In [1]:

```
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
%matplotlib inline
import pandas as pd
```

In [2]:

```
data= pd.read_csv('EDA.csv')
data
```

Out[2]:

	country	Year	Region	CO2 emissions from fuel combustion (MtCO2)	Average CO2 emission factor (tCO2/toe)	CO2 intensity at constant purchasing power parities (kCO2/$15p)	Total energy production (Mtoe)	Total en consump (M
0	Algeria	1990	Arab States	52.981921	2.387855	0.228889	100.112749	22.18
1	Algeria	1991	Arab States	57.149971	2.435026	0.249895	102.586171	23.46
2	Algeria	1992	Arab States	57.262430	2.383517	0.245959	104.205908	24.02
3	Algeria	1993	Arab States	59.198702	2.455335	0.259730	104.928745	24.11
4	Algeria	1994	Arab States	56.347927	2.424649	0.249468	100.612596	23.23
...	...	...	...	...	...	...	...	
1359	Venezuela	2016	South/Latin America	130.244135	2.214743	0.304632	169.124743	58.80
1360	Venezuela	2017	South/Latin America	124.106123	2.222989	0.344218	148.976952	55.82
1361	Venezuela	2018	South/Latin America	113.249553	2.701646	0.390782	115.312748	41.91
1362	Venezuela	2019	South/Latin America	101.200206	2.793624	0.537237	81.895069	36.22
1363	Venezuela	2020	South/Latin America	85.048356	2.916478	0.601990	51.397630	29.16

1364 rows × 21 columns

In [3]:

```
data.nunique()
```

Out[3]:

```
country
44
Year
31
Region
7
CO2 emissions from fuel combustion (MtCO2)
1364
Average CO2 emission factor (tCO2/toe)
1364
CO2 intensity at constant purchasing power parities (kCO2/$15p)
1364
Total energy production (Mtoe)
1364
Total energy consumption (Mtoe)
1364
Share of renewables in electricity production (%)
1302
Share of electricity in total final energy consumption (%)
1364
Oil products domestic consumption (Mt)
1347
Refined oil products production (Mt)
1354
Natural gas production (bcm)
1264
Natural gas domestic consumption (bcm)
1356
Energy intensity of GDP at constant purchasing power parities (koe/$15p)
1364
Electricity production (TWh)
1360
Electricity domestic consumption (TWh)
1359
Coal and lignite domestic consumption (Mt)
1271
Share of wind and solar in electricity production (%)
1054
Crude oil production (Mt)
1253
Coal and lignite production (Mt)
1080
dtype: int64
```

In [4]:

```
data.describe()
```

Out[4]:

	Year	CO2 emissions from fuel combustion (MtCO2)	Average CO2 emission factor (tCO2/toe)	CO2 intensity at constant purchasing power parities (kCO2/$15p)	Total energy production (Mtoe)	Total energy consumption (Mtoe)	Sl renev in ele prod
count	1364.000000	1364.000000	1364.000000	1364.000000	1364.000000	1364.000000	1364.(
mean	2005.000000	550.375178	2.238542	0.339275	229.826912	225.967632	24.(
std	8.947552	1213.057599	0.560594	0.246502	400.993541	445.242818	25.4
min	1990.000000	7.597759	0.412331	0.063650	2.939712	2.622633	0.(
25%	1997.000000	99.947198	2.021666	0.194481	34.720989	46.632538	6.(
50%	2005.000000	221.036281	2.339406	0.255510	102.602180	93.825969	14.8
75%	2013.000000	423.833152	2.521811	0.421460	217.237230	195.807818	33.2
max	2020.000000	9716.772478	3.516984	1.820894	2749.046580	3381.399262	99.8

In [5]:

```
data.corr()
```

Out[5]:

	Year	CO2 emissions from fuel combustion (MtCO2)	Average CO2 emission factor (tCO2/toe)	CO2 intensity at constant purchasing power parities (kCO2/$15p)	Total energy production (Mtoe)	Total energy consumption (Mtoe)	Sl renev ele prod
Year	1.000000	0.076034	-0.049803	-0.229235	0.090650	0.082564	0.(
CO2 emissions from fuel combustion (MtCO2)	0.076034	1.000000	0.203074	0.166831	0.915701	0.990632	-0.
Average CO2 emission factor (tCO2/toe)	-0.049803	0.203074	1.000000	0.489971	0.137877	0.145476	-0.5
CO2 intensity at constant purchasing power parities (kCO2/$15p)	-0.229235	0.166831	0.489971	1.000000	0.173768	0.136194	-0.3
Total energy production (Mtoe)	0.090650	0.915701	0.137877	0.173768	1.000000	0.924969	-0.(
Total energy consumption (Mtoe)	0.082564	0.990632	0.145476	0.136194	0.924969	1.000000	-0.1
Share of renewables in electricity production (%)	0.038103	-0.117377	-0.562855	-0.361654	-0.078679	-0.107352	1.(
Share of electricity in total final energy consumption (%)	0.253455	0.010975	-0.150981	-0.208228	-0.048509	0.008709	0.3
Oil products domestic consumption (Mt)	0.042235	0.883694	0.107419	0.050833	0.822100	0.926530	-0.
Refined oil products production (Mt)	0.055176	0.873999	0.110292	0.054587	0.851515	0.921391	-0.
Natural gas domestic consumption (bcm)	0.100317	0.676902	0.088793	0.114956	0.778109	0.742813	-0.1

	Year	CO2 emissions from fuel combustion (MtCO2)	Average CO2 emission factor (tCO2/toe)	CO2 intensity at constant purchasing power parities (kCO2/$15p)	Total energy production (Mtoe)	Total energy consumption (Mtoe)	Sl renev ele prod
Energy intensity of GDP at constant purchasing power parities (koe/$15p)	-0.248191	0.102395	0.142983	0.922538	0.143330	0.092248	-0.
Electricity production (TWh)	0.121872	0.973465	0.128116	0.084052	0.887897	0.983155	-0.(
Electricity domestic consumption (TWh)	0.121324	0.969436	0.127615	0.077970	0.881486	0.979843	-0.(
Crude oil production (Mt)	0.054966	0.448813	0.027750	0.068146	0.723797	0.486539	-0.(

In [6]:

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1364 entries, 0 to 1363
Data columns (total 21 columns):
 #   Column
Non-Null Count  Dtype
---  ------
--------------  -----
 0   country
1364 non-null   object
 1   Year
1364 non-null   int64
 2   Region
1364 non-null   object
 3   CO2 emissions from fuel combustion (MtCO2)
1364 non-null   float64
 4   Average CO2 emission factor (tCO2/toe)
1364 non-null   float64
 5   CO2 intensity at constant purchasing power parities (kCO2/$15p)
1364 non-null   float64
 6   Total energy production (Mtoe)
1364 non-null   float64
 7   Total energy consumption (Mtoe)
1364 non-null   float64
 8   Share of renewables in electricity production (%)
1364 non-null   float64
 9   Share of electricity in total final energy consumption (%)
1364 non-null   float64
 10  Oil products domestic consumption (Mt)
1364 non-null   float64
 11  Refined oil products production (Mt)
1364 non-null   float64
 12  Natural gas production (bcm)
1364 non-null   object
 13  Natural gas domestic consumption (bcm)
1364 non-null   float64
 14  Energy intensity of GDP at constant purchasing power parities (koe/$1
5p)  1364 non-null   float64
 15  Electricity production (TWh)
1364 non-null   float64
 16  Electricity domestic consumption (TWh)
1364 non-null   float64
 17  Coal and lignite domestic consumption (Mt)
1364 non-null   object
 18  Share of wind and solar in electricity production (%)
1364 non-null   object
 19  Crude oil production (Mt)
1364 non-null   float64
 20  Coal and lignite production (Mt)
1364 non-null   object
dtypes: float64(14), int64(1), object(6)
memory usage: 223.9+ KB
```

In [7]:

```
print(list(data))
```

```
['country', 'Year', 'Region', 'CO2 emissions from fuel combustion (MtCO
2)', 'Average CO2 emission factor (tCO2/toe)', 'CO2 intensity at constant
purchasing power parities (kCO2/$15p)', 'Total energy production (Mtoe)',
'Total energy consumption (Mtoe)', 'Share of renewables in electricity pro
duction (%)', 'Share of electricity in total final energy consumption
(%)', 'Oil products domestic consumption (Mt)', 'Refined oil products prod
uction (Mt)', 'Natural gas production (bcm)', 'Natural gas domestic consum
ption (bcm)', 'Energy intensity of GDP at constant purchasing power pariti
es (koe/$15p)', 'Electricity production (TWh)', 'Electricity domestic cons
umption (TWh)', 'Coal and lignite domestic consumption (Mt)', 'Share of wi
nd and solar in electricity production (%)', 'Crude oil production (Mt)',
'Coal and lignite production (Mt)']
```

In [8]:

```
columns=['Year', 'CO2 emissions from fuel combustion (MtCO2)', 'Average CO2 emission fact
```

In [9]:

```
data.isnull().sum()
```

Out[9]:

```
country
0
Year
0
Region
0
CO2 emissions from fuel combustion (MtCO2)
0
Average CO2 emission factor (tCO2/toe)
0
CO2 intensity at constant purchasing power parities (kCO2/$15p)
0
Total energy production (Mtoe)
0
Total energy consumption (Mtoe)
0
Share of renewables in electricity production (%)
0
Share of electricity in total final energy consumption (%)
0
Oil products domestic consumption (Mt)
0
Refined oil products production (Mt)
0
Natural gas production (bcm)
0
Natural gas domestic consumption (bcm)
0
Energy intensity of GDP at constant purchasing power parities (koe/$15p)
0
Electricity production (TWh)
0
Electricity domestic consumption (TWh)
0
Coal and lignite domestic consumption (Mt)
0
Share of wind and solar in electricity production (%)
0
Crude oil production (Mt)
0
Coal and lignite production (Mt)
0
dtype: int64
```

In [10]:

```
def null_values(data,column):
    #null_values=data[data[column]=='n.a.']
    null_values=data[column].astype(str).replace('n.a.',0,regex=True)
    return null_values
```

In [11]:

```
natural_gas_bcm= null_values(data,'Natural gas production (bcm)')
data['Natural gas production (bcm)'] = natural_gas_bcm
natural_gas_bcm
```

Out[11]:

```
0       45.68403547
1       49.53198889
2       51.67001137
3       52.38599054
4       48.20601995
           ...
1359    24.47599903
1360    24.00755269
1361    23.70998143
1362    19.59998465
1363    14.62158855
Name: Natural gas production (bcm), Length: 1364, dtype: object
```

In [12]:

```
natural_gas_bcm
```

Out[12]:

```
0       45.68403547
1       49.53198889
2       51.67001137
3       52.38599054
4       48.20601995
           ...
1359    24.47599903
1360    24.00755269
1361    23.70998143
1362    19.59998465
1363    14.62158855
Name: Natural gas production (bcm), Length: 1364, dtype: object
```

In [13]:

```
coal_lignite_mt=null_values(data,'Coal and lignite domestic consumption (Mt)')
data['Coal and lignite domestic consumption (Mt)']= coal_lignite_mt
coal_lignite_mt
```

Out[13]:

```
0              0.94
1             1.171
2             1.056
3             1.015
4             0.943
           ...
1359          0.202
1360       0.112222
1361       0.044889
1362    0.029469014
1363    0.019346004
Name: Coal and lignite domestic consumption (Mt), Length: 1364, dtype: obj
ect
```

In [14]:

```
share_wind=null_values(data,'Share of wind and solar in electricity production (%)')
data['Share of wind and solar in electricity production (%)']=share_wind
share_wind
```

Out[14]:

```
0                 0
1                 0
2                 0
3                 0
4                 0
           ...
1359     0.07223493
1360    0.075160838
1361    0.077432701
1362    0.073940761
1363     0.07684631
Name: Share of wind and solar in electricity production (%), Length: 1364,
dtype: object
```

In [15]:

```
coal_lignite_production=null_values(data,'Coal and lignite production (Mt)')
data['Coal and lignite production (Mt)']=coal_lignite_production
coal_lignite_production
```

Out[15]:

```
0                 0
1                 0
2                 0
3                 0
4                 0
           ...
1359            0.9
1360            0.5
1361            0.2
1362     0.29820421
1363    0.195766979
Name: Coal and lignite production (Mt), Length: 1364, dtype: object
```

In [16]:

```
data1= data
data1.drop(['country','Year','Region','Coal and lignite production (Mt)','Share of wind a
```

Out[16]:

	CO2 emissions from fuel combustion (MtCO2)	Average CO2 emission factor (tCO2/toe)	CO2 intensity at constant purchasing power parities (kCO2/$15p)	Total energy production (Mtoe)	Total energy consumption (Mtoe)	Share of renewables in electricity production (%)	Share electricity total fin energ consumptic (%
0	52.981921	2.387855	0.228889	100.112749	22.188078	0.782972	8.3765
1	57.149971	2.435026	0.249895	102.586171	23.469959	1.580793	7.9947
2	57.262430	2.383517	0.245959	104.205908	24.024343	1.022926	8.4900
3	59.198702	2.455335	0.259730	104.928745	24.110238	1.709195	8.2910
4	56.347927	2.424649	0.249468	100.612596	23.239627	0.783869	9.1706
...	...	...	...	...	...	...	
1359	130.244135	2.214743	0.304632	169.124743	58.807791	58.377863	19.3083
1360	124.106123	2.222989	0.344218	148.976952	55.828488	58.379586	19.6585
1361	113.249553	2.701646	0.390782	115.312748	41.918720	58.380532	20.3192
1362	101.200206	2.793624	0.537237	81.895069	36.225425	58.377040	21.2923
1363	85.048356	2.916478	0.601990	51.397630	29.161319	60.944184	23.7875

1364 rows × 15 columns

In [17]:

```
data1
```

Out[17]:

	country	Year	Region	CO2 emissions from fuel combustion (MtCO2)	Average CO2 emission factor (tCO2/toe)	CO2 intensity at constant purchasing power parities (kCO2/$15p)	Total energy production (Mtoe)	Total en consump (M
0	Algeria	1990	Arab States	52.981921	2.387855	0.228889	100.112749	22.18
1	Algeria	1991	Arab States	57.149971	2.435026	0.249895	102.586171	23.46
2	Algeria	1992	Arab States	57.262430	2.383517	0.245959	104.205908	24.02
3	Algeria	1993	Arab States	59.198702	2.455335	0.259730	104.928745	24.11
4	Algeria	1994	Arab States	56.347927	2.424649	0.249468	100.612596	23.23
...	...	...	...	...	...	...	...	
1359	Venezuela	2016	South/Latin America	130.244135	2.214743	0.304632	169.124743	58.80
1360	Venezuela	2017	South/Latin America	124.106123	2.222989	0.344218	148.976952	55.82
1361	Venezuela	2018	South/Latin America	113.249553	2.701646	0.390782	115.312748	41.91
1362	Venezuela	2019	South/Latin America	101.200206	2.793624	0.537237	81.895069	36.22
1363	Venezuela	2020	South/Latin America	85.048356	2.916478	0.601990	51.397630	29.16

1364 rows × 21 columns

In [18]:

```
print(list(data1))
```

```
['country', 'Year', 'Region', 'CO2 emissions from fuel combustion (MtCO
2)', 'Average CO2 emission factor (tCO2/toe)', 'CO2 intensity at constant
purchasing power parities (kCO2/$15p)', 'Total energy production (Mtoe)',
'Total energy consumption (Mtoe)', 'Share of renewables in electricity pro
duction (%)', 'Share of electricity in total final energy consumption
(%)', 'Oil products domestic consumption (Mt)', 'Refined oil products prod
uction (Mt)', 'Natural gas production (bcm)', 'Natural gas domestic consum
ption (bcm)', 'Energy intensity of GDP at constant purchasing power pariti
es (koe/$15p)', 'Electricity production (TWh)', 'Electricity domestic cons
umption (TWh)', 'Coal and lignite domestic consumption (Mt)', 'Share of wi
nd and solar in electricity production (%)', 'Crude oil production (Mt)',
'Coal and lignite production (Mt)']
```

In [19]:

```
col= ['CO2 emissions from fuel combustion (MtCO2)', 'Average CO2 emission factor (tCO2/tc
data1[col]=data1[col].astype('float')
```

In [20]:

```
data1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1364 entries, 0 to 1363
Data columns (total 21 columns):
 #   Column
Non-Null Count  Dtype
---  ------
--------------  -----
 0   country
1364 non-null   object
 1   Year
1364 non-null   int64
 2   Region
1364 non-null   object
 3   CO2 emissions from fuel combustion (MtCO2)
1364 non-null   float64
 4   Average CO2 emission factor (tCO2/toe)
1364 non-null   float64
 5   CO2 intensity at constant purchasing power parities (kCO2/$15p)
1364 non-null   float64
 6   Total energy production (Mtoe)
1364 non-null   float64
 7   Total energy consumption (Mtoe)
1364 non-null   float64
 8   Share of renewables in electricity production (%)
1364 non-null   float64
 9   Share of electricity in total final energy consumption (%)
1364 non-null   float64
 10  Oil products domestic consumption (Mt)
1364 non-null   float64
 11  Refined oil products production (Mt)
1364 non-null   float64
 12  Natural gas production (bcm)
1364 non-null   float64
 13  Natural gas domestic consumption (bcm)
1364 non-null   float64
 14  Energy intensity of GDP at constant purchasing power parities (koe/$1
5p)  1364 non-null   float64
 15  Electricity production (TWh)
1364 non-null   float64
 16  Electricity domestic consumption (TWh)
1364 non-null   float64
 17  Coal and lignite domestic consumption (Mt)
1364 non-null   float64
 18  Share of wind and solar in electricity production (%)
1364 non-null   float64
 19  Crude oil production (Mt)
1364 non-null   float64
 20  Coal and lignite production (Mt)
1364 non-null   float64
dtypes: float64(18), int64(1), object(2)
memory usage: 223.9+ KB
```

In [21]:

```
data1.drop(['Natural gas production (bcm)'],axis=1)
```

Out[21]:

	country	Year	Region	CO2 emissions from fuel combustion (MtCO2)	Average CO2 emission factor (tCO2/toe)	CO2 intensity at constant purchasing power parities (kCO2/$15p)	Total energy production (Mtoe)	Total en consump (M
0	Algeria	1990	Arab States	52.981921	2.387855	0.228889	100.112749	22.18
1	Algeria	1991	Arab States	57.149971	2.435026	0.249895	102.586171	23.46
2	Algeria	1992	Arab States	57.262430	2.383517	0.245959	104.205908	24.02
3	Algeria	1993	Arab States	59.198702	2.455335	0.259730	104.928745	24.11
4	Algeria	1994	Arab States	56.347927	2.424649	0.249468	100.612596	23.23
...	...	...	...	...	...	...	...	
1359	Venezuela	2016	South/Latin America	130.244135	2.214743	0.304632	169.124743	58.80
1360	Venezuela	2017	South/Latin America	124.106123	2.222989	0.344218	148.976952	55.82
1361	Venezuela	2018	South/Latin America	113.249553	2.701646	0.390782	115.312748	41.91
1362	Venezuela	2019	South/Latin America	101.200206	2.793624	0.537237	81.895069	36.22
1363	Venezuela	2020	South/Latin America	85.048356	2.916478	0.601990	51.397630	29.16

1364 rows × 20 columns

In [22]:

