```
import numpy as np
import pandas as pd
import matplotlib as mpl
from matplotlib import pyplot as plt
import seaborn as sns
```

```
df1=pd.read_csv('Social Indicators.csv')
df2=pd.read_csv('Statistical indicators.csv')
df3=pd.read_csv('Environmental indicators..csv')
df4=pd.read_csv('Economic Indicators.csv')
df5=pd.read_csv('QPSD indicators.csv')
```

```
columns_to_sum = ["2021Q1 [YR2021Q1]", "2021Q2 [YR2021Q2]", "2021Q3 [YR2021Q3]", "2021Q4 [YR2021Q4]"]
df5["2021 [YR2021]"] = df5[columns_to_sum].sum(axis=1)
df5.drop(columns=columns_to_sum, inplace=True)

columns_to_sum2 = ["2022Q1 [YR2022Q1]", "2022Q2 [YR2022Q2]", "2022Q3 [YR2022Q3]", "2022Q4 [YR2022Q4]"]
df5["2022 [YR2022]"] = df5[columns_to_sum2].sum(axis=1)
df5.drop(columns=columns_to_sum2, inplace=True)

columns_to_sum3 = ["2023Q1 [YR2023Q1]", "2023Q2 [YR2023Q2]", "2023Q3 [YR2023Q3]", "2023Q4 [YR2023Q4]"]
df5["2023 [YR2023]"] = df5[columns_to_sum3].sum(axis=1)
df5.drop(columns=columns_to_sum3, inplace=True)
```

```
rename_dict = {"Series_type": "series_type"}
df4 = df4.rename(columns=rename_dict)
df5 = df5.rename(columns=rename_dict)
```

```
df = pd.concat([df1, df2, df3, df4, df5], ignore_index=True)
```

```
df = pd.get_dummies(df, columns=['Country Code'], dtype=int)
```

```
df = df[df['2021 [YR2021]'].str.contains(r'\d', na=False)]
df = df[df['2022 [YR2022]'].str.contains(r'\d', na=False)]
df = df[df['2023 [YR2023]'].str.contains(r'\d', na=False)]
```

```
df.to_csv('modified_file.csv', index=False)
```