Submission

Put the ipynb file and html file in the github branch you created in the last assignment and submit the link to the commit in brightspace

In [42]:

```
from plotly.offline import init_notebook_mode
import plotly.io as pio
import plotly.express as px

init_notebook_mode(connected=True)
pio.renderers.default = "plotly_mimetype+notebook"
```

In [43]:

```
#Load data
df = px.data.gapminder()
df.head()
```

Out[43]:

	country	continent	year	lifeExp	pop	gdpPercap	iso_alpha	iso_num
0	Afghanistan	Asia	1952	28.801	8425333	779.445314	AFG	4
1	Afghanistan	Asia	1957	30.332	9240934	820.853030	AFG	4
2	Afghanistan	Asia	1962	31.997	10267083	853.100710	AFG	4
3	Afghanistan	Asia	1967	34.020	11537966	836.197138	AFG	4
4	Afghanistan	Asia	1972	36.088	13079460	739.981106	AFG	4

Question 1:

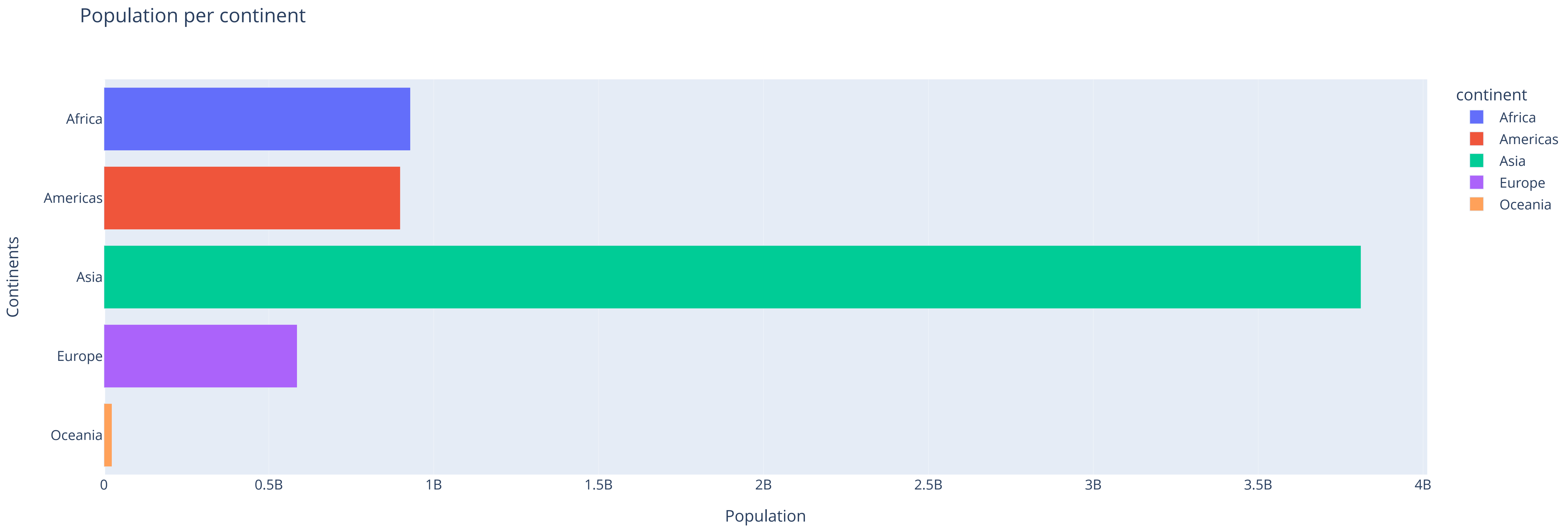
Recreate the barplot below that shows the population of different continents for the year 2007.

Hints:

- Extract the 2007 year data from the dataframe. You have to process the data accordingly
- use plotly bar
- Add different colors for different continents
- Sort the order of the continent for the visualisation. Use axis layout setting
- Add text to each bar that represents the population

In [44]:

```
df_2007 = df.query('year==2007')
dfnew = df_2007.groupby('continent').sum()
fig = px.bar(dfnew, x='pop', title='Population per continent', color=dfnew.index)
fig.update_layout(yaxis_title='Continents', xaxis_title='Population')
fig.show()
```

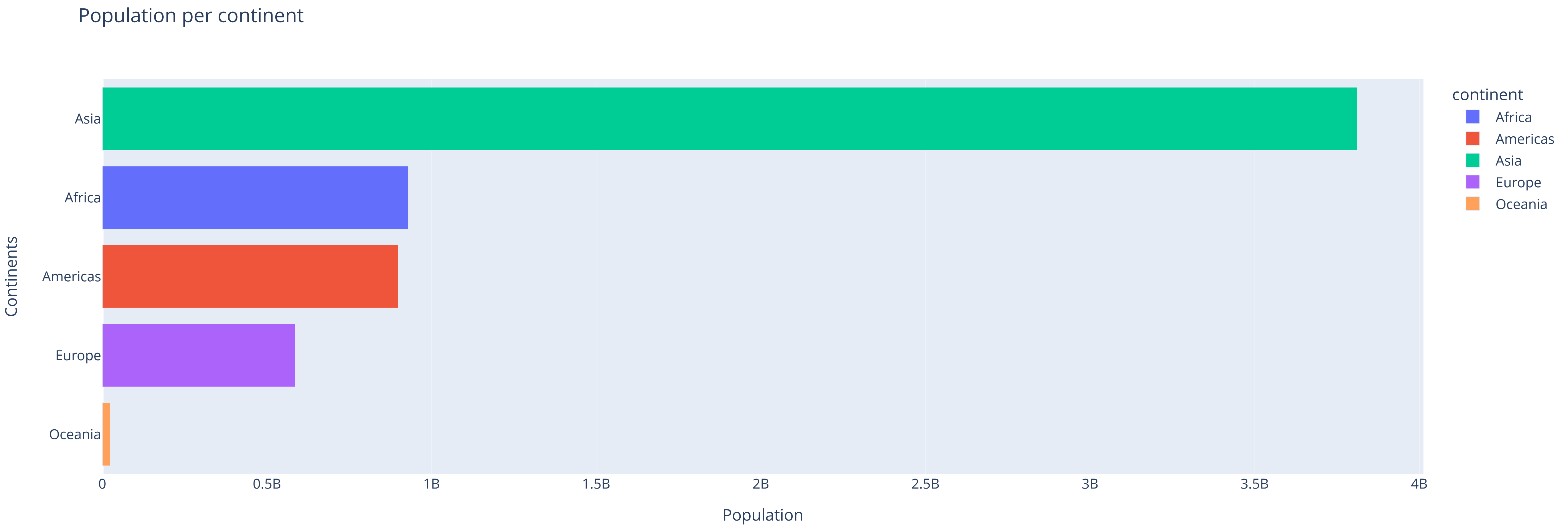


Question 2:

Sort the order of the continent for the visualisation

Hint: Use axis layout setting

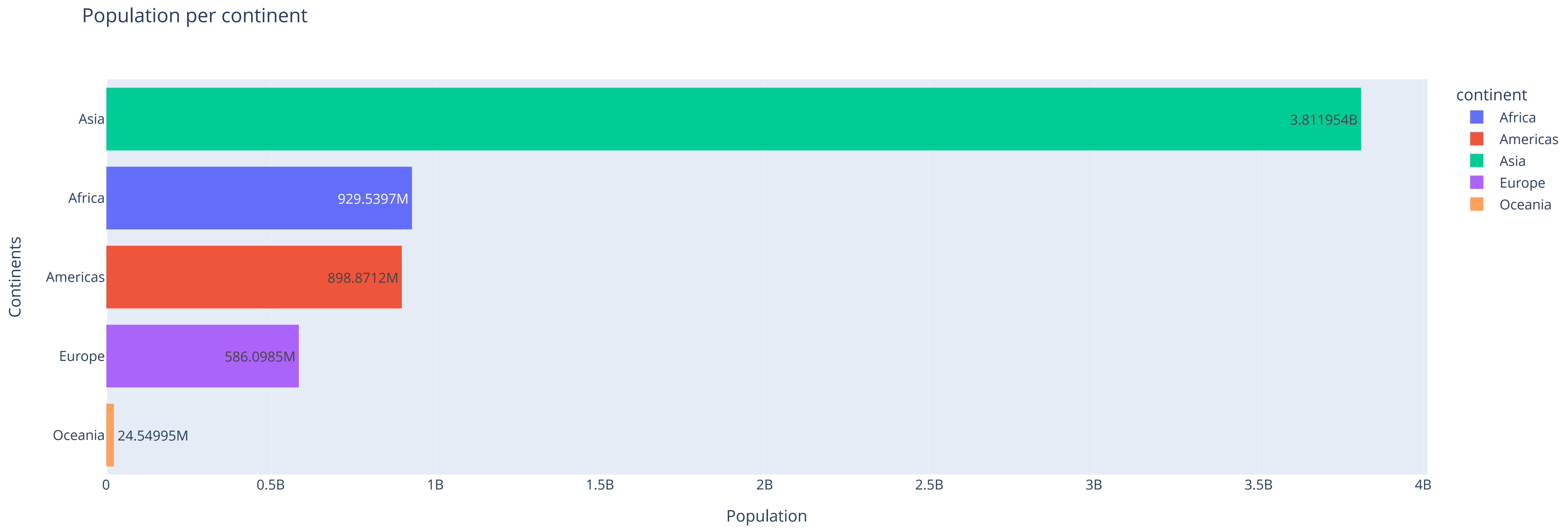
```
In [45]: df_2007 = df.query('year==2007')
         dfnew = df_2007.groupby('continent').sum()
         fig = px.bar(dfnew, x='pop', title='Population per continent', color=dfnew.index)
         fig.update_yaxes(categoryorder="total ascending")
         fig.update_layout(yaxis_title='Continents', xaxis_title='Population')
         fig.show()
```



