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 [1]:
```
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 [2]:
```
#load data
df = px.data.gapminder()
df.head()
```

Out[2]:

	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:

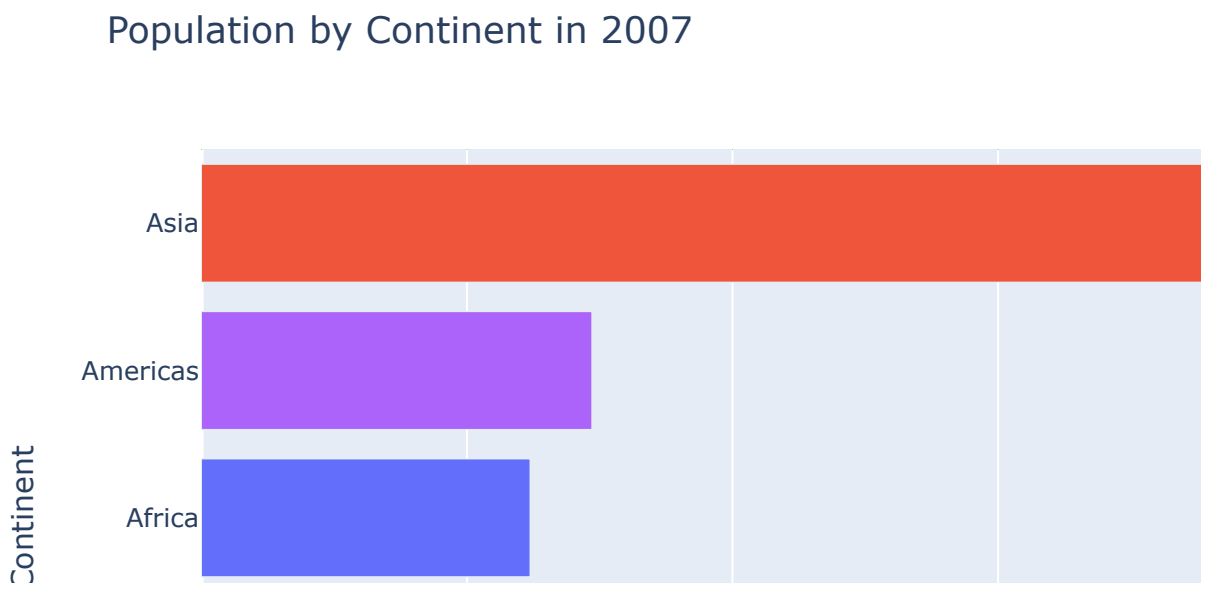
Recreate the barplot below that shows the population of different continents for the year 2007.

Hints:

- Extract the 2007 year data from the dataframe. You have to process the data accordingly
- use plotly bar (https://plotly.com/python-api-reference/generated/plotly.express.bar)
- Add different colors for different continents
- Sort the order of the continent for the visualisation. Use axis layout setting (https://plotly.com/python/reference/layout/xaxis/)
- Add text to each bar that represents the population

In [49]:
```

```

```
# YOUR CODE HERE
import plotly.express as px

continents = ['Africa', 'Asia', 'Europe', 'Americas', 'Oceania']
population_2007 = [6187585961, 30507333901, 6181115304, 7351438499,

fig = px.bar(x=population_2007, y=continents, color=continents,
             title='Population by Continent in 2007',
             orientation='h')

fig.update_layout(
    xaxis_title='Population (Billions)',
    yaxis_title='Continent',
    yaxis={'categoryorder': 'total ascending'},
    showlegend=False )


fig.update_traces(texttemplate='', textposition='outside')

fig.show()
```

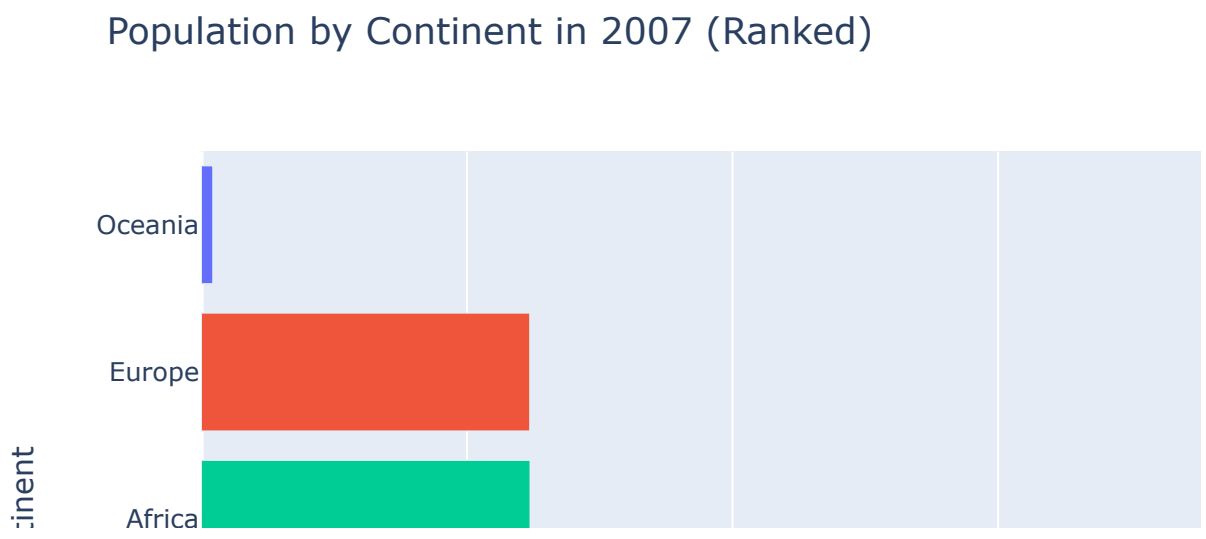


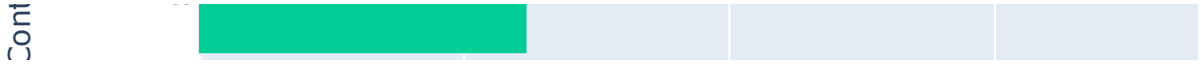
Question 2:

Sort the order of the continent for the visualisation

Hint: Use axis layout setting (https://plotly.com/python/reference/layout/xaxis/)

In [2]:
```
# YOUR CODE HERE
import pandas as pd
import plotly.express as px

data = pd.DataFrame({
    'Continent': ['Africa', 'Asia', 'Europe', 'Americas', 'Oceania'
    'Population_2007': [6187585961, 30507333901, 6181115304, 735143

continent_order = ['Asia', 'Americas', 'Africa', 'Europe', 'Oceania

data['Continent'] = pd.Categorical(data['Continent'], categories=co
data = data.sort_values(by='Continent', ascending=False)

data['Population_2007_Billions'] = data['Population_2007'] / 1e9

fig = px.bar(data, x='Population_2007_Billions', y='Continent', col
             title='Population by Continent in 2007 (Ranked)',
             orientation='h')

fig.update_layout(
    xaxis_title='Population (Billions)',
    yaxis_title='Continent',
    showlegend=False
)

fig.update_traces(texttemplate=None, textposition='outside')

fig.show()
```



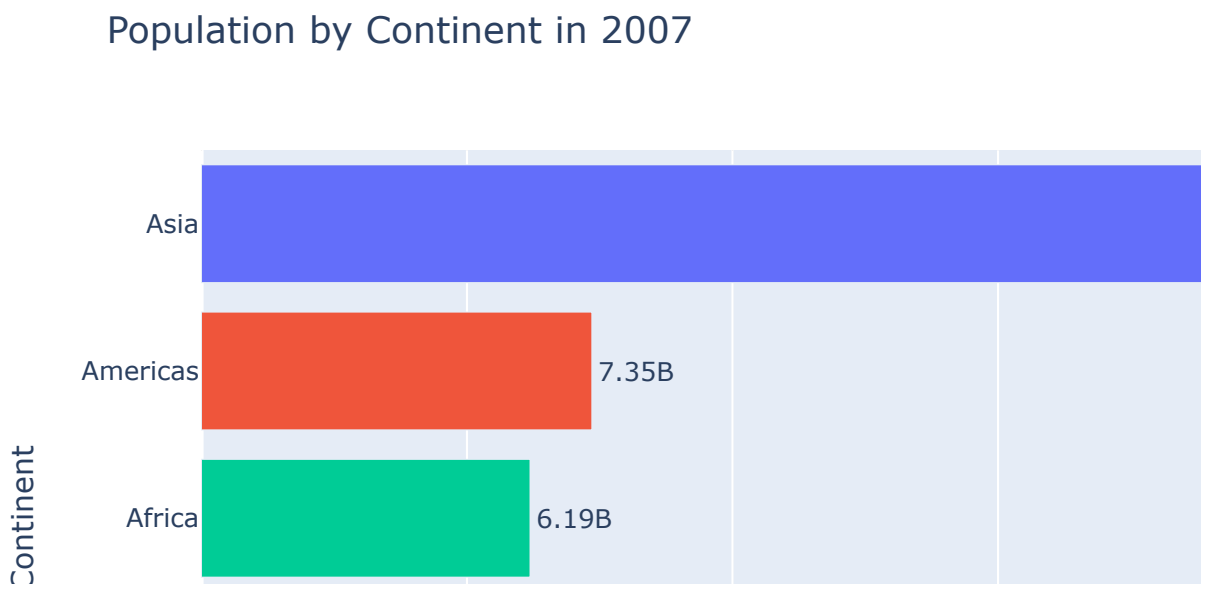


Question 3:

Add text to each bar that represents the population

In [3]:

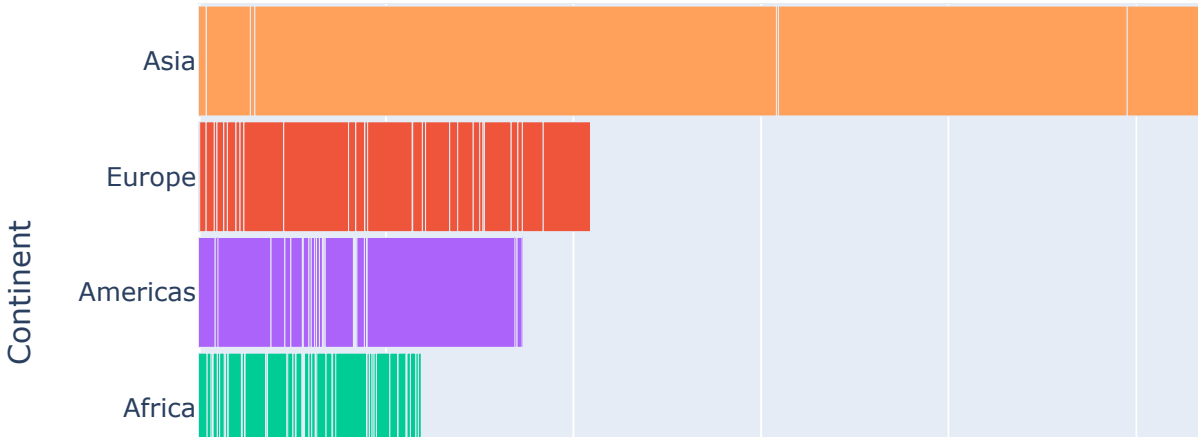
```
# YOUR CODE HERE
import pandas as pd
import plotly.express as px

data = pd.DataFrame({
    'Continent': ['Africa', 'Asia', 'Europe', 'Americas', 'Oceania'
    'Population_2007': [6187585961, 30507333901, 6181115304, 735143
})


continent_order = ['Americas', 'Europe', 'Asia', 'Africa', 'Oceania

data = data.groupby('Continent').sum().reset_index()
data['Continent'] = pd.Categorical(data['Continent'], categories=co
data = data.sort_values(by=['Population_2007'], ascending=[False])

data['Population_2007_Billions'] = data['Population_2007'] / 1e9

fig = px.bar(data, x='Population_2007_Billions', y='Continent', col
             text='Population_2007_Billions', title='Population by
             orientation='h')

fig.update_layout(
    xaxis_title='Population (Billions)',
    yaxis_title='Continent',
    showlegend=False
)

fig.update_traces(texttemplate='%{text:.2f}B', textposition='outsid

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

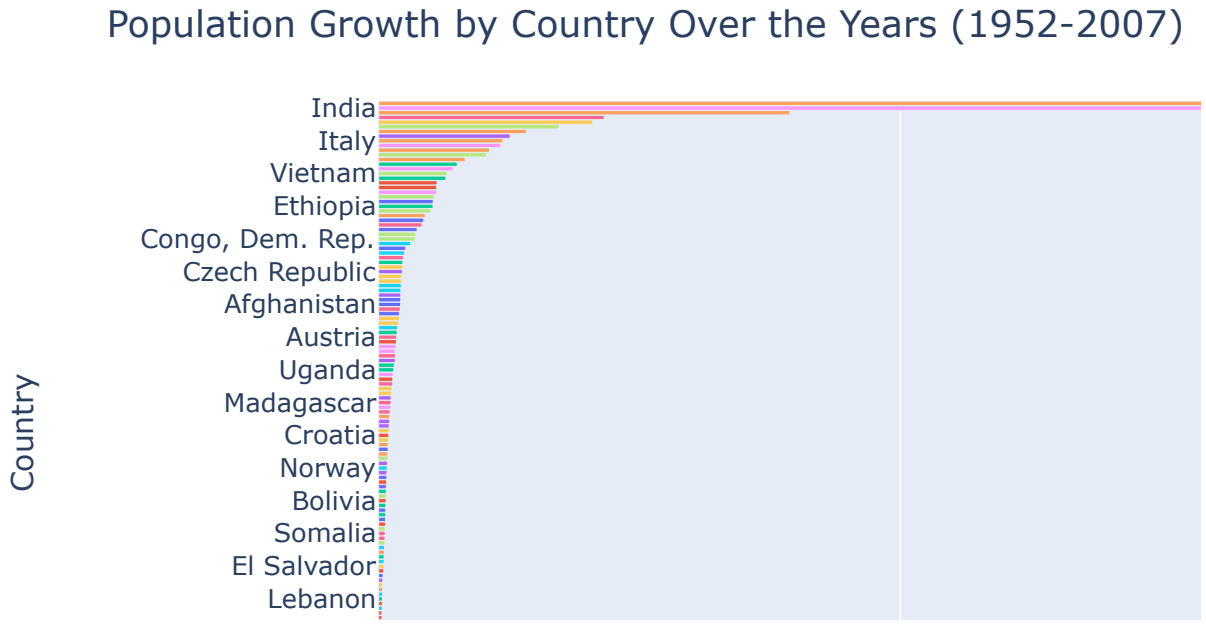
In [9]:
```
# YOUR CODE HERE
import pandas as pd
import plotly.express as px

df = px.data.gapminder()

continent_order = ['Asia', 'Americas', 'Africa', 'Europe', 'Oceania

df['continent'] = pd.Categorical(df['continent'], categories=contin
df = df.sort_values(by=['continent', 'year'], ascending=[False, Tru

df['pop_billions'] = df['pop'] / 1e9

fig = px.bar(df, x='pop_billions', y='continent', color='continent'
             animation_frame='year', range_x=[0, max(df['pop_billio
             title='Population Growth by Continent Over the Years')

fig.update_layout(
    xaxis_title='Population (Billions)',
    yaxis_title='Continent',
    showlegend=False,
    xaxis_tickformat=',.1fB',
    bargap=0.05,
    yaxis={'categoryorder': 'total ascending'})

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

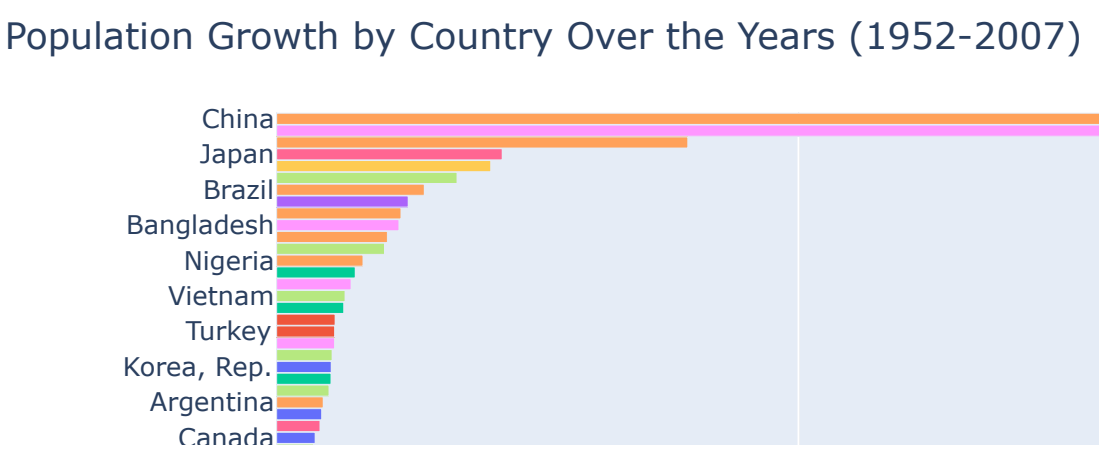
In [5]:
```
# YOUR CODE HERE
import plotly.express as px

df = px.data.gapminder()

fig = px.bar(df, x='pop', y='country', animation_frame='year',
             orientation='h', color='country', text=None)

fig.update_layout(
    yaxis_title='Country',
    xaxis_title='Population',
    title='Population Growth by Country Over the Years (1952-2007)'
    showlegend=False,
    bargap=0.05,
    yaxis={'categoryorder': 'total ascending'})

tickvals = [0, 0.2e9, 0.4e9, 0.6e9, 0.8e9, 1.0e9, 1.2e9, 1.4e9]
ticktext = ['0.0', '0.2', '0.4', '0.6', '0.8', '1.0', '1.2', '1.4']
fig.update_xaxes(
    tickvals=tickvals,
    ticktext=ticktext
)

fig.update_xaxes(tickformat=',.1fB')

fig.show()
```

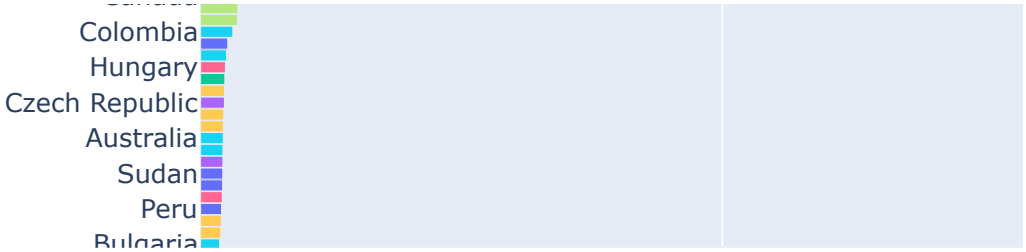


Question 6:

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

In [4]:
```
# YOUR CODE HERE
import plotly.express as px

df = px.data.gapminder()

fig = px.bar(df, x='pop', y='country', animation_frame='year',
             orientation='h', color='country', text=None)

fig.update_layout(
    yaxis_title='Country',
    xaxis_title='Population',
    title='Population Growth by Country Over the Years (1952-2007)'
    showlegend=False,
    bargap=0.05,
    yaxis={'categoryorder': 'total ascending'},
    height=1000
)

tickvals = [0, 0.2e9, 0.4e9, 0.6e9, 0.8e9, 1.0e9, 1.2e9, 1.4e9]
ticktext = ['0.0', '0.2', '0.4', '0.6', '0.8', '1.0', '1.2', '1.4']
fig.update_xaxes(
    tickvals=tickvals,
    ticktext=ticktext
)

fig.update_xaxes(tickformat=',.1fB')

fig.show()
```





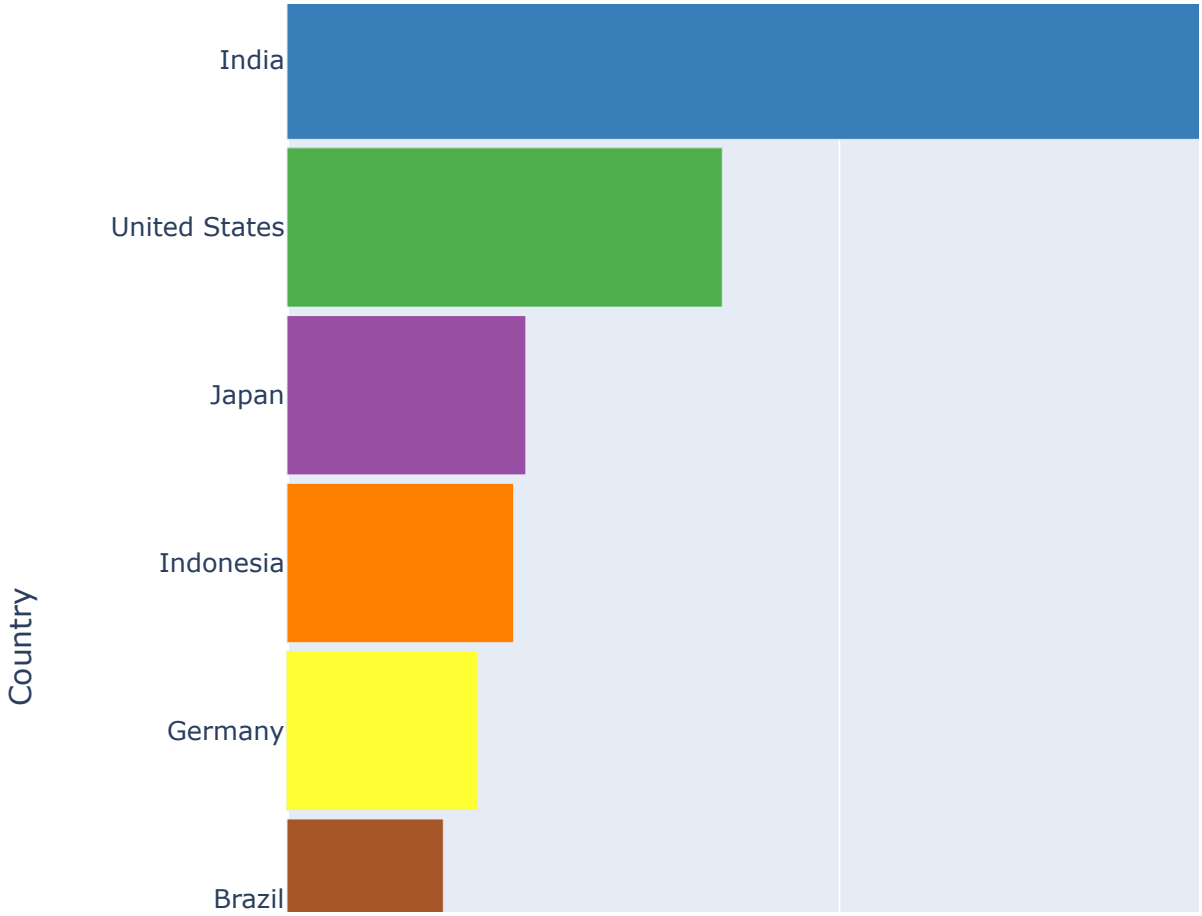
Question 7:

Show only the top 10 countries in the animation

Hint: Use the axis limit to set this.

In [7]:

```
# YOUR CODE HERE
import plotly.express as px

df = px.data.gapminder()

top_10_df = pd.DataFrame()

years = df['year'].unique()
for year in years:
    df_year = df[df['year'] == year]
    df_year_sorted = df_year.sort_values(by='pop', ascending=False)
    top_10_df = pd.concat([top_10_df, df_year_sorted])

fig = px.bar(top_10_df, x='pop', y='country', animation_frame='year
             orientation='h', color='country',
             labels={'country': 'Country'},
             color_discrete_sequence=px.colors.qualitative.Set1)

fig.update_layout(
    xaxis_title='Population (Billions)',
    yaxis_title='Country',
    title='Top 10 Countries by Population Over the Years (1952-2007
    showlegend=False,
    bargap=0.05,
    yaxis={'categoryorder': 'total ascending'},
    height=1000  )


tickvals = [0, 0.2e9, 0.4e9, 0.6e9, 0.8e9, 1.0e9, 1.2e9, 1.4e9]
ticktext = ['0.0', '0.2', '0.4', '0.6', '0.8', '1.0', '1.2', '1.4']
fig.update_xaxes(
    tickvals=tickvals,
    ticktext=ticktext
)

fig.show()
```


India
United States
Japan
Indonesia
Germany
Brazil
Country