Rycroft-Smith & Macey
Research Group Ltd

Maths Week Scotland

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

Stats in the Wild

Activity ①
Creature Features

Data moves

Grouping (using, creating or highlighting subsets)

Filtering (showing/ hiding subsets)

Ordering (sorting into an order)

Summarising (computing or calculating to describe a characteristic of a dataset)

Linking (identifying corresponding case(s) in one representation and another representation)

Inspecting (hovering, clicking on or locating an object to gain information)

Expanding datasets (adding data, merging or joining datasets)

Creating new variables (e.g. rates/ proportions from existing data)

Data forms

Choosing or creating a representation for a purpose

Finding and using relationships or patterns

Adapted from Hudson, R. A., Mojica, G. F., Lee, H. S., & Casey, S. (2024) *Data Moves as a Focusing Lens for Learning to Teach with CODAP.* Computers in the Schools, 1–26. https://doi.org/10.1080/07380569.2024.2411705

Overview | approx. 2-4 hours in total

1 **Explore** decision trees and categorisation schemes

2 **Collect** data in the wild

3 **Create and tidy** a dataset

4 **Create** your own decision trees and categorisation schemes

5 **Evaluate and critique** others' work

You will need | Video notes: bit.ly/StatsWild1

MWS Activity 1 Arthropod Photographs.pdf **(printed)**

MWS Activity 1 Classification schemes.pdf **(printed)**

Camera **(1 per student or group)**

Printer **(preferably colour)**

Ruler or tape measure

Seek app **(pre-downloaded)**

CODAP Software **(optional)**

Rycroft-Smith & Macey
Research Group Ltd

Maths Week
Scotland

ROYAL
STATISTICAL
SOCIETY
DATA | EVIDENCE | DECISIONS

Why do this activity?

In this activity students are encouraged to work with and develop their own **classification schemes** by identifying features of arthropods that they can find in their local environment.

The activity provides students with an opportunity to consider how data is created by reducing an object in the real world to **a set of countable or measurable characteristics**.

Categorisation is an important mathematical idea and the application of categorisation for the purpose of classification is a key idea in **data science**.

Possible approach

1 **Explore** decision trees and categorisation schemes

Encourage students to **look at the images** and **use the categorisation schemes** to try and identify which arthropods are in each image. The activity itself should be fairly straightforward; the focus should be on deciding which of the categorisation schemes are more useful, and why.

In particular, support students in noticing the features of the categorisation schemes that **make them useful**, and any particular possible **points of failure**. For example, the categorisation schemes here are expressly designed to only work with this dataset (the five photographs of these arthropods) – using them with other data would likely result in misclassification.

Other issues might be some **classification decisions are subjective**. One example of this is that people do not see colours in the same way (for example visual impairments such as colour vision impairment), meaning classifying an arthropod by colour might be less user-friendly for some. Another might be the idea of a 'flat' or 'rounded' body, which might mean something different to different users.

Students may also know these arthropods by different names, which is a nice opportunity to discuss biological naming in general and the prevalence of many different 'common' names for species. For example, the species referred to as the European garden spider in the materials (*Araneus diadematus*) is also known as a **cross orbweaver**, **cross spider**, **diadem spider**, **orangie**, **cross spider**, and **crowned orb weaver**, or a **pumpkin spider**. This could prompt an interesting discussion with students regarding the reasoning for using Latin names (more specific and unique) as opposed to common names (easier to pronounce, remember and use).

Rycroft-Smith & Macey
Research Group Ltd

Maths Week
Scotland

ROYAL
STATISTICAL
SOCIETY
DATA | EVIDENCE | DECISIONS

2 **Collect** data in the wild

In this section, students go outside and **find arthropods to take pictures of**. This may require a little planning in terms of finding outdoor areas that are safe, bounded, not too far away and still offer opportunities to see local arthropods, for example under rocks or paving slabs, in compost heaps, on colourful flowers, or in leaf litter.

> ! Students should be instructed to take care **not to disturb** wildlife any more than is necessary, and to **not remove** any creatures from their habitat.

In small groups, students are aiming to take photographs of 5-10 different species, and may need reminding to photograph them next to a ruler or tape measure so they can accurately measure/estimate length, as well as taking multiple photographs of each arthropod so they can best identify its features.

