**UNIVERSITY OF HERTFORDSHIRE**

APPLIED DATA SCIENCE-1 ASSIGNMENT-2

**Analysis of Electric Power Consumption and Population Growth: A Comparative Study of Iceland and Zambia**

|  |  |  |
| --- | --- | --- |
| Name | : | Sanjana Dabbikar |
| Student ID | : | 22084524 |
|  |  |  |

**Introduction**

This research delves into the statistical analysis and trends of Electric Power Consumption in relation to Population Growth, utilizing two distinct public datasets sourced from the World Bank. The study investigated correlations between electric energy consumption and population growth, focusing on Iceland and Zambia. This research also extended its scope to calculate the total power consumption of all countries from 1971 to 2015.

**Dataset**

This work focus on exploring statistics and trends Electric Power Consumption with respect to Population Growth. For this purpose, the 2 different public datasets i.e., Electric Power Consumption and Population Growth have been explored from World Bank. The URL Link for the 2 different datasets are shown below,

<https://data.worldbank.org/indicator/EG.USE.ELEC.KH.PC?end=2019&start=1960&view=chart>

<https://data.worldbank.org/indicator/SP.POP.GROW?view=chart>

**Ingesting & manipulating the data**

The Electric Power Consumption dataset has been ingested into the pandas dataframe by importing necessary data package such as pandas, numpy, etc., The dataset which is in the .csv file has been imported into the pandas dataframe. After importing the datasets, the Electric Power Consumption contains 266 rows and 65 columns and the Population Growth contains 266 rows and 64 columns.

**Electric Power Consumption Data Cleansing**

Null values – The dataset contains 116 rows and 20 columns contains null values and the same has been removed. After removing the null values, the dataset contains 150 rows and 45 columns. In the dataset, the country details are arranged in rows and year are arranged in columns and hence the dataset has to be transposed using transpose () function for converting the columns into rows and rows into columns.

**Exploring Statistical Analysis**

The Electric Power Consumption dataset is statistically analysed for analysing the data and this work will focus on cross-comparing the electric power consumption for the year 2000 with different countries. Based on the same, the research question and hypothesis has been framed as shown below.

**RQ-1:** Does Iceland consumes more Energy than Zambia?

**Null Hypothesis:** Iceland consumes more energy than Zambia.

**Alternate Hypothesis:** Iceland does not consume more energy than Zambia.

For exploring the statistical analysis, the matplotlib package has been imported for comparing the electric power consumption for different countries using bar plot for the year 2000 and 2014. Based on the exploration, it is inferred that, Iceland consumes more energy than Zambia for both the year 2000 and 2014. Hence accepting the Null Hypothesis and rejecting the Alternate Hypothesis.

**Understanding Correlations**

In the Electric Power Consumption and Population Growth dataset, the index has been set as ‘Country Name’ and then data has been selected from the period of 1971 to 2014. To explore and understand correlations, two indicators such as Power Consumption and Energy Consumption has been correlated using Pearson Correlation Co-Efficient. Based on the same, the research question and hypothesis has been framed as shown below.

