BSD2333 DATA WRANGLING

GROUP PROJECT

TITLE:

ROAD TO CLEAN AND RENEW:

SOUTHEAST ASIA'S PATH TOWARD SUSTAINABLE ENERGY

LECTURER: DR. MOHD KHAIRUL BAZLI BIN MOHD AZIZ

GROUP NAME: SEA YOU LATER

NO	MATRIC ID	NAME	SECTION	PHOTO
1	SD23039	CHUA EN HUAN	02G	
2	SD23063	MUHAMMAD DANIAL BIN ISSHAM	02G	
3	SD23048	YIP YOONG ENG	02G	
4	SD23023	TION JIA LE	02G	
5	SD23042	SOH HONG YANG	01G	

TABLE OF CONTENT

LIST OF TABLES **2**

LIST OF FIGURES **3**

1.0 SYNOPSIS **6**

1.1 Project Description 6

1.2 Problem Statement 8

1.3 Project Questions 8

1.4 Objectives 8

1.5 Basic Description of Data 9

2.0 PACKAGES REQUIRED **11**

3.0 DATA PREPARATION **13**

3.1 Data Import 13

3.2 Data Cleaning 15

3.3 Data Preview 43

3.4 Data Description 57

3.5 Data Description 61

4.0 EXPLORATORY DATA ANALYSIS **62**

4.1 Overview of Access to Electricity by Country 62

4.2 Average Electricity Generation Mix by Country 64

4.3 Yearly Trend of Access to Clean Fuels by Country 66

4.4 Renewable Energy Trends: Installed Capacity vs. Consumption Share by Country 68

4.5 CO_2 Emissions vs. Primary Energy Consumption per Capita 71

4.6 CO_2 Emissions vs. Renewable Energy Share in Final Consumption 73

4.7 Energy Intensity by Country with Yearly Variation 75

4.8 3D View: GDP per Capita vs. Renewable Energy Share vs. Energy Intensity Level of Primary Energy 78

5.0 SUMMARY **80**

6.0 REFERENCES **84**

7.0 APPENDIX **85**

LIST OF TABLES

Table 1.5.1	9
Table 2.1.1	12
Table 3.4.1	58
Table 3.4.2	59

LIST OF FIGURES

Figure 3.1.1	13
Figure 3.1.2	13
Figure 3.1.3	14
Figure 3.2.1	15
Figure 3.2.2	16
Figure 3.2.3	16
Figure 3.2.4	17
Figure 3.2.5	17
Figure 3.2.6	18
Figure 3.2.7	19
Figure 3.2.8	19
Figure 3.2.9	21
Figure 3.2.10	21
Figure 3.2.11	22
Figure 3.2.12	22
Figure 3.2.13	22
Figure 3.2.14	23
Figure 3.2.15	24
Figure 3.2.16	25
Figure 3.2.17	26
Figure 3.2.18	26
Figure 3.2.19	26
Figure 3.2.20	27
Figure 3.2.21	28
Figure 3.2.22	28
Figure 3.2.23	29
Figure 3.2.24	30
Figure 3.2.25	30
Figure 3.2.26	31
Figure 3.2.27	31

Figure 3.2.28	32
Figure 3.2.29	32
Figure 3.2.30	34
Figure 3.2.31	34
Figure 3.2.32	34
Figure 3.2.33	35
Figure 3.2.34	35
Figure 3.2.35	35
Figure 3.2.36	36
Figure 3.2.37	36
Figure 3.2.38	36
Figure 3.2.39	37
Figure 3.2.40	37
Figure 3.2.41	37
Figure 3.2.42	38
Figure 3.2.43	38
Figure 3.2.44	38
Figure 3.2.45	39
Figure 3.2.46	40
Figure 3.2.47	40
Figure 3.2.48	41
Figure 3.2.49	41
Figure 3.3.1	43
Figure 3.3.2	44
Figure 3.3.3	44
Figure 3.3.4	45
Figure 3.3.5	45
Figure 3.3.6	46
Figure 3.3.7	47
Figure 3.3.8	48
Figure 3.3.9	48

Figure 3.3.10	49
Figure 3.3.11	49
Figure 3.3.12	50
Figure 3.3.13	50
Figure 3.3.14	51
Figure 3.3.15	51
Figure 3.3.16	52
Figure 3.3.17	52
Figure 3.3.18	53
Figure 3.3.19	53
Figure 3.3.20	54
Figure 3.3.21	54
Figure 3.3.22	55
Figure 3.3.23	55
Figure 3.3.24	56
Figure 3.3.25	56
Figure 3.5.1	61
Figure 4.1.1	62
Figure 4.2.1	64
Figure 4.3.1	66
Figure 4.4.1	68
Figure 4.4.2	68
Figure 4.4.3	69
Figure 4.4.4	69
Figure 4.5.1	71
Figure 4.6.1	73
Figure 4.7.1	75
Figure 4.8.1	78

1.0 SYNOPSIS

1.1 Project Description

Renewable energy or clean energy is considered as a major future energy source due to its various benefits brought by it. It comes in various forms such as solar energy, wind energy, hydropower, and many more and it can be easily obtained as we are surrounded by them. Since renewable energy can be obtained easily and is free of charge, humans have come out with various methods and techniques to harness this energy by converting it into usable forms to fulfill daily needs. Due to its special characteristics and potential in generating large revenues, it has boosted various enterprises that have started their initiative in inventing and innovating new, reliable, portable yet long-lasting renewable energy devices such as engines to grab the opportunity to be the first enterprise to earn it. For example, the large American enterprise, Tesla invested a large amount of money in creating clean energy products such as electric cars and successfully became the world's most valuable car company on June 10, 2020, along with expanding their clean energy market to other sectors in the upcoming years (Berdikeeva, 2023). Here brings question marks regarding why clean energy or renewable energy has become the most desirable slice of the pie in the business world and how it might impact the future ahead of us.

Renewable energy has been used by humans since ancient times without us knowing it. During the Greek and Roman era, Greeks and Romans used solar energy by reflecting them to light torches using concave mirrors called burning mirrors. In ancient China, wind energy was used to pump water with simple windmills, while in the Middle East, vertical axis windmills were used to capture wind to grind grains as early as 200 BC (Gomstyn, 2024). Even back in the days when there was limited technology, our ancestors managed to utilize renewable energy in their daily lives with simple and basic tools. After that, humans are starting to use and harness renewable energy more effectively and efficiently to their benefit with the continuous and fast development of technology. Various improvements and adjustments have been made to make renewable energy smarter, cheaper, and more efficient to be used by everyone. From the invention of the first battery in the 19th century, the same technology is still applied to today's fully electric, gas-free cars to the present where modern solar plants power entire communities, renewable energy has come a long way. For example, in 2017, Jimmy Carter, a former president of the United States of America oversaw the construction of a 1.3-megawatt solar plant to fulfil 50% of the power needs of his hometown of Plains, Georgia (*History of Renewable Energy Timeline | Inspire Clean Energy*, n.d.). This evolution reflects that renewable energy is now slowly taking over or substituting the use of fossil fuels or gas in pursuit of creating a cleaner, more environmentally friendly, and sustainable neighbourhood.

A lot of countries have started their initiatives to slowly increase the usage of renewable energy while reducing dependence on fossil fuels or gases at the same time. Countries from Southeast Asia are also not falling behind. ASEAN countries have come to an agreement where they aim to increase the proportion of renewable energy usage to 23% of the total primary energy supply by 2025 in order to secure energy needs and create favourable conditions for the region's clean energy development (ASEAN, n.d.). Since nuclear-related products are prohibited from being used in this area, ASEAN countries have focused on alternative renewable sources such as solar, wind, hydro, and bioenergy to meet their growing energy demand in a more sustainable manner. For instance, Malaysia whose fossil fuels dominate the country's current primary total energy supply aims to expand the share of renewable energy to 31% of the country's energy demand. In recent years, their energy source from fossil fuels has slowly decreased with solar and hydropower slowly filling in the energy gap due to geographical advantages of the country with abundant sunlight exposure and ample water sources (Koons, 2024). With continuous efforts, ASEAN countries will achieve their goals of having a high share of renewable energy in their total primary energy supply. This ongoing transition also directly supports Sustainable Development Goal 7, which aims to ensure access to affordable, reliable, sustainable, and modern energy for all.

This project explores the regional trends in the generation of electricity from renewable energy sources and its impact on sustainability, emissions and economic aspects. To fully understand the interrelationships of the variables in the dataset, data analysis and interactive visualizations are used to help people with or without data interpretation knowledge to easily understand the findings and outcomes of this project.

1.2 Problem Statement

Despite growing awareness and investment in renewable energy, many countries, especially in developing regions, still face challenges in reducing their dependence on fossil fuels. ASEAN countries are one of the developing regions where most countries greatly rely on fossil fuels as their primary energy source such as the main source in generating electricity for daily usage. The transition to clean energy varies significantly across regions, influenced by various factors such as economic capacity, policy support, and geographic advantages. Furthermore, the link between renewable energy adoption and reductions in CO_2 emissions is not always clearly defined or measured. As ASEAN countries set ambitious renewable energy targets, there is a need to better understand how renewable energy contributes to emissions reduction, energy intensity, and sustainable development. These problems needed to be addressed to have a clearer understanding of the whole situation and come up with solutions to overcome soon-to-be unknown challenges and obstacles.

1.3 Project Questions

Several questions need to be answered in this project. The questions are:

1. How does access to electricity and clean fuels vary among Southeast Asian countries over time?
2. What is the relationship between renewable energy development and CO_2 emissions in Southeast Asia?
3. How do economic factors such as GDP per capita relate to renewable energy use and energy intensity in Southeast Asia?

1.4 Objectives

Several objectives need to be achieved in this project. The objectives are:

1. To analyze the trends of access to electricity and clean fuels among Southeast Asian countries over time.
2. To investigate the relationship between renewable energy development and CO_2 emissions in Southeast Asia.
3. To explore how economic factors like GDP per capita relate to renewable energy usage and energy intensity in Southeast Asia.

1.5 Basic Description of Data

In this project, sustainable energy used in each country will be investigated. A global sustainable energy dataset is obtained through Kaggle. The dataset contains 21 columns of data with both quantitative data and qualitative data. The table below shows the attributes of the dataset.

Quantitative	Qualitative
• Year • Access to electricity (% of population) • Access to clean fuels for cooking (% of population) • Renewable-electricity-generating-capacity-per-capita • Financial flows to developing countries (US $) • Renewable energy share in total final energy consumption (%) • Electricity from fossil fuels (TWh) • Electricity from nuclear (TWh) • Electricity from renewables (TWh) • Low-carbon electricity (% electricity) • Primary energy consumption per capita (kWh/person) • Energy intensity level of primary energy (MJ/$2017 PPP GDP) • Value_co2_emission (metric tons per capita) • Renewables (% equivalent primary energy) • GDP growth (annual %) • GDP per capita • Density (P/Km2) • Land area (Km2) • Latitude • Longitude	• Entity

Table 1.5.1

Year in the dataset represents the year for which the data is reported, ranging from 2000 to 2020. Access to electricity and access to clean fuels for cooking describe the percentage of population in each country with access to electricity and primary reliance on clean fuels respectively. The variables named financial flows to developing countries are measured in US dollars and it indicates the amount of financial aid from developed countries for projects of clean energy.

Furthermore, the variables to describe the source of electricity in each country include electricity from fossil fuels, electricity from nuclear and electricity from renewables where electricity generated from fossil fuels, nuclear energy and renewable energy are measured in terawatt-hours. The variable named low-carbon electricity represents the percentage of electricity from nuclear energy and renewable energy.

In terms of production, Renewable-electricity-generating-capacity-per-capita indicates the amount of electricity generated by renewable energy per capita. For consumption, Renewable energy share in total final energy consumption refers to the percentage of renewable energy used in final energy consumption. Primary energy consumption per capita describes the energy consumed in kilowatt hours for each person. The variable Renewables (% equivalent primary energy) represents the share of renewable sources in the total equivalent primary energy supply. Other than that, the energy intensity level of primary energy refers to energy use per unit of GDP at purchasing power parity. Carbon dioxide (CO_2) emissions for each person are measured in metric tons for the variables value_co2_emissions.

Each country is evaluated through several perspectives which are GDP growth, GDP per capita, Density, Land Area, Latitude and Longitude. GDP growth and GDP per capita describe the annual GDP growth rate based on local currency and gross domestic product per person respectively. Land area and density indicate the total land area in square kilometers and population density in person per square kilometers for each country respectively. Latitude and longitude represent the location of the country on the world map.

2.0 PACKAGES REQUIRED

To carry out this project effectively, several Python libraries were utilized to support data preparation and visualization. Each of the packages plays a specific role in handling the dataset in order to create meaningful visualization with interactive elements.

Packages/ Libraries	Description	Usage
google.colab.files	A module provided in Google Colab that allows file interactions between local machines and Colab environment.	It is used to upload the CSV dataset from the local disk into Google Colab using file.upload() for data processing and analysis.
Pandas	An open-source library for data manipulation and analysis for Python programming language.	It is used to read the datasets in CSV format, clean data, merge datasets, filter rows, and summarize data through group-by operations.
Seaborn	A statistical data visualization library based on Matplotlib that comes with built-in themes and color palettes.	It is used to create interactive and informative plots like box plots with minimal code.
matplotlib.pyplot	A foundational plotting module in Python for 2D visualizations.	It is used to build useful charts like bar charts, scatterplots, and line charts, at the same time to customize visual elements like axis labels, titles and figure sizes.
plotly.express	A high-level wrapper for Plotly that simplifies the creation of interactive visualizations.	It is used to quickly generate interactive charts such as scatter plots, bar charts and choropleth maps with concise syntax.

plotly.graph_objects	A lower-level API from Plotly that offers flexibility for building complex and highly customized visualizations.	It is used to define and combine multiple chart element and control "Year" widget.
bokeh.plotting	A core interface tool for creating plots in Python.	It is used to create new plot by "figure", while "show" displays it in the notebook or browser.
bokeh.io	A module used to handle Bokeh's output settings by controlling the way and the place that Bokeh plots are displayed.	It is used to ensure that the interactive Bokeh plots appear directly within a Colab notebook, instead of separated tabs.
bokeh.models	The module provides the underlying building blocks and tools to enhance plot interactivity.	"ColumnDataSource" is used to act as the linkage between dataset and Bokeh's visual elements to allow efficient interactions. "HoverTool" helps on tooltip feature that shows data values when hovering over a point on the plot to improve user interaction.
bokeh.palettes	A collection of pre-defined colour palettes used for data visualization to show contrast for readability.	It is used to provide uniform colour code to differentiate the data groups in a plot.

Table 2.1.1

3.0 DATA PREPARATION

3.1 Data Import

The dataset downloaded from Kaggle is imported into Google Colab for data cleaning.

```
import pandas as pd
```

```
# Uploading excel data
from google.colab import files
uploaded=files.upload()
```

```
Choose Files  No file chosen        Upload widget is only available when the cell has been executed in the current browser session. Please rerun this cell to enable.
Saving global-data-on-sustainable-energy.csv to global-data-on-sustainable-energy (1).csv
```

Figure 3.1.1

```
# Making data frame from csv file
data = pd.read_csv("global-data-on-sustainable-energy.csv")
data
```

	Entity	Year	Access to electricity (% of population)	Access to clean fuels for cooking	Renewable-electricity-generating-capacity-per-capita	Financial flows to developing countries (US $)	Renewable energy share in the total final energy consumption (%)	Electricity from fossil fuels (TWh)	Electricity from nuclear (TWh)	Electricity from renewables (TWh)	...	Primary energy consumption per capita (kWh/person)
0	Afghanistan	2000	1.61	6.20	9.22	20000.00	44.99	0.16	0.00	0.31	...	302.59
1	Afghanistan	2001	4.07	7.20	8.86	130000.00	45.60	0.09	0.00	0.50	...	236.89
2	Afghanistan	2002	9.41	8.20	8.47	3950000.00	37.83	0.13	0.00	0.56	...	210.86
3	Afghanistan	2003	14.74	9.50	8.09	25970000.00	36.66	0.31	0.00	0.63	...	229.97
4	Afghanistan	2004	20.06	10.90	7.75	NaN	44.24	0.33	0.00	0.56	...	204.23
...	...	...	...	...	...	...	...	...	...	...	...	...
3644	Zimbabwe	2016	42.56	29.80	62.88	30000.00	81.90	3.50	0.00	3.32	...	3227.68
3645	Zimbabwe	2017	44.18	29.80	62.33	5570000.00	82.46	3.05	0.00	4.30	...	3068.01
3646	Zimbabwe	2018	45.57	29.90	82.53	10000.00	80.23	3.73	0.00	5.46	...	3441.99
3647	Zimbabwe	2019	46.78	30.10	81.40	250000.00	81.50	3.66	0.00	4.58	...	3003.66
3648	Zimbabwe	2020	52.75	30.40	80.61	30000.00	81.90	3.40	0.00	4.19	...	2680.13

3649 rows × 21 columns

Figure 3.1.2

Energy intensity level of primary energy (MJ/$2017 PPP GDP)	Value_co2_emissions_kt_by_country	Renewables (% equivalent primary energy)	gdp_growth	gdp_per_capita	Density\n(P/Km2)	Land Area(Km2)	Latitude	Longitude
1.64	760.00	NaN	NaN	NaN	60	652230.00	33.94	67.71
1.74	730.00	NaN	NaN	NaN	60	652230.00	33.94	67.71
1.40	1030.00	NaN	NaN	179.43	60	652230.00	33.94	67.71
1.40	1220.00	NaN	8.83	190.68	60	652230.00	33.94	67.71
1.20	1030.00	NaN	1.41	211.38	60	652230.00	33.94	67.71
...	...	...	...	...	...	...	...	...
10.00	11020.00	NaN	0.76	1464.59	38	390757.00	-19.02	29.15
9.51	10340.00	NaN	4.71	1235.19	38	390757.00	-19.02	29.15
9.83	12380.00	NaN	4.82	1254.64	38	390757.00	-19.02	29.15
10.47	11760.00	NaN	-6.14	1316.74	38	390757.00	-19.02	29.15
10.00	NaN	NaN	-6.25	1214.51	38	390757.00	-19.02	29.15

Figure 3.1.3

3.2 Data Cleaning

Data cleaning is one of the important processes before further analysis. Data cleaning processes ensure the consistency, integrity and accuracy of data which will enhance the quality and efficiency of data-driven decision making.

Step 1: Check unique values in categorical columns

```
# Check unique values in Entity column
data['Entity'].unique()
```

```
array(['Afghanistan', 'Albania', 'Algeria', 'Angola',
       'Antigua and Barbuda', 'Argentina', 'Armenia', 'Aruba',
       'Australia', 'Austria', 'Azerbaijan', 'Bahamas', 'Bahrain',
       'Bangladesh', 'Barbados', 'Belarus', 'Belgium', 'Belize', 'Benin',
       'Bermuda', 'Bhutan', 'Bosnia and Herzegovina', 'Botswana',
       'Brazil', 'Bulgaria', 'Burkina Faso', 'Burundi', 'Cambodia',
       'Cameroon', 'Canada', 'Cayman Islands', 'Central African Republic',
       'Chad', 'Chile', 'China', 'Colombia', 'Comoros', 'Congo',
       'Costa Rica', 'Croatia', 'Cuba', 'Cyprus', 'Czechia', 'Denmark',
       'Djibouti', 'Dominica', 'Dominican Republic', 'Ecuador', 'Egypt',
       'El Salvador', 'Equatorial Guinea', 'Eritrea', 'Estonia',
       'Eswatini', 'Ethiopia', 'Fiji', 'Finland', 'France',
       'French Guiana', 'Gabon', 'Gambia', 'Georgia', 'Germany', 'Ghana',
       'Greece', 'Grenada', 'Guatemala', 'Guinea', 'Guinea-Bissau',
       'Guyana', 'Haiti', 'Honduras', 'Hungary', 'Iceland', 'India',
       'Indonesia', 'Iraq', 'Ireland', 'Israel', 'Italy', 'Jamaica',
       'Japan', 'Jordan', 'Kazakhstan', 'Kenya', 'Kiribati', 'Kuwait',
       'Kyrgyzstan', 'Latvia', 'Lebanon', 'Lesotho', 'Liberia', 'Libya',
       'Lithuania', 'Luxembourg', 'Madagascar', 'Malawi', 'Malaysia',
       'Maldives', 'Mali', 'Malta', 'Mauritania', 'Mauritius', 'Mexico',
       'Mongolia', 'Montenegro', 'Morocco', 'Mozambique', 'Myanmar',
       'Namibia', 'Nauru', 'Nepal', 'Netherlands', 'New Caledonia',
       'New Zealand', 'Nicaragua', 'Niger', 'Nigeria', 'North Macedonia',
       'Norway', 'Oman', 'Pakistan', 'Panama', 'Papua New Guinea',
       'Paraguay', 'Peru', 'Philippines', 'Poland', 'Portugal',
       'Puerto Rico', 'Qatar', 'Romania', 'Rwanda',
       'Saint Kitts and Nevis', 'Saint Lucia',
       'Saint Vincent and the Grenadines', 'Samoa',
       'Sao Tome and Principe', 'Saudi Arabia', 'Senegal', 'Serbia',
       'Seychelles', 'Sierra Leone', 'Singapore', 'Slovakia', 'Slovenia',
       'Solomon Islands', 'Somalia', 'South Africa', 'South Sudan',
       'Spain', 'Sri Lanka', 'Sudan', 'Suriname', 'Sweden', 'Switzerland',
       'Tajikistan', 'Thailand', 'Togo', 'Tonga', 'Trinidad and Tobago',
       'Tunisia', 'Turkey', 'Turkmenistan', 'Tuvalu', 'Uganda', 'Ukraine',
       'United Arab Emirates', 'United Kingdom', 'United States',
       'Uruguay', 'Uzbekistan', 'Vanuatu', 'Yemen', 'Zambia', 'Zimbabwe'],
      dtype=object)
```

Figure 3.2.1

The output listed the countries available in the dataset. By scanning through the output, countries in Southeast Asia such as Indonesia, Laos and Vietnam are all included, spelled correctly, and consistently formatted. There is no noticeable noisy data. This step helps to ensure data accuracy when filtering the dataset.

```
# Check unique values in Year column

data['Year'].unique()
```

```
array([2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010,
       2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019, 2020])
```

Figure 3.2.2

Year as one of the categorical values has also been further confirmed to have unique values. The output shows the list of years available in the dataset, ranging from 2000 to 2020. This confirms that the dataset spans a continuous period of 21 years with no missing or unexpected values. Therefore, any time-series analysis or filtering based on specific years can proceed reliably.

Step 2: Filter Southeast Asian countries

```
# List of Southeast Asian countries to filter
sea_countries = ['Malaysia', 'Singapore', 'Thailand', 'Indonesia',
'Vietnam', 'Philippines', 'Brunei', 'Cambodia', 'Laos', 'Myanmar', 'Timor-
Leste']

# Filter the dataset to include only Southeast Asian countries
sea_data = data[data['Entity'].isin(sea_countries)]

# Display the filtered data
sea_data
```

	Entity	Year	Access to electricity (% of population)	Access to clean fuels for cooking	Renewable-electricity-generating-capacity-per-capita	Financial flows to developing countries (US $)	Renewable energy share in the total final energy consumption (%)	Electricity from fossil fuels (TWh)	Electricity from nuclear (TWh)	Electricity from renewables (TWh)	...	Primary energy consumption per capita (kWh/person)
567	Cambodia	2000	16.60	3.60	0.83	NaN	81.58	0.42	0.00	0.05	...	739.46
568	Cambodia	2001	14.62	4.10	0.82	NaN	80.51	0.48	0.00	0.03	...	769.41
569	Cambodia	2002	18.19	4.70	0.97	NaN	80.97	0.66	0.00	0.03	...	754.14
570	Cambodia	2003	19.30	5.30	0.96	4630000.00	79.92	0.67	0.00	0.04	...	803.82
571	Cambodia	2004	25.30	5.90	1.02	409150000.00	80.69	0.73	0.00	0.03	...	813.56
...	...	...	...	...	...	...	...	...	...	...	...	...
3266	Thailand	2016	99.85	80.10	136.90	420000.00	22.43	161.79	0.00	15.97	...	19981.58
3267	Thailand	2017	99.90	80.80	148.00	1990000.00	22.25	161.88	0.00	19.92	...	20264.65
3268	Thailand	2018	99.82	81.70	163.82	296130000.00	23.70	156.26	0.00	25.84	...	20834.09
3269	Thailand	2019	99.90	82.70	170.30	99160000.00	23.96	162.59	0.00	28.02	...	20792.38
3270	Thailand	2020	100.00	83.60	171.79	NaN	NaN	154.52	0.00	24.73	...	19298.09

147 rows × 21 columns

Figure 3.2.3

Energy intensity level of primary energy (MJ/$2017 PPP GDP)	Value_co2_emissions_kt_by_country	Renewables (% equivalent primary energy)	gdp_growth	gdp_per_capita	Density\n(P/Km2)	Land Area(Km2)	Latitude	Longitude
7.91	1960.00	NaN	9.99	300.61	95	181035.00	12.57	104.99
7.51	2150.00	NaN	8.15	321.15	95	181035.00	12.57	104.99
7.05	2210.00	NaN	6.58	338.99	95	181035.00	12.57	104.99
6.69	2380.00	NaN	8.51	362.34	95	181035.00	12.57	104.99
6.23	2380.00	NaN	10.34	408.51	95	181035.00	12.57	104.99
...	...	...	...	...	...	...	...	...
5.02	261600.01	4.60	3.44	5993.31	137	513120.00	15.87	100.99
4.79	258820.01	5.31	4.18	6593.82	137	513120.00	15.87	100.99
4.50	257049.99	6.38	4.19	7296.88	137	513120.00	15.87	100.99
4.52	267090.00	6.96	2.27	7817.01	137	513120.00	15.87	100.99
NaN	NaN	6.99	-6.10	7186.87	137	513120.00	15.87	100.99

Figure 3.2.4

Since the project focuses on the renewable energy in Southeast Asian countries only, the dataset is filtered by remaining only the data for Southeast Asian countries. From Figure 3.2.3 and Figure 3.2.4, the dataset includes only Southeast Asian countries and is named sea_data.

```
# Check the unique countries in the filtered data
sea_data['Entity'].unique()
```

```
array(['Cambodia', 'Indonesia', 'Malaysia', 'Myanmar', 'Philippines',
       'Singapore', 'Thailand'], dtype=object)
```

Figure 3.2.5

The filtered data undergo checking for unique values to ensure all data in the filtered dataset is valid and accurate before moving on to the next step.

Step 3: Review dataset structure and summary

```
# Check basic info: data types, non-null counts, memory usage
sea_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Index: 147 entries, 567 to 3270
Data columns (total 21 columns):
 #   Column                                                        Non-Null Count  Dtype
---  ------                                                        --------------  -----
 0   Entity                                                        147 non-null    object
 1   Year                                                          147 non-null    int64
 2   Access to electricity (% of population)                       147 non-null    float64
 3   Access to clean fuels for cooking                             147 non-null    float64
 4   Renewable-electricity-generating-capacity-per-capita          147 non-null    float64
 5   Financial flows to developing countries (US $)                102 non-null    float64
 6   Renewable energy share in the total final energy consumption (%)  140 non-null    float64
 7   Electricity from fossil fuels (TWh)                           147 non-null    float64
 8   Electricity from nuclear (TWh)                                105 non-null    float64
 9   Electricity from renewables (TWh)                             147 non-null    float64
 10  Low-carbon electricity (% electricity)                        147 non-null    float64
 11  Primary energy consumption per capita (kWh/person)            147 non-null    float64
 12  Energy intensity level of primary energy (MJ/$2017 PPP GDP)   140 non-null    float64
 13  Value_co2_emissions_kt_by_country                             140 non-null    float64
 14  Renewables (% equivalent primary energy)                      105 non-null    float64
 15  gdp_growth                                                    147 non-null    float64
 16  gdp_per_capita                                                147 non-null    float64
 17  Density\n(P/Km2)                                              147 non-null    object
 18  Land Area(Km2)                                                147 non-null    float64
 19  Latitude                                                      147 non-null    float64
 20  Longitude                                                     147 non-null    float64
dtypes: float64(18), int64(1), object(2)
memory usage: 25.3+ KB
```

Figure 3.2.6

From the output, there are 147 rows and 21 columns. Some column names contain unwanted characters such as hyphens (-), underscores (_), and newline characters (\n). These characters need to be cleared through renaming the columns to ensure consistency, clear and easy to understand. Besides that, data type of Density is recorded as object instead of numeric. Data type for the values recorded in Density columns should be further investigated and converted into the correct data type. Also, the output shows that there are some rows containing missing values which will need to be addressed during the process of handling missing data.

```python
# View summary statistics for numeric columns
sea_data.describe()
```

	Year	Access to electricity (% of population)	Access to clean fuels for cooking	Renewable-electricity-generating-capacity-per-capita	Financial flows to developing countries (US $)	Renewable energy share in the total final energy consumption (%)	Electricity from fossil fuels (TWh)	Electricity from nuclear (TWh)	Electricity from renewables (TWh)	Low-carbon electricity (% electricity)	Primary energy consumption per capita (kWh/person)	Energy intensity level of primary energy (MJ/$2017 PPP GDP)
count	147.00	147.00	147.00	147.00	102.00	140.00	147.00	105.00	147.00	147.00	147.00	140.00
mean	2010.00	82.13	54.39	63.71	99996666.67	34.01	71.06	0.00	12.22	22.21	29379.59	4.50
std	6.08	23.37	35.74	56.05	201433979.77	28.12	63.50	0.00	11.74	21.21	47350.54	1.41
min	2000.00	14.62	2.90	0.82	10000.00	0.33	0.42	0.00	0.00	0.00	739.46	2.05
25%	2005.00	70.11	18.80	28.76	887500.00	4.39	6.83	0.00	2.08	6.89	3200.56	3.45
50%	2010.00	92.94	47.30	51.72	8720000.00	28.62	50.04	0.00	8.95	15.35	7121.82	4.58
75%	2015.00	99.53	96.70	71.11	94385000.00	58.39	120.77	0.00	20.22	34.47	31145.51	5.18
max	2020.00	100.00	100.00	268.76	1056140000.00	85.71	247.39	0.00	52.91	79.07	164801.44	10.52

Figure 3.2.7

Value_co2_emissions_kt_by_country	Renewables (% equivalent primary energy)	gdp_growth	gdp_per_capita	Land Area(Km2)	Latitude	Longitude
140.00	105.00	147.00	147.00	147.00	147.00	147.00
142557.14	5.40	5.59	9135.51	557980.71	9.71	106.20
145160.92	5.24	3.79	15899.00	587616.42	7.73	8.13
1960.00	0.18	-9.57	128.10	716.00	-0.79	95.96
18540.00	2.32	4.27	1135.46	181035.00	1.35	100.99
79300.00	3.53	5.69	3001.04	329847.00	12.57	103.82
229115.00	6.96	7.17	7215.17	676578.00	15.87	113.92
619840.03	18.68	14.53	66679.05	1904569.00	21.92	121.77

Figure 3.2.8

Figure 3.2.7 and Figure 3.2.8 displays the statistical summary for filtered dataset. From the output, Density is not included in the table due to the data type in previous process is recorded as object type instead of numeric type. The statistical summary of the filtered dataset also provides information of statistical measurement for each column which eases the process of detecting outliers based on the minimum value and maximum value. For instance, the maximum value of GDP per capita lies far away from 75th percentile indicates the possible outliers present in the dataset. These outliers require further investigation to determine whether it should be removed.

Step 4: Rename columns

```
sea_data = sea_data.rename(columns={
    'Entity': 'Country',
    'Year': 'Year',
    'Access to electricity (% of population)': 'Access to electricity (% of population)',
    'Access to clean fuels for cooking': 'Access to clean fuels for cooking (% of population)',
    'Renewable-electricity-generating-capacity-per-capita': 'Installed renewable electricity capacity per capita (kW)',
    'Financial flows to developing countries (US $)': 'Financial flows to developing countries (US$)',
    'Renewable energy share in the total final energy consumption (%)': 'Renewable energy share in total final energy consumption (%)',
    'Electricity from fossil fuels (TWh)': 'Electricity from fossil fuels (TWh)',
    'Electricity from nuclear (TWh)': 'Electricity from nuclear (TWh)',
    'Electricity from renewables (TWh)': 'Electricity from renewables (TWh)',
    'Low-carbon electricity (% electricity)': 'Low-carbon electricity (% of electricity)',
    'Primary energy consumption per capita (kWh/person)': 'Primary energy consumption per capita (kWh)',
    'Energy intensity level of primary energy (MJ/$2017 PPP GDP)': 'Energy intensity level of primary energy (MJ/$2017 PPP GDP)',
    'Value_co2_emissions_kt_by_country': 'CO2 emissions per capita (metric tons)',
    'Renewables (% equivalent primary energy)': 'Renewables (% equivalent primary energy)',
    'gdp_growth': 'GDP growth rate (annual %)',
    'gdp_per_capita': 'GDP per capita',
    'Density\\n(P/Km2)': 'Population density (people per km2)',
    'Land Area(Km2)': 'Land area (km2)',
    'Latitude': 'Latitude',
    'Longitude': 'Longitude'
})
sea_data
```

	Country	Year	Access to electricity (% of population)	Access to clean fuels for cooking (% of population)	Installed renewable electricity capacity per capita (kW)	Financial flows to developing countries (US$)	Renewable energy share in total final energy consumption (%)	Electricity from fossil fuels (TWh)	Electricity from nuclear (TWh)	Electricity from renewables (TWh)	...	Primary energy consumption per capita (kWh)	Energy intensity level of primary energy (MJ/$2017 PPP GDP)	CO2 emissions per capita (metric tons)	Renewables (% equivalent primary energy)
567	Cambodia	2000	16.60	3.60	0.83	NaN	81.58	0.42	0.00	0.05	...	739.46	7.91	1960.00	NaN
568	Cambodia	2001	14.62	4.10	0.82	NaN	80.51	0.48	0.00	0.03	...	769.41	7.51	2150.00	NaN
569	Cambodia	2002	18.19	4.70	0.97	NaN	80.97	0.66	0.00	0.03	...	754.14	7.05	2210.00	NaN
570	Cambodia	2003	19.30	5.30	0.96	4630000.00	79.92	0.67	0.00	0.04	...	803.82	6.69	2380.00	NaN
571	Cambodia	2004	25.30	5.90	1.02	409150000.00	80.69	0.73	0.00	0.03	...	813.56	6.23	2380.00	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3266	Thailand	2016	99.85	80.10	136.90	420000.00	22.43	161.79	0.00	15.97	...	19981.58	5.02	261600.01	4.60
3267	Thailand	2017	99.90	80.80	148.00	1990000.00	22.25	161.88	0.00	19.92	...	20264.65	4.79	258820.01	5.31
3268	Thailand	2018	99.82	81.70	163.82	296130000.00	23.70	156.26	0.00	25.84	...	20834.09	4.50	257049.99	6.38
3269	Thailand	2019	99.90	82.70	170.30	99160000.00	23.96	162.59	0.00	28.02	...	20792.38	4.52	267090.00	6.96
3270	Thailand	2020	100.00	83.60	171.79	NaN	NaN	154.52	0.00	24.73	...	19298.09	NaN	NaN	6.99

147 rows × 21 columns

Figure 3.2.9

GDP growth rate (annual %)	GDP per capita	Population density (people per km2)	Land area (km2)	Latitude	Longitude
9.99	300.61	95	181035.00	12.57	104.99
8.15	321.15	95	181035.00	12.57	104.99
6.58	338.99	95	181035.00	12.57	104.99
8.51	362.34	95	181035.00	12.57	104.99
10.34	408.51	95	181035.00	12.57	104.99
...	...	...	...	...	...
3.44	5993.31	137	513120.00	15.87	100.99
4.18	6593.82	137	513120.00	15.87	100.99
4.19	7296.88	137	513120.00	15.87	100.99
2.27	7817.01	137	513120.00	15.87	100.99
-6.10	7186.87	137	513120.00	15.87	100.99

Figure 3.2.10

Figure 3.2.10 and Figure 3.2.10 shows the dataset after renaming the columns. This step removes the unwanted characters in column names and allows user to have a better understanding about the data represented by referring to the column names.

Step 5: Handle incorrect data types

```
# Display unique values to review
sea_data['Population density (people per km2)'].unique()
```

```
array(['95', '151', '99', '83', '368', '8,358', '137'], dtype=object)
```

Figure 3.2.11

The output shows the unique values recorded in the population density column. Values are recorded as string due to the existence of comma among the numbers, such as 8,358. So, the comma needed to be removed before the conversion of data type happens.

```
# Remove commas from the 'Population density (people per km2)' column
sea_data['Population density (people per km2)'] = sea_data['Population
density (people per km2)'].str.replace(',', '')

# Display unique values to review
sea_data['Population density (people per km2)'].unique()
```

```
array(['95', '151', '99', '83', '368', '8358', '137'], dtype=object)
```

Figure 3.2.12

The comma among the numbers has been replaced and Figure 3.2.12 above shows the unique values in population density column of the dataset after being replaced where the comma has been removed but the data type remains as object.

```
# Convert the column to numeric type (float), forcing errors to NaN if any
sea_data['Population density (people per km2)'] =
pd.to_numeric(sea_data['Population density (people per km2)'],
errors='coerce')

# Verify the change

sea_data['Population density (people per km2)'].dtype
```

```
dtype('int64')
```

Figure 3.2.13

The values in the data are converted into numeric type which is float, and the final data type of population density is displayed as int64, which indicates that the data type of population density column is converted into numeric successfully.

Step 6: Handle missing values

```
# Check number of missing values in each column
sea_data.isnull().sum()
```

	0
Country	0
Year	0
Access to electricity (% of population)	0
Access to clean fuels for cooking (% of population)	0
Installed renewable electricity capacity per capita (kW)	0
Financial flows to developing countries (US$)	45
Renewable energy share in total final energy consumption (%)	7
Electricity from fossil fuels (TWh)	0
Electricity from nuclear (TWh)	42
Electricity from renewables (TWh)	0
Low-carbon electricity (% of electricity)	0
Primary energy consumption per capita (kWh)	0
Energy intensity level of primary energy (MJ/$2017 PPP GDP)	7
CO2 emissions per capita (metric tons)	7
Renewables (% equivalent primary energy)	42
GDP growth rate (annual %)	0
GDP per capita	0
Population density (people per km2)	0
Land area (km2)	0
Latitude	0
Longitude	0

dtype: int64

Figure 3.2.14

Figure 3.2.14 displays the number of missing values for each column where six columns contain missing values. These six columns include financial flows to developing countries, renewable energy share in total final energy consumption, electricity from nuclear, energy intensity level of primary energy, CO2 emissions per capita, and renewables.

```
# Calculate the percentage of missing values per column
 (((sea_data.isna().sum())/sea_data.shape[0])*100).sort_values(ascending =
False)
```

	0
Financial flows to developing countries (US$)	30.61
Electricity from nuclear (TWh)	28.57
Renewables (% equivalent primary energy)	28.57
Renewable energy share in total final energy consumption (%)	4.76
CO2 emissions per capita (metric tons)	4.76
Energy intensity level of primary energy (MJ/$2017 PPP GDP)	4.76
Installed renewable electricity capacity per capita (kW)	0.00
Access to electricity (% of population)	0.00
Year	0.00
Country	0.00
Access to clean fuels for cooking (% of population)	0.00
Electricity from fossil fuels (TWh)	0.00
Electricity from renewables (TWh)	0.00
Primary energy consumption per capita (kWh)	0.00
Low-carbon electricity (% of electricity)	0.00
GDP growth rate (annual %)	0.00
GDP per capita	0.00
Population density (people per km2)	0.00
Land area (km2)	0.00
Latitude	0.00
Longitude	0.00

Figure 3.2.15

Figure 3.2.15 shows the percentage of missing values exist in each column. The financial flows to developing countries column has the highest percentage of missing values.

```python
# Columns to check
columns_to_check = [
    'Financial flows to developing countries (US$)',
    'Electricity from nuclear (TWh)',
    'Renewables (% equivalent primary energy)',
    'Renewable energy share in total final energy consumption (%)',
    'CO2 emissions per capita (metric tons)',
    'Energy intensity level of primary energy (MJ/$2017 PPP GDP)'
]

# Calculate missing value percentage by year
missing_by_year = sea_data.groupby('Year')[columns_to_check].apply(
    lambda x: (x.isna().sum() / len(x)) * 100
)

missing_by_year
```

Year	Financial flows to developing countries (US$)	Electricity from nuclear (TWh)	Renewables (% equivalent primary energy)	Renewable energy share in total final energy consumption (%)	CO2 emissions per capita (metric tons)	Energy intensity level of primary energy (MJ/$2017 PPP GDP)
2000	42.86	28.57	28.57	0.00	0.00	0.00
2001	57.14	28.57	28.57	0.00	0.00	0.00
2002	28.57	28.57	28.57	0.00	0.00	0.00
2003	28.57	28.57	28.57	0.00	0.00	0.00
2004	14.29	28.57	28.57	0.00	0.00	0.00
2005	57.14	28.57	28.57	0.00	0.00	0.00
2006	28.57	28.57	28.57	0.00	0.00	0.00
2007	42.86	28.57	28.57	0.00	0.00	0.00
2008	57.14	28.57	28.57	0.00	0.00	0.00
2009	28.57	28.57	28.57	0.00	0.00	0.00
2010	14.29	28.57	28.57	0.00	0.00	0.00
2011	28.57	28.57	28.57	0.00	0.00	0.00
2012	14.29	28.57	28.57	0.00	0.00	0.00
2013	14.29	28.57	28.57	0.00	0.00	0.00
2014	14.29	28.57	28.57	0.00	0.00	0.00
2015	14.29	28.57	28.57	0.00	0.00	0.00
2016	14.29	28.57	28.57	0.00	0.00	0.00
2017	14.29	28.57	28.57	0.00	0.00	0.00
2018	14.29	28.57	28.57	0.00	0.00	0.00
2019	14.29	28.57	28.57	0.00	0.00	0.00
2020	100.00	28.57	28.57	100.00	100.00	100.00

Figure 3.2.16

From Figure 3.2.16, almost all the data in 2020 is not recorded.

```python
# Filter out rows where Year is 2020
sea_data = sea_data[sea_data['Year'] != 2020]
sea_data
```

	Country	Year	Access to electricity (% of population)	Access to clean fuels for cooking (% of population)	Installed renewable electricity capacity per capita (kW)	Financial flows to developing countries (US$)	Renewable energy share in total final energy consumption (%)	Electricity from fossil fuels (TWh)	Electricity from nuclear (TWh)	Electricity from renewables (TWh)	...	Primary energy consumption per capita (kWh)	Energy intensity level of primary energy (MJ/$2017 PPP GDP)	CO2 emissions per capita (metric tons)	Renewables (% equivalent primary energy)
567	Cambodia	2000	16.60	3.60	0.83	NaN	81.58	0.42	0.00	0.05	...	739.46	7.91	1960.00	NaN
568	Cambodia	2001	14.62	4.10	0.82	NaN	80.51	0.48	0.00	0.03	...	769.41	7.51	2150.00	NaN
569	Cambodia	2002	18.19	4.70	0.97	NaN	80.97	0.66	0.00	0.03	...	754.14	7.05	2210.00	NaN
570	Cambodia	2003	19.30	5.30	0.96	4630000.00	79.92	0.67	0.00	0.04	...	803.82	6.69	2380.00	NaN
571	Cambodia	2004	25.30	5.90	1.02	409150000.00	80.69	0.73	0.00	0.03	...	813.56	6.23	2380.00	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3265	Thailand	2015	99.60	78.90	115.96	55520000.00	22.65	153.40	0.00	13.33	...	19661.77	5.07	264000.00	4.19
3266	Thailand	2016	99.85	80.10	136.90	420000.00	22.43	161.79	0.00	15.97	...	19981.58	5.02	261600.01	4.60
3267	Thailand	2017	99.90	80.80	148.00	1990000.00	22.25	161.88	0.00	19.92	...	20264.65	4.79	258820.01	5.31
3268	Thailand	2018	99.82	81.70	163.82	296130000.00	23.70	156.26	0.00	25.84	...	20834.09	4.50	257049.99	6.38
3269	Thailand	2019	99.90	82.70	170.30	99160000.00	23.96	162.59	0.00	28.02	...	20792.38	4.52	267090.00	6.96

140 rows × 21 columns

Figure 3.2.17

Renewable energy share in total final energy consumption (%)	Electricity from fossil fuels (TWh)	Electricity from nuclear (TWh)	Electricity from renewables (TWh)	...	Primary energy consumption per capita (kWh)	Energy intensity level of primary energy (MJ/$2017 PPP GDP)	CO2 emissions per capita (metric tons)	Renewables (% equivalent primary energy)	GDP growth rate (annual %)	GDP per capita	Population density (people per km2)	Land area (km2)	Latitude	Longitude
81.58	0.42	0.00	0.05	...	739.46	7.91	1960.00	NaN	9.99	300.61	95	181035.00	12.57	104.99
80.51	0.48	0.00	0.03	...	769.41	7.51	2150.00	NaN	8.15	321.15	95	181035.00	12.57	104.99
80.97	0.66	0.00	0.03	...	754.14	7.05	2210.00	NaN	6.58	338.99	95	181035.00	12.57	104.99
79.92	0.67	0.00	0.04	...	803.82	6.69	2380.00	NaN	8.51	362.34	95	181035.00	12.57	104.99
80.69	0.73	0.00	0.03	...	813.56	6.23	2380.00	NaN	10.34	408.51	95	181035.00	12.57	104.99
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
22.65	153.40	0.00	13.33	...	19661.77	5.07	264000.00	4.19	3.13	5840.05	137	513120.00	15.87	100.99
22.43	161.79	0.00	15.97	...	19981.58	5.02	261600.01	4.60	3.44	5993.31	137	513120.00	15.87	100.99
22.25	161.88	0.00	19.92	...	20264.65	4.79	258820.01	5.31	4.18	6593.82	137	513120.00	15.87	100.99
23.70	156.26	0.00	25.84	...	20834.09	4.50	257049.99	6.38	4.19	7296.88	137	513120.00	15.87	100.99
23.96	162.59	0.00	28.02	...	20792.38	4.52	267090.00	6.96	2.27	7817.01	137	513120.00	15.87	100.99

Figure 3.2.18

Figure 3.2.17 and Figure 3.2.18 show the dataset after removing data recorded for year 2020.

```python
# Check the unique values for the Year column
sea_data['Year'].unique()
```

```
array([2000, 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009, 2010,
       2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018, 2019])
```

Figure 3.2.19

Figure 3.2.19 shows that the dataset now contains data from year 2000 to 2019 only.

```
# Percentage of missing values in our dataset

(((sea_data.isna().sum())/sea_data.shape[0])*100).sort_values(ascending =
False)
```

Electricity from nuclear (TWh)	28.57
Renewables (% equivalent primary energy)	28.57
Financial flows to developing countries (US$)	27.14
Year	0.00
Country	0.00
Installed renewable electricity capacity per capita (kW)	0.00
Access to clean fuels for cooking (% of population)	0.00
Access to electricity (% of population)	0.00
Renewable energy share in total final energy consumption (%)	0.00
Electricity from renewables (TWh)	0.00
Low-carbon electricity (% of electricity)	0.00
Primary energy consumption per capita (kWh)	0.00
Electricity from fossil fuels (TWh)	0.00
Energy intensity level of primary energy (MJ/$2017 PPP GDP)	0.00
CO2 emissions per capita (metric tons)	0.00
GDP growth rate (annual %)	0.00
GDP per capita	0.00
Population density (people per km2)	0.00
Land area (km2)	0.00
Latitude	0.00
Longitude	0.00

Figure 3.2.20

After removing the data for year 2020, only three columns contain missing values which are electricity from nuclear, renewables (% equivalent primary energy), and financial flows to developing countries.

```
# Drop columns that is not essential
sea_data = sea_data.drop(['Financial flows to developing countries
(US$)'],axis=1)
# Replace missing values in nuclear electricity with 0
sea_data['Electricity from nuclear (TWh)'] = sea_data['Electricity from
nuclear (TWh)'].fillna(0)
# Check number of missing values
sea_data.isnull().sum()
```

	0
Country	0
Year	0
Access to electricity (% of population)	0
Access to clean fuels for cooking (% of population)	0
Installed renewable electricity capacity per capita (kW)	0
Renewable energy share in total final energy consumption (%)	0
Electricity from fossil fuels (TWh)	0
Electricity from nuclear (TWh)	0
Electricity from renewables (TWh)	0
Low-carbon electricity (% of electricity)	0
Primary energy consumption per capita (kWh)	0
Energy intensity level of primary energy (MJ/$2017 PPP GDP)	0
CO2 emissions per capita (metric tons)	0
Renewables (% equivalent primary energy)	40
GDP growth rate (annual %)	0
GDP per capita	0
Population density (people per km2)	0
Land area (km2)	0
Latitude	0
Longitude	0

Figure 3.2.21

The column named financial flows to developing countries is dropped because the data from this column does not relate much to our main analysis. While for the column electricity from nuclear, missing values are replaced by 0 as Southeast Asia are one of the nuclear free zones in fact and missing values only exist in countries that are not in Southeast Asia. Hence, it is reasonable to replace missing values to 0 for Southeast Asian countries. After the processes, only the column renewables (% equivalent primary energy) contain missing values according to Figure 3.2.21.

```python
# Identify rows where the value is missing
missing_renewables = sea_data[sea_data['Renewables (% equivalent primary
energy)'].isna()]

# Display unique countries with missing values
missing_renewables['Country'].unique()
```

```
array(['Cambodia', 'Myanmar'], dtype=object)
```

Figure 3.2.22

The rows with missing values are identified and gathered into a new data frame called missing_renewables. From the new data frame formed, it is found that only Cambodia and Myanmar have missing values in the column Renewables (% equivalent primary energy). Since this indicator involved in our data analysis, a new dataset without these two countries will be created and used when this indicator is involved.

```
# Create a filtered version of the dataset for analysis involving this column
sea_data_renewables = sea_data.dropna(subset=['Renewables (% equivalent primary energy)'])
sea_data_renewables
```

	Country	Year	Access to electricity (% of population)	Access to clean fuels for cooking (% of population)	Installed renewable electricity capacity per capita (kW)	Renewable energy share in total final energy consumption (%)	Electricity from fossil fuels (TWh)	Electricity from nuclear (TWh)	Electricity from renewables (TWh)	Low-carbon electricity (% of electricity)	Primary energy consumption per capita (kWh)	Energy intensity level of primary energy (MJ/$2017 PPP GDP)	CO2 emissions per capita (metric tons)	Renewables (% equivalent primary energy)	GDP growth rate (annual %)
1555	Indonesia	2000	86.30	6.10	25.89	45.62	78.43	0.00	19.60	19.99	5435.44	5.42	280650.00	3.73	4.92
1556	Indonesia	2001	86.26	6.90	21.71	44.33	83.96	0.00	22.19	20.90	5740.43	5.34	302060.00	4.11	3.64
1557	Indonesia	2002	87.60	8.00	21.61	44.66	92.03	0.00	21.00	18.58	5815.39	5.30	305640.01	3.68	4.50
1558	Indonesia	2003	87.94	9.30	21.48	43.00	97.57	0.00	19.82	16.88	6224.94	5.07	333890.01	3.22	4.78
1559	Indonesia	2004	89.01	11.30	21.40	41.46	103.80	0.00	20.97	16.81	6087.68	5.14	341239.99	3.43	5.03
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
3265	Thailand	2015	99.60	78.90	115.96	22.65	153.40	0.00	13.33	7.99	19661.77	5.07	264000.00	4.19	3.13
3266	Thailand	2016	99.85	80.10	136.90	22.43	161.79	0.00	15.97	8.98	19981.58	5.02	261600.01	4.60	3.44
3267	Thailand	2017	99.90	80.80	148.00	22.25	161.88	0.00	19.92	10.96	20264.65	4.79	258820.01	5.31	4.18
3268	Thailand	2018	99.82	81.70	163.82	23.70	156.26	0.00	25.84	14.19	20834.09	4.50	257049.99	6.38	4.19
3269	Thailand	2019	99.90	82.70	170.30	23.96	162.59	0.00	28.02	14.70	20792.38	4.52	267090.00	6.96	2.27

100 rows × 20 columns

Figure 3.2.23

GDP per capita	Population density (people per km2)	Land area (km2)	Latitude	Longitude
780.19	151	1904569.00	-0.79	113.92
748.26	151	1904569.00	-0.79	113.92
900.18	151	1904569.00	-0.79	113.92
1065.65	151	1904569.00	-0.79	113.92
1150.26	151	1904569.00	-0.79	113.92
...	...	...	...	...
5840.05	137	513120.00	15.87	100.99
5993.31	137	513120.00	15.87	100.99
6593.82	137	513120.00	15.87	100.99
7296.88	137	513120.00	15.87	100.99
7817.01	137	513120.00	15.87	100.99

Figure 3.2.24

A filtered dataset named sea_data_renewables is created without involving the two countries with missing values in the column Renewables (% equivalent primary energy).

```
# Check the unique values for the Country column in the filtered dataset
sea_data_renewables['Country'].unique()
```

```
array(['Indonesia', 'Malaysia', 'Philippines', 'Singapore', 'Thailand'],
      dtype=object)
```

Figure 3.2.25

In the new filtered dataset, only five countries in Southeast Asia are involved as Cambodia and Myanmar have missing values in the column Renewables (% equivalent primary energy).

```
# Check number of missing values in the filtered dataset
sea_data_renewables.isnull().sum()
```

	0
Country	0
Year	0
Access to electricity (% of population)	0
Access to clean fuels for cooking (% of population)	0
Installed renewable electricity capacity per capita (kW)	0
Renewable energy share in total final energy consumption (%)	0
Electricity from fossil fuels (TWh)	0
Electricity from nuclear (TWh)	0
Electricity from renewables (TWh)	0
Low-carbon electricity (% of electricity)	0
Primary energy consumption per capita (kWh)	0
Energy intensity level of primary energy (MJ/$2017 PPP GDP)	0
CO2 emissions per capita (metric tons)	0
Renewables (% equivalent primary energy)	0
GDP growth rate (annual %)	0
GDP per capita	0
Population density (people per km2)	0
Land area (km2)	0
Latitude	0
Longitude	0

Figure 3.2.26

The new filtered dataset is checked for the existence of missing values.

```
# Drop "Renewables (% equivalent primary energy)" column
sea_data = sea_data.drop(['Renewables (% equivalent primary
energy)'],axis=1)
# Check number of missing values in the dataset
sea_data.isnull().sum()
```

	0
Country	0
Year	0
Access to electricity (% of population)	0
Access to clean fuels for cooking (% of population)	0
Installed renewable electricity capacity per capita (kW)	0
Renewable energy share in total final energy consumption (%)	0
Electricity from fossil fuels (TWh)	0
Electricity from nuclear (TWh)	0
Electricity from renewables (TWh)	0
Low-carbon electricity (% of electricity)	0
Primary energy consumption per capita (kWh)	0
Energy intensity level of primary energy (MJ/$2017 PPP GDP)	0
CO2 emissions per capita (metric tons)	0
GDP growth rate (annual %)	0
GDP per capita	0
Population density (people per km2)	0
Land area (km2)	0
Latitude	0
Longitude	0

Figure 3.2.27

To avoid confusion, the column named Renewables (% equivalent primary energy) is removed from the sea_data, the main dataset used in this project. Figure 3.2.27 shows that there are no missing values in the sea_data.

Step 7: Check for duplicates

```
# Check for duplicate values in main dataset
sea_data.duplicated().sum()
```

```
np.int64(0)
```

Figure 3.2.28

Figure 3.2.28 shows that sea_data does not contain any duplicates.

```
# Check for duplicate values in the filtered dataset
sea_data_renewables.duplicated().sum()
```

```
np.int64(0)
```

Figure 3.2.29

Figure 3.2.29 shows that sea_data_renewables does not contain any duplicates.

Step 8: Check for outliers using boxplot

```python
# Visualize numeric variables in sea_data using boxplots to detect
potential outliers
import seaborn as sns
import matplotlib.pyplot as plt

cols_to_plot = [
    'Access to electricity (% of population)',
    'Access to clean fuels for cooking (% of population)',
    'Installed renewable electricity capacity per capita (kW)',
    'Renewable energy share in total final energy consumption (%)',
    'Electricity from fossil fuels (TWh)',
    'Electricity from nuclear (TWh)',
    'Electricity from renewables (TWh)',
    'Low-carbon electricity (% of electricity)',
    'Primary energy consumption per capita (kWh)',
    'Energy intensity level of primary energy (MJ/$2017 PPP GDP)',
    'CO2 emissions per capita (metric tons)',
    'GDP growth rate (annual %)',
    'GDP per capita',
    'Population density (people per km2)',
    'Land area (km2)'
]

for col in cols_to_plot:
    plt.figure(figsize=(8, 4))
    sns.boxplot(x=sea_data[col])
    plt.title(f'Boxplot of {col}')
    plt.show()
```

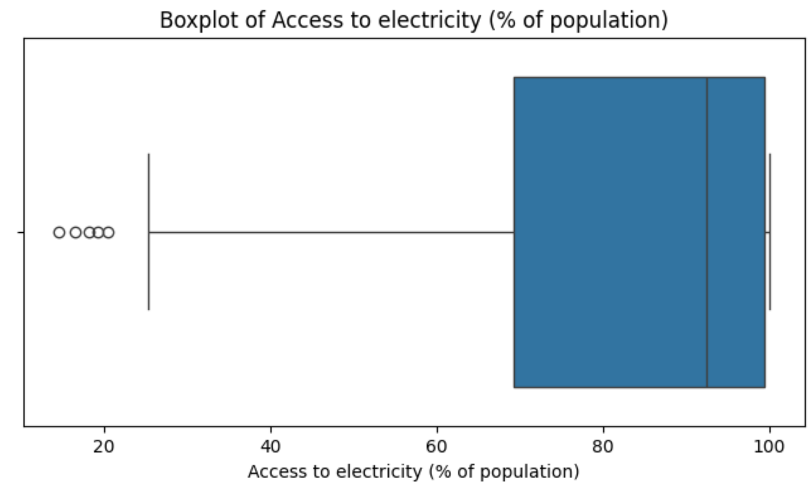

Figure 3.2.30

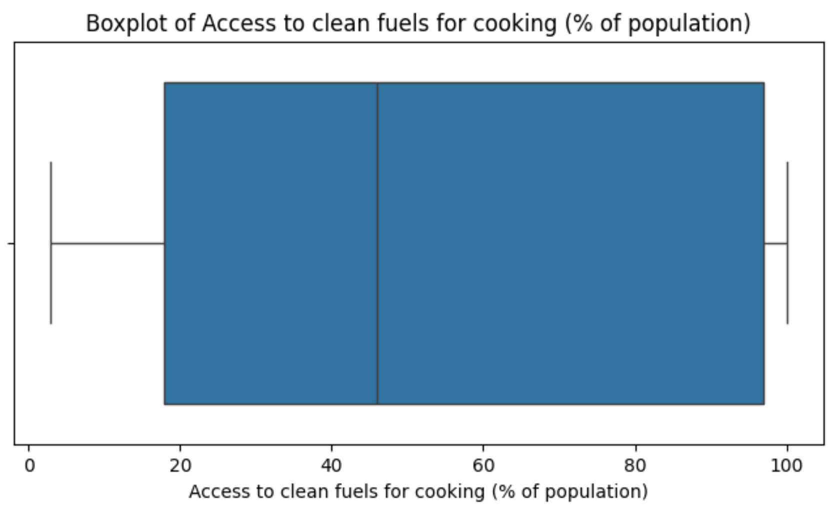

Figure 3.2.31

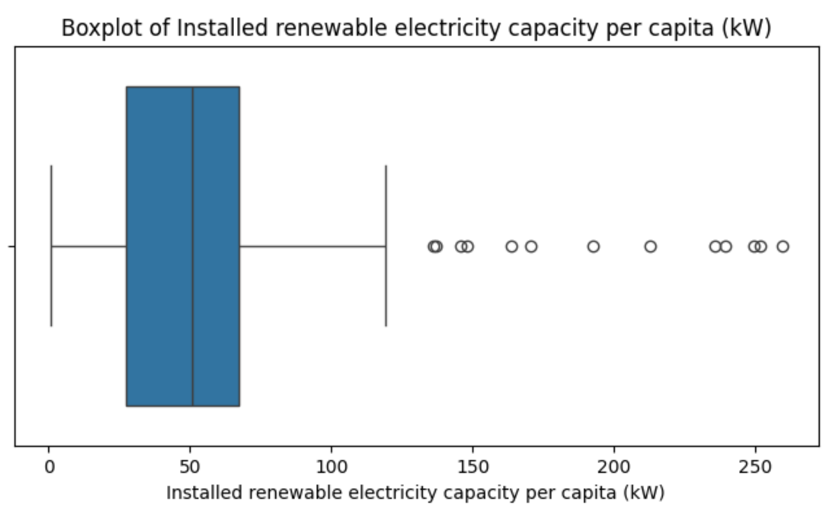

Figure 3.2.32

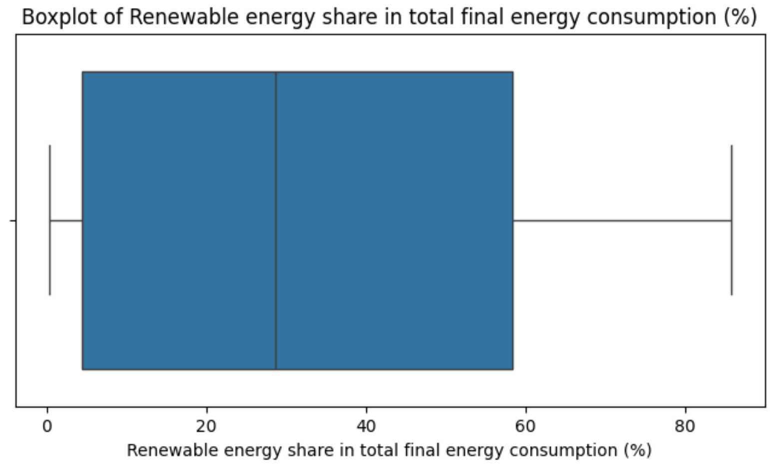

Figure 3.2.33

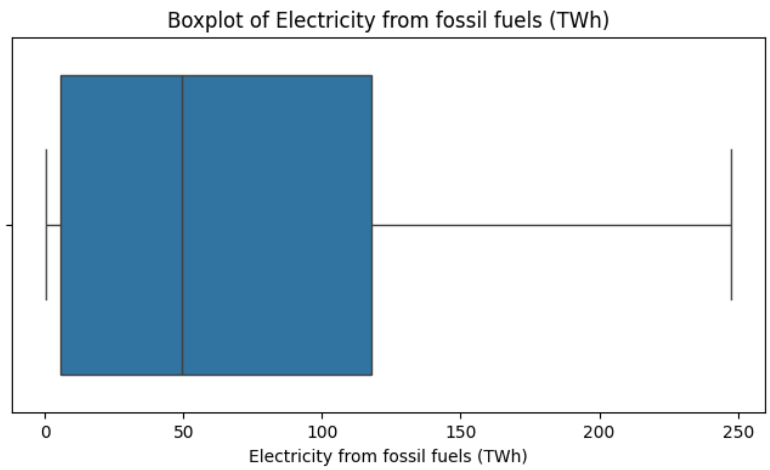

Figure 3.2.34

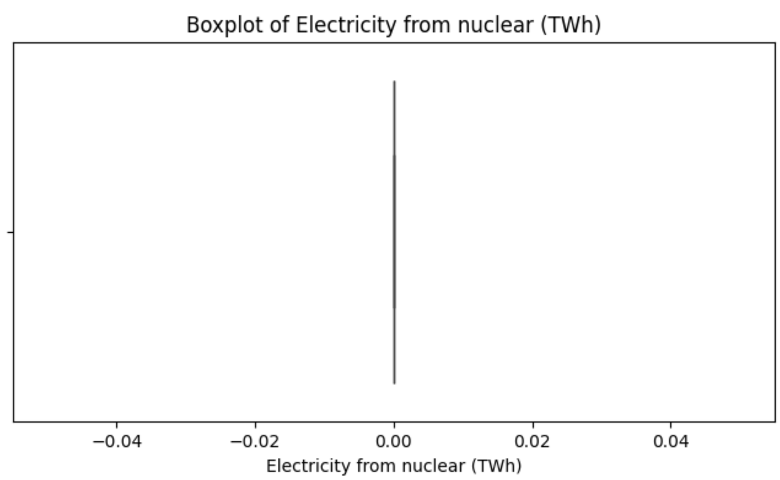

Figure 3.2.35

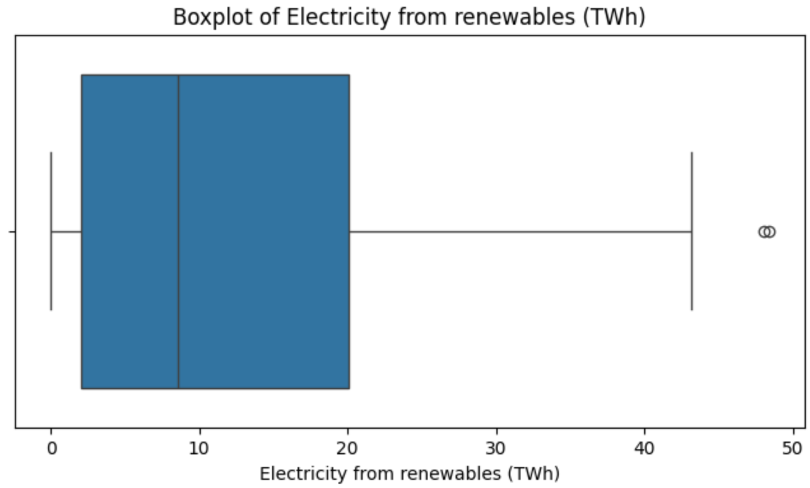

Figure 3.2.36

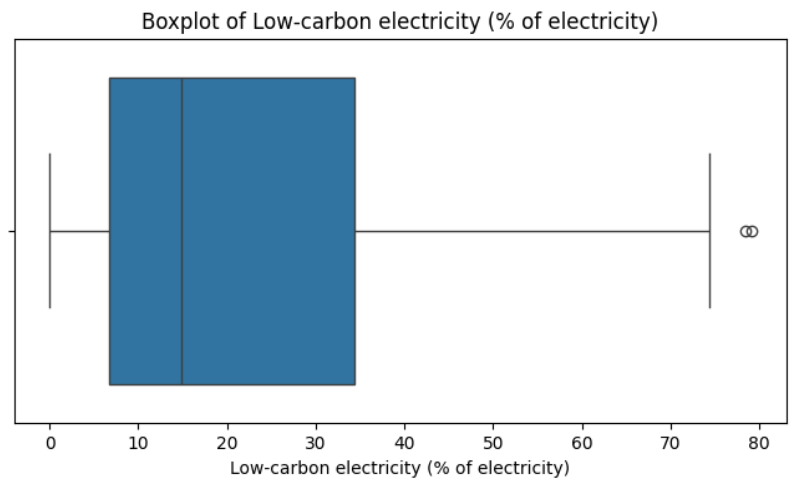

Figure 3.2.37

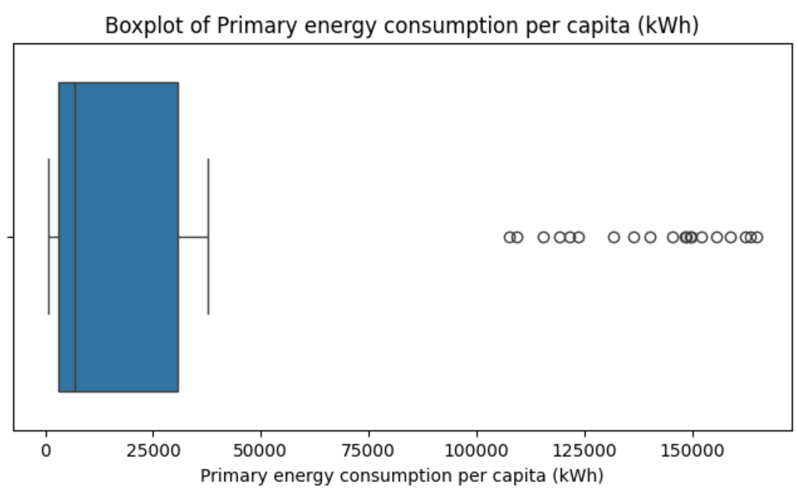

Figure 3.2.38

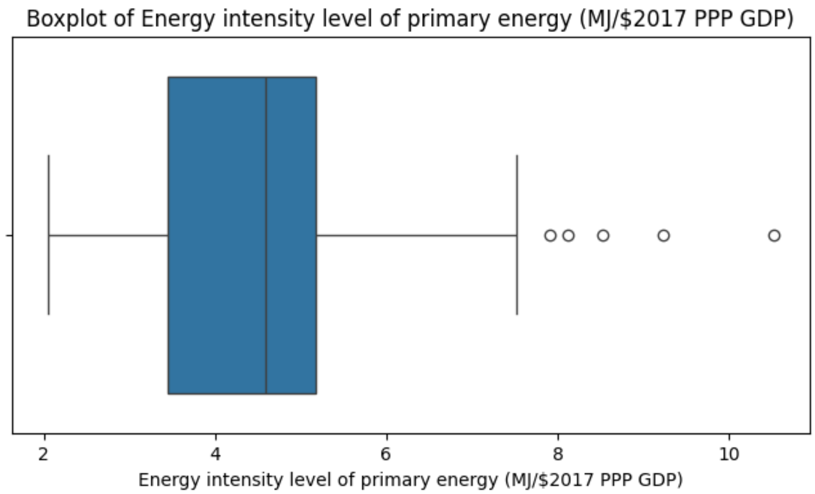

Figure 3.2.39

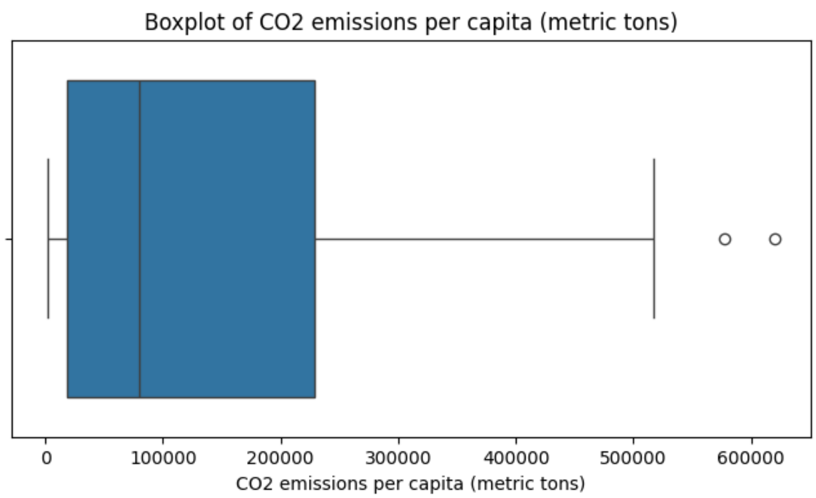

Figure 3.2.40

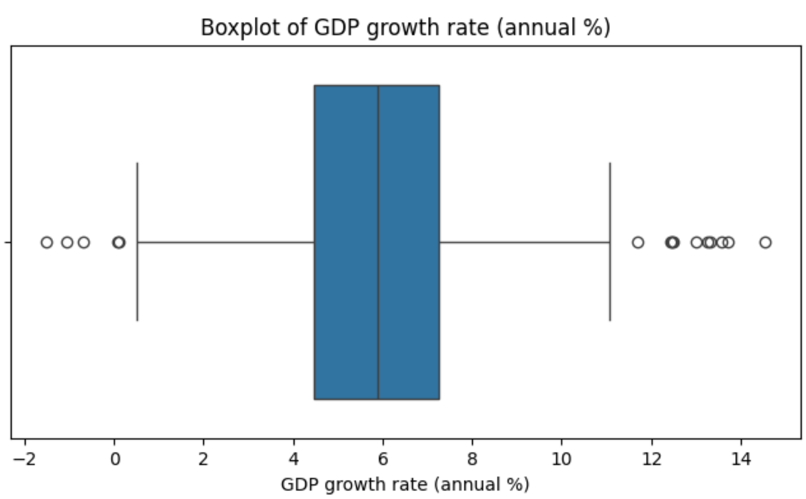

Figure 3.2.41

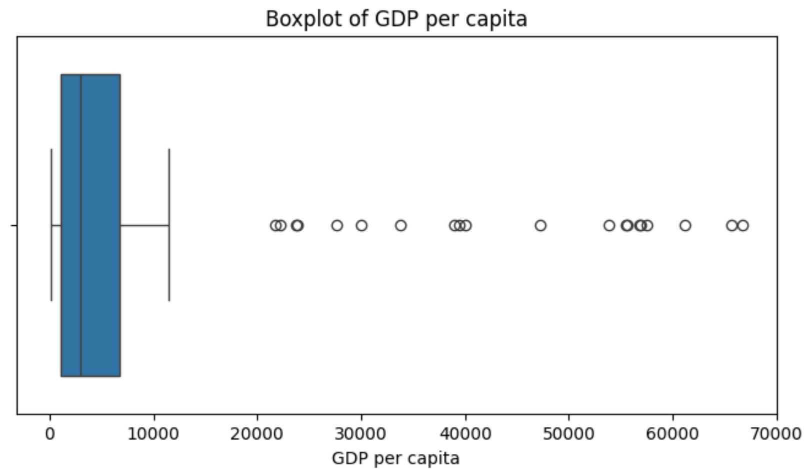

Figure 3.2.42

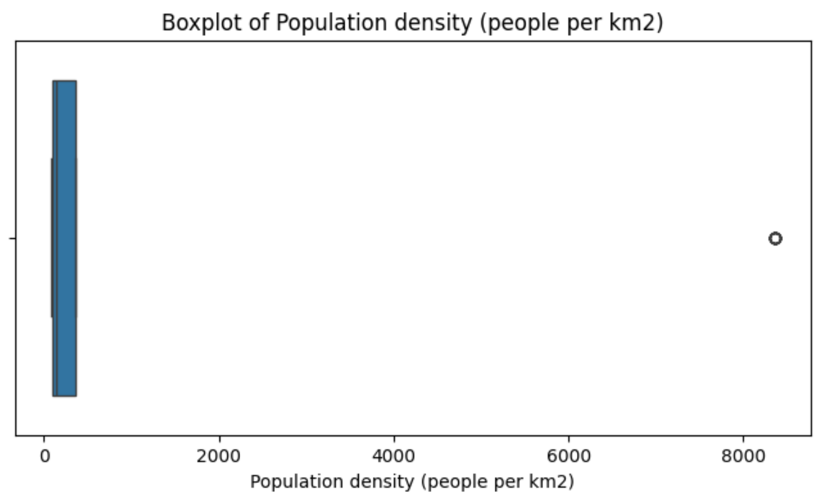

Figure 3.2.43

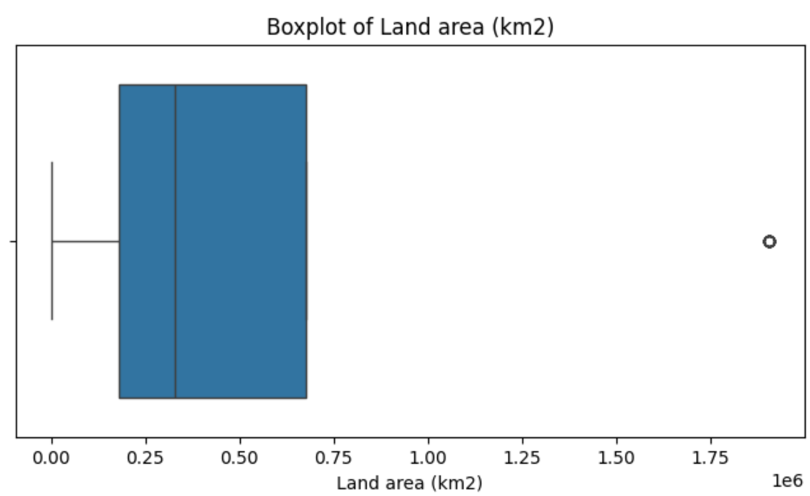

Figure 3.2.44

Figure 3.2.30 to Figure 3.2.44 display boxplots for each column in sea_data. Out of all columns, only 10 contain outliers. However, the outliers remained as they likely represent real variations in the data rather than errors.

```
# Visualize Renewables (% equivalent primary energy) in
sea_data_renewables using boxplot to detect potential outliers

plt.figure(figsize=(8, 4))
sns.boxplot(x=sea_data_renewables['Renewables (% equivalent primary
energy)'])
plt.title('Boxplot of Renewables (% equivalent primary energy)')
plt.show()
```

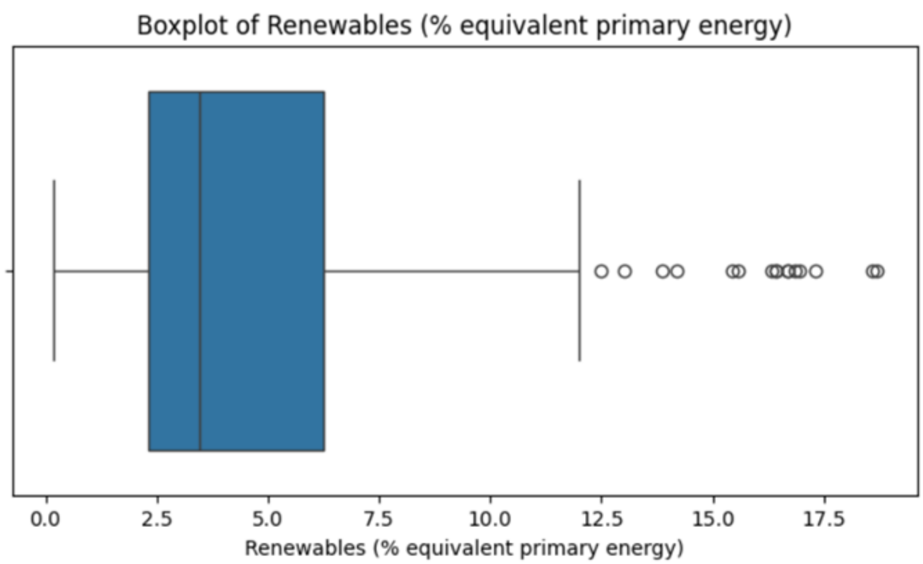

Figure 3.2.45

Figure 3.2.45 displays the boxplot for the column named Renewables (% equivalent primary energy) from sea_data_renewables. The outliers in this column also kept for further analysis.

