

# 7PAM2000 Applied Data Science 2

## Assignment 2

In the given report pandas dataframe has been used to access the extensive world Bank database. Wbgapi has been used to call the data and use and access it. In this assignment the arable land data in hectares for which the indicator is AG.LND.ARBL.HA and the total area covered by forest of three countries namely India, USA and Europe has been accessed over the time period of 10 years.

At first the libraries are being called through the code

```
import pandas as pd
import matplotlib.pyplot as plt
import wbgapi as wb
```

```
wb.source.info()
wb.series.info(q='arable')
wb.series.info(q='forest')
```

The information about the indicators are being collected through these codes.

```
Arable= wb.data.DataFrame('AG.LND.ARBL.HA', wb.region.members('EMU'))
Forest= wb.data.DataFrame('AG.LND.FRST.K2', wb.region.members('EMU'))
```

The data has been arranged in a table through the following code

```
indicator=['AG.LND.ARBL.HA','AG.LND.FRST.K2']
country=['GBR','USA','IND']
my_dataframe=wb.data.DataFrame(indicator,country, mrv=5)
```

t[18]:

		YR2016	YR2017	YR2018	YR2019	YR2020
economy	series					
GBR	AG.LND.ARBL.HA	6026010.5	6082879.4	6037787.4	NaN	NaN
	AG.LND.FRST.K2	31590.0	31640.0	31730.0	31820.0	31900.0
IND	AG.LND.ARBL.HA	158416000.0	158416000.0	158416000.0	NaN	NaN
	AG.LND.FRST.K2	710944.0	713608.0	716272.0	718936.0	721600.0
USA	AG.LND.ARBL.HA	157191000.0	157736800.0	157736800.0	NaN	NaN
	AG.LND.FRST.K2	3100950.0	3097950.0	3097950.0	3097950.0	3097950.0

Data retrieval

```
De = wb.data.DataFrame('AG.LND.FRST.K2', ['GBR','USA','IND'], time=range(2011,2020,1))
De = wb.data.DataFrame('AG.LND.FRST.K2', ['GBR','USA','IND'], time=range(2011,2020,1))
```

Out[20]:

		YR2011	YR2012	YR2013	YR2014	YR2015	YR2016	YR2017	YR2018	YR2019
economy										
GBR	AG.LND.FRST.K2	30782.0	30974.0	31186.0	31368.0	31560.0	31580.0	31640.0	31730.0	31820.0
IND	AG.LND.FRST.K2	667624.0	702880.0	702852.0	706118.0	708280.0	710944.0	713608.0	716272.0	718936.0
USA	AG.LND.FRST.K2	3089360.0	3082700.0	3086450.0	3088200.0	3100950.0	3100950.0	3097950.0	3097950.0	3097950.0

Out[20]:

		YR2011	YR2012	YR2013	YR2014	YR2015	YR2016	YR2017	YR2018	YR2019
economy										
GBR	AG.LND.FRST.K2	3078200.0	3112000.0	3255400.0	3336500.0	3411000.0	3528900.0	3603000.0	3677000.0	NaN
IND	AG.LND.FRST.K2	15842000.0	15842000.0	15842000.0	15842000.0	15842000.0	15842000.0	15842000.0	15842000.0	NaN
USA	AG.LND.FRST.K2	15842000.0	15842000.0	15842000.0	15842000.0	15842000.0	15842000.0	15842000.0	15842000.0	NaN

```
De.describe()
ar.describe()
```

```

YR2011 YR2012 YR2013 YR2014 YR2015 YR2016 YR2017 YR2018 YR2019
count 3.900000e+00 3.000000e+00 3.000000e+00 3.000000e+00 3.000000e+00 3.000000e+00 3.000000e+00 3.000000e+00 3.000000e+00
mean 1.364077e+08 1.136275e+08 1.080886e+08 1.042603e+08 1.064402e+08 1.061462e+08 1.061462e+08 1.061462e+08 1.061462e+08
std 6.856411e+07 6.035407e+07 5.844913e+07 5.850324e+07 5.850324e+07 5.850324e+07 5.850324e+07 5.850324e+07 5.850324e+07
min 6.352300e+06 5.212059e+06 5.281405e+06 5.233500e+06 5.311905e+06 5.320310e+06 5.320310e+06 5.320310e+06 5.320310e+06
25% 8.151233e+07 8.066975e+07 8.066975e+07 8.118043e+07 8.152138e+07 8.152138e+07 8.152138e+07 8.152138e+07 8.152138e+07
50% 1.503000e+08 1.500075e+08 1.500075e+08 1.500000e+08 1.504190e+08 1.504190e+08 1.504190e+08 1.504190e+08 1.504190e+08
75% 1.508750e+08 1.507750e+08 1.508775e+08 1.508711e+08 1.508500e+08 1.508500e+08 1.507574e+08 1.507574e+08 1.507574e+08
max 1.904700e+08 1.884400e+08 1.884420e+08 1.884430e+08 1.884431e+08 1.871810e+08 1.871810e+08 1.871810e+08 1.871810e+08

```

```

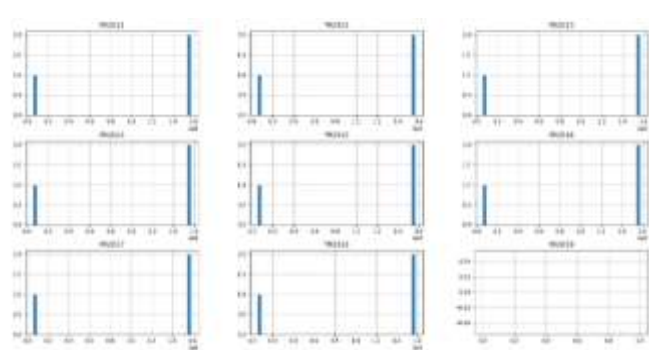
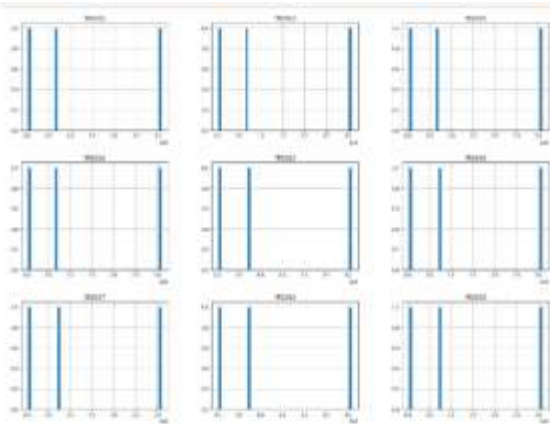
YR2011 YR2012 YR2013 YR2014 YR2015 YR2016 YR2017 YR2018 YR2019
count 3.000000e+00 3.000000e+00 3.000000e+00 3.000000e+00 3.000000e+00 3.000000e+00 3.000000e+00 3.000000e+00 3.000000e+00
mean 1.271700e+08 1.214884e+08 1.214884e+08 1.214884e+08 1.214884e+08 1.214884e+08 1.214884e+08 1.214884e+08 1.214884e+08
std 1.838844e+08 1.638844e+08 1.638844e+08 1.638844e+08 1.638844e+08 1.638844e+08 1.638844e+08 1.638844e+08 1.638844e+08
min 1.570000e+04 1.570000e+04 1.570000e+04 1.570000e+04 1.570000e+04 1.570000e+04 1.570000e+04 1.570000e+04 1.570000e+04
25% 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08
50% 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08
75% 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08 1.645000e+08
max 3.000000e+08 3.000000e+08 3.000000e+08 3.000000e+08 3.000000e+08 3.000000e+08 3.000000e+08 3.000000e+08 3.000000e+08

```

```

De.hist(bins=50, figsize=(20, 15))
plt.savefig('numeric_attributes.png')
plt.show()

```

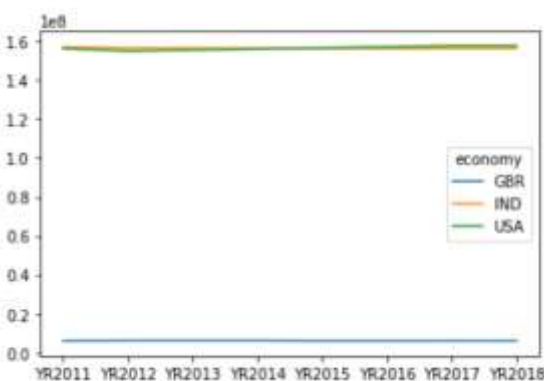


```

Plot
wb.data.DataFrame('AG.LND.ARBL.HA',
                  ['GBR','USA','IND'],
                  time=range(2011,2020,1)).T.plot();

wb.data.DataFrame('AG.LND.FRST.K2',
                  ['GBR','USA','IND'],
                  time=range(2011,2020,1)).T.plot();

```



Arable land



Forest covered

From the above figure it can be observed that for INDIA the forest covered land has not changed much over the period of 10 years. It has not changed for all the three countries.