

Biodiversity data analysis workshop – Day 2

Dominic Henry & Lizanne Roxburgh
Conservation Science Unit

dominich@ewt.org.za
lizanner@ewt.org.za



**ENDANGERED
WILDLIFE TRUST**
Protecting forever, together.



**science
& technology**

Department:
Science and Technology
REPUBLIC OF SOUTH AFRICA



**National
Research
Foundation**



Biodiversity for Life

South African National Biodiversity Institute

Programme – Day 2

08h30 – 10h00	Session 1 – Introduction to R, RStudio, basics of programming
10h00 – 10h30	Tea break
10h30 – 12h15	Session 2 – Data wrangling with the tidyverse
12h15 – 13h30	Lunch
13h30 – 15h00	Session 3 – Data visualisation using ggplot2
15h00 – 15h30	Tea break
15h30 – 17h00	Session 4 – Handling spatial data in R



ENDANGERED
WILDLIFE TRUST

Data visualisation with ggplot2

- The grammar of language helps us construct meaningful sentences out of words
- The Grammar of Graphics helps us to construct graphical figures out of different visual elements.
- This grammar gives us a way to talk about parts of a plot: all the circles, lines, arrows, and words that are combined into a diagram for visualizing data

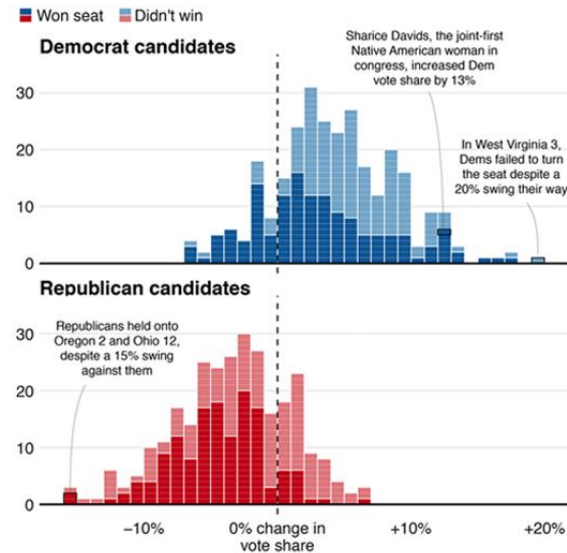
Source: https://uc-r.github.io/ggplot_intro

