author: Flavio Lichtenstein
data update: 2019/01/26 data created: 2016/07/20
local: Molecular Biology/Inovation & CENTD - Butantan Institute

Lecture 01 - dir and files in R

where I am - directory - folder

use getwd() and setwd()

In [1]:

```
getwd()
```

'/media/flalix/disk2t/udemy/R-lectures-biostatistics'

In [2]:

```
my.dir = '/media/flalix/disk2t/udemy/'
setwd(my.dir)
getwd()
```

'/media/flalix/disk2t/udemy'

In [3]:

```
my.dir = '/media/flalix/disk2t/udemy/R-lectures-biostatistics'
setwd(my.dir)
getwd()
```

'/media/flalix/disk2t/udemy/R-lectures-biostatistics'

What is inside this folder?

use dir(), and best list.files()

In [4]:

```
dir()
```

'data' 'handson_01_dir_and_file.ipynb' 'pdf' 'R' 'Untitled.ipynb'

In [8]:

```
list.files()
```

'data' 'handson_01_dir_and_file.ipynb' 'pdf' 'R' 'Untitled.ipynb'

In [14]:

```
list.files(path="data")
```

'femaleControlsPopulation.csv'

In [16]:

```
list.files(path="R", pattern="plot")
```

'ex03_dataframe - ggplot.R' 'ex04_dataframe - bar error plot.R'
'ex05_dataframe - bar error plot.R'

Does a file exists?

In [17]:

```
file.exists("ex04_dataframe - bar error plot.R")
```

FALSE

In [18]:

```
file.exists("R/ex04_dataframe - bar error plot.R")
```

TRUE

Does a folder (subdir) exists?

In [19]:

```
dir.exists("R")
```

TRUE

otherwise, create it!

In [25]:

```
dir.create("cursor")
dir()
```

```
Warning message in dir.create("cursor"):
“'cursor' already exists”
```

'cursor' 'data' 'handson_01_dir_and_file.ipynb' 'pdf' 'R' 'Untitled.ipynb'

In [27]:

```
suppressWarnings(dir.create("cursor"))
dir()
```

'cursor' 'data' 'handson_01_dir_and_file.ipynb' 'pdf' 'R' 'Untitled.ipynb'

R & Linux environment

Lets look which variables, tables and functions are active in R (global) environment

In [49]:

```
ls()
```

'my.dir'

In [50]:

```
a = 3
print(a)
print(a**2)
print(a**3-2)
```

```
[1] 3
[1] 9
[1] 25
```

In [56]:

```
ls()
```

'a' 'my.dir'

You can clean all your environment

In [59]:

```
rm(list=ls())
ls()
```

Where is R and its libraries installed?

In [60]:

```
.libPaths()
```

'/home/flalix/anaconda3/envs/study/lib/R/library'

Is very usefull to know versions and parameteres when an error occurs!

In [61]:

```
sessionInfo()
```

```
R version 3.5.1 (2018-07-02)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Linux Mint 19

Matrix products: default
BLAS: /home/flalix/anaconda3/envs/study/lib/R/lib/libRblas.so
LAPACK: /home/flalix/anaconda3/envs/study/lib/R/lib/libRlapack.so

locale:
 [1] LC_CTYPE=en_US.UTF-8          LC_NUMERIC=pt_BR.UTF-8
 [3] LC_TIME=en_US.UTF-8           LC_COLLATE=en_US.UTF-8
 [5] LC_MONETARY=pt_BR.UTF-8       LC_MESSAGES=en_US.UTF-8
 [7] LC_PAPER=pt_BR.UTF-8          LC_NAME=pt_BR.UTF-8
 [9] LC_ADDRESS=pt_BR.UTF-8        LC_TELEPHONE=pt_BR.UTF-8
[11] LC_MEASUREMENT=pt_BR.UTF-8    LC_IDENTIFICATION=pt_BR.UTF-8

attached base packages:
[1] stats     graphics  grDevices utils     datasets  methods   base


other attached packages:
[1] RevoUtils_11.0.1     RevoUtilsMath_11.0.0

loaded via a namespace (and not attached):
 [1] Rcpp_0.12.18    digest_0.6.15   crayon_1.3.4    IRdisplay_0.5.0
 [5] repr_0.15.0     jsonlite_1.5    magrittr_1.5    evaluate_0.11
 [9] stringi_1.2.4   uuid_0.1-2      IRkernel_0.8.11 tools_3.5.1
[13] stringr_1.3.1   compiler_3.5.1  base64enc_0.1-3 pbdZMQ_0.3-3
[17] htmltools_0.3.6
```

I can call external programs, but I must know the external (Linux) environment PATH

In [62]:

```
Sys.getenv("PATH")
```

'/home/flalix/anaconda3/envs/study/bin:/home/flalix/anaconda3/envs/study/bin:/home/flalix/a

and even, I can change the PATH

In [64]:

```
Sys.setenv("PATH" = paste(Sys.getenv("PATH"), "~/Downloads", sep=":"))
```

In [65]:

```
Sys.getenv("PATH")
```

'/home/flalix/anaconda3/envs/study/bin:/home/flalix/anaconda3/envs/study/bin:/home/flalix/a

Markdown language

markdown (https://help.github.com/articles/basic-writing-and-formatting-syntax/#links)

or

another site (https://guides.github.com/features/mastering-markdown/)

In [48]:

```
### internal table: mtcars
```

In [29]:

```
?mtcars
```

In [30]:

```
mtcars
```

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160.0	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160.0	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108.0	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258.0	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360.0	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225.0	105	2.76	3.460	20.22	1	0	3	1
Duster 360	14.3	8	360.0	245	3.21	3.570	15.84	0	0	3	4
Merc 240D	24.4	4	146.7	62	3.69	3.190	20.00	1	0	4	2
Merc 230	22.8	4	140.8	95	3.92	3.150	22.90	1	0	4	2
Merc 280	19.2	6	167.6	123	3.92	3.440	18.30	1	0	4	4
Merc 280C	17.8	6	167.6	123	3.92	3.440	18.90	1	0	4	4
Merc 450SE	16.4	8	275.8	180	3.07	4.070	17.40	0	0	3	3
Merc 450SL	17.3	8	275.8	180	3.07	3.730	17.60	0	0	3	3
Merc 450SLC	15.2	8	275.8	180	3.07	3.780	18.00	0	0	3	3
Cadillac Fleetwood	10.4	8	472.0	205	2.93	5.250	17.98	0	0	3	4
Lincoln Continental	10.4	8	460.0	215	3.00	5.424	17.82	0	0	3	4
Chrysler Imperial	14.7	8	440.0	230	3.23	5.345	17.42	0	0	3	4
Fiat 128	32.4	4	78.7	66	4.08	2.200	19.47	1	1	4	1
Honda Civic	30.4	4	75.7	52	4.93	1.615	18.52	1	1	4	2
Toyota Corolla	33.9	4	71.1	65	4.22	1.835	19.90	1	1	4	1
Toyota Corona	21.5	4	120.1	97	3.70	2.465	20.01	1	0	3	1
Dodge Challenger	15.5	8	318.0	150	2.76	3.520	16.87	0	0	3	2
AMC Javelin	15.2	8	304.0	150	3.15	3.435	17.30	0	0	3	2
Camaro Z28	13.3	8	350.0	245	3.73	3.840	15.41	0	0	3	4
Pontiac Firebird	19.2	8	400.0	175	3.08	3.845	17.05	0	0	3	2
Fiat X1-9	27.3	4	79.0	66	4.08	1.935	18.90	1	1	4	1
Porsche 914-2	26.0	4	120.3	91	4.43	2.140	16.70	0	1	5	2
Lotus Europa	30.4	4	95.1	113	3.77	1.513	16.90	1	1	5	2
Ford Pantera L	15.8	8	351.0	264	4.22	3.170	14.50	0	1	5	4
Ferrari Dino	19.7	6	145.0	175	3.62	2.770	15.50	0	1	5	6
Maserati Bora	15.0	8	301.0	335	3.54	3.570	14.60	0	1	5	8
Volvo 142E	21.4	4	121.0	109	4.11	2.780	18.60	1	1	4	2

In [31]:

```
head(mtcars, 6)
```

	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
Mazda RX4	21.0	6	160	110	3.90	2.620	16.46	0	1	4	4
Mazda RX4 Wag	21.0	6	160	110	3.90	2.875	17.02	0	1	4	4
Datsun 710	22.8	4	108	93	3.85	2.320	18.61	1	1	4	1
Hornet 4 Drive	21.4	6	258	110	3.08	3.215	19.44	1	0	3	1
Hornet Sportabout	18.7	8	360	175	3.15	3.440	17.02	0	0	3	2
Valiant	18.1	6	225	105	2.76	3.460	20.22	1	0	3	1

If you could not find mtcars, use these commands:

- library(datasets)
- attach(mtcars)

Write a csv/tsv file

csv: comma separated values
tsv: tab separated values

?write.csv

In [34]:

```
write.csv(mtcars, "data/mtcars.csv")
```

In [36]:

```
list.files(path="data")
```

'femaleControlsPopulation.csv' 'mtcars.csv'

In [37]:

```
file.exists("mtcars.csv")
```

FALSE

In [38]:

```
file.exists("data/mtcars.csv")
```

TRUE

Amazing and important dataset: IRIS (Ronald Fischer & Edgar Anderson)

Iris dataset (https://en.wikipedia.org/wiki/Iris_flower_data_set)

In [39]:

```
head(iris)
```

Sepal.Length	Sepal.Width	Petal.Length	Petal.Width	Species
5.1	3.5	1.4	0.2	setosa
4.9	3.0	1.4	0.2	setosa
4.7	3.2	1.3	0.2	setosa
4.6	3.1	1.5	0.2	setosa
5.0	3.6	1.4	0.2	setosa
5.4	3.9	1.7	0.4	setosa

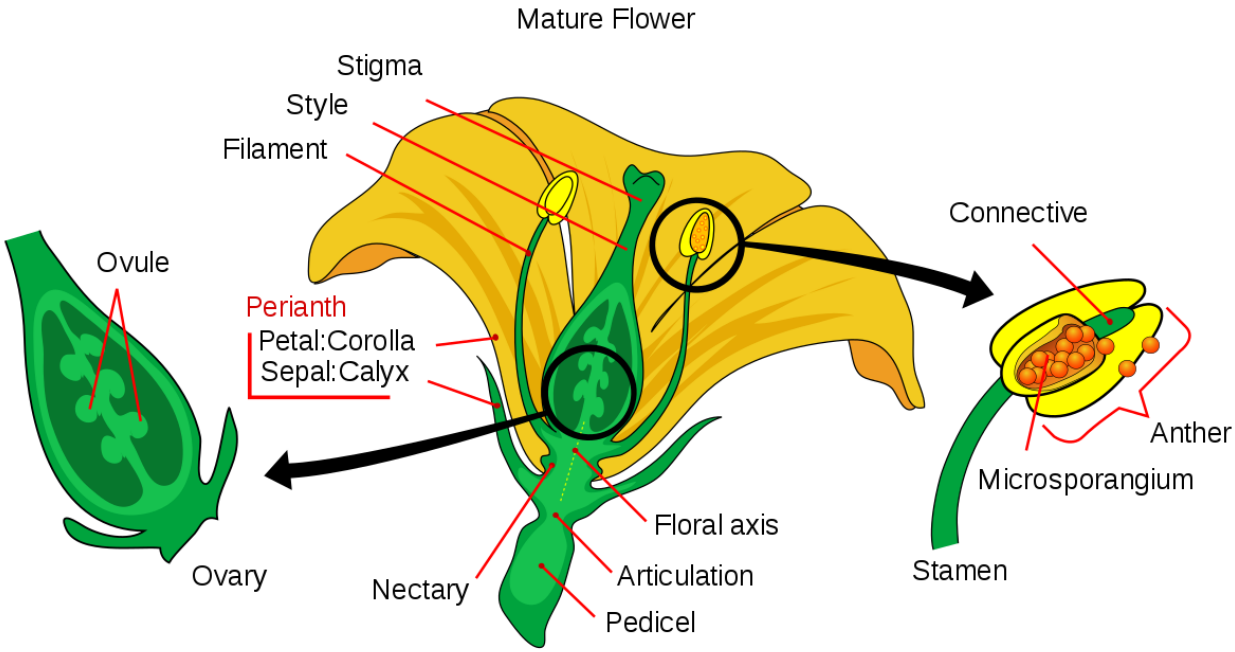
In [43]:

```
write.csv(iris, "data/iris.csv")
dir("data")
```

'femaleControlsPopulation.csv' 'iris.csv' 'mtcars.csv'

How to see an image?

What means sepal and petal?



lets go beyond,

How many species (classes)?

In [45]:

```
table(iris$Species)
```

```
    setosa versicolor  virginica 
        50         50         50 
```

In [ ]:

```
iris$<tab complete>
```

How is the correlation of length and width for each species?

In [46]:

```
pairs(iris)
```