In [1]:
```
!pip install country_converter --upgrade
```

```
Requirement already satisfied: country_converter in c:\users\a\anaconda3\lib\site-packages (1.0.0)
Requirement already satisfied: pandas>=1.0 in c:\users\a\anaconda3\lib\site-packages (from country_c
onverter) (1.4.2)
Requirement already satisfied: python-dateutil>=2.8.1 in c:\users\a\anaconda3\lib\site-packages (fro
m pandas>=1.0->country_converter) (2.8.2)
Requirement already satisfied: numpy>=1.18.5 in c:\users\a\anaconda3\lib\site-packages (from pandas>
=1.0->country_converter) (1.21.5)
Requirement already satisfied: pytz>=2020.1 in c:\users\a\anaconda3\lib\site-packages (from pandas>=
1.0->country_converter) (2021.3)
Requirement already satisfied: six>=1.5 in c:\users\a\anaconda3\lib\site-packages (from python-dateu
til>=2.8.1->pandas>=1.0->country_converter) (1.16.0)
```

In [2]:
```
# Data Analysis

import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt

# Data Visualization

import plotly.express as px
from plotly.subplots import make_subplots
import plotly.graph_objects as go

# World map representation
import country_converter as coco
import datetime
import calendar
```

In [25]:
```
df=pd.read_csv('World Energy Overview.csv')
pf=pd.read_csv('Production_Total.csv')
cf=pd.read_csv('Consumption_Total.csv')
df.head()
```

Out[25]:

	Date	Total Fossil Fuels Production	Nuclear Electric Power Production	Total Renewable Energy Production	Total Primary Energy Production	Primary Energy Imports	Primary Energy Exports	Primary Energy Net Imports	Primary Energy Stock Change and Other	Total Fossil Fuels Consumption	Nuclea Electri Powe Consumptio
0	1973-01-31	4.932632	0.068103	0.403981	5.404715	1.173080	0.125781	1.047299	0.771858	6.747651	0.06810
1	1973-02-28	4.729582	0.064634	0.360900	5.155115	1.168005	0.120883	1.047122	0.390129	6.163095	0.06463
2	1973-03-31	4.946902	0.072494	0.400161	5.419556	1.309473	0.139950	1.169523	-0.067640	6.044647	0.07249
3	1973-04-30	4.716271	0.064070	0.380470	5.160812	1.085169	0.194185	0.890984	-0.110067	5.493184	0.06407
4	1973-05-31	4.956995	0.062111	0.392141	5.411246	1.162804	0.196775	0.966029	-0.305335	5.613551	0.06211

In [26]:
```
df.tail()
```

Out[26]:

	Date	Total Fossil Fuels Production	Nuclear Electric Power Production	Total Renewable Energy Production	Total Primary Energy Production	Primary Energy Imports	Primary Energy Exports	Primary Energy Net Imports	Primary Energy Stock Change and Other	Total Fossil Fuels Consumption	Nuc Ele Pc Consump
594	2022-07-31	6.921426	0.718109	1.132400	8.771935	1.907862	2.294320	-0.386458	0.199327	6.736891	0.71
595	2022-08-31	7.047525	0.718526	1.044026	8.810077	1.821828	2.327090	-0.505262	0.225628	6.755154	0.71
596	2022-09-30	6.915201	0.664673	0.978647	8.558521	1.702048	2.219831	-0.517784	-0.257997	6.144280	0.66
597	2022-10-31	7.126618	0.614741	1.019209	8.760568	1.767733	2.322475	-0.554743	-0.394651	6.179721	0.61
598	2022-11-30	6.875368	0.647029	1.097519	8.619915	1.761697	2.305923	-0.544226	0.210202	6.553431	0.64

In [5]:
```
df.shape
```

Out[5]:
```
(599, 13)
```

In [6]:
```
pf.shape
```

Out[6]:
```
(229, 45)
```

In [7]:
```
cf.shape
```

Out[7]:
```
(230, 45)
```

In [8]:
```
df.columns
```

Out[8]:
```
Index(['Date', 'Total Fossil Fuels Production',
       'Nuclear Electric Power Production',
       'Total Renewable Energy Production', 'Total Primary Energy Production',
       'Primary Energy Imports', 'Primary Energy Exports',
       'Primary Energy Net Imports', 'Primary Energy Stock Change and Other',
       'Total Fossil Fuels Consumption', 'Nuclear Electric Power Consumption',
       'Total Renewable Energy Consumption',
       'Total Primary Energy Consumption'],
      dtype='object')
```

In [9]:
```
df.isnull().sum()
```

Out[9]:
```
Date                                    0
Total Fossil Fuels Production           0
Nuclear Electric Power Production       0
Total Renewable Energy Production       0
Total Primary Energy Production         0
Primary Energy Imports                  0
Primary Energy Exports                  0
Primary Energy Net Imports              0
Primary Energy Stock Change and Other   0
Total Fossil Fuels Consumption          0
Nuclear Electric Power Consumption      0
Total Renewable Energy Consumption      0
Total Primary Energy Consumption        0
dtype: int64
```

In [10]:
```
pf.isnull().sum()
```

Out[10]:
```
Continent           0
Country             0
1980                0
1981                0
1982                0
1983                0
1984                0
1985                0
1986                0
1987                0
1988                0
1989                0
1990                0
1991                0
1992                0
1993                0
1994                0
1995                0
1996                0
1997                0
1998                0
1999                0
2000                0
2001                0
2002                0
2003                0
2004                0
2005                0
2006                0
2007                0
2008                0
2009                0
2010                0
2011                0
2012                0
2013                0
2014                0
2015                0
2016                0
2017                0
2018                0
2019                0
2020                0
2021                0
Total_production    0
dtype: int64
```

In [11]:
```
cf.isnull().sum()
cf.fillna(0, inplace=True)
```

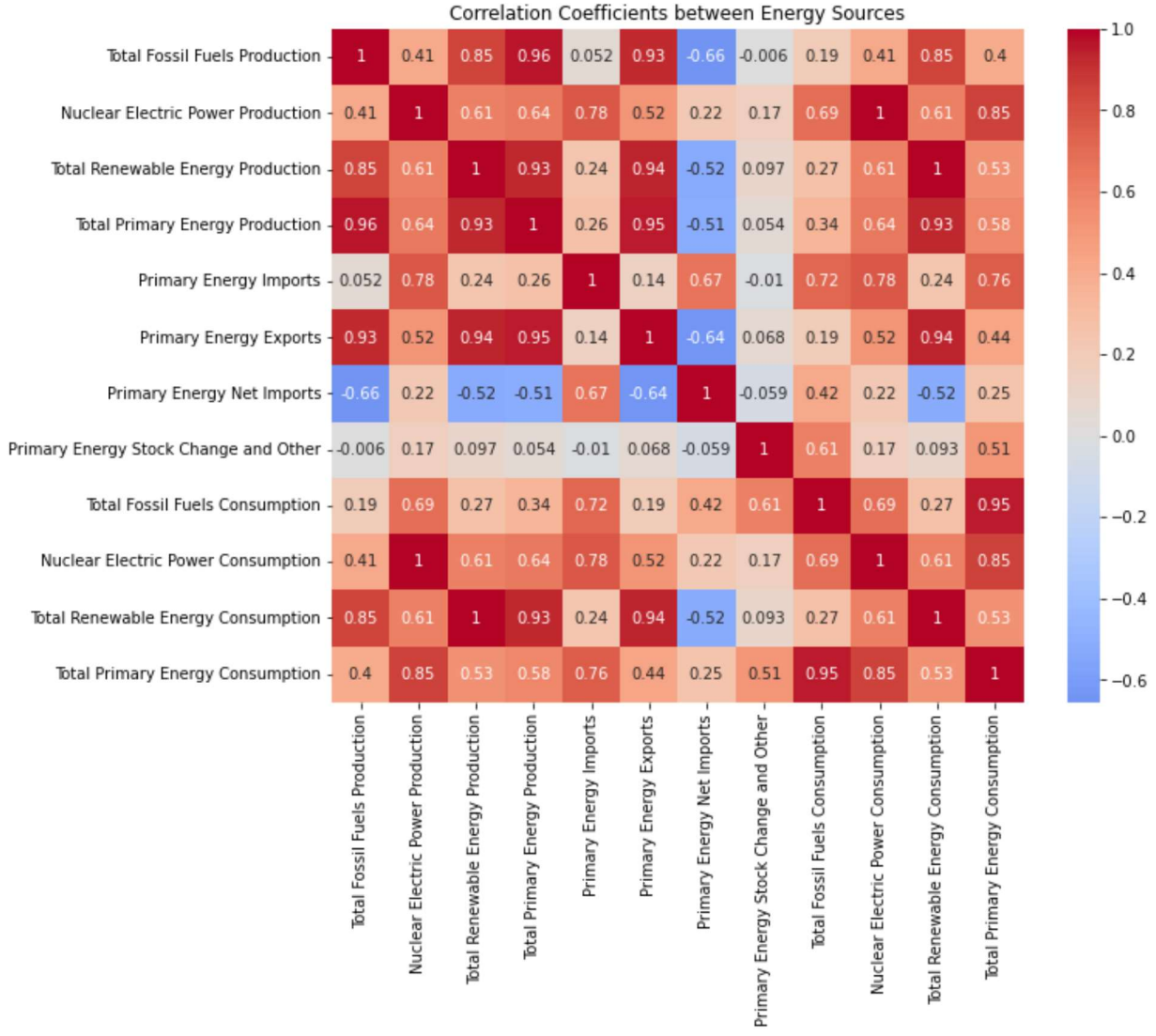
In [12]:
```
print(df.describe())
```

```
       Total Fossil Fuels Production  Nuclear Electric Power Production  \
count                     599.000000                         599.000000
mean                        5.034634                           0.519567
std                         0.610126                           0.202697
min                         3.676065                           0.062111
25%                         4.683559                           0.328635
50%                         4.831601                           0.594293
75%                         5.087384                           0.681056
max                         7.126618                           0.780456

       Total Renewable Energy Production  Total Primary Energy Production  \
count                         599.000000                         599.000000
mean                            0.593709                           6.147909
std                             0.193351                           0.895076
min                             0.304328                           4.306800
25%                             0.467414                           5.590289
50%                             0.527479                           5.906173
75%                             0.685252                           6.290192
max                             1.218790                           8.810077

       Primary Energy Imports  Primary Energy Exports  \
count              599.000000              599.000000
mean                 1.873459                0.611704
std                  0.561286                0.544532
min                  0.710558                0.056798
25%                  1.447091                0.310775
50%                  1.848642                0.373346
75%                  2.281205                0.680119
max                  3.149640                2.386337

       Primary Energy Net Imports  Primary Energy Stock Change and Other  \
count                  599.000000                             599.000000
mean                     1.261755                               0.031835
std                      0.724907                               0.476652
min                     -0.554743                              -0.894627
25%                      0.856081                              -0.327821
50%                      1.200957                              -0.081028
75%                      1.759529                               0.324711
max                      2.741692                               1.551345

       Total Fossil Fuels Consumption  Nuclear Electric Power Consumption  \
count                      599.000000                         599.000000
mean                         6.321391                           0.519567
std                          0.708356                           0.202697
min                          4.783910                           0.062111
25%                          5.798086                           0.328635
50%                          6.338503                           0.594293
75%                          6.789340                           0.681056
max                          8.096323                           0.780456

       Total Renewable Energy Consumption  Total Primary Energy Consumption
count                          599.000000                         599.000000
mean                             0.591620                           7.441499
std                              0.189229                           0.946882
min                              0.304328                           5.435627
25%                              0.467414                           6.659050
50%                              0.527257                           7.617372
75%                              0.683567                           8.111785
max                              1.199383                           9.664299
```

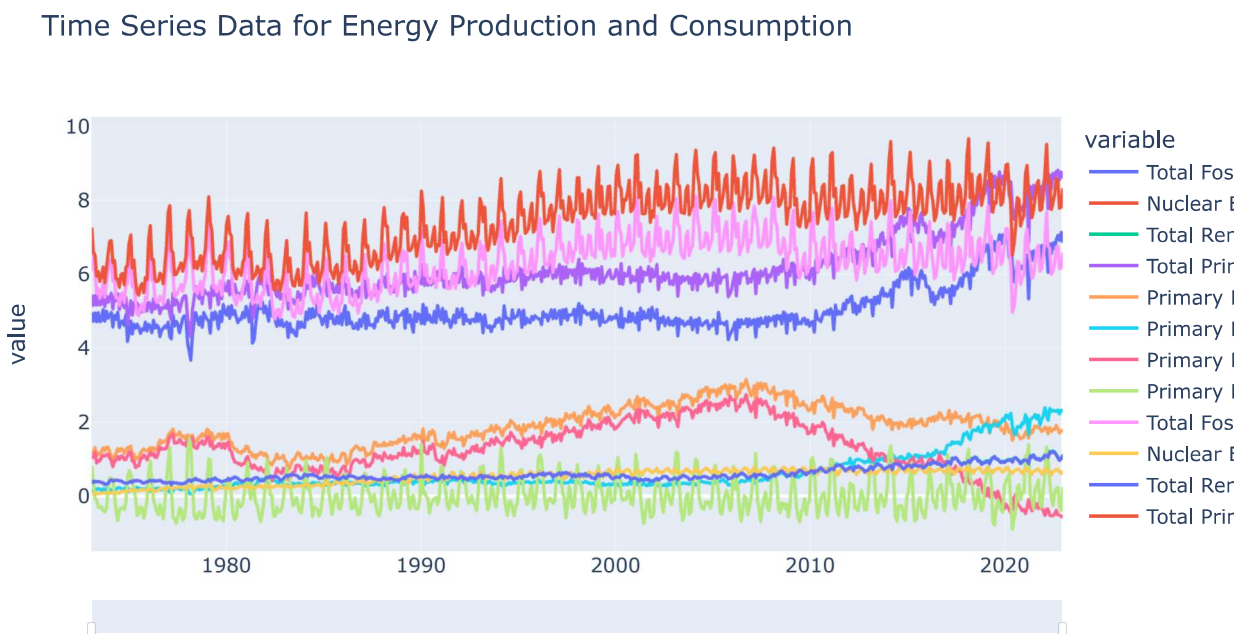
In [13]:
```
# Calculating the correlation matrix
corr_matrix = df.corr()
print(corr_matrix)
```

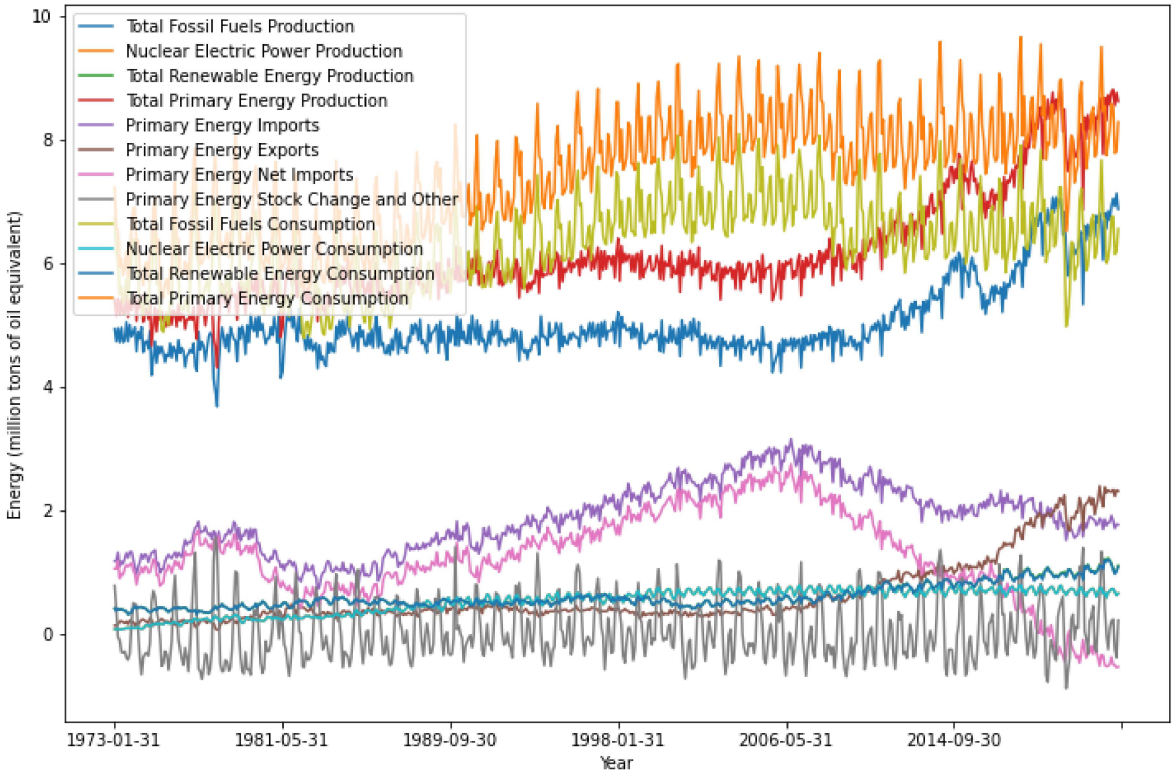
```
                                       Total Fossil Fuels Production  \
Total Fossil Fuels Production                               1.000000
Nuclear Electric Power Production                           0.411684
Total Renewable Energy Production                           0.851062
Total Primary Energy Production                             0.958719
Primary Energy Imports                                      0.051540
Primary Energy Exports                                      0.925480
Primary Energy Net Imports                                 -0.655291
Primary Energy Stock Change and Other                      -0.005999
Total Fossil Fuels Consumption                              0.189429
Nuclear Electric Power Consumption                          0.411684
Total Renewable Energy Consumption                          0.847192
Total Primary Energy Consumption                            0.401573

                                       Nuclear Electric Power Production  \
Total Fossil Fuels Production                                    0.411684
Nuclear Electric Power Production                                1.000000
Total Renewable Energy Production                                0.609987
Total Primary Energy Production                                  0.638849
```

In [14]:
```
# Calculate the correlation coefficients between the energy sources and other factors
corr_cols = ['Total Fossil Fuels Production', 'Nuclear Electric Power Production', 'Total Renewable E
             'Total Primary Energy Production', 'Primary Energy Imports', 'Primary Energy Exports',
             'Primary Energy Net Imports', 'Primary Energy Stock Change and Other', 'Total Fossil Fue
             'Nuclear Electric Power Consumption', 'Total Renewable Energy Consumption', 'Total Prima
corr_df = df[corr_cols].corr()

# Set the size of the plot
plt.figure(figsize=(10, 8))

# Plot the data using a heatmap
sns.heatmap(corr_df, annot=True, cmap='coolwarm', center=0)

# Set plot title
plt.title('Correlation Coefficients between Energy Sources')

# Show the plot
plt.show()
```

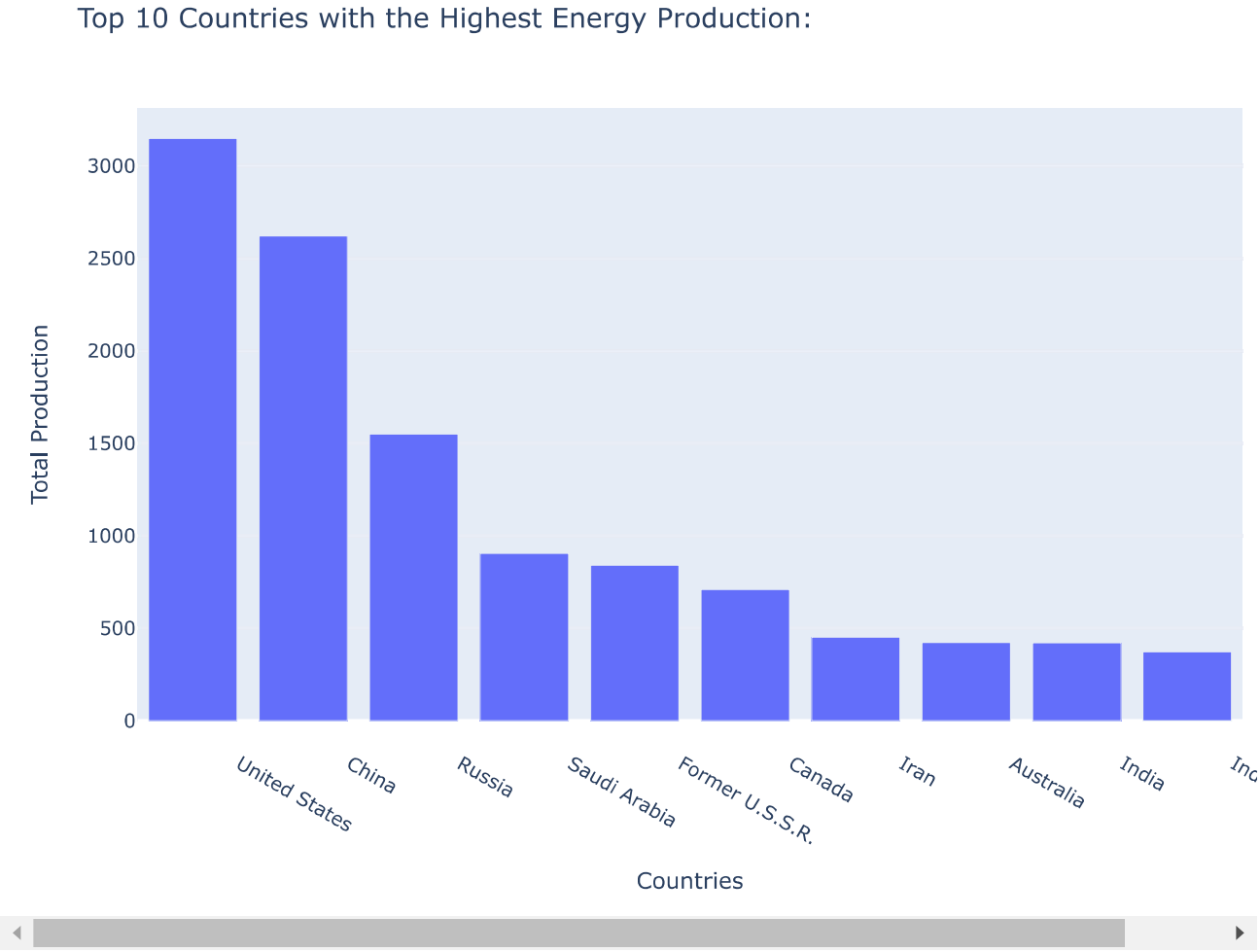


In [15]:
```
# Setting the date column as the index
df.set_index('Date', inplace=True)

# Ploting the data
df[df.columns].plot(figsize=(12,8))

plt.xlabel('Year')
plt.ylabel('Energy (million tons of oil equivalent)')

fig = px.line(df, x=df.index, y=df.columns, title='Time Series Data for Energy Production and Consumpt
fig.update_xaxes(rangeslider_visible=True)


# displays the plot
fig.show()
```



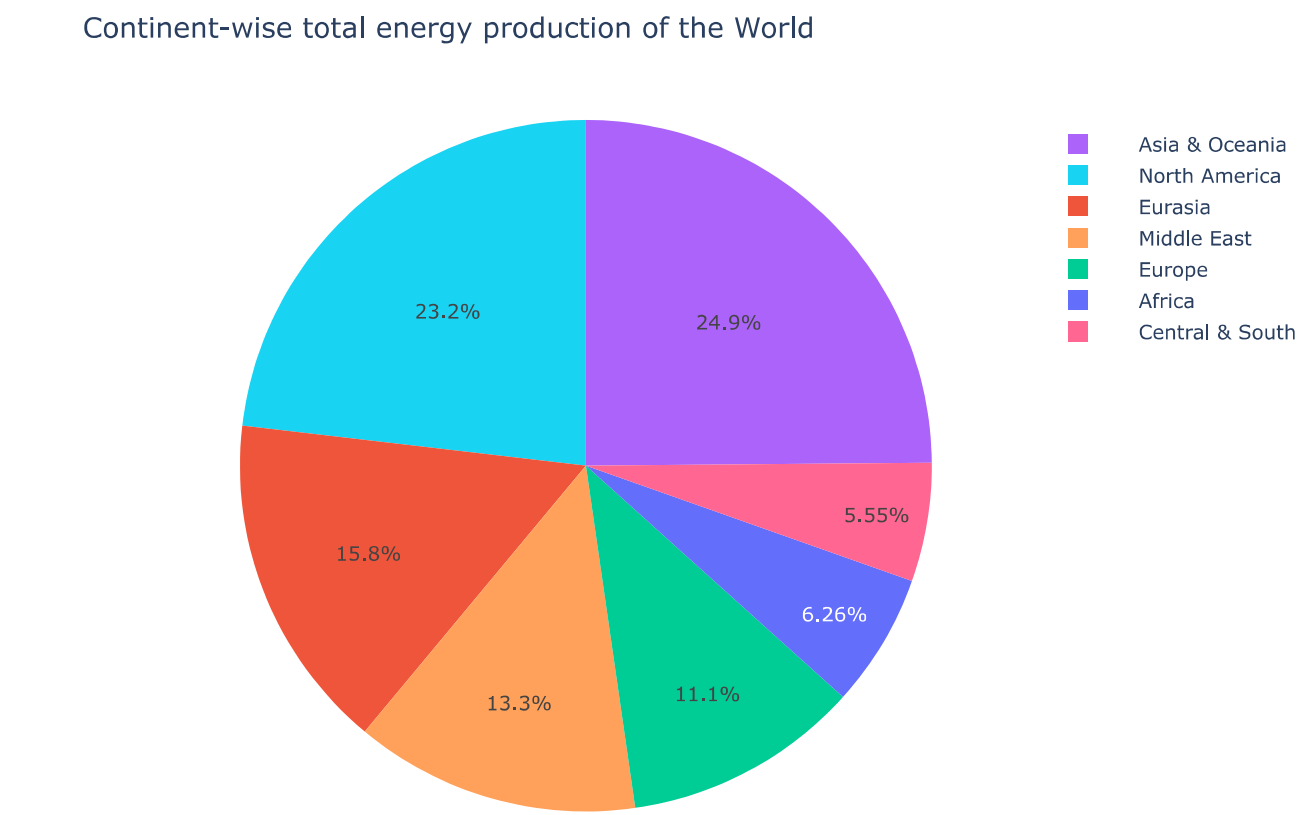

Total Fossil Fuels Production
Nuclear Electric Power Production
Total Renewable Energy Production
Total Primary Energy Production
Primary Energy Imports
Primary Energy Exports
Primary Energy Net Imports
Primary Energy Stock Change and Other
Total Fossil Fuels Consumption
Nuclear Electric Power Consumption
Total Renewable Energy Consumption
Total Primary Energy Consumption
Energy (million tons of oil equivalent)
10
8
6
4
2
0
1973-01-31
1981-05-31
1989-09-30
1998-01-31
2006-05-31
2014-09-30
Year

In [16]:
```
top_pf = pf[['Country', 'Total_production']].sort_values(by=['Total_production'], ascending=False).he

fig = px.bar(top_pf, x='Country', y='Total_production',labels={
                     "Country": "Countries",
                     "Total_production": "Total Production"
                 },title='Top 10 Countries with the Highest Energy Production:', width=850, height=60
fig.show()
```



In [17]:
```
continent_pf =  pf.groupby(by='Continent')['Total_production'].sum()
print(continent_pf)

fig = px.pie(pf, values='Total_production', names='Continent',color='Continent',
             title='Continent-wise total energy production of the World',
             width=850, height=600)
fig.show()
```

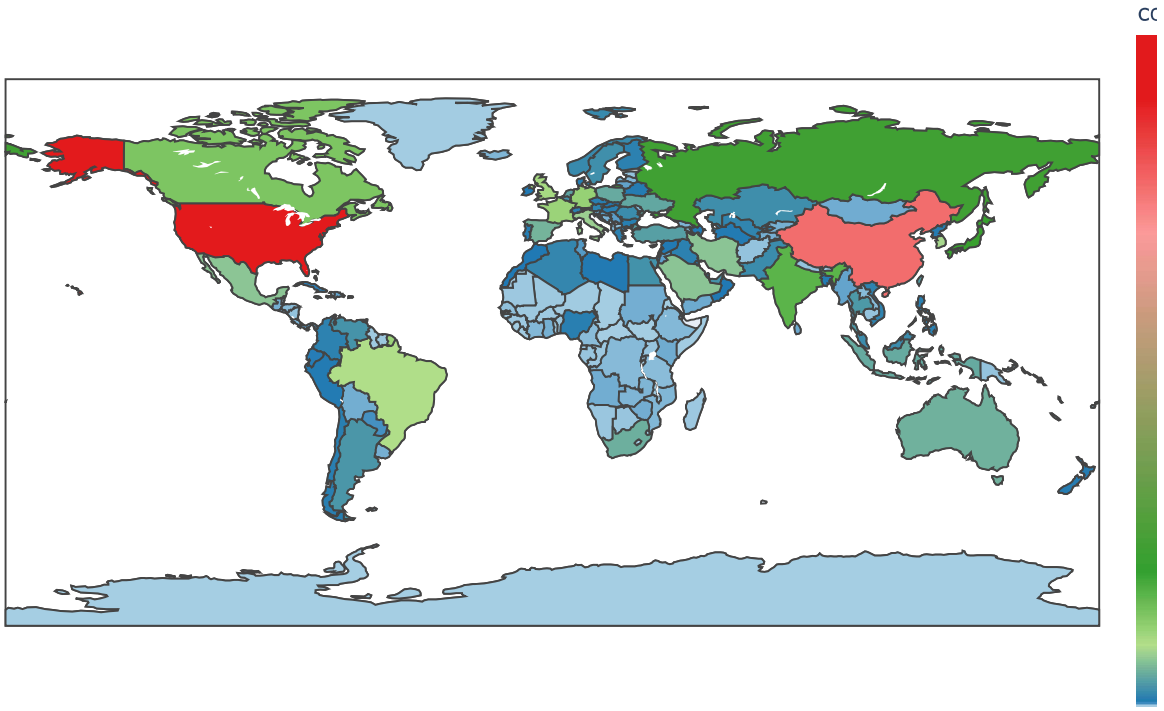
```
Continent
    Africa                      1136.607828
    Asia & Oceania              4513.091785
    Central & South America     1007.329117
    Eurasia                     2871.262313
    Europe                      2006.767697
    Middle East                 2416.979487
    North America               4203.165946
Name: Total_production, dtype: float64
```



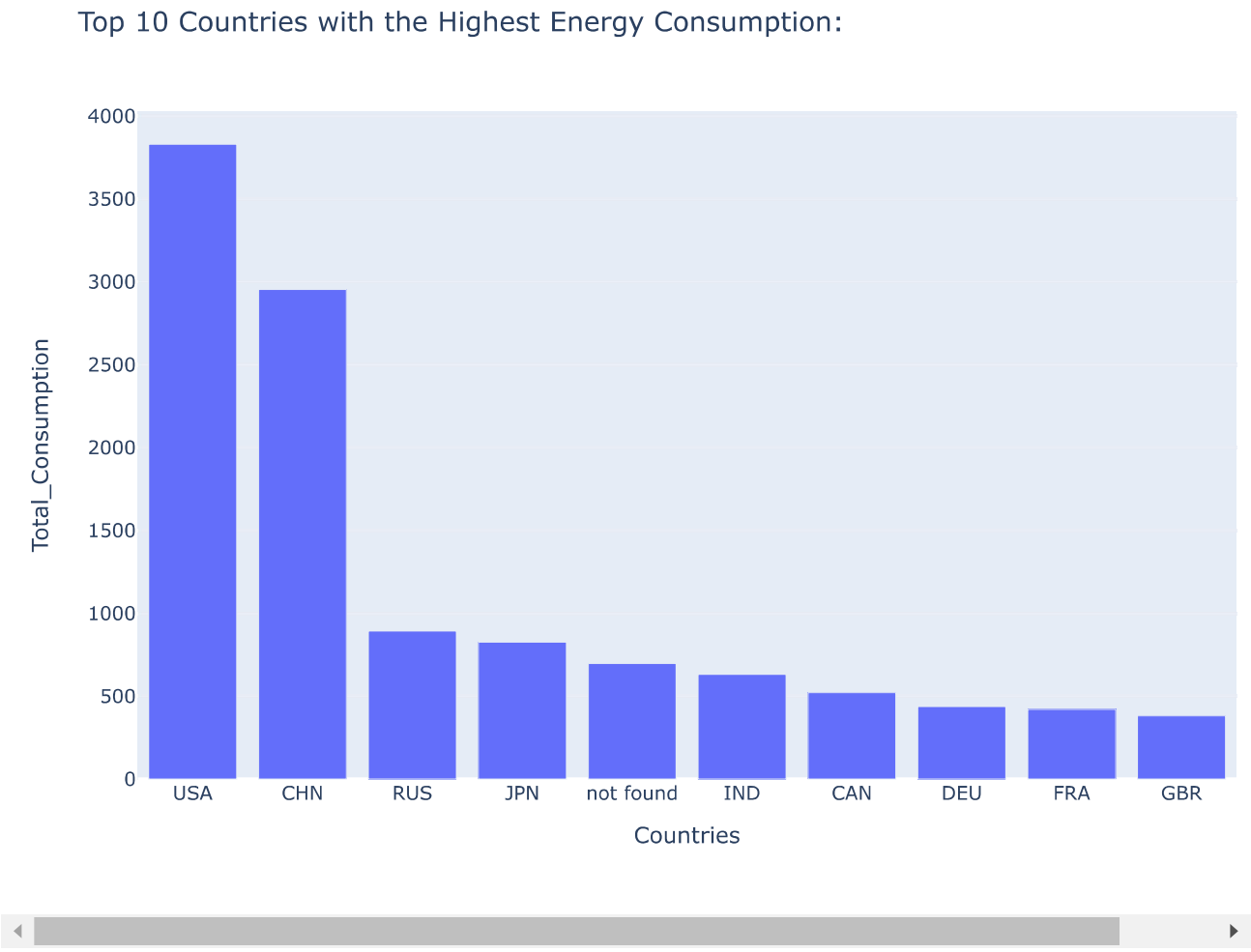
In [18]:
```
country = coco.convert(names = cf['Country'], to = "ISO3")
cf['Country'] = country

fig = px.choropleth(locations = cf.Country,
                    color = cf.Total_Consumption,
                    color_continuous_scale = [[0, 'rgb(166,206,227, 0.5)'],
                      [0.005, 'rgb(31,120,180,0.5)'],
                      [0.09, 'rgb(178,223,138,0.5)'],
                      [0.2, 'rgb(51,160,44,0.5)'],
                      [0.7, 'rgb(251,154,153,0.5)'],
                      [0.9, 'rgb(227,26,28,0.5)'],
                      [1, 'rgb(227,26,28,0.5)']
                     ],
                    title = 'Global Energy Consumption')
fig.update_layout(height=630, width=850)
fig.show()
```

```
Hawaiian Trade Zone not found in regex
Samoa not found in regex
U.S. Pacific Islands not found in regex
Wake Island not found in regex
Netherlands Antilles not found in regex
U.S. Territories not found in regex
Former Serbia and Montenegro not found in regex
Former Yugoslavia not found in regex
Germany, East not found in regex
Germany, West not found in regex
Former U.S.S.R. not found in regex
```

Global Energy Consumption

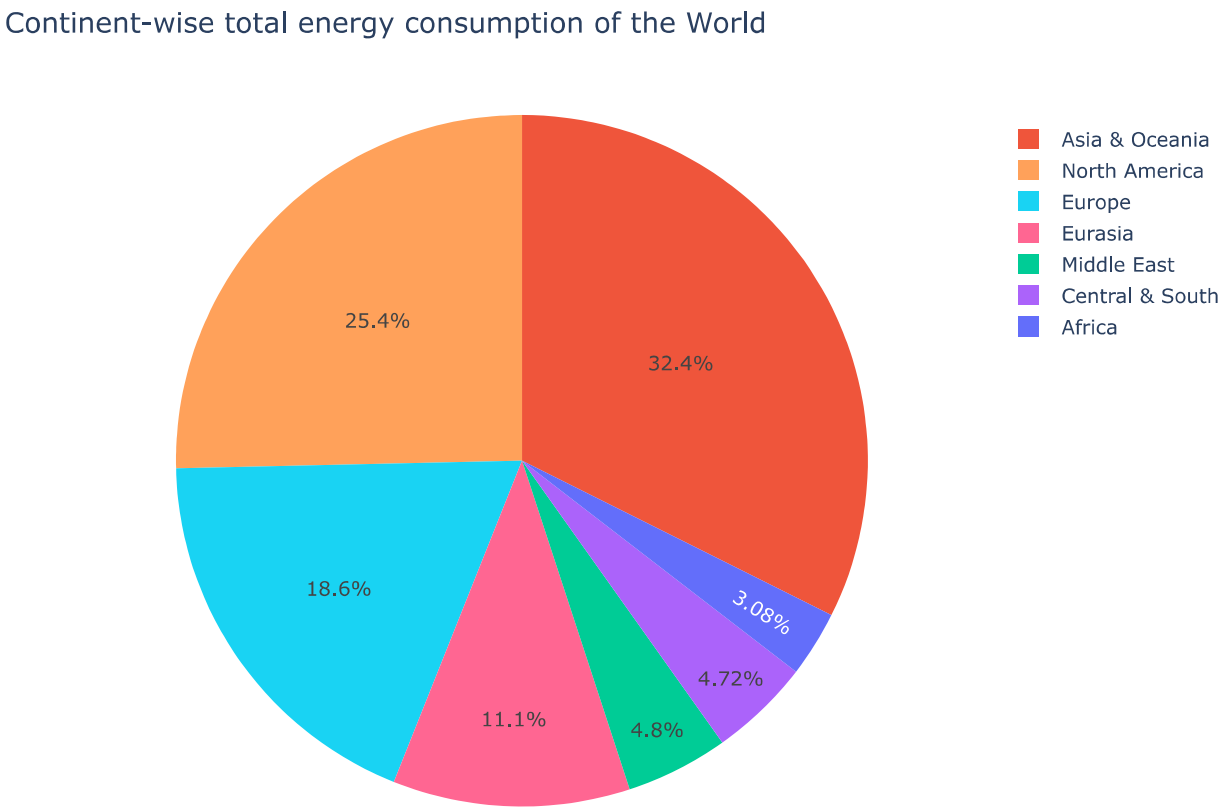


In [19]:
```
top_consumption = cf[['Country', 'Total_Consumption']].sort_values(by=['Total_Consumption'], ascendin

fig = px.bar(top_consumption, x='Country', y='Total_Consumption',labels={
                     "Country": "Countries",
                     "Total_production": "Total Production"
                 },title='Top 10 Countries with the Highest Energy Consumption:', width=850, height=6
fig.show()
```



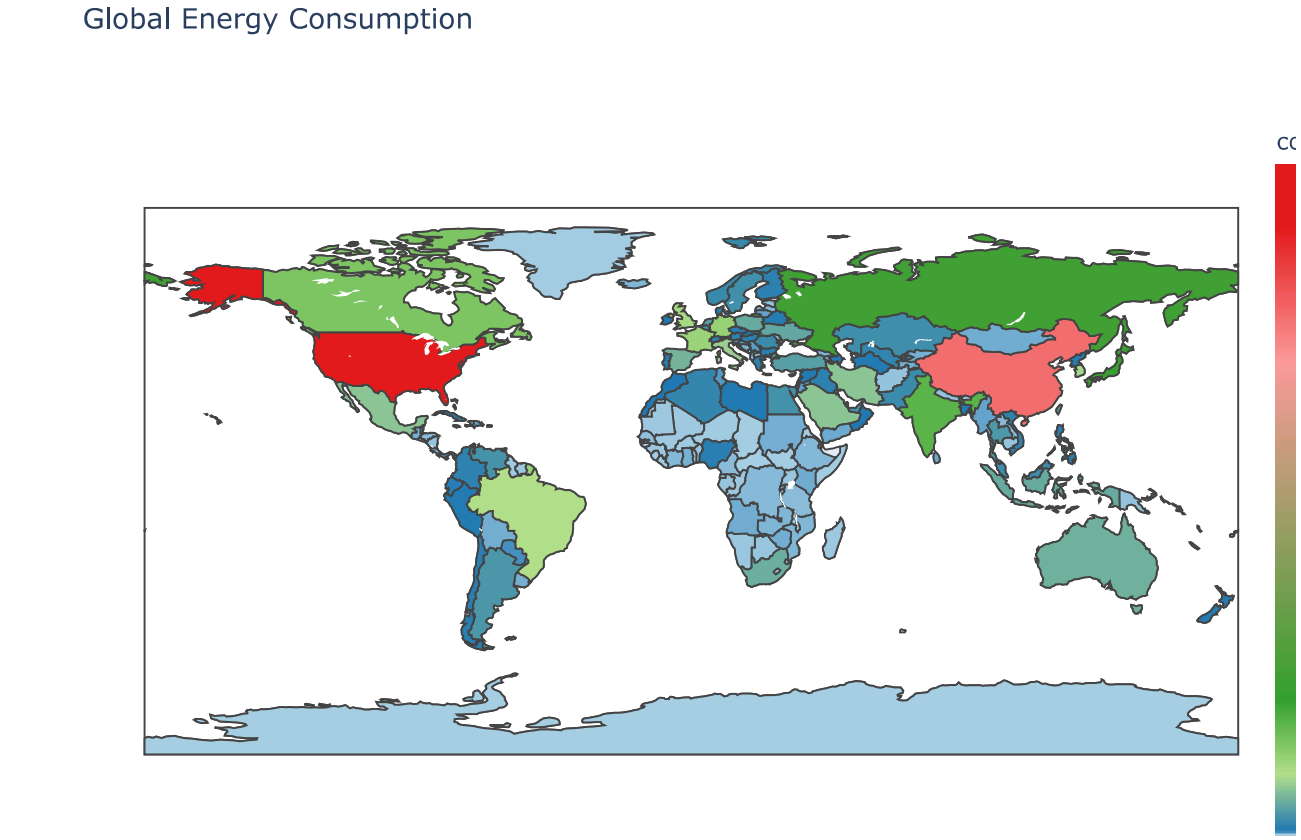
In [20]:
```
continent_consumption =  cf.groupby(by='Continent')['Total_Consumption'].sum()
print(continent_consumption)

fig = px.pie(cf, values='Total_Consumption', names='Continent',color='Continent',
             title='Continent-wise total energy consumption of the World',
             width=850, height=600)
fig.show()
```

```
Continent
Africa                        559.693585
Asia & Oceania               5890.712088
Central & South America       858.645159
Eurasia                      2015.012531
Europe                       3387.844503
Middle East                   873.845333
North America                4614.115501
Name: Total_Consumption, dtype: float64
```

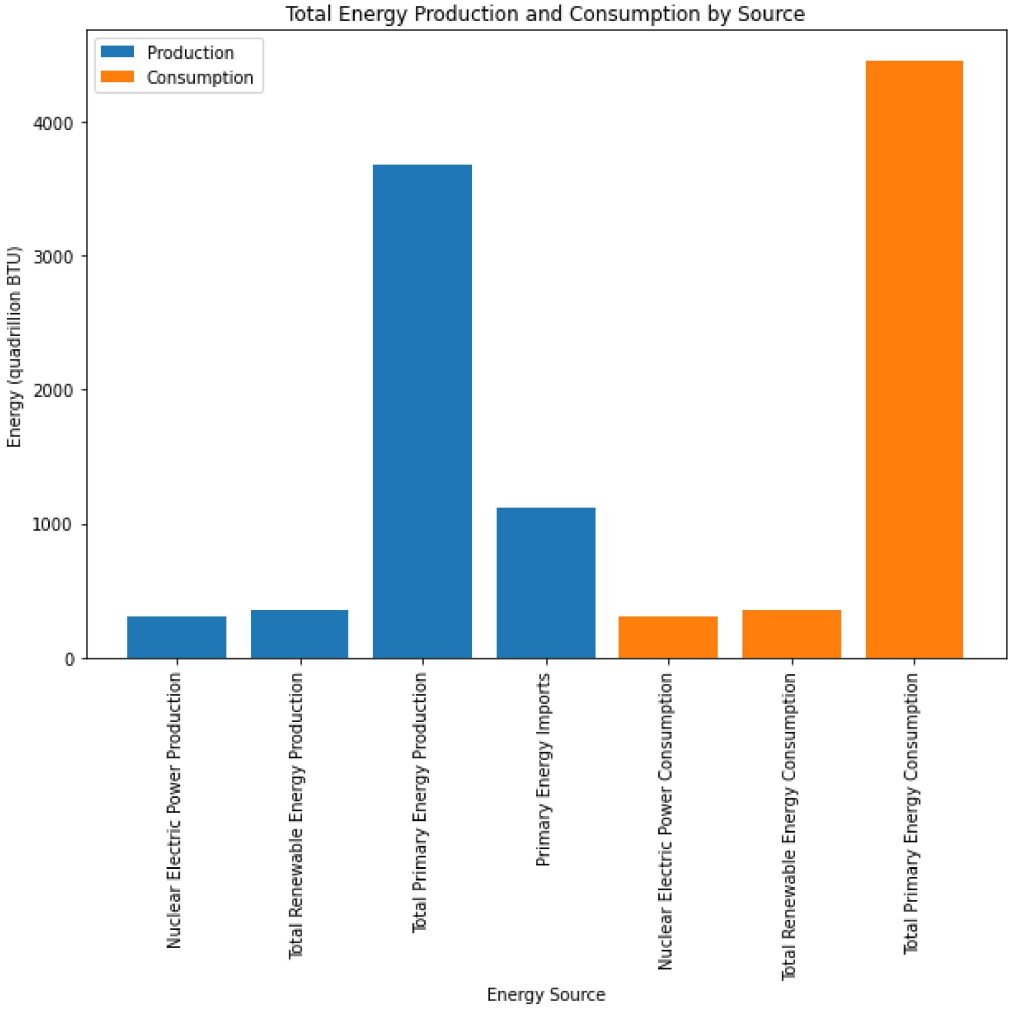


In [21]:
```
country = coco.convert(names = cf['Country'], to = "ISO3")
cf['Country'] = country

fig = px.choropleth(locations = cf.Country,
                    color = cf.Total_Consumption,
                    color_continuous_scale = [[0, 'rgb(166,206,227, 0.5)'],
                      [0.005, 'rgb(31,120,180,0.5)'],
                      [0.09, 'rgb(178,223,138,0.5)'],
                      [0.2, 'rgb(51,160,44,0.5)'],
                      [0.7, 'rgb(251,154,153,0.5)'],
                      [0.9, 'rgb(227,26,28,0.5)'],
                      [1, 'rgb(227,26,28,0.5)']
                     ],
                    title = 'Global Energy Consumption')
fig.update_layout(height=630, width=850)
fig.show()
```

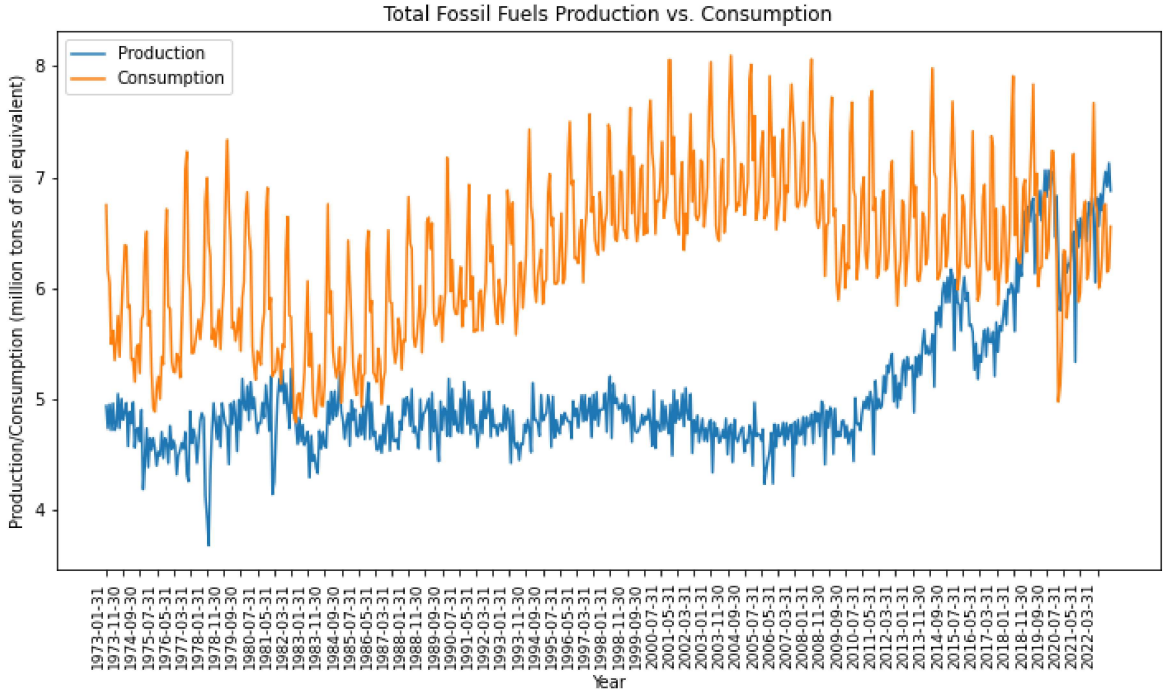
```
not found not found in regex
not found not found in regex
not found not found in regex
not found not found in regex
not found not found in regex
not found not found in regex
not found not found in regex
not found not found in regex
not found not found in regex
not found not found in regex
not found not found in regex
```



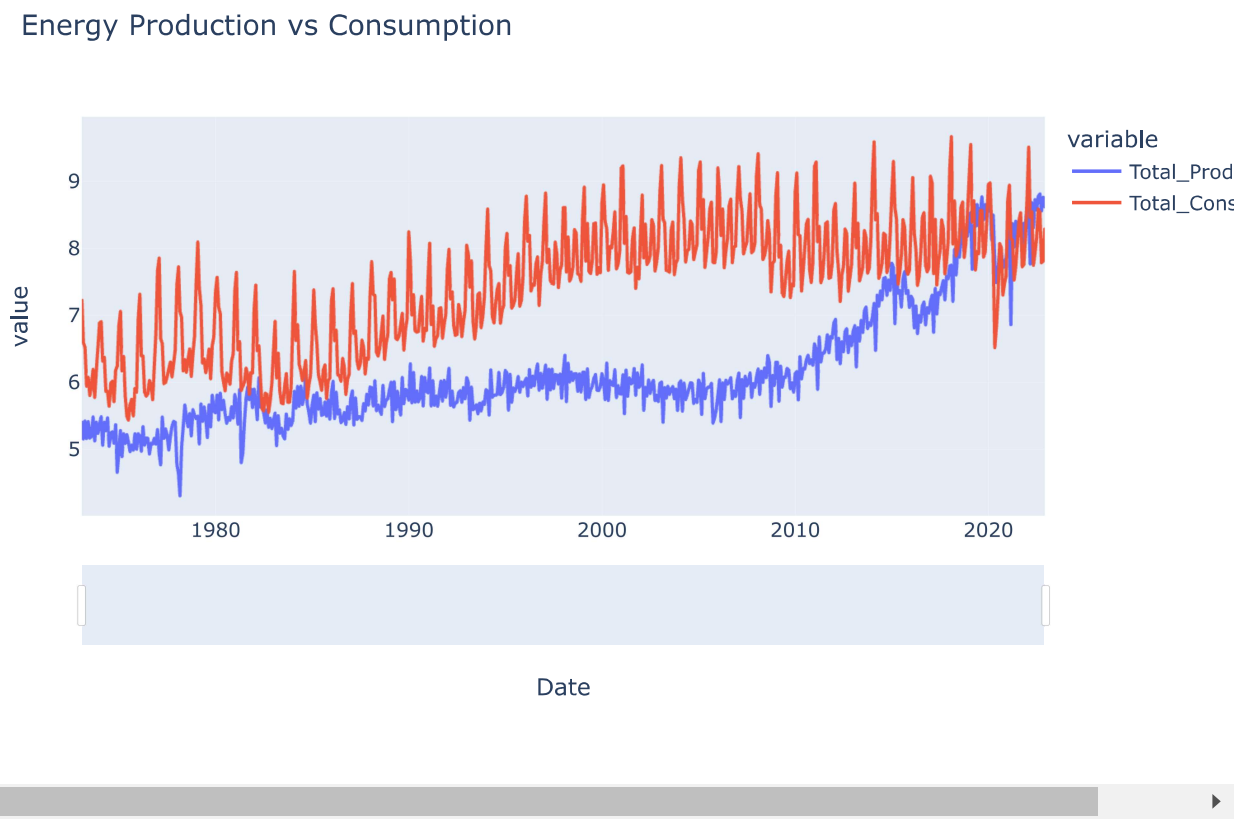
In [22]:
```
# Calculating total production and consumption for each energy source
total_pf = df.iloc[:, 1:5].sum()
total_cf = df.iloc[:, 9:13].sum()

# Creating a bar chart of the total production and consumption for each energy source
fig, ax = plt.subplots(figsize=(10, 7))
ax.bar(total_pf.index, total_pf.values, label='Production')
ax.bar(total_cf.index, total_cf.values, label='Consumption')
plt.xticks(rotation=90, fontsize=10)
plt.yticks(fontsize=10)
plt.title('Total Energy Production and Consumption by Source', fontsize=12)
plt.xlabel('Energy Source', fontsize=10)
plt.ylabel('Energy (quadrillion BTU)', fontsize=10)
plt.legend(fontsize=10)
plt.show()
```

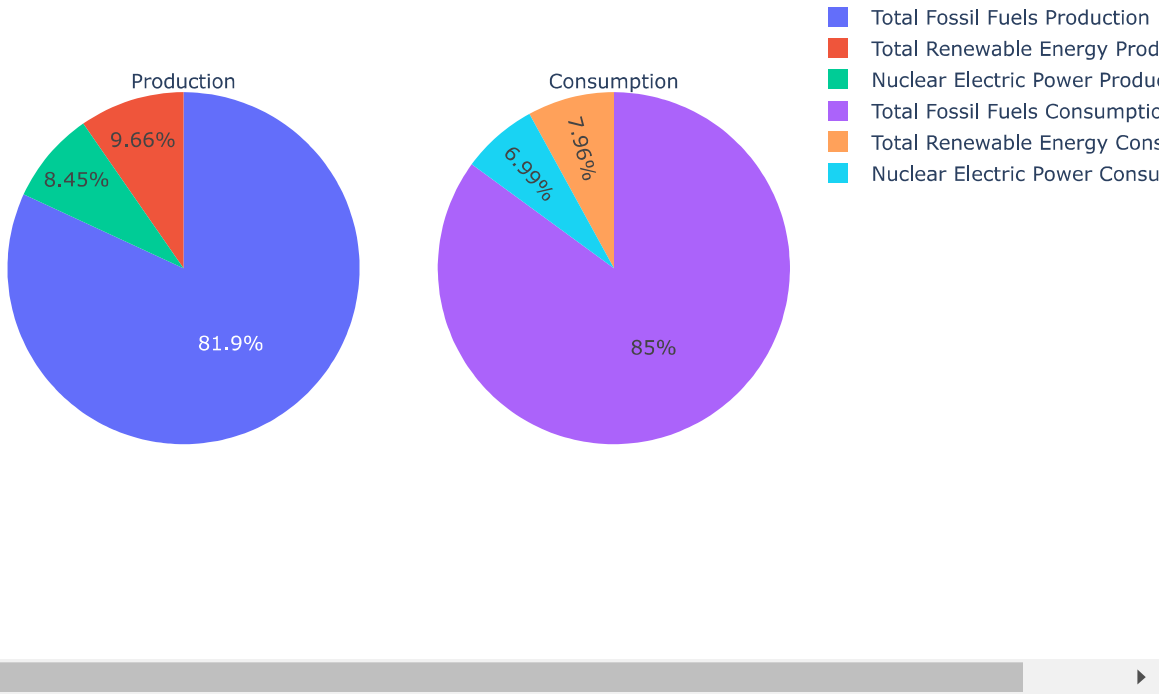


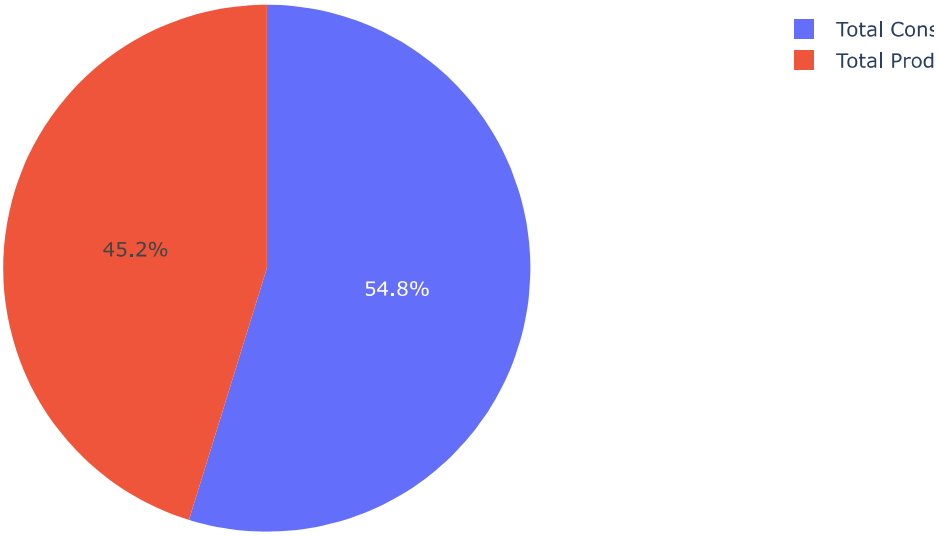
In [23]:
```
# Plotting the Total Fossil Fuels Production vs. Consumption
plt.figure(figsize=(12,6))
plt.plot(df['Total Fossil Fuels Production'], label='Production')
plt.plot(df['Total Fossil Fuels Consumption'], label='Consumption')
plt.title('Total Fossil Fuels Production vs. Consumption')
plt.xlabel('Year')
plt.xticks(df.index[::10], rotation=90, ha='right', fontsize=9)
plt.ylabel('Production/Consumption (million tons of oil equivalent)')
plt.legend()
plt.show()
```



In [27]:
```
datax1 = df.filter(['Total Primary Energy Production','Total Primary Energy Consumption', 'Date'], ax
datax2 = datax1.rename(columns={'Total Primary Energy Production':'Total_Production',
                                'Total Primary Energy Consumption':'Total_Consum
datax2 = datax2.set_index('Date')
datax2.head()

fig = px.line(datax2, x=datax2.index, y=datax2.columns,
              title='Energy Production vs Consumption',width=850, height=500)
fig.update_xaxes(rangeslider_visible=True)
fig.show()

fig = make_subplots(rows=1, cols=2, specs=[[{"type": "pie"}, {"type": "pie"}]])

fig.add_trace(go.Pie(
    values= [3015.745538, 355.6315180000005, 311.220662],
    labels=['Total Fossil Fuels Production', 'Total Renewable Energy Production', 'Nuclear Electric P
    domain=dict(x=[0, 0.5]),
    title='Production'),
    row=1, col=1)

fig.add_trace(go.Pie(
    values= [3786.513071, 354.38061200000004, 311.220662],
    labels=['Total Fossil Fuels Consumption', 'Total Renewable Energy Consumption', 'Nuclear Electric
    domain=dict(x=[0, 0.5]),
    title='Consumption'),
    row=1, col=2)

fig.update_layout(height=500, width=850)
fig.show()


labels = ['Total Production','Total Consumption']
values = [3682.597704, 4457.458121]

fig = go.Figure(data=[go.Pie(labels=labels, values=values)])
fig.update_layout(width=850, height=500)
fig.show()
```




Production
Consumption
9.66%
8.45%
81.9%
7.96%
6.99%
85%
Total Fossil Fuels Production
Total Renewable Energy Prod
Nuclear Electric Power Produ
Total Fossil Fuels Consumptic
Total Renewable Energy Cons
Nuclear Electric Power Consu


45.2%
54.8%
Total Cons
Total Prod

In [ ]: