In [13]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
```

In [14]:
```
csv_df = pd.read_csv('C:/Users/USEER/Documents/Hamoye AI Lab/Data Science - Python/Food
print(csv_df)
```

```
       Area Code      Area  Item Code           Item  Element Code  \
0              4   Algeria       2501     Population           511
1              4   Algeria       2501     Population          5301
2              4   Algeria       2901    Grand Total           664
3              4   Algeria       2901    Grand Total           674
4              4   Algeria       2901    Grand Total           684
...          ...       ...        ...            ...           ...
60938        181  Zimbabwe       2899  Miscellaneous          5142
60939        181  Zimbabwe       2899  Miscellaneous           645
60940        181  Zimbabwe       2899  Miscellaneous           664
60941        181  Zimbabwe       2899  Miscellaneous           674
60942        181  Zimbabwe       2899  Miscellaneous           684

                                    Element             Unit     Y2014  \
0               Total Population - Both sexes     1000 persons  38924.00
1                    Domestic supply quantity      1000 tonnes      0.00
2               Food supply (kcal/capita/day)  kcal/capita/day   3377.00
3      Protein supply quantity (g/capita/day)     g/capita/day     94.90
4          Fat supply quantity (g/capita/day)     g/capita/day     80.06
...                                       ...              ...       ...
60938                                  Food      1000 tonnes     42.00
60939     Food supply quantity (kg/capita/yr)              kg      3.06
60940           Food supply (kcal/capita/day)  kcal/capita/day      3.00
60941  Protein supply quantity (g/capita/day)     g/capita/day      0.10
60942      Fat supply quantity (g/capita/day)     g/capita/day      0.04

          Y2015     Y2016     Y2017     Y2018
0      39728.00  40551.00  41389.00  42228.00
1          0.00      0.00      0.00      0.00
2       3379.00   3372.00   3341.00   3322.00
3         94.35     94.72     92.82     91.83
4         79.36     77.40     80.19     77.28
...         ...       ...       ...       ...
60938     46.00     33.00     19.00     16.00
60939      3.33      2.35      1.33      1.08
60940      4.00      3.00      1.00      1.00
60941      0.11      0.08      0.04      0.04
60942      0.05      0.03      0.02      0.01

[60943 rows x 12 columns]
```

In [15]:
```
csv_df.describe()
```

Out[15]:

	Area Code	Item Code	Element Code	Y2014	Y2015	Y2016	
count	60943.000000	60943.000000	60943.000000	59354.000000	59395.000000	59408.000000	59437.0
mean	134.265576	2687.176706	3814.856456	134.196282	135.235966	136.555222	140.9
std	72.605709	146.055739	2212.007033	1567.663696	1603.403984	1640.007194	1671.8

	Area Code	Item Code	Element Code	Y2014	Y2015	Y2016	
min	4.000000	2501.000000	511.000000	-1796.000000	-3161.000000	-3225.000000	-1582.0
25%	74.000000	2562.000000	684.000000	0.000000	0.000000	0.000000	0.0
50%	136.000000	2630.000000	5142.000000	0.090000	0.080000	0.080000	0.1
75%	195.000000	2775.000000	5511.000000	8.340000	8.460000	8.430000	9.0
max	276.000000	2961.000000	5911.000000	176405.000000	181137.000000	185960.000000	190873.0

In [17]:
```
csv_df.isnull().sum()
```

Out[17]:
```
Area Code          0
Area               0
Item Code          0
Item               0
Element Code       0
Element            0
Unit               0
Y2014           1589
Y2015           1548
Y2016           1535
Y2017           1506
Y2018           1436
dtype: int64
```

In [18]:
```
csv_df.groupby('Item')['Item'].count()
```

Out[18]:
```
Item
Alcohol, Non-Food      283
Alcoholic Beverages    564
Animal Products        135
Animal fats            639
Apples and products    517
                      ...
Vegetables, Other      583
Vegetal Products       135
Wheat and products     653
Wine                   497
Yams                   383
Name: Item, Length: 119, dtype: int64
```

In [19]:
```
csv_df.groupby('Item')['Item'].sum()
```

Out[19]:
```
Item
Alcohol, Non-Food      Alcohol, Non-FoodAlcohol, Non-FoodAlcohol, Non...
Alcoholic Beverages    Alcoholic BeveragesAlcoholic BeveragesAlcoholi...
Animal Products        Animal ProductsAnimal ProductsAnimal ProductsA...
Animal fats            Animal fatsAnimal fatsAnimal fatsAnimal fatsAn...
Apples and products    Apples and productsApples and productsApples a...
                                                   ...
Vegetables, Other      Vegetables, OtherVegetables, OtherVegetables, ...
Vegetal Products       Vegetal ProductsVegetal ProductsVegetal Produc...
Wheat and products     Wheat and productsWheat and productsWheat and ...
Wine                   WineWineWineWineWineWineWineWineWineWineWi...
```

```
Yams                    YamsYamsYamsYamsYamsYamsYamsYamsYamsYamsYa...
Name: Item, Length: 119, dtype: object
```

In [20]:
```
csv_df.describe(include='all')
```

Out[20]:

	Area Code	Area	Item Code	Item	Element Code	Element	Unit	Y2014
count	60943.000000	60943	60943.000000	60943	60943.000000	60943	60943	59354.000000
unique	NaN	49	NaN	119	NaN	18	5	NaN
top	NaN	Kenya	NaN	Milk - Excluding Butter	NaN	Domestic supply quantity	1000 tonnes	NaN
freq	NaN	1560	NaN	1262	NaN	5295	40933	NaN
mean	134.265576	NaN	2687.176706	NaN	3814.856456	NaN	NaN	134.196282
std	72.605709	NaN	146.055739	NaN	2212.007033	NaN	NaN	1567.663696
min	4.000000	NaN	2501.000000	NaN	511.000000	NaN	NaN	-1796.000000
25%	74.000000	NaN	2562.000000	NaN	684.000000	NaN	NaN	0.000000
50%	136.000000	NaN	2630.000000	NaN	5142.000000	NaN	NaN	0.090000
75%	195.000000	NaN	2775.000000	NaN	5511.000000	NaN	NaN	8.340000
max	276.000000	NaN	2961.000000	NaN	5911.000000	NaN	NaN	176405.000000

In [21]:
```
csv_df.groupby('Element')['Element'].count()
```

Out[21]:
```
Element
Domestic supply quantity                5295
Export Quantity                         4403
Fat supply quantity (g/capita/day)      5023
Feed                                    1319
Food                                    4941
Food supply (kcal/capita/day)           5014
Food supply quantity (kg/capita/yr)     4905
Import Quantity                         5139
Losses                                  2009
Other uses (non-food)                   1732
Processing                              2010
Production                              3881
Protein supply quantity (g/capita/day)  5023
Residuals                               4655
Seed                                     762
Stock Variation                         4232
Total Population - Both sexes             45
Tourist consumption                       555
Name: Element, dtype: int64
```

In [26]:
```
csv_df.groupby('Area')['Area'].first().count()
```

Out[26]:
```
49
```

In [30]:
```
csv_df.groupby('Y2018')['Y2018'].first().count()
```

Out[30]:
```
4591
```

In [33]:
```
csv_df
```

Out[33]:

	Area Code	Area	Item Code	Item	Element Code	Element	Unit	Y2014	Y20
0	4	Algeria	2501	Population	511	Total Population - Both sexes	1000 persons	38924.00	39728
1	4	Algeria	2501	Population	5301	Domestic supply quantity	1000 tonnes	0.00	0
2	4	Algeria	2901	Grand Total	664	Food supply (kcal/capita/day)	kcal/capita/day	3377.00	3379
3	4	Algeria	2901	Grand Total	674	Protein supply quantity (g/capita/day)	g/capita/day	94.90	94
4	4	Algeria	2901	Grand Total	684	Fat supply quantity (g/capita/day)	g/capita/day	80.06	79
...	...	...	...	...	...	...	...	...	
60938	181	Zimbabwe	2899	Miscellaneous	5142	Food	1000 tonnes	42.00	46
60939	181	Zimbabwe	2899	Miscellaneous	645	Food supply quantity (kg/capita/yr)	kg	3.06	3
60940	181	Zimbabwe	2899	Miscellaneous	664	Food supply (kcal/capita/day)	kcal/capita/day	3.00	4
60941	181	Zimbabwe	2899	Miscellaneous	674	Protein supply quantity (g/capita/day)	g/capita/day	0.10	0
60942	181	Zimbabwe	2899	Miscellaneous	684	Fat supply quantity (g/capita/day)	g/capita/day	0.04	0

60943 rows × 12 columns

In [35]:
```
area = csv_df.Area
```

In [37]:
```
y2018 = csv_df.Y2018
print(y2018)
```

```
0        42228.00
1            0.00
2         3322.00
3           91.83
4           77.28
```

```
           ...
60938       16.00
60939        1.08
60940        1.00
60941        0.04
60942        0.01
Name: Y2018, Length: 60943, dtype: float64
```

In [41]:
```
element = csv_df.Element
element
```

Out[41]:
```
0                 Total Population - Both sexes
1                      Domestic supply quantity
2                 Food supply (kcal/capita/day)
3         Protein supply quantity (g/capita/day)
4             Fat supply quantity (g/capita/day)
                          ...
60938                                      Food
60939       Food supply quantity (kg/capita/yr)
60940             Food supply (kcal/capita/day)
60941    Protein supply quantity (g/capita/day)
60942        Fat supply quantity (g/capita/day)
Name: Element, Length: 60943, dtype: object
```

In [47]:
```
csv_df.groupby('Element').first().count()
```

Out[47]:
```
Area Code       18
Area            18
Item Code       18
Item            18
Element Code    18
Unit            18
Y2014           18
Y2015           18
Y2016           18
Y2017           18
Y2018           18
dtype: int64
```

In [48]:
```
csv_df.groupby('Y2018').first().count()
```

Out[48]:
```
Area Code       4591
Area            4591
Item Code       4591
Item            4591
Element Code    4591
Element         4591
Unit            4591
Y2014           4591
Y2015           4591
Y2016           4591
Y2017           4591
dtype: int64
```

In [50]:
```
csv_df.groupby('Element').first('Import Quantity').count()
```

Out[50]:
```
Area Code       18
Item Code       18
Element Code    18
Y2014           18
```

```
Y2015           18
Y2016           18
Y2017           18
Y2018           18
dtype: int64
```

In [52]:
```
csv_df.groupby('Element Code').first()
```

Out[52]:

Element Code	Area Code	Area	Item Code	Item	Element	Unit	Y2014	Y2015	Y2016
511	4	Algeria	2501	Population	Total Population - Both sexes	1000 persons	38924.00	39728.00	40551.00
645	4	Algeria	2905	Cereals - Excluding Beer	Food supply quantity (kg/capita/yr)	kg	219.56	219.72	221.98
664	4	Algeria	2901	Grand Total	Food supply (kcal/capita/day)	kcal/capita/day	3377.00	3379.00	3372.00
674	4	Algeria	2901	Grand Total	Protein supply quantity (g/capita/day)	g/capita/day	94.90	94.35	94.72
684	4	Algeria	2901	Grand Total	Fat supply quantity (g/capita/day)	g/capita/day	80.06	79.36	77.40
5072	4	Algeria	2905	Cereals - Excluding Beer	Stock Variation	1000 tonnes	684.00	1663.00	905.00
5123	4	Algeria	2905	Cereals - Excluding Beer	Losses	1000 tonnes	1084.00	1203.00	1148.00
5131	4	Algeria	2905	Cereals - Excluding Beer	Processing	1000 tonnes	35.00	38.00	37.00
5142	4	Algeria	2905	Cereals - Excluding Beer	Food	1000 tonnes	8546.00	8729.00	9001.00
5154	4	Algeria	2905	Cereals - Excluding Beer	Other uses (non-food)	1000 tonnes	787.00	847.00	835.00
5170	4	Algeria	2905	Cereals - Excluding Beer	Residuals	1000 tonnes	0.00	0.00	0.00
5171	7	Angola	2913	Oilcrops	Tourist consumption	1000 tonnes	0.00	0.00	0.00
5301	4	Algeria	2501	Population	Domestic supply quantity	1000 tonnes	0.00	0.00	0.00

Element Code	Area Code	Area	Item Code	Item	Element	Unit	Y2014	Y2015	Y2016
5511	4	Algeria	2905	Cereals - Excluding Beer	Production	1000 tonnes	3436.00	3761.00	3445.00
5521	4	Algeria	2905	Cereals - Excluding Beer	Feed	1000 tonnes	4725.00	5070.00	4838.00
5527	4	Algeria	2905	Cereals - Excluding Beer	Seed	1000 tonnes	177.00	189.00	197.00
5611	4	Algeria	2905	Cereals - Excluding Beer	Import Quantity	1000 tonnes	12604.00	13983.00	13523.00
5911	4	Algeria	2905	Cereals - Excluding Beer	Export Quantity	1000 tonnes	3.00	7.00	7.00

In [60]:
```
csv_df.groupby('Item')['Item'].first('Animal Fat').count()
```

Out[60]:
```
119
```

In [66]:
```
csv_df.groupby('Element').first().sum()
```

Out[66]:
```
Area Code                                                    75
Area            AlgeriaAlgeriaAlgeriaAlgeriaAlgeriaAlgeriaAlge...
Item Code                                                 51478
Item            PopulationCereals - Excluding BeerGrand TotalC...
Element Code                                              72523
Unit            1000 tonnes1000 tonnesg/capita/day1000 tonnes1...
Y2014                                                  74776.52
Y2015                                                  78990.43
Y2016                                                   78253.1
Y2017                                                  78155.03
Y2018                                                  84430.42
dtype: object
```

In [78]:
```
csv_df.groupby('Element')['Element'].first('Y2018').sum()
```

Out[78]:
```
'Domestic supply quantityExport QuantityFat supply quantity (g/capita/day)FeedFoodFood s
upply (kcal/capita/day)Food supply quantity (kg/capita/yr)Import QuantityLossesOther use
s (non-food)ProcessingProductionProtein supply quantity (g/capita/day)ResidualsSeedStock
VariationTotal Population - Both sexesTourist consumption'
```

In [79]:
```
csv_df.groupby('Element')['Element'].count()
```

Out[79]:
```
Element
Domestic supply quantity                  5295
```

```
Export Quantity                          4403
Fat supply quantity (g/capita/day)       5023
Feed                                     1319
Food                                     4941
Food supply (kcal/capita/day)            5014
Food supply quantity (kg/capita/yr)      4905
Import Quantity                          5139
Losses                                   2009
Other uses (non-food)                    1732
Processing                               2010
Production                               3881
Protein supply quantity (g/capita/day)   5023
Residuals                                4655
Seed                                      762
Stock Variation                          4232
Total Population - Both sexes              45
Tourist consumption                       555
Name: Element, dtype: int64
```

In [ ]: