Import Libraries

In [1]:
```
#import Relevant libraries
import pandas as pd
import numpy as np
import matplotlib as mpl
import matplotlib.pyplot as plt
import seaborn as sns
```

Read the datasets

In [2]:
```
#Read the dataset and make a copy of each
prod_data = pd.read_csv('Africa Food Production (2004 - 2013).csv')
prod_data_copy  = prod_data.copy()
sup_data = pd.read_csv('Africa Food Supply (2004 - 2013).csv')
sup_data_copy = sup_data.copy()
#pop_data = pd.read_csv('population.csv')
#pop_data_copy = pop_data.copy()
```

In [3]:
```
#find the first five rows.
print(prod_data.head())
print(sup_data.head())
```

```
   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
   Country  Year  Value
0  Algeria  2004   2987
1  Algeria  2005   2958
2  Algeria  2006   3047
3  Algeria  2007   3041
4  Algeria  2008   3048
```

Exploratory Data Analysis for providing Insights

In [4]:
```
#Convert year to dateobject to enable us work with it and add unit to values
#data['Year'] = pd.to_datetime(data['Year'], format="%Y")
#prod_data['Value'] = [str(value) + "kt" for value in prod_data['Value']]
```

In [5]:
```
prod_data.info()
sup_data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 23110 entries, 0 to 23109
Data columns (total 4 columns):
 #   Column   Non-Null Count  Dtype
---  ------   --------------  -----
 0   Country  23110 non-null  object
 1   Item     23110 non-null  object
 2   Year     23110 non-null  int64
 3   Value    23110 non-null  int64
dtypes: int64(2), object(2)
memory usage: 722.3+ KB
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 450 entries, 0 to 449
Data columns (total 3 columns):
 #   Column   Non-Null Count  Dtype
---  ------   --------------  -----
 0   Country  450 non-null    object
 1   Year     450 non-null    int64
 2   Value    450 non-null    int64
dtypes: int64(2), object(1)
memory usage: 10.7+ KB
```

In [6]:
```
#There are no missing values in any of the columns
print(prod_data.count())
print(sup_data.count())
```

```
Country    23110
Item       23110
Year       23110
Value      23110
dtype: int64
Country    450
Year       450
Value      450
dtype: int64
```

In [7]:
```
#Find the number of rows and columns in the data
print(prod_data.shape, sup_data.shape)
```

```
(23110, 4) (450, 3)
```

In [8]:
```
#Give the summary statistics of both numeric and text data
prod_data.describe(include='all')
```

Out[8]:

	Country	Item	Year	Value
count	23110	23110	23110.000000	23110.000000
unique	45	94	NaN	NaN
top	United Republic of Tanzania	Fats, Animals, Raw	NaN	NaN
freq	790	450	NaN	NaN
mean	NaN	NaN	2008.498269	327.785201
std	NaN	NaN	2.871740	1607.940343
min	NaN	NaN	2004.000000	0.000000
25%	NaN	NaN	2006.000000	3.000000
50%	NaN	NaN	2008.000000	18.000000
75%	NaN	NaN	2011.000000	108.000000
max	NaN	NaN	2013.000000	54000.000000

In [9]:
```
sup_data.describe(include='all')
```

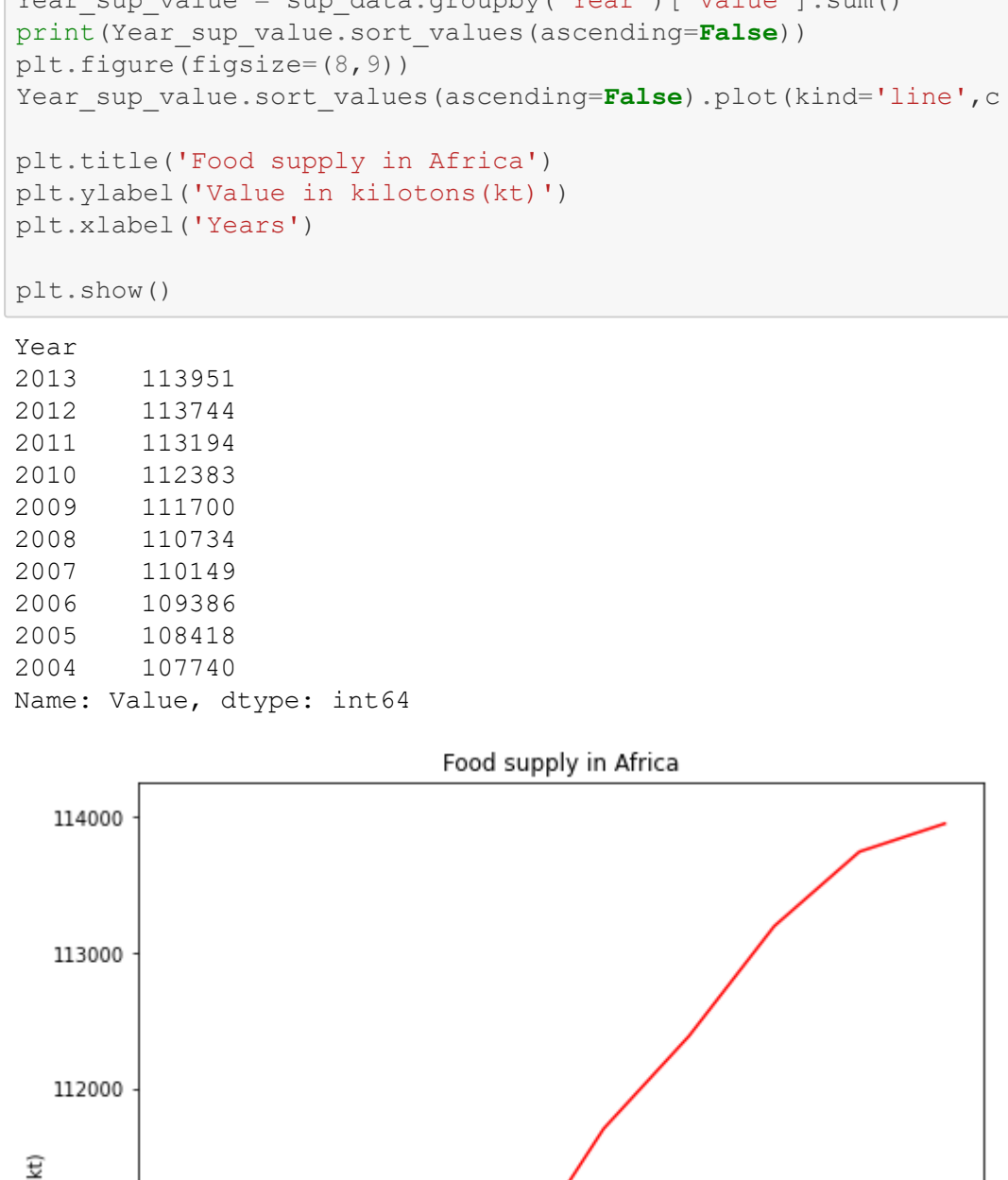
Out[9]:

	Country	Year	Value
count	450	450.000000	450.000000
unique	45	NaN	NaN
top	Botswana	NaN	NaN
freq	10	NaN	NaN
mean	NaN	2008.500000	2469.775556
std	NaN	2.875478	379.140143
min	NaN	2004.000000	1781.000000
25%	NaN	2006.000000	2174.000000
50%	NaN	2008.500000	2376.000000
75%	NaN	2011.000000	2681.750000
max	NaN	2013.000000	3561.000000

All Value are in kt(kilotons)

In [10]:
```
#Find the highest production year and visualize
Year_prod_value = prod_data.groupby('Year')['Value'].sum()
print(Year_prod_value.sort_values(ascending=False))
plt.figure(figsize=(8,9))
Year_prod_value.sort_values(ascending=False).plot(kind='line')

plt.title('Food production in Africa')
plt.ylabel('Value in kilotons(kt)')
plt.xlabel('Years')

plt.show()
```

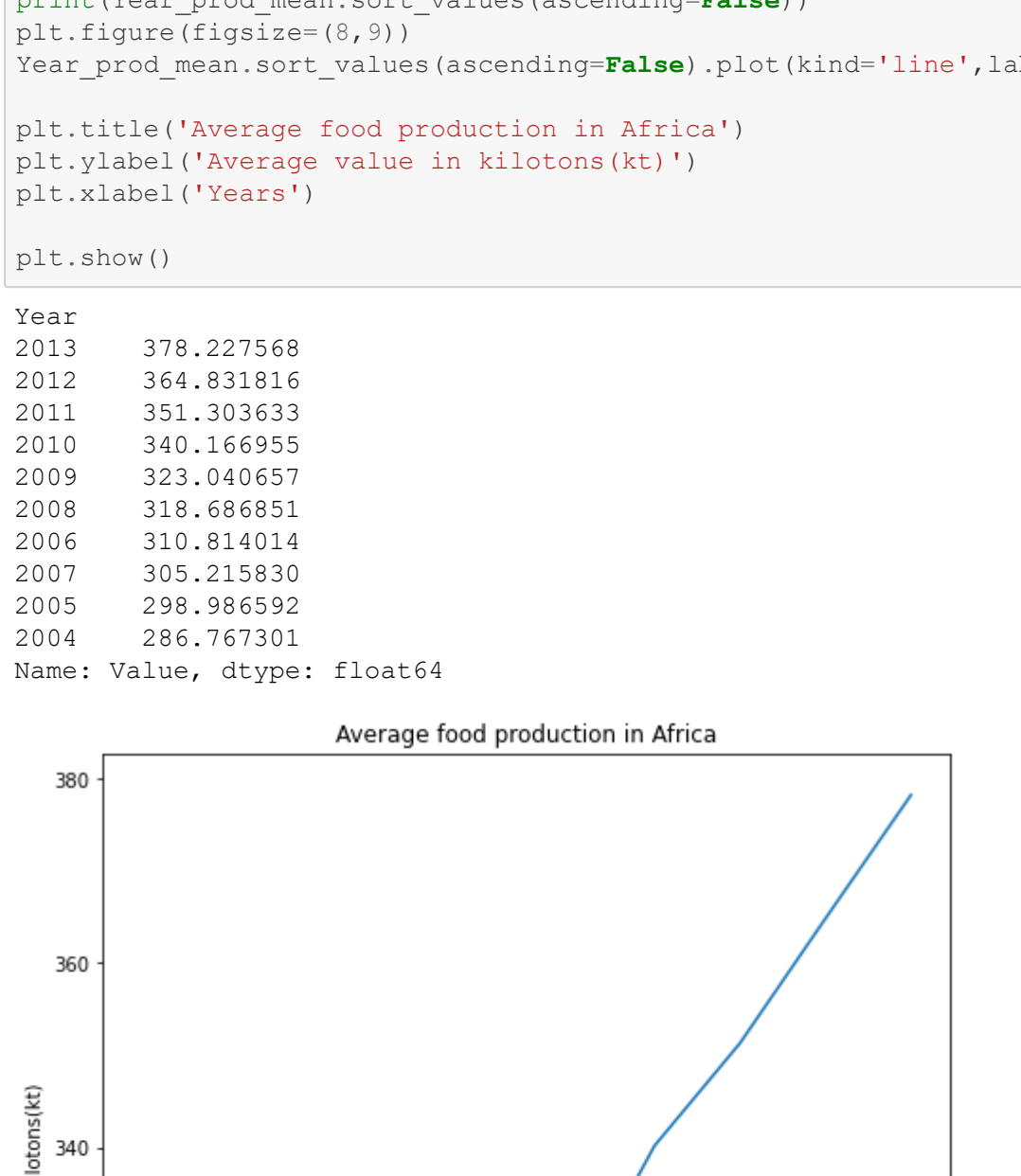
```
Year
2013    872571
2012    841667
2011    812214
2010    786466
2009    746870
2008    736804
2006    718602
2007    705659
2005    691257
2004    663006
Name: Value, dtype: int64
```



There was a slight reduction from 2006 to 2007. Apart from that, food production in africa has always been on the increase for the given data.

