



**File ad uso pubblico
micro.STAT**

Aspetti della vita quotidiana
Periodo di riferimento: anno 2016

Descrizione del file

Anno 2018

INDICE

Introduzione	3
Le misure di protezione adottate.....	4
Riferimenti bibliografici.....	8

Introduzione

I file di microdati ad uso pubblico (mlcro.STAT) sono collezioni di dati elementari, liberamente scaricabile via web (<http://www.istat.it/it/archivio/microdati+ad+uso+pubblico>), per le quali, a partire dal 2013, sono stati sviluppati anche i corrispondenti file per la ricerca (MFR¹). Il contenuto informativo di questi ultimi è superiore rispetto a quello del file ad uso pubblico².

Per la predisposizione del file mlcro.STAT relativo all'indagine *Aspetti della vita quotidiana*, periodo di riferimento 2016, è stata adottata una opportuna metodologia, allo scopo di limitare il rischio di violazione della riservatezza. Gli scenari di intrusione considerati sono stati:

- identificazione attraverso archivi esterni, ovvero attraverso il collegamento con i dati rilasciati da altre fonti pubbliche;
- 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.

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 (<http://www.istat.it/it/archivio/file+per+la+ricerca>).

² I ricercatori, che necessitano di informazioni maggiormente dettagliate, possono, su richiesta motivata e previa autorizzazione del Presidente dell'Istituto, richiedere il file di microdati per la ricerca.

Le misure di protezione adottate

Per la predisposizione del file ad uso pubblico *Aspetti della vita quotidiana*, periodo di riferimento 2016, 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 la variabile indice di massa corporea [*bmi*], con le modalità: sottopeso, normopeso, sovrappeso, obeso.

Soppressione di variabili:

1. Provincia di residenza [*prov*]
2. Comune di residenza [*com*]
3. Ampiezza demografica del comune di residenza [*dom*]
4. Settore di attività economica [*ateco*]
5. Frequenza scolastica [*frsc*]

Ricodifiche di variabili categoriali o quantitative:

1. Regione di residenza [*reg* nel file originario → *regmif* nel file mlcro.STAT]
 Aggregazione delle modalità *Provincia autonoma di Trento* e *Provincia autonoma di Bolzano* in un'unica modalità *Trentino-Alto Adige*
2. 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 – 34 anni
010	35 – 44 anni
011	45 – 54 anni
012	55 - 59 anni
013	60 – 64 anni
014	65 – 74 anni
015	75 anni e più

3. Stato civile [*stciv* → *stcivm*]

Aggregazione delle modalità *separato/a di fatto, separato/a legalmente e divorziato/a*

4. 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, nessun titolo di studio

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

Aggregazione delle modalità originarie nelle seguenti:

Etichetta	Descrizione
1	occupato
2	in cerca di occupazione
3	inattivo

6. 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

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

Aggregazione delle modalità:

Etichetta	Descrizione
01	senza nuclei
02	un nucleo senza isolati coppie
03	un nucleo senza isolati monogenitore
04	un nucleo con isolati coppie
05	un nucleo con isolati monogenitore
06	due nuclei senza isolati
07	due nuclei con isolati
08	tre o più nuclei

Variabili relative al ricovero in ospedale o in istituto di cura

Per la variabile:

- *ggrica*, numero di notti in cui è stato ricoverato negli ultimi tre mesi

i valori 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

Per la variabile:

- *nnrica*, numero di volte in cui è stato ricoverato negli ultimi tre mesi

è stato applicato il metodo di top-coding, in particolare i valori uguali e superiori a 5 sono stati ricodificati in 5, con significato “5 e oltre”.

Variabili relative ai mezzi di trasporto personali

Per le seguenti variabili:

- *nmotor*, numero motorini, scooter;
- *nmoto*, numero di motociclette, moto;
- *nauto*, numero automobili;

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

Variabili relative al consumo di bevande alcoliche

Per le seguenti variabili:

- *bicbirra*, consumo abituale al giorno di birra, in bicchieri;
- *bicvino*, consumo abituale al giorno di vino, in bicchieri;
- *bicaltro*, numero di bicchieri al giorno di aperitivi alcolici, amari o superalcolici;

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

Per le seguenti variabili:

- *nbicalc*, consumo di bevande alcoliche negli ultimi 12 mesi, numero bicchieri;
- *bicfuori*, complessivamente in una settimana quanti bicchieri di vino o alcolici consuma abitualmente fuori dai pasti?

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

Altre variabili quantitative

Per la seguente variabile:

- *nsigar*, quante sigarette fuma in media al giorno;

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

Per la seguente variabile:

- *nccred*, numero di carte di credito;

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

Per la seguente variabile:

- *stanze*, numero di stanze nell’abitazione;

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

Ulteriori interventi di protezione dei dati

In alcuni record sono stati modificati i valori puntuali in corrispondenza della variabile *regione*: essi sono stati sostituiti in alcuni casi con il corrispondente livello ripartizionale, in altri casi con la modalità “non disponibile” (codice 999).

Inoltre, sono stati inseriti valori mancanti in corrispondenza di una o più variabili.

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.