



UNIVERSITÀ  
DEGLI STUDI  
DI TORINO

Corso di **Pedagogia Sperimentale**  
Dipartimento di **Scienze dell'Educazione**  
Indirizzo **Nidi e comunità infantili**  
Professore: Roberto Trincherò

**RELAZIONE TRA L' AVERE AMICI  
FUMATORI E LA PROPRIA SCELTA DI  
FUMARE**

Ricerca empirica di:  
Cecilia Appendino (866073)  
Erika Percassi (865189)  
Ramona Decataldo (849918)

Anno scolastico 2019/2020

## **SOMMARIO:**

1. Premessa
2. Tema di ricerca
3. Problema di ricerca
4. Obiettivo di ricerca
5. Quadro teorico
6. Mappa concettuale
7. Bibliografia e sitografia
8. Ipotesi di ricerca- Individuazione dei fattori e definizione operativa
9. Individuazione del campione e tipologia di campionamento
10. Tecniche e strumenti di rilevazione dei dati
11. Questionario
12. Tecniche di analisi dei dati utilizzati
13. Interpretazione dei risultati
14. Autoriflessione

### **1. Premessa:**

Come argomento della nostra ricerca empirica esamineremo, tramite un questionario sottoposto a un gruppo di ragazzi di Torino e provincia, di età compresa tra i 14 e i 26 anni, come l'influenza del gruppo di amici possa essere un forte predittore della probabilità che ragazzi non fumatori possano iniziare a fumare.

### **2. Tema della ricerca:**

Relazione tra avere amici fumatori e scelta di iniziare a fumare.

### **3. Problema conoscitivo:**

Vi è relazione tra avere amici fumatori e scelta di iniziare a fumare?

### **4. Obiettivo della ricerca:**

Indagare se vi è relazione tra avere amici fumatori e scelta di iniziare a fumare.

### **5. Quadro teorico:**

L'abitudine al fumo è ancora molto diffusa in Italia. I dati relativi al 2018 del Rapporto nazionale sul fumo, presentato dall'Osservatorio Fumo, Alcol e Droga dell'Istituto Superiore di Sanità, ci dicono che in Italia fuma il 23,3% della popolazione; fuma inoltre un ragazzo/a su cinque nella fascia 13-15 anni. Tra i dati che è più interessante analizzare, c'è quello che esamina le ragioni che inducono a cominciare a fumare: la maggior parte dei fumatori (il 59,4%) dichiara di aver iniziato per influenza di compagni e amici in contesti di svago. In pratica, ci si trova in un gruppo dove gli altri fumano e si accende la sigaretta per non sentirsi esclusi.

Quali sono i meccanismi psicologici che favoriscono l'instaurarsi dell'abitudine al fumo? E quali i fattori che rendono così difficile la disassuefazione? OggiScienza lo ha chiesto a Gabriella Pravettoni, che dirige la divisione di Psiconcologia dell'Istituto Europeo di Oncologia di Milano. «Dobbiamo, innanzitutto, esplorare i complessi meccanismi che facilitano la creazione del legame tra individuo e sigaretta», sottolinea Pravettoni. «La componente fisiologica è principalmente legata alla nicotina contenuta nelle sigarette, che è in grado di creare, in un breve arco temporale, un'importante

dipendenza a livello fisico», aggiunge la dottoressa. Ma la dipendenza fisiologica non esaurisce assolutamente la questione: le componenti psicologiche dell'abitudine al fumo sono un ostacolo molto difficile da superare. Quali sono le ragioni che spingono a fumare?

«La sigaretta ricopre diverse funzioni», ricorda Pravettoni. «Da un lato può rappresentare un modo per gestire lo stress correlato agli eventi di vita e più in generale le emozioni negative. Si tratta di quello che gli psicologi definiscono 'meccanismo di coping maladattivo', cioè un modo per fronteggiare un problema che, però, in ultima analisi, peggiora la situazione. Allo stesso tempo la sigaretta può avere anche una funzione attivante, in quanto può essere usata dalla persona che fuma per favorire la concentrazione durante lo svolgimento di un compito, per aumentare la propria performance o per contrastare la sensazione di noia».

C'è poi la sigaretta che, come dicevamo, si accende in compagnia, per evitare di sentirsi esclusi dal gruppo dei pari. «Si tratta di un meccanismo che si può notare soprattutto nei giovani fumatori, nei quali il consumo di sigarette può essere anche favorito dalla tendenza a privilegiare i benefici immediati del fumo, come la percezione di sentirsi grandi, il piacere percettivo e gestuale legato alla sigaretta, perché si ha difficoltà a rappresentarsi le conseguenze future che il fumo ha sulla salute», sottolinea la dottoressa. Non di rado, l'ingresso alla scuola superiore o in una classe successiva coincide con la sperimentazione della sigaretta, che segna appunto il passaggio verso il mondo degli adulti.

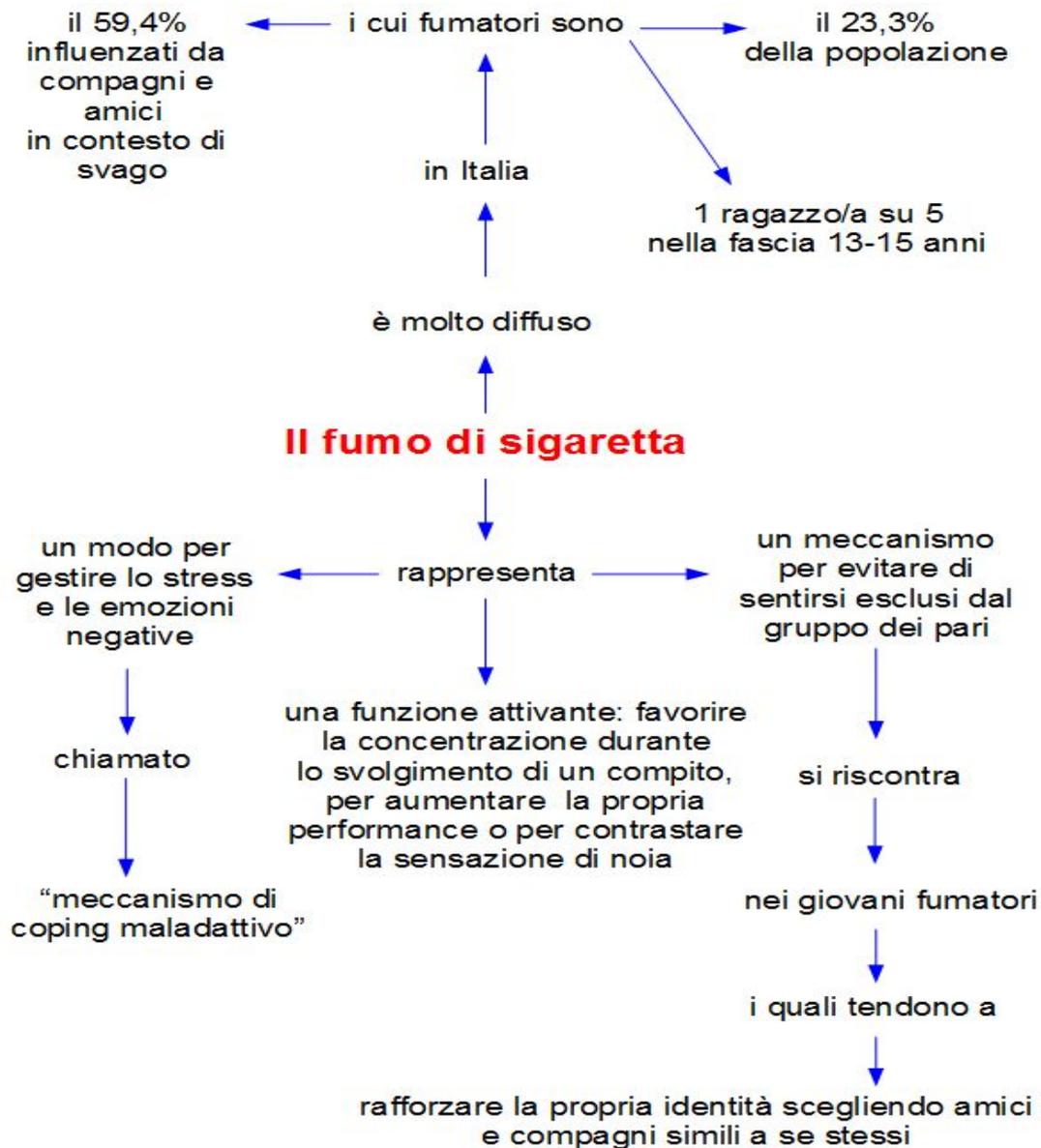
Il fumare rappresenta un comportamento, criticato dal punto di vista della salute ma accettato nel mondo degli adulti, pertanto gli adolescenti ritengono che il fumo sia il modo facile di potersi affermare nella società in qualità di adulti. Vi è un'alta correlazione tra tutti i comportamenti a rischio e ciò significa che non si presentano in forma isolata, ma piuttosto come una costellazione di comportamenti simili che portano a seguire un determinato stile di vita (Bonino, 2005). Nella decisione di iniziare a fumare, fondamentale è l'approvazione del gruppo e il fatto che gli amici fumino. Il fumo in questo frangente non è un comportamento solitario, ma di gruppo, tanto da portare i fumatori a non avere amici che non fumano. Ciò conferma il fatto che i giovani tendono a rafforzare la propria identità scegliendo amici e compagni simili a se stessi. Di conseguenza, i gruppi si costituiscono e

tendono a differenziarsi sempre più in base a questa caratteristica e crescendo i fumatori hanno un numero sempre più alto di amici fumatori. Il fumare inoltre sembra facilitare l'inserimento nel gruppo, al punto che i ragazzi che non fumano si sentono più facilmente tagliati fuori dalle attività svolte dai ragazzi della loro età, temono di non riuscire a farsi degli amici e si sentono socialmente più incerti.

Il fumo è un modo per fare cose da grandi e non più da bambini. In questo senso il fumo viene inteso come un rito di legame, come modalità ritualizzata di entrare in relazione con il gruppo, di unire i partecipanti e di accomunarli. Il fumo infatti condivide molti dei tratti dei comportamenti ritualizzati, caratterizzati da ridondanza, esagerazione e semplificazione del gesto. Pensiamo alla sequenza rituale del fumo: dalla richiesta o dall'offerta, all'accensione, all'inalazione, allo sbuffo, allo scambio della sigaretta accesa.

In conclusione, è chiaro come i giovani cerchino in ogni modo di mettere in atto comportamenti (che andranno a sedimentarsi nell'età adulta) che permettano loro di affermare la propria identità e di costruire una rete di relazioni sociali e affettive. Ci sono giovani che riescono a raggiungere tali obiettivi senza mettere in pericolo la propria vita, mentre altri optano per i comportamenti a rischio.

## 6. Mappa concettuale



## 7. Sitografia

<https://oggiscienza.it/2019/05/29/perche-fumiamo/>

<https://www.milleunadonna.it/benessere/articoli/adolescenti-motivazioni-e-abitudinal-fumo/>

## 8. Ipotesi di ricerca - Individuazione dei fattori e definizione operativa:

Chi ha amici fumatori è maggiormente predisposto ad iniziare a fumare.

Fattore indipendente: avere amici fumatori

Fattore dipendente: iniziare a fumare

FATTORI	INDICATORI	ITEM	VARIABILI
<b>Variabili di sfondo</b>	1. Genere	Sei maschio o femmina?	1.Maschio 2.Femmina
	2. Età	Quanti anni hai?	□□□
	3. Residenza	Vivi in città o in provincia?	1.Città 2.Provincia
	4. Abitazione	Vivi con i genitori?	1.Sì 2.No
<b>Iniziare a fumare</b>	5. Rapporto con il fumo	Sei un fumatore?	1.Sì 2.Fumo raramente 3.No
	6. Abitudine giornaliera	Quante sigarette fumi al giorno?	1.Meno di 10 2.Da 10 a 20 3.Più di un pacchetto
	7. Compagnia	Con chi fumi?	1. Da solo 2. In compagnia 3. Entrambi
	8. Luoghi	Dove sei solito fumare?	1.Pub/discoteche 2.In giro 3.A casa 4.E' indifferente
	9. Origine dell'abitudine	A che età hai iniziato a fumare?	1.Prima dei 14 anni 2.Tra 14 e 18 anni 3.Dopo i 18 anni
	10. Motivazione inizio	Quali sono i fattori che ti hanno spinto ad iniziare a fumare?	1.Provare esperienze nuove 2.Mi aiuta a rilassarmi 3.Mi piace 4.Per essere maggiormente inserito nel gruppo di amici
	11. Percezione del rischio	Secondo te, è davvero pericoloso fumare?	1.Sì, ma solo se ne si fa un uso eccessivo 2.Sì, anche fumando poche sigarette al giorno 3.Non lo so

	12. Volontà/ possibilità di smettere di fumare	Hai mai provato a smettere di fumare?	1.Sì, ma non ci sono riuscito/a 2.No, non voglio 3.Non ci ho mai pensato
<b>Amici fumatori</b>	13. Rapporto con il fumo	I tuoi amici fumano?	1.Sì, la maggior parte 2.Sì, ma solo qualcuno 3.No, nessuno
	14. Abitudine giornaliera	Quante sigarette fumano al giorno?	1.Meno di 10 2.Da 10 a 20 3.Più di un pacchetto 4.Non lo so
	15. Luoghi frequentati	Quali luoghi frequentate solitamente?	1.Locali 2.Discoteche 3.Piazzetta del paese 4.Casa di amici 5.Altro
	16. L'influenza del gruppo sul singolo	Credi che il gruppo di amici possa influenzare il singolo a fare uso di sigarette?	1.Sì 2.No

## **9. Individuazione del campione e tipologia di campionamento:**

Popolazione: ragazzi di età compresa tra i 14 e i 26 anni di Torino e provincia

Campione: un campionamento causale e corretto sarebbe stato ricavato da tutta la popolazione di riferimento di Torino e della sua provincia; noi consideriamo per ragioni economiche e temporali un campione accidentale di 100 soggetti a cui abbiamo somministrato un questionario online mediante il seguente link:

<https://www.farnt.unito.it/trincher/qgen/richiama.asp?codice=18149898>

Tipo di campionamento: campionamento probabilistico randomizzato

## 10. Tecniche e strumenti di rilevazione dei dati

La ricerca presentata è una ricerca standard. Per ottenere i dati necessari abbiamo utilizzato un questionario ad alta strutturazione, totalmente anonimo e con domande a risposta chiusa presentate con linguaggio semplice, partendo dal generale per arrivare al particolare. Il questionario è stato realizzato tramite il programma online Q-gen.

I dati ottenuti sono stati inseriti in una matrice dati su Excel ed elaborati con il programma JsStat.

## 11. Questionario anonimo

INFORMATIZZAZIONE E STRATEGIE DI APPRENDIMENTO.

Chiediamo la tua collaborazione a questa ricerca condotta presso il Dipartimento di Scienze dell'Educazione e della Formazione, Università degli Studi di Torino. Garantiamo che le risposte da te fornite rimarranno assolutamente anonime e verranno utilizzate esclusivamente per elaborazioni statistiche.

1. Sei maschio o femmina?

*maschio*

*femmina*

2. Quanti anni hai?

\_\_|\_\_

3. Dove vivi?

*città*

*provincia*

4. Vivi con i genitori?

*si*

*no*

5. Sei un fumatore?

*si*

*fumo raramente*

*no*

6. Quante sigarette fumi al giorno?

*meno di 10*

*da 10 a 20*

*più di un pacchetto*

7. Con chi fumi di solito?

*da solo*

*in compagnia*

*entrambi*

8. Dove sei solito fumare?

*pub/discoteche*

*in giro*

*a casa*

*è indifferente*

9. A che età hai iniziato a fumare?

*prima dei 14 anni*

*tra i 14 e 18 anni*

*dopo i 18 anni*

10. Quali sono i fattori che ti hanno spinto ad iniziare a fumare?

*per provare esperienze nuove*

*mi aiuta a rilassarmi*

*mi piace*

*per essere maggiormente inserito/a nel gruppo di amici*

11. Secondo te è davvero pericoloso fumare?

*sì, ma solo se ne si fa un uso eccessivo*

*sì, anche fumando poche sigarette al giorno*

*non lo so*

12. Hai mai provato a smettere di fumare?

*sì, e ci sono riuscito/a*

*sì, ma non ci sono riuscito/a*

*no, non voglio*

*non ci ho mai pensato*

13. I tuoi amici fumano?

*sì, la maggior parte*

*sì, ma solo qualcuno*

*no nessuno*

14. Quante sigarette fumano al giorno?

*meno di 10*

*da 10 a 20*

*più di un pacchetto*

*non lo so*

15. Quali luoghi frequentate solitamente?

*locali*

*discoteche*

*piazzetta del paese*

*casa di amici*

*altro*

16. Credi che il gruppo di amici possa influenzare il singolo a fare uso di sigarette?

*si*

*no*

## **12. Tecniche di analisi dei dati utilizzati**

Il questionario online è stato elaborato da noi stesse. Per l'analisi dei dati abbiamo suddiviso il questionario in tre parti:

- Prima parte: domande generali
- Seconda parte: domande personali relative ai destinatari del questionario
- Terza parte: domande personali relative agli amici dei destinatari del questionario

La matrice dati dove abbiamo inserito i dati del questionario è stata riportata nel programma JsStat (<http://www.edurete.org/jsstat/jsstat.htm>) per eseguire l'analisi monovariata delle variabili, individuando la distribuzione di frequenza, gli indici di tendenza centrale e gli indici di dispersione.

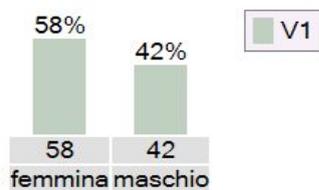
Successivamente abbiamo eseguito un'analisi bivariata su alcune domande incrociando le variabili indipendenti con le variabili dipendenti per poi stabilire tra quali variabili vi fosse relazione.

## Analisi monovariata:

### → *Analisi del sesso del campione*

Distribuzione di frequenza: *Sesso*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>femmina</b>	58	58%	58	58%	48%:68%
<b>maschio</b>	42	42%	100	100%	32%:52%



### **Campione:**

Numero di casi= 100

Indici di tendenza centrale:

Moda = femmina

Mediana = femmina

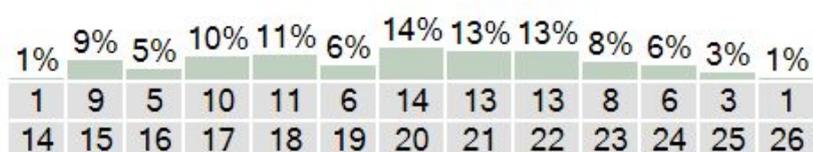
Indici di dispersione:

Squilibrio = 0.51

## → *Analisi dell'età del campione*

**Distribuzione di frequenza: Età**

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
14	1	1%	1	1%	0%:4%
15	9	9%	10	10%	3%:15%
16	5	5%	15	15%	0%:10%
17	10	10%	25	25%	4%:16%
18	11	11%	36	36%	5%:17%
19	6	6%	42	42%	1%:11%
20	14	14%	56	56%	7%:21%
21	13	13%	69	69%	6%:20%
22	13	13%	82	82%	6%:20%
23	8	8%	90	90%	3%:13%
24	6	6%	96	96%	1%:11%
25	3	3%	99	99%	0%:7%
26	1	1%	100	100%	0%:4%



### **Campione:**

Numero di casi= 100

Indici di tendenza centrale:

Moda = 20

Mediana = 20

Media = 19.79

Indici di dispersione:

Squilibrio = 0.1

Campo di variazione = 12

Differenza interquartilica = 4

Scarto tipo = 2.86

Indici di forma:

Asimmetria = -0.08

Curtosi = -0.86

### Popolazione:

Parametro	Int. Fid. 95%
Media	da 19.23 a 20.35
Scarto tipo	da 2.53 a 3.37

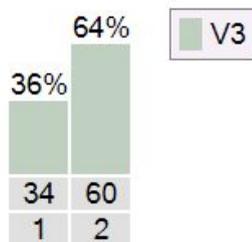
Probabilità di normalità della distribuzione (test di Jarque-Bera): 0.199

## → *Analisi della percentuale di abitazione*

### Distribuzione di frequenza:

*Vivi in città o in provincia?*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
1	34	36%	34	36%	26%:46%
2	60	64%	94	100%	54%:74%



### Campione:

Numero di casi = 94

Indici di tendenza centrale:

Moda = 2

Mediana = 2

Media = 1.64

Indici di dispersione:

Squilibrio = 0.54

Campo di variazione = 1

Differenza interquartilica = 1

Scarto tipo = 0.48

Indici di forma:

Asimmetria = -0.58

Curtosi = -1.67

### Popolazione:

Parametro	Int. Fid. 95%
Media	da 1.54 a 1.74
Scarto tipo	da 0.42 a 0.57

Probabilità di normalità della distribuzione (test di Jarque-Bera): 0

La definizione di indici posizionali ha senso solo quando la variabile in esame è almeno categoriale ordinata (ordinale). La tabella seguente riporta i dati ordinati con il rispettivo centile (il numero indica il punto corrispondente alla riga a sinistra del dato nella cella della tabella). La media e lo scarto tipo hanno senso solo quando calcolati su una variabile cardinale.

### Vivi in città o in provincia?

Indici posizione

Centile	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	43	44	45	46	47	48	49	50		
Dato	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

51	52	53	54	55	56	57	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	93	94	95	96	97	98	99	100						
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

La mediana (punto che lascia alla sua sinistra e alla sua destra lo stesso numero di casi) vale 2. Il primo quartile Q1 (punto che lascia alla sua sinistra il 25 percento dei casi) vale 1. Il terzo quartile Q3 (punto che lascia alla sua sinistra il 75 percento dei casi) vale 2. La differenza interquartilica Q3-Q1 vale 1.

Lo squilibrio è dato dalla somma delle proporzioni al quadrato per ciascuna delle k modalità della variabile, ossia:

$$\sum_{j=1}^k p_j^2 = 0.36^2 + 0.64^2 = 0.54.$$

E' un indice di dispersione dei casi nelle modalità assunte dalla variabile. Se è vicino a 0.5 (ossia 1/k, dove k è il numero delle modalità) i casi sono equidistribuiti nelle categorie corrispondenti alle modalità della variabile. Se è vicino a 1 i casi sono concentrati in un'unica categoria.

Il campo di variazione indica la differenza tra il valore minimo (1) e il valore massimo (2) della distribuzione.

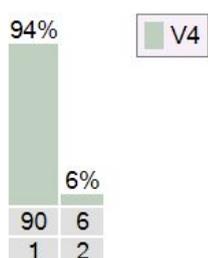


## → *Analisi della percentuale di convivenza coi genitori*

### Distribuzione di frequenza:

*Vivi coi tuoi genitori?*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
1	90	94%	90	94%	89%:99%
2	6	6%	96	100%	1%:11%



### Campione:

Numero di casi = 96

Indici di tendenza centrale:

Moda = 1

Mediana = 1

Media = 1.06

Indici di dispersione:

Squilibrio = 0.88

Campo di variazione = 1

Differenza interquartilica = 0

Scarto tipo = 0.24

Indici di forma:

Asimmetria = 3.61

Curtosi = 11.07

### Popolazione:

Parametro	Int. Fid. 95%
Media	da 1.01 a 1.11
Scarto tipo	da 0.21 a 0.29

Probabilità di normalità della distribuzione (test di Jarque-Bera): 0

La definizione di indici posizionali ha senso solo quando la variabile in esame è almeno categoriale ordinata (ordinale). La tabella seguente riporta i dati ordinati con il rispettivo



$$06)^2+(1-1.06)^2+(1-1.06)^2+(1-1.06)^2+(1-1.06)^2+(1-1.06)^2+(1-1.06)^2+(1-1.06)^2+(1-1.06)^2+(1-1.06)^2+(1-1.06)^2+(2-1.06)^2+(2-1.06)^2+(2-1.06)^2+(2-1.06)^2+(2-1.06)^2+(2-1.06)^2+(2-1.06)^2)/96 = 0.24$$

I parametri illustrati (percentuali semplici delle categorie, media e scarto tipo) sono quelli calcolati sul campione considerato (la matrice di dati di partenza). Di questi viene fornita anche la proiezione sulla popolazione, con intervallo di fiducia 95 percento (Int. Fid. 95%), valida se:

- a) il campione è stato ottenuto mediante estrazione casuale;
- b) la popolazione è normale se il numero dei casi del campione è minore di 30; se è maggiore la forma può essere qualsiasi.

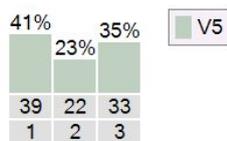
Media e varianza della popolazione vengono supposte ignote. Le proiezioni (stime per intervallo) ci dicono che il 95 percento dei campioni estratti da quella popolazione avranno parametri che si situano all'interno di quell'intervallo.

Per la stima della proporzione viene usata la distribuzione Binomiale quando la numerosità del campione è inferiore o uguale a 30 casi; se superiore viene usata la distribuzione di Poisson se la proporzione della categoria è 5% o inferiore, la distribuzione Normale altrimenti.

## → *Analisi della percentuale dei fumatori e non fumatori:*

**Distribuzione di frequenza:**  
*Sei un fumatore?*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>1</b>	39	41%	39	41%	32%:51%
<b>2</b>	22	23%	61	65%	15%:32%
<b>3</b>	33	35%	94	100%	25%:45%



**Campione:**

Numero di casi= 94

Indici di tendenza centrale:

Moda = 1

Mediana = 2

Media = 1.94

Indici di dispersione:



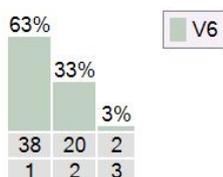


## → *Analisi della percentuale delle sigarette fumate*

### Distribuzione di frequenza:

Quante sigarette fumi al giorno?

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>1</b>	38	63%	38	63%	51%:76%
<b>2</b>	20	33%	58	97%	21%:45%
<b>3</b>	2	3%	60	100%	0%:10%



### Campione:

Numero di casi= 60

Indici di tendenza centrale:

Moda = 1

Mediana = 1

Media = 1.4

Indici di dispersione:

Squilibrio = 0.51

Campo di variazione = 2

Differenza interquartilica = 1

Scarto tipo = 0.55

Indici di forma:

Asimmetria = 0.99

Curtosi = -0.05

### Popolazione:

Parametro	Int. Fid. 95%
Media	da 1.26 a 1.54
Scarto tipo	da 0.47 a 0.69

Probabilità di normalità della distribuzione (test di Jarque-Bera): 0.007



anche la proiezione sulla popolazione, con intervallo di fiducia 95 percento (Int. Fid. 95%), valida se:

- a) il campione è stato ottenuto mediante estrazione casuale;
- b) la popolazione è normale se il numero dei casi del campione è minore di 30; se è maggiore la forma può essere qualsiasi.

Media e varianza della popolazione vengono supposte ignote. Le proiezioni (stime per intervallo) ci dicono che il 95 percento dei campioni estratti da quella popolazione avranno parametri che si situano all'interno di quell'intervallo.

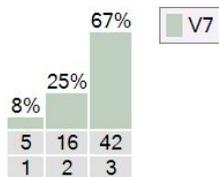
Per la stima della proporzione viene usata la distribuzione Binomiale quando la numerosità del campione è inferiore o uguale a 30 casi; se superiore viene usata la distribuzione di Poisson se la proporzione della categoria è 5% o inferiore, la distribuzione Normale altrimenti.

## *→ Analisi della percentuale del fumare in compagnia o da soli*

### Distribuzione di frequenza:

*Con chi fumi di solito?*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>1</b>	5	8%	5	8%	1%:15%
<b>2</b>	16	25%	21	33%	15%:36%
<b>3</b>	42	67%	63	100%	55%:78%



### Campione:

Numero di casi= 63

Indici di tendenza centrale:

Moda = 3

Mediana = 3

Media = 2.59

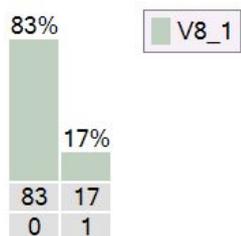
Indici di dispersione:

Squilibrio = 0.52

Campo di variazione = 2







**Campione:**

Numero di casi= 100

Indici di tendenza centrale:

Moda = 0

Mediana = V8\_1

Media = 0.17

Indici di dispersione:

Squilibrio = 0.72

Campo di variazione = 1

Differenza interquartilica = 0

Scarto tipo = 0.38

Indici di forma:

Asimmetria = 1.76

Curtosi = 1.09

**Popolazione:**

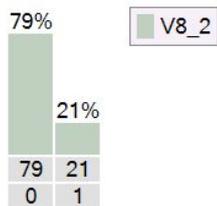
Parametro	Int. Fid. 95%
Media	da 0.1 a 0.24
Scarto tipo	da 0.33 a 0.44

Probabilità di normalità della distribuzione (test di Jarque-Bera): 0

• **Distribuzione di frequenza:**

*In giro*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
0	79	79%	79	79%	71%:87%
1	21	21%	100	100%	13%:29%



**Campione:**

Numero di casi= 100

Indici di tendenza centrale:

Moda = 0

Mediana = V8\_2

Media = 0.21

Indici di dispersione:

Squilibrio = 0.67

Campo di variazione = 1

Differenza interquartilica = 0

Scarto tipo = 0.41

Indici di forma:

Asimmetria = 1.42

Curtosi = 0.03

**Popolazione:**

Parametro	Int. Fid. 95%
Media	da 0.13 a 0.29
Scarto tipo	da 0.36 a 0.48

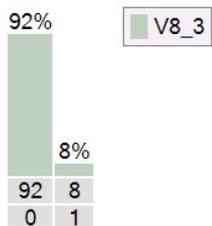
Probabilità di normalità della distribuzione (test di Jarque-Bera): 0

- Distribuzione di frequenza:**

*A casa*

**V8\_3**

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>0</b>	92	92%	92	92%	87%:97%
<b>1</b>	8	8%	100	100%	3%:13%



**Campione:**

Numero di casi= 100

Indici di tendenza centrale:

Moda = 0

Mediana = V8\_3

Media = 0.08

Indici di dispersione:

Squilibrio = 0.85

Campo di variazione = 1

Differenza interquartilica = 0

Scarto tipo = 0.27

Indici di forma:

Asimmetria = 3.1

Curtosi = 7.59

**Popolazione:**

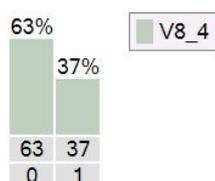
Parametro	Int. Fid. 95%
Media	da 0.03 a 0.13
Scarto tipo	da 0.24 a 0.32

Probabilità di normalità della distribuzione (test di Jarque-Bera): 0

- Distribuzione di frequenza:**

*E' indifferente*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>0</b>	63	63%	63	63%	54%:72%
<b>1</b>	37	37%	100	100%	28%:46%



**Campione:**

Numero di casi= 100

Indici di tendenza centrale:

Moda = 0

Mediana = V8\_4

Media = 0.37

Indici di dispersione:

Squilibrio = 0.53

Campo di variazione = 1

Differenza interquartilica = 1

Scarto tipo = 0.48

Indici di forma:

Asimmetria = 0.54

Curtosi = -1.71

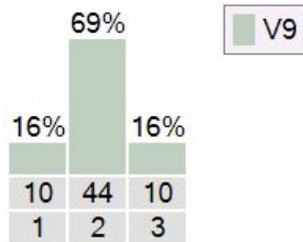
**Popolazione:**

Parametro	Int. Fid. 95%
Media	da 0.28 a 0.46
Scarto tipo	da 0.43 a 0.57

Probabilità di normalità della distribuzione (test di Jarque-Bera): 0

**→ *Analisi dell'età di inizio a fumare*****Distribuzione di frequenza:***A che età hai iniziato a fumare?*

<i>Modalità</i>	<i>Frequenza semplice</i>	<i>Percent. semplice</i>	<i>Frequenza cumulata</i>	<i>Percent. cumulata</i>	<i>Int. Fid. 95%</i>
<b>1</b>	10	16%	10	16%	7%:25%
<b>2</b>	44	69%	54	84%	57%:80%
<b>3</b>	10	16%	64	100%	7%:25%



**Campione:**

Numero di casi= 64

Indici di tendenza centrale:

Moda = 2

Mediana = 2

Media = 2

Indici di dispersione:

Squilibrio = 0.52

Campo di variazione = 2

Differenza interquartilica = 0

Scarto tipo = 0.56

Indici di forma:

Asimmetria = 0

Curtosi = 0.2

**Popolazione:**

Parametro	Int. Fid. 95%
Media	da 1.86 a 2.14
Scarto tipo	da 0.48 a 0.69

Probabilità di normalità della distribuzione (test di Jarque-Bera): 0.948

La definizione di indici posizionali ha senso solo quando la variabile in esame è almeno categoriale ordinata (ordinale). La tabella seguente riporta i dati ordinati con il rispettivo centile (il numero indica il punto corrispondente alla riga a sinistra del dato nella cella della tabella). La media e lo scarto tipo hanno senso solo quando calcolati su una variabile cardinale.

**A che età hai iniziato a fumare?**

Indici posizionali

Centile	0	2	3	5	6	8	9	11	13	14	16	17	19	20	22	23	25	27	28	30	31	33	34	36	38	39	41	42	44	45	47	48	50	
Dato	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2

	52	53	55	56	58	59	61	63	64	66	67	69	70	72	73	75	77	78	80	81	83	84	86	88	89	91	92	94	95	97	98	100	
	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2



b) la popolazione è normale se il numero dei casi del campione è minore di 30; se è maggiore la forma può essere qualsiasi.

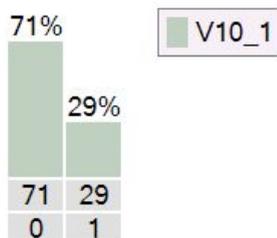
Media e varianza della popolazione vengono supposte ignote. Le proiezioni (stime per intervallo) ci dicono che il 95 percento dei campioni estratti da quella popolazione avranno parametri che si situano all'interno di quell'intervallo.

Per la stima della proporzione viene usata la distribuzione Binomiale quando la numerosità del campione è inferiore o uguale a 30 casi; se superiore viene usata la distribuzione di Poisson se la proporzione della categoria è 5% o inferiore, la distribuzione Normale altrimenti.

**→ *Analisi dei fattori che hanno spinto il campione a fumare***  
*Quali sono i fattori che ti hanno spinto ad iniziare a fumare?*

- **Distribuzione di frequenza:** *Per provare esperienze nuove*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>0</b>	71	71%	71	71%	62%:80%
<b>1</b>	29	29%	100	100%	20%:38%



**Campione:**

Numero di casi= 100

Indici di tendenza centrale:

Moda = 0

Mediana = V10\_1

Media = 0.29

Indici di dispersione:

Squilibrio = 0.59

Campo di variazione = 1

Differenza interquartilica = 1

Scarto tipo = 0.45

Indici di forma:

Asimmetria = 0.93

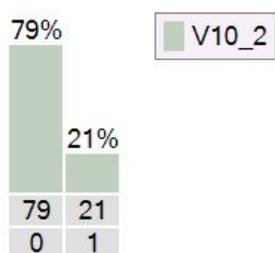
Curtosi = -1.14

**Popolazione:**

Parametro	Int. Fid. 95%
Media	da 0.2 a 0.38
Scarto tipo	da 0.4 a 0.53

- **Distribuzione di frequenza: *mi aiuta a rilassarmi***

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>0</b>	79	79%	79	79%	71%:87%
<b>1</b>	21	21%	100	100%	13%:29%

**Campione:**

Numero di casi= 100

Indici di tendenza centrale:

Moda = 0

Mediana = V10\_2

Media = 0.21

Indici di dispersione:

Squilibrio = 0.67

Campo di variazione = 1

Differenza interquartilica = 0

Scarto tipo = 0.41

Indici di forma:

Asimmetria = 1.42

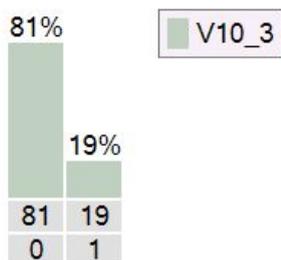
Curtosi = 0.03

**Popolazione:**

Parametro	Int. Fid. 95%
Media	da 0.13 a 0.29
Scarto tipo	da 0.36 a 0.48

- **Distribuzione di frequenza: *mi piace***

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>0</b>	81	81%	81	81%	73%:89%
<b>1</b>	19	19%	100	100%	11%:27%



**Campione:**

Numero di casi= 100

Indici di tendenza centrale:

Moda = 0

Mediana = V10\_3

Media = 0.19

Indici di dispersione:

Squilibrio = 0.69

Campo di variazione = 1

Differenza interquartilica = 0

Scarto tipo = 0.39

Indici di forma:

Asimmetria = 1.58

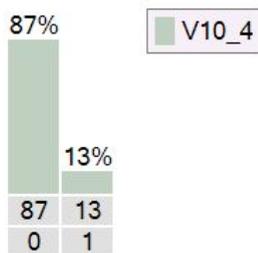
Curtosi = 0.5

**Popolazione:**

Parametro	Int. Fid. 95%
Media	da 0.11 a 0.27
Scarto tipo	da 0.35 a 0.46

- **Distribuzione di frequenza:** per essere maggiormente inserito nel gruppo di amici

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>0</b>	87	87%	87	87%	80%:94%
<b>1</b>	13	13%	100	100%	6%:20%



### Campione:

Numero di casi = 100

Indici di tendenza centrale:

Moda = 0

Mediana = V10\_4

Media = 0.13

Indici di dispersione:

Squilibrio = 0.77

Campo di variazione = 1

Differenza interquartilica = 0

Scarto tipo = 0.34

Indici di forma:

Asimmetria = 2.2

Curtosi = 2.84

### Popolazione:

Parametro	Int. Fid. 95%
-----------	---------------

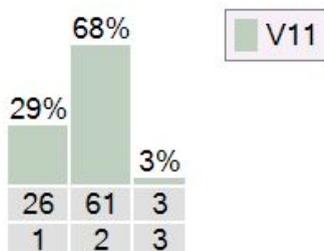
Media	da 0.06 a 0.2
Scarto tipo	da 0.3 a 0.4

**→ Analisi dell'opinione del campione secondo cui è veramente pericoloso fumare**

**Distribuzione di frequenza:**

Secondo te è davvero pericoloso fumare?

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
1	26	29%	26	29%	20%:38%
2	61	68%	87	97%	58%:77%
3	3	3%	90	100%	0%:8%



**Campione:**

Numero di casi= 90

Indici di tendenza centrale:

Moda = 2

Mediana = 2

Media = 1.74

Indici di dispersione:

Squilibrio = 0.54

Campo di variazione = 2

Differenza interquartilica = 1

Scarto tipo = 0.51

Indici di forma:

Asimmetria = -0.32

Curtosi = -0.36

**Popolazione:**

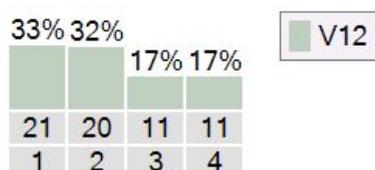




### Distribuzione di frequenza:

*Hai mai provato a smettere di fumare?*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
1	21	33%	21	33%	22%:45%
2	20	32%	41	65%	20%:43%
3	11	17%	52	83%	8%:27%
4	11	17%	63	100%	8%:27%



#### Campione:

Numero di casi= 63

Indici di tendenza centrale:

Moda = 1

Mediana = 2

Media = 2.19

Indici di dispersione:

Squilibrio = 0.27

Campo di variazione = 3

Differenza interquartilica = 2

Scarto tipo = 1.08

Indici di forma:

Asimmetria = 0.44

Curtosi = -1.09

#### Popolazione:

Parametro	Int. Fid. 95%
Media	da 1.92 a 2.46
Scarto tipo	da 0.93 a 1.34

Probabilità di normalità della distribuzione (test di Jarque-Bera): 0.075

La definizione di indici posizionali ha senso solo quando la variabile in esame è almeno categoriale ordinata (ordinale). La tabella seguente riporta i dati ordinati con il rispettivo centile (il numero indica il punto corrispondente alla riga a sinistra del dato nella cella della tabella). La media e lo scarto tipo hanno senso solo quando calcolati su una variabile cardinale.

### Hai mai provato a smettere di fumare?

Indici posizionali

Centile	0	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27	29	30	32	33	35	37	38	40	41	43	44	46	48	49
Dato	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2

51	52	54	56	57	59	60	62	63	65	67	68	70	71	73	75	76	78	79	81	83	84	86	87	89	90	92	94	95	97	98	100
2	2	2	2	2	2	2	2	3	3	3	3	3	3	3	3	3	3	3	3	4	4	4	4	4	4	4	4	4	4	4	4

La mediana (punto che lascia alla sua sinistra e alla sua destra lo stesso numero di casi) vale 2. Il primo quartile Q1 (punto che lascia alla sua sinistra il 25 percento dei casi) vale 1. Il terzo quartile Q3 (punto che lascia alla sua sinistra il 75 percento dei casi) vale 3. La differenza interquartilica Q3-Q1 vale 2.

Lo squilibrio è dato dalla somma delle proporzioni al quadrato per ciascuna delle k modalità della variabile, ossia:

$$\sum_{j=1}^k p_j^2 = 0.33^2 + 0.32^2 + 0.17^2 + 0.17^2 = 0.27.$$

E' un indice di dispersione dei casi nelle modalità assunte dalla variabile. Se è vicino a 0.25 (ossia 1/k, dove k è il numero delle modalità) i casi sono equidistribuiti nelle categorie corrispondenti alle modalità della variabile. Se è vicino a 1 i casi sono concentrati in un'unica categoria.

Il campo di variazione indica la differenza tra il valore minimo (1) e il valore massimo (4) della distribuzione.

La media (aritmetica) è data dalla somma dei valori corrispondenti a ciascun caso divisa per il numero dei casi, ossia:

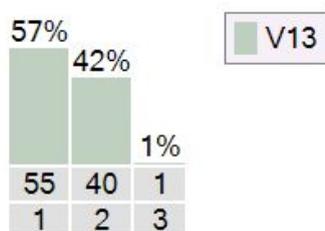


## → *Analisi della percentuale degli amici fumatori*

### Distribuzione di frequenza:

*I tuoi amici fumano?*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>1</b>	55	57%	55	57%	47%:67%
<b>2</b>	40	42%	95	99%	32%:52%
<b>3</b>	1	1%	96	100%	0%:4%



### Campione:

Numero di casi = 96

Indici di tendenza centrale:

Moda = 1

Mediana = 1

Media = 1.44

Indici di dispersione:

Squilibrio = 0.5

Campo di variazione = 2

Differenza interquartilica = 1

Scarto tipo = 0.52

Indici di forma:

Asimmetria = 0.48

Curtosi = -1.25

### Popolazione:

Parametro	Int. Fid. 95%
Media	da 1.33 a 1.54
Scarto tipo	da 0.46 a 0.61



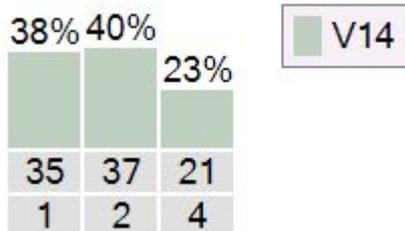


## **→ Analisi della percentuale di sigarette fumate dagli amici**

**Distribuzione di frequenza:**

Quante sigarette fumano al giorno?

Modalità	Frequenza semplice	Percentuale semplice	Frequenza cumulata	Percentuale cumulata	Int. Fid. 95%
<b>1</b>	35	38%	35	38%	28%:47%
<b>2</b>	37	40%	72	77%	30%:50%
<b>4</b>	21	23%	93	100%	14%:31%



**Campione:**

Numero di casi= 93

Indici di tendenza centrale:

Moda = 2

Mediana = 2

Media = 2.08

Indici di dispersione:

Squilibrio = 0.35

Campo di variazione = 3

Differenza interquartilica = 1

Scarto tipo = 1.13

Indici di forma:

Asimmetria = 0.79

Curtosi = -0.78

**Popolazione:**

Parametro	Int. Fid. 95%
Media	da 1.85 a 2.3
Scarto tipo	da 0.99 a 1.34

Probabilità di normalità della distribuzione (test di Jarque-Bera): 0.002

---

La definizione di indici posizionali ha senso solo quando la variabile in esame è almeno categoriale ordinata (ordinale). La tabella seguente riporta i dati ordinati con il rispettivo centile (il numero indica il punto corrispondente alla riga a sinistra del dato nella cella della tabella). La media e lo scarto tipo hanno senso solo quando calcolati su una variabile cardinale.

**Quante sigarette fumano al giorno?**

*Indici posizionali*

Centile	0	1	2	3	4	5	6	8	9	10	11	12	13	14	15	16	17	18	19	20	22	23	24	25	26	27	28	29	30	31	32	33	34	35	37	38	39	40	41	42	43	44	45	46	47	48	49			
Dato	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2

51	52	53	54	55	56	57	58	59	60	61	62	63	65	66	67	68	69	70	71	72	73	74	75	76	77	78	80	81	82	83	84	85	86	87	88	89	90	91	92	94	95	96	97	98	99	100					
2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4

La mediana (punto che lascia alla sua sinistra e alla sua destra lo stesso numero di casi) vale 2. Il primo quartile Q1 (punto che lascia alla sua sinistra il 25 percento dei casi) vale 1. Il terzo quartile Q3 (punto che lascia alla sua sinistra il 75 percento dei casi) vale 2. La differenza interquartilica Q3-Q1 vale 1.



anche la proiezione sulla popolazione, con intervallo di fiducia 95 percento (Int. Fid. 95%), valida se:

- a) il campione è stato ottenuto mediante estrazione casuale;
- b) la popolazione è normale se il numero dei casi del campione è minore di 30; se è maggiore la forma può essere qualsiasi.

Media e varianza della popolazione vengono supposte ignote. Le proiezioni (stime per intervallo) ci dicono che il 95 percento dei campioni estratti da quella popolazione avranno parametri che si situano all'interno di quell'intervallo.

Per la stima della proporzione viene usata la distribuzione Binomiale quando la numerosità del campione è inferiore o uguale a 30 casi; se superiore viene usata la distribuzione di Poisson se la proporzione della categoria è 5% o inferiore, la distribuzione Normale altrimenti.

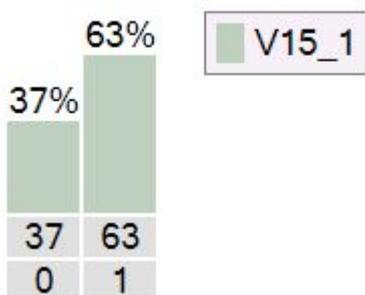
### **→ Analisi della percentuale dei luoghi frequentati solitamente**

*Quali luoghi frequentate solitamente?*

- **Distribuzione di frequenza:**

*locali*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>0</b>	37	37%	37	37%	28%:46%
<b>1</b>	63	63%	100	100%	54%:72%



**Campione:**

Numero di casi= 100

Indici di tendenza centrale:

Moda = 1

Mediana = 1

Media = 0.63

Indici di dispersione:

Squilibrio = 0.53

Campo di variazione = 1

Differenza interquartilica = 1

Scarto tipo = 0.48

Indici di forma:

Asimmetria = -0.54

Curtosi = -1.71

**Popolazione:**

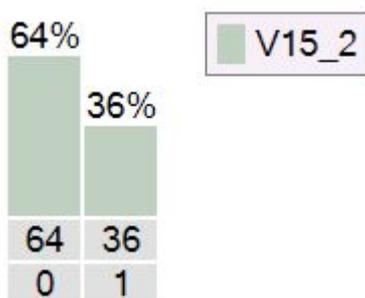
Parametro	Int. Fid. 95%
Media	da 0.54 a 0.72
Scarto tipo	da 0.43 a 0.57

Probabilità di normalità della distribuzione (test di Jarque-Bera): 0

- Distribuzione di frequenza:**

*discoteche*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>0</b>	64	64%	64	64%	55%:73%
<b>1</b>	36	36%	100	100%	27%:45%



**Campione:**

Numero di casi= 100

Indici di tendenza centrale:

Moda = 0

Mediana = V15\_2

Media = 0.36

Indici di dispersione:

Squilibrio = 0.54

Campo di variazione = 1

Differenza interquartilica = 1

Scarto tipo = 0.48

Indici di forma:

Asimmetria = 0.58

Curtosi = -1.66

**Popolazione:**

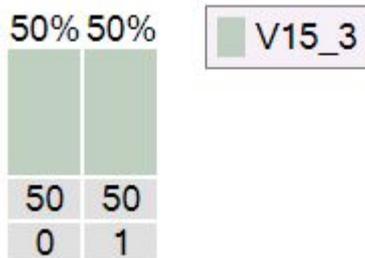
Parametro	Int. Fid. 95%
Media	da 0.27 a 0.45
Scarto tipo	da 0.42 a 0.57

Probabilità di normalità della distribuzione (test di Jarque-Bera): 0

- Distribuzione di frequenza:**

*piazzetta del paese*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>0</b>	50	50%	50	50%	40%:60%
<b>1</b>	50	50%	100	100%	40%:60%

**Campione:**

Numero di casi= 100

Indici di tendenza centrale:

Moda = 0; 1

Mediana = tra V15\_3 e 1

Media = 0.5

Indici di dispersione:

Squilibrio = 0.5

Campo di variazione = 1

Differenza interquartilica = 1

Scarto tipo = 0.5

Indici di forma:

Asimmetria = 0

Curtosi = -2

### Popolazione:

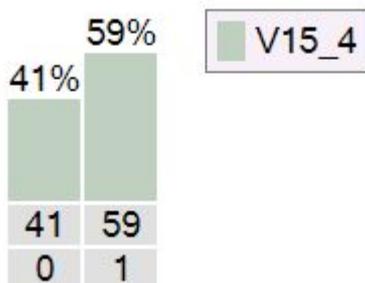
Parametro	Int. Fid. 95%
Media	da 0.4 a 0.6
Scarto tipo	da 0.44 a 0.59

Probabilità di normalità della distribuzione (test di Jarque-Bera): 0

- **Distribuzione di frequenza:**

*casa di amici*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>0</b>	41	41%	41	41%	31%:51%
<b>1</b>	59	59%	100	100%	49%:69%



**Campione:**

Numero di casi= 100

Indici di tendenza centrale:

Moda = 1

Mediana = 1

Media = 0.59

Indici di dispersione:

Squilibrio = 0.52

Campo di variazione = 1

Differenza interquartilica = 1

Scarto tipo = 0.49

Indici di forma:

Asimmetria = -0.37

Curtosi = -1.87

### Popolazione:

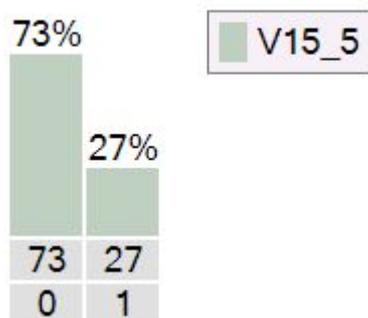
Parametro	Int. Fid. 95%
Media	da 0.49 a 0.69
Scarto tipo	da 0.43 a 0.58

Probabilità di normalità della distribuzione (test di Jarque-Bera): 0

- **Distribuzione di frequenza:**

*altro*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>0</b>	73	73%	73	73%	64%:82%
<b>1</b>	27	27%	100	100%	18%:36%



**Campione:**

Numero di casi= 100

Indici di tendenza centrale:

Moda = 0

Mediana = V15\_5

Media = 0.27

Indici di dispersione:

Squilibrio = 0.61

Campo di variazione = 1

Differenza interquartilica = 1

Scarto tipo = 0.44

Indici di forma:

Asimmetria = 1.04

Curtosi = -0.93

**Popolazione:**

Parametro	Int. Fid. 95%
Media	da 0.18 a 0.36
Scarto tipo	da 0.39 a 0.52

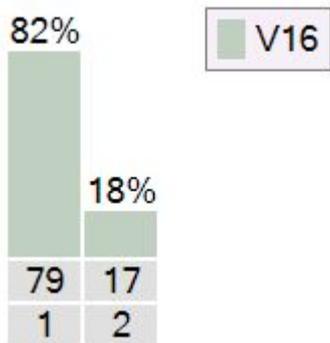
Probabilità di normalità della distribuzione (test di Jarque-Bera): 0

**→ *Analisi dell'opinione del campione secondo cui il gruppo di amici influenza o meno il singolo a fare uso di sigarette***

**Distribuzione di frequenza:**

*Credi che il gruppo di amici possa influenzare il singolo a far uso di sigarette?*

Modalità	Frequenza semplice	Percent. semplice	Frequenza cumulata	Percent. cumulata	Int. Fid. 95%
<b>1</b>	79	82%	79	82%	75%:90%
<b>2</b>	17	18%	96	100%	10%:25%



**Campione:**

Numero di casi= 96

Indici di tendenza centrale:

Moda = 1

Mediana = 1

Media = 1.18

Indici di dispersione:

Squilibrio = 0.71

Campo di variazione = 1

Differenza interquartilica = 0

Scarto tipo = 0.38

Indici di forma:

Asimmetria = 1.69

Curtosi = 0.86

**Popolazione:**

Parametro	Int. Fid. 95%
Media	da 1.1 a 1.25
Scarto tipo	da 0.34 a 0.45

Probabilità di normalità della distribuzione (test di Jarque-Bera): 0

La definizione di indici posizionali ha senso solo quando la variabile in esame è almeno categoriale ordinata (ordinale). La tabella seguente riporta i dati ordinati con il rispettivo centile (il numero indica il punto corrispondente alla riga a sinistra del dato nella cella della tabella). La media e lo scarto tipo hanno senso solo quando calcolati su una variabile cardinale.

**V16**



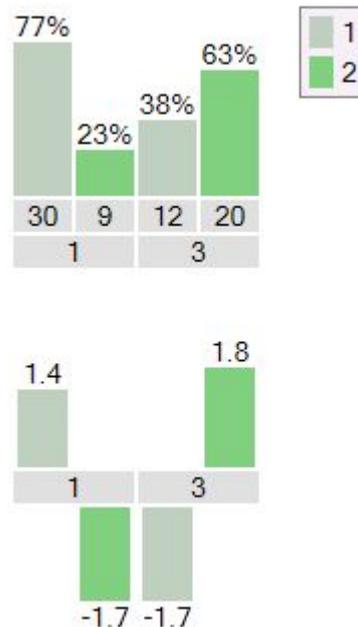


## Analisi bivariata:

### → Relazione tra l'essere un fumatore e l'aver amici fumatori

Tabella a doppia entrata:  
V5 x V13

V13-> V5	1	2	Marginale di riga
1	30 23.3 1.4	9 15.7 -1.7	39
3	12 19.7 -1.7	20 13.3 1.8	33
Marginale di colonna	43	29	72



X quadro = 11.2. Significatività = **0.001**

V di Cramer = 0.39

Probabilità esatta (dal test di Fisher) = **0.014**

Nelle celle della tabella sono indicati:

- la frequenza osservata O
- la frequenza attesa A
- il residuo standardizzato di cella, ossia lo scarto tra frequenza osservata e attesa rapportato alla radice quadrata della frequenza attesa  $(O-A)/\text{rad}q(A)$

La tabella a doppia entrata riporta la distribuzione congiunta delle due variabili. I dati del campione ci danno, per ogni cella:

- La frequenza osservata  $O_i$  ossia il numero di casi che hanno quei dati valori sulle variabili considerate.
- La frequenza attesa  $A_i$ , ossia la frequenza che avremmo osservato nella cella se non vi fosse relazione tra le due variabili. In caso contrario potrebbero essere presenti 'addensamenti' di casi in alcune celle della tabella, dovuti ad 'attrazione' tra determinate modalità delle due variabili. La frequenza attesa deriva da una semplice

proporzione: se non vi è attrazione tra le modalità delle due variabili, il numero di casi in una cella dovrebbe avere la stessa proporzione rispetto al suo marginale di riga che ha il suo marginale di colonna rispetto al totale dei casi, ossia

$A_i$  : marginale di riga = marginale di colonna : totale dei casi  
da cui deriva che

$A_i = (\text{marginale di riga} * \text{marginale di colonna}) / \text{numero di casi}$

Ovviamente quanto più le frequenze osservate si discostano dalle frequenze attese tanto più è probabile che vi sia attrazione tra le singole modalità delle due variabili e quindi vi sia una relazione tra le variabili stesse.

E' possibile, cella per cella, definire un indice dato dalla differenza tra la frequenza osservata

e la frequenza attesa, rapportata alla frequenza attesa  $\frac{(O_i - A_i)}{A_i}$ . Quanto più è alto e positivo questo indice tanto più si può dire vi sia attrazione tra le modalità corrispondenti alla cella. Quanto più è alto e negativo tanto più si può dire che vi sia repulsione tra le modalità corrispondenti alla cella. Questo indice non va utilizzato se la frequenza attesa è inferiore a 1, dato che il valore diventa artificialmente alto perché il denominatore è inferiore a 1. Se ciò si verifica è utile accorpate i casi per evitare di avere frequenze marginali troppo basse (che portano a frequenze attese basse), oppure escludere le modalità corrispondenti dall'elaborazione.

E' possibile poi calcolare un indice complessivo, detto X quadro, come somma, cella per cella, della differenza tra la frequenza osservata e la frequenza attesa (elevata al quadrato per evitare che gli addendi di segno negativo elidano quelli di segno positivo) rapportata alla

frequenza attesa della singola cella:  $\chi^2 = \sum_{i=1}^r \frac{(O_i - A_i)^2}{A_i}$ . Quanto più è alto X quadro, tanto più è forte la relazione tra le due variabili. Per le ragioni illustrate anche questo indice non può essere applicato quando sono presenti frequenze attese inferiori a 1.

In questo caso il valore di X quadro è  $\chi^2 = \sum_{i=1}^r \frac{(O_i - A_i)^2}{A_i} = ((30-23.3)^2)/23.3 + ((9-15.7)^2)/15.7 + ((12-19.7)^2)/19.7 + ((20-13.3)^2)/13.3 = 11.2$ . La probabilità che X quadro sia diverso da zero per effetto del caso è di 0. Il valore è calcolato sulla distribuzione di probabilità Chi quadro con 1 grado/i di libertà, in corrispondenza dell'ascissa 11.2 (area a destra di tale punto).

Quando questo valore di probabilità (detto significatività della relazione) è inferiore a 0,05 si può iniziare a supporre lecitamente che vi sia una relazione significativa tra le due variabili.

### **Vi è quindi relazione tra le due variabili (a livello di fiducia 0,05)**

Il residuo standardizzato di cella indica se la differenza tra la frequenza osservata e la frequenza attesa in quella data cella è piccola o grande. Quando il numero totale dei casi è maggiore di 30, il residuo standardizzato può essere letto come un punteggio z. Se è superiore a 2 (in modulo) indica una differenza significativa (a livello di fiducia 0,05 corrispondente a  $z=1,96$ ) tra la frequenza osservata e attesa nella cella relativa. Se è negativo, nella cella abbiamo meno soggetti di quanti potremmo aspettarci se non vi fosse

relazione tra le due variabili; se è positivo, abbiamo più soggetti di quanti potremmo aspettarci se non vi fosse relazione tra le due variabili.

Il V di Cramer indica la forza della relazione tra le due variabili. Viene ottenuto rapportando X quadro al massimo che può assumere ed estraendo la radice quadrata, ossia

$$V = \sqrt{\frac{\chi^2}{n \cdot [\min(\#r, \#c) - 1]}}$$

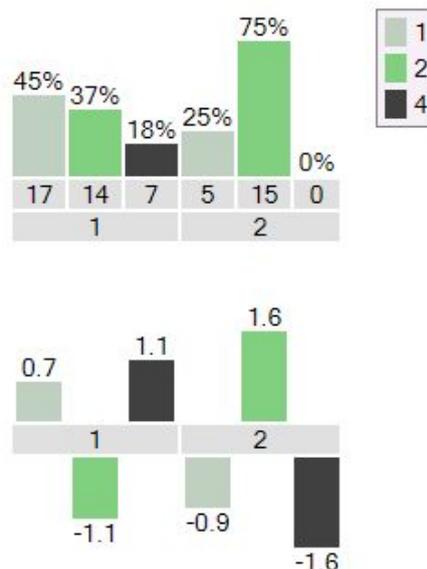
= 0.39, dove N è il numero dei casi, #r il numero di righe, #c il numero di colonne. Varia tra 0 (minima forza della relazione) e 1 (massima forza della relazione).

Il test esatto di Fisher ci consente di calcolare la probabilità esatta che non vi sia relazione tra le due variabili. Può essere applicato solo con tabelle 2x2 (due righe e due colonne), indipendentemente dal valore delle frequenze attese della tabella. In questo caso il valore è 0.01. **La probabilità calcolata con il test esatto di Fisher ci dice che vi è relazione tra le due variabili.**

## → Relazione tra quante sigarette fuma il campione e quante ne fumano gli amici

**Tabella a doppia entrata:  
V6 x V14**

V14-> V6	1	2	4	Marginale di riga
<b>1</b>	17 14.4 0.7	14 19 -1.1	7 4.6 1.1	38
<b>2</b>	5 7.6 -0.9	15 10 1.6	0 2.4 -1.6	20
Marginale di colonna	22	29	7	58



X quadro = 8.85. Significatività = 0.012

V di Cramer = 0.39

Nelle celle della tabella sono indicati:

- la frequenza osservata O
- la frequenza attesa A
- il residuo standardizzato di cella, ossia lo scarto tra frequenza osservata e attesa rapportato alla radice quadrata della frequenza attesa  $(O-A)/\text{radq}(A)$

La tabella a doppia entrata riporta la distribuzione congiunta delle due variabili. I dati del campione ci danno, per ogni cella:

- La frequenza osservata  $O_i$ , ossia il numero di casi che hanno quei dati valori sulle variabili considerate.
- La frequenza attesa  $A_i$ , ossia la frequenza che avremmo osservato nella cella se non vi fosse relazione tra le due variabili. In caso contrario potrebbero essere presenti 'addensamenti' di casi in alcune celle della tabella, dovuti ad 'attrazione' tra determinate modalità delle due variabili. La frequenza attesa deriva da una semplice proporzione: se non vi è attrazione tra le modalità delle due variabili, il numero di casi in una cella dovrebbe avere la stessa proporzione rispetto al suo marginale di riga che ha il suo marginale di colonna rispetto al totale dei casi, ossia  $A_i$  : marginale di riga = marginale di colonna : totale dei casi da cui deriva che  $A_i = (\text{marginale di riga} * \text{marginale di colonna}) / \text{numero di casi}$ . Ovviamente quanto più le frequenze osservate si discostano dalle frequenze attese tanto più è probabile che vi sia attrazione tra le singole modalità delle due variabili e quindi vi sia una relazione tra le variabili stesse.

E' possibile, cella per cella, definire un indice dato dalla differenza tra la frequenza osservata

e la frequenza attesa, rapportata alla frequenza attesa  $\frac{(O_i - A_i)}{A_i}$ . Quanto più è alto e positivo questo indice tanto più si può dire vi sia attrazione tra le modalità corrispondenti alla cella. Quanto più è alto e negativo tanto più si può dire che vi sia repulsione tra le modalità corrispondenti alla cella. Questo indice non va utilizzato se la frequenza attesa è inferiore a 1, dato che il valore diventa artificialmente alto perché il denominatore è inferiore a 1. Se ciò si verifica è utile accorpate i casi per evitare di avere frequenze marginali troppo basse (che portano a frequenze attese basse), oppure escludere le modalità corrispondenti dall'elaborazione.

E' possibile poi calcolare un indice complessivo, detto X quadro, come somma, cella per cella, della differenza tra la frequenza osservata e la frequenza attesa (elevata al quadrato per evitare che gli addendi di segno negativo elidano quelli di segno positivo) rapportata alla

frequenza attesa della singola cella:  $\chi^2 = \sum_{i=1}^r \frac{(O_i - A_i)^2}{A_i}$ . Quanto più è alto X quadro, tanto più è forte la relazione tra le due variabili. Per le ragioni illustrate anche questo indice non può essere applicato quando sono presenti frequenze attese inferiori a 1.

In questo caso il valore di X quadro è  $\chi^2 = \sum_{i=1}^r \frac{(O_i - A_i)^2}{A_i} = ((17-14.4)^2)/14.4 + ((14-19)^2)/19 + ((7-4.6)^2)/4.6 + ((5-7.6)^2)/7.6 + ((15-10)^2)/10 + ((0-2.4)^2)/2.4 = 8.85$ . La probabilità che X quadro sia diverso da zero per effetto del caso è di 0.01. Il valore è calcolato sulla distribuzione di probabilità Chi quadro con 2 grado/i di libertà, in corrispondenza dell'ascissa 8.85 (area a destra di tale punto).

Quando questo valore di probabilità (detto significatività della relazione) è inferiore a 0,05 si può iniziare a supporre lecitamente che vi sia una relazione significativa tra le due variabili.

### Vi è quindi relazione tra le due variabili (a livello di fiducia 0,05)

Il residuo standardizzato di cella indica se la differenza tra la frequenza osservata e la frequenza attesa in quella data cella è piccola o grande. Quando il numero totale dei casi è maggiore di 30, il residuo standardizzato può essere letto come un punteggio z. Se è superiore a 2 (in modulo) indica una differenza significativa (a livello di fiducia 0,05 corrispondente a  $z=1,96$ ) tra la frequenza osservata e attesa nella cella relativa. Se è negativo, nella cella abbiamo meno soggetti di quanti potremmo aspettarci se non vi fosse relazione tra le due variabili; se è positivo, abbiamo più soggetti di quanti potremmo aspettarci se non vi fosse relazione tra le due variabili.

Il V di Cramer indica la forza della relazione tra le due variabili. Viene ottenuto rapportando X quadro al massimo che può assumere ed estraendo la radice quadrata, ossia

$$V = \sqrt{\frac{\chi^2}{n \cdot [\min(\#r, \#c) - 1]}} = 0.39, \text{ dove } N \text{ è il numero dei casi, } \#r \text{ il numero di righe, } \#c \text{ il numero di colonne. Varia tra 0 (minima forza della relazione) e 1 (massima forza della relazione).}$$

### 13. Interpretazione dei risultati:

Il questionario da noi realizzato è stato somministrato a 100 persone di Torino e provincia.

Svolgendo l'*analisi monovariata* abbiamo ottenuto i seguenti risultati:

- ❖ il 58% dei partecipanti è di sesso femminile, il restante 42% è di sesso maschile.
- ❖ la maggioranza dei partecipanti ha 20 anni (14%).
- ❖ i fumatori sono rappresentati da una maggioranza del 41%, rispetto al 23% dei non fumatori.
- ❖ tra i fumatori la maggioranza, con il 63%, fuma meno di 10 sigarette al giorno; il 33% da 10 a 20 e il restante 3% più di un pacchetto al giorno.
- ❖ per il 67% tra coloro che fumano è indifferente fumare in compagnia o da soli; il 25% preferisce fumare in compagnia, mentre solamente l'8% preferisce fumare da solo.
- ❖ tra coloro che fumano, la maggioranza (69%) ha iniziato a fumare tra i 14 e i 18 anni, mentre il 16% ha iniziato prima dei 14 anni e altrettanti dopo i 18 anni.
- ❖ la motivazione principale dell'inizio del vizio è stata quella di provare esperienze nuove, con una percentuale del 29%; il 21% del campione ha affermato, invece, di aver iniziato a fumare perchè aiutava a rilassarsi;

il 19% per puro piacere e il restante 13% per essere maggiormente inserito nel gruppo di amici.

- ❖ il 57% del campione ha amici per la maggior parte fumatori, il 42% ha solo qualche amico fumatore e soltanto l'1% ha affermato di non aver nessun amico fumatore.
- ❖ 82% del campione ha affermato di credere che il gruppo di amici possa influenzare il singolo a fare uso di sigarette, mentre il 18% ha affermato il contrario.

Sulla base dell'*analisi bivariata* possiamo affermare che:

- ❖ è stata riscontrata una relazione tra l'aver amici fumatori (v13) e la propria scelta di fumare(v5): il valore di X quadro (11.2) è significativo.
- ❖ c'è relazione tra il numero di sigarette fumate dal campione (v6) e quelle fumate dagli amici(v14) in quanto il valore di X quadro (8.85) è significativo.

#### **14. Autoriflessione sull'esperienza compiuta**

Il lavoro che abbiamo svolto è stato utile e stimolante per comprendere la complessità e l'importanza di fare ricerca in ambito educativo.

Da questa esperienza abbiamo imparato a realizzare una ricerca empirica utilizzando programmi informatici come JsStat per l'analisi statistica dei dati, Q-gen per la creazione del questionario e Excel per caricare i dati ottenuti sulla sua matrice.

Ci riteniamo soddisfatte del risultato ottenuto in quanto, la popolazione che abbiamo preso in considerazione, risulta avere una buona percentuale di fumatori e, l'ipotesi da noi formulata, è stata confermata dai dati e dalle analisi di ricerca, constatando che avere amici fumatori può incidere sulla scelta di iniziare a fumare.

## **Divisione parti:**

### **CECILIA APPENDINO:**

- Premessa
- Tema di ricerca
- Problema di ricerca
- Ipotesi di ricerca- individuazione dei fattori e definizione operativa
- Tecniche di analisi dei dati utilizzati
- Interpretazione dei risultati

### **ERIKA PERCASSI:**

- Obiettivo di ricerca
- Individuazione del campione e tipologia di campionamento
- Questionario
- Tecniche di analisi dei dati utilizzati
- Interpretazione dei risultati

### **RAMONA DECATALDO:**

- Quadro teorico
- Mappa concettuale
- Bibliografia e sitografia
- Tecniche e strumenti di rilevazione dei dati
- Tecniche di analisi dei dati utilizzati
- Interpretazione dei risultati