In [11]:
```
#Find the highest supply year and visualize
Year_sup_value = sup_data.groupby('Year')['Value'].sum()
print(Year_sup_value.sort_values(ascending=False))
plt.figure(figsize=(8,9))
Year_sup_value.sort_values(ascending=False).plot(kind='line',c ='r')

plt.title('Food supply in Africa')
plt.ylabel('Value in kilotons(kt)')
plt.xlabel('Years')

plt.show()
```

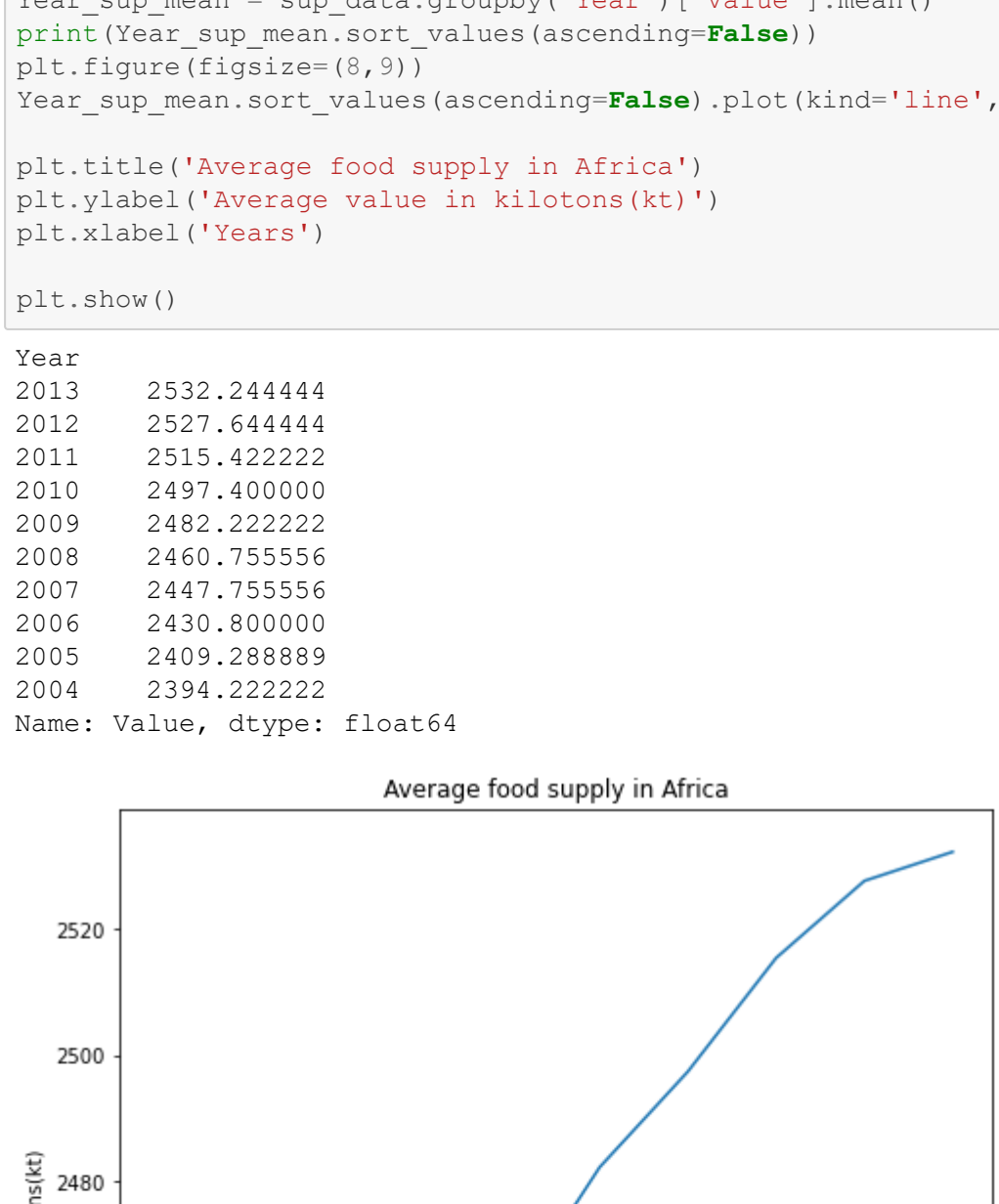
```
Year
2013    113951
2012    113744
2011    113194
2010    112383
2009    111700
2008    110734
2007    110149
2006    109386
2005    108418
2004    107740
Name: Value, dtype: int64
```



There is an upward trend in the supply of food as year increase.

In [12]:
```
#Find the average production by year and visualize
Year_prod_mean = prod_data.groupby('Year')['Value'].mean()
print(Year_prod_mean.sort_values(ascending=False))
plt.figure(figsize=(8,9))
Year_prod_mean.sort_values(ascending=False).plot(kind='line',label="Values of food produced in kt")

plt.title('Average food production in Africa')
plt.ylabel('Average value in kilotons(kt)')
plt.xlabel('Years')

plt.show()
```

```
Year
2013    378.227568
2012    364.831816
2011    351.303633
2010    340.166955
2009    323.040657
2008    318.686851
2006    310.814014
2007    305.215830
2005    298.986592
2004    286.767301
Name: Value, dtype: float64
```



There is positive trend in the average year production accross the countries too except 2007 where there slight decrease.

In [13]:
```
#Find the average production by year and visualize
Year_sup_mean = sup_data.groupby('Year')['Value'].mean()
print(Year_sup_mean.sort_values(ascending=False))
plt.figure(figsize=(8,9))
Year_sup_mean.sort_values(ascending=False).plot(kind='line',label="Values of food supplied in kt")

plt.title('Average food supply in Africa')
plt.ylabel('Average value in kilotons(kt)')
plt.xlabel('Years')

plt.show()
```

```
Year
2013    2532.244444
2012    2527.644444
2011    2515.422222
2010    2497.400000
2009    2482.222222
2008    2460.755556
2007    2447.755556
2006    2430.800000
2005    2409.288889
2004    2394.222222
Name: Value, dtype: float64
```

There is positive trend in the average supply production accross the countries too.

In [14]:
```
# Since the highest production year is 2013
#Let us find what item was the highest produced in the year 2013.
# first set year as index to locate
prod_data.set_index('Year', inplace=True)
```

In [15]:
```
data_2013 = prod_data.loc[2013]
data_2013.head()
data_2013.reset_index()
```

Out[15]:

	Year	Country	Item	Value
0	2013	Algeria	Wheat and products	3299
1	2013	Algeria	Rice (Milled Equivalent)	0
2	2013	Algeria	Barley and products	1499
3	2013	Algeria	Maize and products	1
4	2013	Algeria	Oats	113
...	...	...	...	...
2302	2013	Zimbabwe	Fats, Animals, Raw	7
2303	2013	Zimbabwe	Eggs	30
2304	2013	Zimbabwe	Milk - Excluding Butter	410
2305	2013	Zimbabwe	Freshwater Fish	18
2306	2013	Zimbabwe	Crustaceans	0

2307 rows × 4 columns

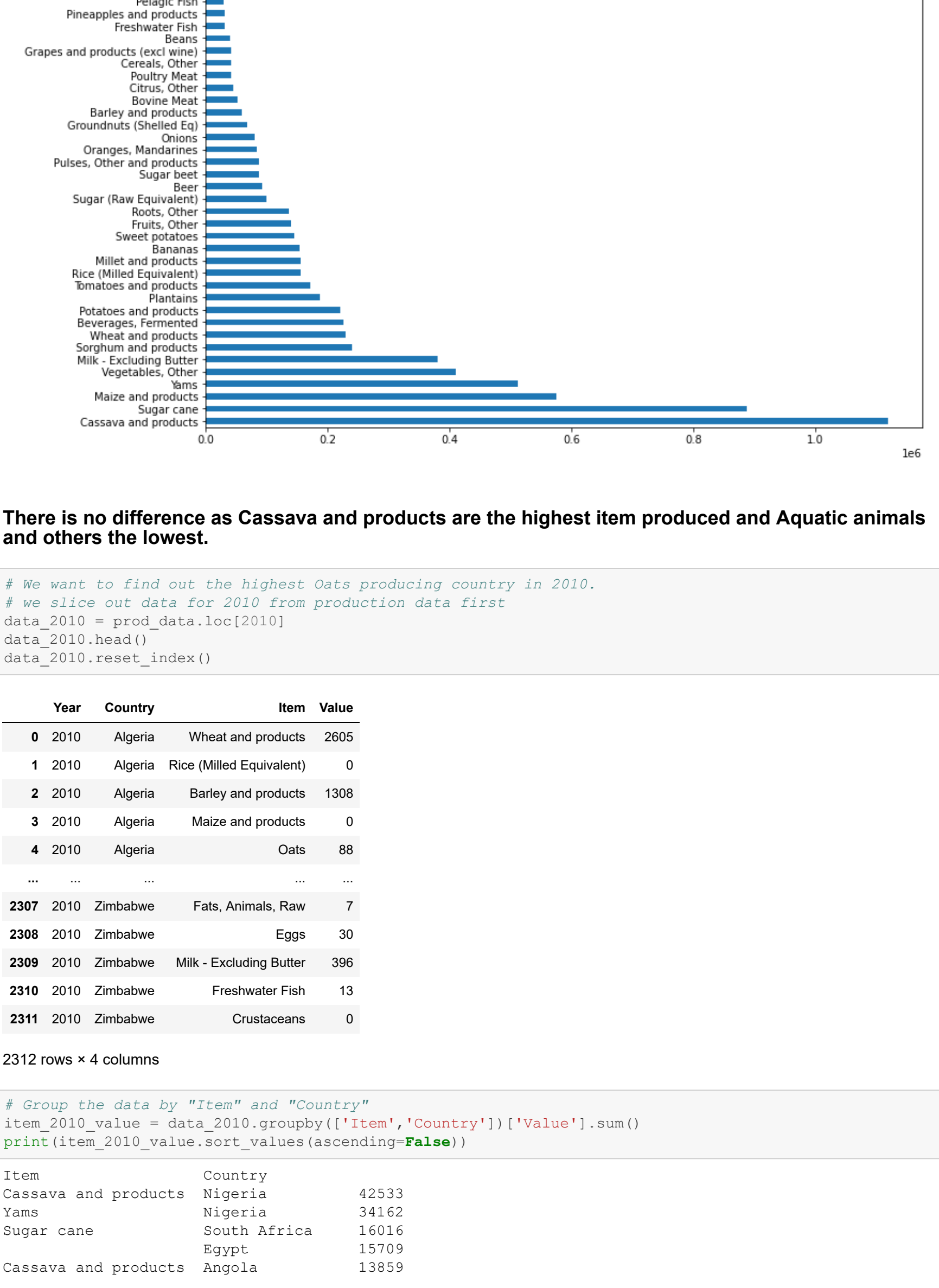
In [16]:
```
item_2013_value = data_2013.groupby('Item')['Value'].sum()
print(item_2013_value.sort_values(ascending=False))
# Cassava and products are the highest item produced in 2013, let us find out the highest produced item
 across the years
#And see if there is a difference.
```

```
Item
Cassava and products       136350
Sugar cane                  95348
Maize and products          70741
Yams                        60532
Vegetables, Other           46243
                            ...
Fish, Body Oil                 30
Molluscs, Other                29
Pepper                         23
Fish, Liver Oil                 4
Aquatic Animals, Others         1
Name: Value, Length: 94, dtype: int64
```

Cassava and products are the highest item produced in 2013, let us find out the highest produced item in all the years.

