CODES\_growth\_analysis2.R

user

2020-05-15

rm(list = ls(all.names = TRUE))

library(Amelia)

## Loading required package: Rcpp

## ##
## ## Amelia II: Multiple Imputation
## ## (Version 1.7.6, built: 2019-11-24)
## ## Copyright (C) 2005-2020 James Honaker, Gary King and Matthew Blackwell
## ## Refer to http://gking.harvard.edu/amelia/ for more information
## ##

library(bestNormalize)
library(openxlsx)
library(lubridate)

##
## Attaching package: 'lubridate'

## The following object is masked from 'package:base':
##
## date

library(purrr)
library(dplyr)

##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:lubridate':
##
## intersect, setdiff, union

## The following objects are masked from 'package:stats':
##
## filter, lag

## The following objects are masked from 'package:base':
##
## intersect, setdiff, setequal, union

library(nlme)

##
## Attaching package: 'nlme'

## The following object is masked from 'package:dplyr':
##
## collapse

library(rsq)
library(lm.beta)
library(tidyverse)

## -- Attaching packages -------------------------------------------------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.0 v readr 1.3.1
## v tibble 2.1.3 v stringr 1.4.0
## v tidyr 1.0.2 v forcats 0.4.0

## -- Conflicts ----------------------------------------------------------------------------------- tidyverse\_conflicts() --
## x lubridate::as.difftime() masks base::as.difftime()
## x nlme::collapse() masks dplyr::collapse()
## x lubridate::date() masks base::date()
## x dplyr::filter() masks stats::filter()
## x lubridate::intersect() masks base::intersect()
## x dplyr::lag() masks stats::lag()
## x lubridate::setdiff() masks base::setdiff()
## x lubridate::union() masks base::union()

library(caret)

## Loading required package: lattice

##
## Attaching package: 'caret'

## The following object is masked from 'package:purrr':
##
## lift

library(leaps)
library(ggplot2)
library(reghelper)

##
## Attaching package: 'reghelper'

## The following object is masked from 'package:base':
##
## beta

library(ggpubr)

## Loading required package: magrittr

##
## Attaching package: 'magrittr'

## The following object is masked from 'package:tidyr':
##
## extract

## The following object is masked from 'package:purrr':
##
## set\_names

library(car)

## Loading required package: carData

##
## Attaching package: 'car'

## The following object is masked from 'package:dplyr':
##
## recode

## The following object is masked from 'package:purrr':
##
## some

library(missMDA)

## Registered S3 methods overwritten by 'lme4':
## method from
## cooks.distance.influence.merMod car
## influence.merMod car
## dfbeta.influence.merMod car
## dfbetas.influence.merMod car

library(psych)

##
## Attaching package: 'psych'

## The following object is masked from 'package:car':
##
## logit

## The following object is masked from 'package:reghelper':
##
## ICC

## The following objects are masked from 'package:ggplot2':
##
## %+%, alpha

setwd("C:/Users/user/Dropbox/BCG-COVID-19/Ver 3/TF\_SG")

CV19 <- read.xlsx("Table\_S1-9.xlsx", sheet=3)

DATA <- CV19[, which (colnames(CV19) %in% c(
"Country\_code",
"BCG.TB",
"m\_cases",
"r\_cases",
"popData2018.x",
#"Positive.cases.per.test",
"temp",
"BCG.type","Current.BCG.vaccination",
"prevalence\_overweight\_2012",
"Fraction\_urban\_population",
"Population\_ages\_65\_and\_above\_2018\_percent",
"TB.incidents.per.10.thosand.mean\_2000\_2018",
"TB.Incidents.100"))]

DATA\_COMPLETE <- na.omit(DATA)
dim(DATA\_COMPLETE)

## [1] 94 13

DATA\_COMPLETE$pop2018\_N <- predict(orderNorm(DATA\_COMPLETE$popData2018.x))
hist(DATA\_COMPLETE$pop2018\_N)



DATA\_COMPLETE$Ob\_N <- predict(orderNorm(DATA\_COMPLETE$prevalence\_overweight\_2012))

## Warning in orderNorm(DATA\_COMPLETE$prevalence\_overweight\_2012): Ties in data, Normal distribution not guaranteed

hist(DATA\_COMPLETE$Ob\_N)



DATA\_COMPLETE$TB\_Z <- scale(DATA\_COMPLETE$TB.incidents.per.10.thosand.mean\_2000\_2018)[,1]
hist(DATA\_COMPLETE$TB\_Z)



DATA\_COMPLETE$temp\_N <- predict(orderNorm(DATA\_COMPLETE$temp))

## Warning in orderNorm(DATA\_COMPLETE$temp): Ties in data, Normal distribution not guaranteed

hist(DATA\_COMPLETE$temp\_N)



DATA\_COMPLETE$frac\_N <- predict(orderNorm(DATA\_COMPLETE$Fraction\_urban\_population))
hist(DATA\_COMPLETE$frac\_N)



DATA\_COMPLETE$age\_N <- predict(orderNorm(DATA\_COMPLETE$Population\_ages\_65\_and\_above\_2018\_percent))
hist(DATA\_COMPLETE$age\_N)



DATA\_COMPLETE$Yy <- ifelse(DATA\_COMPLETE$BCG.TB == "Y\_y", 1, 0)
DATA\_COMPLETE$Yn <- ifelse(DATA\_COMPLETE$BCG.TB == "Y\_n", 1, 0)
DATA\_COMPLETE$Nn <- ifelse(DATA\_COMPLETE$BCG.TB == "N\_n", 1, 0)

DATA\_d <- CV19[, which (colnames(CV19) %in% c(
"Country\_code",
"BCG.TB",
"m\_deaths",
"r\_deaths",
"popData2018.x",
#"Positive.cases.per.test",
"temp",
"BCG.type","Current.BCG.vaccination",
"prevalence\_overweight\_2012",
"Fraction\_urban\_population",
"Population\_ages\_65\_and\_above\_2018\_percent",
"TB.incidents.per.10.thosand.mean\_2000\_2018",
"TB.Incidents.100"))]

DATA\_COMPLETE\_d <- na.omit(DATA\_d)
dim(DATA\_COMPLETE\_d)

## [1] 52 13

DATA\_COMPLETE\_d$pop2018\_N <- predict(orderNorm(DATA\_COMPLETE\_d$popData2018.x))

DATA\_COMPLETE\_d$Ob\_N <- predict(orderNorm(DATA\_COMPLETE\_d$prevalence\_overweight\_2012))

## Warning in orderNorm(DATA\_COMPLETE\_d$prevalence\_overweight\_2012): Ties in data, Normal distribution not guaranteed

DATA\_COMPLETE\_d$TB\_Z <- scale(DATA\_COMPLETE\_d$TB.incidents.per.10.thosand.mean\_2000\_2018)[,1]

DATA\_COMPLETE\_d$temp\_N <- predict(orderNorm(DATA\_COMPLETE\_d$temp))

## Warning in orderNorm(DATA\_COMPLETE\_d$temp): Ties in data, Normal distribution not guaranteed

DATA\_COMPLETE\_d$frac\_N <- predict(orderNorm(DATA\_COMPLETE\_d$Fraction\_urban\_population))

DATA\_COMPLETE\_d$age\_N <- predict(orderNorm(DATA\_COMPLETE\_d$Population\_ages\_65\_and\_above\_2018\_percent))

###Analysis###

DATA\_COMPLETE$m\_cases\_Z <- scale(DATA\_COMPLETE$m\_cases)[,1]
summary(DATA\_COMPLETE$r\_cases)

## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.7013 0.9498 0.9717 0.9618 0.9872 0.9977

cases\_gr\_lm <- lm(m\_cases ~ Current.BCG.vaccination + TB\_Z +
Ob\_N + pop2018\_N + temp\_N + age\_N + frac\_N
, weights=DATA\_COMPLETE$r\_cases, data = DATA\_COMPLETE)

summary(cases\_gr\_lm, corr=T)

##
## Call:
## lm(formula = m\_cases ~ Current.BCG.vaccination + TB\_Z + Ob\_N +
## pop2018\_N + temp\_N + age\_N + frac\_N, data = DATA\_COMPLETE,
## weights = DATA\_COMPLETE$r\_cases)
##
## Weighted Residuals:
## Min 1Q Median 3Q Max
## -0.053457 -0.013445 -0.000435 0.011079 0.048321
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.081005 0.005778 14.019 < 2e-16 \*\*\*
## Current.BCG.vaccinationY -0.020037 0.006775 -2.957 0.0040 \*\*
## TB\_Z -0.001822 0.003184 -0.572 0.5686
## Ob\_N 0.006004 0.003869 1.552 0.1244
## pop2018\_N 0.015869 0.002571 6.172 2.15e-08 \*\*\*
## temp\_N -0.005626 0.003213 -1.751 0.0835 .
## age\_N 0.006335 0.003507 1.806 0.0744 .
## frac\_N -0.001945 0.003771 -0.516 0.6074
## ---
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0225 on 86 degrees of freedom
## Multiple R-squared: 0.4977, Adjusted R-squared: 0.4568
## F-statistic: 12.17 on 7 and 86 DF, p-value: 1.033e-10
##
## Correlation of Coefficients:
## (Intercept) Current.BCG.vaccinationY TB\_Z Ob\_N
## Current.BCG.vaccinationY -0.91
## TB\_Z 0.05 -0.05
## Ob\_N -0.07 0.08 0.32
## pop2018\_N -0.13 0.14 -0.22 0.05
## temp\_N 0.08 -0.08 -0.02 0.24
## age\_N -0.25 0.27 0.23 0.33
## frac\_N -0.10 0.11 0.05 -0.64
## pop2018\_N temp\_N age\_N
## Current.BCG.vaccinationY
## TB\_Z
## Ob\_N
## pop2018\_N
## temp\_N -0.13
## age\_N 0.05 0.51
## frac\_N -0.04 -0.16 -0.30

shapiro.test(resid(cases\_gr\_lm)) #List of residuals

##
## Shapiro-Wilk normality test
##
## data: resid(cases\_gr\_lm)
## W = 0.98712, p-value = 0.4913

plot(density(resid(cases\_gr\_lm))) #A density plot



rsq.partial(cases\_gr\_lm, adj=TRUE)

## $adjustment
## [1] TRUE
##
## $variable
## [1] "Current.BCG.vaccination" "TB\_Z"
## [3] "Ob\_N" "pop2018\_N"
## [5] "temp\_N" "age\_N"
## [7] "frac\_N"
##
## $partial.rsq
## [1] 0.081759778 -0.007788573 0.015918551 0.298919510 0.023196920
## [6] 0.025349039 -0.008509936

lm.beta(cases\_gr\_lm)

##
## Call:
## lm(formula = m\_cases ~ Current.BCG.vaccination + TB\_Z + Ob\_N +
## pop2018\_N + temp\_N + age\_N + frac\_N, data = DATA\_COMPLETE,
## weights = DATA\_COMPLETE$r\_cases)
##
## Standardized Coefficients::
## (Intercept) Current.BCG.vaccinationY TB\_Z
## 0.00000000 -0.27043096 -0.05873103
## Ob\_N pop2018\_N temp\_N
## 0.19319428 0.51073268 -0.18104548
## age\_N frac\_N
## 0.20389760 -0.06258373

Anova(cases\_gr\_lm)

## Anova Table (Type II tests)
##
## Response: m\_cases
## Sum Sq Df F value Pr(>F)
## Current.BCG.vaccination 0.004427 1 8.7464 0.004004 \*\*
## TB\_Z 0.000166 1 0.3276 0.568551
## Ob\_N 0.001218 1 2.4073 0.124442
## pop2018\_N 0.019282 1 38.0942 2.145e-08 \*\*\*
## temp\_N 0.001552 1 3.0661 0.083509 .
## age\_N 0.001651 1 3.2627 0.074369 .
## frac\_N 0.000135 1 0.2659 0.607430
## Residuals 0.043530 86
## ---
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

models\_cases\_gr <- regsubsets(m\_cases ~ Current.BCG.vaccination + TB\_Z +
Ob\_N + temp\_N + age\_N + frac\_N +
pop2018\_N,
weights=DATA\_COMPLETE$r\_cases, data = DATA\_COMPLETE, nvmax = 6)

summary(models\_cases\_gr)

## Subset selection object
## Call: regsubsets.formula(m\_cases ~ Current.BCG.vaccination + TB\_Z +
## Ob\_N + temp\_N + age\_N + frac\_N + pop2018\_N, weights = DATA\_COMPLETE$r\_cases,
## data = DATA\_COMPLETE, nvmax = 6)
## 7 Variables (and intercept)
## Forced in Forced out
## Current.BCG.vaccinationY FALSE FALSE
## TB\_Z FALSE FALSE
## Ob\_N FALSE FALSE
## temp\_N FALSE FALSE
## age\_N FALSE FALSE
## frac\_N FALSE FALSE
## pop2018\_N FALSE FALSE
## 1 subsets of each size up to 6
## Selection Algorithm: exhaustive
## Current.BCG.vaccinationY TB\_Z Ob\_N temp\_N age\_N frac\_N pop2018\_N
## 1 ( 1 ) "\*" " " " " " " " " " " " "
## 2 ( 1 ) "\*" " " " " " " " " " " "\*"
## 3 ( 1 ) "\*" " " " " "\*" " " " " "\*"
## 4 ( 1 ) "\*" " " "\*" " " "\*" " " "\*"
## 5 ( 1 ) "\*" " " "\*" "\*" "\*" " " "\*"
## 6 ( 1 ) "\*" "\*" "\*" "\*" "\*" " " "\*"

res.sum <- summary(models\_cases\_gr)
data.frame(
 Adj.R2 = which.max(res.sum$adjr2),
 CP = which.min(res.sum$cp),
 BIC = which.min(res.sum$bic)
)

