DATA STORYTELLING IN PATTERN RECOGNITION

African Countries Food Production and Supply (2004-2013) Data Exploration

BY OLUWATOMISIN EBUN AROKODARE

PROJECT OVERVIEW

Food production and supply play an important role in providing nutrition and promoting economic growth in Africa. Food security includes the availability, usability, stability, and affordability of food for both people and communities. Human well-being, societal stability, poverty reduction, and sustainable development all rely on food security.

In this project, I will be analyzing two different datasets namely food production and food supply datasets. The dataset was sourced from the Food and Agriculture Organization of United Nations website. The two datasets have information about food production and food supply in 45 African countries between (2004 - 2013) respectively. The Population and Country code dataset was sourced from Kaggle.

Project Goals

The goal of this project is to analyze these datasets, and gain insights on African food shortage problems, and producing visuals that can help in making appropriate decisions and ensuring food security in Africa.

Dataset Overview

- The food production dataset covers 45 African countries, different food items (rice, cassava, wheat, beverage, oats etc.) and years ranging from 2004 to 2013.The Supply dataset also covers 45 African Countries, year ranging from (2004-2013) and Values.

```
In [2]: # loading the food production dataset into a pandas dataframe
        af_prod= pd.read_csv('Africa Food Production (2004 - 2013).csv')
        af_prod
```

Out[2]:

	Country	Item	Year	Value
0	Algeria	Wheat and products	2004	2731
1	Algeria	Wheat and products	2005	2415
2	Algeria	Wheat and products	2006	2688
3	Algeria	Wheat and products	2007	2319
4	Algeria	Wheat and products	2008	1111
...	...	...	...	...
23105	Zimbabwe	Crustaceans	2009	0
23106	Zimbabwe	Crustaceans	2010	0
23107	Zimbabwe	Crustaceans	2011	0
23108	Zimbabwe	Crustaceans	2012	0
23109	Zimbabwe	Crustaceans	2013	0

```
In [15]: # loading the food supply dataset into a pandas dataframe
         af_supply = pd.read_csv('Africa Food Supply (2004 - 2013).csv')
         af_supply
```

Out[15]:

	Country	Year	Value
0	Algeria	2004	2987
1	Algeria	2005	2958
2	Algeria	2006	3047
3	Algeria	2007	3041
4	Algeria	2008	3048
...	...	...	...
445	Zimbabwe	2009	2147
446	Zimbabwe	2010	2168
447	Zimbabwe	2011	2200
448	Zimbabwe	2012	2197
449	Zimbabwe	2013	2110

450 rows × 3 columns

- The Country code dataset consist of countries, GDP(Billions) and Country Code

```
In [83]: # Importing the Country Code dataset

         country_codes = pd.read_csv('2014 World_ GDP_with_Country_Codes.csv')
         country_codes.head()
```

Out[83]:

	COUNTRY	GDP (BILLIONS)	CODE
0	Afghanistan	21.71	AFG
1	Albania	13.40	ALB
2	Algeria	227.80	DZA
3	American Samoa	0.75	ASM
4	Andorra	4.80	AND

- The population dataset consists of country, year, and count of Population.

```
In [89]: # Renaming the country column in the population data
         pop_df.rename(columns = {"Country Name":"Country", "Count": "Population"}, inplace=True)
         pop_df
```

Out[89]:

	Country	Year	Population
0	Aruba	1960	54211
1	Afghanistan	1960	8996973
2	Angola	1960	5454933
3	Albania	1960	1608800
4	Andorra	1960	13411
...	...	...	...
12590	Kosovo	2017	1830700
12591	Yemen, Rep.	2017	27834821
12592	South Africa	2017	57000451
12593	Zambia	2017	16853688
12594	Zimbabwe	2017	14236745

12595 rows × 3 columns

Data Cleaning

1. There are no null values in the Food Production and supply in the dataset (2004-2013)

```
In [6]: ### checking for missing values in the dataframe
        af_prod.isnull().sum()
```

```
Out[6]: Country    0
        Item       0
        Year       0
        Value      0
        dtype: int64
```

```
In [19]: ### checking for missing values in the dataframe
         af_supply.isnull().sum()
```

```
Out[19]: Country    0
         Year       0
         Value      0
         dtype: int64
```

The Final Data frame for Food Production and Food Supply Dataset

```
In [11]: af_prod.head(10)
```
Out[11]:

	Country	Item	Year	Value	Unit
0	Algeria	Wheat and products	2004	2731	kt
1	Algeria	Wheat and products	2005	2415	kt
2	Algeria	Wheat and products	2006	2688	kt
3	Algeria	Wheat and products	2007	2319	kt
4	Algeria	Wheat and products	2008	1111	kt
5	Algeria	Wheat and products	2009	2953	kt
6	Algeria	Wheat and products	2010	2605	kt
7	Algeria	Wheat and products	2011	2555	kt
8	Algeria	Wheat and products	2012	3432	kt
9	Algeria	Wheat and products	2013	3299	kt

```
In [24]: af_supply.head(10)
```
Out[24]:

	Country	Year	Value	Unit
0	Algeria	2004	2987	kcal/person/day
1	Algeria	2005	2958	kcal/person/day
2	Algeria	2006	3047	kcal/person/day
3	Algeria	2007	3041	kcal/person/day
4	Algeria	2008	3048	kcal/person/day
5	Algeria	2009	3110	kcal/person/day
6	Algeria	2010	3142	kcal/person/day
7	Algeria	2011	3217	kcal/person/day
8	Algeria	2012	3272	kcal/person/day
9	Algeria	2013	3296	kcal/person/day

Dataset Features Description

Food Production Dataset:

1.Country: The name of the country.

2.Item: The name of the produce from each country.

3.Year: The year of production.

4.Value: The quantity of each item, measured in kilotons (kt).

Food Supply Dataset:

1.Country: The name of the country.

2.Year: The year of supply.

3.Value: The quantity supplied, measured in kilocalories

per person per day(kcal/person/day).

Exploratory Data Analysis

QUESTION 1. What is the Trend in food production over the years.

Observation: Food production in Africa increases yearly. However, apart from the drop in 2007, there was a progress increase in food production from 2008 to 2013.


Food Production Trends in Africa (2004 - 2013)

