Classification

Preventing excessive drug consumption among young adults is a social responsibility. Publicly sponsored Drug Prevention Programs (DPPs) are one possible preventive measure against drug addiction. However, DPPs are relatively expensive. Budget limits of public institutions restrict the number of people who can participate in DPPs. Many DPPs are targeted at youths with a high risk to become drug addicted. The objective of this exercise is to predict the risk of young adults to become drug addicted based on socio-economic characteristics. These predictions could be used to assign youths to DPPs.

The file `drugs.RData` contains information about drug consumption and socio-economic characteristics for 10'001 young adults in the United States. Table 1 describes the observable variables.

Table 1: Variable descriptions.

Variable	**Description**
Gender	String for female or male
Age	Age categories: 16-17 years, 18-19 years, and 20-24 years
Ethnicity	Ethnicity categories: Afro-American, Hispanic, White, and Other
Employment	String for at least once employed or unemployed during the previous year
Earning	Earnings categories (measured in preceding year): <1k USD, 1-5k USD, 5-10k USD, and >10k USD
Partner_Status	String for single or partner
Delinquency	String for arrested or not arrested in the previous year
Soft_Drug	Logical for soft drug consumption
Hard_Drug	Logical for hard drug consumption

Group Home Assignment (max. 3 points)

The mandatory group home assignment has to be submitted before 12:00 o'clock prior PC-session 5. A sheet with the answers to the three questions below as well as the file with the R code has to be submitted via Canvas.

Download the data set `drugs.RData` from Canvas. Load the data into R. Install and load the packages `rpart` and `rpart.plot`.

1. How large is the share of males who consume soft drugs (in percent)? (1 point)
2. How large is the difference between the share of male and female hard drug consumers (in percentage points)? (1 point)
3. Report the shares of young adults who consume soft drugs for each age group (16-17 years, 18-19 years, and 20-24 years). Is soft drug consumption increasing or decreasing with age? (1 point)