FAMD

Alayna Johnson

2024-09-19

```
library(tidyverse)
```

```
## Warning: package 'tidyverse' was built under R version 4.4.1
```

```
## Warning: package 'stringr' was built under R version 4.4.1
```

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
## ✓ dplyr     1.1.4     ✓ readr     2.1.5
## ✓ forcats   1.0.0     ✓ stringr   1.5.1
## ✓ ggplot2   3.5.1     ✓ tibble    3.2.1
## ✓ lubridate 1.9.3     ✓ tidyr     1.3.1
## ✓ purrr     1.0.2
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to becom
e errors
```

```
library(FactoMineR)
```

```
## Warning: package 'FactoMineR' was built under R version 4.4.1
```

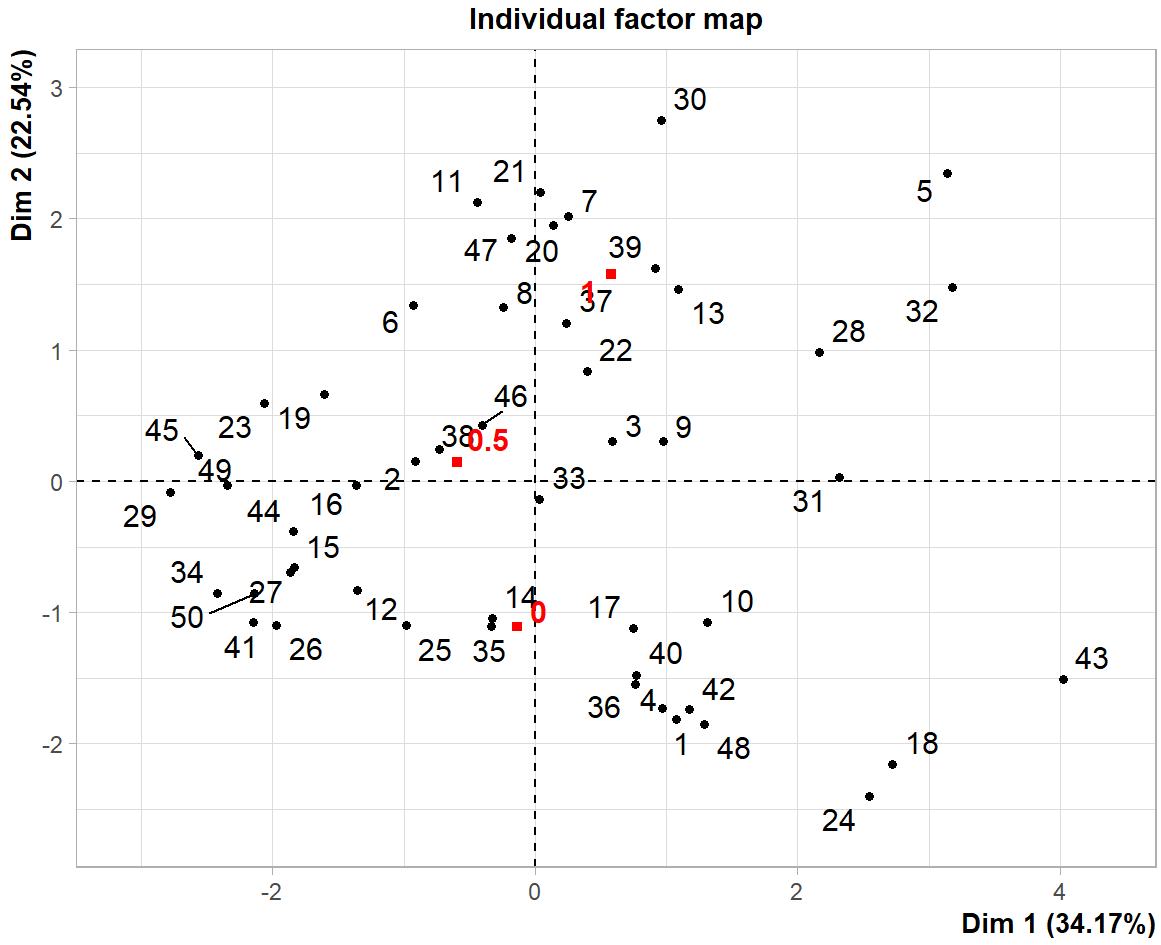
```
ctrl_vars <- read_csv("ctrl_vars.csv")
```

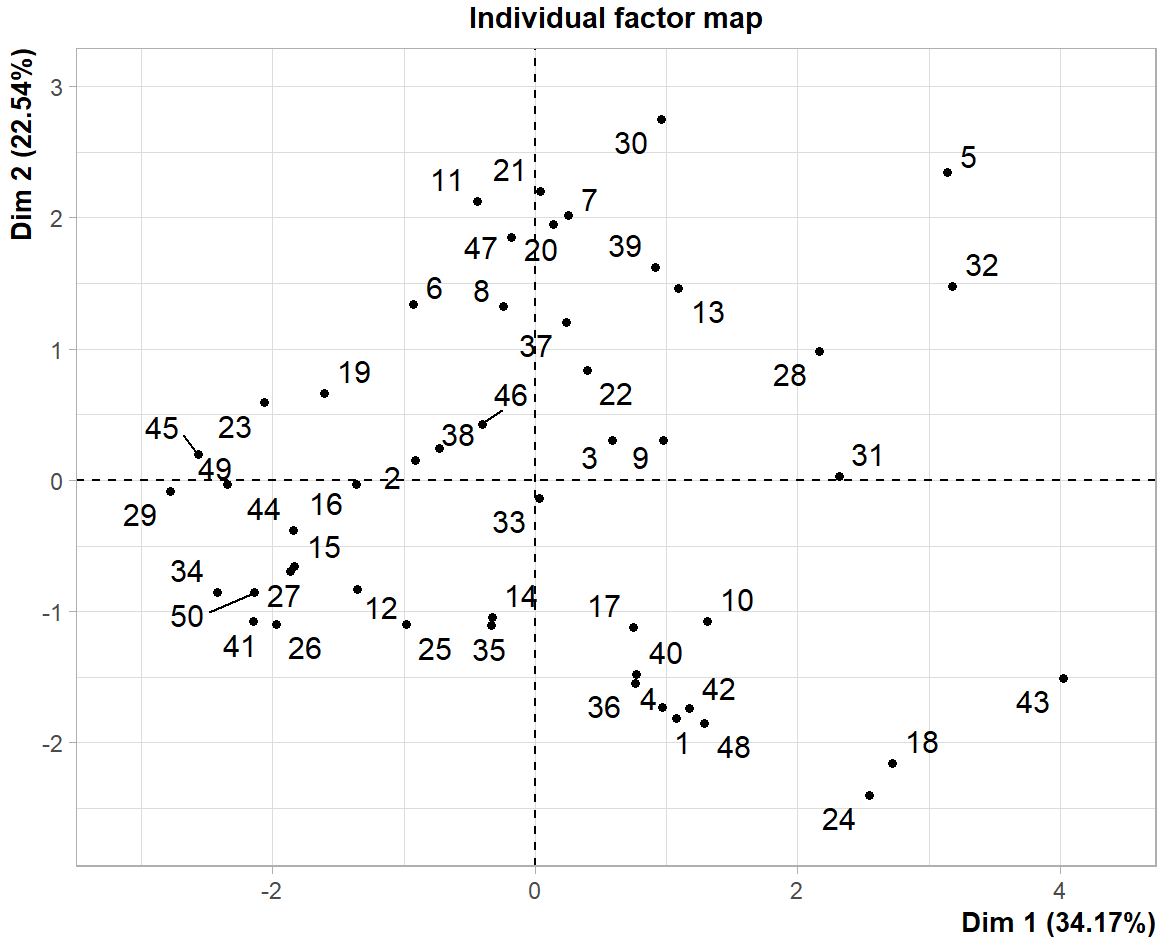
```
## Rows: 50 Columns: 56
## ── Column specification ────────────────────────────────────────────────────────
## Delimiter: ","
## chr  (1): state
## dbl (52): fips, median_age, median_income, population_2022, poverty_perc, wh...
## num  (3): foreign_born, bills_introduced, bills_enacted
##
## ℹ Use `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.
```

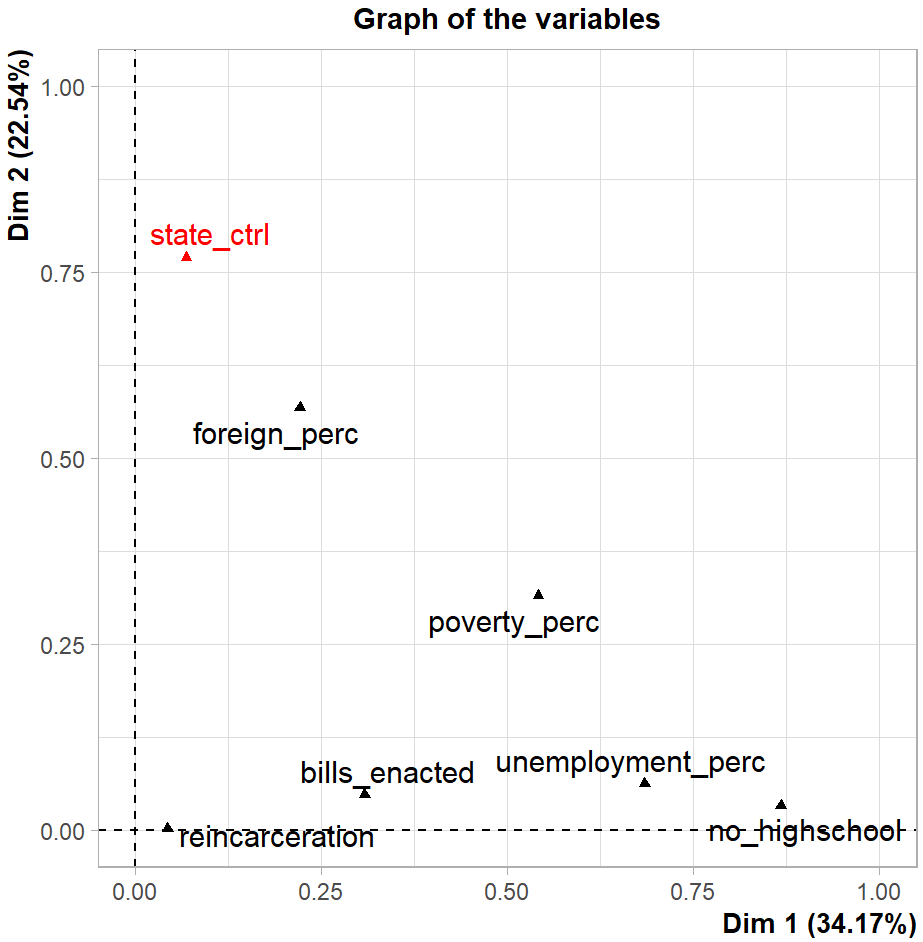
```
ctrl_vars <- ctrl_vars %>%
  mutate(white_perc = (white_total/total_population)) %>%
  mutate(blkAA_perc = (blkAA_total/total_population)) %>%
  mutate(AIAN_perc = (AIAN_total/total_population)) %>%
  mutate(asian_perc = (asian_total/total_population)) %>%
  mutate(PI_perc = (PI_total/total_population)) %>%
  mutate(other_perc = (other_total/total_population)) %>%
  mutate(twomore_perc = (twomore_total/total_population)) %>%
  mutate(foreign_perc = (foreign_born/population_2022)) %>%
  mutate(able_to_work = (population_2022/total_population)) %>%
  select(-white_total, -blkAA_total,-AIAN_total, -asian_total, -PI_total, -other_total, -twomore
_total, -total_population, -foreign_born, -population_2022)
```

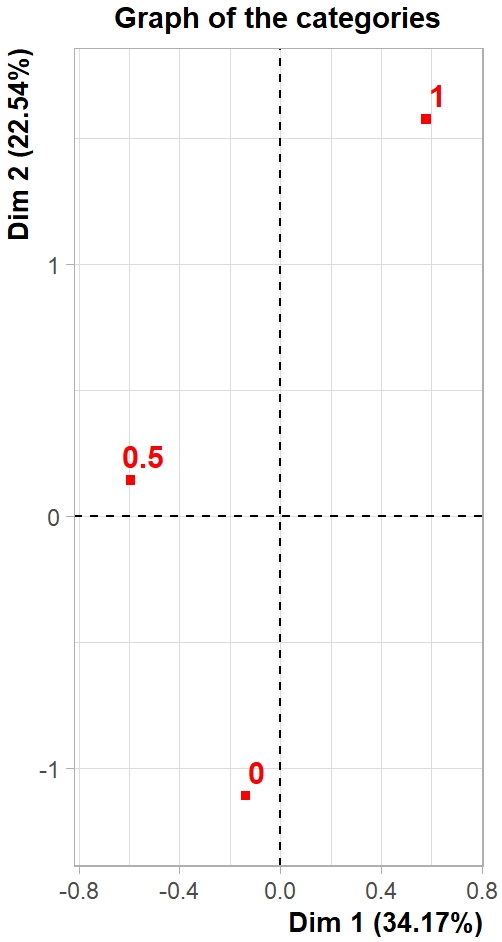
```
clean_vars <- ctrl_vars %>%
  mutate(
    leg_ctrl = as.factor(leg_ctrl),
    gov_party = as.factor(gov_party),
    state_ctrl = as.factor(state_ctrl),
    det_sentencing = as.factor(det_sentencing)
  )
```

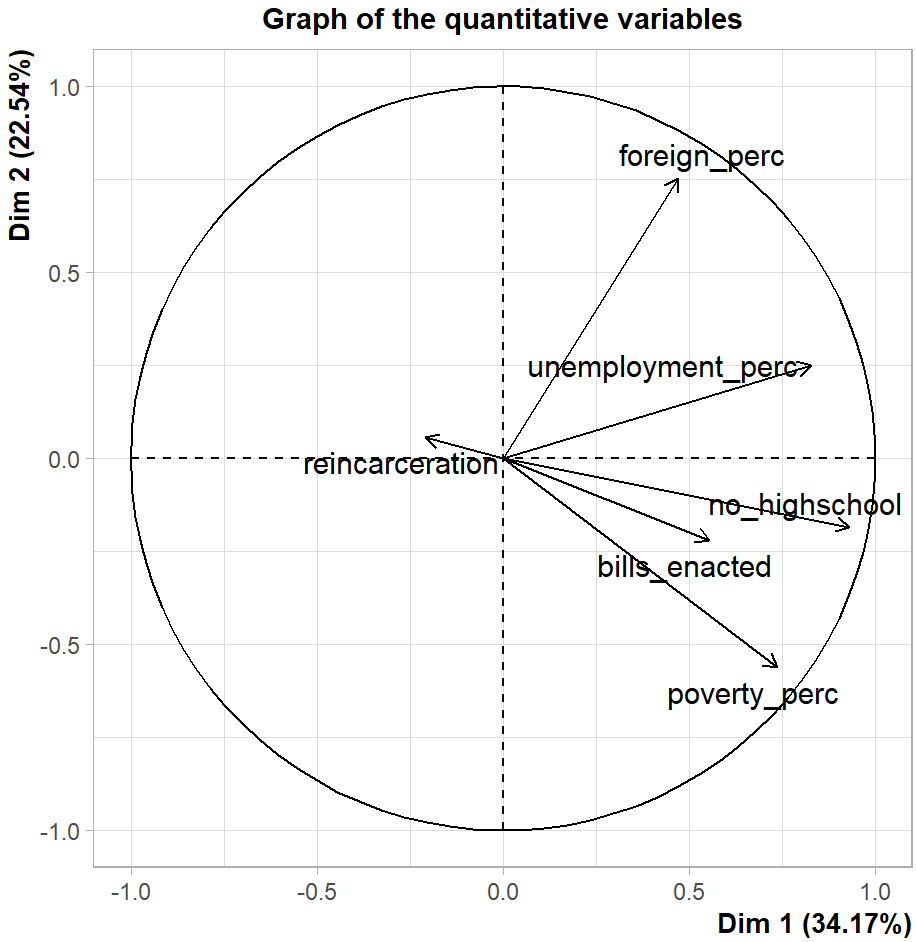
```
small_data <- clean_vars %>%
  select(foreign_perc, poverty_perc, no_highschool, unemployment_perc, state_ctrl, bills_enacte
d, reincarceration)

small_analysis <- FAMD(small_data, ncp = 3)
```


Individual factor map
Dim 2 (22.54%)
Dim 1 (34.17%)
1
0.5
0


Individual factor map
Dim 2 (22.54%)
Dim 1 (34.17%)


Graph of the variables
Dim 2 (22.54%)
1.00
0.75
0.50
0.25
0.00
state_ctrl
foreign_perc
poverty_perc
bills_enacted
unemployment_perc
reincarceration
no_highschool
0.00
0.25
0.50
0.75
1.00
Dim 1 (34.17%)


Graph of the categories
Dim 2 (22.54%)
1
1
0.5
0
-1
0
-0.8
-0.4
0.0
0.4
0.8
Dim 1 (34.17%)




```
summary(small_analysis)
```

```
##
## Call:
## FAMD(base = small_data, ncp = 3)
##
##
## Eigenvalues
##                        Dim.1  Dim.2  Dim.3
## Variance               2.733  1.803  1.154
## % of var.             34.168 22.542 14.422
## Cumulative % of var. 34.168 56.710 71.131
##
## Individuals (the 10 first)
##                       Dist    Dim.1    ctr   cos2    Dim.2    ctr   cos2
## 1                  |  2.260 |  1.076  0.847  0.227 | -1.813  3.644  0.643 |
## 2                  |  3.724 | -0.916  0.614  0.061 |  0.154  0.026  0.002 |
## 3                  |  2.345 |  0.587  0.252  0.063 |  0.304  0.103  0.017 |
## 4                  |  2.464 |  0.966  0.683  0.154 | -1.731  3.324  0.494 |
## 5                  |  4.414 |  3.139  7.208  0.506 |  2.348  6.116  0.283 |
## 6                  |  2.004 | -0.931  0.634  0.216 |  1.343  2.002  0.449 |
## 7                  |  2.396 |  0.255  0.048  0.011 |  2.019  4.521  0.710 |
## 8                  |  2.546 | -0.240  0.042  0.009 |  1.326  1.951  0.271 |
## 9                  |  2.557 |  0.972  0.692  0.145 |  0.301  0.100  0.014 |
## 10                 |  1.980 |  1.309  1.254  0.438 | -1.073  1.276  0.294 |
##                     Dim.3    ctr   cos2
## 1                   0.257  0.114  0.013 |
## 2                   3.052 16.146  0.672 |
## 3                  -0.750  0.976  0.102 |
## 4                   1.326  3.050  0.290 |
## 5                  -0.758  0.997  0.030 |
## 6                   0.024  0.001  0.000 |
## 7                   1.224  2.598  0.261 |
## 8                   2.030  7.140  0.635 |
## 9                  -0.940  1.532  0.135 |
## 10                 -0.613  0.651  0.096 |
##
## Continuous variables
##                        Dim.1    ctr   cos2    Dim.2    ctr   cos2    Dim.3    ctr
## foreign_perc       |  0.470  8.084  0.221 |  0.754 31.520  0.568 | -0.182  2.883
## poverty_perc       |  0.736 19.808  0.541 | -0.562 17.514  0.316 |  0.164  2.339
## no_highschool      |  0.931 31.733  0.867 | -0.184  1.880  0.034 | -0.056  0.271
## unemployment_perc  |  0.827 25.031  0.684 |  0.251  3.495  0.063 |  0.282  6.900
## bills_enacted      |  0.555 11.264  0.308 | -0.221  2.705  0.049 | -0.237  4.878
## reincarceration    | -0.208  1.577  0.043 |  0.055  0.170  0.003 |  0.867 65.129
##                      cos2
## foreign_perc        0.033 |
## poverty_perc        0.027 |
## no_highschool       0.003 |
## unemployment_perc   0.080 |
## bills_enacted       0.056 |
## reincarceration     0.751 |
##
## Categories
```

```
##                      Dim.1    ctr   cos2 v.test    Dim.2    ctr   cos2 v.test
## 0                  | -0.138  0.123  0.012 -0.563 | -1.110 18.172  0.795 -5.557 |
## 0.5                | -0.594  0.944  0.082 -1.257 |  0.142  0.125  0.005  0.371 |
## 1                  |  0.579  1.434  0.104  1.681 |  1.575 24.419  0.769  5.633 |
##                     Dim.3    ctr   cos2 v.test
## 0                   0.158  0.895  0.016  0.987 |
## 0.5                -0.952 13.613  0.211 -3.102 |
## 1                   0.359  3.091  0.040  1.603 |
```

```
factors <- small_analysis$ind$coord
```

```
pc_data <- as.data.frame(factors)

regression <- read.csv("regression.csv")

final_df <- cbind(pc_data, regression)

final_df <- final_df %>%
  select(-state) %>%
  mutate(scaled_fair_chance = fair_chance/1000) %>%
  select(-fair_chance)
```

```
library(betareg)
```

```
## Warning: package 'betareg' was built under R version 4.4.1
```

```
model <- betareg(scaled_fair_chance ~., data = final_df)
summary(model)
```

```
##
## Call:
## betareg(formula = scaled_fair_chance ~ ., data = final_df)
##
## Quantile residuals:
##     Min      1Q  Median      3Q     Max
## -1.4995 -0.7074 -0.3120  0.5511  3.3744
##
## Coefficients (mean model with logit link):
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -3.9848858  0.1977208 -20.154   <2e-16 ***
## Dim.1           -0.0163874  0.0421917  -0.388   0.6977
## Dim.2            0.0693046  0.0340130   2.038   0.0416 *
## Dim.3           -0.0622532  0.0420221  -1.481   0.1385
## tot_pos_index    0.0043455  0.2319998   0.019   0.9851
## neg_combo_index  0.0006890  0.0004442   1.551   0.1209
## licensing_rank  -0.0010029  0.0032442  -0.309   0.7572
##
## Phi coefficients (precision model with identity link):
##       Estimate Std. Error z value Pr(>|z|)
## (phi)   459.71      92.56   4.966 6.82e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Type of estimator: ML (maximum likelihood)
## Log-likelihood: 178.5 on 8 Df
## Pseudo R-squared: 0.1717
## Number of iterations: 61 (BFGS) + 7 (Fisher scoring)
```

```
model_neg <- betareg(scaled_fair_chance ~neg_combo_index, data = final_df)
summary(model_neg)
```

```
##
## Call:
## betareg(formula = scaled_fair_chance ~ neg_combo_index, data = final_df)
##
## Quantile residuals:
##     Min      1Q  Median      3Q     Max
## -1.6047 -0.6467 -0.2930  0.4953  3.8695
##
## Coefficients (mean model with logit link):
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -3.9552059  0.1309475 -30.205   <2e-16 ***
## neg_combo_index  0.0005717  0.0003047   1.876   0.0606 .
##
## Phi coefficients (precision model with identity link):
##       Estimate Std. Error z value Pr(>|z|)
## (phi)   396.82      79.98   4.962 6.99e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Type of estimator: ML (maximum likelihood)
## Log-likelihood: 174.9 on 3 Df
## Pseudo R-squared: 0.05212
## Number of iterations: 341 (BFGS) + 5 (Fisher scoring)
```

```
model_pos <- betareg(scaled_fair_chance ~tot_pos_index, data = final_df)
summary(model_pos)
```

```
##
## Call:
## betareg(formula = scaled_fair_chance ~ tot_pos_index, data = final_df)
##
## Quantile residuals:
##     Min      1Q  Median      3Q     Max
## -1.3981 -0.6898 -0.1694  0.4818  4.2223
##
## Coefficients (mean model with logit link):
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -3.8220     0.1024 -37.315   <2e-16 ***
## tot_pos_index   0.2236     0.2157   1.037      0.3
##
## Phi coefficients (precision model with identity link):
##       Estimate Std. Error z value Pr(>|z|)
## (phi)   382.29      77.07    4.96 7.04e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Type of estimator: ML (maximum likelihood)
## Log-likelihood:   174 on 3 Df
## Pseudo R-squared: 0.01916
## Number of iterations: 515 (BFGS) + 4 (Fisher scoring)
```