 9 Useful R Data Visualization Packages for Any Discipline. Mode. https://mode.com/blog/r-data-visualization-packages/
Kiss, Zoltan. 2018. Data and Data Analysis in Linguistics. http://seas3.elte.hu/lingtheo/notes/gkiss-data-in-lcs.html.
Klemm, Rosemarie; Klemm, Dietrich. 2013.Gold and Gold Mining in Ancient Egypt and Nubia. Heidelberg: Springer. pp.132–136, 144
https://www.r-bloggers.com/2017/07/interactive-sunburst-graphs-in-power-bi-in-5-minutes/ Norman, Jeremy. William Playfair Founds Statistical Graphics, and Invents the Line Chart and Bar Chart. History of Information. https://www.historyofinformation.com/detail.php?id=2527
Open Learn. 2020. Session 3: Discrete and Continuous Data. Openlearn Edu: Uk. https://www.open.edu/openlearn/ocw/mod/oucontent/view.php?id=85587&section=1
https://vocal.media/futurism/most-powerful-star-wars-jedi
Simmons, Ryan. 2016. ENC-2135 Research, Genre & Context. https://dfb1585.wixsite.com/enc2135/single-post/2016/1/10/Blog-Prompt-1
Spiske, Markus. 2020. Unsplash
Stephanie. 2020. Statistics How To: Statistics for the rest of us!. Types of Variable.
Young, Joe. 2019. Star Wars: Characters Dataset. Version 6. Retrieved 1 October 2020. https://www.kaggle.com/jsphyg/star-wars.
Tableau. 2020. Data visualization beginner's guide: a definition, examples, and learning resources. https://www.tableau.com/learn/articles/data-visualization
Taylor, Timory. 2017. William Playfair: Inventor of the Bar Graph, Line Graph, and Pie Chart. Conversable Economist. https://conversableeconomist.blogspot.com/2017/08/william-playfair-inventor-of-bar-graph.html
Tufte, Edward. 1983.The Visual Display of Quantitative Information.
Tufte, Edward. 1990. Envisioning Information.
Tufte, Edward. 2006. Beautiful Evidence.
UC Denver Libraries. Research Guides: auraria Library. https://guides.auraria.edu/c.php?g=528873&p=3964258.
Wickam, Hadley. 2010. A layered grammar of graphics. Journal of Computational and Graphical Statistics.
Wickam, Hadley. 2016. ggplot2: Elegant Graphics for Data Analysis(Use R!). 2nd Edition.
Wiediger, Christian. 2020. Unsplash.