Covid-19 Data Analysis bold text

```
#importing librarires
import pandas as pd
import plotly.express as px
```

```
#importing dataset
df = pd.read_csv('country_wise_latest.csv')
```

```
#checking the dimension
df.shape
```

```
(187, 15)
```

```
#displaying the data
df.head()
```

	Country/Region	Confirmed	Deaths	Recovered	Active	New cases	New deaths	New recovered	Deaths / 100 Cases	Recovered / 100 Cases	Deaths / 100 Recovered	Confirmed last week	1 week change	1 week % increase	WHO Region
0	Afghanistan	36263	1269	25198	9796	106	10	18	3.50	69.49	5.04	35526	737	2.07	Eastern Mediterranean
1	Albania	4880	144	2745	1991	117	6	63	2.95	56.25	5.25	4171	709	17.00	Europe
2	Algeria	27973	1163	18837	7973	616	8	749	4.16	67.34	6.17	23691	4282	18.07	Africa
3	Andorra	907	52	803	52	10	0	0	5.73	88.53	6.48	884	23	2.60	Europe
4	Angola	950	41	242	667	18	1	0	4.32	25.47	16.94	749	201	26.84	Africa

```
#checking the null values
df.isnull().sum()
```

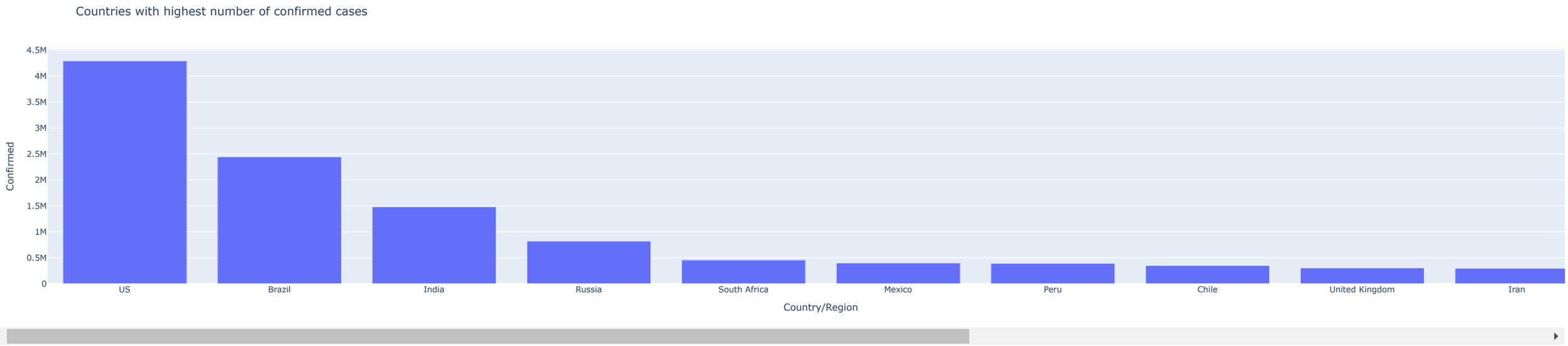
	0
Country/Region	0
Confirmed	0
Deaths	0
Recovered	0
Active	0
New cases	0
New deaths	0
New recovered	0
Deaths / 100 Cases	0
Recovered / 100 Cases	0
Deaths / 100 Recovered	0
Confirmed last week	0
1 week change	0
1 week % increase	0
WHO Region	0

EDA (Exploratory Data Analysis) and Visualization

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

```
hight_confirmed_cases = df.nlargest(10,'Confirmed')
```

```
px.bar(hight_confirmed_cases,x='Country/Region',y='Confirmed',title='Countries with highest number of confirmed cases')
```