## Adj.R2 CP BIC
## 1 5 5 3

coef(models\_cases\_gr, 1:5)

## [[1]]
## (Intercept) Current.BCG.vaccinationY
## 0.09291005 -0.03513929
##
## [[2]]
## (Intercept) Current.BCG.vaccinationY pop2018\_N
## 0.09345174 -0.03588752 0.01177396
##
## [[3]]
## (Intercept) Current.BCG.vaccinationY temp\_N
## 0.08672353 -0.02733035 -0.01008537
## pop2018\_N
## 0.01430956
##
## [[4]]
## (Intercept) Current.BCG.vaccinationY Ob\_N
## 0.081536800 -0.020683199 0.006510296
## age\_N pop2018\_N
## 0.009339825 0.014814382
##
## [[5]]
## (Intercept) Current.BCG.vaccinationY Ob\_N
## 0.080875340 -0.019859529 0.005473334
## temp\_N age\_N pop2018\_N
## -0.005910937 0.006249599 0.015505958

vcov(models\_cases\_gr, 5)

## (Intercept) Current.BCG.vaccinationY Ob\_N
## (Intercept) 3.242274e-05 -3.459063e-05 -3.477133e-06
## Current.BCG.vaccinationY -3.459063e-05 4.445630e-05 4.421166e-06
## Ob\_N -3.477133e-06 4.421166e-06 6.797880e-06
## temp\_N 1.107135e-06 -1.378638e-06 1.735640e-06
## age\_N -5.843202e-06 7.492268e-06 7.137250e-07
## pop2018\_N -1.812391e-06 2.281237e-06 9.886194e-07
## temp\_N age\_N pop2018\_N
## (Intercept) 1.107135e-06 -5.843202e-06 -1.812391e-06
## Current.BCG.vaccinationY -1.378638e-06 7.492268e-06 2.281237e-06
## Ob\_N 1.735640e-06 7.137250e-07 9.886194e-07
## temp\_N 9.893575e-06 5.172342e-06 -1.157541e-06
## age\_N 5.172342e-06 1.027255e-05 8.233142e-07
## pop2018\_N -1.157541e-06 8.233142e-07 6.187012e-06

cases\_gr\_lm\_3 <- lm(m\_cases ~ Current.BCG.vaccination + pop2018\_N + temp\_N,
 weights=DATA\_COMPLETE$r\_cases, data = DATA\_COMPLETE)

summary(cases\_gr\_lm\_3, corr=T)

##
## Call:
## lm(formula = m\_cases ~ Current.BCG.vaccination + pop2018\_N +
## temp\_N, data = DATA\_COMPLETE, weights = DATA\_COMPLETE$r\_cases)
##
## Weighted Residuals:
## Min 1Q Median 3Q Max
## -0.046102 -0.015023 0.001199 0.011668 0.053827
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.086724 0.005423 15.990 < 2e-16 \*\*\*
## Current.BCG.vaccinationY -0.027330 0.006246 -4.376 3.25e-05 \*\*\*
## pop2018\_N 0.014310 0.002522 5.673 1.69e-07 \*\*\*
## temp\_N -0.010085 0.002728 -3.697 0.000375 \*\*\*
## ---
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02299 on 90 degrees of freedom
## Multiple R-squared: 0.4509, Adjusted R-squared: 0.4326
## F-statistic: 24.64 on 3 and 90 DF, p-value: 9.971e-12
##
## Correlation of Coefficients:
## (Intercept) Current.BCG.vaccinationY pop2018\_N
## Current.BCG.vaccinationY -0.90
## pop2018\_N -0.07 0.08
## temp\_N 0.34 -0.37 -0.27

rsq.partial(cases\_gr\_lm\_3, adj=TRUE)

## $adjustment
## [1] TRUE
##
## $variable
## [1] "Current.BCG.vaccination" "pop2018\_N"
## [3] "temp\_N"
##
## $partial.rsq
## [1] 0.1662712 0.2551993 0.1221962

lm.beta(cases\_gr\_lm\_3)

##
## Call:
## lm(formula = m\_cases ~ Current.BCG.vaccination + pop2018\_N +
## temp\_N, data = DATA\_COMPLETE, weights = DATA\_COMPLETE$r\_cases)
##
## Standardized Coefficients::
## (Intercept) Current.BCG.vaccinationY pop2018\_N
## 0.0000000 -0.3688723 0.4605376
## temp\_N
## -0.3245729

#example reported in the Discussion
#confint(cases\_gr\_lm\_3)

log\_f1<-summary(cases\_gr\_lm\_3)$coef[1,1]
log\_f2<-summary(cases\_gr\_lm\_3)$coef[2,1]+summary(cases\_gr\_lm\_3)$coef[1,1]

100\*(10^log\_f1)^15

## [1] 1999.185

100\*(10^log\_f2)^15

## [1] 777.8532

100\*(10^log\_f1)^30

## [1] 39967.39

100\*(10^log\_f2)^30

## [1] 6050.556

#deaths

deaths\_gr\_lm <- lm(m\_deaths~ Current.BCG.vaccination + TB\_Z +
Ob\_N + pop2018\_N + temp\_N + age\_N + frac\_N
, weights=DATA\_COMPLETE$r\_deaths, data = DATA\_COMPLETE\_d)

summary(deaths\_gr\_lm, corr=T)

##
## Call:
## lm(formula = m\_deaths ~ Current.BCG.vaccination + TB\_Z + Ob\_N +
## pop2018\_N + temp\_N + age\_N + frac\_N, data = DATA\_COMPLETE\_d,
## weights = DATA\_COMPLETE$r\_deaths)
##
## Residuals:
## Min 1Q Median 3Q Max
## -0.056076 -0.011394 -0.001503 0.013328 0.044205
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.0789875 0.0060749 13.002 < 2e-16 \*\*\*
## Current.BCG.vaccinationY -0.0276379 0.0082871 -3.335 0.00174 \*\*
## TB\_Z -0.0041246 0.0053288 -0.774 0.44305
## Ob\_N 0.0007251 0.0046380 0.156 0.87648
## pop2018\_N 0.0168782 0.0034310 4.919 1.26e-05 \*\*\*
## temp\_N -0.0003093 0.0042450 -0.073 0.94225
## age\_N 0.0049831 0.0048363 1.030 0.30847
## frac\_N -0.0032970 0.0043840 -0.752 0.45602
## ---
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02193 on 44 degrees of freedom
## Multiple R-squared: 0.4841, Adjusted R-squared: 0.402
## F-statistic: 5.898 on 7 and 44 DF, p-value: 7.192e-05
##
## Correlation of Coefficients:
## (Intercept) Current.BCG.vaccinationY TB\_Z Ob\_N
## Current.BCG.vaccinationY -0.87
## TB\_Z -0.11 0.12
## Ob\_N -0.12 0.14 0.52
## pop2018\_N 0.07 -0.08 -0.25 -0.03
## temp\_N 0.25 -0.28 -0.32 0.02
## age\_N -0.26 0.30 0.33 0.50
## frac\_N -0.17 0.20 0.21 -0.36
## pop2018\_N temp\_N age\_N
## Current.BCG.vaccinationY
## TB\_Z
## Ob\_N
## pop2018\_N
## temp\_N -0.02
## age\_N 0.08 0.28
## frac\_N -0.13 -0.29 -0.33

shapiro.test(resid(deaths\_gr\_lm)) #List of residuals

##
## Shapiro-Wilk normality test
##
## data: resid(deaths\_gr\_lm)
## W = 0.9756, p-value = 0.3592

plot(density(resid(deaths\_gr\_lm))) #A density plot



rsq.partial(deaths\_gr\_lm, adj=TRUE)

## $adjustment
## [1] TRUE
##
## $variable
## [1] "Current.BCG.vaccination" "TB\_Z"
## [3] "Ob\_N" "pop2018\_N"
## [5] "temp\_N" "age\_N"
## [7] "frac\_N"
##
## $partial.rsq
## [1] 0.183638917 -0.008988596 -0.022159428 0.340178549 -0.022603885
## [6] 0.001367733 -0.009747703

lm.beta(deaths\_gr\_lm)

##
## Call:
## lm(formula = m\_deaths ~ Current.BCG.vaccination + TB\_Z + Ob\_N +
## pop2018\_N + temp\_N + age\_N + frac\_N, data = DATA\_COMPLETE\_d,
## weights = DATA\_COMPLETE$r\_deaths)
##
## Standardized Coefficients::
## (Intercept) Current.BCG.vaccinationY TB\_Z
## 0.00000000 -0.47392827 -0.14546056
## Ob\_N pop2018\_N temp\_N
## 0.02550513 0.59374152 -0.01088053
## age\_N frac\_N
## 0.17529518 -0.11598170

Anova(deaths\_gr\_lm)

## Anova Table (Type II tests)
##
## Response: m\_deaths
## Sum Sq Df F value Pr(>F)
## Current.BCG.vaccination 0.0053479 1 11.1227 0.001739 \*\*
## TB\_Z 0.0002881 1 0.5991 0.443055
## Ob\_N 0.0000118 1 0.0244 0.876477
## pop2018\_N 0.0116358 1 24.2003 1.258e-05 \*\*\*
## temp\_N 0.0000026 1 0.0053 0.942245
## age\_N 0.0005104 1 1.0616 0.308474
## frac\_N 0.0002719 1 0.5656 0.456022
## Residuals 0.0211557 44
## ---
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

models\_deaths\_gr <- regsubsets(m\_deaths ~ Current.BCG.vaccination + TB\_Z +
Ob\_N + temp\_N + age\_N + frac\_N +
pop2018\_N,
weights=DATA\_COMPLETE\_d$r\_deaths, data = DATA\_COMPLETE\_d, nvmax = 6)

summary(models\_deaths\_gr)

## Subset selection object
## Call: regsubsets.formula(m\_deaths ~ Current.BCG.vaccination + TB\_Z +
## Ob\_N + temp\_N + age\_N + frac\_N + pop2018\_N, weights = DATA\_COMPLETE\_d$r\_deaths,
## data = DATA\_COMPLETE\_d, nvmax = 6)
## 7 Variables (and intercept)
## Forced in Forced out
## Current.BCG.vaccinationY FALSE FALSE
## TB\_Z FALSE FALSE
## Ob\_N FALSE FALSE
## temp\_N FALSE FALSE
## age\_N FALSE FALSE
## frac\_N FALSE FALSE
## pop2018\_N FALSE FALSE
## 1 subsets of each size up to 6
## Selection Algorithm: exhaustive
## Current.BCG.vaccinationY TB\_Z Ob\_N temp\_N age\_N frac\_N pop2018\_N
## 1 ( 1 ) "\*" " " " " " " " " " " " "
## 2 ( 1 ) "\*" " " " " " " " " " " "\*"
## 3 ( 1 ) "\*" " " " " " " "\*" " " "\*"
## 4 ( 1 ) "\*" "\*" " " " " "\*" " " "\*"
## 5 ( 1 ) "\*" "\*" " " " " "\*" "\*" "\*"
## 6 ( 1 ) "\*" "\*" "\*" " " "\*" "\*" "\*"

res.sum <- summary(models\_deaths\_gr)
data.frame(
 Adj.R2 = which.max(res.sum$adjr2),
 CP = which.min(res.sum$cp),
 BIC = which.min(res.sum$bic)
)

## Adj.R2 CP BIC
## 1 3 3 2

coef(models\_deaths\_gr, 1:5)

## [[1]]
## (Intercept) Current.BCG.vaccinationY
## 0.07828212 -0.02615906
##
## [[2]]
## (Intercept) Current.BCG.vaccinationY pop2018\_N
## 0.08306417 -0.03382166 0.01467054
##
## [[3]]
## (Intercept) Current.BCG.vaccinationY age\_N
## 0.079545670 -0.028275600 0.005367558
## pop2018\_N
## 0.015659641
##
## [[4]]
## (Intercept) Current.BCG.vaccinationY TB\_Z
## 0.078832174 -0.027156358 -0.003251691
## age\_N pop2018\_N
## 0.004267351 0.016438312
##
## [[5]]
## (Intercept) Current.BCG.vaccinationY TB\_Z
## 0.079631742 -0.028407906 -0.004706004
## age\_N frac\_N pop2018\_N
## 0.004594937 -0.003075282 0.016876374

vcov(models\_deaths\_gr, 5)

## (Intercept) Current.BCG.vaccinationY TB\_Z
## (Intercept) 3.235501e-05 -3.721120e-05 1.158637e-06
## Current.BCG.vaccinationY -3.721120e-05 5.869948e-05 -1.823473e-06
## TB\_Z 1.158637e-06 -1.823473e-06 1.646414e-05
## age\_N -7.444574e-06 1.174638e-05 3.772606e-06
## frac\_N -3.708919e-06 5.805503e-06 6.746060e-06
## pop2018\_N 1.331668e-06 -2.180187e-06 -4.139606e-06
## age\_N frac\_N pop2018\_N
## (Intercept) -7.444574e-06 -3.708919e-06 1.331668e-06
## Current.BCG.vaccinationY 1.174638e-05 5.805503e-06 -2.180187e-06
## TB\_Z 3.772606e-06 6.746060e-06 -4.139606e-06
## age\_N 1.514442e-05 -1.519559e-06 1.621843e-06
## frac\_N -1.519559e-06 1.426518e-05 -2.032018e-06
## pop2018\_N 1.621843e-06 -2.032018e-06 1.110111e-05

deaths\_gr\_lm\_2 <- lm(m\_deaths ~ Current.BCG.vaccination + pop2018\_N
,weights=DATA\_COMPLETE\_d$r\_deaths, data = DATA\_COMPLETE\_d)

summary(deaths\_gr\_lm\_2, corr=T)

##
## Call:
## lm(formula = m\_deaths ~ Current.BCG.vaccination + pop2018\_N,
## data = DATA\_COMPLETE\_d, weights = DATA\_COMPLETE\_d$r\_deaths)
##
## Weighted Residuals:
## Min 1Q Median 3Q Max
## -0.05414 -0.01282 -0.00106 0.01454 0.04566
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.083064 0.005027 16.523 < 2e-16 \*\*\*
## Current.BCG.vaccinationY -0.033822 0.006394 -5.289 2.85e-06 \*\*\*
## pop2018\_N 0.014671 0.003116 4.708 2.09e-05 \*\*\*
## ---
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.0212 on 49 degrees of freedom
## Multiple R-squared: 0.4498, Adjusted R-squared: 0.4274
## F-statistic: 20.03 on 2 and 49 DF, p-value: 4.391e-07
##
## Correlation of Coefficients:
## (Intercept) Current.BCG.vaccinationY
## Current.BCG.vaccinationY -0.81
## pop2018\_N 0.20 -0.25

rsq.partial(deaths\_gr\_lm\_2, adj=TRUE)

## $adjustment
## [1] TRUE
##
## $variable
## [1] "Current.BCG.vaccination" "pop2018\_N"
##
## $partial.rsq
## [1] 0.3504684 0.2973944

lm.beta(deaths\_gr\_lm\_2)

##
## Call:
## lm(formula = m\_deaths ~ Current.BCG.vaccination + pop2018\_N,
## data = DATA\_COMPLETE\_d, weights = DATA\_COMPLETE\_d$r\_deaths)
##
## Standardized Coefficients::
## (Intercept) Current.BCG.vaccinationY pop2018\_N
## 0.0000000 -0.5799654 0.5160788

log\_f1\_d<-summary(deaths\_gr\_lm\_2)$coef[1,1]
log\_f2\_d<-summary(deaths\_gr\_lm\_2)$coef[2,1]+summary(deaths\_gr\_lm\_2)$coef[1,1]

100\*(10^log\_f1\_d)^15

## [1] 1761.824

100\*(10^log\_f2\_d)^15

## [1] 547.8196

100\*(10^log\_f1\_d)^30

## [1] 31040.24

100\*(10^log\_f2\_d)^30

## [1] 3001.064

#####
#PCA#

cases\_PCA\_lm <- lm(m\_cases ~ Current.BCG.vaccination + pop2018\_N + TB\_Z +
RC1 +
RC2 + RC3
,
weights=CV19$r\_cases, data = CV19)

summary(cases\_PCA\_lm, corr=T)

##
## Call:
## lm(formula = m\_cases ~ Current.BCG.vaccination + pop2018\_N +
## TB\_Z + RC1 + RC2 + RC3, data = CV19, weights = CV19$r\_cases)
##
## Weighted Residuals:
## Min 1Q Median 3Q Max
## -0.065868 -0.013574 -0.000805 0.011456 0.050092
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.0685336 0.0051855 13.216 < 2e-16 \*\*\*
## Current.BCG.vaccinationY -0.0067856 0.0060741 -1.117 0.2665
## pop2018\_N 0.0086871 0.0043611 1.992 0.0490 \*
## TB\_Z -0.0006346 0.0024379 -0.260 0.7952
## RC1 0.0170544 0.0026634 6.403 4.6e-09 \*\*\*
## RC2 0.0091924 0.0041738 2.202 0.0299 \*
## RC3 -0.0040546 0.0019998 -2.028 0.0452 \*
## ---
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02 on 103 degrees of freedom
## Multiple R-squared: 0.6139, Adjusted R-squared: 0.5914
## F-statistic: 27.29 on 6 and 103 DF, p-value: < 2.2e-16
##
## Correlation of Coefficients:
## (Intercept) Current.BCG.vaccinationY pop2018\_N TB\_Z
## Current.BCG.vaccinationY -0.93
## pop2018\_N 0.16 -0.18
## TB\_Z 0.01 -0.01 -0.27
## RC1 -0.47 0.50 0.05 0.37
## RC2 -0.19 0.21 -0.88 0.14
## RC3 -0.06 0.06 -0.16 -0.11
## RC1 RC2
## Current.BCG.vaccinationY
## pop2018\_N
## TB\_Z
## RC1
## RC2 -0.06
## RC3 -0.05 0.15

shapiro.test(resid(cases\_PCA\_lm)) #List of residuals

##
## Shapiro-Wilk normality test
##
## data: resid(cases\_PCA\_lm)
## W = 0.98387, p-value = 0.207

plot(density(resid(cases\_PCA\_lm))) #A density plot



rsq.partial(cases\_PCA\_lm, adj=TRUE)

