

Model 03

2024-06-17

#Importing the necessary libraries

```
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

library(tidyr)
library(dplyr)
library(caret)

## Warning: package 'caret' was built under R version 4.3.3

## Loading required package: ggplot2

## Warning: package 'ggplot2' was built under R version 4.3.3

## Loading required package: lattice

library(rpart)
library(partykit)

## Warning: package 'partykit' was built under R version 4.3.3

## Loading required package: grid

## Loading required package: libcoin

## Warning: package 'libcoin' was built under R version 4.3.3

## Loading required package: mvtnorm

## Warning: package 'mvtnorm' was built under R version 4.3.3

library(prettyR)
```

Loading the data file from Wave 2 interviews to calculate the BMI

```
load("34921-0001-Data.rda")

da34921.0001 <- da34921.0001 %>%
  mutate(
    OBESITY = case_when(
      ((WEIGHT)/(HEIGHT*HEIGHT) * 703) >= 30.000 ~ 1,
      ((WEIGHT)/(HEIGHT*HEIGHT) * 703) < 30.000 ~ 0
    )
  )

obesity <- da34921.0001 %>% select(ID, OBESITY)
head(obesity)
```

##	ID	OBESITY
## 1	100005	0
## 2	100033	1
## 3	100067	0
## 4	100080	1
## 5	100149	1
## 6	100154	0

Loading and Processing the Independent Social Network Variables to calculate Bridge from WAVE 1.

```
load("20541-0001-Data.rda")
load("20541-0004-Data.rda")

da20541.0001 <- da20541.0001 %>%
  select (ID, HEARN_RECODE, GENDER, AGE, RACE_RECODE, ETHGRP, COMBUILD,
  DEGREE_RECODE, HISPANIC, MARITLST, JOBSTAT_1, PHYSHLTH, MNTLHLTH, ATNDSERV )

da20541.0001 <- da20541.0001 %>%
  mutate(DEGREE_RECODE = if_else(DEGREE_RECODE == "(-2) don't know", NA,
  DEGREE_RECODE),
  HEARN_RECODE = if_else(HEARN_RECODE == "(-2) don't know", NA,
  HEARN_RECODE),
  RACE_RECODE = if_else(RACE_RECODE == "(-2) don't know", NA,
  RACE_RECODE))

head(da20541.0001)
```

##	ID	HEARN_RECODE	GENDER	AGE	RACE_RECODE
## 1	100005	(4) 100k or higher	(2) female	62	(1) white/caucasian
## 2	100033	(2) 25,000-49,999	(2) female	79	(1) white/caucasian
## 3	100080	(3) 50,000-99,999	(1) male	60	(1) white/caucasian
## 4	100154	(2) 25,000-49,999	(2) female	78	(1) white/caucasian
## 5	100203	<NA>	(2) female	61	(1) white/caucasian
## 6	100359	(3) 50,000-99,999	(1) male	75	(1) white/caucasian

```

##          ETHGRP          COMBUILD
DEGREE_RECODE
## 1          (1) white          (3) average          (5)
masters
## 2          (1) white (4) above average (2) high school
diploma/equivalency
## 3          (1) white          (3) average (2) high school
diploma/equivalency
## 4          (1) white          (3) average (2) high school
diploma/equivalency
## 5 (3) hispanic, non-black          (3) average          (1)
none
## 6          (1) white          (3) average (2) high school
diploma/equivalency
##  HISPANIC  MARITLST JOBSTAT_1  PHYSHLTH  MNTLHLTH
## 1  (0) no (1) married  (1) yes (4) very good (4) very good
## 2  (0) no (5) widowed  (0) no (4) very good (4) very good
## 3  (0) no (1) married  (1) yes  (3) good (5) excellent
## 4  (0) no (1) married  (0) no  (3) good  (3) good
## 5  (1) yes (5) widowed  (1) yes  (1) poor  (2) fair
## 6  (0) no (1) married  (0) no  (2) fair  (3) good
##          ATNDSERV
## 1  (3) several times a year
## 2  (1) less than once a year
## 3          (5) every week
## 4  (6) several times a week
## 5          (0) never
## 6  (6) several times a week

nrow(da20541.0001)

## [1] 3005

da20541.0004 <- da20541.0004 %>%
  group_by(ID) %>%
  filter(n() > 2) %>%
  ungroup()

da20541.0004 <- da20541.0004 %>%
  pivot_longer(
    cols = starts_with("TALKFREQ"),
    names_to = "TALKFREQ",
    values_to = "FREQ"
  )

da20541.0004 <- da20541.0004 %>%
  group_by(ID) %>%
  summarize(
    BRIDGE = if_else(any(FREQ == '(0) have never spoken to each other', na.rm
= TRUE), 1, 0),
    HEALTHDISCUSSIONS = if_else(any(HEALTHTALK == '(3) very likely', na.rm =

```

```

TRUE), 1, 0),
  LIVEALONE = if_else(any(LIVewith == '(1) yes -- lives in the same
household', na.rm = TRUE), 0,1))

head(da20541.0004)

## # A tibble: 6 × 4
##   ID      BRIDGE HEALTHDISCUSSIONS LIVEALONE
##   <fct>    <dbl>             <dbl>    <dbl>
## 1 100005      1              1          0
## 2 100033      0              1          0
## 3 100080      1              1          0
## 4 100154      1              1          0
## 5 100203      0              1          0
## 6 100359      0              1          0

nrow(da20541.0004)

## [1] 2522

modeldata <- da20541.0001 %>%
  left_join(da20541.0004, by = "ID")

modeldata <- modeldata %>%
  left_join(obesity, by = "ID")

modeldata <- na.omit(modeldata)
modeldata <- modeldata %>% select(-ID)

modeldata$BRIDGE <- as.factor(modeldata$BRIDGE)
modeldata$HEALTHDISCUSSIONS <- as.factor(modeldata$HEALTHDISCUSSIONS)
modeldata$LIVEALONE <- as.factor(modeldata$LIVEALONE)
modeldata$OBESITY <- as.factor(modeldata$OBESITY)
head(modeldata)

##           HEARN_RECODE      GENDER AGE      RACE_RECODE      ETHGRP
## 1 (4) 100k or higher (2) female  62 (1) white/caucasian (1) white
## 2 (2) 25,000-49,999 (2) female  79 (1) white/caucasian (1) white
## 3 (3) 50,000-99,999 (1) male   60 (1) white/caucasian (1) white
## 4 (2) 25,000-49,999 (2) female  78 (1) white/caucasian (1) white
## 7 (2) 25,000-49,999 (1) male   80 (1) white/caucasian (1) white
## 9 (3) 50,000-99,999 (2) female  59 (1) white/caucasian (1) white
##           COMBUILD           DEGREE_RECODE HISPANIC
## 1 (3) average (5) masters (0) no (1)
married
## 2 (4) above average (2) high school diploma/equivalency (0) no (5)
widowed
## 3 (3) average (2) high school diploma/equivalency (0) no (1)
married
## 4 (3) average (2) high school diploma/equivalency (0) no (1)

```

```

married
## 7 (4) above average (2) high school diploma/equivalency (0) no (5)
widowed
## 9 (3) average (2) high school diploma/equivalency (0) no (1)
married
## JOBSTAT_1 PHYSHLTH MNTLHLTH ATNDSERV
BRIDGE
## 1 (1) yes (4) very good (4) very good (3) several times a year
1
## 2 (0) no (4) very good (4) very good (1) less than once a year
0
## 3 (1) yes (3) good (5) excellent (5) every week
1
## 4 (0) no (3) good (3) good (6) several times a week
1
## 7 (0) no (3) good (3) good (5) every week
0
## 9 (1) yes (4) very good (4) very good (2) about once or twice a year
1
## HEALTHDISCUSSIONS LIVEALONE OBESITY
## 1 1 0 0
## 2 1 0 1
## 3 1 0 1
## 4 1 0 0
## 7 1 1 0
## 9 1 0 0

```

Creating Data Partition for 70% Training Data and 30% Testing Data

```

library(rpart)
library(caret)

set.seed(19032023)

index <- createDataPartition(modeldata$OBESITY,
                             p=0.7,
                             list=FALSE,
                             times = 1
                             )

modeldata.train <- modeldata[index,]
modeldata.test <- modeldata[-index,]

nrow(modeldata.train)

## [1] 995

nrow(modeldata.test)

## [1] 425

```

Applying Logistic Regression on to find the association between Bridge and Obesity.

```
model.lr <- glm(OBESITY ~ ., data = modeldata.train, family = "binomial")

summary.lr <- summary(model.lr)
```

p-value for Bridge variable

```
print(summary.lr)

##
## Call:
## glm(formula = OBESITY ~ ., family = "binomial", data = modeldata.train)
##
## Coefficients: (1 not defined because of singularities)
##
Estimate
## (Intercept)
1.924814
## HEARN_RECODE(1) 0-24,999
-0.053225
## HEARN_RECODE(2) 25,000-49,999
-0.106097
## HEARN_RECODE(3) 50,000-99,999
0.095769
## HEARN_RECODE(4) 100k or higher
-0.171060
## GENDER(2) female
-0.198780
## AGE
-0.050390
## RACE_RECODE(2) black/african american
0.470805
## RACE_RECODE(3) asian, pacific islander, american indian or alaskan native
0.142746
## ETHGRP(2) black
NA
## ETHGRP(3) hispanic, non-black
-0.015314
## ETHGRP(4) other
-0.486292
## COMBUILD(2) below average
0.712499
## COMBUILD(3) average
0.706080
## COMBUILD(4) above average
0.559192
## COMBUILD(5) far above average
1.100602
```

DEGREE_RECODE(2) high school diploma/equivalency
0.256349
DEGREE_RECODE(3) associates
0.238307
DEGREE_RECODE(4) bachelors
0.205095
DEGREE_RECODE(5) masters
-0.313473
DEGREE_RECODE(6) law, md or phd
-0.869972
HISPANIC(1) yes
0.018341
MARITLST(2) living with a partner
0.480805
MARITLST(3) separated
-0.308887
MARITLST(4) divorced
0.422486
MARITLST(5) widowed
0.294636
MARITLST(6) never married
0.352495
JOBSTAT_1(1) yes
0.131454
PHYSHLTH(2) fair
0.056148
PHYSHLTH(3) good
-0.474518
PHYSHLTH(4) very good
-1.033825
PHYSHLTH(5) excellent
-1.500161
MNTLHLTH(2) fair
-0.068229
MNTLHLTH(3) good
-0.278326
MNTLHLTH(4) very good
0.213802
MNTLHLTH(5) excellent
0.137781
ATNDSERV(1) less than once a year
-0.501595
ATNDSERV(2) about once or twice a year
0.002732
ATNDSERV(3) several times a year
0.437711
ATNDSERV(4) about once a month
0.057627
ATNDSERV(5) every week
0.025598

```
## ATNDSERV(6) several times a week
0.210739
## BRIDGE1
-0.306024
## HEALTHDISCUSSIONS1
0.845916
## LIVEALONE1
-0.149124
##
Std. Error
## (Intercept)
1.390537
## HEARN_RECODE(1) 0-24,999
0.271365
## HEARN_RECODE(2) 25,000-49,999
0.265373
## HEARN_RECODE(3) 50,000-99,999
0.261951
## HEARN_RECODE(4) 100k or higher
0.329871
## GENDER(2) female
0.153110
## AGE
0.011473
## RACE_RECODE(2) black/african american
0.203240
## RACE_RECODE(3) asian, pacific islander, american indian or alaskan native
0.465453
## ETHGRP(2) black
NA
## ETHGRP(3) hispanic, non-black
1.507797
## ETHGRP(4) other
0.698808
## COMBUILD(2) below average
0.720302
## COMBUILD(3) average
0.660904
## COMBUILD(4) above average
0.669795
## COMBUILD(5) far above average
0.713536
## DEGREE_RECODE(2) high school diploma/equivalency
0.213526
## DEGREE_RECODE(3) associates
0.244791
## DEGREE_RECODE(4) bachelors
0.273448
## DEGREE_RECODE(5) masters
0.314176
```


DEGREE_RECODE(6) law, md or phd
0.599202
HISPANIC(1) yes
1.474397
MARITLST(2) living with a partner
0.535377
MARITLST(3) separated
0.609400
MARITLST(4) divorced
0.310796
MARITLST(5) widowed
0.272941
MARITLST(6) never married
0.497814
JOBSTAT_1(1) yes
0.165269
PHYSHLTH(2) fair
0.381309
PHYSHLTH(3) good
0.376144
PHYSHLTH(4) very good
0.382791
PHYSHLTH(5) excellent
0.431210
MNTLHLTH(2) fair
0.794135
MNTLHLTH(3) good
0.770150
MNTLHLTH(4) very good
0.767135
MNTLHLTH(5) excellent
0.777149
ATNDSERV(1) less than once a year
0.420240
ATNDSERV(2) about once or twice a year
0.299118
ATNDSERV(3) several times a year
0.274187
ATNDSERV(4) about once a month
0.289693
ATNDSERV(5) every week
0.223775
ATNDSERV(6) several times a week
0.264050
BRIDGE1
0.145639
HEALTHDISCUSSIONS1
0.572021
LIVEALONE1
0.255613

```
##
z value
## (Intercept)
1.384
## HEARN_RECODE(1) 0-24,999
-0.196
## HEARN_RECODE(2) 25,000-49,999
-0.400
## HEARN_RECODE(3) 50,000-99,999
0.366
## HEARN_RECODE(4) 100k or higher
-0.519
## GENDER(2) female
-1.298
## AGE
-4.392
## RACE_RECODE(2) black/african american
2.316
## RACE_RECODE(3) asian, pacific islander, american indian or alaskan native
0.307
## ETHGRP(2) black
NA
## ETHGRP(3) hispanic, non-black
-0.010
## ETHGRP(4) other
-0.696
## COMBUILD(2) below average
0.989
## COMBUILD(3) average
1.068
## COMBUILD(4) above average
0.835
## COMBUILD(5) far above average
1.542
## DEGREE_RECODE(2) high school diploma/equivalency
1.201
## DEGREE_RECODE(3) associates
0.974
## DEGREE_RECODE(4) bachelors
0.750
## DEGREE_RECODE(5) masters
-0.998
## DEGREE_RECODE(6) law, md or phd
-1.452
## HISPANIC(1) yes
0.012
## MARITLST(2) living with a partner
0.898
## MARITLST(3) separated
-0.507
```

```

## MARITLST(4) divorced
1.359
## MARITLST(5) widowed
1.079
## MARITLST(6) never married
0.708
## JOBSTAT_1(1) yes
0.795
## PHYSHLTH(2) fair
0.147
## PHYSHLTH(3) good
-1.262
## PHYSHLTH(4) very good
-2.701
## PHYSHLTH(5) excellent
-3.479
## MNTLHLTH(2) fair
-0.086
## MNTLHLTH(3) good
-0.361
## MNTLHLTH(4) very good
0.279
## MNTLHLTH(5) excellent
0.177
## ATNDSERV(1) less than once a year
-1.194
## ATNDSERV(2) about once or twice a year
0.009
## ATNDSERV(3) several times a year
1.596
## ATNDSERV(4) about once a month
0.199
## ATNDSERV(5) every week
0.114
## ATNDSERV(6) several times a week
0.798
## BRIDGE1
-2.101
## HEALTHDISCUSSIONS1
1.479
## LIVEALONE1
-0.583
##
Pr(>|z|)
## (Intercept)
0.166290
## HEARN_RECODE(1) 0-24,999
0.844503
## HEARN_RECODE(2) 25,000-49,999
0.689300

```

```
## HEARN_RECODE(3) 50,000-99,999
0.714663
## HEARN_RECODE(4) 100k or higher
0.604062
## GENDER(2) female
0.194191
## AGE
1.12e-05
## RACE_RECODE(2) black/african american
0.020531
## RACE_RECODE(3) asian, pacific islander, american indian or alaskan native
0.759086
## ETHGRP(2) black
NA
## ETHGRP(3) hispanic, non-black
0.991896
## ETHGRP(4) other
0.486499
## COMBUILD(2) below average
0.322581
## COMBUILD(3) average
0.285360
## COMBUILD(4) above average
0.403790
## COMBUILD(5) far above average
0.122961
## DEGREE_RECODE(2) high school diploma/equivalency
0.229925
## DEGREE_RECODE(3) associates
0.330298
## DEGREE_RECODE(4) bachelors
0.453235
## DEGREE_RECODE(5) masters
0.318395
## DEGREE_RECODE(6) law, md or phd
0.146533
## HISPANIC(1) yes
0.990075
## MARITLST(2) living with a partner
0.369149
## MARITLST(3) separated
0.612246
## MARITLST(4) divorced
0.174030
## MARITLST(5) widowed
0.280372
## MARITLST(6) never married
0.478891
## JOBSTAT_1(1) yes
0.426386
```

```
## PHYSHLTH(2) fair
0.882934
## PHYSHLTH(3) good
0.207117
## PHYSHLTH(4) very good
0.006918
## PHYSHLTH(5) excellent
0.000503
## MNTLHLTH(2) fair
0.931533
## MNTLHLTH(3) good
0.717807
## MNTLHLTH(4) very good
0.780474
## MNTLHLTH(5) excellent
0.859280
## ATNDSERV(1) less than once a year
0.232638
## ATNDSERV(2) about once or twice a year
0.992714
## ATNDSERV(3) several times a year
0.110400
## ATNDSERV(4) about once a month
0.842323
## ATNDSERV(5) every week
0.908928
## ATNDSERV(6) several times a week
0.424811
## BRIDGE1
0.035619
## HEALTHDISCUSSIONS1
0.139188
## LIVEALONE1
0.559626
##
## (Intercept)
## HEARN_RECODE(1) 0-24,999
## HEARN_RECODE(2) 25,000-49,999
## HEARN_RECODE(3) 50,000-99,999
## HEARN_RECODE(4) 100k or higher
## GENDER(2) female
## AGE
***
## RACE_RECODE(2) black/african american
*
## RACE_RECODE(3) asian, pacific islander, american indian or alaskan native
## ETHGRP(2) black
## ETHGRP(3) hispanic, non-black
## ETHGRP(4) other
## COMBUILD(2) below average
```

```

## COMBUILD(3) average
## COMBUILD(4) above average
## COMBUILD(5) far above average
## DEGREE_RECODE(2) high school diploma/equivalency
## DEGREE_RECODE(3) associates
## DEGREE_RECODE(4) bachelors
## DEGREE_RECODE(5) masters
## DEGREE_RECODE(6) law, md or phd
## HISPANIC(1) yes
## MARITLST(2) living with a partner
## MARITLST(3) separated
## MARITLST(4) divorced
## MARITLST(5) widowed
## MARITLST(6) never married
## JOBSTAT_1(1) yes
## PHYSHLTH(2) fair
## PHYSHLTH(3) good
## PHYSHLTH(4) very good
**
## PHYSHLTH(5) excellent
***
## MNTLHLTH(2) fair
## MNTLHLTH(3) good
## MNTLHLTH(4) very good
## MNTLHLTH(5) excellent
## ATNDSERV(1) less than once a year
## ATNDSERV(2) about once or twice a year
## ATNDSERV(3) several times a year
## ATNDSERV(4) about once a month
## ATNDSERV(5) every week
## ATNDSERV(6) several times a week
## BRIDGE1
*
## HEALTHDISCUSSIONS1
## LIVEALONE1
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##    Null deviance: 1327.1  on 994  degrees of freedom
## Residual deviance: 1208.8  on 951  degrees of freedom
## AIC: 1296.8
##
## Number of Fisher Scoring iterations: 4

names(coef(model.lmr))

## [1] "(Intercept)"
## [2] "HEARN_RECODE(1) 0-24,999"

