



Luglio 2015

**LABORATORIO ADELE:  
LE REGOLE PER IL RILASCIO DEI RISULTATI DELLE  
ELABORAZIONI**

Estratto da:  
Il Laboratorio per l'Analisi dei Dati ELEMENTARI (ADELE)  
Guida all'utenza

Per qualsiasi comunicazione o richiesta di informazioni: [adele@istat.it](mailto:adele@istat.it).

## Scopi del laboratorio

Il Laboratorio ADELE risponde alle esigenze di analisi statistica per finalità di ricerca scientifica che necessitano dell'utilizzo dei dati elementari, laddove questi non siano disponibili in altre forme (*cf.* Allegato 1).

All'interno del laboratorio, la sicurezza dei dati e il segreto statistico sono garantiti dal controllo sia dell'ambiente di lavoro che dei risultati delle analisi condotte dagli utenti.

Una volta concluse tutte le elaborazioni relative al progetto di ricerca, l'output di cui si chiede il rilascio sarà valutato sotto il profilo della riservatezza statistica dagli esperti del Laboratorio ADELE.

## Cosa è possibile ottenere e cosa non è possibile ottenere

Si riportano di seguito alcune regole per i tipi di output più frequenti.

### 1) Statistiche descrittive e tabelle

- Le statistiche descrittive e le tabelle sono particolarmente problematiche sotto il profilo della riservatezza statistica, e costituiscono output rilasciato con carattere di eccezionalità; devono pertanto limitarsi all'essenziale. Sono in genere consentite le comuni statistiche descrittive di supporto ai modelli statistici.
- Ciascun valore riportato nelle tabelle (di frequenza o di intensità), o nelle statistiche descrittive in genere, deve essere riferito ad almeno 10 unità non pesate.

Inoltre:

- statistiche descrittive che riportino dati puntuali sulle singole unità (ad esempio massimo e minimo per variabili continue) non possono essere rilasciate;
- moda, minimo e massimo: possono essere rilasciati se le modalità che individuano sono assunte da almeno 10 unità;
- quantili: la mediana è considerata rilasciabile se riferita ad una distribuzione di almeno 50 unità; gli altri quantili non sono rilasciati salvo casi particolari da concordare;
- medie, rapporti e indicatori: questi output devono essere presentati nella loro forma disaggregata (ad esempio per le medie e i rapporti: separare numeratore e denominatore; medie di variabili dicotomiche: presentare anche il complemento, etc.); ciascun elemento deve essere corredato dal numero di unità (almeno 10) che concorrono a determinarne il valore; ciò vale anche per eventuali complementi, anch'essi da presentare in forma esplicita (ad esempio, se un indicatore riporta il valore del 95%, si deve poter verificare che anche il 5% corrisponda ad almeno 10 unità; stessa cosa per le medie delle variabili dicotomiche, etc.).

### 2) Grafici

- I grafici su variabili non continue devono essere corredati dalla corrispondente tabella di valori che rappresentano; questa sarà valutata secondo quanto specificato al punto precedente. I grafici su variabili continue devono essere salvati come immagini; saranno valutati in merito alla eventuale identificabilità delle unità statistiche rappresentate.

### 3) Regressioni

- Le unità coinvolte nella regressione devono essere almeno pari a 100. Vengono rilasciate solo le stime dei coefficienti e non i loro residui.

## Regole per la presentazione dei risultati delle elaborazioni

- È fortemente sconsigliato produrre risultati senza l'impiego dei pesi di riporto all'universo; tuttavia, ai fini della valutazione, gli utenti devono presentare (anche) le frequenze non pesate delle analisi; l'utente è invitato ad indicare se le proprie elaborazioni fanno uso di pesi standardizzati (normalizzati) e in che modo (se la normalizzazione è rispetto al totale della popolazione o a sottopopolazioni specifiche).
- Il **volume** dell'output può essere considerato esso stesso una ragione di rifiuto al rilascio: l'output di cui si chiede il rilascio deve essere minimale e corrispondere a quanto sarà incluso nel lavoro che si intende divulgare; a titolo indicativo, viene suggerito un numero massimo di 30 pagine (~ 60Kb in ascii text format).
- L'output deve essere preferibilmente fornito in file di testo, oppure in file Word o Excel, ma non nel formato proprietario delle applicazioni statistiche utilizzate; eventuali statistiche descrittive e tabelle devono essere fornite in formato Excel.
- L'output deve essere redatto in modo da poter essere rilasciato così com'è, senza necessità di modifiche da parte del personale che ne effettua la valutazione: in caso di output non rilasciabile sarà necessario proseguire le elaborazioni per rendere l'output rilasciabile.
- L'output deve essere chiaramente ed estesamente documentato secondo la "Scheda per la descrizione dell'output" (*cf.* Allegato 2), nella quale va specificato: lo scopo e le modalità dell'analisi, nome e contenuto dei file di output, i trattamenti effettuati sul data set originario e le eventuali (sub)popolazioni oggetto d'analisi, il significato di ciascuna variabile (per quelle derivate anche la definizione), ed ogni altra informazione utile ad una corretta interpretazione dei file di output. La descrizione dell'output deve essere sufficiente a comprenderlo (non è consentito il riferimento ad altre fonti quali, ad esempio, i file di sintassi utilizzati).
- **Non è consentito il rilascio di output intermedi** (ovvero di elaborazioni che non concludano il progetto).

