In [11]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [13]:
```
df = pd.read_csv("FoodBalanceSheets_E_Africa_NOFLAG.csv", encoding="latin-1")
df
```

Out[13]:

	Area Code	Area	Item Code	Item	Element Code	Element	Unit	Y2014	Y2015	Y2016	Y2017	Y2018
0	4	Algeria	2501	Population	511	Total Population - Both sexes	1000 persons	38924.00	39728.00	40551.00	41389.00	42228.00
1	4	Algeria	2501	Population	5301	Domestic supply quantity	1000 tonnes	0.00	0.00	0.00	0.00	0.00
2	4	Algeria	2901	Grand Total	664	Food supply (kcal/capita/day)	kcal/capita/day	3377.00	3379.00	3372.00	3341.00	3322.00
3	4	Algeria	2901	Grand Total	674	Protein supply quantity (g/capita/day)	g/capita/day	94.90	94.35	94.72	92.82	91.83
4	4	Algeria	2901	Grand Total	684	Fat supply quantity (g/capita/day)	g/capita/day	80.06	79.36	77.40	80.19	77.28
...	...	...	...	...	...	...	...	...	...	...	...	...
60938	181	Zimbabwe	2899	Miscellaneous	5142	Food	1000 tonnes	42.00	46.00	33.00	19.00	16.00
60939	181	Zimbabwe	2899	Miscellaneous	645	Food supply quantity (kg/capita/yr)	kg	3.06	3.33	2.35	1.33	1.08
60940	181	Zimbabwe	2899	Miscellaneous	664	Food supply (kcal/capita/day)	kcal/capita/day	3.00	4.00	3.00	1.00	1.00
60941	181	Zimbabwe	2899	Miscellaneous	674	Protein supply quantity (g/capita/day)	g/capita/day	0.10	0.11	0.08	0.04	0.04
60942	181	Zimbabwe	2899	Miscellaneous	684	Fat supply quantity (g/capita/day)	g/capita/day	0.04	0.05	0.03	0.02	0.01

60943 rows × 12 columns

In [ ]:
```
group = df[["Element", "Y2014", "Y2015", "Y2016","Y2017", "Y2018" ]]
grouped = group.groupby(by="Element").sum()
grouped.loc["Stock Variation"]
```

Out[ ]:
```
Y2014    58749.83
Y2015    34910.99
Y2016    33140.12
Y2017    54316.91
Y2018    20577.91
Name: Stock Variation, dtype: float64
```

In [23]:
```
year_2017 = df["Y2017"]
year_2017.describe()
```

Out[23]:
```
count     59437.000000
mean        140.917765
std        1671.862359
min       -1582.000000
25%           0.000000
50%           0.100000
75%           9.000000
max      190873.000000
Name: Y2017, dtype: float64
```

In [26]:
```
year_2015 = df[df["Area"] == "Madagascar"]
year_2015 = year_2015[["Y2015", "Element"]]
year_2015.groupby(by="Element").sum()
```

Out[26]:

Element	Y2015
Domestic supply quantity	31214.98
Export Quantity	494.75
Fat supply quantity (g/capita/day)	91.85
Feed	2070.22
Food	21120.65
Food supply (kcal/capita/day)	7685.00
Food supply quantity (kg/capita/yr)	871.59
Import Quantity	1721.80
Losses	2176.00
Other uses (non-food)	3957.12
Processing	1803.00
Production	29482.89
Protein supply quantity (g/capita/day)	173.05
Residuals	-616.00
Seed	699.00
Stock Variation	-505.00
Total Population - Both sexes	24234.00

In [32]:
```
"Processing" in df["Item"].values
```

Out[32]:
```
False
```

In [33]:
```
group = df[["Element Code", "Y2014", "Y2015", "Y2016","Y2017", "Y2018" ]]
group.corr()
```

Out[33]:

	Element Code	Y2014	Y2015	Y2016	Y2017	Y2018
Element Code	1.000000	0.024457	0.023889	0.023444	0.024254	0.024279
Y2014	0.024457	1.000000	0.994647	0.996081	0.995230	0.994872
Y2015	0.023889	0.994647	1.000000	0.995739	0.988048	0.988208
Y2016	0.023444	0.996081	0.995739	1.000000	0.992785	0.992757
Y2017	0.024254	0.995230	0.988048	0.992785	1.000000	0.998103
Y2018	0.024279	0.994872	0.988208	0.992757	0.998103	1.000000

In [37]:
```
group2 = df[["Area", "Y2017"]]
group2.groupby(by="Area").sum().sort_values(by="Y2017", ascending=False)
```

Out[37]:

Area	Y2017
Nigeria	1483268.23
Egypt	866379.92
South Africa	517590.54
Ethiopia	448683.76
Morocco	388495.36
Ghana	337599.06
Algeria	325644.27
United Republic of Tanzania	322616.85
Kenya	264660.66
Sudan	239931.92
Cameroon	232030.43
Angola	229159.57
Côte d'Ivoire	224599.01
Uganda	213950.38
Malawi	181098.71
Mozambique	161407.98
Mauritania	156665.46
Mali	149928.33
Madagascar	131197.73
Niger	126707.58
Benin	124771.22
Tunisia	124167.20
Zambia	103223.77
Burkina Faso	101855.07
Guinea	98138.87
Senegal	95681.15
Zimbabwe	75919.34
Rwanda	73663.69
Chad	71594.68
Sierra Leone	55311.33
Eswatini	54343.33
Mauritius	51114.83

Area	Y2017
Togo	49841.88
Congo	41181.68
Central African Republic	29937.00
Namibia	29874.89
Liberia	29342.20
Gabon	27979.64
Gambia	23154.18
Djibouti	22729.91
Botswana	22101.30
Lesotho	21267.96
Guinea-Bissau	19102.77
Cabo Verde	14650.74
Sao Tome and Principe	12662.63
Seychelles	442.34
Comoros	59.84
Ethiopia PDR	0.00
Sudan (former)	0.00

In [39]:
```
unique = df["Area"].unique()
len(unique)
```

Out[39]: 49

In [44]:
```
missing = df["Y2014"].isnull().sum()
missing
```

Out[44]: np.int64(1589)

In [46]:
```
"Wine" in df["Item"]
```

Out[46]: False

In [47]:
```
year_2017 = df[df["Area"] == "Madagascar"]
year_2017 = year_2017[["Y2017", "Element"]]
year_2017.groupby(by="Element").sum()
```

Out[47]:

Element	Y2017
Domestic supply quantity	31927.87
Export Quantity	566.81
Fat supply quantity (g/capita/day)	101.03
Feed	2059.60
Food	21676.21
Food supply (kcal/capita/day)	7768.00
Food supply quantity (kg/capita/yr)	847.75
Import Quantity	3139.79
Losses	2153.00
Other uses (non-food)	4170.06
Processing	1826.00
Production	29267.86
Protein supply quantity (g/capita/day)	175.75
Residuals	-657.00
Seed	695.00
Stock Variation	-91.00
Total Population - Both sexes	25571.00

In [50]:
```
df2 = pd.DataFrame({"name": ["opey", "demi"]}, index= [x for x in range(1,3)])
df2
```

Out[50]:

	name
1	opey
2	demi