summary_final_data_findings

Introduction

The primary goal of the analysis is to analyze how different variables, such as armed conflict and disaster events (droughts and earthquakes), influence mortality rates over time in various countries.

```
library(here)
```

```
here() starts at /Users/annie/Desktop/Version Control/armed_conflict_930
```

```
finaldata <- read.csv(here("data", "finaldata.csv"), header = TRUE)

num_unique_countries <- length(unique(finaldata$country_name))
print(num_unique_countries)
```

```
[1] 186
```

```
column_names <- colnames(finaldata)
print(column_names)
```

```
 [1] "country_name" "ISO"          "region"       "year"         "gdp1000"
 [6] "OECD"         "OECD2023"     "popdens"      "urban"        "agedep"
[11] "male_edu"     "temp"         "rainfall1000" "totdeath"     "armconf1"
[16] "matmor"       "infmor"       "neomor"       "un5mor"       "drought"
[21] "earthquake"
```

The data set provides detailed and cleaned information on various socio-economic and demographic indicators for 186 countries in 21 columns, respectively:

1. **country_name**: Name of the country.
2. **ISO**: ISO code for the country (usually a three-letter abbreviation).
3. **region**: Geographic region of the country (e.g., Asia, Europe).
4. **year**: Year of the data record.
5. **gdp1000**: GDP per capita in thousands (adjusted for inflation).
6. **OECD**: Indicates if the country is a member of the OECD (Organization for Economic Co-operation and Development).
7. **OECD2023**: Likely indicates whether the country is part of the OECD as of 2023.
8. **popdens**: Population density (people per square kilometer).
9. **urban**: Percentage of the population living in urban areas.
10. **agedep**: Age dependency ratio (ratio of dependents to the working-age population).
11. **male_edu**: Education level for males, possibly years of schooling or literacy rate.
12. **temp**: Average temperature (could be in Celsius or Fahrenheit).
13. **rainfall1000**: Annual rainfall in thousands of millimeters or another unit.
14. **totdeath**: Total number of deaths (could be annual or per capita).
15. **armconf1**: Likely represents armed conflict index or a similar measure.
16. **matmor**: Maternal mortality rate (number of deaths per 100,000 live births).
17. **infmor**: Infant mortality rate (number of deaths per 1,000 live births).
18. **neomor**: Neonatal mortality rate (number of deaths within the first 28 days per 1,000 live births).
19. **un5mor**: Under-5 mortality rate (number of deaths per 1,000 live births).
20. **drought**: Drought index or measure of drought occurrences.
21. **earthquake**: Index or measure of earthquake occurrences or risk.

Model

```
library(dplyr)
```

```
Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

    filter, lag

The following objects are masked from 'package:base':

    intersect, setdiff, setequal, union
```

```
model <- lm(matmor ~ drought + earthquake + armconf1 + gdp1000 + popdens + urban + agedep + m

summary(model)
```

```
Call:
lm(formula = matmor ~ drought + earthquake + armconf1 + gdp1000 +
    popdens + urban + agedep + male_edu + temp + rainfall1000,
    data = finaldata)

Residuals:
    Min      1Q  Median      3Q     Max
-445.87  -87.06   -3.86   63.78 1846.61

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  -115.04087   31.75536  -3.623 0.000296 ***
drought         4.64616   11.08292   0.419 0.675085
earthquake    -31.61549   11.19484  -2.824 0.004770 **
armconf1       61.35476    8.53815   7.186 8.26e-13 ***
gdp1000         1.43143    0.22833   6.269 4.11e-10 ***
popdens        -0.20238    0.22258  -0.909 0.363296
urban          -0.02378    0.28157  -0.084 0.932695
agedep          8.72442    0.23571  37.013  < 2e-16 ***
```

```
male_edu      -29.26461    1.91220 -15.304  < 2e-16 ***
temp            0.05475    0.64657   0.085 0.932524
rainfall1000   -4.56087    4.38903  -1.039 0.298811
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 174.7 on 3212 degrees of freedom
  (497 observations deleted due to missingness)
Multiple R-squared:  0.6369,    Adjusted R-squared:  0.6357
F-statistic: 563.3 on 10 and 3212 DF,  p-value: < 2.2e-16
```

Residuals

- **Min**: -445.87
  - This value indicates that the largest underestimation by the model is 445.87 maternal deaths. In practical terms, this means that for at least one observation, the model predicted a much lower maternal mortality than what was actually observed.
- **1Q**: -87.06
  - This statistic means that 25% of the residuals are less than -87.06. This suggests that a significant portion of the predictions underestimates the actual maternal mortality by more than 87.06 deaths.
- **Median**: -3.86
  - The median residual is close to zero, indicating that, on average, the model predictions are fairly close to the observed values. However, since the median is negative, it suggests a slight tendency for the model to underestimate maternal mortality overall.
- **3Q**: 63.78
  - This indicates that 75% of the residuals are less than 63.78. This shows that while the model has some positive residuals, they are generally not as large as the negative ones.
- **Max**: 1846.61
  - The maximum residual shows the largest overestimation by the model. This means that for at least one observation, the model predicted 1846.61 more maternal deaths than were actually observed. This is a substantial overestimation and could be a point of concern, indicating that the model may not fit well for certain countries or years.

The residuals suggest a spread of values around zero, with a wider range indicating variability in predicted values.

The range of residuals from -445.87 to 1846.61 suggests significant variability in model performance across different observations.

The skewness towards negative values (more negative residuals than positive) hints that the model generally tends to underestimate maternal mortality, particularly for observations with high actual values.

Coefficients

The coefficients indicate the estimated change in maternal mortality for a one-unit increase in each predictor, while holding other variables constant. Here are the interpretations for significant predictors:

1. **(Intercept)**: -115.04

The baseline maternal mortality when all predictors are at zero, which may not be meaningful depending on the context.

2. **Drought**: 4.65

Not statistically significant ($p = 0.675$), suggesting droughts do not have a clear effect on maternal mortality in this model.

3. **Earthquake**: -31.62

Statistically significant ($p = 0.0048$). A one-unit increase in earthquake occurrences is associated with a decrease in maternal mortality by about 31.62, holding other factors constant.

4. **Armconf1**: 61.35

Highly significant ($p < 2e\text{-}16$). This indicates that being in a region with significant armed conflict (i.e., at least 25 deaths) is associated with an increase in maternal mortality by approximately 61.35.

5. **GDP1000**: 1.43

Highly significant ($p < 4.11e\text{-}10$). Each additional thousand dollars in GDP per capita is associated with an increase in maternal mortality by about 1.43.

6. **Population Density (popdens)**: -0.20

Not statistically significant ($p = 0.363$), indicating little to no effect on maternal mortality.

7. **Urbanization (urban)**: -0.02

Not statistically significant (p = 0.933), suggesting urbanization doesn't impact maternal mortality significantly in this model.

8. **Age Dependency Ratio (agedep)**: 8.72

Highly significant (p < 2e-16). A one-unit increase in the age dependency ratio is associated with an increase in maternal mortality by approximately 8.72.

9. **Male Education Level (male_edu)**: -29.26

Highly significant (p < 2e-16). Each additional unit increase in male education is associated with a decrease in maternal mortality by approximately 29.26.

10. **Temperature (temp)**: 0.055

Not statistically significant (p = 0.933), indicating temperature does not appear to have a significant impact on maternal mortality.

11. **Rainfall (rainfall1000)**: -4.56

Not statistically significant (p = 0.299), suggesting rainfall does not significantly affect maternal mortality.

Model Fit

- **Multiple R-squared**: 0.6369 indicates that approximately 63.69% of the variance in maternal mortality can be explained by this model.
- **Adjusted R-squared**: 0.6357, which accounts for the number of predictors in the model, remains quite high.
- **F-statistic**: 563.3 (p < 2.2e-16) shows that the model overall is statistically significant.

Key Insights

Significant predictors that positively correlate with maternal mortality include armed conflict and age dependency ratio, while male education and earthquakes correlate negatively.

- Armed Conflict (`armconf1`):

The model indicates a strong positive relationship between armed conflict and maternal mortality. Specifically, the coefficient for `armconf1` is approximately 61.35, suggesting that, on average, countries experiencing armed conflict have significantly higher maternal mortality rates compared to those that do not.

This could be due to several factors such as healthcare disruption and resource allocation.

Armed conflict can lead to the breakdown of healthcare systems, making it difficult for women to access prenatal and postnatal care.

In conflict zones, resources are often diverted from healthcare to military efforts, resulting in inadequate care for pregnant women.

- Age Dependency Ratio (`agedep`):

The positive relationship between the age dependency ratio and maternal mortality indicates that higher proportions of dependents (younger or older individuals not in the labor force) in a population correlate with increased maternal mortality rates. The coefficient of approximately 8.72 suggests that as the age dependency ratio increases, maternal mortality also tends to rise.

This could be due to several factors such as economic strain and healthcare access.

High dependency ratios can strain household resources, making it difficult for families to afford healthcare.

In countries with high dependency ratios, the economic burden may lead to less investment in maternal healthcare services.

- Male Education (`male_edu`):

The negative coefficient for male education (approximately -29.26) suggests that higher levels of education among men are associated with lower maternal mortality rates.

This could be due to reasons such as increased resources and supportive environment.

Higher education levels are often linked to better economic opportunities, allowing families to invest more in healthcare.

Educated men may be more supportive of women's health issues, promoting better maternal health practices within households.

- Earthquakes (`earthquake`):

The negative coefficient for earthquakes (approximately -31.62) implies that, in the context of this model, earthquake occurrences correlate with lower maternal mortality rates. This finding may seem counterintuitive but could be explained by migration effects and focus on rebuilding.

In some cases, disasters can lead to population movements, potentially bringing women into areas with better healthcare access.

Disaster recovery efforts might prioritize maternal health as part of broader public health initiatives.

Conclusion

The results highlight the intricate nature of the factors influencing maternal mortality, pointing out that both social and environmental elements are crucial. By concentrating on the key predictors identified in the model, stakeholders can more effectively design and implement interventions aimed at lowering maternal mortality rates. Policymakers should prioritize educational programs, especially for men, and offer support in regions affected by conflict to enhance maternal health outcomes. Ongoing monitoring of maternal mortality in relation to these predictors will aid in refining interventions and policies, ensuring they address the evolving challenges faced by mothers. The interaction of these variables underscores the necessity for a comprehensive approach to maternal health, which should consider not only access to healthcare but also wider socio-economic issues.