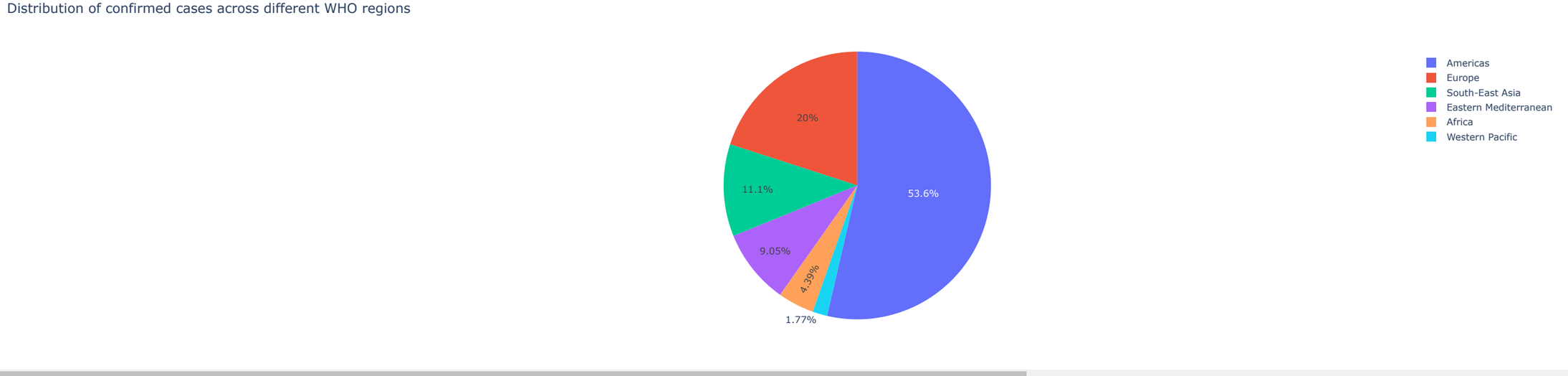


2. What was the distribution of confirmed cases across different WHO regions?

```
df['WHO Region'].unique()
```

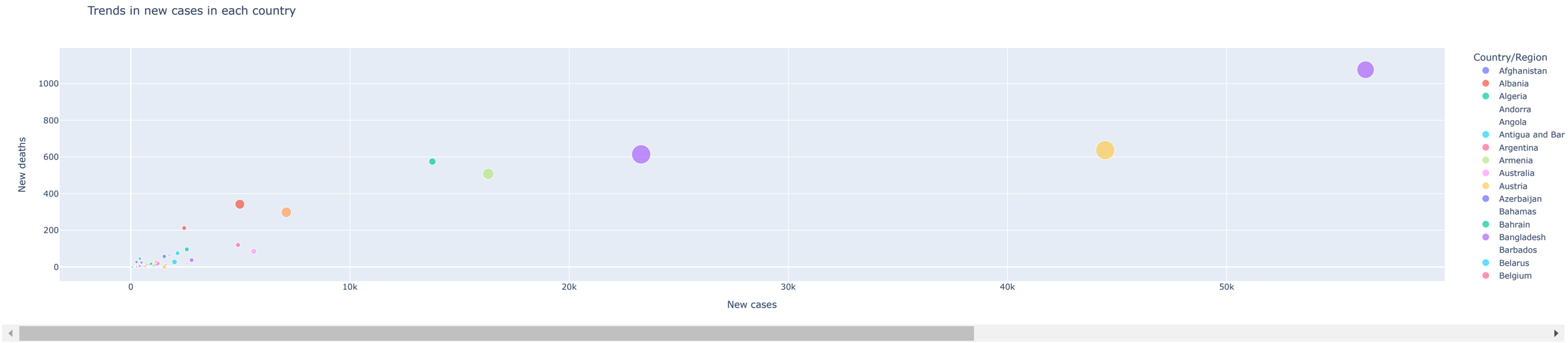
```
array(['Eastern Mediterranean', 'Europe', 'Africa', 'Americas',
       'Western Pacific', 'South-East Asia'], dtype=object)
```

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



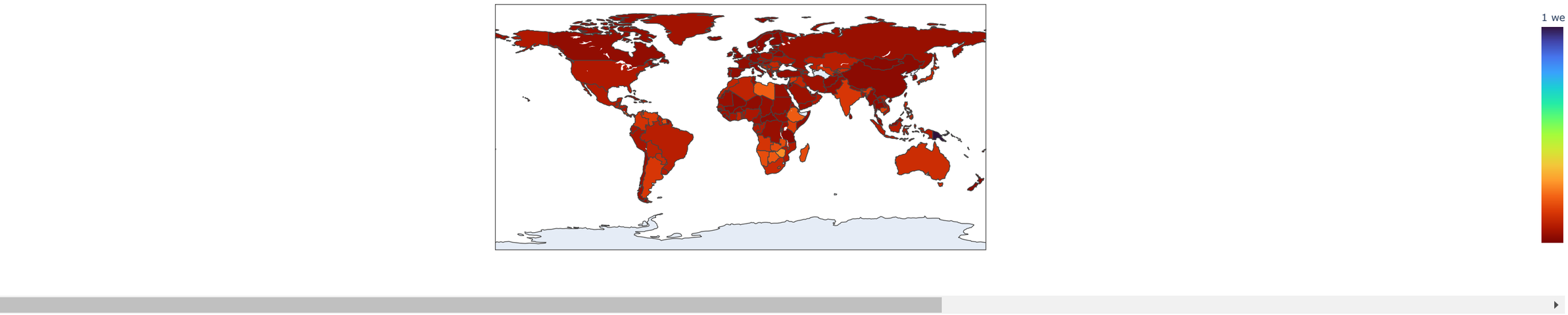
3. What were the trends in new cases, new deaths, and new recoveries in each country?

```
px.scatter(df,x='New cases',y='New deaths',size = 'New recovered', color = 'Country/Region', title='Trends in new cases in each country')
```



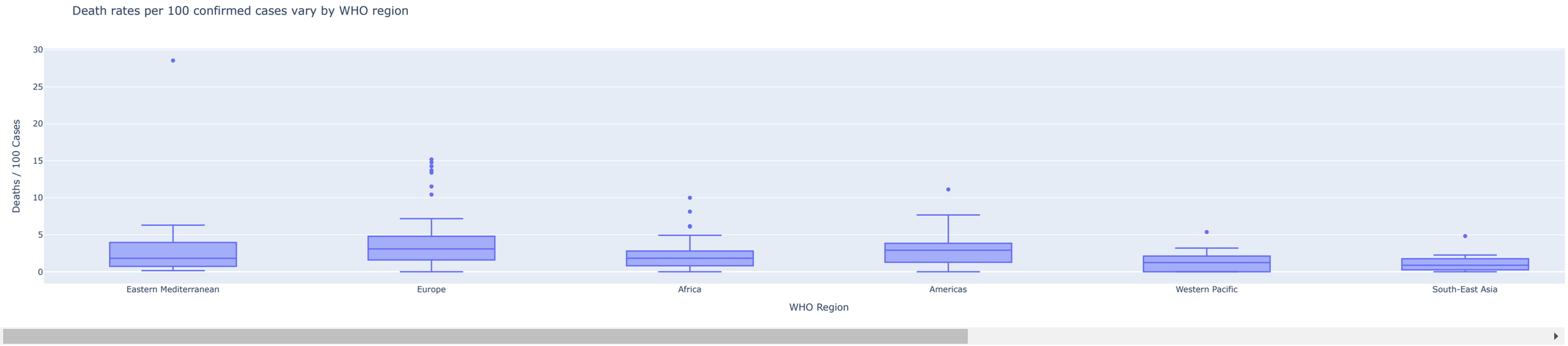
4. What was the 1-week percentage increase in confirmed cases for each country?

```
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 country',
              color_continuous_scale= px.colors.sequential.Turbo_r)
```



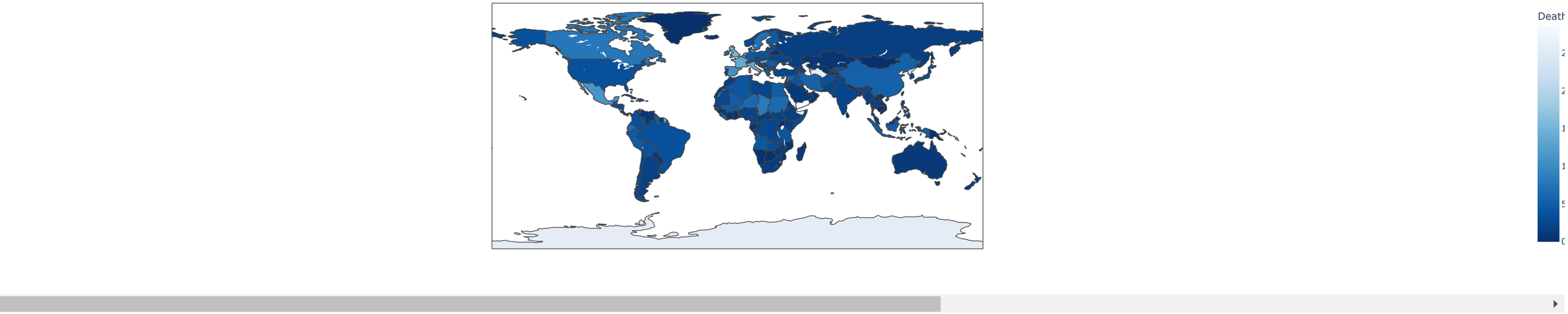
5. How do death rates per 100 confirmed cases vary by WHO region?

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



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

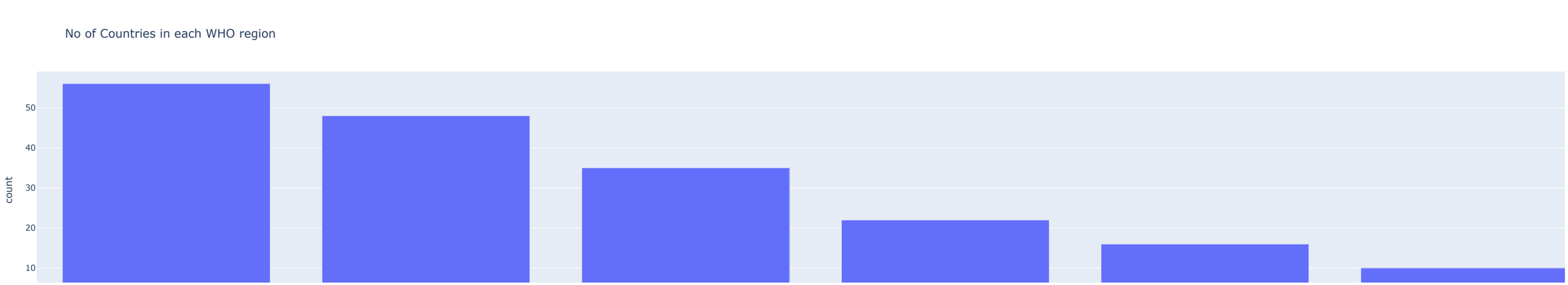
```
px.choropleth(df,
              locations='Country/Region',
              locationmode='country names',
              color='Deaths / 100 Cases',
              hover_name='Country/Region',
              title='Mortality rate (deaths per 100 confirmed cases) for each country',
              color_continuous_scale= px.colors.sequential.R)
```



7. How many countries are in each WHO *region*?

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

```
px.bar(who_region_count,x='WHO Region',y='count',title='No of Countries in each WHO region')
```



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

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