R Notebook

Load libraries

```
library(readxl)
library(rstatix)
library(data.table)
library(chisq.posthoc.test)
library(dplyr)
options(dplyr.summarise.inform = FALSE)
```

Load Data

```
data <- read_excel("pain_v_3_first_1_add_Nov 22 2022 request.xlsx")
```

Data Cleaning

Removing rows where pain is due to cancer

```
# Remove rows where pain is due to cancer

data <- data[-which(data$DrQ6_Cancer == -1),]

dim(data)
```

```
## [1] 541 240
```

Cleaning and analyzing response variable

```
# Setting response variable: 0: No chronic pain (left blank), 1 - having chronic pain (some response av
data$DrQ3 <- ifelse(is.na(data$DrQ3), "No", "Yes")
data$DrQ3 <- as.factor(data$DrQ3)
# Creating table
table(data$DrQ3, useNA = "always")
```

```
##
##   No  Yes <NA>
##  291  250    0
```

As you can see from the above table, 291 people do not suffer from chronic pain, while 250 suffer from chronic pain. In this table, there are no null values.

Cleaning and analyzing explanatory variable

```
# Changing date of birth to age

# Created a new feature called age from birthd_bs
data$age <- as.numeric(difftime(Sys.Date(),data$birthd_bs, units = "weeks"))/52.25


# Unstable housing (1=yes, 0=no)
data$unstabh_1mb <- as.factor(data$unstabh_1mb)
unique(data$unstabh_1mb)
```

```
## [1] 0    1    <NA>
## Levels: 0 1
```

```
table(data$unstabh_1mb, useNA = "always")
```

```
##
##    0    1 <NA>
##  380  160    1
```

```
data$unstabh_1mb[is.na(data$unstabh_1mb)] <- 0


# incar_3m (incarceration past month) (1=yes, 0=no)
data$incar_3m <- as.factor(data$incar_3m)
unique(data$incar_3m)
```

```
## [1] 0 1
## Levels: 0 1
```

```
table(data$incar_3m, useNA = "always")
```

```
##
##    0    1 <NA>
##  484   57    0
```

```
# education CO_scolar (values: 0 to 9)
data$CO_scolar <- as.factor(data$CO_scolar)
unique(data$CO_scolar)
```

```
## [1] 4 8 6 7 5 9 3 2 1
## Levels: 1 2 3 4 5 6 7 8 9
```

```
table(data$CO_scolar, useNA = "always")
```

```
##
##    1    2    3    4    5    6    7    8    9 <NA>
##    2    6   17  179  141   59   66   34   37    0
```

```
# OAT_3m : OAT past 3 months (1=yes, 0=no)
data$OAT_3m <- as.factor(data$OAT_3m)
unique(data$OAT_3m)
```

```
## [1] 0 1
## Levels: 0 1
```

```
table(data$OAT_3m, useNA = "always")
```

```
##
##    0    1 <NA>
##  319  222    0
```

```
# cocaiv_pm (iv cocaine use past month) (1=yes, 0=no)
data$cocaiv_pm <- as.factor(data$cocaiv_pm)
unique(data$cocaiv_pm)
```

```
## [1] 1 0
## Levels: 0 1
```

```
table(data$cocaiv_pm, useNA = "always")
```

```
##
##    0    1 <NA>
##  343  198    0
```

```
# heroiv_pm (iv heroine use past month) (1=yes, 0=no)
data$heroiv_pm <- as.factor(data$heroiv_pm)
unique(data$heroiv_pm)
```

```
## [1] 0 1
## Levels: 0 1
```

```
table(data$heroiv_pm, useNA = "always")
```

```
##
##    0    1 <NA>
##  413  128    0
```

```
# opiiv_pm (iv opioid use past month) (1=yes, 0=no)
data$opiiv_pm <- as.factor(data$opiiv_pm)
unique(data$opiiv_pm)
```

```
## [1] 1 0
## Levels: 0 1
```

```
table(data$opiiv_pm, useNA = "always")
```

```
##
##     0     1  <NA>
##   389   152     0
```

```
# Gender (1 Male, 2 Female, 3 Other)
unique(data$VBA3)
```

```
## [1] 1 2 3
```

```
# Removing others as there frequency is less than 5
data <- data[-which(data$VBA3 == 3),]
data$VBA3 <- as.factor(data$VBA3)
table(data$VBA3, useNA = "always")
```

```
##
##     1     2  <NA>
##   452    87     0
```

```
# NW_jourcons (how many days in total have you consumed all substances listed above in last month)
unique(data$NW_jourcons)
```

```
##  [1] 25 20  1 11  8 30  0 15  6  2  5  4  3  7 28 16 12 26 27 10 21 13 17 29 23
## [26] 18 14 22  9 24
```

```
table(data$NW_jourcons, useNA = "always")
```

```
##
##     0     1     2     3     4     5     6     7     8     9    10    11    12    13    14    15
##    41    10    13     8    21     7    15    11     9     4    19     4     7     3     3    24
##    16    17    18    20    21    22    23    24    25    26    27    28    29    30  <NA>
##     7     2     4    32     4     3     2     2    19     5     4    15     6   235     0
```

```
# CO_njalc1 (how many days in last month did you consume alcohol)
unique(data$CO_njalc1)
```

```
##  [1] 20 12  0  8  3 15 30  1  5  2 10 25 16 26 29 27  4 28 14 21 13  6  7 22 11
## [26] 24 18 17
```

```
table(data$CO_njalc1, useNA = "always")
```

```
##
##     0     1     2     3     4     5     6     7     8    10    11    12    13    14    15    16
##   242    38    21    18    22    15     8     5    10    20     2    11     4     2    18     4
##    17    18    20    21    22    24    25    26    27    28    29    30  <NA>
##     1     1    15     2     1     1     7     3     2     9     2    55     0
```

```
# CO_ihero1m (in last month did you inject heroin)
unique(data$CO_ihero1m)
```

```
## [1] 2 1
```

```
table(data$CO_ihero1m, useNA = "always")
```

```
##
##    1    2 <NA>
##  128  411    0
```

```
# Q22AJ (how many days)
unique(data$Q22AJ)
```

```
##  [1] NA  2  1 30 15  4 10 20 25  5 12  3 28  7 14  6 23  8 29 17
```

```
table(data$Q22AJ, useNA = "always")
```

```
##
##    1    2    3    4    5    6    7    8   10   12   14   15   17   20   23   25
##   18   23    6    9    5    5    8    4    7    1    2    9    1    4    1    2
##   28   29   30 <NA>
##    1    2   20  411
```

```
data$Q22AJ[is.na(data$Q22AJ)] <- 0
```

```
# COMD_nihero1m (on average how many times per day)
unique(data$COMD_nihero1m)
```

```
## [1] NA  1  2  3
```

```
table(data$COMD_nihero1m, useNA = "always")
```

```
##
##    1    2    3 <NA>
##  109   15    4  411
```

```
data$COMD_nihero1m[is.na(data$COMD_nihero1m)] <- 0
```

```
# COMD_fshero1m (in last month did you sniff heroin)
unique(data$COMD_fshero1m)
```

```
## [1] 2 1
```

```
table(data$COMD_fshero1m, useNA = "always")
```

```
##
##    1    2 <NA>
##   10  529    0
```

```
# Q22BJ (how many days)
unique(data$Q22BJ)
```

```
## [1] NA  1  8  5 15 30  2
```

```
table(data$Q22BJ, useNA = "always")
```

```
##
##    1    2    5    8   15   30 <NA>
##    4    1    1    1    2    1  529
```

```
data$Q22BJ[is.na(data$Q22BJ)] <- 0

# COMD_nfshero1m (how many times per day)
unique(data$COMD_nfshero1m)
```

```
## [1] NA  1  2
```

```
table(data$COMD_nfshero1m, useNA = "always")
```

```
##
##    1    2 <NA>
##    8    2  529
```

```
data$COMD_nfshero1m[is.na(data$COMD_nfshero1m)] <- 0

# CO_ipoud1m (in last month did you inject powdered cocaine)
unique(data$CO_ipoud1m)
```

```
## [1] 1 2
```

```
table(data$CO_ipoud1m, useNA = "always")
```

```
##
##    1    2 <NA>
##  192  347    0
```

```
# MD_Q23AJ_1 (how many days)
unique(data$MD_Q23AJ_1)
```

```
##  [1]  4 NA  1 20  3  2 10  5 30 28  8 15 16 12  7  6 13 22 29  9 17 14 25 21 23
```

```
table(data$MD_Q23AJ_1, useNA = "always")
```

```
##
##    1    2    3    4    5    6    7    8    9   10   12   13   14   15   16   17
##   36   31    8   23    8    5    6    8    1   13    3    1    1   11    2    2
##   20   21   22   23   25   28   29   30 <NA>
##    9    1    2    1    3    5    2   10  347
```

```
data$MD_Q23AJ_1[is.na(data$MD_Q23AJ_1)] <- 0
```

```
# COMD_nicoke1m_p (how many times per day)
unique(data$COMD_nicoke1m_p)
```

```
## [1]  4 NA  2  1  3  5
```

```
table(data$COMD_nicoke1m_p, useNA = "always")
```

```
##
##    1    2    3    4    5 <NA>
##  112   52   22    5    1  347
```

```
data$COMD_nicoke1m_p[is.na(data$COMD_nicoke1m_p)] <- 0
```

```
# CO_icrak1m (in last month did you inject crack) (1=yes 2=no)
unique(data$CO_icrak1m)
```

```
## [1] 2 1
```

```
table(data$CO_icrak1m, useNA = "always")
```

```
##
##    1    2 <NA>
##   28  511    0
```

```
data$CO_icrak1m <- as.factor(data$CO_icrak1m)
```

```
# MD_Q23AJ_2 (how many days)
unique(data$MD_Q23AJ_2)
```

```
##  [1] NA  2  4  1 10  5 20  6 11 15  3
```

```
table(data$MD_Q23AJ_2, useNA = "always")
```

```
##
##    1    2    3    4    5    6   10   11   15   20 <NA>
##    8    4    1    4    5    1    2    1    1    1  511
```

```
data$MD_Q23AJ_2[is.na(data$MD_Q23AJ_2)] <- 0
```

```
# COMD_nicoke1m_c (how many times per day) (1= 1-3, 2= 4-6, 3= 6+)
unique(data$COMD_nicoke1m_c)
```

```
## [1] NA  1  3  4  2
```

```r
table(data$COMD_nicoke1m_c, useNA = "always")
```

```
##
##    1    2    3    4 <NA>
##   22    2    3    1  511
```

```r
data$COMD_nicoke1m_c[is.na(data$COMD_nicoke1m_c)] <- 0
data$COMD_nicoke1m_c <- as.factor(data$COMD_nicoke1m_c)

# CO_fpoud1m (in last month did you smoke cocaine powder) (1=yes 2=no)
unique(data$CO_fpoud1m)
```

```
## [1] 2 1
```

```r
table(data$CO_fpoud1m, useNA = "always")
```

```
##
##    1    2 <NA>
##   11  528    0
```

```r
data$CO_fpoud1m <- as.factor(data$CO_fpoud1m)

# MD_Q23BJ_1 (how many days)
unique(data$MD_Q23BJ_1)
```

```
## [1] NA  1  8  2 20 29
```

```r
table(data$MD_Q23BJ_1, useNA = "always")
```

```
##
##    1    2    8   20   29 <NA>
##    6    2    1    1    1  528
```

```r
data$MD_Q23BJ_1[is.na(data$MD_Q23BJ_1)] <- 0

# COMD_nfcoke1m_p (how many times per day) (1= 1-3, 2= 4-6, 3= 6+)
unique(data$COMD_nfcoke1m_p)
```

```
## [1] NA  1  2  3
```

```r
table(data$COMD_nfcoke1m_p, useNA = "always")
```

```
##
##    1    2    3 <NA>
##    9    1    1  528
```

```
data$COMD_nfcoke1m_p[is.na(data$COMD_nfcoke1m_p)] <- 0
data$COMD_nfcoke1m_p <- as.factor(data$COMD_nfcoke1m_p)

# COMD_fcrakfree1m (1=yes 2=no)
unique(data$COMD_fcrakfree1m)
```

```
## [1] 1 2
```

```
table(data$COMD_fcrakfree1m, useNA = "always")
```

```
##
##    1    2 <NA>
##  201  338    0
```

```
data$COMD_fcrakfree1m[is.na(data$COMD_fcrakfree1m)] <- 0
data$COMD_fcrakfree1m <- as.factor(data$COMD_fcrakfree1m)

# MD_Q23BJ_2
unique(data$MD_Q23BJ_2)
```

```
##  [1]  1 NA  2  5  4 15  8 20 30 12  7 14  3 28 18 25 10  6 23 27 11 29 26
```

```
table(data$MD_Q23BJ_2, useNA = "always")
```

```
##
##    1    2    3    4    5    6    7    8   10   11   12   14   15   18   20   23
##   30   25   15    9   18    3    3    8   10    1    5    2   16    4   10    2
##   25   26   27   28   29   30 <NA>
##    6    1    2    3    1   27  338
```

```
data$MD_Q23BJ_2[is.na(data$MD_Q23BJ_2)] <- 0

# COMD_nfcoke1m_cf (1= 1-3, 2= 4-6, 3= 6+)
unique(data$COMD_nfcoke1m_cf)
```

```
## [1]        1       NA 88888888        5        2        3        4
```

```
table(data$COMD_nfcoke1m_cf, useNA = "always")
```

```
##
##         1        2        3        4        5 88888888     <NA>
##        83       54       24       24       15        1      338
```

```
data$COMD_nfcoke1m_cf[is.na(data$COMD_nfcoke1m_cf)] <- 0
data$COMD_nfcoke1m_cf <- as.factor(data$COMD_nfcoke1m_cf)

# CO_scoke1m (1=yes 2=no)
unique(data$CO_scoke1m)
```

```
## [1] 2 1
```

```r
table(data$CO_scoke1m, useNA = "always")
```

```
## 
##    1    2 <NA>
##   49  490    0
```

```r
data$CO_scoke1m <- as.factor(data$CO_scoke1m)

# Q23CJ
unique(data$Q23CJ)
```

```
##  [1] NA  3 30 16  1  4  8  2 28 14  5 10  6 20
```

```r
table(data$Q23CJ, useNA = "always")
```

```
## 
##    1    2    3    4    5    6    8   10   14   16   20   28   30 <NA>
##   15    6    6    4    5    1    2    2    2    1    1    1    3  490
```

```r
data$Q23CJ[is.na(data$Q23CJ)] <- 0

# COMD_nscoke1m (how many times per day) (1= 1-3, 2= 4-6, 3= 6+)
unique(data$COMD_nscoke1m)
```

```
## [1] NA  1  2  3  5  4
```

```r
table(data$COMD_nscoke1m, useNA = "always")
```

```
## 
##    1    2    3    4    5 <NA>
##   26   14    6    2    1  490
```

```r
data$COMD_nscoke1m[is.na(data$COMD_nscoke1m)] <- 0
data$COMD_nscoke1m <- as.factor(data$COMD_nscoke1m)

# CO_ispeed1m (1=yes 2=no)
unique(data$CO_ispeed1m)
```

```
## [1] 2 1
```

```r
table(data$CO_ispeed1m, useNA = "always")
```

```
## 
##    1    2 <NA>
##   27  512    0
```

```
data$CO_ispeed1m <- as.factor(data$CO_ispeed1m)

# Q2223AJ
unique(data$Q2223AJ)
```

```
##  [1] NA  2 10 30  1  3 20  4 25 29  5
```

```
table(data$Q2223AJ, useNA = "always")
```

```
##
##    1    2    3    4    5   10   20   25   29   30 <NA>
##    6    6    2    1    1    3    2    1    2    3  512
```

```
data$Q2223AJ[is.na(data$Q2223AJ)] <- 0

# COMD_nispeed1m (1= 1-3, 2= 4-6, 3= 6+)
unique(data$COMD_nispeed1m)
```

```
## [1] NA  1  2
```

```
table(data$COMD_nispeed1m, useNA = "always")
```

```
##
##    1    2 <NA>
##   25    2  512
```

```
data$COMD_nispeed1m[is.na(data$COMD_nispeed1m)] <- 0
data$COMD_nispeed1m <- as.factor(data$COMD_nispeed1m)

# NH_Q7Ia_suboxIV
unique(data$NH_Q7Ia_suboxIV)
```

```
## [1]  2 NA  1
```

```
table(data$NH_Q7Ia_suboxIV, useNA = "always")
```

```
##
##    1    2 <NA>
##    1  154  384
```

```
data$NH_Q7Ia_suboxIV[is.na(data$NH_Q7Ia_suboxIV)] <- 2
data$NH_Q7Ia_suboxIV <- as.factor(data$NH_Q7Ia_suboxIV)


# NH_Q7Ia_suboxIVj
unique(data$NH_Q7Ia_suboxIVj)
```

```
## [1] NA  1
```

```r
table(data$NH_Q7Ia_suboxIVj, useNA = "always")
```

```
## 
##    1 <NA>
##    1  538
```

```r
data$NH_Q7Ia_suboxIVj[is.na(data$NH_Q7Ia_suboxIVj)] <- 2

# NH_Q7Ib_subox
unique(data$NH_Q7Ib_subox)
```

```
## [1]  2 NA  1
```

```r
table(data$NH_Q7Ib_subox, useNA = "always")
```

```
## 
##    1    2 <NA>
##    2  153  384
```

```r
data$NH_Q7Ib_subox[is.na(data$NH_Q7Ib_subox)] <- 2
data$NH_Q7Ib_subox <- as.factor(data$NH_Q7Ib_subox)

# NH_Q7Ib_suboxj
unique(data$NH_Q7Ib_suboxj)
```

```
## [1] NA  1  5
```

```r
table(data$NH_Q7Ib_suboxj, useNA = "always")
```

```
## 
##    1    5 <NA>
##    1    1  537
```

```r
data$NH_Q7Ib_suboxj[is.na(data$NH_Q7Ib_suboxj)] <- 0

# NH_Q7Ja_methaIV
unique(data$NH_Q7Ja_methaIV)
```

```
## [1]  2 NA  1
```

```r
table(data$NH_Q7Ja_methaIV, useNA = "always")
```

```
## 
##    1    2 <NA>
##    2  153  384
```

```
data$NH_Q7Ja_methaIV[is.na(data$NH_Q7Ja_methaIV)] <- 2
data$NH_Q7Ja_methaIV <- as.factor(data$NH_Q7Ja_methaIV)

# NH_Q7Ja_methaIVj
unique(data$NH_Q7Ja_methaIVj)
```

```
## [1] NA  1  2
```

```
table(data$NH_Q7Ja_methaIVj, useNA = "always")
```

```
##
##    1    2 <NA>
##    1    1  537
```

```
data$NH_Q7Ja_methaIVj[is.na(data$NH_Q7Ja_methaIVj)] <- 0

# NH_Q7Jb_metha
unique(data$NH_Q7Jb_metha)
```

```
## [1]  2 NA  1
```

```
table(data$NH_Q7Jb_metha, useNA = "always")
```

```
##
##    1    2 <NA>
##    6  149  384
```

```
data$NH_Q7Jb_metha[is.na(data$NH_Q7Jb_metha)] <- 2
data$NH_Q7Jb_metha <- as.factor(data$NH_Q7Jb_metha)

# NH_Q7Jb_methaj
unique(data$NH_Q7Jb_methaj)
```

```
## [1] NA  7  2  3  1
```

```
table(data$NH_Q7Jb_methaj, useNA = "always")
```

```
##
##    1    2    3    7 <NA>
##    1    3    1    1  533
```

```
data$NH_Q7Jb_methaj[is.na(data$NH_Q7Jb_methaj)] <- 0

# CO_iopiac1m (1=yes 2=no)
unique(data$CO_iopiac1m)
```

```
## [1] 1 2
```

```r
table(data$CO_iopiac1m, useNA = "always")
```

```
## 
##    1    2 <NA>
##  152  387    0
```

```r
data$CO_iopiac1m <- as.factor(data$CO_iopiac1m)


# Q26BJ
unique(data$Q26BJ)
```

```
##  [1]  1  8 NA 10  2  9 25  5  7 28  4 20 30 29  3 23 15 22  6 26 27 18 12 14 21
## [26] 19
```

```r
table(data$Q26BJ, useNA = "always")
```

```
## 
##    1    2    3    4    5    6    7    8    9   10   12   14   15   18   19   20
##   12   12    5    9   12    4    4    5    1    8    2    1    7    1    1    5
##   21   22   23   25   26   27   28   29   30 <NA>
##    1    1    1    3    2    2    1    2   50  387
```

```r
data$Q26BJ[is.na(data$Q26BJ)] <- 0

# CO_niopiac1m (1= 1-3, 2= 4-6, 3= 6+)
unique(data$CO_niopiac1m)
```

```
## [1]  1  2 NA  3  4
```

```r
table(data$CO_niopiac1m, useNA = "always")
```

```
## 
##    1    2    3    4 <NA>
##   98   42    7    5  387
```

```r
data$CO_niopiac1m[is.na(data$CO_niopiac1m)] <- 0
data$CO_niopiac1m <- as.factor(data$CO_niopiac1m)

# CO_autopiac1m
unique(data$CO_autopiac1m)
```

```
## [1] 2 1
```

```r
table(data$CO_autopiac1m, useNA = "always")
```

```
## 
##    1    2 <NA>
##   33  506    0
```

```r
data$CO_autopiac1m <- as.factor(data$CO_autopiac1m)

# Q26AJ
unique(data$Q26AJ)
```

```
##  [1] NA  4  7 23  1 12  2  3 10 24  6 15 20
```

```r
table(data$Q26AJ, useNA = "always")
```

```
##
##    1    2    3    4    6    7   10   12   15   20   23   24 <NA>
##   14    4    4    1    1    2    1    1    2    1    1    1  506
```

```r
data$Q26AJ[is.na(data$Q26AJ)] <- 0

# NH_Q26AJnf (1= 1-3, 2= 4-6, 3= 6+)
unique(data$NH_Q26AJnf)
```

```
## [1] NA  1  2  3
```

```r
table(data$NH_Q26AJnf, useNA = "always")
```

```
##
##    1    2    3 <NA>
##   12    1    1  525
```

```r
data$NH_Q26AJnf[is.na(data$NH_Q26AJnf)] <- 0
data$NH_Q26AJnf <- as.factor(data$NH_Q26AJnf)

# NMD_CO_iamph1m
unique(data$NMD_CO_iamph1m)
```

```
## [1] 2 1
```

```r
table(data$NMD_CO_iamph1m, useNA = "always")
```

```
##
##    1    2 <NA>
##   57  482    0
```

```r
data$NMD_CO_iamph1m <- as.factor(data$NMD_CO_iamph1m)

# NMD_Q27BJ
unique(data$NMD_Q27BJ)
```

```
##  [1] NA  1  2  5 30  8  4  3 16 10 24 26 14  6 15 25 18 20
```

```
table(data$NMD_Q27BJ, useNA = "always")
```

```
##
##    1    2    3    4    5    6    8   10   14   15   16   18   20   24   25   26
##   13    7    5    4    2    2    3    3    1    2    1    1    2    1    1    1
##   30 <NA>
##    8  482
```

```
data$NMD_Q27BJ[is.na(data$NMD_Q27BJ)] <- 0
```

```
# NMD_CO_niamph1m (1= 1-3, 2= 4-6, 3= 6+)
unique(data$NMD_CO_niamph1m)
```

```
## [1] NA  1  2  4  3
```

```
table(data$NMD_CO_niamph1m, useNA = "always")
```

```
##
##    1    2    3    4 <NA>
##   42   10    2    3  482
```

```
data$NMD_CO_niamph1m[is.na(data$NMD_CO_niamph1m)] <- 0
data$NMD_CO_niamph1m <- as.factor(data$NMD_CO_niamph1m)
```

```
# NMD_COMD_fsautamph1m
unique(data$NMD_COMD_fsautamph1m)
```

```
## [1] 1 2
```

```
table(data$NMD_COMD_fsautamph1m, useNA = "always")
```

```
##
##    1    2 <NA>
##  107  432    0
```

```
data$NMD_COMD_fsautamph1m <- as.factor(data$NMD_COMD_fsautamph1m)
```

```
# NMD_Q27AJ
unique(data$NMD_Q27AJ)
```

```
##  [1] 10 NA  1 15 30 25  7  2  3 14  5  4 28 21  8  6  9 20 24 16
```

```
table(data$NMD_Q27AJ, useNA = "always")
```

```
##
##    1    2    3    4    5    6    7    8    9   10   14   15   16   20   21   24
##   20   13   17    2    6    5    2    3    1   10    1    6    1    2    1    1
##   25   28   30 <NA>
##    2    1   13  432
```

```r
data$NMD_Q27AJ[is.na(data$NMD_Q27AJ)] <- 0

# NMD_COMD_nfsaamph1m (1= 1-3, 2= 4-6, 3= 6+)
unique(data$NMD_COMD_nfsaamph1m)
```

```
## [1]  3 NA  1  4  2
```

```r
table(data$NMD_COMD_nfsaamph1m, useNA = "always")
```

```
##
##    1    2    3    4 <NA>
##   78   20    6    3  432
```

```r
data$NMD_COMD_nfsaamph1m[is.na(data$NMD_COMD_nfsaamph1m)] <- 0
data$NMD_COMD_nfsaamph1m <- as.factor(data$NMD_COMD_nfsaamph1m)

# NW_tranquiliv1m (1=yes 2=no)
unique(data$NW_tranquiliv1m)
```

```
## [1] 2 1
```

```r
table(data$NW_tranquiliv1m, useNA = "always")
```

```
##
##    1    2 <NA>
##    5  534    0
```

```r
data$NW_tranquiliv1m <- as.factor(data$NW_tranquiliv1m)

# NW_Q29BJ
unique(data$NW_Q29BJ)
```

```
## [1] NA  1  3  7
```

```r
table(data$NW_Q29BJ, useNA = "always")
```

```
##
##    1    3    7 <NA>
##    3    1    1  534
```

```r
data$NW_Q29BJ[is.na(data$NW_Q29BJ)] <- 0

# NW_tranquiliv1m_fj (1= 1-3, 2= 4-6, 3= 6+)
unique(data$NW_tranquiliv1m_fj)
```

```
## [1] NA  1
```

```r
table(data$NW_tranquiliv1m_fj, useNA = "always")
```

```
## 
##    1 <NA> 
##    5  534
```

```r
data$NW_tranquiliv1m_fj[is.na(data$NW_tranquiliv1m_fj)] <- 0
data$NW_tranquiliv1m_fj <- as.factor(data$NW_tranquiliv1m_fj)

# NW_tranquil1m
unique(data$NW_tranquil1m)
```

```
## [1] 2 1
```

```r
table(data$NW_tranquil1m, useNA = "always")
```

```
## 
##    1    2 <NA> 
##   57  482    0
```

```r
data$NW_tranquil1m <- as.factor(data$NW_tranquil1m)

# Q29AJ
unique(data$Q29AJ)
```

```
##  [1] NA  1  6  5  4  2 30  3 10 15 20 28 12 27  9  7
```

```r
table(data$Q29AJ, useNA = "always")
```

```
## 
##    1    2    3    4    5    6    7    9   10   12   15   20   27   28   30 <NA> 
##    9    8    7    7    5    2    1    1    3    1    3    1    1    1    7  482
```

```r
data$Q29AJ[is.na(data$Q29AJ)] <- 0

# NW_tranquil1m_fj (1= 1-3, 2= 4-6, 3= 6+)
unique(data$NW_tranquil1m_fj)
```

```
## [1] NA  1  2  3
```

```r
table(data$NW_tranquil1m_fj, useNA = "always")
```

```
## 
##    1    2    3 <NA> 
##   46    7    4  482
```

```r
data$NW_tranquil1m_fj[is.na(data$NW_tranquil1m_fj)] <- 0
data$NW_tranquil1m_fj <- as.factor(data$NW_tranquil1m_fj)

# CO_mari1m
unique(data$CO_mari1m)
```

```
## [1] 1 2
```

```r
table(data$CO_mari1m, useNA = "always")
```

```
##
##    1    2 <NA>
##  298  241    0
```

```r
data$CO_mari1m <- as.factor(data$CO_mari1m)

# Q30AJ
unique(data$Q30AJ)
```

```
##  [1]  7 NA  1  3  2  4 30 15 20  6 18 10 12  9 25 28 27 29 23  8 16  5 26
```

```r
table(data$Q30AJ, useNA = "always")
```

```
##
##    1    2    3    4    5    6    7    8    9   10   12   15   16   18   20   23
##   29   18   14   11   11    7    7    4    2   24    5   16    1    2   18    1
##   25   26   27   28   29   30 <NA>
##    6    1    2    6    2  111  241
```

```r
data$Q30AJ[is.na(data$Q30AJ)] <- 0

# NW_mari1m_fj (1= 1-3, 2= 4-6, 3= 6+)
unique(data$NW_mari1m_fj)
```

```
## [1]        1       NA        3        2 88888888
```

```r
table(data$NW_mari1m_fj, useNA = "always")
```

```
##
##        1        2        3 88888888     <NA>
##      205       60       31        2      241
```

```r
data$NW_mari1m_fj[is.na(data$NW_mari1m_fj)] <- 0
data$NW_mari1m_fj <- as.factor(data$NW_mari1m_fj)

# NW_psychiv1m (1=yes 2=no)
unique(data$NW_psychiv1m)
```

```
## [1] 2 1
```

```r
table(data$NW_psychiv1m, useNA = "always")
```

```
## 
##    1    2 <NA>
##    3  536    0
```

```r
data$NW_psychiv1m <- as.factor(data$NW_psychiv1m)

# NW_Q31BJ
unique(data$NW_Q31BJ)
```

```
## [1] NA  1  3
```

```r
table(data$NW_Q31BJ, useNA = "always")
```

```
## 
##    1    3 <NA>
##    2    1  536
```

```r
data$NW_Q31BJ[is.na(data$NW_Q31BJ)] <- 0

# NW_psychiv1m_fj
unique(data$NW_psychiv1m_fj)
```

```
## [1] NA  1
```

```r
table(data$NW_psychiv1m_fj, useNA = "always")
```

```
## 
##    1 <NA>
##    3  536
```

```r
data$NW_psychiv1m_fj[is.na(data$NW_psychiv1m_fj)] <- 0
data$NW_psychiv1m_fj <- as.factor(data$NW_psychiv1m_fj)

# NW_psych1m (1=yes 2=no)
unique(data$NW_psych1m)
```

```
## [1] 2 1
```

```r
table(data$NW_psych1m, useNA = "always")
```

```
## 
##    1    2 <NA>
##   28  511    0
```

```
data$NW_psych1m <- as.factor(data$NW_psych1m)

# Q31AJ
unique(data$Q31AJ)
```

```
## [1] NA  4  3  2  1 10  5 15
```

```
table(data$Q31AJ, useNA = "always")
```

```
##
##    1    2    3    4    5   10   15 <NA>
##   12    3    3    5    2    2    1  511
```

```
data$Q31AJ[is.na(data$Q31AJ)] <- 0

# NW_psych1m_fj (1= 1-3, 2= 4-6, 3= 6+)
unique(data$NW_psych1m_fj)
```

```
## [1] NA  1  3  2
```

```
table(data$NW_psych1m_fj, useNA = "always")
```

```
##
##    1    2    3 <NA>
##   21    3    4  511
```

```
data$NW_psych1m_fj[is.na(data$NW_psych1m_fj)] <- 0
data$NW_psych1m_fj <- as.factor(data$NW_psych1m_fj)

# NMD_NW_autmediv1m (1=yes 2=no)
unique(data$NMD_NW_autmediv1m)
```

```
## [1] 2 1
```

```
table(data$NMD_NW_autmediv1m, useNA = "always")
```

```
##
##    1    2 <NA>
##    2  537    0
```

```
data$NMD_NW_autmediv1m <- as.factor(data$NMD_NW_autmediv1m)

# NMD_NW_Q32BJ
unique(data$NMD_NW_Q32BJ)
```

```
## [1] NA  4  1
```

```
table(data$NMD_NW_Q32BJ, useNA = "always")
```

```
##
##    1    4 <NA>
##    1    1  537
```

```
data$NMD_NW_Q32BJ[is.na(data$NMD_NW_Q32BJ)] <- 0

# NMD_NW_autmediv1m_fs (1= 1-3, 2= 4-6, 3= 6+)
unique(data$NMD_NW_autmediv1m_fs)
```

```
## [1] NA  2  1
```

```
table(data$NMD_NW_autmediv1m_fs, useNA = "always")
```

```
##
##    1    2 <NA>
##    1    1  537
```

```
data$NMD_NW_autmediv1m_fs[is.na(data$NMD_NW_autmediv1m_fs)] <- 0
data$NMD_NW_autmediv1m_fs <- as.factor(data$NMD_NW_autmediv1m_fs)

# NMD_NW_autmed1m (1=yes 2=no)
unique(data$NMD_NW_autmed1m)
```

```
## [1] 2 1
```

```
table(data$NMD_NW_autmed1m, useNA = "always")
```

```
##
##    1    2 <NA>
##   17  522    0
```

```
data$NMD_NW_autmed1m <- as.factor(data$NMD_NW_autmed1m)

# NMD_NW_Q32AJ
unique(data$NMD_NW_Q32AJ)
```

```
##  [1] NA  2  8 30  5  3 20 10  4  1 16 14
```

```
table(data$NMD_NW_Q32AJ, useNA = "always")
```

```
##
##    1    2    3    4    5    8   10   14   16   20   30 <NA>
##    1    1    2    4    1    1    2    1    1    2    1  522
```

```r
data$NMD_NW_Q32AJ[is.na(data$NMD_NW_Q32AJ)] <- 0


# NMD_NW_autmed1m_fs (1= 1-3, 2= 4-6, 3= 6+)
unique(data$NMD_NW_autmed1m_fs)
```

```
## [1] NA  1  2  3
```

```r
table(data$NMD_NW_autmed1m_fs, useNA = "always")
```

```
##
##    1    2    3 <NA>
##   14    2    1  522
```

```r
data$NMD_NW_autmed1m_fs[is.na(data$NMD_NW_autmed1m_fs)] <- 0
data$NMD_NW_autmed1m_fs <- as.factor(data$NMD_NW_autmed1m_fs)

# CO_iaut11m (1=yes 2=no)
unique(data$CO_iaut11m)
```

```
## [1] 2 1
```

```r
table(data$CO_iaut11m, useNA = "always")
```

```
##
##    1    2 <NA>
##    1  538    0
```

```r
data$CO_iaut11m <- as.factor(data$CO_iaut11m)

# CO_fsaut11m (1=yes 2=no)
unique(data$CO_fsaut11m)
```

```
## [1] 2 1
```

```r
table(data$CO_fsaut11m, useNA = "always")
```

```
##
##    1    2 <NA>
##    3  536    0
```

```r
data$CO_fsaut11m <- as.factor(data$CO_fsaut11m)


# NW_Q21C1 (1- yes, 2-No)
unique(data$NW_Q21C1)
```

```
## [1]  1  2 NA
```

```r
table(data$NW_Q21C1, useNA = "always")
```

```
##
##    1    2 <NA>
##  320  217    2
```

```r
data$NW_Q21C1 <- as.factor(data$NW_Q21C1)


# age_40 (greater than and equal to 40 = 1,  less than 40 = 0)
data$age_40 <- ifelse(data$age >= 40, 1, 0)
data$age_40 <- as.factor(data$age_40)



new_data <- data %>% select(DrQ3, age, age_40, unstabh_1mb, incar_3m, CO_scolar, OAT_3m, cocaiv_pm, her
CO_icrak1m, MD_Q23AJ_2, COMD_nicoke1m_c, CO_fpoud1m, MD_Q23BJ_1, COMD_nfcoke1m_p, COMD_fcrakfree1m, MD_
NMD_NW_autmed1m, NMD_NW_Q32AJ, CO_iaut11m, NMD_NW_autmed1m_fs,CO_fsaut11m, NW_Q21C1)
```

```r
# Creating groups by age
new_data <- new_data %>%
  mutate(
    # Create categories
    age_group = dplyr::case_when(
      age < 30 ~ "0-30",
      age >= 30 & age < 40 ~ "30-40",
      age >= 40 & age < 50 ~ "40-50",
      age >= 50 & age < 60 ~ "50-60",
      age >= 60 & age < 70 ~ "60-70",
      age >= 70           ~ ">=70"
    ),
    # Convert to factor
    age_group = factor(
      age_group,
      level = c("0-30","30-40", "40-50", "50-60", "60-70", ">=70")
    )
  )

table(new_data$age_group)
```

```
##
##  0-30 30-40 40-50 50-60 60-70  >=70
##    16    99   157   169    93     5
```

```r
# Creating groups by CO_njalc1

table(new_data$CO_njalc1)
```

```
##
##    0    1    2    3    4    5    6    7    8   10   11   12   13   14   15   16   17   18   20   21
##  242   38   21   18   22   15    8    5   10   20    2   11    4    2   18    4    1    1   15    2
##   22   24   25   26   27   28   29   30
##    1    1    7    3    2    9    2   55
```

```r
new_data <- new_data %>%
  mutate(
    # Create categories
    CO_njalc1_group = dplyr::case_when(
      CO_njalc1 == 0 ~ "0",
      CO_njalc1 >= 1 & CO_njalc1 <= 10 ~ "1-10",
      CO_njalc1 >= 11 & CO_njalc1 <= 20 ~ "11-20",
      CO_njalc1 >= 21 & CO_njalc1 <= 29 ~ "21-29",
      CO_njalc1 == 30 ~ "30"
    ),
    # Convert to factor
    CO_njalc1_group = factor(
      CO_njalc1_group,
      level = c("0","1-10", "11-20", "21-29", "30")
    )
  )

table(new_data$CO_njalc1_group)
```

```
## 
##      0  1-10 11-20 21-29    30
##    242   157    58    27    55
```

PART 1

```r
# Age summary
summarize(new_data, mean(age), median(age), IQR(age))
```

```
## # A tibble: 1 x 3
##   `mean(age)` `median(age)` `IQR(age)`
##         <dbl>         <dbl>      <dbl>
## 1        49.5          49.9       17.3
```

In this sample, the mean age is 49.47, the median is 49.87, and the interquartile range is 17.26.

```r
# grouping into genders
new_data.group.gender <- group_by(new_data, VBA3)

# Male vs Female median age
summarize(new_data.group.gender, median(age), mean(age), IQR(age))
```

```
## # A tibble: 2 x 4
##   VBA3  `median(age)` `mean(age)` `IQR(age)`
##   <fct>         <dbl>       <dbl>      <dbl>
## 1 1              51.1        50.4       16.7
## 2 2              43.4        44.9       14.1
```

The following are descriptive statistics for men: 51.09 is the median, 50.38 is the mean, and 16.71 is the IQR, while for women: 43.43 is the median, 44.89 is the mean, and 14.13 is the IQR.

```r
# Age proportion from 40 years gender wise
men_greater_39 <- length(which(new_data.group.gender$age >= 40 & new_data.group.gender$VBA3 == 1))
men_less_40 <- length(which(new_data.group.gender$age < 40 & new_data.group.gender$VBA3 == 1))
women_greater_39 = length(which(new_data.group.gender$age >= 40 & new_data.group.gender$VBA3 == 2))
women_less_40 = length(which(new_data.group.gender$age < 40 & new_data.group.gender$VBA3 == 2))


r.men <- c(men_less_40, men_greater_39)
r.women <- c(women_less_40, women_greater_39)

r.df <- as.table(cbind(r.men, r.women))

dimnames(r.df) <- list(division_with_40 = c("less than 40", "greater than and equal to 40"),
                       gender = c("Men", "Women"))

r.df
```

```
##                                gender
## division_with_40              Men Women
##    less than 40                87    28
##    greater than and equal to 40 365    59
```

There are 87 Men who are less than 40 and 365 men greater than and equal to 40. There are 28 Women who are less than 40 and 59 Women who are greater than and equal to 40.

```r
# Unstable housing
table_unstabh_1mb <- table(new_data.group.gender$unstabh_1mb)
addmargins(table_unstabh_1mb)
```

```
##
##   0   1 Sum
## 380 159 539
```

```r
addmargins(prop.table(table_unstabh_1mb))
```

```
##
##         0         1       Sum
## 0.7050093 0.2949907 1.0000000
```

```r
# Unstable housing
table_unstabh_1mb <- table(new_data.group.gender$unstabh_1mb)
addmargins(table_unstabh_1mb)
```

```
##
##   0   1 Sum
## 380 159 539
```

```r
addmargins(prop.table(table_unstabh_1mb))
```

```
##
##         0         1       Sum
## 0.7050093 0.2949907 1.0000000
```

```r
# incar_3m
table_incar_3m <- table(new_data.group.gender$incar_3m)
addmargins(table_incar_3m)
```

```
##
##   0   1 Sum
## 482  57 539
```

```r
addmargins(prop.table(table_incar_3m))
```

```
##
##         0         1       Sum
## 0.8942486 0.1057514 1.0000000
```

```r
# CO_scolar
table_CO_scolar <- table(new_data.group.gender$CO_scolar)
addmargins(table_CO_scolar)
```

```
##
##   1   2   3   4   5   6   7   8   9 Sum
##   2   6  16 179 140  59  66  34  37 539
```

```r
addmargins(prop.table(table_CO_scolar))
```

```
##
##           1           2           3           4           5           6
## 0.003710575 0.011131725 0.029684601 0.332096475 0.259740260 0.109461967
##           7           8           9         Sum
## 0.122448980 0.063079777 0.068645640 1.000000000
```

```r
# OAT
table_OAT <- table(new_data.group.gender$OAT_3m)
addmargins(table_OAT)
```

```
##
##   0   1 Sum
## 317 222 539
```

```r
addmargins(prop.table(table_OAT))
```

```
##
##         0         1       Sum
## 0.5881262 0.4118738 1.0000000
```

```r
# cocaiv_pm
table_cocaiv_pm <- table(new_data.group.gender$cocaiv_pm)
addmargins(table_cocaiv_pm)
```

```
##
##   0   1 Sum
## 341 198 539
```

```
addmargins(prop.table(table_cocaiv_pm))
```

```
##
##         0         1       Sum
## 0.6326531 0.3673469 1.0000000
```

```
# heroiv_pm
table_heroiv_pm <- table(new_data.group.gender$heroiv_pm)
addmargins(table_heroiv_pm)
```

```
##
##   0   1 Sum
## 411 128 539
```

```
addmargins(prop.table(table_heroiv_pm))
```

```
##
##         0         1       Sum
## 0.7625232 0.2374768 1.0000000
```

```
# opiiv_pm
table_opiiv_pm <- table(new_data.group.gender$opiiv_pm)
addmargins(table_opiiv_pm)
```

```
##
##   0   1 Sum
## 387 152 539
```

```
addmargins(prop.table(table_opiiv_pm))
```

```
##
##         0         1       Sum
## 0.7179963 0.2820037 1.0000000
```

PART 2

```
new_data.group.cp.gender <- group_by(new_data, DrQ3, VBA3)
```

```
# Division of cp by age
table_cp_gender <- table(new_data.group.cp.gender$DrQ3, new_data.group.cp.gender$VBA3)
addmargins(table_cp_gender)
```

```
##
##         1   2 Sum
##   No  237  52 289
##   Yes 215  35 250
##   Sum 452  87 539
```

```
addmargins(prop.table(table_cp_gender))
```

```
##
##            1          2        Sum
##   No  0.43970315 0.09647495 0.53617811
##   Yes 0.39888683 0.06493506 0.46382189
##   Sum 0.83858998 0.16141002 1.00000000
```

```
# Summary of chronic pain by age
new_data.group.cp.gender %>% summarise(mean_age = mean(age), median_age = median(age), iqr_age = IQR(age
```

```
## # A tibble: 4 x 5
## # Groups:   DrQ3 [2]
##   DrQ3  VBA3  mean_age median_age iqr_age
##   <fct> <fct>    <dbl>      <dbl>   <dbl>
## 1 No    1         49.3       49.6    17.3
## 2 No    2         43.1       42.0    12.2
## 3 Yes   1         51.6       52.5    15.8
## 4 Yes   2         47.5       45.9    14.1
```

```
# division of cp by unstabh_1mb
table_cp_unstabh_1mb <- table(new_data.group.cp.gender$DrQ3, new_data.group.cp.gender$unstabh_1mb)
addmargins(table_cp_unstabh_1mb)
```

```
##
##         0   1 Sum
##   No   202  87 289
##   Yes  178  72 250
##   Sum  380 159 539
```

```
addmargins(prop.table(table_cp_unstabh_1mb))
```

```
##
##            0         1       Sum
##   No   0.3747681 0.1614100 0.5361781
##   Yes  0.3302412 0.1335807 0.4638219
##   Sum  0.7050093 0.2949907 1.0000000
```

```
# division by cp and incar_3m
table_cp_incar_3m <- table(new_data.group.cp.gender$DrQ3, new_data.group.cp.gender$incar_3m)
addmargins(table_cp_incar_3m)
```

```
##
##         0   1 Sum
##   No   262  27 289
##   Yes  220  30 250
##   Sum  482  57 539
```

```r
addmargins(prop.table(table_cp_incar_3m))
```

```
## 
##              0          1        Sum
##   No  0.48608534 0.05009276 0.53617811
##   Yes 0.40816327 0.05565863 0.46382189
##   Sum 0.89424861 0.10575139 1.00000000
```

```r
# division by cp and CO_scolar
table_cp_CO_scolar <- table(new_data.group.cp.gender$DrQ3, new_data.group.cp.gender$CO_scolar)
addmargins(table_cp_CO_scolar)
```

```
## 
##        1   2   3   4   5   6   7   8   9 Sum
##   No   0   3  11  97  79  30  37  13  19 289
##   Yes  2   3   5  82  61  29  29  21  18 250
##   Sum  2   6  16 179 140  59  66  34  37 539
```

```r
addmargins(prop.table(table_cp_CO_scolar))
```

```
## 
##              1          2          3          4          5          6
##   No  0.000000000 0.005565863 0.020408163 0.179962894 0.146567718 0.055658627
##   Yes 0.003710575 0.005565863 0.009276438 0.152133581 0.113172542 0.053803340
##   Sum 0.003710575 0.011131725 0.029684601 0.332096475 0.259740260 0.109461967
## 
##              7          8          9        Sum
##   No  0.068645640 0.024118738 0.035250464 0.536178108
##   Yes 0.053803340 0.038961039 0.033395176 0.463821892
##   Sum 0.122448980 0.063079777 0.068645640 1.000000000
```

```r
# division by cp and OAT
table_cp_OAT <- table(new_data.group.cp.gender$DrQ3, new_data.group.cp.gender$OAT_3m)
addmargins(table_cp_OAT)
```

```
## 
##         0   1 Sum
##   No  176 113 289
##   Yes 141 109 250
##   Sum 317 222 539
```

```r
addmargins(prop.table(table_cp_OAT))
```

```
## 
##             0         1        Sum
##   No  0.3265306 0.2096475 0.5361781
##   Yes 0.2615955 0.2022263 0.4638219
##   Sum 0.5881262 0.4118738 1.0000000
```

```
# division by cp and cocaiv_pm
table_cp_cocaiv_pm <- table(new_data.group.cp.gender$DrQ3, new_data.group.cp.gender$cocaiv_pm)
addmargins(table_cp_cocaiv_pm)
```

```
##
##          0   1 Sum
##   No   175 114 289
##   Yes  166  84 250
##   Sum  341 198 539
```

```
addmargins(prop.table(table_cp_cocaiv_pm))
```

```
##
##              0         1       Sum
##   No   0.3246753 0.2115028 0.5361781
##   Yes  0.3079777 0.1558442 0.4638219
##   Sum  0.6326531 0.3673469 1.0000000
```

```
# division by cp and heroiv_pm
table_cp_heroiv_pm <- table(new_data.group.cp.gender$DrQ3, new_data.group.cp.gender$heroiv_pm)
addmargins(table_cp_heroiv_pm)
```

```
##
##          0   1 Sum
##   No   214  75 289
##   Yes  197  53 250
##   Sum  411 128 539
```

```
addmargins(prop.table(table_cp_heroiv_pm))
```

```
##
##               0          1         Sum
##   No   0.39703154 0.13914657 0.53617811
##   Yes  0.36549165 0.09833024 0.46382189
##   Sum  0.76252319 0.23747681 1.00000000
```

```
# division by cp and opiiv_pm
table_cp_opiiv_pm  <- table(new_data.group.cp.gender$DrQ3, new_data.group.cp.gender$opiiv_pm)
addmargins(table_cp_opiiv_pm)
```

```
##
##          0   1 Sum
##   No   205  84 289
##   Yes  182  68 250
##   Sum  387 152 539
```

```
addmargins(prop.table(table_cp_opiiv_pm))
```

```
##
##              0           1         Sum
##   No  0.3803340 0.1558442 0.5361781
##   Yes 0.3376623 0.1261596 0.4638219
##   Sum 0.7179963 0.2820037 1.0000000
```

