Covid-19 Impacts Analysis using Python

The Covid-19 pandemic has had a significant impact on virtually every aspect of our lives, from health and economics to social norms and personal relationships. Almost all the countries were impacted negatively by the rise in the cases of Covid- 19.

The first wave of Covid- 19 impacted Global Economic, Health, Education, Politics, environments as the world was never ready for the pandemic. It resulted in a rise in cases, a rise in deaths, a rise in unemployment and a rise in povetry.

This project covers the analysis of the spread of Covid-19 cases and all the impacts of covid-19 on the economy.

Data soource: Kaggle (https://www.kaggle.com/datasets/shashwatwork/impact-of-covid19-pandemic-on-the-global-economy?resource=download)

The data has the following attributes;

- the country code
- name of all the countries
- date of the record
- Human development index of all the countries
- Daily covid-19 cases
- Daily deaths due to covid-19
- stringency index of the countries
- the population of the countries
- GDP per capita of the countries

In [1]:
```
# Importing libraries
import pandas as pd
import plotly.express as px
import plotly.graph_objects as go
```

Importing Covid-19 datasets

In [2]:
```
df = pd.read_csv("transformed_data.csv")
df2 = pd.read_csv("raw_data.csv")
print(df)
```

```
      CODE      COUNTRY        DATE    HDI        TC        TD       STI  \
0      AFG  Afghanistan  2019-12-31  0.498  0.000000  0.000000  0.000000
1      AFG  Afghanistan  2020-01-01  0.498  0.000000  0.000000  0.000000
2      AFG  Afghanistan  2020-01-02  0.498  0.000000  0.000000  0.000000
3      AFG  Afghanistan  2020-01-03  0.498  0.000000  0.000000  0.000000
4      AFG  Afghanistan  2020-01-04  0.498  0.000000  0.000000  0.000000
...    ...          ...         ...    ...       ...       ...       ...
50413  ZWE     Zimbabwe  2020-10-15  0.535  8.994048  5.442418  4.341855
50414  ZWE     Zimbabwe  2020-10-16  0.535  8.996528  5.442418  4.341855
50415  ZWE     Zimbabwe  2020-10-17  0.535  8.999496  5.442418  4.341855
50416  ZWE     Zimbabwe  2020-10-18  0.535  9.000853  5.442418  4.341855
50417  ZWE     Zimbabwe  2020-10-19  0.535  9.005405  5.442418  4.341855

             POP    GDPCAP
0      17.477233  7.497754
1      17.477233  7.497754
2      17.477233  7.497754
3      17.477233  7.497754
4      17.477233  7.497754
...          ...       ...
```

```
50413  16.514381  7.549491
50414  16.514381  7.549491
50415  16.514381  7.549491
50416  16.514381  7.549491
50417  16.514381  7.549491

[50418 rows x 9 columns]
```

we have 50418 rows and 9 columns. It contains the data on covid-19 cases and their impact on GDP from December 31, 2019, to October 10, 2020.

Data Preparation

there are two data files, Using both the dataset for this analysis, because they both contain equally vital information in different columns.

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

Out[3]:

	CODE	COUNTRY	DATE	HDI	TC	TD	STI	POP	GDPCAP
0	AFG	Afghanistan	2019-12-31	0.498	0.0	0.0	0.0	17.477233	7.497754
1	AFG	Afghanistan	2020-01-01	0.498	0.0	0.0	0.0	17.477233	7.497754
2	AFG	Afghanistan	2020-01-02	0.498	0.0	0.0	0.0	17.477233	7.497754
3	AFG	Afghanistan	2020-01-03	0.498	0.0	0.0	0.0	17.477233	7.497754
4	AFG	Afghanistan	2020-01-04	0.498	0.0	0.0	0.0	17.477233	7.497754

In [4]:
```
df2.head()
```

Out[4]:

	iso_code	location	date	total_cases	total_deaths	stringency_index	population	gdp_per_capita	human_development_index	Unnamed: 9	Unnamed: 10	Unnamed: 11	Unnamed: 12	Unnamed: 13
0	AFG	Afghanistan	2019-12-31	0.0	0.0	0.0	38928341	1803.987	0.498	#NUM!	#NUM!	#NUM!	17.477233	7.497754494
1	AFG	Afghanistan	2020-01-01	0.0	0.0	0.0	38928341	1803.987	0.498	#NUM!	#NUM!	#NUM!	17.477233	7.497754494
2	AFG	Afghanistan	2020-01-02	0.0	0.0	0.0	38928341	1803.987	0.498	#NUM!	#NUM!	#NUM!	17.477233	7.497754494
3	AFG	Afghanistan	2020-01-03	0.0	0.0	0.0	38928341	1803.987	0.498	#NUM!	#NUM!	#NUM!	17.477233	7.497754494
4	AFG	Afghanistan	2020-01-04	0.0	0.0	0.0	38928341	1803.987	0.498	#NUM!	#NUM!	#NUM!	17.477233	7.497754494

After having first views of the datasets, there is a need to consolidate both datasets by created a new dataset.

Checking counts of countries available in the dataset.

In [5]:
```
df['COUNTRY'].value_counts()
```

Out[5]:
```
Afghanistan        294
Indonesia          294
Macedonia          294
Luxembourg         294
Lithuania          294
                  ...
Tajikistan         172
Comoros            171
Lesotho            158
Hong Kong           51
Solomon Islands      4
Name: COUNTRY, Length: 210, dtype: int64
```

We can say countries like Afghannistan, Indonesia, Maceconia, Luxembourg, Lithuania recorded high number of cases. But wanna confirm by getting the mode.

