

## Secondo progetto di Social Computing 2020-2021

Francesco Bombassei De Bona (144665)

Andrea Cantarutti (141808)

Alessandro Zanatta (143154)

17 gennaio 2021

# 1 Introduzione

Il seguente elaborato espone il processo di analisi dei dati ottenuti in seguito al dispiegamento di un task di Crowdsourcing, ottenuti attraverso le metodologie offerte dal framework *Crowd Frame*. In particolare, il lavoro è stato suddiviso nelle seguenti fasi:

- [Ottenimento dei dati serializzati su un bucket Amazon S3](#)
- [Analisi dei dati e produzioni di grafici esplicativi](#)

# 2 Recupero dei dati

I dati relativi allo svolgimento del task da parte di ogni singolo worker sono stati serializzati direttamente da *Crowdframe* in diversi documenti JSON all'interno del bucket Amazon S3 predisposto. I dati raccolti sono stati scaricati tramite gli script descritti nell'elaborato precedente.

Il task è stato svolto da un totale di **54 worker**, dei quali 6 appartenenti rispettivamente a due gruppi coinvolti nello svolgimento del medesimo task e 48 reclutati al fine di incrementare la numerosità dei dati acquisiti.

Per ogni worker, è stata scaricata una directory contenente i dati di completamente del task (denominata in base all'id del worker). L'insieme dei dati raccolti è situato all'interno della directory **Data**.

### 3 Analisi dei dati raccolti

L'analisi dei dati è contenuta all'interno della directory `pyAnalysis` ed è stata svolta all'interno di un apposito Jupyter Notebook.

#### 3.1 Osservazione dei dati ottenuti dal questionario

Al fine di rendere possibile una visualizzazione d'insieme in merito alla compilazione del questionario introduttivo, è stato prodotto il seguente diagramma “sunburst”.



Figura 1: Sunburst plot

Inoltre, sono stati prodotti istogrammi e diagrammi a barre relativi alla frequenza relativa di ogni variabile campionata per mezzo del questionario.

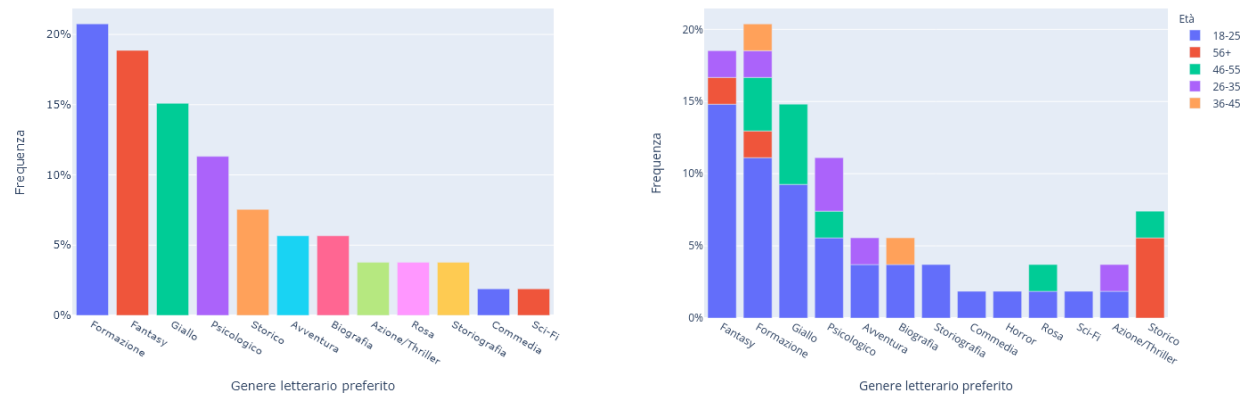


Figura 2: Barplot relativo al genere letterario preferito, anche in relazione all'età

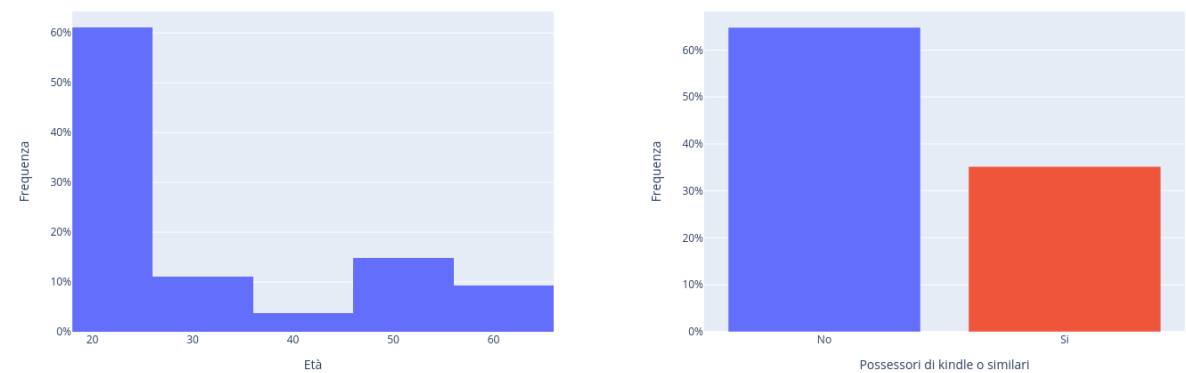


Figura 3: Istogramma in relazione all'età e barplot relativo ai possessori di *eBook Reader*

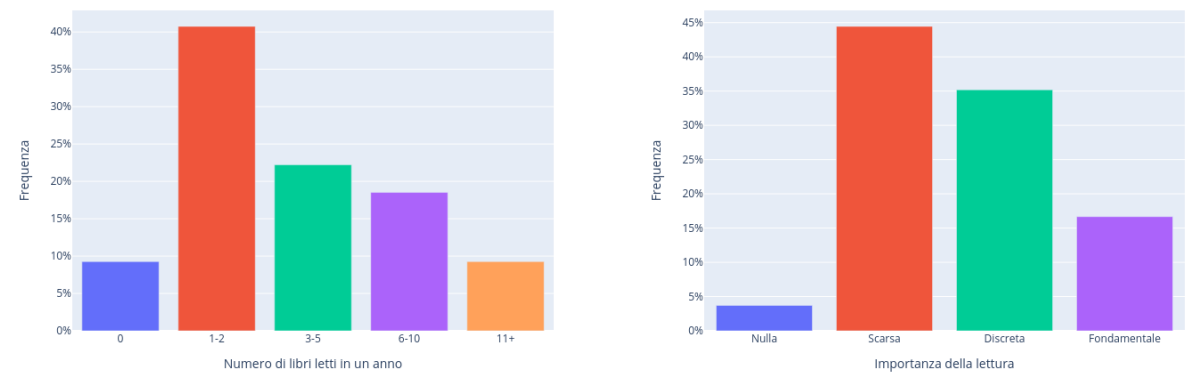


Figura 4: Barplot relativi alle preferenze in relazione alla lettura

### 3.2 Osservazione dei dati relativi alle dimensioni proposte

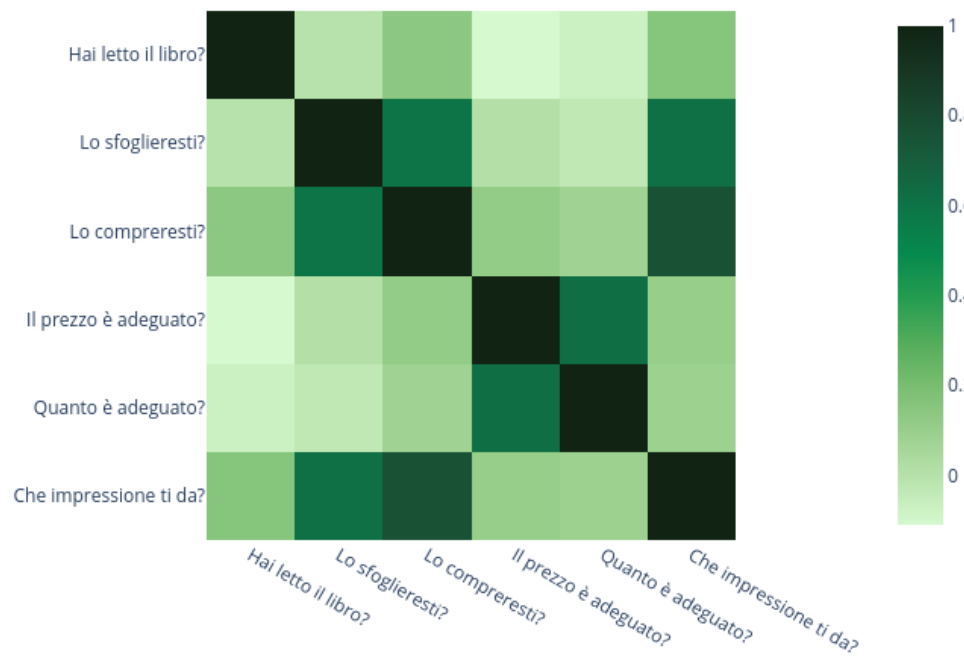
#### 3.2.1 Grado medio di adeguatezza del prezzo

Per ogni edizione presa in analisi, è stato calcolato il grado medio di adeguatezza del prezzo sulla base delle risposte fornite dai worker. In particolare, ogni worker poteva fornire una risposta tramite l'apposito slider con valori inclusi fra 0 e 5 (rispettivamente riferiti alla **totale adeguatezza** e **ineguatezza** del prezzo).

Titolo	Il prezzo è adeguato?	Quanto è adeguato?
Le cronache di Narnia (italiano, cartaceo)	0.944444	3.944444
Le cronache di Narnia (italiano, eBook)	0.777778	3.777778
The Chronicles of Narnia (inglese, cartaceo)	0.833333	3.222222
Assassinio sull'Orient Express (italiano, cartaceo)	0.833333	3.666667
Assassinio sull'Orient Express (italiano, eBook)	0.722222	3.444444
Assassinio sull'Orient Express(inglese, cartaceo)	0.833333	3.888889
Così parlò Zarathustra (italiano, cartaceo)	0.833333	3.666667
Così parlò Zarathustra (italiano, eBook)	0.722222	3.666667
Così parlò Zarathustra (inglese, cartaceo)	0.777778	3.611111

Altra tabella

Metrica	Il prezzo è adeguato?	Quanto è adeguato?
Le cronache di Narnia (italiano, cartaceo)	1.0	4.0
Le cronache di Narnia (italiano, eBook)	1.0	4.0
Le cronache di Narnia (inglese, cartaceo)	1.0	3.0
Assassinio sull'Orient Express (italiano, cartaceo)	1.0	4.0
Assassinio sull'Orient Express (italiano, eBook)	1.0	4.0
Assassinio sull'Orient Express(inglese, cartaceo)	1.0	4.0
Così parlò Zarathustra (italiano, cartaceo)	1.0	3.5
Così parlò Zarathustra (italiano, eBook)	1.0	4.0
Così parlò Zarathustra (inglese, cartaceo)	1.0	3.5



## 4 Conclusioni