Homework 2

Nick Climaco

February 9, 2024

Table of contents

Chapter 3 Exercises (Hyndman and Athanasopoulos)

```
<Figure size 2700x1800 with 0 Axes>
```

```
<Figure size 2700x1800 with 0 Axes>
```

Exercise 1

Consider the GDP information in global_economy.

Plot the GDP per capita for each country over time. Which country has the highest GDP per capita?

```
# read in the data from R generated csv file
df_global_economy = pd.read_csv("data/global_economy.csv",
                                parse_dates=["Year"],
                                index_col=["Year"])

df_global_economy["GDP_Per_Capita"] = (df_global_economy["GDP"] /
df_global_economy["Population"]).round(2)

# calculate gdp per capita
df_gdp_per_capita = df_global_economy[["Country", "GDP_Per_Capita","GDP", "Population"]]

#group by country
df_grouped = df_gdp_per_capita.groupby("Country")
```

Which country has the highest GDP per capita?

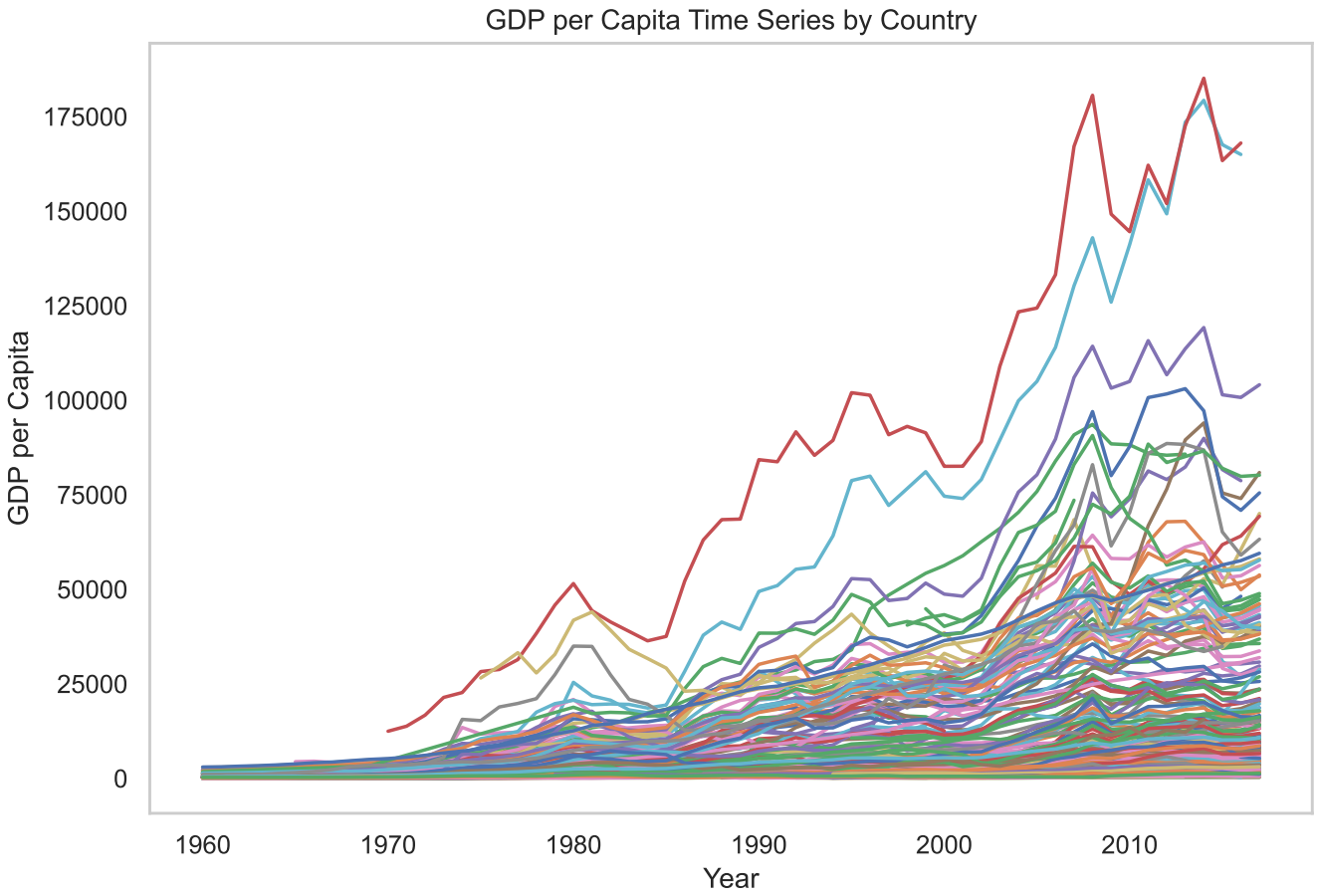
```
# sort vales of GDP per Capita by Country
df_gdp_per_capita.groupby("Country")[["GDP_Per_Capita", "GDP", "Population"]]\
  .max()\
    .reset_index()\
      .sort_values("GDP_Per_Capita", ascending=False)\
        .head(10)
```

	Country	GDP_Per_Capita	GDP	Population
163	Monaco	185152.53	7.060236e+09	38695.0
139	Liechtenstein	179308.08	6.657171e+09	37922.0
144	Luxembourg	119225.38	6.632734e+10	599449.0
180	Norway	103059.25	5.235021e+11	5282223.0
145	Macao SAR, China	94004.39	5.534800e+10	622567.0
22	Bermuda	93605.75	6.109928e+09	65636.0
202	San Marino	90682.58	2.752307e+09	33400.0
114	Isle of Man	89941.64	7.428280e+09	84287.0
197	Qatar	88564.82	2.062247e+11	2639211.0
232	Switzerland	88415.63	7.091826e+11	8466017.0

The country with the highest GDP per Capita is Monaca. It is a city-state that borders the Mediterranean Sea. Moreover, looking at the table we can infer that rich countries with very small population have higher GDP per Capita which makes sense.

How has this changed over time?

```
# plot the time series
plt.figure(figsize=(9,6))

sns.set_style("whitegrid")

# create lines for each country
for country, value in df_grouped:
    sns.lineplot(x=value.index,
                 y = "GDP_Per_Capita",
                 data=value, label = None)

# add labels
plt.title('GDP per Capita Time Series by Country')
plt.xlabel('Year')
plt.ylabel('GDP per Capita')
plt.grid()
plt.show()
```



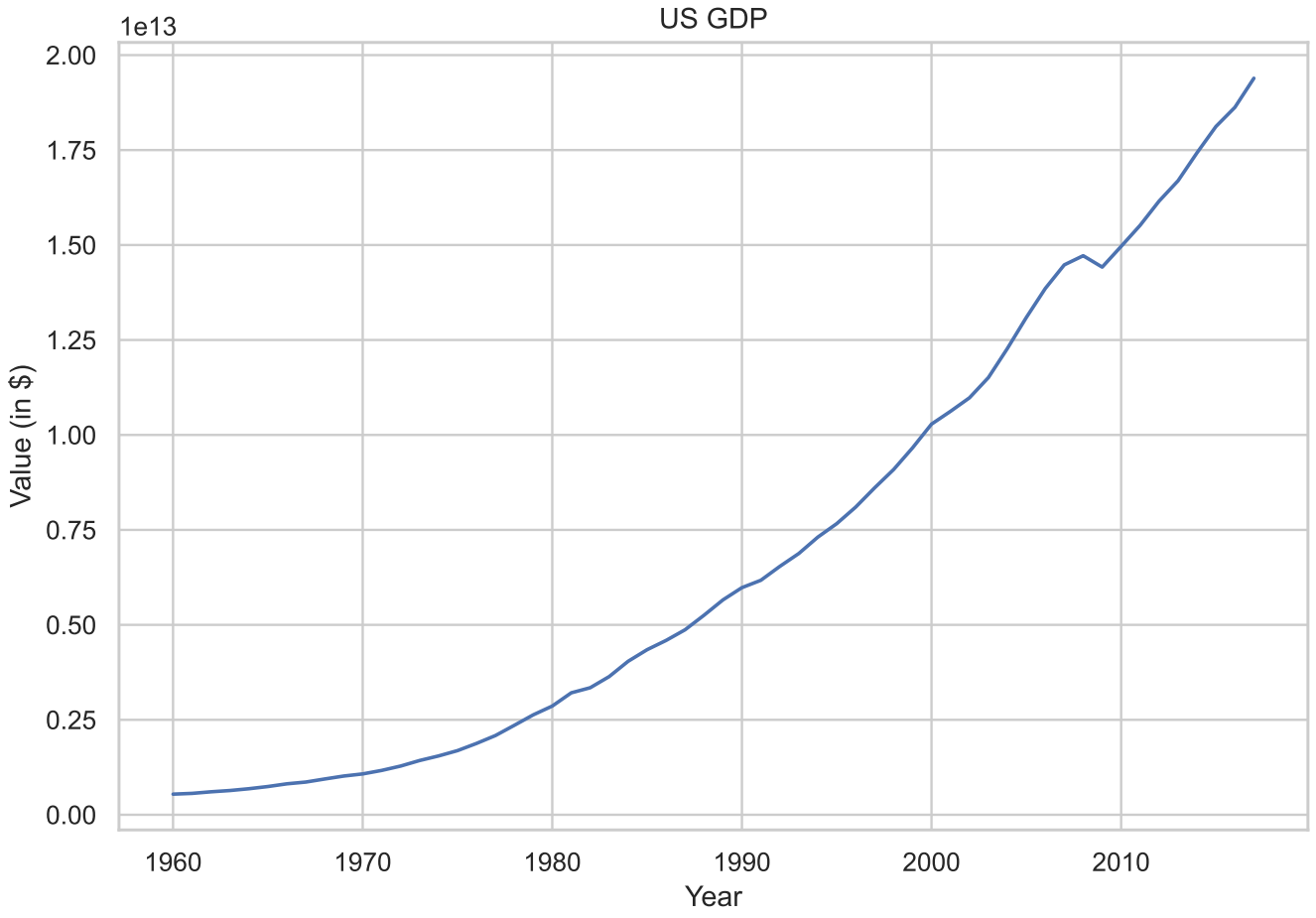
Exercise 2

For each of the following series, make a graph of the data. If transforming seems appropriate, do so and describe the effect.

United States GDP from global_economy.

```
# filter USA from the rest of the data
df_us_gdp = df_global_economy[df_global_economy["Country"] == "United States"]
```

```
#plot
plt.figure(figsize=(9,6))

sns.lineplot(x=df_us_gdp.index,
             y = df_us_gdp.GDP)

# edit labels
plt.xlabel("Year")
plt.ylabel("Value (in $)")
plt.title("US GDP")

plt.show()
```

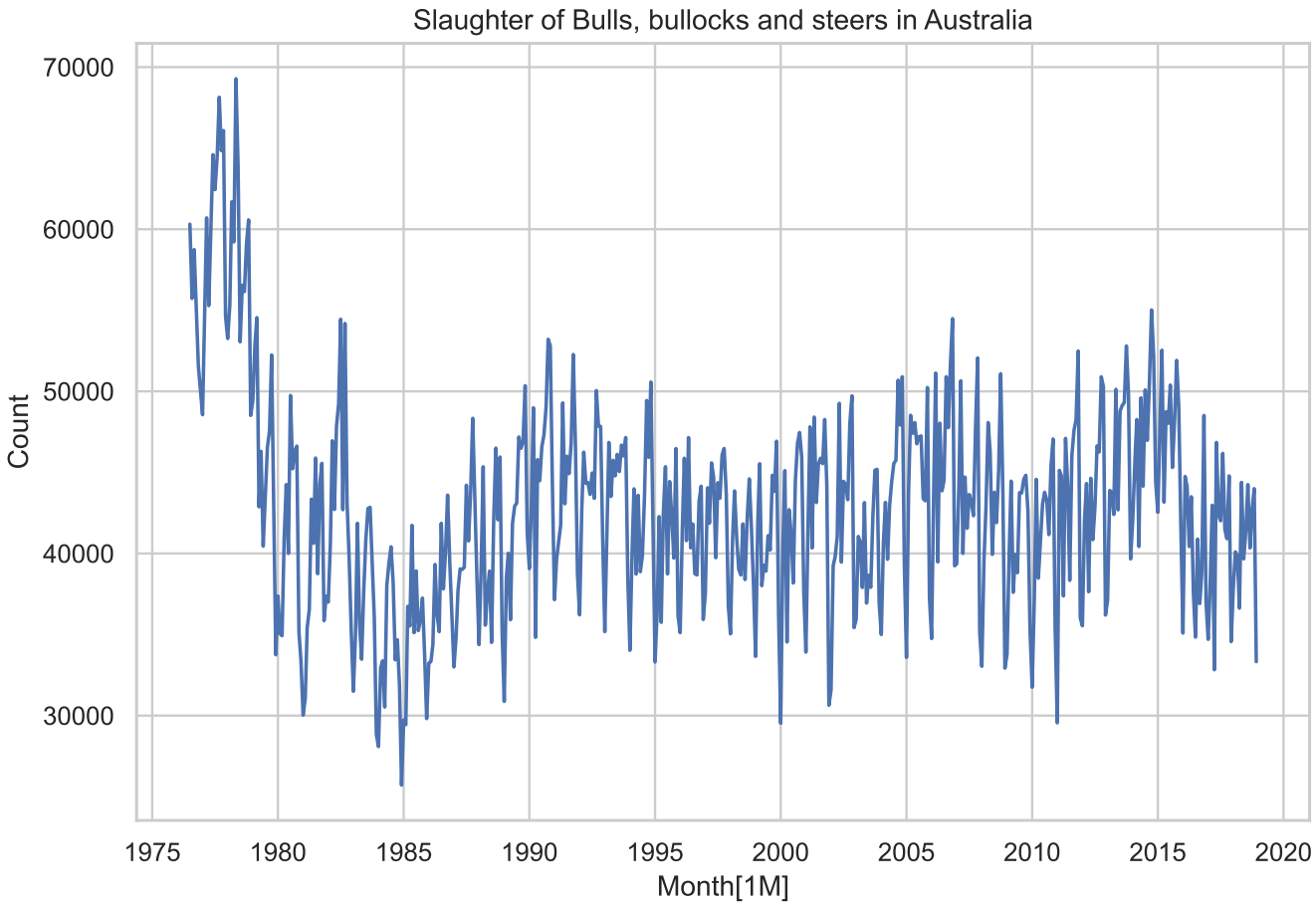


Slaughter of Victorian "Bulls, bullocks and steers" in aus_livestock.

```
# read in data
df_aus_livestock = pd.read_csv("aus_livestock.csv",
                               parse_dates=["Month"],
                               index_col=['Month'])
```

```
# filter animal in animals
df_filtered_livestock = df_aus_livestock[df_aus_livestock["Animal"] == \
  "Bulls, bullocks and steers"]

df_filtered_livestock = df_filtered_livestock[["Count"]]
```

```
# plot the time series
plt.figure(figsize=(9,6))

sns.lineplot(x = df_filtered_livestock.index,
             y = df_filtered_livestock["Count"],
             errorbar=None)

plt.xlabel("Month[1M]")
plt.ylabel("Count")
plt.title("Slaughter of Bulls, bullocks and steers in Australia")

plt.show()
```



Looking the plot, we can see that the variation of the data is all over place. To combat the we must imploy transformations in order to mitigate the variability of the data.

Let us try a log tranform on the data and determine whether it is able to reduce the variability on the data.