

Task 4

2023-10-02

Using R example datasets

1. Describe briefly the content of the CO2 dataset using the help function

```
#load the data
data("CO2")
?CO2
```

The experiment is analyzing the cold tolerance of the grass species *Echinochola crus-galli* by measuring the CO2 response curve for photosynthesis. The data contain the CO2 uptake at several levels of CO2 concentration in the chilled and non-chilled plants.

2. What is the average and median CO2 uptake of the plants from Quebec and Mississippi?

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.3      v readr      2.1.4
## v forcats    1.0.0      v stringr   1.5.0
## v ggplot2    3.4.3      v tibble    3.2.1
## v lubridate  1.9.3      v tidyr     1.3.0
## v purrr      1.0.2
## -- 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(dplyr)
```

```
data(CO2)
#select the plant originated from quebec only
Quebec <- filter(CO2, Type=="Quebec")
summary(Quebec)
```

```
##      Plant      Type      Treatment      conc      uptake
## Qn1       :7  Quebec      :42 nonchilled:21  Min.   : 95  Min.   : 9.30
## Qn2       :7  Mississippi: 0   chilled  :21  1st Qu.: 175  1st Qu.:30.32
## Qn3       :7                                     Median : 350  Median :37.15
## Qc1       :7                                     Mean   : 435  Mean   :33.54
## Qc3       :7                                     3rd Qu.: 675  3rd Qu.:40.15
## Qc2       :7                                     Max.   :1000  Max.   :45.50
## (Other):0
```

```
#select the plant originated from Mississippi only
Mississippi<-filter(CO2, Type=="Mississippi")
summary(Mississippi)
```

```
##      Plant      Type      Treatment      conc      uptake
## Mn3      :7  Quebec      : 0  nonchilled:21  Min.    : 95  Min.    : 7.70
## Mn2      :7  Mississippi:42  chilled   :21  1st Qu.: 175  1st Qu.:13.88
## Mn1      :7                                     Median : 350  Median :19.30
## Mc2      :7                                     Mean   : 435  Mean   :20.88
## Mc3      :7                                     3rd Qu.: 675  3rd Qu.:28.05
## Mc1      :7                                     Max.   :1000  Max.   :35.50
## (Other):0
```

Another way is using tapply

```
tapply(CO2$uptake, CO2$Type, mean)
```

```
##      Quebec Mississippi
## 33.54286    20.88333
```

```
tapply(CO2$uptake, CO2$Type, median)
```

```
##      Quebec Mississippi
## 37.15      19.30
```

[OPTIONAL]In the airway example data from Bioconductor, how many genes are expressed in each sample? How many genes are not expressed in any sample?