

# Assignment\_1\_—*MSB104*—\_Group\_3

Irjan\_&\_Magnus

```
library(modelr)
library(dplyr)
library(ggplot2)
library(tidyverse)
library(vtable)
library(flextable)
library(scales)
library(tinytex)
library(readxl)
```

```
Sys.setlocale("LC_ALL", "Norwegian")
```

```
[1] "LC_COLLATE=Norwegian_Norway.1252;LC_CTYPE=Norwegian_Norway.1252;LC_MONETARY=Norwegian_Norway.1252;LC_NUMERIC=Norwegian_Norway.1252;LC_TIME=Norwegian_Norway.1252"
```

```
library(readxl)
library(dplyr)
library(tidyr)
library(janitor)

# Les inn DEMO som character
demo <- read_excel("DEMO_Ass1.xlsx", sheet = 2, col_types = "text") %>%
  clean_names()

# Les inn GDP som character
gdp <- read_excel("GDP_Ass1.xlsx", sheet = 2, col_types = "text") %>%
  clean_names()
```

```
# ---- Gjør begge "long" ----
demo_long <- demo %>%
  pivot_longer(
    cols = starts_with("x"),
    names_to = "time",
    values_to = "population"
  ) %>%
  mutate(
    time = as.integer(str_remove(time, "^x")),
    population = as.numeric(population)
  )

gdp_long <- gdp %>%
  pivot_longer(
    cols = starts_with("x"),
    names_to = "time",
    values_to = "gdp"
  ) %>%
  mutate(
    time = as.integer(str_remove(time, "^x")),
    gdp = as.numeric(gdp)
  )

# Beregne GDP per innbygger
combined <- demo_long %>%
  left_join(gdp_long, by = join_by(region, time)) %>%
  mutate(gdp_per_capita = gdp / population* 1000000)
```

## Kort gjennomgang av datasett og variabler

Datasette *demo\_r\_pjanggr3* som er hentet fra Eurostat inneholder årlige befolkningestimater på NUTS3-nivå for EU-, EFTA- og kandidatland. Variablene **values** viser totalt antall bosatte personer per 1. januar, målt i antall personer. Hver observasjon identifiseres ved regionkode(**geo**) og år (**time**), samt kjønn (**sex**) og alder (**age**). I denne analysen benyttes kun total befolkning (**sex = T, age = TOTAL**), slik at dataene ikke er splittet etter kjønn eller alder.