

Лабораторна робота 3

Юрій Харченко

11/24/2020

Підключимо потрібні пакети

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
##
## The following objects are masked from 'package:stats':
##
##   filter, lag
##
## The following objects are masked from 'package:base':
##
##   intersect, setdiff, setequal, union
```

```
library(psych)
```

Завантажимо створений в попередній роботі фрейм

```
all <- read.csv(file = "all3.csv")
head(all)
```

```
##   X.1 X iso2c      country SP.POP.TOTL year iso3c      region
## 1  1 1 AR      Argentina  44938712 2019 ARG Latin America & Caribbean
## 2  2 2 BD      Bangladesh 163046161 2019 BGD South Asia
## 3  3 3 BR      Brazil    211049527 2019 BRA Latin America & Caribbean
## 4  4 4 CD Congo, Dem. Rep. 86790567 2019 COD Sub-Saharan Africa
## 5  5 5 CN      China    1397715000 2019 CHN East Asia & Pacific
## 6  6 6 CO      Colombia  50339443 2019 COL Latin America & Caribbean
##      capital longitude latitude      income lending
## 1 Buenos Aires -58.4173 -34.61180 Upper middle income IBRD
## 2      Dhaka    90.4113  23.70550 Lower middle income IDA
## 3    Brasilia -47.9292 -15.78010 Upper middle income IBRD
## 4    Kinshasa  15.3222  -4.32500 Low income IDA
## 5     Beijing 116.2860 40.04950 Upper middle income IBRD
## 6     Bogota  -74.0820  4.60987 Upper middle income IBRD
## NY.GDP.PCAP.PP.CD AG.LND.TOTL.K2 AG.LND.AGRI.K2 SP.RUR.TOTL.ZS
## 1      22947.140      2736690      1487000.0      8.009
## 2      4950.744      130170      91942.1      62.595
## 3     15258.851     8358140     2835460.0     13.176
## 4      1143.453     2267050     262000.0     54.954
## 5     16784.659     9388210     5277330.0     39.692
## 6     15643.731     1109500     446656.0     18.896
```

Функції описової статистики

середнє (mean)

```
# mean
data.frame(
  "Населення" = mean(all$SP.POP.TOTL),
  "Сільське населення відс" = mean(all$SP.RUR.TOTL.ZS, na.rm = TRUE),
  "ВВП на душу долл" = mean(all$NY.GDP.PCAP.PP.CD, na.rm = TRUE),
  "Площа кв.км" = mean(all$AG.LND.TOTL.K2, na.rm = TRUE),
  "Площа сг кв.км" = mean(all$AG.LND.AGRI.K2, na.rm = TRUE)
)

##   Населення  Сільське.населення.відс  ВВП.на.душу.долл  Площа.кв.км  Площа.сг.кв.км
## 1 176049552                40.11503        20129.08      2126901      781560.8
```

медіана (median)

```
# median
data.frame(
  "Населення" = median(all$SP.POP.TOTL),
  "Сільське населення відс" = median(all$SP.RUR.TOTL.ZS, na.rm = TRUE),
  "ВВП на душу долл" = median(all$NY.GDP.PCAP.PP.CD, na.rm = TRUE),
  "Площа кв.км" = median(all$AG.LND.TOTL.K2, na.rm = TRUE),
  "Площа сг кв.км" = median(all$AG.LND.AGRI.K2, na.rm = TRUE)
)

##   Населення  Сільське.населення.відс  ВВП.на.душу.долл  Площа.кв.км  Площа.сг.кв.км
## 1   83132799                33.144        13170.17      828340      368440
```

стандартне відхилення (sd)

```
# sd
data.frame(
  "Населення" = sd(all$SP.POP.TOTL),
  "Сільське населення відс" = sd(all$SP.RUR.TOTL.ZS, na.rm = TRUE),
  "ВВП на душу долл" = sd(all$NY.GDP.PCAP.PP.CD, na.rm = TRUE),
  "Площа кв.км" = sd(all$AG.LND.TOTL.K2, na.rm = TRUE),
  "Площа сг кв.км" = sd(all$AG.LND.AGRI.K2, na.rm = TRUE)
)

##   Населення  Сільське.населення.відс  ВВП.на.душу.долл  Площа.кв.км  Площа.сг.кв.км
## 1 309044393                22.01243        18036.86      3494435      1167531
```

дисперсія (var)

```
# var
data.frame(
  "Населення" = var(all$SP.POP.TOTL),
  "Сільське населення відс" = var(all$SP.RUR.TOTL.ZS, na.rm = TRUE),
  "ВВП на душу долл" = var(all$NY.GDP.PCAP.PP.CD, na.rm = TRUE),
  "Площа кв.км" = var(all$AG.LND.TOTL.K2, na.rm = TRUE),
  "Площа сг кв.км" = var(all$AG.LND.AGRI.K2, na.rm = TRUE)
)
```

```
##      Населення Сільське.населення.відс ВВП.на.душу.долл Площа.кв.км
## 1 9.550844e+16          484.547          325328369 1.221108e+13
##      Площа.сг.кв.км
## 1 1.363129e+12
```

абсолютне відхилення медіани (mad)

```
# mad
data.frame(
  "Населення" = mad(all$SP.POP.TOTL),
  "Сільське населення відс" = mad(all$SP.RUR.TOTL.ZS, na.rm = TRUE),
  "ВВП на душу долл" = mad(all$NY.GDP.PCAP.PP.CD, na.rm = TRUE),
  "Площа кв.км" = mad(all$AG.LND.TOTL.K2, na.rm = TRUE),
  "Площа сг кв.км" = mad(all$AG.LND.AGRI.K2, na.rm = TRUE)
)

##      Населення Сільське.населення.відс ВВП.на.душу.долл Площа.кв.км Площа.сг.кв.км
## 1 48619430          23.96475          12562.16    777208.6    357069.4
```

квартилі (quantile)

```
qu <- c(0.25, 0.5, 0.75)

# quantile
data.frame(

  "Населення" = quantile(all$SP.POP.TOTL, probs=qu),
  "Сільське населення відс" = quantile(all$SP.RUR.TOTL.ZS, probs=qu, na.rm = TRUE),
  "ВВП на душу долл" = quantile(all$NY.GDP.PCAP.PP.CD, probs=qu, na.rm = TRUE),
  "Площа кв.км" = quantile(all$AG.LND.TOTL.K2, probs=qu, na.rm = TRUE),
  "Площа сг кв.км" = quantile(all$AG.LND.AGRI.K2, probs=qu, na.rm = TRUE)
)

##      Населення Сільське.населення.відс ВВП.на.душу.долл Площа.кв.км
## 25% 53309696          19.4955          5350.072    398311
## 50% 83132799          33.1440          13170.165    828340
## 75% 135974532         62.8440          28854.821    1910855
##      Площа.сг.кв.км
## 25% 155360.0
## 50% 368440.0
## 75% 694930.8
```

розмах варіації показників (range)

```
# range
data.frame(
  "Населення" = range(all$SP.POP.TOTL),
  "Сільське населення відс" = range(all$SP.RUR.TOTL.ZS, na.rm = TRUE),
  "ВВП на душу долл" = range(all$NY.GDP.PCAP.PP.CD, na.rm = TRUE),
  "Площа кв.км" = range(all$AG.LND.TOTL.K2, na.rm = TRUE),
  "Площа сг кв.км" = range(all$AG.LND.AGRI.K2, na.rm = TRUE)
)

##      Населення Сільське.населення.відс ВВП.на.душу.долл Площа.кв.км Площа.сг.кв.км
```

```
## 1 42813238 8.009 1143.453 97489 17008
## 2 1397715000 78.775 65118.358 16376870 5277330
```

Функції psych

describe - описові статистики

```
# describe
```

```
all %>%
  select("Населення" = SP.POP.TOTL,
         "Сільське населення відс" = SP.RUR.TOTL.ZS,
         "ВВП на душу долл" = NY.GDP.PCAP.PP.CD,
         "Площа кв.км" = AG.LND.TOTL.K2,
         "Площа сг кв.км" = AG.LND.AGRI.K2) %>%
  describe(na.rm = TRUE)
```

```
##          vars  n      mean      sd      median
## Населення      1 35 176049552.20 309044392.58 83132799.00
## Сільське населення відс 2 35      40.12      22.01      33.14
## ВВП на душу долл 3 34      20129.08      18036.86      13170.17
## Площа кв.км 4 34      2126901.18      3494435.06      828340.00
## Площа сг кв.км 5 35      781560.79      1167531.24      368440.00
##          trimmed      mad      min      max
## Населення      101352626.34 48619429.61 42813238.00 1.397715e+09
## Сільське населення відс      39.57      23.96      8.01 7.878000e+01
## ВВП на душу долл      18144.85      12562.16      1143.45 6.511836e+04
## Площа кв.км      1320498.61      777208.57      97489.00 1.637687e+07
## Площа сг кв.км      520143.34      357069.38      17008.00 5.277330e+06
##          range skew kurtosis      se
## Населення      1.354902e+09 3.37      10.28 52238036.66
## Сільське населення відс 7.077000e+01 0.22      -1.50      3.72
## ВВП на душу долл      6.397491e+04 0.92      -0.46      3093.30
## Площа кв.км      1.627938e+07 2.60      6.49      599290.67
## Площа сг кв.км      5.260322e+06 2.41      5.45      197348.80
```

describeBy - описові статистики по регіонах

```
# describeBy
```

```
all %>%
  select("Населення" = SP.POP.TOTL,
         "Сільське населення відс" = SP.RUR.TOTL.ZS,
         "ВВП на душу долл" = NY.GDP.PCAP.PP.CD,
         "Площа кв.км" = AG.LND.TOTL.K2,
         "Площа сг кв.км" = AG.LND.AGRI.K2) %>%
  describeBy(all$region, na.rm = TRUE)
```

```
##
## Descriptive statistics by group
## group: East Asia & Pacific
##          vars  n      mean      sd      median
## Населення      1 8 271820540.00 460300151.24 102289360.50
## Сільське населення відс 2 8      43.16      20.87      46.66
```

```

## ВВП на душу долл      3 8      19698.30      15138.85      14543.21
## Площа кв.км           4 8      1679254.88      3159477.27      437725.00
## Площа сг кв.км        5 8      812991.00      1812072.13      126000.00
##
##              trimmed      mad      min      max
## Населення      271820540.00 59976892.09 51709098.00 1.397715e+09
## Сільське населення відс      43.16      17.55      8.30 6.915000e+01
## ВВП на душу долл      19698.30      8476.46      5355.27 4.323572e+04
## Площа кв.км      1679254.88      263094.78      97489.00 9.388210e+06
## Площа сг кв.км      812991.00      130757.91      17008.00 5.277330e+06
##
##              range skew kurtosis      se
## Населення      1.346006e+09 1.78      1.50 162740679.16
## Сільське населення відс 6.085000e+01 -0.42      -1.37      7.38
## ВВП на душу долл      3.788045e+04 0.71      -1.39      5352.39
## Площа кв.км      9.290721e+06 1.76      1.45      1117043.90
## Площа сг кв.км      5.260322e+06 1.82      1.62      640664.25
## -----
## group: Europe & Central Asia
##
##              vars n      mean      sd      median      trimmed
## Населення      1 8 74573696.62 31648286.13 66947146.00 74573696.62
## Сільське населення відс 2 8      23.41      4.97      23.50      23.41
## ВВП на душу долл      3 8      38875.74      14191.09      43205.41      38875.74
## Площа кв.км      4 8      2457292.62 5626975.41 523560.49 2457292.62
## Площа сг кв.км      5 8      499089.39      685803.34      274918.59      499089.39
##
##              mad      min      max      range skew
## Населення      24216878.84 44385155.00 144373535.00 99988380.00 1.14
## Сільське населення відс      6.13      16.35      30.53      14.18 0.09
## ВВП на душу долл      14141.62      13341.21      56052.42      42711.21 -0.50
## Площа кв.км      299204.23      241930.00      16376870.00 16134940.00 1.85
## Площа сг кв.км      160639.71      127170.00      2177218.12      2050048.12 1.77
##
##              kurtosis      se
## Населення      0.14 11189358.87
## Сільське населення відс -1.61      1.76
## ВВП на душу долл      -1.28      5017.31
## Площа кв.км      1.69 1989436.23
## Площа сг кв.км      1.51 242468.10
## -----
## group: Latin America & Caribbean
##
##              vars n      mean      sd      median      trimmed
## Населення      1 4 108475802.75 78108827.23 88957486.00 108475802.75
## Сільське населення відс 2 4      14.91      5.42      16.04      14.91
## ВВП на душу долл      3 4      18565.11      3744.96      18027.22      18565.11
## Площа кв.км      4 4      3537070.00 3281993.88 2340320.00 3537070.00
## Площа сг кв.км      5 4      1457869.00 1012848.24 1274680.00 1457869.00
##
##              mad      min      max      range skew
## Населення      61258672.44 44938712.00 211049527.00 166110815.00 0.33
## Сільське населення відс      4.73      8.01      19.56      11.55 -0.27
## ВВП на душу долл      3819.07      15258.85      22947.14      7688.29 0.14
## Площа кв.км      1206235.95 1109500.00 8358140.00 7248640.00 0.66
## Площа сг кв.км      771207.00      446656.02      2835460.00      2388803.98 0.37
##
##              kurtosis      se
## Населення      -2.07 39054413.62
## Сільське населення відс -2.13      2.71
## ВВП на душу долл      -2.27      1872.48
## Площа кв.км      -1.75 1640996.94

```

```

## Площа сг кв.км          -1.89    506424.12
## -----
## group: Middle East & North Africa
##               vars n          mean          sd          median          trimmed
## Населення          1 3 75451677.67 29386897.42 82913906.00 75451677.67
## Сільське населення відс 2 3          36.23          18.25          26.81          36.23
## ВВП на душу долл    3 2    12035.43          304.55    12035.43    12035.43
## Площа кв.км        4 3   1668650.00    694005.33  1628760.00  1668650.00
## Площа сг кв.км     5 3    303493.23    231639.05  413601.99    303493.23
##               mad          min          max          range skew
## Населення          25907199.99 43053054.00 100388073.00 57335019.00 -0.24
## Сільське населення відс 3.26          24.61          57.27          32.66 0.38
## ВВП на душу долл    319.27    11820.09    12250.78          430.69 0.00
## Площа кв.км        938945.41   995450.00   2381740.00  1386290.00 0.06
## Площа сг кв.км     68107.69    37337.70    459540.00   422202.30 -0.37
##               kurtosis          se
## Населення          -2.33 16966533.13
## Сільське населення відс -2.33          10.54
## ВВП на душу долл    -2.75          215.35
## Площа кв.км        -2.33   400684.16
## Площа сг кв.км     -2.33   133736.87
## -----
## group: North America
##               vars n          mean sd          median          trimmed mad
## Населення          1 1 328239523.00 NA 328239523.00 328239523.00 0
## Сільське населення відс 2 1          17.54 NA          17.54          17.54 0
## ВВП на душу долл    3 1    65118.36 NA    65118.36    65118.36 0
## Площа кв.км        4 1   9147420.00 NA   9147420.00   9147420.00 0
## Площа сг кв.км     5 1   4058625.00 NA   4058625.00   4058625.00 0
##               min          max range skew kurtosis se
## Населення          328239523.00 328239523.00 0 NA NA NA
## Сільське населення відс 17.54          17.54 0 NA NA NA
## ВВП на душу долл    65118.36    65118.36 0 NA NA NA
## Площа кв.км        9147420.00   9147420.00 0 NA NA NA
## Площа сг кв.км     4058625.00   4058625.00 0 NA NA NA
## -----
## group: South Asia
##               vars n          mean          sd          median
## Населення          1 3 582009744.33 679844113.94 216565318.00
## Сільське населення відс 2 3          63.74          1.57          63.09
## ВВП на душу долл    3 3    5623.28    1222.35    4950.74
## Площа кв.км        4 3   1291413.33   1491276.93   770880.00
## Площа сг кв.км     5 3    752530.70    915220.69   368440.00
##               trimmed          mad          min          max
## Населення          582009744.33 79347502.17 163046161.00 1.366418e+09
## Сільське населення відс 63.74          0.74          62.59 6.553000e+01
## ВВП на душу долл    5623.28          97.65    4884.88 7.034220e+03
## Площа кв.км        1291413.33   949916.65   130170.00 2.973190e+06
## Площа сг кв.км     752530.70   409935.79    91942.10 1.797210e+06
##               range skew kurtosis          se
## Населення          1.203372e+09 0.38    -2.33 392508182.19
## Сільське населення відс 2.930000e+00 0.34    -2.33          0.91
## ВВП на душу долл    2.149340e+03 0.38    -2.33          705.72
## Площа кв.км        2.843020e+06 0.31    -2.33    860989.14

```

```

## Площа сг кв.км      1.705268e+06 0.35      -2.33      528402.91
## -----
## group: Sub-Saharan Africa
##               vars n      mean      sd      median      trimmed
## Населення      1 8 82006679.2 53492785.21 58281866.50 82006679.2
## Сільське населення відс 2 8      61.8      15.37      65.28      61.8
## ВВП на душу долл 3 8      4434.6      3723.94      3446.60      4434.6
## Площа кв.км     4 7 1006624.3 645023.01 910770.00 1006624.3
## Площа сг кв.км  5 8 474976.5 281233.20 379545.00 474976.5
##               mad      min      max      range      skew
## Населення      21854191.91 42813238.00 200963599.00 158150361.00 1.26
## Сільське населення відс      15.33      33.14      78.78      45.63 -0.59
## ВВП на душу долл      1712.29      1143.45      12999.12      11855.67 1.38
## Площа кв.км     448219.63 200520.00 2267050.00 2066530.00 0.75
## Площа сг кв.км  261634.42 144150.00 968410.00 824260.00 0.48
##               kurtosis      se
## Населення      0.16 18912555.58
## Сільське населення відс -1.12      5.44
## ВВП на душу долл 0.61      1316.61
## Площа кв.км     -0.54 243795.78
## Площа сг кв.км  -1.40 99430.95

```