

load_vaccine_data

OOBr

2025-11-07

Load vaccination microdata from the “basedosdados”

```
# Load vaccination microdata from the "basedosdados" project, specifically the
# COVID-19 vaccination dataset from the Brazilian Ministry of Health, considering
# general population and pregnant and postpartum women.

# Important packages
library(basedosdados)
library(dplyr)
library(questionr)
require("data.table")
require("tidyverse")
require("readxl")
require("writexl")

memory.limit(999999)

# set_billing_id("token")

# Load vaccination microdata from the "basedosdados" project
oi <- bdplyr("basedosdados.br_ms_vacinacao_covid19.micrdados") %>%
  select(sigla_uf, data_aplicacao_vacina, dose_vacina) %>%
  group_by(sigla_uf, data_aplicacao_vacina, dose_vacina) %>%
  summarise(n_casos = n()) %>%
  bd_collect()

# Set vaccine doses
oil <- oi %>%
  mutate(
    num_dose = case_when(
      dose_vacina == "1ª Dose" ~ "1a dose",
      dose_vacina == "1ª Dose Revacinação" ~ "1o reforço",
      dose_vacina == "1º Reforço" ~ "1o reforço",
      dose_vacina == "2ª Dose" ~ "2a dose ou unica",
      dose_vacina == "2ª Dose Revacinação" ~ "2o reforço",
      dose_vacina == "2º Reforço" ~ "2o reforço",
      dose_vacina == "3ª Dose" ~ "1o reforço",
      dose_vacina == "3ª Dose Revacinação" ~ "1o reforço",
      dose_vacina == "3º Reforço" ~ "1o reforço",
      dose_vacina == "4ª Dose" ~ "2o reforço",
      dose_vacina == "Dose Adicional" ~ "1o reforço",
```

```

dose_vacina == "Dose Inicial" ~ "1a dose",
dose_vacina == "Dose Única" ~ "2a dose ou unica",
dose_vacina == "Reforço" ~ "1o reforço",
dose_vacina == "Revacinação" ~ "1o reforço",
dose_vacina == "Única" ~ "2a dose ou unica"
)
)

oi2 <- oi1 %>%
  select(-dose_vacina) %>%
  group_by(sigla_uf, data_aplicacao_vacina, num_dose) %>%
  summarise(n_casos = sum(n_casos)) %>%
  select(sigla_uf, data_aplicacao_vacina, dose_vacina = num_dose, n_casos) %>%
  filter(!is.na(dose_vacina))

# Brazilian population by (federative unity) vaccinated
write_csv(oi2, "Vacinados_porUF_01-02-22.csv")
write_xlsx(oi2, "Vacinados_porUF_01-02-22.xlsx")

dados_br <- oi2 %>%
  group_by(data_aplicacao_vacina, dose_vacina) %>%
  summarise(num_casos = sum(n_casos))

# Brazilian population vaccinated
write_csv(dados_br, "Vacinados_Brasil_01-02-22.csv")
write_xlsx(dados_br, "Vacinados_Brasil_01-02-22.xlsx")

```

Now filtering for pregnant and postpartum women

```

dados_gesta_puerp1 <- dados_gesta_puerp %>%
  mutate(
    num_dose = case_when(
      dose_vacina == "1ª Dose" ~ "1a dose",
      dose_vacina == "1ª Dose Revacinação" ~ "1o reforço",
      dose_vacina == "1º Reforço" ~ "1o reforço",
      dose_vacina == "2ª Dose" ~ "2a dose ou unica",
      dose_vacina == "2ª Dose Revacinação" ~ "2o reforço",
      dose_vacina == "2º Reforço" ~ "2o reforço",
      dose_vacina == "3ª Dose" ~ "1o reforço",
      dose_vacina == "3ª Dose Revacinação" ~ "1o reforço",
      dose_vacina == "3º Reforço" ~ "1o reforço",
      dose_vacina == "4ª Dose" ~ "2o reforço",
      dose_vacina == "Dose Adicional" ~ "1o reforço",
      dose_vacina == "Dose Inicial" ~ "1a dose",
      dose_vacina == "Dose Única" ~ "2a dose ou unica",
      dose_vacina == "Reforço" ~ "1o reforço",
      dose_vacina == "Revacinação" ~ "1o reforço",
      dose_vacina == "Única" ~ "2a dose ou unica"
    )
  )

# Aggregate vaccination counts for pregnant/puerperal women at the state (UF) level.
dados_materna_UF <- dados_gesta_puerp1 %>%
  # Remove the original, more granular dose_vacina label to avoid confusion,
  # since we will now work with the standardized `num_dose`.
  select(-dose_vacina) %>%
  # Group by state, vaccination date, and standardized dose category
  group_by(sigla_uf, data_aplicacao_vacina, num_dose) %>%
  summarise(n_casos = sum(n_casos)) %>%
  select(sigla_uf, data_aplicacao_vacina, dose_vacina = num_dose, n_casos) %>%
  filter(!is.na(dose_vacina))

# Export state-level dataset for pregnant/puerperal women in both CSV and XLSX formats.
# The filenames suggest a data extraction or update up to 01-02-22.
write_csv(dados_materna_UF, "Vacinados_GestaPuerp_porUF_01-02-22.csv")
write_xlsx(dados_materna_UF, "Vacinados_GestaPuerp_porUF_01-02-22.xlsx")

# Build a national-level aggregation (Brazil as a whole) by summing across all states.
dados_materna_Br <- dados_materna_UF %>%
  group_by(data_aplicacao_vacina, dose_vacina) %>%
  summarise(n_casos = sum(n_casos))

# Export national-level dataset in CSV and XLSX formats.
write_csv(dados_materna_Br, "Vacinados_GestaPuerp_Brasil_01-02-22.csv")
write_xlsx(dados_materna_Br, "Vacinados_GestaPuerp_Brasil_01-02-22.xlsx")

#----- The End

# oi1 <- oi %>%
#   mutate(
#     num_dose = case_when(

```

```
#      id %in% str_which(dose_vacina, "1ª Dose") ~ "1a dose",
#      id %in% str_which(dose_vacina, "1ª Dose Revacinação") ~ "1a dose",
#      id %in% str_which(dose_vacina, "2º") ~ "2a dose",
#      id %in% str_which(dose_vacina, "nica") ~ "dose unica",
#      id %in% str_which(dose_vacina, "Única") ~ "dose unica",
#      id %in% str_which(dose_vacina, "DoseA") ~ "dose unica",
#      id %in% str_which(dose_vacina, "Adicional") ~ "dose adicional",
#      id %in% str_which(dose_vacina, "DoseAdicional") ~ "dose adicional",
#      vacina_descricao_dose == "Dose" ~ "dose unica",
#      id %in% str_which(vacina_descricao_dose, "3º") ~ "dose adicional",
#      id %in% str_which(vacina_descricao_dose, "Reforço") ~ "dose adicional"
#    )
#  )
```