

## 1. Install Bar Chart Race

```
In [1]: # Install package in the current Jupyter kernel
import sys
!{sys.executable} -m pip install bar_chart_race
```

Requirement already satisfied: bar\_chart\_race in c:\users\user\appdata\local\programs\python\python38-32\lib\site-packages (0.1.0)
Requirement already satisfied: pandas>=0.24 in c:\users\user\appdata\local\programs\python\python38-32\lib\site-packages (from bar\_chart\_race) (1.0.3)
Requirement already satisfied: matplotlib>=3.1 in c:\users\user\appdata\local\programs\python\python38-32\lib\site-packages (from bar\_chart\_race) (3.2.1)
Requirement already satisfied: python-dateutil>=2.6.1 in c:\users\user\appdata\local\programs\python\python38-32\lib\site-packages (from pandas>=0.24->bar\_chart\_race) (2.8.1)
Requirement already satisfied: numpy>=1.13.3 in c:\users\user\appdata\local\programs\python\python38-32\lib\site-packages (from pandas>=0.24->bar\_chart\_race) (1.18.4)
Requirement already satisfied: pytz>=2017.2 in c:\users\user\appdata\local\programs\python\python38-32\lib\site-packages (from pandas>=0.24->bar\_chart\_race) (2020.1)
Requirement already satisfied: cycler>=0.10 in c:\users\user\appdata\local\programs\python\python38-32\lib\site-packages (from matplotlib>=3.1->bar\_chart\_race) (0.10.0)
Requirement already satisfied: pyparsing!=2.0.4,!2.1.2,!2.1.6,>2.0.1 in c:\users\user\appdata\local\programs\python\python38-32\lib\site-packages (from matplotlib>=3.1->bar\_chart\_race) (2.4.7)
Requirement already satisfied: kiwisolver>=1.0.1 in c:\users\user\appdata\local\programs\python\python38-32\lib\site-packages (from matplotlib>=3.1->bar\_chart\_race) (1.2.0)
Requirement already satisfied: six>=1.5 in c:\users\user\appdata\local\programs\python\python38-32\lib\site-packages (from python-dateutil>=2.6.1->pandas>=0.24->bar\_chart\_race) (1.15.0)

## 2. Importing Libraries

```
In [2]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import bar_chart_race as bcr
from IPython.display import HTML
import warnings
warnings.filterwarnings("ignore")
from IPython.display import Video
```

## 3. Loading Data

```
In [3]: df = pd.read_csv('https://git.io/fjpo3', usecols=['name', 'group', 'year', 'value'])
df.head(10)
```

```
Out[3]:
```

	name	group	year	value
0	Agra	India	1575	200.0
1	Agra	India	1576	212.0
2	Agra	India	1577	224.0
3	Agra	India	1578	236.0
4	Agra	India	1579	248.0
5	Agra	India	1580	260.0
6	Agra	India	1581	272.0
7	Agra	India	1582	284.0
8	Agra	India	1583	296.0
9	Agra	India	1584	308.0

## 4. Inspecting the Dataframe

```
In [4]: df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6252 entries, 0 to 6251
Data columns (total 4 columns):
 #   Column  Non-Null Count  Dtype  
--- 
 0   name    6252 non-null   object  
 1   group   6252 non-null   object  
 2   year    6252 non-null   int64  
 3   value   6252 non-null   float64 
dtypes: float64(1), int64(1), object(2)
memory usage: 146.6+ KB
```

```
In [5]: df.isnull().sum()
```

```
Out[5]:
```

	name	0	group	0	year	0	value	0	dtype	int64
--	------	---	-------	---	------	---	-------	---	-------	-------

## 5. Data Transformations

```
In [6]: current_year = 2018
dff = df[df['year'].eq(current_year)].sort_values(by='value', ascending=True).head(10)
dff
```

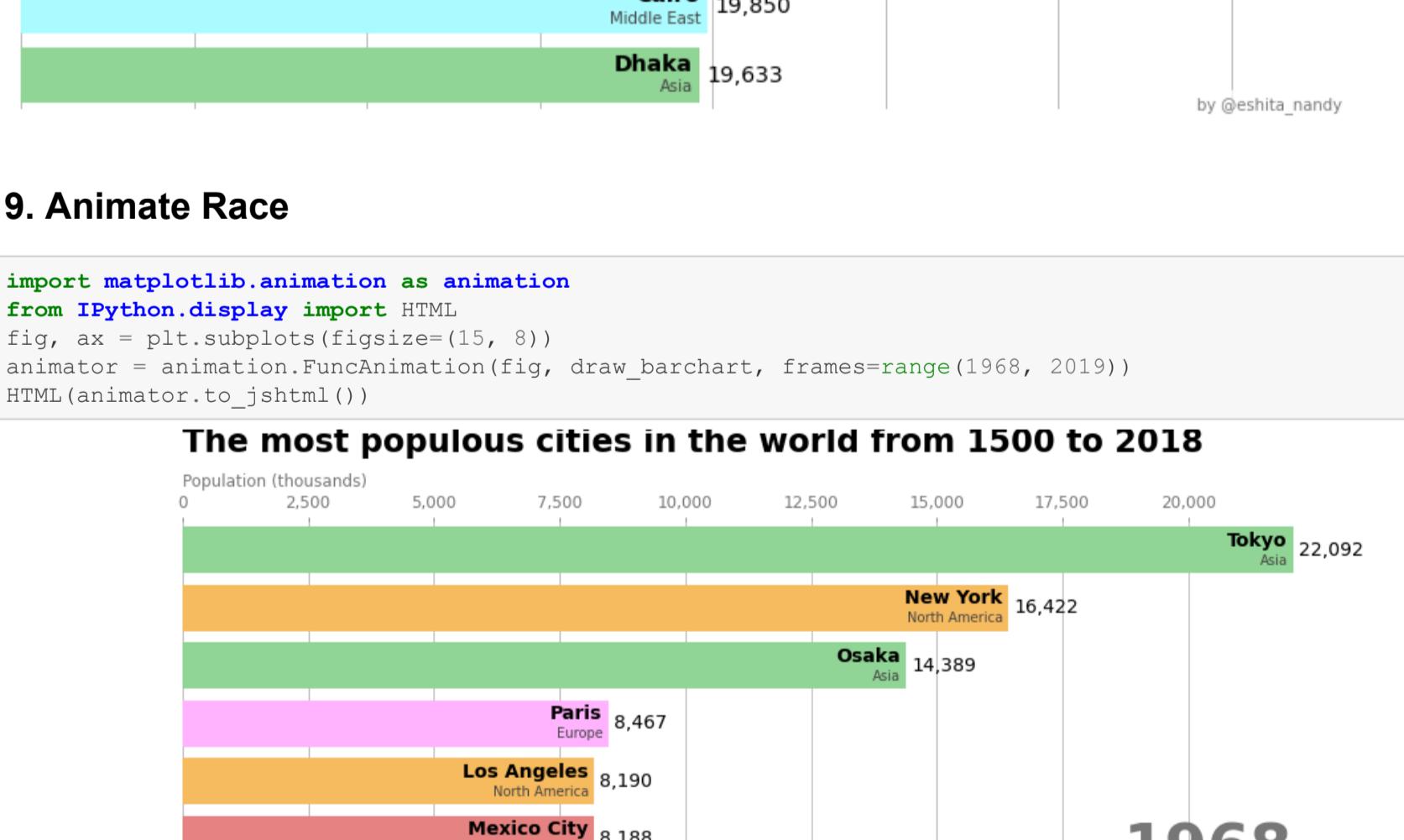
```
Out[6]:
```

	name	group	year	value
2537	Karachi	Asia	2018	18185.2
4327	New York	North America	2018	18713.0
1336	Dhaka	Asia	2018	19632.6
1195	Cairo	Middle East	2018	19849.6
4679	Osaka	Asia	2018	20409.0
3574	Mexico City	Latin America	2018	21520.4
5445	Sao Paulo	Latin America	2018	21697.8
3748	Mumbai	India	2018	22120.0
689	Beijing	Asia	2018	22674.2
5547	Shanghai	Asia	2018	25778.6

## 6. Basic Chart

```
In [7]: fig, ax = plt.subplots(figsize=(15, 8))
ax.bart(dff['name'], dff['value'])
```

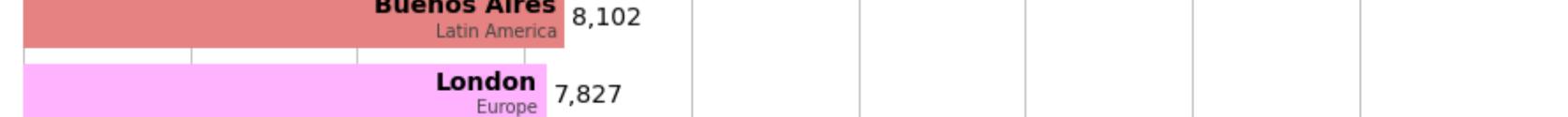
```
Out[7]: <BarContainer object of 10 artists>
```



## 7. Color and Labels

```
In [8]: colors = dict(zip(
    ['India', 'Europe', 'Asia', 'Latin America', 'Middle East', 'North America', 'Africa'],
    ['#adb0ff', '#ffb3ff', '#90d595', '#e48381', '#aafbff', '#f7bb5f', '#eafb50']
))
group_lk = df.set_index('name')[['group']].to_dict()
```

```
In [9]: fig, ax = plt.subplots(figsize=(15, 8))
dff = dff[-1:] # flip values from top to bottom
# pass colors values to 'color='
ax.bart(dff['name'], dff['value'], color=[colors[group_lk[x]] for x in dff['name']])
for i, (value, name) in enumerate(zip(dff['value'], dff['name'])):
    ax.text(value-dx, i, name, ha='right') # Tokyo: name
    ax.text(value, i-.25, group_lk[name], ha='right') # Asia: group name
    ax.text(value, i, value, ha='left') # 38194.2: value
# Add year right middle portion of canvas
ax.text(1, 0.4, current_year, transform=ax.transAxes, size=46, ha='right')
plt.show(1)
```



## 8. Polish Style

```
In [10]: import matplotlib.ticker as ticker
fig, ax = plt.subplots(figsize=(15, 8))
def draw_barchart():
    dff = df[df['year'].eq(year)].sort_values(by='value', ascending=True).tail(10)
    ax.clear()
    ax.bart(dff['name'], dff['value'], color=[colors[group_lk[x]] for x in dff['name']])
    dx = dff['value'].max() / 200
    for i, (value, name) in enumerate(zip(dff['value'], dff['name'])):
        ax.text(value-dx, i, name, size=14, weight=600, ha='right', va='bottom')
        ax.text(value-dx, i-.25, group_lk[name], size=10, color='#444444', ha='right', va='baseline')
        ax.text(value-dx, i, f'{value:.0f}', size=14, ha='left', va='center')
    # ... polished styles
    ax.text(1, 0.4, year, transform=ax.transAxes, size=46, ha='right', weight=800)
    ax.text(0, 1.06, 'Population (thousands)', transform=ax.transAxes, size=12, color="#777777")
    ax.xaxis.set_major_formatter(ticker.StrMethodFormatter('{x:,.0f}'))
    ax.xaxis.set_ticks_position('top')
    ax.set_yticks([])
    ax.margins(0, 0.01)
    ax.grid(which='major', axis='x', linestyle='-' )
    ax.set_axisbelow(True)
    ax.text(1, 0.12, 'The most populous cities in the world from 1500 to 2018',
           transform=ax.transAxes, size=24, weight=600, ha='left')
    ax.text(1, 0.04, 'by @eshita_nandy', transform=ax.transAxes, ha='right',
           color="#777777", bbox=dict(facecolor='white', alpha=0.8, edgecolor='white'))
    plt.box(False)
draw_barchart(2018)
```



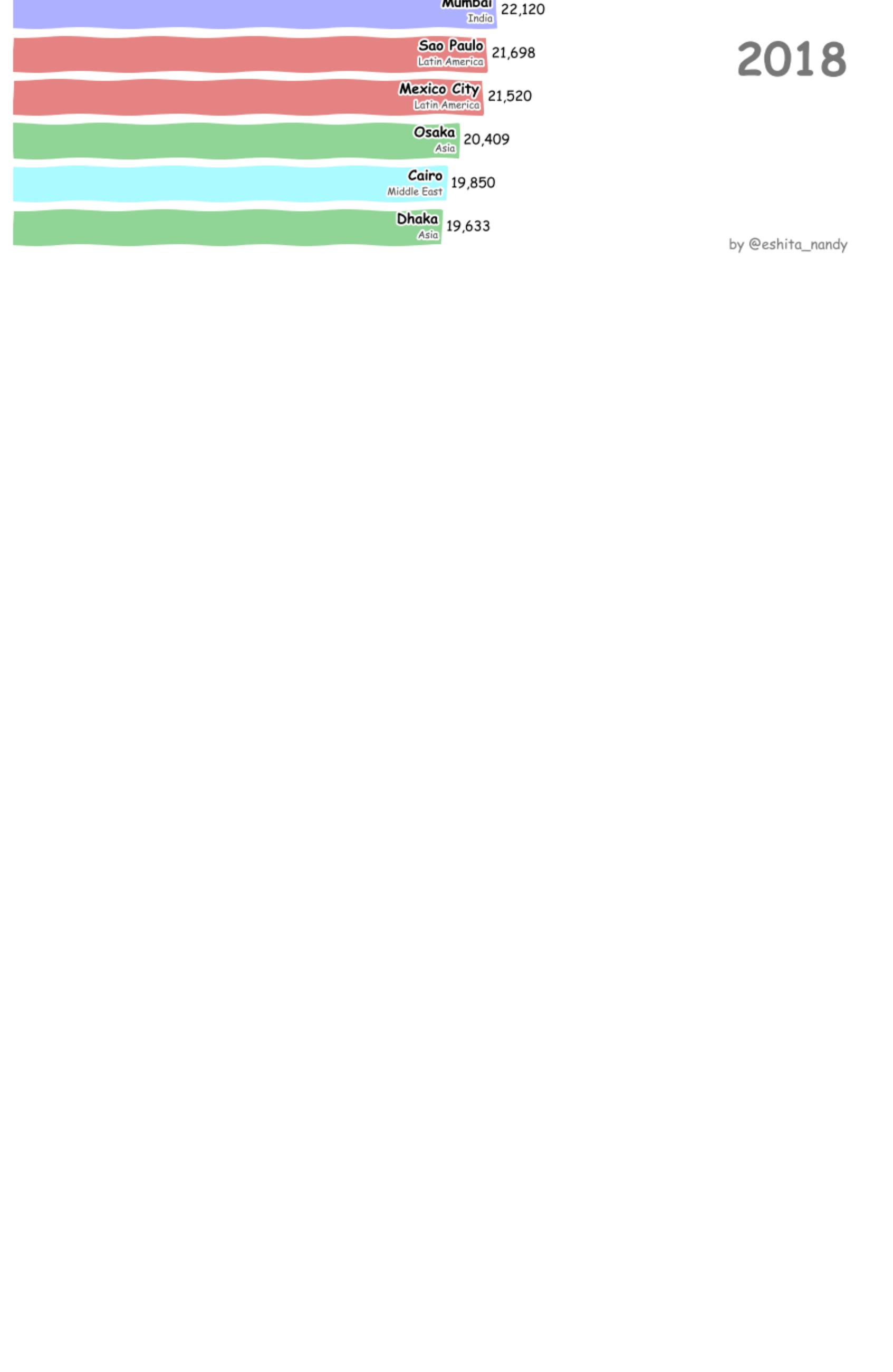
## 9. Animate Race

```
In [11]: import matplotlib.animation as animation
from IPython.display import HTML
fig, ax = plt.subplots(figsize=(15, 8))
anim = animation.FuncAnimation(fig, draw_barchart, frames=range(1968, 2019))
HTML(anim.to_jshtml())
```



```
In [12]: with plt.xkcd():
    fig, ax = plt.subplots(figsize=(15, 8))
    draw_barchart(2018)
```

## The most populous cities in the world from 1500 to 2018



2018

by @eshita\_nandy