

Regression Analysis at PHU Level (includes recent health indicators)

Sofia Bahmutsky

17/06/2020

Data

```
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
```

```
# load the data
```

```
data <- read.csv("/Users/Sofia/Desktop/data-599-capstone-statistics-canada/PHU_analysis/data/econ.csv")
```

```
str(data)
```

```
## 'data.frame': 34 obs. of 86 variables:
```

```
## $ X
```

```
: int 0 1 2 3 4 5 6
```

```
## $ Reporting_PHU
```

```
: Factor w/ 34 levels
```

```
## $ X..of.owner.households.spending.30..or.more.of.its.income.on.shelter.costs
```

```
: num 11.2 14.8 12
```

```
## $ X..of.tenant.households.spending.30..or.more.of.its.income.on.shelter.costs
```

```
: num 44.6 45.4 41
```

```
## $ Apartment.in.a.building.that.has.fewer.than.five.storeys
```

```
: num 5.63 3.42 4.
```

```
## $ Apartment.in.a.building.that.has.five.or.more.storeys
```

```
: num 2.295 2.886
```

```
## $ Average.household.size
```

```
: num 2.2 2.5 2.3
```

```
## $ Bachelor.s.degree
```

```
: num 9442 9602 60
```

```
## $ Bicycle
```

```
: num 0.279 0.378
```

```
## $ Car..truck..van...as.a.driver
```

```
: num 32.5 38.6 35
```

```
## $ Car..truck..van...as.a.passenger
```

```
: num 2.59 3.32 2.
```

```
## $ Earned.doctorate
```

```
: num 297.5 302.5
```

```
## $ Employment.rate
```

```
: num 49 60.8 55.6
```

```
## $ External.migrants
```

```
: num 520 842 612
```

```
## $ Interprovincial.migrants
```

```
: num 0.676 0.528
```

```
## $ Intraprovincial.migrants
```

```
: num 5.95 8.72 4.
```

```
## $ Master.s.degree
```

```
: num 1635 2195 11
```

```
## $ Median.after.tax.income.of.households.in.2015....
```

```
: num 52908 61093
```

```
## $ Median.age.of.the.population
```

```
: num 49.1 41.7 45
```

```
## $ Movable.dwelling
```

```
: num 0.3449 0.014
```

```
## $ No.certificate..diploma.or.degree
```

```
: num 11.94 11.76
```

```
## $ Not.in.the.labour.force
```

```
: num 43175 38230
```

## \$ Other.attached.dwelling	: num	11.47	9.99	8
## \$ Other.method	: num	0.473	0.378	0
## \$ Population..2016	: num	113084	13494	0
## \$ Public.transit	: num	1.163	1.423	0
## \$ Single.detached.house	: num	30.7	26.2	32
## \$ Unemployment.rate	: num	10.4	6.2	7.5
## \$ Walked	: num	2.15	1.73	2.1
## \$ FEMALE	: int	13	68	29
## \$ MALE	: int	10	59	120
## \$ TRANSGENDER	: int	0	0	0
## \$ UNKNOWN	: int	0	0	0
## \$ OTHER	: int	0	0	0
## \$ CONTACT	: int	3	61	130
## \$ NEITHER	: int	9	44	8
## \$ TRAVEL.RELATED	: int	8	18	11
## \$ NOT.RESOLVED	: int	2	7	2
## \$ RESOLVED	: int	21	115	146
## \$ FATAL	: int	0	5	1
## \$ youth	: int	1	8	1
## \$ X20s	: int	3	24	34
## \$ X30s	: int	1	15	49
## \$ X40s	: int	4	25	36
## \$ X50s	: int	4	24	14
## \$ X60s	: int	6	12	12
## \$ X70s	: int	3	10	2
## \$ X80s	: int	0	6	1
## \$ X90s	: int	1	3	0
## \$ TOTAL	: int	23	127	149
## \$ Arthritis..15.years.and.over.	: num	29	23.8	23.9
## \$ Asthma	: num	11.2	11.1	9
## \$ Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese	: num	38.7	38.1	35
## \$ Body.mass.index..adjusted.self.reported..adult..18.years.and.over...overweight:	num	33.2	29.3	35
## \$ Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over.	: num	7.9	6.4	4.9
## \$ Current.smoker..daily	: num	15.6	20.1	11
## \$ Current.smoker..daily.or.occasional	: num	20.9	24.2	14
## \$ Diabetes	: num	11	9	8.6
## \$ Has.a.regular.healthcare.provider	: num	88.7	90.3	92
## \$ Heavy.drinking	: num	23.5	22.3	18
## \$ High.blood.pressure	: num	24.6	18.1	25
## \$ Influenza.immunization.in.the.past.12.months	: num	31.2	31.7	32
## \$ Life.satisfaction..satisfied.or.very.satisfied	: num	92.3	90.9	93
## \$ Mood.disorder	: num	11.9	10.7	10
## \$ Perceived.health..fair.or.poor	: num	15.4	15.2	11
## \$ Perceived.health..very.good.or.excellent	: num	55.2	49	61.9
## \$ Perceived.life.stress..most.days.quite.a.bit.or.extremely.stressful	: num	19.9	17.1	14
## \$ Perceived.mental.health..fair.or.poor	: num	8	9.9	8.2
## \$ Perceived.mental.health..very.good.or.excellent	: num	66.6	67.5	69
## \$ Physical.activity..150.minutes.per.week..adult..18.years.and.over.	: num	55.7	52.9	53
## \$ Physical.activity..average.60.minutes.per.day..youth..12.to.17.years.old.	: num	59.9	55.2	62
## \$ Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong	: num	79	69.4	76.8
## \$ DBPOP	: int	112847	13729	0
## \$ prox_idx_emp	: num	0.0025	0.016	0
## \$ prox_idx_pharma	: num	0.0208	0.022	0
## \$ prox_idx_childcare	: num	0.0144	0.020	0

```
## $ prox_idx_health : num 0.0033 0.004
## $ prox_idx_grocery : num 0.0346 0.041
## $ prox_idx_educpri : num 0.0794 0.100
## $ prox_idx_educsec : num 0.0966 0.049
## $ prox_idx_lib : num 0.0834 0.071
## $ prox_idx_parks : num 0.0278 0.044
## $ prox_idx_transit : num 0.0099 0.000
## $ amenity : num 0.0126 0.084
## $ FATAL_prop : num 0.00 3.64e-03
## $ TOTAL_prop : num 0.000204 0.00
```

```
data[is.na(data)] <- 0

# drop duplicate rows
#data <- data[-c(13, 28), ]

# dropping unneeded cols, cleaning...
data <- data[-c(1, 3:39, 41:49, 61)]
data <- data[-c(26:35)]
#data

# calcualte the proportion of COVID-19 cases per PHU
#data$TOTAL_prop <- data$TOTAL/data$HRPOP
#data$FATAL_prop <- data$FATAL/data$HRPOP

# need to convert all percentages into proportion...
prop <- function(x) x/100
data <- data.frame(data[1:3], lapply(data[4:24], prop), data[25:28])
```

Keep only the predictors and response variables needed for the analysis

```
final_cases <- subset(data, select = -c(Reporting_PHU, FATAL, TOTAL, FATAL_prop, DBPOP))
final_fatalities <- subset(data, select = -c(Reporting_PHU, FATAL, TOTAL, TOTAL_prop, DBPOP))
```

LASSO

```
### LASSO
library(leaps)
library(corrplot)
```

```
## corrplot 0.84 loaded
```

```
library(glmnet)
```

```
## Loading required package: Matrix
```

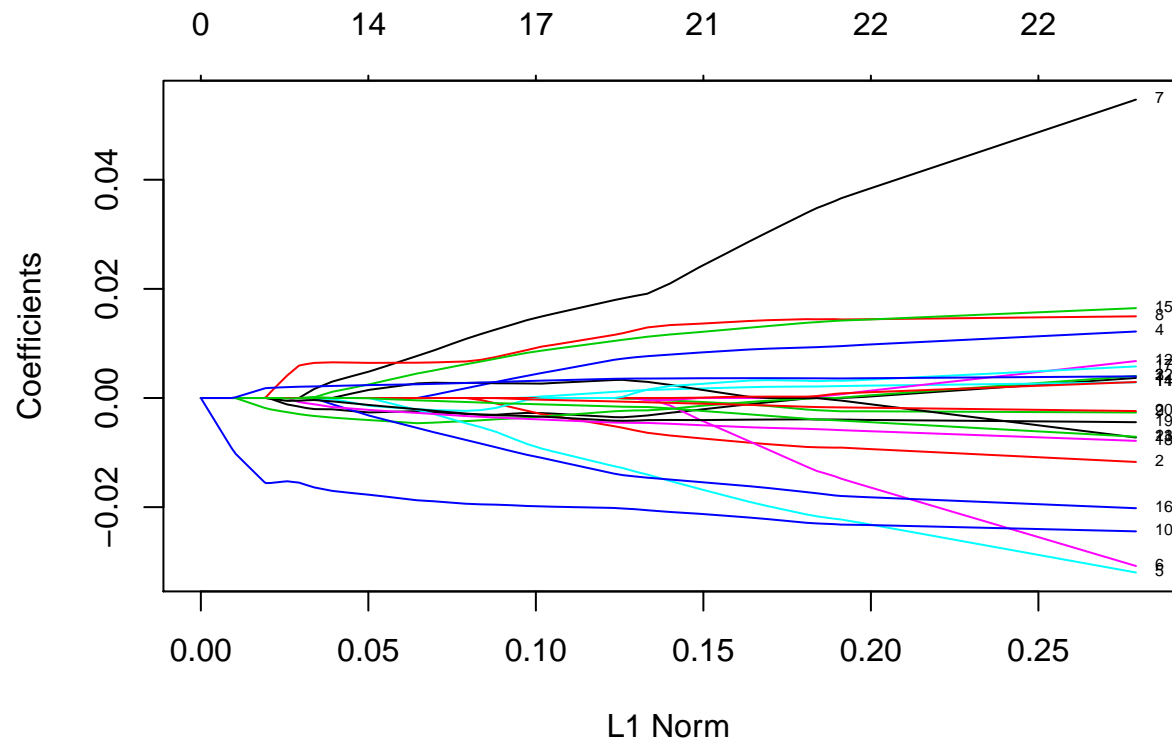
```
## Loaded glmnet 3.0-2
```

```
y=na.omit(final_cases)
x=model.matrix(TOTAL_prop~.,final_cases)[,-1]
rownames(x)=c()
```

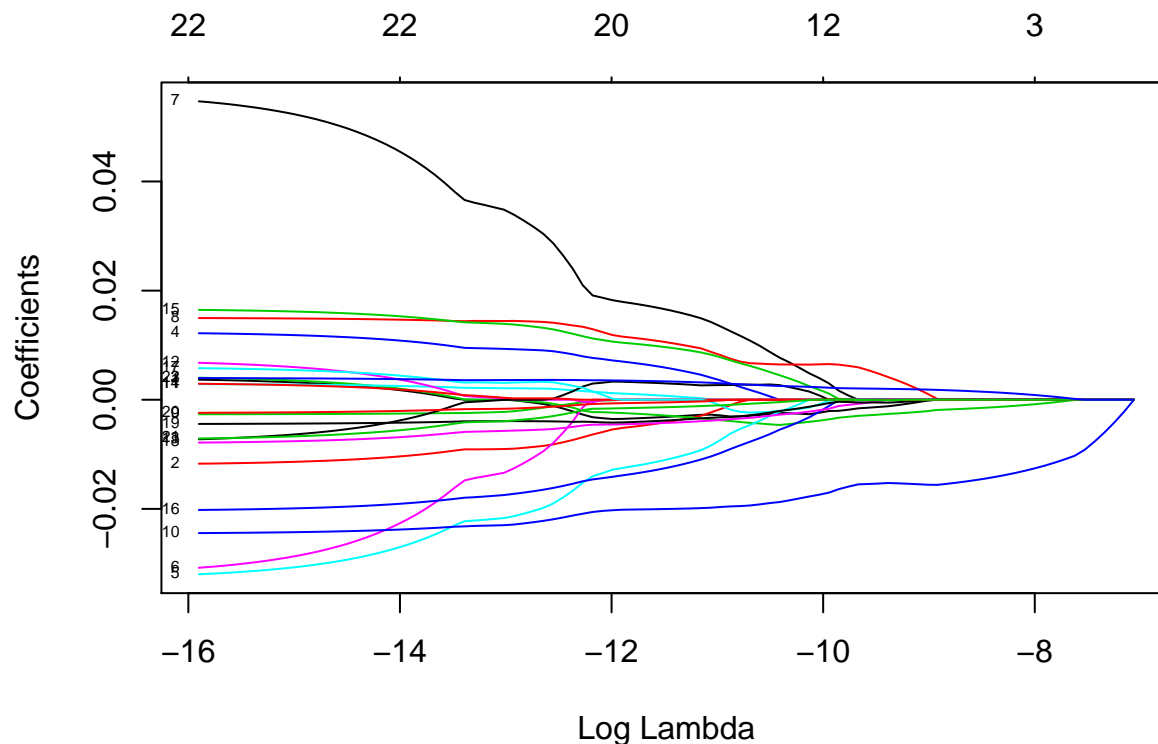
```
y=as.matrix(final_cases$TOTAL_prop)
```

```
lasso <- glmnet(x, y, alpha = 1)
```

```
plot(lasso, label = TRUE)
```



```
plot(lasso, xvar = "lambda", label = TRUE)
```



```
regfit.full=regsubsets(TOTAL_prop ~ ., final_cases, nvmax = 23)
reg.summary = summary(regfit.full)

reg.summary
```

```
## Subset selection object
## Call: regsubsets.formula(TOTAL_prop ~ ., final_cases, nvmax = 23)
## 22 Variables (and intercept)
##
##                                     Forced in
## Arthritis..15.years.and.over.      FALSE
## Asthma                             FALSE
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese  FALSE
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...overweight  FALSE
## Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over.             FALSE
## Current.smoker..daily                                                         FALSE
## Current.smoker..daily.or.occasional                                           FALSE
## Diabetes                                                                      FALSE
## Has.a.regular.healthcare.provider                                             FALSE
## Heavy.drinking                                                                FALSE
## Influenza.immunization.in.the.past.12.months                                FALSE
## Life.satisfaction..satisfied.or.very.satisfied                              FALSE
## Mood.disorder                                                                FALSE
## Perceived.health..fair.or.poor                                                FALSE
## Perceived.health..very.good.or.excellent                                     FALSE
## Perceived.life.stress..most.days.quite.a.bit.or.extremely.stressful          FALSE
```

## Perceived.mental.health..fair.or.poor	FALSE
## Perceived.mental.health..very.good.or.excellent	FALSE
## Physical.activity..150.minutes.per.week..adult..18.years.and.over.	FALSE
## Physical.activity..average.60.minutes.per.day..youth..12.to.17.years.old.	FALSE
## Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong	FALSE
## amenity	FALSE
##	Forced out
## Arthritis..15.years.and.over.	FALSE
## Asthma	FALSE
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese	FALSE
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...overweight	FALSE
## Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over.	FALSE
## Current.smoker..daily	FALSE
## Current.smoker..daily.or.occasional	FALSE
## Diabetes	FALSE
## Has.a.regular.healthcare.provider	FALSE
## Heavy.drinking	FALSE
## Influenza.immunization.in.the.past.12.months	FALSE
## Life.satisfaction..satisfied.or.very.satisfied	FALSE
## Mood.disorder	FALSE
## Perceived.health..fair.or.poor	FALSE
## Perceived.health..very.good.or.excellent	FALSE
## Perceived.life.stress..most.days.quite.a.bit.or.extremely.stressful	FALSE
## Perceived.mental.health..fair.or.poor	FALSE
## Perceived.mental.health..very.good.or.excellent	FALSE
## Physical.activity..150.minutes.per.week..adult..18.years.and.over.	FALSE
## Physical.activity..average.60.minutes.per.day..youth..12.to.17.years.old.	FALSE
## Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong	FALSE
## amenity	FALSE
## 1 subsets of each size up to 22	
## Selection Algorithm: exhaustive	
## Arthritis..15.years.and.over. Asthma	
## 1 (1) " "	" "
## 2 (1) " "	" "
## 3 (1) " "	" "
## 4 (1) " "	" "
## 5 (1) " "	" "
## 6 (1) " "	" "
## 7 (1) " "	" "
## 8 (1) " "	" "
## 9 (1) " "	" "
## 10 (1) " "	" "
## 11 (1) " "	" "
## 12 (1) " "	" "
## 13 (1) " "	" "
## 14 (1) " "	"*"
## 15 (1) " "	"*"
## 16 (1) " "	"*"
## 17 (1) " "	"*"
## 18 (1) " "	"*"
## 19 (1) "*"	"*"
## 20 (1) "*"	"*"
## 21 (1) "*"	"*"
## 22 (1) "*"	"*"

```

##          Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) "*"
## 5 ( 1 ) " "
## 6 ( 1 ) "*"
## 7 ( 1 ) "*"
## 8 ( 1 ) "*"
## 9 ( 1 ) "*"
## 10 ( 1 ) "*"
## 11 ( 1 ) "*"
## 12 ( 1 ) " "
## 13 ( 1 ) " "
## 14 ( 1 ) " "
## 15 ( 1 ) " "
## 16 ( 1 ) " "
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##          Body.mass.index..adjusted.self.reported..adult..18.years.and.over...overweight
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) " "
## 8 ( 1 ) " "
## 9 ( 1 ) " "
## 10 ( 1 ) "*"
## 11 ( 1 ) "*"
## 12 ( 1 ) "*"
## 13 ( 1 ) "*"
## 14 ( 1 ) "*"
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##          Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over.
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) " "

```

```

## 8 ( 1 ) " "
## 9 ( 1 ) "*"
## 10 ( 1 ) "*"
## 11 ( 1 ) "*"
## 12 ( 1 ) "*"
## 13 ( 1 ) "*"
## 14 ( 1 ) "*"
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##
## Current.smoker..daily Current.smoker..daily.or.occasional
## 1 ( 1 ) " " " "
## 2 ( 1 ) " " " "
## 3 ( 1 ) " " "*"
## 4 ( 1 ) " " "*"
## 5 ( 1 ) " " "*"
## 6 ( 1 ) " " "*"
## 7 ( 1 ) " " "*"
## 8 ( 1 ) " " "*"
## 9 ( 1 ) " " "*"
## 10 ( 1 ) " " "*"
## 11 ( 1 ) " " "*"
## 12 ( 1 ) "*" "*"
## 13 ( 1 ) "*" "*"
## 14 ( 1 ) "*" "*"
## 15 ( 1 ) "*" "*"
## 16 ( 1 ) "*" "*"
## 17 ( 1 ) "*" "*"
## 18 ( 1 ) "*" "*"
## 19 ( 1 ) "*" "*"
## 20 ( 1 ) "*" "*"
## 21 ( 1 ) "*" "*"
## 22 ( 1 ) "*" "*"
##
## Diabetes Has.a.regular.healthcare.provider Heavy.drinking
## 1 ( 1 ) " " " " "*"
## 2 ( 1 ) " " " " "*"
## 3 ( 1 ) " " " " "*"
## 4 ( 1 ) " " " " "*"
## 5 ( 1 ) " " " " "*"
## 6 ( 1 ) "*" " " "*"
## 7 ( 1 ) " " " " "*"
## 8 ( 1 ) "*" " " "*"
## 9 ( 1 ) "*" " " "*"
## 10 ( 1 ) "*" " " "*"
## 11 ( 1 ) "*" " " "*"
## 12 ( 1 ) " " " " "*"
## 13 ( 1 ) "*" " " "*"
## 14 ( 1 ) "*" " " "*"
## 15 ( 1 ) "*" "*" "*"

```



```

## 16 ( 1 ) "*"      "*"      "*"
## 17 ( 1 ) "*"      "*"      "*"
## 18 ( 1 ) "*"      "*"      "*"
## 19 ( 1 ) "*"      "*"      "*"
## 20 ( 1 ) "*"      "*"      "*"
## 21 ( 1 ) "*"      "*"      "*"
## 22 ( 1 ) "*"      "*"      "*"
##      Influenza.immunization.in.the.past.12.months
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) " "
## 8 ( 1 ) " "
## 9 ( 1 ) " "
## 10 ( 1 ) " "
## 11 ( 1 ) " "
## 12 ( 1 ) " "
## 13 ( 1 ) " "
## 14 ( 1 ) " "
## 15 ( 1 ) " "
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##      Life.satisfaction..satisfied.or.very.satisfied Mood.disorder
## 1 ( 1 ) " "      " "
## 2 ( 1 ) " "      " "
## 3 ( 1 ) " "      " "
## 4 ( 1 ) " "      " "
## 5 ( 1 ) " "      " "
## 6 ( 1 ) " "      " "
## 7 ( 1 ) " "      " "
## 8 ( 1 ) " "      " "
## 9 ( 1 ) " "      " "
## 10 ( 1 ) " "      " "
## 11 ( 1 ) " "      " "
## 12 ( 1 ) " "      " "
## 13 ( 1 ) " "      " "
## 14 ( 1 ) " "      " "
## 15 ( 1 ) " "      " "
## 16 ( 1 ) " "      " "
## 17 ( 1 ) " "      " "
## 18 ( 1 ) " "      "*"
## 19 ( 1 ) " "      "*"
## 20 ( 1 ) " "      "*"
## 21 ( 1 ) "*"      "*"
## 22 ( 1 ) "*"      "*"
##      Perceived.health..fair.or.poor

```

```

## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) " "
## 8 ( 1 ) " "
## 9 ( 1 ) " "
## 10 ( 1 ) " "
## 11 ( 1 ) " "
## 12 ( 1 ) " "
## 13 ( 1 ) " "
## 14 ( 1 ) " "
## 15 ( 1 ) " "
## 16 ( 1 ) " "
## 17 ( 1 ) " "
## 18 ( 1 ) " "
## 19 ( 1 ) " "
## 20 ( 1 ) " "
## 21 ( 1 ) " "
## 22 ( 1 ) "*"
##
Perceived.health..very.good.or.excellent
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) "*"
## 6 ( 1 ) "*"
## 7 ( 1 ) "*"
## 8 ( 1 ) "*"
## 9 ( 1 ) "*"
## 10 ( 1 ) "*"
## 11 ( 1 ) "*"
## 12 ( 1 ) "*"
## 13 ( 1 ) "*"
## 14 ( 1 ) "*"
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##
Perceived.life.stress..most.days.quite.a.bit.or.extremely.stressful
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) "*"
## 8 ( 1 ) "*"

```

```

## 9 ( 1 ) "*"
## 10 ( 1 ) "*"
## 11 ( 1 ) "*"
## 12 ( 1 ) "*"
## 13 ( 1 ) "*"
## 14 ( 1 ) "*"
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##
## Perceived.mental.health..fair.or.poor
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) " "
## 8 ( 1 ) " "
## 9 ( 1 ) " "
## 10 ( 1 ) " "
## 11 ( 1 ) " "
## 12 ( 1 ) " "
## 13 ( 1 ) " "
## 14 ( 1 ) " "
## 15 ( 1 ) " "
## 16 ( 1 ) " "
## 17 ( 1 ) " "
## 18 ( 1 ) " "
## 19 ( 1 ) " "
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##
## Perceived.mental.health..very.good.or.excellent
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) "*"
## 6 ( 1 ) " "
## 7 ( 1 ) "*"
## 8 ( 1 ) "*"
## 9 ( 1 ) "*"
## 10 ( 1 ) "*"
## 11 ( 1 ) "*"
## 12 ( 1 ) "*"
## 13 ( 1 ) "*"
## 14 ( 1 ) "*"
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"

