

```

# R course for beginners
# Week 2
# assignment by < Itamar > < Ronel >, id < 032702391 >

#### CREATE DATAFRAME ----

# generate vectors
N = 100
subject_id = seq(from = 1, to = N, by = 1)
gender = sample(c('female', 'male'), size = N, replace = TRUE, prob = c(0.5, 0.5))
gender = as.factor(gender)
age = runif(N, 15, 40)
depression = rbinom(N, 1, 0.176)
IQ = rnorm(N, mean = 100, sd = 15)

# create data frame
df = data.frame(subject_id, gender, age, depression, IQ)

# save as .csv
write.csv(df, file = "data/df.csv", row.names = FALSE)

#### DESCRIPTIVE STATS ----

# base package descriptive data
dim(df) # גודל הטבלה
names(df) # שמות המשתנים
range(df$age) # טווח הגיל
mean(df$IQ) # ממוצע IQ
median(df$depression * 100) # חציון אחוזי הדיכאון

# dplyr package piping - UNDER 18
df |>
  filter(age < 18) |>
  mutate(depression = depression * 100) |>
  group_by(gender) |>
  summarize(
    mean_age = mean(age), # גיל ממוצע מתחת ל-18
    mean_iq = mean(IQ), # IQ ממוצע מתחת ל-18
    mean_depression = mean(depression) # ממוצע דיכאון מתחת ל-18
  )

# dplyr package piping - OVER 18
df |>
  filter(age > 18) |>
  mutate(depression = depression * 100) |>
  group_by(gender) |>
  summarize(
    mean_age = mean(age), # גיל ממוצע מעל ל-18
    mean_iq = mean(IQ), # IQ ממוצע מעל ל-18
    mean_depression = mean(depression) # ממוצע דיכאון מעל ל-18
  )

```