Data types and storage (Answers)

Worksheet section	Contents
1	Data types
2	Data structures
3	Extension

This lesson has been created by Effini in partnership with Data Education in Schools, The Data Lab and Data Skills for Work, with funding from the Scottish Government.

1. Data types

Reminder
Integer numbers with no decimal or fractional parts
Floating point numbers that can contain a decimal or fractional part
Boolean can take two possible values such as true/false or yes/no
Date/time the number of days or seconds passed since the 'epoch' date
Character a single text character which can be a letter, number or symbol

Section 1.1

1) State the data type for each of these examples.

Example	Data type	Reason
Number of people at a football match	Integer	Whole numbers with no decimal or fractional parts
Length of the Queensferry crossing (2.7km)	Floating point	Number that can contain a decimal or fractional part
£	Character	A single text character which can be a letter, number or symbol
Are the tickets sold out for a concert? Yes/No	Boolean	Can take two possible values
Your birthday	Date and time	Its a date
@	Character	A single text character which can be a letter, number or symbol
Distance in miles (168.9miles) between Inverness and Glasgow	Floating point	Number that can contain a decimal or fractional part

Section 1.2

2) State the data type and explain the reason behind your choice.

Example	Data type	Reason
Number of 'likes' on a Facebook post	Integer	Whole numbers with no decimal or fractional parts
Weight of bag of flour	Floating point	Number that can contain a decimal or fractional part
$	Character	A single text character which can be a letter, number or symbol
Whether someone voted in an election? Yes/No	Boolean	Can take two possible values
Date of a lesson (DD/MM/YYYY)	Date and time	Its a date
&	Character	A single text character which can be a letter, number or symbol
Average depth (95.0m) of the North Sea	Floating point	Number that can contain a decimal or fractional part

1. Data types

Reminder
Integer numbers with no decimal or fractional parts
Floating point numbers that can contain a decimal or fractional part
Boolean can take two possible values such as true/false or yes/no
Date/time the number of days or seconds passed since the 'epoch' date
Character a single text character which can be a letter, number or symbol

Section 1.3

3) What is meant by the 'epoch' date?

The epoch is the date and time relative to which other date/times are calculated. It is the point that the software considers to be 0 hours, 0 minutes, 0 seconds.

In Unix (as used in Python) it is 1/1/1970, in Microsoft Excel it is 1/1/1900.

4) Fill in the missing values in this table

Reminder: the stored value is the number of days between the 'epoch' date and the displayed date

Software	Epoch date	Stored value	Displayed date
Excel	01/01/1900	44000	19 Jun 2020
Python	01/01/1970	15	16 Jan 1970
Excel	01/01/1900	25569	01 Jan 1970
Excel	01/01/1900	10	11 Jan 1900
Python	01/01/1970	10	11 Jan 1970
Excel	01/01/1900	45	15 Feb 1900
Python	01/01/1970	45	15 Feb 1970

2. Data structures

Reminder
String collection of characters combined to create alphanumeric text
Array structure of a fixed size which can hold items of the same data type
Dataset a two-dimensional structure that has rows and columns

Section 2.1

1) State the most appropriate data structures these examples could be stored in.

Example	Data type	Reason
Your name	Strings	Collection of characters combined to create alphanumeric text
Table, 2 rows by 3 columns containing only integers	Array	Fixed sized structure which can hold items of the same data type
To do list	List	A dynamically sized structure which can contain different data types
Name and address for your customers	Dataset	Each column can hold different data types, but all must contain the same number of items.
Time taken to run a race with the name of the runner and age	Dataset	Each column can hold different data types, but all must contain the same number of items.
Flight number (e.g. BA1254)	Strings	Collection of characters combined to create alphanumeric text

Section 2.2

2) Turn this article into a dataset with the following column names, 'name', 'distance_km', 'finish_position', 'time_to_finish', 'personal_best'

"On Sunday 12th July, 5 competitors ran in the local fun run. The winner 'J. King' finished the 5K course in 15mins 10secs. 4 of the competitors finished the course however 'L. Smith' had to drop out after 3.5 kilometres. The other 3 runners achieved personal bests, R. Gate 16mins 45secs , W. Henry 17mins 1sec, B. Perkins 18mins 49secs."

name	distance_km	finish_position	time_to_finish	personal_best
J. King	5	1	15m 10s	N
L. Smith	3.5	NA	NA	N
R. Gate	5	2	16m 45s	Y
W. Henry	5	3	17m 1s	Y
B. Perkins	5	4	18m 49s	Y

2. Data structures

Reminder
String collection of characters combined to create alphanumeric text
Array structure of a fixed size which can hold items of the same data type
Dataset a two-dimensional structure that has rows and columns

3) Turn this article into a dataset with the following column names, 'tree_name', 'tree_type' and 'height_m'.

"Within the grounds of the house there are a mixture of trees. They are 2 broadleaf trees (oak and silver birch) and a coniferous tree (scots pine). The scots pine is 36.5m tall, the oak is 30.0m and the silver birch is 24.0m."

tree_name	tree_type	height_m
Scots pine	Coniferous	36.5
Oak	Broadleaf	30.0
Silver birch	Broadleaf	24.0

What type of data are in these columns?

tree_name	string
tree_type	string
height_m	floating point

Section 2.3

4) What data structure is this data in? Explain why.

3	3	54
4	15	
5		14
4	4	
4	100	45
5	5	4

It is an array because all the data is the same type (integers) and there are empty cells which are not allowed in datasets.

2. Data structures

Reminder
String collection of characters combined to create alphanumeric text
Array structure of a fixed size which can hold items of the same data type
Dataset a two-dimensional structure that has rows and columns

5) What are the benefits in terms of flexibility, if you choose a dataset rather than an array when storing data?

A dataset allows you to have flexibility in the number of rows and columns. Also you can hold multiple data types (e.g. integer, date/time).

3. Extension

This extension section will cover,
Data types
Display formats
Data Structure
File formats

Section 3.1

You are planning a group outing to the beach and you are asking people for **ideas on where and when** to go.

1) Give an example of how you could collect this data in unstructured and structured way.

Structured	Unstructured
Online survey tool	Text message conversation
Online feedback form	Talking to people
Paper survey	Group chat e.g. Email, instant messaging app

Note for the teacher:

Given the open ended style of this section, this page gives **some examples** of acceptable answers. Other answers may be acceptable too.

2) Can you think of some pros/cons of collecting data in a structured way?

Pros	Cons
Easier to analyse	Restricts the possible answers e.g. survey questions vs. talking to someone
You can make sure you get the answers to specific questions e.g. how many people are coming?	Limits the use of the data to the single purpose.

3) Imagine you are setting up a structured way of holding the ideas. What type of data structure would you use and why?

Dataset, because each column could hold different data types and would be easy to analyse.

3. Extension

This extension section will cover,
Data types
Display formats
Data Structure
File formats

4) You decide to conduct a survey to capture everyone's ideas for the beach trip. Design a question that you could ask for each of these data types, then decide on the best display format.

Data type	Question	Display format
Floating point	How much would you be willing to pay for the trip?	Accounting (£)
Boolean	Could you drive people to the beach?	Yes/No
Date/time	When would you like to go?	DD MM YYYY

5) What file format would you hold this data in?

Database, e.g. xml, csv, tab

6) Can you list 5 data items you might want to collect as part of organising this trip? Then state the data type and the display format you would like to see the data in.

Data item	Data type	Display format
Number of people coming	Integer	Integer
Distance to the bench	Floating point	0 d.p (nearest mile)
Phone numbers of people coming	Character	String (0xxxx xxxxxx)
Planned time to leave	Date/time	hh:mm
Transport type (e.g. bus, car, train)	Character	String
Has the person replied?	Boolean	Yes/No