ENG 014: Impact of Population with Electricity Access on GDP Growth in Kenya

Kimaru Boruett

May 7, 2024

***Abstract*—This report investigates the impact of electricity access on GDP growth in Kenya from 2000 to 2020. Using a comprehensive dataset from Kaggle, the study focuses on the correlation between the percentage of the population with electricity access and GDP growth. The analysis employs Ordinary Least Squares (OLS) regression modeling and data visualization techniques to explore this relationship. Results indicate a strong positive correlation, with increased electrification aligning with significant economic growth. GDP per capita rose from about $400 in 2000 to nearly $1900 in 2020, as electricity access expanded from 17% to 70%. Predictive modeling suggests further economic benefits from continued improvements in electrification, although with notable uncertainties. The findings highlight the critical role of electricity access in fostering economic development while acknowledging the influence of other contributing factors. This study underscores the need for sustained investment in electrical infrastructure and further research to better understand the multifaceted dynamics of economic growth.**

I. Introduction

Electricity access is a critical component of modern economic development, acting as a catalyst for industrialization, technological advancement, and improved quality of life. In developing countries, such as Kenya, the provision of electricity to the population can significantly influence economic growth by enhancing productivity, enabling business operations, and improving educational and health outcomes. This lab report investigates the impact of electricity access on Gross Domestic Product (GDP) growth in Kenya, exploring the hypothesis that increased electrification correlates with higher economic performance.

Kenya, a lower-middle-income country in East Africa, has made substantial progress in expanding its electricity infrastructure over the past two decades. The Kenyan government, through initiatives such as the Kenya Vision 2030 and the Last Mile Connectivity Project, has prioritized expanding electricity access to rural and underserved areas. These efforts aim to support economic activities, reduce poverty, and achieve sustainable development goals. However, the extent to which electricity access directly influences GDP growth remains a subject of empirical investigation.

The relationship between electricity access and economic growth is multifaceted. Electricity can drive growth by powering industries, improving efficiency, and enabling innovation. For households, access to electricity can lead to better educational outcomes, increased income-generating opportunities, and improved health and living standards. Conversely, the lack of reliable electricity can constrain economic activities, limit educational opportunities, and perpetuate poverty cycles.

II. Methods

The dataset for this study was sourced from Kaggle's "Global Energy Trends 2000-2020," which includes comprehensive data on energy access and economic indicators across various countries. For this analysis, data specific to Kenya was extracted, focusing on GDP growth and the percentage of the population with electricity access. The dataset was cleaned and prepared using Pandas and NumPy, ensuring the accuracy and consistency of the data.

To assess the impact of electricity access on GDP growth, an Ordinary Least Squares (OLS) regression analysis was conducted using the statsmodels library in Python. The independent variable was the percentage of the population with electricity access, and the dependent variable was GDP growth. The model was fitted and analyzed to determine the relationship between these variables. Data visualization was performed using matplotlib and seaborn to create scatter plots with regression lines and line plots showing trends over time. These visualizations helped illustrate the correlation and trends, providing a clearer understanding of how electrification may influence economic growth in Kenya.

III. Results

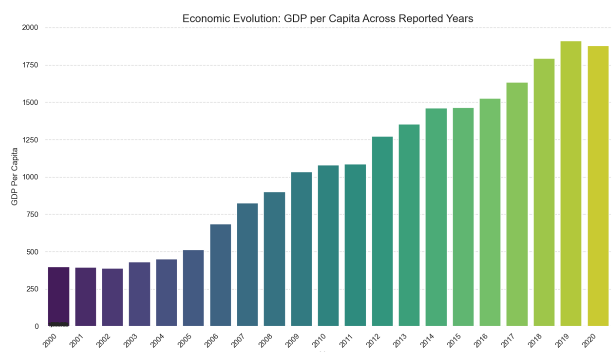


Fig. 1. Kenya GDP Growth reported from 2000 to 2020

Fig. 1 presents the GDP per capita for Kenya over the same period. The graph demonstrates a consistent increase in GDP per capita, from around $400 in 2000 to nearly $1900 in 2020. This growth aligns with the improvements in electricity access, suggesting a potential correlation between these two variables.

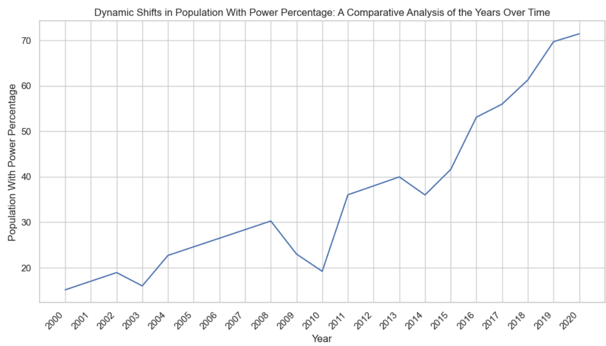


Fig. 2. Shifts in Population with Power Percentage from 2000 to 2020

Fig. 2 illustrates the percentage of the Kenyan population with access to electricity from 2000 to 2020. The data shows a significant upward trend, starting from about 17% in 2000 and reaching approximately 70% by 2020. This increase reflects the substantial efforts made by the Kenyan government and other stakeholders to expand electricity access over the years.

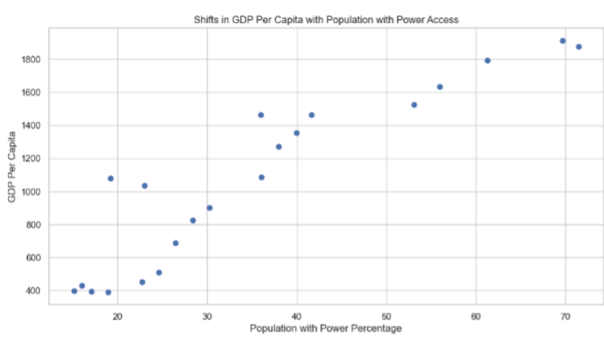


Fig. 3. Shifts in GDP Per Capita with respect to population with power percentage

Fig. 3 depicts a scatter plot of GDP per capita against the percentage of the population with electricity access. The positive correlation is evident, with GDP per capita rising as electricity access increases. This visualization supports the hypothesis that greater electricity access is associated with higher economic performance.

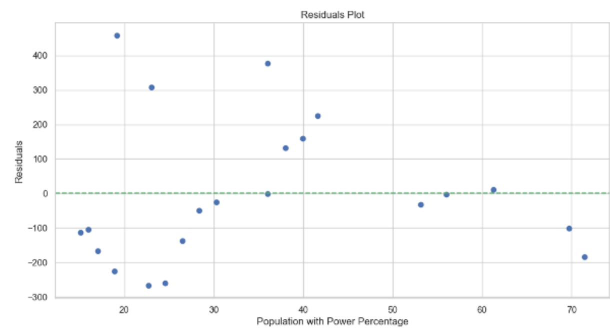


Fig. 4. Plot of Residuals following an OLS model fit on the data

Using the OLS regression model, we predicted GDP per capita for scenarios where 85% and 100% of the population have electricity access. The predictions were $2436.39 and $2849.91, respectively, although these estimates come with large confidence intervals due to being extrapolated from unseen data.

IV. Discussion

The analysis reveals a positive correlation between electricity access and GDP growth in Kenya. Over the past two decades, the proportion of the population with electricity access has increased markedly, from about 17% in 2000 to approximately 70% in 2020. Concurrently, GDP per capita has shown a steady rise, reaching nearly $1900 by 2020. These trends suggest that enhanced electrification has played a significant role in driving economic growth.

The scatter plot further reinforces this relationship, showing a clear upward trajectory in GDP per capita as electricity access increases. The OLS regression model predictions for scenarios with 85% and 100% electricity access, at $2436.39 and $2849.91 respectively, though accompanied by large confidence intervals, indicate potential for continued economic benefits from further electrification efforts. These findings align with the hypothesis that electricity access is a critical enabler of economic performance.

However, the large confidence intervals for the predictions highlight the need for caution in interpreting these extrapolations. They underscore the importance of gathering more empirical data to refine future estimates and enhance predictive accuracy. This suggests that while the trend is promising, policymakers and stakeholders should consider the uncertainties involved and continue to monitor and analyze the impacts of electrification on economic growth.

V. Conclusion

This study has established a strong positive correlation between electricity access and GDP growth in Kenya from 2000 to 2020. The significant increase in the percentage of the population with electricity access, from about 17% in 2000 to approximately 70% in 2020, appears to be closely linked with the rise in GDP per capita, which reached nearly $1900 by 2020. Our analysis, supported by visualizations and OLS regression modeling, suggests that enhanced electrification has played a significant role in driving economic growth. While the results are promising, it is crucial to recognize that other factors, such as government policies, international aid, infrastructure development, education improvements, and technological advancements, may also have contributed to the observed economic growth. Therefore, this study underscores the importance of sustained investment in electrical infrastructure as a strategy for economic development, while also highlighting the need for further research to fully understand the complex dynamics between electrification and economic performance.

References

[1] Global Data on Sustainable Energy (2000-2020). (2023, August 19). Kaggle. https://www.kaggle.com/datasets/anshtanwar/global-data-on-sustainable-energy/code

[2] Everbach et al. Lab 01 -Technical Writing Manual

VI. Appendix

A. Python Code Used for Analysis

The first document attached to the report

`

```
import pandas as pd
import numpy as np
import matplotlib as mpl
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.graph_objects as go
from mpl_toolkits.mplot3d import Axes3D
import statsmodels.api as sm

df = pd.read_csv('global-data-on-sustainable-energy (1).csv')

kenya_data = df[df['Entity'] == 'Kenya']
kenya_data.rename(columns={'Entity':'country_name',
'Year':'reported_year',
'Access to electricity (% of population)':'population_with_power_percentage',
'Access to clean fuels for
cooking':'population_with_cooking_fuel_percentage',
'Renewable-electricity-generating-capacity-per-capita':'renewable_capita',
'Financial flows to developing countries (US $)':'foreign_aid_dollar',
'Renewable energy share in the total final energy consumption
(%)':'renewable_percentage_in_total_consumption',
'Electricity from fossil fuels (TWh)':'fossil_power_generate_per_hour',
'Electricity from nuclear (TWh)':'nuclear_power_generate_per_hour',
'Electricity from renewables (TWh)':'renewable_power_generate_per_hour',
'Low-carbon electricity (% electricity)':'decarbonized_power_percentage',
'Primary energy consumption per capita (kWh/person)':'per_capita_power',
'Energy intensity level of primary energy (MJ/$2017 PPP GDP)':'power_to_gdp',
'Value_co2_emissions_kt_by_country':'co2_per_capita',
'Renewables (% equivalent primary energy)':'renewable_percentage',
'gdp_growth':'gdp_growth',
'gdp_per_capita':'gdp_per_capita',
'Density\\n(P/Km2)':'population_density',
'Land Area(Km2)':'country_area',
'Latitude':'country_latitude',
'Longitude':'country_longitude'}, inplace=True)
kenya_data.drop(columns=['population_with_cooking_fuel_percentage',
'foreign_aid_dollar', 'co2_per_capita', 'gdp_growth',
                             'population_density', 'country_area',
'country_latitude',
'country_longitude','renewable_percentage','nuclear_power_generate_per_hour']
, inplace=True)
kenya_data.describe().transpose()
kenya_data.loc[kenya_data['country_name'] == 'Kenya', 'country_name'] = 1
kenya_data.fillna(0, inplace=True)
```

```
/var/folders/h2/wt54cxk57ls_yd6srq2q5ztc0000gn/T/ipykernel_47615/786491104.py
:15: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame
```

`

```
See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  kenya_data.rename(columns={'Entity':'country_name',
/var/folders/h2/wt54cxk57ls_yd6srq2q5ztc0000gn/T/ipykernel_47615/786491104.py:36: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  kenya_data.drop(columns=['population_with_cooking_fuel_percentage',
'foreign_aid_dollar', 'co2_per_capita', 'gdp_growth',
/var/folders/h2/wt54cxk57ls_yd6srq2q5ztc0000gn/T/ipykernel_47615/786491104.py:40: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  kenya_data.fillna(0, inplace=True)
```

```
kenya_data.corr()
```

```
/var/folders/h2/wt54cxk57ls_yd6srq2q5ztc0000gn/T/ipykernel_47615/1483117826.py:1: FutureWarning: The default value of numeric_only in DataFrame.corr is deprecated. In a future version, it will default to False. Select only valid columns or specify the value of numeric_only to silence this warning.
  kenya_data.corr()

                                           reported_year  \
reported_year                                   1.000000
population_with_power_percentage                0.927035
renewable_capita                                0.831950
renewable_percentage_in_total_consumption      -0.555031
fossil_power_generate_per_hour                 -0.033429
renewable_power_generate_per_hour               0.949399
decarbonized_power_percentage                   0.700675
per_capita_power                                0.901341
power_to_gdp                                   -0.576934
gdp_per_capita                                  0.991008

                                           population_with_power_percentage
\
reported_year                                                      0.927035
population_with_power_percentage                                   1.000000
renewable_capita                                                   0.918468
renewable_percentage_in_total_consumption                         -0.640339
fossil_power_generate_per_hour                                    -0.220153
renewable_power_generate_per_hour                                  0.959917
decarbonized_power_percentage                                      0.740301
per_capita_power                                                   0.893229
power_to_gdp                                                      -0.656556
```

