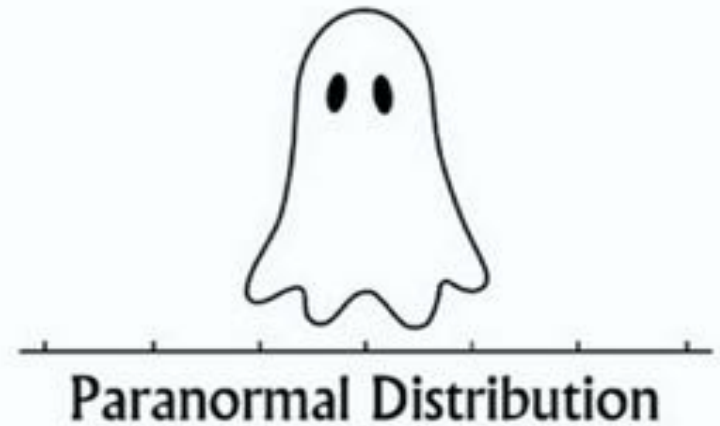
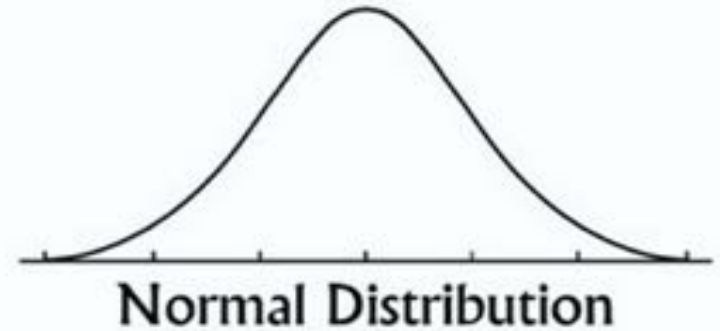


DATA VISUALIZATION: BAR GRAPHS AND HISTOGRAMS

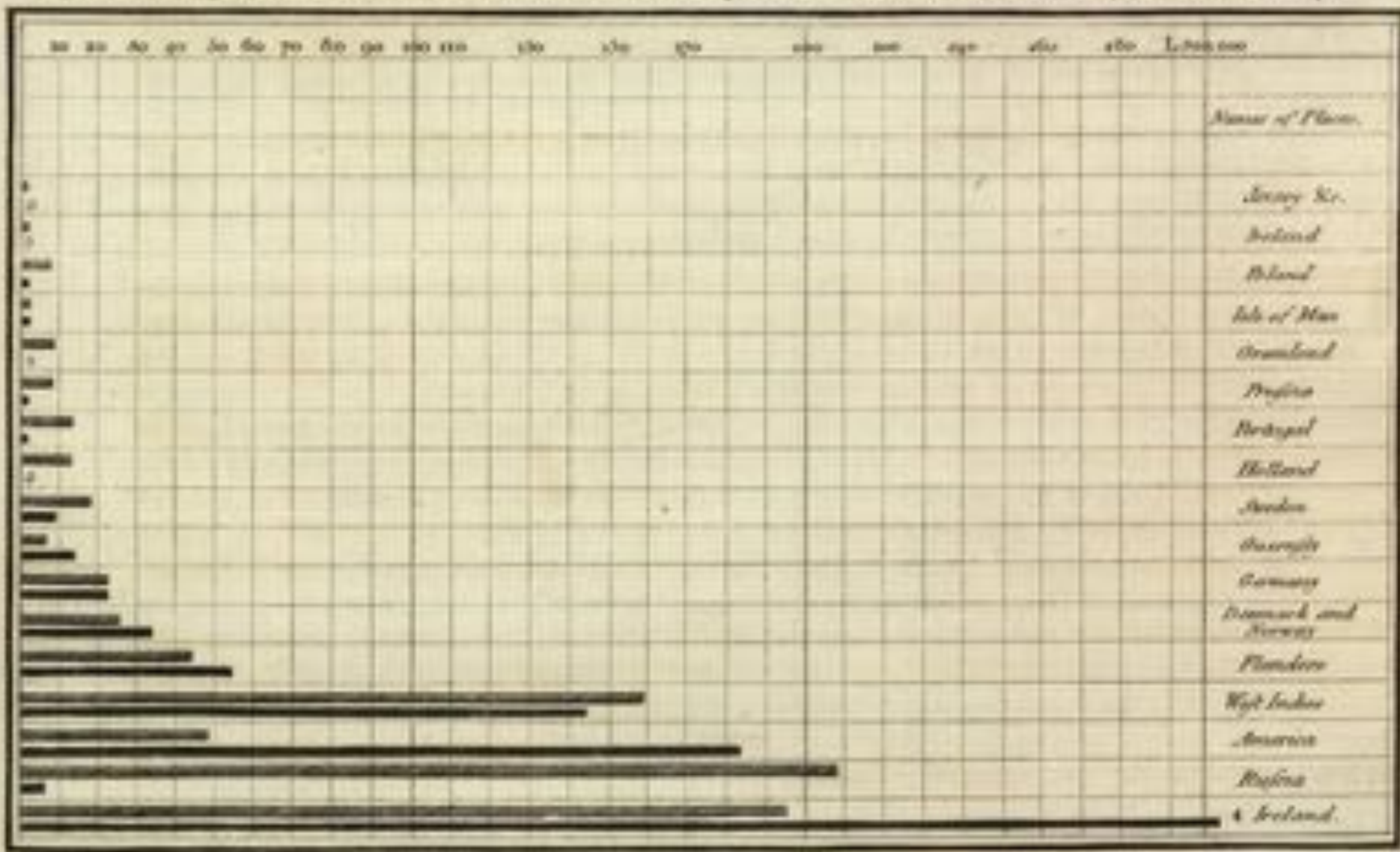
Katie Ireland Kuiper
4 November 2020
katherine.kuiper25@uga.edu



