# Trade___Tariffs

By Hatem Elgenedy

November 7, 2025

[62]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import warnings
warnings.filterwarnings('ignore')
import sklearn
import requests, time, io
from dateutil.relativedelta import relativedelta
```

[63]:

```
data = pd.read_csv('/Users/hatemelgenedy/Desktop/AI and Data Science Microsoft␣
  ↪course/Capstone project 2025 /Project 1/Cleaned CSV FILES/
  ↪economic_freedom_index2019_data.csv', encoding="latin1")

print(data.info())
print(data.head())
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 186 entries, 0 to 185
Data columns (total 34 columns):
 #   Column                       Non-Null Count  Dtype
---  ------                       --------------  -----
 0   CountryID                    186 non-null    int64
 1   Country Name                 186 non-null    object
 2   WEBNAME                      186 non-null    object
 3   Region                       186 non-null    object
 4   World Rank                   180 non-null    float64
 5   Region Rank                  180 non-null    float64
 6   2019 Score                   180 non-null    float64
 7   Property Rights              185 non-null    float64
 8   Judical Effectiveness        185 non-null    float64
 9   Government Integrity         185 non-null    float64
 10  Tax Burden                   180 non-null    float64
 11  Gov't Spending               183 non-null    float64
 12  Fiscal Health                183 non-null    float64
 13  Business Freedom             185 non-null    float64
 14  Labor Freedom                184 non-null    float64
 15  Monetary Freedom             184 non-null    float64
 16  Trade Freedom                182 non-null    float64
 17  Investment Freedom           184 non-null    float64
 18  Financial Freedom            181 non-null    float64
 19  Tariff Rate (%)              182 non-null    float64
 20  Income Tax Rate (%)          183 non-null    float64
 21  Corporate Tax Rate (%)       183 non-null    float64
 22  Tax Burden % of GDP          179 non-null    float64
 23  Gov't Expenditure % of GDP   182 non-null    float64
 24  Country                      186 non-null    object
 25  Population (Millions)        186 non-null    object
 26  GDP (Billions, PPP)          185 non-null    object
 27  GDP Growth Rate (%)          184 non-null    float64
 28  5 Year GDP Growth Rate (%)   183 non-null    float64
 29  GDP per Capita (PPP)         184 non-null    object
 30  Unemployment (%)             181 non-null    object
 31  Inflation (%)                182 non-null    float64
 32  FDI Inflow (Millions)        181 non-null    object
 33  Public Debt (% of GDP)       182 non-null    float64
dtypes: float64(24), int64(1), object(9)
memory usage: 49.5+ KB
None
   CountryID Country Name      WEBNAME                         Region  \
0          1  Afghanistan  Afghanistan                   Asia-Pacific
1          2      Albania      Albania                         Europe
2          3      Algeria      Algeria  Middle East and North Africa
3          4       Angola       Angola             Sub-Saharan Africa
4          5    Argentina    Argentina                       Americas

   World Rank  Region Rank  2019 Score  Property Rights  \
0       152.0         39.0        51.5             19.6
1        52.0         27.0        66.5             54.8
2       171.0         14.0        46.2             31.6
3       156.0         33.0        50.6             35.9
4       148.0         26.0        52.2             47.8

   Judical Effectiveness  Government Integrity  …      Country  \
0                   29.6                  25.2  …  Afghanistan
1                   30.6                  40.4  …      Albania
2                   36.2                  28.9  …      Algeria
3                   26.6                  20.5  …       Angola
4                   44.5                  33.5  …    Argentina

   Population (Millions)  GDP (Billions, PPP)  GDP Growth Rate (%)  \
0                   35.5                $69.6                  2.5
1                    2.9                $36.0                  3.9
2                   41.5               $632.9                  2.0
3                   28.2               $190.3                  0.7
4                   44.1               $920.2                  2.9
```

```
   5 Year GDP Growth Rate (%)  GDP per Capita (PPP)  Unemployment (%)  \
0                         2.9                $1,958               8.8
1                         2.5               $12,507              13.9
2                         3.1               $15,237              10.0
3                         2.9                $6,753               8.2
4                         0.7               $20,876               8.7

   Inflation (%)  FDI Inflow (Millions)  Public Debt (% of GDP)
0            5.0                   53.9                     7.3
1            2.0                1,119.1                    71.2
2            5.6                1,203.0                    25.8
3           31.7               -2,254.5                    65.3
4           25.7               11,857.0                    52.6

[5 rows x 34 columns]
```

[140]:

```
!pip install ydata-profiling
from ydata_profiling import ProfileReport
profile = ProfileReport(data, title="Economic Freedom Index Data Profiling␣
 ↪Report", explorative=True)
profile.to_notebook_iframe()
profile.to_file("economic_profile_report.html")
```

```
Requirement already satisfied: ydata-profiling in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (4.17.0)
Requirement already satisfied: scipy<1.16,>=1.4.1 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from ydata-
profiling) (1.13.1)
Requirement already satisfied: pandas!=1.4.0,<3.0,>1.1 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from ydata-
profiling) (2.3.1)
Requirement already satisfied: matplotlib<=3.10,>=3.5 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from ydata-
profiling) (3.10.0)
Requirement already satisfied: pydantic>=2 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from ydata-profiling) (2.11.7)
Requirement already satisfied: PyYAML<6.1,>=5.0.0 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from ydata-
profiling) (6.0.2)
Requirement already satisfied: jinja2<3.2,>=2.11.1 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from ydata-
profiling) (3.1.6)
Requirement already satisfied: visions<0.8.2,>=0.7.5 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from
visions[type_image_path]<0.8.2,>=0.7.5->ydata-profiling) (0.8.1)
Requirement already satisfied: numpy<2.2,>=1.16.0 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from ydata-
profiling) (1.26.4)
Requirement already satisfied: minify-html>=0.15.0 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from ydata-
profiling) (0.18.1)
Requirement already satisfied: filetype>=1.0.0 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from ydata-profiling) (1.2.0)
Requirement already satisfied: phik<0.13,>=0.11.1 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from ydata-
profiling) (0.12.5)
Requirement already satisfied: requests<3,>=2.24.0 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from ydata-
profiling) (2.32.4)
Requirement already satisfied: tqdm<5,>=4.48.2 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from ydata-profiling) (4.67.1)
Requirement already satisfied: seaborn<0.14,>=0.10.1 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from ydata-
profiling) (0.13.2)
Requirement already satisfied: multimethod<2,>=1.4 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from ydata-
profiling) (1.12)
Requirement already satisfied: statsmodels<1,>=0.13.2 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from ydata-
profiling) (0.14.5)
Requirement already satisfied: typeguard<5,>=3 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from ydata-profiling) (4.4.4)
Requirement already satisfied: imagehash==4.3.1 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from ydata-profiling) (4.3.1)
Requirement already satisfied: wordcloud>=1.9.3 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from ydata-profiling) (1.9.4)
Requirement already satisfied: dacite>=1.8 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from ydata-profiling) (1.9.2)
Requirement already satisfied: numba<=0.61,>=0.56.0 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from ydata-
profiling) (0.61.0)
Requirement already satisfied: PyWavelets in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from imagehash==4.3.1->ydata-profiling)
(1.9.0)
Requirement already satisfied: pillow in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from imagehash==4.3.1->ydata-profiling)
(11.3.0)
Requirement already satisfied: MarkupSafe>=2.0 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from jinja2<3.2,>=2.11.1->ydata-profiling)
(3.0.2)
Requirement already satisfied: contourpy>=1.0.1 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from matplotlib<=3.10,>=3.5->ydata-profiling)
(1.3.1)
Requirement already satisfied: cycler>=0.10 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from matplotlib<=3.10,>=3.5->ydata-profiling)
```

```
(0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from
matplotlib<=3.10,>=3.5->ydata-profiling) (4.55.3)
Requirement already satisfied: kiwisolver>=1.3.1 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from
matplotlib<=3.10,>=3.5->ydata-profiling) (1.4.8)
Requirement already satisfied: packaging>=20.0 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from matplotlib<=3.10,>=3.5->ydata-profiling)
(24.2)
Requirement already satisfied: pyparsing>=2.3.1 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from matplotlib<=3.10,>=3.5->ydata-profiling)
(3.2.0)
Requirement already satisfied: python-dateutil>=2.7 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from
matplotlib<=3.10,>=3.5->ydata-profiling) (2.9.0.post0)
Requirement already satisfied: llvmlite<0.45,>=0.44.0dev0 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from
numba<=0.61,>=0.56.0->ydata-profiling) (0.44.0)
Requirement already satisfied: pytz>=2020.1 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from pandas!=1.4.0,<3.0,>1.1->ydata-profiling)
(2025.2)
Requirement already satisfied: tzdata>=2022.7 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from pandas!=1.4.0,<3.0,>1.1->ydata-profiling)
(2025.2)
Requirement already satisfied: joblib>=0.14.1 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from phik<0.13,>=0.11.1->ydata-profiling)
(1.5.1)
Requirement already satisfied: charset_normalizer<4,>=2 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from
requests<3,>=2.24.0->ydata-profiling) (3.3.2)
Requirement already satisfied: idna<4,>=2.5 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from requests<3,>=2.24.0->ydata-profiling)
(3.7)
Requirement already satisfied: urllib3<3,>=1.21.1 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from
requests<3,>=2.24.0->ydata-profiling) (2.5.0)
Requirement already satisfied: certifi>=2017.4.17 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from
requests<3,>=2.24.0->ydata-profiling) (2025.8.3)
Requirement already satisfied: patsy>=0.5.6 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from statsmodels<1,>=0.13.2->ydata-profiling)
(1.0.2)
Requirement already satisfied: typing_extensions>=4.14.0 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from
typeguard<5,>=3->ydata-profiling) (4.15.0)
Requirement already satisfied: attrs>=19.3.0 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from
visions<0.8.2,>=0.7.5->visions[type_image_path]<0.8.2,>=0.7.5->ydata-profiling)
(24.3.0)
Requirement already satisfied: networkx>=2.4 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from
visions<0.8.2,>=0.7.5->visions[type_image_path]<0.8.2,>=0.7.5->ydata-profiling)
(3.4.2)
Requirement already satisfied: puremagic in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from
visions<0.8.2,>=0.7.5->visions[type_image_path]<0.8.2,>=0.7.5->ydata-profiling)
(1.30)
Requirement already satisfied: annotated-types>=0.6.0 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from
pydantic>=2->ydata-profiling) (0.6.0)
Requirement already satisfied: pydantic-core==2.33.2 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from
pydantic>=2->ydata-profiling) (2.33.2)
Requirement already satisfied: typing-inspection>=0.4.0 in
/opt/anaconda3/envs/anaconda-nlp/lib/python3.11/site-packages (from
pydantic>=2->ydata-profiling) (0.4.0)
Requirement already satisfied: six>=1.5 in /opt/anaconda3/envs/anaconda-
nlp/lib/python3.11/site-packages (from python-
dateutil>=2.7->matplotlib<=3.10,>=3.5->ydata-profiling) (1.17.0)
```

```
<IPython.core.display.HTML object>
```

```
Summarize dataset:   0%|          | 0/5 [00:00<?, ?it/s]
```

```
100%|     | 35/35 [00:00<00:00, 60963.72it/s]
```

```
Generate report structure:   0%|          | 0/1 [00:00<?, ?it/s]
```

```
Render HTML:   0%|          | 0/1 [00:00<?, ?it/s]
```

```
<IPython.core.display.HTML object>
```

```
Export report to file:   0%|          | 0/1 [00:00<?, ?it/s]
```

[85]:
```
df = data
```

[86]:
```
df.isnull().sum()
```

[86]:
```
CountryID                   0
Country Name              186
WEBNAME                   186
Region                    186
World Rank                  6
Region Rank                 6
2019 Score                  6
Property Rights             1
Judical Effectiveness       1
Government Integrity        1
```

```
Tax Burden                       6
Gov't Spending                   3
Fiscal Health                    3
Business Freedom                 1
Labor Freedom                    2
Monetary Freedom                 2
Trade Freedom                    4
Investment Freedom               2
Financial Freedom                5
Tariff Rate (%)                  4
Income Tax Rate (%)              3
Corporate Tax Rate (%)           3
Tax Burden % of GDP              7
Gov't Expenditure % of GDP       4
Country                        186
Population (Millions)            1
GDP (Billions, PPP)              3
GDP Growth Rate (%)              2
5 Year GDP Growth Rate (%)       3
GDP per Capita (PPP)             4
Unemployment (%)                 6
Inflation (%)                    4
FDI Inflow (Millions)            5
Public Debt (% of GDP)           4
TariffGroup                      4
dtype: int64
```

[87]:

```python
df = data
def to_number(x):
    if pd.isna(x):
        return np.nan
    if isinstance(x, (int, float)):
        return float(x)
    s = str(x)

    s = re.sub(r"[\$,()%]", "", s).replace(",", "").strip()

    if isinstance(x, str) and "(" in x and ")" in x and "-" not in x:
        try:
            return -float(s)
        except:
            return np.nan
    try:
        return float(s)
    except:
        return np.nan
```

```python
for col in df.columns:
    try:
        df[col] = df[col].apply(to_number)
    except:
        pass

df.describe(include="all")
```

[87]:

```
        CountryID  Country Name  WEBNAME  Region  World Rank  Region Rank  \
count  186.000000           0.0      0.0     0.0  180.000000   180.000000
mean    93.500000           NaN      NaN     NaN   90.500000    20.538889
std     53.837719           NaN      NaN     NaN   52.105662    12.738611
min      1.000000           NaN      NaN     NaN    1.000000     1.000000
25%     47.250000           NaN      NaN     NaN   45.750000     9.750000
50%     93.500000           NaN      NaN     NaN   90.500000    19.500000
75%    139.750000           NaN      NaN     NaN  135.250000    31.000000
max    186.000000           NaN      NaN     NaN  180.000000    47.000000

       2019 Score  Property Rights  Judical Effectiveness  \
count  180.000000       185.000000             185.000000
mean    60.768333        52.327568              44.899459
std     11.255725        19.608526              18.104745
min      5.900000         7.600000               5.000000
25%     53.950000        37.000000              31.000000
50%     60.750000        50.100000              42.900000
75%     67.800000        65.900000              54.700000
max     90.200000        97.400000              92.400000

       Government Integrity  …  Population (Millions)  GDP (Billions, PPP)  \
count            185.000000  …             185.000000           183.000000
mean              41.470270  …              40.157297           694.233333
std               19.793193  …             145.155754          2421.728981
min                7.900000  …               0.100000             0.200000
25%               27.200000  …               2.700000            25.700000
50%               35.500000  …               8.800000            83.600000
75%               50.300000  …              29.500000           402.550000
max               96.700000  …            1390.100000         23159.100000

       GDP Growth Rate (%)  5 Year GDP Growth Rate (%)  GDP per Capita (PPP)  \
count           184.000000                  183.000000            182.000000
mean              3.470109                    2.984153          20757.324176
std               5.835732                    2.926503          22358.225141
min             -14.000000                  -16.100000            677.000000
25%               1.800000                    1.900000           4479.500000
50%               3.200000                    3.000000          12697.500000
75%               4.650000                    4.450000          29509.250000
max              70.800000                    9.900000         124529.000000
```

```
       Unemployment (%)  Inflation (%)  FDI Inflow (Millions)  \
count        180.000000     182.000000             181.000000
mean           7.426111      10.586264            7911.153039
std            5.684856      80.507501           25984.794434
min            0.100000      -0.900000           -8296.900000
25%            3.775000       1.300000             213.800000
50%            5.750000       2.750000             896.600000
75%            9.425000       5.450000            4046.000000
max           27.300000    1087.500000          275381.000000

       Public Debt (% of GDP)  TariffGroup
count              182.000000          0.0
mean                56.469231          NaN
std                 34.163855          NaN
min                  0.000000          NaN
25%                 34.950000          NaN
50%                 49.900000          NaN
75%                 70.125000          NaN
max                236.400000          NaN

[8 rows x 35 columns]
```

[88]:

```python
coverage = df.isna().mean().sort_values(ascending=False) * 100
coverage.head(15)
```

[88]:

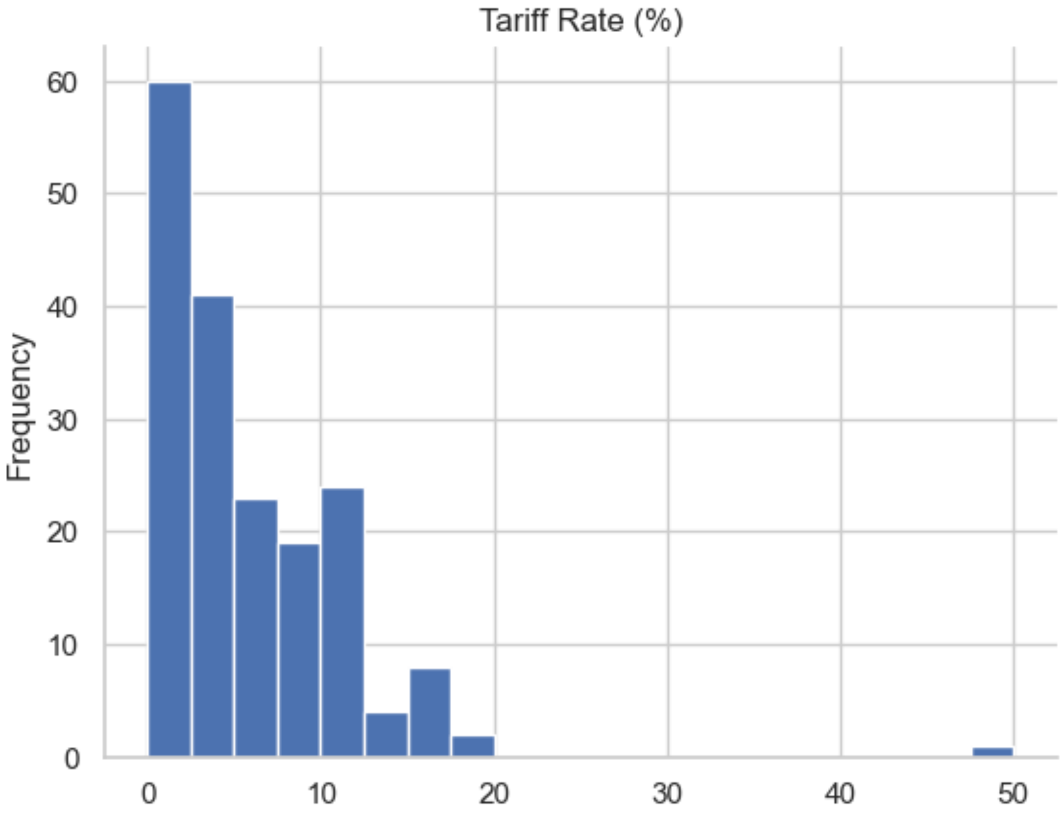
```
TariffGroup                 100.000000
WEBNAME                     100.000000
Region                      100.000000
Country                     100.000000
Country Name                100.000000
Tax Burden % of GDP           3.763441
Tax Burden                    3.225806
Unemployment (%)              3.225806
2019 Score                    3.225806
World Rank                    3.225806
Region Rank                   3.225806
Financial Freedom             2.688172
FDI Inflow (Millions)         2.688172
GDP per Capita (PPP)          2.150538
Trade Freedom                 2.150538
dtype: float64
```

[89]:

```python
df[["Tariff Rate (%)", "Trade Freedom", "GDP per Capita (PPP)", "FDI Inflow␣
↪(Millions)"]].describe()
```

```
[89]:        Tariff Rate (%)  Trade Freedom  GDP per Capita (PPP)  \
      count       182.000000     182.000000            182.000000
      mean          5.986813      74.260989          20757.324176
      std           5.533568      12.261766          22358.225141
      min           0.000000       0.000000            677.000000
      25%           2.000000      66.650000           4479.500000
      50%           4.300000      76.100000          12697.500000
      75%           8.775000      84.300000          29509.250000
      max          50.000000      95.000000         124529.000000

             FDI Inflow (Millions)
      count             181.000000
      mean             7911.153039
      std             25984.794434
      min             -8296.900000
      25%               213.800000
      50%               896.600000
      75%              4046.000000
      max            275381.000000
```

```
[90]: df['Tariff Rate (%)'].plot(kind='hist', bins=20, title='Tariff Rate (%)')
      plt.gca().spines[['top', 'right',]].set_visible(False)
```



[91]:

```
cols = [
    "Share of China's exports affected by punitive tariffs",
    "Share of US exports affected by punitive tariffs"]
```

[92]:

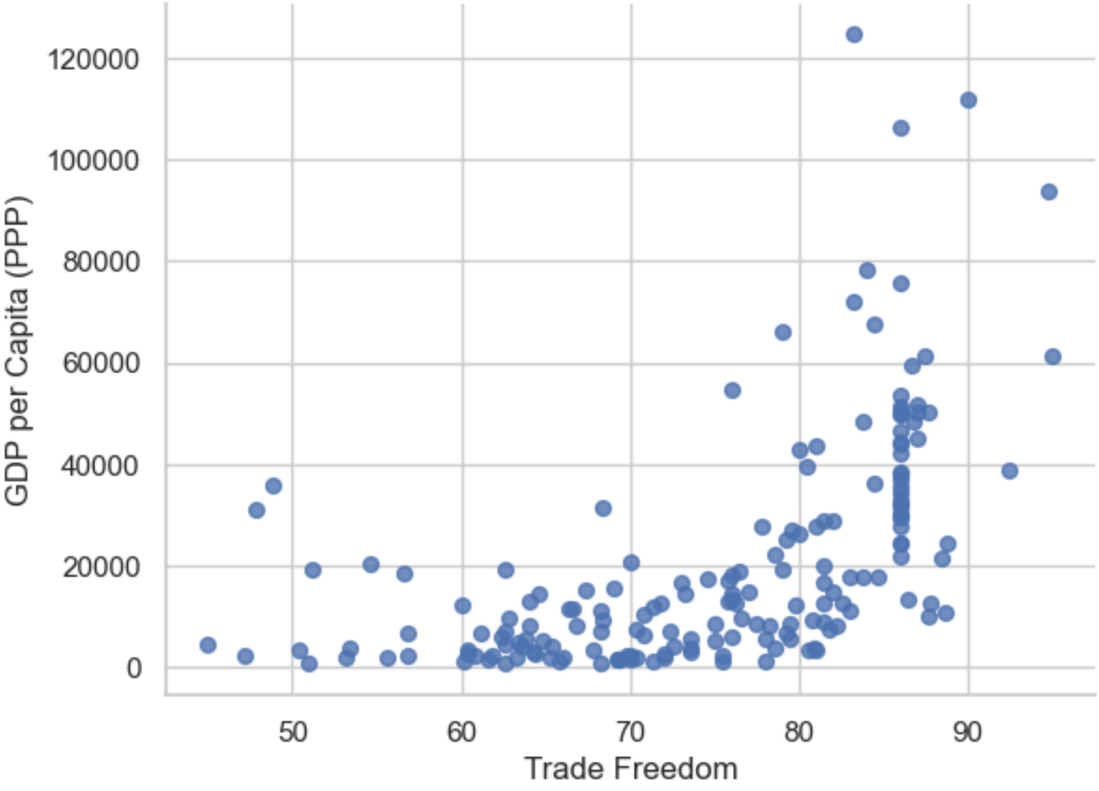
```
corr_cols = [c for c in [
    "Tariff Rate (%)", "Trade Freedom", "GDP per Capita (PPP)",
    "GDP Growth Rate (%)", "5 Year GDP Growth Rate (%)",
    "Inflation (%)", "Unemployment (%)", "FDI Inflow (Millions)",
    "Tax Burden", "Gov't Spending", "Investment Freedom ", "Financial Freedom"
] if c in df.columns]

print("Using columns:", corr_cols)


for c in corr_cols:
    s = df[c].astype(str)
    s = s.str.replace(r"[,\$%]", "", regex=True)
    s = s.str.replace(r"^\((.*)\)$", r"-\1", regex=True)
    df[c] = pd.to_numeric(s, errors="coerce")


corr = df[corr_cols].corr().round(2)
print(corr)
```

```
Using columns: ['Tariff Rate (%)', 'Trade Freedom', 'GDP per Capita (PPP)', 'GDP
Growth Rate (%)', '5 Year GDP Growth Rate (%)', 'Inflation (%)', 'Unemployment
(%)', 'FDI Inflow (Millions)', 'Tax Burden', "Gov't Spending", 'Investment
Freedom ', 'Financial Freedom']
                            Tariff Rate (%)  Trade Freedom  \
Tariff Rate (%)                        1.00          -0.95
Trade Freedom                         -0.95           1.00
GDP per Capita (PPP)                  -0.47           0.56
GDP Growth Rate (%)                   -0.07           0.11
5 Year GDP Growth Rate (%)            -0.00           0.03
Inflation (%)                          0.09          -0.12
Unemployment (%)                      -0.00          -0.01
FDI Inflow (Millions)                 -0.17           0.21
Tax Burden                            -0.27           0.18
Gov't Spending                         0.08          -0.13
Investment Freedom                    -0.46           0.60
Financial Freedom                     -0.50           0.64

                            GDP per Capita (PPP)  GDP Growth Rate (%)  \
Tariff Rate (%)                            -0.47                -0.07
Trade Freedom                               0.56                 0.11
GDP per Capita (PPP)                        1.00                -0.07
GDP Growth Rate (%)                        -0.07                 1.00
5 Year GDP Growth Rate (%)                 -0.15                 0.16
Inflation (%)                              -0.05                -0.20
Unemployment (%)                           -0.17                -0.01
FDI Inflow (Millions)                       0.29                 0.04
Tax Burden                                 -0.09                 0.15
Gov't Spending                             -0.28                -0.05
Investment Freedom                          0.48                -0.11
Financial Freedom                           0.59                 0.06

                            5 Year GDP Growth Rate (%)  Inflation (%)  \
Tariff Rate (%)                                  -0.00           0.09
Trade Freedom                                     0.03          -0.12
GDP per Capita (PPP)                             -0.15          -0.05
GDP Growth Rate (%)                               0.16          -0.20
5 Year GDP Growth Rate (%)                        1.00          -0.28
Inflation (%)                                    -0.28           1.00
Unemployment (%)                                 -0.26           0.01
FDI Inflow (Millions)                             0.01          -0.03
Tax Burden                                        0.21          -0.01
Gov't Spending                                    0.29          -0.01
Investment Freedom                                0.01          -0.23
Financial Freedom                                -0.07          -0.18

                            Unemployment (%)  FDI Inflow (Millions)  \
Tariff Rate (%)                        -0.00                  -0.17
Trade Freedom                          -0.01                   0.21
GDP per Capita (PPP)                   -0.17                   0.29
GDP Growth Rate (%)                    -0.01                   0.04
5 Year GDP Growth Rate (%)             -0.26                   0.01
Inflation (%)                           0.01                  -0.03
Unemployment (%)                        1.00                  -0.10
FDI Inflow (Millions)                  -0.10                   1.00
Tax Burden                             -0.11                  -0.08
Gov't Spending                         -0.14                  -0.03
Investment Freedom                     -0.00                   0.16
Financial Freedom                       0.01                   0.26

                            Tax Burden  Gov't Spending  Investment Freedom   \
Tariff Rate (%)                  -0.27            0.08                -0.46
Trade Freedom                     0.18           -0.13                 0.60
GDP per Capita (PPP)             -0.09           -0.28                 0.48
GDP Growth Rate (%)               0.15           -0.05                -0.11
5 Year GDP Growth Rate (%)        0.21            0.29                 0.01
Inflation (%)                    -0.01           -0.01                -0.23
Unemployment (%)                 -0.11           -0.14                -0.00
FDI Inflow (Millions)            -0.08           -0.03                 0.16
Tax Burden                        1.00            0.39                -0.12
```

```
Gov't Spending                 0.39            1.00                -0.09
Investment Freedom            -0.12           -0.09                 1.00
Financial Freedom             -0.05           -0.13                 0.81

                            Financial Freedom
Tariff Rate (%)                         -0.50
Trade Freedom                            0.64
GDP per Capita (PPP)                     0.59
GDP Growth Rate (%)                      0.06
5 Year GDP Growth Rate (%)              -0.07
Inflation (%)                           -0.18
Unemployment (%)                         0.01
FDI Inflow (Millions)                    0.26
Tax Burden                              -0.05
Gov't Spending                          -0.13
Investment Freedom                       0.81
Financial Freedom                        1.00
```

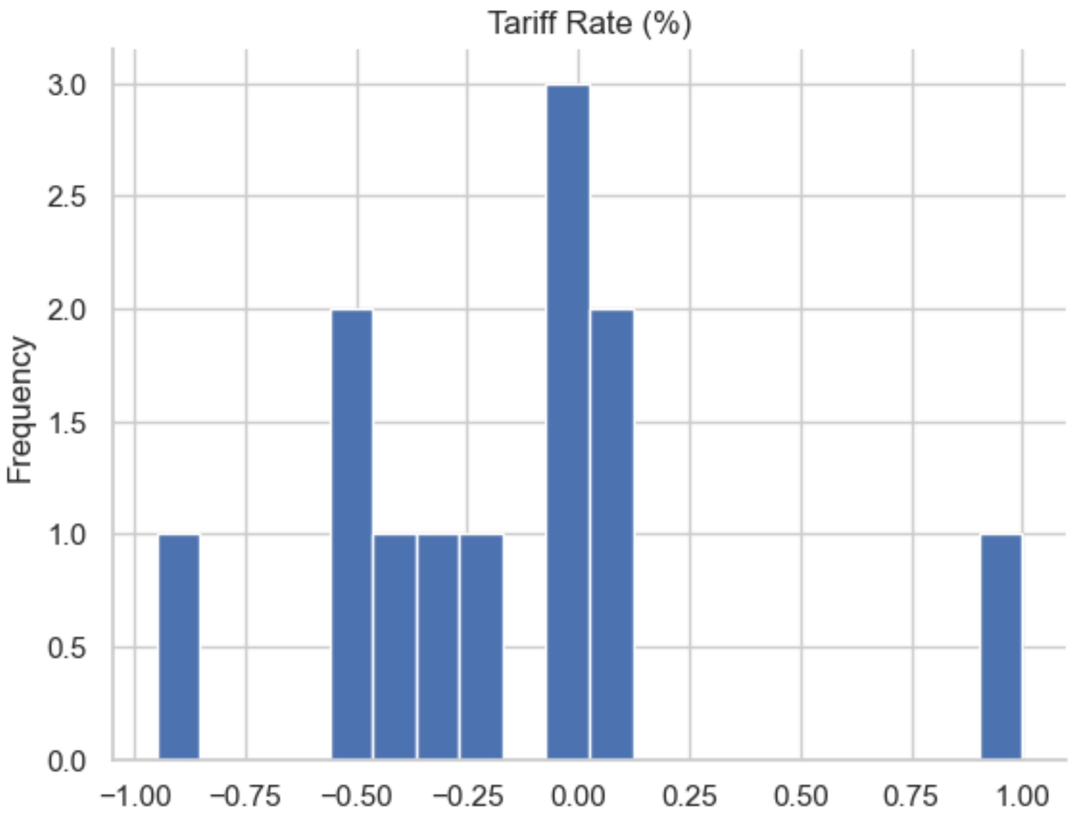
[93]:

```
from matplotlib import pyplot as plt
df.plot(kind='scatter', x='Trade Freedom', y='GDP per Capita (PPP)', s=32,␣
 ↪alpha=.8)
plt.gca().spines[['top', 'right',]].set_visible(False)
```



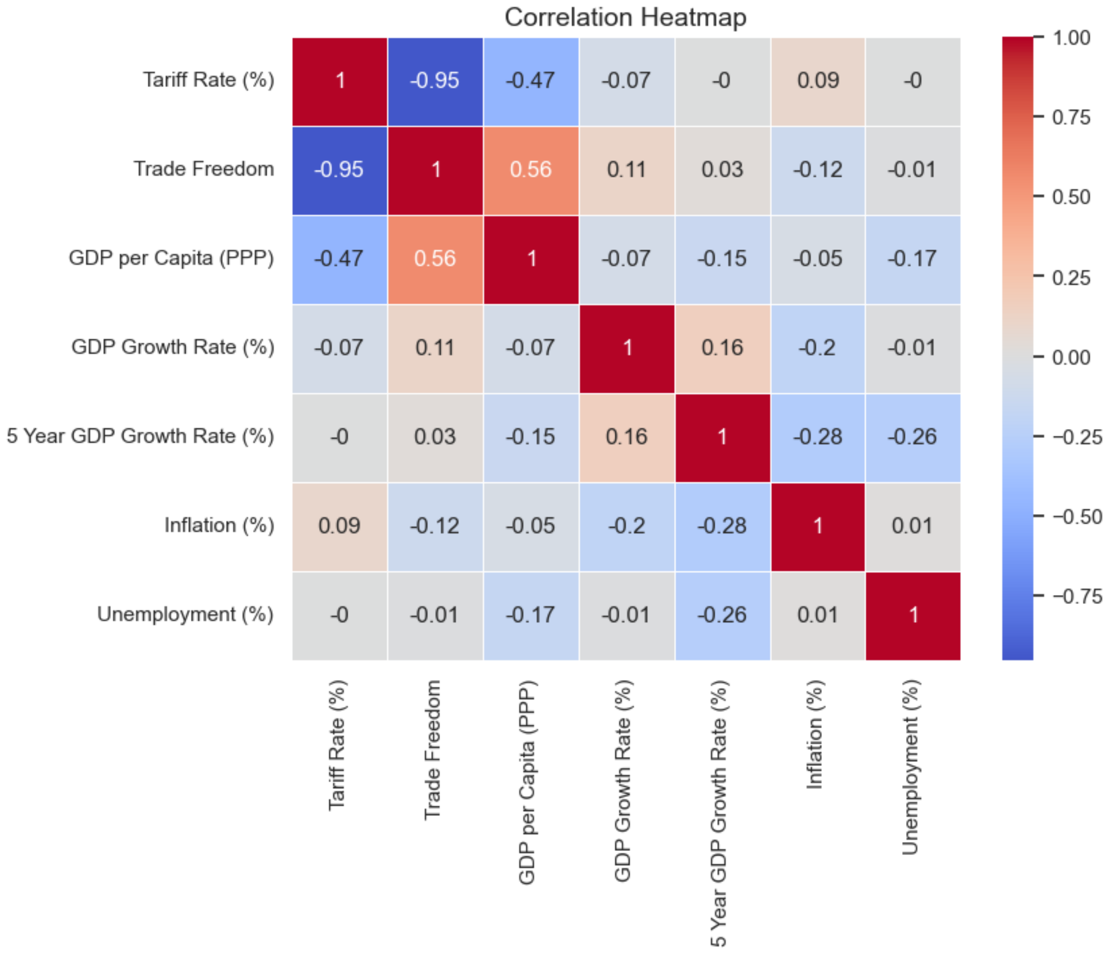
[94]:

```
corr['Tariff Rate (%)'].plot(kind='hist', bins=20, title='Tariff Rate (%)')
plt.gca().spines[['top', 'right',]].set_visible(False)
```



[95]:

```
corr = df[[
    "Tariff Rate (%)", "Trade Freedom", "GDP per Capita (PPP)",
    "GDP Growth Rate (%)", "5 Year GDP Growth Rate (%)",
    "Inflation (%)", "Unemployment (%)"
]].corr().round(2)

corr.style.background_gradient(cmap="coolwarm").format(precision=2)
plt.figure(figsize=(8,6))
sns.heatmap(corr, annot=True, cmap="coolwarm", center=0, linewidths=0.5)
plt.title("Correlation Heatmap", fontsize=14)
plt.show()
```



[96]:
```
col = "Tariff Rate (%)"


mask = df["Country Name"].astype(str).str.contains("Korea", na=False)
korea_rows = df.loc[mask, ["Country Name", col]]

for _, r in korea_rows.iterrows():
    print(r["Country Name"], "→", repr(str(r[col])))
```

[97]:
```
import re
import difflib

country_col = "Country Name" if "Country Name" in df.columns else "Country"

names = df[country_col].astype(str)
```

```
pat = r"(?i)\bkorea\b|dem\.?\s*people|dprk|rep\.?
 ↪\s*of\s*korea|korea,\s*(north|south)"
korea_like = names[names.str.contains(pat, regex=True, na=False)].unique()
print("Candidates containing Korea-like patterns:", korea_like)


if len(korea_like) == 0:
    uniq = names.dropna().unique().tolist()
    close = difflib.get_close_matches("korea", uniq, n=10, cutoff=0.3)
    print("Closest fuzzy matches to 'korea':", close)


print("\nSample of country names:")
print(names.unique()[:50])
```

```
Candidates containing Korea-like patterns: []
Closest fuzzy matches to 'korea': []

Sample of country names:
['nan']
```

[98]:

```
print("Shape:", df.shape)
print("\nColumns:", list(df.columns))

print("\nNon-null counts per column:")
print(df.notna().sum())

print("\nFirst 5 rows:")
display(df.head())
```

```
Shape: (186, 35)

Columns: ['CountryID', 'Country Name', 'WEBNAME', 'Region', 'World Rank',
'Region Rank', '2019 Score', 'Property Rights', 'Judical Effectiveness',
'Government Integrity', 'Tax Burden', "Gov't Spending", 'Fiscal Health',
'Business Freedom', 'Labor Freedom', 'Monetary Freedom', 'Trade Freedom',
'Investment Freedom ', 'Financial Freedom', 'Tariff Rate (%)', 'Income Tax Rate
(%)', 'Corporate Tax Rate (%)', 'Tax Burden % of GDP', "Gov't Expenditure % of
GDP ", 'Country', 'Population (Millions)', 'GDP (Billions, PPP)', 'GDP Growth
Rate (%)', '5 Year GDP Growth Rate (%)', 'GDP per Capita (PPP)', 'Unemployment
(%)', 'Inflation (%)', 'FDI Inflow (Millions)', 'Public Debt (% of GDP)',
'TariffGroup']

Non-null counts per column:
CountryID                      186
Country Name                     0
WEBNAME                          0
Region                           0
```

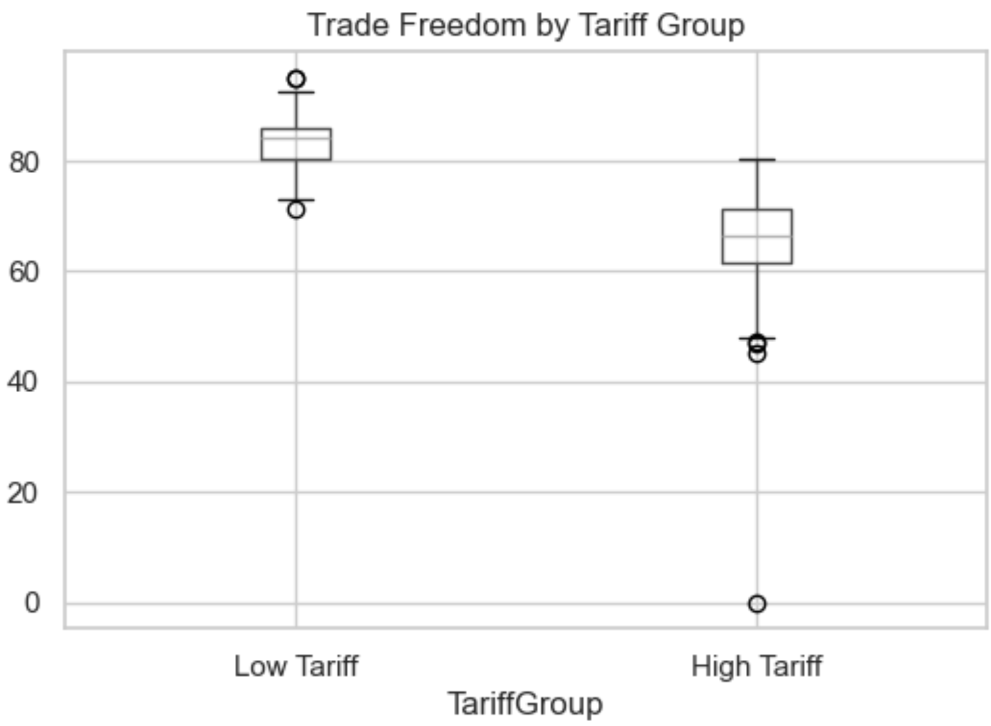
```
World Rank                     180
Region Rank                    180
2019 Score                     180
Property Rights                185
Judical Effectiveness          185
Government Integrity           185
Tax Burden                     180
Gov't Spending                 183
Fiscal Health                  183
Business Freedom               185
Labor Freedom                  184
Monetary Freedom               184
Trade Freedom                  182
Investment Freedom             184
Financial Freedom              181
Tariff Rate (%)                182
Income Tax Rate (%)            183
Corporate Tax Rate (%)         183
Tax Burden % of GDP            179
Gov't Expenditure % of GDP     182
Country                          0
Population (Millions)          185
GDP (Billions, PPP)            183
GDP Growth Rate (%)            184
5 Year GDP Growth Rate (%)     183
GDP per Capita (PPP)           182
Unemployment (%)               180
Inflation (%)                  182
FDI Inflow (Millions)          181
Public Debt (% of GDP)         182
TariffGroup                      0
dtype: int64

First 5 rows:
   CountryID  Country Name  WEBNAME  Region  World Rank  Region Rank  \
0        1.0           NaN      NaN     NaN       152.0         39.0
1        2.0           NaN      NaN     NaN        52.0         27.0
2        3.0           NaN      NaN     NaN       171.0         14.0
3        4.0           NaN      NaN     NaN       156.0         33.0
4        5.0           NaN      NaN     NaN       148.0         26.0

   2019 Score  Property Rights  Judical Effectiveness  Government Integrity  \
0        51.5             19.6                   29.6                  25.2
1        66.5             54.8                   30.6                  40.4
2        46.2             31.6                   36.2                  28.9
3        50.6             35.9                   26.6                  20.5
4        52.2             47.8                   44.5                  33.5
```

```
   …  Population (Millions)  GDP (Billions, PPP)  GDP Growth Rate (%)  \
0  …                   35.5                 69.6                  2.5
1  …                    2.9                 36.0                  3.9
2  …                   41.5                632.9                  2.0
3  …                   28.2                190.3                  0.7
4  …                   44.1                920.2                  2.9

   5 Year GDP Growth Rate (%)  GDP per Capita (PPP)  Unemployment (%)  \
0                         2.9                1958.0               8.8
1                         2.5               12507.0              13.9
2                         3.1               15237.0              10.0
3                         2.9                6753.0               8.2
4                         0.7               20876.0               8.7

   Inflation (%)  FDI Inflow (Millions)  Public Debt (% of GDP)  TariffGroup
0            5.0                   53.9                     7.3          NaN
1            2.0                 1119.1                    71.2          NaN
2            5.6                 1203.0                    25.8          NaN
3           31.7                -2254.5                    65.3          NaN
4           25.7                11857.0                    52.6          NaN

[5 rows x 35 columns]
```

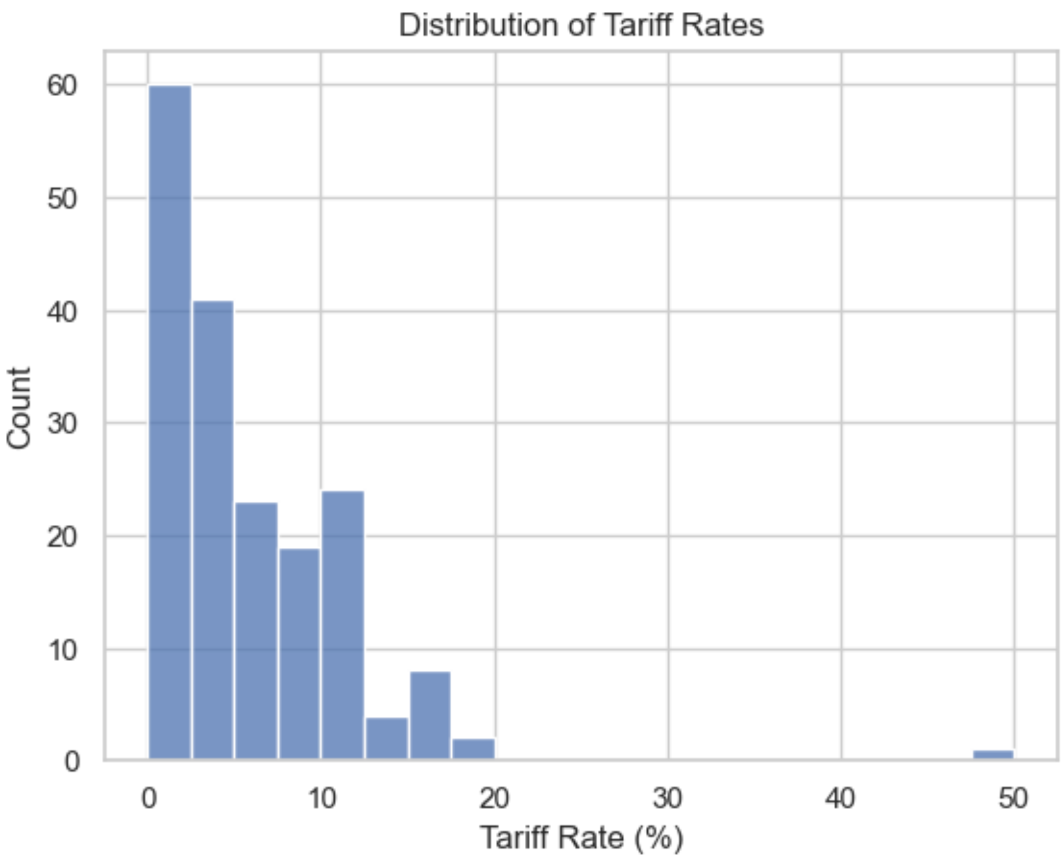
```
[100]: def iqr_outliers(series, k=1.5):
           q1, q3 = np.percentile(series.dropna(), [25, 75])
           iqr = q3 - q1
           lower, upper = q1 - k*iqr, q3 + k*iqr
           return df[(series < lower) | (series > upper)][["Country Name", series.
        ↪name]]

       outliers_tariff = iqr_outliers(df["Tariff Rate (%)"])
       outliers_tariff
```

```
[100]:     Country Name  Tariff Rate (%)
       88           NaN             50.0
```

```
[101]: df["TariffGroup"] = pd.qcut(df["Tariff Rate (%)"], q=2, labels=["Low Tariff",␣
        ↪"High Tariff"])

       df.groupby("TariffGroup")[["Trade Freedom", "GDP per Capita (PPP)", "GDP Growth␣
        ↪Rate (%)"]].mean()

       df.boxplot(column="Trade Freedom", by="TariffGroup", figsize=(6,4))
       plt.title("Trade Freedom by Tariff Group")
       plt.suptitle("")
       plt.show()
```



[102]:
```
# Distribution
sns.histplot(df["Tariff Rate (%)"].dropna(), bins=20)
plt.title("Distribution of Tariff Rates")
plt.show()
```



```
[83]: df.head()
```

```
[83]:    CountryID  Country Name  WEBNAME  Region  World Rank  Region Rank  \
      0        1.0           NaN      NaN     NaN       152.0         39.0
      1        2.0           NaN      NaN     NaN        52.0         27.0
      2        3.0           NaN      NaN     NaN       171.0         14.0
      3        4.0           NaN      NaN     NaN       156.0         33.0
      4        5.0           NaN      NaN     NaN       148.0         26.0

         2019 Score  Property Rights  Judical Effectiveness  Government Integrity  \
      0        51.5             19.6                   29.6                  25.2
      1        66.5             54.8                   30.6                  40.4
      2        46.2             31.6                   36.2                  28.9
      3        50.6             35.9                   26.6                  20.5
      4        52.2             47.8                   44.5                  33.5

         …  Population (Millions)  GDP (Billions, PPP)  GDP Growth Rate (%)  \
      0  …                   35.5                 69.6                  2.5
      1  …                    2.9                 36.0                  3.9
```

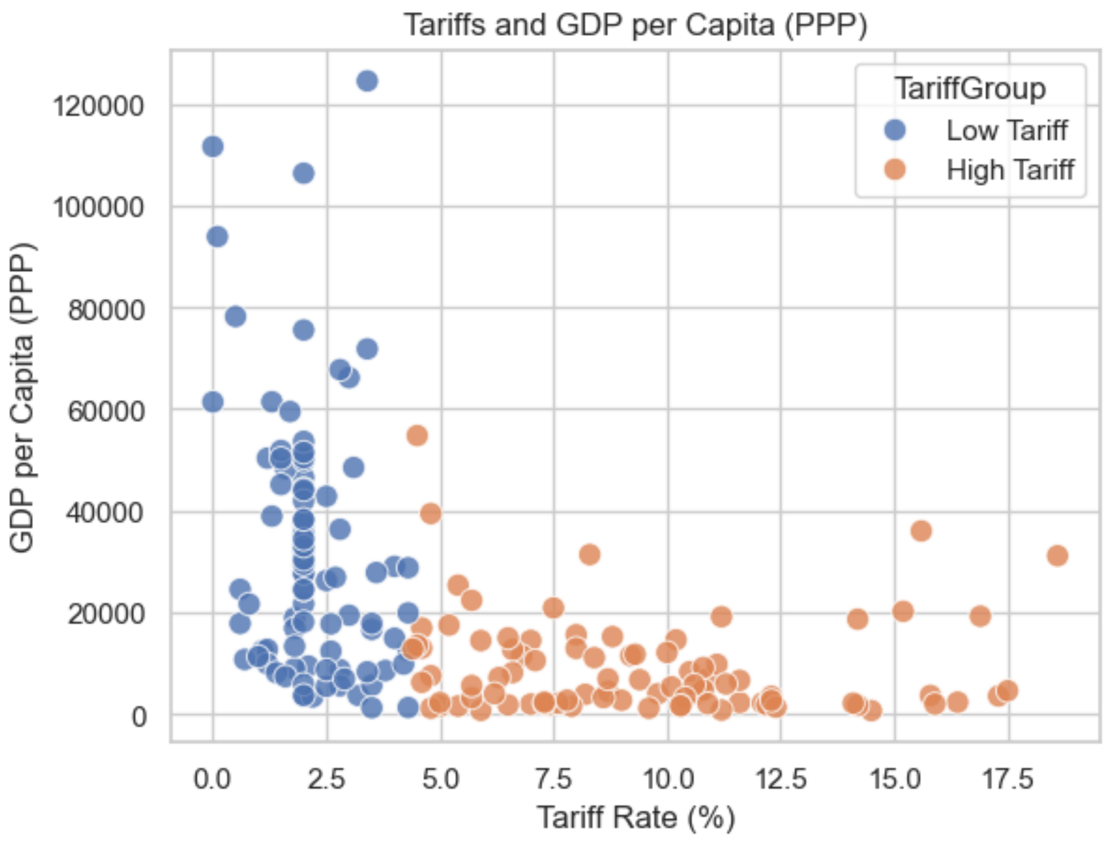
```
2  …                   41.5                632.9                   2.0
3  …                   28.2                190.3                   0.7
4  …                   44.1                920.2                   2.9

   5 Year GDP Growth Rate (%)  GDP per Capita (PPP)  Unemployment (%)  \
0                         2.9                1958.0               8.8
1                         2.5               12507.0              13.9
2                         3.1               15237.0              10.0
3                         2.9                6753.0               8.2
4                         0.7               20876.0               8.7

   Inflation (%)  FDI Inflow (Millions)  Public Debt (% of GDP)  TariffGroup
0            5.0                   53.9                     7.3  High Tariff
1            2.0                 1119.1                    71.2   Low Tariff
2            5.6                 1203.0                    25.8  High Tariff
3           31.7                -2254.5                    65.3  High Tariff
4           25.7                11857.0                    52.6  High Tariff

[5 rows x 35 columns]
```

[84]:
```
df.info()
```

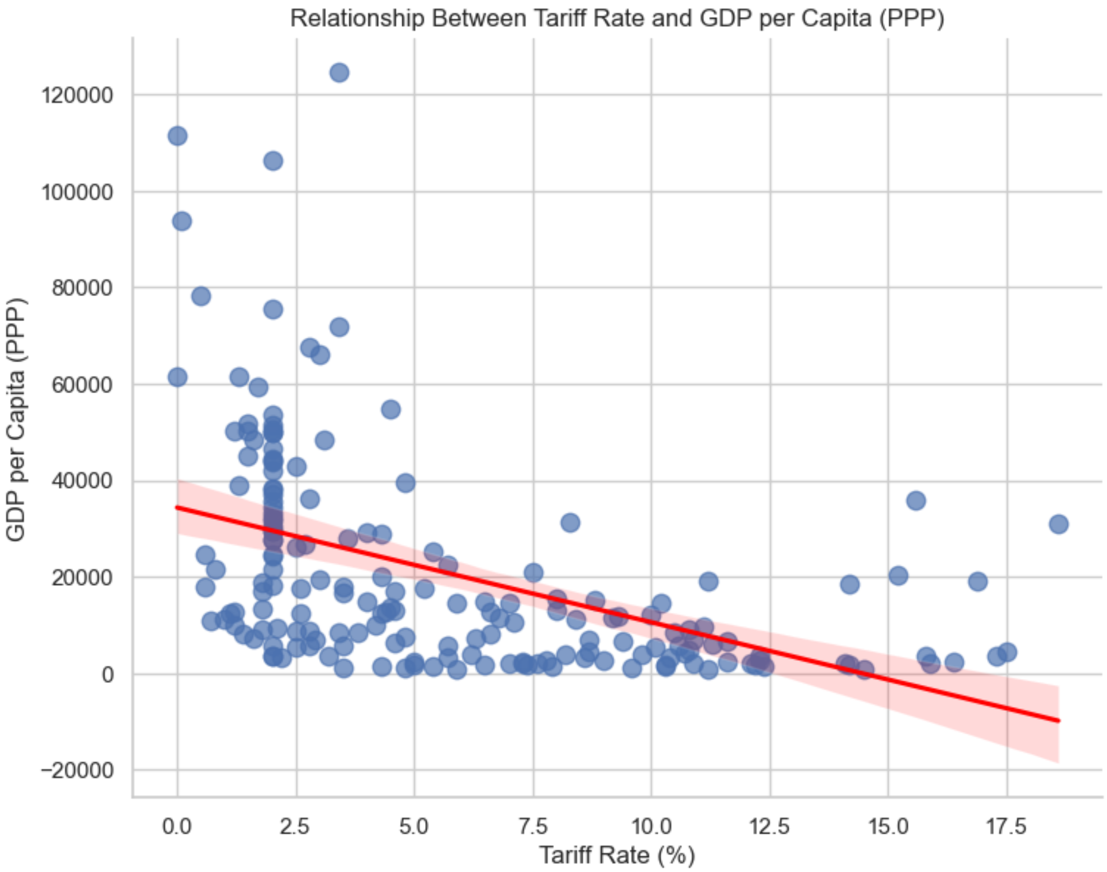
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 186 entries, 0 to 185
Data columns (total 35 columns):
 #   Column                       Non-Null Count  Dtype
---  ------                       --------------  -----
 0   CountryID                    186 non-null    float64
 1   Country Name                 0 non-null      float64
 2   WEBNAME                      0 non-null      float64
 3   Region                       0 non-null      float64
 4   World Rank                   180 non-null    float64
 5   Region Rank                  180 non-null    float64
 6   2019 Score                   180 non-null    float64
 7   Property Rights              185 non-null    float64
 8   Judical Effectiveness        185 non-null    float64
 9   Government Integrity         185 non-null    float64
 10  Tax Burden                   180 non-null    float64
 11  Gov't Spending               183 non-null    float64
 12  Fiscal Health                183 non-null    float64
 13  Business Freedom             185 non-null    float64
 14  Labor Freedom                184 non-null    float64
 15  Monetary Freedom             184 non-null    float64
 16  Trade Freedom                182 non-null    float64
 17  Investment Freedom           184 non-null    float64
 18  Financial Freedom            181 non-null    float64
 19  Tariff Rate (%)              182 non-null    float64
 20  Income Tax Rate (%)          183 non-null    float64
 21  Corporate Tax Rate (%)       183 non-null    float64
 22  Tax Burden % of GDP          179 non-null    float64
 23  Gov't Expenditure % of GDP   182 non-null    float64
 24  Country                      0 non-null      float64
 25  Population (Millions)        185 non-null    float64
 26  GDP (Billions, PPP)          183 non-null    float64
 27  GDP Growth Rate (%)          184 non-null    float64
 28  5 Year GDP Growth Rate (%)   183 non-null    float64
 29  GDP per Capita (PPP)         182 non-null    float64
 30  Unemployment (%)             180 non-null    float64
 31  Inflation (%)                182 non-null    float64
 32  FDI Inflow (Millions)        181 non-null    float64
 33  Public Debt (% of GDP)       182 non-null    float64
 34  TariffGroup                  182 non-null    category
dtypes: category(1), float64(34)
memory usage: 49.8 KB
```

[117]:

```
sns.set(style="whitegrid")

plot_df = df.dropna(subset=["Tariff Rate (%)", "GDP per Capita (PPP)"])

sns.scatterplot(
    data=plot_df,
    x="Tariff Rate (%)",
    y="GDP per Capita (PPP)",
    hue="TariffGroup",
    s=80,
    alpha=0.8
)

plt.title("Tariffs and GDP per Capita (PPP)")
plt.xlabel("Tariff Rate (%)")
plt.ylabel("GDP per Capita (PPP)")
plt.show()
```

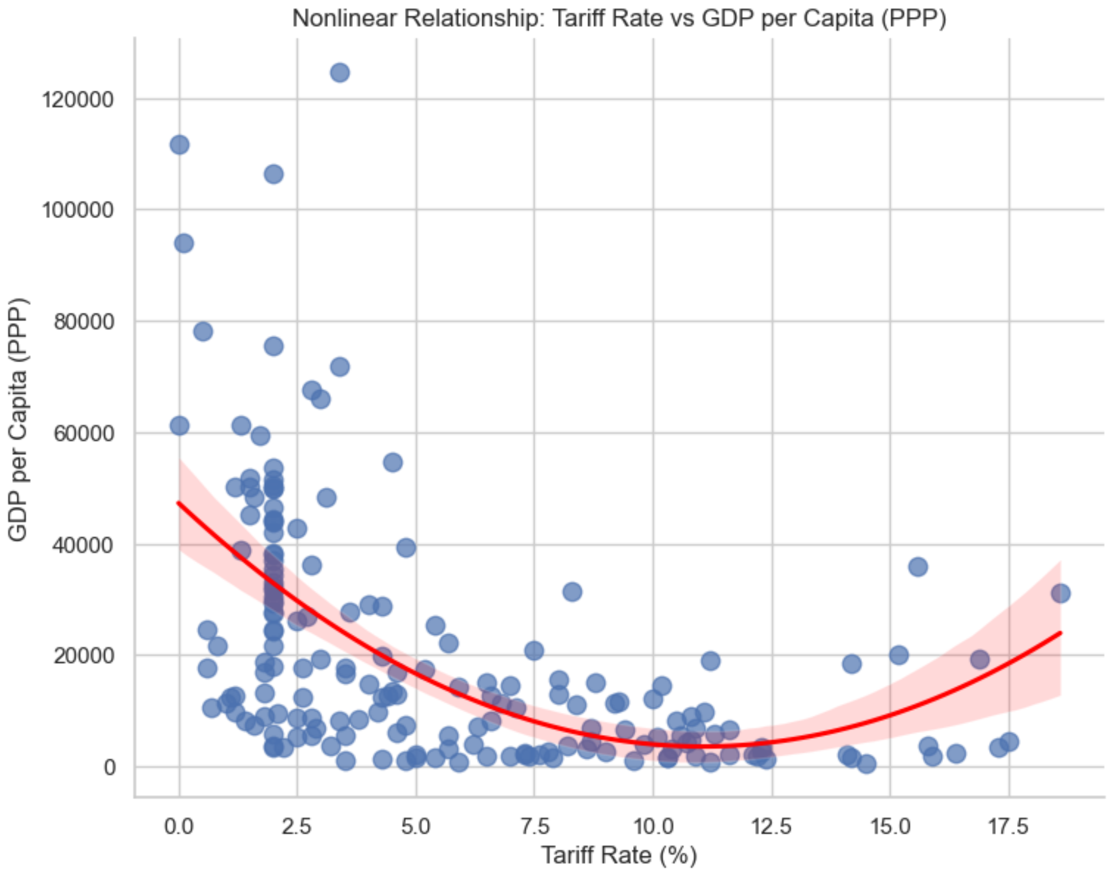


```
[ ]: sns.set(style="whitegrid")

     sns.lmplot(
         x="Tariff Rate (%)",
         y="GDP per Capita (PPP)",
         data=df,
         height=6,
         aspect=1.3,
         scatter_kws={"s": 80, "alpha": 0.7},
         line_kws={"color": "red"}
     )

     plt.title("Relationship Between Tariff Rate and GDP per Capita (PPP)")
     plt.xlabel("Tariff Rate (%)")
     plt.ylabel("GDP per Capita (PPP)")
     plt.show()
```



[120]:

```
sns.lmplot(
    x="Tariff Rate (%)",
    y="GDP per Capita (PPP)",
    data=df,
    order=2,                # quadratic
    height=6,
    aspect=1.3,
    scatter_kws={"s": 80, "alpha": 0.7},
    line_kws={"color": "red"}
)
plt.title("Nonlinear Relationship: Tariff Rate vs GDP per Capita (PPP)")
plt.show()
```



```
[123]: df_subset = df[["CountryID", "Trade Freedom", "Tariff Rate (%)"]].dropna()
       df_subset.head(10)
```

```
[123]:    CountryID  Trade Freedom  Tariff Rate (%)
       0        1.0           66.0              7.0
       1        2.0           87.8              1.1
       2        3.0           67.4              8.8
       3        4.0           61.2              9.4
       4        5.0           70.0              7.5
       5        6.0           80.8              2.1
       6        7.0           87.6              1.2
       7        8.0           86.0              2.0
       8        9.0           74.6              5.2
       9       10.0           47.8             18.6
```

```
[ ]: from sklearn.linear_model import LinearRegression

     clean = df[["Tariff Rate (%)", "Trade Freedom"]].dropna()
```

```
X = clean[["Tariff Rate (%)"]]      # 2D
y = clean["Trade Freedom"]          # 1D

print(clean.shape)  # just to see how many rows are left

model = LinearRegression()
model.fit(X, y)
```

```
(182, 2)
```

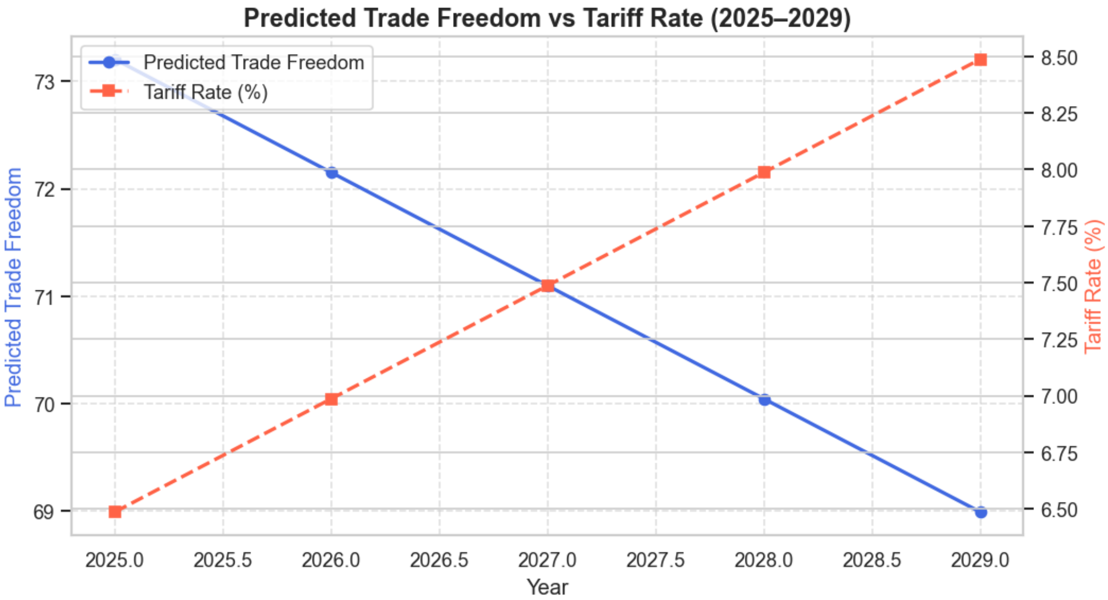
```
[ ]: LinearRegression()
```

```
[127]: start_tariff = clean["Tariff Rate (%)"].mean()

future_tariffs = np.array([start_tariff + 0.5*i for i in range(1, 6)]).
 ↪reshape(-1, 1)

future_trade = model.predict(future_tariffs)

future_years = list(range(2025, 2030))

for year, t, tf in zip(future_years, future_tariffs.flatten(), future_trade):
    print(f"Year {year}: Tariff ~ {t:.2f}%, Predicted Trade Freedom ~ {tf:.2f}")
```

```
Year 2025: Tariff ~ 6.49%, Predicted Trade Freedom ~ 73.21
Year 2026: Tariff ~ 6.99%, Predicted Trade Freedom ~ 72.15
Year 2027: Tariff ~ 7.49%, Predicted Trade Freedom ~ 71.10
Year 2028: Tariff ~ 7.99%, Predicted Trade Freedom ~ 70.04
Year 2029: Tariff ~ 8.49%, Predicted Trade Freedom ~ 68.99
```

```
[130]: plt.figure(figsize=(9, 5), dpi=120)

ax1 = plt.gca()
ax2 = ax1.twinx()


ax1.plot(future_years, future_trade, marker='o', color='royalblue',␣
 ↪linewidth=2, label="Predicted Trade Freedom")


ax2.plot(future_years, future_tariffs.flatten(), marker='s', linestyle='--',␣
 ↪color='tomato', linewidth=2, label="Tariff Rate (%)")


ax1.set_xlabel("Year", fontsize=12)
ax1.set_ylabel("Predicted Trade Freedom", color="royalblue", fontsize=12)
ax2.set_ylabel("Tariff Rate (%)", color="tomato", fontsize=12)
```

```
plt.title("Predicted Trade Freedom vs Tariff Rate (2025-2029)", fontsize=14,␣
 ↪weight="bold")


lines, labels = ax1.get_legend_handles_labels()
lines2, labels2 = ax2.get_legend_handles_labels()
plt.legend(lines + lines2, labels + labels2, loc="upper left")


ax1.grid(True, linestyle="--", alpha=0.6)
plt.tight_layout()
plt.show()
```



[132]:

```
from sklearn.metrics import r2_score, mean_absolute_error, mean_squared_error
from sklearn.model_selection import train_test_split
```

[133]:

```
clean = df[["Tariff Rate (%)", "Trade Freedom"]].dropna()
```

[134]:

```
X = clean[["Tariff Rate (%)"]]
y = clean["Trade Freedom"]
```

[135]:

```
X_train , X_test, y_train, y_test = train_test_split(X, y, test_size=0.2,␣
 ↪random_state=42)
model = LinearRegression()
model.fit(X_train, y_train)
```

[135]:

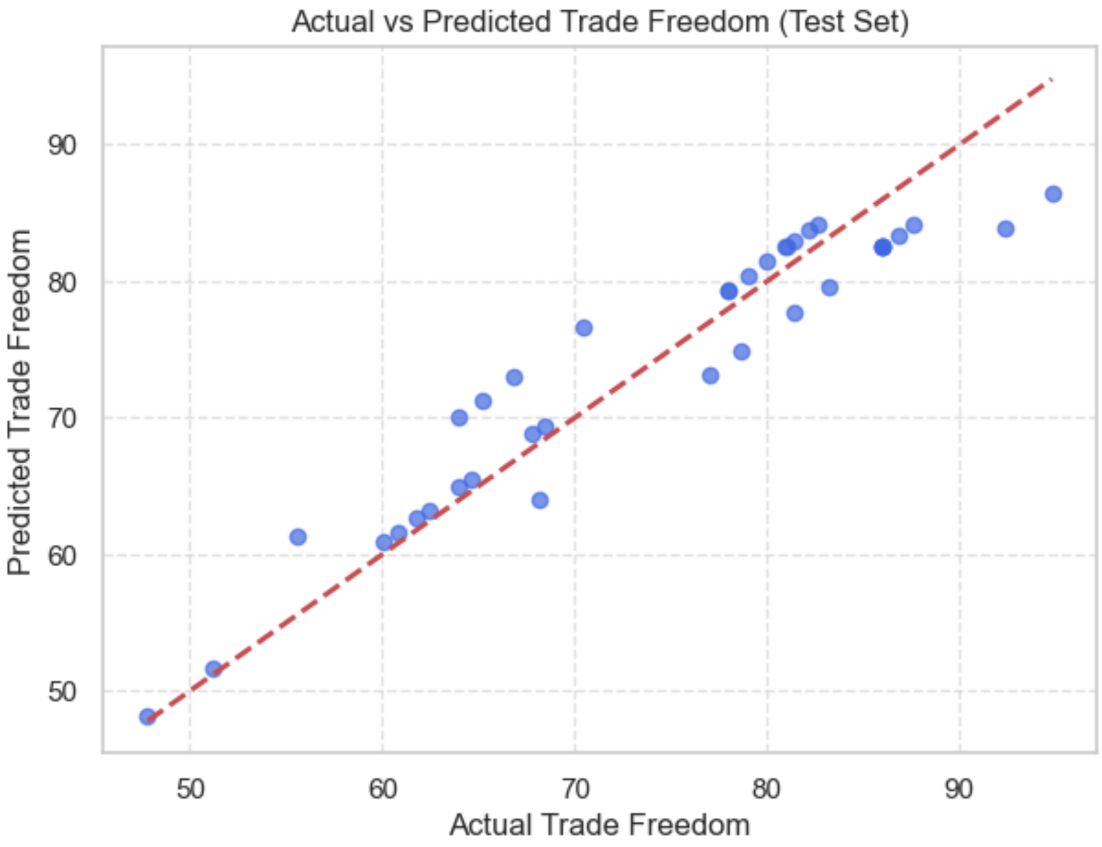
```
LinearRegression()
```

[136]:
```
y_pred = model.predict(X_test)
```

[137]:
```
r2 = r2_score(y_test, y_pred)
mae = mean_absolute_error(y_test, y_pred)
rmse = np.sqrt(mean_squared_error(y_test, y_pred))

print(f"R² Score: {r2:.3f}")
print(f"Mean Absolute Error: {mae:.3f}")
print(f"Root Mean Squared Error: {rmse:.3f}")
```

```
R² Score: 0.900
Mean Absolute Error: 2.933
Root Mean Squared Error: 3.675
```

[138]:
```
plt.figure(figsize=(7,5))
plt.scatter(y_test, y_pred, color='royalblue', alpha=0.7)
plt.plot([y_test.min(), y_test.max()], [y_test.min(), y_test.max()], 'r--',␣
 ↪lw=2)
plt.xlabel("Actual Trade Freedom")
plt.ylabel("Predicted Trade Freedom")
plt.title("Actual vs Predicted Trade Freedom (Test Set)")
plt.grid(True, linestyle="--", alpha=0.6)
plt.show()
```


Actual vs Predicted Trade Freedom (Test Set)
Predicted Trade Freedom
Actual Trade Freedom