

Indagine quantitativa dell'impatto delle misure di
contenimento da Covid-19 sulla socialità presso la
Residenza universitaria Torrescalla tramite tecniche
di analisi dati su R

Carlo Ghiglione

21-02-2021

Indice

1	Introduzione	3
2	Dataset	4
3	Analisi frequenze di incontro	5
3.1	Matrice delle frequenze di incontro	5
3.2	Visualizzazione grafica	5
4	Score d'incontro	7
4.1	Definizione e calcolo dello score di incontro	7
4.2	Visualizzazione score di incontro	7
5	Correlazione tra score di incontro e turni	8
5.1	Correlazione	8
5.2	Visualizzazione risultati	8
6	Grafo degli incontri	9
6.1	Visualizzazione grafica	9
6.2	Interpretazione risultati	9
7	Conclusione	10

1 Introduzione

L'obiettivo di questo progetto è di indagare l'impatto delle misure di contenimento del contagio da Covid-19 durante l'Anno Accademico 2020-21 presso la Residenza Universitaria Torrescalla di Milano della Fondazione RUI, facente parte della Conferenza Collegi Universitari di Merito.

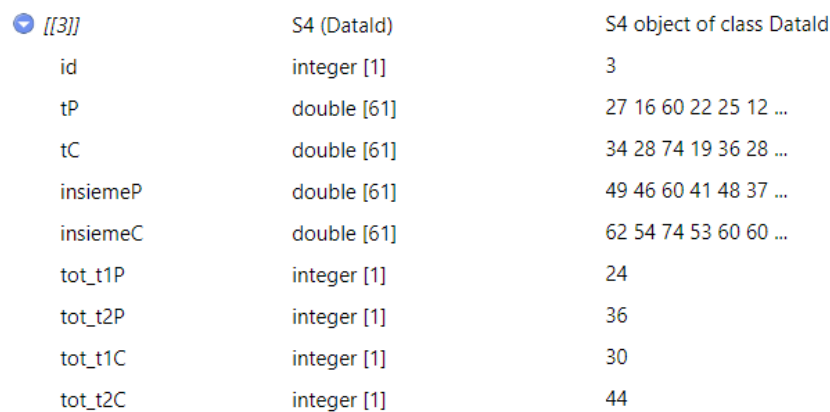
In particolare, si è indagato come l'introduzione del doppio turno ai pasti abbia modificato le abitudini dei residenti, ponendo particolare attenzione sul fatto se ciò possa limitare o meno le occasioni di incontro e socializzazione durante il pranzo e la cena. A tal fine, si è deciso di analizzare i dati del registro pranzi, piattaforma informatica dove i residenti prenotano e segnano la loro presenza ai pasti e, a partire da questi risultati, fare alcune considerazioni su come l'esistenza del doppio turno abbia o no impattato la socialità tra i ragazzi di Torrescalla.

2 Dataset

Il dataset su cui è stata svolta l'analisi si concentra sul periodo tra il 25-09-2020 e il 21-12-2020, tenendo conto solo dei residenti che hanno passato a Torrescalla più di 15 giorni, considerandone così un totale di 61.

Per ciascun residente, sono state estrapolate dal dataset le seguenti informazioni, rispettivamente a pranzo e a cena:

1. codice identificativo;
2. un array contenente il numero totale di volte in cui è stato presente lo stesso giorno e allo stesso turno insieme a ogni altro residente;
3. un array contenente il numero di volte in cui è stato presente lo stesso giorno insieme a ogni altro residente, senza tener conto se fossero o meno allo stesso turno;
4. numero totale di presenze rispettivamente al primo e secondo turno.



[[3]]	S4 (DataId)	S4 object of class DataId
id	integer [1]	3
tP	double [61]	27 16 60 22 25 12 ...
tC	double [61]	34 28 74 19 36 28 ...
insiemeP	double [61]	49 46 60 41 48 37 ...
insiemeC	double [61]	62 54 74 53 60 60 ...
tot_t1P	integer [1]	24
tot_t2P	integer [1]	36
tot_t1C	integer [1]	30
tot_t2C	integer [1]	44

Figura 1: Struttura dati con le informazioni estrapolate per ogni residente.

3 Analisi frequenze di incontro

3.1 Matrice delle frequenze di incontro

A partire dalle informazioni estrapolate dal dataset, è stata ricavata una matrice delle frequenze di incontro di ogni residente con tutti gli altri rispettivamente per il pranzo e per la cena.

La frequenza di incontro f_{ij} del residente i con il residente j è calcolata al seguente modo:

$$f_{ij} = \frac{t_{ij}}{v_{ij}} \quad (1)$$

- t_{ij} numero di volte che il residente i è stato presente lo stesso giorno e allo stesso turno col residente j ;
- v_{ij} numero di volte che il residente i è stato presente lo stesso giorno (ma non necessariamente allo stesso turno) col residente j ;

Le frequenze così calcolate tengono conto esclusivamente degli incontri avvenuti tra residenti presenti lo stesso giorno, evitando di considerare come mancati incontri le giornate in cui i ragazzi non erano entrambi presenti in collegio. In tal modo, si evita che i risultati siano distorti dalla inevitabile differenza di presenze totali in residenza che sussiste tra le persone, concentrando l'attenzione esclusivamente sulle abitudini dei ragazzi durante la loro permanenza a Torrescalla.

3.2 Visualizzazione grafica

Le matrici delle frequenze così prodotte, rispettivamente per il pranzo e per la cena, possono essere visualizzate con una heatmap alla quale è stato applicato un algoritmo di riordinamento automatico per mettere in rilievo eventuali cluster.

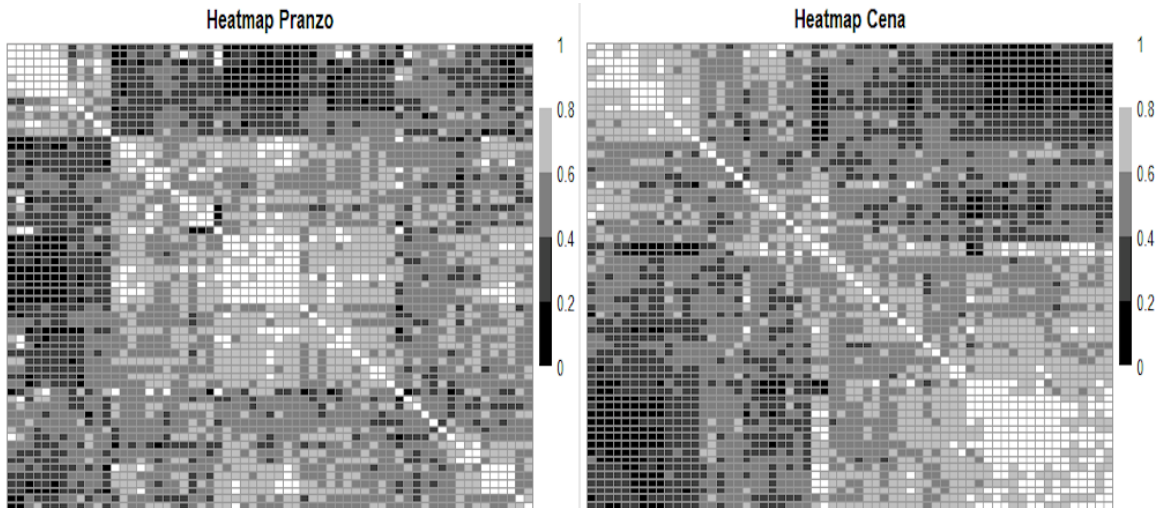


Figura 2: Heatmap delle frequenze di incontro a pranzo (sx) e a cena (dx).

Da una prima osservazione dei grafici, si vede come sia a pranzo che a cena esistano gruppi di residenti che tendono a incontrarsi con una elevata frequenza tra di loro (in bianco, tra 0.8 e 1.0) e poco con altri (in nero, tra 0.0 e 0.2). In particolare si vede che:

1. a pranzo la distribuzione delle frequenze è più omogenea ed c'è un solo cluster isolato di dimensione ridotte (circa una dozzina di persone);
2. a cena le frequenze sono distribuite in due macro aree, ciascuna delle quali ha una zona con valori di frequenze molto elevate formata da circa 15 persone e una zona con valori intermedi (in grigio, tra 0.4 e 0.8) altrettanto estesa, determinando una evidente configurazione "a farfalla" del grafico.

Da questa analisi, sembra emergere che sia piuttosto forte l'abitudine dei residenti a cenare sempre allo stesso orario, determinando una situazione in cui si formano due cluster di circa 15 residenti che hanno una frequenza di incontro molto elevata tra di loro ed estremamente bassa con quelli dell'altro gruppo. Per la restante parte dei residenti invece, la frequenza di incontro sembra essere abbastanza equilibrata con tutti, benché una predilezione per uno dei due gruppi sia evidente.

4 Score d'incontro

4.1 Definizione e calcolo dello score di incontro

Per misurare quantitativamente la tendenza generale di un residente a vedersi frequentemente con gli altri ragazzi, a ciascuno è stato assegnato uno score di incontro s_i calcolato al seguente modo:

$$s_i = 10 \cdot \frac{n_i}{n_{tot}} \quad (2)$$

- n_i numero di persone con cui il residente i ha una frequenza di incontro maggiore di 0.25;
- n_{tot} numero totale di residenti (61).

In tal modo, se un residente ha un punteggio alto, vuol dire che riesce a incontrarsi frequentemente con un elevato numero di persone. Viceversa, un residente con un punteggio basso tenderà a incontrarsi poco con gli altri.

4.2 Visualizzazione score di incontro

Gli score di incontro sono stati calcolati per ogni residenti e visualizzati con un istogramma, rispettivamente per il pranzo e la cena. Si osserva che ad entrambi i pasti lo score medio sia circa 9 e la sua distribuzione è molto simile tra i due. I valori mediamente elevati dello score mostrano che i residenti tendono a incontrarsi abbastanza frequentemente nel complesso, nonostante la divisione in cluster che emerge dalla precedente analisi.

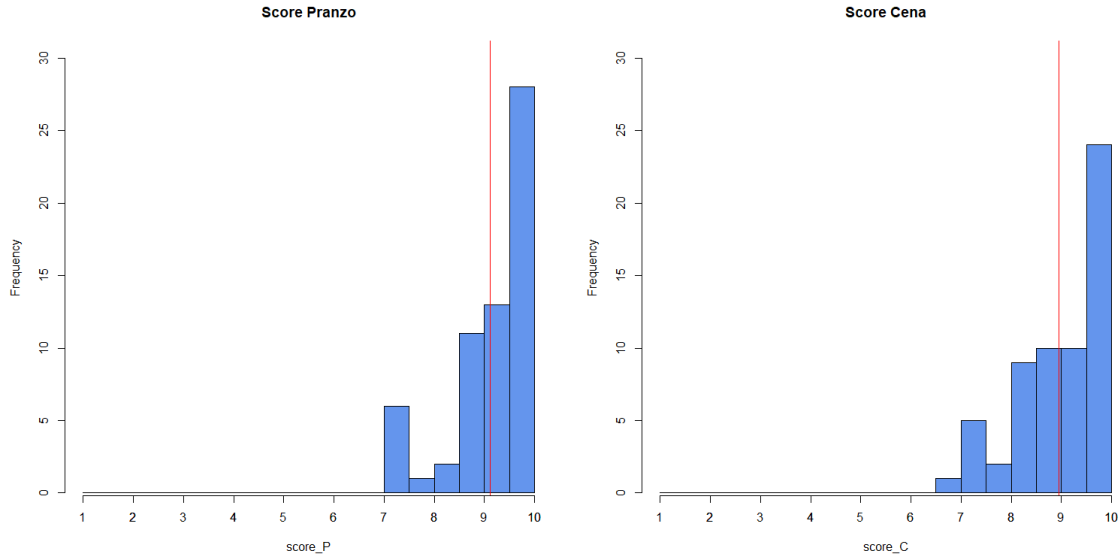


Figura 3: Istogramma degli score di incontro a pranzo (sx) e cena (dx).

5 Correlazione tra score di incontro e turni

5.1 Correlazione

In questa sezione si vuole studiare se esiste una qualche correlazione tra lo score d'incontro e la frequenza con cui i residenti si prenotano al primo e al secondo turno. Ci si aspetta che i residenti che frequentano in modo equilibrato i due turni tendano a incontrare gli altri residenti con maggiore frequenza di quelli che, invece, hanno l'abitudine a segnarsi sempre allo stesso.

E' stato quindi studiato il seguente modello di regressione:

$$y = ax^2 + bx + c + \varepsilon \quad \varepsilon \sim \mathcal{N}(0, \sigma^2) \quad (3)$$

- y score di incontro;
- x percentuale di frequentazione del primo turno ($1 - x$ per il secondo turno).

5.2 Visualizzazione risultati

Eseguendo il modello di regressione sia per il pranzo che la cena, si osserva una correlazione evidente (le ipotesi di normalità dei residui risultano sufficientemente soddisfatte). Il modello per il pranzo ha infatti un indice R^2 di circa 0.86, mentre per la cena è di 0.83 e, in entrambi i casi, i tre coefficienti sono significativamente non nulli.

La supposizione precedente è corretta: i residenti con alto score di incontro sono proprio quelli che, sia a pranzo che a cena, frequentano i due turni in modo equilibrato. Viceversa, i residenti che hanno punteggi più bassi sono quelli che hanno abitudini fortemente squilibrate verso uno dei due turni.

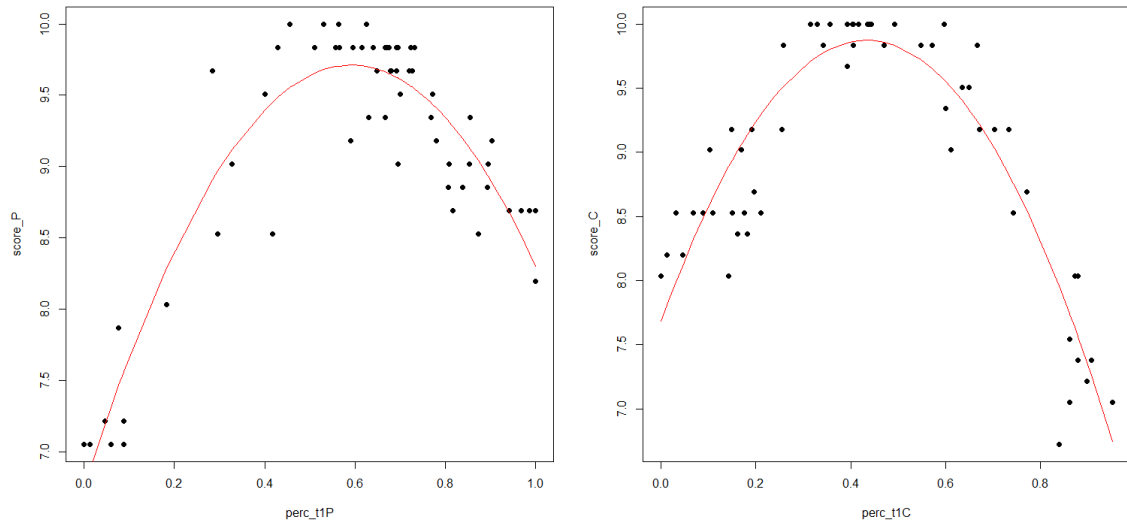


Figura 4: Score incontri vs frequenza al primo turno al pranzo (sx) e alla cena (dx).

6 Grafo degli incontri

6.1 Visualizzazione grafica

Per visualizzare in modo unitario i tre fattori di frequenza reciproca di incontro, frequentazione dei turni e score di incontro, è stata realizzata un grafo, rispettivamente per il pranzo e la cena, nel seguente modo:

1. la dimensione dei vertici è proporzionale allo score di incontro di ogni residente;
2. gli archi sono posti solo tra residenti che hanno una frequenza di incontro molto elevata (superiore a 0.80);
3. il colore dei vertici varia da rosso (forte frequentazione del primo turno) a blu (forte frequentazione del secondo turno).

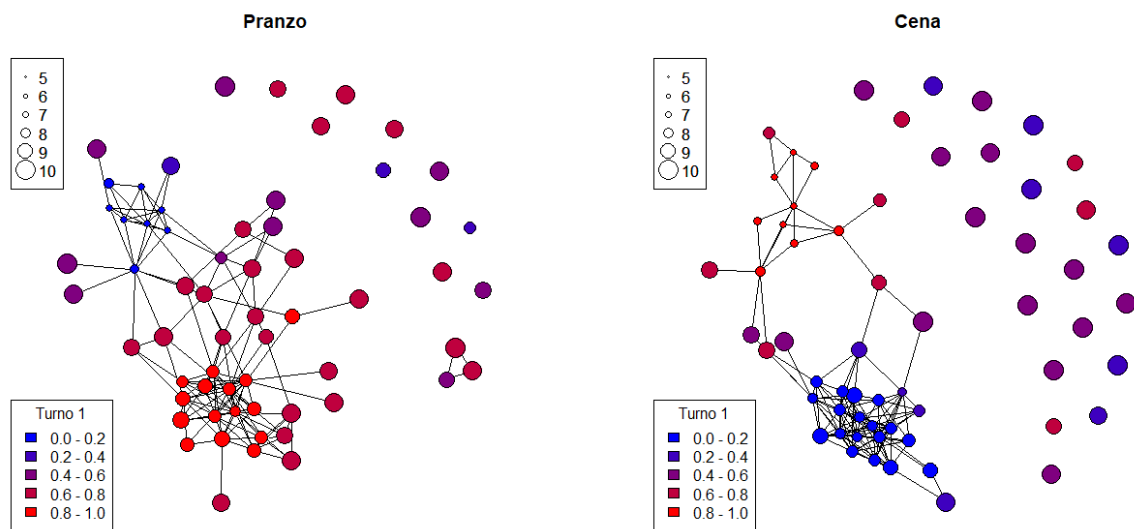


Figura 5: Grafo degli incontri per il pranzo (sx) e per la cena (dx).

6.2 Interpretazione risultati

I grafi mettono alla luce, sia a pranzo che a cena, i seguenti fatti:

1. è identificabile un gruppo di residenti che tendono fortemente a incontrarsi tra di loro, come era stato evidenziato dall'analisi della heatmap;
2. si vede bene che le persone nel cluster, soprattutto nel caso della cena, abbiano in media score di incontro più bassi degli altri;
3. le persone con bassi punteggi nel cluster sono quelle con frequentazione dei turni molto sbilanciata, come era stato evidenziato dall'analisi della correlazione;
4. il cluster in realtà è suddivisibile in due sotto gruppi con una scarsa comunicazione reciproca; tali gruppi coincidono esattamente con le persone che frequentano fortemente lo stesso turno.

7 Conclusione

Alla luce delle precedenti analisi, si può affermare che l'introduzione del doppio turno ai pasti abbia determinato effettivamente una situazione in cui i residenti tendono a incontrare più frequentemente quelli con le loro stesse abitudini e più raramente gli altri. L'analisi dei grafi e delle heatmap hanno messo in evidenza la presenza di gruppi di residenti che faticano a incontrarsi tra loro con una buona frequenza ma, essendo i punteggi degli score di incontro comunque mediamente alti, tale fenomeno non raggiunge dimensioni preoccupanti e generalizzate.

Per venire incontro a tale situazione, suggeriamo di incoraggiare quanto più possibile i residenti a segnarsi in maniera equilibrata tra i due turni per far sì che la loro tendenza generale a incontrarsi migliori, come l'analisi della correlazione ha mostrato. Sicuramente, in tal modo, la socialità della Residenza Torrescalla ne trarrebbe beneficio.