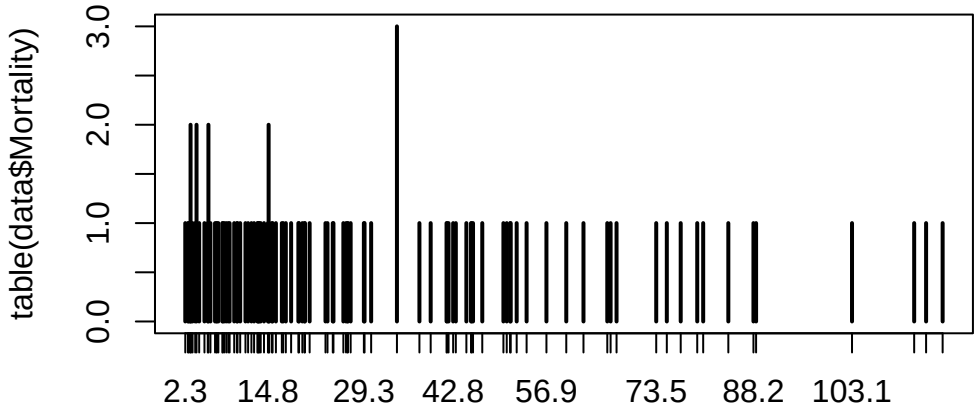
Ordinal, Generalized ordinal, Poisson Regression Models

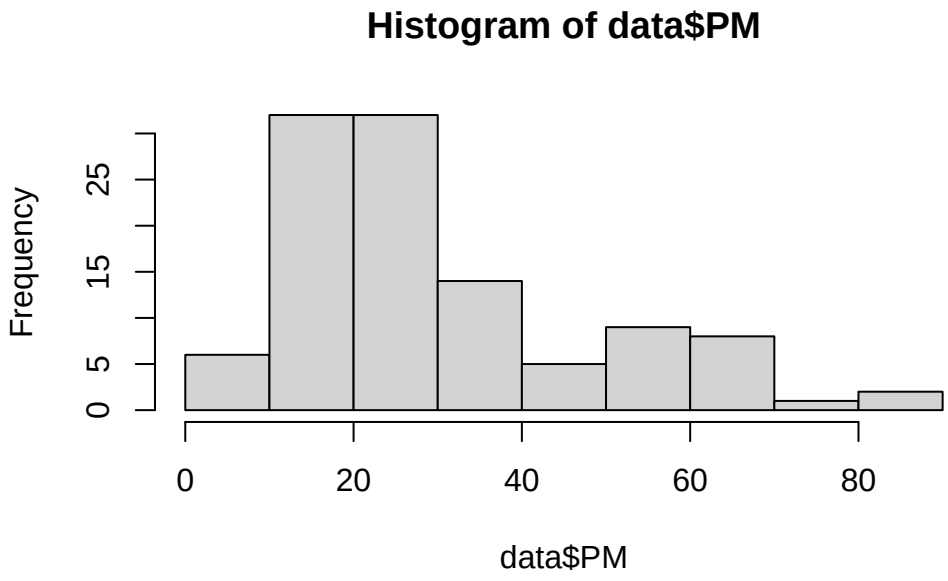
```r
# Install Library
library(readr)

# Ordinal regression packages
library(VGAM)

# Install and load the "MASS" package if not already installed
if (!require("MASS")) {
  install.packages("MASS")
  library("MASS")
}
```

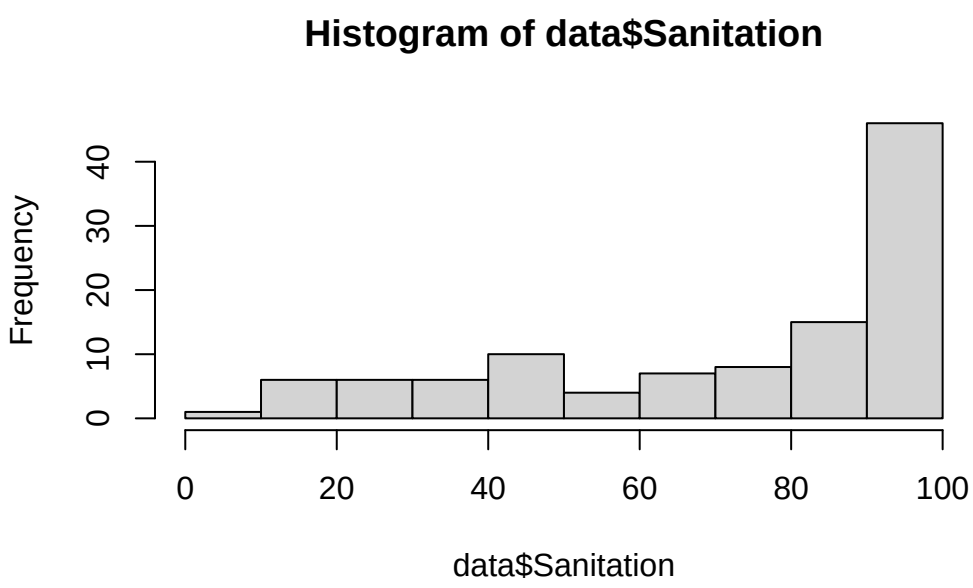
```
Rows: 109 Columns: 38
-- Column specification --------------------------------------------------------
Delimiter: ","
chr  (7): Country, CC, pm_category, pm_ordinal, Region, Mortality_Category, ...
dbl (31): CF, GDP, RE, CHE, CHE_GDP, CO2, CO2m, IMD, IMH, IMM, F, Sanitation...

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

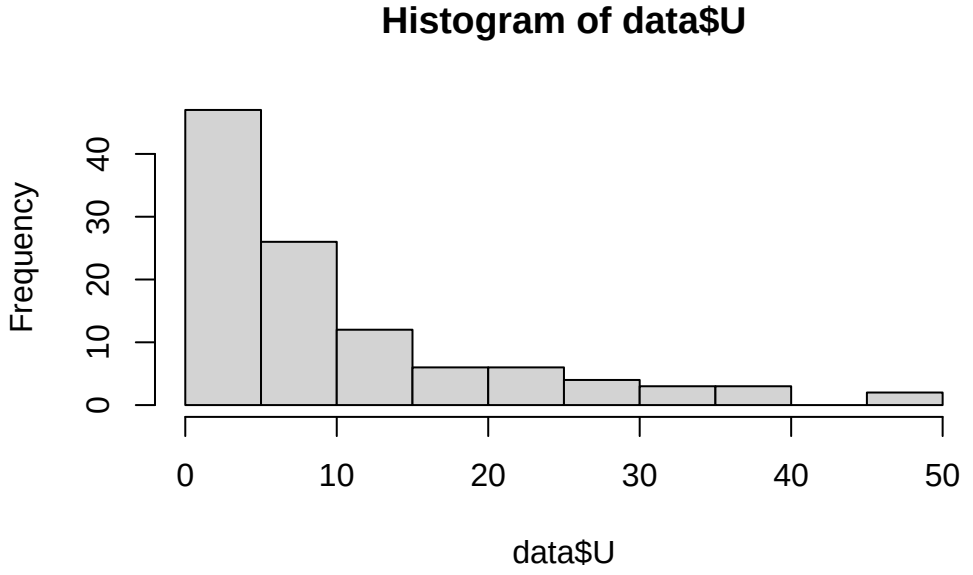
Data Summaries of potential final model variables

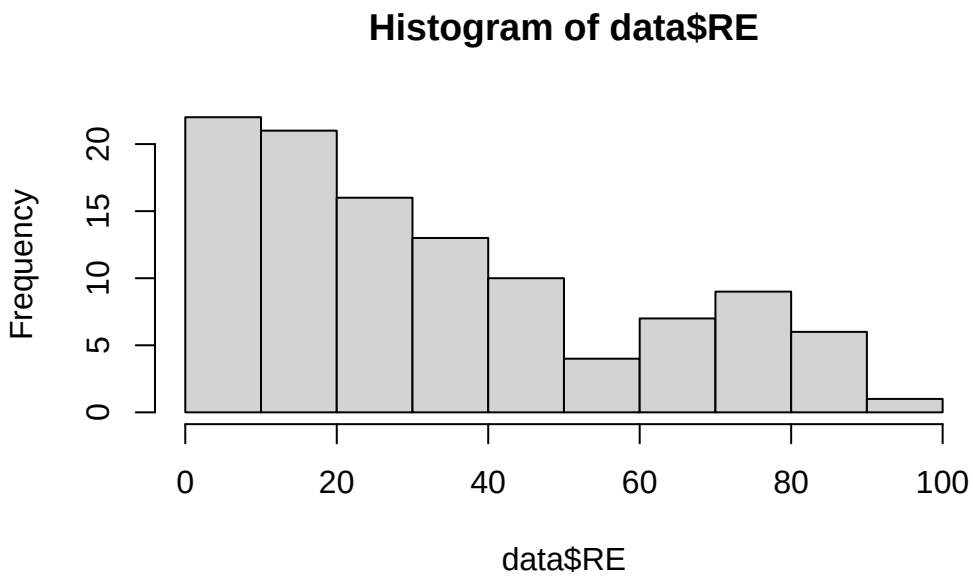


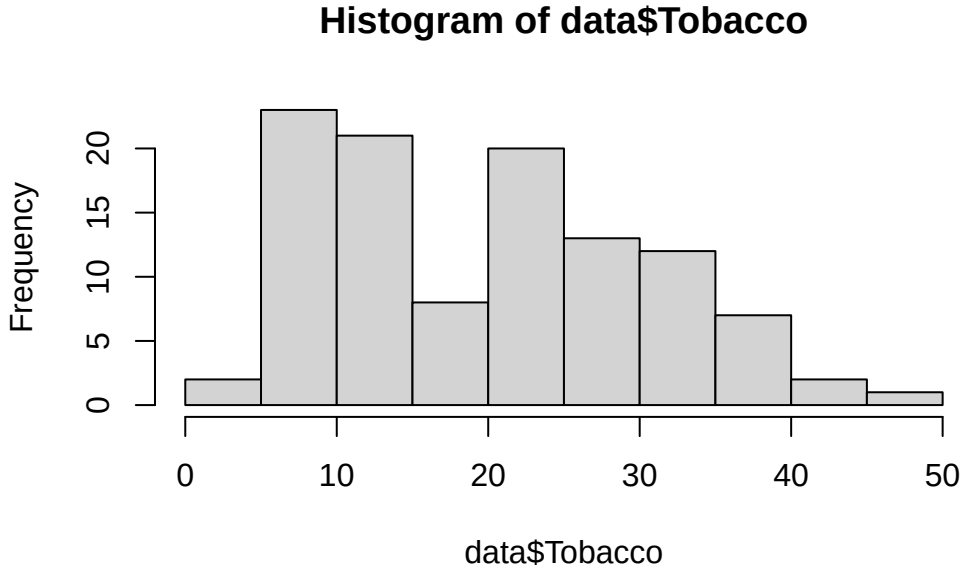
2.3	2.7	2.8	3	3.1	3.2	3.4	3.8	4	4.4	5.2	5.7	5.8
1	1	1	1	2	1	1	1	2	1	1	1	2
6.1	6.8	7	7.1	7.3	7.9	8	8.2	8.5	8.8	9	9.7	10.1
1	1	1	1	1	1	1	1	1	1	1	1	1
10.2	10.6	11.4	11.8	12.3	12.7	13.2	13.3	13.4	13.5	13.7	14.2	14.8
1	1	1	1	1	1	1	1	1	1	1	1	1
14.9	15	15.4	15.5	16	16.9	17.1	17.5	18.3	19.4	19.5	20	20.3
2	1	1	1	1	1	1	1	1	1	1	1	1
20.4	21.1	23.5	23.8	24.6	24.7	26.2	26.6	26.8	26.9	27.3	29.3	29.4
1	1	1	1	1	1	1	1	1	1	1	1	1
30.4	34.3	37.7	39.4	41.8	41.9	42.1	42.8	43.2	44.8	45.5	45.6	45.8
1	3	1	1	1	1	1	1	1	1	1	1	1
47.2	50.4	50.9	51.4	51.5	52.4	53.9	56.9	59.9	62.5	66.1	66.6	67.5
1	1	1	1	1	1	1	1	1	1	1	1	1
73.5	75.1	77.2	79.7	80.6	84.4	88.2	88.6	103.1	112.5	114.3	116.8	
1	1	1	1	1	1	1	1	1	1	1	1	

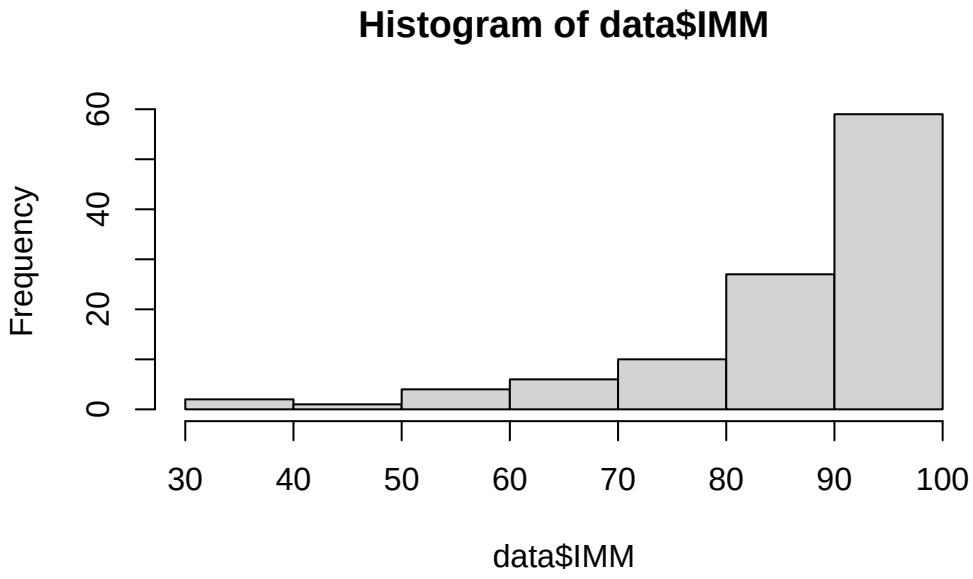



Histogram of data$Sanitation
Frequency
0
10
20
30
40
0
20
40
60
80
100
data$Sanitation


Histogram of data$U
Frequency
0
10
20
30
40
0
10
20
30
40
50
data$U


Histogram of data$RE
Frequency
0
5
10
15
20
0
20
40
60
80
100
data$RE


Histogram of data$Tobacco
Frequency
0
5
10
15
20
0
10
20
30
40
50
data$Tobacco



(5b) cont'd

Ordinal

The ordinal model is not a good fit for the data, as demonstrated below.

Ordinal data formulation:

Refer to the section titled: *Data Formulation To fit a Logistic, Multinomial, Ordinal, Generalized Ordinal, Poisson*

```
Call:
vglm(formula = Mortality_Ordinal ~ PM, family = cumulative(parallel = TRUE,
    reverse = FALSE), data = data)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.03076    0.46224  -4.393 1.12e-05 ***
PM           0.05921    0.01436   4.124 3.72e-05 ***
```

```
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Name of linear predictor: logitlink(P[Y<=1])

Residual deviance: 126.2788 on 107 degrees of freedom

Log-likelihood: -63.1394 on 107 degrees of freedom

Number of Fisher scoring iterations: 4

No Hauck-Donner effect found in any of the estimates


Exponentiated coefficients:
      PM
1.061003



Call:
vglm(formula = Mortality_Ordinal ~ PM + PM_squared, family = cumulative(parallel = TRUE,
    reverse = FALSE), data = data)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.1703822  0.9596585  -2.262   0.0237 *
PM           0.0689166  0.0598978   1.151   0.2499
PM_squared  -0.0001317  0.0007853  -0.168   0.8668
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Name of linear predictor: logitlink(P[Y<=1])

Residual deviance: 126.2512 on 106 degrees of freedom

Log-likelihood: -63.1256 on 106 degrees of freedom

Number of Fisher scoring iterations: 5

No Hauck-Donner effect found in any of the estimates
```

```
Exponentiated coefficients:
        PM PM_squared
 1.0713468  0.9998683


Call:
vglm(formula = Mortality_Ordinal ~ PM + CO2m + PM * CO2m, family = cumulative(parallel = TRU
    reverse = FALSE), data = data)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -0.4862332  0.8303266  -0.586   0.5581
PM           0.0523685  0.0250454   2.091   0.0365 *
CO2m        -0.5615484  0.3570660  -1.573   0.1158
PM:CO2m      0.0009029  0.0113598   0.079   0.9367
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Name of linear predictor: logitlink(P[Y<=1])

Residual deviance: 98.8368 on 105 degrees of freedom

Log-likelihood: -49.4184 on 105 degrees of freedom

Number of Fisher scoring iterations: 6

No Hauck-Donner effect found in any of the estimates


Exponentiated coefficients:
       PM      CO2m   PM:CO2m
1.0537639 0.5703253 1.0009033


Call:
vglm(formula = Mortality_Ordinal ~ PM + CO2m + GDP + GDP * CO2m,
    family = cumulative(parallel = TRUE, reverse = FALSE), data = data)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  1.618e+00  9.204e-01   1.758   0.0788 .
PM           2.354e-02  1.866e-02   1.261   0.2072
```

```
CO2m        -2.525e-01  1.864e-01  -1.355   0.1755
GDP         -5.228e-04  1.733e-04      NA       NA
CO2m:GDP     2.458e-05  1.058e-05   2.323   0.0202 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Name of linear predictor: logitlink(P[Y<=1])

Residual deviance: 83.4882 on 104 degrees of freedom

Log-likelihood: -41.7441 on 104 degrees of freedom

Number of Fisher scoring iterations: 7

Warning: Hauck-Donner effect detected in the following estimate(s):
'GDP'


Exponentiated coefficients:
       PM      CO2m       GDP  CO2m:GDP
1.0238153 0.7768465 0.9994773 1.0000246



Call:
vglm(formula = Mortality_Ordinal ~ PM + Sanitation + U + +IMM,
    family = cumulative(parallel = TRUE, reverse = FALSE), data = data)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  5.37484    2.82730   1.901  0.05730 .
PM           0.03668    0.01848   1.985  0.04713 *
Sanitation  -0.06028    0.01882  -3.203  0.00136 **
U            0.07496    0.05117   1.465  0.14296
IMM         -0.03325    0.02743  -1.212  0.22552
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Name of linear predictor: logitlink(P[Y<=1])

Residual deviance: 70.4662 on 104 degrees of freedom

Log-likelihood: -35.2331 on 104 degrees of freedom
```

```
Number of Fisher scoring iterations: 6

No Hauck-Donner effect found in any of the estimates


Exponentiated coefficients:
        PM Sanitation          U        IMM
 1.0373585  0.9415007  1.0778434  0.9672985



Call:
vglm(formula = Mortality_Ordinal ~ PM + Tobacco + Sanitation +
    IMM + RE + CO2 + GDP, family = cumulative(parallel = TRUE,
    reverse = FALSE), data = data)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  5.489e+00  3.238e+00   1.695   0.0900 .
PM           3.233e-02  2.233e-02   1.448   0.1477
Tobacco     -1.110e-02  3.185e-02  -0.349   0.7273
Sanitation  -4.636e-02  2.064e-02  -2.247   0.0247 *
IMM         -2.860e-02  2.806e-02  -1.019   0.3080
RE           1.305e-02  1.717e-02   0.760   0.4472
CO2         -1.917e-07  5.977e-07  -0.321   0.7484
GDP         -1.737e-04  1.250e-04      NA       NA
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Name of linear predictor: logitlink(P[Y<=1])

Residual deviance: 67.6773 on 101 degrees of freedom

Log-likelihood: -33.8386 on 101 degrees of freedom

Number of Fisher scoring iterations: 7

Warning: Hauck-Donner effect detected in the following estimate(s):
'GDP'


Exponentiated coefficients:
```

```
       PM    Tobacco Sanitation        IMM         RE        CO2        GDP
1.0328540  0.9889580  0.9546942  0.9718014  1.0131316  0.9999998  0.9998263
```

Generalized ordinal

The generalized ordinal model is not a good fit for the data, as demonstrated below.