`

```
gdp_per_capita                                                  0.920508

                                            renewable_capita  \
reported_year                                       0.831950
population_with_power_percentage                    0.918468
renewable_capita                                    1.000000
renewable_percentage_in_total_consumption          -0.646471
fossil_power_generate_per_hour                     -0.415768
renewable_power_generate_per_hour                   0.933132
decarbonized_power_percentage                       0.754138
per_capita_power                                    0.901957
power_to_gdp                                       -0.614252
gdp_per_capita                                      0.842990


renewable_percentage_in_total_consumption  \
reported_year                                                              -
0.555031
population_with_power_percentage                                           -
0.640339
renewable_capita                                                           -
0.646471
renewable_percentage_in_total_consumption
1.000000
fossil_power_generate_per_hour
0.370184
renewable_power_generate_per_hour                                          -
0.629216
decarbonized_power_percentage                                              -
0.507894
per_capita_power                                                           -
0.486187
power_to_gdp
0.984857
gdp_per_capita                                                             -
0.543882

                                            fossil_power_generate_per_hour  \
reported_year                                                    -0.033429
population_with_power_percentage                                 -0.220153
renewable_capita                                                 -0.415768
renewable_percentage_in_total_consumption                         0.370184
fossil_power_generate_per_hour                                    1.000000
renewable_power_generate_per_hour                                -0.317263
decarbonized_power_percentage                                    -0.672004
per_capita_power                                                 -0.137320
power_to_gdp                                                      0.324683
gdp_per_capita                                                   -0.009278
```

