Course: Visual Analytics of large-scale biological data

Kay Nieselt

Center for Bioinformatics Tübingen
University of Tübingen

HANDS-ON BASIC VIS

Overview

- A short and painless intro to using R
- R on the web: Shiny
- ColorBrewer, Color Tool

Using R

- Box plots
- Bar charts
- Scatter plots
- ggplot

R scripts

- A text file (e.g. lab1.r) that contains all your R code
- Scientific method: complete record of your analyses
- Reproducible: rerunning your code is easy for you or someone else
- Easily modified and rerun
- In RStudio, select code and type <ctrl +enter> to run the code in the R console
- SAVE YOUR SCRIPTS

In-class exercise 1

- Use R to do the following. Create a new script to save your work, and remember to use driver & navigator roles.
- Now use R to compute:

  - $1 + 2(3 + 4)$
  - $\ln(4^3+3^{2+1})$
  - $\sqrt{(4+3)(2+1)}$
  - $\left(\frac{1+2}{3+4}\right)^2$

In-class exercise 2

- Create vectors using `seq()` and `rep()`. Only use `c()` if absolutely necessary
  - Positive integers from 1 to 99
  - Odd integers between 1 and 99
  - The numbers 1,1,1, 2,2,2, 3,3,3
  - The numbers 1,2,3,4,5,4,3,2,1,0
  - The fractions 1, 1/2, 1/3, 1/4, …, 1/10
  - The cubes 1, 8, 27, 64, 125, 216

A small R quiz

1.

```
> qplot(Sepal.Width, Sepal.Length, data = iris)
Error: could not find function "qplot“
```

Why?

2.

```
> make_list <- function() {
            list(date = Sys.Date(),
                 time = Sys.time(),
                 timezone = Sys.timezone())
}
> make_list$time
Error in make_list$time : object of type 'closure' is not subsettable
```

Why?

A small R quiz

3.

```
> lst <- list(numbers = 1:10, letters = letters, boolean = c(TRUE, FALSE))
```

- How can you access the numbers of lst?

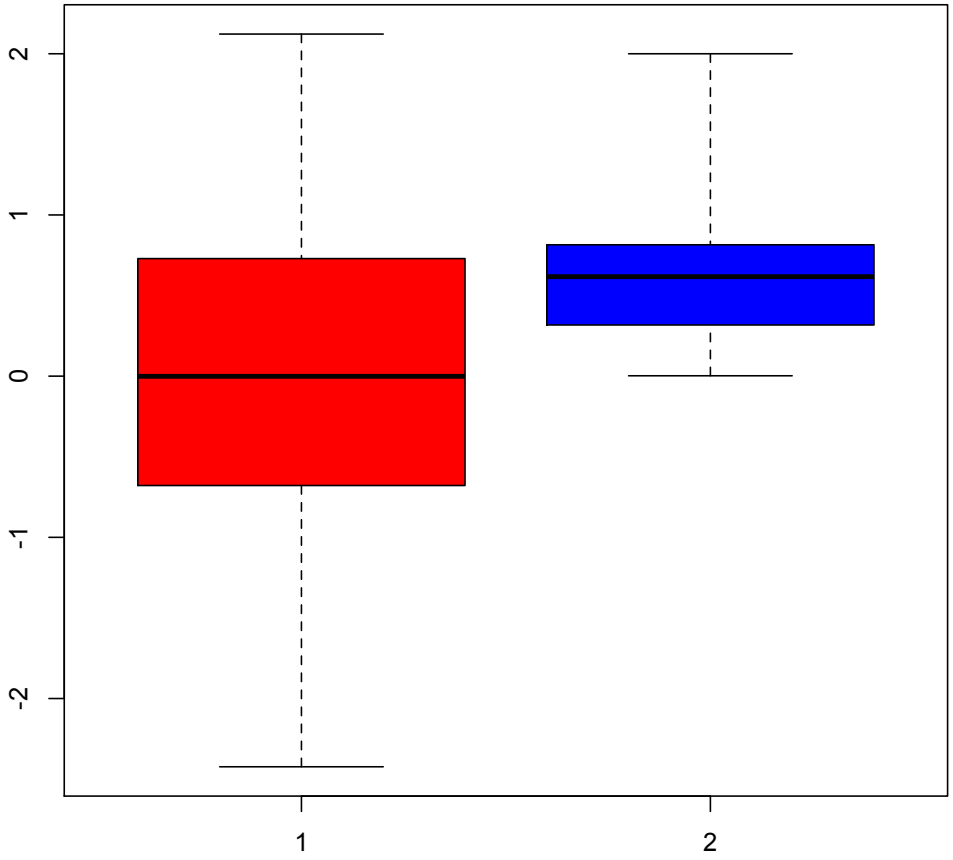
A small R quiz

4. Here is a data frame called pressure that comes with R. How can you calculate the sum of its temperature column?

```
> pressure
##    temperature pressure
## 1            0   0.0002
## 2           20   0.0012
## 3           40   0.0060
## 4           60   0.0300
## 5           80   0.0900
## 6          100   0.2700
## 7          120   0.7500
## 8          140   1.8500
## 9          160   4.2000
## 10         180   8.8000
## 11         200  17.3000
## 12         220  32.1000
## 13         240  57.0000
## 14         260  96.0000
## 15         280 157.0000
## 16         300 247.0000
## 17         320 376.0000
## 18         340 558.0000
## 19         360 806.0000
```

Boxplot and R

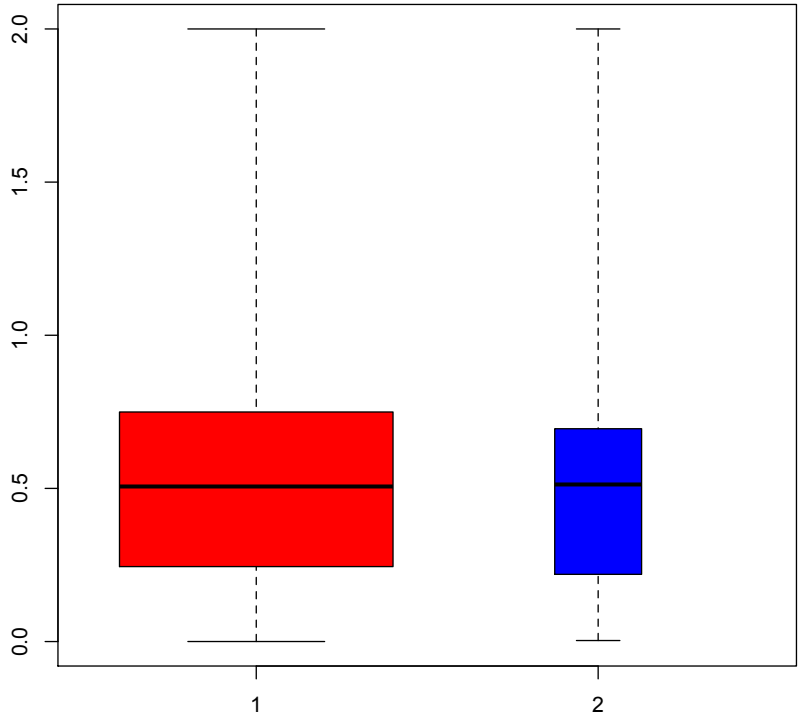

```
y <- c(runif(100),2.0)
x <- c(rnorm(100),2.0)
```



Boxplot and R

Sample size differences can be assessed by scaling the box plot width in proportion to √n, the factor by which the precision of the sample's estimate of population statistics improves as sample size is increased.

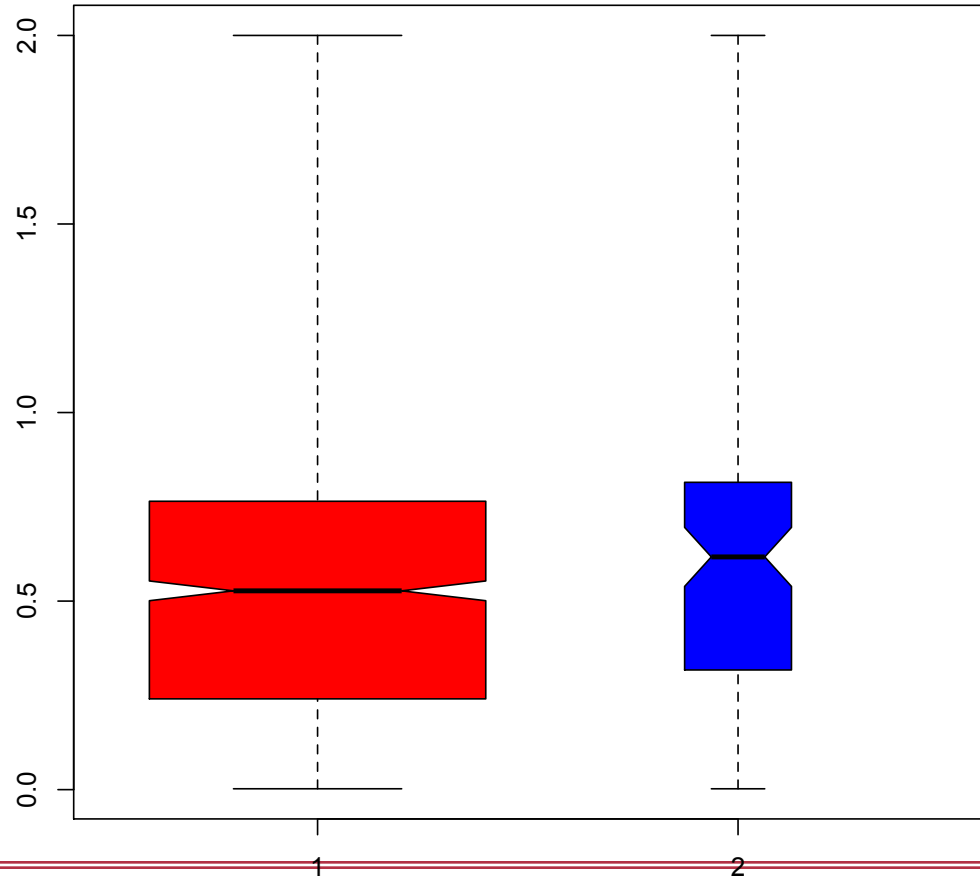
```
y <- c(runif(100),2.0)
x <- c(runif(1000),2.0)
```



Boxplot and R

A *notch* can be used to show the 95% confidence interval for the median m, given by

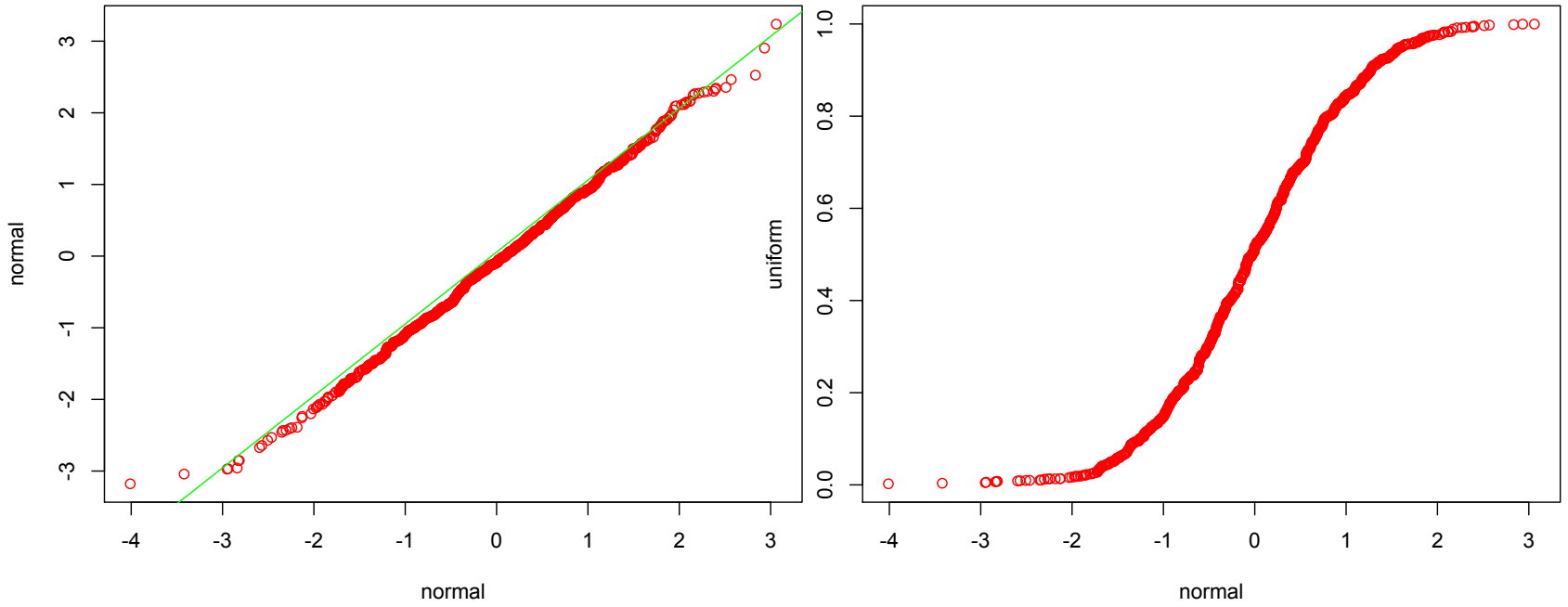
$$m \pm 1.58 \times IQR/\sqrt{n}$$



Boxplots – Online tool

- Boxplot online
- Users can upload data, create and label the plot and export in common formats.
- Play around and produce your favorite plot

QQ plots: example



ggplot2


```
install.packages(„ggplot2“)
library(ggplot2)


?mpg
head(mpg) ## shows top of data
str(mpg) ## shows structure of data
summary(mpg) ## summarizes data, alternative to str
boxplot(mpg$displ)
qplot(displ, hwy, data = mpg) # qplot=quick plot
```

ggplot2


```
ggplot(mtcars, aes(x=disp, y=mpg)) + geom_point()
```

Continue if you wish on your own.

Bar charts

- http://www.cookbook-r.com/Graphs/Bar_and_line_graphs_(ggplot2)/

```
> library(RColorBrewer)
> data <- read.delim("ExampleBarchart.txt“
> sequential <- brewer.pal(5, "BuGn“)
> barplot(as.matrix(data[,2:6]),col=sequential)
> barplot(as.matrix(t(data[,
2:6])),beside=TRUE,col=sequential)
```

Shiny

- A web application framework for R
  http://shiny.rstudio.com
- With Tutorials etc., very nice and easy to get started with.

Color chooser

- ColorBrewer: http://colorbrewer2.org

- Color Tool: http://colorusage.arc.nasa.gov/ColorTool.php#1

Feedback on Color Chooser Tools

- Much easier to use ColorBrewer if you just want a couple of colors!
- ColorTool is certainly powerful, but needs to time to get to know it.
- Starting it was also a bit hindered.