Overview of "teengamb" Dataset from "faraway" Package

Julian Hatwell

January 3, 2016

This document provides a brief overview of the teengamb dataset in the faraway R package.

```
##       sex            status          income           verbal      
##  Min.   :0.0000   Min.   :18.00   Min.   : 0.600   Min.   : 1.00  
##  1st Qu.:0.0000   1st Qu.:28.00   1st Qu.: 2.000   1st Qu.: 6.00  
##  Median :0.0000   Median :43.00   Median : 3.250   Median : 7.00  
##  Mean   :0.4043   Mean   :45.23   Mean   : 4.642   Mean   : 6.66  
##  3rd Qu.:1.0000   3rd Qu.:61.50   3rd Qu.: 6.210   3rd Qu.: 8.00  
##  Max.   :1.0000   Max.   :75.00   Max.   :15.000   Max.   :10.00  
##      gamble       
##  Min.   :  0.0    
##  1st Qu.:  1.1    
##  Median :  6.0    
##  Mean   : 19.3    
##  3rd Qu.: 19.4    
##  Max.   :156.0
```

From the summary, and the associated help (not shown), the following observations can be made:

The dataframe contains 47 rows and 5 columns.

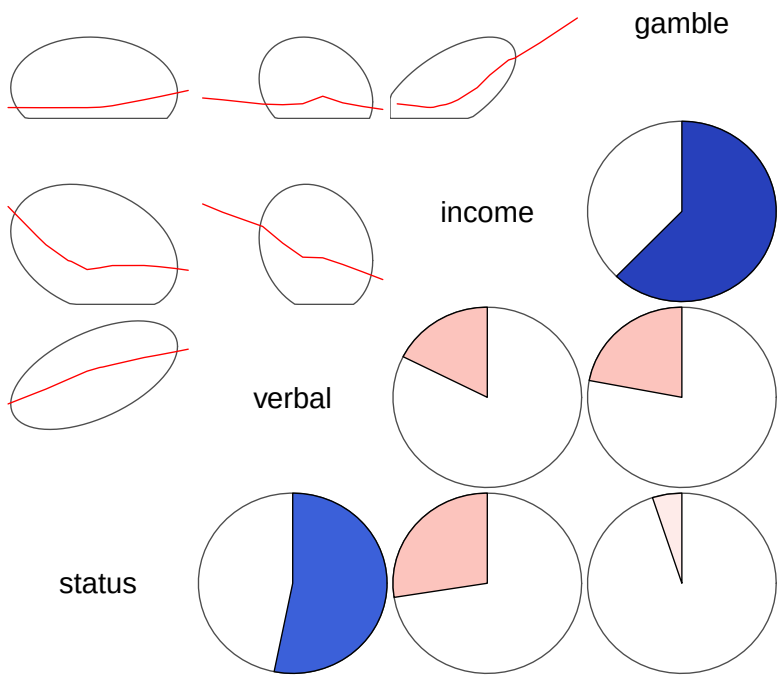

Figure 1: Correlogram

```
##
##  Welch Two Sample t-test
##
## data:  gamble by sex
## t = 3.6227, df = 28.503, p-value = 0.001123
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  11.27085 40.54758
## sample estimates:
##   mean in group male mean in group female
##            29.775000             3.865789
```

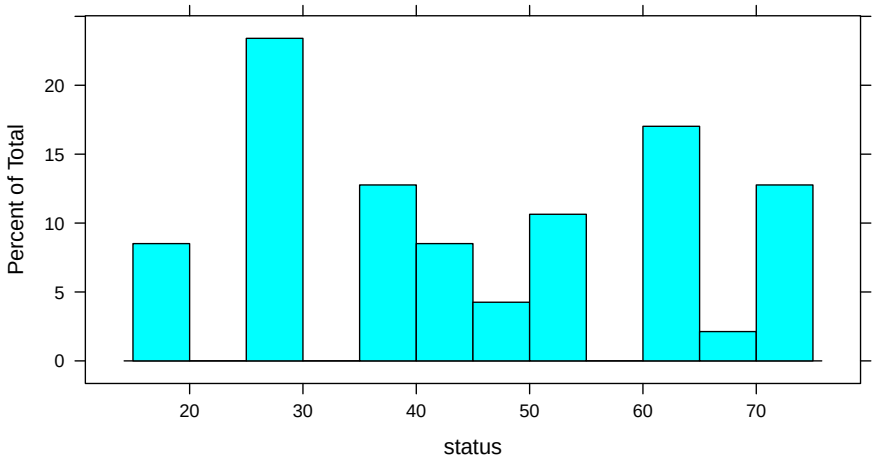

Figure 2: Histogram of the status variable

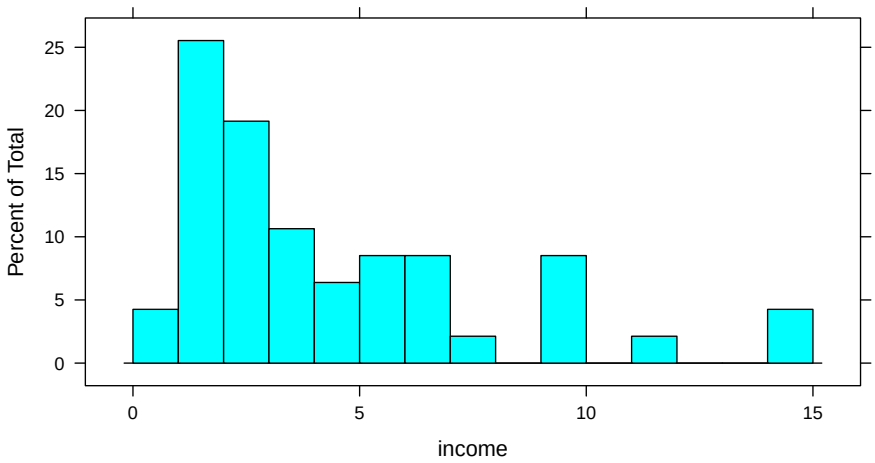

Figure 3: Histogram of the income variable

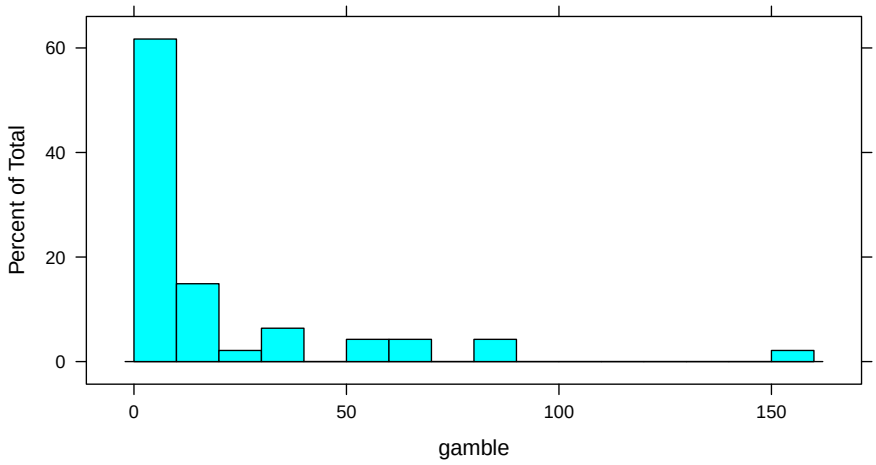

Figure 4: Histogram of the gamble variable

```
# Sex has been coded as integer values 0
# and 1. 40% of the observations are
# female.

# Parents socioeconomic status is likely
# to be a percentage ranging between 18
# and 75.

# Verbal may be an indicator of education
# with discrete levels 1 - 10 (maximum of
# possible 12 indicated in the help).

# There don't appear to be any missing
# values. Where zeros appear, they seem
# to be reasonable values.

# set factors correctly
df <- mutate(df, sex = factor(sex, labels = c("male",
    "female")), verbal = factor(verbal, levels = 1:12))
```

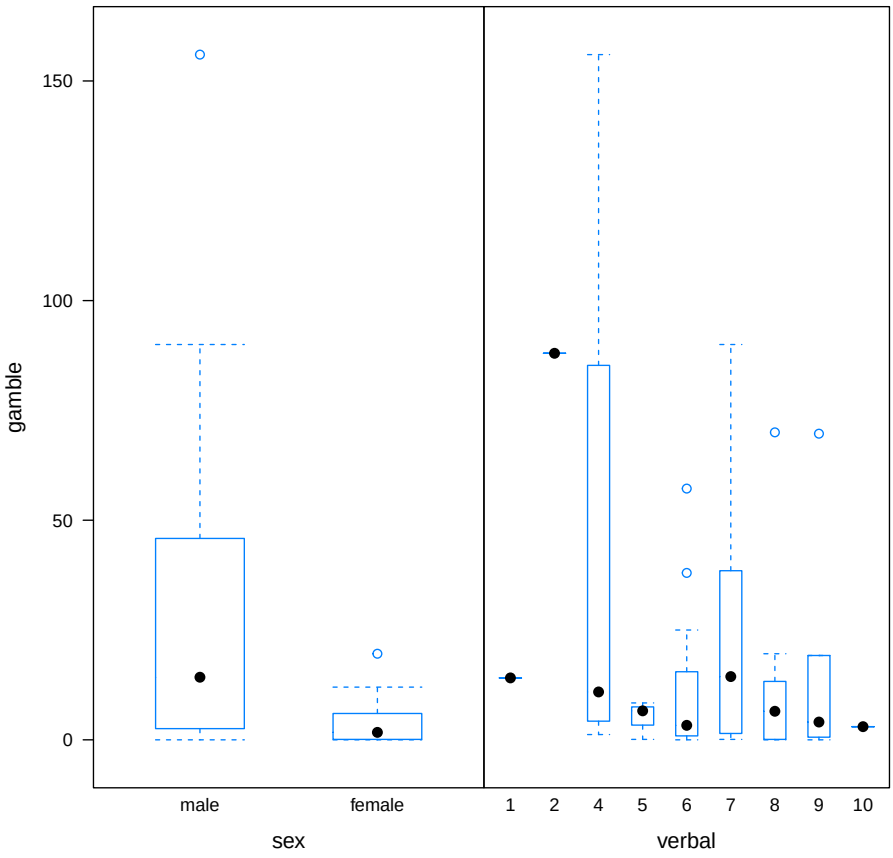

Figure 5: Boxplot of the dependent variable gamble by each factor variable

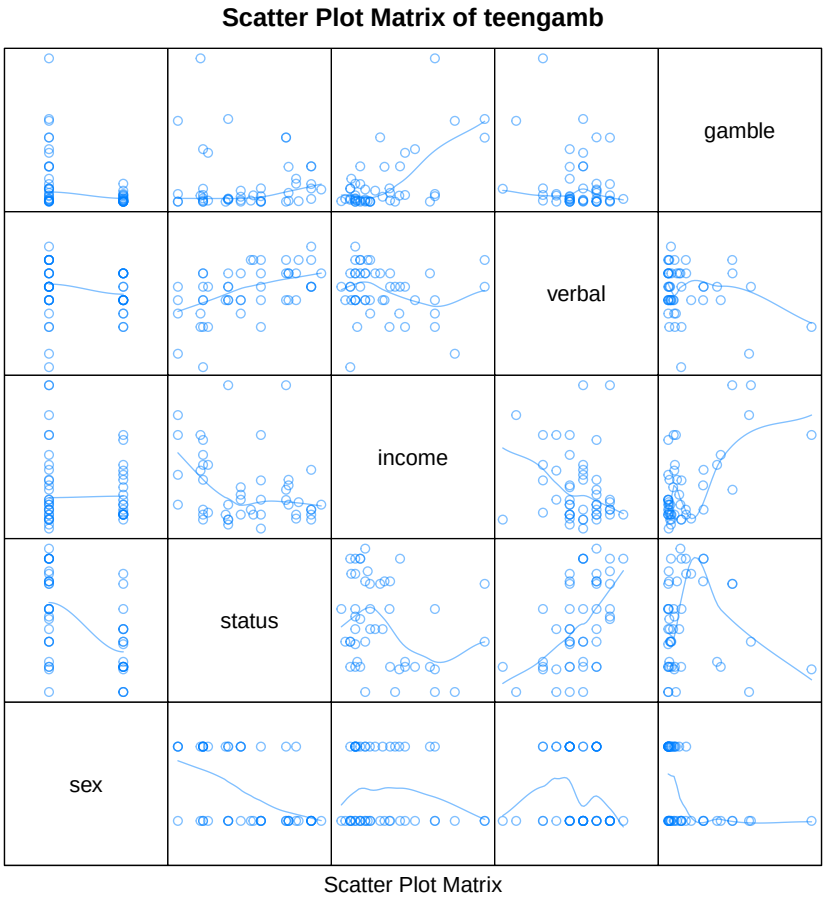

Figure 6: multi-variate comparisons