```

```

## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##      Physical.activity..150.minutes.per.week..adult..18.years.and.over.
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) " "
## 8 ( 1 ) " "
## 9 ( 1 ) " "
## 10 ( 1 ) " "
## 11 ( 1 ) "*"
## 12 ( 1 ) "*"
## 13 ( 1 ) "*"
## 14 ( 1 ) "*"
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##      Physical.activity..average.60.minutes.per.day..youth..12.to.17.years.old.
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) " "
## 8 ( 1 ) " "
## 9 ( 1 ) " "
## 10 ( 1 ) " "
## 11 ( 1 ) " "
## 12 ( 1 ) "*"
## 13 ( 1 ) "*"
## 14 ( 1 ) "*"
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##      Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong
## 1 ( 1 ) " "

```

```
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) " "
## 8 ( 1 ) " "
## 9 ( 1 ) " "
## 10 ( 1 ) " "
## 11 ( 1 ) " "
## 12 ( 1 ) "*"
## 13 ( 1 ) "*"
## 14 ( 1 ) "*"
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##          amenity
## 1 ( 1 ) " "
## 2 ( 1 ) "*"
## 3 ( 1 ) "*"
## 4 ( 1 ) "*"
## 5 ( 1 ) "*"
## 6 ( 1 ) "*"
## 7 ( 1 ) "*"
## 8 ( 1 ) "*"
## 9 ( 1 ) "*"
## 10 ( 1 ) "*"
## 11 ( 1 ) "*"
## 12 ( 1 ) "*"
## 13 ( 1 ) "*"
## 14 ( 1 ) "*"
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"

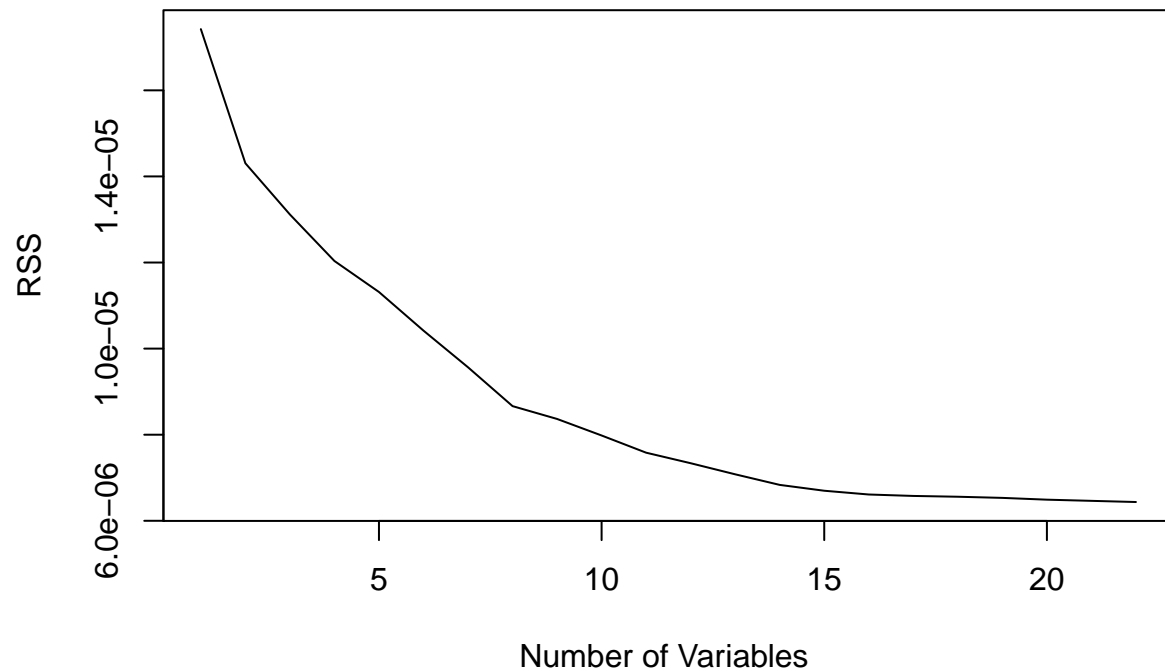
```

```
reg.summary$RSS
```

```
## [1] 1.742203e-05 1.430581e-05 1.311628e-05 1.203591e-05 1.131552e-05
## [6] 1.041765e-05 9.565955e-06 8.663833e-06 8.365882e-06 7.981724e-06
## [11] 7.582398e-06 7.337840e-06 7.080007e-06 6.833929e-06 6.698241e-06
## [16] 6.611192e-06 6.577814e-06 6.558809e-06 6.532418e-06 6.490623e-06
## [21] 6.463716e-06 6.436578e-06

```

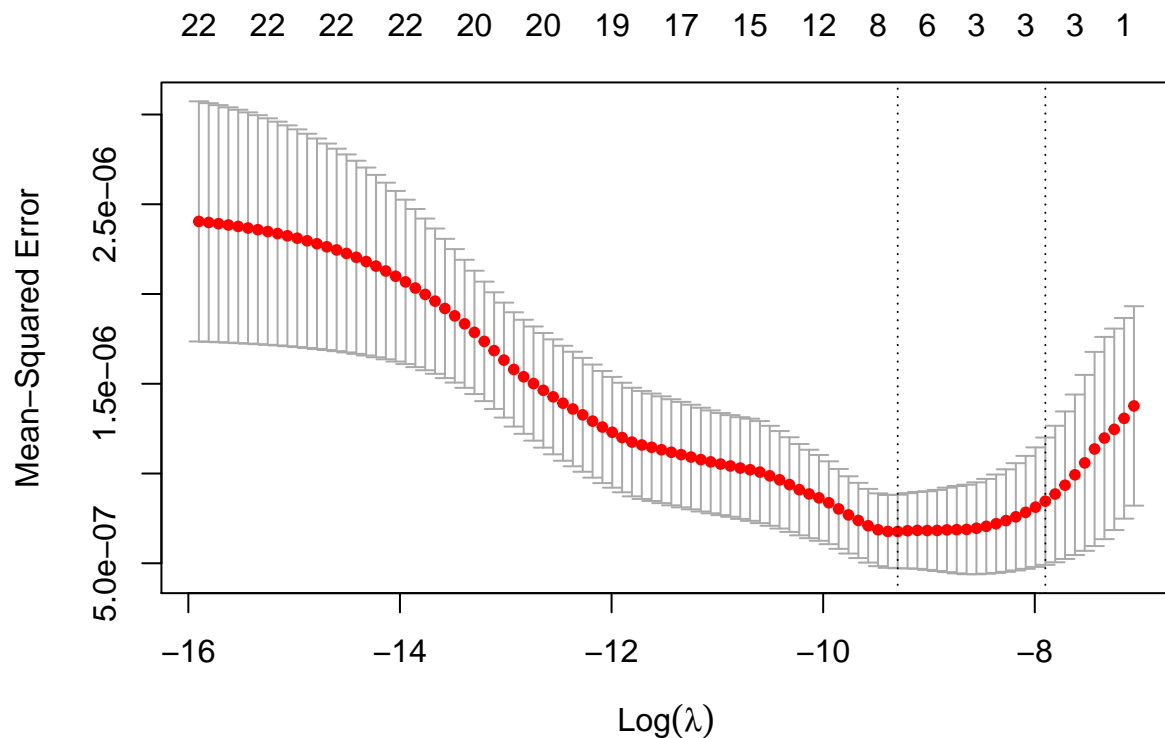
```
plot(reg.summary$rss ,xlab="Number of Variables ",ylab="RSS", type="l")
```



```
# plot suggests to keep ~8 predictors

set.seed(123)
cv.lasso <- cv.glmnet(x, y, alpha = 1)
# Fit the final model on the training data
model <- glmnet(x, y, alpha = 1, lambda = cv.lasso$lambda.min)

plot(cv.lasso)
```



```
cv.lasso$lambda.min
```

```
## [1] 9.184336e-05
```

```
coef(cv.lasso, cv.lasso$lambda.min)
```

```
## 23 x 1 sparse Matrix of class "dgCMatrix"
##
## (Intercept) 5.398123e-03
## Arthritis..15.years.and.over. -9.799490e-04
## Asthma .
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese -2.498881e-03
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...overweight .
## Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over. .
## Current.smoker..daily .
## Current.smoker..daily.or.occasional .
## Diabetes 3.516130e-03
## Has.a.regular.healthcare.provider .
## Heavy.drinking -1.531532e-02
## Influenza.immunization.in.the.past.12.months .
## Life.satisfaction..satisfied.or.very.satisfied .
## Mood.disorder .
## Perceived.health..fair.or.poor .
## Perceived.health..very.good.or.excellent .
## Perceived.life.stress..most.days.quite.a.bit.or.extremely.stressful .
```

```
## Perceived.mental.health..fair.or.poor .
## Perceived.mental.health..very.good.or.excellent .
## Physical.activity..150.minutes.per.week..adult..18.years.and.over. -4.452048e-04
## Physical.activity..average.60.minutes.per.day..youth..12.to.17.years.old. .
## Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong -3.515305e-05
## amenity 1.971229e-03
```

looks like the appropriate value for predictors is 4.

beta regression model...

```
library(betareg)
beta_model <- betareg(TOTAL_prop~Arthritis..15.years.and.over.+Body.mass.index..adjusted.self.reported.
summary(beta_model)
```

```
##
## Call:
## betareg(formula = TOTAL_prop ~ Arthritis..15.years.and.over. + Body.mass.index..adjusted.self.reported.
## Heavy.drinking + amenity, data = final_cases, link = "loglog")
##
## Standardized weighted residuals 2:
##      Min      1Q  Median      3Q      Max
## -2.0545 -0.7306 -0.0993  0.5537  2.2037
##
## Coefficients (mean model with loglog link):
##
## (Intercept) Estimate
## Arthritis..15.years.and.over. -1.46624
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese -0.57619
## Heavy.drinking -0.06477
## amenity -1.43391
## Std. Error 0.14896
## (Intercept) 0.09078
## Arthritis..15.years.and.over. 0.46487
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese 0.37109
## Heavy.drinking 0.49521
## amenity 0.09812
## z value
## (Intercept) -16.151
## Arthritis..15.years.and.over. -1.239
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese -0.175
## Heavy.drinking -2.896
## amenity 1.518
## Pr(>|z|)
## (Intercept) < 2e-16
## Arthritis..15.years.and.over. 0.21518
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese 0.86144
## Heavy.drinking 0.00379
## amenity 0.12897
##
## (Intercept) ***
## Arthritis..15.years.and.over.
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese
```

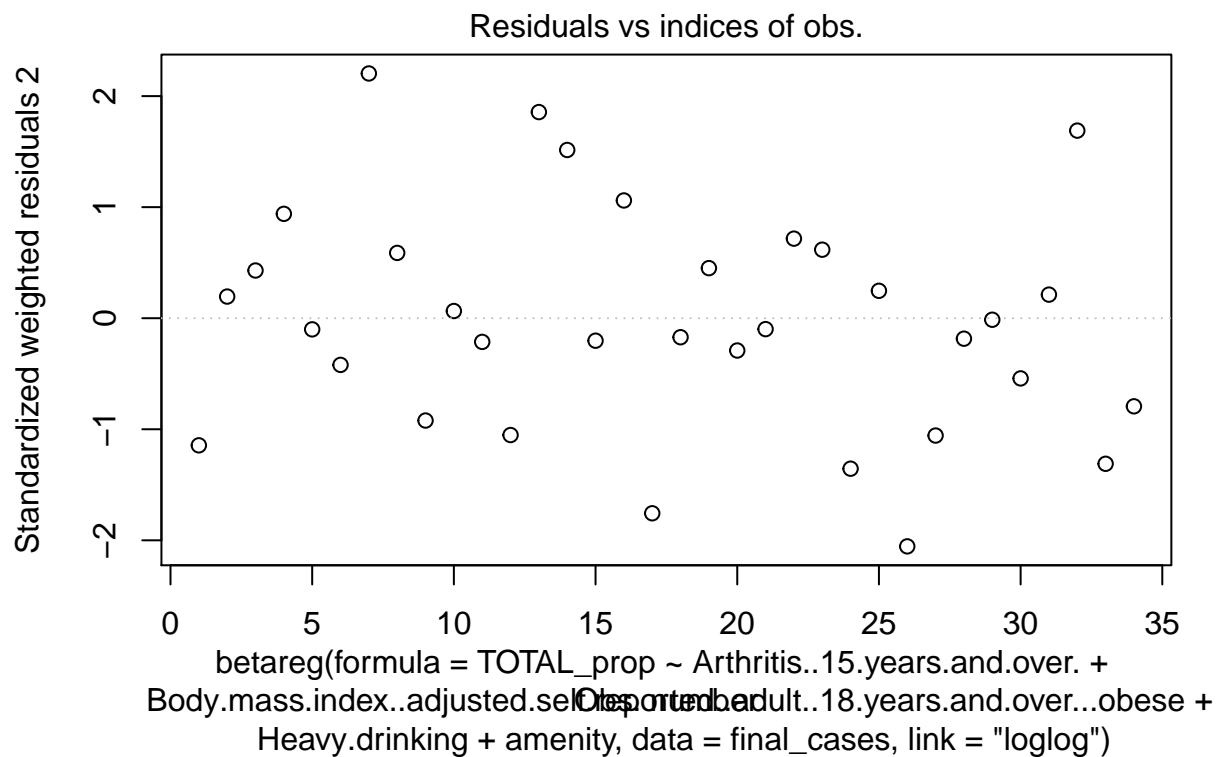


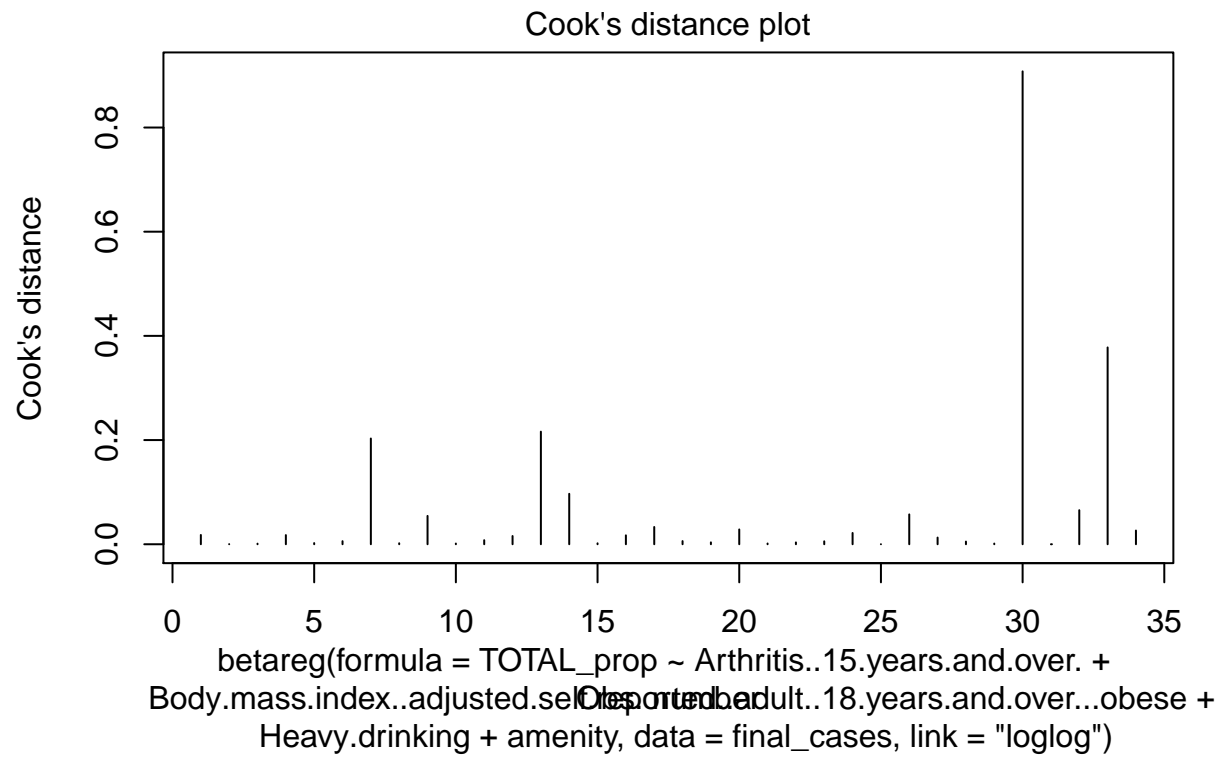
```
## Heavy.drinking                                     **
## amenity
##
## Phi coefficients (precision model with identity link):
##      Estimate Std. Error z value Pr(>|z|)
## (phi)   3063.9      766.3   3.998 6.39e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Type of estimator: ML (maximum likelihood)
## Log-likelihood: 207.6 on 6 Df
## Pseudo R-squared: 0.5884
## Number of iterations: 3568 (BFGS) + 10 (Fisher scoring)

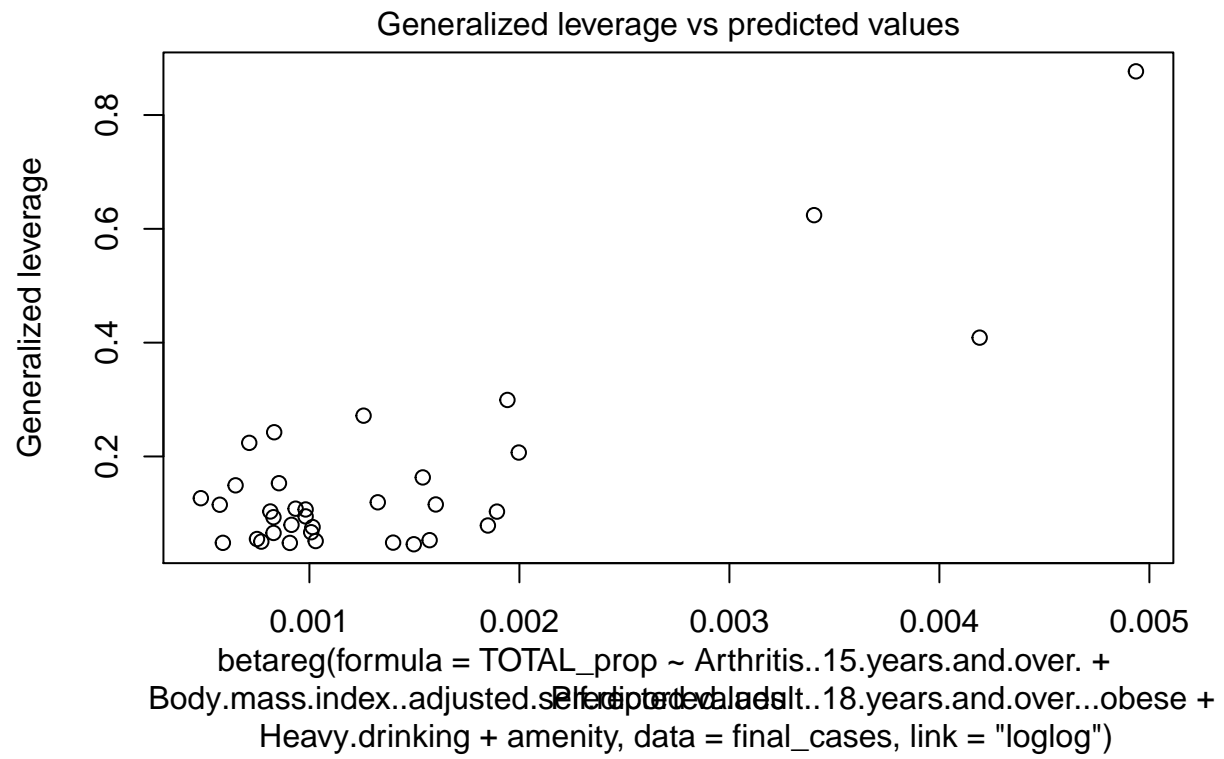
# checking the links.
sapply(c("logit", "probit", "cauchit", "loglog", "cloglog"),
function(x) logLik(update(beta_model, link = x)))

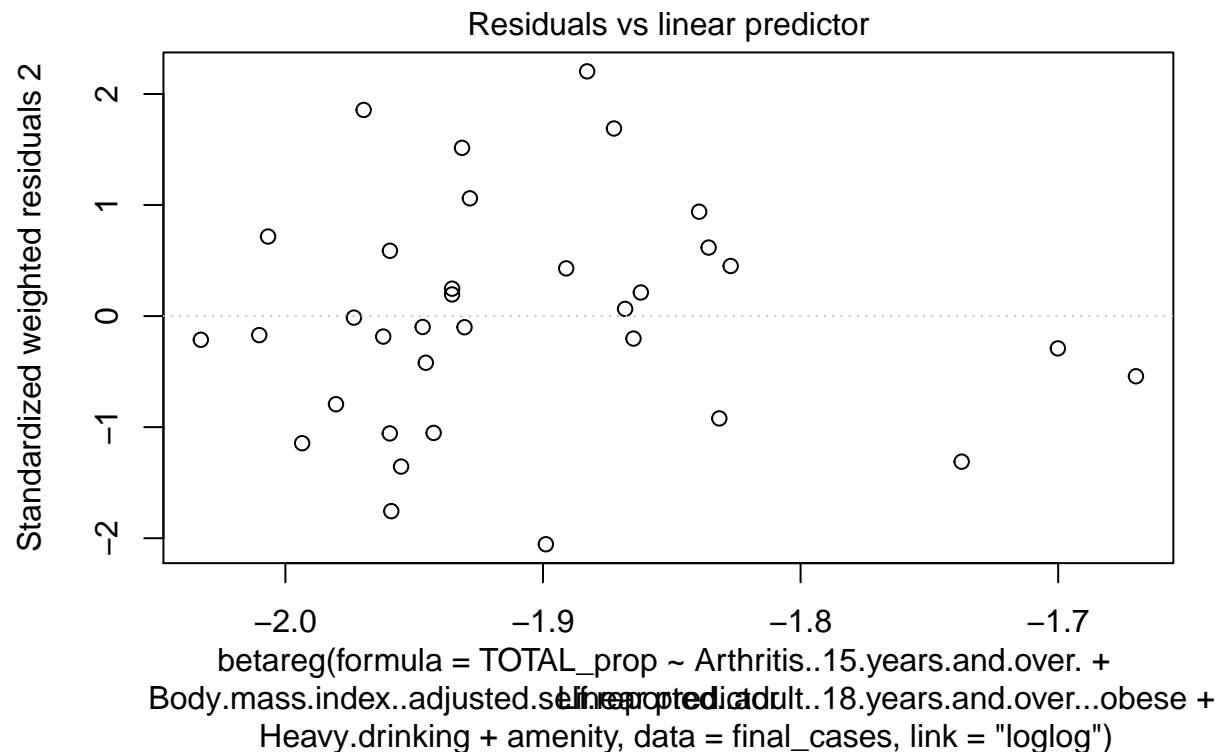
##      logit      probit      cauchit      loglog      cloglog
## 207.0923 207.3729 204.3610 207.5823 207.0895

# diagnostic plots.
plot(beta_model)
```









```
library(dplyr)
library(ggplot2)
library(ggrepel)
library(ggiraphExtra)
library(plotly)
```

```
## Warning: package 'plotly' was built under R version 3.6.2
```

```
##
```

```
## Attaching package: 'plotly'
```

```
## The following object is masked from 'package:ggplot2':
```

```
##
```

```
## last_plot
```

```
## The following object is masked from 'package:stats':
```

```
##
```

```
## filter
```

```
## The following object is masked from 'package:graphics':
```

```
##
```

```
## layout
```

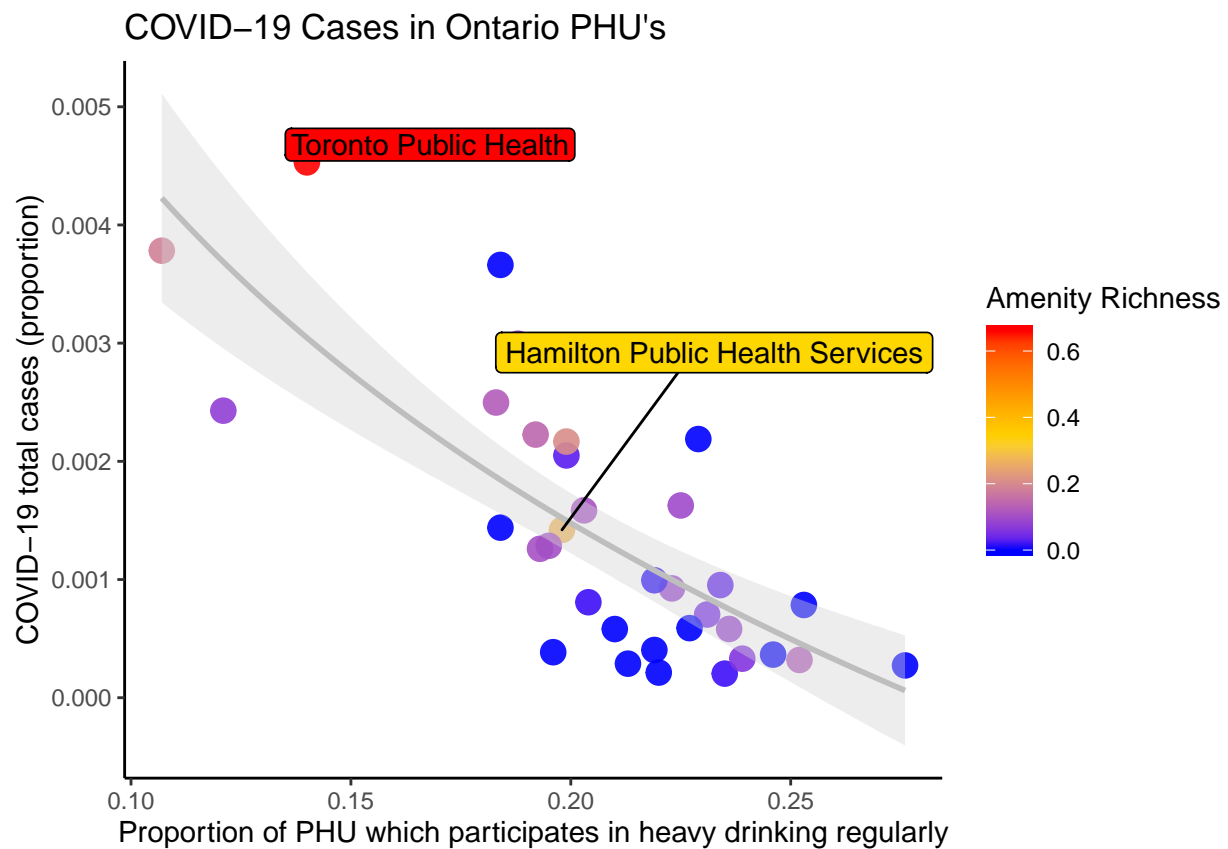
```

ggplot(final_cases, aes(x=Heavy.drinking, y=TOTAL_prop, col = amenity)) +
  theme_classic()+
  geom_point(alpha = 0.9, aes(colour=amenity), size=4) +
  stat_smooth(method = "glm", family= "beta", formula = y ~ log(x), se = T, col = "gray", fill="lightgray",
    xlab("Proportion of PHU which participates in heavy drinking regularly")+
    ylab("COVID-19 total cases (proportion)")+
    labs(col = "Amenity Richness")+
    ggtitle("COVID-19 Cases in Ontario PHU's")+
    scale_color_gradientn(colors = c("blue", "gold", "red")) +

ggrepel::geom_label_repel(data=data %>% filter(Heavy.drinking == 0.198), aes(label = Reporting_PHU, fill = amenity),
  geom_label( data=data %>% filter(TOTAL_prop> 0.004),aes(label=Reporting_PHU, fill = factor(amenity)),
  guides(fill=FALSE)+
  scale_fill_manual(values = setNames(c("gold", "red"), levels(data$amenity)))

```

Warning: Ignoring unknown parameters: family



LASSO fatalities

```

### LASSO
library(leaps)
library(corrplot)
library(glmnet)

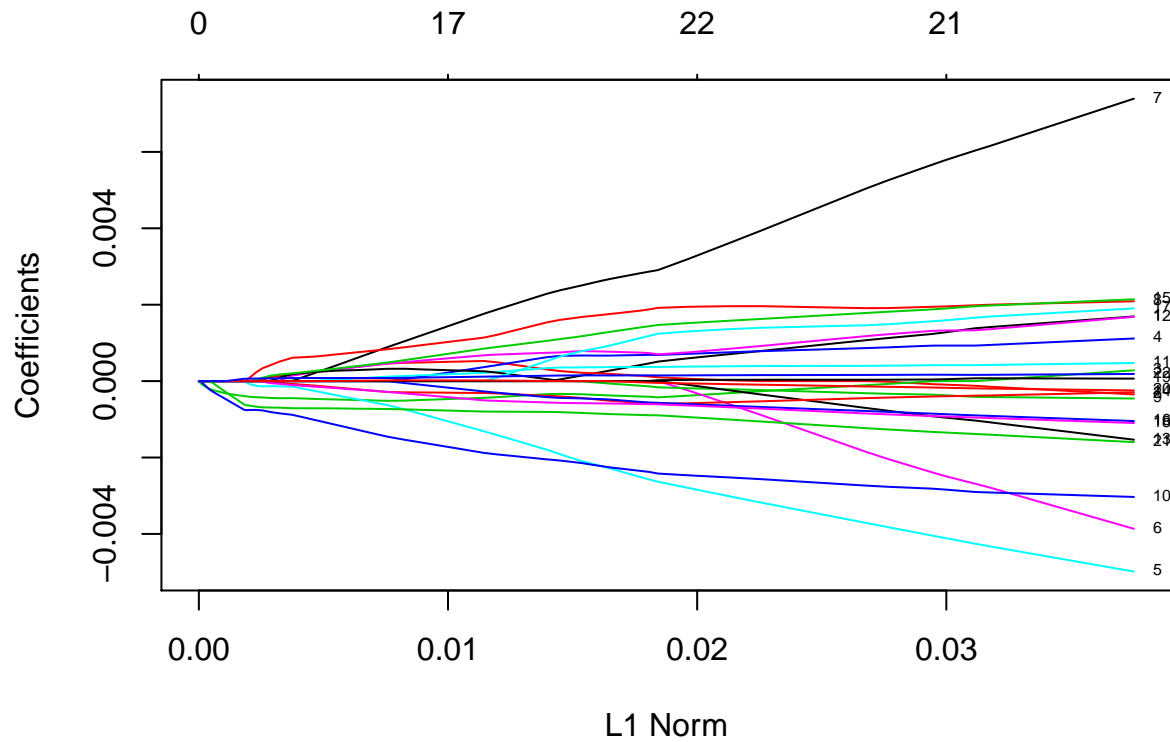
```

```

y=na.omit(final_fatalities)
x=model.matrix(FATAL_prop~.,final_fatalities)[,-1]
rownames(x)=c()
y=as.matrix(final_fatalities$FATAL_prop)

lasso <- glmnet(x, y, alpha = 1)
plot(lasso, label = TRUE)

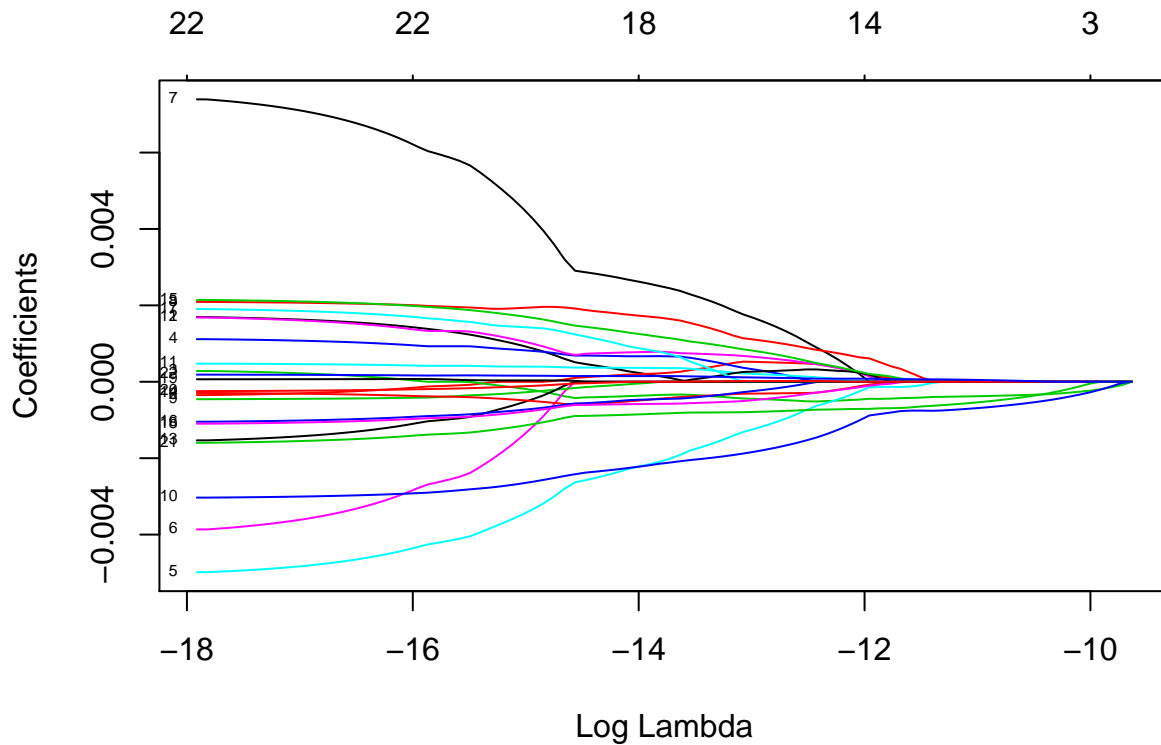
```



```

plot(lasso, xvar = "lambda", label = TRUE)

```



```
regfit.full=regsubsets(FATAL_prop~.,final_fatalities, nvmax = 23)
reg.summary = summary(regfit.full)

reg.summary
```

```
## Subset selection object
## Call: regsubsets.formula(FATAL_prop ~ ., final_fatalities, nvmax = 23)
## 22 Variables (and intercept)
##
##                                     Forced in
## Arthritis..15.years.and.over.      FALSE
## Asthma                             FALSE
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese  FALSE
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...overweight  FALSE
## Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over.             FALSE
## Current.smoker..daily                FALSE
## Current.smoker..daily.or.occasional    FALSE
## Diabetes                             FALSE
## Has.a.regular.healthcare.provider      FALSE
## Heavy.drinking                        FALSE
## Influenza.immunization.in.the.past.12.months  FALSE
## Life.satisfaction..satisfied.or.very.satisfied  FALSE
## Mood.disorder                         FALSE
## Perceived.health..fair.or.poor         FALSE
## Perceived.health..very.good.or.excellent  FALSE
## Perceived.life.stress..most.days.quite.a.bit.or.extremely.stressful  FALSE
```

