

Year 9 Topic 2 – Data science

Lesson	Can you?
Lesson 1: Delving into data science	Define data science Explain how visualising data can help identify patterns and trends in order to help us gain insights Use an appropriate software tool to visualise data sets and look for patterns or trends
Lesson 2: Global data	Recognise examples of where large data sets are used Select criteria to investigate predictions Evaluate findings to support arguments for prediction
Lesson 3: Statistical state of mind	Define the terms 'correlation' and 'outliers' Identify the steps of the investigative cycle Solve a problem by implementing steps of the investigative cycle on a data set Use findings to support a recommendation
Lesson 4: Data for action	Identify the steps of the investigative cycle Identify the data needed to answer a question defined by the learner Create a data capture form
Lesson 5: Clean it up	Describe the need for data cleansing Apply data cleansing techniques to a data set Visualise a data set
Lesson 6: Make a change	Visualise a data set Analyse visualisations to identify patterns, trends, and outliers Draw conclusions and report findings

Useful websites

- www.datawrapper.de
- www.youtube.com/watch?v=f_6IEKqS2l0
- www.gapminder.org
- www.berkeleyearth.lbl.gov/country-list
- www.codap.concord.org
- www.datashine.org.uk
- www.naei.beis.gov.uk/emissionsapp
- www.gaugemap.co.uk



KNOWLEDGE ORGANISER
Key Stage 3 - COMPUTING

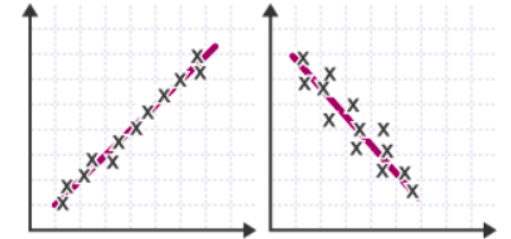
This is a list of data items:

- 42
- rabbits
- 16:00
- 76
- apples
- 09743245530
- £40
- seaside

Spreadsheets are commonly used to turn data into information.



A **line of best fit** is a sensible straight line that goes as centrally as possible through the coordinates plotted. It should also follow the same steepness of the crosses.



Positive correlation means as one variable increases, so does the other variable. They have a positive connection.

The investigative cycle

So far we have spent time investigating data sets to see patterns or to extract meaning.

The **PPDAC** cycle is a framework for us to follow when asking and answering real-world problems using data.

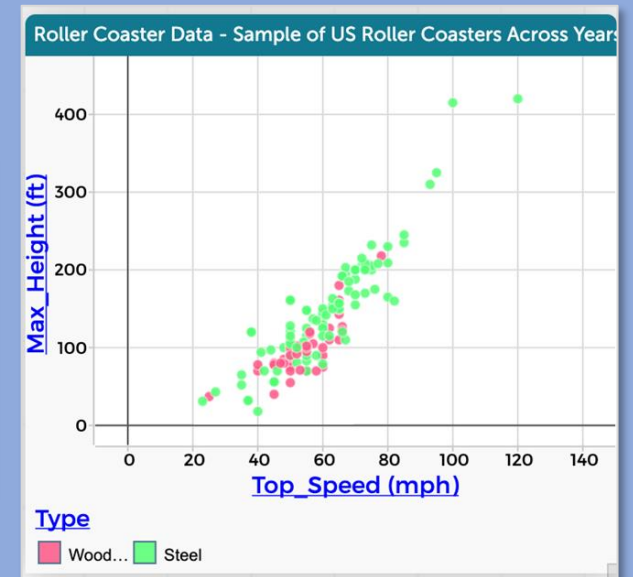
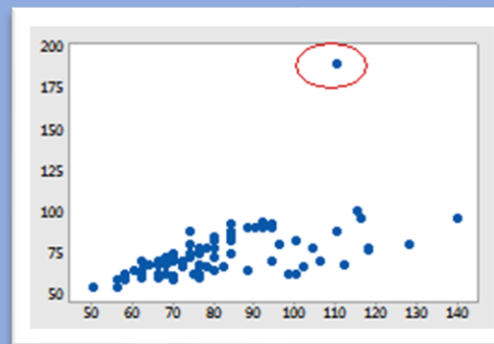


Outliers on scatter graphs

Scatter plots often have a pattern. We call a data point an **outlier** if it doesn't fit the pattern.

The scatter graph below shows data for students on a hiking trip.

Each student is carrying a backpack and each point on the graph represents a student.



Negative correlation means as one variable increases, the other variable decreases. They have a negative connection.