```
from google.colab import files
uploaded = files.upload()
```

Choose Files FoodBalanc…OFLAG.csv

- **FoodBalanceSheets_E_Africa_NOFLAG.csv**(text/csv) - 8603007 bytes, last modified: 11/21/2022 - 100% done

Saving FoodBalanceSheets_E_Africa_NOFLAG.csv to FoodBalanceSheets_E_Africa_NOFLAG.csv

```
import pandas as pd
import numpy as np



df = pd.read_csv('FoodBalanceSheets_E_Africa_NOFLAG.csv', encoding='latin-1')


df.head()
```

	Area Code	Area	Item Code	Item	Element Code	Element	Unit	Y2014	Y201
0	4	Algeria	2501	Population	511	Total Population - Both sexes	1000 persons	38924.00	39728.0
1	4	Algeria	2501	Population	5301	Domestic supply quantity	1000 tonnes	0.00	0.0
2	4	Algeria	2901	Grand Total	664	Food supply (kcal/capita/day)	kcal/capita/day	3377.00	3379.0

```
# Question 1 - Select columns ‘Y2017’ and ‘Area’, Perform a groupby operation on ‘Area’.  Whi

df.groupby('Area')['Y2017'].sum()
```

```
Area
Algeria                         325644.27
Angola                          229159.57
Benin                           124771.22
Botswana                         22101.30
Burkina Faso                    101855.07
Cabo Verde                       14650.74
Cameroon                        232030.43
Central African Republic         29937.00
Chad                             71594.68
Comoros                             59.84
Congo                            41181.68
Côte d'Ivoire                   224599.01
Djibouti                         22729.91
Egypt                           866379.92
Eswatini                         54343.33
Ethiopia                        448683.76
Ethiopia PDR                         0.00
Gabon                            27979.64
```

```
Gambia                          23154.18
Ghana                          337599.06
Guinea                          98138.87
Guinea-Bissau                   19102.77
Kenya                          264660.66
Lesotho                         21267.96
Liberia                         29342.20
Madagascar                     131197.73
Malawi                         181098.71
Mali                           149928.33
Mauritania                     156665.46
Mauritius                       51114.83
Morocco                        388495.36
Mozambique                     161407.98
Namibia                         29874.89
Niger                          126707.58
Nigeria                       1483268.23
Rwanda                          73663.69
Sao Tome and Principe           12662.63
Senegal                         95681.15
Seychelles                        442.34
Sierra Leone                    55311.33
South Africa                   517590.54
Sudan                          239931.92
Sudan (former)                      0.00
Togo                            49841.88
Tunisia                        124167.20
Uganda                         213950.38
United Republic of Tanzania    322616.85
Zambia                         103223.77
Zimbabwe                        75919.34
Name: Y2017, dtype: float64
```

```
# Question 2 - What is the total sum of Wine produced in 2015 and 2018 respectively?
df.groupby('Item')['Y2015'].sum()
```

```
Item
Alcohol, Non-Food        2180.00
Alcoholic Beverages     98783.72
Animal Products         11811.73
Animal fats            200675.72
Apples and products     10559.15
                           ...
Vegetables, Other      158104.08
Vegetal Products       107064.17
Wheat and products     234710.51
Wine                     4251.81
Yams                   203151.78
Name: Y2015, Length: 119, dtype: float64
```

```
df.groupby('Item')['Y2018'].sum()
```

```
Item
Alcohol, Non-Food        2293.00
```

```
Alcoholic Beverages     97847.27
Animal Products         11578.61
Animal fats            269648.27
Apples and products      9640.51
                          ...
Vegetables, Other      163987.21
Vegetal Products       107775.39
Wheat and products     242645.19
Wine                     4039.32
Yams                   221272.09
Name: Y2018, Length: 119, dtype: float64
```

```
# Question 3 -Select columns ‘Y2017’ and ‘Area’, Perform a groupby operation on ‘Area’.  Whic
```

```
df.groupby('Area')['Y2017'].sum()
```

```
Area
Algeria                         325644.27
Angola                          229159.57
Benin                           124771.22
Botswana                         22101.30
Burkina Faso                    101855.07
Cabo Verde                       14650.74
Cameroon                        232030.43
Central African Republic         29937.00
Chad                             71594.68
Comoros                             59.84
Congo                            41181.68
Côte d'Ivoire                   224599.01
Djibouti                         22729.91
Egypt                           866379.92
Eswatini                         54343.33
Ethiopia                        448683.76
Ethiopia PDR                         0.00
Gabon                            27979.64
Gambia                           23154.18
Ghana                           337599.06
Guinea                           98138.87
Guinea-Bissau                    19102.77
Kenya                           264660.66
Lesotho                          21267.96
Liberia                          29342.20
Madagascar                      131197.73
Malawi                          181098.71
Mali                            149928.33
Mauritania                      156665.46
Mauritius                        51114.83
Morocco                         388495.36
Mozambique                      161407.98
Namibia                          29874.89
Niger                           126707.58
Nigeria                        1483268.23
Rwanda                           73663.69
Sao Tome and Principe            12662.63
Senegal                          95681.15
```

```
Seychelles                      442.34
Sierra Leone                  55311.33
South Africa                 517590.54
Sudan                        239931.92
Sudan (former)                    0.00
Togo                          49841.88
Tunisia                      124167.20
Uganda                       213950.38
United Republic of Tanzania  322616.85
Zambia                       103223.77
Zimbabwe                      75919.34
Name: Y2017, dtype: float64
```

```
# Question 5 - How would you assign element 8 from the list to a variable x?

y = [(2, 4), (7, 8), (1, 5, 9)]
y[1][-1]
```

```
8
```

```
# Question 6 - Which year had the least correlation with 'Element Code'?

df.corr()
```

	Area Code	Item Code	Element Code	Y2014	Y2015	Y2016	Y2017
Area Code	1.000000	-0.005159	-0.000209	0.006164	0.005472	0.005247	0.005006
Item Code	-0.005159	1.000000	-0.024683	0.021722	0.020857	0.020109	0.021494
Element Code	-0.000209	-0.024683	1.000000	0.024457	0.023889	0.023444	0.024254
Y2014	0.006164	0.021722	0.024457	1.000000	0.994647	0.996081	0.995230
Y2015	0.005472	0.020857	0.023889	0.994647	1.000000	0.995739	0.988048
Y2016	0.005247	0.020109	0.023444	0.996081	0.995739	1.000000	0.992785
Y2017	0.005006	0.021494	0.024254	0.995230	0.988048	0.992785	1.000000
Y2018	0.005665	0.021314	0.024279	0.994872	0.988208	0.992757	0.998103

```
# Question 7
my_tuppy = (1,2,5,8)
my_tuppy[2] = 6
```

```
---------------------------------------------------------------------------
TypeError                                 Traceback (most recent call last)
<ipython-input-100-bd8c85ffe01b> in <module>
      1 # Question 7
      2 my_tuppy = (1,2,5,8)
----> 3 my_tuppy[2] = 6
```

```
# Question 8
array  = ([[94, 89, 63], [93, 92, 48], [92, 94, 56]])

array[:1,1:]
```

```
---------------------------------------------------------------------------
TypeError                                 Traceback (most recent call last)
<ipython-input-101-fbda66c6ecab> in <module>
      2 array  = ([[94, 89, 63], [93, 92, 48], [92, 94, 56]])
      3
----> 4 array[:1,1:]

TypeError: list indices must be integers or slices, not tuple
```

SEARCH STACK OVERFLOW

```
# Question 9 - Perform a groupby operation on ‘Element’.  What is the total number of the sum
df.groupby('Element').sum()
```

