

REPORT SCIENTIFICO – ESAME DEL 30 MAGGIO 2020

ANALISI DEL PENSIERO POLITICO DEGLI ITALIANI PRIMA DELLE  
ELEZIONI DEL 2018

TRAMITE L'IMPIGO DI ACP E ANALISI DEI CLUSTER



ANNO ACCADEMICO 2019 – 2020

## DESCRIZIONE DEL DATA SET

L'elaborato di seguito proposto mira ad analizzare il pensiero politico degli italiani prima delle elezioni dell'anno 2018.

I dati oggetto di studio sono stati scaricati dal sito [itanes.org](http://itanes.org), forniti dall'associazione Italian national election Studies.

Il dataset scelto per l'analisi è "Itanes 2018" relativi a delle interviste effettuate ad un campione di 2573 soggetti prima e dopo le elezioni politiche dell'anno 2018.

L'obiettivo del nostro studio è analizzare il pensiero dei cittadini italiani in merito a quattro tematiche:

- Economia (2 items)
- Immigrazione (3 items)
- Europa (2 items)
- Interesse per la politica (3 items)

Verrà dunque da prima svolta un'analisi in componenti principali per creare delle componenti che siano in grado di riassumere gli item di ogni singola tematica e poi verranno utilizzate le componenti estratte per effettuare una Cluster Analysis utilizzando prima una tecnica gerarchica e poi una tecnica non gerarchica.

In primo luogo, è necessario effettuare un'iniziale analisi descrittiva per verificare se sia o meno necessario effettuare data cleaning eliminando una o più osservazioni anomale.

Dapprima osserviamo le **variabili socio – demografiche**:

		Genere			
		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	Uomo	1208	46,9	46,9	46,9
	Donna	1365	53,1	53,1	100,0
	Totale	2573	100,0	100,0	

		Classe d'età			
		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	18-34 anni	505	19,6	19,6	19,6
	35-54 anni	1253	48,7	48,7	68,3
	55+ anni	815	31,7	31,7	100,0
	Totale	2573	100,0	100,0	

### Regione di residenza

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	Piemonte	155	6,0	6,0	6,0
	Valle d'Aosta	6	,2	,2	6,3
	Lombardia	408	15,9	15,9	22,1
	Trentino Alto Adige	24	,9	,9	23,0
	Veneto	223	8,7	8,7	31,7
	Friuli Venezia Giulia	61	2,4	2,4	34,1
	Liguria	60	2,3	2,3	36,4
	Emilia Romagna	223	8,7	8,7	45,1
	Toscana	167	6,5	6,5	51,6
	Umbria	31	1,2	1,2	52,8
	Marche	84	3,3	3,3	56,0
	Lazio	259	10,1	10,1	66,1
	Abruzzo	64	2,5	2,5	68,6
	Molise	18	,7	,7	69,3
	Campania	238	9,2	9,2	78,5
	Puglia	184	7,2	7,2	85,7
	Basilicata	27	1,0	1,0	86,7
	Calabria	53	2,1	2,1	88,8
	Sicilia	219	8,5	8,5	97,3
	Sardegna	69	2,7	2,7	100,0
Totale		2573	100,0	100,0	

### Livello di istruzione

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	Licenza elementare / nessun titolo	21	,8	,8	,8
	Licenza media	182	7,1	7,1	7,9
	Diploma qualifica professionale (2-3 anni)	164	6,4	6,4	14,3
	Diploma maturità professionale (compreso istituto d'arte)	134	5,2	5,2	19,5
	Diploma maturità tecnica	652	25,3	25,3	44,8
	Diploma maturità liceo classico o scientifico	284	11,0	11,0	55,8

Altro diploma maturità (istituto magistrale, liceo linguistico, liceo artistico, liceo socio-psicopedagogico)	186	7,2	7,2	63,1
Laurea Scientifica (3/4/5 anni, triennale, specialistica) (include medicina, biologia ed economia)	413	16,1	16,1	79,1
Laurea Umanistica (3/4/5 anni, triennale, specialistica) (include psicologia, sociologia e scienza politica)	293	11,4	11,4	90,5
Master/scuola di specializzazione post laurea	201	7,8	7,8	98,3
Dottorato di ricerca	43	1,7	1,7	100,0
<b>Totale</b>	<b>2573</b>	<b>100,0</b>	<b>100,0</b>	

### Stato di occupazione

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	Occupato	1730	67,2	67,2	67,2
	Non Occupato	843	32,8	32,8	100,0
	<b>Totale</b>	<b>2573</b>	<b>100,0</b>	<b>100,0</b>	

Dopodiché osserviamo anche le **statistiche descrittive relative alle nostre 10 variabili di interesse:**

### La situazione economica in Italia nell'ultimo anno è

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	Molto migliorata	29	1,1	1,1	1,1
	Abbastanza migliorata	328	12,7	12,9	14,0
	Rimasta uguale	1144	44,5	45,0	59,0
	Abbastanza peggiorata	689	26,8	27,1	86,1
	Molto peggiorata	354	13,8	13,9	100,0
	<b>Totale</b>	<b>2544</b>	<b>98,9</b>	<b>100,0</b>	
Mancante	Non saprei	29	1,1		
<b>Totale</b>		<b>2573</b>	<b>100,0</b>		

### Nei prossimi dodici mesi la situazione economica dell'Italia

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	Migliorerà molto	17	,7	,7	,7
	Migliorerà abbastanza	388	15,1	15,9	16,6
	Rimarrà uguale	1221	47,5	49,9	66,5
	Peggiorerà abbastanza	614	23,9	25,1	91,6
	Peggiorerà molto	205	8,0	8,4	100,0
	Totale	2445	95,0	100,0	
Mancante	Non saprei	128	5,0		
Totale		2573	100,0		

**Alcuni dicono che riceviamo troppi immigrati. Altri dicono che va bene come è adesso. Altri ancora dicono che potremmo facilmente accoglierne di più.**

### Lei dove collocherebbe la sua opinione?

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	1 = Riceviamo troppi immigrati	1027	39,9	41,3	41,3
	2	266	10,3	10,7	52,0
	3	298	11,6	12,0	63,9
	4	402	15,6	16,2	80,1
	5	294	11,4	11,8	91,9
	6	133	5,2	5,3	97,3
	7 = Potremmo accogliere facilmente molti più immigrati	68	2,6	2,7	100,0
	Totale	2488	96,7	100,0	
Mancante	Non saprei	85	3,3		
Totale		2573	100,0		

### Secondo Lei in generale gli immigrati sono un bene o un male per l'economia italiana?

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	Un bene per l'economia italiana	90	3,5	3,7	3,7
	1	58	2,3	2,4	6,0

2		121	4,7	4,9	11,0
3		173	6,7	7,1	18,0
4		187	7,3	7,6	25,7
5		429	16,7	17,5	43,2
6		202	7,9	8,2	51,4
7		276	10,7	11,3	62,7
8		251	9,8	10,2	73,0
9		120	4,7	4,9	77,9
Un male per l'economia italiana		542	21,1	22,1	100,0
Totale		2449	95,2	100,0	
Mancante	Non saprei	124	4,8		
Totale		2573	100,0		

**Secondo Lei in generale gli immigrati rappresentano una minaccia o un arricchimento per la cultura italiana?**

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	Una minaccia per la cultura italiana	493	19,2	19,8	19,8
	1	130	5,1	5,2	25,1
	2	173	6,7	7,0	32,0
	3	186	7,2	7,5	39,5
	4	181	7,0	7,3	46,8
	5	374	14,5	15,0	61,8
	6	271	10,5	10,9	72,7
	7	255	9,9	10,3	83,0
	8	175	6,8	7,0	90,0
	9	91	3,5	3,7	93,7
	Una arricchimento per la cultura italiana	157	6,1	6,3	100,0
	Totale	2486	96,6	100,0	
Mancante	Non saprei	87	3,4		
Totale		2573	100,0		

**Secondo Lei il fatto che l'Italia appartenga all'Unione Europea è un bene, un male, o né un bene né un male?**

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
--	--	-----------	-------------	--------------------	------------------------

Valido	Un bene	1096	42,6	45,6	45,6
	Né un bene né un male	785	30,5	32,6	78,2
	Un male	524	20,4	21,8	100,0
	Totale	2405	93,5	100,0	
Mancante	Non saprei	168	6,5		
Totale		2573	100,0		

### In generale, Lei pensa che avere l'Euro è un bene o un male per l'Italia?

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	Un bene	982	38,2	40,5	40,5
	Né un bene né un male	585	22,7	24,2	64,7
	Un male	855	33,2	35,3	100,0
	Totale	2422	94,1	100,0	
Mancante	Non saprei	151	5,9		
Totale		2573	100,0		

### In generale, quanto si interessa di politica?

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	Molto	365	14,2	14,5	14,5
	Abbastanza	1123	43,6	44,7	59,2
	Poco	832	32,3	33,1	92,3
	Per niente	194	7,5	7,7	100,0
	Totale	2514	97,7	100,0	
Mancante	Non saprei/Preferisco non rispondere	59	2,3		
Totale		2573	100,0		

### Quanto è interessato alla campagna elettorale in corso?

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	Molto interessato	457	17,8	18,2	18,2
	Abbastanza interessato	1038	40,3	41,3	59,5
	Poco interessato	751	29,2	29,9	89,4
	Per niente interessato	266	10,3	10,6	100,0
	Totale	2512	97,6	100,0	
Mancante	Non saprei	61	2,4		
Totale		2573	100,0		

### Nell'ultima settimana, quanto spesso le è capitato di parlare di politica?

		Frequenza	Percentuale	Percentuale valida	Percentuale cumulativa
Valido	Mai	373	14,5	15,3	15,3
	1 giorno	419	16,3	17,2	32,5
	2 giorni	530	20,6	21,8	54,3
	3-4 giorni	523	20,3	21,5	75,8
	5-6 giorni	163	6,3	6,7	82,5
	Ogni giorno	426	16,6	17,5	100,0
	Totale	2434	94,6	100,0	
Mancante	Non saprei	139	5,4		
Totale		2573	100,0		

Tutte le 10 variabili di interesse sono misurate su scala Likert. Le risposte "Non saprei" sono state correttamente impostate come mancanti.

Non risulta quindi necessario provvedere all'eliminazione di alcuna unità statistica: procederemo effettuando l'analisi sui 2573 soggetti rilevati.

Andiamo innanzitutto ad osservare la **matrice di correlazione dei 10 item**.

### Correlazioni

		ECO1	ECO2	IMM1	IMM2	IMM3	EURO1	EURO2	POL1	POL2	POL3
ECO1	Correlazione di Pearson	1	,643**	-,254**	,288**	-,249**	,348**	,378**	,105**	,074**	,005
	Sign. (a due code)		,000	,000	,000	,000	,000	,000	,000	,000	,812
	N	2544	2435	2469	2433	2468	2390	2405	2491	2494	2416
ECO2	Correlazione di Pearson	,643**	1	-,199**	,236**	-,209**	,318**	,319**	,109**	,125**	-,039
	Sign. (a due code)	,000		,000	,000	,000	,000	,000	,000	,000	,062
	N	2435	2445	2376	2344	2376	2302	2313	2395	2396	2323
IMM1	Correlazione di Pearson	-,254**	-,199**	1	-,680**	,720**	-,394**	-,404**	-,107**	-,046*	,061**
	Sign. (a due code)	,000	,000		,000	,000	,000	,000	,000	,022	,003
	N	2469	2376	2488	2424	2448	2349	2362	2442	2441	2369
IMM2	Correlazione di Pearson	,288**	,236**	-,680**	1	-,694**	,423**	,420**	,095**	,018	-,043*
	Sign. (a due code)	,000	,000	,000		,000	,000	,000	,000	,379	,040
	N	2469	2376	2488	2424	2448	2349	2362	2442	2441	2369

	N	2433	2344	2424	2449	2429	2325	2337	2405	2405	2338
IMM3	Correlazione di Pearson	-,249**	-,209**	,720**	-,694**	1	-,421**	-,402**	-,079**	-,025	,012
	Sign. (a due code)	,000	,000	,000	,000		,000	,000	,000	,226	,555
	N	2468	2376	2448	2429	2486	2355	2367	2445	2441	2373
EURO1	Correlazione di Pearson	,348**	,318**	-,394**	,423**	-,421**	1	,750**	,055**	,055**	,000
	Sign. (a due code)	,000	,000	,000	,000	,000		,000	,008	,007	,995
	N	2390	2302	2349	2325	2355	2405	2339	2368	2368	2298
EURO2	Correlazione di Pearson	,378**	,319**	-,404**	,420**	-,402**	,750**	1	,124**	,085**	-,062**
	Sign. (a due code)	,000	,000	,000	,000	,000	,000		,000	,000	,003
	N	2405	2313	2362	2337	2367	2339	2422	2386	2385	2314
POL1	Correlazione di Pearson	,105**	,109**	-,107**	,095**	-,079**	,055**	,124**	1	,695**	-,585**
	Sign. (a due code)	,000	,000	,000	,000	,000	,008	,000		,000	,000
	N	2491	2395	2442	2405	2445	2368	2386	2514	2497	2410
POL2	Correlazione di Pearson	,074**	,125**	-,046*	,018	-,025	,055**	,085**	,695**	1	-,552**
	Sign. (a due code)	,000	,000	,022	,379	,226	,007	,000	,000		,000
	N	2494	2396	2441	2405	2441	2368	2385	2497	2512	2410
POL3	Correlazione di Pearson	,005	-,039	,061**	-,043*	,012	,000	-,062**	-,585**	-,552**	1
	Sign. (a due code)	,812	,062	,003	,040	,555	,995	,003	,000	,000	
	N	2416	2323	2369	2338	2373	2298	2314	2410	2410	2434

\*\* . La correlazione è significativa a livello 0,01 (a due code).

\* . La correlazione è significativa a livello 0,05 (a due code).

Notiamo come vi siano moltissimi item significativamente correlati tra di loro.

Ciò è anche normale essendo gli item divisi in 4 tematiche specifiche.

È dunque necessario effettuare un'analisi delle componenti principali.

## ANALISI DELLE COMPONENTI PRINCIPALI (ACP)

Per l'analisi in componenti principali inseriremo tutti e 10 gli item di interesse, sapendo già che questi dovrebbero andare a costituire quattro componenti (in quanto il questionario è stato strutturato in questa maniera): si tratterà perciò di un'analisi fattoriale confermativa.

E' stata necessaria una rotazione per rendere più chiaro il pattern dei dati: la rotazione scelta è stata Oblimin in quanto quest'ultima non impone l'ortogonalità delle componenti.

### Test di KMO e Bartlett

Misura di Kaiser-Meyer-Olkin di adeguatezza del campionamento.		,755
Test della sfericità di Bartlett	Appross. Chi-quadrato	10491,150
	gl	45
	Sign.	,000

Il valore dell'indice KMO è buono ( $0.755 > 0.6$  ovvero della soglia di accettabilità minima). Inoltre il test di sfericità di Bartlett rifiuta l'ipotesi nulla di sfericità consentendoci così di procedere con la nostra analisi.

### Comunalità

	Iniziale	Estrazione
La situazione economica in Italia nell'ultimo anno è	1,000	,816
Nei prossimi dodici mesi la situazione economica dell'Italia	1,000	,829
Alcuni dicono che riceviamo troppi immigrati. Altri dicono che va bene come è adesso. Altri ancora dicono che potremmo facilmente accoglierne di più. Lei dove collocherebbe la sua opinione?	1,000	,806
Secondo Lei in generale gli immigrati sono un bene o un male per l'economia italiana?	1,000	,779
Secondo Lei in generale gli immigrati rappresentano una minaccia o un arricchimento per la cultura italiana?	1,000	,813
Secondo Lei il fatto che l'Italia appartenga all'Unione Europea è un bene, un male, o né un bene né un male?	1,000	,877
In generale, Lei pensa che avere l'Euro è un bene o un male per l'Italia?	1,000	,875
In generale, quanto si interessa di politica?	1,000	,789
Quanto è interessato alla campagna elettorale in corso?	1,000	,766
Nell'ultima settimana, quanto spesso le è capitato di parlare di politica?	1,000	,682

Metodo di estrazione: Analisi dei componenti principali.

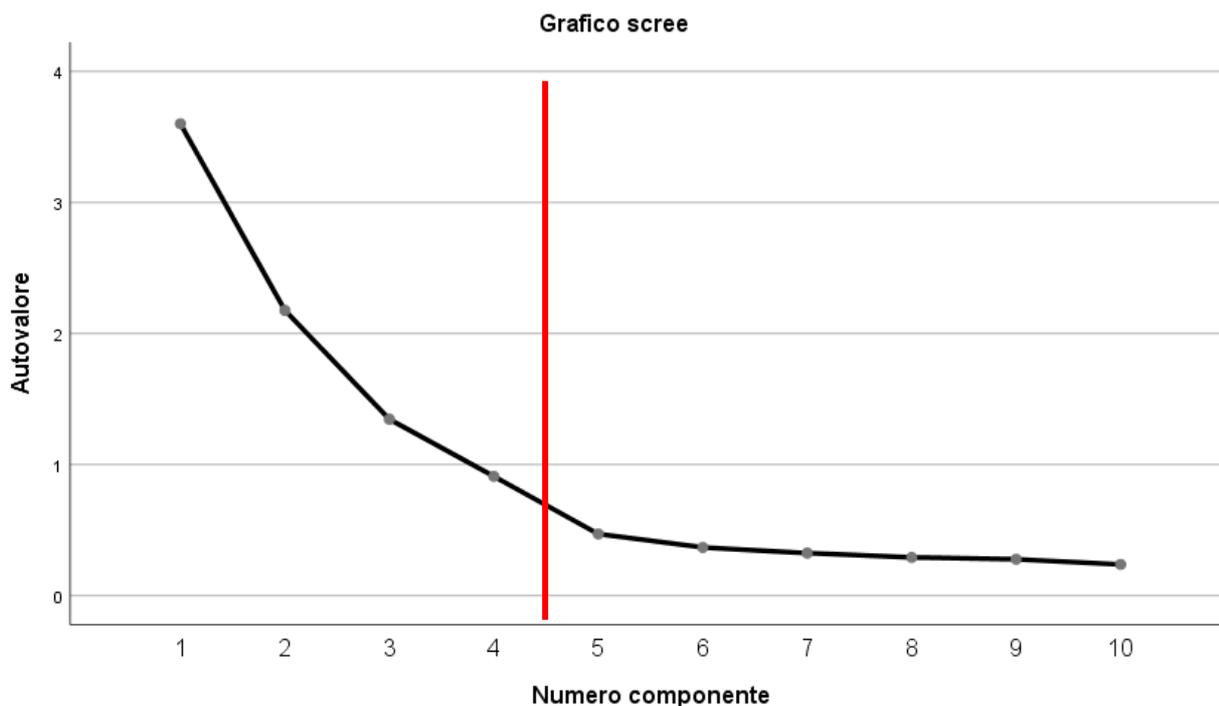
Tutti i singoli item risultano adeguatamente spiegati dalle componenti estratte, infatti tutti i valori delle comunalità sono piuttosto alti (il più basso si ha per l'item "Nell'ultima settimana, quanto spesso le è capitato di parlare di politica?" pari a 0.68).

### Varianza totale spiegata

Componente	Autovalori iniziali			Caricamenti somme dei quadrati di estrazione			Caricamenti somme dei quadrati di rotazione <sup>a</sup>
	Totale	% di varianza	% cumulativa	Totale	% di varianza	% cumulativa	
1	3,601	36,007	36,007	3,601	36,007	36,007	2,969
2	2,176	21,764	57,771	2,176	21,764	57,771	2,261
3	1,346	13,458	71,228	1,346	13,458	71,228	2,146
4	,909	9,094	80,323	,909	9,094	80,323	2,607
5	,470	4,701	85,024				
6	,367	3,673	88,697				
7	,324	3,242	91,939				
8	,292	2,917	94,856				
9	,277	2,769	97,625				
10	,238	2,375	100,000				

Metodo di estrazione: Analisi dei componenti principali.

a. Quando i componenti sono correlati, i caricamenti delle somme dei quadrati non possono essere aggiunti per ottenere una varianza totale.



Sono state estratte 4 componenti così come ipotizzato in partenza. La quarta componente avrebbe un autovalore leggermente inferiore alla soglia (ovvero leggermente inferiore ad 1) ma viene comunque estratta: anche lo scree-plot ci conferma la correttezza dell'estrazione della quarta componente, infatti il “gomito” si ha in corrispondenza della quinta componente. La varianza complessivamente spiegata dalle quattro componenti è molto alta, pari circa all'80% della variabilità complessiva dei 10 item originali.

### Matrice dei componenti<sup>a</sup>

	Componente			
	1	2	3	4
Secondo Lei in generale gli immigrati sono un bene o un male per l'economia italiana?	,766		-,354	
Secondo Lei in generale gli immigrati rappresentano una minaccia o un arricchimento per la cultura italiana?	-,757		,408	
Alcuni dicono che riceviamo troppi immigrati. Altri dicono che va bene come è adesso. Altri ancora dicono che potremmo facilmente accoglierne di più. Lei dove collocherebbe la sua opinione?	-,754		,417	
In generale, Lei pensa che avere l'Euro è un bene o un male per l'Italia?	,744			,541
Secondo Lei il fatto che l'Italia appartenga all'Unione Europea è un bene, un male, o né un bene né un male?	,731			,550
Quanto è interessato alla campagna elettorale in corso?		,848		
In generale, quanto si interessa di politica?		,843		
Nell'ultima settimana, quanto spesso le è capitato di parlare di politica?		-,798		
Nei prossimi dodici mesi la situazione economica dell'Italia	,538		,659	-,324
La situazione economica in Italia nell'ultimo anno è	,587		,619	

Metodo di estrazione: Analisi dei componenti principali.

a. 4 componenti estratti.

La matrice delle componenti risulta di difficile comprensione, nonostante siano stati nascosti i coefficienti piccoli (<0,3): risulta dunque indispensabile la rotazione Oblimin.

### Matrice del modello<sup>a</sup>

	Componente			
	1	2	3	4
Alcuni dicono che riceviamo troppi immigrati. Altri dicono che va bene come è adesso. Altri ancora dicono che potremmo facilmente accoglierne di più. Lei dove collocherebbe la sua opinione?	-,907			
Secondo Lei in generale gli immigrati rappresentano una minaccia o un arricchimento per la cultura italiana?	-,903			
Secondo Lei in generale gli immigrati sono un bene o un male per l'economia italiana?	,863			
In generale, quanto si interessa di politica?		,881		
Quanto è interessato alla campagna elettorale in corso?		,870		
Nell'ultima settimana, quanto spesso le è capitato di parlare di politica?		-,829		
Nei prossimi dodici mesi la situazione economica dell'Italia			,922	
La situazione economica in Italia nell'ultimo anno è			,883	
Secondo Lei il fatto che l'Italia appartenga all'Unione Europea è un bene, un male, o né un bene né un male?				,939
In generale, Lei pensa che avere l'Euro è un bene o un male per l'Italia?				,930

Metodo di estrazione: Analisi dei componenti principali.

Metodo di rotazione: Oblimin con normalizzazione Kaiser.

a. Convergenza per la rotazione eseguita in 5 iterazioni.

Avendo effettuato la rotazione è ora molto più facile comprendere la struttura delle componenti. Le 4 componenti ipotizzate sono perfettamente rappresentate nella matrice del modello e vanno a rappresentare rispettivamente:

Componente 1: Contrarietà all'immigrazione

Componente 2: Disinteresse per la politica

Componente 3: Pessimismo per l'economia

Componente 4: Antieuropeismo

Nota bene: alcuni factor loadings hanno segno negativo poiché alcuni quesiti sono stati posti con polarità inversa.

### Matrice di correlazione dei componenti

Componente	1	2	3	4
1	1,000	,066	,284	,487
2	,066	1,000	,090	,073
3	,284	,090	1,000	,396
4	,487	,073	,396	1,000

Metodo di estrazione: Analisi dei componenti principali.

Metodo di rotazione: Oblimin con normalizzazione Kaiser.

Dalla matrice di correlazione delle componenti notiamo infine come alcune delle componenti risultino tra di loro notevolmente correlate, ad ulteriore conferma della necessità di una rotazione NON ortogonale, come ad esempio quella da noi effettuata (Oblimin).

A questo punto si procede con la cluster analysis (gerarchica e non gerarchica) effettuata sulle quattro componenti estratte al fine di individuare dei cluster omogenei al loro interno e diversi tra di loro relativamente alle quattro tematiche di interesse dell'indagine.

## CLUSTER ANALYSIS GERARCHICA

Per effettuare la prima cluster analysis viene utilizzata la tecnica gerarchica del legame medio tra gruppi.

### Riepilogo elaborazione casi<sup>a,b</sup>

Valido		Casi Mancante		Totale	
N	Percentuale	N	Percentuale	N	Percentuale
2033	79,0	540	21,0	2573	100,0

a. Distanza euclidea al quadrato utilizzata

b. Legame medio (tra gruppi)

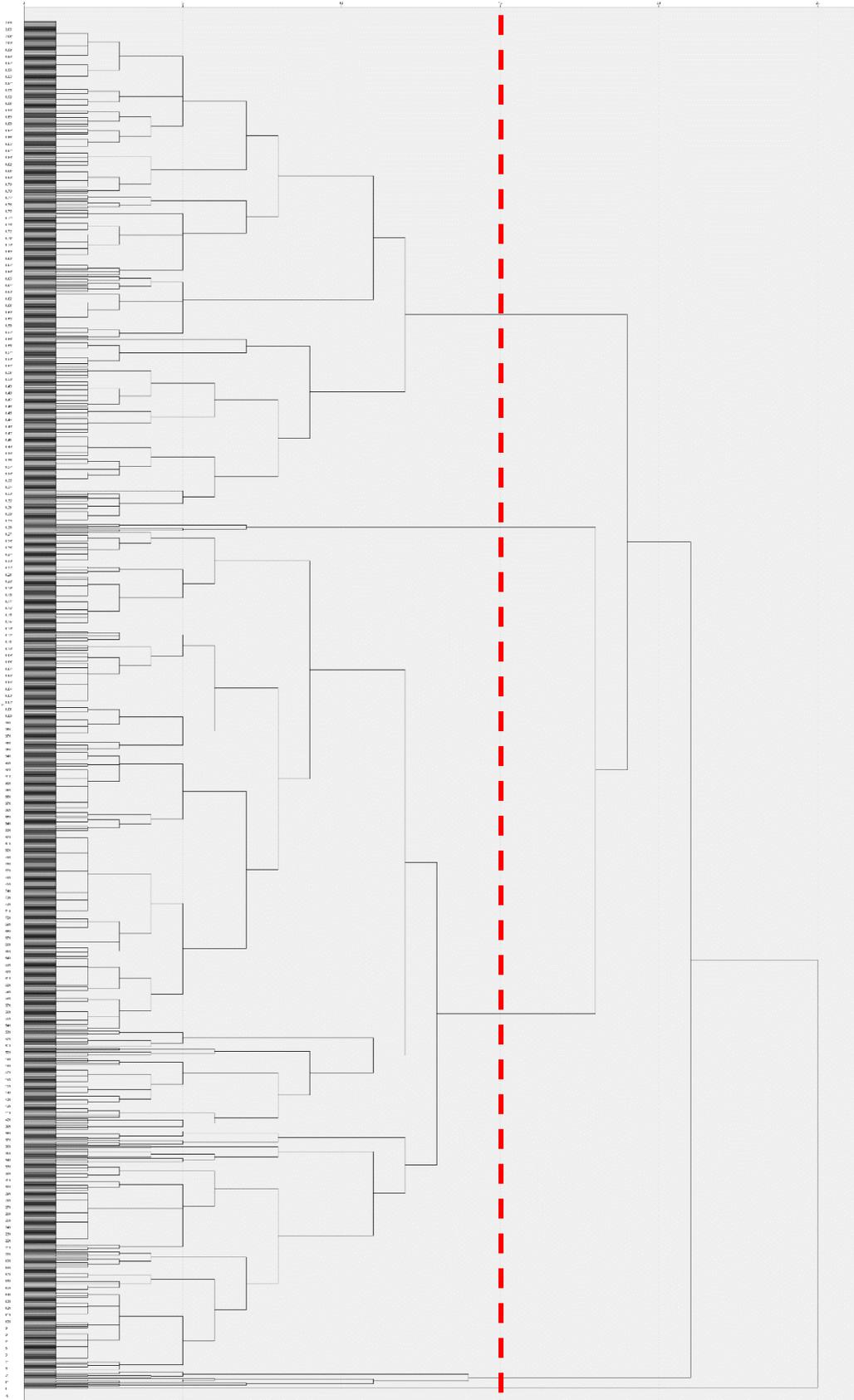
Osserviamo come solamente 2033 dei 2573 soggetti vengono utilizzati poiché ci sono 540 soggetti per i quali, con la precedente analisi in componenti principali, non è stato possibile calcolare i factor score a causa delle risposte mancanti ai singoli item.

Evitiamo di incollare interamente la tabella con la pianificazione di agglomerazione in quanto contiene 2032 righe ed è dunque esageratamente grande. Riportiamo solamente le prime e le ultime righe che contengono le agglomerazioni di nostro interesse utili ad individuare il “salto” per decidere il numero ottimale di cluster.

### Pianificazione di agglomerazione

Stadio	Combinato in cluster		Coefficienti	Stadio prima apparizione cluster		Stadio successivo
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
1	2118	2536	,000	0	0	21
2	2193	2529	,000	0	0	15
3	1825	2500	,000	0	0	298
4	1869	2495	,000	0	0	35
...						
2022	49	226	6,072	2009	1993	2028
2023	1	102	6,186	2010	1990	2026
2024	2	25	6,522	2018	2021	2027
2025	19	29	6,576	2020	2014	2027
2026	1	13	6,669	2023	2017	2030
2027	2	19	7,177	2024	2025	2029
<b>2028</b>	<b>49</b>	<b>474</b>	<b>7,865</b>	<b>2022</b>	<b>1987</b>	<b>2031</b>
2029	2	466	10,065	2027	2003	2030
2030	1	2	10,614	2026	2029	2031
2031	1	49	11,535	2030	2028	2032
2032	1	506	14,257	2031	1786	0

**Dendrogramma che utilizza il legame medio (tra gruppi)**  
Combinazione cluster distanza con scala modificata



Notiamo come il “salto” si abbia in corrispondenza di 5 gruppi, così come evidente anche dal piano di agglomerazione, dove il “salto” nei coefficienti si ha alla quint’ultima riga.

## CLUSTER ANALYSIS NON GERARCHICA (K-MEANS)

Così come suggerito nelle linee guida andiamo ora a fare una cluster non gerarchica con la tecnica delle k-medie impostando, impostando 5 come numero di cluster, che è risultato essere il numero ottimale di cluster dal dendrogramma della cluster gerarchica.

### Centri cluster iniziali

	Cluster				
	1	2	3	4	5
Contrarietà all’immigrazione	-1,49286	-2,40713	-,87827	,75954	1,35174
Disinteresse per la politica	-1,79783	-1,39986	2,06733	2,09904	-1,83797
Pessimismo per l’economia	-2,90528	2,15600	2,12653	-1,60760	,84967
Antieuropeismo	-1,10793	-1,08499	1,48644	-1,10032	1,49073

### Cronologia delle iterazioni<sup>a</sup>

Iterazione	Modifica nei centri del cluster				
	1	2	3	4	5
1	2,134	2,119	1,887	2,027	1,628
2	,130	,155	,383	,086	,168
3	,042	,105	,206	,038	,076
4	,016	,090	,086	,023	,040
5	,022	,068	,051	,021	,025
6	,007	,019	,019	,016	,008
7	,002	,008	,000	,003	,000
8	,003	,007	,000	,006	,000
9	,005	,009	,000	,005	,003
10	,000	,005	,009	,011	,003

a. Le iterazioni sono state arrestate perché è stato eseguito il numero massimo di iterazioni. Convergenza delle iterazioni non riuscita. La modifica di coordinata assoluta massima per un centro è ,010. L’iterazione corrente è 10. La distanza minima tra i centri iniziali è 4,580.

### Centri finali del cluster

	Cluster				
	1	2	3	4	5
Contrarietà all'immigrazione	-,72920	-1,09297	,69533	-,00410	,85553
Disinteresse per la politica	-,75839	-,33959	1,01868	,63083	-,80348
Pessimismo per l'economia	-1,02267	,61947	,75147	-,41911	,43636
Antieuropeismo	-,82889	-,51802	,98200	-,54313	,92267

Essendo quattro fattori, così come suggerito nelle linee guida, non andiamo ad interpretare i centroidi: andremo alla fine a valutare la distribuzione delle 10 variabili originali sui 5 cluster creati.

### ANOVA

	Cluster		Errore		F	Sign.
	Media	gl	Media	gl		
	quadratica		quadratica			
Contrarietà all'immigrazione	272,079	4	,486	2028	559,420	,000
Disinteresse per la politica	281,247	4	,392	2028	716,715	,000
Pessimismo per l'economia	237,799	4	,561	2028	423,899	,000
Antieuropeismo	315,058	4	,396	2028	794,814	,000

I test F devono essere utilizzati solo per scopi descrittivi perché i cluster sono stati scelti per massimizzare le differenze tra i casi in cluster differenti. I livelli di significatività osservati non sono corretti per tale motivo e, pertanto, non possono essere interpretati come test dell'ipotesi che le medie dei cluster siano uguali.

Il test ANOVA che ci indica quali variabili hanno maggiormente contribuito all'individuazione dei cluster, ci conferma che tutte e quattro le componenti hanno contribuito in maniera determinante ( $p < 0,001$  per tutte e quattro le componenti).

### Numero di casi in ciascun cluster

Cluster	1	461,000
	2	279,000
	3	331,000
	4	484,000
	5	478,000
Valido		2033,000
Mancante		540,000

Possiamo notare come tutti e cinque i cluster contengano un numero consistente di soggetti. Il cluster più piccolo è il 2 (279 soggetti) mentre il più grande è il 4 (484 soggetti).

Così come indicato nelle linee guida andremo a valutare e a definire i cluster non sui centroidi delle componenti, bensì sulle variabili originali (10 item di partenza).

Salviamo quindi una nuova variabile per il cluster di appartenenza e andiamo ad osservare le caratteristiche dei cluster.

- Pessimismo sull'economia

		Numero cluster del caso				
		1	2	3	4	5
		% di casi per colonna	% di casi per colonna	% di casi per colonna	% di casi per colonna	% di casi per colonna
La situazione economica in Italia nell'ultimo anno è	Molto migliorata	3,3%	0,0%	0,3%	1,0%	0,6%
	Abbastanza migliorata	43,2%	0,4%	0,9%	12,6%	2,1%
	Rimasta uguale	51,2%	26,9%	26,9%	68,6%	36,6%
	Abbastanza peggiorata	2,4%	55,2%	39,9%	17,1%	34,5%
	Molto peggiorata	0,0%	17,6%	32,0%	0,6%	26,2%
Nei prossimi dodici mesi la situazione economica dell'Italia	Migliorerà molto	2,8%	0,0%	0,0%	0,2%	0,2%
	Migliorerà abbastanza	49,0%	0,4%	0,3%	16,3%	9,2%
	Rimarrà uguale	46,2%	39,1%	37,8%	72,7%	42,9%
	Peggiorerà abbastanza	2,0%	49,5%	41,1%	10,7%	33,1%
	Peggiorerà molto	0,0%	11,1%	20,8%	0,0%	14,6%

Relativamente alla visione economica dell'Italia notiamo come nel cluster 1 vi siano i soggetti più ottimisti relativamente alla situazione economica dell'Italia, mentre nel cluster 2, 3 e 5 i più pessimisti. Il cluster 4 invece contiene principalmente soggetti che ritengono che la situazione sia inalterata.

- Visione dell'immigrazione

		Numero cluster del caso				
		1	2	3	4	5
		% di casi per colonna				
Alcuni dicono che riceviamo troppi immigrati. Altri dicono che va bene come è adesso. Altri ancora dicono che potremmo facilmente accoglierne di più. Lei dove collocherebbe la sua opinione?	1 = Riceviamo troppi immigrati	12,1%	5,4%	70,7%	34,7%	76,8%
	2	6,9%	3,6%	11,2%	15,9%	10,9%
	3	12,1%	11,1%	8,5%	21,5%	6,9%
	4	26,0%	24,0%	5,7%	18,6%	3,8%
	5	26,0%	25,4%	3,6%	7,0%	1,3%
	6	13,0%	16,5%	0,3%	1,9%	0,4%
	7 = Potremmo accogliere facilmente molti più immigrati	3,7%	14,0%	0,0%	0,4%	0,0%
Secondo Lei in generale gli immigrati sono un bene o un male per l'economia italiana?	Un bene per l'economia italiana	7,2%	13,6%	0,0%	0,6%	0,2%
	1	5,0%	7,9%	0,3%	0,8%	0,2%
	2	9,5%	12,2%	1,5%	3,1%	1,0%
	3	13,2%	14,7%	1,8%	6,0%	1,7%
	4	13,0%	9,0%	2,7%	11,4%	1,3%
	5	22,8%	20,4%	12,1%	22,9%	6,7%
	6	9,5%	8,6%	5,7%	12,8%	3,8%
	7	9,8%	9,0%	10,0%	15,3%	9,6%
	8	5,2%	3,9%	16,9%	11,8%	14,2%
	9	2,2%	0,4%	5,4%	6,4%	9,4%
Un male per l'economia italiana	2,6%	0,4%	43,5%	8,9%	51,9%	
Secondo Lei in generale gli immigrati rappresentano una minaccia o un arricchimento per la cultura italiana?	Una minaccia per la cultura italiana	2,8%	0,7%	40,5%	9,7%	45,8%
	1	2,0%	0,0%	8,5%	5,4%	9,8%
	2	3,7%	1,1%	8,2%	8,3%	11,7%
	3	5,2%	2,9%	10,6%	8,7%	9,0%
	4	6,1%	4,7%	5,7%	10,3%	6,3%
	5	13,7%	14,0%	15,1%	20,0%	7,3%
	6	17,1%	10,8%	5,1%	14,0%	5,2%
	7	18,0%	16,8%	2,7%	13,6%	2,1%
	8	12,8%	17,2%	2,4%	5,2%	2,1%
	9	8,7%	8,2%	0,6%	2,1%	0,4%
Un arricchimento per la cultura italiana	10,0%	23,7%	0,6%	2,7%	0,2%	

Relativamente alla visione sull'immigrazione notiamo come nel cluster 1 e 2 vi siano i soggetti più favorevoli all'immigrazione, mentre nel cluster 3 e 5 i più contrari. Il cluster 4 anche in questo caso contiene principalmente soggetti con posizioni intermedie.

- Europeismo

		Numero cluster del caso				
		1	2	3	4	5
		% di casi per colonna	% di casi per colonna	% di casi per colonna	% di casi per colonna	% di casi per colonna
Secondo Lei il fatto che l'Italia appartenga all'Unione Europea è un bene, un male, o né un bene né un male?	Un bene	83,7%	67,7%	5,4%	65,5%	6,3%
	Né un bene né un male	15,4%	26,2%	43,8%	32,9%	39,7%
	Un male	0,9%	6,1%	50,8%	1,7%	54,0%
In generale, Lei pensa che avere l'Euro è un bene o un male per l'Italia?	Un bene	80,0%	60,9%	1,5%	57,2%	6,3%
	Né un bene né un male	15,0%	26,5%	18,7%	34,3%	22,0%
	Un male	5,0%	12,5%	79,8%	8,5%	71,8%

Relativamente alla visione sull'Europa notiamo come nel cluster 1, 2 e 4 vi siano i soggetti più europeisti, mentre nel cluster 3 e 5 i più antieuropeisti.

- Interesse per la politica

		Numero cluster del caso				
		1	2	3	4	5
		% di casi per colonna	% di casi per colonna	% di casi per colonna	% di casi per colonna	% di casi per colonna
In generale, quanto si interessa di politica?	Molto	31,9%	17,2%	0,3%	0,4%	27,6%
	Abbastanza	63,6%	59,1%	11,5%	29,8%	65,7%
	Poco	4,3%	22,6%	64,7%	61,8%	6,7%
	Per niente	0,2%	1,1%	23,6%	8,1%	0,0%
Quanto è interessato alla campagna elettorale in corso?	Molto interessato	36,4%	19,4%	0,3%	1,0%	36,6%
	Abbastanza interessato	54,7%	50,5%	12,7%	33,1%	57,9%
	Poco interessato	8,5%	25,1%	56,8%	53,3%	5,2%
	Per niente interessato	0,4%	5,0%	30,2%	12,6%	0,2%
Nell'ultima settimana, quanto spesso le è capitato di parlare di politica?	Mai	2,2%	4,3%	32,0%	25,4%	0,4%
	1 giorno	4,3%	12,9%	35,6%	28,7%	6,1%
	2 giorni	21,0%	21,1%	21,1%	28,3%	18,8%
	3-4 giorni	34,1%	27,6%	8,2%	14,0%	28,0%
	5-6 giorni	10,6%	10,8%	1,2%	2,1%	11,5%
	Ogni giorno	27,8%	23,3%	1,8%	1,4%	35,1%

Relativamente all'interesse per la politica notiamo come nel cluster 1 e 5 vi siano i soggetti più interessati, mentre nel cluster 3 e 4 i più disinteressati. Il cluster 2 risulta invece principalmente composto da soggetti mediamente interessati alla politica.

## CONCLUSIONI

Volendo dunque riassumere abbiamo ottenuto i seguenti cluster:

CLUSTER N°1 Costituito da 461 soggetti (22.6% dei casi validi), contiene i soggetti ottimisti a livello economico, favorevoli all'immigrazione, totalmente europeisti e interessati alla politica.

→ EUROPEISTI OTTIMISTI

CLUSTER N°2 Costituito da 279 soggetti (13.7% dei casi validi), contiene i soggetti pessimisti a livello economico, favorevoli all'immigrazione, europeisti e abbastanza interessati alla politica.

→ EUROPEISTI PESSIMISTI

CLUSTER N°3 Costituito da 331 soggetti (16.3% dei casi validi), contiene i soggetti pessimisti a livello economico, contrari all'immigrazione, totalmente antieuropeisti e molto disinteressati alla politica.

→ POPULISTI DISINTERESSATI

CLUSTER N°4 Costituito da 484 soggetti (23.8% dei casi validi), contiene i soggetti né ottimisti né pessimisti a livello economico, con posizioni intermedie sull'immigrazione, abbastanza europeisti e abbastanza disinteressati alla politica.

→ MODERATI

CLUSTER N°5 Costituito da 478 soggetti (23.5% dei casi validi), contiene i soggetti pessimisti a livello economico, molto contrari all'immigrazione, totalmente antieuropeisti e molto interessati alla politica.

→ POPULISTI INTERESSATI

## SITOGRAFIA

Sito dal quale ho reperito il dataset

<http://www.itanes.org/>

Esempi di applicazione della Cluster Analysis e ACP reperiti nella pagina e – learning del corso  
“*Analisi quantitative di Mercato*”

<https://elearning.unimib.it/course/view.php?id=23581>

Manuale SPSS Statistics Base 25, in

[ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/25.0/it/client/Manuals/IBM\\_SPSS\\_Statistics\\_Base.pdf](ftp://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/25.0/it/client/Manuals/IBM_SPSS_Statistics_Base.pdf)