**RQ-2:** Is there is a correlation between electric energy consumption and population growth in Iceland?

**Null Hypothesis:** There is a correlation between population growth and energy consumption in Iceland

**Alternate Hypothesis:** There is no correlation between population growth and energy consumption in Iceland.

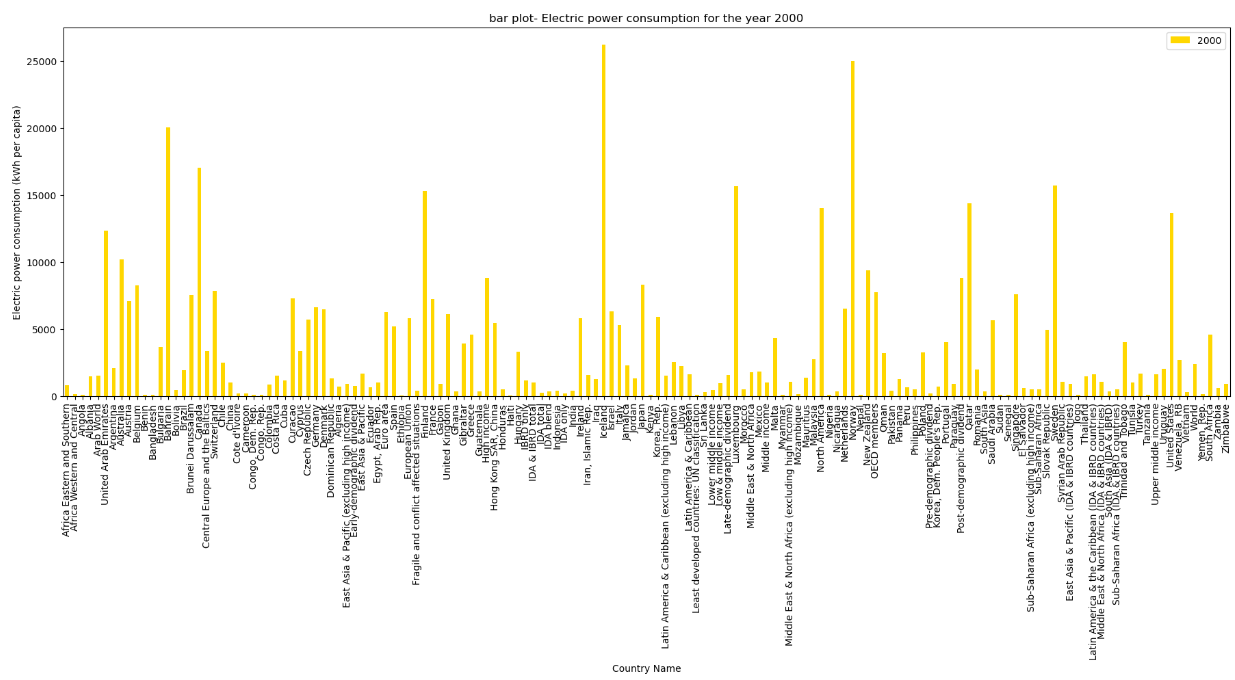
Using Pearson Correlation Co-Efficient, the correlation between electric energy consumption and population growth have been analysed in two different countries such as Iceland and Zambia.

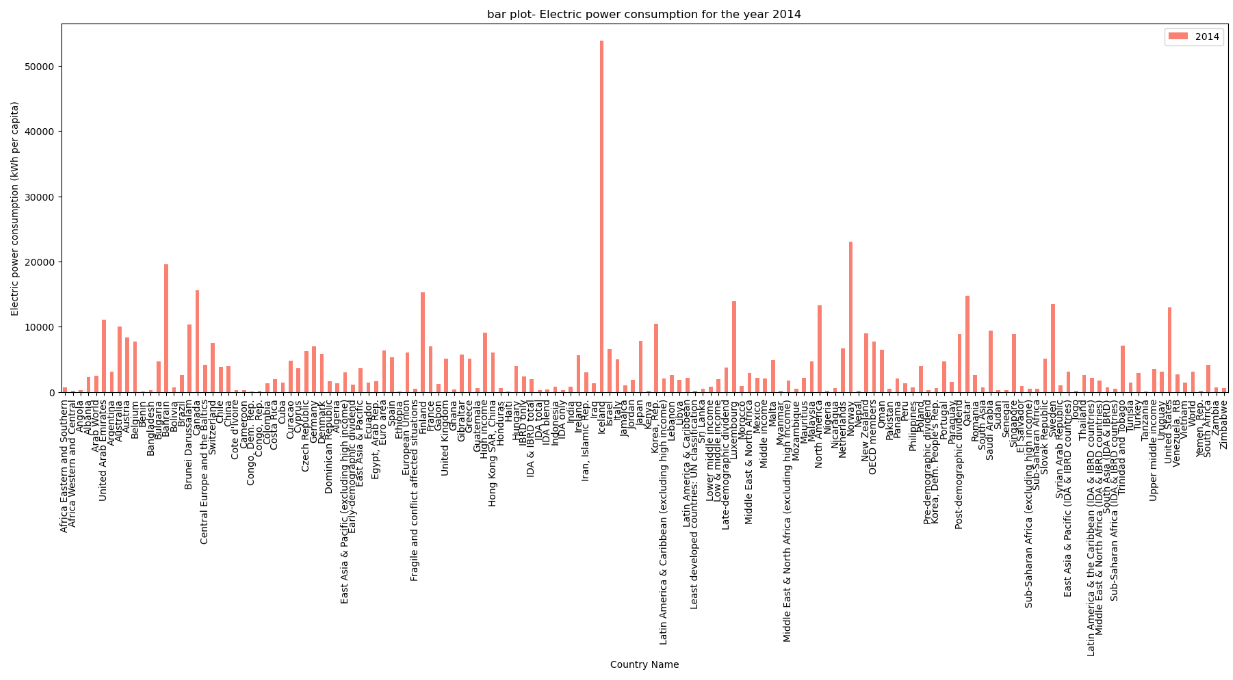
The Pearson Correlation value for Iceland country is -0.145127704602779 with a P-value of = 0.34724639981877736. Since the Pearson correlation value is -0.145 (Close to Zero). Hence accepting Alternate Hypothesis (There is no correlation between population growth and energy consumption in Iceland) and rejecting Null Hypothesis.

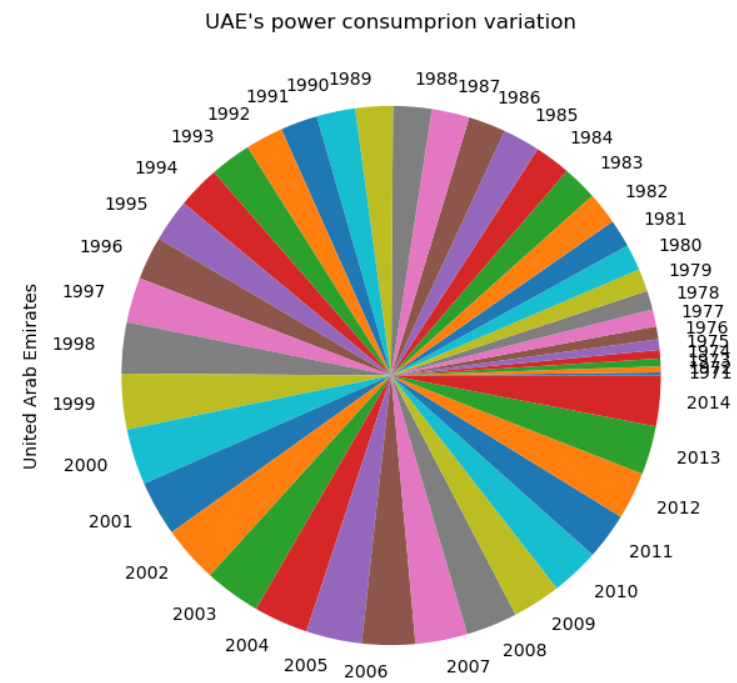
The Pearson Correlation value for Zambia country is 0.8418798187137676 with a P-value of = 8.085954566413688e-13. Since the Pearson correlation value is 0.84 (Close to One). Hence accepting Null Hypothesis (There is correlation between population growth and energy consumption in Zambia) and rejecting Alternate Hypothesis.

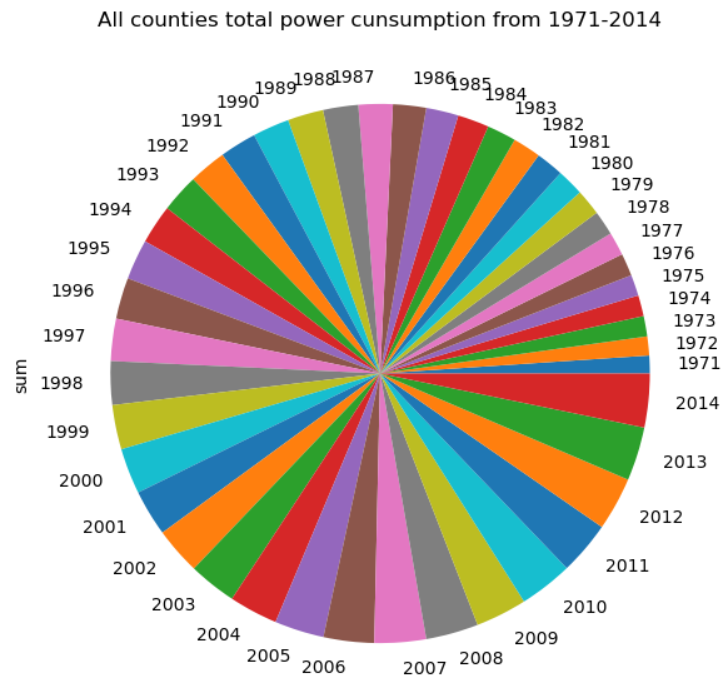
**Results & Interpretation**

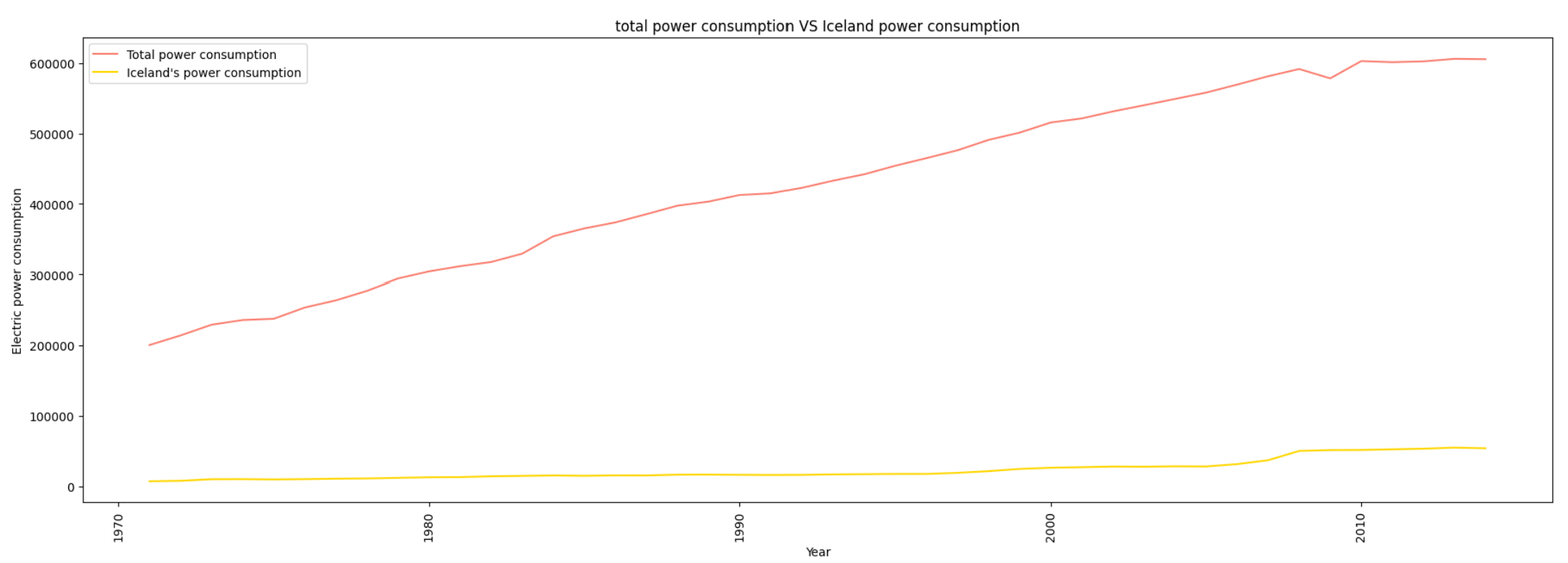
This work is further explored to analyse and calculate the total power consumption of all countries in the year 2000 and the second plot explains total power consumption of all countries in the year 2014. The next pie chart shows the power consumption of UAE’s 1971 to 2014 where the power consumption of the year 2014 is higher. The second pie chart shows the powder consumption sum of all the countries in every year with 2014 having the highest consumption. This work is further explored to analyse and calculate the total power consumption of all countries in the year 2000 over a period of time from 1971 to 2015. The coding for calculating the total power consumption of all countries is shown in the Appendix – D. Further this work has contributed in analysing and visualizing the total power consumption vs Iceland power consumption using time series plot and the visualization is as shown below.



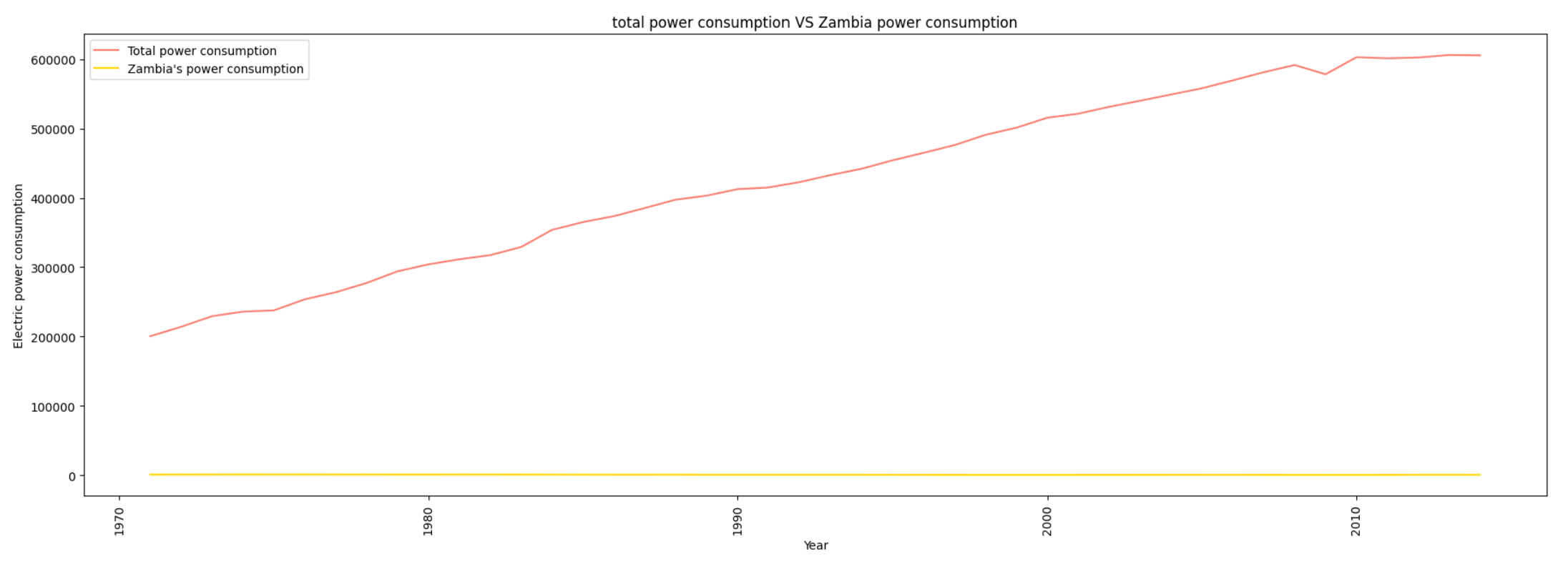








The visualization of total power consumption vs Zambia power consumption using time series plot and the visualization is as shown below.



**Key Findings**

* Through exploratory analysis using Bar plot, the Iceland consumes more energy than Zambia for both the year 2000 and 2014.
* The Pearson Correlation value for Iceland country is -0.145 (Close to Zero), which shows that there is no correlation between population growth and energy consumption in Iceland.
* The Pearson Correlation value for Zambia country is 0.84 (Close to One), which shows that there is correlation between population growth and energy consumption in Zambia.
* Using Time series plot, the total power consumption vs Iceland power consumption, the Iceland power consumption increases year after year.
* Using Time series plot, the total power consumption vs Zambia power consumption, there is no increase or decrease in Zambia power consumption even when there is change in period.