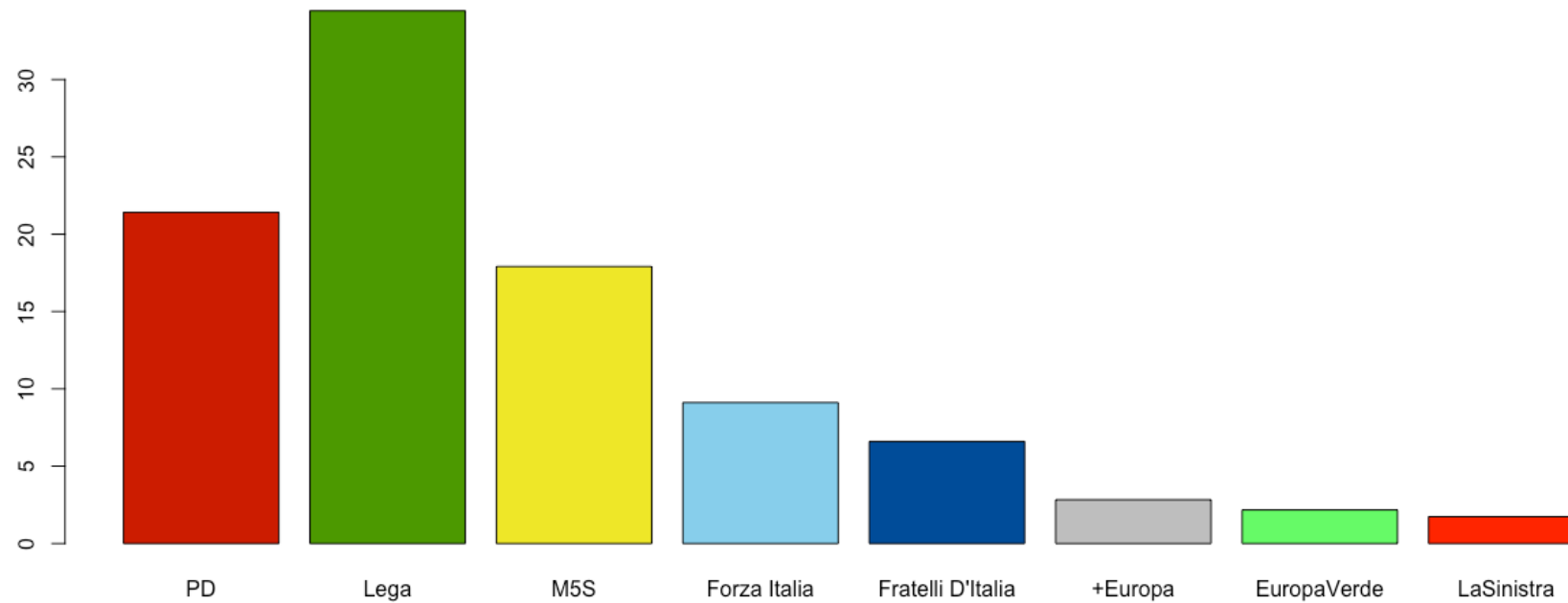


Risultati Elezioni Europee 26/05/2019 in Italia



PROGETTO DI INFERENZA STATISTICA: ANALISI DELL'ASTENSIONISMO IN ITALIA

Elezioni Europee del 26/05/2019



OBIETTIVO DELL'ANALISI

- Analizzare il fenomeno dell'astensionismo alle consultazioni elettorali, individuandone le cause.
- Verificare l'esistenza di un'eventuale localizzazione geografica dell'astensionismo nel nostro Paese.



STRUTTURA DEL DATASET

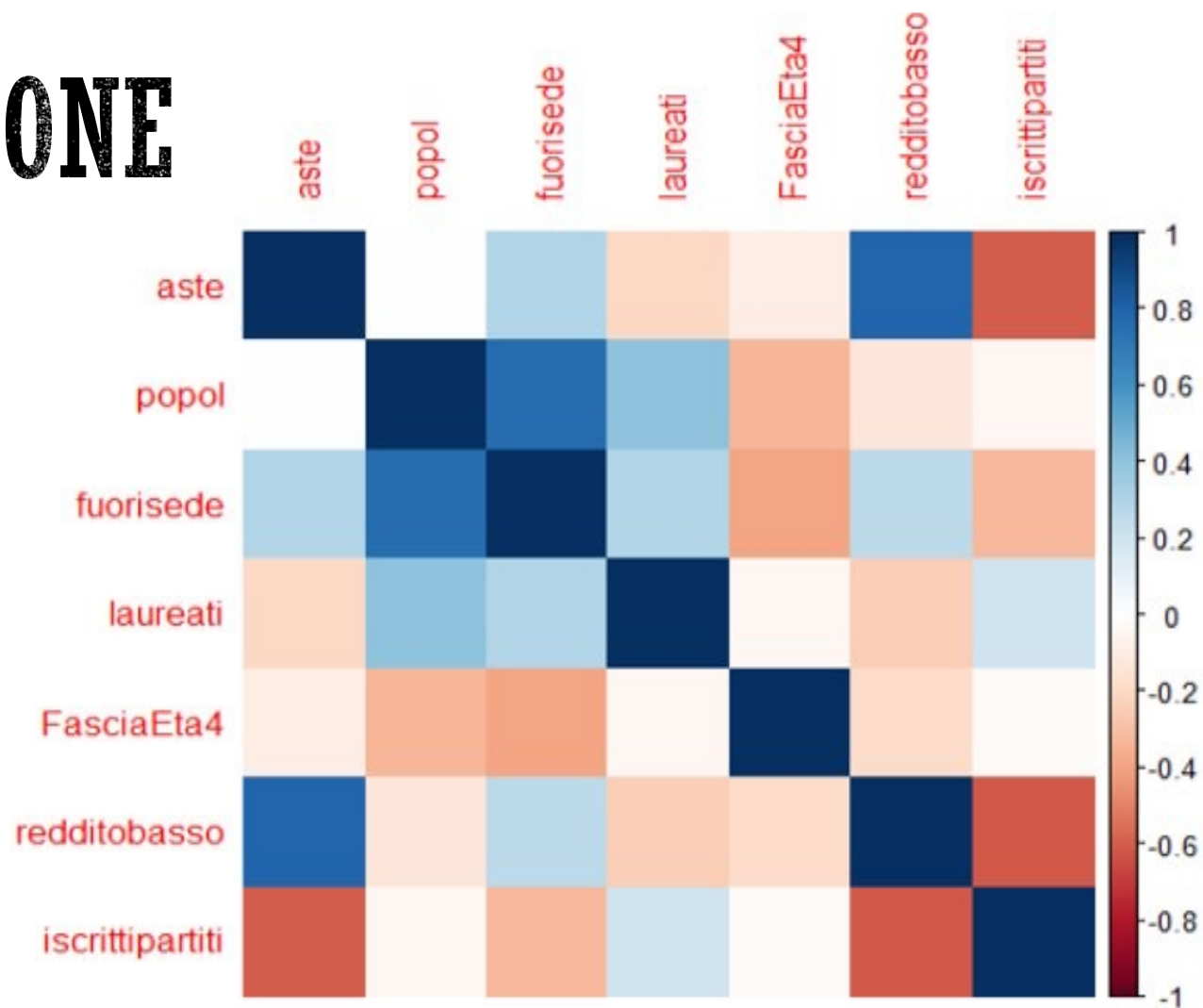
X1,...,X8	Risultato elettorale del Partito X_i in %
popol	Popolazione della provincia
aste	% di astenuti sugli aventi diritto al voto
redditobasso	% di persone con reddito basso (< 15000 euro)
disoccup	Tasso di disoccupazione
FasciaEta	% popol. in fasce d'età: {(0-18), (19-35), (36-50), (51-70), (70-oltre)}
Fuorisede	% degli studenti che studiano in una regione diversa da quella di residenza
LivIstruzione	% della popolazione con titolo pari o superiore alla laurea triennale
Iscritti	% della popolazione iscritta a partiti politici
Zone	Zona della provincia: {Nord, Centro, Sud}

Unità di analisi: province

Totale osservazioni: 107



HEAT MAP DELLA CORRELAZIONE



MODELLO LINEARE

Variabile risposta: % degli astenuti sugli aventi diritto al voto

Covariate	Fuorisede	Liv. Istruzione *	FasciaEtà4 **	Redditobasso ***	Iscritti **
Estimate (β)	1.3960	-0.7873	1.0975	0.4117	-3.1856
P-value	0.5042	0.0471	0.0054	0.0007	0.0037
VIF	1.7315	1.4275	1.3536	3.5734	2.0478

Variabile 'Fuorisede'
non significativa!



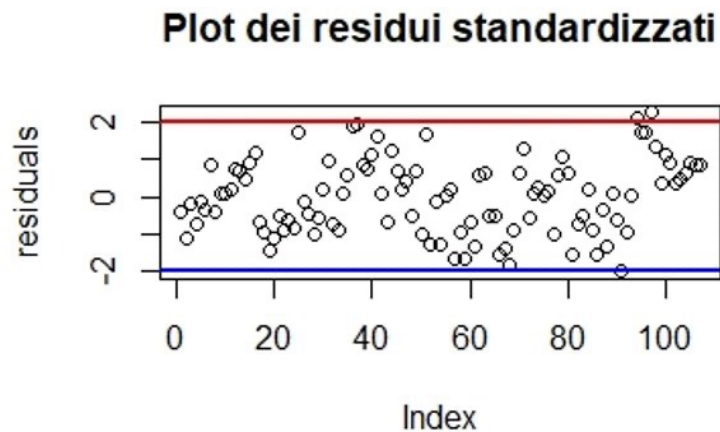
R^2	0.7161
R^2_{adj}	0.6991
AIC	697.3968



MODELLO LINEARE

Variabile risposta: % degli astenuti sugli aventi diritto al voto

Covariate	Liv. Istruzione *	FasciaEtà4 **	Redditobasso ***	Iscritti **
Estimate (β)	-0.7015	0.9529	0.4027	-3.4656
P-value	0.0403	0.0069	0.0008	0.0006
VIF	1.2750	1.1549	3.5272	1.7359



Shapiro test per la normalità dei residui:

P_value = 0.2021

R^2	0.7148
R^2_{adj}	0.7007
AIC	695.8766



CROSS-VALIDAZIONE DEL MODELLO

E' stata svolta una cross validazione per validare la stabilità e il GOF del modello.

Dividiamo K volte il dataset in:

- **Training** set – 80%
- **Test** set – 20%

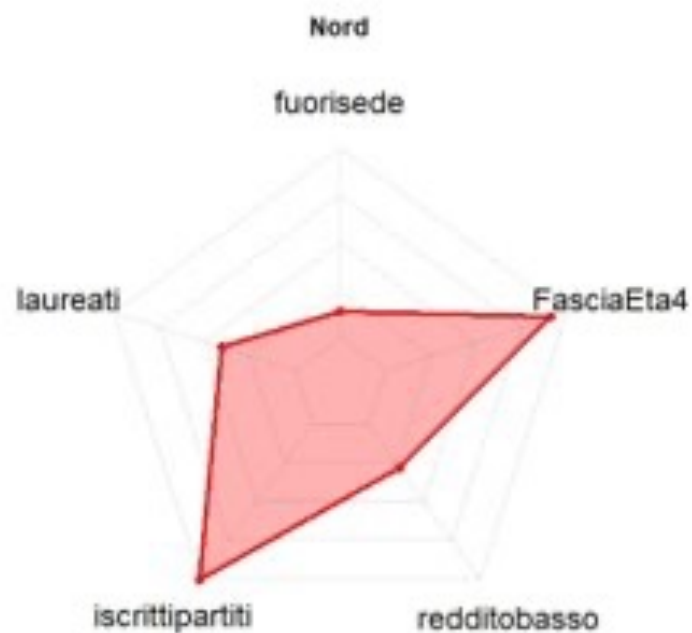
La partizione è randomica.

K	R^2_{adj}
5	0.6905
10	0.6868
20	0.6927
40	0.6951

R^2_{adj} modello originario	0.7007
-----------------------------------	--------



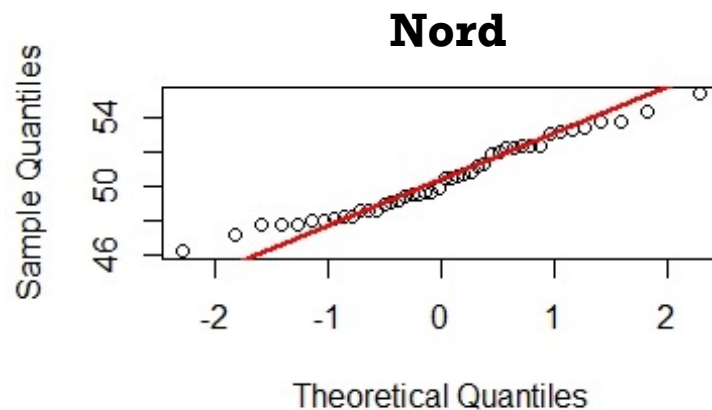
DIAGRAMMA A RAGNATELA



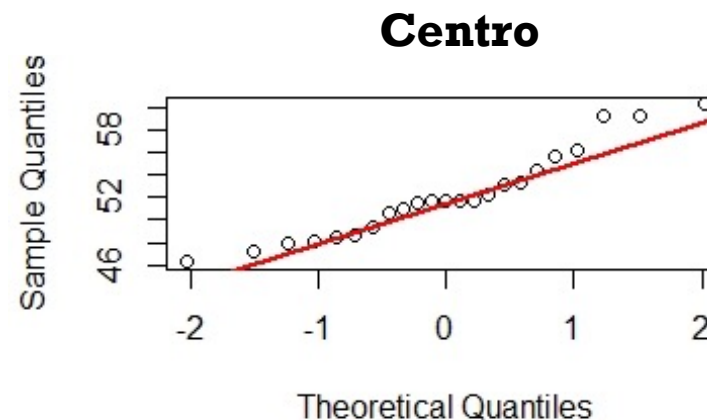
TEST ANOVA

Domanda: la media dell'astensionismo è la stessa tra nord, sud e centro?

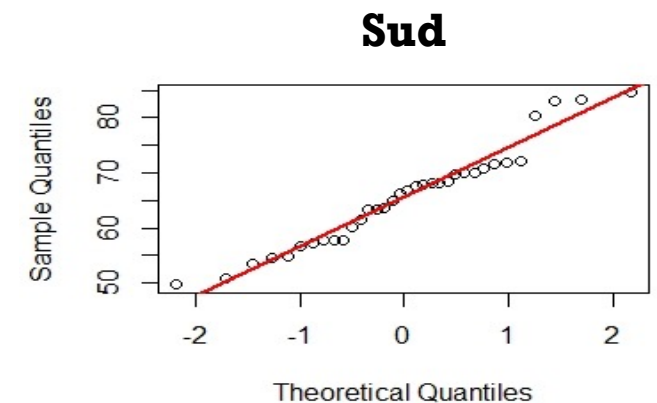
Verifichiamo gaussianità ed omoschedasticità dei dati



P-Value 0.2185



0.1625



0.2152

P-Value test di Bartlett: 0.0644



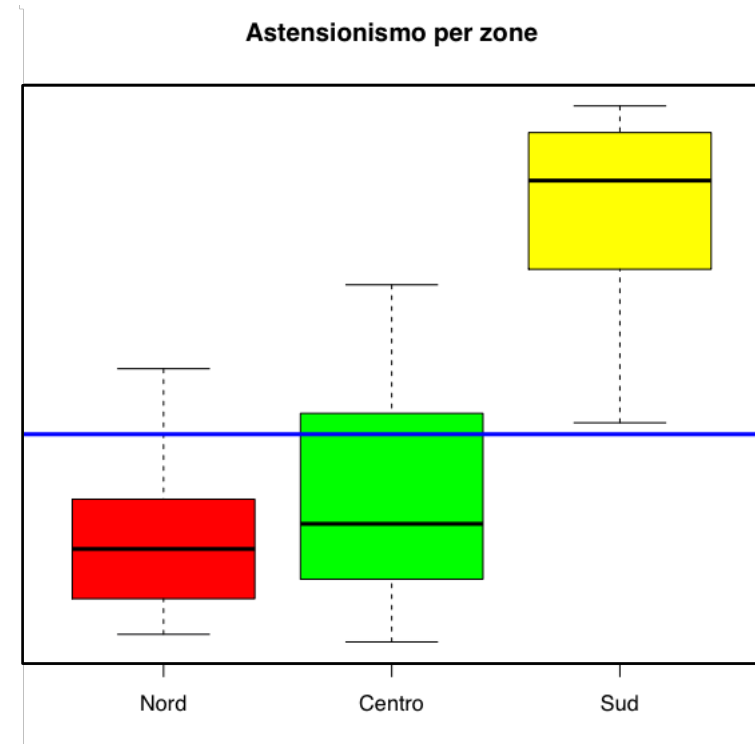
TEST ANOVA

H_0 : media dell'astensionismo tra Nord, Centro e Sud uguali

vs

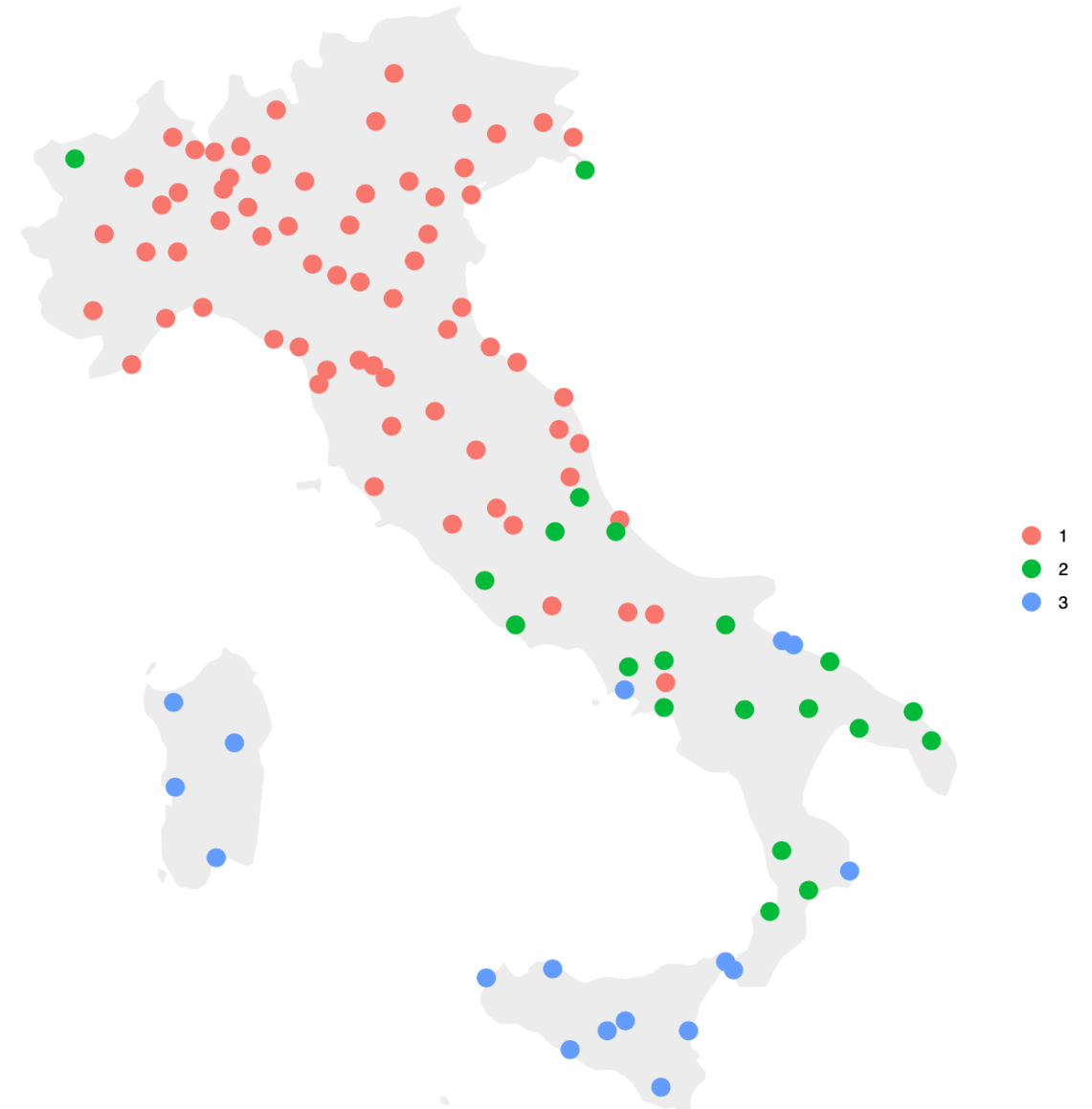
H_1 : media dell'astensionismo tra Nord, Centro e Sud diverse