Question 3:

Add text to each bar that represents the population

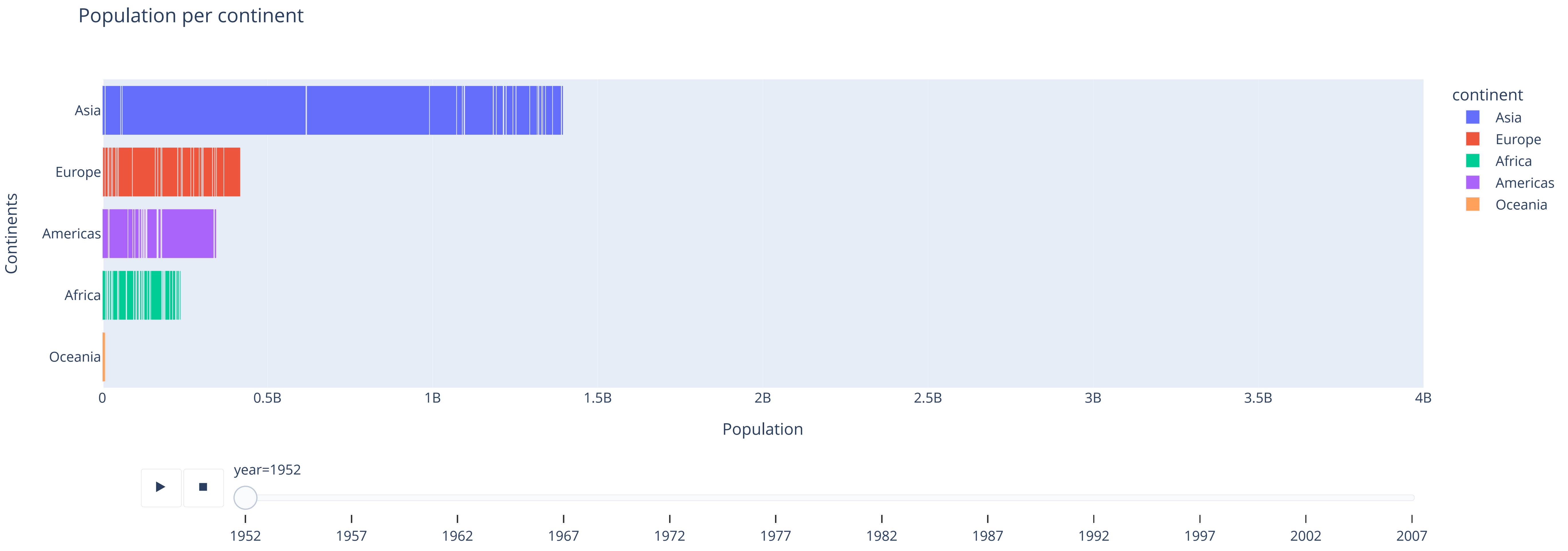
In [46]:
```
df_2007 = df.query('year==2007')
dfnew = df_2007.groupby('continent').sum()
fig = px.bar(dfnew, x='pop', title='Population per continent', color=dfnew.index, text_auto=True)
fig.update_yaxes(categoryorder="total ascending")
fig.update_layout(yaxis_title='Continents', xaxis_title='Population')
fig.show()
```



Question 4:

Thus far we looked at data from one year (2007). Lets create an animation to see the population growth of the continents through the years

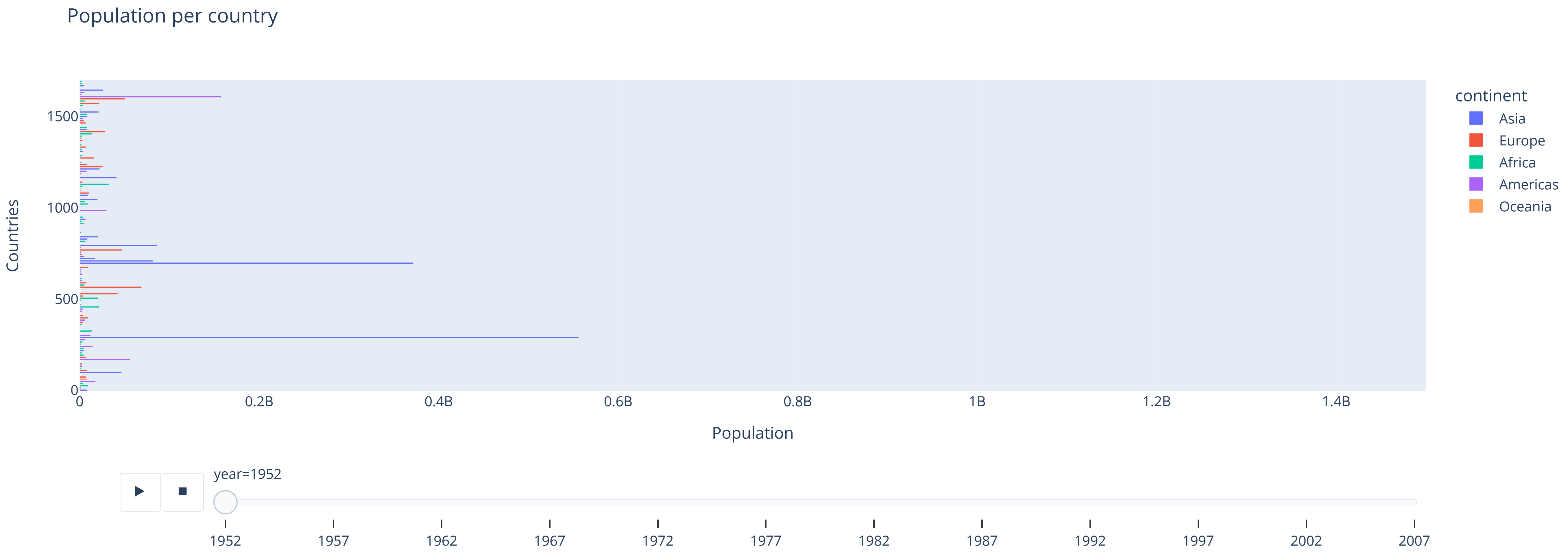
```
In [68]: fig = px.bar(df, x='pop', y='continent', title='Population per continent',
                     color='continent',
                     animation_frame='year',)
         fig.update_xaxes(range=[0, 4*10**9])
         fig.update_yaxes(categoryorder="total ascending")
         fig.update_layout(yaxis_title='Continents', xaxis_title='Population')
         fig.show()
```



Question 5:

Instead of the continents, lets look at individual countries. Create an animation that shows the population growth of the countries through the years

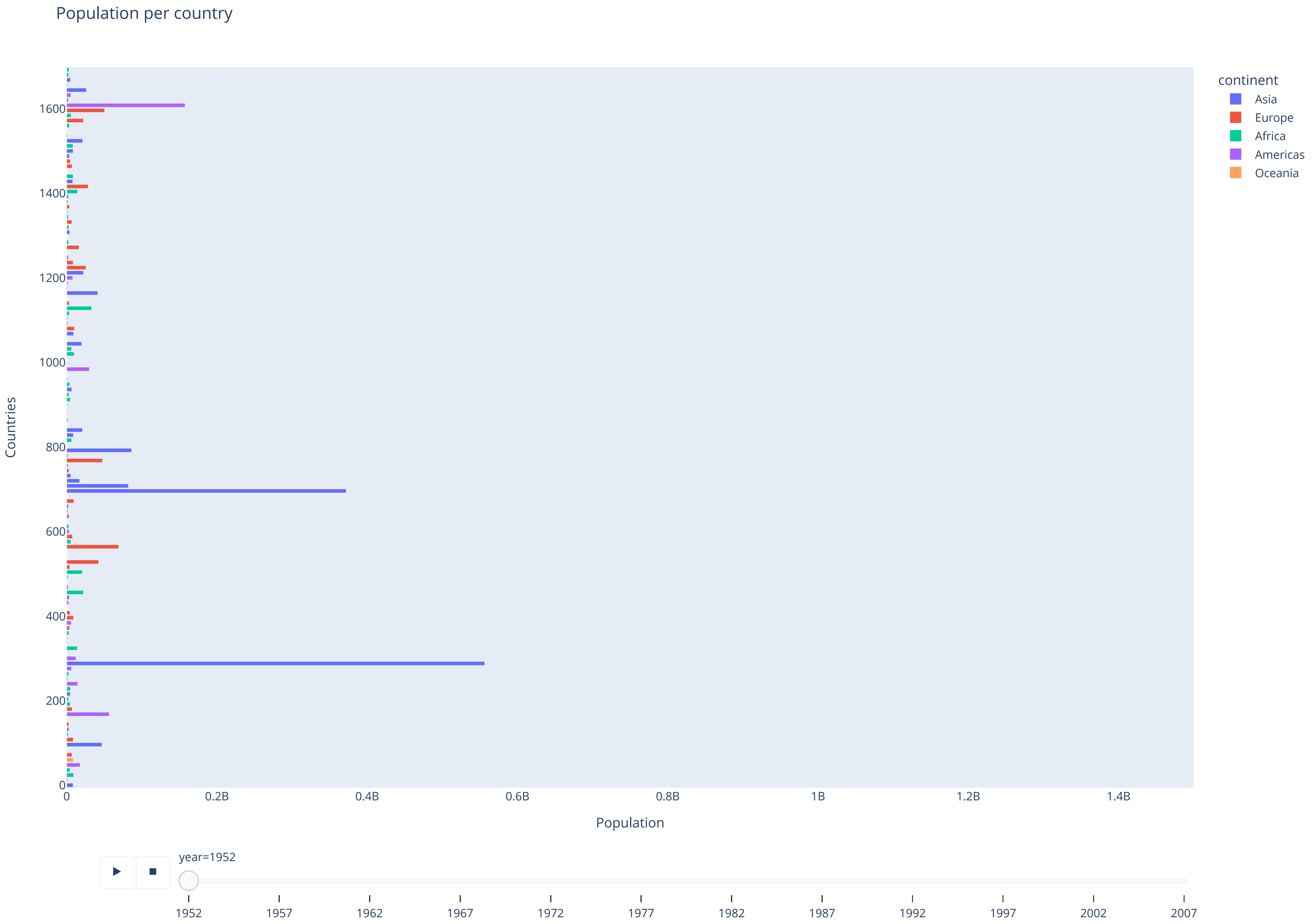
In [64]:
```
fig = px.bar(df, x='pop', title='Population per country', color='continent', animation_frame='year')
fig.update_xaxes(range=[0, 1.5*10**9])
fig.update_yaxes(categoryorder="total ascending")
fig.update_layout(yaxis_title='Countries', xaxis_title='Population')
fig.show()
```



Question 6:

Clean up the country animation. Set the height size of the figure to 1000 to have a better view of the animation

```
In [65]: fig = px.bar(df, x='pop', title='Population per country', color='continent', animation_frame='year', height=1000)
         fig.update_xaxes(range=[0, 1.5*10**9])
         fig.update_yaxes(categoryorder='total ascending')
         fig.update_layout(yaxis_title='Countries', xaxis_title='Population')
         fig.show()
```


Population per country
Countries
1600
1400
1200
1000
800
600
400
200
0
0
0.2B
0.4B
0.6B
0.8B
1B
1.2B
1.4B
Population
continent
Asia
Europe
Africa
Americas
Oceania
year=1952
1952
1957
1962
1967
1972
1977
1982
1987
1992
1997
2002
2007

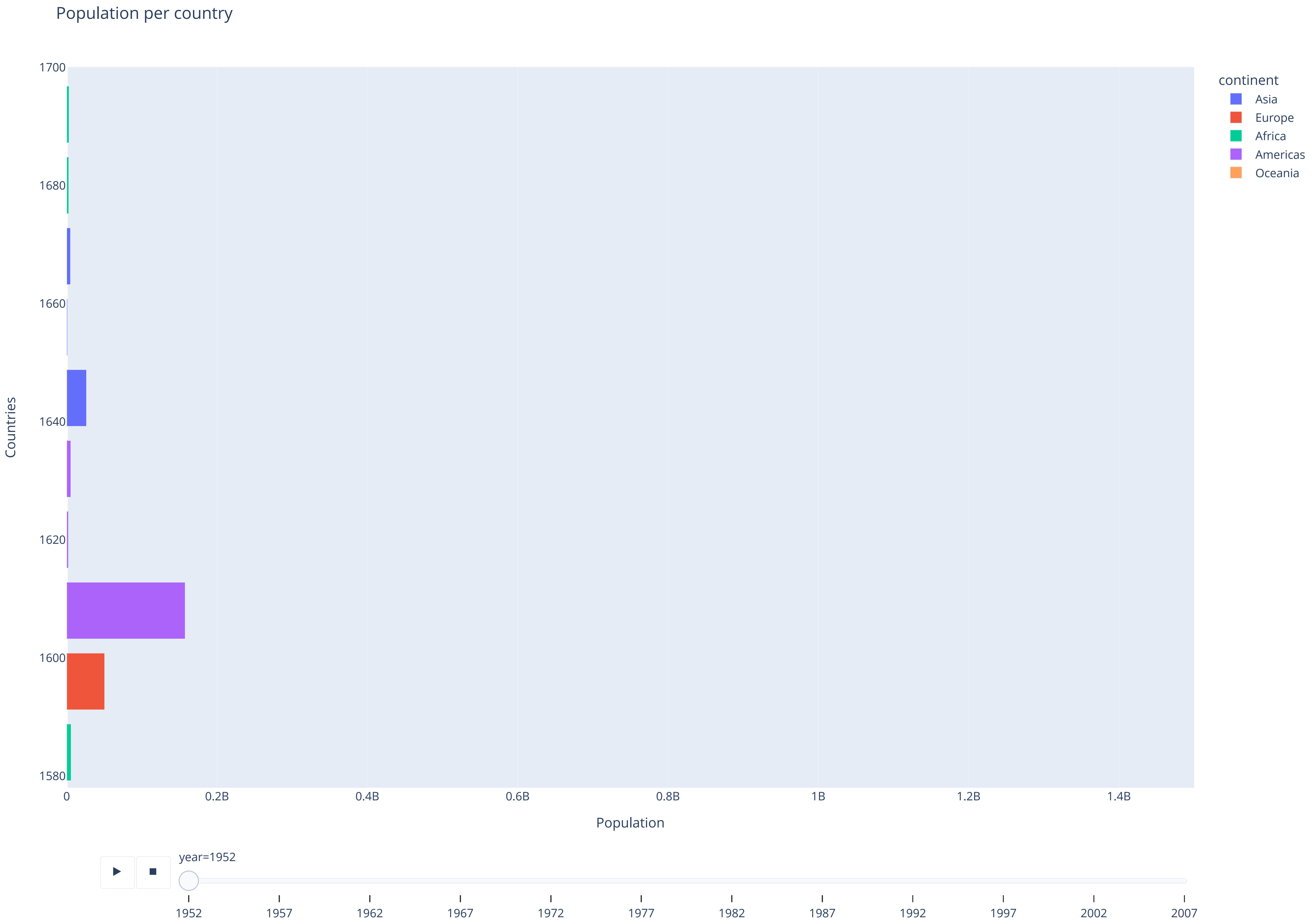
Question 7:

Show only the top 10 countries in the animation

Hint: Use the axis limit to set this.

In [84]:

```
fig = px.bar(df, x='pop', title='Population per country', color='continent', animation_frame='year', height=1000)
fig.update_xaxes(range=[0, 1.5*10**9])
fig.update_yaxes(categoryorder='total ascending', range=(1578, 1700))
#For some reason the categoryorder('total ascending') does not work here, but this code shows the 10 countries at the top
#If categoryorder does work, these would be the top 10 countries
fig.update_layout(yaxis_title='Countries', xaxis_title='Population')
fig.show()
```


Population per country
1700
1680
1660
1640
1620
1600
1580
Countries
0
0.2B
0.4B
0.6B
0.8B
1B
1.2B
1.4B
Population
continent
🟦 Asia
🟥 Europe
🟩 Africa
🟪 Americas
🟧 Oceania
year=1952
▶
■
1952
1957
1962
1967
1972
1977
1982
1987
1992
1997
2002
2007

```
In [ ]:
```