



Università Ca'Foscari Venezia

Analisi degli studi e dell'occupazione in Italia ed Europa

Filippo Bisconcini, 852144

28 Maggio 2018

Indice

1	Premessa	2
2	Italia	3
2.1	Disoccupazione post-laurea	3
2.2	Motivazioni di trasferimento	7
3	Europa	10
4	Conclusioni e considerazioni personali	13

Capitolo 1

Premessa

Durante la stesura del progetto è nata la volontà nel sottoscritto di concentrarsi su un unico argomento, da trattare mediante dataset multipli rispetto all'analisi di tre o più situazioni scollegate tra di loro. Tra le diverse proposte sono stati scelti i tre seguenti dataset, di cui riportiamo la data di costruzione del campione:

- La qualità della vita delle province italiane: 2003/2004
- Laureati e lavoro: 2015
- Regioni Europee: 2016

Essendo di nostro interesse confrontare il grado di disoccupazione in Italia con la disoccupazione post-laurea saremmo costretti a confrontare due dataset costruiti a dodici anni di distanza. Ciò non è accettabile, quindi provvediamo ad includere un quarto dataset messo a disposizione dall'ISTAT dal quale possiamo estrarre i dati relativi al 2015.

- Tasso di disoccupazione: 2015

Utilizzeremo il primo dataset (2003) solo per confronti temporali con i dati aggiornati (2015).

Capitolo 2

Italia

Ci interessa analizzare la situazione del nostro Paese in fatto di occupazione e, in particolare, della relazione con le modalità di studio di un campione di laureati.

2.1 Disoccupazione post-laurea

Cominciamo con l'analizzare la disoccupazione dei laureati ad un anno dalla conclusione del loro percorso di studi.

La percentuale nazionale di disoccupati dopo un anno dalla laurea è del 36%.

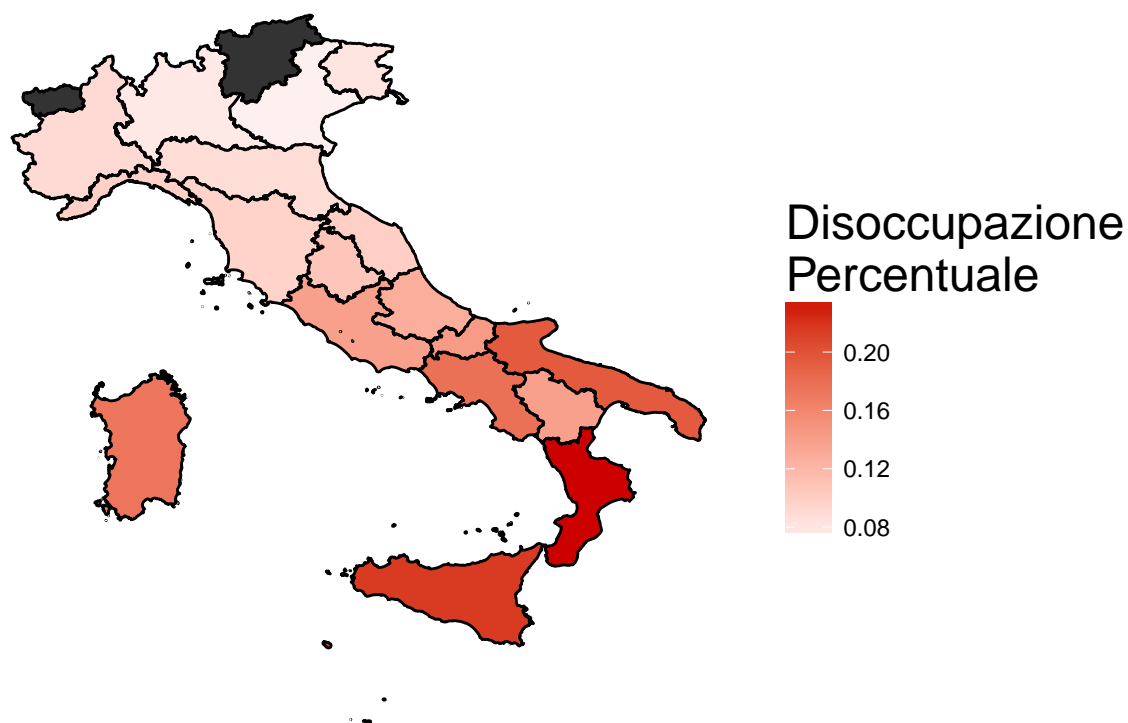


Figura 2.1.2: Percentuale di disoccupazione regionale nel 2015.

Fatta eccezione per le regioni Valle d'Aosta e Trentino Alto Adige di cui non abbiamo informazione, lo scenario appare graficamente molto simile allo studio precedente, con un Sud Italia più disoccupato con picco in Calabria.

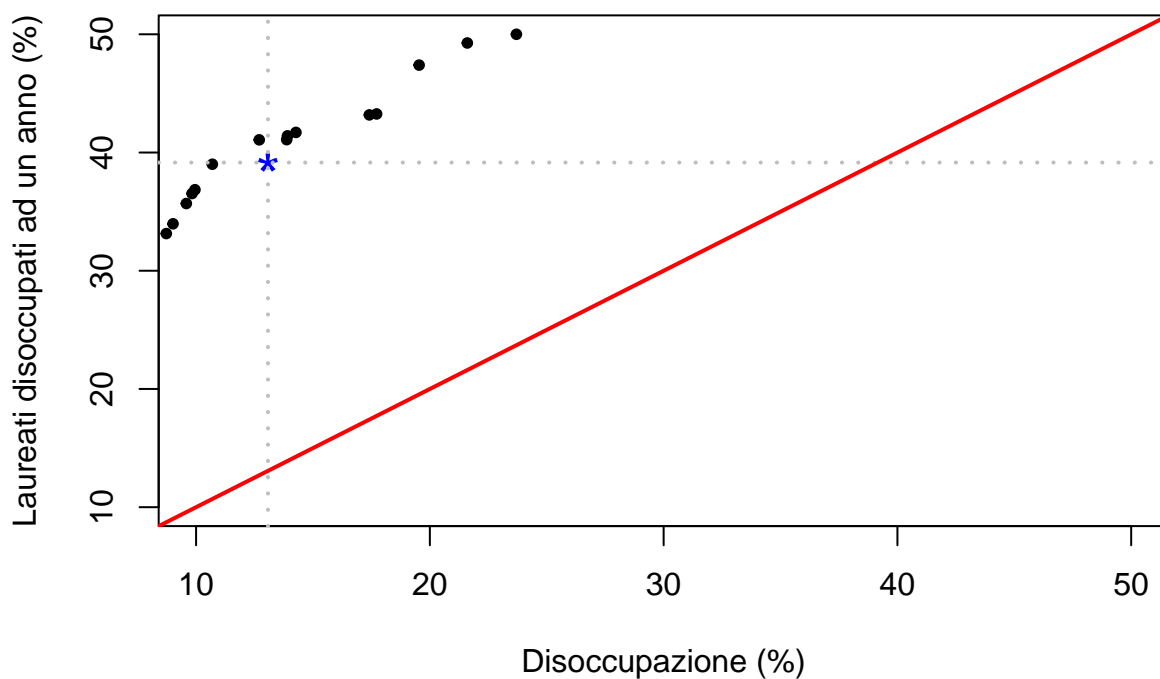


Figura 2.1.3: Disoccupazione in Italia confrontata con neolaureati.

Nella figura 2.1.3 notiamo la simmetria tra la disoccupazione e la percentuale di neolaureati disoccupati, in quanto le due distribuzioni si dispongono tendenzialmente su due rette parallele. Il fatto che ci sia un *offset* tra le due è motivato dal fatto che stiamo analizzando due variabili di diversa natura, su campioni diversi costruiti però nel medesimo luogo.

Avendo a disposizione due dataset costruiti a dodici anni di distanza, possiamo calcolarci la variazione nella percentuale di disoccupazione. Rappresentiamo le due variabili su un piano cartesiano per mostrare la distribuzione dell'indice appena calcolato.

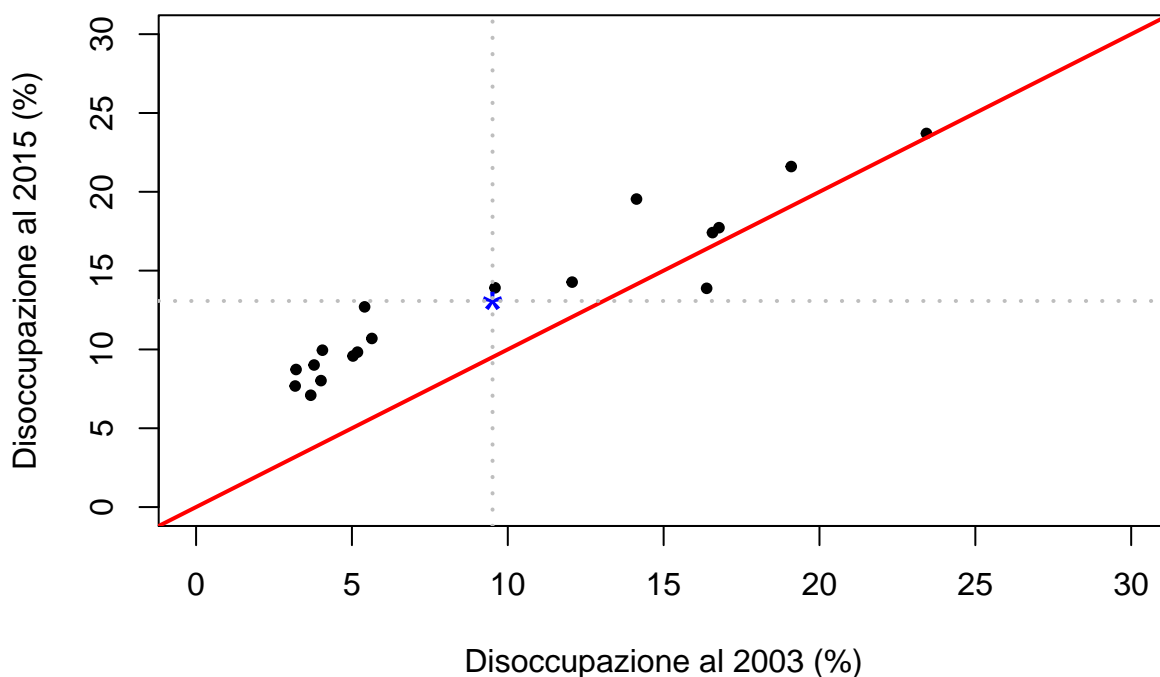


Figura 2.1.4: Disoccupazione in italia 2003/2015.

La figura 2.1.4 mostra la correlazione tra la disoccupazione nelle due annate 2003 e 2015. Nelle regioni che compaiono al di sopra della diagonale, la disoccupazione nel 2015 risulta diminuita rispetto alla rilevazione effettuata nel 2003.

Proponiamo una rappresentazione geografica di quanto appena analizzato.