Food Production trends in Africa (2004 -2013)
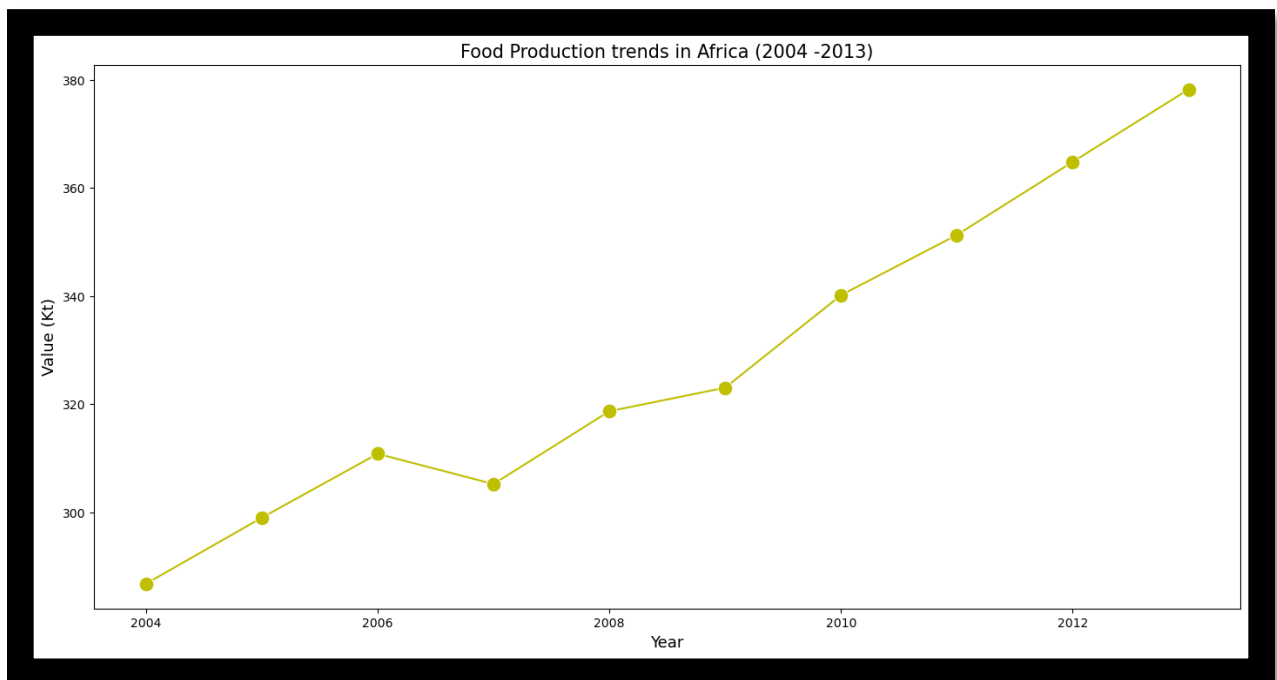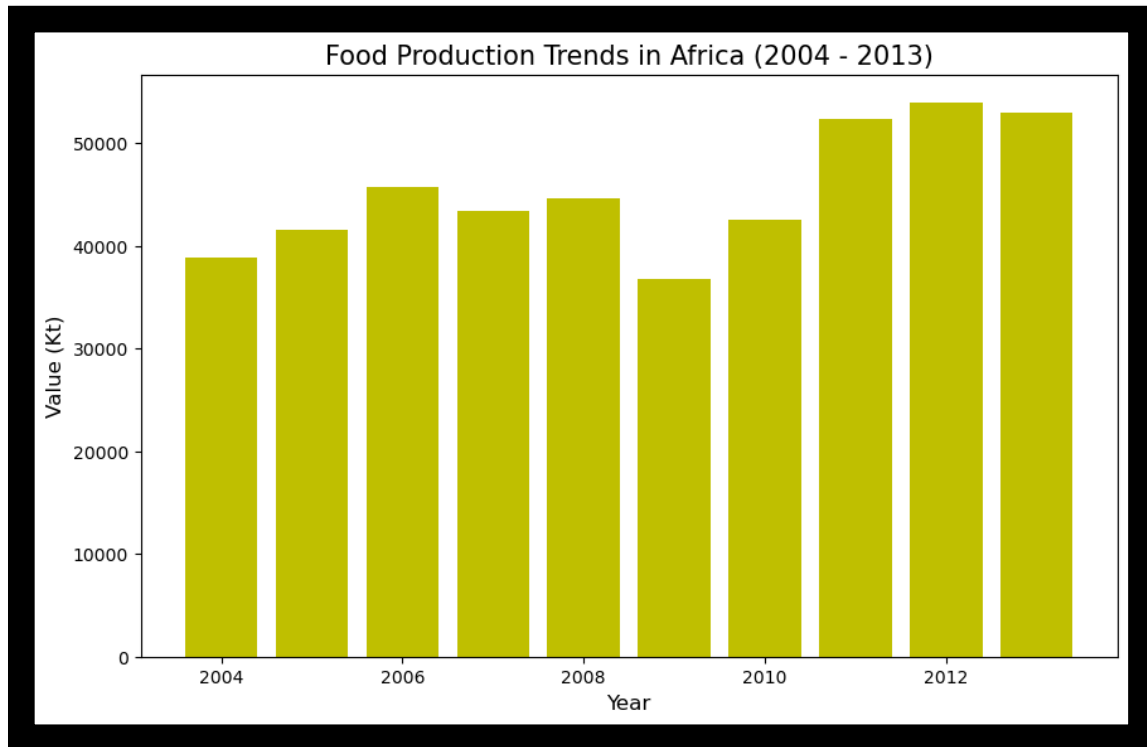
QUESTION 2. What is the Trend in food supply over the years?

**Observation:** There is a progressive increase in food supply in African countries from 2004 to 2013.

Food supply Trends in Africa (2004 - 2013)

Food Supply trends in Africa (2004 -2013)
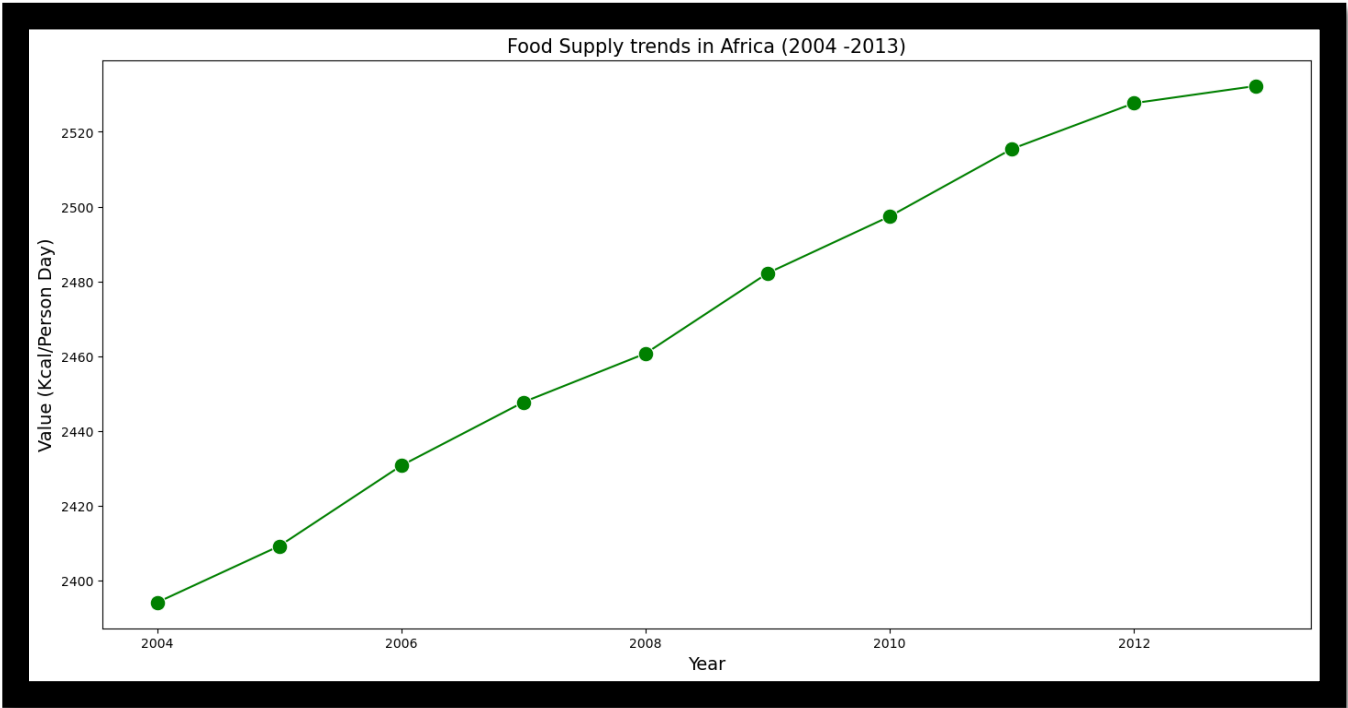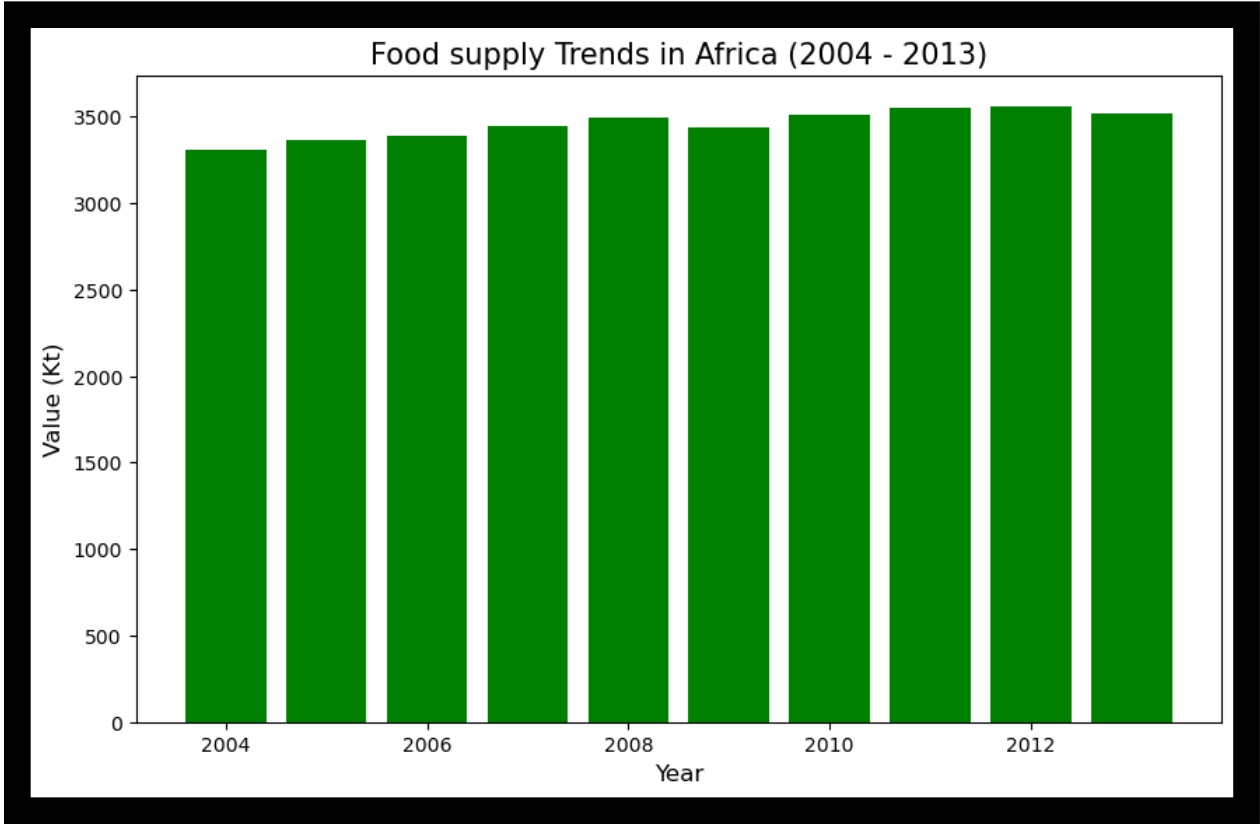
QUESTION 3. Which countries belong to the Top and Bottom 3 with the highest and Least volume of food production (2004-2013)?

Observation: The **TOP 3** African countries with the highest volume of food production within (2004-2013) are Nigeria, Egypt, & South Africa. However, the **BOTTOM 3** African countries Cabo Verde, Sao Tome & Principe and Djibouti produced the least volume of food.

Volume of Food Produced by African Countries (2004 - 2013)
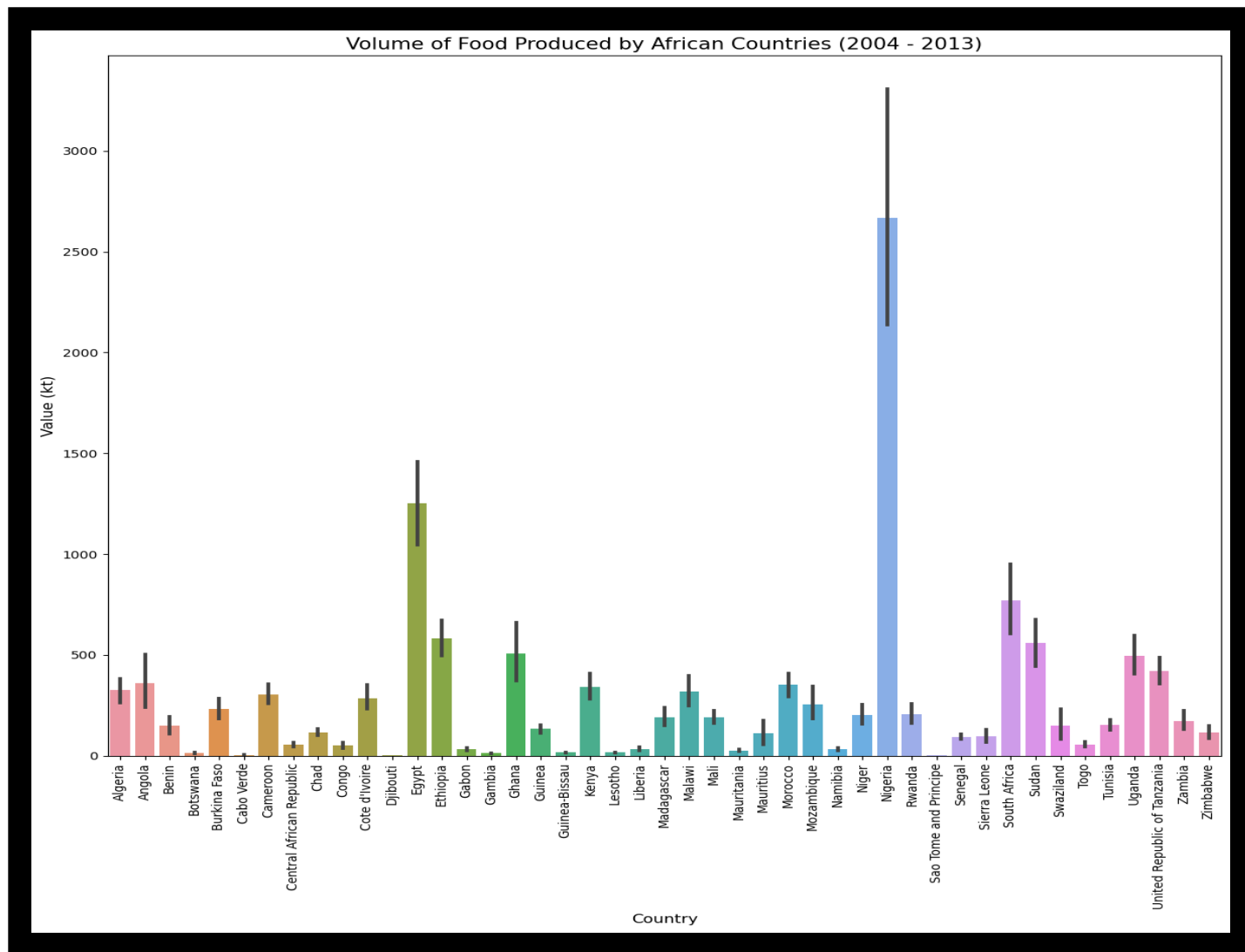
QUESTION 4. Which Countries belong to the Top and Bottom 3 with the highest and Least volume of food Supply (2004-2013)?

Observation: The TOP 3 African countries with the highest volume of food supply within (2004-2013) are Egypt, Tunisia, & Morocco. However, the BOTTOM 3 African countries Chad, Ethiopia and Zambia supplied the least volume of food.

Volume of Food Supplied by African Countries (2004 - 2013)
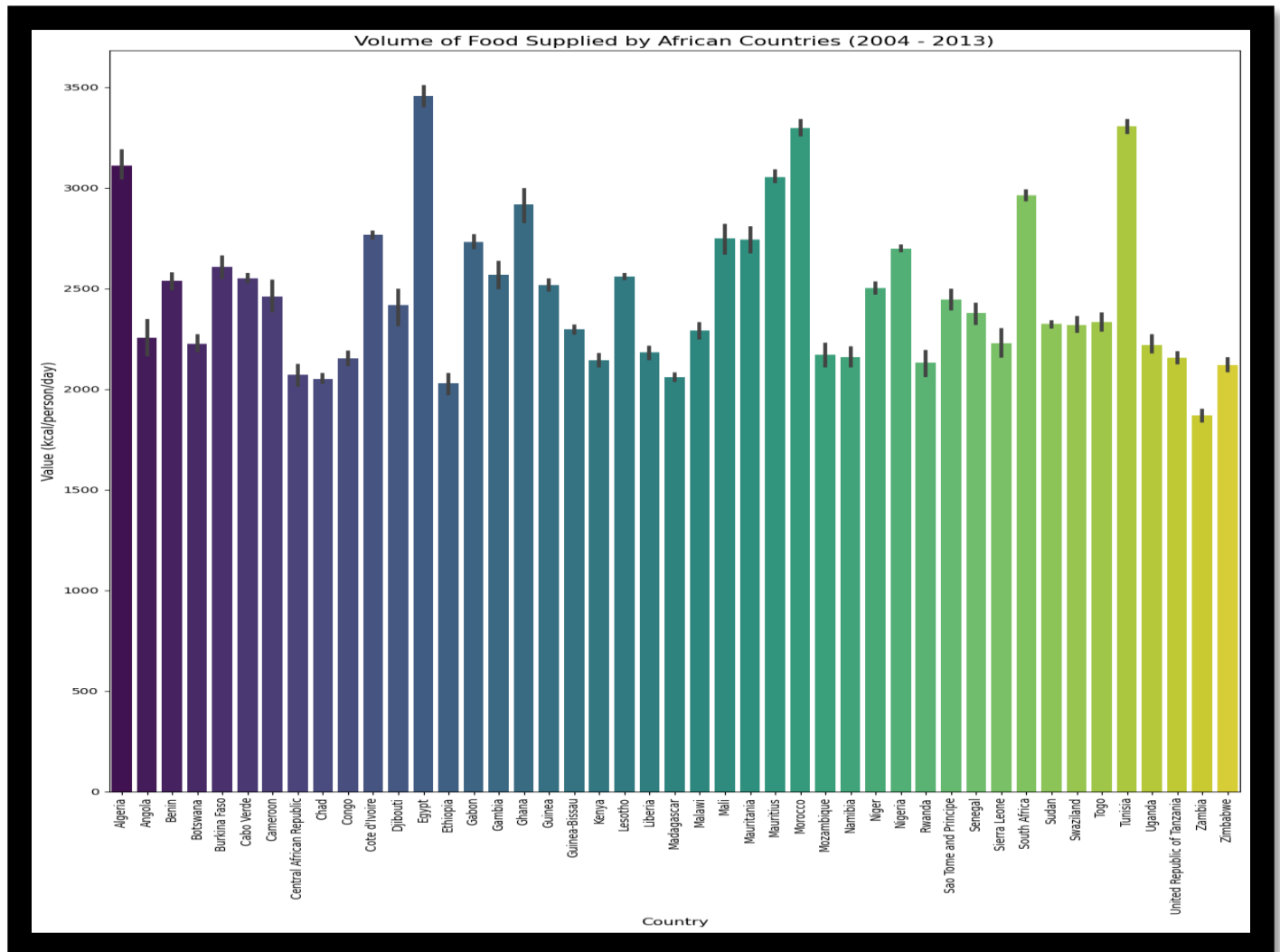
QUESTION 5. What Food items was produced the most by African countries within (2004-2013)

Observation: The Top 5 most produced food in Africa between 2004 and 2013 are Cassava, Sugar Cane, Maize and products, Yam, and Vegetables.
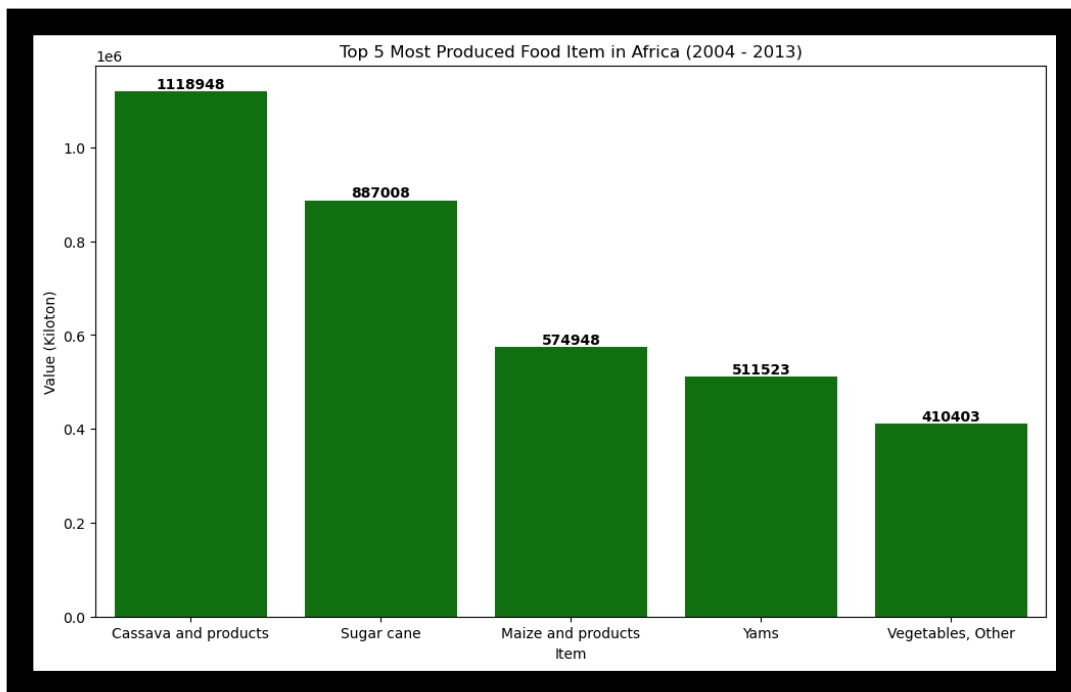
QUESTION 6. What Food items was least produced the most by African countries within (2004-2013)?

Observation: The Least 5 most produced food in Africa between 2004 and 2013 are Cloves, Molluscs,Pepper, Fish, Liver oil and Aquatic Animals
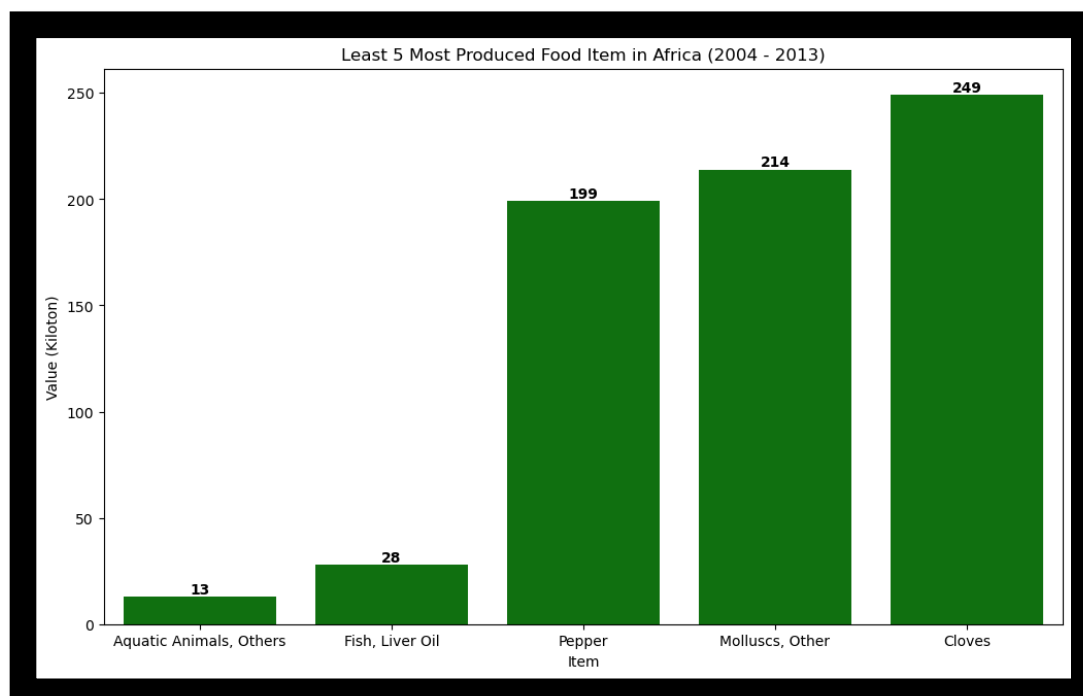
QUESTION 7. What are the most popular Food Items that the Top 3 producing Africa Countries Produced?

