

Lecture 18 - R Software

Ashutosh Rajput
Department of Mathematics

Rajdhani College
University of Delhi

Introduction to Graphical Analysis

Introduction: Graphs are a powerful way to present data and results in a concise manner. Irrespective of the data given, there is a way to illustrate it graphically. A graph is more readily understandable than words and numbers, and producing good graphs is a vital skill.

★ BOX-WHISKER PLOTS

The box-whisker plot (often abbreviated to boxplot) is a useful way to visualize complex data where you have multiple samples.

Features:-

- Box-whisker plot shows the median value, the quartiles (or hinges), and the max/min values.
- A lot of information can be represented in a compact manner.
- We can use the `boxplot()` command to create box-whisker plots.

Types of Boxplots

1. Basic Boxplots
2. Customizing Boxplots
3. Horizontal Boxplots

Basic Boxplots

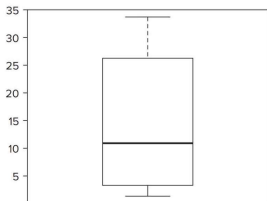
The following example shows a simple data frame composed of two columns:

```
> fw
```

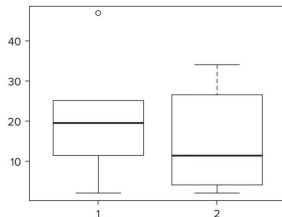
	count	speed
Taw	9	2
Torr ridge	25	3
Ouse	15	5
Exe	2	9
Lyn	14	14
Brook	25	24
Ditch	24	29
Fal	47	34

We can use the `boxplot()` command to visualize one of the variables here:

```
> boxplot(fw$speed)
```



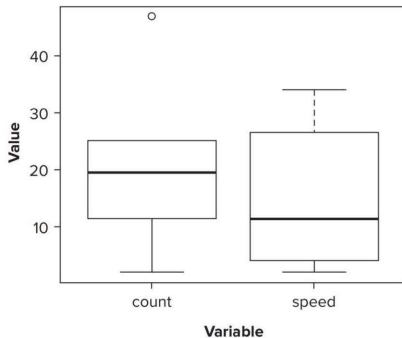
```
> boxplot(fw$count, fw$speed)
```



Customizing Boxplots

We can label the plots as `xlab` and `ylab` instructions to label the axes. We can use the `names` instruction to set the labels.

```
> boxplot(fw$count, fw$speed, names = c('count', 'speed'))  
> title(xlab = 'Variable', ylab = 'Value')
```

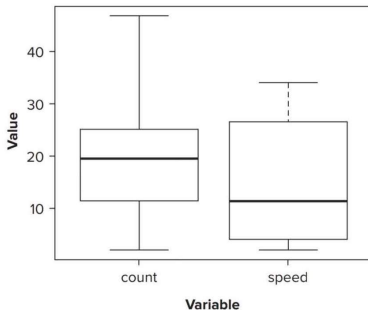


Note: The whiskers of the count sample do not extend to the top, and that we appear to have a separate point displayed.

However, we can determine how far out the whiskers extend.

Remark: 1. By default this is 1.5 times the interquartile range. 2. We can alter this by using the `range = instruction`, if you specify `range = 0`, then the whiskers extend to the maximum and minimum values.

```
> boxplot(fw$count, fw$speed, names
= c('count', 'speed'), range = 0, xlab
= 'Variable', ylab = 'Value')
```



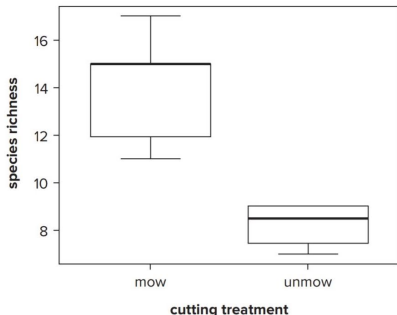
Suppose, the data be given as:

```
> grass
   rich  graze
1   12   mow
2   15   mow
3   17   mow
4   11   mow
5   15   mow
6    8 unmov
```

```
7 9 unmow
8 7 unmow
9 9 unmow
```

The following command gives:

```
> boxplot(rich ~ graze, data = grass, range = 0)
> title(xlab = 'cutting treatment', ylab = 'species richness')
```



Note: We use the \sim symbol to separate the response(dependent) variable to the left and the predictor(independent) variable to the right.

- We can choose, to add the axis labels separately with the `title()` command.
- The samples are automatically labeled, the command takes the names of the samples from the levels of the factor, presented in alphabetical order.

Horizontal Boxplots

- ```
> boxplot(rich ~ graze, data = grass, range = 0, horizontal = TRUE)
> title(ylab = 'cutting treatment', xlab = 'species richness')
```

