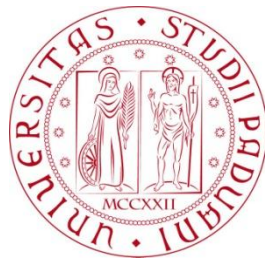


Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Triennale in
Statistica per l'Economia e l'Impresa



RELAZIONE FINALE
**Studio sulla relazione tra spesa farmaceutica e stile di
vita in Italia**

Relatore Prof. Ruli Erlis
Dipartimento di Scienze Statistiche

Laureando: Giovanni Zerbo
Matricola N 1235636

Anno Accademico 2021/2022

Indice

1	Introduzione	4
2	Spesa farmaceutica in Italia	5
2.1	La spesa sanitaria e farmaceutica.	5
2.2	Raccolta dei dati	6
2.3	Composizione della spesa farmaceutica	7
2.4	I farmaci.	9
2.4.1	Classi di farmaci	9
2.4.2	Tipologia di farmaci	10
2.4.3	Farmaci più venduti	13
2.4.4	Concentrazione della spesa	14
2.5	Disuguaglianze territoriali.	16
2.5.1	Vitamina D	18
2.5.2	Inibitori della pompa acida.	21
2.5.3	Ormoni Tiroidei	23
3	Stile di vita	25
3.1	Raccolta dei dati	25
3.2	Obesità.	26
3.3	Attività fisica.	27
3.4	Fumo	29
4	Spesa convenzionata e stile di vita	31
4.1	Correlazioni	31
4.2	Analisi delle Correlazioni Canoniche	33
4.3	Clustering.	35

5 Conclusioni.....	40
6 Bibliografia.....	41

1 Introduzione

”La spesa sanitaria non è un costo, è un investimento nella riduzione della povertà, nell’occupazione, nella produttività, nella crescita economica inclusiva e in società più sane, più sicure e più giuste”.

Queste le parole usate nel 2019 da Tedros Adhanom Ghebreyesus, direttore generale dell’OMS¹ per commentare la relazione dell’organizzazione sulla spesa sanitaria globale. La salute, quindi, oltre a essere riconosciuta come un diritto fondamentale dell’uomo², viene considerata un aspetto sul quale è conveniente investire, sia a livello di spesa pubblica che a livello individuale con lo stile di vita quotidiano.

L’obiettivo di questa tesi è fornire, quindi, una panoramica sulla distribuzione della spesa farmaceutica in Italia e sulle abitudini degli italiani che più influenzano la salute. In seguito si cercherà anche di capire come questi due aspetti variano da regione a regione, e quali relazioni ci siano tra di esse.

¹Organizzazione Mondiale della Sanità, istituto specializzato dell’ONU per la salute

²Art. 25 Dichiarazione Universale dei Diritti Umani

2 Spesa farmaceutica in Italia

2.1 La spesa sanitaria e farmaceutica

Nel 2019 la spesa sanitaria pro capite in Italia (pubblica e privata) a parità di potere d'acquisto, si attestava a 2473 euro. Un valore al di sotto della media OCSE (2572€) e di molto inferiore a quello di grandi paesi europei come Francia e Germania (rispettivamente 3644€ e 4504€). Le cause di ciò possono essere ricercate in vari ambiti, uno dei quali è il fatto che l'Italia è stato uno dei pochi paesi ad aver effettuato tagli alla sanità durante la crisi finanziaria del 2008.

Così come la spesa sanitaria, anche la spesa farmaceutica pubblica pro capite è inferiore (circa il 20% in meno) agli altri principali paesi europei. Essa ha costituito, nel 2021, il 17,4% della spesa sanitaria pubblica, in aumento rispetto agli anni precedenti. In questo caso, però, la differenza con gli altri paesi Europei è dovuta anche al fatto che nel nostro paese i medicinali hanno un prezzo minore, e tempi di accesso molto più lunghi (fino a 24 mesi) a causa della doppia natura del sistema sanitario (nazionale e regionale).

L'ente preposto all'attività regolatoria dei farmaci in Italia

è l'AIFA (Agenzia italiana del farmaco), che elabora annualmente un rapporto dove vengono registrate le spese farmaceutiche. Ogni spesa è caratterizzata dalla tipologia di farmaco, la regione in cui è stato venduto, il tipo di spesa effettuato e il numero di confezioni vendute. Per le analisi mostrate in seguito si farà uso di questi database, presi negli anni da 2016 a 2020.

2.2 Raccolta dei dati

I database elaborati da AIFA sono consultabili a libero accesso sul corrispondente sito governativo. Essi sono stati stilati annualmente, dal 2016 al 2020, nell'ambito dei processi di monitoraggio della spesa, tramite l'utilizzo dei rapporti **OsMed** (Osservatorio Nazionale sull'Impiego dei Medicinali). I rapporti OsMed riportano il flusso informativo delle prestazioni farmaceutiche erogate attraverso le farmacie, pubbliche e private, convenzionate con il SSN³. Sono raccolti mensilmente su base provinciale dalle associazioni di FederFarma (per quanto concerne le farmacie private) e AssoFarm (farmacie pubbliche), che aggregano conseguentemente i dati a livello regionale.

³Sistema Sanitario Nazionale

2.3 Composizione della spesa farmaceutica

I dati raccolti sulla spesa farmaceutica nel rapporto OsMed incorporano due tipi distinti di spesa: la **spesa convenzionata lorda** e la **spesa rilevata dal flusso della tracciabilità del farmaco**. Il principale oggetto di studio di questa tesi sarà la spesa convenzionata lorda, la quale si identifica con la spesa per Farmaci SSN erogati dalle farmacie nel normale regime convenzionale. Questa spesa, stando ai dati raccolti negli ultimi anni, è per la maggior parte (86-87%) pubblica, mentre il restante 13-14% è una quota di compartecipazione a carico del cittadino⁴. La spesa convenzionata lorda (che oscilla tra i 10 e gli 11 miliardi annui) rappresenta circa un terzo della spesa farmaceutica totale (pubblica e privata) in Italia (Figura 1), la quale nel 2019 raggiungeva il valore di 30,8 miliardi di euro.

⁴Ticket regionali e differenza tra il prezzo del medicinale a brevetto scaduto e il corrispondente prezzo di riferimento

Spesa Farmaceutica in Italia, 2019

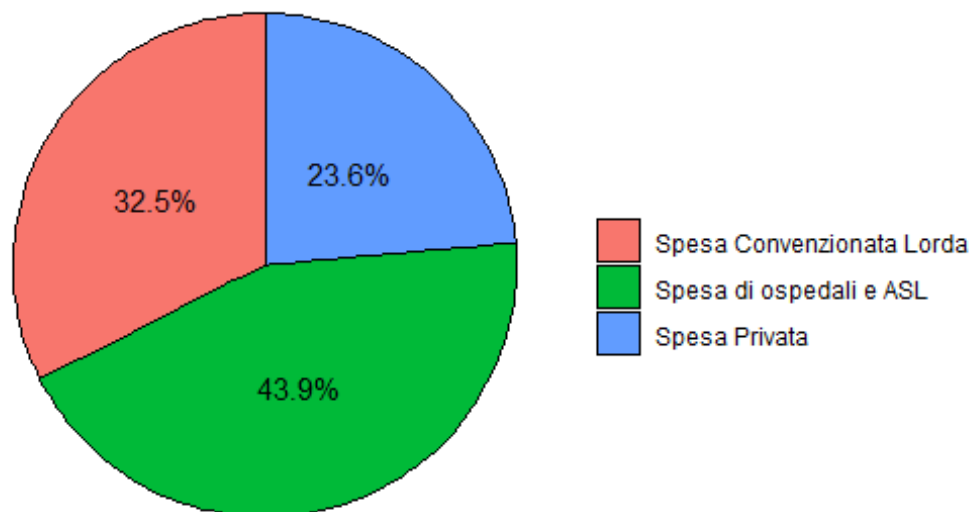


Figura 1

La spesa rilevata dal flusso della tracciabilità del farmaco, invece, include la spesa farmaceutica ospedaliera, la distribuzione diretta (spesa per farmaci acquistati dalle ASL⁵ e consegnati direttamente ai cittadini dalle ASL stesse) e la distribuzione per conto (spesa per farmaci acquistati dalle ASL e distribuiti dalle farmacie). Rimane esclusa dall'equazione la spesa privata non convenzionata, di cui non sono disponibili dati dettagliati, e che negli ultimo decennio è in leggero ma costante aumento e si aggira sui 7 miliardi di euro annui⁶.

⁵Azienda Sanitaria Locale

⁶Fonte: ISS (Istituto Superiore di Sanità)

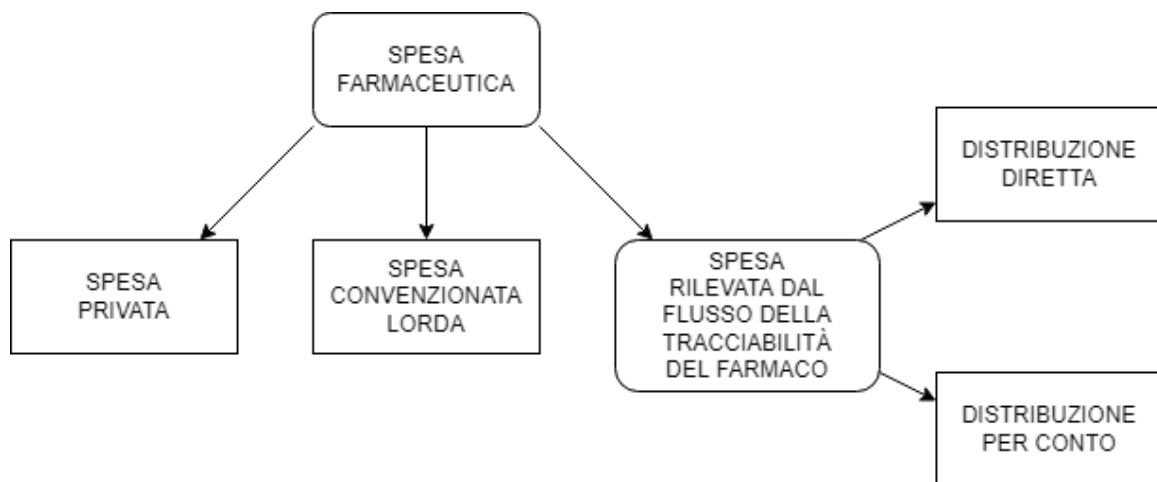


Figura 2

2.4 I farmaci

2.4.1 Classi di farmaci

Il Sistema Sanitario Nazionale italiano suddivide i farmaci in tre macrocategorie (classi), a seconda della loro modalità di vendita: classe A, classe C e classe H (vedi Figura 3).

Classe A	Rimborsati dal SSN (medicinali essenziali e farmaci per malattie croniche)
	<ul style="list-style-type: none"> • totalmente rimborsabili • parzialmente rimborsabili per i quali il cittadino paga una differenza (ticket) • rimborsabili, totalmente o parzialmente, se con nota Aifa (altrimenti sono classe C)
Classe C	Farmaci non rimborsabili dal SSN e pagati interamente dal cittadino ECCETTO nei casi che prevedono il codice di esenzione
Classe H	Farmaci a carico del SSN solo se dispensati da strutture SSN come ospedali e poliambulatori, venduti a pagamento in farmacia.

Figura 3
Fonte: Ordine dei Farmacisti, Bologna

La Spesa Farmaceutica Convenzionata, per definizione, consiste quasi esclusivamente in spesa per farmaci di classe A (che sommano in media a circa il 99.7% della spesa). Fanno eccezione alcuni farmaci di classe C e H, come ad esempio i vaccini per l'influenza e particolari farmaci per il sistema nervoso, per i quali annualmente vengono spesi circa 20 milioni di euro.

2.4.2 Tipologia di farmaci

Un altro modo per classificare i farmaci, secondo le direttive dell'Organizzazione Mondiale della Sanità, è per tipologia, tramite il codice ATC (Anatomical Therapeutic Chemical classification

system). Questo sistema assegna ad ogni farmaco un codice alfanumerico, che riassume in 5 livelli gerarchici (dal primo, più generale, al quinto, più specifico) le caratteristiche del farmaco e i principi attivi in esso contenuti.

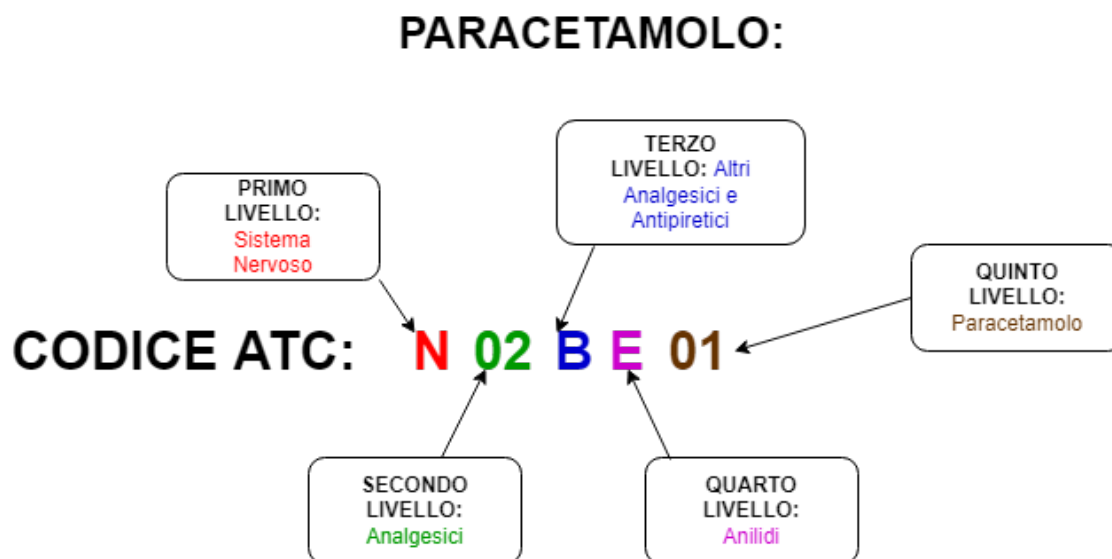


Figura 4: Codice ATC del paracetamolo

Il primo livello del codice ATC suddivide i farmaci in 14 categorie, indicandone il gruppo anatomico principale. Possiamo vedere, nella figura 5, come la spesa farmaceutica convenzionata si distribuisca in queste macrocategorie. Considerando i livelli successivi, le categorie analizzate nel database AIFA nei vari anni sono tra le 400 e le 500 (Tabella 1).

LIVELLO ATC	ANNO				
	2016	2017	2018	2019	2020
ATC1	14	14	14	14	14
ATC2	86	86	86	86	86
ATC3	204	202	199	197	199
ATC4	474	468	450	447	452

Tabella 1:
Numero di modalità presenti nel database AIFA, per ogni livello del codice ATC, 2016-2020

Spesa Pro-Capite per tipologia di Farmaco, Italia

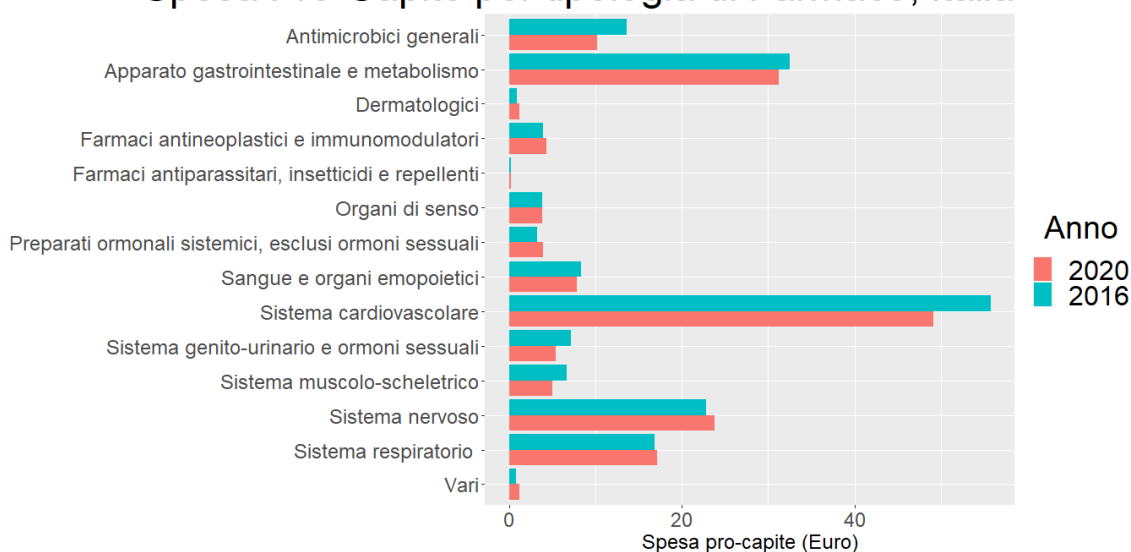


Figura 5

Si nota che per la maggior parte delle categorie è avvenuta una diminuzione della spesa, dal 2016 al 2020. Spiccano gli Antimicrobici, i farmaci per il sistema genito-urinario e i farmaci per il sistema muscolo-scheletrico, che hanno visto la corrispettiva spesa scendere di circa il 25% in 5 anni. In controtendenza, invece, ci sono i farmaci dermatologici, antineoplastici, per il

sistema nervoso e per il sistema respiratorio, dove nel 2020 si è speso di più rispetto al 2016.

2.4.3 Farmaci più venduti

Grazie al codice ATC, espresso nel database OsMed fino al quarto livello, è possibile calcolare quali siano i tipi di farmaci più venduti in Italia, sempre per quanto riguarda la spesa convenzionata lorda. Al primo posto si trovano, con ampio distacco e per tutto l'arco temporale considerato, i farmaci Inibitori della pompa acida, cioè quei farmaci atti a contrastare il reflusso gastroesofageo e l'eccessiva produzione di acido da parte delle pareti dell'esofago e dello stomaco. Tra il resto dei farmaci più venduti possiamo trovare farmaci contro il colesterolo (Inibitori della HMG CoA riduttasi), farmaci per l'abbassamento della pressione sanguigna (Betabloccanti, selettivi), farmaci anti-ipertensivi (ACE inibitori) e farmaci usati per ridurre la possibilità della formazione di trombi nel sangue (Anti-aggreganti piastrinici).

Categorie di Farmaco più vendute tramite la spesa convenzionata, Italia 2020

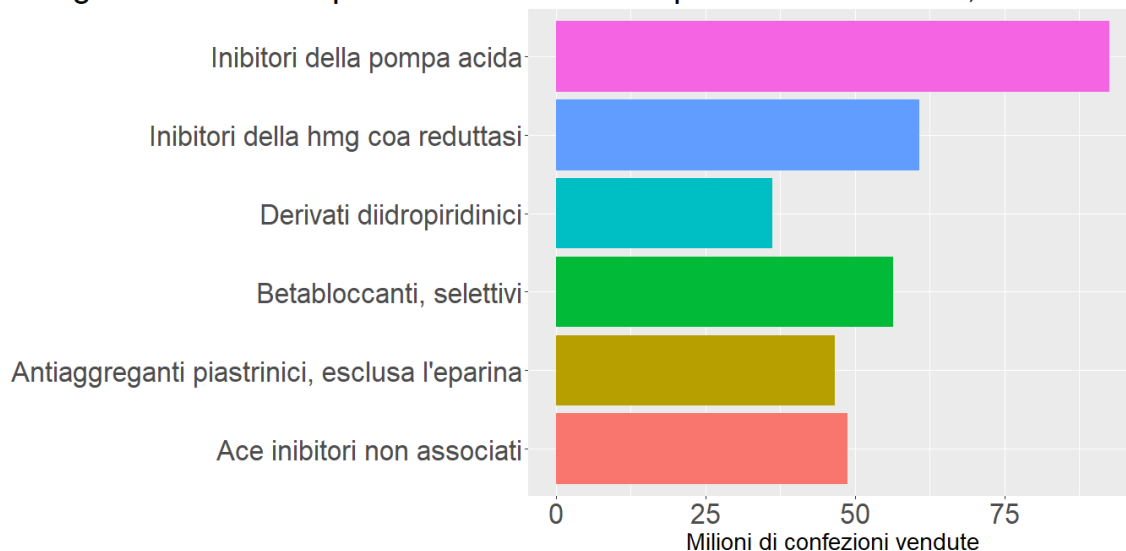


Figura 6

2.4.4 Concentrazione della spesa

Per capire se la spesa convenzionata lorda è composta più o meno uniformemente da ogni categoria di farmaco, si può fare ricorso a delle misure di concentrazione, come l'indice di Gini o la curva di Lorenz.

Un indice di Gini prossimo allo 0 indica che tutti i farmaci (raggruppati secondo il codice ATC fino al quarto livello) intervengono in maniera simile nella spesa totale. Viceversa, un indice vicino a 1 indica che pochi farmaci contribuiscono alla maggior parte della spesa.

Si evidenzia come, tra il 2016 e il 2020, l'indice di Gini della

spesa si mantenga tra 0.879 e 0.865 (Tabella 2), manifestando quindi un'ampia diseguaglianza nella distribuzione della spesa. Analizzando nel dettaglio questa distribuzione, possiamo notare come, nel 2020, per le 17 categorie di farmaci che hanno contribuito alla spesa in modo più sostanzioso, sia stata spesa una cifra pari a quella delle altre 435 categorie. Questo fenomeno è ben illustrato dalla curva di Lorentz, che mostra come la percentuale di spesa rispetto alla spesa totale cresca (asse Y), man mano che cresce la percentuale di categorie ATC considerate (Asse X). La situazione è simile negli anni precedenti, dove circa una quindicina di categorie costituiscono una spesa equivalente alle restanti.

INDICE DI GINI				
2016	2017	2018	2019	2020
0.879	0.875	0.867	0.865	0.867

Tabella 2

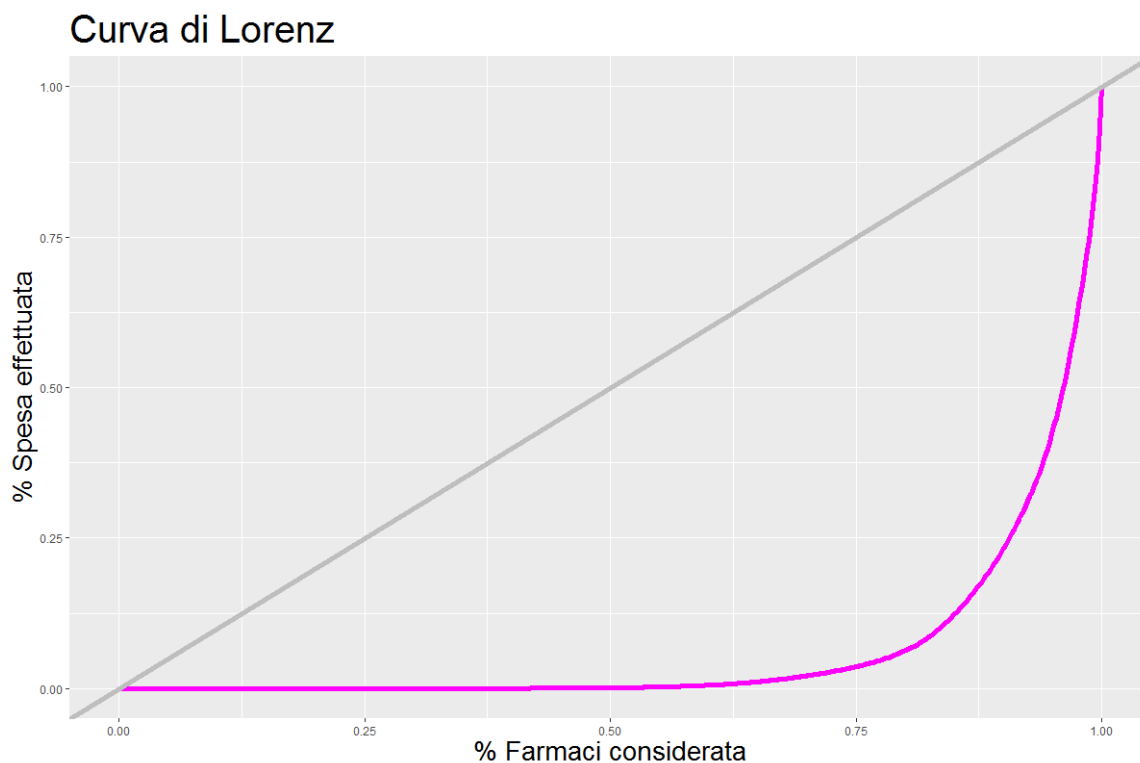


Figura 7: Curva di Lorenz della Spesa Farmaceutica Convenzionata, 2020

2.5 Disuguaglianze territoriali

Per evidenziare le differenze a livello geografico di quanto viene speso in farmacia tramite la spesa convenzionata, si calcola la spesa convenzionata pro-capite, dividendo la spesa totale effettuata in una certa area geografica, in un anno, per la popolazione residente nell'area, nello stesso anno (dati ISTAT). La spesa pro-capite media italiana nel 2016 era di 177€, ed è gradualmente diminuita di anno in anno fino ad arrivare, nel 2020, a 165€. Facendo una prima distinzione territoriale tra le tre

grandi macroaree italiane (Nord, Centro, Sud + Isole), si può notare come nel Nord la spesa media pro-capite rimanga al di sotto della media, mentre nel Sud Italia e nelle isole avviene l'opposto. La spesa nel Centro Italia rimane costantemente di poco sopra alla media nazionale, fino ad assumerne un valore molto simile nel 2020. Considerando l'andamento generale, si nota (Figura 8) che, in ogni macroarea, dal 2016 al 2020 la spesa convenzionata lorda pro capite sia andata gradualmente decrescendo.

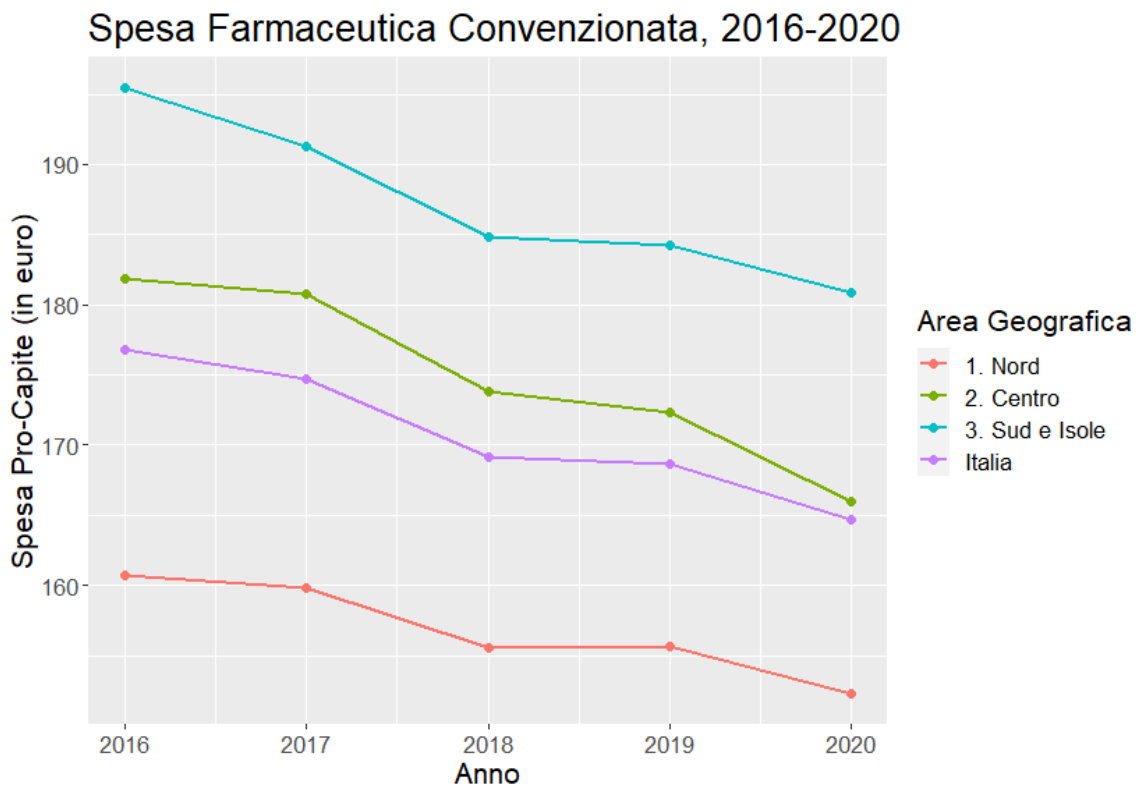


Figura 8

2.5.1 Vitamina D

La Vitamina D è un ormone liposolubile essenziale all'organismo umano. Essa regola l'assorbimento intestinale di calcio e fosforo, ed è quindi necessaria al metabolismo dell'apparato scheletrico. Una sua carenza in età infantile può causare rachitismo, mentre nei soggetti più anziani è usata per contrastare l'osteoporosi. Il corpo umano produce la Vitamina D principalmente a livello della cute, tramite l'esposizione solare, e in minor quantità dagli alimenti. Quando ciò non è sufficiente entrano in gioco gli integratori (Codice ATC A11CC), di cui si dispone dei dati di vendita. Le regioni dove in media si fa più uso di integratori risultano essere, nell'ordine, Abruzzo, Puglia e Basilicata; quelle dove se ne assumono di meno Sardegna e Toscana. Nel periodo 2016-2020 l'andamento del numero annuale di confezioni acquistate pro-capite si diversifica in base alla regione analizzata. Esso raggiunge un suo massimo nel 2018 o nel 2019 per la maggior parte delle regioni, mentre è prevalentemente decrescente in altre, come Lazio ed Emilia Romagna.

Un dato interessante da sottolineare, comune a tutte e 20 le regioni, è come il numero di confezioni vendute pro-capite nel 2020 sia sempre inferiore a quello del 2019, indipendentemente

dai trend regionali. Quasi sicuramente questo cambiamento ha una natura politica, infatti nell'autunno del 2019 l'AIFA ha emanato una nota che limitava la prescrizione dei farmaci contenenti Vitamina D, consentendone l'acquisto convenzionato solo in determinate condizioni mediche.

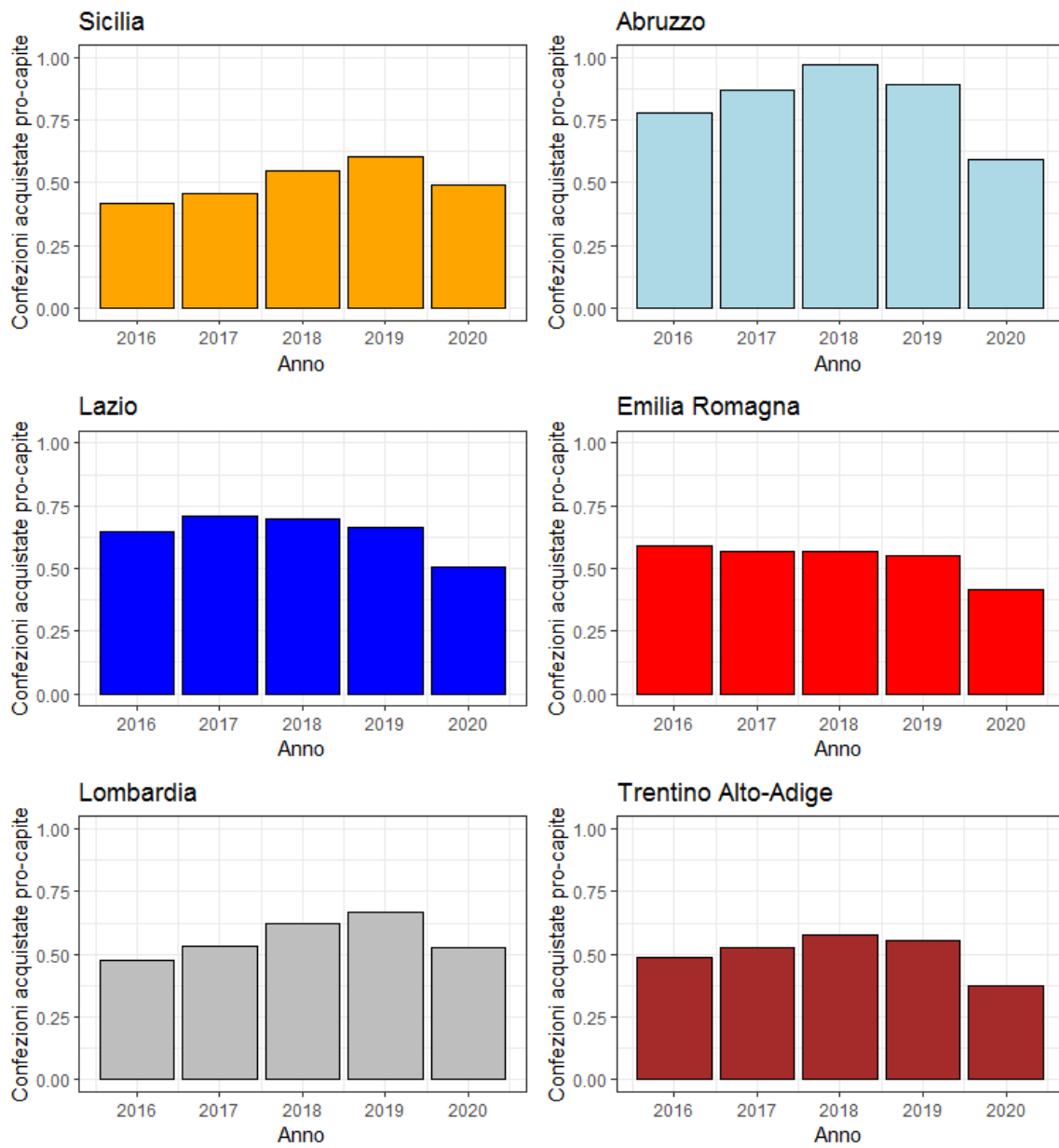


Figura 9: Vitamina D, confronto tra regioni Italiane

2.5.2 Inibitori della pompa acida

Come visto in precedenza (v.p. 2.4.3), i farmaci inibitori della pompa acida (Codice ATC A02BC) sono di gran lunga i più venduti tramite il canale della spesa convenzionata. Infatti, ogni anno in Italia in media viene acquistata più di una confezione per ogni italiano. Questo si può ipotizzare sia dovuto a molteplici fattori, come il fatto che una delle malattie contro cui questi farmaci vengono usati, il reflusso gastro-esofageo, sia notoriamente recidiva e in Europa colpisca circa il 10-20% della popolazione⁷. Il trend osservato delle vendite degli inibitori della pompa acida è, in ogni zona d'Italia, decrescente durante il periodo 2016-2020. Nel 2016, infatti, in 15 regioni su 20, si acquistavano in media più di 2 confezioni per abitante, mentre nel 2020 le uniche regioni dove ciò avviene sono Basilicata, Calabria e Campania. Considerando il dato nazionale, nel giro di 5 anni il numero annuale di confezioni vendute pro-capite ha rilevato un calo del 30.3%.

⁷Fonte: Humanitas Research Hospital

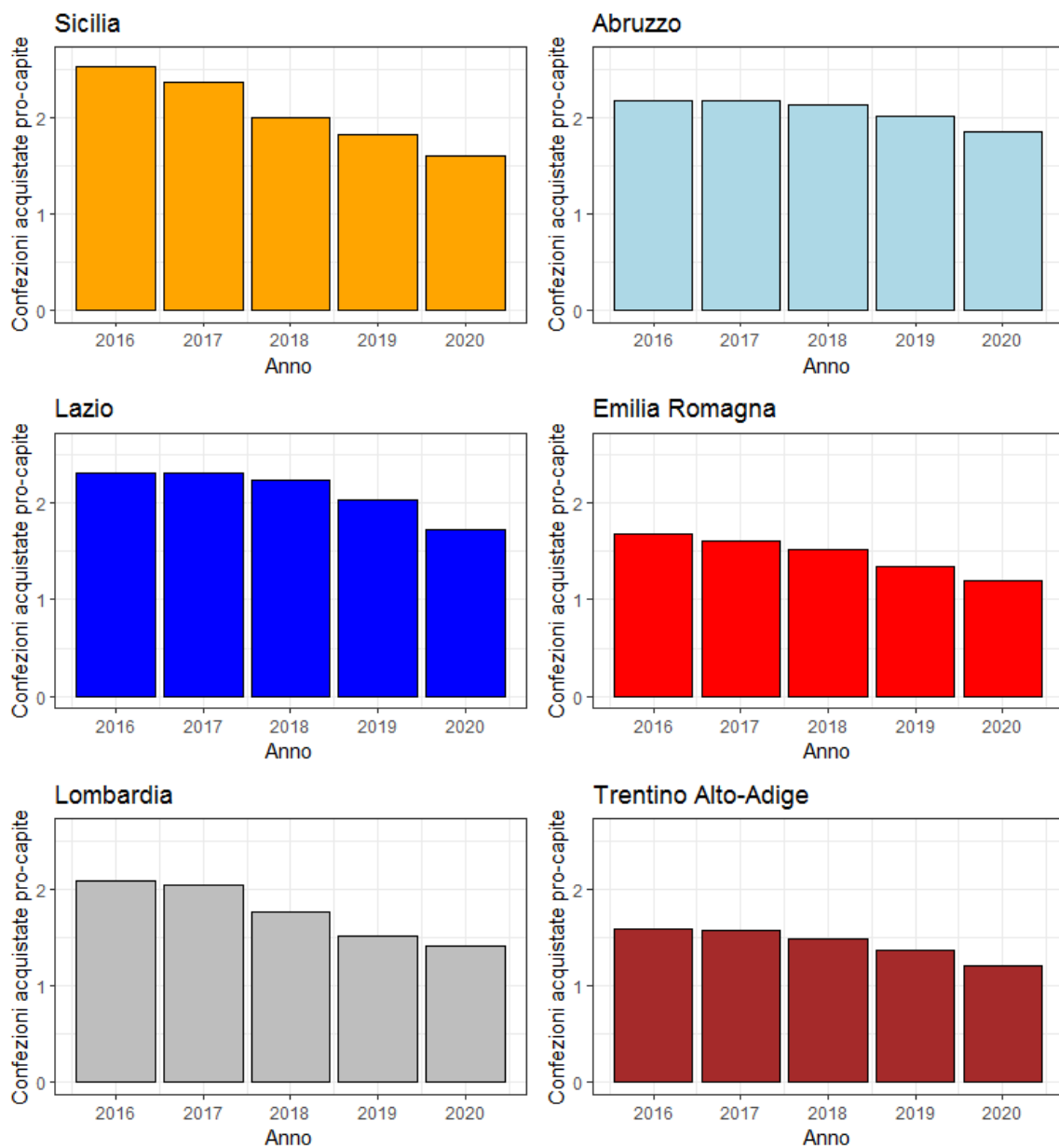


Figura 10: Inibitori della pompa acida, confronto tra regioni Italiane

2.5.3 Ormoni Tiroidei

L'ultima categoria di farmaco analizzata è quella degli ormoni tiroidei (Codice ATC H03AA). I farmaci che rientrano in questa categoria vengono prescritti principalmente in caso di ipotiroidismo, una condizione nella quale la tiroide (uno dei principali organi del sistema endocrino) non produce una quantità sufficiente di ormoni. Gli ormoni tiroidei sono farmaci con un utilizzo minore rispetto a quelli precedentemente analizzati, infatti in nessuna regione si supera il valore medio di 1 confezione acquistata ogni 2 abitanti. Il trend delle vendite dal 2016 al 2020 è comunque interessante, dato che in tutte le regioni italiane il numero annuale di confezioni acquistate pro-capite è andato in leggero ma costante aumento (+12,6% per quanto riguarda l'acquisto su base nazionale).

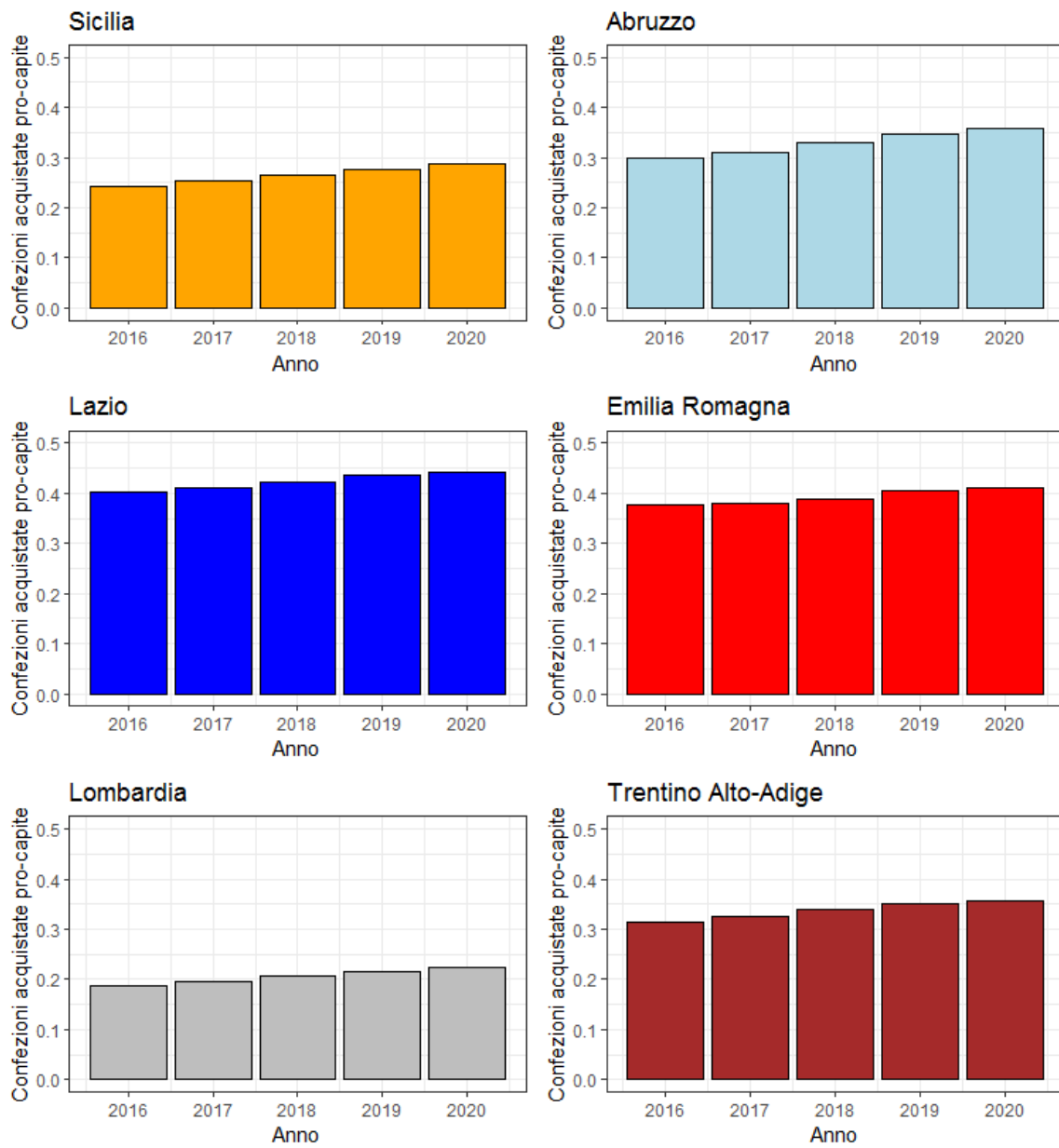


Figura 11: Ormoni Tiroidei, confronto tra regioni Italiane

3 Stile di vita

3.1 Raccolta dei dati

Ogni anno, l'ISTAT svolge in Italia un insieme di indagini sociali, chiamate Indagini Multiscopo sulle famiglie, al fine di ottenere le rilevazioni statistiche necessarie al paese. Una di queste è l'Indagine sugli "Aspetti della vita quotidiana", che raccoglie, tramite questionari familiari e individuali, informazioni sulle abitudini dei cittadini riguardo alla vita di tutti i giorni e allo stile di vita dal punto di vista sociale e personale. Il campione esaminato conta circa 25.000 unità familiari sparse in tutta Italia, e i risultati delle indagini sono consultabili liberamente in forma aggregata per età, sesso o regione nel database online Istat.

Verranno analizzati in questa tesi tre fattori rappresentativi dello stile di vita, i cui dati sono stati ricavati proprio da queste indagini negli anni da 2016 a 2020, collegati alla salute personale: l'obesità, l'attività sportiva e il fumo.

3.2 Obesità

Negli ultimi 70 anni, l'aumento della ricchezza nei paesi industrializzati e in via di sviluppo ha portato con sé anche alcune problematiche, oltre agli innumerevoli benefici. Una di queste è proprio l'obesità, una delle principali cause della riduzione dell'aspettativa di vita nei paesi occidentali, in quanto fattore di rischio di note patologie quali diabete mellito e ipertensione. I dati per quanto riguarda l'Italia, rispetto al resto del mondo, sono comunque positivi. Nel 2017 solo Corea del Sud e Giappone contavano una quota di persone obese minore rispetto al nostro paese⁸, dove si stima che attorno al 10.7% della popolazione appartenga a questa categoria.

Il trend degli ultimi anni, tuttavia, non si distingue per segno rispetto alle nazioni estere, con un incremento dell'obesità in Italia, dal 2016 al 2020, che si attesta sull'11%. Ai maschi, in Italia, è associato mediamente un tasso di obesità più alto rispetto a quello femminile (quasi 2 punti percentuali in più), mentre la classe d'età maggiormente affetta da obesità risulta essere quella tra 65 e 74 anni. Analizzando le differenze a livello territoriale (figura 13) si nota che solo in 3 regioni (Puglia,

⁸dati OCSE

Sardegna, e Abruzzo) la quota di persone obese è diminuita, mentre nel 2020 nel Lazio si è arrivati fino a un aumento di oltre il 25% della popolazione obesa, rispetto al 2016. Ragionando in termini assoluti Basilicata, Campania e Abruzzo contano il maggior numero di persone obese sulla rispettiva popolazione nel 2020 (rispettivamente 14.9%, 14.3%, 13.9%), contrariamente a Trentino Alto-Adige e Liguria (8.5% e 9.5%).

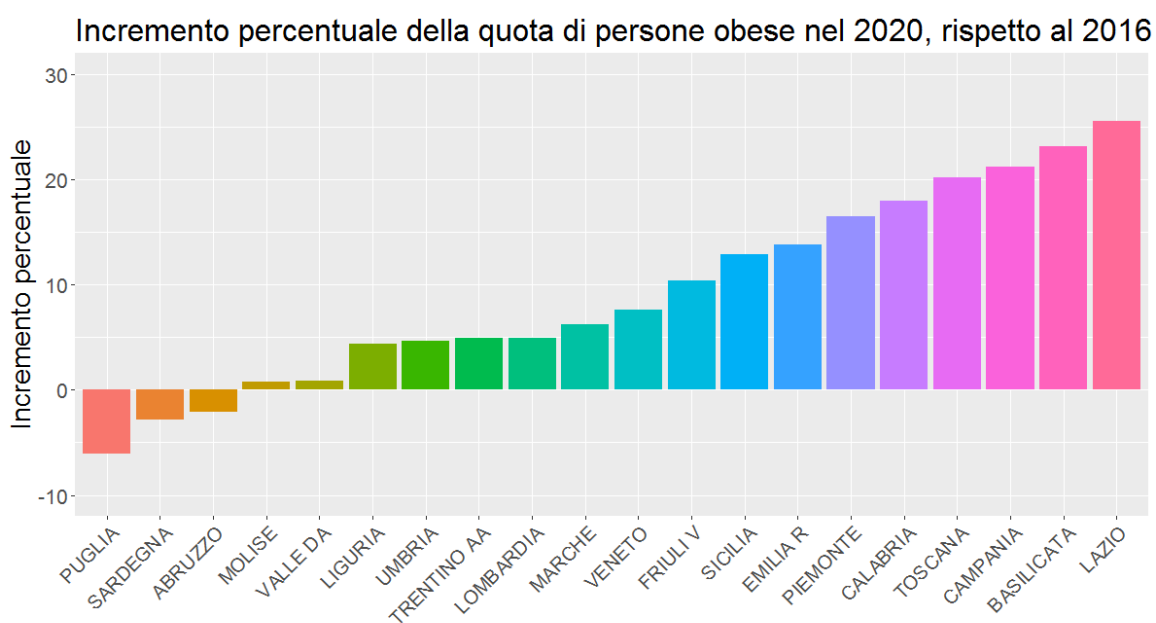


Figura 12

3.3 Attività fisica

È ben documentato come praticare attività fisica o sport comporti innumerevoli benefici per la salute, migliorando il be-

nessere fisico e mentale e riducendo il rischio di molteplici malattie.

Se nell'ultimo periodo l'obesità si può ritenere un fenomeno che sempre maggiormente condizionerà in negativo la salute degli italiani, non si può dire lo stesso dello svolgimento di attività fisica. Sempre secondo i dati Istat, coloro che si ritengono sedentari, cioè che non praticano alcun tipo di sport o attività fisica, sono in diminuzione negli ultimi anni. Nel 2020, infatti, circa 35 persone su 100 dichiaravano di essere sedentarie, contro le 39 su 100 del 2016. Bisogna comunque tenere a mente che questo dato è molto variabile rispetto all'età. Infatti, tutte le fasce d'età sotto i 55 anni contengono una quota di persone maggiore rispetto alla media nazionale che svolgono un qualche tipo di attività fisica, mentre per le fasce d'età sopra i 55 anni la quota è di gran lunga inferiore.

A livello regionale (Figura 14) si può notare come solo in Trentino Alto-Adige nel 2020 si pratici meno attività fisica rispetto al 2016, con un incremento del 2.6% delle persone sedentarie. Dal lato opposto risulta subito all'occhio la regione Liguria, dove si è passati da circa 38 persone sedentarie su 100 nel 2016 a solo 26 nel 2020.

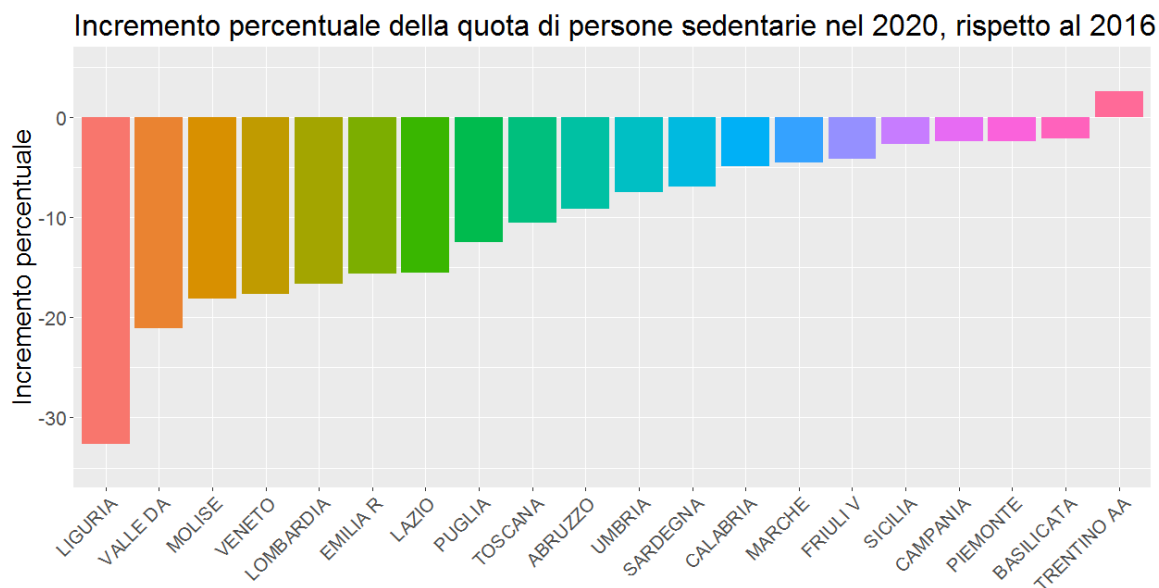


Figura 13

3.4 Fumo

Un altro dato interessante messo a disposizione dall'Istat riguarda la percentuale di fumatori, calcolata sulla popolazione di 14 anni o più. Il fumo è una tra le abitudini che notoriamente ha un impatto maggiore sulla salute, essendo un fattore di rischio considerevole di vari tumori e malattie dell'apparato cardio-circolatorio.

La tendenza a fumare in Italia è di poco diminuita tra il 2016 e il 2020, passando da un 19.8% di fumatori a un 18.6%. I maschi nel 2020 sono coloro che fumano di più: 22.2% di fumatori contro il 15.2% delle femmine. Per quanto riguarda l'età,

invece, la pratica del fumo si mantiene abbastanza costante tra le varie classi della popolazione adulta, riducendosi solamente nella popolazione più giovane o più anziana. Le abitudini delle varie realtà regionali si sono abbastanza modificate negli ultimi anni: nel 2016 la regione dove si fumava di più era la Campania, che però nel giro di 5 anni ha ridotto la sua percentuale di fumatori di quasi il 25%, lasciando il primato di questa classifica all'Emilia Romagna. I luoghi dove si fuma di meno, invece, nel 2020, sono Friuli Venezia-Giulia (14.7% di fumatori), Valle d'Aosta (16%) e Veneto(16.1%).

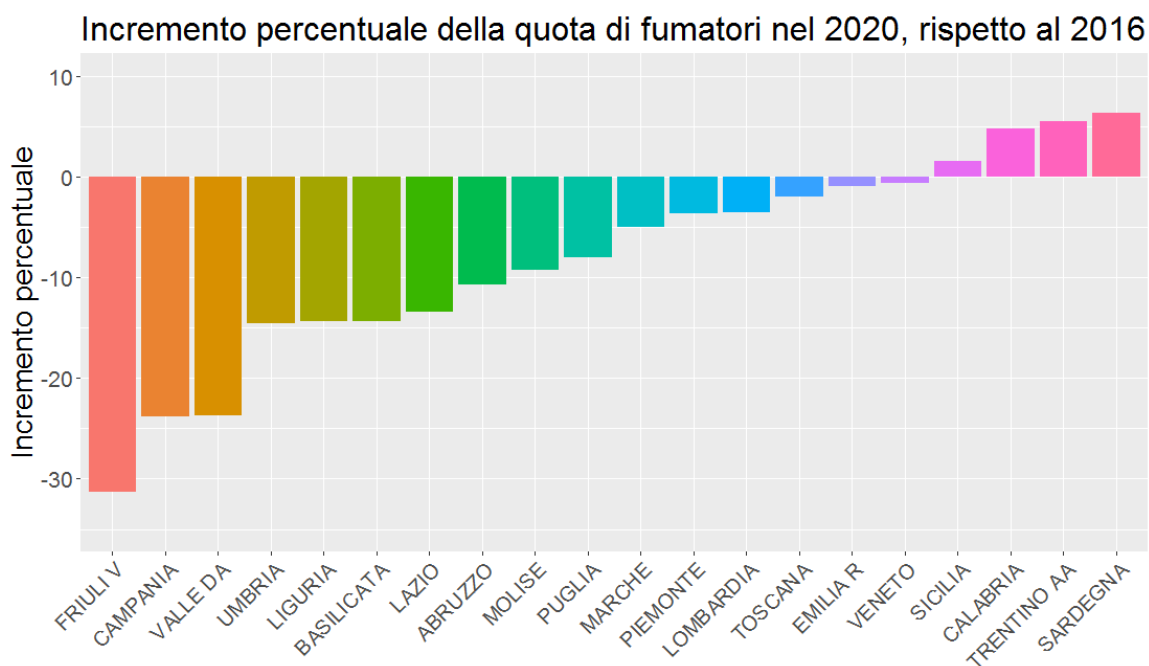


Figura 14

4 Spesa convenzionata e stile di vita

4.1 Correlazioni

Dopo aver analizzato singolarmente le caratteristiche della spesa farmaceutica convenzionata e delle abitudini di vita quotidiana che più influiscono sulla salute degli italiani, è interessante capire come tutte queste variabili si comportano se confrontate l'una con l'altra. Tramite l'utilizzo di un diagramma a dispersione (Figura 16) si possono notare le relazioni 2 a 2 tra spesa farmaceutica convenzionata pro-capite, percentuale di persone obese, percentuale di sedentari e percentuale di fumatori. Le unità statistiche considerate sono le 20 regioni italiane, i cui dati si riferiscono all'anno 2020.

	Spesa F. Conv.	Obesità	Sedentarietà	Abitudine al Fumo
Spesa F. Conv.	1	0.472	0.746	0.077
Obesità	0.472	1	0.685	-0.036
Sedentarietà	0.746	0.685	1	0.182
Abitudine al Fumo	0.077	-0.036	0.182	1

Tabella 3: Matrice di correlazione

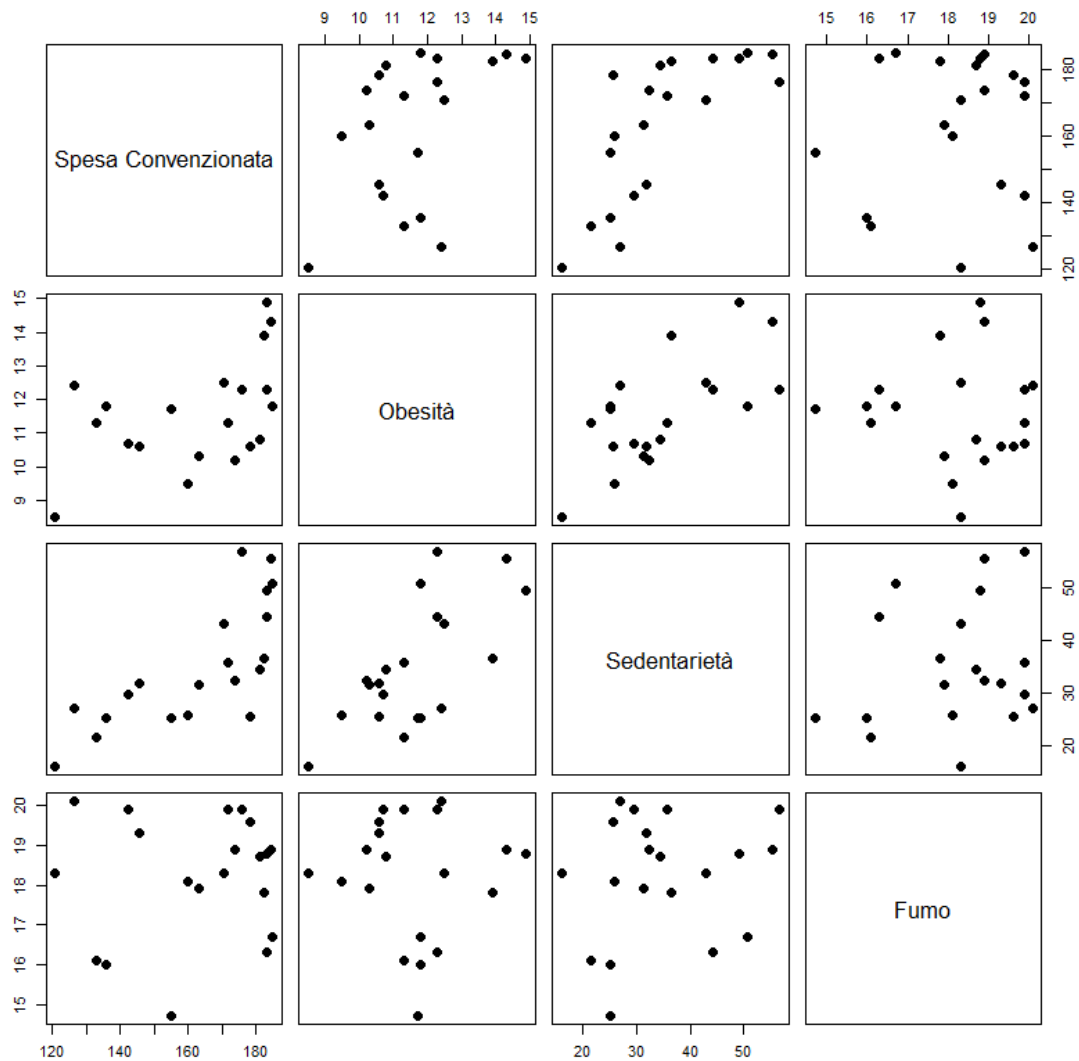


Figura 15: Valori calcolati sulle regioni italiane

Calcolando inoltre la matrice di correlazione (Tabella 3) si possono intuire quali sono le relazioni lineari più e meno forti tra le variabili. L'abitudine al fumo presenta in tutti e 3 i confronti delle correlazioni molto basse, mentre la spesa convenzionata

pro-capite, l'obesità e la sedentarietà sono tra loro correlate positivamente.

4.2 Analisi delle Correlazioni Canoniche

Per avere un'idea più generale della correlazione tra spesa farmaceutica convenzionata e stile di vita si può usare lo strumento della CCA (Canonical-Correlation Analysis) campionaria. Questa tecnica viene impiegata quando si vuole ottenere una misura globale della correlazione tra due gruppi di variabili (nel nostro caso, appunto, spesa farmaceutica convenzionata e stile di vita), e non ci è sufficiente conoscere le correlazioni tra le singole variabili.

Riferendosi sempre ai dati sulle regioni italiane nel 2020 si definiscono le matrici $X^{(1)}$ (20×1), con i dati della spesa convenzionata, e $X^{(2)}$ (20×3), con i dati di obesità, sedentarietà e abitudine al fumo. Lo scopo dell'analisi è determinare delle coppie di variabili canoniche campionarie $(\hat{u}_1, \hat{v}_1), \dots, (\hat{u}_p, \hat{v}_p)$ formate ognuna da due vettori che siano combinazioni lineari⁹ delle colonne, rispettivamente, di $X^{(1)}$ e $X^{(2)}$. In questo modo che i vettori possono rappresentare in una singola dimensione i due

⁹Si cerca, in particolare, la combinazione lineare che massimizza la correlazione

gruppi di variabili, così da poterne calcolare la correlazione.

In questa circostanza il primo gruppo racchiude soltanto una variabile (la spesa farmaceutica convenzionata), quindi $p=1$ e sarà possibile calcolare soltanto una coppia di variabili canoniche (\hat{u}_1 e \hat{v}_1). Il risultato che si ottiene è una correlazione tra \hat{u}_1 e \hat{v}_1 di **0.751**. Per interpretarlo dobbiamo capire in che modo \hat{u}_1 e \hat{v}_1 siano associate alle nostre variabili originarie. Il primo vettore rappresenta esclusivamente i valori della spesa farmaceutica condizionata, dato che essa è l'unica variabile nel gruppo $X^{(1)}$. Per quanto riguarda il secondo si ottiene:

```
> cc.dat1$scores$corr.Y.yscores
      [,1]
Obesità      0.6280407
Sedentarietà 0.9925967
Fumo         0.1024630
> |
```

Figura 16: Correlazione tra \hat{v}_1 e $X^{(2)}$

ciò significa che \hat{v}_1 rappresenta l'obesità in modo sostanziale, la sedentarietà quasi totalmente e l'abitudine al fumo in modo meno evidente. Possiamo concludere quindi, mettendo insieme le informazioni della Figura 17 e il risultato ottenuto di 0.751, che la spesa farmaceutica convenzionata pro-capite è positiva-

mente correlata ad uno stile di vita poco salutare.

4.3 Clustering

Il clustering (anche detto analisi dei gruppi o classificazione) è una procedura statistica che ha come obiettivo quello di suddividere le osservazioni (unità statistiche) in gruppi (cluster), in modo che le unità appartenenti allo stesso gruppo siano simili tra loro, ma ci siano differenze significative tra i vari gruppi. Nel nostro caso può essere utile capire come vengano suddivise in cluster le regioni italiane, sulla conseguenza dei dati raccolti dalla spesa farmaceutica e lo stile di vita. Sarà poi interessante analizzare se i cluster così formati corrisponderanno o meno a delle macroaree a livello geografico.

Per farlo si utilizza l'algoritmo K-Means, un algoritmo ben adatto a raggruppare le osservazioni in base a delle variabili quantitative. K-Means è un algoritmo iterativo non gerarchico, ciò significa che è necessario scegliere il numero di gruppi da formare prima di eseguire l'algoritmo. Una volta fatto questo, l'algoritmo assegna casualmente ad ogni gruppo le unità statistiche, calcola le medie di ogni gruppo, e poi sposta ogni unità nel gruppo più vicino. Questi ultimi 2 passi vengono poi ripetuti

fino a quando non è più necessario spostare nessun osservazione da un gruppo all'altro. Applicando K-Means ai nostri dati, scegliendo di partizionare le regioni italiane in 3 cluster nell'anno di riferimento 2020, si ottiene una classificazione soddisfacente, con la devianza tra i gruppi (betweenSS) che spiega l'83.9% della devianza totale.

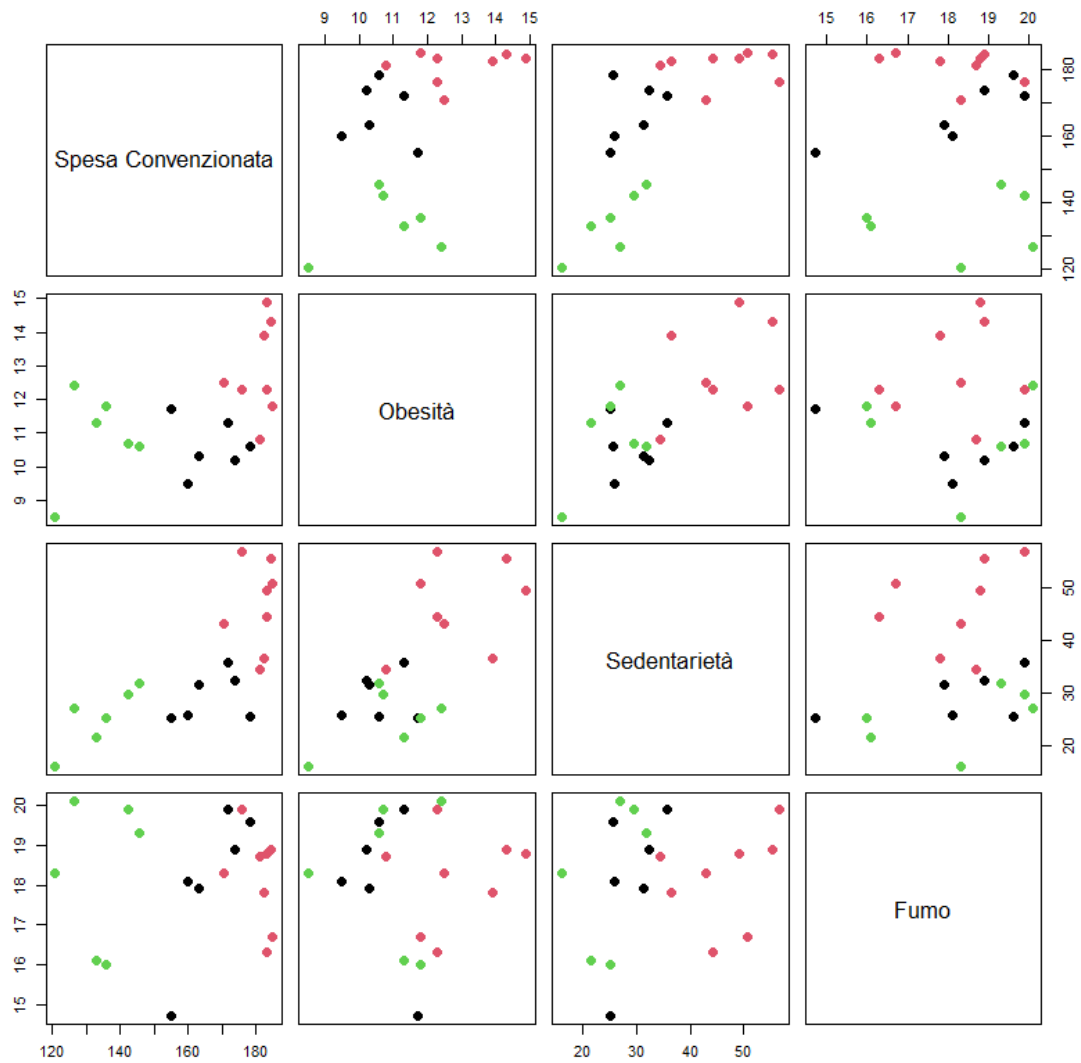


Figura 17: Valori calcolati sulle regioni italiane, suddivise in 3 cluster secondo l'algoritmo K-means

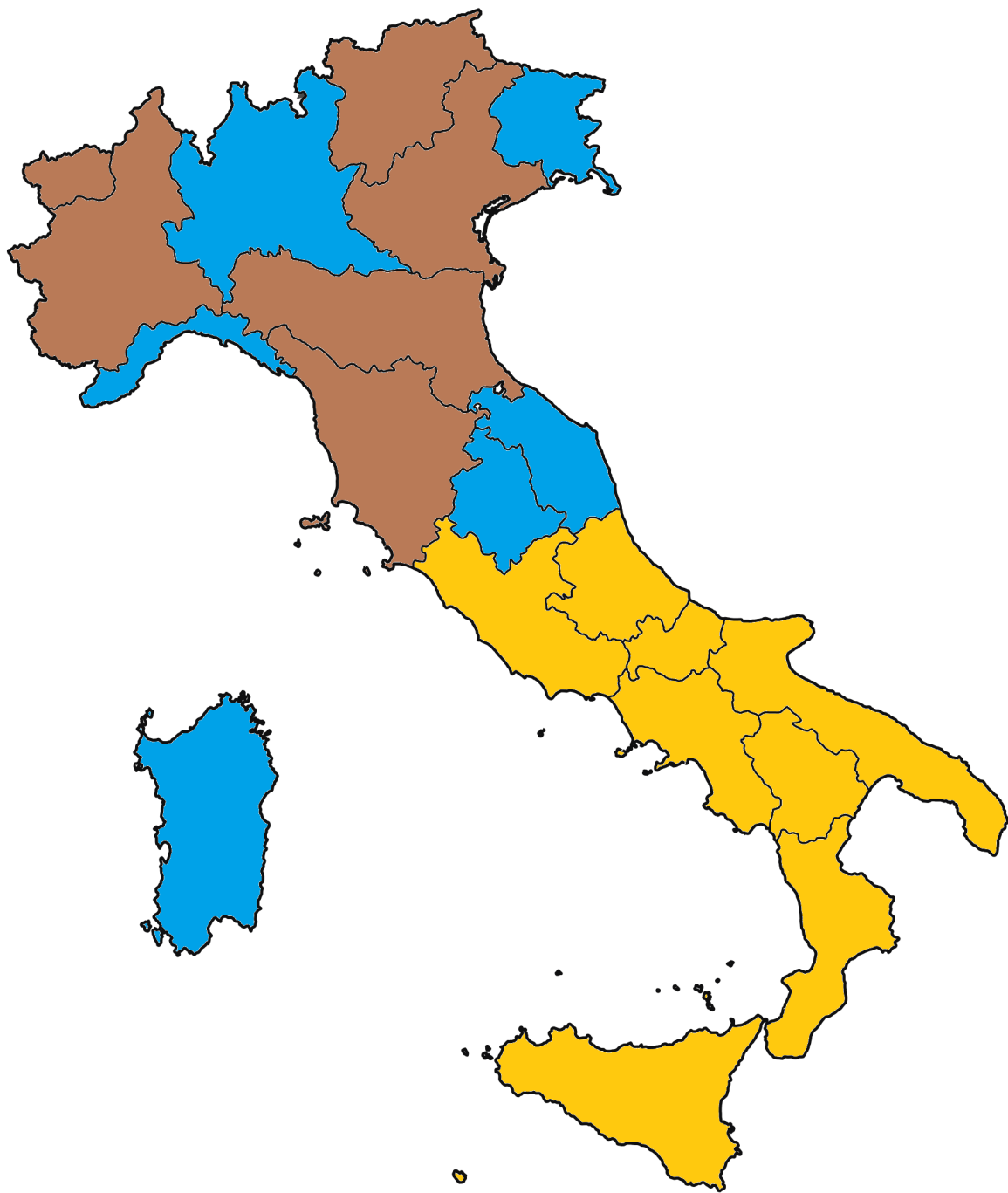


Figura 18: Regioni italiane suddivise in 3 gruppi dall'algoritmo K-Means

Il primo gruppo (colorato in marrone, Figura 18) comprende la maggior parte delle regioni del Nord (Valle d’Aosta, Piemonte, Trentino Alto-Adige, Veneto, Emilia Romagna) e la Toscana. È il gruppo con la spesa convenzionata più bassa (con una media tra le spese regionali pro-capite di 133€) e la minor percentuale di persone obese (media del 10.6%). Segue un secondo gruppo (colorato in giallo) che presenta un’alta spesa convenzionata, un’alta prevalenza dell’obesità e un’alta percentuale di persone sedentarie. Esso si identifica con il Lazio e tutte le regioni del Sud Italia.

	Spesa F. Conv.	% Obesità	% Sedentarietà	% Abitudine al Fumo
Gruppo 1	133.87€	10.88	25.13	18.28
Gruppo 2	180.64€	12.85	46.33	18.18
Gruppo 3	166.95€	10.60	29.32	18.18

Tabella 4: Medie dei singoli cluster

Rimane poi un ultimo gruppo (in blu), che può definirsi ”misto”, composto da Sardegna, Umbria, Marche, Liguria, Lombardia e Friuli Venezia Giulia, con una spesa convenzionata intermedia rispetto agli altri gruppi ma un maggior numero di persone che mediamente praticano attività fisica. Non è sorprendente il fatto che la percentuale media di fumatori sia pressoché identica in tutti e 3 i cluster: come visto in precedenza (Tabella 3), questa variabile ha una bassa correlazione con tutte le al-

tre, di conseguenza non ci si aspetta che a forti differenze tra i gruppi nelle altre 3 variabili corrisponda una forte differenza tra i gruppi nella percentuale di fumatori.

5 Conclusioni

In Italia negli ultimi anni la spesa farmaceutica è in costante crescita, ma non in ogni ambito. Infatti l'importante aumento della spesa privata (+7.2% dal 2018 al 2019) e di ospedali e ASL (+10.9%) si contrappone alla decrescita della spesa convenzionata lorda in farmacia. Quest'ultima non è uniforme in tutto il territorio nazionale, ma è mediamente più bassa nelle regioni del Nord Italia, e riguarda principalmente farmaci contro il reflusso gastro-esofageo, farmaci per l'abbassamento del livello di colesterolo e farmaci atti a ridurre la pressione arteriosa.

Per quanto riguarda, invece, le abitudini degli italiani, nei luoghi dove la spesa farmaceutica convenzionata è maggiore si tende ad avere uno stile di vita meno salutare, specialmente prendendo in considerazione gli indici di obesità e attività fisica. Poche informazioni sono state ricavate dall'abitudine al fumo, che è molto variabile in tutta Italia.

6 Bibliografia

<https://www.aifa.gov.it/open-data>

<https://www.dati.istat.it>

<https://www.aifa.gov.it/documents/20142/1728113/nota-96.pdf>

<https://www.istat.it/it/archivio/91926>

<https://www.epicentro.iss.it/farmaci/osmed-rapporto-2019/>

<https://www.aifa.gov.it/documents/20142/1205984/rapporto-osmed-2019.pdf/>

<https://www.oecd.org/health/health-systems/Obesity-Update-2017.pdf/>

<https://www.oecd.org/health/health-systems/Obesity-Update-2017.pdf/>

<https://www.humanitas.it/malattie>

[https://www.who.int/publications/i/item/WHO-HIS-HGF-HFWorkingPaper-18.](https://www.who.int/publications/i/item/WHO-HIS-HGF-HFWorkingPaper-18)

3