SDG6reproducibility

# Data import

# Data exploration

## Data exploration

Text

Rows: 16,576  
Columns: 9  
$ source <chr> "MICS03", "NRVA03", "NRVS05", "NRVA08", "MICS11", "N…  
$ type <chr> "Survey", "Survey", "Survey", "Survey", "Survey", "S…  
$ year <dbl> 2003, 2003, 2005, 2008, 2011, 2012, 2014, 2015, 2017…  
$ var\_short <chr> "s\_imp\_r", "s\_imp\_r", "s\_imp\_r", "s\_imp\_r", "s\_imp\_r…  
$ value <dbl> 26.00000, 28.55865, 31.17534, 30.10000, 44.20000, 36…  
$ iso3 <chr> "AFG", "AFG", "AFG", "AFG", "AFG", "AFG", "AFG", "AF…  
$ var\_long <chr> "Improved", "Improved", "Improved", "Improved", "Imp…  
$ residence <chr> "rural", "rural", "rural", "rural", "rural", "rural"…  
$ san\_service\_chain <fct> user interface, user interface, user interface, user…

# A tibble: 6 × 9  
 source type year var\_short value iso3 var\_long residence san\_service\_chain  
 <chr> <chr> <dbl> <chr> <dbl> <chr> <chr> <chr> <fct>   
1 MICS03 Survey 2003 s\_imp\_r 26 AFG Improved rural user interface   
2 NRVA03 Survey 2003 s\_imp\_r 28.6 AFG Improved rural user interface   
3 NRVS05 Survey 2005 s\_imp\_r 31.2 AFG Improved rural user interface   
4 NRVA08 Survey 2008 s\_imp\_r 30.1 AFG Improved rural user interface   
5 MICS11 Survey 2011 s\_imp\_r 44.2 AFG Improved rural user interface   
6 NRVA12 Survey 2012 s\_imp\_r 36.3 AFG Improved rural user interface

# A tibble: 6 × 9  
 source type year var\_short value iso3 var\_long residence san\_service\_cha…  
 <chr> <chr> <dbl> <chr> <dbl> <chr> <chr> <chr> <fct>   
1 WWT15 Admin 2015 s\_treat\_wt… 47.7 ZWE Treated… national WW treatment   
2 <NA> <NA> NA <NA> NA MAC <NA> <NA> <NA>   
3 <NA> <NA> NA <NA> NA VAT <NA> <NA> <NA>   
4 <NA> <NA> NA <NA> NA IMN <NA> <NA> <NA>   
5 <NA> <NA> NA <NA> NA SPM <NA> <NA> <NA>   
6 <NA> <NA> NA <NA> NA ESH <NA> <NA> <NA>

# Methods

## JMP Methods report

# First linear model

* Model for Uganda, rural, improved
* Data is not linear

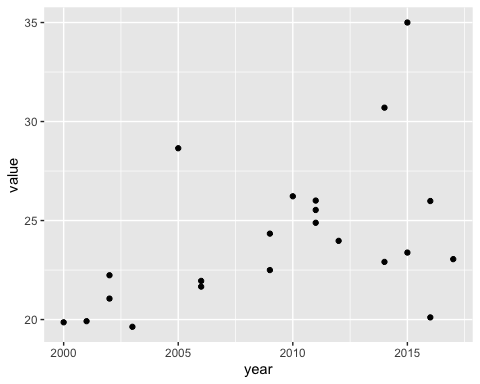
## Observations

Call:  
lm(formula = value ~ year, data = .)  
  
Residuals:  
 Min 1Q Median 3Q Max   
-6.1676 -1.4535 -0.7901 0.8256 9.0535   
  
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) -638.8793 274.7207 -2.326 0.0307 \*  
year 0.3299 0.1367 2.413 0.0255 \*  
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 3.395 on 20 degrees of freedom  
Multiple R-squared: 0.2255, Adjusted R-squared: 0.1868   
F-statistic: 5.823 on 1 and 20 DF, p-value: 0.02553

## Reference papers

* LINK

**Gaps identified**



# A tibble: 22 × 9  
 source type year var\_short value iso3 var\_long residence san\_service\_cha…  
 <chr> <chr> <dbl> <chr> <dbl> <chr> <chr> <chr> <fct>   
 1 NHS00 Survey 2000 s\_imp\_r 19.9 UGA Improved rural user interface   
 2 DHS01 Survey 2001 s\_imp\_r 19.9 UGA Improved rural user interface   
 3 CEN02 Census 2002 s\_imp\_r 21.1 UGA Improved rural user interface   
 4 NHS02 Survey 2002 s\_imp\_r 22.2 UGA Improved rural user interface   
 5 NHS03 Survey 2003 s\_imp\_r 19.6 UGA Improved rural user interface   
 6 DHS05 Survey 2005 s\_imp\_r 28.7 UGA Improved rural user interface   
 7 NHS06 Survey 2006 s\_imp\_r 21.7 UGA Improved rural user interface   
 8 DHS06 Survey 2006 s\_imp\_r 22.0 UGA Improved rural user interface   
 9 NHS09 Survey 2009 s\_imp\_r 22.5 UGA Improved rural user interface   
10 MIS09 Survey 2009 s\_imp\_r 24.3 UGA Improved rural user interface   
# … with 12 more rows