UNIVERSITY OF
GEORGIA

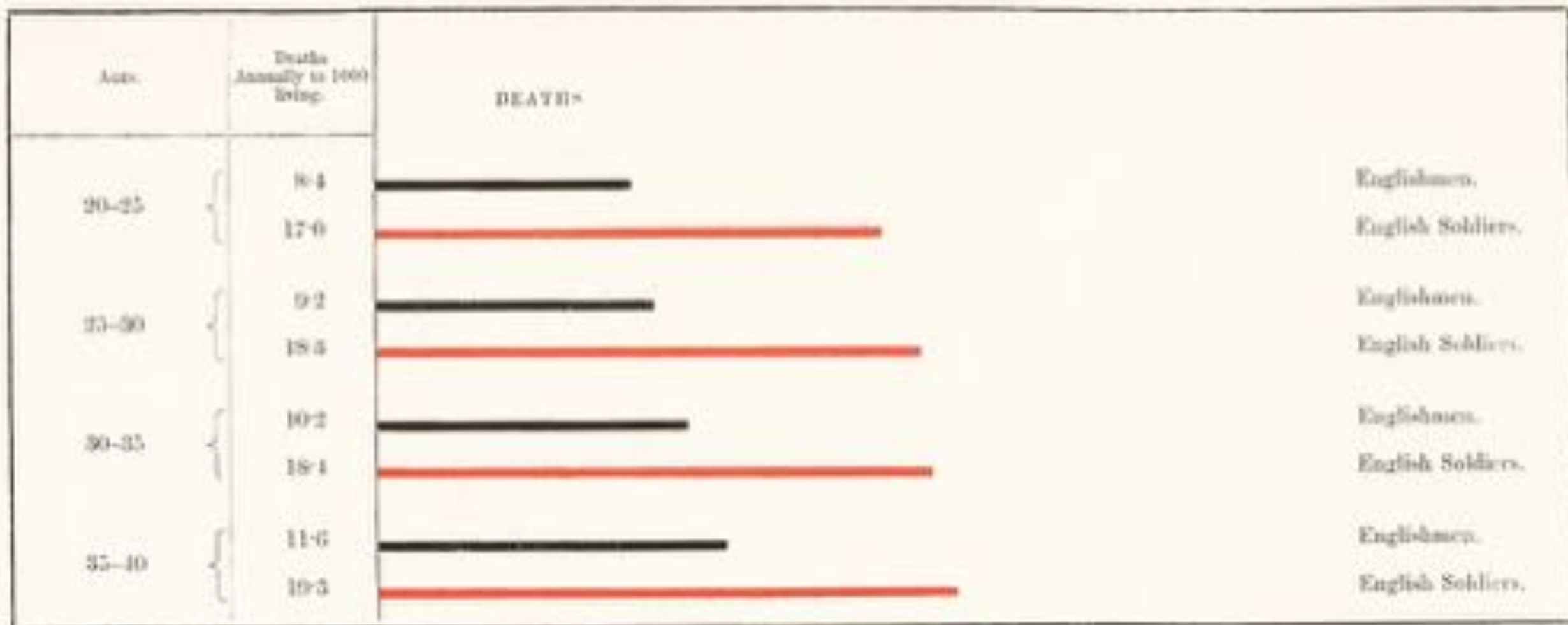


Exports and Imports of SCOTLAND to and from different parts for one Year from Christmas 1760 to Christmas 1761



The Upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed Lines Imports

Representing the Relative Mortality of the Army at Home and of the English Male Population at corresponding Ages.



From Lewis, &c.

BAR: TERM FROM
12TH CENTURY
REFERENCING A
STAKE OR ROD OF
IRON FROM OLD
FRENCH AND LATIN
BARRE, BARRA

BAR VS. COLUMN CHART

- Depending on the program you are using, these are often referencing the same type of chart.
- Or refer to differences in orientation: horizontal vs. vertical columns.

DATA TYPE(S)

- Before embarking on the creation of your bar chart, consider what type of data you are using and need to visualize
 - display discrete categorical data, compare values across categories
 - Also works with time series data to illustrate change over time

KEEP IT SIMPLE!

- X or Y axis displays a category, and the length or height of the bars are proportional to their representative values
- Make sure that it is to scale ie proportional with data values
- Consider arranging in descending or ascending order
- Color, contrast, and names of axes

THE DATA!



- Let's use this cool star wars dataset for free from kaggle.com!
- If you don't already have an account with them, you can go there now to <https://www.kaggle.com/jsphyg/star-wars>. and download this dataset
- Now we're ready!

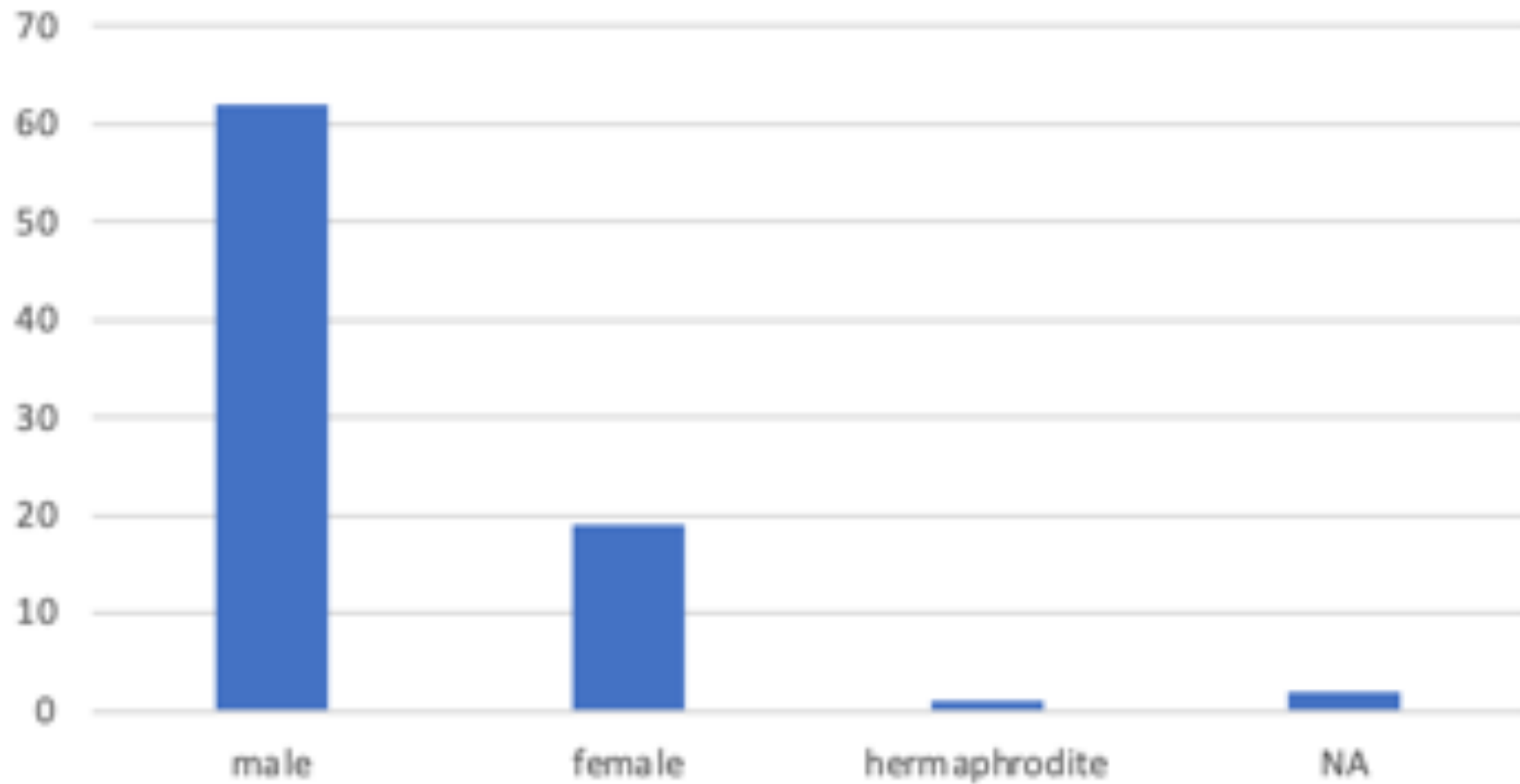
MAKE SURE YOU HAVE

- Excel
- [Tableau public](#)
- [R](#)
- [R Studio](#)

- Excel has specific naming conventions, calling vertical graphs *column charts* and horizontal graphs *bar charts*.
- Let's make them in excel now!
- Go ahead and open your dataset in excel.



Star Wars Characters by Gender





- Manually: create two new columns in your excel spreadsheet, one called gender, another called count
 - List out the possible options in the new gender column, and then in the count column we will use this formula to obtain the values from the different character genders: =COUNTIF(H:H, O2), repeat for each gender so that the item O2 changes to match each gender option
 - Select each manually created column, (and only where the values are located) go to the Excel menu Insert -> chart _> column
 - Option to also use Sort & Filter to sort it by descending or ascending values! Click sort by ascending or descending
 - Next, highlight the columns and go to the menu Insert ->chart -> column
 - Alternatively, highlight the original gender column, under the sub menu bar select Ideas and then count by gender chart
 - This will open it up in a new pivot table sheet, you can then select the values there and go to go to the menu Insert ->chart -> column

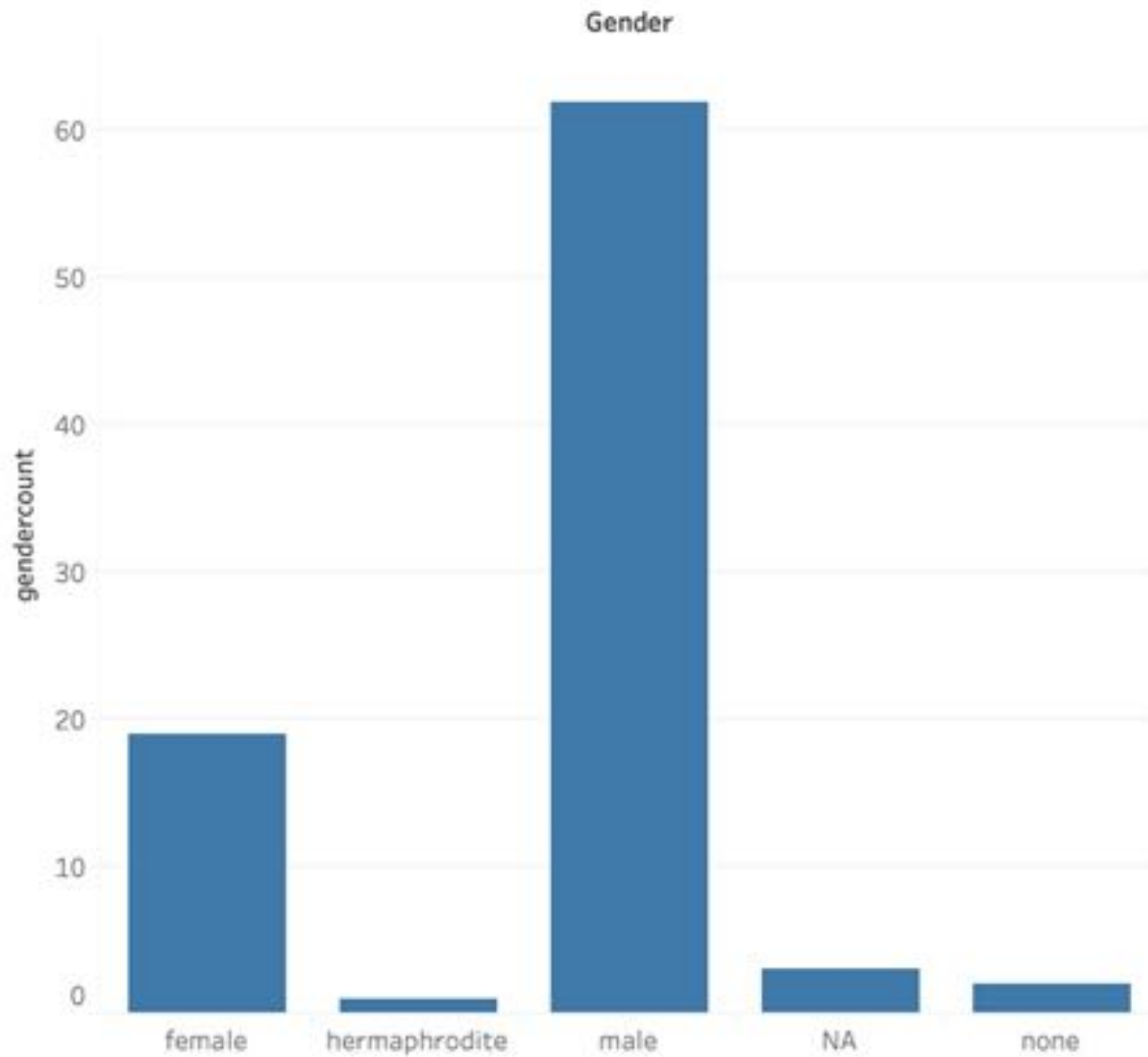
TO ADJUST.... COLORS

- Go to menu -> chart design
- Adjust colors
- To change the orientation to be vertical go to Menu Insert, Chart, Bar
- You can click into the chart names and axis names to change or delete them!



- Open tableau public
- Open text file – csv CharStarWars
- Go to Sheet 1 (bottom left corner)
- Click on Gender Under the Data menu (left side); select duplicate
- Then click on the gender (copy) and select Create -> Calculated Field
- In the formula section, delete what is written and change it to `Count([Gender])`
- You will then have a new calculated field called calculation 1, you can change the name to gendercount if you'd like by highlighting it.
- To make a horizontal bar chart, move the gendercount calculation to columns, and the original gender to rows

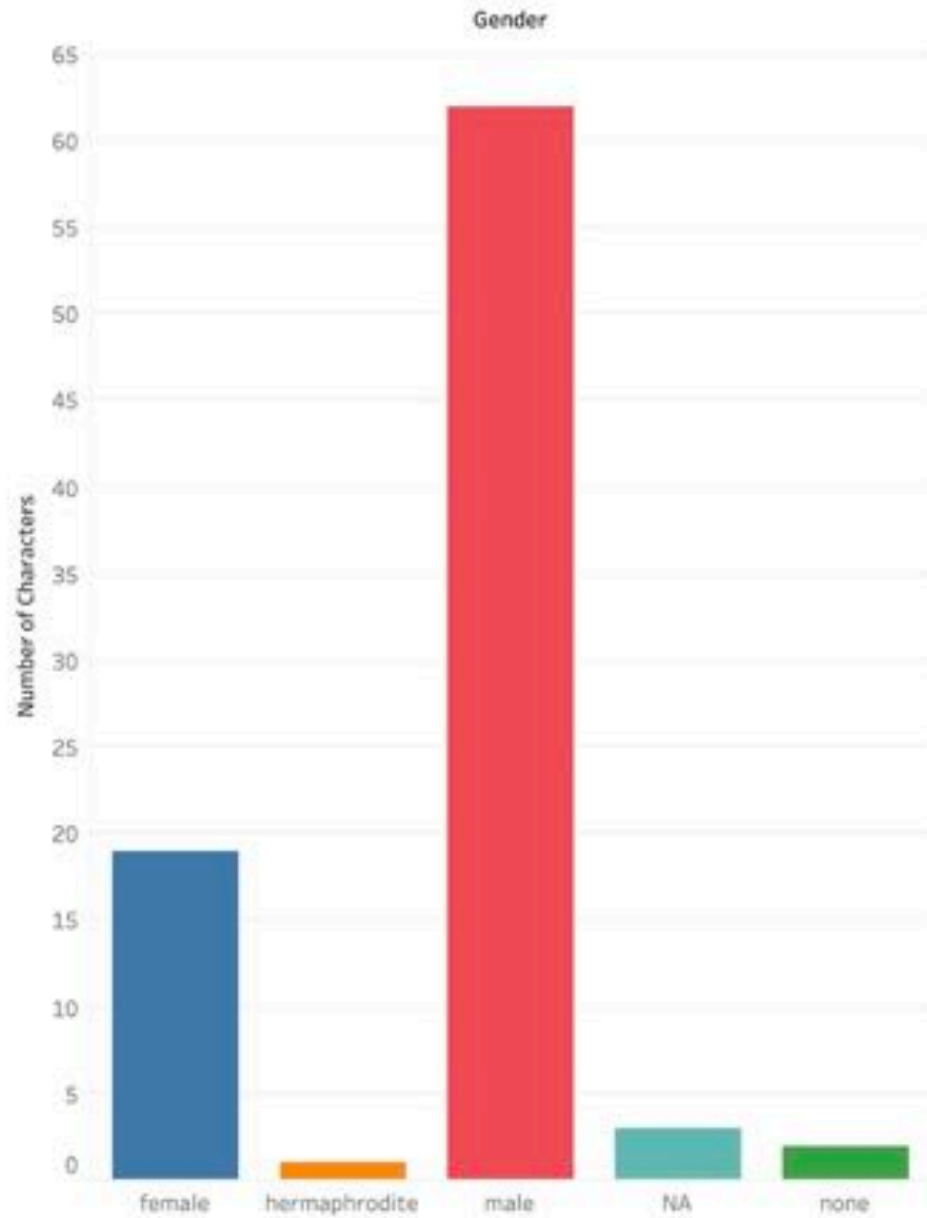
Star Wars Characters by Gender



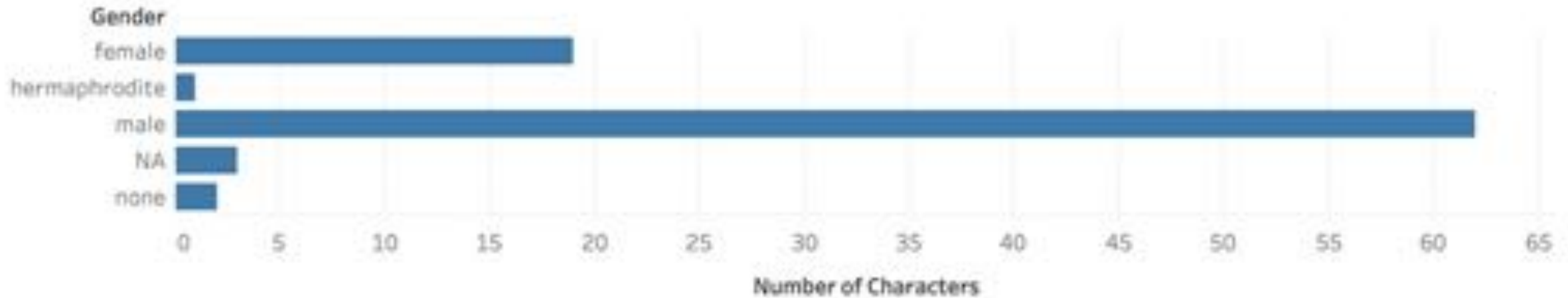


- To make vertical bar graph, move the original gender table to columns, and gender count to rows
- To adjust colors by gender you can drag the gender table to colors.
- To adjust the length of any axis, move your cursor to the edge of the axis til it becomes an arrow, click and drag the cursor to adjust this
- To change the names, click into the axis names and adjust

Star Wars Characters by Gender



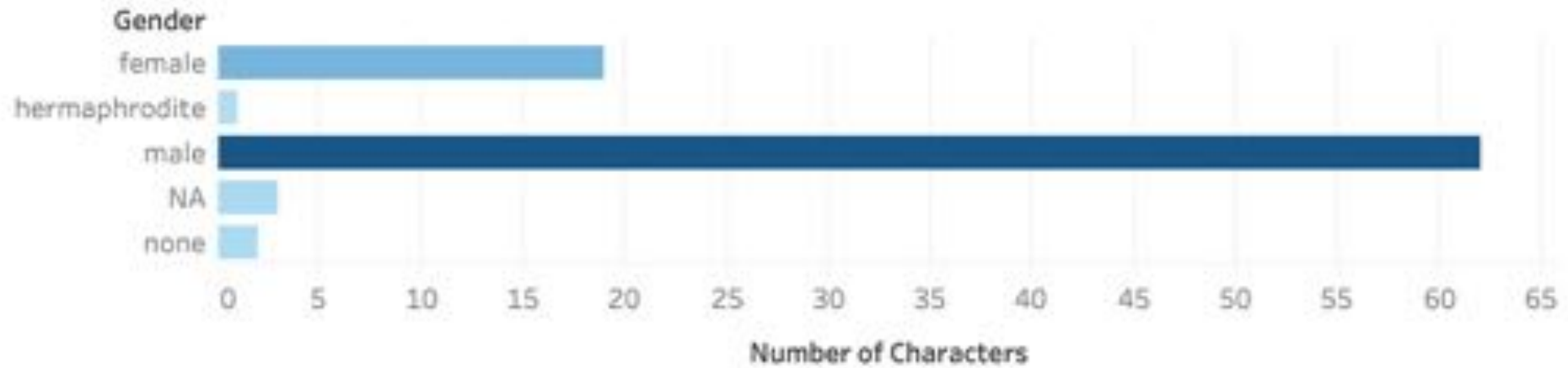
Star Wars Characters by Gender



MAKE IT HORIZONTAL!

- Remove everything from the columns and rows list!
- Now move the gendercount calculated field to columns, and the original gender table to rows

Star Wars Characters by Gender



CHANGE THE COLORS!

- Drag the original gender table onto the colors. You can also drag it onto size, etc to change different things!
- Here is shown what happens if I drag the gendercount object onto colors!
- If you click into the colors, you can change the actual color palette!!
- Have a play with it and see what happens!!



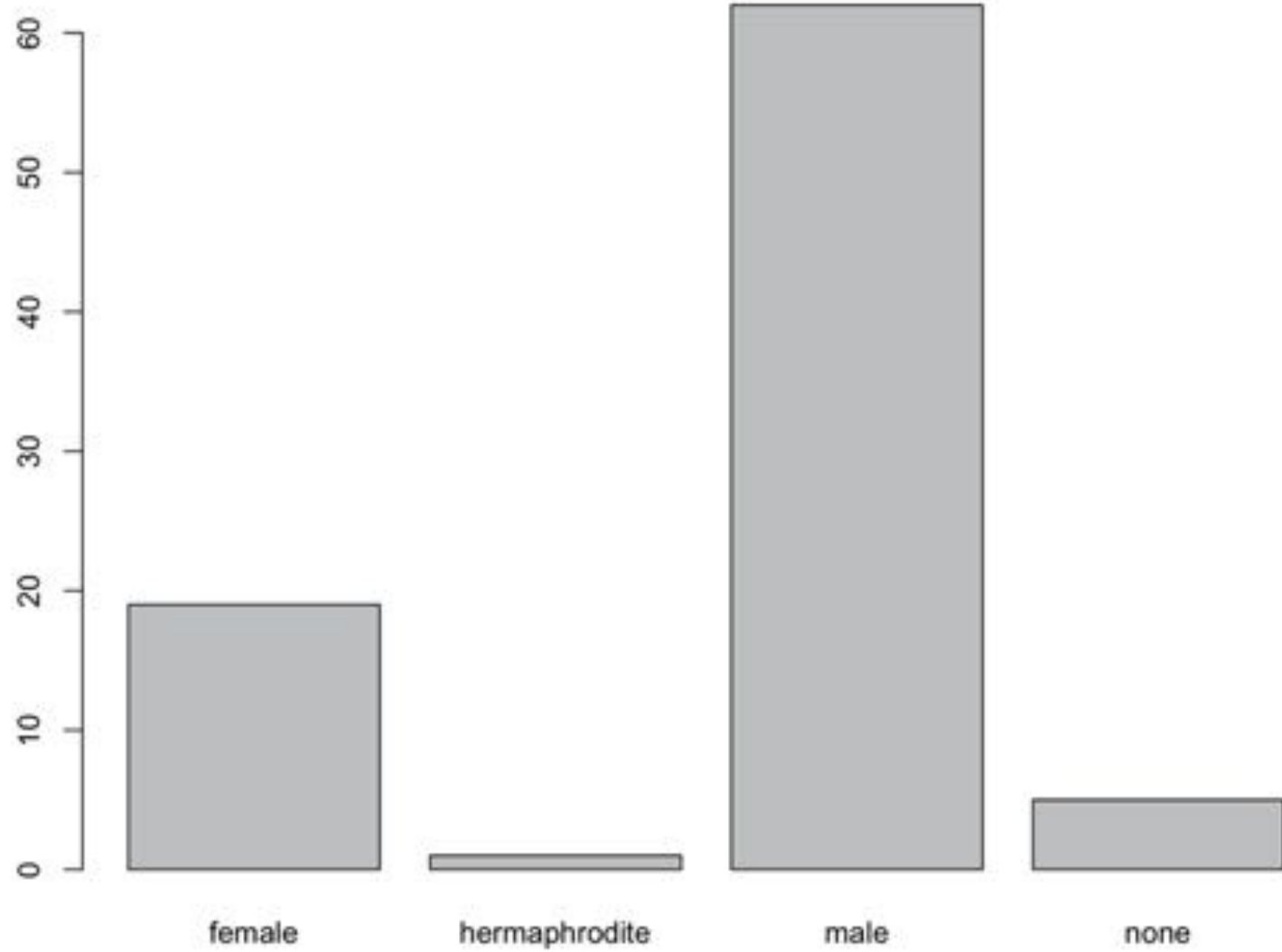
- Other chart settings can be further adjusted from the Marks Menu
- Try out selecting different marks and see what happens!
- Color, Size, Label, Detail, and Tooltip can all be adjusted there!
- If you want to publish your graph on tableau public you can save it through there. If you'd rather just screen shot the image, then it won't be automatically saved to your tableau public account.



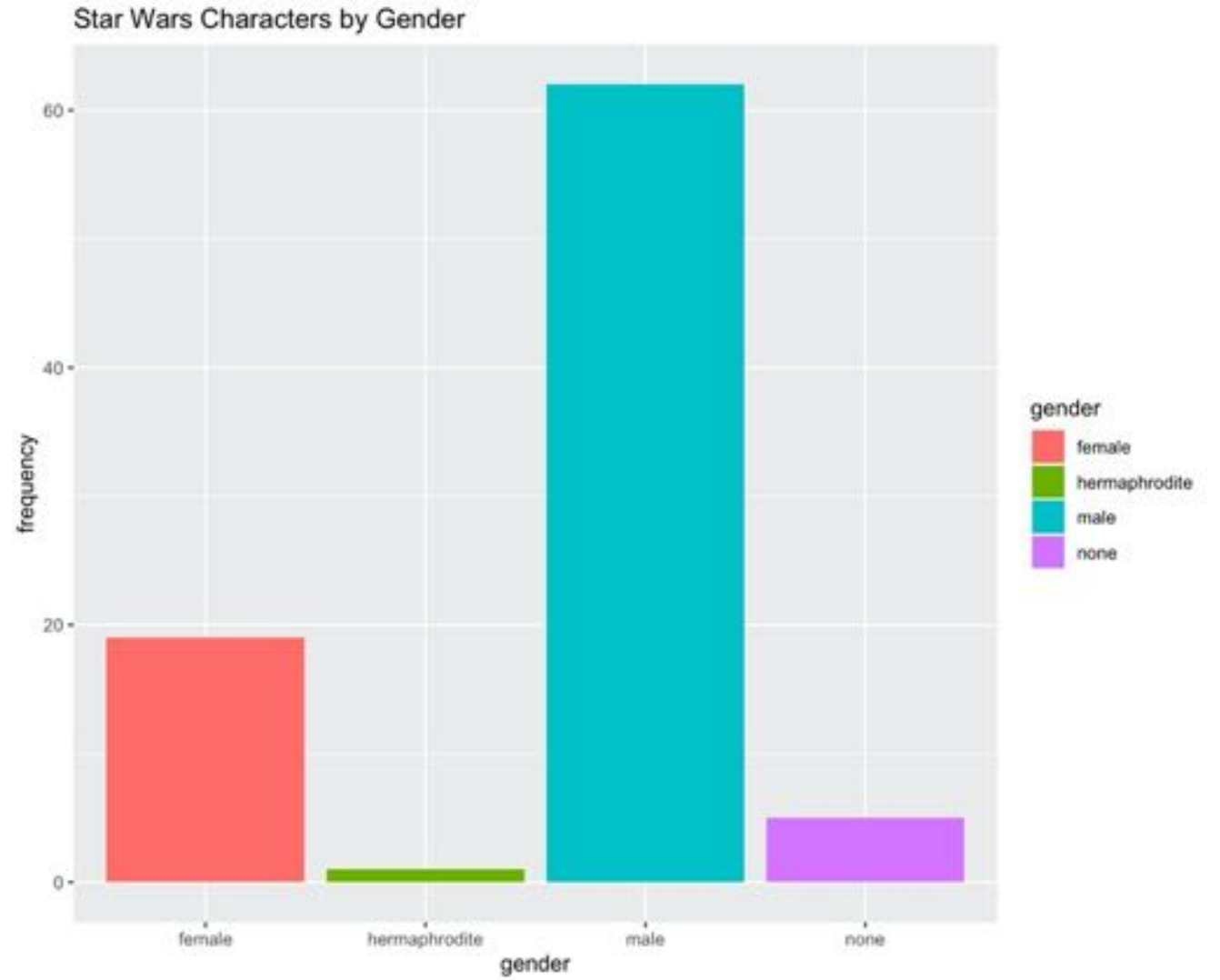
- R has a variety of options for making cool plots!! It is probably the most flexible of the three options, but also has the steepest learning curve
- Barplot and ggplot (from tidyverse) are great options for what we will make today!



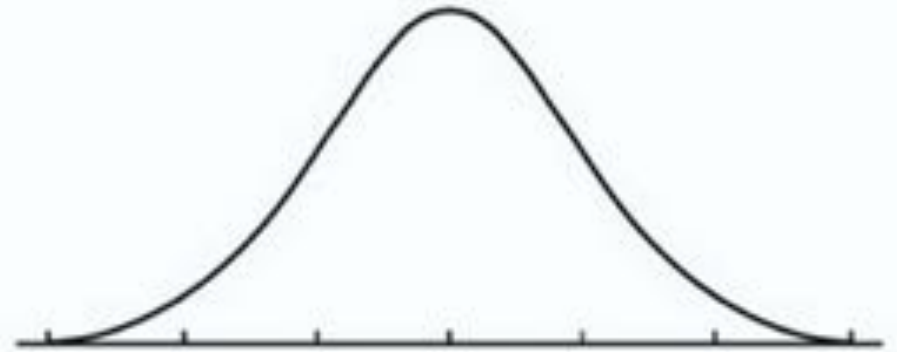
- Check out our handout to copy and paste the R code to create the bar charts!



WITH
GGPLOT2



HISTOGRAMS



Normal Distribution



Paranormal Distribution

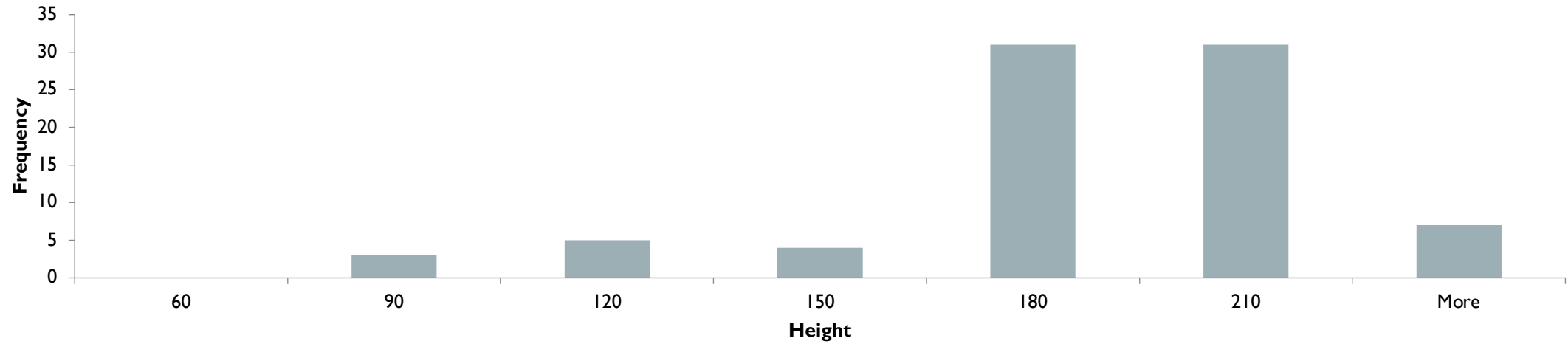
HISTOGRAMS

- Histograms look very similar to vertical bar plots, but they are different!
- Histograms are useful for:
- Visualizing the frequency distribution of continuous data (is it normally distributed?)
- Inspecting datasets
- This requires separating out the continuous data into Bins to prepare the data for visualization!
- This is easy to accomplish in each of the programs we are working with today.



- Highlight, copy and paste star wars character height and move into a new column
- Remove NAs from the new column
- Create a new column called bins and input the values according to whatever increments you would prefer, I'll be using the values in increments of 30: 60, 90, 120, 150, 180, 210
- Highlight the values in these columns, go to data analysis and select Histogram. Make sure that chart output is selected. You will have to input the values from the cells either with drag and drop or type in the values
- (If data analysis is not showing up go to Data -> Excel Add-ins -> Select Analysis Toolpack
- Edit colors and axis labels the same way as with bar charts!

