Extracting and Plotting Features

A train/test scenario: Butterflies

Training images

Polygonia c-album
(Comma)

Maniola jurtina
(Meadow brown)

Pyronia tithonus
(Gatekeeper)

Test image

What label to give
this butterfly?

How to describe a butterfly?

How to describe a butterfly?

Colour?

Number of…
- Spots?
- Antennae?
- Legs?

Patterns?

Location captured?

Shape?

Size?

How to describe a butterfly?

Colour?

Number of…
- Spots?
- Antennae?
- Legs?

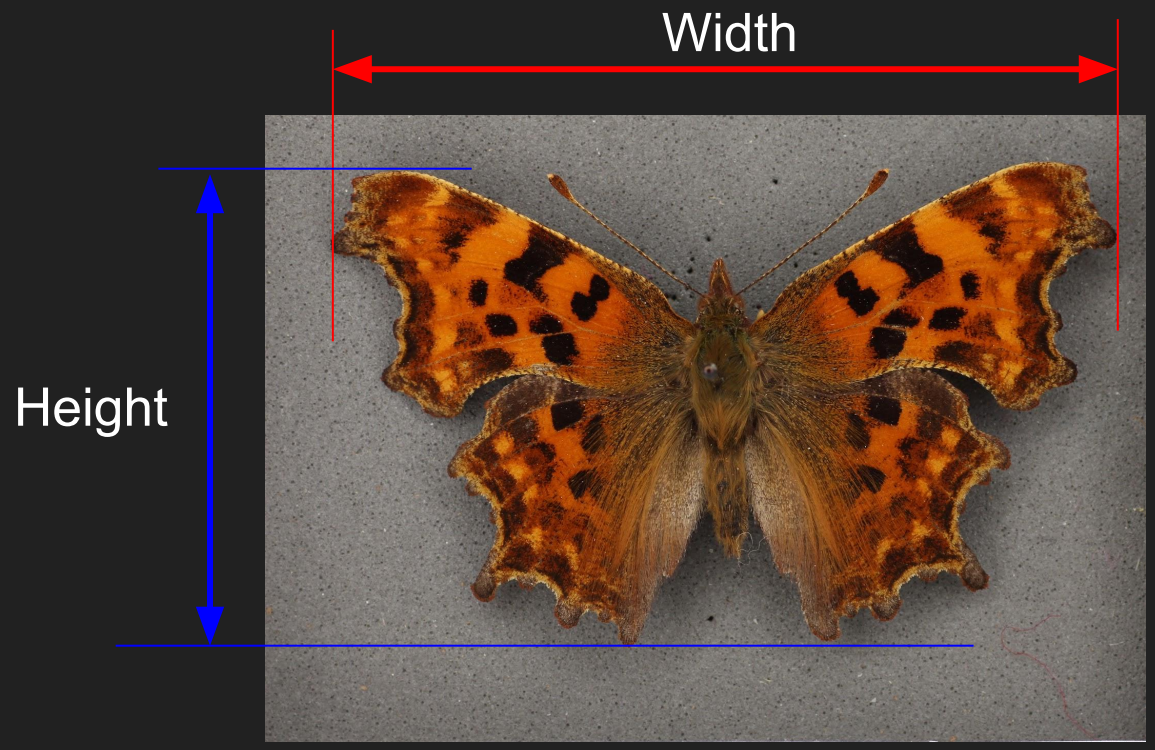
Patterns?

Location captured?

Shape?

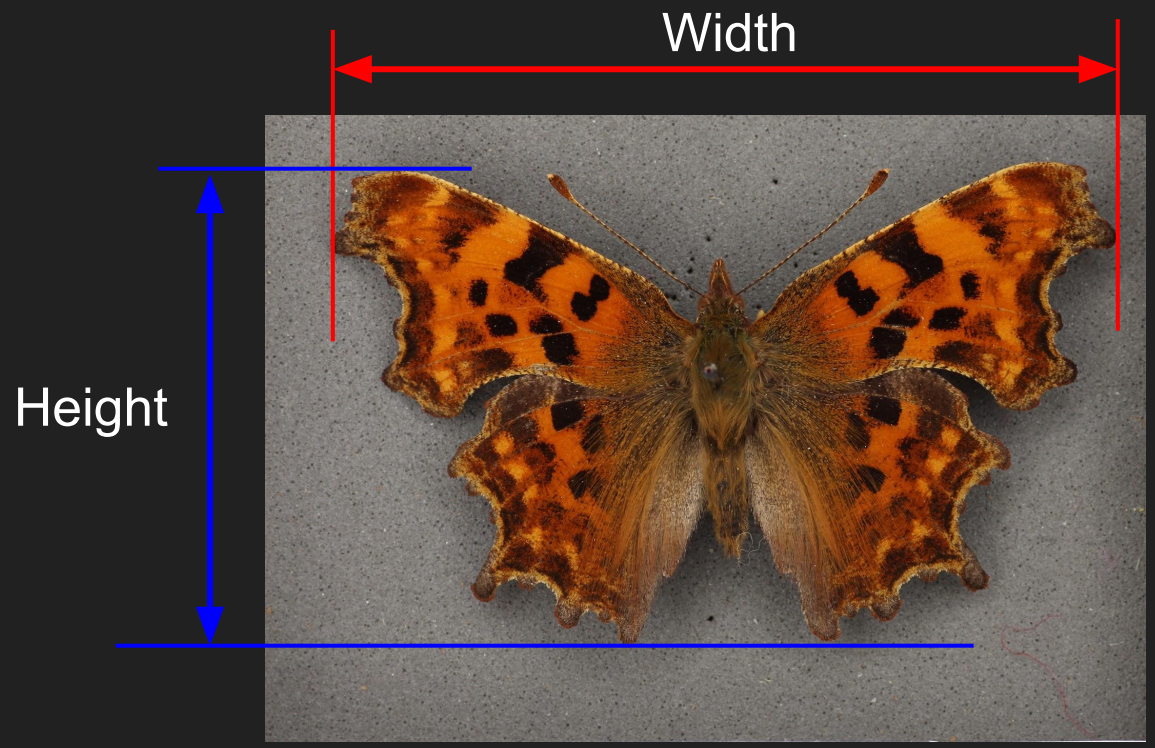
Size?

How to describe a butterfly?



Size?

How to describe a butterfly?



$\mathbf{x}$ = (width, height)

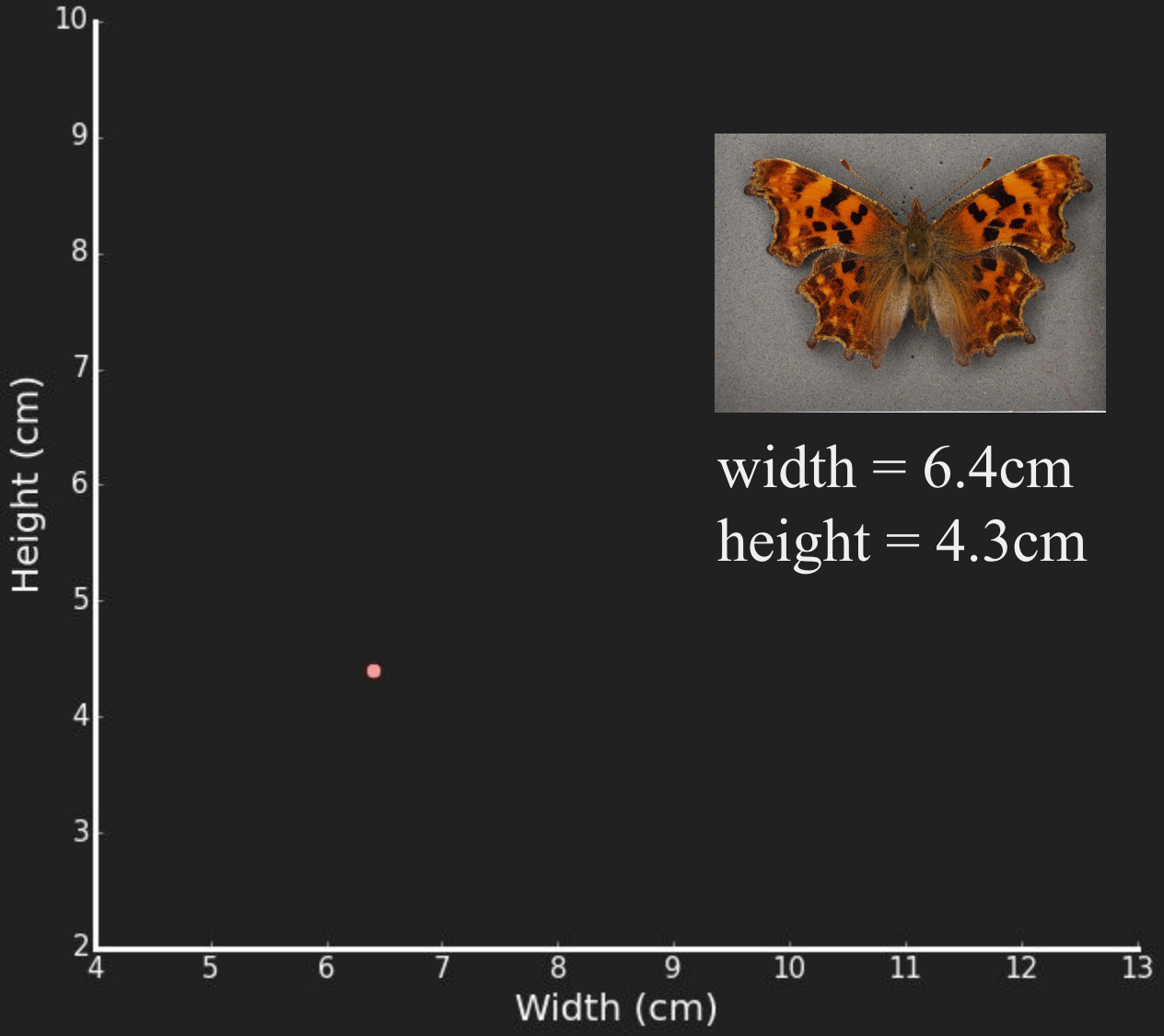
$\mathbf{y}$ = species name

Feature vector

Feature (descriptor)

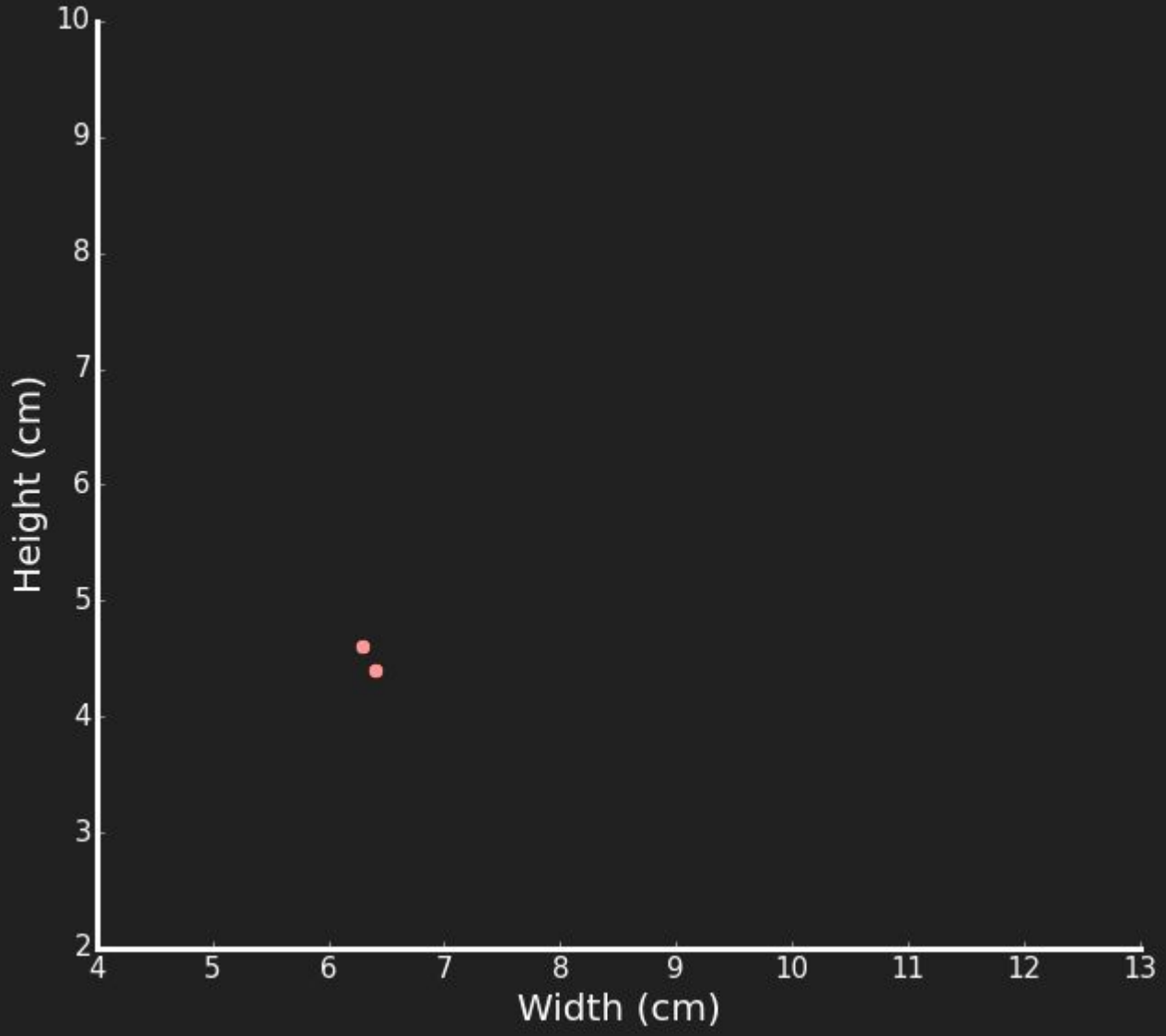
Target variable (label, class)

Visualising features

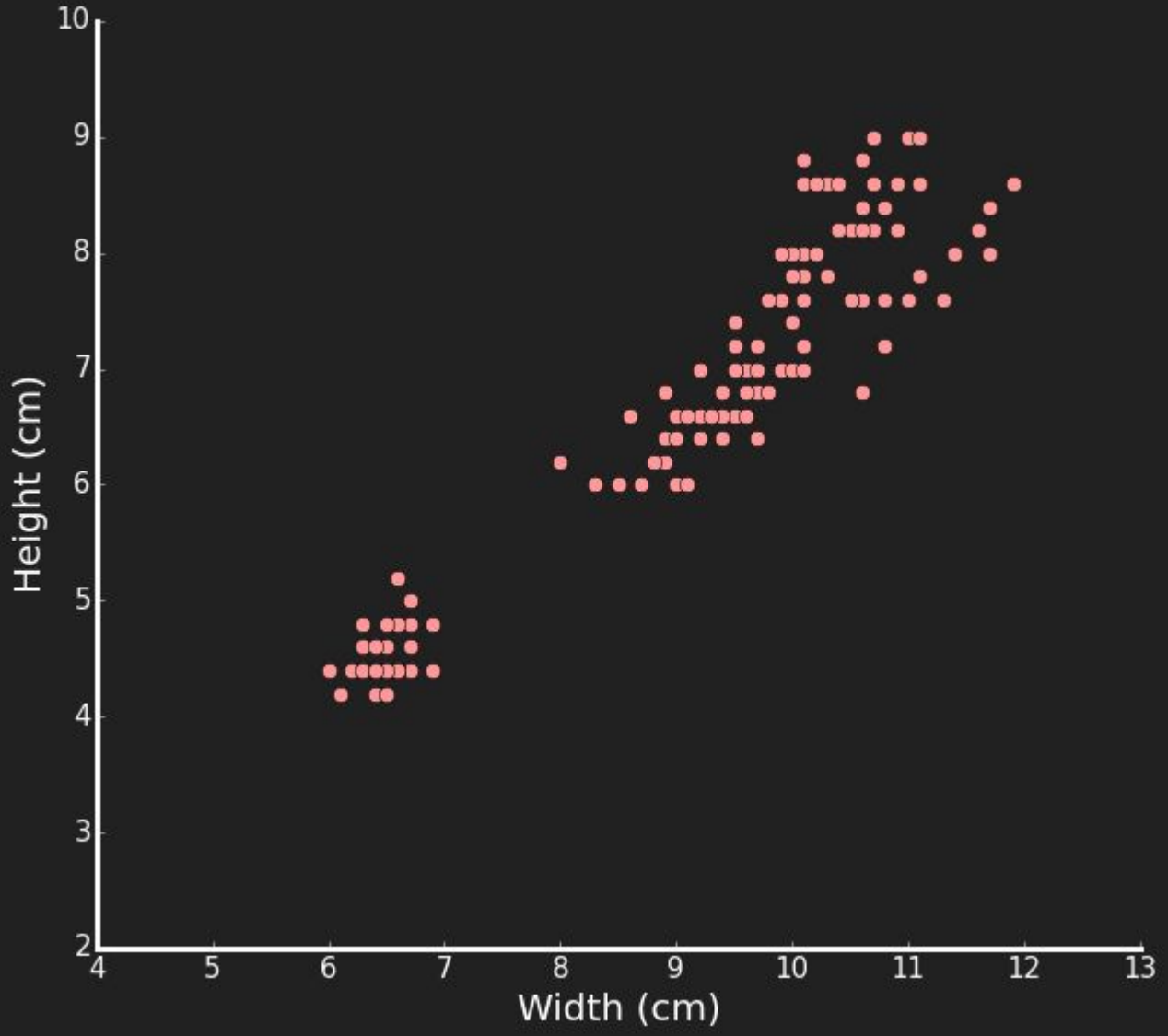


width = 6.4cm
height = 4.3cm

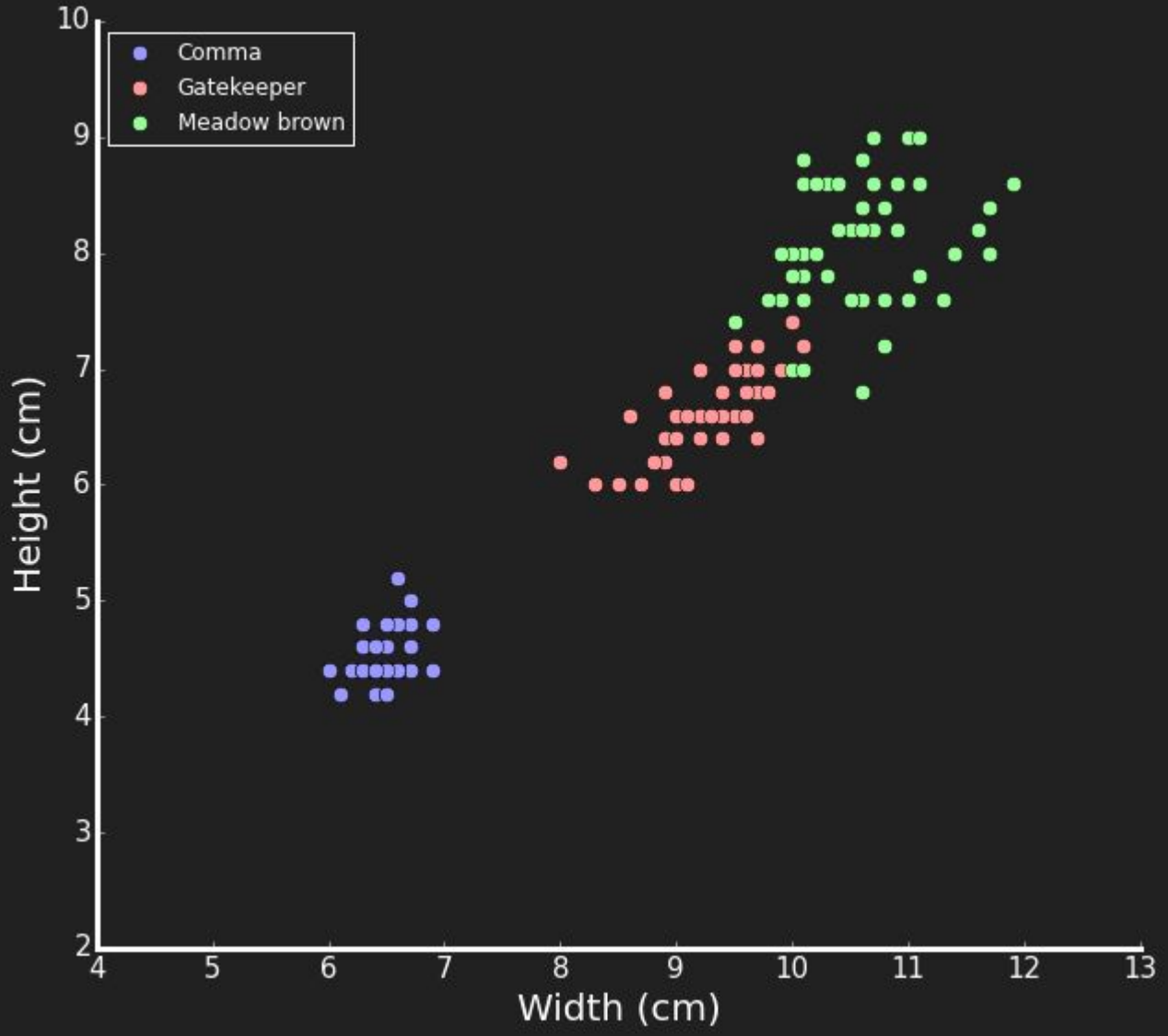
Visualising features



Visualising features



Visualising features



Longer feature vectors

$$\mathbf{x} = \underbrace{(\text{width}, \text{height})}_{\text{2 dimensions}}$$

$$\mathbf{x} = \underbrace{(\text{width}, \text{height}, \text{wing\_area}, \text{latitude}, \text{longitude})}_{\text{5 dimensions}}$$

- We can’t plot this very easily…
- But everything else we show still applies

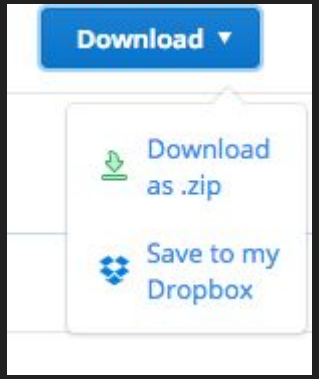
Getting started with the R practicals

1) Go to:

http://www0.cs.ucl.ac.uk/staff/O.MacAodha/ml_intro

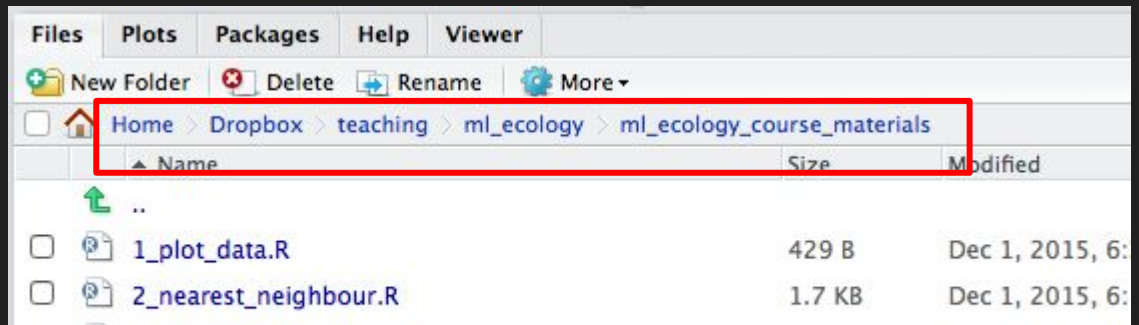
2) Follow the dropbox link

3) Download as zip (or save to your dropbox):

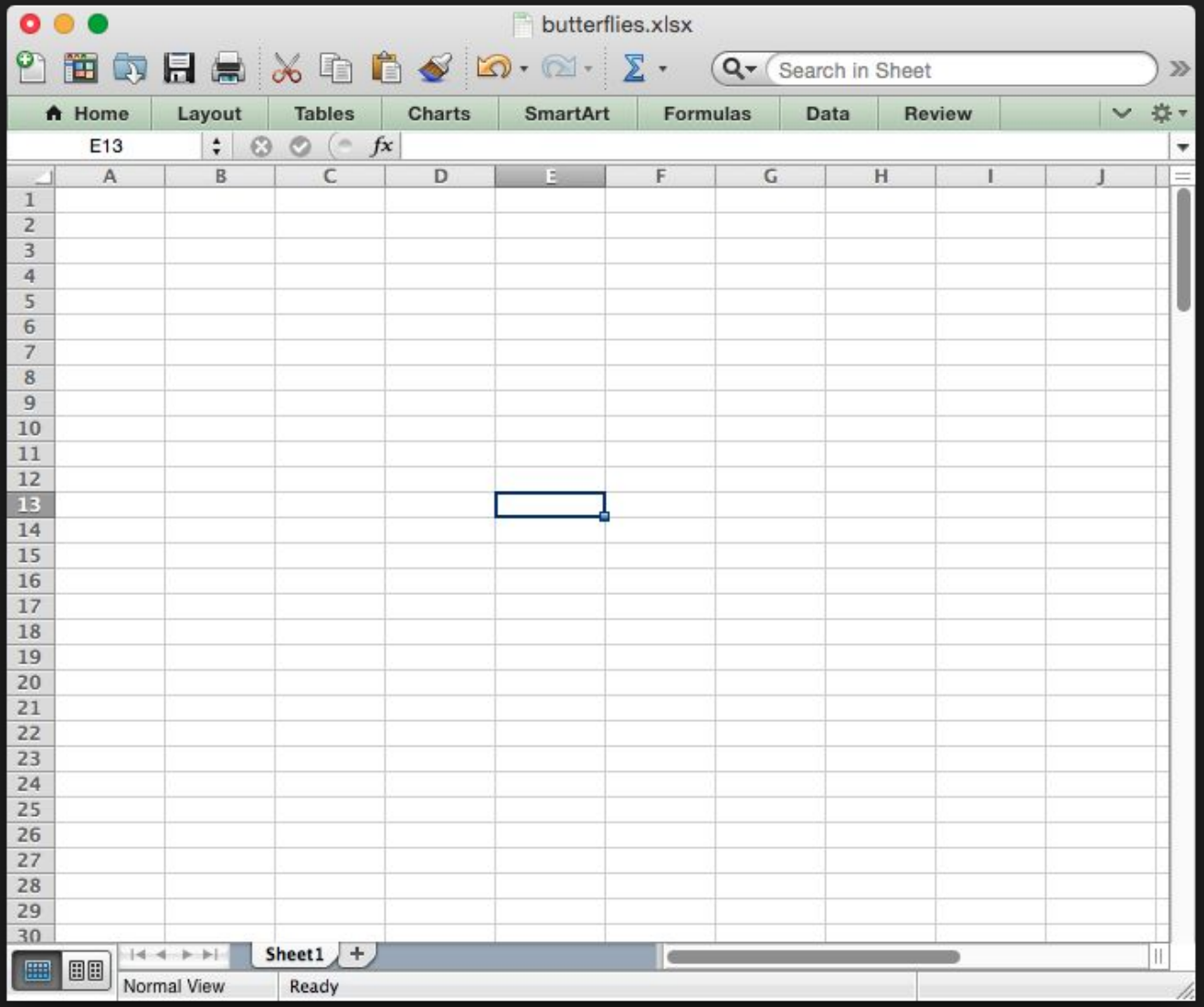


4) Open RStudio

5) Navigate to the folder where the files are:

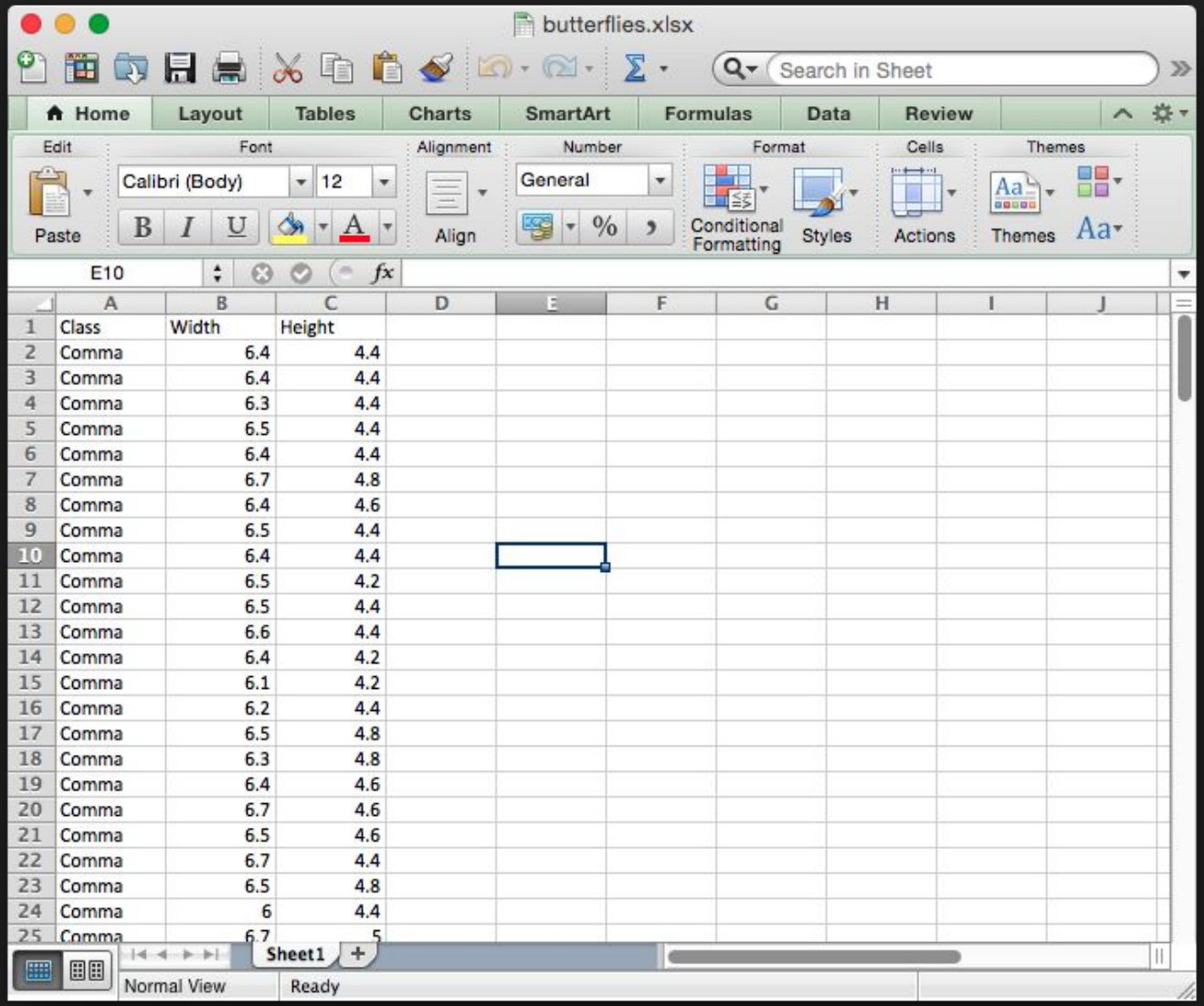


Loading and saving feature vectors



flower_data <- read.csv('data/flower_dataset.csv')

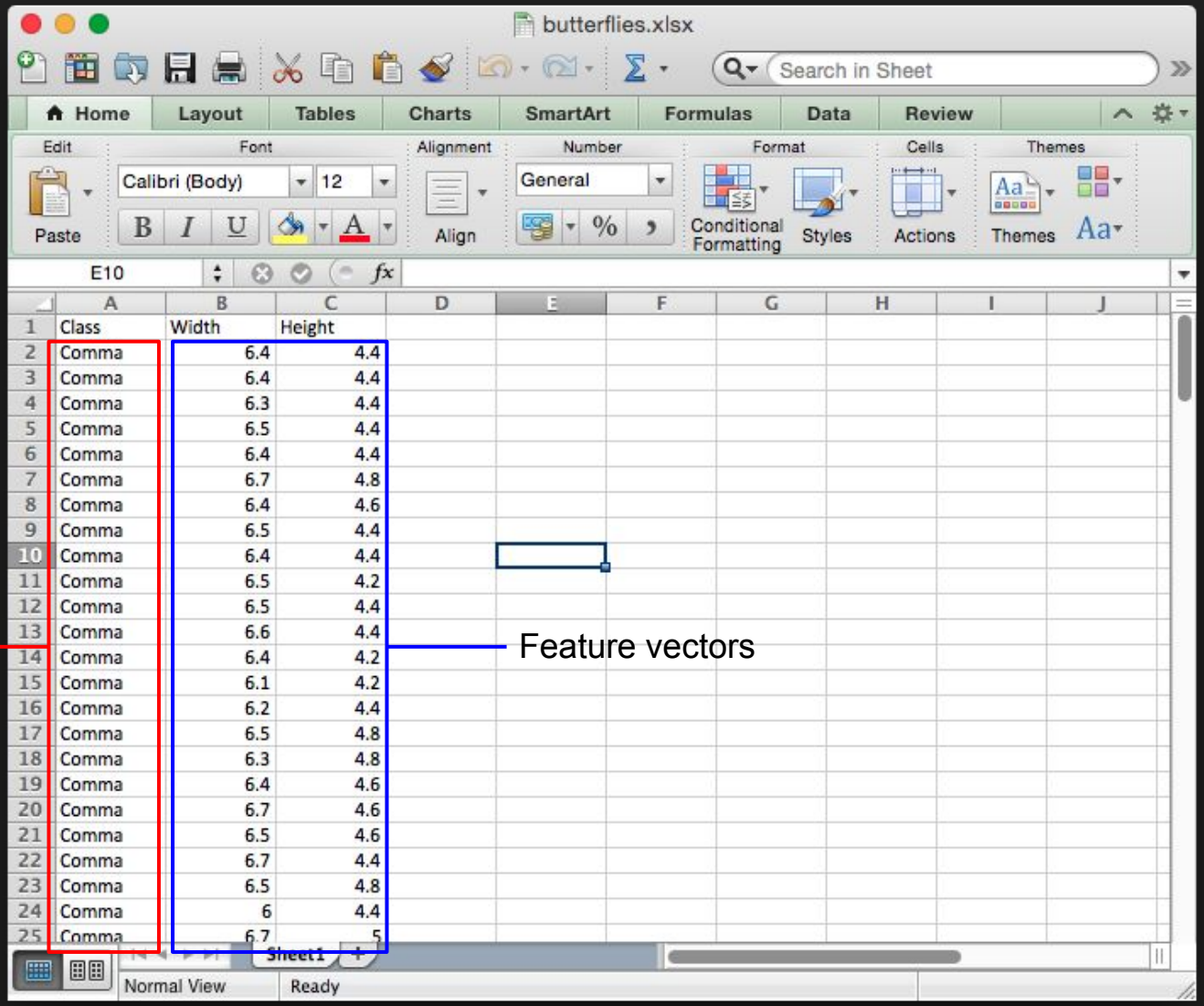
Loading and saving feature vectors



	A	B	C
1	Class	Width	Height
2	Comma	6.4	4.4
3	Comma	6.4	4.4
4	Comma	6.3	4.4
5	Comma	6.5	4.4
6	Comma	6.4	4.4
7	Comma	6.7	4.8
8	Comma	6.4	4.6
9	Comma	6.5	4.4
10	Comma	6.4	4.4
11	Comma	6.5	4.2
12	Comma	6.5	4.4
13	Comma	6.6	4.4
14	Comma	6.4	4.2
15	Comma	6.1	4.2
16	Comma	6.2	4.4
17	Comma	6.5	4.8
18	Comma	6.3	4.8
19	Comma	6.4	4.6
20	Comma	6.7	4.6
21	Comma	6.5	4.6
22	Comma	6.7	4.4
23	Comma	6.5	4.8
24	Comma	6	4.4
25	Comma	6.7	5

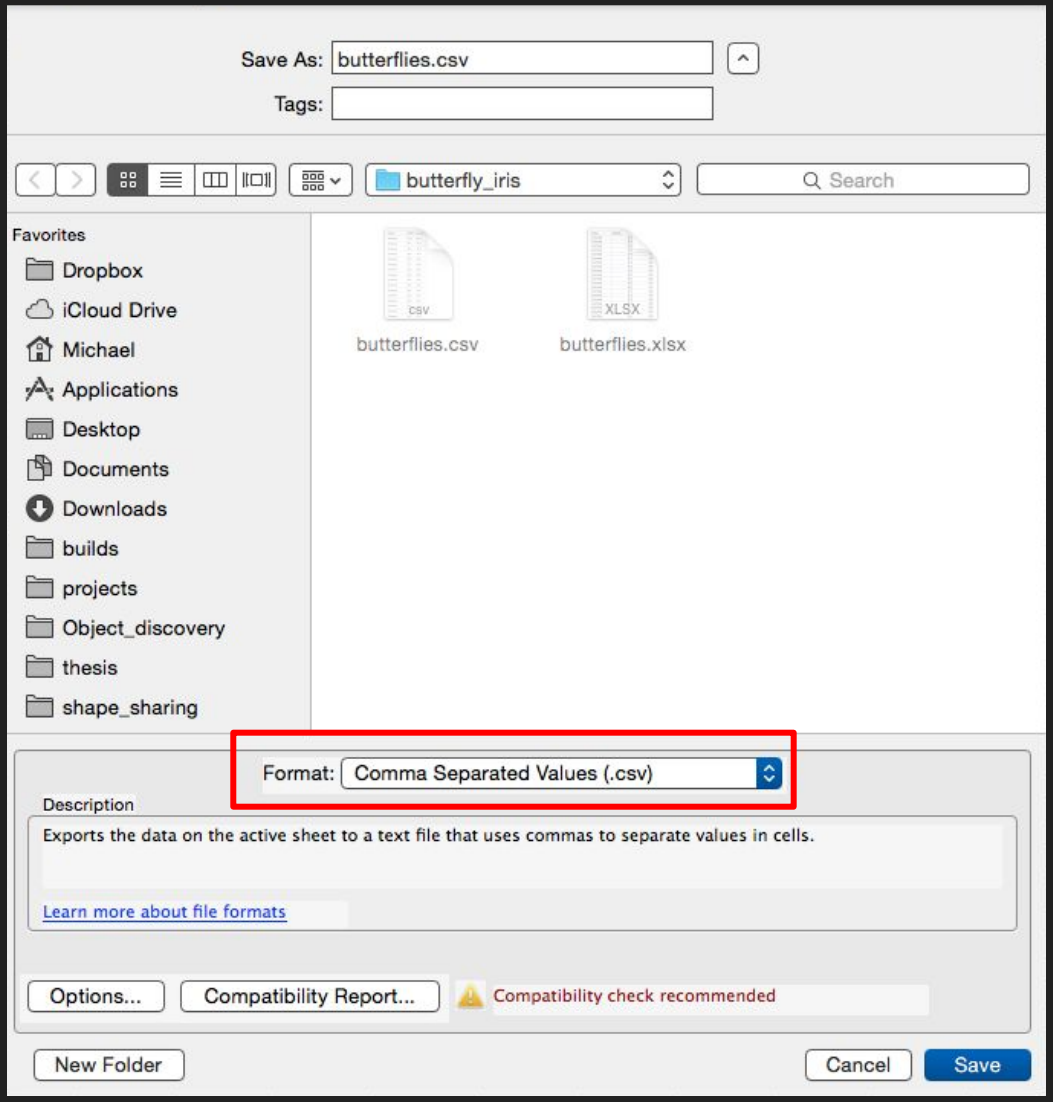
flower_data <- read.csv('data/flower_dataset.csv')

Loading and saving feature vectors

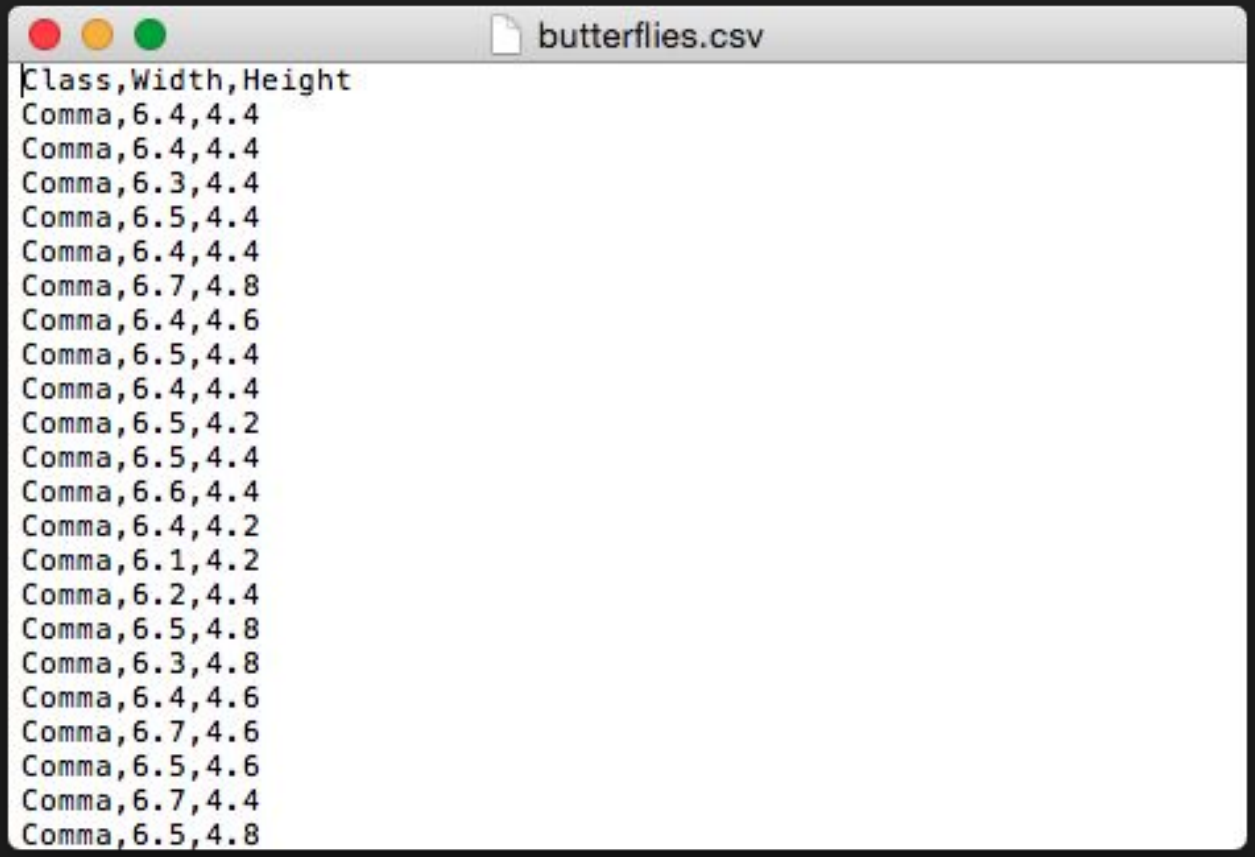


	A	B	C
1	Class	Width	Height
2	Comma	6.4	4.4
3	Comma	6.4	4.4
4	Comma	6.3	4.4
5	Comma	6.5	4.4
6	Comma	6.4	4.4
7	Comma	6.7	4.8
8	Comma	6.4	4.6
9	Comma	6.5	4.4
10	Comma	6.4	4.4
11	Comma	6.5	4.2
12	Comma	6.5	4.4
13	Comma	6.6	4.4
14	Comma	6.4	4.2
15	Comma	6.1	4.2
16	Comma	6.2	4.4
17	Comma	6.5	4.8
18	Comma	6.3	4.8
19	Comma	6.4	4.6
20	Comma	6.7	4.6
21	Comma	6.5	4.6
22	Comma	6.7	4.4
23	Comma	6.5	4.8
24	Comma	6	4.4
25	Comma	6.7	5

Loading and saving feature vectors



Loading and saving feature vectors

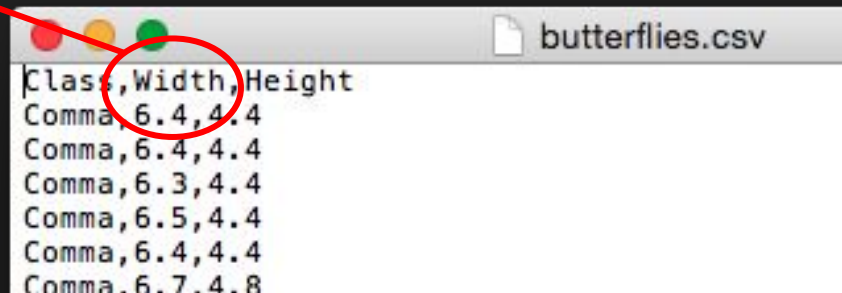


OSX: TextEdit
Windows: Notepad

csv = comma separated values

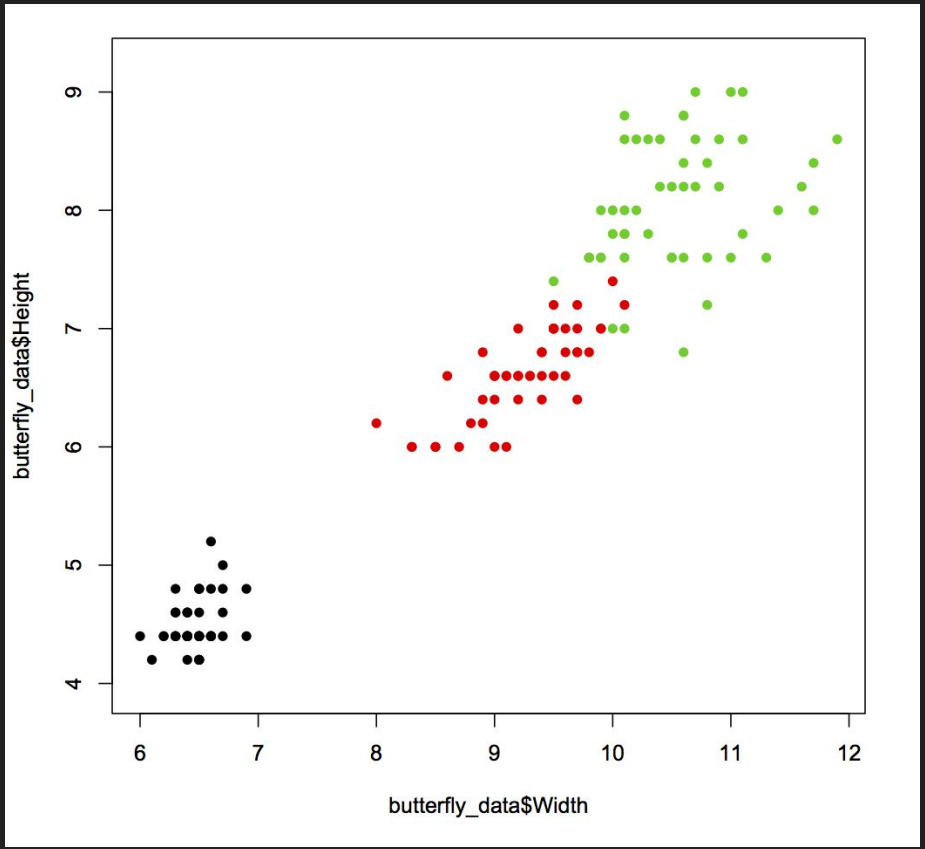
Loading CSV files into R

```
> butterfly_data <- read.csv('butterflies.csv')

> head(butterfly_data)

  Class Width Height
1 Comma   6.4    4.4
2 Comma   6.4    4.4
3 Comma   6.3    4.4
4 Comma   6.5    4.4
5 Comma   6.4    4.4
6 Comma   6.7    4.8

> print(butterfly_data$Width)
```



Plotting data in R

```
> plot(butterfly_data$Width, butterfly_data$Height,
col=butterfly_data$Class, pch=16, asp=1)
```



Practical example

1_plot_data.R