UN Regions

Evangeline Reynolds

12/30/2017

```r
if (!dir.exists("../raw_data")) {
  dir.create("../raw_data")
}

if (!file.exists("../raw_data/un_region_members_2017-12-30.html"))  {
  download.file("https://www.un.org/depts/DGACM/RegionalGroups.shtml",
                destfile=paste0("../raw_data/un_region_members_",
                                Sys.Date(), ".html"))
}
```

```r
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
```

```r
library(stringr)
library(tidyr)
library(readr)
country <- readLines("../raw_data/un_region_members_2017-12-30.html",
                     encoding="UTF-8")
```

```
## Warning in readLines("../raw_data/un_region_members_2017-12-30.html",
## encoding = "UTF-8"): incomplete final line found on '../raw_data/
## un_region_members_2017-12-30.html'
```

```r
df = data_frame(country) %>%
  filter(str_detect(country, "<li>|strong")) %>%
  filter(!str_detect(country, "a href")) %>%
  mutate(un_regional_group = ifelse(str_detect(country,"strong"), country, NA)) %>%
  fill(un_regional_group) %>%
  filter(un_regional_group != country) %>%
  mutate(country = str_replace_all(country, "\\<.+?\\>", "")) %>%
  mutate(country = str_replace_all(country, " ", "")) %>%
  mutate(un_regional_group =
           str_replace_all(un_regional_group, "\\<.+?\\>", "")) %>%
  mutate(un_regional_group = str_trim(un_regional_group))%>%
  mutate(country = str_trim(country))

df %>% filter(str_detect(country, "\\*"))
```

```
## # A tibble: 4 x 2
##                 country                   un_regional_group
```

```
##                              <chr>                                    <chr>
## 1                          Turkey*                      Asia-Pacific Group
## 2                   Israel* Western    European and Others Group (WEOG)
## 3                   Turkey* Western    European and Others Group (WEOG)
## 4 United States of America* Western    European and Others Group (WEOG)
```

```r
notes = data_frame(
  country = c("Kiribati", "Israel","Turkey", "United States of America"),
  un_region_notes = c(
"As of 2010, Kiribati (geographically in Oceania) is not a member of any regional group, despite other (
"In May 2000 Israel became a WEOG full member, on a temporary basis (subject to renewal), in WEOG's head
"Turkey, participates fully in both WEOG and Asian Group, but for electoral purposes is considered a men
"The United States of America is not a member of any regional group, but attends
meetings of the Western Europe and Other States Group (WEOG) as an observer and is
considered to be a member of that group for electoral purposes."))

df = df %>%
  mutate(country = str_replace(country, "\\*", "")) %>%
  filter(!(country == "Turkey" & un_regional_group == "Asia-Pacific Group")) %>%
  full_join(notes)
```

```
## Joining, by = "country"
```

```r
head(df)
```

```
## # A tibble: 6 x 3
##         country un_regional_group un_region_notes
##           <chr>             <chr>           <chr>
## 1     Algeria    African Group           <NA>
## 2      Angola    African Group           <NA>
## 3       Benin    African Group           <NA>
## 4    Botswana    African Group           <NA>
## 5 Burkina Faso    African Group           <NA>
## 6     Burundi    African Group           <NA>
```

```r
tail(df)
```

```
## # A tibble: 6 x 3
##                     country                       un_regional_group
##                       <chr>                                   <chr>
## 1             Spain Western    European and Others Group (WEOG)
## 2            Sweden Western    European and Others Group (WEOG)
## 3       Switzerland Western    European and Others Group (WEOG)
## 4            Turkey Western    European and Others Group (WEOG)
## 5    United Kingdom Western    European and Others Group (WEOG)
## 6 United States of America Western    European and Others Group (WEOG)
## # ... with 1 more variables: un_region_notes <chr>
```

Save product

```r
if (!dir.exists("../product_data")) {
  dir.create("../product_data")
}
```

```r
write_csv(df, "../product_data/un_regional_groups.csv")
```

session info

```r
sessionInfo()
```

```
## R version 3.4.1 (2017-06-30)
## Platform: x86_64-apple-darwin15.6.0 (64-bit)
## Running under: macOS Sierra 10.12.6
##
## Matrix products: default
## BLAS: /Library/Frameworks/R.framework/Versions/3.4/Resources/lib/libRblas.0.dylib
## LAPACK: /Library/Frameworks/R.framework/Versions/3.4/Resources/lib/libRlapack.dylib
##
## locale:
## [1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
##
## attached base packages:
## [1] stats     graphics  grDevices utils     datasets  methods   base
##
## other attached packages:
## [1] bindrcpp_0.2  readr_1.1.1   tidyr_0.6.3   stringr_1.2.0 dplyr_0.7.3
##
## loaded via a namespace (and not attached):
##  [1] Rcpp_0.12.12     digest_0.6.12    rprojroot_1.2    assertthat_0.2.0
##  [5] R6_2.2.2         backports_1.1.0  magrittr_1.5     evaluate_0.10.1
##  [9] rlang_0.1.2      stringi_1.1.5    rmarkdown_1.6    tools_3.4.1
## [13] glue_1.1.1       hms_0.3          yaml_2.1.14      compiler_3.4.1
## [17] pkgconfig_2.0.1  htmltools_0.3.6  bindr_0.1        knitr_1.16
## [21] tibble_1.3.4
```