

```
In [1]: #importing libraries
import pandas as pd
import numpy as np
import plotly.express as px
```

```
In [2]: #importing dataset
df= pd.read_csv("country_wise_latest.csv")
```

```
In [3]: #checking the dimension
df.shape
```

Out[3]: (187, 15)

```
In [4]: #displaying the data
df.head()
```

Out[4]:

	Country/Region	Confirmed	Deaths	Recovered	Active	New cases	New deaths	New recovered
0	Afghanistan	36263	1269	25198	9796	106	10	18
1	Albania	4880	144	2745	1991	117	6	63
2	Algeria	27973	1163	18837	7973	616	8	749
3	Andorra	907	52	803	52	10	0	0
4	Angola	950	41	242	667	18	1	0

```
In [5]: #checking the null values
df.isna().sum()
```

```
Out[5]: Country/Region      0
Confirmed      0
Deaths         0
Recovered      0
Active         0
New cases      0
New deaths     0
New recovered   0
Deaths / 100 Cases      0
Recovered / 100 Cases   0
Deaths / 100 Recovered  0
Confirmed last week     0
1 week change           0
1 week % increase       0
WHO Region             0
dtype: int64
```

EDA (Exploratory Data Analysis) and Visualization

1. Which countries had the highest number of confirmed cases?

```
In [6]: high_confirm_cases= df.nlargest (10,'Confirmed')
```

```
In [7]: high_confirm_cases.head()
```

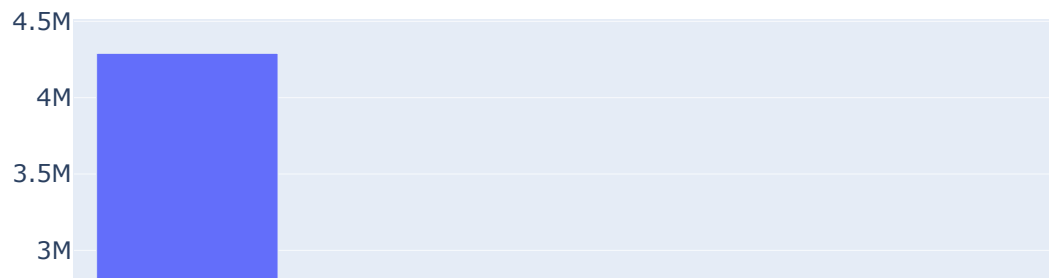
Out[7]:

	Country/Region	Confirmed	Deaths	Recovered	Active	New cases	New deaths	New recoveries
173	US	4290259	148011	1325804	2816444	56336	1076	2794
23	Brazil	2442375	87618	1846641	508116	23284	614	3372
79	India	1480073	33408	951166	495499	44457	637	3359
138	Russia	816680	13334	602249	201097	5607	85	307
154	South Africa	452529	7067	274925	170537	7096	298	984



```
In [8]: px.bar(high_confirm_cases ,
               x= 'Country/Region',
               y='Confirmed',
               title='Top 10 countries with the highest number of confirmed cases')
```

Top 10 countries with the highest number of confirmed cas



2. What was the distribution of confirmed cases across different WHO regions?

```
In [9]: df['WHO Region'].unique()
```

```
Out[9]: array(['Eastern Mediterranean', 'Europe', 'Africa', 'Americas',  
              'Western Pacific', 'South-East Asia'], dtype=object)
```

```
In [10]: px.pie(df,  
               values= 'Confirmed',  
               names= 'WHO Region',  
               title= 'Distribution of confirmed cases across different WHO regions')
```

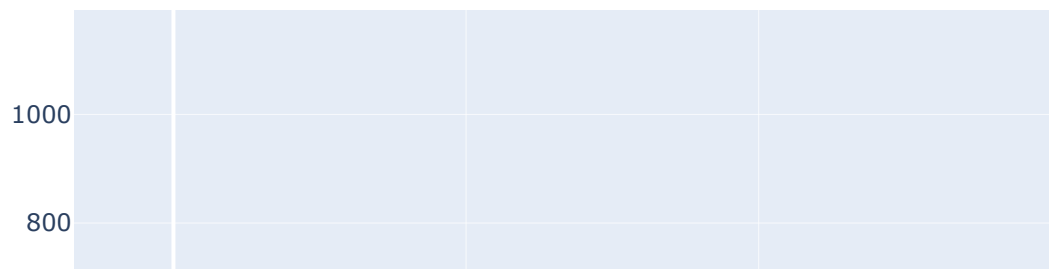
Distribution of confirmed cases across different WHO region



3. What were the trends in new cases, new deaths, and new recoveries in each country?

```
In [11]: px.scatter(df,
    x= 'New cases',
    y= 'New deaths',
    size= 'New recovered',
    color= 'Country/Region',
    title= 'Trends in new cases, new deaths, and new recoveries in each c
```

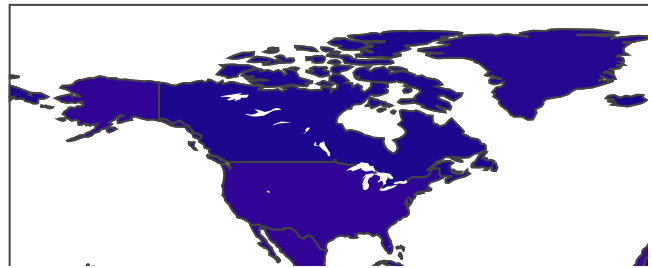
Trends in new cases, new deaths, and new recoveries in ea



4. What was the 1 -week percentage increase in confirmed cases for each country?

```
In [12]: px.choropleth(df,
    locations= 'Country/Region',
    locationmode= 'country names',
    color= '1 week % increase',
    hover_name= 'Country/Region',
    title= '1-week percentage increase in confirmed cases for each cou
    color_continuous_scale= px.colors.sequential.Plasma)
```

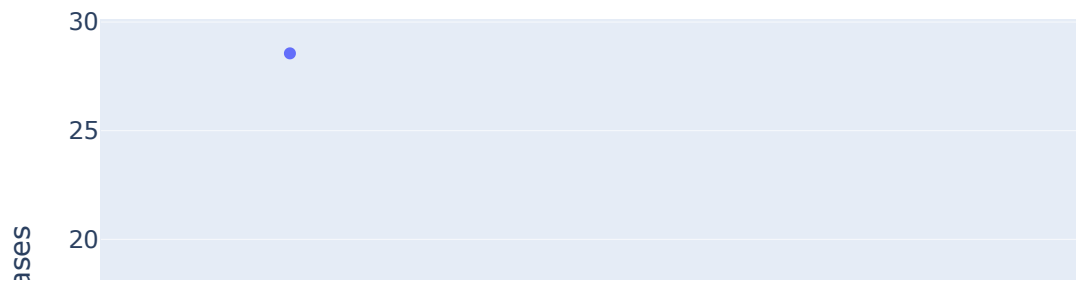
1-week percentage increase in confirmed cases for each co



5. How do death rates per 100 confirmed cases vary by WHO region?

```
In [13]: px.box(df,
             x= 'WHO Region',
             y= 'Deaths / 100 Cases',
             title= 'Death rates per 100 confirmed cases vary by WHO region')
```

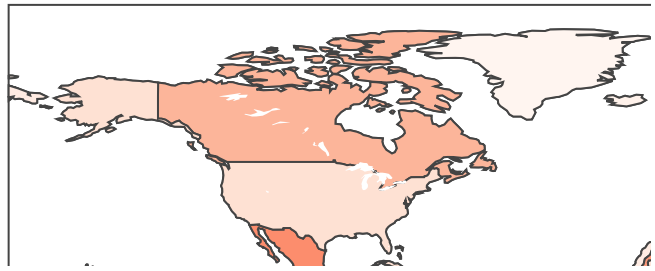
Death rates per 100 confirmed cases vary by WHO region



6. What was the mortality rate (deaths per 100 confirmed cases) for each country?

```
In [14]: px.choropleth(df,
    locations= 'Country/Region',
    locationmode= 'country names',
    color= 'Deaths / 100 Cases',
    hover_name= 'Country/Region',
    title= 'Mortality rate (deaths per 100 confirmed cases) by country',
    color_continuous_scale= px.colors.sequential.Reds)
```

Mortality rate (deaths per 100 confirmed cases) by country



7. How many countries are in each WHO region?

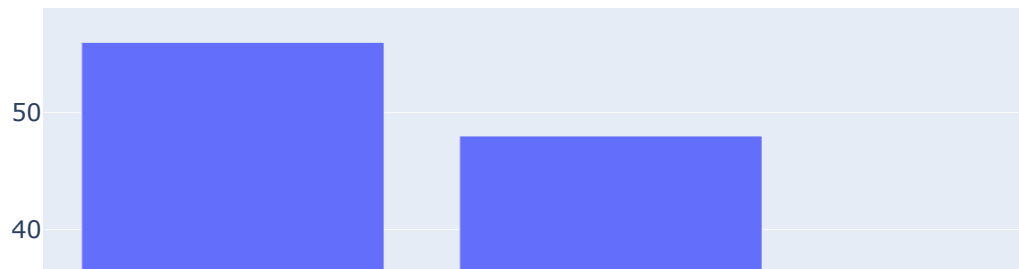
```
In [15]: who_region_count= df['WHO Region'].value_counts().reset_index()  
who_region_count
```

```
Out[15]:
```

	WHO Region	count
0	Europe	56
1	Africa	48
2	Americas	35
3	Eastern Mediterranean	22
4	Western Pacific	16
5	South-East Asia	10

```
In [16]: px.bar(who_region_count,  
                x='WHO Region',  
                y='count',  
                title='No. of countries in each WHO Region')
```

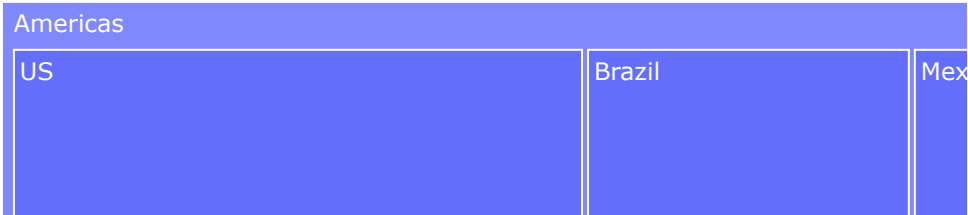

No. of countries in each WHO Region



8. Treemap of COVID-19 confirmed cases by WHO region and country

```
In [17]: px.treemap(df,  
                path= ['WHO Region', 'Country/Region'],  
                values= 'Confirmed',  
                title= 'Treemap of COVID-19 confirmed cases by WHO region and country')
```

Treemap of COVID-19 confirmed cases by WHO region and



In []: