

Dipartimento di Impresa e Management

Cattedra di Informatica

IL POTERE DEI BIG DATA:

COME PREDIRE I RISULTATI DELLE ELEZIONI  
EUROPEE

Luigi Laura

---

RELATORE

208861

---

FRANCESCA CELIO

*Ai miei genitori, che ci hanno  
creduto prima ancora che lo  
facessi io*

## INDICE

<b>CAPITOLO 1:</b> .....	<b>6</b>
<b>PREMESSA: “LA MACCHINA DELLA VERITÀ”, DI SETH STEPHENS DAVIDOWITZ</b> .....	<b>6</b>
1.1 IL LIBRO: “LA MACCHINA DELLA VERITÀ” .....	6
1.1.1 <i>L’idea dietro il libro</i> .....	7
1.2 LO SCOPO DE “LA MACCHINA DELLA VERITÀ” E COSA CI INSEGNA	
10	
1.2.1 <i>Offrire nuovi tipi di dati</i> .....	13
1.2.2 <i>Fornire dati onesti</i> .....	14
1.2.3 <i>Zoomare su piccoli gruppi di individui</i> .....	16
1.2.4 <i>Fare esperimenti causali</i> .....	16
1.2.5 <i>Limiti dei Big Data</i> .....	17
<b>CAPITOLO 2: SITUAZIONE POLITICA IN ITALIA</b> .....	<b>20</b>
2.1 QUADRO POLITICO: ELEZIONI EUROPEE 2019 .....	20
2.1.1 <i>Situazione politica in Italia</i> .....	21
2.2 IL TASSO DI INTERESSE SU GOOGLE È UN PARAMETRO UTILE A PREDIRE	
I RISULTATI DELLE ELEZIONI EUROPEE? .....	26

2.2.1 <i>Affluenza effettiva alle elezioni europee del 2019</i> .....	30
2.2.2 <i>Astensionismo</i> .....	33
2.2.3 <i>Standard Eurobarometer</i> .....	46
<b>CAPITOLO 3:</b> .....	<b>50</b>
<b>PREDIRE I RISULTATI DELLE ELEZIONI EUROPEE DEL 26</b>	
<b>MAGGIO 2019</b> .....	<b>50</b>
3.1 STRUMENTI UTILIZZATI .....	51
3.2 PUNTO DI PARTENZA.....	54
3.3 USARE GOOGLE TRENDS .....	55
3.4 TURNING POINT.....	58
3.5 SITUAZIONE POLITICA IN ITALIA PRIMA DELLE ELEZIONI EUROPEE	
MOSTRATA DA GOOGLE TRENDS.....	60
3.6 POPOLARITÀ DI MATTEO SALVINI.....	61
3.7 FUN FACTS.....	64
3.7.1 <i>Matteo Salvini e Ultimo</i> .....	64
3.7.2 <i>Matteo Salvini e Weather today</i> .....	66
3.7.3 <i>Casi realmente correlati o Correlazione Spuria?</i> .....	67
<b>CONCLUSIONI</b> .....	<b>71</b>

<b>RINGRAZIAMENTI.....</b>	<b>75</b>
<b>BIBLIOGRAFIA.....</b>	<b>77</b>
<b>SITOGRAFIA .....</b>	<b>79</b>

## **Capitolo 1:**

### **Premessa: “La Macchina della Verità”, di Seth Stephens Davidowitz**

#### **1.1 Il Libro: “la Macchina della Verità”**

L'idea alla base di questa tesi è nata dal libro “la Macchina della Verità”, scritto da Seth Stephens-Davidowitz, il quale ha questa capacità unica di spiegare concetti statistici, anche abbastanza complessi, senza essere troppo tecnico e riuscendo ad essere facilmente comprensibile. Uno dei punti cardine del libro è quello di essere curiosi, interrogandosi su questioni riguardanti la vita di tutti i giorni, e studiando in che direzione si stia muovendo il mondo, attraverso i meravigliosi strumenti che Google rende disponibili ed accessibili a tutti. Altro punto fondamentale, che spinge S.S. Davidowitz ad utilizzare questi strumenti, è la verità dietro alle ricerche su Google, infatti le persone che rispondono ad un sondaggio, nonostante venga garantita l'anonimità dello stesso, tendono a mentire nelle loro risposte, mentre si affidano a Google per qualunque problematica o qualunque dubbio. Oltre il 50% degli intervistati nei sondaggi mente; in America, hanno sottoposto donne e uomini ad un sondaggio in cui si chiedeva ad ogni intervistato quanti preservativi usasse per i rapporti eterosessuali ogni anno. Il numero dichiarato da entrambi i sessi, chiaramente, dovrebbe essere lo stesso, invece, gli uomini hanno affermato di usare 1.6 miliardi di preservativi l'anno, mentre, le donne solo 1.1 miliardi. Chi mente? Entrambi. La

risposta viene fornita da Nielsen, le cui ricerche attestano che vengono venduti solamente 600 milioni di preservativi all'anno negli Stati Uniti.

Con queste premesse e con la solita base del lavoro di Davidowitz, in cui ha investigato sulle elezioni presidenziali americane, ho provato a fare la stessa cosa con le elezioni europee, datate 26 maggio 2019, in Italia. Alcuni dei risultati e delle correlazioni scoperte sono decisamente inaspettate.

### ***1.1.1 L'idea dietro il libro***

Tutto ha avuto inizio nel 2008 con il lungo dibattito che ha caratterizzato le elezioni presidenziali di quell'anno. Si è discusso di quanto fosse importante e di quanto condizionasse gli elettori il colore della pelle del candidato democratico Afro-Americano Barack Obama, si è inoltre parlato del pregiudizio razziale, ancora, purtroppo, diffuso nel Paese. In quel periodo furono fatti molti sondaggi e una percentuale variegata di cittadini americani fu intervistata, in maniera anonima, sulla questione. I risultati furono positivi, in quanto la maggior parte degli intervistati dichiarò di non considerare il colore della pelle di Obama un fattore determinante nella propria scelta di voto. Nel 2008, S.S. Davidowitz, ha scoperto Google Trends, uno strumento che indica agli utenti con quale frequenza, in un determinato Paese o in una zona di esso (nel caso dell'Italia la divisione è su base regionale, per quanto riguarda il continente americano si divide il territorio in Stati), parole e frasi vengono cercate su Google, in luoghi e tempi diversi. Durante la notte delle elezioni, Davidowitz,

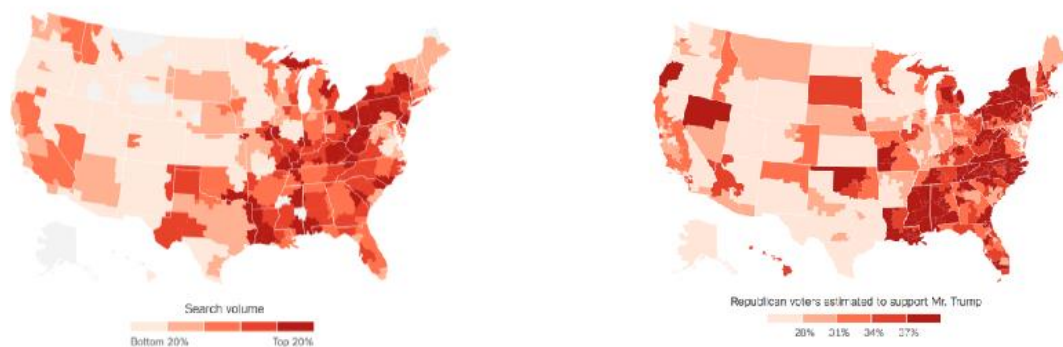
utilizzando Google Trends, ha notato che una ricerca, su cento che contenevano il termine “Obama”, presentava anche le parole “Ku Klux Klan” o “Nigga”. Allo stesso tempo in alcuni Stati le ricerche per “Nigga President” superavano in numero quelle per “First Black President”. Nella stessa notte, inoltre, le iscrizioni al sito nazionalista, convinto assertore della supremazia dei bianchi, “Stormfront” sono state dieci volte più numerose rispetto al solito. Le informazioni appena elencate possono in parte aiutare a capire il successo, otto anni dopo, che ha avuto Donald Trump, 45° Presidente degli Stati Uniti, che ha racchiuso all’interno della sua campagna elettorale concetti molto cari agli utenti che hanno a cuore concetti nazionalisti.

All’inizio lo studio e le ricerche di S.S. Davidowitz furono rigettati da ben cinque giornali accademici, i quali sostenevano fermamente l’impossibilità di un odio razziale così diffuso in America, ma con l’elezione di Donald Trump la sua ricerca è sembrata molto più veritiera e credibile. Donald Trump era stato dato dai sondaggi con qualche punto percentuale in meno rispetto a quelli che realmente gli sono poi appartenuti in seguito allo scrutinio dei voti. Nate Silver, uno statista e scrittore americano, definì impossibile la vittoria di Trump alle presidenziali, così, dopo aver visto i risultati delle elezioni decise di analizzare i dati per provare a capire cosa stesse succedendo su tutto il territorio americano. Notò così che l’area geografica in cui Trump ha avuto più successo formava un disegno particolare.



Racist Search on Google

Support for Trump in Primary



**Figura 1. Scoperta di Nate Silver – “La macchina della verità” – S.S. Davidowitz**

Trump stava andando bene in alcune zone del Nord-Est, del Sud e del Midwest industriale.

Silver notò, quindi, che il fattore maggiormente correlato al supporto di Trump era lo stesso che aveva trovato Davidowitz in precedenza: le zone in cui il Presidente americano aveva ottenuto il maggior numero di voti erano le stesse in cui la parola “nigger” era stata cercata con più frequenza su Google dagli utenti. Questo è un altro esempio concreto che mostra quanto le persone tendano a mentire nei sondaggi ma mai alle ricerche effettuate su internet, forse anche perché si sentono meno giudicate. Per questo motivo S.S. Davidowitz ha definito Google come il “digital truth serum”, perché nessuno mente mai nelle proprie ricerche.

## **1.2 Lo scopo de “La macchina della verità” e cosa ci insegna**

In “La macchina della verità”, Davidowitz, affronta il tema attuale dell’uso dei Big Data per dare nuove spiegazioni ad alcuni comportamenti degli esseri umani e se i risultati mostrino esattamente quello che ci saremmo aspettati o, al contrario, se si traggano conclusioni lontane dall’immaginario collettivo. Il suo libro è un’analisi molto interessante di come si possono utilizzare i Big Data per studiare cosa le persone dicono di fare e di pensare, contro quello che realmente fanno e dicono. Come già detto, tendiamo a mentire nei questionari o nelle interviste, soprattutto quando si tratta di pensieri o comportamenti di cui ci si vergogna, perché magari negativi o imbarazzanti, ma a Google non mentiamo mai, anzi, confessiamo le cose più strane. Nella barra bianca di ricerca non inseriamo solo le parole degli argomenti su cui vogliamo sapere di più, ma confessiamo anche le nostre passioni segrete, le nostre preoccupazioni, desideri, problemi, insicurezze, i nostri schieramenti politici e le nostre scelte religiose. Insomma, Google sa tutto di noi e delle nostre ricerche. Le ricerche di informazioni degli utenti sono esse stesse delle informazioni. In quest’ottica si potrebbe pensare che anche i social media, attraverso lo studio dei likes, dei follow, dei luoghi in cui ci si tagga o dell’uso delle gif, possano fornirci delle informazioni utili sui pensieri degli utenti, ma in realtà non è così. Sui social gli utenti mostreranno sempre e solo la parte migliore della loro vita e mentiranno ogni giorno per provare a sembrare più “attraenti” ed essere più invidiati agli occhi di chi li guarda. Al pari di quanto fatto con Google, Davidowitz ha dato una definizione anche a Facebook e a tutti i social media in

generale: “Digital brag to my friends about how good my life is - Serum”. Quanto appena detto è stato facile da dimostrare; è bastato cercare le prime cinque combinazioni di parole più cercate su Google inserendo la voce “my husband is” (mio marito è) e confrontarlo con le combinazioni delle stesse parole più frequentemente utilizzate all’interno dei post sui social media, in questo caso si è utilizzato Facebook.

## Top descriptions of husbands

<i>Social Media Posts</i>	<i>Searches</i>
the best	gay
my best friend	a jerk
amazing	amazing
the greatest	annoying
so cute	mean

**Figura 2. Ricerche contenenti le parole “mio marito è” su Google e sui post sui social**

Come si nota, nella prima colonna, le combinazioni di parole, associate a “mio marito è”, più utilizzate nei post sui social media, sono tutte parole con connotazioni positive: “the best”, “my best friend”, “amazing”, “the greatest” e “so cute”. Questo perché come

detto prima se tendiamo a mentire nei questionari, figuriamoci sui social in cui la nostra vita è esposta alle critiche delle persone che sono a noi più vicine. A Google invece non si mente ed infatti l'unica voce che resta uguale rispetto ai post su Facebook è “amazing”, per il resto le ricerche hanno meno connotazioni positive e sono più sincere: “gay”, “a jerk”, “annoying” e “mean”. È sorprendente pensare che questa enorme differenza sia data semplicemente dal luogo in cui si digitano le ricerche (dati), e a come la voglia di fare bella figura sia più forte della verità.

In un solo giorno vengono prodotti in media 2,5 trilioni di byte di dati e ogni byte è un'informazione diversa. Da uno strumento utilizzato per conoscere il mondo, adesso Google è diventato uno strumento per conoscere l'uomo. Ci sono moltissime fonti di dati, che solo 15 anni fa non esistevano, su una gamma di argomenti vastissimi, che ci offrono finestre su aree su cui in passato potevamo soltanto provare ad indovinare.

“La macchina della verità” ci insegna come fare per poter sfruttare appieno i Big Data, senza perderci dentro un quantitativo di informazioni così elevato e dando un senso ad essi, cercando di analizzarli e traendo conclusioni.

S.S. Davidowitz ha elencato quattro grandi poteri dei Big Data:

1. Offrire nuovi tipi di dati
2. Fornire dati onesti
3. Permetterci di zoomare su piccoli gruppi di individui
4. Permetterci di fare esperimenti causali

### *1.2.1 Offrire nuovi tipi di dati*

Il primo potere dei Big Data è ampliare le conoscenze, negli anni le categorie di informazioni sono cambiate e sono anche aumentate in numero, in complessità e in facilità nel reperirle. L'avvento dei Big Data ha reso le nuove tipologie di informazioni accessibili a tutti, il loro potere non conta solo delle loro dimensioni, ma anche del fatto che si tratti di una vera e propria novità. I Big Data forniscono anche un nuovo modo di approcciarsi allo studio di fenomeni sociali di tutti i giorni, infatti Davidowitz ha utilizzato le ricerche su Google per provare a calcolare indici come il tasso di disoccupazione. Il ragionamento utilizzato è il seguente: i disoccupati hanno molto tempo libero che non sanno come impiegare. L'analista quindi ha deciso di provare a studiare le ricerche effettuate su Google, tra il 2004 e il 2011. Utilizzando Google Correlates si è reso conto che le principali ricerche collegate all'argomento disoccupazione, non riguardavano l'ufficio di collocamento o posizioni di lavoro aperte, ma "Slutload", un sito porno, e "Spider Solitaire", un gioco di carte per computer. Questo perché oltre a cercare un nuovo lavoro, i disoccupati cercano anche un modo per passare il tempo. Nessuno prima di Davidowitz aveva mai utilizzato le ricerche sugli hobby delle persone per predire il tasso di disoccupazione, partendo da argomenti completamente diversi. Questi dati da soli magari non sono precisi al 100% ma danno un'idea della situazione attuale e di sicuro questo metodo deve essere inserito all'interno di un modello per calcolare questi valori sociali.

### *1.2.2 Fornire dati onesti*

Come già detto in precedenza: tutti mentono! Si mente su cose poco importanti come quante volte si va in palestra, quanto si pesa e quanti anni si hanno, figuriamoci quanto possano mentire le persone quando si fanno domande un po' più di spessore. Le persone non riferiscono i loro reali comportamenti, né tantomeno i loro pensieri più intimi ed imbarazzanti, perché vogliono fare bella figura, nonostante con i sondaggi sia garantita l'anonimità, a loro interessa il modo in cui appaiono. Questo fenomeno si chiama “bias della desiderabilità sociale”. Negli anni cinquanta uscì un articolo che forniva prove sufficienti per dimostrare quanto i sondaggi siano vittime di bias. I ricercatori chiesero, ovviamente specificando l'anonimità della risposta, ai cittadini di Denver alcuni dati sulla propria vita: se fossero iscritti al registro elettorale, se fossero andati a votare nelle ultime elezioni, se possedessero una tessera della biblioteca e se avessero donato dei soldi in beneficenza nell'ultimo periodo. In seguito, i ricercatori raccolsero dei dati su questi argomenti e confrontandoli con le risposte del sondaggio si stupirono nello scoprire che i dati non corrispondevano. Questo grafico mostra i risultati ottenuti:

	Riportato dal Sondaggio	Dati reali
Iscrizione al registro elettorale	83%	69%
Voto nelle ultime elezioni presidenziali	73%	61%
Voto nelle ultime elezioni comunali	63%	36%
Possesso della tessera di una biblioteca	20%	13%
Donazione recente a una campagna di beneficenza	67%	33%

**Figura 3. Risposte positive al sondaggio e dati reali (in percentuale)**

In alcuni casi le percentuali reali risultano essere addirittura inferiori del 50% rispetto a quelle dichiarati dagli intervistati. Il secondo potere dei Big Data è quindi quello di fornire dei dati onesti, ci lasciano vedere quello che gli utenti veramente pensano, vogliono e fanno, e non quello che dicono di pensare, volere e fare.

Google ci mostra la realtà, che risulta essere drasticamente differente rispetto a quella esibita sui social media.

