

SN242

Rilevazione Continua sulle Forze di Lavoro - Annuale (2022)

Italian Labour Force Survey - Annual file (2022)

Istat

Versione/Version: 1.0 - 13/04/2022



UniData

Bicocca Data Archive

Website: www.unidata.unimib.it

E-mail: unidata@unimib.it

Tel.: +39 02 6448 7513

Fax: +39 02 6448 7561

La presente documentazione è distribuita da UniData secondo la [licenza CC-BY 3.0](#).
La fonte che ha prodotto i dati e UniData che li ha distribuiti non rispondono per alcun
utilizzo improprio dei dati e delle elaborazioni pubblicate.

*This documentation is distributed by UniData under the [CC-BY 3.0 License](#).
Neither the depositor nor UniData bear any responsibility for the analysis or
interpretation of the data produced by the user.*



Università degli Studi di Milano-Bicocca
Via Bicocca degli Arcimboldi 8
20126 - Milano (Italia)

Note metodologiche

Methodological notes

INDICE

TABLE OF CONTENTS

Aspetti metodologici dell'indagine <i>Survey's methodological issues</i>	p. 3
Descrizione del file <i>File description</i>	p. 16
Documentazione tecnica relativa al file di microdati per utenti MICRO.STAT <i>Technical documentation on microdata file for MICRO.STAT users</i>	p. 25

Aspetti metodologici dell'indagine
Survey's methodological issues



<Rilevazione sulle Forze di Lavoro>

<dal primo trimestre 2021>

Aspetti metodologici dell'indagine

<Edizione: Marzo 2022>

INDICE

1. Introduzione	3
2. La popolazione di riferimento.....	3
3. Il disegno campionario.....	3
4. La rilevazione e il trattamento dei dati	5
5. La metodologia di calcolo dei pesi campionari	6
6. La diffusione dei risultati dell'indagine	7
7. Glossario	9
8. Riferimenti bibliografici.....	10
9. Contatti.....	11

1. Introduzione

L'indagine ha la finalità di ottenere informazioni sulla situazione lavorativa, sulla ricerca di lavoro e sugli atteggiamenti verso il mercato del lavoro della popolazione in età lavorativa.

Il regolamento europeo 2019/1700, entrato in vigore a partire dall'edizione del primo trimestre 2021, prevede requisiti più precisi e vincolanti a favore dell'armonizzazione tra i paesi relativamente alla raccolta dei dati necessari a definire i principali indicatori.

In questa nota sono riportati in sintesi gli aspetti salienti dell'indagine. Per una trattazione più ampia e dettagliata si faccia riferimento al volume: ISTAT, *La rilevazione sulle forze di lavoro: contenuti, metodologie, organizzazione*, Metodi e Norme, Roma 2006

2. La popolazione di riferimento

La popolazione di interesse è costituita da tutti i componenti delle famiglie residenti in Italia, anche se temporaneamente emigrati all'estero, mentre esclude i membri permanenti delle convivenze (ospizi, istituti religiosi, caserme, eccetera). Il regolamento europeo 2019/1700 impone una nuova definizione di famiglia (di fatto): persone che coabitano abitualmente nella stessa casa, che condividono il cibo e altri beni primari, oltre a reddito e spese familiari.

Rispetto alla vecchia definizione, che come condizione necessaria per far parte della famiglia considerava il legame di parentela e/o affettivo, la nuova modifica la composizione familiare solo per le famiglie in cui uno o più membri non sono legati da vincoli di parentela.

I parametri di interesse dell'indagine sono espressi in termini di numero di individui che possiedono un dato attributo z e riguardano i seguenti aggregati di popolazione: forze di lavoro, occupati, persone in cerca di occupazione.

Relativamente a ciascuno di tali aggregati, i principali parametri oggetto di stima sono il numero di individui che presentano l'attributo z nel trimestre t e nell'anno a (ottenuto come media dei quattro trimestri dell'anno a) e la differenza tra il numero di individui che presentano l'attributo z nel trimestre t e il numero di individui che presentano il medesimo attributo in un trimestre precedente t' .

I domini territoriali di studio sono l'intero territorio nazionale, le ripartizioni geografiche e le regioni; a partire dal 1993, vengono considerate anche le province unicamente per la costruzione di stime medie annue.

