

# Project Report

2024-03-29

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.3.3
```

```
library(tidyverse)
```

```
## Warning: package 'forcats' was built under R version 4.3.3
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --  
## v dplyr      1.1.4      v readr      2.1.5  
## v forcats    1.0.0      v stringr    1.5.1  
## v lubridate  1.9.3      v tibble     3.2.1  
## v purrr      1.0.2      v tidyr      1.3.1
```

```
## -- Conflicts ----- tidyverse_conflicts() --
```

```
## x dplyr::filter() masks stats::filter()
```

```
## x dplyr::lag()     masks stats::lag()
```

```
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(dplyr)
```

```
library(showtext)
```

```
## Warning: package 'showtext' was built under R version 4.3.3
```

```
## Loading required package: sysfonts
```

```
## Warning: package 'sysfonts' was built under R version 4.3.3
```

```
## Loading required package: showtextdb
```

```
## Warning: package 'showtextdb' was built under R version 4.3.3
```

```
font_add_google("Open Sans", "Open Sans")  
showtext_auto()
```

```
df <- read.csv("data/owid-energy.csv")
```

## Question 1

```

df_2021 = df[df$year == 2021, ]

df_2021_top_5 <- df_2021[df_2021$country == "China" | df_2021$country == "India" | df_2021$country == "Germany" | df_2021$country == "Japan" | df_2021$country == "United States", ]

df_2021_top_5_important_metric <- df_2021_top_5[,c("country", "population", "gdp", "biofuel_share_elec", "coal_share_elec", "fossil_share_elec", "gas_share_elec", "hydro_share_elec", "low_carbon_share_elec", "nuclear_share_elec", "oil_share_elec", "other_renewables_share_elec", "renewables_share_elec", "solar_share_elec", "wind_share_elec")]

df_2021_top_5_important_metric

##           country population gdp biofuel_share_elec coal_share_elec
## 4349          China 1425893504 NA           2.003           62.932
## 7731          Germany 83408560 NA           8.057           28.253
## 9287          India 1407563904 NA           2.070           74.173
## 10185         Japan 124612528 NA           3.851           32.510
## 20674 United States 336997632 NA           1.307           21.624
##           fossil_share_elec gas_share_elec hydro_share_elec low_carbon_share_elec
## 4349           66.289           3.213           15.323           33.711
## 7731           48.487           16.354           3.377           51.513
## 9287           78.053           3.745           9.356           21.947
## 10185          71.002           35.119           8.256           28.998
## 20674          60.509           38.037           5.936           39.491
##           nuclear_share_elec oil_share_elec other_renewables_share_elec
## 4349           4.803           0.144           2.003
## 7731           11.873           3.880           8.098
## 9287           2.563           0.135           2.070
## 10185           6.387           3.373           4.166
## 20674          18.742           0.848           1.746
##           renewables_share_elec solar_share_elec wind_share_elec
## 4349           28.908           3.854           7.727
## 7731           39.640           8.474          19.691
## 9287           19.384           3.986           3.973
## 10185           22.611           9.254           0.935
## 20674           20.750           3.960           9.108

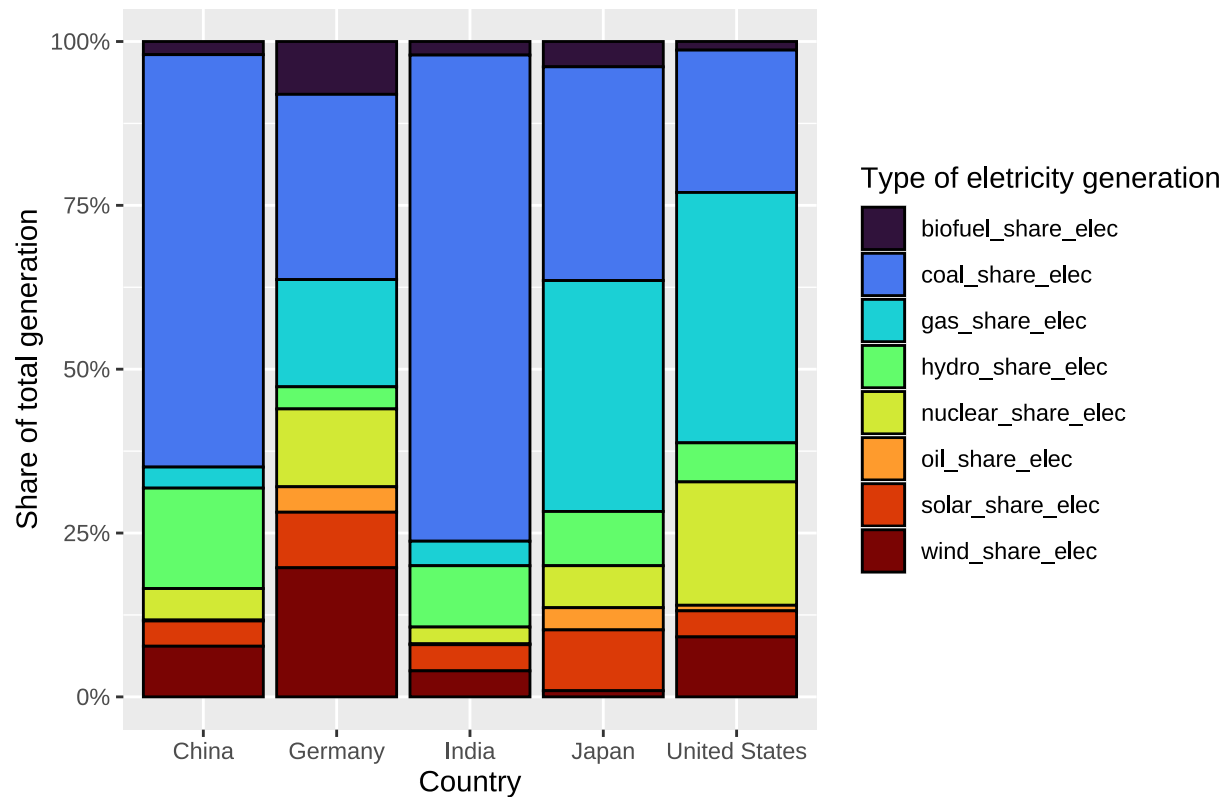
df_2021_top_5_important_metric$total = df_2021_top_5_important_metric$biofuel_share_elec + df_2021_top_5_important_metric$coal_share_elec +
df_2021_top_5_important_metric$gas_share_elec +
df_2021_top_5_important_metric$hydro_share_elec +
df_2021_top_5_important_metric$nuclear_share_elec +
df_2021_top_5_important_metric$oil_share_elec +
df_2021_top_5_important_metric$solar_share_elec +
df_2021_top_5_important_metric$wind_share_elec

df_2021_top_5_important_metric <- pivot_longer(df_2021_top_5_important_metric, cols = c("biofuel_share_elec", "coal_share_elec", "gas_share_elec", "hydro_share_elec", "low_carbon_share_elec", "nuclear_share_elec", "oil_share_elec", "other_renewables_share_elec", "renewables_share_elec", "solar_share_elec", "wind_share_elec"),
  value_name = "percentage")

ggplot(df_2021_top_5_important_metric, aes(fill=electric_share_type, y=percentage, x=country)) +
  geom_bar(position="fill", stat="identity", color = "black") +
  scale_fill_viridis_d(option = "turbo") +
  scale_y_continuous(labels = scales::percent) +
  labs(title = "Electricity generation constitution of 5 most highest GDP in 2021", x = "Country", y = "percentage")

```

Electricity generation constitution of 5 most highest GDP in 2021



```
df_2021_categorize <- df_2021[df_2021$country == "High-income countries" | df_2021$country == "Low-income countries"]

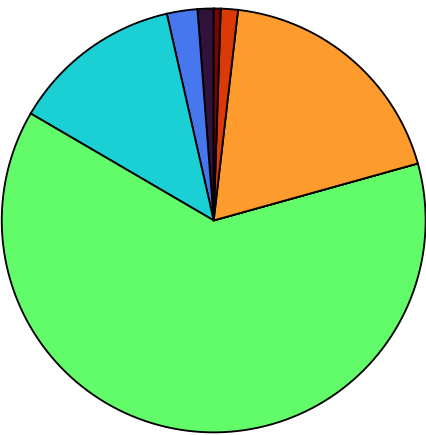
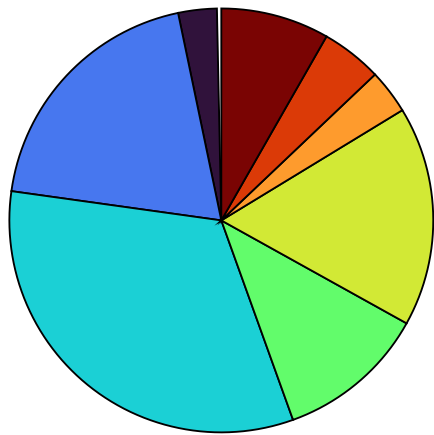
df_2021_categorize_important_metric <- pivot_longer(df_2021_categorize, cols = c("biofuel_share_elec", "coal_share_elec", "gas_share_elec", "hydro_share_elec", "nuclear_share_elec", "oil_share_elec", "solar_share_elec", "wind_share_elec"), values_to = "percentage")

ggplot(df_2021_categorize_important_metric, aes(x = "", y=percentage, fill=electric_share_type)) + geom_bar() +
  scale_fill_viridis_d(option = "turbo") +
  facet_wrap(~ country, nrow = 2) +
  theme_void() +
  theme(legend.position = "bottom", plot.title = element_text(hjust = .5), text=element_text(face="bold", size=12)) +
  coord_polar("y", start=0) +
  labs(title = "Electricity generation constitution of countries categorized by income in 2023", fill = "electric_share_type")
```

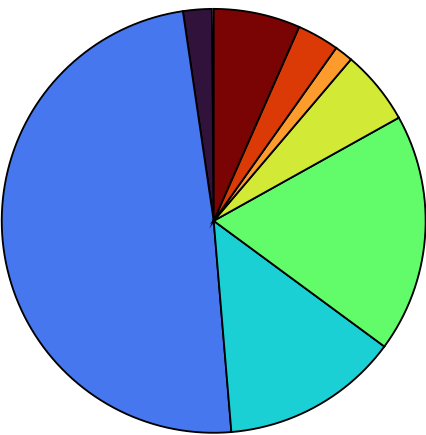
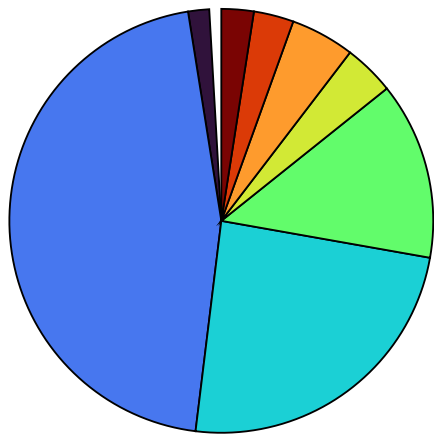
Eletricity generation constitution of countries categorized by income in 2023

High-income countries

Low-income countries



Lower-middle-income countries Upper-middle-income countries



Eletricity generation type

biofuel_share_elec	gas_share_elec	nuclear_share_elec	solar_share_elec
coal_share_elec	hydro_share_elec	oil_share_elec	wind_share_elec

Question 2