## $adjustment
## [1] TRUE
##
## $variable
## [1] "Current.BCG.vaccination" "pop2018\_N"
## [3] "TB\_Z" "RC1"
## [5] "RC2" "RC3"
##
## $partial.rsq
## [1] 0.002379050 0.027744889 -0.009044981 0.277785837 0.035703537
## [6] 0.029043498

lm.beta(cases\_PCA\_lm)

##
## Call:
## lm(formula = m\_cases ~ Current.BCG.vaccination + pop2018\_N +
## TB\_Z + RC1 + RC2 + RC3, data = CV19, weights = CV19$r\_cases)
##
## Standardized Coefficients::
## (Intercept) Current.BCG.vaccinationY pop2018\_N
## 0.00000000 -0.08701214 0.27234122
## TB\_Z RC1 RC2
## -0.01991867 0.53532097 0.28853994
## RC3
## -0.12726854

kruskal.test(RC1 ~ Current.BCG.vaccination, data = CV19)

##
## Kruskal-Wallis rank sum test
##
## data: RC1 by Current.BCG.vaccination
## Kruskal-Wallis chi-squared = 43.811, df = 1, p-value = 3.618e-11

dataset\_1 <- na.omit(CV19[
, which (colnames(CV19) %in% c(
"m\_deaths","r\_deaths","Current.BCG.vaccination","TB\_Z",
"pop2018\_N","RC1","RC2","RC3"))])

dataset\_1<-na.omit(dataset\_1)

deaths\_PCA\_lm <- lm(m\_deaths ~ Current.BCG.vaccination + pop2018\_N + TB\_Z +
RC1 +
RC2 + RC3
,
weights=dataset\_1$r\_deaths, data = dataset\_1)

summary(deaths\_PCA\_lm, corr=T)

##
## Call:
## lm(formula = m\_deaths ~ Current.BCG.vaccination + pop2018\_N +
## TB\_Z + RC1 + RC2 + RC3, data = dataset\_1, weights = dataset\_1$r\_deaths)
##
## Weighted Residuals:
## Min 1Q Median 3Q Max
## -0.057673 -0.012646 0.001347 0.013572 0.039870
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.065476 0.008678 7.545 8.46e-10 \*\*\*
## Current.BCG.vaccinationY -0.024369 0.009274 -2.628 0.0114 \*
## pop2018\_N 0.005228 0.008096 0.646 0.5214
## TB\_Z -0.006319 0.005356 -1.180 0.2436
## RC1 0.003482 0.005911 0.589 0.5585
## RC2 0.011622 0.007359 1.579 0.1206
## RC3 0.004962 0.003468 1.431 0.1587
## ---
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.02139 on 50 degrees of freedom
## Multiple R-squared: 0.5049, Adjusted R-squared: 0.4455
## F-statistic: 8.499 on 6 and 50 DF, p-value: 2.244e-06
##
## Correlation of Coefficients:
## (Intercept) Current.BCG.vaccinationY pop2018\_N TB\_Z
## Current.BCG.vaccinationY -0.91
## pop2018\_N 0.30 -0.34
## TB\_Z -0.08 0.10 -0.44
## RC1 -0.73 0.68 -0.30 0.44
## RC2 -0.36 0.31 -0.91 0.35
## RC3 -0.16 0.15 -0.19 -0.26
## RC1 RC2
## Current.BCG.vaccinationY
## pop2018\_N
## TB\_Z
## RC1
## RC2 0.29
## RC3 0.06 0.17

shapiro.test(resid(deaths\_PCA\_lm)) #List of residuals

##
## Shapiro-Wilk normality test
##
## data: resid(deaths\_PCA\_lm)
## W = 0.96261, p-value = 0.07548

plot(density(resid(deaths\_PCA\_lm))) #A density plot



rsq.partial(deaths\_PCA\_lm, adj=TRUE)

## $adjustment
## [1] TRUE
##
## $variable
## [1] "Current.BCG.vaccination" "pop2018\_N"
## [3] "TB\_Z" "RC1"
## [5] "RC2" "RC3"
##
## $partial.rsq
## [1] 0.103771524 -0.011562909 0.007630868 -0.012970443 0.028468282
## [6] 0.020116255

lm.beta(deaths\_PCA\_lm)

##
## Call:
## lm(formula = m\_deaths ~ Current.BCG.vaccination + pop2018\_N +
## TB\_Z + RC1 + RC2 + RC3, data = dataset\_1, weights = dataset\_1$r\_deaths)
##
## Standardized Coefficients::
## (Intercept) Current.BCG.vaccinationY pop2018\_N
## 0.00000000 -0.39787416 0.17014828
## TB\_Z RC1 RC2
## -0.16968330 0.09572056 0.38469102
## RC3
## 0.15945227