HW 3

Gaea Daniel

10/4/2017

## Summary Statistics for the Health Evaluation and Linkage to Primary Care (HELP) dataset

Because the data are not normally distributed and are not expected to be, non-parametric statistics are appropriate for data interpretation. Furthermore, Levene's test shows significance indicating homoscedasticity and the appropriateness of non-parametric statistics.

### Table 1 (Baseline Data Overall)

######Creating Table 1 for baseline data overall#####  
#Load data  
help.spss <- read\_csv("helpmkh.csv")

## Parsed with column specification:  
## cols(  
## .default = col\_integer(),  
## racegrp = col\_character(),  
## pcs = col\_double(),  
## mcs = col\_double(),  
## substance = col\_character(),  
## pcs1 = col\_double(),  
## mcs1 = col\_double(),  
## pcs2 = col\_double(),  
## mcs2 = col\_double(),  
## pcs3 = col\_double(),  
## mcs3 = col\_double(),  
## pcs4 = col\_double(),  
## mcs4 = col\_double()  
## )

## See spec(...) for full column specifications.

data(help.spss)

## Warning in data(help.spss): data set 'help.spss' not found

#Check variables  
head(help.spss)

## # A tibble: 6 × 88  
## `﻿id` treat age female pss\_fr racegrp homeless a15a a15b d1 e2b  
## <int> <int> <int> <int> <int> <chr> <int> <int> <int> <int> <int>  
## 1 1 1 37 0 0 black 0 0 0 3 NA  
## 2 2 1 37 0 1 white 1 2 3 22 NA  
## 3 3 0 26 0 13 black 0 0 0 0 NA  
## 4 4 0 39 1 11 white 0 0 0 2 1  
## 5 5 0 32 0 10 black 1 15 0 12 1  
## 6 6 1 47 1 5 black 0 0 0 1 NA  
## # ... with 77 more variables: g1b <int>, i1 <int>, i2 <int>, pcs <dbl>,  
## # mcs <dbl>, f1a <int>, f1b <int>, f1c <int>, f1d <int>, f1e <int>,  
## # f1f <int>, f1g <int>, f1h <int>, f1i <int>, f1j <int>, f1k <int>,  
## # f1l <int>, f1m <int>, f1n <int>, f1o <int>, f1p <int>, f1q <int>,  
## # f1r <int>, f1s <int>, f1t <int>, cesd <int>, indtot <int>,  
## # drugrisk <int>, sexrisk <int>, satreat <int>, substance <chr>,  
## # drinkstatus <int>, daysdrink <int>, anysubstatus <int>,  
## # daysanysub <int>, linkstatus <int>, dayslink <int>, e2b1 <int>,  
## # g1b1 <int>, i11 <int>, pcs1 <dbl>, mcs1 <dbl>, cesd1 <int>,  
## # indtot1 <int>, drugrisk1 <int>, sexrisk1 <int>, pcrec1 <int>,  
## # e2b2 <int>, g1b2 <int>, i12 <int>, pcs2 <dbl>, mcs2 <dbl>,  
## # cesd2 <int>, indtot2 <int>, drugrisk2 <int>, sexrisk2 <int>,  
## # pcrec2 <int>, e2b3 <int>, g1b3 <int>, i13 <int>, pcs3 <dbl>,  
## # mcs3 <dbl>, cesd3 <int>, indtot3 <int>, drugrisk3 <int>,  
## # sexrisk3 <int>, pcrec3 <int>, e2b4 <int>, g1b4 <int>, i14 <int>,  
## # pcs4 <dbl>, mcs4 <dbl>, cesd4 <int>, indtot4 <int>, drugrisk4 <int>,  
## # sexrisk4 <int>, pcrec4 <int>

#Make categorical variables factors  
varsToFactor <- c("treat","female","homeless","f1a","f1b","f1c","f1d","f1e","f1f","f1g","f1h","f1i","f1j","f1k","f1l","f1m","f1n","f1o","f1p","f1q","f1r","f1s","f1t","satreat","substance","drinkstatus","anysubstatus","linkstatus","g1b1")  
  
#Create a variable list  
dput(names(help.spss))

## c("﻿id", "treat", "age", "female", "pss\_fr", "racegrp", "homeless",   
## "a15a", "a15b", "d1", "e2b", "g1b", "i1", "i2", "pcs", "mcs",   
## "f1a", "f1b", "f1c", "f1d", "f1e", "f1f", "f1g", "f1h", "f1i",   
## "f1j", "f1k", "f1l", "f1m", "f1n", "f1o", "f1p", "f1q", "f1r",   
## "f1s", "f1t", "cesd", "indtot", "drugrisk", "sexrisk", "satreat",   
## "substance", "drinkstatus", "daysdrink", "anysubstatus", "daysanysub",   
## "linkstatus", "dayslink", "e2b1", "g1b1", "i11", "pcs1", "mcs1",   
## "cesd1", "indtot1", "drugrisk1", "sexrisk1", "pcrec1", "e2b2",   
## "g1b2", "i12", "pcs2", "mcs2", "cesd2", "indtot2", "drugrisk2",   
## "sexrisk2", "pcrec2", "e2b3", "g1b3", "i13", "pcs3", "mcs3",   
## "cesd3", "indtot3", "drugrisk3", "sexrisk3", "pcrec3", "e2b4",   
## "g1b4", "i14", "pcs4", "mcs4", "cesd4", "indtot4", "drugrisk4",   
## "sexrisk4", "pcrec4")

vars <- c("treat","age","female","pss\_fr","racegrp","homeless","a15a","a15b","d1","e2b","g1b","i1","i2","pcs","mcs","f1a","f1b","f1c","f1d","f1e","f1f","f1g","f1h","f1i","f1j","f1k","f1l","f1m","f1n","f1o","f1p","f1q","f1r","f1s","f1t","cesd","indtot","drugrisk","sexrisk","satreat","substance","drinkstatus","daysdrink","anysubstatus","daysanysub","linkstatus","e2b1","g1b1","i11")  
  
#Create Table 1  
tableOne <- CreateTableOne(vars = vars, data = help.spss)  
  
tableOne

##   
## Overall   
## n 453   
## treat (mean (sd)) 0.50 (0.50)   
## age (mean (sd)) 35.65 (7.71)   
## female (mean (sd)) 0.24 (0.43)   
## pss\_fr (mean (sd)) 6.71 (4.00)   
## racegrp (%)   
## black 211 (46.6)   
## hispanic 50 (11.0)   
## other 26 ( 5.7)   
## white 166 (36.6)   
## homeless (mean (sd)) 0.46 (0.50)   
## a15a (mean (sd)) 14.21 (38.93)  
## a15b (mean (sd)) 9.63 (27.84)  
## d1 (mean (sd)) 3.06 (6.19)   
## e2b (mean (sd)) 2.50 (2.52)   
## g1b (mean (sd)) 0.28 (0.45)   
## i1 (mean (sd)) 17.91 (20.02)  
## i2 (mean (sd)) 22.65 (27.30)  
## pcs (mean (sd)) 48.05 (10.78)  
## mcs (mean (sd)) 31.68 (12.84)  
## f1a (mean (sd)) 1.63 (1.11)   
## f1b (mean (sd)) 1.39 (1.12)   
## f1c (mean (sd)) 1.92 (1.09)   
## f1d (mean (sd)) 1.57 (1.19)   
## f1e (mean (sd)) 1.70 (1.09)   
## f1f (mean (sd)) 2.02 (1.06)   
## f1g (mean (sd)) 1.73 (1.10)   
## f1h (mean (sd)) 1.48 (1.16)   
## f1i (mean (sd)) 1.88 (1.13)   
## f1j (mean (sd)) 1.53 (1.16)   
## f1k (mean (sd)) 2.04 (1.11)   
## f1l (mean (sd)) 1.08 (1.01)   
## f1m (mean (sd)) 1.51 (1.08)   
## f1n (mean (sd)) 2.01 (1.07)   
## f1o (mean (sd)) 1.22 (1.12)   
## f1p (mean (sd)) 1.34 (1.14)   
## f1q (mean (sd)) 1.07 (1.16)   
## f1r (mean (sd)) 1.94 (1.01)   
## f1s (mean (sd)) 1.23 (1.14)   
## f1t (mean (sd)) 1.53 (1.09)   
## cesd (mean (sd)) 32.85 (12.51)  
## indtot (mean (sd)) 35.73 (7.15)   
## drugrisk (mean (sd)) 1.89 (4.34)   
## sexrisk (mean (sd)) 4.64 (2.80)   
## satreat (mean (sd)) 0.28 (0.45)   
## substance (%)   
## alcohol 177 (39.1)   
## cocaine 152 (33.6)   
## heroin 124 (27.4)   
## drinkstatus (mean (sd)) 0.67 (0.47)   
## daysdrink (mean (sd)) 93.63 (85.40)  
## anysubstatus (mean (sd)) 0.77 (0.42)   
## daysanysub (mean (sd)) 75.31 (79.24)  
## linkstatus (mean (sd)) 0.38 (0.49)   
## e2b1 (mean (sd)) 2.13 (1.85)   
## g1b1 (mean (sd)) 0.11 (0.31)   
## i11 (mean (sd)) 17.19 (25.53)

summary (help.spss)

## ﻿id treat age female   
## Min. : 1.0 Min. :0.0000 Min. :19.00 Min. :0.0000   
## 1st Qu.:119.0 1st Qu.:0.0000 1st Qu.:30.00 1st Qu.:0.0000   
## Median :233.0 Median :0.0000 Median :35.00 Median :0.0000   
## Mean :233.4 Mean :0.4967 Mean :35.65 Mean :0.2362   
## 3rd Qu.:348.0 3rd Qu.:1.0000 3rd Qu.:40.00 3rd Qu.:0.0000   
## Max. :470.0 Max. :1.0000 Max. :60.00 Max. :1.0000   
##   
## pss\_fr racegrp homeless a15a   
## Min. : 0.000 Length:453 Min. :0.0000 Min. : 0.00   
## 1st Qu.: 3.000 Class :character 1st Qu.:0.0000 1st Qu.: 0.00   
## Median : 7.000 Mode :character Median :0.0000 Median : 0.00   
## Mean : 6.706 Mean :0.4614 Mean : 14.21   
## 3rd Qu.:10.000 3rd Qu.:1.0000 3rd Qu.: 3.00   
## Max. :14.000 Max. :1.0000 Max. :180.00   
##   
## a15b d1 e2b g1b   
## Min. : 0.000 Min. : 0.00 Min. : 1.000 Min. :0.0000   
## 1st Qu.: 0.000 1st Qu.: 1.00 1st Qu.: 1.000 1st Qu.:0.0000   
## Median : 0.000 Median : 2.00 Median : 2.000 Median :0.0000   
## Mean : 9.631 Mean : 3.06 Mean : 2.505 Mean :0.2804   
## 3rd Qu.: 3.000 3rd Qu.: 3.00 3rd Qu.: 3.000 3rd Qu.:1.0000   
## Max. :180.000 Max. :100.00 Max. :21.000 Max. :1.0000   
## NA's :239   
## i1 i2 pcs mcs   
## Min. : 0.00 Min. : 0.00 Min. :14.07 Min. : 6.763   
## 1st Qu.: 3.00 1st Qu.: 3.00 1st Qu.:40.38 1st Qu.:21.676   
## Median : 13.00 Median : 15.00 Median :48.88 Median :28.602   
## Mean : 17.91 Mean : 22.65 Mean :48.05 Mean :31.677   
## 3rd Qu.: 26.00 3rd Qu.: 32.00 3rd Qu.:56.95 3rd Qu.:40.941   
## Max. :142.00 Max. :184.00 Max. :74.81 Max. :62.175   
##   
## f1a f1b f1c f1d   
## Min. :0.000 Min. :0.000 Min. :0.000 Min. :0.000   
## 1st Qu.:1.000 1st Qu.:0.000 1st Qu.:1.000 1st Qu.:0.000   
## Median :2.000 Median :1.000 Median :2.000 Median :1.000   
## Mean :1.634 Mean :1.391 Mean :1.923 Mean :1.565   
## 3rd Qu.:3.000 3rd Qu.:2.000 3rd Qu.:3.000 3rd Qu.:3.000   
## Max. :3.000 Max. :3.000 Max. :3.000 Max. :3.000   
##   
## f1e f1f f1g f1h   
## Min. :0.000 Min. :0.000 Min. :0.00 Min. :0.000   
## 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:1.00 1st Qu.:0.000   
## Median :2.000 Median :2.000 Median :2.00 Median :1.000   
## Mean :1.695 Mean :2.018 Mean :1.73 Mean :1.481   
## 3rd Qu.:3.000 3rd Qu.:3.000 3rd Qu.:3.00 3rd Qu.:3.000   
## Max. :3.000 Max. :3.000 Max. :3.00 Max. :3.000   
## NA's :1   
## f1i f1j f1k f1l   
## Min. :0.000 Min. :0.000 Min. :0.000 Min. :0.000   
## 1st Qu.:1.000 1st Qu.:0.000 1st Qu.:1.000 1st Qu.:0.000   
## Median :2.000 Median :2.000 Median :2.000 Median :1.000   
## Mean :1.883 Mean :1.525 Mean :2.035 Mean :1.084   
## 3rd Qu.:3.000 3rd Qu.:3.000 3rd Qu.:3.000 3rd Qu.:2.000   
## Max. :3.000 Max. :3.000 Max. :3.000 Max. :3.000   
## NA's :2   
## f1m f1n f1o f1p   
## Min. :0.000 Min. :0.000 Min. :0.000 Min. :0.000   
## 1st Qu.:1.000 1st Qu.:1.000 1st Qu.:0.000 1st Qu.:0.000   
## Median :1.000 Median :2.000 Median :1.000 Median :1.000   
## Mean :1.507 Mean :2.009 Mean :1.219 Mean :1.338   
## 3rd Qu.:2.000 3rd Qu.:3.000 3rd Qu.:2.000 3rd Qu.:2.000   
## Max. :3.000 Max. :3.000 