In [17]:
```
item_prod_value = prod_data.groupby('Item').sum()
#print(item_prod_value.sort_values(ascending=False))
item_prod_value['Value'].sort_values(ascending=False).plot(kind='barh', figsize=(12,20))
```

Out[17]: `<matplotlib.axes._subplots.AxesSubplot at 0x2a7fe85f648>`



There is no difference as Cassava and products are the highest item produced and Aquatic animals and others the lowest.

In [18]:
```
# We want to find out the highest Oats producing country in 2010.
# we slice out data for 2010 from production data first
data_2010 = prod_data.loc[2010]
data_2010.head()
data_2010.reset_index()
```

Out[18]:

	Year	Country	Item	Value
0	2010	Algeria	Wheat and products	2605
1	2010	Algeria	Rice (Milled Equivalent)	0
2	2010	Algeria	Barley and products	1308
3	2010	Algeria	Maize and products	0
4	2010	Algeria	Oats	88
...	...	...	...	...
2307	2010	Zimbabwe	Fats, Animals, Raw	7
2308	2010	Zimbabwe	Eggs	30
2309	2010	Zimbabwe	Milk - Excluding Butter	396
2310	2010	Zimbabwe	Freshwater Fish	13
2311	2010	Zimbabwe	Crustaceans	0

2312 rows × 4 columns

In [19]:
```
# Group the data by "Item" and "Country"
item_2010_value = data_2010.groupby(['Item','Country'])['Value'].sum()
print(item_2010_value.sort_values(ascending=False))
```

```
Item                  Country
Cassava and products  Nigeria          42533
Yams                  Nigeria          34162
Sugar cane            South Africa     16016
                      Egypt            15709
Cassava and products  Angola           13859
                                       ...
Fish, Body Oil        Guinea               0
                      Ghana                0
                      Gambia               0
                      Egypt                0
                      Cote d'Ivoire        0
Name: Value, Length: 2312, dtype: int64
```

In [20]:
```
#Slice out oats and find the highest producing company
Oats = item_2010_value.loc['Oats']
print(Oats)
```

```
Country
Algeria         88
Ethiopia        48
Ghana            0
Kenya            3
Lesotho          0
Morocco         37
South Africa    34
Tunisia          2
Zimbabwe         1
Name: Value, dtype: int64
```

Algeria is highest producing country of Oats in 2010 and Zimbabwe is the lowest

In [21]:
```
# I want to find the least suppply countries from 2004 to 2006(inclusive)
# we first set as index year to slice out the years
sup_data.set_index('Year', inplace=True)
sup_year_range = sup_data.loc[[2004, 2005, 2006]]
sup_year_range.head()
sup_year_range.reset_index()
```

Out[21]:

	Year	Country	Value
0	2004	Algeria	2987
1	2004	Angola	2030
2	2004	Benin	2461
3	2004	Botswana	2191
4	2004	Burkina Faso	2505
...	...	...	...
130	2006	Tunisia	3273
131	2006	Uganda	2276
132	2006	United Republic of Tanzania	2141
133	2006	Zambia	1838
134	2006	Zimbabwe	2115

135 rows × 3 columns

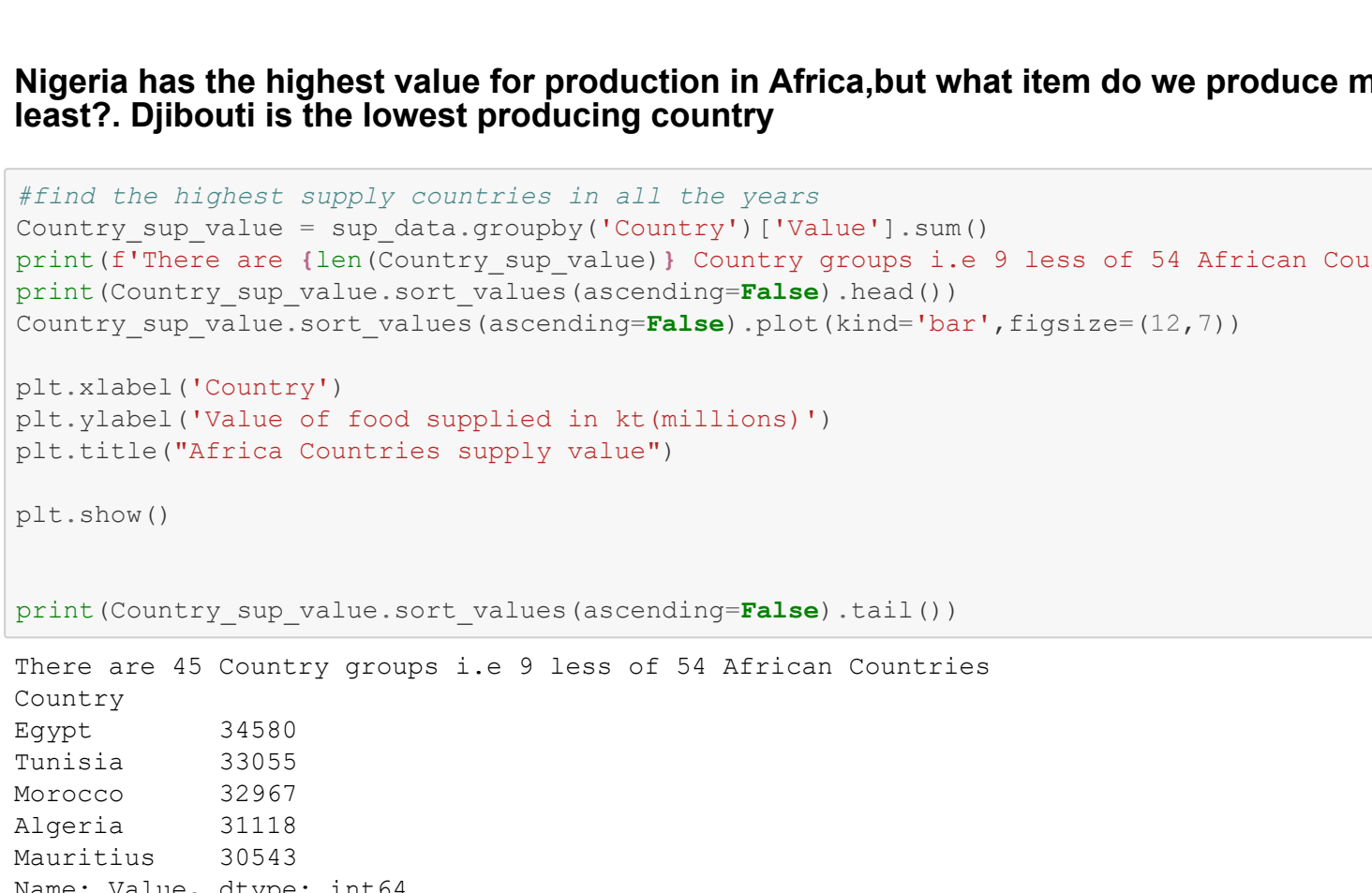
In [22]:
```
sup_range_value = sup_year_range.groupby('Country')['Value'].sum()
print(sup_range_value.sort_values(ascending=False).tail(10))
```

```
Country
Sierra Leone                6261
Angola                      6226
Mozambique                  6208
Zimbabwe                    6184
Madagascar                  6123
Chad                        6082
Rwanda                      6050
Central African Republic    6044
Ethiopia                    5827
Zambia                      5577
Name: Value, dtype: int64
```

The least supplier of food countries in the year range (2004 - 2006) in least order are "Zambia", "Ethiopia" and "CAR" as shown above

In [23]:
```
#find the highest production countries in all the years
Country_prod_value = prod_data.groupby('Country')['Value'].sum()
print(f'There are {len(Country_prod_value)} Country groups i.e 9 less of 54 African Countries')
print(Country_prod_value.sort_values(ascending=False).head())


Country_prod_value.sort_values(ascending=False).plot(kind='bar',figsize=(12,7))
plt.xlabel('Country')
plt.ylabel('Value of food produced in kt(millions)')
plt.title("Africa Countries production value")

plt.show()


print(Country_prod_value.sort_values(ascending=False).tail())
```

```
There are 45 Country groups i.e 9 less of 54 African Countries
Country
Nigeria                        1628030
Egypt                           877498
South Africa                    579592
Ethiopia                        350693
United Republic of Tanzania     332802
Name: Value, dtype: int64
```

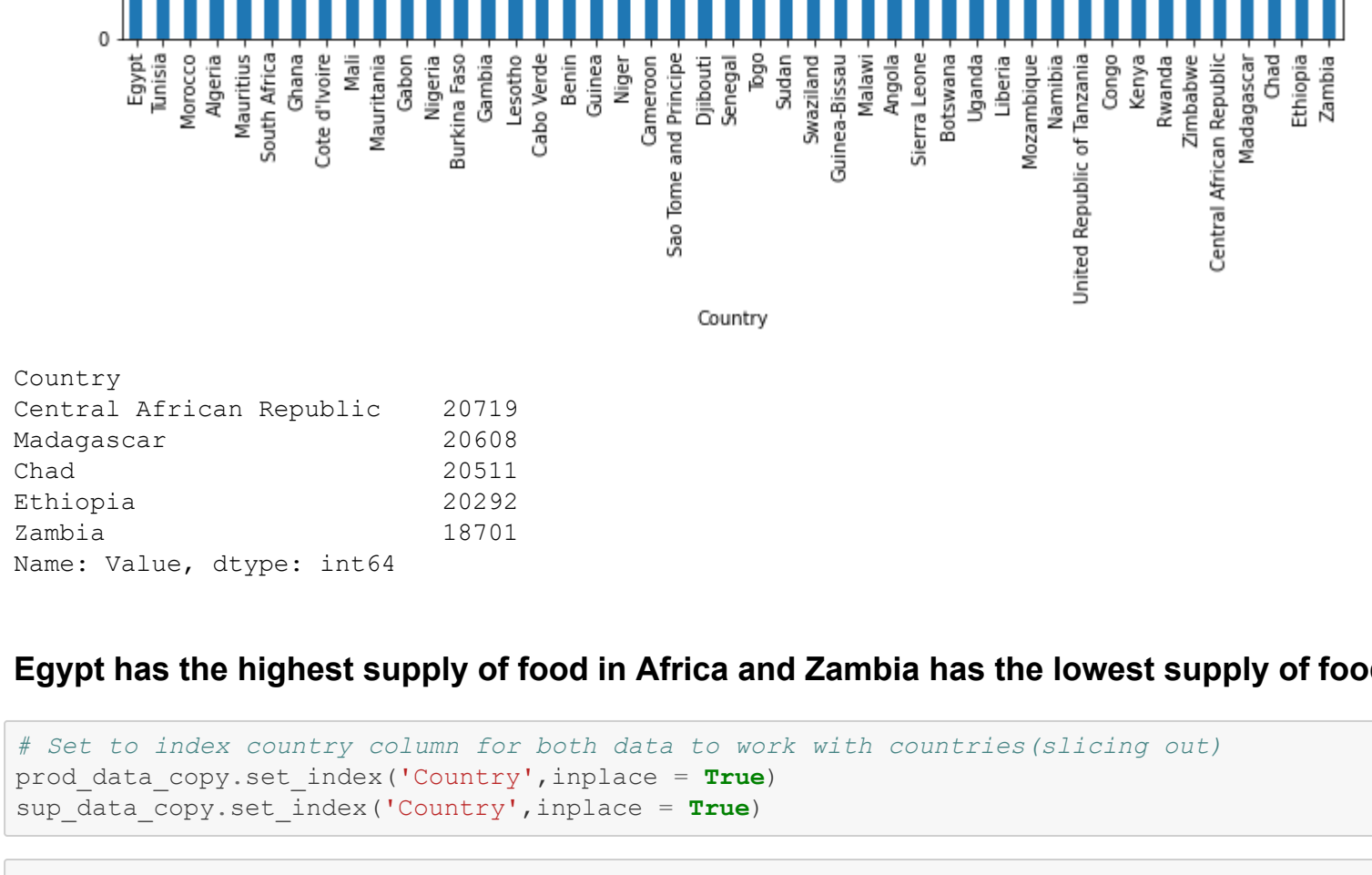


```
Country
Gambia                     4699
Lesotho                    3785
Cabo Verde                 1658
Sao Tome and Principe      1152
Djibouti                    666
Name: Value, dtype: int64
```

Nigeria has the highest value for production in Africa,but what item do we produce most and the least?. Djibouti is the lowest producing country

In [24]:
```
#find the highest supply countries in all the years
Country_sup_value = sup_data.groupby('Country')['Value'].sum()
print(f'There are {len(Country_sup_value)} Country groups i.e 9 less of 54 African Countries')
print(Country_sup_value.sort_values(ascending=False).head())
Country_sup_value.sort_values(ascending=False).plot(kind='bar',figsize=(12,7))

plt.xlabel('Country')
plt.ylabel('Value of food supplied in kt(millions)')
plt.title("Africa Countries supply value")

plt.show()


print(Country_sup_value.sort_values(ascending=False).tail())
```

```
There are 45 Country groups i.e 9 less of 54 African Countries
Country
Egypt        34580
Tunisia      33055
Morocco      32967
Algeria      31118
Mauritius    30543
Name: Value, dtype: int64
```



```
Country
Central African Republic    20719
Madagascar                  20608
Chad                        20511
Ethiopia                    20292
Zambia                      18701
Name: Value, dtype: int64
```

Egypt has the highest supply of food in Africa and Zambia has the lowest supply of food in Africa

In [25]:
```
# Set to index country column for both data to work with countries(slicing out)
prod_data_copy.set_index('Country',inplace = True)
sup_data_copy.set_index('Country',inplace = True)
```

In [26]:
```
print(prod_data_copy.head())
print(sup_data_copy.head())
```

```
                       Item  Year  Value
Country
Algeria  Wheat and products  2004   2731
Algeria  Wheat and products  2005   2415
Algeria  Wheat and products  2006   2688
Algeria  Wheat and products  2007   2319
Algeria  Wheat and products  2008   1111
         Year  Value
Country
Algeria  2004   2987
Algeria  2005   2958
Algeria  2006   3047
Algeria  2007   3041
Algeria  2008   3048
```

In [27]:
```
#locate Nigeria(the highest producing country), Djibouti (the lowest producing country) and Egypt (the
 highest supply country)
#in production data
Nigeria = prod_data_copy.loc['Nigeria']
Djibouti = prod_data_copy.loc['Djibouti']
Egypt = prod_data_copy.loc['Egypt']
```

In [28]:
```
#Find the highest item Nigeria produced and the least
Nga_prod = Nigeria.groupby('Item')['Value'].sum()
Nga_prod.sort_values(ascending= False)
```

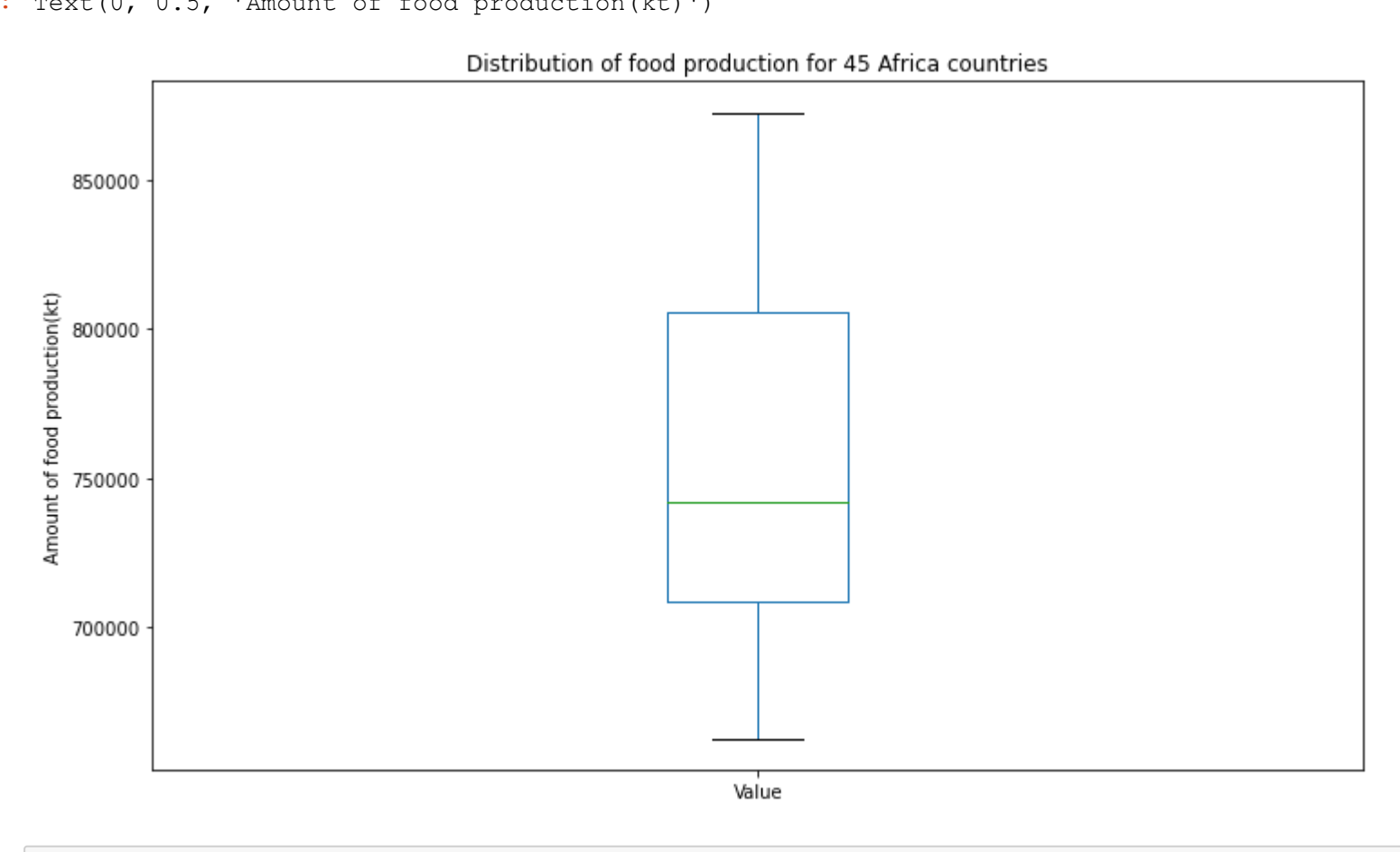
Out[28]:
```
Item
Cassava and products    452881
Yams                    349519
Beverages, Fermented     93839
Vegetables, Other        83335
Sorghum and products     78915
                          ...
Sesameseed Oil              35
Coffee and products         33
Soyabean Oil                32
Molluscs, Other             32
Cephalopods                  0
Name: Value, Length: 61, dtype: int64
```

In [29]:
```python
#Find the highest item Djibouti(lowest producing country) produced and the least
Dji_prod = Djibouti.groupby('Item')['Value'].sum()
Dji_prod.sort_values(ascending= False)
```

Out[29]:
```
Item
Vegetables, Other              295
Milk - Excluding Butter        153
Bovine Meat                     53
Mutton & Goat Meat              50
Beans                           20
Offals, Edible                  19
Lemons, Limes and products      18
Fruits, Other                   17
Tomatoes and products           13
Meat, Other                     10
Pelagic Fish                     8
Demersal Fish                    6
Pimento                          3
Marine Fish, Other               1
Sugar (Raw Equivalent)           0
Maize and products               0
Oranges, Mandarines              0
Fats, Animals, Raw               0
Dates                            0
Crustaceans                      0
Cephalopods                      0
Sugar cane                       0
Name: Value, dtype: int64
```

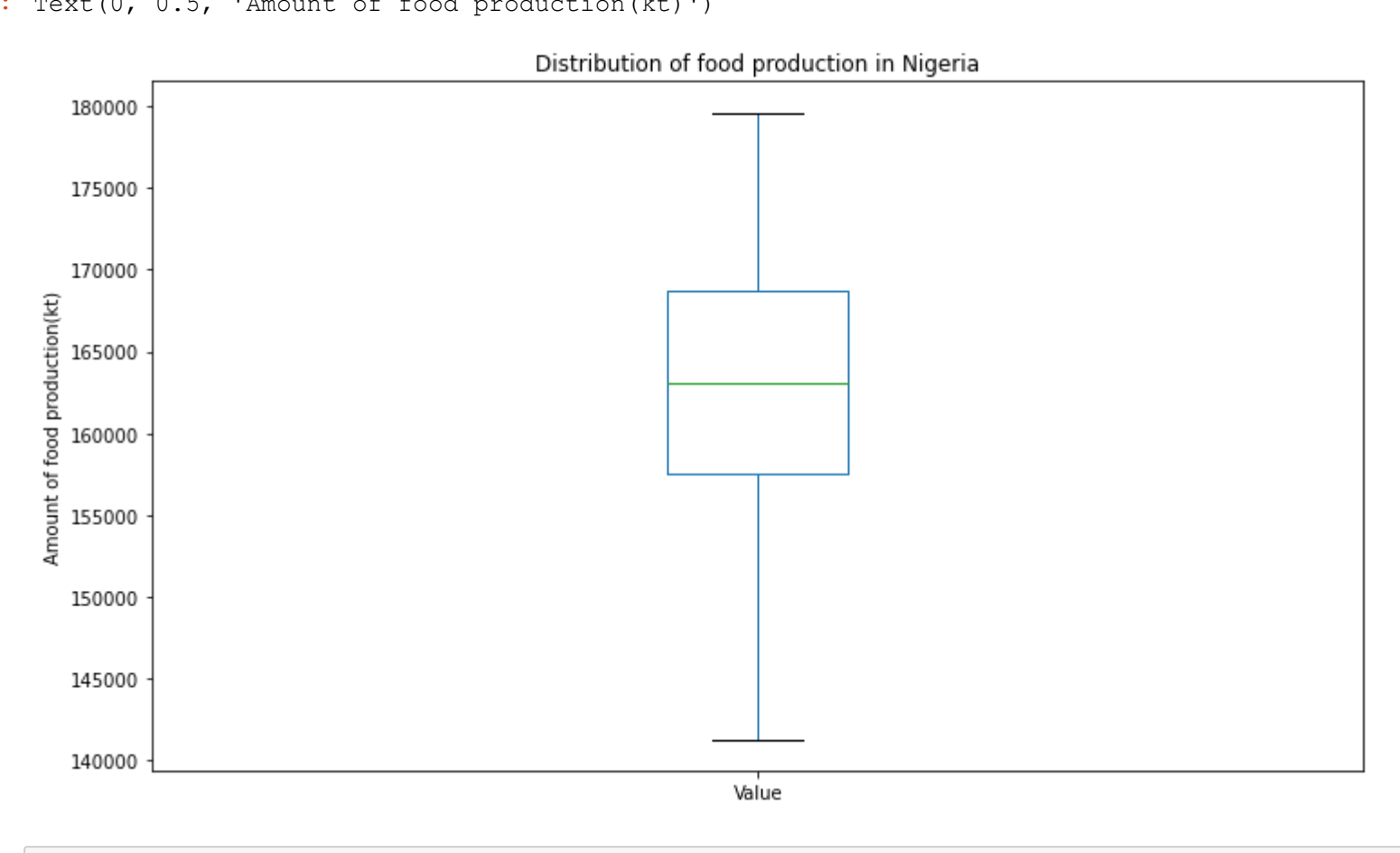
In [30]:
```python
#find the spread of the production data using boxplot
Year_prod_value.plot(kind='box', figsize=(12,7))

plt.title('Distribution of food production for 45 Africa countries')
plt.ylabel('Amount of food production(kt)')
```

Out[30]: Text(0, 0.5, 'Amount of food production(kt)')



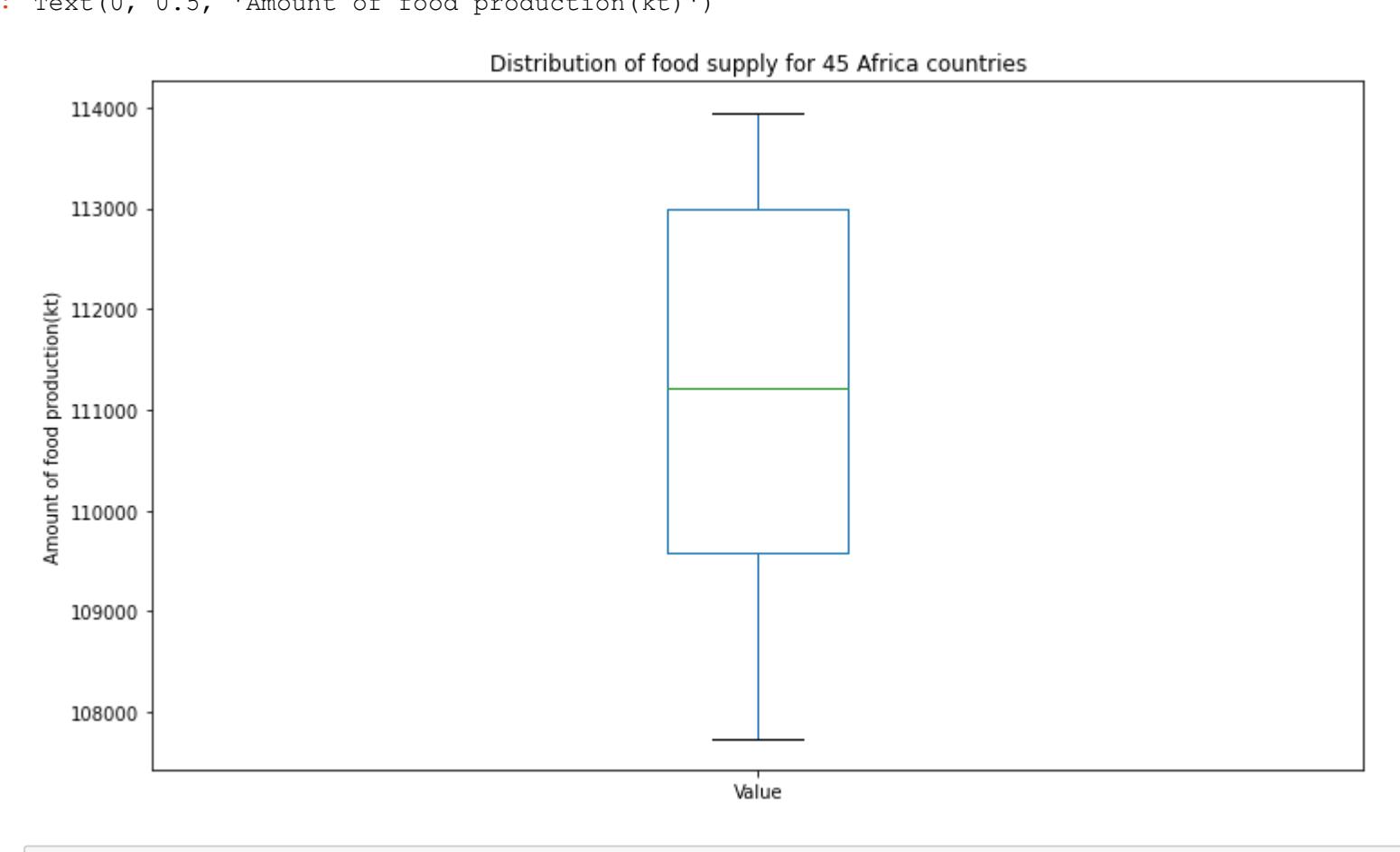
In [31]:
```python
#find the spread of food production in Nigeria
Nga_year = Nigeria.groupby('Year')['Value'].sum()
Nga_year.sort_values(ascending= False)
Nga_year.plot(kind='box', figsize=(12,7))

plt.title('Distribution of food production in Nigeria')
plt.ylabel('Amount of food production(kt)')
```

Out[31]: Text(0, 0.5, 'Amount of food production(kt)')



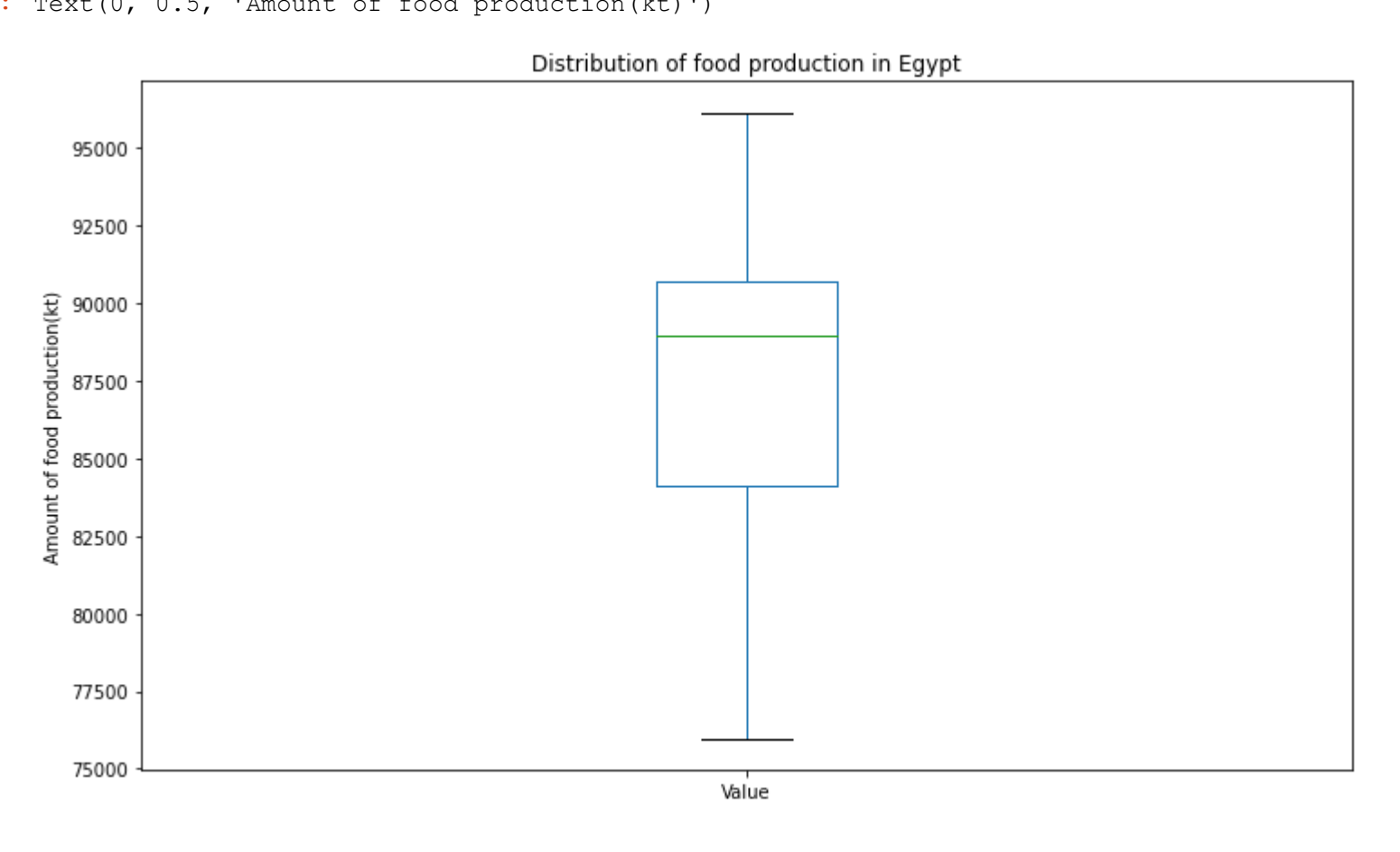
In [32]:
```python
#find the spread of the supply data using boxplot
Year_sup_value.plot(kind='box', figsize=(12,7))

plt.title('Distribution of food supply for 45 Africa countries')
plt.ylabel('Amount of food production(kt)')
```

Out[32]: Text(0, 0.5, 'Amount of food production(kt)')



In [33]:
```python
#find the spread of food production in Egypt and compare to Nigeria above
Egy_year = Egypt.groupby('Year')['Value'].sum()
Egy_year.sort_values(ascending= False)
Egy_year.plot(kind='box', figsize=(12,7))

plt.title('Distribution of food production in Egypt')
plt.ylabel('Amount of food production(kt)')
```

Out[33]: Text(0, 0.5, 'Amount of food production(kt)')



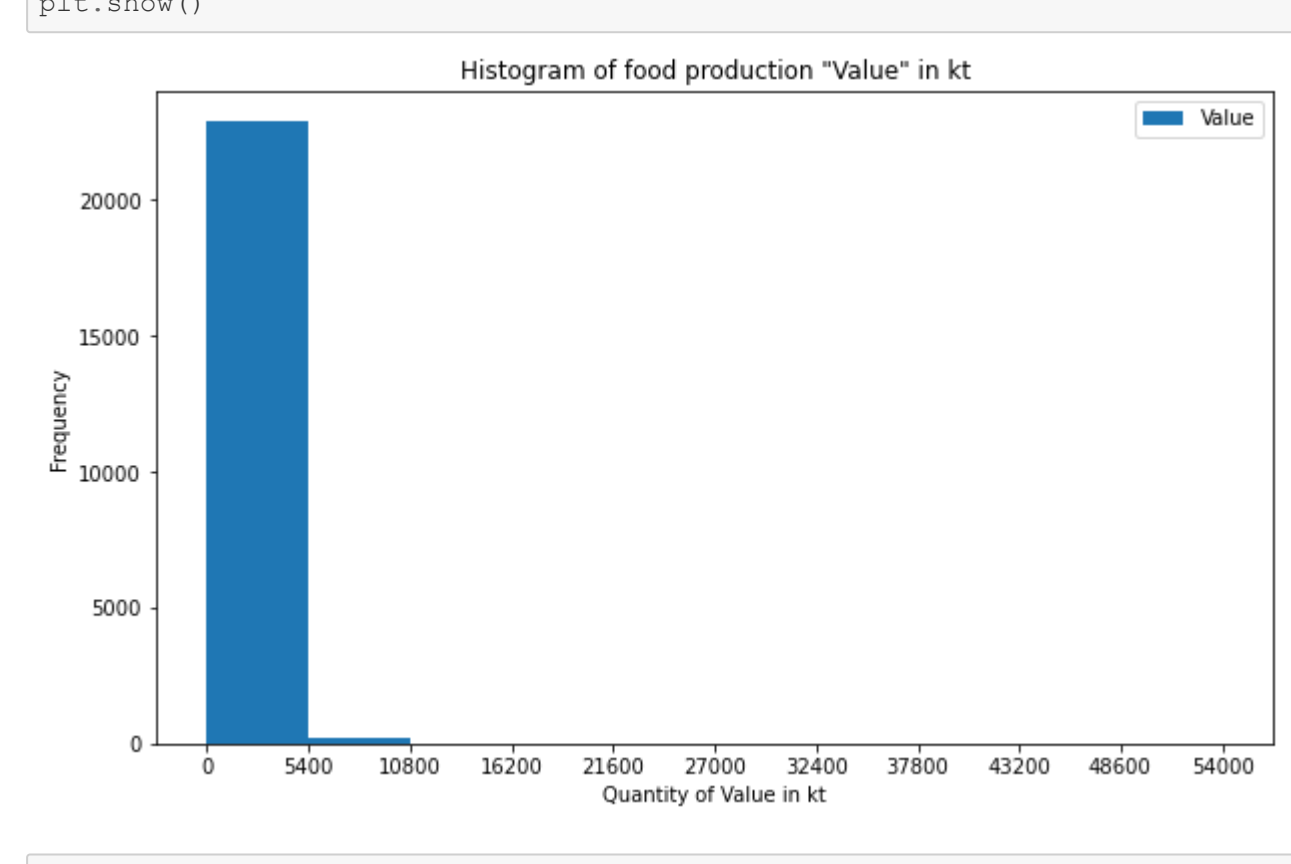
Eqypt food production is more spreadout than Nigeria.

In [34]:
```python
#Find the distribution of Value of the production data
count, bin_edges = np.histogram(prod_data['Value'])
print(count)
print(bin_edges)
```

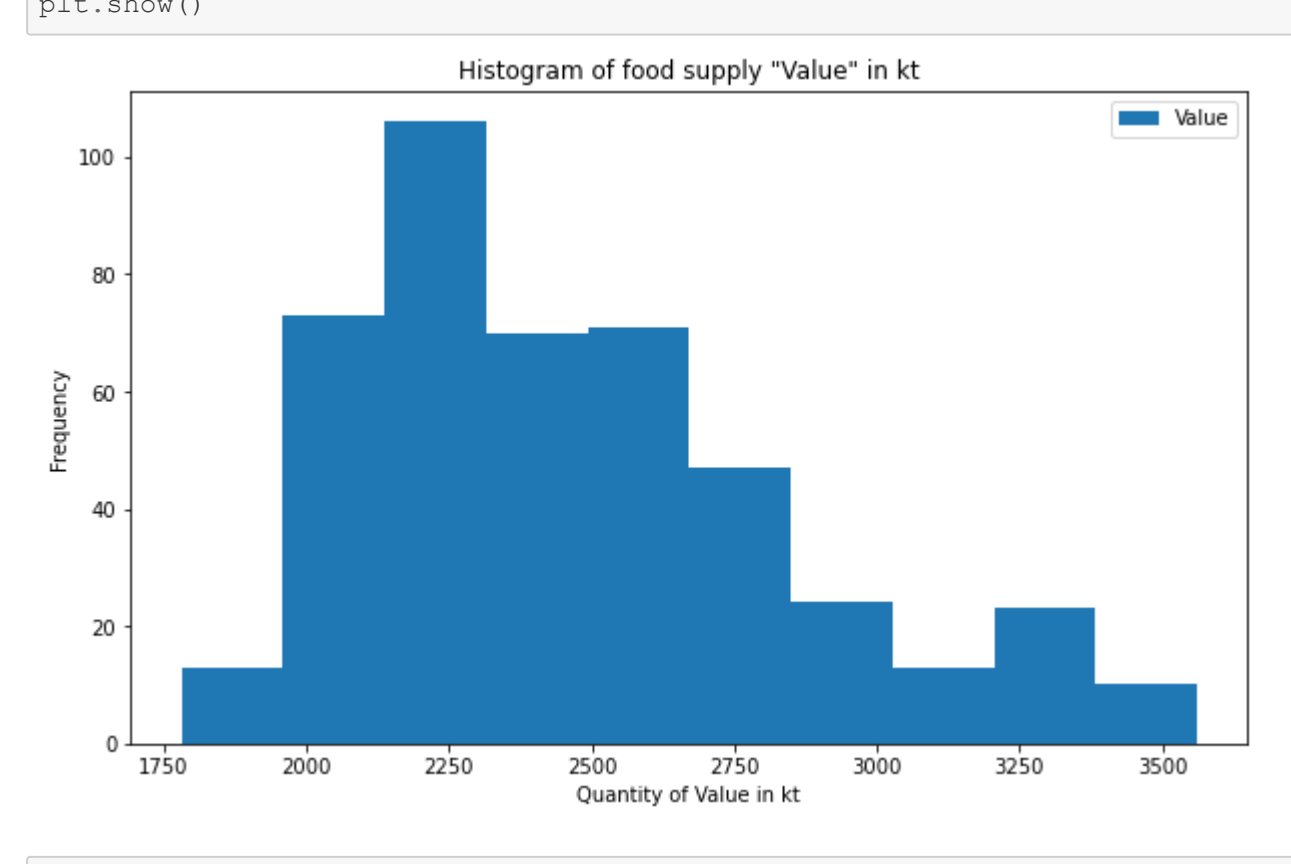
```
[22880   174    21    15     0     3     6     5     3     3]
[    0.  5400. 10800. 16200. 21600. 27000. 32400. 37800. 43200. 48600.
 54000.]
```

Values of food production in the range 0. to 5400. is 22,880 . As the range increases, food production quantity decrease drastically

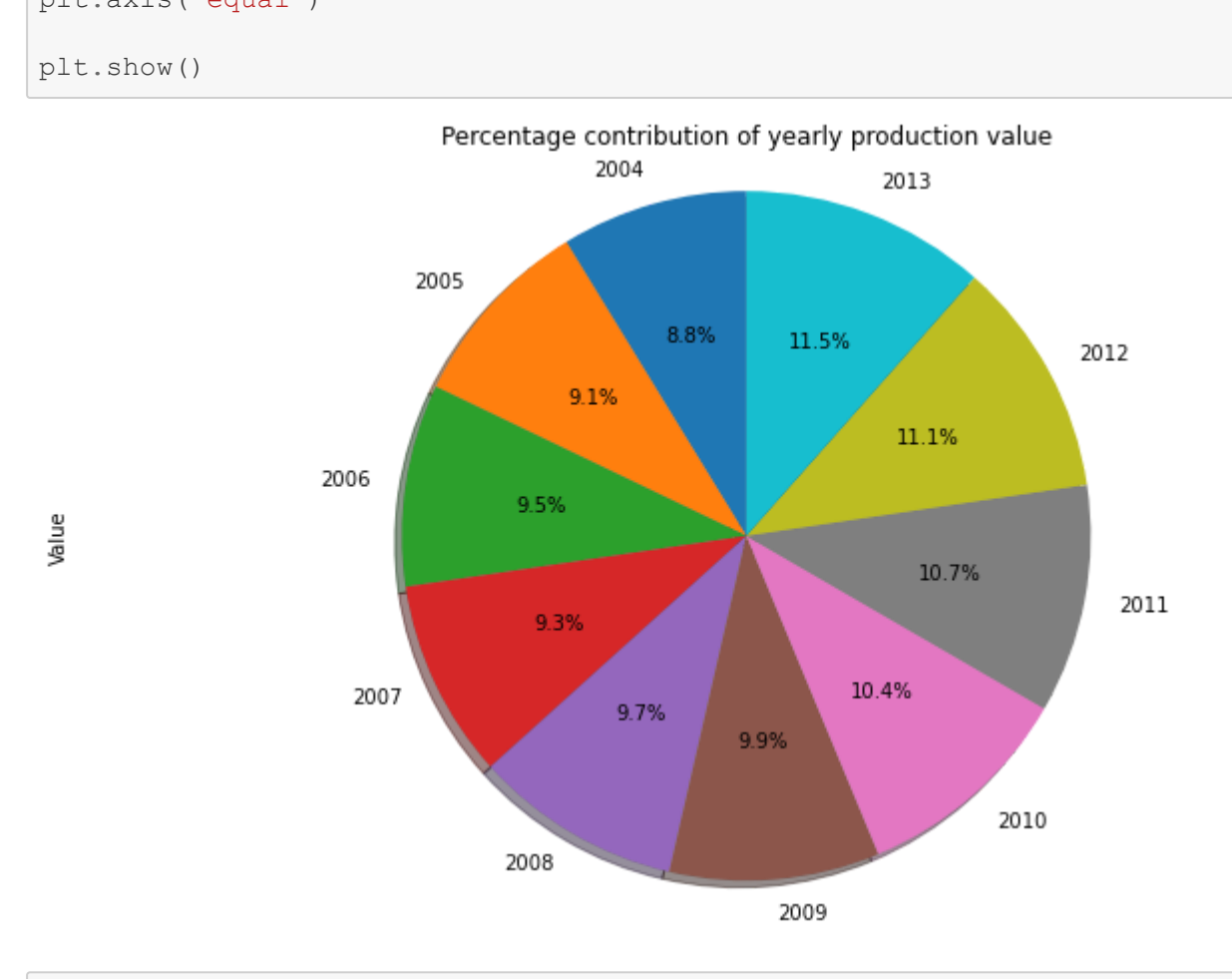
In [35]:
```python
prod_data.plot(kind='hist', figsize=(10,6), xticks=bin_edges)

plt.title('Histogram of food production "Value" in kt')
plt.ylabel('Frequency')
plt.xlabel('Quantity of Value in kt')

plt.show()
```



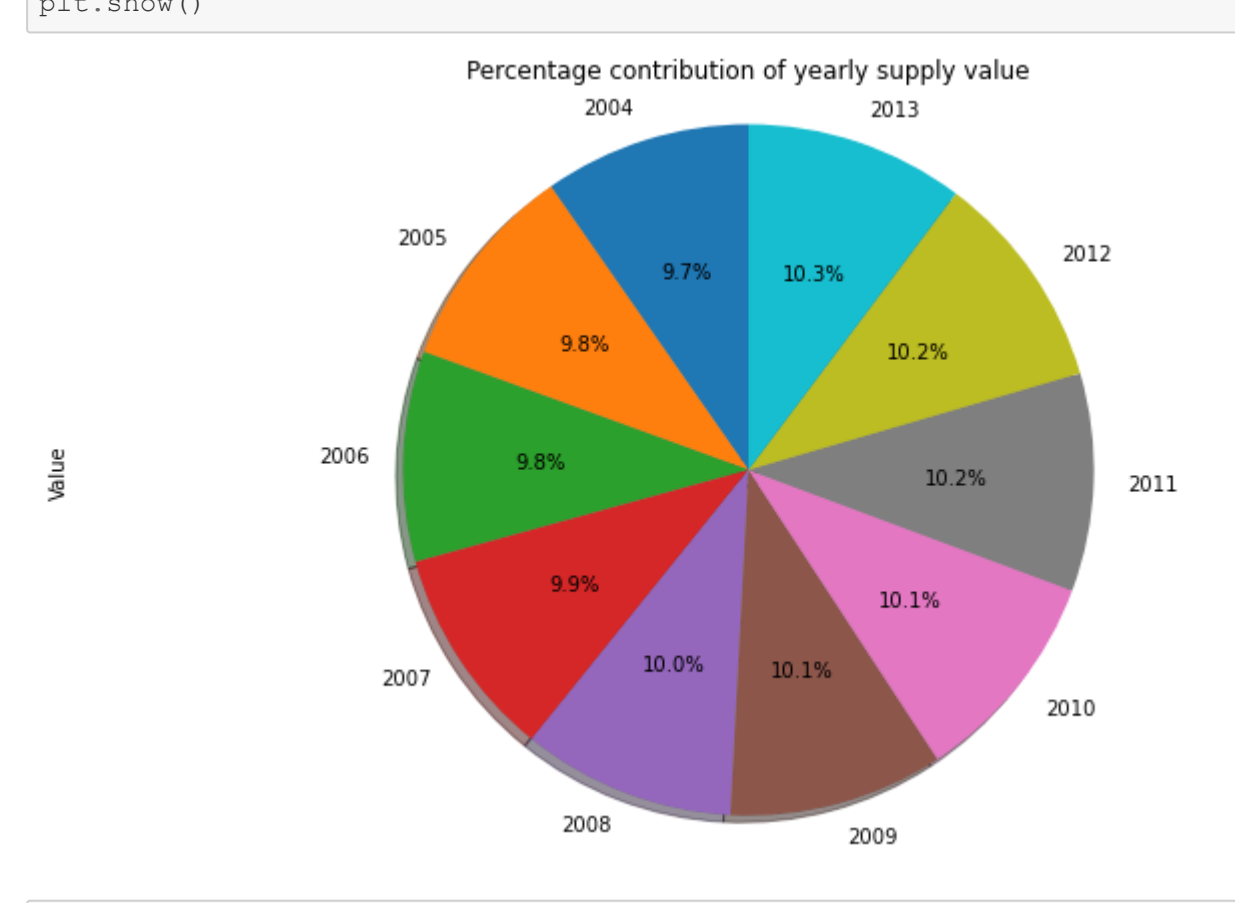
In [36]:
```python
#count, bin_edges = np.histogram(sup_data['Value'])
#print(count)
#print(bin_edges)
sup_data.plot(kind='hist', figsize=(10,6))

plt.title('Histogram of food supply "Value" in kt')
plt.ylabel('Frequency')
plt.xlabel('Quantity of Value in kt')

plt.show()
```



In [37]:
```python
#We find the percentage contribution of food production in  the given years
#autopct create percentage.

Year_prod_value.plot(kind='pie',figsize=(12, 7),autopct='%1.1f%%',startangle=90,
                     shadow=True)

plt.title("Percentage contribution of yearly production value")
plt.axis('equal')

plt.show()
```

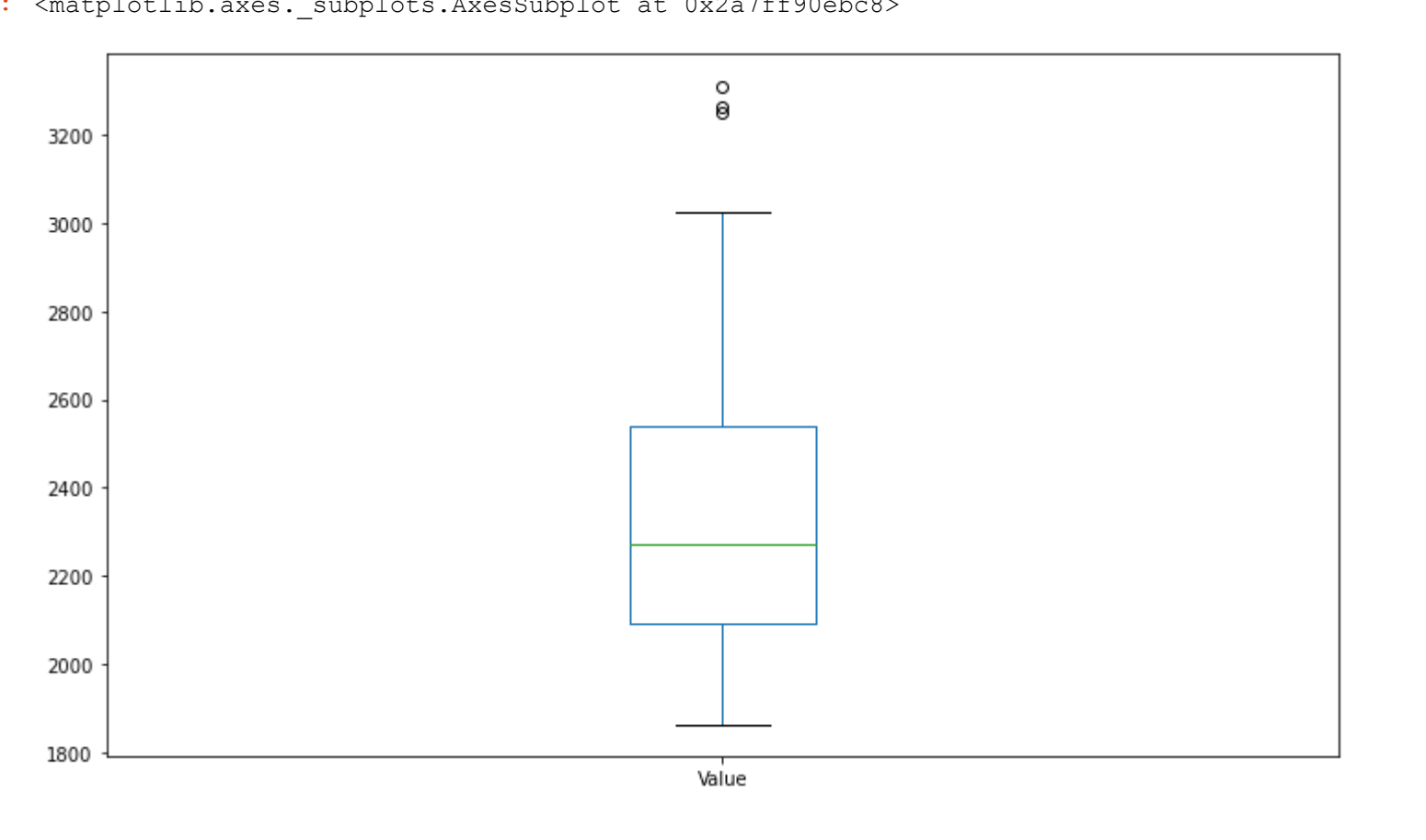


In [38]:
```python
#We find the percentage contribution of food supply in the given years
#autopct create percentage.

Year_sup_value.plot(kind='pie',figsize=(12, 7),autopct='%1.1f%%',startangle=90,
                    shadow=True)

plt.title("Percentage contribution of yearly supply value")
plt.axis('equal')

plt.show()
```



In [39]:
```python
# find the spread  of supply of food in 2004
sup_2004 = sup_data.loc[2004]
sup_2004.plot(kind='box', figsize=(12,7))
```

Out[39]: <matplotlib.axes._subplots.AxesSubplot at 0x2a7ff90ebc8>



There are outliers present, we can find the countries having those outliers

In [40]:
```python
sup_2004.head()
sup_2004.reset_index().sort_values(by="Value",ascending=False)
```

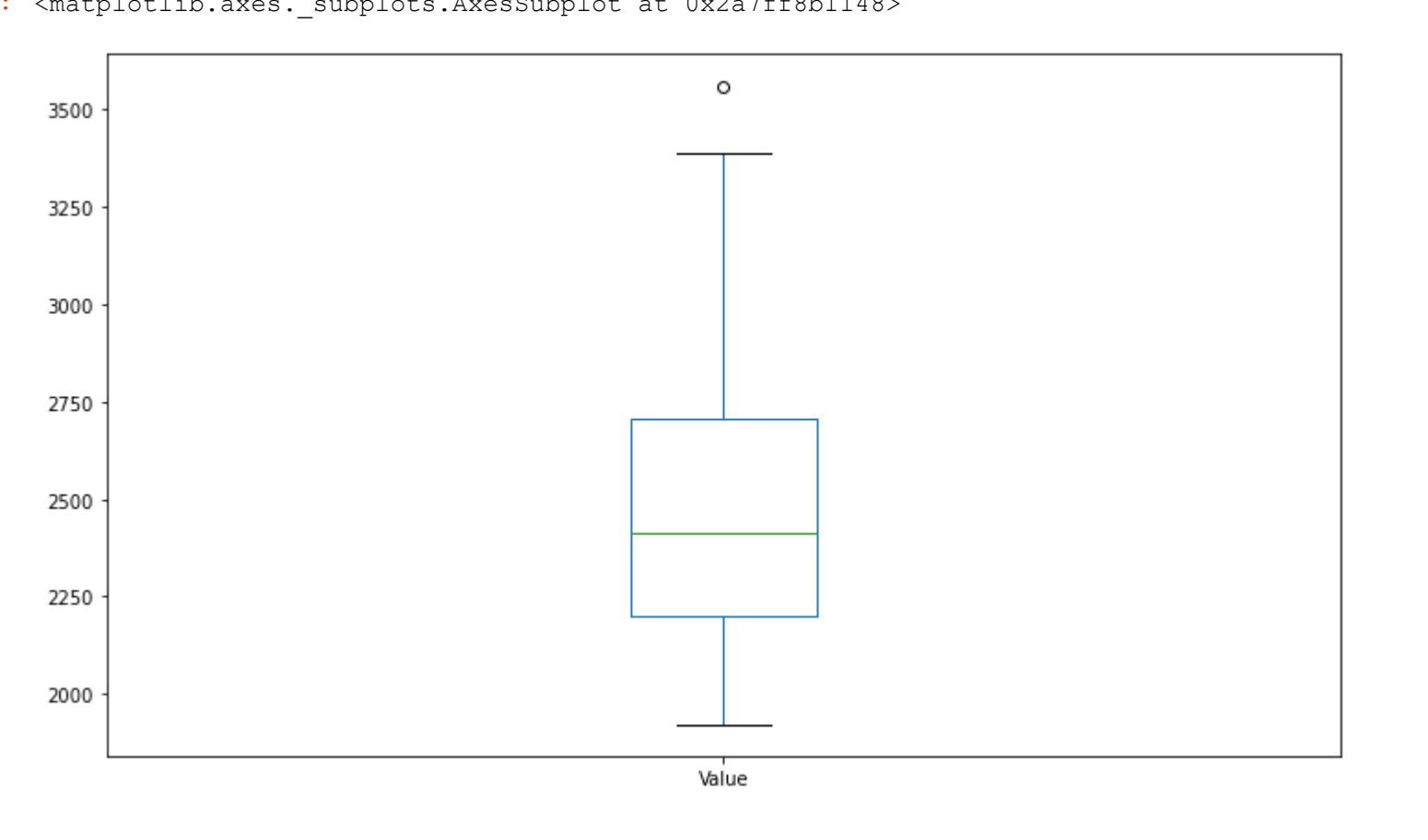
Out[40]:

	Year	Country	Value
12	2004	Egypt	3309
27	2004	Morocco	3263
40	2004	Tunisia	3248
26	2004	Mauritius	3023
0	2004	Algeria	2987
36	2004	South Africa	2940
10	2004	Cote d'Ivoire	2738
16	2004	Ghana	2682
14	2004	Gabon	2676
31	2004	Nigeria	2655
25	2004	Mauritania	2599
20	2004	Lesotho	2541
24	2004	Mali	2540
5	2004	Cabo Verde	2523
33	2004	Sao Tome and Principe	2513
4	2004	Burkina Faso	2505
38	2004	Swaziland	2469
15	2004	Gambia	2465
2	2004	Benin	2461
30	2004	Niger	2453
17	2004	Guinea	2426
41	2004	Uganda	2351
37	2004	Sudan	2271
29	2004	Namibia	2268
9	2004	Congo	2263
18	2004	Guinea-Bissau	2257
34	2004	Senegal	2248
39	2004	Togo	2246
6	2004	Cameroon	2246
3	2004	Botswana	2191
23	2004	Malawi	2166
11	2004	Djibouti	2136
42	2004	United Republic of Tanzania	2098
21	2004	Liberia	2092
28	2004	Mozambique	2059
44	2004	Zimbabwe	2044
1	2004	Angola	2030
8	2004	Chad	2026
35	2004	Sierra Leone	2024
19	2004	Kenya	2022
7	2004	Central African Republic	1989
22	2004	Madagascar	1980
32	2004	Rwanda	1969
13	2004	Ethiopia	1882
43	2004	Zambia	1866

The countries are Egypt, Morocco and Tunisia(they have high values compare to others).

In [41]:
```python
# Find the spread of supply of food in 2012
sup_2012 = sup_data.loc[2012]
sup_2012.head()
sup_2012.reset_index()
sup_2012.plot(kind='box', figsize=(12,7))
```

Out[41]: <matplotlib.axes._subplots.AxesSubplot at 0x2a7ff8b1148>



There is an outlier present , we can find the outlier with the code below and "Eqypt" is the outlier. We can find for the remaining years too.

In [42]:
```python
sup_2012_country =sup_2012 .groupby('Country')['Value'].sum()
print(sup_2012_country.sort_values(ascending=False))
```

```
Country
Egypt                          3561
Tunisia                        3390
Morocco                        3366
Algeria                        3272
Ghana                          3043
South Africa                   3014
Mauritius                      3014
Mauritania                     2873
Mali                           2860
Gabon                          2795
Cote d'Ivoire                  2780
Burkina Faso                   2707
Nigeria                        2665
Gambia                         2652
Cameroon                       2622
Benin                          2610
Cabo Verde                     2590
Niger                          2567
Guinea                         2561
Lesotho                        2547
Djibouti                       2545
Senegal                        2443
Togo                           2414
Angola                         2384
Sierra Leone                   2374
Botswana                       2342
Sao Tome and Principe          2338
Malawi                         2333
Guinea-Bissau                  2331
Sudan                          2301
Mozambique                     2298
Swaziland                      2293
Rwanda                         2270
Liberia                        2200
Zimbabwe                       2197
Kenya                          2192
United Republic of Tanzania    2192
Uganda                         2160
Congo                          2158
Namibia                        2145
Chad                           2124
Ethiopia                       2124
Central African Republic       2116
Madagascar                     2058
Zambia                         1923
Name: Value, dtype: int64
```

In [43]:
```python
# We want to find the country with the highest production of rice in the given years
# we set as index "Item" column
prod_data.set_index('Item', inplace=True)
```

In [44]:
```python
prod_data
```

Out[44]:

Item	Country	Value
Wheat and products	Algeria	2731
Wheat and products	Algeria	2415
Wheat and products	Algeria	2688
Wheat and products	Algeria	2319
Wheat and products	Algeria	1111
...	...	...
Crustaceans	Zimbabwe	0
Crustaceans	Zimbabwe	0
Crustaceans	Zimbabwe	0
Crustaceans	Zimbabwe	0
Crustaceans	Zimbabwe	0

23110 rows × 2 columns

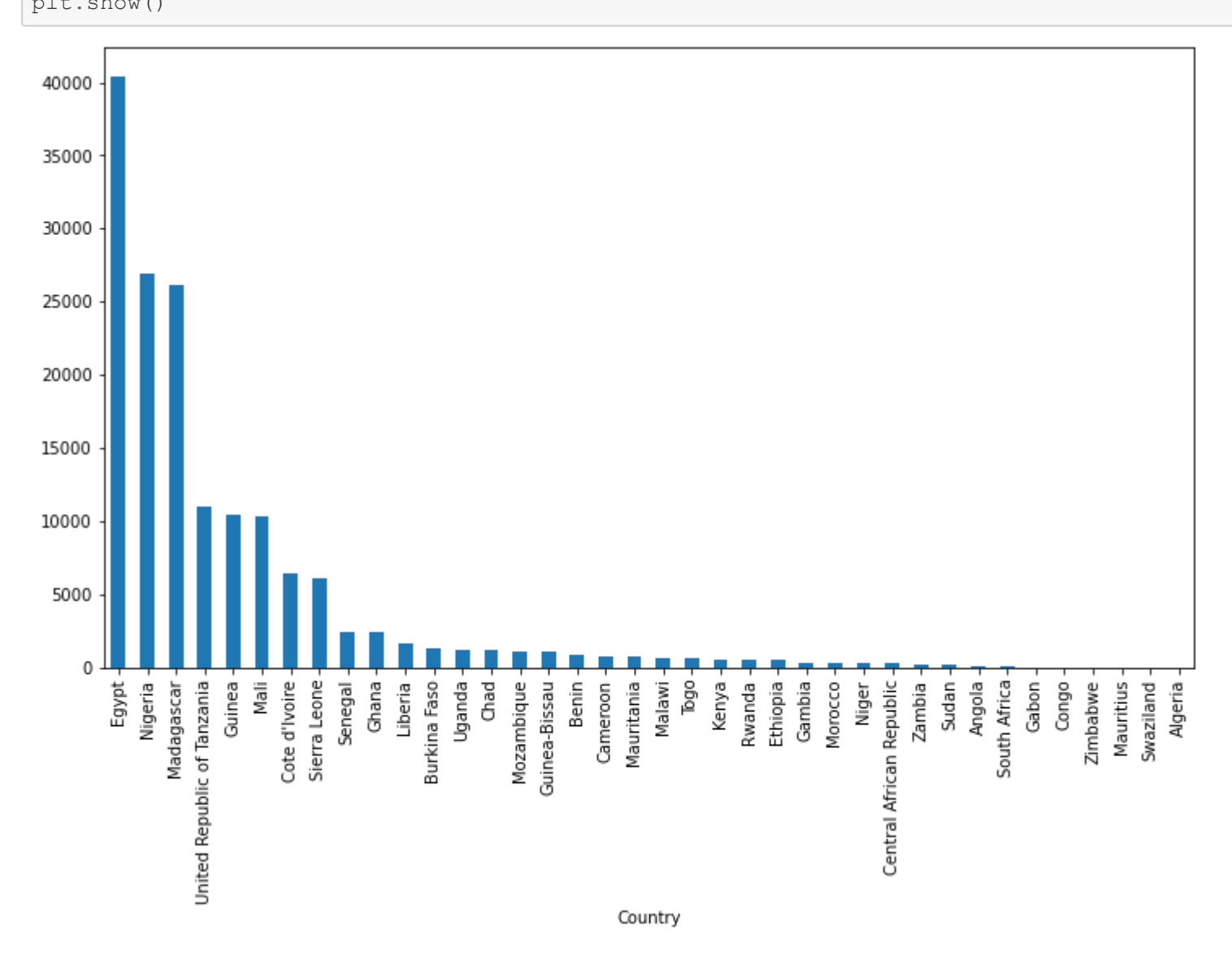
In [45]:
```python
rice_prod = prod_data.loc['Rice (Milled Equivalent)']
rice_prod.head()
rice_prod.reset_index()
```

Out[45]:

	Item	Country	Value
0	Rice (Milled Equivalent)	Algeria	0
1	Rice (Milled Equivalent)	Algeria	0
2	Rice (Milled Equivalent)	Algeria	0
3	Rice (Milled Equivalent)	Algeria	0
4	Rice (Milled Equivalent)	Algeria	0
...	...	...	...
375	Rice (Milled Equivalent)	Zimbabwe	0
376	Rice (Milled Equivalent)	Zimbabwe	0
377	Rice (Milled Equivalent)	Zimbabwe	0
378	Rice (Milled Equivalent)	Zimbabwe	0
379	Rice (Milled Equivalent)	Zimbabwe	0

380 rows × 3 columns

In [46]:
```python
rice = rice_prod.groupby('Country')['Value'].sum()
rice.sort_values(ascending=False).plot(kind='bar',figsize=(12,7))
plt.show()
```



Egypt, Nigeria and Madagascar are the highest producer of rice in the given years, while Algeria is the lowest.

In [ ]: