

Analisi degli studi e dell'occupazione in Italia ed Europa

Filippo Bisconcini, 852144

24 May 2018

Indice

1	Premessa	2
2	Italia	3
2.1	Disoccupazione post Laurea	3
2.2	Riduzione dimensionale	6
2.3	Riduzione dimensionale delle motivazioni di traferimento	6
2.4	Riduzione dimensionale 2	8
2.5	3	9
3	Europa	11
4	Conclusioni	14
4.1	Considerazioni personali	14

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:

- 1) La qualità della vita delle province italiane: 2003/2004
- 2) Laureati e lavoro: 2015
- 3) 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 sarebbe statisticamente accettabile, provvediamo quindi ad includere un quarto dataset messo a disposizione dall'ISTAT dalla quale possiamo estrarre i dati relativi al 2015.

- 4) 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%.

2.1.1 Analisi per regione

Una volta analizzata la disoccupazione ad un anno dalla laurea, ci chiediamo se essa sia legata al livello di disoccupazione della regione.

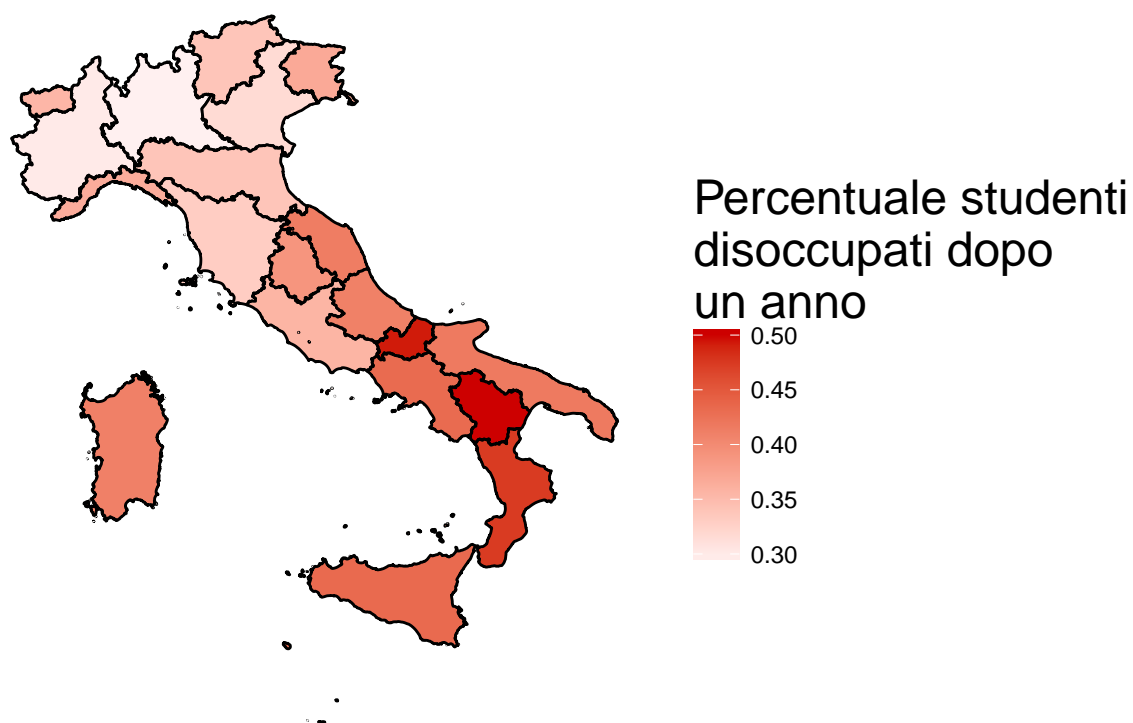


Figura 2.1.1: Percentuale studenti disoccupati dopo un anno.

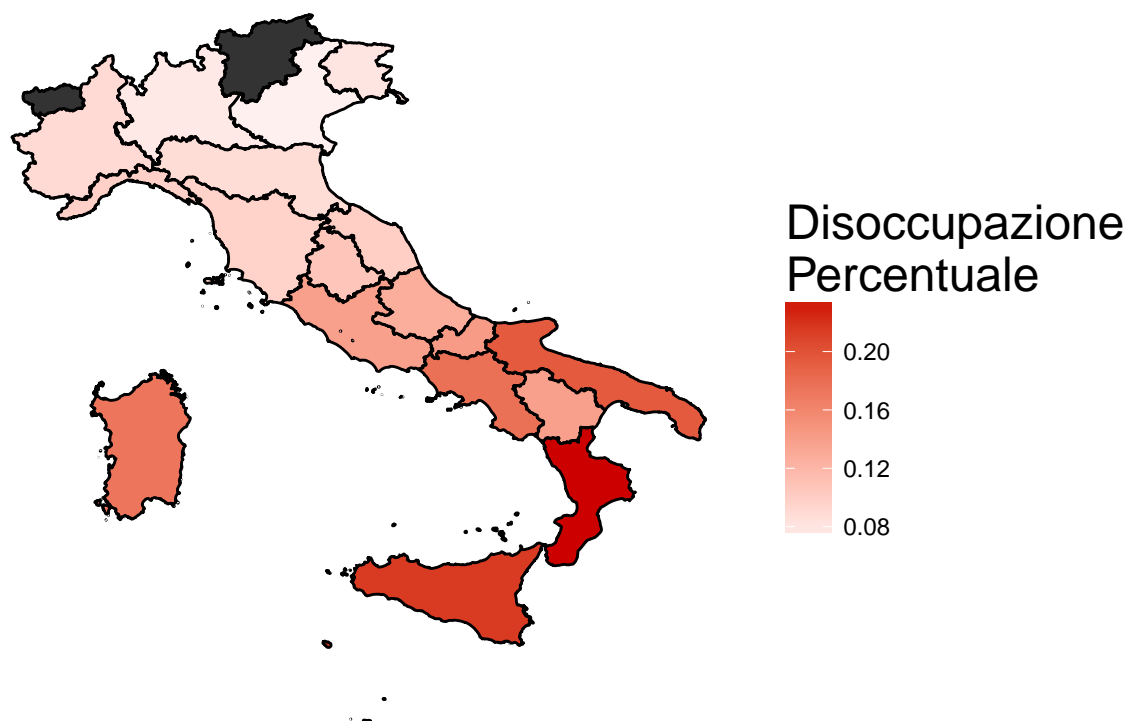
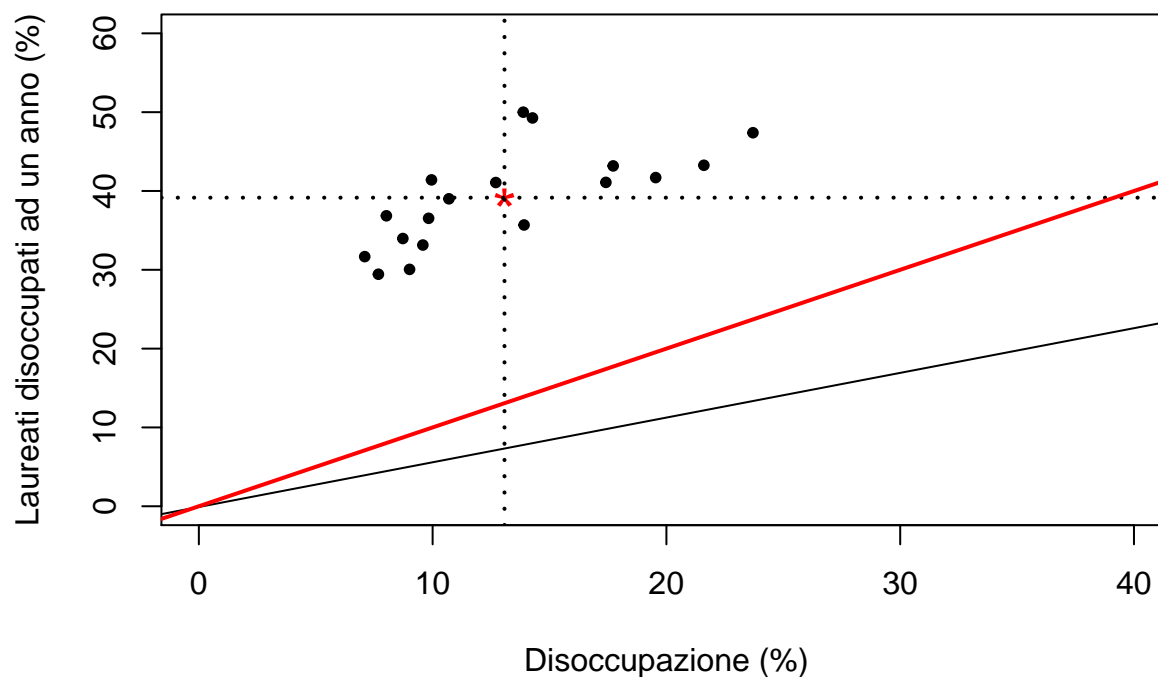


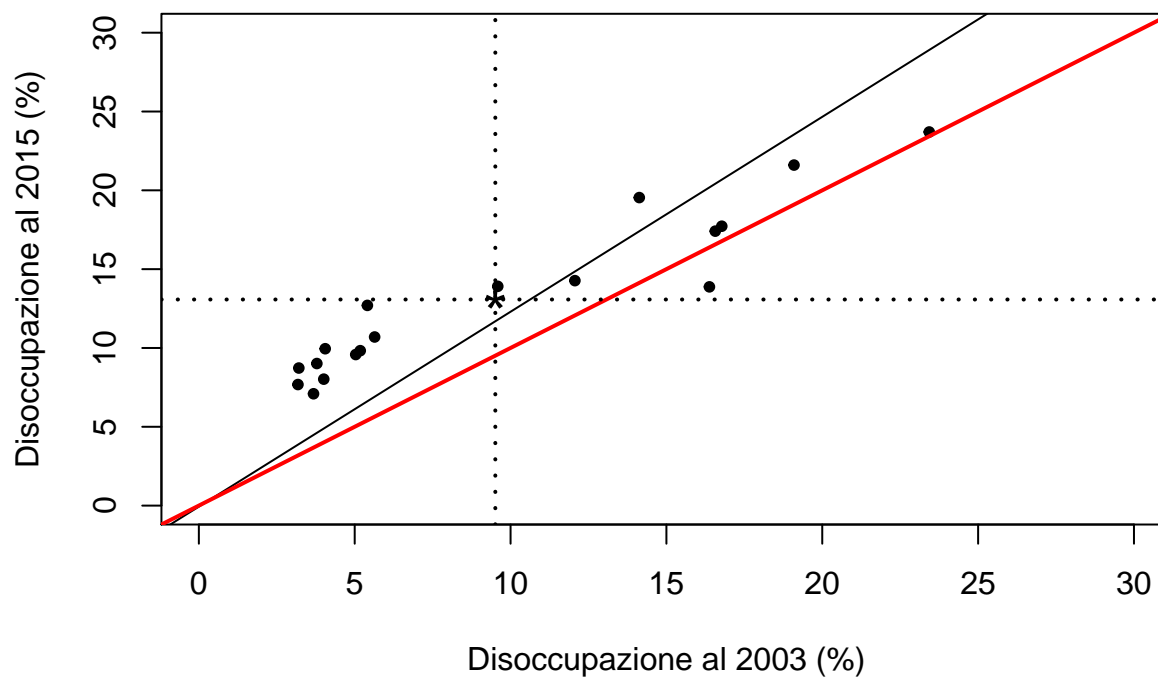
Figura 2.1.2: Percentuale di disoccupazione regionale nel 2015.

Disoccupazione in Italia



Avendo a disposizione due dataset costruiti a dodici anni di distanza, possiamo calcolarci la variazione nella percentuale di disoccupazione. Utilizziamo la solita rappresentazione grafica per mostrare la distribuzione dell'indice appena calcolato.

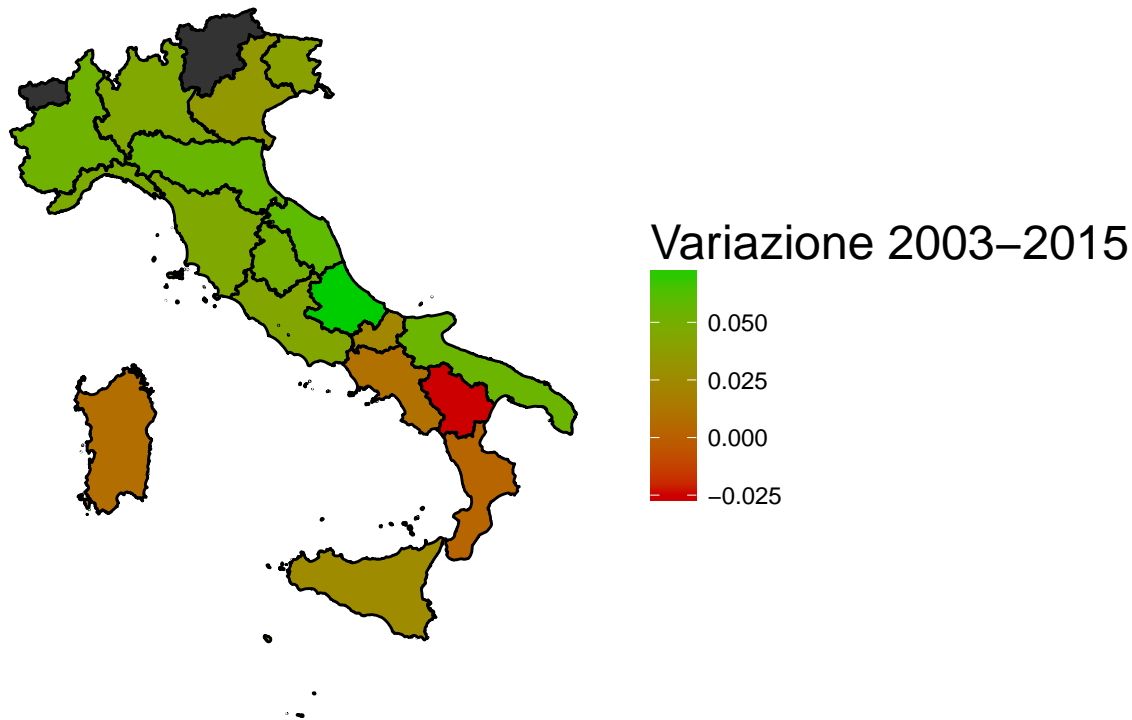
Disoccupazione in Italia



Il grafico soprastante indica la correlazione tra la disoccupazione nelle due annate 2003 e 2015. Punti al di

sopra della diagonale indicano regioni la cui disoccupazione nel 2015 risulta minore della rilevazione eseguita nel 2003.

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 %



2.2 Riduzione dimensionale

Dividiamo il campione di laureati, raggruppato per ambito, in due macro ambiti (*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

2.3 Riduzione dimensionale delle motivazioni di trasferimento

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

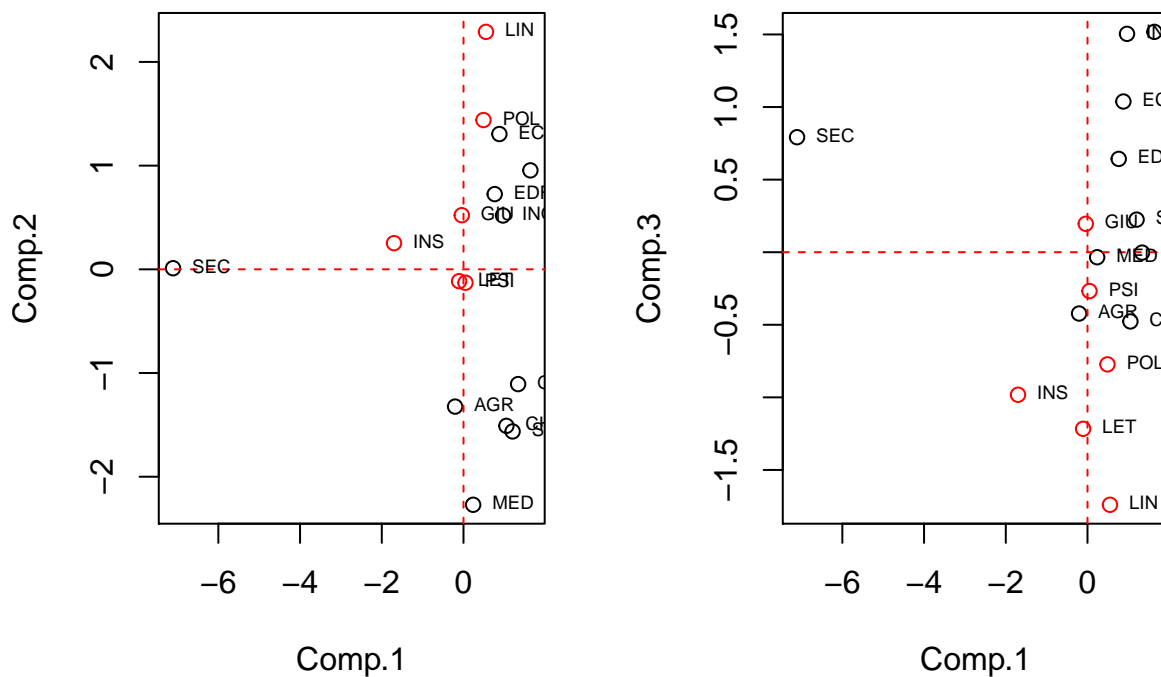
```
## Importance of components:
##               Comp.1   Comp.2   Comp.3   Comp.4   Comp.5
## Standard deviation   1.9904609 1.2233502 0.9124714 0.6733466 0.41110734
## Proportion of Variance 0.5659907 0.2137980 0.1189434 0.0647708 0.02414418
## Cumulative Proportion 0.5659907 0.7797886 0.8987321 0.9635029 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
```

Constatiamo che con le prime tre componenti principali riusciamo a riassumere il 90% della varianza del campione e più del 75% con le prime due.

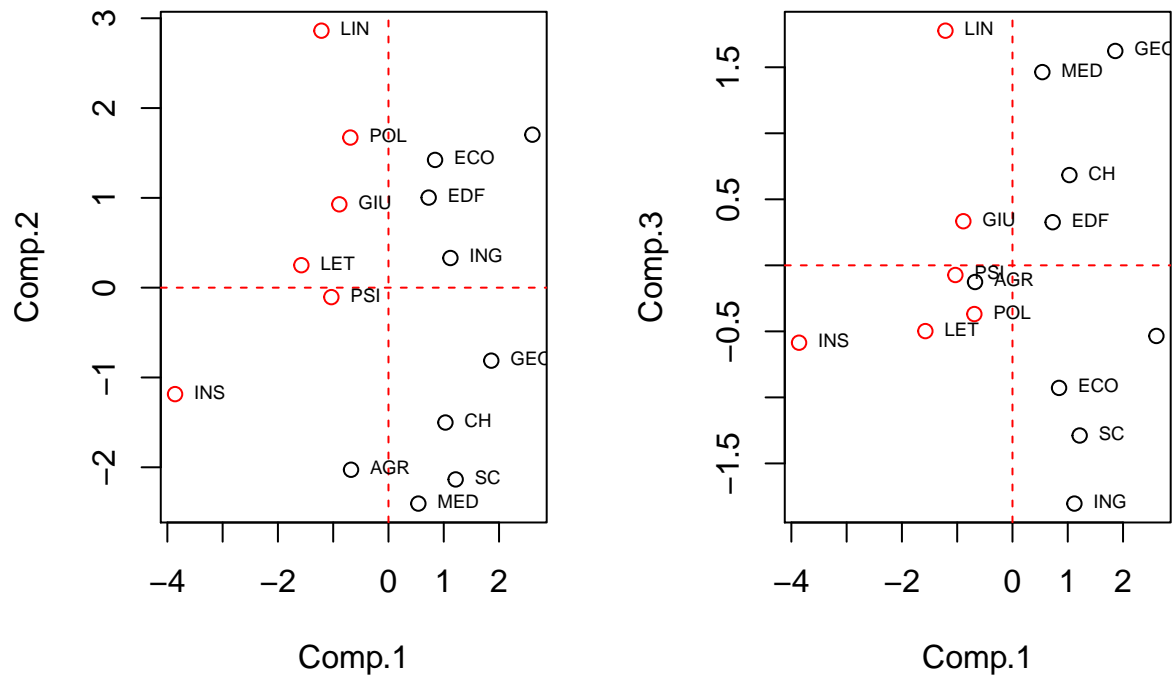
##	Comp.1	Comp.2	Comp.3	Comp.4
## occasioni_studio_formazione_scientifica	-0.40	0.81	0.11	-0.37
## precedenti_esperienze_studio_lavoro	-0.39	-0.62	0.66	-0.11
## piu_opportunita_lavoro	-0.96	-0.19	0.08	-0.06
## lavoro_maggiormente_qualificato	-0.98	-0.04	-0.05	-0.07
## lavoro_maggiormente_retribuito	-0.52	-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 risulta essere fortemente correlata con la variabile “Lavoro maggiormente qualificato” (-0.98) e con “Più opportunità di lavoro” (-0.96).

La seconda è dominata da “Occasioni di studio e formazione scientifica” (0.81).



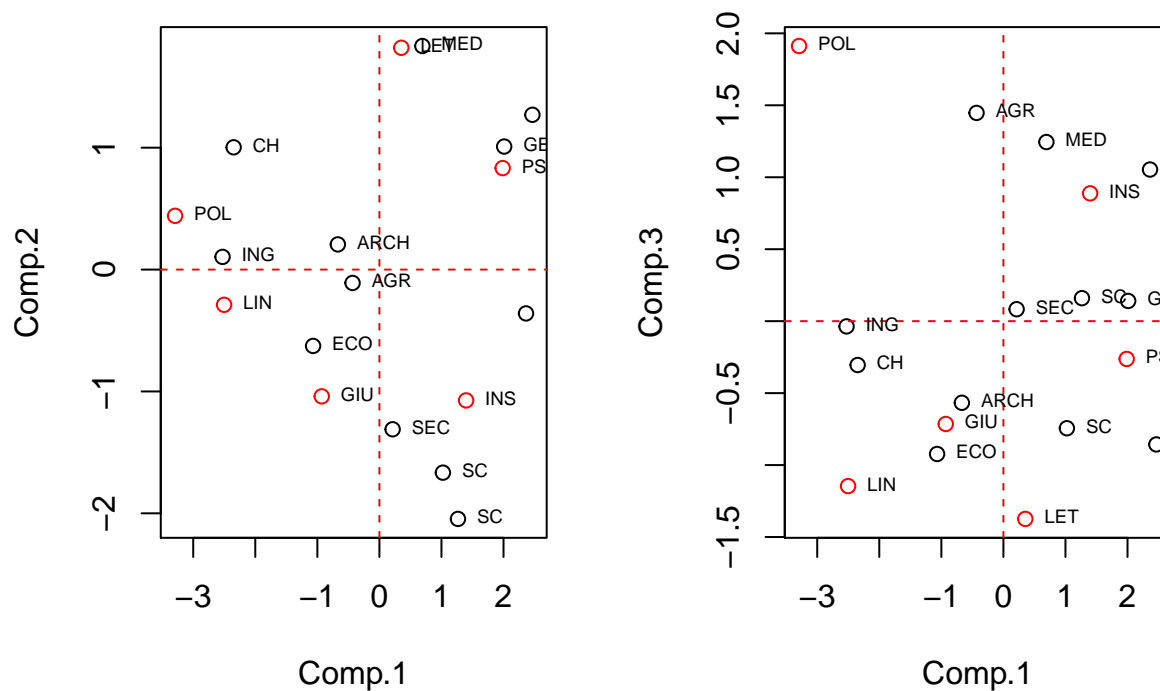
L’ambito “Difesa e Sicurezza” risulta essere fortemente sbilanciato, motivo che ci spinge a riprovare una seconda analisi nella quale esso verrà escluso.



2.4 Riduzione dimensionale 2

```
## Importance of components:
##                               Comp.1   Comp.2   Comp.3   Comp.4   Comp.5
## Standard deviation      1.7725909 1.1251756 0.9305506 0.9064704 0.59652249
## Proportion of Variance  0.4488684 0.1808600 0.1237035 0.1173841 0.05083415
## Cumulative Proportion  0.4488684 0.6297284 0.7534319 0.8708159 0.92165010
##                               Comp.6   Comp.7
## Standard deviation      0.5624976 0.4817113
## Proportion of Variance  0.0452005 0.0331494
## Cumulative Proportion  0.9668506 1.0000000

##                               Comp.1 Comp.2 Comp.3 Comp.4
## voto_medio                0.55  -0.58  0.18  0.48
## laureati                   -0.59   0.06  0.77 -0.04
## pa                         0.85   0.07 -0.11  0.32
## lascerebbero_italia       0.47  -0.54 -0.10 -0.66
## mag.qual                   -0.82   0.03 -0.34  0.06
## delta_03_15               -0.51  -0.77  0.06  0.06
## dis                        0.79   0.22  0.31 -0.21
```

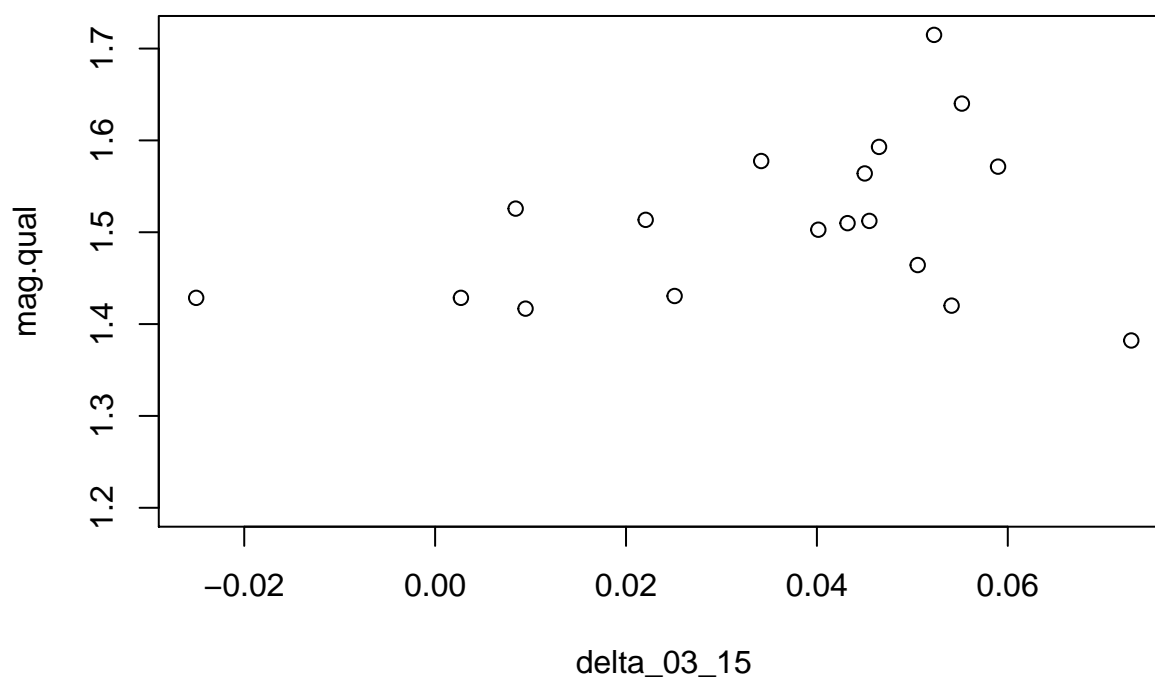


2.5 3

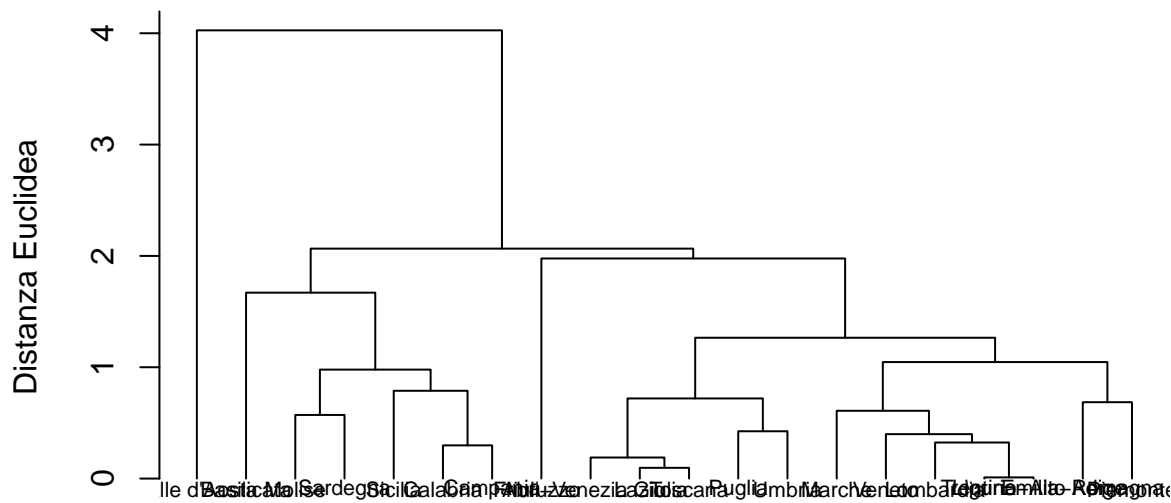
Avendo scoperto che la ricerca di lavoro maggiormente qualificato riassumeva molto bene tutte le ragioni per la quale un lareato sarebbe andato all'estero, utilizziamo quest'ultima per confrontarla con

...