```
data1
```

Out[22]:

	country	Year	Region	CO2 emissions from fuel combustion (MtCO2)	Average CO2 emission factor (tCO2/toe)	CO2 intensity at constant purchasing power parities (kCO2/$15p)	Total energy production (Mtoe)	Total en consump (M
0	Algeria	1990	Arab States	52.981921	2.387855	0.228889	100.112749	22.18
1	Algeria	1991	Arab States	57.149971	2.435026	0.249895	102.586171	23.46
2	Algeria	1992	Arab States	57.262430	2.383517	0.245959	104.205908	24.02
3	Algeria	1993	Arab States	59.198702	2.455335	0.259730	104.928745	24.11
4	Algeria	1994	Arab States	56.347927	2.424649	0.249468	100.612596	23.23
...	...	...	...	...	...	...	...	
1359	Venezuela	2016	South/Latin America	130.244135	2.214743	0.304632	169.124743	58.80
1360	Venezuela	2017	South/Latin America	124.106123	2.222989	0.344218	148.976952	55.82
1361	Venezuela	2018	South/Latin America	113.249553	2.701646	0.390782	115.312748	41.91
1362	Venezuela	2019	South/Latin America	101.200206	2.793624	0.537237	81.895069	36.22
1363	Venezuela	2020	South/Latin America	85.048356	2.916478	0.601990	51.397630	29.16

1364 rows × 21 columns

In [ ]:

```

```

In [23]:

```
31+31+62+93
```

Out[23]:

217

In [24]:

```
def outliers(data1,column,q1=0.15,q3=0.95,thershole=1.5):
    iqr=data1[column].quantile(q3)-data1[column].quantile(q1)
    outliers=data1[(data1[column]<data1[column].quantile(q3)-thershole*iqr)|(data1[columr
    return outliers
```

In [25]:

```
outlire1 = outliers(data1,'CO2 emissions from fuel combustion (MtCO2)')
print('1',len(outlire1))
outlire2 = outliers(data1,'Average CO2 emission factor (tCO2/toe)')
print('2',len(outlire2))
outlire3 = outliers(data1,'CO2 intensity at constant purchasing power parities (kCO2/$15p
print('3',len(outlire3))
outlire4 = outliers(data1,'Total energy production (Mtoe)')
print('4',len(outlire4))
outlire5 = outliers(data1,'Total energy consumption (Mtoe)')
print('5',len(outlire5))
outlire6 = outliers(data1,'Share of renewables in electricity production (%)')
print('6',len(outlire6))
outlire7 = outliers(data1,'Share of electricity in total final energy consumption (%)')
print('7',len(outlire7))
outlire8 = outliers(data1,'Oil products domestic consumption (Mt)')
print('8',len(outlire8))
outlire9 = outliers(data1,'Refined oil products production (Mt)')
print('9',len(outlire9))
#outlire10 = outliers(data1,'Natural gas production (bcm)')
#print('10',len(outlire10))
outlire11 = outliers(data1,'Natural gas domestic consumption (bcm)')
print('11',len(outlire11))
outlire12 = outliers(data1,'Energy intensity of GDP at constant purchasing power parities
print('12',len(outlire12))
outlire13 = outliers(data1,'Electricity production (TWh)')
print('13',len(outlire13))
outlire14 = outliers(data1,'Electricity domestic consumption (TWh)')
print('14',len(outlire14))
#outlire15 = outliers(data1,'Coal and lignite domestic consumption (Mt)')
#print('15',len(outlire15))
#outlire16 = outliers(data1,'Share of wind and solar in electricity production (%)')
#print('16',len(outlire16))
outlire17 = outliers(data1,'Crude oil production (Mt)')
print('17',len(outlire17))
#outlire18 = outliers(data1,'Coal and lignite production (Mt)')
#print('18',len(outlire18))
```

```
1 53
2 39
3 29
4 25
5 49
6 0
7 63
8 48
9 45
11 63
12 32
13 50
14 49
17 8
```

In [26]:

```
a = 'outlire,'*15
```

In [27]:

```
a
```

Out[27]:

```
'outlire,outlire,outlire,outlire,outlire,outlire,outlire,outlire,outlire,o
utlire,outlire,outlire,outlire,outlire,outlire,'
```

In [28]:

```
outliress = pd.concat([outlire1,outlire2,outlire3,outlire4,outlire5,outlire6,outlire7,out
```

In [29]:

```
outliress
```

Out[29]:

	country	Year	Region	CO2 emissions from fuel combustion (MtCO2)	Average CO2 emission factor (tCO2/toe)	CO2 intensity at constant purchasing power parities (kCO2/$15p)	Total energy production (Mtoe)	Total energ consumptio (Mtoe
225	China	1998	Asia & Pacific	3084.756807	2.859900	0.809915	1079.266896	1078.62400
227	China	2000	Asia & Pacific	3140.310907	2.779681	0.705847	1123.707793	1129.73793
228	China	2001	Asia & Pacific	3274.685894	2.802829	0.679390	1176.292128	1168.35033
229	China	2002	Asia & Pacific	3534.084557	2.836422	0.671862	1226.935300	1245.96583
230	China	2003	Asia & Pacific	4108.377553	2.893023	0.709807	1375.185474	1420.09841
...	...	...	...	...	...	...	...	.
1296	United States	2015	North America	5044.075416	2.300877	0.276853	2027.113489	2192.24065
1298	United States	2017	North America	4890.655983	2.263182	0.258558	1998.078309	2160.96430
1299	United States	2018	North America	5055.832670	2.260882	0.259689	2177.940569	2236.22090
1300	United States	2019	North America	4932.637347	2.228277	0.248002	2304.593149	2213.65559
1301	United States	2020	North America	4404.709314	2.153145	0.229491	2190.194556	2045.70974

553 rows × 21 columns

In [30]:

```
outliress.duplicated().sum()
```

Out[30]:

```
359
```

In [31]:

```
outliress.drop_duplicates(inplace=True)
data1.drop(outliress.index,inplace=True)
data1
```

Out[31]:

	country	Year	Region	CO2 emissions from fuel combustion (MtCO2)	Average CO2 emission factor (tCO2/toe)	CO2 intensity at constant purchasing power parities (kCO2/$15p)	Total energy production (Mtoe)	Total en consump (M
0	Algeria	1990	Arab States	52.981921	2.387855	0.228889	100.112749	22.18
1	Algeria	1991	Arab States	57.149971	2.435026	0.249895	102.586171	23.46
2	Algeria	1992	Arab States	57.262430	2.383517	0.245959	104.205908	24.02
3	Algeria	1993	Arab States	59.198702	2.455335	0.259730	104.928745	24.11
4	Algeria	1994	Arab States	56.347927	2.424649	0.249468	100.612596	23.23
...	...	...	...	...	...	...	...	
1359	Venezuela	2016	South/Latin America	130.244135	2.214743	0.304632	169.124743	58.80
1360	Venezuela	2017	South/Latin America	124.106123	2.222989	0.344218	148.976952	55.82
1361	Venezuela	2018	South/Latin America	113.249553	2.701646	0.390782	115.312748	41.91
1362	Venezuela	2019	South/Latin America	101.200206	2.793624	0.537237	81.895069	36.22
1363	Venezuela	2020	South/Latin America	85.048356	2.916478	0.601990	51.397630	29.16

1170 rows × 21 columns

In [32]:

```
data1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1170 entries, 0 to 1363
Data columns (total 21 columns):
 #   Column
Non-Null Count  Dtype
---  ------
--------------  -----
 0   country
1170 non-null   object
 1   Year
1170 non-null   int64
 2   Region
1170 non-null   object
 3   CO2 emissions from fuel combustion (MtCO2)
1170 non-null   float64
 4   Average CO2 emission factor (tCO2/toe)
1170 non-null   float64
 5   CO2 intensity at constant purchasing power parities (kCO2/$15p)
1170 non-null   float64
 6   Total energy production (Mtoe)
1170 non-null   float64
 7   Total energy consumption (Mtoe)
1170 non-null   float64
 8   Share of renewables in electricity production (%)
1170 non-null   float64
 9   Share of electricity in total final energy consumption (%)
1170 non-null   float64
 10  Oil products domestic consumption (Mt)
1170 non-null   float64
 11  Refined oil products production (Mt)
1170 non-null   float64
 12  Natural gas production (bcm)
1170 non-null   float64
 13  Natural gas domestic consumption (bcm)
1170 non-null   float64
 14  Energy intensity of GDP at constant purchasing power parities (koe/$1
5p)  1170 non-null   float64
 15  Electricity production (TWh)
1170 non-null   float64
 16  Electricity domestic consumption (TWh)
1170 non-null   float64
 17  Coal and lignite domestic consumption (Mt)
1170 non-null   float64
 18  Share of wind and solar in electricity production (%)
1170 non-null   float64
 19  Crude oil production (Mt)
1170 non-null   float64
 20  Coal and lignite production (Mt)
1170 non-null   float64
dtypes: float64(18), int64(1), object(2)
memory usage: 201.1+ KB
```

10 bcm of natural gas = 1.2 ktoe 5 Mtoe = 5,000 ktoe

1 ktoe = 0.001 Mtoe

1 bcm of natural gas = 920.2 ktoe / 1000 = 0.92 Mtoe

1 Mtoe of natural gas: approximately 1,832 kton CO2

1 Terawatt hour (TWh) is equal to approximately 0.086 Mtoe (million tons of oil equivalent)

In [33]:

```
1832*0.15
```

Out[33]:

```
274.8
```

In [34]:

```
data1['Natural gas production (bcm)'] = data1['Natural gas production (bcm)']*0.9
```

In [35]:

```
data1['Natural gas domestic consumption (bcm)']=data1['Natural gas domestic consumption (
```

In [36]:

```
data1['CO2 intensity at constant purchasing power parities (kCO2/$15p)']=data1['CO2 inter
```

In [37]:

```
data1['Energy intensity of GDP at constant purchasing power parities (koe/$15p)'] = data1
```

In [38]:

```
data1['Electricity production (TWh)'] = data1['Electricity production (TWh)']*0.086
data1['Electricity domestic consumption (TWh)'] = data1['Electricity domestic consumptior
```

In [39]:

```
colum={'CO2 intensity at constant purchasing power parities (kCO2/$15p)':'CO2 intensity a
data1.rename(columns=colum,inplace=True)
```

In [40]:

```
data1.nunique()
```

Out[40]:

```
country
40
Year
31
Region
7
CO2 emissions from fuel combustion (MtCO2)
1170
Average CO2 emission factor (tCO2/toe)
1170
CO2 intensity at constant purchasing power parities (MtCO2/$274.8)
1170
Total energy production (Mtoe)
1170
Total energy consumption (Mtoe)
1170
Share of renewables in electricity production (%)
1108
Share of electricity in total final energy consumption (%)
1170
Oil products domestic consumption (Mt)
1159
Refined oil products production (Mt)
1163
Natural gas production (Mtoe)
1077
Natural gas domestic consumption (Mtoe)
1162
Energy intensity of GDP at constant purchasing power parities (Mtoe/$274.
8)    1170
Electricity production (Mt)
1166
Electricity domestic consumption (Mt)
1167
Coal and lignite domestic consumption (Mt)
1092
Share of wind and solar in electricity production (%)
909
Crude oil production (Mt)
1065
Coal and lignite production (Mt)
907
dtype: int64
```

In [41]:

```
data1.Region.value_counts()
```

Out[41]:

```
Europe                 418
Asia & Pacific         321
South/Latin America    186
Middle east            152
Arab States             31
North America           31
Africa                  31
Name: Region, dtype: int64
```

In [42]:

```
data1.Region.unique()
```

Out[42]:

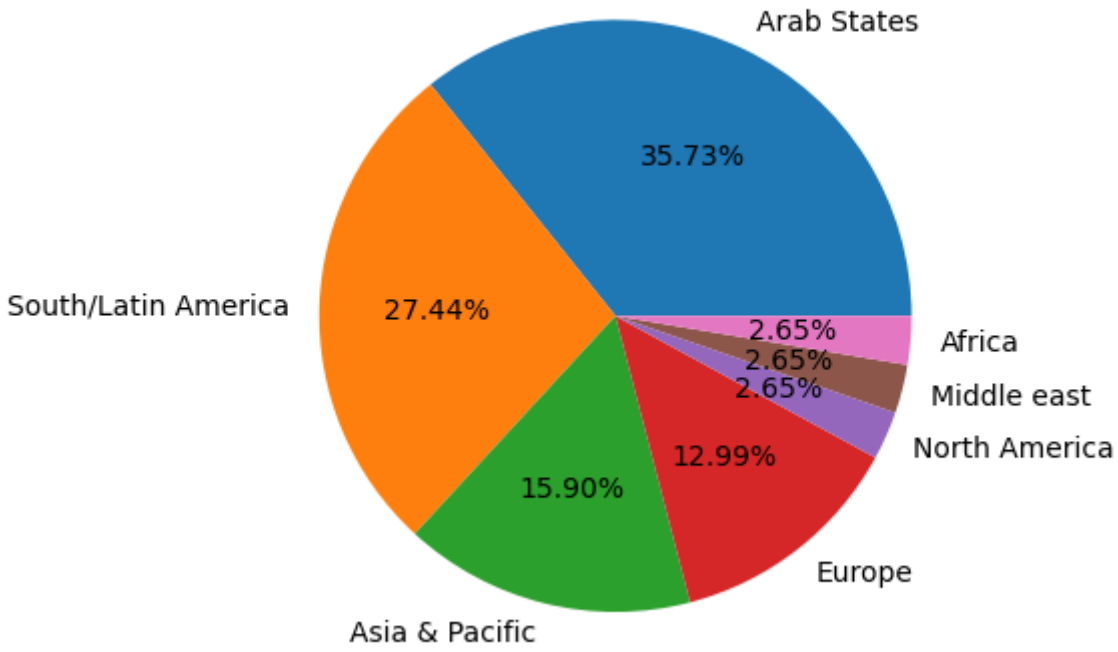
```
array(['Arab States', 'South/Latin America', 'Asia & Pacific', 'Europe',
       'North America', 'Middle east', 'Africa'], dtype=object)
```

In [43]:

```
plt.pie(data1.Region.value_counts(),autopct='%0.2f%%',labels=data1.Region.unique())
```

Out[43]:

```
([<matplotlib.patches.Wedge at 0x29639da0190>,
  <matplotlib.patches.Wedge at 0x29639dac7c0>,
  <matplotlib.patches.Wedge at 0x29639dacee0>,
  <matplotlib.patches.Wedge at 0x29639dc4640>,
  <matplotlib.patches.Wedge at 0x29639dc4d60>,
  <matplotlib.patches.Wedge at 0x29639dd04c0>,
  <matplotlib.patches.Wedge at 0x29639dd0be0>],
 [Text(0.4768918896621537, 0.9912487707808068, 'Arab States'),
  Text(-1.0993299133951526, 0.03838934115879665, 'South/Latin America'),
  Text(-0.26611391971962334, -1.06732534015241, 'Asia & Pacific'),
  Text(0.6772277914839663, -0.8668116972224992, 'Europe'),
  Text(1.0060978274955157, -0.4447101994656787, 'North America'),
  Text(1.065880877427387, -0.27184178327590525, 'Middle east'),
  Text(1.0961914381128786, -0.09145671658232211, 'Africa')],
 [Text(0.26012284890662923, 0.5406811476986217, '35.73%'),
  Text(-0.5996344982155377, 0.020939640632070898, '27.44%'),
  Text(-0.14515304711979454, -0.5821774582649509, '15.90%'),
  Text(0.3693969771730725, -0.47280638030318134, '12.99%'),
  Text(0.5487806331793722, -0.24256919970855198, '2.65%'),
  Text(0.5813895695058473, -0.14827733633231194, '2.65%'),
  Text(0.5979226026070247, -0.049885481772175694, '2.65%')])
```

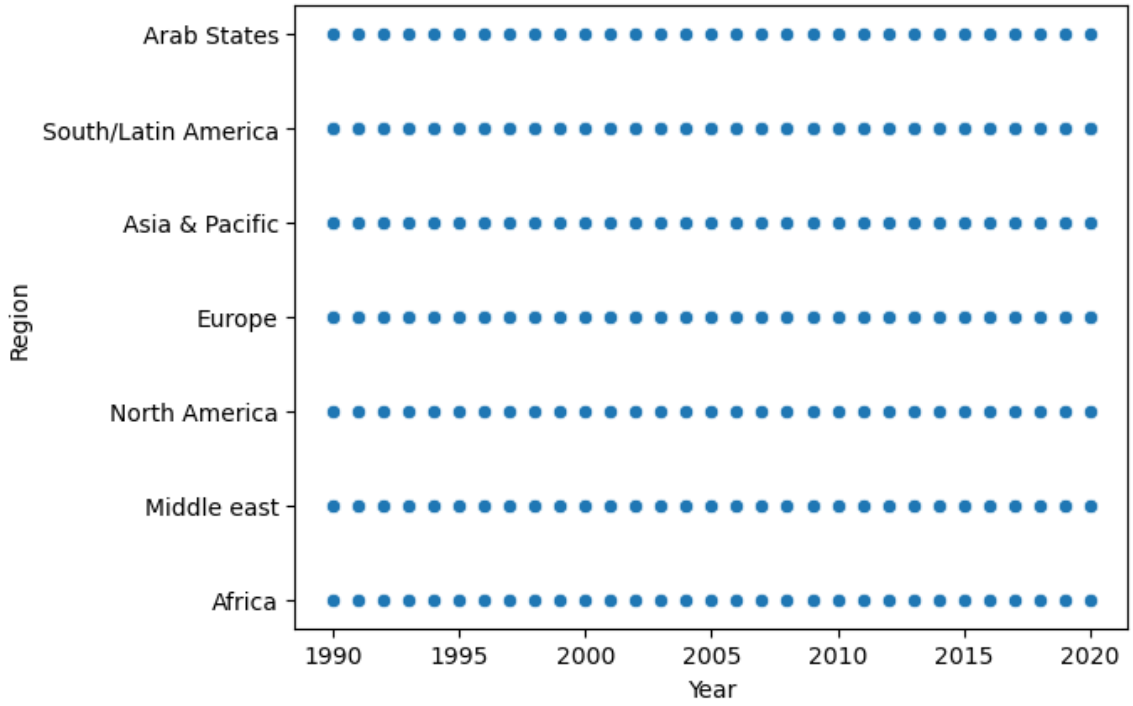


In [44]:

```
sns.scatterplot(y=data1.Region,x=data1.Year,palette='Greys')
plt.figure(figsize=(25,15))
```

Out[44]:

```
<Figure size 2500x1500 with 0 Axes>
```



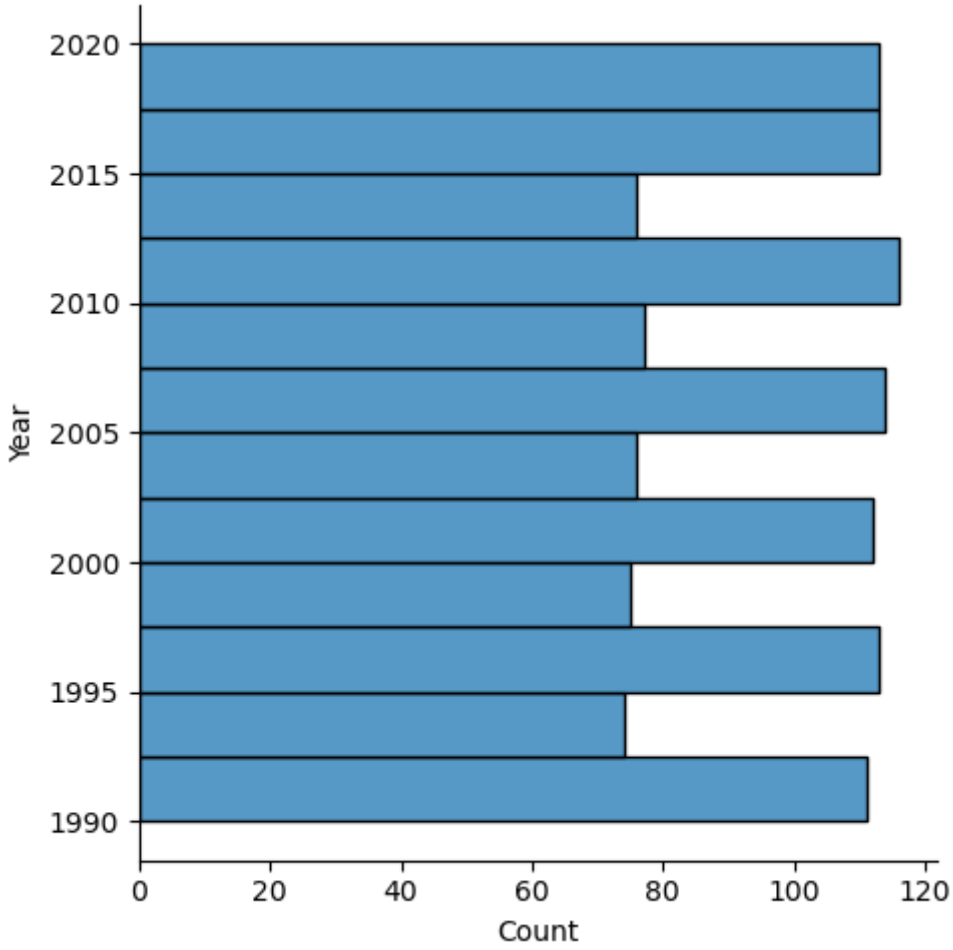

```
<Figure size 2500x1500 with 0 Axes>
```

In [45]:

```
sns.displot(y=data1.Year)
```

Out[45]:

```
<seaborn.axisgrid.FacetGrid at 0x29639e114c0>
```

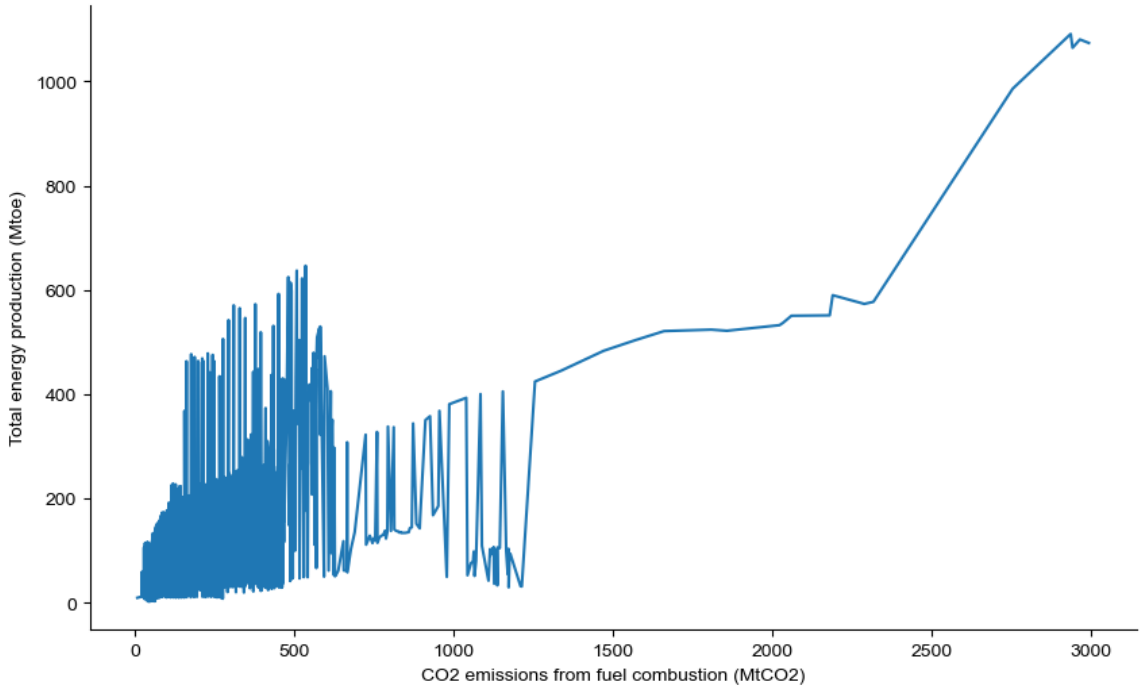


In [46]:

```
def creat_plots(plot,data1,column_x,column_y,palette):
    plt.figure(figsize=(10, 6))
    if plot=='barplot':
        plots=sns.barplot(x=column_x,y=column_y,data=data1)

    elif plot=='scatterplot':
        plots=sns.scatterplot(x=column_x,y=column_y,data=data1)

    elif plot=='boxplot':
        plots=sns.boxplot(x=column_x,y=column_y,data=data1)

    elif plot=='lineplot':
        plots=sns.lineplot(x=column_x,y=column_y,data=data1)

    elif plot=='displot':
        plots=sns.displot(x=column_x,y=column_y,data=data1)


    elif plot=='countplot':
        plots=sns.countplot(x=column_x,y=column_y,data=data1)


    elif plot=='histplot':
        plots=sns.histplot(x=column_x,y=column_y,data=data1)

    elif plot=='lmplot':
        plots=sns.lmplot(x=column_x,y=column_y,data=data1,palette=palette)

    elif plot=='violinplot':
        plots=sns.violinplot(x=column_x,y=column_y,data=data1,palette=palette)

    elif plot=='swarmplot':
        plots=sns.swarmplot(x=column_x,y=column_y,data=data1,palette=palette)

    elif plot=='factorplot':
        plots=sns.factorplot(x=column_x,y=column_y,data=data1)


    elif plot=='stripplot':
        plots=sns.stripplot(x=column_x,y=column_y,data=data1)
    else:
        print('wrong plot')
    sns.despine()
    sns.set_style("dark")
    plt.show()
    return plots
```

In [47]:

```
creat_plots('lineplot',data1,'CO2 emissions from fuel combustion (MtCO2)','Total energy p
```

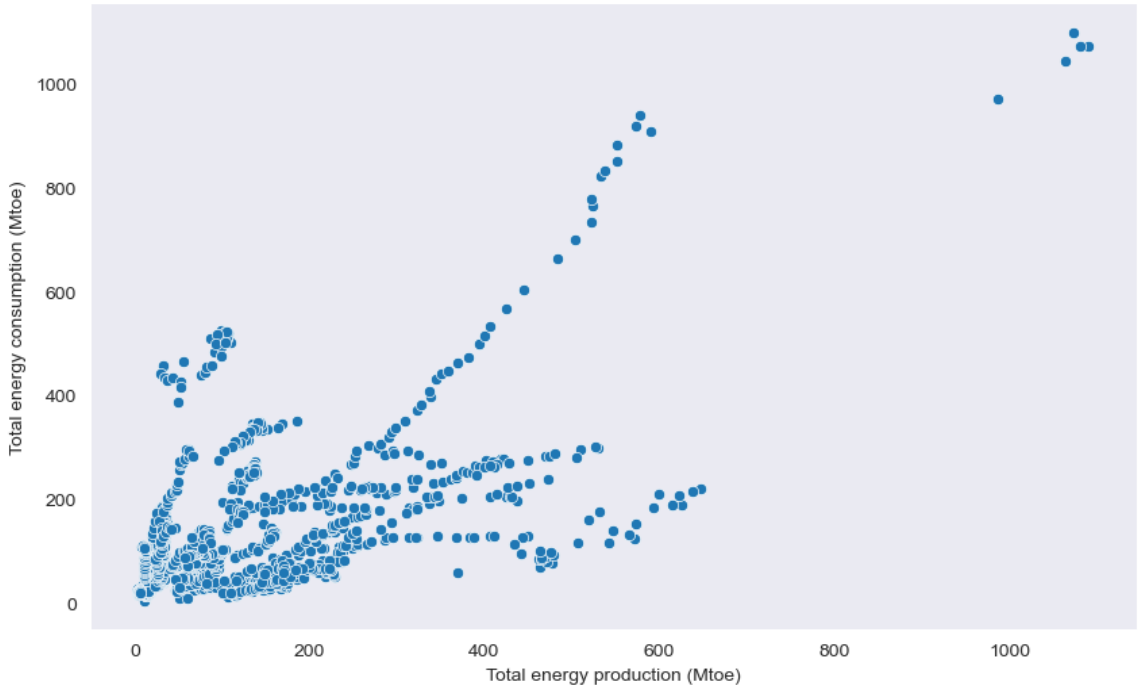


Out[47]:

```
<AxesSubplot:xlabel='CO2 emissions from fuel combustion (MtCO2)', ylabel
='Total energy production (Mtoe)'>
```

In [48]:

```
creat_plots('scatterplot',data1,'Total energy production (Mtoe)','Total energy consumptic
```

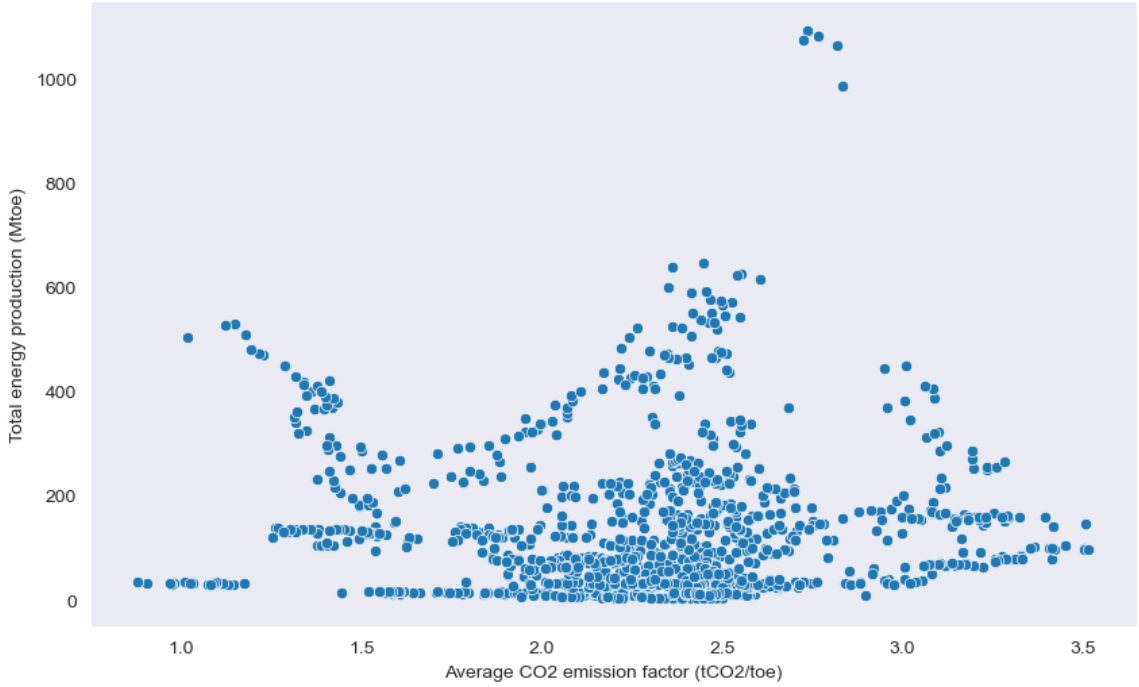


Out[48]:

```
<AxesSubplot:xlabel='Total energy production (Mtoe)', ylabel='Total energy
consumption (Mtoe)'>
```

In [49]:

```
creat_plots('scatterplot',data1,'Average CO2 emission factor (tCO2/toe)','Total energy pr
```



Out[49]:

```
<AxesSubplot:xlabel='Average CO2 emission factor (tCO2/toe)', ylabel='Tota
l energy production (Mtoe)'>
```

In [50]:

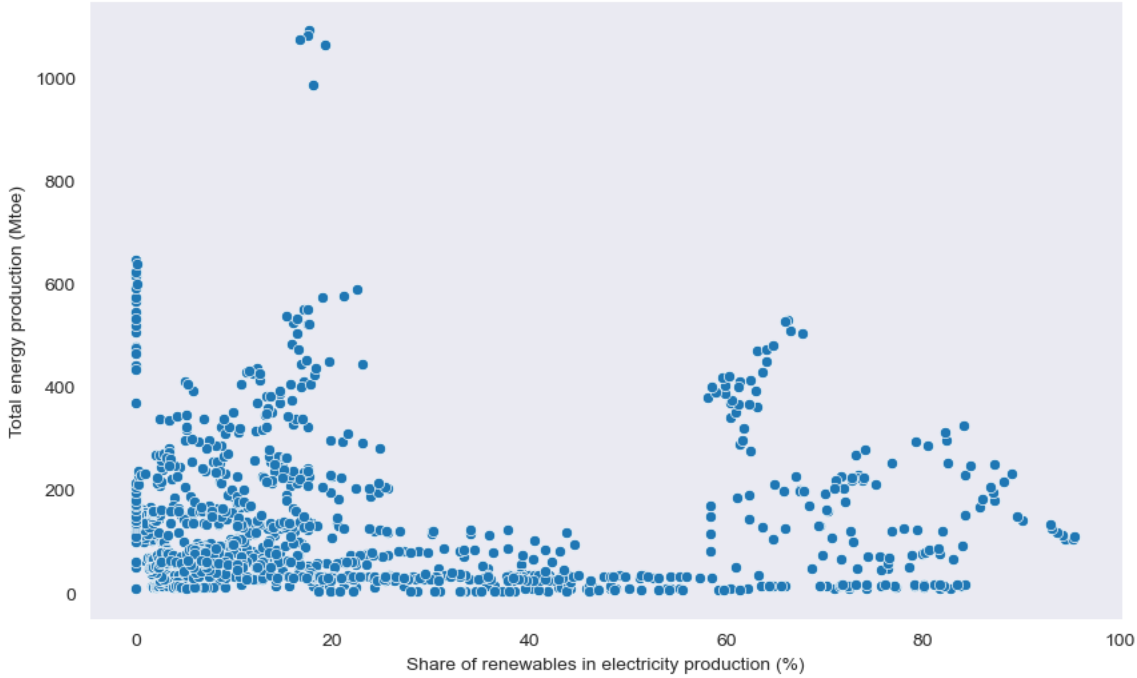
```
creat_plots('scatterplot',data1,'CO2 intensity at constant purchasing power parities (MtC
```

Out[50]:

```
<AxesSubplot:xlabel='CO2 intensity at constant purchasing power parities
(MtCO2/$274.8)', ylabel='Total energy production (Mtoe)'>
```

In [51]:

```
creat_plots('scatterplot',data1,'Share of renewables in electricity production (%)','Tota
```



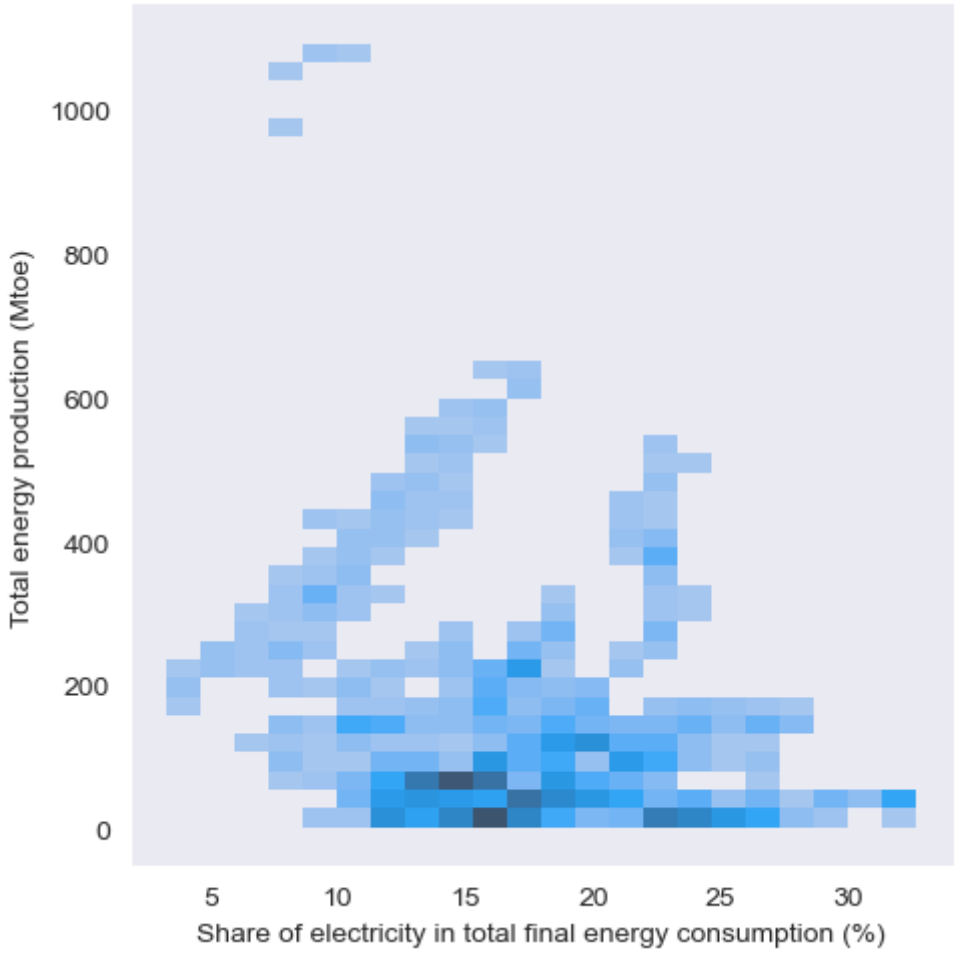
Out[51]:

```
<AxesSubplot:xlabel='Share of renewables in electricity production (%)', y
label='Total energy production (Mtoe)'>
```

In [52]:

```
creat_plots('displot',data1,'Share of electricity in total final energy consumption (%)',
```

```
<Figure size 1000x600 with 0 Axes>
```



Out[52]:

```
<seaborn.axisgrid.FacetGrid at 0x2963a23dcd0>
```

In [53]:

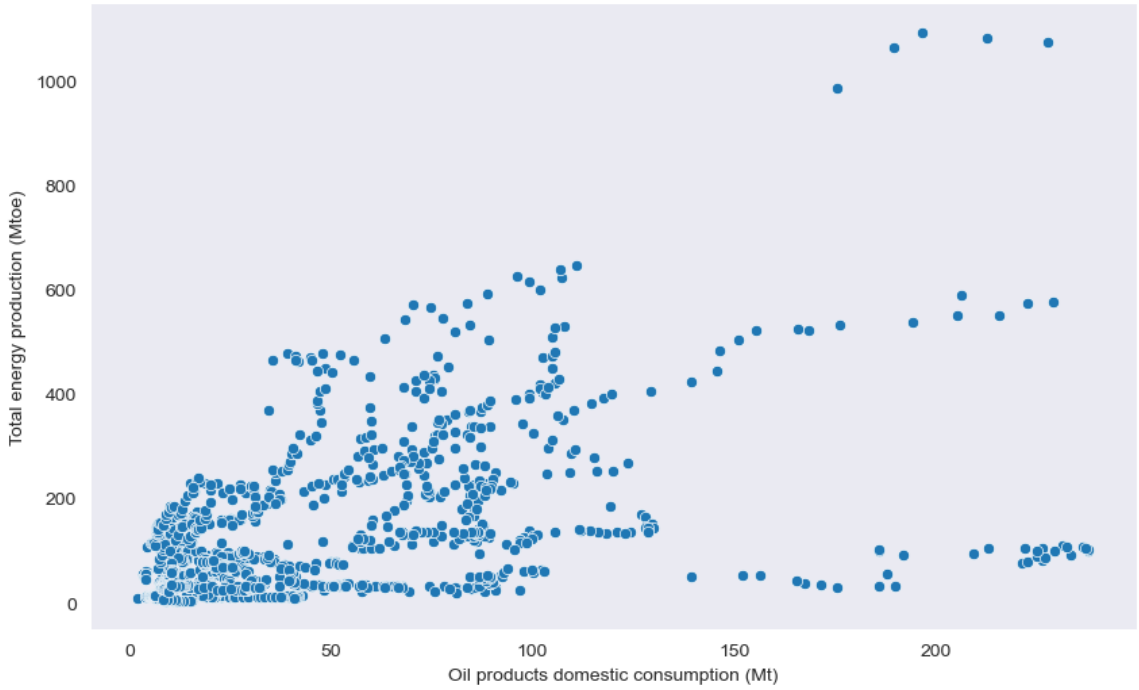
```
creat_plots('scatterplot',data1,'Share of electricity in total final energy consumption (
```

Out[53]:

```
<AxesSubplot:xlabel='Share of electricity in total final energy consumptio
n (%)', ylabel='Total energy production (Mtoe)'>
```

In [54]:

```
creat_plots('scatterplot',data1,'Oil products domestic consumption (Mt)','Total energy pr
```

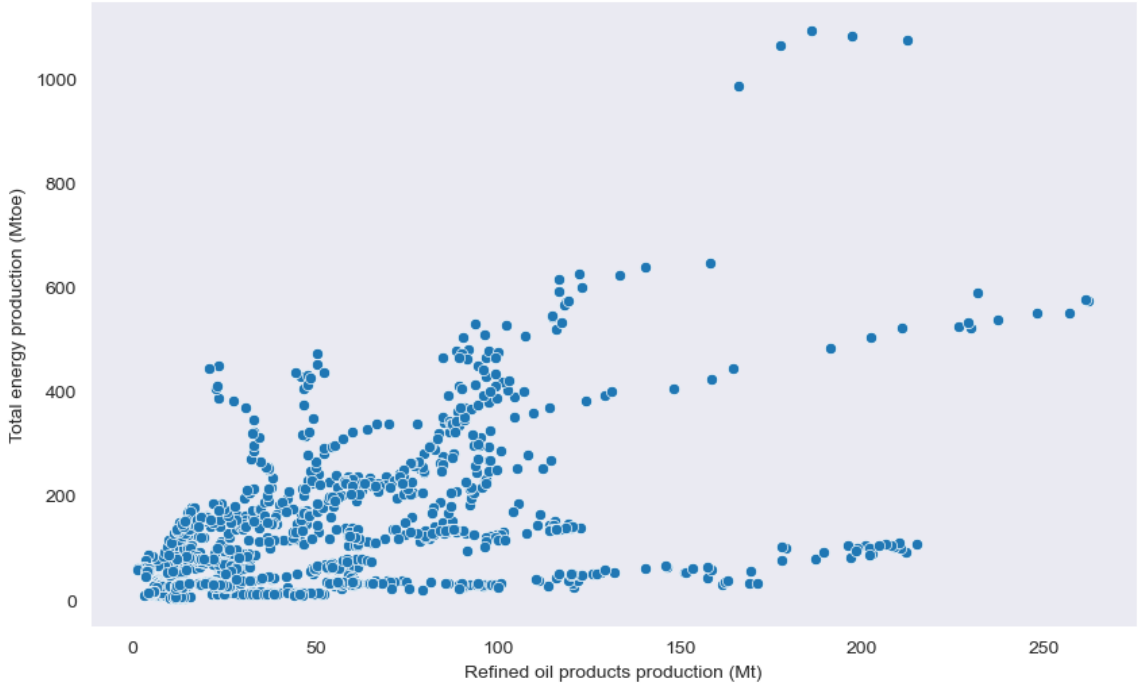


Out[54]:

```
<AxesSubplot:xlabel='Oil products domestic consumption (Mt)', ylabel='Tota
l energy production (Mtoe)'>
```

In [55]:

```
creat_plots('scatterplot',data1,'Refined oil products production (Mt)','Total energy proc
```

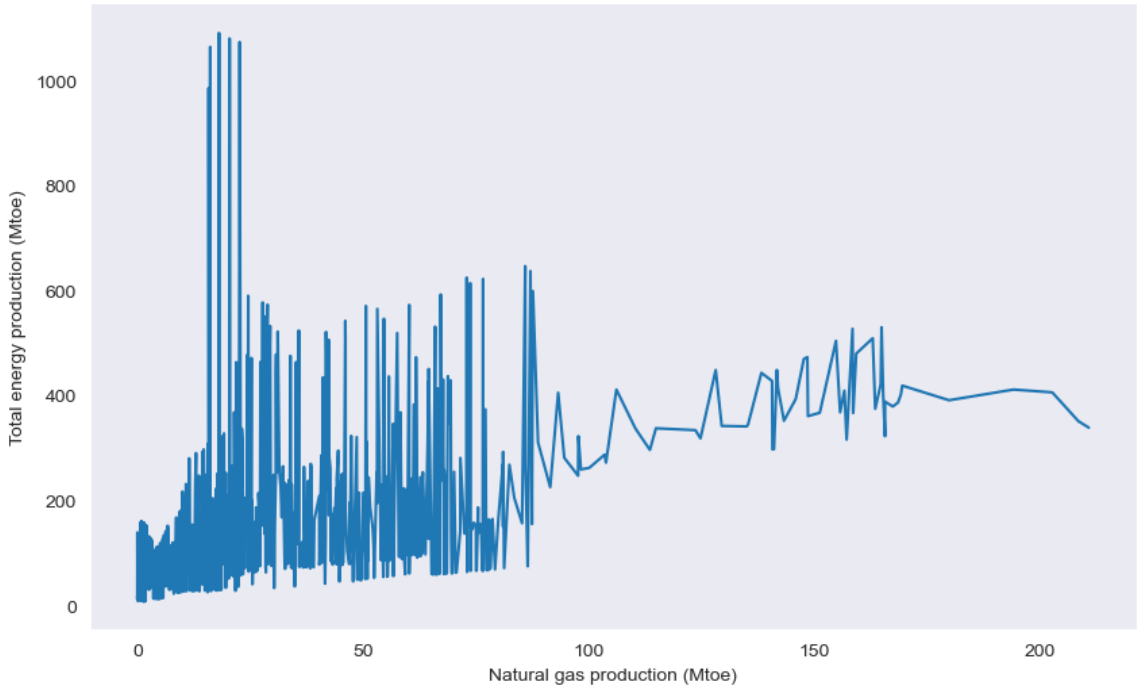


Out[55]:

```
<AxesSubplot:xlabel='Refined oil products production (Mt)', ylabel='Total
energy production (Mtoe)'>
```

In [56]:

```
creat_plots('lineplot',data1,'Natural gas production (Mtoe)','Total energy production (Mt
```

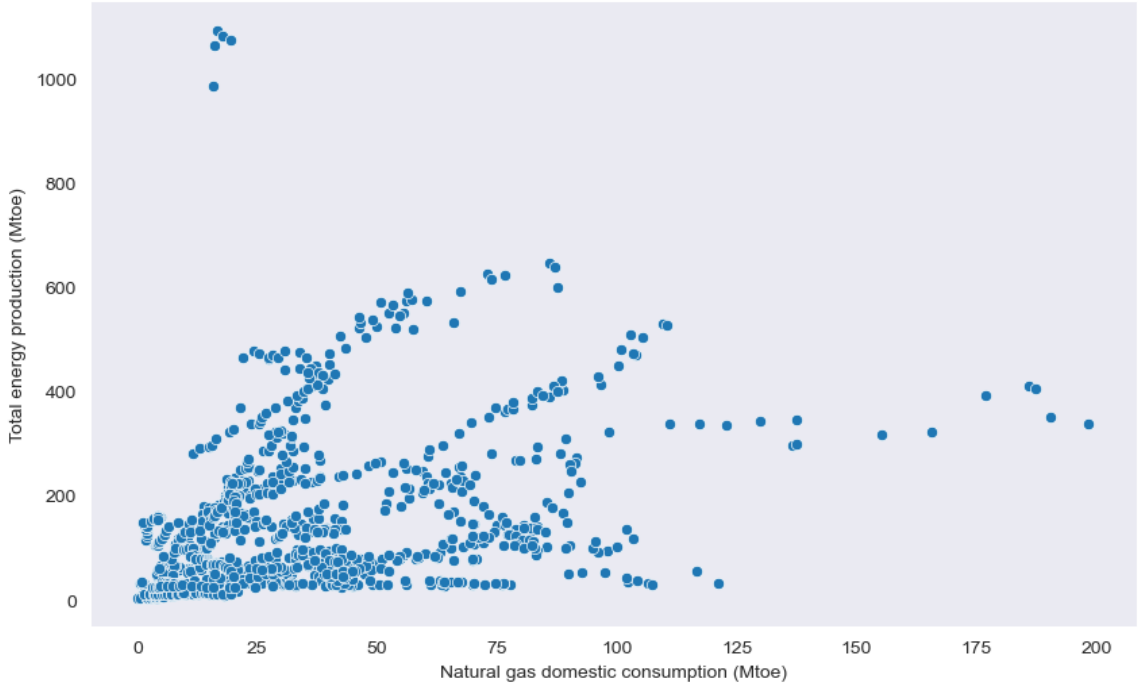


Out[56]:

```
<AxesSubplot:xlabel='Natural gas production (Mtoe)', ylabel='Total energy
production (Mtoe)'>
```

In [57]:

```
creat_plots('scatterplot',data1,'Natural gas domestic consumption (Mtoe)','Total energy p
```

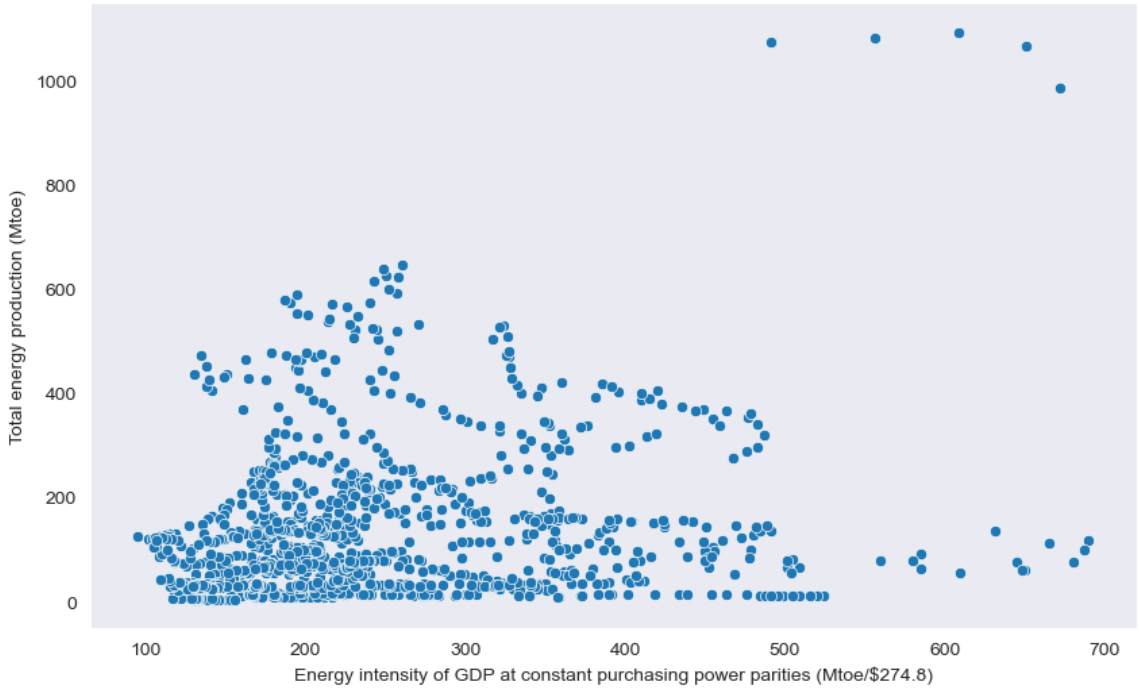


Out[57]:

```
<AxesSubplot:xlabel='Natural gas domestic consumption (Mtoe)', ylabel='Tot
al energy production (Mtoe)'>
```

In [58]:

```
creat_plots('scatterplot',data1,'Energy intensity of GDP at constant purchasing power par
```

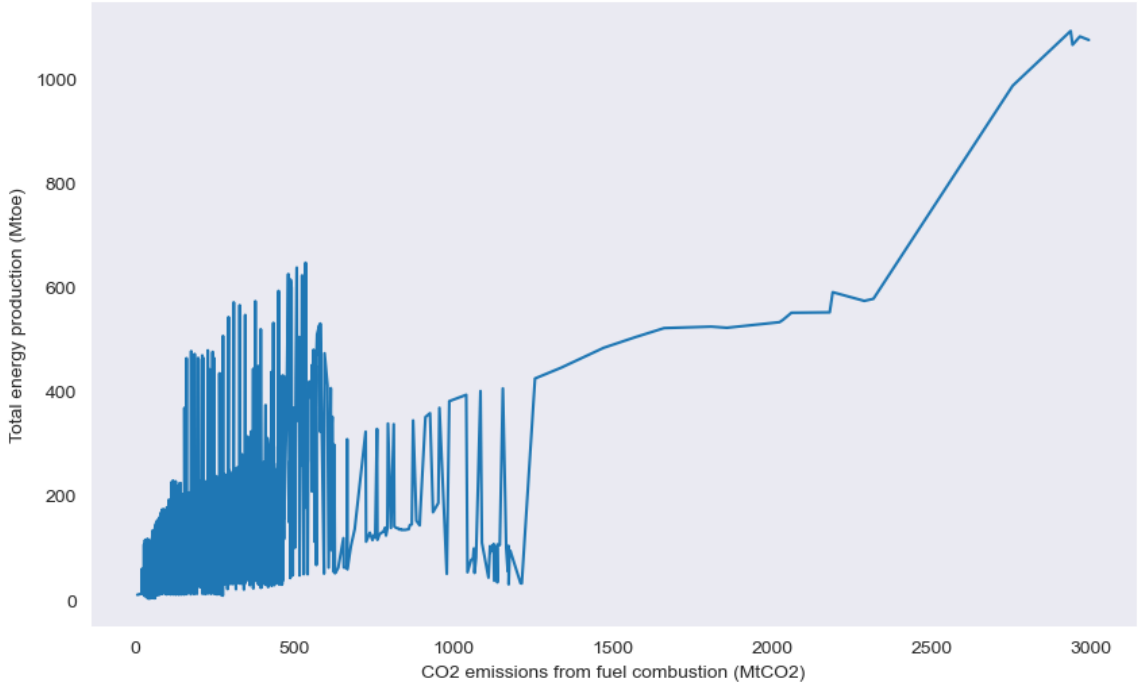


Out[58]:

```
<AxesSubplot:xlabel='Energy intensity of GDP at constant purchasing power
parities (Mtoe/$274.8)', ylabel='Total energy production (Mtoe)'>
```

In [59]:

```
creat_plots('lineplot',data1,'CO2 emissions from fuel combustion (MtCO2)','Total energy p
```

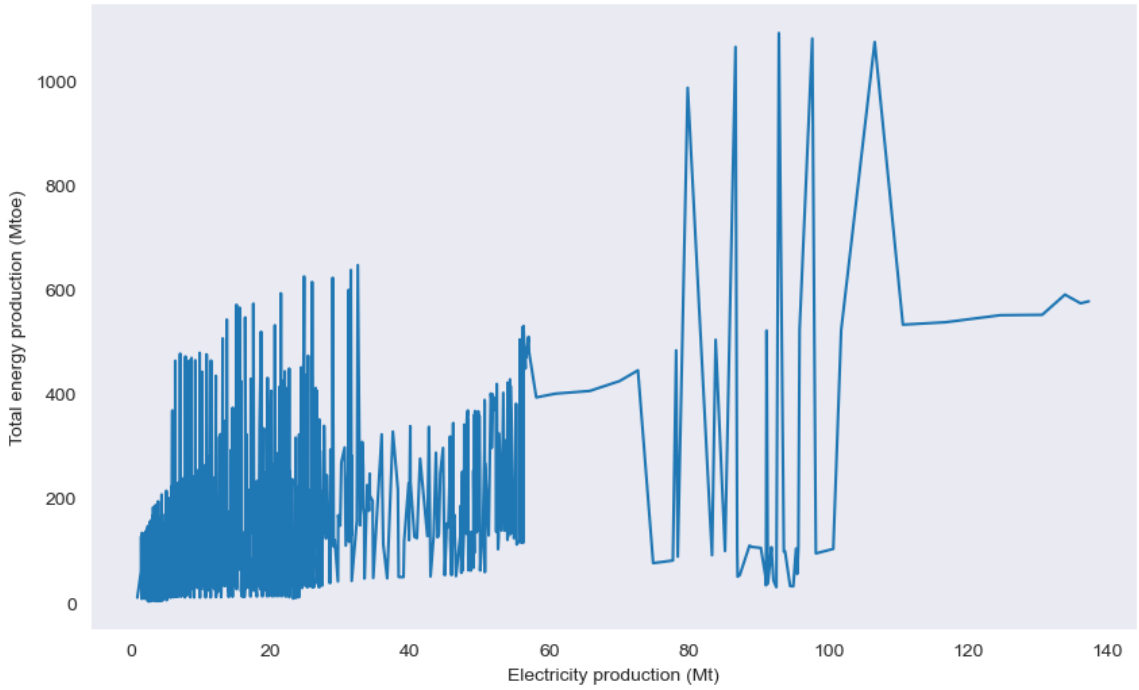


Out[59]:

```
<AxesSubplot:xlabel='CO2 emissions from fuel combustion (MtCO2)', ylabel
='Total energy production (Mtoe)'>
```

In [60]:

```
creat_plots('lineplot',data1,'Electricity production (Mt)','Total energy production (Mtoe
```

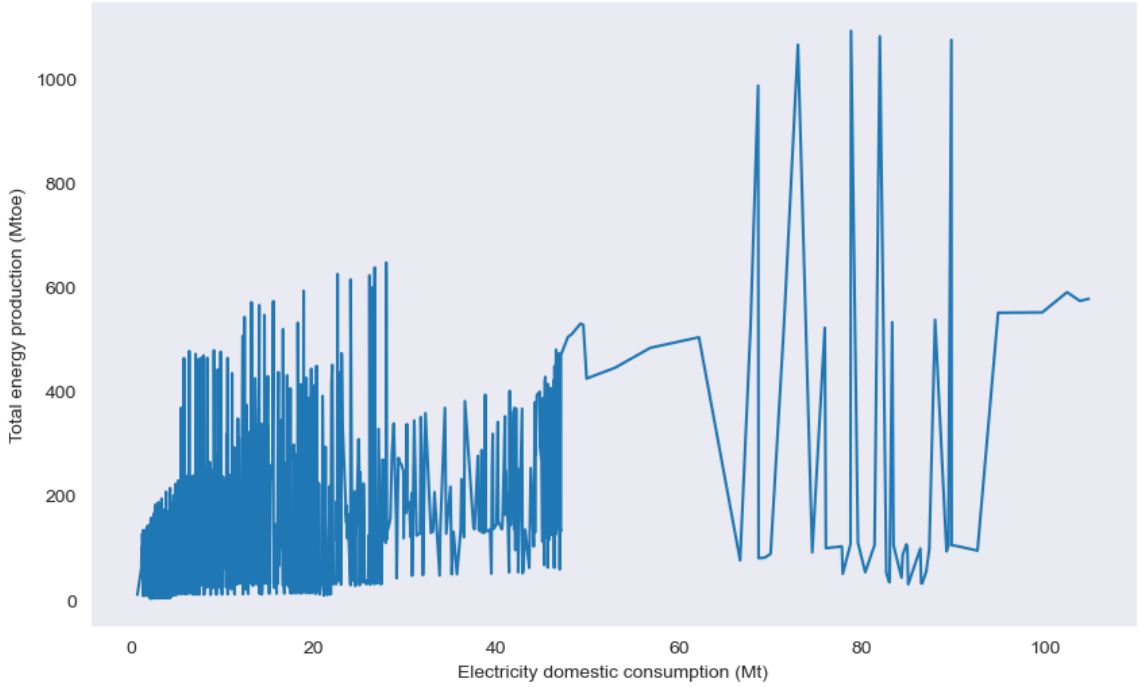


Out[60]:

```
<AxesSubplot:xlabel='Electricity production (Mt)', ylabel='Total energy pr
oduction (Mtoe)'>
```

In [61]:

```
creat_plots('lineplot',data1,'Electricity domestic consumption (Mt)','Total energy produc
```

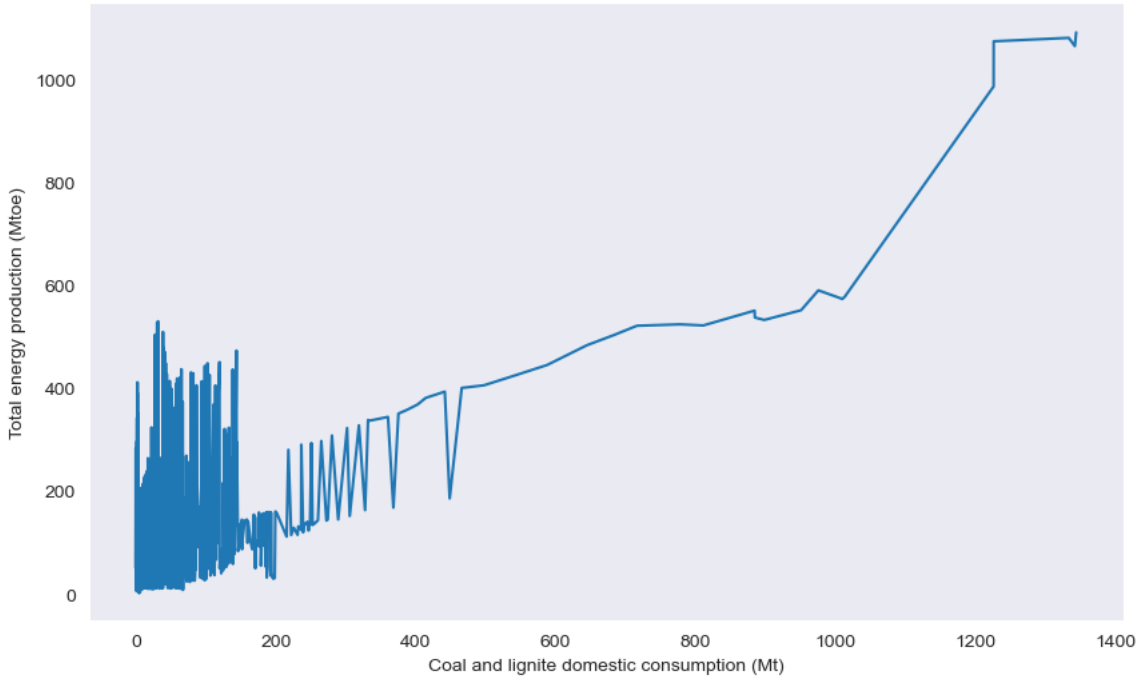


Out[61]:

```
<AxesSubplot:xlabel='Electricity domestic consumption (Mt)', ylabel='Total
energy production (Mtoe)'>
```

In [62]:

```
creat_plots('lineplot',data1,'Coal and lignite domestic consumption (Mt)','Total energy p
```

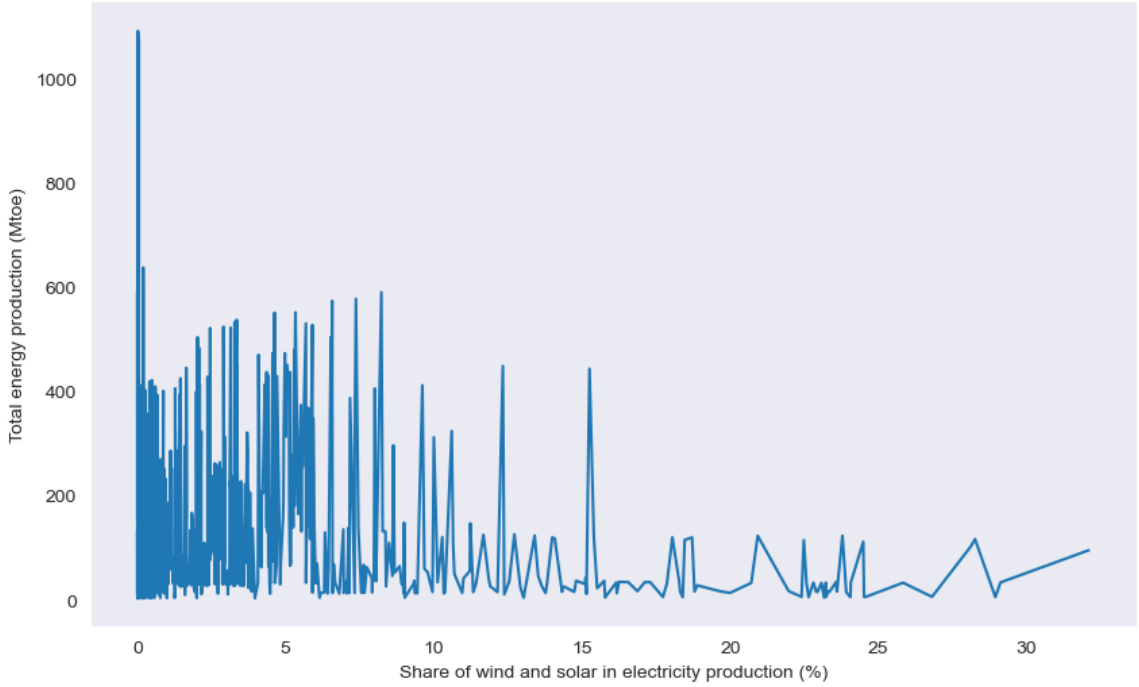


Out[62]:

```
<AxesSubplot:xlabel='Coal and lignite domestic consumption (Mt)', ylabel
='Total energy production (Mtoe)'>
```

In [63]:

```
creat_plots('lineplot',data1,'Share of wind and solar in electricity production (%)','Tot
```

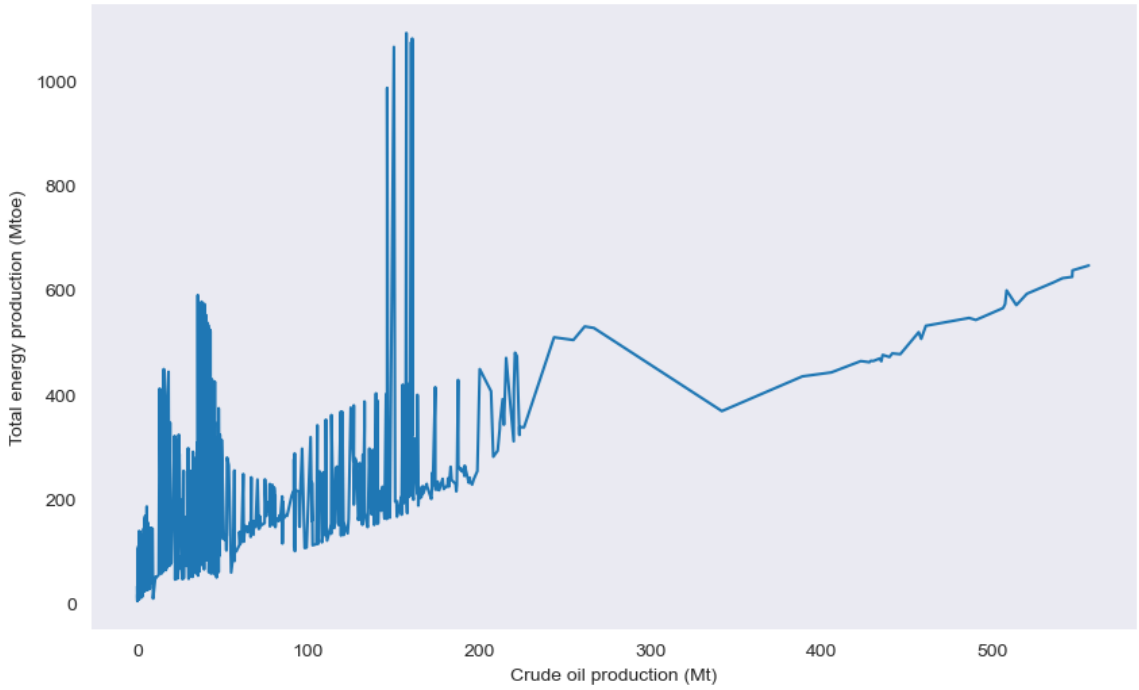


Out[63]:

```
<AxesSubplot:xlabel='Share of wind and solar in electricity production
(%)', ylabel='Total energy production (Mtoe)'>
```

In [64]:

```
creat_plots('lineplot',data1,'Crude oil production (Mt)','Total energy production (Mtoe)'
```

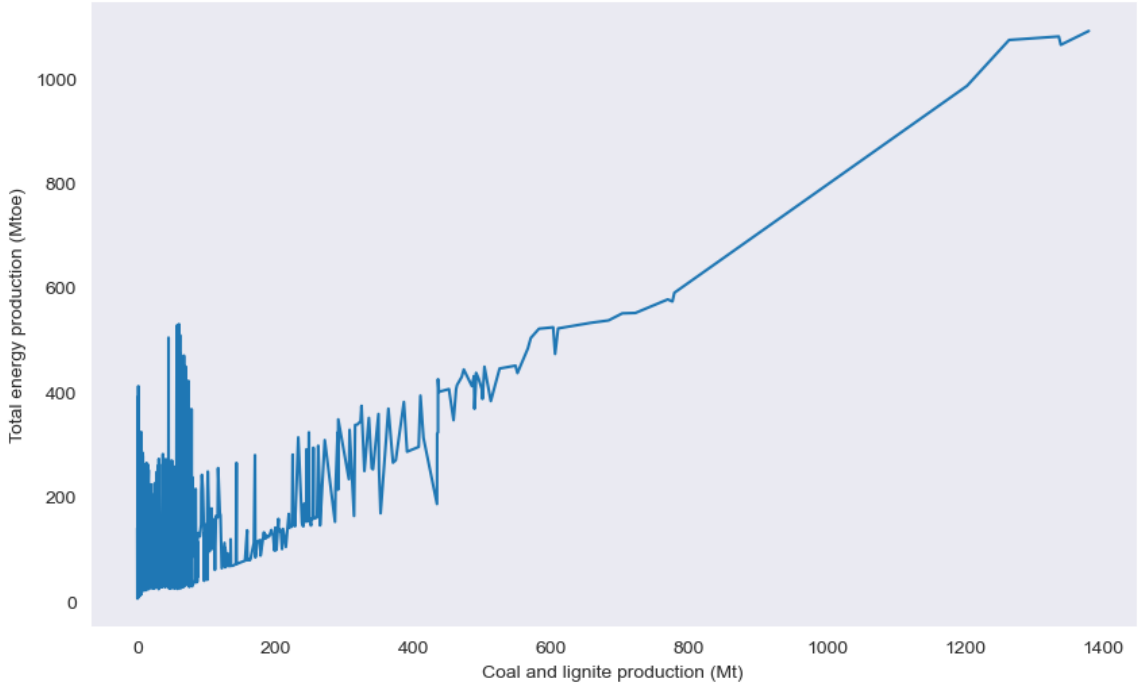


Out[64]:

```
<AxesSubplot:xlabel='Crude oil production (Mt)', ylabel='Total energy prod
uction (Mtoe)'>
```

In [65]:

```
creat_plots('lineplot',data1,'Coal and lignite production (Mt)','Total energy production
```



Out[65]:

```
<AxesSubplot:xlabel='Coal and lignite production (Mt)', ylabel='Total ener
gy production (Mtoe)'>
```

In [66]:

```
data1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1170 entries, 0 to 1363
Data columns (total 21 columns):
 #   Column
Non-Null Count  Dtype
---  ------
--------------  -----
 0   country
1170 non-null   object
 1   Year
1170 non-null   int64
 2   Region
1170 non-null   object
 3   CO2 emissions from fuel combustion (MtCO2)
1170 non-null   float64
 4   Average CO2 emission factor (tCO2/toe)
1170 non-null   float64
 5   CO2 intensity at constant purchasing power parities (MtCO2/$274.8)
1170 non-null   float64
 6   Total energy production (Mtoe)
1170 non-null   float64
 7   Total energy consumption (Mtoe)
1170 non-null   float64
 8   Share of renewables in electricity production (%)
1170 non-null   float64
 9   Share of electricity in total final energy consumption (%)
1170 non-null   float64
 10  Oil products domestic consumption (Mt)
1170 non-null   float64
 11  Refined oil products production (Mt)
1170 non-null   float64
 12  Natural gas production (Mtoe)
1170 non-null   float64
 13  Natural gas domestic consumption (Mtoe)
1170 non-null   float64
 14  Energy intensity of GDP at constant purchasing power parities (Mtoe/
$274.8)  1170 non-null   float64
 15  Electricity production (Mt)
1170 non-null   float64
 16  Electricity domestic consumption (Mt)
1170 non-null   float64
 17  Coal and lignite domestic consumption (Mt)
1170 non-null   float64
 18  Share of wind and solar in electricity production (%)
1170 non-null   float64
 19  Crude oil production (Mt)
1170 non-null   float64
 20  Coal and lignite production (Mt)
1170 non-null   float64
dtypes: float64(18), int64(1), object(2)
memory usage: 233.4+ KB
```

In [ ]:

```
```

In [67]:

```
data_group=data1.groupby('Region').sum()
```

In [68]:

```
data_group
```

Out[68]:

Region	Year	CO2 emissions from fuel combustion (MtCO2)	Average CO2 emission factor (tCO2/toe)	CO2 intensity at constant purchasing power parities (MtCO2/$274.8)	Total energy production (Mtoe)	Total energy consumption (Mtoe)
Africa	62155	11142.015898	93.635351	38409.807378	4544.077618	3663.103098
Arab States	62155	2763.776588	74.342492	14137.277922	4305.540015	1156.444038
Asia & Pacific	643750	150706.605921	762.100349	219718.651400	49265.604899	63549.168067
Europe	838061	114341.778017	939.417832	204218.638980	26450.170217	50439.518114
Middle east	304726	32216.989233	379.730184	96415.426608	35324.430351	13012.680135
North America	62155	16247.117248	41.009333	16288.320116	12497.632166	8047.297411
South/Latin America	372930	34380.914781	383.966055	71179.067139	23907.312216	17605.594699

In [ ]:

```
```

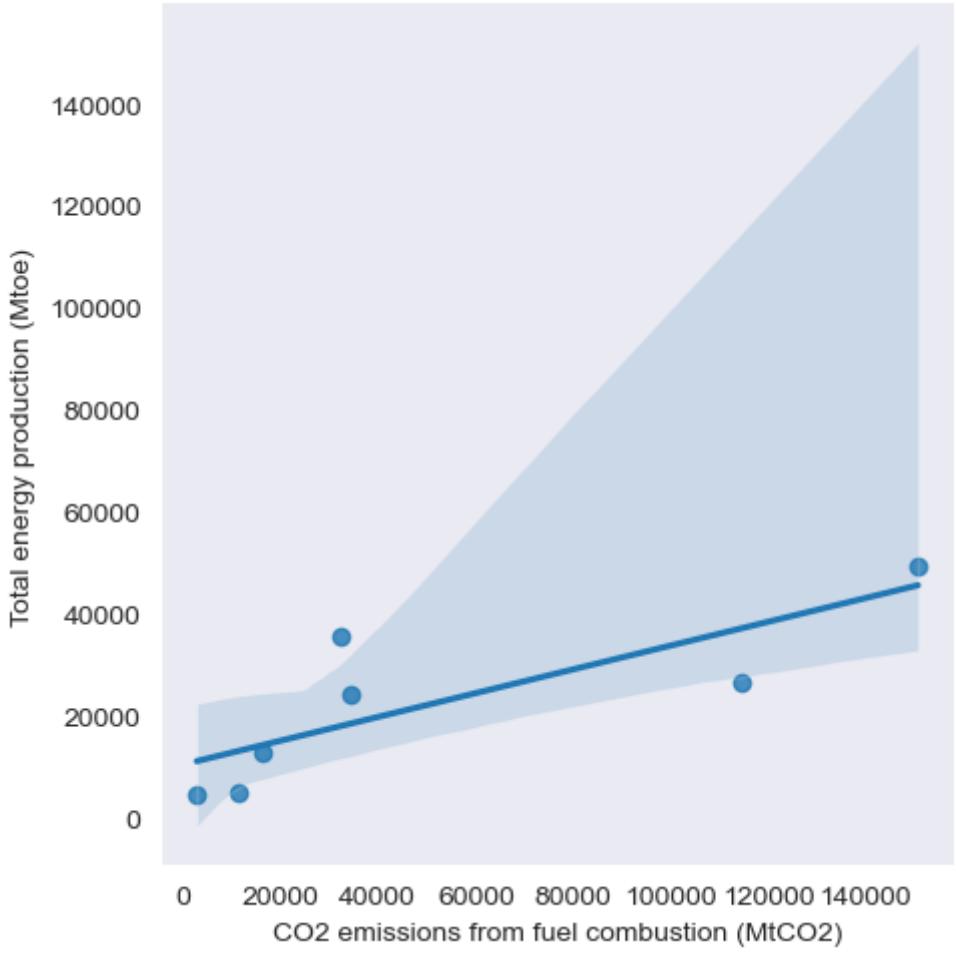
In [ ]:

```
```

In [69]:

```
creat_plots('lmplot',data_group,'CO2 emissions from fuel combustion (MtCO2)','Total energ
```

```
<Figure size 1000x600 with 0 Axes>
```

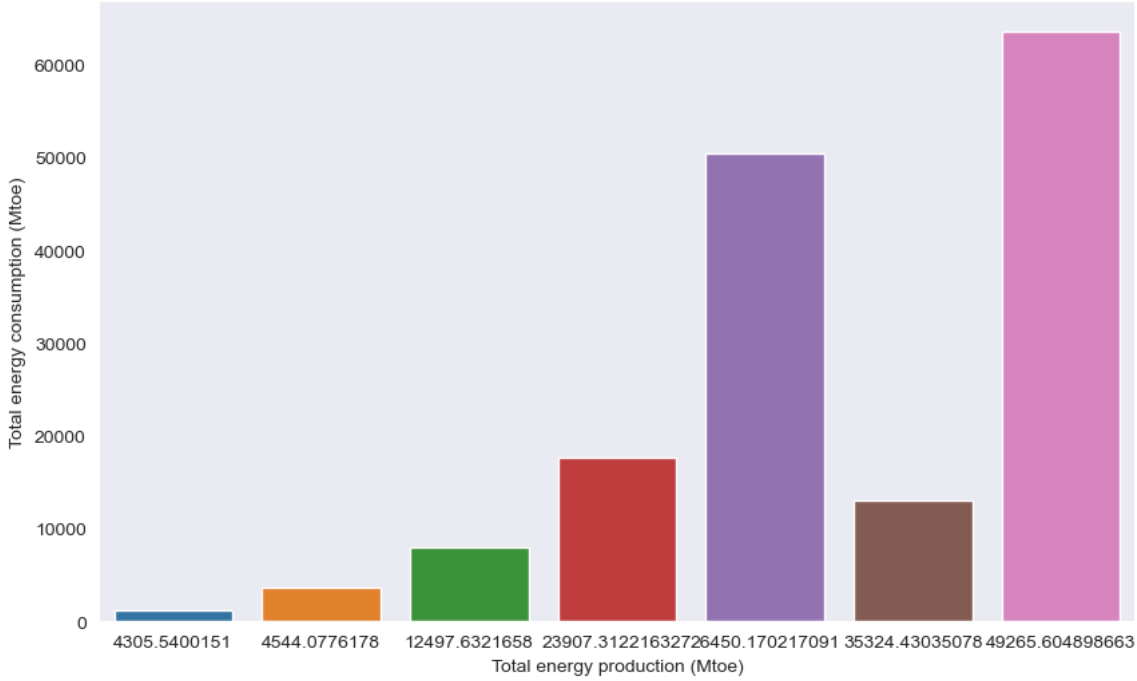


Out[69]:

```
<seaborn.axisgrid.FacetGrid at 0x2963e0efb50>
```

In [70]:

```
creat_plots('barplot',data_group,'Total energy production (Mtoe)','Total energy consumpti
```



Out[70]:

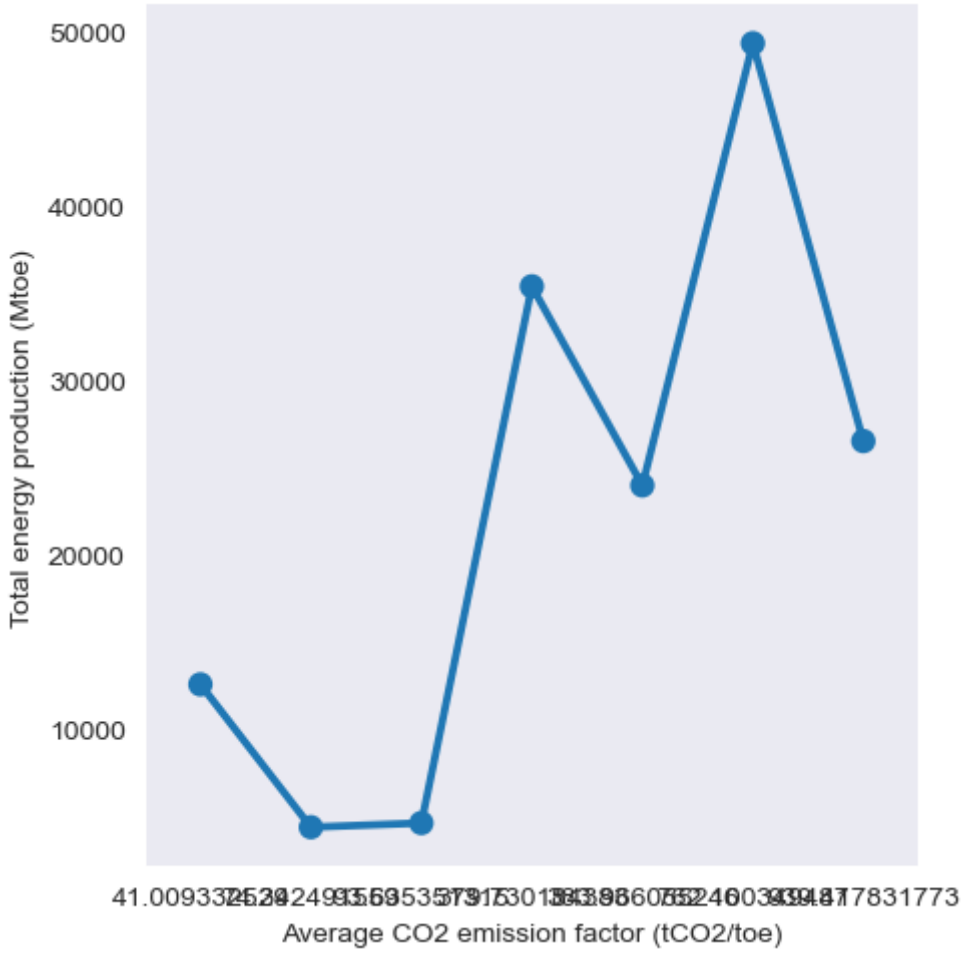
```
<AxesSubplot:xlabel='Total energy production (Mtoe)', ylabel='Total energy
consumption (Mtoe)'>
```

In [71]:

```
creat_plots('factorplot',data_group,'Average CO2 emission factor (tCO2/toe)','Total energ
```

```
C:\Users\user\anaconda3\lib\site-packages\seaborn\categorical.py:3717: Use
rWarning: The `factorplot` function has been renamed to `catplot`. The ori
ginal name will be removed in a future release. Please update your code. N
ote that the default `kind` in `factorplot` (`'point'`) has changed `'stri
p'` in `catplot`.
  warnings.warn(msg)
```

```
<Figure size 1000x600 with 0 Axes>
```



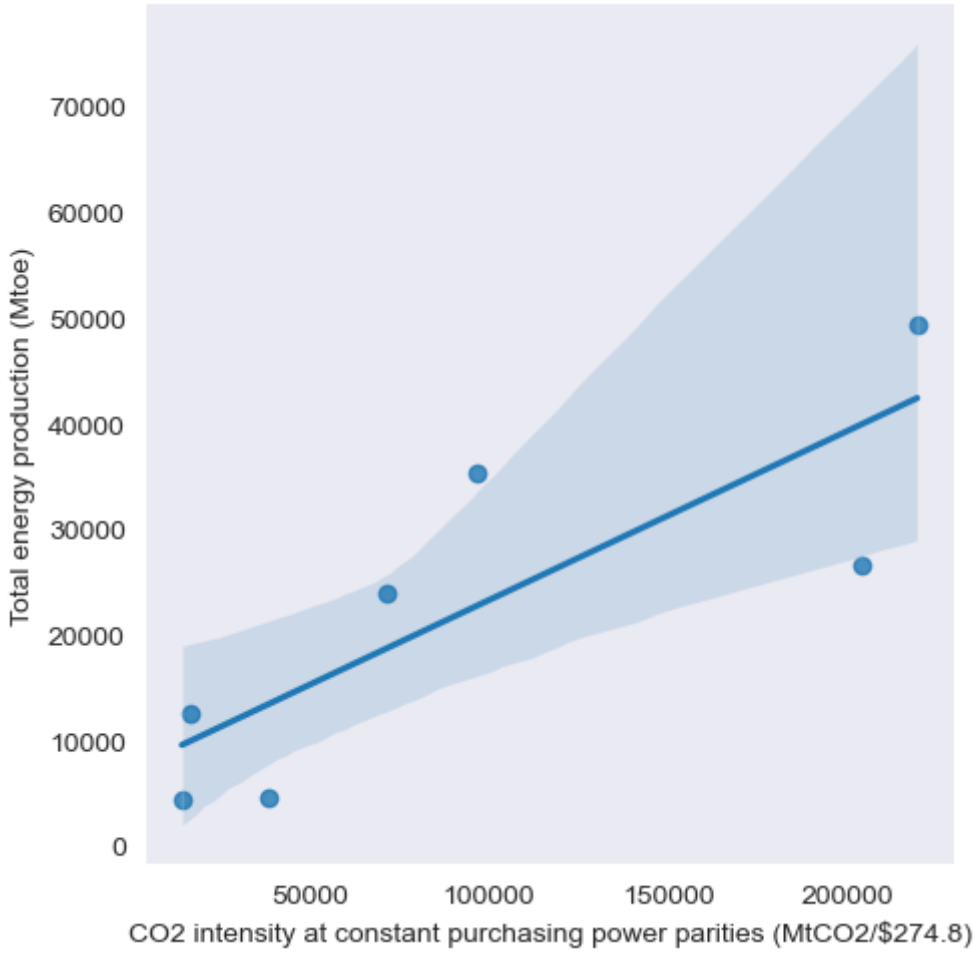
Out[71]:

```
<seaborn.axisgrid.FacetGrid at 0x2963e4b5a30>
```

In [72]:

```
creat_plots('lmplot',data_group,'CO2 intensity at constant purchasing power parities (MtC
```

<Figure size 1000x600 with 0 Axes>

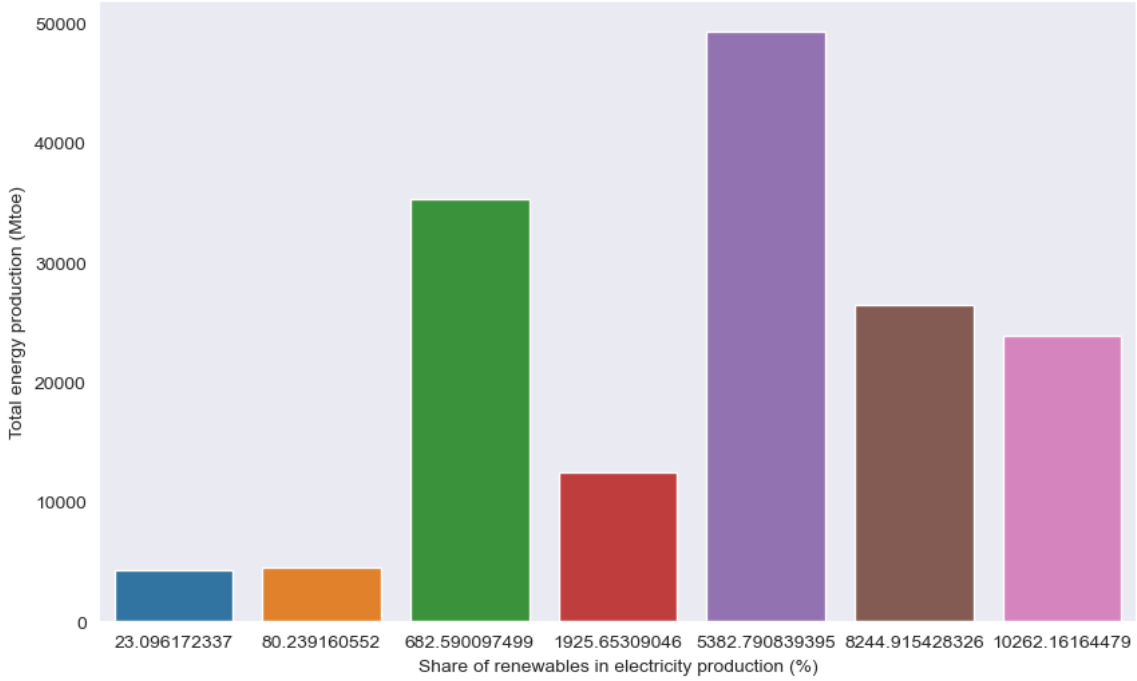


Out[72]:

<seaborn.axisgrid.FacetGrid at 0x2963e0dab80>

In [73]:

```
creat_plots('barplot',data_group,'Share of renewables in electricity production (%)','Tot
```



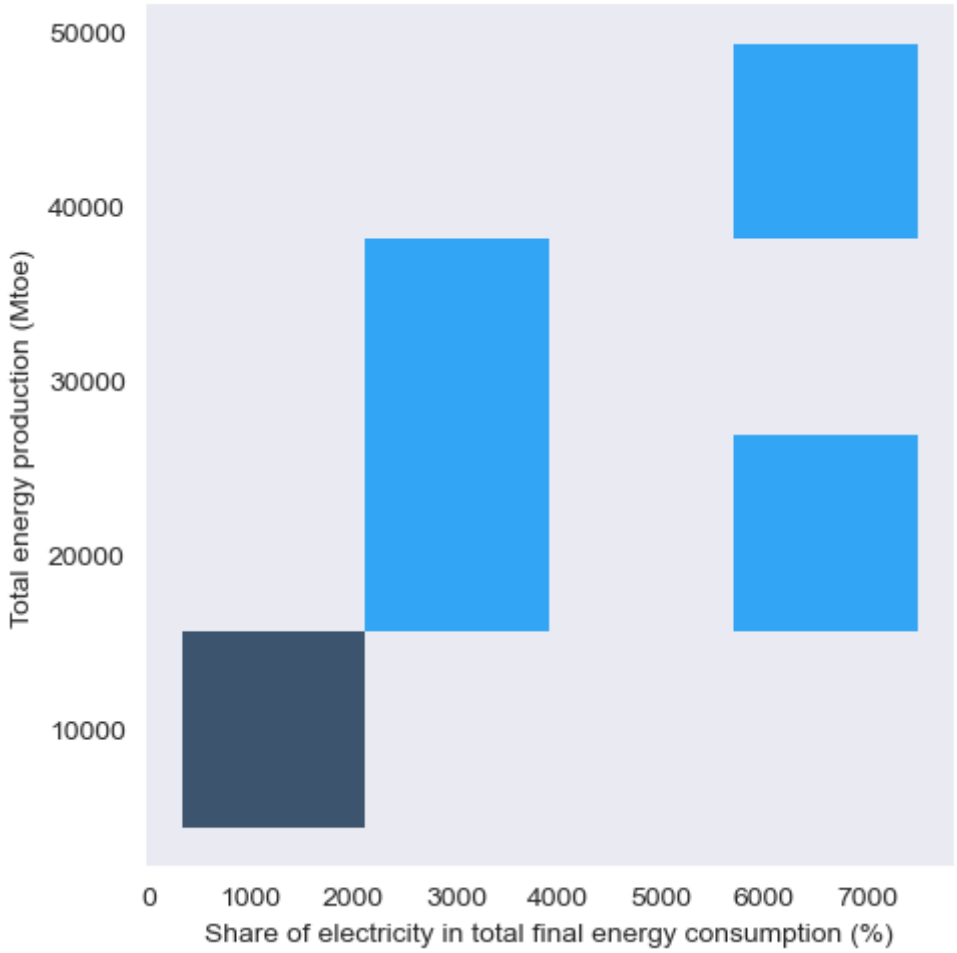
Out[73]:

```
<AxesSubplot:xlabel='Share of renewables in electricity production (%)', y
label='Total energy production (Mtoe)'>
```

In [74]:

```
creat_plots('displot',data_group,'Share of electricity in total final energy consumption
```

```
<Figure size 1000x600 with 0 Axes>
```



Out[74]:

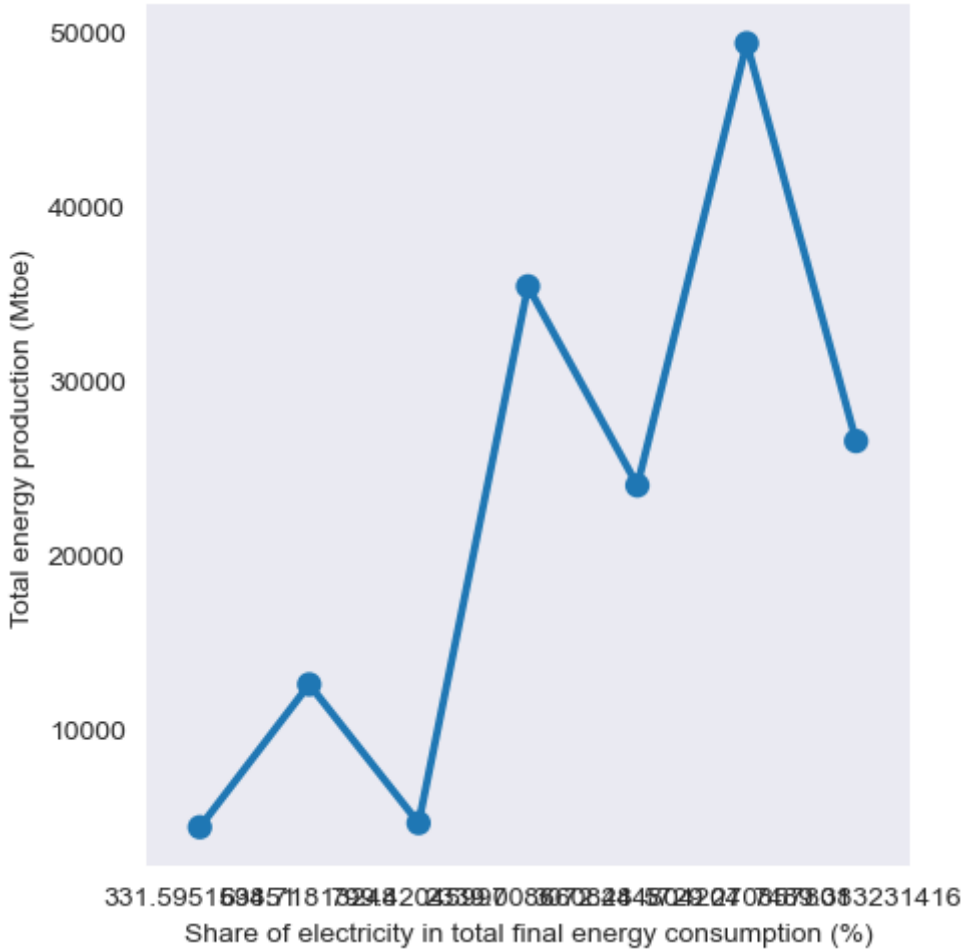
```
<seaborn.axisgrid.FacetGrid at 0x2963e6672e0>
```

In [75]:

```
creat_plots('factorplot',data_group,'Share of electricity in total final energy consumpti
```

```
C:\Users\user\anaconda3\lib\site-packages\seaborn\categorical.py:3717: Use
rWarning: The `factorplot` function has been renamed to `catplot`. The ori
ginal name will be removed in a future release. Please update your code. N
ote that the default `kind` in `factorplot` (`'point'`) has changed `'stri
p'` in `catplot`.
  warnings.warn(msg)
```

```
<Figure size 1000x600 with 0 Axes>
```

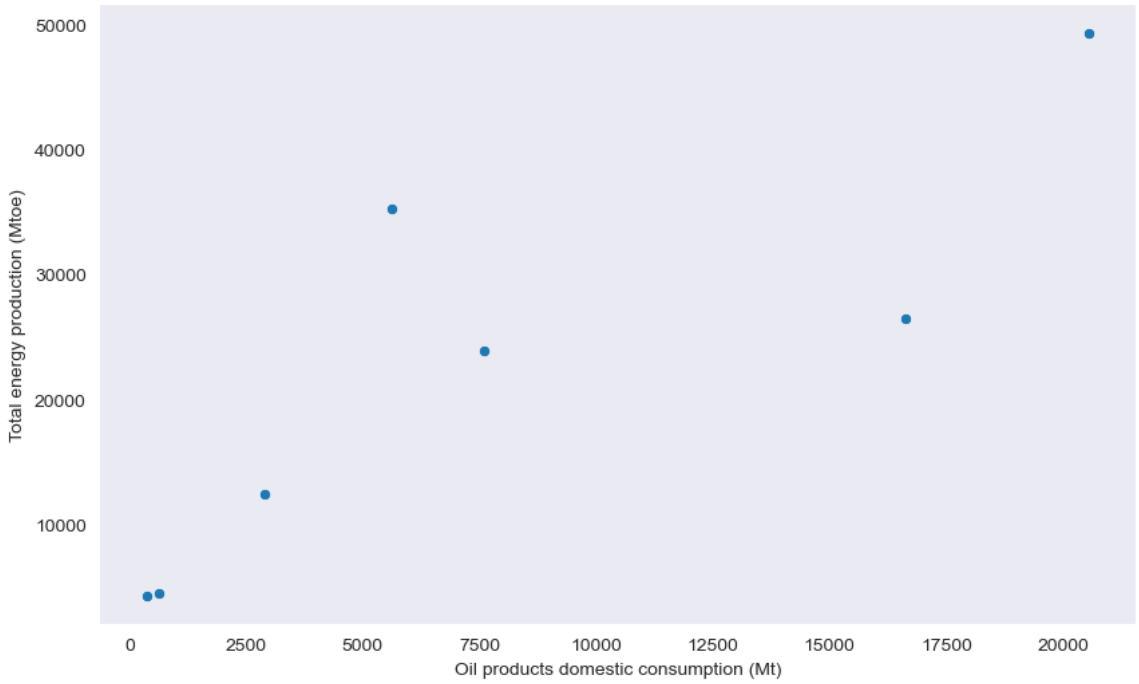


Out[75]:

```
<seaborn.axisgrid.FacetGrid at 0x2963e9471f0>
```

In [76]:

```
creat_plots('scatterplot',data_group,'Oil products domestic consumption (Mt)','Total ener
```

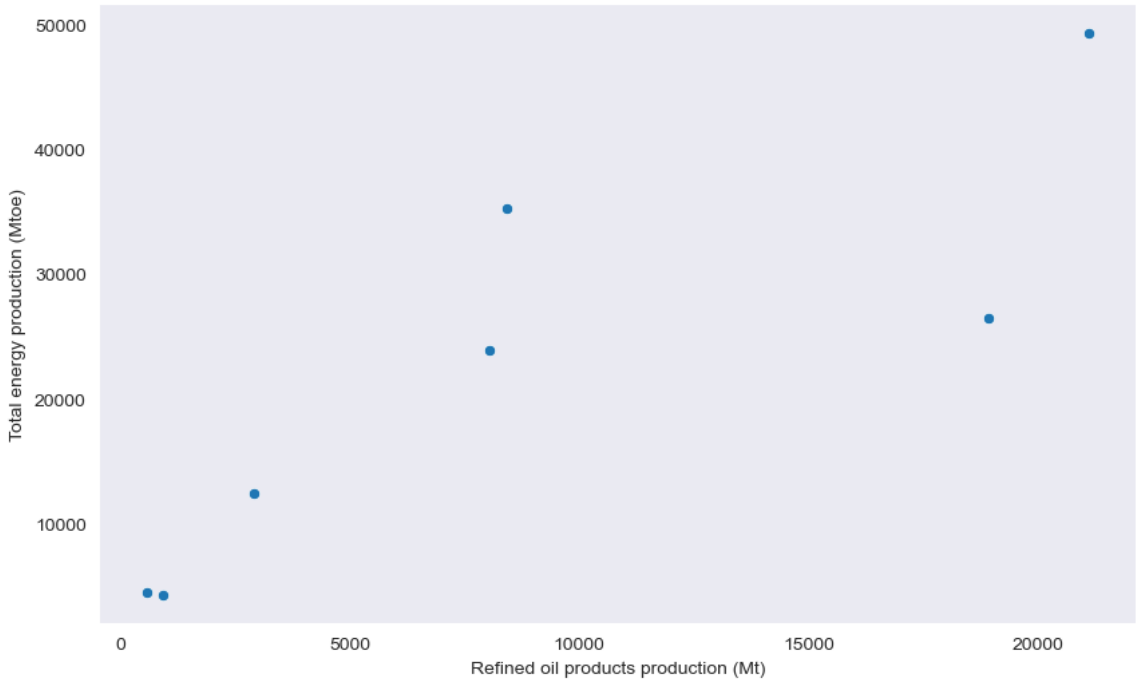


Out[76]:

```
<AxesSubplot:xlabel='Oil products domestic consumption (Mt)', ylabel='Tota
l energy production (Mtoe)'>
```

In [77]:

```
creat_plots('scatterplot',data_group,'Refined oil products production (Mt)','Total energy
```



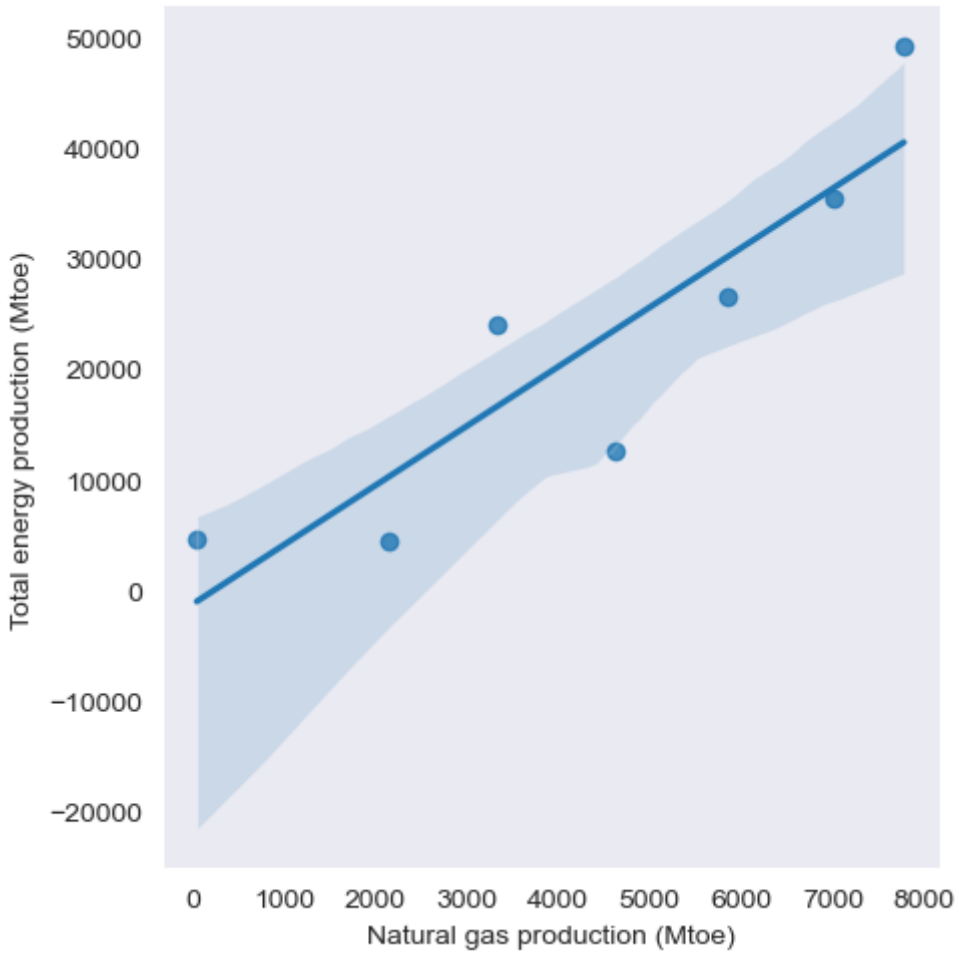
Out[77]:

```
<AxesSubplot:xlabel='Refined oil products production (Mt)', ylabel='Total
energy production (Mtoe)'>
```

In [78]:

```
creat_plots('lmplot',data_group,'Natural gas production (Mtoe)','Total energy production
```

```
<Figure size 1000x600 with 0 Axes>
```

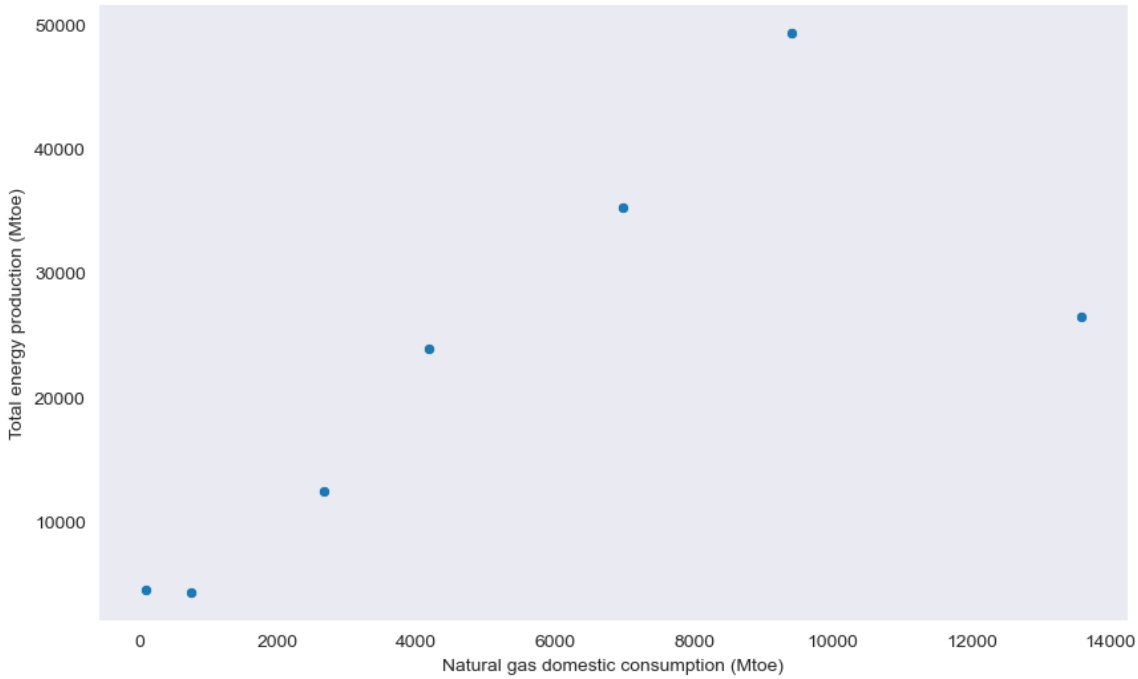


Out[78]:

```
<seaborn.axisgrid.FacetGrid at 0x2963fd6f370>
```

In [79]:

```
creat_plots('scatterplot',data_group,'Natural gas domestic consumption (Mtoe)','Total ene
```

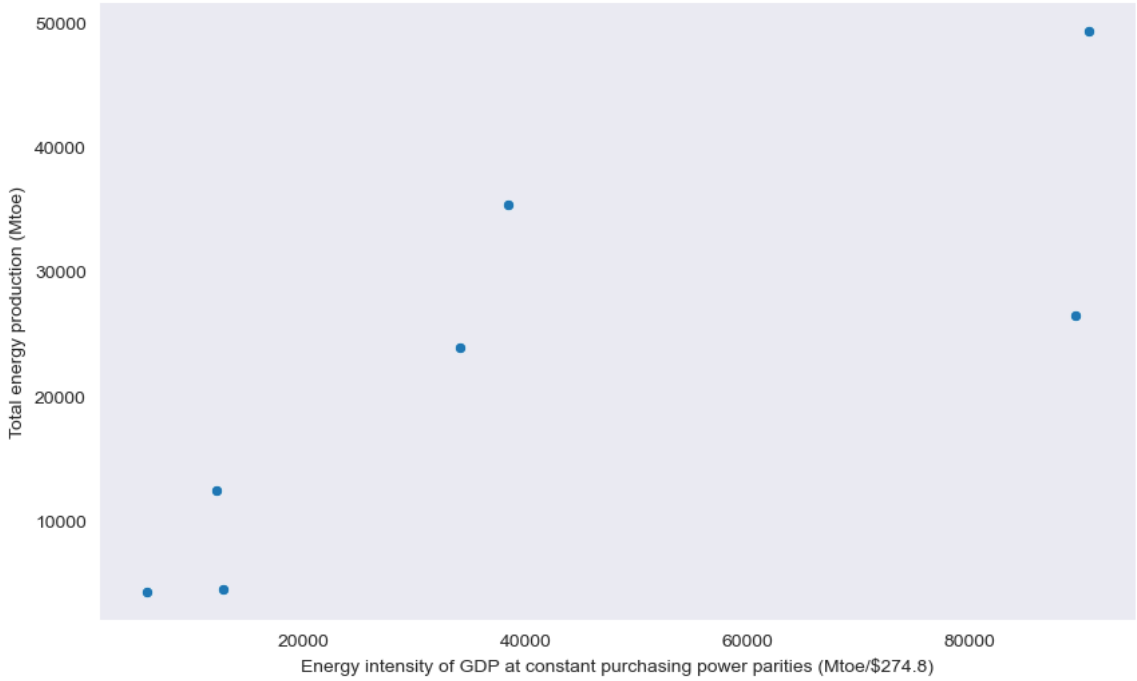


Out[79]:

```
<AxesSubplot:xlabel='Natural gas domestic consumption (Mtoe)', ylabel='Tot
al energy production (Mtoe)'>
```

In [80]:

```
creat_plots('scatterplot',data_group,'Energy intensity of GDP at constant purchasing powe
```

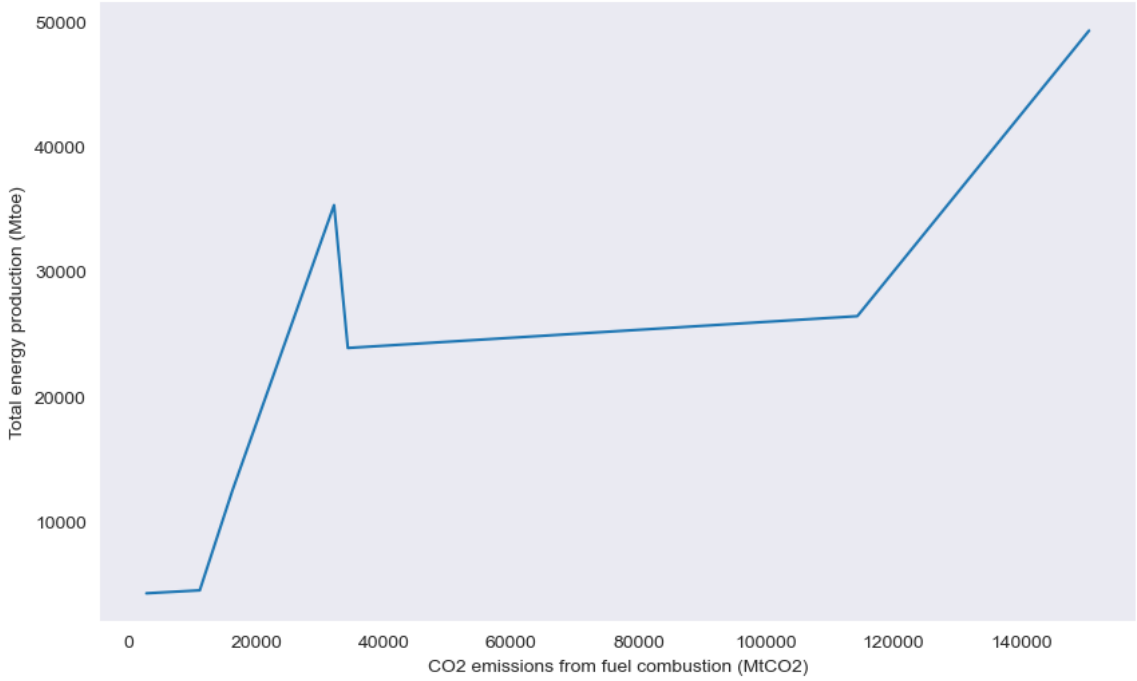


Out[80]:

```
<AxesSubplot:xlabel='Energy intensity of GDP at constant purchasing power
parities (Mtoe/$274.8)', ylabel='Total energy production (Mtoe)'>
```

In [81]:

```
creat_plots('lineplot',data_group,'CO2 emissions from fuel combustion (MtCO2)','Total ene
```

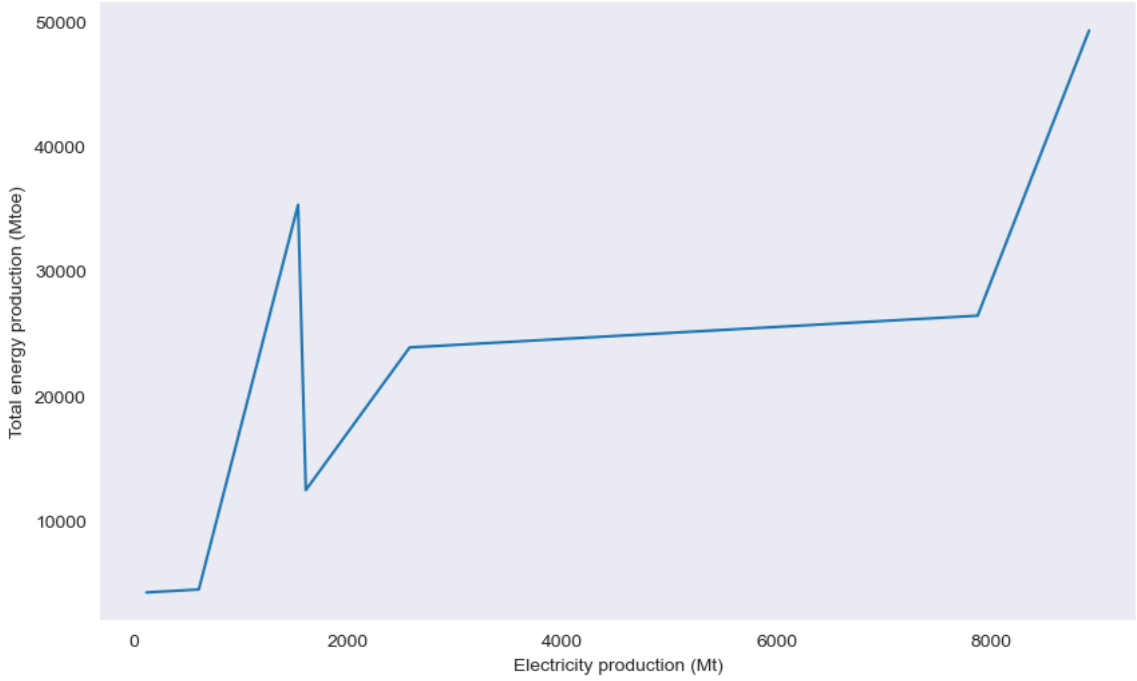


Out[81]:

```
<AxesSubplot:xlabel='CO2 emissions from fuel combustion (MtCO2)', ylabel
='Total energy production (Mtoe)'>
```

In [82]:

```
creat_plots('lineplot',data_group,'Electricity production (Mt)','Total energy production
```

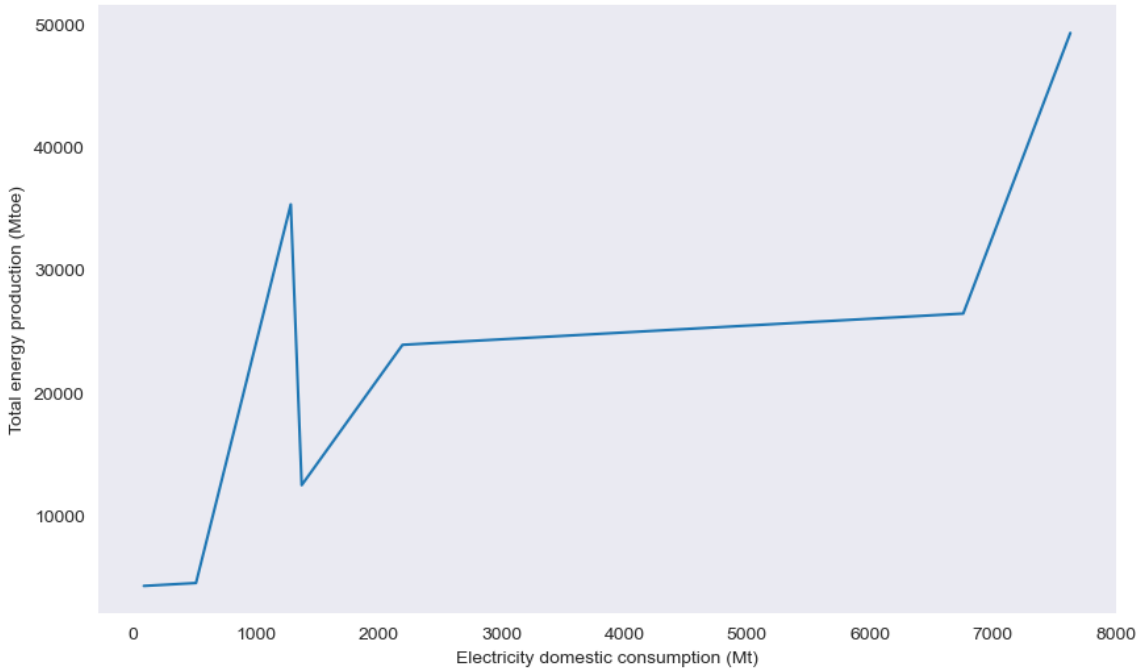


Out[82]:

```
<AxesSubplot:xlabel='Electricity production (Mt)', ylabel='Total energy pr
oduction (Mtoe)'>
```

In [83]:

```
creat_plots('lineplot',data_group,'Electricity domestic consumption (Mt)','Total energy p
```

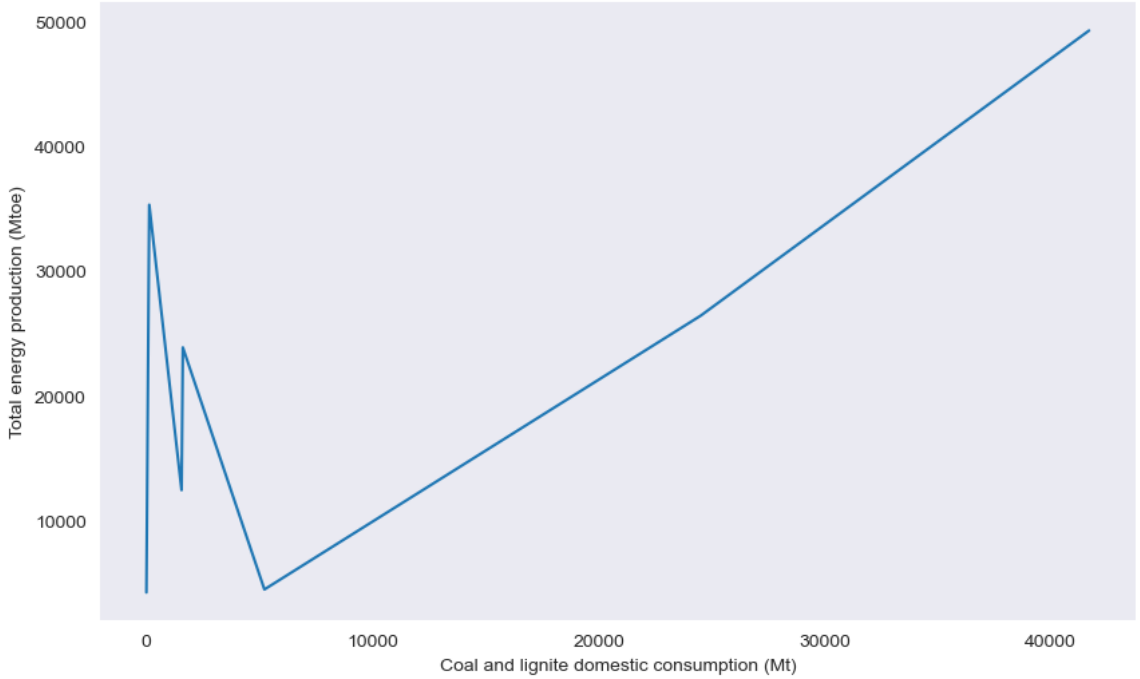


Out[83]:

```
<AxesSubplot:xlabel='Electricity domestic consumption (Mt)', ylabel='Total
energy production (Mtoe)'>
```

In [84]:

```
creat_plots('lineplot',data_group,'Coal and lignite domestic consumption (Mt)','Total ene
```

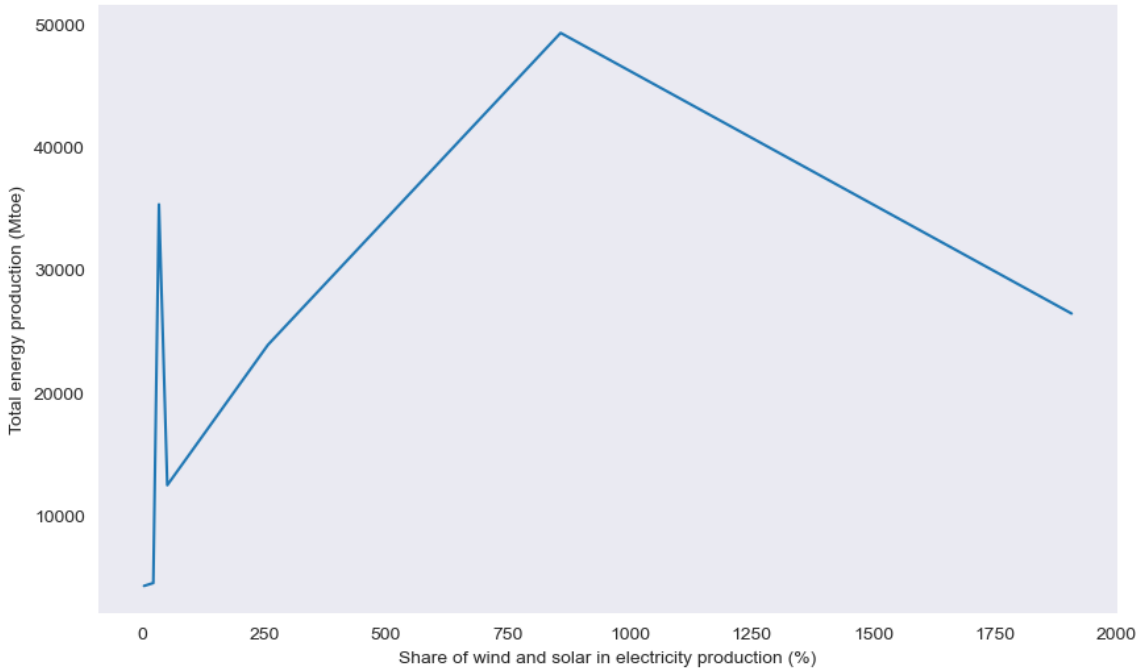


Out[84]:

```
<AxesSubplot:xlabel='Coal and lignite domestic consumption (Mt)', ylabel
='Total energy production (Mtoe)'>
```

In [85]:

```
creat_plots('lineplot',data_group,'Share of wind and solar in electricity production (%)'
```

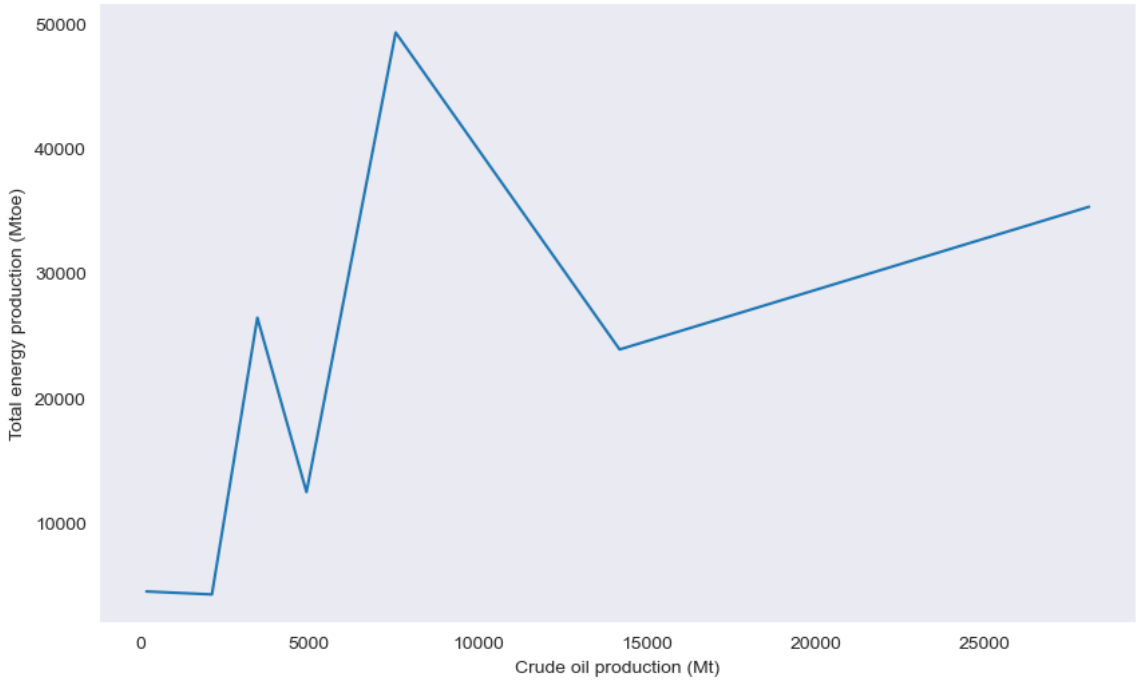


Out[85]:

```
<AxesSubplot:xlabel='Share of wind and solar in electricity production
(%)', ylabel='Total energy production (Mtoe)'>
```

In [86]:

```
creat_plots('lineplot',data_group,'Crude oil production (Mt)','Total energy production (M
```

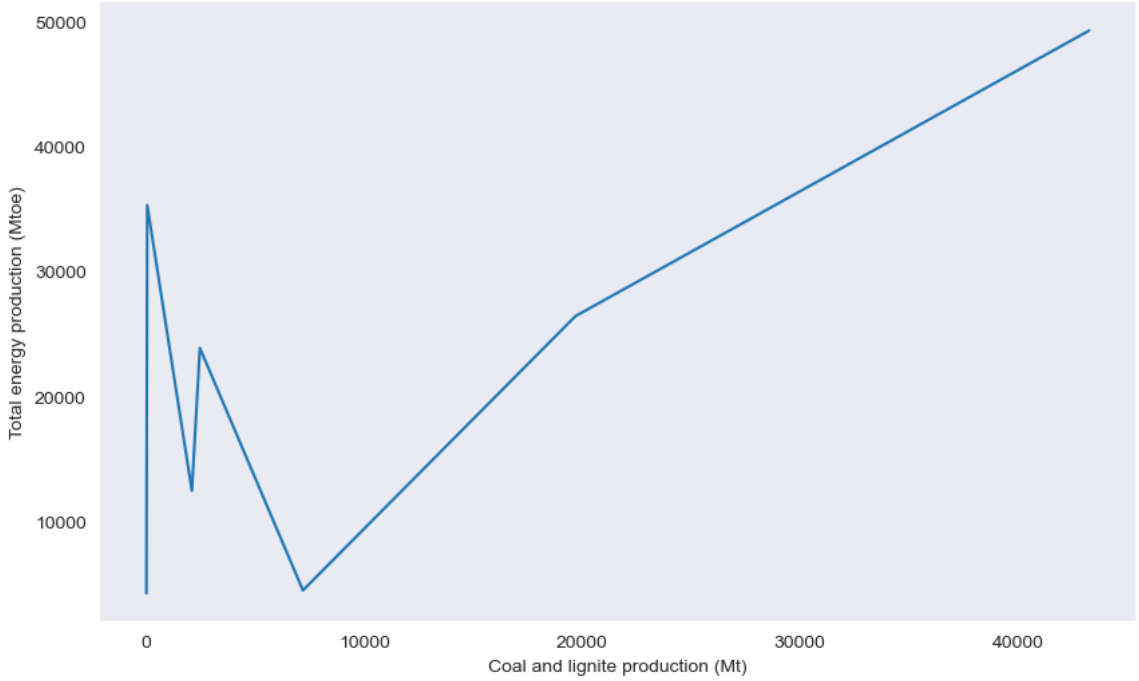


Out[86]:

```
<AxesSubplot:xlabel='Crude oil production (Mt)', ylabel='Total energy prod
uction (Mtoe)'>
```

In [87]:

```
creat_plots('lineplot',data_group,'Coal and lignite production (Mt)','Total energy produc
```



Out[87]:

```
<AxesSubplot:xlabel='Coal and lignite production (Mt)', ylabel='Total ener
gy production (Mtoe)'>
```

In [88]:

```
sns.pairplot(data1,hue='Region')
```

Out[88]:

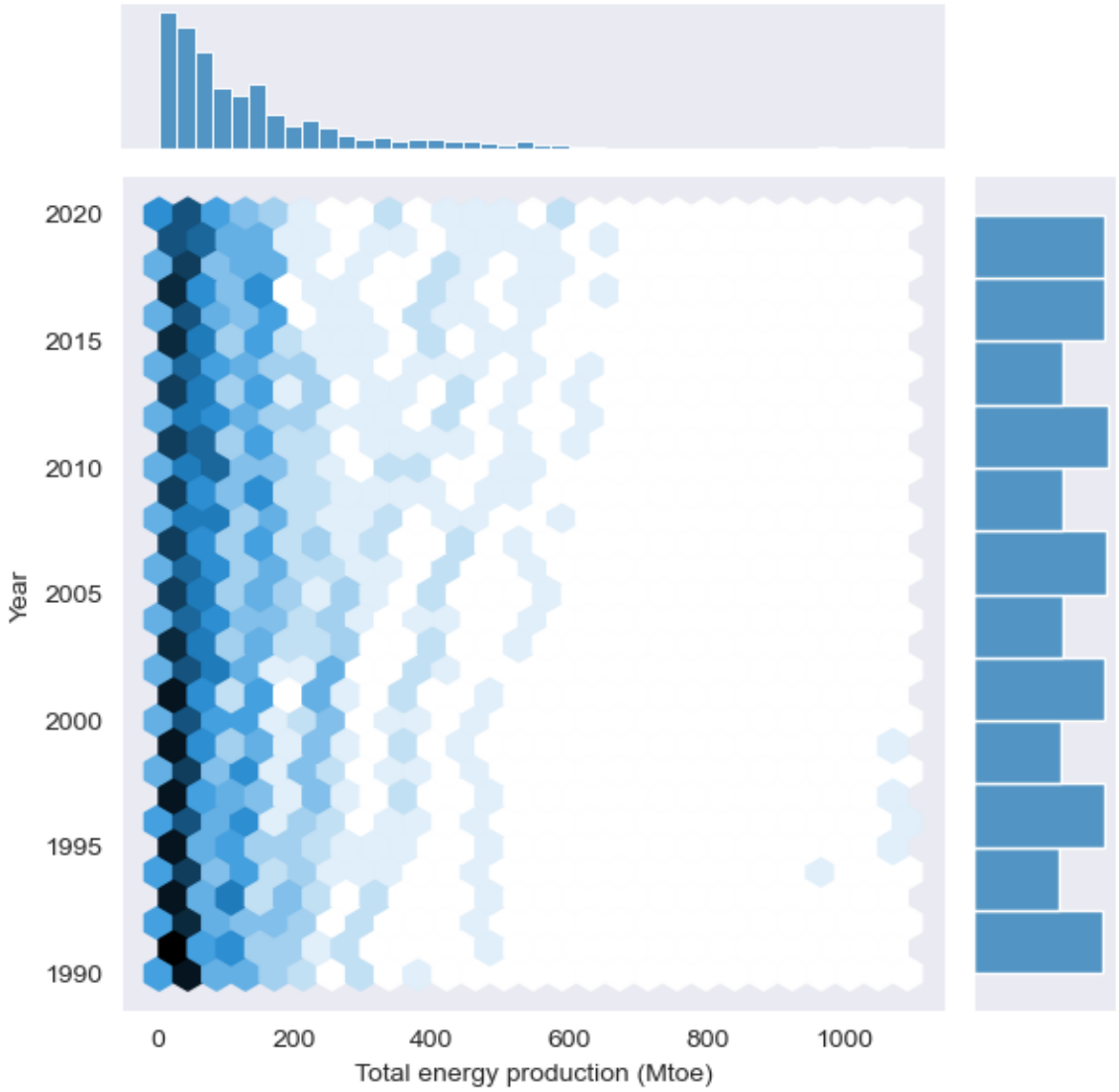
```
<seaborn.axisgrid.PairGrid at 0x29641411220>
```

In [89]:

```
sns.jointplot(x='Total energy production (Mtoe)',y='Year',data=data1,kind='hex')
```

Out[89]:

```
<seaborn.axisgrid.JointGrid at 0x29641045130>
```



In [ ]:

```

```

In [ ]:

```
!pip install pandas-profiling
```

In [ ]:

```
from pandas_profiling import ProfileReport
data_overviwe = ProfileReport(data1,explorative=True,dark_mode=True)
data_overviwe
```