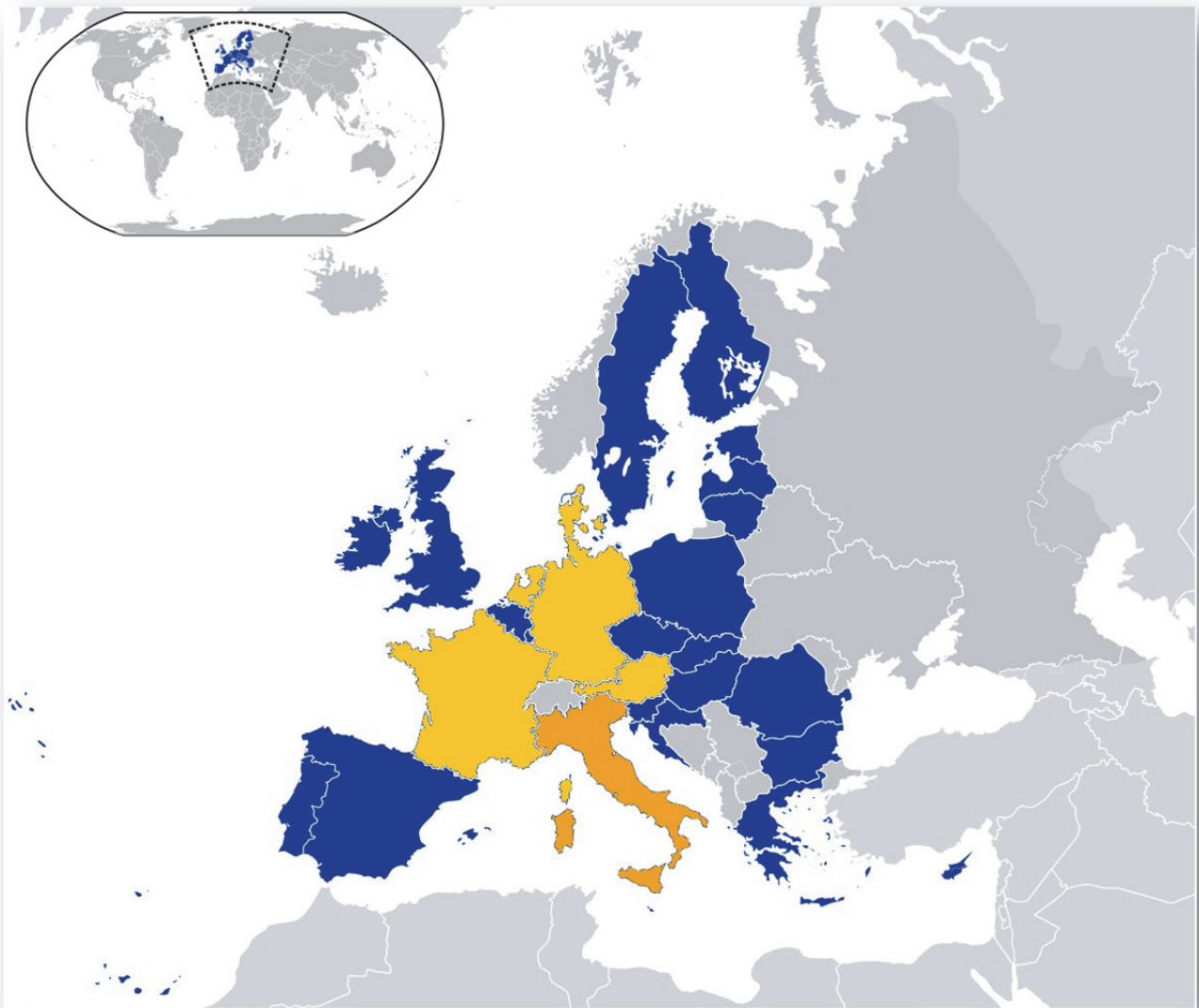


Indicatori socio – economici: un confronto fra il Veneto e altre regioni dell'UE.



Consiglio regionale del Veneto

Servizio Studi documentazione biblioteca

Dirigente Capo: Claudio Giulio Rizzato



Osservatorio della spesa e delle politiche pubbliche

Realizzato da: Ilaria Capuzzo, Ilaria Destro (stagiste presso l'Ufficio per l'analisi della spesa regionale)

Documento ultimato a settembre 2013

Indicatori socio – economici
(un confronto fra il Veneto e altre regioni NUTS 2)

INDICE

1. Economia.....	pag. 1
1.1. Pil a parità di potere d'acquisto (PPS).....	pag. 1
1.2. Pil a parità di potere d'acquisto (PPS) pro capite.....	pag. 2
1.3. Reddito disponibile pro capite.....	pag. 2
2. Demografia e dipendenza.....	pag. 4
2.1. Popolazione, area e densità abitativa.....	pag. 4
2.2. Popolazione per fasce d'età.....	pag. 6
2.3. Tassi di dipendenza.....	pag. 9
2.4. Tasso di natalità.....	pag. 11
2.5. Tasso di fecondità.....	pag. 12
2.6. Tasso di mortalità.....	pag. 12
2.7. Aspettativa di vita maschile e femminile.....	pag. 13
3. Lavoro.....	pag. 15
3.1. Tasso di occupazione.....	pag. 15
3.2. Tasso di occupazione femminile.....	pag. 26
3.3. Tasso di disoccupazione.....	pag. 17
3.4. Tasso di disoccupazione femminile.....	pag. 18
3.5. Tasso di disoccupazione giovanile.....	pag. 19
3.6. Disoccupazione di lunga durata.....	pag. 20
3.7. Posizioni apicali, confronto fra generi.....	pag. 20
4. Capitale umano.....	pag. 22
4.1. Studenti.....	pag. 23
4.2. Studenti di scuola secondaria e post-secondaria non terziaria.....	pag. 23
4.3. Studenti di scuola terziaria.....	pag. 24
4.4. Educazione terziaria, fascia d'età: 25-64 anni.....	pag. 25
4.5. Educazione terziaria, fascia d'età: 30-34 anni.....	pag. 26
4.6. Tasso NEET (<i>not in employment, education or training</i>).....	pag. 27

5. Ricerca e sviluppo.....pag. 29

5.1. Spesa R&S in percentuale al Pil.....	pag. 29
5.2. Spesa R&S pro capite.....	pag. 30
5.3. Ricercatori e personale R&S.....	pag. 31
5.4. Ricercatori.....	pag. 31
5.5. Occupazione nei settori ad alta tecnologia.....	pag. 32
5.6. Brevetti presentati all'EPO (<i>European Patent Office</i>).....	pag. 33

6. Altre informazioni..... pag. 35

6.1. Persone a rischio povertà.....	pag. 35
6.2. Accesso ad internet.....	pag. 36
6.3. Utilizzo di internet.....	pag. 36
6.4. Estraneità ad internet.....	pag. 37
6.5. Inglese.....	pag. 38

7. Altre statistiche di confronto (rispetto alle regioni italiane)..... pag. 40

7.1. Cittadini e minori stranieri.....	pag. 40
7.2. Situazione imprenditoriale: numero di imprese (per 1.000 abitanti) e numero medio addetti per impresa.....	pag. 41
7.3. Natimortalità delle imprese.....	pag. 43
7.4. Esportazioni.....	pag. 44

8. Allegato tecnico..... pag. 47

GLOSSARIO

Per una maggior comprensione dell'elaborato, si presenta la seguente tabella nella quale si riportano i nomi delle aree geografiche che sono coinvolte nella comparazione.

<i>Regioni o aree geografiche (denominazione Eurostat)</i>	<i>Stato</i>	<i>Regioni o aree geografiche (denominazione italiana)</i>
UE27	UE	Unione Europea a 27 paesi
Bourgogne	FRANCIA	Borgogna
Centre (FR)	FRANCIA	Centro
Champagne – Ardenne	FRANCIA	Sciampagna – Ardena
Friuli – Venezia Giulia	ITALIA	Friuli – Venezia Giulia
Italy	ITALIA	Italia
Kärnten	AUSTRIA	Carinzia
Midtjylland	DANIMARCA	Jutland centrale
Noord – Nederland	PAESI BASSI	Paesi Bassi del Nord
Oberösterreich	AUSTRIA	Alta Austria
Piemonte	ITALIA	Piemonte
Provincia autonoma di Trento	ITALIA	Provincia autonoma di Trento
Rheinland – Pfalz	GERMANIA	Renania – Palatinato
Saarland	GERMANIA	Saarland
Salzburg	AUSTRIA	Salisburgo
Schleswig – Holstein	GERMANIA	Schleswig – Holstein
Steiermark	AUSTRIA	Stiria
Syddanmark	DANIMARCA	Danimarca meridionale
Toscana	ITALIA	Toscana
Veneto	ITALIA	Veneto
Zuid – Nederland	PAESI BASSI	Paesi Bassi del Sud
Baden-Württemberg	GERMANIA	Baden-Württemberg
Bayern	GERMANIA	Baviera
Cataluña	SPAGNA	Catalogna
Île de France	FRANCIA	Île de France
Lombardia	ITALIA	Lombardia
South East (UK)	REGNO UNITO	Sud – Est dell'Inghilterra

1. ECONOMIA

1.1. Pil a parità di potere d'acquisto

Per quanto riguarda l'economia, il primo indicatore a cui si fa riferimento è il Prodotto Interno Lordo. Riportiamo di seguito i dati del Pil (a livello regionale) espresso in PPS, ossia a parità di potere d'acquisto. Il Veneto, con un Pil pari a 140.371 milioni di PPS, rappresenta l'**1,14%** del Pil europeo, la quota più alta all'interno del cluster di riferimento. È seguito dal Piemonte, con un Pil pari a 118.610 milioni di PPS, che rappresenta lo 0,97% del Pil europeo.

Nella parte destra dell'istogramma, troviamo invece le aree geografiche con le quali il Veneto viene posto a confronto da altri accreditati istituti (ad esempio Unioncamere). Queste ultime sono regioni che non abbiamo preso in considerazione per il nostro dossier, essendo risultate fuori dalla nostra classe di riferimento rilevata tramite *cluster analysis*.¹ Tuttavia si possono percepire le differenze che intercorrono fra il Veneto e queste ultime anche da questo primo grafico: per quanto riguarda la quota di Pil europeo infatti, l'Ile de France rappresenta il 4,25%, il Bayern il 3,37%, il Baden-Württemberg il 2,80%, la Lombardia il 2,60%, il Sud-Est inglese l'1,97% e la Catalogna l'1,70%.

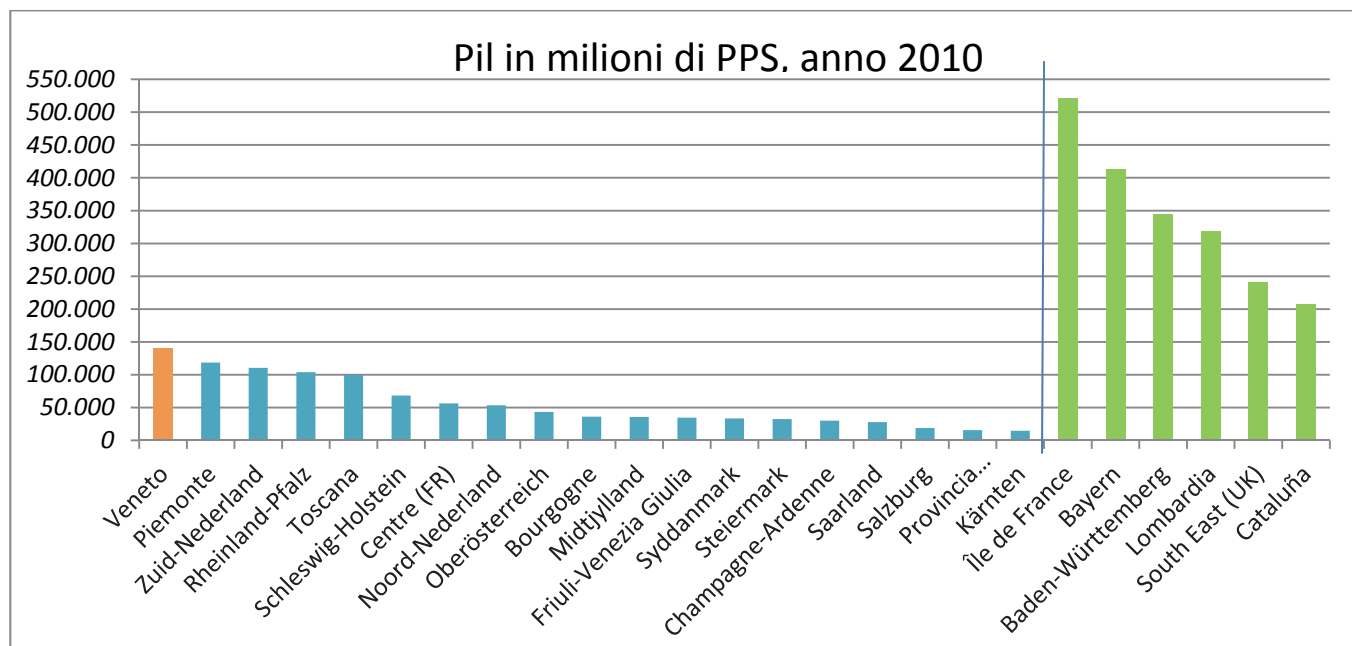


Grafico 1 - Pil in milioni di PPS (Purchasing Power Standard). Elaborazione su dati Eurostat.

¹ Si rimanda all'appendice tecnica per un'illustrazione della procedura che ha portato alla definizione del *benchmark*. Oltre al Pil e alla sua dinamica, il cluster a cui appartiene il Veneto risulta omogeneo anche per fattori quali l'intensità occupazionale, la domanda di lavoratori a qualifica medio-bassa, le disponibilità di lavoratori qualificati e la struttura demografica. Nonostante ciò permangono notevoli differenze tra le regioni-*benchmark* e questo lavoro ha lo proprio scopo di illustrarle.

1.2. Pil a parità di potere d'acquisto pro capite

Per quanto riguarda il Pil pro capite, osserviamo per il Veneto un valore pari a **28.500** a parità di potere d'acquisto. Tenuto conto della popolazione, il Veneto ha un Pil pro capite maggiore rispetto all'Italia e all'Unione Europea a 27 paesi, che hanno rispettivamente 24.700 e 24.500. Tuttavia anche per questo indicatore, abbiamo tenuto divise le aree del cluster da quelle che si è soliti confrontare con il Veneto.

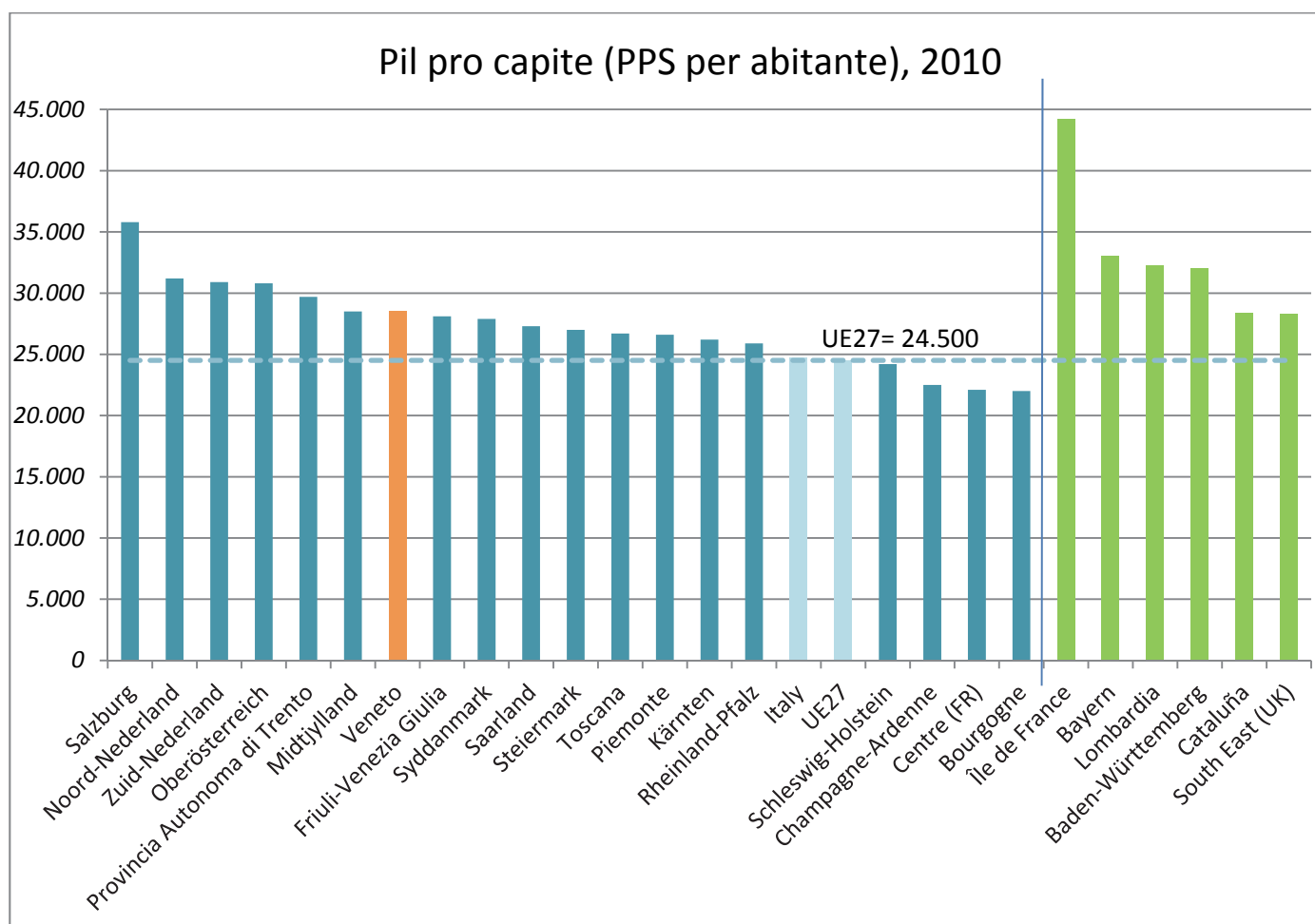


Grafico 2 - Pil pro capite, PPS per abitante. Elaborazione su dati Eurostat.

1.3. Reddito disponibile pro capite

Il reddito disponibile è il reddito al netto dei trasferimenti che operano le amministrazioni pubbliche per favorire la redistribuzione. È un indicatore importante perché ci consente di capire il reddito che le famiglie possono destinare al consumo, al risparmio o agli

investimenti. Il reddito disponibile medio per l'area dell'Unione europea a 27 paesi è pari a 14.544 €, mentre per il Veneto è **19.199 €**.

Le aree che più sembrano svantaggiate da questo punto di vista sembrano le due regioni dei Paesi bassi, rispettivamente lo Zuid-Nederland e il Noord-Nederland, e le due regioni della Danimarca, rispettivamente il Midtjylland e il Syddanmark. In queste aree la tassazione è notoriamente molto elevata, e, a differenza invece di altre aree con regimi di tassazione simili, i servizi offerti ai cittadini hanno una qualità ben maggiore.

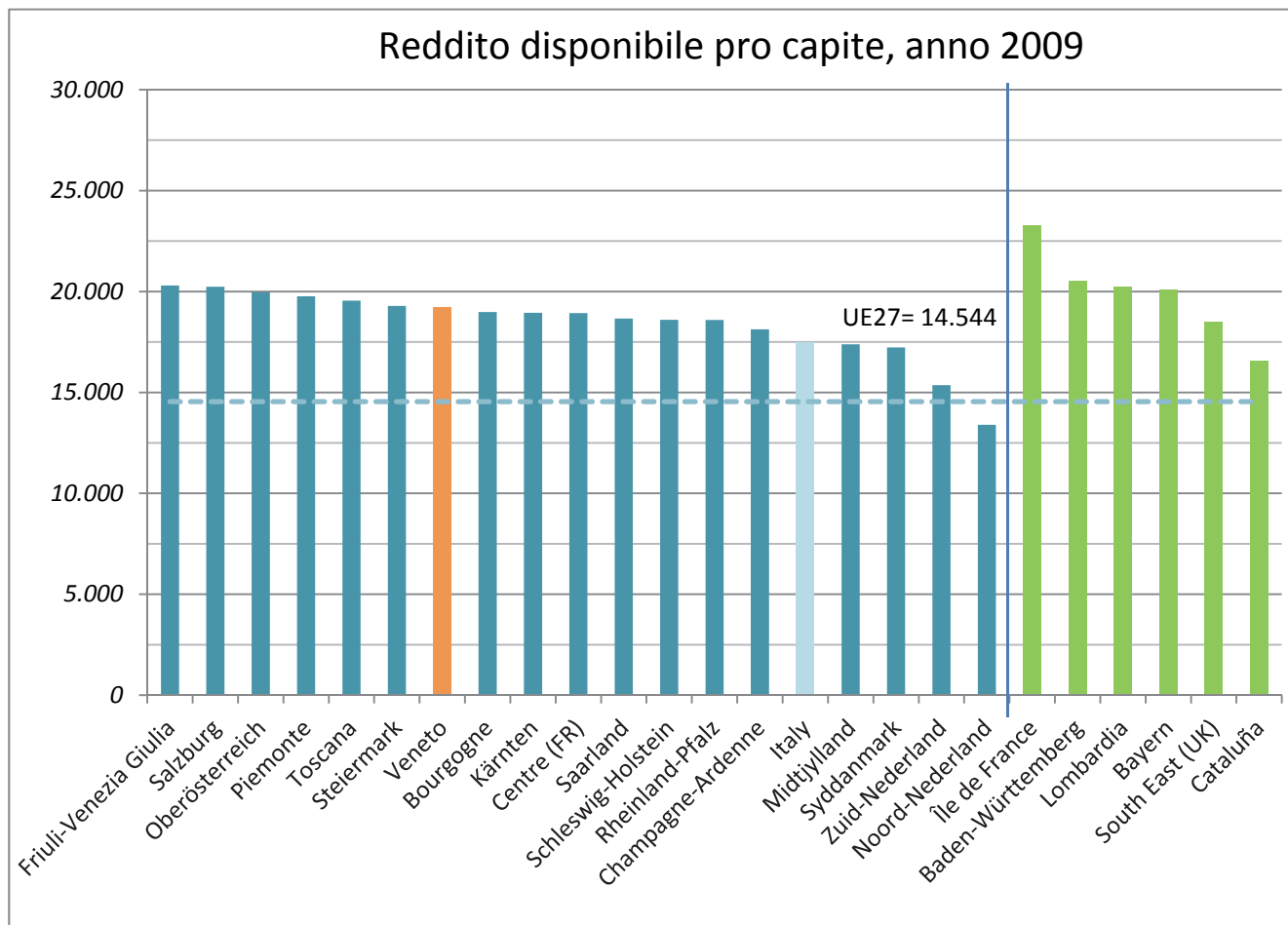


Grafico 3 - Reddito disponibile pro capite. Elaborazione su dati Eurostat e Istat.

2. DEMOGRAFIA E DIPENDENZA

2.1. Popolazione, area e densità abitativa

La struttura demografica di un paese è molto importante perché consente di anticipare alcune dinamiche che poi si ripercuotono nel mercato del lavoro e nella composizione della spesa pubblica (si pensi, ad esempio, alle pensioni e alla sanità). Ci informa quindi di importanti fenomeni sociali, quali l'incidenza delle persone con età inferiore ai 14 anni, l'incidenza delle persone con età maggiore ai 65 anni, il tasso di dipendenza, il tasso di natalità e di mortalità, l'aspettativa di vita alla nascita e la determinazione dei flussi migratori che contribuiscono a questi indici.

	Area (kmq)	Popolazione		Var. %
		2007	2012	2007 - 2012
UE27	4.403.357	495.291.925	503.663.601	1,69%
Italy	301.336	59.131.287	60.820.696	2,86%
Syddanmark	12.257	1.189.817	1.201.342	0,97%
Midtjylland	13.000	1.227.428	1.266.682	3,20%
Rheinland-Pfalz	19.853	4.052.860	3.999.117	-1,33%
Saarland	2.569	1.043.167	1.013.352	-2,86%
Schleswig-Holstein	15.800	2.834.254	2.837.641	0,12%
Champagne-Ardenne	25.606	1.339.487	1.333.163	-0,47%
Centre (FR)	39.151	2.526.919	2.562.227	1,40%
Bourgogne	31.582	1.633.891	1.646.600	0,78%
Piemonte	25.403	4.352.828	4.464.896	2,57%
Provincia Autonoma di Trento	6.207	507.030	533.394	5,20%
Veneto	18.399	4.773.554	4.957.082	3,84%
Friuli-Venezia Giulia	7.858	1.212.602	1.236.103	1,94%
Toscana	22.994	3.638.211	3.761.616	3,39%
Noord-Nederland	:	1.702.020	1.718.896	0,99%
Zuid-Nederland	:	3.546.847	3.586.761	1,13%
Oberösterreich	11.980	1.403.663	1.416.772	0,93%
Steiermark	16.401	1.202.483	1.213.255	0,90%
Kärnten	9.538	559.393	557.773	-0,29%
Salzburg	7.156	526.048	534.122	1,53%

Tabella 1 - Area, popolazione e variazione della popolazione (anni 2007 - 2012). Elaborazione su dati Eurostat.

Var. popolazione % (2007 - 2012)

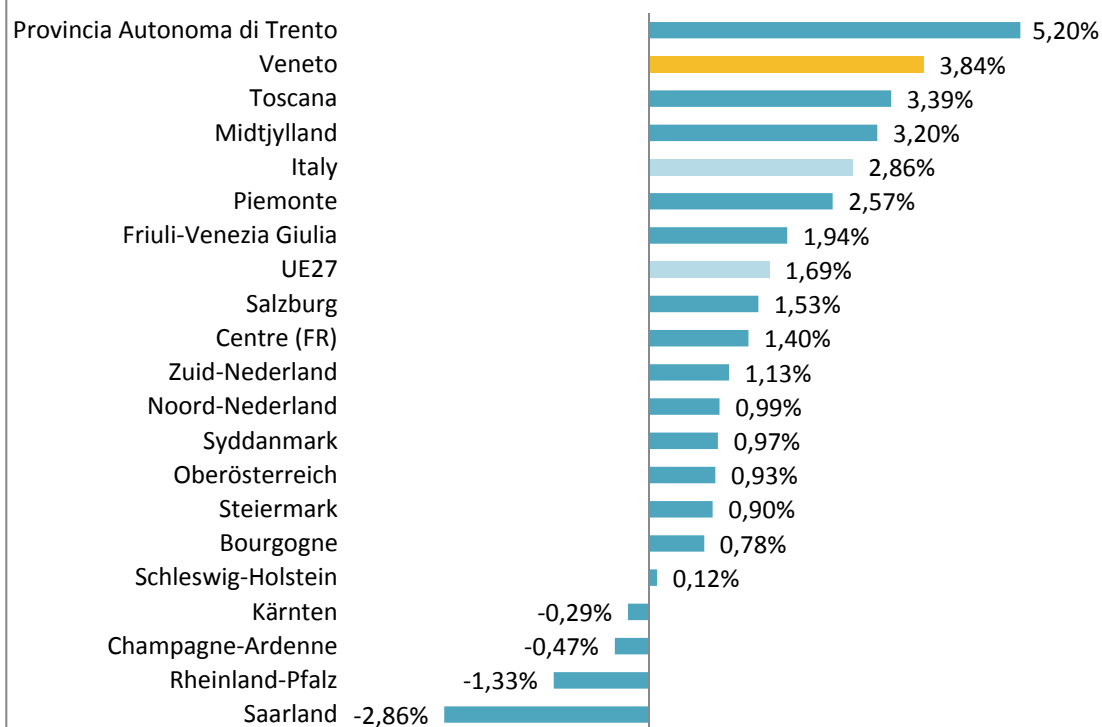


Grafico 4 – Variazione della popolazione (2007 – 2012). Elaborazione su dati Eurostat.

Densità abitativa (ab/kmq), anno 2012

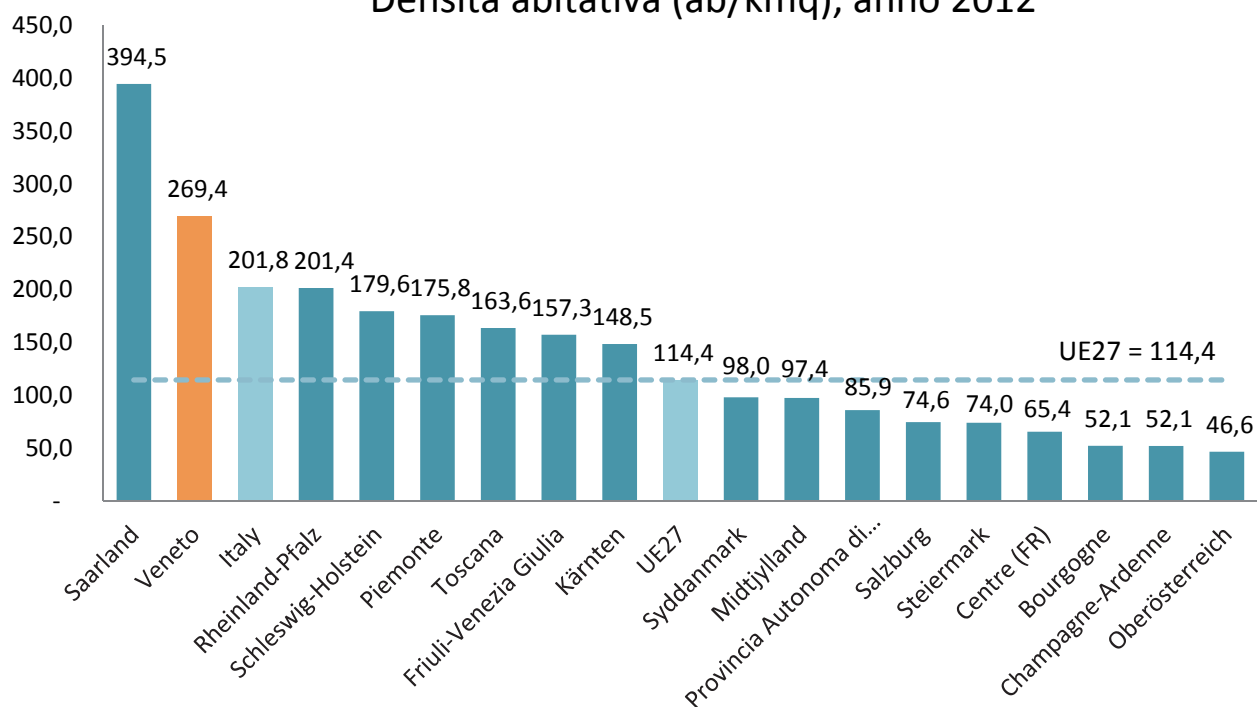


Grafico 5 - Densità abitativa (abitanti per kmq). Elaborazione su dati Eurostat.

Il Veneto con **4.957.082 abitanti** appare come la regione più popolosa all'interno del *cluster*; altre regioni simili per popolazione (in termini assoluti) sono il Piemonte, il Rheinland-Pfalz, la Toscana e il Sud dei Paesi Bassi. Oltre a ciò, il Veneto ha una dinamica demografica superiore a quasi tutte le altre regioni di confronto, con una variazione della popolazione (dal 2007 al 2012) del 3,84%, seconda solo alla Provincia autonoma di Trento che presenta una variazione del 5,20%. In tutto ciò non possiamo non menzionare l'incidenza positiva che apportano i flussi migratori, che esamineremo in seguito.

Per quanto riguarda la densità abitativa, la media dell'Unione europea a 27 paesi presenta un valore di 114,4 abitanti per kmq; il Veneto si discosta ampiamente presentando un valore di **269,4** abitanti per kmq, ad indicare una elevata concentrazione demografica, seconda solo alla regione del Saarland che presenta un valore 394,5 abitanti per kmq (questo dovuto anche al fatto che la regione si estende lungo una superficie abbastanza limitata). Le aree all'interno del *cluster* che presentano una densità abitativa minore sono le tre regioni della Francia, le tre regioni dell'Austria, le due della Danimarca e la Provincia autonoma di Trento.

2.2. Popolazione per fasce d'età

Per quanto riguarda l'incidenza dei giovani (\leq a 14 anni) sulla popolazione, l'Unione europea a 27 paesi presenta un valore pari al 15,6%, superiore sia all'Italia sia al Veneto, che presentano un'incidenza dei giovani sulla popolazione rispettivamente pari al 14,02% e al **14,22%**. Le altre aree del cluster, all'infuori delle regioni italiane, presentano quasi tutte dei valori più elevati o in linea rispetto alla media europea.

Per quanto riguarda invece la popolazione con età maggiore o uguale ai 65 anni, si osserva che mentre l'Unione europea a 27 paesi ha un'incidenza di anziani pari al 17,8% sulla popolazione, l'Italia e il Veneto presentano delle incidenze ben più elevate, rispettivamente pari al 20,6% e al **20,3%**. Tuttavia le aree all'interno del cluster che presentano l'incidenza maggiore sono le regioni Friuli Venezia Giulia (23,9%), Toscana (23,6%) e Piemonte (23,2%). Se guardiamo l'andamento delle incidenze rispetto all'anno 2007, osserviamo per il Veneto che sia l'incidenza dei giovani sia l'incidenza degli anziani sono aumentate, rispettivamente dello 0,25% e dello 0,93%. Questo indica però una dinamica non incoraggiante per i prossimi decenni, che va monitorata guardando soprattutto i tassi di dipendenza.

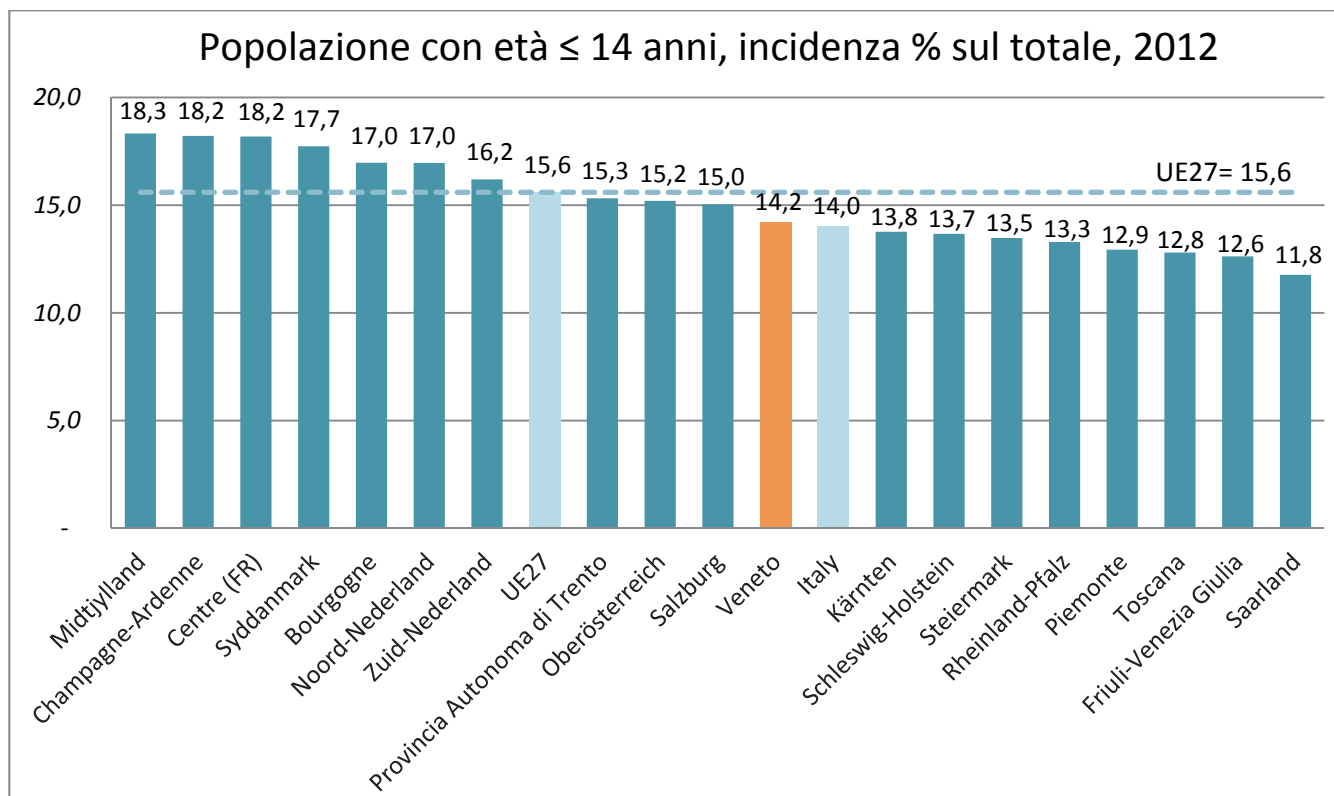


Grafico 6 - Incidenza della popolazione di età minore o uguale a 14 anni sul totale della popolazione. Elaborazione su dati Eurostat.

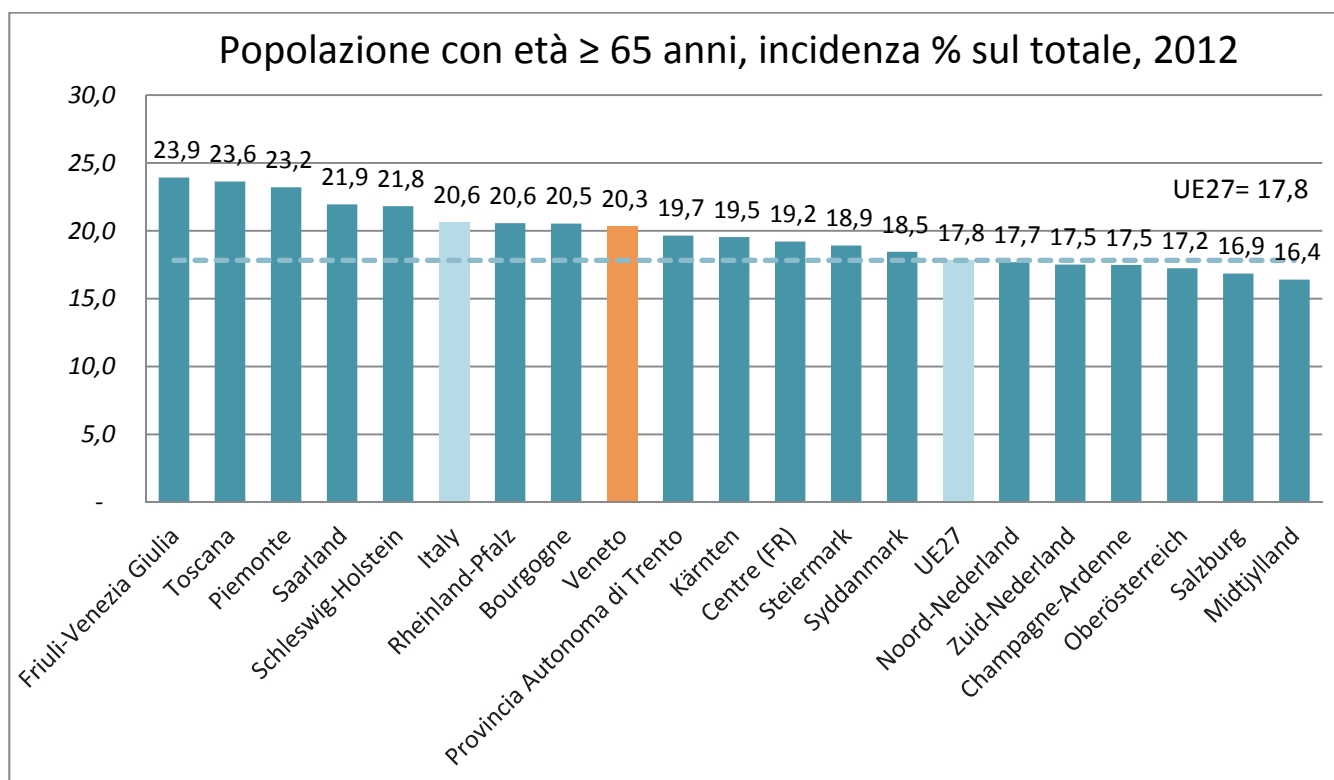


Grafico 7 - Incidenza della popolazione di età maggiore o uguale a 65 anni sul totale della popolazione. Elaborazione su dati Eurostat.

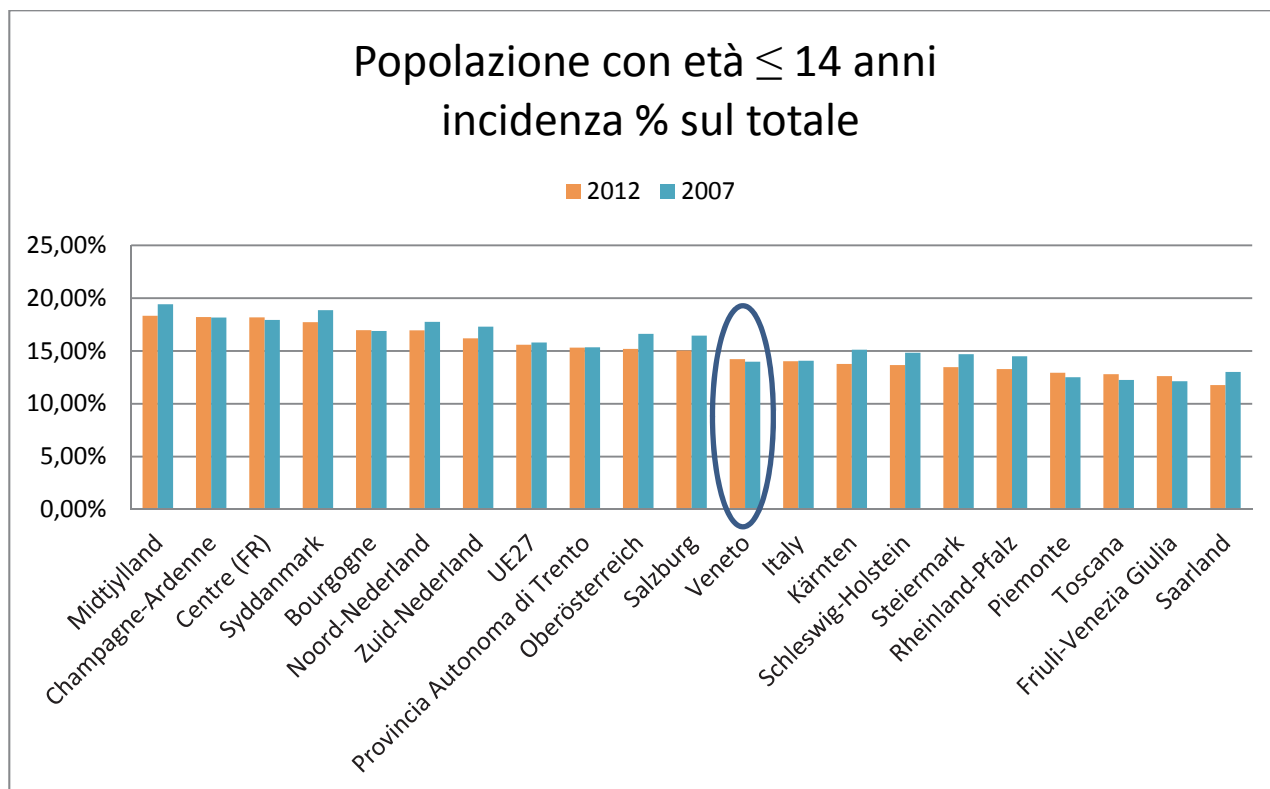


Grafico 8 - Incidenza della popolazione con età minore o uguale a 14 anni sul totale della popolazione, anni 2007 e 2012. Elaborazione su dati Eurostat.

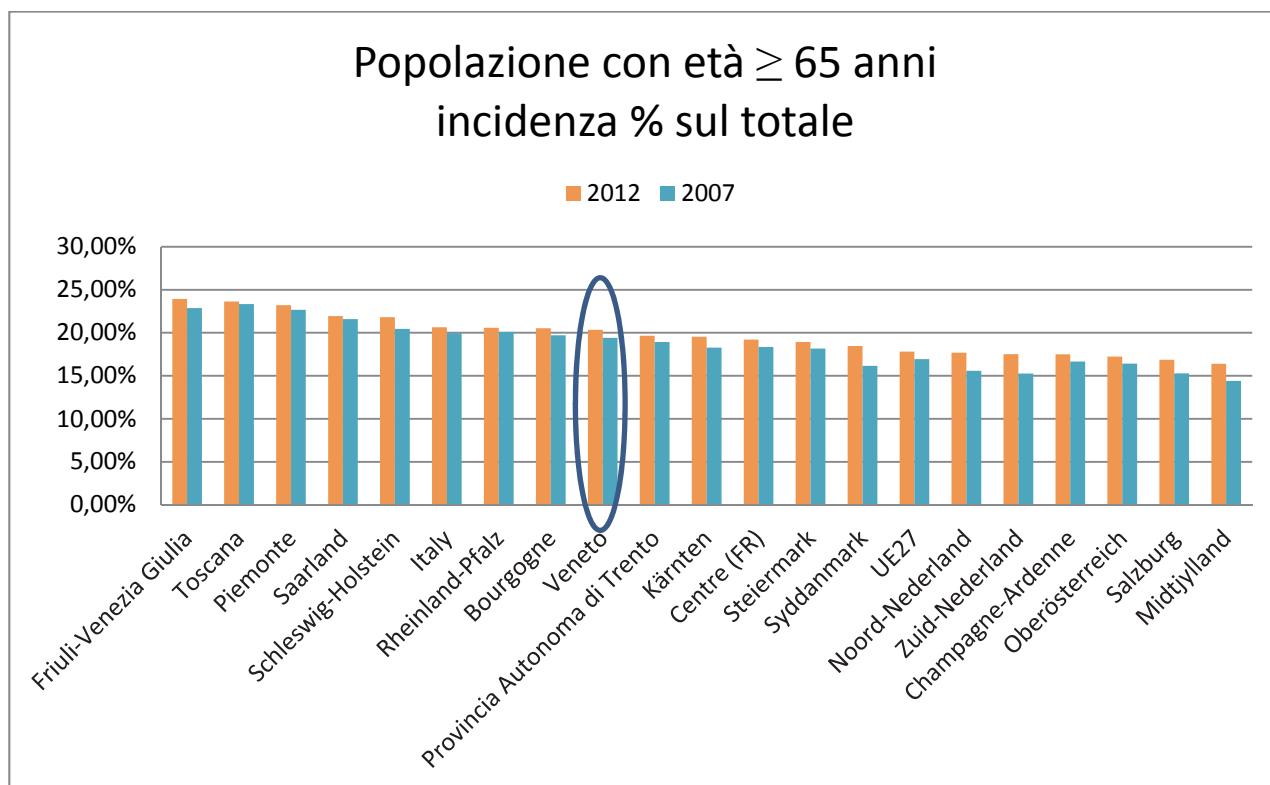


Grafico 9 - Incidenza della popolazione con età maggiore o uguale a 65 anni sul totale della popolazione, anni 2007 e 2012. Elaborazione su dati Eurostat.

2.3. Tassi di dipendenza

Il tasso di dipendenza è una misura della sostenibilità della popolazione. Esso calcola appunto quanti individui ci sono in età non attiva (\leq a 14 anni; \geq a 65 anni, le parti “dipendenti”) ogni 100 individui in età attiva. L'indice misura il carico economico e sociale teorico della popolazione in età attiva. Esso è composto dalla dipendenza giovanile e dalla dipendenza senile; ovviamente è a quest'ultima che va rivolta gran parte della nostra attenzione. Un indice di dipendenza superiore a 50 indica uno squilibrio generazionale.

L'Unione europea a 27 paesi presenta un tasso di dipendenza pari al 50,18%, così composto: 26,76% di dipendenza senile e il restante 23,42% di dipendenza giovanile. Il Veneto invece è caratterizzato da un tasso di dipendenza totale pari al **52,84%**, di cui **31,10%** è la dipendenza senile mentre **21,74%** è la dipendenza giovanile. Se ne osserviamo l'andamento, vediamo che nel 2007 il tasso di dipendenza per il Veneto era pari a 50,15%. Il tasso è quindi cresciuto di 269 punti base, una situazione non incoraggiante che dovrebbe farci capire l'importanza di agire presto in tal senso.

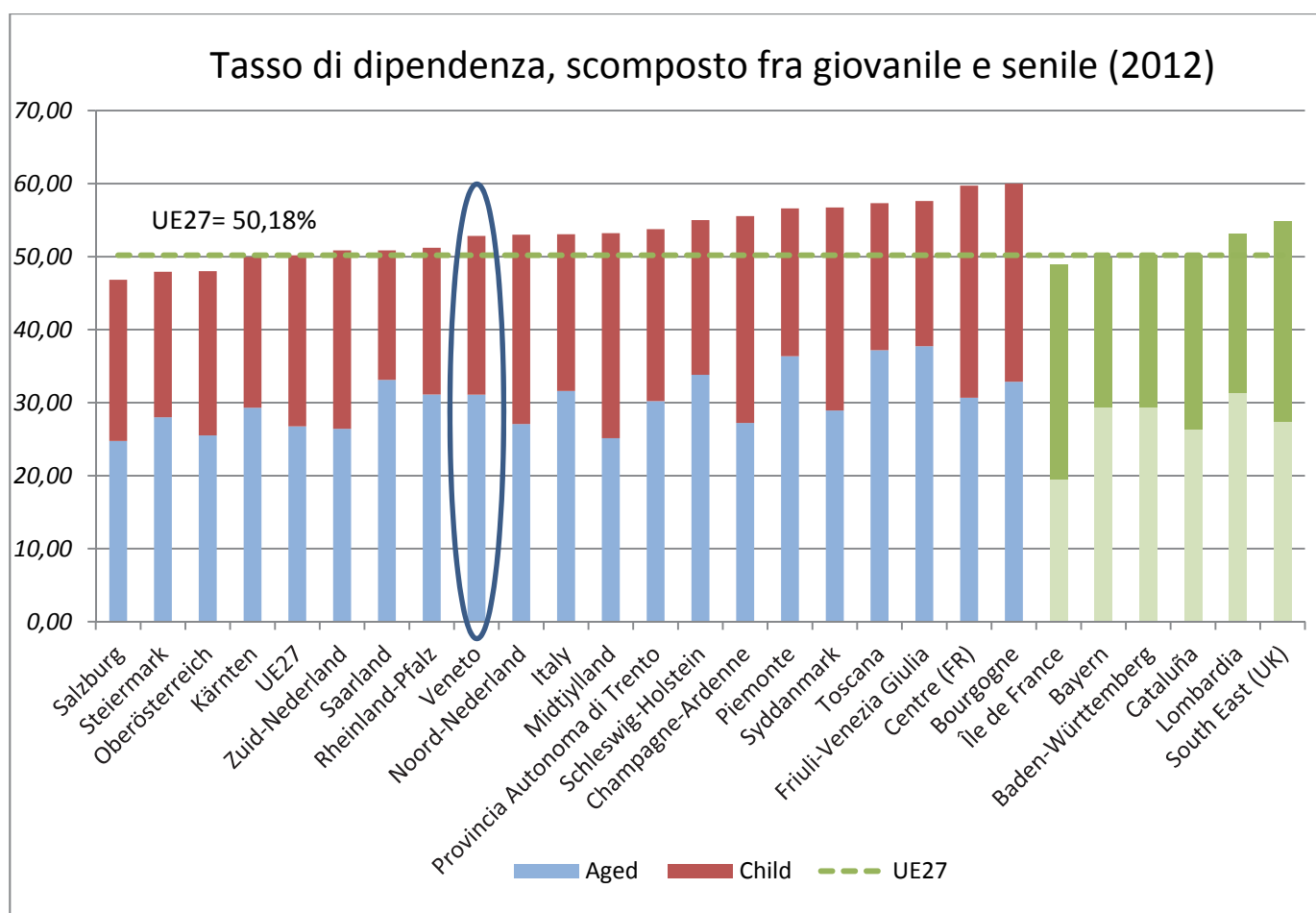


Grafico 10 - Tasso di dipendenza totale, giovanile e senile. Elaborazione su dati Eurostat.

All'interno del cluster osserviamo tuttavia che altre aree versano in condizioni più preoccupanti della nostra (ma non che sia questo un buon motivo per non agire, abbiamo superato comunque la soglia di parità). In particolare ci sembrano casi limite le situazioni del Friuli Venezia Giulia, della Toscana e del Piemonte, che registrano rispettivamente per quanto riguarda il tasso di dipendenza senile un 37,73%, un 37,18% e un 36,35%, ad indicare la presenza dei forti squilibri generazionali.

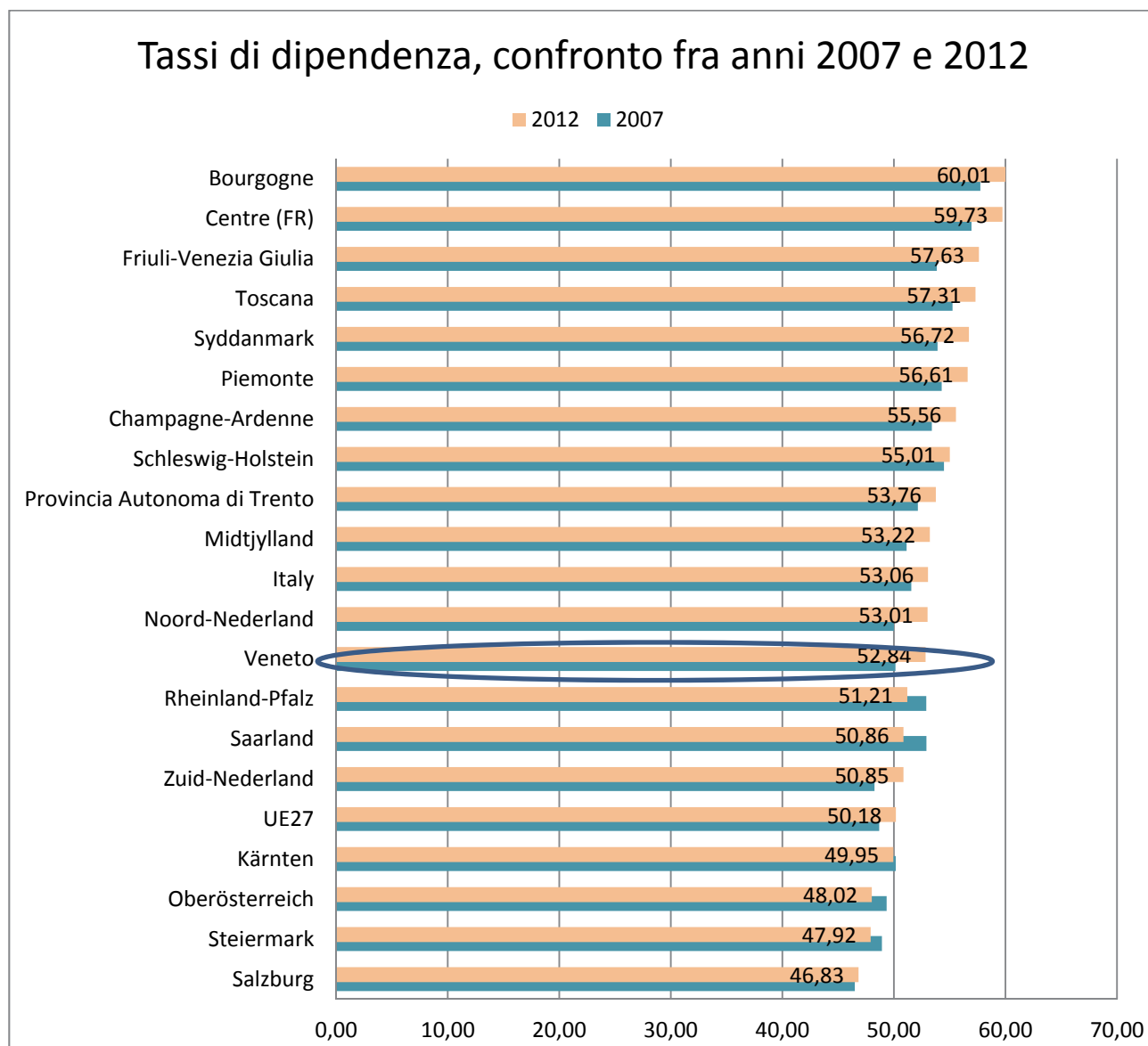


Grafico 11 - Tasso di dipendenza totale, anni 2007 e 2012. Elaborazione su dati Eurostat.

2.4. Tasso di natalità

Il tasso di natalità esprime il rapporto fra il numero dei nati vivi in un periodo di riferimento e la media della popolazione totale. Esso va letto a completamento alla struttura demografica di un determinato territorio. Per quanto riguarda il Veneto, osserviamo per l'anno 2012 un valore pari a **9,2** nati ogni 1.000 abitanti. Nel 2007 esso era pari a 10. Questo indice analizzato assieme al tasso di indipendenza, suggerisce che siamo un popolo che sta invecchiando. L'Unione europea a 27 paesi ha un tasso di natalità pari a 10,4‰, anche se si è abbassato anch'esso rispetto al 2007, anno in cui era pari a 10,7‰.

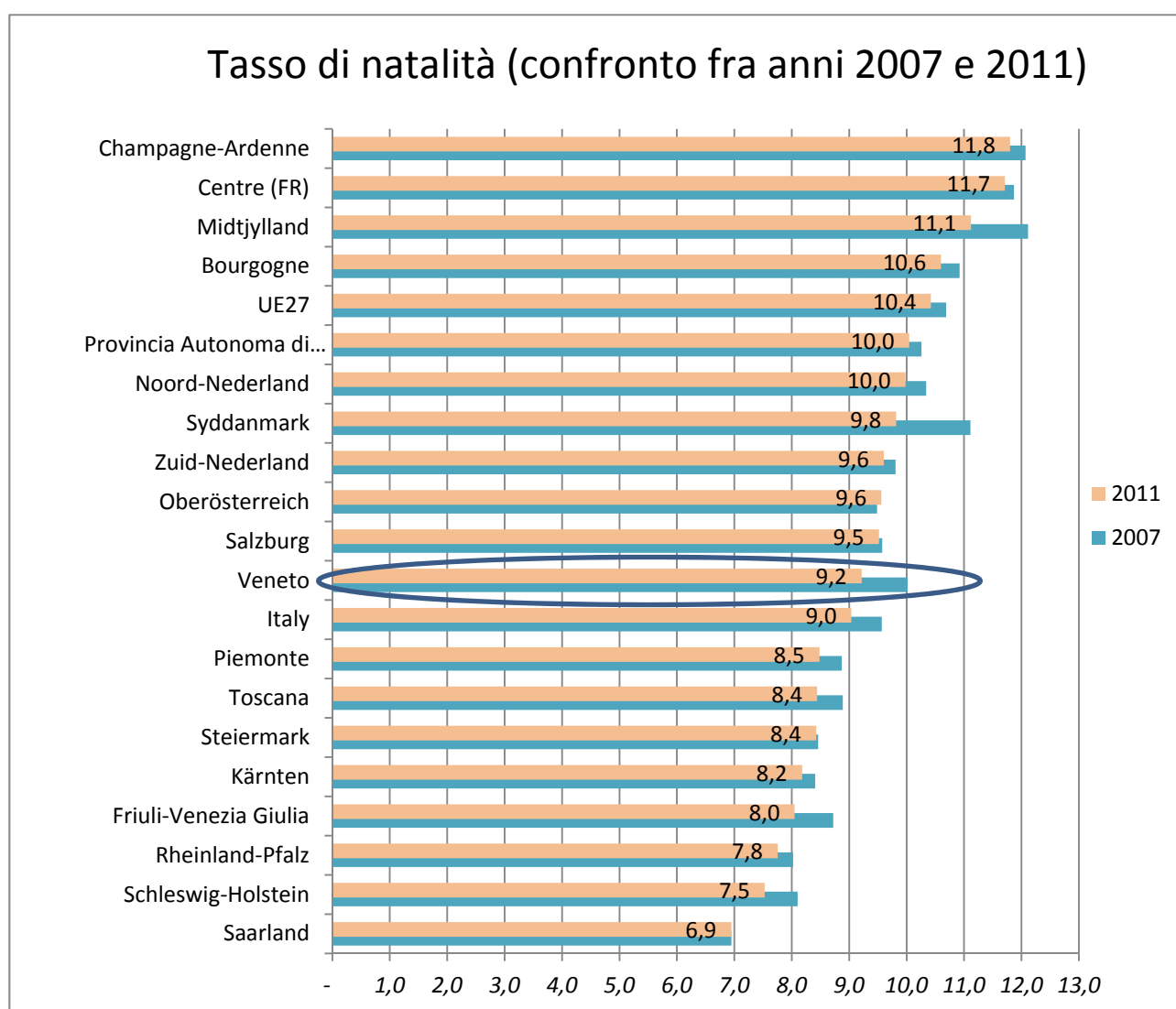


Grafico 12 - Tasso di natalità, anni 2007 e 2011. Elaborazione su dati Eurostat.

2.5. Il tasso di fecondità

Un altro indicatore che ci fa capire il problema dell'invecchiamento della popolazione (*ageing society*) è il tasso di fecondità, ossia il numero medio di figli per donna. Il tasso di fecondità pari a 2,1 è quello necessario a stabilizzare idealmente la popolazione (in assenza di immigrazione). Per quanto riguarda il Veneto, abbiamo un valore pari a **1,44**, valore poco superiore alla media italiana (tasso di fecondità: 1,40).

Si osserva che le regioni con il tasso di fecondità più elevato sono le tre aree della Francia, e ciò non stupisce vista l'attenzione che questo paese ha sempre dedicato alle politiche familiari e di natalità.

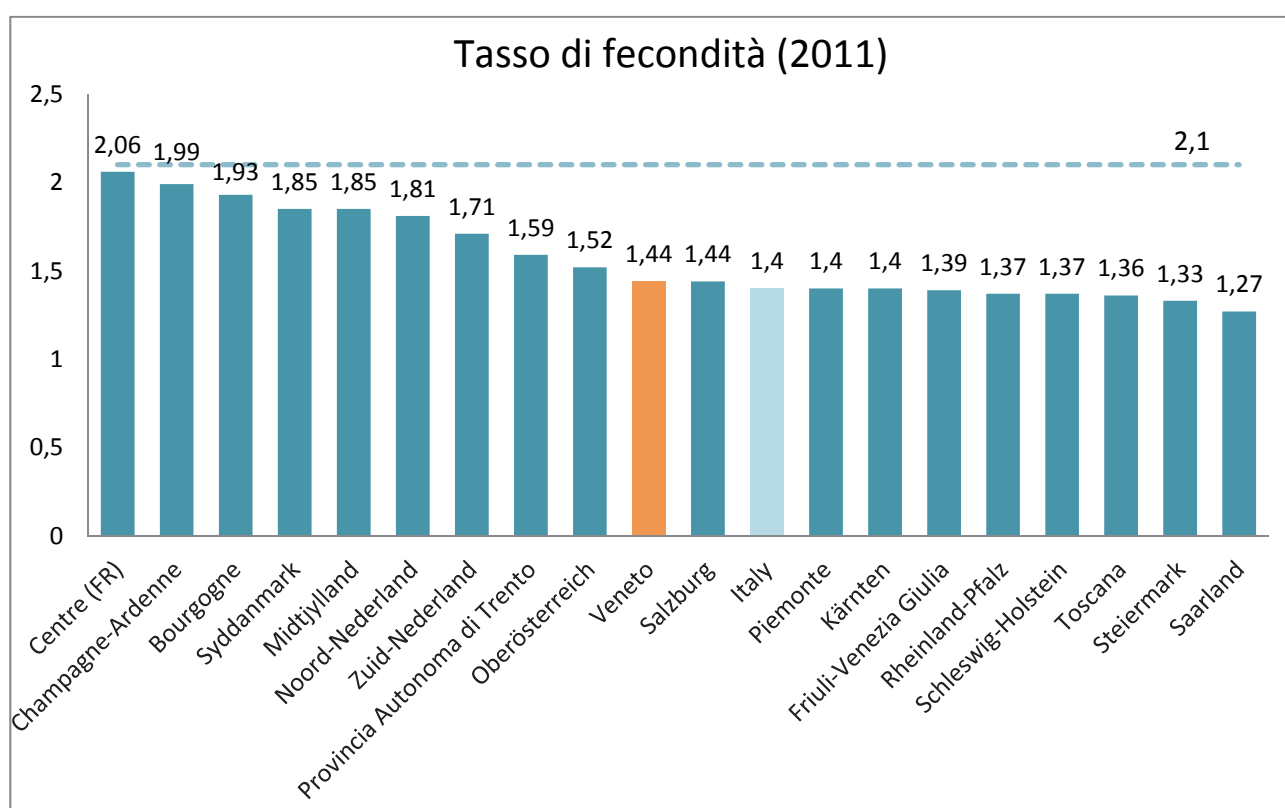


Grafico 13 - Tasso di fecondità, anno 2011. Elaborazione su dati Eurostat.

2.6. Il tasso di mortalità

Il tasso di mortalità indica il numero medio dei decessi in una popolazione di 1.000 abitanti, in un dato periodo. Questo indicatore ci informa sullo sviluppo di una popolazione, e ad un valore elevato dell'indicatore corrisponderà un negativo sviluppo della popolazione. Per quanto riguarda il Veneto osserviamo per il 2010 un tasso pari a **9,1‰**, un valore minore

rispetto a quello dell'Unione Europea a 27 paesi, che presenta un tasso del 9,7‰. In entrambi i casi, il valore era uguale anche nel 2007.

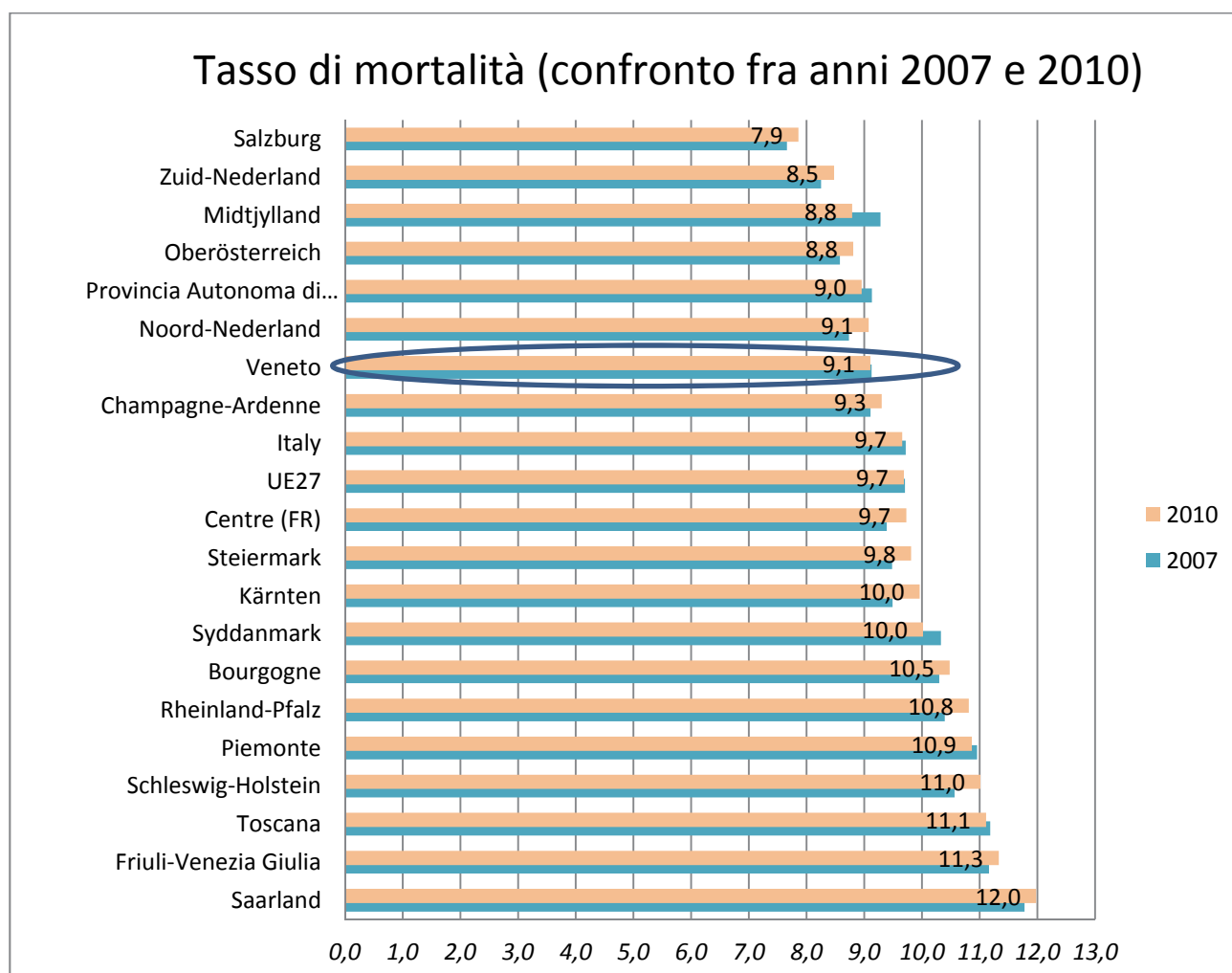


Grafico 14 - Tasso di mortalità, anni 2007 e 2010. Elaborazione su dati Eurostat.

2.7. Aspettativa di vita

Per quanto riguarda l'aspettativa di vita, la regione Veneto con un'aspettativa di vita maschile pari a **80,2** anni e un'aspettativa di vita femminile pari a **85,7** anni si posiziona sicuramente ai primi posti, seconda solo alla Provincia autonoma di Trento.

Con riferimento all'Unione Europea a 27 paesi, osserviamo come l'aspettativa di vita sia maggiore nel Veneto, con una differenza per quanto riguarda il 2010 di 3 anni circa, sia per l'aspettativa maschile sia per la femminile.

Aspettativa di vita alla nascita						
Aree	Maschi			Femmine		
	2000	2005	2010	2000	2005	2010
UE27	:	75,4	77 *	:	81,6	82,9 *
Italy	76,9	78	79,8	82,8	83,6	85
Syddanmark	:	:	77,4	:	:	81,5
Midtjylland	:	:	78,2	:	:	82,1
Rheinland-Pfalz	:	76,8	78,2	:	81,7	83
Saarland	:	75,7	76,9	:	81,1	82,2
Schleswig-Holstein	:	76,9	77,9	:	81,8	82,6
Champagne – Ardenne	74,5	75,6	76,9	82,4	83,3	84,6
Centre (FR)	75,7	77	78	83,3	84,1	85,1
Bourgogne	74,7	76,4	77,6	83,1	83,6	85,3
Piemonte	76,6	78	79,7	82,6	83,5	84,8
Provincia Autonoma di Trento	77,2	78,6	80,4	83,6	84,9	86,1
Veneto	77	78,5	80,2	83,8	84,6	85,7
Friuli-Venezia Giulia	76,6	77,9	79,6	82,9	83,9	85,1
Toscana	77,7	79,1	80,5	83,4	84,4	85,4
Noord – Nederland	:	77,3	78,5	:	82	82,8
Zuid – Nederland	:	77	78,8	:	81,6	82,9
Oberösterreich	75,6	76,9	78,2	81,6	82,4	83,8
Steiermark	75,1	76,6	77,9	81,3	82,8	83,9
Kärnten	75,7	76,8	77,8	81,6	82,6	84
Salzburg	75,9	76,9	78,8	81,5	82,5	84,3

Tabella 2 - Aspettativa di vita alla nascita maschile e femminile, anni 2000, 2005 e 2010. Elaborazione su dati Eurostat.

3. LAVORO

3.1. Tasso di occupazione

I tassi relativi all'occupazione, insieme ad alcuni indicatori su educazione, R&S, cambi climatici e povertà, rientrano fra i target che l'Unione Europea si è prefissata di raggiungere nel programma "Europa 2020".

Per quanto riguarda il tasso di occupazione, il Veneto presenta un tasso del **65%**, un valore superiore sia all'UE a 27 (64,1%) sia all'Italia, che con un tasso del 56,8% risulta essere l'area con il più basso tasso di occupazione tra quelli considerati.

Tuttavia il Veneto, pur avendo uno 0,9% in più rispetto all'Unione Europea a 27 paesi, presenta un valore più basso rispetto a tutte le aree del *cluster*, all'infuori delle regioni italiane e francesi.

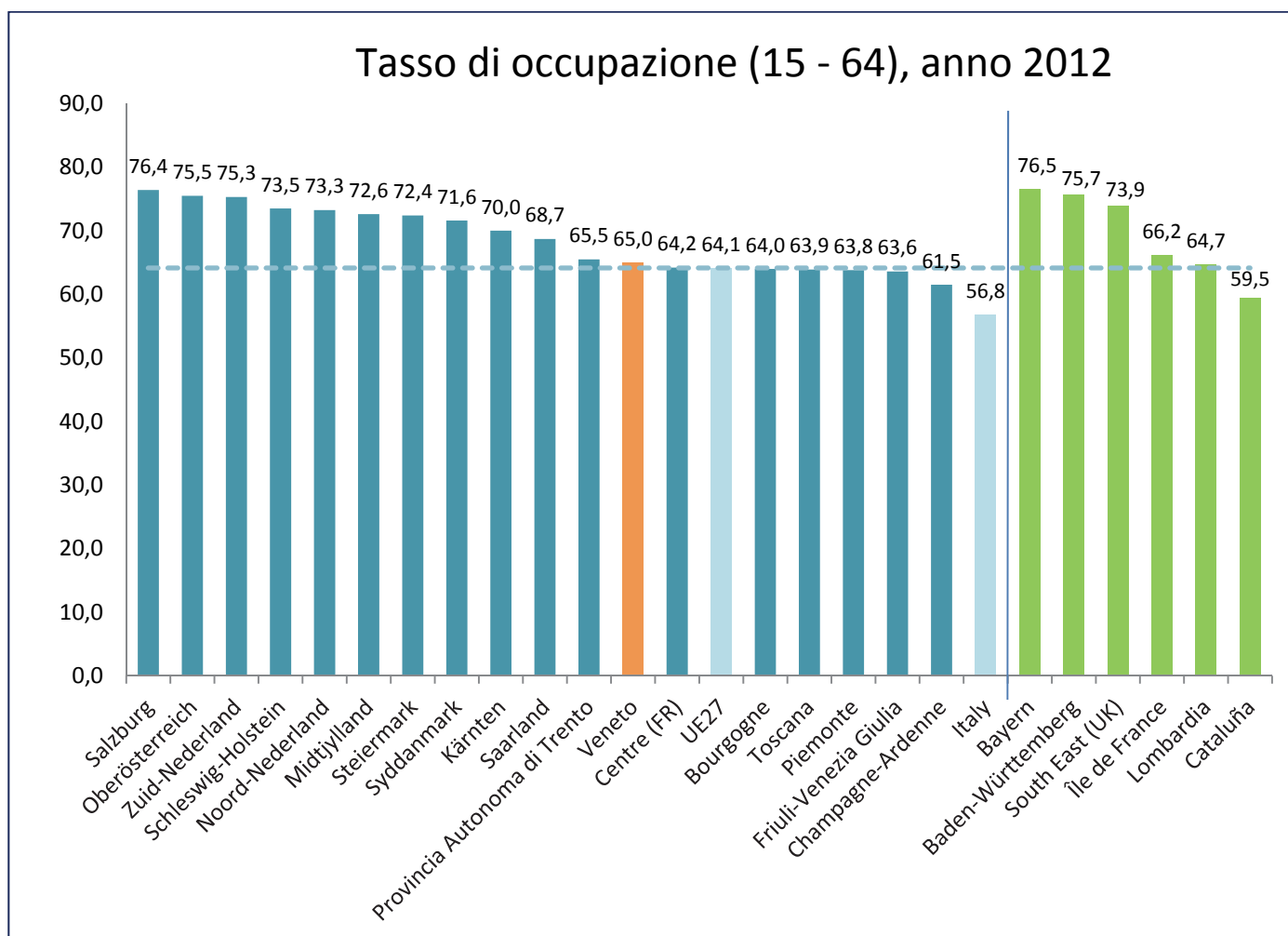


Grafico 15 - Tasso di occupazione (15 - 64 anni), anno 2012. Elaborazione su dati Eurostat.

3.2. Tasso di occupazione femminile

Il tasso di occupazione femminile in Italia, pari a 47,1%, risulta essere nettamente inferiore rispetto al tasso dell'Unione Europea a 27 paesi (che ha un valore di 58,6% per l'anno 2012) e più in generale rispetto a tutte le altre aree del *cluster*. Il Veneto è purtroppo la regione all'interno del *cluster* con il tasso di occupazione femminile più basso, pari a **54,9%**. Si osserva come la parte inferiore del grafico riguardi prevalentemente le aree italiane, sottolineando un certo ritardo per quanto riguarda le politiche di sostegno familiare e le politiche che mirano sempre di più verso una parità fra generi.

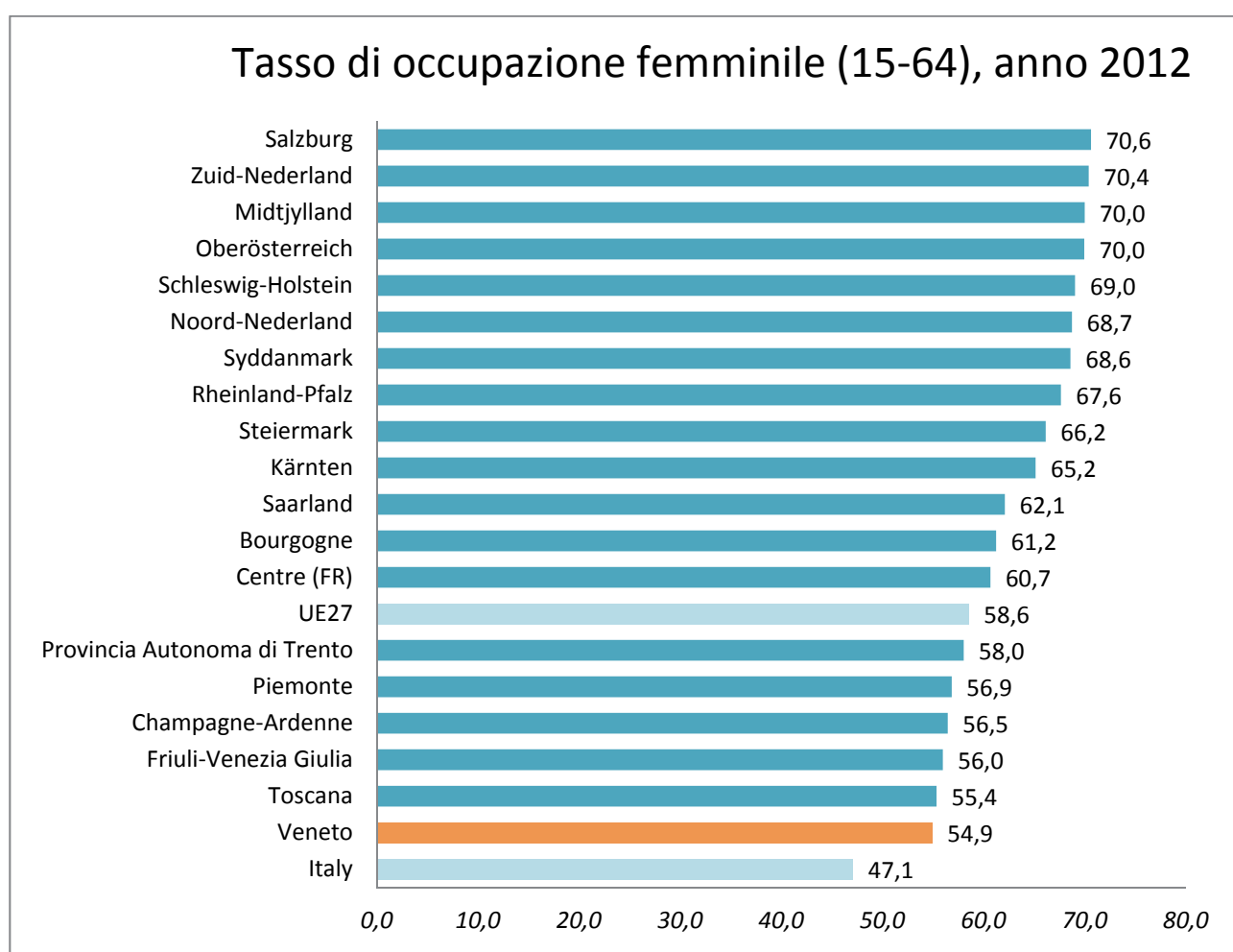


Grafico 16 - Tasso di occupazione femminile (15 - 64 anni), anno 2012. Elaborazione su dati Eurostat.

3.3. Tasso di disoccupazione

Il tasso di disoccupazione è il rapporto percentuale fra le persone che sono in cerca di una occupazione e la forza lavoro (che è pari alla somma delle persone in cerca di lavoro e degli occupati). Questo valore ci informa delle persone che pur essendo in cerca di lavoro, non riescono a trovarlo. E' noto ormai come questo indicatore sottostimi la disoccupazione: infatti non è considerato disoccupato chi ha rinunciato a cercare lavoro (i cosiddetti "scoraggiati"), chi ha lavorato anche solo un'ora a settimana, o chi non ha mai lavorato (che in questo caso viene considerato "inoccupato").

Osservando l'istogramma possiamo affermare che, con un tasso del **6,6%**, il Veneto è pienamente in media con le altre regioni del *cluster*, dove troviamo dei tassi elevati per le aree francesi e dei tassi più bassi per le aree della Germania, dei Paesi Bassi e dell'Austria. Con riferimento all'Italia e all'Unione Europea a 27 paesi, che hanno un tasso rispettivamente pari a 10,7% e a 10,4%, ci troviamo in una situazione migliore. Se osserviamo le regioni che non fanno parte del *cluster* di riferimento, la situazione appare abbastanza analoga, ad eccezione della Catalogna che per il 2012 presenta un tasso del 22,6%, un valore decisamente elevato principalmente dovuto al duro colpo che l'economia spagnola ha subito durante l'attuale crisi economica.

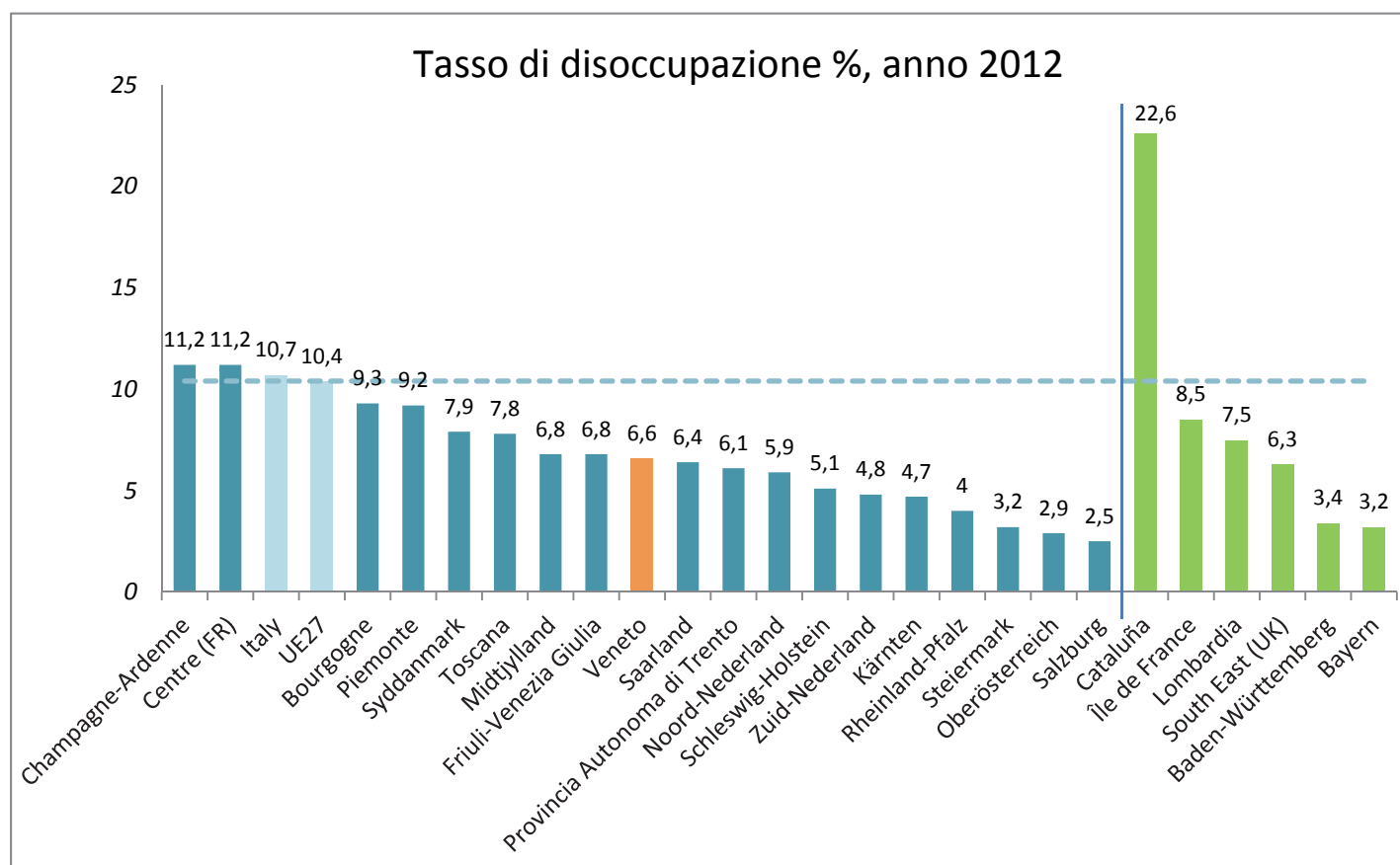


Grafico 17 - Tasso di disoccupazione, anno 2012. Elaborazione su dati Eurostat.

3.4. Tasso di disoccupazione femminile

Il tasso di disoccupazione femminile presenta le stesse problematiche del precedente indicatore. È un indice molto importante che va interpretato però per il significato che ha. In particolare, per quanto riguarda il mercato del lavoro femminile, troviamo un tasso di disoccupazione per il Veneto pari a **7,8%**, un valore medio rispetto a quelli delle aree del *cluster*, e un valore decisamente inferiore se paragonato al tasso dell'Unione Europea a 27 paesi, che presenta un valore del 10,5%.

Ricordiamo che, per costruzione, il tasso di disoccupazione non include le donne che non hanno mai percepito una retribuzione, come ad esempio può essere il caso delle casalinghe, che rientrano nella categoria della popolazione non attiva.

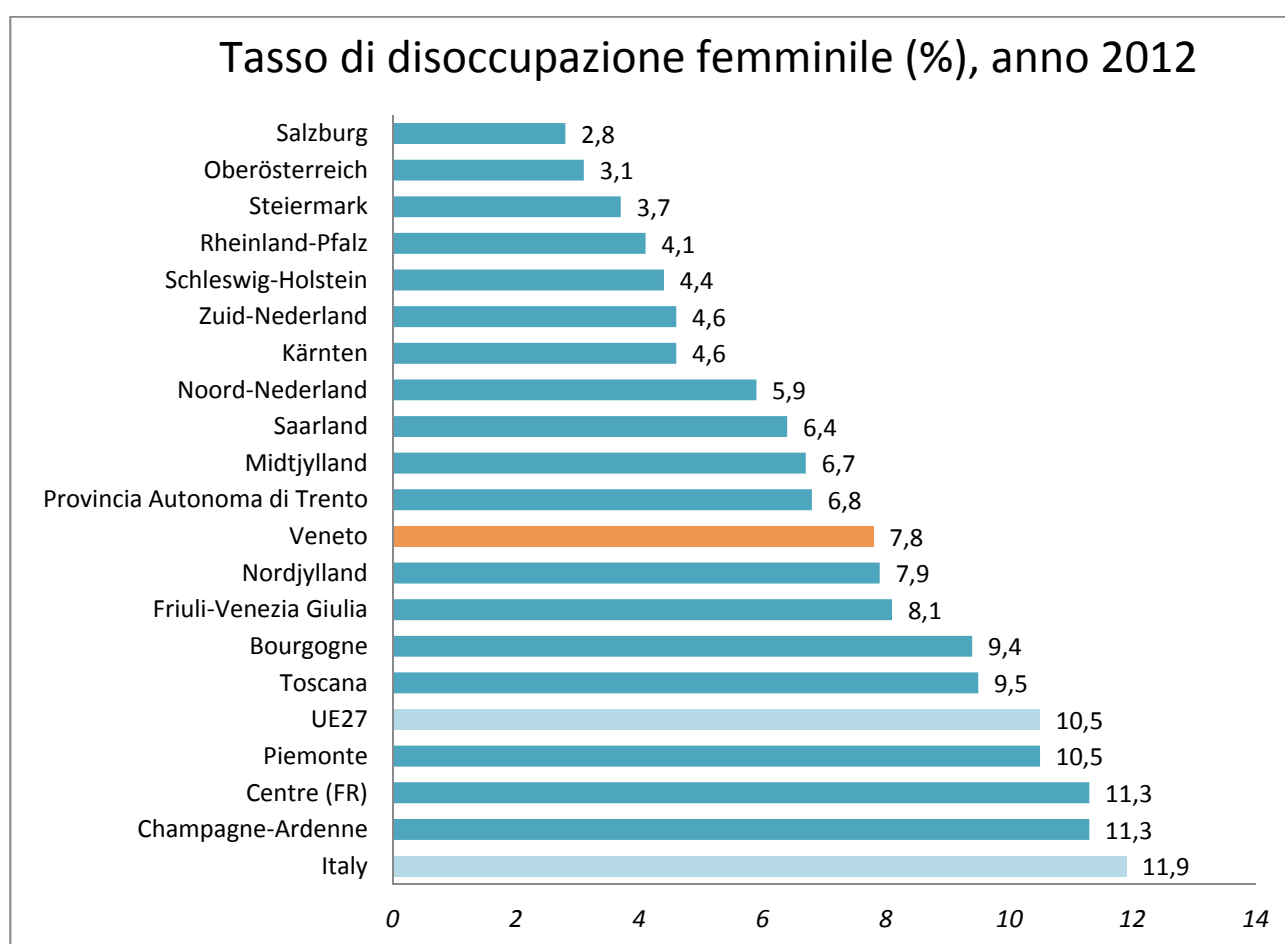


Grafico 18 - Tasso di disoccupazione femminile, anno 2012. Elaborazione su dati Eurostat.

3.5. Tasso di disoccupazione giovanile

Il tasso di disoccupazione giovanile è il rapporto percentuale fra le persone in cerca di una occupazione in età 15-24 anni e la forza lavoro della medesima classe di età. Presentando sempre valori molto elevati, questo tasso rappresenta sempre più un problema per l'intera area europea, a cui l'Europa, le singole nazioni e le regioni vogliono e cercano di dare soluzioni.

Per quanto riguarda le aree considerate, la regione Veneto presenta un tasso del **23,7%**, un valore molto elevato che però appare come valore medio all'interno del *cluster*. Le regioni che hanno i tassi più elevati sono quelle italiane e quelle francesi.

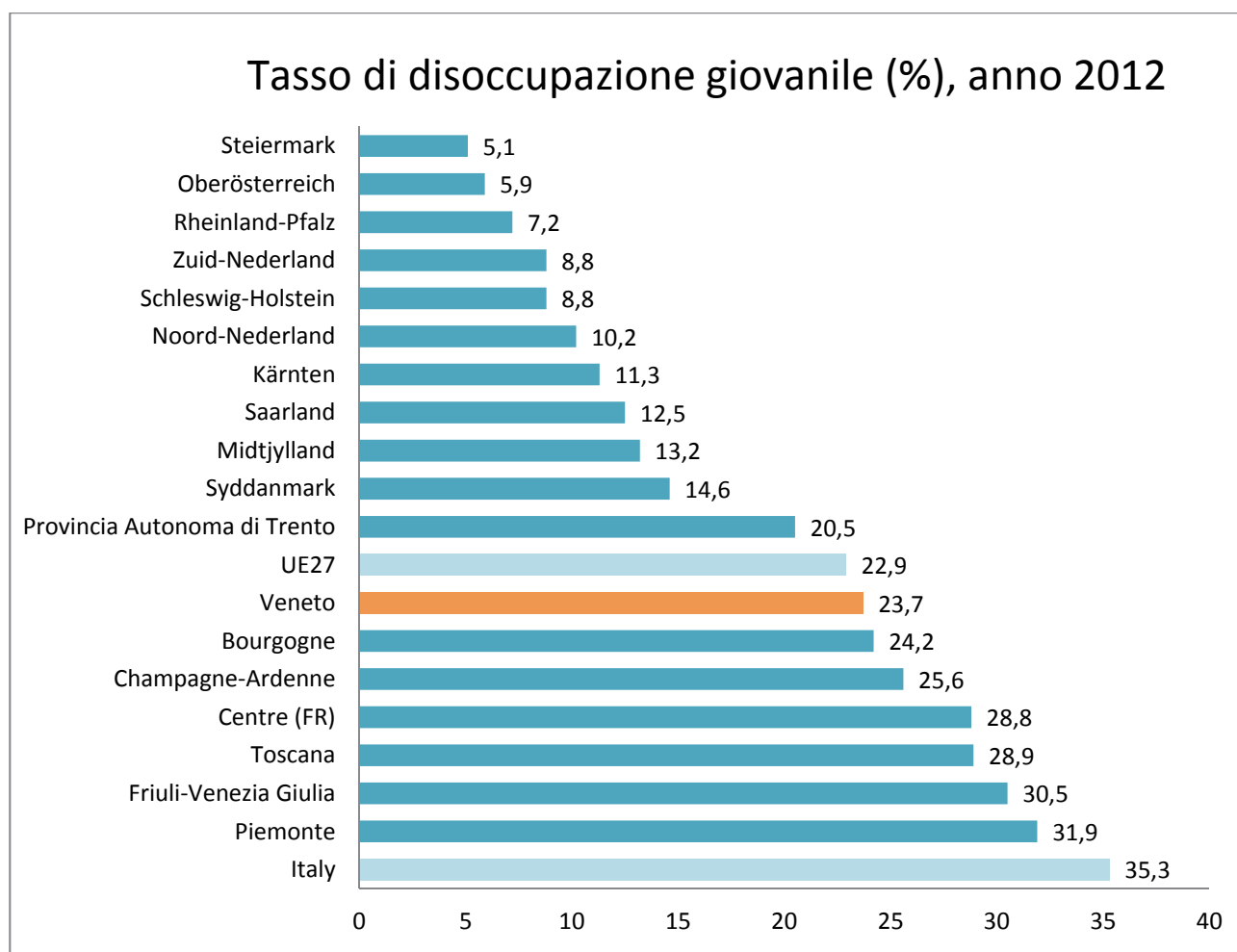


Grafico 19 - Tasso di disoccupazione giovanile (15 - 24 anni), anno 2012. Elaborazione su dati Eurostat.

3.6. Disoccupazione di lunga durata

La disoccupazione di lunga durata è il rapporto percentuale fra coloro che cercano lavoro da più di 12 mesi e il totale dei disoccupati. Anche per quanto riguarda questo indice il Veneto, con un tasso del **38,9%**, si colloca a metà fra i paesi facenti parte del *cluster*. Le regioni che presentano i tassi più elevati sono quelle dell'Italia (in particolare di Piemonte e Toscana) e della Francia (Champagne – Ardenne e Bourgogne), che oscillano fra il 45 e il 55 %.

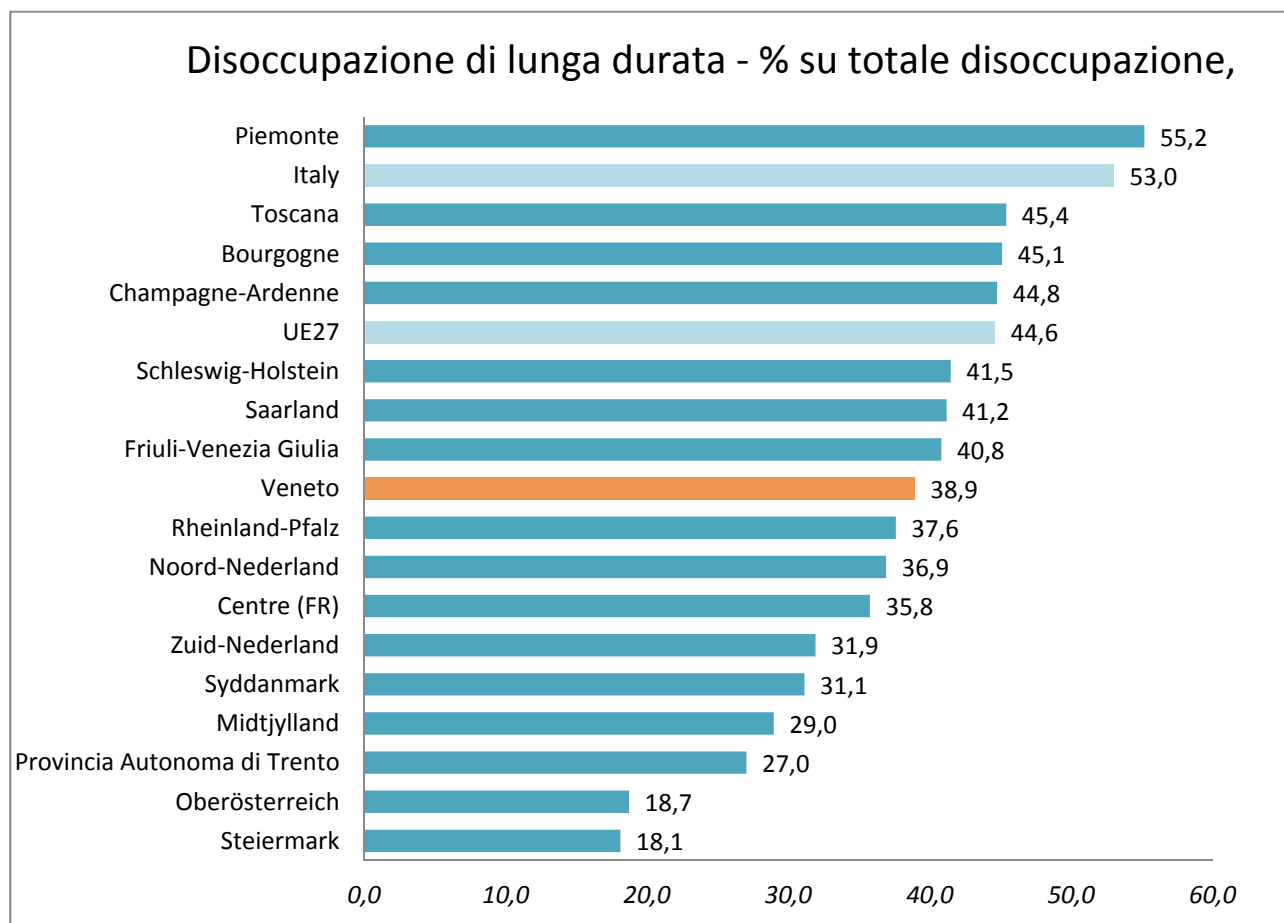


Grafico 20 - Disoccupazione di lunga durata, anno 2012. Elaborazione su dati Eurostat.

3.7. Posizioni apicali, confronto fra generi

Per quanto riguarda questi indici, abbiamo reperito gli ultimi dati Istat nell'Indagine Forza Lavoro (2009). I dati riguardano quindi le regioni italiane e non il cluster di riferimento. Per posizioni dirigenziali si intendono le posizioni di imprenditore, di liberi professionisti e dirigenti. L'indice è costruito come rapporto fra le posizioni dirigenziali occupate dalle donne (o dagli uomini) e il totale delle occupate (o degli occupati).

Per quanto riguarda Il Veneto, osserviamo un tasso di posizioni dirigenziali femminili pari a **3,5%** (3,5 donne su 100 occupate ricoprono ruoli apicali). Il Veneto pertanto si posiziona all'ultimo posto per quanto riguarda la parità fra generi, ben al di sotto della media nazionale, che vede un tasso pari a 5,5%.

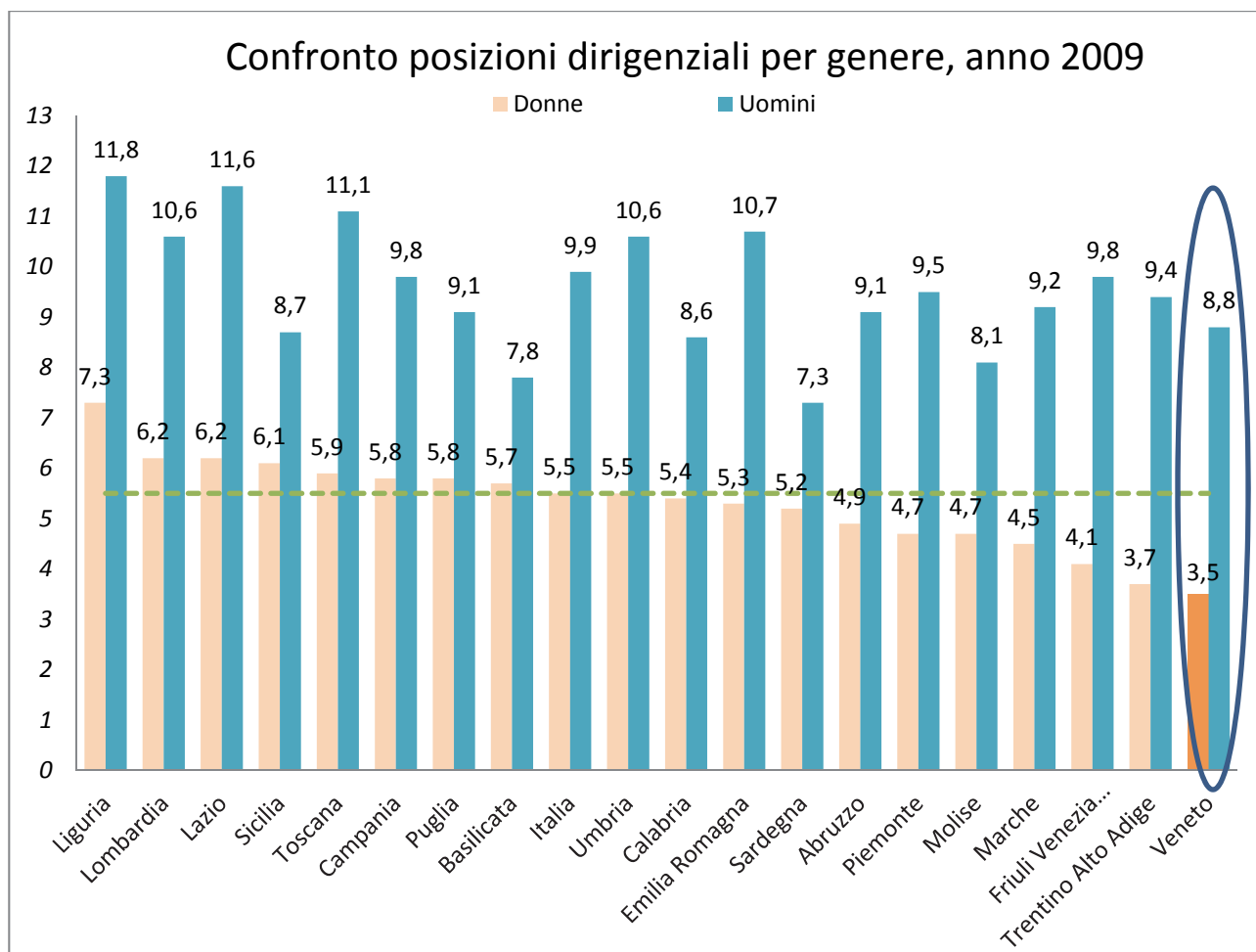


Grafico 21 - Posizioni dirigenziali per genere, anno 2009. Dati relativi alle regioni d'Italia, elaborazione su dati Istat (indagine Forze Lavoro).

4. CAPITALE UMANO

Il settore della formazione è alla base della crescita economica (e non solo) di un paese: riteniamo quindi importante un confronto di vari indicatori, fra cui ritroviamo anche gli indici di formazione terziaria.

Per le statistiche che vengono riportate successivamente, si useranno i seguenti termini per individuare le fasce della formazione determinate dall'Unesco come sistema internazionale di classificazione dei corsi di studio e dei relativi titoli:

- Isced 0: istruzione pre – elementare.
Nel sistema d'istruzione italiano corrisponde alla scuola dell'infanzia.
- Isced 1: istruzione elementare
Nel sistema d'istruzione italiano corrisponde alla scuola primaria.
- Isced 2: istruzione secondaria inferiore.
Nel sistema d'istruzione italiano corrisponde alla scuola secondaria di primo grado.
- Isced 3: istruzione secondaria superiore.
Nel sistema d'istruzione italiano corrisponde alla scuola secondaria di secondo grado.
- Isced 4: istruzione post – secondaria non terziaria.
In Italia ne sono degli esempi i corsi pre – universitari o i corsi brevi professionali.
- Isced 5: primo stadio dell'istruzione terziaria.
Nel sistema d'istruzione italiano corrisponde alla laurea e alla laurea magistrale.
- Isced 6: secondo stadio dell'istruzione terziaria.
Nel sistema d'istruzione italiano corrisponde al dottorato di ricerca.

4.1. Studenti

Di seguito riportiamo il rapporto percentuale fra il totale degli alunni e degli studenti (di tutti i corsi di studio, Isced 0 – 6) e il totale della popolazione. Osserviamo che il Veneto, con un tasso del **17,3%**, risulta essere una delle aree con il minor numero di studenti di tutto il *cluster*, assieme alle altre regioni italiane. Si nota in particolare la differenza con le altre aree, soprattutto con riferimento alle due regioni della Danimarca, che risultano avere una percentuale di alunni e studenti pari, rispettivamente, al 26,2 e al 27,9% della popolazione.

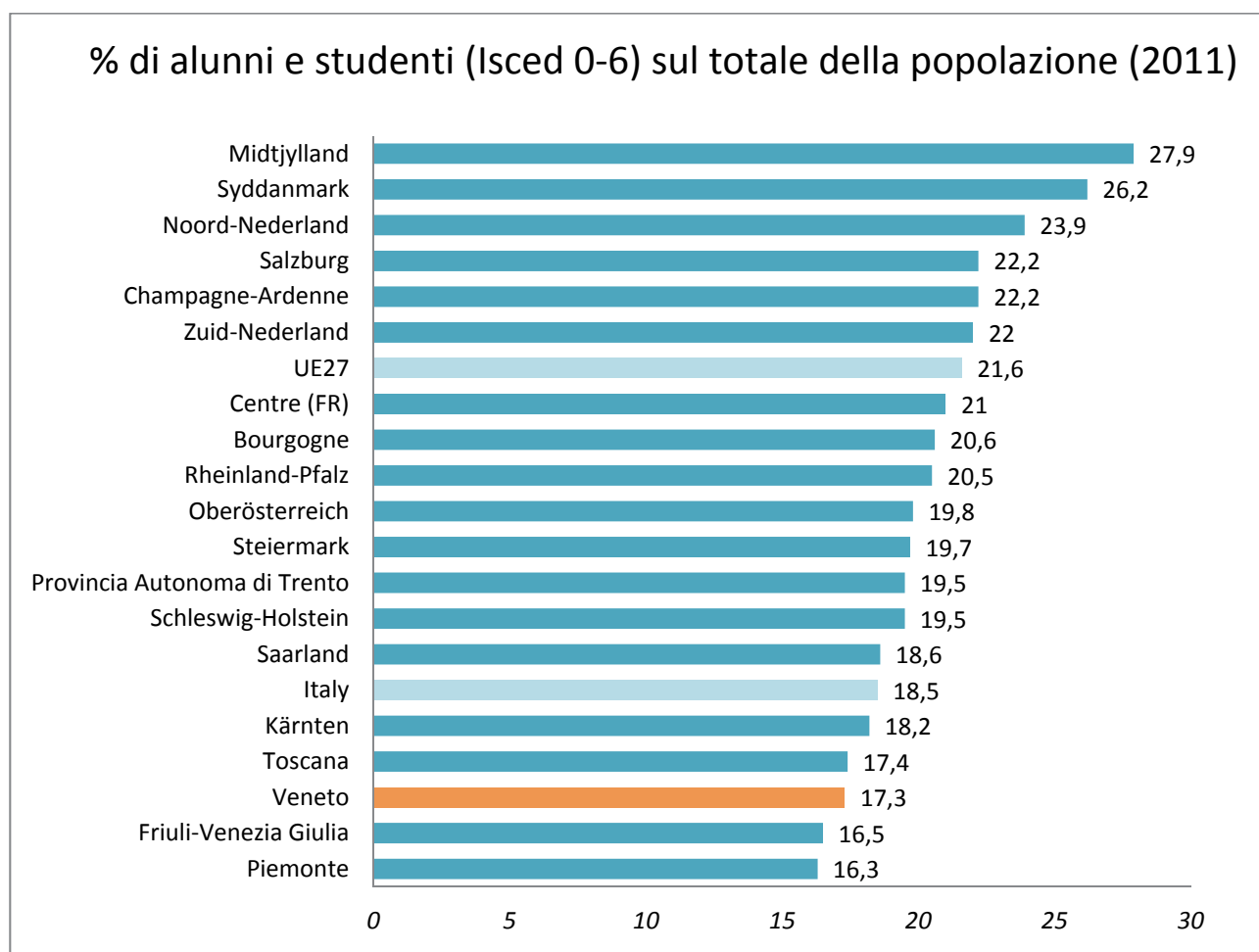


Grafico 22 - Alunni e studenti sul totale della popolazione, anno 2011. Elaborazione su dati Eurostat.

4.2. Studenti di scuola secondaria e post-secondaria non terziaria

Per quanto riguarda gli studenti della scuola secondaria e post-secondaria, che in Italia corrispondono ai percorsi formativi di scuola superiore e ai corsi pre – universitari o ai corsi brevi professionali, il Veneto e in generale l'Italia, con dei tassi rispettivamente pari a **47,4%** e

a 47,2%, si collocano ben al di sopra rispetto alla media dell'Unione europea a 27 paesi, che vede un rapporto fra alunni e studenti della scuola secondaria e post – secondaria e il totale della popolazione in fascia di età 15 – 24 pari a 36,4%.

Questo grafico, messo in relazione con il primo relativo all'istruzione, ci porta ad intuire qual è la classe di istruzione in cui siamo particolarmente carenti: l'istruzione terziaria.

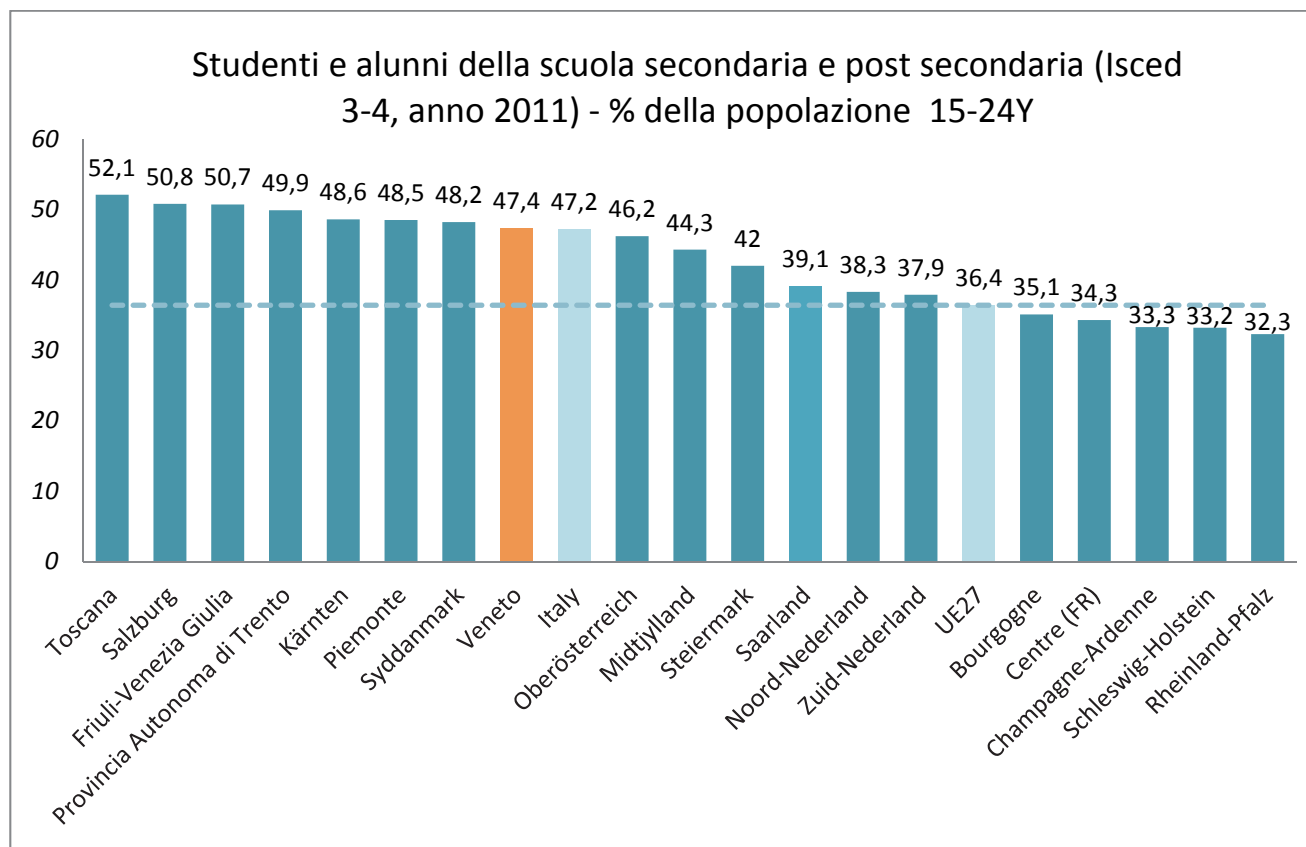


Grafico 23 - Studenti e alunni della scuola secondaria e post secondaria, anno 2011. Elaborazione su dati Eurostat.

4.3. Studenti di scuola terziaria

Per quanto riguarda il rapporto percentuale fra gli studenti della scuola terziaria e il totale della popolazione in classe di età 20 – 24 anni, nel Veneto si osserva un tasso del **49,6%**, un valore ben al di sotto della media dell'Unione europea a 27 paesi e dell'Italia, che registrano un tasso rispettivamente pari a 63,8% e a 62,8%. Elevato il tasso della Toscana, che vede 80 ragazzi su 100 di età 20 – 24 anni impegnati in corsi di studio di livello terziario.

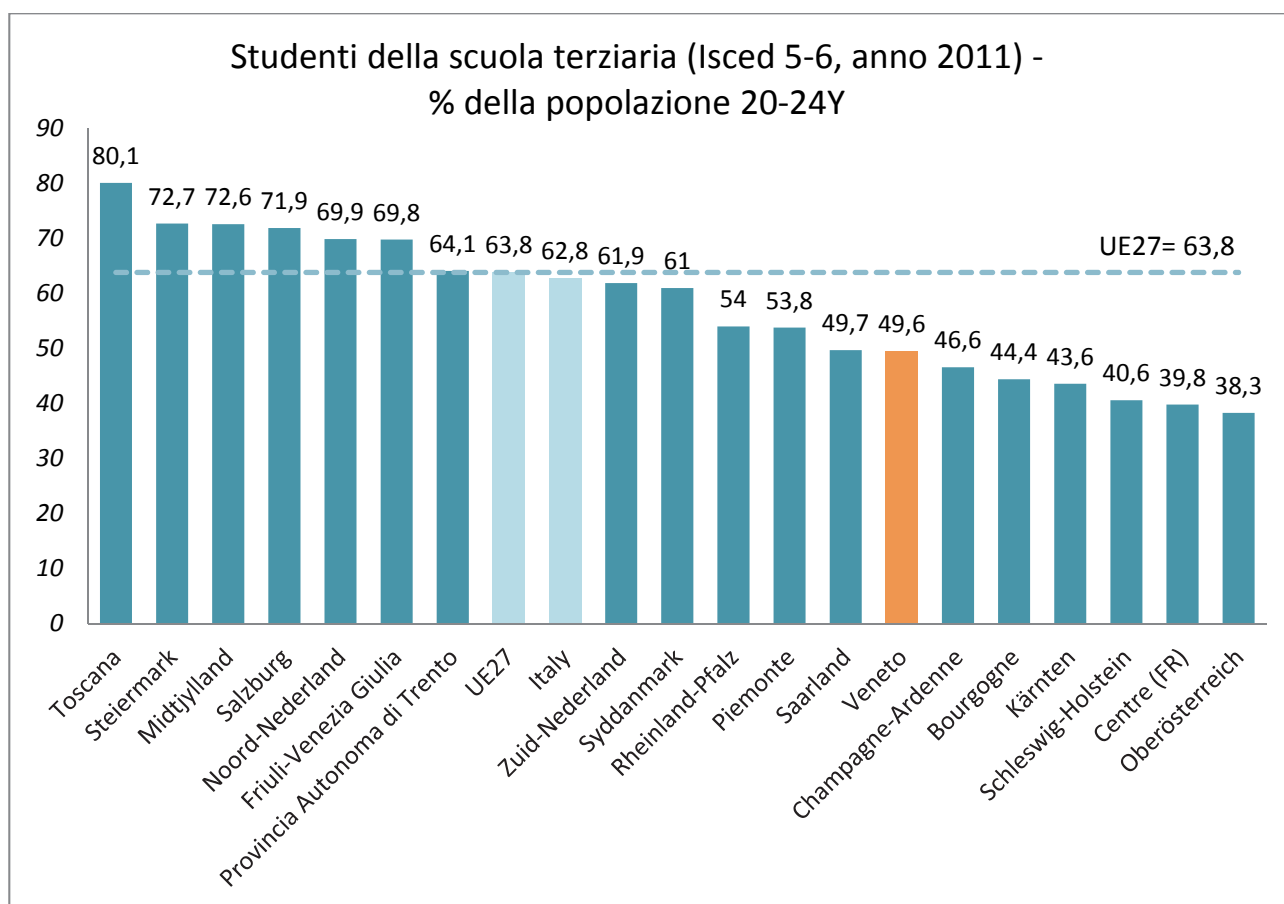


Grafico 24 - Studenti della classe d'istruzione terziaria, anno 2011. Elaborazione su dati Eurostat.

4.4. Educazione terziaria, fascia d'età 25 – 64 anni

Il prossimo grafico illustra la percentuale della popolazione con classe di età 25 – 64 anni che ha ottenuto una educazione terziaria, e cioè che ha conseguito una laurea, o una laurea magistrale o un dottorato di ricerca. La regione Veneto, con un tasso del **13,7%**, risulta essere l'area geografica con il minor tasso di laureati, sia all'interno del *cluster* sia rispetto alle aree più virtuose. Si nota molto la differenza delle due diverse classi geografiche: in particolare notiamo come all'interno del *cluster* la percentuale sia molto più bassa rispetto a quella presente nelle aree più ricche (si guardi per esempio il 42,4% del Sud est dell'Inghilterra o il 41,4% dell'Ile de France), e come invece le aree del nostro *cluster* (ad eccezione delle regioni italiane ed austriache) siano mediamente in linea con la media dell'Unione europea a 27 paesi, che registra un tasso del 27,7%.

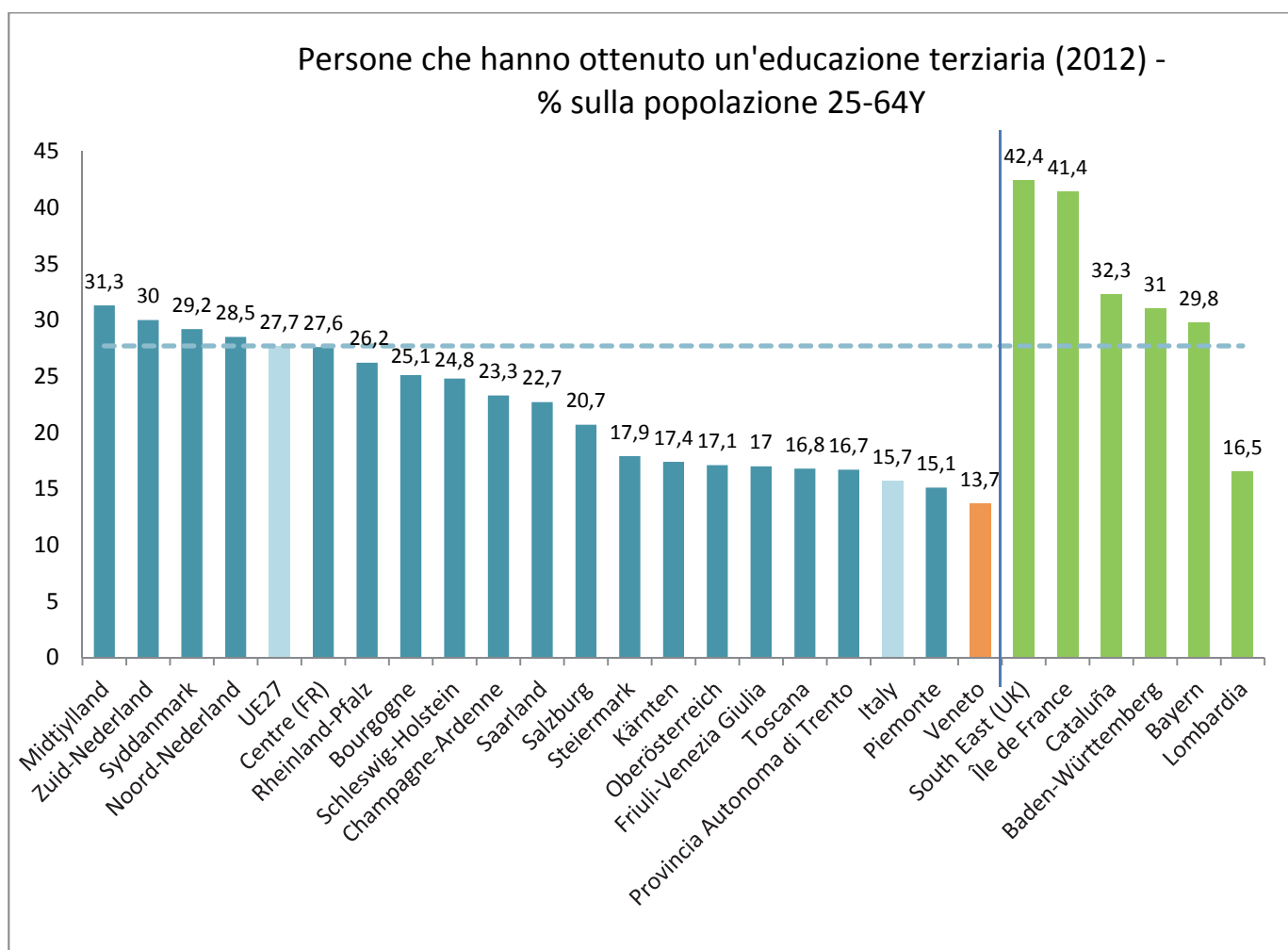


Grafico 25 - Popolazione che ha ottenuto un'educazione di tipo terziario (25 - 64 anni), anno 2012. Elaborazione su dati Eurostat.

4.5. Educazione terziaria, fascia d'età 30 – 34 anni

Nel prossimo grafico illustriamo la percentuale della popolazione in classe d'età 30 – 34 anni che ha ottenuto una laurea, o una laurea magistrale o un dottorato di ricerca. La situazione non cambia molto rispetto al grafico precedente, ad eccezione delle percentuali che sono generalmente un po' più elevate. Osserviamo per il Veneto un tasso del **21,4%**, un valore che, anche per quanto riguarda questo indice, posiziona la nostra regione in coda rispetto a tutte le altre aree, sia del *cluster* di riferimento sia delle aree più virtuose. Anche per quanto riguarda questo indice si nota la differenza fra le due diverse aree geografiche: le aree più ricche hanno infatti dei tassi più alti delle aree del cluster, ad eccezione della Lombardia che, non a caso, è una regione italiana.

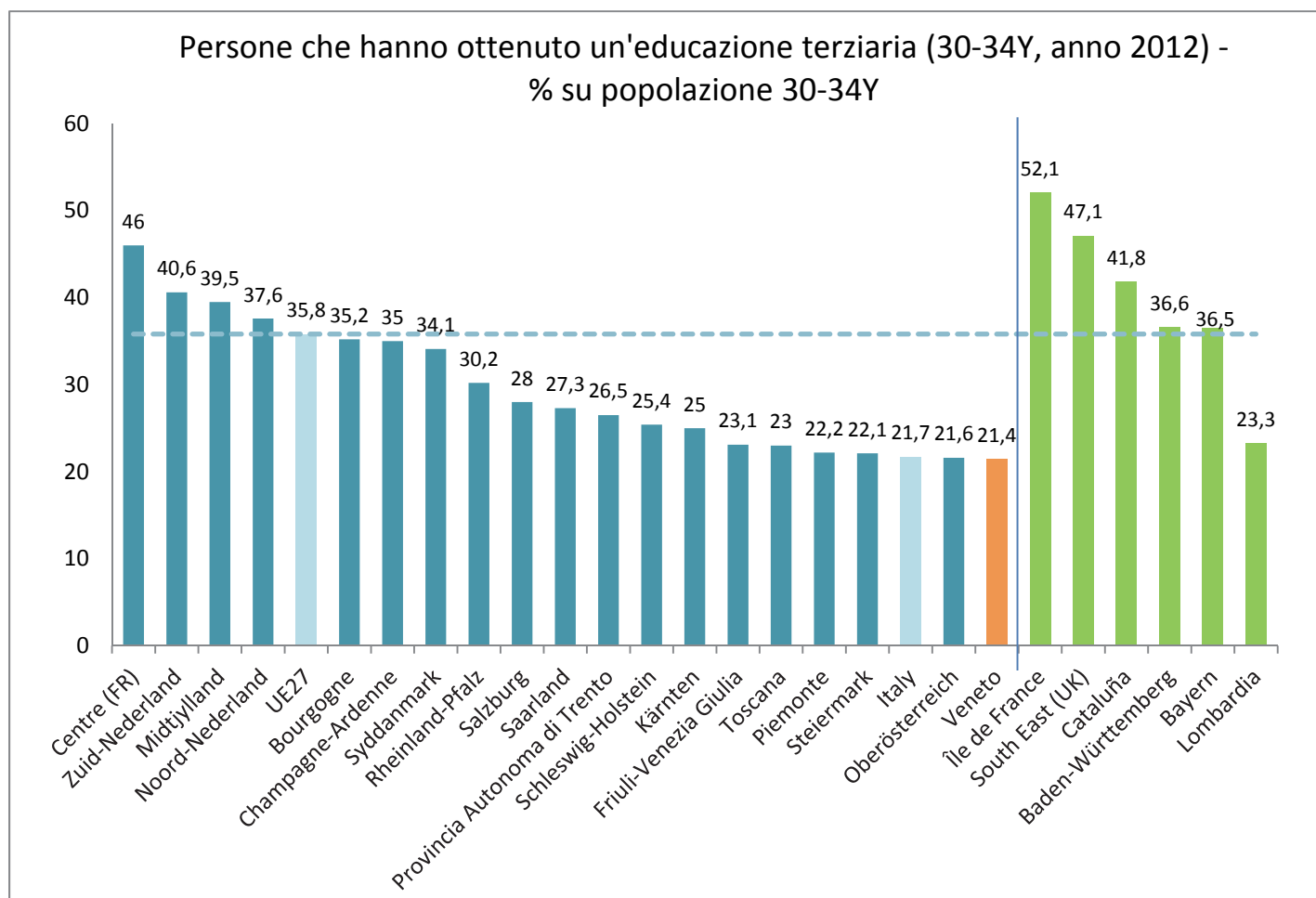


Grafico 26 - Popolazione che ha ottenuto un'educazione terziaria (30 - 34 anni), anno 2012. Elaborazione su dati Eurostat.

4.6. Tasso NEET (“Not in Education, Employment or Training”)

Il tasso *NEET* ci informa sulla percentuale dei giovani (classe di età 18 – 24 anni) che non lavorano, non studiano e non sono inseriti in nessun programma di formazione. Il dato è abbastanza preoccupante, poiché osserviamo per il Veneto un tasso *neet* pari a **20,7%**, inferiore comunque all'Italia che presenta un tasso del 27%. Questo dato non rassicura e osserviamo una grande differenza fra il tasso *neet* che riguarda l'Italia e le regioni italiane, e il tasso *neet* che invece riguarda le regioni dell'Austria, dei Paesi Bassi e della Germania.

**Persone che non lavorano e non studiano (18-24Y, anno 2012) -
% sul totale 18-24Y**

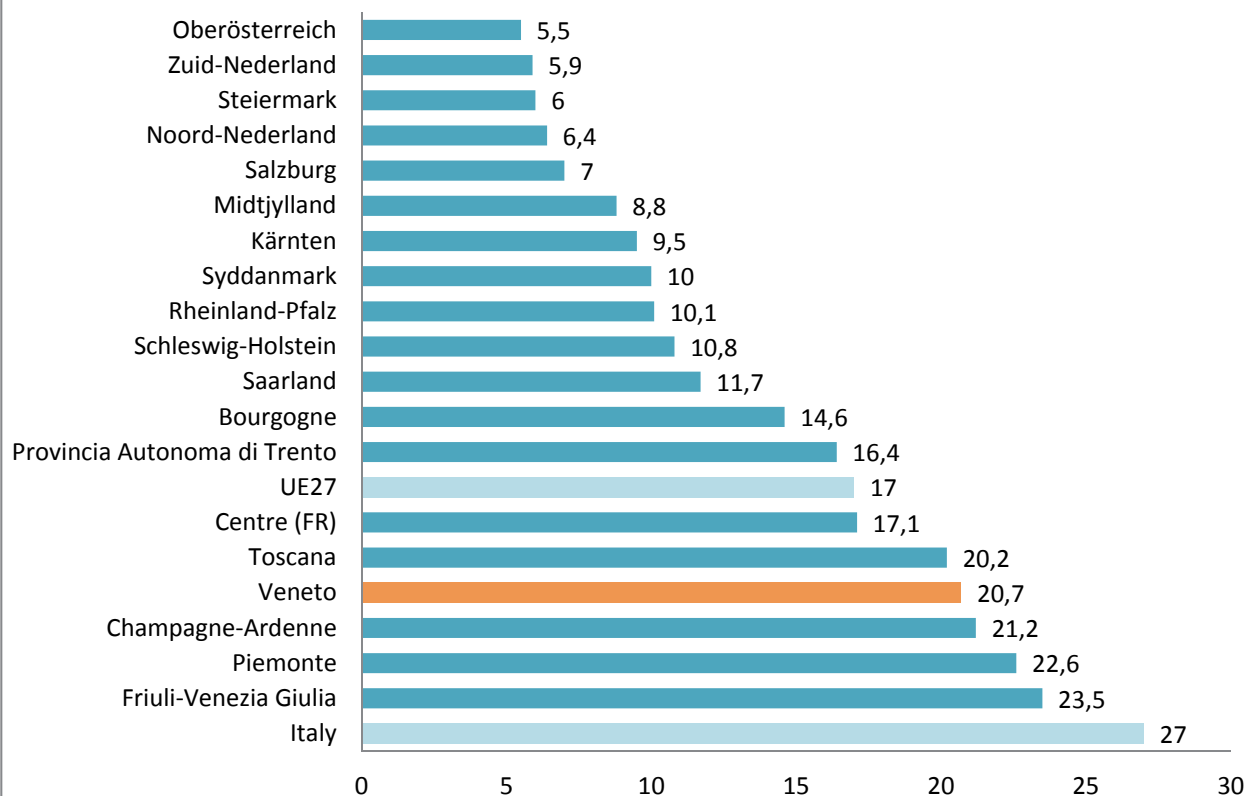


Grafico 27 - Popolazione che non lavora e non studia, classe di età 18 - 24 anni. Anno 2012. Elaborazione su dati Eurostat.

5. RICERCA E SVILUPPO

5.1. Spesa R&S in percentuale a Pil

La spesa per ricerca e sviluppo di un determinato paese ci informa sulla quota di Pil che viene destinata per creare innovazione nel futuro. Essa può essere sia pubblica sia privata. L'indicatore viene anche chiamato *GERD* (acronimo di *Gross domestic expenditure on R&D*) come percentuale del Pil, ed include tutti i settori (il settore delle imprese, il settore pubblico, il settore dell'educazione terziaria e il settore del non-profit privato).

Osserviamo che il Veneto destina l'**1,07%** del suo Pil alla R&S, quota nettamente inferiore rispetto alle aree del cluster di riferimento e rispetto alle quote investite nelle regioni più virtuose. L'Unione Europea a 27 paesi presenta invece una quota del 2,02%, mentre l'Italia destina solo l'1,26% del Pil in R&S.

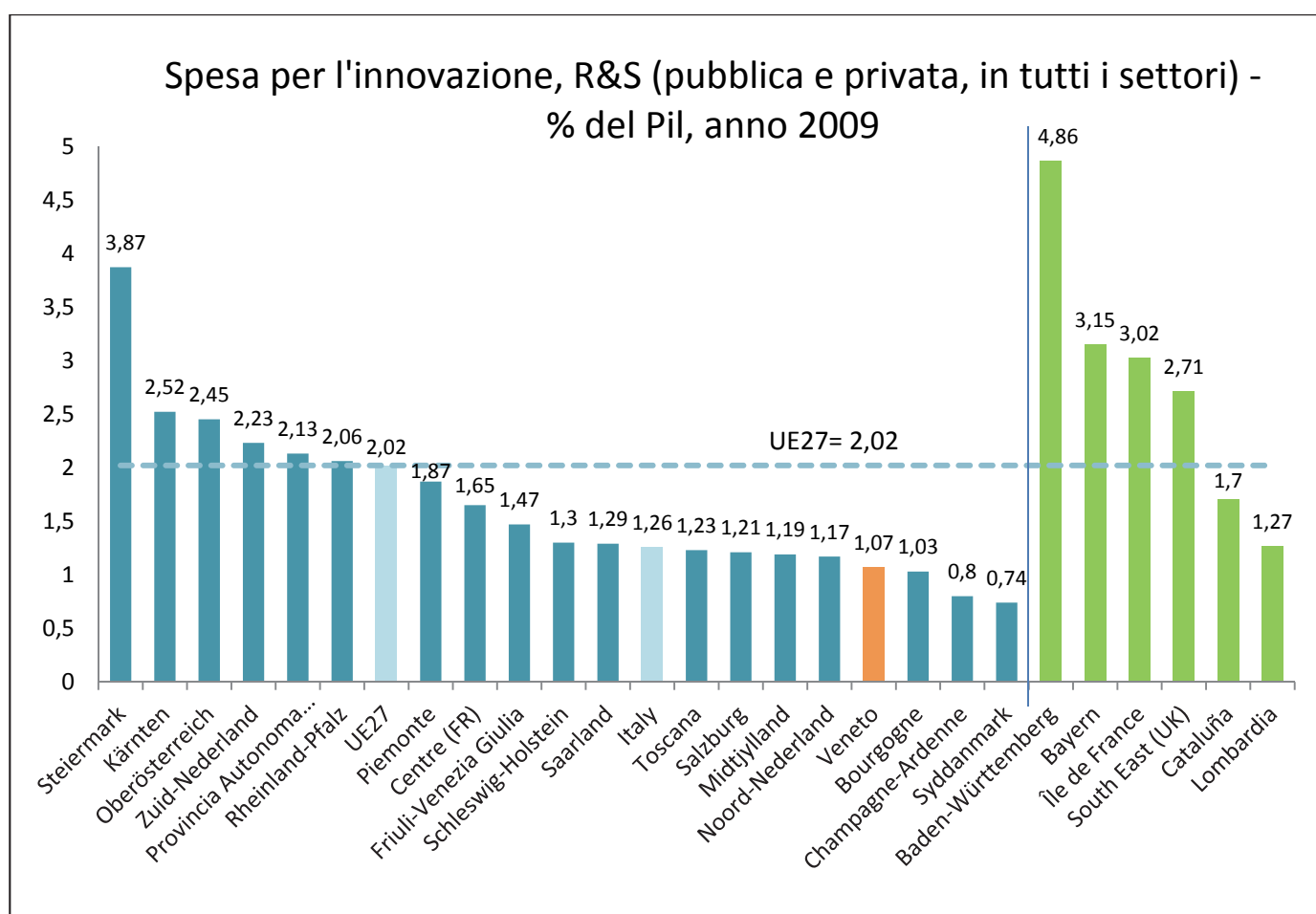


Grafico 28 - Spesa per ricerca e sviluppo, % del Pil, anno 2009. Elaborazione su dati Eurostat.

5.2. Spesa R&S pro capite

Il grafico successivo illustra la spesa per ricerca e sviluppo pro capite. La situazione non è molto diversa da quella prospettata dal grafico precedente, e osserviamo per il Veneto un valore di spesa per R&S di **313€** per abitante, al di sotto rispetto alla linea tratteggiata che illustra la situazione dell'Unione Europea a 27 paesi, la quale registra un valore pari a 474€ per abitante. Anche qui osserviamo che la quota destinata alla ricerca e allo sviluppo nella nostra regione appare nettamente minore rispetto a quelle che si registrano nelle regioni del cluster e nelle regioni più ricche.

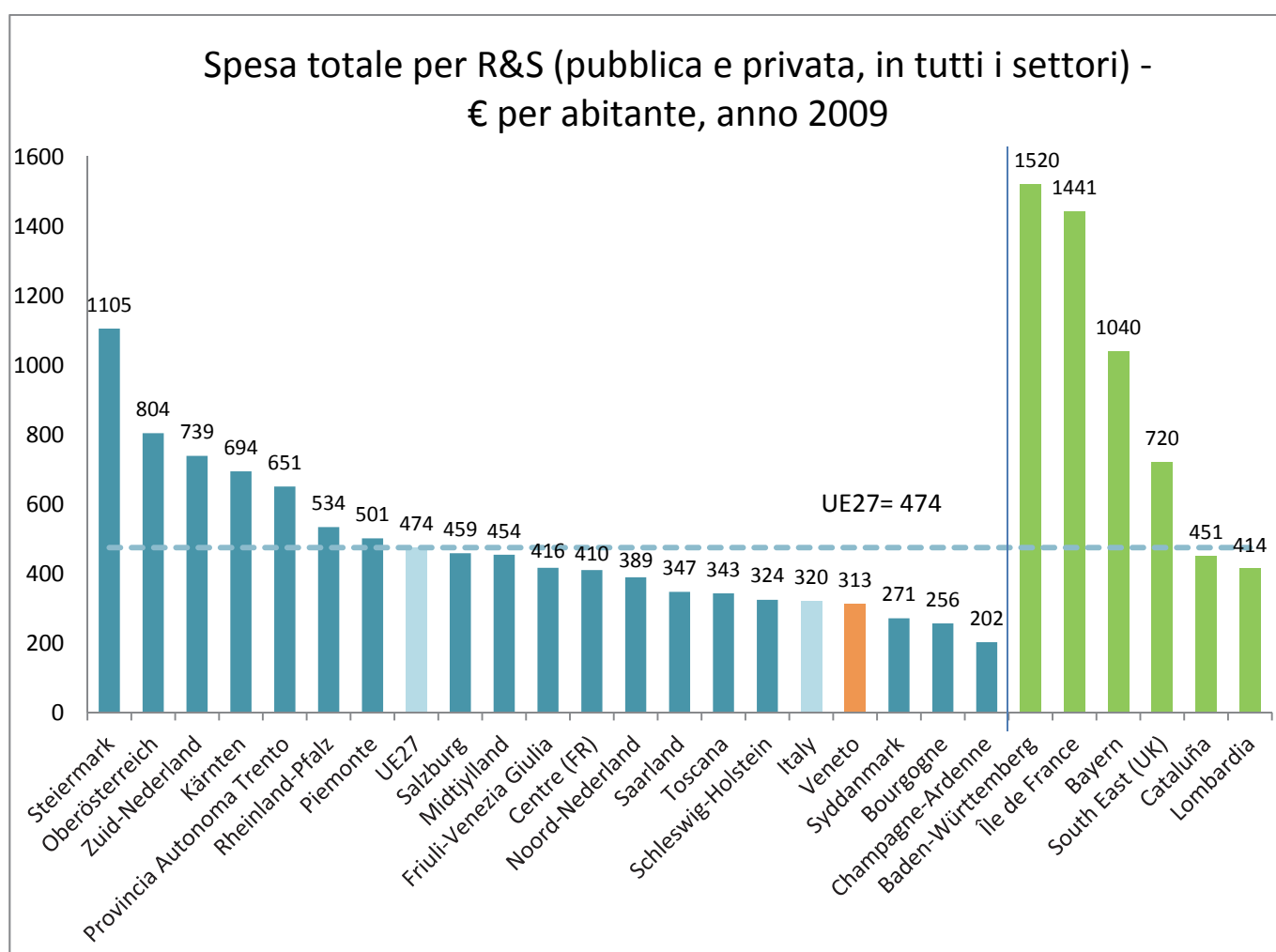


Grafico 29 - Spesa per ricerca e sviluppo, € per abitante, anno 2009. Elaborazione su dati Eurostat.

5.3. Ricercatori e personale R&S

Il prossimo grafico illustra la situazione occupazionale di ricercatori e personale R&S, come percentuale dell'occupazione totale. La posizione del Veneto in questo caso è in linea con la situazione dell'Unione Europea a 27 paesi, registrando un valore pari a **1,71%**. Anche in questo grafico si nota la differenza fra le due diverse aree: all'interno del cluster la maggior parte delle regioni prese in considerazione hanno una percentuale inferiore rispetto all'Unione Europea a 27 paesi, mentre non è lo stesso per le aree più virtuose (ad eccezione della Lombardia che registra un tasso dell'1,65%).

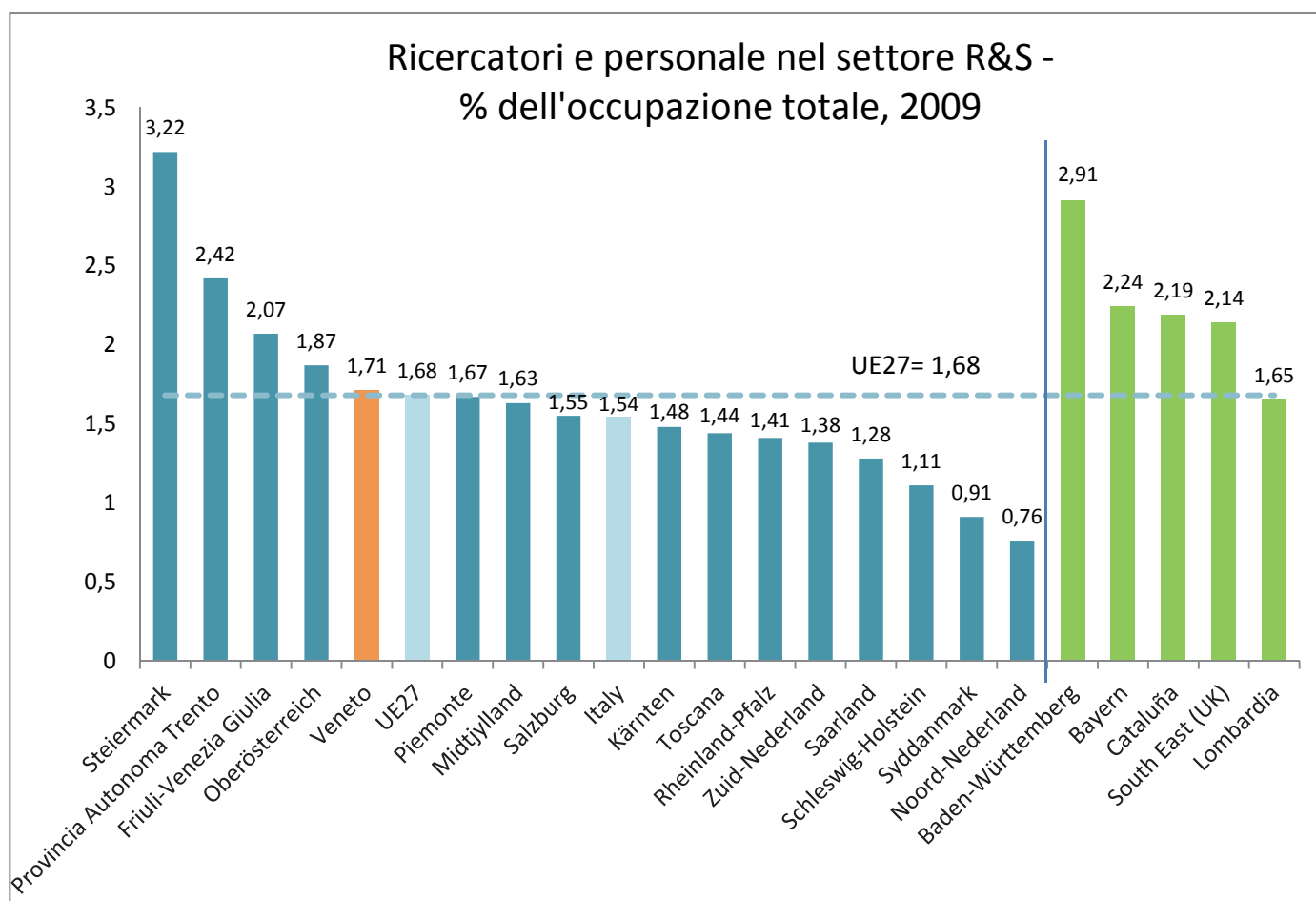


Grafico 30 - Occupazione di ricercatori e personale del settore R&S, anno 2009. Elaborazione su dati Eurostat.

5.4. Ricercatori

Per quanto riguarda il dato occupazionale dei ricercatori, la situazione è diversa rispetto a quella prospettata dal grafico precedente. In particolare per quanto riguarda il Veneto osserviamo una quota occupazionale di ricercatori pari allo **0,39%** dell'occupazione totale,

una delle quote più basse con riferimento sia alle regioni appartenenti al cluster sia alle regioni più ricche.

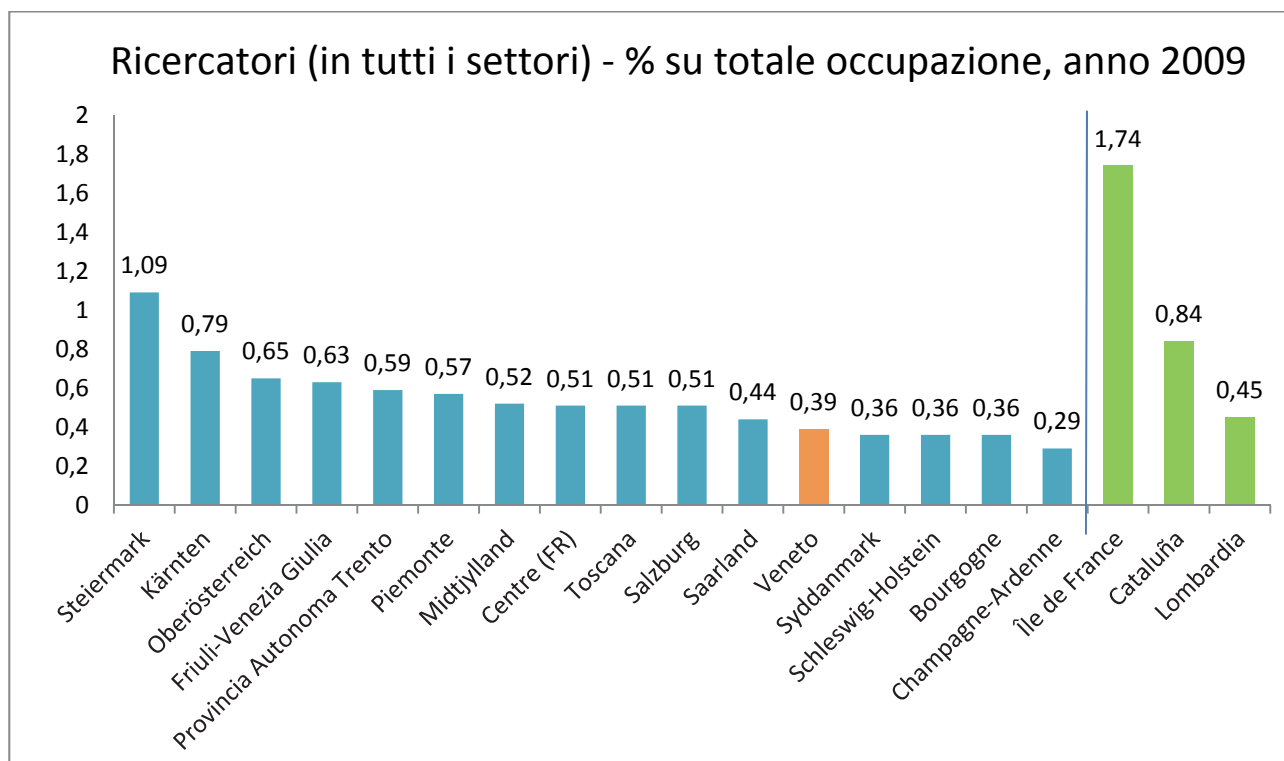


Grafico 31 - Situazione occupazionale dei ricercatori, anno 2009. Elaborazione su dati Eurostat.

5.5. Occupazione nei settori ad alta tecnologia

Per settori ad alta tecnologia si intendono sia i settori manifatturieri ad alta tecnologia (che comprendono il settore aerospaziale, farmaceutico, informatico, elettronico, delle telecomunicazioni e quello degli strumenti scientifici) sia i settori dei servizi a conoscenza intensiva (i cosiddetti *KIS, knowledge-intensive services*), che ricomprendono il settore delle telecomunicazioni, dell'informatica, della ricerca e sviluppo, dei trasporti d'acqua e d'aria, finanziario, immobiliare, assicurativo, dell'educazione, della salute e delle attività ricreative, culturali e sportive. Con riferimento a queste fattispecie la regione Veneto con un tasso del **4,65%** si presenta come la seconda regione all'interno del cluster di riferimento, seconda solo al Piemonte che presenta una percentuale del 5,11%.

Queste sono comunque quote in linea con i tassi osservati nei paesi del cluster, ad eccezione di due regioni della Francia, Bourgogne e Champagne – Ardenne, che risultano essere in coda con dei tassi rispettivamente pari a 2,33% e 2,30%.

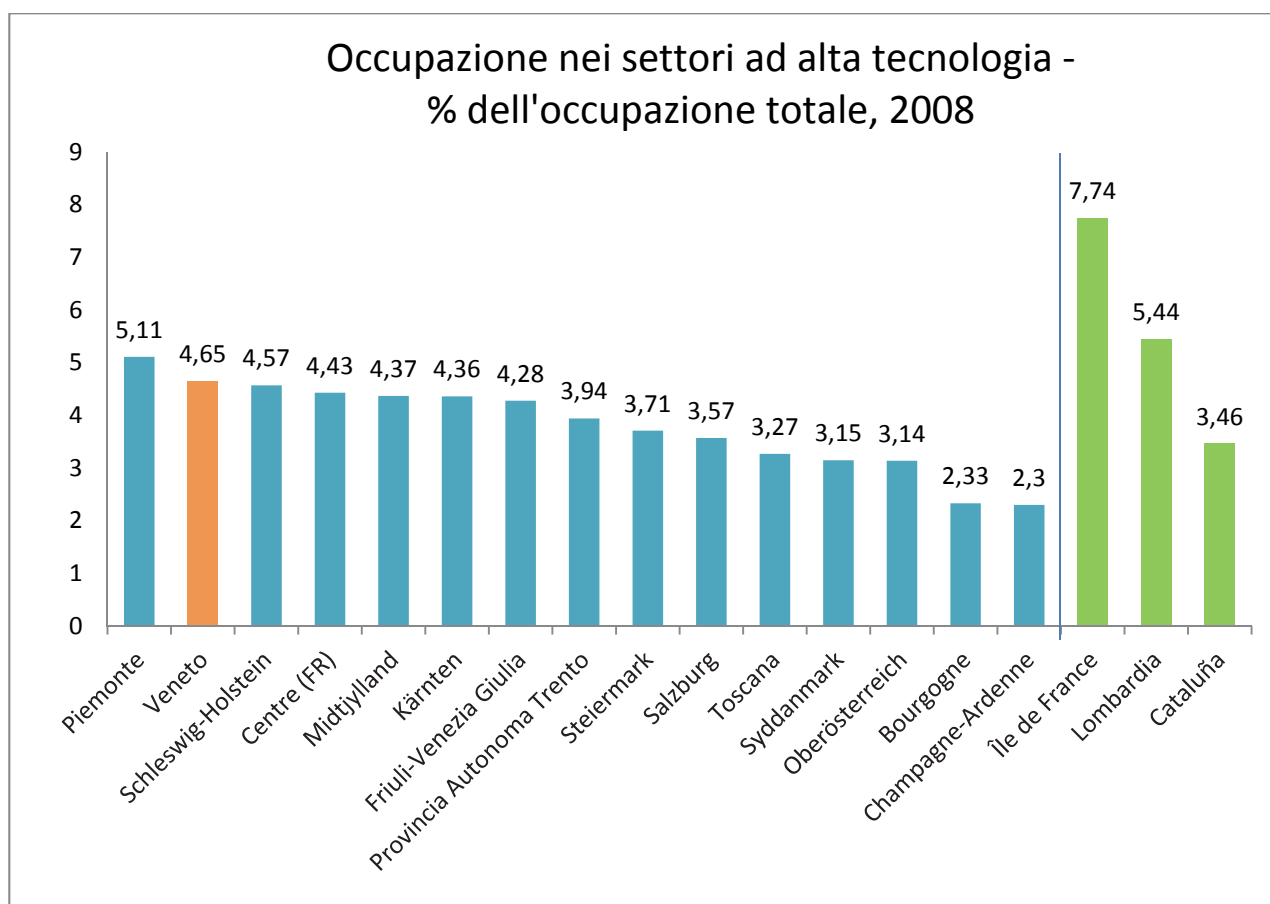


Grafico 32 - Occupazione nei settori ad alta tecnologia, anno 2008. Elaborazione su dati Eurostat.

5.6. Brevetti presentati all'EPO (European Patent Office)

Con riferimento alle richieste che sono pervenute all'ufficio brevetti europeo, la regione Veneto non si posiziona fra le regioni più "innovatrici", e presenta un numero pari a circa **101** richieste per un milione di abitanti. Le regioni austriache risultano essere quelle più dinamiche, con un numero di richieste di brevetti (per un milione di abitanti) che va da 222 a 162. Per quanto riguarda le aree che non fanno parte del cluster, non si registrano grandi differenze.

Con riferimento alla variazione delle richieste rispetto all'anno 2005, osserviamo che la maggior parte delle regioni hanno registrato pesanti flessioni negative, al di fuori delle quattro regioni austriache, che invece hanno visto un aumento delle richieste nell'anno 2009 rispetto all'anno 2005. Il Veneto non è comunque in una buona situazione, avendo una variazione rispetto al 2005 del **-28,7%**.

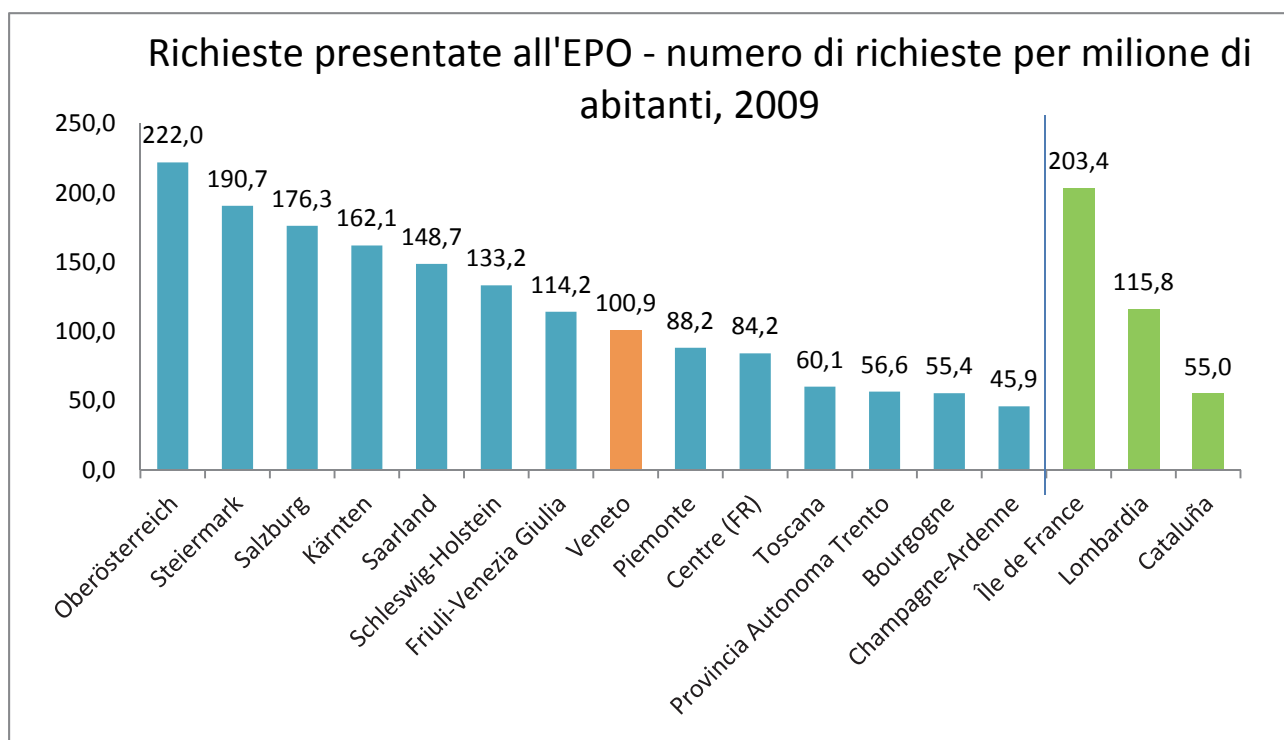


Grafico 33 - Numero di richieste brevetti per un milione di abitanti, anno 2009. Elaborazione su dati Eurostat.

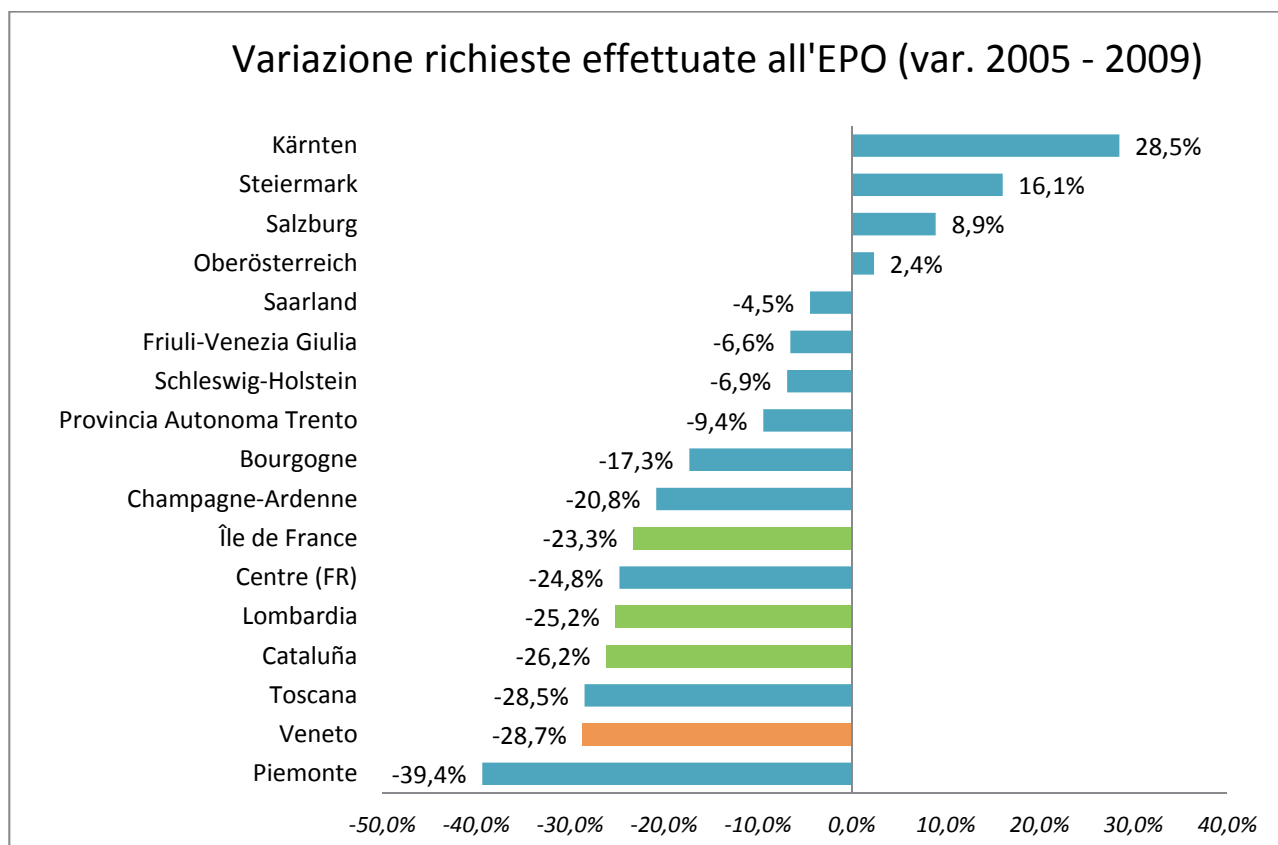


Grafico 34 - Variazione delle richieste effettuate all'EPO, anno 2009 rispetto al 2005. Elaborazione su dati Eurostat.

6. ALTRE INFORMAZIONI

6.1. Persone a rischio povertà

Per tasso di rischio di povertà si intende la percentuale della popolazione con un reddito disponibile uguale o inferiore alla soglia di rischio di povertà (pari al 60% del reddito disponibile mediano equivalente nazionale, espresso in standard di potere d'acquisto). Per quanto riguarda il Veneto, osserviamo un tasso del **15,9%**, un tasso non troppo elevato se paragonato alle altre aree del *cluster*. Si rileva per quanto riguarda l'Italia un tasso molto elevato, pari a 28,2%. Questo valore è giustificato da una presenza molto forte del rischio di povertà nelle regioni del Sud, dove il rischio di povertà medio è pari a 44,9% e in alcune regioni arriva fino a 54,6% (Sicilia).

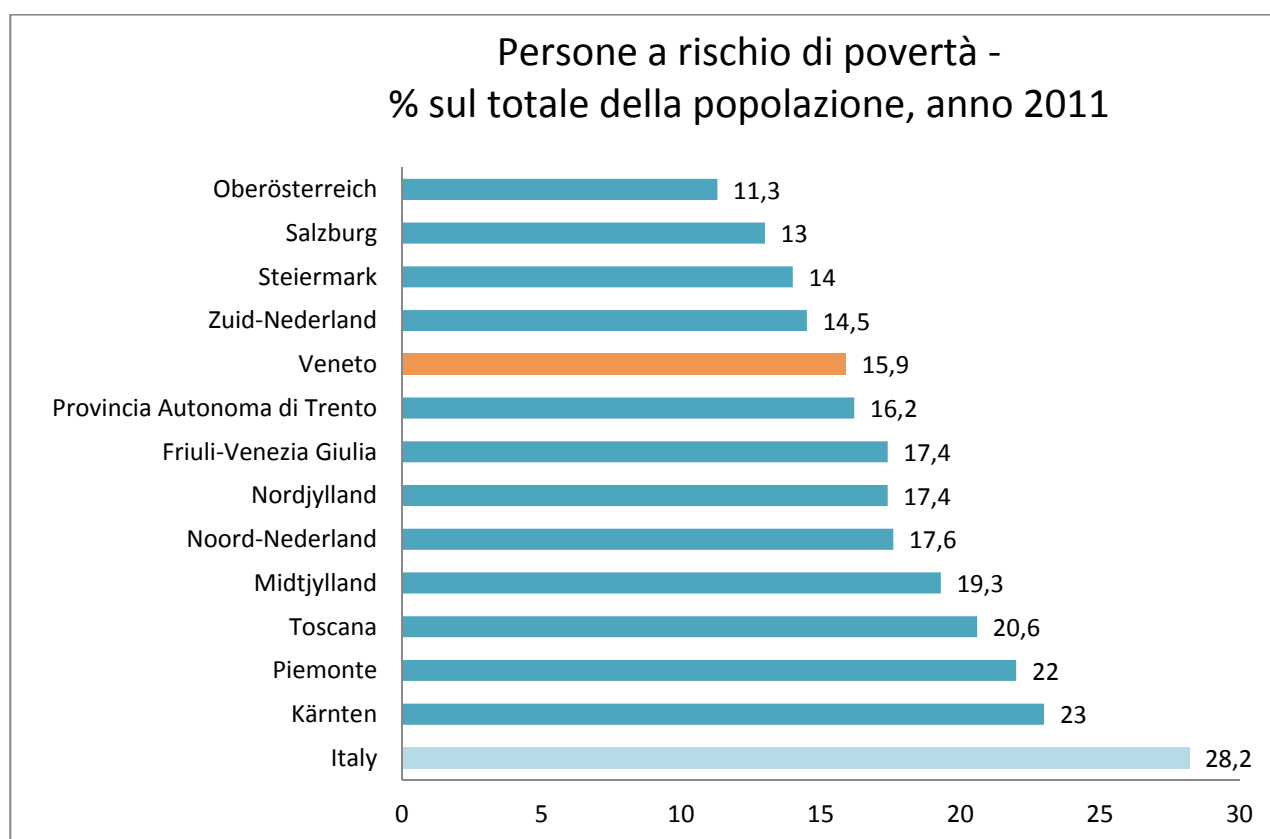


Grafico 35 - Popolazione a rischio di povertà, anno 2011. Elaborazione su dati Eurostat.

6.2. Accesso ad internet

Con riferimento ad internet, osserviamo che la percentuale delle unità abitative che hanno un accesso alla rete per la regione Veneto è ancora abbastanza bassa, pari a **63%**, un tasso in linea con l'Italia. I tassi più elevati li troviamo nelle due regioni della Danimarca, Midtjylland e Syddanmark, che presentano dei tassi rispettivamente pari a 91% e 86%.

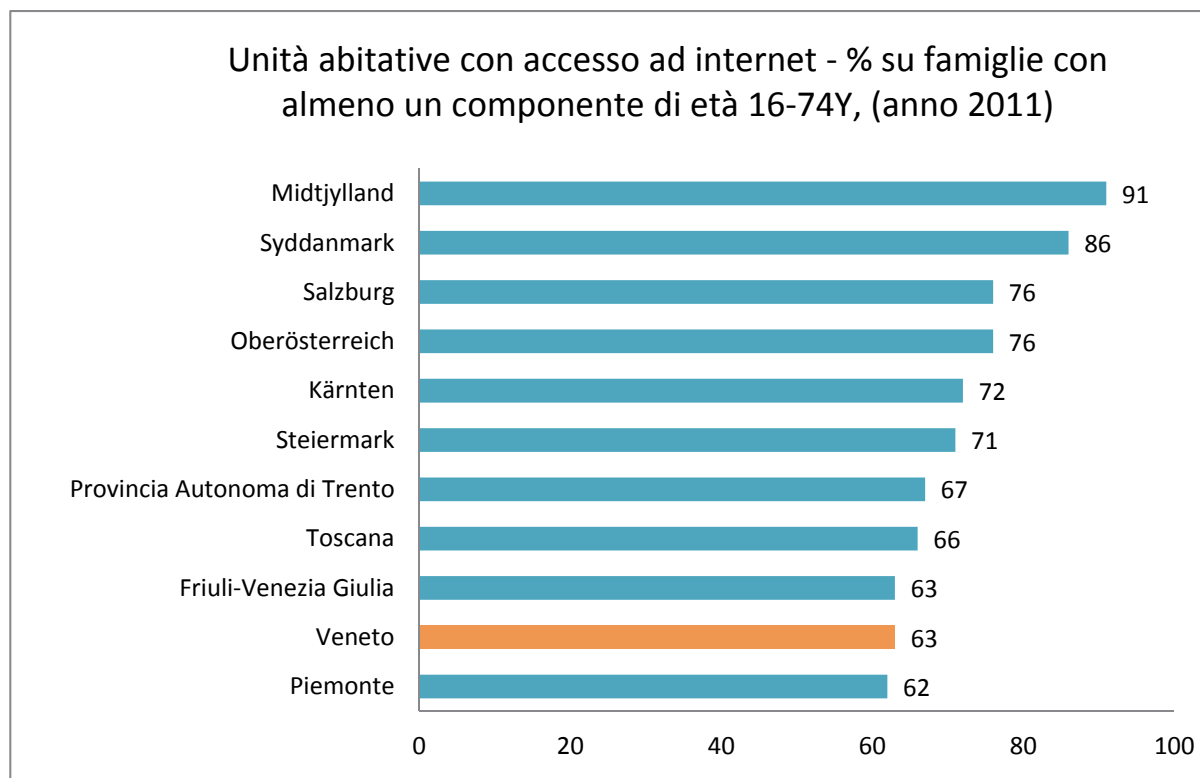


Grafico 36 - Unità abitative con accesso ad internet, anno 2011. Elaborazione su dati Eurostat.

6.3. Utilizzo di internet

Per quanto riguarda l'utilizzo della rete, nel prossimo grafico si illustrano le percentuali per regione delle persone che abitualmente usano la rete (per utilizzo regolare si intende almeno una volta a settimana). Si avverte fin da subito la differenza fra le percentuali che caratterizzano le regioni italiane e tutte le altre regioni del cluster. In particolare per quanto riguarda la regione Veneto, si osserva un tasso del **58%**, un valore che ci vede ben lontani dalla situazione dei Paesi Bassi, caratterizzata dalla quasi totalità della popolazione che usa regolarmente la rete, e da tutti gli altri paesi.

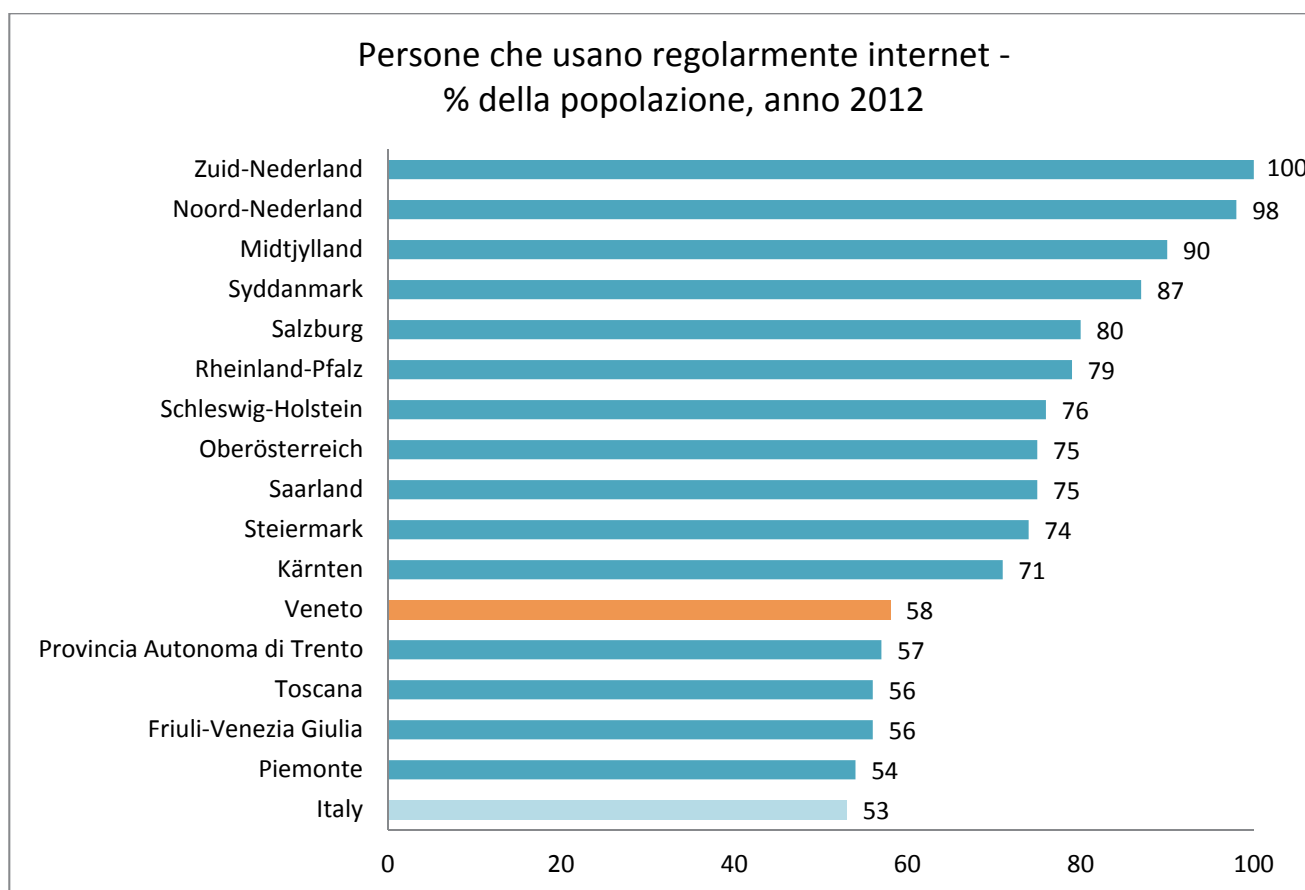


Grafico 37 - Popolazione che usa internet regolarmente (almeno una volta a settimana), anno 2012. Elaborazione su dati Eurostat.

6.4. Estraneità ad internet

Anche per quanto riguarda il grafico successivo si è in una situazione analoga a quella del precedente. Nel Veneto il **33%** della popolazione (una persona su 3!) non ha mai utilizzato internet. Per quanto riguarda l'Italia il dato è peggiore, si sale al 37% della popolazione. Anche in questo caso, e forse ancor di più in questo caso, si può cogliere immediatamente la profonda e netta differenza che intercorre fra le regioni italiane e le altre regioni del cluster.

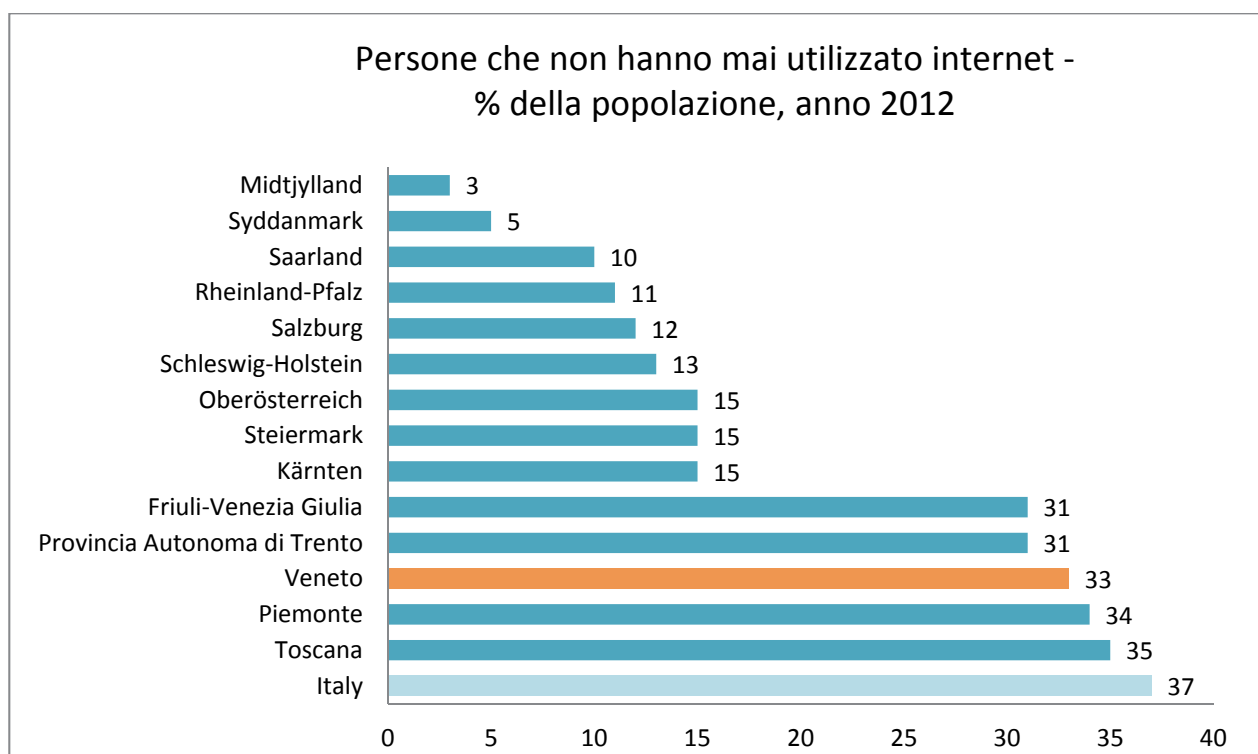


Grafico 38 - Popolazione che non ha mai utilizzato internet, anno 2012. Elaborazione su dati Eurostat.

6.5. Inglese

Per quanto riguarda la lingua inglese, non siamo riusciti a trovare dei dati relativi alle aree del nostro cluster, quindi proponiamo delle statistiche riferite ai 25 paesi d'Europa (si escludono dal confronto l'Irlanda e il Regno Unito dove l'inglese è lingua ufficiale) e alle regioni d'Italia. Con la dicitura "Unione Europea a 27 paesi" si intendono dei valori ottenuti sempre escludendo le persone di madrelingua inglese.

Per quanto riguarda la conversazione, il 34% degli italiani pensa di essere in grado di sostenere una conversazione di inglese, un valore molto basso se pensiamo anche al fatto che questo valore proviene da autodichiarazioni. L'Unione Europea a 27 paesi presenta un tasso del 38%, mentre per quanto riguarda alcuni paesi, in particolare Finlandia, Austria, Cipro, Svezia, Danimarca, Malta e Paesi Bassi, le percentuali aumentano molto, e si è in un range del 70-90%.

Per quanto riguarda la lettura, osserviamo che il 26% degli italiani dichiara di comprendere l'inglese leggendo giornali o articoli in lingua, a fronte di un tasso pari al 25% per l'Unione Europea a 27 paesi.

Entrambe le statistiche fanno comprendere il nostro stato per quanto riguarda la lingua inglese, lingua di cui la conoscenza è essenziale soprattutto se l'obiettivo è attrarre fondi comunitari per la realizzazione di progetti nel nostro territorio.

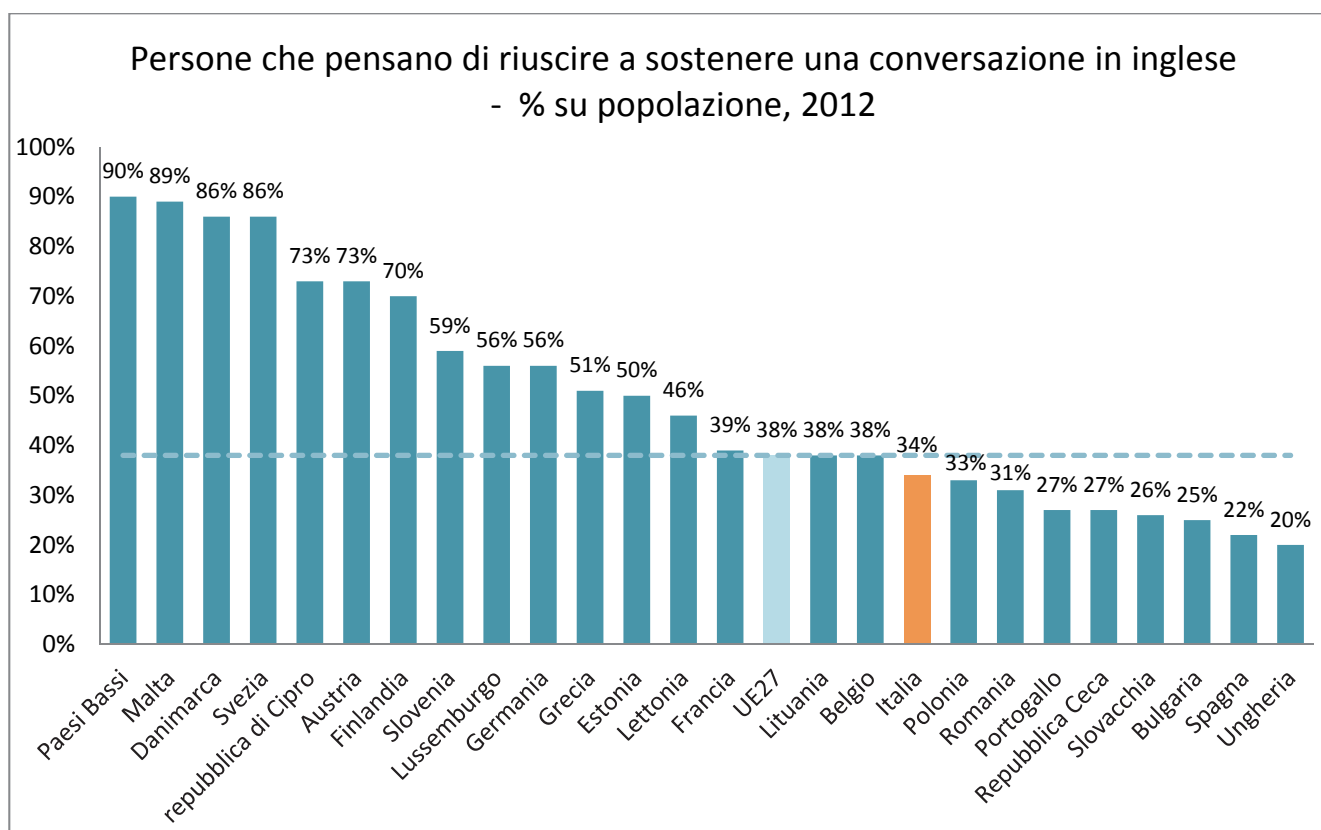


Grafico 39 - Persone che pensano di essere in grado di sostenere una conversazione in inglese, anno 2012.
Elaborazione su dati Eurostat (Eurobarometer).

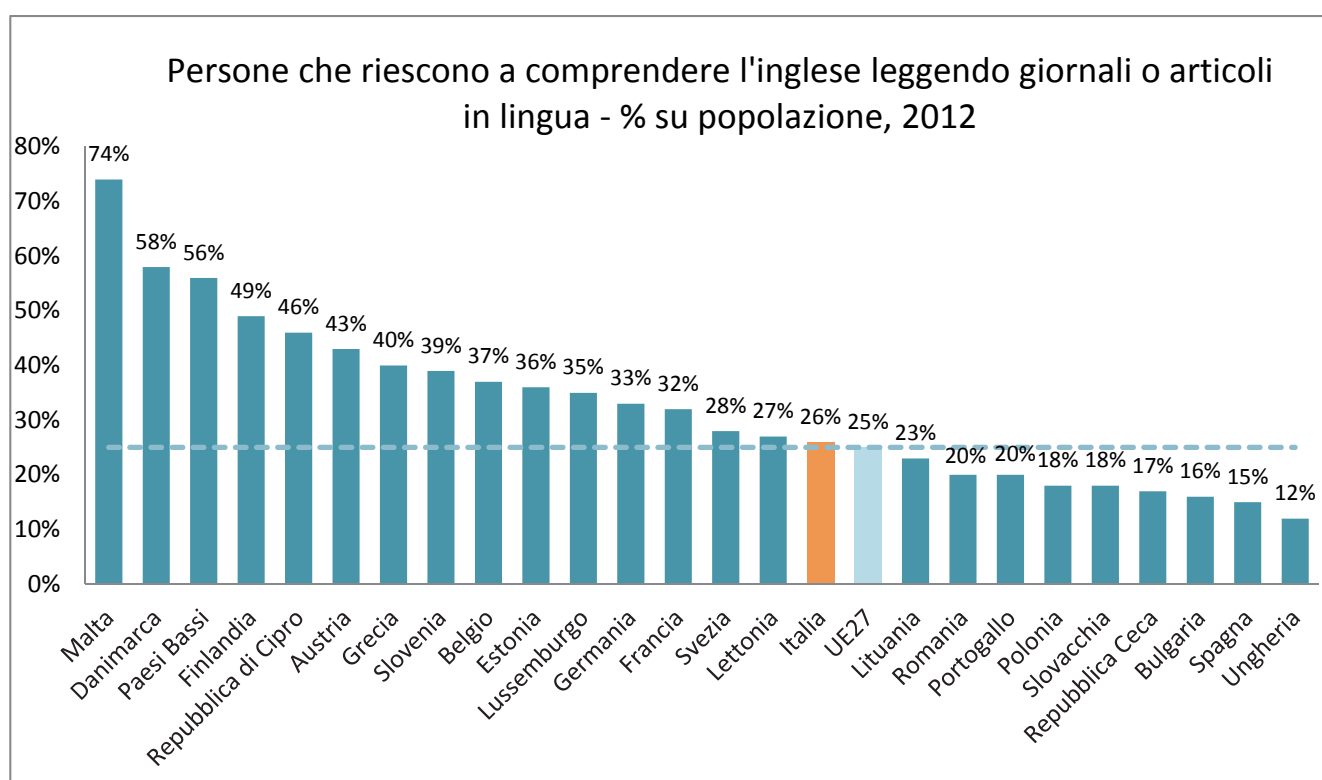


Grafico 40 - Persone che riescono a comprendere la lingua inglese leggendo giornali o articoli, anno 2012.
Elaborazione su dati Eurostat (Eurobarometer).

7. Altre statistiche di confronto

Si propongono di seguito alcune statistiche per le quali non è stato possibile reperire il dato relativo alle regioni del cluster, ma bensì esse sono riferite ai paesi dell'Unione Europea a 27 paesi o alle regioni d'Italia.

7.1. Cittadini e minori stranieri

Per quanto riguarda i cittadini stranieri residenti come percentuale sulla popolazione residente, l'Italia presenta un tasso del 7,5%. Per la regione Veneto si osserva un valore pari a **10,2%**, e questo ci porta a concludere che siamo una delle regioni con la concentrazione maggiore di cittadini stranieri, assieme a Lombardia, Umbria ed Emilia Romagna.

Se guardiamo il grafico successivo però, vediamo anche qual è la percentuale di minori sugli stranieri residenti, e per quanto riguarda il Veneto, osserviamo un valore pari a **24%**. Questi due grafici vanno letti insieme alla variazione della popolazione (che non a caso vede il Veneto come una delle regioni con una maggior dinamica positiva) e insieme ai tassi di dipendenza, anche se il tasso di dipendenza che deve preoccuparci di più non è quello infantile, ma quello senile.

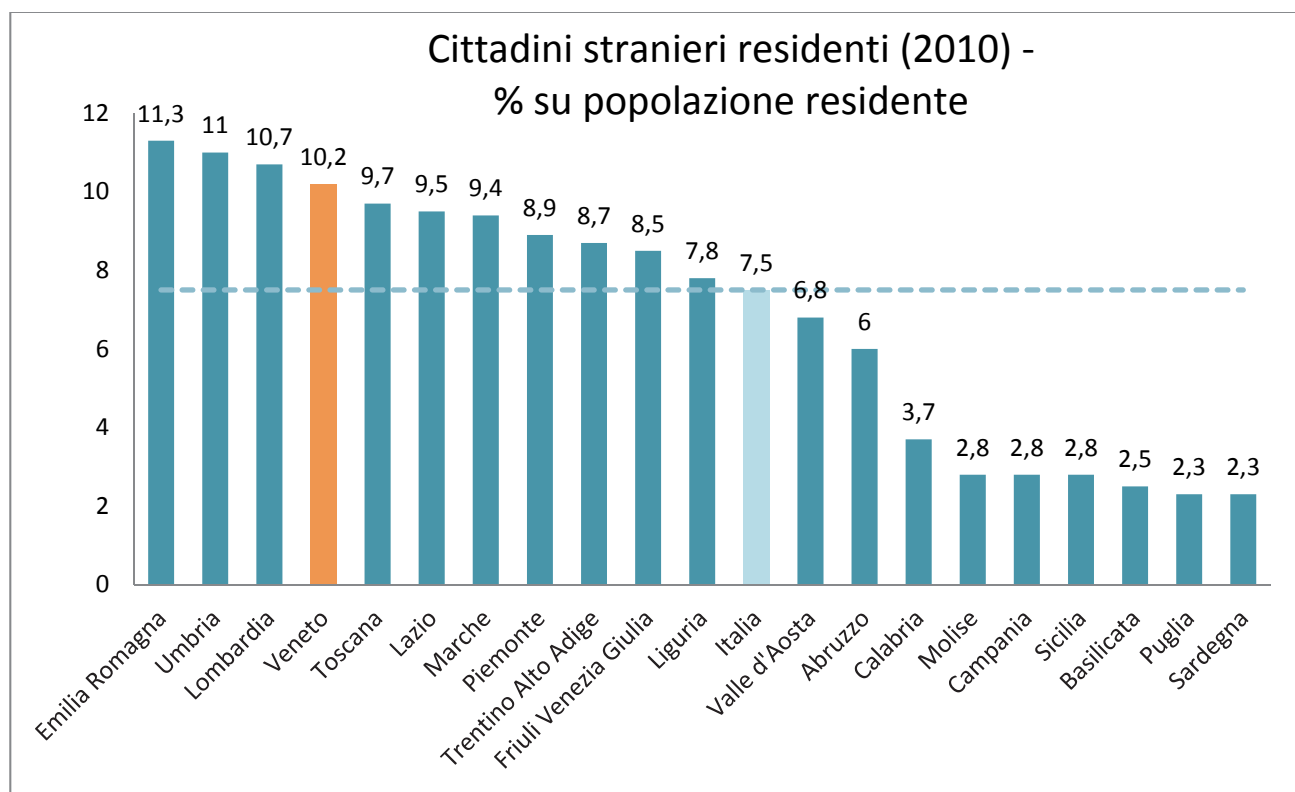


Grafico 41 - Cittadini stranieri residenti su popolazione residente, anno 2010. Elaborazione su dati Istat, Sisreg.

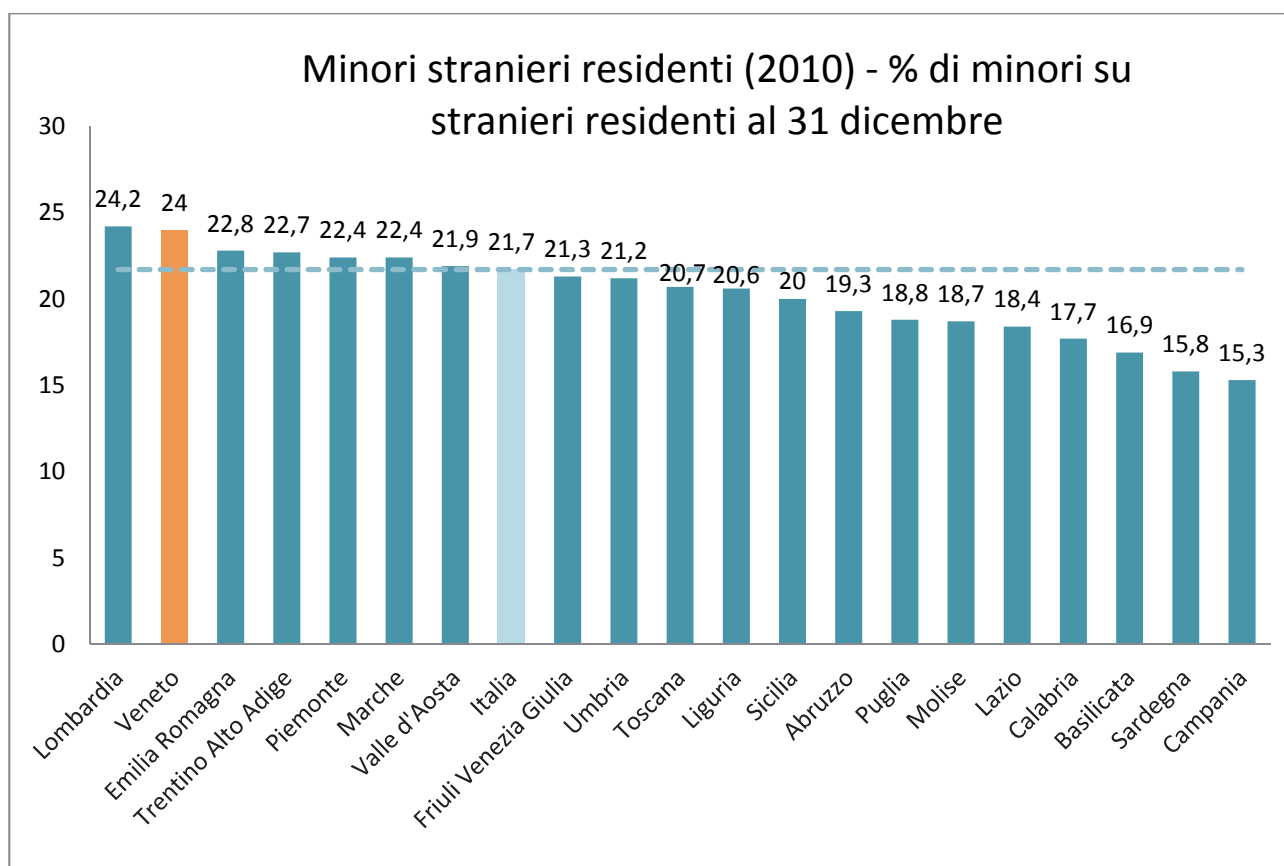


Grafico 42 - Minori stranieri residenti su totale stranieri residenti, anno 2010. Elaborazione su dati Istat, Sisreg.

7.2. Situazione imprenditoriale: numero di imprese (per 1.000 abitanti) e numero medio addetti per impresa

Le imprese rivestono un ruolo chiave nell'Unione Europea, soprattutto perché esse sono uno stimolo alla concorrenza e quindi all'innovazione. Un indicatore che ci fa percepire la situazione delle imprese all'interno dell'UE è il numero di imprese per 1.000 abitanti. I paesi che hanno un numero maggiore di imprese risultano avere (con qualche eccezione) anche il numero medio di addetti più basso, ad indicare una forte presenza di piccole imprese. Questi paesi (che non raggiungono in media i 4 addetti per impresa) sono la Grecia, la Repubblica Ceca, la Slovacchia, il Portogallo e l'Italia.

Il valore che si osserva per l'Italia, pari a 63,5 imprese ogni 1.000 abitanti, è un dato che testimonia da un lato un grande spirito d'iniziativa privata, dall'altro lato, il grado di frammentazione della produzione. Il dato italiano infatti, testimonia la prevalenza di tante imprese di piccolissime dimensioni, sottolineato anche dal numero medio di addetti per impresa, che è pari a 3,9.

Ben diversa appare la realtà tedesca, che registra una presenza di 25,4 imprese ogni 1.000 abitanti (meno della metà delle italiane) e un numero medio di addetti per impresa pari a 12. Essa è infatti la nazione con il maggior numero medio di addetti per impresa.

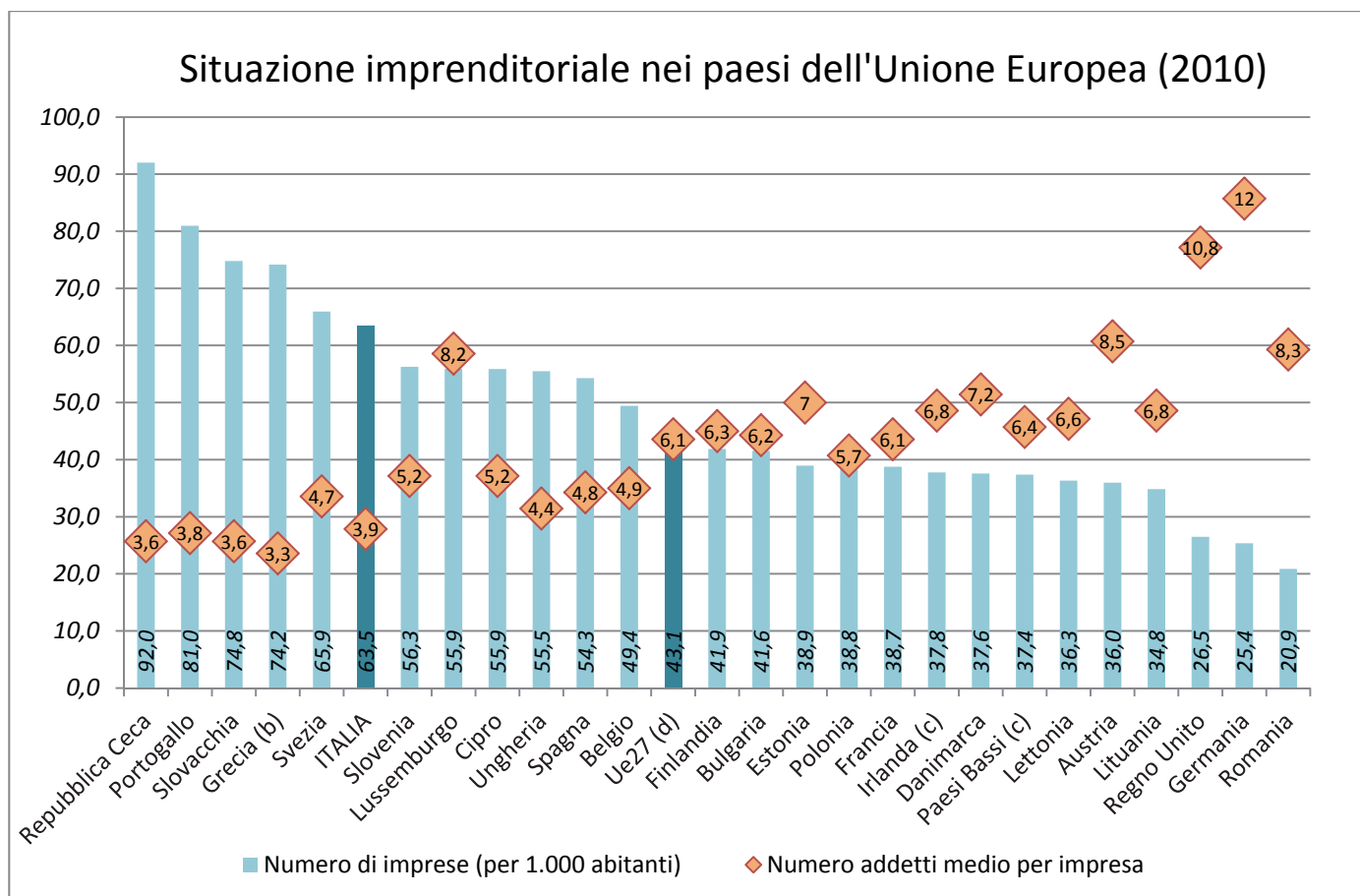


Grafico 43 - Numero di imprese (per 1.000 abitanti) e numero medio di addetti per impresa, anno 2010.

Elaborazione su dati Eurostat, Istat (Noi Italia).

Legenda: (b) ultimo dato disponibile 2008; (c) ultimo dato disponibile 2009; (d) valore stimato.

Il grafico successivo ci illustra la situazione per quanto riguarda le regioni d'Italia o le aree territoriali: osserviamo per il Veneto un numero pari a **71,6** imprese ogni 1.000 abitanti, un valore molto elevato; e mentre per il grafico riguardante l'UE il numero ridotto di imprese significava una presenza di imprese di grandi dimensioni, per quanto riguarda la realtà italiana ci sembra di poter affermare che il numero ridotto di imprese significa (solo per quanto riguarda l'Italia) un basso spirito di iniziativa privata, anche dovuto a ragioni storiche.

Per quanto riguarda il numero medio di addetti per impresa, osserviamo per il Veneto una media pari a **4,2**, un valore che ci porta comunque a concludere che abbiamo un territorio molto frammentato, e che anche per questo troveremo sempre più difficoltà a competere con gli altri paesi dell'Unione Europea.

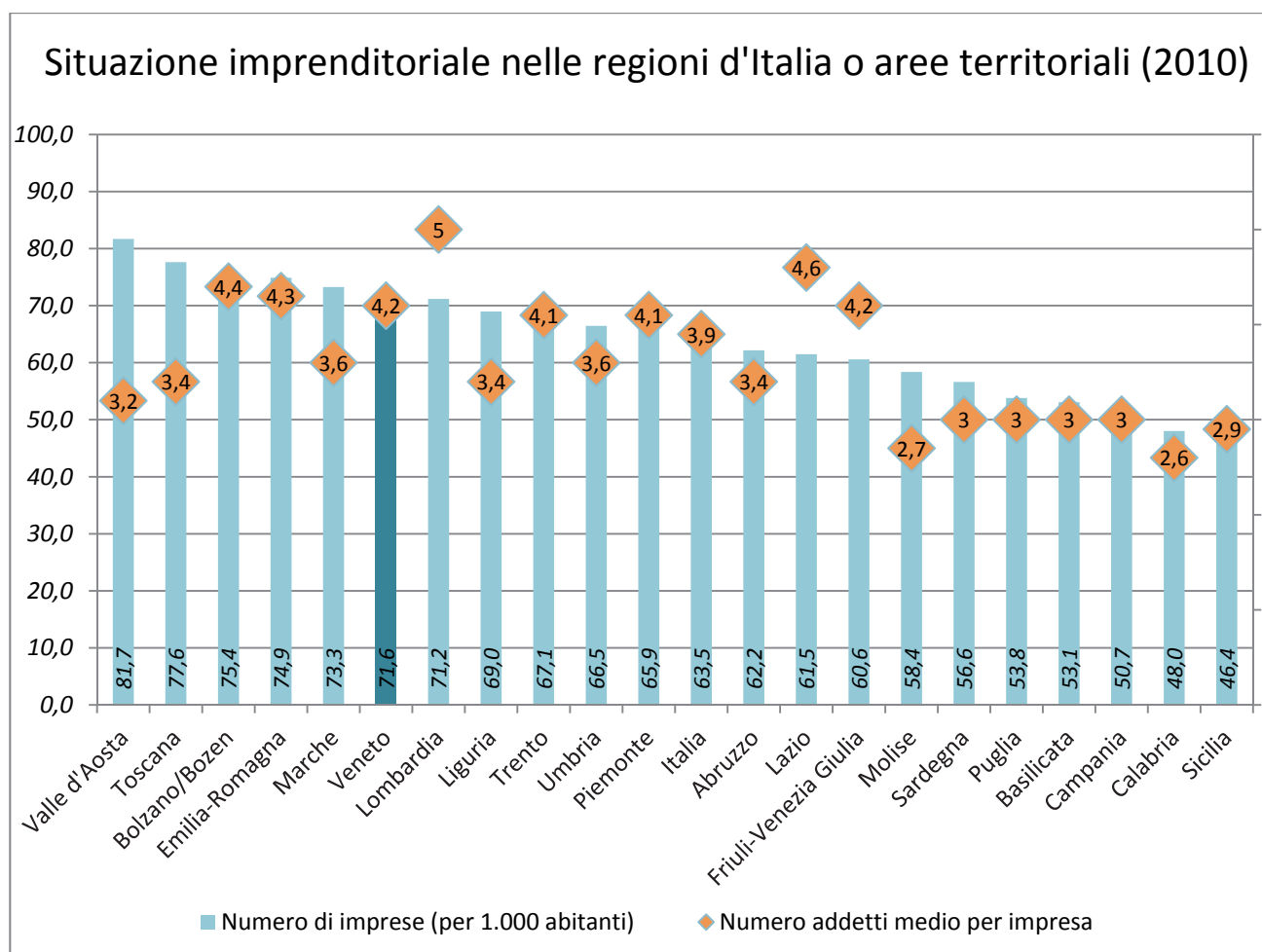


Grafico 44 - Numero di imprese (per 1.000 abitanti) e numero medio di addetti per impresa, anno 2010.
Elaborazione su dati Istat, Registro statistico delle imprese attive (Noi Italia).

7.3. Natimortalità delle imprese

Il tasso di natimortalità delle imprese è calcolato come le nuove imprese meno le imprese cessate sul totale delle imprese registrate nell'anno precedente (moltiplicato per 100). È un indicatore demografico d'impresa, che deve essere interpretato anche come “campanello d'allarme”.

Per quanto riguarda le regioni italiane, osserviamo per il Veneto un tasso di natimortalità netto del **-0,24%**, un tasso che ci informa sul fatto che purtroppo sono maggiori le imprese cessate rispetto a quelle nate. Per l'Italia osserviamo un tasso del -0,04%, mentre sono presenti alcune regioni che registrano un tasso turnover positivo, quali il Lazio, la Campania, la Liguria, l'Abruzzo, le Marche e l'Emilia Romagna.

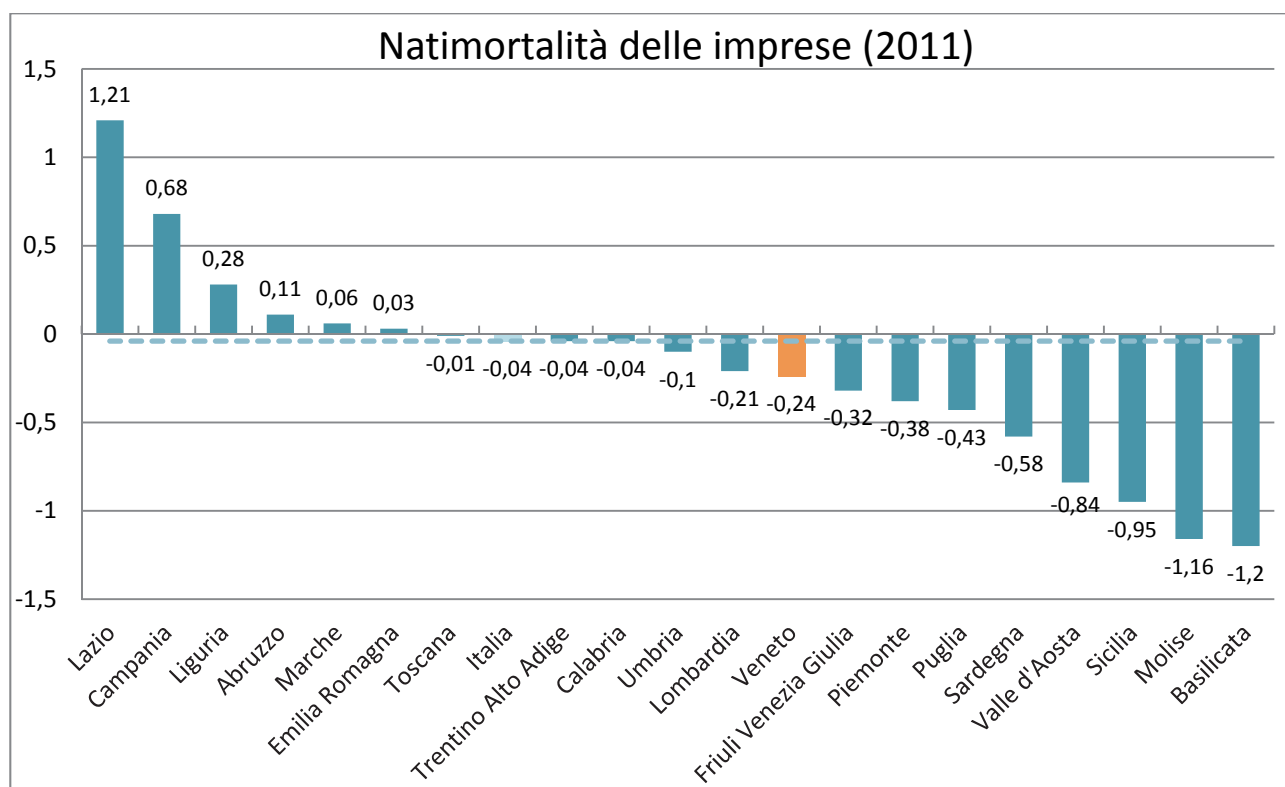


Grafico 45 - Natimortalità delle imprese, anno 2011. Elaborazioni Istat su dati Infocamere.

7.4. Esportazioni

Con riferimento alle esportazioni, osserviamo che l'Italia ha visto calare la sua quota di mercato, sia per quanto riguarda le esportazioni intra – UE sia per quanto riguarda quelle extra – UE. La Germania invece, ha visto incrementare entrambe le quote.

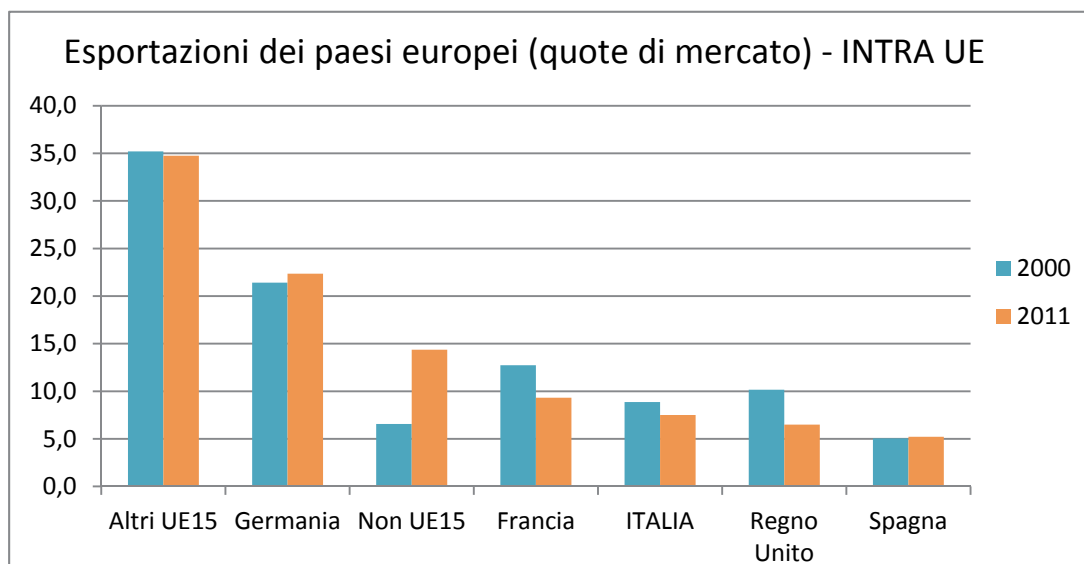


Grafico 46 - Esportazioni all'interno dell'Unione Europea dei paesi europei, anni 2000 e 2011.
Elaborazioni Istat (Noi Italia) su dati Eurostat.

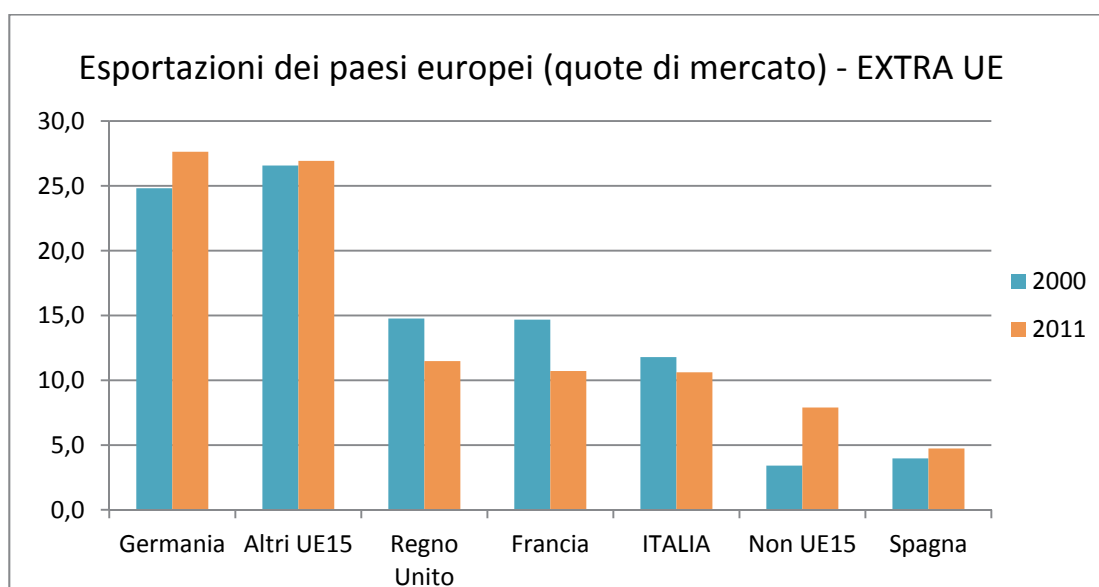


Grafico 47 - Esportazioni all'esterno dell'Unione Europea dei paesi europei, anno 2000 e 2011.
Elaborazioni Istat (Noi Italia) su dati Eurostat.

Per quanto riguarda le regioni d'Italia, osserviamo che la regione che detiene la quota maggiore dell'export italiano è la Lombardia, con una quota pari a 27,7%. Il Veneto la segue, con un valore però nettamente più basso, pari a **13,4%**. Altre regioni simili per quota detenuta sono le regioni Emilia – Romagna, Piemonte e Toscana, che detengono rispettivamente delle quote pari a 12,8%, 10,3% e 8,1%.

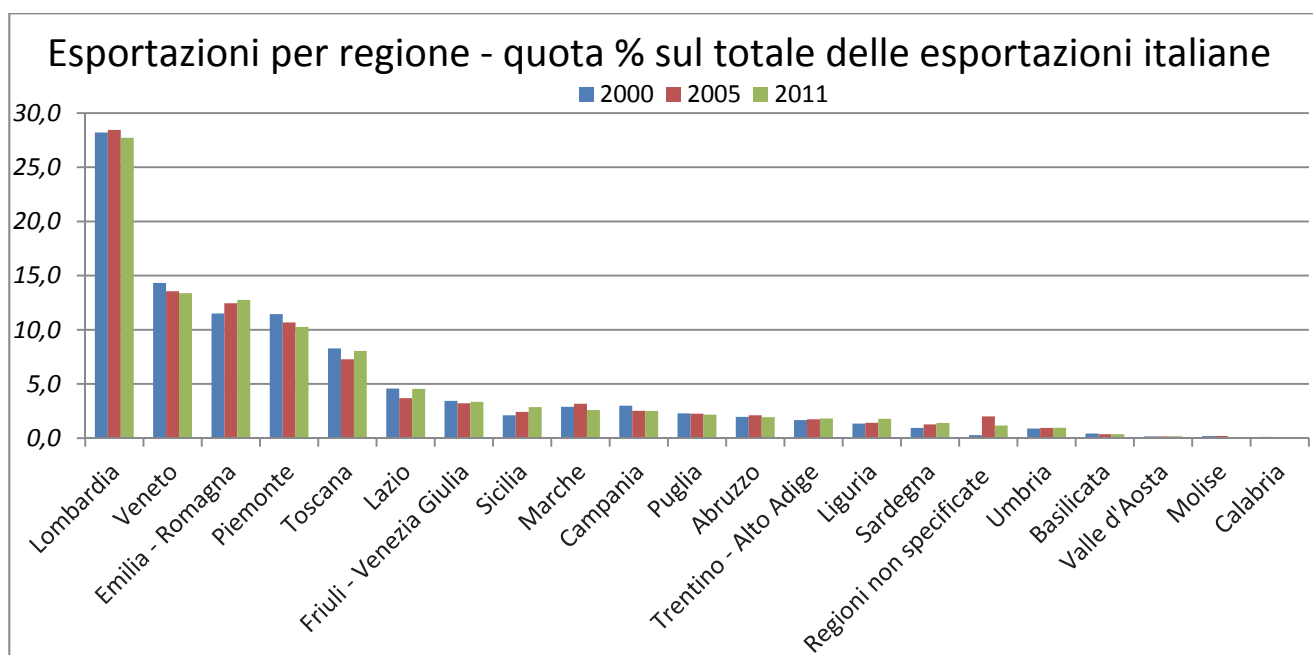


Grafico 45 - Quota di esportazioni per regione sul totale delle esportazioni italiane, anni 2000, 2005 e 2011.
Elaborazione su dati Istat, Statistiche del commercio con l'estero (Noi Italia).

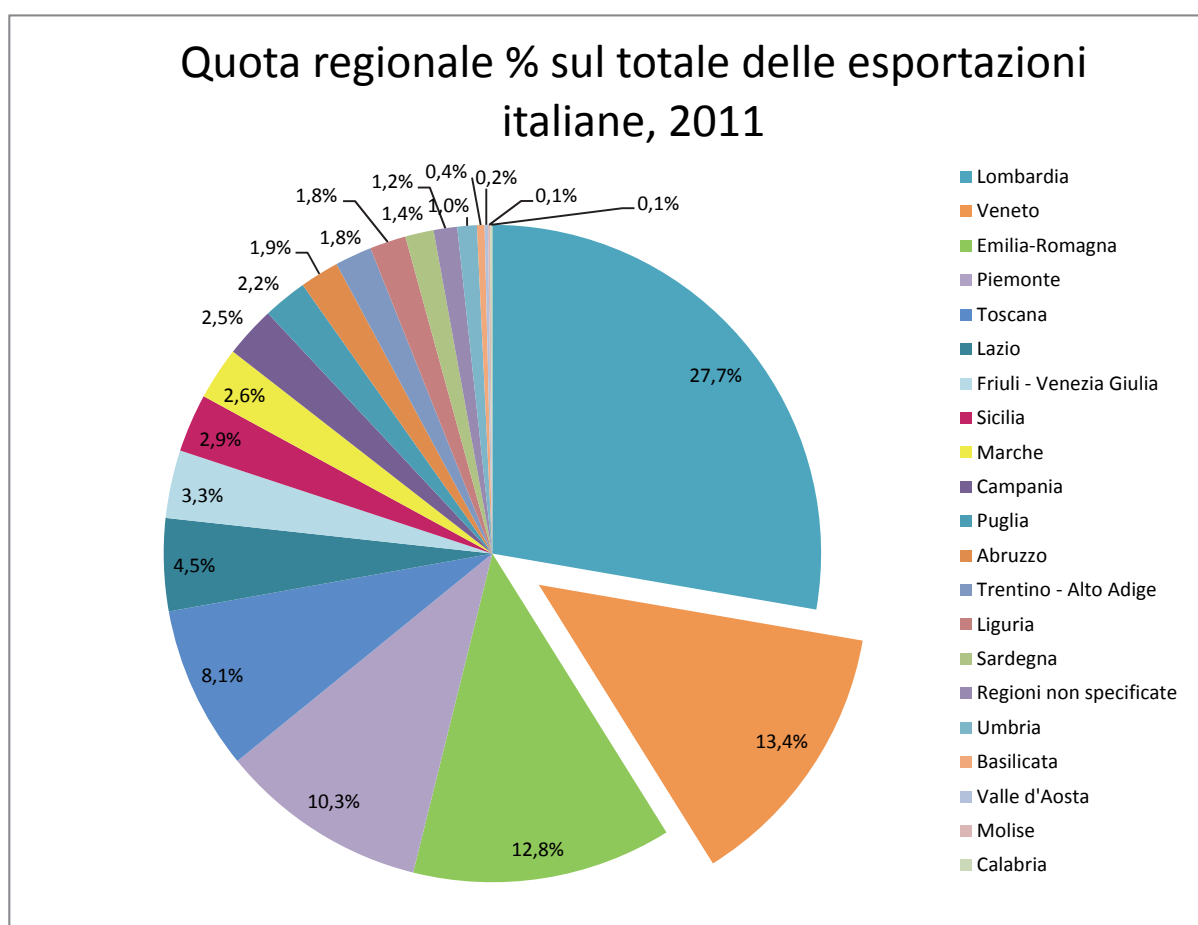


Grafico 46 - Quota di esportazioni per regione sul totale delle esportazioni italiane, anno 2011. Elaborazione su dati Istat, Statistiche del commercio con l'estero (Noi Italia).

8. ALLEGATO TECNICO

L'obiettivo generale di questo lavoro è l'individuazione di gruppi di regioni dell'Unione Europea che, anche se non necessariamente geograficamente confinanti, costituiscano al loro interno un sistema omogeneo in termini di alcune variabili di tipo economico, sociale e demografico ritenute rilevanti.

In questo documento illustreremo gli *step* seguiti per arrivare alla suddivisione di tutte le regioni appartenenti ai 27 stati dell'Unione Europea in 8 gruppi distinti, ciascuno dei quali – come si vedrà meglio in seguito – presenta caratteristiche diverse da quelle degli altri.

I passi che saranno illustrati sono i seguenti:

- ✚ raccolta dei dati ed elaborazione del *dataset* di partenza;
- ✚ una prima *cluster analysis* in base al PIL ed al suo tasso di crescita;
- ✚ analisi fattoriale, al fine di ridurre il numero delle variabili da prendere in considerazione;
- ✚ delle *cluster analysis* successive – ciascuna per ogni gruppo determinato nella c.a. precedente – in base ai fattori definiti in corrispondenza del secondo punto.

Sia le procedure di *cluster analysis* che le procedure di analisi fattoriale sono implementate con il software statistico SPSS 16.0.

DATASET:

I dati utilizzati sono ricavati dal sito dell'Eurostat (sezione "*Statistics*", "*General and regional statistics*", "*Regions and cities (including metropolitan regions)*", "*Data*", "*Database*").

I 27 paesi dell'Unione Europea sono stati suddivisi in 203 regioni, determinate in base ai criteri NUTS1 e NUTS2 (la Nomenclatura delle Unità Territoriali Statistiche) i quali identificano la ripartizione del territorio dell'Unione Europea a fini statistici, mentre alcuni stati sono stati considerati per intero. Nella tabella sottostante è riportato il criterio di suddivisione utilizzato per ciascun paese:

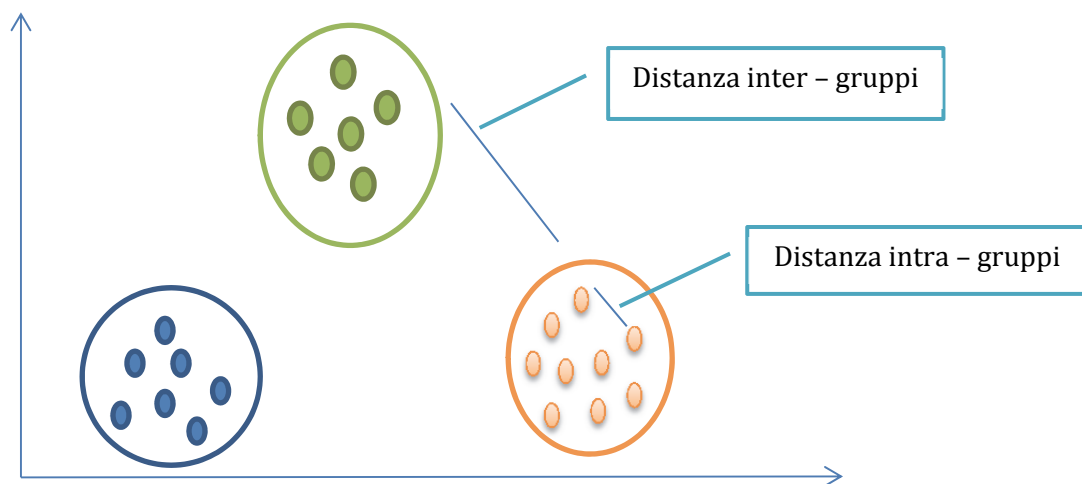
STATO	CLASSIFICAZIONE
Austria	NUTS2
Belgio	NUTS1
Bulgaria	NUTS2
Cipro	INTERO
Repubblica Ceca	NUTS2

STATO	CLASSIFICAZIONE
Danimarca	NUTS2
Estonia	INTERO
Finlandia	NUTS2
Francia	NUTS2
Germania	NUTS1
Grecia	NUTS2
Irlanda	NUTS2
Ungheria	NUTS2
Italia	NUTS2
Lettonia	INTERO
Lituania	INTERO
Lussemburgo	INTERO
Malta	INTERO
Paesi Bassi	NUTS1
Polonia	NUTS2
Portogallo	NUTS2
Romania	NUTS2
Slovenia	NUTS2
Slovacchia	NUTS2
Spagna	NUTS2
Svezia	NUTS2
Regno Unito	NUTS1

CLUSTER ANALYSIS:

- Definizione ed obiettivi della *cluster analysis*:

La *cluster analysis*, letteralmente analisi dei gruppi, comprende un insieme di tecniche statistiche il cui scopo è l'individuazione di raggruppamenti di unità aventi le seguenti caratteristiche: massima somiglianza tra gli elementi degli stessi gruppi (viene minimizzata la distanza intra – gruppi) e massima differenza tra gli elementi di gruppi diversi (viene massimizzata la distanza inter – gruppi).



- Le fasi della *cluster analysis*:

Le fasi di un processo di *cluster analysis* possono essere sintetizzate come riportato di seguito:

1. scelta delle unità di osservazione e delle variabili;
2. scelta della misura di distanza (o di similarità) tra i dati e decisione di standardizzare o meno le variabili;
3. scelta dell'algoritmo di classificazione;
4. scelta del numero dei gruppi;
5. implementazione dell'analisi in seguito alle scelte precedenti.

Per realizzare la classificazione è necessario individuare dei criteri di similarità e dissimilarità tra coppie; ciò si ottiene definendo degli indici che forniscono delle indicazioni importanti al fine di individuare le unità omogenee che costituiranno i *clusters*.

La distanza che verrà utilizzata in questo lavoro è la **distanza euclidea**, definita nel modo seguente:

$$D_{ij} = \sqrt{\sum_{k=1}^n (x_{ki} - x_{kj})^2}$$

dove D_{ij} = distanza tra le unità i e j ;

x_{ki} = valore della variabile x_k per il caso i .

Dal momento che ad unità di misura differenti corrispondono pesi differenti nel calcolo della distanza (unità di misura più elevate in valore assoluto incidono maggiormente sul valore finale di D_{ij}), è opportuno standardizzare² le variabili di partenza.

Per quanto riguarda invece i possibili algoritmi di classificazione, essi si distinguono in “gerarchici” e “non gerarchici”. I primi producono raggruppamenti successivi ordinabili secondo livelli crescenti o decrescenti della distanza (o, viceversa, della similarità); si tratta di procedure iterative che considerano tutti i livelli di distanza ed i gruppi che si ottengono ad un certo livello sono contenuti nei gruppi ottenuti ad un livello di distanza inferiore. Gli algoritmi di questo tipo non richiedono alcuna decisione riguardo al numero di *clusters*.

Le procedure “non gerarchiche”, invece, richiedono che il numero di gruppi sia specificato a priori e forniscono come output un’unica partizione, considerata ottimale rispetto al criterio adottato, degli elementi iniziali nel numero di *cluster* fornito in input.

L’approccio che verrà seguito in questo lavoro consiste in una combinazione di un metodo “gerarchico” e di un metodo “non gerarchico”:

1. in primo luogo verrà usato un metodo “gerarchico” per definire il numero di *cluster* da ottenere;
2. in secondo luogo verrà usato un metodo “non gerarchico” per definire le regioni componenti di ciascun *cluster*.

Nello specifico, il metodo “gerarchico” scelto nel primo *step* è il **metodo di Ward**: è un algoritmo iterativo che riunisce, ad ogni tappa del processo, i due gruppi dalla cui fusione deriva il minimo incremento possibile della devianza intra essi. Il tipo di distanza di *default* per questo metodo è il quadrato della distanza euclidea.

Per quanto riguarda invece il metodo “non gerarchico” utilizzato al secondo *step*, esso è il **metodo delle k - medie** : il suo obiettivo è la minimizzazione della varianza totale intra - cluster. Ogni *cluster* viene identificato mediante un centroide o punto medio; l’algoritmo segue una procedura iterativa. Inizialmente crea K partizioni ed assegna ad ogni partizione i punti d’ingresso o casualmente o usando alcune informazioni euristiche; quindi calcola il centroide di ogni gruppo e costruisce poi una nuova partizione associando ogni punto d’ingresso al *cluster* il cui centroide è più vicino ad esso. Quindi vengono ricalcolati i centroidi per i nuovi *clusters* e così via, finché l’algoritmo non converge. Tale metodo ha soprattutto il pregio di

² Standardizzazione z della variabile x: $z = \frac{x - media(x)}{deviazione\ standard(x)}$

essere particolarmente veloce: infatti si è osservato che generalmente il numero di iterazioni è minore del numero di punti.

- Cluster analysis in base al Pil:

Il primo passo dell'analisi consiste nel suddividere le 203 regioni dell'Unione Europea in un certo numero di *clusters* determinati in base al PIL pro capite medio ed al suo tasso di crescita medio. I dati considerati riguardano i tre anni più recenti di informazioni riguardo al Pil, ossia l'arco temporale 2007 – 2009.

Dal momento che le due variabili usate per determinare i gruppi sono espresse in unità di misura piuttosto differenti (il Pil pro capite è un numero “puro”, mentre il tasso di crescita è misurato in punti percentuali), si procede ad una standardizzazione del PIL pro capite; in questo modo le due variabili non avranno pesi diversi nell'imminente *cluster analysis* (l'intenzione era comunque di attribuire un peso leggermente maggiore al PIL pro capite).

A questo punto si inizia con una *cluster analysis* gerarchica implementata con il metodo di Ward per determinare il numero di *clusters* più opportuno da definire.

A tale scopo si utilizza la scheda agglomerativa (riportata come output da SPSS): la colonna di maggior interesse è quella relativa ai “coefficienti”, nella quale compare il valore della distanza utilizzata per formare i *clusters*. Da tali valori si capisce quanto siano diversi tra di loro i *clusters* che vengono combinati ad ogni passo. Dal momento che in questo lavoro si utilizza una misura di distanza (e non di similarità), coefficienti bassi indicano che vengono assemblati gruppi abbastanza omogenei, mentre coefficienti elevati indicano unioni di gruppi dissimili tra loro. Ci si ferma nella formazione di *clusters* quando l'aumento tra due coefficienti successivi è elevato³.

³ In questo procedimento, come è facilmente intuibile, subentra comunque sempre una componente soggettiva.

Stage	Cluster Combined		Coefficients	Stage Cluster First Appears		Next Stage
	Cluster 1	Cluster 2		Cluster 1	Cluster 2	
...
183	6	124	,841	175	157	189
184	3	19	,941	178	161	193
185	118	144	1,049	168	150	196
186	22	25	1,204	176	163	195
187	10	27	1,360	173	136	191
188	2	20	1,620	177	166	195
189	4	6	1,927	140	183	196
190	8	13	2,300	174	182	198
191	10	28	2,766	187	126	197
192	18	23	3,461	181	180	197
193	3	40	4,317	184	179	198
194	1	120	5,400	0	0	199
195	2	22	6,678	188	186	200
196	4	118	7,969	189	185	201
197	10	18	13,830	191	192	199
198	3	8	19,925	193	190	200
199	1	10	34,245	194	197	202
200	2	3	54,943	195	198	201
201	2	4	102,667	200	196	202
202	1	2	202,138	199	201	0

Il numero di *clusters* scelto è determinato dalla differenza tra il numero totale di osservazioni (in questo caso 203) ed il punto in corrispondenza del quale si verifica l'aumento significativo del coefficiente: perciò si scelgono $203 - 200 = 3$ *clusters*⁴.

Allo scopo di definire quali regioni appartengono a ciascuno di questi 3 insiemi si procede ora con una partizione con il metodo delle k – medie: il risultato porta alla seguente suddivisione:

⁴ È stata riportata solamente l'ultima parte della tabella, perché è la parte più interessante ai nostri fini.

Cluster 1 (n=14)	Cluster 2 (n=96)	Cluster 3 (n=93)
Région de Bruxelles – Capitale (BE)	Vlaams Gewest (BE)	Région wallonne (BE)
Praha (CZ)	Sjælland (DK)	Severozapaden (BG)
Hovedstaden (DK)	Syddanmark (DK)	Severen tsentralen (BG)
Bremen (DE)	Midtjylland (DK)	Severoiztochen (BG)
Hamburg (DE)	Nordjylland (DK)	Yugoiztochen (BG)
Southern and Eastern (IE)	Baden-Württemberg (DE)	Yugozapaden (BG)
Île de France (FR)	Bayern (DE)	Yuzhen tsentralen (BG)
Provincia Autonoma di Bolzano (IT)	Berlin (DE)	Strední Cechy (CZ)
Luxembourg (LU)	Hessen (DE)	Jihozápad (CZ)
Wien (AT)	Niedersachsen (DE)	Severozápad (CZ)
Bratislavský kraj (SK)	Nordrhein-Westfalen (DE)	Severovýchod (CZ)
Åland (FI)	Rheinland-Pfalz (DE)	Jihovýchod (CZ)
Stockholm (SE)	Saarland (DE)	Strední Morava (CZ)
London (UK)	Sachsen (DE)	Moravskoslezsko (CZ)
	Schleswig-Holstein (DE)	Brandenburg (DE)
	Border, Midland and Western (IE)	Mecklenburg-Vorpommern (DE)
	Attiki (EL)	Sachsen-Anhalt (DE)
	Notio Aigaio (EL)	Thüringen (DE)
	Galicía (ES)	Estonia (EE)
	Principado de Asturias (ES)	Anatoliki Makedonia, Thraki (EL)
	Cantabria (ES)	Kentriki Makedonia (EL)
	País Vasco (ES)	Dytiki Makedonia (EL)
	Comunidad Foral de Navarra (ES)	Thessalia (EL)
	La Rioja (ES)	Ipeiros (EL)
	Aragón (ES)	Ionia Nisia (EL)
	Comunidad de Madrid (ES)	Dytiki Ellada (EL)
	Castilla y León (ES)	Stereia Ellada (EL)
	Cataluña (ES)	Peloponnisos (EL)
	Comunidad Valenciana (ES)	Voreio Aigaio (EL)
	Illes Balears (ES)	Kriti (EL)
	Región de Murcia	Castilla-la Mancha (ES)
	Ciudad Autónoma de Ceuta (ES)	Extremadura (ES)
	Ciudad Autónoma de Melilla (ES)	Andalucía (ES)

Canarias (ES)	Picardie (FR)
Champagne-Ardenne (FR)	Basse-Normandie (FR)
Haute-Normandie (FR)	Lorraine (FR)
Centre (FR)	Franche-Comté (FR)
Bourgogne (FR)	Poitou-Charentes (FR)
Nord - Pas-de-Calais (FR)	Limousin (FR)
Alsace (FR)	Languedoc-Roussillon (FR)
Pays de la Loire (FR)	Corse (FR)
Bretagne (FR)	Abruzzo (IT)
Aquitaine (FR)	Molise (IT)
Midi-Pyrénées (FR)	Campania (IT)
Rhône-Alpes (FR)	Puglia (IT)
Auvergne (FR)	Basilicata (IT)
Provence-Alpes-Côte d'Azur (FR)	Calabria (IT)
Piemonte (IT)	Sicilia (IT)
Valle d'Aosta/Vallée d'Aoste (IT)	Sardegna (IT)
Liguria (IT)	Latvia (LV)
Lombardia (IT)	Lithuania (LT)
Provincia Autonoma di Trento (IT)	Közép-Dunántúl (HU)
<i>Veneto (IT)</i>	Nyugat-Dunántúl (HU)
Friuli-Venezia Giulia (IT)	Dél-Dunántúl (HU)
Emilia-Romagna (IT)	Észak-Magyarország (HU)
Toscana (IT)	Észak-Alföld (HU)
Umbria (IT)	Dél-Alföld (HU)
Marche (IT)	Malta (MT)
Lazio (IT)	Burgenland (AT)
Cyprus (CY)	Lódzkie (PL)
Közép-Magyarország (HU)	Malopolskie (PL)
Noord-Nederland (NL)	Slaskie (PL)
Oost-Nederland (NL)	Lubelskie (PL)
West-Nederland (NL)	Podkarpackie (PL)
Zuid-Nederland (NL)	Swietokrzyskie (PL)
Niederösterreich (AT)	Podlaskie (PL)
Kärnten (AT)	Wielkopolskie (PL)

Steiermark (AT)	Zachodniopomorskie (PL)
Oberösterreich (AT)	Lubuskie (PL)
Salzburg (AT)	Dolnoslaskie (PL)
Tirol (AT)	Opolskie (PL)
Vorarlberg (AT)	Kujawsko-Pomorskie (PL)
Mazowieckie (PL)	Warminsko-Mazurskie (PL)
Lisboa (PT)	Pomorskie (PL)
Região Autónoma da Madeira (PT)	Norte (PT)
Bucuresti – Ilfov (RO)	Algarve (PT)
Zahodna Slovenija (SI)	Centro (PT)
Itä-Suomi (FI)	Alentejo (PT)
Etelä-Suomi (FI)	Região Autónoma dos Açores (PT)
Länsi-Suomi (FI)	Nord-Vest (RO)
Pohjois-Suomi (FI)	Centru (RO)
Östra Mellansverige (SE)	Nord-Est (RO)
Småland med öarna (SE)	Sud-Est (RO)
Sydsverige (SE)	Sud – Muntenia (RO)
Västsverige (SE)	Sud-Vest Oltenia (RO)
Norra Mellansverige (SE)	Vest (RO)
Mellersta Norrland (SE)	Vzhodna Slovenija (SI)
Övre Norrland (SE)	Západné Slovensko (SK)
North West (UK)	Stredné Slovensko (SK)
Yorkshire and The Humber (UK)	Východné Slovensko (SK)
East Midlands (UK)	North East (UK)
West Midlands (UK)	Wales (UK)
East of England (UK)	Northern Ireland (UK)
South East (UK)	
South West (UK)	
Scotland (UK)	

Di seguito riportiamo la tabella ANOVA (che illustra il contributo di ciascuna variabile nella formazione dei 3 insiemi ottenuti), i centri finali dei *clusters* e le medie del Pil pro capite nei 3 diversi gruppi:

ANOVA						
	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
media var_gdp	0,008	2	0,001	200	12,586	0
Zscore (media_pro_capite)	76,162	2	0,248	200	306,629	0

Final Cluster Centers			
	Cluster		
	1	2	3
media var_gdp	-0,01172	-0,02822	-0,01063
Zscore(media_pro_capite)	2,31957	0,43321	-0,79637

Cluster	Media pro capite
1	43240,48
2	26578,82
3	15718,28

Dalle tabelle soprastanti si nota innanzitutto che entrambe le variabili hanno contribuito alla definizione dei gruppi: in entrambi i casi infatti abbiamo livelli di significatività molto bassi, circa pari a 0. Il primo cluster è caratterizzato da un reddito mediamente più elevato e (dal fatto di rappresentare regioni sede di capitali o comunque di grandi città che attirano lavoratori residenti anche nelle regioni circostanti); il secondo da un livello di reddito “medio” ed il terzo da un reddito medio più basso. Riguardo al tasso di crescita medio del PIL, è interessante sottolineare che è negativo per tutti i 3 gruppi (primi segni della crisi); tale “recessione” è stata più marcata per il secondo gruppo e leggermente meno marcata per il terzo gruppo.

Il profilo generale che si può tracciare da questa prima suddivisione è il seguente:

1. al cluster 1 appartengono le regioni “d’elite”, cioè le regioni mediamente più ricche all’interno dell’UE, caratterizzate negli anni considerati da un decremento del Pil “medio”;

2. al cluster 2 appartengono le regioni “medie”, cioè mediamente ricche, caratterizzate però da un decremento del Pil più elevato;
3. al cluster 3 appartengono invece le regioni più povere, caratterizzate però da un decremento minore del Pil, quindi da una “stabilità” leggermente maggiore.

Il maggior divario nel reddito fra le regioni “d’elite” e le quelle più “povere” è illustrato anche dalla matrice delle distanze finali tra i centri dei clusters.

Il Veneto appartiene al cluster 2.

Abbiamo effettuato questa prima divisione delle regioni in base al Pil perché è l’indicatore sintetico più efficace per esprimere la ricchezza di un certo territorio. A questo punto, per arrivare a degli insiemi di regioni meno numerosi ed il più omogenei possibile, la nostra indagine continua attraverso la determinazione delle differenze all’interno degli stessi *clusters*; in altre parole, si vuole vedere come si differenziano tra di loro regioni con reddito abbastanza simile, in base ad altre variabili, che scegliamo fra gli ambiti sociali e demografici. Dal momento che le variabili prese in considerazione sono abbastanza numerose, si desidera diminuirne il numero, raggruppandone possibilmente alcune correlate tra di loro: a tale scopo si decide di utilizzare l’analisi fattoriale, che verrà descritta accuratamente nei paragrafi seguenti.

L’ANALISI FATTORIALE: definizione e obiettivi

L’analisi fattoriale consiste in una classe di tecniche statistiche utilizzate per ridurre e sintetizzare i dati relativi a più variabili in un insieme più piccolo di *fattori*: questi ultimi sono variabili *latenti*, vale a dire non direttamente osservabili, e sono indotti dalle variabili osservate.

In altre parole questo tipo di analisi si pone l’obiettivo di semplificare un set di variabili interdipendenti in pochi fattori sottostanti, mantenendo il più possibile lo stesso livello di informazione iniziale.

Da un punto di vista matematico ciascuna variabile osservata può essere espressa come una combinazione lineare dei fattori:

$$x_j = a_{j1} f_1 + a_{j2} f_2 + \dots + a_{jq} f_q + u_j s_j$$

dove: x_j = j – esima variabile osservata;

f_i ($i = 1, \dots, q$) = fattore comune i – esimo (è chiamato così perché è presente in tutte le possibili equazioni aventi come variabile dipendente ciascuna variabile osservata);
 a_{ji} = coefficienti fattoriali (*factor loadings* o *saturazioni*) della j – esima variabile sul fattore comune i – esimo;
 s_j = fattore specifico (o unico) di x_j (è chiamato così perché è proprio della variabile x_j) ;
 u_j = coefficiente di regressione della j – esima variabile sul j – esimo fattore.

Ogni fattore comune è una combinazione lineare di tutte le variabili osservate:

$$f_i = \sum_j w_{ji} x_j \quad (j=1, \dots, p; i=1, \dots, q)$$

dove: w_{ji} = peso (o *coefficiente fattoriale* o *factor score coefficient*) corrispondente alla variabile x_j nella combinazione f_i .

La determinazione dei pesi fattoriali consente di valutare l'influenza di ciascun fattore espressa in termini di apporto relativo alla variabilità complessiva del sistema, così da individuare quali possano ritenersi statisticamente rilevanti.

In questo modello si assume correlazione nulla tra i fattori comuni, tra fattori comuni e fattori unici e tra i fattori unici.

Grazie a tali assunzioni è possibile esprimere la varianza totale del modello come la somma di due diverse componenti: la varianza comune (*comunalità*), che è la parte di varianza spiegata dai fattori comuni, e la varianza unica (*specificità*), che è la parte di varianza specifica per ciascuna variabile e non spiegata dai fattori.

In formule:

$$\sigma_j^2 = 1 = \text{Var}(\sum_i a_{ji} f_i + u_j c_j) = \sum_i a_{ji}^2 \lambda_i + u_j^2 \text{Var}(s_j)$$

$$h_j^2 = \sum_i a_{ji}^2$$

dove h_j^2 = comunalità.

LE FASI DELL'ANALISI FATTORIALE:

Per implementare un procedimento di analisi fattoriale i passi da seguire sono i seguenti:

- ✓ Identificazione degli obiettivi specifici, scelta delle variabili osservate e loro eventuale trasformazione:

È una parte molto soggettiva, in quanto dipende dalla teoria e dal giudizio della persona che effettua l'analisi. Se le variabili non sono espresse nella stessa unità di misura, è opportuno standardizzarle, in modo che le variabilità presentino un medesimo ordine di grandezza.

- ✓ Costruzione della matrice di correlazione:

Nel definire l'analisi fattoriale si è indicato come punto di partenza un "set di variabili interdipendenti". Questa interdipendenza è misurata dalla matrice di correlazione tra le variabili originarie. Infatti affinché abbia senso applicare l'analisi fattoriale è necessario che le variabili osservate siano significativamente correlate.

In particolare, è interessante il caso in cui per un sottoinsieme di variabili ci sia alta correlazione tra di esse e bassa correlazione con le altre: in questo modo ci si aspetta che le variabili altamente correlate misurino un certo fattore.

Per verificare se è opportuno procedere o meno con l'analisi fattoriale, oltre alla matrice di correlazione possono essere presi in considerazione altre statistiche o dei test: in questo elaborato si sceglie di effettuare il **test di sfericità di Bartlett**: in esso l'ipotesi nulla corrisponde all'incorrelazione tra tutte le variabili nel campione di riferimento. Quindi perché sia possibile usare l'analisi fattoriale è necessario che il test conduca al rifiuto dell'ipotesi nulla.

- ✓ Scelta del metodo di stima della matrice di varianza e covarianza:

Dopo aver verificato la fattibilità dell'analisi fattoriale, bisogna scegliere un metodo idoneo.

La scelta in questo lavoro ricade sul **metodo delle componenti principali**. Esso estrae le componenti (= i fattori, espressi come combinazioni lineari delle variabili di partenza) in ordine di varianza spiegata: la prima componente spiega la varianza massima e le successive spiegano parti di varianza sempre minori e sono tutte incorrelate tra loro.

- ✓ Determinazione del numero dei fattori:

Nell'introduzione all'analisi fattoriale è stato affermato che l'informazione contenuta nelle variabili osservate originarie deve essere sintetizzata in un certo numero q di fattori, numero

minore del numero di variabili osservate. Per determinare p esistono diversi metodi possibili: in questa sede viene utilizzato il **criterio degli autovalori**: si mantengono cioè solamente i fattori che presentano autovalori maggiori di 1. Un singolo autovalore illustra l'ammontare della varianza dell' i – esimo fattore e, dal momento che l'analisi è condotta sulla matrice di correlazione, la somma degli autovalori è pari al numero delle variabili osservate.

✓ Rotazione dei fattori:

In seguito alla stima della matrice di varianza e covarianza vengono prodotti i *factor loadings*, che rappresentano la correlazione tra i fattori e le variabili. Se un coefficiente ha un valore assoluto elevato significa che il fattore e la variabile corrispondenti sono fortemente correlati. Nella pratica però raramente la matrice contenente i *factor loadings* permette un'interpretazione chiara dei risultati, perché spesso un fattore è correlato con intensità simile con molte variabili, invece che molto intensamente con poche.

Per ovviare a questo problema si effettua una trasformazione in una matrice di più facile interpretazione tramite la cosiddetta **rotazione dei fattori**: tale rotazione modifica la struttura della matrice dei coefficienti fattoriali, avvicinandosi alla situazione ideale nella quale tutti i pesi fattoriali sono prossimi a 0 (correlazione molto bassa) o a 1 (correlazione molto elevata).

Le rotazioni possono essere di due tipi: ortogonali, che mantengono l'incorrelazione tra i fattori dopo la rotazione; oppure oblique, che conducono invece a fattori correlati.

In questo elaborato si parte da una rotazione obliqua di tipo "**Oblimin diretto**", con $\Delta=0$ (soluzione massimamente obliqua). Se la matrice di correlazione tra i fattori risultanti presenterà valori sufficientemente bassi (in linea di massima $<|0.30|$), si procederà con una rotazione ortogonale, precisamente di tipo "**Varimax**": quest'ultima minimizza il numero di variabili con alte saturazioni⁵ su ciascun fattore e facilita particolarmente l'interpretazione dei fattori.

✓ Interpretazione dei fattori:

Una volta eseguita la rotazione ed identificate le variabili maggiormente correlate con ciascuno dei fattori trovati, è opportuno assegnare dei nomi o delle "etichette" ai fattori stessi.

⁵ Non esistono regole universali per fissare il limite oltre il quale una saturazione è significativa.

✓ Stima dei punteggi fattoriali:

I punteggi fattoriali (*factor scores*) per l'*i* – esimo fattore sono stimati nel modo seguente:

$$f_i = \sum_j w_{ji} x_j \quad (j=1,...,p; i=1,...,q)$$

dove w_{ji} = pesi ottenuti dalla matrice dei coefficienti.

In questo lavoro si sceglie di calcolare i punteggi fattoriali col **metodo della regressione**: i punteggi ottenuti con tale metodo hanno media nulla e varianza pari al quadrato della correlazione multipla tra punteggi stimati e veri valori dei fattori.

L'ANALISI FATTORIALE NEI 3 CLUSTERS:

Si è detto che, una volta determinati i tre *clusters* in base al Pil ed al suo tasso di crescita, si desidera capire se e quali siano le differenze tra le regioni che rientrano nella medesima fascia di reddito; in altre parole si cerca di effettuare un'ulteriore suddivisione, tramite una seconda *cluster analysis*, in base ad altre variabili, che rientrano nell'ambito socio – demografico.

Le variabili di partenza sono nello specifico le seguenti:

✚ la popolazione suddivisa in tre fasce d'età:

- tra gli 0 ed i 14 anni;
- tra i 15 ed i 64 anni;
- dai 65 anni in su.

✚ il tasso di occupazione 15 – 24 (giovanile) e 25 – 64, suddiviso per sesso;

✚ il tasso di disoccupazione 15 – 24 (giovanile) e dai 25 anni in su, suddiviso per sesso;

✚ la percentuale degli occupati rientranti nella fascia d'età 25 – 64 anni e distinti per livello massimo di istruzione ottenuta sul totale della popolazione attiva regionale.

In particolare il livello di istruzione è suddiviso in tre gruppi:

- pre – primaria, primaria o secondaria inferiore (livelli 0 – 2 Isced);
- secondaria superiore e post – secondaria non terziaria (livelli 3 – 4 Isced);
- terziaria (livelli 5 – 6 Isced).

Allo scopo di un'identificazione più immediata, tali variabili sono state etichettate con i nomi rispettivamente di “capitale umano di base”, capitale umano medio” e “capitale umano avanzato”. Anche i valori di queste variabili sono stati suddivisi in base al sesso;

🇮🇹 le risorse umane in scienza e tecnologia espresse come percentuale della popolazione attiva totale regionale.

Al fine di avere un'informazione trasversale di più anni, anche queste variabili (come quelle riguardanti il Pil) sono state espresse come medie dei corrispondenti valori nell'arco di tempo compreso tra il 2007 ed il 2010.

Per avere una *cluster analysis* più "agevole", e dal momento che alcune di queste variabili possono essere supposte correlate tra di loro (ad esempio le tre variabili sulla popolazione nelle diverse fasce d'età, oppure i tassi di occupazione e di disoccupazione), si decide di effettuare un'analisi fattoriale; essa terminerà con il calcolo dei punteggi fattoriali, che saranno le variabili discriminanti nella seconda fase di *cluster analysis*.

Per avere variabili espresse in termini di grandezze simili, i valori delle tre variabili sulla componente demografica vengono standardizzate.

Seguendo le fasi descritte in precedenza, si parte dall'analisi della matrice di correlazione:

								Correlation Matrix												
		cap_um_p	cap_um_g	cap_um_s	cap_um_t	cap_um_t	emp_1524	emp_1524	unemp_1	unemp_1	unemp_2	unemp_2	hrst_activ	emp_2564	emp_2564	Zscore: m	Zscore: m	Zscore: m		
Correlation	cap_um_p	1	0,938	-0,852	-0,78	-0,301	-0,147	-0,174	-0,267	0,313	0,477	0,073	0,359	-0,406	0,072	-0,41	-0,106	-0,115	-0,076	
	cap_um_g	0,938	1	-0,779	-0,726	-0,286	-0,241	-0,019	-0,121	0,171	0,31	-0,035	0,197	-0,436	0,146	-0,256	-0,068	-0,091	-0,057	
	cap_um_s	-0,852	-0,779	1	0,954	-0,209	-0,301	0,028	0,029	-0,381	-0,41	-0,259	-0,404	-0,014	-0,108	0,239	-0,081	-0,057	-0,065	
	cap_um_t	-0,78	-0,726	0,954	1	-0,242	-0,434	0,06	0,043	-0,46	-0,452	-0,298	-0,421	-0,013	-0,028	0,193	-0,026	0,015	0,032	
	cap_um_t	-0,301	-0,286	-0,209	-0,242	1	0,819	0,392	0,545	-0,113	-0,291	-0,001	-0,181	0,866	0,294	0,459	0,367	0,331	0,279	
	cap_um_t	-0,147	-0,241	-0,301	-0,434	0,819	1	0,119	0,296	0,164	-0,042	0,155	-0,042	0,711	0,048	0,32	0,183	0,137	0,067	
	emp_1524	-0,174	-0,019	0,028	0,06	0,392	0,119	1	0,949	-0,615	-0,607	-0,337	-0,456	0,325	0,518	0,583	0,142	0,117	0,131	
	emp_1524	-0,267	-0,121	0,029	0,043	0,545	0,296	0,949	1	-0,553	-0,666	-0,272	-0,502	0,483	0,465	0,666	0,23	0,2	0,195	
	unemp_1	0,313	0,171	-0,381	-0,46	-0,113	0,164	-0,615	-0,553	1	0,815	0,7	0,747	-0,241	-0,515	-0,57	-0,089	-0,1	-0,143	
	unemp_1	0,477	0,31	-0,41	-0,452	-0,291	-0,042	-0,607	-0,666	0,815	1	0,483	0,796	-0,409	-0,345	-0,701	-0,214	-0,213	-0,219	
	unemp_2	0,073	-0,035	-0,259	-0,298	-0,001	0,155	-0,337	-0,272	0,7	0,483	1	0,782	-0,112	-0,604	-0,363	-0,013	0,013	-0,008	
	unemp_2	0,359	0,197	-0,404	-0,421	-0,181	-0,042	-0,456	-0,502	0,747	0,796	0,782	1	-0,293	-0,434	-0,659	-0,133	-0,105	-0,1	
	hrst_activ	-0,406	-0,436	-0,014	-0,013	0,866	0,711	0,325	0,483	-0,241	-0,409	-0,112	-0,293	1	0,281	0,496	0,313	0,296	0,263	
	emp_2564	0,072	0,146	-0,108	-0,028	0,294	0,048	0,518	0,465	-0,515	-0,345	-0,604	-0,434	0,281	1	0,523	0,076	0,069	0,078	
	emp_2564	-0,41	-0,256	0,239	0,193	0,459	0,32	0,583	0,666	-0,57	-0,701	-0,363	-0,659	0,496	0,523	1	0,116	0,096	0,093	
	Zscore: m	-0,106	-0,068	-0,081	-0,026	0,367	0,183	0,142	0,23	-0,089	-0,214	-0,013	-0,133	0,313	0,076	0,116	1	0,983	0,925	
	Zscore: m	-0,115	-0,091	-0,057	0,015	0,331	0,137	0,117	0,2	-0,1	-0,213	0,013	-0,105	0,296	0,069	0,096	0,983	1	0,967	
	Zscore: m	-0,076	-0,057	-0,065	0,032	0,279	0,067	0,131	0,195	-0,143	-0,219	-0,008	-0,1	0,263	0,078	0,093	0,925	0,967	1	

Come si può notare, nella matrice sono presenti alcuni valori di correlazione tra coppie di variabili piuttosto elevati; ciò è confermato anche dai livelli di significatività corrispondenti a ciascun valore:

Sig. (1-tail	cap_um_pr_m	0	0	0	0	0,018	0,007	0	0	0	0,149	0	0	0,152	0	0,065	0,051	0,14
	cap_um_p	0	0	0	0	0	0,393	0,042	0,007	0	0,31	0,002	0	0,019	0	0,169	0,098	0,209
	cap_um_s	0	0	0	0,001	0	0,346	0,34	0	0	0	0	0,421	0,062	0	0,127	0,208	0,18
	cap_um_s	0	0	0	0	0	0,196	0,27	0	0	0	0	0,425	0,348	0,003	0,355	0,415	0,327
	cap_um_t	0	0	0,001	0	0	0	0	0,053	0	0,492	0,005	0	0	0	0	0	0
	cap_um_t	0,018	0	0	0	0	0,046	0	0,01	0,277	0,013	0,274	0	0,25	0	0,005	0,026	0,171
	emp_1524	0,007	0,393	0,346	0,196	0	0,046	0	0	0	0	0	0	0	0	0,021	0,049	0,031
	emp_1524	0	0,042	0,34	0,27	0	0	0	0	0	0	0	0	0	0	0	0,002	0,003
	unemp_15	0	0,007	0	0	0,053	0,01	0	0	0	0	0	0	0	0	0,104	0,079	0,021
	unemp_15	0	0	0	0	0	0,277	0	0	0	0	0	0	0	0	0,001	0,001	0,001
	unemp_25	0,149	0,31	0	0	0,492	0,013	0	0	0	0	0	0,056	0	0	0,426	0,426	0,453
	unemp_25	0	0,002	0	0	0,005	0,274	0	0	0	0	0	0	0	0	0,03	0,068	0,078
	hrst_activ	0	0	0,421	0,425	0	0	0	0	0	0,056	0	0	0	0	0	0	0
	emp_2564	0,152	0,019	0,062	0,348	0	0,25	0	0	0	0	0	0	0	0	0,14	0,165	0,135
	emp_2564	0	0	0	0,003	0	0	0	0	0	0	0	0	0	0	0,049	0,086	0,092
	Zscore: m	0,065	0,169	0,127	0,355	0	0,005	0,021	0	0,104	0,001	0,426	0,03	0	0,14	0,049	0	0
	Zscore: m	0,051	0,098	0,208	0,415	0	0,026	0,049	0,002	0,079	0,001	0,426	0,068	0	0,165	0,086	0	0
	Zscore: m	0,14	0,209	0,18	0,327	0	0,171	0,031	0,003	0,021	0,001	0,453	0,078	0	0,135	0,092	0	0

Come ulteriore verifica della fattibilità dell'analisi fattoriale viene riportato l'output riguardante il test di sfericità di Bartlett:

Bartlett's Test of Sphericity	Approx. Chi-Square	7710
	df	153
	Sig.	0

Bartlett's Test of Sphericity	Approx. Chi-Square	7710
	df	153
	Sig.	0

Come si può vedere, il valore della statistica test è elevato ed il livello di significatività è molto basso (prossimo allo 0); si deduce perciò che la matrice di correlazione delle variabili osservate è diversa dalla matrice identità, evidenziando correlazioni significative tra alcune coppie di variabili.

Dalla struttura della matrice di correlazione e dal risultato di tale test sembra quindi indicato procedere con un'analisi fattoriale.

Di seguito vengono riportati gli output risultanti dall'estrazione dei fattori mediante il metodo delle componenti principali. Inizialmente la rotazione scelta è stata di tipo "Oblimin diretto", e la matrice di correlazione tra i fattori risultante era la seguente:

	Component Correlation Matrix		
Component	1	2	3
1	1	-0,184	0,155
2	-0,184	1	-0,014
3	0,155	-0,014	1

Poiché le correlazioni tra le componenti risultavano essere abbastanza basse in valore assoluto, si è potuto passare ad una rotazione ortogonale di tipo Varimax⁶:

Total Variance Explained									
Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	6,378	35,432	35,432	6,378	35,432	35,432	5,019	27,885	27,885
2	3,692	20,514	55,946	3,692	20,514	55,946	3,861	21,449	49,333
3	2,857	15,873	71,819	2,857	15,873	71,819	3,336	18,531	67,864
4	2,287	12,703	84,522	2,287	12,703	84,522	2,998	16,658	84,522
5	0,953	5,293	89,816						
6	0,576	3,202	93,018						
7	0,412	2,288	95,306						
8	0,315	1,748	97,054						
9	0,152	0,847	97,901						
10	0,12	0,669	98,57						
11	0,097	0,539	99,108						
12	0,063	0,35	99,458						
13	0,05	0,277	99,735						
14	0,023	0,128	99,864						
15	0,018	0,102	99,965						
16	0,005	0,03	99,995						
17	0,001	0,004	100						
18	6,41E-05	0	100						

Component Matrix(a)				
	Component			
	1	2	3	4

⁶ La normalizzazione utilizzata di default da SPSS è quella di Kaiser.

cap_um_pr_m	-0,621	0,432	-0,577	0,255
cap_um_pr_f	-0,482	0,392	-0,676	0,333
cap_um_sec_m	0,44	-0,795	0,388	0,022
cap_um_sec_f	0,46	-0,781	0,346	0,164
cap_um_ter_m	0,539	0,63	0,186	-0,447
cap_um_ter_f	0,249	0,611	0,207	-0,623
emp_1524_m	0,696	0,181	-0,397	-0,032
emp_1524_f	0,767	0,273	-0,263	-0,123
unemp_1524_m	-0,78	0,248	0,347	-0,228
unemp_1524_f	-0,869	0,139	0,126	-0,094
unemp_25_m	-0,547	0,238	0,492	-0,262
unemp_25_f	-0,79	0,216	0,28	-0,13
hrst_active	0,626	0,437	0,239	-0,421
emp_2564_m	0,512	0,181	-0,555	0,078
emp_2564_f	0,795	0,059	-0,198	-0,205
Zscore: media_pop<15	0,355	0,552	0,454	0,571
Zscore: media_pop15-64	0,344	0,523	0,484	0,605
Zscore: media_pop65+	0,334	0,493	0,436	0,644

a. 4 components extracted.

Rotated Component Matrix(a)				
	Component			
	1	2	3	4
cap_um_pr_m	-0,132	0,933	-0,283	-0,049
cap_um_pr_f	0,051	0,917	-0,332	-0,026
cap_um_sec_m	0,171	-0,944	-0,229	-0,067
cap_um_sec_f	0,233	-0,9	-0,321	0,023
cap_um_ter_m	0,199	0,012	0,914	0,215
cap_um_ter_f	-0,067	0,09	0,924	0,02
emp_1524_m	0,766	0,056	0,289	0,052
emp_1524_f	0,722	-0,007	0,463	0,114
unemp_1524_m	-0,872	0,259	0,098	-0,074
unemp_1524_f	-0,782	0,378	-0,142	-0,16
unemp_25_m	-0,78	0,06	0,234	0,026
unemp_25_f	-0,822	0,296	-0,008	-0,056
hrst_active	0,253	-0,181	0,828	0,188
emp_2564_m	0,733	0,245	0,111	0,007
emp_2564_f	0,704	-0,209	0,419	-0,006
Zscore: media_pop<15	0,067	0,004	0,162	0,966
Zscore: media_pop15-64	0,048	-0,025	0,124	0,987
Zscore: media_pop65+	0,076	-0,006	0,064	0,974

Rotation Method: Varimax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

Dalla prima tabella di output si vede che il numero di fattori estratti col metodo degli autovalori, ossia considerando solo gli autovalori > 1 , è pari a 4. Questi 4 fattori spiegano l'84,52% della varianza totale, ripartita rispettivamente in 27,88%, 21,45%, 18,53% e 16,66%.

La seconda tabella illustra la matrice contenente i *factor loadings* prima della rotazione (quindi non è molto interessante e interpretabile chiaramente), mentre la terza illustra la matrice contenente i *factor loadings* dopo la rotazione. Si nota che in essa compaiono poche variabili con alte saturazioni su ciascun fattore – che si evidenziano in giallo per chiarezza – e perciò, essendo essa più facilmente interpretabile, si possono assegnare dei nomi ai quattro fattori individuati.

I risultati ottenuti sono riassumibili come segue:

- ✚ il fattore 1 è altamente correlato con le variabili: tassi di occupazione giovanile e non (con segno della correlazione positivo), tassi di disoccupazione giovanile e non (con segno negativo);
- ✚ il fattore 2 è altamente correlato con: capitale umano primario (con segno positivo) e capitale umano secondario (con segno negativo);
- ✚ il fattore 3 è altamente correlato con: capitale umano terziario e risorse umane in Scienza e Tecnologia (con segno positivo);
- ✚ il fattore 4 è altamente correlato con le tre variabili riguardanti la struttura della popolazione.

Sulla base di questi risultati di output, si riportano di seguito i nomi scelti per i quattro fattori: *“Intensità occupazionale”*; *“Mercato del lavoro per qualifiche medio basse”*; *“Disponibilità di capitale umano altamente qualificato”*; *“Struttura demografica”*.

Proprio in base a questi quattro fattori, dopo aver stimato i punteggi fattoriali, vengono effettuate altre tre *clusters analysis*, ciascuna delle quali ha come dati di partenza rispettivamente le 14 regioni del primo gruppo, le 96 del secondo e le 93 del terzo.

L'approccio utilizzato è sempre metodo gerarchico di Ward per determinare il numero di gruppi più plausibile, seguito poi da un'analisi col metodo non gerarchico delle k – medie per definire gli insiemi nello specifico.

CLUSTER 1: le regioni “d’elite”

Si riporta di seguito la scheda agglomerativa:

Cluster Combined				Stage Cluster First Appears		
Stage	Cluster 1	Cluster 2	Coefficients	Cluster 1	Cluster 2	Next Stage
1	27	135	0,106	0	0	3
2	10	175	0,247	0	0	6
3	27	28	0,477	1	0	5
4	18	184	0,905	0	0	9
5	27	120	1,503	3	0	6
6	10	27	2,668	2	5	8
7	41	198	4,122	0	0	9
8	10	183	6,163	6	0	11
9	18	41	10,039	4	7	10
10	1	18	17,239	0	9	12
11	10	100	24,857	8	0	13
12	1	74	36,094	10	0	13
13	1	10	54,747	12	11	0

Il “salto” maggiore si verifica in corrispondenza del passaggio dalla penultima all’ultima riga; perciò il numero di sottogruppi più indicato sembra essere 2.

Questi due insiemi definiti col metodo delle k – medie sono così strutturati:

Sottogruppo 1.1 (n=6)	Sottogruppo 1.2 (n=8)
Région de Bruxelles-Capitale	Praha
Hovedstaden	Bremen
Southern and Eastern	Hamburg
Île de France	Provincia Autonoma di Bolzano
Stockholm	Luxembourg
London	Wien
	Bratislavský kraj
	Åland

In cosa differiscono tra di loro questi due sottoinsiemi?

Questo può essere compreso osservando la tabella ANOVA, i centri finali dei *clusters* e le medie di ciascuna delle variabili osservate che hanno determinato i quattro fattori:

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Intensità occupazionale	2,944229733	1	0,909927209	12	3,235676	0,097231
Mercato del lavoro per qualifica medio-basse	1,363467735	1	0,289201843	12	4,714589	0,050676
Disponibilità di capitale umano altamente qualificato	8,55497244	1	0,509822805	12	16,78029	0,001482
Struttura demografica	5,789815646	1	1,298915177	12	4,457424	0,056405

Final Cluster Centers		
	1.1	1.2
Intensità occupazionale	-0,08112	0,84556
Mercato del lavoro per qualifica medio-basse	0,2341	-0,39652
Disponibilità di capitale umano altamente qualificato	2,09356	0,51394
Struttura demografica	0,58765	-0,71184

Cluster	media 1.1	media 1.2
cap_um_pr_m	0,17	0,17
cap_um_pr_f	0,15	0,15
cap_um_sec_m	0,35	0,51
cap_um_sec_f	0,31	0,52
cap_um_ter_m	0,39	0,27
cap_um_ter_f	0,46	0,29
emp_1524_m	0,39	0,40
emp_1524_f	0,39	0,35
unemp_1524_m	0,22	0,12
unemp_1524_f	0,19	0,11
unemp_25_m	0,08	0,05
unemp_25_f	0,07	0,04
hrst_active	0,52	0,46
emp_2564_m	0,81	0,83
emp_2564_f	0,69	0,70
Zscore: media_pop<15	1,23	-0,67
Zscore: media_pop15-64	0,90	-0,64
Zscore: media_pop65+	0,28	-0,60

Dalla tabella ANOVA si evince che il fattore che ha maggiormente contribuito alla definizione dei *clusters* è quello relativo al capitale umano altamente qualificato, mentre quello che ha contribuito di meno è il fattore relativo al mercato del lavoro.

La maggiore significatività del terzo fattore è confermata anche nelle altre due tabelle: infatti la distanza maggiore tra i centri finali dei due sottogruppi si verifica proprio in corrispondenza di tale fattore (a favore del gruppo 1.1). Analizzando più approfonditamente le variabili che hanno indotto tale fattore si vede che c'è una differenza nelle medie di 12 e di 17 punti percentuali per quanto riguarda il capitale umano rispettivamente maschile e femminile nei due gruppi.

CLUSTER 2: le regioni sviluppate

Tale *cluster* è di particolare interesse: infatti al suo interno vi troviamo anche la Regione del Veneto. Si parte ancora una volta dalla scheda agglomerativa:

Stage	Cluster Combined			Stage Cluster First Appears		
	Cluster 1	Cluster 2	Coefficients	Cluster 1	Cluster 2	Next Stage
1	71	72	0,003	0	0	68
2	130	132	0,015	0	0	75
3	75	80	0,029	0	0	15
4	19	22	0,043	0	0	21
5	134	137	0,059	0	0	27
6	190	191	0,084	0	0	47
7	140	141	0,11	0	0	17
8	55	63	0,144	0	0	54
9	194	196	0,181	0	0	49
10	181	187	0,22	0	0	33
11	20	129	0,262	0	0	36
12	57	60	0,304	0	0	40
13	103	107	0,347	0	0	28
14	185	188	0,407	0	0	33
15	75	77	0,468	3	0	37
16	83	92	0,53	0	0	53
17	138	140	0,607	0	7	24
18	2	193	0,686	0	0	30
19	197	200	0,775	0	0	34
20	86	88	0,865	0	0	45
21	19	21	0,959	4	0	36
22	96	105	1,055	0	0	44
23	78	85	1,15	0	0	56
24	138	139	1,268	17	0	60
25	121	170	1,389	0	0	63

26	102	104	1,511	0	0	44
27	134	136	1,641	5	0	60
28	103	106	1,773	13	0	43
29	186	189	1,925	0	0	48
30	2	91	2,086	18	0	49
31	179	182	2,248	0	0	70
32	51	94	2,421	0	0	50
33	181	185	2,597	10	14	47
34	195	197	2,777	0	19	51
35	33	37	2,961	0	0	58
36	19	20	3,148	21	11	52
37	75	98	3,348	15	0	73
38	68	70	3,549	0	0	54
39	59	61	3,766	0	0	55
40	56	57	4,004	0	12	55
41	23	24	4,255	0	0	76
42	35	143	4,522	0	0	63
43	101	103	4,793	0	28	71
44	96	102	5,068	22	26	85
45	86	89	5,365	20	0	56
46	131	199	5,694	0	0	78
47	181	190	6,037	33	6	62
48	174	186	6,383	0	29	69
49	2	194	6,778	30	9	59
50	51	81	7,172	32	0	65
51	195	202	7,617	34	0	74
52	19	117	8,067	36	0	75
53	40	83	8,531	0	16	62
54	55	68	9,077	8	38	64
55	56	59	9,649	40	39	80
56	78	86	10,231	23	45	73
57	66	67	10,848	0	0	67
58	33	34	11,532	35	0	79
59	2	29	12,226	49	0	74
60	134	138	12,972	27	24	87
61	53	97	13,723	0	0	71
62	40	181	14,608	53	47	69
63	35	121	15,5	42	25	83
64	55	161	16,431	54	0	80
65	51	108	17,366	50	0	81
66	58	180	18,341	0	0	77
67	62	66	19,401	0	57	81
68	71	73	20,485	1	0	90
69	40	174	21,748	62	48	84
70	25	179	23,22	0	31	77

71	53	101	24,77	61	43	82
72	31	99	26,349	0	0	76
73	75	78	27,958	37	56	79
74	2	195	29,744	59	51	78
75	19	130	31,848	52	2	87
76	23	31	34,12	41	72	86
77	25	58	36,816	70	66	84
78	2	131	39,53	74	46	88
79	33	75	42,262	58	73	83
80	55	56	45,643	64	55	90
81	51	62	49,593	65	67	88
82	53	164	53,552	71	0	85
83	33	35	58,014	79	63	89
84	25	40	63,774	77	69	89
85	53	96	69,74	82	44	91
86	23	32	79,265	76	0	95
87	19	134	90,66	75	60	91
88	2	51	102,418	78	81	92
89	25	33	114,387	84	83	92
90	55	71	127,737	80	68	93
91	19	53	154,339	87	85	94
92	2	25	187,67	88	89	93
93	2	55	230,472	92	90	94
94	2	19	290,163	93	91	95
95	2	23	368,838	94	86	0

Il primo “salto” consistente si verifica tra la terzultima e la penultima riga, perciò il numero di sottogruppi più opportuno in questo caso sembra essere pari a tre.

Procedendo con il metodo delle k-medie si ottiene:

Sottogruppo 2.1 (n=55)	Sottogruppo 2.2 (n=6)	Sottogruppo 2.3 (n=35)
Vlaams Gewest	Baden-Württemberg	Sjælland
Berlin	Bayern	Syddanmark
Hessen	Niedersachsen	Midtjylland
Sachsen	Nordrhein-Westfalen	Nordjylland
Border, Midland and Western	Lombardia	Rheinland-Pfalz
Attiki	South East (UK)	Saarland
Galicia		Schleswig-Holstein
Principado de Asturias		Notio Aigaio
Cantabria		Champagne-Ardenne

País Vasco	Centre (FR)
Comunidad Foral de Navarra	Bourgogne
La Rioja	Piemonte
Aragón	Valle d'Aosta/Vallée d'Aoste
Comunidad de Madrid	Liguria
Castilla y León	Provincia Autonoma Trento
Cataluña	<i>Veneto</i>
Comunidad Valenciana	Friuli-Venezia Giulia
Illes Balears	Emilia-Romagna
Región de Murcia	Toscana
Ciudad Autónoma de Ceuta (ES)	Umbria
Ciudad Autónoma de Melilla (ES)	Marche
Canarias (ES)	Noord-Nederland
Haute-Normandie	Oost-Nederland
Nord - Pas-de-Calais	Zuid-Nederland
Alsace	Niederösterreich
Pays de la Loire	Kärnten
Bretagne	Steiermark
Aquitaine	Oberösterreich
Midi-Pyrénées	Salzburg
Rhône-Alpes	Tirol
Auvergne	Vorarlberg
Provence-Alpes-Côte d'Azur	Região Autónoma da Madeira (PT)
Lazio	Zahodna Slovenija
Cyprus	Småland med öarna
Közép-Magyarország	Norra Mellansverige
West-Nederland	
Mazowieckie	
Lisboa	
Bucuresti - Ilfov	
Itä-Suomi	
Etelä-Suomi	
Länsi-Suomi	
Pohjois-Suomi	

Östra Mellansverige
 Sydsverige
 Västsverige
 Mellersta Norrland
 Övre Norrland
 North West (UK)
 Yorkshire and The Humber
 East Midlands (UK)
 West Midlands (UK)
 East of England
 South West (UK)
 Scotland

In base a questa analisi, le regioni con le quali ha senso confrontare il Veneto perché simili ad esso riguardo le principali variabili economiche e socio – demografiche sono le regioni rientranti nel sottogruppo 2.3.

Si desidera ora comprendere nello specifico in cosa differiscono questi tre sottoinsiemi:

	Cluster		Error			
	Mean Square	df	Mean Square	df	F	Sig.
Intensità occupazionale	16,291	2	0,668	93	24,406	0
Mercato del lavoro per qualifica medio-basse	1,275	2	0,542	93	2,354	0,101
Disponibilità di capitale umano altamente qualificato	16,063	2	0,458	93	35,05	0
Struttura demografica	39,933	2	0,717	93	55,724	0

Final Cluster Centers			
	Cluster		
	2.1	2.2	2.3
Intensità occupazionale	-0,14578	0,83688	1,06002
Mercato del lavoro per qualifica medio-basse	0,33929	-0,22201	0,0891
Disponibilità di capitale umano altamente qualificato	0,83015	-0,25657	-0,35248
Struttura demografica	0,15772	3,6289	-0,31339

Distances between Final Cluster Centers			
Cluster	2.1	2.2	2.3
2.1		3,809	1,771
2.2	3,809		3,962
2.3	1,771	3,962	

Cluster	media 2.1	media 2.2	media 2.3
cap_um_pr_m	0,25	0,16	0,25
cap_um_pr_f	0,21	0,17	0,23
cap_um_sec_m	0,39	0,50	0,50
cap_um_sec_f	0,35	0,54	0,48
cap_um_ter_m	0,29	0,29	0,20
cap_um_ter_f	0,36	0,24	0,23
emp_1524_m	0,38	0,47	0,48
emp_1524_f	0,34	0,43	0,41
unemp_1524_m	0,23	0,12	0,13
unemp_1524_f	0,23	0,11	0,15
unemp_25_m	0,07	0,05	0,04
unemp_25_f	0,08	0,05	0,05
hrst_active	0,40	0,41	0,35
emp_2564_m	0,80	0,84	0,82
emp_2564_f	0,66	0,69	0,67
Zscore: media_pop<15	0,28	3,26	-0,33
Zscore: media_pop15-64	0,25	3,52	-0,33
Zscore: media_pop65+	0,21	3,94	-0,21

Innanzitutto si nota che il fattore che è stato meno significativo nella determinazione dei gruppi è il secondo, cioè quello riguardante il capitale umano a qualifica medio – bassa, benché le regioni 2.2 presentino una quota di lavoratori con capitale umano basso nettamente inferiore a quella delle regioni 2.1 e 2.3.

Osservando la matrice delle distanze tra i centri finali dei sottoinsiemi, si nota che le differenze medie più elevate sono quelle che intercorrono tra il sottogruppo 2.2 e gli altri due. Dalle altre due tabelle sembra che il gruppo 2.2 si distacchi dagli altri due principalmente per la popolazione significativamente più elevata in media. Gli insiemi 2.1 e 2.3 sembrano invece differire per quanto riguarda la situazione del mercato del lavoro ed il grado di qualificazione del personale richiesto in esso (con le regioni del gruppo 2.1 che presentano una maggiore

disponibilità di capitale umano ad alta specializzazione ma altresì con un mercato che assorbe anche molti con qualificazioni inferiori).

CLUSTER 3: le regioni con un livello di sviluppo inferiore

Si riporta di seguito la scheda agglomerativa:

Cluster Combined				Stage Cluster First Appears		
Stage	Cluster 1	Cluster 2	Coefficients	Cluster 1	Cluster 2	Next Stage
1	11	12	0	0	0	52
2	4	152	0,003	0	0	22
3	142	144	0,017	0	0	29
4	122	123	0,035	0	0	45
5	7	9	0,053	0	0	45
6	5	148	0,072	0	0	39
7	146	147	0,092	0	0	14
8	17	176	0,114	0	0	41
9	79	84	0,147	0	0	21
10	14	16	0,182	0	0	33
11	42	48	0,22	0	0	24
12	124	177	0,265	0	0	25
13	111	115	0,319	0	0	60
14	146	156	0,377	7	0	37
15	125	126	0,436	0	0	49
16	47	49	0,496	0	0	55
17	36	118	0,56	0	0	30
18	159	162	0,624	0	0	56
19	127	155	0,697	0	0	37
20	26	38	0,784	0	0	46
21	79	90	0,875	9	0	63
22	4	154	0,967	2	0	50
23	192	201	1,061	0	0	42
24	42	46	1,158	11	0	57
25	124	178	1,257	12	0	49
26	76	82	1,359	0	0	31
27	168	169	1,463	0	0	58
28	165	172	1,57	0	0	38
29	142	157	1,688	3	0	47
30	30	36	1,808	0	17	51
31	76	87	1,936	26	0	63
32	52	110	2,064	0	0	53
33	14	15	2,195	10	0	52
34	45	54	2,328	0	0	61
35	95	109	2,464	0	0	53

36	3	93	2,632	0	0	80
37	127	146	2,799	19	14	50
38	165	171	2,98	28	0	58
39	5	6	3,163	6	0	43
40	113	116	3,348	0	0	48
41	13	17	3,534	0	8	67
42	192	203	3,728	23	0	75
43	5	149	3,932	39	0	70
44	151	153	4,143	0	0	65
45	7	122	4,379	5	4	66
46	8	26	4,65	0	20	64
47	142	150	4,924	29	0	62
48	113	114	5,206	40	0	82
49	124	125	5,511	25	15	83
50	4	127	5,818	22	37	65
51	30	119	6,137	30	0	71
52	11	14	6,476	1	33	76
53	52	95	6,822	32	35	73
54	160	163	7,179	0	0	72
55	47	50	7,543	16	0	77
56	128	159	7,924	0	18	77
57	42	44	8,31	24	0	69
58	165	168	8,703	38	27	68
59	64	65	9,118	0	0	80
60	111	112	9,606	13	0	79
61	43	45	10,125	0	34	69
62	142	145	10,68	47	0	78
63	76	79	11,275	31	21	75
64	8	39	11,879	46	0	71
65	4	151	12,496	50	44	70
66	7	173	13,124	45	0	74
67	13	166	13,787	41	0	74
68	165	167	14,463	58	0	85
69	42	43	15,226	57	61	73
70	4	5	16,182	65	43	78
71	8	30	17,352	64	51	87
72	158	160	18,574	0	54	84
73	42	52	19,966	69	53	82
74	7	13	21,405	66	67	81
75	76	192	23,048	63	42	87
76	11	133	24,823	52	0	81
77	47	128	26,61	55	56	84
78	4	142	28,782	70	62	83
79	69	111	32,019	0	60	88
80	3	64	35,374	36	59	86

81	7	11	39,307	74	76	85
82	42	113	43,415	73	48	86
83	4	124	47,658	78	49	89
84	47	158	53,376	77	72	91
85	7	165	59,152	81	68	89
86	3	42	65,898	80	82	88
87	8	76	73,742	71	75	90
88	3	69	91,994	86	79	91
89	4	7	112,556	83	85	90
90	4	8	144,748	89	87	92
91	3	47	178,678	88	84	92
92	3	4	280,485	91	90	0

Anche in questo caso il primo “salto” significativo corrisponde al passaggio tra la terzultima e la penultima riga, perciò il numero di sotto – *clusters* più indicato sembra essere tre.

Nello specifico, con il metodo delle k – medie, possiamo distinguerli così:

Sottogruppo 3.1 (n=18)	Sottogruppo 3.2 (n=62)	Sottogruppo 3.3 (n=13)
Anatoliki Makedonia, Thraki	Severozapaden	Région wallonne
Thessalia	Severen tsentralen	Kentriki Makedonia
Ipeiros	Severoiztochen	Dytiki Makedonia
Ionia Nisia	Yugoiztochen	Extremadura
Dytiki Ellada	Yugozapaden	Andalucía
Stereia Ellada	Yuzhen tsentralen	Languedoc-Roussillon
Peloponnisos	Strední Cechy	Molise
Voreio Aigaio	Jihozápad	Campania
Kriti	Severozápad	Puglia
Castilla-la Mancha	Severovýchod	Basilicata
Corse	Jihovýchod	Calabria
Abruzzo	Strední Morava	Sicilia
Malta	Moravskoslezsko	Sardegna
Norte	Brandenburg	
Algarve	Mecklenburg-Vorpommern	
Centro (PT)	Sachsen-Anhalt	
Alentejo	Thüringen	
Região Autónoma dos	Estonia	
Açores (PT)	Picardie	

Basse-Normandie
Lorraine
Franche-Comté
Poitou-Charentes
Limousin
Latvia
Lithuania
Közép-Dunántúl
Nyugat-Dunántúl
Dél-Dunántúl
Észak-Magyarország
Észak-Alföld
Dél-Alföld
Burgenland (AT)
Łódzkie
Małopolskie
Ślaskie
Lubelskie
Podkarpackie
Świętokrzyskie
Podlaskie
Wielkopolskie
Zachodniopomorskie
Lubuskie
Dolnośląskie
Opolskie
Kujawsko-Pomorskie
Warmińsko-Mazurskie
Pomorskie
Nord-Vest
Centru
Nord-Est
Sud-Est
Sud - Muntenia

Sud-Vest Oltenia
 Vest
 Vzhodna Slovenija
 Západné Slovensko
 Stredné Slovensko
 Východné Slovensko
 North East (UK)
 Wales
 Northern Ireland (UK)

Per capire in che cosa differiscono questi tre sottoinsiemi si analizzano i seguenti output:

	Cluster		Error		F	Sig.
	Mean Square	df	Mean Square	df		
Intensità occupazionale	10,253	2	0,45	90	22,783	0
Mercato del lavoro per qualifica medio-basse	51,987	2	0,355	90	146,475	0
Disponibilità di capitale umano altamente qualificato	2,363	2	0,58	90	4,072	0,02
Struttura demografica	2,81	2	0,233	90	12,053	0

Final Cluster Centers			
	Cluster		
	3.1	3.2	3.3
Intensità occupazionale	0,00175	-0,32562	-1,5558
Mercato del lavoro per qualifica medio-basse	1,64554	-0,92319	0,68754
Disponibilità di capitale umano altamente qualificato	-0,9808	-0,40599	-0,43303
Struttura demografica	-0,49695	-0,20993	0,35773

Distances between Final Cluster Centers			
Cluster	3.1	3.2	3.3
3.1		2,668	2,091
3.2	2,668		2,105
3.3	2,091	2,105	

Cluster	media 3.1	media 3.2	media 3.3
cap_um_pr_m	0,52656	0,11487	0,41184
cap_um_pr_f	0,40426	0,12126	0,26719
cap_um_sec_m	0,26441	0,63132	0,32699
cap_um_sec_f	0,26865	0,55099	0,33002
cap_um_ter_m	0,15033	0,17838	0,16881
cap_um_ter_f	0,21719	0,24829	0,26258
emp_1524_m	0,33789	0,32644	0,23348
emp_1524_f	0,22042	0,25333	0,1495
unemp_1524_m	0,22004	0,20245	0,30881
unemp_1524_f	0,31047	0,20987	0,38671
unemp_25_m	0,06012	0,07363	0,09133
unemp_25_f	0,10796	0,07798	0,13892
hrst_active	0,23735	0,30558	0,31019
emp_2564_m	0,81972	0,76049	0,73028
emp_2564_f	0,55498	0,62581	0,44631
Zscore: media_pop<15	-0,6378	-0,25833	0,176534
Zscore: media_pop15-64	-0,64013	-0,22673	0,168598
Zscore: media_pop65+	-0,54947	-0,2761	0,164499

In questo caso sembra che i tre sottogruppi differiscano più uniformemente tra di loro per tutti i quattro fattori.