```

```
## [3] "HEARN_RECODE(2) 25,000-49,999"
## [4] "HEARN_RECODE(3) 50,000-99,999"
## [5] "HEARN_RECODE(4) 100k or higher"
## [6] "GENDER(2) female"
## [7] "AGE"
## [8] "RACE_RECODE(2) black/african american"
## [9] "RACE_RECODE(3) asian, pacific islander, american indian or alaskan
native"
## [10] "ETHGRP(2) black"
## [11] "ETHGRP(3) hispanic, non-black"
## [12] "ETHGRP(4) other"
## [13] "COMBUILD(2) below average"
## [14] "COMBUILD(3) average"
## [15] "COMBUILD(4) above average"
## [16] "COMBUILD(5) far above average"
## [17] "DEGREE_RECODE(2) high school diploma/equivalency"
## [18] "DEGREE_RECODE(3) associates"
## [19] "DEGREE_RECODE(4) bachelors"
## [20] "DEGREE_RECODE(5) masters"
## [21] "DEGREE_RECODE(6) law, md or phd"
## [22] "HISPANIC(1) yes"
## [23] "MARITLST(2) living with a partner"
## [24] "MARITLST(3) separated"
## [25] "MARITLST(4) divorced"
## [26] "MARITLST(5) widowed"
## [27] "MARITLST(6) never married"
## [28] "JOBSTAT_1(1) yes"
## [29] "PHYSHLTH(2) fair"
## [30] "PHYSHLTH(3) good"
## [31] "PHYSHLTH(4) very good"
## [32] "PHYSHLTH(5) excellent"
## [33] "MNTLHLTH(2) fair"
## [34] "MNTLHLTH(3) good"
## [35] "MNTLHLTH(4) very good"
## [36] "MNTLHLTH(5) excellent"
## [37] "ATNDSERV(1) less than once a year"
## [38] "ATNDSERV(2) about once or twice a year"
## [39] "ATNDSERV(3) several times a year"
## [40] "ATNDSERV(4) about once a month"
## [41] "ATNDSERV(5) every week"
## [42] "ATNDSERV(6) several times a week"
## [43] "BRIDGE1"
## [44] "HEALTHDISCUSSIONS1"
## [45] "LIVEALONE1"
```

Odds Ratio and 95% Confidence Interval

```
odds_ratio <- exp(coef(model.lm)[ "BRIDGE1" ])
print(odds_ratio)
```

```

## BRIDGE1
## 0.7363692

conf_int <- exp(confint(model.lr, "BRIDGE1"))

## Waiting for profiling to be done...

print(conf_int)

##      2.5 %      97.5 %
## 0.5530377 0.9791486

predicted.prob.lr <- predict(model.lr, modeldata.test, type = "response")

## Warning in predict.lm(object, newdata, se.fit, scale = 1, type = if (type
== :
## prediction from rank-deficient fit; attr(*, "non-estim") has doubtful
cases

predicted.obesity.lr <- ifelse(predicted.prob.lr > 0.5, 1, 0)

actual.obesity.lr <- modeldata.test$OBESITY
conf.matrix.lr <- table(Predicted = predicted.obesity.lr, Actual =
actual.obesity.lr)

print(conf.matrix.lr)

##           Actual
## Predicted   0   1
##           0 204 111
##           1   57  53

confusionMatrix(factor(predicted.obesity.lr), factor(modeldata.test$OBESITY),
positive = as.character(1))

## Confusion Matrix and Statistics
##
##           Reference
## Prediction   0   1
##           0 204 111
##           1   57  53
##
##           Accuracy : 0.6047
##           95% CI : (0.5565, 0.6515)
##           No Information Rate : 0.6141
##           P-Value [Acc > NIR] : 0.6741
##
##           Kappa : 0.1116
##
##           Mcnemar's Test P-Value : 4.332e-05
##
##           Sensitivity : 0.3232

```



```
##          Specificity : 0.7816
##          Pos Pred Value : 0.4818
##          Neg Pred Value : 0.6476
##          Prevalence : 0.3859
##          Detection Rate : 0.1247
##          Detection Prevalence : 0.2588
##          Balanced Accuracy : 0.5524
##
##          'Positive' Class : 1
##
```

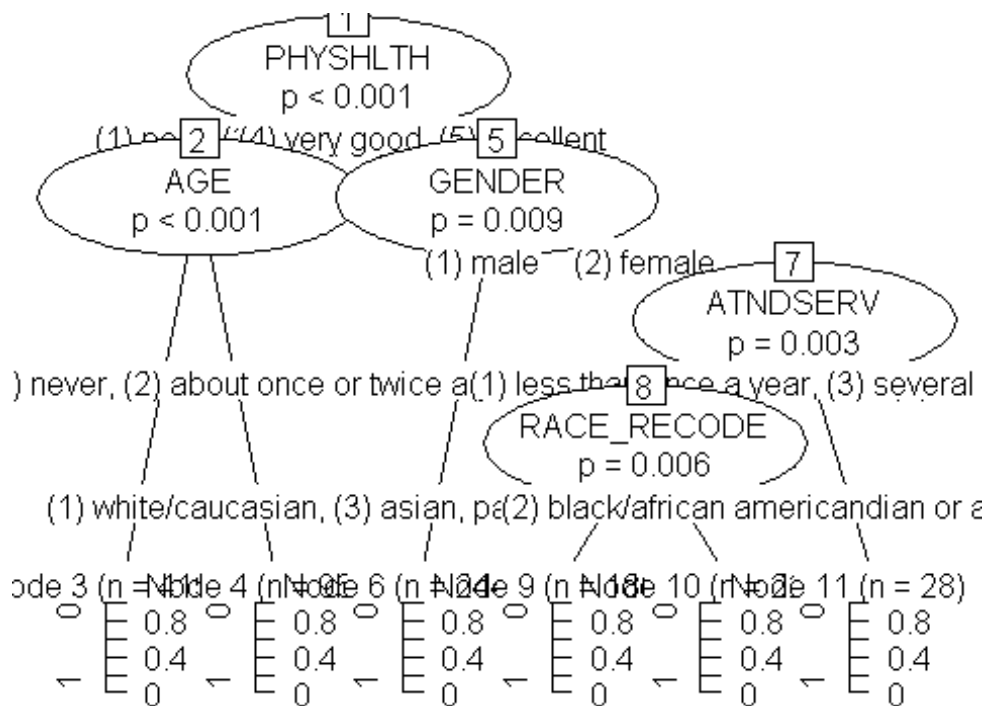
Decision Tree

Conditional Inference Tree implementation using ctree

```
set.seed(123)
```

```
model.dt <- ctree(OBESITY ~ .,
                  data=modeldata.train)
```

```
plot(model.dt)
```



Classification and Regression Tree implementation using rpart

```
rpart.tree <- rpart(OBESITY ~ ., data = modeldata.train, method = "class",
parms = list(split = "information"))
rpart.tree

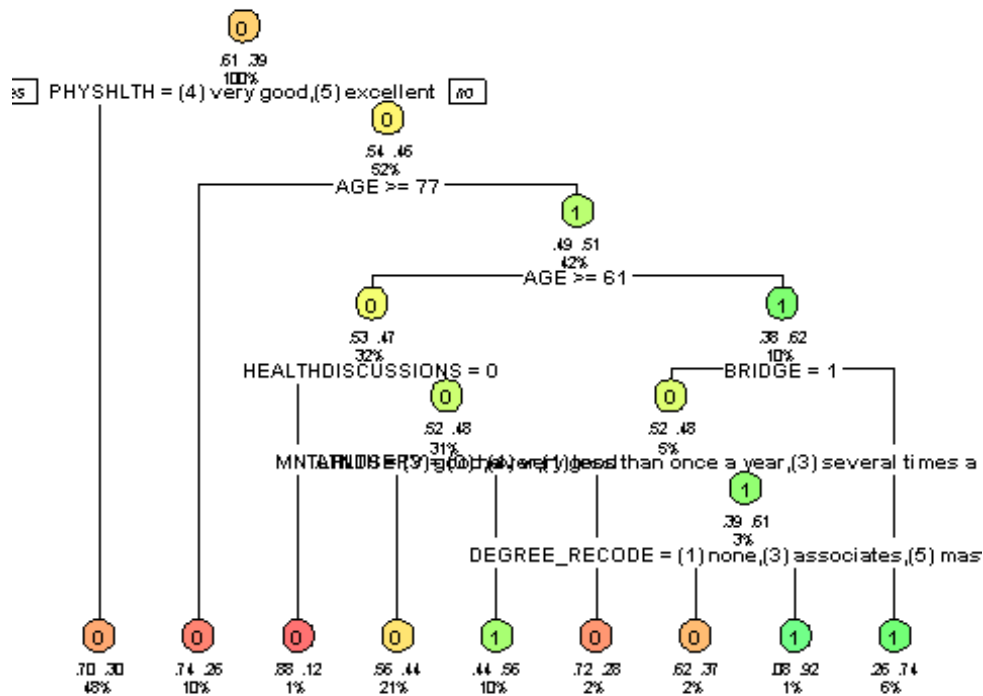
## n= 995
##
## node), split, n, loss, yval, (yprob)
##      * denotes terminal node
##
##  1) root 995 384 0 (0.61407035 0.38592965)
##    2) PHYSHLTH=(4) very good,(5) excellent 480 145 0 (0.69791667
0.30208333) *
##    3) PHYSHLTH=(1) poor,(2) fair,(3) good 515 239 0 (0.53592233
0.46407767)
##      6) AGE>=76.5 96  25 0 (0.73958333 0.26041667) *
##      7) AGE< 76.5 419 205 1 (0.48926014 0.51073986)
##    14) AGE>=60.5 316 150 0 (0.52531646 0.47468354)
##      28) HEALTHDISCUSSIONS=0 8  1 0 (0.87500000 0.12500000) *
##      29) HEALTHDISCUSSIONS=1 308 149 0 (0.51623377 0.48376623)
##    58) MNTLHLTH=(3) good,(4) very good 207  92 0 (0.55555556
0.44444444) *
##    59) MNTLHLTH=(1) poor,(2) fair,(5) excellent 101  44 1
(0.43564356 0.56435644) *
##    15) AGE< 60.5 103  39 1 (0.37864078 0.62135922)
##    30) BRIDGE=1 46  22 0 (0.52173913 0.47826087)
##    60) ATNDSERV=(0) never,(1) less than once a year,(3) several
times a year 18  5 0 (0.72222222 0.27777778) *
##    61) ATNDSERV=(2) about once or twice a year,(4) about once a
month,(5) every week,(6) several times a week 28  11 1 (0.39285714
0.60714286)
##    122) DEGREE_RECODE=(1) none,(3) associates,(5) masters 16  6
0 (0.62500000 0.37500000) *
##    123) DEGREE_RECODE=(2) high school diploma/equivalency,(4)
bachelors 12  1 1 (0.08333333 0.91666667) *
##    31) BRIDGE=0 57  15 1 (0.26315789 0.73684211) *

library(rpart.plot)

## Warning: package 'rpart.plot' was built under R version 4.3.3

rpart.plot(
  rpart.tree,
  type = 2,           # Show split labels at all nodes
  extra = 104,        # Display both percentage and number of observations
  under = TRUE,       # Show the prediction under the node
  cex = 0.6,         # Increase font size
  #tweak = 1.1,       # Fine-tune the overall size of the plot
  box.palette = "RdYlGn", # Add some color for better visualization
```

```
compress = TRUE      # Compress the tree for better fit
)
```



```
library(dplyr)
```

```
importances <- varImp(rpart.tree) %>%
  arrange(desc(Overall))
```

```
importances
```

```
##           Overall
## AGE          25.512077
## PHYSHLTH     23.509638
## DEGREE_RECODE 16.627622
## MNTNLHLTH    14.194606
## BRIDGE       9.781547
## RACE_RECODE  9.413427
## ATNDSERV     7.604941
## ETHGRP       6.454945
## HEARN_RECODE 6.089666
## GENDER       4.015550
## HEALTHDISCUSSIONS 2.288146
## MARITLST     1.691373
## COMBUILD     0.000000
## HISPANIC     0.000000
## JOBSTAT_1    0.000000
## LIVEALONE    0.000000
```