```
# Dieses Script liest eine CSV-Datendatei in GNU R ein.
# Beim Einlesen werden für alle Variablen Beschriftungen (comment) angelegt.
# Die Beschriftungen für Werte wird ebenfalls als Attribute (attr) abgelegt.

ds_file = file.choose()
# setwd("./")
# ds_file = "rdata_iGEMsurvey_2026-07-28_11-46.csv"

options(encoding = "UTF-8")
ds = read.delim(
  file=ds_file, encoding="UTF-8", fileEncoding="UTF-8",
  header = FALSE, sep = "\t", quote = "\"",
  dec = ".", row.names = NULL,
  col.names = c(
    "CASE","SERIAL","REF","QUESTNNR","MODE","STARTED","AU02_01","AU02_02","AU02_03",
    "AU02_04","AU02_05","AU02_06","AU02_07","AU03","AU04","DE03","DE04","DE06",
    "DE07_01","DE07_02","DE08","WA01","WA02_01","WA02_02","WA02_03","WA02_04",
    "WA02_05","WA02_06","WA02_07","WA02_08","WA02_09","WA02_10","WA02_11","WA02_12",
    "WA03","TIME001","TIME002","TIME003","TIME004","TIME_SUM","MAILSENT","LASTDATA",
    "STATUS","FINISHED","Q_VIEWER","LASTPAGE","MAXPAGE","MISSING","MISSREL",
    "TIME_RSI"
  ),
  as.is = TRUE,
  colClasses = c(
    CASE="numeric", SERIAL="character", REF="character", QUESTNNR="character",
    MODE="factor", STARTED="POSIXct", AU02_01="numeric", AU02_02="numeric",
    AU02_03="numeric", AU02_04="numeric", AU02_05="numeric", AU02_06="numeric",
    AU02_07="numeric", AU03="numeric", AU04="numeric", DE03="numeric",
    DE04="numeric", DE06="numeric", DE07_01="character", DE07_02="character",
    DE08="numeric", WA01="numeric", WA02_01="numeric", WA02_02="numeric",
    WA02_03="numeric", WA02_04="numeric", WA02_05="numeric", WA02_06="numeric",
    WA02_07="numeric", WA02_08="numeric", WA02_09="numeric", WA02_10="numeric",
    WA02_11="numeric", WA02_12="numeric", WA03="numeric", TIME001="integer",
    TIME002="integer", TIME003="integer", TIME004="integer", TIME_SUM="integer",
    MAILSENT="POSIXct", LASTDATA="POSIXct", STATUS="character",
    FINISHED="logical", Q_VIEWER="logical", LASTPAGE="numeric",
    MAXPAGE="numeric", MISSING="numeric", MISSREL="numeric", TIME_RSI="numeric"
  ),
  skip = 1,
  check.names = TRUE, fill = TRUE,
  strip.white = FALSE, blank.lines.skip = TRUE,
  comment.char = "",
  na.strings = ""
)

row.names(ds) = ds$CASE

rm(ds_file)

attr(ds, "project") = "iGEMsurvey"
attr(ds, "description") = "Survey on the distribution of tasks in scientific/voluntary teams"
attr(ds, "date") = "2026-07-28 11:46:12"
attr(ds, "server") = "https://soscisurvey.scc.kit.edu"

# Variable und Value Labels
ds$AU03 = factor(ds$AU03,
    levels=c("1","2","3","4","5","-9"),
    labels=c("Never","Rarely","Sometimes","Often","Very often","[NA] Not answered"),
    ordered=FALSE)
ds$AU04 = factor(ds$AU04,
    levels=c("1","2","3","4","5","-9"),
    labels=c("Never","Rarely","Sometimes","Often","Very often","[NA] Not answered"),
    ordered=FALSE)
ds$DE03 = factor(ds$DE03,
    levels=c("1","2","3","4","-9"),
    labels=c("Male","Female","Non-binary/ other","Prefer not to say","[NA] Not answered"),
    ordered=FALSE)
ds$DE04 = factor(ds$DE04,
    levels=c("1","2","3","4","5","6","7","8","-9"),
    labels=c("Undergraduate student","Graduate student","PhD student","Postdoc","Professor","Technical Assistant","Employee","Other","[NA] Not answered"),
    ordered=FALSE)
ds$DE06 = factor(ds$DE06,
    levels=c("1","2","3","4","-9"),
    labels=c("iGEM team","University student group","Research group","Other","[NA] Not answered"),
    ordered=FALSE)
ds$DE08 = factor(ds$DE08,

levels=c("1","2","3","4","5","6","7","8","9","10","11","12","13","14","15","16","17","18","19","20","21","22","23","24","25","26","27","28","29","30","31","32","33","34","35","36","37","38","39","40","41","42","43","44","45","46","47","48","49","50","51","52","53","54","55","56","57","58","59","60","61","62","63","64","65","66","67","68","69","70","71","72","73","74","75","76","77","78","79","80","81","82","83","84","85","86","87","88","89","90","91","92","93","94","95","96","97","98","99","-9"),

labels=c("1","2","3","4","5","6","7","8","9","10","11","12","13","14","15","16","17","18","19","20","21","22","23","24","25","26","27","28","29","30","31","32","33","34","35","36","37","38","39","40","41","42","43","44","45","46","47","48","49","50","51","52","53","54","55","56","57","58","59","60","61","62","63","64","65","66","67","68","69","70","71","72","73","74","75","76","77","78","79","80","81","82","83","84","85","86","87","88","89","90","91","92","93","94","95","96","97","98","99","[NA] Not answered"),
    ordered=FALSE)
ds$WA01 = factor(ds$WA01,
    levels=c("1","2","3","4","5","-9"),
    labels=c("Never","Rarely","Sometimes","Often","Very often","[NA] Not answered"),
    ordered=FALSE)
ds$WA03 = factor(ds$WA03,
    levels=c("1","2","3","4","5","-9"),
    labels=c("Mostly women","About equal numbers of women and men","Mostly men","Only women","Only men","[NA] Not answered"),
    ordered=FALSE)
attr(ds$AU02_01,"1") = "Never"
attr(ds$AU02_01,"2") = "Rarely"
attr(ds$AU02_01,"3") = "Sometimes"
attr(ds$AU02_01,"4") = "Often"
attr(ds$AU02_01,"5") = "Very Often"
attr(ds$AU02_02,"1") = "Never"
attr(ds$AU02_02,"2") = "Rarely"
attr(ds$AU02_02,"3") = "Sometimes"
attr(ds$AU02_02,"4") = "Often"
attr(ds$AU02_02,"5") = "Very Often"
attr(ds$AU02_03,"1") = "Never"
attr(ds$AU02_03,"2") = "Rarely"
attr(ds$AU02_03,"3") = "Sometimes"
attr(ds$AU02_03,"4") = "Often"
attr(ds$AU02_03,"5") = "Very Often"
attr(ds$AU02_04,"1") = "Never"
attr(ds$AU02_04,"2") = "Rarely"
attr(ds$AU02_04,"3") = "Sometimes"
attr(ds$AU02_04,"4") = "Often"
attr(ds$AU02_04,"5") = "Very Often"
attr(ds$AU02_05,"1") = "Never"
attr(ds$AU02_05,"2") = "Rarely"
attr(ds$AU02_05,"3") = "Sometimes"
attr(ds$AU02_05,"4") = "Often"
attr(ds$AU02_05,"5") = "Very Often"
attr(ds$AU02_06,"1") = "Never"
attr(ds$AU02_06,"2") = "Rarely"
attr(ds$AU02_06,"3") = "Sometimes"
attr(ds$AU02_06,"4") = "Often"
attr(ds$AU02_06,"5") = "Very Often"
attr(ds$AU02_07,"1") = "Never"
attr(ds$AU02_07,"2") = "Rarely"
attr(ds$AU02_07,"3") = "Sometimes"
attr(ds$AU02_07,"4") = "Often"
attr(ds$AU02_07,"5") = "Very Often"
attr(ds$WA02_01,"1") = "Strongly disagree"
attr(ds$WA02_01,"2") = "Disagree"
attr(ds$WA02_01,"3") = "Neutral"
attr(ds$WA02_01,"4") = "Agree"
attr(ds$WA02_01,"5") = "Strongly agree"
attr(ds$WA02_02,"1") = "Strongly disagree"
attr(ds$WA02_02,"2") = "Disagree"
attr(ds$WA02_02,"3") = "Neutral"
attr(ds$WA02_02,"4") = "Agree"
attr(ds$WA02_02,"5") = "Strongly agree"
attr(ds$WA02_03,"1") = "Strongly disagree"
attr(ds$WA02_03,"2") = "Disagree"
attr(ds$WA02_03,"3") = "Neutral"
attr(ds$WA02_03,"4") = "Agree"
attr(ds$WA02_03,"5") = "Strongly agree"
attr(ds$WA02_04,"1") = "Strongly disagree"
attr(ds$WA02_04,"2") = "Disagree"
attr(ds$WA02_04,"3") = "Neutral"
attr(ds$WA02_04,"4") = "Agree"
attr(ds$WA02_04,"5") = "Strongly agree"
attr(ds$WA02_05,"1") = "Strongly disagree"
attr(ds$WA02_05,"2") = "Disagree"
attr(ds$WA02_05,"3") = "Neutral"
attr(ds$WA02_05,"4") = "Agree"
attr(ds$WA02_05,"5") = "Strongly agree"
attr(ds$WA02_06,"1") = "Strongly disagree"
attr(ds$WA02_06,"2") = "Disagree"
attr(ds$WA02_06,"3") = "Neutral"
attr(ds$WA02_06,"4") = "Agree"
attr(ds$WA02_06,"5") = "Strongly agree"
attr(ds$WA02_07,"1") = "Strongly disagree"
attr(ds$WA02_07,"2") = "Disagree"
attr(ds$WA02_07,"3") = "Neutral"
attr(ds$WA02_07,"4") = "Agree"
attr(ds$WA02_07,"5") = "Strongly agree"
attr(ds$WA02_08,"1") = "Strongly disagree"
attr(ds$WA02_08,"2") = "Disagree"
attr(ds$WA02_08,"3") = "Neutral"
attr(ds$WA02_08,"4") = "Agree"
attr(ds$WA02_08,"5") = "Strongly agree"
attr(ds$WA02_09,"1") = "Strongly disagree"
attr(ds$WA02_09,"2") = "Disagree"
attr(ds$WA02_09,"3") = "Neutral"
attr(ds$WA02_09,"4") = "Agree"
attr(ds$WA02_09,"5") = "Strongly agree"
attr(ds$WA02_10,"1") = "Strongly disagree"
attr(ds$WA02_10,"2") = "Disagree"
attr(ds$WA02_10,"3") = "Neutral"
attr(ds$WA02_10,"4") = "Agree"
attr(ds$WA02_10,"5") = "Strongly agree"
attr(ds$WA02_11,"1") = "Strongly disagree"
attr(ds$WA02_11,"2") = "Disagree"
attr(ds$WA02_11,"3") = "Neutral"
attr(ds$WA02_11,"4") = "Agree"
attr(ds$WA02_11,"5") = "Strongly agree"
attr(ds$WA02_12,"1") = "Strongly disagree"
attr(ds$WA02_12,"2") = "Disagree"
attr(ds$WA02_12,"3") = "Neutral"
attr(ds$WA02_12,"4") = "Agree"
attr(ds$WA02_12,"5") = "Strongly agree"
attr(ds$STATUS,"complete") = "Complete case"
attr(ds$STATUS,"finished") = "Interview finished (got to last page)"
attr(ds$STATUS,"quality fail") = "Flagged as invalid record"
attr(ds$STATUS,"quota full") = "Rejected due to quota stop"
attr(ds$STATUS,"screenout") = "Rejected due to screening criteria"
attr(ds$FINISHED,"F") = "Canceled"
attr(ds$FINISHED,"T") = "Finished"
attr(ds$Q_VIEWER,"F") = "Respondent"
attr(ds$Q_VIEWER,"T") = "Spectator"
comment(ds$SERIAL) = "Serial number (if provided)"
comment(ds$REF) = "Reference (if provided in link)"
comment(ds$QUESTNNR) = "Questionnaire that has been used in the interview"
comment(ds$MODE) = "Interview mode"
comment(ds$STARTED) = "Time the interview has started (Europe/Berlin)"
comment(ds$AU02_01) = "Aufgaben Häufigkeit: Coordination"
comment(ds$AU02_02) = "Aufgaben Häufigkeit: Social"
comment(ds$AU02_03) = "Aufgaben Häufigkeit: Operational"
comment(ds$AU02_04) = "Aufgaben Häufigkeit: Scientific Conception"
comment(ds$AU02_05) = "Aufgaben Häufigkeit: Creative Design"
comment(ds$AU02_06) = "Aufgaben Häufigkeit: External Communication"
comment(ds$AU02_07) = "Aufgaben Häufigkeit: Leadership / Mentoring"
comment(ds$AU03) = "Gedanken außerhalb Arbeitszeit"
comment(ds$AU04) = "Mehrarbeit"
comment(ds$DE03) = "Geschlecht"
comment(ds$DE04) = "Rolle"
comment(ds$DE06) = "Gruppe"
comment(ds$DE07_01) = "Teamzugehörigkeit: Month(s) ... "
comment(ds$DE07_02) = "Teamzugehörigkeit: /Year(s) ... "
comment(ds$DE08) = "Teamgröße"
comment(ds$WA01) = "Aufgaben aus Pflichtgefühl"
comment(ds$WA02_01) = "Likert-Zustimmung Aufgabenwahrnehmung: Research and lab tasks receive more recognition than organizational tasks."
comment(ds$WA02_02) = "Likert-Zustimmung Aufgabenwahrnehmung: The effort required for coordination and social tasks is underestimated within the team."
comment(ds$WA02_03) = "Likert-Zustimmung Aufgabenwahrnehmung: The visibility of work differs depending on the task area."
comment(ds$WA02_04) = "Likert-Zustimmung Aufgabenwahrnehmung: I actively decide which task areas I am involved in."
comment(ds$WA02_05) = "Likert-Zustimmung Aufgabenwahrnehmung: I have little say in which tasks are assigned to me."
comment(ds$WA02_06) = "Likert-Zustimmung Aufgabenwahrnehmung: My tasks align with my interests and strengths."
comment(ds$WA02_07) = "Likert-Zustimmung Aufgabenwahrnehmung: I consider the task distribution in my team to be fair."
comment(ds$WA02_08) = "Likert-Zustimmung Aufgabenwahrnehmung: I am satisfied with my areas of responsibility within the team."
comment(ds$WA02_09) = "Likert-Zustimmung Aufgabenwahrnehmung: Task distribution is discussed openly within the team."
comment(ds$WA02_10) = "Likert-Zustimmung Aufgabenwahrnehmung: I enjoy the tasks I do within the team."
comment(ds$WA02_11) = "Likert-Zustimmung Aufgabenwahrnehmung: I receive adequate recognition for my contributions to the team."
comment(ds$WA02_12) = "Likert-Zustimmung Aufgabenwahrnehmung: Certain tasks are expected of specific genders, whether spoken or not."
comment(ds$WA03) = "Geschlechterverteilung gesamt"
comment(ds$TIME001) = "Time spent on page 1"
comment(ds$TIME002) = "Time spent on page 2"
comment(ds$TIME003) = "Time spent on page 3"
comment(ds$TIME004) = "Time spent on page 4"
comment(ds$TIME_SUM) = "Time spent overall (except outliers)"
comment(ds$MAILSENT) = "Time when the invitation mailing was sent (personally identifiable recipients, only)"
comment(ds$LASTDATA) = "Time when the data was most recently updated"
comment(ds$STATUS) = "Interview status marker"
comment(ds$FINISHED) = "Has the interview been finished (reached last page)?"
comment(ds$Q_VIEWER) = "Did the respondent only view the questionnaire, omitting mandatory questions?"
comment(ds$LASTPAGE) = "Last page that the participant has handled in the questionnaire"
comment(ds$MAXPAGE) = "Hindmost page handled by the participant"
comment(ds$MISSING) = "Missing answers in percent"
comment(ds$MISSREL) = "Missing answers (weighted by relevance)"
comment(ds$TIME_RSI) = "Completion Speed (relative)"



# Assure that the comments are retained in subsets
as.data.frame.avector = as.data.frame.vector
`[.avector` <- function(x,i,...) {
  r <- NextMethod("[")
  mostattributes(r) <- attributes(x)
  r
}
ds_tmp = data.frame(
  lapply(ds, function(x) {
    structure( x, class = c("avector", class(x)) )
  } )
)
mostattributes(ds_tmp) = attributes(ds)
ds = ds_tmp
rm(ds_tmp)
```