## Perceived.mental.health..fair.or.poor	FALSE
## Perceived.mental.health..very.good.or.excellent	FALSE
## Physical.activity..150.minutes.per.week..adult..18.years.and.over.	FALSE
## Physical.activity..average.60.minutes.per.day..youth..12.to.17.years.old.	FALSE
## Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong	FALSE
## amenity	FALSE
##	Forced out
## Arthritis..15.years.and.over.	FALSE
## Asthma	FALSE
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese	FALSE
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...overweight	FALSE
## Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over.	FALSE
## Current.smoker..daily	FALSE
## Current.smoker..daily.or.occasional	FALSE
## Diabetes	FALSE
## Has.a.regular.healthcare.provider	FALSE
## Heavy.drinking	FALSE
## Influenza.immunization.in.the.past.12.months	FALSE
## Life.satisfaction..satisfied.or.very.satisfied	FALSE
## Mood.disorder	FALSE
## Perceived.health..fair.or.poor	FALSE
## Perceived.health..very.good.or.excellent	FALSE
## Perceived.life.stress..most.days.quite.a.bit.or.extremely.stressful	FALSE
## Perceived.mental.health..fair.or.poor	FALSE
## Perceived.mental.health..very.good.or.excellent	FALSE
## Physical.activity..150.minutes.per.week..adult..18.years.and.over.	FALSE
## Physical.activity..average.60.minutes.per.day..youth..12.to.17.years.old.	FALSE
## Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong	FALSE
## amenity	FALSE
## 1 subsets of each size up to 22	
## Selection Algorithm: exhaustive	
## Arthritis..15.years.and.over. Asthma	
## 1 (1) " "	" "
## 2 (1) " "	" "
## 3 (1) " "	" "
## 4 (1) " "	"*"
## 5 (1) " "	"*"
## 6 (1) " "	"*"
## 7 (1) " "	"*"
## 8 (1) " "	"*"
## 9 (1) "*"	"*"
## 10 (1) "*"	"*"
## 11 (1) " "	" "
## 12 (1) " "	" "
## 13 (1) "*"	" "
## 14 (1) "*"	" "
## 15 (1) "*"	" "
## 16 (1) "*"	" "
## 17 (1) "*"	" "
## 18 (1) "*"	" "
## 19 (1) "*"	" "
## 20 (1) "*"	"*"
## 21 (1) "*"	"*"
## 22 (1) "*"	"*"


```

##          Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese
## 1  ( 1 )  "*"
## 2  ( 1 )  " "
## 3  ( 1 )  "*"
## 4  ( 1 )  "*"
## 5  ( 1 )  "*"
## 6  ( 1 )  "*"
## 7  ( 1 )  "*"
## 8  ( 1 )  "*"
## 9  ( 1 )  "*"
## 10 ( 1 )  "*"
## 11 ( 1 )  " "
## 12 ( 1 )  " "
## 13 ( 1 )  " "
## 14 ( 1 )  " "
## 15 ( 1 )  " "
## 16 ( 1 )  " "
## 17 ( 1 )  " "
## 18 ( 1 )  " "
## 19 ( 1 )  "*"
## 20 ( 1 )  "*"
## 21 ( 1 )  "*"
## 22 ( 1 )  "*"
##          Body.mass.index..adjusted.self.reported..adult..18.years.and.over...overweight
## 1  ( 1 )  " "
## 2  ( 1 )  " "
## 3  ( 1 )  " "
## 4  ( 1 )  " "
## 5  ( 1 )  " "
## 6  ( 1 )  " "
## 7  ( 1 )  " "
## 8  ( 1 )  " "
## 9  ( 1 )  " "
## 10 ( 1 )  " "
## 11 ( 1 )  "*"
## 12 ( 1 )  "*"
## 13 ( 1 )  "*"
## 14 ( 1 )  "*"
## 15 ( 1 )  "*"
## 16 ( 1 )  "*"
## 17 ( 1 )  "*"
## 18 ( 1 )  "*"
## 19 ( 1 )  "*"
## 20 ( 1 )  "*"
## 21 ( 1 )  "*"
## 22 ( 1 )  "*"
##          Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over.
## 1  ( 1 )  " "
## 2  ( 1 )  " "
## 3  ( 1 )  " "
## 4  ( 1 )  " "
## 5  ( 1 )  " "
## 6  ( 1 )  " "
## 7  ( 1 )  "*"

```

```

## 8 ( 1 ) "*"
## 9 ( 1 ) "*"
## 10 ( 1 ) "*"
## 11 ( 1 ) "*"
## 12 ( 1 ) "*"
## 13 ( 1 ) "*"
## 14 ( 1 ) "*"
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##      Current.smoker..daily Current.smoker..daily.or.occasional
## 1 ( 1 ) " " " "
## 2 ( 1 ) " " " "
## 3 ( 1 ) " " " "
## 4 ( 1 ) "*" " "
## 5 ( 1 ) "*" " "
## 6 ( 1 ) "*" " "
## 7 ( 1 ) " " "*"
## 8 ( 1 ) " " "*"
## 9 ( 1 ) " " "*"
## 10 ( 1 ) " " "*"
## 11 ( 1 ) " " "*"
## 12 ( 1 ) " " "*"
## 13 ( 1 ) "*" "*"
## 14 ( 1 ) "*" "*"
## 15 ( 1 ) "*" "*"
## 16 ( 1 ) "*" "*"
## 17 ( 1 ) "*" "*"
## 18 ( 1 ) "*" "*"
## 19 ( 1 ) "*" "*"
## 20 ( 1 ) "*" "*"
## 21 ( 1 ) "*" "*"
## 22 ( 1 ) "*" "*"
##      Diabetes Has.a.regular.healthcare.provider Heavy.drinking
## 1 ( 1 ) " " " " " "
## 2 ( 1 ) " " " " "*"
## 3 ( 1 ) " " " " "*"
## 4 ( 1 ) " " " " "*"
## 5 ( 1 ) " " " " "*"
## 6 ( 1 ) " " " " "*"
## 7 ( 1 ) " " " " "*"
## 8 ( 1 ) " " " " "*"
## 9 ( 1 ) " " " " "*"
## 10 ( 1 ) " " " " "*"
## 11 ( 1 ) "*" " " "*"
## 12 ( 1 ) "*" " " "*"
## 13 ( 1 ) "*" " " "*"
## 14 ( 1 ) "*" "*" "*"
## 15 ( 1 ) "*" "*" "*"

```

```

## 16 ( 1 ) "*"      " "      "*"
## 17 ( 1 ) "*"      " "      "*"
## 18 ( 1 ) "*"      "*"      "*"
## 19 ( 1 ) "*"      "*"      "*"
## 20 ( 1 ) "*"      "*"      "*"
## 21 ( 1 ) "*"      "*"      "*"
## 22 ( 1 ) "*"      "*"      "*"
##      Influenza.immunization.in.the.past.12.months
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) " "
## 8 ( 1 ) " "
## 9 ( 1 ) " "
## 10 ( 1 ) " "
## 11 ( 1 ) "*"
## 12 ( 1 ) "*"
## 13 ( 1 ) " "
## 14 ( 1 ) " "
## 15 ( 1 ) " "
## 16 ( 1 ) " "
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##      Life.satisfaction..satisfied.or.very.satisfied Mood.disorder
## 1 ( 1 ) " "      " "
## 2 ( 1 ) " "      " "
## 3 ( 1 ) " "      " "
## 4 ( 1 ) " "      " "
## 5 ( 1 ) " "      " "
## 6 ( 1 ) " "      " "
## 7 ( 1 ) " "      " "
## 8 ( 1 ) " "      " "
## 9 ( 1 ) " "      " "
## 10 ( 1 ) "*"      " "
## 11 ( 1 ) " "      " "
## 12 ( 1 ) " "      " "
## 13 ( 1 ) "*"      " "
## 14 ( 1 ) "*"      " "
## 15 ( 1 ) "*"      " "
## 16 ( 1 ) "*"      "*"
## 17 ( 1 ) "*"      "*"
## 18 ( 1 ) "*"      "*"
## 19 ( 1 ) "*"      "*"
## 20 ( 1 ) "*"      "*"
## 21 ( 1 ) "*"      "*"
## 22 ( 1 ) "*"      "*"
##      Perceived.health..fair.or.poor

```

```

## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) " "
## 8 ( 1 ) " "
## 9 ( 1 ) " "
## 10 ( 1 ) " "
## 11 ( 1 ) "*"
## 12 ( 1 ) "*"
## 13 ( 1 ) " "
## 14 ( 1 ) " "
## 15 ( 1 ) " "
## 16 ( 1 ) " "
## 17 ( 1 ) " "
## 18 ( 1 ) " "
## 19 ( 1 ) " "
## 20 ( 1 ) " "
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##
## Perceived.health..very.good.or.excellent
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) "*"
## 6 ( 1 ) "*"
## 7 ( 1 ) "*"
## 8 ( 1 ) "*"
## 9 ( 1 ) "*"
## 10 ( 1 ) "*"
## 11 ( 1 ) "*"
## 12 ( 1 ) "*"
## 13 ( 1 ) "*"
## 14 ( 1 ) "*"
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##
## Perceived.life.stress..most.days.quite.a.bit.or.extremely.stressful
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) " "
## 8 ( 1 ) " "

```

```

## 9 ( 1 ) " "
## 10 ( 1 ) " "
## 11 ( 1 ) " "
## 12 ( 1 ) " "
## 13 ( 1 ) "*"
## 14 ( 1 ) "*"
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##
## Perceived.mental.health..fair.or.poor
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) " "
## 8 ( 1 ) " "
## 9 ( 1 ) " "
## 10 ( 1 ) " "
## 11 ( 1 ) "*"
## 12 ( 1 ) "*"
## 13 ( 1 ) " "
## 14 ( 1 ) " "
## 15 ( 1 ) " "
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##
## Perceived.mental.health..very.good.or.excellent
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) "*"
## 7 ( 1 ) " "
## 8 ( 1 ) "*"
## 9 ( 1 ) "*"
## 10 ( 1 ) "*"
## 11 ( 1 ) " "
## 12 ( 1 ) "*"
## 13 ( 1 ) "*"
## 14 ( 1 ) "*"
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"

```

```

## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##      Physical.activity..150.minutes.per.week..adult..18.years.and.over.
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) " "
## 8 ( 1 ) " "
## 9 ( 1 ) " "
## 10 ( 1 ) " "
## 11 ( 1 ) " "
## 12 ( 1 ) " "
## 13 ( 1 ) " "
## 14 ( 1 ) " "
## 15 ( 1 ) " "
## 16 ( 1 ) " "
## 17 ( 1 ) " "
## 18 ( 1 ) " "
## 19 ( 1 ) " "
## 20 ( 1 ) " "
## 21 ( 1 ) " "
## 22 ( 1 ) "*"
##      Physical.activity..average.60.minutes.per.day..youth..12.to.17.years.old.
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) " "
## 8 ( 1 ) " "
## 9 ( 1 ) " "
## 10 ( 1 ) " "
## 11 ( 1 ) " "
## 12 ( 1 ) " "
## 13 ( 1 ) " "
## 14 ( 1 ) "*"
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##      Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong
## 1 ( 1 ) " "

```

```

## 2 ( 1 ) "*"
## 3 ( 1 ) "*"
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) "*"
## 8 ( 1 ) "*"
## 9 ( 1 ) "*"
## 10 ( 1 ) "*"
## 11 ( 1 ) "*"
## 12 ( 1 ) "*"
## 13 ( 1 ) "*"
## 14 ( 1 ) "*"
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"
##          amenity
## 1 ( 1 ) " "
## 2 ( 1 ) " "
## 3 ( 1 ) " "
## 4 ( 1 ) " "
## 5 ( 1 ) " "
## 6 ( 1 ) " "
## 7 ( 1 ) " "
## 8 ( 1 ) " "
## 9 ( 1 ) " "
## 10 ( 1 ) " "
## 11 ( 1 ) "*"
## 12 ( 1 ) "*"
## 13 ( 1 ) "*"
## 14 ( 1 ) " "
## 15 ( 1 ) "*"
## 16 ( 1 ) "*"
## 17 ( 1 ) "*"
## 18 ( 1 ) "*"
## 19 ( 1 ) "*"
## 20 ( 1 ) "*"
## 21 ( 1 ) "*"
## 22 ( 1 ) "*"

```

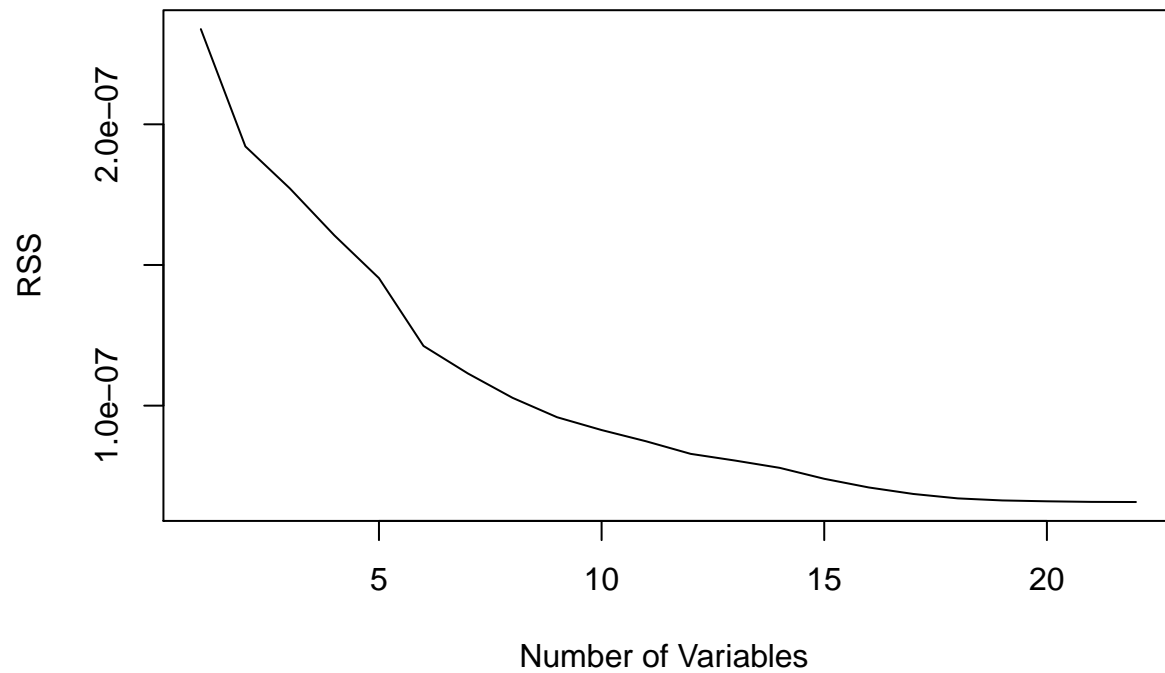
```
reg.summary$RSS
```

```

## [1] 2.338390e-07 1.921443e-07 1.772018e-07 1.605087e-07 1.453314e-07
## [6] 1.212239e-07 1.114120e-07 1.027774e-07 9.589068e-08 9.132671e-08
## [11] 8.732585e-08 8.286091e-08 8.047718e-08 7.789055e-08 7.399657e-08
## [16] 7.091652e-08 6.860342e-08 6.703113e-08 6.630044e-08 6.599838e-08
## [21] 6.579273e-08 6.574797e-08

```

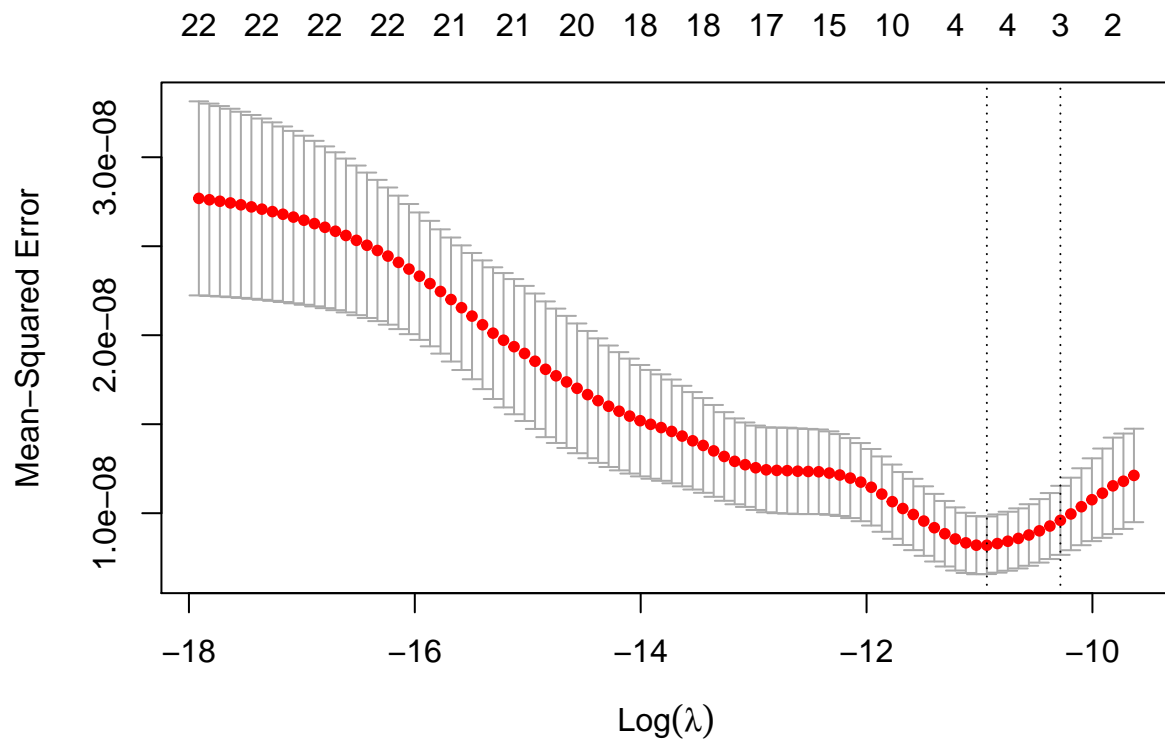
```
plot(reg.summary$rss ,xlab="Number of Variables ",ylab="RSS", type="l")
```



```
# plot suggests to keep ~8 predictors

set.seed(123)
cv.lasso <- cv.glmnet(x, y, alpha = 1)
# Fit the final model on the training data
model <- glmnet(x, y, alpha = 1, lambda = cv.lasso$lambda.min)

plot(cv.lasso)
```

```
cv.lasso$lambda.min
```

```
## [1] 1.782782e-05
```

```
coef(cv.lasso, cv.lasso$lambda.min)
```

```
## 23 x 1 sparse Matrix of class "dgCMatrix"
##
## (Intercept) 1
## Arthritis..15.years.and.over. 0.0007364359
## Asthma .
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese -0.0003737596
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...overweight .
## Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over. .
## Current.smoker..daily .
## Current.smoker..daily.or.occasional .
## Diabetes .
## Has.a.regular.healthcare.provider .
## Heavy.drinking -0.0006730223
## Influenza.immunization.in.the.past.12.months .
## Life.satisfaction..satisfied.or.very.satisfied .
## Mood.disorder .
## Perceived.health..fair.or.poor .
## Perceived.health..very.good.or.excellent .
## Perceived.life.stress..most.days.quite.a.bit.or.extremely.stressful .
```

```
## Perceived.mental.health..fair.or.poor .
## Perceived.mental.health..very.good.or.excellent .
## Physical.activity..150.minutes.per.week..adult..18.years.and.over. .
## Physical.activity..average.60.minutes.per.day..youth..12.to.17.years.old. .
## Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong -0.0005205899
## amenity 0.0000444202
```

```
# looks like the appropriate value for predictors is around 4 or 5.
```

```
#beta regression model...
```

```
library(betareg)
final_fatalities$FATAL_prop <- final_fatalities$FATAL_prop+0.00001

beta_model2 <- betareg(FATAL_prop~Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese +
  Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over. +
  Heavy.drinking + Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong +
  amenity, data = final_fatalities, link = "loglog")
summary(beta_model2)
```

```
##
## Call:
## betareg(formula = FATAL_prop ~ Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese +
##   Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over. +
##   Heavy.drinking + Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong +
##   amenity, data = final_fatalities, link = "loglog")
##
## Standardized weighted residuals 2:
##      Min      1Q  Median      3Q      Max
## -1.9186 -0.9142 -0.0037  0.7229  2.0430
##
## Coefficients (mean model with loglog link):
##
##                                     Estimate
## (Intercept)                        -1.230843
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese -0.258279
## Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over.          -1.199511
## Heavy.drinking                      -0.446035
## Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong      -1.038029
## amenity                             0.009807
##                                     Std. Error
## (Intercept)                        0.251221
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese  0.324577
## Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over.            0.888778
## Heavy.drinking                     0.465326
## Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong       0.345623
## amenity                           0.094368
##                                     z value
## (Intercept)                       -4.899
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese -0.796
## Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over.            -1.350
## Heavy.drinking                     -0.959
## Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong      -3.003
## amenity                           0.104
##                                     Pr(>|z|)
## (Intercept)                       9.61e-07
```

```

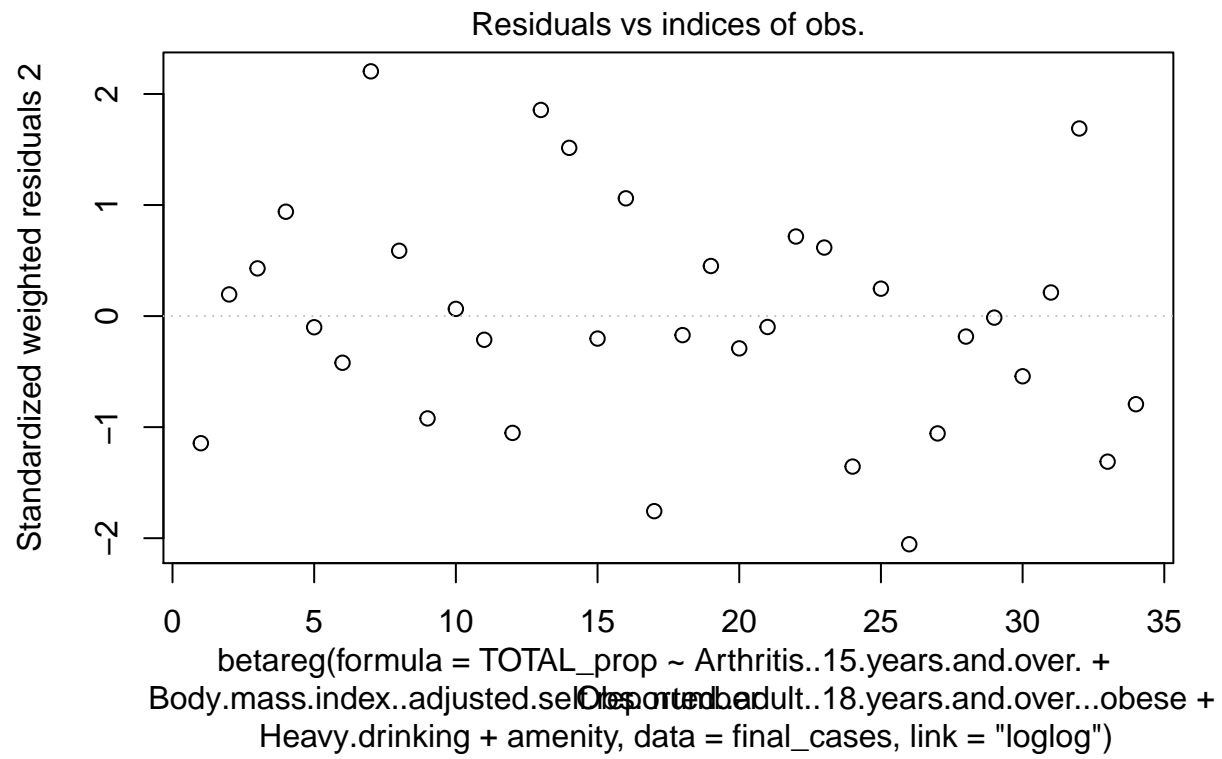
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese 0.42618
## Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over. 0.17714
## Heavy.drinking 0.33779
## Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong 0.00267
## amenity 0.91723
##
## (Intercept) ***
## Body.mass.index..adjusted.self.reported..adult..18.years.and.over...obese
## Chronic.obstructive.pulmonary.disease..COPD..35.years.and.over.
## Heavy.drinking
## Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong **
## amenity
##
## Phi coefficients (precision model with identity link):
##      Estimate Std. Error z value Pr(>|z|)
## (phi)    20824      5413   3.847 0.00012 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Type of estimator: ML (maximum likelihood)
## Log-likelihood: 288.2 on 7 Df
## Pseudo R-squared: 0.5583
## Number of iterations: 5000 (BFGS) + 16 (Fisher scoring)

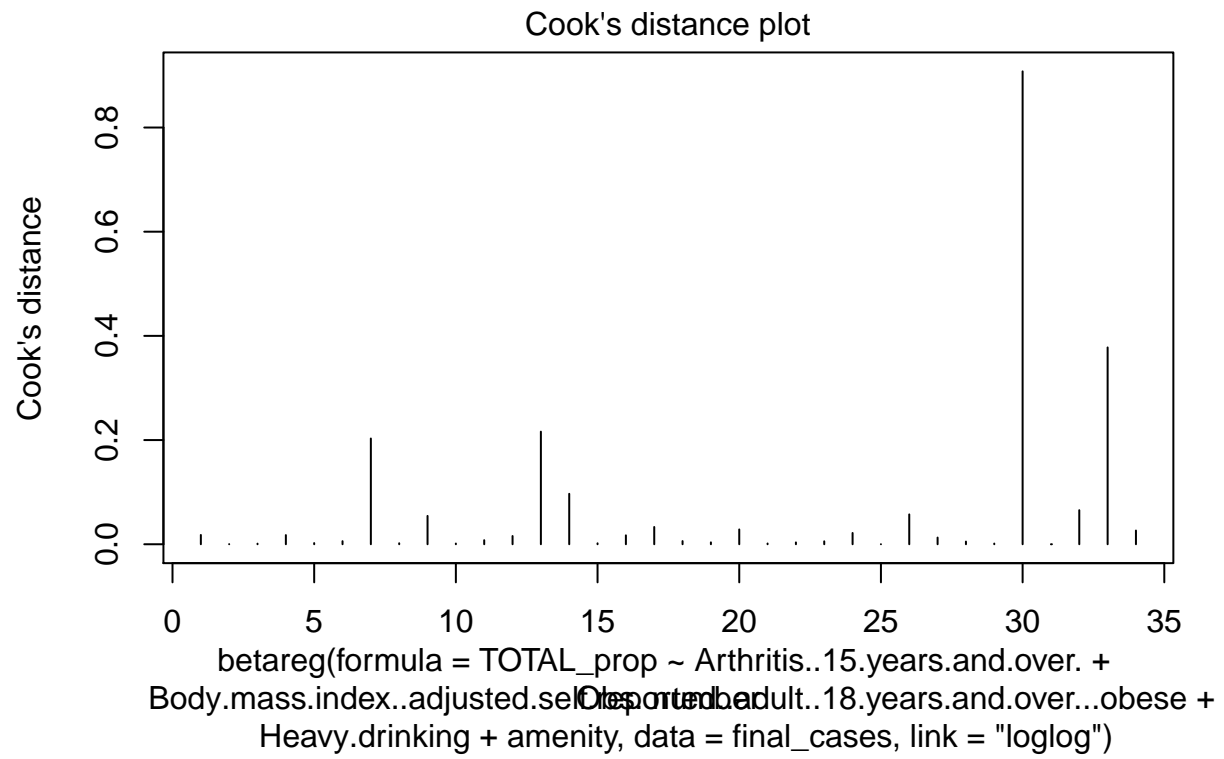
# checking the links.
sapply(c("logit", "probit", "cauchit", "loglog", "cloglog"),
function(x) logLik(update(beta_model, link = x)))

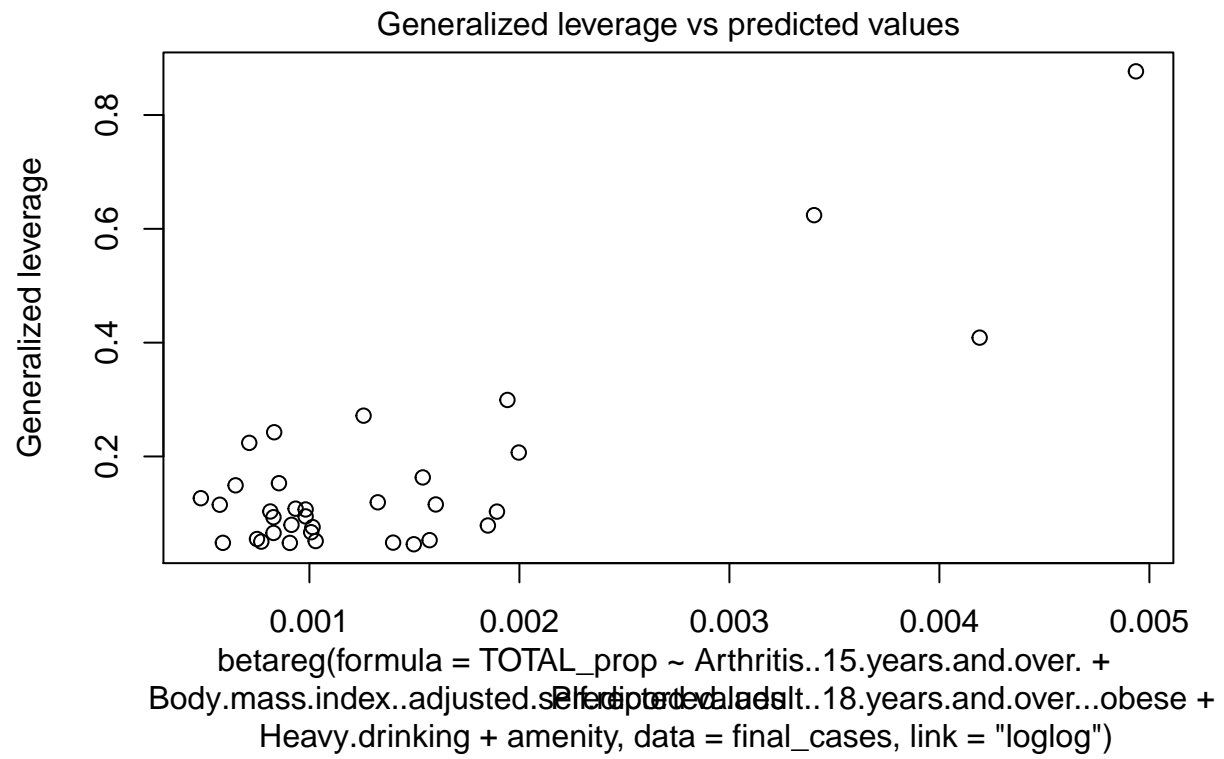
##      logit      probit    cauchit    loglog    cloglog
## 207.0923 207.3729 204.3610 207.5823 207.0895

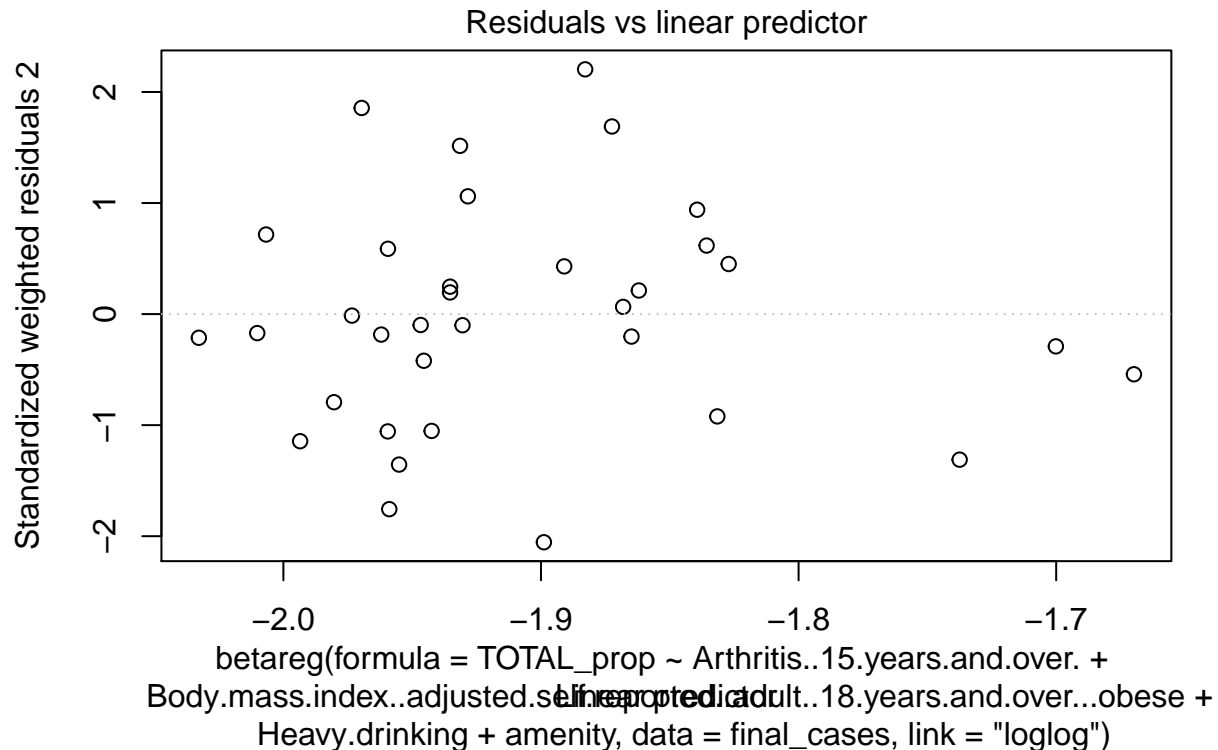
# the highest Loglikelihood is when using the log-log link. Let us use that one.
# diagnostic plots.
plot(beta_model)

```









```
ggplot(final_fatalities, aes(x=Sense.of.belonging.to.local.community..somewhat.strong.or.very.strong, y=
  theme_classic()+
  geom_point(alpha = 0.9, aes(colour=amenity), size=4) +
  xlab("Proportion of PHU which reports a strong sense of belonging to their community")+
  ylab("COVID-19 fatalities (proportion)")+
  labs(col = "Amenity Richness")+
  ggtitle("COVID-19 Fatalities in Ontario PHU's")+
  scale_color_gradientn(colors = c("blue", "gold", "red")) +
  stat_smooth(method = "glm", formula = y ~ log(x), se = T, col = "gray", fill="lightgray") +
  #ggrepel::geom_label_repel(data=data %>% filter(hbp.prop < 0.16), aes(label = Location, fill=factor(amenity)))
  #geom_label_repel( data=data %>% filter(FATAL_prop > 0.00035),aes(label=Reporting_PHU, fill = factor(amenity)))
  # geom_label_repel( data=data %>% filter(hbp.prop < 0.16 & hbp.prop > 0.14),aes(label=Location, fill = factor(amenity)))
  #geom_label_repel( data=data %>% filter(hbp.prop == 0.179),aes(label=Location, fill = factor(amenity)))
  guides(fill=FALSE)+
  scale_fill_manual(values = setNames(c( "red"), levels(data$amenity)))
```

