

Things you will need

- Google account (if you have gmail you will already have this)
- Sign up for an account with <https://datawrapper.de/>
- Make a copy of this folder (click add to drive) (<http://bit.ly/DJtutorial>) in your google drive

Do this now!

Sports injuries report

Make note of some questions such a report might be able to answer, eg

- What is the most dangerous sport in Australia, which is the least?
- Which causes the most injuries?
- Which sport causes the most severe injuries?
- Keep an eye out for interesting trends, anything unexpected

NB

Original PDF here: <http://www.aihw.gov.au/publication-detail/?id=60129549100>

Converted to spreadsheet using: <http://tabula.technology/>



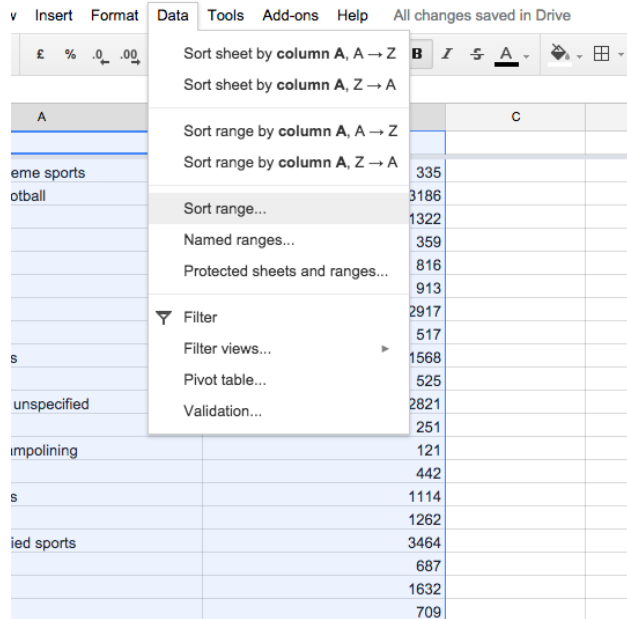
Sorting and ranking

Highlight the entire range of cells you want to sort

	A	B
1	Type of sport	Number of injuries
2	Adventure and extreme sports	335
3	Australian Rules football	3186
4	Basketball	1322
5	Boating sports	359
6	Combative sports	816
7	Cricket	913
8	Cycling	2917
9	Dancing	517
10	Equestrian activities	1568
11	Fishing	525
12	Football, other and unspecified	2821
13	Golf	251
14	Gymnastics and trampolining	121
15	Hockey	442
16	Ice and snow sports	1114
17	Netball	1262
18	Other and unspecified sports	3464
19	Racquet sports	687
20	Roller sports	1632
21	Rugby League	709
22	Rugby Union	262
23	Rugby, unspecified	1650
24	Soccer	2962
25	Touch football	644
26	Walking and running	878
27	Water sports (individual and team)	2143
28	Wheeled motor sports	2737
29		

Sorting and ranking

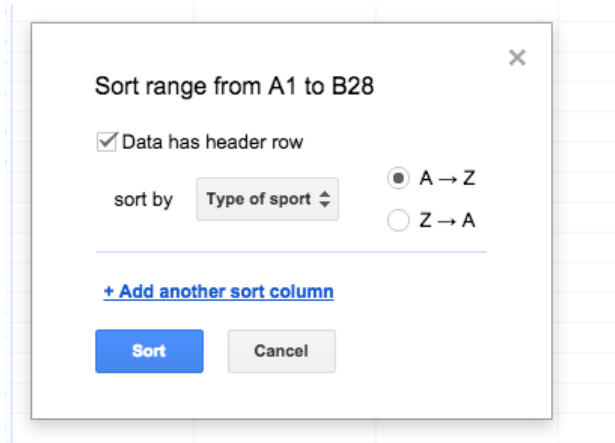
Select sort range from the data menu



The screenshot shows the Google Sheets interface with the 'Data' menu open. The 'Sort range...' option is highlighted. The spreadsheet data is as follows:

A	C
eme sports	335
otball	3186
	1322
	359
	816
	913
	2917
	517
s	1568
	525
unspecified	2821
	251
impolining	121
	442
s	1114
	1262
ied sports	3464
	687
	1632
	709

Tick "data has header row" then sort!



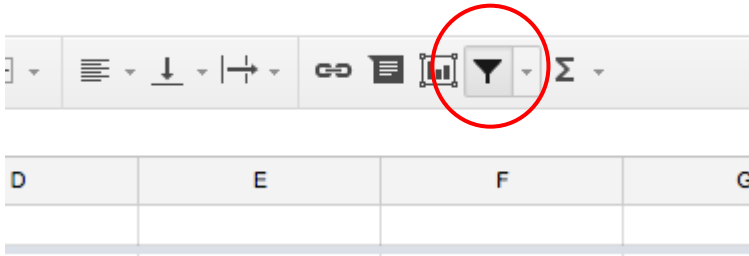
The dialog box shows the following configuration:

- Sort range from A1 to B28
- Data has header row
- sort by: Type of sport
- A → Z
- Z → A
- [+ Add another sort column](#)
- Buttons: Sort, Cancel

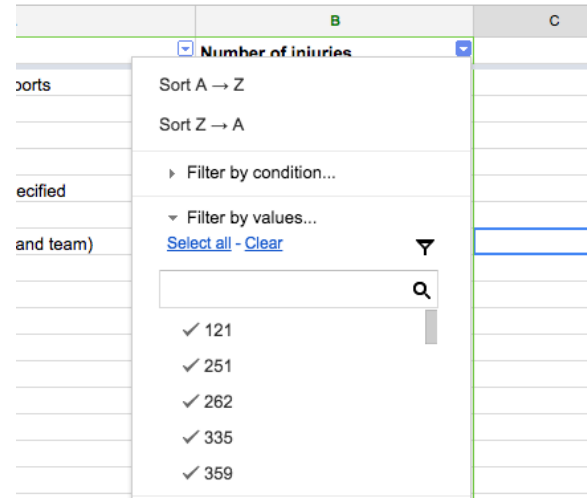
Sorting and ranking

OR

If you'll be sorting multiple times by different columns, use the filter button



Then use the arrows at the top of your columns



Spreadsheet functions and formulas

- Starting an entry with "=" makes it a formula or a function
 - eg. "=A1+B1" will sum cells A1 and B1
 - Can do most calculations, like times, division, subtraction, etc.
 - Built in spreadsheet functions also that do a lot more
-

Useful functions

=SUM()

Sums a range of cells or column

=AVERAGE()

Gives the average of a range of cells or column

=(B2/\$B\$30)*100

Gives a percentage, where cell B30 is your total. Dollar signs allow you to cut/paste function elsewhere, and it will always refer to B30

Exercise 1

1. Add a filter to the sports injuries figures and sort it by the sport with the highest injury
 2. Calculate the total number of injuries
 3. Calculate the average number of injuries
 4. Calculate the percentage of injuries for each sport
-

Exercise 1

- Which sport has the highest number of injuries?
 - Which are above average?
 - Can you describe your results in plain language?
-

Exercise 1

Hypothetical first two pars:

Australian rules football is responsible for the largest number of sporting injuries, according to a new report.

New figures released by the Australian Institute of Health and Welfare (AIHW) show that Aussie rules accounted for 3,186 injuries, closely followed by soccer, which resulted in 2,962.

The importance of rates

- Totals, rankings and simple percentages only one part to the story
 - Our work so far can't tell us which sport is most **likely** to cause injury, only which causes the **most injuries**
 - Is soccer really more dangerous than rugby?
-

The importance of rates

- No, it's not - it's just that more people play soccer
 - We need to adjust the number of injuries by the number of people who play the sport to get a **participation rate**
 - Most commonly you'll use this to adjust for population disparity - eg. crime figures
-

Exercise 2

1. Use the filter button to explore the new rates and other statistics.
 2. Find anything interesting?
-

Exercise 2

- Sorting by participation rates shows us most dangerous sports, ie. if you play these you're more likely to be injured
 - Looks more like what you'd expect
 - Other interesting things - cycling (high threat to life), golf (mean days in hospital)
-

Sports injuries

Datablog

Which sport is Australia's most dangerous?

The findings of a new report on sporting injuries are both surprising (who would have thought golf is a hidden killer?) and expected (stay away from quadbikes), but how do the major codes stack up?

Nick Evershed

 @NickEvershed

Tuesday 4 November 2014 13.57 AEDT



 Shares  Comments

801

41



[Which sport is Australia's most dangerous?](#)

Exercise 3

- We can see some interesting trends in our spreadsheet, but we need to make it easier to interpret
 - Graphing the results is the best way to do this
 - You can make graphs in google sheets or datawrapperr (or many other programs)
-

Exercise 3

Paste in the data you want to graph

1 Upload Data

2 Check & Describe

3 Visualize

4 Publish & Embed

It all starts with your data...

Select your data (including header row/column) in Excel or OpenOffice and paste it in the text field on the right. You can also upload a CSV file from your computer. [Learn more about how to upload your data.](#)

If you just want to try Datawrapper, here's a list of some example datasets you can use:

Select a sample dataset

Sport	Number of cases	Rate per 100,000	Participation-based rate per 100,000
Adventure and extreme sports	335	2	
All Rugby	2621	15	1292
Australian Rules football	3186	18	1319
Basketball	1322	8	373
Combative sports	602	4	168
Cricket	913	5	254
Cycling	2917	16	214
Dancing	517	3	199
Equestrian activities*	1568	9	902

or upload a CSV file

Upload and continue >

Exercise 3

Add a source and link

Make sure the data looks right

Please make sure that Datawrapper interprets your data correctly. In the table number columns should be shown in blue, dates in green and text in black.

First row as label

Where did you get the data?

Who published the data in the first place?

If possible, please provide a link the source data

Checking the square at the top of columns allows you to edit column types and add units

	B	
	Number of cases	Rate p
rts	335	
	2,621	
	3,186	

Exercise 3

You could, for example, add a % sign or show numbers to nearest million

Column options for Number of cases

Column type

auto (Number)

Number format

Divide numbers by

Round numbers to

keep original

Prepend/Append

#

Hide column

Hide column from visualization

< Back

Revert

Proceed >

Chart type is fairly self explanatory - use a bar chart for the sport injury stats

The screenshot shows a chart configuration interface with the following elements:

- Navigation tabs: **Chart type** (selected), Refine, Annotate, Design.
- Chart type grid:
 - Bar Chart (selected)
 - Column Chart
 - Grouped Column Chart
 - Stacked Column Chart
 - Line Chart
 - Pie chart
 - Election Donut
 - Donut chart
 - Data Table
 - Map (beta)
- Hint text: **Hint:** In case the visualization doesn't look like you expected, you should try to ↕ [transpose the data](#).
- Bottom navigation: < Back, Proceed >

Exercise 3

Check "automatically sort bars"

Chart type Refine Annotate Design

Base color █ customize colors...

Automatically sort bars

Reverse order

Use the same scale for all columns

Filter missing values

◀ Back Proceed ▶

Give your chart a title. Maybe highlight soccer as it's a key part of our story

Chart type Refine Annotate Design

TITLE

DESCRIPTION

NOTES

Highlight the most important elements (optional)

- select element - ▾

✕ Soccer

Where did you get the data?

