

Incontri di promozione della cultura statistica
Iniziativa Utenti
Aprile, 2024

Sintesi in tabella

Monica Bailot

DCCI - CIC - Servizio di coordinamento della promozione della cultura statistica e della diffusione dei dati a livello territoriale

Indice

- **Il dato e la forma**
- **Statistica descrittiva**
- **Tabelle di frequenza semplici e doppie**
- **Relazioni tra variabili**

La statistica e la rappresentazione della realtà

Statisticare una realtà

vuol dire *approssimarla*,

fornendo una sua *rappresentazione virtuale*,

ma veritiera

Ripartiamo dai metadati

titolo

Istituzioni non profit e dipendenti per ripartizione geografica. Anno 2021 (valori assoluti) (dati definitivi)

intestazioni
di riga

Ripartizioni	Istituzioni non profit	Dipendenti
Nord-Ovest	100.537	296.021
Nord-Est	81.075	211.548
Centro	79.757	201.114
Sud	65.237	113.778
Isole	34.019	71.280
Italia	360.625	893.741

intestazioni di
colonna

Risponde alle 5W



fonte: Censimento permanente delle istituzioni non profit

Il dato e la forma

Come rendere una tabella leggibile e informativa?

E' importante che la tabella sia comprensibile anche se estratta dal suo contesto

Istituzioni non profit per forma giuridica e periodo di costituzione. Italia.

Anno 2021 (valori assoluti) (dati definitivi)

La tabella grezza

	Periodo di costituzione					
Forma giuridica	Fino al 1984	1985-2004	2005-2014	2015-2019	2020-2021	Totale
Associazione riconosciuta e non riconosciuta	21573,00	102409,00	104074,00	65475,00	12716,00	306247,00
Cooperativa sociale	725,00	4651,00	5272,00	3552,00	769,00	14969,00
Fondazione	1954,00	2537,00	2548,00	1040,00	277,00	8356,00
Altra forma	9862,00	11303,00	4129,00	4247,00	1512,00	31053,00
Totale	34114,00	120900,00	116023,00	74314,00	15274,00	360625,00

Fonte: Censimento permanente delle istituzioni non profit

Alcuni accorgimenti

La tabella leggibile

Istituzioni non profit per forma giuridica e periodo di costituzione. Italia. Anno 2021 (valori assoluti) (dati definitivi)

Forma giuridica	Periodo di costituzione					Totale
	Fino al 1984	1985-2004	2005-2014	2015-2019	2020-2021	
Associazione riconosciuta e non riconosciuta	21.573	102.409	104.074	65.475	12.716	306.247
Cooperativa sociale	725	4.651	5.272	3.552	769	14.969
Fondazione	1.954	2.537	2.548	1.040	277	8.356
Altra forma	9.862	11.303	4.129	4.247	1.512	31.053
Totale	34.114	120.900	116.023	74.314	15.274	360.625

Fonte: Censimento permanente delle istituzioni non profit

Per aiutare la lettura di una tabella è importante:

- Allineare i numeri a destra
- Allineare le modalità in fiancata a sinistra
- Allineare le modalità in testata con i numeri
- Utilizzare il separatore delle migliaia
- Evitare di riempire la tabella di righe orizzontali e verticali

Alcuni accorgimenti

Istituzioni non profit per forma giuridica. Italia
Anno 2021 (valori assoluti e composizione %) (dati definitivi)

Forma giuridica	Istituzioni non profit	
	valori assoluti	valori percentuali
Associazione	306.247	84,9
Cooperativa sociale	14.969	4,2
Fondazione	8.356	2,3
Altra forma giuridica	31.053	8,6
Totale	360.625	100

Fonte: Censimento permanente delle istituzioni non profit

Per aiutare la lettura di una tabella è importante:

- Proporre un solo decimale
- Una tavola con percentuali dovrebbe contenere anche i valori assoluti
- Ordinare righe e colonne secondo un criterio (alfabetico, temporale, graduatoria ecc.)
- Le modalità come: *altro, non so* vanno inserite in fondo (*altra forma giuridica*)

Alcuni accorgimenti

Istituzioni non profit che non hanno utilizzato tecnologie digitali per motivi del mancato utilizzo e forma giuridica. Italia. Anno 2021 (valori assoluti) (dati definitivi)

MOTIVI DEL MANCATO UTILIZZO DELLE TECNOLOGIE DIGITALI	FORMA GIURIDICA				Totale istituzioni non profit che non utilizzano tecnologie digitali (b)
	Associazione	Cooperativa sociale	Fondazione	Altra forma Giuridica	
Per carenza di personale qualificato in materia ICT	7.644	344	142	1.071	
Per mancanza di adeguata formazione in materia ICT	3.155	61	28	302	
Per mancanza di risorse finanziarie	17.331	433	156	1.562	
Per l'assenza di investimenti in innovazione tecnologica	5.686	128	207	678	
Per la necessità di riorganizzare prima processi ed infrastrutture	2.489	40	72	293	
Per la scarsa cultura digitale	10.378	186	78	1.084	
Per la presenza di altre sfide/problematiche più urgenti	8.446	382	139	1.077	10.044
Non rilevante per le attività svolte	19.051	383	331	2.201	21.966
Altro motivo	1.314	86	41	52	1.494
Totale (a)	63.946	1.830	1.083	6.950	73.808

- Righe e colonne spaziate in modo regolare
- Colonne non troppo lontane tra loro
- Inserire le note vicino alla fonte

Fonte: Censimento permanente delle istituzioni non profit
(a) Eventuali differenze tra le somme ed il relativo totale sono da attribuire esclusivamente ad arrotondamenti nei decimali
(b) Il quesito prevede più modalità di risposta pertanto la somma delle singole modalità non corrisponde al totale delle istituzioni non profit che non utilizzano tecnologie digitali

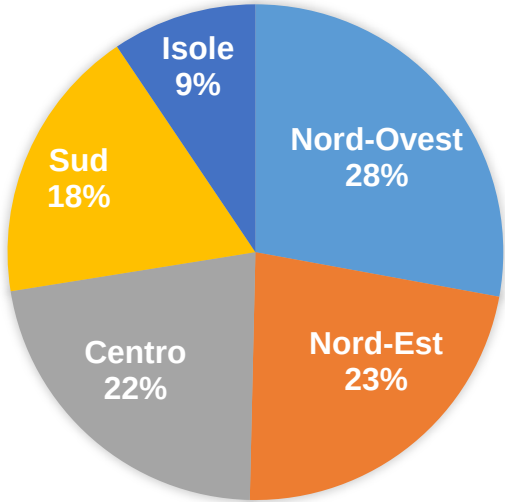
Statistica descrittiva

fotografa il fenomeno di interesse,
mettendone in luce le caratteristiche salienti
e sintetizzandole con
tabelle, grafici e indicatori



Istituzioni non profit e dipendenti per ripartizione geografica.
Anno 2021 (valori assoluti) (dati definitivi)

Ripartizioni	Istituzioni non profit	Dipendenti
Nord-Ovest	100.537	296.021
Nord-Est	81.075	211.548
Centro	79.757	201.114
Sud	65.237	113.778
Isole	34.019	71.280
Italia	360.625	893.741



Istituzioni non profit in Italia:
360 mila

Fonte: Censimento permanente delle istituzioni non profit

Analisi descrittiva

L'approccio può essere distinto in due tipologie di analisi:
monovariata e multivariata

Variabile di interesse: forma giuridica

Istituzioni non profit per forma giuridica. Italia
Anno 2021 (valori assoluti) (dati definitivi)

Forma giuridica	Istituzioni non profit
Associazione	306.247
Cooperativa sociale	14.969
Fondazione	8.356
Altra forma giuridica	31.053
Totale	360.625

Fonte: Censimento permanente delle istituzioni non profit

Variabili di interesse:

- forma giuridica
- periodo di costituzione

Istituzioni non profit per forma giuridica e periodo di costituzione. Italia
Anno 2021 (valori assoluti) (dati definitivi)

Forma giuridica	Periodo di costituzione					Totale
	Fino al 1984	1985-2004	2005-2014	2015-2019	2020-2021	
Associazione riconosciuta e non riconosciuta	21.573	102.409	104.074	65.475	12.716	306.247
Cooperativa sociale	725	4.651	5.272	3.552	769	14.969
Fondazione	1.954	2.537	2.548	1.040	277	8.356
Altra forma	9.862	11.303	4.129	4.247	1.512	31.053
Totale	34.114	120.900	116.023	74.314	15.274	360.625

Fonte: Censimento permanente delle istituzioni non profit

Matrice di dati

Una tabella «primitiva»
Distribuzione **unitaria multipla**

carattere o variabile statistica

Progressivo del rispondente	Cittadinanza	Regione della scuola	Sesso dello studente	Scuola media/scuola superiore	Sei nato in Italia o all'estero?	Quanti anni avevi quando sei arrivato in Italia?	In che lingua pensi?	Parli italiano in famiglia?
progr	cittad	reg	sezzo	tiposcuola	a1	a3	a6	a6p1
1		1	21	1	1	1		
2		1	21	2	2	1		
3		1	21	2	2	1		
4		1	21	1	1	1		
unità statistica		1	21	1	1	1		
		2	21	1	2	1		2
7		1	21	1	2	1		
8		1	21	2	2	1		
9		1	21	2	2	1		
10		1	21	1	2	1		

Analisi descrittiva: monovariata

Primo passo:
analizzare una variabile alla volta

Tabella di frequenza semplice

Variabile di interesse: forma giuridica

Istituzioni non profit per forma giuridica. Italia
Anno 2021 (valori assoluti) (dati definitivi)

