

Impactos da Inovação e da Liberdade Econômica na Renda Per Capita: Uma Análise Regressiva

A amostra está bastante diversa, contando com a seguinte quantidade de países por continente: América: 21, Europa: 44, Oceania: 2, África: 30 e Ásia: 32 países.

Lembrando: os dados de Renda per capita aqui utilizados são o PIB per Capita por Paridade do Poder de Compra.

o PIB per Capita PPC calcula o "valor de todos os finais de bens e serviços produzidos no âmbito de uma nação em determinado ano dividido pela população média para o mesmo ano."

- Coreia do Norte: A última estimativa desse dado para esse país é de 2015. Considerando que ele vem tendo resultados econômicos péssimos no últimos anos, preferi utilizar desse dado estimado em 2015 mesmo.

Bibliotecas

```
import pandas as pd
import gspread
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns
import numpy as np
import statsmodels.api as sm
```

#Acessar a planilha com os dados dos países

```
url =
"https://docs.google.com/spreadsheets/d/1pKewi3DlY2hadaFdPWB8KSfu7MSK6XFA/export?format=csv"
dados = pd.read_csv(url)
```

dados

```
{
  "summary": {
    "name": "dados",
    "rows": 129,
    "fields": [
      {
        "column": "country",
        "properties": {
          "dtype": "string",
          "num_unique_values": 129,
          "samples": [
            "Mauritius",
            "Greece",
            "Luxembourg"
          ],
          "semantic_type": "",
          "description": ""
        },
        "column": "xI",
        "properties": {
          "dtype": "number",
          "std": 14.108839798690616,
          "min": 12.5,
          "max": 67.6,
          "num_unique_values": 111,
          "samples": [
            30.4,
            56.0,
            61.5
          ],
          "semantic_type": "",
          "description": ""
        },
        "column": "xL",
        "properties": {
          "dtype": "number",
          "std": 9.546156297088876,
          "min": 38.2,
          "max": 83.5,
          "num_unique_values": 109,
          "samples": [
            62.9,
            62.5,
            83.5
          ],
          "semantic_type": "",
          "description": ""
        },
        "column": "Y",
        "properties": {
          "dtype": "number",
          "std": 28960,
          "min": 836,
          "max": 142214,
          "num_unique_values": 129,
          "samples": [
            26906,
            36835,
            142214
          ],
          "semantic_type": "",
          "description": ""
        },
        "column": "continente",
        "properties": {
          "dtype": "category",
          "num_unique_values": 8,
          "samples": [
            "Am\u00e9rica do Norte",
            "Africa",
            "Europa"
          ],
          "semantic_type": "",
          "description": ""
        }
      ]
    },
    "type": "dataframe",
    "variable_name": "dados"
  }
}
```

#Ajeitando os dados para evitar problemas:

#Converter a coluna Y, renda per capita, para num\u00e9rico

```
dados['Y'] = pd.to_numeric(dados['Y'], errors='coerce').round()
```

Remover a coluna Unnamed, se existir

```
dados = dados.loc[:, ~dados.columns.str.contains('^Unnamed')]
```

quero somar quantos pa\u00edses eu tenho na amostra por continente:

```
paises_por_continente = dados.groupby('continente')['country'].count()
print(paises_por_continente)
```

```
continente
Am\u00e9rica Central      8
Am\u00e9rica do Central    1
Am\u00e9rica do Norte      3
```

```
América do Sul      9
Europa              44
Oceania             2
África             30
Ásia                32
Name: country, dtype: int64
```

Visualizar os dados

Inovação vs Renda por país

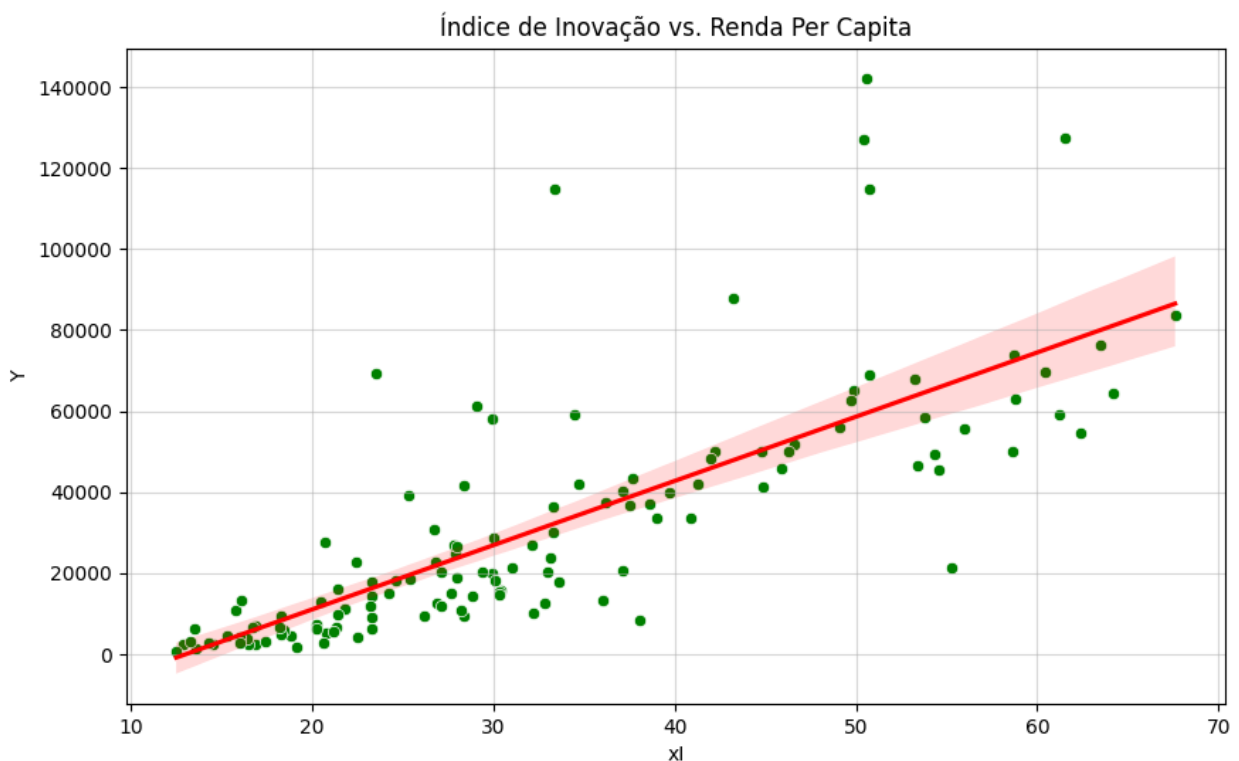
```
# Visualizar xI vs Y em um gráfico de dispersão feito no seaborn
plt.figure(figsize=(10, 6))

sns.scatterplot(x='xI', y='Y', data=dados, color = 'green')

plt.title('Índice de Inovação vs. Renda Per Capita')
plt.xlabel('Inovação')
plt.ylabel('Renda Per Capita')
plt.grid(True, alpha = 0.5)

# Adicionando uma linha de tendência
sns.regplot(x='xI', y='Y', data=dados, scatter=False, color='red')

plt.show()
```

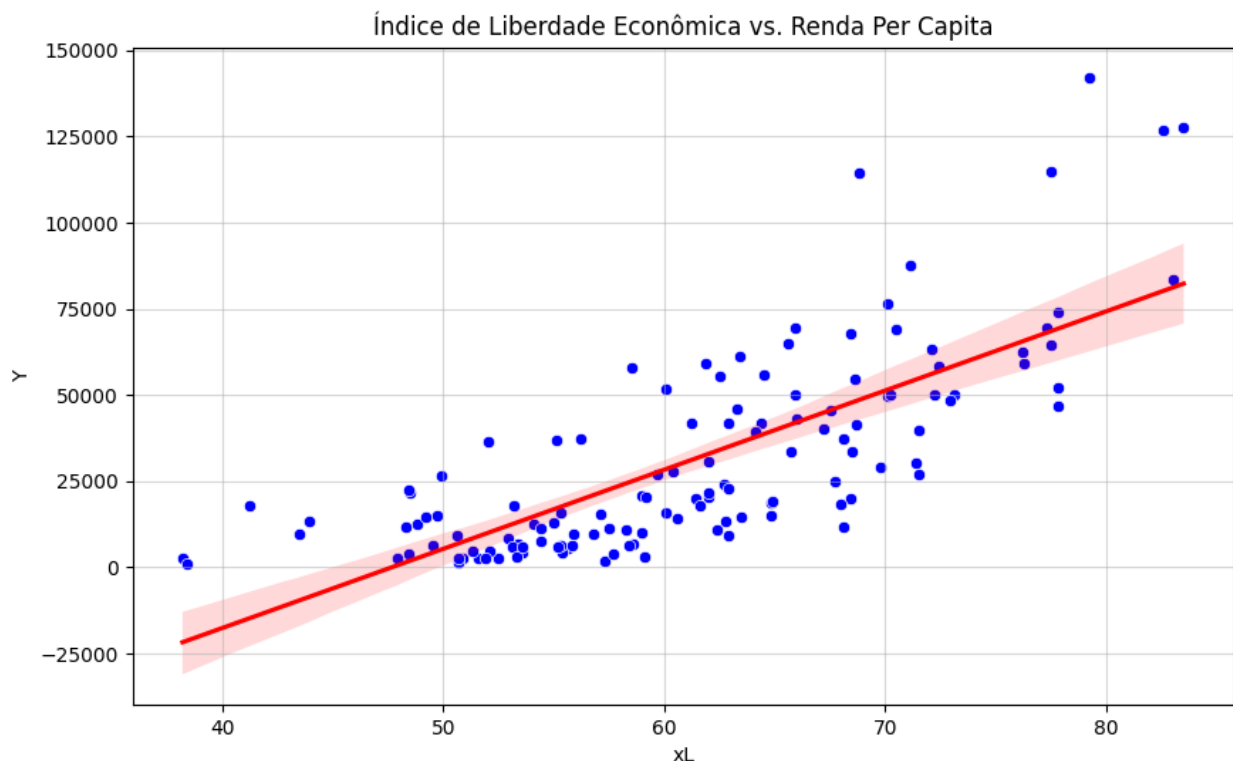


Liberdade econômica vs Renda por país

```
# Visualizar xL vs Y em um gráfico de dispersão feito no seaborn
plt.figure(figsize=(10, 6))

sns.scatterplot(x='xL', y='Y', data=dados, color = 'blue')
plt.title('Índice de Liberdade Econômica vs. Renda Per Capita')
plt.xlabel('Liberdade Econômica')
plt.ylabel('Renda Per Capita')
plt.grid(True, alpha = 0.5)
# Adicionando uma linha de tendência
sns.regplot(x='xL', y='Y', data=dados, scatter=False, color='red')

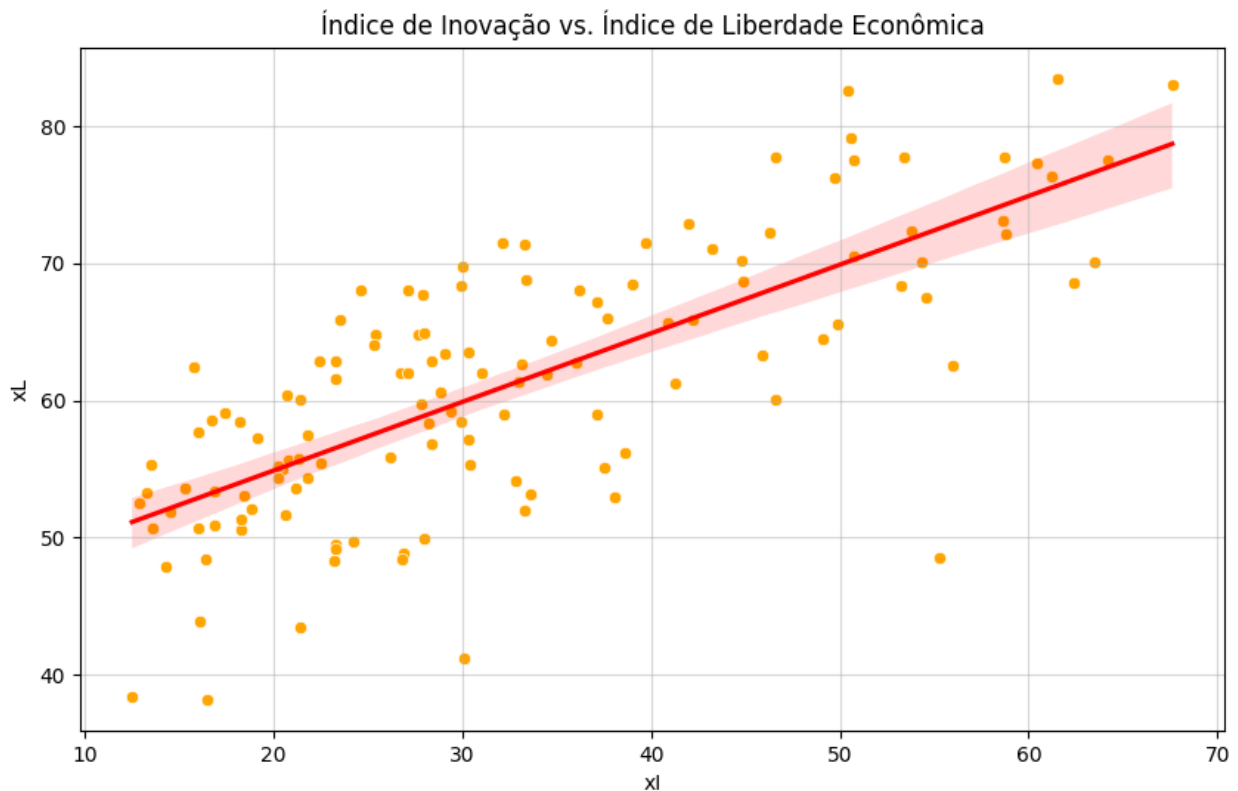
plt.show()
```



Inovação vs Liberdade econômica por país

```
# Visualizar xI vs xL em um gráfico de dispersão feito no seaborn
plt.figure(figsize=(10, 6))
sns.scatterplot(x='xI', y='xL', data=dados, color = 'orange')
plt.title('Índice de Inovação vs. Índice de Liberdade Econômica')
plt.grid(True, alpha = 0.5)
plt.xlabel('Inovação')
plt.ylabel('Liberdade Econômica')
#Adicionando uma linha de tendência
```

```
sns.regplot(x='xI', y='xL', data=dados, scatter=False, color='red')
plt.show()
```



Vou calcular algumas estatísticas interessantes com esses dados:

1. Estatísticas Descritivas

```
# 1. Estatísticas Descritivas:
estatisticas_descritivas = dados[['Y', 'xI', 'xL']].describe()
estatisticas_descritivas_por_continente = dados.groupby('continente')
[['Y', 'xI', 'xL']].describe()
print("Estatísticas Descritivas da amostra:")
print(estatisticas_descritivas)
print()
print()
print("Estatísticas Descritivas por Continente:")
print(estatisticas_descritivas_por_continente)
```

Estatísticas Descritivas da amostra:

	Y	xI	xL
count	129.000000	129.000000	129.000000
mean	31129.937984	32.655814	61.222481
std	28960.040757	14.108840	9.546156
min	836.000000	12.500000	38.200000
25%	9519.000000	21.400000	54.100000
50%	20377.000000	29.400000	61.400000
75%	48397.000000	42.000000	68.100000
max	142214.000000	67.600000	83.500000

Estatísticas Descritivas por Continente:

Y						
\		count	mean	std	min	
25%						
continente						
América Central	8.0	17416.625000	11841.800852	6741.0		
9832.25						
América do Central	1.0	22834.000000	NaN	22834.0		
22834.00						
América do Norte	3.0	52103.666667	27979.966982	21512.0		
39956.00						
América do Sul	9.0	19688.444444	7308.577959	9684.0		
15048.00						
Europa	44.0	49011.500000	28028.042289	12671.0		
31932.50						
Oceania	2.0	57296.000000	7536.344074	51967.0		
54631.50						
África	30.0	6941.533333	6049.543988	836.0		
2718.50						
Ásia	32.0	32523.250000	32306.620777	4725.0		
9495.75						
		xI				
\						...
		50%	75%	max	count	mean
75%						...
continente						...
América Central	11459.0	25636.75	39280.0	8.0	21.525000	...
25.750						
América do Central	22834.0	22834.00	22834.0	1.0	22.400000	...
22.400						
América do Norte	58400.0	67399.50	76399.0	3.0	49.433333	...
58.650						
América do Sul	17822.0	26505.00	30209.0	9.0	27.255556	...
30.000						

Europa	44547.0	60057.75	142214.0	44.0	43.240909	...
50.700						
Oceania	57296.0	59960.50	62625.0	2.0	48.150000	...
48.925						
África	4308.5	9410.00	26906.0	30.0	19.273333	...
22.325						
Ásia	16364.0	49649.25	127565.0	32.0	32.728125	...
37.350						

xL						
\	max	count	mean	std	min	25%
50%						
continente						
América Central	27.9	8.0	61.137500	5.531194	53.4	57.55
61.40						
América do Central	22.4	1.0	62.900000	NaN	62.9	62.90
62.90						
América do Norte	63.5	3.0	68.166667	5.462905	62.0	66.05
70.10						
América do Sul	33.6	9.0	58.544444	9.193219	43.5	53.20
59.20						
Europa	67.6	44.0	67.040909	8.113306	48.4	61.90
66.60						
Oceania	49.7	2.0	77.000000	1.131371	76.2	76.60
77.00						
África	32.1	30.0	53.606667	7.022080	38.2	50.70
53.45						
Ásia	61.5	32.0	59.446875	8.821454	41.2	52.70
59.00						

	75%	max
continente		
América Central	65.000	68.1
América do Central	62.900	62.9
América do Norte	71.250	72.4
América do Sul	64.800	71.4
Europa	72.125	83.0
Oceania	77.400	77.8
África	57.175	71.5
Ásia	64.050	83.5

[8 rows x 24 columns]

Maiores e menores índices de inovação da amostra:

```
#Calcular o maior valor da coluna xI (maior índice de inovação)
dados[dados['xI'] == dados['xI'].max()]
print("Dados do país com o maior índice de inovação:")
print(dados[dados['xI'] == dados['xI'].max()])

Dados do país com o maior índice de inovação:
   country  xI  xL  Y continente
0  Switzerland  67.6  83.0  83598  Europa

# calcular o menor valor da coluna xI
menor_valor = dados['xI'].min()

# encontrar o país com o menor valor da coluna xI
print('Dados do país com o menor índice de inovação da amostra:')
print(dados[dados['xI'] == dados['xI'].min()])

Dados do país com o menor índice de inovação da amostra:
   country  xI  xL  Y continente
84  Burundi  12.5  38.4  836  África
```

Maiores e menores índices de liberdade econômica da amostra:

```
#Calculando os maiores e os menores índices de inovação e liberdade econômica das amostra:
dados[dados['xL'] == dados['xL'].max]
print("Dados do país com o maior índice de liberdade econômica:")
print(dados[dados['xL'] == dados['xL'].max()])
print()
print()
dados[dados['xL'] == dados['xL'].min()]
print("Dados do país com o menor índice de liberdade econômica:")
print(dados[dados['xL'] == dados['xL'].min()])

Dados do país com o maior índice de liberdade econômica:
   country  xI  xL  Y continente
4  Singapore  61.5  83.5  127565  Ásia

Dados do país com o menor índice de liberdade econômica:
   country  xI  xL  Y continente
81  Zimbabwe  16.5  38.2  2531  África
```

2. Média e Mediana de Renda per capita

```
# 2. Média e Mediana de Y
media_Y = dados['Y'].mean()
```



```

mediana_Y = dados['Y'].median()
media_Y_por_continente = dados.groupby('continente')['Y'].mean()
mediana_Y_por_continente = dados.groupby('continente')['Y'].median()

print("Média da Renda per capita da amostra:", media_Y)
print("Mediana da Renda per capita da amostra:", mediana_Y)
print()
print()
print("Média da Renda per capita por Continente:")
print(round(media_Y_por_continente))
print()
print()
print("Mediana da Renda per capita por Continente:")
print(mediana_Y_por_continente)

```

Média da Renda per capita da amostra: 31129.937984496122
Mediana da Renda per capita da amostra: 20377.0

Média da Renda per capita por Continente:

continente	
América Central	17417.0
América do Central	22834.0
América do Norte	52104.0
América do Sul	19688.0
Europa	49012.0
Oceania	57296.0
África	6942.0
Ásia	32523.0

Name: Y, dtype: float64

Mediana da Renda per capita por Continente:

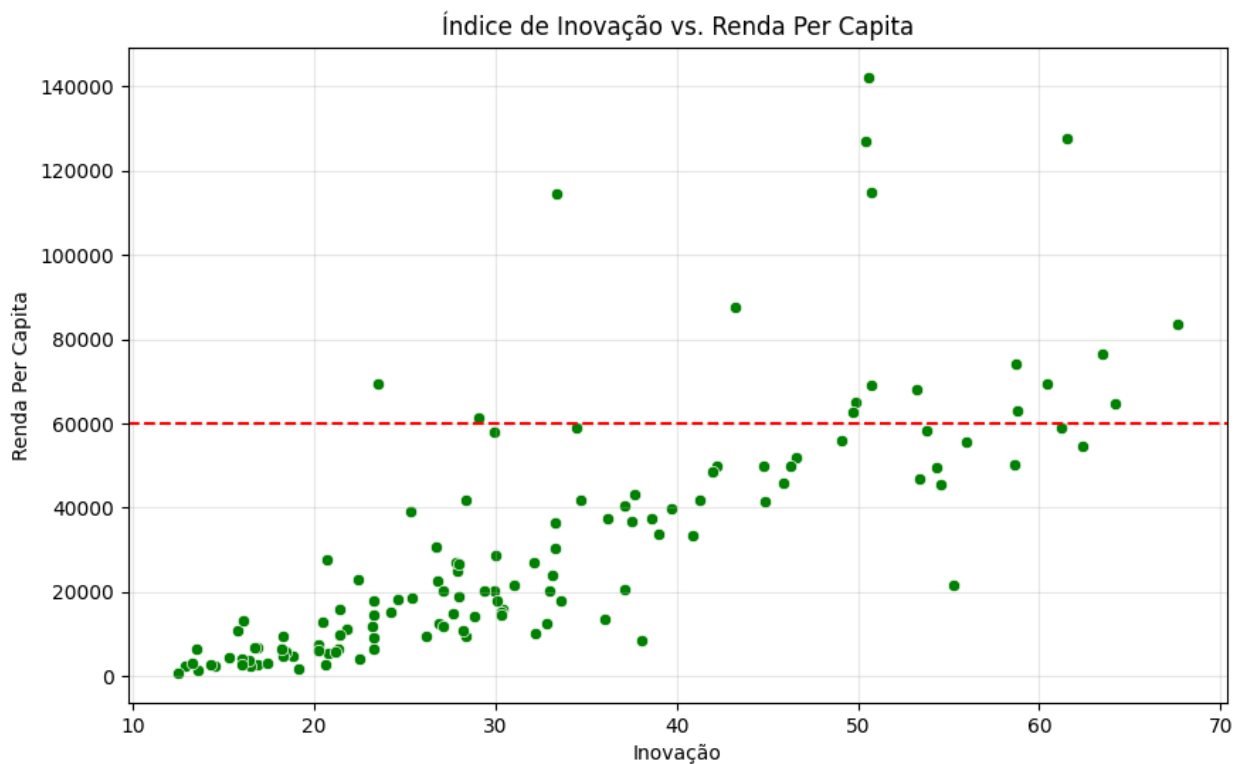
continente	
América Central	11459.0
América do Central	22834.0
América do Norte	58400.0
América do Sul	17822.0
Europa	44547.0
Oceania	57296.0
África	4308.5
Ásia	16364.0

Name: Y, dtype: float64

Medindo a renda per capita média dos países com índice de inovação acima da média e comparar com os de índice abaixo da média:

Visualizar os dados de toda a amostra:

```
#vou montar um gráfico com os dados de xI e Y  
plt.figure(figsize=(10, 6))  
sns.scatterplot(x='xI', y='Y', data=dados, color = 'green')  
plt.title('Índice de Inovação vs. Renda Per Capita')  
plt.xlabel('Inovação')  
plt.ylabel('Renda Per Capita')  
plt.grid(True, alpha = 0.3)  
#linha constante na faixa de renda de 30 mil  
plt.axhline(y=60000, color='red', linestyle='--')  
plt.show()
```



A maioria dos países da amostra têm renda per capita ppp abaixo da faixa dos US\$60 mil.

Países com índice de inovação abaixo da média:

```
print('Média de inovação pelo mundo:')
print(dados['xI'].mean())
print()
print()
abaixo_da_medI = dados[dados['xI'] < dados['xI'].mean()]
print("Países com índice de inovação abaixo da média:")
print(abaixo_da_medI)
```

Média de inovação pelo mundo:
32.65581395348837

Países com índice de inovação abaixo da média:

	country	xI	xL	Y	continente
54	Philippines	32.2	59.0	10133	Ásia
55	Mauritius	32.1	71.5	26906	África
56	Kuwait	29.9	58.5	58056	Ásia
57	Georgia	29.9	68.4	20113	Europa
58	Costa Rica	27.9	67.7	24923	América Central
...
124	Uruguay	30.0	69.8	28842	América do Sul
125	Armenia	28.0	64.9	18942	Europa
126	Argentina	28.0	49.9	26505	América do Sul
127	Benin	16.0	57.7	4056	África
128	Uganda	16.0	50.7	2694	África

[75 rows x 5 columns]

Países com índice de inovação acima da média

```
acima_da_medI = dados[dados['xI'] > dados['xI'].mean()]
print("Países com índice de inovação acima da média:")
print(acima_da_medI)
```

Países com índice de inovação acima da média:

	country	xI	xL	Y	continente
0	Switzerland	67.6	83.0	83598	Europa
1	Sweden	64.2	77.5	64578	Europa
2	United States	63.5	70.1	76399	América do Norte
3	United Kingdom	62.4	68.6	54603	Europa
4	Singapore	61.5	83.5	127565	Ásia
5	Finland	61.2	76.3	59027	Europa
6	Netherlands	60.4	77.3	69577	Europa
7	Germany	58.8	72.1	63150	Europa

8	Denmark	58.7	77.8	74005	Europa
9	South Korea	58.6	73.1	50070	Ásia
10	France	56.0	62.5	55493	Europa
11	China	55.3	48.5	21476	Ásia
12	Japan	54.6	67.5	45573	Ásia
13	Israel	54.3	70.1	49509	Ásia
14	Canada	53.8	72.4	58400	América do Norte
15	Estonia	53.4	77.8	46697	Europa
16	Austria	53.2	68.4	67936	Europa
17	Norway	50.7	77.5	114899	Europa
18	Iceland	50.7	70.5	69081	Europa
19	Luxembourg	50.6	79.2	142214	Europa
20	Ireland	50.4	82.6	126905	Europa
21	Belgium	49.9	65.6	65027	Europa
22	Australia	49.7	76.2	62625	Oceania
23	Malta	49.1	64.5	55928	Europa
24	Italy	46.6	60.1	51865	Europa
25	New Zealand	46.6	77.8	51967	Oceania
26	Cyprus	46.3	72.2	49931	Europa
27	Spain	45.9	63.3	45825	Europa
28	Portugal	44.9	68.7	41452	Europa
29	Czech Republic	44.8	70.2	49946	Europa
30	United Arab Emirates	43.2	71.1	87729	Ásia
31	Slovenia	42.2	65.9	50032	Europa
32	Lithuania	42.0	72.9	48397	Europa
33	Hungary	41.3	61.2	41907	Europa
34	Malaysia	40.9	65.7	33434	Ásia
35	Latvia	39.7	71.5	39956	Europa
36	Bulgaria	39.0	68.5	33582	Europa
37	Türkiye	38.6	56.2	37274	Europa
38	India	38.1	52.9	8379	Ásia
39	Poland	37.7	66.0	43269	Europa
40	Greece	37.5	55.1	36835	Europa
41	Thailand	37.1	59.0	20672	Ásia
42	Croatia	37.1	67.2	40380	Europa
43	Slovakia	36.2	68.1	37459	Europa
44	Vietnam	36.0	62.8	13457	Ásia
45	Romania	34.7	64.4	41888	Europa
46	Saudi Arabia	34.5	61.9	59065	Ásia
47	Brazil	33.6	53.2	17822	América do Sul
48	Qatar	33.4	68.8	114648	Ásia
49	Russia	33.3	52.0	36485	Europa
50	Chile	33.3	71.4	30209	América do Sul
51	Serbia	33.1	62.7	23911	Europa
52	North Macedonia	33.0	61.4	20162	Europa
53	Ukraine	32.8	54.1	12671	Europa

Vamos calcular a renda média dos países com índice de inovação abaixo da média:

```
renda_media_abaixo_da_medI = abaixo_da_medI['Y'].mean()
print("Renda média dos países com índice de inovação abaixo da média: US$", round(renda_media_abaixo_da_medI))
```

Renda média dos países com índice de inovação abaixo da média: US\$ 14544

Vamos fazer o mesmo com os países com inovação acima da média:

```
acima_da_medI_renda_media = acima_da_medI['Y'].mean()
print("Renda média dos países com índice de inovação acima da média: US$", round(acima_da_medI_renda_media))
```

Renda média dos países com índice de inovação acima da média: US\$ 54166

Diferença percentual de renda entre esses países

```
diferença = (round(acima_da_medI_renda_media) - round(renda_media_abaixo_da_medI))/round(acima_da_medI_renda_media)
percenteI = round(diferença * 100)
print("Diferença percentual média de renda entre os países com o índice de inovação acima da média para os países com índice de inovação abaixo da média:", percenteI, '%')
```

Diferença percentual média de renda entre os países com o índice de inovação acima da média para os países com índice de inovação abaixo da média: 73 %

veja que existe uma diferença muito grande entre a renda dos países com melhor desempenho em inovação, para os de pior desempenho.

Quero visualizar os dados de renda vs inovação para cada grupo:

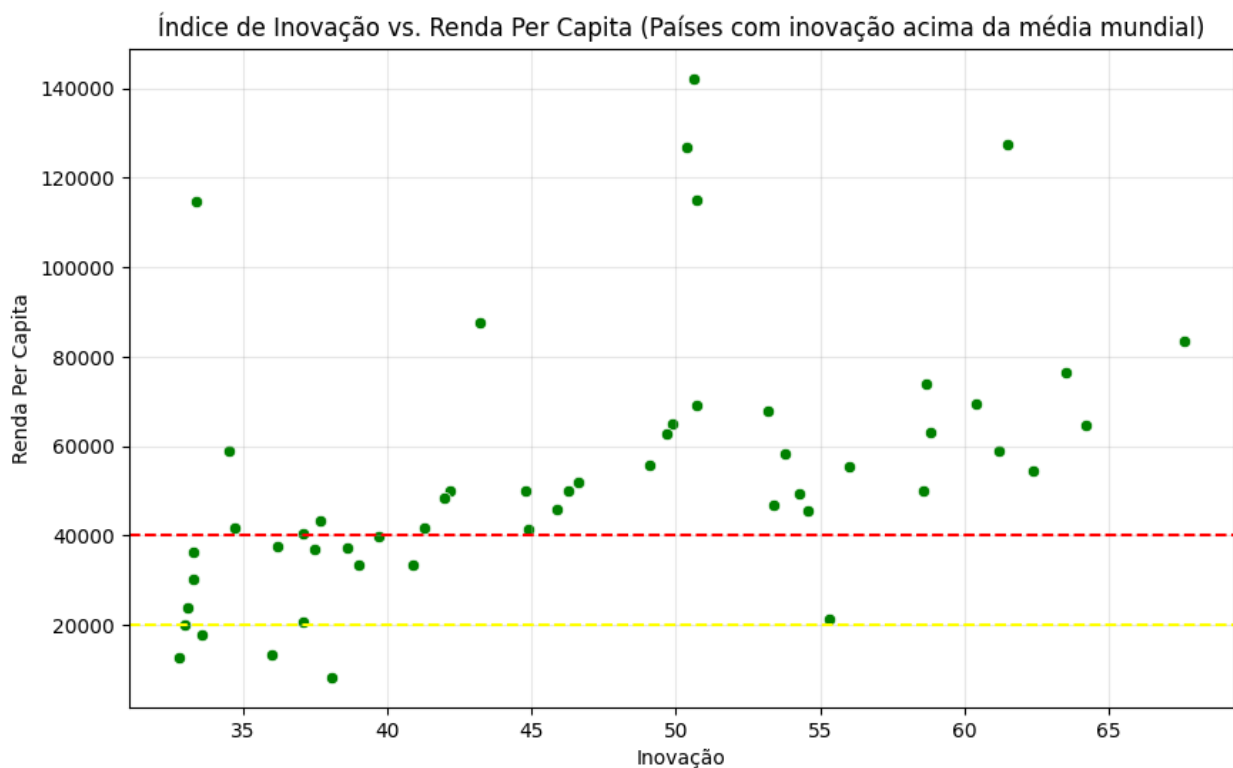
Gráfico dos países com melhor desempenho em inovação:

```
# Gráfico dos países com melhor desempenho em inovação:
plt.figure(figsize=(10, 6))
sns.scatterplot(x='xI', y='Y', data=acima_da_medI, color = 'green')
```

```
plt.title('Índice de Inovação vs. Renda Per Capita (Países com
inovação acima da média mundial)')
plt.xlabel('Inovação')
plt.ylabel('Renda Per Capita')

#vou colocar uma linha constante na faixa de renda per capita de US$
40 mil
plt.axhline(y=40000, color='red', linestyle='--')

#vou colocar outra linha na faixa dos US$20 mil
plt.axhline(y=20000, color='yellow', linestyle='--')
plt.grid(True, alpha = 0.3)
plt.show()
```



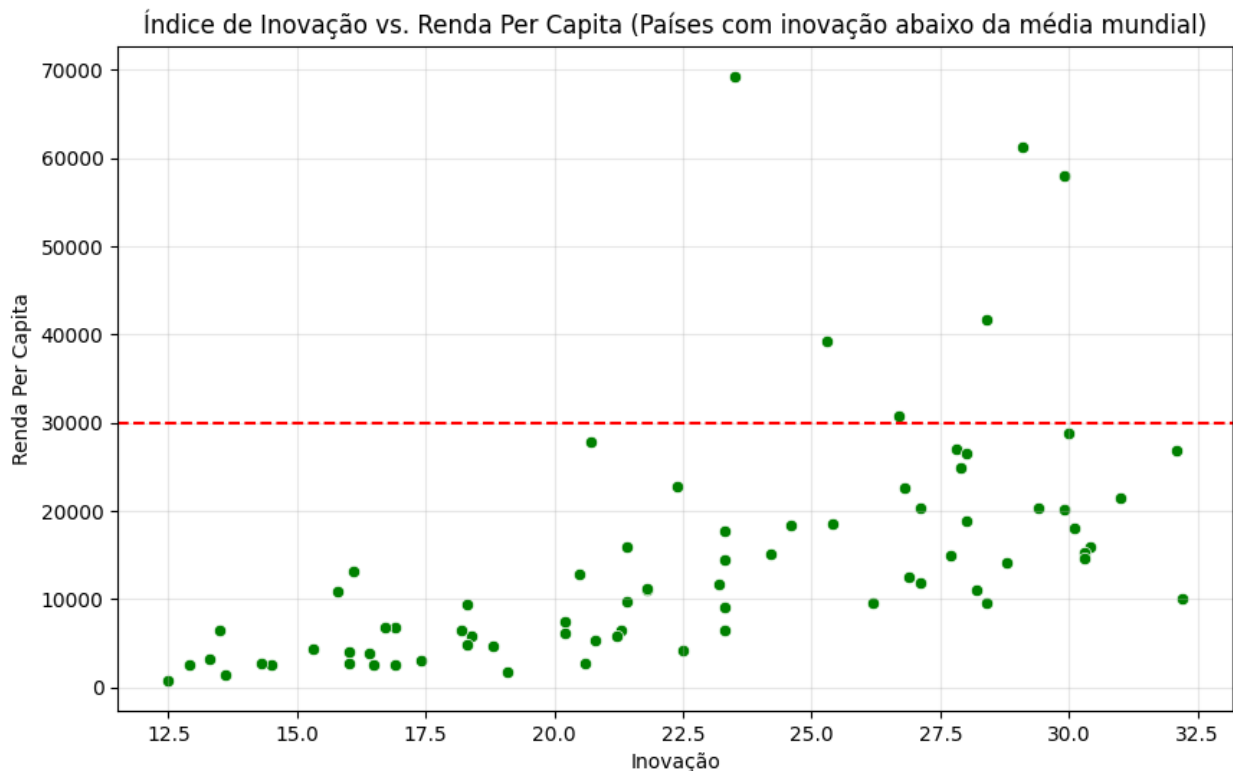
Veja que a maioria desses países estão com renda acima dos US\$ 40 mil (acima da linha tracejada em vermelha)

Além disso, que apenas quatro deles possuem renda abaixo da faixa dos US\$20 mil (linha tracejada em amarelo)

Inovação vs. Renda per Capita: visualizando isso para países com menor inovação

```
plt.figure(figsize=(10, 6))
sns.scatterplot(x='xI', y='Y', data=abaixo_da_medI, color = 'green')
plt.title('Índice de Inovação vs. Renda Per Capita (Países com
inovação abaixo da média mundial)')
plt.xlabel('Inovação')
plt.ylabel('Renda Per Capita')
plt.grid(True, alpha = 0.3)

#vou colocar uma linha constante na faixa de renda per capita de US$
40 mil
plt.axhline(y=30000, color='red', linestyle='--')
plt.show()
```



```
#vou calcular os países com Y acima do 30000 do grupo de países
abaixo_da_medI
print('Países com inovação abaixo da média, mas com elevada renda per
capita:')
print(abaixo_da_medI[abaixo_da_medI['Y'] > 30000])
```

Países com inovação abaixo da média, mas com elevada renda per capita:

	country	xI	xL	Y	continente
56	Kuwait	29.9	58.5	58056	Ásia
72	Kazakhstan	26.7	62.0	30810	Ásia
78	Brunei	23.5	65.9	69275	Ásia
87	Oman	28.4	62.9	41724	Ásia
98	Panama	25.3	64.1	39280	América Central
118	Bahrain	29.1	63.4	61228	Ásia

A semelhança entre eles? São todos, com o excessão de Omã e o Panamá, grandes produtores de petróleo! O petróleo tem compensado a falta de inovação desses países.

Conclusão da relação entre Inovação e Renda:

- países com índice de inovação acima da média são mais ricos!
 - países sem grandes reservas de petróleo precisam ser inovadores, caso contrário dificilmente conseguem ter um grande Pib per Capita
-

Vou fazer a mesma análise anterior agora considerando a Liberdade Econômica:

Países com índice de liberdade econômica abaixo da média:

```
abaixo_da_medL = dados[dados['xL'] < dados['xL'].mean()]
print("Países com índice de liberdade econômica abaixo da média:")
print(abaixo_da_medL)
```

Países com índice de liberdade econômica abaixo da média:

	country	xI	xL	Y	continente
--	---------	----	----	---	------------

11	China	55.3	48.5	21476	Ásia
24	Italy	46.6	60.1	51865	Europa
33	Hungary	41.3	61.2	41907	Europa
37	Türkiye	38.6	56.2	37274	Europa
38	India	38.1	52.9	8379	Ásia
...
121	Madagascar	19.1	57.3	1774	África
122	Algeria	16.1	43.9	13210	África
126	Argentina	28.0	49.9	26505	América do Sul
127	Benin	16.0	57.7	4056	África
128	Uganda	16.0	50.7	2694	África

[64 rows x 5 columns]

Países com liberdade econômica acima da média:

```
acima_da_medL = dados[dados['xL'] > dados['xL'].mean()]
print("Países com índice de liberdade econômica acima da média:")
print(acima_da_medL)
```

Países com índice de liberdade econômica acima da média:

	country	xI	xL	Y	continente
0	Switzerland	67.6	83.0	83598	Europa
1	Sweden	64.2	77.5	64578	Europa
2	United States	63.5	70.1	76399	América do Norte
3	United Kingdom	62.4	68.6	54603	Europa
4	Singapore	61.5	83.5	127565	Ásia
...
119	Bosnia and Herzegovina	27.1	62.0	20377	Europa
120	Jamaica	27.1	68.1	11822	América Central
123	Mexico	31.0	62.0	21512	América do Norte
124	Uruguay	30.0	69.8	28842	América do Sul
125	Armenia	28.0	64.9	18942	Europa

[65 rows x 5 columns]

Renda média para os países com liberdade econômica abaixo da média

```
renda_media_abaixo_da_medL = abaixo_da_medL['Y'].mean()
print("Renda média dos países com índice de liberdade econômica abaixo da média: US$", round(renda_media_abaixo_da_medL))
```

Renda média dos países com índice de liberdade econômica abaixo da média: US\$ 13021

Renda média dos países com índice de liberdade econômica acima da média:

```
acima_da_medL_renda_media = acima_da_medL['Y'].mean()
print("Renda média dos países com índice de liberdade econômica acima da média: US$", round(acima_da_medL_renda_media))
```

Renda média dos países com índice de liberdade econômica acima da média: US\$ 48960

Diferença percentual de renda

```
# vou calcular a diferença percentual entre esses países
diferença_L = (round(acima_da_medL_renda_media) - round(renda_media_abaixo_da_medL))/round(acima_da_medL_renda_media)
print("Diferença percentual média de renda entre os países com o índice de liberdade econômica acima da média para os países com índice de liberdade econômica abaixo da média:", round(diferença_L * 100), '%')
```

Diferença percentual média de renda entre os países com o índice de liberdade econômica acima da média para os países com índice de liberdade econômica abaixo da média: 73 %

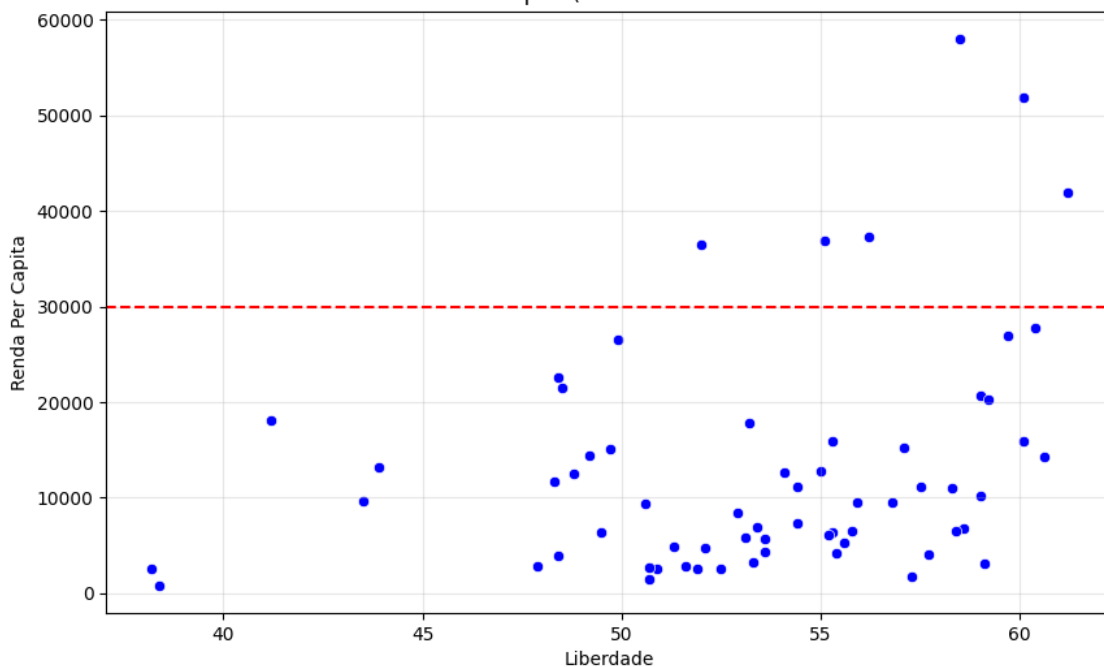
Veja que a diferença percentual média da renda entre os mais livres economicamente e os menos livre é igual a diferença entre os países mais inovadores e os menos inovadores, dessa amostra

Liberdade econômica vs. Renda per Capita: Visualizando o caso dos países com menor liberdade econômica

```
plt.figure(figsize=(10, 6))
sns.scatterplot(x='xL', y='Y', data=abaixo_da_medL, color = 'blue')
plt.title('Índice de Liberdade econômica vs. Renda Per Capita (Países com Liberdade econômica acima da média mundial)')
plt.xlabel('Liberdade')
plt.ylabel('Renda Per Capita')
plt.grid(True, alpha = 0.3)

#vou colocar uma linha na faixa de renda dos us$ 30 mil
plt.axhline(y=30000, color='red', linestyle='--')
plt.show()
```

Índice de Liberdade econômica vs. Renda Per Capita (Países com Liberdade econômica acima da média mundial)



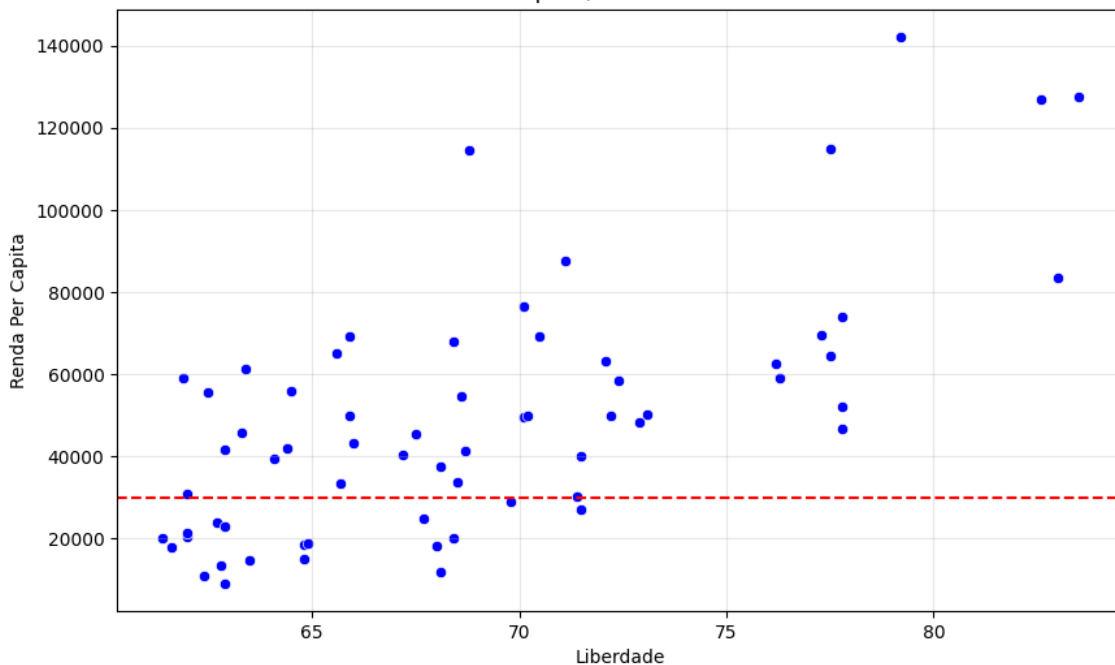
A maioria dos países menos livres tem renda mais baixa.

Visualizar agora para os países mais livres economicamente:

```
plt.figure(figsize=(10, 6))
sns.scatterplot(x='xL', y='Y', data=acima_da_medL, color = 'blue')
plt.title('Índice de Liberdade econômica vs. Renda Per Capita (Países com Liberdade econômica acima da média mundial)')
plt.xlabel('Liberdade')
plt.ylabel('Renda Per Capita')
plt.grid(True, alpha = 0.3)

#vou colocar uma linha na faixa de renda dos us$ 30 mil
plt.axhline(y=30000, color='red', linestyle='--')
plt.show()
```

Índice de Liberdade econômica vs. Renda Per Capita (Países com Liberdade econômica acima da média mundial)



Note que a quantidade de países acima dessa faixa do US\$30 mil é bem maior para esse grupo de países. Porém, que também há vários países que mesmo livres, têm renda abaixo dessa faixa.

Vamos tirar algumas conclusões acerca da relação entre Liberdade Econômica e Renda per Capita

3. Variância e Desvio Padrão

```
desvio_padrao_Y = dados['Y'].std()
desvio_padrao_xI = dados['xI'].std()
desvio_padrao_xL = dados['xL'].std()

print("Desvio Padrão de Y:", desvio_padrao_Y)
print()
print("Desvio Padrão de xI:", desvio_padrao_xI)
print()
print("Desvio Padrão de xL:", desvio_padrao_xL)
```

```
Desvio Padrão de Y: 28960.040757091036
```

```
Desvio Padrão de xI: 14.108839798690616
```

```
Desvio Padrão de xL: 9.546156297088876
```

Disparidade de renda entre os países com maior liberdade econômica

```
acima_da_medL['Y']  
dp_Y_acima_da_medL = acima_da_medL['Y'].std()  
print("Desvio Padrão de Renda para os países com índice de liberdade  
econômica acima da média: US$", round(dp_Y_acima_da_medL))
```

```
Desvio Padrão de Renda para os países com índice de liberdade  
econômica acima da média: US$ 29656
```

Veja que as diferenças de renda entre os países com maiores índices de liberdade econômica é alta. Dessa forma, pode-se notar que existe uma alta disparidade entre as rendas de países com elevada liberdade econômica.

Vejamos o caso da disparidade de renda entre os países com menor liberdade econômica

```
abaixo_da_medL['Y']  
dp_Y_abaixo_da_medL = abaixo_da_medL['Y'].std()  
print("Desvio Padrão de Renda para os países com índice de liberdade  
econômica abaixo da média: US$", round(dp_Y_abaixo_da_medL))
```

```
Desvio Padrão de Renda para os países com índice de liberdade  
econômica abaixo da média: US$ 12225
```

Veja que a diferença de renda entre os menos livres é menor.

4. Correlação

```
correlacao_Y_xI = dados[['Y', 'xI']].corr().iloc[0, 1]  
correlacao_Y_xL = dados[['Y', 'xL']].corr().iloc[0, 1]  
print("Correlação entre Renda e Inovação:", correlacao_Y_xI)  
print()  
print("Correlação entre Renda e Liberdade Econômica:",  
correlacao_Y_xL)  
print()
```

```
correlação_xI_xL = dados[['xI', 'xL']].corr().iloc[0, 1]
print("Correlação entre Inovação e Liberdade Econômica:",
correlação_xI_xL)
```

Correlação entre Renda e Inovação: 0.7723983547872895

Correlação entre Renda e Liberdade Econômica: 0.7573607448104864

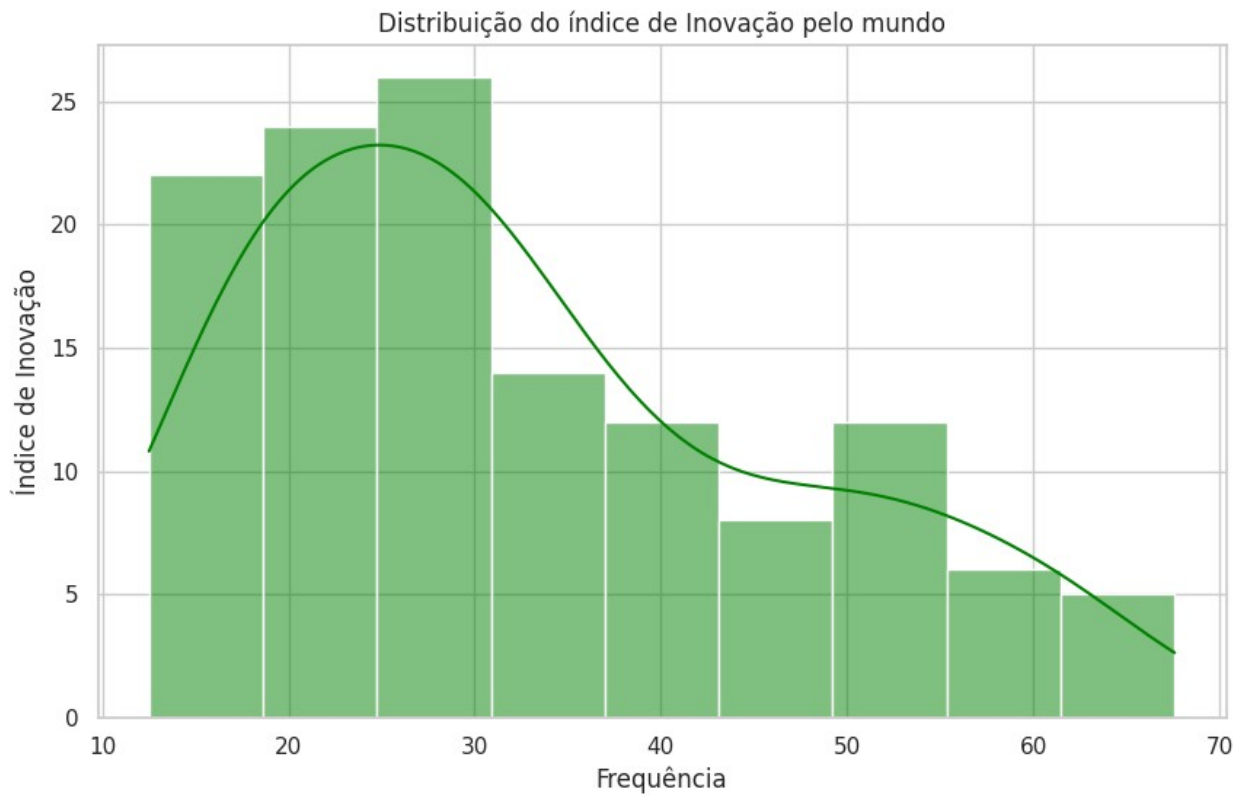
Correlação entre Inovação e Liberdade Econômica: 0.7401587587124917

Veja que existe uma correlação forte entre todas essas três variáveis.

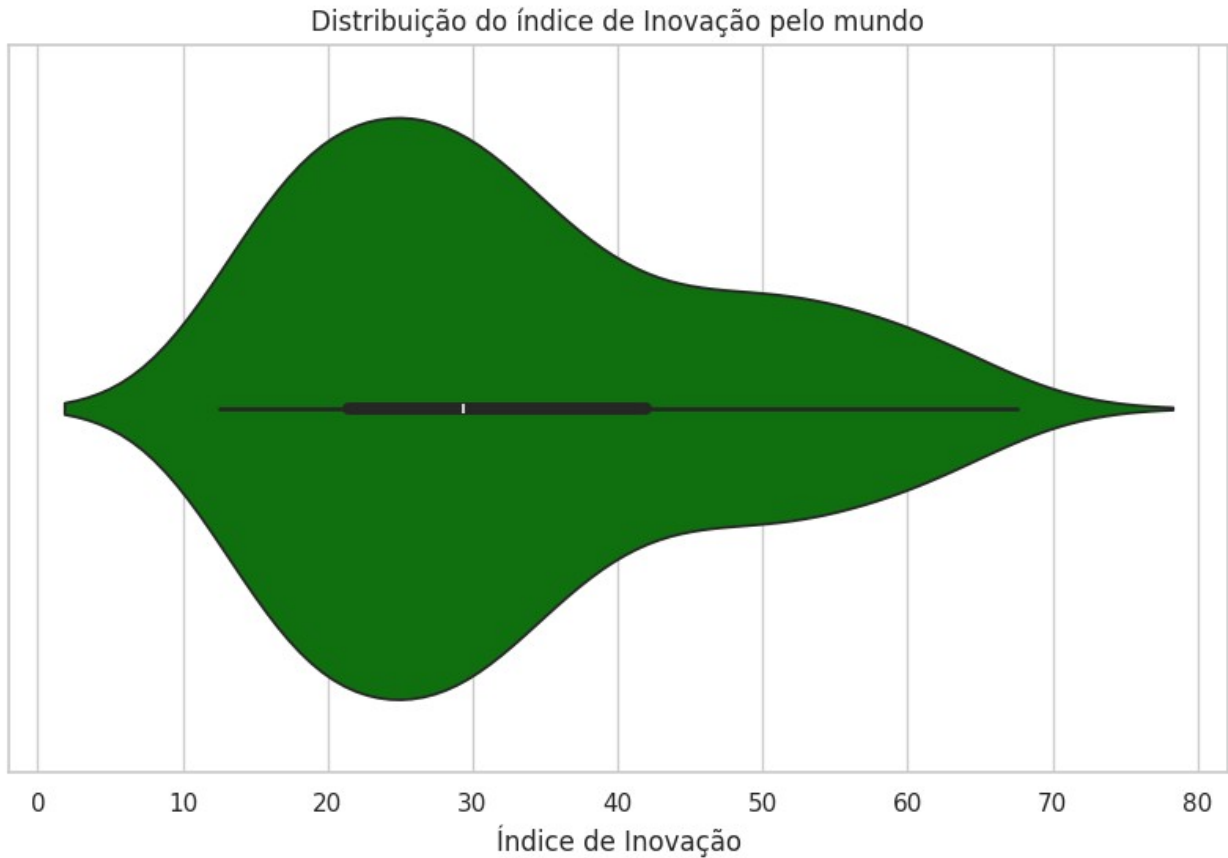
Distribuição e outras visualizações interessantes para verificar a inovação pelo mundo:

Distribuição inovação ao redor mundo

```
#Crio uma figura
sns.set(style="whitegrid")
plt.figure(figsize=(10, 6))
sns.histplot(x=dados['xI'], kde=True, color = 'green')
plt.title('Distribuição do índice de Inovação pelo mundo')
plt.xlabel('Frequência')
plt.ylabel('Índice de Inovação')
plt.show()
```



```
sns.set(style="whitegrid")
plt.figure(figsize=(10, 6))
sns.violinplot(x=dados['xI'], color = 'green')
plt.title('Distribuição do índice de Inovação pelo mundo')
plt.xlabel('Índice de Inovação')
plt.show()
```



Essas visualizações evidenciam que o índice de inovação pelo mundo está a maior parte concentrada ali na faixa entre 10 e 30 e poucos pontos. O Brasil por exemplo, tem um índice de inovação de 33,6 pontos no Índice. Estando em 48º posição.

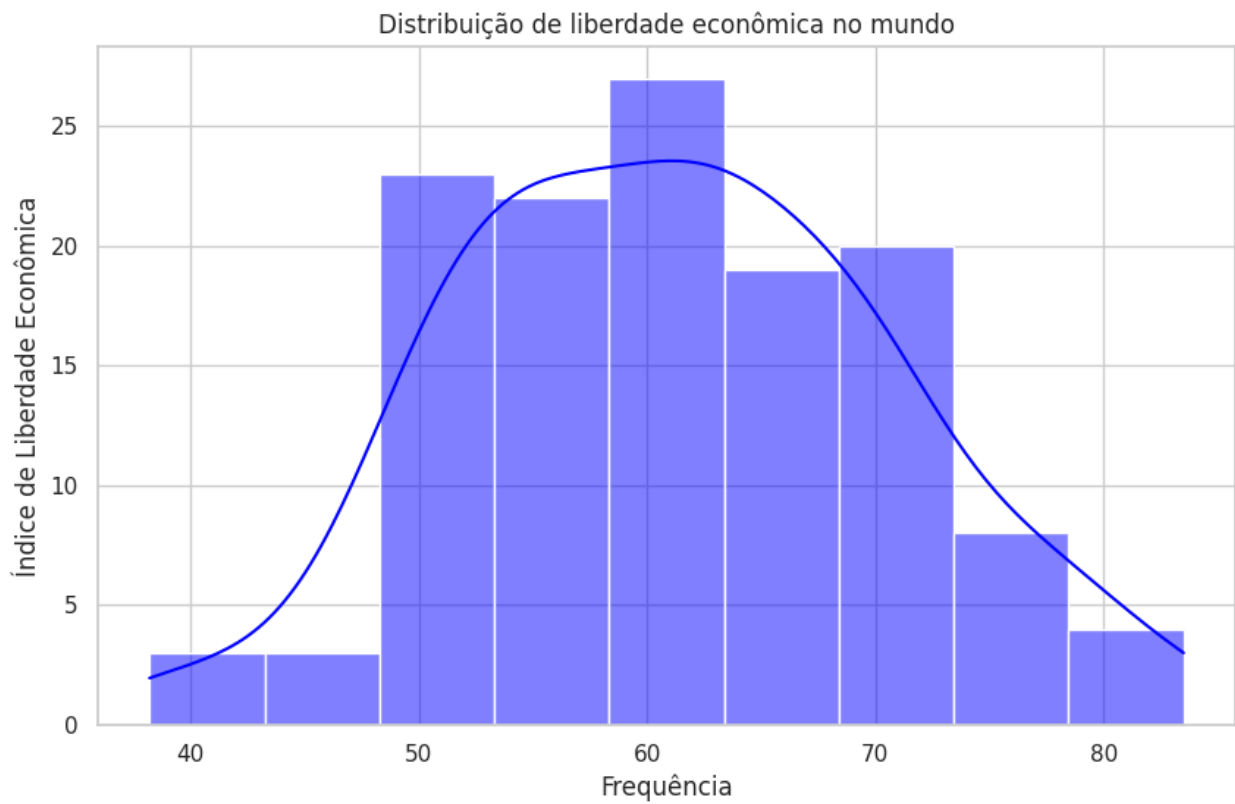
Distribuição e outras visualizações interessantes para verificar a liberdade econômica pelo mundo:

Distribuição liberdade econômica ao redor mundo

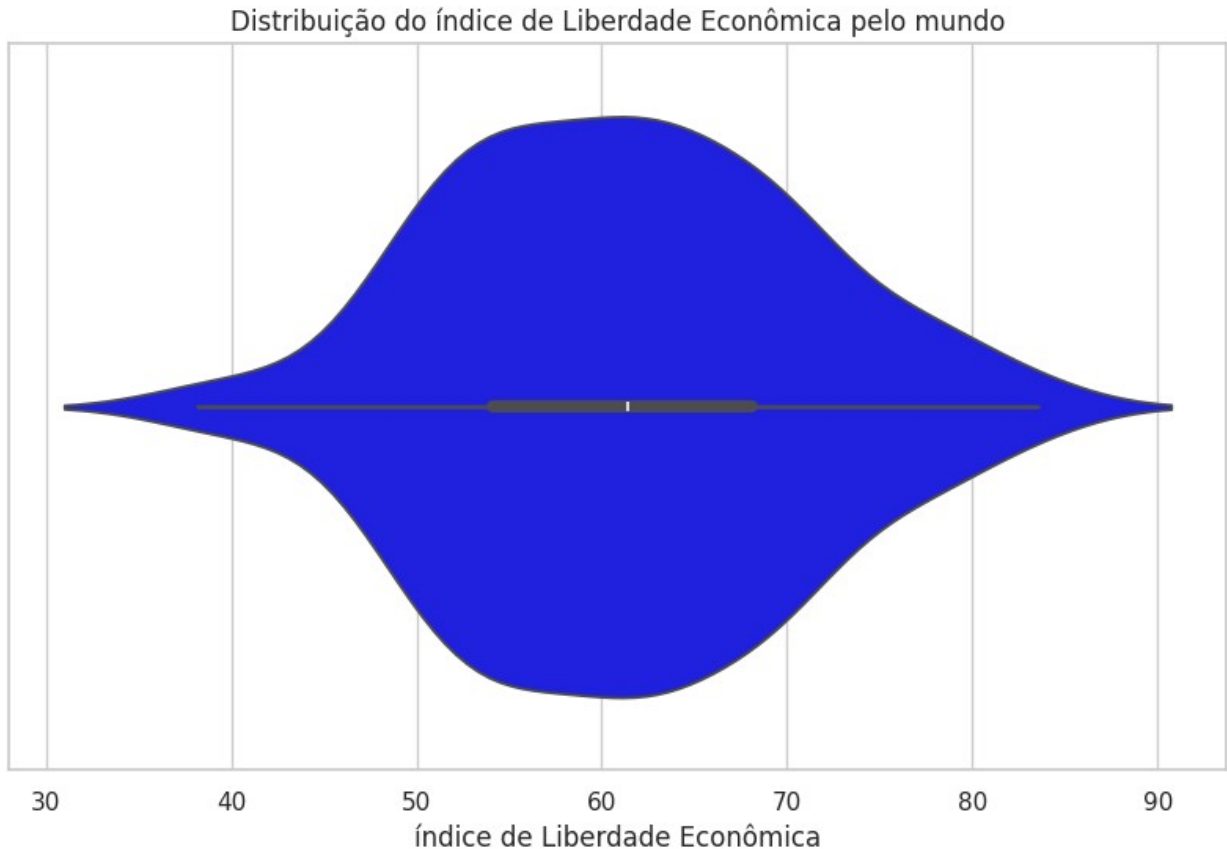
```
#Crio uma figura
sns.set(style="whitegrid")
plt.figure(figsize=(10, 6))
sns.histplot(x=dados['xL'], kde=True, color = 'blue')
plt.title('Distribuição de liberdade econômica no mundo')
plt.xlabel('Frequência')
```



```
plt.ylabel('Índice de Liberdade Econômica')  
plt.show()
```



```
sns.set(style="whitegrid")  
plt.figure(figsize=(10, 6))  
sns.violinplot(x=dados['xL'], color = 'blue')  
plt.title('Distribuição do índice de Liberdade Econômica pelo mundo')  
plt.xlabel('índice de Liberdade Econômica')  
plt.show()
```



Modelo de Regressão:

```
# Função para ajustar o modelo de regressão e retorná-lo
def ajustar_modelo(y, X):
    X = sm.add_constant(X) # Adicionar constante (intercepto) ao
    modelo
    modelo = sm.OLS(y, X).fit()
    return modelo

# 1. Regressão: Inovação (xI) vs. Renda (Y)
modelo_inovacao = ajustar_modelo(dados['Y'], dados['xI'])
print("Regressão: Inovação (xI) vs. Renda (Y)")
print(modelo_inovacao.summary())
```

Regressão: Inovação (xI) vs. Renda (Y)
OLS Regression Results

```
=====
=====
Dep. Variable:          Y    R-squared:
```

```

0.597
Model: OLS Adj. R-squared:
0.593
Method: Least Squares F-statistic:
187.8
Date: Mon, 01 Jul 2024 Prob (F-statistic):
8.38e-27
Time: 18:24:44 Log-Likelihood:
-1449.3
No. Observations: 129 AIC:
2903.
Df Residuals: 127 BIC:
2908.
Df Model: 1

Covariance Type: nonrobust

=====
=====
              coef      std err          t      P>|t|      [0.025
0.975]
-----
-----
const      -2.064e+04    4112.765     -5.019     0.000    -2.88e+04    -
1.25e+04
xI          1585.4378     115.684     13.705     0.000     1356.519
1814.356
=====
=====
Omnibus:      81.255    Durbin-Watson:
1.712
Prob(Omnibus):    0.000    Jarque-Bera (JB):
401.208
Skew:          2.279    Prob(JB):
7.57e-88
Kurtosis:      10.339    Cond. No.
90.0
=====
=====

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.

```

Gráfico dos dados com a linha de regressão

```

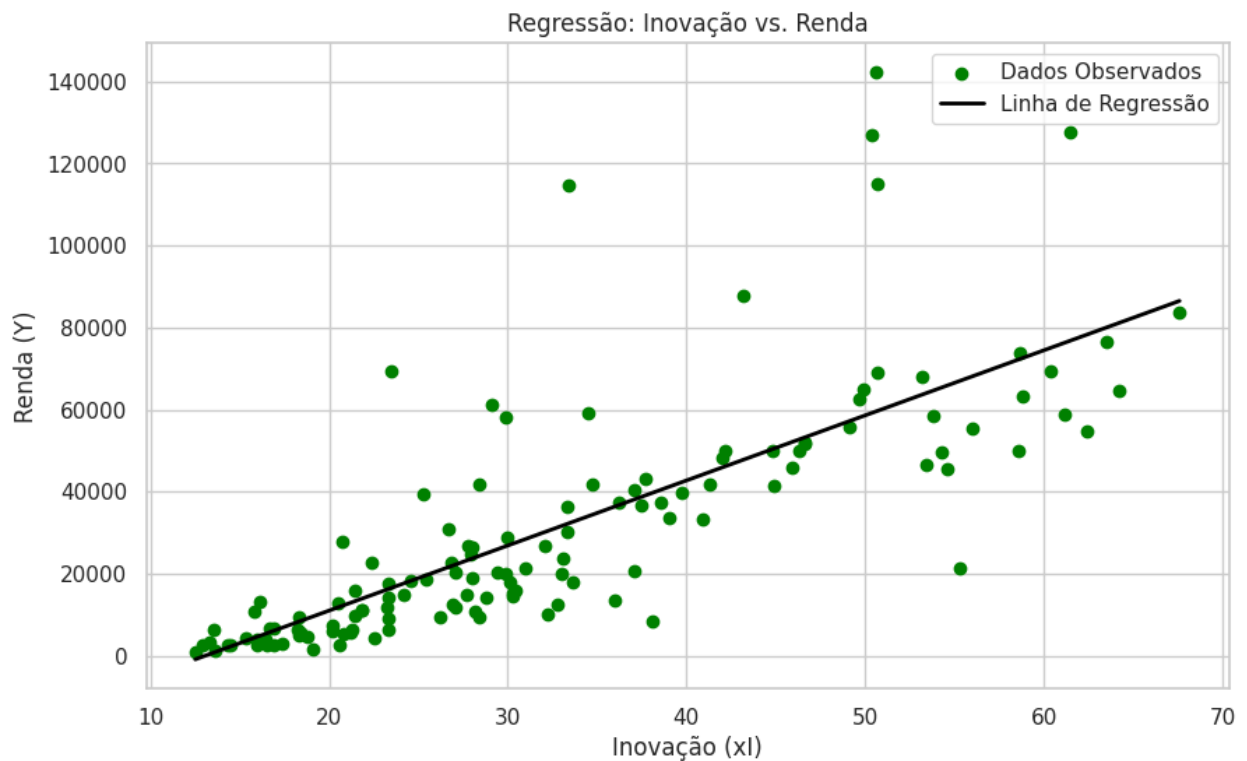
# Dados para plotagem da linha de regressão
x_range = np.linspace(dados['xI'].min(), dados['xI'].max(), 100)
y_pred = modelo_inovacao.predict(sm.add_constant(x_range))

```

```

# Plotagem da curva de regressão
plt.figure(figsize=(10, 6))
plt.scatter(dados['xI'], dados['Y'], color='green', label='Dados Observados') # Pontos de dados
plt.plot(x_range, y_pred, color='black', linewidth=2, label='Linha de Regressão') # Linha de regressão
plt.xlabel('Inovação (xI)')
plt.ylabel('Renda (Y)')
plt.title('Regressão: Inovação vs. Renda')
plt.legend()
plt.grid(True)
plt.show()

```



Interpretando os resultados da Regressão Simples

Inovação vs. Renda per Capita ppc:

No modelo, a variável Índice de Inovação (xl) apresenta um coeficiente positivo significativo de 1585.4378 ($p < 0.001$), indicando que, para cada unidade adicional de Inovação, a renda per capita aumenta em média aproximadamente em US\$ 1585. Esse coeficiente revela a magnitude e a direção da relação entre Inovação e Renda, demonstrando que níveis mais altos de inovação estão associados a aumentos substanciais na renda per capita.

O R^2 ajustado é de 0.593, o que significa que aproximadamente 59.3% da variabilidade na renda per capita pode ser explicada pela variável Inovação dentro do modelo utilizado. Em outras palavras, cerca de 59.3% das diferenças observadas na renda per capita entre os países ou regiões estudadas podem ser atribuídas às variações na Inovação, após ajustar para o número de preditores no modelo. Esse valor de R^2 ajustado indica um nível relativamente alto de explicação da variabilidade, sugerindo que a Inovação é um fator crucial para a determinação da renda per capita, embora outros fatores também possam influenciar significativamente e devam ser considerados para uma análise mais completa.

Regressão Liberdade Econômica vs. Renda per Capita PPC

```
# 2. Regressão: Liberdade Econômica (xL) vs. Renda (Y)
modelo_liberdade = ajustar_modelo(dados['Y'], dados['xL'])
print("\nRegressão: Liberdade Econômica (xL) vs. Renda (Y)")
print(modelo_liberdade.summary())
```

Regressão: Liberdade Econômica (xL) vs. Renda (Y)
OLS Regression Results

```
=====
=====
Dep. Variable:                Y    R-squared:
0.574
Model:                OLS    Adj. R-squared:
0.570
Method:                Least Squares    F-statistic:
170.8
```

```

Date: Mon, 01 Jul 2024 Prob (F-statistic):
2.89e-25
Time: 18:25:01 Log-Likelihood:
-1452.9
No. Observations: 129 AIC:
2910.
Df Residuals: 127 BIC:
2915.
Df Model: 1

Covariance Type: nonrobust

=====
=====
              coef      std err          t      P>|t|      [0.025
0.975]
-----
-----
const      -1.095e+05    1.09e+04    -10.057      0.000    -1.31e+05
-8.8e+04
xL          2297.5947    175.784     13.071      0.000     1949.750
2645.440
=====
=====
Omnibus:                26.725    Durbin-Watson:
1.572
Prob(Omnibus):           0.000    Jarque-Bera (JB):
37.482
Skew:                    1.085    Prob(JB):
7.26e-09
Kurtosis:                4.506    Cond. No.
404.
=====
=====

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.

```

Visualizar agora a regressão xL vs Y

```

# Dados para plotagem da linha de regressão
x2_range = np.linspace(dados['xL'].min(), dados['xL'].max(), 100)
y2_pred = modelo_liberdade.predict(sm.add_constant(x2_range))

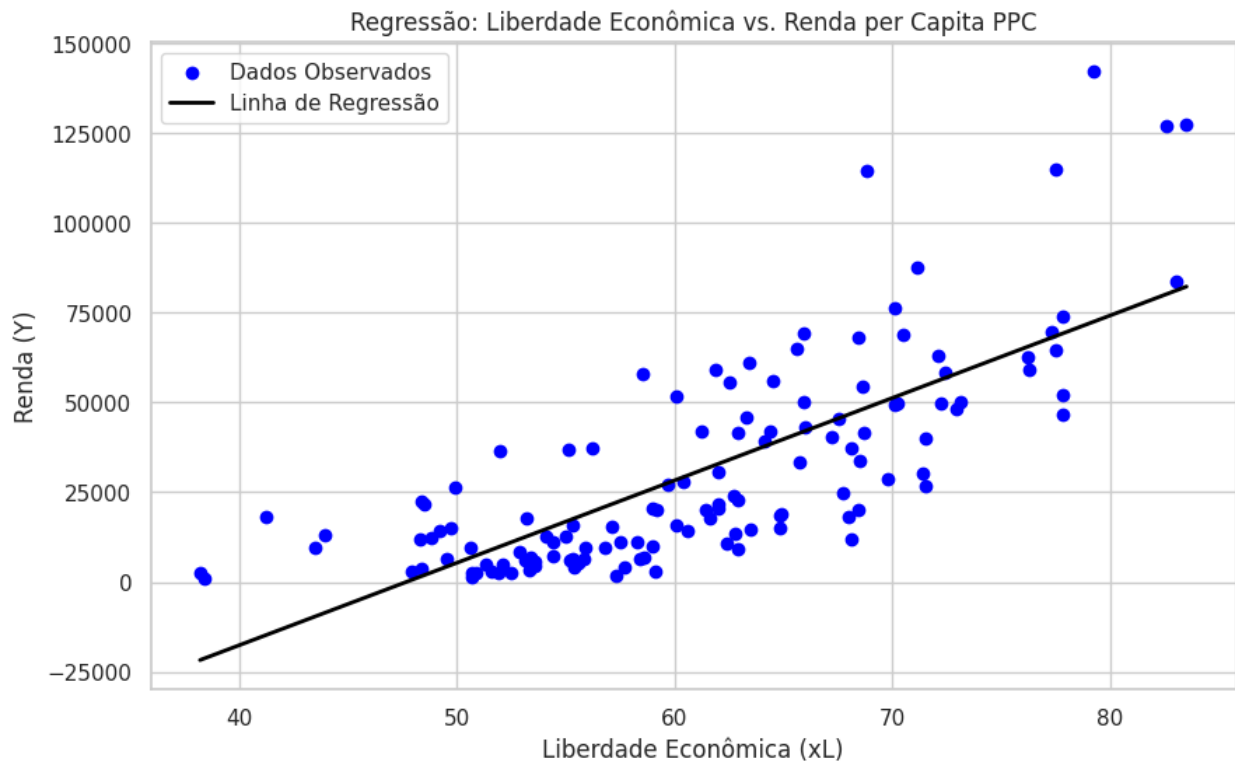
# Plotagem da curva de regressão
plt.figure(figsize=(10, 6))
plt.scatter(dados['xL'], dados['Y'], color='blue', label='Dados
Observados')

```

```

plt.plot(x2_range, y2_pred, color='black', linewidth=2, label='Linha
de Regressão')
plt.xlabel('Liberdade Econômica (xL)')
plt.ylabel('Renda (Y)')
plt.title('Regressão: Liberdade Econômica vs. Renda per Capita PPC')
plt.legend()
plt.grid(True)
plt.show()

```



Interpretando os resultados da Regressão Simples

Liberdade Econômica vs. Renda per Capita ppc:

A variável Liberdade Econômica (x_L) mostra um coeficiente positivo significativo de 2297.5947 ($p < 0.001$), indicando que, para cada unidade adicional de Liberdade Econômica, a renda per capita aumenta em média US\$2297. Esse coeficiente reflete a magnitude e a direção da associação entre Liberdade Econômica e Renda, demonstrando que maiores níveis de liberdade econômica tendem a resultar em maiores níveis de renda per capita.

O R^2 ajustado é de 0.57, o que significa que aproximadamente 57% da variabilidade na renda per capita pode ser explicada pela variável Liberdade Econômica dentro do modelo utilizado. Esse valor de R^2 ajustado (coeficiente de determinação) um nível relativamente alto de explicação da variabilidade, sugerindo que a Liberdade Econômica é um fator crucial para a determinação da renda per capita, embora outros fatores também possam influenciar significativamente e devam ser considerados para uma análise mais completa.

Regressão Múltipla:

```
# 3. Regressão: Inovação (xI) + Liberdade Econômica (xL) vs. Renda (Y)
modelo_inovacao_liberdade = ajustar_modelo(dados['Y'], dados[['xI',
'xL']])
print("\nRegressão: Inovação (xI) + Liberdade Econômica (xL) vs. Renda
(Y)")
print(modelo_inovacao_liberdade.summary())
```

Regressão: Inovação (xI) + Liberdade Econômica (xL) vs. Renda (Y)
OLS Regression Results

```
=====
=====
Dep. Variable:                Y    R-squared:
0.673
Model:                        OLS    Adj. R-squared:
0.668
Method:                        Least Squares    F-statistic:
129.6
Date:                          Mon, 01 Jul 2024    Prob (F-statistic):
```



```

2.69e-31
Time: 18:26:00 Log-Likelihood:
-1435.8
No. Observations: 129 AIC:
2878.
Df Residuals: 126 BIC:
2886.
Df Model: 2

Covariance Type: nonrobust

=====
=====
              coef      std err          t      P>|t|      [0.025
0.975]
-----
-----
const      -7.653e+04    1.1e+04    -6.980    0.000    -9.82e+04    -
5.48e+04
xI           961.6140    155.546     6.182    0.000     653.794
1269.434
xL          1245.6609    229.890     5.419    0.000     790.715
1700.607
=====
=====
Omnibus: 67.842 Durbin-Watson:
1.693
Prob(Omnibus): 0.000 Jarque-Bera (JB):
237.148
Skew: 1.991 Prob(JB):
3.19e-52
Kurtosis: 8.316 Cond. No.
528.
=====
=====

Notes:
[1] Standard Errors assume that the covariance matrix of the errors is
correctly specified.

```

Gráfico da Regressão Múltipla:

```

# Dados para plotagem da linha de regressão
x3_range = np.linspace(dados[['xI', 'xL']].min().min(), dados[['xI',
'xL']].max().max(), 100)
x3_grid = np.meshgrid(x3_range, x3_range)
x3_pred =
modelo_inovacao_liberdade.predict(sm.add_constant(np.stack(x3_grid,
axis=-1).reshape(-1, 2)))

```

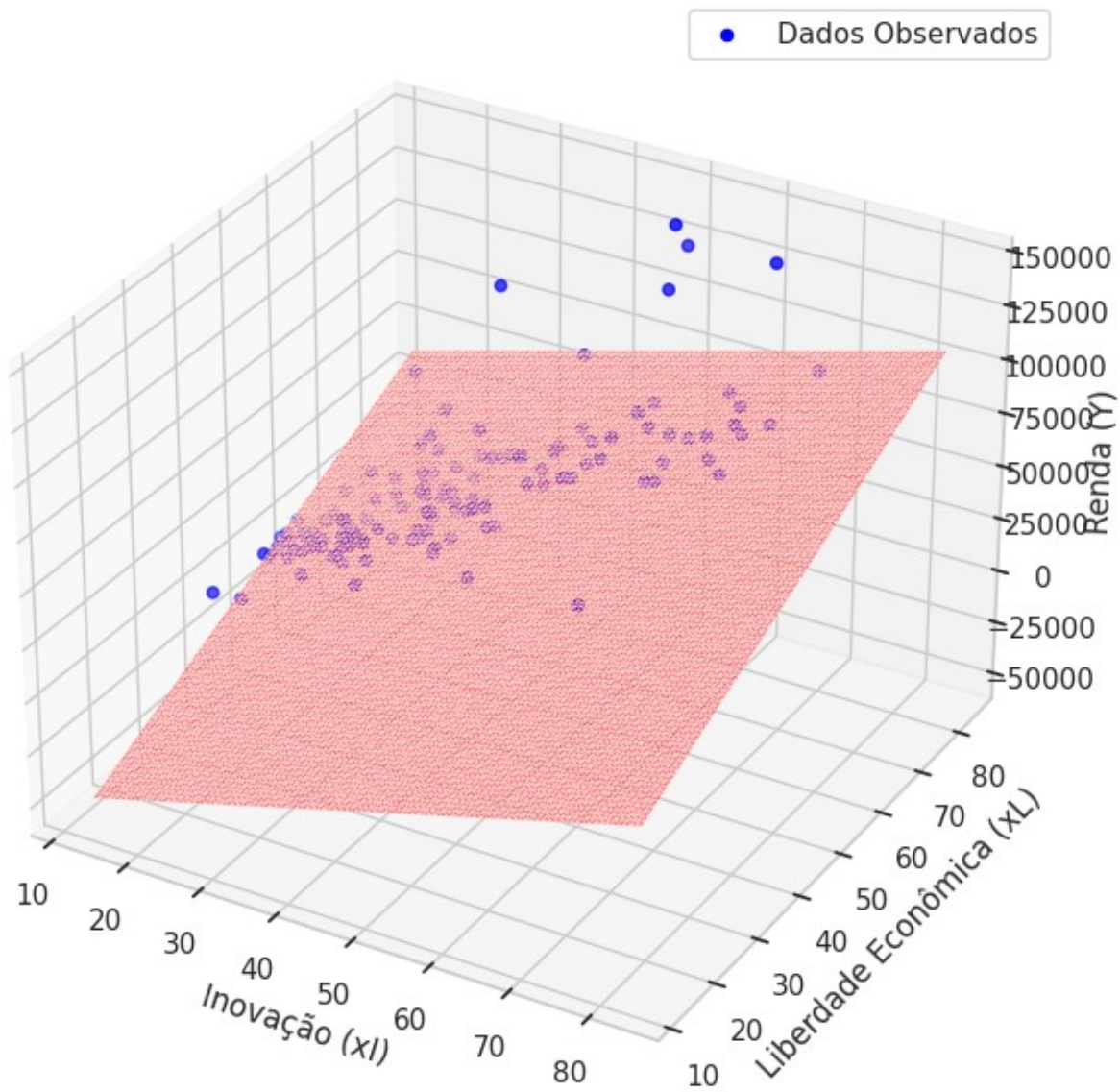
```
# Plotagem da curva de regressão
fig = plt.figure(figsize=(12, 8))
ax = fig.add_subplot(111, projection='3d')

ax.scatter(dados['xI'], dados['xL'], dados['Y'], c='blue', marker='o',
label='Dados Observados')
ax.plot_trisurf(np.ravel(x3_grid[0]), np.ravel(x3_grid[1]),
np.ravel(x3_pred), color='red', alpha=0.5, linewidth=0.2)

ax.set_xlabel('Inovação (xI)')
ax.set_ylabel('Liberdade Econômica (xL)')
ax.set_zlabel('Renda (Y)')
ax.set_title('Regressão: Inovação e Liberdade Econômica vs. Renda')

plt.legend()
plt.show()
```

Regressão: Inovação e Liberdade Econômica vs. Renda



Interpretando os dados da regressão múltipla entre inovação + liberdade econômica vs. Renda per capita ppc:

Aqui estão os principais pontos interpretativos:

1 - Coeficientes (coef): São os valores estimados para os coeficientes da regressão. No seu caso:

- Para a constante (intercepto): -76530 (aproximadamente -76,530).

- Para x_I (Inovação): 961.6140.

- Para x_L (Liberdade Econômica): 1245.6609.

2 - $P > |t|$: Indica o valor p associado ao teste de hipótese nula de que o coeficiente é igual a zero (nenhum efeito). Valores p pequenos (menores que 0.05) indicam que é improvável que o coeficiente seja zero, sugerindo uma relação estatisticamente significativa com a variável de resposta (Renda, Y).

3 - R-quadrado (R-squared): Mede a proporção da variabilidade na variável dependente (Renda, Y) que é explicada pelas variáveis independentes (Inovação e Liberdade Econômica). Um R-quadrado de 0.673 significa que aproximadamente 67.3% da variabilidade na Renda é explicada pelas variáveis explicativas no modelo.

4 - F-statistic: Teste de significância global do modelo. Um valor alto (129.6 neste caso) com um valor p baixo (2.69×10^{-31}) sugere que pelo menos uma das variáveis explicativas é significativa para explicar a variabilidade em Y.

5 - Omnibus, Skewness, Kurtosis: Testes de normalidade dos resíduos. Aqui, o valor baixo de Prob(Omnibus) (0.000) sugere que os resíduos não são normalmente distribuídos, o que pode indicar que o modelo pode não capturar completamente todos os padrões nos dados.

Em resumo, o modelo indica que tanto a Inovação quanto a Liberdade Econômica têm impacto significativo na variável de resposta Renda, com um bom ajuste geral do modelo (R-quadrado de 0.673).

Teste de Hipóteses do Modelo de Regressão Linear:

Vou testar uma hipótese do Modelo de Regressão Linear chamada de Multicolinearidade - ela pressupõe que pode ocorrer um viés no modelo caso aja uma grande correlação linear entre as variáveis explicativas

O não cumprimento dessa hipótese, falando resumidamente, é que o não cumprimento dele pode tornar muitas vezes as estimativas dos coeficientes dos parâmetros (β 's) insignificantes.

Carrego essa funcionalidade da biblioteca statsmodels para testar se a hipótese de Multicolinearidade está sendo cumprida

```
from statsmodels.stats.outliers_influence import  
variance_inflation_factor  
  
df = pd.DataFrame(dados)
```

```

# Adicione a constante ao modelo
X = sm.add_constant(df[['xI', 'xL']])

# Calcula o VIF para cada variável explicativa
vif_data = pd.DataFrame()
vif_data['Variavel'] = X.columns
vif_data['VIF'] = [variance_inflation_factor(X.values, i) for i in
range(X.shape[1])]

print(vif_data)

```

	Variavel	VIF
0	const	55.638083
1	xI	2.211582
2	xL	2.211582

Análise do resultado da VIF:

A constante (intercepto) apresenta um VIF de 55.638083. Isso indica que há uma alta correlação entre a constante e as outras variáveis explicativas no modelo.

No entanto, percebe-se que os VIFs das variáveis explicativas x_1 e x_L , que são de aproximadamente 2.21 cada, estão abaixo de 5. VIFs abaixo de 5 geralmente indicam que a multicolinearidade não é um problema significativo para essas variáveis, o que é um bom sinal.

Com base nesse cálculo pôde-se concluir que a Multicolinearidade não é um problema dos grandes aqui.

Conclusão do Trabalho:

A despeito de todas as limitações desse estudo, tem-se que os resultados obtidos destacam que tanto a Inovação quanto a Liberdade Econômica têm impactos positivos significativos na determinação da renda, com a inovação aparentando ter um impacto ainda maior. Este estudo também revelou que a combinação desses dois fatores no mesmo modelo explica de forma mais abrangente a variabilidade na renda, evidenciando a importância desses elementos sendo utilizados em conjunto e de forma inteligente para a dinâmica econômica.

Os dados corroboram a ideia de que a inovação é crucial e que um ambiente econômico baseado no livre mercado é fundamental para um país que deseja melhorar o nível de renda per capita. Esses resultados sustentam o ideal liberal de que a liberdade econômica é relevante e tem um impacto predominantemente positivo para uma nação. Além disso, reforçam um conceito amplamente aceito entre os economistas sobre os efeitos benéficos da inovação na economia.

Portanto, os achados deste breve estudo são significativos e podem gerar insights valiosos para a elaboração de políticas públicas e estratégias de desenvolvimento que visem aumentar a prosperidade econômica.

END.