

Session one

Mees van der Ent

10/12/2021

Chapter 1

1.1: Common exercise

Create two arrays x and y and plot as coordinates in scatterplot.

```
x <- c(2, 4, 7, 10)
y <- c(1, 10, 11, 18)
```

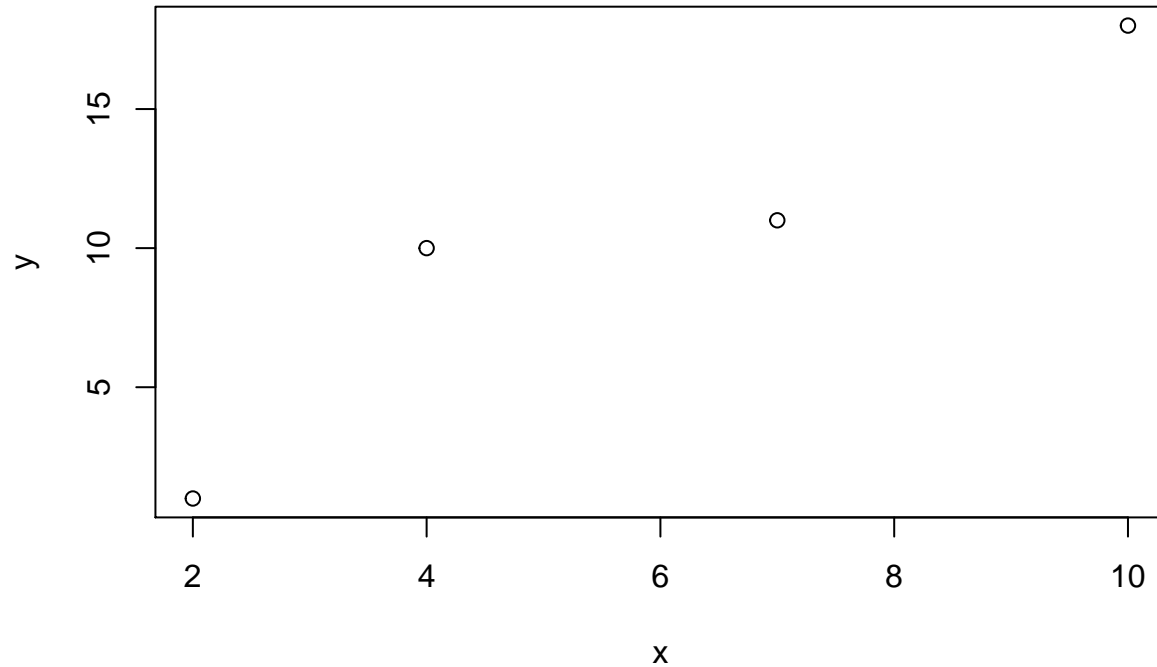
```
x
```

```
## [1]  2  4  7 10
```

```
y
```

```
## [1]  1 10 11 18
```

```
plot(x, y)
```



Chapter 2

2.1: Vectors

- Create the object test1 with numbers 1.5, 0.7, 45.6.
- Create a vector y1 with the numbers from 1 to 10.
- Create a logical vector y2 from y1. An element of y2 should be TRUE if the corresponding element of y1 is larger than 5.
- How many elements from y1 has a value larger than 5?

```
test1 <- c(1.5, 0.7, 45.6)
y1 <- 1:10

y2 <- y1>5
sum(y2)
```

```
## [1] 5
```

2.2: creating sequences

- Create a vector x with elements (1, 2, 3, . . . , 100)

```
x <- 1:100
x
```

```
## [1] 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
## [19] 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36
## [37] 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54
## [55] 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72
## [73] 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
## [91] 91 92 93 94 95 96 97 98 99 100
```

- Create a vector y with elements (0, 5, 10, 15, . . . , 500)

```
y <- seq(from=0, to=500, by=5)
y
```

```
## [1] 0 5 10 15 20 25 30 35 40 45 50 55 60 65 70 75 80 85
## [19] 90 95 100 105 110 115 120 125 130 135 140 145 150 155 160 165 170 175
## [37] 180 185 190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265
## [55] 270 275 280 285 290 295 300 305 310 315 320 325 330 335 340 345 350 355
## [73] 360 365 370 375 380 385 390 395 400 405 410 415 420 425 430 435 440 445
## [91] 450 455 460 465 470 475 480 485 490 495 500
```

- Create a vector z1 with elements (1, 1, 1, 2, 2, 2, . . . , 50, 50, 50)

```
z1 <- rep(1:50, rep(3, 50))
z1
```

```
## [1] 1 1 1 2 2 2 3 3 3 4 4 4 5 5 5 6 6 6 7 7 7 8 8 8 9
## [26] 9 9 10 10 10 11 11 11 12 12 12 13 13 13 14 14 14 15 15 15 16 16 16 17 17
## [51] 17 18 18 18 19 19 19 20 20 20 21 21 21 22 22 22 23 23 23 24 24 24 25 25 25
## [76] 26 26 26 27 27 27 28 28 28 29 29 29 30 30 30 31 31 31 32 32 32 33 33 33 34
## [101] 34 34 35 35 35 36 36 36 37 37 37 38 38 38 39 39 39 40 40 40 41 41 41 42 42
## [126] 42 43 43 43 44 44 44 45 45 45 46 46 46 47 47 47 48 48 48 49 49 49 50 50 50
```

- Create a vector z2 with elements (1, 2, 2, 3, 3, 3, . . . , 10)

```
z2 <- rep(1:10, 1:10)
z2
```

```
## [1] 1 2 2 3 3 3 4 4 4 4 5 5 5 5 5 6 6 6 6 6 6 7 7 7 7
## [26] 7 7 7 8 8 8 8 8 8 8 9 9 9 9 9 9 9 9 9 10 10 10 10 10
## [51] 10 10 10 10 10
```

e. Create a vector `z3` with elements (1, 2, 2, 3, 4, 4, 5, 6, 6, . . . , 50, 50)

```
z3 <- rep(1:50, rep(1:2, 25))
z3
```

```
## [1] 1 2 2 3 4 4 5 6 6 7 8 8 9 10 10 11 12 12 13 14 14 15 16 16 17
## [26] 18 18 19 20 20 21 22 22 23 24 24 25 26 26 27 28 28 29 30 30 31 32 32 33 34
## [51] 34 35 36 36 37 38 38 39 40 40 41 42 42 43 44 44 45 46 46 47 48 48 49 50 50
```

2.3: Matrix creation

a. Create a vector with 100 random normal numbers and use that to generate a 10 by 10 matrix. Call this matrix `mat1`. Hint: to generate a random normal vector, use the function `rnorm()`.

```
rand_dist <- rnorm(100)
mat1 <- matrix(rand_dist, nrow=10, ncol=10)
mat1
```

```
##           [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
## [1,] -1.0093561 -0.6511146 -0.98315521 1.64491901 -0.70103884 -1.3864666
## [2,] -1.3996639 1.0962518 -0.02609903 0.32340504 -0.02210597 0.6252110
## [3,] 2.5167278 0.2990437 0.52921646 0.29508841 -0.23769729 -0.3850374
## [4,] 1.2329209 1.6141044 -1.22504049 0.08597380 1.74097374 -0.6051858
## [5,] 0.9286965 -0.6777476 -0.09677723 -0.11033188 -1.05121637 -1.6343930
## [6,] -0.5779331 0.3801407 -2.26270651 -0.85493448 -0.42085476 -0.3412104
## [7,] 0.7562599 0.3510269 0.17792586 0.29739992 0.41334865 -0.5418801
## [8,] -0.1721980 0.9330015 -1.54322833 -0.70274700 -0.09129967 0.1044389
## [9,] -1.0759659 -3.1226393 -0.12902678 -0.09110797 -2.14557624 0.5508179
## [10,] -1.3338819 0.5643475 1.07403905 -0.47221462 0.69271174 -0.9970182
##           [,7]      [,8]      [,9]      [,10]
## [1,] 0.91506539 0.2588703 -0.55160974 0.3669499
## [2,] 0.13654562 0.0979607 0.06097416 0.9894304
## [3,] -0.59362113 -0.7315607 -0.38453290 -0.7658987
## [4,] -0.86549366 -0.2647643 -0.55371397 1.6994142
## [5,] -0.41261408 0.2986073 1.52177414 1.5226990
## [6,] -0.18730439 2.1060032 1.49923261 -1.2260024
## [7,] 1.16922909 -1.4167714 1.57778154 -0.1922548
## [8,] 0.04826081 2.3458081 -0.23864330 -0.8774377
## [9,] -1.41422192 -1.1722660 0.47737310 1.0584519
## [10,] 1.02055429 1.4643358 -0.85598102 -0.2706132
```

b. Add an extra row to `mat1` with the numbers 1 to 10 which will be the new first row. Also add the row with numbers 10 to 1 which will be the last row.

```
mat2 <- rbind(1:10, mat1, 10:1)
mat2
```

```
##           [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
## [1,] 1.0000000 2.0000000 3.00000000 4.00000000 5.00000000 6.0000000
## [2,] -1.0093561 -0.6511146 -0.98315521 1.64491901 -0.70103884 -1.3864666
## [3,] -1.3996639 1.0962518 -0.02609903 0.32340504 -0.02210597 0.6252110
## [4,] 2.5167278 0.2990437 0.52921646 0.29508841 -0.23769729 -0.3850374
## [5,] 1.2329209 1.6141044 -1.22504049 0.08597380 1.74097374 -0.6051858
## [6,] 0.9286965 -0.6777476 -0.09677723 -0.11033188 -1.05121637 -1.6343930
```

```
## [7,] -0.5779331 0.3801407 -2.26270651 -0.85493448 -0.42085476 -0.3412104
## [8,] 0.7562599 0.3510269 0.17792586 0.29739992 0.41334865 -0.5418801
## [9,] -0.1721980 0.9330015 -1.54322833 -0.70274700 -0.09129967 0.1044389
## [10,] -1.0759659 -3.1226393 -0.12902678 -0.09110797 -2.14557624 0.5508179
## [11,] -1.3338819 0.5643475 1.07403905 -0.47221462 0.69271174 -0.9970182
## [12,] 10.0000000 9.0000000 8.00000000 7.00000000 6.00000000 5.0000000
##      [,7]      [,8]      [,9]     [,10]
## [1,] 7.00000000 8.0000000 9.00000000 10.0000000
## [2,] 0.91506539 0.2588703 -0.55160974 0.3669499
## [3,] 0.13654562 0.0979607 0.06097416 0.9894304
## [4,] -0.59362113 -0.7315607 -0.38453290 -0.7658987
## [5,] -0.86549366 -0.2647643 -0.55371397 1.6994142
## [6,] -0.41261408 0.2986073 1.52177414 1.5226990
## [7,] -0.18730439 2.1060032 1.49923261 -1.2260024
## [8,] 1.16922909 -1.4167714 1.57778154 -0.1922548
## [9,] 0.04826081 2.3458081 -0.23864330 -0.8774377
## [10,] -1.41422192 -1.1722660 0.47737310 1.0584519
## [11,] 1.02055429 1.4643358 -0.85598102 -0.2706132
## [12,] 4.00000000 3.0000000 2.00000000 1.0000000
```

c. Add an extra column to the matrix obtained in step b with the number 1 to 12 (as first column).

```
mat3 <- cbind(1:12, mat2)
mat3
```

```
##      [,1]      [,2]      [,3]      [,4]      [,5]      [,6]      [,7]
## [1,] 1 1.0000000 2.0000000 3.00000000 4.00000000 5.00000000 6.0000000
## [2,] 2 -1.0093561 -0.6511146 -0.98315521 1.64491901 -0.70103884 -1.3864666
## [3,] 3 -1.3996639 1.0962518 -0.02609903 0.32340504 -0.02210597 0.6252110
## [4,] 4 2.5167278 0.2990437 0.52921646 0.29508841 -0.23769729 -0.3850374
## [5,] 5 1.2329209 1.6141044 -1.22504049 0.08597380 1.74097374 -0.6051858
## [6,] 6 0.9286965 -0.6777476 -0.09677723 -0.11033188 -1.05121637 -1.6343930
## [7,] 7 -0.5779331 0.3801407 -2.26270651 -0.85493448 -0.42085476 -0.3412104
## [8,] 8 0.7562599 0.3510269 0.17792586 0.29739992 0.41334865 -0.5418801
## [9,] 9 -0.1721980 0.9330015 -1.54322833 -0.70274700 -0.09129967 0.1044389
## [10,] 10 -1.0759659 -3.1226393 -0.12902678 -0.09110797 -2.14557624 0.5508179
## [11,] 11 -1.3338819 0.5643475 1.07403905 -0.47221462 0.69271174 -0.9970182
## [12,] 12 10.0000000 9.0000000 8.00000000 7.00000000 6.00000000 5.0000000
##      [,8]      [,9]      [,10]     [,11]
## [1,] 7.00000000 8.0000000 9.00000000 10.0000000
## [2,] 0.91506539 0.2588703 -0.55160974 0.3669499
## [3,] 0.13654562 0.0979607 0.06097416 0.9894304
## [4,] -0.59362113 -0.7315607 -0.38453290 -0.7658987
## [5,] -0.86549366 -0.2647643 -0.55371397 1.6994142
## [6,] -0.41261408 0.2986073 1.52177414 1.5226990
## [7,] -0.18730439 2.1060032 1.49923261 -1.2260024
## [8,] 1.16922909 -1.4167714 1.57778154 -0.1922548
## [9,] 0.04826081 2.3458081 -0.23864330 -0.8774377
## [10,] -1.41422192 -1.1722660 0.47737310 1.0584519
## [11,] 1.02055429 1.4643358 -0.85598102 -0.2706132
## [12,] 4.00000000 3.0000000 2.00000000 1.0000000
```

2.4: Working with data frames

a. Install and load the package *reshape*.

```
library(reshape)
```

```
## Warning: package 'reshape' was built under R version 3.4.4
```

b. Check the data description of this data frame.

```
?tips
```

c. Ask for the names of the variables in this data frame.

```
names(tips)
```

```
## [1] "total_bill" "tip"          "sex"          "smoker"       "day"
## [6] "time"       "size"
```

d. Take a subset of data tips which contains the observations from 1 until 20 and only the variables *tip*, *sex* and *day*.

```
tips_subset <- tips[1:20, c("tip", "sex", "day")]
tips_subset
```

```
##      tip    sex day
## 1  1.01 Female Sun
## 2  1.66   Male Sun
## 3  3.50   Male Sun
## 4  3.31   Male Sun
## 5  3.61 Female Sun
## 6  4.71   Male Sun
## 7  2.00   Male Sun
## 8  3.12   Male Sun
## 9  1.96   Male Sun
## 10 3.23   Male Sun
## 11 1.71   Male Sun
## 12 5.00 Female Sun
## 13 1.57   Male Sun
## 14 3.00   Male Sun
## 15 3.02 Female Sun
## 16 3.92   Male Sun
## 17 1.67 Female Sun
## 18 3.71   Male Sun
## 19 3.50 Female Sun
## 20 3.35   Male Sat
```