Income Analysis OBPD

```
#list of files in dir.
contigency.table.dir <- "/Users/omachowda/Google Drive/StatCom 3/OBPD/Contingency table
s/Income folder"
cont.tables.list <- list.files(contigency.table.dir)
#take only .csv files
cont.tables.list <- cont.tables.list[grep(cont.tables.list,pattern = ".csv")]
#get table names
table.name <- gsub(cont.tables.list,pattern = ".csv",replacement = "");
#assign table to name
for(i in 1:length(table.name)){
  tab <- read.csv(cont.tables.list[i]);
  tab <- tab[-c(nrow(tab)),-c(1,ncol(tab))]; #get rid of row and column sums and first 2
 rows
  assign(table.name[i],tab)
}

SIS <- dget("/Users/omachowda/Google Drive/StatCom 3/OBPD/Robin/Sequential Importance Sa
mpling Function.R") #import SIS function
```

```
## Warning: package 'Rmpfr' was built under R version 3.2.5
```

```
## Loading required package: gmp
```

```
## Warning: package 'gmp' was built under R version 3.2.5
```

```
##
## Attaching package: 'gmp'
```

```
## The following objects are masked from 'package:base':
##
##     %*%, apply, crossprod, matrix, tcrossprod
```

```
## C code of R package 'Rmpfr': GMP using 64 bits per limb
```

```
##
## Attaching package: 'Rmpfr'
```

```
## The following objects are masked from 'package:stats':
##
##     dbinom, dnorm, dpois, pnorm
```

```
## The following objects are masked from 'package:base':
##
##     cbind, pmax, pmin, rbind
```

```
#facilities to age output
```

Example of how table is simulated

```
SIS(get(table.name[1]),dist = "hyper")$upper
```

```
##      [,1] [,2] [,3] [,4]
## [1,]   15    7    4    2
## [2,]   58   34   16    9
## [3,]   61   25   17   29
```

```
SIS(get(table.name[1]),dist = "hyper")$lower
```

```
##      [,1] [,2] [,3] [,4]
## [1,]    0    0    0    0
## [2,]    0    0    0    0
## [3,]    0    0    0    0
```

```
SIS(get(table.name[1]),dist = "hyper")$X
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]    6    3    3    3    0
## [2,]   26   16    8    4    4
## [3,]   34   17   10   13   53
## [4,]   27   10    5   13   60
```

```
SIS(get(table.name[1]),dist = "hyper")$p.value
```

```
## 1 'mpfr' number of precision  128   bits
## [1] 0
```

```
#reference character items as objects
SIS.results <- lapply(sapply(table.name,get),function(x) SIS(x,dist = "hyper")$p.value)
```

SIS P-values for all questions based on income

```
SIS.results
```

```
## $`income and facilities`
## 1 'mpfr' number of precision  128   bits
## [1] 0
##
## $`income and oversee sports core`
## 1 'mpfr' number of precision  128   bits
## [1] 1
##
## $`income and park areas analysis complete`
## 1 'mpfr' number of precision  128   bits
## [1] 0.10631787240688266260983289763525813178
##
## $`income and recreation programs`
## 1 'mpfr' number of precision  128   bits
## [1] 1
##
## $`income and satisfaction with district`
## 1 'mpfr' number of precision  128   bits
## [1] 1
##
## $`income and satisfaction with facilities`
## 1 'mpfr' number of precision  128   bits
## [1] 1
##
## $`income and satisfaction with maintenence`
## 1 'mpfr' number of precision  128   bits
## [1] 0
##
## $`income and satisfaction with outdoor amenities and parks`
## 1 'mpfr' number of precision  128   bits
## [1] 1
##
## $`income and satisfaction with programs`
## 1 'mpfr' number of precision  128   bits
## [1] 1
##
## $`income and satisfaction with staff`
## 1 'mpfr' number of precision  128   bits
## [1] 0.95678479101662067830760735458283456550 49
##
## $`income and special events`
## 1 'mpfr' number of precision  128   bits
## [1] 0
##
## $`income and work with sports core`
## 1 'mpfr' number of precision  128   bits
## [1] 0
```

```
##### M^2 Test ######

corr_calc <- function(df){
  columns = length(df)
  new_df = data.frame()
  v1= c()
  v2= c()
  z = 0
  for (i in 1:4){
    for (j in 1:columns){
      new_df <- rbind(new_df, c(i,j, df[i,j]))
      z = z + df[i,j]
    }
  }

  for (i in 1:(4*columns)){
    v1 = append(v1, c(rep(new_df[i,1],new_df[i,3])))
  }

  for (i in 1:(4*columns)){
    v2 = append(v2, c(rep(new_df[i,2],new_df[i,3])))
  }

  fit = cor(x = v1,y = v2)
  M = (z-1)*(fit)
  p = pchisq(M,12)
  return (p)
}

resulted = lapply(sapply(table.name,get),FUN= corr_calc)
```

M^2 p-values for all questions based on income

```
resulted
```

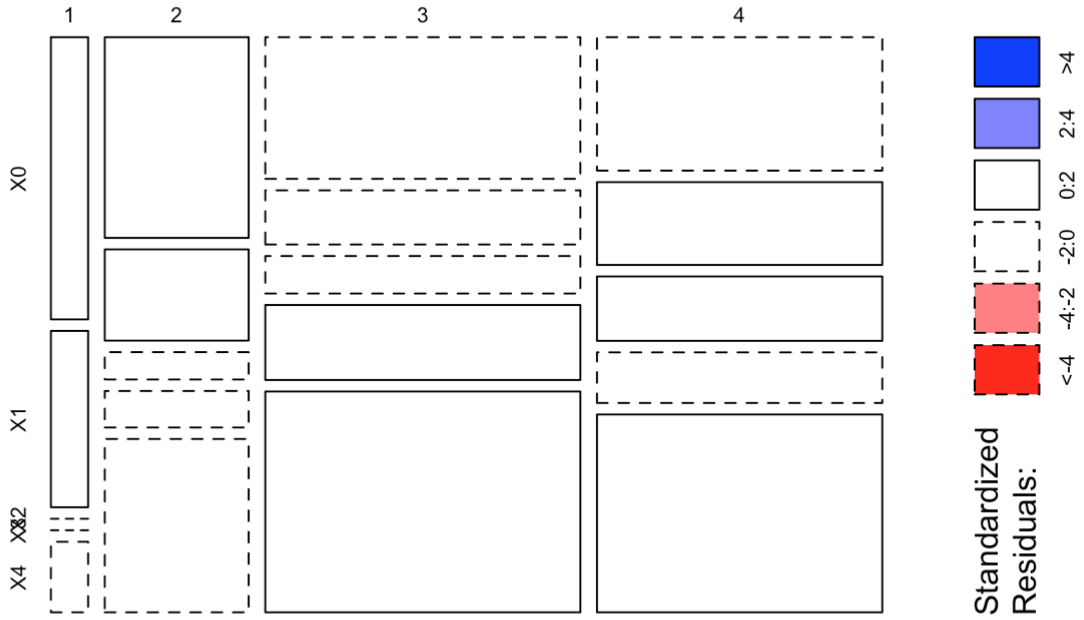
```
## $`income and facilities`
## [1] 0.9999831
##
## $`income and oversee sports core`
## [1] 0.9999992
##
## $`income and park areas analysis complete`
## [1] 0.9999997
##
## $`income and recreation programs`
## [1] 0.9987597
##
## $`income and satisfaction with district`
## [1] 0.831943
##
## $`income and satisfaction with facilities`
## [1] 0.9999978
##
## $`income and satisfaction with maintenence`
## [1] 0.03205391
##
## $`income and satisfaction with outdoor amenities and parks`
## [1] 0.9999755
##
## $`income and satisfaction with programs`
## [1] 0
##
## $`income and satisfaction with staff`
## [1] 0.9439812
##
## $`income and special events`
## [1] 0.06394473
##
## $`income and work with sports core`
## [1] 0.9216534
```

```
par(mfrow=c(1,1))
```

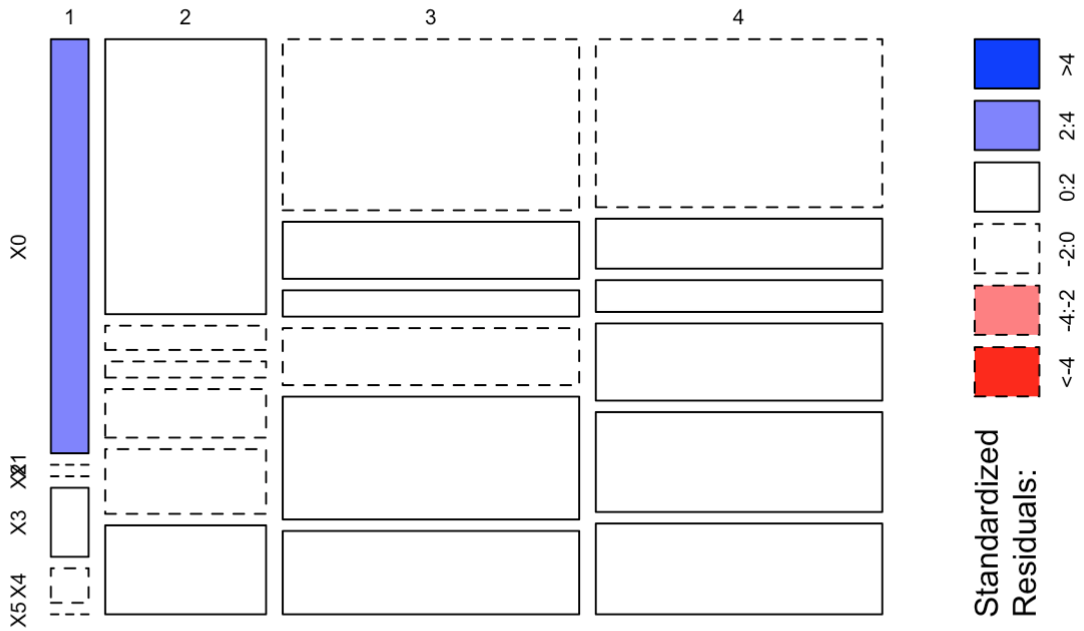
Mosaic Plots for questions based on income

```
for( i in 1:12){
  mosaicplot(x = get(table.name[i]), shade = TRUE,color = TRUE, main= table.name[i])
}
```

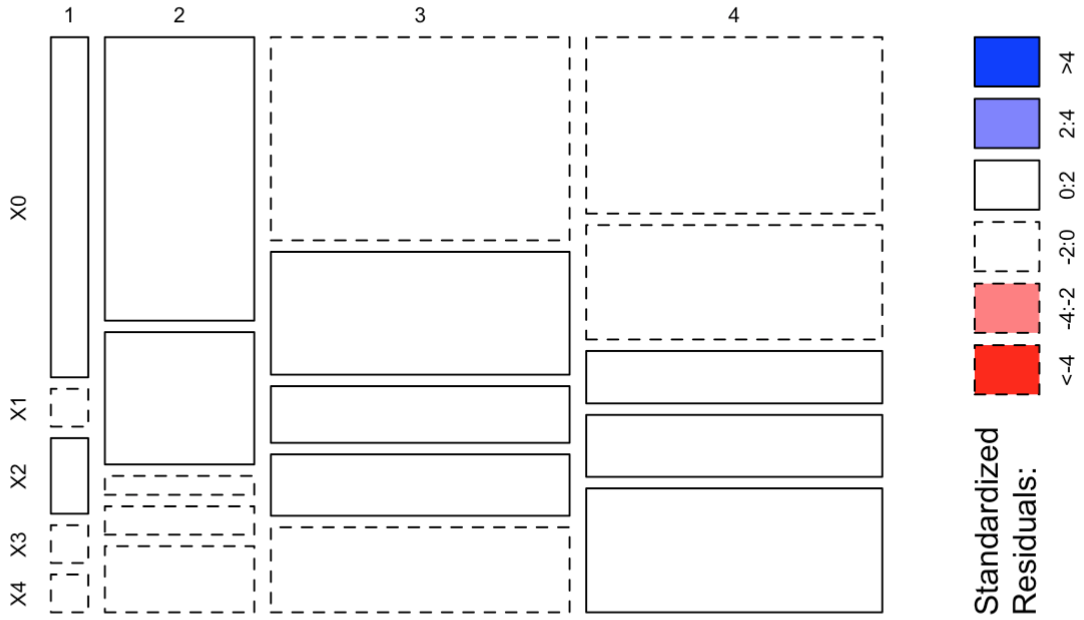
income and facilities



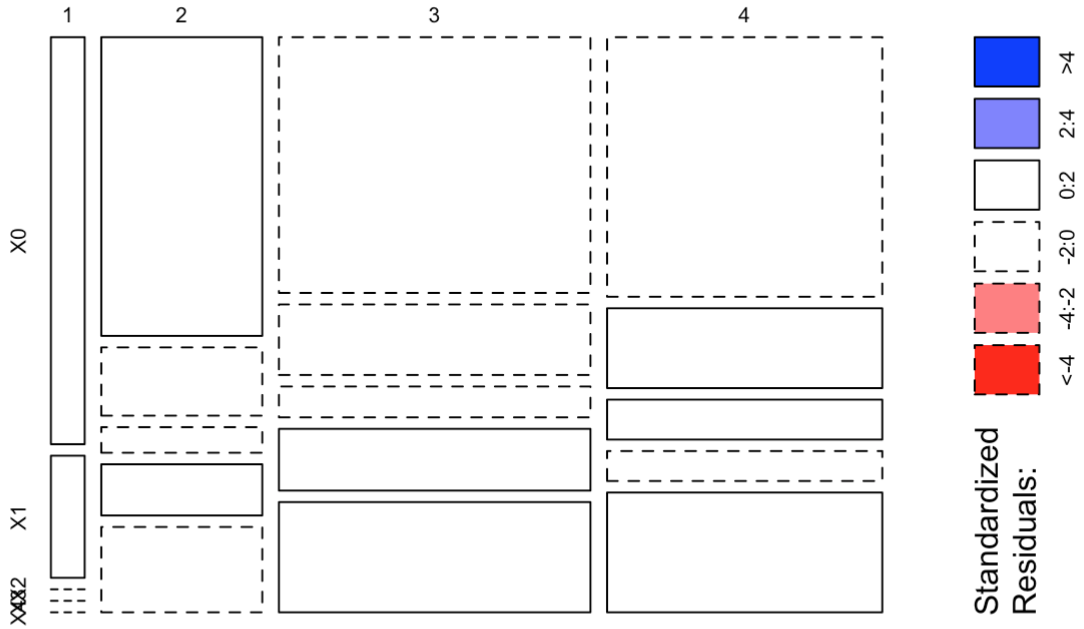
income and oversee sports core



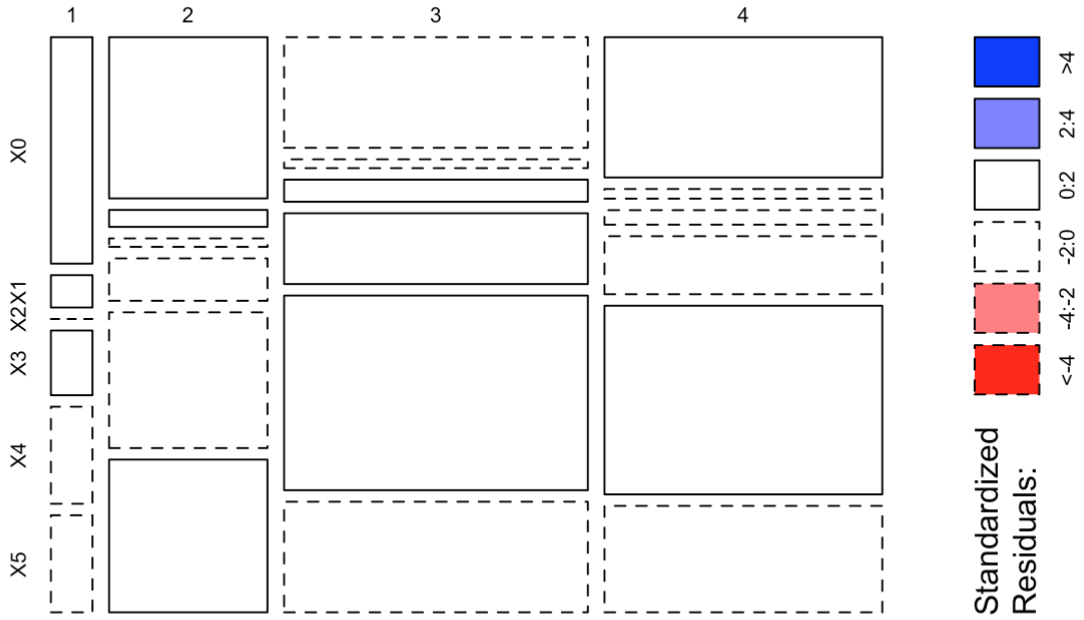
income and park areas analysis complete



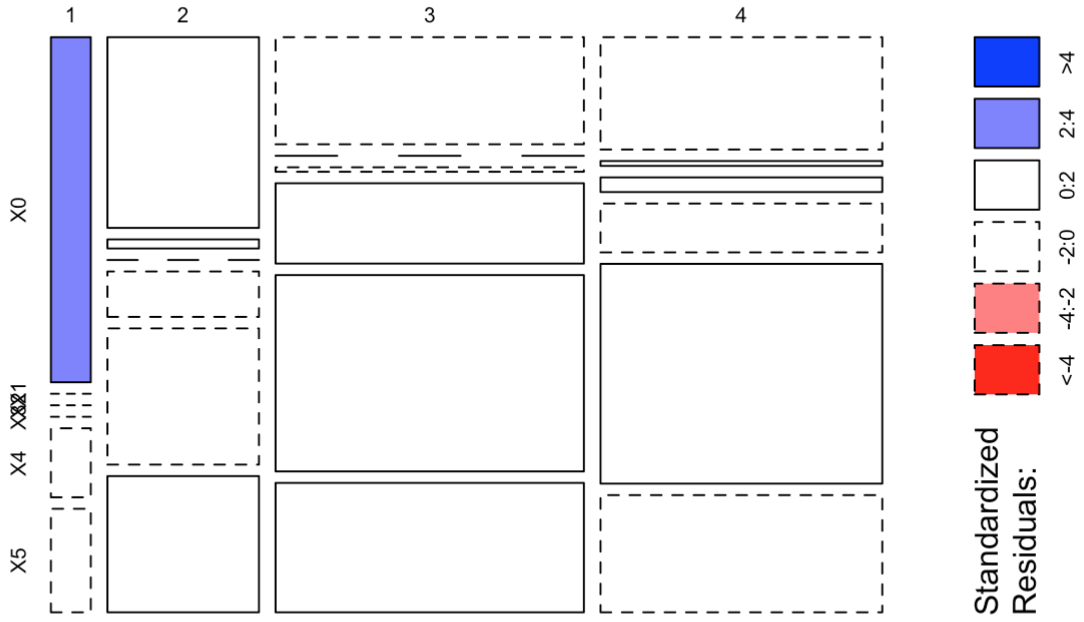
income and recreation programs



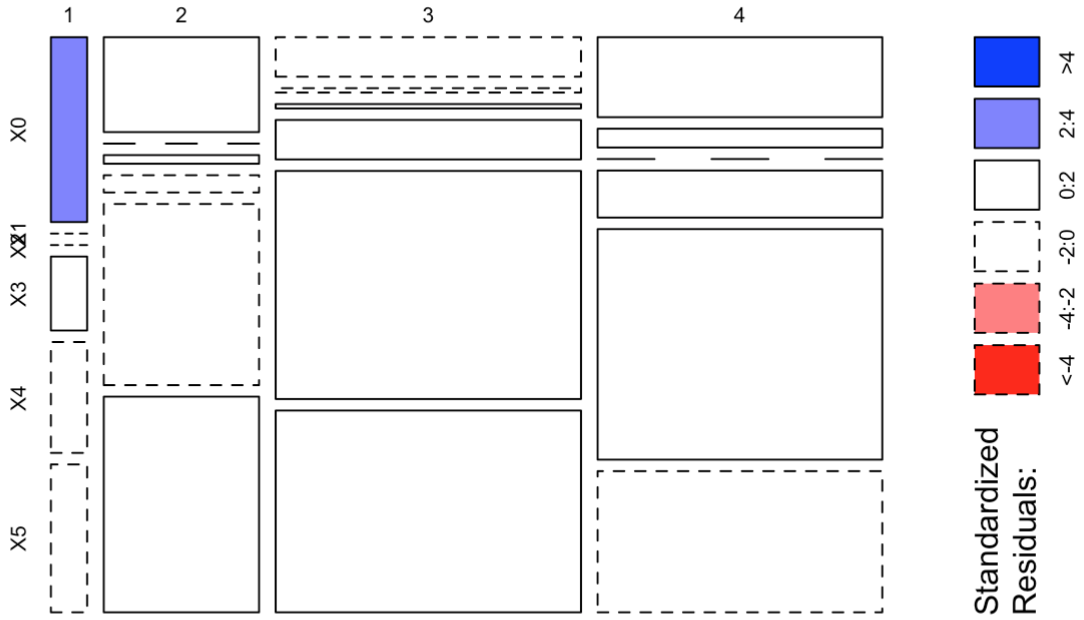
income and satisfaction with district



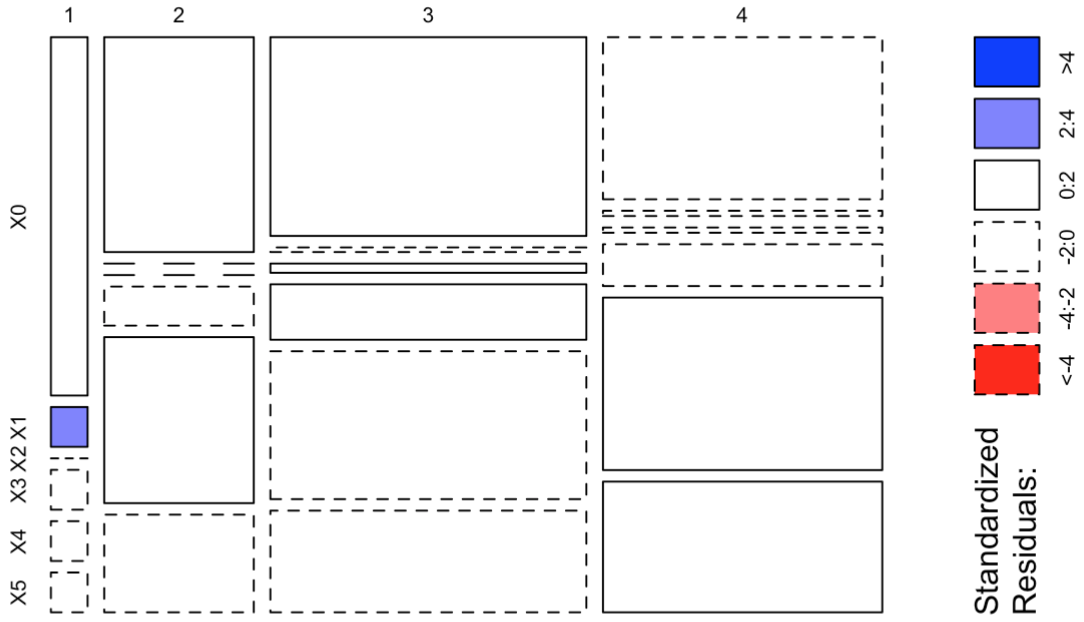
income and satisfaction with facilities



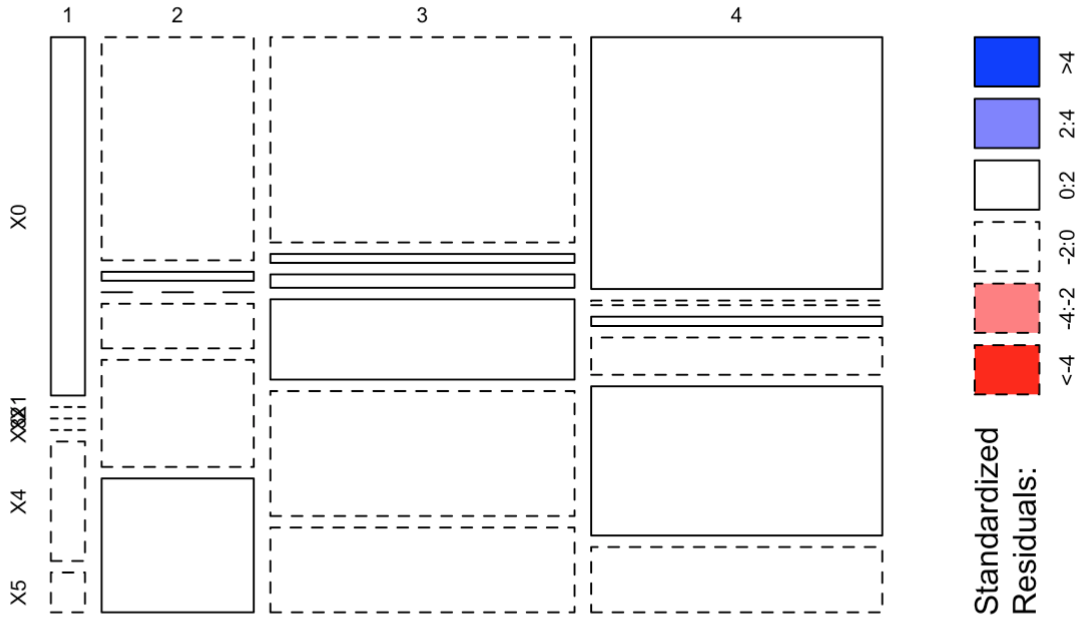
income and satisfaction with maintenence



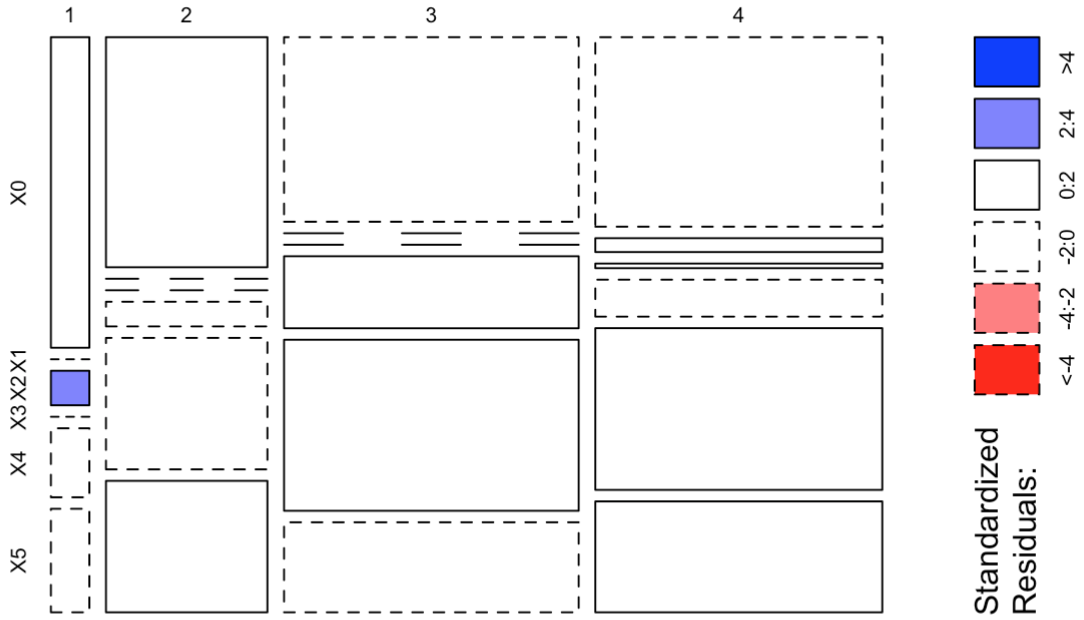
income and satisfaction with outdoor amenities and parks



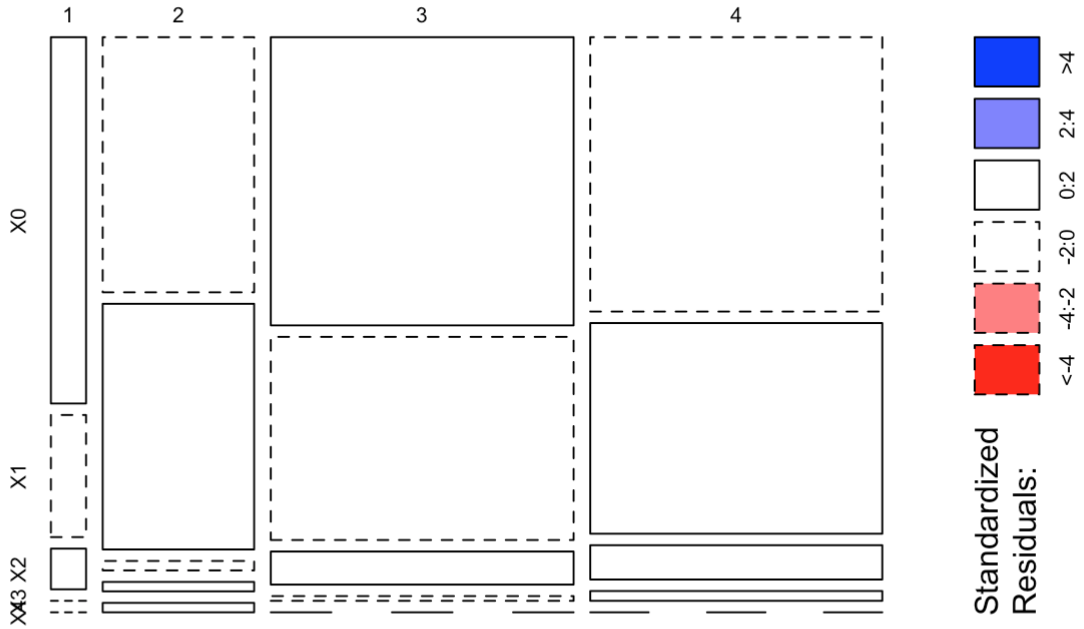
income and satisfaction with programs



income and satisfaction with staff



income and special events



income and work with sports core