

## Assignment On Case-Study --- Tutorial 26

### Annual population of neighborhood country of Pakistan


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

### Read CSV file using pandas

```
In [ ]: fd = pd.read_csv("Annual_Population.csv")
fd.head(4)
# fd.to_numpy()
```

Out [ ]:

	Domain Code	Domain	Area Code (FAO)	Area	Element Code	Element	Item Code	Item	Year Code	Year	Unit
0	OA	Annual population	2	Afghanistan	511	Total Population - Both sexes	3010	Population - Est. & Proj.	2014	2014	1 pers
1	OA	Annual population	2	Afghanistan	512	Total Population - Male	3010	Population - Est. & Proj.	2014	2014	1 pers
2	OA	Annual population	2	Afghanistan	513	Total Population - Female	3010	Population - Est. & Proj.	2014	2014	1 pers
3	OA	Annual population	2	Afghanistan	551	Rural population	3010	Population - Est. & Proj.	2014	2014	1 pers



### Drop coloumn from dataframe for unwanted data

```
In [ ]: population = fd.drop(["Domain Code", "Domain", "Item", "Unit", "Year Code", "Flag"], axis=1)
population = population.drop(["Flag Description", "Item Code"], axis=1)
population.head(3)
```

Out [ ]:

	Area Code (FAO)	Area	Element Code	Element	Year	Value	Note
0	2	Afghanistan	511	Total Population - Both sexes	2014	33370.794	NaN
1	2	Afghanistan	512	Total Population - Male	2014	17138.803	NaN
2	2	Afghanistan	513	Total Population - Female	2014	16232.001	NaN

### Change/Replace the value to short meaningful name

```
In [ ]: population["Element"].replace({"Urban population": "Urban", "Rural population": "Rural"}, inplace=True)
population["Note"].replace({"UNDESA, Population Division ? World Urbanization Prospects": "UNDESA, Population Division ? World Urbanization Prospects"}, inplace=True)
population.head()
```

```
Out[ ]:
```

	Area Code (FAO)	Area	Element Code	Element	Year	Value	Note
0	2	Afghanistan	511	Both_Sex	2014	33370.794	NaN
1	2	Afghanistan	512	Male	2014	17138.803	NaN
2	2	Afghanistan	513	Female	2014	16232.001	NaN
3	2	Afghanistan	551	Rural	2014	24703.798	NaN
4	2	Afghanistan	561	Urban	2014	8054.222	UNDESA

### dataframe to excel format

```
In [ ]: population.to_excel("Neighbourhoob Annual Population.xlsx")
```

### Find means population and count catogries

```
In [ ]: population.mean()
print(population.value_counts( population["Element"] == "Urban" ))
print(population.value_counts( population["Area"] == "Pakistan" ))
```

```
Element
False    100
True      25
dtype: int64
Area
False    100
True      25
dtype: int64
```

### Finding means of each group

```
In [ ]: population.groupby(["Area", "Value"]).mean()
population.groupby(["Element", "Value"]).mean()

population[population["Note"] == "UNDESA"].groupby(["Area", "Element", "Value"]).mean
```

```
Out[ ]:
```

