

Exploring Worldwide Petrol/Gas Prices: An Analytical Journey

In this notebook, our exploration of worldwide petrol/gas prices will be guided by a structured approach:

Importing Data Science Libraries: We'll begin by importing the necessary libraries such as pandas, matplotlib, and seaborn to aid in data manipulation and visualization.

Loading the Dataset: Next, we'll load the dataset containing information about petrol/gas prices worldwide, world share, and consumption per capita.

Exploratory Data Analysis (EDA): Through EDA, we'll delve into the dataset to gain insights and understanding. This involves examining statistical summaries, visualizing distributions, and identifying any patterns or anomalies.

Overall, by following this structured approach, we aim to uncover meaningful insights into worldwide petrol/gas prices and related factors, shedding light on global consumption patterns and economic indicators.

Importing the necessary libraries

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns
import plotly.express as px
import plotly.graph_objects as go
```

Loading the Dataset

```
In [2]: df = pd.read_csv(r"C:\Users\hp\Desktop\Project 2\Petrol Dataset June 23 2022 -- Version 2.csv", encoding='latin1')

Out[2]:
```

	S#	Country	Daily Oil Consumption (Barrels)	World Share	Yearly Gallons Per Capita	Price Per Gallon (USD)	Price Per Liter (USD)	Price Per Liter (PKR)	GDP Per Capita (USD)	Gallons GDP Per Capita Can Buy	xTimes Yearly Gallons Per Capita Buy
0	1	United States	19,687,287	20%	934.3	5.19	1.37	289.97	63,414	12,218	13
1	2	China	12,791,553	13%	138.7	5.42	1.43	302.87	10,435	1,925	14
2	3	India	4,443,000	5%	51.4	5.05	1.33	281.93	1,901	376	7
3	4	Japan	4,012,877	4%	481.5	4.69	1.24	262.05	40,193	8,570	18
4	5	Russia	3,631,287	4%	383.2	3.41	0.90	190.56	10,127	2,970	8
...
176	177	Belize	4,001	0%	166.5	6.68	1.76	373.09	4,436	664	4
177	178	Niue	51	0%	484.4	11.43	3.02	638.73	15,586	1,364	3
178	179	Saint Pierre & Miquelon	660	0%	1705.1	8.27	2.19	462.13	34,900	4,220	2
179	180	Montserrat	400	0%	1231.1	4.57	1.21	255.07	12,589	2,755	2
180	181	Tonga	899	0%	136.3	16.20	4.28	905.22	4,903	303	2

181 rows x 11 columns

Exploratory Data Analysis

In this step we will understand our data by going into its deep.

```
In [3]: df.head()

Out[3]:
```

	S#	Country	Daily Oil Consumption (Barrels)	World Share	Yearly Gallons Per Capita	Price Per Gallon (USD)	Price Per Liter (USD)	Price Per Liter (PKR)	GDP Per Capita (USD)	Gallons GDP Per Capita Can Buy	xTimes Yearly Gallons Per Capita Buy
0	1	United States	19,687,287	20%	934.3	5.19	1.37	289.97	63,414	12,218	13
1	2	China	12,791,553	13%	138.7	5.42	1.43	302.87	10,435	1,925	14
2	3	India	4,443,000	5%	51.4	5.05	1.33	281.93	1,901	376	7
3	4	Japan	4,012,877	4%	481.5	4.69	1.24	262.05	40,193	8,570	18
4	5	Russia	3,631,287	4%	383.2	3.41	0.90	190.56	10,127	2,970	8

```
In [4]: df.tail()

Out[4]:
```

	S#	Country	Daily Oil Consumption (Barrels)	World Share	Yearly Gallons Per Capita	Price Per Gallon (USD)	Price Per Liter (USD)	Price Per Liter (PKR)	GDP Per Capita (USD)	Gallons GDP Per Capita Can Buy	xTimes Yearly Gallons Per Capita Buy
176	177	Belize	4,001	0%	166.5	6.68	1.76	373.09	4,436	664	4
177	178	Niue	51	0%	484.4	11.43	3.02	638.73	15,586	1,364	3
178	179	Saint Pierre & Miquelon	660	0%	1705.1	8.27	2.19	462.13	34,900	4,220	2
179	180	Montserrat	400	0%	1231.1	4.57	1.21	255.07	12,589	2,755	2
180	181	Tonga	899	0%	136.3	16.20	4.28	905.22	4,903	303	2

```
In [5]: rows, columns = df.shape
print(f'There are {rows} rows and {columns} columns in petrol/gas price worldwide dataset')

There are 181 rows and 11 columns in petrol/gas price worldwide dataset

In [6]: df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 181 entries, 0 to 180
Data columns (total 11 columns):
#   Column                Non-Null Count  Dtype  ---
0   S#                     181 non-null   int64  
1   Country               181 non-null   object 
2   Daily Oil Consumption (Barrels)  181 non-null   object 
3   World Share           181 non-null   object 
4   Yearly Gallons Per Capita  181 non-null   float64 
5   Price Per Gallon (USD)    181 non-null   float64 
6   Price Per Liter (USD)     181 non-null   float64 
7   Price Per Liter (PKR)     181 non-null   float64 
8   GDP Per Capita ( USD )    181 non-null   object 
9   Gallons GDP Per Capita Can Buy  181 non-null   object 
10  xTimes Yearly Gallons Per Capita Buy  181 non-null   int64  
dtypes: float64(4), int64(2), object(5)
memory usage: 15.7+ KB
```

Statistical Summary

We check descriptive statistics of our data which summarize include central tendency, dispersion and shape of a dataset's distribution.

```
In [7]: df.describe()

Out[7]:
```

	S#	Yearly Gallons Per Capita	Price Per Gallon (USD)	Price Per Liter (USD)	Price Per Liter (PKR)	xTimes Yearly Gallons Per Capita Buy
count	181.000000	181.000000	181.000000	181.000000	181.000000	181.000000
mean	91.000000	332.06630	5.695691	1.505138	318.219227	14.204420
std	52.394338	436.558735	4.370484	1.154575	244.192081	48.613866
min	1.000000	2.200000	0.080000	0.020000	4.650000	1.000000
25%	46.000000	53.900000	4.150000	1.100000	232.020000	6.000000
50%	91.000000	180.700000	5.280000	1.400000	295.940000	9.000000
75%	136.000000	424.600000	6.760000	1.790000	377.740000	12.000000
max	181.000000	3679.500000	54.890000	14.500000	3066.750000	654.000000

- The minimum fuel prices per liter (PKR) is 4.65 rupees

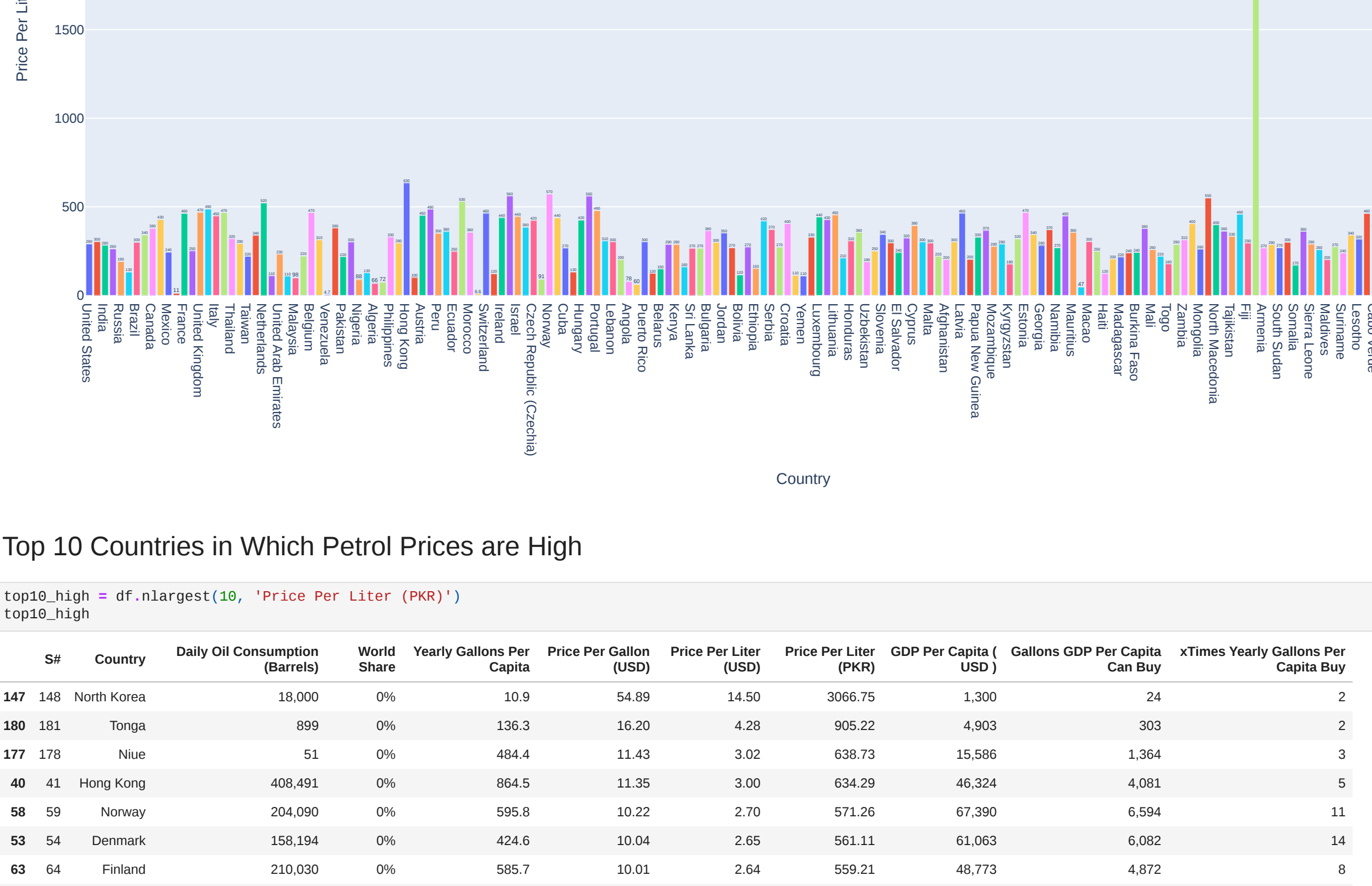
- The maximum fuel prices per liter (PKR) is 3067 rupees

- The average fuel price is 318 rupees

Per Liter (PKR) in each country (Interactive Visualization)

```
In [8]: fig = px.bar(df, y='Price Per Liter (PKR)', x='Country', text_auto='.2s', color = 'Country',
               title='Price Per Liter (PKR) in each country',
               hover_data = ['GDP Per Capita ( USD )', 'Gallons GDP Per Capita Can Buy', 'xTimes Yearly Gallons Per Capita Buy'],)

fig.update_traces(textfont_size=12, textangle=0, textposition='outside', cliponaxis=False)
fig.update_layout(
    height = 800,
    width=1600
)
fig.show()
```



Top 10 Countries in Which Petrol Prices are High

```
In [9]: top10_high = df.nlargest(10, 'Price Per Liter (PKR)')

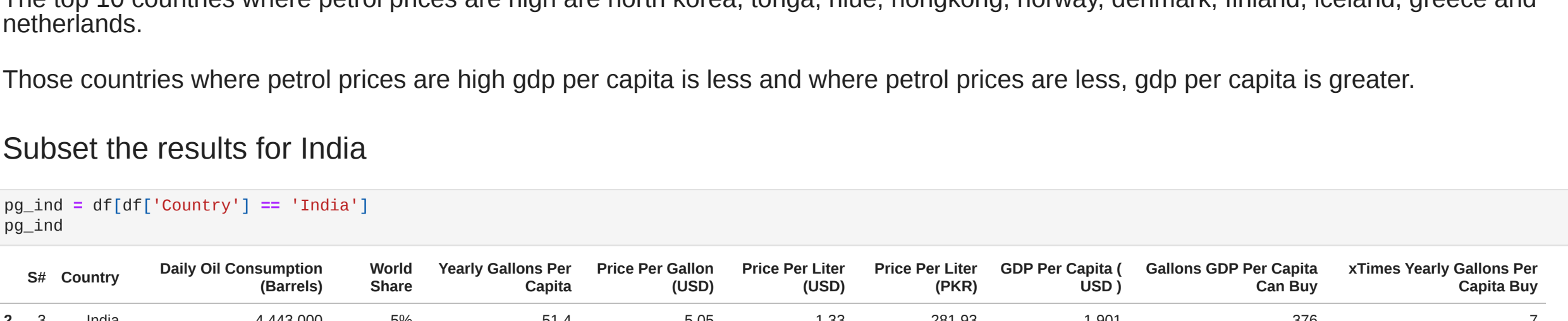
Out[9]:
```

	S#	Country	Daily Oil Consumption (Barrels)	World Share	Yearly Gallons Per Capita	Price Per Gallon (USD)	Price Per Liter (USD)	Price Per Liter (PKR)	GDP Per Capita (USD)	Gallons GDP Per Capita Can Buy	xTimes Yearly Gallons Per Capita Buy
147	148	North Korea	18,000	0%	10.9	54.89	14.50	3066.75	1,300	24	2
180	181	Tonga	899	0%	136.3	16.20	4.28	905.22	4,903	303	2
177	178	Niue	51	0%	484.4	11.43	3.02	638.73	15,586	1,364	3
40	41	Hong Kong	408,491	0%	864.5	11.35	3.00	634.29	46,324	4,081	5
58	59	Norway	204,090	0%	595.8	10.22	2.70	571.26	67,390	6,594	11
53	54	Denmark	158,194	0%	424.6	10.04	2.65	561.11	61,063	6,082	14
64	64	Finland	210,030	0%	585.7	10.01	2.64	559.21	48,773	4,872	8
142	142	Iceland	19,090	0%	880.9	9.83	2.60	549.48	59,270	6,030	7
47	48	Greece	296,101	0%	427.6	9.49	2.51	530.02	17,623	1,857	4
22	23	Netherlands	937,098	1%	846.0	9.33	2.47	521.35	52,397	5,616	7

Interactive visualization of top 10 countries

```
In [10]: fig = px.bar(top10_high, x='Country', y='Price Per Liter (USD)',
                    hover_data = ['GDP Per Capita ( USD )', 'Gallons GDP Per Capita Can Buy', 'xTimes Yearly Gallons Per Capita Buy'],
                    color = 'Country')

fig.show()
```



The top 10 countries where petrol prices are high are north korea, tonga, niue, hongkong, norway, denmark, finland, iceland, greece and netherlands.

Those countries where petrol prices are high gdp per capita is less and where petrol prices are less, gdp per capita is greater.

Subset the results for India

```
In [11]: pg_ind = df[df['Country'] == 'India']

Out[11]:
```

	S#	Country	Daily Oil Consumption (Barrels)	World Share	Yearly Gallons Per Capita	Price Per Gallon (USD)	Price Per Liter (USD)	Price Per Liter (PKR)	GDP Per Capita (USD)	Gallons GDP Per Capita Can Buy	xTimes Yearly Gallons Per Capita Buy
2	3	India	4,443,000	5%	51.4	5.05	1.33	281.93	1,901	376	7

- India has the daily oil consumption in barrels is 4,443,000

- India can buy the petrol about 51.4 yearly gallons per capita

- India GDP per Capita is 1,901 dollars

- India can buy the fuel 7 times years gallons per capita

Comparison of fuel prices in Pakistan with US, China, India, and Bangladesh

