

COVID-19: ANALISI DELLE DETERMINANTI DELLA DIFFUSIONE DEL CONTAGIO.

Bianchino Antonella¹, Daniela Fusco², Maria Liguori³

SOMMARIO

La diffusione del contagio da SARS-COVID19 non è stata omogenea su tutto il territorio italiano. Se è pur vero che la maggiore concentrazione in Lombardia è dovuta alla presenza del primo focolaio, esistono domande aperte sui motivi che hanno determinato questa forte difformità interregionale.

Dare risposta a queste domande può consentire alle regioni di mettere in piedi delle scelte che possano ridurre la potenziale diffusione di nuovi rischi biologici.

Per identificare le determinanti del contagio è necessario fare riferimento ai dati oggettivi raccolti in Italia nel periodo di diffusione della pandemia. Lo studio analizza, a livello regionale, indicatori chiave quali l'età della popolazione, la densità abitativa, i rapporti commerciali con l'estero, il pendolarismo e l'inquinamento. I risultati raggiunti con l'utilizzo dell'analisi fattoriale mostrano in che misura tali fenomeni abbiano contribuito alla diffusione del contagio rendendo alcune regioni più esposte rispetto ad altre.

¹ Istat, Dipartimento per lo sviluppo di metodi e tecnologie per la produzione statistica, Potenza, e-mail: bianchin@istat.it.

² Istat, Dipartimento per lo sviluppo di metodi e tecnologie per la produzione statistica, Napoli, e-mail: dafusco@istat.it (corresponding author).

³ Istat, Dipartimento per lo sviluppo di metodi e tecnologie per la produzione statistica, Napoli, e-mail: liguori@istat.it.

1. Introduzione

I primi due casi di Coronavirus in Italia, una coppia di turisti cinesi, sono stati confermati il 30 gennaio dall'Istituto Spallanzani, dove sono stati ricoverati in isolamento dal 29 gennaio. Il primo caso di trasmissione secondaria si è verificato a Codogno, Comune della Lombardia in provincia di Lodi, il 18 febbraio 2020. Il 10 marzo, giorno di entrata in vigore del d.p.c.m. del 08/03/2020 che decreta l'inizio del lockdown, i casi arrivano a 10.149. Al termine della fase 1, 3 maggio 2020, sono stati registrati 210.717 contagi ufficiali e 28.884 decessi.

La diffusione del virus non è stata omogenea su tutto il territorio italiano e le peculiarità regionali sono rese evidenti non solo dalla curva dei contagi, ma soprattutto dalle curve che esprimono la velocità e l'accelerazione del contagio.

Se da un punto di vista medico terapie, immunità, origine restano ad oggi domande aperte, da un punto di vista politico la comprensione dei fattori che hanno determinato la diffusione del contagio consentirebbero alle regioni di mettere in piedi delle scelte che possono ridurre la potenziale diffusione di nuovi rischi biologici.

Per identificare le determinanti del contagio è necessario fare riferimento ai dati oggettivi raccolti in Italia nel periodo di diffusione della pandemia. Secondo tali dati la popolazione più a rischio è quella di sesso maschile con più di 61 anni (ISS, 2020). Inoltre è assodato che l'indice R_0 , ovvero il "numero di riproduzione di base", aumenta in proporzione ai contatti tra la popolazione infetta e popolazione sana. Non ultimo, studi recenti hanno dimostrato che l'esposizione a fattori inquinanti determina un incremento di morbosità e mortalità umana legata ad affezioni delle vie respiratorie (WHO, 2016).

Lo studio analizza, a livello regionale, indicatori chiave quali l'età della popolazione, la densità abitativa, i rapporti commerciali con l'estero, il pendolarismo e l'inquinamento. I risultati raggiunti con l'utilizzo dell'analisi fattoriale mostrano in che misura tali elementi abbiano contribuito alla diffusione del contagio rendendo alcune regioni più esposte rispetto ad altre.

2. Covid-19: il contesto italiano

Al 1° gennaio 2019 si stima che la popolazione italiana ammonti a 60 milioni 391mila residenti, con una contrazione dell'1,5 per mille rispetto all'anno precedente (Istat, 2020).

L'Italia si conferma un Paese che sta invecchiando: nel 2018 si conteggiano 449mila nascite, ossia 9mila in meno del precedente minimo registrato nel 2017. Di conseguenza si registra un saldo naturale negativo (-187mila), il secondo livello più basso nella storia dopo quello del 2017 (-191mila), una media di figli per donna di 1,3 ben al di sotto del ricambio generazionale.

I decessi sono 636mila, 13mila in meno del 2017. In rapporto al numero di residenti, nel 2018 sono deceduti 10,5 individui ogni mille abitanti, contro i 10,7 del 2017.

L'età media della popolazione italiana è di 46 anni, con una varianza territoriale (NUTS3) di 2,2. Il 29% della popolazione totale ha più di 60 anni.

L'indice di vecchiaia (Tav. 1), calcolato come rapporto tra popolazione di 65 anni e più e popolazione di età 0-14 anni moltiplicato per 100, è aumentato progressivamente negli anni fino ad arrivare ad un valore di 177,9 nel 2020.

Le aree dove il numero di persone anziane è più elevato sono quelle del Nord, in particolare il nord-ovest.

Più giovani invece le regioni del sud Italia. In base ai valori stimati per il 2020, la regione più vecchia si conferma essere la Liguria, la più giovane è la Campania.

Tavola 1 – Indice di vecchiaia ed età metà della popolazione italiana per Regione, anni 2017-2020

Regione	2017		2018		2019		2020*	
	Indice di vecchiaia	Età media	Indice di vecchiaia	Età media	Indice di vecchiaia	Età media	Indice di vecchiaia	Età media
Piemonte	197,6	46,6	201,3	46,8	205,9	47,0	210,3	47,2
Valle d'Aosta	171,1	45,6	176,0	45,9	181,6	46,1	187,2	46,4
Lombardia	159,1	44,8	162,2	45,0	165,5	45,2	169,1	45,4
Trentino-Alto Adige	133,5	43,3	135,8	43,5	138,5	43,7	142,4	44,0
Veneto	163,6	45,1	167,7	45,4	172,1	45,6	177,6	45,9
Friuli-Venezia Giulia	208,8	47,1	212,9	47,3	217,2	47,5	223,2	47,7
Liguria	249,8	48,7	252,4	48,8	255,8	49,0	260,4	49,1
Emilia-Romagna	177,8	45,9	180,1	46,1	182,6	46,2	185,6	46,3
Toscana	198,6	46,7	201,4	46,8	204,6	47,0	209,1	47,2
Umbria	195,9	46,5	199,4	46,7	204,2	47,0	209,6	47,2
Marche	187,6	46,1	191,3	46,3	196,2	46,6	201,8	46,9
Lazio	155,5	44,5	158,5	44,8	162,6	45,1	167,3	45,3
Abruzzo	183,9	45,7	187,6	45,9	191,8	46,2	197,6	46,4
Molise	206,9	46,3	211,2	46,5	217,5	46,8	224,8	47,1
Campania	121,6	42,1	125,2	42,3	129,8	42,7	134,7	43,0
Puglia	157,3	44,1	162,5	44,4	168,6	44,7	175,0	45,1
Basilicata	181,2	45,2	186,7	45,5	193,2	45,8	200,3	46,1
Calabria	155,0	43,9	158,5	44,2	163,3	44,5	168,3	44,8
Sicilia	145,8	43,5	149,3	43,7	153,7	44,0	158,6	44,3
Sardegna	195,5	46,1	202,7	46,4	212,0	46,8	221,2	47,2
ITALIA	165,3	44,9	168,9	45,2	173,1	45,4	177,9	45,7

*dato provvisorio

Fonte Istat

In merito alla densità della popolazione, l'Italia si conferma un paese popoloso con 200 abitanti per chilometro quadrato in media. Le regioni più densamente popolate, con le loro specificità territoriali, sono la Campania, con 424 abitanti per Km², e la Lombardia, con 422. La regione meno popolosa è la valle d'Aosta con 38 abitanti per Km², seguita da Molise e Sardegna, entrambe con 68.

Ogni anno i decessi rappresentano circa l'1% della popolazione totale. Nel 2017 (ultimo anno disponibile) la causa principale di morte sono state le malattie del sistema circolatorio, seguite dai tumori e dalle malattie del sistema respiratorio. Queste ultime includono malattie come la polmonite e l'influenza. Esse hanno rappresentato da sole il 71% del totale dei decessi per affezione delle vie respiratorie.

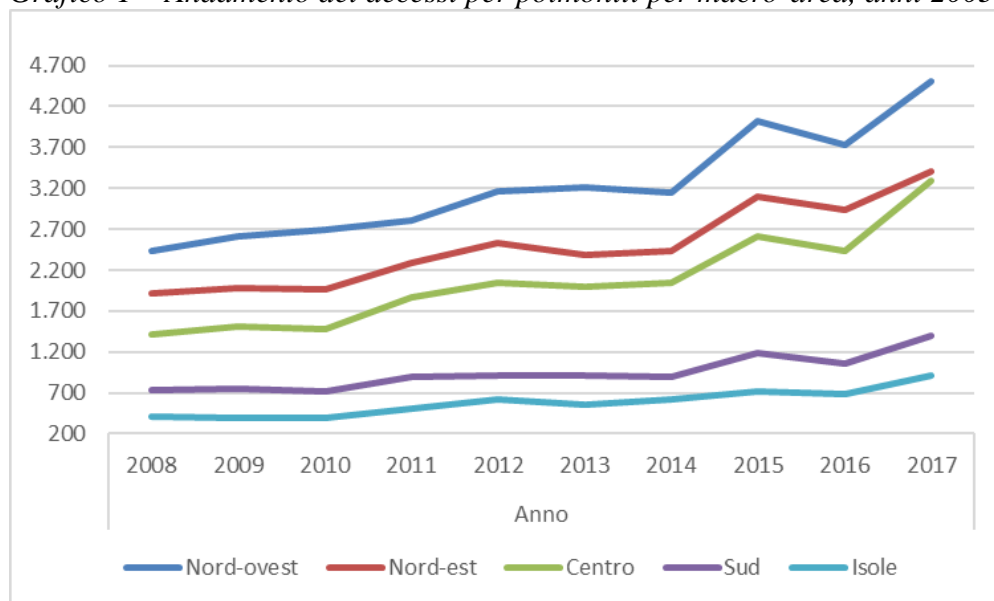
A fronte di un quantitativo di decessi per malattie del sistema circolatorio piuttosto altalenante, c'è stato un progressivo aumento di morti per tumore e per malattie del sistema respiratorio. In particolare queste ultime sono incrementate del 28% tra il 2013 e il 2017.

Focalizzando l'attenzione sulle malattie del sistema respiratorio, esse sono rappresentate per la quasi totalità da malattie croniche delle basse vie respiratorie (49%) e altre malattie croniche delle basse vie respiratorie (48%). Le polmoniti hanno rappresentato nel 2017 il 25% di questo sottogruppo di malattie con 13.516 decessi. Tra il 2013 e il 2017 esse hanno subito un incremento del 49%, dato ben al di sopra della media di crescita delle altre patologie di questa macro-area.

Da un punto di vista geografico, il grafico 1 mostra come l'incremento maggiore nella diffusione delle polmoniti ci sia stato nel nord Italia, in particolare nel nord-Ovest, sebbene a livello percentuale gli incrementi siano notevoli in tutte le ripartizioni geografiche.

In ogni caso, nel 2017 il 59% dei decessi per polmonite si sono registrati nel nord Italia, il 24% al centro e il 17% al sud Italia e isole. Nello stesso anno il totale dei decessi del nord ha rappresentato il 46% del totale Italia, quelli del centro il 21% e quelli del mezzogiorno il 33%.

Grafico 1 – Andamento dei decessi per polmoniti per macro-area, anni 2003-2017



Fonte Istat

Il 3 maggio 2020 sono 210.717 i casi totali dall'inizio della pandemia. Le persone guarite sono 81.654 e 28.884 deceduti. In termini percentuali (tavola 2) vuol dire che Italia è stato colpito il 3,5 per mille della popolazione. Il tasso di mortalità è del 14% (calcolata sui contagiati), quello di sopravvivenza è del 39%.

In termini assoluti la regione più colpita è la Lombardia, seguita dal Piemonte e dall'Emilia Romagna. In riferimento al totale della popolazione, con 9 contagiati ogni 1.000 abitanti è la Valle d'Aosta a detenere il triste primato, seguita da Lombardia e provincia di Trento, entrambe con 8 contagiati su 1.000, e Piemonte (6).

Le regioni meno colpite sono quelle del Mezzogiorno, in primis (in valori relativi) la Calabria con 0,57 contagi ogni 1.000 abitanti.

Sempre facendo riferimento alla popolazione, la regione in cui sono stati fatti il maggior numero di tamponi è la provincia di Bolzano (83 per mille), seguita dal Veneto (77). Il minor numero di tamponi è invece stato fatto in Campania (15).

In riferimento al totale dei casi è l'Umbria (82%) ad avere registrato il maggior numero di guariti, seguita dalla Valle d'Aosta (78%). Il minor numero di guarigioni si registrano invece in Puglia (18%) e in Sicilia (24%).

La maggiore mortalità c'è stata in Lombardia (18%), seguita da Liguria, Marche ed Emilia Romagna (14%). All'opposto, la minore mortalità si è avuta in Umbria (5) e in Basilicata (6).

Considerando la curva dei contagi⁴ nel periodo tra il 24 febbraio e il 3 maggio, il picco è stato raggiunto nelle settimane a cavallo tra la fine di marzo e l'inizio di aprile (Grafico 2), periodo in cui i casi sono più che raddoppiati.

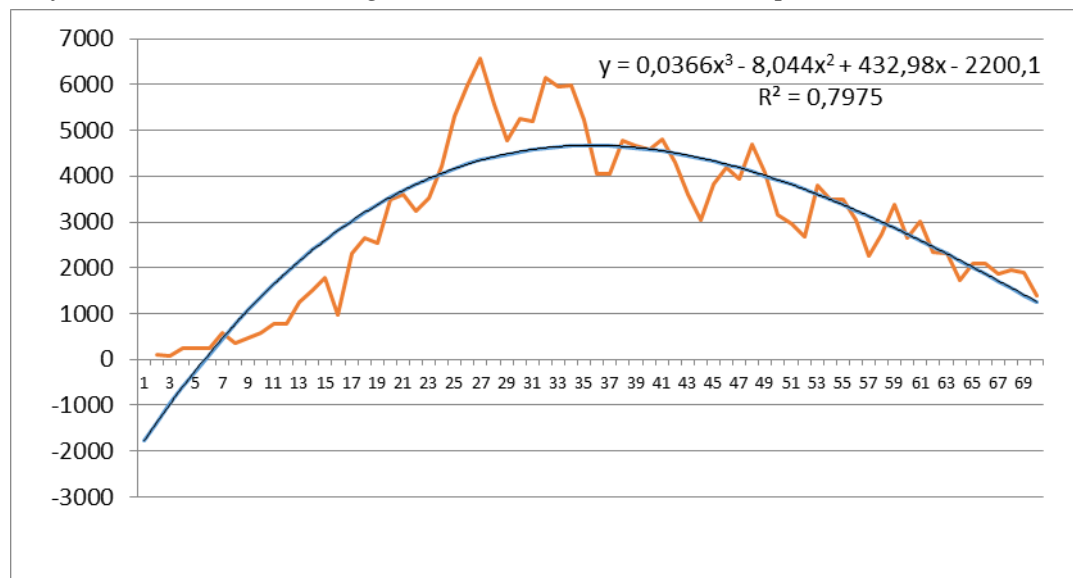
⁴ Intesa come linea di tendenza polinomiale di ordine sesto relativa alle variazioni dei contagi giornalieri

Tavola 2 – Diffusione del contagio in Italia al 3 maggio 2020 per Regione e totale Italia

Regione	03-mag				Totale casi (su 1.000 ab)	incidenze percentuali su tot casi		Guariti/dec essi	Tamponi (su 1.000 ab)
	Totale casi	Guariti	Decessi	Tamponi		Guariti	Decessi		
Piemonte	27.430	8.640	3.152	172.208	6,30	31,50	11,49	2,74	39,53
Valle d'Aosta	1.142	895	138	8.100	9,09	78,37	12,08	6,49	64,46
Lombardia	77.528	26.371	14.231	410.857	7,71	34,01	18,36	1,85	40,84
P.A. Bolzano	2.536	1.590	281	44.240	4,77	62,70	11,08	5,66	83,29
P.A. Trento	4.247	2.571	429	41.095	7,85	60,54	10,10	5,99	75,95
Veneto	18.318	9.503	1.516	378.202	3,73	51,88	8,28	6,27	77,09
Friuli Venezia Giulia	3.072	1.688	297	74.990	2,53	54,95	9,67	5,68	61,71
Liguria	8.359	3.599	1.209	54.492	5,39	43,06	14,46	2,98	35,14
Emilia-Romagna	26.016	13.329	3.642	197.075	5,83	51,23	14,00	3,66	44,19
Toscana	9.563	3.363	872	150.914	2,56	35,17	9,12	3,86	40,46
Umbria	1.394	1.143	68	38.823	1,58	81,99	4,88	16,81	44,02
Marche	6.319	2.194	927	64.412	4,14	34,72	14,67	2,37	42,23
Lazio	6.809	1.916	508	150.912	1,16	28,14	7,46	3,77	25,67
Abruzzo	2.996	798	330	40.699	2,28	26,64	11,01	2,42	31,03
Molise	301	98	22	7.075	0,98	32,56	7,31	4,45	23,15
Campania	4.484	1.394	364	86.498	0,77	31,09	8,12	3,83	14,91
Puglia	4.144	765	424	66.443	1,03	18,46	10,23	1,80	16,49
Basilicata	386	167	25	14.210	0,69	43,26	6,48	6,68	25,25
Calabria	1.114	324	88	38.835	0,57	29,08	7,90	3,68	19,94
Sicilia	3.240	795	242	85.955	0,65	24,54	7,47	3,29	17,19
Sardegna	1.319	511	119	27.737	0,80	38,74	9,02	4,29	16,92
Italia	210.717	81.654	28.884	2.153.772	3,49	38,75	13,71	2,83	35,68

Fonte: ISS

Grafico 2 – Curva dei contagi da SARS-COVID19 in Italia, periodo 24/02/20 03/05/20

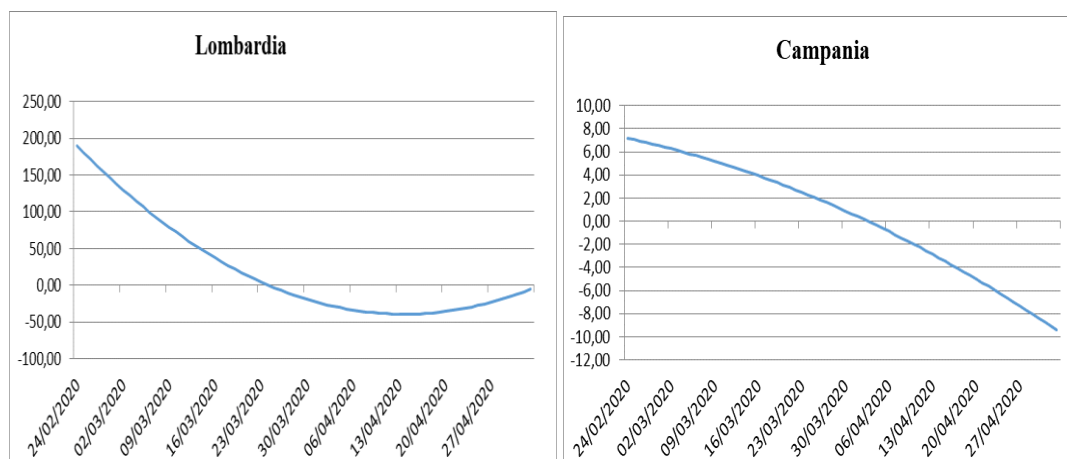


Fonte: nostra elaborazione su dati ISS

Tuttavia, come già precisato in precedenza la situazione differisce in ciascuna regione italiana. Questo non solo in relazione alla curva dei contagi, ma anche in riferimento alla velocità di crescita del contagio (intesa come derivata prima della funzione di variazione dei contagi) e alla accelerazione della stessa (intesa come derivata seconda della funzione di variazione dei contagi).

Il grafico 3 mostra le curve relative alla velocità della crescita del contagio in Lombardia e in Campania, due grandi regioni con andamenti molto differenti. Come si osserva, nella prima regione, dopo un rallentamento progressivo, nella seconda settimana di aprile c'è stata una notevole accelerata. Invece in Campania c'è stato un progressivo rallentamento in tutto il periodo considerato.

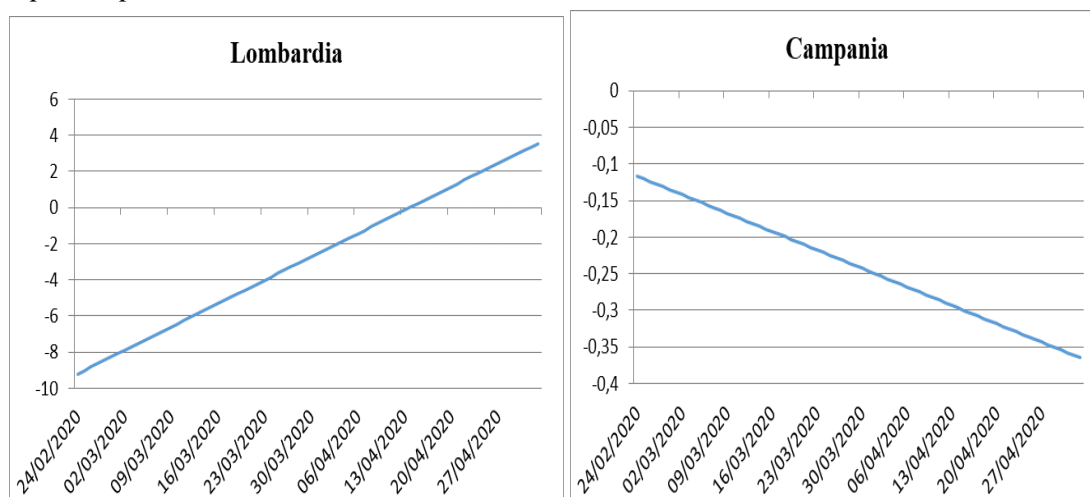
Grafico 3 – Velocità di crescita del contagio da SARS-COVID19 in Lombardia e in Campania, periodo 24/02/20 03/05/20



Fonte: nostra elaborazione su dati ISS

Il grafico 4 mostra invece la curva di accelerazione della crescita del contagio ed esprime molto chiaramente quanto sia stata rapida in Lombardia rispetto alla Campania.

Grafico 4 – Accelerazione di crescita del contagio da SARS-COVID19 in Lombardia e in Campania, periodo 24/02/20 03/05/20



Fonte: nostra elaborazione su dati ISS

3. Le determinanti del contagio

Da un punto di vista medico terapie, sintomi, immunità, origine sono ancora domande aperte.

I numeri ci dicono che esiste una fascia di popolazione più a rischio (maggiore di 60 anni), che gli assembramenti facilitano la diffusione del contagio, così come gli spostamenti. Sappiamo inoltre che il virus ha avuto origine in Cina. Inoltre, secondo studi recenti (Dominici et al., 2020), da un punto di vista statistico,

l'esposizione alle polveri sottili per lungo tempo comporta un aumento del rischio di contagio da COVID e del tasso di mortalità.

In assenza di altre informazioni certe, si è deciso di procedere all'analisi dei dati a disposizione che esprimono i fattori sopraelencati. L'analisi è stata effettuata per ciascuna regione allo scopo di valutare se tali fattori siano stati determinanti nelle differenze riscontrate tra la diffusione del virus in ciascuna regione. Sono stati quindi analizzati:

- Età della popolazione
- Densità abitativa
- Pendolarismo
- Commercio con l'estero
- Inquinamento

Tali fattori sono stati messi in relazione con i dati del contagio diffusi dall'Istituto superiore di sanità.

4. I fattori

4.1. Indicatori demografici

Alla luce dei primi risultati provvisori dell'Istat, il 2019 è un anno nel quale le tendenze demografiche risultano in linea con quelle mediamente espresse negli anni più recenti. Si registra un ulteriore aumento della speranza di vita alla nascita. A livello nazionale gli uomini sfiorano gli 81 anni, le donne gli 85,3. Per gli uni come per le altre l'incremento sul 2018 è pari a 0,1 decimi di anno, corrispondente a un mese di vita in più.

L'Italia è tra i Paesi a più elevata longevità, in particolare l'area del Nordest, dove si riscontrano condizioni di sopravvivenza assai favorevoli. Gli uomini residenti in questa ripartizione geografica possono infatti contare su una speranza di vita alla nascita pari a 81,6 anni, le donne pari a 85,9. Il Mezzogiorno, al contrario, gode di condizioni di sopravvivenza meno favorevoli, in virtù di una speranza di vita alla nascita di 80,2 anni tra gli uomini e di 84,5 tra le donne. Il primato regionale tra gli uomini compete alla Provincia di Trento (82,2 anni), seguono Umbria (81,9), Marche (81,8) e Provincia di Bolzano (81,8). Trento rappresenta l'area più favorevole per la sopravvivenza anche per le donne, grazie a una vita media di 86,6 anni, dato che costituisce peraltro il più alto livello di speranza di vita alla nascita mai toccato nella storia del Paese per una singola regione.

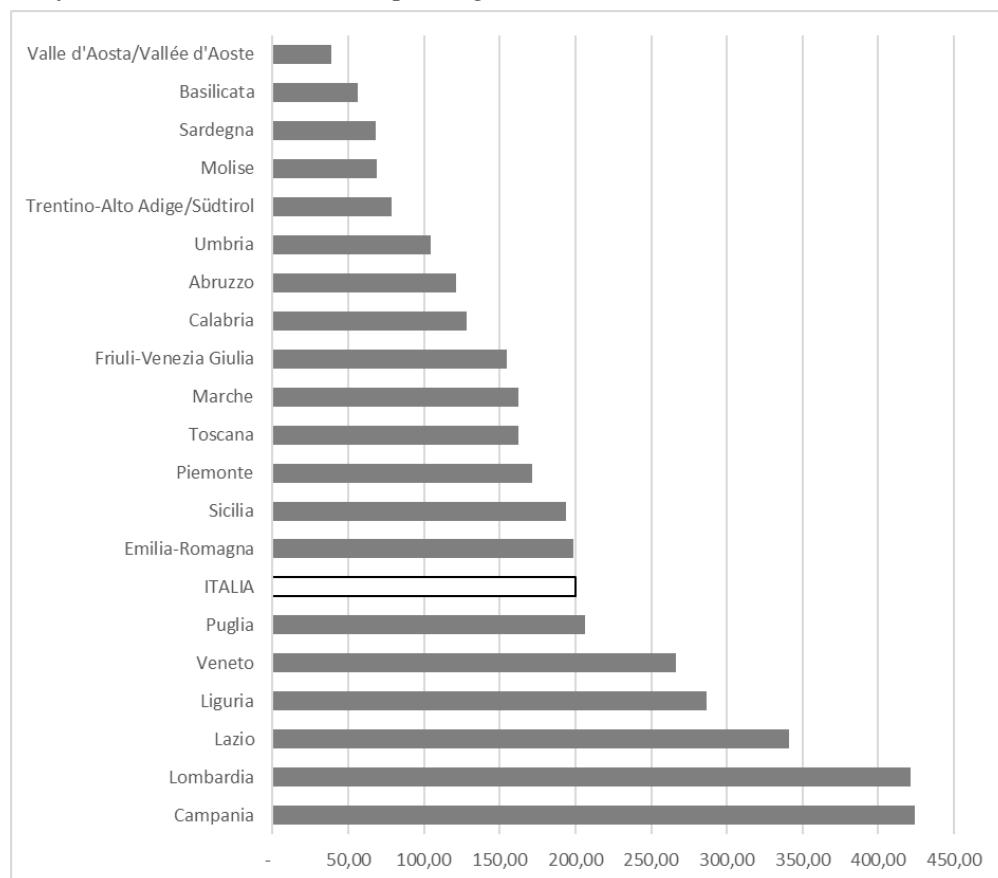
Come conseguenza della diminuzione delle nascite e dell'aumento della speranza di vita, la struttura per età della popolazione prosegue il suo lento ma costante scivolamento verso le età più anziane. Ancora nel 2020, il Mezzogiorno presenta una popolazione più giovane rispetto al Centro-nord. Ad esempio, la popolazione ultrasessantacinquenne incide per il 21,6% del totale, quando nel Nord e nel Centro risulta rispettivamente pari al 23,9% e al 23,8%. Così come, prendendo a riferimento un indicatore sintetico quale l'età media della popolazione, si può rilevare come per il Mezzogiorno (44,6 anni) risulti di oltre un anno e mezzo inferiore rispetto a quella del Centro-nord (46,2 anni).

Accanto all'indice di vecchiaia e all'età media, l'altro indicatore preso in considerazione nello studio è la densità abitativa. Questo perché, oltre ad essere la popolazione anziana quella a maggior rischio, è dimostrato che la vicinanza fisica è direttamente proporzionale al rischio di contagio. Si è partiti quindi dall'ipotesi che la densità abitativa influenza la vicinanza fisica nella misura in cui determina un maggiore affollamento di servizi pubblici di trasporto locale, attività ludico ricreative, ecc.

Considerando le macro aree, la zona più densamente popolata è il nord-ovest con 272,16 abitanti per kmq, seguito dal nord-est con 183,67 e dal centro con 199,72. Il Mezzogiorno è il meno popoloso con 166,65 abitanti per kmq.

Tuttavia, come mostra il grafico 5, la regione più densamente abitata si trova nel sud Italia ed è la Campania, seguita dalla Lombardia e dal Lazio. Superiori alla media nazionale anche i valori raggiunti nell'ordine da Liguria, Veneto e Puglia.

Grafico 5 – Densità abitativa per regione, anno 2019



Fonte: nostra elaborazione su dati Istat

4.2. Indicatori economici

L'economia di per sé non ha potuto comportare la diffusione del contagio, i contatti con il resto del mondo probabilmente sì. L'Italia è il nono paese esportatore al mondo con una quota di mercato del 2,9 per cento nel 2018 (ICE, 2019). È dal 2010 che le vendite di prodotti Made in Italy registrano tassi di crescita positivi e nei primi sette mesi del 2019 hanno registrato, se paragonate ai valori conseguiti durante lo stesso periodo degli anni precedenti, l'ennesimo record. Anche dal lato della bilancia commerciale si è conseguito un risultato estremamente positivo (+29,7 miliardi di euro), che è andato a migliorare il precedente primato risalente al 2016. I prodotti targati Made in Italy conquistano il Nord America, l'Africa sub-sahariana, l'Oceania e la buona parte del continente asiatico (MISE, 2019). In termini relativi il risultato più cospicuo si è conseguito nella regione centrale ed orientale dell'Asia (+6,4%), dove a svolgere un ruolo determinante sono risultati mercati come la Cina (+6,6%), con tassi tuttavia che dal 2015 si stanno attestando sotto la soglia del 7%, l'India (+6,8%) e i cosiddetti Asean53 (+5,2%).

Con riferimento alle finalità strategiche della produzione estera delle imprese, la decisione di delocalizzare l'attività produttiva ha come principale motivazione l'accesso a nuovi mercati nel caso di investimenti nei paesi dell'Area Euro (20,8% del totale delle imprese che hanno delocalizzato), in Nord e in Centro-Sud America (4,4 e 3,9% rispettivamente). Il contenimento del costo del lavoro prevale invece tra le imprese che hanno scelto come mercato di delocalizzazione gli altri paesi non euro (10,9%) ed extra Ue (9,2%), così come Cina (14,3%), India (6,4%) e gli altri paesi dell'area asiatica (4,9%).

Dal punto di vista territoriale, considerando il volume degli scambi commerciali con l'estero (Import ed export), la regione che su tutte incide di più sul bilancio nazionale è la Lombardia (29,42%). Seguono il Veneto con 12,61 ed Emilia Romagna con il 11,25%. Per le finalità dello studio è stato necessario focalizzare l'attenzione sugli scambi con l'Asia orientale, la cui principale quota parte è costituita dalla Cina. Questo perché si presuppone che il virus sia arrivato dal gigante asiatico. Anche in questo caso, ad avere i

maggiori rapporti commerciali con la Cina è la Lombardia che da sola copre il 38,19% del volume degli scambi commerciali con questa area del mondo. Segue sempre il Veneto con il 12,59 e l'Emilia Romagna con il 12,79%.

Considerando invece l'incidenza dei rapporti commerciali con l'Asia orientale sul totale delle bilance commerciali regionali, compare la "piccola" Campania che con il suo 12,19% è seconda alla Lombardia (13,08%).

L'altro aspetto considerato per valutare quanto i fattori economici abbiano potuto incidere sulla diffusione del Sars-Covid19 è la densità delle unità locali (numero di unità locali per Km²). Questo perché lo sviluppo delle imprese conduce ad un maggiore dinamismo sociale, e quindi maggiore contatto umano.

Il quadro generale mostra che i due terzi delle imprese (821 mila, pari al 79,5% del totale) sono microimprese (con 3-9 addetti in organico), 187 mila (pari al 18,2%) sono di piccole dimensioni (10-49 addetti), mentre le medie (con 50-249 addetti) e le grandi imprese (con 250 addetti e oltre) rappresentano il 2,3% delle imprese osservate (24 mila unità), di cui 3mila grandi. Più della metà delle imprese è attiva al Nord (il 29,2% nel Nord-ovest e il 23,4% nel Nord-est), il 21,4% al Centro e il 26,0% nel Mezzogiorno (Istat, 2020).

L'estensione del contesto competitivo ha connotazioni fortemente territoriali: opera in ambito comunale o regionale circa il 70 e il 60%, rispettivamente, delle imprese con almeno 10 addetti di Sardegna e Sicilia, e circa la metà di quelle con sede in Calabria, Puglia e Liguria. Al contrario, un'operatività riferita a mercati sovranazionali caratterizza soprattutto le unità produttive di molte regioni settentrionali, in particolare quelle di Friuli-Venezia Giulia, Lombardia, Veneto, Trentino-Alto Adige e Piemonte, in misura sempre compresa tra il 38 e il 40%.

Nella classifica nazionale, la regione ad avere la maggiore concentrazione di unità locali è la Lombardia con 37,33 unità per Km², seguita da Campania (26,95) e Lazio (26,94), che superano altre regioni del nord come Liguria (25,08) e Veneto (23,56). Le regioni più piccole da questo punto di vista risultano essere Basilicata (3,77), Valle d'Aosta (3,79) e Sardegna (4,68).

4.3. Indicatori di pendolarismo

Durante la così detta fase 1 sono stati vietati gli spostamenti tra i comuni, ciò al fine di evitare il contagio tra le diverse aree geografiche. Si presuppone quindi che il fenomeno del pendolarismo abbia contribuito alla diffusione del virus dal primo focolaio individuato nel mese di febbraio a tutto il territorio nazionale. Nell'analisi è stato quindi analizzato anche questo elemento.

Il fenomeno di coloro che si spostano ogni giorno e percorrono lunghe distanze per motivi di studio o di lavoro interessa circa la metà della popolazione italiana, soprattutto quella residente nel Nord.

Secondo il rapporto Censis-Michelin 2018, un pendolare in Italia fa circa 28,8 km al giorno in auto e impiega 57,5 minuti.

Ad utilizzare il servizio ferroviario regionale sono invece ogni giorno 2milioni e 874mila persone (Legambiente, 2018), di cui 1,425 milioni che utilizzano i convogli di Trenitalia (in deciso aumento rispetto agli 1,377 milioni dello scorso rilevamento) e 1,449 milioni quelli degli altri 20 concessionari (tra cui 750mila Trenord in Lombardia, 205mila CTI in Emilia-Romagna, 190mila per Atac nel Lazio, 145mila per Eav in Campania). Per i treni regionali dal 2010 ad oggi abbiamo un aumento del 6,8% dei passeggeri, mentre l'offerta è aumentata dello 0,2% con un notevole sovraffollamento degli stessi.

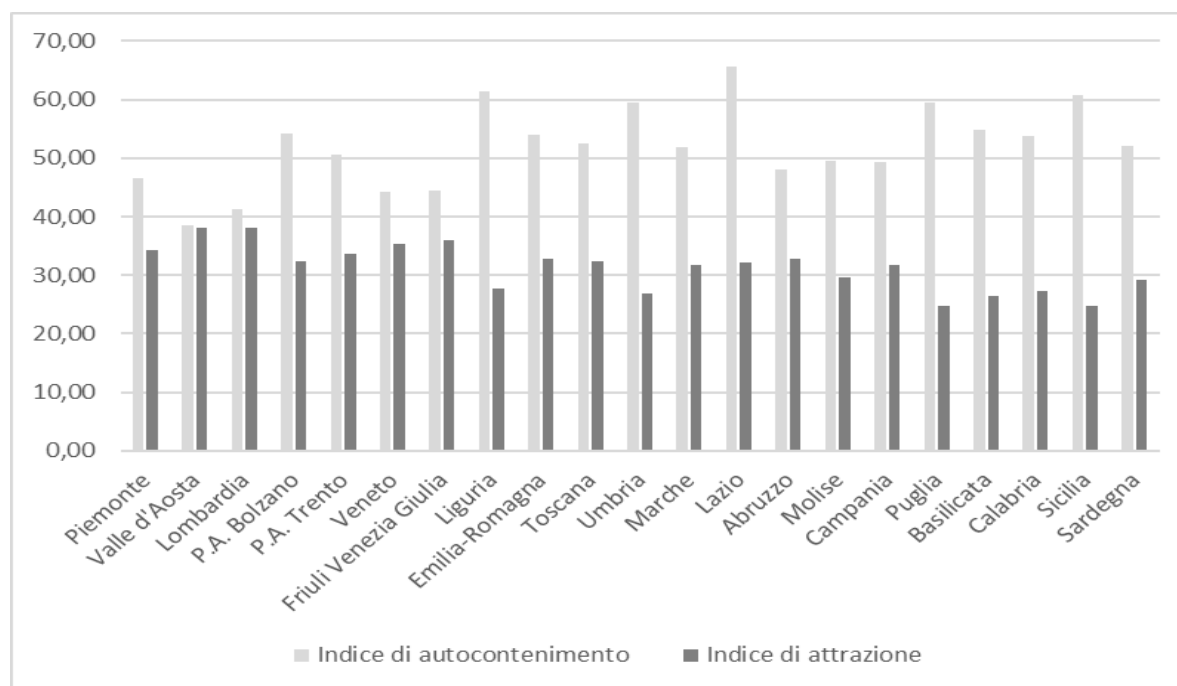
L'Istat misura l'indice di attrazione, dato dai flussi in entrata nell'area in rapporto al totale dei flussi di mobilità, e l'indice di autocontenimento, dato dai flussi interni all'area in rapporto alla somma di flussi interni e flussi in uscita.

Il grafico 6 mostra come le regioni più attrattive siano la Lombardia (38,15%), la Valle d'Aosta (38,10%), il Friuli Venezia Giulia (35,9%), il Veneto (35,37%) e il Piemonte (34,27%).

Le meno attrattive sono Puglia e Basilicata con il 24%.

Le regioni per cui influiscono maggiormente i flussi interni sul totale degli spostamenti sono invece il Lazio (65,58%), la Liguria (61,31%), la Sicilia (60,77%), l'Umbria e la Puglia (entrambe con 59,6%)

Grafico 6– Indice di autocontenimento e indice di attrazione delle regioni italiane, anno 2015



Fonte: nostra elaborazione su dati Istat

4.4. Indicatori ambientali: qualità dell'aria

Circa il 91% della popolazione mondiale vive in luoghi in cui i livelli di qualità dell'aria superano i limiti stabiliti dall'Organizzazione Mondiale della Salute (OMS) (WHO, 2018).

L'esposizione agli inquinanti atmosferici può influire sulla salute dell'uomo in vari modi, con conseguente aumento della mortalità e morbidità. L'inquinamento atmosferico uccide ogni anno circa sette milioni di persone in tutto il mondo, in gran parte a causa di un aumento della mortalità per ictus, malattie cardiache, malattie polmonari ostruttive croniche, cancro ai polmoni e infezioni respiratorie acute. I dati dell'OMS mostrano che 9 persone su 10 respirano aria contenente alti livelli di inquinanti. Dallo smog sospeso sulle città al fumo all'interno della casa, l'inquinamento atmosferico rappresenta una grave minaccia per la salute.

Le principali fonti di inquinamento all'aperto includono veicoli, produzione di energia, sistemi di riscaldamento degli edifici, agricoltura, incenerimento dei rifiuti e industria.

Secondo un recente studio condotto dai ricercatori dell'università di Harvard molte delle condizioni preesistenti che aumentano il rischio di mortalità nelle persone con COVID-19 sono le stesse malattie che sono interessate dall'esposizione a lungo termine all'inquinamento atmosferico.

Lo studio ha confrontato i tassi di mortalità causati dal COVID-19 nei paesi con caratteristiche socio-economiche simili giungendo alla conclusione che piccoli aumenti dell'esposizione a lungo termine alle polveri sottili porta ad un grande aumento del tasso di mortalità COVID-19.

In Italia nel 2018 sono stati superati i limiti giornalieri previsti per le polveri sottili o per l'ozono (35 giorni per il Pm10 e 25 per l'ozono) in ben 55 capoluoghi di provincia (Legambiente, 2019). La val Padana è l'area più inquinata d'Europa e l'aspetto orografico del suolo, la carenza di ventilazione, contribuiscono alla stagnazione delle polveri sottili.

In 24 dei 55 capoluoghi il limite è stato superato per entrambi i parametri, con la conseguenza diretta, per i cittadini, di aver dovuto respirare aria inquinata per circa 4 mesi nell'anno.

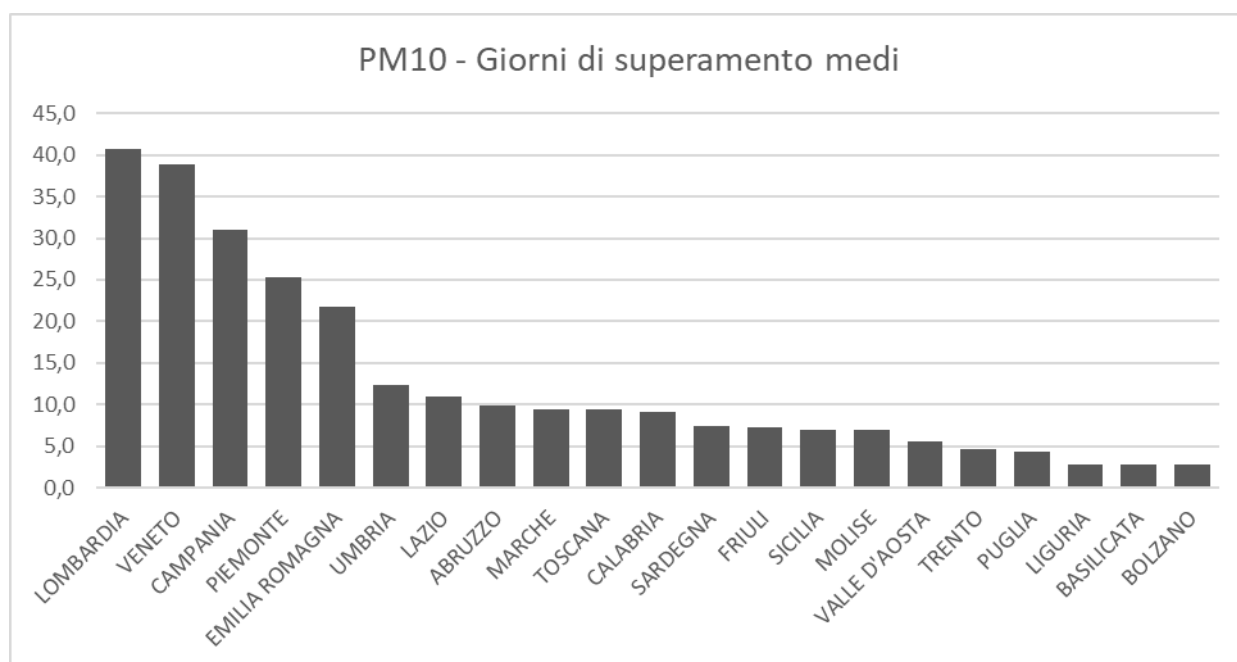
Capofila delle città che nel 2018 hanno registrato il maggior numero di giornate fuorilegge è Brescia con 150 giorni (47 per il Pm10 e 103 per l'ozono), seguita da Lodi con 149 (78 per il Pm10 e 71 per l'ozono), Monza (140), Venezia (139), Alessandria (136), Milano (135), Torino (134), Padova (130), Bergamo e Cremona (127) e Rovigo (121). Tutte le città capoluogo di provincia dell'area padana (ad eccezione di Cuneo, Novara, Verbania e Belluno) hanno superato almeno uno dei due limiti

Il termine PM10 identifica le particelle di diametro aerodinamico inferiore o uguale a 10 µm (Ispra, 2019). Queste sono caratterizzate da lunghi tempi di permanenza in atmosfera e possono, quindi, essere trasportate anche a grande distanza dal punto di emissione, hanno una natura chimica particolarmente complessa e variabile, sono in grado di penetrare nell'albero respiratorio umano, con effetti negativi sulla salute.

In alcune delle città elencate l'inquinamento è causato da alcune peculiarità. Ad esempio a Brescia, nelle vicinanze della centralina Villaggio Sereno c'è anche l'autostrada e l'area, in generale, ricade in una zona industriale importante o Lodi e, in generale, molte città lombarde ed emiliane, dove il contributo dell'agricoltura e degli allevamenti intensivi sicuramente accresce quello che è il contributo delle polveri secondarie.

Considerando il numero di giorni di superamento del limite massimo consentito in ciascuna stazione di rilevazione, è possibile calcolare i giorni di superamento medi regionali nel 2018 (Grafico 7). Anche in questo caso si conferma il triste primato della Lombardia con 40 giorni nel 2018, seguita da Veneto, Campania, Piemonte ed Emilia Romagna.

Grafico 6: PM 10 – Giorni di superamento Medi del valore limite giornaliero per la protezione della salute (2018)



Fonte: elaborazione su dati Ispra, 2019

Per quanto riguarda le emissioni di gas serra, i trasporti risultano, insieme alle industrie energetiche, i settori maggiormente responsabili delle emissioni. In particolare gli autoveicoli incidono sull'aumento del particolato (PM10), Ozono (O3), Biossido di Azoto (NO2), Benzene (C6H6) e Biossido di Zolfo (SO2).

Negli ultimi anni, grazie ai miglioramenti tecnologici apportati ai veicoli, si è assistito ad un calo delle emissioni di inquinanti atmosferici prodotte dal trasporto stradale. Ma non è ancora sufficiente: la qualità dell'aria nelle grandi aree urbane non rispetta ancora i valori limite stabiliti dalla normativa europea. In Italia

ci sono ancora 38 milioni di auto private (corrispondenti al 17% dell'intero parco circolante continentale) che soddisfano complessivamente il 65,3% degli spostamenti.

Osservando il parco auto degli italiani, si riscontra che la percentuale di auto circolanti al di sotto dello standard Euro 4 è ancora elevato: in particolare nelle regioni del mezzogiorno quasi la metà delle auto è della classe compresa tra Euro 0 e Euro 3. In particolare Campania, Basilicata, Calabria e Sicilia hanno valori superiori al 50% di presenza di auto altamente inquinanti.

5. La scelta del modello: l'analisi fattoriale

Il numero considerevole di variabili, che concorrono a determinare la diversa diffusione dei contagi nelle regioni italiane ha condotto alla scelta dell'utilizzo dell'analisi fattoriale allo scopo di ridurre la multidimensionalità del fenomeno e facilitare l'interpretazione dei contributi di ciascuno degli aspetti finora analizzati.

L'analisi fattoriale (Fabbris, 2003) consente, a partire dalla matrice di correlazione o di varianza e covarianza, di determinare un certo numero di variabili "latenti" (fattori non direttamente misurabili nella realtà) più ristretto e riassuntivo rispetto al numero di variabili di partenza garantendo la minima perdita di informazione. In altri termini, l'obiettivo è quello di ridurre la dimensionalità dell'insieme dei dati eliminando la ridondanza di informazioni, risultato di p variabili altamente correlate, e sostituendo a queste ultime un minor numero h ($h < p$) di nuove variabili tra loro non correlate e legate linearmente alle variabili di partenza. Le nuove variabili oltre ad essere non correlate sono ordinate rispetto alla percentuale di variabilità presente nei dati originali.

La prima componente principale (o il primo fattore) è la combinazione lineare delle p variabili di partenza avente massima varianza; la seconda componente principale è la combinazione lineare delle p variabili con varianza immediatamente inferiore, soggetta al vincolo di essere ortogonale alla componente precedente, e così via.

Per la scelta del numero di fattori da estrarre è possibile utilizzare tre criteri:

- il criterio di Kaiser, in base al quale si considerano tutti i fattori il cui autovalore sia superiore o uguale a 1;
- il criterio di Cattell, o analisi dello scree plot, che stabilisce il numero di fattori attraverso un grafico, lo scree plot, che rappresenta gli autovalori in ordinata e le dimensioni, progressivamente dalla prima, in ascissa andando a ricercare il punto di flesso del grafico;
- si fissa una soglia minima di variabilità spiegata o si definisce a priori un numero massimo di fattori da estrarre.

Il risultato può venire poi sottoposto a rotazione con vari metodi. I metodi che usano rotazioni ortogonali preservano l'indipendenza dei fattori, mentre quelli che usano rotazioni oblique rilasciano il vincolo di indipendenza dei fattori per migliorarne l'interpretabilità.

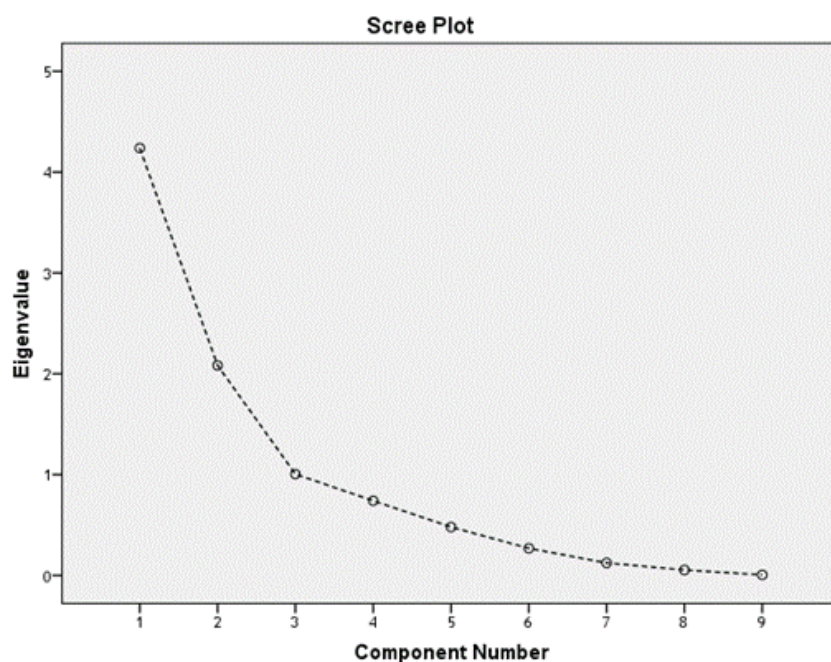
L'interpretazione di ogni componente principale, con la conseguente scelta dell'etichetta verbale più adeguata per la stessa, è l'ultima scelta che spetta al ricercatore. Per questa scelta non esistono regole statistiche ma entrano in gioco capacità del ricercatore, esperienza e sensibilità: una serie di elementi non formalizzabili statisticamente.

6. Risultati

L'analisi della matrice di correlazione ha consentito di valutare l'ammontare di varianza in comune fra le variabili: le variabili più correlate fra di loro hanno poi concorso a spiegare lo stesso fattore.

Per la scelta dei fattori è stato utilizzato il criterio di Kaiser estraendo i fattori con autovalore maggiore di 1 e mantenere un livello di variabilità spiegata elevato. Anche lo Scree plot (Figura 1) suggerisce di orientarsi sulla scelta dei primi tre fattori che insieme spiegano l'81,4% della varianza.

Figura 1 - Scree Plot delle componenti principali



Fonte: nostra elaborazione

Sebbene le quote di varianza riprodotta dalla terza componente (11,164) siano inferiori rispetto a quelle delle prime due (47,093 e 23,137) si è deciso di considerare anche il terzo fattore per dimostrare in che modo aspetti che erano ritenuti a inizio analisi molto influenti si sono rivelati meno esplicativi di altri.

Dall'analisi delle comunalità risulta l'assenza di variabili penalizzate, per le quali cioè la quota di varianza spiegata dai fattori estratti è bassa.

Una volta definiti i fattori, è stato necessario effettuare una rotazione degli assi componenziali per concentrare le saturazioni elevate di ciascuna variabile su un solo fattore. Con questa operazione, si è ridotto il valore dei pesi componenziali che nella fase precedente erano più piccoli e si è incrementato il peso di quelli già alti. Effettuata la rotazione di tipo Varimax, che permette che le componenti ruotate restino ortogonali fra di loro, è stata ottenuta la tavola con i pesi componenziali iniziali e ruotati. La varianza spiegata cumulativa resta naturalmente invariata (tavola 3).

Tavola 3 – Matrice ruotata delle componenti

	Componenti		
	1	2	3
I_autocontenimento2015	-0,159	0,85	-0,167
I_attrazione2015	0,324	-0,901	-0,054
I_vecchiaia2019	-0,117	0,056	0,957
Dens_ab_1_1_19	0,924	0,198	-0,172
Inc_Italia_Mondo	0,81	-0,318	-0,049
Inc_reg_AsiaOr	0,726	-0,274	0,135
GG_superma_PM10	0,852	-0,285	-0,089
Autovettura_inq2018	0,093	0,797	0,188
Dens_UL2015	0,944	0,053	-0,091

Fonte: nostra elaborazione

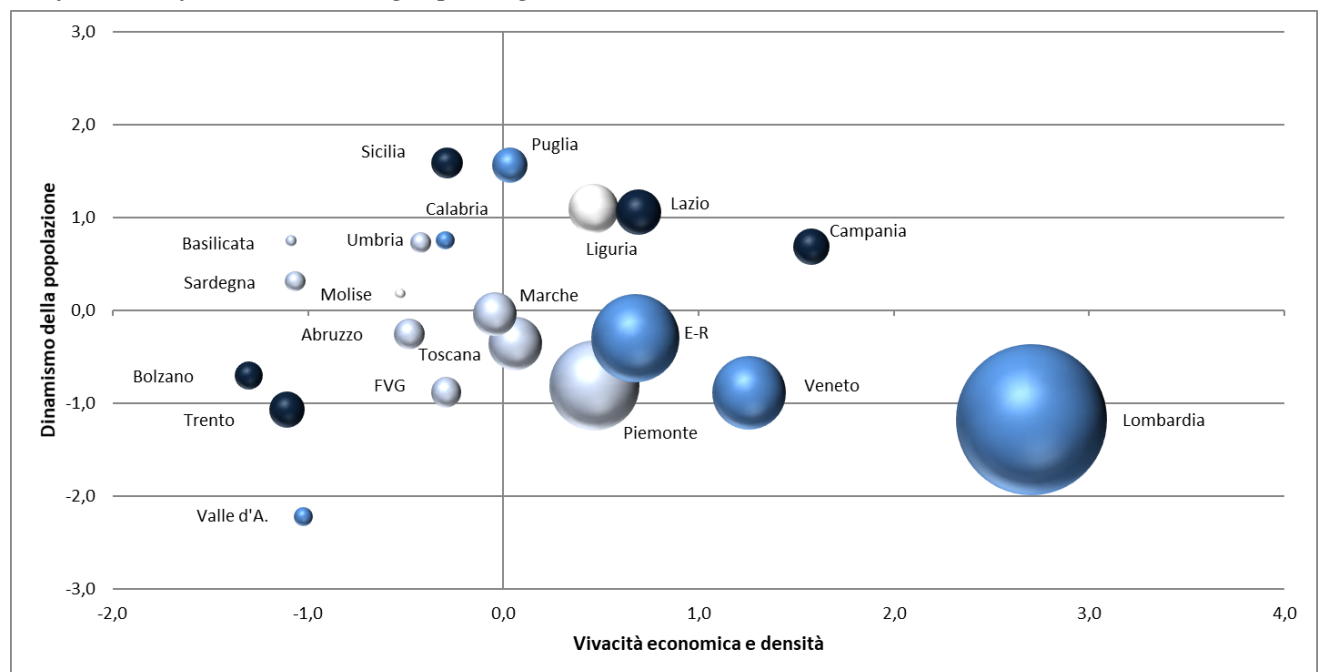
Ricordando che ogni fattore è stato ottenuto come combinazione lineare di tutte le variabili e quindi ogni variabile ha dato un contributo alla formazione di ciascun fattore c'è da precisare che alcune variabili hanno contribuito in modo più significativo e altre in modo molto marginale alla costruzione di ciascun fattore.

Come indicato nella tavola 3, sulla costruzione del primo fattore hanno contribuito la densità abitativa (Dens_ab_1_1_19), incidenza percentuale degli scambi commerciali con il mondo (Inc_Italia_Mondo), e incidenza regionale degli scambi con l'Asia orientale (Inc_reg_AsiaOr), i giorni di superamento del PM10 (GG_superma_PM10) e la densità di unità locali (Dens_UL2015); sul secondo fattore l'autocontenimento (I_autocontenimento2015), l'indice di attrazione (I_attrazione2015) negativamente e la presenza di autovetture inquinanti (Autovettura_inq2018); il terzo fattore è determinato solo dall'indice di vecchiaia (I_vecchiaia2019).

Pertanto il primo asse è stato definito “vivacità economica e densità abitativa” poiché riassume tutte le variabili economiche considerate, inclusa la conseguenza dell'industrializzazione, ossia la presenza di polveri sottili, oltre che la densità abitativa; il secondo asse è stato definito “dinamismo della popolazione” in quanto è determinato dal pendolarismo della popolazione e dal fattore inquinante dato dalle autovetture di vecchia produzione. Il terzo fattore è ovviamente stato definito “vecchiaia” poiché determinato solo dall'omonimo indicatore. Occorre da subito sottolineare che sebbene l'età sia stata identificata da principio un fattore importante nel contagio, essa non ha avuto un ruolo determinante nell'analisi fattoriale.

I tre fattori sono stati rappresentati nel grafico 7: sull'asse delle ascisse è rappresentato il primo fattore, sulle y il secondo; le bolle rappresentano le regioni; il colore delle bolle è determinato dal terzo fattore, più sono scure più è basso l'indice di vecchiaia; l'ampiezza delle bolle è data dal numero dei contagi da Covid19 al 3 maggio 2020.

Grafico 7 – I fattori del contagio per regione



Fonte: nostra elaborazione

Dal grafico è chiaro che le regioni con il maggior numero di contagi si collocano su valori elevati del primo fattore: Lombardia, Veneto, Emilia Romagna e Piemonte. La Lombardia, che si posiziona nella parte più estrema dell'asse, è quella le cui caratteristiche sono esclusivamente dovute alla “vivacità economica e densità”. È importante sottolineare che l'unica regione tra quelle indicate ad avere un alto indice di vecchiaia è il Piemonte. Le regioni con il minor numero di contagi si posizionano invece nel terzo e quarto quadrante. Ciò vuol dire che dipendono negativamente dal primo fattore: non hanno quindi caratteristiche che rappresentano una spiccata economia internazionale e produttiva.

Nel primo quadrante abbiamo anche Marche, Toscana e Puglia, che però si posizionano quasi esattamente sull'asse "Dinamismo della popolazione", e Liguria e Lazio, anch'esse molto vicine al secondo fattore.

È utile sottolineare che il secondo fattore si caratterizza positivamente per l'indicatore di autocontenimento e per la presenza di vetture inquinanti, mentre negativamente per l'indice di attrazione. In sintesi le regioni che si posizionano sull'asse delle y nel grafico 7 sono regioni in cui prevale il pendolarismo interno e non quello esterno in entrata.

L'unica regione a presentare elevati livelli di vivacità economica e densità senza però avere un elevato numero di contagi è la Campania. Essa è però fortemente caratterizzata anche dal fattore che esprime il dinamismo della popolazione, oltre ad essere la regione più giovane d'Italia.

7. Conclusioni

La diffusione geografica dell'epidemia di Covid-19 è stata eterogenea tra le regioni italiane. Il primo rapporto sulla mortalità della popolazione residente a cura di Iss e Istat (Istat, 2020) parla delle tre Italie: nelle Regioni del Sud e nelle isole, la diffusione delle infezioni è stata molto contenuta, in quelle del Centro, è stata mediamente più elevata rispetto al Mezzogiorno mentre in quelle del Nord la circolazione del virus è stata molto elevata.

Dall'inizio del 2020 si osserva un aumento dei decessi: se si assume come riferimento il periodo che va dal primo decesso Covid-19 riportato al Sistema di Sorveglianza integrata (20 febbraio) fino al 31 marzo, i decessi passano da 65.592 (media periodo 2015-2019) a 90.946, nel 2020.

Il maggior costo in termini di vite è stato pagato dalle aree ad "alto" contagio, ossia le province con valori superiori ai 450 casi per 100mila residenti, un totale di 36 province in gran parte dell'Italia settentrionale.

I risultati del lavoro realizzato mostrano che, da un punto di vista statistico, le differenze della diffusione del contagio si evincono anche dalla velocità di crescita e dall'accelerazione del contagio. Come è stato illustrato nei precedenti paragrafi l'epidemia si è diffusa in maniera più o meno rapida nelle diverse regioni. L'obiettivo della ricerca è stato quello di individuare quali sono state le determinanti di tale differenza. Come mostra il grafico 7 i fattori determinanti sono riconducibili alla vivacità e densità economica e alla densità della popolazione. Sembrerebbe che il dinamismo della popolazione influisca in maniera meno incisiva sulla diffusione del contagio.

Nel determinare la vivacità economica si ricorda che sono stati presi a riferimento aspetti strettamente strutturali, la presenza di unità produttive, e i rapporti commerciali con l'estero e in particolare con l'Asia orientale. Si è considerato anche l'effetto dell'industrializzazione, ossia la presenza di polveri sottili, che studi precedentemente citati hanno dimostrato essere un fattore di indebolimento delle vie respiratorie. Relativamente alla vicinanza fisica tra le persone, che dovrebbe determinare l'aumento del così detto indice R_0 , lo studio dimostra che la densità abitativa abbia inciso sulla espansione del virus. Non ha invece inciso allo stesso modo il pendolarismo. La presenza di fattori inquinanti dovuti dalla presenza di autoveicoli di vecchia produzione, pare che non sia stato determinante come la presenza di polveri sottili nell'atmosfera. Infine, sebbene il virus Sars-covid19 colpisca maggiormente la popolazione anziana, non sembra l'età aver inciso sull'allargamento del contagio.

In definitiva il primo aspetto su cui sarebbe necessario intervenire al fine di ridurre rischi simili nel futuro è l'adozione di misure strutturali capaci di abbattere drasticamente le concentrazioni di inquinamento presenti e, di conseguenza, di riportare l'aria a livelli qualitativamente accettabili.

8. Bibliografia

Xiao Wu, Rachel C Nethery, M Benjamin Sabath, Danielle Braun, Francesca Dominici (2020). *Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study*. MedRxiv, USA.

- Legambiente (2020). *Rapporto pendolaria. La situazione e gli scenari del trasporto pendolare ferroviario in Italia*. Dossier di Legambiente – 8 febbraio 2020.
- Istat (2020). *Indicatori demografici anno 2019*. Statistica report - 11 febbraio 2020. Roma
- Istat (2020). *Censimento permanente delle imprese 2019: i primi risultati*. Comunicato stampa – 7 febbraio 2020. Roma
- ISS (2020). *Epidemia Covid-19*. Aggiornamento nazionale 30 giugno 2020 – ore 11:00. Roma – 3 luglio 2020.
- Istat, ISS (2020). *Impatto dell'epidemia Covid-19 sulla mortalità totale della popolazione residente. Primo trimestre 2020*. Dossier Istat ISS – 4 maggio 2020.
- Legambiente (2019). *Mal'Aria di città 2019, la mobilità urbana al centro del risanamento*. Dossier di Legambiente - 22 gennaio 2019
- Ispra (2019). *Annuario dei dati ambientali 2019*. ISPRA, Stato dell'ambiente 89/2020. Roma.
- World Health Organization (2018). *Ambient air pollution: a global assessment of exposure and burden disease*. WHO press, Ginevra.
- Istat, ICE (2019). *Annuario Commercio estero e attività internazionali delle imprese*. Edizione 2019. Roma.
- Ministero degli affari esteri (2019). *Scambi con l'estero - Anno 2019*. Bollettino n.3 Note di Aggiornamento Anno 26° - n. 3 – 2019
- Censis (2018). *Rapporto Censis-Michelin sulla mobilità degli italiani 2018*. Fondazione Censis - 18 ottobre 2018. Roma.
- Fabbris L. (2003) - *Statistica multivariata. Analisi esplorativa dei dati*. McGraw-Hill Ed.

ABSTRACT

The level of diffusion of SARS-COVID19 has not been homogeneous throughout the Italian territory. If it is true that the greatest concentration in Lombardy is due to the presence of the first virus site, there are open questions on the reasons that led to this strong interregional difference.

The answers are important to aid the regions to put in place choices that can reduce the potential diffusion of new biological risks.

To identify the determinants of the infection, it is necessary to refer to the objective data collected in Italy during the period of spread of the pandemic. The study analyzes, at regional level, key indicators such as the age of the population, population density, the daily population movement, the international trade and relationships with foreign countries, pollution and vehicle typology. The results achieved with the use of factorial analysis show the contribute of these phenomena to the virus diffusion and why some regions are more exposed than others.