3. Il disegno campionario

Il disegno di campionamento adottato in ciascun trimestre è un disegno a due stadi con stratificazione delle unità di primo stadio; le unità di primo stadio sono i comuni e le unità di secondo stadio sono le famiglie. All'interno di ciascuna provincia i comuni sono suddivisi in due sottoinsiemi: i comuni la cui dimensione demografica è superiore a una prefissata soglia sono detti comuni auto rappresentativi (Ar); i rimanenti comuni vengono denominati non auto rappresentativi (Nar). Ciascun comune Ar costituisce strato a sé stante e viene incluso con certezza nel campione; i comuni Nar, invece, vengono stratificati sulla base della dimensione demografica e da ogni strato così definito viene estratto un comune con probabilità proporzionale alla dimensione demografica. Dalla lista anagrafica di ogni comune campione viene selezionato, mediante scelta sistematica, un campione di famiglie; tutti gli individui appartenenti alle famiglie estratte vengono intervistati.

Con riferimento a ciascun trimestre la rilevazione è di tipo continuo e viene condotta tutte le settimane. Ogni famiglia campione viene intervistata una sola volta in una specifica settimana.

I campioni relativi a trimestri differenti sono parzialmente sovrapposti in base a uno schema di rotazione secondo cui una famiglia è inclusa nel campione per due rilevazioni successive e, dopo una pausa di due trimestri, viene reinserita nel campione per altre due rilevazioni. Ciò comporta una sovrapposizione del campione teorico del 50 per cento a un trimestre di distanza, del 25% a tre trimestri, del 50 per cento a quattro trimestri e del 25 per cento a cinque trimestri. Viene utilizzato, quindi, uno schema di rotazione trimestrale in cui le famiglie vengono intervistate per due trimestri consecutivi, escluse per due trimestri e successivamente re-intervistate per altri due trimestri¹. Nei quattro trimestri di rilevazione in cui ciascuna famiglia campione viene intervistata, la settimana di rilevazione è sempre la medesima, ad esempio se una famiglia viene intervistata per la prima volta nella seconda settimana del trimestre, anche nei successivi trimestri in cui è coinvolta verrà intervistata nella seconda settimana.

La sovrapposizione dei campioni trimestrali risponde all'esigenza di ridurre le fluttuazioni delle stime di livello riferite a trimestri differenti. Inoltre, la specifica tecnica di rotazione permette di ridurre la stima delle variazioni nette tra trimestri consecutivi e tra trimestri a un anno di distanza.

¹ Tenendo conto che lo schema trimestrale di rotazione delle famiglie è del tipo (2-2-2) e del fatto che nei quattro trimestri di rilevazione in cui viene intervistata ciascuna famiglia campione la settimana di rilevazione è sempre la medesima (ad esempio la famiglia viene intervistata sempre nella seconda settimana di ogni trimestre), le distanze effettive tra la prima intervista e le successive sono rispettivamente di 1, 4 e 5 trimestri (intendendo per trimestre un periodo di 13 settimane consecutive).

Il disegno di campionamento adottato per ciascuna rilevazione trimestrale si basa su una stratificazione temporale di tipo mensile; infatti, il campione trimestrale di unità finali di campionamento (le famiglie) viene suddiviso in tre gruppi distinti ciascuno dei quali viene assegnato casualmente a un mese del trimestre in modo tale che ciascuno dei tre gruppi costituisca un campione rappresentativo della popolazione di riferimento nel mese considerato.

A differenza delle famiglie, che ruotano secondo lo schema sopra introdotto, i comuni campione rimangono sempre gli stessi nel tempo.

Complessivamente, in ciascuna rilevazione trimestrale vengono coinvolti circa 1.400 comuni per un totale di circa 70 mila famiglie (circa 78 mila fino al 2011).

4. La rilevazione e il trattamento dei dati

La prima intervista di ogni famiglia presente nel campione è di tipo diretto, condotta mediante tecnica Capi (intervista faccia a faccia assistita da computer) mentre le tre successive sono da effettuarsi, in linea teorica, mediante intervista telefonica con tecnica Cati (con l'eccezione delle famiglie straniere e di quelle che non dispongono di un telefono fisso).

L'attività di rilevazione si interrompe in alcune settimane dell'anno in cui la reperibilità delle famiglie è piuttosto scarsa. In particolare, l'interruzione delle interviste è prevista per le prime tre settimane di agosto e per il periodo tra Natale e Capodanno². Le interviste da svolgersi durante l'interruzione vengono recuperate in un periodo successivo. In tale frangente, per evitare il sovraccarico della rete di rilevazione Capi, che altrimenti si troverebbe a gestire un numero troppo elevato di interviste con il rischio di un aumento del tasso di caduta, tutte le famiglie in prima intervista che dispongono di telefono vengono recuperate con tecnica Cati.

Le innovazioni introdotte in termini di tecnica di rilevazione e l'importante contenuto informatico a supporto comportano, una strategia di controllo e correzione dati diversa rispetto a quella usata per le indagini tradizionali. In particolare, il dato "grezzo" ha un livello qualitativo estremamente più elevato rispetto alle indagini eseguite con tecniche non assistite da computer. Il dato grezzo presenta, infatti, un numero estremamente ridotto di incongruenze tra le informazioni raccolte, in quanto queste ultime vengono sottoposte ad un piano di verifica implementato nel questionario elettronico che agisce sin dal momento

² Nei periodi indicati i rilevatori Capi generalmente fruiscono delle ferie.

della rilevazione. Tale piano si basa sul controllo dei domini delle variabili, dei percorsi del questionario e delle incongruenze logiche tra le informazioni raccolte.

La strategia adottata tratta essenzialmente due tipologie di errori: gli errori strutturali e quelli logico-formali; i primi derivano dai percorsi strutturali del questionario mentre i secondi dalle incongruenze logico formali delle informazioni raccolte. Nel questionario elettronico, ovviamente, sono implementate tutte le regole (denominate hard) che determinano i percorsi e il flusso del questionario stesso e gran parte di quelle logico-formali (denominate soft)³. In assenza di errori strutturali del questionario, quindi, le sole tipologie di correzione che vengono effettuate riguardano le incongruenze logico-formali derivanti da forzature di regole soft del questionario o di incongruenze non rilevate.

Infine i dati sottoposti ai trattamenti appena descritti passano al vaglio della “procedura famiglie”. Tale procedura permette, dopo aver controllato e corretto le notizie anagrafiche dei componenti, di costruire i nuclei presenti nella famiglia di fatto intervistata, e classificarla secondo un set di tipologie familiari.

5. La metodologia di calcolo dei pesi campionari

Lo stimatore scelto per la Rcfl è uno stimatore di ponderazione vincolata⁴, cioè uno stimatore in cui i pesi finali consentono di ottenere nell’ambito di diversi domini territoriali (regioni, province autonome di Trento e Bolzano, province, grandi comuni) stime di popolazione residente (per sesso e classi di età) uguali ai corrispondenti totali noti di fonte anagrafica. I pesi sono calcolati a livello familiare e ciò implica che a ciascun individuo appartenente alla stessa famiglia venga attribuito un medesimo peso di riporto all’universo; l’unicità del peso a livello familiare assicura la coerenza delle stime familiari con le stime individuali.

La realizzazione del coefficiente finale di riporto all’universo si articola in tre fasi distinte. La prima fase prevede il calcolo del coefficiente base ottenuto come reciproco della probabilità che ogni famiglia ha di far parte del campione. Per la determinazione del coefficiente base, si tiene conto della probabilità di estrarre un generico Comune dallo strato di appartenenza, e della probabilità di estrarre le famiglie campione dalla lista anagrafica di tale Comune. Ogni famiglia del campione, appartenente allo stesso Comune, ha pertanto lo stesso coefficiente base.

³ A differenza della regola hard, quella soft può essere disattivata dal rilevatore in maniera forzata qualora non riesca a rilevare informazioni coerenti.

⁴ Nella letteratura in lingua inglese tale stimatore è noto con il nome di *calibration estimator*; per ulteriori informazioni si veda Deville & Sarndal (1992).

Successivamente, per tenere conto delle interviste non realizzate in ciascun Comune, si determina un elemento correttivo del coefficiente base detto coefficiente per mancata risposta totale. In genere, il numero di famiglie intervistate per ciascun Comune del campione è inferiore al numero teorico previsto dal disegno campionario. Al fine di calcolare l'effettiva proporzione della correzione per mancata risposta totale. Attraverso un processo iterativo di ottimizzazione matematica che tiene conto dei totali noti, si determina infine il coefficiente finale di riporto all'universo. L'ultima fase riguarda il vincolo delle stime ottenute con il coefficiente base corretto per mancata risposta a "totali noti" desunti da fonti anagrafiche. La Rilevazione sulle Forze di Lavoro utilizza come totali noti: la popolazione residente per regione, sesso e quattordici classi d'età (0-14, 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65-69, 70-74, 75 e più), e la popolazione residente per provincia, sesso e cinque classi d'età (0-14, 15-29, 30-49, 50-64, 65 e più).

6. La diffusione dei risultati dell'indagine

La diffusione dei dati dell'indagine tiene conto delle molteplici esigenze degli utenti e si articola pertanto su diversi piani. Il primo è associato alla diffusione trimestrale del Comunicato stampa; il secondo al rilascio dei dati elementari; il terzo alle elaborazioni ad hoc; il quarto alla diffusione dei risultati dettagliati in media d'anno.

- Diffusione trimestrale

Dal I trimestre 2011 insieme all'uscita, in base al calendario prestabilito, del Comunicato stampa i risultati dell'indagine vengono diffusi attraverso il data-warehouse dell'Istituto I.Stat, raggiungibile all'indirizzo web: <http://dati.istat.it/>. Nella struttura gerarchica organizzata per aree tematiche del sistema I.Stat i risultati dell'indagine sulle forze di lavoro sono consultabili nella sezione dedicata al Lavoro. I dati sono presentati in forma aggregata in tavole multidimensionali, con la possibilità di comporre tabelle e grafici personalizzati, agendo sulle variabili, i periodi di riferimento e la disposizione di testate e fiancate. Il sistema è interrogabile secondo le seguenti modalità: per parola chiave, per tema e per territorio.

Fino al IV trimestre del 2010 i risultati venivano diffusi attraverso una serie di tavole in formato excel. Tali tavole, ancora disponibili nell'archivio dei comunicati stampa o nella sezione del data-warehouse I.Stat "*Esportazioni – File pronti all'uso*", riportano le serie storiche dei principali indicatori del mercato del lavoro. Si tratta dei file denominati:

1. popregio: che contiene i dati della popolazione distinti tra le forze di lavoro (occupati, persone in cerca di occupazione) e le non forze di lavoro (per grado di

attachement al mercato del lavoro, con meno di 15 anni, con più di 64 anni). I dati, distinti per genere, sono articolati a livello regionale e riportano anche i principali indicatori (tasso di attività, di occupazione, di disoccupazione).

2. setregio: contenente i dati relativi agli occupati -dipendenti, indipendenti, totale per macro settore di attività (agricoltura, industria in senso stretto, costruzioni, servizi). I dati, distinti per genere, sono articolati a livello regionale.
3. ripartiz: con i dati articolati per ripartizione geografica (Nord-ovest, Nord-est, Centro, Mezzogiorno) della popolazione, delle forze di lavoro, dell'occupazione, dei tassi di occupazione, delle persone in cerca di occupazione, dei tassi di disoccupazione, delle non forze di lavoro. Diverse tavole sono articolate oltre che per ripartizione e genere anche per classe di età, ovvero per settore di attività, per carattere dell'occupazione, per titolo di studio.
4. ripartiz-stranieri: con le principali tavole del file ripartiz elaborate con riguardo esclusivamente ai cittadini stranieri. Per alcune stime viene segnalata la presenza di un elevato errore campionario tale da sconsigliarne l'utilizzo.

Sempre nell'ambito della diffusione trimestrale viene fornito un file che riporta una breve serie storica dei principali aggregati dell'indagine e un apposito cartogramma che presenta una disaggregazione dei dati dei principali aggregati di riferimento dell'indagine: occupati, persone in cerca di occupazione, inattivi.

- Dati elementari

Il rilascio dei dati elementari avviene attraverso file distinti, ognuno orientato ad uno specifico gruppo di utenti:

1. Il file Sistan, in virtù di una specifica regolamentazione tra Istat e soggetti Sistan, contiene un'informazione particolarmente dettagliata. In particolare:
 - permette elaborazioni per famiglia mediante l'uso di un codice familiare;
 - adotta la classificazione a 4 digit delle attività economiche;
 - adotta la classificazione a 4 digit delle professioni;
 - i dati per singolo anno di età;
 - rilascia i codici sia delle regioni sia delle province e dei comuni.

Solo i dati relativi ad alcuni quesiti sono riservati e pertanto non vengono rilasciati.

È disponibile a partire dal I trimestre 2004.

2. Il file standard (ottenuto dopo il trattamento del contenuto informativo del file originale per la riduzione del rischio di violazione della riservatezza) contiene:
 - alcune variabili ricostruite al fine di facilitare le elaborazioni;

- i codici delle regioni ma non delle province e dei comuni;
- la classificazione a 2 digit delle attività economiche;
- la classificazione a 3 digit delle professioni;
- i dati per classi di età;
- l'oscuramento di alcune variabili familiari per motivi di riservatezza.

I dati relativi ad alcuni quesiti sono riservati e pertanto non vengono diffusi.

La produzione del file Standard è iniziata con il I trimestre 2004 ed è cessata con l'edizione relativa al IV trimestre 2013.

3. Il file per la ricerca che contiene:

- i codici delle regioni, delle province e dei grandi comuni;
- la classificazione a 4 digit delle attività economiche;
- la classificazione a 4 digit delle professioni (con un'aggregazione del primo grande gruppo);
- i dati per singolo anno di età;
- i dati per tutte le elaborazioni familiari.

Solo i dati relativi ad alcuni quesiti sono riservati e pertanto non vengono diffusi.

È disponibile a partire dal I trimestre 2009.

4. Il file mlcro.STAT, disponibile sul sito Istat, che presenta le seguenti caratteristiche:

- grazie ad una appropriata metodologia statistica volta a tutelare la riservatezza dei rispondenti, i mlcro.STAT possono essere scaricati liberamente e direttamente dal sito Istat;
- è ottenuto attraverso tecniche di anonimizzazione applicate al file per la ricerca per cui il contenuto informativo è minore rispetto al corrispondente file per la ricerca;
- precisione e accuratezza delle stime risultano identiche rispetto a quelle ottenute utilizzando i microdati originali oppure il corrispondente file per la ricerca

Il file ad uso pubblico è disponibile a partire dal I trimestre 2014.

Peraltro, il mondo della ricerca scientifica può avvalersi, anche con riguardo alla Rilevazione sulle forze di lavoro, dei servizi offerti dal Laboratorio per l'Analisi dei Dati ELEmentari (Laboratorio ADELE). In tale laboratorio i ricercatori possono, in un ambiente controllato, effettuare le proprie elaborazioni dei microdati dell'indagine. In questo caso i controlli relativi alla tutela della riservatezza sono preliminari al rilascio dei risultati prodotti.

- Elaborazioni ad hoc

Nonostante l'elevato numero delle elaborazioni effettuate, e rese disponibili nelle tavole diffuse correntemente, la vastità delle informazioni raccolte dall'indagine stimola l'ampia richiesta di elaborazioni ad hoc da parte degli utenti, singoli o istituzionali. Si tratta nella gran parte dei casi della richiesta di elaborazioni riferite a specifici livelli di dettaglio settoriale, territoriale, lavorativo, ecc. Il rilascio dei risultati delle elaborazioni ad hoc è comunque soggetto alla verifica della validità statistica dei risultati stessi, con riguardo in particolare al rilievo dell'errore campionario.

- Dati di media

A circa un mese dall'uscita dei risultati del quarto trimestre (marzo di ogni anno) vengono diffusi in rete, accompagnati da un comunicato stampa, i dati della media relativi all'anno precedente. Si diffondono i risultati dettagliati per la popolazione, le forze di lavoro, gli occupati, le persone in cerca di occupazione e le non forze di lavoro. Vi sono poi informazioni relative alla frequenza di attività di istruzione e/o formazione professionale, disaggregate in base alla condizione professionale, al sesso, all'età e al territorio. Infine, nei risultati di media sono presenti anche quelli riferiti alle province.