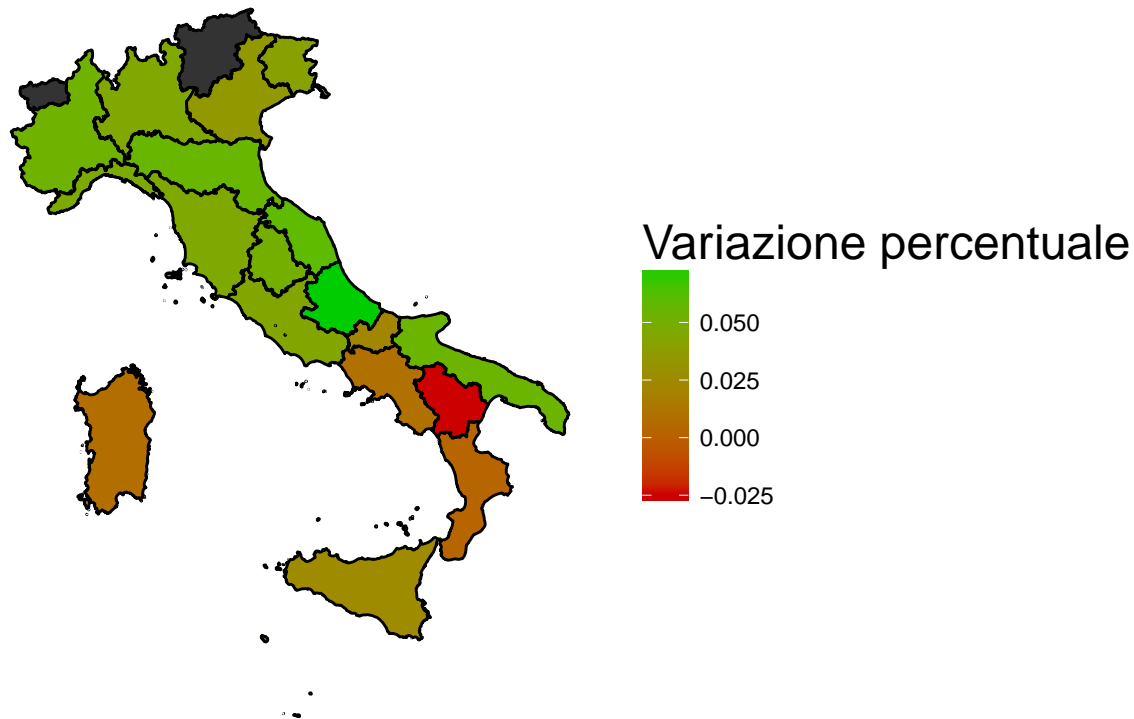


Figura 2.1.5: Variazione della disoccupazione percentuale dal 2003 al 2015.

Risulta evidente un aumento del 2.5% della disoccupazione in Basilicata negli ultimi dodici anni. Ad esclusione di quest'ultima, tutte le altre regioni hanno subito una diminuzione media del 4% di disoccupazione.

Risulta curiosa la marcata differenza tra Puglia e Basilicata, praticamente agli estremi opposti della scala nonostante la vicinanza geografica che, come notato dai grafici precedenti, sembra comportare una similarità nei valori raccolti.

2.2 Motivazioni di trasferimento

Dividiamo il campione di laureati, raggruppato per ambito, in due macro settori (*Scientifico* ed *Umanistico*) così suddivisi:

- Scientifico: Scientifico, Chimico-farmaceutico, Geo-biologico, Medico, Ingegneria, Architettura, Agrario, Economico-statistico, Educazione fisica, Difesa e sicurezza
- Umanistico: Politico-sociale, Giuridico, Letterario, Linguistico, Insegnamento, Psicologico

Procediamo ad analizzare le componenti principali delle proprietà inerenti al punteggio che hanno dato i laureati alle diverse ragioni per cui migrerebbero in un altro Stato.

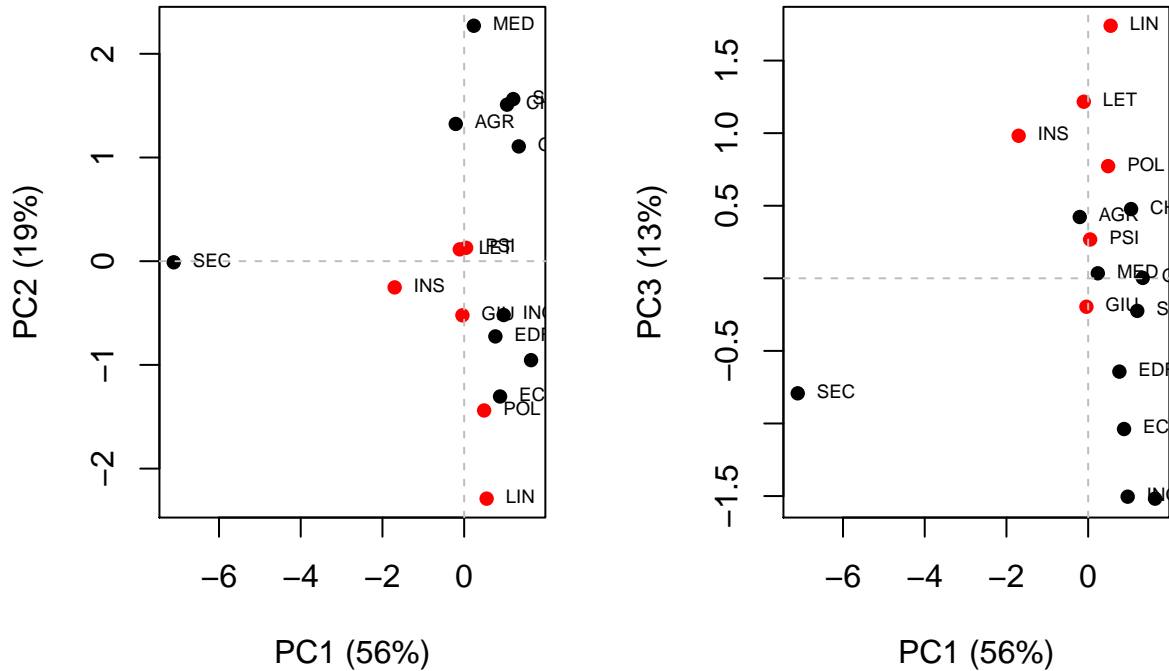
```
## Importance of components:
##               Comp.1   Comp.2   Comp.3   Comp.4   Comp.5
## Standard deviation  1.9904609 1.2233502 0.9124714 0.6733466 0.41110734
## Proportion of Variance 0.5659907 0.2137980 0.1189434 0.0647708 0.02414418
## Cumulative Proportion 0.5659907 0.7797886 0.8987321 0.9635029 0.98764703
##               Comp.6   Comp.7
## Standard deviation  0.243591568 0.164723781
## Proportion of Variance 0.008476693 0.003876275
## Cumulative Proportion 0.996123725 1.000000000
```


Si constata che le prime tre componenti principali riassumono il 90% della varianza del campione, mentre le prime due riassumono il 75%.

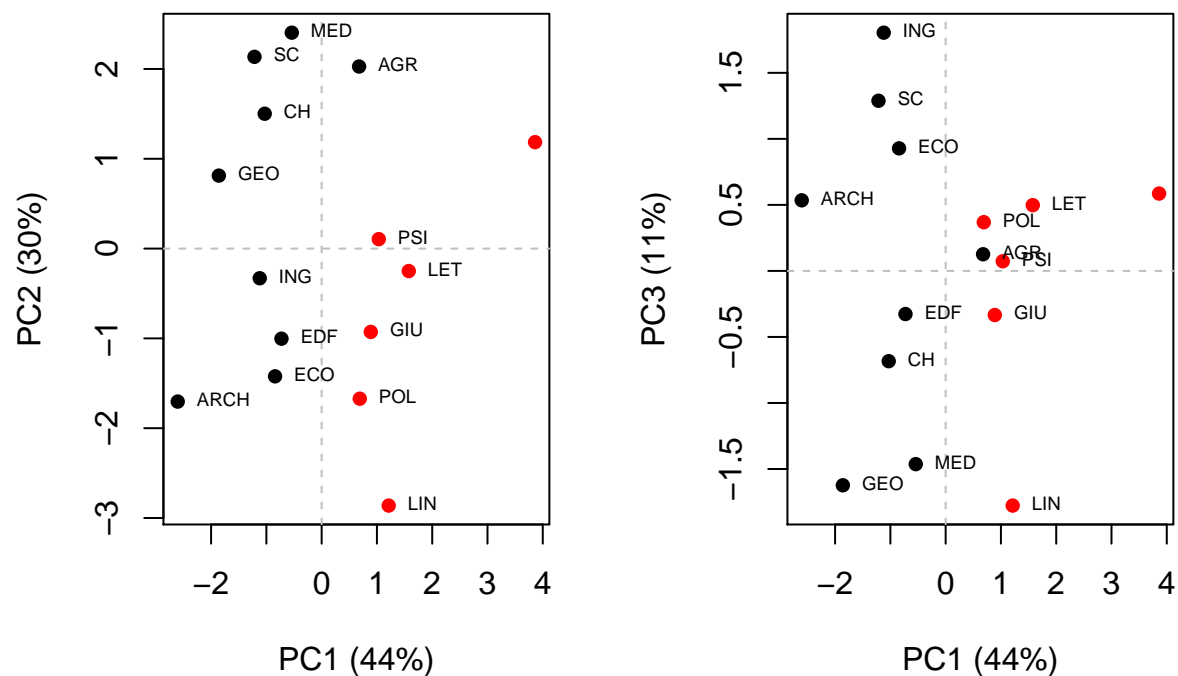
##	Comp.1	Comp.2	Comp.3	Comp.4
## occasioni_studio_formazione_scientifica	0.40	0.81	0.11	-0.37
## precedenti_esperienze_studio_lavoro	0.39	-0.62	0.66	-0.11
## piu_opportunita_lavoro	0.96	-0.19	0.08	-0.06
## lavoro_maggiormente_qualificato	0.98	-0.04	-0.05	-0.07
## lavoro_maggiormente_retribuito	0.52	-0.54	-0.60	-0.26
## paese_avanguardia_settore_interesse	0.90	0.34	0.06	0.06
## accordi_paesi	-0.84	-0.12	0.10	-0.48

La prima componente principale risulta essere fortemente correlata con la variabile “Più opportunità di lavoro” (-0.97) e con “Lavoro maggiormente qualificato” (-0.96) .

La seconda è dominata da “Motivi familiari e personali” (0.79) e da “Precedenti esperienze di studio e/o lavoro” (0.70).



L’ambito “Difesa e Sicurezza” risulta essere fortemente sbilanciato, motivo che ci spinge a riprovare una seconda analisi nella quale esso verrà escluso.

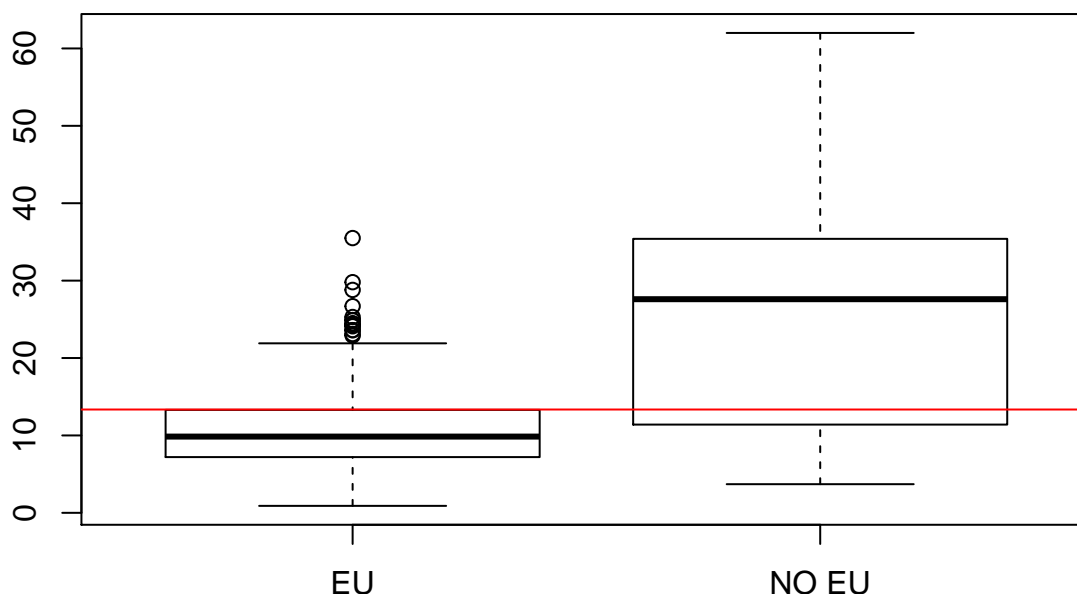


Nei grafici PC1/PC2 potremo quindi trovare ambiti i cui laureati sono motivati dalla ricerca di più lavoro maggiormente qualificato con l'avvicinarsi al lato sinistro mentre l'altezza esprime l'aumento di ricerca di occasioni di studio e di formazione scientifica con l'avvicinarsi al fondo scala negativo. Si può quindi notare come le ragioni appena elencate combacino con gli ambiti scientifici.

Capitolo 3

Europa

Ci chiediamo se l'appartenenza o meno all'Unione Europea abbia una qualche relazione con la percentuale di studenti che interrompono gli studi prematuramente.



Allo scopo di farci un'idea della posizione del nostro Paese, rappresentiamo la relativa media con una linea rossa. Ci accorgiamo di essere vicini al terzo quartile della distribuzione EU e al primo quartile della distribuzione NO-EU.

Notiamo una notevole differenza tra i due gruppi di stati: i paesi appartenenti all'Unione Europea risultano avere una minore percentuale di abbandoni prematuri degli studi, **ipoteticamente** dovuta a politiche di educazioni più motivanti (Es: Erasmus) o classifiche periodiche sulla qualità dell'insegnamento.

Eseguiamo un confronto tra la disoccupazione e la percentuale di abbandono degli studi per ogni Stato.

Analizziamo ora le distanze tra le nazioni europee in base alle due variabili *disoccupazione* e *abbandono prematuro*. Tali valori si possono rappresentare con un dendrogramma, uno schema che permette di individuare il numero di cluster da utilizzare per avere una certa distanza massima tra le nazioni che vi appartengono.

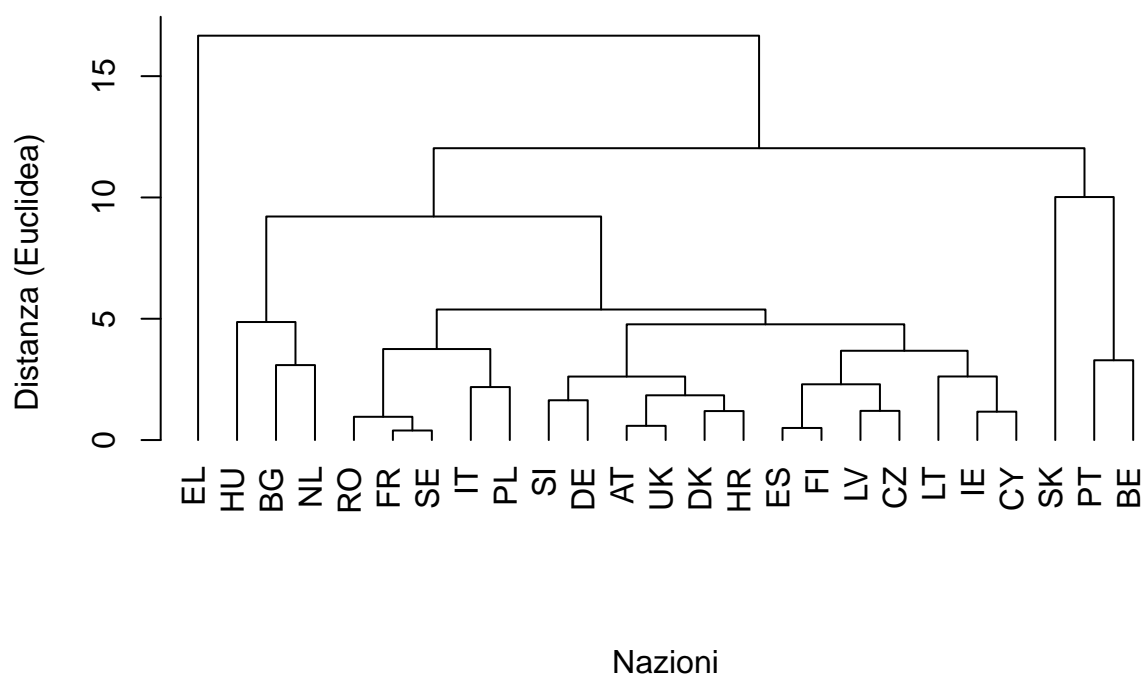


Figura 3.0.1: Distanza tra le nazioni europee.

Notiamo subito un'importante distanza della Grecia dal resto delle nazioni europee, seguita dalla Slovacchia. Il grafico ci permette di vedere che la grande maggioranza degli stati (18 su 25) risulta essere molto concentrata, con una distanza massima prossima al valore 5.

Proviamo ora a suddividere l'insieme di Nazioni in 3 cluster

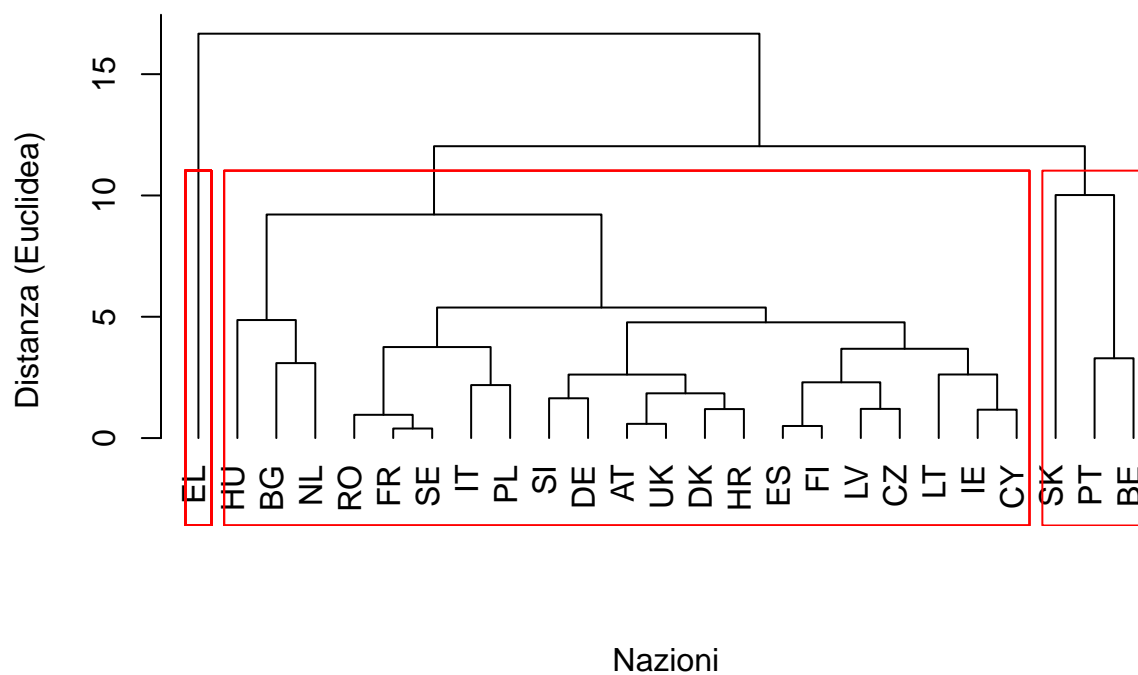


Figura 3.0.2: Cluster delle nazioni.

Scegliendo questi 3 cluster, possiamo rappresentare graficamente le nazioni in base alla loro distanza euclidea nel piano cartesiano delle variabili *disoccupazione* e *abbandono prematuro*.

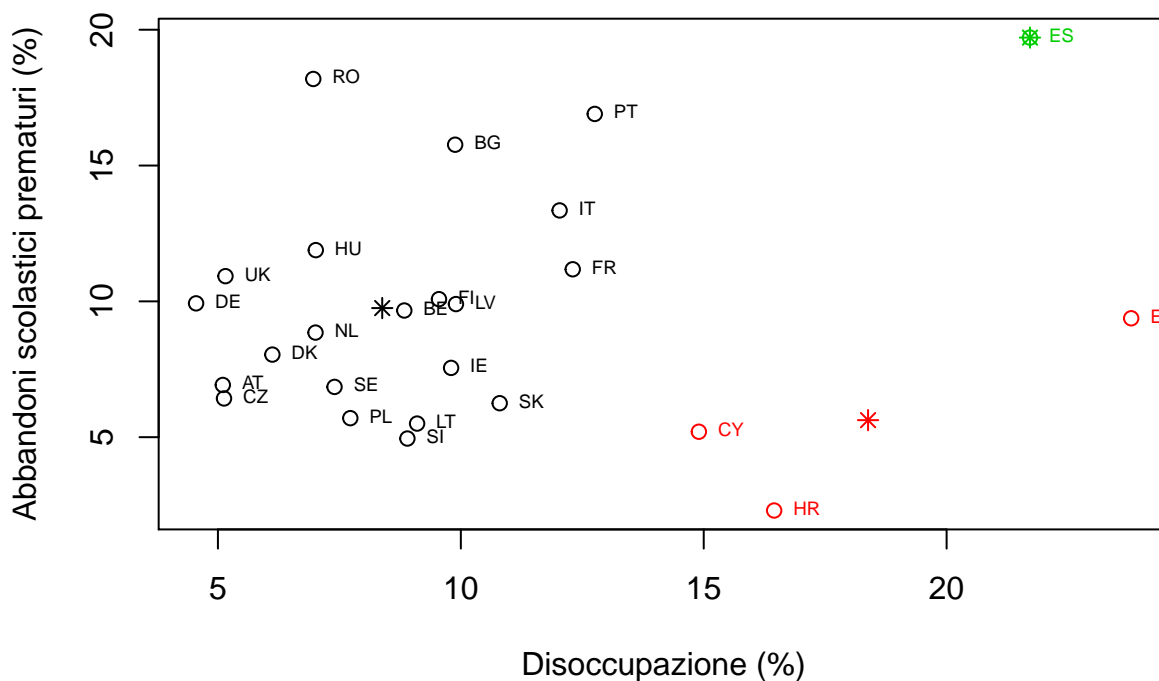


Figura 3.0.3: Disoccupazione e abbandoni scolastici prematuri.

Capitolo 4

Conclusioni e considerazioni personali

Non è stato semplice gestire la mole di dati analizzata, in particolare il dataset relativo ai laureati. La quantità di variabili, codificate tramite valori non intuitivi non ha semplificato l'analisi.

Sono rimasto soddisfatto del risultato dell'analisi, in quanto i valori ottenuti sembrano realistici e concordi con conoscenze in mio possesso o riflessioni personali.