Taking photographs of arthropods is hard! They are often small, move around, and holding a ruler in the shot can pull camera focus from the arthropod itself. Encourage students to work together to take the best photos they can, and to discuss how zoomed-in, well-focused and clear the photos need to be to be used in this activity.

Students can then use the *Seek* or Google Lens apps (either in the field if pre-downloaded, or back in the classroom) to **identify** each arthropod and record this information. *This app does not require registration or user details and is for users aged 12+.*

When students return to the classroom, it is important that they double-check the names of the arthropods they have photographed by researching them online, or using other sources of information such as books, since apps can misclassify and it is good practice to double-check classifications.

3 **Create and tidy** a dataset

Students then choose to **print 5 or 6 of their photographs** (preferably in colour). They should choose ones which are best suited to making a classification diagram – you might suggest properties such as a clear photograph where features are visible. They then group their photographs by considering which features are the same and which are different, for example:

Length	Main colour	Whether their bodies are split into sections or not
Whether they have wings or not	How many legs they (usually) have	Whether they have antennae or not

As they do this grouping and sorting, **encourage them to notice** differences in particular, as this is most helpful for making a classification diagram. **Encourage them to sketch** ideas on paper or a whiteboard, and think about questions such as 'if there are 5 arthropods, how many sections/decisions might need to be made to classify them?'

Rycroft-Smith & Macey
Research Group Ltd

Maths Week
Scotland

ROYAL
STATISTICAL
SOCIETY
DATA | EVIDENCE | DECISIONS

4   **Create** your own decision trees and categorisation schemes

Students then **create a categorisation scheme** so that someone else could correctly identify each of the arthropods they have photographed and chosen. At this stage they are aiming to design a draft categorisation scheme for someone else to use.

Give them the **three types of categorisation scheme** used in part 1 of this activity to remind them and to use as inspiration – they can copy one of these ideas, or try and come up with their own type.

Encourage them to **use the correct biological vocabulary** for features of arthropods, and to question ways that their categorisation scheme might fail.

5   **Evaluate and critique** others' work

Finally, groups should try using **another group's categorisation scheme** to identify their arthropods. Remember that each classification scheme should only be used with the arthropods photographed by that group (ie its 'training data').

Students should consider what they found useful, and less useful, about the categorisation scheme they are trialling, and make suggestions of changes to make it easier to use.

Finally, groups could refine and finalise their classification scheme based on using (or rejecting, with reasoning) **this feedback**.

> 💡 This activity could make a **beautiful display!**

Example classroom dialogue

> *This isn't an arthropod – it doesn't have legs, it's a slug!*

> *I followed the classification scheme, but this isn't a woodlouse – what went wrong?*

> *I know this creature by a different name – we call it a clock leddy...*

> *We took two photographs of spiders but we don't know if they are the same species...*

> *I have a photo of an arthropod, but I can't clearly count the legs on it...*

> *This diagram looks like a tree with branches for answers to questions.*

> *I think with five photographs there are four decision points.. I wonder what would happen with ten photos?*

Rycroft-Smith & Macey
Research Group Ltd

Maths Week
Scotland

ROYAL
STATISTICAL
SOCIETY
DATA | EVIDENCE | DECISIONS

References and further reading

Macey, D., & Rycroft-Smith, L., (2022). *What Does Research Suggest About Teaching Statistics Using Rich Data Sets?* Cambridge Mathematics Espresso. https://www.cambridgemaths.org/for-teachers-and-practitioners/espresso/view/rich-data-sets/

Macey, D., & Hornby, W. (2018). *Teaching Statistics.* Cambridge University Press. https://www.cambridgebookshop.co.uk/products/teaching-statistics?srsltid=AfmBOooZpGVsg-riGqi3aTJu2A09X_wGSR9o9KezHWFauEAQdHGcSvYn

Bargagliotti, A., Franklin, C. A., Arnold, P., Gould, R., Johnson, S., Perez, L., & Spangler, D. A. (2020). *Pre-K-12 guidelines for assessment and instruction in statistics education II (GAISE II).* American Statistical Association and National Council of Teachers of Mathematics. https://www.amstat.org/asa/education/Guidelines-for-Assessment-and-Instruction-in-Statistics-Education-Reports.aspx