Class4_EDA

Quarto

Quarto enables you to weave together content and executable code into a finished document. To learn more about Quarto see https://quarto.org.

Running Code

When you click the **Render** button a document will be generated that includes both content and the output of embedded code. You can embed code like this:

```
1 + 1
```

```
[1] 2
```

You can add options to executable code like this

```
[1] 4
```

The `echo: false` option disables the printing of code (only output is displayed).

```
library(tidyverse)
```

```
-- Attaching core tidyverse packages ---------------------- tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.5
v forcats   1.0.0     v stringr   1.5.1
v ggplot2   3.4.4     v tibble    3.2.1
v lubridate 1.9.3     v tidyr     1.3.0
v purrr     1.0.2
-- Conflicts --------------------------------------- tidyverse_conflicts() --
```

```
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becom
```

```r
library(dplyr)
library(here)
```

```
here() starts at /Users/hamdaaltaf/Desktop/Version Control/maternal_armed_conflict
```

```r
here()
```

```
[1] "/Users/hamdaaltaf/Desktop/Version Control/maternal_armed_conflict"
```

```r
analyticaldata <- read.csv(here("analytical_data", "analyticaldata.csv"), header = TRUE)
```

```r
analyticaldata |>
  head()
```

```
  country_name ISO        region year    gdp1000 OECD OECD2023  popdens    urban
1  Afghanistan AFG Southern Asia 2000         NA    0        0 14.13654 16.25324
2  Afghanistan AFG Southern Asia 2001         NA    0        0 14.23156 16.25661
3  Afghanistan AFG Southern Asia 2002 0.1835328    0        0 14.32270 16.42654
4  Afghanistan AFG Southern Asia 2003 0.2004626    0        0 14.40691 16.60701
5  Afghanistan AFG Southern Asia 2004 0.2216576    0        0 15.21947 16.71367
6  Afghanistan AFG Southern Asia 2005 0.2550551    0        0 15.33619 16.85096
    agedep  male_edu     temp rainfall1000 totaldeath conflict Matmor Infantmor
1 108.3466 2.762086 12.69959    0.2763704          0       NA   1450      90.5
2 108.9899 2.856936 12.85570    0.2793079        832        1   1390      87.9
3 109.3472 2.954241 12.71081    0.3805710        441        1   1300      85.3
4 109.4475 3.054121 12.16592    0.4288939        200        1   1240      82.7
5 109.2868 3.156706 13.04643    0.3754336         60        1   1180      80.0
6 107.9646 3.262133 12.23141    0.4415680         42        1   1140      77.3
  Neonatmor Under5mor drought earthquake
1      60.9     129.2       1          0
2      59.7     125.2       0          1
3      58.5     121.1       0          1
4      57.2     116.9       0          1
5      55.9     112.6       0          1
6      54.6     108.4       0          1
```

```
analyticaldata |>
  glimpse()
```

```
Rows: 3,720
Columns: 21
$ country_name <chr> "Afghanistan", "Afghanistan", "Afghanistan", "Afghanistan~
$ ISO          <chr> "AFG", "AFG", "AFG", "AFG", "AFG", "AFG", "AFG", "AFG", "~
$ region       <chr> "Southern Asia", "Southern Asia", "Southern Asia", "South~
$ year         <int> 2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 200~
$ gdp1000      <dbl> NA, NA, 0.1835328, 0.2004626, 0.2216576, 0.2550551, 0.274~
$ OECD         <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~
$ OECD2023     <int> 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, ~
$ popdens      <dbl> 14.13654, 14.23156, 14.32270, 14.40691, 15.21947, 15.3361~
$ urban        <dbl> 16.25324, 16.25661, 16.42654, 16.60701, 16.71367, 16.8509~
$ agedep       <dbl> 108.34663, 108.98989, 109.34716, 109.44753, 109.28682, 10~
$ male_edu     <dbl> 2.762086, 2.856936, 2.954241, 3.054121, 3.156706, 3.26213~
$ temp         <dbl> 12.69959, 12.85570, 12.71081, 12.16592, 13.04643, 12.2314~
$ rainfall1000 <dbl> 0.2763704, 0.2793079, 0.3805710, 0.4288939, 0.3754336, 0.~
$ totaldeath   <int> 0, 832, 441, 200, 60, 42, 89, 87, 105, 189, 170, 157, 138~
$ conflict     <int> NA, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1,~
$ Matmor       <int> 1450, 1390, 1300, 1240, 1180, 1140, 1120, 1090, 1030, 993~
$ Infantmor    <dbl> 90.5, 87.9, 85.3, 82.7, 80.0, 77.3, 74.6, 71.9, 69.2, 66.~
$ Neonatmor    <dbl> 60.9, 59.7, 58.5, 57.2, 55.9, 54.6, 53.2, 51.7, 50.3, 48.~
$ Under5mor    <dbl> 129.2, 125.2, 121.1, 116.9, 112.6, 108.4, 104.1, 99.9, 95~
$ drought      <int> 1, 0, 0, 0, 0, 0, 1, 0, 1, 0, 0, 1, 0, 0, 0, 0, 0, 0, 1, ~
$ earthquake   <int> 0, 1, 1, 1, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1, 0, 1, 0, 0, 0, ~
```

```
analyticaldata |>
  slice_sample(n=6)
```

```
    country_name ISO                            region year    gdp1000 OECD
1   Turkmenistan TKM                      Central Asia 2012  6.4418866    0
2        Albania ALB                   Southern Europe 2016  4.1240554    0
3      Sri Lanka LKA                     Southern Asia 2013  3.7406168    0
4        Lebanon LBN                      Western Asia 2000  3.9948611    0
5         Brazil BRA Latin America and the Caribbean 2014 12.0711582    0
6          Niger NER                Sub-Saharan Africa 2001  0.2035265    0
  OECD2023  popdens    urban   agedep male_edu     temp rainfall1000 totaldeath
1        0 14.07009 11.18034 51.24511 9.740386 15.52978    0.1694034          0
2        0 46.10236 37.88628 46.55289 9.805501 14.21273    1.1425628          0
3        0 44.25050 53.33719 50.52147 8.451629 26.95073    1.6838666          0
```

```
4           0 50.99654 46.67895 58.91263 8.275673 16.93017   0.6402004           0
5           0 74.00543 47.12970 43.99996 7.103932 23.78938   1.4528032           4
6           0  9.53415  9.94137 98.55614 1.109623 28.76158   0.4667785           0
  conflict Matmor Infantmor Neonatmor Under5mor drought earthquake
1       NA      9      36.3      22.8      42.0       0          0
2       NA     16       8.4       6.3       9.4       0          0
3       NA     37       8.4       5.9       9.8       0          0
4       NA     28      17.0      11.7      19.9       0          0
5        0     62      14.6       9.7      16.3       1          0
6       NA    803      93.7      41.6     215.2       1          0
```

```
analyticaldata |>
  summary()
```

```
 country_name            ISO               region               year
 Length:3720        Length:3720        Length:3720        Min.   :2000
 Class :character   Class :character   Class :character   1st Qu.:2005
 Mode  :character   Mode  :character   Mode  :character   Median :2010
                                                          Mean   :2010
                                                          3rd Qu.:2014
                                                          Max.   :2019

    gdp1000             OECD           OECD2023           popdens
 Min.   :  0.1105   Min.   :0.000   Min.   :0.0000   Min.   : 0.00
 1st Qu.:  1.2383   1st Qu.:0.000   1st Qu.:0.0000   1st Qu.:14.79
 Median :  4.0719   Median :0.000   Median :0.0000   Median :27.52
 Mean   : 11.4917   Mean   :0.171   Mean   :0.1882   Mean   :30.57
 3rd Qu.: 13.1531   3rd Qu.:0.000   3rd Qu.:0.0000   3rd Qu.:40.72
 Max.   :123.6787   Max.   :1.000   Max.   :1.0000   Max.   :99.86
 NA's   :62                                          NA's   :20
     urban             agedep          male_edu           temp
 Min.   : 0.1025   Min.   : 16.17   Min.   : 1.067   Min.   :-2.405
 1st Qu.:17.2872   1st Qu.: 47.94   1st Qu.: 5.904   1st Qu.:12.928
 Median :30.2535   Median : 55.51   Median : 8.368   Median :21.958
 Mean   :30.6948   Mean   : 61.94   Mean   : 8.258   Mean   :19.625
 3rd Qu.:41.6558   3rd Qu.: 77.11   3rd Qu.:10.849   3rd Qu.:25.869
 Max.   :93.4135   Max.   :111.48   Max.   :14.441   Max.   :29.676
 NA's   :20                         NA's   :20       NA's   :20
  rainfall1000        totaldeath         conflict           Matmor
 Min.   :0.01993   Min.   :    0.00   Min.   :0.0000   Min.   :   2.0
 1st Qu.:0.59146   1st Qu.:    0.00   1st Qu.:0.0000   1st Qu.:  17.0
 Median :1.01288   Median :    0.00   Median :0.0000   Median :  66.0
```

```
Mean   :1.20216   Mean   :   56.59   Mean   :0.4943   Mean   : 210.6
3rd Qu.:1.68706   3rd Qu.:    1.00   3rd Qu.:1.0000   3rd Qu.: 299.8
Max.   :4.71081   Max.   :48183.00   Max.   :1.0000   Max.   :2480.0
NA's   :20                           NA's   :2757     NA's   :426
   Infantmor          Neonatmor          Under5mor           drought
Min.   :  1.60    Min.   : 0.80    Min.   :  2.00    Min.   :0.00000
1st Qu.:  7.60    1st Qu.: 4.90    1st Qu.:  9.00    1st Qu.:0.00000
Median : 18.90    Median :12.10    Median : 22.20    Median :0.00000
Mean   : 28.90    Mean   :16.18    Mean   : 40.50    Mean   :0.08737
3rd Qu.: 44.52    3rd Qu.:25.32    3rd Qu.: 61.33    3rd Qu.:0.00000
Max.   :138.10    Max.   :60.90    Max.   :224.90    Max.   :1.00000
NA's   :20        NA's   :20       NA's   :20
   earthquake
Min.   :0.00000
1st Qu.:0.00000
Median :0.00000
Mean   :0.08333
3rd Qu.:0.00000
Max.   :1.00000
```

