Istat
Istituto Nazionale di Statistica

Indagine sulle spese delle famiglie

Periodo di riferimento: anno 2015

Descrizione del file

Anno 2017

INDICE

Introduzione 3
Le misure di protezione adottate 4
Riferimenti bibliografici 6

Introduzione

I file di microdati per scopi di ricerca scientifica (MFR[1]) 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 mIcro.STAT, liberamente scaricabile via web (http://www.istat.it/it/archivio/microdati+ad+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 sulle spese delle famiglie nel 2015 è stata adottata una metodologia che limitasse 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 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 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 di richiesta

http://www.istat.it/it/files/2010/10/modulo_ricerca.pdf

[1] MFR è l'acronimo per il file per la ricerca. La documentazione metodologica di tali file è consultabile sul sito Istat (http://www.istat.it/it/archivio/file+per+la+ricerca).

Le misure di protezione adottate

Si riportano qui di seguito le misure di protezione adottate nella predisposizione del file per la ricerca scientifica per l'Indagine sulle spese delle famiglie nel 2015.

Soppressione di variabili

1. Codice famiglia
2. Codice fiscale dei componenti la famiglia
3. Percentuale di reddito percepito dai componenti la famiglia sul totale dei redditi familiari: sostituita dal principale percettore
4. Comune e provincia di residenza.

Ricodifiche di variabili categoriali o quantitative discrete

1. Tipologia di comune:
   1= centro area metropolitana
   2= periferia area metropolitana e comuni con 50.001 abitanti e più
   3= altri comuni fino a 50.000 ab. (diversi dai comuni periferia area metropolitana)
2. Età dei componenti la famiglia:
   [0,6), [6,15), [15,18), [18,25), [25,30), [30,35), [35,40), [40,45), [45,50), [50,55), [55,60), [60,65), [65,70), [70,75), [75 e oltre)
3. Numero di componenti: top-coding dal sesto componente (6 e più componenti). Tutte le variabili individuali dal settimo componente in poi sono state soppresse.
4. Relazione di parentela dei componenti la famiglia:
   1=persona di riferimento (PR)
   2=coniuge/convivente di PR
   3=figlio di PR
   4=altra relazione
5. Titolo di studio dei componenti la famiglia:
   1=nessun titolo
   2=scuola primaria (elementari)
   3=scuola secondaria di primo grado
   4= scuola secondaria di secondo grado
   5=laurea e post laurea
6. Attività economica dei componenti la famiglia:
   1=settore primario
   2=settore secondario
   3=settore terziario

7. Posizione nella professione dei componenti la famiglia:
   1=dirigenti e impiegati
   2=operai e assimilati
   3=imprenditori e liberi professionisti
   4=lavoratori in proprio
8. Fonte principale di reddito dei componenti la famiglia:
   1=reddito da lavoro dipendente
   2=reddito da lavoro autonomo
   3=reddito da pensione
   4=altri redditi
   5=mantenimento da parte di persone non conviventi
   6=mantenimento da parte di familiari conviventi (nessun reddito)
   7=non dispone di reddito, utilizza i risparmi o fa ricorso a prestiti
9. Professione dei componenti la famiglia:
   1=imprenditori e alta dirigenza
   2=professioni della conoscenza
   3=professioni esecutive di ufficio
   4=professioni qualificate nel commercio e nei servizi;
   5=professioni manuali qualificate
   6=professioni manuali non qualificate
   7=forze armate

Ricodifiche di variabili quantitative continue

1. Raggruppamento e top-coding delle variabili sulle spese delle famiglie relative a variabili sensibili:
   - spesa per occhiali e lenti da vista, spesa per apparecchi acustici, spesa per protesi dentarie e apparecchi ortodontici, spesa per sedie e carrozzine e spesa per riparazione di attrezzature sanitarie sono state raggruppate insieme;
   - spesa per medico generico (certificati, vaccinazioni o altre prestazioni a pagamento), spesa per visite specialistiche, spesa per ricoveri in strutture pubbliche; e spesa per ricoveri in strutture private sono state raggruppate insieme;
   - spesa per analisi cliniche e esami diagnostici (analisi del sangue, radiografie, ecografie, elettrocardiogrammi, ecc.), spesa per infermieri e altri servizi para-medici (fisioterapisti, podologi, logopedisti, agopunturisti), spesa per cure termali,

ginnastica correttiva e spesa per noleggio di attrezzature sanitarie o servizi di ambulanza sono state raggruppate insieme.

2. Per tutte le variabili relative alle spese sanitarie è stato effettuato un top-coding che seguisse la metodologia di Hubert e Vandervieren (2008).

Soppressioni locali

Sono stati inseriti valori mancanti in corrispondenza di tutte le variabili individuali dal settimo componente in poi. A causa delle soppressioni locali si possono verificare scostamenti rispetto ai dati pubblicati dall'Istat.

Riferimenti bibliografici

Hubert, M., Vandervieren, E. (2008). An adjusted boxplot for skewed distributions. Computational Statistics and Data Analysis 52, 5186-5201.

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

Andrea Cutillo

Flavio Foschi

Letizia Valdoni

Luigi Virgili

Il file di microdati per la ricerca è stato creato da Letizia Valdoni