

Population Pyramids

M: 941519

May 16, 2022

Contents

1	Cos'è una "population pyramid"	2
2	I dati necessari per costruire una "population pyramid"	3
3	Esempio di costruzione utilizzando python	3
4	Cosa ci dice una "population pyramid"	5
4.1	Un'appendice sull'età	5
4.1.1	Perché l'età è un indicatore così importante	5

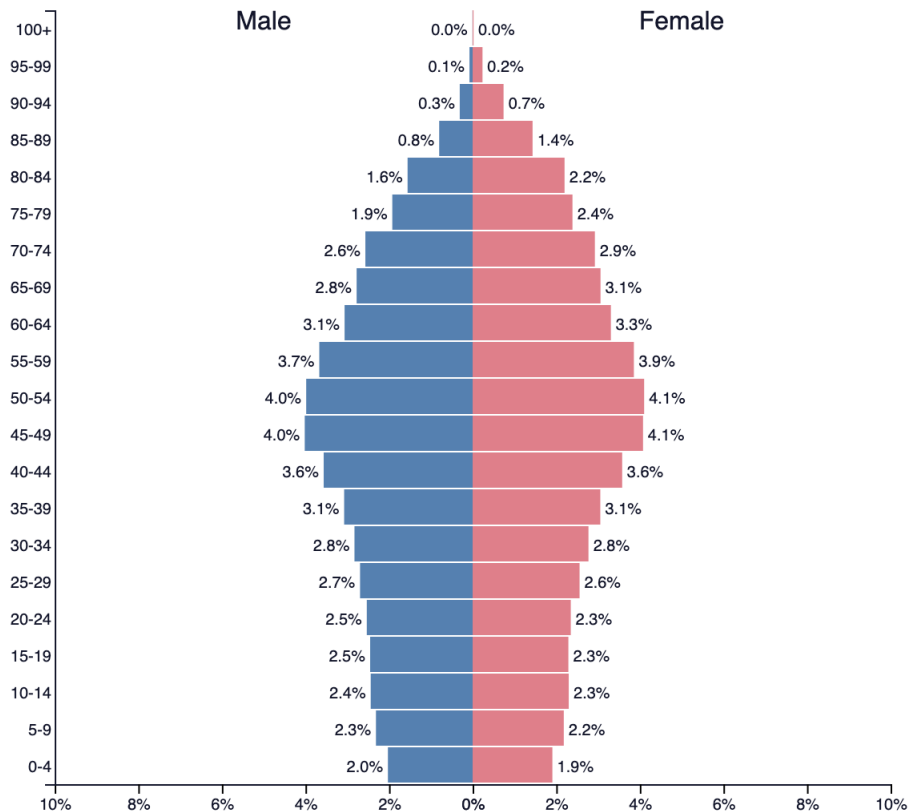
1 Cos'è una "population pyramid"

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



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 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:

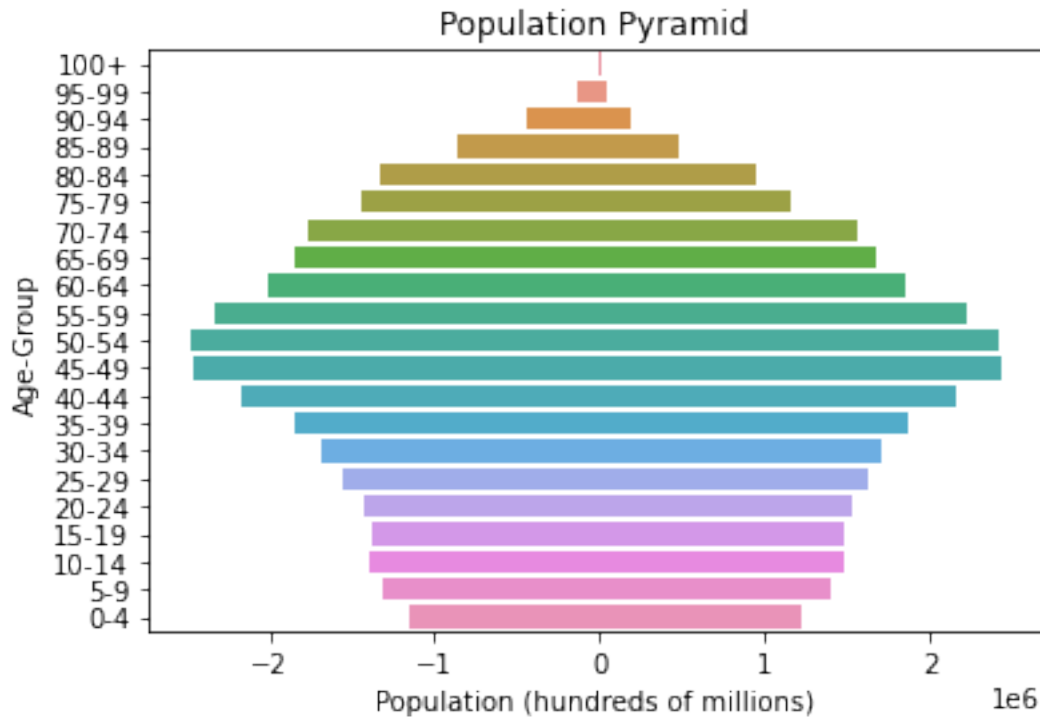
```
In [122]: import pandas as pd
import matplotlib.pyplot as plt
import numpy as np
import seaborn as sns

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

L'età