Graphic Design with ggplot2

Group Projects:

"Solutions"

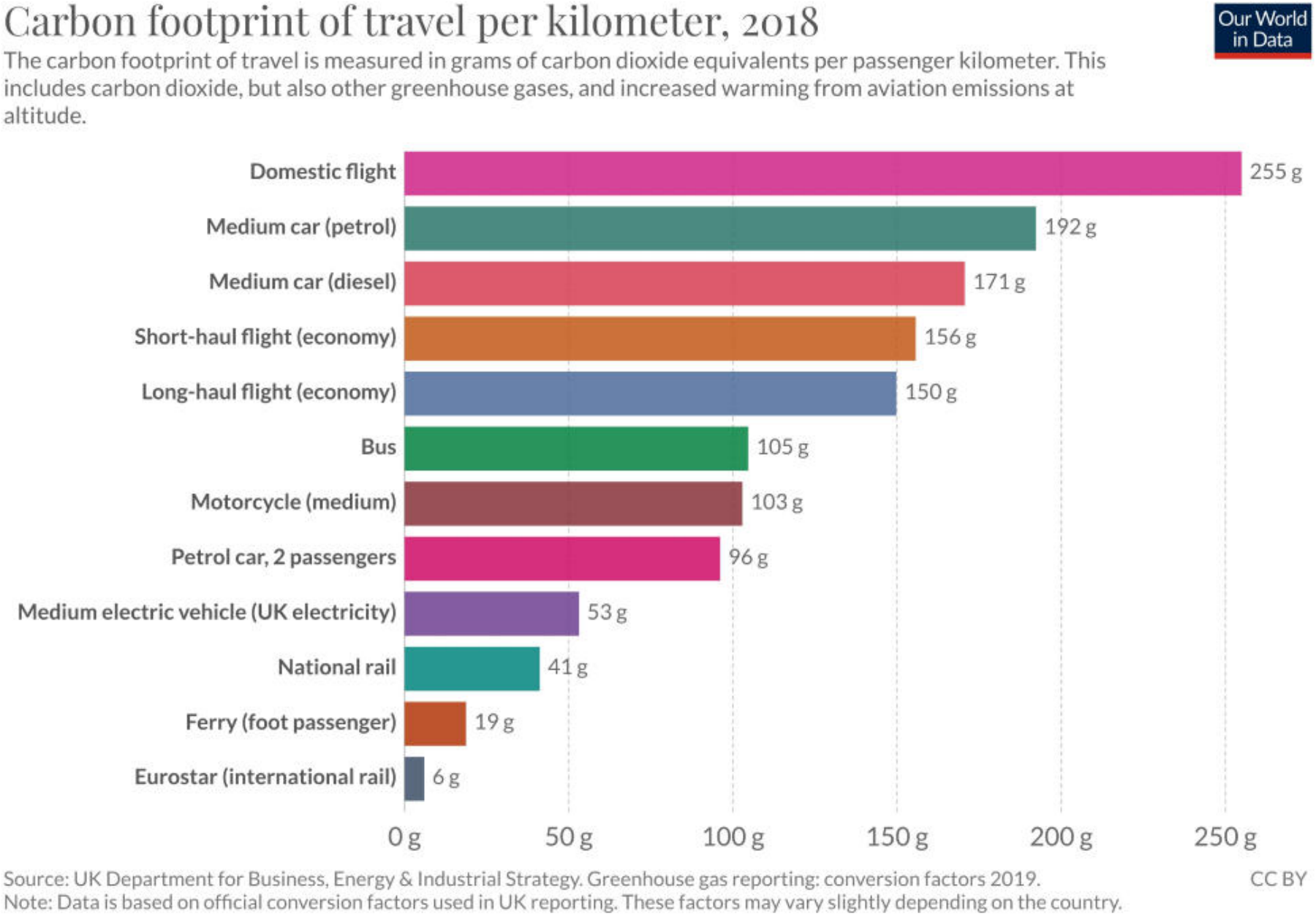
Cédric Scherer // rstudio::conf // July 2022

Group Projects

- Form groups and work one of the following suggested projects:
  - Carbon Footprint of Travel (OWID / UK.gov)
  - Spending Before and During the Pandemic (JP Morgan Chase)
  - Speed of Languages (Economist / Coupé et al.)
  - US Drought Patterns (Drought Monitor)

Carbon Footprint of Travel

Cédric Scherer // rstudio::conf // July 2022

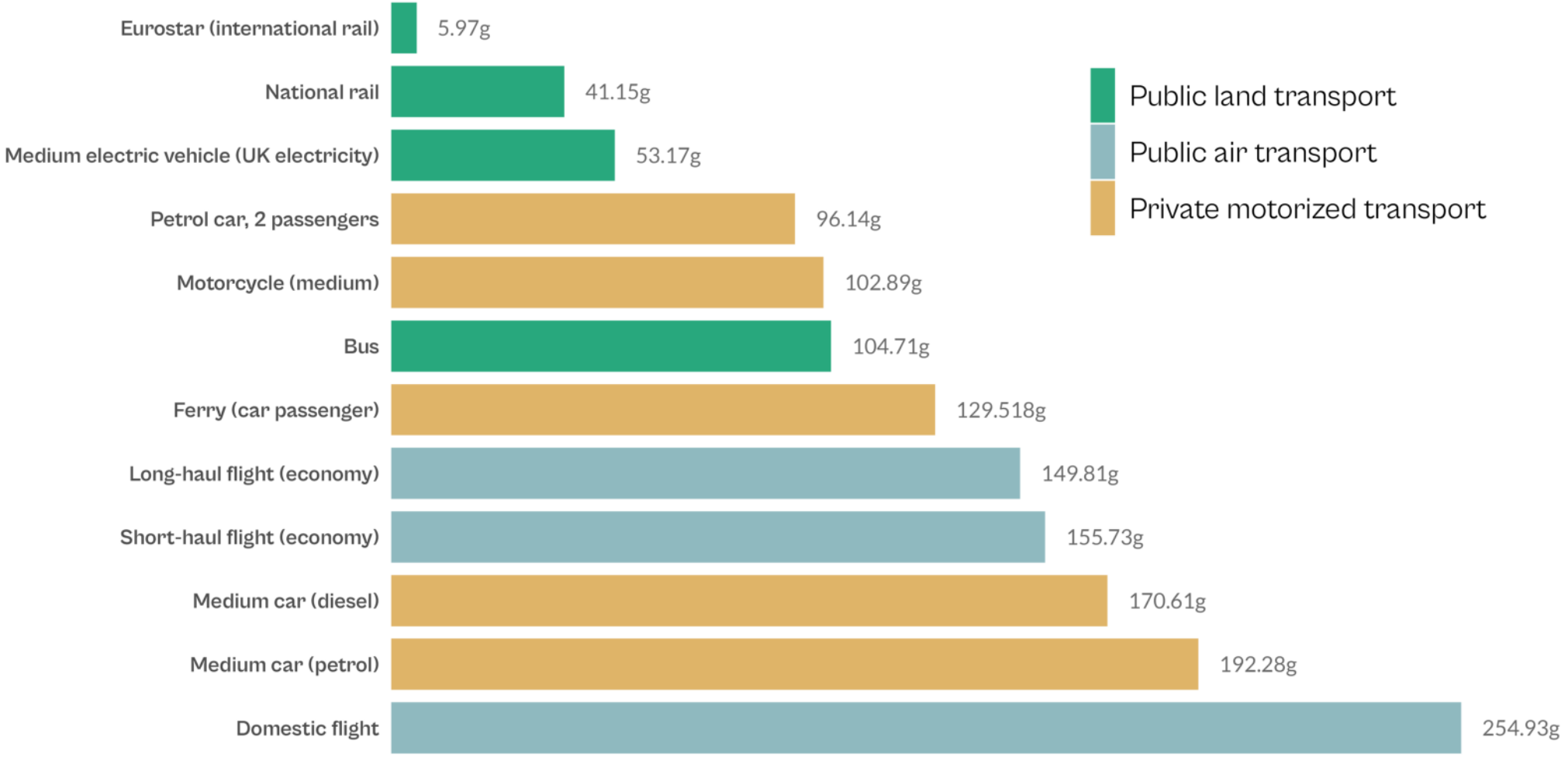


Graphic Source: Our World in Data

```r
library(tidyverse)

data <- read_csv(here::here("data", "carbon-footprint-travel.csv"))

data %>%
  mutate(
    type = case_when(
      str_detect(entity, "car|Motorcycle") ~ "Private motorized transport",
      str_detect(entity, "flight") ~ "Public air transport",
      str_detect(entity, "Ferry") ~ "Public water transport",
      TRUE ~ "Public land transport"
    )
  ) %>%
  ggplot(
    aes(x = emissions,
        y = forcats::fct_reorder(entity, -emissions),
        fill = type)
  ) +
  geom_col(orientation = "y", width = .8) +
```

Carbon footprint of travel per kilometer, 2018

The carbon footprint of travel is measured in grams of carbon dioxide equivalents per passenger kilometer.
This includes carbon dioxide, but also other greenhouse gases, and increased warming from aviation emissions at altitude.

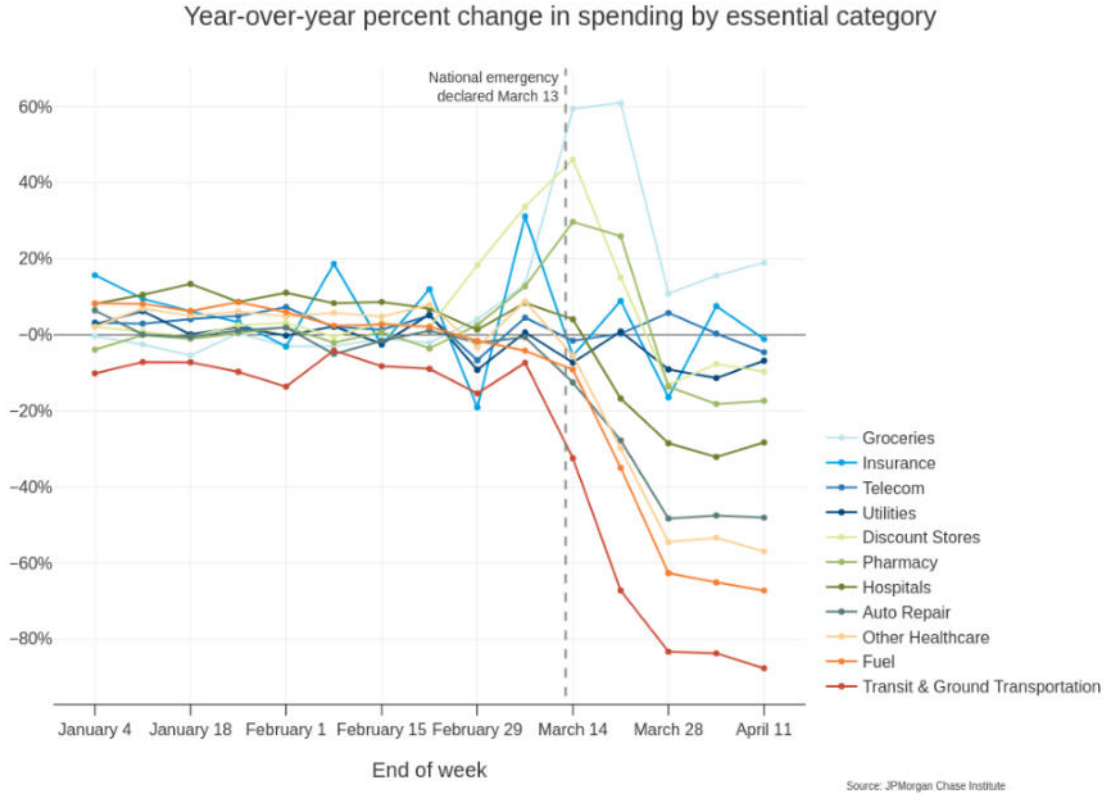


Source: UK Department for Business, Energy & Industrial Grenhouse gas reporting: conversion factors 2019.

Note: Data is based on official conversion factors used in UK reporting. These factors may vary slightly depending on the country.

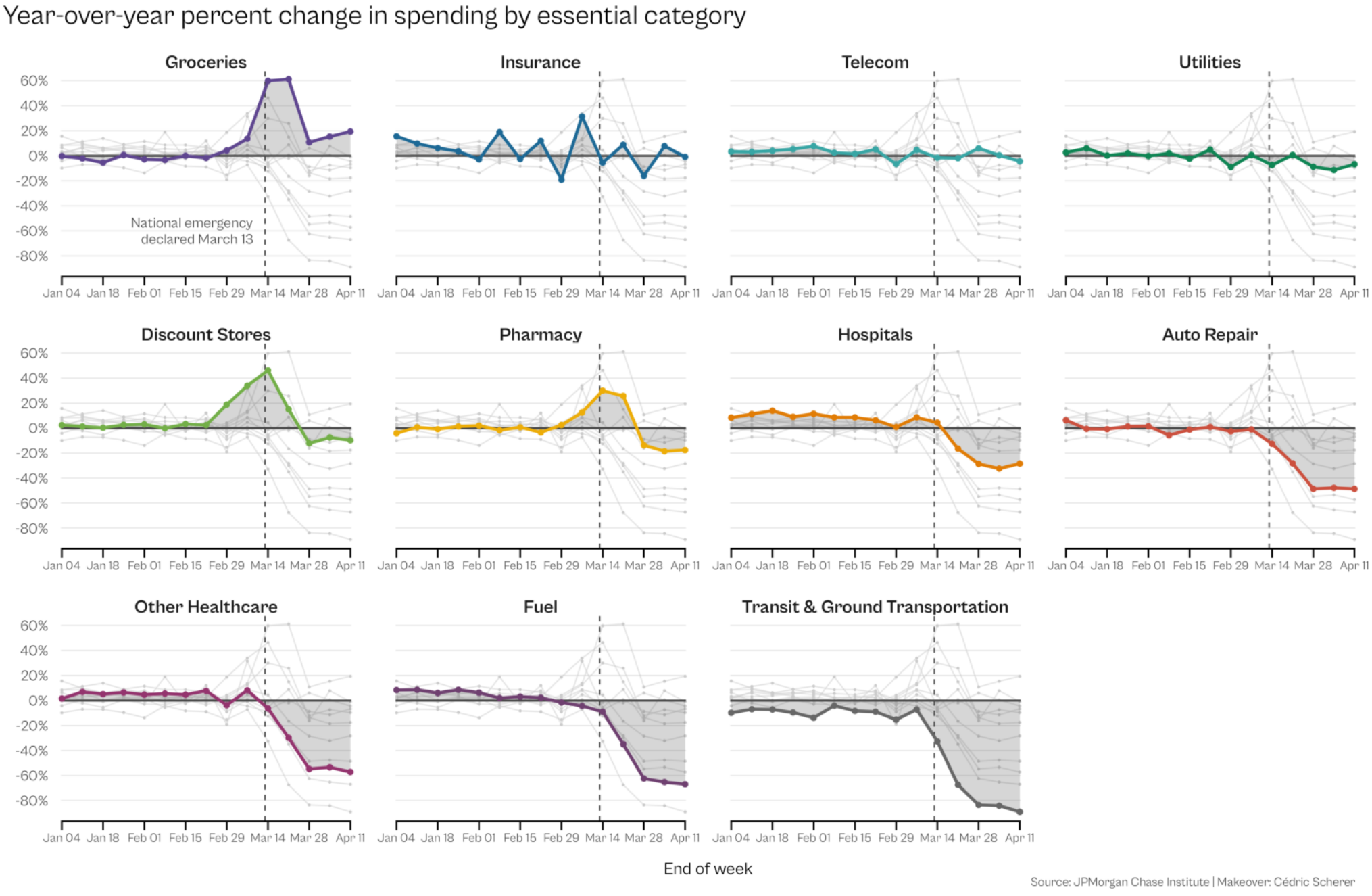
Original visualization by Hannah Ritchie, OurWorldInData.org | Makeover by Cédric Scherer

Spending Before and During the Pandemic

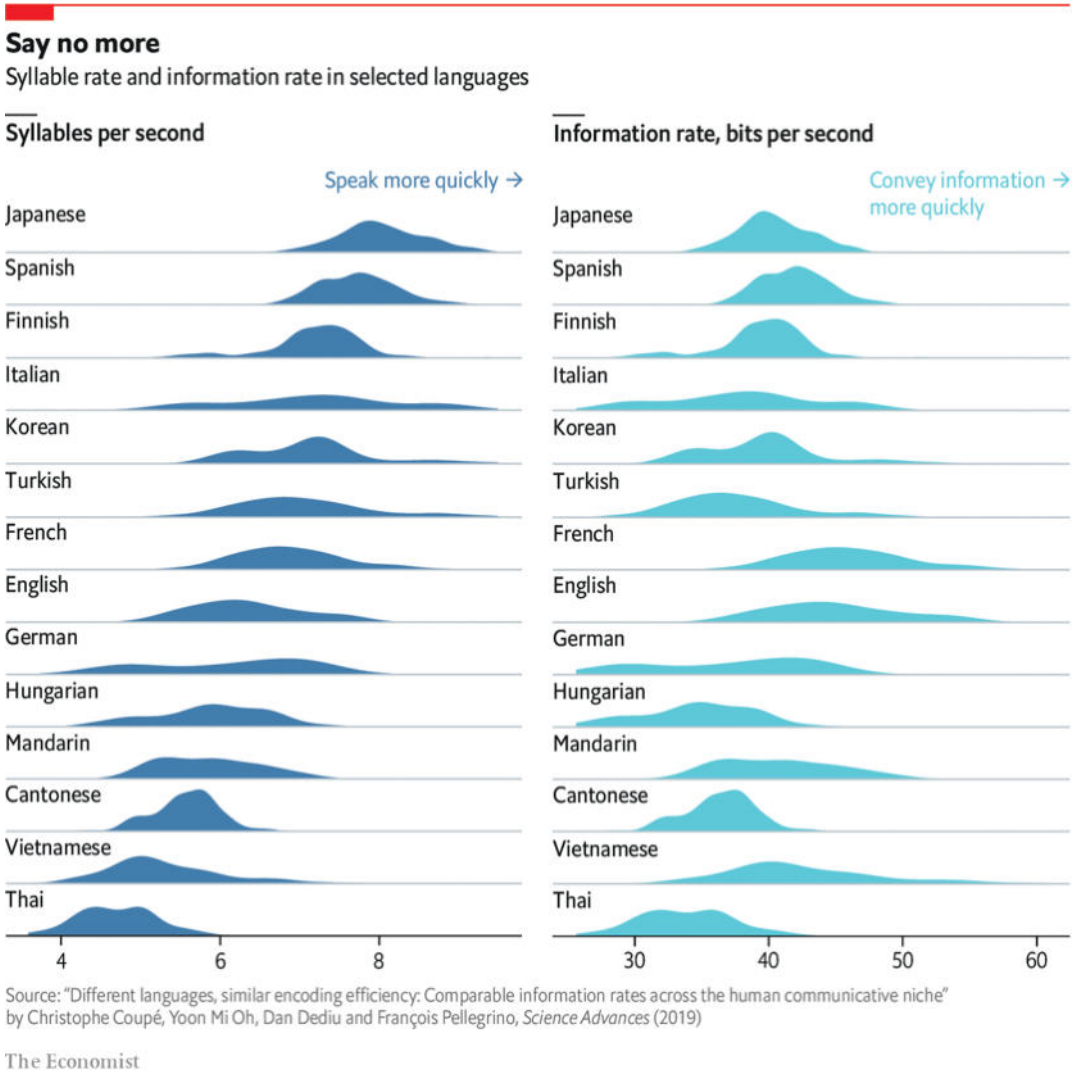


Graphic Source: JPMorgan Chase Institute

```r
library(tidyverse)
library(gghighlight)
library(lubridate)

invisible(Sys.setlocale("LC_TIME", "C"))

data <- read_csv(here::here("data", "spending-jpmorgan.csv")) %>%
  mutate(category = fct_inorder(category))

label_df <-
  tibble(
    date = ymd("2020-03-13"),
    change = -60,
    label = "National emergency\ndeclared March 13",
    category = factor("Groceries", levels = levels(data$category))
  )

ggplot(data, aes(date, change, color = category)) +
    geom_point() +
```


Year-over-year percent change in spending by essential category
Groceries
Insurance
Telecom
Utilities
Discount Stores
Pharmacy
Hospitals
Auto Repair
Other Healthcare
Fuel
Transit & Ground Transportation
60%
40%
20%
0%
-20%
-40%
-60%
-80%
National emergency declared March 13
Jan 04 Jan 18 Feb 01 Feb 15 Feb 29 Mar 14 Mar 28 Apr 11
End of week
Source: JPMorgan Chase Institute | Makeover: Cédric Scherer

Speed of Languages

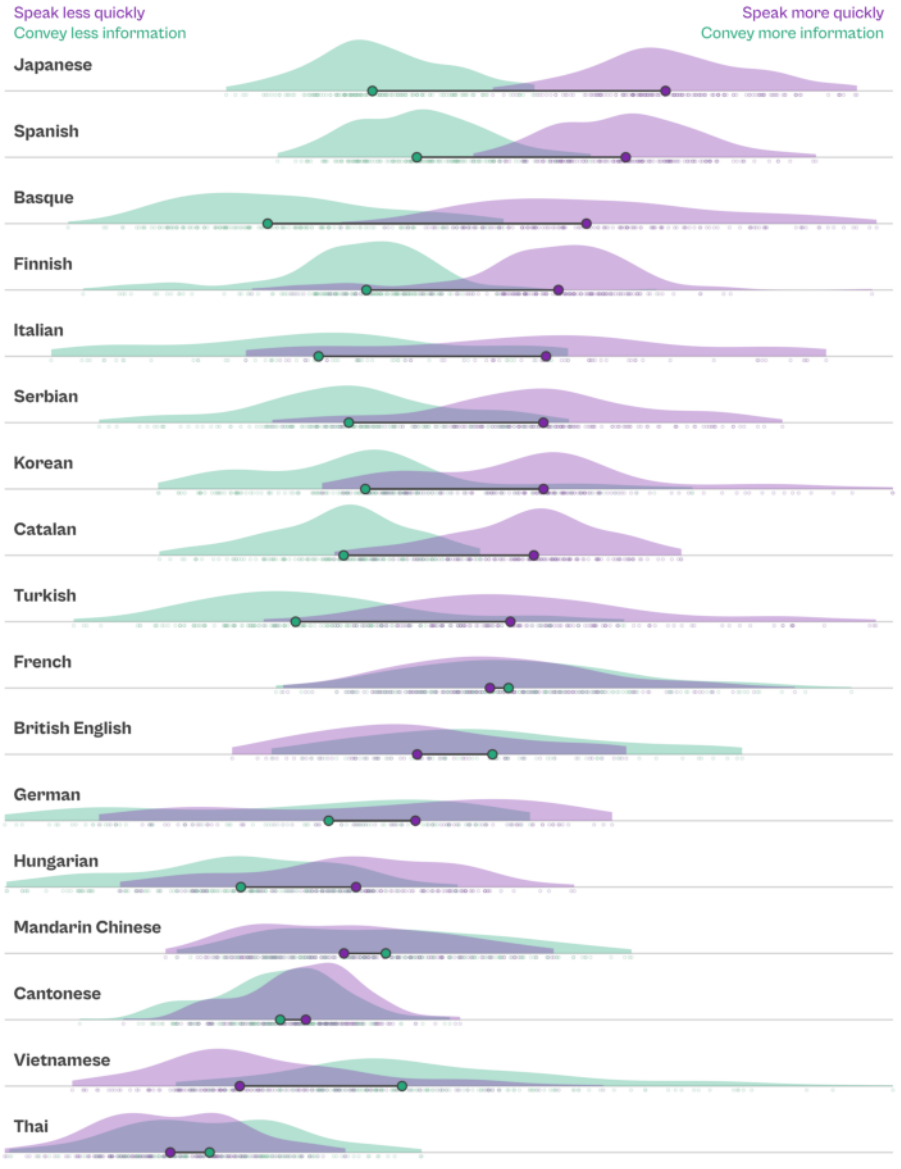


Graphic Source: The Economist

```r
library(tidyverse)
library(ggtext)
library(colorspace)

data <-
  read_csv(here::here("data", "information-speech.csv")) %>%
  group_by(language) %>%
  mutate(
    avg_sr = mean(speech_rate),
    avg_ir = mean(info_rate)
  ) %>%
  ungroup() %>%
  mutate(
    language = fct_reorder(language, avg_sr),
    language_long = fct_reorder(language_long, avg_sr)
  )

systemfonts::register_variant(
  name = "Cabinet Grotesk ExtraBold",
```

Communicating fast doesn't necessarily mean communicating more

Variation in speech and information rates across languages, shown as normalized rates for direct comparison. While there are stark cross-linguistic differences in speech rates, information rates are more similar.



Source: Coupé *et al.* 2019 *Science Advances* 5(9). DOI: 10.1126/sciadv.aaw2594
Graphic: Cédric Scherer • Dots show the median rates for each language.