### ***1.2.3 Zoomare su piccoli gruppi di individui***

Il terzo potere dei Big Data consiste nella possibilità di focalizzarsi su una piccola fetta della popolazione, riuscendo a dividere gli utenti in categorie. I gruppi possono essere di vario tipo, in base al dato più utile per la ricerca specifica; si va dalla zona di residenza, all'età, al tempo in cui si effettua la ricerca ecc..

S.S. Davidowitz ha risposto alla domanda “È l’America il Paese delle opportunità?”. Una persona i cui genitori appartengono al 20% della popolazione con reddito più basso, può arrivare a far parte del 20% con guadagni più alto? La risposta è stata data da Raj Chetty, un economista americano e professore ad Harvard, ha zoomato su alcune zone geografiche e ha scoperto che la probabilità varia a seconda dell’area di appartenenza. Focalizzandosi su alcuni gruppi di individui ha calcolato la probabilità che i loro redditi crescano e ha trovato una risposta veritiera alla domanda: né sì né no, le possibilità di crescita e quindi le opportunità dipendono dalla zona di partenza.

### ***1.2.4 Fare esperimenti causali***

Il quarto e ultimo potere individuato da Davidowitz è la possibilità di fare esperimenti causali rapidi ed incontrollati. Gli ingegneri di Google sono stati i primi a capire l’importanza dello studio dei dati, hanno compreso che un’analisi dei dati risultava molto più conveniente in termini economici, di tempo, di impiego di risorse e anche di quantità e qualità delle informazioni raccolte, rispetto ad un questionario o ad un sondaggio. Per questi esperimenti non è nemmeno necessario reclutare dei partecipanti,



come nel caso delle interviste anonime, perché gli utenti di Google partecipano agli esperimenti utilizzando normalmente i propri dispositivi, generando così delle informazioni utili ai fini della valutazione. Gli esperimenti randomizzati sono stati chiamati “Test A/B”, sono in grado di scoprire effetti davvero causali.

Come scrive Davidowitz nel libro “nell’era dei Big Data tutto il mondo è un laboratorio”, perché basta semplicemente essere connessi ad internet per prender parte ad un esperimento. Solo nel 2011 gli ingegneri di Google hanno sviluppato settemila test A/B e il numero, ovviamente, continua a salire vertiginosamente con il passare del tempo.

### ***1.2.5 Limiti dei Big Data***

Ne “la macchina della verità” Davidowitz non mostra solamente le potenzialità dei Big Data, ma anche i limiti e i modi in cui non possiamo e non dobbiamo utilizzarli.

Come detto in precedenza i Big Data sono molto utili per raccogliere nuove informazioni su campi ancora non analizzati in questo modo; ma su campi già studiati risultano essere poco proficui. È molto più semplice capire meglio il razzismo, l’aborto e gli abusi sui minori di quanto non sia possibile ottenere nuovi dati su un argomento come il mercato degli affari, perché qualunque dettaglio, in questo campo, è già stato ampiamente analizzato. Davidowitz e Lawrence Summers, un economista e politico americano, insieme hanno provato a capire se i Big Data possano consentire di fare previsioni sul mercato azionario.

La risposta in breve è: no, non possono.

Si immagini che il mercato delle azioni sia collegato ad una moneta fortunata, per sceglierla si numerino mille monete e si tiri una moneta al giorno per due anni. Nel 70,3% delle volte in cui la moneta 391 è caduta dal lato della testa, l'indice S&P è salito. Il rapporto è molto significativo dal punto di vista statistico, questo vuol dire che la moneta 391 è la nostra moneta fortunata. A questo punto si provi a lanciare la moneta 391 ogni giorno, secondo quanto detto prima, con una correlazione così alta si dovrebbe guadagnare se si comprano azioni quando esce testa. Ovviamente questa situazione non si verifica, perché i due fenomeni non sono collegati, però analizzando solo i Big Data si potrebbe pensare che la correlazione sia veritiera. Davidowitz definisce questo errore come “la maledizione della dimensionalità”, essa può verificarsi quando ci sono tante variabili (mille monete tra cui scegliere quella fortunata, perché le probabilità che si verifichi la situazione della 391, sono più elevate) e un numero limitato di osservazioni (504 giorni di mercato azionario in due anni). Infatti, se si diminuiscono le variabili, proviamo con 100 monete, le probabilità che una di esse risulti fortunata sono molto più basse. Allo stesso modo, se si aumentano le osservazioni (monitorare l'indice S&P per 20 anni) si nota come la teoria sulla moneta fortunata non riesca a tenere il passo con la realtà. Il problema dei Big Data è che non sempre quello che possiamo misurare è anche realmente utile ed importate per noi. Per esempio, Facebook può misurare ed analizzare dati come i like, i follow e le condivisioni, ma per sapere se il servizio offerto realmente soddisfa i consumatori ha bisogno di sottoporre agli utenti dei questionari di valutazione/soddisfazione.

Facebook quindi utilizza un approccio tradizionale in questo caso, perché anche i Big Data hanno bisogno di “Small Data” per avere una visione completa del mondo.

## Capitolo 2: Situazione politica in Italia

### 2.1 Quadro politico: Elezioni Europee 2019

Il 26 maggio gli italiani sono andati alle urne per votare per l'elezione quinquennale dei membri del Parlamento europeo. Dal 1979 i 751 deputati vengono eletti in modo diretto, in base al risultato delle elezioni il consiglio europeo sceglie il candidato per la commissione europea che dovrà ottenere la fiducia della maggioranza del Parlamento<sup>1</sup> e formare la commissione. L'Italia ha dovuto votare per eleggere 73 deputati. Il sistema elettorale è di tipo proporzionale<sup>2</sup>, secondo la legge elettorale italiana per il Parlamento europeo del 1979<sup>3</sup>, aggiornata successivamente nel 2009 con l'introduzione della soglia di sbarramento fissata al 4% e nel 2014 per la possibilità di esprimere voti di preferenza.

Il territorio italiano è diviso in cinque circoscrizioni plurinominali:

- Italia nord-occidentale (Piemonte, Valle d'Aosta, Liguria e Lombardia)
- Italia nord-orientale (Veneto, Trentino-Alto Adige, Friuli-Venezia Giulia ed Emilia-Romagna)

<sup>1</sup> “Note tematiche sull’Unione Europea”. Eeva Pavy. Maggio, 2019. Articolo Online:

<http://www.europarl.europa.eu/factsheets/it/sheet/19/il-parlamento-europeo-competenze>

<sup>2</sup> Come funziona la legge elettorale per le europee. Il Post, 16 maggio 2019. Articolo Online:

<https://www.ilpost.it/2019/05/16/leggi-elettorali-europee/>

<sup>3</sup> “Elezioni europee 2019 – Manuale elettorale”. Studi Camera. 26 giugno,2019. Articolo Online:

<https://temi.camera.it/leg18/temi/elezioni-europee-2019.html>

- Italia centrale (Toscana, Umbria, Marche e Lazio)
- Italia Meridionale (Abruzzo, Molise, Campania, Puglia, Basilicata e Calabria)
- Italia insulare (Sicilia e Sardegna)

Ognuna di queste può eleggere un numero prestabilito di deputati<sup>4</sup>, e il totale sarà sempre 73.

### ***2.1.1 Situazione politica in Italia***

L'Italia sta vivendo una profonda fase di cambiamento, negli ultimi tempi abbiamo visto, dalle statistiche politiche risultati diversi tra loro e completamente inaspettati, il tutto però con una differenza di anni minima. Il declino dei partiti tradizionali e la crescita di quelli di destra e di quelli populistici di centro è stata la condanna verso il baratro dei partiti socialisti, popolari e storicamente di sinistra. I primi due hanno creato il Governo, con un “accordo di governo”, che, nonostante i continui e normali attriti, ha governato nell'ultimo anno.

Per l'Italia le elezioni del Parlamento europeo sono state un vero e proprio banco di prova per misurare i consensi ottenuti dal Governo, composto dalla Lega Nord e dal Movimento 5 Stelle, che si è insediato a Palazzo Chigi nel giugno 2018, e hanno messo

<sup>4</sup>“Decreto del Presidente della Repubblica”. Gazzetta Ufficiale della Repubblica Italiana, Istituto Poligrafico e Zecca dello Stato. 22 Marzo, 2019. Articolo Online:  
<https://www.gazzettaufficiale.it/eli/id/2019/03/25/19A02051/sg>

in discussione la solidità del Governo stesso, perché a seconda dei risultati si sarebbero potuti prevedere scenari diversi. Le opzioni erano di 2 tipi:

- Se le percentuali tra i due partiti fossero rimaste invariate o fossero variate di poco rispetto alle elezioni del 4 marzo 2018, il Governo non avrebbe avuto grandi cambiamenti
- Se le percentuali di consensi fossero cambiate in maniera importante si sarebbero potuti aprire vari scenari: il più preoccupante era quello in cui il partito con più consensi avrebbe potuto volontariamente far cadere il Governo per poi, formarne un altro o con altri alleati (nel caso della Lega Nord, si sarebbe potuto pensare agli altri due partiti che si erano presentati nella stessa coalizione alle elezioni Politiche del 4 marzo 2018: Fratelli d'Italia e Forza Italia) o da soli, nel caso in cui si raggiungesse la maggioranza

In tutta la storia della Seconda Repubblica Italiana, dopo il passaggio da un sistema elettorale proporzionale a uno maggioritario, lo scenario politico in questo Paese è sempre stato un bipolarismo di due differenti coalizioni, due gruppi in concorrenza tra loro per avere la maggioranza al Governo e al Parlamento. Recentemente, durante le elezioni del 2013, è stato determinato l'avvento del tripolarismo dal Movimento 5 Stelle, un partito di centro fortemente contrario a qualunque tipo di alleanza o coalizione tra partiti. Nel 2013, quindi, la politica italiana ha dovuto affrontare una situazione di tripolarismo e l'impatto che questo cambiamento repentino ha avuto è stato molto forte. In aggiunta alla crisi economica scoppiata nel 2008, il Bel Paese si è

trovato anche nel mezzo di una crisi istituzionale, per via dell'incertezza generata dalle elezioni, perché nessuna delle tre coalizioni fu in grado di ottenere una vittoria netta sulle altre, il Parlamento ebbe problemi a dare fiducia ad un Governo e il mandato del Governo uscente stava giungendo al termine. Per risolvere tale situazione di incertezza istituzionale furono applicate delle procedure speciali:

- Nomina di una commissione di 10 “saggi”, chiamati dal Presidente della Repubblica, per elaborare un programma di riforme attorno a cui riunire una maggioranza parlamentare che sostenesse un nuovo governo
- Rielezione del Presidente della Repubblica uscente, Giorgio Napolitano
- Stipula di un largo accordo di governo (Enrico Letta come Presidente del Consiglio), per riuscire ad avere la maggioranza in entrambe le camere

Le conseguenze di queste misure portarono al voto l'Italia alle successive elezioni, il 4 marzo 2018, con un sistema elettorale completamente diverso. Dal 2013 al 2015 vi è stato il tentativo di varare una riforma elettorale, prima con la legge Calderoli ed in seguito con l'Italicum (valida solo per la Camera dei Deputati), ma ambedue furono successivamente pronunciate incostituzionali da parte della Corte Costituzionale. La legge Rosato (da Ettore Rosato), fu approvata definitivamente il 26 ottobre 2017<sup>5</sup>, ma si ebbe la sua prima applicazione per le elezioni del 4 marzo 2018, conosciuta come

<sup>5</sup> “Decreto del Presidente della Repubblica”. Gazzetta Ufficiale della Repubblica Italiana, Istituto Poligrafico e Zecca dello Stato. 22 Marzo, 2019. Articolo Online: <https://www.gazzettaufficiale.it/eli/id/2019/03/25/19A02051/sg>

“Rosatellum”<sup>6</sup>, disciplina l’elezione sia della Camera dei Deputati che del Senato della Repubblica. L’impianto della legge si configura come un sistema elettorale misto a separazione completa, i due sistemi non sono inter-comunicanti tra di loro, quindi non vi è la possibilità di compensazione tra le due modalità di attribuzione dei seggi.

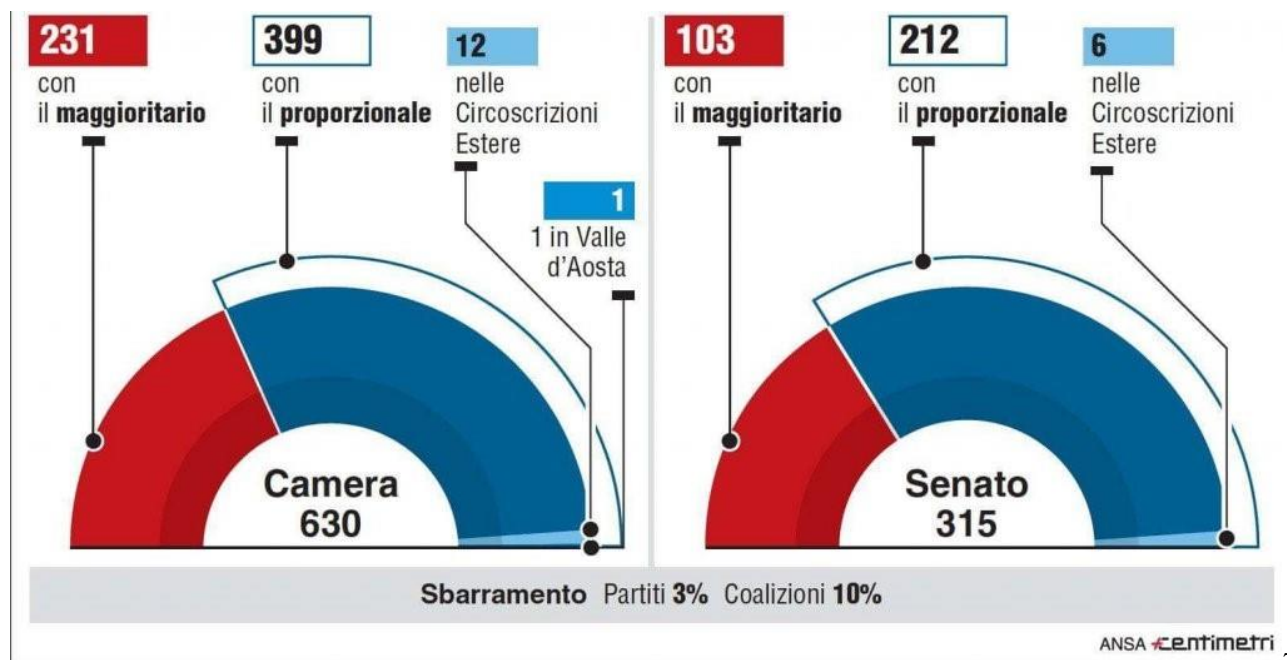
Vi sono tre modalità di assegnazione dei seggi:

- Il 37% dei seggi viene ripartito con un sistema maggioritario in altrettanti collegi uninominali: per ognuno di essi viene eletto il candidato più votato
- Il 61% viene assegnato in maniera proporzionale tra le coalizioni e le liste che abbiano superato la soglia di sbarramento, la ripartizione dei seggi è effettuata a livello regionale per il Senato e nazionale per la Camera
- Il 2% viene assegnato dal voto degli italiani in territorio estero, con un sistema proporzionale che prevede il voto di preferenza

<sup>6</sup>“Rosatellum bis: la nuova legge elettorale. Ecco come funziona (la scheda)”. Sgherza Alessio. 26 ottobre, 2017. Articolo Online:

[https://www.repubblica.it/politica/2017/09/22/news/rosatellum\\_bis\\_scheda-176193741/?refresh\\_ce](https://www.repubblica.it/politica/2017/09/22/news/rosatellum_bis_scheda-176193741/?refresh_ce)





**Figura 4. Divisione dei seggi in Camera e Senato con l'applicazione del Rosatellum**

I due principali problemi connessi al tripolarismo sono:

1. la governabilità: il tripolarismo incrementa il rischio che dopo le elezioni non si venga a creare una maggioranza parlamentare, e di conseguenza, anche al Governo; presupposti per una nuova crisi istituzionale
2. la sopravvivenza del tripolarismo: è un sistema non consolidato, che non è mai stato sperimentato prima; un terzo polo, formato solo da un partito di centro, che sta nascendo e crescendo e i due poli storici (centro - destra e centro - sinistra) che contemporaneamente si stanno evolvendo.

<sup>7</sup>“Rosatellum 2.0, come sarebbero assegnati i seggi a Camera e Senato”. Ansa-Centimetri. 22 luglio, 2019. Articolo Online: [https://www.tgcom24.mediaset.it/politica/infografica/rosatellum-2-0-come-sarebbero-assegnati-i-seggi-a-camera-e-senato\\_1001044-2017.shtml](https://www.tgcom24.mediaset.it/politica/infografica/rosatellum-2-0-come-sarebbero-assegnati-i-seggi-a-camera-e-senato_1001044-2017.shtml)

I partiti storici hanno continuato a proporre lo stesso schema, dando poca importanza al Movimento 5 Stelle, terzo polo emergente, che ha dimostrato, il 4 marzo 2018, di essere in realtà il primo partito d'Italia per numero di consensi. Il Paese è uscito dalle elezioni del 4 marzo 2018 con una situazione completamente inaspettata: il Governo è stato formato con un'alleanza di Movimento 5 Stelle e Lega Nord (capogruppo della coalizione di destra). Il primo ha ottenuto, da solo, come partito, il 32% dei voti, la coalizione del secondo invece ha ottenuto il 37%. Con queste elezioni si è superato lo storico, e ormai antico, concetto della contrapposizione tra destra e sinistra.

## 2.2 Il tasso di interesse su Google è un parametro utile a predire i risultati delle elezioni europee?

Per rispondere a questa domanda occorre prima considerare il tasso di astensionismo che caratterizza tutte le elezioni italiane, per farlo ho provato a studiare le ultime tre elezioni europee.

Cercando “Elezioni Europee” come argomento su Google Trends è apparso un grafico con una funzione perennemente piatta, ma con dei picchi a cadenza regolare durante i periodi delle elezioni stesse (ogni 5 anni).

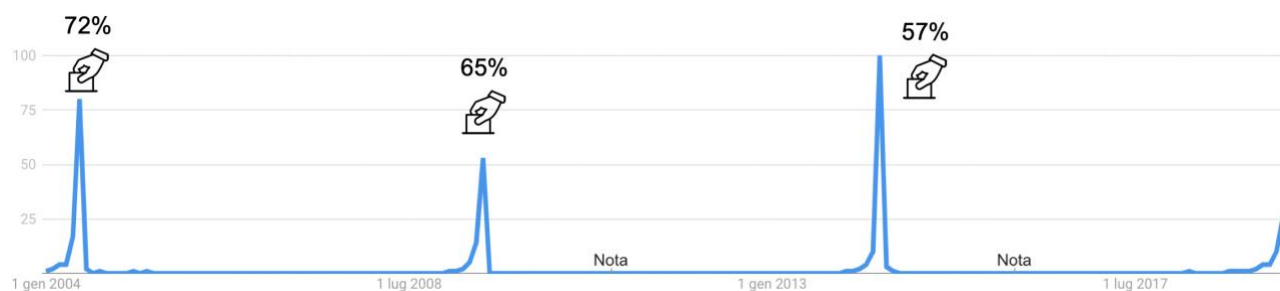


Figura 5. Interesse verso le scorse elezioni europee e percentuali di partecipazione al voto

Questa (figura 5) è la situazione che è apparsa a maggio, poco prima delle elezioni europee del giorno 26. Si può notare, in basso a destra del grafico, che la curva è in crescita quasi esponenziale, nettamente più inclinata rispetto altri tre anni in cui gli italiani sono stati chiamati al voto. Si nota come il 2014 sia stato l'anno in cui l'interesse ha raggiunto il suo punto massimo, anche se l'affluenza è diminuita.

Se il trend continuasse in questo senso allora ci si potrebbe aspettare che quest'anno il volume di ricerche sull'argomento sia ancora maggiore, almeno del 40%, se non di più.

In alto, su ogni picco, sono state inserite le percentuali di affluenza al voto per ogni anno, è evidente come con il passare del tempo essa sia diminuita in maniera importante. È quindi evidente che il tasso di interesse non sia un parametro accettabile per predire i risultati delle elezioni europee. Le persone cercano informazioni sulle elezioni e sui modi per esprimere le proprie preferenze, ma questo non vuol dire che successivamente andranno a votare. D'altronde non c'è un modo affidabile per riuscire a capire, prima della votazione stessa, quante persone andranno a votare. Come già detto, anche chiedendo agli elettori se abbiano intenzione di votare o meno, non si avrebbe un dato certo, perché tenderebbero a mentire.

Roger Tourangeau, professore dell'università del Michigan, nei suoi studi e successivamente, nel suo libro "The Psychology of Survey Response"<sup>8</sup>, ha affermato

<sup>8</sup> Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press

che si tende a mentire ai sondaggi anonimi per via della debolezza del genere umano per le “white lies” (bugie bianche). Per circa un terzo delle giornate le persone mentono nella vita di tutti i giorni e questo processo si ripete, ovviamente, anche nelle interviste. In aggiunta a questo, c'è anche l'abitudine che abbiamo di mentire a noi stessi, siamo abbastanza riluttanti a mostrarci a noi stessi o agli altri per chi siamo realmente, per paura di essere giudicati. Sicuramente non vogliamo sentirci giudicati dall'intervistatore, ma la paura più grande è quella di deludere noi stessi e quindi non più quella di essere criticati da un terzo. Ecco spiegato perché i dati sono quanto più attendibili quanto più l'intervista o il sondaggio sono impersonali. I questionari online sono più sicuri delle chiamate al telefono, che a loro volta sono più sicure delle interviste porta a porta o personali. Il problema, secondo Tourangeau, è che le persone non hanno alcun incentivo a dire la verità nei questionari.

Quando non c'erano fonti di dati ufficiali, prima dell'avvento dei Big Data, era veramente difficile provare a capire cosa le persone realmente pensassero e facessero. Questo secondo Davidowitz è proprio il secondo potere dei Big Data: alcuni siti spingono le persone a dire cose che in altre circostanze non ammetterebbero mai. Nelle ricerche su Google, considerate miglior questionario possibile, gli utenti sono online, a casa, da soli e non c'è un intervistatore. Oltre a questo vi sono anche degli incentivi che vengono offerti agli utenti: se si hanno dei problemi, nessuno sarebbe mai incentivato a raccontarli in un sondaggio, mentre più informazioni si danno a Google su un determinato argomento e più corretta sarà la soluzione che esso trova, quindi è interesse dell'utente non mentire.

In qualunque caso in tema di astensionismo il potere dei Big Data è limitato, perché nemmeno Google può prevedere quante persone saranno pigre il giorno delle votazioni o quante avranno degli imprevisti che impedirà loro di esprimere la propria preferenza. Questo è il motivo per cui il tasso di interesse, quindi il numero di volte in cui le persone hanno cercato informazioni sulle elezioni, può essere un parametro utilizzabile ma non affidabile al 100%.



**Figura 6. Interesse verso le scorse elezioni europee, comprese le ultime: 26 maggio 2019**

Il primo grafico (figura5) mostrato in questo paragrafo indicava il trend di ricerca di “Elezioni Europee” prima che esse avessero luogo, questo perché la prima analisi da me fatta risale ai giorni precedenti le elezioni stesse. Quest’ultimo grafico invece mostra i dati aggiornati a dopo le elezioni. Come si può notare il numero di ricerche di quest’anno, non solo supera quello degli anni precedenti, ma addirittura doppia il numero di ricerche effettuate nel 2014, proprio come ci aspettavamo dall’analisi svolta in precedenza.

### ***2.2.1 Affluenza effettiva alle elezioni europee del 2019***

Nonostante l'interesse dimostrato nelle ricerche effettuate su Google dagli italiani sia aumentato notevolmente, il tasso di affluenza alle urne, continuando a muoversi come negli altri anni, ha continuato la sua decrescita. Solo il 56.10% degli italiani ha espresso la propria preferenza alle urne, circa il 2% in meno rispetto agli altri anni<sup>9</sup>. Si è votato di più solo in Trentino Alto Adige, mentre nel resto d'Italia la percentuale di votanti è diminuita ovunque, soprattutto in Sardegna e in Sicilia<sup>10</sup>. In questo grafico si osservano, negli anni, le varie percentuali di affluenza alle urne in Italia per l'elezione del Senato europeo<sup>11</sup>.

<sup>9</sup> "Europee: affluenza in netto aumento, la più alta da 20 anni". Redazione ANSA.

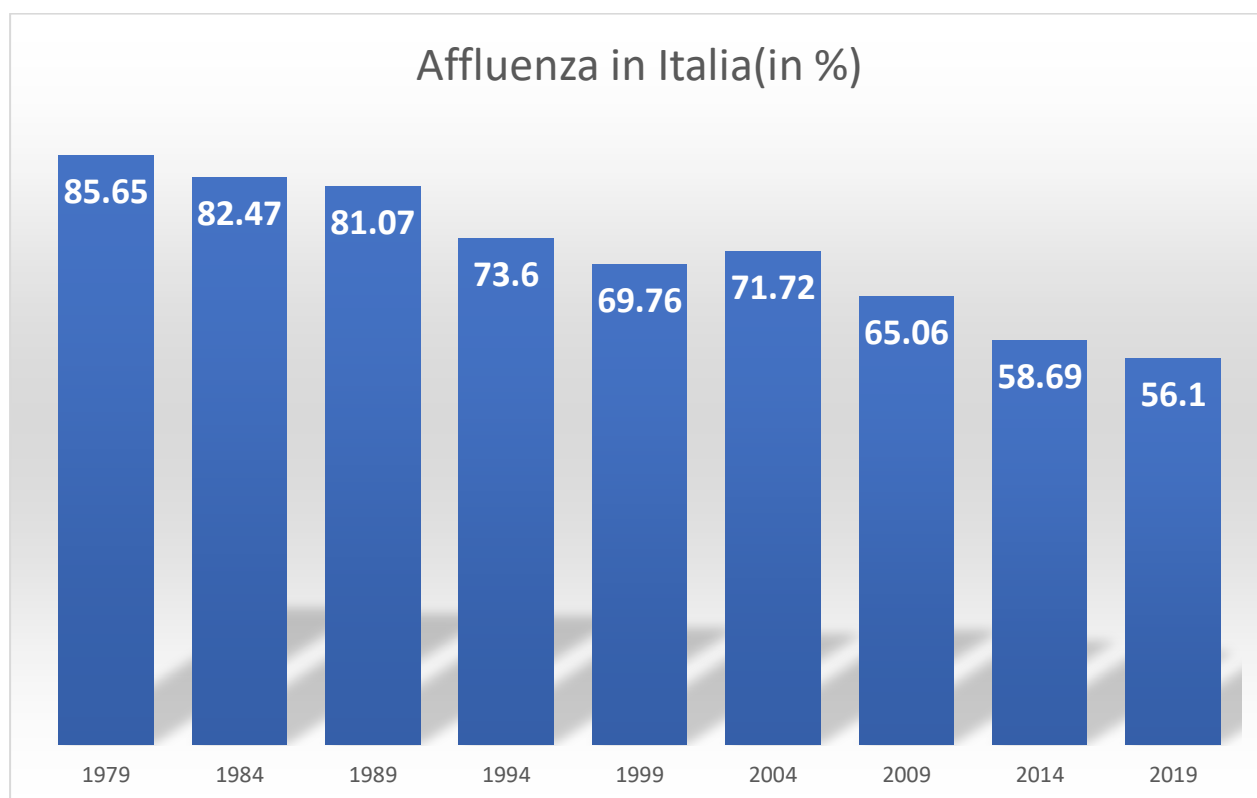
27 maggio, 2019. Articolo Online: [http://www.ansa.it/europee\\_2019/notizie/2019/05/25/italia-al-voto-alle-europee-in-gioco-il-governo\\_bc75d6b0-918f-4388-8830-2a888c7b4942.html](http://www.ansa.it/europee_2019/notizie/2019/05/25/italia-al-voto-alle-europee-in-gioco-il-governo_bc75d6b0-918f-4388-8830-2a888c7b4942.html)

<sup>10</sup> "Elezioni Europee 2019, in Italia affluenza al 56%: in calo sul 2015 soprattutto al Sud. In Ue mai così alta da 20 anni". Il Fatto Quotidiano. 27 maggio 2019. Articolo Online:

<https://www.ilfattoquotidiano.it/2019/05/27/elezioni-europee-2019-in-italia-affluenza-al-56-in-calo-sul-2014-soprattutto-al-sud-in-ue-mai-cosi-alta-da-20-anni/5208070/>

<sup>11</sup> "Risultati delle elezioni europee". Parlamento europeo. 2014. Articolo Online:

<http://www.europarl.europa.eu/elections2014-results/it/turnout.html>



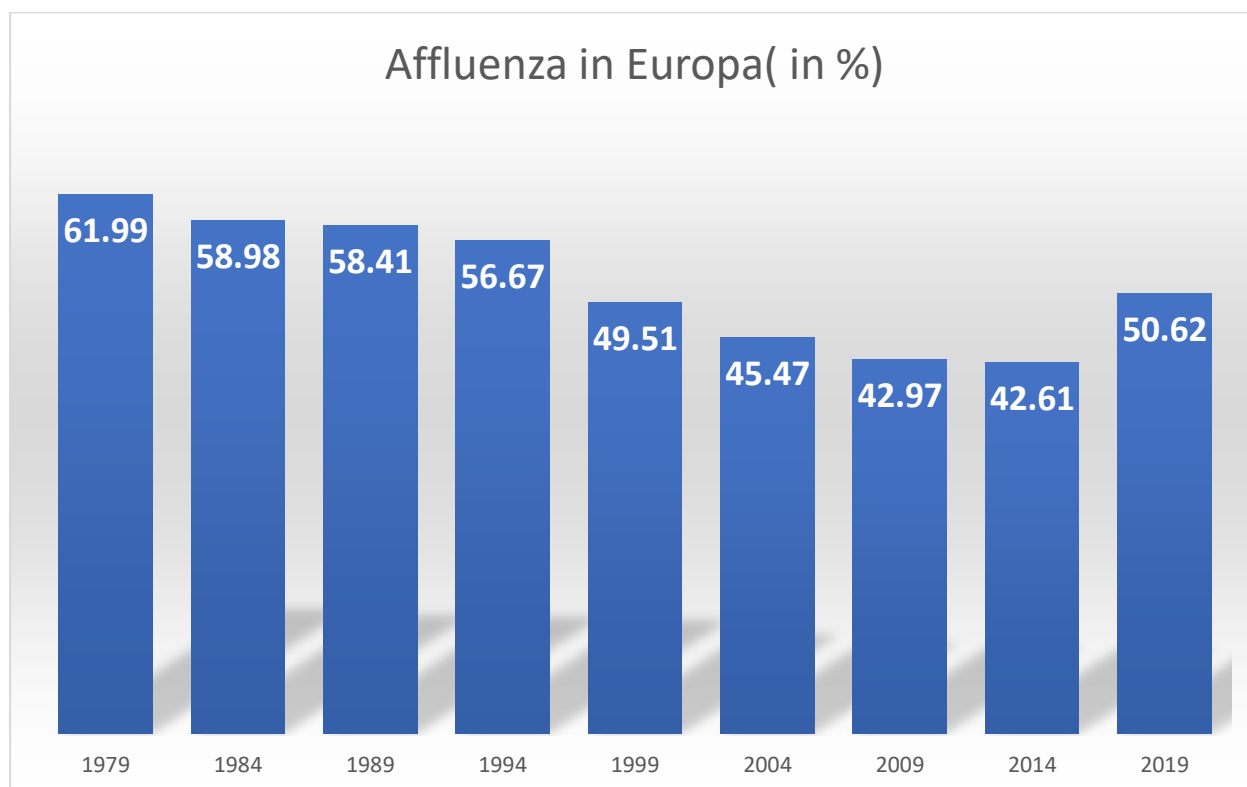
**Figura 7. Affluenza alle elezioni europee in Italia (in percentuale)**

Come si nota dal grafico la percentuale, in Italia, è drasticamente calata.

La maggior parte dei quotidiani ha inserito in prima pagina l'importanza di queste elezioni, in quanto in Europa l'affluenza è nettamente aumentata rispetto agli altri anni, infatti, non si registrava una percentuale così elevata di votanti da 20 anni. Quest'anno la percentuale di affluenza è stata del 50,62%, il paese con più votanti è stato il Belgio (89%), l'Italia si piazza al nono posto rispetto agli altri paesi europei<sup>12</sup>.

<sup>12</sup>“Europee: affluenza in netto aumento, la più alta da 20 anni”. Redazione ANSA. 27 maggio, 2019. Articolo Online: [http://www.ansa.it/europee\\_2019/notizie/2019/05/25/litalia-al-voto-alle-europee-in-gioco-il-governo\\_bc75d6b0-918f-4388-8830-2a888c7b4942.html](http://www.ansa.it/europee_2019/notizie/2019/05/25/litalia-al-voto-alle-europee-in-gioco-il-governo_bc75d6b0-918f-4388-8830-2a888c7b4942.html)

Nel grafico seguente sono indicate le percentuali di affluenza degli ultimi 40 anni<sup>13</sup>:



**Figura 8. Affluenza alle elezioni europee in Europa (in percentuale)**

<sup>13</sup> “Risultati delle elezioni europee del 2019”. Parlamento europeo in collaborazione con Kantar. 24 giugno, 2019. Articolo Online: <https://risultati-elezioni.eu/affluenza/>



### *2.2.2 Astensionismo*

In Italia il contrasto tra il tasso di interesse mostrato dalle ricerche su Google e l'affluenza può essere spiegato dal malcontento generale della popolazione italiana, che ha dimostrato di volersi informare e di cercare informazioni su come votare e su chi votare, ma che in seguito non ha espresso la propria preferenza di voto. Questo per via dei continui scandali che scoppiano sui politici ogni giorno, o forse perché non ci si vuole accontentare di votare “il meno peggio” e non un candidato in cui realmente si ripone fiducia. L'astensionismo è un fenomeno in via di sviluppo in molti paesi europei, anche in quelli in cui votare è obbligatorio. Figure politiche importanti, come Matteo Renzi e Giorgio Napolitano, hanno sostenuto la legittimità del non voto, come diritto di espressione dei cittadini.

Gianfranco Pasquino, laureato in Scienze Politiche ed ex Senatore, ha studiato le possibili cause che portano al non voto in Italia, ha individuato tre motivi principali che portano all'astensionismo<sup>14</sup>:

- I. Ha notato che i cittadini tendono a partecipare alle tornate elettorali ritenute più importanti: in genere l'affluenza alle amministrative o alle europee è molto più bassa rispetto a quella misurata alle politiche. Per questo motivo accorpare più elezioni in una sola giornata può essere una soluzione vincente,

<sup>14</sup>“Perché le persone non vanno a votare? Le cause principali dell'astensionismo”. Openpolis. 23 giugno 2016. Articolo online: <https://blog.openpolis.it/2016/06/23/perche-le-persone-non-vanno-votare-cause-astensionismo/8870>

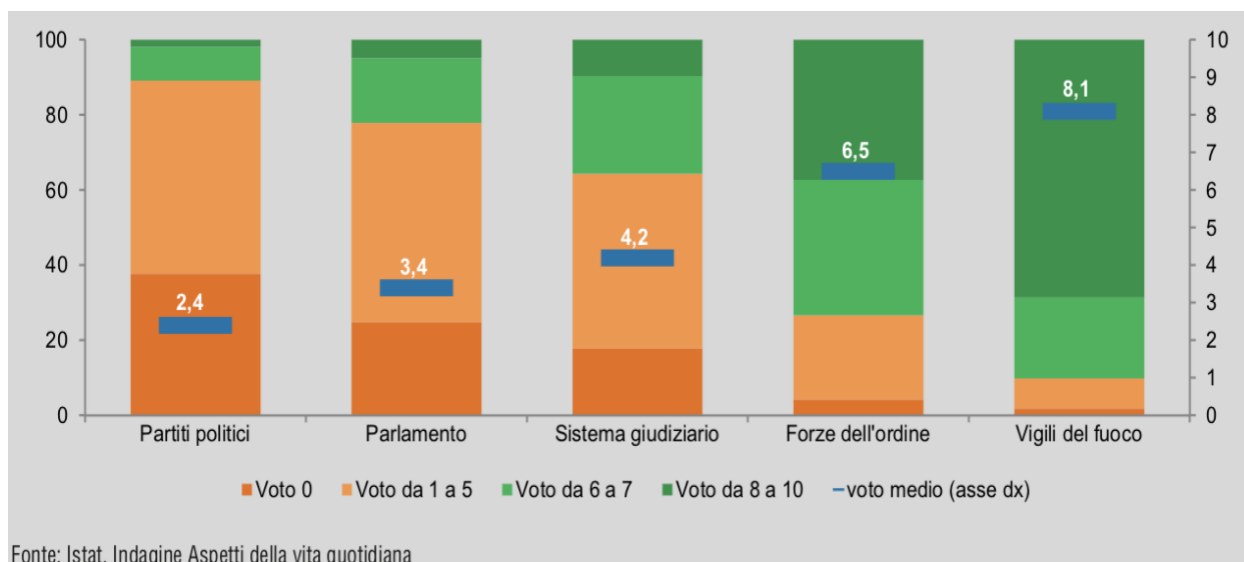
sia dal punto di vista economico, che da quello politico (es. il 4 marzo 2018, nel Lazio, oltre a votare per l'elezione delle due Camere, Senato della Repubblica e Camera dei Deputati, al seggio era presente una terza urna, per l'elezione del Presidente della Regione)

- II. Programmi molto simili tra di loro, poco chiari e definiti rendono l'elettore quasi indifferente rispetto al candidato da eleggere, la vittoria di uno o dell'altro avrebbe uno scarso impatto sulla sua vita di tutti i giorni
- III. La crisi dei partiti, che non riescono più a portare gli elettori alle urne, perché forse mossi da scarso interesse

In Italia l'opzione più influente sembra essere la terza, con una sfiducia generale nei confronti di partiti ed istituzioni. È risultato molto utile il “Rapporto sul benessere equo e sostenibile (BES)” dell'Istat<sup>15</sup>, risalente al 2018, per concretizzare queste teorie con dei dati. Il documento presenta una parte dedicata alla politica e alle istituzioni<sup>16</sup>, verso cui il paese prova una molto diffusa insoddisfazione.

<sup>15</sup> ISTAT, C. (2018). “Rapporto Bes 2018”. 18 dicembre, 2018. Online: [https://www.istat.it/it/files//2018/12/Bes\\_2018.pdf](https://www.istat.it/it/files//2018/12/Bes_2018.pdf)

<sup>16</sup> ISTAT, C. (2018). “Rapporto Bes 2018”. 18 dicembre, 2018. Pag. 85-93. Barbara Baldazzi. Online: <https://www.istat.it/it/files//2018/12/BES2018-cap-06.pdf>



**Figura 9. Voto di fiducia verso le diverse istituzioni. Anno 2017**

La fiducia mostrata verso le Forze dell'Ordine e dei Vigili del Fuoco è più che positiva:

- Il 73,3% assegna un voto maggiore di 6 alle Forze dell'ordine; voto medio: 6,5
- Il 90,2% assegna un voto maggiore di 6 ai Vigili del Fuoco; voto medio: 8,1

Molto al di sotto della sufficienza (voti calcolati da 1 a 10) invece, la fiducia per il Sistema giudiziario, i Partiti Politici e il Parlamento:

- Il Sistema Giudiziario ha meno del 40% di voti superiori al 6; voto medio: 4,2
- Il Parlamento ha poco meno del 80% di voti inferiori al 5; voto medio: 3,4
- I Partiti politici hanno circa il 90% di voti inferiori 5, con il 40% addirittura pari a 0; voto medio: 2,4.

In questo grafico<sup>17</sup> sono mostrati alcuni parametri studiati a livello regionale e in fondo ad esso, a livello nazionale.

REGIONI RIPARTIZIONI GEOGRAFICHE	Partecipazione elettorale (a)	Fiducia nel Parlamento italiano (b)	Fiducia nel Sistema giudiziario (b)	Fiducia nei partiti (b)	Fiducia in altri tipi di istituzioni (b)
	2014	2017	2017	2017	2017
Piemonte	67,4	3,2	4,1	2,3	7,4
Valle d'Aosta/Vallée d'Aoste	49,6	3,2	4,2	2,3	7,3
Liguria	60,7	3,6	4,2	2,4	7,5
Lombardia	66,4	3,4	4,0	2,4	7,4
Trentino-Alto Adige/Südtirol	52,7	3,7	4,6	3,1	7,6
<i>Bolzano/Bozen</i>	<i>52,3</i>	<i>4,0</i>	<i>4,9</i>	<i>3,8</i>	<i>7,4</i>
<i>Trento</i>	<i>53,1</i>	<i>3,3</i>	<i>4,3</i>	<i>2,4</i>	<i>7,8</i>
Veneto	63,9	2,8	3,6	2,1	7,5
Friuli-Venezia Giulia	57,6	3,5	4,3	2,5	7,6
Emilia-Romagna	70,0	3,4	4,1	2,4	7,3
Toscana	66,7	3,6	4,3	2,6	7,5
Umbria	70,5	3,4	4,1	2,3	7,3
Marche	65,6	3,0	3,7	2,1	7,1
Lazio	56,4	3,7	4,4	2,7	7,3
Abruzzo	64,1	3,3	3,9	2,2	7,5
Molise	54,8	3,1	3,9	2,2	7,0
Campania	51,1	3,6	4,4	2,6	6,8
Puglia	51,5	3,4	4,2	2,3	7,0
Basilicata	49,5	3,6	4,3	2,5	7,0
Calabria	45,8	3,6	4,8	2,7	7,1
Sicilia	42,9	3,1	4,3	2,1	6,9
Sardegna	42,0	2,9	4,3	2,0	7,4
Nord	65,4	3,3	4,0	2,3	7,4
Centro	61,8	3,6	4,3	2,6	7,4
Mezzogiorno	48,8	3,3	4,3	2,3	7,0
<b>Italia</b>	<b>58,7</b>	<b>3,4</b>	<b>4,2</b>	<b>2,4</b>	<b>7,3</b>

**Figura 10. Indicatori di partecipazione e di fiducia calcolati a livello regionale e nazionale**

<sup>17</sup> ISTAT, C. (2018). “Rapporto Bes 2018”. 18 dicembre, 2018. Pag. 85-93. Barbara Baldazzi.

Online: <https://www.istat.it/it/files//2018/12/BES2018-cap-06.pdf>

Legenda:

(a) Per 100 (%) aventi diritto

(b) Espressa su una scala da 1 a 10 da persone con più di 14 anni

Gli indicatori/parametri studiati in questa tabella sono i seguenti<sup>18</sup>:

1. Partecipazione elettorale: persone che hanno votato alle ultime elezioni del Parlamento europeo in percentuale, sul totale degli aventi diritto

Fonte: Ministero dell'Interno

2. Fiducia nel Parlamento italiano: Fiducia nel Parlamento italiano, in una scala da 0 a 10, espresso da persone con più di 14 anni

Fonte: Istat, Indagine Aspetti della vita quotidiana

3. Fiducia nel Sistema giudiziario: Fiducia nel Sistema giudiziario, in una scala da 0 a 10, espresso da persone con più di 14 anni

Fonte: Fonte: Istat, Indagine Aspetti della vita quotidiana

4. Fiducia nei Partiti: Fiducia nei partiti politici, in una scala da 0 a 10, espresso da persone con più di 14 anni

Fonte: Istat, Indagine Aspetti della vita quotidiana

<sup>18</sup> ISTAT, C. (2018). "Rapporto Bes 2018". 18 dicembre, 2018. Pag. 91. Barbara Baldazzi. Online: <https://www.istat.it/it/files//2018/12/BES2018-cap-06.pdf>

5. Fiducia in altri tipi di istituzioni: Fiducia nelle Forze dell'Ordine e nei Vigili del fuoco, in una scala da 0 a 10, espresso da persone con più di 14 anni

Fonte: Istat, Indagine Aspetti della vita quotidiana

La fiducia dei cittadini verso le istituzioni sopra citate (Parlamento, sistema giudiziario e partiti) è bassa in tutto il territorio italiano, ma raggiunge punti di minimo nel Nord.

L'Italia storicamente ha sempre mantenuto un tasso di partecipazione elettorale elevato, il crollo di quest'ultimo è avvenuto dopo lo scandalo Tangentopoli e la fine della Prima Repubblica.

Un'inchiesta del 2015 di Cmr Intesa Sanpaolo per La Stampa ha rivelato che il 52% degli italiani non si sentiva parte di alcun partito. Questo perché la maggior parte di essi pensa che i politici non si interessino ai problemi della gente e che votare sia inutile.

Sono state rilevate dai sondaggi le seguenti percentuali:

- Persone che andranno sicuramente a votare (57,3%)
- Persone che dichiarano la sicurezza del non voto (15,7%)
- Indecisi (27%)

Gli appartenenti al 15,7%, quindi gli assenteisti, hanno dichiarato i motivi per cui non esprimeranno il proprio voto<sup>19</sup>:

1. I politici non sono interessati alle problematiche dei cittadini (37,4%)
2. Votare non sarebbe utile, per via della staticità del Paese (27,5%)
3. I partiti non soddisfano le aspettative (15,2%)

L'immagine seguente spiega come, secondo Demopolis, Istituto Nazionale di Ricerche, circa 20 milioni di italiani potenzialmente si sarebbero potuti non recare alle urne lo scorso 26 maggio.



**Figura 11. Astensionismo: segni di incertezza sul voto delle elezioni europee**

<sup>19</sup> “Elezioni, l’Italia dell’astensionismo. Il 52% non si riconosce in nessun partito”. Il fatto quotidiano. 18 maggio 2015. Articolo online: [www.ilfattoquotidiano.it/2015/05/18/elezioni-litalia-dellastensionismo-il-52-non-si-riconosce-in-nessun-partito/1694292/](http://www.ilfattoquotidiano.it/2015/05/18/elezioni-litalia-dellastensionismo-il-52-non-si-riconosce-in-nessun-partito/1694292/)

È possibile parlare di disaffezione alla politica, in particolare da parte delle nuove generazioni, dovuta ad una società che promette, ma non mantiene la parola data<sup>20</sup>. L'assenza di prospettive, l'alto tasso di disoccupazione (soprattutto tra i giovani), la mancanza di investimenti sul futuro da parte dello Stato e l'instabilità della vita in generale portano alla diffusione di un'attitudine anti-sistemica, che alcuni partiti politici hanno provato a sfruttare, ponendosi come promotori "dell'onda del cambiamento". Questo è esattamente quello che ha fatto il Movimento 5 Stelle dalla sua fondazione, avvenuta circa nel 2013, infatti in soli 5 anni (4 marzo 2018) è riuscito a diventare il partito più votato in Italia, per poi nel corso di questo ultimo anno deludere le aspettative di molti elettori, che ora lo considerano come gli altri (come dimostrato dalle percentuali di consensi ottenuti il 26 maggio 2019). Nel corso degli anni quindi le persone che non si identificavano con un partito in particolare sono aumentate sensibilmente<sup>21 22</sup>, questo grafico<sup>23</sup> mostra le percentuali fino al 2008:

<sup>20</sup> Galli, G., & Capecchi, V. (1968). *Il comportamento elettorale in Italia: una indagine ecologica sulle elezioni in Italia fra il 1946 e il 1963* (Vol. 1). Bologna, Il Mulino

<sup>21</sup> Biorcio, R. (2010), *Gli antecedenti politici della scelta di voto: l'identificazione di partito e l'autocollocazione sinistra-destra*, in Bellucci, P. e Segatti, P. (a cura di), *Votare in Italia: 1968-2008. Dall'appartenenza alla scelta*, Bologna, il Mulino, pp. 187-211

<sup>22</sup> Dalton, R. e Watterberg, M. (2000), *Parties without Partisans. Political Change in Advanced Industrial Democracies*, Oxford, Oxford University Press

<sup>23</sup> "La Partecipazione elettorale in Italia". Quaderni di sociologia pag. 17 -39. Maurizio Cerruto. 2012. Articolo online: <https://journals.openedition.org/qds/537>



	1968	1972	1975	1990	2001	2006	2008
Vicino: molto o abbastanza	77,8	64,5	56,8	25,8	33,5	38,5	23,3
Vicino ma solo simpatizzante	5,4	7,0	12,0	23,5	21,3	16,9	27,9
Non vicino	16,8	28,5	31,2	50,7	45,2	44,6	48,8

**Figura 12. Individui che si identificano in un partito (1968 – 2008)**

In aggiunta a questa disaffezione generale c'è la nuova tendenza ad essere politicamente passivi<sup>24</sup>, non essere coinvolti con le istituzioni politiche, una sorta di presenza-assenza (dovuta alla debole identificazione nei partiti) che può spiegare in parte l'apatia degli elettori<sup>25</sup>.

Secondo Maurizio Cerruto, sociologo presso l'Università di Cagliari e autore dello studio “La partecipazione elettorale in Italia”<sup>26</sup>, esistono due tipi di astensionismo:

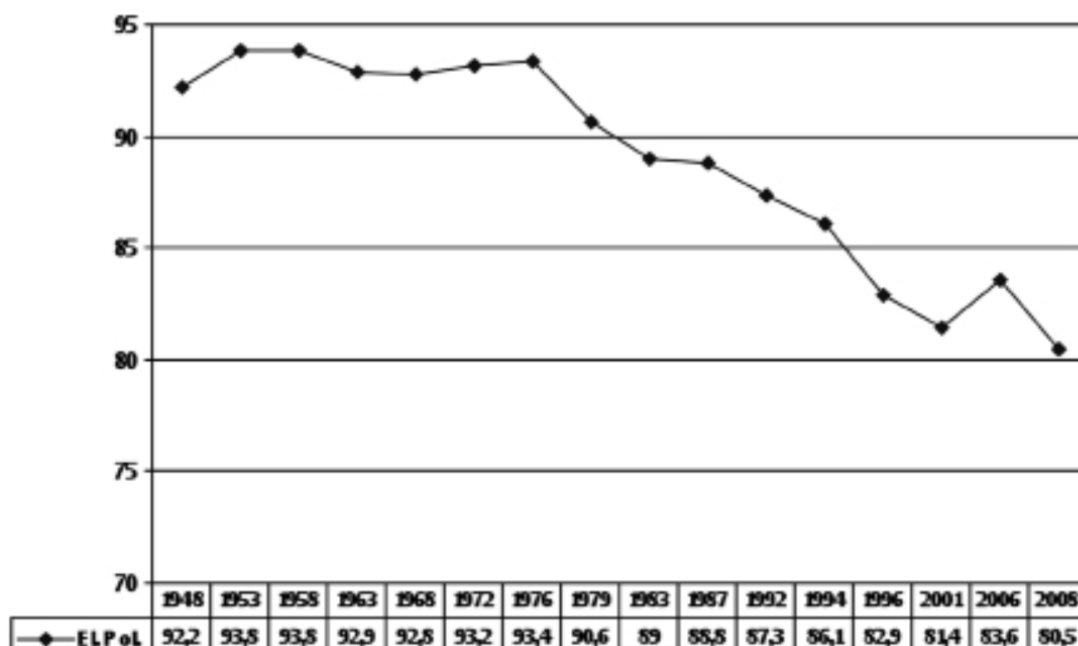
1. Astensionismo da apatia, che si concretizza nella distanza tra l'elettore e l'offerta politica. Questo tipo ha le sue origini nella posizione poco centrale che la politica ha nelle menti degli elettori
2. Astensionismo di protesta, come espressione di un'insoddisfazione da parte dell'elettore, che manifesta sfiducia e in alcuni casi ostilità verso le politiche attuali

<sup>24</sup> Agosta, A. (2006), *Come leggere la partecipazione elettorale*, «il Mulino», 55, 3, pp. 465-469

<sup>25</sup> Tuorto, D. (2006), *Apatia o protesta? L'astensionismo elettorale in Italia*, Bologna, il Mulino

<sup>26</sup> “La Partecipazione elettorale in Italia”. Quaderni di sociologia pag. 17 -39. Maurizio Cerruto. 2012. Articolo online: <https://journals.openedition.org/qds/537>

Il fenomeno più diffuso è chiaramente il primo, che denota la marginalità che la politica ha nella vita dell'elettore.



**Figura 13.. Partecipazione (in percentuale) alle Elezione Camera dei Deputati (1948 – 2008)**

Come detto prima, l'interesse per le elezioni della Camera dei Deputati e del Senato della Repubblica è molto superiore rispetto a quello per le altre votazioni, questo grafico (figura 13) serve solo a far capire come negli anni l'assenteismo sia aumentato sensibilmente<sup>27</sup>.

Il divario tra le generazioni elettorali diventa più evidente, se oltre alle elezioni politiche si considerano le percentuali di astensionismo delle europee. Tra il 1979 e il 2009 il calo è stato di circa il 20%, da quasi l'86% nel 1979 a 65% nel 2009. Si può notare che è molto aumentato anche il divario del livello partecipazione tra le elezioni

<sup>27</sup> Caciagli, M. e Spreafico, A. (a cura di) (1990), *Vent'anni di elezioni in Italia. 1968-1987*, Padova, Liviana

politiche, considerate le più importanti, e di quelle “secondarie” (europee, regionali, referendum, amministrative).

Quando si parla di voti inespressi però, non si considerano solamente gli elettori che non si sono recati alle urne, ma si prendono in considerazione anche coloro che hanno espresso voto nullo o non hanno espresso alcuna preferenza (scheda bianca). In questo caso si parla di astensionismo ampio, o attivo, perché si considerano dei cittadini che hanno espresso la propria volontà di votare, ma che non sono soddisfatti o sono sfiduciati dai candidati e dai partiti. Come si nota nel grafico sottostante <sup>28</sup> le percentuali di astensionismo sono fortemente influenzate dai voti nulli e dalle schede bianche:

<sup>28</sup> “La Partecipazione elettorale in Italia”. Quaderni di sociologia pag. 17 -39. Maurizio Cerruto. 2012. Articolo online: <https://journals.openedition.org/qds/537>

	1992	1994	1996	2001	2006	2008	Media
<i>Partecipazione</i>							
Nord	91,1	90,7	87,7	84,8	86,6	83,2	87,3
Centro	92,1	91,4	89,0	87,0	88,1	84,6	88,7
Sud	82,2	80,3	76,5	76,4	79,4	76,7	78,6
Italia	87,3	86,1	82,9	81,4	83,6	80,5	83,4
<i>Astensionismo</i>							
Nord	8,9	9,3	12,3	15,2	13,4	16,8	12,6
Centro	7,9	8,6	11,0	13,0	11,9	15,4	11,3
Sud	17,8	19,7	23,5	23,6	20,6	23,3	21,4
Italia	12,7	13,9	17,1	18,6	16,4	19,5	16,4
<i>Voti nulli</i>							
Nord	2,6	2,5	3,4	2,9	1,6	2,1	2,5
Centro	2,5	2,3	2,6	2,2	1,5	2,0	2,2
Sud	3,9	4,9	5,6	3,8	2,1	3,0	3,8
Italia	2,8	3,5	4,1	3,2	1,8	2,5	3,0
<i>Schede bianche</i>							
Nord	1,9	2,3	2,6	3,0	0,9	0,8	1,9
Centro	2,2	2,2	2,3	2,8	1,0	0,9	1,9
Sud	2,3	5,0	3,9	5,9	1,4	1,9	3,4
Italia	1,9	3,4	3,1	4,2	1,1	1,3	2,5
<i>Totale voti inespressi</i>							
Nord	13,4	14,1	18,2	21,1	16,0	19,8	17,1
Centro	12,6	13,2	15,9	18,0	14,3	18,3	15,4
Sud	24,0	29,6	32,9	33,3	24,1	28,2	28,7
Italia	17,4	20,8	24,3	26,0	19,3	23,2	21,8

**Figura 14. Percentuali di partecipazione e di voto inespresso – Elezione Camera dei Deputati (1992 – 2008)**

Si può, tra l'altro, notare un dualismo molto evidente tra Nord e Sud Italia riguardo l'astensionismo in senso ampio<sup>29 30</sup>, quindi compreso di voti nulli o inespressi (scheda bianca). La percentuale media degli astenuti in Italia è del 16,4%, valore influenzato positivamente dal Sud che spinge questo valore verso l'alto, con il suo 21,4% di assenti al voto, e negativamente dal Centro e dal Nord, rispettivamente con 11,3% e 12,6%.

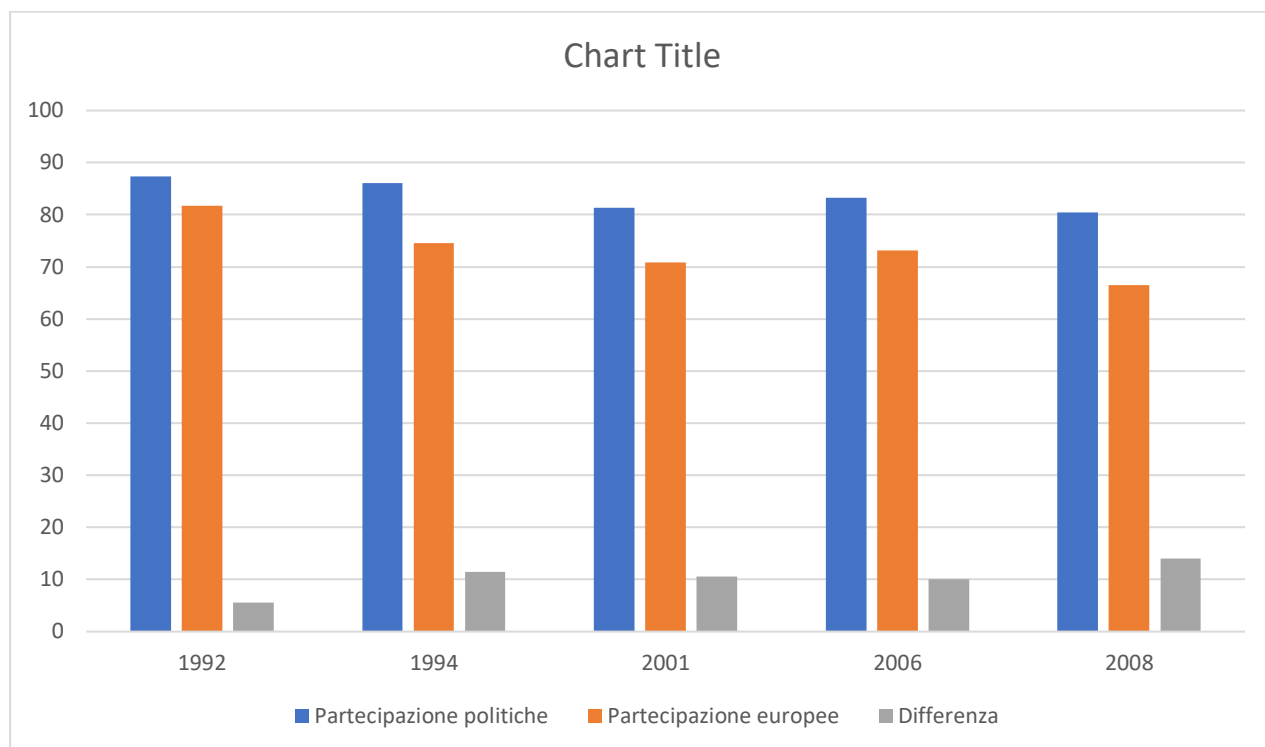
Si può notare come il Sud sia in testa anche in termini di voti nulli (3,8%) e schede bianche (3,4%), mentre il Centro e il Nord siano sotto di circa 1,5 punti percentuale rispetto al meridione.

Provando a confrontare l'affluenza delle elezioni politiche e di quelle europee, si nota subito che le seconde sono considerate di "seconda classe" da molti italiani<sup>31</sup>, non solo l'affluenza negli anni è diminuita in entrambi i casi, ma il divario tra le due è aumentato notevolmente con il passare del tempo.

<sup>29</sup> Fabrizio, D., Feltrin, P. e Marcone, L. (2006), *La decisività del voto nelle regioni meridionali: un'analisi degli orientamenti elettorali negli ultimi dieci anni*, «Polena», 3, 2, pp. 9-35

<sup>30</sup> Nuvoli, P. (1989), *Il dualismo elettorale Nord-Sud in Italia: persistenza o progressiva riduzione?*, «Quaderni dell'osservatorio elettorale»

<sup>31</sup> Corbetta, P., Parisi, A. e Schadee, H. (1996), *Elezioni in Italia. Struttura e tipologia delle consultazioni politiche*, Bologna, il Mulino



**Figura 15. Percentuali votanti alle politiche, alle europee e differenza tra le due**

Le elezioni europee in particolare hanno sempre registrato un basso tasso di partecipazione, forse per via di una distanza troppo grande percepita tra le istituzioni europee e la vita di tutti i giorni.

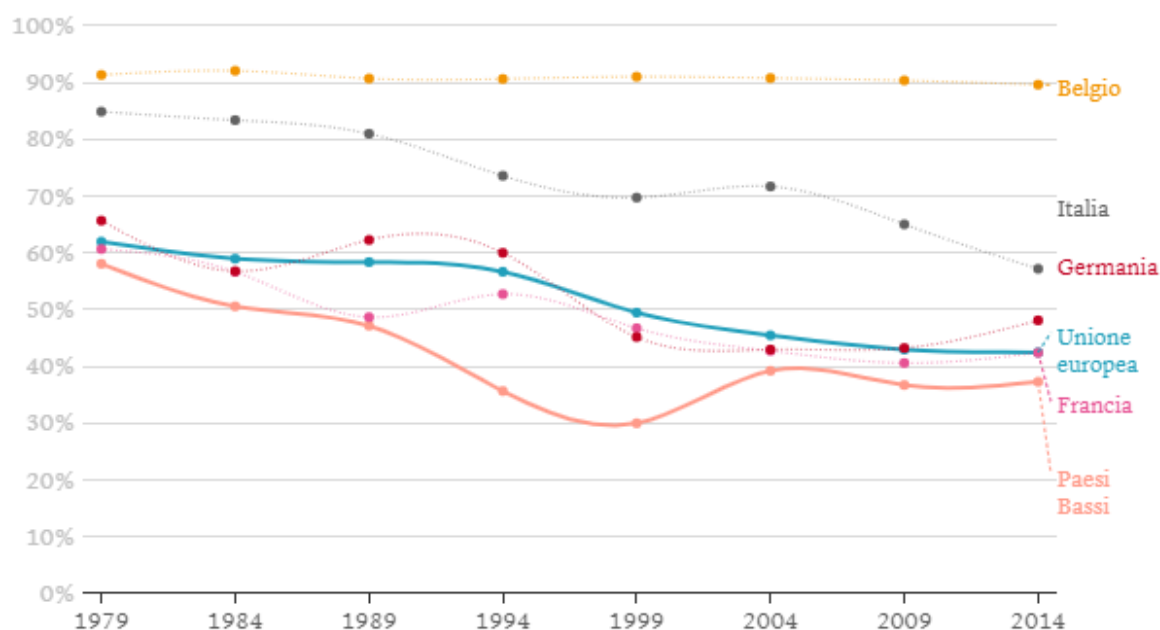
### ***2.2.3 Standard Eurobarometer***

Lo standard Eurobarometer, inserito nel 1974, si occupa di svolgere ed analizzare delle interviste anonime poste ai cittadini europei, per avere la loro opinione sulle istituzioni. Ogni survey conta circa 1000 interviste per ogni paese, effettuate faccia a faccia. L'Eurobarometer ha rivelato che solo il 48% dei cittadini europei pensa che la propria voce abbia un peso all'interno dell'Unione stessa, ovviamente questo dato va

considerato con le dovute differenze tra i vari paesi (in Svezia il 90% dei cittadini pensa che il proprio voto conti, mentre in Italia solo il 24%).

Questo, e tutto ciò di cui si è discusso prima, è in grado di spiegare il perché in alcuni paesi l'affluenza sia inferiore rispetto ad altri.

### L'affluenza alle elezioni europee



**Figura 16. Affluenza internazionale alle elezioni europee**

L'Italia ha comunque una media di affluenza alle elezioni europee superiore alla media in tutta Europa. Il Belgio è il primo paese in assoluto e la sua curva di affluenza è rimasta più o meno costante negli anni.

In qualunque caso non si tratta solo di fiducia che il singolo cittadino ha nei confronti delle istituzioni europee. L'Eurobarometer mostra che in media i cittadini europei tendono ad avere più fiducia nel Parlamento dell'Unione che in quello del proprio Stato. Detto questo, perché allora solo una parte della popolazione vota alle elezioni

europee? Questo paradosso può essere spiegato grazie a due fattori, come uno studioso delle politiche europee, Alberto Alemanno, afferma.

Il primo esplicita il concetto secondo cui le elezioni europee non siano altro che la “somma” di quelle nazionali, e non un evento politico transnazionale animato da reali sentimenti europei. Il secondo, invece, espone l’idea per cui non esista una sfera comune pubblica in cui si possa parlare del sistema politico europeo.

In qualunque caso, l’affluenza delle elezioni europee è molto difficile da analizzare nel corso del tempo, perché, come nota Jules Baley dell’università SciencesPo di Parigi, non è possibile confrontare la percentuale di partecipazione attuale (28 Stati con diverse culture e tradizioni democratiche) con quella del 1979, quando la Comunità Europea era formata solo da nove Paesi, per di più tutti appartenenti all’Europa del Ovest. Nel momento in cui si parla dell’affluenza alle elezioni europee, bisogna considerare il fatto che alcuni cittadini abbiano problemi a votare perché magari si sono trasferiti in un altro Paese membro.

Più del 10% di rumeni, bulgari, croati, lettoni e portoghesi vive in uno Stato membro, che non corrisponde a quello di nascita. Ovviamente hanno il diritto di votare nella città in cui vivono, ma nel 2014, il 95% di essi ha dichiarato di non esser andato a votare per via di problemi legati alla lingua, all’indecisione riguardo la lista da votare, alla burocrazia del Paese in cui vive e ad ostacoli politici<sup>32</sup>.

<sup>32</sup>Ottaviani, Jacopo and Ferrari, Lorenzo. What happens with abstentionism in Europe. Internazionale, May 16, 2019. Online article: <https://www.internazionale.it/notizie/jacopo-ottaviani/2019/05/16/elezioni-europee-astensionismo>



La creazione di un solo Collegio elettorale europeo permetterebbe di presentare delle liste e dei partiti transnazionali, e questo potrebbe essere una soluzione all'astensionismo legato alla problematica che riscontrano i cittadini che, cambiando paese e non conoscendo lo scenario politico e i partiti del luogo di residenza, non si recano a votare. In questo modo i partiti sarebbero spinti a concentrarsi di più su argomenti di interesse europeo, e non solo nazionale, come invece fanno adesso. Forse in questa maniera i dibattiti politici potranno coinvolgere più profondamente gli elettori, che così andranno a votare.

## **Capitolo 3:**

### **Predire i risultati delle Elezioni Europee del 26 maggio 2019**

I capitoli precedenti contengono solo un brevissimo riassunto di quello che S.S. Davidowitz ha spiegato nel suo libro, ma dovrebbe essere sufficiente per comprendere le motivazioni che mi hanno spinto ad intraprendere questa ricerca e il metodo che ho utilizzato nel provare a predire i risultati delle elezioni Europee del 26 maggio 2019. Insieme al mio team abbiamo provato a riprodurre e, ad adattare al nostro caso, lo studio che Davidowitz aveva fatto con le elezioni presidenziali in America. Davidowitz, studiando, ha quindi capito che Google possiede e può fornire molte informazioni che i sondaggi non riescono a cogliere.

Più del 50% dei cittadini che non votano dichiarano, prima delle elezioni, nei questionari o nelle interviste anonime che andranno a votare, dando così alle stime dei sondaggi una visione molto distante da quella che è la realtà. Invece, le ricerche effettuate su Google “come votare?” o “dove votare?”, o quesiti che riguardano informazioni sul voto in generale, permettono di avere una visione molto più precisa dell’affluenza ai seggi, dividendoli anche secondo alcune categorie, per esempio le Regioni o gli Stati.

### **3.1 Strumenti utilizzati**

Per questa ricerca gli strumenti utilizzati sono stati Google Trends e Google Correlate, che permettono di studiare grandi quantitativi di dati. Il Data Mining è il processo che permette di analizzare una grande quantità di dati per trovare eventuali correlazioni tra di loro. Esso è un processo di selezione ed analisi dei dati che porta all'estrazione di informazioni più o meno utili, attraverso metodi automatici o semi-automatici (es. machine learning). Questo processo viene utilizzato per cercare correlazioni tra più variabili, relativamente ai singoli elementi; in pratica è l'analisi, da un punto di vista matematico, eseguita su database di grandi dimensioni. L'aumento della quantità di dati disponibile e lo sviluppo tecnologico, che ha portato alla crescita delle capacità di calcolo dei computer, sono i due fattori principali che hanno portato al potenziamento di nuove metodologie per l'analisi dei dati.

Nella società iper-connessa in cui viviamo oggi il quantitativo dei dati disponibili è enorme, quindi non solo serve un algoritmo capace di estrarre le informazioni significative dalle varie sorgenti a disposizione, ma una volta estratti, questi dati devono essere analizzati coerentemente alle loro dimensioni. Il Data mining ci aiuta nella fase di estrapolazione; il processo di analisi dei dati si divide in sei fasi:

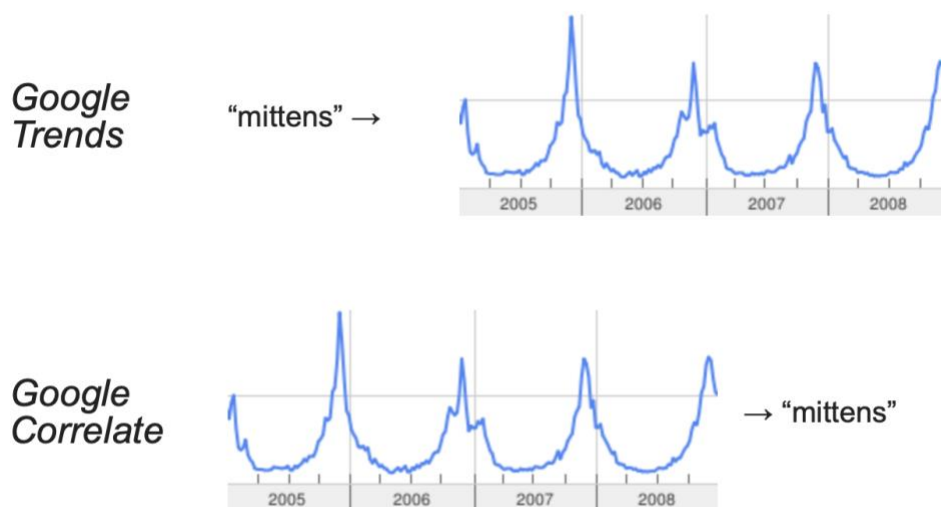
- a) Identificazione e selezione dei dati
- b) Acquisizione dei dati
- c) Trasformazione dei dati
- d) Studio e comprensione dei dati

e) Analisi dei risultati

f) Valutazioni finali

Uno strumento che permette di fare queste analisi, ma che le rende più semplici e comprensibili, è stato lanciato da Google e si tratta proprio di Google Correlate, esso permette di scoprire i trend più correlati tra di loro. Google Correlate analizza gli andamenti del numero di ricerche di un determinato argomento, o parola chiave, e li confronta con quelli cercati dall'utente.

Inoltre, permette, attraverso l'inserimento di un file "csv" di Excel la ricerca dei trend più correlati con i dati del file caricato. Con Google Correlate si inserisce una serie di dati (target) e si ricevono indietro *queries* la cui frequenza segue un pattern simile.



**Figura 17. Differenze tra Google Trends e Google Correlate<sup>33</sup>**

<sup>33</sup> "Google Correlate Tutorial". Online: <https://www.google.com/trends/correlate/tutorial>

Quando si inserisce un data set, Google Correlate calcola l'indice di correlazione di Pearson, indicato con "r" (anche detto coefficiente di correlazione lineare), che esprime un'eventuale relazione lineare tra due variabili statistiche. L'intervallo dell'indice di correlazione varia da  $r = -1.0$  a  $r = +1.0$ . I risultati che Google Correlate mostra sono quelli con il più alto valore di questo indice, quindi il più possibile vicino a  $+1.0$ .

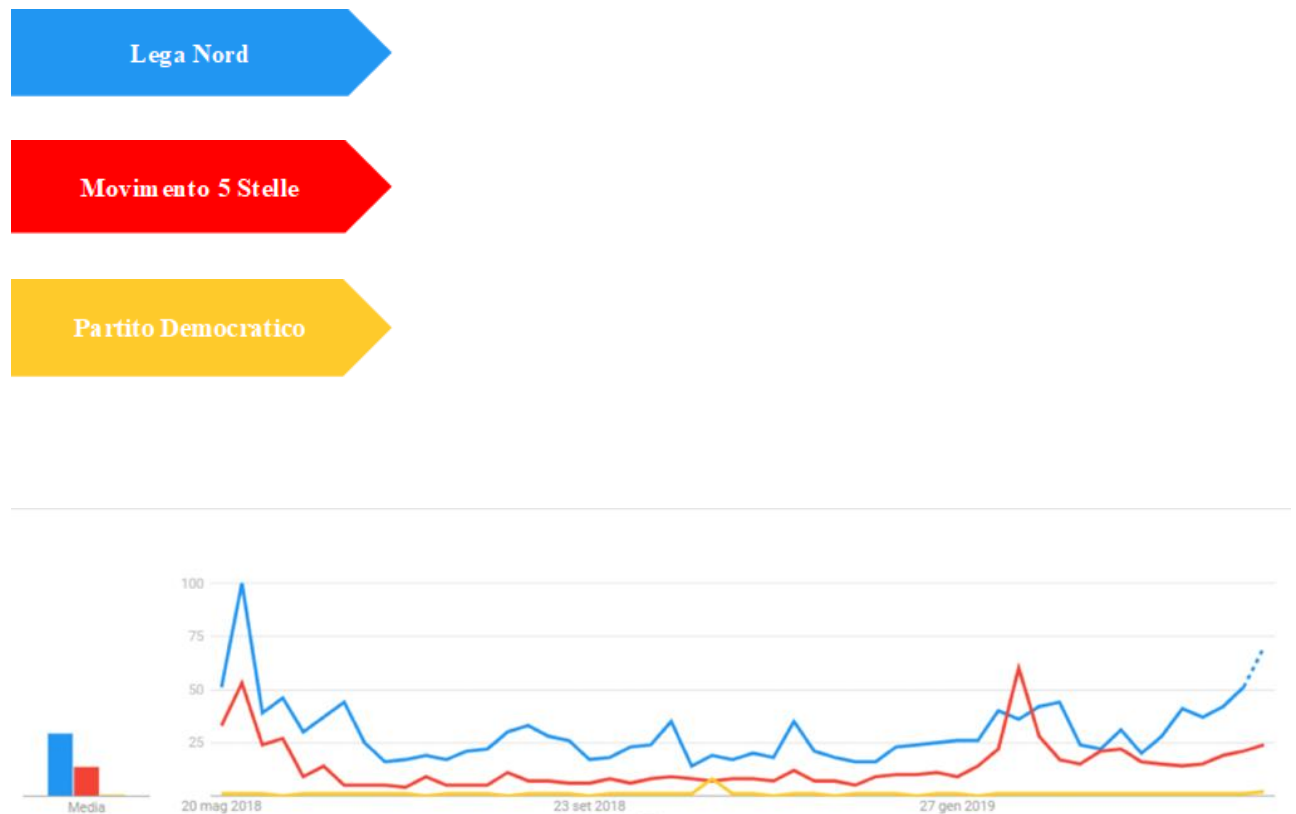
Questo strumento ha quindi aiutato nel capire quali fossero gli argomenti/termini più correlati agli ambiti di nostro interesse.

Google Trends, invece è come se fosse, l'inverso di Google Correlate, si digita una *query* e si ottiene la frequenza con cui essa viene cercata dagli utenti, anche in questo caso si può selezionare il periodo, l'area geografica e la categoria a cui appartiene l'argomento. Google Trends, è uno strumento che permette di calcolare la frequenza delle ricerche su Google di una parola o di uno specifico argomento. I trends vengono mostrati con un grafico che sintetizza, nel tempo, la sua popolarità. Importante sapere che i dati forniti non sono assoluti, quindi non corrispondono ad un numero preciso di utenti, ma sono relativi e calcolati in percentuale sul totale delle ricerche, che comunque non è noto, quindi si può trattare di qualche decina di migliaia di quesiti posti a Google, come di milioni. Vi sono due sezioni nella schermata di Trends: la prima è "researches' tendency" in cui è possibile visualizzare i trend più ricercati del momento, invece nella seconda è presente la classifica delle ricerche più cercate in assoluto. Nel 2018 la sua grafica è stata aggiornata per essere più moderna e più facilmente utilizzabile da tutti gli utenti, non solo da quelli più esperti.

### 3.2 Punto di partenza

Utilizzando Google Trends ho cercato di capire l'andamento delle ricerche dei tre principali partiti politici in Italia.

Il grafico seguente mostra proprio il trend delle ricerche per Lega Nord, Movimento 5 Stelle e Partito Democratico, da maggio 2018 fino a poco prima delle elezioni europee, quindi intorno al 20 maggio 2019.

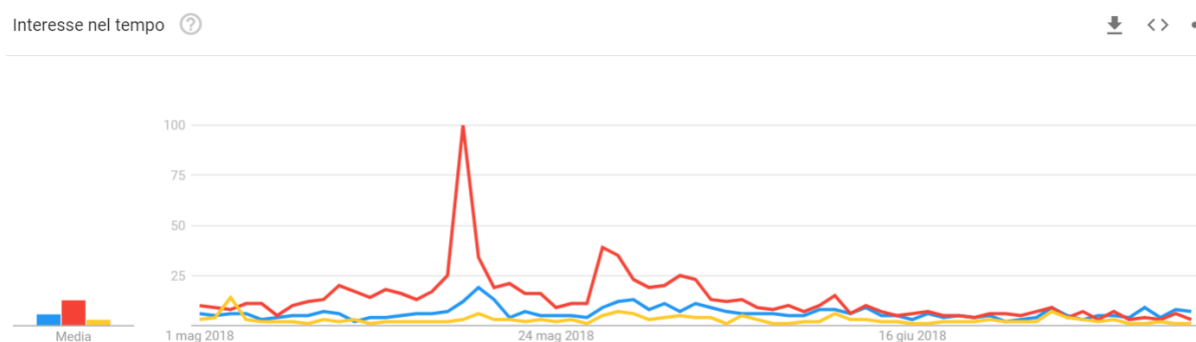


**Figura 18. Trend politici dei 3 principali partiti italiani – Google Trends**

In basso a sinistra del grafico si può notare la media totale delle ricerche effettuate per ogni termine inserito, è evidente come Lega Nord sia risultata in media il più cercato.

Il Movimento 5 Stelle, risultando il secondo più cercato, ha aperto una serie di domande su come le percentuali di consensi fossero cambiate nell'ultimo anno di Governo.

Se prima il Movimento 5 Stelle, non solo era il partito più votato, ma anche quello più cercato online, come dimostra il grafico qui sotto (figura 19), adesso si può notare proprio il sorpasso della Lega Nord su di esso, nel numero di ricerche effettuate.



**Figura 19. Trend politici dal 1 maggio 2018 a luglio 2019**

Nel 2018 il primo partito per ricerche in Italia, come viene suggerito anche dai risultati elettorali, è proprio il Movimento 5 Stelle.

### 3.3 Usare Google Trends

È realmente possibile predire quale candidato/partito gli elettori voteranno in base solo alle loro ricerche su Google?

S.S. Davidowitz ha dimostrato che non possiamo limitarci a vedere chi è stato cercato di più su Google, perché il candidato viene cercato sia da persone che lo appoggiano, sia il contrario, perché gli utenti magari sono in totale disaccordo con le sue proposte, o addirittura solo per avere notizie di gossip.

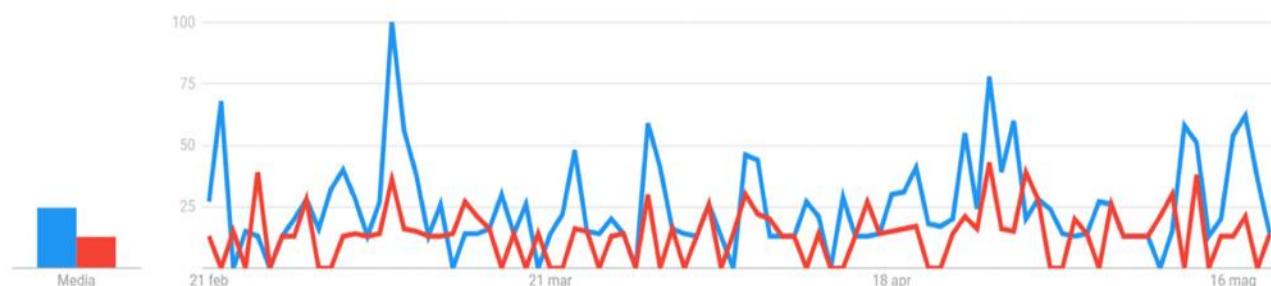
Quindi Davidowitz e Stuart Gabriel, un professore di finanza alla California University, hanno scoperto (analizzando i dati forniti dalle elezioni presidenziali americane) un indicatore importante che collega il modo in cui le persone cercano informazioni sui candidati su Google e il modo in cui poi realmente voteranno. Un grande quantitativo di ricerche riguardanti le elezioni contiene entrambi i nomi dei candidati: durante le presidenziali in America, il 12 % delle ricerche che conteneva la parola “Trump” includeva anche “Clinton” e più di un quarto delle ricerche che contenevano il termine “Clinton” includevano anche il nome dell’attuale presidente degli Stati Uniti. Queste informazioni possono darci degli indicatori su chi sia il candidato realmente sostenuto da una persona, è sufficiente studiare la sequenza in cui i nomi compaiono nella ricerca.

Lo studio di S.S. Davidowitz e S. Gabriel mostra che il modo in cui le persone inseriscono le parole da ricercare online non è casuale: in una ricerca con entrambi i nomi dei candidati, gli elettori inseriscono come primo termine il nome di quello che sostengono/preferiscono. Hanno dimostrato che, confermato dalle tre ultime elezioni americane, il nuovo Presidente della Casa Bianca era sempre colui che appariva per primo nel maggior numero di ricerche. Questa scoperta può darci informazioni che altrimenti non avremmo mai ottenuto.

Uno dei primi metodi utilizzati nel provare a predire i risultati delle elezioni europee è stato proprio questo: comparare le ricerche. Comparando le *queries* che includevano i nomi dei leader dei due principali partiti italiani (Matteo Salvini per la Lega Nord e Luigi Di Maio per il Movimento 5 Stelle) sono apparsi questi risultati:



- “Salvini Di Maio”
- “Di Maio Salvini”



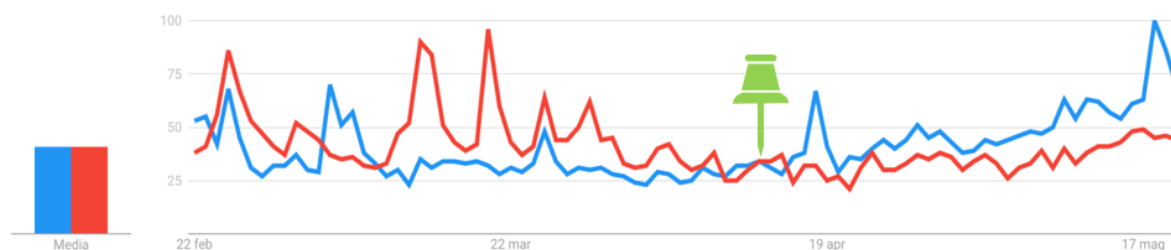
**Figura 20. Ricerche comparate – Google Trends**

Secondo il metodo identificato da S.S. Davidowitz e S. Gabriel non ci sono dubbi su quale dei due leader sarebbe uscito vincitore dalle elezioni europee, Matteo Salvini, perché le ricerche con il suo nome digitato prima di quello del rivale erano in numero superiore.

### 3.4 Turning point

Come già dimostrato, Salvini ed il suo partito non sono sempre stati i preferiti dagli italiani: Google Trends ha mostrato chiaramente un “turning point” in cui il trend è cambiato, passando da una preferenza generale per Di Maio a una per Salvini.

- “Salvini”
- “Di Maio”



**Figura 21. Turning point: Cambio del trend delle ricerche- Google Trends**

L'immagine seguente (figura 22) mostra, invece, regione per regione, le ricerche effettuate durante il periodo delle ultime elezioni politiche (4 marzo 2018). È evidente come il partito maggiormente cercato fosse il Movimento 5 Stelle, senza alcuna eccezione su tutto il territorio.

● Lega Nord ● Movimento 5 Stelle



**Figura 22. Ricerche su Google durante il periodo delle ultime elezioni politiche (3 marzo 2018) – Google Trends**

Quest'altra immagine (figura 23), simile alla precedente ma con colori opposti, mostra invece il momento in cui è cambiato il trend, quindi il periodo più vicino alle elezioni europee (26 maggio 2019). È evidente come in questo caso il partito più ricercato dagli italiani sia la Lega Nord.

● Lega Nord ● Movimento 5 Stelle



**Figura 23. Ricerche su Google durante il periodo delle ultime elezioni europee (26 maggio 2019)**

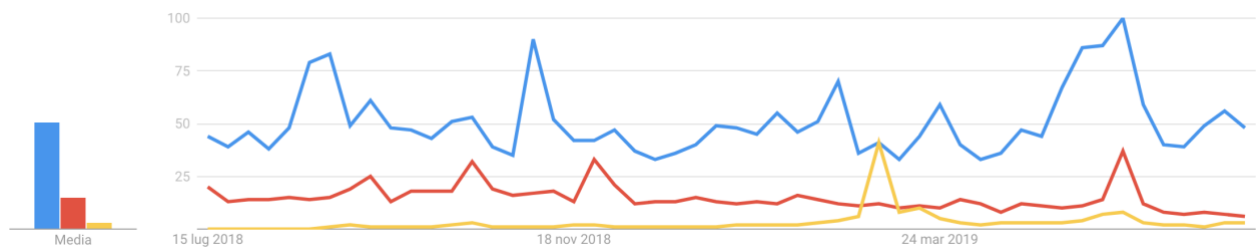
– Google Trends

### **3.5 Situazione politica in Italia prima delle elezioni europee mostrata da Google Trends**



**Figura 24. Leader dei tre principali Partiti italiani**

In questo grafico sono mostrati i trend delle ricerche che comprendono i nomi dei leader dei tre principali partiti.



**Figura 25. Trend delle ricerche dei nomi dei leader dei tre principali Partiti italiani**

È evidente la superiore popolarità di Matteo Salvini rispetto a Luigi di Maio e Nicola Zingaretti, che resta sempre il più cercato, tranne per un picco del capogruppo del Partito Democratico, intorno al 3 marzo 2019, giorno in cui si sono tenute le primarie del gruppo politico. Ovviamente il numero di ricerche non indica per forza il consenso ottenuto, si potrebbe cercare una persona anche solo per avere notizie sulle novità del Paese o per trovare delle barzellette o deridere lo stesso.

Fermo restando questo, la popolarità di Salvini è risultata comunque un indicatore del suo possibile successo alle elezioni europee.

### **3.6 Popolarità di Matteo Salvini**

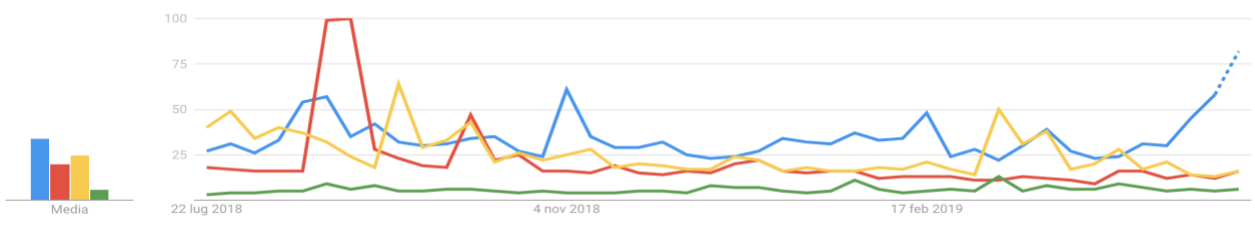
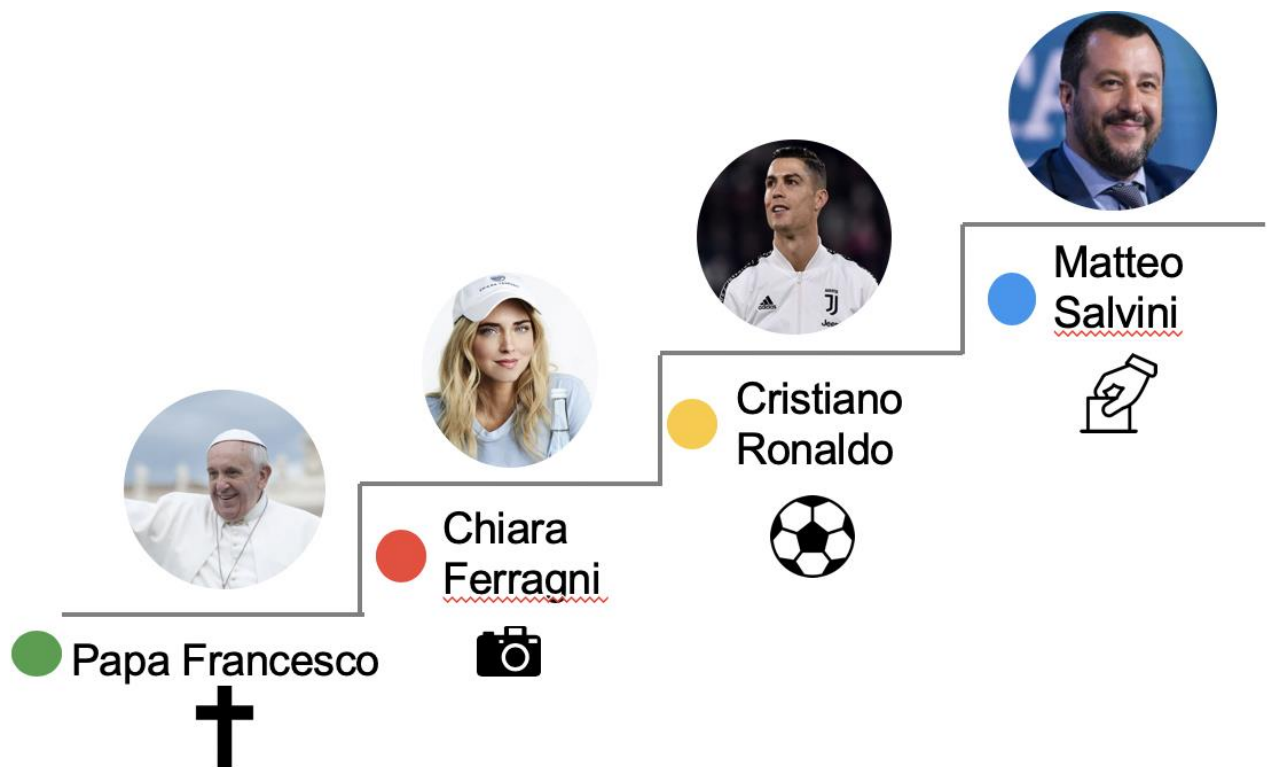
Il successo della figura di Matteo Salvini si può dedurre anche dalla sua popolarità sui social media: dopo le elezioni del 4 marzo 2018, il leader della Lega Nord ha acquisito 3 milioni di followers in più su Facebook, superando in numero Angela Merkel, primo ministro tedesco (2.5 milioni), Marine Le Pen, presidente del Front National (1.5 milioni) e tutti gli altri leader europei. Anche Donald Trump, che conta 22 milioni di followers, viene battuto da Salvini in termini di engagement e involvement con la sua

audience: 2.6 milioni in una settimana contro 1 milioni e mezzo, considerando anche che la popolazione americana è 5 volte quella italiana.

In Italia nell'ottobre 2018 Salvini è risultato essere il più seguito su Facebook (con 3.233.000 milioni di followers), seguito da Luigi Di Maio (2.100.000 milioni); un ottimo risultato, considerando, che solo a luglio dell'anno prima era poco conosciuto dal pubblico social.

Per dimostrare la popolarità di Salvini proviamo ad utilizzare Google Trends e a confrontare il numero di ricerche contenenti il suo nome con quelle su altre tre persone molto famose in Italia:

- Cristiano Ronaldo: calciatore che gioca nella Juventus e capitano della nazionale di calcio portoghese
- Chiara Ferragni: fashion blogger, influencer, imprenditrice digitale, modella e proprietaria del brand che porta il suo nome. Nel 2017, inoltre, è stata nominata da Forbes “the most important fashion influencer of the world”
- Papa Francesco: 266° papa della Chiesa cattolica e Vescovo di Roma



**Figura 26. Popolarità di Matteo Salvini paragonata a quella di altri personaggi famosi**

Come si nota dal grafico in cui compaiono le medie delle ricerche, in basso a sinistra, Matteo Salvini è risultato essere il più cercato tra tutti. L'unico punto in cui viene ampiamente superato è quello in cui la fashion blogger, Chiara Ferragni, con 16,8 milioni di followers su Instagram, si è sposata con il cantante e influencer Fedez. Avendo i due reso il matrimonio uno dei più social di tutti i tempi era prevedibile che il numero di ricerche in quel periodo aumentasse a dismisura. Dalle ricerche e dai vari

tentativi effettuati su Google Trends sembra che nessun altro personaggio in Italia riesca a battere la popolarità del Ministro Matteo Salvini.

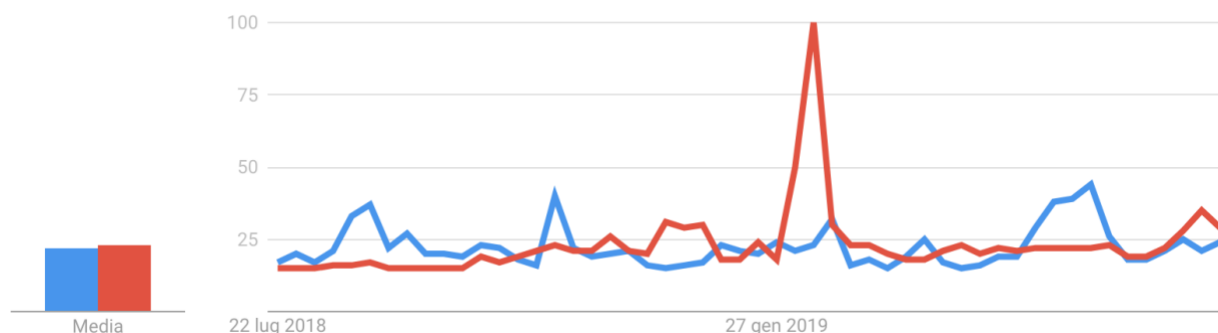
### 3.7 Fun Facts

Di seguito verranno analizzate alcune correlazioni abbastanza particolari trovate con Google Trends:

- Matteo Salvini e Ultimo
- Matteo Salvini e Weather today

#### 3.7.1 Matteo Salvini e Ultimo

Confrontando su Google Trends la popolarità di Matteo Salvini (in blu) con quella di Ultimo (in rosso), nome d'arte di Niccolò Moriconi, cantautore italiano e vincitore del Festival di Sanremo 2018 nella categoria “nuove proposte”, è risultato evidente come nell'anno passato il loro trend di ricerche sia risultato simile.

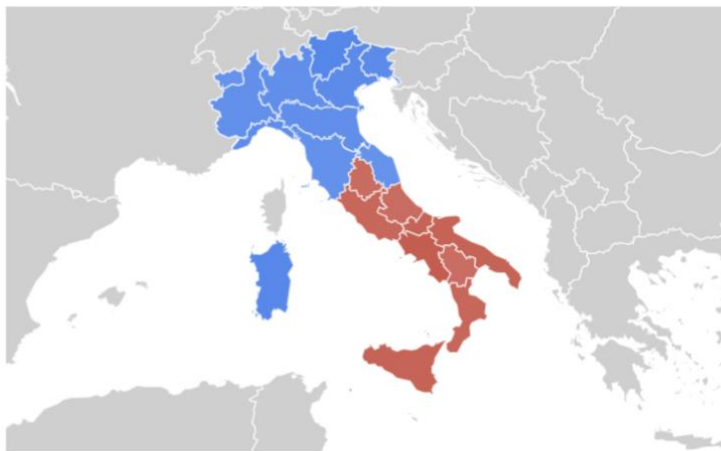


**Figura 27. Trend di ricerca per “Matteo Salvini” e “Ultimo” – Google Trends – 18 luglio 2019**



Il picco rosso risale al periodo in cui Ultimo, con non pochi scandali, si è aggiudicato il secondo posto al Festival di Sanremo. Se si allunga il periodo di ricerca, risulta evidente la superiorità in termini di popolarità di Matteo Salvini, che risulta più cercato di Ultimo se non si considera il picco durante il periodo del Festival di Sanremo. La cosa interessante da notare è la divisione delle ricerche per area geografica:

● Matteo Salvini ● Ultimo



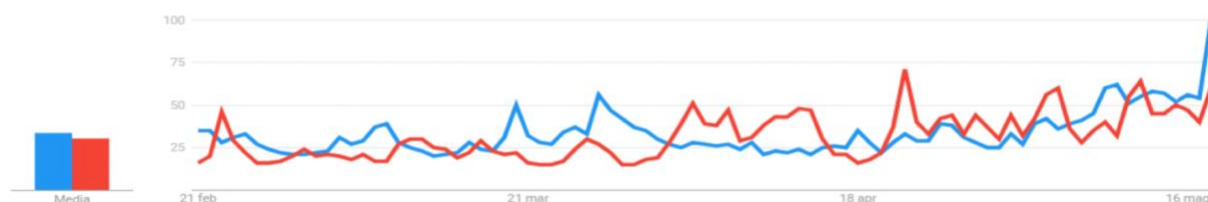
**Figura 28. Trend di ricerca per “Matteo Salvini” e “Ultimo”, divisione regionale – Google Trends – 18 luglio 2019**

Come si nota dal grafico la divisione è netta: il centro-sud Italia preferisce cercare più spesso notizie sul cantante romano, mentre il centro-nord e la Sardegna hanno effettuato un maggior numero di ricerche per il politico Matteo Salvini.

### 3.7.2 Matteo Salvini e Weather today

Nel tentativo di analizzare il successo di Matteo Salvini e per sostenere la tesi secondo cui sarebbe uscito vincitore da queste elezioni europee, è stato utilizzato Google Trends che ha mostrato una correlazione interessante con il termine di ricerca “meteo oggi”.

- “Matteo Salvini”
- “Weather today”



**Figura 29. Correlazione tra Matteo Salvini e le ricerche Google sul meteo**

Sembra che le ricerche effettuate sul tempo vengano effettuate dagli italiani durante le prime ore del mattino, appena svegli, e pare che subito dopo cerchino su Google informazioni su Matteo Salvini. È quasi evidente il perché le persone cerchino le condizioni meteo la mattina presto: i cambiamenti climatici, negli ultimi anni, hanno portato ad un clima inusuale e aperto a veloci alterazioni, e noi, la mattina, quando dobbiamo decidere cosa indossare per la giornata, siamo preoccupati di una possibile pioggia o di temperature troppo rigide.

Ma perché gli individui cercano Matteo Salvini con la stessa frequenza con cui si informano sulle condizioni meteo? Con una settimana di distanza il tasso di correlazione presente tra le ricerche sul meteo e sul leader della Lega Nord risulta

essere di 0,58. Molto spesso Google Trends evidenzia correlazioni non significative, o semplicemente frutto di casualità.

### ***3.7.3 Casi realmente correlati o Correlazione Spuria?***

Uno dei problemi più diffusi quando si analizzano dei dati alla ricerca di correlazioni è la possibilità di incontrare delle correlazioni spurie (in inglese si può anche chiamare “nonsense correlation”, termine che spiega molto bene il concetto)<sup>34</sup>. In statistica una correlazione indica un legame tra due variabili, che può essere sia positivo che negativo.

Una correlazione spuria avviene quando si prendono in esame due serie che analizzano due fenomeni diversi e non collegati tra loro, ma che presentano delle caratteristiche, nell'andamento della curva, simili.

Lo statistico inglese George Udny Yule<sup>35</sup>, nel 1926, descrisse questa associazione per la prima volta, perché aveva notato in un suo studio una correlazione positiva tra il tasso di mortalità e il tasso di matrimoni celebrati in chiesa.

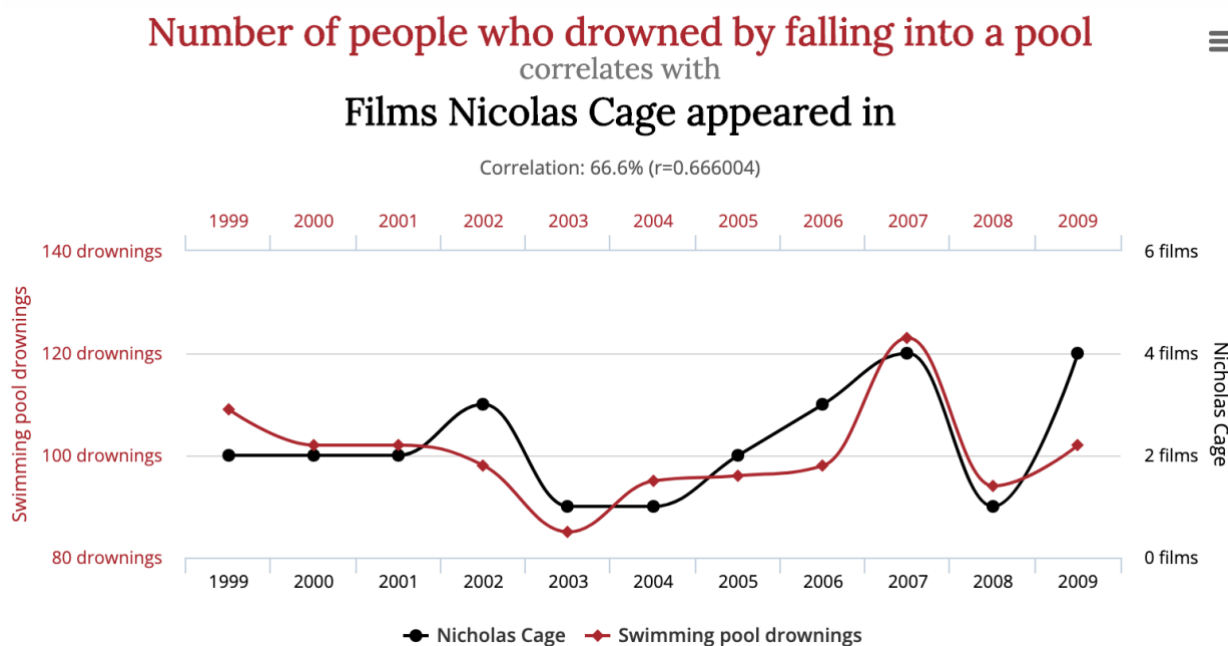
<sup>34</sup> “Correlazione Spuria (nonsense correlation correlation illusoire)”. Giovanni Visco. 2006.

Articolo Online: <http://www.cma4ch.org/chemo/heritage/correlazione/rslide24.html>

<sup>35</sup> Udny Yule, G. (1926), *Why Do We Get Some Nonsense Correlations Between Time Series? A Study in Sampling and the Nature of Time Serie*, journal of the royal statistical society. Online: [https://www.jstor.org/stable/2341482?seq=1#metadata\\_info\\_tab\\_contents](https://www.jstor.org/stable/2341482?seq=1#metadata_info_tab_contents)

Secondo il principio logico statistico, l'assunzione più corretta da fare sarebbe stata quella che affermava che un numero maggiore di matrimoni religiosi avrebbe portato ad una vita più breve. Se due fenomeni sono, secondo i dati, statisticamente collegati non è obbligatoriamente vero asserire che esista un legame di causa-effetto che li colleghi, in quanto la suddetta correlazione potrebbe essere casuale, e quindi dipendere da un'altra variabile in comune.

Per spiegare bene il concetto risulta più facile illustrarlo con dei grafici e degli esempi<sup>36</sup>.

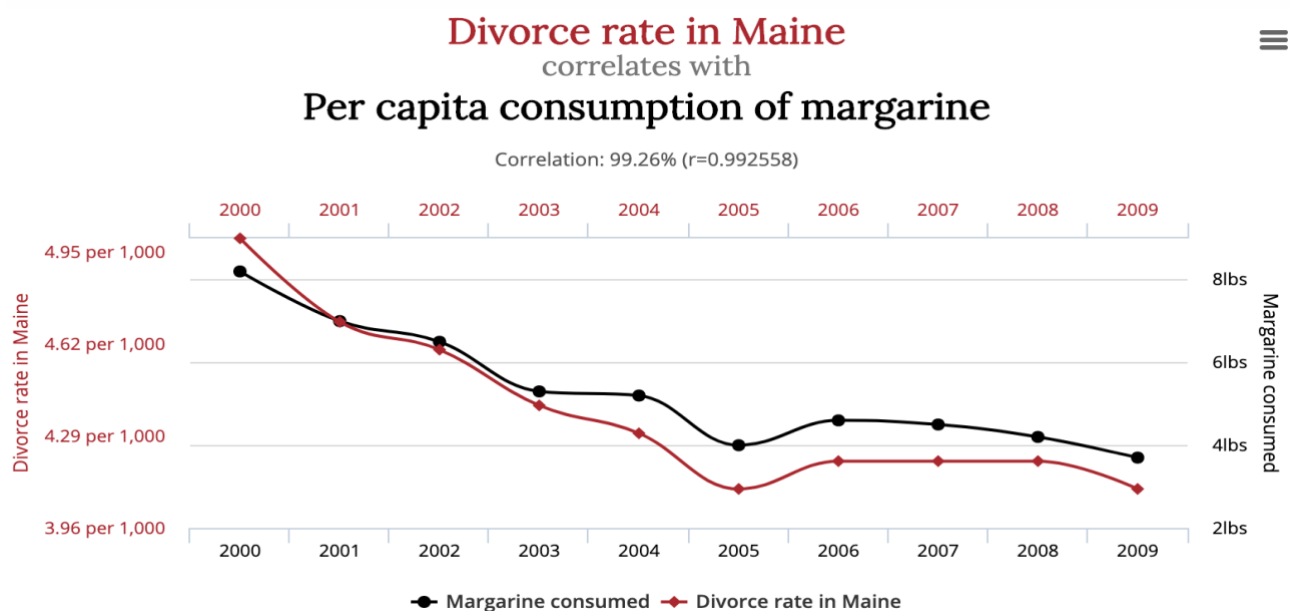


**Figura 30. Esempio di correlazione spuria: Film di Nicolas Cage (in ordine di uscita) e numero di persone affogate in piscina**

<sup>36</sup> “Spurious Correlations”. Articolo Online: <http://tylervigen.com/spurious-correlations>

Nel grafico (figura 30) le due curve sono: la nera è la linea che cattura l'andamento dell'uscita dei film di Nicholas Cage negli anni e la rossa identifica il numero di morti per annegamento nelle piscine. La correlazione è del 66,6% ( $r = 0,666004$ ), sufficientemente alta per poter dire che al crescere dell'uno cresce anche l'altro. Ovviamente non si può asserire causalità, quindi anche se collegati dal punto di vista statistico, non lo sono da un punto di vista logico.

Un altro esempio può essere fornito dal seguente grafico:



**Figura 31. Esempio di correlazione spuria: Tasso di divorzi nel Maine e Consumo di margarina pro-capite**

Come si vede la correlazione in questo caso è molto più forte rispetto all' esempio precedente, qui raggiunge il 99,26% ( $r = 0,992558$ ). Un indice di correlazione pari ad 1 tra due variabili vuol dire che i due fenomeni si muovono sempre insieme, si potrebbe asserire che siano la stessa cosa, quindi un valore così alto come in questo caso indica

una correlazione tra i due eventi fortissima, che però si rivela essere un caso, perché a rigor di logica non sarebbe possibile dire che i due siano collegati.

## Conclusioni

Lo scopo di questa tesi è riassumibile in due obiettivi principali: il primo è quello di analizzare e predire i risultati delle Elezioni Europee del 26 maggio 2019, utilizzando gli strumenti offerti da Google, mentre, il secondo è quello di dimostrare come i Big Data possano essere utilizzati in vario modo ai giorni nostri e come essi abbiano un potere che in molti casi si tende a sottovalutare.

Nel corso di questa tesi sono state condotte delle analisi volte allo studio della popolarità e dei consensi acquisiti dai principali leader e partiti politici in Italia. È stato dimostrato, inoltre, come le persone siano in grado di mentire nelle interviste e nei questionari, nonostante essi siano anonimi, perché meno intermediari sono presenti tra gli intervistati e il risultato finale del sondaggio e più questi ultimi si sentono tranquilli e non giudicati nelle loro risposte, fattori che in caso contrario li porterebbero a mentire pur di fare bella figura. Una fonte per ottenere risposte veritiere è, quindi, sicuramente Google, strumento cui tutti si rivolgono per esprimere pareri, risolvere problematiche o anche solo per curiosità. Google è una fonte di dati affidabile, in quanto, non presenta alcun intermediario tra l'utente e il risultato della ricerca richiesto, cosa che porterà gli utilizzatori a confidarsi apertamente. Una volta appurato l'enorme fonte di dati che Google offre, sono state effettuate delle analisi volte allo studio della popolarità e dei consensi acquisiti dai principali Partiti e leader politici in Italia. Per predire i risultati delle Elezioni Europee prima ancora che esse prendessero luogo sono stati utilizzati due strumenti: Google Trends e Google Correlate.

Dalle analisi svolte è risultata evidente l'intenzione dei cittadini italiani di volersi informare sulle elezioni, il numero di ricerche con oggetto queste elezioni ha addirittura doppiato quello degli anni passati. Purtroppo la realtà ha però dimostrato che la percentuale di astensionisti è aumentata sensibilmente nel corso degli anni, questo ha aperto una parentesi importante sulla sfiducia e sulla poca considerazione che la popolazione italiana ha dei partiti e della politica nel suo complesso.

In seguito ad una previsione di partecipazione inferiore rispetto agli altri anni, bisognava capire in che modo i votanti avrebbero espresso il proprio parere nelle elezioni europee del 26 maggio 2019.

Studiando i trend delle ricerche effettuate su Google Trends è emersa una superiorità in termini di numero di click effettuati sugli argomenti riguardanti la Lega Nord rispetto agli altri due partiti leader in Italia, Movimento 5 Stelle e Partito Democratico, e quindi un maggiore interesse dimostrato nei confronti del partito leghista.

Nel corso del tempo si è assistito ad un vero e proprio sorpasso del Partito capeggiato da Matteo Salvini rispetto a quello di Luigi Di Maio. Se durante le elezioni politiche del 4 marzo 2018 il Movimento 5 Stelle, non solo risultava primo partito d'Italia ai seggi elettorali, ma era anche il più cercato su Google, successivamente all'accordo di Governo firmato con la Lega Nord, si nota un vero e proprio disinteresse nelle ricerche su Google riguardanti i pentastellati. Questo trend negativo viene affiancato ad un aumento di curiosità per il gruppo Lega Nord, che in poco meno di un anno è riuscito a superare l'avversario sia in termini di ricerche su Google, che nei sondaggi e successivamente alle urne.



È risultato evidente, inoltre, come Matteo Salvini abbia superato gli avversari politici in termini di popolarità, egli infatti è stato addirittura più cercato su Google di alcuni personaggi famosi in Italia, come Chiara Ferragni, Cristiano Ronaldo e Papa Francesco. La popolarità di Matteo Salvini non si limita a questo però, egli è riuscito meglio di chiunque altro a far identificare il proprio nome con quello del Partito di appartenenza, quando si pensa al termine “Lega Nord” lo si associa in automatico al nome del capogruppo.

La conclusione più evidente in seguito a queste analisi è proprio quella per cui la Lega Nord avrebbe ottenuto una percentuale alle elezioni nettamente superiore, rispetto a quella di un anno e mezzo addietro (4 marzo 2018), il Movimento 5 Stelle avrebbe visto i suoi consensi calare drasticamente e il Partito Democratico sarebbe riuscito ad ottenere un lieve aumento di preferenze. Questa situazione si è poi verificata in pieno nelle elezioni europee del 26 maggio 2019.

È stato inoltre dimostrato come in molti casi i Big Data mostrino delle correlazioni, anche abbastanza evidenti, che in realtà non sono altro che spurie, quindi non veritiere. Pertanto nell'estrazione di informazioni occorre prestare attenzione a non commettere errori e ritenere collegati degli eventi che in realtà potrebbero non esserlo, come nel caso mostrato nella tesi della correlazione tra Matteo Salvini e le ricerche sul meteo.

Gli andamenti previsti in questa tesi hanno trovato riscontro nella realtà delle elezioni europee tenutesi il 26 maggio 2019. Infatti i risultati resi pubblici il 27 maggio 2019 hanno rispecchiato in pieno le previsioni proposte in questo elaborato.

Le precedenti analisi hanno dimostrato quanto il potere dei Big Data sia enorme e quanto, chiunque possa usarlo a proprio vantaggio.

Si pensi alle aziende e a come esse potrebbero cercare correlazioni tra beni o servizi ricercati su Google, o la frequenza e la periodicità con cui vengono cercati.

Questi dati sono una fonte enorme di informazioni e utilizzati nella maniera corretta e da mani esperte sono in grado di dirci molto più di quello che si può scoprire da un sondaggio online o da un'intervista anonima, che presentano anche dei costi più elevati.

I Big Data e le tecnologie in generale dovrebbero essere considerati maggiormente come un supporto ed un aiuto nelle attività, nelle ricerche e negli studi, probabilmente il miglior supporto che potremmo mai avere. Senza esse non vivremmo in una società così moderna e civile; ma è fondamentale non dimenticare la nostra realtà di esseri umani e utilizzare le tecnologie come aiuto e supporto, e non come nostro sostituto, perché se siamo arrivati ad un tale livello di progresso è solo merito dell'intelletto umano.

## Ringraziamenti

Questa tesi è nata da un progetto di gruppo supervisionato dal Professor Luigi Laura, che vorrei ringraziare per questa meravigliosa occasione e per il supporto e la disponibilità sempre dimostrati durante il mio percorso universitario.



In ordine: Emma Gamba, io, Antonio Miloso, Seth Stephens-Davidowitz e il professor Luigi Laura.

Gli altri componenti del team sono stati di fondamentale importanza nella realizzazione di questo progetto, un sentito ringraziamento va anche a loro.

Questa foto risale al giorno in cui abbiamo avuto l'incredibile opportunità di presentare il nostro progetto a Seth Stephens-Davidowitz, autore del libro "Everybody lies".

Vorrei inoltre ringraziare Marco Iecher che per me è stato oltre che un capo, il migliore che si potesse desiderare, anche un amico. Grazie a lui e al Professor Luigi Laura ho avuto l'opportunità di entrare nel mondo del Luiss LOFT, un laboratorio di cose favolose, e di persone fantastiche.

A Lorenzo e Claudia, amici e compagni di vita, devo porgere i miei ringraziamenti per tutte quelle volte in cui mi hanno supportata e sopportata, nei bei momenti con una risata e in quelli brutti con una carezza.

I ringraziamenti più importanti, infine, vanno ai miei genitori per quello che hanno sempre fatto per me e per tutte le possibilità che mi hanno dato; è merito vostro ogni mio piccolo traguardo.

## Bibliografia

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press

Galli, G., & Capecchi, V. (1968). *Il comportamento elettorale in Italia: una indagine ecologica sulle elezioni in Italia fra il 1946 e il 1963* (Vol. 1). Bologna, Il Mulino

Biorcio, R. (2010), *Gli antecedenti politici della scelta di voto: l'identificazione di partito e l'autocollocazione sinistra-destra*, in Bellucci, P. e Segatti, P. (a cura di), *Votare in Italia: 1968-2008*. Dall'appartenenza alla scelta, Bologna, il Mulino, pp. 187-211

Dalton, R. e Watterberg, M. (2000), *Parties without Partisans. Political Change in Advanced Industrial Democracies*, Oxford, Oxford University Press

Agosta, A. (2006), *Come leggere la partecipazione elettorale*, «il Mulino»,  
55, 3, pp. 465-469

Tuorto, D. (2006), *Apatia o protesta? L'astensionismo elettorale in Italia*,  
Bologna, il Mulino

Caciagli, M. e Spreafico, A. (a cura di) (1990), *Vent'anni di elezioni in  
Italia. 1968-1987*, Padova, Liviana

Fabrizio, D., Feltrin, P. e Marcone, L. (2006), *La decisività del voto nelle  
regioni meridionali: un'analisi degli orientamenti elettorali negli ultimi  
dieci anni*, «Polena», 3, 2, pp. 9-35

Nuvoli, P. (1989), *Il dualismo elettorale Nord-Sud in Italia: persistenza o  
progressiva riduzione?*, «Quaderni dell'osservatorio elettorale»

Corbetta, P., Parisi, A. e Schadee, H. (1996), *Elezioni in Italia. Struttura e  
tipologia delle consultazioni politiche*, Bologna, il Mulino

## Sitografia

“Note tematiche sull’Unione Europea”. Eeva Pavy. Maggio, 2019.

Articolo Online: <http://www.europarl.europa.eu/factsheets/it/sheet/19/il-parlamento-europeo-competenze>

Come funziona la legge elettorale per le europee. Il Post, 16 maggio 2019.

Articolo Online: <https://www.ilpost.it/2019/05/16/leggi-elettorali-europee/>

“Elezioni europee 2019 – Manuale elettorale”. Studi Camera. 26

giugno,2019. Articolo Online: <https://temi.camera.it/leg18/temi/elezioni-europee-2019.html>

“Decreto del Presidente della Repubblica”. Gazzetta Ufficiale della

Repubblica Italiana, Istituto Poligrafico e Zecca dello Stato. 22 Marzo,

2019. Articolo Online:

<https://www.gazzettaufficiale.it/eli/id/2019/03/25/19A02051/sg>

“Decreto del Presidente della Repubblica”. Gazzetta Ufficiale della Repubblica Italiana, Istituto Poligrafico e Zecca dello Stato. 22 Marzo, 2019. Articolo Online:

<https://www.gazzettaufficiale.it/eli/id/2019/03/25/19A02051/sg>

“Rosatellum bis: la nuova legge elettorale. Ecco come funziona (la scheda)”. Sgherza Alessio. 26 ottobre, 2017. Articolo Online:

[https://www.repubblica.it/politica/2017/09/22/news/rosatellum\\_bis\\_scheda-176193741/?refresh\\_ce](https://www.repubblica.it/politica/2017/09/22/news/rosatellum_bis_scheda-176193741/?refresh_ce)

“Rosatellum 2.0, come sarebbero assegnati i seggi a Camera e Senato”.

Ansa-Centimetri. 22 luglio, 2019. Articolo Online:

[https://www.tgcom24.mediaset.it/politica/infografica/rosatellum-2-0-come-sarebbero-assegnati-i-seggi-a-camera-e-senato\\_1001044-2017.shtml](https://www.tgcom24.mediaset.it/politica/infografica/rosatellum-2-0-come-sarebbero-assegnati-i-seggi-a-camera-e-senato_1001044-2017.shtml)



“Europee: affluenza in netto aumento, la più alta da 20 anni”. Redazione ANSA. 27 maggio, 2019. Articolo Online:

[http://www.ansa.it/europee\\_2019/notizie/2019/05/25/litalia-al-voto-alle-europee-in-gioco-il-governo\\_bc75d6b0-918f-4388-8830-2a888c7b4942.html](http://www.ansa.it/europee_2019/notizie/2019/05/25/litalia-al-voto-alle-europee-in-gioco-il-governo_bc75d6b0-918f-4388-8830-2a888c7b4942.html)

“Elezioni Europee 2019, in Italia affluenza al 56%: in calo sul 2015 soprattutto al Sud. In Ue mai così alta da 20 anni”. Il Fatto Quotidiano. 27 maggio 2019. Articolo Online:

<https://www.ilfattoquotidiano.it/2019/05/27/elezioni-europee-2019-in-italia-affluenza-al-56-in-calo-sul-2014-soprattutto-al-sud-in-ue-mai-cosi-alta-da-20-anni/5208070/>

“Risultati delle elezioni europee”. Parlamento europeo. 2014. Articolo Online: <http://www.europarl.europa.eu/elections2014-results/it/turnout.html>

“Risultati delle elezioni europee del 2019” . Parlamento europeo in collaborazione con Kantar. 24 giugno, 2019. Articolo Online:

<https://risultati-elezioni.eu/affluenza/>

“Perché le persone non vanno a votare? Le cause principali dell’astensionismo”. Openpolis. 23 giugno 2016. Articolo online:

<https://blog.openpolis.it/2016/06/23/perche-le-persone-non-vanno-votare-cause-astensionismo/8870>

ISTAT, C. (2018). “Rapporto Bes 2018”. 18 dicembre, 2018. Online:

[https://www.istat.it/it/files//2018/12/Bes\\_2018.pdf](https://www.istat.it/it/files//2018/12/Bes_2018.pdf)

ISTAT, C. (2018). “Rapporto Bes 2018”. 18 dicembre, 2018. Pag. 85-93.

Barbara Baldazzi. Online: <https://www.istat.it/it/files//2018/12/BES2018-cap-06.pdf>

ISTAT, C. (2018). “Rapporto Bes 2018”. 18 dicembre, 2018. Pag. 91.

Barbara Baldazzi. Online: <https://www.istat.it/it/files//2018/12/BES2018-cap-06.pdf>

“Elezioni, l’Italia dell’astensionismo. Il 52% non si riconosce in nessun partito”. Il fatto quotidiano. 18 maggio 2015. Articolo online:

[www.ilfattoquotidiano.it/2015/05/18/elezioni-litalia-dellastensionismo-il-52-non-si-riconosce-in-nessun-partito/1694292/](http://www.ilfattoquotidiano.it/2015/05/18/elezioni-litalia-dellastensionismo-il-52-non-si-riconosce-in-nessun-partito/1694292/)

“La Partecipazione elettorale in Italia”. Quaderni di sociologia pag. 17 -39.

Maurizio Cerruto. 2012. Articolo online:

<https://journals.openedition.org/qds/537>

Ottaviani, Jacopo and Ferrari, Lorenzo. What happens with abstentionism in Europe. Internazionale, May 16, 2019. Online article:

<https://www.internazionale.it/notizie/jacopo-ottaviani/2019/05/16/elezioni-europee-astensionismo>

“Google Correlate Tutorial”. Online:

<https://www.google.com/trends/correlate/tutorial>

“Correlazione Spuria (nonsense correlation correlation illusoire)”.

Giovanni Visco. 2006. Articolo Online:

<http://www.cma4ch.org/chemo/heritage/correlazione/rslide24.html>

Udny Yule, G. (1926), *Why Do We Get Some Nonsense Correlations Between Time Series? A Study in Sampling and the Nature of Time Serie*, journal of the royal statistical society. Online:

[https://www.jstor.org/stable/2341482?seq=1#metadata\\_info\\_tab\\_contents](https://www.jstor.org/stable/2341482?seq=1#metadata_info_tab_contents)

“Spurious Correlations”. Articolo Online: <http://tylervigen.com/spurious-correlations>