7. Glossario

Forze di lavoro: comprendono gli occupati e le persone in cerca di occupazione.

Occupati: comprendono le persone di 15 anni e più che nella settimana di riferimento:

- hanno svolto almeno un'ora di lavoro in una qualsiasi attività che preveda un corrispettivo monetario o in natura;
- hanno svolto almeno un'ora di lavoro non retribuito nella ditta di un familiare nella quale collaborano abitualmente;
- sono assenti dal lavoro (ad esempio, per ferie o malattia). I dipendenti assenti dal lavoro sono considerati occupati se l'assenza non supera tre mesi. Gli indipendenti assenti dal lavoro, ad eccezione dei coadiuvanti familiari, sono considerati occupati se, durante il periodo di assenza, mantengono l'attività. I coadiuvanti familiari sono considerati occupati se l'assenza non supera tre mesi.

Persone in cerca di occupazione: comprendono le persone non occupate tra 15 e 74 anni che:

- hanno effettuato almeno un'azione attiva di ricerca di lavoro nei trenta giorni che precedono l'intervista e sono disponibili a lavorare (o ad avviare un'attività autonoma) entro le due settimane successive all'intervista;

– oppure, inizieranno un lavoro entro tre mesi dalla data dell'intervista e sono disponibili a lavorare (o ad avviare un'attività autonoma) entro le due settimane successive all'intervista, qualora fosse possibile anticipare l'inizio del lavoro.

Inattivi: comprendono le persone che non fanno parte delle forze di lavoro, ovvero quelle non classificate come occupate o in cerca di occupazione.

Tasso di attività: rapporto tra le persone appartenenti alle forze di lavoro e la corrispondente popolazione di riferimento. La somma del tasso di attività e del tasso di inattività è pari al 100 per cento.

Tasso di occupazione: rapporto tra gli occupati e la corrispondente popolazione di riferimento.

Tasso di disoccupazione: rapporto tra le persone in cerca di occupazione e le corrispondenti forze di lavoro.

Tasso di disoccupazione di lunga durata: rapporto tra le persone in cerca di occupazione da dodici mesi e oltre e le forze di lavoro.

Tasso di inattività: rapporto tra le persone non appartenenti alle forze di lavoro e la corrispondente popolazione di riferimento. La somma del tasso di inattività e del tasso di attività è pari al 100 per cento.

Dato destagionalizzato: dato depurato dalla stagionalità.

Variazione congiunturale: variazione rispetto al trimestre precedente.

Variazione tendenziale: variazione rispetto allo stesso periodo dell'anno precedente.

Settimana di riferimento: settimana a cui fanno riferimento le informazioni raccolte.

8. Riferimenti bibliografici

- ISTAT, *“La rilevazione sulle forze di lavoro: contenuti, metodologie, organizzazione.”* Collana Metodi e norme n. 32 - 2006 – Roma
- Deville J. C., Särndal C. E., (1992), *“Calibration Estimator in Survey Sampling”*, Journal of the American Statistical Association, vol. 87, pp.376-382.
- ISTAT, *“La progettazione e lo sviluppo informatico del sistema Capi sulle forze di lavoro.”* Collana Metodi e norme n. 36 - 2008 – Roma

9. Contatti

Servizio Formazione Lavoro – e.mail: infolav@istat.it

Alessandro Martini – SWB – e.mail: alemartini@istat.it

Descrizione del file
File description



**File ad uso pubblico
micro.STAT**

Rilevazione continua sulle Forze di Lavoro
Periodo di riferimento: Dal primo Trimestre 2021

Descrizione del file

Maggio 2022

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, scaricabili via web (<http://www.istat.it/it/archivio/microdati-ad-uso-pubblico>).

Per la predisposizione del file mlcro.STAT relativo all'indagine continua sulle Forze di Lavoro, a partire dal primo trimestre 2021, 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:

- individuano in maniera univoca le unità statistiche di rilevazione/analisi (come ad esempio indirizzo e codice fiscale);
- permettono di circoscrivere la popolazione alla quale appartengono i rispondenti e, da sole o in combinazione con altre, possono portare alla re-identificazione di uno o più record.

Mentre le prime sono cancellate dal file, le seconde vengono trattate statisticamente riducendone il contenuto informativo. Per il significato delle variabili e delle rispettive modalità si rinvia al questionario di indagine contenuto nella cartella dei metadati.

Le misure di protezione adottate

Per la predisposizione del file ad uso pubblico dell'indagine continua sulle Forze di Lavoro periodo di riferimento, a partire dal primo trimestre 2021, sono state adottate le misure di protezione appresso specificate.

L'identificativo univoco della famiglia a livello indagine [CHIAVE_FAMIGLIA] è stato sostituito con un numero progressivo generato in maniera casuale.

Soppressione di variabili:

1. Provincia di residenza [prov]
2. Comune di residenza [com]
3. Grande Comune [gracom]
4. Settore di attività a 2 e 4 digit
5. Professione a 3 digit

Ricodifiche di variabili categoriali:

1. Classi di età: [0,2],[3,5],[6,10],[11,14), [15,19), [20,24), [25,29), [30,34), [35,39), [40,44), [45,49), [50,54), [55,59), [60,64), [65,69), [70,74), [75,120)
2. Residenza: la provincia di residenza è stata sostituita dalla regione con accorpamento di Piemonte e Valle d'Aosta e delle modalità Provincia autonoma di Trento e Provincia autonoma di Bolzano in un'unica modalità Trentino-Alto Adige
3. Titolo di studio: aggregazione delle modalità originarie nelle seguenti:

Etichetta	Descrizione
1	Nessun titolo
2	Licenza elementare
3	Licenza media-Diploma di istruzione secondaria di I grado
4	Diploma di qualifica professionale di scuola secondaria superiore (di II grado) di 2-3 anni
5	Diploma di maturità
6	ITS, Diploma universitario di due-tre anni
7	Laurea biennale specialistica-magistrale, Laurea a ciclo unico 4-6 anni
8	Dottorato

4. Tipologia familiare: aggregazione delle modalità originarie nelle seguenti:

Etichetta	Descrizione
1	Persona sola under 35;
2	Persona sola over 35;
3	Coppie senza figli senza isolati;
4	Coppie con figli senza isolati;

5	Monogenitore maschio senza isolati;
6	Monogenitore femmina senza isolati;
7	Altre famiglie

Ricodifiche di variabili quantitative:

1. Ore lavorate ORELAV→ORELAVMCR: aggregazione delle modalità originarie nelle seguenti:

Ore lavorate contrattuali QC31→QC31MCR

Ore lavorative abituali QC32→QC32MCR

Ore lavorate mediamente ultime quattro settimane QC33→QC33MCR

Ore lavorative desiderate QC35→QC35MCR

Ore lavorative settimana di riferimento QC44→QC44MCR

Ore lavorative abituali attività secondaria QD08→QD08MCR

Ore lavorative desiderate attività principale e secondaria QD11→QD11MCR

Etichetta	Descrizione
1	0-5 ore
2	6-10 ore
3	11-20 ore
4	21-30 ore
5	31-40 ore
6	41-50 ore
7	51-60 ore
8	Oltre 60 ore
9	n.d.

2. Ore lavorate extra QC43→QC43MCR

Ore lavorate attività secondaria nella settimana di riferimento QD09→QD09MCR

Etichetta	Descrizione
0	0 ore
1	1-5 ore
2	6-10 ore
3	11-15 ore
4	16-20 ore
5	21-30 ore
6	Oltre 30 ore
9	n.d.

3. Durata della Non occupazione DURRIC→DRICMCR

Durata della Non occupazione DURNO→DROCMCR

Durata della Non occupazione DURAD→DRADMCR

Etichetta	Descrizione
1	0-12 mesi
2	13-24 mesi
3	25-36 mesi
4	37-60 mesi
5	Oltre 60 mesi
9	n.d.