Step 9: Standardize inconsistent data

For the main dataset (sea_data):

```
# List of columns to exclude from rounding
exclude_cols = ['Year', 'Population density (people per km2)', 'Latitude',
'Longitude']

# Select only float columns excluding those in exclude_cols
cols_to_round = [col for col in
sea_data.select_dtypes(include='float').columns if col not in
exclude_cols]

# Round selected columns to 2 decimal places
sea_data[cols_to_round] = sea_data[cols_to_round].round(2)
# Display the standardized dataset
sea_data
```

	Country	Year	Access to electricity (% of population)	Access to clean fuels for cooking (% of population)	Installed renewable electricity capacity per capita (kW)	Renewable energy share in total final energy consumption (%)	Electricity from fossil fuels (TWh)	Electricity from nuclear (TWh)	Electricity from renewables (TWh)	Low-carbon electricity (% of electricity)	Primary energy consumption per capita (kWh)	Energy intensity level of primary energy (MJ/$2017 PPP GDP)	CO2 emissions per capita (metric tons)	GDP growth rate (annual %)	GDP per capita
0	Cambodia	2000	16.60	3.60	0.83	81.58	0.42	0.00	0.05	10.64	739.46	7.91	1960.00	9.99	300.61
1	Cambodia	2001	14.62	4.10	0.82	80.51	0.48	0.00	0.03	5.88	769.41	7.51	2150.00	8.15	321.15
2	Cambodia	2002	18.19	4.70	0.97	80.97	0.66	0.00	0.03	4.35	754.14	7.05	2210.00	6.58	338.99
3	Cambodia	2003	19.30	5.30	0.96	79.92	0.67	0.00	0.04	5.63	803.82	6.69	2380.00	8.51	362.34
4	Cambodia	2004	25.30	5.90	1.02	80.69	0.73	0.00	0.03	3.95	813.56	6.23	2380.00	10.34	408.51
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
135	Thailand	2015	99.60	78.90	115.96	22.65	153.40	0.00	13.33	7.99	19661.77	5.07	264000.00	3.13	5840.05
136	Thailand	2016	99.85	80.10	136.90	22.43	161.79	0.00	15.97	8.98	19981.58	5.02	261600.01	3.44	5993.31
137	Thailand	2017	99.90	80.80	148.00	22.25	161.88	0.00	19.92	10.96	20264.65	4.79	258820.01	4.18	6593.82
138	Thailand	2018	99.82	81.70	163.82	23.70	156.26	0.00	25.84	14.19	20834.09	4.50	257049.99	4.19	7296.88
139	Thailand	2019	99.90	82.70	170.30	23.96	162.59	0.00	28.02	14.70	20792.38	4.52	267090.00	2.27	7817.01

140 rows × 19 columns

Figure 3.2.46

Population density (people per km2)	Land area (km2)	Latitude	Longitude
95	181035.00	12.57	104.99
95	181035.00	12.57	104.99
95	181035.00	12.57	104.99
95	181035.00	12.57	104.99
95	181035.00	12.57	104.99
...	...	...	...
137	513120.00	15.87	100.99
137	513120.00	15.87	100.99
137	513120.00	15.87	100.99
137	513120.00	15.87	100.99
137	513120.00	15.87	100.99

Figure 3.2.47

Figure 3.2.46 and Figure 3.2.47 show that the float data in sea_data has been rounded to two decimal points, except for the columns Year, Latitude, Longitude and Population Density.

For the filtered dataset (sea_data_renewables):

```
# List of columns to exclude from rounding
exclude_cols = ['Year', 'Population density (people per km2)', 'Latitude',
'Longitude']

# Select only float columns excluding those in exclude_cols
cols_to_round_renewables = [col for col in
sea_data_renewables.select_dtypes(include='float').columns if col not in
exclude_cols]

# Round selected columns to 2 decimal places
sea_data_renewables.loc[:, cols_to_round_renewables] =
sea_data_renewables[cols_to_round_renewables].round(2)

# Display the standardized dataset
sea_data_renewables
```

	Country	Year	Access to electricity (% of population)	Access to clean fuels for cooking (% of population)	Installed renewable electricity capacity per capita (kW)	Renewable energy share in total final energy consumption (%)	Electricity from fossil fuels (TWh)	Electricity from nuclear (TWh)	Electricity from renewables (TWh)	Low-carbon electricity (% of electricity)	Primary energy consumption per capita (kWh)	Energy intensity level of primary energy (MJ/$2017 PPP GDP)	CO2 emissions per capita (metric tons)	Renewables (% equivalent primary energy)	GDP growth rate (annual %)
0	Indonesia	2000	86.30	6.10	25.89	45.62	78.43	0.00	19.60	19.99	5435.44	5.42	280650.00	3.73	4.92
1	Indonesia	2001	86.26	6.90	21.71	44.33	83.96	0.00	22.19	20.90	5740.43	5.34	302060.00	4.11	3.64
2	Indonesia	2002	87.60	8.00	21.61	44.66	92.03	0.00	21.00	18.58	5815.39	5.30	305640.01	3.68	4.50
3	Indonesia	2003	87.94	9.30	21.48	43.00	97.57	0.00	19.82	16.88	6224.94	5.07	333890.01	3.22	4.78
4	Indonesia	2004	89.01	11.30	21.40	41.46	103.80	0.00	20.97	16.81	6087.68	5.14	341239.99	3.43	5.03
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
95	Thailand	2015	99.60	78.90	115.96	22.65	153.40	0.00	13.33	7.99	19661.77	5.07	264000.00	4.19	3.13
96	Thailand	2016	99.85	80.10	136.90	22.43	161.79	0.00	15.97	8.98	19981.58	5.02	261600.01	4.60	3.44
97	Thailand	2017	99.90	80.80	148.00	22.25	161.88	0.00	19.92	10.96	20264.65	4.79	258820.01	5.31	4.18
98	Thailand	2018	99.82	81.70	163.82	23.70	156.26	0.00	25.84	14.19	20834.09	4.50	257049.99	6.38	4.19
99	Thailand	2019	99.90	82.70	170.30	23.96	162.59	0.00	28.02	14.70	20792.38	4.52	267090.00	6.96	2.27

100 rows × 20 columns

Figure 3.2.48

GDP per capita	Population density (people per km2)	Land area (km2)	Latitude	Longitude
780.19	151	1904569.00	-0.79	113.92
748.26	151	1904569.00	-0.79	113.92
900.18	151	1904569.00	-0.79	113.92
1065.65	151	1904569.00	-0.79	113.92
1150.26	151	1904569.00	-0.79	113.92
...	...	...	...	...
5840.05	137	513120.00	15.87	100.99
5993.31	137	513120.00	15.87	100.99
6593.82	137	513120.00	15.87	100.99
7296.88	137	513120.00	15.87	100.99
7817.01	137	513120.00	15.87	100.99

Figure 3.2.49

Figure 3.2.48 and Figure 3.2.49 show the similar process for the dataset sea_data_renewables.

Both sea_data and sea_data_renewables datasets are exported for further analysis.

```
# Export sea_data as a new csv

sea_data.to_csv('sea_data.csv', index=False)
files.download('sea_data.csv')

# Export sea_data_renewables as a new csv

sea_data_renewables.to_csv('sea_data_renewables.csv', index=False)
files.download('sea_data_renewables.csv')
```

3.3 Data Preview

After the data is cleaned, it is crucial to conduct a data preview to inspect its structure, content, and format before proceeding for data visualization or analysis. This includes checking functions such as head(), shape, info(), unique(), and describe(). These functions help ensure that the dataset is correctly structured, contains the expected data types, and is free from unexpected anomalies before deeper analysis is performed. The preview of our data is shown below.

For the main dataset (sea_data):

Step 1: Load the dataset

```
import pandas as pd

# Making data frame from csv file
sea_data = pd.read_csv("sea_data.csv")
# Display the data
sea_data
```

	Country	Year	Access to electricity (% of population)	Access to clean fuels for cooking (% of population)	Installed renewable electricity capacity per capita (kW)	Renewable energy share in total final energy consumption (%)	Electricity from fossil fuels (TWh)	Electricity from nuclear (TWh)	Electricity from renewables (TWh)
0	Cambodia	2000	16.60	3.60	0.83	81.58	0.42	0.00	0.05
1	Cambodia	2001	14.62	4.10	0.82	80.51	0.48	0.00	0.03
2	Cambodia	2002	18.19	4.70	0.97	80.97	0.66	0.00	0.03
3	Cambodia	2003	19.30	5.30	0.96	79.92	0.67	0.00	0.04
4	Cambodia	2004	25.30	5.90	1.02	80.69	0.73	0.00	0.03
...	...	...	...	...	...	...	...	...	...
135	Thailand	2015	99.60	78.90	115.96	22.65	153.40	0.00	13.33
136	Thailand	2016	99.85	80.10	136.90	22.43	161.79	0.00	15.97
137	Thailand	2017	99.90	80.80	148.00	22.25	161.88	0.00	19.92
138	Thailand	2018	99.82	81.70	163.82	23.70	156.26	0.00	25.84
139	Thailand	2019	99.90	82.70	170.30	23.96	162.59	0.00	28.02

Figure 3.3.1

Low-carbon electricity (% of electricity)	Primary energy consumption per capita (kWh)	Energy intensity level of primary energy (MJ/$2017 PPP GDP)	CO2 emissions per capita (metric tons)	GDP growth rate (annual %)	GDP per capita	Population density (people per km2)	Land area (km2)	Latitude	Longitude
10.64	739.46	7.91	1960.00	9.99	300.61	95	181035.00	12.57	104.99
5.88	769.41	7.51	2150.00	8.15	321.15	95	181035.00	12.57	104.99
4.35	754.14	7.05	2210.00	6.58	338.99	95	181035.00	12.57	104.99
5.63	803.82	6.69	2380.00	8.51	362.34	95	181035.00	12.57	104.99
3.95	813.56	6.23	2380.00	10.34	408.51	95	181035.00	12.57	104.99
...	...	...	...	...	...	...	...	...	...
7.99	19661.77	5.07	264000.00	3.13	5840.05	137	513120.00	15.87	100.99
8.98	19981.58	5.02	261600.01	3.44	5993.31	137	513120.00	15.87	100.99
10.96	20264.65	4.79	258820.01	4.18	6593.82	137	513120.00	15.87	100.99
14.19	20834.09	4.50	257049.99	4.19	7296.88	137	513120.00	15.87	100.99
14.70	20792.38	4.52	267090.00	2.27	7817.01	137	513120.00	15.87	100.99

Figure 3.3.2

Step 2: Preview of first 10 rows

```
# Preview of first 10 rows
sea_data.head(10)
```

	Country	Year	Access to electricity (% of population)	Access to clean fuels for cooking (% of population)	Installed renewable electricity capacity per capita (kW)	Renewable energy share in total final energy consumption (%)	Electricity from fossil fuels (TWh)	Electricity from nuclear (TWh)	Electricity from renewables (TWh)
0	Cambodia	2000	16.60	3.60	0.83	81.58	0.42	0.00	0.05
1	Cambodia	2001	14.62	4.10	0.82	80.51	0.48	0.00	0.03
2	Cambodia	2002	18.19	4.70	0.97	80.97	0.66	0.00	0.03
3	Cambodia	2003	19.30	5.30	0.96	79.92	0.67	0.00	0.04
4	Cambodia	2004	25.30	5.90	1.02	80.69	0.73	0.00	0.03
5	Cambodia	2005	20.50	6.70	1.02	79.24	0.86	0.00	0.04
6	Cambodia	2006	32.41	7.60	1.35	78.01	0.98	0.00	0.13
7	Cambodia	2007	35.97	8.40	1.40	74.79	1.25	0.00	0.16
8	Cambodia	2008	26.40	9.40	1.47	74.13	1.23	0.00	0.16
9	Cambodia	2009	43.11	10.50	1.53	68.04	1.02	0.00	0.15

Figure 3.3.3

Low-carbon electricity (% of electricity)	Primary energy consumption per capita (kWh)	Energy intensity level of primary energy (MJ/$2017 PPP GDP)	CO2 emissions per capita (metric tons)	GDP growth rate (annual %)	GDP per capita	Population density (people per km2)	Land area (km2)	Latitude	Longitude
10.64	739.46	7.91	1960.00	9.99	300.61	95	181035.00	12.57	104.99
5.88	769.41	7.51	2150.00	8.15	321.15	95	181035.00	12.57	104.99
4.35	754.14	7.05	2210.00	6.58	338.99	95	181035.00	12.57	104.99
5.63	803.82	6.69	2380.00	8.51	362.34	95	181035.00	12.57	104.99
3.95	813.56	6.23	2380.00	10.34	408.51	95	181035.00	12.57	104.99
4.44	909.82	5.73	2660.00	13.25	474.11	95	181035.00	12.57	104.99
11.71	968.00	5.31	2960.00	10.77	539.75	95	181035.00	12.57	104.99
11.35	1116.15	5.03	3660.00	10.21	631.53	95	181035.00	12.57	104.99
11.51	1152.20	4.76	4060.00	6.69	745.61	95	181035.00	12.57	104.99
12.82	1399.93	5.07	4750.00	0.09	738.05	95	181035.00	12.57	104.99

Figure 3.3.4

Figure 3.3.1 and Figure 3.3.2 display the structure of the sea_data dataset, while Figure 3.3.3 and Figure 3.3.4 display the first ten rows after the dataset was cleaned and filtered. The dataset was filtered to include only countries located in Southeast Asia. To maintain the overall accuracy and integrity of the dataset, columns with the highest proportion of missing values were removed. Data from the year 2020 was also removed from the dataset as the data collected for this year was limited and it might affect the accuracy of the analysis. Null values in the column "Electricity from Nuclear (TWh)" were all filled with 0, as not all countries used nuclear energy to generate electricity.

Step 3: Check the dimensions of the dataset

```
# Check the shape of the dataset
sea_data.shape
```

```
(140, 19)
```

Figure 3.3.5

Figure 3.3.5 shows the number of rows and columns in the cleaned main dataset, which is 140 rows and 19 columns.

Step 4: Preview dataset structure and data types

```
# Display dataset structure, non-null counts, and column data types
sea_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 140 entries, 0 to 139
Data columns (total 19 columns):
 #   Column                                                 Non-Null Count  Dtype
---  ------                                                 --------------  -----
 0   Country                                                140 non-null    object
 1   Year                                                   140 non-null    int64
 2   Access to electricity (% of population)                140 non-null    float64
 3   Access to clean fuels for cooking (% of population)    140 non-null    float64
 4   Installed renewable electricity capacity per capita (kW)  140 non-null    float64
 5   Renewable energy share in total final energy consumption (%)  140 non-null    float64
 6   Electricity from fossil fuels (TWh)                    140 non-null    float64
 7   Electricity from nuclear (TWh)                         140 non-null    float64
 8   Electricity from renewables (TWh)                      140 non-null    float64
 9   Low-carbon electricity (% of electricity)              140 non-null    float64
 10  Primary energy consumption per capita (kWh)            140 non-null    float64
 11  Energy intensity level of primary energy (MJ/$2017 PPP GDP)  140 non-null    float64
 12  CO2 emissions per capita (metric tons)                 140 non-null    float64
 13  GDP growth rate (annual %)                             140 non-null    float64
 14  GDP per capita                                         140 non-null    float64
 15  Population density (people per km2)                    140 non-null    int64
 16  Land area (km2)                                        140 non-null    float64
 17  Latitude                                               140 non-null    float64
 18  Longitude                                              140 non-null    float64
dtypes: float64(16), int64(2), object(1)
memory usage: 20.9+ KB
```

Figure 3.3.6

The figure above shows the amount of non-null data in each column with their respective data types.

Step 5: Preview of unique values in each column

```
# Display unique values for each column
for col in sea_data.columns:
    print(f"Unique values in '{col}':")
    print(sea_data[col].unique())
    print("-" * 40)
```

```
Unique values in 'Country':
['Cambodia' 'Indonesia' 'Malaysia' 'Myanmar' 'Philippines' 'Singapore'
 'Thailand']
----------------------------------------
Unique values in 'Year':
[2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013
 2014 2015 2016 2017 2018 2019]
----------------------------------------
Unique values in 'Access to electricity (% of population)':
[ 16.6   14.62  18.19  19.3   25.3   20.5   32.41  35.97  26.4   43.11
  31.1   50.31  53.93  57.57  56.1   64.91  69.54  89.07  81.08  84.
  86.3   86.26  87.6   87.94  89.01  85.04  90.62  91.1   92.73  93.55
  94.15  94.83  96.    96.46  97.01  97.54  97.62  98.14  98.51  98.85
  99.1   99.13  99.14  99.15  99.16  99.17  99.18  99.2   99.3   99.37
  99.47  99.8   99.6   99.65  99.78  99.9   99.95  99.99 100.    41.83
  42.94  47.    45.14  46.23  47.32  48.42  49.51  50.62  51.73  48.8
  54.01  55.17  56.35  52.    60.5   55.6   69.81  66.23  68.36  74.66
  75.72  76.78  76.6   79.92  80.18  82.01  83.3   84.3   85.4   87.2
  87.08  87.5   89.72  89.08  92.07  93.    94.45  95.65  82.1   90.45
  91.08  91.7   92.32  92.94  94.19  98.96  99.7   99.35  99.11  99.44
  99.46  99.85  99.82]
----------------------------------------
Unique values in 'Access to clean fuels for cooking (% of population)':
[  3.6    4.1    4.7    5.3    5.9    6.7    7.6    8.4    9.4   10.5
  11.7   13.1   14.5   16.2   18.    20.3   22.7   25.5   28.9   32.6
   6.1    6.9    8.     9.3   11.3   14.1   17.9   22.5   28.5   34.7
  41.3   47.3   52.9   58.    62.8   67.1   71.2   75.1   78.7   81.85
  98.    98.1   98.2   97.7   97.6   97.4   97.2   97.1   96.8   96.6
  96.5   96.1   95.8    2.9    3.1    3.3    4.6    5.4    6.2    7.05
   9.8   11.5   13.4   15.1   17.5   19.6   21.8   24.    26.4   37.9
  39.15  39.6   40.1   40.9   41.4   40.8   40.7   40.6   42.    42.2
  43.3   44.1   45.4   46.5  100.    56.9   58.6   61.1   63.    64.6
  66.4   68.2   69.6   71.1   72.6   73.6   74.7   75.9   77.1   77.9
  78.9   80.1   80.8   81.7   82.7 ]
----------------------------------------
Unique values in 'Installed renewable electricity capacity per capita (kW)':
[  0.83   0.82   0.97   0.96   1.02   1.35   1.4    1.47   1.53   1.51
  14.97  17.85  47.59  63.67  62.79  62.07  66.17  86.74  91.27  25.89
  21.71  21.61  21.48  21.4   23.61  24.78  24.93  24.94  28.08  28.35
  29.16  30.14  32.81  32.99  33.18  34.65  35.74  36.64  38.06 119.04
 115.99 115.13 113.41 111.19 107.28 106.98 105.86 104.09 102.42  99.15
 135.95 145.91 192.63 212.85 249.45 259.33 235.74 239.17 251.85   7.36
   7.4    7.89   7.86   9.38  15.04  15.24  15.8   16.31  29.21  44.59
  50.71  52.92  55.26  58.21  62.02  63.07  63.69  63.47  63.87  48.53
  50.08  54.3   53.26  52.65  51.72  50.86  50.06  49.67  51.19  50.41
  49.88  49.12  52.54  55.08  60.51  61.08  61.82  62.55  37.95  37.5
  37.25  37.03  36.6   35.84  34.73  33.39  32.07  26.16  25.61  25.27
  25.37  31.58  37.77  44.91  54.31  57.36  65.66  84.47  56.77  56.24
  55.92  55.6   57.08  57.07  61.55  66.31  69.67  70.66  71.57  74.96
  84.    98.33 108.21 115.96 136.9  148.   163.82 170.3 ]
```

Figure 3.3.7

```
Unique values in 'Renewable energy share in total final energy consumption (%)':
[81.58 80.51 80.97 79.92 80.69 79.24 78.01 74.79 74.13 68.04 64.82 63.99
 64.36 64.97 63.7  60.63 58.01 56.25 56.38 53.36 45.62 44.33 44.66 43.
 41.46 41.57 40.07 40.01 41.12 38.18 34.81 30.92 28.78 29.7  28.32 26.69
 28.24 26.15 21.05 19.09  4.44  4.26  3.57  3.33  3.22  3.01  3.13  2.85
  3.04  2.27  1.96  2.16  2.52  2.75  3.03  3.41  4.43  5.22  5.31  5.11
 80.17 82.48 83.12 79.86 80.47 79.82 81.46 80.7  85.71 85.48 84.93 84.08
 78.46 75.96 72.43 70.43 69.25 59.92 59.55 57.85 33.42 32.12 31.42 30.67
 29.96 31.17 33.24 32.57 33.91 33.98 32.59 33.69 34.06 33.05 32.33 30.75
 28.45 27.68 27.51 26.73  0.33  0.6   0.58  0.53  0.54  0.52  0.5   0.48
  0.47  0.51  0.56  0.59  0.65  0.68  0.67  0.73  0.84 21.82 19.89 19.91
 20.21 20.11 20.53 21.61 22.63 22.83 22.96 23.56 23.21 24.41 22.65 22.43
 22.25 23.7  23.96]
----------------------------------------
Unique values in 'Electricity from fossil fuels (TWh)':
[  0.42   0.48   0.66   0.67   0.73   0.86   0.98   1.25   1.23   1.02
   0.81   0.51   0.47   1.03   2.1    2.62   3.7    2.96   3.94  78.43
  83.96  92.03  97.57 103.8  110.22 116.8  124.1  129.55 136.05 142.88
 161.41 177.83 189.66 203.11 209.71 217.97 222.64 235.41 247.39  62.91
  69.71  74.78  77.25  89.3   91.79  96.31 101.52 105.01 107.66 117.16
 117.43 123.23 127.86 133.16 135.17 135.56 132.7  142.27 149.7    2.8
   2.45   2.55   2.86   2.79   2.66   2.47   2.39   2.19   1.47   2.02
   2.73   3.18   4.99   6.12   7.53   9.18   9.78  12.78  24.29  27.72
  29.29  33.03  34.83  35.94  34.12  38.29  37.75  39.1   46.55  49.04
  51.82  55.06  57.07  60.82  68.44  71.03  76.05  83.3   29.    32.05
  33.53  34.21  35.56  36.84  38.03  39.64  40.27  40.47  43.05  43.66
  44.51  45.5   46.68  47.45  48.58  49.16  49.87  50.96  83.15  88.97
  93.51 100.61 109.46 115.58 119.41 122.12 127.43 128.09 141.72 135.31
 143.73 148.29 149.26 153.4  161.79 161.88 156.26 162.59]
----------------------------------------
Unique values in 'Electricity from nuclear (TWh)':
[0.]
----------------------------------------
Unique values in 'Electricity from renewables (TWh)':
[5.000e-02 3.000e-02 4.000e-02 1.300e-01 1.600e-01 1.500e-01 1.200e-01
 1.800e-01 8.600e-01 1.260e+00 1.960e+00 2.160e+00 2.810e+00 3.070e+00
 5.000e+00 4.510e+00 1.960e+01 2.219e+01 2.100e+01 1.982e+01 2.097e+01
 2.266e+01 2.118e+01 2.429e+01 2.634e+01 2.679e+01 3.463e+01 3.046e+01
 3.111e+01 3.550e+01 3.441e+01 3.356e+01 3.958e+01 4.317e+01 4.838e+01
 4.804e+01 7.560e+00 6.640e+00 5.810e+00 5.510e+00 6.150e+00 6.600e+00
 6.940e+00 6.720e+00 8.610e+00 8.280e+00 7.620e+00 9.600e+00 1.076e+01
 1.304e+01 1.431e+01 1.494e+01 2.109e+01 2.794e+01 2.831e+01 2.876e+01
 2.110e+00 2.050e+00 2.450e+00 2.520e+00 2.600e+00 3.160e+00 3.490e+00
 3.810e+00 4.240e+00 5.340e+00 5.240e+00 7.630e+00 7.940e+00 8.990e+00
 1.002e+01 1.127e+01 1.261e+01 1.095e+01 1.937e+01 1.750e+01 1.722e+01
 1.764e+01 1.882e+01 1.825e+01 2.039e+01 1.880e+01 2.056e+01 2.026e+01
 1.815e+01 2.009e+01 2.106e+01 2.018e+01 2.159e+01 2.238e+01 2.335e+01
 2.373e+01 2.278e+01 0.000e+00 1.000e-02 2.800e-01 4.100e-01 5.300e-01
 5.500e-01 6.700e-01 7.800e-01 6.380e+00 6.760e+00 8.070e+00 8.360e+00
 7.420e+00 9.820e+00 1.020e+01 8.950e+00 9.090e+00 8.580e+00 1.183e+01
 1.342e+01 1.233e+01 1.368e+01 1.333e+01 1.597e+01 1.992e+01 2.584e+01
 2.802e+01]
```

Figure 3.3.8

```
Unique values in 'Low-carbon electricity (% of electricity)':
[1.064e+01 5.880e+00 4.350e+00 5.630e+00 3.950e+00 4.440e+00 1.171e+01
 1.135e+01 1.151e+01 1.282e+01 1.290e+01 1.818e+01 6.277e+01 7.283e+01
 6.555e+01 5.070e+01 5.175e+01 4.535e+01 6.281e+01 5.337e+01 1.999e+01
 2.090e+01 1.858e+01 1.688e+01 1.681e+01 1.705e+01 1.535e+01 1.637e+01
 1.690e+01 1.645e+01 1.951e+01 1.588e+01 1.489e+01 1.577e+01 1.449e+01
 1.380e+01 1.537e+01 1.624e+01 1.626e+01 1.073e+01 8.700e+00 7.210e+00
 6.660e+00 6.440e+00 6.710e+00 6.720e+00 6.210e+00 7.580e+00 7.140e+00
 6.110e+00 7.560e+00 8.030e+00 9.250e+00 9.700e+00 9.950e+00 1.346e+01
 1.739e+01 1.660e+01 1.612e+01 4.297e+01 4.556e+01 4.900e+01 4.684e+01
 4.824e+01 5.430e+01 5.856e+01 6.145e+01 6.594e+01 7.841e+01 7.139e+01
 7.907e+01 7.441e+01 7.387e+01 6.431e+01 6.107e+01 5.709e+01 5.511e+01
 5.632e+01 4.614e+01 4.437e+01 3.870e+01 3.702e+01 3.481e+01 3.508e+01
 3.368e+01 3.741e+01 3.293e+01 3.526e+01 3.413e+01 2.805e+01 2.906e+01
 2.890e+01 2.682e+01 2.612e+01 2.620e+01 2.464e+01 2.474e+01 2.378e+01
 2.147e+01 0.000e+00 2.000e-02 3.500e-01 6.000e-01 8.600e-01 1.080e+00
 1.110e+00 1.330e+00 1.510e+00 7.130e+00 7.060e+00 7.940e+00 7.670e+00
 6.520e+00 6.030e+00 7.600e+00 7.710e+00 6.560e+00 6.630e+00 5.710e+00
 8.040e+00 8.540e+00 7.680e+00 8.400e+00 7.990e+00 8.980e+00 1.096e+01
 1.419e+01 1.470e+01]
----------------------------------------
Unique values in 'Primary energy consumption per capita (kWh)':
[   739.46    769.41    754.14    803.82    813.56    909.82    968.
   1116.15   1152.2    1399.93   1591.32   1623.92   1783.72   1866.12
   2031.75   2299.25   2828.74   2950.3    3674.93   4069.46   5435.44
   5740.43   5815.39   6224.94   6087.68   6186.44   6211.29   6579.23
   6569.1    6662.82   7121.82   7429.86   7574.46   7225.18   7235.56
   7266.89   7212.07   7388.52   8026.89   8474.48  26639.93  26582.33
  27303.57  28867.15  30459.57  32213.14  31831.44  33221.49  33712.13
  32003.47  32487.75  32524.59  35010.12  36062.54  35862.22  35878.01
  37199.58  37162.89  37333.03  37823.5    1055.13    963.84   1107.38
   1236.49   1285.36   1642.79   1538.98   1648.79   1504.84   1377.9
   1386.88   1505.2    1779.6    2042.98   2368.49   2192.52   2501.6
   3203.9    3270.8    3197.22   3945.09   3776.74   3717.37   3758.51
   3771.8    3713.32   3471.66   3535.24   3587.58   3563.59   3601.58
   3608.63   3668.37   3865.76   4002.74   4314.06   4629.12   4989.89
   5040.64   5104.39 109289.08 121494.56 115456.87 107484.38 119262.51
 123551.64 131721.67 136257.9  140055.12 145354.05 149763.22 151959.9
 149476.58 148317.19 148692.52 155468.98 162248.22 164801.44 163526.77
 158657.89  11764.55  12293.46  13208.06  14080.67  15161.76  15521.24
  15733.9   16149.87  16027.42  16344.1   17299.5   17644.1   18817.51
  18809.14  19310.58  19661.77  19981.58  20264.65  20834.09  20792.38]
----------------------------------------
Unique values in 'Energy intensity level of primary energy (MJ/$2017 PPP GDP)':
[ 7.91  7.51  7.05  6.69  6.23  5.73  5.31  5.03  4.76  5.07  5.05  4.81
  4.68  4.47  4.5   4.58  4.74  4.59  4.61  5.42  5.34  5.3   5.14  4.94
  4.79  4.48  4.3   4.34  4.21  4.04  3.79  3.47  3.44  3.26  3.19  3.2
  3.16  5.48  5.67  5.56  5.63  5.68  5.85  5.57  5.59  5.65  5.24  5.19
  4.99  5.38  5.18  4.72  4.69  4.28  4.25 10.52  9.23  8.52  8.12  7.52
  6.63  5.93  5.49  4.1   3.67  3.53  3.59  3.48  3.41  3.45  3.72  3.52
  3.58  4.36  4.26  4.06  3.78  3.7   3.49  3.33  3.35  3.22  3.14  3.06
  2.99  2.89  2.9   2.96  2.93  2.81  2.68  3.46  3.99  3.81  4.41  4.82
  3.07  3.03  2.56  3.05  2.48  2.05  2.31  2.39  2.52  2.66  2.8   2.51
  2.57  4.93  4.9   5.1   5.15  4.95  5.    5.28  5.2   5.02  4.52]
```

Figure 3.3.9

```
Unique values in 'CO2 emissions per capita (metric tons)':
[  1960.      2150.      2210.      2380.      2660.      2960.      3660.
   4060.      4750.      5140.      5360.      5670.      5740.      6950.
   8430.     11070.     12690.     14580.     16180.    280650.    302060.
 305640.01 333890.01 341239.99 342149.99 364470.    379959.99 376140.01
 391079.99 415519.99 475310.    481510.01 447940.    483910.    488549.99
 482510.01 517320.01 576989.99 619840.03 124360.    129460.    136380.
 144490.01 158270.    167420.    174100.01 189940.    202910.    181929.99
 200220.    202740.01 205810.    223660.    236649.99 236539.99 235960.01
 227620.    244410.    253270.      9440.      8350.      8300.     10490.
  10150.     10720.      9970.     10390.      7900.      7490.      8130.
   8690.     11960.     13600.     17040.     19040.     21860.     32500.
  32940.     36720.     72100.     70480.     71570.     73630.     75140.
  76670.     69470.     73220.     75150.     76300.     81930.     82630.
  86180.     95500.    101820.    111010.    121960.    133500.    138570.01
 145420.     42120.     42030.     41290.     38160.     39620.     36900.
  37170.     38290.     38420.     38830.     42410.     44770.     43690.
  43910.     44400.     45430.     46100.     49140.     47400.     47380.
 164490.    173160.    184240.01 191929.99 210190.    217770.    219880.
 224590.    227580.    220259.99 234380.    233600.01 250679.99 260700.01
 256799.99 264000.    261600.01 258820.01 257049.99 267090.  ]
----------------------------------------
Unique values in 'GDP growth rate (annual %)':
[ 9.99  8.15  6.58  8.51 10.34 13.25 10.77 10.21  6.69  0.09  5.96  7.07
  7.31  7.36  7.14  6.97  6.93  7.    7.47  7.05  4.92  3.64  4.5   4.78
  5.03  5.69  5.5   6.35  6.01  4.63  6.22  6.17  6.03  5.56  5.01  4.88
  5.07  5.17  5.02  8.86  0.52  5.39  5.79  6.78  5.33  5.58  6.3   4.83
 -1.51  7.42  5.29  5.47  4.69  5.09  4.45  5.81  4.84  4.44 12.42 12.47
 11.7  12.99 13.7  13.57 13.31 12.5  11.07 10.41 10.07  7.52  6.49  7.9
  8.2   3.28 10.51  5.75  6.4   6.75  4.38  3.05  3.72  6.57  4.94  5.32
  6.52  4.34  1.45  7.33  3.86  6.9   7.15  6.34  6.12  9.04 -1.07  3.91
  4.54  9.82  9.01  9.02  1.87  0.12 14.53  4.46  3.94  2.99  3.33  4.52
  3.5   1.35  3.44  6.15  7.19  6.29  4.19  4.97  5.44  1.73 -0.69  7.51
  0.84  7.24  2.69  0.98  3.13  4.18  2.27]
----------------------------------------
Unique values in 'GDP per capita':
[  300.61   321.15   338.99   362.34   408.51   474.11   539.75   631.53
   745.61   738.05   785.5    882.28   950.88  1013.42  1093.5   1162.9
  1269.59  1385.26  1512.13  1643.12   780.19   748.26   900.18  1065.65
  1150.26  1263.29  1589.8   1860.    2166.85  2261.25  3122.36  3643.05
  3694.36  3623.93  3491.64  3331.7   3562.82  3837.58  3893.86  4135.2
  4043.66  3913.43  4165.73  4461.85  4952.21  5587.02  6209.13  7243.46
  8474.59  7292.49  9040.57 10399.37 10817.43 10970.1  11319.06  9955.24
  9817.79 10259.3  11380.08 11432.82   146.6    131.72   128.1    161.06
   193.37   216.31   240.62   314.2    460.91   586.17   746.95  1061.34
  1134.3   1168.17  1210.1   1196.74  1136.61  1151.11  1250.17  1271.11
  1072.8    990.57  1036.16  1048.01  1121.49  1244.35  1452.44  1744.64
  1998.03  1904.2   2217.47  2450.74  2694.31  2871.43  2959.65  3001.04
  3073.65  3123.25  3252.11  3485.34 23852.33 21700.02 22159.69 23730.15
 27608.54 29961.26 33769.15 39432.94 40007.47 38927.21 47236.96 53890.43
 55546.49 56967.43 57562.53 55646.62 56848.18 61176.46 66679.05 65640.71
  2007.74  1893.26  2096.19  2359.12  2660.13  2894.06  3369.54  3973.02
  4379.66  4213.01  5076.34  5492.12  5860.58  6168.26  5951.88  5840.05
  5993.31  6593.82  7296.88  7817.01]
```

Figure 3.3.10

```
----------------------------------------
Unique values in 'Population density (people per km2)':
[  95  151   99   83  368 8358  137]
----------------------------------------
Unique values in 'Land area (km2)':
[1.810350e+05 1.904569e+06 3.298470e+05 6.765780e+05 3.000000e+05
 7.160000e+02 5.131200e+05]
----------------------------------------
Unique values in 'Latitude':
[12.565679 -0.789275  4.210484 21.916221 12.879721  1.352083 15.870032]
----------------------------------------
Unique values in 'Longitude':
[104.990963 113.921327 101.975766  95.955974 121.774017 103.819836
 100.992541]
```

Figure 3.3.11

Figure 3.3.7, Figure 3.3.8, Figure 3.3.9, Figure 3.3.10, and Figure 3.3.11 show all unique values contained in the cleaned main dataset.

Step 6: Preview of descriptive statistics

```
# Review the descriptive statistics of the dataset
sea_data.describe()
```

	Year	Access to electricity (% of population)	Access to clean fuels for cooking (% of population)	Installed renewable electricity capacity per capita (kW)	Renewable energy share in total final energy consumption (%)	Electricity from fossil fuels (TWh)	Electricity from nuclear (TWh)	Electricity from renewables (TWh)	Low-carbon electricity (% of electricity)
count	140.00	140.00	140.00	140.00	140.00	140.00	140.00	140.00	140.00
mean	2009.50	81.59	53.68	61.23	34.01	69.77	0.00	11.77	22.05
std	5.79	23.71	35.98	53.75	28.12	62.35	0.00	11.26	21.36
min	2000.00	14.62	2.90	0.82	0.33	0.42	0.00	0.00	0.00
25%	2004.75	69.25	17.80	27.60	4.39	5.84	0.00	2.03	6.72
50%	2009.50	92.53	45.95	50.78	28.62	49.52	0.00	8.59	14.79
75%	2014.25	99.46	96.88	67.15	58.39	117.93	0.00	20.11	34.30
max	2019.00	100.00	100.00	259.33	85.71	247.39	0.00	48.38	79.07

Figure 3.3.12

Primary energy consumption per capita (kWh)	Energy intensity level of primary energy (MJ/$2017 PPP GDP)	CO2 emissions per capita (metric tons)	GDP growth rate (annual %)	GDP per capita	Population density (people per km2)	Land area (km2)	Latitude	Longitude
140.00	140.00	140.00	140.00	140.00	140.00	140.00	140.00	140.00
29212.93	4.50	142557.14	6.08	8966.74	1327.29	557980.71	9.71	106.20
47162.82	1.41	145160.92	3.06	15674.59	2882.02	587717.07	7.73	8.13
739.46	2.05	1960.00	-1.51	128.10	83.00	716.00	-0.79	95.96
3135.49	3.45	18540.00	4.46	1114.49	95.00	181035.00	1.35	100.99
6892.32	4.58	79300.00	5.88	2926.86	137.00	329847.00	12.57	103.82
30802.54	5.18	229115.00	7.26	6756.23	368.00	676578.00	15.87	113.92
164801.44	10.52	619840.03	14.53	66679.05	8358.00	1904569.00	21.92	121.77

Figure 3.3.13

Figure 3.3.12 and Figure 3.3.13 display the descriptive statistics of the cleaned main dataset.

For the filtered dataset (sea_data_renewables):

Step 1: Load the dataset

```
# Making data frame from csv file
sea_data_renewables = pd.read_csv("sea_data_renewables.csv")
# Display the data
sea_data_renewables
```

	Country	Year	Access to electricity (% of population)	Access to clean fuels for cooking (% of population)	Installed renewable electricity capacity per capita (kW)	Renewable energy share in total final energy consumption (%)	Electricity from fossil fuels (TWh)	Electricity from nuclear (TWh)	Electricity from renewables (TWh)	Low-carbon electricity (% of electricity)
0	Indonesia	2000	86.30	6.10	25.89	45.62	78.43	0.00	19.60	19.99
1	Indonesia	2001	86.26	6.90	21.71	44.33	83.96	0.00	22.19	20.90
2	Indonesia	2002	87.60	8.00	21.61	44.66	92.03	0.00	21.00	18.58
3	Indonesia	2003	87.94	9.30	21.48	43.00	97.57	0.00	19.82	16.88
4	Indonesia	2004	89.01	11.30	21.40	41.46	103.80	0.00	20.97	16.81
...	...	...	...	...	...	...	...	...	...	...
95	Thailand	2015	99.60	78.90	115.96	22.65	153.40	0.00	13.33	7.99
96	Thailand	2016	99.85	80.10	136.90	22.43	161.79	0.00	15.97	8.98
97	Thailand	2017	99.90	80.80	148.00	22.25	161.88	0.00	19.92	10.96
98	Thailand	2018	99.82	81.70	163.82	23.70	156.26	0.00	25.84	14.19
99	Thailand	2019	99.90	82.70	170.30	23.96	162.59	0.00	28.02	14.70

Figure 3.3.14

Primary energy consumption per capita (kWh)	Energy intensity level of primary energy (MJ/$2017 PPP GDP)	CO2 emissions per capita (metric tons)	Renewables (% equivalent primary energy)	GDP growth rate (annual %)	GDP per capita	Population density (people per km2)	Land area (km2)	Latitude	Longitude
5435.44	5.42	280650.00	3.73	4.92	780.19	151	1904569.00	-0.79	113.92
5740.43	5.34	302060.00	4.11	3.64	748.26	151	1904569.00	-0.79	113.92
5815.39	5.30	305640.01	3.68	4.50	900.18	151	1904569.00	-0.79	113.92
6224.94	5.07	333890.01	3.22	4.78	1065.65	151	1904569.00	-0.79	113.92
6087.68	5.14	341239.99	3.43	5.03	1150.26	151	1904569.00	-0.79	113.92
...	...	...	...	...	...	...	...	...	...
19661.77	5.07	264000.00	4.19	3.13	5840.05	137	513120.00	15.87	100.99
19981.58	5.02	261600.01	4.60	3.44	5993.31	137	513120.00	15.87	100.99
20264.65	4.79	258820.01	5.31	4.18	6593.82	137	513120.00	15.87	100.99
20834.09	4.50	257049.99	6.38	4.19	7296.88	137	513120.00	15.87	100.99
20792.38	4.52	267090.00	6.96	2.27	7817.01	137	513120.00	15.87	100.99

Figure 3.3.15

Step 2: Preview of first 10 rows

```
# Preview of first 10 rows
sea_data_renewables.head(10)
```

	Country	Year	Access to electricity (% of population)	Access to clean fuels for cooking (% of population)	Installed renewable electricity capacity per capita (kW)	Renewable energy share in total final energy consumption (%)	Electricity from fossil fuels (TWh)	Electricity from nuclear (TWh)	Electricity from renewables (TWh)	Low-carbon electricity (% of electricity)
0	Indonesia	2000	86.30	6.10	25.89	45.62	78.43	0.00	19.60	19.99
1	Indonesia	2001	86.26	6.90	21.71	44.33	83.96	0.00	22.19	20.90
2	Indonesia	2002	87.60	8.00	21.61	44.66	92.03	0.00	21.00	18.58
3	Indonesia	2003	87.94	9.30	21.48	43.00	97.57	0.00	19.82	16.88
4	Indonesia	2004	89.01	11.30	21.40	41.46	103.80	0.00	20.97	16.81
5	Indonesia	2005	85.04	14.10	23.61	41.57	110.22	0.00	22.66	17.05
6	Indonesia	2006	90.62	17.90	24.78	40.07	116.80	0.00	21.18	15.35
7	Indonesia	2007	91.10	22.50	24.93	40.01	124.10	0.00	24.29	16.37
8	Indonesia	2008	92.73	28.50	24.94	41.12	129.55	0.00	26.34	16.90
9	Indonesia	2009	93.55	34.70	28.08	38.18	136.05	0.00	26.79	16.45

Figure 3.3.16

Primary energy consumption per capita (kWh)	Energy intensity level of primary energy (MJ/$2017 PPP GDP)	CO2 emissions per capita (metric tons)	Renewables (% equivalent primary energy)	GDP growth rate (annual %)	GDP per capita	Population density (people per km2)	Land area (km2)	Latitude	Longitude
5435.44	5.42	280650.00	3.73	4.92	780.19	151	1904569.00	-0.79	113.92
5740.43	5.34	302060.00	4.11	3.64	748.26	151	1904569.00	-0.79	113.92
5815.39	5.30	305640.01	3.68	4.50	900.18	151	1904569.00	-0.79	113.92
6224.94	5.07	333890.01	3.22	4.78	1065.65	151	1904569.00	-0.79	113.92
6087.68	5.14	341239.99	3.43	5.03	1150.26	151	1904569.00	-0.79	113.92
6186.44	4.94	342149.99	3.53	5.69	1263.29	151	1904569.00	-0.79	113.92
6211.29	4.79	364470.00	3.22	5.50	1589.80	151	1904569.00	-0.79	113.92
6579.23	4.48	379959.99	3.37	6.35	1860.00	151	1904569.00	-0.79	113.92
6569.10	4.30	376140.01	3.59	6.01	2166.85	151	1904569.00	-0.79	113.92
6662.82	4.34	391079.99	3.66	4.63	2261.25	151	1904569.00	-0.79	113.92

Figure 3.3.17

Figure 3.3.14 and Figure 3.3.15 show the data frame made from the sea_data_renewables and its structure. Figure 3.3.16 and Figure 3.3.17 displays the first 10 rows of the dataset.

Step 3: Check the dimensions of the dataset

```
# Check the shape of the dataset
sea_data_renewables.shape
```

```
(100, 20)
```

Figure 3.3.18

Figure 3.3.18 shows the total number of rows and columns in the dataset, which are 100 rows and 20 columns.

Step 4: Preview dataset structure and data types

```
# Display dataset structure, non-null counts, and column data types
sea_data_renewables.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 100 entries, 0 to 99
Data columns (total 20 columns):
 #   Column                                                   Non-Null Count  Dtype
---  ------                                                   --------------  -----
 0   Country                                                  100 non-null    object
 1   Year                                                     100 non-null    int64
 2   Access to electricity (% of population)                  100 non-null    float64
 3   Access to clean fuels for cooking (% of population)      100 non-null    float64
 4   Installed renewable electricity capacity per capita (kW) 100 non-null    float64
 5   Renewable energy share in total final energy consumption (%) 100 non-null    float64
 6   Electricity from fossil fuels (TWh)                      100 non-null    float64
 7   Electricity from nuclear (TWh)                           100 non-null    float64
 8   Electricity from renewables (TWh)                        100 non-null    float64
 9   Low-carbon electricity (% of electricity)                100 non-null    float64
 10  Primary energy consumption per capita (kWh)              100 non-null    float64
 11  Energy intensity level of primary energy (MJ/$2017 PPP GDP) 100 non-null    float64
 12  CO2 emissions per capita (metric tons)                   100 non-null    float64
 13  Renewables (% equivalent primary energy)                 100 non-null    float64
 14  GDP growth rate (annual %)                               100 non-null    float64
 15  GDP per capita                                           100 non-null    float64
 16  Population density (people per km2)                      100 non-null    int64
 17  Land area (km2)                                          100 non-null    float64
 18  Latitude                                                 100 non-null    float64
 19  Longitude                                                100 non-null    float64
dtypes: float64(17), int64(2), object(1)
memory usage: 15.8+ KB
```

Figure 3.3.19

The figure above shows the amount of non-null data in each column with their respective data types.

Step 5: Preview of unique values in each column

```
# Display unique values for each column
for col in sea_data_renewables.columns:
    print(f"Unique values in '{col}':")
    print(sea_data_renewables[col].unique())
    print("-" * 40)
```

```
Unique values in 'Country':
['Indonesia' 'Malaysia' 'Philippines' 'Singapore' 'Thailand']
----------------------------------------
Unique values in 'Year':
[2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013
 2014 2015 2016 2017 2018 2019]
----------------------------------------
Unique values in 'Access to electricity (% of population)':
[ 86.3   86.26  87.6   87.94  89.01  85.04  90.62  91.1   92.73  93.55
  94.15  94.83  96.    96.46  97.01  97.54  97.62  98.14  98.51  98.85
  99.1   99.13  99.14  99.15  99.16  99.17  99.18  99.2   99.3   99.37
  99.47  99.8   99.6   99.65  99.78  99.9   99.95  99.99 100.    74.66
  75.72  76.78  76.6   79.92  80.18  82.01  83.3   84.3   85.4   87.2
  87.08  87.5   89.72  89.08  92.07  93.    94.45  95.65  82.1   90.45
  91.08  91.7   92.32  92.94  94.19  98.96  99.7   99.35  99.11  99.44
  99.46  99.85  99.82]
```

Figure 3.3.20

```
Unique values in 'Access to clean fuels for cooking (% of population)':
[  6.1    6.9    8.     9.3   11.3   14.1   17.9   22.5   28.5   34.7
  41.3   47.3   52.9   58.    62.8   67.1   71.2   75.1   78.7   81.85
  98.    98.1   98.2   97.7   97.6   97.4   97.2   97.1   96.8   96.6
  96.5   96.1   95.8   37.9   39.15  39.6   40.1   40.9   41.4   40.8
  40.7   40.6   42.    42.2   43.3   44.1   45.4   46.5  100.    56.9
  58.6   61.1   63.    64.6   66.4   68.2   69.6   71.1   72.6   73.6
  74.7   75.9   77.1   77.9   78.9   80.1   80.8   81.7   82.7 ]
----------------------------------------
Unique values in 'Installed renewable electricity capacity per capita (kW)':
[ 25.89  21.71  21.61  21.48  21.4   23.61  24.78  24.93  24.94  28.08
  28.35  29.16  30.14  32.81  32.99  33.18  34.65  35.74  36.64  38.06
 119.04 115.99 115.13 113.41 111.19 107.28 106.98 105.86 104.09 102.42
  99.15 135.95 145.91 192.63 212.85 249.45 259.33 235.74 239.17 251.85
  48.53  50.08  54.3   53.26  52.65  51.72  50.86  50.06  49.67  51.19
  50.41  49.88  49.12  52.54  55.08  60.51  61.08  61.82  62.55  37.95
  37.5   37.25  37.03  36.6   35.84  34.73  33.39  32.07  26.16  25.61
  25.27  25.37  31.58  37.77  44.91  54.31  57.36  65.66  84.47  56.77
  56.24  55.92  55.6   57.08  57.07  61.55  66.31  69.67  70.66  71.57
  74.96  84.    98.33 108.21 115.96 136.9  148.   163.82 170.3 ]
----------------------------------------
Unique values in 'Renewable energy share in total final energy consumption (%)':
[45.62 44.33 44.66 43.   41.46 41.57 40.07 40.01 41.12 38.18 34.81 30.92
 28.78 29.7  28.32 26.69 28.24 26.15 21.05 19.09  4.44  4.26  3.57  3.33
  3.22  3.01  3.13  2.85  3.04  2.27  1.96  2.16  2.52  2.75  3.03  3.41
  4.43  5.22  5.31  5.11 33.42 32.12 31.42 30.67 29.96 31.17 33.24 32.57
 33.91 33.98 32.59 33.69 34.06 33.05 32.33 30.75 28.45 27.68 27.51 26.73
  0.33  0.6   0.58  0.53  0.54  0.52  0.5   0.48  0.47  0.51  0.56  0.59
  0.65  0.68  0.67  0.73  0.84 21.82 19.89 19.91 20.21 20.11 20.53 21.61
 22.63 22.83 22.96 23.56 23.21 24.41 22.65 22.43 22.25 23.7  23.96]
----------------------------------------
Unique values in 'Electricity from fossil fuels (TWh)':
[ 78.43  83.96  92.03  97.57 103.8  110.22 116.8  124.1  129.55 136.05
 142.88 161.41 177.83 189.66 203.11 209.71 217.97 222.64 235.41 247.39
  62.91  69.71  74.78  77.25  89.3   91.79  96.31 101.52 105.01 107.66
 117.16 117.43 123.23 127.86 133.16 135.17 135.56 132.7  142.27 149.7
  24.29  27.72  29.29  33.03  34.83  35.94  34.12  38.29  37.75  39.1
  46.55  49.04  51.82  55.06  57.07  60.82  68.44  71.03  76.05  83.3
  29.    32.05  33.53  34.21  35.56  36.84  38.03  39.64  40.27  40.47
  43.05  43.66  44.51  45.5   46.68  47.45  48.58  49.16  49.87  50.96
  83.15  88.97  93.51 100.61 109.46 115.58 119.41 122.12 127.43 128.09
 141.72 135.31 143.73 148.29 149.26 153.4  161.79 161.88 156.26 162.59]
----------------------------------------
Unique values in 'Electricity from nuclear (TWh)':
[0.]
----------------------------------------
Unique values in 'Electricity from renewables (TWh)':
[1.960e+01 2.219e+01 2.100e+01 1.982e+01 2.097e+01 2.266e+01 2.118e+01
 2.429e+01 2.634e+01 2.679e+01 3.463e+01 3.046e+01 3.111e+01 3.550e+01
 3.441e+01 3.356e+01 3.958e+01 4.317e+01 4.838e+01 4.804e+01 7.560e+00
 6.640e+00 5.810e+00 5.510e+00 6.150e+00 6.600e+00 6.940e+00 6.720e+00
 8.610e+00 8.280e+00 7.620e+00 9.600e+00 1.076e+01 1.304e+01 1.431e+01
 1.494e+01 2.109e+01 2.794e+01 2.831e+01 2.876e+01 1.937e+01 1.750e+01
 1.722e+01 1.764e+01 1.882e+01 1.825e+01 2.039e+01 1.880e+01 2.056e+01
 2.026e+01 1.815e+01 2.009e+01 2.106e+01 2.018e+01 2.159e+01 2.238e+01
 2.335e+01 2.373e+01 2.278e+01 0.000e+00 1.000e-02 1.600e-01 2.800e-01
 4.100e-01 5.300e-01 5.500e-01 6.700e-01 7.800e-01 6.380e+00 6.760e+00
 8.070e+00 8.360e+00 7.630e+00 7.420e+00 9.820e+00 1.020e+01 8.950e+00
 9.090e+00 8.580e+00 1.183e+01 1.342e+01 1.233e+01 1.368e+01 1.333e+01
 1.597e+01 1.992e+01 2.584e+01 2.802e+01]
```

Figure 3.3.21

```
Unique values in 'Low-carbon electricity (% of electricity)':
[1.999e+01 2.090e+01 1.858e+01 1.688e+01 1.681e+01 1.705e+01 1.535e+01
 1.637e+01 1.690e+01 1.645e+01 1.951e+01 1.588e+01 1.489e+01 1.577e+01
 1.449e+01 1.380e+01 1.537e+01 1.624e+01 1.626e+01 1.073e+01 8.700e+00
 7.210e+00 6.660e+00 6.440e+00 6.710e+00 6.720e+00 6.210e+00 7.580e+00
 7.140e+00 6.110e+00 7.560e+00 8.030e+00 9.250e+00 9.700e+00 9.950e+00
 1.346e+01 1.739e+01 1.660e+01 1.612e+01 4.437e+01 3.870e+01 3.702e+01
 3.481e+01 3.508e+01 3.368e+01 3.741e+01 3.293e+01 3.526e+01 3.413e+01
 2.805e+01 2.906e+01 2.890e+01 2.682e+01 2.612e+01 2.620e+01 2.464e+01
 2.474e+01 2.378e+01 2.147e+01 0.000e+00 2.000e-02 3.500e-01 6.000e-01
 8.600e-01 1.080e+00 1.110e+00 1.330e+00 1.510e+00 7.130e+00 7.060e+00
 7.940e+00 7.670e+00 6.520e+00 6.030e+00 7.600e+00 7.710e+00 6.560e+00
 6.630e+00 5.710e+00 8.040e+00 8.540e+00 7.680e+00 8.400e+00 7.990e+00
 8.980e+00 1.096e+01 1.419e+01 1.470e+01]
----------------------------------------
Unique values in 'Primary energy consumption per capita (kWh)':
[  5435.44   5740.43   5815.39   6224.94   6087.68   6186.44   6211.29
   6579.23   6569.1    6662.82   7121.82   7429.86   7574.46   7225.18
   7235.56   7266.89   7212.07   7388.52   8026.89   8474.48  26639.93
  26582.33  27303.57  28867.15  30459.57  32213.14  31831.44  33221.49
  33712.13  32003.47  32487.75  32524.59  35010.12  36062.54  35862.22
  35878.01  37199.58  37162.89  37333.03  37823.5    3945.09   3776.74
   3717.37   3758.51   3771.8    3713.32   3471.66   3535.24   3587.58
   3563.59   3601.58   3608.63   3668.37   3865.76   4002.74   4314.06
   4629.12   4989.89   5040.64   5104.39 109289.08 121494.56 115456.87
 107484.38 119262.51 123551.64 131721.67 136257.9  140055.12 145354.05
 149763.22 151959.9  149476.58 148317.19 148692.52 155468.98 162248.22
 164801.44 163526.77 158657.89  11764.55  12293.46  13208.06  14080.67
  15161.76  15521.24  15733.9   16149.87  16027.42  16344.1   17299.5
  17644.1   18817.51  18809.14  19310.58  19661.77  19981.58  20264.65
  20834.09  20792.38]
----------------------------------------
Unique values in 'Energy intensity level of primary energy (MJ/$2017 PPP GDP)':
[5.42 5.34 5.3  5.07 5.14 4.94 4.79 4.48 4.3  4.34 4.21 4.04 3.79 3.47
 3.44 3.26 3.19 3.2  3.16 5.48 5.67 5.56 5.63 5.68 5.85 5.57 5.59 5.65
 5.24 5.19 4.99 5.38 5.18 4.72 4.69 4.28 4.5  4.25 4.36 4.26 4.06 3.78
 3.7  3.49 3.33 3.35 3.22 3.14 3.06 2.99 2.89 2.9  2.96 2.93 2.81 2.68
 3.46 3.99 3.81 4.41 4.82 3.07 3.03 2.56 3.05 2.48 2.05 2.31 2.39 2.52
 2.66 2.8  2.51 2.57 4.93 4.9  5.1  5.15 5.03 4.95 5.   5.05 5.28 5.2
 5.02 4.52]
----------------------------------------
Unique values in 'CO2 emissions per capita (metric tons)':
[280650.   302060.   305640.01 333890.01 341239.99 342149.99 364470.
 379959.99 376140.01 391079.99 415519.99 475310.   481510.01 447940.
 483910.   488549.99 482510.01 517320.01 576989.99 619840.03 124360.
 129460.   136380.   144490.01 158270.   167420.   174100.01 189940.
 202910.   181929.99 200220.   202740.01 205810.   223660.   236649.99
 236539.99 235960.01 227620.   244410.   253270.    72100.    70480.
  71570.    73630.    75140.    76670.    69470.    73220.    75150.
  76300.    81930.    82630.    86180.    95500.   101820.   111010.
 121960.   133500.   138570.01 145420.    42120.    42030.    41290.
  38160.    39620.    36900.    37170.    38290.    38420.    38830.
  42410.    44770.    43690.    43910.    44400.    45430.    46100.
  49140.    47400.    47380.   164490.   173160.   184240.01 191929.99
 210190.   217770.   219880.   224590.   227580.   220259.99 234380.
 233600.01 250679.99 260700.01 256799.99 264000.   261600.01 258820.01
 257049.99 267090.  ]
```

Figure 3.3.22

```
----------------------------------------
Unique values in 'Renewables (% equivalent primary energy)':
[ 3.73  4.11  3.68  3.22  3.43  3.53  3.37  3.59  3.66  4.33  3.39  3.47
  4.48  3.79  5.48  5.28  7.39  7.87  3.14  2.59  2.25  2.32  2.29  2.37
  2.14  2.62  2.33  2.86  2.91  3.32  3.56  3.65  4.85  6.29  6.25  6.18
 18.6  16.95 16.69 16.42 16.85 16.32 18.68 16.43 17.31 16.68 14.19 15.43
 15.56 13.85 13.01 12.5  11.98 11.5  11.2  10.23  0.18  0.3   0.31  0.34
  0.29  0.28  0.25  0.24  0.23  0.27  2.57  2.82  2.72  2.31  2.28  2.98
  2.94  2.95  3.05  2.81  3.46  3.77  3.86  4.26  4.19  4.6   5.31  6.38
  6.96]
----------------------------------------
Unique values in 'GDP growth rate (annual %)':
[ 4.92  3.64  4.5   4.78  5.03  5.69  5.5   6.35  6.01  4.63  6.22  6.17
  6.03  5.56  5.01  4.88  5.07  5.17  5.02  8.86  0.52  5.39  5.79  6.78
  5.33  5.58  6.3   4.83 -1.51  7.42  5.29  5.47  4.69  5.09  4.45  5.81
  4.84  4.44  4.38  3.05  3.72  6.57  4.94  5.32  6.52  4.34  1.45  7.33
  3.86  6.9   6.75  7.15  6.93  6.34  6.12  9.04 -1.07  3.91  4.54  9.82
  7.36  9.01  9.02  1.87  0.12 14.53  4.46  3.94  2.99  3.33  4.52  3.5
  1.35  3.44  6.15  7.19  6.29  4.19  4.97  5.44  1.73 -0.69  7.51  0.84
  7.24  2.69  0.98  3.13  4.18  2.27]
----------------------------------------
Unique values in 'GDP per capita':
[  780.19   748.26   900.18  1065.65  1150.26  1263.29  1589.8   1860.
  2166.85  2261.25  3122.36  3643.05  3694.36  3623.93  3491.64  3331.7
  3562.82  3837.58  3893.86  4135.2   4043.66  3913.43  4165.73  4461.85
  4952.21  5587.02  6209.13  7243.46  8474.59  7292.49  9040.57 10399.37
 10817.43 10970.1  11319.06  9955.24  9817.79 10259.3  11380.08 11432.82
  1072.8    990.57  1036.16  1048.01  1121.49  1244.35  1452.44  1744.64
  1998.03  1904.2   2217.47  2450.74  2694.31  2871.43  2959.65  3001.04
  3073.65  3123.25  3252.11  3485.34 23852.33 21700.02 22159.69 23730.15
 27608.54 29961.26 33769.15 39432.94 40007.47 38927.21 47236.96 53890.43
 55546.49 56967.43 57562.53 55646.62 56848.18 61176.46 66679.05 65640.71
  2007.74  1893.26  2096.19  2359.12  2660.13  2894.06  3369.54  3973.02
  4379.66  4213.01  5076.34  5492.12  5860.58  6168.26  5951.88  5840.05
  5993.31  6593.82  7296.88  7817.01]
----------------------------------------
Unique values in 'Population density (people per km2)':
[ 151   99  368 8358  137]
----------------------------------------
Unique values in 'Land area (km2)':
[1.904569e+06 3.298470e+05 3.000000e+05 7.160000e+02 5.131200e+05]
----------------------------------------
Unique values in 'Latitude':
[-0.789275  4.210484 12.879721  1.352083 15.870032]
----------------------------------------
Unique values in 'Longitude':
[113.921327 101.975766 121.774017 103.819836 100.992541]
```

Figure 3.3.23

Figure 3.3.20 to Figure 3.3.23 show the unique values for each column in the dataset.

Step 6: Preview of descriptive statistics

```
# Review the descriptive statistics of the dataset
sea_data_renewables.describe()
```

	Year	Access to electricity (% of population)	Access to clean fuels for cooking (% of population)	Installed renewable electricity capacity per capita (kW)	Renewable energy share in total final energy consumption (%)	Electricity from fossil fuels (TWh)	Electricity from nuclear (TWh)	Electricity from renewables (TWh)	Low-carbon electricity (% of electricity)
count	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
mean	2009.50	94.75	70.09	73.37	18.45	96.56	0.00	15.00	13.19
std	5.80	6.80	28.94	56.62	14.66	54.02	0.00	11.65	10.96
min	2000.00	74.66	6.10	21.40	0.33	24.29	0.00	0.00	0.00
25%	2004.75	90.97	41.85	35.49	2.83	46.65	0.00	6.63	6.50
50%	2009.50	99.10	75.50	53.78	22.04	91.91	0.00	14.00	9.47
75%	2014.25	99.90	98.03	99.97	30.79	133.66	0.00	21.28	17.14
max	2019.00	100.00	100.00	259.33	45.62	247.39	0.00	48.38	44.37

Figure 3.3.24

Primary energy consumption per capita (kWh)	Energy intensity level of primary energy (MJ/$2017 PPP GDP)	CO2 emissions per capita (metric tons)	Renewables (% equivalent primary energy)	GDP growth rate (annual %)	GDP per capita	Population density (people per km2)	Land area (km2)	Latitude	Longitude
100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
40188.54	4.16	195413.40	5.32	4.99	12248.79	1822.60	609650.40	6.70	108.50
51931.43	1.06	140288.99	5.29	2.33	17515.78	3285.52	671363.26	6.56	8.12
3471.66	2.05	36900.00	0.18	-1.51	748.26	99.00	716.00	-0.79	100.99
6205.08	3.16	73527.50	2.32	4.12	2427.84	137.00	300000.00	1.35	101.98
16821.80	4.35	183085.00	3.46	5.03	4150.47	151.00	329847.00	4.21	103.82
35924.14	5.08	257492.49	6.26	6.29	10503.89	368.00	513120.00	12.88	113.92
164801.44	5.85	619840.03	18.68	14.53	66679.05	8358.00	1904569.00	15.87	121.77

Figure 3.3.25

Figure 3.3.24 and Figure 3.3.25 shows the descriptive statistics for each column in the dataset.

3.4 Data Description

After the process of cleaning and filtering, the dataset is now left with only non-null values along with countries that are needed to be investigated to fulfill the objectives of this project. The table below presents the description of cleaned and filtered dataset.

Data Type	Column	Description
Quantitative	Year	Year from 2000 to 2019
	Access to electricity (% of population)	Percentage of population having access to electricity
	Access to clean fuels for cooking (% of population)	Percentage of population having access to clean fuels for cooking
	Installed renewable electricity capacity per capita (kW)	Installed Renewable energy capacity per person
	Renewable energy share in total final energy consumption (%)	Percentage of renewable energy in final energy consumption
	Electricity from fossil fuels (TWh)	Electricity generated from fossil fuels (coal, oil, gas) in terawatt-hours
	Electricity from nuclear (TWh)	Electricity generated from nuclear power in terawatt-hours
	Electricity from renewables (TWh)	Electricity generated from renewable sources (hydro, solar, wind, etc.) in terawatt-hours
	Low-carbon electricity (% of electricity)	Percentage of electricity from low-carbon sources (nuclear and renewables)
	Primary energy consumption per capita (kWh)	Energy consumption per person in kilowatt-hours
	Energy intensity level of primary energy (MJ/$2017 PPP GDP)	Energy use per unit of GDP at purchasing power parity
	CO2 emissions per capita (metric tons)	Carbon dioxide emissions per person in metric tons

	GDP growth rate (annual %)	Annual GDP growth rate based on constant local currency
	GDP per capita	Gross domestic products per person.
	Population density (people per km2)	Population density in persons per square kilometer
	Land area (km2)	Total land area in square kilometers
	Latitude	Latitude of country
	Longitude	Longitude of country
Qualitative	Country	Name of countries in Southeast Asia (Cambodia, Indonesia, Malaysia, Myanmar, Philippines, Singapore, and Thailand)

Table 3.4.1

Table 3.4.1 presents the overall description for each column under the main dataset, sea_data. Two columns were removed which are "Renewables (% equivalent primary energy)" and "Financial flows to developing countries (US $)" as these two columns contained a high percentage of null values in them which might affect the overall accuracy of the data when they are visualized. Other columns were filtered displaying only data from the countries in Southeast Asia. Null values in "Electricity from nuclear" were filled with 0 as countries from Southeast Asia do not use nuclear energy to generate electricity.

Data Type	Column	Description
Quantitative	Year	Year from 2000 to 2019
	Access to electricity (% of population)	Percentage of population having access to electricity
	Access to clean fuels for cooking (% of population)	Percentage of population having access to clean fuels for cooking
	Installed renewable electricity capacity per capita (kW)	Installed Renewable energy capacity per person

	Renewable energy share in total final energy consumption (%)	Percentage of renewable energy in final energy consumption
	Electricity from fossil fuels (TWh)	Electricity generated from fossil fuels (coal, oil, gas) in terawatt-hours
	Electricity from nuclear (TWh)	Electricity generated from nuclear power in terawatt-hours
	Electricity from renewables (TWh)	Electricity generated from renewable sources (hydro, solar, wind, etc.) in terawatt-hours
	Low-carbon electricity (% of electricity)	Percentage of electricity from low-carbon sources (nuclear and renewables)
	Primary energy consumption per capita (kWh)	Energy consumption per person in kilowatt-hours
	Energy intensity level of primary energy (MJ/$2017 PPP GDP)	Energy use per unit of GDP at purchasing power parity
	CO2 emissions per capita (metric tons)	Carbon dioxide emissions per person in metric tons
	Renewables (% equivalent primary energy)	Equivalent primary energy that is derived from renewable sources
	GDP growth rate (annual %)	Annual GDP growth rate based on constant local currency
	GDP per capita	Gross domestic products per person.
	Population density (people per km2)	Population density in persons per square kilometer
	Land area (km2)	Total land area in square kilometers
	Latitude	Latitude of country
	Longitude	Longitude of country
Qualitative	Country	Name of countries in Southeast Asia (Indonesia, Malaysia, Philippines, Singapore, and Thailand)

Table 3.4.2

Table 3.4.2 displays the overall description for each column under the filtered dataset, sea_data_renewables. For this dataset, the column “Renewables (% equivalent primary energy)” is included to investigate the proportion of renewable energy used as primary energy in countries from Southeast Asia. Only countries with values in this column (Indonesia, Malaysia, Philippines, Singapore, and Thailand) are included, while those without values (Cambodia and Myanmar) have been excluded. The column “Financial flows to developing countries (US $)” was excluded as it is insignificant in achieving the objectives of this project.

3.5 Data Description

Figure 3.5.1 below shows the flowchart of the complete data preparation process, starting from downloading and importing the dataset, followed by data cleaning, data preview, and ending with data description.

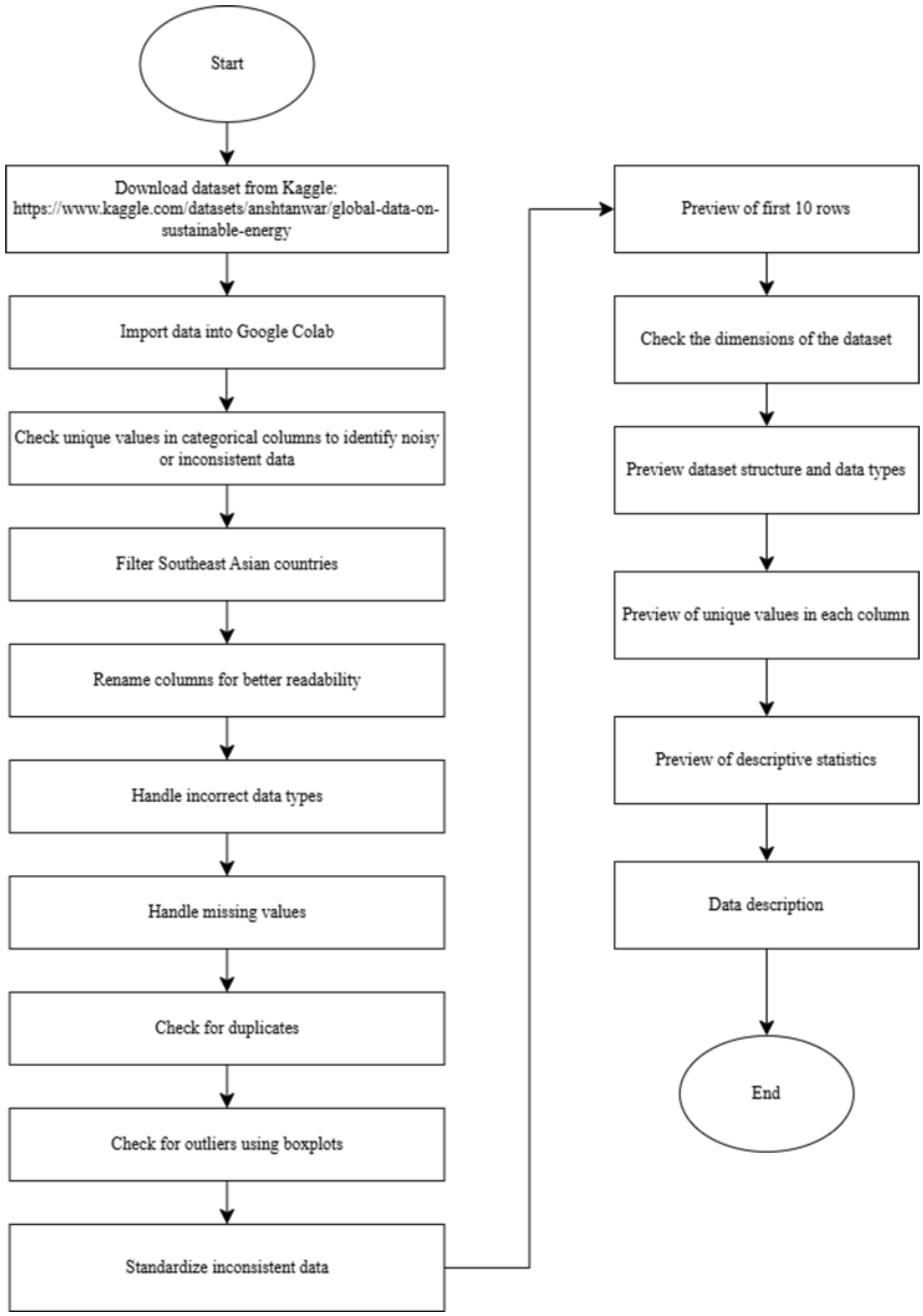

Figure 3.5.1

4.0 EXPLORATORY DATA ANALYSIS

4.1 Overview of Access to Electricity by Country

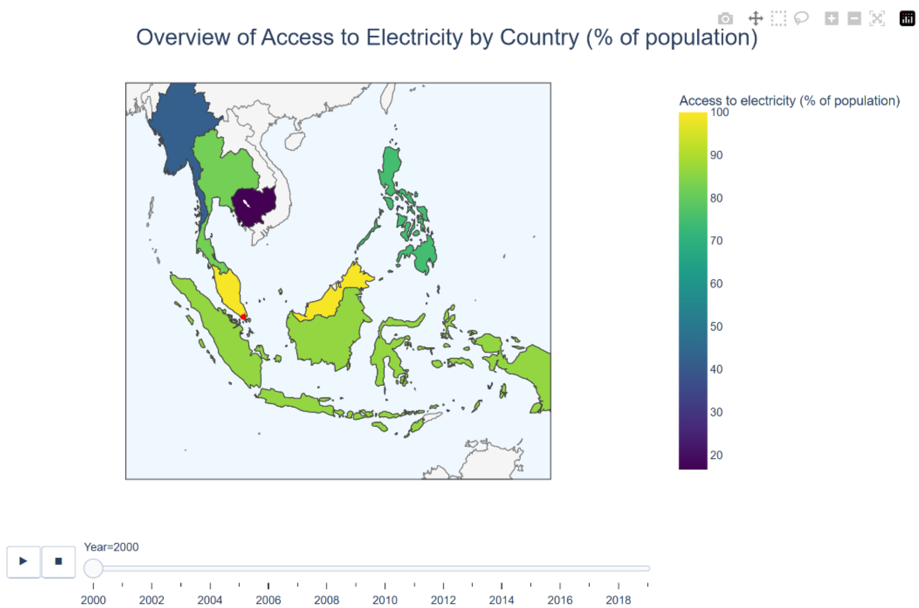

Figure 4.1.1

This animated choropleth illustrates the progression of access to electricity across Southeast Asian countries over time. Using a slider, viewers can select a specific year to see how each country's population access to electricity changes. The colour gradient of the map ranges from dark purple to bright yellow, with brighter shades representing higher access to electricity and more citizens connected to the electric power grid. The darker shades are the opposite to the brighter ones. This means that countries shaded in dark purple have significantly poorer access compared to those in lighter colors, falling behind relative to those countries in the same year. The colour scale dynamically adjusts each year based on the highest and lowest measurement among all countries. This allows for comparisons within the same time frame. For instance, even if a country's electricity access improves over time, it may still appear dark if other countries improve more significantly in that same year.

From the visualization, Singapore stands out with 100% electricity access, represented by the red dot throughout the years. This reflects the country's advanced infrastructure and the fact that it has maintained near-universal access over the period. Malaysia follows a similar trend, with the country appearing in the yellow zone for most of the years, which indicates stable and near-complete access to electricity over time.

Thailand also shows increase in access, turning from green to yellow shades, which suggests a high and consistent level of electricity coverage. In contrast, Indonesia and Philippines display more moderate improvements. Indonesia transitions from a darker green to a lighter yellow tone, indicating gradual progress, but still lags countries like Singapore and Malaysia in terms of full access. The Philippines, while improving, remains in darker green tones, which reflects a slower progress in increasing electricity access.

Myanmar and Cambodia exhibit the slowest rates of progress. Cambodia, initially represented in the purple zone, shifts to a lighter shade over the years, indicating gradual improvement but still limited access. Myanmar shows a similar trend. While access has improved over time, it is still the lowest compared to the six other countries by 2019. This suggests that efforts to expand electricity access in Myanmar are ongoing but comparatively slower.

In conclusion, the map illustrates significant progress across Southeast Asia in expanding electricity access, with countries like Singapore, Malaysia, and Thailand leading the way. However, countries like Indonesia, Philippines, Myanmar, and Cambodia are still working toward full coverage, with varied rates of improvement. This highlights the continued need for infrastructure development and energy access initiatives across the region.

4.2 Average Electricity Generation Mix by Country

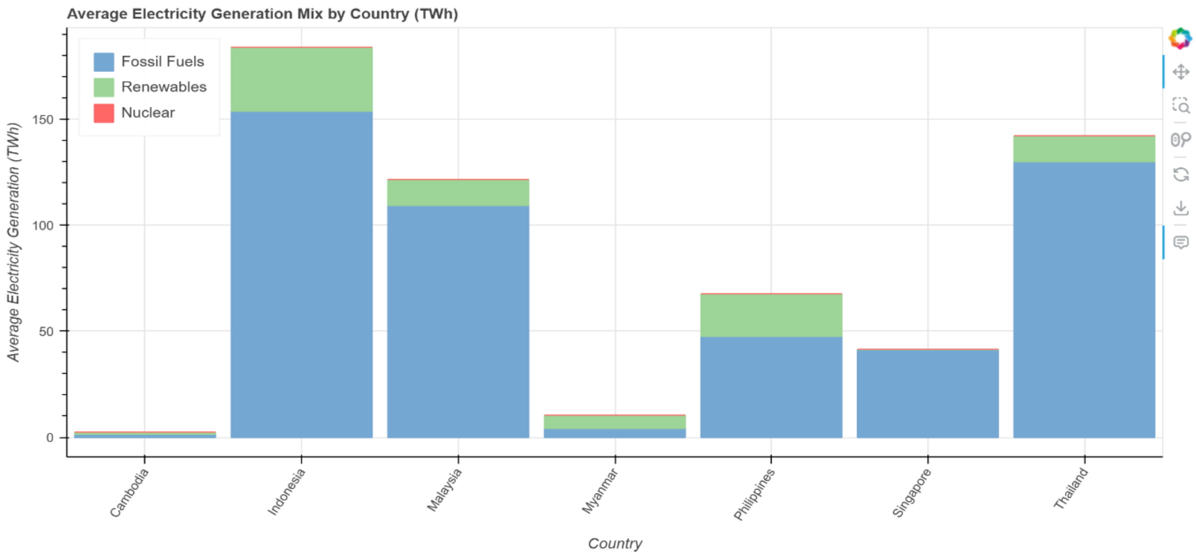

Figure 4.2.1

The bar chart visualizes the average electricity generation mix by country in Southeast Asia from the years 2000 to 2019. It shows the average amount of electricity generated from fossil fuels (blue), renewables (green), and nuclear (red). Each country is represented on the x-axis, and the total average electricity generated is represented on the y-axis in terawatt-hours (TWh).

In the chart, Indonesia stands out as the largest electricity generator in the region, with a dominant reliance on fossil fuels (blue), while also being the top generator of renewable energy (green) among the countries shown. This indicates that despite its high fossil fuel consumption, Indonesia is making considerable progress in incorporating renewable energy into its mix. Thailand follows as the second-largest electricity generator, with a significant portion of its electricity from fossil fuels but also a notable share from renewables. This reflects Thailand's efforts to diversify its energy sources, although its renewable energy share is smaller compared to Indonesia.

Malaysia also generates a large portion of its electricity from fossil fuels, yet it shows a moderate share of renewable energy, highlighting its commitment to sustainable energy solutions. Philippines shows the second-largest share of renewables, with a larger green portion compared to most other countries in Southeast Asia. This suggests that the Philippines is making a more significant effort to incorporate renewable energy sources into its electricity mix, possibly due to its abundant natural resources for renewables like solar and wind energy.

In contrast, Myanmar has a very low total electricity generation, but interestingly, the generation from renewable energy (green) slightly surpasses that from fossil fuels (blue). This suggests that Myanmar is focusing more on renewable sources, despite its low overall generation capacity. On the other hand, Singapore generates a moderate amount of electricity compared to other countries, with no visible green segment. This suggests that it does not rely on renewable energy sources. Instead, the country heavily depends on fossil fuels for most of its electricity production. Finally, Cambodia has the lowest total electricity generation among all the countries. Its generation is minimal, and its reliance on both fossil fuels and renewable energy is almost equal, reflecting a nascent energy sector that still has significant room for development.

Notably, no Southeast Asian country is generating electricity from nuclear energy, as shown by the red portion of the bars, which is absent across all countries. This suggests that while nuclear energy might be considered in some regions, it has not yet been adopted as a major energy source in Southeast Asia.

This graph highlights the region's heavy reliance on fossil fuels for electricity generation, with limited use of renewables and no nuclear energy across Southeast Asia. While countries like the Philippines and Indonesia are making strides in incorporating renewable energy, much of the region's energy production still depends on traditional fossil fuels. The absence of nuclear energy across the region indicates that nuclear power has not yet been a significant part of Southeast Asia's energy strategy, potentially due to concerns over safety, costs, or lack of infrastructure. This chart underscores the challenges and opportunities Southeast Asia faces in transitioning to a more sustainable and diversified energy mix in the coming years.

4.3 Yearly Trend of Access to Clean Fuels by Country

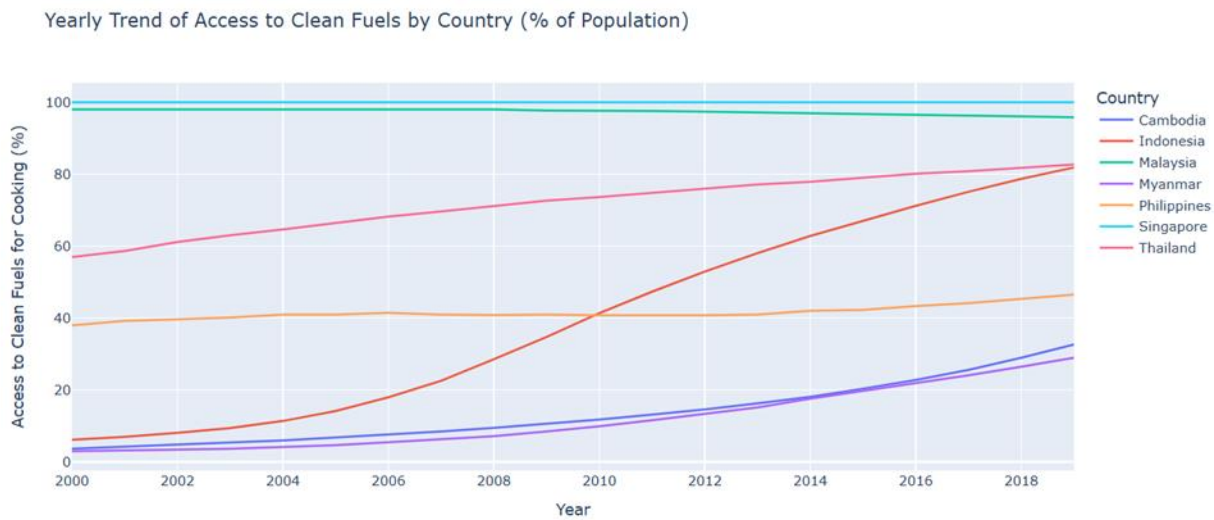

Figure 4.3.1

According to World Health Organization (WHO), clean fuel can be differentiated into many types such as Liquified petroleum gas (LPG), electric, biogas, alcohol and natural gas. Figure 4.3.1 shows the trend in access to clean fuels for cooking by country in Southeast Asia region. From 2000 to 2019, the overall trend in access to clean fuels increased for all countries except for Malaysia. Access to clean fuels in Malaysia decreases slightly across the year but is still high overall, 95% and above due to subsidies by the local government. The slight drop might be due to partial fuel conversion for the purpose of enhancing the food aroma instead of a systemic decline in energy services.

Singapore remains the top performer in ASEAN with 100% of access to clean fuel for cooking. This is because Singapore adopted three types of clean fuel which are piped town gas, electricity for induction stove, and LPG. Singapore government stated that almost 62% of residents rely on the town gas generated by a combination of hydrogen and methane to cook and heat water. Also, electricity supply for induction stoves in Singapore is generated through natural gas which further explains why access to clean fuels in Singapore remains the highest among ASEAN countries.

Cambodia and Myanmar are ASEAN countries with low access to clean fuel for cooking. Economic conditions and unstable political conditions might be the primary factors that lead to low access to clean fuels in these two countries. Residents in Cambodia and Myanmar mostly cook using energy generated by biomass including wood and sawdust because of the availability of these resources for them and low cost. According to the World Bank in 2018, approximately 66.7% of households in Cambodia applied biomass stoves using firewood and coals as their primary stove while 32.9% of households are using clean fuel stoves. Among this 32.9% of households, 30.9% of households use LPG for cooking, and the remaining are using electricity and biogas.

Thailand, Indonesia and Philippines are ASEAN countries that have middle performance in terms of access to clean fuel for cooking. This situation is caused by the economic differences in urban and rural areas. From Figure 4.3.1, these countries are putting effort to improve the access to clean fuel for cooking in their respective countries. For example, the government of Thailand is giving subsidies for LPG, making it affordable to encourage their people using LPG for cooking (Praiwan, 2024b).

In conclusion, the line graph illustrates the trend of access to clean fuels across Southeast Asia countries. Malaysia and Singapore have high accessibility due to uniform urban development and effective government subsidies. Thailand, Myanmar, Indonesia, Cambodia and Philippines show steady growth over the years as the government are trying to reduce the economic difference between urban and rural areas by providing subsidies and improving the available infrastructure.

4.4 Renewable Energy Trends: Installed Capacity vs. Consumption Share by Country

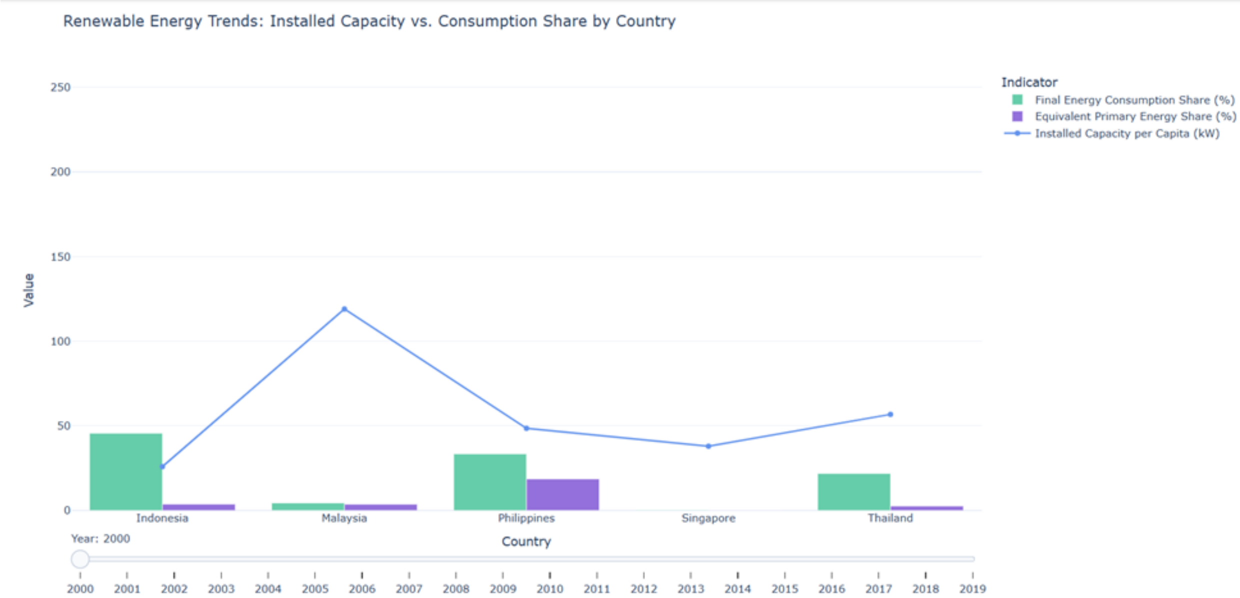

Figure 4.4.1

Figure 4.4.1 shows the renewable energy performance in five ASEAN countries from 2000 to 2019. The figure above compares three key indicators which are final energy consumption, equivalent primary energy share and installed capacity per capita. Philippines and Thailand are the countries with the most stable and moderate performance among these five ASEAN countries. Both have higher final energy consumption than equivalent primary energy share which means that renewable energy is generated into other forms of energy before being used.

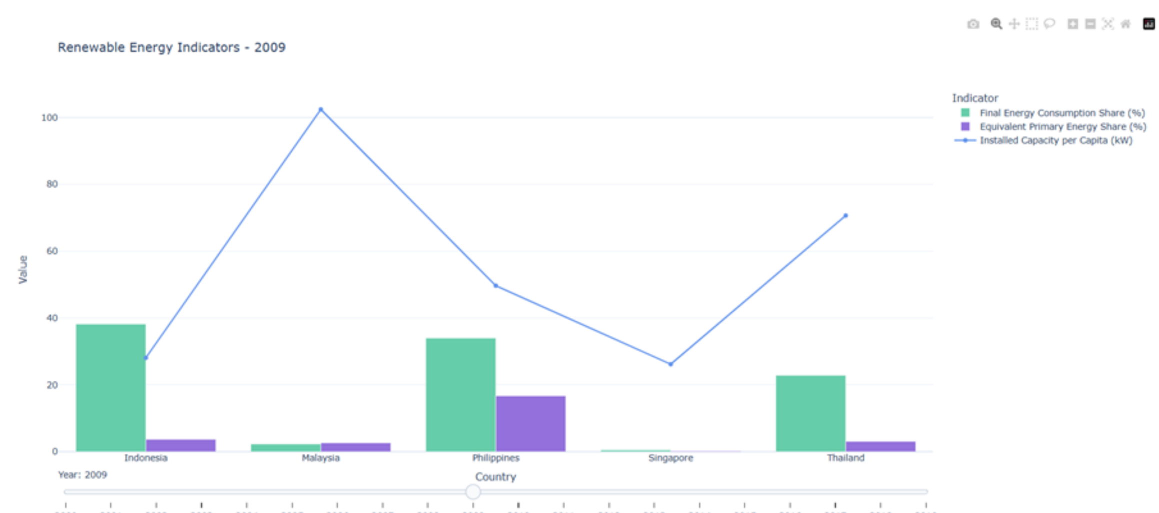

Figure 4.4.2

Among the five ASEAN countries, Malaysia has the most installed renewable energy plant compared to other four countries for all time, but the final energy consumption share considered low which means Malaysia still rely on non-renewable energy such as fossil fuel as their main energy source. However, the Malaysian government always encourages the use of renewable energy which can be shown since 2009 where equivalent primary energy share was higher than final energy consumption used in the end level as shown in Figure 4.4.2. Renewable energy sources such as solar power and hydropower are slowly replacing the traditional energy sources in generating energy.

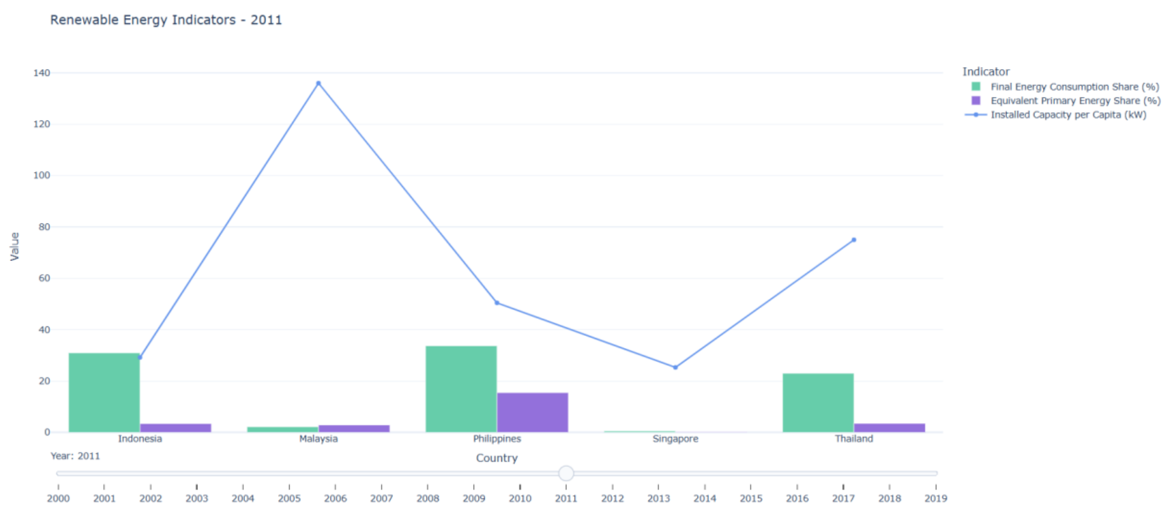

Figure 4.4.3

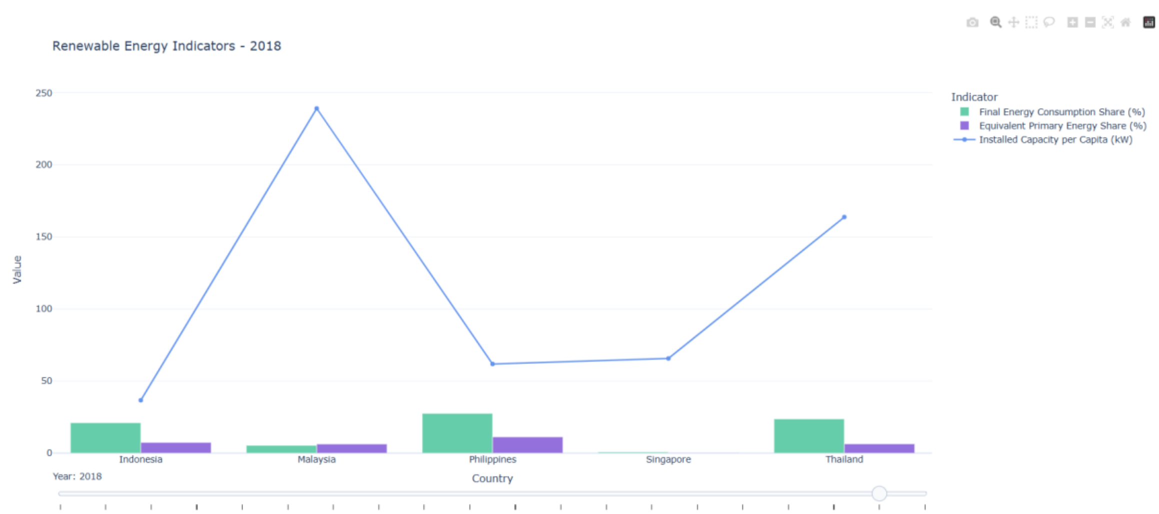

Figure 4.4.4

Indonesia has low installed capacity per capita due to its large population but highest in final energy consumption from renewable sources until 2010 as shown in Figure 4.4.3. Indonesia shows strong reliance on renewable sources as their energy sources like hydropower, geothermal and bioenergy in end-use sectors. In contrast, Singapore has the lowest final energy consumption from renewable and the lowest equivalent primary energy share until 2018 as shown in Figure 4.4.4. This might be due to the limitations of land area available in Singapore which affect the plan of building renewable energy plants in the country. Although installed capacity per capita increases, Singapore remain the lowest in final energy consumption and equivalent primary energy share from renewables. This indicates that Singapore still mainly relies on non-renewables such as natural gas in their energy consumption.

In short, the graph displays the status and consumption of renewable energy among Southeast Asian countries. Across the years, Malaysia had effectively invested in renewable energy which leads to its higher equivalent primary energy share than final energy consumption. While for other countries, installed capacity per capita is increasing, indicating that they are slowly replacing their energy consumption with renewables instead of fossil fuels.

4.5 CO_2 Emissions vs. Primary Energy Consumption per Capita

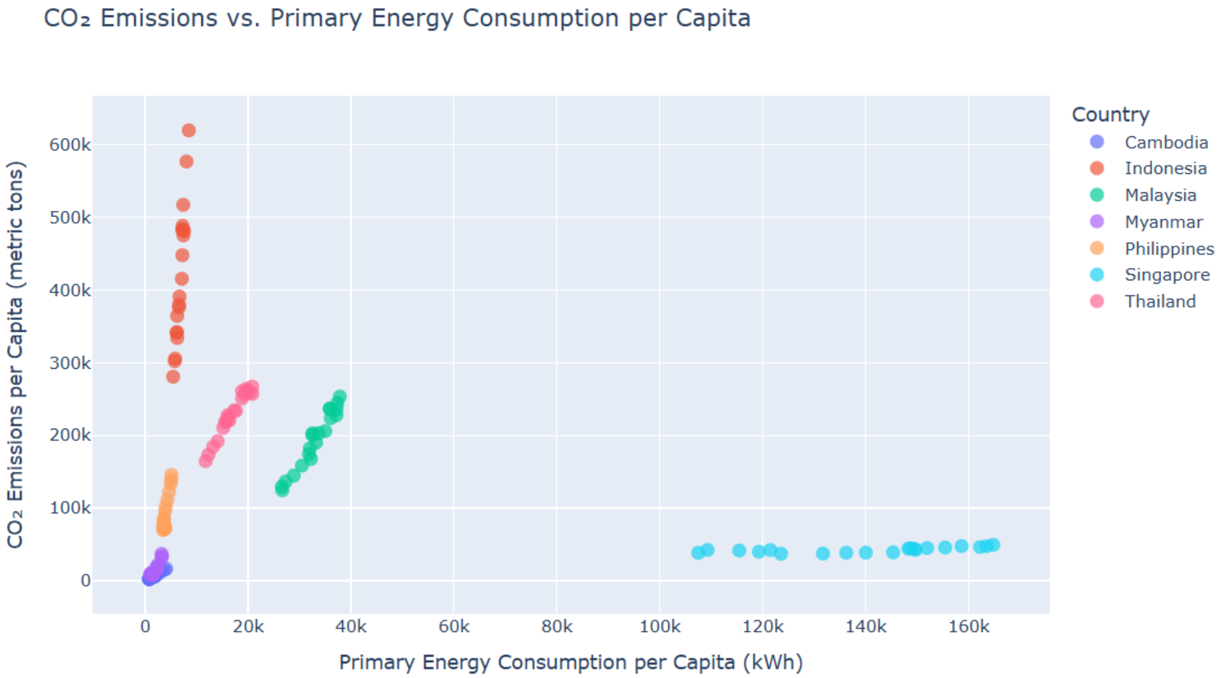

Figure 4.5.1

The scatter plot visualizes CO_2 emissions per capita (y-axis) versus primary energy consumption per capita (x-axis) for seven Southeast Asian countries, each represented by a different color. Each data point corresponds to a specific year, which allows us to observe changes over time by hovering over them. From the plot, we can see that both primary energy consumption and CO_2 emissions per capita have generally increased over the years in every country, although the rate and scale of change differ widely. This shows that total energy demand is increasing, likely due to population growth, urbanization, and industrial development.

By looking at individual countries, Singapore stands out prominently. It has extremely high primary energy consumption per capita, far exceeding the others. Yet, its CO_2 emissions per capita are relatively low compared to its energy use. This suggests that Singapore likely maintains high energy efficiency, possibly due to advanced technology, strict environmental policies, or the use of lower-emission fossil fuels like natural gas. This demonstrates that emissions can be managed effectively even with high energy demand, depending on how energy is produced and consumed. This observation also aligns with the trend shown in the line chart in Section 4.3: Yearly Trend of

Access to Clean Fuels by Country, where Singapore has consistently maintained 100% access to clean fuels for cooking.

On the other hand, Indonesia shows a contrasting trend. Although it has lower energy consumption per capita, its CO_2 emissions are significantly higher. This indicates inefficient energy use or heavy reliance on fossil fuels, especially coal. Malaysia and Thailand fall in the middle range for both energy consumption and emissions per capita, suggesting a moderate level of energy demand likely supported by a mixed energy portfolio that still includes a substantial share of fossil fuels. In contrast, the Philippines, Myanmar, and Cambodia are clustered in the lower left corner of the plot, with low energy consumption and low CO_2 emissions per capita. This reflects their lower levels of industrialization and limited access to energy infrastructure, although these values may increase over time as economic development progresses.

The scatter plot reveals that there is no consistent linear relationship between primary energy consumption and CO_2 emissions per capita across Southeast Asian countries. While most countries show that higher energy consumption tends to correspond with higher CO_2 emissions, Singapore stands out as an exception. This highlights the critical role of energy sources, efficiency, and technology in shaping a country's environmental impact. As economic growth and energy demand continue to rise across the region, these factors will become increasingly important in efforts to reduce emissions and promote sustainable development.

4.6 CO_2 Emissions vs. Renewable Energy Share in Final Consumption

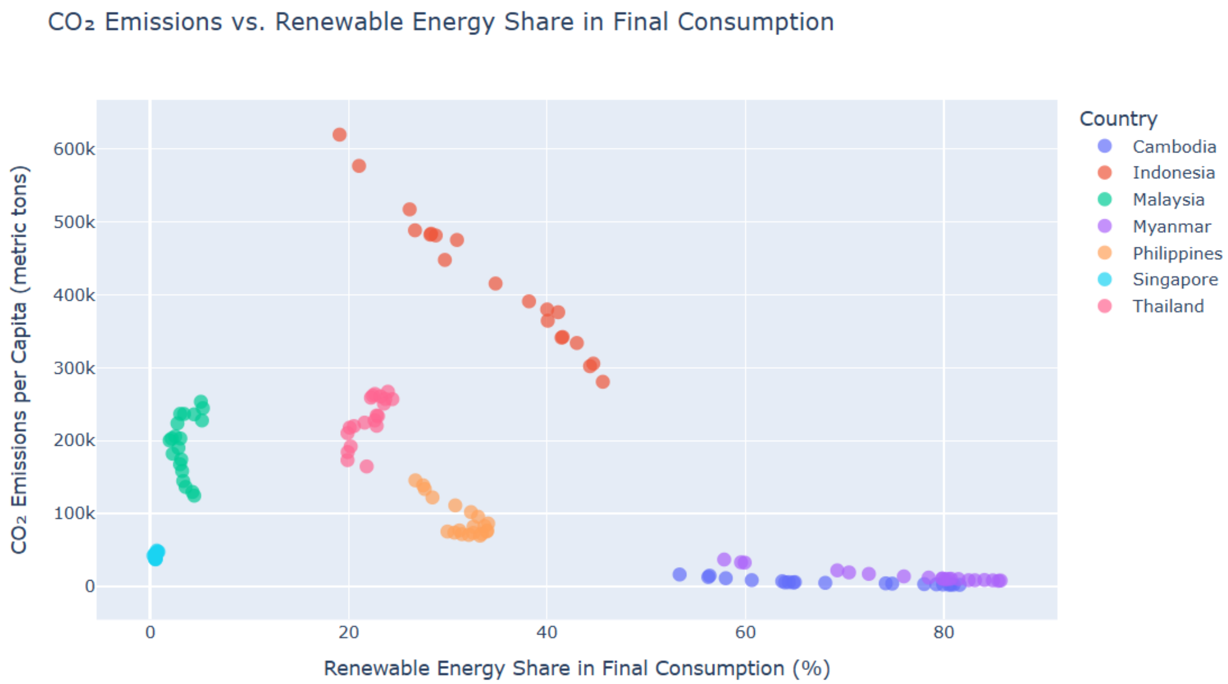

Figure 4.6.1

The scatter plot visualizes CO_2 emissions per capita (y-axis) versus renewable energy share in final energy consumption (x-axis) for seven Southeast Asian countries, each represented by a different color. Each data point corresponds to a specific year, which allows us to observe changes over time by hovering over them. From the plot, a clear negative correlation can be observed overall, which is as the renewable energy share increases, CO_2 emissions per capita tend to decrease. This suggests that greater reliance on renewable energy is associated with lower carbon emissions, which means less CO_2 is released into the atmosphere, helping to reduce the greenhouse effect and mitigate global warming.

Looking at individual countries, Indonesia clearly demonstrates this trend. Its data points show that as the share of renewable energy increases, CO_2 emissions per capita steadily decrease. This serves as strong evidence that transitioning to renewables can effectively reduce emissions, even in countries with growing energy demands. Besides that, Cambodia and Myanmar have very high renewable energy shares and very low CO_2 emissions per capita. This indicates that their energy mix is already dominated by renewables, such as hydropower or biomass, which helps keep their carbon footprint low.

In contrast, Singapore presents an interesting outlier. It has almost no renewable energy share, yet its CO_2 emissions per capita remain relatively low. This shows that while Singapore does not depend on renewable energy, it likely applies efficient energy systems, cleaner burning fuels, and strict environmental regulations. Therefore, CO_2 emissions not only can be reduced through the usage of renewable energy, but also through how the energy used is managed.

Malaysia, Thailand, and the Philippines show unstable or non-linear trends in the plot, with no consistent reduction in CO_2 emissions despite moderate or rising renewable energy shares. This may reflect a transitional phase in their energy journey, where renewables are being adopted, but fossil fuels still dominate. It also suggests that increasing renewable energy alone is not enough because other factors like energy efficiency, fuel type, industrial output, and policy enforcement also significantly influence emission levels.

Overall, the scatter plot highlights a general negative correlation between renewable energy share and CO_2 emissions per capita. However, the variations among individual countries show that the relationship might be influenced by multiple factors beyond just renewable energy adoption. This emphasizes that while expanding renewable energy is crucial, a comprehensive approach including energy efficiency improvements and regulatory frameworks is necessary to effectively contribute to a more sustainable environment.

4.7 Energy Intensity by Country with Yearly Variation

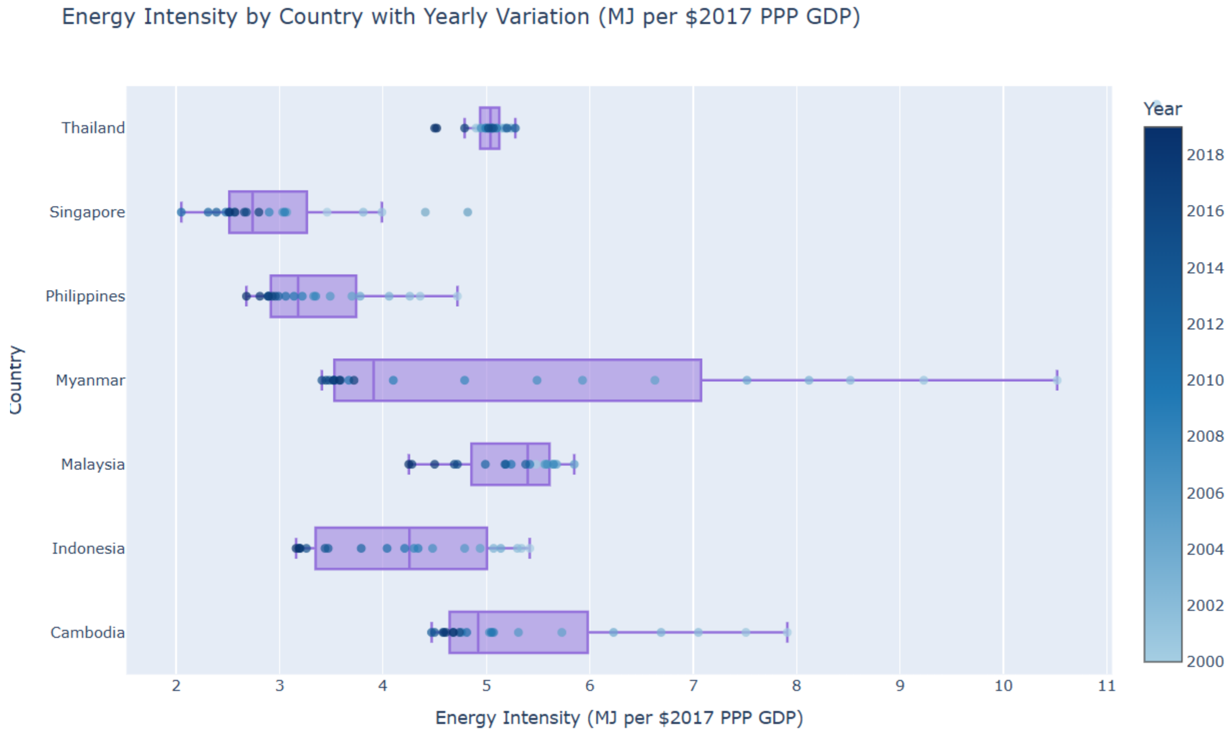

Figure 4.7.1

The chart above shows a horizontal boxplot that visualizes energy intensity by country, with data points spanning from 2000 to 2008. Each country has a boxplot that summarizes the variation in energy intensity by year. Energy intensity is a measure of the energy efficiency of a country's economy, and it is calculated as the amount of energy consumed per unit of gross domestic product (GDP) where it is typically expressed in megajoules (MJ) per dollar of GDP, adjusted for purchasing power parity (PPP) and inflation with unit $2017 PPP GDP. Higher energy intensity indicates that more energy is used to produce one unit of the economic output, meaning lower energy efficiency. In contrast, lower energy intensity suggests higher energy efficiency and a more energy-productive economy.

To dive into the insights of each country, each boxplot will be explained in detail. Thailand shows a relatively high energy intensity values, with tight cluster between 4.79 and 5.28 MJ/$2017 PPP GDP. The narrow box suggests that energy efficiency remained fairly consistent, from 2000 to 2017, without large changes. The dark-coloured points indicate that recent years have hovered around the range, suggesting there is not much significant improvement in energy efficiency. However, there exist two outliers for the years 2018 and 2019 located above the upper whisker of the box. This means that during the two years, Thailand recorded an unusually lower energy intensity value compared to the previous years, indicating Thailand experienced a more energy-efficient economy in 2018 and 2019.

For Singapore, this country stands out for having the lowest energy intensity among all countries in the chart. Its values range from 2.05 to 3.99 MJ/$2017 PPP GDP, which indicates high energy efficiency in Singapore from 2000 to 2019 except for 2003 and 2004 which had slightly higher energy intensity. The boxplot is right skewed, indicating Singapore has strong performance in terms of energy usage relative to economic output. In most recent years, data points tend to cluster to the left, indicating a slightly improving trend in energy efficiency.

Philippines has a moderate energy intensity, between 2.68 to 4.72 MJ/$2017 PPP GDP with a slight upward trend over time. The data points show some variability but there are no outliers in the distribution. The distribution suggests that the country has slight improvement from 2000 to 2019 with the darkest datapoint are distributed at the left. The energy intensity is positively skewed with a median of 3.18 MJ/$2017 PPP GDP, which suggests that the country is trying to make some progress when they are still facing moderate energy efficiency challenges.

Next, Myanmar exhibits the highest spread and variation in energy intensity values over the years, ranging from 3.41 to 10.52 MJ/$2017 PPP GDP. The highest energy intensity occurred in the year 2000, which indicates instability and inefficiencies in energy use. The earlier year tends to show higher energy intensity, but more recent years tend to reflect a downward trend, explaining that Myanmar has seen improvements recently, but still remains less efficient compared to other countries.

Furthermore, Malaysia's energy intensity is concentrated between 4.25 and 5.85 MJ/$2017 PPP GDP, with limited spread. This suggests a stable but unchanging pattern and a consistent growth from 2000 to 2019. It shows a consistent increase towards energy efficiency in Malaysia.

Indonesia shows a moderate energy intensity, ranging from 3.16 to 5.42 MJ/$2017 PPP GDP. The spread of data points and box indicate significant variation over the years. The boxplot is almost symmetric but slightly left skewed, suggesting that energy intensity values are fairly evenly distributed, with gradual changes rather than sudden increase in energy efficiency.

Last but not least, Cambodia has a widespread in energy intensity values, from 4.47 to 7.91 MJ/$2017 PPP GDP, with a median of 4.92 MJ/$2017 PPP GDP. The right-skewed boxplot with most light-coloured points in the higher range indicates that energy intensity was very high in the early 2000s. However, the values have gradually declined over time, indicating that Cambodia has made progress in improving energy efficiency in recent years.

In short, Singapore consistently outperforms other Southeast Asian countries in terms of energy efficiency, by recording the lowest energy intensity. Myanmar and Cambodia show a large spread in energy intensity values, but both have demonstrated a general improvement from the past to more recent years. Thailand and Malaysia have comparatively higher medians with smaller spreads, reflecting both countries rely on energy-intensive industries such as manufacturing sector. This analysis highlights the necessity for each country to improve their energy infrastructure for long term economic growth.

4.8 3D View: GDP per Capita vs. Renewable Energy Share vs. Energy Intensity Level of Primary Energy

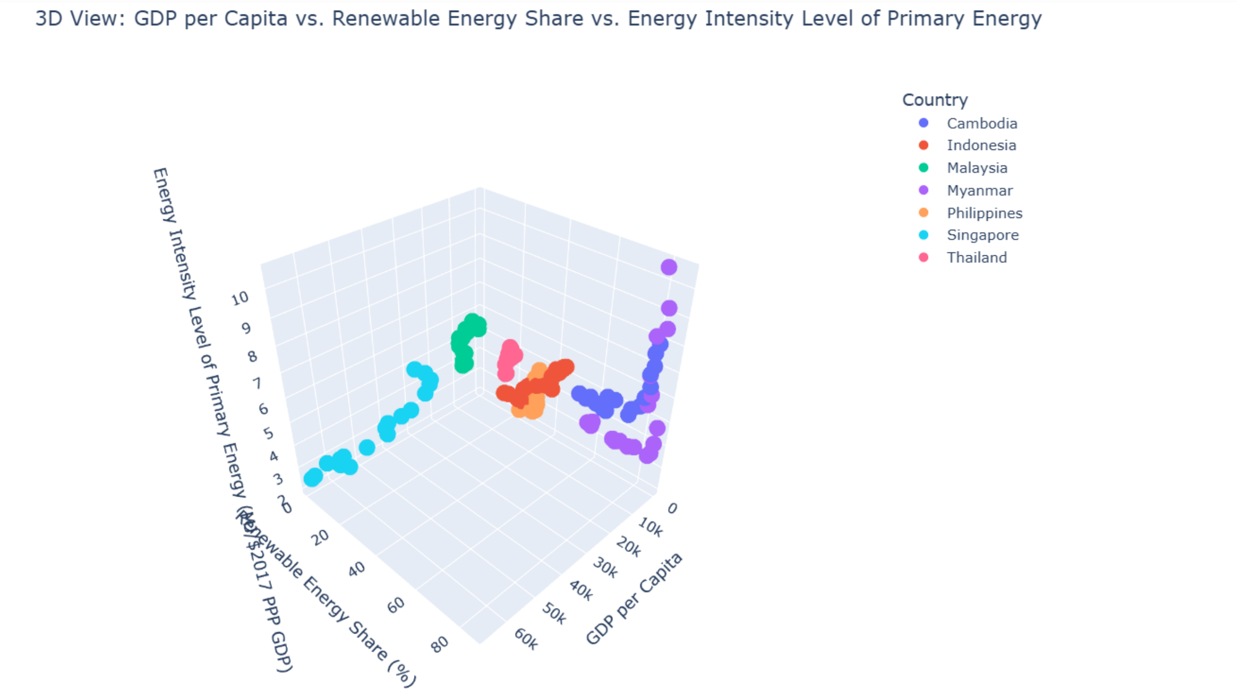

Figure 4.8.1

The 3D scatterplot illustrates the relationship between GDP per capita, renewable energy share in total energy consumption, and the energy intensity level of primary energy across the Southeast Asian countries. The visualization reveals the clear difference in energy profiles and the respective economic development among these nations. Each point in the 3D space represents a country achievement in a specific year, summarizing the relationship of these three variables in each country over time.

One of the obvious observations is the outstanding performance of Singapore, with the highest GDP per capita among all Southeast Asian countries, which signifies a highly developed economy. However, it has a low energy intensity level, which indicates that energy efficiency is ensured in producing each unit of economic output. The renewable energy share of Singapore is the lowest among all countries, suggesting that Singapore has lower adoption of renewable energy. This could be due to the limitations in natural renewable resources for implementation of renewable energy in the country.

In contrast, Myanmar and Cambodia show a totally different pattern. Both countries have low GDP per capita, reflecting limited economic development, but they show a relatively high share of renewable energy in total consumption. This can be related with the potential of energy sources in underdeveloped areas. At the same time, their energy intensity levels remain high, demonstrating the inefficient energy usage that might be caused by outdated technologies and energy losses in economic sector even with high usage of renewable energy.

Next, middle-income countries like Malaysia, Thailand, and Indonesia fall within a more balanced zone on the plot. They display moderate GDP per capita, intermediate renewable energy shares and moderate energy intensity levels. These countries are currently in transition, trying to balance both economic growth and sustainable energy consumption. Malaysia and Thailand are actively increasing the renewable energy share while also focusing on energy efficiency measures, which reflects an ongoing progress toward sustainable development.

To summarize, the 3D scatterplot shows the relationship between development in economic growth, renewable energy use, and energy intensity from 2000 to 2019. Singapore demonstrates better energy efficiency for greater economic growth, without reliance on renewable energy. In contrast, less developed countries might rely on renewables but face challenges in improving efficiency. This visualization suggests each country's unique contexts in balancing renewables and energy efficiency in relation to economic growth. The authorities can tailor specific energy policies and planning to address the current challenges each Southeast Asian country faces.

5.0 SUMMARY

This section gives a summary of the key integer attributes used in the analysis. Each variable is described based on its role in the visualizations, its range, and its significance in understanding energy trends across Southeast Asian countries. The summaries help highlight patterns and insights derived from the data.

Access to Electricity (% of population) represents the percentage of the population in each country that has access to electricity. It is an integer value ranging from 0 to 100, and it was the key variable in the animated choropleth map (Visualization 1) in Section 4.1 Overview of Access to Electricity by Country. Most Southeast Asian countries have shown a steady increase in access over the years, with Singapore reaching full access (100%) and Malaysia nearing full access, while countries like Indonesia, Philippines, Myanmar, and Cambodia have gradually improved but still face challenges in achieving universal coverage. This metric directly reflects progress in infrastructure development and energy equity across the region.

Electricity from Fossil Fuels, Renewables, and Nuclear (TWh) were the three integer attributes used in the stacked bar chart (Visualization 2) of Section 4.2 Average Electricity Generation Mix by Country. The values are measured in terawatt-hours (TWh), indicating the volume of electricity produced from each source. Most of the Southeast Asia countries like Indonesia and Thailand rely heavily on fossil fuels to generate electricity while renewables also contribute as one of the sources to generate electricity. Myanmar is an exception, showing a higher proportion of electricity generated from renewables compared to fossil fuels despite its low total generation capacity. Nuclear generation is non-existent in all the Southeast Asian countries.

Access to Clean Fuels for Cooking (% of population) is visualized in the line chart (Visualization 3) of Section 4.3, showing the proportion of people using clean cooking fuels like LPG and electricity instead of biomass or coal. Overall, Southeast Asian countries have shown steady improvement from 2000 to 2019, with Singapore and Malaysia maintaining high access levels. Other countries demonstrate gradual growth due to government efforts to improve infrastructure and reduce urban-rural disparities. This trend reflects ongoing progress in addressing energy poverty and promoting cleaner, healthier cooking solutions in the region.

Renewable Energy Share in Final Energy Consumption (%) was used in multiple visualizations (Visualizations 4 and 6) in Sections 4.4 Renewable Energy Trends: Installed Capacity vs. Consumption Share by Country and 4.6 CO_2 Emissions vs. Renewable Energy Share in Final Consumption. This percentage indicates how much of a country's total final energy use comes from renewable sources. Higher values suggest better integration of clean energy into everyday energy consumption, which reflects a stronger commitment to sustainability. For example, Indonesia has the highest percentage in final energy consumption while Singapore has the lowest.

Renewables (% Equivalent Primary Energy) was another percentage-based metric used in Visualization 4. It measures the share of renewables in total primary energy supply, adjusted for equivalent energy content. It offers a broader perspective than final energy share, as it accounts for upstream energy losses and conversion efficiencies. Philippines is the best-performing country among Southeast Asia countries with the highest percentage of equivalent primary energy while Singapore has the lowest. This reflects the varying effectiveness of using renewables in generating energy.

Installed Renewable Electricity Capacity per Capita (kW) used in the same visualization (Visualization 4), shows how much renewable electricity capacity is available per person in each country. It reflects a country's infrastructure readiness for renewable energy. Countries with small populations but heavy investment in renewables tend to have higher per capita values. For instance, Malaysia has the highest installed renewable electricity capacity per capita for all time. In contrast, Indonesia has the lowest due to its large population.

Primary Energy Consumption per Capita (kWh) was used in Visualization 5 to compare with CO_2 emissions in Section 4.5 CO_2 Emissions vs. Primary Energy Consumption per Capita. This metric indicates the energy use per person and reflects the level of energy demand and industrial activity in a country. Higher values suggest greater energy consumption, which is often linked to economic development. For example, Singapore has the highest energy use per capita, showing its advanced infrastructure and energy-intensive lifestyle, while countries like Cambodia and Myanmar have the lowest due to lower industrialization and limited energy access.

CO_2 Emissions per Capita (metric tons) was used in both scatter plots (Visualizations 5 and 6), to show each country's carbon emissions per person. This variable serves as a key indicator of a country's environmental impact and energy sustainability. The visualizations revealed varied emission levels across Southeast Asian countries. While higher energy use often corresponds with increased emissions, countries like Singapore demonstrate that emissions can be managed effectively through advanced technologies and cleaner fuels. Additionally, the negative correlation between renewable energy share and emissions highlights the importance of transitioning to renewables, though energy management and policy also play crucial roles in reducing carbon footprints.

Energy Intensity Level of Primary Energy (MJ/$2017 PPP GDP), used in Visualization 7 and 8 in Sections 4.7 Energy Intensity by Country with Yearly Variation (MJ per $2017 PPP GDP) and 4.8 3D View: GDP per Capita vs. Renewable Energy Share vs. Energy Intensity Level of Primary Energy measures how much energy is used to generate one dollar of economic output (PPP adjusted). A lower value indicates better energy efficiency. The boxplot and 3D scatterplot highlight countries with unusually high or low energy intensities and track their variation by year. Singapore has the lowest energy intensity, indicating the highest energy efficiency while Myanmar and Cambodia have outstanding improvement from the past to recent.

GDP per Capita although typically treated as a continuous variable, was used as an integer in the 3D plot (Visualization 8). It provides economic context when evaluating energy consumption and emissions. Countries with stronger economics tend to have lower energy intensity or higher energy efficiency, while lower income developing countries might use a high percentage of renewable energy but lack energy efficiency. However, the relationship is not always straightforward as it depends on each country's policy and priorities.

The analysis provides a comprehensive overview of energy and sustainability indicators across Southeast Asian countries, revealing important interconnections between access, consumption, generation sources, and emissions. From tracking electricity access and clean cooking fuel usage to examining energy generation mixes and renewable energy shares, the findings highlight both progress and disparities within the region. Countries like Singapore and Malaysia consistently lead in energy access and efficiency, while others like Myanmar and Cambodia show promising improvement trajectories. Notably, the data reveals a strong relationship between higher renewable energy adoption and lower CO_2 emissions, supporting the case for cleaner energy investments. In addition, measures like energy intensity and GDP per capita help show how well countries use energy to support their economies.

Overall, the findings highlight the need for strategies that fit each country's situation to support fair access to energy, cleaner energy use, and better efficiency as Southeast Asia moves toward a more sustainable energy future. ASEAN countries should come out with policies such as subsidies such as tax reductions to encourage private sectors or organizations to help the effort in reducing the usage of non-renewable energy while increasing the dependence on renewable energy such as solar power. ASEAN countries can also work together by continuously cooperating and communicating together on renewable energy projects to hasten the phase of energy transitions for all Southeast Asia countries. Other than that, ASEAN countries can look for more potential investments outside the region to gain more funds for more flexibility to build and improve renewable energy systems, support new ideas, and move faster toward clean and reliable energy for the future.

6.0 REFERENCES

Association of Southeast Asian Nations. (n.d.). *Priority areas of cooperation.* ASEAN. https://asean.org/our-communities/economic-community/asean-energy-cooperation/priority-areas-of-cooperation/

Berdikeeva, S. (2023, March 16). *Tesla and the Future of Clean Energy | SaveOnEnergy.com*. SaveOnEnergy.com. https://www.saveonenergy.com/green-energy/tesla-future-of-clean-energy/

Defining clean fuels and technologies. (n.d.). https://www.who.int/tools/clean-household-energy-solutions-toolkit/module-7-defining-clean

Energy and utilities. (n.d.). Base. https://www.singstat.gov.sg/publications/reference/ebook/industry/energy-and-utilities

Gomstyn, A. (2024, August 21). *The History of Renewable Energy | IBM*. https://www.ibm.com/think/topics/renewable-energy-history

History of Renewable Energy Timeline | Inspire Clean Energy. (n.d.). https://www.inspirecleanenergy.com/blog/clean-energy-101/earth-day-technology-timeline

Koons, E. (2024, April 5). *Renewable energy in Malaysia 2024 - From oil to sustainability*. Energy Tracker Asia. https://energytracker.asia/renewable-energy-in-malaysia/

Praiwan, Y. (2024b, June 27). Cooking gas subsidies extended to Sept 30. Bangkok Post. https://www.bangkokpost.com/business/general/2818945/cooking-gas-subsidies-extended-to-sept-30?

Rutu Dave, Sandra Keller, Bryan Bonsuk Koo, and Gina Fleurantin, Elisa Portale and Dana Rysankova. (2018, June). Beyond Connections Cambodia Energy Access Diagnostic Report Based on the Muilti-Tier Framework. www.worldbank.org

7.0 APPENDIX

Google Drive link:

https://drive.google.com/drive/folders/1LGbGAE3h12kJiUKtRmDIGQIm8gOMHklB?usp=sharing

The Google Drive link above provides access to the following files:

- Original dataset – *global-data-on-sustainable-energy.csv*
- Cleaned main dataset – *sea_data.csv*
- Cleaned filtered dataset – *sea_data_renewables.csv*
- Google Colab notebook for data preparation – *Data_Preparation.ipynb*
- Google Colab notebook for data visualization – *Data_Visualization.ipynb*