WDI

Amwaj Almalki

## Plan

Health Services Research.

1. Choose data
   1. Pick a good outcome (Health outcome). A measure of health care quality
   2. Make sure the data set has good predictors of health.
   3. Identify the intervention.
   4. Identify moderators.
   5. Identify mediators.
   6. Identify confounders.
   7. Need a good mix is continuous and categorical and dichotomous.
2. Analyse the data
   1. ggplot (grammar of graphics)
   2. gtsummary (grammar of tables)

## Data Source

Will use the World Data Bank. Note that there is an app call WDI.

## Sample Analysis

Data downloaded from (<https://databank.worldbank.org/source/education-statistics-%5e-all-indicators#>)

Two series were downloaded:

* Mortality rate under 5 per 1000
* Percent of students enrolled in Health and Welfare programs in tertiary education who are female

First there is a search engine.

r <- WDI::WDIsearch("mortality")  
r %>% as\_tibble() %>% head

## # A tibble: 6 x 2  
## indicator name   
## <chr> <chr>   
## 1 5.51.01.03.mortal Child mortality   
## 2 HF.DYN.IMRT.IN Mortality rate, infant (per 1,000 live births)   
## 3 HF.DYN.IMRT.IN.Q1 Mortality rate, infant (per 1,000 live births): Q1 (lowest)  
## 4 HF.DYN.IMRT.IN.Q2 Mortality rate, infant (per 1,000 live births): Q2   
## 5 HF.DYN.IMRT.IN.Q3 Mortality rate, infant (per 1,000 live births): Q3   
## 6 HF.DYN.IMRT.IN.Q4 Mortality rate, infant (per 1,000 live births): Q4

Extract the data from WDI for the same indicators.

#h <- WDI::WDI(indicator = c("UIS.FEP.56.F900",'SH.DYN.MORT'))

store the data in the package

#usethis::use\_data(h,overwrite = TRUE)  
#usethis::use\_data(m,overwrite = TRUE)

h %>% head

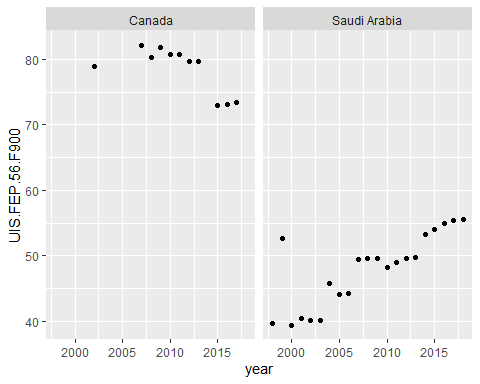
## iso2c country year UIS.FEP.56.F900 SH.DYN.MORT  
## 1 Global Partnership for Education 1970 NA NA  
## 2 Global Partnership for Education 1971 NA NA  
## 3 Global Partnership for Education 1972 NA NA  
## 4 Global Partnership for Education 1973 NA NA  
## 5 Global Partnership for Education 1974 NA NA  
## 6 Global Partnership for Education 1975 NA NA

h %>% count(country) %>%   
 arrange(-n) %>%   
 head

## country n  
## 1 Afghanistan 61  
## 2 Albania 61  
## 3 Algeria 61  
## 4 American Samoa 61  
## 5 Andorra 61  
## 6 Angola 61

Mortality rate under 5 per 1000

h %>%   
 na.omit() %>%   
 filter(country %in% c('Canada','Saudi Arabia')) %>%   
 ggplot(aes(x=year, y = UIS.FEP.56.F900))+  
 geom\_point() +  
 facet\_wrap(~country)



h %>%   
 filter(country %in% c('Canada','Saudi Arabia')) %>%   
 select(country,UIS.FEP.56.F900,SH.DYN.MORT) %>%   
 janitor::clean\_names() %>%   
gtsummary::tbl\_summary(h, by = 'country')

## Table printed with {flextable}, not {gt}. Learn why at  
## http://www.danieldsjoberg.com/gtsummary/articles/rmarkdown.html  
## To suppress this message, include `message = FALSE` in the code chunk header.

| Characteristic | Canada, N = 611 | Saudi Arabia, N = 611 |
| --- | --- | --- |
| Percentage of students enrolled in Health and Welfare programmes in tertiary education who are female (%) | 80 (76, 81) | 49 (44, 53) |
| Unknown | 50 | 40 |
| Mortality rate, under-5 (per 1,000 live births) | 8 (6, 17) | 29 (14, 72) |
| Unknown | 1 | 13 |
| 1Statistics presented: median (IQR) | | |