della regione



Genitori e Figli



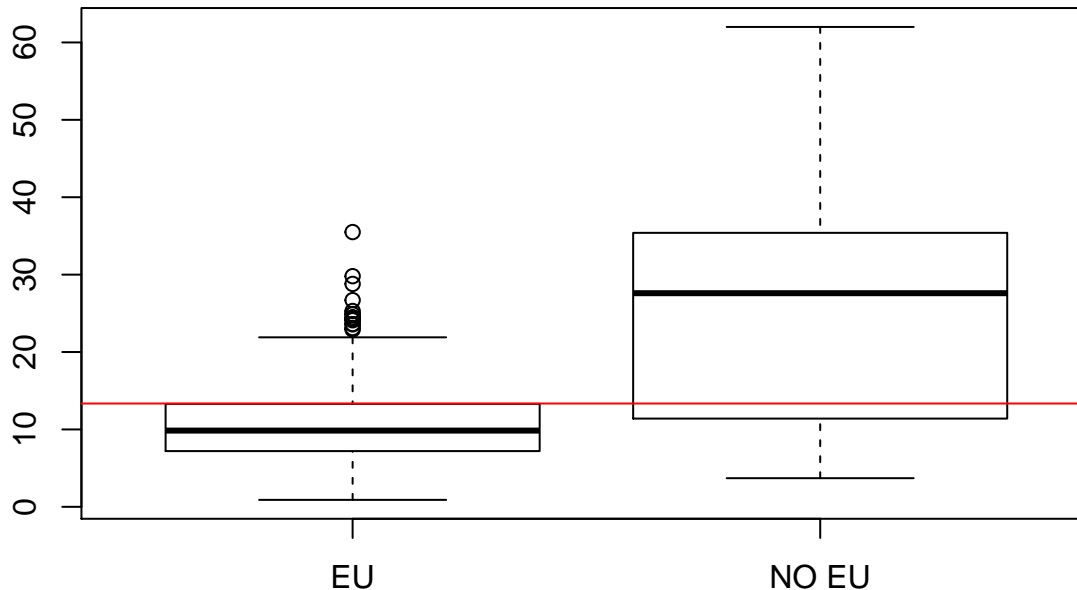
Unita...
Legame Medio

##	[,1]	[,2]
## [1,]	-8	-17
## [2,]	-7	-16
## [3,]	-6	2
## [4,]	-3	-4
## [5,]	-9	1
## [6,]	-20	5
## [7,]	-13	-18
## [8,]	-11	-14
## [9,]	-10	6
## [10,]	-5	-12
## [11,]	3	7
## [12,]	-15	4
## [13,]	8	12
## [14,]	9	10
## [15,]	11	14
## [16,]	-2	13
## [17,]	-1	15
## [18,]	16	17
## [19,]	-19	18

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.



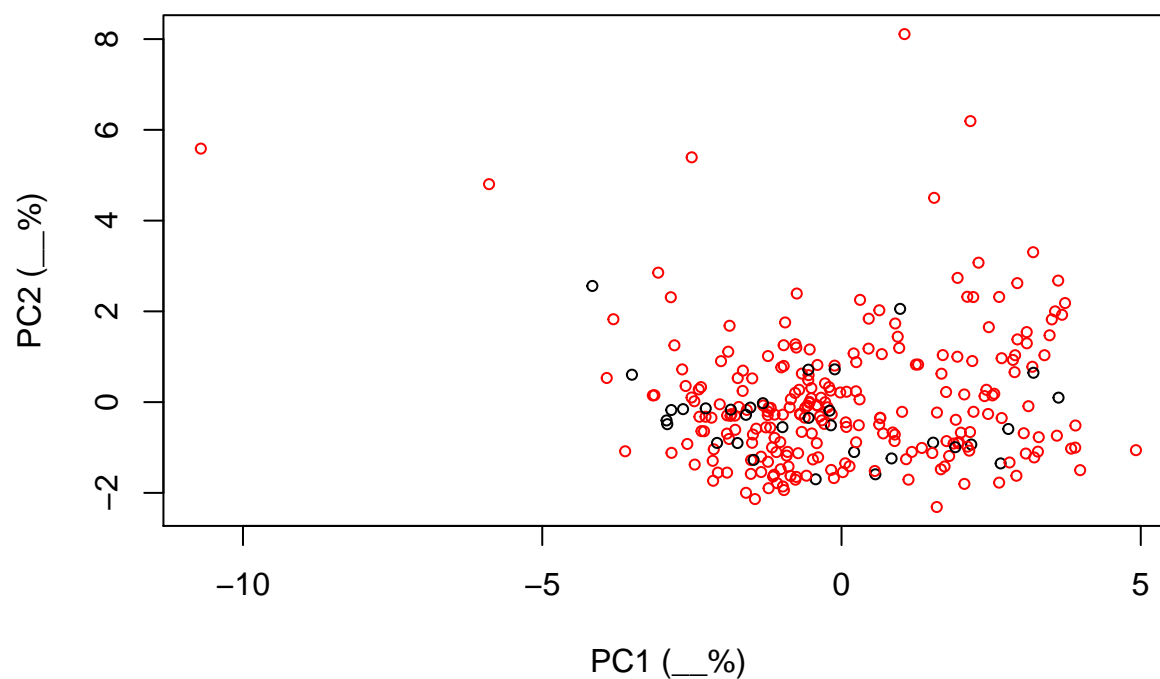
Allo scopo di farci un'idea della posizione del nostro Paese, rappresentiamo la media dei nostri abbandoni con una linea rossa. Ci accorgiamo di essere vicini al terzo quartile della distribuzione EU e al primo quartile della distribuzione NO-EU.

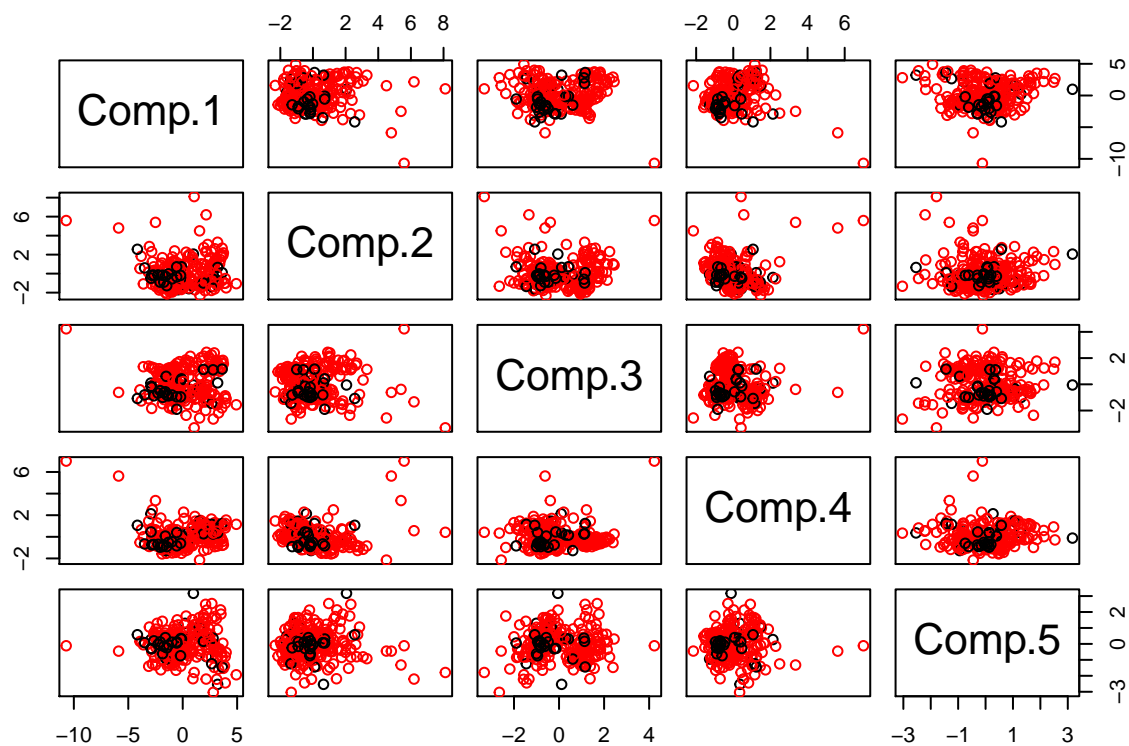
Notiamo una notevole differenza tra i due gruppi di stati: i paesi appartenenti all'Unione Europea risultano avere una minore percentuale di abbandoni prematuri degli studi.

```
## Importance of components:
##               Comp.1   Comp.2   Comp.3   Comp.4   Comp.5
## Standard deviation  2.0683688 1.4352009 1.1385331 0.99663440 0.90982810
## Proportion of Variance 0.3889227 0.1872547 0.1178416 0.09029819 0.07525338
## Cumulative Proportion 0.3889227 0.5761774 0.6940190 0.78431718 0.85957056
##               Comp.6   Comp.7   Comp.8   Comp.9
## Standard deviation  0.67631492 0.66834668 0.50485993 0.44301502
## Proportion of Variance 0.04158199 0.04060793 0.02317123 0.01784203
## Cumulative Proportion 0.90115254 0.94176048 0.96493171 0.98277374
##               Comp.10   Comp.11
```

```
## Standard deviation      0.325539809 0.288985640
## Proportion of Variance 0.009634197 0.007592064
## Cumulative Proportion  0.992407936 1.000000000

##      Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6
## dens   -0.42  0.57  -0.02  0.60  -0.18  0.06
## fert   -0.41  0.25  -0.70  -0.40  -0.12  0.11
## nmigr   -0.67  -0.37  0.15  -0.02  -0.29  -0.40
## natg    -0.53  0.64  -0.37  -0.06  0.10  0.06
## life    -0.48  0.37  0.50  -0.56  0.13  0.01
## gdp_pc  -0.80  0.18  0.33  0.28  -0.04  0.18
## inc_pc  -0.84  -0.01  0.39  -0.16  -0.13  0.19
## empl    -0.75  -0.53  -0.20  -0.06  -0.09  0.03
## unempl   0.60  0.68  0.25  -0.10  0.13  -0.16
## hedu    -0.71  0.31  -0.16  0.09  0.23  -0.42
## ledu     0.44  0.37  0.03  -0.15  -0.76  -0.07
```





Capitolo 4

Conclusioni

4.1 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.