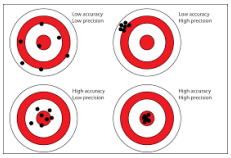
# Devul Assignment 1

Q1. **What are the <u>main features</u> of the variables in the data set?**

1.1Provide a compact description of the variables involved, 1.2their nature (categorical/numerical), 1.3special features, and 1.4their mutual relationship. 1.5Complement with appropriate visualization.



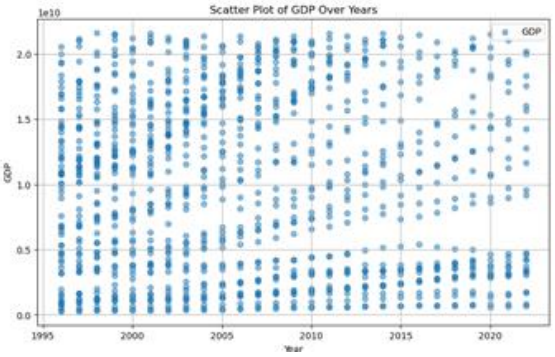
1.1 The variables are evenly distributed: All countries have an equal number of years in the table, although some cells contain missing values. It includes both categorical (qualitative) and continuous (numerical-quantitative) data, with continuous variables ranging from floating point numbers to large integers for GDP. There are multiple instances of missing information marked as NA, which will affect the quality and accuracy of the final data analysis. The front infographic is worth considering in all types of data analysis.

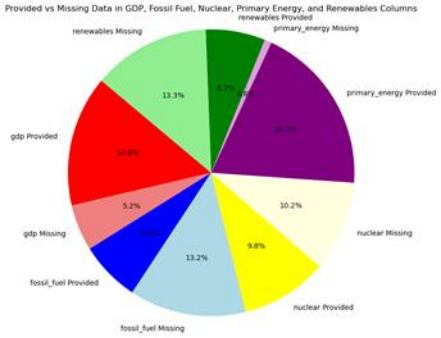
1.2 Population, primary_energy, fossil_fuel, nuclear, renewables, and GDP are numerical (quantitative). Year as Discreate Categorical, iso_code and name of countries cases these are nominal categorical data sets (qualitative discrete). While name of countries can be sorted alphabetically, but firstly there's no natural order even on the map secondly some lists change order for ease of their applications. There's no natural less or greater relation between countries so, it's not ordinal like small, medium and big.

1.3 population is medium integer, GPDs are large integers and primary_energy is floating point data type. In many numerical cases, N/A has shown up which is a categorical data. Table has been sorted by the alphabetical orders of the name of countries (and in many cases, iso_code except HRV and GBR). The 2nd layer order is the year for each country. The dataset contains some columns with missing values, such as fossil_fuel, nuclear, and renewables. Additionally,
the primary_energy column has negative values, which might indicate data errors or special cases. The year column provides temporal data, allowing for time-series analysis.

1.4 Generally, countries with larger populations tend to have higher GDPs. Higher GDPs often correlate with higher energy consumption across all energy types. It looks like data of nuclear energy assumption isn't provided enough as we see in the pie chart, maybe because of being noble, or even because of its sensitivity accessing to these type of data is harder.

1.5 Visualization



Q2. **How is the quality of the data of low GDP countries?**

**From the main data set**, 2.1create a new data set containing only the measurements of the countries with gdp smaller than the first quartile.  Prof: this is a hint

2.2Provide a compact description of outliers, 2.3unusual values, and 2.4missing values of the variables in the new data set. 2.5Do you see any quality issues? 2.6Complement with appropriate visualization.

2.2 Identify outliers is by using the Interquartile Range (IQR) method: 1.5 coefficient for outliers and 3 or more for extreme outliers. Outliers in the dataset have been identified in several columns. In the population column, outliers are present in the following rows: Afghanistan from 1997 to 2003, Madagascar in 2003, Mozambique from 2000 to 2012, and Niger from 2012 to 2019. For the fossil fuel column, outliers are found in Trinidad and Tobago for the years 1996 to 1997, and in Turkmenistan for the year 1996. In the primary energy column, outliers are observed in Armenia for the year 1996, Bahrain from 1996 to 1999, and Zimbabwe from 2005 to 2011. Outliers were found in Afghanistan, Madagascar, Mozambique, and Niger for population; Trinidad and Tobago, and Turkmenistan for fossil fuel; and Armenia, Bahrain, and Zimbabwe for primary energy. No outliers were detected for GDP, nuclear, and renewables. There were no extreme outliers for population, GDP, nuclear, and renewables, while fossil fuel had 1 extreme outlier and primary energy had 20 extreme outliers.

2.3  I saw a -1 in Afghanistan row in 1996 at the beginning for primary energy. It's nonsense to have a negative number for primary energy, while all the positive or missing data are totally acceptable.

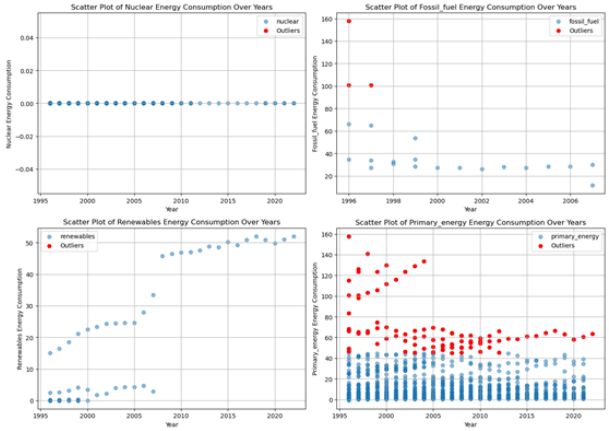
2.4 It's transparent that countries, iso codes, and years are those basic data so we don't expect missing because we have defined our enery.csv based on them. Happily, there's no missing in populations, but in other columns which means fossil_fuel, nuclear, primary_energy, and renewables, we see NA (missing) values. Please consider that I wrote my code such that I omitted NA in GDP columns at the beginning, because as we learned in week 2 and 3, this makes data analysis much less crowded of data and effective.

2.5 Sadly, lots of it! Based on the former questions and the following framework from module materials, I can response carefully:

**Missing values**: We have lots of them as I pointed out in 2.4.**Outliers**: As they existed as pointed above. **Measurement or data entry errors:** First data has -1 for primary enegy as I pointed out.

**Data coding errors**: I utilized CoPilot for a faster review. By having 2 smaller charts, I saw no issue.

2.6 Visualization



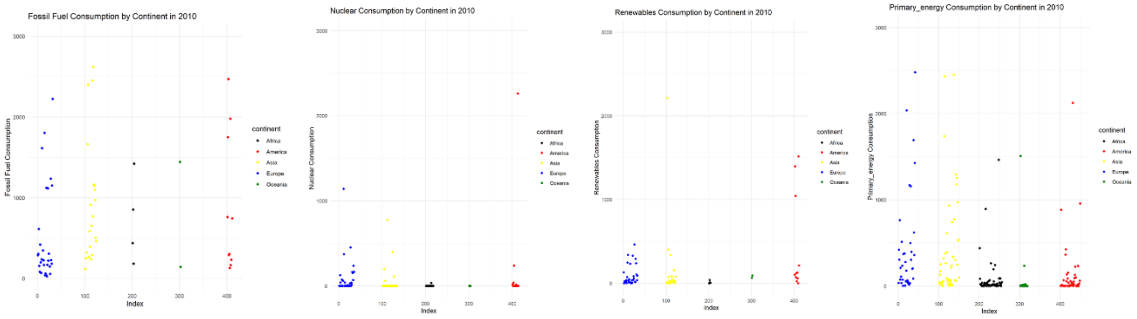
Q3. **<u>How</u> did the world source its energy in 2010?**

3.1Extract from your data set the data relevant to the year 2010 and 3.2provide a visual overview of energy consumption in the world at this time.

In particular, 3.3discuss possible regional features (similar for all countries in the same area) as well as 3.4features that are specific to countries that are located in different areas.

3.5 Give an interpretation of your findings and complement with appropriate visualization.

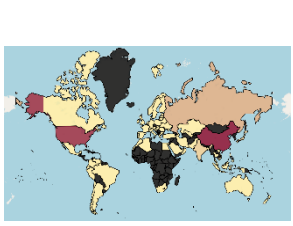
3.2 Scattered Plotting organizing continents to make easier data analysis



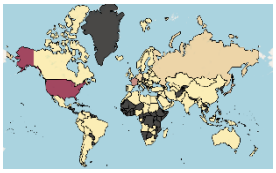

3.3 For some older resources of energy, including fossil fuel, and primary energy, consumption has a high variance as it's clear in the plots. But, for those which are newer, like nuclear energy and renewable energy, Europe and America has more variance and positive amounts rather than other continents. Moreover, these 2 resources of energy haven't appeared a lot in comparison with 2 other resources, as we see lesser points as a result of more NAs.

3.4 Europe has the highest variance Absolut value in all the 3 plots. African has a small variance in 2 new resources of energy which are nuclear and renewable or has a high variance as in fossil fuel and primary energy, no middle scenario occurred.

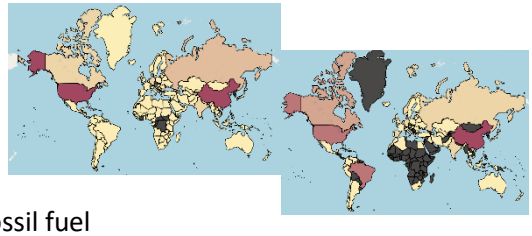
3.5 Below, we see a heatmap



Nuclear



Primary



Fossil fuel Renewable

**My findings show that Primary is the widest utilized resource of energy. It's hard to determine the least utilized form of the energy because missing data may change the game completely. Europe has the highest reported data, while Africa plays a vice versa rule.**

Q4. **Are there <u>effects of population and economy</u> on energy consumption?**

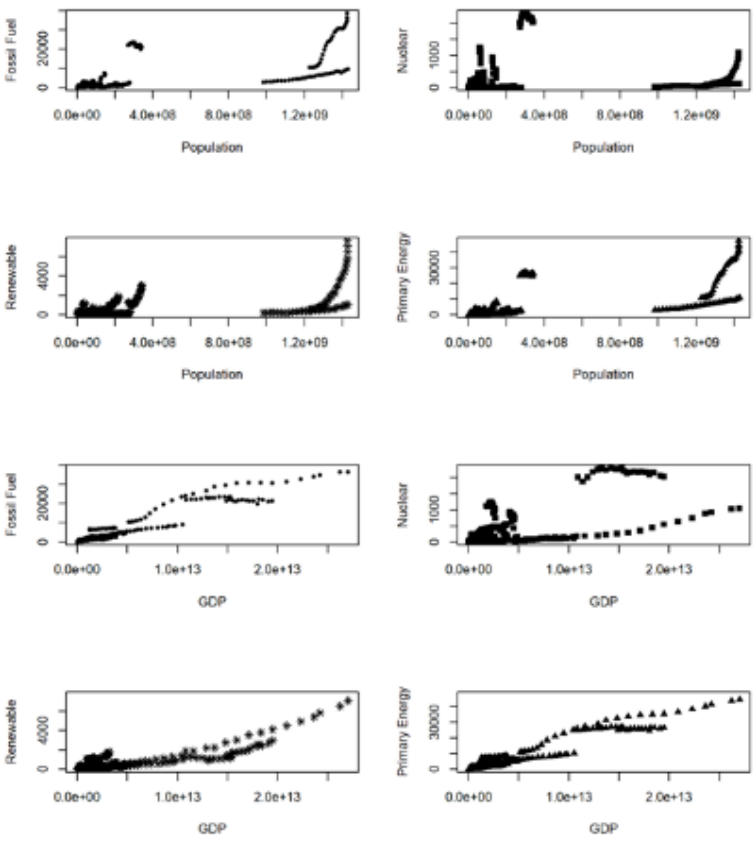
4.1Explore the association of the energy consumption variables with population and gdp.

4.2Give an interpretation of your findings and 4.3complement with appropriate visualization.

4.4Discuss limitations and implications for future analysis.

4.2 Having correlation syntax gives NA for all 4*2 results, but as we observe on the plot, there are sub relations there. When population is low, consuming is low or in another scenario, there's a group of data which shows a higher intention of usage. In GDP relationship, a general trend is direct relation between increasing GDP and increasing energy consumption.

4.4 The first restriction is NA values, which totally can change the density of the data. The second one is the development level of economy. Consider a country which is categorized as a developing country, so its data behavior for categorizing is clear. But, we don't know if and when it'd move into developed ones to change the analysis of this data. As we may see on the plot, in lower population countries sometime consumption is low sometime high. One of the factors of it is this unpredictable factor.



Q5. **How do energy consumption trends in South Korea and Japan compare over the years? Trend changing**

5.1Consider the data of all types of energy consumption in South Korea and Japan for all the years available.

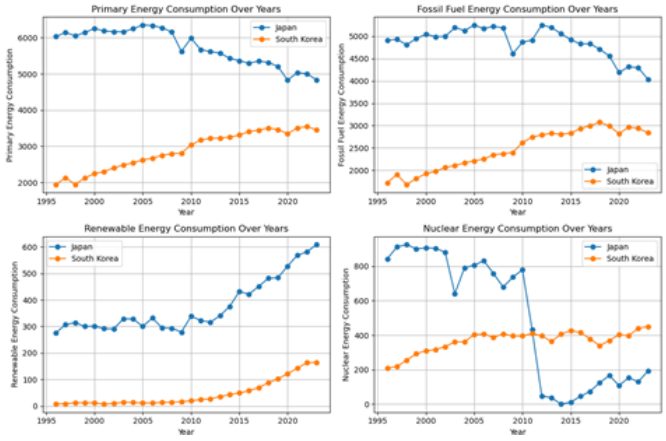
5.2Compare the **corresponding trends** giving an interpretation of your findings and
5.3complementing with appropriate visualization.

5.4Discuss limitations and implications for future analysis.

5.2 For visualizing, time series work here the best. Firstly, I analyze countries over 4 energy sources. Secondly, I compare each energy source between two countries. Japan declined fossil fuel, primary energy and Nuclear energy specifically after Fukushima disaster in 2010 and instead, increased Renewable energy. South Korea kept a general increment in utilizing all sorts of energy. For both countries, we see a great corresponding in trends of fossil fuel and primary energy usage, the trend of for instance fossil fuel and Primary Energy of Japan looks similar in changes over years.

In all types of energy utilizing except Nuclear energy, Japan has a total of more consumption. Fossil fuel and primary energy utilizing in both countries declined in 2020 due to pandemic.

5.3 Visualize




5.4 I can point out to potential tremendous changes like Fukushima disaster for prediction of utilizing Nuclear energy. In addition, the huge economical changes like pandemic affects the utilizing trend of energy. Including additional variables such as government policies, economic indicators, and technological advancements could provide a more comprehensive understanding of the factors influencing energy consumption.

**Note of Professor:**

**How do you analyze How do you show As name of Devul**

**You can visualize with more than 1, but up to 4 not more than that.**

**Appendix**

2.1

```
# Read the CSV file
> data <- read.csv("C:/Users/Parian/Desktop/energy.csv")
# Remove rows with NA values in the GDP column
> data <- data[!is.na(data$gdp), ]
# Calculate the first quartile (Q1) of the GDP column, removing any remaining missing values
> q1 <- quantile(data$gdp, probs = 0.25, na.rm = TRUE)
# Filter the data to include only rows where GDP is less than Q1
> filtered_data <- data[data$gdp < q1, , drop = FALSE]
> print(filtered_data)

# Load the dataset
> data <- read.csv("C:/Users/Parian/Desktop/energy.csv")
# Function to identify outliers
> identify_outliers <- function(df, column) {
  Q1 <- quantile(df[[column]], 0.25, na.rm = TRUE)
  Q3 <- quantile(df[[column]], 0.75, na.rm = TRUE)
  IQR <- Q3 - Q1
  lower_bound <- Q1 - 1.5 * IQR
  upper_bound <- Q3 + 1.5 * IQR
  return(df[df[[column]] < lower_bound | df[[column]] > upper_bound, ])
}
# Columns to check for outliers
```

```
columns_to_check <- c("population", "gdp", "fossil_fuel", "nuclear", "primary_energy",
"renewables")

# Identify outliers in each column

outliers <- lapply(columns_to_check, function(col) identify_outliers(data, col))

names(outliers) <- columns_to_check


# Display outliers

for (col in columns_to_check) {

  cat("Outliers in", col, ":\n")

  print(outliers[[col]])

}
```

Extreme Outliers

```
# Load the dataset

data <- read.csv("filtered_energy 2.csv")


# Function to identify outliers

identify_outliers <- function(df, column) {

  Q1 <- quantile(df[[column]], 0.25, na.rm = TRUE)

  Q3 <- quantile(df[[column]], 0.75, na.rm = TRUE)

  IQR <- Q3 - Q1

  lower_bound <- Q1 - 1.5 * IQR

  upper_bound <- Q3 + 1.5 * IQR

  return(df[df[[column]] < lower_bound | df[[column]] > upper_bound, ])

}


# Identify outliers in the primary_energy column

primary_energy_outliers <- identify_outliers(data, "primary_energy")


# Display the rows with outliers in the primary_energy column
```

```
print(primary_energy_outliers)



#Nuclear energy scattered plot
# Load necessary library
library(ggplot2)

# Read the CSV file
data <- read.csv("Filtered_data.csv")

# Filter out rows where nuclear energy consumption is NA
filtered_data <- data[!is.na(data$nuclear), ]

# Create the plot
ggplot(filtered_data, aes(x = year, y = nuclear)) +
  geom_point() +
  labs(title = "Nuclear Energy Consumption Over Years",
       x = "Year",
       y = "Nuclear Energy Consumption") +
  theme_minimal()




3.1
data <- read.csv("C:/Users/Parian/Desktop/energy.csv")
data_2010 <- data[data$year == 2010, ]
# Print the filtered data
print(data_2010)
write.csv(data_2010, "C:/Users/Parian/Desktop/year2010_energy.csv", row.names = FALSE)
```

```
4.1
data <- read.csv("C:/Users/Parian/Desktop/energy.csv")
par(mfrow =c(4,2))

#fossil fuel with population
cor(y=data$fossil_fuel, x=data$population)
plot(y=data$fossil_fuel, x=data$population, pch=20,
    ylab='Fossil Fuel',
    xlab='Population')
cor(y=data$nuclear, x=data$population)
plot(y=data$nuclear, x=data$population, pch=15,
    ylab='Nuclear',
    xlab='Population')
cor(y=data$renewable, x=data$population)
plot(y=data$renewable, x=data$population, pch=8,
    ylab='Renewable',
    xlab='Population')
cor(y=data$primary_energy, x=data$population)
plot(y=data$primary_energy, x=data$population, pch=17,
    ylab='Primary Energy',
    xlab='Population')

#GDP
cor(y=data$fossil_fuel, x=data$gdp)
plot(y=data$fossil_fuel, x=data$gdp, pch=20,
    ylab='Fossil Fuel',
    xlab='GDP')
cor(y=data$nuclear, x=data$gdp)
plot(y=data$nuclear, x=data$gdp, pch=15,
    ylab='Nuclear',
    xlab='GDP')
```

```
cor(y=data$renewable, x=data$gdp)
plot(y=data$renewable, x=data$gdp, pch=8,
    ylab='Renewable',
    xlab='GDP')
cor(y=data$primary_energy, x=data$gdp)
plot(y=data$primary_energy, x=data$gdp, pch=17,
    ylab='Primary Energy',
    xlab='GDP')
```

5.1

```
data <- read.csv("C:/Users/Parian/Desktop/energy.csv")
# Assuming your data has a column named 'country'
energy_data_sk_jp <- data[data$country %in% c("South Korea", "Japan"), ]
# Print the filtered data
print(energy_data_sk_jp)
write.csv(filtered_data, "C:/Users/Parian/Desktop/kj.csv", row.names = FALSE)
```

**3.1**

**For fossil_fuel**

```
# Load necessary libraries
library(ggplot2)
library(dplyr)

data <- read.csv("C:/Users/Parian/Desktop/year2010_energy.csv")

# Omit rows with NA values in the fossil_fuel column
data <- data[!is.na(data$fossil_fuel), ]

# Define a comprehensive mapping of countries to continents
continent_mapping <- c(
```

**'Afghanistan' = 'Asia', 'Albania' = 'Europe', 'Algeria' = 'Africa', 'American Samoa' = 'Oceania', 'Angola' = 'Africa',**

**'Antarctica' = 'Antarctica', 'Antigua and Barbuda' = 'America', 'Argentina' = 'America', 'Armenia' = 'Asia', 'Aruba' = 'America',**

**'Australia' = 'Oceania', 'Austria' = 'Europe', 'Azerbaijan' = 'Asia', 'Bahamas' = 'America', 'Bahrain' = 'Asia',**

**'Bangladesh' = 'Asia', 'Barbados' = 'America', 'Belarus' = 'Europe', 'Belgium' = 'Europe', 'Belize' = 'America',**

**'Benin' = 'Africa', 'Bermuda' = 'America', 'Bhutan' = 'Asia', 'Bolivia' = 'America', 'Bosnia and Herzegovina' = 'Europe',**

**'Botswana' = 'Africa', 'Brazil' = 'America', 'British Virgin Islands' = 'America', 'Brunei' = 'Asia', 'Bulgaria' = 'Europe',**

**'Burkina Faso' = 'Africa', 'Burundi' = 'Africa', 'Cambodia' = 'Asia', 'Cameroon' = 'Africa', 'Canada' = 'America',**

**'Cape Verde' = 'Africa', 'Cayman Islands' = 'America', 'Central African Republic' = 'Africa', 'Chad' = 'Africa', 'Chile' = 'America',**

