

# Raw exit poll data preparation

Part of the final project for AQMSS II

Stepan Polikanov

Vera Okisheva

```
# Prepping raw exit poll data
# Dependencies: exitpoll_rawdata.xlsx
```

```
# Packages
source(here::here("utilities", "check_packages.R"))
```

```
# Load exit poll raw data
ep_raw <- read_excel(here("data", "data_raw", "exitpoll_individual",
                          "exitpoll_rawdata.xlsx"),
                    sheet = 4, guess_max = 69262)
```

```
ep_raw_clean <- ep_raw |>
  separate(col = " ", into = c("countryname_ru", "countryname_en"),
    sep = " / ") |>
  separate(col = " ", into = c("city_ru", "city_en"),
    sep = " / ") |>
  mutate(countryname_en = if_else(countryname_en == "Czech",
                                   "Czechia", countryname_en),
    countrycode_n = countrycode(countryname_en,
                                  origin = "country.name",
                                  destination = "iso3n"),
    countrycode_c = countrycode(countrycode_n,
                                  origin = "iso3n",
                                  destination = "iso3c"),
    voting_station = as.character(` `),
    city_en = if_else(str_detect(city_ru, "№") == T,
                      paste(city_en, str_sub(city_ru, -1,
                                              nchar(city_ru))),
                      city_en)) |>
  transmute(volunteer_id = `ID`,
    data_source = case_when(
      ` ` == "QR" ~ "One-off QR e-form",
      ` ` == " " ~ "General e-form",
      ` ` == " " ~ "Tokio e-form",
      ` ` == " " ~ "Stockholm e-form",
      ` ` == " " ~ "Prague e-form",
      ` ` == " " ~ "Dubai e-form",
      ` ` == " " ~ "Wellington e-form",
      ` ` == " " ~ "Sydney e-form"),
    upload_time = ymd_hms(` `),
    countryname_en, countryname_ru, city_ru, city_en, countrycode_c, countrycode_n,
```

```

voting_station,
vote = case_when(
  ~ ?^ == " . . " ~ "Davankov",
  ~ ?^ == " . . " ~ "Slutsky",
  ~ ?^
  == " ( , )"
  ~ "Spoiled ballot",
  ~ ?^ == " / / "
  ~ "Tore up/took",
  ~ ?^ == " . . " ~ "Putin",
  ~ ?^ == " "
  ~ "Declined to answer",
  ~ ?^ == " . . " ~ "Haritonov"
),
sex = case_when(~ ?^ == " " ~ "Male",
  ~ ?^ == " " ~ "Female",
  ~ ?^ == " " ~ "Other",
  ~ ?^ == " " ~ "Declined to answer",
  .default = NA),
age_bin = case_when(
  ~ ?^ == " " ~ "Declined to answer",
  .default = ?^ ),
out_of_Russia_time = case_when(
  ~ ?^ == " 6 " ~ "< 6 months",
  ~ ?^ == "6 - 2 ( 2022)"
  ~ "6 months - 2 years",
  ~ ?^ == " 2 " ~ "< 2 years",
  ~ ?^ == "2 - 5 " ~ "2 - 5 years",
  ~ ?^ == " 5 " ~ "> 5 years",
  ~ ?^ == "6-10 " ~ "6 - 10 years",
  ~ ?^ == " 10 " ~ "> 10 years",
  ~ ?^ == " "
  ~ "Declined to answer",
  ~ ?^ == " / ( )"
  ~ "Tourist (lives in Russia)", .default = NA),
time_to_vs = case_when(
  ~ ?^ == "<30 "
  ~ "<30 minutes" ,
  ~ ?^ == "30 - 1 "
  ~ "30 minutes - 1 hour" ,
  ~ ?^ == "1 - 2 "
  ~ "1 - 2 hours",
  ~ ?^ == "2 - 3 "
  ~ "2 - 3 hours",
  ~ ?^ == "3 - 4 "
  ~ "3 - 4 hours",
  ~ ?^ == " 2 "
  ~ "> 2 hours" ,
  ~ ?^
  == " 4 ( )"
  ~ "> 4 hours (staying for the night)",
  ~ ?^ == " "
  ~ "Declined to answer", .default = NA),
result_trust = case_when(
  ~ ?^ %in% c(" ", " ")
  ~ "Definitely no",
  ~ ?^ == " "
  ~ "Probably no",
  ~ ?^ == " "
  ~ "Probably yes",

```

```

~                                     ?` %in% c("      ", "      ")
~ "Definitely yes",
~                                     ?` == "      "
~ "Don't know",
~                                     ?` == "      "
~ "Declined to answer", .default = NA),
uik_closetime = ` - (CET)`,
uik_nvoters = ` - `,
uik_didntvote_closed = ` - `,
id = row_number())

```

```

saveRDS(ep_raw_clean, file = here("data", "data_built", "ep_raw_clean.rds"))

```