```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
```

```
df_WB_read_csv('C:\Users\Dagmara Tarka\Documents\Metodyka\Projekty_Excel\_WPI_DP_FDI_TOTL_ZGS_en_csv_v2_4750824.csv')
```

```
df.head()
```

	Country Code	Region	IncomeGroup	SpecialNotes	TableName	Unnamed: 5
0	ABW	Latin America & Caribbean	High income	NaN	Aruba	NaN
1	AFG	SAS	Low income	24 countries, stretching from the Red Sea t...	Africa Eastern and Southern	NaN
2	AFG	South Asia	Low income	The reporting period for national accounts dat...	Afghanistan	NaN
3	ATG	LAC	High income	15 countries, stretching from the westernmost ...	Africa Western and Central	NaN
4	AGO	Sub-Saharan Africa	Lower middle income	The World Bank systematically assesses the app...	Angola	NaN

```
df.tail()
```

	Country Code	Region	IncomeGroup	SpecialNotes	TableName	Unnamed: 5
200	XKX	Europe & Central Asia	Upper middle income	NaN	Kosovo	NaN
201	YEM	Middle East & North Africa	Low income	The World Bank systematically assesses the app...	Yemen, Rep.	NaN
202	ZAF	Sub-Saharan Africa	Upper middle income	Fiscal year end March 31; reporting period for...	South Africa	NaN
203	ZMB	Sub-Saharan Africa	Lower middle income	National accounts data were released in 2011; ...	Zambia	NaN
204	ZWE	Sub-Saharan Africa	Lower middle income	National Accounts data are reported in Zimbab...	Zimbabwe	NaN

```
df.info()
```

```
<class pandas.core.frame.DataFrame>
RangeIndex: 205 entries, 0 to 204
Data columns (total 6 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   Country Code  205 non-null    object
 1   Region        205 non-null    object
 2   IncomeGroup   205 non-null    object
 3   SpecialNotes  128 non-null    object
 4   TableName     205 non-null    object
 5   Unnamed: 5    0 non-null      float64
dtype: float64, memory usage: 9.7 KB
```

```
df.shape
```

```
(205, 6)
```

```
df.isnull().sum()
```

```
Country Code    0
Region          0
IncomeGroup     0
SpecialNotes   77
TableName       0
Unnamed: 5    205
dtype: int64
```

```
df.drop(columns=['Unnamed: 5'],inplace=True)
df.isnull().sum()
```

```
Country Code    0
Region          0
IncomeGroup     0
SpecialNotes   77
TableName       0
dtype: int64
```

```
df.duplicated().sum()
```

```
0
```

```
df.duplicate_rows=df.duplicated()
print(df.duplicate_rows.sum)
print('No duplicate rows found.')
```

```
<bound_method_NDFrame.sum of 0    False
1    False
2    False
3    False
4    False
200  False
201  False
202  False
203  False
204  False
dtype: bool>
No duplicate rows found.
```

```
df.isnull()
```

	Country Code	Region	IncomeGroup	SpecialNotes	TableName	Unnamed: 5
0	False	False	False	True	False	True
1	False	False	False	False	False	True
2	False	False	False	False	False	True
3	False	False	False	False	False	True
4	False	False	False	False	False	True
...	...	...	...	...	...	...
200	False	False	False	True	False	True
201	False	False	False	False	False	True
202	False	False	False	False	False	True
203	False	False	False	False	False	True
204	False	False	False	False	False	True

205 rows × 6 columns

```
df[df.Region == 'Europe']
```

	Country Code	Region	IncomeGroup	SpecialNotes	TableName	Unnamed: 5
8	ADR	Latin America & Caribbean	High income	NaN	Aruba	NaN
3	ATG	South Asia	Low income	The reporting period for national accounts dat...	Afghanistan	NaN
4	AGO	Sub-Saharan Africa	Lower middle income	The World Bank systematically assesses the app...	Angola	NaN
200	XKX	Europe & Central Asia	Upper middle income	The reporting period for national accounts dat...	Kosovo	NaN
201	YEM	Middle East & North Africa	Low income	The World Bank systematically assesses the app...	Yemen, Rep.	NaN
202	ZAF	Sub-Saharan Africa	Upper middle income	Fiscal year end March 31; reporting period for...	South Africa	NaN
203	ZMB	Sub-Saharan Africa	Lower middle income	National accounts data were released in 2011; ...	Zambia	NaN
204	ZWE	Sub-Saharan Africa	Lower middle income	National Accounts data are reported in Zimbab...	Zimbabwe	NaN

205 rows × 6 columns

```
df[df.Region == 'Europe']
```

```
Series([], dtype: int64)
```

```
df[df.Region == 'Europe'].count()
```

```
Country Code    0
Region          0
IncomeGroup     0
SpecialNotes    0
TableName       0
Unnamed: 5      0
dtype: int64
```

```
df.groupby('Region').count()
```

Region	Country Code	IncomeGroup	SpecialNotes	TableName	Unnamed: 5
East Asia & Pacific	37	37	29	37	0
Europe & Central Asia	58	58	54	58	0
Latin America & Caribbean	42	42	35	42	0
Middle East & North Africa	21	21	19	21	0
North America	2	2	2	2	0
South Asia	8	8	7	8	0
Sub-Saharan Africa	37	37	32	37	0

```
df.groupby('Region').count()[['Country Code']]
```

Region	Country Code
East Asia & Pacific	37
Europe & Central Asia	58
Latin America & Caribbean	42
Middle East & North Africa	21
North America	2
South Asia	8
Sub-Saharan Africa	37

```
df['Region'].groupby
```

```
<pandas.core.groupby.generic.SeriesGroupBy object at 0x7fb4f8f8a0>
Name: Region, dtype: object
Length: 205, dtype: int64, kind: O, nlevels: 1, sort: True, dropna: True, observed: False, in_axis: False, exclusions: frozenset(), obj_shape: (205,), ngroups: 7>
```

```
df['Country Code'].groupby
```

```
<pandas.core.groupby.generic.SeriesGroupBy object at 0x7fb4f8f8a0>
Name: Country Code, dtype: object
Length: 205, dtype: int64, kind: O, nlevels: 1, sort: True, dropna: True, observed: False, in_axis: False, exclusions: frozenset(), obj_shape: (205,), ngroups: 205>
```

```
df['Region'].groupby.count()
```

```
Region
East Asia & Pacific       37
Europe & Central Asia     58
Latin America & Caribbean 42
Middle East & North Africa 21
North America             2
South Asia                8
Sub-Saharan Africa        37
Name: Region, dtype: int64
```

```
df['Country Code'].groupby.count()
```

```
Country Code
ABW    1
AFG    1
AGO    1
ALB    1
ARE    1
      ..
XKX    1
YEM    1
ZAF    1
ZMB    1
ZWE    1
Name: Country Code, dtype: int64
```

```
df['Region'].groupby.count()[['Country Code']]
```