**'China' = 'Asia', 'Colombia' = 'America', 'Comoros' = 'Africa', 'Congo' = 'Africa', 'Cook Islands' = 'Oceania',**

**'Costa Rica' = 'America', 'Cote d\'Ivoire' = 'Africa', 'Croatia' = 'Europe', 'Cuba' = 'America', 'Cyprus' = 'Europe',**

**'Czechia' = 'Europe', 'Democratic Republic of Congo' = 'Africa', 'Denmark' = 'Europe', 'Djibouti' = 'Africa', 'Dominica' = 'America',**

**'Dominican Republic' = 'America', 'East Timor' = 'Asia', 'Ecuador' = 'America', 'Egypt' = 'Africa', 'El Salvador' = 'America',**

**'Equatorial Guinea' = 'Africa', 'Eritrea' = 'Africa', 'Estonia' = 'Europe', 'Eswatini' = 'Africa', 'Ethiopia' = 'Africa',**

**'Falkland Islands' = 'America', 'Faroe Islands' = 'Europe', 'Fiji' = 'Oceania', 'Finland' = 'Europe', 'France' = 'Europe',**

**'French Guiana' = 'America', 'French Polynesia' = 'Oceania', 'Gabon' = 'Africa', 'Gambia' = 'Africa', 'Georgia' = 'Asia',**

**'Germany' = 'Europe', 'Ghana' = 'Africa', 'Gibraltar' = 'Europe', 'Greece' = 'Europe', 'Greenland' = 'Europe',**

**'Grenada' = 'America', 'Guadeloupe' = 'America', 'Guam' = 'Oceania', 'Guatemala' = 'America', 'Guinea' = 'Africa',**

**'Guinea-Bissau' = 'Africa', 'Guyana' = 'America', 'Haiti' = 'America', 'Honduras' = 'America', 'Hong Kong' = 'Asia',**

**'Hungary' = 'Europe', 'Iceland' = 'Europe', 'India' = 'Asia', 'Indonesia' = 'Asia', 'Iran' = 'Asia',**

**'Iraq' = 'Asia', 'Ireland' = 'Europe', 'Israel' = 'Asia', 'Italy' = 'Europe', 'Jamaica' = 'America',**

**'Japan' = 'Asia', 'Jordan' = 'Asia', 'Kazakhstan' = 'Asia', 'Kenya' = 'Africa', 'Kiribati' = 'Oceania',**

**'Kuwait' = 'Asia', 'Kyrgyzstan' = 'Asia', 'Laos' = 'Asia', 'Latvia' = 'Europe', 'Lebanon' = 'Asia',**

**'Lesotho' = 'Africa', 'Liberia' = 'Africa', 'Libya' = 'Africa', 'Lithuania' = 'Europe', 'Luxembourg' = 'Europe',**

**'Macao' = 'Asia', 'Madagascar' = 'Africa', 'Malawi' = 'Africa', 'Malaysia' = 'Asia', 'Maldives' = 'Asia',**

**'Mali' = 'Africa', 'Malta' = 'Europe', 'Martinique' = 'America', 'Mauritania' = 'Africa', 'Mauritius' = 'Africa',**

**'Mexico' = 'America', 'Micronesia (country)' = 'Oceania', 'Moldova' = 'Europe', 'Mongolia' = 'Asia', 'Montenegro' = 'Europe',**

**'Montserrat' = 'America', 'Morocco' = 'Africa', 'Mozambique' = 'Africa', 'Myanmar' = 'Asia', 'Namibia' = 'Africa',**

**'Nauru' = 'Oceania', 'Nepal' = 'Asia', 'Netherlands' = 'Europe', 'Netherlands Antilles' = 'America', 'New Caledonia' = 'Oceania',**

**'New Zealand' = 'Oceania', 'Nicaragua' = 'America', 'Niger' = 'Africa', 'Nigeria' = 'Africa', 'Niue' = 'Oceania',**

**'North Korea' = 'Asia', 'North Macedonia' = 'Europe', 'Northern Mariana Islands' = 'Oceania', 'Norway' = 'Europe', 'Oman' = 'Asia',**

**'Pakistan' = 'Asia', 'Palau' = 'Oceania', 'Palestine' = 'Asia', 'Panama' = 'America', 'Papua New Guinea' = 'Oceania',**

**'Paraguay' = 'America', 'Peru' = 'America', 'Philippines' = 'Asia', 'Poland' = 'Europe', 'Portugal' = 'Europe',**

**'Puerto Rico' = 'America', 'Qatar' = 'Asia', 'Reunion' = 'Africa', 'Romania' = 'Europe', 'Russia' = 'Europe',**

**'Rwanda' = 'Africa', 'Saint Helena' = 'Africa', 'Saint Kitts and Nevis' = 'America', 'Saint Lucia' = 'America', 'Saint Pierre and Miquelon' = 'America',**

**'Saint Vincent and the Grenadines' = 'America', 'Samoa' = 'Oceania', 'Sao Tome and Principe' = 'Africa', 'Saudi Arabia' = 'Asia', 'Senegal' = 'Africa',**

**'Serbia' = 'Europe', 'Seychelles' = 'Africa', 'Sierra Leone' = 'Africa', 'Singapore' = 'Asia', 'Slovakia' = 'Europe',**

**'Slovenia' = 'Europe', 'Solomon Islands' = 'Oceania', 'Somalia' = 'Africa', 'South Africa' = 'Africa', 'South Korea' = 'Asia',**

**'South Sudan' = 'Africa', 'Spain' = 'Europe', 'Sri Lanka' = 'Asia', 'Sudan' = 'Africa', 'Suriname' = 'America',**

**'Sweden' = 'Europe', 'Switzerland' = 'Europe', 'Syria' = 'Asia', 'Taiwan' = 'Asia', 'Tajikistan' = 'Asia',**

**'Tanzania' = 'Africa', 'Thailand' = 'Asia', 'Togo' = 'Africa', 'Tonga' = 'Oceania', 'Trinidad and Tobago' = 'America',**

```
  'Tunisia' = 'Africa', 'Turkey' = 'Asia', 'Turkmenistan' = 'Asia', 'Turks and Caicos Islands' = 'America',
'Tuvalu' = 'Oceania',

  'Uganda' = 'Africa', 'Ukraine' = 'Europe', 'United Arab Emirates' = 'Asia', 'United Kingdom' =
'Europe', 'United States' = 'America',

  'United States Virgin Islands' = 'America', 'Uruguay' = 'America', 'Uzbekistan' = 'Asia', 'Vanuatu' =
'Oceania', 'Vatican City' = 'Europe',

  'Venezuela' = 'America', 'Vietnam' = 'Asia', 'Western Sahara' = 'Africa', 'Yemen' = 'Asia', 'Zambia' =
'Africa', 'Zimbabwe' = 'Africa'

)# Map countries to continents

data$continent <- continent_mapping[data$country]


# Define colors for continents

continent_colors <- c(

  'Europe' = '#0000FF',  # Blue

  'Asia' = '#FFFF00',    # Yellow

  'Africa' = '#000000',  # Black

  'Oceania' = '#008000', # Green

  'America' = '#FF0000'  # Red

)



# Create a new column for the X-axis positions based on continent groups

data <- data %>%

  group_by(continent) %>%

  mutate(x_pos = row_number())


# Adjust the X-axis positions to have spaces between continents

continent_offsets <- c(

  'Europe' = 0,

  'Asia' = 100,

  'Africa' = 200,

  'Oceania' = 300,
```

```
  'America' = 400
)
data$x_pos <- data$x_pos + continent_offsets[data$continent]

# Create scatter plot for fossil fuel consumption, categorizing by continent
ggplot(data, aes(x = x_pos, y = fossil_fuel, color = continent)) +
  geom_point(shape = 16) +
  scale_color_manual(values = continent_colors) +
  labs(x = 'Index', y = 'Fossil Fuel Consumption', title = 'Fossil Fuel Consumption by Continent in 2010') +
  theme_minimal() +
  ylim(0, 3000)  # Set Y-axis range to be visible for all data
```

for nuclear

```
# Load necessary libraries
library(ggplot2)
library(dplyr)

data <- read.csv("C:/Users/Parian/Desktop/year2010_energy.csv")

# Omit rows with NA values in the nuclear column
data <- data[!is.na(data$nuclear), ]

# Define a comprehensive mapping of countries to continents
continent_mapping <- c(
  'Afghanistan' = 'Asia', 'Albania' = 'Europe', 'Algeria' = 'Africa', 'American Samoa' = 'Oceania', 'Angola' = 'Africa',
  'Antarctica' = 'Antarctica', 'Antigua and Barbuda' = 'America', 'Argentina' = 'America', 'Armenia' = 'Asia', 'Aruba' = 'America',
```

'Australia' = 'Oceania', 'Austria' = 'Europe', 'Azerbaijan' = 'Asia', 'Bahamas' = 'America', 'Bahrain' = 'Asia',

'Bangladesh' = 'Asia', 'Barbados' = 'America', 'Belarus' = 'Europe', 'Belgium' = 'Europe', 'Belize' = 'America',

'Benin' = 'Africa', 'Bermuda' = 'America', 'Bhutan' = 'Asia', 'Bolivia' = 'America', 'Bosnia and Herzegovina' = 'Europe',

'Botswana' = 'Africa', 'Brazil' = 'America', 'British Virgin Islands' = 'America', 'Brunei' = 'Asia', 'Bulgaria' = 'Europe',

'Burkina Faso' = 'Africa', 'Burundi' = 'Africa', 'Cambodia' = 'Asia', 'Cameroon' = 'Africa', 'Canada' = 'America',

'Cape Verde' = 'Africa', 'Cayman Islands' = 'America', 'Central African Republic' = 'Africa', 'Chad' = 'Africa', 'Chile' = 'America',

'China' = 'Asia', 'Colombia' = 'America', 'Comoros' = 'Africa', 'Congo' = 'Africa', 'Cook Islands' = 'Oceania',

'Costa Rica' = 'America', 'Cote d\'Ivoire' = 'Africa', 'Croatia' = 'Europe', 'Cuba' = 'America', 'Cyprus' = 'Europe',

'Czechia' = 'Europe', 'Democratic Republic of Congo' = 'Africa', 'Denmark' = 'Europe', 'Djibouti' = 'Africa', 'Dominica' = 'America',

'Dominican Republic' = 'America', 'East Timor' = 'Asia', 'Ecuador' = 'America', 'Egypt' = 'Africa', 'El Salvador' = 'America',

'Equatorial Guinea' = 'Africa', 'Eritrea' = 'Africa', 'Estonia' = 'Europe', 'Eswatini' = 'Africa', 'Ethiopia' = 'Africa',

'Falkland Islands' = 'America', 'Faroe Islands' = 'Europe', 'Fiji' = 'Oceania', 'Finland' = 'Europe', 'France' = 'Europe',

'French Guiana' = 'America', 'French Polynesia' = 'Oceania', 'Gabon' = 'Africa', 'Gambia' = 'Africa', 'Georgia' = 'Asia',

'Germany' = 'Europe', 'Ghana' = 'Africa', 'Gibraltar' = 'Europe', 'Greece' = 'Europe', 'Greenland' = 'Europe',

'Grenada' = 'America', 'Guadeloupe' = 'America', 'Guam' = 'Oceania', 'Guatemala' = 'America', 'Guinea' = 'Africa',

'Guinea-Bissau' = 'Africa', 'Guyana' = 'America', 'Haiti' = 'America', 'Honduras' = 'America', 'Hong Kong' = 'Asia',

'Hungary' = 'Europe', 'Iceland' = 'Europe', 'India' = 'Asia', 'Indonesia' = 'Asia', 'Iran' = 'Asia',

'Iraq' = 'Asia', 'Ireland' = 'Europe', 'Israel' = 'Asia', 'Italy' = 'Europe', 'Jamaica' = 'America',

'Japan' = 'Asia', 'Jordan' = 'Asia', 'Kazakhstan' = 'Asia', 'Kenya' = 'Africa', 'Kiribati' = 'Oceania',

'Kuwait' = 'Asia', 'Kyrgyzstan' = 'Asia', 'Laos' = 'Asia', 'Latvia' = 'Europe', 'Lebanon' = 'Asia',

**'Lesotho' = 'Africa', 'Liberia' = 'Africa', 'Libya' = 'Africa', 'Lithuania' = 'Europe', 'Luxembourg' = 'Europe',**

**'Macao' = 'Asia', 'Madagascar' = 'Africa', 'Malawi' = 'Africa', 'Malaysia' = 'Asia', 'Maldives' = 'Asia',**

**'Mali' = 'Africa', 'Malta' = 'Europe', 'Martinique' = 'America', 'Mauritania' = 'Africa', 'Mauritius' = 'Africa',**

**'Mexico' = 'America', 'Micronesia (country)' = 'Oceania', 'Moldova' = 'Europe', 'Mongolia' = 'Asia', 'Montenegro' = 'Europe',**

**'Montserrat' = 'America', 'Morocco' = 'Africa', 'Mozambique' = 'Africa', 'Myanmar' = 'Asia', 'Namibia' = 'Africa',**

**'Nauru' = 'Oceania', 'Nepal' = 'Asia', 'Netherlands' = 'Europe', 'Netherlands Antilles' = 'America', 'New Caledonia' = 'Oceania',**

**'New Zealand' = 'Oceania', 'Nicaragua' = 'America', 'Niger' = 'Africa', 'Nigeria' = 'Africa', 'Niue' = 'Oceania',**

**'North Korea' = 'Asia', 'North Macedonia' = 'Europe', 'Northern Mariana Islands' = 'Oceania', 'Norway' = 'Europe', 'Oman' = 'Asia',**

**'Pakistan' = 'Asia', 'Palau' = 'Oceania', 'Palestine' = 'Asia', 'Panama' = 'America', 'Papua New Guinea' = 'Oceania',**

**'Paraguay' = 'America', 'Peru' = 'America', 'Philippines' = 'Asia', 'Poland' = 'Europe', 'Portugal' = 'Europe',**

**'Puerto Rico' = 'America', 'Qatar' = 'Asia', 'Reunion' = 'Africa', 'Romania' = 'Europe', 'Russia' = 'Europe',**

**'Rwanda' = 'Africa', 'Saint Helena' = 'Africa', 'Saint Kitts and Nevis' = 'America', 'Saint Lucia' = 'America', 'Saint Pierre and Miquelon' = 'America',**

**'Saint Vincent and the Grenadines' = 'America', 'Samoa' = 'Oceania', 'Sao Tome and Principe' = 'Africa', 'Saudi Arabia' = 'Asia', 'Senegal' = 'Africa',**

**'Serbia' = 'Europe', 'Seychelles' = 'Africa', 'Sierra Leone' = 'Africa', 'Singapore' = 'Asia', 'Slovakia' = 'Europe',**

**'Slovenia' = 'Europe', 'Solomon Islands' = 'Oceania', 'Somalia' = 'Africa', 'South Africa' = 'Africa', 'South Korea' = 'Asia',**

**'South Sudan' = 'Africa', 'Spain' = 'Europe', 'Sri Lanka' = 'Asia', 'Sudan' = 'Africa', 'Suriname' = 'America',**

**'Sweden' = 'Europe', 'Switzerland' = 'Europe', 'Syria' = 'Asia', 'Taiwan' = 'Asia', 'Tajikistan' = 'Asia',**

**'Tanzania' = 'Africa', 'Thailand' = 'Asia', 'Togo' = 'Africa', 'Tonga' = 'Oceania', 'Trinidad and Tobago' = 'America',**

**'Tunisia' = 'Africa', 'Turkey' = 'Asia', 'Turkmenistan' = 'Asia', 'Turks and Caicos Islands' = 'America', 'Tuvalu' = 'Oceania',**

```
  'Uganda' = 'Africa', 'Ukraine' = 'Europe', 'United Arab Emirates' = 'Asia', 'United Kingdom' = 'Europe', 'United States' = 'America',
  'United States Virgin Islands' = 'America', 'Uruguay' = 'America', 'Uzbekistan' = 'Asia', 'Vanuatu' = 'Oceania', 'Vatican City' = 'Europe',
  'Venezuela' = 'America', 'Vietnam' = 'Asia', 'Western Sahara' = 'Africa', 'Yemen' = 'Asia', 'Zambia' = 'Africa', 'Zimbabwe' = 'Africa'
)# Map countries to continents
data$continent <- continent_mapping[data$country]

# Define colors for continents
continent_colors <- c(
  'Europe' = '#0000FF',  # Blue
  'Asia' = '#FFFF00',    # Yellow
  'Africa' = '#000000',  # Black
  'Oceania' = '#008000', # Green
  'America' = '#FF0000'  # Red
)


# Create a new column for the X-axis positions based on continent groups
data <- data %>%
  group_by(continent) %>%
  mutate(x_pos = row_number())

# Adjust the X-axis positions to have spaces between continents
continent_offsets <- c(
  'Europe' = 0,
  'Asia' = 100,
  'Africa' = 200,
  'Oceania' = 300,
  'America' = 400
)
```

```
data$x_pos <- data$x_pos + continent_offsets[data$continent]

# Create scatter plot for nuclear consumption, categorizing by continent
ggplot(data, aes(x = x_pos, y = nuclear, color = continent)) +
  geom_point(shape = 16) +
  scale_color_manual(values = continent_colors) +
  labs(x = 'Index', y = 'Nuclear Consumption', title = 'Nuclear Consumption by Continent in 2010') +
  theme_minimal() +
  ylim(0, 3000)  # Set Y-axis range to be visible for all data

for renewables

# Load necessary libraries
library(ggplot2)
library(dplyr)

data <- read.csv("C:/Users/Parian/Desktop/year2010_energy.csv")

# Omit rows with NA values in the renewables column
data <- data[!is.na(data$renewables), ]

# Define a comprehensive mapping of countries to continents
continent_mapping <- c(
  'Afghanistan' = 'Asia', 'Albania' = 'Europe', 'Algeria' = 'Africa', 'American Samoa' = 'Oceania',
'Angola' = 'Africa',
  'Antarctica' = 'Antarctica', 'Antigua and Barbuda' = 'America', 'Argentina' = 'America', 'Armenia' =
'Asia', 'Aruba' = 'America',
  'Australia' = 'Oceania', 'Austria' = 'Europe', 'Azerbaijan' = 'Asia', 'Bahamas' = 'America', 'Bahrain' =
'Asia',
  'Bangladesh' = 'Asia', 'Barbados' = 'America', 'Belarus' = 'Europe', 'Belgium' = 'Europe', 'Belize' =
'America',
  'Benin' = 'Africa', 'Bermuda' = 'America', 'Bhutan' = 'Asia', 'Bolivia' = 'America', 'Bosnia and
Herzegovina' = 'Europe',
```

**'Botswana' = 'Africa', 'Brazil' = 'America', 'British Virgin Islands' = 'America', 'Brunei' = 'Asia', 'Bulgaria' = 'Europe',**

**'Burkina Faso' = 'Africa', 'Burundi' = 'Africa', 'Cambodia' = 'Asia', 'Cameroon' = 'Africa', 'Canada' = 'America',**

**'Cape Verde' = 'Africa', 'Cayman Islands' = 'America', 'Central African Republic' = 'Africa', 'Chad' = 'Africa', 'Chile' = 'America',**

**'China' = 'Asia', 'Colombia' = 'America', 'Comoros' = 'Africa', 'Congo' = 'Africa', 'Cook Islands' = 'Oceania',**

**'Costa Rica' = 'America', 'Cote d\'Ivoire' = 'Africa', 'Croatia' = 'Europe', 'Cuba' = 'America', 'Cyprus' = 'Europe',**

**'Czechia' = 'Europe', 'Democratic Republic of Congo' = 'Africa', 'Denmark' = 'Europe', 'Djibouti' = 'Africa', 'Dominica' = 'America',**

**'Dominican Republic' = 'America', 'East Timor' = 'Asia', 'Ecuador' = 'America', 'Egypt' = 'Africa', 'El Salvador' = 'America',**

**'Equatorial Guinea' = 'Africa', 'Eritrea' = 'Africa', 'Estonia' = 'Europe', 'Eswatini' = 'Africa', 'Ethiopia' = 'Africa',**

**'Falkland Islands' = 'America', 'Faroe Islands' = 'Europe', 'Fiji' = 'Oceania', 'Finland' = 'Europe', 'France' = 'Europe',**

**'French Guiana' = 'America', 'French Polynesia' = 'Oceania', 'Gabon' = 'Africa', 'Gambia' = 'Africa', 'Georgia' = 'Asia',**

**'Germany' = 'Europe', 'Ghana' = 'Africa', 'Gibraltar' = 'Europe', 'Greece' = 'Europe', 'Greenland' = 'Europe',**

**'Grenada' = 'America', 'Guadeloupe' = 'America', 'Guam' = 'Oceania', 'Guatemala' = 'America', 'Guinea' = 'Africa',**

**'Guinea-Bissau' = 'Africa', 'Guyana' = 'America', 'Haiti' = 'America', 'Honduras' = 'America', 'Hong Kong' = 'Asia',**

**'Hungary' = 'Europe', 'Iceland' = 'Europe', 'India' = 'Asia', 'Indonesia' = 'Asia', 'Iran' = 'Asia',**

**'Iraq' = 'Asia', 'Ireland' = 'Europe', 'Israel' = 'Asia', 'Italy' = 'Europe', 'Jamaica' = 'America',**

**'Japan' = 'Asia', 'Jordan' = 'Asia', 'Kazakhstan' = 'Asia', 'Kenya' = 'Africa', 'Kiribati' = 'Oceania',**

**'Kuwait' = 'Asia', 'Kyrgyzstan' = 'Asia', 'Laos' = 'Asia', 'Latvia' = 'Europe', 'Lebanon' = 'Asia',**

**'Lesotho' = 'Africa', 'Liberia' = 'Africa', 'Libya' = 'Africa', 'Lithuania' = 'Europe', 'Luxembourg' = 'Europe',**

**'Macao' = 'Asia', 'Madagascar' = 'Africa', 'Malawi' = 'Africa', 'Malaysia' = 'Asia', 'Maldives' = 'Asia',**

**'Mali' = 'Africa', 'Malta' = 'Europe', 'Martinique' = 'America', 'Mauritania' = 'Africa', 'Mauritius' = 'Africa',**

**'Mexico' = 'America', 'Micronesia (country)' = 'Oceania', 'Moldova' = 'Europe', 'Mongolia' = 'Asia', 'Montenegro' = 'Europe',**

**'Montserrat' = 'America', 'Morocco' = 'Africa', 'Mozambique' = 'Africa', 'Myanmar' = 'Asia', 'Namibia' = 'Africa',**

**'Nauru' = 'Oceania', 'Nepal' = 'Asia', 'Netherlands' = 'Europe', 'Netherlands Antilles' = 'America', 'New Caledonia' = 'Oceania',**

**'New Zealand' = 'Oceania', 'Nicaragua' = 'America', 'Niger' = 'Africa', 'Nigeria' = 'Africa', 'Niue' = 'Oceania',**

**'North Korea' = 'Asia', 'North Macedonia' = 'Europe', 'Northern Mariana Islands' = 'Oceania', 'Norway' = 'Europe', 'Oman' = 'Asia',**

**'Pakistan' = 'Asia', 'Palau' = 'Oceania', 'Palestine' = 'Asia', 'Panama' = 'America', 'Papua New Guinea' = 'Oceania',**

**'Paraguay' = 'America', 'Peru' = 'America', 'Philippines' = 'Asia', 'Poland' = 'Europe', 'Portugal' = 'Europe',**

**'Puerto Rico' = 'America', 'Qatar' = 'Asia', 'Reunion' = 'Africa', 'Romania' = 'Europe', 'Russia' = 'Europe',**

**'Rwanda' = 'Africa', 'Saint Helena' = 'Africa', 'Saint Kitts and Nevis' = 'America', 'Saint Lucia' = 'America', 'Saint Pierre and Miquelon' = 'America',**

**'Saint Vincent and the Grenadines' = 'America', 'Samoa' = 'Oceania', 'Sao Tome and Principe' = 'Africa', 'Saudi Arabia' = 'Asia', 'Senegal' = 'Africa',**

**'Serbia' = 'Europe', 'Seychelles' = 'Africa', 'Sierra Leone' = 'Africa', 'Singapore' = 'Asia', 'Slovakia' = 'Europe',**

**'Slovenia' = 'Europe', 'Solomon Islands' = 'Oceania', 'Somalia' = 'Africa', 'South Africa' = 'Africa', 'South Korea' = 'Asia',**

**'South Sudan' = 'Africa', 'Spain' = 'Europe', 'Sri Lanka' = 'Asia', 'Sudan' = 'Africa', 'Suriname' = 'America',**

**'Sweden' = 'Europe', 'Switzerland' = 'Europe', 'Syria' = 'Asia', 'Taiwan' = 'Asia', 'Tajikistan' = 'Asia',**

**'Tanzania' = 'Africa', 'Thailand' = 'Asia', 'Togo' = 'Africa', 'Tonga' = 'Oceania', 'Trinidad and Tobago' = 'America',**

**'Tunisia' = 'Africa', 'Turkey' = 'Asia', 'Turkmenistan' = 'Asia', 'Turks and Caicos Islands' = 'America', 'Tuvalu' = 'Oceania',**

**'Uganda' = 'Africa', 'Ukraine' = 'Europe', 'United Arab Emirates' = 'Asia', 'United Kingdom' = 'Europe', 'United States' = 'America',**

**'United States Virgin Islands' = 'America', 'Uruguay' = 'America', 'Uzbekistan' = 'Asia', 'Vanuatu' = 'Oceania', 'Vatican City' = 'Europe',**

**'Venezuela' = 'America', 'Vietnam' = 'Asia', 'Western Sahara' = 'Africa', 'Yemen' = 'Asia', 'Zambia' = 'Africa', 'Zimbabwe' = 'Africa'**

**)# Map countries to continents**

**data$continent <- continent_mapping[data$country]**

```
# Define colors for continents
continent_colors <- c(
  'Europe' = '#0000FF',  # Blue
  'Asia' = '#FFFF00',    # Yellow
  'Africa' = '#000000',  # Black
  'Oceania' = '#008000', # Green
  'America' = '#FF0000'  # Red
)


# Create a new column for the X-axis positions based on continent groups
data <- data %>%
  group_by(continent) %>%
  mutate(x_pos = row_number())

# Adjust the X-axis positions to have spaces between continents
continent_offsets <- c(
  'Europe' = 0,
  'Asia' = 100,
  'Africa' = 200,
  'Oceania' = 300,
  'America' = 400
)
data$x_pos <- data$x_pos + continent_offsets[data$continent]

# Create scatter plot for renewables consumption, categorizing by continent
ggplot(data, aes(x = x_pos, y = renewables, color = continent)) +
  geom_point(shape = 16) +
  scale_color_manual(values = continent_colors) +
```

```
  labs(x = 'Index', y = 'Renewables Consumption', title = 'Renewables Consumption by Continent in 2010') +
  theme_minimal() +
  ylim(0, 3000)  # Set Y-axis range to be visible for all data
```

for primary energy

```
# Load necessary libraries
library(ggplot2)
library(dplyr)

data <- read.csv("C:/Users/Parian/Desktop/year2010_energy.csv")

# Omit rows with NA values in the primary_energy column
data <- data[!is.na(data$primary_energy), ]

# Define a comprehensive mapping of countries to continents
continent_mapping <- c(
  'Afghanistan' = 'Asia', 'Albania' = 'Europe', 'Algeria' = 'Africa', 'American Samoa' = 'Oceania', 'Angola' = 'Africa',
  'Antarctica' = 'Antarctica', 'Antigua and Barbuda' = 'America', 'Argentina' = 'America', 'Armenia' = 'Asia', 'Aruba' = 'America',
  'Australia' = 'Oceania', 'Austria' = 'Europe', 'Azerbaijan' = 'Asia', 'Bahamas' = 'America', 'Bahrain' = 'Asia',
  'Bangladesh' = 'Asia', 'Barbados' = 'America', 'Belarus' = 'Europe', 'Belgium' = 'Europe', 'Belize' = 'America',
  'Benin' = 'Africa', 'Bermuda' = 'America', 'Bhutan' = 'Asia', 'Bolivia' = 'America', 'Bosnia and Herzegovina' = 'Europe',
  'Botswana' = 'Africa', 'Brazil' = 'America', 'British Virgin Islands' = 'America', 'Brunei' = 'Asia', 'Bulgaria' = 'Europe',
  'Burkina Faso' = 'Africa', 'Burundi' = 'Africa', 'Cambodia' = 'Asia', 'Cameroon' = 'Africa', 'Canada' = 'America',
  'Cape Verde' = 'Africa', 'Cayman Islands' = 'America', 'Central African Republic' = 'Africa', 'Chad' = 'Africa', 'Chile' = 'America',
```

'China' = 'Asia', 'Colombia' = 'America', 'Comoros' = 'Africa', 'Congo' = 'Africa', 'Cook Islands' = 'Oceania',

'Costa Rica' = 'America', 'Cote d\'Ivoire' = 'Africa', 'Croatia' = 'Europe', 'Cuba' = 'America', 'Cyprus' = 'Europe',

'Czechia' = 'Europe', 'Democratic Republic of Congo' = 'Africa', 'Denmark' = 'Europe', 'Djibouti' = 'Africa', 'Dominica' = 'America',

'Dominican Republic' = 'America', 'East Timor' = 'Asia', 'Ecuador' = 'America', 'Egypt' = 'Africa', 'El Salvador' = 'America',

'Equatorial Guinea' = 'Africa', 'Eritrea' = 'Africa', 'Estonia' = 'Europe', 'Eswatini' = 'Africa', 'Ethiopia' = 'Africa',

'Falkland Islands' = 'America', 'Faroe Islands' = 'Europe', 'Fiji' = 'Oceania', 'Finland' = 'Europe', 'France' = 'Europe',

'French Guiana' = 'America', 'French Polynesia' = 'Oceania', 'Gabon' = 'Africa', 'Gambia' = 'Africa', 'Georgia' = 'Asia',

'Germany' = 'Europe', 'Ghana' = 'Africa', 'Gibraltar' = 'Europe', 'Greece' = 'Europe', 'Greenland' = 'Europe',

'Grenada' = 'America', 'Guadeloupe' = 'America', 'Guam' = 'Oceania', 'Guatemala' = 'America', 'Guinea' = 'Africa',

'Guinea-Bissau' = 'Africa', 'Guyana' = 'America', 'Haiti' = 'America', 'Honduras' = 'America', 'Hong Kong' = 'Asia',

'Hungary' = 'Europe', 'Iceland' = 'Europe', 'India' = 'Asia', 'Indonesia' = 'Asia', 'Iran' = 'Asia',

'Iraq' = 'Asia', 'Ireland' = 'Europe', 'Israel' = 'Asia', 'Italy' = 'Europe', 'Jamaica' = 'America',

'Japan' = 'Asia', 'Jordan' = 'Asia', 'Kazakhstan' = 'Asia', 'Kenya' = 'Africa', 'Kiribati' = 'Oceania',

'Kuwait' = 'Asia', 'Kyrgyzstan' = 'Asia', 'Laos' = 'Asia', 'Latvia' = 'Europe', 'Lebanon' = 'Asia',

'Lesotho' = 'Africa', 'Liberia' = 'Africa', 'Libya' = 'Africa', 'Lithuania' = 'Europe', 'Luxembourg' = 'Europe',

'Macao' = 'Asia', 'Madagascar' = 'Africa', 'Malawi' = 'Africa', 'Malaysia' = 'Asia', 'Maldives' = 'Asia',

'Mali' = 'Africa', 'Malta' = 'Europe', 'Martinique' = 'America', 'Mauritania' = 'Africa', 'Mauritius' = 'Africa',

'Mexico' = 'America', 'Micronesia (country)' = 'Oceania', 'Moldova' = 'Europe', 'Mongolia' = 'Asia', 'Montenegro' = 'Europe',

'Montserrat' = 'America', 'Morocco' = 'Africa', 'Mozambique' = 'Africa', 'Myanmar' = 'Asia', 'Namibia' = 'Africa',

'Nauru' = 'Oceania', 'Nepal' = 'Asia', 'Netherlands' = 'Europe', 'Netherlands Antilles' = 'America', 'New Caledonia' = 'Oceania',

'New Zealand' = 'Oceania', 'Nicaragua' = 'America', 'Niger' = 'Africa', 'Nigeria' = 'Africa', 'Niue' = 'Oceania',

```
  'North Korea' = 'Asia', 'North Macedonia' = 'Europe', 'Northern Mariana Islands' = 'Oceania',
'Norway' = 'Europe', 'Oman' = 'Asia',

  'Pakistan' = 'Asia', 'Palau' = 'Oceania', 'Palestine' = 'Asia', 'Panama' = 'America', 'Papua New
Guinea' = 'Oceania',

  'Paraguay' = 'America', 'Peru' = 'America', 'Philippines' = 'Asia', 'Poland' = 'Europe', 'Portugal' =
'Europe',

  'Puerto Rico' = 'America', 'Qatar' = 'Asia', 'Reunion' = 'Africa', 'Romania' = 'Europe', 'Russia' =
'Europe',

  'Rwanda' = 'Africa', 'Saint Helena' = 'Africa', 'Saint Kitts and Nevis' = 'America', 'Saint Lucia' =
'America', 'Saint Pierre and Miquelon' = 'America',

  'Saint Vincent and the Grenadines' = 'America', 'Samoa' = 'Oceania', 'Sao Tome and Principe' =
'Africa', 'Saudi Arabia' = 'Asia', 'Senegal' = 'Africa',

  'Serbia' = 'Europe', 'Seychelles' = 'Africa', 'Sierra Leone' = 'Africa', 'Singapore' = 'Asia', 'Slovakia' =
'Europe',

  'Slovenia' = 'Europe', 'Solomon Islands' = 'Oceania', 'Somalia' = 'Africa', 'South Africa' = 'Africa',
'South Korea' = 'Asia',

  'South Sudan' = 'Africa', 'Spain' = 'Europe', 'Sri Lanka' = 'Asia', 'Sudan' = 'Africa', 'Suriname' =
'America',

  'Sweden' = 'Europe', 'Switzerland' = 'Europe', 'Syria' = 'Asia', 'Taiwan' = 'Asia', 'Tajikistan' = 'Asia',

  'Tanzania' = 'Africa', 'Thailand' = 'Asia', 'Togo' = 'Africa', 'Tonga' = 'Oceania', 'Trinidad and Tobago'
= 'America',

  'Tunisia' = 'Africa', 'Turkey' = 'Asia', 'Turkmenistan' = 'Asia', 'Turks and Caicos Islands' = 'America',
'Tuvalu' = 'Oceania',

  'Uganda' = 'Africa', 'Ukraine' = 'Europe', 'United Arab Emirates' = 'Asia', 'United Kingdom' =
'Europe', 'United States' = 'America',

  'United States Virgin Islands' = 'America', 'Uruguay' = 'America', 'Uzbekistan' = 'Asia', 'Vanuatu' =
'Oceania', 'Vatican City' = 'Europe',

  'Venezuela' = 'America', 'Vietnam' = 'Asia', 'Western Sahara' = 'Africa', 'Yemen' = 'Asia', 'Zambia' =
'Africa', 'Zimbabwe' = 'Africa'

)# Map countries to continents

data$continent <- continent_mapping[data$country]


# Define colors for continents

continent_colors <- c(

  'Europe' = '#0000FF',  # Blue

  'Asia' = '#FFFF00',    # Yellow
```

```
  'Africa' = '#000000',  # Black
  'Oceania' = '#008000', # Green
  'America' = '#FF0000'  # Red
)


# Create a new column for the X-axis positions based on continent groups
data <- data %>%
  group_by(continent) %>%
  mutate(x_pos = row_number())

# Adjust the X-axis positions to have spaces between continents
continent_offsets <- c(
  'Europe' = 0,
  'Asia' = 100,
  'Africa' = 200,
  'Oceania' = 300,
  'America' = 400
)
data$x_pos <- data$x_pos + continent_offsets[data$continent]

# Create scatter plot for primary_energy consumption, categorizing by continent
ggplot(data, aes(x = x_pos, y = primary_energy, color = continent)) +
  geom_point(shape = 16) +
  scale_color_manual(values = continent_colors) +
  labs(x = 'Index', y = 'Primary_energy Consumption', title = 'Primary_energy Consumption by
Continent in 2010') +
  theme_minimal() +
  ylim(0, 3000)  # Set Y-axis range to be visible for all data
```