## for\_presentation

## Sara Shao

## 4/11/2021

```
library(tidyverse)
## -- Attaching packages ------ tidyverse 1.3.0 --
## v tibble 3.1.0
                       v purrr
                                 0.3.4
## v tidyr
             1.1.3
                       v dplyr
                                 1.0.5
## v readr
             1.4.0
                       v forcats 0.5.1
## -- Conflicts ----- tidyverse_conflicts() --
## x lubridate::as.difftime() masks base::as.difftime()
## x lubridate::date()
    masks base::date()
## x dplyr::filter()
                            masks stats::filter()
## x readr::guess_encoding() masks rvest::guess_encoding()
## 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(nnet)
library(varhandle)
us_19 <- read_csv('~/df_data/US/us_19.csv')
#col_names <- colnames(us_19)</pre>
#us_19 <- lapply(us_19%>% select(-DEM_AGE), as.factor)
\#us_19 \leftarrow data.frame(matrix(unlist(us_19), nrow=length(us_19), byrow=TRUE))
\#us_19 \leftarrow data.frame(t(us_19))
#colnames(us_19) <- col_names</pre>
set.seed(2021)
us_19_mod <- us_19 %>%
  sample_n(3000) %>%
  mutate(DAST_binary = if_else(DAST_SUM < 3, 0, 1))</pre>
us_19_mod <- us_19_mod %>%
  select(ends_with("NMUYR"), starts_with("DEM"), TOB_LIFE, DAST_CAT, DAST_binary) %>%
  subset(select = -c(DEM_PREG, DEM_PREG_WK, DEM_STDNT_PROGRAM, DEM_STDNT_UNDER, DEM_VET_SERV, DEM_HEALT.
us_19_mod
## # A tibble: 3,000 x 86
      BHYD_NMUYR BUP_NMUYR COD_NMUYR DIHY_NMUYR ELU_NMUYR FENT_NMUYR GAB_NMUYR
##
                               <dbl> <lgl>
                    <dbl>
                                                    <dbl>
                                                               <dbl>
```

