

# CHL8010: Statistical Programming and Computation in Health Data

2024-10-07

```
head(Final_data,10)
```

	country_name	ISO	region	gdp1000	OECD	OECD2023	popdens	urban	
1	Afghanistan	AFG	Southern Asia	NA	0	0	14.13654	16.25324	
2	Afghanistan	AFG	Southern Asia	NA	0	0	14.23156	16.25661	
3	Afghanistan	AFG	Southern Asia	0.1835328	0	0	14.32270	16.42654	
4	Afghanistan	AFG	Southern Asia	0.2004626	0	0	14.40691	16.60701	
5	Afghanistan	AFG	Southern Asia	0.2216576	0	0	15.21947	16.71367	
6	Afghanistan	AFG	Southern Asia	0.2550551	0	0	15.33619	16.85096	
7	Afghanistan	AFG	Southern Asia	0.2740005	0	0	15.43982	16.98105	
8	Afghanistan	AFG	Southern Asia	0.3750781	0	0	15.65217	17.12259	
9	Afghanistan	AFG	Southern Asia	0.3878492	0	0	15.74447	17.26919	
10	Afghanistan	AFG	Southern Asia	0.4438452	0	0	15.83043	17.43508	
	agedep	male_edu	temp	rainfall1000	Year	Totdeath	Conflict	MatMor	NeoMor
1	108.3466	2.762086	12.69959	0.2763704	2000	5065	1	1450	60.9
2	108.9899	2.856936	12.85570	0.2793079	2001	5394	1	1390	59.7
3	109.3472	2.954241	12.71081	0.3805710	2002	5553	1	1300	58.5
4	109.4475	3.054121	12.16592	0.4288939	2003	1157	1	1240	57.2
5	109.2868	3.156706	13.04643	0.3754336	2004	944	1	1180	55.9
6	107.9646	3.262133	12.23141	0.4415680	2005	817	1	1140	54.6
7	106.3262	3.370551	12.96153	0.4437097	2006	1711	1	1120	53.2
8	108.3381	3.482112	12.47451	0.4092555	2007	4982	1	1090	51.7
9	109.2404	3.596977	12.63527	0.3901204	2008	7020	1	1030	50.3
10	106.8458	3.715306	12.61764	0.4808727	2009	5660	1	993	48.9
	InfMor	Und5Mor	drought	earthquake					
1	90.5	129.2	1	0					
2	87.9	125.2	0	2					
3	85.3	121.1	0	3					

4	82.7	116.9	0	1
5	80.0	112.6	0	1
6	77.3	108.4	0	2
7	74.6	104.1	1	1
8	71.9	99.9	0	0
9	69.2	95.7	1	0
10	66.7	91.7	0	1

```
# Restrict countries to 2000 & remove unnecessary columns
data_2000 <- Final_data %>%
  filter(Year == 2000) %>%
  select(-country_name, -region, -ISO, -OECD2023, -Year)

# Make Conflict a factor and label levels
data_2000$Conflict <- factor(data_2000$Conflict,
                             levels = c(0, 1),
                             labels = c("No Conflict in Year 2000",
                                           "Conflict in Year 2000"))

data_2000$OECD <- factor(data_2000$OECD,
                         levels = c(0, 1),
                         labels = c("No", "Yes"))

data_2000$drought <- factor(data_2000$drought,
                            levels = c(0, 1),
                            labels = c("No", "Yes"))

data_2000$earthquake <- factor(data_2000$earthquake,
                               levels = c(0, 1),
                               labels = c("No", "Yes"))

# Apply labels to outcomes and covariates
label(data_2000$MatMor) <- "Maternal mortality ratio per 100,000 live births"
label(data_2000$NeoMor) <- "Neonatal mortality rate per 1,000 live births"
label(data_2000$InfMor) <- "Infant mortality rate per 1,000 live births"
label(data_2000$Und5Mor) <- "Under-5 mortality rate per 1,000 live births"
label(data_2000$gdp1000) <- "GDP per capita"
label(data_2000$OECD) <- "OECD member"
label(data_2000$popdens) <- "Population density"
label(data_2000$urban) <- "Urban residence"
label(data_2000$agedep) <- "Age dependency ratio"
label(data_2000$male_edu) <- "Male education"
```

```

label(data_2000$temp) <- "Temperature"
label(data_2000$rainfall1000) <- "Rainfall"
label(data_2000$earthquake) <- "Earthquakes"
label(data_2000$drought) <- "Droughts"

#Create Table
caption <- "Table of Armed Conflict Demographics"
table1_2000_latex <- table1(~ MatMor + NeoMor + InfMor + Und5Mor + gdp1000 + OECD +
                             popdens + urban + agedep + male_edu + temp +
                             rainfall1000 + earthquake + drought | Conflict,
                             data = data_2000,
                             caption = caption,
                             overall = c(left = "Total"),
                             render.continuous = c(.= "Median [Min, Max]"))

# Convert to kable for LaTeX compatibility
table_latex <- kable(table1_2000_latex, "latex", booktabs = TRUE, caption = caption) %>%
  kable_styling(latex_options = c("HOLD_position", "scale_down"))
print(table_latex)

```

Table 1: Table of Armed Conflict Demographics

	No Conflict in Year 2000	Conflict in Year 2000	Total
	(N=147)	(N=39)	(N=186)
Maternal mortality ratio per 100,000 live births			
Median [Min, Max]	57.0 [3.00, 1730]	553 [13.0, 2480]	77.0 [3.00, 2480]
Missing	3 (2.0%)	0 (0%)	3 (1.6%)
Neonatal mortality rate per 1,000 live births			
Median [Min, Max]	12.9 [1.60, 56.0]	36.6 [7.80, 60.9]	16.8 [1.60, 60.9]
Missing	1 (0.7%)	0 (0%)	1 (0.5%)
Infant mortality rate per 1,000 live births			
Median [Min, Max]	21.0 [3.00, 112]	66.7 [10.9, 138]	27.4 [3.00, 138]
Missing	1 (0.7%)	0 (0%)	1 (0.5%)
Under-5 mortality rate per 1,000 live births			
Median [Min, Max]	24.8 [3.90, 225]	98.5 [12.6, 225]	31.5 [3.90, 225]
Missing	1 (0.7%)	0 (0%)	1 (0.5%)
GDP per capita			
Median [Min, Max]	2.19 [0.137, 48.7]	0.558 [0.123, 4.80]	1.77 [0.123, 48.7]
Missing	3 (2.0%)	2 (5.1%)	5 (2.7%)
OECD member			
No	118 (80.3%)	38 (97.4%)	156 (83.9%)
Yes	29 (19.7%)	1 (2.6%)	30 (16.1%)
Population density			
Median [Min, Max]	27.3 [0, 99.8]	21.3 [0, 71.7]	25.4 [0, 99.8]
Missing	1 (0.7%)	0 (0%)	1 (0.5%)
Urban residence			
Median [Min, Max]	28.9 [0.106, 91.6]	24.1 [3.80, 49.3]	28.0 [0.106, 91.6]
Missing	1 (0.7%)	0 (0%)	1 (0.5%)
Age dependency ratio			
Median [Min, Max]	60.2 [30.0, 108]	84.4 [44.2, 111]	63.5 [30.0, 111]
Male education			
Median [Min, Max]	7.91 [1.07, 14.0]	4.94 [1.69, 11.8]	7.14 [1.07, 14.0]
Missing	1 (0.7%)	0 (0%)	1 (0.5%)
Temperature			
Median [Min, Max]	21.0 [-1.21, 28.6]	24.0 [5.09, 28.5]	21.4 [-1.21, 28.6]
Missing	1 (0.7%)	0 (0%)	1 (0.5%)
Rainfall			
Median [Min, Max]	0.998 [0.0480, 4.71]	1.07 [0.191, 3.03]	1.00 [0.0480, 4.71]
Missing	1 (0.7%)	0 (0%)	1 (0.5%)
Earthquakes			
No	134 (91.2%)	34 (87.2%)	168 (90.3%)
Yes	9 (6.1%)	3 (7.7%)	12 (6.5%)
Missing	4 (2.7%)	2 (5.1%)	6 (3.2%)
Droughts			
No	128 (87.1%)	36 (92.3%)	164 (88.2%)
Yes	17 (11.6%)	3 (7.7%)	20 (10.8%)
Missing	2 (1.4%)	0 (0%)	2 (1.1%)

#Maternal Mortality Trends between 2000 - 2017

```

Maternal_Mortality_Plot_Data <- Final_data %>%
  filter(Year == 2000 | Year == 2017) %>% # Filter only 2000 & 2017
  group_by(ISO) %>% # Group by ISO
  mutate(
    MatMor_2000 = ifelse(Year == 2000, MatMor, NA),
    MatMor_2017 = ifelse(Year == 2017, MatMor, NA)
  ) %>%
  fill(MatMor_2000, .direction = "downup") %>% # Fill in the 2000 value
  fill(MatMor_2017, .direction = "downup") %>% # Fill in the 2017 value
  filter(MatMor_2017 > MatMor_2000) %>% # Only keep ISO if 2017 > 2000
  ungroup() # Ungroup

print(Maternal_Mortality_Plot_Data)

```

```

# A tibble: 26 x 23
  country_name      ISO region  gdp1000  OECD OECD2023 popdens urban agedep
  <chr>            <chr> <chr>    <dbl> <int>   <int>   <dbl> <dbl>  <dbl>
1 Brunei          BRN  South-e~ 18.0     0       0    17.2  57.5   56.3
2 Brunei          BRN  South-e~ 28.2     0       0    22.1  58.3   39.1
3 Canada          CAN  Norther~ 24.3     1       1    66.2  56.1   46.3
4 Canada          CAN  Norther~ 45.1     1       1    70.4  59.6   49.1
5 Dominican Republic DOM Latin A~  2.85     0       0    44.7  42.4   66.1
6 Dominican Republic DOM Latin A~  7.51     0       0    50.0  48.4   53.4
7 Haiti           HTI  Latin A~  0.815    0       0    31.3  39.7   78.8
8 Haiti           HTI  Latin A~  1.38     0       0    44.0  42.7   60.7
9 Jamaica         JAM  Latin A~  3.45     0       0    23.6  38.0   64.7
10 Jamaica        JAM  Latin A~  5.27     0       0    23.2  40.8   40.7
# i 16 more rows
# i 14 more variables: male_edu <dbl>, temp <dbl>, rainfall1000 <dbl>,
#   Year <int>, Totdeath <int>, Conflict <int>, MatMor <int>, NeoMor <dbl>,
#   InfMor <dbl>, Und5Mor <dbl>, drought <int>, earthquake <int>,
#   MatMor_2000 <int>, MatMor_2017 <int>

```

```

#Q2 Plotting Maternal Mortality Increasing Trends between 2000-2017
ggplot(Maternal_Mortality_Plot_Data, aes(Year, MatMor)) +
  geom_line(aes(group = ISO), alpha = 2/5) +
  labs(title = "Maternal Mortality Trends by ISO Code
  (Only Included Countries that Increased)", x = "Year",
    y = "Maternal Mortality") +
  scale_y_log10() +
  geom_smooth(se = FALSE) +

```

```
theme(plot.title = element_text(hjust = 0.5)) +
geom_text(data = Maternal_Mortality_Plot_Data %>% filter(Year == 2017),
  aes(label = ISO),
  hjust = -0.1,
  vjust = 0.5,
  size = 2)
```

`geom\_smooth()` using method = 'loess' and formula = 'y ~ x'

Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,  
: pseudoinverse used at 1999.9

Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,  
: neighborhood radius 17.085

Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,  
: reciprocal condition number 0

Warning in simpleLoess(y, x, w, span, degree = degree, parametric = parametric,  
: There are other near singularities as well. 291.9

