

Population Pyramids

M: 941519

June 1, 2022

Contents

1	Cos'è una "population pyramid"	2
1.1	Perché è chiamata una "piramide"?	3
2	I tipi di population pyramid	3
3	I dati necessari per costruire una "population pyramid"	4
4	Esempio di costruzione utilizzando python	4
4.1	Costruzione di diagramma a barre	4
4.1.1	Costruzione della population pyramid	5
4.2	Dependancy ratio	7
5	Cosa ci dice una "population pyramid"	9
5.1	Un'appendice sull'età	9
5.1.1	Perché l'età è un indicatore così importante	9
6	Estensione delle population pyramids	10

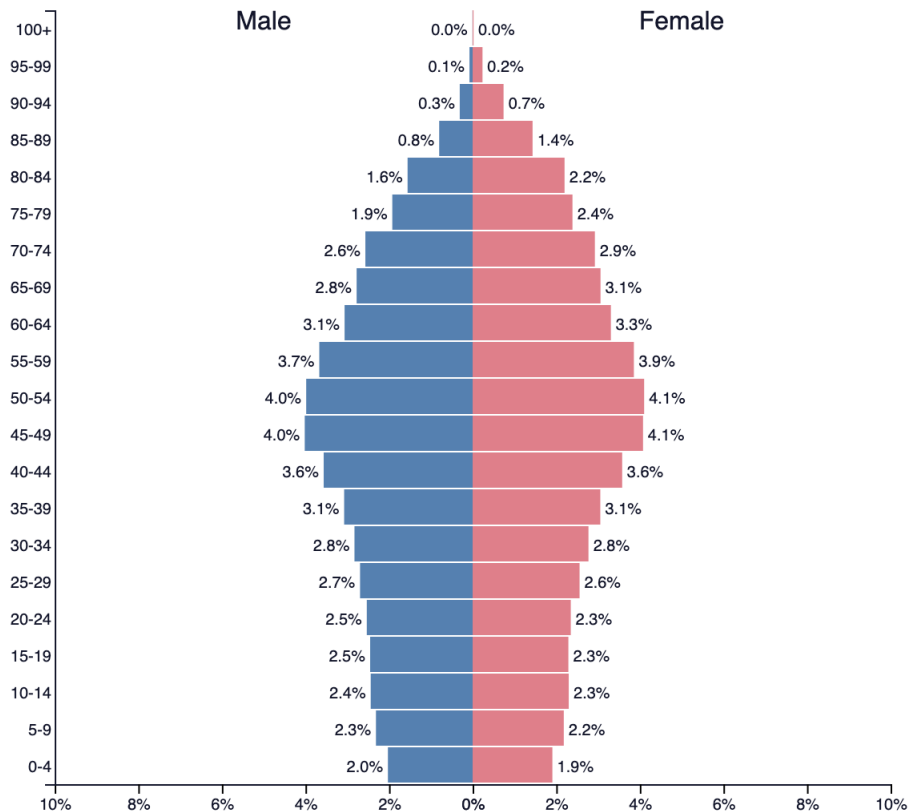
1 Cos'è una "population pyramid"

Una "population pyramid" è una illustrazione grafica che ci permette di visualizzare la distribuzione di una popolazione (che può essere un paese in particolare, una regione del mondo, o un continente). E' formata da due istogrammi che rappresentano le distribuzioni maschili e femminili di un paese, disposti simmetricamente intorno ad un'asse verticale che rappresenta le età.

La population pyramid è formata quindi da una serie di barre ad istogramma, la popolazione è rappresentata in proporzione o in valore assoluto dalla grandezza dell'asse orizzontale, ed le fasce di anni sono rappresentati dall'asse y .

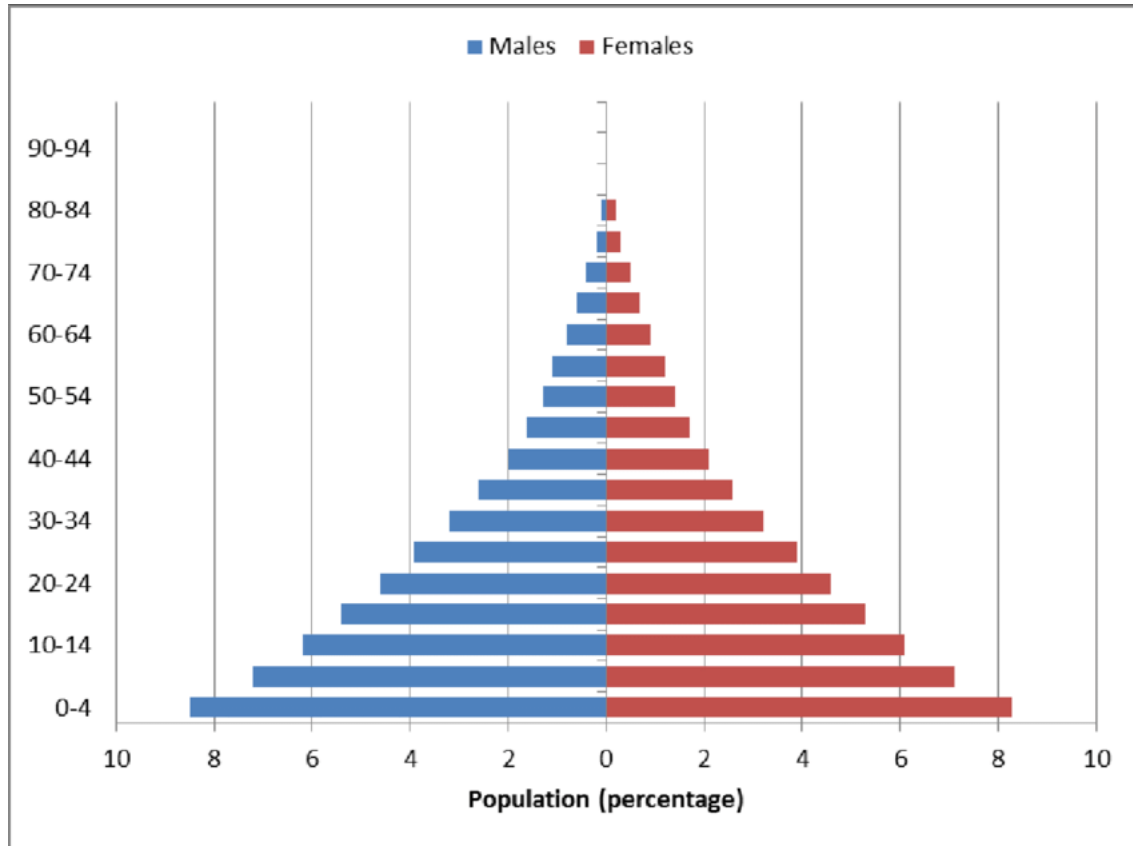
Italy ▼
2019

Population: 60,550,092



1.1 Perché è chiamata una "piramide"?

L'origine del nome deriva dalla forma che questo grafico ha assunto per la maggior parte degli ultimi 200 anni, ovvero quella della piramide:



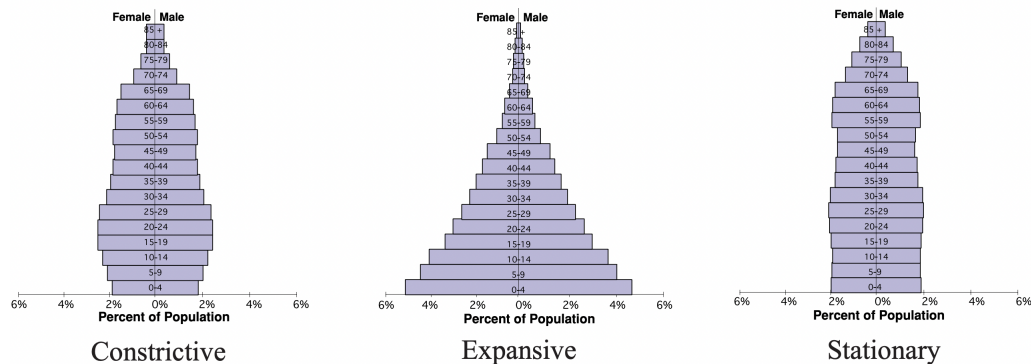
Caratterizzata da una forma chiaramente piramidale, senza buoni contraccettivi e con una popolazione poco educata, la popolazione cresce man mano che ci spostiamo verso il basso del grafico, le famiglie sono quindi composte da molti bambini, che in gran parte non raggiunge neanche la maggiore età, e la popolazione diminuisce progressivamente per la mancanza di buona medicina e prevenzione. Quella comunque, è la "population pyramid" dell'Africa Subsahariana.

2 I tipi di population pyramid

- Constrictive o in compressione: rappresenta una piramide che ha meno persone nella fascia di persone giovani, ed è il tipo di forma assunta dalle parte delle società "sviluppate", con un alto tasso di sviluppo, un tasso di mortalità infantile e natalità basso, con una popolazione in via di compressione.

- **Expansive** o espansiva: rappresenta la piramide con forma piramidale, è caratterizzata da una fascia di persone giovani nettamente più grandi delle altre, e distingue paesi in cui i tassi di mortalità infantile e natalità sono alti, principalmente è la forma assunta dalle società in via di sviluppo.
- **Stationary** o stazionaria: rappresenta una piramide pressoché bilanciata in tutte le fasce, questo va a caratterizzare società in cui i tassi di natalità sono bassi, ma pressoché costanti, con un alta qualità di vita.

Qualcuno ha voglia di indovinare che tipo di piramide avranno i seguenti paesi: Burkina Faso (espansiva), Stati Uniti (stazionaria), Italia (in compressione), Norvegia (stazionaria)



3 I dati necessari per costruire una "population pyramid"

Per costruire una population pyramid abbiamo bisogno prima di tutto della popolazione totale in uno specifico anno, nel nostro esempio abbiamo preso l'Italia nell'anno 2019, e la popolazione totale è di 60.550.092.

Il secondo dato fondamentale che ci serve è la proporzione, rispetto alla popolazione totale, della popolazione per ogni "age bracket". Questo dato va a rispondere alla domanda: *"quante persone, maschio e femmina, vi erano, in percentuale, rispetto alla popolazione totale ?"*. Questo dato ci permette di dimensionare le varie barre del nostro grafico

L'ultimo dato fondamentale è la proporzione tra maschi e femmine, rispetto ad ogni "age bracket", che ci permette di dimensionare, in proporzione, i maschi e le femmine.

4 Esempio di costruzione utilizzando python

4.1 Costruzione di diagramma a barre

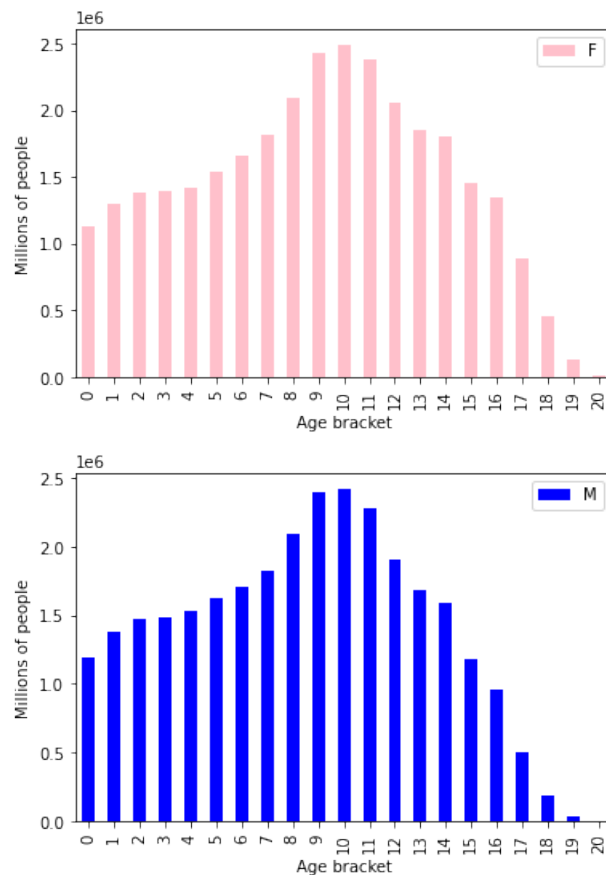
Siccome il dataset raccoglie già i valori di popolazione per age bracket, il primo grafico che possiamo andare a costruire è il bar chart. Il bar chart è un grafico formato da rettangoli la cui altezza è dimensionata in base a valore che vanno a rappresentare.

Siccome vogliamo lavorare con un subset dei dati presenti, effettuiamo uno slicing sul dataframe iniziale, e chiamiamo la funzione plot, con argument *kind* = 'bar' per costruire il diagramma a barre.

```
males = dff.loc[:, 'Age':'M']
females = ['Age', 'F']
females = pd.DataFrame(dff, columns=females)

males.plot(kind = 'bar', xlabel = "Age bracket", ylabel = "Millions of people", color='blue')
females.plot(kind = 'bar', xlabel = "Age bracket", ylabel = "Millions of people", color= 'pink')
```

Coloriamo i grafici di blue e rosso rispettivamente in base al sesso:



4.1.1 Costruzione della population pyramid

Utilizziamo un dataset fornito da <https://www.populationpyramid.net/>: Andiamo ad importare il dataset all'interno del notebook:

```
In [93]: dff = pd.read_csv("italy-2019.csv")
dff.head(10)
```

```
Out[93]:
```

	Age	M	F
0	0-4	1227140	1155405
1	5-9	1400399	1320632
2	10-14	1476556	1392257
3	15-19	1484452	1387067
4	20-24	1533272	1422049
5	25-29	1630731	1549174
6	30-34	1711693	1678215
7	35-39	1861234	1849601
8	40-44	2157716	2166599
9	45-49	2432025	2468825

Costruiamo due liste, una per i maschi ed una per le femmine, raccogliendo i valori in base all'age bracket:

```
In [133]: males = dff.loc[:, 'Age':'M']

females = ['Age', 'F']
females = pd.DataFrame(dff, columns=females)
females['F'] = females['F'].apply(lambda x: x * -1) # Utilizziamo un
                                                    # lambda per invertire
                                                    # i valori di females

females = females['F'].values.tolist()
males = males['M'].values.tolist()
```

Creiamo un dataframe con headers: "Age", "Males", "Females", e utilizziamo le liste precedentemente create come input. utilizziamo quindi la libreria seaborn (basata su matplotlib) per creare la nostra population pyramid:

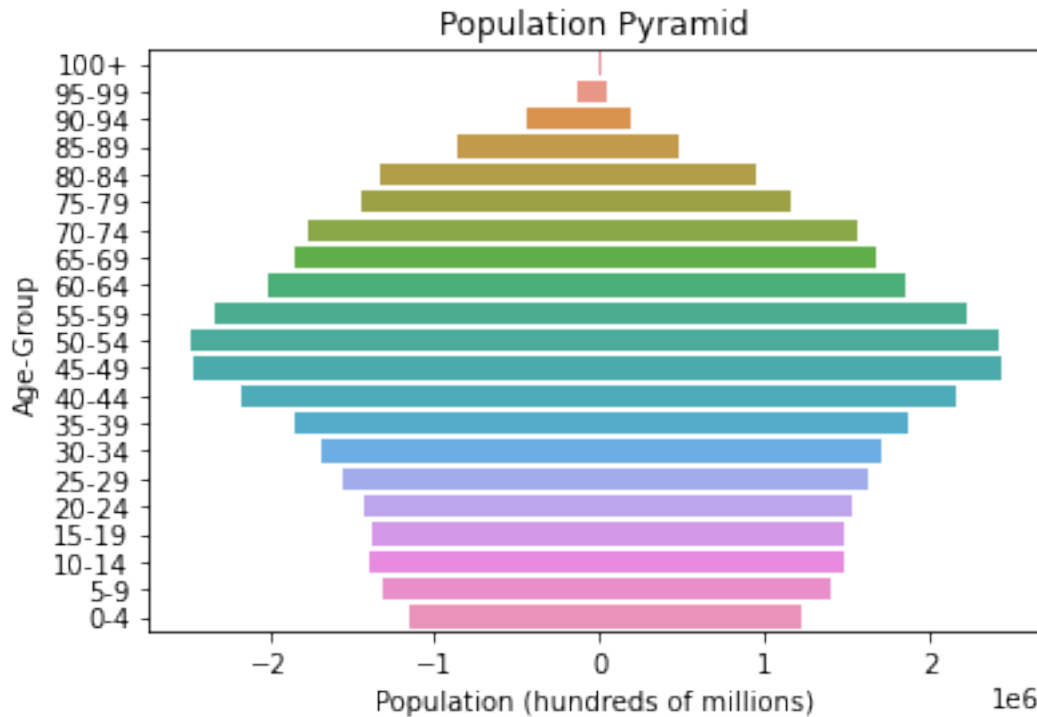
```
In [122]: import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns

df = pd.DataFrame({'Age': ['0-4', '5-9', '10-14', '15-19', '20-24', '25-29', '30-34', '35-39', '40-44',
                           '45-49', '50-54', '55-59', '60-64', '65-69', '70-74', '75-79', '80-84', '85-89',
                           '90-94', '95-99', '100+'],
                  'Males': males,
                  'Females': females})
AgeClass = ['100+', '95-99', '90-94', '85-89', '80-84', '75-79', '70-74', '65-69', '60-64', '55-59',
            '50-54', '45-49', '40-44', '35-39', '30-34', '25-29', '20-24', '15-19', '10-14', '5-9', '0-4']

bar_plot = sns.barplot(x='Male', y='Age', data=df, order=AgeClass, lw = 0)
bar_plot = sns.barplot(x='Female', y='Age', data=df, order=AgeClass, lw = 0)

bar_plot.set(xlabel="Population (hundreds of millions)", ylabel="Age-Group", title = "Population Pyramid")
```

Il risultato finale è il seguente:



4.2 Dependency ratio

Il dependency ratio è un indicatore che va a definire il rapporto tra la popolazione "dipendente" e quella "non dipendente". Per popolazione dipendente si intende la popolazione con età compresa tra 0 e 14, e da 65+. Questa suddivisione è fatta per misurare la pressione sulla popolazione produttiva, e quella non produttiva. Un dependency ratio più basso -di norma- si materializza in pensioni più alte, ed un sistema economico più salutare di uno con un dependency ratio più alto, che solitamente si traduce in un maggiore stress sulla popolazione lavoratrice.

Dependency ratio o indice di dipendenza è un indicatore demografico che va a rappresentare il rapporto tra le persone non più presenti nella manodopera di un paese, e quelli nella manodopera. E' quindi un indicatore che va a rappresentare la pressione sulla popolazione produttiva, higher is worse. Un basso tasso di dipendenza -solitamente- significa che ci sono meno persone da supportare, e quindi migliori pensioni, migliore sanità, più reinvestimenti nella società.

L'equazione della dependency ratio è così definita:

$$(\text{Total}) \text{ Dependency ratio} = \frac{(\text{number of people aged 0 to 14}) + (\text{number of people aged 65 and over})}{\text{number of people aged 15 to 64}} \times 100$$

Andiamo a costruire un grafico dell'andamento del dependency ratio dell'Italia dal 2000 al 2020: Andiamo a ciclare sui vari dataset, facendo un slicing per mettere in risalto i valori maschili e femmine di porzioni attive ed inattive. Utilizziamo poi la formula per calcolare il dependency ratio dell'anno e lo inseriamo nella lista

```
dependencies = []
years = []
for i in range(2000, 2021):
    years.append(i)
    location = "population/italy-" + str(i) + ".csv"
    dff = pd.read_csv(location)
    males = dff.loc[:, 'Age':'M']
    females = ['Age', 'F']
    females = pd.DataFrame(dff, columns=females)

    laborMales = males.iloc[0:3]
    laborMales = laborMales.append(males.iloc[13:20])
    inactiveMales = males.iloc[3:13]

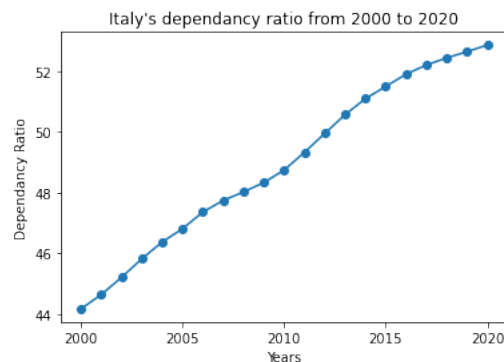
    laborFemales = females.iloc[0:3]
    laborFemales = laborFemales.append(females.iloc[13:20])
    inactiveFemales = females.iloc[3:13]

    dependency = ((laborMales['M'].sum() + laborFemales['F'].sum()) /
                  (inactiveMales['M'].sum() + inactiveFemales['F'].sum()))
    * 100
    dependencies.append(dependency)
```

Prepariamo il line chart mettendo dei pallini per mettere in risalto i valori per anno.

```
plt.plot(years, dependencies, marker = 'o')
plt.title("Italy's dependency ratio from 2000 to 2020")
plt.xticks(np.arange(min(years), max(years)+5, 5.0))
plt.xlabel('Years')
plt.ylabel('Dependency Ratio')
```

Ed ecco il grafico finale:



5 Cosa ci dice una "population pyramid"

5.1 Un'appendice sull'età

Qual'è il miglior indicatore per prevedere il comportamento di una "persona media" ed il suo contributo alla società?

Ci sono molte risposte a questa domanda, il background culturale, la professione, le condizioni sociali, la propensione al risparmio di una società. L'europeo medio ad esempio risparmia di più rispetto ad un americano medio, il finanziere che guadagna 100.000\$ l'anno investe di più (anche solo perché ha più disposable income) di un meccanico che raggiunge a malapena il fine mese, e così via.

C'è tuttavia un dato che ci permette di prevedere, in maniera pressoché accurata, il contributo di una persona al proprio paese: l'età.

5.1.1 Perché l'età è un indicatore così importante

Utilizziamo la nostra "population pyramid", e consideriamo la bracket di popolazione con età superiore a 65 anni circa. Questa categoria è rappresentata solitamente dai pensionati, che hanno terminato la propria carriera lavorativa, spendono poco, investono in maniera prudente e non contribuiscono più alle casse dello stato.

Il secondo gruppo che ci interessa è la bracket di popolazione con età compresa tra i 30 e i 65 anni. Questo bracket di popolazione rappresenta il "motore" di un paese, rappresenta il lavoratore medio che percepisce uno stipendio, spende, risparmia, e contribuisce alle casse dello stato. Durante questi anni la persona media raggiunge il massimo potenziale economico, facendo carriera, comprando case, macchine, e mettendo su famiglia. E' questo quindi il motivo per cui possiamo considerare questo bracket di popolazione come la risorsa più importante di un paese, più importante anche delle risorse naturali, e della posizione geografica.

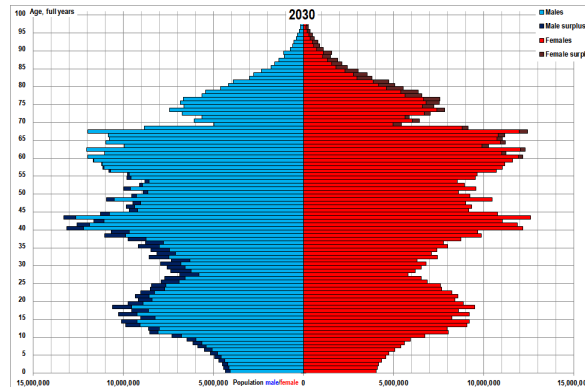
Il terzo gruppo che ci interessa è la bracket di popolazione con età compresa tra i 15 ed i 30 anni. Questo bracket di popolazione è principalmente composta da persone che stanno per terminare la propria carriera formativa, e che stanno per entrare nel mondo del lavoro. Durante questi anni la persona media non guadagna molto, vive solitamente con i genitori, e ha bisogno di prestiti per finanziare rette universitarie, o per comprare la prima casa.

L'ultimo gruppo che ci interessa è la bracket di popolazione con età inferiore ai 15 anni. Questo bracket di popolazione è la "più inutile" per un paese, formata da studenti e bambini che non contribuiscono ad una economicamente ad una società, ed invece costano molto alle famiglie sotto forma di mantenimento, pannolini etc.

Chiaramente la caratterizzazione che abbiamo non può essere applicato ad una qualsiasi società nel mondo e dipende strettamente dal livello di sviluppo di un paese, dalla geografia, dal livello di sviluppo della sanità, economia, e da una miriade di altri fattori.

6 Estensione delle population pyramids

- Population pyramid colorate maschio e femmina
L'abbiamo vista in precedenza, ed è la più semplice delle varianti, sono population pyramids in cui le barre sono colorate a seconda del sesso della popolazione.
- Population pyramid con surplus
Sono un'estensione della precedente, in cui oltre al colorare il sesso delle barre, coloriamo in maniera differente il surplus delle barre rispetto all'altro sesso, in questo modo possiamo visualizzare in maniera semplice la quantità di maschi, o femmine, in più in un paese.



- Animazioni con population pyramids
Andiamo a costruire un'animazione con il dataset che abbiamo. Siccome il nostro dataset va dal 2000 al 2020, estendiamo il programmino creato in precedenza aggiungendo un ciclo range, e salviamo con nome ogni grafico prodotto, dal 2000 al 2020.

```
images = []
for i in range(2000, 2021):
    location = "population/China-" + str(i) + ".csv"
    dff = pd.read_csv(location)
    males = dff.loc[:, 'Age':'M']

    output = "output" + str(i)
    bar_plot.figure.savefig("china/" + output)
```

```

png_dir = 'china/'
images = []
for file_name in sorted(os.listdir(png_dir)):
    if file_name.endswith('.png'):
        file_path = os.path.join(png_dir, file_name)
        images.append(imageio.imread(file_path))
imageio.mimsave('chinaMovie.gif', images)
print("Done")

```

Con le immagini generate utilizziamo un piccolo script basato su *imageio* per generare una .gif.

Questo tipo di animazione ci permette di vedere in maniera dinamica lungo gli anni l'andamento della popolazione, in particolare possiamo vedere come la popolazione -nei paesi in via di sviluppo-, si sposta tendenzialmente verso l'alto allontanandosi sempre più da una piramide, ed avvicinandosi ad una campana; mentre nei paesi il cui l'indice di sviluppo è basso il grafico rimane una piramide, indicatore di come lo sviluppo non stia tenendo passo.

- Predizioni sul futuro con population pyramids
Un qualcosa che ci stanno pregando di fare le animazioni precedenti è quello di estenderle, per darci una visione del futuro. In base all'andamento demografico di un paese è quindi possibile estendere il grafico lungo gli anni.