Data Visualization in Stata (II.)

Data Exploratory Analysis.

Tomáš Oleš
Department of Economic Policy
Faculty of Economics and Finance

January 29, 2025

Agenda

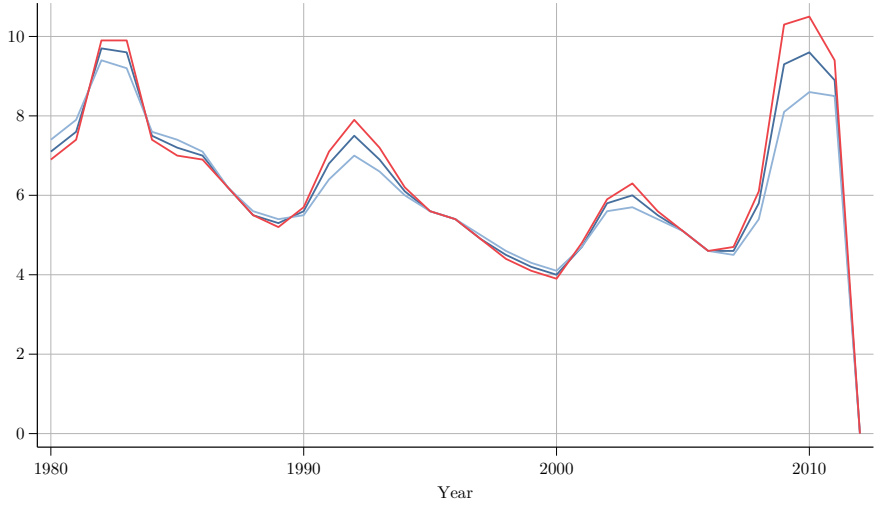
- Introduction to data exploratory analysis.
- Gain hands-on experience in creating advanced visualizations using STATA, including line plots, histograms, scatter plots, and box plots.

What Was the Key Principle of Visualization?

What Was the Key Principle of Visualization?

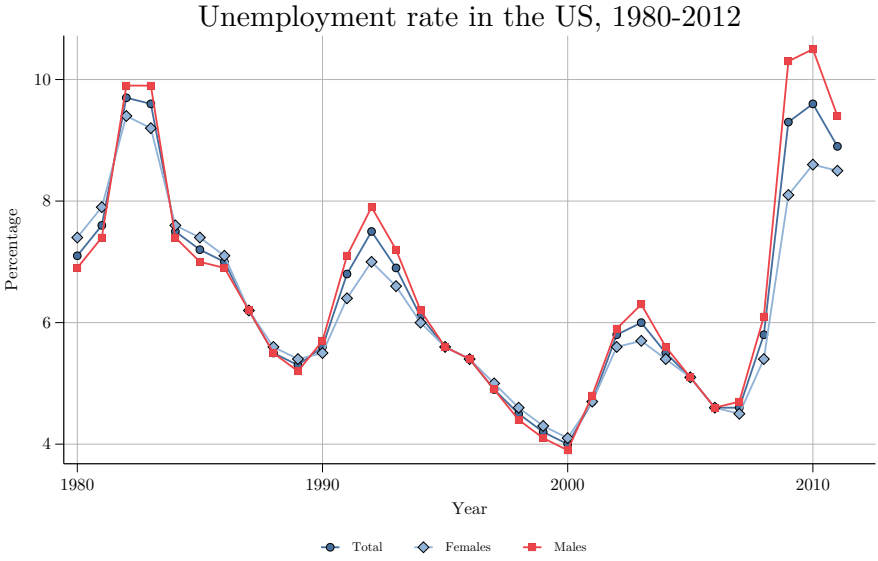
- "Visualization gives you answers to questions you didn't know you had." – Ben Schneiderman

Line Plots

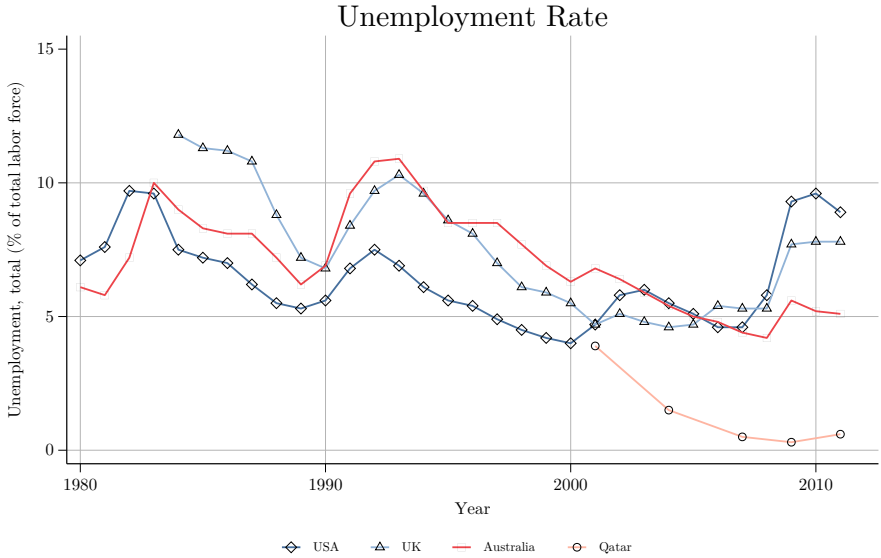

10
8
6
4
2
0
1980
1990
2000
2010
Year
Unemployment, total (% of total labor force)
Unemployment, male (% of male labor force)
Unemployment, female (% of female labor force)

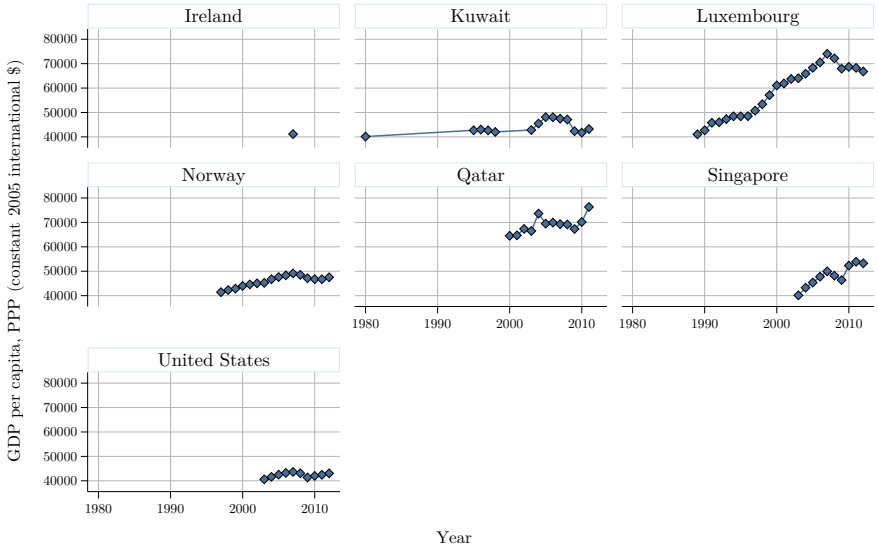
10
8
6
4
1980
1990
2000
2010
Year
Unemployment, total (% of total labor force)
Unemployment, male (% of male labor force)
Unemployment, female (% of female labor force)

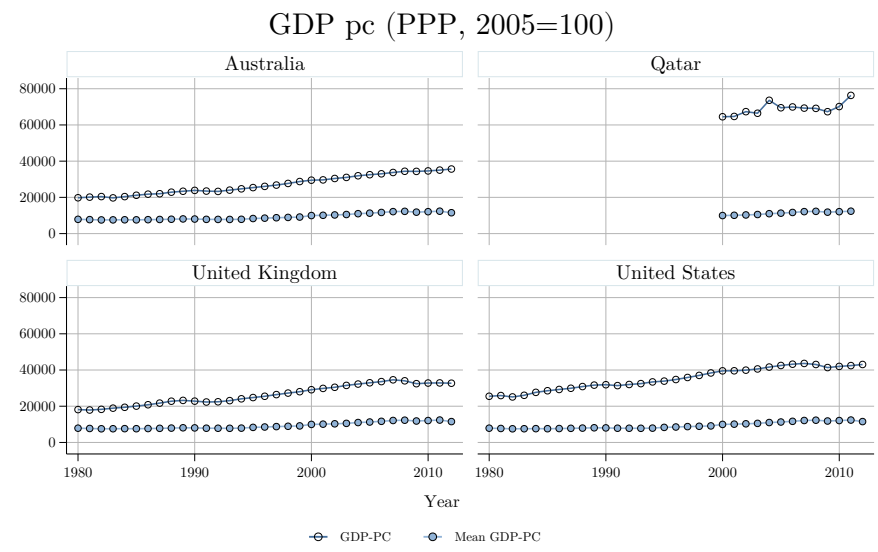
Unemployment rate in the US, 1980-2012
Percentage
10
8
6
4
1980
1990
2000
2010
Year
Total
Females
Males

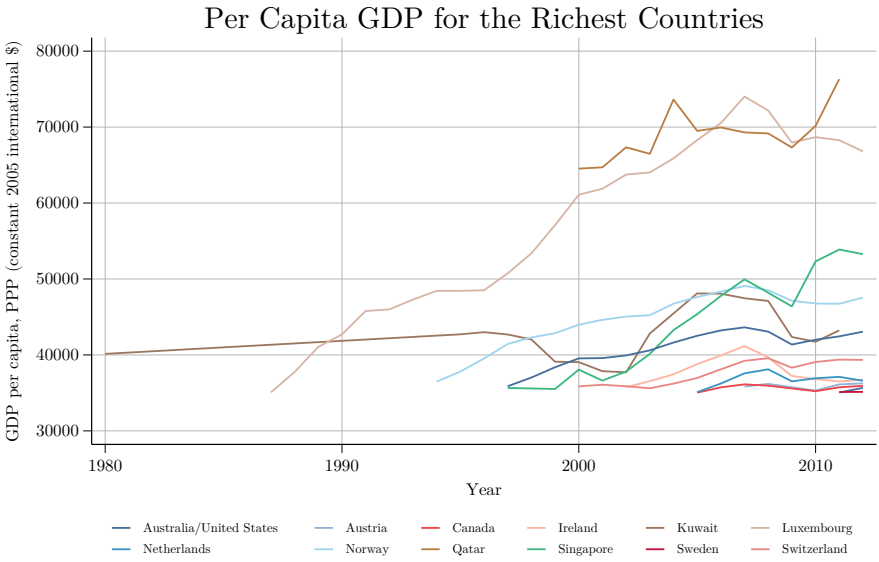

Unemployment rate in the US, 1980-2012
Percentage
10
8
6
4
1980
1990
2000
2010
Year
Total
Females
Males

Unemployment Rate
Australia
Qatar
United Kingdom
United States
Unemployment, total (% of total labor force)
0
5
10
15
1980
1990
2000
2010
Year
Graphs by Country Name

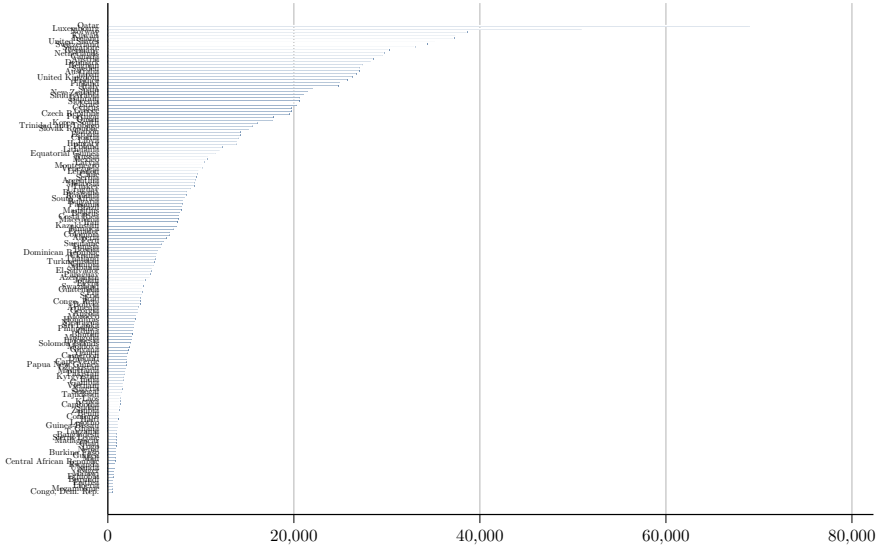

Unemployment Rate
Unemployment, total (% of total labor force)
15
10
5
0
1980
1990
2000
2010
Year
USA
UK
Australia
Qatar

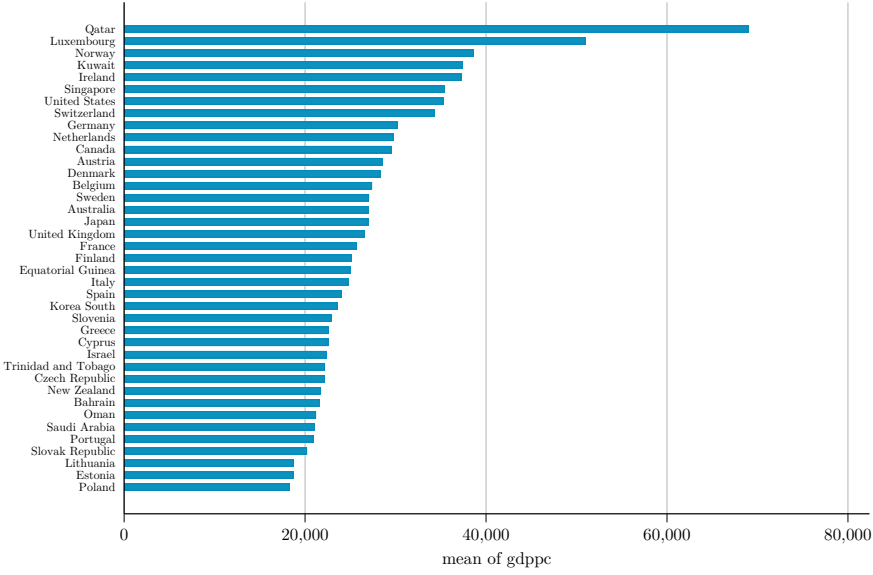

Ireland
Kuwait
Luxembourg
Norway
Qatar
Singapore
United States
GDP per capita, PPP (constant 2005 international $)
80000
70000
60000
50000
40000
1980
1990
2000
2010
Year
Graphs by Country Name

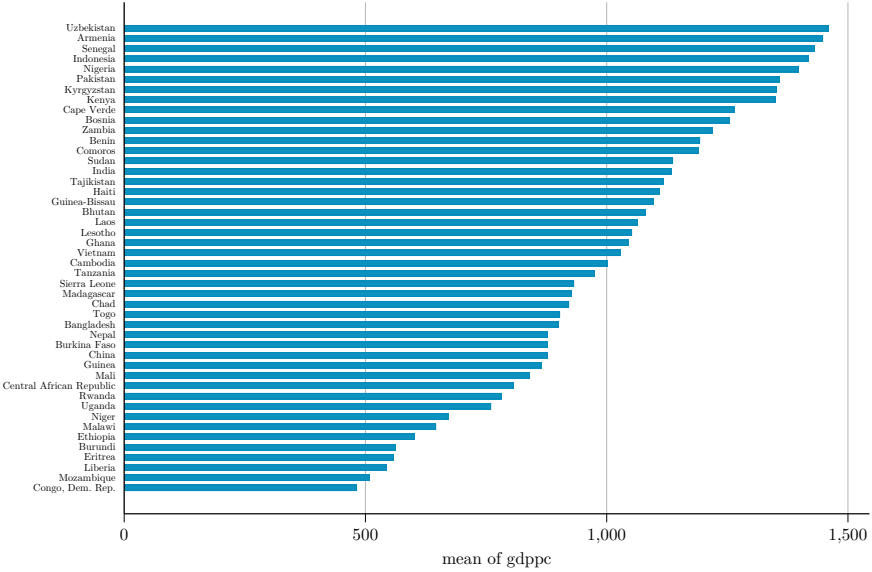

GDP pc (PPP, 2005=100)
Australia
Qatar
United Kingdom
United States
80000
60000
40000
20000
0
1980
1990
2000
2010
Year
GDP-PC
Mean GDP-PC
Graphs by Country Name

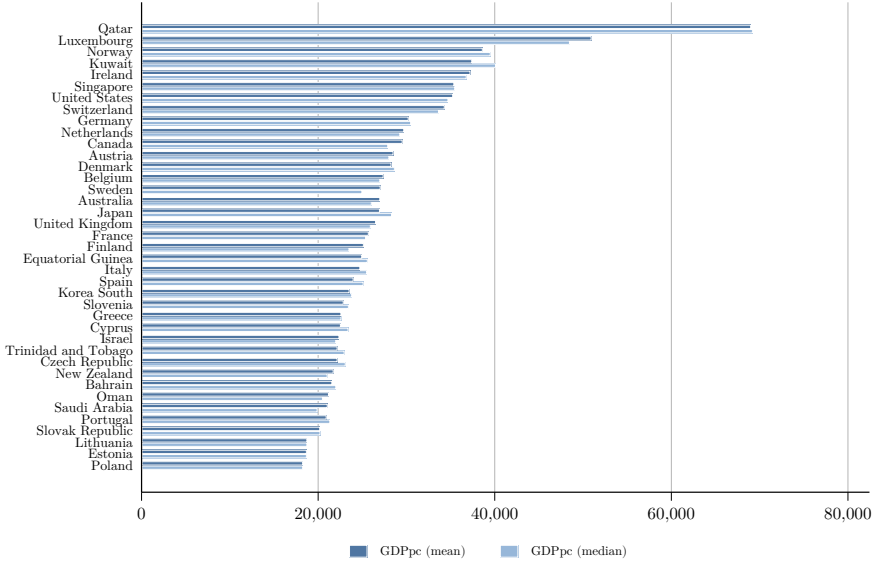

Per Capita GDP for the Richest Countries
GDP per capita, PPP (constant 2005 international $)
80000
70000
60000
50000
40000
30000
1980
1990
2000
2010
Year
Australia/United States
Austria
Canada
Ireland
Kuwait
Luxembourg
Netherlands
Norway
Qatar
Singapore
Sweden
Switzerland

Bar Plots

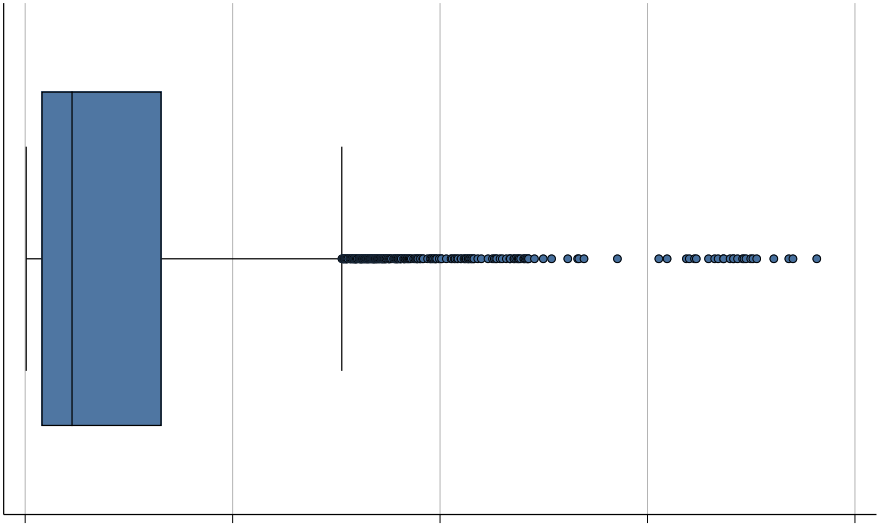


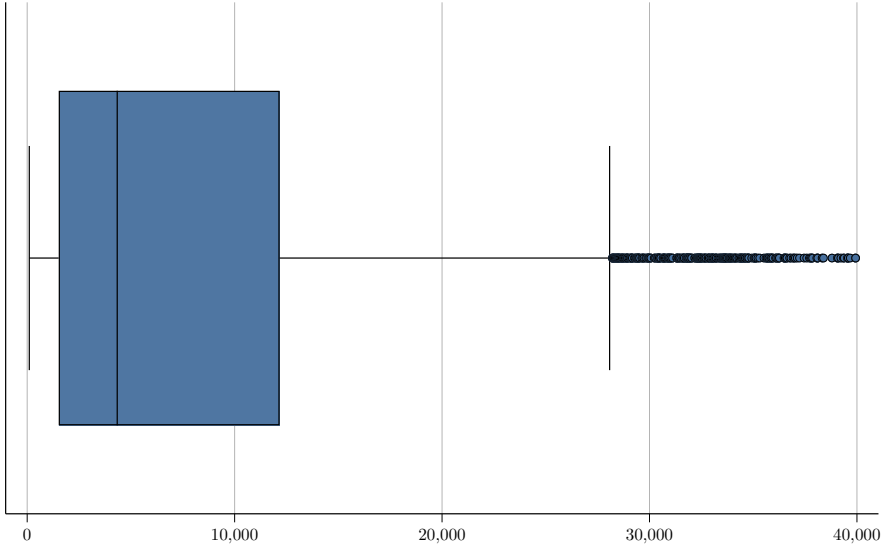

Qatar
Luxembourg
Norway
Kuwait
Ireland
Singapore
United States
Switzerland
Germany
Netherlands
Canada
Austria
Denmark
Belgium
Sweden
Australia
Japan
United Kingdom
France
Finland
Equatorial Guinea
Italy
Spain
Korea South
Slovenia
Greece
Cyprus
Israel
Trinidad and Tobago
Czech Republic
New Zealand
Bahrain
Oman
Saudi Arabia
Portugal
Slovak Republic
Lithuania
Estonia
Poland
0
20,000
40,000
60,000
80,000
mean of gdppc

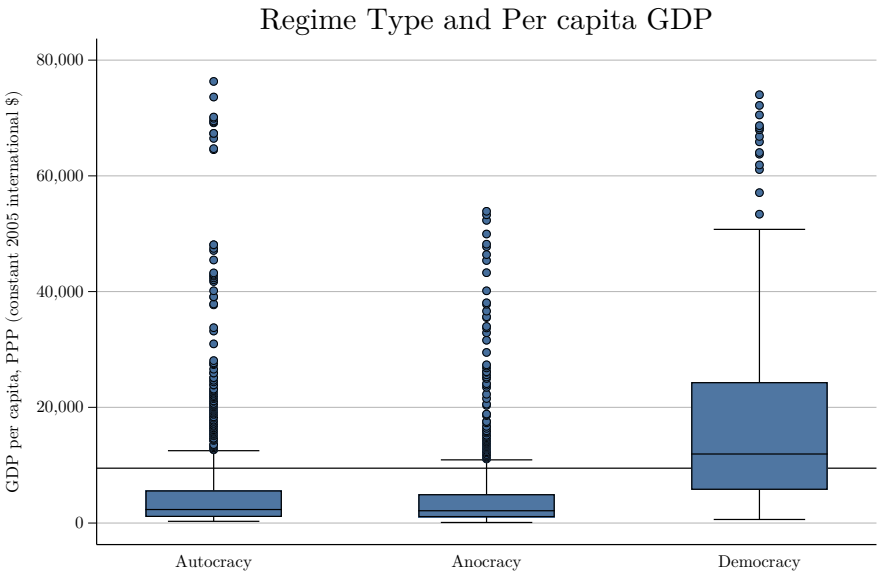

Uzbekistan
Armenia
Senegal
Indonesia
Nigeria
Pakistan
Kyrgyzstan
Kenya
Cape Verde
Bosnia
Zambia
Benin
Comoros
Sudan
India
Tajikistan
Haiti
Guinea-Bissau
Bhutan
Laos
Lesotho
Ghana
Vietnam
Cambodia
Tanzania
Sierra Leone
Madagascar
Chad
Togo
Bangladesh
Nepal
Burkina Faso
China
Guinea
Mali
Central African Republic
Rwanda
Uganda
Niger
Malawi
Ethiopia
Burundi
Eritrea
Liberia
Mozambique
Congo, Dem. Rep.
0
500
1,000
1,500
mean of gdppc

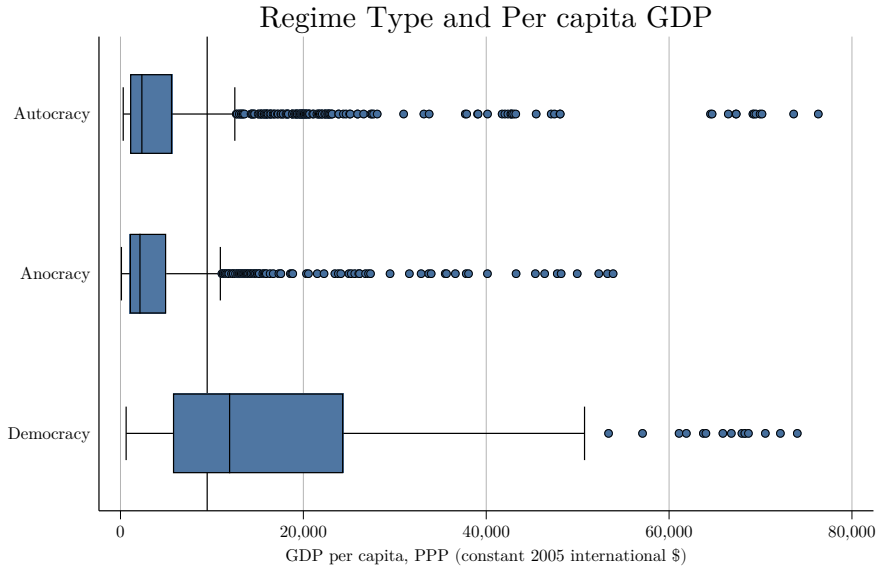

Qatar
Luxembourg
Norway
Kuwait
Ireland
Singapore
United States
Switzerland
Germany
Netherlands
Canada
Austria
Denmark
Belgium
Sweden
Australia
Japan
United Kingdom
France
Finland
Equatorial Guinea
Italy
Spain
Korea South
Slovenia
Greece
Cyprus
Israel
Trinidad and Tobago
Czech Republic
New Zealand
Bahrain
Oman
Saudi Arabia
Portugal
Slovak Republic
Lithuania
Estonia
Poland
0
20,000
40,000
60,000
80,000
GDPpc (mean)
GDPpc (median)

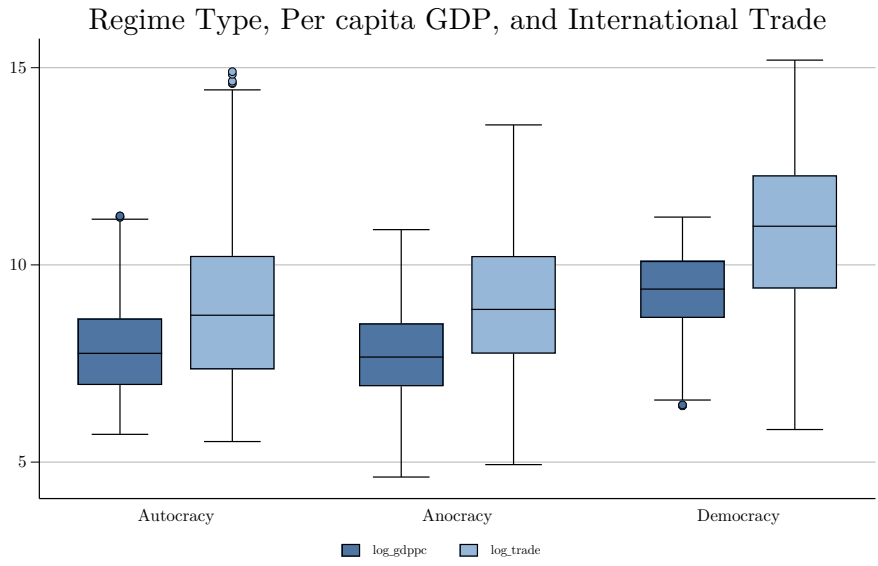
Box Plots


0
20,000
40,000
60,000
80,000
GDP per capita, PPP (constant 2005 international $)

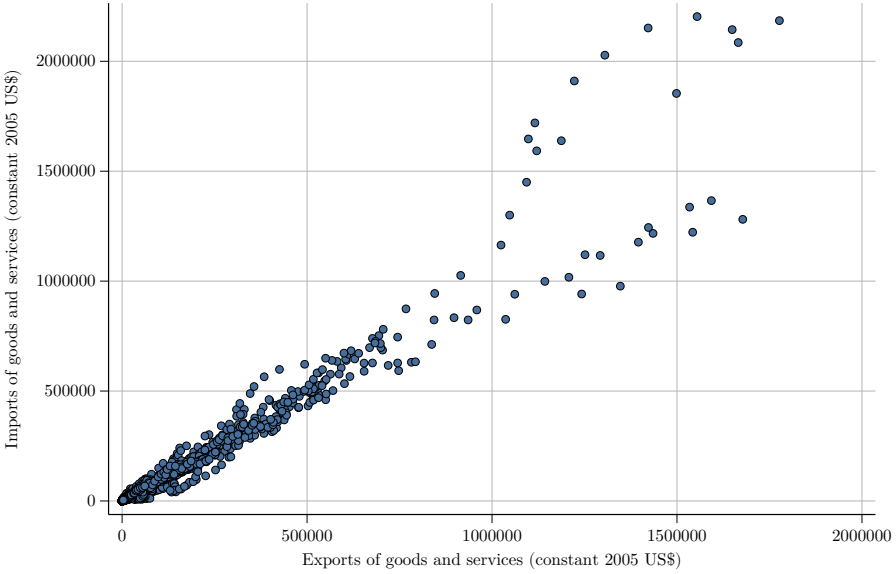

0
10,000
20,000
30,000
40,000
GDP per capita, PPP (constant 2005 international $)

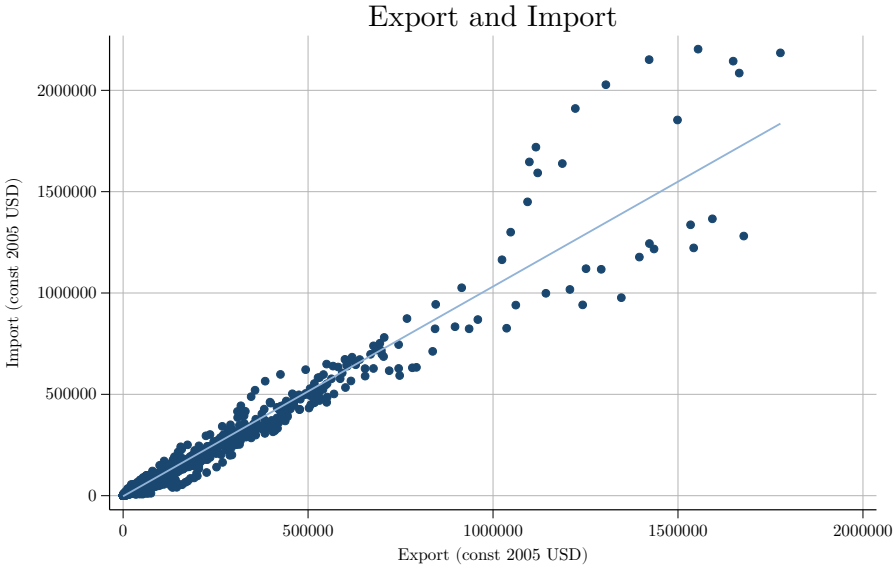

Regime Type and Per capita GDP
GDP per capita, PPP (constant 2005 international $)
80,000
60,000
40,000
20,000
0
Autocracy
Anocracy
Democracy

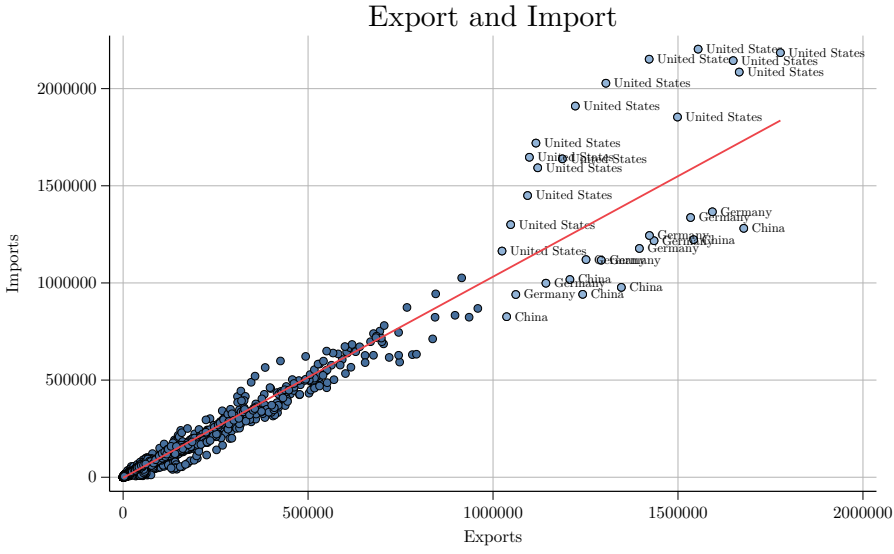

Regime Type and Per capita GDP
Autocracy
Anocracy
Democracy
0
20,000
40,000
60,000
80,000
GDP per capita, PPP (constant 2005 international $)


Regime Type, Per capita GDP, and International Trade
15
10
5
Autocracy
Anocracy
Democracy
log_gdppc
log_trade

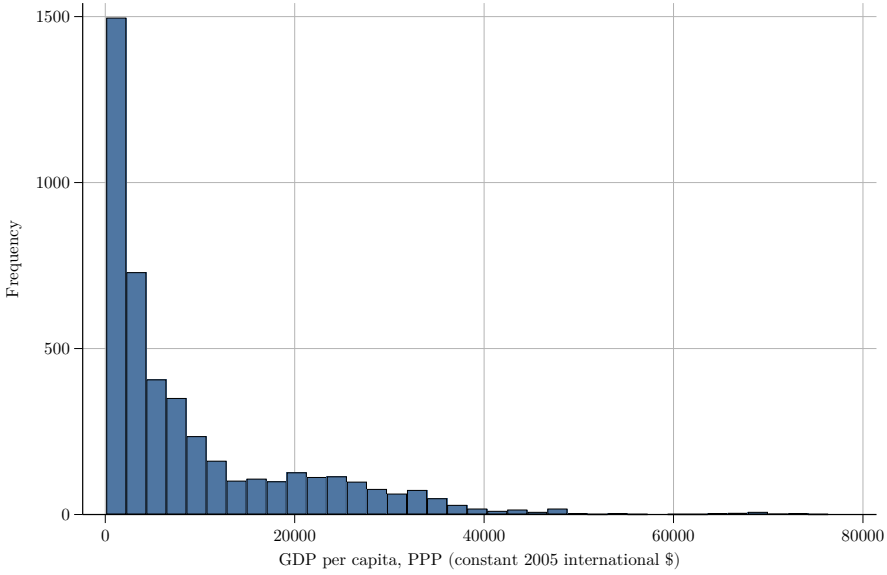
Scatter Plots

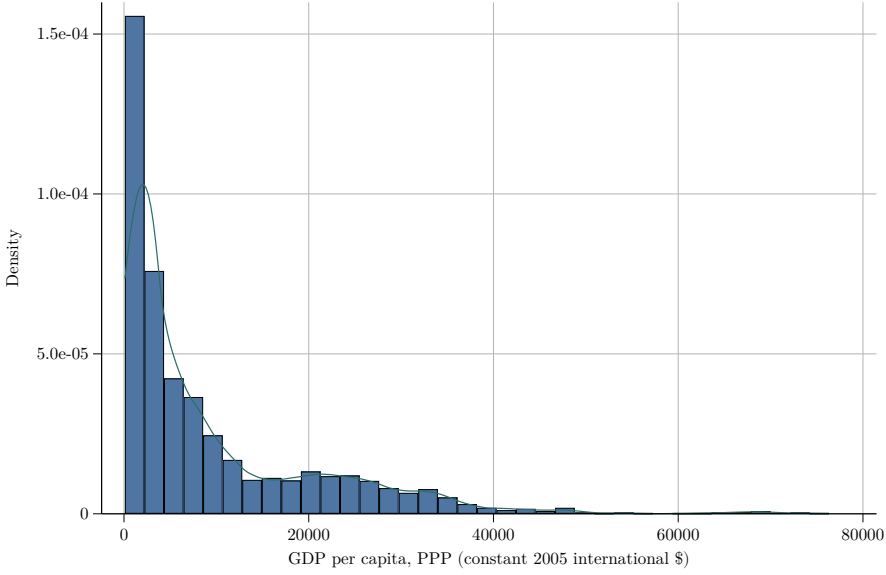


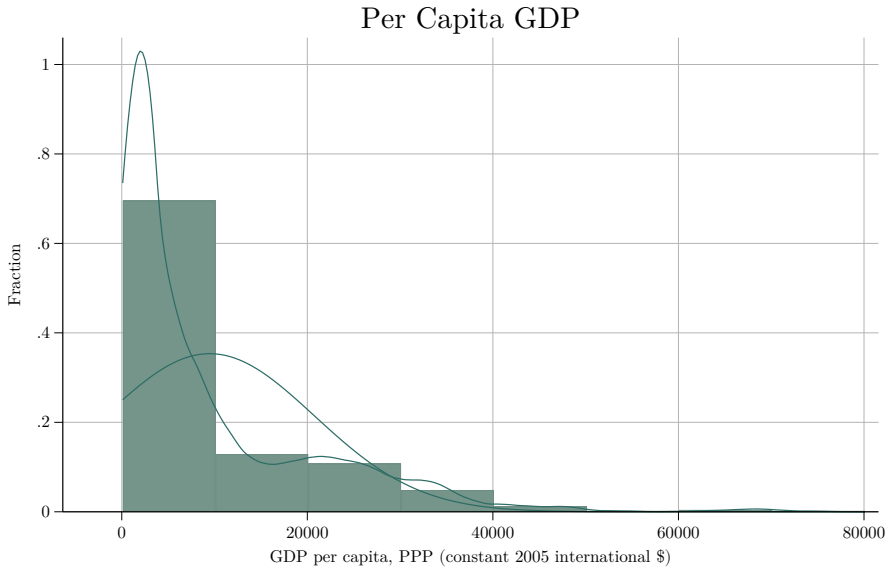

Export and Import
Import (const 2005 USD)
Export (const 2005 USD)
0
500000
1000000
1500000
2000000

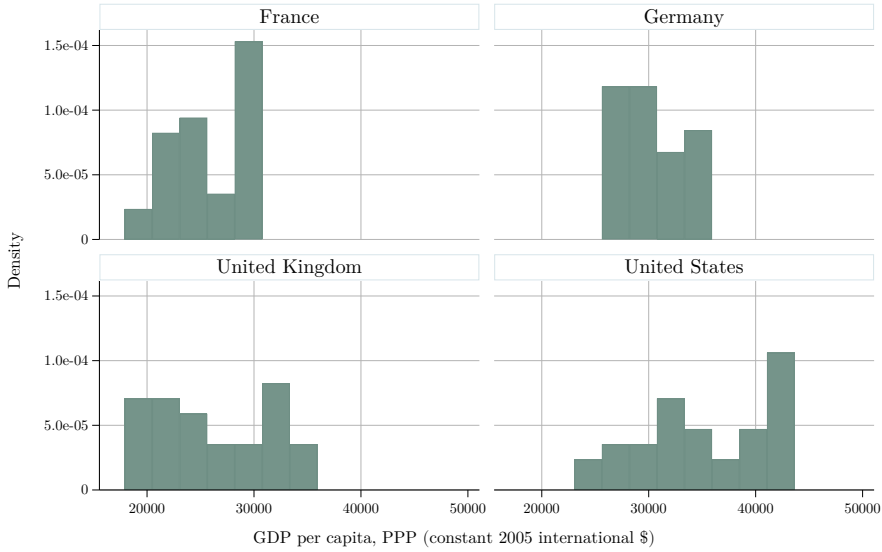

Export and Import
Imports
Exports
0
500000
1000000
1500000
2000000
United States
Germany
China
Constant values, 2005, millions US$

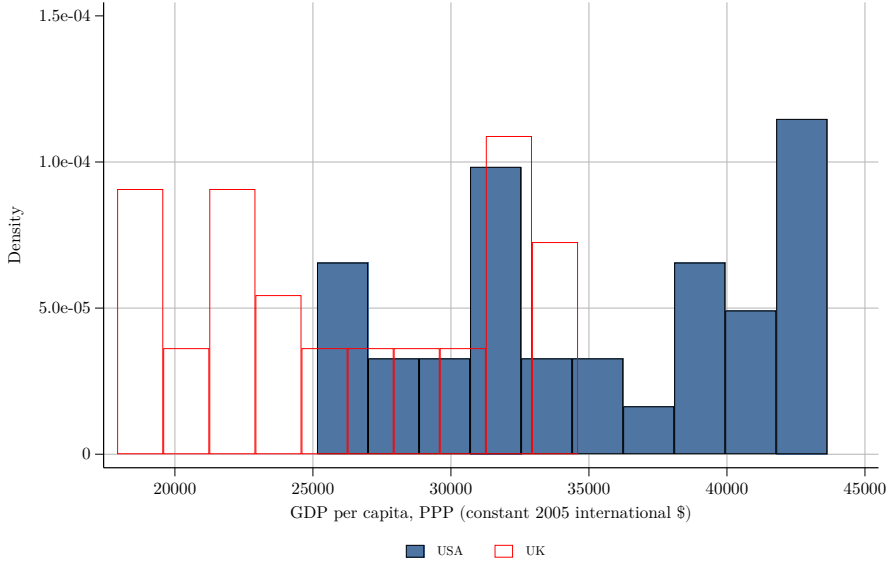
Histograms


1500
1000
500
0
Frequency
0
20000
40000
60000
80000
GDP per capita, PPP (constant 2005 international $)


1.5e-04
1.0e-04
5.0e-05
0
Density
0
20000
40000
60000
80000
GDP per capita, PPP (constant 2005 international $)


Per Capita GDP
Fraction
1
.8
.6
.4
.2
0
0
20000
40000
60000
80000
GDP per capita, PPP (constant 2005 international $)


France
Germany
United Kingdom
United States
Density
1.5e-04
1.0e-04
5.0e-05
0
20000
30000
40000
50000
GDP per capita, PPP (constant 2005 international $)
Graphs by Country Name


1.5e-04
1.0e-04
5.0e-05
0
Density
20000
25000
30000
35000
40000
45000
GDP per capita, PPP (constant 2005 international $)
USA
UK