

IMPORTANZA dell'INFORMAZIONE e MODALITA' di acquisizione



Indice

INTRODUZIONE

- ✓ Oggetto dell'indagine
- ✓ Campo di indagine
- ✓ Metodologie di raccolta dati

ANALISI UNI-BIVARIATA

- ✓ Dati socio-demografici (età, sesso, professione, tempo libero)
- ✓ Item principali

ANALISI MULTIVARIATA

- ✓ Analisi fattoriale
- ✓ Analisi di dipendenza

CONCLUSIONI

Oggetto

Informarsi è necessario per tutti noi: trattasi di un'attività che quotidianamente e quasi inconsciamente compiamo; è una priorità, un bisogno (perché viviamo in un mondo Complesso dove gli eventi si succedono con rapidità), ma anche un passatempo, un piacere, un "hobbie".

Quindi svariate sono le motivazioni che ci spingono ad acquisire informazione e con l'avvento della tecnologia e la diffusione di internet captare informazioni è sempre più facile ed immediato.

Per tutte queste ragioni riteniamo che l'informazione sia una tematica interessante da trattare, perché ci riguarda da vicino, accompagnando la nostra quotidianità.

Dunque compiremo un'analisi approfondita di questo fenomeno servendoci dei dati raccolti mediante l'elaborazione e la somministrazione di un questionario a un campione di riferimento. Questo ci permetterà di comprendere quali siano i principali argomenti informativi d'interesse e soprattutto quali siano i canali e le fonti a cui si attinge per informarsi e infine quanto si è soddisfatti di tali sorgenti informative.

Campo d'indagine

✓ Universo di riferimento:

il Campione prelevato risiede in provincia di Milano e Varese.

✓ Periodo di rilevazione

i questionari sono stati somministrati al campione nei mesi di ottobre, novembre e dicembre.

✓ Dimensioni del campione

185 intervistati.

✓ Metodo di Campionamento

abbiamo utilizzato un metodo di campionamento probabilistico casuale semplice. I dati a nostra disposizione sono stati raccolti attraverso interviste fatte a familiari, conoscenti, ragazzi universitari della LIUC, e persone intervistate presso supermercati e centri commerciali.

Abbiamo scelto come campione di riferimento persone appartenenti a varie classi di età per poter analizzare le modalità di acquisizione dell'informazione in generale e successivamente le differenze tra le varie fasce d'età.

Questionario

Il questionario è stato diviso in 7 parti:

- Informazioni SOCIO-DEMOGRAFICHE
- Valutazione sull' IMPORTANZA dell' INFORMAZIONE e sugli ARGOMENTI di MAGGIORE INTERESSE
- MOMENTI, LUOGHI, e MOTIVI per informarsi
- Informazione sui CANALI PREFERITI
- Informazione sui MOTIVI di SCELTA di ogni CANALE e relativo GRADO di AFFIDABILITA'
- FREQUENZA di UTILIZZO e COSTI dei vari CANALI
- Valutazione sulla VOLONTA' di SOSTITUIRE L'INFORMAZIONE da FORMATO CARTACEO a VIRTUALE.

Elaborazione dati e analisi

Una volta raccolti i dati e sintetizzati in un data-entry, si è provveduto a suddividere gli intervistati in 3 fasce di età per poterne rilevare le differenze comportamentali. (ETA₁' = fino ai 23 anni; ETA₂' = dai 24 ai 45; ETA₃' = dai 46 agli 86)

Successivamente i dati sono stati trattati tramite opportune tecniche di analisi attraverso l'utilizzo del software SAS; in particolare:

- A) Analisi uni/bivariata
- B) Analisi fattoriale
- C) Analisi di dipendenza

Pertanto nelle slides di seguito andremo, in primis, ad affrontare in modo dettagliato l'analisi univariata e l'analisi bivariata.

La prima si occupa dello studio di una sola variabile aleatoria; mentre la seconda ricerca eventuali correlazioni o qualche relazione tra due diverse variabili aleatorie.

dati anagrafici: età (1)

Inferiori		Superiori	
Valore	Oss	Valore	Oss
13	76	70	4
15	110	70	136
16	126	71	176
16	91	75	101
17	163	86	108

Posizione	
Media	34.341
Mediana	28
Moda	22
Skewness	0.883
Kurtosis	-0.287

Dall'analisi dei dati anagrafici notiamo che il campione su cui è stata compiuta l'indagine è costituito da persone con età minima di 13 anni e massima di 86.

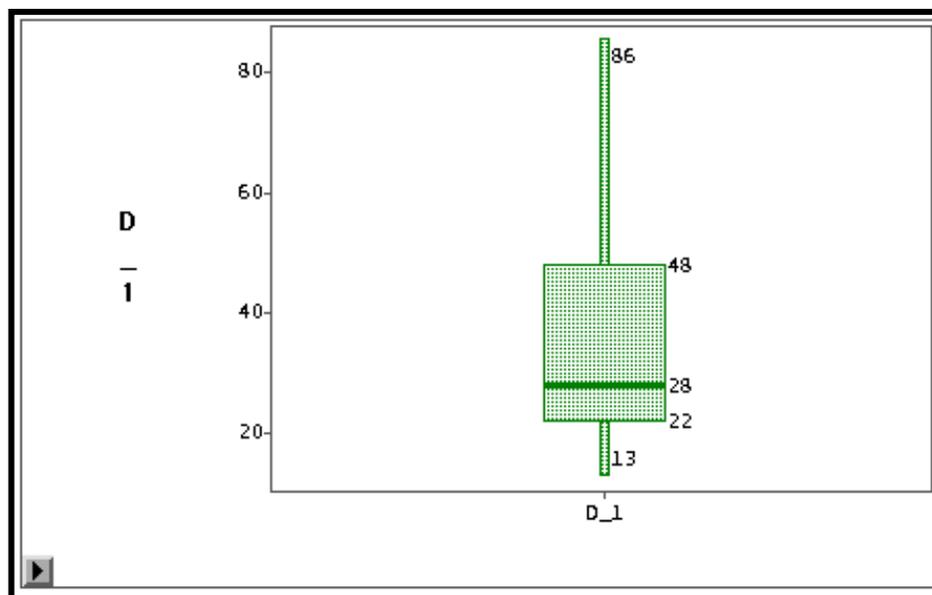
L'età della popolazione di riferimento si assesta a 34 anni circa, con un valore centrale (mediana) di 28 e un valore con più alta frequenza (moda) di 22. Da ciò deduciamo di aver intervistato una popolazione che può approssimativamente definirsi "giovane".

Per quanto riguarda la forma della distribuzione possiamo affermare che sia obliqua a destra (asimmetria positiva), in quanto la media è maggiore della mediana e ciò è poi confermato dall'indicatore di Skewness, che è maggiore di zero. La nostra distribuzione ha quindi una coda che si estende verso valori positivi.

L'indicatore di Kurtosis è minore di 3: questo sta invece a significare che la distribuzione è iponormale, cioè per valori molto distanti dalla media ha densità di frequenza minore rispetto alla normale standard.

dati anagrafici: età (2)

QUANTILI	
Quantile	Stima
100% Max	86
99%	75
95%	61
90%	59
75% Q3	48
50% Mediana	28
25% Q1	22
10%	19
5%	17
1%	15
0% Min	13



La tabella dei quantili e il box plot sintetizzano in modo alternativo quanto già dedotto nella slide precedente. Troviamo così conferma circa i valori estremi (13 e 86), la mediana (28) e la moda(22).

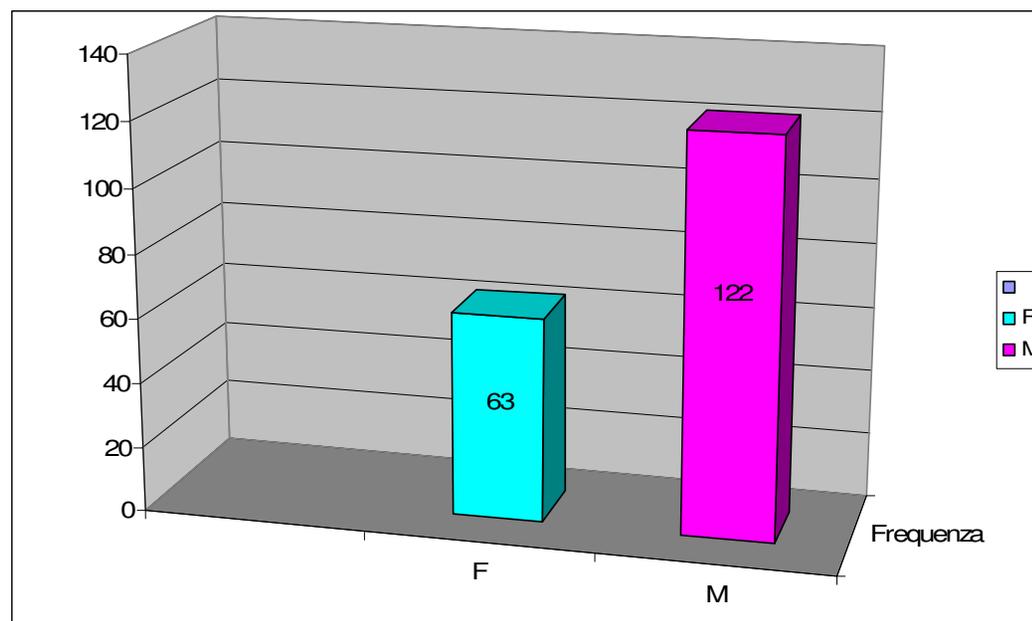
dati anagrafici: età (3)

ETA'					ETA'				
ETA'	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata	ETA'	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
13	1	0.54	1	0.54	44	2	1.08	131	70.81
15	1	0.54	2	1.08	45	4	2.16	135	72.97
16	2	1.08	4	2.16	46	1	0.54	136	73.51
17	8	4.32	12	6.49	47	2	1.08	138	74.59
18	2	1.08	14	7.57	48	2	1.08	140	75.68
19	9	4.86	23	12.43	49	2	1.08	142	76.76
20	8	4.32	31	16.76	50	1	0.54	143	77.3
21	4	2.16	35	18.92	51	5	2.7	148	80
22	17	9.19	52	28.11	52	8	4.32	156	84.32
23	10	5.41	62	33.51	53	2	1.08	158	85.41
24	13	7.03	75	40.54	55	3	1.62	161	87.03
25	10	5.41	85	45.95	56	2	1.08	163	88.11
26	5	2.7	90	48.65	57	1	0.54	164	88.65
28	8	4.32	98	52.97	59	3	1.62	167	90.27
29	4	2.16	102	55.14	60	3	1.62	170	91.89
30	4	2.16	106	57.3	61	6	3.24	176	95.14
31	5	2.7	111	60	66	1	0.54	177	95.68
33	5	2.7	116	62.7	67	2	1.08	179	96.76
35	3	1.62	119	64.32	68	1	0.54	180	97.3
38	2	1.08	121	65.41	70	2	1.08	182	98.38
39	2	1.08	123	66.49	71	1	0.54	183	98.92
40	2	1.08	125	67.57	75	1	0.54	184	99.46
41	3	1.62	128	69.19	86	1	0.54	185	100
42	1	0.54	129	69.73					

Dalla tabella notiamo che circa il 50% dei casi ha un'età compresa tra i 13 e i 26 anni. Da 28 anni in poi si evidenzia una distribuzione piuttosto omogenea delle diverse età.

dati anagrafici: sesso

SESSO				
SESSO	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
F	63	34.05	63	34.05
M	122	65.95	185	100



Analizzando le caratteristiche del target di riferimento notiamo poi che il nostro campione presenta una netta prevalenza di sesso maschile su quello femminile.

Su 185 osservazioni, infatti, abbiamo 122 uomini che rappresentano il 65.95% del campione, mentre 63 donne che rappresentano solo il restante 34.05%

dati anagrafici: professione

PROFESSIONE				
PROFESSIONE	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
STUDENTE	59	31.89	59	31.89
IMPIEGATO	42	22.7	101	54.59
OPERAIO	20	10.81	121	65.4
PENSIONATO	12	6.49	133	71.89
LAVORATORE AUTOMO	10	5.41	143	77.3
IMPRENDITORE	9	4.86	152	82.16
LIBERO PROFESSIONISTA	9	4.86	161	87.02
INSEGNANTE	8	4.32	169	91.34
ARTIGIANO	7	3.78	176	95.12
CASALINGA	5	2.7	181	97.82
DISOCCUPATO	4	2.16	185	100

La professione prevalente è quella dello studente con il 31.89%. Ciò è in linea con quanto già dedotto dai dati sull'età, giustifica cioè la giovinezza del nostro campione.

tempo libero

Valore%	PER NIENTE(1-2)			POCO(3-5)			ABBASTANZA(6-8)			MOLTO(9-10)		
HOBBIES	ETA' 1	ETA' 2	ETA' 3	ETA' 1	ETA' 2	ETA' 3	ETA' 1	ETA' 2	ETA' 3	ETA' 1	ETA' 2	ETA' 3
SPORT	4.84	17.81	40	32.26	28.77	12	38.71	28.76	32	24.19	24.66	16
MUSICA	1.61	13.7	12	35.49	42.46	54	48.74	36.99	30	14.16	6.85	4
CINEMA	14.52	9.59	30	30.64	43.83	32	51.61	42.47	32	3.23	4.11	6
VIAGGI	16.13	12.33	18	32.26	32.88	34	33.87	45.2	38	17.74	9.59	10
TEATRO	59.68	30.14	36	29.03	46.57	28	9.68	19.18	18	1.61	4.11	18
MOTORI	38.71	31.51	40	25.81	34.24	24	22.58	23.29	24	12.9	10.96	12
MODELLISMO/NUMISMATICA	61.29	72.6	82	16.13	12.33	6	19.35	12.33	10	3.23	2.74	2
TECNOLOGIA	12.9	24.66	58	19.36	32.87	22	59.68	32.88	16	8.06	9.59	4

Suddividendo il campione in tre fasce d'età abbiamo potuto analizzarne le differenze/somiglianze, attraverso un accurato confronto. La tabella mostra quanto tempo ogni classe dedica alla cura dei diversi hobbies. Evidenziando così i valori più alti associati a queste due variabili emergono alcuni aspetti su cui riflettere.

La differenza più importante è quella relativa alla tematica della tecnologia che inevitabilmente dipende in modo inverso dal fattore età. I più anziani hanno generalmente una "condotta" di vita più tradizionale e legata al passato e si avvicinano alla tecnologia con maggiore difficoltà e spesso perplessità.

Con riferimento agli altri interessi notiamo un comportamento sostanzialmente omogeneo e in particolare rileviamo una scarsa dedizione del campione ad attività quali teatro, motori e modellismo.

importanza informazione

Alla domanda Circa l'importanza di mantenere aggiornato il proprio livello informativo, 168 persone su 185, rappresentanti il 90.81% della popolazione ha dato un giudizio sufficiente (≥ 6).

Dunque la quasi totalità del campione ritiene sufficientemente importante informarsi.

Se poi diamo un occhio agli indicatori statistici di base riscontriamo una media molto alta (8 su 10). Ciò ci permette di concludere con sicurezza che il nostro campione ritenga mediamente molto importante aggiornare il proprio livello informativo.

Importanza dell'informazione				
importanza infomaz	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
1	1	0.54	1	0.54
3	1	0.54	2	1.08
4	3	1.62	5	2.7
5	12	6.49	17	9.19
6	21	11.35	38	20.54
7	28	15.14	66	35.68
8	51	27.57	117	63.24
9	27	14.59	144	77.84
10	41	22.16	185	100

Misure statistiche di base			
Posizione		Variabilità	
Media	7.886486	Deviazione std	1.69817
Mediana	8	Varianza	2.88378
Moda	8	Intervallo	9
		Intervallo interquartile	2

importanza informazione rispetto al sesso

impo info	sesso		Totale
	F	M	
1	0	1	1
	0	0.54	0.54
	0	100	
	0	0.82	
3	1	0	1
	0.54	0	0.54
	100	0	
	1.59	0	
4	2	1	3
	1.08	0.54	1.62
	66.67	33.33	
	3.17	0.82	
5	2	10	12
	1.08	5.41	6.49
	16.67	83.33	
	3.17	8.2	
6	9	12	21
	4.86	6.49	11.35
	42.86	57.14	
	14.29	9.84	
7	14	14	28
	7.57	7.57	15.14
	50	50	
	22.22	11.48	
8	19	32	51
	10.27	17.3	27.57
	37.25	62.75	
	30.16	26.23	
9	6	21	27
	3.24	11.35	14.59
	22.22	77.78	
	9.52	17.21	
10	10	31	41
	5.41	16.76	22.16
	24.39	75.61	
	15.87	25.41	
Totale	63	122	185
	34.05	65.95	100

Anche classificando la popolazione in base al sesso di appartenenza si nota come entrambi esprimano una preferenza più che positiva nei confronti dell'importanza dell'aggiornamento informativo.

importanza informazione rispetto all'istruzione

importanza dell'informazione									
tipo di istruzione	1	3	4	5	6	7	8	9	10
Nessuno	0	0	0	0	1	0	0	0	0
Elementare	0	0	1	1	5	1	0	1	2
Media	0	1	0	1	7	10	5	3	8
Superiore	0	0	2	10	6	11	34	16	17
Laurea	0	0	0	0	2	6	12	7	14
Frequenza dei valori mancanti = 1									

Ci siamo poi occupati di incrociare il grado di istruzione con la valutazione espressa circa l'importanza di aggiornare il proprio bagaglio informativo.

Anche in questi caso possiamo affermare che l'importanza di informarsi prescinda dal grado di formazione, infatti qualsiasi sia il titolo di studio, sono comunque rari i casi in cui si esprime una valutazione inferiore a 6/10. Ciò dipende da quanto concluso in precedenza, e cioè dal fatto che i dati relativi all'importanza dell'aggiornamento informativo abbiano una media molto alta e che quindi il campione ritenga mediamente piuttosto importante informarsi.

motivazione

Alla domanda relativa alla motivazione che spinge ad acquisire nuove informazioni, il 40% dei casi ha risposto di informarsi esclusivamente per conoscenza personale, mentre il restante 60% si informa oltre che per conoscenza personale anche per studio e ragioni lavorative.

MOTIVAZIONE	conoscenza personale	motivazioni varie
Percentuale	40.00%	60.00%

argomenti di interesse

Analizzando gli argomenti di maggiore interesse notiamo che l'attualità è quello cui è associata la media più elevata: mediamente dunque l'attualità è l'argomento che riscuote maggiore successo.

Notiamo poi che ad argomenti quali economia e finanza, attualità, notizie locali e meteo sono associate rispettivamente mode di 8, 8, 7 e 10. Ciò significa che rispetto a questi argomenti la tendenza è quella di nutrire un interesse positivo o più che positivo.

ARGOMENTI	MEDIA	MEDIANA	MODA
cronaca nera	5.71	6	6
Cronaca rosa/gossip	4.32	4	1
economia e finanza	5.84	6	8
sport	6.17	6	6
attualità	6.95	7	8
politica	6.35	7	7
notizie locali	6.69	7	7
eventi culturali	6.25	6	5
meteo	6.64	7	10

argomenti di interesse rispetto al sesso

ARGOMENTI	F	M
cronaca nera	6.4	5.3
cronaca rosa/gossip	6.19	3.36
economia e finanza	5.44	6
sport	4.88	6.83
attualità	7.17	6.83
politica	5.63	6.7
notizie locali	6.2	6.9
eventi culturali	6.65	6.04
meteo	7.22	6.34

In generale notiamo che le donne siano maggiormente interessate degli uomini, rispetto agli argomenti a cui hanno risposto.

L'unica eccezione evidente è quella dello sport rispetto al quale l'interesse nutrito è più forte nella componente maschile.

argomenti di interesse rispetto all' età

Valore%	PER NIENTE(1-2)			POCO(3-5)			ABBASTANZA(6-8)			MOLTO(9-10)		
	ETA' 1	ETA' 2	ETA' 3	ETA' 1	ETA' 2	ETA' 3	ETA' 1	ETA' 2	ETA' 3	ETA' 1	ETA' 2	ETA' 3
CRONACA NERA	11.29	6.85	14	33.87	26.03	32	43.55	56.16	36	11.29	10.96	18
CRONACA ROSA/ GOSSIP	40.32	26.03	46	22.58	39.72	26	22.58	24.66	12	14.52	9.59	16
ECONOMIA E FINANZA	16.13	6.85	10	30.64	36.99	30	38.71	47.94	40	14.52	8.22	20
SPORT	8.06	6.85	28	24.2	20.55	32	38.71	49.31	18	29.03	23.29	22
ATTUALITA'	1.61	2.74	6	24.2	10.96	22	41.93	69.86	52	32.26	16.44	20
POLITICA	6.45	5.48	16	32.26	24.66	22	37.1	49.31	40	24.19	20.55	22
NOTIZIE LOCALI	8.06	4.11	6	30.65	15.07	8	50	63.01	44	11.29	17.81	42
EVENTI CULTURALI	4.84	4.11	2	41.93	39.73	22	46.78	41.09	42	6.45	15.07	34
METEO	12.9	13.7	6	19.36	10.96	24	45.16	52.05	20	22.58	23.29	50

La tabella mostra l' interesse espresso da ciascuna classe di età nei confronti dei vari argomenti. Questa sintesi ci permette di cogliere le eventuali uguaglianze/ differenze tra persone con età differente. Possiamo affermare che in generale tutte e 3 le categorie ritengono abbastanza interessanti i seguenti argomenti: cronaca nera, economia e finanza, attualità, politica, notizie locali ed eventi culturali. Al contrario il gossip è un argomento che suscita un interesse nullo o scarso.

Per quanto invece concerne le differenze, rileviamo due casi in cui gli appartenenti alla fascia di età 3 si discostano in modo evidente dalle altre due fasce. Il primo è l'argomento dello sport che suscita un più basso interesse da parte di questi intervistati rispetto alla fascia 1 e 2; il secondo è il meteo a cui la terza fascia appare decisamente più interessata rispetto a quanto non lo siano le altre due.

momenti della giornata di acquisizione dell'informazione

momento	intervallo	frequenza	percentuale
mattino	0%	62	33.51%
	da 1 a 25%	51	27.57%
	da 26 a 50%	57	30.81%
	da 51 a 75	8	4.33%
	da 76 a 100%	7	3.78%
pranzo	0%	107	57.84%
	da 1 a 25%	49	26.48%
	da 26 a 50%	25	13.52%
	da 51 a 75	3	1.62%
	da 76 a 100%	1	0.54%
primo pomeriggio	0%	131	70.81%
	da 1 a 25%	43	23.25%
	da 26 a 50%	11	5.94%
	da 51 a 75	0	0.00%
	da 76 a 100%	0	0.00%
tardo pomeriggio	0%	108	58.38%
	da 1 a 25%	56	29.57%
	da 26 a 50%	18	10.43%
	da 51 a 75	2	1.08%
	da 76 a 100%	1	0.54%
cena	0%	25	13.51%
	da 1 a 25%	33	17.84%
	da 26 a 50%	79	42.70%
	da 51 a 75	22	11.89%
	da 76 a 100%	26	14.06%
notte	0%	130	70.27%
	da 1 a 25%	39	21.08%
	da 26 a 50%	8	4.33%
	da 51 a 75	1	0.54%
	da 76 a 100%	7	3.78%

dell'informazione

Dalla tabella si evince che i momenti della giornata ove maggiormente si acquisisce informazione sono: al primo posto la cena (infatti solo il 13.51% non si informa) e al secondo il mattino (con il 33.51%). In particolare solo per quanto riguarda le cena si raggiunge un picco di informazione di 79 su 185. Al contrario riscontriamo che il primo pomeriggio e la notte siano momenti di scarsa informazione con frequenza pressoché omogenea.

frequenza di acquisizione dell'informazione

Posizione	
Media	3.46
Mediana	3
Moda	2

Notiamo che la quasi totalità del campione (83.24%) si informa almeno 4 volte al giorno; mentre 35.68% è la frequenza più elevata che corrisponde ad un'acquisizione di informazione pari a due volte al giorno. Quanto detto è confermato altresì dagli indicatori di posizione. Analizzando gli estremi (0 e 30) è interessante notare come solo 1 caso su 185 non si informi mai.

frequenza di acquisizione informazione				
n di volte/g	Frequenza	Percentuale	Frequenza cumulata	Percentuale cumulata
0	1	0.54	1	0.54
1	19	10.27	20	10.81
2	66	35.68	86	46.49
3	44	23.78	130	70.27
4	24	12.97	154	83.24
5	9	4.86	163	88.11
6	6	3.24	169	91.35
7	6	3.24	175	94.59
8	1	0.54	176	95.14
10	6	3.24	182	98.38
20	2	1.08	184	99.46
30	1	0.54	185	100

tempo utilizzato

Quantile	Stima
100% Max	120
99%	120
95%	120
90%	60
75% Q3	40
50% Mediana	30
25% Q1	15
10%	10
5%	10
1%	3
0% Min	3

Posizione	
Media	35.31351
Mediana	30
Moda	30

Il campione intervistato spende mediamente circa 35 minuti ogniqualvolta si informa.

Moda e mediana si posizionano sui 30 minuti.

Dalla tabella dei quantili evidenziamo i casi estremi, e cioè che il tempo d'informazione è compreso tra un minimo di 3 min a un massimo di 120.

Tempo medio speso per professione

Incrociando il tipo di professione svolto dal Campione di riferimento e il tempo impiegato ogniqualvolta ci si informa, notiamo tre categorie che mediamente spendono più tempo in questa attività: impiegato, lavoratore autonomo e disoccupato. Le ragioni sono legate all'attività lavorativa svolta dai primi due e dal tempo libero dell'ultimo.

PROFESSIONE	MEDIA (min)
artigiano	24.28
casalinga	21
disoccupato	42.5
impiegato	43.67
imprenditore	35
insegnante	23.62
lavoratore autonomo	42
libero professionista	30
operaio	35
pensionato	28.75
studente	34.15

canale preferito

Fatto cento il totale, il campione mediamente distribuisce la preferenza di canale come indicato nella seconda colonna; in particolare da questa spicca la televisione come il canale mediamente preferito, che in media si utilizza più di una volta al giorno, e segue internet il cui utilizzo medio settimanale è di circa 7 volte.

CANALE	MEDIA	UTILIZZO MEDIO SETTIMANALE
internet	25.7	7.18
quotidiani	17.51	4.32
periodici	4.65	1.11
televisione	42.64	8.96
radio	8.01	3.79
convegni	1.45	0.17

I convegni invece rappresentano un canale scarsamente apprezzato dal nostro campione, tanto che questi non vengono frequentati mediamente neanche una volta a settimana. Poco utilizzati dal campione sono anche i periodici, che vengono mediamente letti solamente circa una volta a settimana.

utilizzo medio settimanale per canale

CANALE	MEDIA	MEDIANA	MODA
internet	7.18	7	7
quotidiani	4.32	3	0
periodici	1.11	0	0
televisione	8.96	7	7
radio	3.79	2	0
convegni	0.17	0	0

La tabella sintetizza quanto già osservato nel canale precedente.

Televisione e internet sono i canali mediamente più utilizzati a settimana: nei casi più ricorrenti l'utilizzo medio è di una volta al giorno. Mentre convegni e periodici sono i canali meno sfruttati, tanto è vero che nei casi più ricorrenti si nota che il campione affermi di non utilizzarli neanche una volta a settimana.

Soddisfazione canale

CANALE	MEDIA	MODA
internet	6.87	8
quotidiani	6.57	7
periodici	4.95	1
televisione	6.57	7
radio	6.52	7
convegni	3.45	1

Il campione si ritiene mediamente abbastanza soddisfatto dall'utilizzo di internet: questo canale è quello che registra il maggior grado di soddisfazione (si esprime infatti una preferenza più che positiva nella maggior parte dei casi); a questo canale seguono quotidiani e televisione nei confronti dei quali il campione si esprime in modo omogeneo. I convegni invece sono il canale verso cui il campione si esprime piuttosto negativamente, tanto è vero che nella maggior parte dei casi la votazione espressa è di 1 su 10.

Soddisfazione canale in base all'età

Valore%	PER NIENTE(1-2)			POCO(3-5)			ABBASTANZA(6-8)			MOLTO(9-10)		
CANALE	ETA' 1	ETA' 2	ETA' 3	ETA' 1	ETA' 2	ETA' 3	ETA' 1	ETA' 2	ETA' 3	ETA' 1	ETA' 2	ETA' 3
INTERNET	8.33	10.61	27.5	8.34	15.15	17.5	55	37.88	32.5	28.33	36.36	22.5
QUOTIDIANI	6.67	8.96	8.7	11.66	17.91	17.39	65	56.71	50	16.67	16.42	23.91
PERIODICI	29.31	18.75	27.5	34.48	31.25	17.5	31.04	42.19	40	5.17	7.81	15
TELEVISIONE	9.68	4.23	10.2	20.97	19.71	20.41	56.45	60.57	40.82	12.9	15.49	28.57
RADIO	21.67	12.12	20	26.66	27.27	25	46.67	45.46	47.5	5	15.15	7.5
CONVEGNI	58.93	29.03	45.95	19.64	38.71	29.73	21.43	27.42	18.91	0	4.84	5.41

L'analisi ci permette di mettere a confronto la soddisfazione del campione per fasce d'età. Non notiamo sostanziali differenze: infatti per quasi tutti i canali, con eccezione dei convegni, le tre fasce si mostrano mediamente soddisfatte. Il caso dei convegni è anomalo in quanto rispetto a questo canale tutte e tre le fasce presentano una soddisfazione piuttosto bassa. Unico piccolo scostamento è rappresentato dalla soddisfazione dei giovani nei confronti dei periodici.

fattori determinanti nella scelta del canale

CRITERIO	MEDIA	MEDIANA	MODA
SEMPLICITA'	5.99	6	6
COSTO	4.83	5	1
VELOCITA' DI ACQUISIZIONE	6.88	7	6
COMODITA'	7.32	7	7
TEMPO DI AGGIORNAMENTO	6.97	7	8

Dall'analisi emerge che la comodità è il fattore che principalmente influenza la scelta del canale informativo; seguono poi velocità di acquisizione e tempo di aggiornamento, fattori altrettanto determinanti.

utilizzo medio canale rispetto

PROFESSIONE	INTERNET	QUOTIDIANI	PERIODICI	TELEVISIONE	RADIO	CONVEGNI
ARTIGIANO	8	2.42	0.42	7	6.57	0
CASALINGA	0	4.2	3.8	7	1.2	0
DISOCCUPATO	3.5	3.5	0	3.5	0	0.5
IMPIEGATO	6.38	3.3	1.41	6.76	4.42	0.22
IMPRENDITORE	5.44	3.77	1	5.66	3.88	0.44
INSEGNANTE	2.625	2.875	1.125	6.125	2.25	0.125
LAVORATORE AUTONOMO	4.4	4.3	1.8	5.7	4	0.1
LIBERO PROFESSIONISTA	8.5	6.16	1.5	7	3	1
OPERAIO	5.15	2.1	0.25	9.2	4.1	0.15
PENSIONATO	1.75	3.83	0.75	7	0.75	0.166
STUDENTE	11.52	6.28	1.05	13.37	4.25	0.67

alla
profes-
sione

In questo caso andiamo a verificare se l'approccio informativo, e in particolare l'utilizzo medio settimanale di ciascun canale, sia influenzato da caratteristiche personali come la professione. Dall'analisi emerge che internet è maggiormente utilizzato dallo studente, il quale ne fa uso almeno due volte al giorno. Peraltro lo studente utilizza molto anche televisione, circa più di due volte al giorno, e legge i quotidiani quasi tutti i giorni. Mediamente i convegni sono frequentati una volta a settimana solo dal libero professionista. Per quanto riguarda i periodici, la categoria che maggiormente li utilizza risulta quella delle casalinghe, che ne fa uso mediamente ogni due giorni. Infine ascoltano mediamente tutti i giorni la radio solo gli artigiani.

fonti utilizzate per canale

La tabella illustra il numero di fonti mediamente utilizzato dal campione per ciascun tipo di canale. Focalizzandoci sui dati più rilevanti possiamo affermare che generalmente il campione intervistato si informa attraverso il canale televisivo, guardando mediamente 4 diversi canali; mentre naviga su internet utilizzando in media più di 3 siti differenti.

CANALE	MEDIA
internet	3.778
quotidiani	1.63
periodici	0.827
televisione	4.227
radio	1.178
convegni	1.2

fattori determinanti nella scelta della fonte

CRITERIO	MEDIA
orientamento politico	5.092
temi trattati	7.416
area geografica di interesse	5.113
direttore	4.062
formato/stile	5.448
con chi vivi	4.454
redazione	3.978
giornalisti/speaker	5.286
qualità servizi	7.102

Dall'analisi emerge chiaramente che il campione è mediamente piuttosto influenzato nella scelta della fonte informativa sia dai temi trattati, che dalla qualità dei servizi prestati. Sono invece mediamente poco influenti direttore e redazione.

affidabilità delle fonti per fascia d'età

Valore%	PER NIENTE(1-2)			POCO(3-5)			ABBASTANZA(6-8)			MOLTO(9-10)		
CANALE	ETA' 1	ETA' 2	ETA' 3	ETA' 1	ETA' 2	ETA' 3	ETA' 1	ETA' 2	ETA' 3	ETA' 1	ETA' 2	ETA' 3
INTERNET	6.56	10.45	32.5	6.55	16.42	10	65.58	64.17	32.5	21.31	8.96	25
QUOTIDIANI	10	4.48	14.89	8.33	14.92	29.79	65	68.66	38.3	16.67	11.94	17.02
PERIODICI	19.64	10.77	33.33	28.57	33.85	28.57	42.86	52.3	28.58	8.93	3.08	9.52
TELEVISIONE	8.06	7.14	8.16	22.59	20	24.49	59.67	65.72	44.9	9.68	7.14	22.45
RADIO	17.24	7.46	17.07	32.76	14.93	24.39	44.83	71.64	43.91	5.17	5.97	14.63

Ad esclusione dei convegni, canale che dalle analisi precedenti abbiamo dedotto essere trascurabile,

con questa tabella andiamo ad indagare quanto ogni fascia d'età, in cui è stato suddiviso il nostro campione, ritenga affidabili le informazioni acquisite dalle fonti utilizzate.

Tutte e tre le classi ritengono abbastanza affidabili i quotidiani, la televisione e la radio. Per internet e i periodici dobbiamo fare qualche considerazione in più. Infatti nel caso del primo canale, mentre le prime due fasce si esprimono nello stesso modo ritenendo abbastanza affidabile questo canale, la terza fascia risulta un po' più critica nei confronti dello stesso, dato che un buon 32.5% si esprime più che negativamente nei confronti di internet. Un'ulteriore differenza si riscontra per il canale periodici: infatti mentre le prime due fasce sono concordi nel considerare abbastanza affidabile questo canale, la terza fascia si discosta da queste.

la spesa per l'informazione

Tra i canali informativi quelli che comportano una spesa sono:

- a) quotidiani/ periodici--> costo d'acquisto;
- b) televisione/ internet--> abbonamenti e servizi aggiuntivi.

Nella tabella si sintetizza quanto mediamente il campione intervistato spenda al mese per usufruire dei vari canali informativi:

CANALE	SPESA MEDIA (€)
internet	17.875
quotidiani	14.951
periodici	6.783
televisione	13.881

L'analisi fattoriale

Abbiamo finora visto in modo dettagliato come vi siano svariate modalità di acquisizione dell'informazione: possiamo infatti scegliere tra i vari CANALI (internet, quotidiani, periodici, televisione, radio, convegni) e all'interno di ogni singolo canale tra una pluralità di FONTI a cui attingere (siti web, Canali tv, stazioni radio,...). Il nostro obiettivo è capire quali fattori influenzano la scelta sia del canale che della rispettiva fonte. Ci serviamo a tale scopo dell'**analisi fattoriale**. Si tratta di una tecnica statistica multivariata che permette di sintetizzare l'informazione condivisa tra più Variabili correlate tra loro: questo porta ad una minor perdita di informazioni rilevanti e allo stesso tempo ad una semplificazione notevole della ricerca.

a) METODO UTILIZZATO: METODO DELLE COMPONENTI PRINCIPALI

b) VARIABILI: le variabili considerate sono i parametri che influenzano la scelta del canale e quelli che influenzano la scelta relativa al tipo di fonte

DESCRIZIONE VARIABILI	NOME VAR
critério di scelta:	
SEMPLICITÀ	D_17_s
COSTO	D_17_c
VELOCITÀ DI ACQUISIZIONE	D_17_v
COMODITÀ	D_17_com
TEMPO DI AGGIORNAMENTO	D_17_tda
ORIENTAMENTO POLITICO	D_20_orp
TEMI TRATTATI	D_20_tt
AREA GEOGRAFICA DI INTERESSE	D_20_ag
DIRETTORE	D_20_d
FORMATO / STILE	D_20_fs
CON CHI VIVI	D_20_ccv
REDAZIONE	D_20_r
GIORNALISTI/SPEAKER	D_20_gs
QUALITÀ SERVIZI	D_20_qs

scelta del numero di fattori da considerare(1)



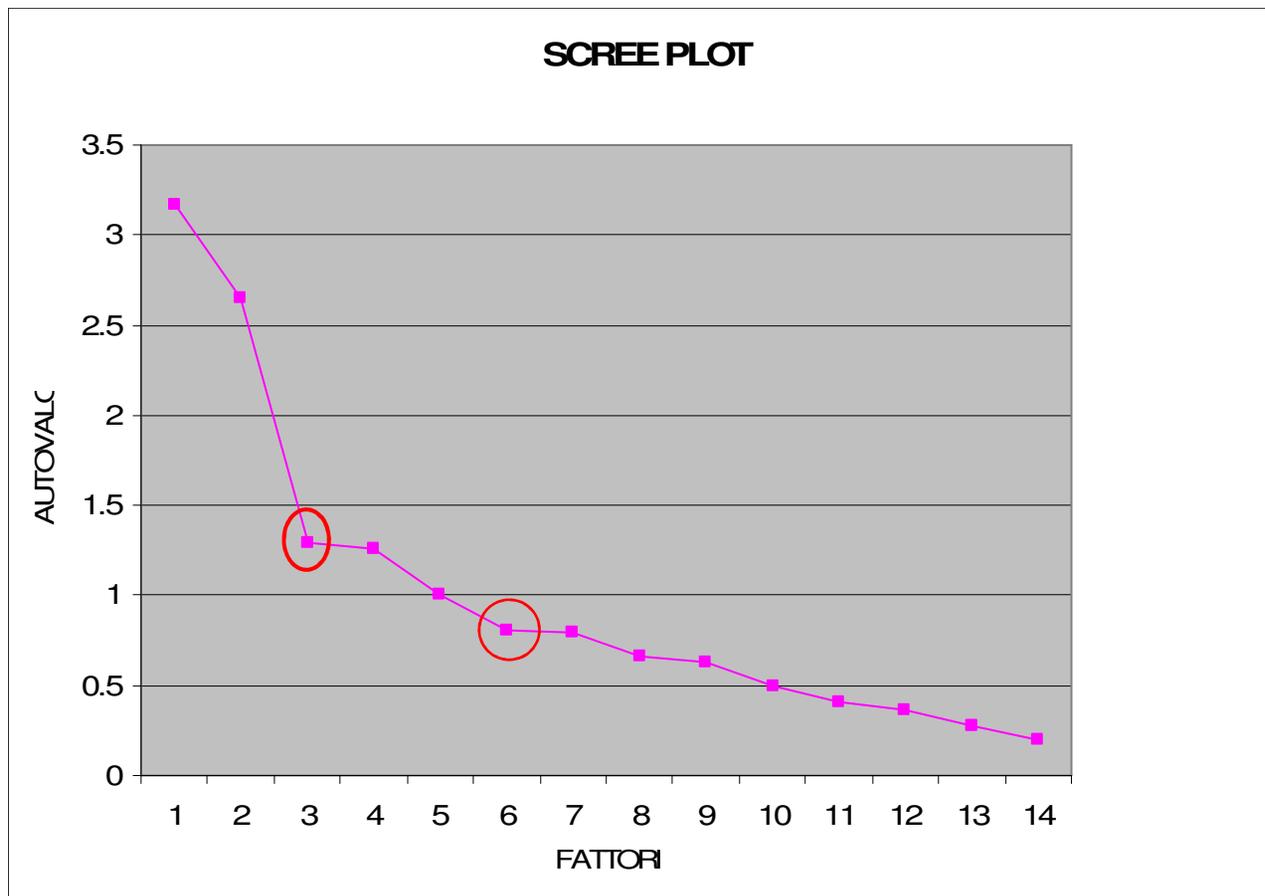
Autovalori della matrice di correlazione: Totale				
= 14 Media = 1				
	Autovalore	Differenza	Proporzione	Cumulata
1	3.16944223	0.52227941	0.2264	0.2264
2	2.64716281	1.35701039	0.1891	0.4155
3	1.29015243	0.02599489	0.0922	0.5076
4	1.26415754	0.2604549	0.0903	0.5979
5	1.00370264	0.20036187	0.0717	0.6696
6	0.80334077	0.01216326	0.0574	0.727
7	0.79117751	0.13231428	0.0565	0.7835
8	0.65886322	0.03460029	0.0471	0.8306
9	0.62426293	0.12396136	0.0446	0.8752
10	0.50030158	0.09138333	0.0357	0.9109
11	0.40891825	0.04258591	0.0292	0.9401
12	0.36633234	0.09276211	0.0262	0.9663
13	0.27357023	0.07495472	0.0195	0.9858
14	0.19861552	0	0.0142	1

PERCENTUALE DI VARIANZA SPIEGATA:

Grazie a questo metodo si può analizzare la Variabilità Cumulata spiegata dalle Variabili.

Come si può vedere dall'output dei fattori all'interno della tabella, si possono considerare da un minimo di 3 ad un massimo di 6. Con 3 si arriva a spiegare il 50.76% della Variabilità del fenomeno, mentre con 6 elementi si ha più Varianza spiegata (72,7%), ma considerando più Variabili si rischia di avere una maggiore complessità.

scelta del numero di fattori da considerare(2)



METODO

DELLO SCREE- PLOT:

permette di vedere graficamente la Variabilità assoluta spiegata da Ciascun fattore. In questo caso la curva presenta un “gomito” in corrispondenza del 3° e del 6° fattore, ove si riscontra una evidente variazione di pendenza.

scelta del numero di fattori da considerare(3)

Autovalori della matrice di correlazione: Totale				
= 14 Media = 1				
	Autovalore	Differenza	Proporzione	Cumulata
1	3.16944223	0.52227941	0.2264	0.2264
2	2.64716281	1.35701039	0.1891	0.4155
3	1.29015243	0.02599489	0.0922	0.5076
4	1.26415754	0.2604549	0.0903	0.5979
5	1.00370264	0.20036187	0.0717	0.6696
6	0.80334077	0.01216326	0.0574	0.727
7	0.79117751	0.13231428	0.0565	0.7835
8	0.65886322	0.03460029	0.0471	0.8306
9	0.62426293	0.12396136	0.0446	0.8752
10	0.50030158	0.09138333	0.0357	0.9109
11	0.40891825	0.04258591	0.0292	0.9401
12	0.36633234	0.09276211	0.0262	0.9663
13	0.27357023	0.07495472	0.0195	0.9858
14	0.19861552	0	0.0142	1

METODO DEGLI AUTOVALORI:

Il metodo degli autovalori indica 5 come numero ideale di fattori, in quanto dopo il quinto fattore la Varianza è minore di 1, valore ideale che si avrebbe nel caso in cui tutte le Variabili fossero correlate ugualmente.

confronto tra le situazioni scelte

In base a quanto detto si procederà dunque con l'analisi a 3, 5 e 6 fattori.

CONFRONTO CUMUNALITA' FINALI:

VARIABILI	DESCRIZIONE	CUMUNALITA'		
		n=3	n=5	n=6
D_17_s	semplicità	0.416	0.6951	0.6967
D_17_c	costo	0.5719	0.7324	0.7335
D_17_v	velocità di acquisizione	0.7034	0.8154	0.8367
D_17_com	comodità	0.647	0.6589	0.7054
D_17_tda	tempo di aggiornamento	0.567	0.7333	0.807
D_20_orp	orientamento politico	0.4078	0.7971	0.805
D_20_tt	temi trattati	0.3535	0.6494	0.6971
D_20_ag	area geografica di interesse	0.4339	0.6814	0.8443
D_20_d	direttore	0.5959	0.7226	0.7251
D_20_fs	formato/stile	0.3562	0.5612	0.6345
D_20_ccv	con chi vivi	0.3719	0.5363	0.5586
D_20_r	redazione	0.6518	0.6988	0.7273
D_20_gs	giornalisti/speaker	0.6092	0.6341	0.6652
D_20_qs	qualità servizi	0.4211	0.4586	0.7417

Dal confronto tra cumunalità si nota come la soluzione a 3 fattori sia poco adeguata perchè ha una capacità esplicativa rispetto ad alcuni fattori molto bassa.

La soluzione a 5 fattori fornisce una spiegazione meno adeguata di quattro variabili, aventi probabilmente un alto contenuto di specificità.

La soluzione a 6 fattori fornisce, da un lato, una capacità esplicativa maggiore sulle quattro variabili, mentre, dall'altro lato, apporta maggiore complessità.

In conclusione optiamo per la soluzione a 5 privilegiando una maggiore sintesi pur a discapito della capacità esplicativa. Tuttavia questo sacrificio non è così pesante : se si guarda infatti alla Varianza Complessiva spiegata dalle due soluzioni, si passa da un 72.20% a un 66.96%, che rappresenta una quota di Variabilità spiegata più che significativa.

schema fattoriale

		Factor1	Factor2	Factor3	Factor4	Factor5
D_17_s	semplicità	0.56626	.	.	.	0.46051
D_17_c	costo	0.35685	.	0.65469	.	0.3875
D_17_v	velocità	0.75292
D_17_com	comodità	0.68764	-0.36206	.	.	.
D_17_tda	tempo di aggiornamento	0.5326	-0.43612	.	.	-0.38524
D_20_orp	orientamento politico	.	0.54298	.	0.53024	.
D_20_tt	temi trattati	0.41299	.	.	0.53419	.
D_20_ag	area geografica di interesse	.	.	-0.5248	.	0.38026
D_20_d	direttore	.	0.74874	.	.	.
D_20_fs	formato/stile	0.38261	.	.	-0.43544	.
D_20_ccv	con chi vivi	.	0.50515	.	.	.
D_20_r	redazione	.	0.72899	.	.	.
D_20_gs	giornalisti/speaker	0.58604	0.49902	.	.	.
D_20_qs	qualità servizi	0.63683
I valori minori di 0.35 non sono stampati.						

Lo schema fattoriale a 5 fattori, così come si presenta, è di difficile interpretazione; per questo risulta opportuno ruotare i fattori attraverso un apposito metodo (VARIMAX).

interpretazione dei fattori

SCHEMA FATTORIALE RUOTATO		Factor1	Factor2	Factor3	Factor4	Factor5
D_17_v	velocità	0.8578
D_17_tda	tempo di aggiornamento	0.7885
D_17_com	comodità	0.70345	.	.	0.39398	.
D_20_qs	qualità servizio	0.53133
D_20_r	rapidità	.	0.74824	.	.	.
D_20_fs	formato/stile	.	0.71171	.	.	.
D_20_ccv	con chi vivi	.	0.70059	.	.	.
D_20_gs	giornalisti/speaker	.	0.62098	0.36737	.	.
D_20_orp	orientamento politico	.	.	0.8923	.	.
D_20_d	direttore	.	.	0.77647	.	.
D_17_c	costo	.	.	.	0.83334	.
D_17_s	semplicità	.	.	.	0.65037	0.45187
D_20_ag	area geografica di interesse	0.7622
D_20_tt	temi trattati	0.71198

I valori minori di 0.35 non sono stampati.

Rapidità di acquisizione e qualità del servizio offerto.

Esposizione dell'informazione

Affinità politica/ideologica

Accessibilità al servizio

Attrattività argomenti trattati

L'analisi di dipendenza: regressione lineare(1)

Vogliamo cercare di spiegare l'influenza di determinate Variabili sull'importanza di mantenere aggiornato il livello informativo della popolazione di riferimento.

La Variabile dipendente è dunque : D_9 = importanza di mantenere il proprio livello informativo aggiornato.

Il modello scelto per l'analisi di dipendenza in questione è il **modello di regressione lineare**. Questo è dovuto principalmente alla natura della Variabile dipendente, che nel caso osservato è un dato quantitativo riferito ad una scala di intervalli. Dunque il modello ha l'obiettivo di rappresentare la relazione tra l'importanza di aggiornamento informativo e i regressori selezionati tra quelli a nostra disposizione.

regressione lineare(2)

Due indicatori di bontà complessiva del modello:

◊ Il TEST F, un test che valuta la significatività congiunta del set di variabili: poiché il p-value è molto piccolo possiamo affermare che vi sia almeno un coefficiente tra le variabili indipendenti diverso da zero e quindi che almeno una variabile è significativamente legata con la variabile dipendente.

Analisi della varianza					
Origine	DF	Somma dei quadrati	Media dei quadrati	Valore F	Pr > F
Modello	71	322.55368	4.54301	5.05	<.0001
Errore	69	62.0988	0.89998		
Totale corretto	140	384.65248			

◊ L' R quadro valuta la capacità esplicativa del modello, cioè, più si avvicina ad 1 e più il modello è significativo. In questo caso il valore è molto alto, in quanto lontano dalla soglia di 0.2, dunque il modello è più che buono.

Radice MSE	0.9487	R-quadro	0.8386
Media dipendente	8.0497	R-quadro corr	0.6724
Coeff var	11.785		

Stime dei parametri		Stime dei parametri		Stime dei parametri	
Variabile	Pr > t	Variabile	Pr > t	Variabile	Pr > t
Interc.	0.1566	D_16_c		D_21_r	0.0523
D_10_cn	0.5002	D_17_s	0.0197	D_21_c	0.3146
D_10_gos	0.3095	D_17_c	0.0209	D_22_i	<.0001
D_10_ec_fi	0.0229	D_17_v	0.9535	D_22_q	0.1843
D_10_s	0.1604	D_17_com	0.3615	D_22_p	0.8859
D_10_at	0.4059	D_17_tda	0.0072	D_22_t	0.9111
D_10_p	0.1518	D_18_i	0.0046	D_22_r	0.5905
D_10_ni	0.0034	D_18_q	0.7688	D_22_c	0.1065
D_10_ec	<.0001	D_18_p	0.467	D_23_i	0.9714
D_10_m	0.7036	D_18_t	0.089	D_23_q	0.0213
D_11_m	0.2349	D_18_r	0.0249	D_23_p	0.947
D_11_pr	0.0463	D_18_c	0.2903	D_23_t	0.5922
D_11_pp	0.2651	D_19_i	0.0003	D_23_r	
D_11_tp	0.005	D_19_q	0.3346	D_23_c	0.0091
D_11_c	0.0492	D_19_p	0.0929		
D_11_n		D_19_t	0.5632		
D_12	0.7031	D_19_r	0.6461		
D_13	0.6931	D_19_c	0.2949		
D_14_cp	0.1422	D_20_orp	0.272		
D_14_st	0.1434	D_20_tt	0.6785		
D_14_lav	0.1435	D_20_ag	0.1181		
D_15_c	0.0342	D_20_d	0.0008		
D_15_uf	0.1312	D_20_fs	0.2468		
D_15_uni	0.0698	D_20_ccv	0.1012		
D_15_sa	0.5821	D_20_r	0.1613		
D_15_aa		D_20_gs	0.0191		
D_16_i	0.0776	D_20_qs	0.1355		
D_16_q	0.126	D_21_i	0.0055		
D_16_p	0.0161	D_21_q	0.1533		
D_16_t	0.1311	D_21_p	0.5441		
D_16_r	0.1808	D_21_t	0.0659		

regressione lineare(3)

◊ Per un'analisi più approfondita andiamo a considerare anche il TEST T, che valuta la significatività dei singoli coefficienti. Questo test dunque valuta l'adeguatezza di ogni singolo regressore vedendo se ciascuno è statisticamente legato alla Variabile dipendente.

Teniamo in considerazione i regressori a cui è associato un p-value piccolo, cioè minore di 0.05: 20 regressori.

regressione lineare(4)

Variabile	coefficiente di correlazione	Variabile	coefficiente di correlazione	Variabile	coefficiente di correlazione
D_10_cn	-0.13002	D_16_i	0.17063	D_20_ag	0.02448
D_10_gos	-0.09117	D_16_q	-0.03527	D_20_d	-0.03404
D_10_ec	0.33503	D_16_p	-0.13416	D_20_fs	0.09767
D_10_s	0.02426	D_16_t	-0.15116	D_20_ccv	-0.07823
D_10_at	0.20068	D_16_r	0.10774	D_20_r	-0.0258
D_10_p	0.1918	D_16_c	0.23187	D_20_gs	-0.02583
D_10_nl	0.09487	D_17_s	-0.1281	D_20_qs	0.0996
D_10_ec	0.308	D_17_c	-0.1045	D_21_i	0.08562
D_10_m	0.03396	D_17_v	0.06697	D_21_q	0.0266
D_11_m	0.26065	D_17_com	-0.08666	D_21_p	0.04534
D_11_pr	-0.22892	D_17_tda	0.23141	D_21_t	-0.16245
D_11_pp	0.06862	D_18_i	0.18359	D_21_r	0.08253
D_11_tp	-0.14609	D_18_q	0.03723	D_21_c	0.08099
D_11_c	-0.0578	D_18_p	0.00767	D_22_i	0.14728
D_11_n	0.06052	D_18_t	0.02645	D_22_q	-0.04447
D_12	0.08031	D_18_r	0.09717	D_22_p	0.03015
D_13	0.04111	D_18_c	0.11493	D_22_t	-0.14127
D_14_cp	-0.21438	D_19_i	0.1252	D_22_r	0.17302
D_14_st	0.05747	D_19_q	0.03987	D_22_c	0.2295
D_14_lav	0.20953	D_19_p	0.1112	D_23_i	0.25352
D_15_c	-0.0841	D_19_t	-0.12952	D_23_q	0.03357
D_15_uf	0.16987	D_19_r	0.06981	D_23_p	0.06773
D_15_uni	0.05319	D_19_c	0.17405	D_23_t	0.02446
D_15_sa	0.03404	D_20_orp	-0.03408	D_23_r	.
D_15_aa	-0.21174	D_20_tf	0.23965	D_23_c	0.15282

◊ Un ulteriore indicatore puntuale, che mostra la relazione tra ogni singola Variabile indipendente e quella dipendente è il coefficiente di correlazione. Quindi prendiamo in considerazione solo i regressori il cui coefficiente di correlazione è $> |0.15|$.

regressione lineare(5)

sintesi di selezione stepwise									
Step	Variabile immessa	Variabile rimossa	Descrizione	Numero var in	R-quadro parziale	R-quadro del modello	C(p)	Valore F	Pr > F
1	D_10_ec		interesse per eventi culturali	1	0.1601	0.1601	41.5837	26.69	<.0001
2	D_15_aa		informazione all' aria aperta	2	0.0642	0.2244	29.8468	11.51	0.0009
3	D_11_pr		informazione durante pranzo	3	0.0491	0.2734	21.3545	9.32	0.0027
4	D_22_r		informazione alla radio	4	0.0448	0.3182	13.7747	9	0.0032
5	D_10_p		interesse per la politica	5	0.0214	0.3396	11.2087	4.4	0.0378

Abbiamo poi utilizzato il metodo di regressione automatica STEPWISE per scegliere i migliori regressori tra quelli selezionati nella slide precedente. Tale metodo, più preciso e accurato, ha scelto 5 regressori.

regressione lineare(6)

Presentiamo di seguito la regressione lineare (i) effettuata con i 5 migliori regressori, e l'analisi di multicollinearità (ii)

(i)

Stime dei parametri							
Variabile	Descrizione	DF	Stima dei parametri	Errore standard	Valore t	Pr > t	Stima standardizzata
Interc.	Interc.	1	5.41771	0.50587	10.71	<.0001	0
D_10_p	interesse per la politica	1	0.10426	0.04972	2.1	0.0378	0.14837
D_10_ec	interesse per eventi culturali	1	0.26179	0.05176	5.06	<.0001	0.36071
D_11_pr	informazione durante pranzo	1	-0.02204	0.00634	-3.48	0.0007	-0.24626
D_15_aa	informazione all' aria aperta	1	-0.03065	0.00788	-3.89	0.0002	-0.27235
D_22_r	informazione alla radio	1	0.13987	0.0501	2.79	0.006	0.19629

Radice MSE	1.36828	R-quadro	0.3396
Media dipendente	8.05634	R-quadro corr	0.3153
Coeff var	16.98391		

attenzione:

Il valore R-QUADRO non è ottimale in quanto piuttosto basso. Il modello non spiega il fenomeno in maniera esaustiva, sono state privilegiate chiarezza interpretativa e correlazione tra i regressori al fine di rendere l'analisi più comprensibile.

regressione lineare(7)

(ii)

Stime dei parametri								
Variabile	Descrizione	DF	Stima dei parametri	Errore standard	Valore t	Pr > t	Stima standardizzata	Inflazione varianza
Interc.	Interc.	1	5.55852	0.49918	11.14	<.0001	0	0
D_10_p	interesse per la politica	1	0.10591	0.04921	2.15	0.0329	0.14733	1.05293
D_10_ec	interesse per eventi culturali	1	0.24754	0.05218	4.74	<.0001	0.32437	1.05044
D_11_pr	informazione durante pranzo	1	-0.01977	0.00637	-3.1	0.0023	-0.21048	1.0332
D_15_aa	informazione all' aria aperta	1	-0.03195	0.00766	-4.17	<.0001	-0.27944	1.00811
D_22_r	informazione alla radio	1	0.09929	0.04893	2.03	0.0441	0.13796	1.03877

Poiché il VIF non supera la soglia di due non riscontriamo multicollinearità tra i regressori estratti: questi non sono cioè fortemente correlati tra loro.

- In sintesi:**
- ◇ 5 regressori
 - ◇ bassa correlazione tra regressori (assenza di multicollinearità)
 - ◇ p-Value associato a ogni regressore < 0.05
 - ◇ buona influenza dei regressori sulla Variabile dipendente

regressione lineare(8)

Variabile	Descrizione	DF	Stima dei parametri	Errore standard	Valore t	Pr > t	Stima standardizzata
Interc.	Interc.	1	5.57506	0.51543	10.82	<.0001	0
D_10_p	interesse per la politica	1	0.1512	0.05497	2.75	0.0069	0.22164
D_10_ec	interesse per eventi culturali	1	0.18475	0.05749	3.21	0.0017	0.25294
D_11_pr	informazione durante pranzo	1	-0.02855	0.00764	-3.74	0.0003	-0.28359
D_22_r	informazione alla radio	1	0.13685	0.05156	2.65	0.009	0.20453

Manca ancora un elemento che garantisca la bontà del modello risultato: trattasi dell' analisi di influenza (iii).

Radice MSE	1.24517	R-quadro	0.3187
Media dipendente	8.24603	R-quadro corr	0.2962
Coeff var	15.10021		

Questo processo ci consente di individuare delle situazioni particolari, osservazioni anomale che possono, per la loro natura, inflazionare il modello medesimo.

Osservazioni influenti sono quelle che soddisfano almeno una di queste due condizioni:

- DISTANZA di COOK (D) > 1
- H di LEVERAGE > $2*(p+1)/n$

(alleghiamo il file excel cartellone2_out ove si rilevano le osservazioni influenti)



Foglio di lavoro di Microsoft Excel

Una volta individuate, le osservazioni influenti vengono eliminate dai regressori di input considerati nel processo di selezione automatica di stepwise, che verrà nuovamente fatto girare. Nella tabella abbiamo i 4 migliori regressori selezionati.

regressione lineare(9)

Stime dei parametri								
Variabile	Descrizione	DF	Stima dei parametri	Errore standard	Valore t	Pr > t	Stima standardizzata	Inflazione varianza
Interc.	Interc.	1	5.85393	0.52051	11.25	<.0001	0	0
D_10_p	interesse per la politica	1	0.14866	0.05555	2.68	0.0083	0.21373	1.16584
D_10_ec	interesse per eventi culturali	1	0.16836	0.0584	2.88	0.0046	0.22368	1.10034
D_11_pr	informazione durante pranzo	1	-0.0256	0.00772	-3.31	0.0012	-0.248	1.02453
D_22_r	informazione alla radio	1	0.07528	0.05229	1.44	0.1522	0.11077	1.08214

- in sintesi:**
- ◇ assenza di osservazioni influenti
 - ◇ 4 regressori
 - ◇ bassa correlazione tra regressori (assenza di multicollinearità)
 - ◇ p-value associato a ogni regressore < 0.05
 - ◇ R-quadro "accettabile"(anche se non molto alto)
 - ◇ buona influenza dei regressori sulla variabile dipendente

regressione lineare(10)

variabile	descrizione	stima standardizzata
D_10_p	interesse per la politica	0.14733
D_10_ec	interesse per eventi culturali	0.32437
D_11_pr	informazione durante pranzo	-0.21048
D_22_r	informazione alla radio	0.13796

Dall'analisi degli output possiamo concludere che coloro i quali sono interessati alla politica come anche agli eventi culturali, ritengono più importante mantenere il proprio livello di informazione aggiornato (es.: se l'interesse per la cultura aumenta di un'unità, si ha un aumento pari al 32.43% dell'esigenza di aggiornamento informativo).

Questo legame positivo con la variabile dipendente si riscontra anche per coloro che ritenendo molto affidabile la radio, si informano principalmente usufruendo di questo canale. Compare poi una fascia giornaliera in cui gli intervistati hanno dichiarato di informarsi: c'è una relazione negativa tra l'essere informati e chi predilige il primo pomeriggio come momento di informazione.

conclusioni(1)

L'obiettivo del nostro lavoro è stato quello di indagare ed analizzare la rilevanza dell'informazione e le diverse modalità di acquisizione della stessa. Per raggiungere questo fine sono state utilizzate diverse tipologie di analisi statistiche, procedendo per step graduali.

Si è dapprima analizzato le più importanti variabili del nostro questionario, divise in categorie, attraverso analisi univariate. Queste analisi sono state utili anche per valutare la correttezza dei dati in input.

Il passo successivo è stato quello di indagare, attraverso analisi bivariate, le relazioni più importanti esistenti tra le variabili dei dati personali e quelle costituenti la parte principale del questionario:

- ◊ L'importanza di mantenere aggiornato il proprio livello informativo in base al sesso;
- ◊ L'importanza dell'informazione rispetto al grado d'istruzione;
- ◊ L'argomento principale di interesse per sesso;
- ◊ L'argomento principale di interesse rispetto alle fasce d'età individuate;
- ◊ Il tempo speso ogniqualvolta ci si informa in rapporto alla professione svolta;
- ◊ La soddisfazione del canale scelto in rapporto alle fasce d'età individuate;
- ◊ L'utilizzo medio settimanale del canale in base al tipo di professione svolta;
- ◊ L'affidabilità delle fonti informative per le diverse fasce d'età individuate.

Dall'analisi sono emersi due spunti di riflessione :

1) Le maggiori differenze, per quanto riguarda sia la **scelta del canale** che la sua **affidabilità**, non sono riscontrabili né tra i due sessi né tra le diverse professioni, ma tra le diverse **classi di età** individuate.

conclusioni(2)

2) Per quanto riguarda invece il **tempo speso** e la **frequenza** con la quale si acquisisce informazione, la Variabile più discriminante è risultata la **professione**, probabilmente in virtù del diverso tempo a disposizione dalle diverse categorie.

La nostra indagine è, quindi, proseguita con un'analisi fattoriale. Questa è stata molto importante in quanto ci ha permesso di individuare, sommando e riassumendo l'informazione comune presente nelle diverse variabili, quali sono i fattori che maggiormente influenzano la scelta del canale informativo e della fonte a cui attingere le notizie. I fattori risultanti sono i seguenti:

A) RAPIDITA' DI ACQUISIZIONE DELL'INFORMAZIONE E QUALITA' DEL SERVIZIO OFFERTO

B) ESPOSIZIONE DELL'INFORMAZIONE

C) AFFINITA' POLITICA /IDEOLOGICA CON CHI DIVULGA L'INFORMAZIONE

D) ACCESSIBILITA' DEL SERVIZIO

E) ATTRATTIVITA' ARGOMENTI TRATTATI.

Il processo è stato poi completato con un'analisi di dipendenza (attraverso il modello di regressione lineare) che individua le variabili che maggiormente impattano (positivamente o negativamente) sull'importanza di aggiornare il proprio livello informativo, argomento che rappresenta il cuore della nostra indagine.

conclusioni. (3)

-implicazioni economico/manageriali-

Dunque dai dati rileviamo quanto fondamentale sia informarsi, per questa ragione lo si fa quotidianamente sfruttando le molteplici occasioni e mezzi di cui si può beneficiare. In particolare notiamo come negli ultimi anni lo sviluppo tecnologico abbia avvicinato sempre più i giovani all'utilizzo di internet, che è diventato uno fra i principali canali informativi. D'altro canto osserviamo una diffusione sempre minore dei convegni, forse per la loro caratteristica più specialistica e per gli argomenti di nicchia trattati.

Dalle evidenze emerse dunque ci sentiamo di formulare le seguenti proposte di business:

-utilizzo di internet come canale informativo sempre più approfondito, che possa essere utilizzato sia come fonte principale, sfruttando la sua rapidità di accesso all'informazione, la sua interattività e continuo aggiornamento, sia come canale di "supporto", ove approfondire tematiche che sono trattate superficialmente dagli altri canali.

-migliore organizzazione dei convegni: dai dati emerge una scarsa frequentazione di questi da parte del nostro campione e dunque dal pubblico giovane; riteniamo però possano essere un'occasione interessante di vivace scambio culturale, d'incontro, e confronto. Perciò pensiamo che questa risorsa debba essere meglio sfruttata. Per esempio sarebbe utile affiancare ai percorsi didattici tradizionali nelle università, la partecipazione obbligatoria ai convegni-organizzati in loco, come modalità alternativa di approfondimento di alcune tematiche legate agli argomenti d'esame, coinvolgendo in prima linea gli studenti, attribuendo loro un ruolo attivo (possibilità di intervistare alcuni "ospiti"-relatori ...)