A total of 543 participants (median age of 45 years, 13.7% women) were included. The prevalence of chronic pain was 46% (n=252). 108/252 (n=43%) noted using alcohol, cannabis, or an illegal drug to help manage their chronic pain. A total of 87/252 (35%) participants were prescribed opioid (n=41, 47%) and/or non-opioid medications (n=59, 68%). 24% (n= 21) of participants reported using their prescribed medication at an increased dose or frequency than that recommended; 23% (n=20) of participants consumed their prescribed medication by a route different from the recommended mode of consumption; 23% (n=20) of participants reported using medication that was not prescribed to them to help manage their pain. Overall, prescribed pain medication was misused in at least one of the above manners by 30% of the participants who reported chronic pain (n=75). Of the participants who received a chronic pain diagnosis (n=156, 62%), 32% (n=50) visited a physician in the last 3 months due to their pain, and 23/31 (74%) were prescribed pain medication when requested.

PART 3

Chi-square test for association (independence) tests

```
# chi-sq for cp and age_group
table_cp_age_group <- table(new_data.group.cp.gender$DrQ3, new_data.group.cp.gender$age_group)
(chisq_age_group <- chisq.test(table_cp_age_group))
```

```
## Warning in chisq.test(table_cp_age_group): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_age_group
## X-squared = 11.02, df = 5, p-value = 0.05099
```

```
chisq_age_group$observed
```

```
##
##       0-30 30-40 40-50 50-60 60-70 >=70
##   No    13    60    88    83    43    2
##   Yes    3    39    69    86    50    3
```

```
chisq_age_group$expected
```

```
##
##           0-30    30-40    40-50   50-60    60-70     >=70
##   No  8.57885 53.08163 84.17996 90.6141 49.86456 2.680891
##   Yes 7.42115 45.91837 72.82004 78.3859 43.13544 2.319109
```

```
chisq.posthoc.test(table_cp_age_group)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension     Value      0-30     30-40      40-50     50-60     60-70
## 1        No Residuals  2.250032  1.543206  0.7261901 -1.417551 -1.569164
## 2        No  p values  0.293363  1.000000  1.0000000  1.000000  1.000000
## 3       Yes Residuals -2.250032 -1.543206 -0.7261901  1.417551  1.569164
## 4       Yes  p values  0.293363  1.000000  1.0000000  1.000000  1.000000
##            >=70
## 1 -0.6134595
## 2  1.0000000
## 3  0.6134595
## 4  1.0000000
```

The p-value of the above ch-square test is < 0.1. Therefore it is statistically significant.

```
# chi-sq for cp and age_40
table_cp_age_40 <- table(new_data.group.cp.gender$DrQ3, new_data.group.cp.gender$age_40)
(chisq_age <- chisq.test(table_cp_age_40))
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_age_40
## X-squared = 5.2226, df = 1, p-value = 0.0223
```

```
chisq_age$observed
```

```
##
##        0   1
##   No   73 216
##   Yes  42 208
```

```
chisq_age$expected
```

```
##
##             0         1
##   No  61.66048 227.3395
##   Yes 53.33952 196.6605
```

```
chisq.posthoc.test(table_cp_age_40)
```

```
##    Dimension     Value        0         1
## 1        No Residuals  2.390710 -2.390710
## 2        No  p values  0.067263  0.067263
## 3       Yes Residuals -2.390710  2.390710
## 4       Yes  p values  0.067263  0.067263
```

The p-value of the above ch-square test is < 0.1. Therefore it is statistically significant.

```
# Table for Age group and Alcohol frequency
(table_age_group_alchol_freq <- table(new_data.group.cp.gender$age_group, new_data.group.cp.gender$CO_n_
```

```
##
##          0 1-10 11-20 21-29 30
##   0-30   6    8     0     1  1
##   30-40 50   27    12     2  8
##   40-50 67   48    18     6 18
##   50-60 79   46    23     6 15
##   60-70 40   23     5    12 13
##   >=70   0    5     0     0  0
```

```
(chisq_age_group_alchol_freq <- chisq.test(table_age_group_alchol_freq))
```

```
## Warning in chisq.test(table_age_group_alchol_freq): Chi-squared approximation
## may be incorrect
```

```
##
##   Pearson's Chi-squared test
##
## data:  table_age_group_alchol_freq
## X-squared = 39.687, df = 20, p-value = 0.005471
```

```
chisq_age_group_alchol_freq$observed
```

```
##
##          0 1-10 11-20 21-29 30
##   0-30   6    8     0     1  1
##   30-40 50   27    12     2  8
##   40-50 67   48    18     6 18
##   50-60 79   46    23     6 15
##   60-70 40   23     5    12 13
##   >=70   0    5     0     0  0
```

```
chisq_age_group_alchol_freq$expected
```

```
##
##                   0       1-10      11-20     21-29        30
##   0-30    7.183673  4.660482  1.7217069 0.8014842  1.6326531
##   30-40 44.448980 28.836735 10.6530612 4.9591837 10.1020408
##   40-50 70.489796 45.730983 16.8942486 7.8645640 16.0204082
##   50-60 75.877551 49.226345 18.1855288 8.4656772 17.2448980
##   60-70 41.755102 27.089054 10.0074212 4.6586271  9.4897959
##   >=70   2.244898  1.456401  0.5380334 0.2504638  0.5102041
```

```
chisq.posthoc.test(table_age_group_alchol_freq)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

34

```
##      Dimension     Value          0        1-10      11-20      21-29         30
## 1        0-30 Residuals -0.6039741  1.8654110 -1.4100840  0.2309673 -0.5304371
## 2        0-30  p values  1.0000000  1.0000000  1.0000000  1.0000000  1.0000000
## 3       30-40 Residuals  1.2414411 -0.4496806  0.4835045 -1.5090175 -0.7724606
## 4       30-40  p values  1.0000000  1.0000000  1.0000000  1.0000000  1.0000000
## 5       40-50 Residuals -0.6651441  0.4734322  0.3382771 -0.8103303  0.6199737
## 6       40-50  p values  1.0000000  1.0000000  1.0000000  1.0000000  1.0000000
## 7       50-60 Residuals  0.5828387 -0.6592775  1.4424508 -1.0494416 -0.6885437
## 8       50-60  p values  1.0000000  1.0000000  1.0000000  1.0000000  1.0000000
## 9       60-70 Residuals -0.4022450 -1.0259249 -1.8420518  3.8364953  1.3219121
## 10      60-70  p values  1.0000000  1.0000000  1.0000000  0.0037440  1.0000000
## 11       >=70 Residuals -2.0278611  3.5042139 -0.7800999 -0.5158883 -0.7572991
## 12       >=70  p values  1.0000000  0.0137390  1.0000000  1.0000000  1.0000000
```

The p-value of the above ch-square test is < 0.1. Therefore it is statistically significant.

```
# chi-sq for cp and NW_Q21C1
table_cp_NW_Q21C1 <- table(new_data.group.cp.gender$DrQ3, new_data.group.cp.gender$NW_Q21C1)
(chisq_NW_Q21C1 <- chisq.test(table_cp_NW_Q21C1))
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_NW_Q21C1
## X-squared = 3.1849, df = 1, p-value = 0.07432
```

```
chisq_NW_Q21C1$observed
```

```
##
##         1    2
##   No  161 127
##   Yes 159  90
```

```
chisq_NW_Q21C1$expected
```

```
##
##            1         2
##   No  171.6201 116.3799
##   Yes 148.3799 100.6201
```

```
chisq.posthoc.test(table_cp_NW_Q21C1)
```

```
##    Dimension     Value         1         2
## 1         No Residuals -1.872793  1.872793
## 2         No  p values  0.244388  0.244388
## 3        Yes Residuals  1.872793 -1.872793
## 4        Yes  p values  0.244388  0.244388
```

The p-value of the above ch-square test is < 0.1. Therefore it is statistically significant.

```
# cp and alcohol frequency group (CO_njalc1_group)
(table_cp_alcohol_group <- table(new_data$DrQ3, new_data$CO_njalc1_group))
```

```
##
##         0 1-10 11-20 21-29  30
##   No  143   83    23    11  29
##   Yes  99   74    35    16  26
```

```
chisq.test(table_cp_alcohol_group)
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_alcohol_group
## X-squared = 9.3151, df = 4, p-value = 0.05369
```

```
(chisq_cp_alcohol_group <- chisq.test(table_cp_alcohol_group))
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_alcohol_group
## X-squared = 9.3151, df = 4, p-value = 0.05369
```

```
chisq_cp_alcohol_group$observed
```

```
##
##         0 1-10 11-20 21-29  30
##   No  143   83    23    11  29
##   Yes  99   74    35    16  26
```

```
chisq_cp_alcohol_group$expected
```

```
##
##            0      1-10    11-20    21-29      30
##   No  129.7551 84.17996 31.09833 14.47681 29.4898
##   Yes 112.2449 72.82004 26.90167 12.52319 25.5102
```

```
chisq.posthoc.test(table_cp_alcohol_group)
```

```
##   Dimension     Value         0       1-10      11-20      21-29         30
## 1      No Residuals  2.299995 -0.2243113 -2.257216 -1.376665 -0.1397576
## 2      No  p values  0.214485  1.0000000  0.239946  1.000000  1.0000000
## 3     Yes Residuals -2.299995  0.2243113  2.257216  1.376665  0.1397576
## 4     Yes  p values  0.214485  1.0000000  0.239946  1.000000  1.0000000
```

The p-value of the above ch-square test is < 0.1. Therefore it is statistically significant.

```
# chisq for cp and Gender
(table_cp_VBA3 <- table(data$DrQ3, data$VBA3))
```

```
##
##          1    2
##    No  237   52
##    Yes 215   35
```

```
chisq.test(table_cp_VBA3)
```

```
##
##   Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_VBA3
## X-squared = 1.2978, df = 1, p-value = 0.2546
```

```
chisq.posthoc.test(table_cp_VBA3)
```

```
##    Dimension     Value        1          2
## 1        No Residuals -1.256587   1.256587
## 2        No  p values  0.835613   0.835613
## 3       Yes Residuals  1.256587  -1.256587
## 4       Yes  p values  0.835613   0.835613
```

```
# chi-sq for cp and unstabh_1mb
chisq.test(table_cp_unstabh_1mb)
```

```
##
##   Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_unstabh_1mb
## X-squared = 0.055841, df = 1, p-value = 0.8132
```

```
chisq.posthoc.test(table_cp_unstabh_1mb)
```

```
##    Dimension     Value        0          1
## 1        No Residuals -0.3310061  0.3310061
## 2        No  p values  1.0000000  1.0000000
## 3       Yes Residuals  0.3310061 -0.3310061
## 4       Yes  p values  1.0000000  1.0000000
```

```
# chi-sq for cp and incar_3m
chisq.test(table_cp_incar_3m)
```

```
##
##   Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_incar_3m
## X-squared = 0.73971, df = 1, p-value = 0.3898
```

```
chisq.posthoc.test(table_cp_incar_3m)
```

```
##   Dimension    Value         0          1
## 1        No Residuals  1.000497 -1.000497
## 2        No  p values  1.000000  1.000000
## 3       Yes Residuals -1.000497  1.000497
## 4       Yes  p values  1.000000  1.000000
```

```
# chi-sq for cp and CO_scolar
chisq.test(table_cp_CO_scolar)
```

```
## Warning in chisq.test(table_cp_CO_scolar): Chi-squared approximation may be
## incorrect
```

```
##
##   Pearson's Chi-squared test
##
## data:  table_cp_CO_scolar
## X-squared = 7.937, df = 8, p-value = 0.4397
```

```
chisq.posthoc.test(table_cp_CO_scolar)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension    Value         1          2          3          4          5
## 1        No Residuals -1.523355 -0.178699  1.232183  0.1878178  0.7751141
## 2        No  p values  1.000000  1.000000  1.000000  1.0000000  1.0000000
## 3       Yes Residuals  1.523355  0.178699 -1.232183 -0.1878178 -0.7751141
## 4       Yes  p values  1.000000  1.000000  1.000000  1.0000000  1.0000000
##           6         7        8          9
## 1 -0.4521731  0.424808 -1.85817 -0.2864581
## 2  1.0000000  1.000000  1.00000  1.0000000
## 3  0.4521731 -0.424808  1.85817  0.2864581
## 4  1.0000000  1.000000  1.00000  1.0000000
```

```
# chi-sq for cp and OAT_3m
chisq.test(new_data$DrQ3, new_data$OAT_3m)
```

```
##
##   Pearson's Chi-squared test with Yates' continuity correction
##
## data:  new_data$DrQ3 and new_data$OAT_3m
## X-squared = 0.94234, df = 1, p-value = 0.3317
```

```
chisq.posthoc.test(table_cp_OAT)
```

```
##   Dimension    Value         0          1
## 1        No Residuals  1.058489 -1.058489
## 2        No  p values  1.000000  1.000000
## 3       Yes Residuals -1.058489  1.058489
## 4       Yes  p values  1.000000  1.000000
```

```
# chi-sq for cp and cocaiv
chisq.test(table_cp_cocaiv_pm)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_cocaiv_pm
## X-squared = 1.7279, df = 1, p-value = 0.1887
```

```
chisq.posthoc.test(table_cp_cocaiv_pm)
```

```
##   Dimension      Value        0        1
## 1       No Residuals -1.404073  1.404073
## 2       No  p values  0.641189  0.641189
## 3      Yes Residuals  1.404073 -1.404073
## 4      Yes  p values  0.641189  0.641189
```

```
# chi-sq for cp and heroiv_pm
chisq.test(table_cp_heroiv_pm)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_heroiv_pm
## X-squared = 1.4192, df = 1, p-value = 0.2335
```

```
chisq.posthoc.test(table_cp_heroiv_pm)
```

```
##   Dimension      Value        0        1
## 1       No Residuals -1.292776  1.292776
## 2       No  p values  0.784354  0.784354
## 3      Yes Residuals  1.292776 -1.292776
## 4      Yes  p values  0.784354  0.784354
```

```
# chi-sq for cp and opiiv_pm
chisq.test(table_cp_opiiv_pm)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_opiiv_pm
## X-squared = 0.14752, df = 1, p-value = 0.7009
```

```
chisq.posthoc.test(table_cp_opiiv_pm)
```

```
##   Dimension      Value         0          1
## 1       No Residuals -0.4800518  0.4800518
## 2       No  p values  1.0000000  1.0000000
## 3      Yes Residuals  0.4800518 -0.4800518
## 4      Yes  p values  1.0000000  1.0000000
```

```r
# cp and alcohol frequecy
(table_cp_alcohol <- table(new_data$DrQ3, new_data$CO_njalc1))
```

```
## 
##         0   1   2   3   4   5   6   7   8  10  11  12  13  14  15  16  17  18
##   No  143  23   9   8  11  10   3   4   4  11   1   6   1   0   7   2   0   1
##   Yes  99  15  12  10  11   5   5   1   6   9   1   5   3   2  11   2   1   0
## 
##        20  21  22  24  25  26  27  28  29  30
##   No    5   0   1   0   3   2   0   4   1  29
##   Yes  10   2   0   1   4   1   2   5   1  26
```

```r
chisq.test(table_cp_alcohol)
```

```
## Warning in chisq.test(table_cp_alcohol): Chi-squared approximation may be
## incorrect
```

```
## 
##  Pearson's Chi-squared test
## 
## data:  table_cp_alcohol
## X-squared = 26.617, df = 27, p-value = 0.4846
```

```r
chisq.posthoc.test(table_cp_alcohol)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension     Value         0          1          2          3          4
## 1        No Residuals  2.299995  0.8857705 -1.008668 -0.7937986 -0.3474371
## 2        No  p values  1.000000  1.0000000  1.000000  1.0000000  1.0000000
## 3       Yes Residuals -2.299995 -0.8857705  1.008668  0.7937986  0.3474371
## 4       Yes  p values  1.000000  1.0000000  1.000000  1.0000000  1.0000000
##           5          6         7          8         10         11          12
## 1  1.027819 -0.9210178  1.188473 -0.8716531  0.1263174 -0.1027869  0.06233398
## 2  1.000000  1.0000000  1.000000  1.0000000  1.0000000  1.0000000  1.00000000
## 3 -1.027819  0.9210178 -1.188473  0.8716531 -0.1263174  0.1027869 -0.06233398
## 4  1.000000  1.0000000  1.000000  1.0000000  1.0000000  1.0000000  1.00000000
##          13        14        15         16        17        18        20
## 1 -1.152003 -1.523355 -1.274537 -0.1456341 -1.076173  0.9309457 -1.597747
## 2  1.000000  1.000000  1.000000  1.0000000  1.000000  1.0000000  1.000000
## 3  1.152003  1.523355  1.274537  0.1456341  1.076173 -0.9309457  1.597747
## 4  1.000000  1.000000  1.000000  1.0000000  1.000000  1.0000000  1.000000
##          21        22        24         25         26        27        28
## 1 -1.523355  0.9309457 -1.076173 -0.5746411  0.4544801 -1.523355 -0.5565142
## 2  1.000000  1.0000000  1.000000  1.0000000  1.0000000  1.000000  1.0000000
## 3  1.523355 -0.9309457  1.076173  0.5746411 -0.4544801  1.523355  0.5565142
## 4  1.000000  1.0000000  1.000000  1.0000000  1.0000000  1.000000  1.0000000
##           29         30
## 1 -0.1027869 -0.1397576
## 2  1.0000000  1.0000000
## 3  0.1027869  0.1397576
## 4  1.0000000  1.0000000
```

```
# cp and CO_ihero1m (1=yes 2=no)
table_cp_CO_ihero1m <- table(new_data$DrQ3, new_data$CO_ihero1m)
chisq.test(table_cp_CO_ihero1m)
```

```
##
##   Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_CO_ihero1m
## X-squared = 1.4192, df = 1, p-value = 0.2335
```

```
chisq.posthoc.test(table_cp_CO_ihero1m)
```

```
##    Dimension      Value         1          2
## 1        No Residuals   1.292776  -1.292776
## 2        No  p values   0.784354   0.784354
## 3       Yes Residuals  -1.292776   1.292776
## 4       Yes  p values   0.784354   0.784354
```

```
# cp and Q22AJ freq.
table_cp_Q22AJ <- table(new_data$DrQ3, new_data$Q22AJ)
chisq.test(table_cp_Q22AJ)
```

```
## Warning in chisq.test(table_cp_Q22AJ): Chi-squared approximation may be
## incorrect
```

```
##
##   Pearson's Chi-squared test
##
## data:  table_cp_Q22AJ
## X-squared = 17.212, df = 19, p-value = 0.5755
```

```
chisq.posthoc.test(table_cp_Q22AJ)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension      Value          0          1          2          3          4
## 1        No Residuals  -1.292776  0.6484175 -0.1419187 -0.178699 -0.5565142
## 2        No  p values   1.000000  1.0000000  1.0000000  1.000000  1.0000000
## 3       Yes Residuals   1.292776 -0.6484175  0.1419187  0.178699  0.5565142
## 4       Yes  p values   1.000000  1.0000000  1.0000000  1.000000  1.0000000
##            5          6          7          8        10         12         14
## 1  0.2875069   1.188473   1.221839 -0.1456341   0.95113  -1.076173   1.317781
## 2  1.0000000   1.000000   1.000000  1.0000000   1.00000   1.000000   1.000000
## 3 -0.2875069  -1.188473  -1.221839  0.1456341  -0.95113   1.076173  -1.317781
## 4  1.0000000   1.000000   1.000000  1.0000000   1.00000   1.000000   1.000000
##           15         17         20         23         25         28         29
## 1  0.1175558  0.9309457   1.867104 -1.076173 -0.1027869 -1.076173 -1.523355
## 2  1.0000000  1.0000000   1.000000  1.000000  1.0000000  1.000000  1.000000
## 3 -0.1175558 -0.9309457  -1.867104  1.076173  0.1027869  1.076173  1.523355
## 4  1.0000000  1.0000000   1.000000  1.000000  1.0000000  1.000000  1.000000
```

```
##          30
## 1  0.5832641
## 2  1.0000000
## 3 -0.5832641
## 4  1.0000000
```

```
# cp and COMD_nihero1m freq.
table_cp_COMD_nihero1m <- table(new_data$DrQ3, new_data$COMD_nihero1m)
chisq.test(table_cp_COMD_nihero1m)
```

```
## Warning in chisq.test(table_cp_COMD_nihero1m): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_COMD_nihero1m
## X-squared = 4.5694, df = 3, p-value = 0.2062
```

```
chisq.posthoc.test(table_cp_COMD_nihero1m)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension      Value        0         1         2         3
## 1         No Residuals -1.292776  0.7648039  0.5027058  1.867104
## 2         No  p values  1.000000  1.0000000  1.0000000  0.495096
## 3        Yes Residuals  1.292776 -0.7648039 -0.5027058 -1.867104
## 4        Yes  p values  1.000000  1.0000000  1.0000000  0.495096
```

```
# cp and COMD_fshero1m freq
table_cp_COMD_fshero1m <- table(new_data$DrQ3, new_data$COMD_fshero1m)
chisq.test(table_cp_COMD_fshero1m)
```

```
## Warning in chisq.test(table_cp_COMD_fshero1m): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_COMD_fshero1m
## X-squared = 0.0078272, df = 1, p-value = 0.9295
```

```
chisq.posthoc.test(table_cp_COMD_fshero1m)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension      Value         1          2
## 1         No Residuals  0.4085132 -0.4085132
## 2         No  p values  1.0000000  1.0000000
## 3        Yes Residuals -0.4085132  0.4085132
## 4        Yes  p values  1.0000000  1.0000000
```

```r
# cp and Q22BJ freq
table_cp_Q22BJ <- table(new_data$DrQ3, new_data$Q22BJ)
chisq.test(table_cp_Q22BJ)
```

```
## Warning in chisq.test(table_cp_Q22BJ): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_Q22BJ
## X-squared = 6.8016, df = 6, p-value = 0.3396
```

```r
chisq.posthoc.test(table_cp_Q22BJ)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension      Value          0          1          2          5          8
## 1        No Residuals -0.4085132 -1.152003  0.9309457  0.9309457 -1.076173
## 2        No  p values  1.0000000  1.000000  1.0000000  1.0000000  1.000000
## 3       Yes Residuals  0.4085132  1.152003 -0.9309457 -0.9309457  1.076173
## 4       Yes  p values  1.0000000  1.000000  1.0000000  1.0000000  1.000000
##          15        30
## 1  1.317781  0.9309457
## 2  1.000000  1.0000000
## 3 -1.317781 -0.9309457
## 4  1.000000  1.0000000
```

```r
# cp and COMD_nfshero1m
table_cp_COMD_nfshero1m <- table(new_data$DrQ3, new_data$COMD_nfshero1m)
chisq.test(table_cp_COMD_nfshero1m)
```

```
## Warning in chisq.test(table_cp_COMD_nfshero1m): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_COMD_nfshero1m
## X-squared = 0.26741, df = 2, p-value = 0.8748
```

```r
chisq.posthoc.test(table_cp_COMD_nfshero1m)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension      Value          0          1          2
## 1        No Residuals -0.4085132  0.5075537 -0.1027869
## 2        No  p values  1.0000000  1.0000000  1.0000000
## 3       Yes Residuals  0.4085132 -0.5075537  0.1027869
## 4       Yes  p values  1.0000000  1.0000000  1.0000000
```

```
# cp and CO_ipoud1m
table_cp_CO_ipoud1m <- table(new_data$DrQ3, new_data$CO_ipoud1m)
chisq.test(table_cp_CO_ipoud1m)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_CO_ipoud1m
## X-squared = 1.3973, df = 1, p-value = 0.2372
```

```
chisq.posthoc.test(table_cp_CO_ipoud1m)
```

```
##   Dimension     Value        1        2
## 1       No Residuals  1.272249 -1.272249
## 2       No  p values  0.813138  0.813138
## 3      Yes Residuals -1.272249  1.272249
## 4      Yes  p values  0.813138  0.813138
```

```
# cp and MD_Q23AJ_1
table_cp_MD_Q23AJ_1 <- table(new_data$DrQ3, new_data$MD_Q23AJ_1)
chisq.test(table_cp_MD_Q23AJ_1)
```

```
## Warning in chisq.test(table_cp_MD_Q23AJ_1): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_MD_Q23AJ_1
## X-squared = 25.079, df = 24, p-value = 0.4015
```

```
chisq.posthoc.test(table_cp_MD_Q23AJ_1)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension     Value        0          1          2          3          4
## 1       No Residuals -1.272249  0.9332642  0.1404083  0.5075537 -0.5692606
## 2       No  p values  1.000000  1.0000000  1.0000000  1.0000000  1.0000000
## 3      Yes Residuals  1.272249 -0.9332642 -0.1404083 -0.5075537  0.5692606
## 4      Yes  p values  1.000000  1.0000000  1.0000000  1.0000000  1.0000000
##              5          6          7          8          9         10         12
## 1   0.5075537  0.2875069 -1.825173 -0.206732  0.9309457  0.01671209  1.615451
## 2   1.0000000  1.0000000  1.000000  1.000000  1.0000000  1.00000000  1.000000
## 3  -0.5075537 -0.2875069  1.825173  0.206732 -0.9309457 -0.01671209 -1.615451
## 4   1.0000000  1.0000000  1.000000  1.000000  1.0000000  1.00000000  1.000000
##             13         14         15         16         17         20         21
## 1   0.9309457  0.9309457  1.894953 -1.523355 -0.1027869  1.465696  0.9309457
## 2   1.0000000  1.0000000  1.000000  1.000000  1.0000000  1.000000  1.0000000
## 3  -0.9309457 -0.9309457 -1.894953  1.523355  0.1027869 -1.465696 -0.9309457
## 4   1.0000000  1.0000000  1.000000  1.000000  1.0000000  1.000000  1.0000000
```

```
##              22         23        25        28         29        30
## 1 -1.523355 -1.076173 -0.7064904  1.188473 -0.1027869 -0.23157
## 2  1.000000  1.000000  1.0000000  1.000000  1.0000000  1.00000
## 3  1.523355  1.076173  0.7064904 -1.188473  0.1027869  0.23157
## 4  1.000000  1.000000  1.0000000  1.000000  1.0000000  1.00000
```

```
# cp and COMD_nicoke1m_p
table_cp_COMD_nicoke1m_p <- table(new_data$DrQ3, new_data$COMD_nicoke1m_p)
chisq.test(table_cp_COMD_nicoke1m_p)
```

```
## Warning in chisq.test(table_cp_COMD_nicoke1m_p): Chi-squared approximation may
## be incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_COMD_nicoke1m_p
## X-squared = 4.5876, df = 5, p-value = 0.4683
```

```
chisq.posthoc.test(table_cp_COMD_nicoke1m_p)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension     Value         0         1         2          3         4
## 1         No Residuals -1.272249  0.4147074  1.204931  0.08908642  1.188473
## 2         No  p values  1.000000  1.0000000  1.000000  1.00000000  1.000000
## 3        Yes Residuals  1.272249 -0.4147074 -1.204931 -0.08908642 -1.188473
## 4        Yes  p values  1.000000  1.0000000  1.000000  1.00000000  1.000000
##            5
## 1 -1.076173
## 2  1.000000
## 3  1.076173
## 4  1.000000
```

```
# cp and CO_icrak1m
table_cp_CO_icrak1m <- table(new_data$DrQ3, new_data$CO_icrak1m)
chisq.test(table_cp_CO_icrak1m)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_CO_icrak1m
## X-squared = 0.93693, df = 1, p-value = 0.3331
```

```
chisq.posthoc.test(table_cp_CO_icrak1m)
```

```
##    Dimension     Value        1         2
## 1         No Residuals  1.162551 -1.162551
## 2         No  p values  0.980048  0.980048
## 3        Yes Residuals -1.162551  1.162551
## 4        Yes  p values  0.980048  0.980048
```

```
# cp and MD_Q23AJ_2
table_cp_MD_Q23AJ_2 <- table(new_data$DrQ3, new_data$MD_Q23AJ_2)
chisq.test(table_cp_MD_Q23AJ_2)
```

```
## Warning in chisq.test(table_cp_MD_Q23AJ_2): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_MD_Q23AJ_2
## X-squared = 7.2969, df = 10, p-value = 0.6971
```

```
chisq.posthoc.test(table_cp_MD_Q23AJ_2)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension    Value        0         1         2         3         4
## 1       No Residuals -1.162551  1.221839  0.8607351 -1.076173 -0.1456341
## 2       No  p values  1.000000  1.000000  1.0000000  1.000000  1.0000000
## 3      Yes Residuals  1.162551 -1.221839 -0.8607351  1.076173  0.1456341
## 4      Yes  p values  1.000000  1.000000  1.0000000  1.000000  1.0000000
##           5         6        10        11        15        20
## 1  0.2875069  0.9309457 -0.1027869  0.9309457 -1.076173  0.9309457
## 2  1.0000000  1.0000000  1.0000000  1.0000000  1.000000  1.0000000
## 3 -0.2875069 -0.9309457  0.1027869 -0.9309457  1.076173 -0.9309457
## 4  1.0000000  1.0000000  1.0000000  1.0000000  1.000000  1.0000000
```

```
# cp and COMD_nicoke1m_c
table_cp_COMD_nicoke1m_c <- table(new_data$DrQ3, new_data$COMD_nicoke1m_c)
chisq.test(table_cp_COMD_nicoke1m_c)
```

```
## Warning in chisq.test(table_cp_COMD_nicoke1m_c): Chi-squared approximation may
## be incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_COMD_nicoke1m_c
## X-squared = 2.0391, df = 4, p-value = 0.7286
```

```
chisq.posthoc.test(table_cp_COMD_nicoke1m_c)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension    Value        0         1         2         3         4
## 1       No Residuals -1.162551  0.9621334 -0.1027869  0.4544801  0.9309457
## 2       No  p values  1.000000  1.0000000  1.0000000  1.0000000  1.0000000
## 3      Yes Residuals  1.162551 -0.9621334  0.1027869 -0.4544801 -0.9309457
## 4      Yes  p values  1.000000  1.0000000  1.0000000  1.0000000  1.0000000
```

```r
# cp and CO_fpoud1m
table_cp_CO_fpoud1m <- table(new_data$DrQ3, new_data$CO_fpoud1m)
chisq.test(table_cp_CO_fpoud1m)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_CO_fpoud1m
## X-squared = 0.059099, df = 1, p-value = 0.8079
```

```r
chisq.posthoc.test(table_cp_CO_fpoud1m)
```

```
##   Dimension     Value          1          2
## 1        No Residuals -0.548539   0.548539
## 2        No  p values  1.000000   1.000000
## 3       Yes Residuals  0.548539  -0.548539
## 4       Yes  p values  1.000000   1.000000
```

```r
# cp and MD_Q23BJ_1
table_cp_MD_Q23BJ_1 <- table(new_data$DrQ3, new_data$MD_Q23BJ_1)
chisq.test(table_cp_MD_Q23BJ_1)
```

```
## Warning in chisq.test(table_cp_MD_Q23BJ_1): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_MD_Q23BJ_1
## X-squared = 5.906, df = 5, p-value = 0.3155
```

```r
chisq.posthoc.test(table_cp_MD_Q23BJ_1)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension     Value          0          1          2          8         20
## 1        No Residuals  0.548539  0.6445382 -1.523355  0.9309457 -1.076173
## 2        No  p values  1.000000  1.0000000  1.000000  1.0000000  1.000000
## 3       Yes Residuals -0.548539 -0.6445382  1.523355 -0.9309457  1.076173
## 4       Yes  p values  1.000000  1.0000000  1.000000  1.0000000  1.000000
##          29
## 1 -1.076173
## 2  1.000000
## 3  1.076173
## 4  1.000000
```

```r
# cp and COMD_nfcoke1m_p
table_cp_COMD_nfcoke1m_p <- table(new_data$DrQ3, new_data$COMD_nfcoke1m_p)
chisq.test(table_cp_COMD_nfcoke1m_p)
```

```
## Warning in chisq.test(table_cp_COMD_nfcoke1m_p): Chi-squared approximation may
## be incorrect
```

```
##
##   Pearson's Chi-squared test
##
## data:  table_cp_COMD_nfcoke1m_p
## X-squared = 2.3317, df = 3, p-value = 0.5065
```

```
chisq.posthoc.test(table_cp_COMD_nfcoke1m_p)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension      Value         0          1          2          3
## 1       No Residuals  0.548539 -0.5565142  0.9309457 -1.076173
## 2       No  p values  1.000000  1.0000000  1.0000000  1.000000
## 3      Yes Residuals -0.548539  0.5565142 -0.9309457  1.076173
## 4      Yes  p values  1.000000  1.0000000  1.0000000  1.000000
```

```
# cp and COMD_fcrakfree1m
table_cp_COMD_fcrakfree1m <- table(new_data$DrQ3, new_data$COMD_fcrakfree1m)
chisq.test(table_cp_COMD_fcrakfree1m)
```

```
##
##   Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_COMD_fcrakfree1m
## X-squared = 0.09528, df = 1, p-value = 0.7576
```

```
chisq.posthoc.test(table_cp_COMD_fcrakfree1m)
```

```
##   Dimension      Value          1          2
## 1       No Residuals  0.3979808 -0.3979808
## 2       No  p values  1.0000000  1.0000000
## 3      Yes Residuals -0.3979808  0.3979808
## 4      Yes  p values  1.0000000  1.0000000
```

```
# cp and MD_Q23BJ_2
table_cp_MD_Q23BJ_2 <- table(new_data$DrQ3, new_data$MD_Q23BJ_2)
chisq.test(table_cp_MD_Q23BJ_2)
```

```
## Warning in chisq.test(table_cp_MD_Q23BJ_2): Chi-squared approximation may be
## incorrect
```

```
##
##   Pearson's Chi-squared test
##
## data:  table_cp_MD_Q23BJ_2
## X-squared = 12.521, df = 22, p-value = 0.9457
```

```
chisq.posthoc.test(table_cp_MD_Q23BJ_2)
```

Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect

```
##    Dimension     Value          0          1          2          3          4
## 1        No Residuals -0.3979808 -0.03215243 -0.1661041  0.5027058 -0.5565142
## 2        No  p values  1.0000000  1.00000000  1.0000000  1.0000000  1.0000000
## 3       Yes Residuals  0.3979808  0.03215243  0.1661041 -0.5027058  0.5565142
## 4       Yes  p values  1.0000000  1.00000000  1.0000000  1.0000000  1.0000000
##            5          6         7          8          10         11         12
## 1  0.6484175  0.4544801  1.615451  0.5075537 -0.8716531  0.9309457  0.2875069
## 2  1.0000000  1.0000000  1.000000  1.0000000  1.0000000  1.0000000  1.0000000
## 3 -0.6484175 -0.4544801 -1.615451 -0.5075537  0.8716531 -0.9309457 -0.2875069
## 4  1.0000000  1.0000000  1.000000  1.0000000  1.0000000  1.0000000  1.0000000
##           14         15         18         20         23         25         26
## 1 -1.523355  0.2143337 -0.1456341  1.048596 -0.1027869 -1.001936  0.9309457
## 2  1.000000  1.0000000  1.0000000  1.000000  1.0000000  1.000000  1.0000000
## 3  1.523355 -0.2143337  0.1456341 -1.048596  0.1027869  1.001936 -0.9309457
## 4  1.000000  1.0000000  1.0000000  1.000000  1.0000000  1.000000  1.0000000
##           27         28         29         30
## 1 -0.1027869  0.4544801 -1.076173 -0.1887955
## 2  1.0000000  1.0000000  1.000000  1.0000000
## 3  0.1027869 -0.4544801  1.076173  0.1887955
## 4  1.0000000  1.0000000  1.000000  1.0000000
```

```r
# cp and COMD_nfcoke1m_cf
table_cp_COMD_nfcoke1m_cf <- table(new_data$DrQ3, new_data$COMD_nfcoke1m_cf)
chisq.test(table_cp_COMD_nfcoke1m_cf)
```

Warning in chisq.test(table_cp_COMD_nfcoke1m_cf): Chi-squared approximation may
be incorrect

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_COMD_nfcoke1m_cf
## X-squared = 7.9336, df = 6, p-value = 0.243
```

```r
chisq.posthoc.test(table_cp_COMD_nfcoke1m_cf)
```

Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect

```
##    Dimension     Value          0         1          2          3          4
## 1        No Residuals -0.3979808 -0.598916  0.5886853  0.8926604  1.730161
## 2        No  p values  1.0000000  1.000000  1.0000000  1.0000000  1.000000
## 3       Yes Residuals  0.3979808  0.598916 -0.5886853 -0.8926604 -1.730161
## 4       Yes  p values  1.0000000  1.000000  1.0000000  1.0000000  1.000000
##            5  88888888
## 1 -1.597747 -1.076173
## 2  1.000000  1.000000
## 3  1.597747  1.076173
## 4  1.000000  1.000000
```

```
# cp and CO_scoke1m
table_cp_CO_scoke1m <- table(new_data$DrQ3, new_data$CO_scoke1m)
chisq.test(table_cp_CO_scoke1m)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_CO_scoke1m
## X-squared = 0.0539, df = 1, p-value = 0.8164
```

```
chisq.posthoc.test(table_cp_CO_scoke1m)
```

```
##   Dimension     Value          1          2
## 1        No Residuals -0.3823873  0.3823873
## 2        No  p values  1.0000000  1.0000000
## 3       Yes Residuals  0.3823873 -0.3823873
## 4       Yes  p values  1.0000000  1.0000000
```

```
# cp and Q23CJ
table_cp_Q23CJ <- table(new_data$DrQ3, new_data$Q23CJ)
chisq.test(table_cp_Q23CJ)
```

```
## Warning in chisq.test(table_cp_Q23CJ): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_Q23CJ
## X-squared = 14.736, df = 13, p-value = 0.3242
```

```
chisq.posthoc.test(table_cp_Q23CJ)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension     Value          0           1          2          3          4
## 1        No Residuals  0.3823873 -0.02240743  0.6445382 -1.001936  0.8607351
## 2        No  p values  1.0000000  1.00000000  1.0000000  1.000000  1.0000000
## 3       Yes Residuals -0.3823873  0.02240743 -0.6445382  1.001936 -0.8607351
## 4       Yes  p values  1.0000000  1.00000000  1.0000000  1.000000  1.0000000
##           5         6         8        10        14         16        20
## 1 -1.514426 -1.076173 -1.523355  1.317781  1.317781  0.9309457 -1.076173
## 2  1.000000  1.000000  1.000000  1.000000  1.000000  1.0000000  1.000000
## 3  1.514426  1.076173  1.523355 -1.317781 -1.317781 -0.9309457  1.076173
## 4  1.000000  1.000000  1.000000  1.000000  1.000000  1.0000000  1.000000
##          28        30
## 1 -1.076173  0.4544801
## 2  1.000000  1.0000000
## 3  1.076173 -0.4544801
## 4  1.000000  1.0000000
```

```
# cp and COMD_nscoke1m
table_cp_COMD_nscoke1m <- table(new_data$DrQ3, new_data$COMD_nscoke1m)
chisq.test(table_cp_COMD_nscoke1m)
```

Warning in chisq.test(table_cp_COMD_nscoke1m): Chi-squared approximation may be
incorrect

##
Pearson's Chi-squared test
##
data: table_cp_COMD_nscoke1m
X-squared = 1.4182, df = 5, p-value = 0.9223

```
chisq.posthoc.test(table_cp_COMD_nscoke1m)
```

Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect

Dimension   Value   0   1   2   3   4
1   No Residuals   0.3823873 -0.3791736 0.2679868 -0.178699 -0.1027869
2   No p values   1.0000000 1.0000000 1.0000000 1.000000 1.0000000
3   Yes Residuals -0.3823873 0.3791736 -0.2679868 0.178699 0.1027869
4   Yes p values   1.0000000 1.0000000 1.0000000 1.000000 1.0000000
5
1 -1.076173
2 1.000000
3 1.076173
4 1.000000

```
# cp and CO_ispeed1m
table_cp_CO_ispeed1m <- table(new_data$DrQ3, new_data$CO_ispeed1m)
chisq.test(table_cp_CO_ispeed1m)
```

##
Pearson's Chi-squared test with Yates' continuity correction
##
data: table_cp_CO_ispeed1m
X-squared = 8.4321e-05, df = 1, p-value = 0.9927

```
chisq.posthoc.test(table_cp_CO_ispeed1m)
```

Dimension   Value   1   2
1   No Residuals   0.2071609 -0.2071609
2   No p values   1.0000000 1.0000000
3   Yes Residuals -0.2071609 0.2071609
4   Yes p values   1.0000000 1.0000000

```
# cp and Q2223AJ
table_cp_Q2223AJ <- table(new_data$DrQ3, new_data$Q2223AJ)
chisq.test(table_cp_Q2223AJ)
```

```
## Warning in chisq.test(table_cp_Q2223AJ): Chi-squared approximation may be
## incorrect


##
##  Pearson's Chi-squared test
##
## data:  table_cp_Q2223AJ
## X-squared = 11.436, df = 10, p-value = 0.3246

chisq.posthoc.test(table_cp_Q2223AJ)


## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect


##   Dimension     Value         0         1         2         3         4
## 1        No Residuals -0.2071609  0.6445382  0.6445382  1.317781  0.9309457
## 2        No  p values  1.0000000  1.0000000  1.0000000  1.000000  1.0000000
## 3       Yes Residuals  0.2071609 -0.6445382 -0.6445382 -1.317781 -0.9309457
## 4       Yes  p values  1.0000000  1.0000000  1.0000000  1.000000  1.0000000
##           5        10        20        25        29        30
## 1  0.9309457  0.4544801 -0.1027869 -1.076173 -1.523355 -1.867461
## 2  1.0000000  1.0000000  1.0000000  1.000000  1.000000  1.000000
## 3 -0.9309457 -0.4544801  0.1027869  1.076173  1.523355  1.867461
## 4  1.0000000  1.0000000  1.0000000  1.000000  1.000000  1.000000
```

```r
# cp and COMD_nispeed1m
table_cp_COMD_nispeed1m <- table(new_data$DrQ3, new_data$COMD_nispeed1m)
chisq.test(table_cp_COMD_nispeed1m)
```

```
## Warning in chisq.test(table_cp_COMD_nispeed1m): Chi-squared approximation may be
## incorrect


##
##  Pearson's Chi-squared test
##
## data:  table_cp_COMD_nispeed1m
## X-squared = 0.069723, df = 2, p-value = 0.9657
```

```r
chisq.posthoc.test(table_cp_COMD_nispeed1m)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect


##   Dimension     Value         0         1         2
## 1        No Residuals -0.2071609  0.2445845 -0.1027869
## 2        No  p values  1.0000000  1.0000000  1.0000000
## 3       Yes Residuals  0.2071609 -0.2445845  0.1027869
## 4       Yes  p values  1.0000000  1.0000000  1.0000000
```

```r
# cp and NH_Q7Ia_suboxIV
table_cp_NH_Q7Ia_suboxIV <- table(new_data$DrQ3, new_data$NH_Q7Ia_suboxIV)
chisq.test(table_cp_NH_Q7Ia_suboxIV)
```

```
## Warning in chisq.test(table_cp_NH_Q7Ia_suboxIV): Chi-squared approximation may
## be incorrect


##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_NH_Q7Ia_suboxIV
## X-squared = 2.0861e-27, df = 1, p-value = 1
```

```
chisq.posthoc.test(table_cp_NH_Q7Ia_suboxIV)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect


##    Dimension     Value          1          2
## 1        No Residuals  0.9309457 -0.9309457
## 2        No  p values  1.0000000  1.0000000
## 3       Yes Residuals -0.9309457  0.9309457
## 4       Yes  p values  1.0000000  1.0000000
```

```
# cp and NH_Q7Ia_suboxIVj
table_cp_NH_Q7Ia_suboxIVj <- table(new_data$DrQ3, new_data$NH_Q7Ia_suboxIVj)
chisq.test(table_cp_NH_Q7Ia_suboxIVj)
```

```
## Warning in chisq.test(table_cp_NH_Q7Ia_suboxIVj): Chi-squared approximation may
## be incorrect


##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_NH_Q7Ia_suboxIVj
## X-squared = 2.0861e-27, df = 1, p-value = 1
```

```
chisq.posthoc.test(table_cp_NH_Q7Ia_suboxIVj)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect


##    Dimension     Value          1          2
## 1        No Residuals  0.9309457 -0.9309457
## 2        No  p values  1.0000000  1.0000000
## 3       Yes Residuals -0.9309457  0.9309457
## 4       Yes  p values  1.0000000  1.0000000
```

```
# cp and NH_Q7Ib_subox
table_cp_NH_Q7Ib_subox <- table(new_data$DrQ3, new_data$NH_Q7Ib_subox)
chisq.test(table_cp_NH_Q7Ib_subox)
```

```
## Warning in chisq.test(table_cp_NH_Q7Ib_subox): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_NH_Q7Ib_subox
## X-squared = 0.36905, df = 1, p-value = 0.5435
```

```
chisq.posthoc.test(table_cp_NH_Q7Ib_subox)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension      Value         1         2
## 1       No Residuals  1.317781 -1.317781
## 2       No  p values  0.750308  0.750308
## 3      Yes Residuals -1.317781  1.317781
## 4      Yes  p values  0.750308  0.750308
```

```
# cp and NH_Q7Ib_suboxj
table_cp_NH_Q7Ib_suboxj <- table(new_data$DrQ3, new_data$NH_Q7Ib_suboxj)
chisq.test(table_cp_NH_Q7Ib_suboxj)
```

```
## Warning in chisq.test(table_cp_NH_Q7Ib_suboxj): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_NH_Q7Ib_suboxj
## X-squared = 1.7365, df = 2, p-value = 0.4197
```

```
chisq.posthoc.test(table_cp_NH_Q7Ib_suboxj)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension      Value         0          1          5
## 1       No Residuals -1.317781  0.9309457  0.9309457
## 2       No  p values  1.000000  1.0000000  1.0000000
## 3      Yes Residuals  1.317781 -0.9309457 -0.9309457
## 4      Yes  p values  1.000000  1.0000000  1.0000000
```

```
# cp and NH_Q7Ja_methaIV
table_cp_NH_Q7Ja_methaIV <- table(new_data$DrQ3, new_data$NH_Q7Ja_methaIV)
chisq.test(table_cp_NH_Q7Ja_methaIV)
```

```
## Warning in chisq.test(table_cp_NH_Q7Ja_methaIV): Chi-squared approximation may
## be incorrect
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_NH_Q7Ja_methaIV
## X-squared = 0.66108, df = 1, p-value = 0.4162
```

```
chisq.posthoc.test(table_cp_NH_Q7Ja_methaIV)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension     Value         1         2
## 1       No Residuals -1.523355  1.523355
## 2       No  p values  0.510680  0.510680
## 3      Yes Residuals  1.523355 -1.523355
## 4      Yes  p values  0.510680  0.510680
```

```
# cp and NH_Q7Ja_methaIVj
table_cp_NH_Q7Ja_methaIVj <- table(new_data$DrQ3, new_data$NH_Q7Ja_methaIVj)
chisq.test(table_cp_NH_Q7Ja_methaIVj)
```

```
## Warning in chisq.test(table_cp_NH_Q7Ja_methaIVj): Chi-squared approximation may
## be incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_NH_Q7Ja_methaIVj
## X-squared = 2.3206, df = 2, p-value = 0.3134
```

```
chisq.posthoc.test(table_cp_NH_Q7Ja_methaIVj)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension     Value         0         1         2
## 1       No Residuals  1.523355 -1.076173 -1.076173
## 2       No  p values  0.766019  1.000000  1.000000
## 3      Yes Residuals -1.523355  1.076173  1.076173
## 4      Yes  p values  0.766019  1.000000  1.000000
```

```
# cp and NH_Q7Jb_metha
table_cp_NH_Q7Jb_metha <- table(new_data$DrQ3, new_data$NH_Q7Jb_metha)
chisq.test(table_cp_NH_Q7Jb_metha)
```

```
## Warning in chisq.test(table_cp_NH_Q7Jb_metha): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_NH_Q7Jb_metha
## X-squared = 0.054252, df = 1, p-value = 0.8158
```

```
chisq.posthoc.test(table_cp_NH_Q7Jb_metha)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension      Value           1          2
## 1        No Residuals  0.6445382 -0.6445382
## 2        No  p values  1.0000000  1.0000000
## 3       Yes Residuals -0.6445382  0.6445382
## 4       Yes  p values  1.0000000  1.0000000
```

```
# cp and NH_Q7Jb_methaj
table_cp_NH_Q7Jb_methaj <- table(new_data$DrQ3, new_data$NH_Q7Jb_methaj)
chisq.test(table_cp_NH_Q7Jb_methaj)
```

```
## Warning in chisq.test(table_cp_NH_Q7Jb_methaj): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_NH_Q7Jb_methaj
## X-squared = 3.0961, df = 4, p-value = 0.5419
```

```
chisq.posthoc.test(table_cp_NH_Q7Jb_methaj)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension      Value           0          1          2          3          7
## 1        No Residuals -0.6445382  0.9309457  0.4544801  0.9309457 -1.076173
## 2        No  p values  1.0000000  1.0000000  1.0000000  1.0000000  1.000000
## 3       Yes Residuals  0.6445382 -0.9309457 -0.4544801 -0.9309457  1.076173
## 4       Yes  p values  1.0000000  1.0000000  1.0000000  1.0000000  1.000000
```

```
# cp and CO_iopiac1m
table_cp_CO_iopiac1m <- table(new_data$DrQ3, new_data$CO_iopiac1m)
chisq.test(table_cp_CO_iopiac1m)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_CO_iopiac1m
## X-squared = 0.14752, df = 1, p-value = 0.7009
```

```
chisq.posthoc.test(table_cp_CO_iopiac1m)
```

```
##   Dimension      Value           1          2
## 1        No Residuals  0.4800518 -0.4800518
## 2        No  p values  1.0000000  1.0000000
## 3       Yes Residuals -0.4800518  0.4800518
## 4       Yes  p values  1.0000000  1.0000000
```

```
# cp and Q26BJ
table_cp_Q26BJ <- table(new_data$DrQ3, new_data$Q26BJ)
chisq.test(table_cp_Q26BJ)
```

```
## Warning in chisq.test(table_cp_Q26BJ): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_Q26BJ
## X-squared = 29.088, df = 25, p-value = 0.2602
```

```
chisq.posthoc.test(table_cp_Q26BJ)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##     Dimension     Value         0          1          2          3          4
## 1          No Residuals -0.4800518  0.3312679  0.3312679 -1.514426  0.1175558
## 2          No  p values  1.0000000  1.0000000  1.0000000  1.000000  1.0000000
## 3         Yes Residuals  0.4800518 -0.3312679 -0.3312679  1.514426 -0.1175558
## 4         Yes  p values  1.0000000  1.0000000  1.0000000  1.000000  1.0000000
##            5          6          7          8          9         10         12
## 1 -0.2541531  1.867104 -0.1456341 -0.6134595 -1.076173 -0.206732  1.317781
## 2  1.0000000  1.000000  1.0000000  1.0000000  1.000000  1.000000  1.000000
## 3  0.2541531 -1.867104  0.1456341  0.6134595  1.076173  0.206732 -1.317781
## 4  1.0000000  1.000000  1.0000000  1.0000000  1.000000  1.000000  1.000000
##           14         15         18         19         20         21         22
## 1 -1.076173  2.476901  0.9309457 -1.076173 -0.6134595  0.9309457 -1.076173
## 2  1.000000  0.689149  1.0000000  1.000000  1.0000000  1.0000000  1.000000
## 3  1.076173 -2.476901 -0.9309457  1.076173  0.6134595 -0.9309457  1.076173
## 4  1.000000  0.689149  1.0000000  1.000000  1.0000000  1.0000000  1.000000
##           23         25         26         27         28         29         30
## 1 -1.076173  0.4544801  1.317781  1.317781  0.9309457 -1.523355 -0.2408366
## 2  1.000000  1.0000000  1.000000  1.000000  1.0000000  1.000000  1.0000000
## 3  1.076173 -0.4544801 -1.317781 -1.317781 -0.9309457  1.523355  0.2408366
## 4  1.000000  1.0000000  1.000000  1.000000  1.0000000  1.000000  1.0000000
```

```r
# cp and CO_niopiac1m
table_cp_CO_niopiac1m <- table(new_data$DrQ3, new_data$CO_niopiac1m)
chisq.test(table_cp_CO_niopiac1m)
```

```
## Warning in chisq.test(table_cp_CO_niopiac1m): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_CO_niopiac1m
## X-squared = 4.6893, df = 4, p-value = 0.3207
```

```
chisq.posthoc.test(table_cp_CO_niopiac1m)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension       Value           0          1           2          3          4
## 1        No Residuals -0.4800518  1.445432 -0.8118436  0.1882445 -1.514426
## 2        No  p values  1.0000000  1.000000  1.0000000  1.0000000  1.000000
## 3       Yes Residuals  0.4800518 -1.445432  0.8118436 -0.1882445  1.514426
## 4       Yes  p values  1.0000000  1.000000  1.0000000  1.0000000  1.000000
```

```
# cp and CO_autopiac1m
table_cp_CO_autopiac1m <- table(new_data$DrQ3, new_data$CO_autopiac1m)
chisq.test(table_cp_CO_autopiac1m)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_CO_autopiac1m
## X-squared = 0.42341, df = 1, p-value = 0.5152
```

```
chisq.posthoc.test(table_cp_CO_autopiac1m)
```

```
##    Dimension       Value           1          2
## 1        No Residuals  0.8308342 -0.8308342
## 2        No  p values  1.0000000  1.0000000
## 3       Yes Residuals -0.8308342  0.8308342
## 4       Yes  p values  1.0000000  1.0000000
```

```
# cp and Q26AJ
table_cp_Q26AJ <- table(new_data$DrQ3, new_data$Q26AJ)
chisq.test(table_cp_Q26AJ)
```

```
## Warning in chisq.test(table_cp_Q26AJ): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_Q26AJ
## X-squared = 12.41, df = 12, p-value = 0.4134
```

```
chisq.posthoc.test(table_cp_Q26AJ)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension       Value           0          1           2          3          4
## 1        No Residuals -0.8308342  0.8110127  0.8607351  1.867104  0.9309457
## 2        No  p values  1.0000000  1.0000000  1.0000000  1.000000  1.0000000
## 3       Yes Residuals  0.8308342 -0.8110127 -0.8607351 -1.867104 -0.9309457
## 4       Yes  p values  1.0000000  1.0000000  1.0000000  1.000000  1.0000000
##              6          7         10         12         15         20         23
## 1 -1.076173 -0.1027869 -1.076173 -1.076173 -0.1027869 -1.076173 -1.076173
## 2  1.000000  1.0000000  1.000000  1.000000  1.0000000  1.000000  1.000000
## 3  1.076173  0.1027869  1.076173  1.076173  0.1027869  1.076173  1.076173
```

```
## 4  1.000000  1.0000000  1.000000  1.000000  1.0000000  1.000000  1.000000
##            24
## 1  0.9309457
## 2  1.0000000
## 3 -0.9309457
## 4  1.0000000
```

```r
# cp and NH_Q26AJnf
table_cp_NH_Q26AJnf <- table(new_data$DrQ3, new_data$NH_Q26AJnf)
chisq.test(table_cp_NH_Q26AJnf)
```

```
## Warning in chisq.test(table_cp_NH_Q26AJnf): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_NH_Q26AJnf
## X-squared = 3.0493, df = 3, p-value = 0.3841
```

```r
chisq.posthoc.test(table_cp_NH_Q26AJnf)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension     Value        0          1          2          3
## 1         No Residuals  1.361091 -0.8395741 -1.076173 -1.076173
## 2         No  p values  1.000000  1.0000000  1.000000  1.000000
## 3        Yes Residuals -1.361091  0.8395741  1.076173  1.076173
## 4        Yes  p values  1.000000  1.0000000  1.000000  1.000000
```

```r
# cp and NMD_CO_iamph1m
table_cp_NMD_CO_iamph1m <- table(new_data$DrQ3, new_data$NMD_CO_iamph1m)
chisq.test(table_cp_NMD_CO_iamph1m)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_NMD_CO_iamph1m
## X-squared = 0.00030473, df = 1, p-value = 0.9861
```

```r
chisq.posthoc.test(table_cp_NMD_CO_iamph1m)
```

```
##    Dimension     Value         1          2
## 1         No Residuals -0.1578909  0.1578909
## 2         No  p values  1.0000000  1.0000000
## 3        Yes Residuals  0.1578909 -0.1578909
## 4        Yes  p values  1.0000000  1.0000000
```

```
# cp and NMD_Q27BJ
table_cp_NMD_Q27BJ <- table(new_data$DrQ3, new_data$NMD_Q27BJ)
chisq.test(table_cp_NMD_Q27BJ)
```

```
## Warning in chisq.test(table_cp_NMD_Q27BJ): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_NMD_Q27BJ
## X-squared = 15.176, df = 17, p-value = 0.5828
```

```
chisq.posthoc.test(table_cp_NMD_Q27BJ)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension     Value         0          1        2         3         4
## 1         No Residuals  0.1578909  0.01671209  0.95113  0.2875069 -1.152003
## 2         No  p values  1.0000000  1.00000000  1.00000  1.0000000  1.000000
## 3        Yes Residuals -0.1578909 -0.01671209 -0.95113 -0.2875069  1.152003
## 4        Yes  p values  1.0000000  1.00000000  1.00000  1.0000000  1.000000
##           5         6         8         10        14         15        16
## 1 -1.523355  1.317781  0.4544801 -0.7064904 -1.076173 -0.1027869  0.9309457
## 2  1.000000  1.000000  1.0000000  1.0000000  1.000000  1.0000000  1.0000000
## 3  1.523355 -1.317781 -0.4544801  0.7064904  1.076173  0.1027869 -0.9309457
## 4  1.000000  1.000000  1.0000000  1.0000000  1.000000  1.0000000  1.0000000
##           18        20        24        25        26        30
## 1  0.9309457 -1.523355  0.9309457 -1.076173  0.9309457 -0.206732
## 2  1.0000000  1.000000  1.0000000  1.000000  1.0000000  1.000000
## 3 -0.9309457  1.523355 -0.9309457  1.076173 -0.9309457  0.206732
## 4  1.0000000  1.000000  1.0000000  1.000000  1.0000000  1.000000
```

```
# cp and NMD_CO_niamph1m
table_cp_NMD_CO_niamph1m <- table(new_data$DrQ3, new_data$NMD_CO_niamph1m)
chisq.test(table_cp_NMD_CO_niamph1m)
```

```
## Warning in chisq.test(table_cp_NMD_CO_niamph1m): Chi-squared approximation may
## be incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_NMD_CO_niamph1m
## X-squared = 0.98635, df = 4, p-value = 0.9119
```

```
chisq.posthoc.test(table_cp_NMD_CO_niamph1m)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension    Value         0          1          2          3          4
## 1        No Residuals  0.1578909  0.1548362 -0.8716531 -0.1027869  0.4544801
## 2        No  p values  1.0000000  1.0000000  1.0000000  1.0000000  1.0000000
## 3       Yes Residuals -0.1578909 -0.1548362  0.8716531  0.1027869 -0.4544801
## 4       Yes  p values  1.0000000  1.0000000  1.0000000  1.0000000  1.0000000
```

```r
# cp and NMD_COMD_fsautamph1m
table_cp_NMD_COMD_fsautamph1m <- table(new_data$DrQ3, new_data$NMD_COMD_fsautamph1m)
chisq.test(table_cp_NMD_COMD_fsautamph1m)
```

```
## 
##  Pearson's Chi-squared test with Yates' continuity correction
## 
## data:  table_cp_NMD_COMD_fsautamph1m
## X-squared = 0.059759, df = 1, p-value = 0.8069
```

```r
chisq.posthoc.test(table_cp_NMD_COMD_fsautamph1m)
```

```
##   Dimension    Value        1         2
## 1        No Residuals  0.352725 -0.352725
## 2        No  p values  1.000000  1.000000
## 3       Yes Residuals -0.352725  0.352725
## 4       Yes  p values  1.000000  1.000000
```

```r
# cp and NMD_Q27AJ
table_cp_NMD_Q27AJ <- table(new_data$DrQ3, new_data$NMD_Q27AJ)
chisq.test(table_cp_NMD_Q27AJ)
```

```
## Warning in chisq.test(table_cp_NMD_Q27AJ): Chi-squared approximation may be
## incorrect
```

```
## 
##  Pearson's Chi-squared test
## 
## data:  table_cp_NMD_Q27AJ
## X-squared = 14.887, df = 19, p-value = 0.7298
```

```r
chisq.posthoc.test(table_cp_NMD_Q27AJ)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension    Value         0          1         2         3         4
## 1        No Residuals -0.352725 -0.3306293  1.142689  1.425759  1.317781
## 2        No  p values  1.000000  1.0000000  1.000000  1.000000  1.000000
## 3       Yes Residuals  0.352725  0.3306293 -1.142689 -1.425759 -1.317781
## 4       Yes  p values  1.000000  1.0000000  1.000000  1.000000  1.000000
##            5          6         7          8         9         10        14
## 1 -0.178699 -0.6134595 -1.523355 -0.7064904 -1.076173  0.4085132 -1.076173
## 2  1.000000  1.0000000  1.000000  1.0000000  1.000000  1.0000000  1.000000
## 3  0.178699  0.6134595  1.523355  0.7064904  1.076173 -0.4085132  1.076173
```

```
## 4  1.000000  1.0000000  1.000000  1.0000000  1.000000  1.0000000  1.000000
##           15         16         20         21         24         25         28
## 1 -0.178699 -1.076173 -0.1027869  0.9309457  0.9309457 -0.1027869 -1.076173
## 2  1.000000  1.000000  1.0000000  1.0000000  1.0000000  1.0000000  1.000000
## 3  0.178699  1.076173  0.1027869 -0.9309457 -0.9309457  0.1027869  1.076173
## 4  1.000000  1.000000  1.0000000  1.0000000  1.0000000  1.0000000  1.000000
##           30
## 1  0.01671209
## 2  1.00000000
## 3 -0.01671209
## 4  1.00000000
```

```
# cp and NMD_COMD_nfsaamph1m
table_cp_NMD_COMD_nfsaamph1m <- table(new_data$DrQ3, new_data$NMD_COMD_nfsaamph1m)
chisq.test(table_cp_NMD_COMD_nfsaamph1m)
```

```
## Warning in chisq.test(table_cp_NMD_COMD_nfsaamph1m): Chi-squared approximation
## may be incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_NMD_COMD_nfsaamph1m
## X-squared = 4.349, df = 4, p-value = 0.3608
```

```
chisq.posthoc.test(table_cp_NMD_COMD_nfsaamph1m)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##     Dimension     Value         0          1         2         3          4
## 1         No Residuals -0.352725 -0.6927984  1.954104 -0.178699  0.4544801
## 2         No  p values  1.000000  1.0000000  0.506889  1.000000  1.0000000
## 3        Yes Residuals  0.352725  0.6927984 -1.954104  0.178699 -0.4544801
## 4        Yes  p values  1.000000  1.0000000  0.506889  1.000000  1.0000000
```

```
# cp and NW_tranquiliv1m
table_cp_NW_tranquiliv1m <- table(new_data$DrQ3, new_data$NW_tranquiliv1m)
chisq.test(table_cp_NW_tranquiliv1m)
```

```
## Warning in chisq.test(table_cp_NW_tranquiliv1m): Chi-squared approximation may
## be incorrect
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_NW_tranquiliv1m
## X-squared = 0.026561, df = 1, p-value = 0.8705
```

```
chisq.posthoc.test(table_cp_NW_tranquiliv1m)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension     Value              1          2
## 1        No Residuals  -0.6134595  0.6134595
## 2        No  p values   1.0000000  1.0000000
## 3       Yes Residuals   0.6134595 -0.6134595
## 4       Yes  p values   1.0000000  1.0000000
```

```
# cp and NW_Q29BJ
table_cp_NW_Q29BJ <- table(new_data$DrQ3, new_data$NW_Q29BJ)
chisq.test(table_cp_NW_Q29BJ)
```

```
## Warning in chisq.test(table_cp_NW_Q29BJ): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_NW_Q29BJ
## X-squared = 2.5209, df = 3, p-value = 0.4715
```

```
chisq.posthoc.test(table_cp_NW_Q29BJ)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension     Value              0          1          3          7
## 1        No Residuals   0.6134595 -0.7064904 -1.076173  0.9309457
## 2        No  p values   1.0000000  1.0000000  1.000000  1.0000000
## 3       Yes Residuals  -0.6134595  0.7064904  1.076173 -0.9309457
## 4       Yes  p values   1.0000000  1.0000000  1.000000  1.0000000
```

```
# cp and NW_tranquiliv1m_fj
table_cp_NW_tranquiliv1m_fj <- table(new_data$DrQ3, new_data$NW_tranquiliv1m_fj)
chisq.test(table_cp_NW_tranquiliv1m_fj)
```

```
## Warning in chisq.test(table_cp_NW_tranquiliv1m_fj): Chi-squared approximation
## may be incorrect
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_NW_tranquiliv1m_fj
## X-squared = 0.026561, df = 1, p-value = 0.8705
```

```
chisq.posthoc.test(table_cp_NW_tranquiliv1m_fj)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension     Value              0          1
## 1        No Residuals   0.6134595 -0.6134595
## 2        No  p values   1.0000000  1.0000000
## 3       Yes Residuals  -0.6134595  0.6134595
## 4       Yes  p values   1.0000000  1.0000000
```

```r
# cp and NW_tranquil1m
table_cp_NW_tranquil1m <- table(new_data$DrQ3, new_data$NW_tranquil1m)
chisq.test(table_cp_NW_tranquil1m)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_NW_tranquil1m
## X-squared = 0.33547, df = 1, p-value = 0.5625
```

```r
chisq.posthoc.test(table_cp_NW_tranquil1m)
```

```
##   Dimension      Value          1          2
## 1        No Residuals -0.7196282  0.7196282
## 2        No  p values  1.0000000  1.0000000
## 3       Yes Residuals  0.7196282 -0.7196282
## 4       Yes  p values  1.0000000  1.0000000
```

```r
# cp and Q29AJ
table_cp_Q29AJ <- table(new_data$DrQ3, new_data$Q29AJ)
chisq.test(table_cp_Q29AJ)
```

```
## Warning in chisq.test(table_cp_Q29AJ): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_Q29AJ
## X-squared = 11.775, df = 15, p-value = 0.696
```

```r
chisq.posthoc.test(table_cp_Q29AJ)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension      Value          0          1          2          3          4
## 1        No Residuals  0.7196282  0.1175558 -0.206732 -1.337527 -0.5746411
## 2        No  p values  1.0000000  1.0000000  1.000000  1.000000  1.0000000
## 3       Yes Residuals -0.7196282 -0.1175558  0.206732  1.337527  0.5746411
## 4       Yes  p values  1.0000000  1.0000000  1.000000  1.000000  1.0000000
##           5          6          7          9         10         12         15
## 1  0.2875069 -0.1027869  0.9309457  0.9309457 -1.867461 -1.076173 -0.7064904
## 2  1.0000000  1.0000000  1.0000000  1.0000000  1.000000  1.000000  1.0000000
## 3 -0.2875069  0.1027869 -0.9309457 -0.9309457  1.867461  1.076173  0.7064904
## 4  1.0000000  1.0000000  1.0000000  1.0000000  1.000000  1.000000  1.0000000
##          20         27         28         30
## 1  0.9309457  0.9309457  0.9309457  0.1882445
## 2  1.0000000  1.0000000  1.0000000  1.0000000
## 3 -0.9309457 -0.9309457 -0.9309457 -0.1882445
## 4  1.0000000  1.0000000  1.0000000  1.0000000
```

```
# cp and NW_tranquil1m_fj
table_cp_NW_tranquil1m_fj <- table(new_data$DrQ3, new_data$NW_tranquil1m_fj)
chisq.test(table_cp_NW_tranquil1m_fj)
```

```
## Warning in chisq.test(table_cp_NW_tranquil1m_fj): Chi-squared approximation may
## be incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_NW_tranquil1m_fj
## X-squared = 3.5847, df = 3, p-value = 0.3099
```

```
chisq.posthoc.test(table_cp_NW_tranquil1m_fj)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##   Dimension      Value         0          1        2          3
## 1       No Residuals  0.7196282 -1.441911  0.95113  0.8607351
## 2       No  p values  1.0000000  1.000000  1.00000  1.0000000
## 3      Yes Residuals -0.7196282  1.441911 -0.95113 -0.8607351
## 4      Yes  p values  1.0000000  1.000000  1.00000  1.0000000
```

```
# cp and CO_mari1m
table_cp_CO_mari1m <- table(new_data$DrQ3, new_data$CO_mari1m)
chisq.test(table_cp_CO_mari1m)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_CO_mari1m
## X-squared = 0.049527, df = 1, p-value = 0.8239
```

```
chisq.posthoc.test(table_cp_CO_mari1m)
```

```
##   Dimension      Value          1          2
## 1       No Residuals -0.3094069  0.3094069
## 2       No  p values  1.0000000  1.0000000
## 3      Yes Residuals  0.3094069 -0.3094069
## 4      Yes  p values  1.0000000  1.0000000
```

```
# cp and Q30AJ
table_cp_Q30AJ <- table(new_data$DrQ3, new_data$Q30AJ)
chisq.test(table_cp_Q30AJ)
```

```
## Warning in chisq.test(table_cp_Q30AJ): Chi-squared approximation may be
## incorrect
```

```
## 
##  Pearson's Chi-squared test
## 
## data:  table_cp_Q30AJ
## X-squared = 18.608, df = 22, p-value = 0.6694
```

```
chisq.posthoc.test(table_cp_Q30AJ)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension     Value          0         1         2         3         4
## 1         No Residuals  0.3094069  1.703812  0.1676788 -0.818065  0.673207
## 2         No  p values  1.0000000  1.000000  1.0000000  1.000000  1.000000
## 3        Yes Residuals -0.3094069 -1.703812 -0.1676788  0.818065 -0.673207
## 4        Yes  p values  1.0000000  1.000000  1.0000000  1.000000  1.000000
##           5          6          7         8          9         10         12
## 1 -0.548539 -0.5746411 -2.100412  0.8607351 -0.1027869 -0.7823403  0.2875069
## 2  1.000000  1.0000000  1.000000  1.0000000  1.0000000  1.0000000  1.0000000
## 3  0.548539  0.5746411  2.100412 -0.8607351  0.1027869  0.7823403 -0.2875069
## 4  1.000000  1.0000000  1.000000  1.0000000  1.0000000  1.0000000  1.0000000
##           15         16         18         20         23         25         26
## 1 -0.8035154 -1.076173 -0.1027869  0.6484175 -1.076173 -0.178699 -1.076173
## 2  1.0000000  1.000000  1.0000000  1.0000000  1.000000  1.000000  1.000000
## 3  0.8035154  1.076173  0.1027869 -0.6484175  1.076173  0.178699  1.076173
## 4  1.0000000  1.000000  1.0000000  1.0000000  1.000000  1.000000  1.000000
##           27         28         29         30
## 1 -0.1027869 -1.825173 -0.1027869  0.7441966
## 2  1.0000000  1.000000  1.0000000  1.0000000
## 3  0.1027869  1.825173  0.1027869 -0.7441966
## 4  1.0000000  1.000000  1.0000000  1.0000000
```

```r
# cp and NW_mari1m_fj
table_cp_NW_mari1m_fj <- table(new_data$DrQ3, new_data$NW_mari1m_fj)
chisq.test(table_cp_NW_mari1m_fj)
```

```
## Warning in chisq.test(table_cp_NW_mari1m_fj): Chi-squared approximation may be
## incorrect
```

```
## 
##  Pearson's Chi-squared test
## 
## data:  table_cp_NW_mari1m_fj
## X-squared = 0.51938, df = 4, p-value = 0.9716
```

```
chisq.posthoc.test(table_cp_NW_mari1m_fj)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension     Value          0         1          2          3   88888888
## 1         No Residuals  0.3094069  0.1927694 -0.3214855 -0.6015534 -0.1027869
## 2         No  p values  1.0000000  1.0000000  1.0000000  1.0000000  1.0000000
## 3        Yes Residuals -0.3094069 -0.1927694  0.3214855  0.6015534  0.1027869
## 4        Yes  p values  1.0000000  1.0000000  1.0000000  1.0000000  1.0000000
```

```r
# cp and NMD_NW_autmediv1m
table_cp_NMD_NW_autmediv1m <- table(new_data$DrQ3, new_data$NMD_NW_autmediv1m)
chisq.test(table_cp_NMD_NW_autmediv1m)
```

```
## Warning in chisq.test(table_cp_NMD_NW_autmediv1m): Chi-squared approximation may
## be incorrect
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_NMD_NW_autmediv1m
## X-squared = 2.5783e-28, df = 1, p-value = 1
```

```r
chisq.posthoc.test(table_cp_NMD_NW_autmediv1m)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension      Value           1          2
## 1         No Residuals  -0.1027869  0.1027869
## 2         No  p values   1.0000000  1.0000000
## 3        Yes Residuals   0.1027869 -0.1027869
## 4        Yes  p values   1.0000000  1.0000000
```

```r
# cp and NMD_NW_Q32BJ
table_cp_NMD_NW_Q32BJ <- table(new_data$DrQ3, new_data$NMD_NW_Q32BJ)
chisq.test(table_cp_NMD_NW_Q32BJ)
```

```
## Warning in chisq.test(table_cp_NMD_NW_Q32BJ): Chi-squared approximation may be
## incorrect
```

```
##
##  Pearson's Chi-squared test
##
## data:  table_cp_NMD_NW_Q32BJ
## X-squared = 2.0211, df = 2, p-value = 0.364
```

```r
chisq.posthoc.test(table_cp_NMD_NW_Q32BJ)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect
```

```
##    Dimension      Value          0          1          4
## 1         No Residuals   0.1027869  0.9309457 -1.076173
## 2         No  p values   1.0000000  1.0000000  1.000000
## 3        Yes Residuals  -0.1027869 -0.9309457  1.076173
## 4        Yes  p values   1.0000000  1.0000000  1.000000
```

```r
# cp and NMD_NW_autmediv1m_fs
table_cp_NMD_NW_autmediv1m_fs <- table(new_data$DrQ3, new_data$NMD_NW_autmediv1m_fs)
chisq.test(table_cp_NMD_NW_autmediv1m_fs)
```

```
## Warning in chisq.test(table_cp_NMD_NW_autmediv1m_fs): Chi-squared approximation
## may be incorrect


##
##  Pearson's Chi-squared test
##
## data:  table_cp_NMD_NW_autmediv1m_fs
## X-squared = 2.0211, df = 2, p-value = 0.364
```

```
chisq.posthoc.test(table_cp_NMD_NW_autmediv1m_fs)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect


##    Dimension     Value          0          1          2
## 1         No Residuals  0.1027869  0.9309457 -1.076173
## 2         No  p values  1.0000000  1.0000000  1.000000
## 3        Yes Residuals -0.1027869 -0.9309457  1.076173
## 4        Yes  p values  1.0000000  1.0000000  1.000000
```

```
# cp and NMD_NW_autmed1m
table_cp_NMD_NW_autmed1m <- table(new_data$DrQ3, new_data$NMD_NW_autmed1m)
chisq.test(table_cp_NMD_NW_autmed1m)
```

```
##
##  Pearson's Chi-squared test with Yates' continuity correction
##
## data:  table_cp_NMD_NW_autmed1m
## X-squared = 0.092384, df = 1, p-value = 0.7612
```

```
chisq.posthoc.test(table_cp_NMD_NW_autmed1m)
```

```
##    Dimension     Value         1          2
## 1         No Residuals -0.551049  0.551049
## 2         No  p values  1.000000  1.000000
## 3        Yes Residuals  0.551049 -0.551049
## 4        Yes  p values  1.000000  1.000000
```

```
# cp and NMD_NW_Q32AJ
table_cp_NMD_NW_Q32AJ <- table(new_data$DrQ3, new_data$NMD_NW_Q32AJ)
chisq.test(table_cp_NMD_NW_Q32AJ)
```

```
## Warning in chisq.test(table_cp_NMD_NW_Q32AJ): Chi-squared approximation may be
## incorrect


##
##  Pearson's Chi-squared test
##
## data:  table_cp_NMD_NW_Q32AJ
## X-squared = 10.297, df = 11, p-value = 0.5039
```

```
chisq.posthoc.test(table_cp_NMD_NW_Q32AJ)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect

##    Dimension    Value        0         1         2         3         4
## 1        No Residuals  0.551049  0.9309457  0.9309457 -0.1027869 -1.152003
## 2        No  p values  1.000000  1.0000000  1.0000000  1.0000000  1.000000
## 3       Yes Residuals -0.551049 -0.9309457 -0.9309457  0.1027869  1.152003
## 4       Yes  p values  1.000000  1.0000000  1.0000000  1.0000000  1.000000
##            5         8        10        14        16        20        30
## 1 -1.076173 -1.076173  1.317781  0.9309457 -1.076173 -0.1027869 -1.076173
## 2  1.000000  1.000000  1.000000  1.0000000  1.000000  1.0000000  1.000000
## 3  1.076173  1.076173 -1.317781 -0.9309457  1.076173  0.1027869  1.076173
## 4  1.000000  1.000000  1.000000  1.0000000  1.000000  1.0000000  1.000000
```

```
# cp and NMD_NW_autmed1m_fs
table_cp_NMD_NW_autmed1m_fs <- table(new_data$DrQ3, new_data$NMD_NW_autmed1m_fs)
chisq.test(table_cp_NMD_NW_autmed1m_fs)
```

```
## Warning in chisq.test(table_cp_NMD_NW_autmed1m_fs): Chi-squared approximation
## may be incorrect

##
##  Pearson's Chi-squared test
##
## data:  table_cp_NMD_NW_autmed1m_fs
## X-squared = 3.5475, df = 3, p-value = 0.3146
```

```
chisq.posthoc.test(table_cp_NMD_NW_autmed1m_fs)
```

```
## Warning in chisq.test(x, ...): Chi-squared approximation may be incorrect

##    Dimension    Value        0         1         2         3
## 1        No Residuals  0.551049 -0.818065  1.317781 -1.076173
## 2        No  p values  1.000000  1.000000  1.000000  1.000000
## 3       Yes Residuals -0.551049  0.818065 -1.317781  1.076173
## 4       Yes  p values  1.000000  1.000000  1.000000  1.000000
```

As the p-Value of this test is high (greater than 0.1), so, we failed to reject the null hypothesis. So, from the chi-squared test, it is evident that the variables are independent of each other.

Training and Test Division

```
set.seed(1) # Set Seed so that same sample can be reproduced in future also
# Now Selecting 75% of data as sample from total 'n' rows of the data
sample <- sample.int(n = nrow(new_data), size = floor(.75*nrow(new_data)), replace = F)
TrainingSet <- new_data[sample, ]
TestingSet  <- new_data[-sample, ]
```

Dividing our sample into training (75 %) and testing (25%).

Univariate Logistic regression

```
Model_age_group <- glm(DrQ3 ~ age_group, family = binomial, data = new_data)
summary(Model_age_group)
```

```
##
## Call:
## glm(formula = DrQ3 ~ age_group, family = binomial, data = new_data)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.354  -1.076  -1.001   1.162   1.830
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -1.4663     0.6405  -2.289   0.0221 *
## age_group30-40   1.0356     0.6727   1.539   0.1237
## age_group40-50   1.2231     0.6604   1.852   0.0640 .
## age_group50-60   1.5018     0.6587   2.280   0.0226 *
## age_group60-70   1.6172     0.6734   2.401   0.0163 *
## age_group>=70    1.8718     1.1152   1.679   0.0932 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 732.90  on 533  degrees of freedom
## AIC: 744.9
##
## Number of Fisher Scoring iterations: 4
```

```
# exponentiate the coefficients (Odds ratio)
exp(coef(Model_age_group))
```

```
##    (Intercept) age_group30-40 age_group40-50 age_group50-60 age_group60-70
##      0.2307692      2.8166667      3.3977273      4.4899598      5.0387597
##  age_group>=70
##      6.5000000
```

```
# 95% CI (odds ratio)
exp(confint(Model_age_group))
```

```
## Waiting for profiling to be done...
```

```
##                     2.5 %     97.5 %
## (Intercept)    0.05290604   0.715872
## age_group30-40 0.84184910 12.859959
## age_group40-50 1.04545672 15.243156
## age_group50-60 1.38709299 20.097089
## age_group60-70 1.50508997 23.038152
## age_group>=70  0.76697860 71.144690
```

Here, p value is less than 0.1 for groups like age_group<=30, age_group40-50, age_group50-60, age_group60-70, age_group>=70, therefore it is statistically significant.

```
Model_age_40 <- glm(DrQ3 ~  age_40, family = binomial, data = new_data)
summary(Model_age_40)
```

```
##
## Call:
## glm(formula = DrQ3 ~ age_40, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min      1Q   Median      3Q      Max
## -1.1614  -1.1614  -0.9534   1.1935   1.4193
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.5528     0.1937  -2.854  0.00431 **
## age_401       0.5150     0.2167   2.377  0.01745 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 738.60  on 537  degrees of freedom
## AIC: 742.6
##
## Number of Fisher Scoring iterations: 4
```

```
# exponentiate the coefficients (Odds ratio)
exp(coef(Model_age_40))
```

```
## (Intercept)     age_401
##   0.5753425   1.6737213
```

```
# 95% CI (odds ratio)
exp(confint(Model_age_40))
```

```
## Waiting for profiling to be done...
```

```
##                 2.5 %     97.5 %
## (Intercept) 0.3906065 0.8364935
## age_401     1.0993145 2.5748320
```

Here, p-value < 0.05, therefore the above features turns out to be significant.

```
Model_unstabh_1mb <- glm(DrQ3 ~ unstabh_1mb , family = binomial, data = new_data)
summary(Model_unstabh_1mb)
```

```
##
## Call:
```

```
## glm(formula = DrQ3 ~ unstabh_1mb, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.124   -1.124   -1.098    1.232    1.259
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.12648    0.10280  -1.230    0.219
## unstabh_1mb1 -0.06276    0.18961  -0.331    0.741
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 744.28  on 537  degrees of freedom
## AIC: 748.28
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_unstabh_1mb))
```

```
##  (Intercept) unstabh_1mb1
##    0.8811881    0.9391709
```

```
exp(confint(Model_unstabh_1mb))
```

```
## Waiting for profiling to be done...
```

```
##                    2.5 %    97.5 %
## (Intercept)   0.7199494 1.077614
## unstabh_1mb1  0.6467733 1.361133
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_OAT_3m <- glm(DrQ3 ~ OAT_3m -1, family = binomial, data = new_data)
summary(Model_OAT_3m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ OAT_3m - 1, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.162   -1.085   -1.085    1.273    1.273
##
## Coefficients:
##         Estimate Std. Error z value Pr(>|z|)
## OAT_3m0 -0.22172    0.11302  -1.962   0.0498 *
## OAT_3m1 -0.03604    0.13425  -0.268   0.7884
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 747.21  on 539  degrees of freedom
## Residual deviance: 743.27  on 537  degrees of freedom
## AIC: 747.27
##
## Number of Fisher Scoring iterations: 3
```

```r
exp(coef(Model_OAT_3m))
```

```
##   OAT_3m0   OAT_3m1
## 0.8011364 0.9646018
```

```r
exp(confint(Model_OAT_3m))
```

```
## Waiting for profiling to be done...
```

```
##             2.5 %     97.5 %
## OAT_3m0 0.6412233 0.9991187
## OAT_3m1 0.7409952 1.2551613
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant. But, the value of intercept is significant as p < 0.05.

```r
Model_incar_3m <- glm(DrQ3 ~ incar_3m , family = binomial, data = new_data)
summary(Model_incar_3m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ incar_3m, family = binomial, data = new_data)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.222  -1.104  -1.104   1.252   1.252
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.17472    0.09145  -1.911   0.0561 .
## incar_3m1    0.28008    0.28059   0.998   0.3182
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 743.39  on 537  degrees of freedom
## AIC: 747.39
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_incar_3m))
```

```
## (Intercept)    incar_3m1
##   0.8396947    1.3232323
```

```
exp(confint(Model_incar_3m))
```

```
## Waiting for profiling to be done...
```

```
##                 2.5 %    97.5 %
## (Intercept) 0.7014996 1.004174
## incar_3m1   0.7632360 2.304178
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant. But, the value of intercept is significant as p $\sim<= 0.05$.

```
Model_opiiv_pm <- glm(DrQ3 ~ opiiv_pm , family = binomial, data = new_data)
summary(Model_opiiv_pm)
```

```
##
## Call:
## glm(formula = DrQ3 ~ opiiv_pm, family = binomial, data = new_data)
##
## Deviance Residuals:
##    Min     1Q  Median     3Q     Max
## -1.127  -1.127  -1.089   1.228   1.268
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.11900    0.10185  -1.168    0.243
## opiiv_pm1   -0.09231    0.19231  -0.480    0.631
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 744.16  on 537  degrees of freedom
## AIC: 748.16
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_opiiv_pm))
```

```
## (Intercept)    opiiv_pm1
##   0.8878049    0.9118263
```

```
exp(confint(Model_opiiv_pm))
```

```
## Waiting for profiling to be done...
```

```
##                2.5 %   97.5 %
## (Intercept) 0.7267450 1.083701
## opiiv_pm1    0.6244625 1.328270
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_heroiv_pm <- glm(DrQ3 ~ heroiv_pm , family = binomial, data = new_data)
summary(Model_heroiv_pm)
```

```
##
## Call:
## glm(formula = DrQ3 ~ heroiv_pm, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min      1Q  Median      3Q     Max
## -1.142  -1.142  -1.034   1.213   1.328
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.08277    0.09874  -0.838    0.402
## heroiv_pm1  -0.26442    0.20482  -1.291    0.197
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 742.71  on 537  degrees of freedom
## AIC: 746.71
##
## Number of Fisher Scoring iterations: 4
```

```
exp(coef(Model_heroiv_pm))
```

```
## (Intercept)  heroiv_pm1
##   0.9205607   0.7676481
```

```
exp(confint(Model_heroiv_pm))
```

```
## Waiting for profiling to be done...
```

```
##                2.5 %   97.5 %
## (Intercept) 0.7582787 1.116995
## heroiv_pm1  0.5121248 1.144518
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_cocaiv_pm <- glm(DrQ3 ~ cocaiv_pm , family = binomial, data = new_data)
summary(Model_cocaiv_pm)
```

```
##
## Call:
```

```
## glm(formula = DrQ3 ~ cocaiv_pm, family = binomial, data = new_data)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.155  -1.155  -1.051   1.200   1.310
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.0528     0.1083  -0.487    0.626
## cocaiv_pm1   -0.2526     0.1800  -1.403    0.161
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 742.41  on 537  degrees of freedom
## AIC: 746.41
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_cocaiv_pm))
```

```
## (Intercept)   cocaiv_pm1
##   0.9485714    0.7767914
```

```
exp(confint(Model_cocaiv_pm))
```

```
## Waiting for profiling to be done...
```

```
##                  2.5 %    97.5 %
## (Intercept) 0.7667859 1.172987
## cocaiv_pm1  0.5450034 1.104475
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_CO_fsaut11m <- glm(DrQ3 ~ CO_fsaut11m , family = binomial, data = new_data)
summary(Model_CO_fsaut11m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ CO_fsaut11m, family = binomial, data = new_data)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.121  -1.121  -1.121   1.235   1.235
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -14.57     509.65  -0.029    0.977
## CO_fsaut11m2   14.43     509.65   0.028    0.977
##
## (Dispersion parameter for binomial family taken to be 1)
```

```
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 740.63  on 537  degrees of freedom
## AIC: 744.63
##
## Number of Fisher Scoring iterations: 13
```

```
exp(coef(Model_CO_fsaut11m))
```

```
##  (Intercept) CO_fsaut11m2
## 4.721034e-07 1.851556e+06
```

```
exp(confint(Model_CO_fsaut11m))
```

```
## Waiting for profiling to be done...
```

```
##                      2.5 %      97.5 %
## (Intercept)             NA 3.604543e+22
## CO_fsaut11m2 2.447943e-23          NA
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_CO_iaut11m <- glm(DrQ3 ~ CO_iaut11m , family = binomial, data = new_data)
summary(Model_CO_iaut11m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ CO_iaut11m, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.115   -1.115   -1.115    1.241    1.241
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)     13.57     535.41   0.025     0.98
## CO_iaut11m2    -13.72     535.41  -0.026     0.98
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 742.85  on 537  degrees of freedom
## AIC: 746.85
##
## Number of Fisher Scoring iterations: 12
```

```
exp(coef(Model_CO_iaut11m))
```

```
##  (Intercept) CO_iaut11m2
## 7.792339e+05 1.105691e-06
```

```
exp(confint(Model_CO_iaut11m))
```

```
## Waiting for profiling to be done...
```

```
##                     2.5 %        97.5 %
## (Intercept) 3.163471e-42           NA
## CO_iaut11m2           NA 2.714965e+41
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_NMD_NW_autmed1m <- glm(DrQ3 ~ NMD_NW_autmed1m , family = binomial, data = new_data)
summary(Model_NMD_NW_autmed1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ NMD_NW_autmed1m, family = binomial, data = new_data)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.228  -1.113  -1.113   1.243   1.243
##
## Coefficients:
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept)        0.1178     0.4859   0.242    0.808
## NMD_NW_autmed1m2  -0.2713     0.4938  -0.550    0.583
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 744.09  on 537  degrees of freedom
## AIC: 748.09
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_NMD_NW_autmed1m))
```

```
##      (Intercept) NMD_NW_autmed1m2
##        1.1250000        0.7623567
```

```
exp(confint(Model_NMD_NW_autmed1m))
```

```
## Waiting for profiling to be done...
```

```
##                      2.5 %   97.5 %
## (Intercept)      0.4301253 2.996435
## NMD_NW_autmed1m2 0.2821425 2.023818
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_NMD_NW_autmediv1m <- glm(DrQ3 ~ NMD_NW_autmediv1m , family = binomial, data = new_data)
summary(Model_NMD_NW_autmediv1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ NMD_NW_autmediv1m, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min      1Q  Median      3Q     Max
## -1.177  -1.116  -1.116   1.240   1.240
##
## Coefficients:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)       -4.039e-14  1.414e+00   0.000    1.000
## NMD_NW_autmediv1m2 -1.455e-01  1.417e+00  -0.103    0.918
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 744.38  on 537  degrees of freedom
## AIC: 748.38
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_NMD_NW_autmediv1m))
```

```
##       (Intercept) NMD_NW_autmediv1m2
##         1.0000000          0.8645833
```

```
exp(confint(Model_NMD_NW_autmediv1m))
```

```
## Waiting for profiling to be done...
```

```
##                      2.5 %    97.5 %
## (Intercept)       0.03954417 25.28818
## NMD_NW_autmediv1m2 0.03407155 21.93923
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_NW_psych1m <- glm(DrQ3 ~ NW_psych1m , family = binomial, data = new_data)
summary(Model_NW_psych1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ NW_psych1m, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min      1Q  Median      3Q     Max
## -1.126  -1.126  -1.126   1.229   1.435
##
```

```
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.5878     0.3944  -1.490    0.136
## NW_psych1m2   0.4663     0.4042   1.154    0.249
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 743.01  on 537  degrees of freedom
## AIC: 747.01
##
## Number of Fisher Scoring iterations: 4
```

```
exp(coef(Model_NW_psych1m))
```

```
## (Intercept) NW_psych1m2
##   0.5555556   1.5940959
```

```
exp(confint(Model_NW_psych1m))
```

```
## Waiting for profiling to be done...
```

```
##                 2.5 %   97.5 %
## (Intercept) 0.2466935 1.180906
## NW_psych1m2 0.7350746 3.652772
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_NW_psychiv1m <- glm(DrQ3 ~ NW_psychiv1m , family = binomial, data = new_data)
summary(Model_NW_psychiv1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ NW_psychiv1m, family = binomial, data = new_data)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.118  -1.118  -1.118   1.238   1.482
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.6931     1.2242  -0.566    0.571
## NW_psychiv1m2   0.5511     1.2273   0.449    0.653
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 744.18  on 537  degrees of freedom
## AIC: 748.18
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_NW_psychiv1m))
```

```
##    (Intercept) NW_psychiv1m2
##      0.5000006     1.7351897
```

```
exp(confint(Model_NW_psychiv1m))
```

```
## Waiting for profiling to be done...
```

```
##                   2.5 %   97.5 %
## (Intercept)   0.0232356  5.22020
## NW_psychiv1m2 0.1652575 37.47044
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_CO_mari1m <- glm(DrQ3 ~ CO_mari1m , family = binomial, data = new_data)
summary(Model_CO_mari1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ CO_mari1m, family = binomial, data = new_data)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.127  -1.127  -1.104   1.229   1.252
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.12095    0.11607  -1.042    0.297
## CO_mari1m2  -0.05376    0.17377  -0.309    0.757
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 744.29  on 537  degrees of freedom
## AIC: 748.29
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_CO_mari1m))
```

```
## (Intercept)  CO_mari1m2
##   0.8860759   0.9476554
```

```
exp(confint(Model_CO_mari1m))
```

```
## Waiting for profiling to be done...
```

```
##                 2.5 %   97.5 %
## (Intercept) 0.7052432 1.112114
## CO_mari1m2  0.6737790 1.332095
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_NW_tranquil1m <- glm(DrQ3 ~ NW_tranquil1m , family = binomial, data = new_data)
summary(Model_NW_tranquil1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ NW_tranquil1m, family = binomial, data = new_data)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.192  -1.108  -1.108   1.249   1.249
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)     0.03509    0.26495   0.132    0.895
## NW_tranquil1m2 -0.20145    0.28027  -0.719    0.472
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 743.87  on 537  degrees of freedom
## AIC: 747.87
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_NW_tranquil1m))
```

```
##    (Intercept) NW_tranquil1m2
##      1.0357143      0.8175453
```

```
exp(confint(Model_NW_tranquil1m))
```

```
## Waiting for profiling to be done...
```

```
##                    2.5 %   97.5 %
## (Intercept)    0.6148746 1.747366
## NW_tranquil1m2 0.4704697 1.418513
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_NW_tranquiliv1m <- glm(DrQ3 ~ NW_tranquiliv1m , family = binomial, data = new_data)
summary(Model_NW_tranquiliv1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ NW_tranquiliv1m, family = binomial, data = new_data)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.354  -1.114  -1.114   1.242   1.242
```

```
##
## Coefficients:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)       0.4055     0.9128   0.444    0.657
## NW_tranquiliv1m2 -0.5556     0.9169  -0.606    0.545
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 744.01  on 537  degrees of freedom
## AIC: 748.01
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_NW_tranquiliv1m))
```

```
##      (Intercept) NW_tranquiliv1m2
##        1.4999999        0.5737515
```

```
exp(confint(Model_NW_tranquiliv1m))
```

```
## Waiting for profiling to be done...
```

```
##                       2.5 %     97.5 %
## (Intercept)      0.24852661 11.388431
## NW_tranquiliv1m2 0.07512011  3.489101
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_NMD_COMD_fsautamph1m <- glm(DrQ3 ~ NMD_COMD_fsautamph1m , family = binomial, data = new_data)
summary(Model_NMD_COMD_fsautamph1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ NMD_COMD_fsautamph1m, family = binomial,
##     data = new_data)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.123  -1.123  -1.091   1.233   1.266
##
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -0.20634    0.19438  -1.062    0.288
## NMD_COMD_fsautamph1m2  0.07652    0.21698   0.353    0.724
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 744.26  on 537  degrees of freedom
## AIC: 748.26
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_NMD_COMD_fsautamph1m))
```

```
##            (Intercept) NMD_COMD_fsautamph1m2
##              0.8135593             1.0795290
```

```
exp(confint(Model_NMD_COMD_fsautamph1m))
```

```
## Waiting for profiling to be done...
```

```
##                              2.5 %    97.5 %
## (Intercept)              0.5537894 1.189221
## NMD_COMD_fsautamph1m2 0.7063308 1.656307
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_NMD_CO_iamph1m <- glm(DrQ3 ~ NMD_CO_iamph1m , family = binomial, data = new_data)
summary(Model_NMD_CO_iamph1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ NMD_CO_iamph1m, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.133  -1.115  -1.115    1.242    1.242
##
## Coefficients:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -0.1054     0.2653  -0.397    0.691
## NMD_CO_iamph1m2  -0.0443     0.2806  -0.158    0.875
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 744.36  on 537  degrees of freedom
## AIC: 748.36
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_NMD_CO_iamph1m))
```

```
##     (Intercept) NMD_CO_iamph1m2
##        0.900000        0.956671
```

```
exp(confint(Model_NMD_CO_iamph1m))
```

```
## Waiting for profiling to be done...
```

```
##                    2.5 %    97.5 %
## (Intercept)     0.5322572 1.514561
## NMD_CO_iamph1m2 0.5517869 1.665632
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_CO_autopiac1m <- glm(DrQ3 ~ CO_autopiac1m , family = binomial, data = new_data)
summary(Model_CO_autopiac1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ CO_autopiac1m, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min      1Q  Median      3Q     Max
## -1.124  -1.124  -1.001   1.232   1.365
##
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -0.4308     0.3562  -1.209    0.227
## CO_autopiac1m2   0.3041     0.3672   0.828    0.408
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 743.69  on 537  degrees of freedom
## AIC: 747.69
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_CO_autopiac1m))
```

```
##   (Intercept) CO_autopiac1m2
##     0.6500001      1.3554474
```

```
exp(confint(Model_CO_autopiac1m))
```

```
## Waiting for profiling to be done...
```

```
##                   2.5 %   97.5 %
## (Intercept)   0.3154544 1.293382
## CO_autopiac1m2 0.6664102 2.849168
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_NH_Q7Jb_metha <- glm(DrQ3 ~ NH_Q7Jb_metha , family = binomial, data = new_data)
summary(Model_NH_Q7Jb_metha)
```

```
##
## Call:
## glm(formula = DrQ3 ~ NH_Q7Jb_metha, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min      1Q  Median      3Q     Max
## -1.119  -1.119  -1.119   1.237   1.482
```

```
## 
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)     -0.6931     0.8660  -0.800    0.423
## NH_Q7Jb_metha2   0.5541     0.8704   0.637    0.524
## 
## (Dispersion parameter for binomial family taken to be 1)
## 
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 743.96  on 537  degrees of freedom
## AIC: 747.96
## 
## Number of Fisher Scoring iterations: 4
```

```
exp(coef(Model_NH_Q7Jb_metha))
```

```
##    (Intercept) NH_Q7Jb_metha2
##       0.500000       1.740351
```

```
exp(confint(Model_NH_Q7Jb_metha))
```

```
## Waiting for profiling to be done...
```

```
##                     2.5 %    97.5 %
## (Intercept)    0.06930964  2.561686
## NH_Q7Jb_metha2 0.33674181 12.631050
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_NH_Q7Ib_subox <- glm(DrQ3 ~ NH_Q7Ib_subox , family = binomial, data = new_data)
summary(Model_NH_Q7Ib_subox)
```

```
## 
## Call:
## glm(formula = DrQ3 ~ NH_Q7Ib_subox, family = binomial, data = new_data)
## 
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.119  -1.119  -1.119   1.237   1.237
## 
## Coefficients:
##                Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -14.57     624.19  -0.023    0.981
## NH_Q7Ib_subox2    14.43     624.19   0.023    0.982
## 
## (Dispersion parameter for binomial family taken to be 1)
## 
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 741.89  on 537  degrees of freedom
## AIC: 745.89
## 
## Number of Fisher Scoring iterations: 13
```

```
exp(coef(Model_NH_Q7Ib_subox))
```

```
##    (Intercept) NH_Q7Ib_subox2
##   4.721034e-07   1.845104e+06
```

```
exp(confint(Model_NH_Q7Ib_subox))
```

```
## Waiting for profiling to be done...
```

```
##                        2.5 %      97.5 %
## (Intercept)              NA 1.730103e+36
## NH_Q7Ib_subox2 5.082249e-37         NA
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_NH_Q7Ia_suboxIV <- glm(DrQ3 ~ NH_Q7Ia_suboxIV , family = binomial, data = new_data)
summary(Model_NH_Q7Ia_suboxIV)
```

```
##
## Call:
## glm(formula = DrQ3 ~ NH_Q7Ia_suboxIV, family = binomial, data = new_data)
##
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.118  -1.118  -1.118   1.238   1.238
##
## Coefficients:
##                 Estimate Std. Error z value Pr(>|z|)
## (Intercept)       -13.57     535.41  -0.025     0.98
## NH_Q7Ia_suboxIV2   13.42     535.41   0.025     0.98
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 743.14  on 537  degrees of freedom
## AIC: 747.14
##
## Number of Fisher Scoring iterations: 12
```

```
exp(coef(Model_NH_Q7Ia_suboxIV))
```

```
##     (Intercept) NH_Q7Ia_suboxIV2
##     1.283312e-06     6.764183e+05
```

```
exp(confint(Model_NH_Q7Ia_suboxIV))
```

```
## Waiting for profiling to be done...
```

```
##                          2.5 %      97.5 %
## (Intercept)                NA 3.161083e+41
## NH_Q7Ia_suboxIV2 2.754769e-42         NA
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_COMD_nispeed1m <- glm(DrQ3 ~ COMD_nispeed1m , family = binomial, data = new_data)
summary(Model_COMD_nispeed1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ COMD_nispeed1m, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min      1Q  Median      3Q     Max
## -1.177  -1.118  -1.118   1.238   1.281
##
## Coefficients:
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept)      -0.14086    0.08861  -1.590    0.112
## COMD_nispeed1m1  -0.10030    0.41254  -0.243    0.808
## COMD_nispeed1m2   0.14086    1.41699   0.099    0.921
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 744.32  on 536  degrees of freedom
## AIC: 750.32
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_COMD_nispeed1m))
```

```
##     (Intercept) COMD_nispeed1m1 COMD_nispeed1m2
##       0.8686131       0.9045618       1.1512605
```

```
exp(confint(Model_COMD_nispeed1m))
```

```
## Waiting for profiling to be done...
```

```
##                      2.5 %     97.5 %
## (Intercept)     0.72979666   1.033102
## COMD_nispeed1m1 0.39428166   2.025603
## COMD_nispeed1m2 0.04536141  29.218718
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_CO_ispeed1m <- glm(DrQ3 ~ CO_ispeed1m , family = binomial, data = new_data)
summary(Model_CO_ispeed1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ CO_ispeed1m, family = binomial, data = new_data)
##
## Deviance Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -1.118  -1.118  -1.118   1.238   1.274
##
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)   -0.22314    0.38730  -0.576    0.565
## CO_ispeed1m2   0.08229    0.39730   0.207    0.836
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 744.35  on 537  degrees of freedom
## AIC: 748.35
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_CO_ispeed1m))
```

```
##  (Intercept) CO_ispeed1m2
##     0.800000     1.085766
```

```
exp(confint(Model_CO_ispeed1m))
```

```
## Waiting for profiling to be done...
```

```
##                   2.5 %    97.5 %
## (Intercept)   0.3670411 1.706307
## CO_ispeed1m2  0.4992310 2.410582
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_COMD_fcrakfree1m <- glm(DrQ3 ~ COMD_fcrakfree1m , family = binomial, data = new_data)
summary(Model_COMD_fcrakfree1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ COMD_fcrakfree1m, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min      1Q  Median      3Q     Max
## -1.128  -1.128  -1.098   1.228   1.259
##
## Coefficients:
##                   Estimate Std. Error z value Pr(>|z|)
## (Intercept)       -0.18962    0.14170  -1.338    0.181
## COMD_fcrakfree1m2  0.07114    0.17876   0.398    0.691
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 744.23  on 537  degrees of freedom
## AIC: 748.23
##
## Number of Fisher Scoring iterations: 3
```

```r
exp(coef(Model_COMD_fcrakfree1m))
```

```
##     (Intercept) COMD_fcrakfree1m2
##       0.8272727        1.0737307
```

```r
exp(confint(Model_COMD_fcrakfree1m))
```

```
## Waiting for profiling to be done...
```

```
##                      2.5 %    97.5 %
## (Intercept)      0.6256124 1.091272
## COMD_fcrakfree1m2 0.7566139 1.525588
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```r
Model_CO_fpoud1m <- glm(DrQ3 ~ CO_fpoud1m , family = binomial, data = new_data)
summary(Model_CO_fpoud1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ CO_fpoud1m, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min      1Q  Median      3Q     Max
## -1.256  -1.114  -1.114   1.243   1.243
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)   0.1823     0.6055   0.301    0.763
## CO_fpoud1m2  -0.3341     0.6118  -0.546    0.585
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 744.09  on 537  degrees of freedom
## AIC: 748.09
##
## Number of Fisher Scoring iterations: 3
```

```r
exp(coef(Model_CO_fpoud1m))
```

```
## (Intercept) CO_fpoud1m2
##   1.2000000   0.7159624
```

```r
exp(confint(Model_CO_fpoud1m))
```

```
## Waiting for profiling to be done...
```

```
##                 2.5 %    97.5 %
## (Intercept) 0.3614086 4.163874
## CO_fpoud1m2 0.2041007 2.405159
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_CO_icrak1m <- glm(DrQ3 ~ CO_icrak1m , family = binomial, data = new_data)
summary(Model_CO_icrak1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ CO_icrak1m, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min      1Q  Median      3Q     Max
## -1.126  -1.126  -1.126   1.229   1.435
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.5878     0.3944  -1.490    0.136
## CO_icrak1m2   0.4663     0.4042   1.154    0.249
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 743.01  on 537  degrees of freedom
## AIC: 747.01
##
## Number of Fisher Scoring iterations: 4
```

```
exp(coef(Model_CO_icrak1m))
```

```
## (Intercept) CO_icrak1m2
##   0.5555556   1.5940959
```

```
exp(confint(Model_CO_icrak1m))
```

```
## Waiting for profiling to be done...
```

```
##                  2.5 %    97.5 %
## (Intercept) 0.2466935 1.180906
## CO_icrak1m2 0.7350746 3.652772
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_CO_ipoud1m <- glm(DrQ3 ~ CO_ipoud1m , family = binomial, data = new_data)
summary(Model_CO_ipoud1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ CO_ipoud1m, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min      1Q  Median      3Q     Max
## -1.151  -1.151  -1.055   1.204   1.304
```

```
## 
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.5241     0.3109  -1.686   0.0919 .
## CO_ipoud1m    0.2303     0.1812   1.271   0.2036
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## 
## (Dispersion parameter for binomial family taken to be 1)
## 
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 742.77  on 537  degrees of freedom
## AIC: 746.77
## 
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_CO_ipoud1m))
```

```
## (Intercept)  CO_ipoud1m
##   0.5920878   1.2590271
```

```
exp(confint(Model_CO_ipoud1m))
```

```
## Waiting for profiling to be done...
```

```
##                 2.5 %    97.5 %
## (Intercept) 0.3205082 1.086184
## CO_ipoud1m  0.8835098 1.798506
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant. But, the value of intercept is significant as p $\sim<= 0.01$

```
Model_COMD_fshero1m <- glm(DrQ3 ~ COMD_fshero1m , family = binomial, data = new_data)
summary(Model_COMD_fshero1m)
```

```
## 
## Call:
## glm(formula = DrQ3 ~ COMD_fshero1m, family = binomial, data = new_data)
## 
## Deviance Residuals:
##    Min      1Q  Median      3Q     Max
## -1.119  -1.119  -1.119   1.237   1.354
## 
## Coefficients:
##               Estimate Std. Error z value Pr(>|z|)
## (Intercept)    -0.6708     1.2938  -0.518    0.604
## COMD_fshero1m   0.2653     0.6513   0.407    0.684
## 
## (Dispersion parameter for binomial family taken to be 1)
## 
##     Null deviance: 744.39  on 538  degrees of freedom
```

```
## Residual deviance: 744.22  on 537  degrees of freedom
## AIC: 748.22
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_COMD_fshero1m))
```

```
##   (Intercept) COMD_fshero1m
##     0.5112918     1.3038869
```

```
exp(confint(Model_COMD_fshero1m))
```

```
## Waiting for profiling to be done...
```

```
##                     2.5 %    97.5 %
## (Intercept)   0.03325573 6.301740
## COMD_fshero1m 0.36825476 5.149621
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_CO_ihero1m <- glm(DrQ3 ~ CO_ihero1m , family = binomial, data = new_data)
summary(Model_CO_ihero1m)
```

```
##
## Call:
## glm(formula = DrQ3 ~ CO_ihero1m, family = binomial, data = new_data)
##
## Deviance Residuals:
##    Min     1Q  Median     3Q     Max
## -1.142  -1.142  -1.034   1.213   1.328
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -0.6116     0.3722  -1.643    0.100
## CO_ihero1m    0.2644     0.2048   1.291    0.197
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 742.71  on 537  degrees of freedom
## AIC: 746.71
##
## Number of Fisher Scoring iterations: 4
```

```
exp(coef(Model_CO_ihero1m))
```

```
## (Intercept)  CO_ihero1m
##   0.5424713   1.3026803
```

```
exp(confint(Model_CO_ihero1m))
```

```
## Waiting for profiling to be done...
```

```
##                 2.5 %    97.5 %
## (Intercept) 0.2593845 1.119622
## CO_ihero1m  0.8737306 1.952649
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

NW_Q21C1

```
Model_NW_Q21C1 <- glm(DrQ3 ~ NW_Q21C1 -1, family = binomial, data = new_data)
summary(Model_NW_Q21C1)
```

```
##
## Call:
## glm(formula = DrQ3 ~ NW_Q21C1 - 1, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min      1Q  Median      3Q     Max
## -1.172  -1.172  -1.035   1.183   1.327
##
## Coefficients:
##           Estimate Std. Error z value Pr(>|z|)
## NW_Q21C11  -0.0125     0.1118  -0.112   0.9110
## NW_Q21C12  -0.3444     0.1378  -2.499   0.0124 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.44  on 537  degrees of freedom
## Residual deviance: 738.09  on 535  degrees of freedom
##   (2 observations deleted due to missingness)
## AIC: 742.09
##
## Number of Fisher Scoring iterations: 4
```

```
exp(coef(Model_NW_Q21C1))
```

```
## NW_Q21C11 NW_Q21C12
## 0.9875776 0.7086614
```

```
exp(confint(Model_NW_Q21C1))
```

```
## Waiting for profiling to be done...
```

```
##              2.5 %    97.5 %
## NW_Q21C11 0.7930210 1.229743
## NW_Q21C12 0.5395985 0.926831
```

Here, p-value < 0.1, therefore the above features i.e NW_Q21C12 turns out to be significant.

```
Model_CO_njalc1 <- glm(DrQ3 ~ CO_njalc1 , family = binomial, data = new_data)
summary(Model_CO_njalc1)
```

```
##
## Call:
## glm(formula = DrQ3 ~ CO_njalc1, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min      1Q  Median      3Q     Max
## -1.279  -1.073  -1.066   1.271   1.293
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.267130   0.105916  -2.522   0.0117 *
## CO_njalc1    0.016732   0.008334   2.008   0.0447 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 740.33  on 537  degrees of freedom
## AIC: 744.33
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_CO_njalc1))
```

```
## (Intercept)   CO_njalc1
##   0.7655733   1.0168727
```

```
exp(confint(Model_CO_njalc1))
```

```
## Waiting for profiling to be done...
```

```
##                   2.5 %    97.5 %
## (Intercept) 0.6214312 0.9415669
## CO_njalc1   1.0004492 1.0337353
```

Here, p-value < 0.05, therefore the above features turns out to be significant.

```
Model_CO_njalc1_group <- glm(DrQ3 ~ CO_njalc1_group , family = binomial, data = new_data)
summary(Model_CO_njalc1_group)
```

```
##
## Call:
## glm(formula = DrQ3 ~ CO_njalc1_group, family = binomial, data = new_data)
##
## Deviance Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -1.360  -1.129  -1.026   1.227   1.337
##
## Coefficients:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)          -0.3677     0.1307  -2.813  0.00491 **
## CO_njalc1_group1-10   0.2529     0.2065   1.225  0.22067
## CO_njalc1_group11-20  0.7876     0.2986   2.638  0.00834 **
## CO_njalc1_group21-29  0.7424     0.4129   1.798  0.07218 .
## CO_njalc1_group30     0.2585     0.3001   0.862  0.38893
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 735.06  on 534  degrees of freedom
## AIC: 745.06
##
## Number of Fisher Scoring iterations: 4
```

```
exp(coef(Model_CO_njalc1_group))
```

```
##          (Intercept)  CO_njalc1_group1-10 CO_njalc1_group11-20
##            0.6923077            1.2878179            2.1980676
## CO_njalc1_group21-29    CO_njalc1_group30
##            2.1010101            1.2950192
```

```
exp(confint(Model_CO_njalc1_group))
```

```
## Waiting for profiling to be done...
```

```
##                            2.5 %    97.5 %
## (Intercept)            0.5345495 0.8930514
## CO_njalc1_group1-10    0.8589755 1.9316572
## CO_njalc1_group11-20   1.2315310 3.9878902
## CO_njalc1_group21-29   0.9433685 4.8399921
## CO_njalc1_group30      0.7163218 2.3331954
```

Here, p-value < 0.1, therefore the above features turns out to be significant. Here intercept i.e CO_njalc1_group (0), CO_njalc1_group (11-20), CO_njalc1_group (21-29) are significant as there $p < 0.1$.

```
Model_NW_jourcons <- glm(DrQ3 ~ NW_jourcons , family = binomial, data = new_data)
summary(Model_NW_jourcons)
```

```
##
## Call:
## glm(formula = DrQ3 ~ NW_jourcons, family = binomial, data = new_data)
##
## Deviance Residuals:
```

```
##     Min      1Q  Median      3Q     Max
## -1.129  -1.125  -1.092   1.227   1.265
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.203521   0.173859  -1.171    0.242
## NW_jourcons  0.002955   0.007609   0.388    0.698
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 744.24  on 537  degrees of freedom
## AIC: 748.24
##
## Number of Fisher Scoring iterations: 3
```

```
exp(coef(Model_NW_jourcons))
```

```
## (Intercept) NW_jourcons
##   0.8158534   1.0029596
```

```
exp(confint(Model_NW_jourcons))
```

```
## Waiting for profiling to be done...
```

```
##                  2.5 %    97.5 %
## (Intercept) 0.5790440 1.146044
## NW_jourcons 0.9881267 1.018075
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
(Model_VBA3 <- glm(DrQ3 ~ VBA3 , family = binomial, data = new_data))
```

```
##
## Call:  glm(formula = DrQ3 ~ VBA3, family = binomial, data = new_data)
##
## Coefficients:
## (Intercept)        VBA32
##    -0.09742     -0.29847
##
## Degrees of Freedom: 538 Total (i.e. Null);  537 Residual
## Null Deviance:        744.4
## Residual Deviance: 742.8      AIC: 746.8
```

```
summary(Model_VBA3)
```

```
##
## Call:
## glm(formula = DrQ3 ~ VBA3, family = binomial, data = new_data)
##
```

```
## Deviance Residuals:
##     Min       1Q  Median      3Q      Max
## -1.136  -1.136  -1.015   1.219    1.349
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept) -0.09742    0.09418  -1.034    0.301
## VBA32       -0.29847    0.23806  -1.254    0.210
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 742.80  on 537  degrees of freedom
## AIC: 746.8
##
## Number of Fisher Scoring iterations: 4
```

```
exp(coef(Model_VBA3))
```

```
## (Intercept)       VBA32
##   0.9071730   0.7419499
```

```
exp(confint(Model_VBA3))
```

```
## Waiting for profiling to be done...
```

```
##                   2.5 %    97.5 %
## (Intercept) 0.7539512 1.090929
## VBA32       0.4623500 1.178685
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

```
Model_CO_scolar <- glm(DrQ3 ~ CO_scolar , family = binomial, data = new_data)
summary(Model_CO_scolar)
```

```
##
## Call:
## glm(formula = DrQ3 ~ CO_scolar, family = binomial, data = new_data)
##
## Deviance Residuals:
##     Min       1Q  Median      3Q      Max
## -1.387  -1.107  -1.070   1.250    1.525
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)    14.57     624.19   0.023    0.981
## CO_scolar2    -14.57     624.19  -0.023    0.981
## CO_scolar3    -15.35     624.19  -0.025    0.980
## CO_scolar4    -14.73     624.19  -0.024    0.981
## CO_scolar5    -14.82     624.19  -0.024    0.981
## CO_scolar6    -14.60     624.19  -0.023    0.981
```

```
## CO_scolar7    -14.81      624.19  -0.024      0.981
## CO_scolar8    -14.09      624.19  -0.023      0.982
## CO_scolar9    -14.62      624.19  -0.023      0.981
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 744.39  on 538  degrees of freedom
## Residual deviance: 735.64  on 530  degrees of freedom
## AIC: 753.64
##
## Number of Fisher Scoring iterations: 13
```

```
exp(coef(Model_CO_scolar))
```

```
##  (Intercept)   CO_scolar2   CO_scolar3   CO_scolar4   CO_scolar5   CO_scolar6
## 2.118180e+06 4.721034e-07 2.145925e-07 3.990977e-07 3.645356e-07 4.563666e-07
##   CO_scolar7   CO_scolar8   CO_scolar9
## 3.700270e-07 7.626286e-07 4.472559e-07
```

```
exp(confint(Model_CO_scolar))
```

```
## Waiting for profiling to be done...
```

```
##                   2.5 %       97.5 %
## (Intercept) 5.780007e-37          NA
## CO_scolar2            NA 5.854700e+35
## CO_scolar3            NA 5.672290e+35
## CO_scolar4            NA 1.422055e+36
## CO_scolar5            NA 1.288447e+36
## CO_scolar6            NA 1.536440e+36
## CO_scolar7            NA 1.256460e+36
## CO_scolar8            NA 2.374301e+36
## CO_scolar9            NA 1.431565e+36
```

Here, p-value > 0.05, therefore the above features turns out to be insignificant.

We found age_group, age_40, CO_njalc1_group, NW_Q21C1 features as statistically significant in chi-square tests for independence.

The above univariate logistic regression shows that variables age_40, age_group, CO_njalc1, CO_njalc1_group, NW_Q21C12 are statistically significant, and for OAT_3m, CO_ipoud1m, and incar_3m there were statistically significant intercepts only, so will include them for further research. These features will be used in multivariate logistic regression.

We will not include NW_Q21C12 as it has two values (1 - Yes, drink alcohol and 2 - No, do not drink alcohol) which is already covered in CO_njalc1_group as group 0 covers all samples who do not drink achohol and remaining other groups drink alcohol.

Multivaraite logistic regression

CASE 1:

Performing multivariate logistic regression on:
age_40, CO_njalc1_group, OAT_3m, CO_ipoud1m and incar_3m

MODEL 1:

```
Model_multivariate <- glm(DrQ3 ~ age_40 +  CO_njalc1_group + OAT_3m +  CO_ipoud1m + incar_3m, family =

summary(Model_multivariate)
```

```
##
## Call:
## glm(formula = DrQ3 ~ age_40 + CO_njalc1_group + OAT_3m + CO_ipoud1m +
##      incar_3m, family = binomial, data = TrainingSet)
##
## Deviance Residuals:
##     Min      1Q   Median      3Q      Max
## -1.6781  -1.1085  -0.7875   1.1686   1.6294
##
## Coefficients:
##                     Estimate Std. Error z value Pr(>|z|)
## (Intercept)          -1.5287     0.4674  -3.270  0.00107 **
## age_401               0.7159     0.2656   2.695  0.00703 **
## CO_njalc1_group1-10   0.2000     0.2422   0.826  0.40899
## CO_njalc1_group11-20  0.7492     0.3525   2.126  0.03354 *
## CO_njalc1_group21-29  1.1067     0.5238   2.113  0.03463 *
## CO_njalc1_group30     0.0386     0.3541   0.109  0.91319
## OAT_3m1               0.1850     0.2124   0.871  0.38382
## CO_ipoud1m            0.3242     0.2171   1.494  0.13530
## incar_3m1             0.3984     0.3498   1.139  0.25476
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 558.86  on 403  degrees of freedom
## Residual deviance: 539.44  on 395  degrees of freedom
## AIC: 557.44
##
## Number of Fisher Scoring iterations: 4
```

```
tr_prob_1 <- predict(Model_multivariate, newdata = TrainingSet, type = "response")

tr_pred_1 <- ifelse(tr_prob_1 > 0.5, "Yes", "No")
tr_acc_1 <- mean(tr_pred_1 == TrainingSet$DrQ3)

test_prob_1 <- predict(Model_multivariate, newdata = TestingSet, type = "response")
test_pred_1 <- ifelse(test_prob_1 > 0.5, "Yes", "No")
test_acc_1 <- mean(test_pred_1 == TestingSet$DrQ3)


cat("Model 1 Training Accuracy:   ", tr_acc_1, "\n",
    "Model 1 Validation Accuracy: ", test_acc_1, "\n\n"
)
```

```
## Model 1 Training Accuracy:    0.5965347
##  Model 1 Validation Accuracy:  0.5555556
```

```
exp(coef(Model_multivariate))
```

```
##          (Intercept)              age_401  CO_njalc1_group1-10
##            0.2168253            2.0461217            1.2214235
## CO_njalc1_group11-20 CO_njalc1_group21-29     CO_njalc1_group30
##            2.1153993            3.0243477            1.0393565
##              OAT_3m1             CO_ipoud1m             incar_3m1
##            1.2031901            1.3829408            1.4894175
```

```
exp(confint(Model_multivariate))
```

```
## Waiting for profiling to be done...
```

```
##                           2.5 %     97.5 %
## (Intercept)          0.08540945 0.5355447
## age_401              1.22527660 3.4818167
## CO_njalc1_group1-10  0.75947295 1.9655467
## CO_njalc1_group11-20 1.06829660 4.2842831
## CO_njalc1_group21-29 1.12369445 9.0677533
## CO_njalc1_group30    0.51405415 2.0744206
## OAT_3m1              0.79408475 1.8276019
## CO_ipoud1m           0.90521916 2.1223190
## incar_3m1            0.75135815 2.9862947
```

From the output above, the coefficients table shows the beta coefficient estimates and their significance levels. Columns are:

1. Estimate: the intercept (b0) and the beta coefficient estimates associated to each predictor variable

2. Std.Error: the standard error of the coefficient estimates. This represents the accuracy of the coefficients. The larger the standard error, the less confident we are about the estimate.

3. z value: the z-statistic, which is the coefficient estimate (column 2) divided by the standard error of the estimate (column 3)

4. Pr(>|z|): The p-value corresponding to the z-statistic. The smaller the p-value, the more significant the estimate is.

It can be seen that only few predictors are significantly associated to the outcome. These include: age_400, age_401, CO_njalc1_group11-20, CO_njalc1_group21-29, OAT_3m1. To avoid over fitting we will remove the worst performing features i.e. OAT_3m, incar_3m and CO_ipoud1m and try again.

MODEL 2:

```
Model_multivariate_1 <- glm(DrQ3 ~ age_40 +  CO_njalc1_group, family = binomial, data = TrainingSet)

summary(Model_multivariate_1)
```

```
##
## Call:
## glm(formula = DrQ3 ~ age_40 + CO_njalc1_group, family = binomial,
##     data = TrainingSet)
##
## Deviance Residuals:
##     Min      1Q  Median      3Q     Max
## -1.571  -1.111  -0.849   1.242   1.546
##
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)          -0.834992   0.259356  -3.219  0.00128 **
## age_401               0.675688   0.261080   2.588  0.00965 **
## CO_njalc1_group1-10   0.190736   0.239334   0.797  0.42548
## CO_njalc1_group11-20  0.731250   0.349595   2.092  0.03646 *
## CO_njalc1_group21-29  1.048719   0.519888   2.017  0.04367 *
## CO_njalc1_group30     0.008727   0.349181   0.025  0.98006
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 558.86  on 403  degrees of freedom
## Residual deviance: 543.46  on 398  degrees of freedom
## AIC: 555.46
##
## Number of Fisher Scoring iterations: 4
```

```r
tr_prob_1 <- predict(Model_multivariate_1, newdata = TrainingSet, type = "response")

tr_pred_1 <- ifelse(tr_prob_1 > 0.5, "Yes", "No")
tr_acc_1 <- mean(tr_pred_1 == TrainingSet$DrQ3)

test_prob_1 <- predict(Model_multivariate_1, newdata = TestingSet, type = "response")
test_pred_1 <- ifelse(test_prob_1 > 0.5, "Yes", "No")
test_acc_1 <- mean(test_pred_1 == TestingSet$DrQ3)


cat("Model 2 Training Accuracy:   ", tr_acc_1, "\n",
    "Model 2 Validation Accuracy: ", test_acc_1, "\n\n"
)
```

```
## Model 2 Training Accuracy:    0.5742574
##  Model 2 Validation Accuracy:  0.5333333
```

```r
exp(coef(Model_multivariate_1))
```

```
##          (Intercept)               age_401  CO_njalc1_group1-10
##            0.4338778             1.9653853            1.2101397
## CO_njalc1_group11-20  CO_njalc1_group21-29     CO_njalc1_group30
##            2.0776761             2.8539938            1.0087652
```

```
exp(confint(Model_multivariate_1))
```

```
## Waiting for profiling to be done...
```

```
##                            2.5 %    97.5 %
## (Intercept)            0.2574449 0.7141776
## age_401                1.1870463 3.3138344
## CO_njalc1_group1-10    0.7566774 1.9359295
## CO_njalc1_group11-20   1.0551992 4.1844532
## CO_njalc1_group21-29   1.0681986 8.4931285
## CO_njalc1_group30      0.5036125 1.9937690
```

CASE 2:

Performing multivariate logistic regression on:
age_group, CO_njalc1_group, OAT_3m, CO_ipoud1m and incar_3m

MODEL 3

```
Model_multivariate_2 <- glm(DrQ3 ~ age_group +  CO_njalc1_group + OAT_3m +  CO_ipoud1m + incar_3m, famil

summary(Model_multivariate_2)
```

```
##
## Call:
## glm(formula = DrQ3 ~ age_group + CO_njalc1_group + OAT_3m + CO_ipoud1m +
##     incar_3m, family = binomial, data = TrainingSet)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.7989  -1.0812  -0.7917   1.1663   1.6204
##
## Coefficients:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)          -1.85808    0.78177  -2.377   0.0175 *
## age_group30-40        0.27558    0.73241   0.376   0.7067
## age_group40-50        0.67496    0.71160   0.949   0.3429
## age_group50-60        1.18425    0.70824   1.672   0.0945 .
## age_group60-70        1.05276    0.72289   1.456   0.1453
## age_group>=70         0.21249    1.41619   0.150   0.8807
## CO_njalc1_group1-10   0.25253    0.24659   1.024   0.3058
## CO_njalc1_group11-20  0.75818    0.35577   2.131   0.0331 *
## CO_njalc1_group21-29  1.13768    0.53374   2.132   0.0330 *
## CO_njalc1_group30     0.05322    0.35489   0.150   0.8808
## OAT_3m1               0.23314    0.21977   1.061   0.2888
## CO_ipoud1m            0.34996    0.21906   1.598   0.1101
## incar_3m1             0.41847    0.35380   1.183   0.2369
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 558.86  on 403  degrees of freedom
## Residual deviance: 535.06  on 391  degrees of freedom
## AIC: 561.06
##
## Number of Fisher Scoring iterations: 4
```

```
tr_prob_1 <- predict(Model_multivariate_2, newdata = TrainingSet, type = "response")

tr_pred_1 <- ifelse(tr_prob_1 > 0.5, "Yes", "No")
tr_acc_1 <- mean(tr_pred_1 == TrainingSet$DrQ3)

test_prob_1 <- predict(Model_multivariate_2, newdata = TestingSet, type = "response")
test_pred_1 <- ifelse(test_prob_1 > 0.5, "Yes", "No")
test_acc_1 <- mean(test_pred_1 == TestingSet$DrQ3)


cat("Model 3 Training Accuracy:   ", tr_acc_1, "\n",
    "Model 3 Validation Accuracy: ", test_acc_1, "\n\n"
)
```

```
## Model 3 Training Accuracy:    0.5990099
##  Model 3 Validation Accuracy:  0.5037037
```

```
exp(coef(Model_multivariate_2))
```

```
##           (Intercept)         age_group30-40         age_group40-50
##             0.1559724              1.3172935              1.9639528
##        age_group50-60         age_group60-70          age_group>=70
##             3.2682389              2.8655373              1.2367504
##  CO_njalc1_group1-10 CO_njalc1_group11-20 CO_njalc1_group21-29
##             1.2872795              2.1343946              3.1195134
##     CO_njalc1_group30               OAT_3m1            CO_ipoud1m
##             1.0546575              1.2625570              1.4190044
##             incar_3m1
##             1.5196342
```

```
exp(confint(Model_multivariate_2))
```

```
## Waiting for profiling to be done...
```

```
##                           2.5 %     97.5 %
## (Intercept)           0.02899420  0.6719509
## age_group30-40        0.33804907  6.5405615
## age_group40-50        0.52783809  9.4518684
## age_group50-60        0.88532363 15.6533394
## age_group60-70        0.75124970 14.0233196
## age_group>=70         0.04602126 19.3645103
## CO_njalc1_group1-10   0.79397696  2.0903822
## CO_njalc1_group11-20  1.07096843  4.3503095
```

```
## CO_njalc1_group21-29 1.13655377  9.5268990
## CO_njalc1_group30     0.52089591  2.1084739
## OAT_3m1               0.82178747  1.9470948
## CO_ipoud1m            0.92538205  2.1865627
## incar_3m1             0.76063260  3.0700653
```

Again we see only few predictors are significantly associated to the outcome. These include: age_group0-30, age_group40-50, age_group50-60, age_group60-70, CO_njalc1_group11-20, CO_njalc1_group21-29 To avoid over fitting we will remove the worst performing features i.e. incar_3m, OAT_3m and CO_ipoud1m and try again.

MODEL 4:

```
Model_multivariate_3 <- glm(DrQ3 ~ age_group +  CO_njalc1_group, family = binomial, data = TrainingSet)
```

```
summary(Model_multivariate_3)
```

```
##
## Call:
## glm(formula = DrQ3 ~ age_group + CO_njalc1_group, family = binomial,
##     data = TrainingSet)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.6723  -1.1209  -0.8618   1.1632   1.7033
##
## Coefficients:
##                    Estimate Std. Error z value Pr(>|z|)
## (Intercept)        -1.18339    0.69716  -1.697   0.0896 .
## age_group30-40      0.38413    0.73255   0.524   0.6000
## age_group40-50      0.76855    0.71301   1.078   0.2811
## age_group50-60      1.21686    0.71091   1.712   0.0870 .
## age_group60-70      1.04892    0.72378   1.449   0.1473
## age_group>=70       0.25591    1.41113   0.181   0.8561
## CO_njalc1_group1-10  0.23433    0.24313   0.964   0.3351
## CO_njalc1_group11-20 0.72656    0.35242   2.062   0.0392 *
## CO_njalc1_group21-29 1.08110    0.52767   2.049   0.0405 *
## CO_njalc1_group30    0.01437    0.35072   0.041   0.9673
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 558.86  on 403  degrees of freedom
## Residual deviance: 539.82  on 394  degrees of freedom
## AIC: 559.82
##
## Number of Fisher Scoring iterations: 4
```

```r
probabilities <- predict(Model_multivariate_3,
                         newdata = TestingSet,
                         type = "response")


tr_prob_1 <- predict(Model_multivariate_3,
                     newdata = TrainingSet,
                     type = "response")

tr_pred_1 <- ifelse(tr_prob_1 > 0.5, "Yes", "No")
tr_acc_1 <- mean(tr_pred_1 == TrainingSet$DrQ3)

test_prob_1 <- predict(Model_multivariate_3, newdata = TestingSet, type = "response")
test_pred_1 <- ifelse(test_prob_1 > 0.5, "Yes", "No")
test_acc_1 <- mean(test_pred_1 == TestingSet$DrQ3)


cat("Model 4 Training Accuracy:   ", tr_acc_1, "\n",
    "Model 4 Validation Accuracy: ", test_acc_1, "\n\n"
)
```

```
## Model 4 Training Accuracy:    0.5816832
##  Model 4 Validation Accuracy:  0.4666667
```

```r
exp(coef(Model_multivariate_3))
```

```
##         (Intercept)      age_group30-40      age_group40-50
##           0.3062401           1.4683326           2.1566338
##      age_group50-60      age_group60-70       age_group>=70
##           3.3765771           2.8545553           1.2916400
##  CO_njalc1_group1-10 CO_njalc1_group11-20 CO_njalc1_group21-29
##           1.2640566           2.0679545           2.9479310
##    CO_njalc1_group30
##           1.0144743
```

```r
exp(confint(Model_multivariate_3))
```

```
## Waiting for profiling to be done...
```

```
##                          2.5 %    97.5 %
## (Intercept)           0.06483814  1.100013
## age_group30-40        0.37738445  7.302572
## age_group40-50        0.57892299 10.416612
## age_group50-60        0.91126712 16.264912
## age_group60-70        0.74812190 14.006038
## age_group>=70         0.04838874 20.040689
## CO_njalc1_group1-10   0.78480241  2.038142
## CO_njalc1_group11-20  1.04440109  4.187426
## CO_njalc1_group21-29  1.08610041  8.895923
## CO_njalc1_group30     0.50498502  2.011218
```

Here, age_group0-30, age_group40-50, age_group50-60, age_group60-70, age_group>=70, CO_njalc1_group11-20, CO_njalc1_group21-29, OAT_3m are statistically significant as there p < 0.1.

Model 1 Training Accuracy: 0.5965347 Model 1 Validation Accuracy: 0.5555556

Model 2 Training Accuracy: 0.5742574 Model 2 Validation Accuracy: 0.5333333

Model 3 Training Accuracy: 0.5990099 Model 3 Validation Accuracy: 0.5037037

Model 4 Training Accuracy: 0.5816832 Model 4 Validation Accuracy: 0.4666667

Model 2 Interpretation

The odds of the patient having chronic pain decreased by e^(0.834992) = 0.4338778 times for each additional patient under age of 40, whilst controlling for every other variable.

The odds of the patient having chronic pain increased by e^(0.675688) = 1.9653853 times for each additional patient over age of 40, whilst controlling for every other variable.

The odds of the patient having chronic pain increased by e^(0.731250) = 2.0776761 times for each additional patient CO_njalc1_group11-20, whilst controlling for every other variable.

The odds of the patient having chronic pain increased by e^(1.048719) = 2.8539938 times for each additional patient in CO_njalc1_group21-29, whilst controlling for every other variable.

Model 4 Interpretation

The odds of the patient having chronic pain decreased by e^(1.18339) = 0.3062401 times for each additional patient under age of 30, whilst controlling for every other variable.

The odds of the patient having chronic pain increased by e^(1.21686) = 3.3765771 times for each additional patient in age group of 50-60, whilst controlling for every other variable.

The odds of the patient having chronic pain increased by e^(0.72656) = 2.0679545 times for each additional patient CO_njalc1_group11-20, whilst controlling for every other variable.

The odds of the patient having chronic pain increased by e^(1.08110) = 2.9479310 times for each additional patient in CO_njalc1_group21-29, whilst controlling for every other variable.

Model 1 has the highest accuracy on the validation set of 55.5 percentage but it also has some insignificant features. Therefore, we will choose Model 2 which has only significant features with the second highest accuracy on validation dataset as 53.3 percentage.

Not considering model 4 as it has the lowest accuracy of 46.6 percentage.