

## Appendix 1: Data Cleaning

```
library(tidyverse)
library(tsibble)
library(lubridate)

# Read the unclean data
fires <- read_csv(here::here("data-raw", "amazon.csv"))

# Fixing the Problem Month
problem_term <- fires %>%
  filter(stringr::str_detect(month, "[^a-zA-Z]")) %>%
  slice(1) %>%
  pull(month)

fires_date <- fires %>%
  mutate(month_clean = case_when(
    month == "Abril" ~ "April",
    month == "Agosto" ~ "August",
    month == "Dezembro" ~ "December",
    month == "Fevereiro" ~ "February",
    month == "Janeiro" ~ "January",
    month == "Julho" ~ "July",
    month == "Junho" ~ "June",
    month == "Maio" ~ "May",
    month == problem_term ~ "March",
    month == "Novembro" ~ "November",
    month == "Outubro" ~ "October",
    month == "Setembro" ~ "September",
    TRUE ~ NA_character_
  ),
  day = 1) %>%
  mutate(date_month = lubridate::ymd(glue::glue("{year}-{month_clean}-{day}"))) %>%
  mutate(date_index = yearmonth(date_month)) %>%
  mutate(state = ifelse(str_detect(state, "Par\\xe1"), "Para", state))

# Creating the clean fire object
fires_clean <-
  fires_date %>%
  filter(!(state %in% c("Mato Grosso", "Rio", "Paraíba"))) %>%
  filter(duplicated(.) == FALSE) %>%
  select(-c(month, date, day, date_month)) %>%
  rename(date = date_index, month = month_clean, fires = number) %>%
  mutate(state = ifelse(str_detect(state, "Par\\xe1"), "Para", state)) %>%
  as_tsibble(index = date, key = state)

fires_clean_all <-
  fires_date %>%
  filter(!(state %in% c("Mato Grosso", "Rio", "Paraíba"))) %>%
  filter(duplicated(.) == FALSE) %>%
```

```
group_by(date_index) %>%
summarise(fires = sum(number)) %>%
as_tsibble(index = date_index) %>%
rename(date = date_index)

# Save the clean data
write_rds(fires_clean, here::here("data", "fires_clean.rds"))
write_rds(fires_clean_all, here::here("data", "fires_clean_all.rds"))

write_rds(fires_date, here::here("data-raw", "fires_date.rds"))
```