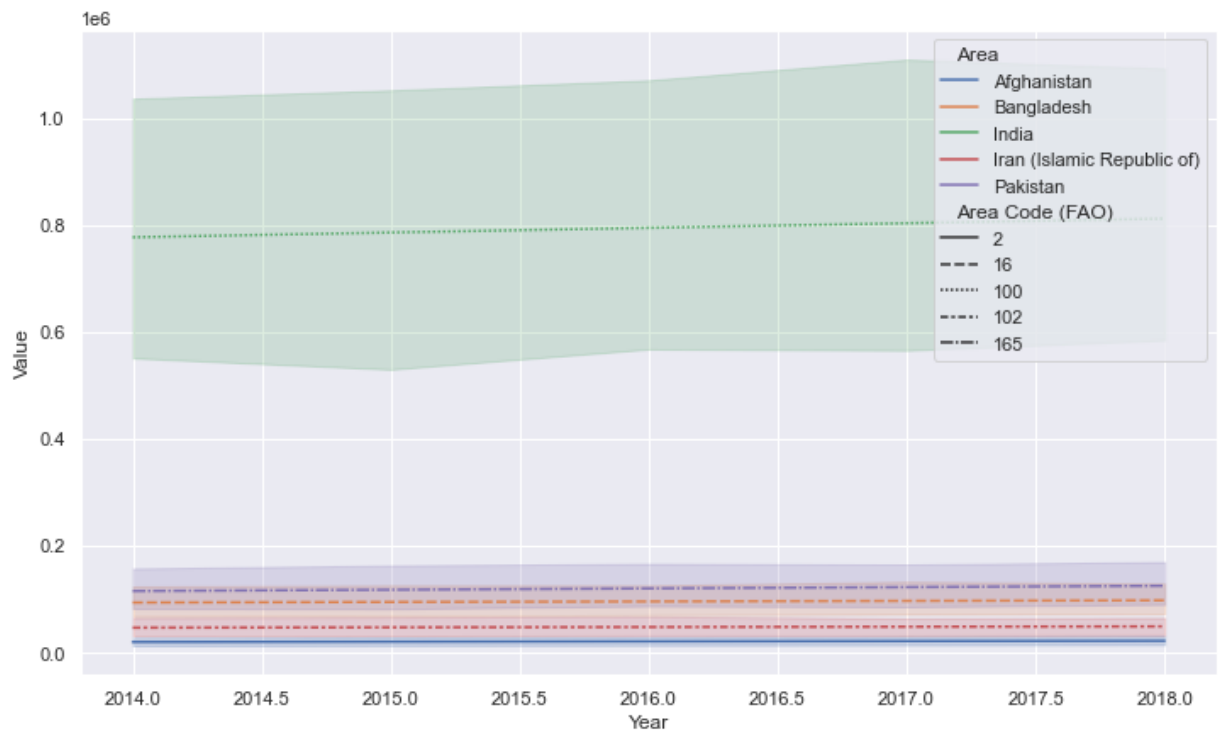
	Area Code (FAO)	Element Code	Year
	Area	Element	Value
Afghanistan	Urban	8054.222	2 561 2014
		8367.571	2 561 2015
		8670.939	2 561 2016
		8971.472	2 561 2017
		9273.302	2 561 2018
Bangladesh	Urban	53455.788	16 561 2014
		55305.132	16 561 2015
		57168.030	16 561 2016
		59046.818	16 561 2017
		60944.245	16 561 2018
India	Urban	419001.673	100 561 2014

			Area Code (FAO)	Element Code	Year
Area	Element	Value			
Iran (Islamic Republic of)	Urban	429069.459	100	561	2015
		439391.699	100	561	2016
		449963.381	100	561	2017
		460779.764	100	561	2018
	Rural	57106.555	102	561	2014
		58217.032	102	561	2015
		59308.845	102	561	2016
		60380.188	102	561	2017
		61425.055	102	561	2018
		66461.625	165	561	2014
Pakistan	Urban	68226.783	165	561	2015
		70005.271	165	561	2016
		71795.700	165	561	2017
		73630.430	165	561	2018

## Plots

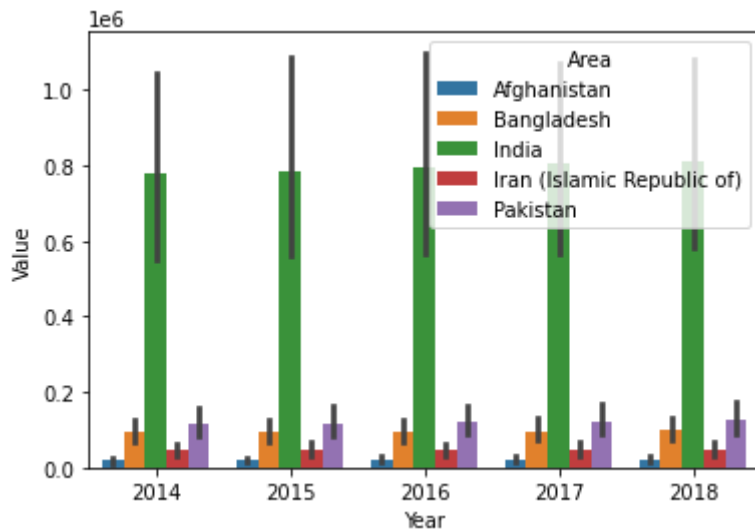
In [ ]:

```
plt.figure(figsize= ( 12, 7))
sns.lineplot(x="Year", y="Value",
             hue="Area", style="Area Code (FAO)",
             data=population)
plt.show()
```

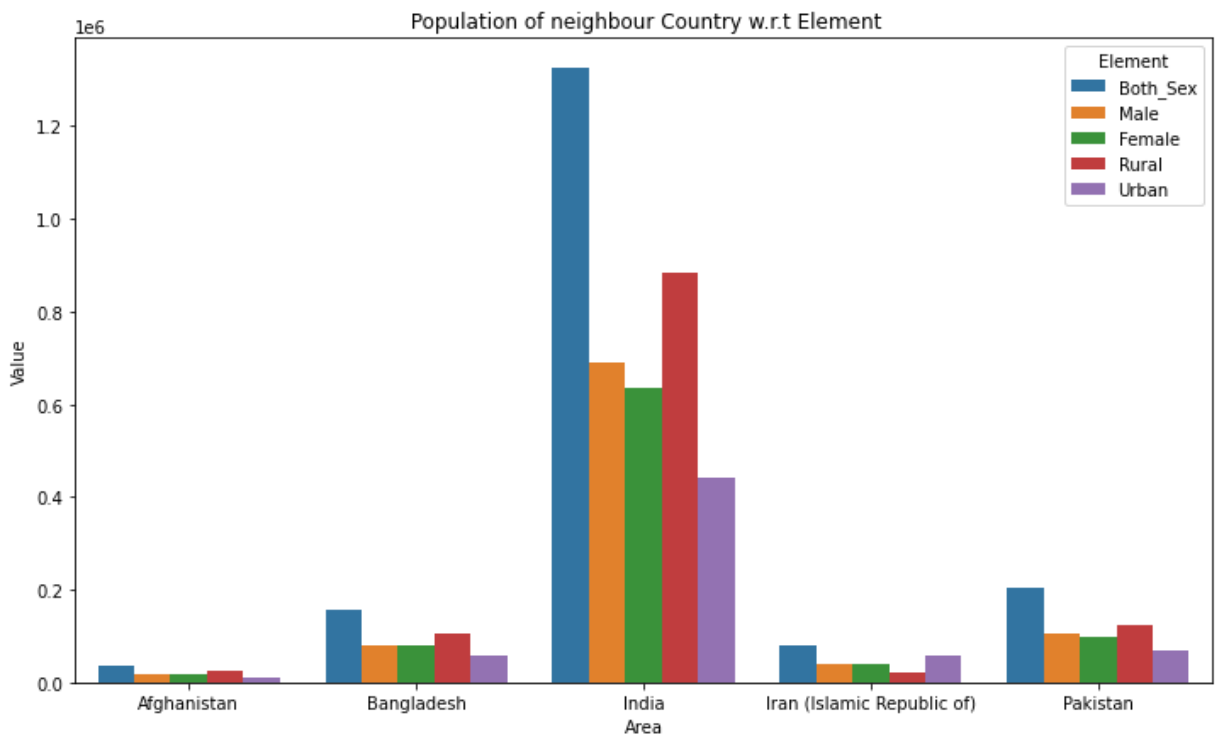


In [ ]:

```
sns.barplot( x = 'Year', y = 'Value' , data = population , hue = 'Area' ) # order ch
plt.show()
```



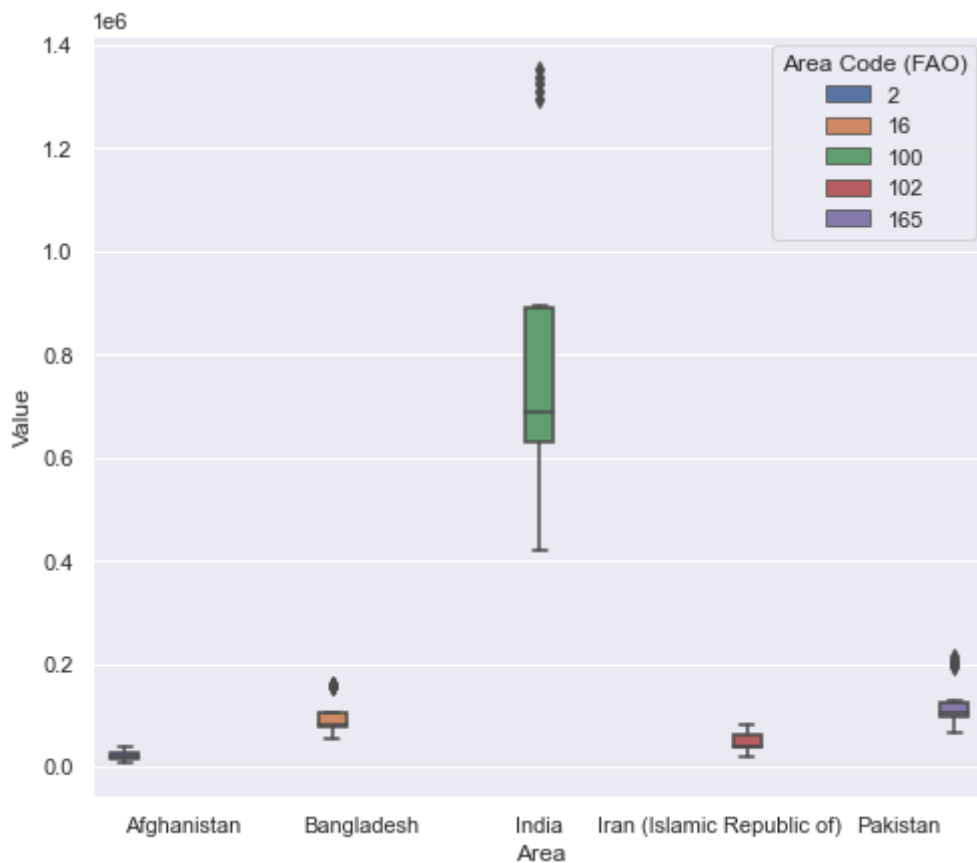
```
In [ ]: plt.figure(figsize= ( 12, 7))
sns.barplot( x = 'Area', y ='Value' , data = population , hue = 'Element' , ci=False)
plt.title(" Population of neighbour Country w.r.t Element")
plt.show()
```



```
In [ ]: sns.set(style='darkgrid')
population.head(2)
```

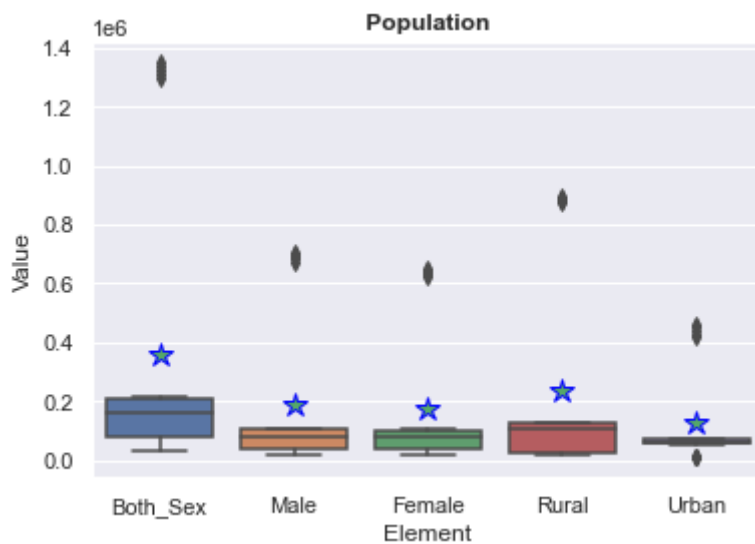
```
Out[ ]:   Area Code (FAO)  Area  Element Code  Element  Year  Value  Note
0          2  Afghanistan      511  Both_Sex  2014  33370.794  NaN
1          2  Afghanistan      512    Male    2014  17138.803  NaN
```

```
In [ ]: plt.figure(figsize= ( 8, 7))
sns.boxplot( x = 'Area' , y= 'Value', hue= "Area Code (FAO)" , data = population)
plt.show()
```



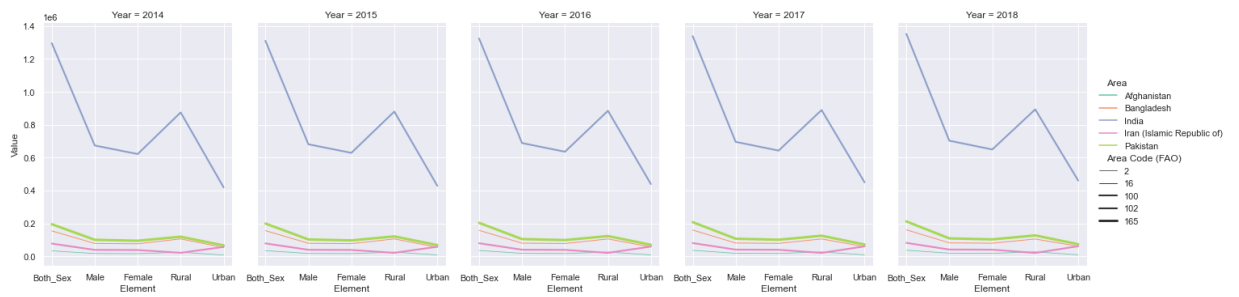
```
In [ ]: sns.boxplot( x = 'Element' , y = 'Value', data = population, showmeans =True,
                    meanprops={"marker":"*", "markersize":"12", "markeredgecolor":"blue"})

plt.title(" Population ",weight ="bold")
plt.show()
```

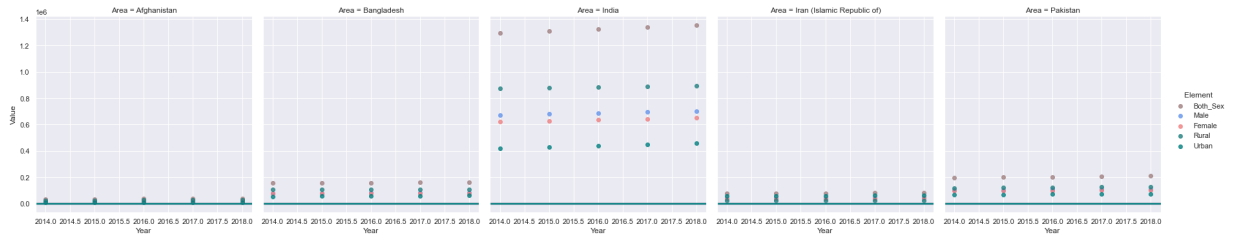


```
In [ ]: sns.relplot(
    data=population,
    x="Element", y="Value",
    hue="Area", size="Area Code (FAO)", col="Year",
    kind="line", size_order=["T1", "T2"], palette="Set2",
    height=5, aspect=.75, facet_kws=dict(sharex=False),
)
```

Out[ ]: <seaborn.axisgrid.FacetGrid at 0x2d4bb7bc160>



```
In [ ]: # Make a custom palette with gendered colors
pal = dict(Male="#6495ED", Female="#F08080", Both_Sex="#A08080",Urban="#008080",Rural="#008080")
g = sns.lmplot(x="Year", y="Value", col="Area", hue="Element", data=population,
              palette=pal, y_jitter=.02, logistic=True, truncate=False)
```



```
In [ ]: sns.pairplot(population, hue="Year")
plt.show()
```