```
Call:
vglm(formula = Mortality_Ordinal ~ PM, family = cumulative(parallel = FALSE,
    reverse = T), data = data)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  2.03076    0.46224   4.393 1.12e-05 ***
PM          -0.05921    0.01436  -4.124 3.72e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Name of linear predictor: logitlink(P[Y>=2])

Residual deviance: 126.2788 on 107 degrees of freedom

Log-likelihood: -63.1394 on 107 degrees of freedom

Number of Fisher scoring iterations: 4

No Hauck-Donner effect found in any of the estimates


Exponentiated coefficients:
       PM
0.9425043


Call:
vglm(formula = Mortality_Ordinal ~ PM + PM_squared, family = cumulative(parallel = FALSE,
    reverse = T), data = data)

Coefficients:
```

```
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  2.1703822  0.9596585   2.262   0.0237 *
PM          -0.0689166  0.0598978  -1.151   0.2499
PM_squared   0.0001317  0.0007853   0.168   0.8668
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Name of linear predictor: logitlink(P[Y>=2])

Residual deviance: 126.2512 on 106 degrees of freedom

Log-likelihood: -63.1256 on 106 degrees of freedom

Number of Fisher scoring iterations: 5

No Hauck-Donner effect found in any of the estimates


Exponentiated coefficients:
        PM PM_squared
 0.9334045  1.0001317



Call:
vglm(formula = Mortality_Ordinal ~ PM + CO2m + PM * CO2m, family = cumulative(parallel = FALS
    reverse = T), data = data)

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept)  0.4862332  0.8303266   0.586   0.5581
PM          -0.0523685  0.0250454  -2.091   0.0365 *
CO2m         0.5615484  0.3570660   1.573   0.1158
PM:CO2m     -0.0009029  0.0113598  -0.079   0.9367
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Name of linear predictor: logitlink(P[Y>=2])

Residual deviance: 98.8368 on 105 degrees of freedom

Log-likelihood: -49.4184 on 105 degrees of freedom
```

```
Number of Fisher scoring iterations: 6

No Hauck-Donner effect found in any of the estimates


Exponentiated coefficients:
       PM      CO2m   PM:CO2m
0.9489791 1.7533854 0.9990975



Call:
vglm(formula = Mortality_Ordinal ~ PM + CO2m + GDP + GDP * CO2m,
    family = cumulative(parallel = FALSE, reverse = T), data = data)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.618e+00  9.204e-01  -1.758   0.0788 .
PM          -2.354e-02  1.866e-02  -1.261   0.2072
CO2m         2.525e-01  1.864e-01   1.355   0.1755
GDP          5.228e-04  1.733e-04      NA       NA
CO2m:GDP    -2.458e-05  1.058e-05  -2.323   0.0202 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Name of linear predictor: logitlink(P[Y>=2])

Residual deviance: 83.4882 on 104 degrees of freedom

Log-likelihood: -41.7441 on 104 degrees of freedom

Number of Fisher scoring iterations: 7

Warning: Hauck-Donner effect detected in the following estimate(s):
'GDP'


Exponentiated coefficients:
       PM      CO2m       GDP  CO2m:GDP
0.9767387 1.2872556 1.0005229 0.9999754



Call:
```

```
vglm(formula = Mortality_Ordinal ~ PM + Sanitation + U + +IMM,
    family = cumulative(parallel = FALSE, reverse = T), data = data)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -5.37484    2.82730  -1.901  0.05730 .
PM          -0.03668    0.01848  -1.985  0.04713 *
Sanitation   0.06028    0.01882   3.203  0.00136 **
U           -0.07496    0.05117  -1.465  0.14296
IMM          0.03325    0.02743   1.212  0.22552
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Name of linear predictor: logitlink(P[Y>=2])

Residual deviance: 70.4662 on 104 degrees of freedom

Log-likelihood: -35.2331 on 104 degrees of freedom

Number of Fisher scoring iterations: 6

No Hauck-Donner effect found in any of the estimates


Exponentiated coefficients:
        PM Sanitation          U        IMM
 0.9639869  1.0621341  0.9277785  1.0338071



Call:
vglm(formula = Mortality_Ordinal ~ PM + Tobacco + Sanitation +
    IMM + RE + CO2, family = cumulative(parallel = FALSE, reverse = T),
    data = data)

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -5.254e+00  3.165e+00  -1.660 0.096954 .
PM          -4.401e-02  2.063e-02  -2.134 0.032865 *
Tobacco      2.143e-03  3.020e-02   0.071 0.943432
Sanitation   6.459e-02  1.894e-02   3.409 0.000651 ***
IMM          2.716e-02  2.740e-02   0.991 0.321601
RE          -1.686e-02  1.676e-02  -1.006 0.314553
```

```
CO2          2.842e-07  6.169e-07   0.461 0.645036
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Name of linear predictor: logitlink(P[Y>=2])

Residual deviance: 71.2439 on 102 degrees of freedom

Log-likelihood: -35.622 on 102 degrees of freedom

Number of Fisher scoring iterations: 6

No Hauck-Donner effect found in any of the estimates


Exponentiated coefficients:
        PM    Tobacco Sanitation        IMM         RE        CO2
 0.9569404  1.0021450  1.0667174  1.0275278  0.9832856  1.0000003

Warning in pchisq(deviance(mod.gen.ordinal.4) - deviance(mod.gen.ordinal.3), :
NaNs produced

[1] NaN
```

(5c)

Poisson

Poisson data formulation:

Refer to the section titled: *Data Formulation To fit a Logistic, Multinomial, Ordinal, Generalized Ordinal, Poisson*

Outline of analysis setup:

1. **Data Preparation**:

- Load and clean our dataset, ensuring that it contains the variables of interest: PM2.5 air quality, under-5 mortality (Mortality) rate, Sanitation rate, Undernourishment rate, Renewable Energy Access, Immunization (IMM) rate, and Tobacco usage.

2. **Exploratory Data Analysis (EDA)**:
    - Start with an exploratory data analysis to understand the distributions and relationships between variables. This includes summary statistics, correlation analysis, and data visualization.
3. **Model Setup**:
    - Define our outcome variable (dependent variable):
        - Under-5 mortality (Mortality)
    - Identify potential confounding variables:
        - Sanitation
        - Undernourishment
        - Renewable Energy
        - Immunization (IMM)
        - Tobacco
    - Define our predictor (independent variable):
        - PM2.5 air quality (PM)
    - Create a data frame containing the variables of interest.
4. **Assumptions Check**:
    - Check the assumptions of the poisson regression model, including mean = variance assumption.
5. **Multiple Linear Regression Model**:
    - Set up a poisson multiple linear regression model.
6. **Model Fitting**:
    - Fit the multiple linear regression model to your data.
7. **Interpretation**:

- Examine the coefficients, p-values, and confidence intervals for each predictor in the model.
- Focus on the coefficient for PM2.5 to assess its effect on Mortality while controlling for potential confounders.
- If the coefficient for PM2.5 is statistically significant, it suggests that PM2.5 has an effect on Mortality.

8. **Control for Confounding**:
    - If there is evidence of confounding, you can further explore interactions between variables or consider additional control variables.
9. **Model Assessment**:
    - Evaluate the overall goodness of fit for the model.

Having done the data preparation, exploratory data analysis and model setup, we now check for whether the assumptions of the poisson regression holds.

Checking Mean = Variance Assumption

The mean is:

```
[1] 30.62661
```

The Variance is:

```
var(data$Mortality)
```

```
[1] 771.6242
```

The Dispersion parameter (= residual deviance/df) is:

```
# Dispersion parameter (= residual deviance/df)
qpois <- glm(Mortality ~ PM, data=data, family = quasipoisson)
summary(qpois)$dispersion
```

```
[1] 17.69821
```

The Mean = Variance Assumption does not hold. The variance of this (presumably) Poisson outcome variable is much greater than the mean. Our data exhibit overdispersion. Hence, our data is not truly supportive of a Poisson modelling

```
[1] 15.93499
```

```
[1] 5.962113
```

```
[1] 5.894287
```

```
Call:
glm(formula = Mortality ~ PM + Sanitation + U + Tobacco + Tobacco_Squared +
    RE + RE_Squared + IMM + IMM_Squared + CO2, family = poisson(),
    data = data)

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)      2.006e+00  3.661e-01   5.480 4.25e-08 ***
PM               1.331e-02  1.045e-03  12.737  < 2e-16 ***
Sanitation      -1.248e-02  1.189e-03 -10.495  < 2e-16 ***
U                1.177e-02  2.348e-03   5.013 5.35e-07 ***
Tobacco         -1.891e-02  8.766e-03  -2.157 0.031031 *
Tobacco_Squared  3.859e-04  1.944e-04   1.985 0.047168 *
RE               1.011e-02  2.916e-03   3.466 0.000528 ***
RE_Squared      -8.424e-05  2.921e-05  -2.884 0.003922 **
IMM              4.990e-02  9.548e-03   5.226 1.74e-07 ***
IMM_Squared     -3.609e-04  6.612e-05  -5.459 4.80e-08 ***
CO2             -8.646e-08  3.238e-08  -2.670 0.007582 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 2462.55  on 108  degrees of freedom
Residual deviance:  584.55  on  98  degrees of freedom
AIC: Inf

Number of Fisher Scoring iterations: 5
```

```
[1] 5.96477
```

We might consider a negative binomial model as a possible alternative since our data exhibits over dispersion. Below we fit a negative binomial:

Negative Binomial Models

```
Call:
glm.nb(formula = Mortality ~ PM, data = data, init.theta = 1.740385005,
    link = log)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) 2.554498   0.146631   17.42  < 2e-16 ***
PM          0.025683   0.004169    6.16 7.26e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(1.7404) family taken to be 1)

    Null deviance: 153.98  on 108  degrees of freedom
Residual deviance: 117.31  on 107  degrees of freedom
AIC: 936.9

Number of Fisher Scoring iterations: 1

              Theta:  1.740
          Std. Err.:  0.238

 2 x log-likelihood:  -930.900
```

```
Call:
glm.nb(formula = Mortality ~ PM + PM_squared, data = data, init.theta = 1.816027869,
    link = log)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  1.9563598  0.2881134   6.790 1.12e-11 ***
PM           0.0669913  0.0175271   3.822 0.000132 ***
PM_squared  -0.0005363  0.0002153  -2.490 0.012758 *
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(1.816) family taken to be 1)

    Null deviance: 160.09  on 108  degrees of freedom
Residual deviance: 117.03  on 106  degrees of freedom
AIC: 934.11

Number of Fisher Scoring iterations: 1

              Theta:  1.816
          Std. Err.:  0.250

 2 x log-likelihood:  -926.107


Call:
glm.nb(formula = Mortality ~ PM + PM_squared + CO2m + PM * CO2m +
    PM_squared * CO2m, data = data, init.theta = 2.74925526,
    link = log)

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)      2.269e+00  3.364e-01   6.746 1.52e-11 ***
PM               7.158e-02  1.960e-02   3.652  0.00026 ***
PM_squared      -6.544e-04  2.352e-04  -2.782  0.00540 **
CO2m            -6.986e-02  7.029e-02  -0.994  0.32029
PM:CO2m         -4.691e-03  4.613e-03  -1.017  0.30923
PM_squared:CO2m  6.855e-05  6.202e-05   1.105  0.26901
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(2.7493) family taken to be 1)

    Null deviance: 232.07  on 108  degrees of freedom
Residual deviance: 116.13  on 103  degrees of freedom
AIC: 897.08

Number of Fisher Scoring iterations: 1

              Theta:  2.749
          Std. Err.:  0.414
```

```
Warning while fitting theta: alternation limit reached

 2 x log-likelihood:  -883.075


Call:
vglm(formula = Mortality_Ordinal ~ PM + PM_squared + GDP + CO2m +
    GDP * CO2m, family = cumulative(parallel = FALSE, reverse = T),
    data = data)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -2.249e+00  1.540e+00  -1.461    0.144
PM           1.865e-02  8.438e-02   0.221    0.825
PM_squared  -5.437e-04  1.078e-03  -0.504    0.614
GDP          5.329e-04  1.746e-04      NA       NA
CO2m         2.244e-01  1.926e-01   1.165    0.244
GDP:CO2m    -2.399e-05  1.070e-05  -2.242    0.025 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Name of linear predictor: logitlink(P[Y>=2])

Residual deviance: 83.2124 on 103 degrees of freedom

Log-likelihood: -41.6062 on 103 degrees of freedom

Number of Fisher scoring iterations: 7

Warning: Hauck-Donner effect detected in the following estimate(s):
'GDP'


Exponentiated coefficients:
        PM PM_squared        GDP       CO2m   GDP:CO2m
 1.0188244  0.9994564  1.0005330  1.2515252  0.9999760


Call:
vglm(formula = Mortality_Ordinal ~ PM + CO2m + GDP + GDP * CO2m,
    family = cumulative(parallel = FALSE, reverse = T), data = data)
```

```
Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.618e+00  9.204e-01  -1.758   0.0788 .
PM          -2.354e-02  1.866e-02  -1.261   0.2072
CO2m         2.525e-01  1.864e-01   1.355   0.1755
GDP          5.228e-04  1.733e-04      NA       NA
CO2m:GDP    -2.458e-05  1.058e-05  -2.323   0.0202 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Name of linear predictor: logitlink(P[Y>=2])

Residual deviance: 83.4882 on 104 degrees of freedom

Log-likelihood: -41.7441 on 104 degrees of freedom

Number of Fisher scoring iterations: 7

Warning: Hauck-Donner effect detected in the following estimate(s):
'GDP'


Exponentiated coefficients:
       PM      CO2m       GDP  CO2m:GDP
0.9767387 1.2872556 1.0005229 0.9999754



Call:
glm.nb(formula = Mortality ~ PM + PM_squared + Sanitation + U +
    RE + RE_Squared + IMM + IMM_Squared, data = data, init.theta = 5.150795942,
    link = log)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  1.303e+00  1.073e+00   1.215  0.22428
PM           2.363e-02  1.260e-02   1.875  0.06080 .
PM_squared  -1.496e-04  1.503e-04  -0.995  0.31963
Sanitation  -1.734e-02  3.356e-03  -5.168 2.36e-07 ***
U            1.032e-02  6.946e-03   1.485  0.13747
RE           1.614e-03  7.327e-03   0.220  0.82567
RE_Squared  -8.930e-07  8.100e-05  -0.011  0.99120
IMM          7.070e-02  2.734e-02   2.586  0.00972 **
```

```
IMM_Squared -4.831e-04  1.860e-04  -2.597  0.00940 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(5.1508) family taken to be 1)

    Null deviance: 393.75  on 108  degrees of freedom
Residual deviance: 115.52  on 100  degrees of freedom
AIC: 845.07

Number of Fisher Scoring iterations: 1

              Theta:  5.151
          Std. Err.:  0.903

 2 x log-likelihood:  -825.074


Call:
glm.nb(formula = Mortality ~ PM + Sanitation + U + RE + RE_Squared +
    IMM + IMM_Squared + CO2, data = data, init.theta = 5.229787106,
    link = log)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  1.482e+00  1.054e+00   1.406  0.15964
PM           1.230e-02  2.907e-03   4.230 2.33e-05 ***
Sanitation  -1.764e-02  3.267e-03  -5.399 6.69e-08 ***
U            1.119e-02  6.863e-03   1.631  0.10297
RE          -1.716e-04  6.980e-03  -0.025  0.98039
RE_Squared   1.088e-05  7.883e-05   0.138  0.89021
IMM          7.112e-02  2.726e-02   2.609  0.00909 **
IMM_Squared -4.819e-04  1.856e-04  -2.596  0.00943 **
CO2         -7.848e-08  5.288e-08  -1.484  0.13778
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(5.2298) family taken to be 1)

    Null deviance: 398.58  on 108  degrees of freedom
Residual deviance: 115.58  on 100  degrees of freedom
AIC: 843.84
```

```
Number of Fisher Scoring iterations: 1

              Theta:  5.230
          Std. Err.:  0.921

 2 x log-likelihood:  -823.844


Call:
glm.nb(formula = Mortality ~ PM + Sanitation + U + Tobacco +
    Tobacco_Squared + RE + RE_Squared + IMM + IMM_Squared + GDP,
    data = data, init.theta = 7.917885787, link = log)

Coefficients:
                  Estimate Std. Error z value Pr(>|z|)
(Intercept)      2.801e+00  9.237e-01   3.033  0.00242 **
PM               7.310e-03  2.582e-03   2.831  0.00464 **
Sanitation      -1.158e-02  2.965e-03  -3.904 9.48e-05 ***
U                9.504e-03  6.084e-03   1.562  0.11825
Tobacco         -3.874e-02  1.968e-02  -1.969  0.04896 *
Tobacco_Squared  6.751e-04  4.321e-04   1.562  0.11822
RE              -4.523e-03  6.178e-03  -0.732  0.46405
RE_Squared       4.709e-05  6.933e-05   0.679  0.49703
IMM              5.060e-02  2.346e-02   2.157  0.03102 *
IMM_Squared     -3.513e-04  1.592e-04  -2.206  0.02738 *
GDP             -5.088e-05  8.107e-06  -6.277 3.46e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(7.9179) family taken to be 1)

    Null deviance: 548.46  on 108  degrees of freedom
Residual deviance: 109.41  on  98  degrees of freedom
AIC: 808.3

Number of Fisher Scoring iterations: 1

              Theta:  7.92
          Std. Err.:  1.48

 2 x log-likelihood:  -784.301
```

```
Call:
glm.nb(formula = Mortality ~ PM + Sanitation + IMM + IMM_Squared +
    GDP, data = data, init.theta = 7.26843058, link = log)

Coefficients:
              Estimate Std. Error z value Pr(>|z|)
(Intercept)  2.682e+00  8.708e-01   3.080  0.00207 **
PM           7.757e-03  2.510e-03   3.091  0.00200 **
Sanitation  -1.526e-02  2.215e-03  -6.890 5.56e-12 ***
IMM          5.097e-02  2.364e-02   2.156  0.03111 *
IMM_Squared -3.593e-04  1.613e-04  -2.228  0.02589 *
GDP         -4.928e-05  8.040e-06  -6.129 8.82e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Negative Binomial(7.2684) family taken to be 1)

    Null deviance: 514.64  on 108  degrees of freedom
Residual deviance: 113.20  on 103  degrees of freedom
AIC: 808.68

Number of Fisher Scoring iterations: 1

              Theta:  7.27
          Std. Err.:  1.37

 2 x log-likelihood:  -794.681

Likelihood ratio tests of Negative Binomial Models

Response: Mortality
                                                                                                    Model
1                                                   PM + Sanitation + IMM + IMM_Squared + GDP
2 PM + Sanitation + U + Tobacco + Tobacco_Squared + RE + RE_Squared + IMM + IMM_Squared + GDP
     theta Resid. df    2 x log-lik.   Test    df LR stat.    Pr(Chi)
1 7.268431       103       -794.6807
2 7.917886        98       -784.3011 1 vs 2     5 10.37962 0.06516624
```