Rycroft-Smith & Macey
Research Group Ltd

Maths Week Scotland

ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS

Stats in the Wild

Activity ②
Wings and Things

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	2	3	4
Explore a dataset	**Create** graphs of different variables	**Investigate** patterns in the data	**Communicate** your findings

You will need   Video notes: bit.ly/StatsWild2

CODAP software
codap.concord.org

MWS Activity 2 BirdsData.csv

Internet access/search engine

Audio/video recording equipment
(optional)

Art materials
(optional)

Rycroft-Smith & Macey
Research Group Ltd

Maths Week
Scotland

ROYAL
STATISTICAL
SOCIETY
DATA | EVIDENCE | DECISIONS

Why do this activity?

This activity is an introduction to an approach to learning statistics and data handling known as **exploratory data analysis**.

In exploratory data analysis, students are given rich, multivariate, datasets of real or realistic data. They are encouraged to explore the data to look for patterns or features, make hypotheses and pose questions, and manipulate the data using techniques such as sub-setting the data using a categorical variable.

Exploratory data analysis activities are best supported by software tools such as **CODAP** which allow students to quickly try out different charts and generate statistical values so they can focus on interpretation. This activity is an opportunity for students and teachers to become familiar with the CODAP software through a series of activities and acts as a starting point for teachers to explore this approach further by using their own data with CODAP in future lessons.

Possible approach

1   **Explore** a dataset

The first task is designed to encourage students to familiarise themselves with the dataset they will be exploring, and begin to play with the software. CODAP is a free online data exploration tool that does not require you to download or install anything on a computer. The **video notes** provide an overview of how to import data into CODAP as well as some features that will be useful in the later tasks.
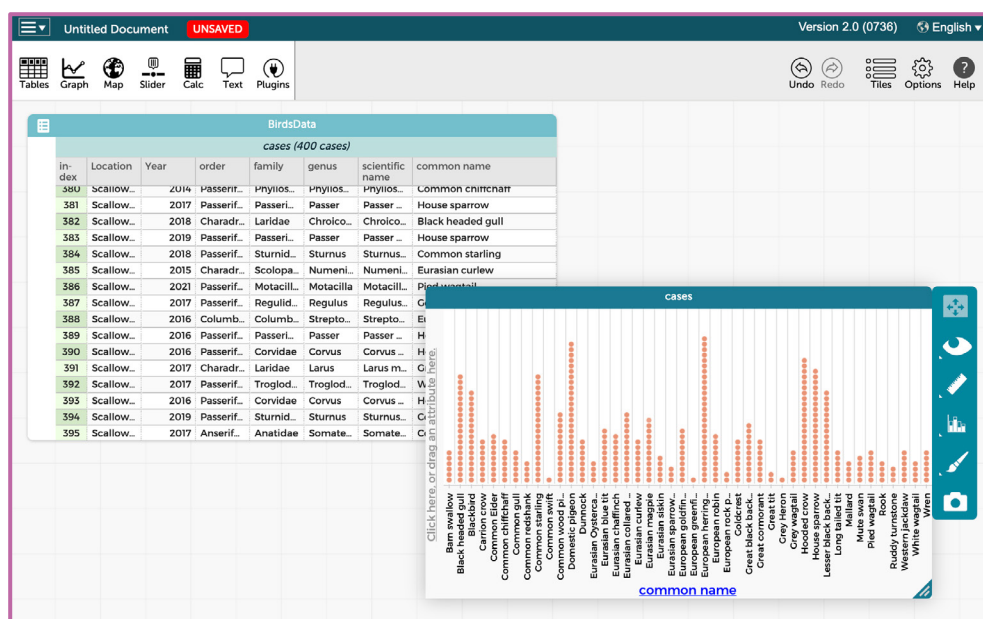
CODAP software workspace

You will need to ensure that the students have **access to the internet** and to the *BirdsData.csv file* that will be used in the activity.

Once students have launched CODAP (https://codap.concord.org) and imported the data, they have **5 questions** to work through and discuss. This could be done in small groups or as a teacher-led discussion.

Column name	Description	Data Type
Index	The numerical reference generated by CODAP	
Location	The place where the data was recorded	Categorical
Year	The year the data was recorded	Quantitative, discrete, time series
Order	A hierarchical term representing a group of families	Categorical
Family	A hierarchical term representing a group of genera	Categorical
Genus	A hierarchical term representing a group of species	Categorical
Scientific name	The scientific name for a species	Categorical
Common name	A common name for the species	Categorical

2 **Create** graphs of different variables

In the second part of the activity students will explore the software and experiment with the **graphing tools**. The focus is not on making specific graphs but to see what effect different actions have. Students should be encouraged to create "nonsense graphs" – graphs that represent data, but don't necessarily have any obvious purpose. This could include identifying actions that appear to cause the graphing tool to give an error and trying to work out why.

One approach would be to allow students time to create some graphs and identify students who have created examples of the following skills: creating a dot plot; creating a bar chart; showing proportions rather than counts; changing the title of a chart; and creating linked charts.

Once all students have had a chance to experiment, you could bring the class back together and ask the students you previously identified to **demonstrate how they made their graph** to the rest of the class.

3 **Investigate** patterns in the data

In this section students will **explore the dataset** and:

- Look for patterns in the data and try to explain them

  and/or

- Make a prediction and see if the data backs it up

An example of these might relate to sea birds. For example, students might:

- Notice that Scalloway has sightings of European rock pipits but Glasgow does not and look them up on the internet to see if they can explain why (e.g. European rock pipits inhabit rocky coastal areas - they may notice that Scalloway is on a small island surrounded by sea)

- Predict that Scalloway will have more sea birds because it is on a small island, and suggest that if this is the case then there will be more sightings of gulls for Scalloway than Glasgow, then check the data to verify their prediction

Give students time to explore the data using CODAP and discuss their ideas.

4 **Communicate** your findings

In the final section, encourage students to **explain their findings** to others. They could be asked to identify a single bird species that stood out to them and discuss what they noticed. For example they could produce a poster about the European rock pipit with a picture and some facts about the species, along with commentary on the data, what it suggests, and what their conclusions or predictions were.

They should be encouraged to include images of any **graphs they created** with that bird highlighted in order to demonstrate the use of data to support claims or arguments.

Example classroom dialogue

If I drag the year onto the y axis, everything gets messy!

When I click on a dot, the table moves… what's it doing that for?

How are Scalloway and Glasgow different?

Glasgow is a big city, so I predict it will have more pigeons…

There aren't any long-tailed tits in Scalloway, I wonder why…

Rycroft-Smith & Macey
Research Group Ltd

Maths Week Scotland

ROYAL
STATISTICAL
SOCIETY
DATA | EVIDENCE | DECISIONS

References and further reading

Rycroft-Smith, L., & Macey, D., (2020). *What Does Research Suggest About Teaching Statistics Using Exploratory Data Analysis (EDA)?* Cambridge Mathematics Espresso. https://www.cambridgemaths.org/for-teachers-and-practitioners/espresso/view/exploratory-data-analysis/

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