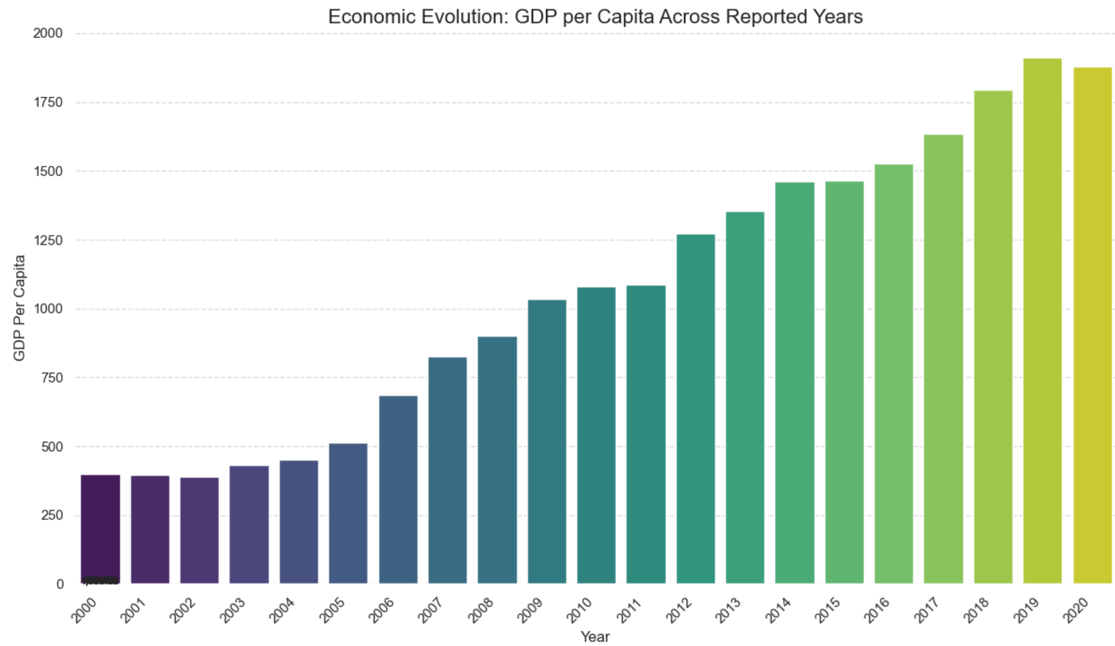
`

```
                                         renewable_power_generate_per_hour
\
reported_year                                                     0.949399
population_with_power_percentage                                  0.959917
renewable_capita                                                  0.933132
renewable_percentage_in_total_consumption                        -0.629216
fossil_power_generate_per_hour                                   -0.317263
renewable_power_generate_per_hour                                 1.000000
decarbonized_power_percentage                                     0.848969
per_capita_power                                                  0.918134
power_to_gdp                                                     -0.635507
gdp_per_capita                                                    0.936177

                                         decarbonized_power_percentage  \
reported_year                                                 0.700675
population_with_power_percentage                              0.740301
renewable_capita                                              0.754138
renewable_percentage_in_total_consumption                    -0.507894
fossil_power_generate_per_hour                               -0.672004
renewable_power_generate_per_hour                             0.848969
decarbonized_power_percentage                                 1.000000
per_capita_power                                              0.661255
power_to_gdp                                                 -0.515806
gdp_per_capita                                                0.655320

                                         per_capita_power  power_to_gdp  \
reported_year                                    0.901341     -0.576934
population_with_power_percentage                 0.893229     -0.656556
renewable_capita                                 0.901957     -0.614252
renewable_percentage_in_total_consumption       -0.486187      0.984857
fossil_power_generate_per_hour                  -0.137320      0.324683
renewable_power_generate_per_hour                0.918134     -0.635507
decarbonized_power_percentage                    0.661255     -0.515806
per_capita_power                                 1.000000     -0.462825
power_to_gdp                                    -0.462825      1.000000
gdp_per_capita                                   0.907692     -0.561351

                                         gdp_per_capita
reported_year                                  0.991008
population_with_power_percentage               0.920508
renewable_capita                               0.842990
renewable_percentage_in_total_consumption     -0.543882
fossil_power_generate_per_hour                -0.009278
renewable_power_generate_per_hour              0.936177
decarbonized_power_percentage                  0.655320
per_capita_power                               0.907692
power_to_gdp                                  -0.561351
gdp_per_capita                                 1.000000
```

`

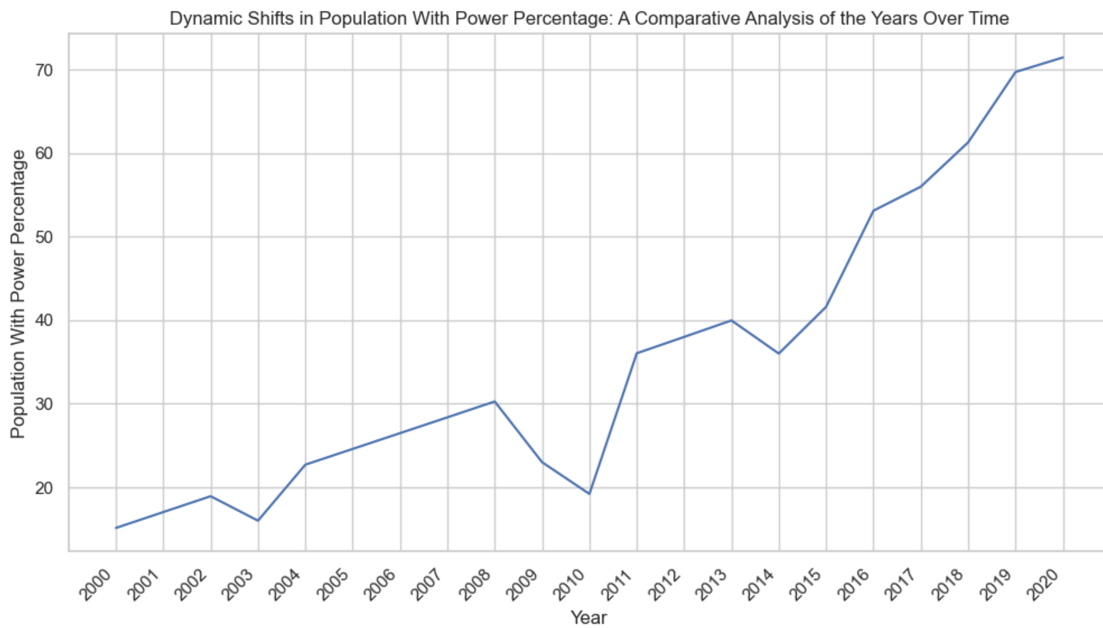
```
# Set the figure size
plt.figure(figsize=(15, 8))

# Create a bar plot with seaborn for a different column (e.g.,
'renewable_capita')
sns.barplot(x='reported_year', y='gdp_per_capita', data=kenya_data,
palette='viridis')

# Rotate x-axis labels for better readability
plt.xticks(rotation=45, ha='right')

# Add data labels on top of each bar
for p in plt.gca().patches:
    plt.gca().annotate(f"{p.get_height():,.2f}",
                       (p.get_x() + p.get_width() / 2., p.get_height()),
                       ha='center', va='center',
                       xytext=(0, 10),
                       fontsize=8)

# Add titles and axis labels
plt.title('Economic Evolution: GDP per Capita Across Reported Years',
fontsize=16)
plt.xlabel('Year', fontsize=12)
plt.ylabel('GDP Per Capita', fontsize=12)

# Add a grid for better readability
plt.grid(axis='y', linestyle='--', alpha=0.7)

# Customize the appearance of the plot
sns.despine(left=True, bottom=True)

# Display the plot
plt.show()
```

`

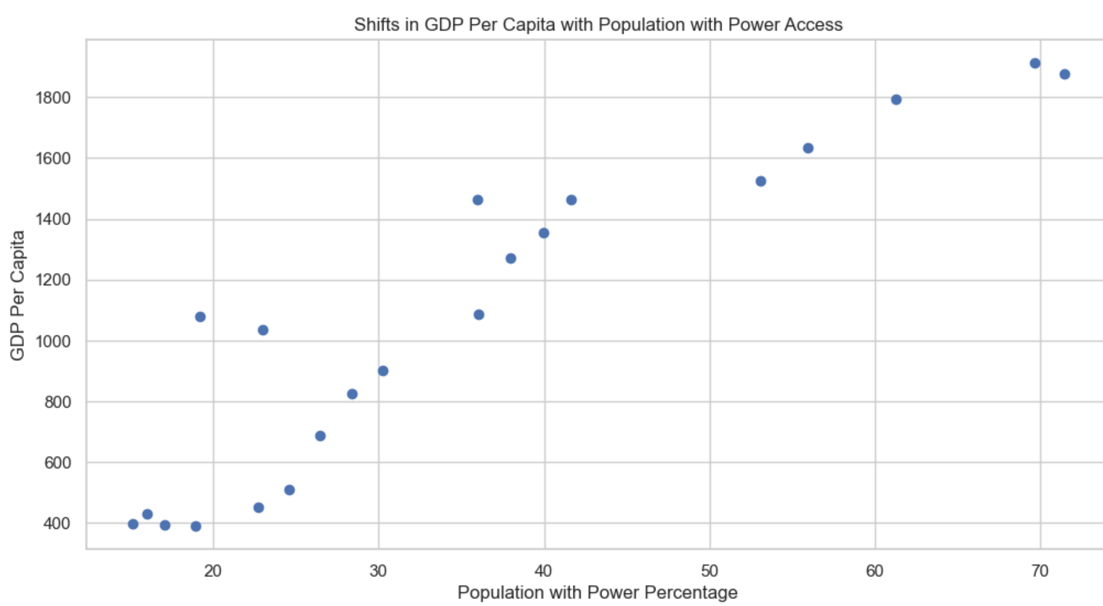



```
# Assuming 'input_dataframe' has a 'year' column
sns.set(style="whitegrid")
plt.figure(figsize=(12, 6))

# Calculate the mean renewable percentage for each country
mean_renewable =
kenya_data.groupby('reported_year')['population_with_power_percentage'].mean(
).sort_values(ascending=False).reset_index()

# Plot a line chart
sns.lineplot(x='reported_year', y='population_with_power_percentage',
data=mean_renewable)

plt.title('Dynamic Shifts in Population With Power Percentage: A Comparative
Analysis of the Years Over Time', fontsize=12)
plt.xlabel('Year', fontsize=12)
plt.ylabel('Population With Power Percentage', fontsize=12)

# Set the x-axis ticks to show all years as integers
plt.xticks(mean_renewable['reported_year'].unique(), rotation=45, ha='right')

# Uncomment the following line if you want a legend
# plt.legend(title='Country', bbox_to_anchor=(1.05, 1), loc='upper left')

plt.show()
```

`



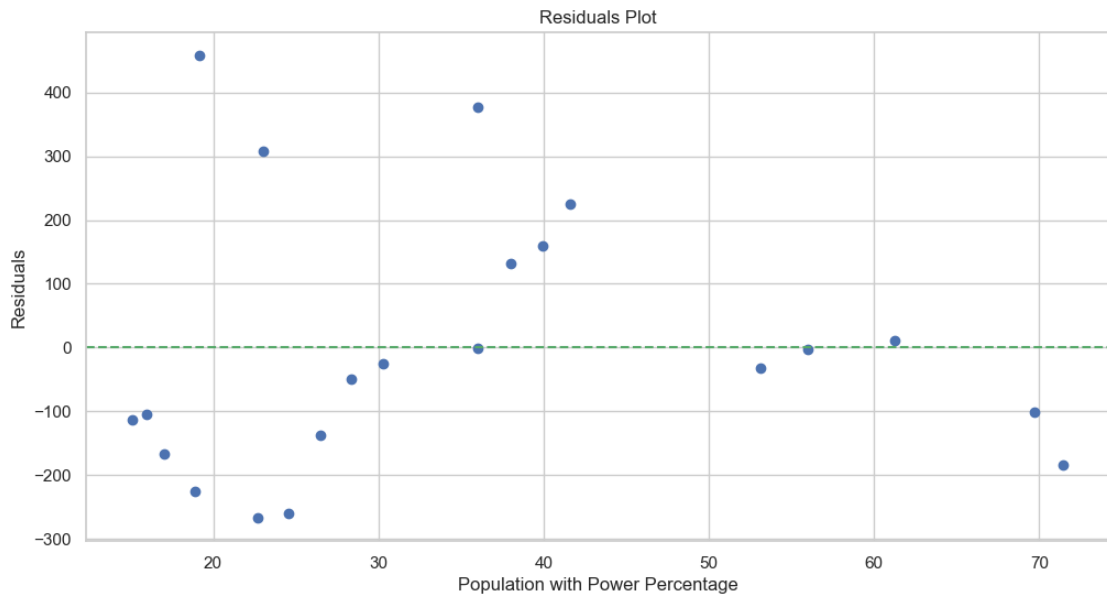

```
plt.figure(figsize=(12, 6))
pp_with_power = kenya_data["population_with_power_percentage"]
gdp_per_capita_pp = kenya_data["gdp_per_capita"]

plt.scatter(pp_with_power,gdp_per_capita_pp)
plt.xlabel("Population with Power Percentage")
plt.ylabel("GDP Per Capita")
plt.title("Shifts in GDP Per Capita with Population with Power Access")
plt.show()
```




```
plt.figure(figsize=(12, 6))
```

`

```
X = sm.add_constant(pp_with_power)
model = sm.OLS(gdp_per_capita_pp,X).fit()
# print(model.summary())

residuals = model.resid
plt.scatter(pp_with_power,residuals)
plt.xlabel("Population with Power Percentage")
plt.ylabel("Residuals")
plt.title("Residuals Plot")
plt.axhline(y=0,color='g',linestyle="--")
plt.show()
```




```
gdp_prediction = np.array([85,100])
X_new = sm.add_constant(gdp_prediction)
predictions = model.predict(X_new).round(4)
conf_int_predictions = model.get_prediction(X_new).conf_int().round(3)

print("\nPrediction for ",gdp_prediction[0] ," percent access:")
print("Predicted GDP Per Capita:",predictions[0])
print("95% Confidence Interval: [",conf_int_predictions[0][0], ",",
conf_int_predictions[0][1],"]")



print("\nPrediction for ",gdp_prediction[1] ," percent access:")
print("Predicted GDP Per Capita:",predictions[1])
print("95% Confidence Interval: [",conf_int_predictions[1][0], ",",
conf_int_predictions[1][1],"]")
```

```
Prediction for  85  percent access:
```

`

```
Predicted GDP Per Capita: 2436.3924
95% Confidence Interval: [ 2141.817 , 2730.968 ]

Prediction for  100  percent access:
Predicted GDP Per Capita: 2849.9111
95% Confidence Interval: [ 2474.687 , 3225.135 ]
```

```
#Performing Model Selection to Optimize Results
```