p-value = 0

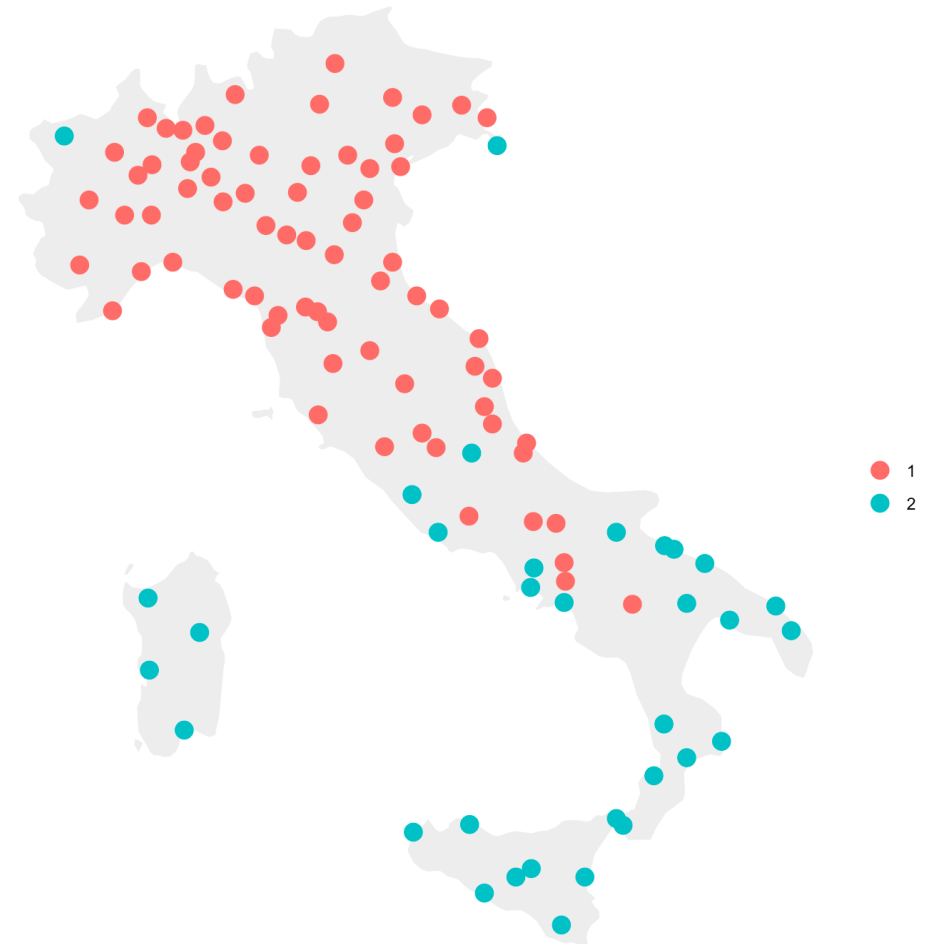
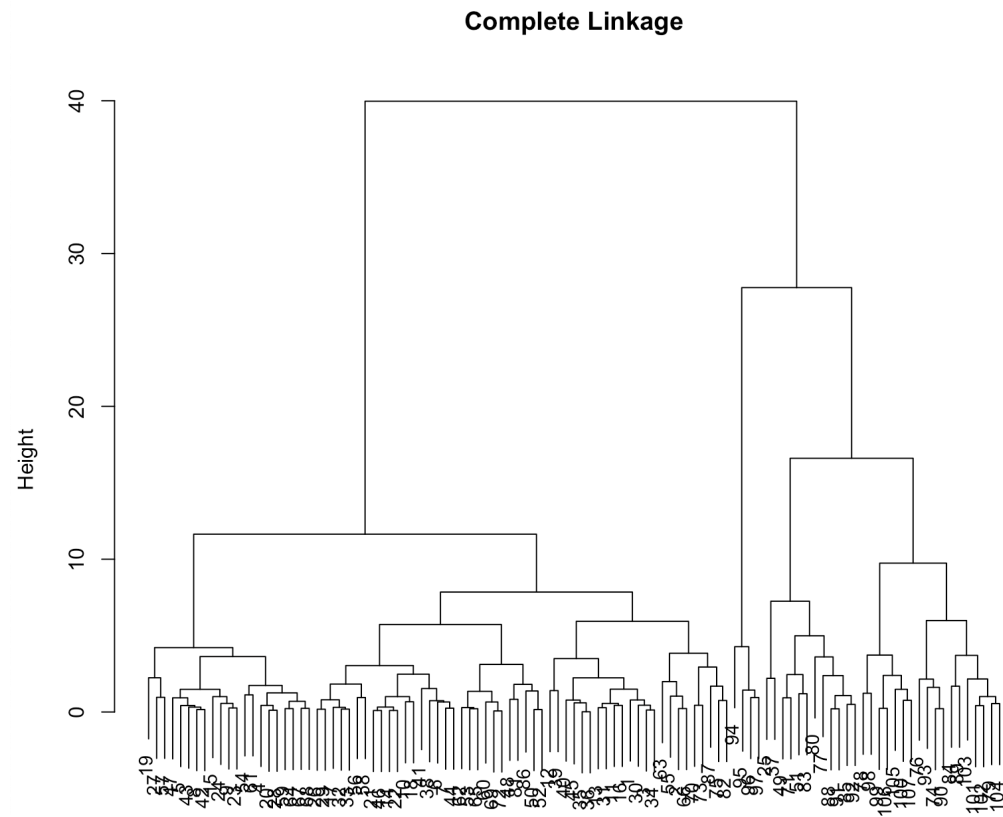


K-MEANS CLUSTERING

Domanda: dai dati dell'astensionismo emerge un'evidente divisione geografica?



HIERARCHICAL CLUSTERING



TEST ANOVA

H_0 : media dell'astensionismo tra Nord e Centro uguali

vs

H_1 : media dell'astensionismo tra Nord e Centro diverse

p-value = 0.107

Astensionismo per zone

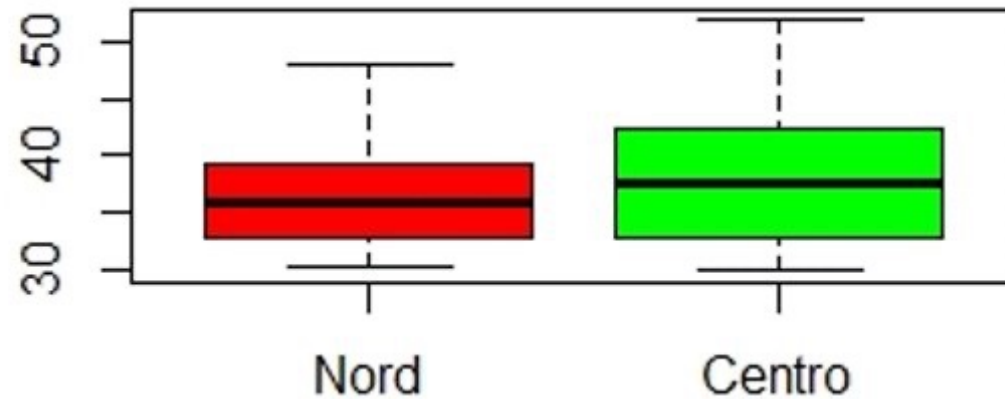


DIAGRAMMA A RAGNATELA

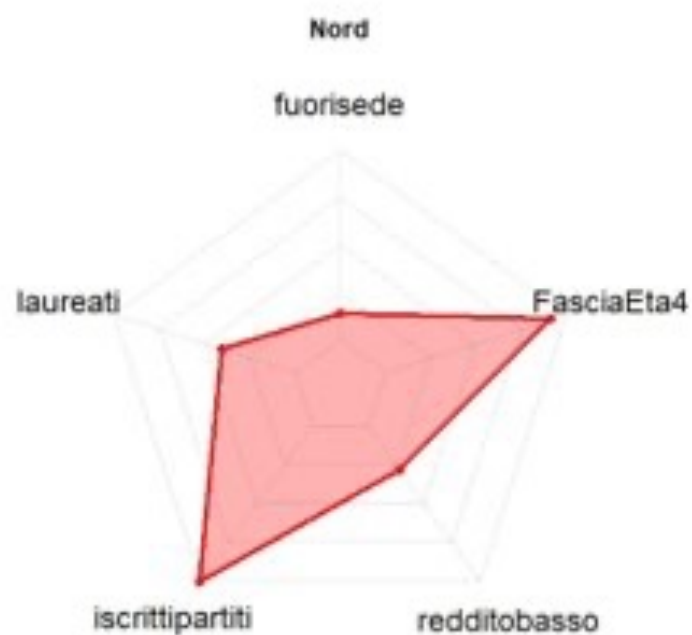
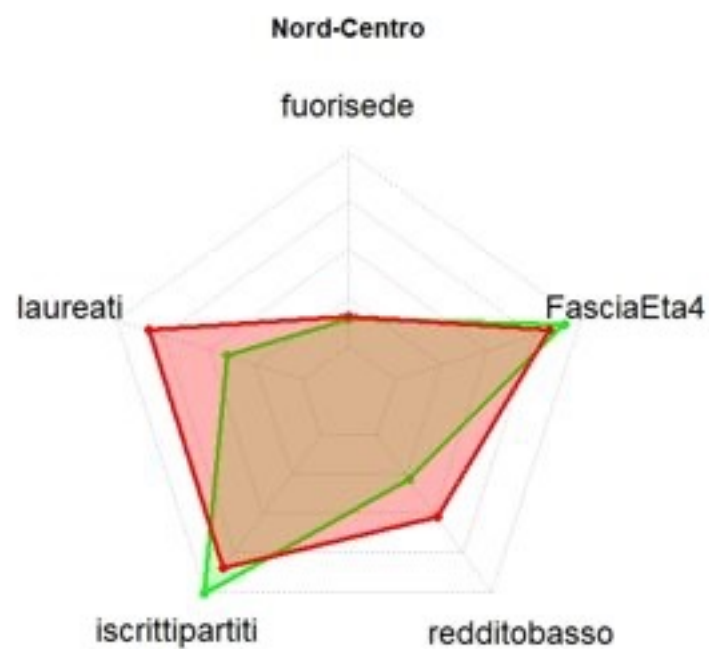


DIAGRAMMA A RAGNATELA



CONCLUSIONI

- Si è osservato che i fattori che influenzano l'astensionismo sono:

- livello di istruzione della popolazione,



- età anagrafica,



- condizione socio-economica,



- partecipazione all'attività politica.



- Sugeriamo alla classe dirigente di incentivare i giovani e le fasce di popolazione meno abbienti al voto, cercando anche di promuovere la partecipazione di tutti i cittadini all'attività politica.



BIBLIOGRAFIA

- Fonti per la costruzione del dataset:
 - Ministero dell'interno
 - ISTAT
 - Il Sole 24Ore
 - Repubblica
 - ANAV
- James, Witten, Hastie, Tibshirani; An introduction to Statistical Learning; Springer.
- Casella, Berger; Statistical Inference; Pearson.





**GRAZIE PER
L'ATTENZIONE!**



**FRANCESCO TRIPOLI
MASSIMO RIZZUTO
MICHELE MASSARIA**