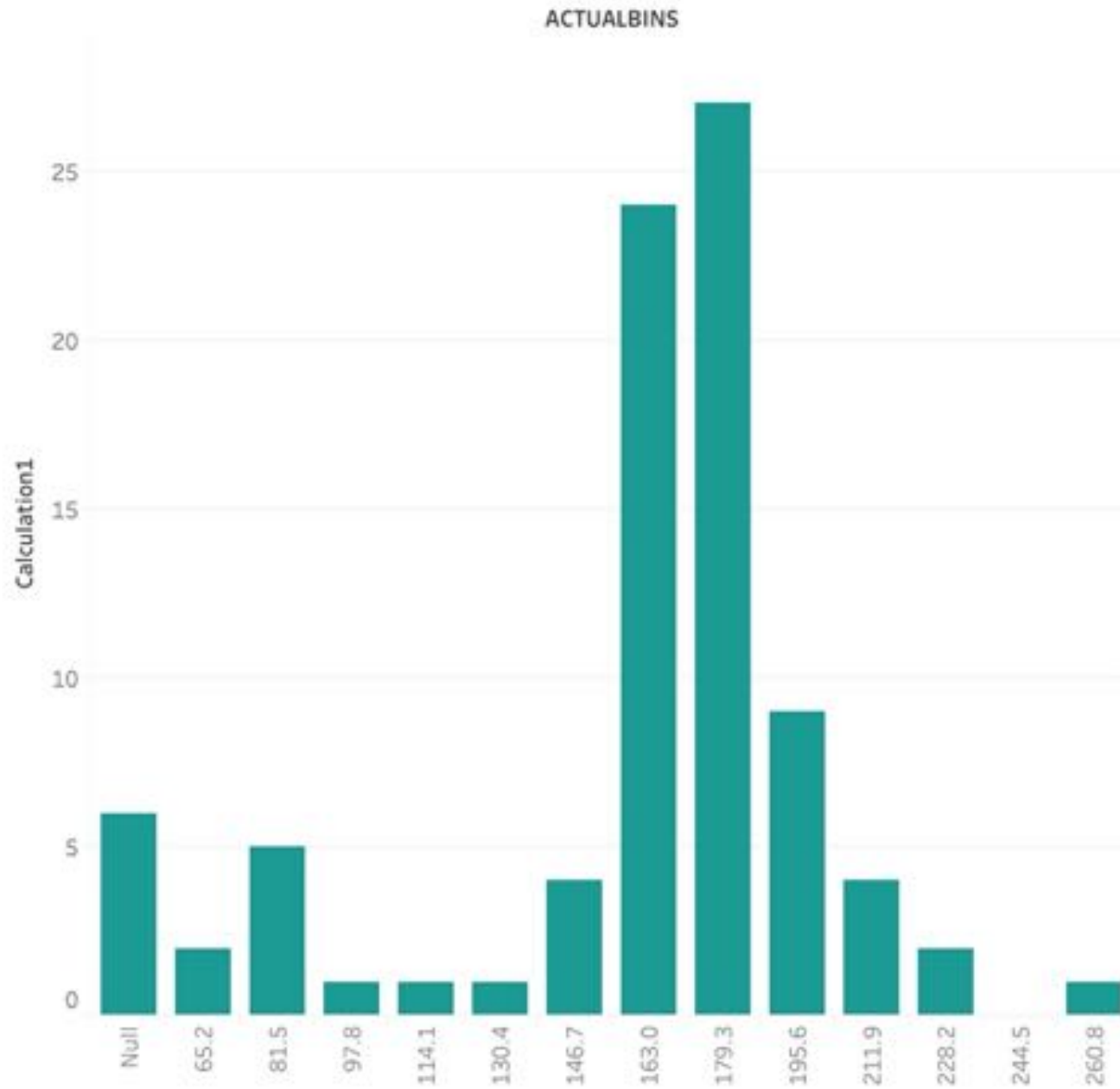
Character Heights in Star Wars





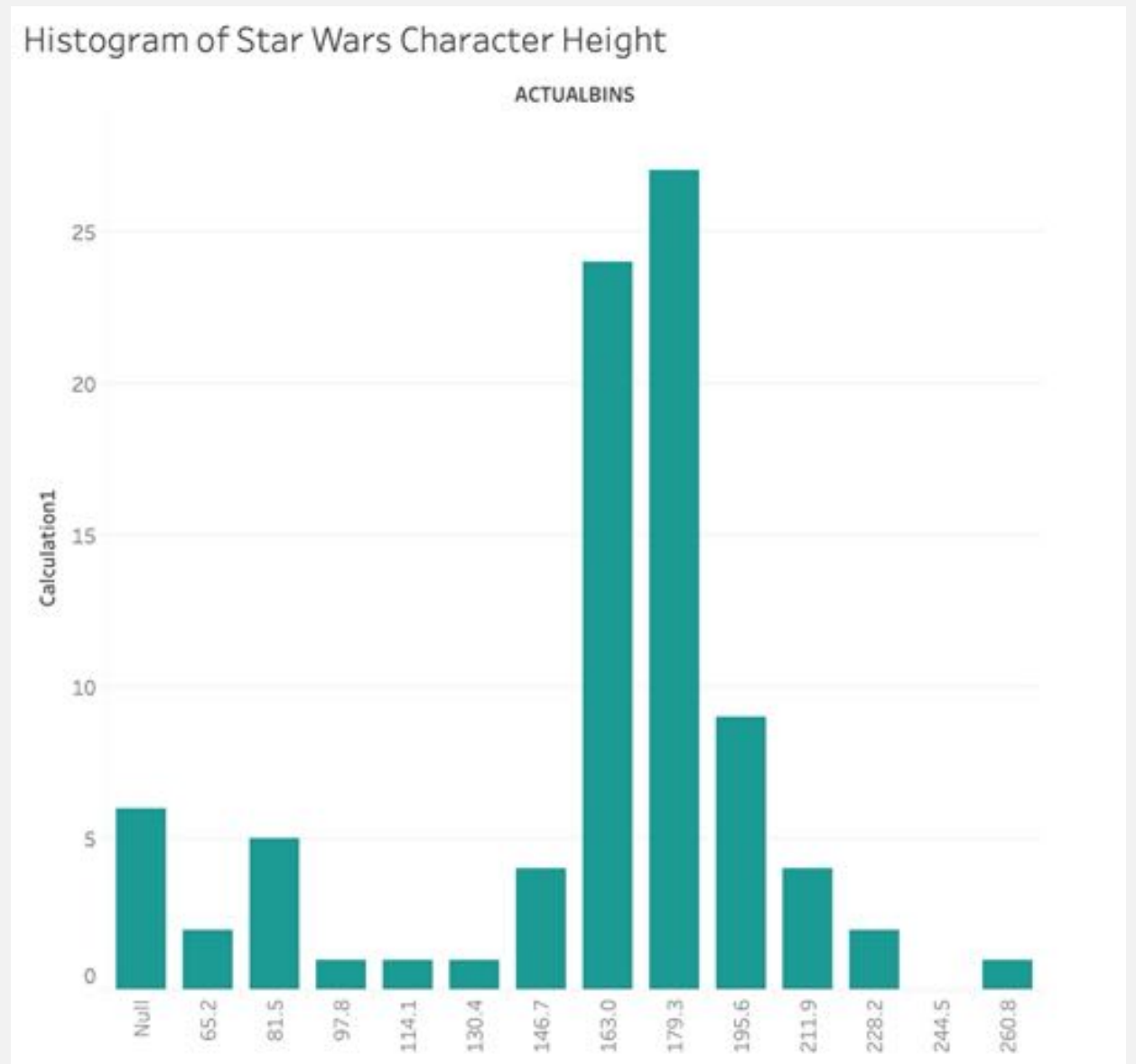
- Let's make a histogram of character heights in tableau!!
- Duplicate the height table data object.
- Make it continuous, then create...bins
- You can change the bin levels (or not); for this, I'll let it automatic!
- Then duplicate the height object once again, Create Calculated Field Count(Height), rename it to Count of Heights
- Then drag the Bins to columns and calculated field to rows!

Histogram of Star Wars Character Height





- The easier way!!
- Click on your calculated field count of heights, then in the show me column, select histogram!
- You can adjust colors, titles, etc. the same ways as with the bar charts!





- Check out the R code for what to do!!



CAREFULLY CONSIDER

Color, contrast, clutter, and scale

ADDITIONAL RESOURCES

- Kaggle.com for other datasets.
- Tableau also has freely available tutorials.
- Tutorials from the Digilab here at UGA: excel, R, and Tableau!
- R Ladies Athens!!
- [Ggplot2 cheatsheet](#) and [ggplot2: Elegant Graphics for Data Analysis](#) by Wickam

COMING UP NEXT...

11 Nov. Scatterplots

In Excel, Tableau, and R

in Excel, Tableau, and R

18 Nov. Area Graphs

Two pencils, one light blue and one dark blue, are positioned diagonally on the left side of the slide against a bright yellow background. The text is centered in a dark grey rectangular box with a white border.

THANKS FOR LISTENING!

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](https://columbiaviz.github.io)
- Cieslak, Jack. 13 Powerful ways to visualize your data. [Sisense.com](https://sisense.com)
- Durcevel, Sandra. Designing Charts and Graphs: How to Choose the Right Data visualization Types. May 2nd, 2019. [datapine](https://www.datapine.com).
<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](https://www.datapine.com).
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 Finds 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§ion=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/2016/11/10/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.