Tidyverse Introduction

Pietro Franceschi

Introduction

This .Rmd file is a concise summary of a set of advanced R libraries and ideas which can be used for the analysis of large scale datasets.

In particular:

- Pipes (`%>%`)
- Tabular data
  - `data.frame`, `tibble`, `data.table`
  - data carpentry (`tidyverse`, `data.table`)
  - *long* and *wide* tables
  - modeling (`broom`)
- Vectorizing operations (`purrr`)

```
library(tidyverse)     ## the full tidyverse ecosystem for seamless working with tables
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.2     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.2     v tibble    3.2.1
## v lubridate 1.9.2     v tidyr     1.3.0
## v purrr     1.0.1
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(broom)         ## a broom to tidy the outcomes of modeling
library(data.table)    ## a less flexible (bit more fast) approach to the manipulation of tabular data
```

```
##
## Attaching package: 'data.table'
##
## The following objects are masked from 'package:lubridate':
##
##     hour, isoweek, mday, minute, month, quarter, second, wday, week,
##     yday, year
##
## The following objects are masked from 'package:dplyr':
##
##     between, first, last
##
## The following object is masked from 'package:purrr':
##
##     transpose
```

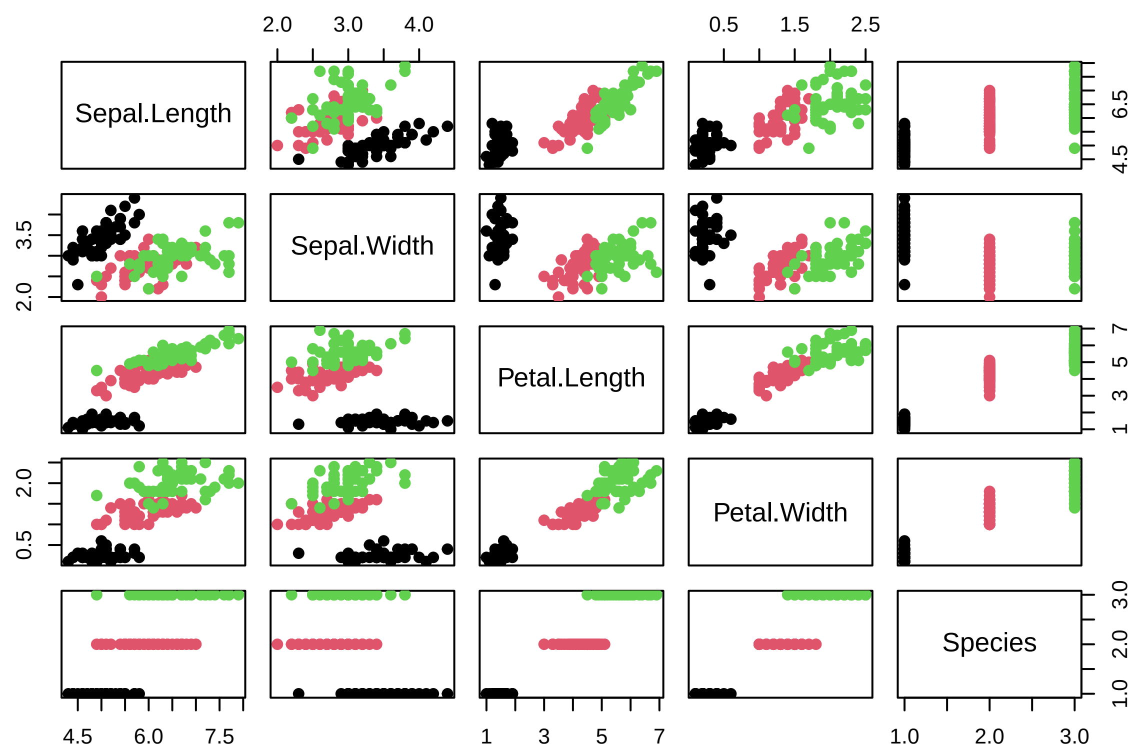
Piping

The overall idea behind piping is to make easy to read a chain of functions. Pipes `%>%` have been introduced in the `magrittr` package but are now a core tool of tidyverse

```
## sequence going from 1 to the square root of 100
one_old <-seq(1,sqrt(100))

## The old style work have to be read in an "onion" fashion

one_pipe <- sqrt(100) %>%
  seq(1,.)

## one_old and one_pipe are exactly equivalent, but the second is by far more easy to read
```

When putting a function in a pile, you refer to "what is coming from the pipe" with a dot `.`

```
## Pipes can be used also to produce plots!

data(iris) ## iris dataset

## note the dot to refer to the iris data.frame which reach the pairs function from the pipe
iris %>% pairs(., col = factor(iris$Species), pch = 19)
```



Advantages

- Clear writing
- No need of cluttering the workspace with intermediate objects

Disadvantages

- no big disadvantages, even if I normally rely on them when I use the console, while in programming tasks I prefer to rely on the "old" onion approach

Tabular Data

In R tabular data are commonly treated with three different classes of objects

- `data.frame`: old, faithful and the father of almost everything else. A data frame is a list
- `data.table`: basically still a data.frame which have been optimized for efficiency
- `tibble`: is the tidyverse form of data.frames, less efficient than data.table, but more flexyble since it is integrated in the tidy environment

Both data.tables and tibble retains the characteristics of dfs (indexing with square brackets, possible use of $ to get the columns). Importantly, in both cases the row.names attribute has been removed.

These three box of code allow to benchmark the efficiency of the three solution sin reading a relatively big dataset (35 MB)

```
## Base R
system.time(read.csv("athlete_events.csv"))
```

```
##    user  system elapsed
##     1.7     0.1     1.8
```

```
# data.table
system.time(fread("athlete_events.csv"))
```

```
##    user  system elapsed
##    0.21    0.01    0.14
```

```
# tidyverse
system.time(read_csv("athlete_events.csv"))
```

```
## Rows: 271116 Columns: 15
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (10): Name, Sex, Team, NOC, Games, Season, City, Sport, Event, Medal
## dbl  (5): ID, Age, Height, Weight, Year
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
##    user  system elapsed
##    1.08    0.14    0.74
```

As it can be seen data.table is almost ten times faster than base R, while tidyverse stays somehow in the middle

tibbles and data.tables also own an improved print method that allows for a more relaxed visualization of the content of the table

```
## read the three tables

baseR <- read.csv("athlete_events.csv")
datat <- fread("athlete_events.csv")
tidyv <- read_csv("athlete_events.csv")
```

```
## Rows: 271116 Columns: 15
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (10): Name, Sex, Team, NOC, Games, Season, City, Sport, Event, Medal
## dbl  (5): ID, Age, Height, Weight, Year
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
## this is the data.table printout
datat
```

```
##             ID                     Name Sex Age Height Weight           Team
##      1:      1                A Dijiang   M  24    180     80          China
##      2:      2                 A Lamusi   M  23    170     60          China
##      3:      3      Gunnar Nielsen Aaby   M  24     NA     NA        Denmark
##      4:      4     Edgar Lindenau Aabye   M  34     NA     NA Denmark/Sweden
##      5:      5 Christine Jacoba Aaftink   F  21    185     82    Netherlands
##     ---
## 271112: 135569               Andrzej ya   M  29    179     89       Poland-1
## 271113: 135570                 Piotr ya   M  27    176     59         Poland
## 271114: 135570                 Piotr ya   M  27    176     59         Poland
## 271115: 135571       Tomasz Ireneusz ya   M  30    185     96         Poland
## 271116: 135571       Tomasz Ireneusz ya   M  34    185     96         Poland
##         NOC       Games Year Season           City         Sport
##      1: CHN 1992 Summer 1992 Summer      Barcelona    Basketball
##      2: CHN 2012 Summer 2012 Summer         London          Judo
##      3: DEN 1920 Summer 1920 Summer      Antwerpen      Football
##      4: DEN 1900 Summer 1900 Summer          Paris    Tug-Of-War
##      5: NED 1988 Winter 1988 Winter        Calgary Speed Skating
##     ---
## 271112: POL 1976 Winter 1976 Winter      Innsbruck          Luge
## 271113: POL 2014 Winter 2014 Winter          Sochi   Ski Jumping
## 271114: POL 2014 Winter 2014 Winter          Sochi   Ski Jumping
## 271115: POL 1998 Winter 1998 Winter         Nagano     Bobsleigh
## 271116: POL 2002 Winter 2002 Winter Salt Lake City     Bobsleigh
##                                           Event Medal
##      1:             Basketball Men's Basketball  <NA>
##      2:            Judo Men's Extra-Lightweight  <NA>
##      3:                 Football Men's Football  <NA>
##      4:             Tug-Of-War Men's Tug-Of-War  Gold
##      5:        Speed Skating Women's 500 metres  <NA>
##     ---
## 271112:              Luge Mixed (Men)'s Doubles  <NA>
## 271113: Ski Jumping Men's Large Hill, Individual  <NA>
## 271114:       Ski Jumping Men's Large Hill, Team  <NA>
## 271115:                    Bobsleigh Men's Four  <NA>
## 271116:                    Bobsleigh Men's Four  <NA>
```

```
## this is the tidyverse printout
tidyv
```

```
## # A tibble: 271,116 x 15
##        ID Name     Sex     Age Height Weight Team  NOC   Games  Year Season City
##     <dbl> <chr>    <chr> <dbl>  <dbl>  <dbl> <chr> <chr> <chr> <dbl> <chr>  <chr>
##  1      1 A Dijia~ M        24    180     80 China CHN   1992~  1992 Summer Barc~
##  2      2 A Lamusi M        23    170     60 China CHN   2012~  2012 Summer Lond~
```

```
##   3      3 Gunnar ~ M        24      NA      NA Denm~ DEN   1920~  1920 Summer Antw~
##   4      4 Edgar L~ M        34      NA      NA Denm~ DEN   1900~  1900 Summer Paris
##   5      5 Christi~ F        21     185      82 Neth~ NED   1988~  1988 Winter Calg~
##   6      5 Christi~ F        21     185      82 Neth~ NED   1988~  1988 Winter Calg~
##   7      5 Christi~ F        25     185      82 Neth~ NED   1992~  1992 Winter Albe~
##   8      5 Christi~ F        25     185      82 Neth~ NED   1992~  1992 Winter Albe~
##   9      5 Christi~ F        27     185      82 Neth~ NED   1994~  1994 Winter Lill~
##  10      5 Christi~ F        27     185      82 Neth~ NED   1994~  1994 Winter Lill~
## # i 271,106 more rows
## # i 3 more variables: Sport <chr>, Event <chr>, Medal <chr>
```

Data carpentry intidyverse and `data.table`

Filtering rows on condition

It is helpful to summarize the data.table slicing approach:

`DT[i, j, by]` "Take DT, subset rows using i, then calculate j grouped by by"

```
## filtering with data.table
iris %>%
  data.table(.) %>%  ## this is needed to transform the iris data.frame to a data.table
  .[Species %in% c("setosa","versicolor"),]
```

```
##      Sepal.Length Sepal.Width Petal.Length Petal.Width    Species
##   1:          5.1         3.5          1.4         0.2     setosa
##   2:          4.9         3.0          1.4         0.2     setosa
##   3:          4.7         3.2          1.3         0.2     setosa
##   4:          4.6         3.1          1.5         0.2     setosa
##   5:          5.0         3.6          1.4         0.2     setosa
##   6:          5.4         3.9          1.7         0.4     setosa
##   7:          4.6         3.4          1.4         0.3     setosa
##   8:          5.0         3.4          1.5         0.2     setosa
##   9:          4.4         2.9          1.4         0.2     setosa
##  10:          4.9         3.1          1.5         0.1     setosa
##  11:          5.4         3.7          1.5         0.2     setosa
##  12:          4.8         3.4          1.6         0.2     setosa
##  13:          4.8         3.0          1.4         0.1     setosa
##  14:          4.3         3.0          1.1         0.1     setosa
##  15:          5.8         4.0          1.2         0.2     setosa
##  16:          5.7         4.4          1.5         0.4     setosa
##  17:          5.4         3.9          1.3         0.4     setosa
##  18:          5.1         3.5          1.4         0.3     setosa
##  19:          5.7         3.8          1.7         0.3     setosa
##  20:          5.1         3.8          1.5         0.3     setosa
##  21:          5.4         3.4          1.7         0.2     setosa
##  22:          5.1         3.7          1.5         0.4     setosa
##  23:          4.6         3.6          1.0         0.2     setosa
##  24:          5.1         3.3          1.7         0.5     setosa
##  25:          4.8         3.4          1.9         0.2     setosa
##  26:          5.0         3.0          1.6         0.2     setosa
##  27:          5.0         3.4          1.6         0.4     setosa
##  28:          5.2         3.5          1.5         0.2     setosa
##  29:          5.2         3.4          1.4         0.2     setosa
##  30:          4.7         3.2          1.6         0.2     setosa
```

```
##  31:          4.8         3.1          1.6         0.2     setosa
##  32:          5.4         3.4          1.5         0.4     setosa
##  33:          5.2         4.1          1.5         0.1     setosa
##  34:          5.5         4.2          1.4         0.2     setosa
##  35:          4.9         3.1          1.5         0.2     setosa
##  36:          5.0         3.2          1.2         0.2     setosa
##  37:          5.5         3.5          1.3         0.2     setosa
##  38:          4.9         3.6          1.4         0.1     setosa
##  39:          4.4         3.0          1.3         0.2     setosa
##  40:          5.1         3.4          1.5         0.2     setosa
##  41:          5.0         3.5          1.3         0.3     setosa
##  42:          4.5         2.3          1.3         0.3     setosa
##  43:          4.4         3.2          1.3         0.2     setosa
##  44:          5.0         3.5          1.6         0.6     setosa
##  45:          5.1         3.8          1.9         0.4     setosa
##  46:          4.8         3.0          1.4         0.3     setosa
##  47:          5.1         3.8          1.6         0.2     setosa
##  48:          4.6         3.2          1.4         0.2     setosa
##  49:          5.3         3.7          1.5         0.2     setosa
##  50:          5.0         3.3          1.4         0.2     setosa
##  51:          7.0         3.2          4.7         1.4 versicolor
##  52:          6.4         3.2          4.5         1.5 versicolor
##  53:          6.9         3.1          4.9         1.5 versicolor
##  54:          5.5         2.3          4.0         1.3 versicolor
##  55:          6.5         2.8          4.6         1.5 versicolor
##  56:          5.7         2.8          4.5         1.3 versicolor
##  57:          6.3         3.3          4.7         1.6 versicolor
##  58:          4.9         2.4          3.3         1.0 versicolor
##  59:          6.6         2.9          4.6         1.3 versicolor
##  60:          5.2         2.7          3.9         1.4 versicolor
##  61:          5.0         2.0          3.5         1.0 versicolor
##  62:          5.9         3.0          4.2         1.5 versicolor
##  63:          6.0         2.2          4.0         1.0 versicolor
##  64:          6.1         2.9          4.7         1.4 versicolor
##  65:          5.6         2.9          3.6         1.3 versicolor
##  66:          6.7         3.1          4.4         1.4 versicolor
##  67:          5.6         3.0          4.5         1.5 versicolor
##  68:          5.8         2.7          4.1         1.0 versicolor
##  69:          6.2         2.2          4.5         1.5 versicolor
##  70:          5.6         2.5          3.9         1.1 versicolor
##  71:          5.9         3.2          4.8         1.8 versicolor
##  72:          6.1         2.8          4.0         1.3 versicolor
##  73:          6.3         2.5          4.9         1.5 versicolor
##  74:          6.1         2.8          4.7         1.2 versicolor
##  75:          6.4         2.9          4.3         1.3 versicolor
##  76:          6.6         3.0          4.4         1.4 versicolor
##  77:          6.8         2.8          4.8         1.4 versicolor
##  78:          6.7         3.0          5.0         1.7 versicolor
##  79:          6.0         2.9          4.5         1.5 versicolor
##  80:          5.7         2.6          3.5         1.0 versicolor
##  81:          5.5         2.4          3.8         1.1 versicolor
##  82:          5.5         2.4          3.7         1.0 versicolor
##  83:          5.8         2.7          3.9         1.2 versicolor
##  84:          6.0         2.7          5.1         1.6 versicolor
```

```
## 85:          5.4         3.0          4.5         1.5 versicolor
## 86:          6.0         3.4          4.5         1.6 versicolor
## 87:          6.7         3.1          4.7         1.5 versicolor
## 88:          6.3         2.3          4.4         1.3 versicolor
## 89:          5.6         3.0          4.1         1.3 versicolor
## 90:          5.5         2.5          4.0         1.3 versicolor
## 91:          5.5         2.6          4.4         1.2 versicolor
## 92:          6.1         3.0          4.6         1.4 versicolor
## 93:          5.8         2.6          4.0         1.2 versicolor
## 94:          5.0         2.3          3.3         1.0 versicolor
## 95:          5.6         2.7          4.2         1.3 versicolor
## 96:          5.7         3.0          4.2         1.2 versicolor
## 97:          5.7         2.9          4.2         1.3 versicolor
## 98:          6.2         2.9          4.3         1.3 versicolor
## 99:          5.1         2.5          3.0         1.1 versicolor
## 100:         5.7         2.8          4.1         1.3 versicolor
##      Sepal.Length Sepal.Width Petal.Length Petal.Width    Species
```

```
## Note here I'm fitting a data.table slicing into a tidyverse like piping style. The . before the squa
## data table coming from the pipe
```

The syntax of tidyverse is slightly more verbose but extremely easy to read

```
iris %>%
  tibble() %>%
  filter(Species == "setosa") ## the function filter selects rows on condition
```

```
## # A tibble: 50 x 5
##    Sepal.Length Sepal.Width Petal.Length Petal.Width Species
##           <dbl>       <dbl>        <dbl>       <dbl> <fct>
##  1          5.1         3.5          1.4         0.2 setosa
##  2          4.9         3            1.4         0.2 setosa
##  3          4.7         3.2          1.3         0.2 setosa
##  4          4.6         3.1          1.5         0.2 setosa
##  5          5           3.6          1.4         0.2 setosa
##  6          5.4         3.9          1.7         0.4 setosa
##  7          4.6         3.4          1.4         0.3 setosa
##  8          5           3.4          1.5         0.2 setosa
##  9          4.4         2.9          1.4         0.2 setosa
## 10          4.9         3.1          1.5         0.1 setosa
## # i 40 more rows
```

```
## Note: as we discussed in the lecture, the tibble() call is not necessary, since the coercion to a ti
## is performed under the hood by the fact that you are using pipes
```

For the large majority of situations data.tables and tibbles can be used interchangeably, but remember that everything for data table has been optimized for speed. This is true for calculations, selections and sorting

Selecting columns

If you consider the previous general syntax, you see that selecting columns in dt is fast. Remember the comma!

```
## Extracting several columns as a data.table
iris %>%
  data.table(.) %>%
  .[,list(Species,Sepal.Width)]
```

```
##        Species Sepal.Width
##   1:    setosa         3.5
##   2:    setosa         3.0
##   3:    setosa         3.2
##   4:    setosa         3.1
##   5:    setosa         3.6
##  ---
## 146: virginica         3.0
## 147: virginica         2.5
## 148: virginica         3.0
## 149: virginica         3.4
## 150: virginica         3.0
```

```
## extracting one column as as vector
iris %>%
  data.table(.) %>%
  .[,Species]
```

```
##   [1] setosa     setosa     setosa     setosa     setosa     setosa
##   [7] setosa     setosa     setosa     setosa     setosa     setosa
##  [13] setosa     setosa     setosa     setosa     setosa     setosa
##  [19] setosa     setosa     setosa     setosa     setosa     setosa
##  [25] setosa     setosa     setosa     setosa     setosa     setosa
##  [31] setosa     setosa     setosa     setosa     setosa     setosa
##  [37] setosa     setosa     setosa     setosa     setosa     setosa
##  [43] setosa     setosa     setosa     setosa     setosa     setosa
##  [49] setosa     setosa     versicolor versicolor versicolor versicolor
##  [55] versicolor versicolor versicolor versicolor versicolor versicolor
##  [61] versicolor versicolor versicolor versicolor versicolor versicolor
##  [67] versicolor versicolor versicolor versicolor versicolor versicolor
##  [73] versicolor versicolor versicolor versicolor versicolor versicolor
##  [79] versicolor versicolor versicolor versicolor versicolor versicolor
##  [85] versicolor versicolor versicolor versicolor versicolor versicolor
##  [91] versicolor versicolor versicolor versicolor versicolor versicolor
##  [97] versicolor versicolor versicolor versicolor virginica  virginica
## [103] virginica  virginica  virginica  virginica  virginica  virginica
## [109] virginica  virginica  virginica  virginica  virginica  virginica
## [115] virginica  virginica  virginica  virginica  virginica  virginica
## [121] virginica  virginica  virginica  virginica  virginica  virginica
## [127] virginica  virginica  virginica  virginica  virginica  virginica
## [133] virginica  virginica  virginica  virginica  virginica  virginica
## [139] virginica  virginica  virginica  virginica  virginica  virginica
## [145] virginica  virginica  virginica  virginica  virginica  virginica
## Levels: setosa versicolor virginica
```

Obviously, the selection of many columns can be performed by using another dot … just to make life more clear ..

```
## Extracting several columns as a data.table
iris %>%
  data.table(.) %>%
  .[,.(Species,Sepal.Width)]
```

```
##        Species Sepal.Width
##   1:    setosa         3.5
##   2:    setosa         3.0
```

```
##   3:    setosa         3.2
##   4:    setosa         3.1
##   5:    setosa         3.6
##  ---
## 146: virginica        3.0
## 147: virginica        2.5
## 148: virginica        3.0
## 149: virginica        3.4
## 150: virginica        3.0
```

Let's call this, dot deluge . . . ;-)

In tidyverse the extraction of a single column as a vector is performed by the `pull` function, while the selection of one or more column resulting in a smaller tibble is performed by the `select` function

```
## pull one column as vector
iris %>%
  pull(Species)
```

```
##   [1] setosa     setosa     setosa     setosa     setosa     setosa
##   [7] setosa     setosa     setosa     setosa     setosa     setosa
##  [13] setosa     setosa     setosa     setosa     setosa     setosa
##  [19] setosa     setosa     setosa     setosa     setosa     setosa
##  [25] setosa     setosa     setosa     setosa     setosa     setosa
##  [31] setosa     setosa     setosa     setosa     setosa     setosa
##  [37] setosa     setosa     setosa     setosa     setosa     setosa
##  [43] setosa     setosa     setosa     setosa     setosa     setosa
##  [49] setosa     setosa     versicolor versicolor versicolor versicolor
##  [55] versicolor versicolor versicolor versicolor versicolor versicolor
##  [61] versicolor versicolor versicolor versicolor versicolor versicolor
##  [67] versicolor versicolor versicolor versicolor versicolor versicolor
##  [73] versicolor versicolor versicolor versicolor versicolor versicolor
##  [79] versicolor versicolor versicolor versicolor versicolor versicolor
##  [85] versicolor versicolor versicolor versicolor versicolor versicolor
##  [91] versicolor versicolor versicolor versicolor versicolor versicolor
##  [97] versicolor versicolor versicolor versicolor virginica  virginica
## [103] virginica  virginica  virginica  virginica  virginica  virginica
## [109] virginica  virginica  virginica  virginica  virginica  virginica
## [115] virginica  virginica  virginica  virginica  virginica  virginica
## [121] virginica  virginica  virginica  virginica  virginica  virginica
## [127] virginica  virginica  virginica  virginica  virginica  virginica
## [133] virginica  virginica  virginica  virginica  virginica  virginica
## [139] virginica  virginica  virginica  virginica  virginica  virginica
## [145] virginica  virginica  virginica  virginica  virginica  virginica
## Levels: setosa versicolor virginica
```

```
## select two columns and return a tibble
iris %>%
  select(Sepal.Length,Species)
```

```
##     Sepal.Length    Species
## 1            5.1     setosa
## 2            4.9     setosa
## 3            4.7     setosa
## 4            4.6     setosa
## 5            5.0     setosa
```

```
## 6              5.4     setosa
## 7              4.6     setosa
## 8              5.0     setosa
## 9              4.4     setosa
## 10             4.9     setosa
## 11             5.4     setosa
## 12             4.8     setosa
## 13             4.8     setosa
## 14             4.3     setosa
## 15             5.8     setosa
## 16             5.7     setosa
## 17             5.4     setosa
## 18             5.1     setosa
## 19             5.7     setosa
## 20             5.1     setosa
## 21             5.4     setosa
## 22             5.1     setosa
## 23             4.6     setosa
## 24             5.1     setosa
## 25             4.8     setosa
## 26             5.0     setosa
## 27             5.0     setosa
## 28             5.2     setosa
## 29             5.2     setosa
## 30             4.7     setosa
## 31             4.8     setosa
## 32             5.4     setosa
## 33             5.2     setosa
## 34             5.5     setosa
## 35             4.9     setosa
## 36             5.0     setosa
## 37             5.5     setosa
## 38             4.9     setosa
## 39             4.4     setosa
## 40             5.1     setosa
## 41             5.0     setosa
## 42             4.5     setosa
## 43             4.4     setosa
## 44             5.0     setosa
## 45             5.1     setosa
## 46             4.8     setosa
## 47             5.1     setosa
## 48             4.6     setosa
## 49             5.3     setosa
## 50             5.0     setosa
## 51             7.0 versicolor
## 52             6.4 versicolor
## 53             6.9 versicolor
## 54             5.5 versicolor
## 55             6.5 versicolor
## 56             5.7 versicolor
## 57             6.3 versicolor
## 58             4.9 versicolor
## 59             6.6 versicolor
```

```
## 60          5.2 versicolor
## 61          5.0 versicolor
## 62          5.9 versicolor
## 63          6.0 versicolor
## 64          6.1 versicolor
## 65          5.6 versicolor
## 66          6.7 versicolor
## 67          5.6 versicolor
## 68          5.8 versicolor
## 69          6.2 versicolor
## 70          5.6 versicolor
## 71          5.9 versicolor
## 72          6.1 versicolor
## 73          6.3 versicolor
## 74          6.1 versicolor
## 75          6.4 versicolor
## 76          6.6 versicolor
## 77          6.8 versicolor
## 78          6.7 versicolor
## 79          6.0 versicolor
## 80          5.7 versicolor
## 81          5.5 versicolor
## 82          5.5 versicolor
## 83          5.8 versicolor
## 84          6.0 versicolor
## 85          5.4 versicolor
## 86          6.0 versicolor
## 87          6.7 versicolor
## 88          6.3 versicolor
## 89          5.6 versicolor
## 90          5.5 versicolor
## 91          5.5 versicolor
## 92          6.1 versicolor
## 93          5.8 versicolor
## 94          5.0 versicolor
## 95          5.6 versicolor
## 96          5.7 versicolor
## 97          5.7 versicolor
## 98          6.2 versicolor
## 99          5.1 versicolor
## 100         5.7 versicolor
## 101         6.3  virginica
## 102         5.8  virginica
## 103         7.1  virginica
## 104         6.3  virginica
## 105         6.5  virginica
## 106         7.6  virginica
## 107         4.9  virginica
## 108         7.3  virginica
## 109         6.7  virginica
## 110         7.2  virginica
## 111         6.5  virginica
## 112         6.4  virginica
## 113         6.8  virginica
```

```
## 114          5.7  virginica
## 115          5.8  virginica
## 116          6.4  virginica
## 117          6.5  virginica
## 118          7.7  virginica
## 119          7.7  virginica
## 120          6.0  virginica
## 121          6.9  virginica
## 122          5.6  virginica
## 123          7.7  virginica
## 124          6.3  virginica
## 125          6.7  virginica
## 126          7.2  virginica
## 127          6.2  virginica
## 128          6.1  virginica
## 129          6.4  virginica
## 130          7.2  virginica
## 131          7.4  virginica
## 132          7.9  virginica
## 133          6.4  virginica
## 134          6.3  virginica
## 135          6.1  virginica
## 136          7.7  virginica
## 137          6.3  virginica
## 138          6.4  virginica
## 139          6.0  virginica
## 140          6.9  virginica
## 141          6.7  virginica
## 142          6.9  virginica
## 143          5.8  virginica
## 144          6.8  virginica
## 145          6.7  virginica
## 146          6.7  virginica
## 147          6.3  virginica
## 148          6.5  virginica
## 149          6.2  virginica
## 150          5.9  virginica
```

An interesting and useful characteristic of `select` is the possibility of using a series of selection helpers to identify columns on the base of their properties. See the help of select for a more detailed description

```
## extract all the column with a name starting with sepal
iris %>%
  tibble() %>%
  select(starts_with("Sepal"))
```

```
## # A tibble: 150 x 2
##    Sepal.Length Sepal.Width
##           <dbl>       <dbl>
##  1          5.1         3.5
##  2          4.9         3
##  3          4.7         3.2
##  4          4.6         3.1
##  5          5           3.6
##  6          5.4         3.9
```

```
##  7          4.6         3.4
##  8          5           3.4
##  9          4.4         2.9
## 10          4.9         3.1
## # i 140 more rows
```

```
## interesting! getting only numeric columns
iris %>%
  tibble() %>%
  select(where(~is.numeric(.x)))
```

```
## # A tibble: 150 x 4
##    Sepal.Length Sepal.Width Petal.Length Petal.Width
##           <dbl>       <dbl>        <dbl>       <dbl>
##  1          5.1         3.5          1.4         0.2
##  2          4.9         3            1.4         0.2
##  3          4.7         3.2          1.3         0.2
##  4          4.6         3.1          1.5         0.2
##  5          5           3.6          1.4         0.2
##  6          5.4         3.9          1.7         0.4
##  7          4.6         3.4          1.4         0.3
##  8          5           3.4          1.5         0.2
##  9          4.4         2.9          1.4         0.2
## 10          4.9         3.1          1.5         0.1
## # i 140 more rows
```

Unfortunately we have another dot …

Note: the writing `~is.numeric(x)` could seem wired. This is a special shortcut to construct *functionals*. Tidysomething will transform formulas starting with `~` into functions.

There are shorthands to refer to their arguments. For functions with one argument you can use the dot! For one or two (`.x` and `.y`), for an arbitrary number of arguments `..1`,`..2`, `..3`, etc.

So in our case, the following three constructs are equivalent

```
iris %>%
  select(where(function(c) is.numeric(c)))
```

```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width
## 1            5.1         3.5          1.4         0.2
## 2            4.9         3.0          1.4         0.2
## 3            4.7         3.2          1.3         0.2
## 4            4.6         3.1          1.5         0.2
## 5            5.0         3.6          1.4         0.2
## 6            5.4         3.9          1.7         0.4
## 7            4.6         3.4          1.4         0.3
## 8            5.0         3.4          1.5         0.2
## 9            4.4         2.9          1.4         0.2
## 10           4.9         3.1          1.5         0.1
## 11           5.4         3.7          1.5         0.2
## 12           4.8         3.4          1.6         0.2
## 13           4.8         3.0          1.4         0.1
## 14           4.3         3.0          1.1         0.1
## 15           5.8         4.0          1.2         0.2
## 16           5.7         4.4          1.5         0.4
## 17           5.4         3.9          1.3         0.4
## 18           5.1         3.5          1.4         0.3
```

```
## 19          5.7         3.8          1.7         0.3
## 20          5.1         3.8          1.5         0.3
## 21          5.4         3.4          1.7         0.2
## 22          5.1         3.7          1.5         0.4
## 23          4.6         3.6          1.0         0.2
## 24          5.1         3.3          1.7         0.5
## 25          4.8         3.4          1.9         0.2
## 26          5.0         3.0          1.6         0.2
## 27          5.0         3.4          1.6         0.4
## 28          5.2         3.5          1.5         0.2
## 29          5.2         3.4          1.4         0.2
## 30          4.7         3.2          1.6         0.2
## 31          4.8         3.1          1.6         0.2
## 32          5.4         3.4          1.5         0.4
## 33          5.2         4.1          1.5         0.1
## 34          5.5         4.2          1.4         0.2
## 35          4.9         3.1          1.5         0.2
## 36          5.0         3.2          1.2         0.2
## 37          5.5         3.5          1.3         0.2
## 38          4.9         3.6          1.4         0.1
## 39          4.4         3.0          1.3         0.2
## 40          5.1         3.4          1.5         0.2
## 41          5.0         3.5          1.3         0.3
## 42          4.5         2.3          1.3         0.3
## 43          4.4         3.2          1.3         0.2
## 44          5.0         3.5          1.6         0.6
## 45          5.1         3.8          1.9         0.4
## 46          4.8         3.0          1.4         0.3
## 47          5.1         3.8          1.6         0.2
## 48          4.6         3.2          1.4         0.2
## 49          5.3         3.7          1.5         0.2
## 50          5.0         3.3          1.4         0.2
## 51          7.0         3.2          4.7         1.4
## 52          6.4         3.2          4.5         1.5
## 53          6.9         3.1          4.9         1.5
## 54          5.5         2.3          4.0         1.3
## 55          6.5         2.8          4.6         1.5
## 56          5.7         2.8          4.5         1.3
## 57          6.3         3.3          4.7         1.6
## 58          4.9         2.4          3.3         1.0
## 59          6.6         2.9          4.6         1.3
## 60          5.2         2.7          3.9         1.4
## 61          5.0         2.0          3.5         1.0
## 62          5.9         3.0          4.2         1.5
## 63          6.0         2.2          4.0         1.0
## 64          6.1         2.9          4.7         1.4
## 65          5.6         2.9          3.6         1.3
## 66          6.7         3.1          4.4         1.4
## 67          5.6         3.0          4.5         1.5
## 68          5.8         2.7          4.1         1.0
## 69          6.2         2.2          4.5         1.5
## 70          5.6         2.5          3.9         1.1
## 71          5.9         3.2          4.8         1.8
## 72          6.1         2.8          4.0         1.3
```

```
## 73          6.3          2.5          4.9          1.5
## 74          6.1          2.8          4.7          1.2
## 75          6.4          2.9          4.3          1.3
## 76          6.6          3.0          4.4          1.4
## 77          6.8          2.8          4.8          1.4
## 78          6.7          3.0          5.0          1.7
## 79          6.0          2.9          4.5          1.5
## 80          5.7          2.6          3.5          1.0
## 81          5.5          2.4          3.8          1.1
## 82          5.5          2.4          3.7          1.0
## 83          5.8          2.7          3.9          1.2
## 84          6.0          2.7          5.1          1.6
## 85          5.4          3.0          4.5          1.5
## 86          6.0          3.4          4.5          1.6
## 87          6.7          3.1          4.7          1.5
## 88          6.3          2.3          4.4          1.3
## 89          5.6          3.0          4.1          1.3
## 90          5.5          2.5          4.0          1.3
## 91          5.5          2.6          4.4          1.2
## 92          6.1          3.0          4.6          1.4
## 93          5.8          2.6          4.0          1.2
## 94          5.0          2.3          3.3          1.0
## 95          5.6          2.7          4.2          1.3
## 96          5.7          3.0          4.2          1.2
## 97          5.7          2.9          4.2          1.3
## 98          6.2          2.9          4.3          1.3
## 99          5.1          2.5          3.0          1.1
## 100         5.7          2.8          4.1          1.3
## 101         6.3          3.3          6.0          2.5
## 102         5.8          2.7          5.1          1.9
## 103         7.1          3.0          5.9          2.1
## 104         6.3          2.9          5.6          1.8
## 105         6.5          3.0          5.8          2.2
## 106         7.6          3.0          6.6          2.1
## 107         4.9          2.5          4.5          1.7
## 108         7.3          2.9          6.3          1.8
## 109         6.7          2.5          5.8          1.8
## 110         7.2          3.6          6.1          2.5
## 111         6.5          3.2          5.1          2.0
## 112         6.4          2.7          5.3          1.9
## 113         6.8          3.0          5.5          2.1
## 114         5.7          2.5          5.0          2.0
## 115         5.8          2.8          5.1          2.4
## 116         6.4          3.2          5.3          2.3
## 117         6.5          3.0          5.5          1.8
## 118         7.7          3.8          6.7          2.2
## 119         7.7          2.6          6.9          2.3
## 120         6.0          2.2          5.0          1.5
## 121         6.9          3.2          5.7          2.3
## 122         5.6          2.8          4.9          2.0
## 123         7.7          2.8          6.7          2.0
## 124         6.3          2.7          4.9          1.8
## 125         6.7          3.3          5.7          2.1
## 126         7.2          3.2          6.0          1.8
```

```
## 127          6.2         2.8          4.8         1.8
## 128          6.1         3.0          4.9         1.8
## 129          6.4         2.8          5.6         2.1
## 130          7.2         3.0          5.8         1.6
## 131          7.4         2.8          6.1         1.9
## 132          7.9         3.8          6.4         2.0
## 133          6.4         2.8          5.6         2.2
## 134          6.3         2.8          5.1         1.5
## 135          6.1         2.6          5.6         1.4
## 136          7.7         3.0          6.1         2.3
## 137          6.3         3.4          5.6         2.4
## 138          6.4         3.1          5.5         1.8
## 139          6.0         3.0          4.8         1.8
## 140          6.9         3.1          5.4         2.1
## 141          6.7         3.1          5.6         2.4
## 142          6.9         3.1          5.1         2.3
## 143          5.8         2.7          5.1         1.9
## 144          6.8         3.2          5.9         2.3
## 145          6.7         3.3          5.7         2.5
## 146          6.7         3.0          5.2         2.3
## 147          6.3         2.5          5.0         1.9
## 148          6.5         3.0          5.2         2.0
## 149          6.2         3.4          5.4         2.3
## 150          5.9         3.0          5.1         1.8
```

```
iris %>%
  select(where(~is.numeric(.x)))
```

```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width
## 1            5.1         3.5          1.4         0.2
## 2            4.9         3.0          1.4         0.2
## 3            4.7         3.2          1.3         0.2
## 4            4.6         3.1          1.5         0.2
## 5            5.0         3.6          1.4         0.2
## 6            5.4         3.9          1.7         0.4
## 7            4.6         3.4          1.4         0.3
## 8            5.0         3.4          1.5         0.2
## 9            4.4         2.9          1.4         0.2
## 10           4.9         3.1          1.5         0.1
## 11           5.4         3.7          1.5         0.2
## 12           4.8         3.4          1.6         0.2
## 13           4.8         3.0          1.4         0.1
## 14           4.3         3.0          1.1         0.1
## 15           5.8         4.0          1.2         0.2
## 16           5.7         4.4          1.5         0.4
## 17           5.4         3.9          1.3         0.4
## 18           5.1         3.5          1.4         0.3
## 19           5.7         3.8          1.7         0.3
## 20           5.1         3.8          1.5         0.3
## 21           5.4         3.4          1.7         0.2
## 22           5.1         3.7          1.5         0.4
## 23           4.6         3.6          1.0         0.2
## 24           5.1         3.3          1.7         0.5
## 25           4.8         3.4          1.9         0.2
## 26           5.0         3.0          1.6         0.2
```

```
## 27          5.0         3.4          1.6         0.4
## 28          5.2         3.5          1.5         0.2
## 29          5.2         3.4          1.4         0.2
## 30          4.7         3.2          1.6         0.2
## 31          4.8         3.1          1.6         0.2
## 32          5.4         3.4          1.5         0.4
## 33          5.2         4.1          1.5         0.1
## 34          5.5         4.2          1.4         0.2
## 35          4.9         3.1          1.5         0.2
## 36          5.0         3.2          1.2         0.2
## 37          5.5         3.5          1.3         0.2
## 38          4.9         3.6          1.4         0.1
## 39          4.4         3.0          1.3         0.2
## 40          5.1         3.4          1.5         0.2
## 41          5.0         3.5          1.3         0.3
## 42          4.5         2.3          1.3         0.3
## 43          4.4         3.2          1.3         0.2
## 44          5.0         3.5          1.6         0.6
## 45          5.1         3.8          1.9         0.4
## 46          4.8         3.0          1.4         0.3
## 47          5.1         3.8          1.6         0.2
## 48          4.6         3.2          1.4         0.2
## 49          5.3         3.7          1.5         0.2
## 50          5.0         3.3          1.4         0.2
## 51          7.0         3.2          4.7         1.4
## 52          6.4         3.2          4.5         1.5
## 53          6.9         3.1          4.9         1.5
## 54          5.5         2.3          4.0         1.3
## 55          6.5         2.8          4.6         1.5
## 56          5.7         2.8          4.5         1.3
## 57          6.3         3.3          4.7         1.6
## 58          4.9         2.4          3.3         1.0
## 59          6.6         2.9          4.6         1.3
## 60          5.2         2.7          3.9         1.4
## 61          5.0         2.0          3.5         1.0
## 62          5.9         3.0          4.2         1.5
## 63          6.0         2.2          4.0         1.0
## 64          6.1         2.9          4.7         1.4
## 65          5.6         2.9          3.6         1.3
## 66          6.7         3.1          4.4         1.4
## 67          5.6         3.0          4.5         1.5
## 68          5.8         2.7          4.1         1.0
## 69          6.2         2.2          4.5         1.5
## 70          5.6         2.5          3.9         1.1
## 71          5.9         3.2          4.8         1.8
## 72          6.1         2.8          4.0         1.3
## 73          6.3         2.5          4.9         1.5
## 74          6.1         2.8          4.7         1.2
## 75          6.4         2.9          4.3         1.3
## 76          6.6         3.0          4.4         1.4
## 77          6.8         2.8          4.8         1.4
## 78          6.7         3.0          5.0         1.7
## 79          6.0         2.9          4.5         1.5
## 80          5.7         2.6          3.5         1.0
```

```
## 81           5.5         2.4          3.8         1.1
## 82           5.5         2.4          3.7         1.0
## 83           5.8         2.7          3.9         1.2
## 84           6.0         2.7          5.1         1.6
## 85           5.4         3.0          4.5         1.5
## 86           6.0         3.4          4.5         1.6
## 87           6.7         3.1          4.7         1.5
## 88           6.3         2.3          4.4         1.3
## 89           5.6         3.0          4.1         1.3
## 90           5.5         2.5          4.0         1.3
## 91           5.5         2.6          4.4         1.2
## 92           6.1         3.0          4.6         1.4
## 93           5.8         2.6          4.0         1.2
## 94           5.0         2.3          3.3         1.0
## 95           5.6         2.7          4.2         1.3
## 96           5.7         3.0          4.2         1.2
## 97           5.7         2.9          4.2         1.3
## 98           6.2         2.9          4.3         1.3
## 99           5.1         2.5          3.0         1.1
## 100          5.7         2.8          4.1         1.3
## 101          6.3         3.3          6.0         2.5
## 102          5.8         2.7          5.1         1.9
## 103          7.1         3.0          5.9         2.1
## 104          6.3         2.9          5.6         1.8
## 105          6.5         3.0          5.8         2.2
## 106          7.6         3.0          6.6         2.1
## 107          4.9         2.5          4.5         1.7
## 108          7.3         2.9          6.3         1.8
## 109          6.7         2.5          5.8         1.8
## 110          7.2         3.6          6.1         2.5
## 111          6.5         3.2          5.1         2.0
## 112          6.4         2.7          5.3         1.9
## 113          6.8         3.0          5.5         2.1
## 114          5.7         2.5          5.0         2.0
## 115          5.8         2.8          5.1         2.4
## 116          6.4         3.2          5.3         2.3
## 117          6.5         3.0          5.5         1.8
## 118          7.7         3.8          6.7         2.2
## 119          7.7         2.6          6.9         2.3
## 120          6.0         2.2          5.0         1.5
## 121          6.9         3.2          5.7         2.3
## 122          5.6         2.8          4.9         2.0
## 123          7.7         2.8          6.7         2.0
## 124          6.3         2.7          4.9         1.8
## 125          6.7         3.3          5.7         2.1
## 126          7.2         3.2          6.0         1.8
## 127          6.2         2.8          4.8         1.8
## 128          6.1         3.0          4.9         1.8
## 129          6.4         2.8          5.6         2.1
## 130          7.2         3.0          5.8         1.6
## 131          7.4         2.8          6.1         1.9
## 132          7.9         3.8          6.4         2.0
## 133          6.4         2.8          5.6         2.2
## 134          6.3         2.8          5.1         1.5
```

```
## 135          6.1         2.6          5.6         1.4
## 136          7.7         3.0          6.1         2.3
## 137          6.3         3.4          5.6         2.4
## 138          6.4         3.1          5.5         1.8
## 139          6.0         3.0          4.8         1.8
## 140          6.9         3.1          5.4         2.1
## 141          6.7         3.1          5.6         2.4
## 142          6.9         3.1          5.1         2.3
## 143          5.8         2.7          5.1         1.9
## 144          6.8         3.2          5.9         2.3
## 145          6.7         3.3          5.7         2.5
## 146          6.7         3.0          5.2         2.3
## 147          6.3         2.5          5.0         1.9
## 148          6.5         3.0          5.2         2.0
## 149          6.2         3.4          5.4         2.3
## 150          5.9         3.0          5.1         1.8
```

```
iris %>%
  select(where(~is.numeric(.)))
```

```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width
## 1            5.1         3.5          1.4         0.2
## 2            4.9         3.0          1.4         0.2
## 3            4.7         3.2          1.3         0.2
## 4            4.6         3.1          1.5         0.2
## 5            5.0         3.6          1.4         0.2
## 6            5.4         3.9          1.7         0.4
## 7            4.6         3.4          1.4         0.3
## 8            5.0         3.4          1.5         0.2
## 9            4.4         2.9          1.4         0.2
## 10           4.9         3.1          1.5         0.1
## 11           5.4         3.7          1.5         0.2
## 12           4.8         3.4          1.6         0.2
## 13           4.8         3.0          1.4         0.1
## 14           4.3         3.0          1.1         0.1
## 15           5.8         4.0          1.2         0.2
## 16           5.7         4.4          1.5         0.4
## 17           5.4         3.9          1.3         0.4
## 18           5.1         3.5          1.4         0.3
## 19           5.7         3.8          1.7         0.3
## 20           5.1         3.8          1.5         0.3
## 21           5.4         3.4          1.7         0.2
## 22           5.1         3.7          1.5         0.4
## 23           4.6         3.6          1.0         0.2
## 24           5.1         3.3          1.7         0.5
## 25           4.8         3.4          1.9         0.2
## 26           5.0         3.0          1.6         0.2
## 27           5.0         3.4          1.6         0.4
## 28           5.2         3.5          1.5         0.2
## 29           5.2         3.4          1.4         0.2
## 30           4.7         3.2          1.6         0.2
## 31           4.8         3.1          1.6         0.2
## 32           5.4         3.4          1.5         0.4
## 33           5.2         4.1          1.5         0.1
## 34           5.5         4.2          1.4         0.2
```

```
## 35          4.9         3.1          1.5         0.2
## 36          5.0         3.2          1.2         0.2
## 37          5.5         3.5          1.3         0.2
## 38          4.9         3.6          1.4         0.1
## 39          4.4         3.0          1.3         0.2
## 40          5.1         3.4          1.5         0.2
## 41          5.0         3.5          1.3         0.3
## 42          4.5         2.3          1.3         0.3
## 43          4.4         3.2          1.3         0.2
## 44          5.0         3.5          1.6         0.6
## 45          5.1         3.8          1.9         0.4
## 46          4.8         3.0          1.4         0.3
## 47          5.1         3.8          1.6         0.2
## 48          4.6         3.2          1.4         0.2
## 49          5.3         3.7          1.5         0.2
## 50          5.0         3.3          1.4         0.2
## 51          7.0         3.2          4.7         1.4
## 52          6.4         3.2          4.5         1.5
## 53          6.9         3.1          4.9         1.5
## 54          5.5         2.3          4.0         1.3
## 55          6.5         2.8          4.6         1.5
## 56          5.7         2.8          4.5         1.3
## 57          6.3         3.3          4.7         1.6
## 58          4.9         2.4          3.3         1.0
## 59          6.6         2.9          4.6         1.3
## 60          5.2         2.7          3.9         1.4
## 61          5.0         2.0          3.5         1.0
## 62          5.9         3.0          4.2         1.5
## 63          6.0         2.2          4.0         1.0
## 64          6.1         2.9          4.7         1.4
## 65          5.6         2.9          3.6         1.3
## 66          6.7         3.1          4.4         1.4
## 67          5.6         3.0          4.5         1.5
## 68          5.8         2.7          4.1         1.0
## 69          6.2         2.2          4.5         1.5
## 70          5.6         2.5          3.9         1.1
## 71          5.9         3.2          4.8         1.8
## 72          6.1         2.8          4.0         1.3
## 73          6.3         2.5          4.9         1.5
## 74          6.1         2.8          4.7         1.2
## 75          6.4         2.9          4.3         1.3
## 76          6.6         3.0          4.4         1.4
## 77          6.8         2.8          4.8         1.4
## 78          6.7         3.0          5.0         1.7
## 79          6.0         2.9          4.5         1.5
## 80          5.7         2.6          3.5         1.0
## 81          5.5         2.4          3.8         1.1
## 82          5.5         2.4          3.7         1.0
## 83          5.8         2.7          3.9         1.2
## 84          6.0         2.7          5.1         1.6
## 85          5.4         3.0          4.5         1.5
## 86          6.0         3.4          4.5         1.6
## 87          6.7         3.1          4.7         1.5
## 88          6.3         2.3          4.4         1.3
```

```
## 89          5.6         3.0          4.1         1.3
## 90          5.5         2.5          4.0         1.3
## 91          5.5         2.6          4.4         1.2
## 92          6.1         3.0          4.6         1.4
## 93          5.8         2.6          4.0         1.2
## 94          5.0         2.3          3.3         1.0
## 95          5.6         2.7          4.2         1.3
## 96          5.7         3.0          4.2         1.2
## 97          5.7         2.9          4.2         1.3
## 98          6.2         2.9          4.3         1.3
## 99          5.1         2.5          3.0         1.1
## 100         5.7         2.8          4.1         1.3
## 101         6.3         3.3          6.0         2.5
## 102         5.8         2.7          5.1         1.9
## 103         7.1         3.0          5.9         2.1
## 104         6.3         2.9          5.6         1.8
## 105         6.5         3.0          5.8         2.2
## 106         7.6         3.0          6.6         2.1
## 107         4.9         2.5          4.5         1.7
## 108         7.3         2.9          6.3         1.8
## 109         6.7         2.5          5.8         1.8
## 110         7.2         3.6          6.1         2.5
## 111         6.5         3.2          5.1         2.0
## 112         6.4         2.7          5.3         1.9
## 113         6.8         3.0          5.5         2.1
## 114         5.7         2.5          5.0         2.0
## 115         5.8         2.8          5.1         2.4
## 116         6.4         3.2          5.3         2.3
## 117         6.5         3.0          5.5         1.8
## 118         7.7         3.8          6.7         2.2
## 119         7.7         2.6          6.9         2.3
## 120         6.0         2.2          5.0         1.5
## 121         6.9         3.2          5.7         2.3
## 122         5.6         2.8          4.9         2.0
## 123         7.7         2.8          6.7         2.0
## 124         6.3         2.7          4.9         1.8
## 125         6.7         3.3          5.7         2.1
## 126         7.2         3.2          6.0         1.8
## 127         6.2         2.8          4.8         1.8
## 128         6.1         3.0          4.9         1.8
## 129         6.4         2.8          5.6         2.1
## 130         7.2         3.0          5.8         1.6
## 131         7.4         2.8          6.1         1.9
## 132         7.9         3.8          6.4         2.0
## 133         6.4         2.8          5.6         2.2
## 134         6.3         2.8          5.1         1.5
## 135         6.1         2.6          5.6         1.4
## 136         7.7         3.0          6.1         2.3
## 137         6.3         3.4          5.6         2.4
## 138         6.4         3.1          5.5         1.8
## 139         6.0         3.0          4.8         1.8
## 140         6.9         3.1          5.4         2.1
## 141         6.7         3.1          5.6         2.4
## 142         6.9         3.1          5.1         2.3
```

```
## 143          5.8         2.7          5.1         1.9
## 144          6.8         3.2          5.9         2.3
## 145          6.7         3.3          5.7         2.5
## 146          6.7         3.0          5.2         2.3
## 147          6.3         2.5          5.0         1.9
## 148          6.5         3.0          5.2         2.0
## 149          6.2         3.4          5.4         2.3
## 150          5.9         3.0          5.1         1.8
```

```
iris %>%
  select(where(~is.numeric(..1)))
```

```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width
## 1            5.1         3.5          1.4         0.2
## 2            4.9         3.0          1.4         0.2
## 3            4.7         3.2          1.3         0.2
## 4            4.6         3.1          1.5         0.2
## 5            5.0         3.6          1.4         0.2
## 6            5.4         3.9          1.7         0.4
## 7            4.6         3.4          1.4         0.3
## 8            5.0         3.4          1.5         0.2
## 9            4.4         2.9          1.4         0.2
## 10           4.9         3.1          1.5         0.1
## 11           5.4         3.7          1.5         0.2
## 12           4.8         3.4          1.6         0.2
## 13           4.8         3.0          1.4         0.1
## 14           4.3         3.0          1.1         0.1
## 15           5.8         4.0          1.2         0.2
## 16           5.7         4.4          1.5         0.4
## 17           5.4         3.9          1.3         0.4
## 18           5.1         3.5          1.4         0.3
## 19           5.7         3.8          1.7         0.3
## 20           5.1         3.8          1.5         0.3
## 21           5.4         3.4          1.7         0.2
## 22           5.1         3.7          1.5         0.4
## 23           4.6         3.6          1.0         0.2
## 24           5.1         3.3          1.7         0.5
## 25           4.8         3.4          1.9         0.2
## 26           5.0         3.0          1.6         0.2
## 27           5.0         3.4          1.6         0.4
## 28           5.2         3.5          1.5         0.2
## 29           5.2         3.4          1.4         0.2
## 30           4.7         3.2          1.6         0.2
## 31           4.8         3.1          1.6         0.2
## 32           5.4         3.4          1.5         0.4
## 33           5.2         4.1          1.5         0.1
## 34           5.5         4.2          1.4         0.2
## 35           4.9         3.1          1.5         0.2
## 36           5.0         3.2          1.2         0.2
## 37           5.5         3.5          1.3         0.2
## 38           4.9         3.6          1.4         0.1
## 39           4.4         3.0          1.3         0.2
## 40           5.1         3.4          1.5         0.2
## 41           5.0         3.5          1.3         0.3
## 42           4.5         2.3          1.3         0.3
```

```
## 43          4.4         3.2          1.3         0.2
## 44          5.0         3.5          1.6         0.6
## 45          5.1         3.8          1.9         0.4
## 46          4.8         3.0          1.4         0.3
## 47          5.1         3.8          1.6         0.2
## 48          4.6         3.2          1.4         0.2
## 49          5.3         3.7          1.5         0.2
## 50          5.0         3.3          1.4         0.2
## 51          7.0         3.2          4.7         1.4
## 52          6.4         3.2          4.5         1.5
## 53          6.9         3.1          4.9         1.5
## 54          5.5         2.3          4.0         1.3
## 55          6.5         2.8          4.6         1.5
## 56          5.7         2.8          4.5         1.3
## 57          6.3         3.3          4.7         1.6
## 58          4.9         2.4          3.3         1.0
## 59          6.6         2.9          4.6         1.3
## 60          5.2         2.7          3.9         1.4
## 61          5.0         2.0          3.5         1.0
## 62          5.9         3.0          4.2         1.5
## 63          6.0         2.2          4.0         1.0
## 64          6.1         2.9          4.7         1.4
## 65          5.6         2.9          3.6         1.3
## 66          6.7         3.1          4.4         1.4
## 67          5.6         3.0          4.5         1.5
## 68          5.8         2.7          4.1         1.0
## 69          6.2         2.2          4.5         1.5
## 70          5.6         2.5          3.9         1.1
## 71          5.9         3.2          4.8         1.8
## 72          6.1         2.8          4.0         1.3
## 73          6.3         2.5          4.9         1.5
## 74          6.1         2.8          4.7         1.2
## 75          6.4         2.9          4.3         1.3
## 76          6.6         3.0          4.4         1.4
## 77          6.8         2.8          4.8         1.4
## 78          6.7         3.0          5.0         1.7
## 79          6.0         2.9          4.5         1.5
## 80          5.7         2.6          3.5         1.0
## 81          5.5         2.4          3.8         1.1
## 82          5.5         2.4          3.7         1.0
## 83          5.8         2.7          3.9         1.2
## 84          6.0         2.7          5.1         1.6
## 85          5.4         3.0          4.5         1.5
## 86          6.0         3.4          4.5         1.6
## 87          6.7         3.1          4.7         1.5
## 88          6.3         2.3          4.4         1.3
## 89          5.6         3.0          4.1         1.3
## 90          5.5         2.5          4.0         1.3
## 91          5.5         2.6          4.4         1.2
## 92          6.1         3.0          4.6         1.4
## 93          5.8         2.6          4.0         1.2
## 94          5.0         2.3          3.3         1.0
## 95          5.6         2.7          4.2         1.3
## 96          5.7         3.0          4.2         1.2
```

```
## 97           5.7          2.9           4.2          1.3
## 98           6.2          2.9           4.3          1.3
## 99           5.1          2.5           3.0          1.1
## 100          5.7          2.8           4.1          1.3
## 101          6.3          3.3           6.0          2.5
## 102          5.8          2.7           5.1          1.9
## 103          7.1          3.0           5.9          2.1
## 104          6.3          2.9           5.6          1.8
## 105          6.5          3.0           5.8          2.2
## 106          7.6          3.0           6.6          2.1
## 107          4.9          2.5           4.5          1.7
## 108          7.3          2.9           6.3          1.8
## 109          6.7          2.5           5.8          1.8
## 110          7.2          3.6           6.1          2.5
## 111          6.5          3.2           5.1          2.0
## 112          6.4          2.7           5.3          1.9
## 113          6.8          3.0           5.5          2.1
## 114          5.7          2.5           5.0          2.0
## 115          5.8          2.8           5.1          2.4
## 116          6.4          3.2           5.3          2.3
## 117          6.5          3.0           5.5          1.8
## 118          7.7          3.8           6.7          2.2
## 119          7.7          2.6           6.9          2.3
## 120          6.0          2.2           5.0          1.5
## 121          6.9          3.2           5.7          2.3
## 122          5.6          2.8           4.9          2.0
## 123          7.7          2.8           6.7          2.0
## 124          6.3          2.7           4.9          1.8
## 125          6.7          3.3           5.7          2.1
## 126          7.2          3.2           6.0          1.8
## 127          6.2          2.8           4.8          1.8
## 128          6.1          3.0           4.9          1.8
## 129          6.4          2.8           5.6          2.1
## 130          7.2          3.0           5.8          1.6
## 131          7.4          2.8           6.1          1.9
## 132          7.9          3.8           6.4          2.0
## 133          6.4          2.8           5.6          2.2
## 134          6.3          2.8           5.1          1.5
## 135          6.1          2.6           5.6          1.4
## 136          7.7          3.0           6.1          2.3
## 137          6.3          3.4           5.6          2.4
## 138          6.4          3.1           5.5          1.8
## 139          6.0          3.0           4.8          1.8
## 140          6.9          3.1           5.4          2.1
## 141          6.7          3.1           5.6          2.4
## 142          6.9          3.1           5.1          2.3
## 143          5.8          2.7           5.1          1.9
## 144          6.8          3.2           5.9          2.3
## 145          6.7          3.3           5.7          2.5
## 146          6.7          3.0           5.2          2.3
## 147          6.3          2.5           5.0          1.9
## 148          6.5          3.0           5.2          2.0
## 149          6.2          3.4           5.4          2.3
## 150          5.9          3.0           5.1          1.8
```

Creating new columns or mutating existing ones

Creating new columns on the bases of the one present in our dataset is one of the most useful and common tasks of data carpentry. If you look to it in abstract, also mutating the content of an existing column fits in the previous reasoning: I'm creating a new column with a name which is identical to the old one …

In DT, new columns are created by using the `:=` operator in the second "placeholder" of the call

```
## Create a column with the ration between sepal lenght and sepal width
new_dt_col <- iris %>%
  data.table(.) %>%
  .[,myratio := Sepal.Length/Sepal.Width]


## Create multiple columns

new_dt_col <- iris %>%
  data.table(.) %>%
  .[,c("myratio","myratio1") := list(Sepal.Length/Sepal.Width, Petal.Length/Petal.Width)]
```

In TB there is specific function `mutate` which can be piped to create or manipulate the columns

```
## here the creation!
nef_tb_col <- iris %>%
  mutate(myratio = Sepal.Length/Sepal.Width)

## multiple columns can be created inside the same mutate call by using commas
```

As usual the DT syntax is more compact, the TB syntax is more easy to read. But DT is by far more efficient!

In tidyverse, the combination of selectors and mutate can be used to apply some sort of transformation to a bunch of columns, To do that, `mutate` have to be combined with `across`. The following example clearly shows the idea:

```
## Suppose I want to calculate the logarithm of all the numeric columns in the iris dataset ...

iris %>%
  mutate(across(where(~is.numeric(.x)),~log10(.x), .names = "log_{.col}"))
```

```
##     Sepal.Length Sepal.Width Petal.Length Petal.Width    Species
## 1            5.1         3.5          1.4         0.2     setosa
## 2            4.9         3.0          1.4         0.2     setosa
## 3            4.7         3.2          1.3         0.2     setosa
## 4            4.6         3.1          1.5         0.2     setosa
## 5            5.0         3.6          1.4         0.2     setosa
## 6            5.4         3.9          1.7         0.4     setosa
## 7            4.6         3.4          1.4         0.3     setosa
## 8            5.0         3.4          1.5         0.2     setosa
## 9            4.4         2.9          1.4         0.2     setosa
## 10           4.9         3.1          1.5         0.1     setosa
## 11           5.4         3.7          1.5         0.2     setosa
## 12           4.8         3.4          1.6         0.2     setosa
## 13           4.8         3.0          1.4         0.1     setosa
## 14           4.3         3.0          1.1         0.1     setosa
## 15           5.8         4.0          1.2         0.2     setosa
## 16           5.7         4.4          1.5         0.4     setosa
## 17           5.4         3.9          1.3         0.4     setosa
## 18           5.1         3.5          1.4         0.3     setosa
```

```
## 19          5.7         3.8          1.7         0.3     setosa
## 20          5.1         3.8          1.5         0.3     setosa
## 21          5.4         3.4          1.7         0.2     setosa
## 22          5.1         3.7          1.5         0.4     setosa
## 23          4.6         3.6          1.0         0.2     setosa
## 24          5.1         3.3          1.7         0.5     setosa
## 25          4.8         3.4          1.9         0.2     setosa
## 26          5.0         3.0          1.6         0.2     setosa
## 27          5.0         3.4          1.6         0.4     setosa
## 28          5.2         3.5          1.5         0.2     setosa
## 29          5.2         3.4          1.4         0.2     setosa
## 30          4.7         3.2          1.6         0.2     setosa
## 31          4.8         3.1          1.6         0.2     setosa
## 32          5.4         3.4          1.5         0.4     setosa
## 33          5.2         4.1          1.5         0.1     setosa
## 34          5.5         4.2          1.4         0.2     setosa
## 35          4.9         3.1          1.5         0.2     setosa
## 36          5.0         3.2          1.2         0.2     setosa
## 37          5.5         3.5          1.3         0.2     setosa
## 38          4.9         3.6          1.4         0.1     setosa
## 39          4.4         3.0          1.3         0.2     setosa
## 40          5.1         3.4          1.5         0.2     setosa
## 41          5.0         3.5          1.3         0.3     setosa
## 42          4.5         2.3          1.3         0.3     setosa
## 43          4.4         3.2          1.3         0.2     setosa
## 44          5.0         3.5          1.6         0.6     setosa
## 45          5.1         3.8          1.9         0.4     setosa
## 46          4.8         3.0          1.4         0.3     setosa
## 47          5.1         3.8          1.6         0.2     setosa
## 48          4.6         3.2          1.4         0.2     setosa
## 49          5.3         3.7          1.5         0.2     setosa
## 50          5.0         3.3          1.4         0.2     setosa
## 51          7.0         3.2          4.7         1.4 versicolor
## 52          6.4         3.2          4.5         1.5 versicolor
## 53          6.9         3.1          4.9         1.5 versicolor
## 54          5.5         2.3          4.0         1.3 versicolor
## 55          6.5         2.8          4.6         1.5 versicolor
## 56          5.7         2.8          4.5         1.3 versicolor
## 57          6.3         3.3          4.7         1.6 versicolor
## 58          4.9         2.4          3.3         1.0 versicolor
## 59          6.6         2.9          4.6         1.3 versicolor
## 60          5.2         2.7          3.9         1.4 versicolor
## 61          5.0         2.0          3.5         1.0 versicolor
## 62          5.9         3.0          4.2         1.5 versicolor
## 63          6.0         2.2          4.0         1.0 versicolor
## 64          6.1         2.9          4.7         1.4 versicolor
## 65          5.6         2.9          3.6         1.3 versicolor
## 66          6.7         3.1          4.4         1.4 versicolor
## 67          5.6         3.0          4.5         1.5 versicolor
## 68          5.8         2.7          4.1         1.0 versicolor
## 69          6.2         2.2          4.5         1.5 versicolor
## 70          5.6         2.5          3.9         1.1 versicolor
## 71          5.9         3.2          4.8         1.8 versicolor
## 72          6.1         2.8          4.0         1.3 versicolor
```

```
## 73          6.3         2.5          4.9         1.5 versicolor
## 74          6.1         2.8          4.7         1.2 versicolor
## 75          6.4         2.9          4.3         1.3 versicolor
## 76          6.6         3.0          4.4         1.4 versicolor
## 77          6.8         2.8          4.8         1.4 versicolor
## 78          6.7         3.0          5.0         1.7 versicolor
## 79          6.0         2.9          4.5         1.5 versicolor
## 80          5.7         2.6          3.5         1.0 versicolor
## 81          5.5         2.4          3.8         1.1 versicolor
## 82          5.5         2.4          3.7         1.0 versicolor
## 83          5.8         2.7          3.9         1.2 versicolor
## 84          6.0         2.7          5.1         1.6 versicolor
## 85          5.4         3.0          4.5         1.5 versicolor
## 86          6.0         3.4          4.5         1.6 versicolor
## 87          6.7         3.1          4.7         1.5 versicolor
## 88          6.3         2.3          4.4         1.3 versicolor
## 89          5.6         3.0          4.1         1.3 versicolor
## 90          5.5         2.5          4.0         1.3 versicolor
## 91          5.5         2.6          4.4         1.2 versicolor
## 92          6.1         3.0          4.6         1.4 versicolor
## 93          5.8         2.6          4.0         1.2 versicolor
## 94          5.0         2.3          3.3         1.0 versicolor
## 95          5.6         2.7          4.2         1.3 versicolor
## 96          5.7         3.0          4.2         1.2 versicolor
## 97          5.7         2.9          4.2         1.3 versicolor
## 98          6.2         2.9          4.3         1.3 versicolor
## 99          5.1         2.5          3.0         1.1 versicolor
## 100         5.7         2.8          4.1         1.3 versicolor
## 101         6.3         3.3          6.0         2.5  virginica
## 102         5.8         2.7          5.1         1.9  virginica
## 103         7.1         3.0          5.9         2.1  virginica
## 104         6.3         2.9          5.6         1.8  virginica
## 105         6.5         3.0          5.8         2.2  virginica
## 106         7.6         3.0          6.6         2.1  virginica
## 107         4.9         2.5          4.5         1.7  virginica
## 108         7.3         2.9          6.3         1.8  virginica
## 109         6.7         2.5          5.8         1.8  virginica
## 110         7.2         3.6          6.1         2.5  virginica
## 111         6.5         3.2          5.1         2.0  virginica
## 112         6.4         2.7          5.3         1.9  virginica
## 113         6.8         3.0          5.5         2.1  virginica
## 114         5.7         2.5          5.0         2.0  virginica
## 115         5.8         2.8          5.1         2.4  virginica
## 116         6.4         3.2          5.3         2.3  virginica
## 117         6.5         3.0          5.5         1.8  virginica
## 118         7.7         3.8          6.7         2.2  virginica
## 119         7.7         2.6          6.9         2.3  virginica
## 120         6.0         2.2          5.0         1.5  virginica
## 121         6.9         3.2          5.7         2.3  virginica
## 122         5.6         2.8          4.9         2.0  virginica
## 123         7.7         2.8          6.7         2.0  virginica
## 124         6.3         2.7          4.9         1.8  virginica
## 125         6.7         3.3          5.7         2.1  virginica
## 126         7.2         3.2          6.0         1.8  virginica
```

```
## 127          6.2         2.8          4.8         1.8  virginica
## 128          6.1         3.0          4.9         1.8  virginica
## 129          6.4         2.8          5.6         2.1  virginica
## 130          7.2         3.0          5.8         1.6  virginica
## 131          7.4         2.8          6.1         1.9  virginica
## 132          7.9         3.8          6.4         2.0  virginica
## 133          6.4         2.8          5.6         2.2  virginica
## 134          6.3         2.8          5.1         1.5  virginica
## 135          6.1         2.6          5.6         1.4  virginica
## 136          7.7         3.0          6.1         2.3  virginica
## 137          6.3         3.4          5.6         2.4  virginica
## 138          6.4         3.1          5.5         1.8  virginica
## 139          6.0         3.0          4.8         1.8  virginica
## 140          6.9         3.1          5.4         2.1  virginica
## 141          6.7         3.1          5.6         2.4  virginica
## 142          6.9         3.1          5.1         2.3  virginica
## 143          5.8         2.7          5.1         1.9  virginica
## 144          6.8         3.2          5.9         2.3  virginica
## 145          6.7         3.3          5.7         2.5  virginica
## 146          6.7         3.0          5.2         2.3  virginica
## 147          6.3         2.5          5.0         1.9  virginica
## 148          6.5         3.0          5.2         2.0  virginica
## 149          6.2         3.4          5.4         2.3  virginica
## 150          5.9         3.0          5.1         1.8  virginica
##     log_Sepal.Length log_Sepal.Width log_Petal.Length log_Petal.Width
## 1          0.7075702       0.5440680       0.14612804     -0.69897000
## 2          0.6901961       0.4771213       0.14612804     -0.69897000
## 3          0.6720979       0.5051500       0.11394335     -0.69897000
## 4          0.6627578       0.4913617       0.17609126     -0.69897000
## 5          0.6989700       0.5563025       0.14612804     -0.69897000
## 6          0.7323938       0.5910646       0.23044892     -0.39794001
## 7          0.6627578       0.5314789       0.14612804     -0.52287875
## 8          0.6989700       0.5314789       0.17609126     -0.69897000
## 9          0.6434527       0.4623980       0.14612804     -0.69897000
## 10         0.6901961       0.4913617       0.17609126     -1.00000000
## 11         0.7323938       0.5682017       0.17609126     -0.69897000
## 12         0.6812412       0.5314789       0.20411998     -0.69897000
## 13         0.6812412       0.4771213       0.14612804     -1.00000000
## 14         0.6334685       0.4771213       0.04139269     -1.00000000
## 15         0.7634280       0.6020600       0.07918125     -0.69897000
## 16         0.7558749       0.6434527       0.17609126     -0.39794001
## 17         0.7323938       0.5910646       0.11394335     -0.39794001
## 18         0.7075702       0.5440680       0.14612804     -0.52287875
## 19         0.7558749       0.5797836       0.23044892     -0.52287875
## 20         0.7075702       0.5797836       0.17609126     -0.52287875
## 21         0.7323938       0.5314789       0.23044892     -0.69897000
## 22         0.7075702       0.5682017       0.17609126     -0.39794001
## 23         0.6627578       0.5563025       0.00000000     -0.69897000
## 24         0.7075702       0.5185139       0.23044892     -0.30103000
## 25         0.6812412       0.5314789       0.27875360     -0.69897000
## 26         0.6989700       0.4771213       0.20411998     -0.69897000
## 27         0.6989700       0.5314789       0.20411998     -0.39794001
## 28         0.7160033       0.5440680       0.17609126     -0.69897000
## 29         0.7160033       0.5314789       0.14612804     -0.69897000
```

```
## 30         0.6720979     0.5051500     0.20411998    -0.69897000
## 31         0.6812412     0.4913617     0.20411998    -0.69897000
## 32         0.7323938     0.5314789     0.17609126    -0.39794001
## 33         0.7160033     0.6127839     0.17609126    -1.00000000
## 34         0.7403627     0.6232493     0.14612804    -0.69897000
## 35         0.6901961     0.4913617     0.17609126    -0.69897000
## 36         0.6989700     0.5051500     0.07918125    -0.69897000
## 37         0.7403627     0.5440680     0.11394335    -0.69897000
## 38         0.6901961     0.5563025     0.14612804    -1.00000000
## 39         0.6434527     0.4771213     0.11394335    -0.69897000
## 40         0.7075702     0.5314789     0.17609126    -0.69897000
## 41         0.6989700     0.5440680     0.11394335    -0.52287875
## 42         0.6532125     0.3617278     0.11394335    -0.52287875
## 43         0.6434527     0.5051500     0.11394335    -0.69897000
## 44         0.6989700     0.5440680     0.20411998    -0.22184875
## 45         0.7075702     0.5797836     0.27875360    -0.39794001
## 46         0.6812412     0.4771213     0.14612804    -0.52287875
## 47         0.7075702     0.5797836     0.20411998    -0.69897000
## 48         0.6627578     0.5051500     0.14612804    -0.69897000
## 49         0.7242759     0.5682017     0.17609126    -0.69897000
## 50         0.6989700     0.5185139     0.14612804    -0.69897000
## 51         0.8450980     0.5051500     0.67209786     0.14612804
## 52         0.8061800     0.5051500     0.65321251     0.17609126
## 53         0.8388491     0.4913617     0.69019608     0.17609126
## 54         0.7403627     0.3617278     0.60205999     0.11394335
## 55         0.8129134     0.4471580     0.66275783     0.17609126
## 56         0.7558749     0.4471580     0.65321251     0.11394335
## 57         0.7993405     0.5185139     0.67209786     0.20411998
## 58         0.6901961     0.3802112     0.51851394     0.00000000
## 59         0.8195439     0.4623980     0.66275783     0.11394335
## 60         0.7160033     0.4313638     0.59106461     0.14612804
## 61         0.6989700     0.3010300     0.54406804     0.00000000
## 62         0.7708520     0.4771213     0.62324929     0.17609126
## 63         0.7781513     0.3424227     0.60205999     0.00000000
## 64         0.7853298     0.4623980     0.67209786     0.14612804
## 65         0.7481880     0.4623980     0.55630250     0.11394335
## 66         0.8260748     0.4913617     0.64345268     0.14612804
## 67         0.7481880     0.4771213     0.65321251     0.17609126
## 68         0.7634280     0.4313638     0.61278386     0.00000000
## 69         0.7923917     0.3424227     0.65321251     0.17609126
## 70         0.7481880     0.3979400     0.59106461     0.04139269
## 71         0.7708520     0.5051500     0.68124124     0.25527251
## 72         0.7853298     0.4471580     0.60205999     0.11394335
## 73         0.7993405     0.3979400     0.69019608     0.17609126
## 74         0.7853298     0.4471580     0.67209786     0.07918125
## 75         0.8061800     0.4623980     0.63346846     0.11394335
## 76         0.8195439     0.4771213     0.64345268     0.14612804
## 77         0.8325089     0.4471580     0.68124124     0.14612804
## 78         0.8260748     0.4771213     0.69897000     0.23044892
## 79         0.7781513     0.4623980     0.65321251     0.17609126
## 80         0.7558749     0.4149733     0.54406804     0.00000000
## 81         0.7403627     0.3802112     0.57978360     0.04139269
## 82         0.7403627     0.3802112     0.56820172     0.00000000
## 83         0.7634280     0.4313638     0.59106461     0.07918125
```

```
## 84           0.7781513    0.4313638    0.70757018    0.20411998
## 85           0.7323938    0.4771213    0.65321251    0.17609126
## 86           0.7781513    0.5314789    0.65321251    0.20411998
## 87           0.8260748    0.4913617    0.67209786    0.17609126
## 88           0.7993405    0.3617278    0.64345268    0.11394335
## 89           0.7481880    0.4771213    0.61278386    0.11394335
## 90           0.7403627    0.3979400    0.60205999    0.11394335
## 91           0.7403627    0.4149733    0.64345268    0.07918125
## 92           0.7853298    0.4771213    0.66275783    0.14612804
## 93           0.7634280    0.4149733    0.60205999    0.07918125
## 94           0.6989700    0.3617278    0.51851394    0.00000000
## 95           0.7481880    0.4313638    0.62324929    0.11394335
## 96           0.7558749    0.4771213    0.62324929    0.07918125
## 97           0.7558749    0.4623980    0.62324929    0.11394335
## 98           0.7923917    0.4623980    0.63346846    0.11394335
## 99           0.7075702    0.3979400    0.47712125    0.04139269
## 100          0.7558749    0.4471580    0.61278386    0.11394335
## 101          0.7993405    0.5185139    0.77815125    0.39794001
## 102          0.7634280    0.4313638    0.70757018    0.27875360
## 103          0.8512583    0.4771213    0.77085201    0.32221929
## 104          0.7993405    0.4623980    0.74818803    0.25527251
## 105          0.8129134    0.4771213    0.76342799    0.34242268
## 106          0.8808136    0.4771213    0.81954394    0.32221929
## 107          0.6901961    0.3979400    0.65321251    0.23044892
## 108          0.8633229    0.4623980    0.79934055    0.25527251
## 109          0.8260748    0.3979400    0.76342799    0.25527251
## 110          0.8573325    0.5563025    0.78532984    0.39794001
## 111          0.8129134    0.5051500    0.70757018    0.30103000
## 112          0.8061800    0.4313638    0.72427587    0.27875360
## 113          0.8325089    0.4771213    0.74036269    0.32221929
## 114          0.7558749    0.3979400    0.69897000    0.30103000
## 115          0.7634280    0.4471580    0.70757018    0.38021124
## 116          0.8061800    0.5051500    0.72427587    0.36172784
## 117          0.8129134    0.4771213    0.74036269    0.25527251
## 118          0.8864907    0.5797836    0.82607480    0.34242268
## 119          0.8864907    0.4149733    0.83884909    0.36172784
## 120          0.7781513    0.3424227    0.69897000    0.17609126
## 121          0.8388491    0.5051500    0.75587486    0.36172784
## 122          0.7481880    0.4471580    0.69019608    0.30103000
## 123          0.8864907    0.4471580    0.82607480    0.30103000
## 124          0.7993405    0.4313638    0.69019608    0.25527251
## 125          0.8260748    0.5185139    0.75587486    0.32221929
## 126          0.8573325    0.5051500    0.77815125    0.25527251
## 127          0.7923917    0.4471580    0.68124124    0.25527251
## 128          0.7853298    0.4771213    0.69019608    0.25527251
## 129          0.8061800    0.4471580    0.74818803    0.32221929
## 130          0.8573325    0.4771213    0.76342799    0.20411998
## 131          0.8692317    0.4471580    0.78532984    0.27875360
## 132          0.8976271    0.5797836    0.80617997    0.30103000
## 133          0.8061800    0.4471580    0.74818803    0.34242268
## 134          0.7993405    0.4471580    0.70757018    0.17609126
## 135          0.7853298    0.4149733    0.74818803    0.14612804
## 136          0.8864907    0.4771213    0.78532984    0.36172784
## 137          0.7993405    0.5314789    0.74818803    0.38021124
```

```
## 138       0.8061800     0.4913617     0.74036269     0.25527251
## 139       0.7781513     0.4771213     0.68124124     0.25527251
## 140       0.8388491     0.4913617     0.73239376     0.32221929
## 141       0.8260748     0.4913617     0.74818803     0.38021124
## 142       0.8388491     0.4913617     0.70757018     0.36172784
## 143       0.7634280     0.4313638     0.70757018     0.27875360
## 144       0.8325089     0.5051500     0.77085201     0.36172784
## 145       0.8260748     0.5185139     0.75587486     0.39794001
## 146       0.8260748     0.4771213     0.71600334     0.36172784
## 147       0.7993405     0.3979400     0.69897000     0.27875360
## 148       0.8129134     0.4771213     0.71600334     0.30103000
## 149       0.7923917     0.5314789     0.73239376     0.36172784
## 150       0.7708520     0.4771213     0.70757018     0.25527251
```

```
## Here:
## where is used to select the columns which are numeric
## across is used to mutate on all these columns
## and there is a beautiful ~ and . deluge ;-)
## the .names argument allows you to specify a set of new names ... {.col} refers to the old names ...
```

Perform operations on subgroups of samples (lines)

The last type of operations I want to touch on this flyby, are the one meant to calculate some quantity from groups of samples (rows). This is normally handy when you want to calculate summary statistics over a large table of samples.

In DT this operation is performed combining what we heve done before with the `by` argument

```
## Calculate the average of sepal length on the three species

## Summarising the output as a data.table
iris_mean_dt <- iris %>%
  data.table(.) %>%
  .[,list(myavg = mean(Sepal.Length)), by = Species]


## Creating a new column with the separate averages "recycled". I.e the columns of averages is of full
iris_mean_newcol <- iris %>%
  data.table(.) %>%
  .[,myavg := mean(Sepal.Length), by = Species]
```

in the case of TB, "by group" operations are performed by using the `group_by` function, often combined with `summarize`

```
## this does what we have just done ...
iris %>%
  group_by(Species) %>%
  summarise(myratio = mean(Sepal.Width), sd = sd(Sepal.Width))
```

```
## # A tibble: 3 x 3
##   Species    myratio    sd
##   <fct>        <dbl> <dbl>
## 1 setosa        3.43 0.379
## 2 versicolor    2.77 0.314
## 3 virginica     2.97 0.322
```

```
## note that I have here two summary functions
```

Group_by can also be combined with `mutate()` to mirror the "recycling" behavior of dt

```
## here, for example, I'm adding a column with the number of samples for each group.

iris %>%
  group_by(Species) %>%
  mutate(nsamples = length(Species))
```

```
## # A tibble: 150 x 6
## # Groups:   Species [3]
##    Sepal.Length Sepal.Width Petal.Length Petal.Width Species nsamples
##           <dbl>       <dbl>        <dbl>       <dbl> <fct>      <int>
##  1          5.1         3.5          1.4         0.2 setosa        50
##  2          4.9         3            1.4         0.2 setosa        50
##  3          4.7         3.2          1.3         0.2 setosa        50
##  4          4.6         3.1          1.5         0.2 setosa        50
##  5          5           3.6          1.4         0.2 setosa        50
##  6          5.4         3.9          1.7         0.4 setosa        50
##  7          4.6         3.4          1.4         0.3 setosa        50
##  8          5           3.4          1.5         0.2 setosa        50
##  9          4.4         2.9          1.4         0.2 setosa        50
## 10          4.9         3.1          1.5         0.1 setosa        50
## # i 140 more rows
```