

Name : Emebet Worku

Id:1558b526e5c1f000

```
In [2]: import numpy as np
import pandas as pd

df = pd.read_csv(r'C:\Users\user 1\Downloads\FoodBalanceSheets_E_Africa_NOFLAG.csv', encoding='utf-8')
df.head()
```

Out[2]:

	Area Code	Area	Item Code	Item	Element Code	Element	Unit	Y2014	Y2015	Y2016	Y2017
0	4	Algeria	2501	Population	511	Total Population - Both sexes	1000 persons	38924.00	39728.00	40551.00	41380.00
1	4	Algeria	2501	Population	5301	Domestic supply quantity	1000 tonnes	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	3340.00
3	4	Algeria	2901	Grand Total	674	Protein supply quantity (g/capita/day)	g/capita/day	94.90	94.35	94.72	94.00
4	4	Algeria	2901	Grand Total	684	Fat supply quantity (g/capita/day)	g/capita/day	80.06	79.36	77.40	78.00

question 11

```
In [4]: df_item = df.groupby('Item')
```

Out[4]: <pandas.core.groupby.generic.DataFrameGroupBy object at 0x000001E92CB5BF70>

```
In [6]: df_item['Item'].unique()
```

Out[6]:

Item	
Alcohol, Non-Food	[Alcohol, Non-Food]
Alcoholic Beverages	[Alcoholic Beverages]
Animal Products	[Animal Products]
Animal fats	[Animal fats]
Apples and products	[Apples and products]
...	
Vegetables, Other	[Vegetables, Other]
Vegetal Products	[Vegetal Products]
Wheat and products	[Wheat and products]
Wine	[Wine]
Yams	[Yams]

Name: Item, Length: 119, dtype: object

```
In [10]: sum2014 = df_item.get_group('Animal fats')['Y2014'].sum()
sum2014
```

Out[10]: 209460.54

```
In [11]: sum2017 = df_item.get_group('Animal fats')['Y2017'].sum()
sum2017
```

Out[11]: 269617.53

In [15]: sum2014, sum2017

Out[15]: (209460.54, 269617.53)

question 12

In [16]: round(df['Y2015'].mean(),3), round(df['Y2015'].std(),3)

Out[16]: (135.236, 1603.404)

In [17]: df.describe()

Out[17]:

	Area Code	Item Code	Element Code	Y2014	Y2015	Y2016	Y2017
count	60943.000000	60943.000000	60943.000000	59354.000000	59395.000000	59408.000000	59437.000000
mean	134.265576	2687.176706	3814.856456	134.196282	135.235966	136.555222	140.917765
std	72.605709	146.055739	2212.007033	1567.663696	1603.403984	1640.007194	1671.862359
min	4.000000	2501.000000	511.000000	-1796.000000	-3161.000000	-3225.000000	-1582.000000
25%	74.000000	2562.000000	684.000000	0.000000	0.000000	0.000000	0.000000
50%	136.000000	2630.000000	5142.000000	0.090000	0.080000	0.080000	0.100000
75%	195.000000	2775.000000	5511.000000	8.340000	8.460000	8.430000	9.000000
max	276.000000	2961.000000	5911.000000	176405.000000	181137.000000	185960.000000	190873.000000

question 13

In [18]: df['Y2016'].isnull().sum(), round(df['Y2016'].isnull().sum()*100/len(df['Y2016']),2)

Out[18]: (1535, 2.52)

question 14

In [19]: df.corr()['Element Code']

Out[19]:

Area Code	-0.000209
Item Code	-0.024683
Element Code	1.000000
Y2014	0.024457
Y2015	0.023889
Y2016	0.023444
Y2017	0.024254
Y2018	0.024279

Name: Element Code, dtype: float64

question 15

In [20]: df_element = df.groupby('Element')
df['Element'].unique()

```
Out[20]: array(['Total Population - Both sexes', 'Domestic supply quantity',  
        'Food supply (kcal/capita/day)', 'Protein supply quantity (g/capita/day)',  
        'Fat supply quantity (g/capita/day)', 'Production',  
        'Import Quantity', 'Stock Variation', 'Export Quantity', 'Feed',  
        'Seed', 'Losses', 'Processing', 'Other uses (non-food)',  
        'Residuals', 'Food', 'Food supply quantity (kg/capita/yr)',  
        'Tourist consumption'], dtype=object)
```

```
In [21]: years = ['Y2014', 'Y2015', 'Y2016', 'Y2017']  
for year in years:  
    print(df_element.get_group('Import Quantity')[year].sum())
```

```
274144.48  
267018.45999999996  
286582.78  
294559.09
```

question 16

```
In [22]: df['Element'].unique()
```

```
Out[22]: array(['Total Population - Both sexes', 'Domestic supply quantity',  
        'Food supply (kcal/capita/day)', 'Protein supply quantity (g/capita/day)',  
        'Fat supply quantity (g/capita/day)', 'Production',  
        'Import Quantity', 'Stock Variation', 'Export Quantity', 'Feed',  
        'Seed', 'Losses', 'Processing', 'Other uses (non-food)',  
        'Residuals', 'Food', 'Food supply quantity (kg/capita/yr)',  
        'Tourist consumption'], dtype=object)
```

```
In [23]: df_element.get_group('Production')['Y2014'].sum()
```

```
Out[23]: 1931287.75
```

question 17

```
In [25]: df_sum = df_element['Y2018'].sum()  
df_sum
```

```
Out[25]: Element  
Domestic supply quantity      2161192.10  
Export Quantity               181594.80  
Fat supply quantity (g/capita/day)    10258.69  
Feed                          233489.68  
Food                          1303841.28  
Food supply (kcal/capita/day)    455261.00  
Food supply quantity (kg/capita/yr)   49056.85  
Import Quantity               287997.09  
Losses                       163902.00  
Other uses (non-food)          91300.97  
Processing                    308429.00  
Production                    2075072.89  
Protein supply quantity (g/capita/day)  11833.56  
Residuals                     34864.00  
Seed                          25263.14  
Stock Variation                20577.91  
Total Population - Both sexes    1140605.00  
Tourist consumption             90.00  
Name: Y2018, dtype: float64
```

```
In [26]: df_sum.sort_values(ascending=False)
```

```
Out[26]:
```

Element	
Domestic supply quantity	2161192.10
Production	2075072.89
Food	1303841.28
Total Population - Both sexes	1140605.00
Food supply (kcal/capita/day)	455261.00
Processing	308429.00
Import Quantity	287997.09
Feed	233489.68
Export Quantity	181594.80
Losses	163902.00
Other uses (non-food)	91300.97
Food supply quantity (kg/capita/yr)	49056.85
Residuals	34864.00
Seed	25263.14
Stock Variation	20577.91
Protein supply quantity (g/capita/day)	11833.56
Fat supply quantity (g/capita/day)	10258.69
Tourist consumption	90.00

Name: Y2018, dtype: float64

question 18

```
In [27]: df_sum.sort_values(ascending=True)
```

```
Out[27]:
```

Element	
Tourist consumption	90.00
Fat supply quantity (g/capita/day)	10258.69
Protein supply quantity (g/capita/day)	11833.56
Stock Variation	20577.91
Seed	25263.14
Residuals	34864.00
Food supply quantity (kg/capita/yr)	49056.85
Other uses (non-food)	91300.97
Losses	163902.00
Export Quantity	181594.80
Feed	233489.68
Import Quantity	287997.09
Processing	308429.00
Food supply (kcal/capita/day)	455261.00
Total Population - Both sexes	1140605.00
Food	1303841.28
Production	2075072.89
Domestic supply quantity	2161192.10

Name: Y2018, dtype: float64

question 19

```
In [28]: df.columns
```

```
Out[28]:
```

Index(['Area Code', 'Area', 'Item Code', 'Item', 'Element Code', 'Element', 'Unit', 'Y2014', 'Y2015', 'Y2016', 'Y2017', 'Y2018'], dtype='object')

```
In [29]: df['Element'].unique()
```

```
Out[29]:
```

array(['Total Population - Both sexes', 'Domestic supply quantity', 'Food supply (kcal/capita/day)', 'Protein supply quantity (g/capita/day)', 'Fat supply quantity (g/capita/day)', 'Production', 'Import Quantity', 'Stock Variation', 'Export Quantity', 'Feed', 'Seed', 'Losses', 'Processing', 'Other uses (non-food)', 'Residuals', 'Food', 'Food supply quantity (kg/capita/yr)', 'Tourist consumption'], dtype=object)

```
In [30]: df_element.get_group('Import Quantity')[df['Area']=='Algeria']['Y2018'].sum()
```

```
C:\Users\user 1\AppData\Local\Temp\ipykernel_25256\1432886645.py:1: UserWarning: Boolean Series key will be reindexed to match DataFrame index.
```

```
df_element.get_group('Import Quantity')[df['Area']=='Algeria']['Y2018'].sum()
```

```
Out[30]: 36238.29
```

question 20

```
In [32]: len(df['Area'].unique())
```

```
Out[32]: 49
```

```
In [ ]:
```