

case_study_assign

January 13, 2022

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

```
[ ]: psrp = pd.read_csv("PSRP.csv")
psrp.head(2)
```

```
[ ]: Domain Code          Domain Area Code      Area Element Code \
0      FBS  Food Balances (2010-)      165  Pakistan      5511
1      FBS  Food Balances (2010-)      165  Pakistan      5511

      Element Item Code  Item Year Code Year      Unit Value Flag \
0  Production      2805  Rice      2010  2010  1000 tonnes  7235  Im
1  Production      2805  Rice      2011  2011  1000 tonnes  9241  Im
```

```
Flag Description
0  FAO data based on imputation methodology
1  FAO data based on imputation methodology
```

```
[ ]: # Saving data frame into excel file
psrp.to_csv("PSRP.xlsx")
```

```
[ ]: #basic statistics or summery
psrp.describe()
```

```
[ ]: Area Code  Element Code  Item Code  Year Code  Year \
count      20.0         20.0  20.000000  20.000000  20.000000
mean       165.0        5511.0 2670.500000 2014.500000 2014.500000
std         0.0         0.0   137.994088    2.946898    2.946898
min        165.0        5511.0 2536.000000 2010.000000 2010.000000
25%        165.0        5511.0 2536.000000 2012.000000 2012.000000
50%        165.0        5511.0 2670.500000 2014.500000 2014.500000
75%        165.0        5511.0 2805.000000 2017.000000 2017.000000
max        165.0        5511.0 2805.000000 2019.000000 2019.000000
```

Value

```
count      20.000000
mean      37551.800000
std       29126.440339
min       7235.000000
25%      10418.750000
50%      30274.000000
75%      65831.500000
max       83333.000000
```

```
[ ]: pd.read_csv("PSRP.csv").head(2)
```

```
[ ]:      Domain Code      Domain Area Code      Area Element Code \
0      FBS Food Balances (2010-)      165 Pakistan      5511
1      FBS Food Balances (2010-)      165 Pakistan      5511

      Element Item Code Item Year Code Year      Unit Value Flag \
0 Production      2805 Rice      2010 2010 1000 tonnes 7235 Im
1 Production      2805 Rice      2011 2011 1000 tonnes 9241 Im

      Flag Description
0 FAO data based on imputation methodology
1 FAO data based on imputation methodology
```

```
[ ]: #dropping few columns and make a new data set
new_data= psrp.drop(["Area Code","Domain Code","Element Code", 'Item_
↳Code',"Flag Description","Flag","Year Code", "Domain"], axis =1)
new_data.head()
```

```
[ ]:      Area      Element Item Year      Unit Value
0 Pakistan Production Rice 2010 1000 tonnes 7235
1 Pakistan Production Rice 2011 1000 tonnes 9241
2 Pakistan Production Rice 2012 1000 tonnes 8304
3 Pakistan Production Rice 2013 1000 tonnes 10467
4 Pakistan Production Rice 2014 1000 tonnes 10504
```

```
[ ]: new_data.mean()
```

C:\Users\Epazz\AppData\Local\Temp\ipykernel_5776\109624642.py:1: FutureWarning: Dropping of nuisance columns in DataFrame reductions (with 'numeric_only=None') is deprecated; in a future version this will raise TypeError. Select only valid columns before calling the reduction.

```
new_data.mean()
```

```
[ ]: Year      2014.5
Value      37551.8
dtype: float64
```

```
[ ]: #grouping a data set
new_data.groupby(["Item"]).mean()
```

```
[ ]:
```

| | Year | Value |
|------------|--------|---------|
| Item | | |
| Rice | 2014.5 | 9932.0 |
| Sugar cane | 2014.5 | 65171.6 |

```
[ ]: new_data[new_data['Year']<2018].groupby(["Item"]).mean()
```

```
[ ]:
```

| | Year | Value |
|------------|--------|----------|
| Item | | |
| Rice | 2013.5 | 9675.25 |
| Sugar cane | 2013.5 | 64707.75 |

```
[ ]:
```

```
[ ]:
```

```
[ ]:
```

```
[ ]:
```