Introduction to Computer Programming Lecture 10.1:

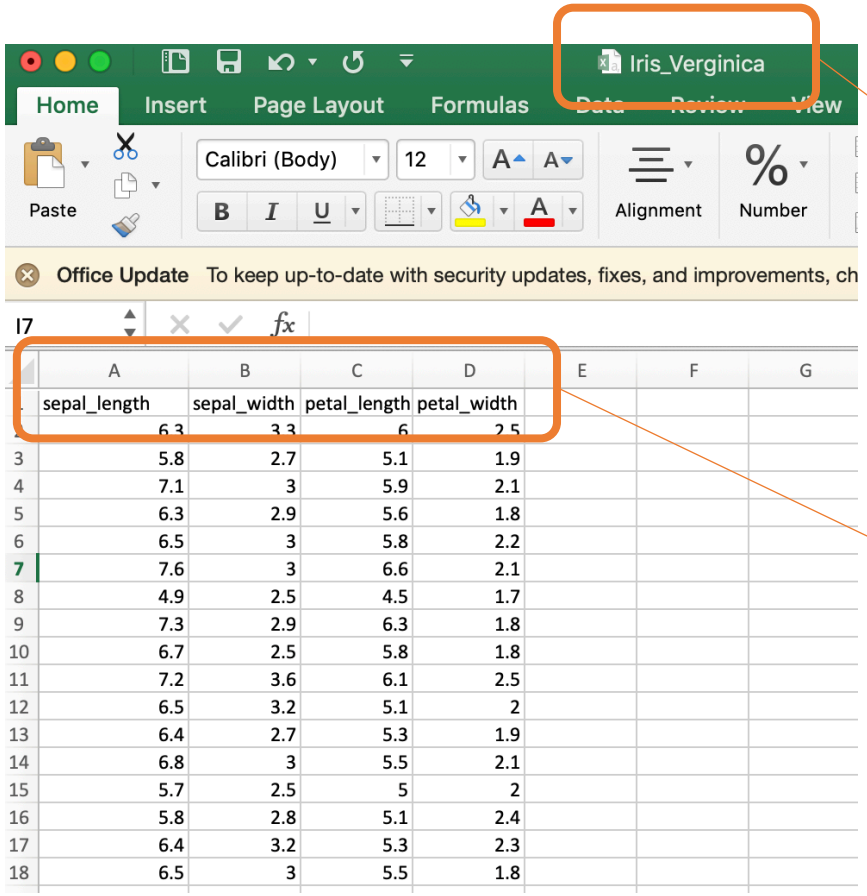
Writing Programs: Working with data

Hemma Philamore

Department of Engineering Mathematics

Example: Iris data set

You are working in a lab where data is collected on iris flower samples.



	A	B	C	D
	sepal_length	sepal_width	petal_length	petal_width
2	6.3	3.3	6	2.5
3	5.8	2.7	5.1	1.9
4	7.1	3	5.9	2.1
5	6.3	2.9	5.6	1.8
6	6.5	3	5.8	2.2
7	7.6	3	6.6	2.1
8	4.9	2.5	4.5	1.7
9	7.3	2.9	6.3	1.8
10	6.7	2.5	5.8	1.8
11	7.2	3.6	6.1	2.5
12	6.5	3.2	5.1	2
13	6.4	2.7	5.3	1.9
14	6.8	3	5.5	2.1
15	5.7	2.5	5	2
16	5.8	2.8	5.1	2.4
17	6.4	3.2	5.3	2.3
18	6.5	3	5.5	1.8

.csv file for each iris species

Four measurements for each sample:

- Sepal length
- Sepal width
- Petal length
- Petal width

Example: Iris data set

You must write a program that can extract data about a species from the .csv file.

The data about different species should be extracted within the same program so that values for one species can be plotted against another.

Start

↓

Open file

↓

Contents to data structure

↓

Separate headings and numerical data

↓

Data to numpy array

↓

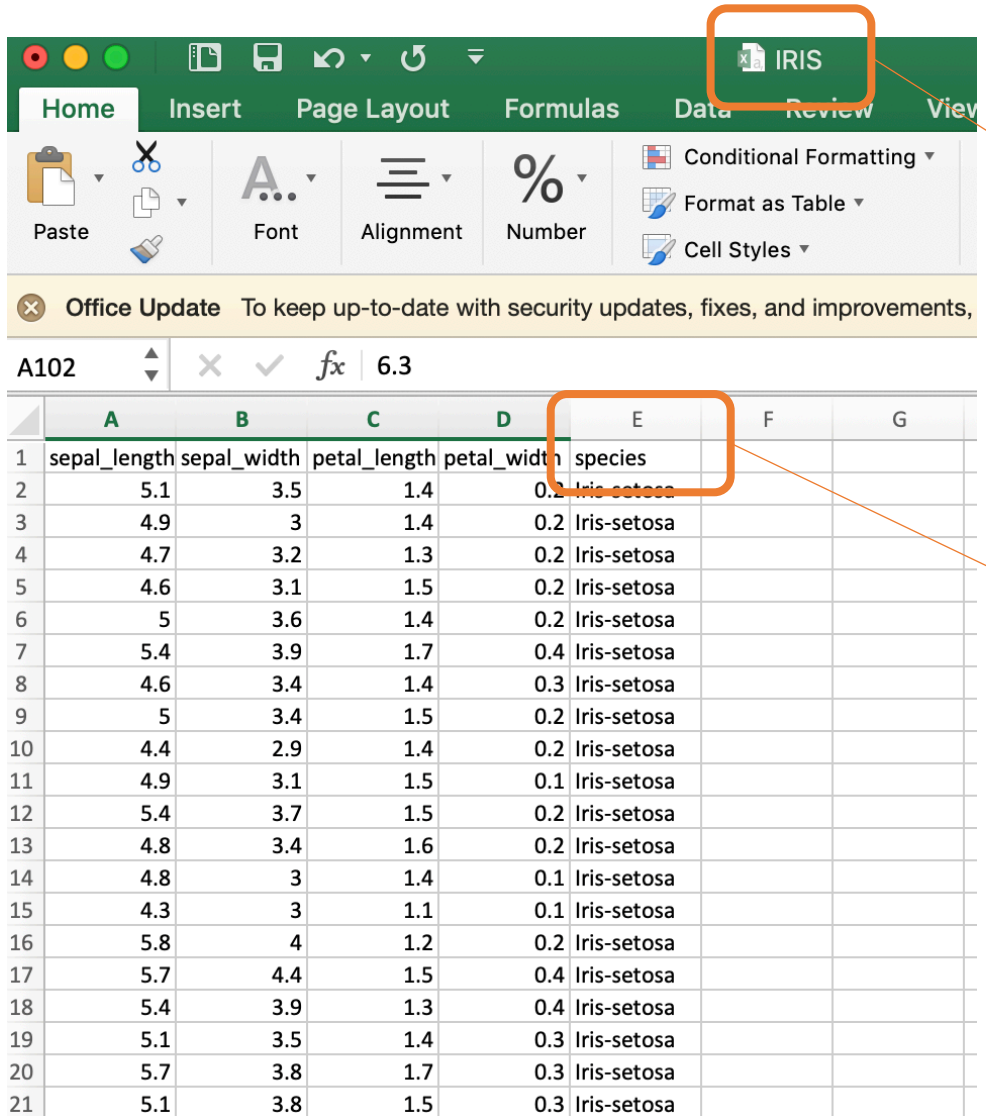
<u>Extract data:</u>

- Sepal width
- Petal length
- Average sepal width
- Maximum sepal length

Example: Iris data set

The lab changes its file system for storing data.

Update your program so that it accepts the new data format.



	A	B	C	D	E	F	G
1	sepal_length	sepal_width	petal_length	petal_width	species		
2	5.1	3.5	1.4	0.2	Iris-setosa		
3	4.9	3	1.4	0.2	Iris-setosa		
4	4.7	3.2	1.3	0.2	Iris-setosa		
5	4.6	3.1	1.5	0.2	Iris-setosa		
6	5	3.6	1.4	0.2	Iris-setosa		
7	5.4	3.9	1.7	0.4	Iris-setosa		
8	4.6	3.4	1.4	0.3	Iris-setosa		
9	5	3.4	1.5	0.2	Iris-setosa		
10	4.4	2.9	1.4	0.2	Iris-setosa		
11	4.9	3.1	1.5	0.1	Iris-setosa		
12	5.4	3.7	1.5	0.2	Iris-setosa		
13	4.8	3.4	1.6	0.2	Iris-setosa		
14	4.8	3	1.4	0.1	Iris-setosa		
15	4.3	3	1.1	0.1	Iris-setosa		
16	5.8	4	1.2	0.2	Iris-setosa		
17	5.7	4.4	1.5	0.4	Iris-setosa		
18	5.4	3.9	1.3	0.4	Iris-setosa		
19	5.1	3.5	1.4	0.3	Iris-setosa		
20	5.7	3.8	1.7	0.3	Iris-setosa		
21	5.1	3.8	1.5	0.3	Iris-setosa		

The data for all species are now stored in a single file

There is an additional column which shows the species of each iris sample