Versatility of ggplot2

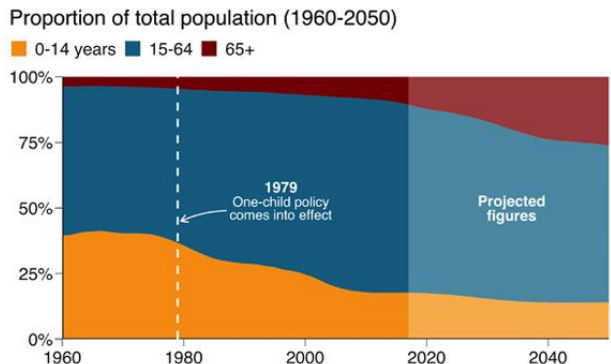


ENDANGERED
WILDLIFE TRUST

Blue wave

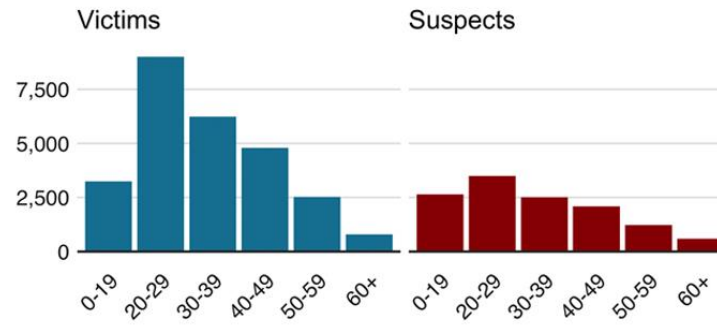


Breakdown of China's population by age group

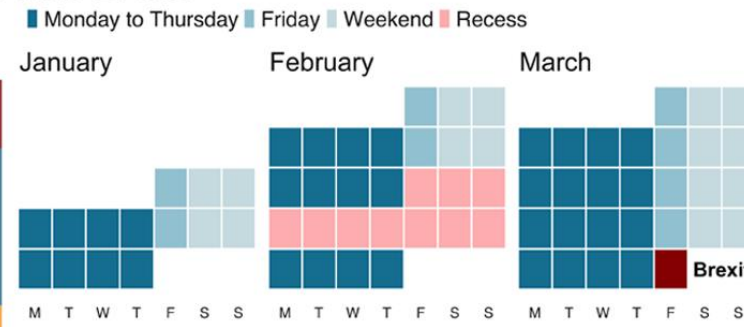


Homophobic hate crimes are mainly committed by young people on young people

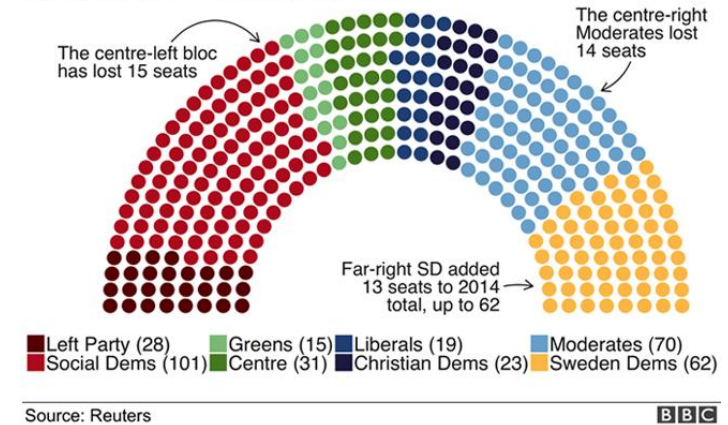
Number in each age group 2014 - 2017



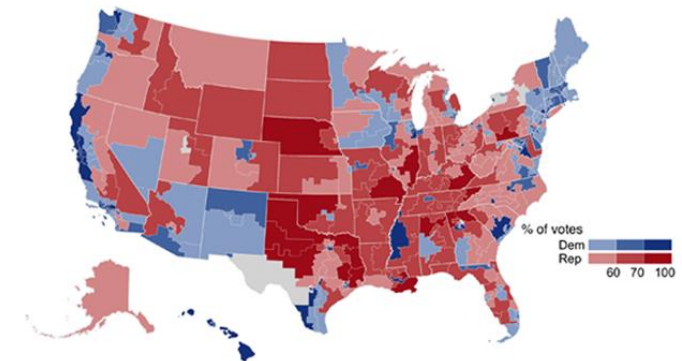
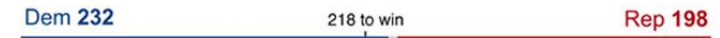
The Commons has 36 normal working days until Brexit



Results of the 2018 election



Democrats take the House





Components of a plot

- the data being plotted
- the geometric objects (circles, lines, etc.) that appear on the plot
- a set of mappings from variables in the data to the aesthetics (appearance) of the geometric objects
- a statistical transformation used to calculate the data values used in the plot
- a position adjustment for locating each geometric object on the plot
- a scale (e.g., range of values) for each aesthetic mapping used
- a coordinate system used to organize the geometric objects
- the facets or groups of data shown in different plots
- the themes control the display of all non-data elements of the plot



ENDANGERED
WILDLIFE TRUST

ggplot2 basics

- Call the `ggplot()` function which creates a blank canvas
- Specify aesthetic mappings, which specifies how you want to map variables to visual aspects (which variables relate to the x- and y-axes)
- Add new layers that are geometric objects which will show up on the plot (e.g., use `geom_point` to add a layer with points to represent the data).

Source: https://uc-r.github.io/ggplot_intro

Demo



ENDANGERED
WILDLIFE TRUST

Link to Rmd file