

Modello epidemiologico ad agenti basato su multi-layer network

1.0 _Raccolta dati

La prima parte del progetto volge sulla creazione di più strati sulla popolazione di Milano.

Per quanto riguarda i dati, il comune di Milano mette a disposizione i dati del censimento¹, insieme ad indicatori² che permettono di ricostruire i nuclei familiari di ogni sezione della città. Ad oggi (Novembre 2020), i dati più recenti fanno riferimento al censimento del 2011.

1.1 _Household layer

1.1.1 _Indicatori

I soli dati del censimento dividono la città in poco più di **6100 sezioni**. A partire da queste ultime, è possibile ricostruire le **ACE** (Aree di censimento), la tipologia di località (centro abitato, nucleo abitato, località produttiva e aree sparse) ed infine, il numero di abitanti per ogni sezione come dato aggiornato al 2010 (**POP₂₀₁₀**).

Il numero degli abitanti di ogni sezione in sé però non ci permette di fare stime sulla divisione e sul numero di nuclei familiari della sezione stessa: il comune di Milano però, sempre insieme ai dati del censimento, mette a disposizione **un set di 43 differenti indicatori per ogni sezione** che permettono di ricostruire quest'ultima a livello socio-demografico. Di questi 43 indicatori, 13 si rivelano particolarmente utili per ricostruire in modo approssimativo i nuclei familiari:

- Numero medio di residenti per edificio ad uso residenziale **(1)**
- Numero medio di abitazioni per edificio ad uso residenziale **(2)**
- Numero medio di abitazioni occupate da residenti per edificio ad uso residenziale **(3)**
- Percentuale di bambini in età prescolare **(8)**
- Percentuale di abitazioni non occupate **(13)**
- Percentuale di famiglie con un solo genitore **(14)**
- Percentuale di over 65 che vivono da soli **(17)**
- Giovani (0-14 anni) ogni 100 in età attiva (15-65 anni) **(25)**
- Over 65 ogni 100 in età attiva (15-64 anni) **(26)**
- Percentuale di popolazione di 15+ anni che percepisce reddito sul totale della popolazione 15+ **(27)**
- Percentuale di popolazione residente che si sposta giornalmente per studio o lavoro **(34)**
- Numero medio di componenti delle famiglie **(39)**
- Percentuale di famiglie di un componente **(40)**

¹ https://dati.comune.milano.it/it/dataset/ds98_infogeo_sezioni_censimento_localizzazione_2011c

² <https://dati.comune.milano.it/it/dataset/ds385-popolazione-indicatori-popolazione-famiglie-alloggi-edifici-sezione-censimento-2011>

A partire da questi indicatori, possiamo **ricavarne di nuovi** per migliorare la stima dei nuclei familiari.

Numero di edifici:

$$\frac{POP_{2010}}{\text{Numero medio di residenti per edificio}}$$

Numero di appartamenti:

$$\text{Numero di edifici} * \text{Numero medio di abitazioni occupate da residenti per edificio}$$

Numero di abitanti per appartamento:

$$POP_{2010} * \text{Numero di appartamenti}$$

Abitanti con meno di 14 anni:

$$\frac{\text{Under14 ogni 100 attivi}}{\left(\frac{100 + \text{Under14 ogni 100 attivi}}{100}\right)} * \left(\frac{POP_{2010}}{100}\right)$$

Abitanti con più di 64 anni:

$$\frac{\text{Over64 ogni 100 attivi}}{\left(\frac{100 + \text{Over64 ogni 100 attivi}}{100}\right)} * \left(\frac{POP_{2010}}{100}\right)$$

Abitanti nella fascia di età 15-64:

$$POP_{2010} - (\text{Abitanti con meno di 14 anni} + \text{Abitanti con più di 64 anni})$$

Con questi indicatori, possiamo ricostruire approssimativamente le fasce di età di ogni sezione, considerando gli abitanti con meno di 14 anni come scolari e quelli over 65 come pensionati, escludendo gli ultimi dal secondo strato, ovvero quello del **Workplace Layer** e inserendo i primi nel terzo, ovvero quello dello **School Layer**.

1.1.1 _Generazione dei nuclei familiari

Sapere il numero di abitanti over 65 ci permette **isolare i nuclei familiari composti da soli over 65 che vivono da soli**, calcolando la percentuale dell'indice 17 (Percentuale di over 65 che vivono da soli) sugli abitanti totali over 65, identificando una prima fetta dei nuclei familiari.

Tutto ciò che dobbiamo fare è applicare la percentuale degli over 65 che vivono da soli sul totale degli over 65 di ogni singola sezione e creare gli n-nuclei familiari che risultano dalla percentuale, **assegnando ad ogni nucleo (ovvero ad ogni persona) un appartamento**.

Una volta fatto ciò avremo i primi n nuclei familiari e i rimanenti (appartamenti-n) appartamenti. Procediamo ora a creare le famiglie composte da un solo genitore.

Sapere il numero degli abitanti under 14 ci permette, insieme all'indice 14 (Percentuale di famiglie con un solo genitore), di **avere una stima delle famiglie composte da un solo genitore**. L'indice 14 ci dice la percentuale di famiglie con un solo genitore e uno o più figli. Per decidere il numero medio di figli per famiglia si utilizza l'indice 39 (Numero medio di componenti per famiglia), in modo da far corrispondere il numero totale delle famiglie con il numero di appartamenti di ogni sezione.

Calcoliamo il **numero massimo di famiglie composte da un solo genitore** come:

$$\frac{\text{Giovani}}{100} * \text{Percentuale di famiglie con un solo genitore}$$

dove "Giovani" corrisponde al numero degli under 14, calcolato come numero massimo di famiglie con genitori e figli (nell'improbabile caso siano tutte famiglie con figli unici, ma ci permette di avere un upper limit).

Dividiamo il risultato per il numero medio di componenti per appartamento per calcolare il **numero effettivo di famiglie con un solo genitore**

$$\frac{\frac{\text{Giovani}}{100} * \text{Percentuale di famiglie con un solo genitore}}{\frac{POP_{2010}}{\text{Numero di appartamenti}}}$$

In quanto stiamo effettivamente dividendo il numero massimo di famiglie per la dimensione media delle famiglie nella sezione ispezionata.

Ottenuto questo numero che è una stima delle famiglie composte da un solo genitore, possiamo ora calcolare il **numero di figli per famiglia** come il numero totale di under 14 in famiglie con un solo genitore fratto il numero totale delle famiglie, distribuendo equamente quindi i figli in tutte le famiglie.

$$\frac{\frac{\frac{\text{Giovani}}{100} * \text{Percentuale di famiglie con un solo genitore}}{\frac{\text{Giovani}}{100} * \text{Percentuale di famiglie con un solo genitore}}}{\frac{POP_{2010}}{\text{Numero di appartamenti}}}$$

che corrisponde al numero medio di componenti per famiglia.

Fatto ciò, abbiamo ottenuto il primo nucleo di famiglie che comprende

- **Over 65 che vivono da soli**
- **Famiglie composte da un solo genitore**

Iniziamo generando le **famiglie con entrambi i genitori** (che hanno una dimensione minima di 3 elementi), per poi andare a riempire gli appartamenti vuoti con le persone ancora non assegnate a nuclei familiari fino ad aver assegnato tutti i POP₂₀₁₀ cittadini.