Max. :3.000 Max. :3.000   
## NA's :1 NA's :1 NA's :1   
## f1q f1r f1s f1t   
## Min. :0.000 Min. :0.000 Min. :0.000 Min. :0.00   
## 1st Qu.:0.000 1st Qu.:1.000 1st Qu.:0.000 1st Qu.:1.00   
## Median :1.000 Median :2.000 Median :1.000 Median :1.00   
## Mean :1.073 Mean :1.945 Mean :1.225 Mean :1.53   
## 3rd Qu.:2.000 3rd Qu.:3.000 3rd Qu.:2.000 3rd Qu.:3.00   
## Max. :3.000 Max. :3.000 Max. :3.000 Max. :3.00   
## NA's :1   
## cesd indtot drugrisk sexrisk   
## Min. : 1.00 Min. : 4.00 Min. : 0.000 Min. : 0.000   
## 1st Qu.:25.00 1st Qu.:32.00 1st Qu.: 0.000 1st Qu.: 3.000   
## Median :34.00 Median :38.00 Median : 0.000 Median : 4.000   
## Mean :32.85 Mean :35.73 Mean : 1.887 Mean : 4.642   
## 3rd Qu.:41.00 3rd Qu.:41.00 3rd Qu.: 1.000 3rd Qu.: 6.000   
## Max. :60.00 Max. :45.00 Max. :21.000 Max. :14.000   
## NA's :1   
## satreat substance drinkstatus daysdrink   
## Min. :0.0000 Length:453 Min. :0.0000 Min. : 0.00   
## 1st Qu.:0.0000 Class :character 1st Qu.:0.0000 1st Qu.: 9.00   
## Median :0.0000 Mode :character Median :1.0000 Median : 63.00   
## Mean :0.2848 Mean :0.6707 Mean : 93.63   
## 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:181.00   
## Max. :1.0000 Max. :1.0000 Max. :270.00   
## NA's :207 NA's :209   
## anysubstatus daysanysub linkstatus dayslink   
## Min. :0.0000 Min. : 0.00 Min. :0.0000 Min. : 2.0   
## 1st Qu.:1.0000 1st Qu.: 5.00 1st Qu.:0.0000 1st Qu.: 74.0   
## Median :1.0000 Median : 33.00 Median :0.0000 Median :361.0   
## Mean :0.7724 Mean : 75.31 Mean :0.3782 Mean :255.6   
## 3rd Qu.:1.0000 3rd Qu.:164.25 3rd Qu.:1.0000 3rd Qu.:365.0   
## Max. :1.0000 Max. :268.00 Max. :1.0000 Max. :456.0   
## NA's :207 NA's :209 NA's :22 NA's :22   
## e2b1 g1b1 i11 pcs1   
## Min. : 1.000 Min. :0.0000 Min. : 1.00 Min. :19.70   
## 1st Qu.: 1.000 1st Qu.:0.0000 1st Qu.: 4.00 1st Qu.:42.84   
## Median : 1.500 Median :0.0000 Median : 10.00 Median :53.32   
## Mean : 2.129 Mean :0.1098 Mean : 17.19 Mean :50.16   
## 3rd Qu.: 2.000 3rd Qu.:0.0000 3rd Qu.: 19.25 3rd Qu.:58.17   
## Max. :11.000 Max. :1.0000 Max. :216.00 Max. :71.44   
## NA's :337 NA's :207 NA's :349 NA's :207   
## mcs1 cesd1 indtot1 drugrisk1   
## Min. : 6.677 Min. : 0.00 Min. : 0.00 Min. : 0.000   
## 1st Qu.:30.214 1st Qu.:11.00 1st Qu.: 6.25 1st Qu.: 0.000   
## Median :42.441 Median :22.00 Median : 46.50 Median : 0.000   
## Mean :40.983 Mean :22.72 Mean : 47.46 Mean : 1.115   
## 3rd Qu.:52.734 3rd Qu.:34.75 3rd Qu.: 79.00 3rd Qu.: 0.000   
## Max. :69.939 Max. :56.00 Max. :121.00 Max. :19.000   
## NA's :207 NA's :207 NA's :211 NA's :209   
## sexrisk1 pcrec1 e2b2 g1b2   
## Min. : 0.000 Min. :0.0000 Min. :1.000 Min. :0.0000   
## 1st Qu.: 1.000 1st Qu.:0.0000 1st Qu.:1.000 1st Qu.:0.0000   
## Median : 4.000 Median :1.0000 Median :2.000 Median :0.0000   
## Mean : 3.555 Mean :0.8041 Mean :1.871 Mean :0.1053   
## 3rd Qu.: 5.000 3rd Qu.:2.0000 3rd Qu.:2.000 3rd Qu.:0.0000   
## Max. :15.000 Max. :2.0000 Max. :7.000 Max. :1.0000   
## NA's :208 NA's :208 NA's :352 NA's :244   
## i12 pcs2 mcs2 