##assign 0s to NAs for conflict data

```
analyticaldata$conflict[is.na(analyticaldata$conflict)] <- 0
```

```
analyticaldata %>%
  group_by(conflict) %>%
 summarise(mean = mean(Matmor, na.rm = TRUE)) %>%
  arrange(desc(mean))
```

```
# A tibble: 2 x 2
  conflict  mean
     <dbl> <dbl>
1        1  451.
2        0  175.
```

The mean for the maternal mortality is higher for an armed conflict compared to no conflict.
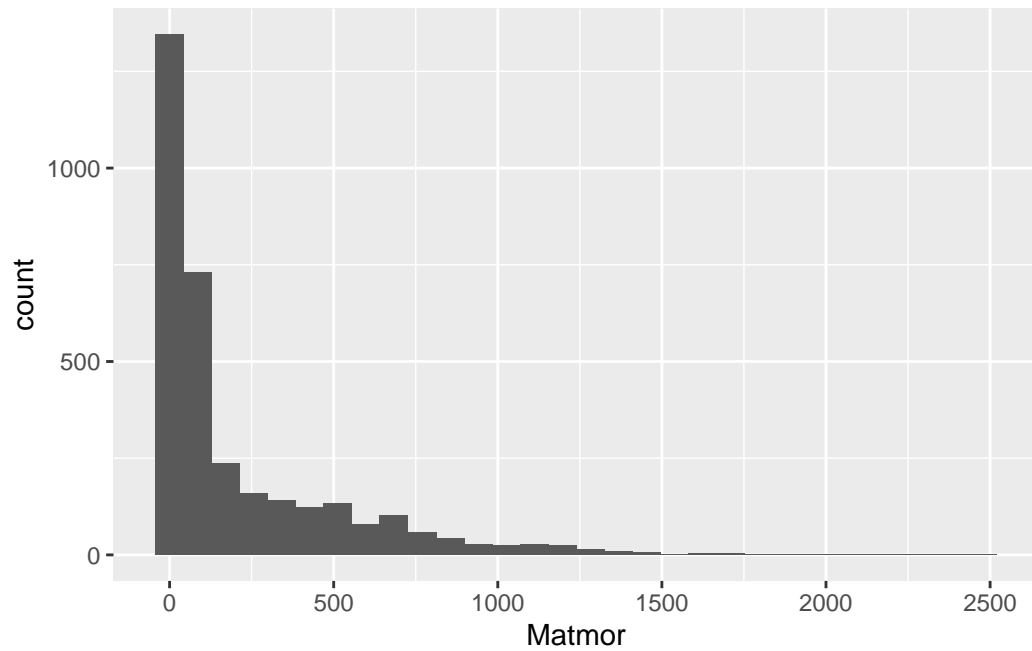
```
analyticaldata %>%
  group_by(conflict) %>%
 summarise(mean = mean(Infantmor, na.rm = TRUE)) %>%
  arrange(desc(mean))
```

```
# A tibble: 2 x 2
  conflict  mean
     <dbl> <dbl>
1        1  50.6
2        0  25.7
```

The mean for the infant mortality is higher for an armed conflict compared to no conflict.

```
analyticaldata %>%
  group_by(conflict) %>%
 summarise(mean = mean(Neonatmor, na.rm = TRUE)) %>%
  arrange(desc(mean))
```

```
# A tibble: 2 x 2
  conflict  mean
     <dbl> <dbl>
1        1  27.5
2        0  14.5
```

The mean for the neonatal mortality is higher for an armed conflict compared to no conflict.

```
analyticaldata %>%
  group_by(conflict) %>%
 summarise(mean = mean(Under5mor, na.rm = TRUE)) %>%
  arrange(desc(mean))
```

```
# A tibble: 2 x 2
  conflict  mean
     <dbl> <dbl>
1        1  73.9
2        0  35.6
```

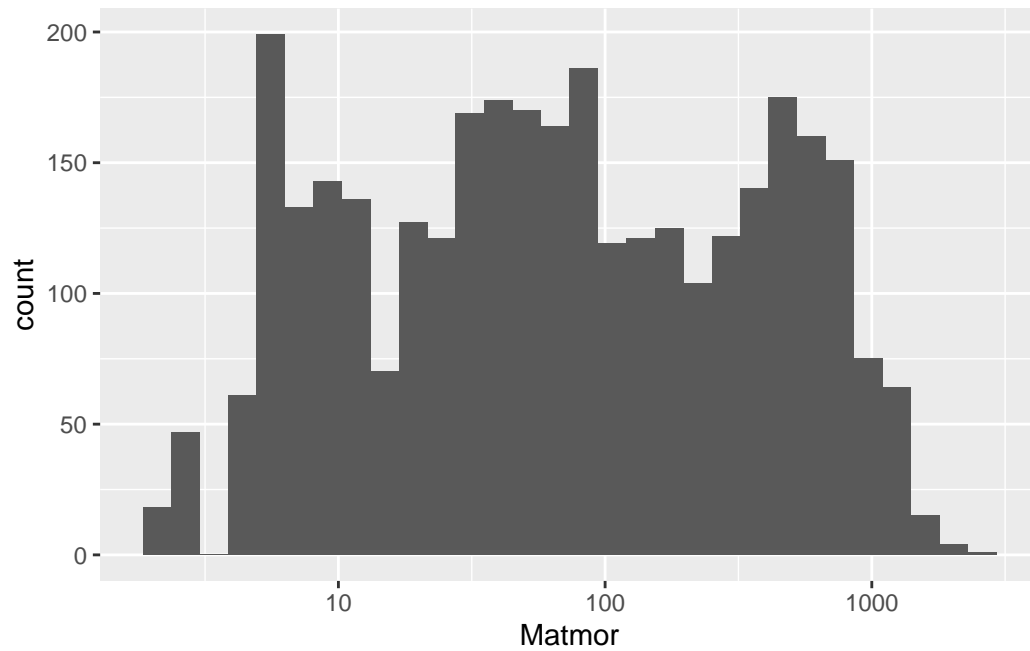The mean for the under 5 mortality is higher for an armed conflict compared to no conflict.

```
analyticaldata |>
  ggplot(aes(x = Matmor)) +
  geom_histogram(bins = 30)
```

```
Warning: Removed 426 rows containing non-finite values (`stat_bin()`).
```

```
analyticaldata |>
  ggplot(aes(x = Matmor)) +
  geom_histogram(bins = 30) +
  scale_x_log10()
```

Warning: Removed 426 rows containing non-finite values (`stat_bin()`).

```
analyticaldata |>
  ggplot() +
  geom_col(
    aes(
      x = conflict,
      y = Matmor,
      fill = region
    ),
    position = "dodge"
  )
```

Warning: Removed 426 rows containing missing values (`geom_col()`).