11_CL_Tiolet

Group K

2025-09-19

#Display Dataset content

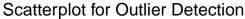
```
## # A tibble: 6 x 29
           DataId Indicator Value Precision DHS CountryCode CountryName SurveyYear
     ISO3
     <chr> <chr> <chr>
                             <chr> <chr>
                                                                          <chr>>
                                             <chr>>
                                                              <chr>
## 1 #coun~ #meta~ #indicat~ #ind~ #indicat~ <NA>
                                                             #country+n~ #date+year
## 2 ZAF
           795762 Househol~ 50.1 1
                                                             South Afri~ 1998
                                             ZA
## 3 ZAF
           795768 Househol~ 38.3 1
                                                             South Afri~ 1998
## 4 ZAF
           795760 Househol~ 31.2 1
                                             ZA
                                                             South Afri~ 1998
## 5 ZAF
            795764 Househol~ 6
                                             ZA
                                                             South Afri~ 1998
           795765 Househol~ 11.6 1
## 6 ZAF
                                                             South Afri~ 1998
                                             ZA
## # i 21 more variables: SurveyId <chr>, IndicatorId <chr>, IndicatorOrder <dbl>,
       IndicatorType <chr>, CharacteristicId <dbl>, CharacteristicOrder <dbl>,
## #
       CharacteristicCategory <chr>, CharacteristicLabel <chr>,
       ByVariableId <chr>, ByVariableLabel <chr>, IsTotal <dbl>,
## #
      IsPreferred <dbl>, SDRID <chr>, RegionId <lgl>, SurveyYearLabel <dbl>,
      SurveyType <chr>, DenominatorWeighted <dbl>, DenominatorUnweighted <dbl>,
## #
      CILow <lgl>, CIHigh <lgl>, LevelRank <lgl>
```

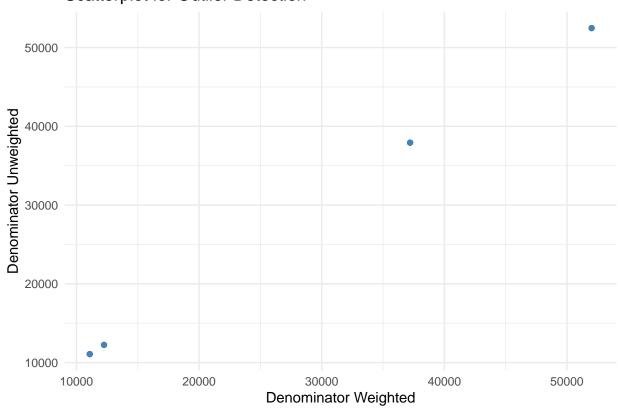
#check for unique values

```
## # A tibble: 29 x 3
##
      column
                      n_unique sample_values
##
      <chr>
                         <int> <chr>
##
  1 ISO3
                            1 ZAF
                            46 795762, 795768, 795760
## 2 DataId
## 3 Indicator
                            32 Households with an improved sanitation facility, Ho~
## 4 Value
                            37 50.1, 38.3, 31.2
## 5 Precision
                             2 1, 0
## 6 DHS CountryCode
                             1 ZA
## 7 CountryName
                             1 South Africa
## 8 SurveyYear
                             2 1998, 2016
                             2 ZA1998DHS, ZA2016DHS
## 9 SurveyId
## 10 IndicatorId
                            32 WS_TLET_H_IMP, WS_TLET_H_NIM, WS_TLET_H_NPT
## # i 19 more rows
```

Drop the countries only onw unquue value: reason, there is no useful information - county is also always za

Assumed pattern, the missing values can be filled with the previous non missing value in the opposite attribute





Boxplot of Denominator Weighted

