Week-10: Diary entry

Dawn Cheung

2023-10-30

1. What is the question that you are going to answer? (Answer: One sentence that ends with a question mark that could act like the title of your data story),

How does unequal access to the internet affect socio-economic outcomes?

2. Why is this an important question? (Answer: 3 sentences, each of which has some evidence, e.g., “According to the United Nations…” to justify why the question you have chosen is important),

According to the world economic forum, “a third of the world’s population (some 2.9 billion people) suffers from the digital divide”, and this significant portion of the world hence do not have access to digital tools, technologies, and information that may greatly improve their quality of life and standard of living. The UN also states that “Social well-being will not be possible without a digital transformation”, particularly during crisies like COVID-19. Hence, unequal access to the internet excaberates social and economic inequality, and this creates a vicious cycle where countries lack the resources to build public ICT infrastructure, reducing people’s ability to utilize digital resources to increase their material/non-material standard of living, hence GDP per capita falls, so these countries become poorer and more unable to implement public ICT infrastructure

Sources:

https://www.weforum.org/agenda/2022/05/how-to-counter-the-global-digital-divide/ (https://www.weforum.org/agenda/2022/05/how-to-counter-the-global-digital-divide/)

https://press.un.org/en/2021/soc4890.doc.htm (https://press.un.org/en/2021/soc4890.doc.htm)

3. Which rows and columns of the dataset will be used to answer this question? (Answer: Actual names of the variables in the dataset that you plan to use).

PLANNING TO INCLUDE:

% population access to the internet, based on country

internet speeds (both broadband and mobile), based on country

GINI coefficient

3i survey data

- Being able to use the internet to access information and express opinions should be a human right? 73% agree
- The internet has improved life in my country 66% agree
- Using the internet has helped me pursue my education or that of my household members 67% agree

CONSIDERING, BUT UNLIKELY TO INCLUDE:

ICT development ranking index (only 2017 from what il could find)

DESI index (only EU)

Include the challenges and errors that you faced and how you overcame them.

One challenge i faced was world maps. I had to try multiple methods, including the sf library (didn’t work). Then, geom_polygon() worked, but some countries went missing. It was funny, the map had holes and Russia was gone. This was because the names of the countries used in the dataset did not correspond with the maps library’s countries’ name (eg USA vs US vs United States etc, Congo Dem. Rep. vs Democratic Republic of Congo). So my solution was to use the ISO codes because they were consistent (still had problems with ISO-2 and ISO-3) but it was resolved; used ISO-3 in the end. Hence, the challenge was overcome using the internet and sleep deprevation.

```
library(tidyverse)
```

```
## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──
## ✓ dplyr     1.1.2     ✓ readr     2.1.4
## ✓ forcats   1.0.0     ✓ stringr   1.5.0
## ✓ ggplot2   3.4.4     ✓ tibble    3.2.1
## ✓ lubridate 1.9.2     ✓ tidyr     1.3.0
## ✓ purrr     1.0.2
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
## ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(ggplot2)
library(maps)
```

```
##
## Attaching package: 'maps'
##
## The following object is masked from 'package:purrr':
##
##     map
```

```
library(sf)
```

```
## Linking to GEOS 3.11.2, GDAL 3.7.2, PROJ 9.3.0; sf_use_s2() is TRUE
```

```
library(dplyr)
library(countrycode)

# Import the data with coordinates
world_map <- map_data("world")
world_map$region <- countrycode(world_map$region, "country.name", "iso3c")
```

```
## Warning: Some values were not matched unambiguously: Ascension Island, Azores, Barbuda, Bonaire, Canary Islands, Chagos A
rchipelago, Grenadines, Heard Island, Kosovo, Madeira Islands, Micronesia, Saba, Saint Martin, Siachen Glacier, Sint Eustati
us, Virgin Islands
```

```
accesstonet <- read.csv("accesstointernet.csv")

imstupid <- accesstonet %>%
   pivot_longer(
    cols = X1990:X2022,
    names_to = "Xyear",
    values_to = "percentaccess") %>%
  rename(region = Country.Code)

yearspecs <- imstupid %>% filter(Xyear == "X2010")
sotired_map <- left_join(world_map, yearspecs, by = "region")

#group = group connects the points in the correct order
ggplot(sotired_map, aes(long, lat, group = group))+
  geom_polygon(aes(fill = percentaccess), color = "white", size = 0.03)+
  scale_fill_viridis_c(option = "B")
```

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
## ℹ Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```