```
In [12]: # Make a subset of countries
pak = df[df['Country'] == 'Pakistan']
us = df[df['Country'] == 'United States']
china = df[df['Country'] == 'China']
india = df[df['Country'] == 'India']
bangladesh = df[df['Country'] == 'Bangladesh']

# Concatenate to make a DataFrame to compare the countries
df_concat = pd.concat([pak, us, china, india, bangladesh])
df_concat
```

```
Out[12]:
```

	S#	Country	Daily Oil Consumption (Barrels)	World Share	Yearly Gallons Per Capita	Price Per Gallon (USD)	Price Per Liter (USD)	Price Per Liter (PKR)	GDP Per Capita (USD)	Gallons GDP Per Capita Can Buy	xTimes Yearly Gallons Per Capita Buy
32	33	Pakistan	556,000	1%	41.9	3.90	1.03	217.85	1,194	306	7
0	1	United States	19,687,287	20%	934.3	5.19	1.37	289.97	63,414	12,218	13
1	2	China	12,791,553	13%	138.7	5.42	1.43	302.87	10,435	1,925	14
2	3	India	4,443,000	5%	51.4	5.05	1.33	281.93	1,901	376	7
67	68	Bangladesh	113,000	0%	11.0	3.50	0.95	200.50	1,969	548	50

```
In [13]: fig = px.bar(df_concat, y='Price Per Liter (PKR)', x='Country', text_auto='.2s', color = 'Country',
               title='Comparison of fuel prices',
               hover_data = ['GDP Per Capita ( USD )', 'Gallons GDP Per Capita Can Buy', 'xTimes Yearly Gallons Per Capita Buy'],)

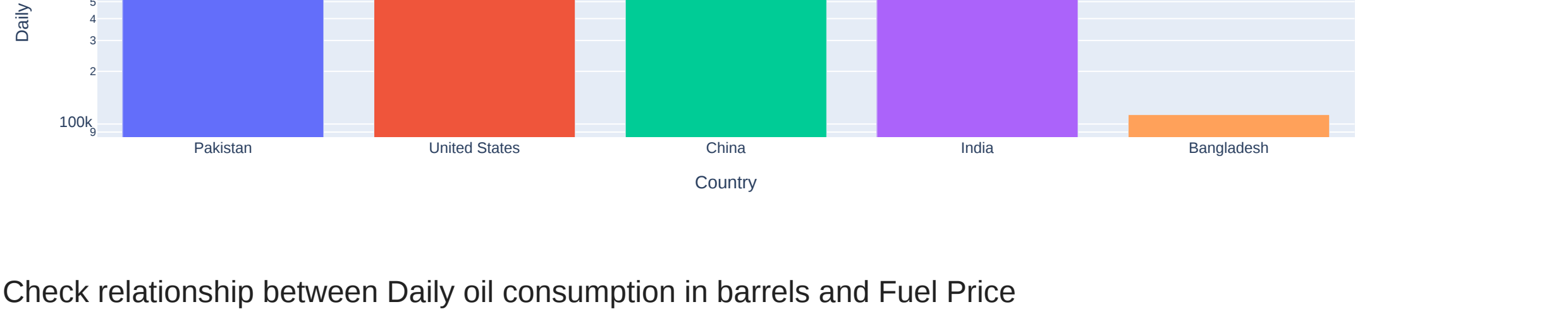
fig.update_traces(textfont_size=12, textangle=0, textposition='outside', cliponaxis=False)
fig.show()
```



Visualize Daily Oil Consumptions Of above Countries

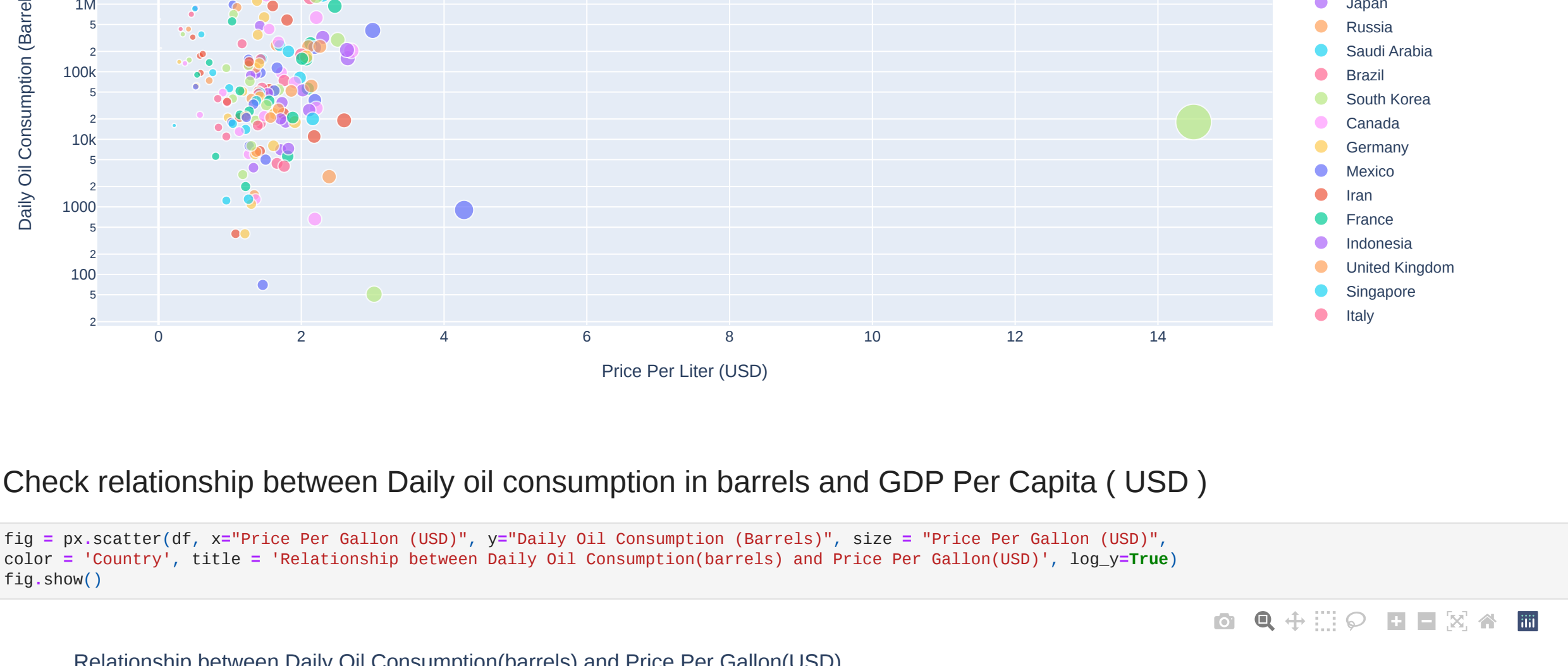
```
In [14]: fig = px.bar(df_concat, y='Daily Oil Consumption (Barrels)', x='Country', color = 'Country',
               title='Daily Oil Consumption In Barrels',
               hover_data = ['GDP Per Capita ( USD )', 'Gallons GDP Per Capita Can Buy', 'xTimes Yearly Gallons Per Capita Buy'],
               log_y = True)

fig.show()
```



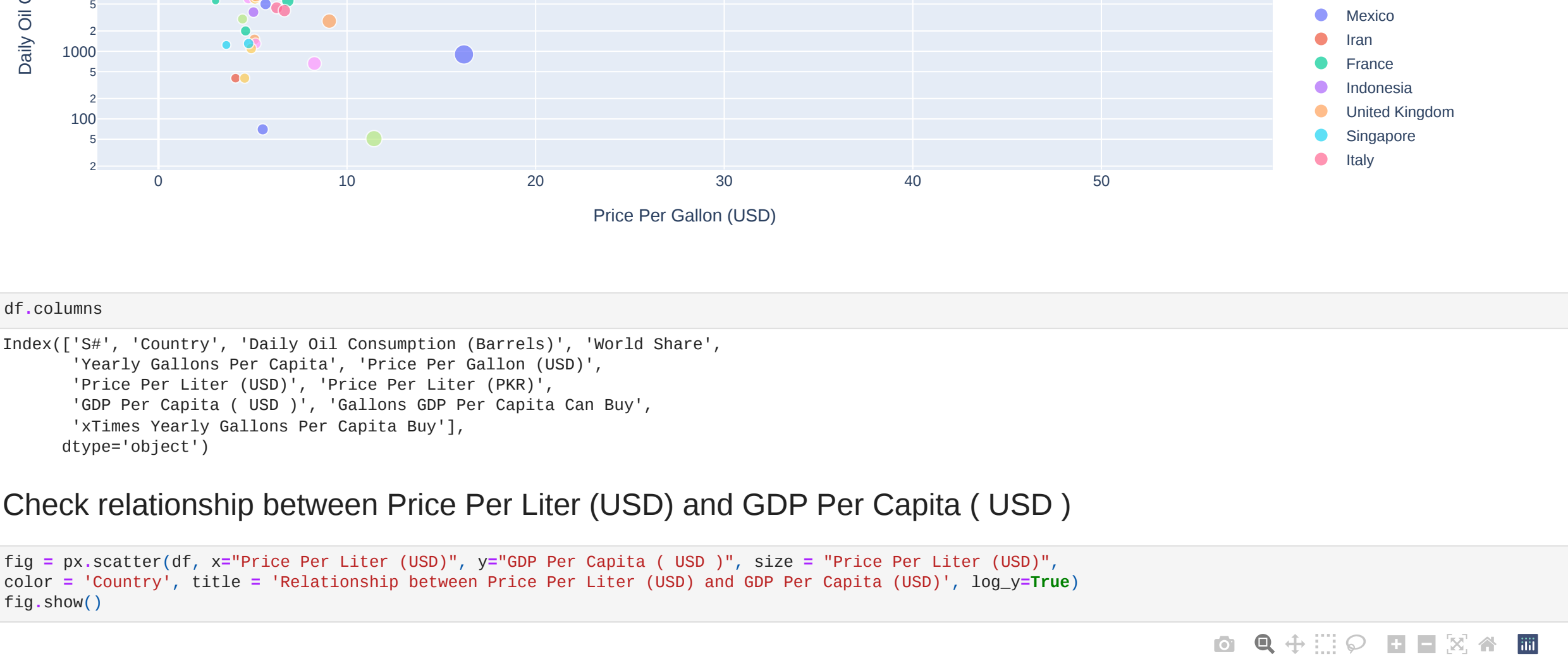
Check relationship between Daily oil consumption in barrels and Fuel Price

```
In [15]: fig = px.scatter(df, x='Price Per Liter (USD)', y='Daily Oil Consumption (Barrels)', size='Price Per Liter (USD)',
               color = 'Country', title = 'Relationship between Daily Oil Consumption(barrels) and Price Per Liter(USD)', log_y=True)
fig.show()
```



Check relationship between Daily oil consumption in barrels and GDP Per Capita (USD)

```
In [16]: fig = px.scatter(df, x='Price Per Gallon (USD)', y='Daily Oil Consumption (Barrels)', size = 'Price Per Gallon (USD)',
               color = 'Country', title = 'Relationship between Daily Oil Consumption(barrels) and Price Per Gallon(USD)', log_y=True)
fig.show()
```

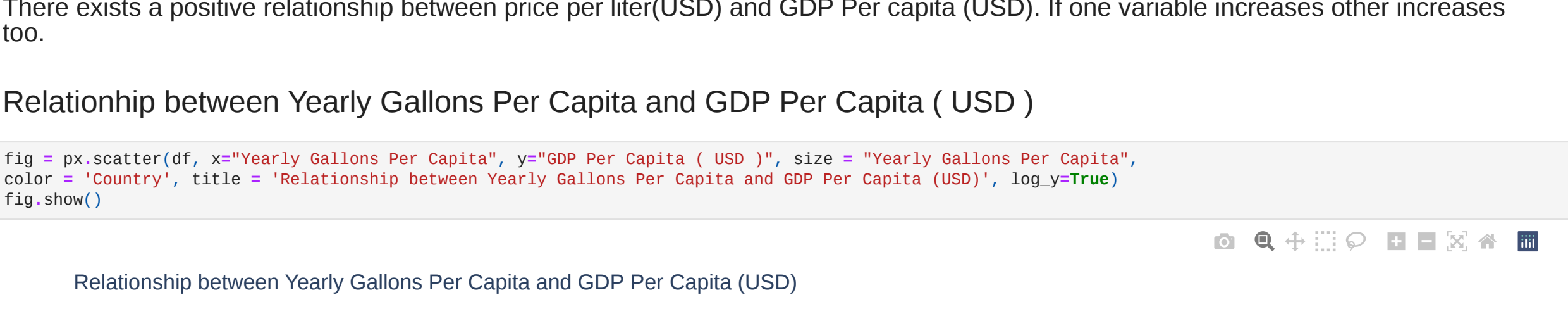


```
In [17]: df.columns

Out[17]:
Index(['S#', 'Country', 'Daily Oil Consumption (Barrels)', 'World Share',
      'Yearly Gallons Per Capita', 'Price Per Gallon (USD)', 'Price Per Liter (USD)',
      'Price Per Liter (PKR)', 'GDP Per Capita ( USD )', 'Gallons GDP Per Capita Can Buy',
      'xTimes Yearly Gallons Per Capita Buy'],
      dtype='object')
```

Check relationship between Price Per Liter (USD) and GDP Per Capita (USD)

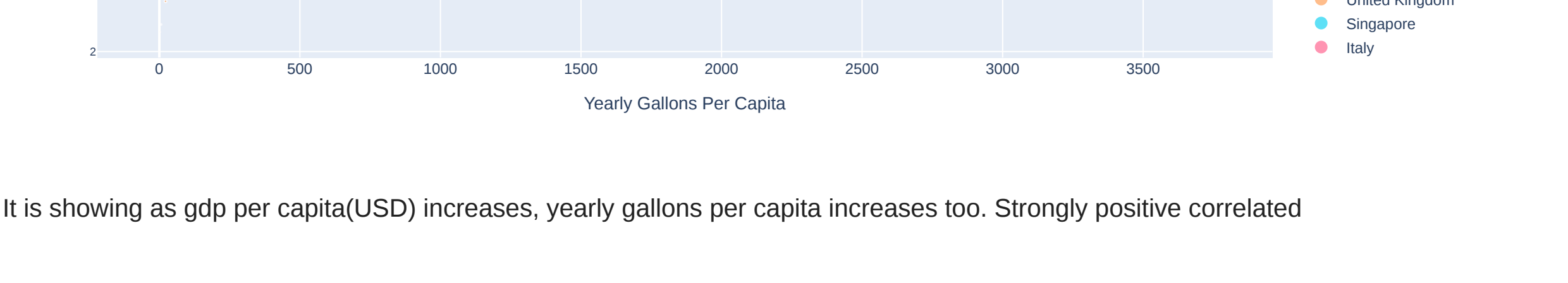
```
In [18]: fig = px.scatter(df, x='Price Per Liter (USD)', y='GDP Per Capita ( USD )', size = 'Price Per Liter (USD)',
               color = 'Country', title = 'Relationship between Price Per Liter (USD) and GDP Per Capita (USD)', log_y=True)
fig.show()
```



There exists a positive relationship between price per liter(USD) and GDP per capita (USD). If one variable increases other increases too.

Relationship between Yearly Gallons Per Capita and GDP Per Capita (USD)

```
In [19]: fig = px.scatter(df, x='Yearly Gallons Per Capita', y='GDP Per Capita ( USD )', size = 'Yearly Gallons Per Capita',
               color = 'Country', title = 'Relationship between Yearly Gallons Per Capita and GDP Per Capita (USD)', log_y=True)
fig.show()
```



It is showing as gdp per capita(USD) increases, yearly gallons per capita increases too. Strongly positive correlated