Lesson 5: Clean it up

Computing

Data science

Ben Garside

Materials from the Teach Computing Curriculum created by the National Centre for Computing Education

Task 1 - Clean it up

The following link takes you to a data set about zoo animals:

ncce.io/dsci-5-a0-hd

Spend some time looking at the data and see whether it is ready for analysis.

Check to see if any data needs correcting. If so, how might you correct it?

Name	Animal Type	Weight (Kg)	Height (m)	Age	Gender
Echo	Elephant	5900	3.3	45	Female
Stretch	Giraffe	800	5.9		Male
Yakov	Meerkat	0.73	0.32	11	Male
Maiya	Meerkat	0.67	0.33	11	Female
Vassily	Meerkat	0.69	0.32	4	Male
Bogdan	Meerkat	0.76	3	1	Male
Spot	Giraffe	820	5.1	1	Male
Batyr	Elephant	6000	3.2	32	Male
Sher	Lion		1.1	7	Female
Lavi	Lion	130	1.2	4	Female
Sarabi	Lion	124	1.3	4	Female
Drona	Elephant	0	3.1	9	Male
Alexander	Meerkat	0.71	3	4	Male
Scar	Lion	190	1.3	8	Male
Beo	Giraffe	750	4.9	5	Female
Alexander	Meerkat	0.71	3	4	Male

Task 2 - Data upload

- Save your spreadsheet as a CSV file
- Open the CODAP website (codap.concord.org)
- Create a new document by selecting Try CODAP
- Upload the CSV file

If you don’t have a data set to use, you can use the following data. This data set is a collection of data recorded about litter dropped in a school.

ncce.io/dsci-5-rs

Task 3 - Data analysis 1

Now that you have uploaded your data to CODAP, it's time to analyse it and create a visualisation that helps you to answer **one** of the questions that you posed last lesson.

- Write one of the questions that you posed last week into the table on the next slide
- Insert an image of a visualisation that helps to answer the question
- Write a few sentences to describe what you have learnt from the visualisation and whether it helps you to answer the question

Task 2 - Data analysis 1

Question	

Findings: *write a few sentences describing what you have learnt from this visualisation*	

Paste a screenshot of your visualisation in the box below