R tidyverse

Testind tidyverse

reference website: https://www.tidyverse.org/index.html

reference e-book: https://r4ds.had.co.nz/

Tidyverse is an opinionated collection of R packages designed for data science. All packages share an underlying design philosophy, grammar, and data structures.

Install the complete tidyverse with:

```
#install.packages("tidyverse")
library(tidyverse)
```

Basic tidyverse packages include dyplr, tibble and ggplot2. Additional tidyverse packages that help with general programming challenges: purr (loops), and magrittr (pipe operator)

magrittr %>% pipe operator %>%

magrittr provides the pipe operator %>% used throughout the tidyverse. It also provide a number of more specialised piping operators (like %$% and %<>%) that can be useful in other places. basic meaning of piping:

- x %>% f is equivalent to f(x)
- x %>% f(y) is equivalent to f(x, y)
- x %>% f %>% g %>% h is equivalent to h(g(f(x)))

In the following example we calculate a new variable (addetti_tot), filter only the observations that have more than 500 addetti_tot and province in FVG, finally sort.

```
path = './../../_DATA/imprese'
#load a dataset
large_companies <- read.csv( paste0(path,"/t_imprese.csv"), sep = "|") %>%
                   transform(addetti_tot = addetti_indip + addetti_dip  ) %>%
                   subset(addetti_tot > 500) %>%
                   subset(prov %in% c("GO","TS","UD","PN")) %>%
                   arrange(-addetti_tot) %>%
                   select(c("denominazione", "addetti_tot"))
large_companies %>% head(5)
```

```
##                                                                denominazione
## 1 EURO&PROMOS FM SOCIETA  PER AZIONI  IN FORMA ABBREVIATA  EURO&PROMOS FM S.P.A.
## 2                                                          IDEALSERVICE SOC. COOP.
## 3         COOP NONCELLO   SOCIETA  COOPERATIVA SOCIALE   IMPRESA SOCIALE   ONLUS
```

```
## 4                                                          FINCANTIERI S.P.A.
## 5                                                         BOFROST  ITALIA SPA
##   addetti_tot
## 1      210829
## 2      160182
## 3       55776
## 4       36715
## 5       32058
```

```
# define classes and count the number of companies by class
classes <- large_companies %>%
            mutate(groups = case_when(    addetti_tot >= 2000 ~ 'XL',
                   addetti_tot >= 1000  & addetti_tot <= 1999 ~ 'L',
                   addetti_tot >= 500   & addetti_tot <= 999  ~ 'M')) %>%
           group_by(groups) %>%
           summarise(n())
classes
```

```
## # A tibble: 3 x 2
##   groups `n()`
##   <chr>  <int>
## 1 L         49
## 2 M        118
## 3 XL        54
```