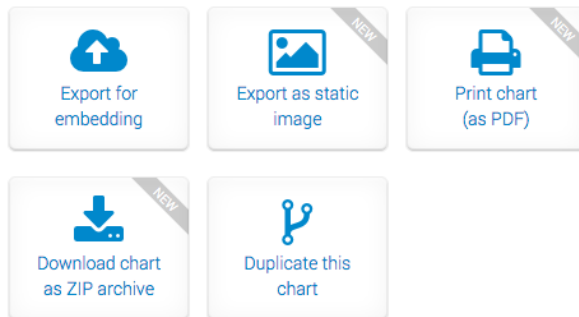
SOURCE NAME SOURCE URL

◀ Back Proceed ▶

Exercise 3

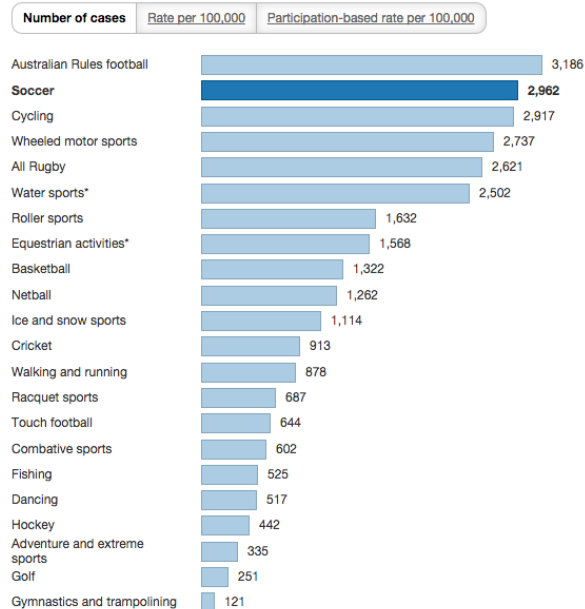
The export image button doesn't work very well. Make a screenshot instead (shift + command + 4)

What do you want to do with this chart?



Exercise 3

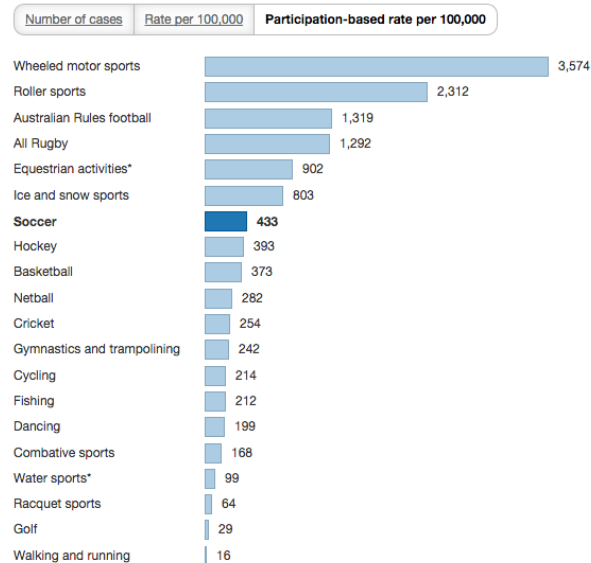
Sporting injuries in Australia



Source: [AIHW Get the data](#)

Created with [Datawrapper](#)

Sporting injuries in Australia



Source: [AIHW Get the data](#)

Created with [Datawrapper](#)

Bonus: some useful functions

=trim()

Removes whitespace

=split()

Splits text based on a character

=concatenate()

Joins text

=vlookup()

Super function! Joins two datasets based in common column

Reference for all functions: <https://support.google.com/docs/table/25273?hl=en>

Bonus: other free visualisation sites

Plot.ly - <https://plot.ly/>

Highcharts - <https://cloud.highcharts.com/>

CartoDB - <https://cartodb.com/>

Mapbox - <https://www.mapbox.com/>

Infogram - <https://infogr.am/>

Tableau - <https://public.tableau.com/s/>