NΑ

NA

NA

NA

NA

NA

NA NA

NA NA

NA

NA

## 1 NA

## 2 NA

```
3 NA
                             NA NA
                                                                   0
##
                    NA
                                               NA
                                                         NA
##
   4 NA
                    NA
                             NA NA
                                                NA
                                                         NA
                                                                  NA
##
   5 NA
                    NA
                             NA NA
                                                NA
                                                         NA
                                                                  NA
##
   6 NA
                    NA
                              O NA
                                                NA
                                                         NA
                                                                  NA
##
   7 NA
                    NA
                              O NA
                                                NA
                                                         NA
                                                                  NA
##
   8 NA
                    NA
                             NA NA
                                                NA
                                                         NA
                                                                  NA
##
   9 NA
                    NA
                             NA NA
                                                NA
                                                         NA
                                                                  NA
## 10 NA
                    NA
                             NA NA
                                                NA
                                                         NA
                                                                  NA
    ... with 2,990 more rows, and 79 more variables: HYD_NMUYR <dbl>,
      HYDM_NMUYR <dbl>, KTM_NMUYR <dbl>, METH_NMUYR <dbl>, MORPH_NMUYR <dbl>,
## #
      OXY_NMUYR <dbl>, OXYM_NMUYR <dbl>, PREG_NMUYR <dbl>, SUF_NMUYR <dbl>,
      TAP_NMUYR <1gl>, TRAM_NMUYR <dbl>, ALP_NMUYR <dbl>, BAC_NMUYR <dbl>,
## #
## #
      CHL_NMUYR <dbl>, CLOB_NMUYR <dbl>, CLON_NMUYR <dbl>, CLOR_NMUYR <dbl>,
      DIA_NMUYR <dbl>, EST_NMUYR <dbl>, ESZ_NMUYR <dbl>, FLUR_NMUYR <dbl>,
## #
      LORA_NMUYR <dbl>, MID_NMUYR <dbl>, OXA_NMUYR <dbl>, QUA_NMUYR <lgl>,
## #
## #
      TEM_NMUYR <dbl>, TRI_NMUYR <dbl>, ZAL_NMUYR <dbl>, ZOL_NMUYR <dbl>,
      AMPH_NMUYR <dbl>, ATOM_NMUYR <dbl>, MPHEN_NMUYR <dbl>, MOD_NMUYR <dbl>,
## #
## #
      CANN NMUYR <dbl>, DRON NMUYR <dbl>, NAB NMUYR <dbl>, ACE NMUYR <dbl>,
      ASP_NMUYR <dbl>, DEX_NMUYR <dbl>, DIPH_NMUYR <dbl>, IBU_NMUYR <dbl>,
## #
## #
      LOP_NMUYR <dbl>, NAP_NMUYR <dbl>, OTCOTH_NMUYR <dbl>, PAINREL_NMUYR <dbl>,
## #
      SED_NMUYR <dbl>, STIM_NMUYR <dbl>, THC_NMUYR <dbl>, OP_NMUYR <dbl>,
      GABA_NMUYR <dbl>, DEM_GENDER <dbl>, DEM_AGE <dbl>, DEM_POSTAL <chr>,
## #
      DEM_REGION <dbl>, DEM_INCOME <dbl>, DEM_HOME <dbl>, DEM_GENHEALTH <dbl>,
## #
      DEM_LIMIT <dbl>, DEM_HISPANIC <dbl>, DEM_RACE_AIAN <dbl>,
## #
## #
      DEM_RACE_ASIAN <dbl>, DEM_RACE_BLACK <dbl>, DEM_RACE_NHPI <dbl>,
      DEM_RACE_WHITE <dbl>, DEM_RACE_OTH <dbl>, DEM_MARITAL <dbl>, DEM_EDU <dbl>,
## #
      DEM_STDNT <dbl>, DEM_VET <dbl>, DEM_HEALTH <dbl>, DEM_EMPLOY <dbl>,
## #
## #
      DEM_INSUR <dbl>, DEM_HOSPSTAY <dbl>, DEM_ZIP <dbl>, DEM_STATE <chr>,
      DEM_AGE10 <dbl>, TOB_LIFE <dbl>, DAST_CAT <dbl>, DAST_binary <dbl>
us_19_mod <- us_19_mod %>%
 mutate(DEM_GENDER = factor(DEM_GENDER),
       DEM_REGION = factor(DEM_REGION),
       DEM_INCOME = factor(DEM_INCOME),
       TOB_LIFE = factor(TOB_LIFE),
       DAST_binary = factor(DAST_binary),
       PAINREL_NMUYR = factor(PAINREL_NMUYR),
       STIM_NMUYR = factor(STIM_NMUYR),
       SED NMUYR = factor(SED NMUYR),
       THC_NMUYR = factor(THC_NMUYR),
       OP_NMUYR = factor(OP_NMUYR),
       GAB_NMUYR = factor(GAB_NMUYR))
glimpse(us_19_mod)
## Rows: 3,000
## Columns: 86
## $ BHYD_NMUYR
                 ## $ BUP_NMUYR
                 ## $ COD_NMUYR
                 <dbl> NA, NA, NA, NA, NA, O, O, NA, NA, NA, NA, NA, NA, NA, O~
## $ DIHY_NMUYR
                 ## $ ELU NMUYR