cesd2   
## Min. : 1.00 Min. :14.44 Min. :11.65 Min. : 0.00   
## 1st Qu.: 6.00 1st Qu.:42.99 1st Qu.:29.71 1st Qu.:11.00   
## Median : 11.00 Median :53.37 Median :41.59 Median :24.00   
## Mean : 15.92 Mean :50.02 Mean :40.36 Mean :23.58   
## 3rd Qu.: 21.00 3rd Qu.:58.14 3rd Qu.:53.51 3rd Qu.:36.00   
## Max. :115.00 Max. :68.93 Max. :64.79 Max. :59.00   
## NA's :348 NA's :242 NA's :242 NA's :244   
## indtot2 drugrisk2 sexrisk2 pcrec2   
## Min. : 0.00 Min. : 0.000 Min. : 0.000 Min. :0.000   
## 1st Qu.: 7.75 1st Qu.: 0.000 1st Qu.: 1.000 1st Qu.:0.000   
## Median : 48.00 Median : 0.000 Median : 4.000 Median :1.000   
## Mean : 47.49 Mean : 1.442 Mean : 3.553 Mean :1.033   
## 3rd Qu.: 82.25 3rd Qu.: 0.000 3rd Qu.: 5.000 3rd Qu.:2.000   
## Max. :114.00 Max. :21.000 Max. :13.000 Max. :2.000   
## NA's :249 NA's :245 NA's :245 NA's :243   
## e2b3 g1b3 i13 pcs3   
## Min. : 1.000 Min. :0.00000 Min. : 1.00 Min. :18.52   
## 1st Qu.: 1.000 1st Qu.:0.00000 1st Qu.: 4.00 1st Qu.:41.44   
## Median : 2.000 Median :0.00000 Median : 12.00 Median :51.80   
## Mean : 2.216 Mean :0.08907 Mean : 15.92 Mean :49.10   
## 3rd Qu.: 3.000 3rd Qu.:0.00000 3rd Qu.: 19.00 3rd Qu.:57.48   
## Max. :13.000 Max. :1.00000 Max. :256.00 Max. :71.29   
## NA's :356 NA's :206 NA's :331 NA's :205   
## mcs3 cesd3 indtot3 drugrisk3   
## Min. :13.39 Min. : 0.00 Min. : 0.00 Min. : 0.000   
## 1st Qu.:30.66 1st Qu.:11.00 1st Qu.: 2.00 1st Qu.: 0.000   
## Median :43.57 Median :21.00 Median : 47.00 Median : 0.000   
## Mean :41.89 Mean :22.07 Mean : 46.35 Mean : 1.298   
## 3rd Qu.:54.55 3rd Qu.:32.00 3rd Qu.: 79.00 3rd Qu.: 0.000   
## Max. :66.10 Max. :60.00 Max. :121.00 Max. :22.000   
## NA's :205 NA's :205 NA's :208 NA's :205   
## sexrisk3 pcrec3 e2b4 g1b4   
## Min. : 0.000 Min. :0.000 Min. : 1.000 Min. :0.00000   
## 1st Qu.: 2.000 1st Qu.:0.000 1st Qu.: 1.000 1st Qu.:0.00000   
## Median : 4.000 Median :2.000 Median : 1.000 Median :0.00000   
## Mean : 3.731 Mean :1.129 Mean : 2.643 Mean :0.07895   
## 3rd Qu.: 5.000 3rd Qu.:2.000 3rd Qu.: 3.000 3rd Qu.:0.00000   
## Max. :13.000 Max. :2.000 Max. :34.000 Max. :1.00000   
## NA's :208 NA's :204 NA's :355 NA's :187   
## i14 pcs4 mcs4 cesd4   
## Min. : 1.00 Min. :15.41 Min. : 6.241 Min. : 0.00   
## 1st Qu.: 5.00 1st Qu.:43.25 1st Qu.:31.614 1st Qu.: 7.00   
## Median : 8.00 Median :52.87 Median :46.099 Median :19.00   
## Mean : 14.87 Mean :49.76 Mean :43.367 Mean :20.14   
## 3rd Qu.: 19.00 3rd Qu.:57.38 3rd Qu.:55.566 3rd Qu.:29.00   
## Max. :113.00 Max. :69.72 Max. :66.910 Max. :52.00   
## NA's :330 NA's :187 NA's :187 NA's :187   
## indtot4 drugrisk4 sexrisk4 pcrec4   
## Min. : 0.00 Min. : 0.000 Min. : 0.000 Min. :0.000   
## 1st Qu.: 2.00 1st Qu.: 0.000 1st Qu.: 1.000 1st Qu.:0.000   
## Median : 41.00 Median : 0.000 Median : 4.000 Median :2.000   
## Mean : 41.99 Mean : 1.135 Mean : 3.389 Mean :1.275   
## 3rd Qu.: 72.00 3rd Qu.: 0.000 3rd Qu.: 5.000 3rd Qu.:2.000   
## Max. :126.00 Max. :20.000 Max. :13.000 Max. :2.000   
## NA's :192 NA's :187 NA's :191 NA's :184

### Table 1 (Stratification by Treatment Group)

######Creating Table 1 for stratification by treatment group#####  
#Load data  
help.spss <- read\_csv("helpmkh.csv")

## Parsed with column specification:  
## cols(  
## .default = col\_integer(),  
## racegrp = col\_character(),  
## pcs = col\_double(),  
## mcs = col\_double(),  
## substance = col\_character(),  
## pcs1 = col\_double(),  
## mcs1 = col\_double(),  
## pcs2 = col\_double(),  
## mcs2 = col\_double(),  
## pcs3 = col\_double(),  
## mcs3 = col\_double(),  
## pcs4 = col\_double(),  
## mcs4 = col\_double()  
## )

## See spec(...) for full column specifications.

data(help.spss)

## Warning in data(help.spss): data set 'help.spss' not found

#Check variables  
head(help.spss)

## # A tibble: 6 × 88  
## `﻿id` treat age female pss\_fr racegrp homeless a15a a15b d1 e2b  
## <int> <int> <int> <int> <int> <chr> <int> <int> <int> <int> <int>  
## 1 1 1 37 0 0 black 0 0 0 3 NA  
## 2 2 1 37 0 1 white 1 2 3 22 NA  
## 3 3 0 26 0 13 black 0 0 0 0 NA  
## 4 4 0 39 1 11 white 0 0 0 2 1  
## 5 5 0 32 0 10 black 1 15 0 12 1  
## 6 6 1 47 1 5 black 0 0 0 1 NA  
## # ... with 77 more variables: g1b <int>, i1 <int>, i2 <int>, pcs <dbl>,  
## # mcs <dbl>, f1a <int>, f1b <int>, f1c <int>, f1d <int>, f1e <int>,  
## # f1f <int>, f1g <int>, f1h <int>, f1i <int>, f1j <int>, f1k <int>,  
## # f1l <int>, f1m <int>, f1n <int>, f1o <int>, f1p <int>, f1q <int>,  
## # f1r <int>, f1s <int>, f1t <int>, cesd <int>, indtot <int>,  
## # drugrisk <int>, sexrisk <int>, satreat <int>, substance <chr>,  
## # drinkstatus <int>, daysdrink <int>, anysubstatus <int>,  
## # daysanysub <int>, linkstatus <int>, dayslink <int>, e2b1 <int>,  
## # g1b1 <int>, i11 <int>, pcs1 <dbl>, mcs1 <dbl>, cesd1 <int>,  
## # indtot1 <int>, drugrisk1 <int>, sexrisk1 <int>, pcrec1 <int>,  
## # e2b2 <int>, g1b2 <int>, i12 <int>, pcs2 <dbl>, mcs2 <dbl>,  
## # cesd2 <int>, indtot2 <int>, drugrisk2 <int>, sexrisk2 <int>,  
## # pcrec2 <int>, e2b3 <int>, g1b3 <int>, i13 <int>, pcs3 <dbl>,  
## # mcs3 <dbl>, cesd3 <int>, indtot3 <int>, drugrisk3 <int>,  
## # sexrisk3 <int>, pcrec3 <int>, e2b4 <int>, g1b4 <int>, i14 <int>,  
## # pcs4 <dbl>, mcs4 <dbl>, cesd4 <int>, indtot4 <int>, drugrisk4 <int>,  
## # sexrisk4 <int>, pcrec4 <int>

#Make categorical variables factors  
varsToFactor <- c("treat","female","homeless","f1a","f1b","f1c","f1d","f1e","f1f","f1g","f1h","f1i","f1j","f1k","f1l","f1m","f1n","f1o","f1p","f1q","f1r","f1s","f1t","satreat","substance","drinkstatus","anysubstatus","linkstatus","g1b1")  
  
#Create a variable list  
dput(names(help.spss))

## c("﻿id", "treat", "age", "female", "pss\_fr", "racegrp", "homeless",   
## "a15a", "a15b", "d1", "e2b", "g1b", "i1", "i2", "pcs", "mcs",   
## "f1a", "f1b", "f1c", "f1d", "f1e", "f1f", "f1g", "f1h", "f1i",   
## "f1j", "f1k", "f1l", "f1m", "f1n", "f1o", "f1p", "f1q", "f1r",   
## "f1s", "f1t", "cesd", "indtot", "drugrisk", "sexrisk", "satreat",   
## "substance", "drinkstatus", "daysdrink", "anysubstatus", "daysanysub",   
## "linkstatus", "dayslink", "e2b1", "g1b1", "i11", "pcs1", "mcs1",   
## "cesd1", "indtot1", "drugrisk1", "sexrisk1", "pcrec1", "e2b2",   
## "g1b2", "i12", "pcs2", "mcs2", "cesd2", "indtot2", "drugrisk2",   
## "sexrisk2", "pcrec2", "e2b3", "g1b3", "i13", "pcs3", "mcs3",   
## "cesd3", "indtot3", "drugrisk3", "sexrisk3", "pcrec3", "e2b4",   
## "g1b4", "i14", "pcs4", "mcs4", "cesd4", "indtot4", "drugrisk4",   
## "sexrisk4", "pcrec4")

vars <- c("treat","age","female","pss\_fr","racegrp","homeless","a15a","a15b","d1","e2b","g1b","i1","i2","pcs","mcs","f1a","f1b","f1c","f1d","f1e","f1f","f1g","f1h","f1i","f1j","f1k","f1l","f1m","f1n","f1o","f1p","f1q","f1r","f1s","f1t","cesd","indtot","drugrisk","sexrisk","satreat","substance","drinkstatus","daysdrink","anysubstatus","daysanysub","linkstatus","e2b1","g1b1","i11")  
  
#Create Table 1 stratified by treat  
tableOne <- CreateTableOne(vars = vars, strata = c("treat"), data = help.spss)  
  
tableOne

## Stratified by treat  
## 0 1 p test  
## n 228 225   
## treat (mean (sd)) 0.00 (0.00) 1.00 (0.00) <0.001   
## age (mean (sd)) 36.30 (8.14) 35.00 (7.21) 0.071   
## female (mean (sd)) 0.24 (0.43) 0.23 (0.42) 0.800   
## pss\_fr (mean (sd)) 6.82 (3.90) 6.60 (4.09) 0.558   
## racegrp (%) 0.785   
## black 106 (46.5) 105 (46.7)   
## hispanic 24 (10.5) 26 (11.6)   
## other 11 ( 4.8) 15 ( 6.7)   
## white 87 (38.2) 79 (35.1)   
## homeless (mean (sd)) 0.48 (0.50) 0.44 (0.50) 0.474   
## a15a (mean (sd)) 14.46 (38.58) 13.95 (39.37) 0.888   
## a15b (mean (sd)) 11.45 (32.52) 7.79 (22.03) 0.161   
## d1 (mean (sd)) 3.64 (8.12) 2.48 (3.13) 0.046   
## e2b (mean (sd)) 2.46 (2.79) 2.55 (2.20) 0.806   
## g1b (mean (sd)) 0.26 (0.44) 0.30 (0.46) 0.413   
## i1 (mean (sd)) 19.00 (20.59) 16.80 (19.41) 0.241   
## i2 (mean (sd)) 23.19 (25.31) 22.10 (29.22) 0.672   
## pcs (mean (sd)) 47.83 (10.66) 48.27 (10.93) 0.669   
## mcs (mean (sd)) 31.65 (13.12) 31.70 (12.58) 0.970   
## f1a (mean (sd)) 1.66 (1.11) 1.60 (1.11) 0.580   
## f1b (mean (sd)) 1.35 (1.14) 1.43 (1.11) 0.448   
## f1c (mean (sd)) 1.98 (1.09) 1.87 (1.10) 0.279   
## f1d (mean (sd)) 1.56 (1.19) 1.57 (1.20) 0.884   
## f1e (mean (sd)) 1.76 (1.08) 1.63 (1.10) 0.214   
## f1f (mean (sd)) 2.02 (1.06) 2.01 (1.07) 0.931   
## f1g (mean (sd)) 1.80 (1.11) 1.66 (1.08) 0.182   
## f1h (mean (sd)) 1.48 (1.17) 1.48 (1.15) 0.982   
## f1i (mean (sd)) 1.79 (1.14) 1.98 (1.11) 0.063   
## f1j (mean (sd)) 1.47 (1.18) 1.58 (1.15) 0.283   
## f1k (mean (sd)) 1.99 (1.13) 2.08 (1.09) 0.351   
## f1l (mean (sd)) 1.08 (1.01) 1.08 (1.02) 0.991   
## f1m (mean (sd)) 1.57 (1.06) 1.45 (1.10) 0.240   
## f1n (mean (sd)) 2.00 (1.12) 2.01 (1.03) 0.930   
## f1o (mean (sd)) 1.17 (1.14) 1.27 (1.11) 0.323   
## f1p (mean (sd)) 1.43 (1.12) 1.25 (1.15) 0.095   
## f1q (mean (sd)) 1.10 (1.18) 1.04 (1.14) 0.607   
## f1r (mean (sd)) 1.97 (1.03) 1.92 (1.00) 0.605   
## f1s (mean (sd)) 1.22 (1.15) 1.23 (1.12) 0.978   
## f1t (mean (sd)) 1.59 (1.13) 1.47 (1.06) 0.257   
## cesd (mean (sd)) 32.86 (12.64) 32.84 (12.42) 0.990   
## indtot (mean (sd)) 35.87 (7.09) 35.59 (7.23) 0.676   
## drugrisk (mean (sd)) 1.62 (3.87) 2.16 (4.75) 0.191   
## sexrisk (mean (sd)) 4.48 (2.89) 4.80 (2.70) 0.221   
## satreat (mean (sd)) 0.31 (0.46) 0.26 (0.44) 0.292   
## substance (%) 0.043   
## alcohol 102 (44.7) 75 (33.3)   
## cocaine 68 (29.8) 84 (37.3)   
## heroin 58 (25.4) 66 (29.3)   
## drinkstatus (mean (sd)) 0.69 (0.47) 0.66 (0.48) 0.614   
## daysdrink (mean (sd)) 89.07 (85.28) 97.63 (85.63) 0.436   
## anysubstatus (mean (sd)) 0.77 (0.42) 0.77 (0.42) 0.957   
## daysanysub (mean (sd)) 72.35 (78.74) 77.90 (79.88) 0.586   
## linkstatus (mean (sd)) 0.17 (0.37) 0.58 (0.50) <0.001   
## e2b1 (mean (sd)) 2.13 (1.80) 2.13 (1.90) 0.999   
## g1b1 (mean (sd)) 0.09 (0.28) 0.13 (0.34) 0.286   
## i11 (mean (sd)) 14.70 (14.22) 19.17 (31.76) 0.377

## Statistical Analyses

The sample was stratified based on whether adults received the usual (0) treatment (treat) or the HELP clinic (1) treatment (treat) from an inpatient detoxification unit. The treat variable was recoded as factors Usual and HELP to analyze the data. The two treatment groups were compared to see if and what statistically significant differences existed. Two of the three significant categories [How many times hospitalized for medical problems (d1) and Post-detox linkage to primary care (linkstatus)] were analyzed. The drug type from which the adult was detoxing (substance) was not analyzed because the variable is categorical. Histograms were created to visualize the distribution of the data, and Levene’s test was conducted to analyze variance. Kruskal-Wallis tests were conducted to determine statistically significant differences between treat and d1 and treat and linkstatus. The Two-Sample Wilcoxon Test (also known as Wilcoxon rank sum test or Mann-Whitney test) was also conducted to compare the independent groups whose data are not normally distributed. R Commander (Rcmdr) version 2.4-0 within RStudion version 1.0.153 was used to conduct the analyses. The code can be found at <https://github.com/gdaniel99/N736Fall2017_HELPdataset.git> in the HW 3.Rmd file.

#To access R-commander  
library(Rcmdr)

## Loading required package: splines

## Loading required package: RcmdrMisc

## Loading required package: car

##   
## Attaching package: 'car'

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

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

## Loading required package: sandwich

## Loading required package: effects

##   
## Attaching package: 'effects'

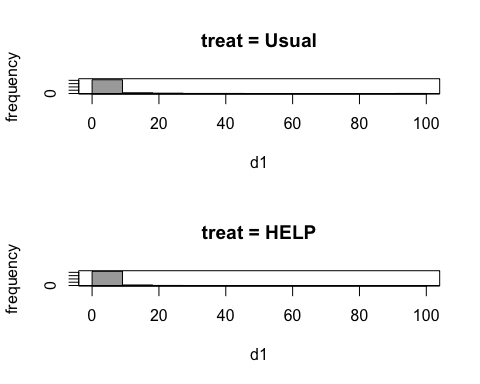
## The following object is masked from 'package:car':  
##   
## Prestige

## The Commander GUI is launched only in interactive sessions

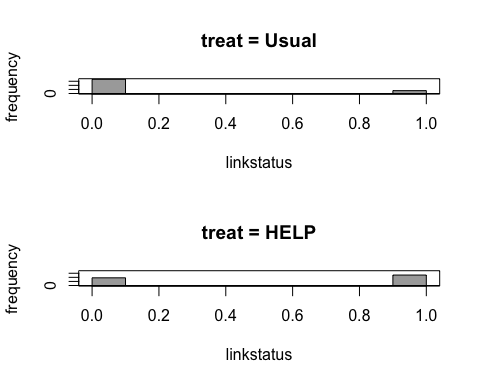
library(splines)  
library(RcmdrMisc)  
library(car)

#To change treat to factor  
help.spss <- within(help.spss, {  
 treat <- factor(treat, labels=c('Usual','HELP'))  
})

#Histogram; lifetime hospitalizations by treatment grp  
with(help.spss, Hist(d1, groups=treat, scale="frequency",   
 breaks="Sturges", col="darkgray"))



#Histogram; post-detox link to primary care by treatment grp  
with(help.spss, Hist(linkstatus, groups=treat, scale="frequency",  
 breaks="Sturges", col="darkgray"))



#Levene's test for variance; lifetime hospitalizations by treatment grp  
with(help.spss, tapply(d1, treat, var, na.rm=TRUE))

## Usual HELP   
## 65.915353 9.777302

leveneTest(d1 ~ treat, data=help.spss, center="median")

## Levene's Test for Homogeneity of Variance (center = "median")  
## Df F value Pr(>F)   
## group 1 3.7661 0.05292 .  
## 451   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#Levene's test for variance; post-detox link to primary care by treatment grp  
with(help.spss, tapply(linkstatus, treat, var, na.rm=TRUE))

## Usual HELP   
## 0.1400902 0.2452407

leveneTest(linkstatus ~ treat, data=help.spss, center="median")

## Levene's Test for Homogeneity of Variance (center = "median")  
## Df F value Pr(>F)   
## group 1 36.306 3.633e-09 \*\*\*  
## 429   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

#Kruskal-Wallis Test; lifetime hospitalizations by treatment grp  
with(help.spss, tapply(d1, treat, median, na.rm=TRUE))

## Usual HELP   
## 2 2

kruskal.test(d1 ~ treat, data=help.spss)

##   
## Kruskal-Wallis rank sum test  
##   
## data: d1 by treat  
## Kruskal-Wallis chi-squared = 1.5427, df = 1, p-value = 0.2142

#Kruskal-Wallis Test; post-detox link to primary care by treatment grp  
with(help.spss, tapply(linkstatus, treat, median, na.rm=TRUE))

## Usual HELP   
## 0 1

kruskal.test(linkstatus ~ treat, data=help.spss)

##   
## Kruskal-Wallis rank sum test  
##   
## data: linkstatus by treat  
## Kruskal-Wallis chi-squared = 76.442, df = 1, p-value < 2.2e-16

#Two-sample Wilcoxon Test; lifetime hospitalizations by treatment grp  
with(help.spss, tapply(d1, treat, median, na.rm=TRUE))

## Usual HELP   
## 2 2

wilcox.test(d1 ~ treat, alternative="two.sided", data=help.spss)

##   
## Wilcoxon rank sum test with continuity correction  
##   
## data: d1 by treat  
## W = 27348, p-value = 0.2143  
## alternative hypothesis: true location shift is not equal to 0

#Two-sample Wilcoxon Test; post-detox link to primary care by treatment grp  
with(help.spss, tapply(linkstatus, treat, median, na.rm=TRUE))

## Usual HELP   
## 0 1

wilcox.test(linkstatus ~ treat, alternative="two.sided",   
 data=help.spss)

##   
## Wilcoxon rank sum test with continuity correction  
##   
## data: linkstatus by treat  
## W = 13708, p-value < 2.2e-16  
## alternative hypothesis: true location shift is not equal to 0

## Results

The overall sample included 453 adults ranging in age from 19 to 60 years old. Consisting mostly of males (76%), half of the sample was homeless with the majority of the sample being non-white (63%). The sample was divided into two groups: 228 adults received the usual treatment and 225 adults received the HELP clinic (intervention) treatment. Of all the variables observed, how many times hospitalized for medical problems (d1), post-detox linkage to primary care (linkstatus), and the drug type from which the adult was detoxing (substance) showed a significantly statistical difference between the treatment groups. After further analyzing d1 and linkstatus, it can be concluded that when adults are in an inpatient detoxification unit and the HELP clinic treatment is implemented, the adult is more likely to see a primary care provider after detoxification.