Observation: Nigeria highest food produces are Cassava and its products & Yams.
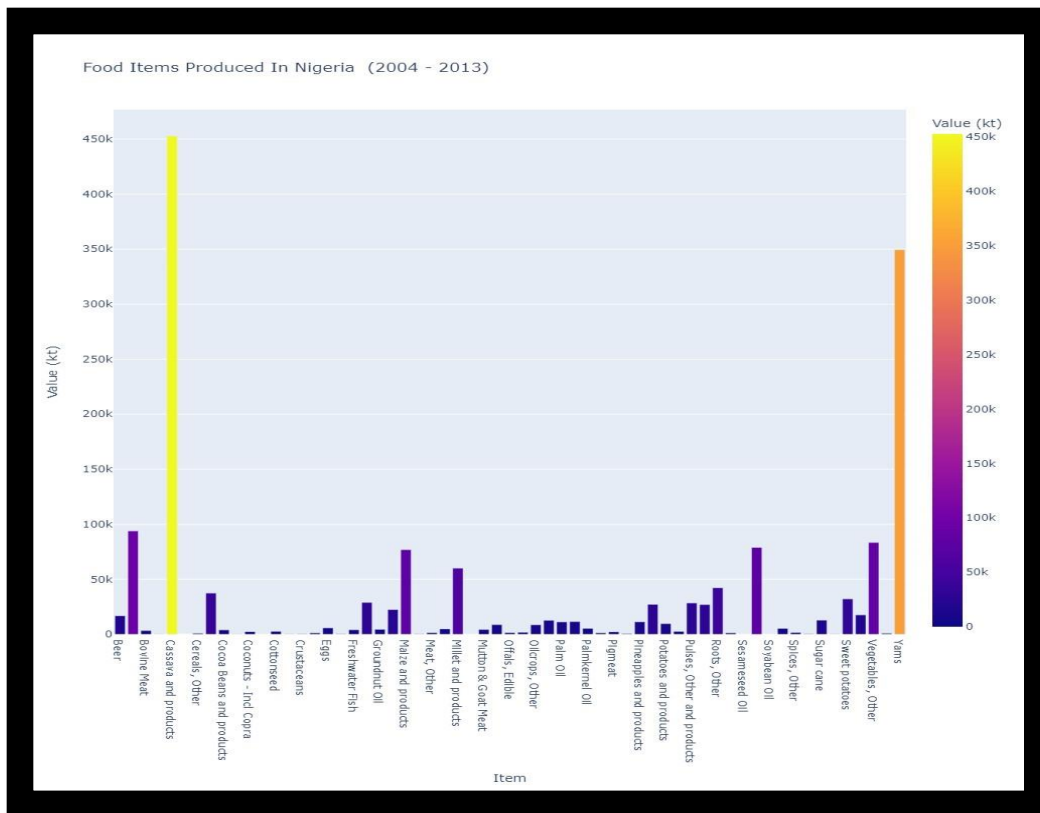
Observation: Egypt highest food produces is Sugarcane although there is a relatively fair spread in their food production including Products like Maize & Products, Tomatoes, Vegetables, Wheat, Milk (excluding butter) unlike Nigeria.
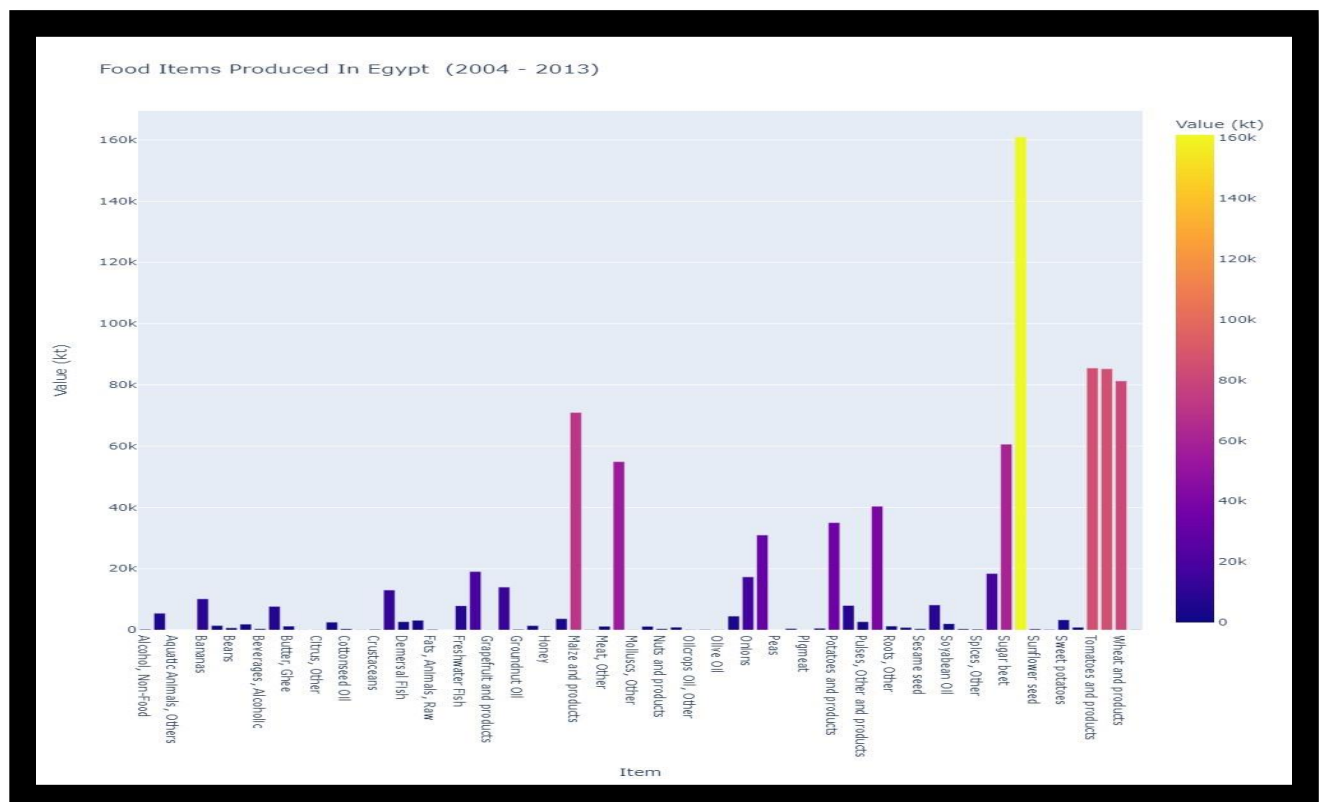
Observation: South Africa highest food produces is Sugarcane and Maize.


Food Items Produced In South Africa (2004 – 2013)
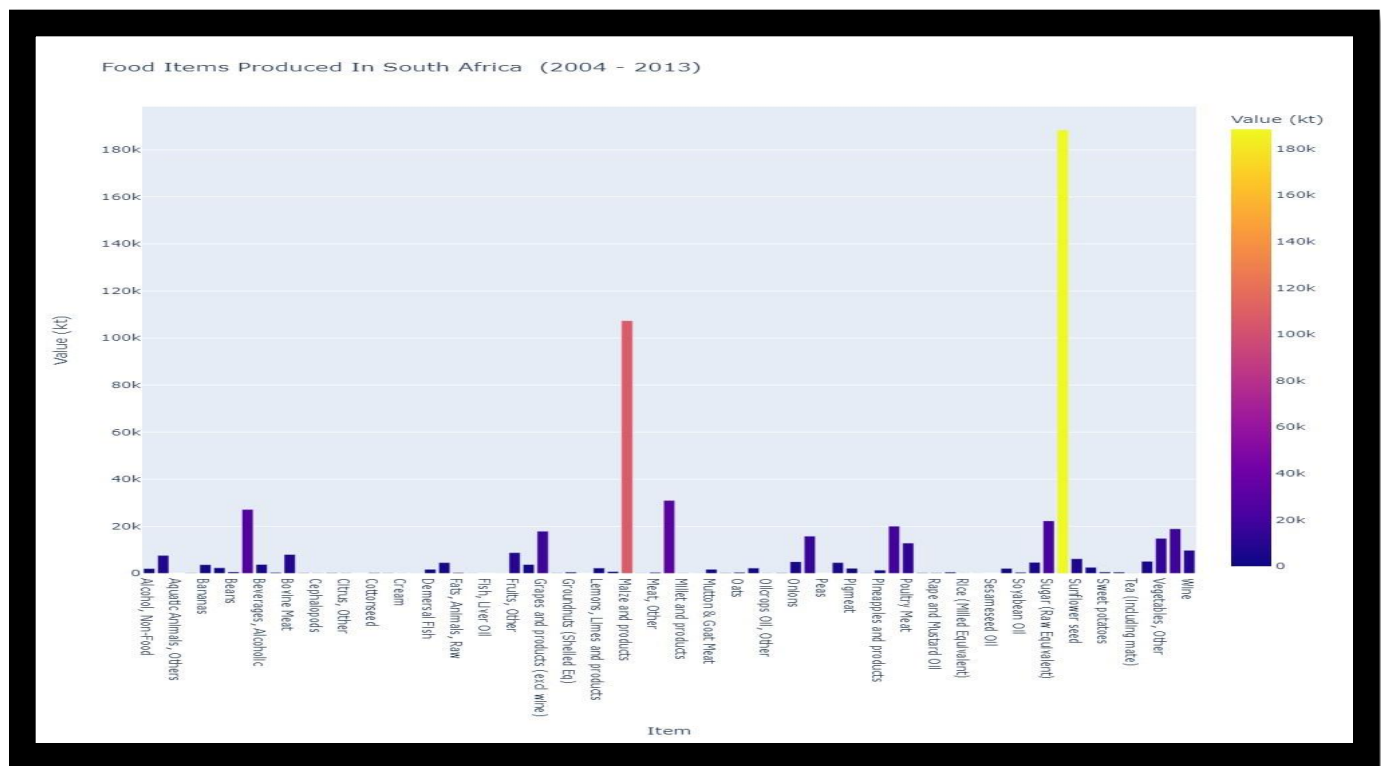
QUESTION 8. Are there any outlier countries in the food production dataset between 2004 and 2013?

Observation:39 countries in total produced consistently high amount of food between 2004 and 2013 and they are regarded as outlier countries. This Country include (Algeria, Angola, Benin, Burkina Faso, Cameroon, Central African Republic, Chad, Congo, Cote d'Ivoire, Egypt, Ethiopia, Gabon, Ghana, Guinea, Guinea-Bissau, Kenya, Liberia, Madagascar, Malawi, Mali, Mauritania, Mauritius, Morocco, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Sierra Leone, South Africa, Sudan, Swaziland, Togo, Tunisia, Uganda, United Republic of Tanzania, Zambia, and Zimbabwe).


Food Producing Countries with Outliers
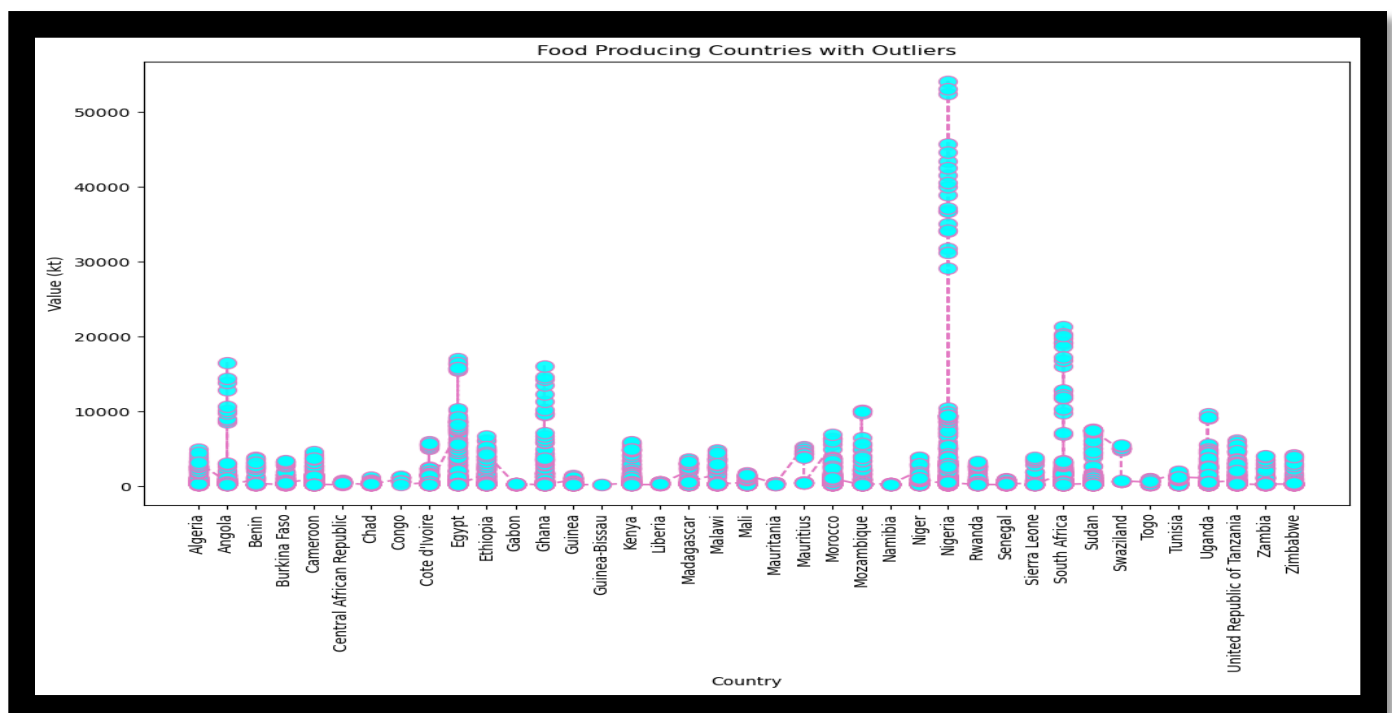
QUESTION 9. Are there any outlier countries in the food Supply dataset between 2004 and 2013?

Observation: Egypt is the African country with highest food supply outliers because the supply value is higher than the upper fence value. It is also evident from the Visualization that Egypt has a substantially bigger food supply accessible for consumption in 2012 compared to other African countries.
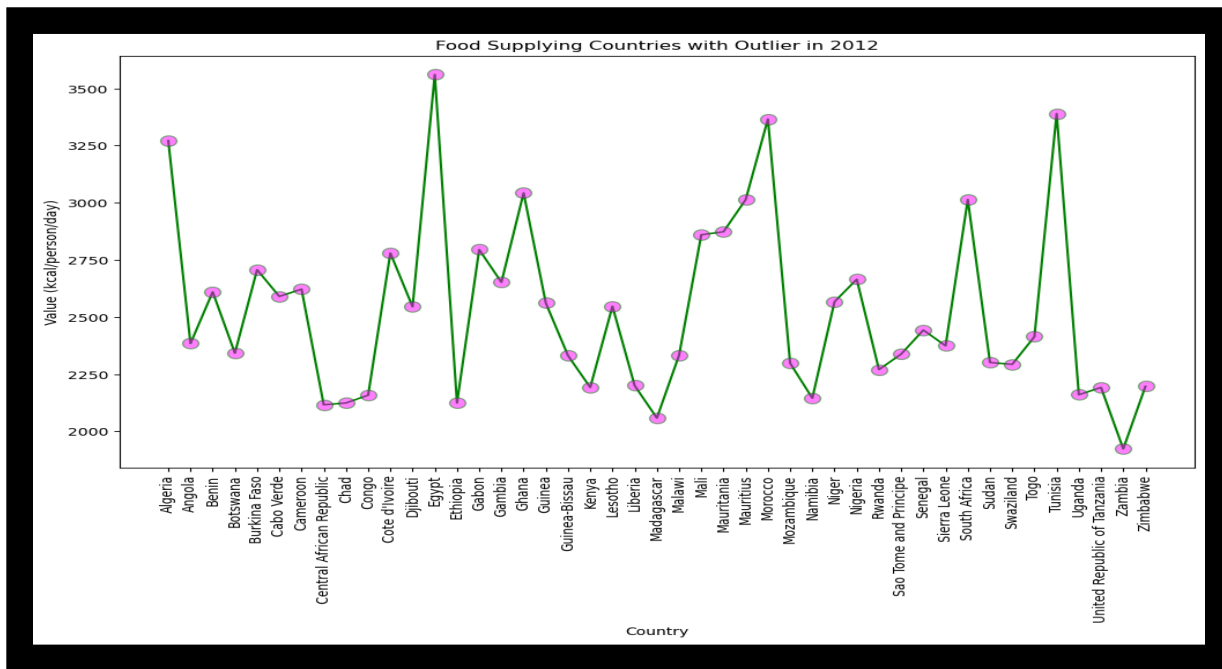

QUESTION 10. Is there any relationship between Average Food Production and Food Supply per year?

Observation: In 2007, the average food production dropped to 305.2158kt and the average food supplied was 2447.75kcal/person/day this means that more food was supplied to meet the need of the population. This situation highlights the fact that food production alone does not determine the availability of food for a population. There is a need for African Countries to improve their food productions capacities and reduce dependence on external sources of food.
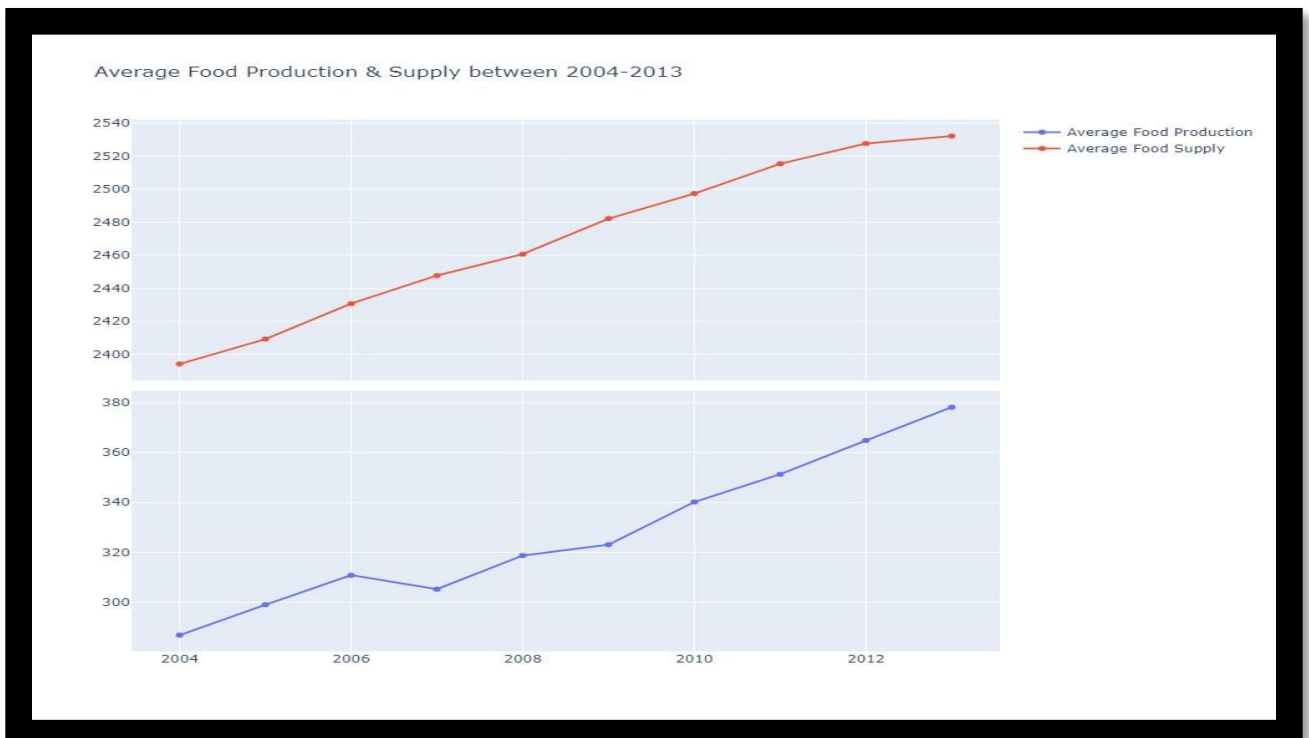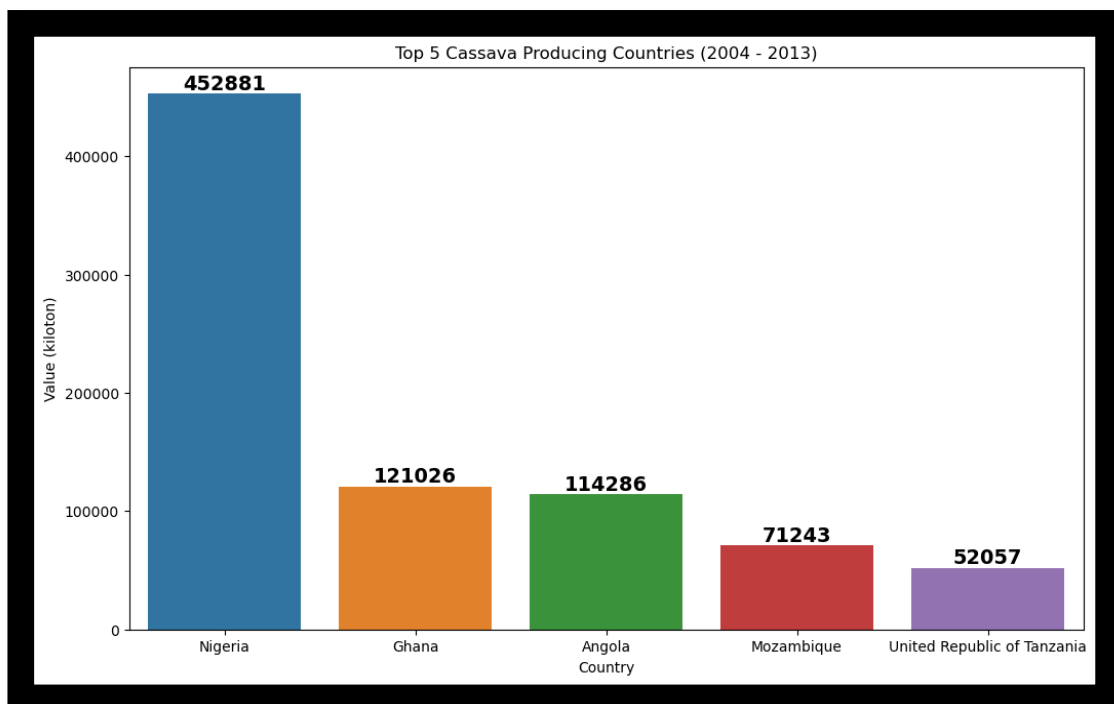
QUESTION 11. Which African countries belong to the Top 5 OATS Producers?

Observation: The Top 5 Cassava Producing Countries that has the highest total quantity of Oats from 2004 to 2013 are: Algeria, Ethiopia, South Africa, Morrocco, and Kenya. However, Algeria is the top leading producer of Oat products.
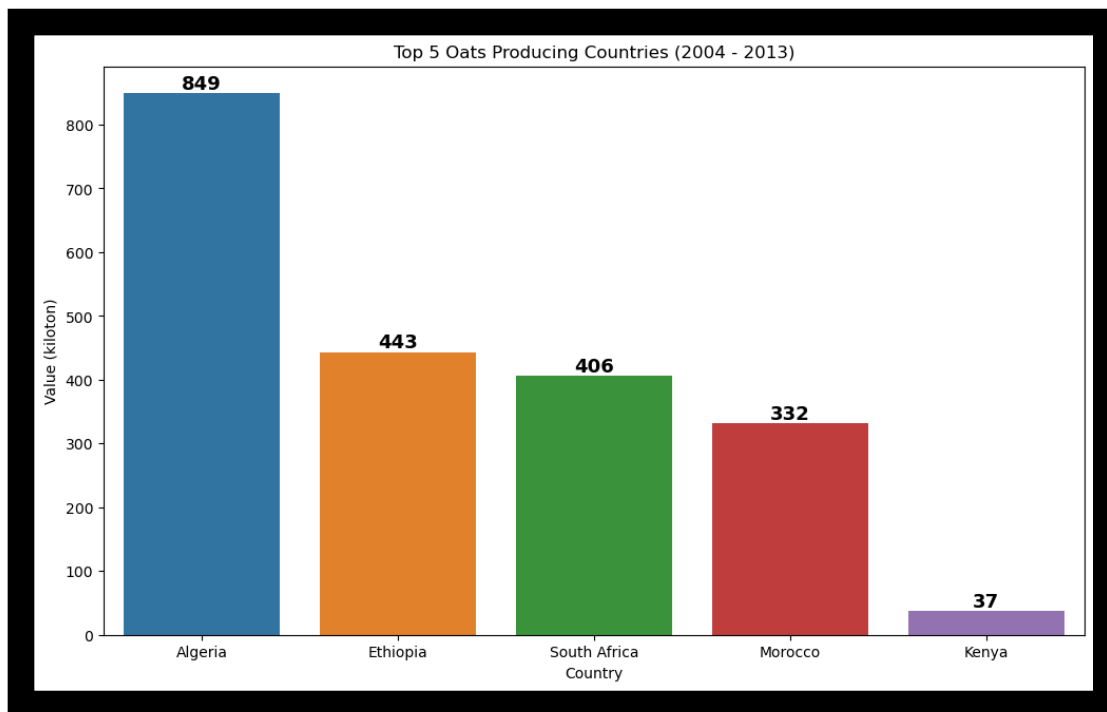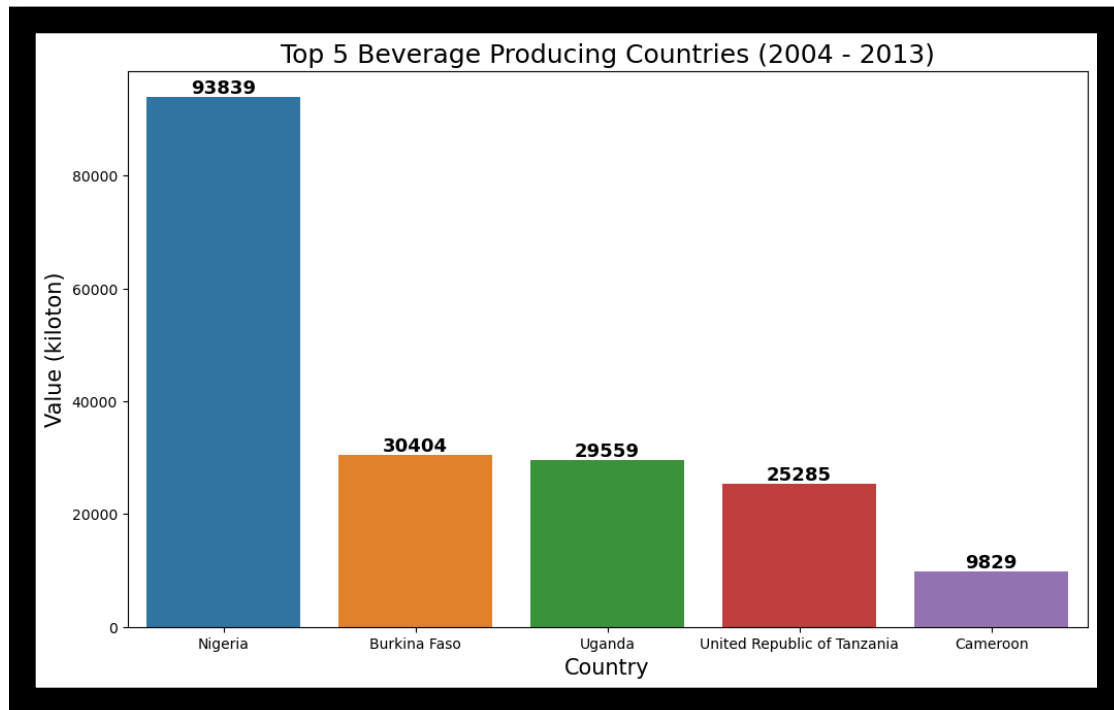

QUESTION 12. Which African countries belong to the Top 5 Cassava Producers?

Observation: The Top 5 Cassava Producing Countries that has the highest total quantity of Cassava from 2004 to 2013 are: Nigeria, Ghana, Angola, Mozambique, and United Republic of Tanzania. However, Nigeria is the top leading producer of Cassava.

QUESTION 13. Which African Countries belong to the Top 5 Beverage Producers?

Observation: The Top 5 Producing Countries that has the highest total quantity of Beverage from 2004 to 2013 are: Nigeria, Burkina Faso, Uganda, United Republic of Tanzania, and Cameroon. However, Nigeria is the top leading producer of beverage.

QUESTION 14. Which African Countries belong to the Top 5 Rice Producers?

Observation: The Top 5 countries that have produced the highest total quantity of rice from 2004 to 2013 are: Egypt, Nigeria, Madagascar, United Republic of Tanzania, and Guinea. However, Egypt is the top leading producer of Rice.
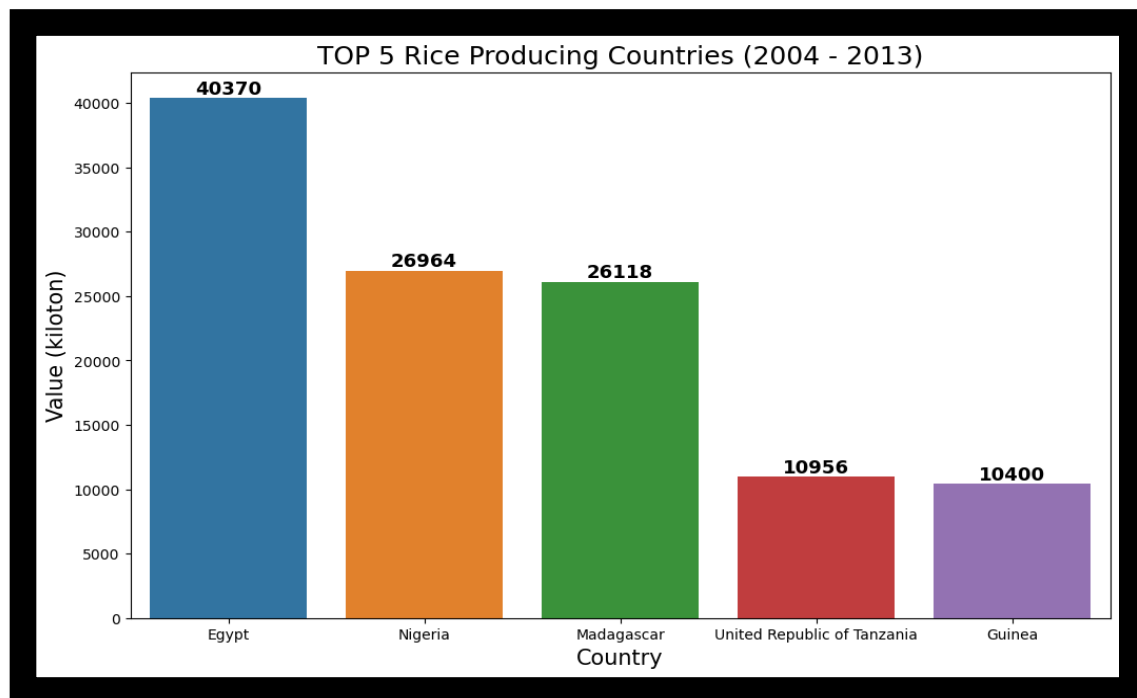
This is a link to the interactive map- file:///C:/Users/aroko/rice_production_map.html. that shows the Top 5 Rice Producing Countries.

QUESTION 15: Is there a relationship between Africa's population and the total amount of rice produced in African countries?

TEST OF HYPOTHESIS

Population growth may have an impact on Africa's high food production between 2004 and 2013.According to the FAO, rice was the third most produced food in the world in 2014.Rice (Milled Equivalent), a staple meal in most of the African Countries, may be used in hypothesis testing to gauge this.

The World Bank population dataset, the Country Code dataset, and the Rice (Milled Equivalent) production dataset for African Countries from 2004 to 2013 will be combined to provide test data for hypothesis measurement.

To analyze the linear relationship between our variables, X and Y, I will utilize the Pearson Correlation Coefficient.

> X: **The Annual Volume of rice produced in African Countries**

> **Y: The Average annual population of African Countries.**

The hypothesis test statements: H0 <= 0(NULL HYPOTHESIS), H1 > 0(ALTERNATIVE HYPOTHESIS)

NULL HYPOTHESIS: No significant correlation between population average in African countries and rice production in a specific year. i.e. no relationship between the two variables.

ALTERNATIVE HYPOTHESIS: African countries population and rice production are significantly correlated. i.e there is a relationship between the two variables.
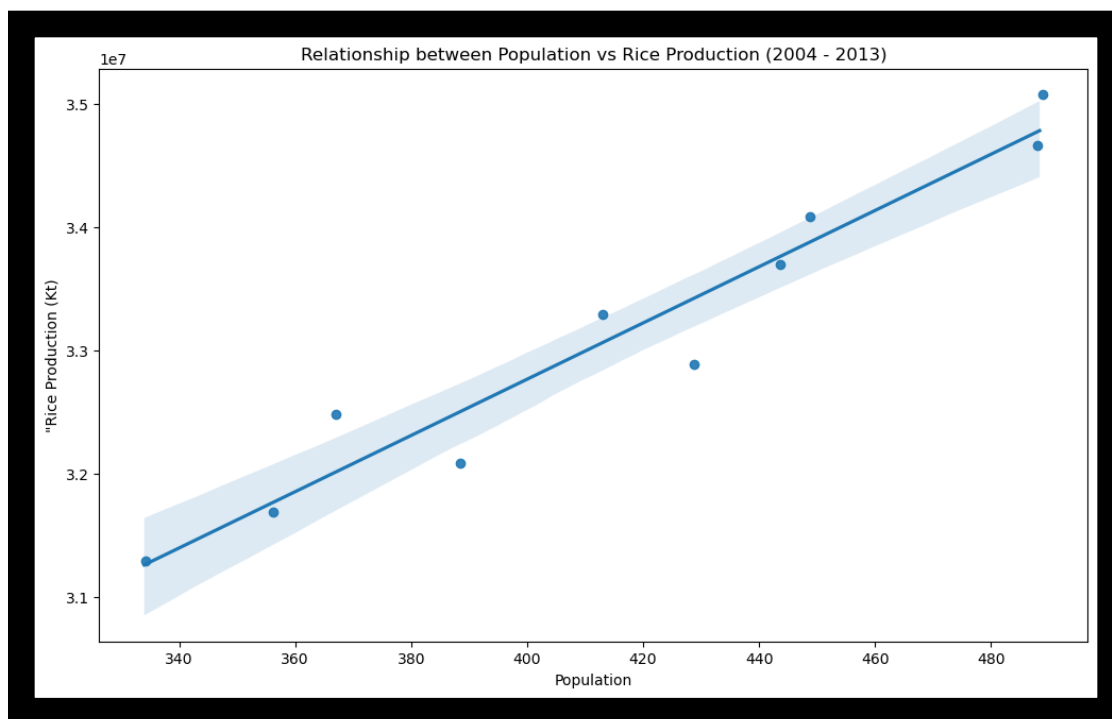
Observation: The correlation hypothesis shows a significant positive relationship between rice production and African countries population during a specific year, the small p value shows that the null hypothesis (H0) is rejected at the 5% level.

The correlation coefficient "0.97" shows a strong relationship between average rice production each year and average African population size in the same year.

However, despite a significant relationship between population and rice production, correlation doesn't imply causation. This does not necessarily imply that rice production is what causes population growth.

CONCLUSION

According to the Hypothesis testing, there was a bit significant relationship between population, food supply, and production in Africa between 2004 and 2013. Between 2004 and 2013, African countries experienced a potential food shortage due to a low food supply compared to increased production of food and population.

However, the population doesn't seem to explain the possible cause of the food shortage in Africa during this time other elements like price inflation, exports, and inadequate storage facilities may be reasons for the shortage because countries with high food production have smaller food supply compared to their production.

African countries must embark on a more deliberate transition from subsistence to commercial agriculture, this will enhance both the food supply and the countries total GDP. It will also help Africa country decrease or eradicate global hunger.