Forma giuridica	Istituzioni non profit
Associazione	306.247
Cooperativa sociale	14.969
Fondazione	8.356
Altra forma giuridica	31.053
Totale	360.625



**frequenze
assolute**

Le frequenze assolute:

dipendono dalla numerosità della popolazione di riferimento

non permettono il confronto tra due o più popolazioni di numerosità differente

non offrono una visione immediata dell'importanza di ogni modalità

Fonte: Censimento permanente delle istituzioni non profit

Analisi descrittiva: monovariata

Primo passo:
Analizzare una variabile alla volta

Tabella di frequenza semplice

Variabile di interesse: forma giuridica

Istituzioni non profit per forma giuridica. **Italia**
Anno 2021 (valori assoluti e composizione %) (dati definitivi)

Forma giuridica	Istituzioni non profit	
	valori assoluti	valori percentuali
Associazione	306.247	84,9
Cooperativa sociale	14.969	4,2
Fondazione	8.356	2,3
Altra forma giuridica	31.053	8,6
Totale	360.625	100

frequenze percentuali

Fonte: Censimento permanente delle istituzioni non profit

Come confrontare popolazioni diverse?

Si calcolano le frequenze relative e/o le frequenze percentuali

Istituzioni non profit per forma giuridica. **Isole**
Anno 2021 (valori assoluti e composizione %) (dati definitivi)

Forma giuridica	Istituzioni non profit	
	valori assoluti	valori percentuali
Associazione	28.946	85,1
Cooperativa sociale	2.745	8,1
Fondazione	406	1,2
Altra forma giuridica	1.922	5,6
Totale	34.019	100

Fonte: Censimento permanente delle istituzioni non profit

Analisi descrittiva: bivariata

Secondo passo:
individuare eventuali associazioni tra le variabili

Tabella a doppia entrata o di contingenza

Istituzioni non profit per forma giuridica e periodo di costituzione. Italia
Anno 2021 (valori assoluti) (dati definitivi)

Forma giuridica	Periodo di costituzione					Totale
	Fino al 1984	1985-2004	2005-2014	2015-2019	2020-2021	
Associazione riconosciuta e non riconosciuta	21.573	102.409	104.074	65.475	12.716	306.247
Cooperativa sociale	725	4.651	5.272	3.552	769	14.969
Fondazione	1.954	2.537	2.548	1.040	277	8.356
Altra forma	9.862	11.303	4.129	4.247	1.512	31.053
Totale	34.114	120.900	116.023	74.314	15.274	360.625

Fonte: Censimento permanente delle istituzioni non profit

tabelle a doppia entrata
rappresentazione della
distribuzione doppia di frequenze

Tabella di contingenza:
almeno una delle due variabili è
qualitativa

Tabella di correlazione:
entrambe le variabili sono
quantitative

Associazione

L'analisi dell'associazione è legata ai concetti di **dipendenza** o **indipendenza assoluta** o **interdipendenza**, cioè il legame reciproco tra due variabili.

Il metodo utilizzato per valutare l'associazione dipende dagli obiettivi del ricercatore e dalla tipologia delle variabili

Metodo	Tipo di variabili		
	Entrambe qualitative	Una qualitativa e una quantitativa	Entrambe quantitative
Connessione	Sì	Sì	Sì
Dipendenza in media	No	Sì	Sì
Correlazione e regressione	No	No	Sì

Connessione. Esempio n. 1



Esiste un'associazione tra il genere e la soddisfazione per il lavoro svolto?

Persone tra 25 e 44 anni secondo la soddisfazione per il lavoro e genere.
Italia. Anno 2022 (valori in migliaia)

Soddisfazione	Genere		Totale
	Maschio	Femmina	
molto	971	738	1.709
abbastanza	3.239	2.376	5.615
poco	770	646	1.416
per niente	154	108	262
Totale	5.134	3.868	9.002

Fonte: Indagine multiscopo «Aspetti della vita quotidiana»

Indici di connessione

misurano la forza del legame associativo tra due variabili

Indice di Cramer

(o di connessione normalizzato)

L'indice **C** ha le seguenti proprietà:

- vale 0 in caso di indipendenza
- vale 1 in caso di massima connessione

in tutti gli altri casi assume un valore compreso tra 0 e 1

Indice di Cramer. Esempio n. 1



Primo passo: calcoliamo le frequenze teoriche attese in caso di indipendenza

Indipendenza assoluta: frequenze teoriche attese

Soddisfazione	Genere		Totale
	Maschio	Femmina	
molto	975	734	1.709
abbastanza	3.202	2.413	5.615
poco	808	608	1.416
per niente	149	113	262
Totale	5.134	3.868	9.002

Come si calcolano:

$$\frac{1.709 \times 5.134}{9.002} = 975$$

$$\frac{1.709 \times 3.868}{9.002} = 734$$

Indice di Cramer. Esempio n. 1



Secondo passo: confrontiamo le frequenze osservate con quelle teoriche attese in caso di indipendenza

Frequenze osservate

Soddisfazione	Genere		Totale
	Maschio	Femmina	
molto	971	738	1.709
abbastanza	3.239	2.376	5.615
poco	770	646	1.416
per niente	154	108	262
Totale	5.134	3.868	9.002

Frequenze teoriche attese

Soddisfazione	Genere		Totale
	Maschio	Femmina	
molto	975	734	1.709
abbastanza	3.202	2.413	5.615
poco	808	608	1.416
per niente	149	113	262
Totale	5.134	3.868	9.002

Indice di Cramer. Esempio n. 1



L'indice di connessione di Cramer si basa sulla distanza tra frequenze osservate e attese

Soddisfazione	Genere	
	Maschio	Femmina
molto	0,014	0,018
abbastanza	0,420	0,557
poco	1,748	2,320
per niente	0,140	0,186
Somma delle distanze	5,403	
Indice di Cramer	0,024	

Come si calcolano le distanze:

$$\frac{(f.osservata - f.teorica)^2}{f.teorica}$$

$$\frac{(971 - 975)^2}{975} = 0,014$$

Indice di Cramer. Esempio n. 1



L'indice di connessione di Cramer si basa sulla distanza tra frequenze osservate e attese

Soddisfazione	Genere	
	Maschio	Femmina
molto	0,014	0,018
abbastanza	0,420	0,557
poco	1,748	2,320
per niente	0,140	0,186
Somma delle distanze	5,403	
Indice di Cramer	0,024	

Esiste un legame tra le due variabili ma non è molto forte

Come si calcola Cramer:

$$\sqrt{\frac{\sum \frac{(f.osservata - f.teorica)^2}{f.teorica}}{N \times \min[(n.righe - 1), (n.colonne - 1)]}}$$

$$\sqrt{\frac{5,403}{9.002 \times \min[(4 - 1), (2 - 1)]}}$$

$$\sqrt{\frac{5,403}{9.002 \times 1}} = 0,024$$

Connessione. Esempio n. 2



Esiste un’associazione tra il genere e il regime orario del lavoro?

Occupati tra 35 e 49 anni secondo il regime orario e genere. Italia.
Anno 2023 (valori in migliaia)

Regime orario	Genere		Totale
	Maschi	Femmine	
Tempo pieno	4.683	2.559	7.242
Tempo parziale	303	1.246	1.549
Totale	4.986	3.806	8.791

Fonte: Indagine continua sulle forze di lavoro

Indipendenza assoluta

Regime orario	Genere		Totale
	Maschi	Femmine	
Tempo pieno	4.107	3.135	7.242
Tempo parziale	879	671	1.549
Totale	4.986	3.806	8.791

Indice di Cramer. Esempio n. 2



Esiste un'associazione tra il genere e il regime orario del lavoro?

Regime orario	Genere	
	Maschi	Femmine
Tempo pieno	81	106
Tempo parziale	377	494
Somma delle distanze	1058	
Indice di Cramer	0,347	

L'indice di Cramer **evidenzia un certo legame** tra le due variabili

Anche dalle frequenze assolute si nota una **prevalenza del tempo parziale per le donne**

Esiste una relazione lineare tra la percentuale di Neet e l'andamento del mercato del lavoro?

La fonte è la banca dati

a misura di Comune

La percentuale di Neet è stata messa in relazione con due indicatori del mercato del lavoro per tutte le province d'Italia (vedi tema *Istruzione - Tavola 4.2*, e tema *Lavoro - Tavola 2.2*).

Regressione. Definizioni delle variabili

Percentuale di Neet

(giovani tra 15 e 29 anni)

$$\frac{\text{Giovani che non studiano e non lavorano}}{\text{Giovani}} \times 100$$

Tasso di disoccupazione

(persone di 15 anni e più)

$$\frac{\text{Disoccupati}}{\text{Occupati} + \text{Disoccupati}} \times 100$$

Tasso di occupazione

(persone di 15 anni e più)

$$\frac{\text{Occupati}}{\text{Popolazione residente}} \times 100$$

Regressione. Esempio n. 1

Percentuale di Neet e tasso di disoccupazione per provincia. Anno 2019

Provincia	% Neet	tasso disoccupazione
Alessandria	23,9	11,0
Asti	21,5	9,7
Biella	22,8	9,2
Cuneo	18,5	6,6
Novara	22,5	10,3
Torino	22,3	10,8
Verbano-Cusio-Ossola	24,9	11,4
Vercelli	22,0	10,2

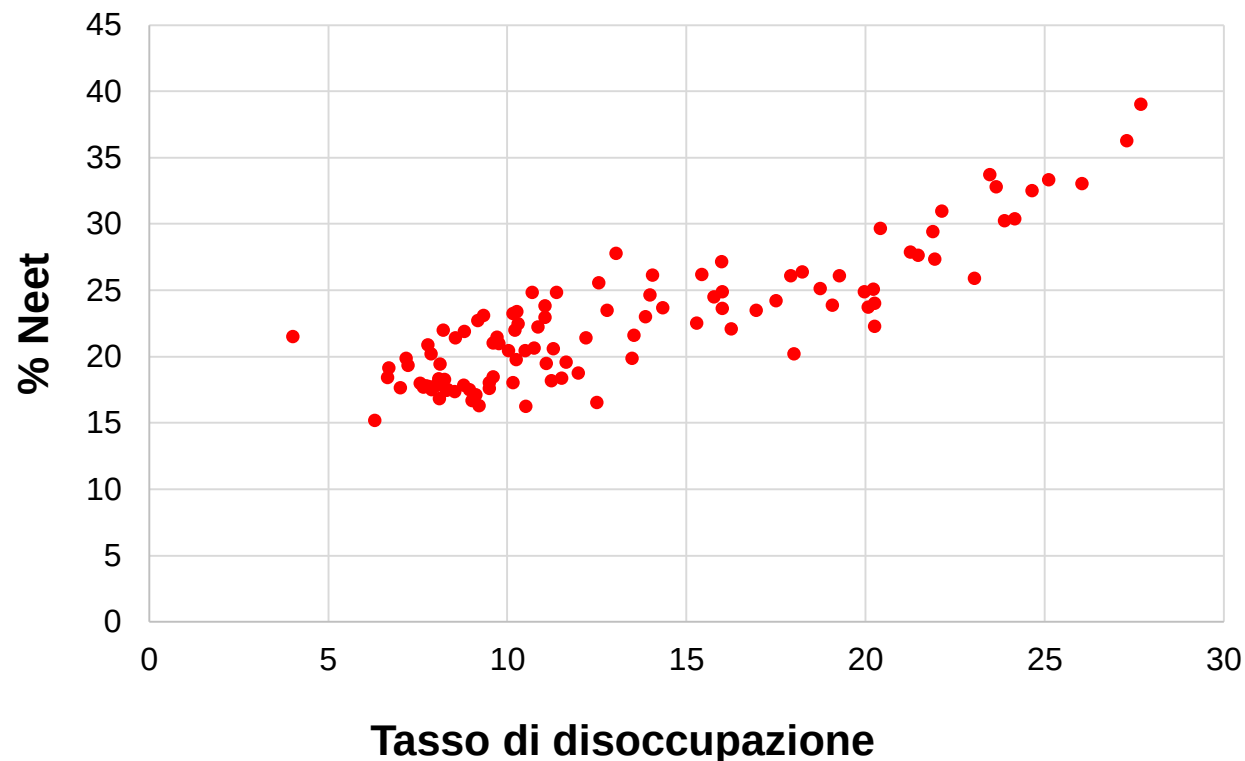
Esiste una relazione lineare tra la percentuale di Neet e il tasso di disoccupazione?

La tavola contiene 107 righe, una per ogni provincia

Fonte: Istat - Censimento permanente della popolazione e delle abitazioni, Condizioni socio-economiche delle famiglie - ARCH.I.M.E.DE (fonti amministrative integrate)

Regressione. Esempio n. 1

Primo passo: il diagramma a dispersione



Dal diagramma a dispersione si possono osservare:

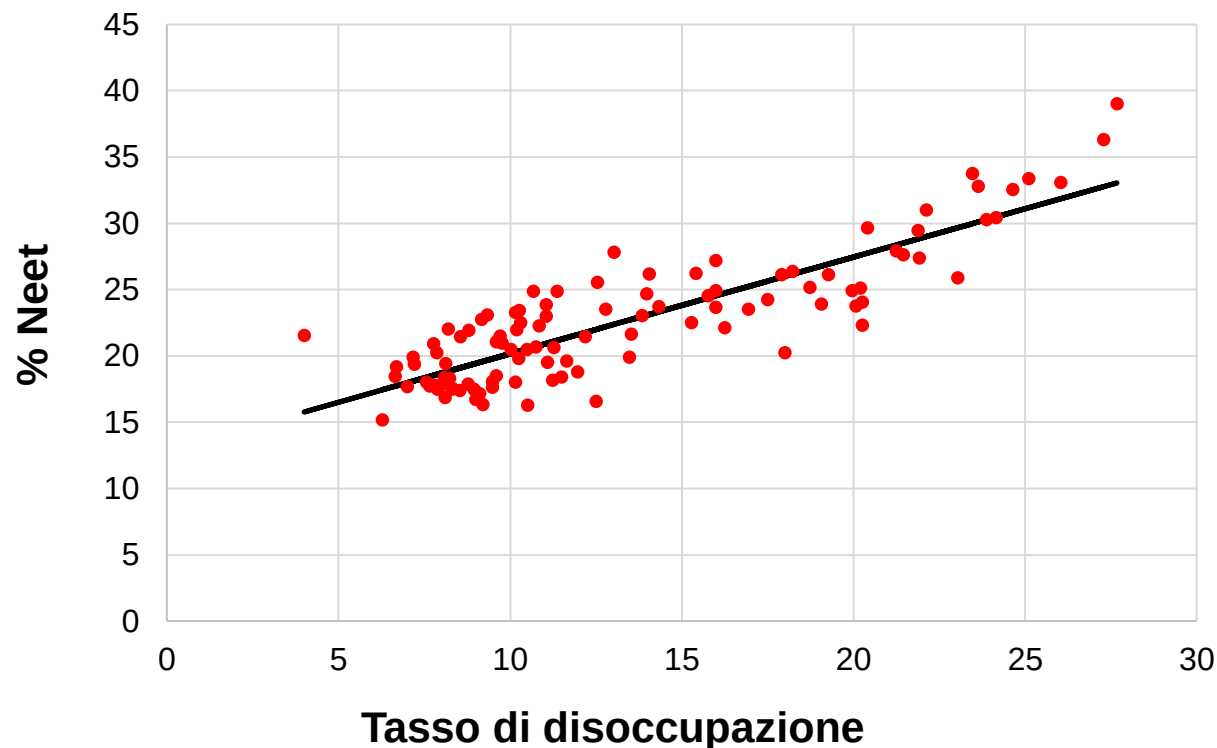
- **direzione** della relazione
- **forma** della relazione
- **forza** della relazione

Nell'esempio esiste una **associazione positiva**, al crescere del tasso di disoccupazione, cresce la percentuale di Neet.

Fonte: Istat - Censimento permanente della popolazione e delle abitazioni, Condizioni socio-economiche delle famiglie - ARCH.I.M.E.DE (fonti amministrative integrate)

Regressione. Esempio n. 1

Secondo passo: la regressione lineare



Fonte: Istat - Censimento permanente della popolazione e delle abitazioni, Condizioni socio-economiche delle famiglie - ARCH.I.M.E.DE (fonti amministrative integrate)

Obiettivo:

comprendere come la percentuale di Neet dipenda dal tasso di disoccupazione

Modello:

$\%Neet = \text{costante} + b \times \text{tasso di disoccupazione}$

Bontà di adattamento:

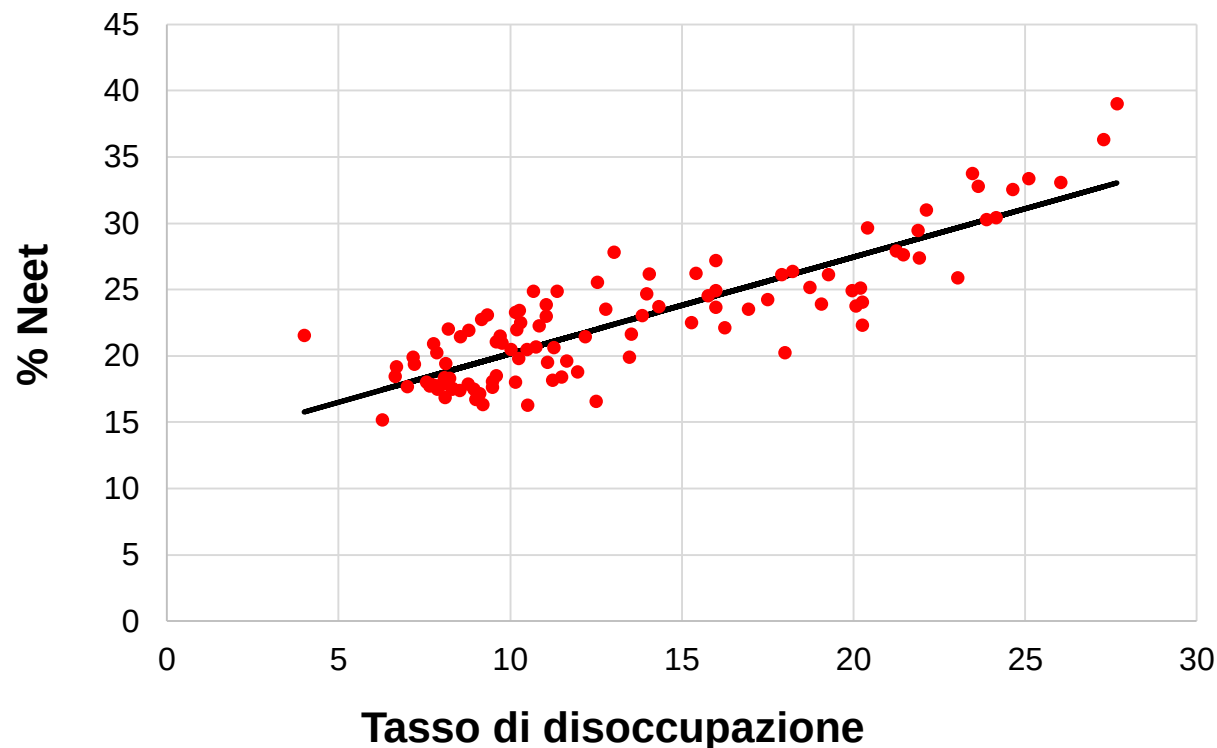
La retta riesce a riprodurre i dati osservati? Quanto sono distanti i dati stimati da quelli osservati?

Si misura con il:

Coefficiente di determinazione R^2

Regressione. Esempio n. 1

Secondo passo: la regressione lineare



Fonte: Istat - Censimento permanente della popolazione e delle abitazioni, Condizioni socio-economiche delle famiglie - ARCH.I.M.E.DE (fonti amministrative integrate)

Modello:

$\%Neet \text{ stimata} = 12,9 + 0,73 \times \text{tasso disoccupazione}$

Bontà di adattamento: 0,75

Coefficiente di determinazione R^2
ha le seguenti proprietà:

- Prossimo a 0, la retta non descrive la variabilità osservata
- Vicino a 0,5, la retta descrive circa la metà della variabilità osservata
- Superiore a 0,7, il grado di adattamento è elevato. **La retta è un buon modello per descrivere il legame tra le due variabili**

Regressione. Esempio n. 2

Percentuale di Neet e tasso di occupazione per provincia. Anno 2019

Provincia	% Neet	Tasso occupazione
Alessandria	23,9	46,6
Asti	21,5	48,0
Biella	22,8	46,5
Cuneo	18,5	52,4
Novara	22,5	49,3
Torino	22,3	48,0
Verbano-Cusio-Ossola	24,9	44,1
Vercelli	22,0	47,2

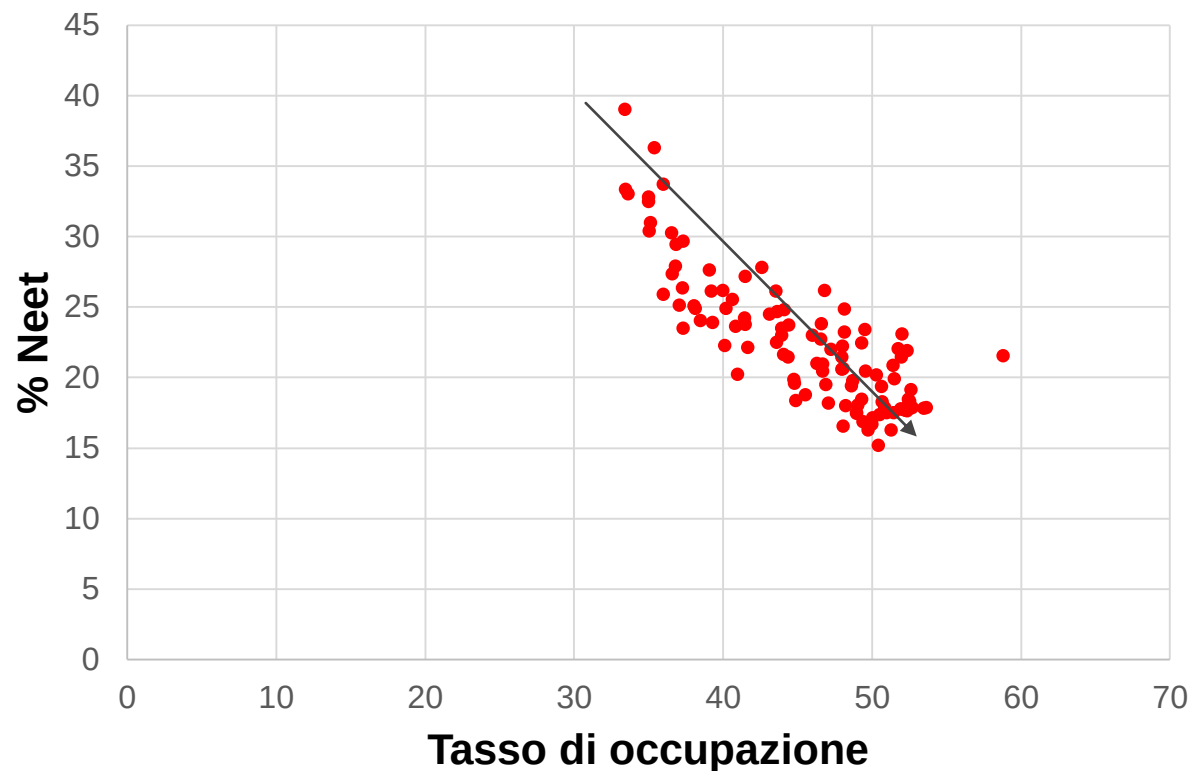
Esiste una relazione lineare tra la percentuale di Neet e il tasso di occupazione?

La tavola contiene 107 righe, una per ogni provincia

Fonte: Istat - Censimento permanente della popolazione e delle abitazioni, Condizioni socio-economiche delle famiglie - ARCH.I.M.E.DE (fonti amministrative integrate)

Regressione. Esempio n. 1

Primo passo: il diagramma a dispersione

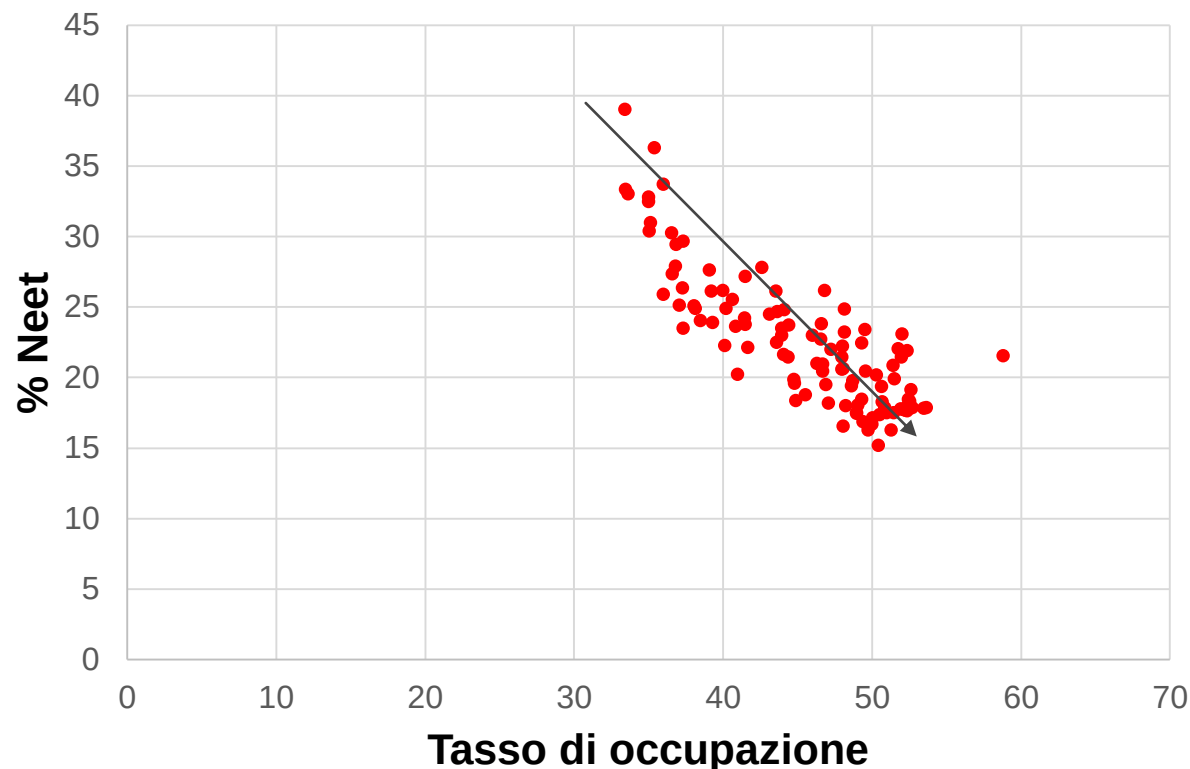


Nell'esempio si ha una **associazione negativa**, al crescere del tasso di occupazione, decresce la percentuale di Neet.

Fonte: Istat - Censimento permanente della popolazione e delle abitazioni, Condizioni socio-economiche delle famiglie - ARCH.I.M.E.DE (fonti amministrative integrate)

Regressione. Esempio n. 2

Secondo passo: la regressione lineare



Modello:

%Neet stimata = $53,1 - 0,67 \times$ tasso occupazione

Bontà di adattamento: 0,7

Il grado di adattamento è buono.

La retta è un buon modello per descrivere il legame tra le due variabili

Fonte: Istat - Censimento permanente della popolazione e delle abitazioni, Condizioni socio-economiche delle famiglie - ARCH.I.M.E.DE (fonti amministrative integrate)

Bibliografia e sitografia

(2009) Antolini F., Truglia F., La statistica e le statistiche: dal dato amministrativo al dato statistico. Fonti e strumenti di analisi. Milano. LED

(2012) Bocci L., Mingo I., Statistiche, istruzioni per l'uso. Roma. Edizioni Nuova Cultura

(2014) Corbetta P. Metodologia e tecniche della ricerca sociale. Bologna. Il Mulino

Ventura L., Valutazione dei fattori di rischio per la salute: correlazione e regressione. Unipd. Dipartimento di Scienze Statistiche (<https://www.stat.unipd.it/terza-missione/statistica-tutti-0>)

Zuccolotto P., Analisi dell'associazione tra caratteri. Unibs. Scienze economiche e statistiche

IstatData. <https://esploradati.istat.it/databrowser/>

A misura di comune: <https://www.istat.it/it/archivio/279229>

Censimento permanente delle istituzioni non profit: <https://www.istat.it/it/archivio/296339>

Grazie

utenti.territorio@istat.it