# Introduction

The data set which will be used throughout the project is retrieved from the Kaggle website, the data set contains information regarding “Global data on sustainable energy” from the year 2000 to 2020, the data set contains information about key sustainable indicators across all countries it also contains information regarding, electricity access, renewable energy, carbon emission and economic growth indicators of the countries for the year 2000 to 2020.

Data set had multiple columns containing information of countries socio-economic attributes and key indicators on renewable energy and the overall energy production of the country. The variable name and description of each variable can be found in the below variable definition table.

Table 1: Variable Description table

|  |  |  |
| --- | --- | --- |
| **Variable** | **Definition** | **Source** |
| Entity | The name of the country or region for which data is reported | (Tanwar, 2023) |
| Year | The year for which data is reported ranging from 2000 to 2020 |
| Access to electricity (% of population) | The percentage of population with access to electricity |
| Access to clean fuels for cooking (% of population) | The percentage of the population with primary reliance on clean fuels. |
| Renewable-electricity-generating-capacity-per-capita | Installed Renewable energy capacity per person |
| Financial flows to developing countries (US $) | Aid and assistance from developed countries for clean energy projects. |
| 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 |
| Low-carbon electricity (% electricity) | Percentage of electricity from low-carbon sources (nuclear and renewables) |
| Primary energy consumption per capita (kWh/person) | Energy consumption per person in kilowatt-hours. |
| Energy intensity level of primary energy (MJ/$2011 PPP GDP) | Energy use per unit of GDP at purchasing power parity |
| Value\_co2\_emissions (metric tons per capita) | Carbon dioxide emissions per person in metric tons |
| Renewables (% equivalent primary energy) | Equivalent primary energy that is derived from renewable sources |
| GDP growth (annual %) | Annual GDP growth rate based on constant local currency. |
| GDP per capita | Gross domestic product per person. |
| Density (P/Km2) | Population density in persons per square kilometer. |
| Land Area (Km2) | Total land area in square kilometers |
| Latitude: | Latitude of the country's centroid in decimal degrees |
| Longitude: | Longitude of the country's centroid in decimal degrees |

# Data Transformation and modeling

The selected data set on renewable energy will undergo power BI power query editor to make the necessary transformation to the data set, which will be used throughout this project. Using built in power query functionality extract-transform-Load (ETL) will be performed on the selected data set to prepare the data to be used in building interactive dashboards.

Initial data transformation started with setting the right rows as the headers of the column along with the optimal data type which fits the column’s attributes, most of the data set in the columns were numerical, almost most of the data columns were recorded with the correct data type, the only change made was to the “Year” column which was changed from integer type to date type, to represent the years. Once the necessary data types have been set appropriately to match the recorded data instances, we focused on column renaming, which will be beneficial to the team during dashboard building and data analysis. Almost all the column’s names represented prior to transformation were clear and understandable, the only renaming change we made was to the ’Entity’ column, later it was renamed to ‘Countries’ to match the data instanced recorded in that specific column.

Power query M (used to rename columns) - = Table.RenameColumns(#"Changed Type",{{"Entity", "Countries"}})

Later we analyzed the data set to identify any unwanted columns which does not add any contribution to building the dashboard. We identified “longitude” and “Latitude” columns that were deemed unnecessary in this project, since power BI is smart enough to detect the countries by the name, it does not require precise longitudinal and latitude values of the country.

There were certain columns with null values associated with it, the data set we are working on signifies the energy generation from each country, certain economic key indicators. This type of data is not available for every country due to disparity in resource availability between developed countries and other developing nations. During data analysis in order to tackle any data type issues, we simply replaced the ‘null’ values with ‘0’, which will be beneficial during calculations made to be displayed in the interactive dashboard.

The data set contains values accumulated over two decades from the year 2000 to 2020. In order to reduce the computational complexity during the process of interactive dashboard building. We planned on building a dashboard to represent the data from the year 2016 to 2019, our initial plan was to build the interactive dashboard for 2020 as well, since the data set we selected did not have sufficient amount of updated data on renewable energy, and electricity generation of the countries in the year 2020.

Power query M (used to filter the columns by selected year) - = Table.SelectRows(#"Replaced Value", each [Year] = #date(2016, 1, 1) or [Year] = #date(2017, 1, 1) or [Year] = #date(2018, 1, 1) or [Year] = #date(2019, 1, 1))

In the final stages of our data transformation, we added new columns named “organization” which represents whether the country falls under the group of SAARC nations, since our aim is to create a dashboard with certain overall key indicators which will be analyzed worldwide and provide a drill down report on South Asian countries. This column will be used to filter the data to make detailed reports on south Asian countries. Right after creating the column, using the power BI transform window, we created another table which only contains the information of countries in SAARC nations list. This was done in order to ease the filtration process when building the dashboard on south Asian countries’ total renewable energy generation in (TWh), and total fossil fuel generation of south Asian countries.

if List.Contains({"India", "Nepal", "Bangladesh", "Pakistan", "Sri Lanka", "Bhutan", "Maldives", "Afghanistan"}, [Countries]) then "SAARC Country" else "Non-SAARC Country"

During the modelling process we combined both nuclear energy generation columns and renewable energy generation columns, to sum and combine them into a single column naming sustainable energy, which will be used in further model building processes. Once all the necessary data table and column is transformed using power query in power BI, many to many relationships between the energy table and SAARC table are created to enable cross slicer function to filter the graph as per user’s desire.

Table 2: Table relationship

A screenshot of a computer

Description automatically generated

# Explanation of the design choices and overall analysis of the dashboard

## Page 1- world overview

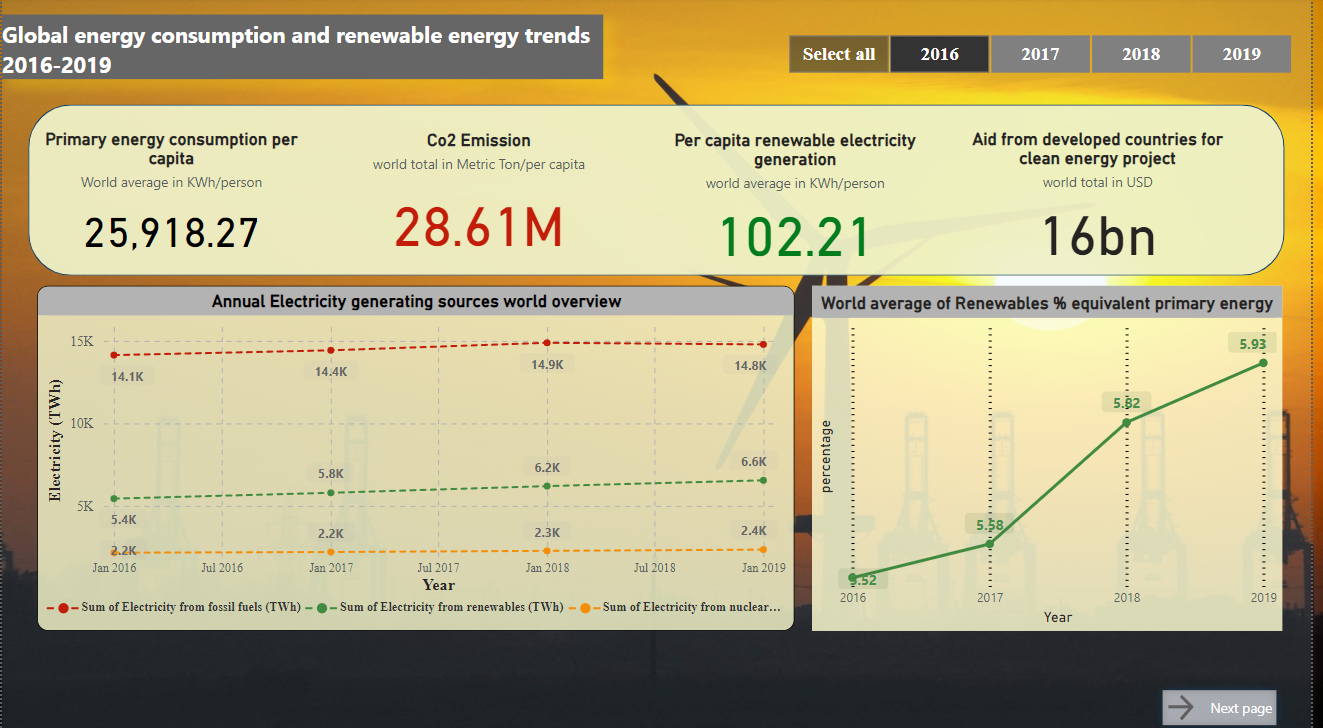


Figure 1: Dashboard page 1

The first page of this dashboard contains insight from the world statistics from renewable energy generated per capita, Co2 emission per capita, primary energy consumption per capita and aid flowed out of developed countries to developing countries for clean energy projects.

In order to display the world average energy consumption per capita (KWh/person), world total CO2 emission (Metric ton/person), per capita renewable energy electricity generation (KWh/person) and finally summation of total aid for renewable project in USD were displayed in the dashboard using key cards, which is often t used to highlight the key metrics, the reason we went on with key card is, since the page one of the dashboard deals with the insights from world energy generation, we want the user to gain quick summary of the above mentioned key metrics, to indicate the world’s stance on sustainable energy goals. We inserted a slicer at the top right corner where users have the ability to change the years to look for numbers generated during those specific years. Users also have the ability to multi-select to combine multiple years to view the number generated for the above-mentioned attributes for that specific year.

We have used line charts to visualize the annual world overview of sources used to generate electricity, and the world average of renewable energy percentage equivalent to primary energy. The reason we went on with line chart is, to handle time series data, to effectively showcase how data points are evolved over time, line charts are typically used to identify trends, so we can identify any upward or downwards trends in electricity generating sources over the years worldwide, it can even be used to compared different category on the same chart , in our case we can compares the trends in fossil fuel electricity generation and renewable energy sources such as (nuclear, solar, wind and hydro). This gives a clear understanding on which source contributes to highest electricity generation and from those sources we can identify whether the world relies heavily on renewable sources or fossil fuels which are harmful to the environment. From our data points it is quite clear that the world still heavily relies on fossil fuel as their primary source of electricity generation, which highlights the importance of work that needs to be done in order to make the switch to renewable energy sources. The second line chart which contains insight of the world's renewable energy percentage equivalent to total primary energy shows gradual increase throughout the years, which ensures there has been an upward trend in renewable energy generation worldwide**.**

## Page 2- SAARC Nations

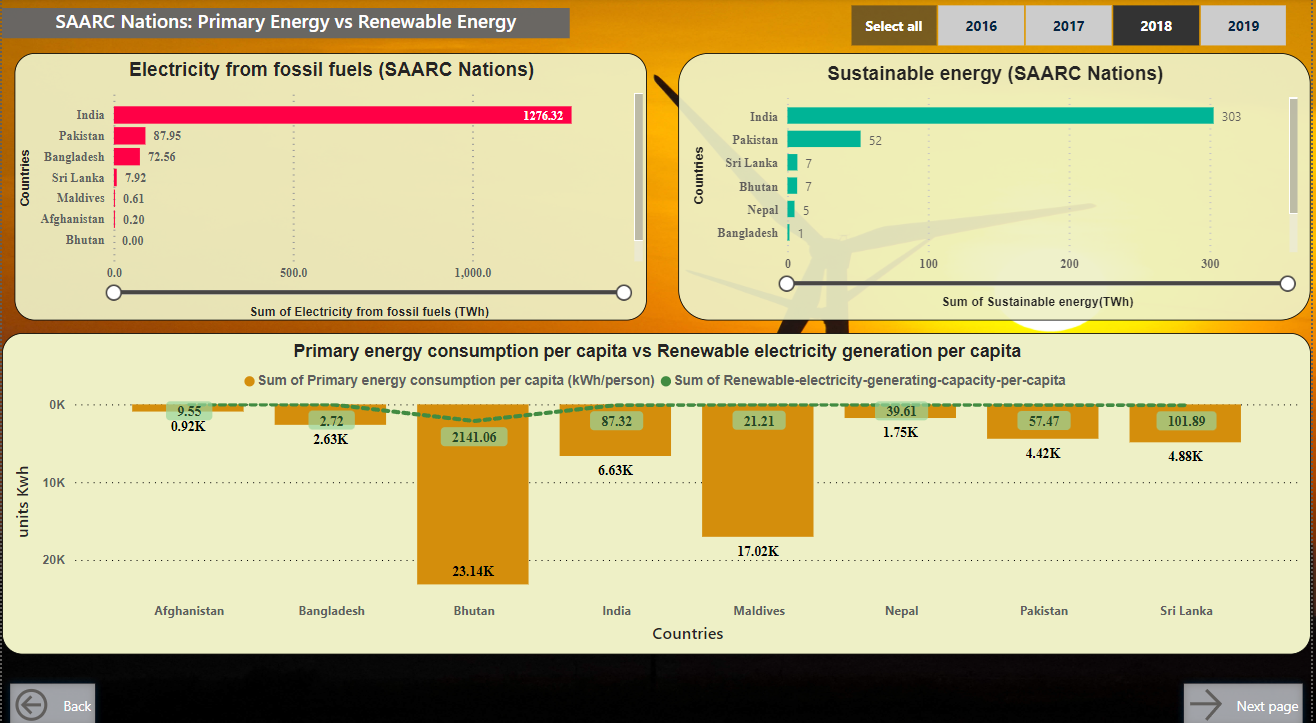


Figure 2: Dashboard page 2

The main aim of this dashboard page is to visualize the energy generated from fossil fuel sources and sustainable sources (nuclear and renewable) by South Asian countries which are a part of SAARC organization. The main visualization tools we used here are stacked bar charts, and line and stacked column charts. Stacked bar chart shows the detailed breakdown of the energy generated in (Twh) using the mentioned sources. From the stacked bar chart, it is quite clear that India heavily relies on fossil fuel as their main source of generating electricity. From the sustainable energy generation still, India tops the chart since it is the only nation among the other SAARC countries which has invested significantly in green energy to achieve their sustainable goals. We used line and stacked column charts to explain the trend between primary energy consumption per capita vs renewable energy generation per capita. We used this graph to easily obtain which countries balance off their total energy consumption per individual with the amount of sustainable energy generated per individual. The graph shows that all the SAARC nations need to explore further options to bring more electricity from sustainable energy to their main electricity grid and find different sources of renewable energy to increase the renewable electricity generating capacity per capita.

## Page 3 – sustainable map

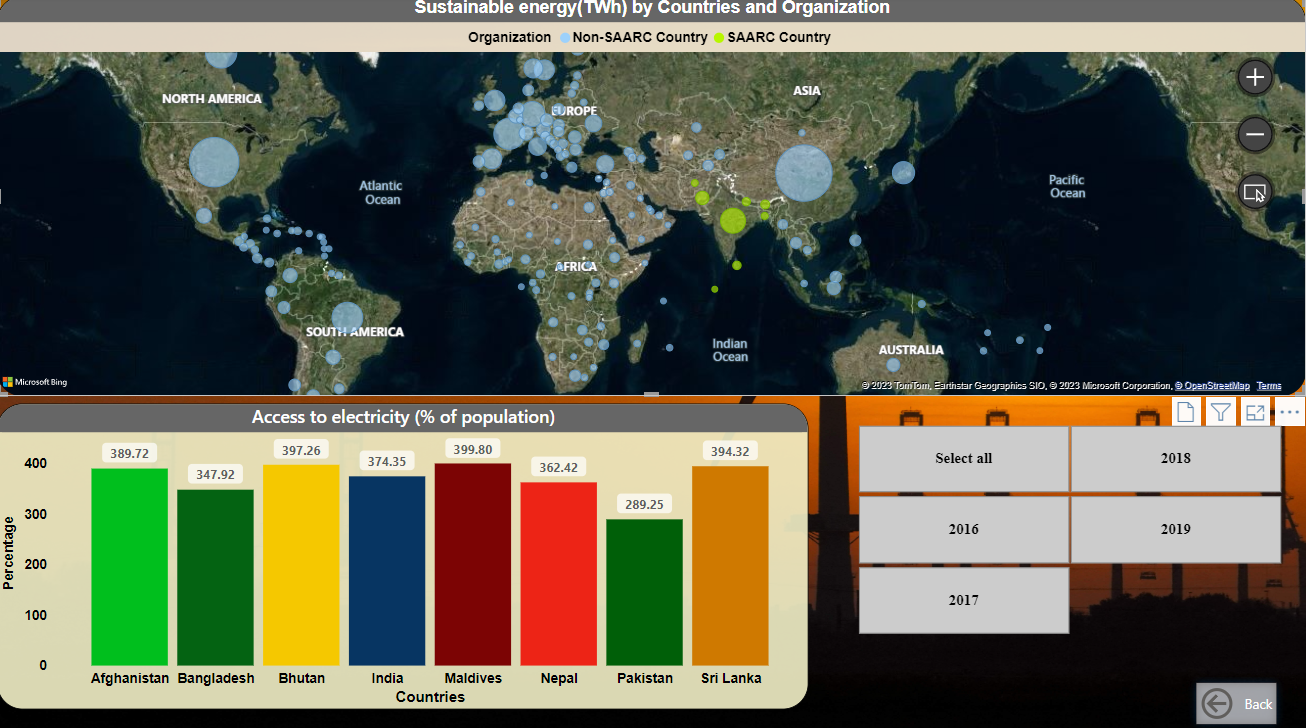


Figure 3: Dashboard page 3

In this page our aim is to use the newly constructed columns in our data transformation phase, sustainable energy column (nuclear sources + renewable energy sources). To display the country with the highest sustainable energy generation we used a world map to assign output units of (Twh) to a bubble, the bigger the bubble the countries generated more electricity from sustainable sources. To differentiate between the south Asian countries and the rest of the world we used organization as a legend to differentiate the bubble colours accordingly. Finally, we finished off our dashboard with the column chart to display the SAARC countries percentage of population with access to electricity,

## Dashboard conclusion

From this dashboard it is quite clear that form the global overview fossil fuel as a source still remains the main source of electricity generation, which highlights the urgency for a shift towards a renewable energy source for electricity generation, even with the domination of fossil fuels, which is bad for the environment, there has been a gradual increase in renewable energy generation from worldwide.

Form the SAARC Nation’s dashboard we can see, countries under the SAARC organization heavily relies on fossil fuel for their primary electricity generation and it is quite clear SAARC countries needs to bring up policies and sustainable strategic plans to improves the south Asian countries renewable energy generation and it also implies that SAARC nations needs to look for ways to balance the primary energy consumption vs renewable energy generation per capita to balance the consumption of electricity with sustainable energy.