Week 8- TASK

November 15, 2021

1 Global Food & Agriculture Statistics

Data Source: Global Food Dataset from Kaggle Name: Mohammad Aziz, Student Id: 20192233

[1]:
```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
```

[3]:
```
df = pd.read_csv('fao_data_production_indices_data.csv')
```

[4]:
```
df.head()
```

[4]:

	country_or_area	element_code	element	year	unit	value	value_footnotes	category
0	Afghanistan	152	Gross Production 1999-2001 (1000 I$)	2007.0	1000 Int. $	2486910.0	A	agriculture_pin
1	Afghanistan	152	Gross Production 1999-2001 (1000 I$)	2006.0	1000 Int. $	2278516.0	A	agriculture_pin
2	Afghanistan	152	Gross Production 1999-2001 (1000 I$)	2005.0	1000 Int. $	2524097.0	A	agriculture_pin
3	Afghanistan	152	Gross Production 1999-2001 (1000 I$)	2004.0	1000 Int. $	2226346.0	A	agriculture_pin
4	Afghanistan	152	Gross Production 1999-2001 (1000 I$)	2003.0	1000 Int. $	2289434.0	A	agriculture_pin

[ ]:
```

```

2 TASK: Do we have missing data in this dataset if so how many? and for which features?

[11]:
```
df.isnull().sum()
```

[11]:
```
country_or_area     0
element_code        0
element            10
```

```
year                10
unit                10
value               10
value_footnotes     10
category             0
dtype: int64
```

[ ]:
```
## yes we have missing data
```

[21]:
```
print("Total missing data is {}.".format(sum(df.isnull().sum())))
```

```
Total missing data is 50.
```

3 TASK: How many Unique Geographical locations are captured in this dataset

[24]:
```
print(len(df['country_or_area'].unique()))
```

```
254
```

4 What are the Element codes for the following Geographical locations?:

- Belgium
- Africa +
- Cayman Islands
- Georgia
- France

[ ]:
```

```

[138]:
```
qwerty=df.loc[df["country_or_area"].isin(["Belgium", "Africa +", "Cayman␣
↪Islands", "Georgia", "France"]), ["country_or_area", "element_code"]]
```

[140]:
```
qwerty.drop_duplicates()
```

[140]:

	country_or_area	element_code
282	Africa +	152
329	Africa +	154
376	Africa +	432
423	Africa +	434
470	Africa +	436
517	Africa +	438
5082	Belgium	152
5090	Belgium	154
5098	Belgium	432

```
5106          Belgium          434
5114          Belgium          436
5122          Belgium          438
10254  Cayman Islands          152
10301  Cayman Islands          154
10348  Cayman Islands          432
10395  Cayman Islands          434
10442  Cayman Islands          436
10489  Cayman Islands          438
20150          France          152
20197          France          154
20244          France          432
20291          France          434
20338          France          436
20385          France          438
21560         Georgia          152
21576         Georgia          154
21592         Georgia          432
21608         Georgia          434
21624         Georgia          436
21640         Georgia          438
```

5 How many categories do we have? and what are they?

[109]:
```
print("Thet are total",len(df["category"].unique()))
df.category.unique()
```

```
Thet are total 5
```

[109]:
```
array(['agriculture_pin', 'crops_pin', 'food_pin', 'livestock_pin',
       'non_food_pin'], dtype=object)
```

[ ]:

[ ]:

6 What is the total value of livestock in France?

hint .sum()

[63]:
```
a=df.loc[(df["country_or_area"]=="France") & (df["category"]==␣
↪"livestock_pin"), ["country_or_area", "category", "value"]]
```

[71]:
```
print("The value of livestock in france is {}".format(sum(a["value"])))
```

```
The value of livestock in france is 1404737251.0
```

[ ]:

7 Explore the data

[ ]: