**Name:** PANKAJ BHATNAGAR

**Registration Number:** 202002287

**CHAPTERISATION**

|  |  |  |
| --- | --- | --- |
| S.no. | Topic | Page No. |
| 1. | Declaration | 2 |
| 2 | Analysis Covid-19 | 3 - 11 |
| 2.a) | Introduction and User Requirements | 3 |
| 2.b) | Analysis of Work Done and Design: | 4 |
| 2.c) | Data Preparation | 4 |
| 2.d) | Analyzing the Spread of Covid-19 using Graphical Representation | 7 |
| 2.e) | Analyzing Covid-19 knock on Economy | 11 |
| 3. | Conclusion | 14 |
| 4. | Bibliography / References | 14 |
| 5. | PLAGIARISM Report | 15 |

**DECLARATION BY THE LEARNER**

This is to declare that I have carried out this project work myself in part fulfilment of the **POST GRADUATE DIPLOMA IN DATA SCIENCE (PGDDS)** Program of SCDL. The work is original, has not been copied from anywhere else and has not been submitted to any other University/Institute for an award of any degree/diploma.

**Name:** PANKAJ BHATNAGAR

**Signature:** PANKAJ BHATNAGAR

**Place:** NEW DELHI

**Date:** 26th December 2022

# ANALYSIS COVID-19

# (Using Python)

**Name:** PANKAJ BHATNAGAR

**Registration Number:** 202002287

**Introduction:**

The outbreak of Covid-19 resulted in some of regulations which ended in such a lot of influences on the global economic machine. almost all international locations have been impacted negatively through the rise inside the instances of covid-19. The primary wave of covid-19 impacted the worldwide economic system as the sector turned into in no way equipped for the pandemic. it resulted in a upward thrust in instances, a upward thrust in deaths, a upward thrust in unemployment and a upward push in poverty, ensuing in an monetary slowdown. right here, you're required to investigate the spread of covid-19 instances and all of the impacts of covid-19 at the financial system.

The dataset we are the usage of to investigate the impacts of covid-19 is downloaded from kaggle. It consists of statistics approximately:

1. The country code
2. Name of all the countries
3. Date of the record
4. Human development index of all the countries
5. Daily covid-19 cases
6. Daily deaths due to covid-19
7. Stringency index of the countries
8. The population of the countries
9. GDP per capita of the countries

**Dataset -**

**Direct Links :-**

**Data:** <https://raw.githubusercontent.com/PankajBhat/Analysis-COVID-19/main/Data.csv>

**RawData:** <https://raw.githubusercontent.com/PankajBhat/Analysis-COVID-19/main/Raw_Data.csv>

**User Requirements:**

* Hardware requirements:
  + Operating system- Windows 10, 11
  + Processor- Dual core 2.4 GHz (i5 or i7 series Intel processor or equivalent AMD)
  + RAM - Atleast 4GB
* Software Requirements:
  + Python 3.8
  + PyCharm
  + PIP
  + Kaggle/Jupyter Notebook
  + Chrome

**Analysis of Work Done and Design:**

Begin the undertaking of covid-19 affects analysis by way of importing the important python libraries and the [dataset](https://raw.githubusercontent.com/PankajBhat/Analysis-COVID-19/main/Data.csv):

**Using Python Liberaries:**

**Code:**

import pandas as pd

import plotly.express as px

import plotly.graph\_objects as go

**Dataset:**

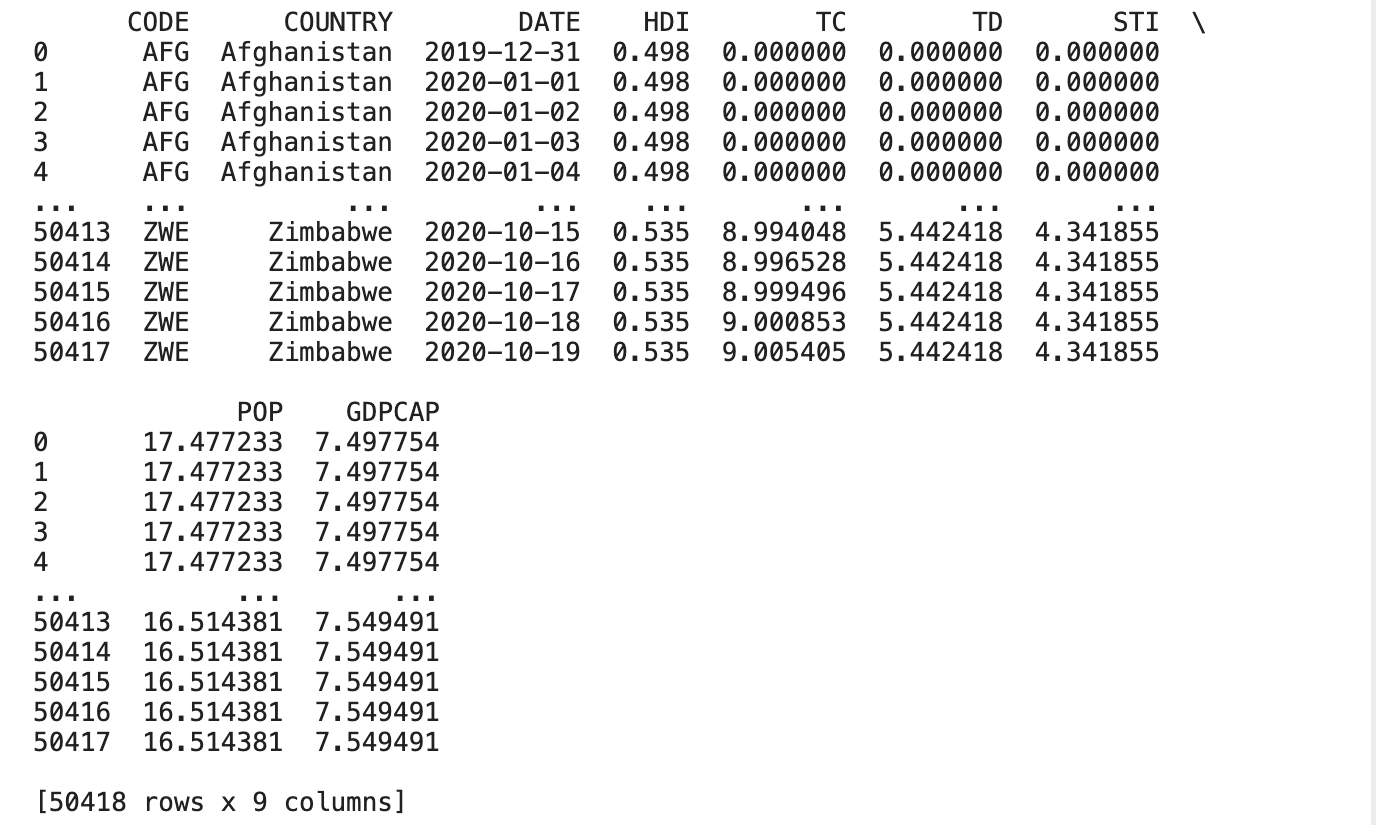
**Code:**

data = pd.read\_csv("/kaggle/input/dataset/Data.csv")

rawdata = pd.read\_csv("/kaggle/input/rawdata/Raw\_Data.csv")

print(data)

#### **Output:**



#### The records we're using incorporates the records on covid-19 cases and their effect on GDP from **December 31, 2019, to October 10, 2020**.

#### **Data Preparation:**

Dataset that i am using for the project and analysis Covid 19 have 2 files:

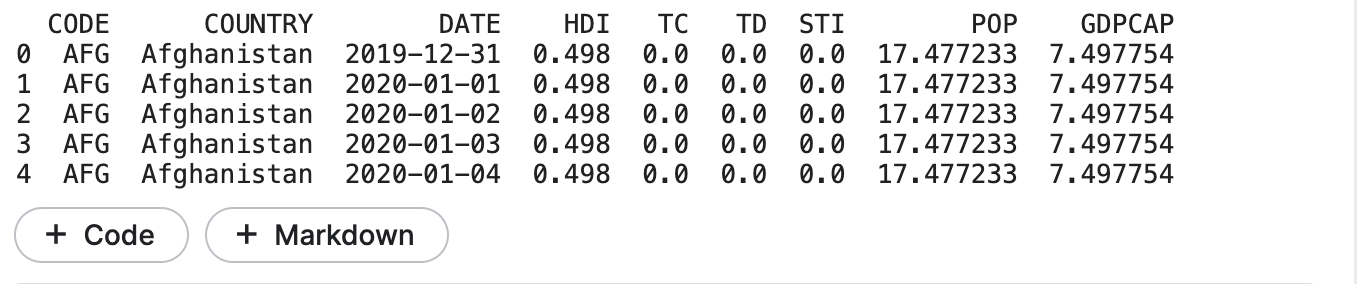
* Raw Data
* Transformed Data

Both of these file are in CSV format and uploaded on Kaggle for better performance.

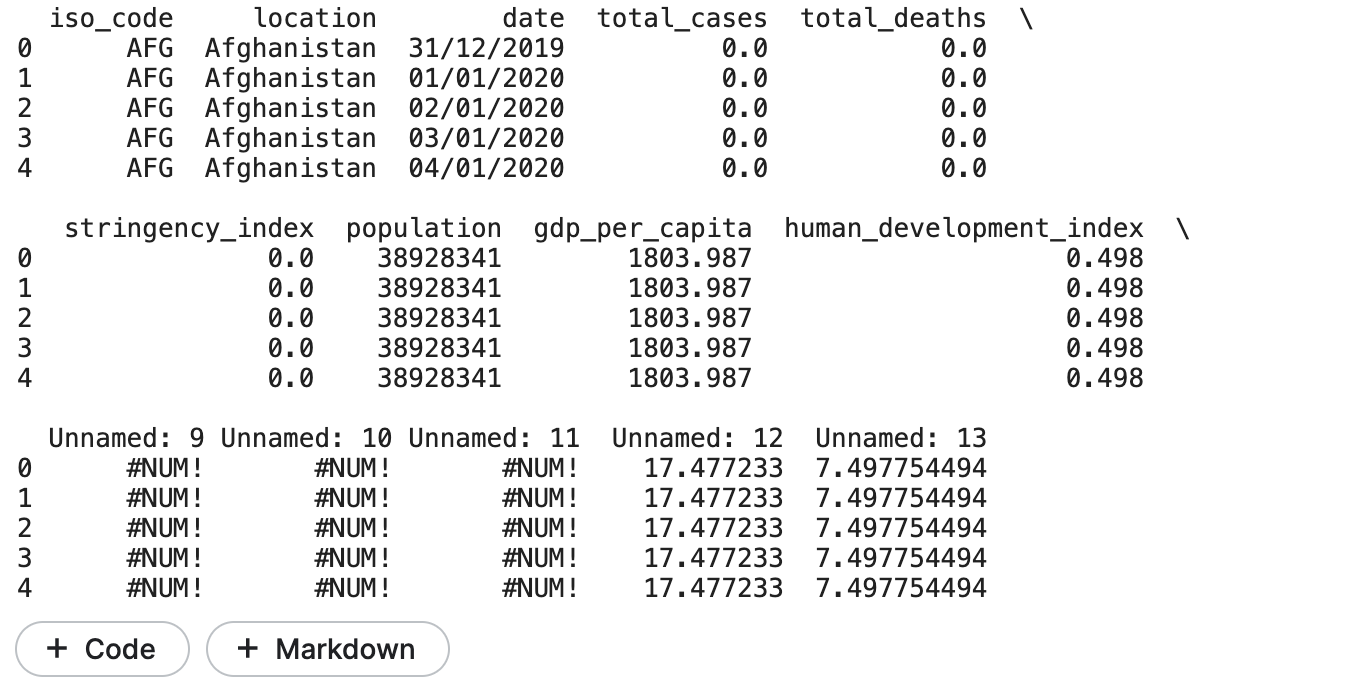
So let’s check both the datasets:

**Code:** print(data.head())

**Output:**

**Code:** print(rawdata.head())

**Output:**

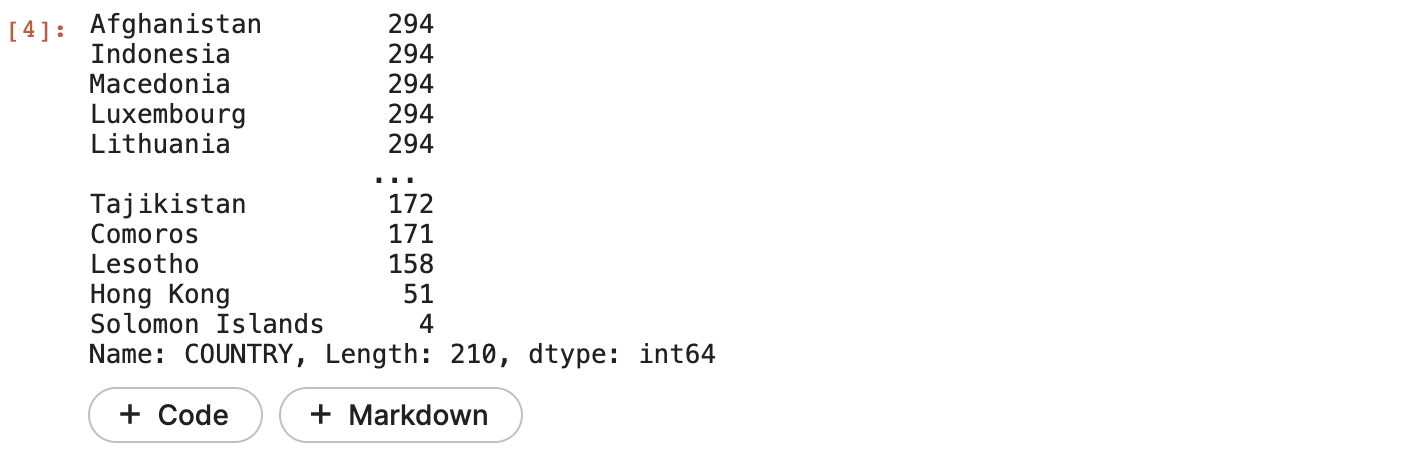


After having preliminary impressions of each datasets,

I found that we must integrate both datasets by creating a brand new dataset. However before we create a new dataset, let's see how many samples are present in the dataset:

**Code:** data["COUNTRY"].value\_counts()

**Output:**

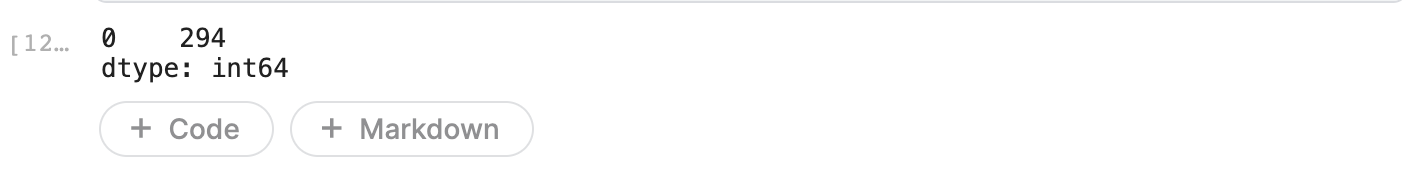


So, we do not have an identical wide variety of samples in the dataset.

Let’s have a look at the mode value:

**Code:** data["COUNTRY"].value\_counts().mode()

**Outoput:**



So, we have 294 mode value. We will need to use it for dividing the sum of all the samples related to the human development index, GDP per capita, and the population. Now let’s create a new dataset by combining the necessary columns from both the datasets:

**Code:**

code = data["CODE"].unique().tolist()

country = data["COUNTRY"].unique().tolist()

hdi = []

tc = []

td = []

sti = []

population = data["POP"].unique().tolist()

gdp = []

for i in country:

hdi.append((data.loc[data["COUNTRY"] == i, "HDI"]).sum()/294)

tc.append((rawdata.loc[rawdata["location"] == i, "total\_cases"]).sum())

td.append((rawdata.loc[rawdata["location"] == i, "total\_deaths"]).sum())

sti.append((data.loc[data["COUNTRY"] == i, "STI"]).sum()/294)

population.append((rawdata.loc[rawdata["location"] == i, "population"]).sum()/294)

aggregated\_data = pd.DataFrame(list(zip(code, country, hdi, tc, td, sti, population)),

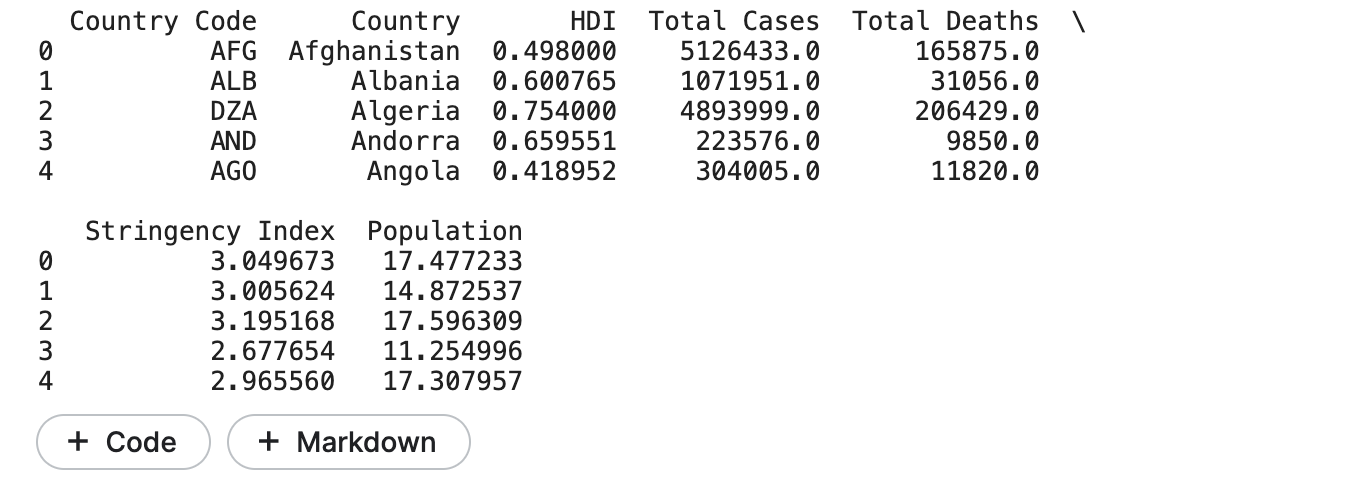
columns = ["Country Code", "Country", "HDI",

"Total Cases", "Total Deaths",

"Stringency Index", "Population"])

print(aggregated\_data.head())

**Output:**



I have now not included the GDP according to capita column but I didn't discover the correct figures for GDP in line with capita inside the dataset. So it is going to be higher to manually collect the data approximately the GDP per capita of the nations.

As we've got so many countries on this statistics, it'll not be smooth to manually acquire the records about the gdp per capita of all the countries. so allow's select a subsample from this dataset.

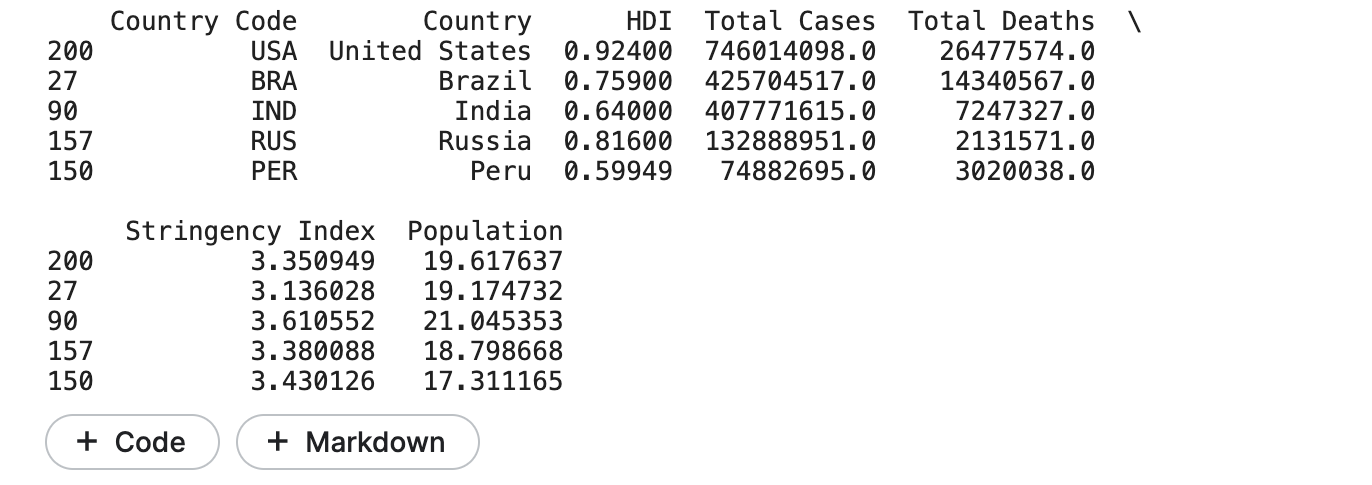
**Code:**

**# Sorting Data According to Total Cases**

**data = aggregated\_data.sort\_values(by=["Total Cases"], ascending=False)**

**print(data.head())**

**Output:**

To create a subsample from this dataset, I may be selecting the pinnacle 10 international locations with the best wide variety of covid-19 cases.Now here’s how we can select the top 10 countries with the highest number of cases:

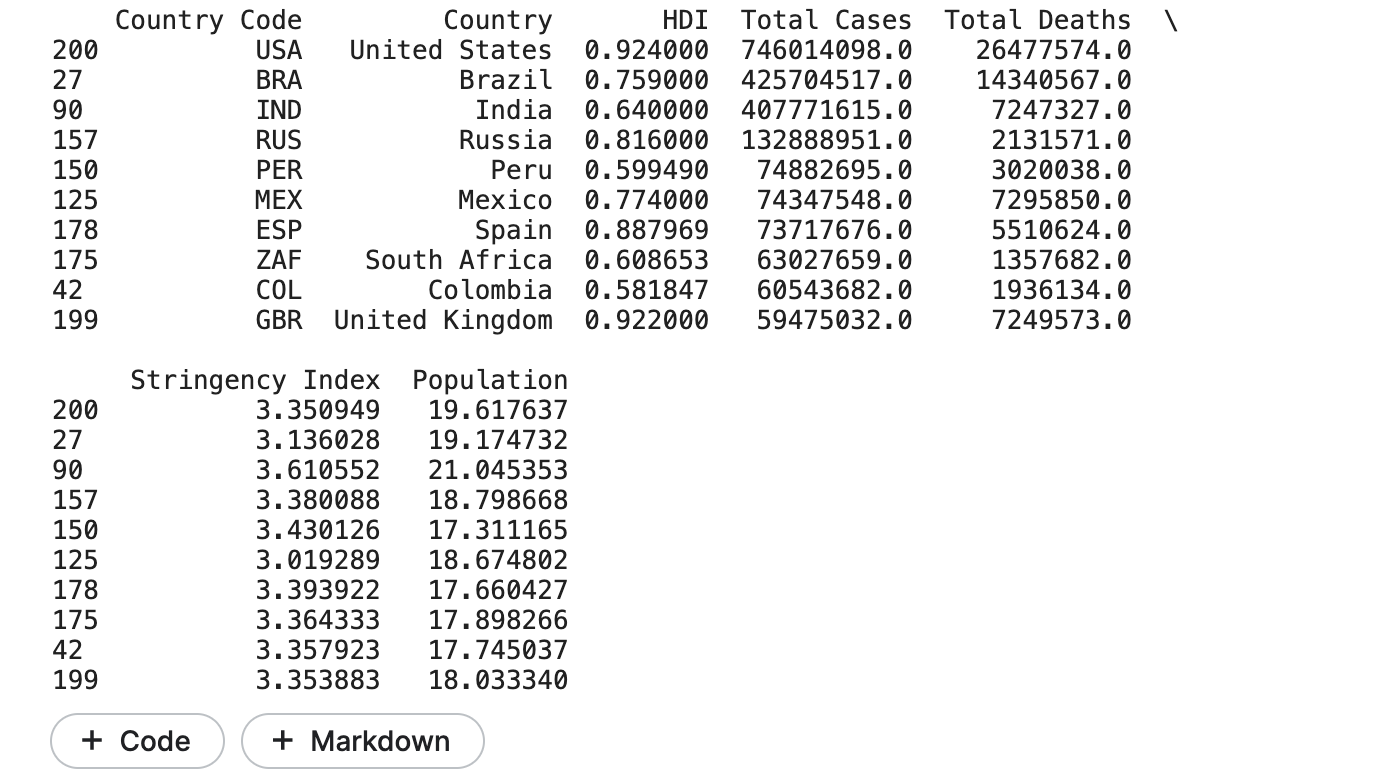
**Code:**

**# Top 10 Countries with Highest Covid Cases**

**data = data.head(10)**

**print(data)**

**Output:**



Now I will add two more columns****(GDP per capita before Covid-19, GDP per capita during Covid-19)**** to this dataset:

**Code:**

data["GDP Before Covid"] = [65279.53, 8897.49, 2100.75,

11497.65, 7027.61, 9946.03,

29564.74, 6001.40, 6424.98, 42354.41]

