



File di microdati per la ricerca

Aspetti della vita quotidiana

Periodo di riferimento: anno 2017

Descrizione del file

INDICE

Introduzione	3
Le misure di protezione adottate.....	4
Riferimenti bibliografici.....	9

Introduzione

I file di microdati per scopi di ricerca scientifica (MFR¹) possono essere rilasciati su richiesta motivata e previa autorizzazione del Presidente dell'Istituto, purché siano resi anonimi e privi di ogni riferimento che ne permetta il collegamento con le unità intervistate.

Si segnala che relativamente a questa stessa indagine è disponibile anche il file mlcro.STAT, liberamente scaricabile via web (https://www.istat.it/it/dati-analisi-e-prodotti/microdati#file_uso_pubblico). Il contenuto informativo di quest'ultimo è ridotto rispetto al file MFR a causa della maggiore protezione necessaria in mancanza di vincoli contrattuali.

Per la predisposizione del file MFR relativo all'indagine *Aspetti della vita quotidiana*, periodo di riferimento 2017, è stata adottata una opportuna metodologia, allo scopo di limitare il rischio di violazione della riservatezza. Lo scenario di intrusione considerato è stato quello relativo all'identificazione spontanea, ovvero conseguente a conoscenze *a priori* dell'utente che potrebbero permettere di attribuire correttamente i dati rilasciati alle unità della popolazione rilevata.

Le variabili coinvolte nel processo di protezione sono quelle che possono consentire l'associazione tra le informazioni e i rispondenti, ossia:

- gli *identificativi diretti*, che individuano in maniera univoca le unità statistiche di *rilevazione/analisi* (come ad esempio indirizzo e codice fiscale);
- gli *identificativi indiretti*, o *variabili chiave* (come ad esempio il numero di componenti della famiglia), che permettono di circoscrivere la popolazione cui appartengono i rispondenti e che, da sole o in combinazione con altre, possono portare alla re-identificazione di uno o più record.

Mentre i primi vengono cancellati dal file, i secondi vengono trattati allo scopo di limitare il rischio di violazione della riservatezza. Per il significato delle variabili e delle rispettive modalità si rinvia al questionario di indagine contenuto nella cartella metadati.

Si ricorda che gli utenti del file di microdati per scopi di ricerca scientifica sono obbligati ad adottare tutte le misure per la tutela della riservatezza cui si sono già impegnati all'atto della sottoscrizione del modulo per la proposta di ricerca:
<https://www.istat.it/it/files//2010/10/Allegato3-proposta-ricerca.pdf>

¹ MFR è l'acronimo per il file per la ricerca (Microdata File for Research). La documentazione metodologica di tali file è consultabile sul sito Istat (https://www.istat.it/it/dati-analisi-e-prodotti/microdati#file_ricerca).

Le misure di protezione adottate

Per la predisposizione del file per la ricerca scientifica *Aspetti della vita quotidiana*, periodo di riferimento 2017, sono state adottate le misure di protezione appresso specificate.

Il progressivo famiglia univoco a livello indagine [*profam*] è stato sostituito con un numero fittizio.

Le variabili peso [*peso*] e statura [*staf*] sono state sostituite con le variabili indice di massa corporea:

1. Indice di massa corporea per le persone di 18 anni e più [*bmi*], con le modalità *sottopeso*, *normopeso*, *sovrappeso*, *obeso*;
2. Persone di 6-17 anni in eccesso di peso [*bmimin*], con le modalità *non in eccesso di peso*, *in eccesso di peso*.

Soppressione di variabili

1. Provincia di residenza [*prov*]
2. Comune di residenza [*com*]

Ricodifiche di variabili categoriali

1. Regione di residenza [*reg* nel file originario → *regmf* nel file MFR]
Aggregazione delle modalità *Provincia autonoma di Trento* e *Provincia autonoma di Bolzano* in un'unica modalità *Trentino-Alto Adige*
2. Ampiezza demografica del comune di residenza [*dom* → *dommf*]
Aggregazione delle modalità originarie nelle seguenti:

Etichetta	Descrizione
1	area metropolitana
3	comuni aventi fino a 10.000 abitanti
5	comuni con oltre 10.000 abitanti

3. Età [*eta* → *etam*]

Aggregazione dell'età in anni compiuti nelle seguenti classi di età:

Etichetta	Descrizione
001	0-2 anni
002	3-5 anni
003	6-10 anni
004	11-13 anni
005	14-15 anni

006	16-17 anni
007	18-19 anni
008	20-24 anni
009	25-29 anni
010	30-34 anni
011	35-39 anni
012	40-44 anni
013	45-49 anni
014	50-54 anni
015	55-59 anni
016	60-64 anni
017	65-74 anni
018	75 anni e più

4. Stato civile [*stciv* → *stcivm*]

Aggregazione delle modalità *separato/a di fatto*, *separato/a legalmente*, *divorziato/a già in unione civile (per scioglimento unione)*

5. Titolo di studio [*istr* → *istrm*]

Aggregazione delle modalità originarie nelle seguenti:

Etichetta	Descrizione
01	laurea e post-laurea
07	diploma
09	licenza di scuola media
10	licenza di scuola elementare
11	nessun titolo di studio

6. Condizione occupazionale [*cond* → *condm*]

Aggregazione delle modalità *inabile al lavoro* e *altra condizione*

7. Posizione nella professione [*posiz* → *posizm*]

Aggregazione delle modalità originarie nelle seguenti:

Etichetta	Descrizione
1	dirigente, autonomo come imprenditore o libero professionista
2	direttivo, quadro o impiegato
3	capo operaio, operaio subalterno e assimilati, apprendista o lavorante a domicilio per conto d'impresa
4	lavoratore in proprio, socio cooperativa Produzione Beni e/o prestazioni di servizio, coadiuvante, collaborazione coordinata e continuativa (con o senza progetto) o prestazione d'opera occasionale

8. Settore di attività economica [*ateco* → *atecom*]

Aggregazione delle modalità originarie nelle seguenti:

Etichetta	Descrizione
01	agricoltura, silvicoltura e pesca
02	totale industria escluse costruzioni
03	costruzioni
04	totale servizi

9. Frequenza scolastica [*frsc* → *frscm*]

Aggregazione delle modalità originarie nelle seguenti:

Etichetta	Descrizione
01	NO
02	Dottorato di ricerca
03	Corsi di laurea o AFAM
06	Scuola secondaria di secondo grado
13	Scuola secondaria di primo grado
14	Scuola primaria
15	Scuola dell'infanzia e Asilo nido

10. Tipologia familiare [*tipfa2* → *tipfa2m*]

Aggregazione delle modalità:

Etichetta	Descrizione
01	persona sola
02	genitore con figli non celibi o nubili
03	insieme di parenti
04	parenti ed altri
05	persone non parenti
06	coppia coniugata senza figli, senza isolati
07	coppia non coniugata senza figli senza isolati
08	coppia coniugata con figli senza isolati
09	coppia non coniugata con figli senza isolati
10	monogenitore maschio celibe senza isolati
11	monogenitore maschio separato di fatto senza isolati; monogenitore maschio separato legalmente senza isolati; monogenitore maschio divorziato senza isolati
14	monogenitore maschio vedovo senza isolati
15	monogenitore femmina nubile senza isolati
16	monogenitore femmina separata di fatto senza isolati; monogenitore femmina separata legalmente senza isolati; monogenitore femmina divorziata senza isolati
19	monogenitore femmina vedova senza isolati

20	coppia coniugata senza figli con isolati
21	coppia non coniugata senza figli con isolati
22	coppia coniugata con figli con isolati
23	coppia non coniugata con figli con isolati
24	monogenitore maschio celibe con isolati
25	monogenitore maschio separato di fatto con isolati; monogenitore maschio separato legalmente con isolati; monogenitore maschio divorziato con isolati
28	monogenitore maschio vedovo con isolati
29	monogenitore femmina nubile con isolati
30	monogenitore femmina separata di fatto con isolati; monogenitore femmina separata legalmente con isolati; monogenitore femmina divorziata con isolati
33	monogenitore femmina vedova con isolati
34	due nuclei a due generazioni senza isolati
35	due nuclei di tipo fraterno senza isolati
36	due nuclei di altro tipo senza isolati
37	due nuclei a due generazioni con isolati
38	due nuclei di tipo fraterno con isolati
39	due nuclei di altro tipo con isolati
40	tre o più nuclei senza isolati

Ricodifiche di variabili quantitative

Per le variabili:

- numero di notti in cui è stato ricoverato negli ultimi tre mesi [*ggrica* → *ggricam*]
- numero di volte in cui ha fatto ricorso all'assistenza domiciliare [*nassdo* → *nassdom*]
- numero di visite specialistiche (ultimi 12 mesi) [*nvismed12* → *nvismed12m*]
- numero esami del sangue (ultimi 12 mesi) [*nansang12* → *nansang12m*]
- numero esami specialistici (ultimi 12 mesi) [*naccer12* → *naccer12m*]

i valori numerici sono stati aggregati secondo la seguente classificazione:

Etichetta	Descrizione
puntuali	Per i valori minore e uguali a 15
16	16 – 20
17	21 – 30
18	31 – 40
19	41 – 50
20	51 e oltre

Top coding di variabili quantitative

Per le seguenti variabili è stato applicato il metodo di *top-coding*, che consiste nel lasciare i valori puntuali fino a una certa soglia e ricodificare i valori puntuali più alti in un'unica classe, secondo i seguenti valori:

- numero motorini, scooter [*nmotor* → *nmotorm*]
- numero di motociclette, moto [*nmoto* → *nmotom*]
- numero automobili [*nauto* → *nautom*]
- numero di carte di credito [*nccred* → *nccredm*]

i valori superiori a 3 sono stati ricodificati in 4, con significato “4 e oltre”;

- numero di volte in cui è stato ricoverato negli ultimi tre mesi [*nnrica* → *nnricam*]

i valori superiori a 4 sono stati ricodificati in 5, con significato “5 e oltre”;

- consumo abituale al giorno di birra, in bicchieri [*bicbirra* → *bicbirram*]
- consumo abituale al giorno di vino, in bicchieri [*bicvino* → *bicvinom*]
- numero di bicchieri al giorno di aperitivi alcolici, amari o superalcolici [*bicaltro* → *bicaltrom*]

i valori superiori a 5 sono stati ricodificati in 6, con significato “6 e oltre”;

- numero di stanze nell'abitazione [*stanze* → *stanzem*]

i valori superiori a 9 sono stati ricodificati in 10, con significato “10 e oltre”;

- consumo di bevande alcoliche negli ultimi 12 mesi, numero bicchieri [*nbicalc* → *nbicalcm*]
- complessivamente in una settimana quanti bicchieri di vino o alcolici consuma abitualmente fuori dai pasti? [*bicfuori* → *bicfuorim*]
- quante sigarette fuma in media al giorno [*nsigar* → *nsigarm*]

i valori superiori a 20 sono stati ricodificati in 21, con significato “21 e oltre”.

Ulteriori interventi di protezione dei dati

In alcuni record sono stati modificati i valori puntuali in corrispondenza delle variabili *regione* e *tipologia comunale*: essi sono stati sostituiti in alcuni casi con il corrispondente livello ripartizionale, in altri casi con la modalità “non disponibile” (codice 999 per la regione; codice 9 per la tipologia comunale).

Ai fini della tutela della riservatezza, le variabili individuate dal D.Lgs 30/06/03 sono state trattate mediante permutazione casuale di una parte delle osservazioni.

A causa delle misure di protezione adottate, si possono verificare scostamenti rispetto ai dati pubblicati dall'Istat.

Riferimenti bibliografici

Hundepool, A., Domingo-Ferrer, J., Franconi, L., Giessing, S., Nordholt, E. S., Spicer, K. E. de Wolf, P.-P. (2012). Statistical Disclosure Control. Wiley.

Willenborg, L. e de Waal, T. (1996). Statistical Disclosure Control in Practice. Lecture Notes in Statistics, 111, New York: Springer-Verlag.

Willenborg, L. e de Waal, T. (2000). Elements of statistical disclosure control. Lecture Notes in Statistics, 115, New York: Springer-Verlag.

Curatori

Il documento è stato redatto da Ludovica Ioppolo.