

# Population Pyramids

M: 941519

May 20, 2022

## Contents

<b>1</b>	<b>Cos'è una "population pyramid"</b>	<b>2</b>
1.1	Perché è chiamata una "piramide"? . . . . .	3
<b>2</b>	<b>I dati necessari per costruire una "population pyramid"</b>	<b>3</b>
<b>3</b>	<b>Esempio di costruzione utilizzando python</b>	<b>4</b>
<b>4</b>	<b>Cosa ci dice una "population pyramid"</b>	<b>6</b>
4.1	Un'appendice sull'età . . . . .	6
4.1.1	Perché l'età è un indicatore così importante . . . . .	6
<b>5</b>	<b>Estensione delle population pyramids</b>	<b>7</b>

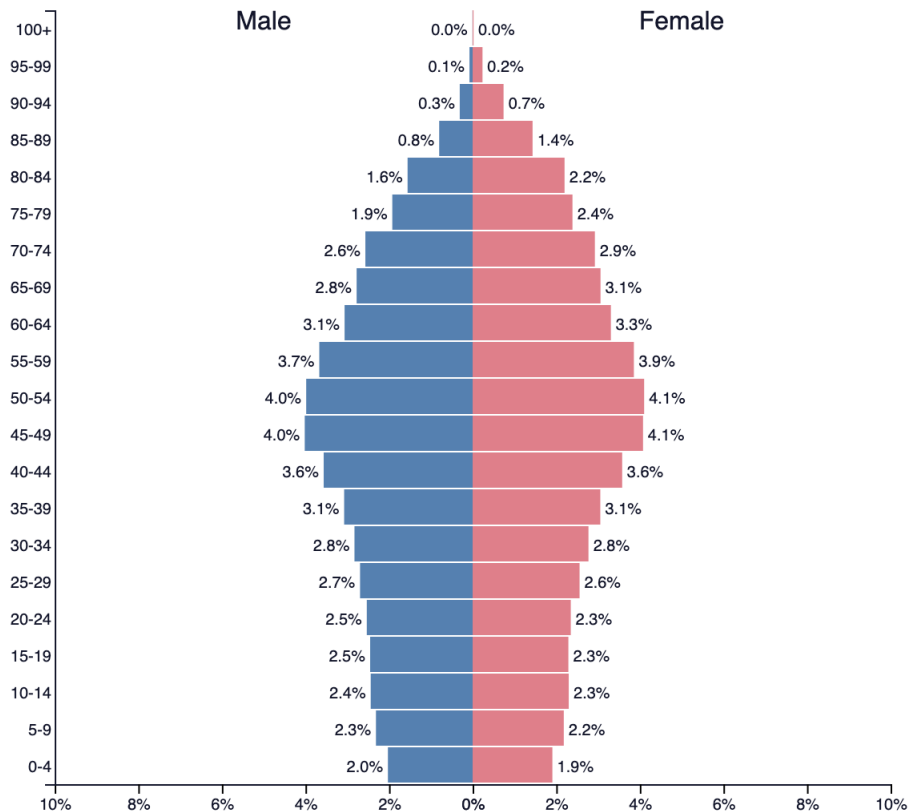
## 1 Cos'è una "population pyramid"

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), dividendoli in bracket di anni ed in spezzandoli in segmenti in base alla proporzione tra maschi e femmine.

La population pyramid è formata quindi da una serie di barre ad istogramma, la popolazione è rappresentata dalla grandezza dell'asse orizzontale, ed i bracket di anni sono rappresentati dall'asse  $y$ .

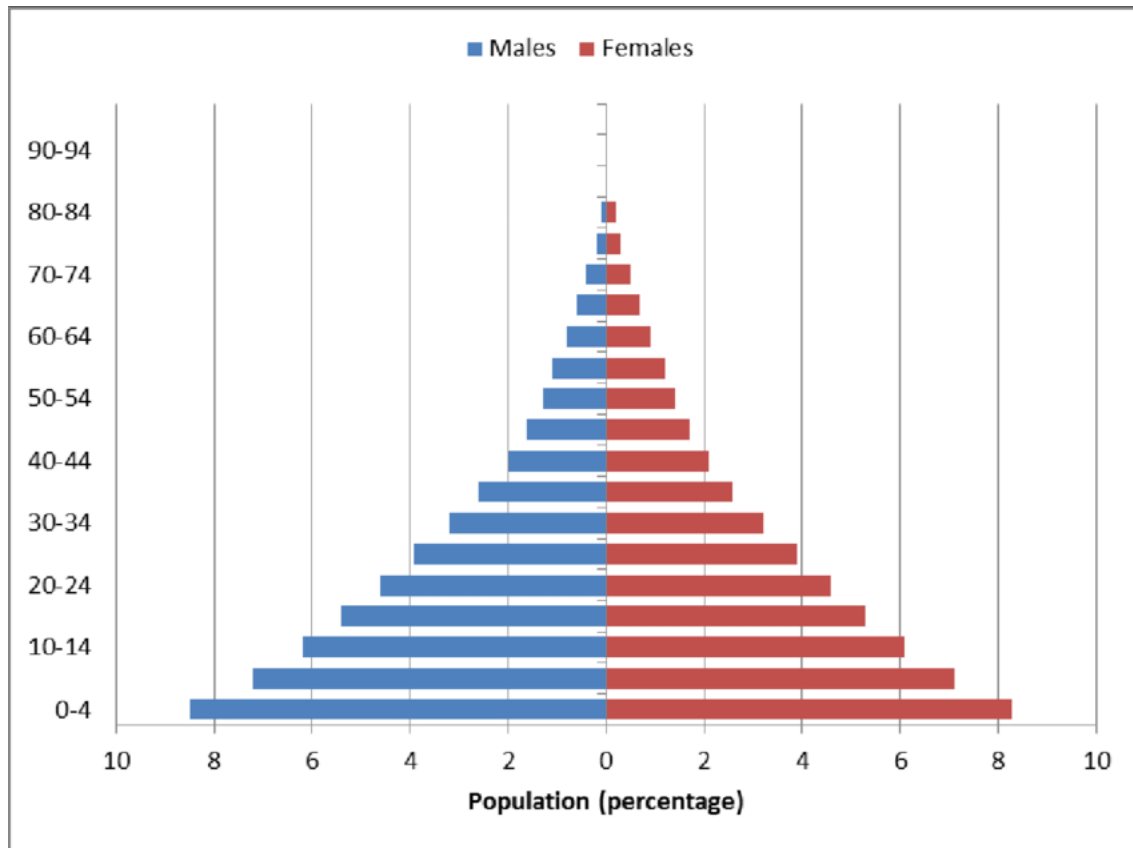
Italy ▼  
2019

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 della vita dell'uomo, ovvero quella della piramide:



Formata 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 composte da molti bambini, che in parte non raggiungono neanche i vent'anni, e la popolazione diminuisce progressivamente per la mancanza di buona medicina e prevenzione. Quella comunque, è la "population pyramid" dell'Africa Subsahariana.

## 2 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.

### 3 Esempio di costruzione utilizzando python

Utilizziamo un dataset fornito da <https://www.populationpyramid.net/>:

Age	M	F
0-4	1227140	1155405
5-9	1400399	1320632
10-14	1476556	1392257
15-19	1484452	1387067
20-24	1533272	1422049
25-29	1630731	1549174
30-34	1711693	1678215
35-39	1861234	1849601
40-44	2157716	2166599
45-49	2432025	2468825
50-54	2410929	2484894
55-59	2221334	2336011
60-64	1854790	2003108
65-69	1680438	1852790
70-74	1553702	1771151
75-79	1163633	1446562
80-84	942098	1331834
85-89	484087	868523
90-94	186445	448312
95-99	45405	143036
100	2956	13012

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:

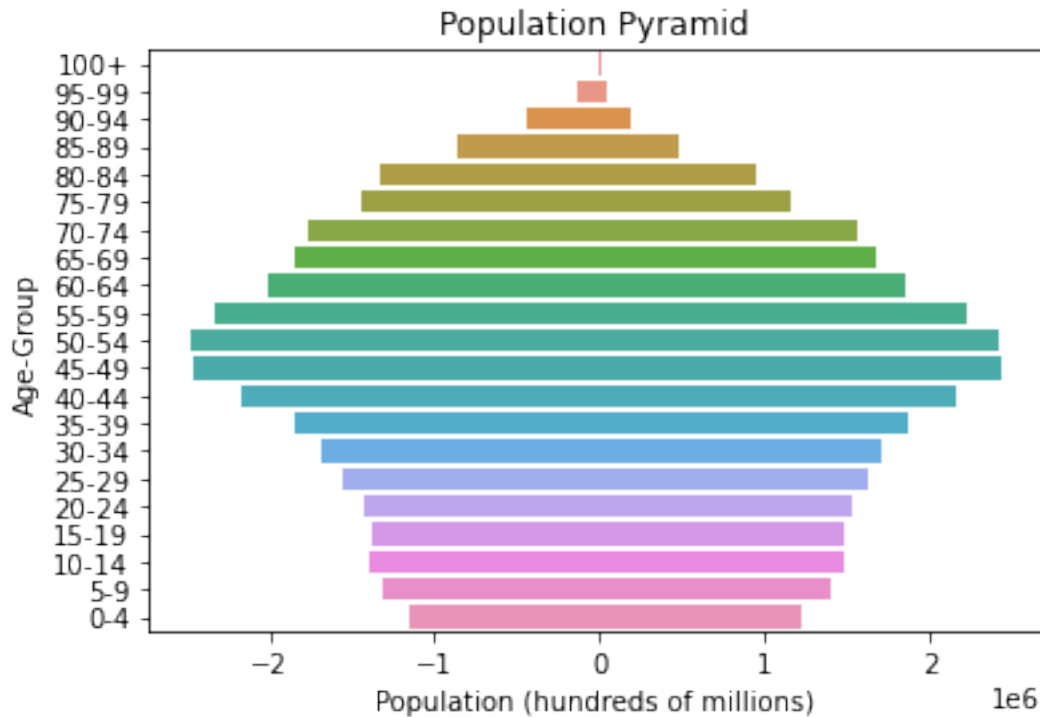
```
In [122]: import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns

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 Cosa ci dice una "population pyramid"

### 4.1 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, le norme e gli usi e costumi di un paese. 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.

#### 4.1.1 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.

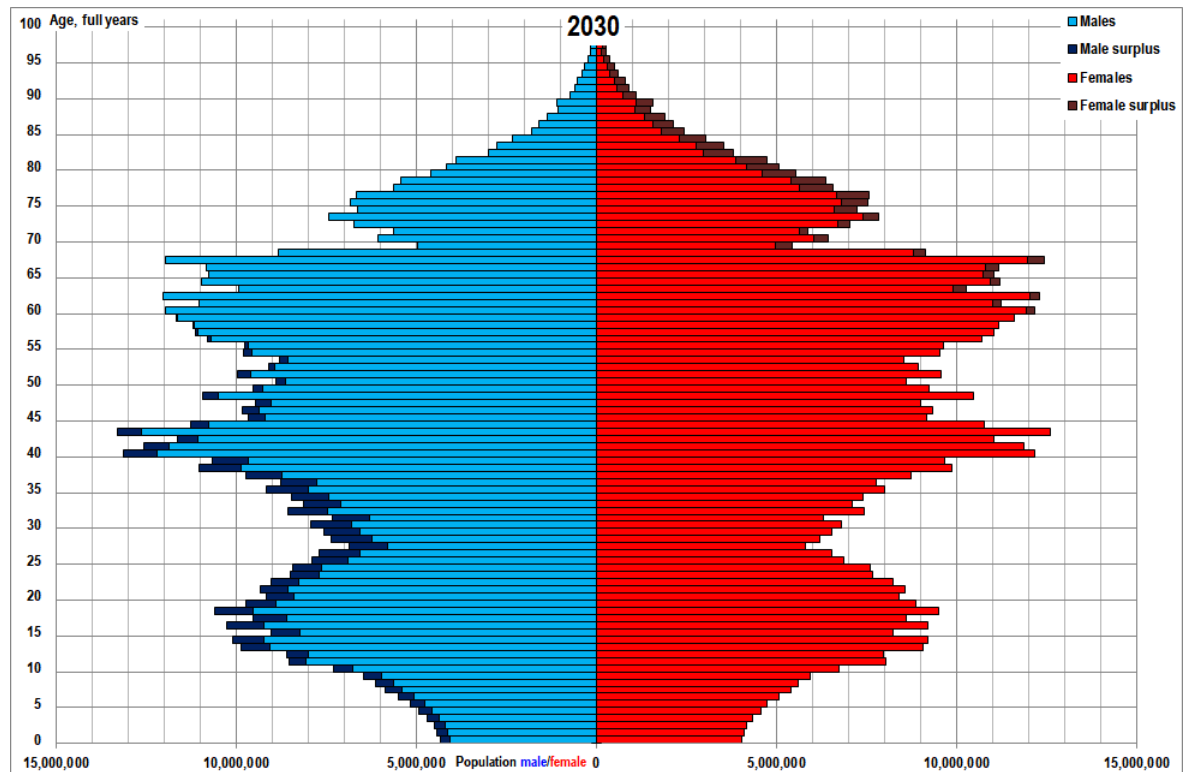
Il secondo gruppo che ci interessa è la bracket di popolazione con età compresa tra i 30 e i 65 anni. Questo bracket di popolazione rappresenta il "motore" di un paese, rappresenta il lavoratore medio che percepisce uno stipendio, spende, risparmia, e contribuisce alle casse dello stato. Durante questi 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.

## 5 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.