Element	Area Code	Item Code	Element Code	Y2014	Y2015	Y2016	Y2
Domestic supply quantity	708993	14197445	28068795	1996716.35	2021493.55	2044842.70	2088198
Export Quantity	599910	11840553	26026133	150020.64	157614.47	151920.46	182338

```
# Question 10 - What is the total Protein supply quantity in Madagascar in 2015?

df.groupby('Area')['Y2015'].sum()
```

```
Area
Algeria                        324058.40
Angola                         209565.67
Benin                          108181.13
Botswana                        23024.85
Burkina Faso                   102701.22
Cabo Verde                      14538.93
Cameroon                       225220.72
Central African Republic        28885.34
Chad                            65835.88
Comoros                            41.34
Congo                           40255.51
Côte d'Ivoire                  211591.29
Djibouti                        23690.16
Egypt                          868218.73
Eswatini                        54072.07
Ethiopia                       429139.33
Ethiopia PDR                        0.00
Gabon                           28367.30
Gambia                          19665.39
Ghana                          311092.13
Guinea                          92803.39
Guinea-Bissau                   17990.06
Kenya                          265506.19
Lesotho                         17965.84
Liberia                         28698.10
Madagascar                     126674.90
Malawi                         167989.09
Mali                           132456.58
Mauritania                      61909.99
Mauritius                       53372.42
Morocco                        412473.80
Mozambique                     158231.17
Namibia                         30687.90
Niger                          117834.65
Nigeria                       1414362.83
Rwanda                          70138.83
Sao Tome and Principe           12386.93
Senegal                         91939.14
Seychelles                        358.20
Sierra Leone                    62523.14
```

```
South Africa                    468352.18
Sudan                           230652.54
Sudan (former)                       0.00
Togo                             47834.59
Tunisia                         128211.41
Uganda                          220615.72
United Republic of Tanzania     341969.46
Zambia                           96214.00
Zimbabwe                         74041.79
Name: Y2015, dtype: float64
```

```
# Question 12 - number rows and columns
df.shape
```

```
(60943, 12)
```

```
# Question 14 - How do you create a pandas DataFrame using this list
lst = [[35, 'Portugal', 94], [33, 'Argentina', 93], [30 , 'Brazil', 92]]

col = ['Age','Nationality','Overall']

pd.DataFrame(lst, columns = col, index = [1,2,3])
```

	Age	Nationality	Overall
1	35	Portugal	94
2	33	Argentina	93
3	30	Brazil	92

```
# Question 15 - What is the mean and standard deviation across the whole dataset for the year
df.describe()
```

	Area Code	Item Code	Element Code	Y2014	Y2015	Y2(
count	60943.000000	60943.000000	60943.000000	59354.000000	59395.000000	59408.000(
mean	134.265576	2687.176706	3814.856456	134.196282	135.235966	136.555(
std	72.605709	146.055739	2212.007033	1567.663696	1603.403984	1640.007
min	4.000000	2501.000000	511.000000	-1796.000000	-3161.000000	-3225.000(
25%	74.000000	2562.000000	684.000000	0.000000	0.000000	0.000(
50%	136.000000	2630.000000	5142.000000	0.090000	0.080000	0.080(
75%	195.000000	2775.000000	5511.000000	8.340000	8.460000	8.430(
max	276.000000	2961.000000	5911.000000	176405.000000	181137.000000	185960.000(

```
# Question 16 - Perform a groupby operation on 'Element'.  What year has the highest sum of S

df.groupby('Element').sum()
```

Element	Area Code	Item Code	Element Code	Y2014	Y2015	Y2016	Y2(
Domestic supply quantity	708993	14197445	28068795	1996716.35	2021493.55	2044842.70	2088198
Export Quantity	599910	11840553	26026133	150020.64	157614.47	151920.46	182338
Fat supply quantity (g/capita/day)	675050	13535000	3435732	10225.56	10235.74	10102.77	10253
Feed	176272	3538507	7282199	216927.89	225050.22	228958.65	223705
Food	663295	13285035	25406622	1212332.49	1232361.10	1247022.17	1258888
Food supply (kcal/capita/day)	674057	13511060	3329296	454257.00	453383.00	451810.00	454681
Food supply quantity (kg/capita/yr)	658446	13185401	3163725	49650.63	49345.13	48985.28	48690
Import Quantity	688174	13795966	28834929	274144.48	267018.46	286582.78	294559
Losses	274353	5424803	10292107	153223.00	155439.00	157787.00	160614
Other uses (non-food)	235554	4729749	8926728	78718.13	66254.41	69563.68	91645
Processing	271940	5350416	10313310	282923.00	287929.00	280631.00	292836
Production	526751	10450053	21388191	1931287.75	1947019.39	1943537.15	2030056
Protein supply quantity (g/capita/day)	675050	13535000	3385502	11836.46	11833.95	11779.69	11842

```
# Question 17 - What is the total number of unique countries in the dataset?

# counting unique values
len(pd.unique(df['Area']))
```

```
49
```

```
# Question 18 - What is the total number and percentage of missing data in 2014 to 3 decimal
df.isnull().sum()
```

```
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
```

```
# Question 20 - What would be the output for?
S = [['him', 'sell'], [90, 28, 43]]

S[0][1][1]
```

```
'e'
```