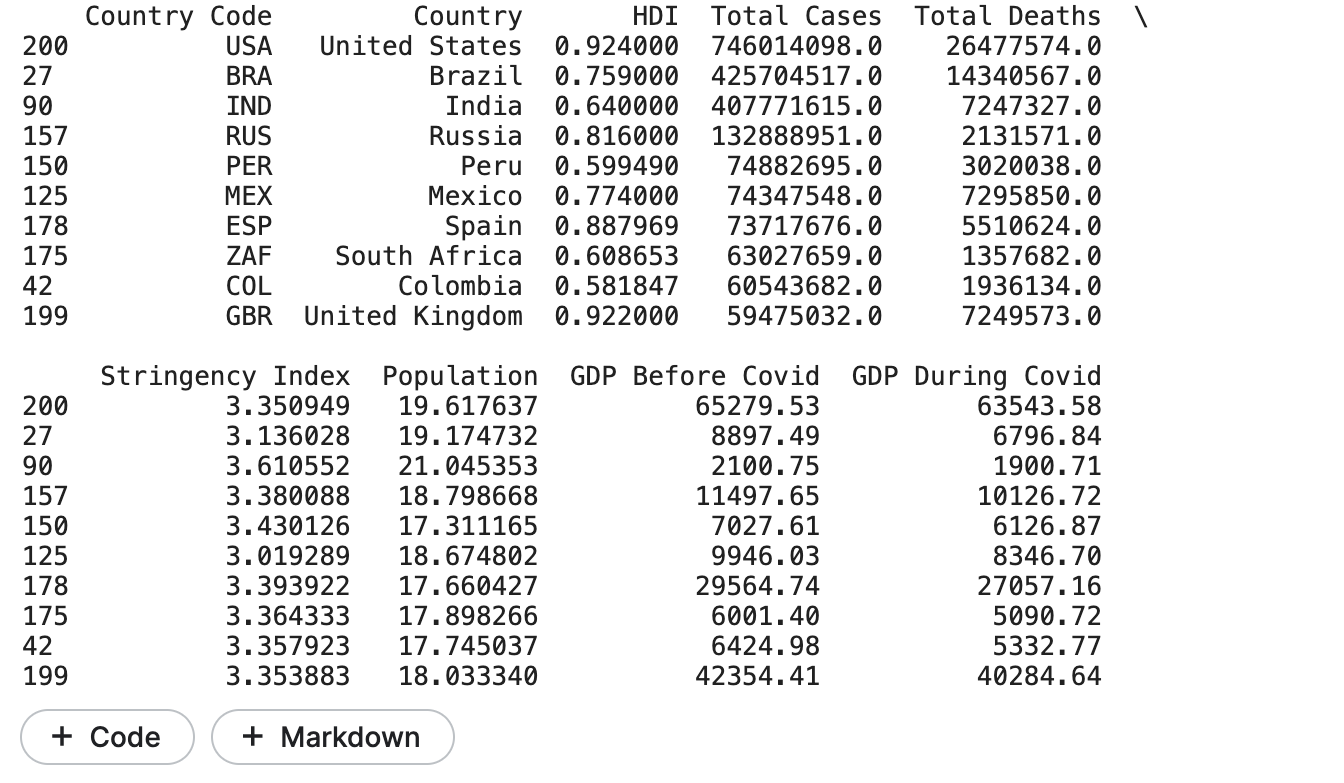
data["GDP During Covid"] = [63543.58, 6796.84, 1900.71,

10126.72, 6126.87, 8346.70,

27057.16, 5090.72, 5332.77, 40284.64]

print(data)

**Output:**



****Note:** The data about the GDP per capita manually collected**.****

#### **Analyzing the Spread of Covid-19 using Graphical Representation-**

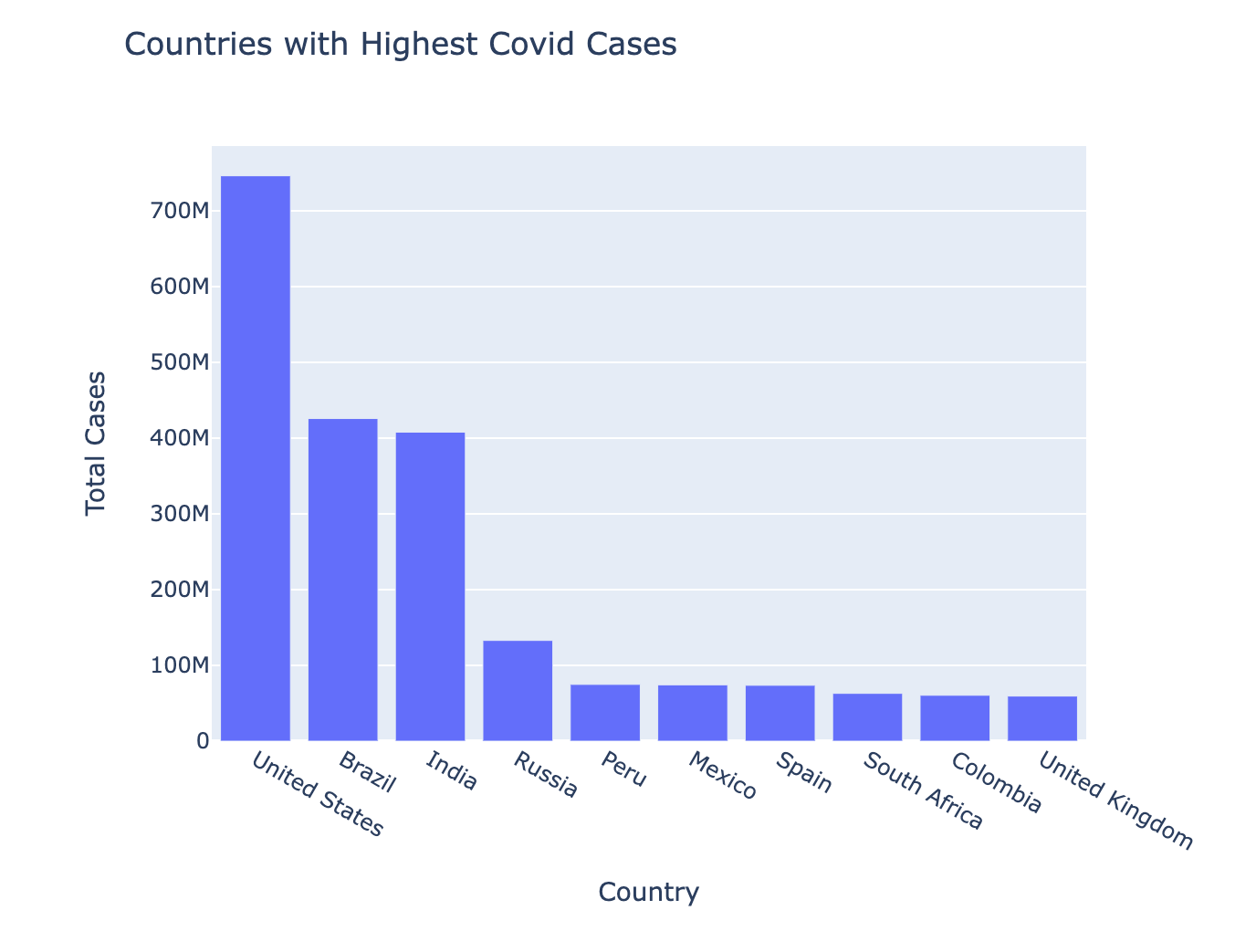
Now start by reading the unfold of covid-19 in all the international locations with the very best number of covid-19 cases. i will first have a take a look at all the international locations with the very best range of covid-19 instances:

**Code: figure = px.bar(data, y='Total Cases', x='Country',**

**title="Countries with Highest Covid Cases")**

**figure.show()**

**Output:**



We are able to see that the us is comparatively having a completely excessive quantity of covid-19 instances as compared to brazil and india in the 2d and 1/3 positions. now let's examine the full wide variety of deaths some of the international locations with the highest quantity of covid-19 instances:

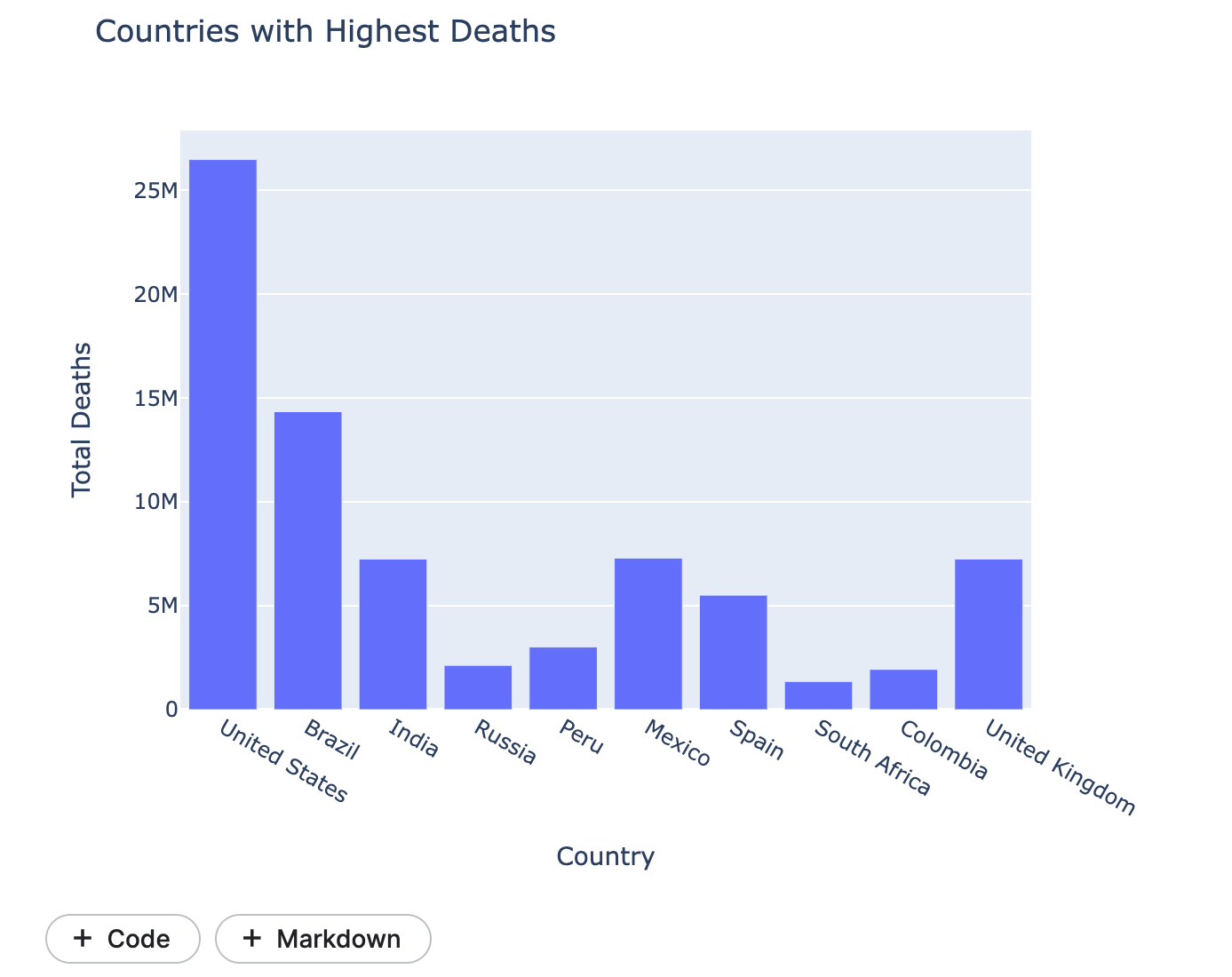
**Code:**

figure = px.bar(data, y='Total Deaths', x='Country',

title="Countries with Highest Deaths")

figure.show()

**Output:**



just like the overall wide variety of covid-19 cases, the usa is leading within the deaths, with brazil and india in the 2nd and 1/3 positions. one component to note right here is that the dying rate in india, russia, and south africa is comparatively low according to the total quantity of instances. Now permit's compare the full variety of instances and overall deaths in these types of international locations:

**Code:**

fig = go.Figure()

fig.add\_trace(go.Bar(

x=data["Country"],

y=data["Total Cases"],

name='Total Cases',

marker\_color='indianred'

))

fig.add\_trace(go.Bar(

x=data["Country"],

y=data["Total Deaths"],

name='Total Deaths',

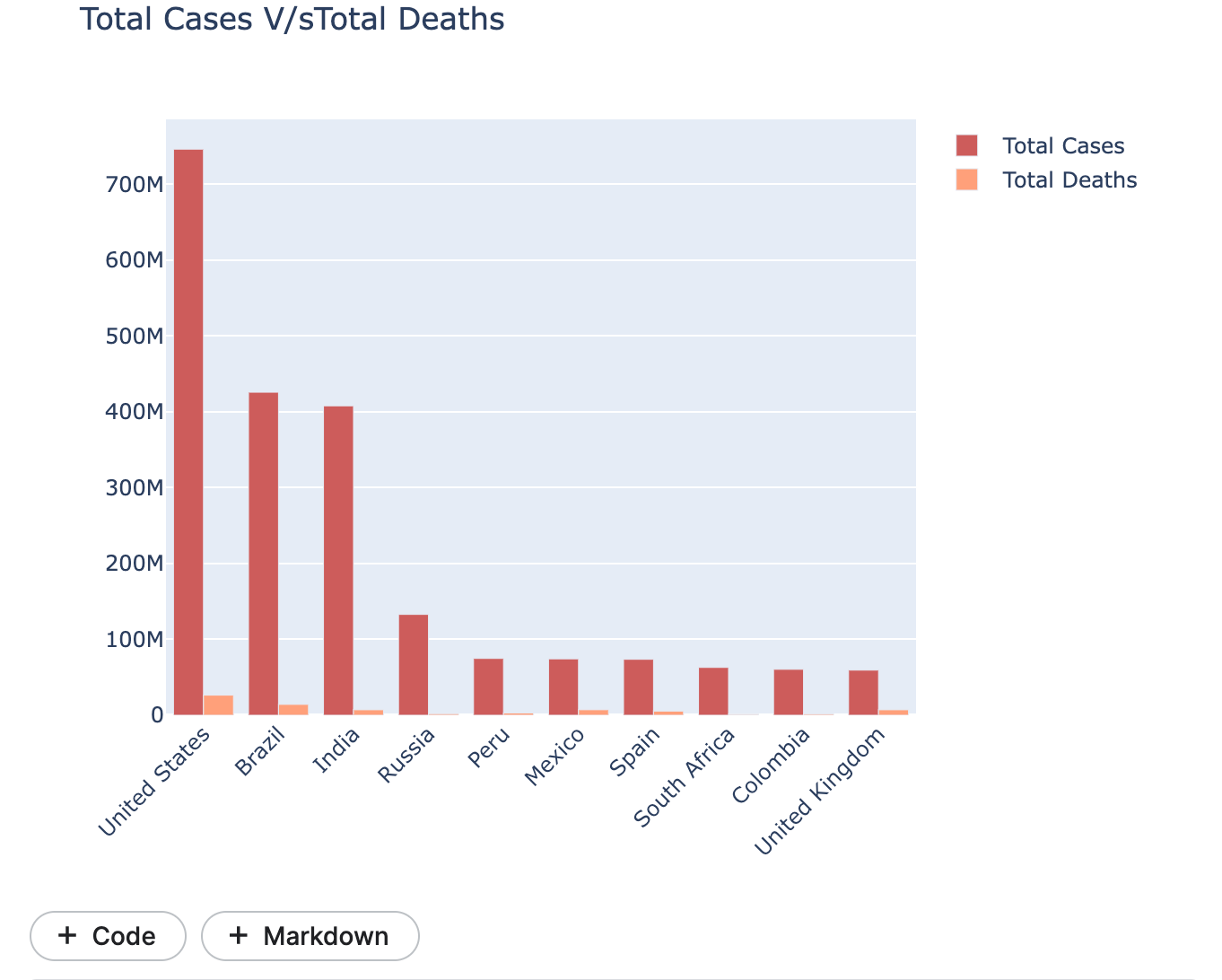
marker\_color='lightsalmon'

))

fig.update\_layout(barmode='group', xaxis\_tickangle=-45,title="Total Cases V/sTotal Deaths")

fig.show()

**Ouput:**



Now let’s have a look at the percentage of total deaths and total cases among all the countries with the highest number of covid-19 cases:

**Code:**

# Percentage of Total Cases and Deaths

cases = data["Total Cases"].sum()

deceased = data["Total Deaths"].sum()

labels = ["Total Cases", "Total Deaths"]

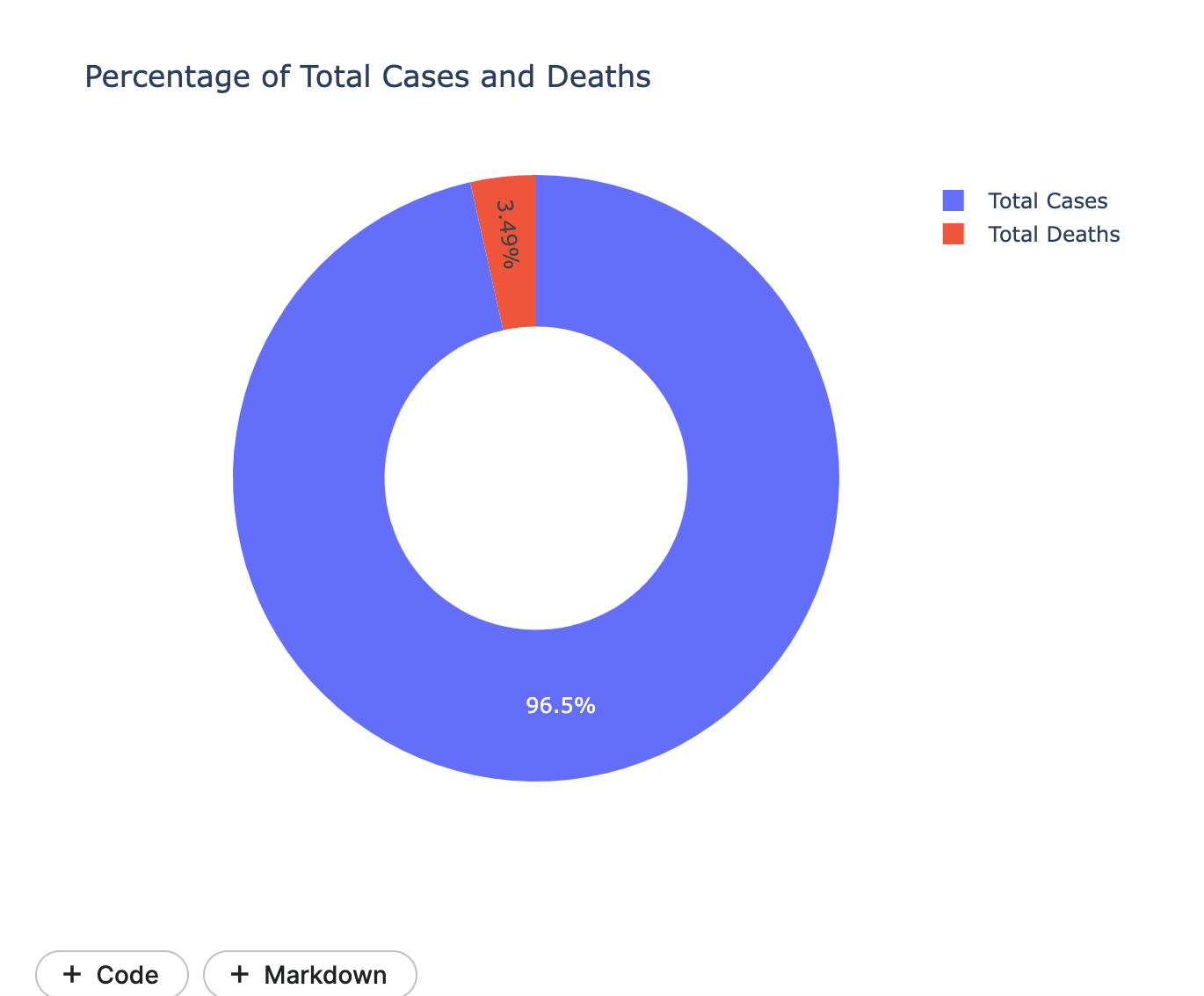
values = [cases, deceased]

fig = px.pie(data, values=values, names=labels,

title='Percentage of Total Cases and Deaths', hole=0.5)

fig.show()

**Output:**



Below is how you can calculate the death rate of Covid-19 cases:

**Code:**

death\_rate = (data["Total Deaths"].sum() / data["Total Cases"].sum()) \* 100

print("Death Rate = ", death\_rate)

**Output:**



Any other essential column in this dataset is the stringency index. it is a composite measure of reaction indicators, consisting of faculty closures, administrative center closures, and travel bans. it indicates how strictly international locations are following these measures to control the unfold of covid-19:

**Code:**

fig = px.bar(data, x='Country', y='Total Cases',

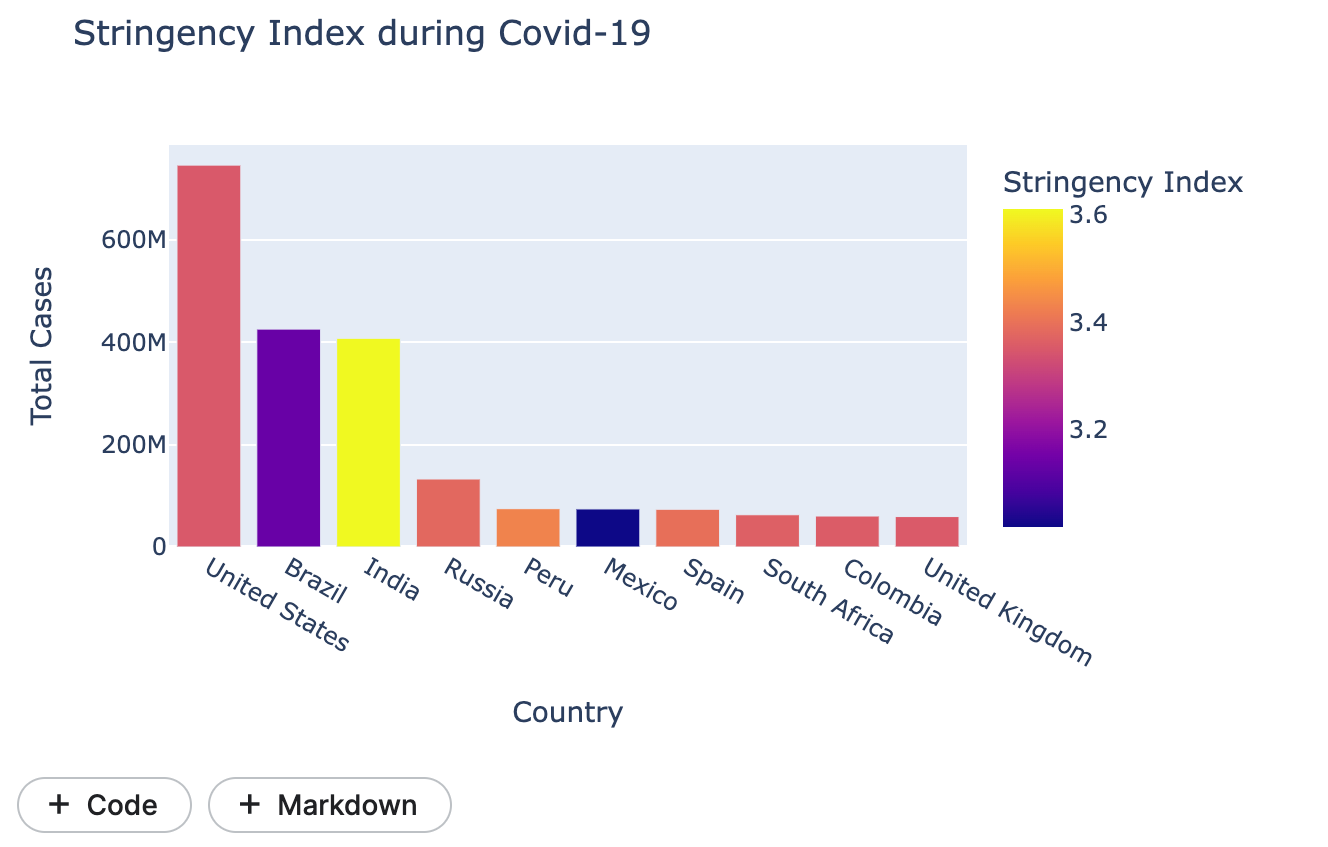
hover\_data=['Population', 'Total Deaths'],

color='Stringency Index', height=400,

title= "Stringency Index during Covid-19")

fig.show()

**Ouput:**



it is a composite measure of reaction indicators, consisting of faculty closures, administrative center closures, and travel bans:.

**Analyzing Covid-19 knock on Economy:**

Now permit's flow to research the affects of covid-19 on the economy. here the gdp in keeping with capita is the number one component for reading the monetary slowdowns brought on because of the outbreak of covid-19.

**Code:**

fig = px.bar(data, x='Country', y='Total Cases',

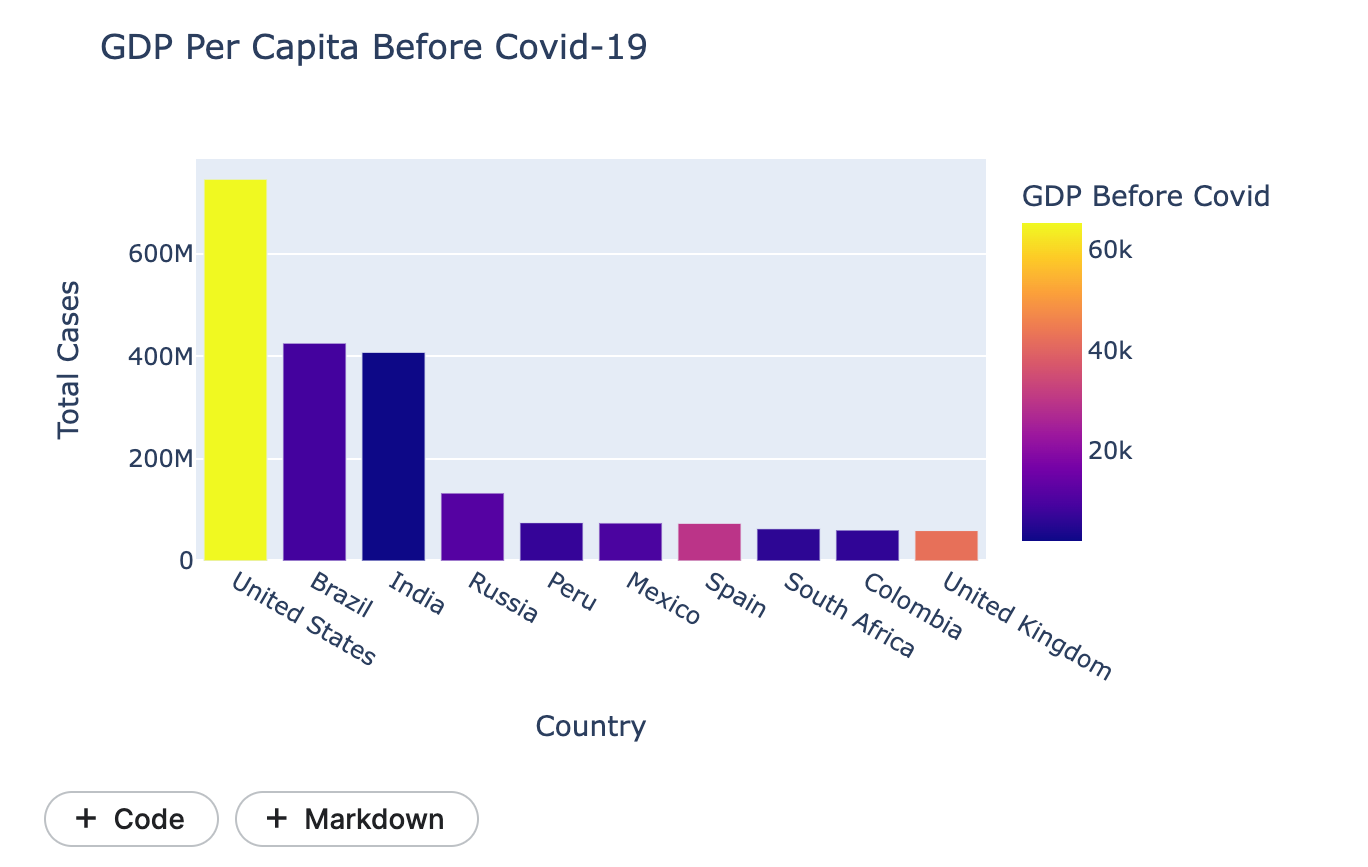
hover\_data=['Population', 'Total Deaths'],

color='GDP Before Covid', height=400,

title="GDP Per Capita Before Covid-19")

fig.show()

**Output:**

  
Now let's have a look at the gdp in line with capita at some point of the rise in the instances of covid-19:

**Code:**

fig = px.bar(data, x='Country', y='Total Cases',

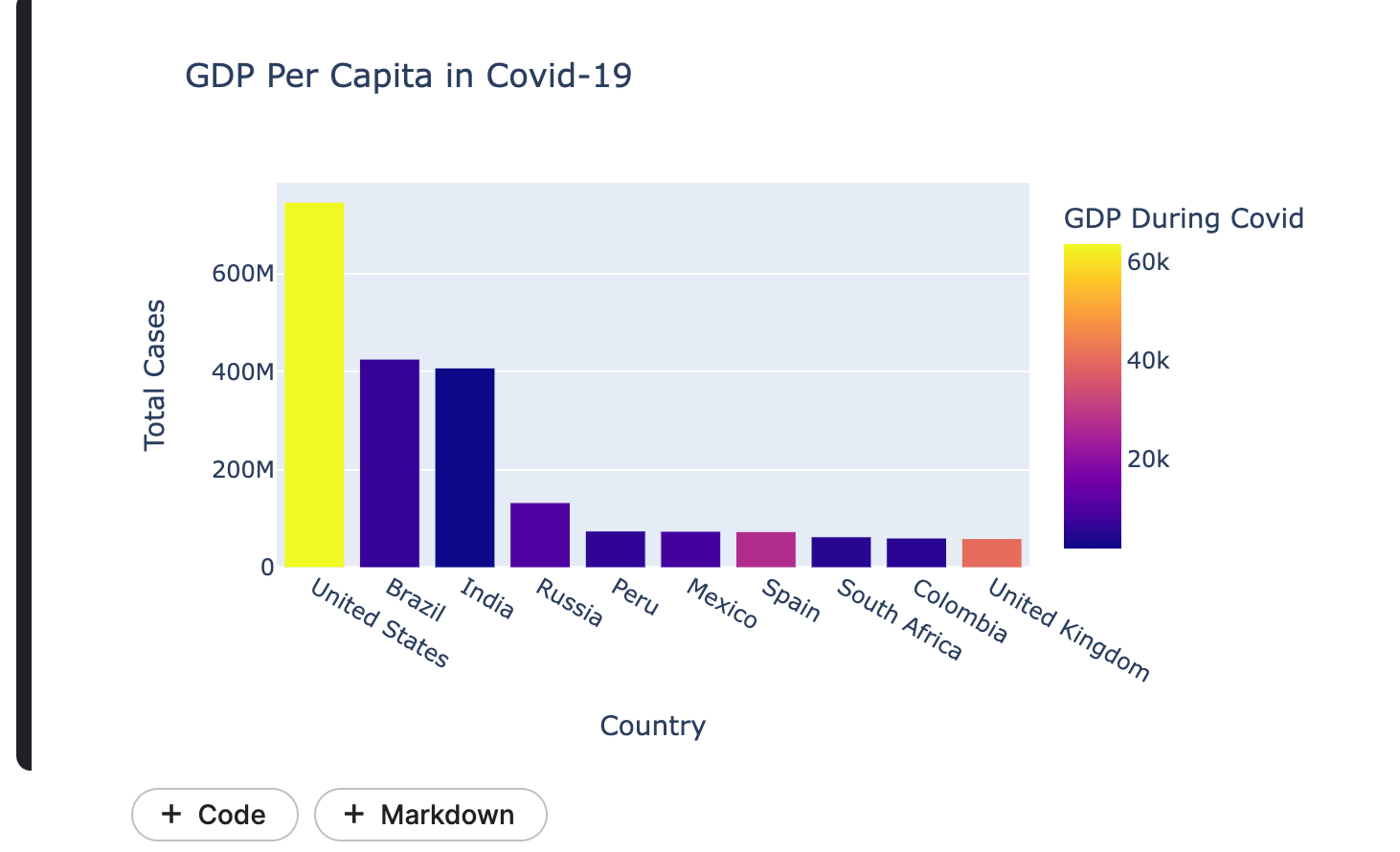
hover\_data=['Population', 'Total Deaths'],

color='GDP During Covid', height=400,

title="GDP Per Capita in Covid-19")

fig.show()

**Output:**



Now allow's evaluate the gdp in line with capita before covid-19 and at some point of covid-19 to have a study the impact of covid-19 on the gdp in keeping with capita:

**Code:**

fig = go.Figure()

fig.add\_trace(go.Bar(

x=data["Country"],

y=data["GDP Before Covid"],

name='GDP Per Capita Before Covid-19',

marker\_color='indianred'

))

fig.add\_trace(go.Bar(

x=data["Country"],

y=data["GDP During Covid"],

name='GDP Per Capita in Covid-19',

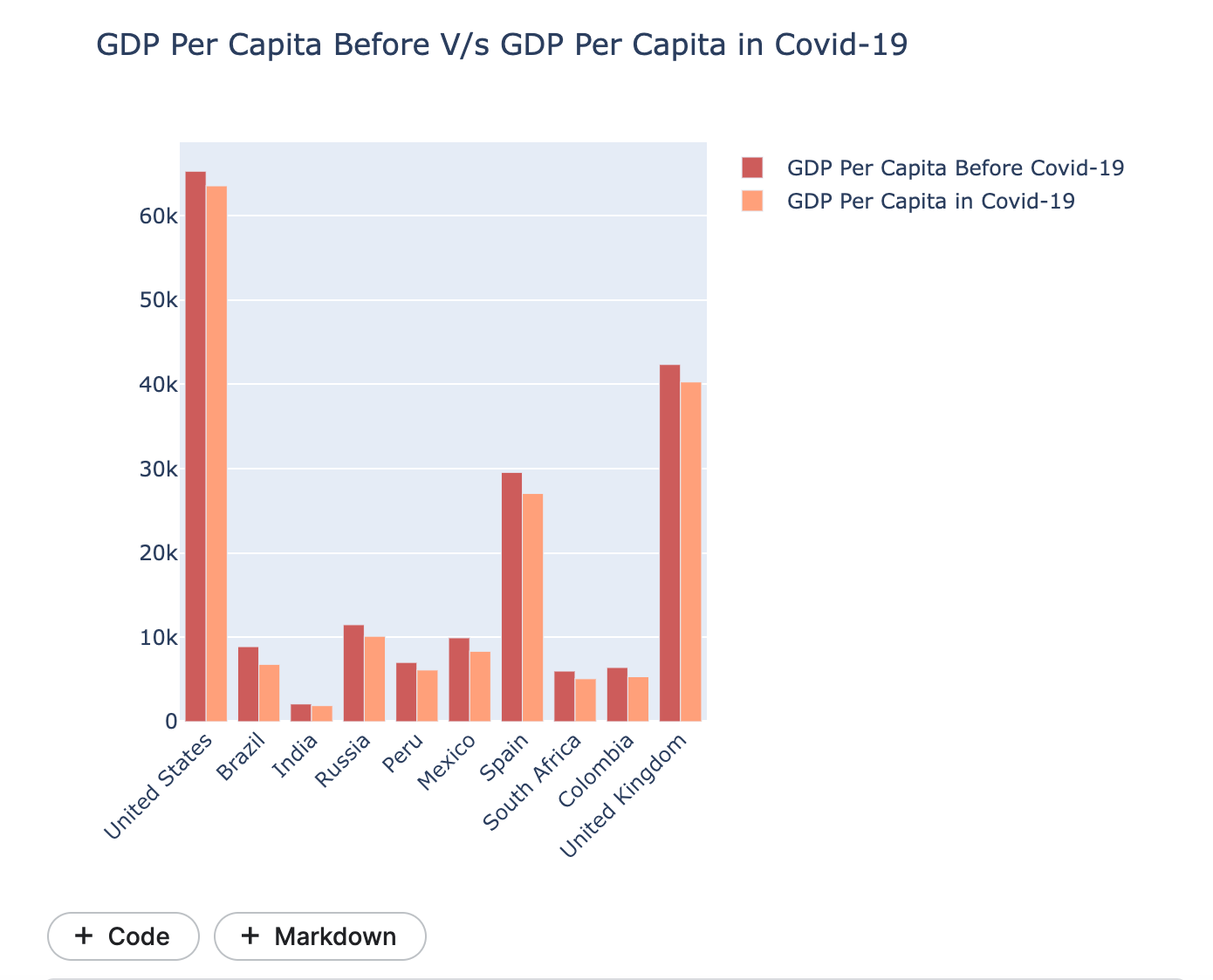
marker\_color='lightsalmon'

))

fig.update\_layout(barmode='group', xaxis\_tickangle=-45,title ="GDP Per Capita Before V/s GDP Per Capita in Covid-19")

fig.show()

**Output:**



You can see a drop in GDP in line with capita in all the nations with the best wide variety of covid-19 cases.

A different essential economic issue is human development index. it is a statistic composite index of life expectancy, training, and consistent with capita signs. let's have a look at how many nations have been spending their finances at the human development:

**Code:**

fig = px.bar(data, x='Country', y='Total Cases',

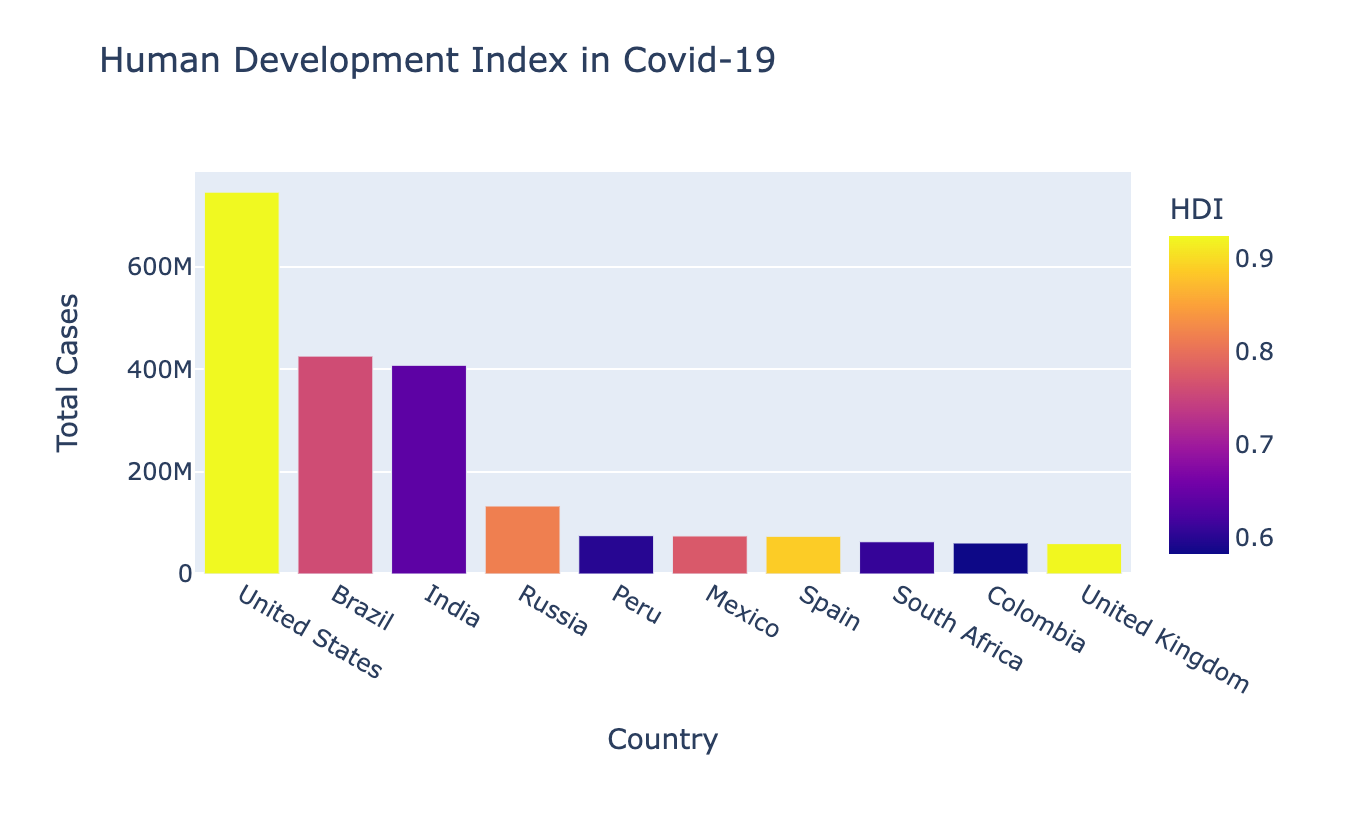
hover\_data=['Population', 'Total Deaths'],

color='HDI', height=400,

title="Human Development Index during Covid-19")

fig.show()

**Output:**



So this is how we can examine the unfold of covid-19 and its impact at the economy.

**Conclusion:**

In this venture, we studied the spread of covid-19 the various nations and its effect on the global economic system. we noticed that the outbreak of covid-19 resulted in the maximum quantity of covid-19 cases and deaths inside the united states. one principal reason at the back of this is the stringency index of the USA. it's far relatively low according to the population. we additionally analyzed how the GDP is in line with the per capita of each that was affected during the outbreak of covid-19.

**Bibliography / References:**

* COVID 19 Impact on Economy: [2020 Year in Review: The impact of COVID-19 in 12 charts](https://blogs.worldbank.org/voices/2020-year-review-impact-covid-19-12-charts" \l ":~:text=The%20June%20edition%20of%20the,as%20well%20as%20per%20capita)
* Kaggle - <https://www.kaggle.com/pankajbhatnagar/analysis-covid-19-using-python/edit>
* Google - <https://www.google.com/>
* Matplotlib - <https://matplotlib.org/>
* Seaborn - <https://seaborn.pydata.org/>
* Plotly- <https://plotly.com/>
* Python - <https://www.python.org/>
* Github - <https://github.com/PankajBhat/Analysis-COVID-19>

**PLAGIARISM Check**: <https://www.duplichecker.com/>

**Report Attached:**

1. **<https://drive.google.com/drive/u/0/my-drive>**
2. **<https://drive.google.com/drive/u/0/my-drive>**