4. Numero di componenti della famiglia QSF04 → QSF04MCR

La modalità "7" indica 7 o più componenti

5. Anno di arrivo in Italia QSF24 → QSF24MCR

Anno di inizio lavoro attuale QC58 → QC58MCR

Anno di inizio primo lavoro QC64 → QC64MCR

Etichetta	Descrizione
1	0-5 anni fa
2	6-10 anni fa
3	11-20 anni fa
4	Più di 20 anni fa
9	n.d.

6. Durata del contratto QC24 → QC24MCR

Etichetta	Descrizione
1	0-6 mesi
2	7-12 mesi
3	13-24 mesi
4	25-36 mesi
5	Oltre tre anni
8	Non è specificata la durata/scadenza
9	Non sa

7. Anno di interruzione del lavoro QE02 → QE02MCR

Etichetta	Descrizione
1	0-2 anni fa
2	3-7 anni fa
3	Oltre 7 anni
9	n.d.

8. Anno di conseguimento titolo di studio QH04 → QH04MCR

Etichetta	Descrizione
1	0-2 anni fa
2	3-5 anni fa
3	6-10 anni fa
4	11-20 anni fa
5	21-30 anni fa
6	Oltre 30 anni fa
9	n.d.

9. Durata dello stage o tirocinio QH33 → QH33MCR

Etichetta	Descrizione
1	0-3 mesi
2	4-6 mesi
3	7-12 mesi
4	Oltre 12 mesi
9	Non sa

10. Tempo intercorso da ultima ricerca di lavoro QF06 → QF06MCR

Etichetta	Descrizione
1	0-3 mesi
2	4-6 mesi
3	7-12 mesi
4	13-24 mesi
5	Oltre due anni
8	Non ha ancora iniziato a cercare lavoro
9	Non sa

11. Tempo intercorso vivendo in Italia senza allontanarsi per più di un anno

QSF27 → QSF27MCR

Etichetta	Descrizione
1	0-4 anni
2	5-10 anni
3	11-20 anni
4	Oltre 20 anni
9	Non sa

12. Anno di matrimonio AMATRI → AMATRIMCR

Etichetta	Descrizione
1	0-5 anni
2	6-10 anni
3	11-20 anni
4	21-40 anni
5	Oltre 40 anni
9	Non sa

Ulteriori interventi di protezione dei dati:

Soppressioni locali: in alcuni record sono stati inseriti valori mancanti in corrispondenza di una o più variabili.

Alcuni record, relativi a individui non in età lavorativa, inoltre, sono stati soppressi.

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.

**Documentazione tecnica relativa al file di
microdati per utenti MICRO.STAT**
*Technical documentation on microdata file for
MICRO.STAT users*



Istituto Nazionale di Statistica

**Rilevazione
sulle Forze di Lavoro**

A grey laptop is shown open, with a screen displaying a data visualization. The laptop is resting on a stack of papers. The text is overlaid on this image.

**Documentazione
tecnica relativa
al file di microdati
per utenti
MICRO .STAT**

MEDIA 2022

Roma, Marzo 2022

AVVERTENZE

Per effettuare elaborazioni in media annua occorre tener conto di quanto segue:

- 1. Caricare nel proprio software di elaborazione i file dei microdati dei quattro trimestri separatamente utilizzando per ciascuno di essi il proprio tracciato record;**
- 2. Utilizzando il proprio software formare un solo file costituito dall'insieme dei record individuali dei singoli trimestri (per l'anno 2020 il numero di record del file di media è pari a $134679+128142+122830+127222=512873$);**
- 3. Calcolare (in una nuova variabile) il coefficiente di riporto per le stime di media annua dividendo per 4 il coefficiente utilizzato per le stime trimestrali;**
- 4. Il file di media annua permette elaborazioni per famiglia mediante l'uso di un codice familiare formato dal concatenamento delle variabili: ANNO, TRIMESTRE e MCRFAM;**
- 5. Per la valutazione della precisione delle stime di media annua utilizzare il file "Errori_campionari_2022_media.xls" (contiene i parametri dei modelli da utilizzare per il calcolo degli errori campionari a livello nazionale, ripartizionale e regionale).**

Informazioni:

Servizio Sistema integrato lavoro, istruzione e formazione

Via Cesare Balbo, 16 - 00184 – Roma

e-mail: infolav@istat.it