In [6]:
```
df['COUNTRY'].value_counts().mode()
```

Out[6]:
```
0    294
dtype: int64
```

I was right with 294 as the mode value. Gonna use it to divide the sum of all the smaples related to the human development index, GDP per capital, and the pupolation.

merging the two datasets, by combining necessary columns from both datasets.

In [7]:
```
# Aggregeting the data

code = df["CODE"].unique().tolist()
country = df["COUNTRY"].unique().tolist()
hdi = []
tc = []
td = []
sti = []
population = df["POP"].unique().tolist()
gdp = []

for i in country:
    hdi.append((df.loc[df["COUNTRY"] == i, "HDI"]).sum()/294)
    tc.append((df2.loc[df2["location"] == i, "total_cases"]).sum())
    td.append((df2.loc[df2["location"] == i, "total_deaths"]).sum())
    sti.append((df.loc[df["COUNTRY"] == i, "STI"]).sum()/294)
    population.append((df2.loc[df2["location"] == i, "population"]).sum()/294)

aggregated_data = pd.DataFrame(list(zip(code, country, hdi, tc, td, sti, population)),
                               columns = ["Country Code", "Country", "HDI",
                                          "Total Cases", "Total Deaths",
                                          "Stringency Index", "Population"])
print(aggregated_data.head())
```

```
  Country Code      Country       HDI  Total Cases  Total Deaths  \
0          AFG  Afghanistan  0.498000    5126433.0      165875.0
1          ALB      Albania  0.600765    1071951.0       31056.0
2          DZA      Algeria  0.754000    4893999.0      206429.0
3          AND      Andorra  0.659551     223576.0        9850.0
4          AGO       Angola  0.418952     304005.0       11820.0

   Stringency Index  Population
0          3.049673   17.477233
1          3.005624   14.872537
2          3.195168   17.596309
3          2.677654   11.254996
4          2.965560   17.307957
```

Note: GDP per capital is not included in the column yet. No correct figures for it in the dataset. Probally, it will be better to compute the GDP per capital manually for the countries. But doing that for all the countries will be time consuming, so, selecting a subsmaple from the dataset by selecting the top 10 countries with the highest number of covid-19 cases.

In [8]:
```
# Sorting the data in descending order to get top 10 countries with high cases
data = aggregated_data.sort_values(by=["Total Cases"], ascending=False)
```

Top 10 Countries with highest Covis-19 Cases

In [9]:
```
df = data.head(10)
print(df)
```

```
    Country Code         Country       HDI  Total Cases  Total Deaths  \
200          USA   United States  0.924000  746014098.0    26477574.0
```

```
27          BRA          Brazil  0.759000  425704517.0    14340567.0
90          IND           India  0.640000  407771615.0     7247327.0
157         RUS          Russia  0.816000  132888951.0     2131571.0
150         PER            Peru  0.599490   74882695.0     3020038.0
125         MEX          Mexico  0.774000   74347548.0     7295850.0
178         ESP           Spain  0.887969   73717676.0     5510624.0
175         ZAF    South Africa  0.608653   63027659.0     1357682.0
42          COL        Colombia  0.581847   60543682.0     1936134.0
199         GBR  United Kingdom  0.922000   59475032.0     7249573.0

     Stringency Index  Population
200          3.350949   19.617637
27           3.136028   19.174732
90           3.610552   21.045353
157          3.380088   18.798668
150          3.430126   17.311165
125          3.019289   18.674802
178          3.393922   17.660427
175          3.364333   17.898266
42           3.357923   17.745037
199          3.353883   18.033340
```

Adding two more columns; GDP per capital before pandemic, and during covid-19 pandemic.

In [10]:

```
df["GDP Before Covid"] = [65279.53, 8897.49, 2100.75,
                          11497.65, 7027.61, 9946.03,
                          29564.74, 6001.40, 6424.98, 42354.41]
df["GDP During Covid"] = [63543.58, 6796.84, 1900.71,
                          10126.72, 6126.87, 8346.70,
                          27057.16, 5090.72, 5332.77, 40284.64]
print(df)
```

```
    Country Code         Country       HDI  Total Cases  Total Deaths  \
200          USA   United States  0.924000  746014098.0    26477574.0
27           BRA          Brazil  0.759000  425704517.0    14340567.0
90           IND           India  0.640000  407771615.0     7247327.0
157          RUS          Russia  0.816000  132888951.0     2131571.0
150          PER            Peru  0.599490   74882695.0     3020038.0
125          MEX          Mexico  0.774000   74347548.0     7295850.0
178          ESP           Spain  0.887969   73717676.0     5510624.0
175          ZAF    South Africa  0.608653   63027659.0     1357682.0
42           COL        Colombia  0.581847   60543682.0     1936134.0
199          GBR  United Kingdom  0.922000   59475032.0     7249573.0

     Stringency Index  Population  GDP Before Covid  GDP During Covid
200          3.350949   19.617637          65279.53          63543.58
27           3.136028   19.174732           8897.49           6796.84
90           3.610552   21.045353           2100.75           1900.71
157          3.380088   18.798668          11497.65          10126.72
150          3.430126   17.311165           7027.61           6126.87
125          3.019289   18.674802           9946.03           8346.70
178          3.393922   17.660427          29564.74          27057.16
175          3.364333   17.898266           6001.40           5090.72
42           3.357923   17.745037           6424.98           5332.77
199          3.353883   18.033340          42354.41          40284.64
C:\Users\USER-PC\AppData\Local\Temp/ipykernel_16996/516469135.py:1: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df["GDP Before Covid"] = [65279.53, 8897.49, 2100.75,
C:\Users\USER-PC\AppData\Local\Temp/ipykernel_16996/516469135.py:4: SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
```

```
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
  df["GDP During Covid"] = [63543.58, 6796.84, 1900.71,
```

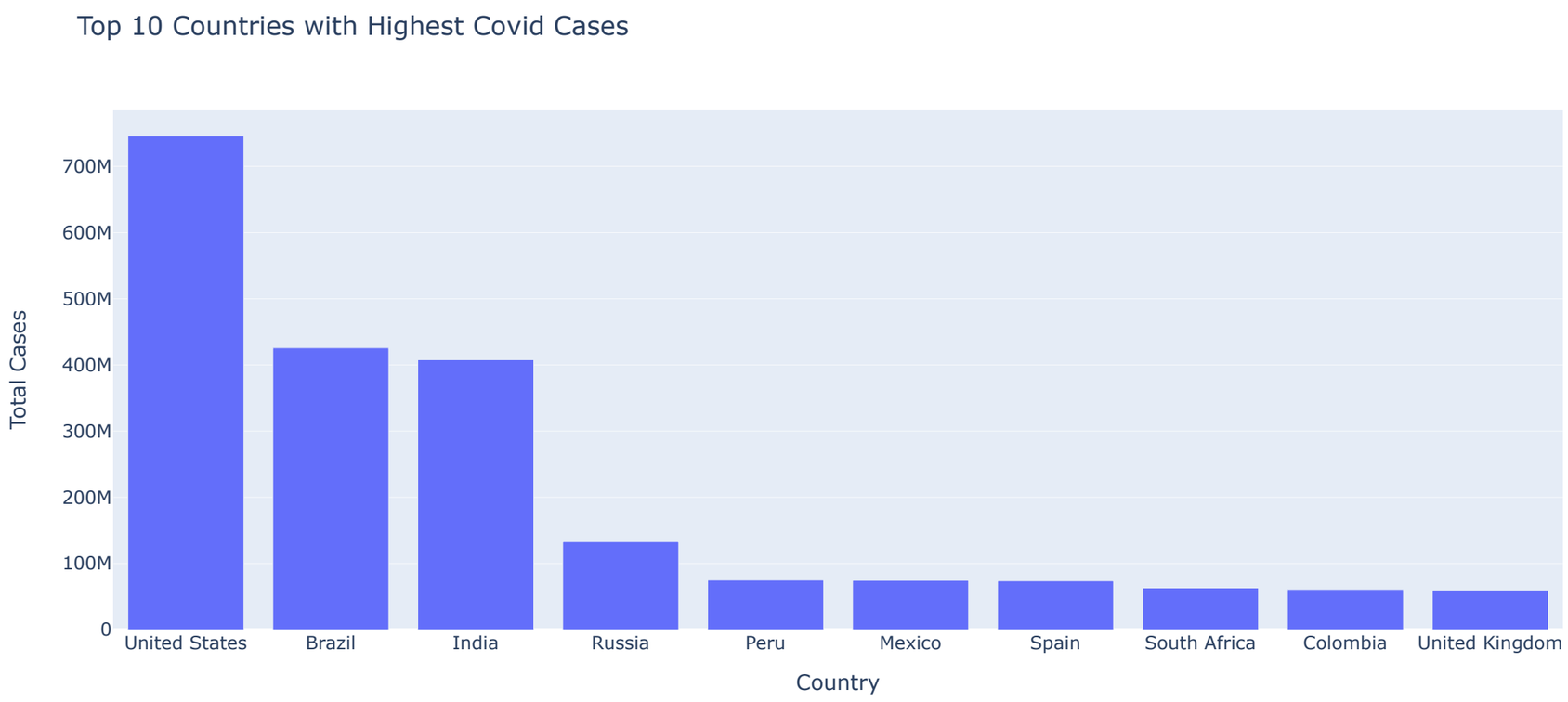
Note; The data about the GDP per Capital is inputted manually. Data source is online

Info Graph Analysis of the spread of Covid-19

Basically, doing my analysis on the selected countries with highest number of recorded cases. so, visualizing the countries.

Decided to use of Plotly tool for my visualization instead of seaborn and matplotlib

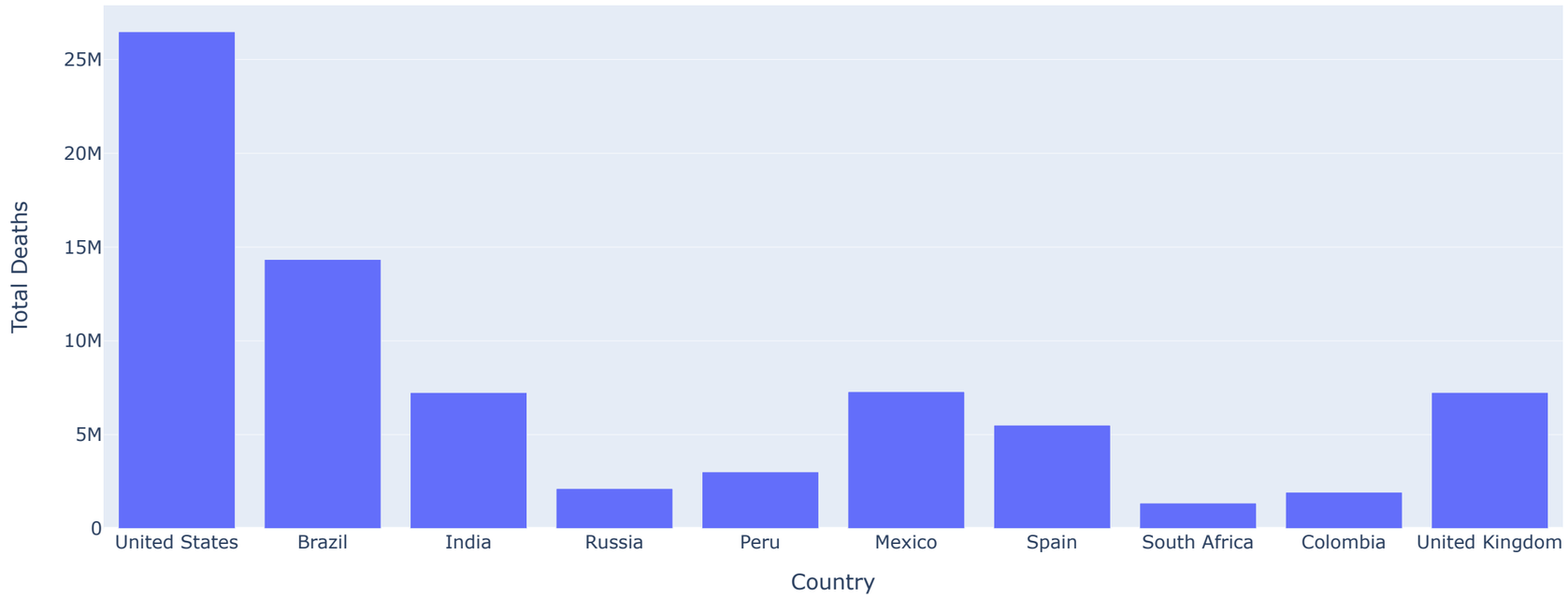
In [11]:
```
fig = px.bar(df, y = 'Total Cases', x= 'Country', title='Top 10 Countries with Highest Covid Cases')
fig.show()
```



From the data, I can see that USA is having a highnumber of recorded covid-19 cases as compared to the rest of the countries.While UK, Columbia have the least number of covid-19 cases among the top 10 countries selected.

Now looking at the total number of dealths among the countries with the highest number of dealths among the top 10 countries with the highest number of covid-19 cases

In [25]:
```
figure = px.bar(df, y='Total Deaths', x='Country',
           title="Countries with Highest Deaths")
figure.show()
```

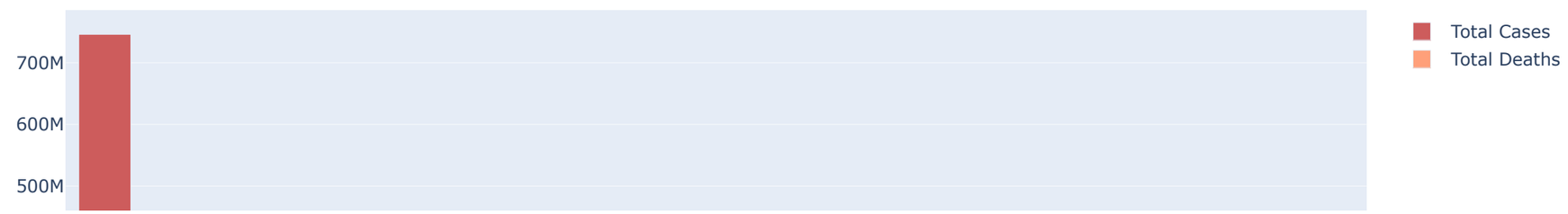


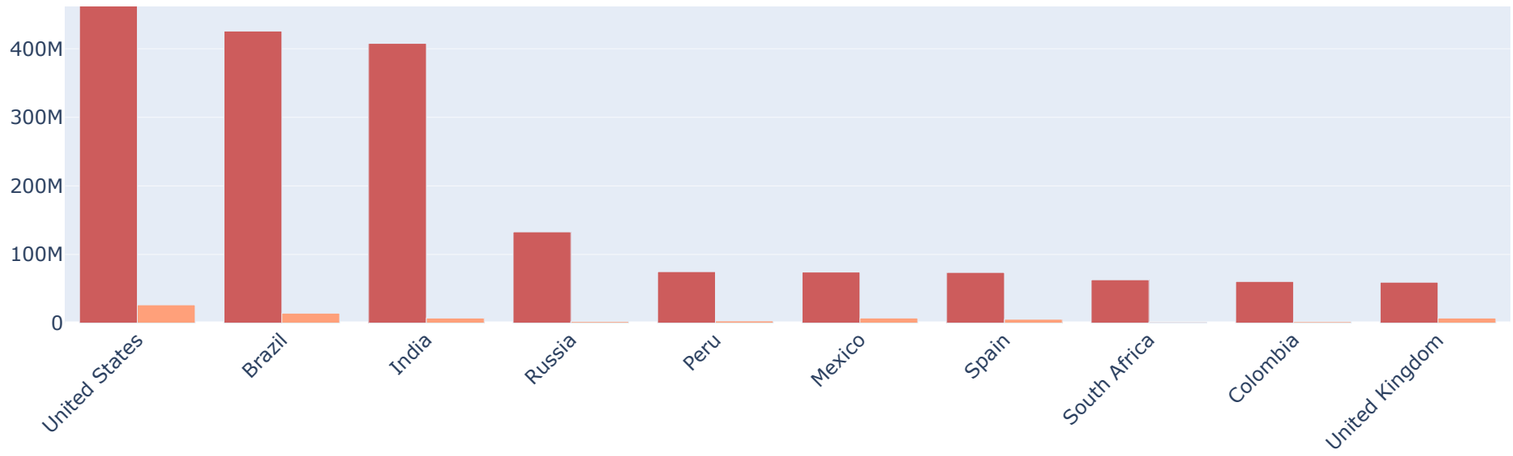
USA still leading in the deaths recorded, with Brazil and India in the second and third positions. One thing to notice here is that the death rate in India, Russia, and South Africa is correlating according to the total number of cases observed on the above chart.

Plotting the total number of cases and total dealths in all these countries together for comparison.

In [13]:

```
fig = go.Figure()
fig.add_trace(go.Bar(
    x=df["Country"],
    y=df["Total Cases"],
    name='Total Cases',
    marker_color='indianred'
))
fig.add_trace(go.Bar(
    x=df["Country"],
    y=df["Total Deaths"],
    name='Total Deaths',
    marker_color='lightsalmon'
))
fig.update_layout(barmode='group', xaxis_tickangle=-45)
fig.show()
```





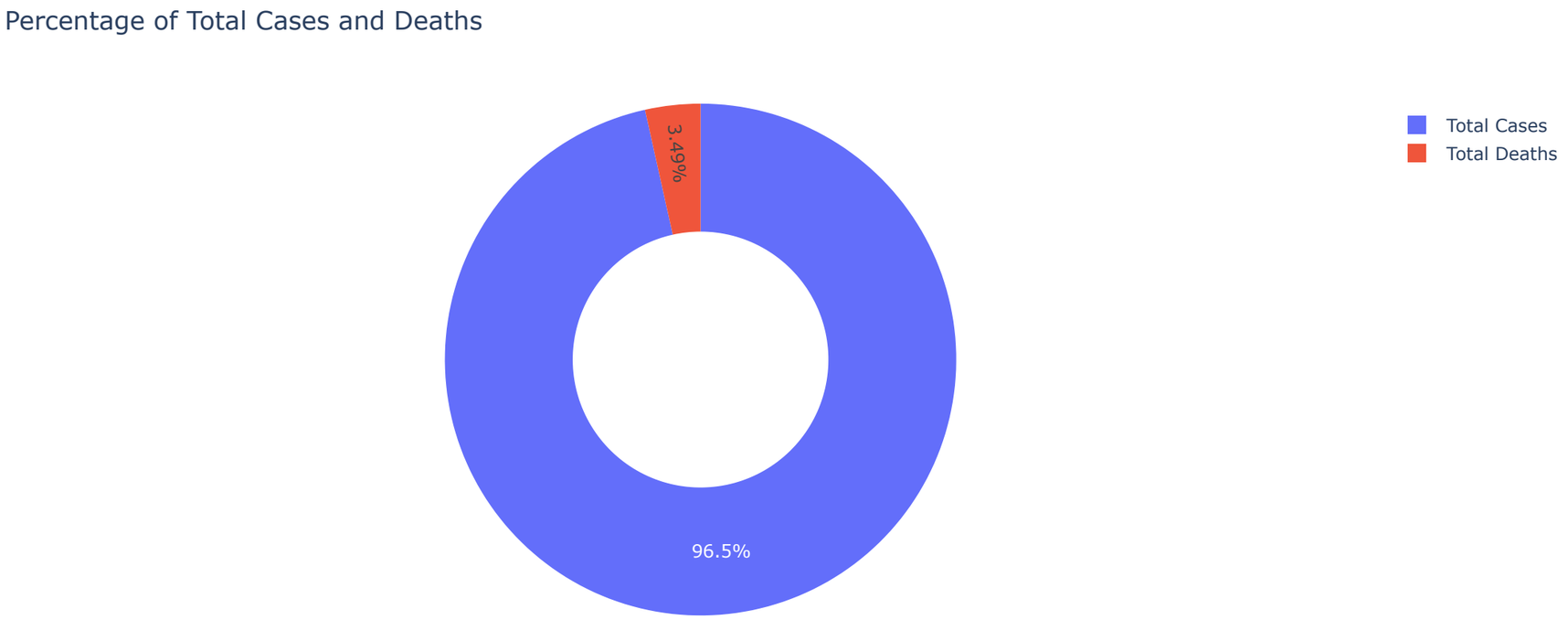
This can lead us to a question that, what happened in Inida, russia and south Africa to have high number of cases but low dealth rates.

This brought up a question to know the percentage of total deaths and total cases among all the countries with highest number covid-19 cases

In [14]:
```
# Percentage of Total Cases and Deaths
cases = df["Total Cases"].sum()
deceased = df["Total Deaths"].sum()

labels = ["Total Cases", "Total Deaths"]
values = [cases, deceased]

fig = px.pie(data, values=values, names=labels,
             title='Percentage of Total Cases and Deaths', hole=0.5)
fig.show()
```



There was only 3.49% of total deaths with 96.5% total cases.

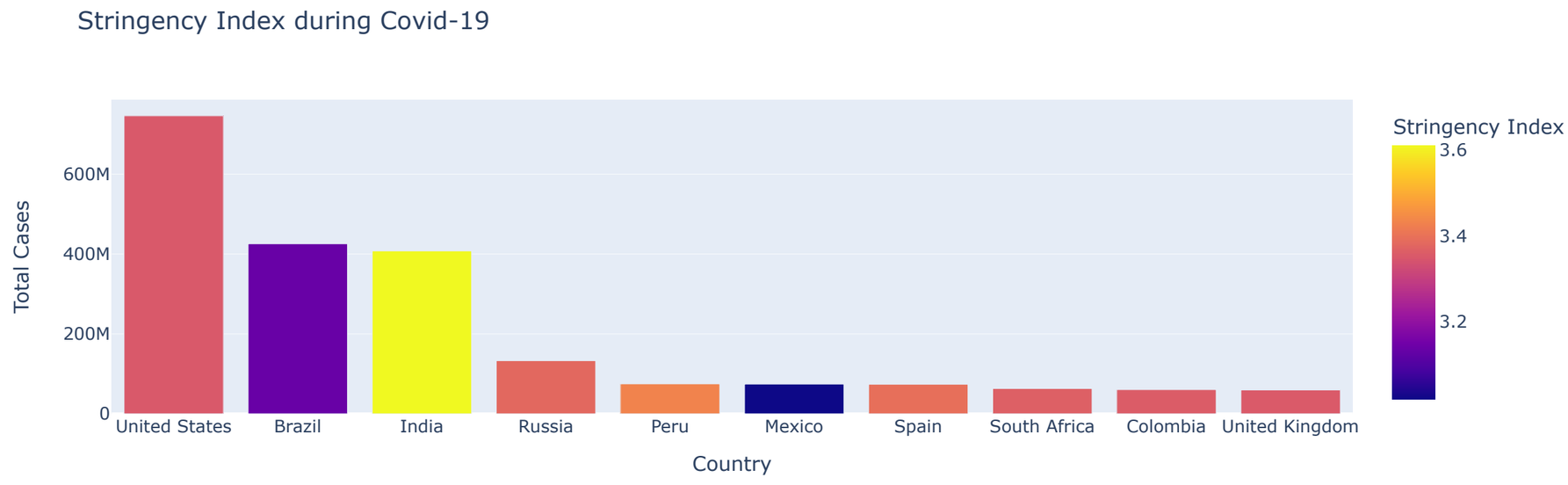
Death rate of Covid-19 cases can be calculated manually below;

In [15]:
```
death_rate = (df["Total Deaths"].sum() / df["Total Cases"].sum()) * 100
print("Death Rate = ", death_rate)
```

```
Death Rate =  3.6144212045653767
```

Stringency index, talks about response indicators measurement. Including schood closures, workplace closures, and travel bans. It measures how the countries followed these precautions to control the spread of Covid-19

In [16]:
```
fig = px.bar(df, x='Country', y='Total Cases',
             hover_data=['Population', 'Total Deaths'],
             color='Stringency Index', height=400,
             title= "Stringency Index during Covid-19")
fig.show()
```

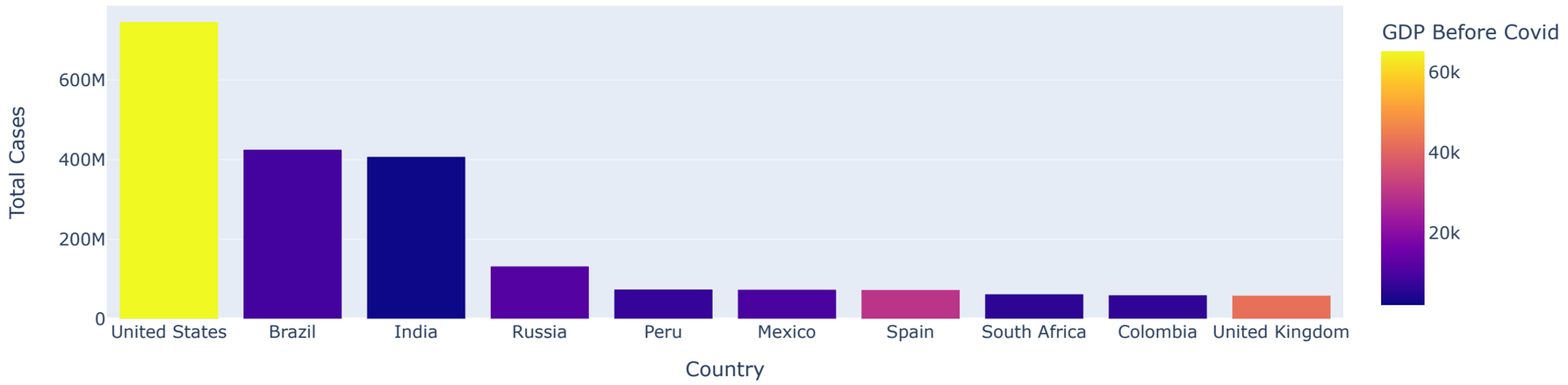


This shows India performed well in the taken strictly measure to avoid the spread of covid-19. It has the highest stringency index of 3.6

Analyzing the Impacts on Global Economy.

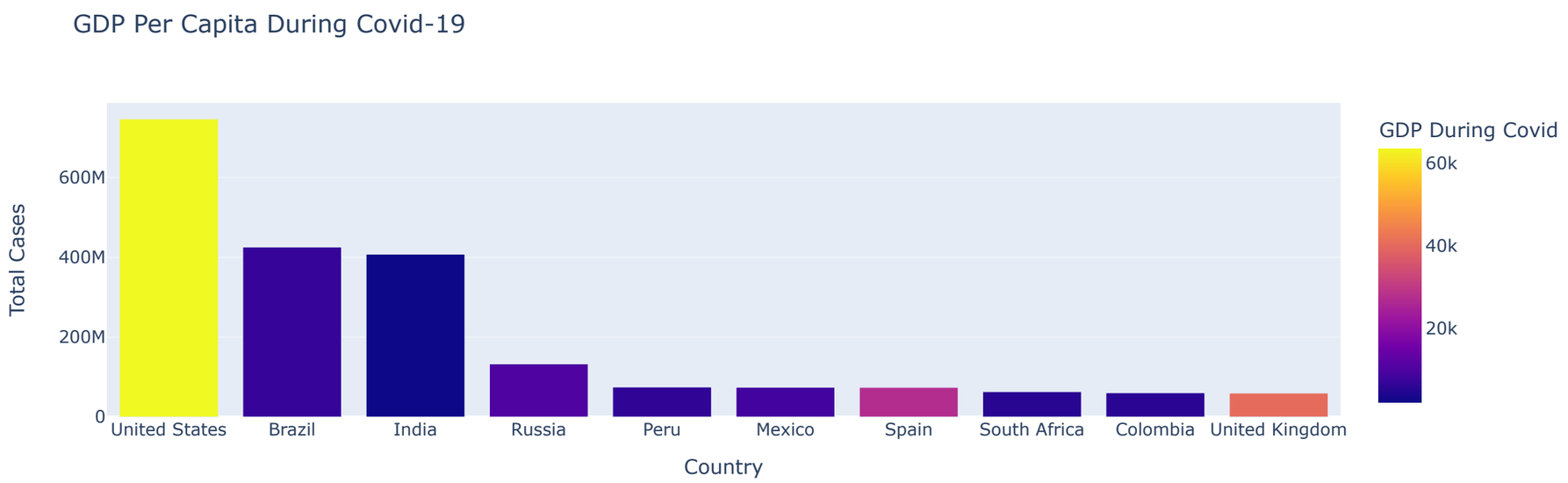
As GDP per capital is the only primary factor for analyzing the economic retrogression caused due to the covid-19 pandemic. Visually analyzing GDp before and during the outbreak among the countries selected.

In [17]:
```
fig = px.bar(df, x='Country', y='Total Cases',
             hover_data=['Population', 'Total Deaths'],
             color='GDP Before Covid', height=400,
             title="GDP Per Capita Before Covid-19")
fig.show()
```



checking for GDP per capital during the outbreak;

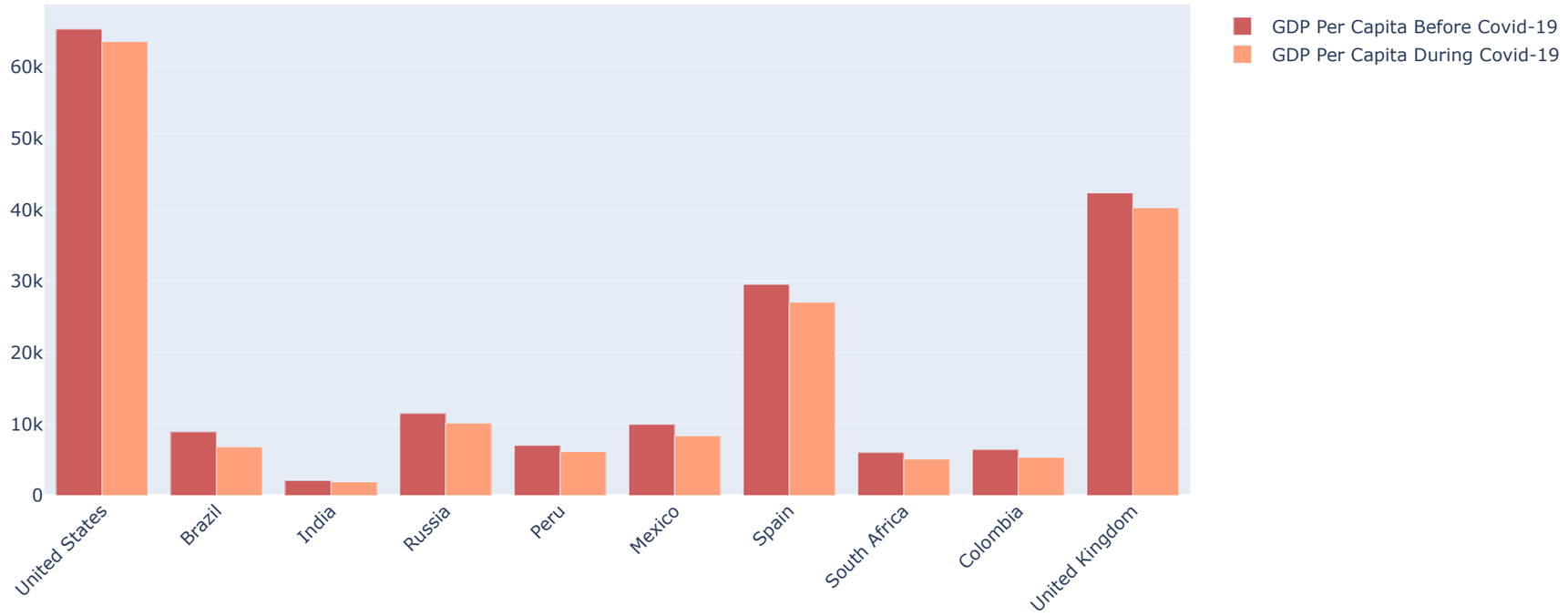
In [18]:
```
fig = px.bar(df, x='Country', y='Total Cases',
             hover_data=['Population', 'Total Deaths'],
             color='GDP During Covid', height=400,
             title="GDP Per Capita During Covid-19")
fig.show()
```



Overlaying Before and during GPD per capital to see their differences. to have a look at the impact of covid-19.

In [19]:
```
fig = go.Figure()
fig.add_trace(go.Bar(
    x=df["Country"],
    y=df["GDP Before Covid"],
    name='GDP Per Capita Before Covid-19',
```

```
    marker_color='indianred'
))
fig.add_trace(go.Bar(
    x=df["Country"],
    y=df["GDP During Covid"],
    name='GDP Per Capita During Covid-19',
    marker_color='lightsalmon'
))
fig.update_layout(barmode='group', xaxis_tickangle=-45)
fig.show()
```

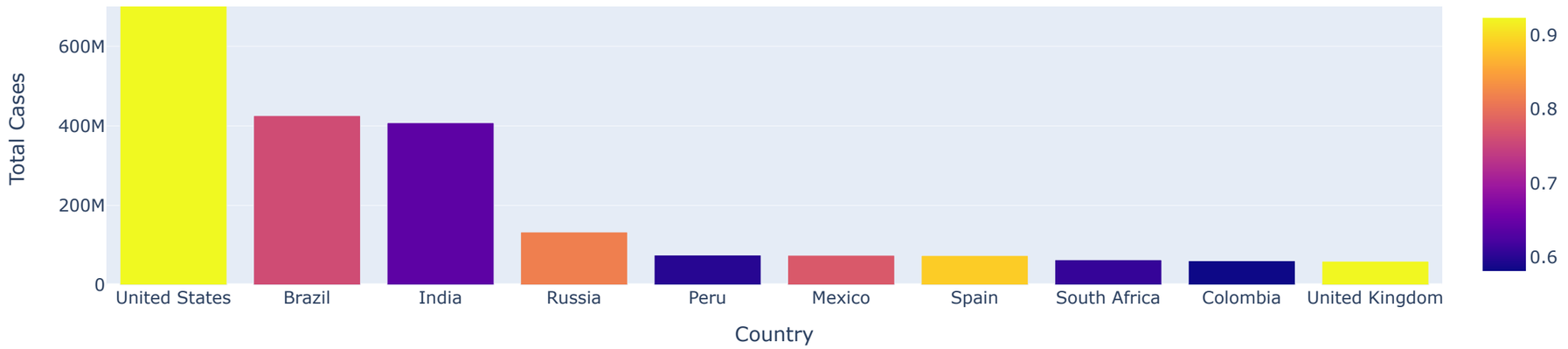


It's definitely obvious that, GDP per capital dropped in all the countries with the highest number of covid-19 cases.

Another economic factor is Human Development Index

It is a statistic composite index of life expectancy, education, and per capita indicators. Let's have a look at how many countries were spending their budget on the human development:

In [20]:
```
fig = px.bar(df, x='Country', y='Total Cases',
             hover_data=['Population', 'Total Deaths'],
             color='HDI', height=400,
             title="Human Development Index during Covid-19")
fig.show()
```



This shows that USA and Uk spent their budgets more on the human development than other countries.

Conclusion

The Covid-19 pandemic has affected almost all countries in the world, with varying degrees of severity. The total number of cases and deaths are correlated, suggesting that countries with higher cases are likely to have higher deaths as well. The stringency index of a country, which measures the strictness of Covid-19 control measures, is negatively correlated with the number of cases and deaths, suggesting that stricter measures have been effective in reducing the spread of the virus.The outbreak of covid-19 resulted in the highest number of covid-19 cases and deaths in the united states. One major reason behind this is the stringency index of the United States. It is comparatively low according to the population. All the countries GDP per capital were affected during the outbreak of covid-19.

Recommendations

- Governments and health authorities should continue to prioritize the implementation of effective Covid-19 control measures, such as increased testing, contact tracing, and vaccination campaigns, in order to reduce the number of cases and deaths.
- Countries with lower levels of development and income should be supported with additional resources and funding to improve their capacity to manage the pandemic and limit its impacts on their populations.
- Further research should be conducted to explore the specific factors that contribute to the varying impacts of Covid-19 on different countries, in order to inform more targeted interventions and policies.