NS125 PCW Session 7

2022-09-26

Using packages and inspecting data

```
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(ggplot2)
library(readxl)
library(ellipse)
```

```
##
## Attaching package: 'ellipse'

## The following object is masked from 'package:graphics':
##
##     pairs
```

```
df <- read_excel("data.xlsx", skip = 1, col_names = TRUE, na="NIL")
head(df)
```

```
## # A tibble: 6 x 43
##     WP5 Country    ISO   Region Sub_r~1 Sampl~2 Sampl~3 Sampl~4 Popul~5  WPCIAS
##   <dbl> <chr>      <chr> <chr>  <chr>     <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
## 1    70 Afghanistan AFG   Asia   Wester~       0    1010    1010  1.37e7 6.43
## 2    56 Angola     AGO   Africa Centra~       0    1000    1000  9.55e6 0.563
## 3    87 Argentina  ARG   Latin~ Southe~    1000    1000    2000  2.96e7 9.07
## 4    88 Armenia    ARM   Europe Easter~    1000    1000    2000  2.36e6 0.00485
## 5    47 Australia  AUS   Europ~ Wester~       0    1005    1005  1.73e7 6.98
## 6    89 Austria    AUT   Europe Wester~       0    1001    1001  7.07e6 0.012
## # ... with 33 more variables: SPCIAS <dbl>, APCIAS <dbl>, HDI <dbl>,
## #   CO2emi <dbl>, GDPperUS <dbl>, glo_tot <dbl>, Efcon <dbl>, TotBioCap <dbl>,
## #   VA <dbl>, PS <dbl>, GovE <dbl>, RQ <dbl>, RL <dbl>, CC <dbl>, WGI <dbl>,
## #   Aware <dbl>, Unaware <dbl>, RF_aware <dbl>, AUC_aware <dbl>,
## #   CI_aware <dbl>, Top_aware <chr>, Top2_aware <chr>, Top3_aware <chr>,
## #   Ratio_aware <chr>, Serious <dbl>, Not_serious <dbl>, RF_serious <dbl>,
## #   AUC_serious <dbl>, CITree_serious <dbl>, Top1_serious <chr>, ...
```

```
unique(df$Region)
```

```
## [1] "Asia"                  "Africa"
```

```
## [3] "Latin America & Caribbean" "Europe"
## [5] "Europe (Oceania)"          "Europe (North America)"
```

```r
unique(df$Sub_region)
```

```
##  [1] "Western Asia"                "Central Africa"
##  [3] "Southern South America"      "Eastern Europe"
##  [5] "Western Europe (AustraliaNZ)" "Western Europe"
##  [7] "Coastal West Africa"         "Sahelian Africa"
##  [9] "Southern Asia"               "Central America"
## [11] "Andean South America"        "Northern South America"
## [13] "Southern Africa"             "Western Europe (North America)"
## [15] "Northeast Asia"              "East Africa"
## [17] "Caribbean Islands"           "North Africa"
## [19] "Southeast Asia"              "Middle East"
```

Factorizing columns

```r
factor_cols <- c("Top_aware", "Top2_aware", "Top3_aware", "Top1_serious", "Top2_serious", "Top3_serious
df[,factor_cols] <- lapply(df[,factor_cols], factor)
head(df)
```

```
## # A tibble: 6 x 43
##     WP5 Country     ISO   Region Sub_r~1 Sampl~2 Sampl~3 Sampl~4 Popul~5  WPCIAS
##   <dbl> <chr>       <chr> <fct>  <fct>     <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
## 1    70 Afghanistan AFG   Asia   Wester~       0    1010    1010  1.37e7 6.43
## 2    56 Angola      AGO   Africa Centra~       0    1000    1000  9.55e6 0.563
## 3    87 Argentina   ARG   Latin~ Southe~    1000    1000    2000  2.96e7 9.07
## 4    88 Armenia     ARM   Europe Easter~    1000    1000    2000  2.36e6 0.00485
## 5    47 Australia   AUS   Europ~ Wester~       0    1005    1005  1.73e7 6.98
## 6    89 Austria     AUT   Europe Wester~       0    1001    1001  7.07e6 0.012
## # ... with 33 more variables: SPCIAS <dbl>, APCIAS <dbl>, HDI <dbl>,
## #   CO2emi <dbl>, GDPperUS <dbl>, glo_tot <dbl>, Efcon <dbl>, TotBioCap <dbl>,
## #   VA <dbl>, PS <dbl>, GovE <dbl>, RQ <dbl>, RL <dbl>, CC <dbl>, WGI <dbl>,
## #   Aware <dbl>, Unaware <dbl>, RF_aware <dbl>, AUC_aware <dbl>,
## #   CI_aware <dbl>, Top_aware <fct>, Top2_aware <fct>, Top3_aware <fct>,
## #   Ratio_aware <chr>, Serious <dbl>, Not_serious <dbl>, RF_serious <dbl>,
## #   AUC_serious <dbl>, CITree_serious <dbl>, Top1_serious <fct>, ...
```

Corregram plot

```
##                 Population.2008 WPCIAS SPCIAS APCIAS    HDI CO2emi GDPperUS
## Population.2008           1.000  0.944  0.279  0.630  0.002  0.002   -0.053
## WPCIAS                    0.944  1.000  0.181  0.481 -0.030 -0.019   -0.069
## SPCIAS                    0.279  0.181  1.000  0.160  0.042 -0.004    0.016
## APCIAS                    0.630  0.481  0.160  1.000 -0.115 -0.076   -0.094
## HDI                       0.002 -0.030  0.042 -0.115  1.000  0.564    0.693
## CO2emi                    0.002 -0.019 -0.004 -0.076  0.564  1.000    0.660
## GDPperUS                 -0.053 -0.069  0.016 -0.094  0.693  0.660    1.000
## WGI                      -0.036 -0.051  0.036 -0.061  0.791  0.486    0.801
## Aware                    -0.043 -0.061  0.037 -0.167  0.850  0.454    0.661
## Serious                  -0.163 -0.203  0.029  0.048 -0.277 -0.255   -0.365
##                    WGI  Aware Serious
## Population.2008 -0.036 -0.043  -0.163
## WPCIAS          -0.051 -0.061  -0.203
## SPCIAS           0.036  0.037   0.029
```
```

```
## APCIAS          -0.061 -0.167   0.048
## HDI              0.791  0.850  -0.277
## CO2emi           0.486  0.454  -0.255
## GDPperUS         0.801  0.661  -0.365
## WGI              1.000  0.728  -0.312
## Aware            0.728  1.000  -0.329
## Serious         -0.312 -0.329   1.000
```

```
library(corrgram)
```

```
corrgram(R, order = NULL, lower.panel = panel.shade, upper.panel = NULL, text.panel = panel.txt,
    main = "Predictors of Climate Change Corrgram Plot")
```

Predictors of Climate Change Corrgram Plot