                 ## $ FENT_NMUYR
                 ## $ GAB NMUYR
                 <dbl> NA, NA, NA, NA, NA, 1, 0, NA, NA, NA, NA, NA, NA, NA, NA
## $ HYD_NMUYR
## $ HYDM_NMUYR
```

```
## $ KTM NMUYR
         ## $ METH_NMUYR
         <dbl> NA, NA, NA, NA, NA, O, NA, NA, NA, NA, NA, NA, NA, NA, NA, ~
## $ MORPH NMUYR
         <dbl> NA, NA, NA, NA, NA, O, O, NA, NA, NA, NA, NA, NA, NA, NA, N~
         <dbl> NA, NA, NA, NA, NA, 1, NA, NA, NA, NA, NA, NA, NA, NA, ~
## $ OXY_NMUYR
## $ OXYM NMUYR
         ## $ PREG NMUYR
         ## $ SUF NMUYR
         ## $ TAP NMUYR
## $ TRAM NMUYR
         <dbl> NA, NA, NA, NA, NA, NA, O, NA, NA, NA, NA, NA, NA, NA, ~
## $ ALP_NMUYR
         ## $ BAC_NMUYR
         ## $ CHL_NMUYR
## $ CLOB_NMUYR
         ## $ CLON_NMUYR
         ## $ CLOR_NMUYR
## $ DIA_NMUYR
         ## $ EST_NMUYR
         ## $ ESZ NMUYR
         ## $ FLUR_NMUYR
## $ LORA NMUYR
         ## $ MID_NMUYR
         ## $ OXA NMUYR
         ## $ QUA_NMUYR
## $ TEM NMUYR
         ## $ TRI NMUYR
         ## $ ZAL_NMUYR
         ## $ ZOL_NMUYR
## $ AMPH_NMUYR
         ## $ ATOM_NMUYR
         ## $ MPHEN_NMUYR
         ## $ MOD_NMUYR
         ## $ CANN_NMUYR
         ## $ DRON_NMUYR
         ## $ NAB_NMUYR
         <dbl> NA, O, NA, O, O, O, O, O, NA, O, O, O, O, 1, NA, O, ~
## $ ACE NMUYR
## $ ASP_NMUYR
         <dbl> NA, O, 1, NA, O, O, O, NA, O, O, O, NA, NA, NA, NA, NA, ~
## $ DEX NMUYR
         <dbl> NA, NA, 1, 0, NA, NA, NA, NA, NA, NA, NA, O, O, NA, NA,~
## $ DIPH_NMUYR
         <dbl> NA, 0, 1, 0, 0, 0, 0, 0, NA, 0, NA, NA, 0, NA, 0, NA~
## $ IBU NMUYR
## $ LOP_NMUYR
         <dbl> NA, NA, NA, NA, NA, NA, NA, NA, NA, O, NA, NA, NA, NA, ~
## $ NAP NMUYR
         <dbl> NA, O, NA, O, O, O, NA, NA, O, O, O, NA, NA, NA, NA, O,~
## $ OTCOTH NMUYR
         <dbl> NA, NA, 1, NA, NA, NA, NA, NA, NA, NA, NA, O, O, NA, NA~
## $ PAINREL NMUYR
         <fct> 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
         ## $ SED_NMUYR
## $ STIM_NMUYR
         ## $ THC_NMUYR
         ## $ OP_NMUYR
         <fct> 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
## $ GABA_NMUYR
         ## $ DEM_GENDER
         <fct> 1, 1, 2, 2, 1, 1, 1, 1, 2, 1, 1, 2, 2, 2, 2, 2, 1, 2, 2~
         <dbl> 38, 57, 40, 54, 35, 60, 21, 29, 56, 66, 79, 58, 55, 42,~
## $ DEM_AGE
         <chr> "133", "806", "930", "194", "040", "551", "748", "330",~
## $ DEM_POSTAL
## $ DEM_REGION
         <fct> 1, 4, 4, 1, 1, 2, 3, 3, 3, 3, 3, 3, 3, 3, 1, 1, 4, 4~
## $ DEM INCOME
         <fct> 4, 2, 4, 4, 4, 4, 2, 1, 1, 2, 2, 3, 1, 5, 2, 3, 3, 5, 3~
## $ DEM HOME
         <dbl> 6, 1, 5, 1, 2, 3, 2, 3, 1, 1, 4, 2, 2, 4, 4, 1, 2, 2, 2~
```

```
## $ DEM_GENHEALTH <db1> 3, 3, 4, 4, 3, 2, 3, 4, 4, 3, 3, 5, 1, 4, 2, 4, 4, 2, 5~
               <dbl> 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0~
## $ DEM LIMIT
## $ DEM HISPANIC
               <dbl> 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
## $ DEM_RACE_AIAN <db1> 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0
## $ DEM RACE OTH
               ## $ DEM_MARITAL
               <dbl> 4, 3, 1, 5, 5, 1, 5, 1, 2, 5, 1, 1, 1, 1, 5, 2, 1, 1, 1~
## $ DEM_EDU
               <dbl> 6, 8, 3, 6, 3, 6, 2, 3, 5, 6, 2, 3, 2, 6, 2, 3, 8, 6, 7~
## $ DEM_STDNT
               <dbl> 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
## $ DEM_VET
               <dbl> 0, 0, 0, 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0~
               ## $ DEM_HEALTH
## $ DEM_EMPLOY
               <dbl> 0, 1, 0, 1, 1, 0, 0, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 1, 1~
## $ DEM_INSUR
               <dbl> 0, 1, 0, 1, 0, 1, 1, 0, 1, 1, 1, 1, 0, 1, 1, 1, 1, 1, 1
               <dbl> 0, 0, 0, 0, 0, 1, 1, 0, 0, 0, 1, 0, 1, 0, 0, 0, 0~
## $ DEM_HOSPSTAY
## $ DEM ZIP
               <dbl> 133, 806, 930, 194, 40, 551, 748, 330, 723, 731, 224, 3~
               <chr> "NY", "CO", "CA", "PA", "ME", "MN", "OK", "FL", "AR", "~
## $ DEM_STATE
               <dbl> 3, 5, 3, 4, 3, 5, 1, 2, 5, 6, 6, 5, 5, 3, 6, 6, 5, 4, 3~
## $ DEM AGE10
## $ TOB_LIFE
               <fct> 3, 3, 1, 3, 3, 3, 3, 1, 3, 3, 3, 3, 4, 3, 3, 3, 3~
## $ DAST CAT
               <dbl> 2, 2, 2, 1, 2, 4, 2, 1, 1, 1, 2, 2, 1, 1, 1, 1, 1, 1, 1~
## $ DAST_binary
               <fct> 0, 0, 0, 0, 0, 1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0~
binary_model <- glm(DAST_binary ~ THC_NMUYR + OP_NMUYR + GABA_NMUYR + PAINREL_NMUYR + STIM_NMUYR + SED_
tidy(binary_model) %>%
 kable(digits = 3)
```

term	estimate	std.error	statistic	p.value
(Intercept)	-0.354	0.429	-0.826	0.409
THC_NMUYR1	-0.221	0.465	-0.475	0.635
OP_NMUYR1	-0.133	0.746	-0.178	0.859
GABA_NMUYR	-0.444	0.516	-0.859	0.390
PAINREL_NMUYR1	1.097	0.749	1.466	0.143
STIM_NMUYR1	0.903	0.370	2.439	0.015
SED_NMUYR1	1.267	0.297	4.270	0.000
DEM_GENDER2	-0.593	0.193	-3.072	0.002
DEM_AGE	-0.034	0.006	-5.313	0.000
DEM_REGION2	-0.057	0.328	-0.173	0.863
DEM_REGION3	0.221	0.270	0.819	0.413
DEM_REGION4	0.207	0.295	0.702	0.483
DEM_INCOME2	-0.252	0.255	-0.988	0.323
DEM_INCOME3	-0.184	0.277	-0.663	0.507
DEM_INCOME4	-1.103	0.356	-3.096	0.002
DEM_INCOME5	-0.636	0.319	-1.997	0.046
TOB_LIFE2	-0.038	0.257	-0.146	0.884
TOB_LIFE3	-1.436	0.221	-6.507	0.000
TOB_LIFE4	-1.585	0.617	-2.570	0.010

```
int_only_binary_model <- glm(DAST_binary ~ 1, data = us_19_mod, family = binomial)
final_binary_model <- step(binary_model, scope = formula(int_only_binary_model), direction = "backward"</pre>
```

```
## Start: AIC=936.35
## DAST_binary ~ THC_NMUYR + OP_NMUYR + GABA_NMUYR + PAINREL_NMUYR +
      STIM NMUYR + SED NMUYR + DEM GENDER + DEM AGE + DEM REGION +
      DEM_INCOME + TOB_LIFE
##
##
##
                   Df Deviance
                                  AIC
## - DEM REGION
                    3
                       899.91 931.91
## - OP NMUYR
                        898.39 934.39
                    1
## - THC_NMUYR
                    1
                        898.58 934.58
## - GABA_NMUYR
                    1
                        899.11 935.11
## - PAINREL_NMUYR 1
                        900.33 936.33
                        898.35 936.35
## <none>
## - STIM_NMUYR
                        904.08 940.08
                    1
## - DEM_INCOME
                      911.32 941.32
                    4
## - DEM_GENDER
                       907.99 943.99
                    1
## - SED_NMUYR
                    1
                       915.53 951.53
## - DEM_AGE
                        927.99 963.99
                    1
## - TOB LIFE
                    3
                      952.29 984.29
##
## Step: AIC=931.91
## DAST_binary ~ THC_NMUYR + OP_NMUYR + GABA_NMUYR + PAINREL_NMUYR +
      STIM_NMUYR + SED_NMUYR + DEM_GENDER + DEM_AGE + DEM_INCOME +
##
      TOB_LIFE
##
##
                   Df Deviance
                                  AIC
## - OP NMUYR
                   1
                        899.96 929.96
## - THC_NMUYR
                        900.16 930.16
                    1
                        900.86 930.86
## - GABA_NMUYR
                   1
## <none>
                        899.91 931.91
## - PAINREL_NMUYR 1
                        902.07 932.07
## - STIM_NMUYR
                    1
                        905.76 935.76
## - DEM_INCOME
                    4
                        913.12 937.12
## - DEM_GENDER
                    1
                        909.46 939.46
                        917.01 947.01
## - SED_NMUYR
                    1
## - DEM AGE
                    1
                        929.79 959.79
## - TOB_LIFE
                    3
                        954.61 980.61
##
## Step: AIC=929.96
## DAST binary ~ THC NMUYR + GABA NMUYR + PAINREL NMUYR + STIM NMUYR +
      SED_NMUYR + DEM_GENDER + DEM_AGE + DEM_INCOME + TOB_LIFE
##
##
##
                   Df Deviance
                                  AIC
                        900.21 928.21
## - THC NMUYR
                   1
## - GABA_NMUYR
                        900.86 928.86
                    1
                        899.96 929.96
## <none>
## - STIM_NMUYR
                        905.82 933.82
                    1
## - DEM_INCOME
                    4
                        913.21 935.21
## - DEM_GENDER
                        909.52 937.52
                    1
## - PAINREL_NMUYR 1
                        912.14 940.14
## - SED_NMUYR
                    1
                        917.09 945.09
## - DEM_AGE
                    1
                        929.94 957.94
                    3
## - TOB_LIFE
                        954.78 978.78
##
## Step: AIC=928.21
```

```
## DAST_binary ~ GABA_NMUYR + PAINREL_NMUYR + STIM_NMUYR + SED_NMUYR +
       DEM_GENDER + DEM_AGE + DEM_INCOME + TOB_LIFE
##
##
##
                  Df Deviance
                                 AIC
## - GABA_NMUYR
                   1 901.10 927.10
## <none>
                       900.21 928.21
## - STIM NMUYR
                       905.84 931.84
                   1
## - DEM_INCOME
                       913.47 933.47
                   4
## - DEM_GENDER
                   1
                       909.58 935.58
## - PAINREL_NMUYR 1
                       912.15 938.15
## - SED_NMUYR
                   1
                       917.10 943.10
## - DEM_AGE
                       929.94 955.94
                   1
## - TOB_LIFE
                       954.83 976.83
##
## Step: AIC=927.1
## DAST_binary ~ PAINREL_NMUYR + STIM_NMUYR + SED_NMUYR + DEM_GENDER +
##
       DEM_AGE + DEM_INCOME + TOB_LIFE
##
##
                  Df Deviance
                                 AIC
## <none>
                       901.10 927.10
                       906.89 930.89
## - STIM_NMUYR
                   1
## - DEM_INCOME
                       914.55 932.55
## - DEM_GENDER
                       910.56 934.56
                   1
## - PAINREL NMUYR 1
                       912.17 936.17
## - SED NMUYR
                       917.14 941.14
                   1
## - DEM AGE
                   1
                       930.82 954.82
## - TOB_LIFE
                    3 955.30 975.30
tidy(final_binary_model, conf.int = TRUE) %>%
kable(digits = 3)
```

term	estimate	std.error	statistic	p.value	conf.low	conf.high
(Intercept)	-0.245	0.374	-0.657	0.511	-0.986	0.481
PAINREL_NMUYR1	0.920	0.266	3.458	0.001	0.387	1.431
STIM_NMUYR1	0.881	0.359	2.451	0.014	0.166	1.578
SED_NMUYR1	1.188	0.289	4.118	0.000	0.616	1.749
DEM_GENDER2	-0.585	0.192	-3.043	0.002	-0.966	-0.211
$\mathrm{DEM}\_\mathrm{AGE}$	-0.034	0.006	-5.319	0.000	-0.047	-0.022
DEM_INCOME2	-0.233	0.254	-0.915	0.360	-0.731	0.269
DEM_INCOME3	-0.173	0.276	-0.627	0.530	-0.719	0.367
DEM_INCOME4	-1.112	0.356	-3.124	0.002	-1.841	-0.437
DEM_INCOME5	-0.638	0.319	-2.001	0.045	-1.281	-0.024
TOB_LIFE2	-0.023	0.257	-0.088	0.930	-0.534	0.475
TOB_LIFE3	-1.429	0.220	-6.496	0.000	-1.862	-0.997
TOB_LIFE4	-1.592	0.616	-2.585	0.010	-3.035	-0.537