*Al termine di ciascun progetto agli utenti viene chiesto di rispondere facoltativamente a un breve questionario finalizzato a valutare gli aspetti del servizio dal punto di vista dell'utente. Il modulo non fa riferimento a dati personali; le informazioni raccolte sono utilizzate esclusivamente per produrre dei report sulla qualità del servizio e non sono in alcun modo diffuse associandole a dati personali sugli utenti. Tale questionario verrà inviato via email all'utente al termine del progetto.*

## **ALLEGATO 1 : I servizi e i prodotti per l'utenza specialistica offerti dall'Istat**

L'Istituto nazionale di statistica è impegnato in un costante sforzo volto ad assicurare la fruibilità dei risultati delle rilevazioni statistiche nella maniera più semplice e completa possibile. Tale impegno si traduce in una molteplicità di canali di accesso ai dati ed in numerosi servizi e prodotti offerti all'utenza.

Per gli utenti che necessitino di elaborazioni particolari o debbano poter utilizzare i dati elementari d'indagine per specifiche esigenze di ricerca, l'Istat mette a disposizione i File di microdati ed i servizi delle Elaborazioni personalizzate e del Laboratorio ADELE.

### ***Elaborazioni personalizzate***

Il servizio delle Elaborazioni personalizzate ([www.istat.it/it/informazioni/per-gli-utenti](http://www.istat.it/it/informazioni/per-gli-utenti)) consente di soddisfare quanti abbiano esigenze conoscitive non previste dai normali piani di diffusione (ad esempio ricercatori che necessitino di particolari tabelle non pubblicate) e che richiedano l'esecuzione di elaborazioni sui microdati. L'utente deve in questo caso illustrare le proprie necessità al personale addetto e contribuire economicamente a sostenere il costo dell'elaborazione. In generale, le elaborazioni non potranno essere troppo complesse ed il risultato sarà disponibile all'utente secondo costi e tempi che l'ufficio competente avrà cura di comunicare all'utente.

### ***File di microdati***

I File di microdati ([www.istat.it/it/prodotti/microdati](http://www.istat.it/it/prodotti/microdati)) sono collezioni di dati elementari d'indagine, nei quali l'anonimità delle unità statistiche, ovviamente già prive di identificativi diretti, viene tutelata tramite l'applicazione di diverse metodologie statistiche che rendono altamente improbabile l'identificazione indiretta delle unità statistiche.

I requisiti e le condizioni per il rilascio dei file variano a seconda dei soggetti che li richiedono e sono subordinati alla sottoscrizione di precisi accordi di utilizzo. Questi file si distinguono in:

- *File mIcro.STAT*: file ad uso pubblico scaricabili direttamente dal sito Istat: sono sviluppati a partire dai rispettivi *File per la ricerca*;
- *File standard*: sono rilasciati a seguito di richiesta motivata, , per finalità di studio e ricerca e sono prodotti per alcune rilevazioni dell'Istituto su individui e famiglie;
- *File per la ricerca*: vengono prodotti per rilevazioni statistiche riguardanti sia individui e famiglie sia imprese; sono realizzati specificatamente per esigenze di ricerca scientifica e contengono un maggiore livello di dettaglio informativo rispetto ai File standard. Il rilascio di tali file è soggetto alla sussistenza di alcuni requisiti sia soggettivi, relativi all'organizzazione di appartenenza, sia oggettivi, relativi alle caratteristiche del progetto di ricerca per le cui finalità viene richiesto il file.

## ALLEGATO 2

### SCHEDA PER LA DESCRIZIONE DELL'OUTPUT

#### DATI UTILIZZATI

Specificare, tra i dati forniti, quelli effettivamente utilizzati nelle elaborazioni di cui si chiede il rilascio: indicare il nome ed il periodo di riferimento della/e rilevazione/i utilizzate e specificare eventuali file di dati esterni impiegati nell'elaborazione.

#### DESCRIZIONE DELLE VARIABILI / INDICATORI

Riportare il nome ed una breve descrizione delle variabili utilizzate. Nel caso di variabili non presenti nelle basi di dati originarie (riclassificazioni effettuate dall'utente, variabili esterne etc.) oltre al nome ed alla descrizione, riportare il significato delle modalità assunte (o il procedimento di costruzione, soprattutto nel caso in cui la variabile assuma valori in funzione di altre variabili).

#### FILE DI OUTPUT

Riportare il nome e la struttura (esempio: file excel con un foglio per anno considerato) dei file di output dei quali si richiede il rilascio, fornendo una descrizione sintetica del contenuto.

#### ELABORAZIONI EFFETTUATE

Descrivere le singole elaborazioni effettuate, fornendone una descrizione breve ma esauriente.

È utile associare una denominazione a ciascuna elaborazione e riportarla nel file di output, così da poterne garantire una non equivoca identificazione ed interpretazione.

#### FILTRI SULLE UNITÀ

Per ciascuna elaborazione (o gruppo di elaborazioni) specificare i filtri applicati alla popolazione di partenza e la numerosità delle osservazioni coinvolte.

Notare che è necessario specificare esattamente la numerosità effettiva in ogni elaborazione, anche nel caso di riduzioni della numerosità dovute alla presenza di valori mancanti in una o più delle variabili adoperate.

#### SISTEMA DI PESI

Specificare il sistema di pesi eventualmente utilizzato e se questo varia tra le diverse elaborazioni.

Nel caso si faccia uso di pesi standardizzati (normalizzati), specificare se la normalizzazione è rispetto al totale della popolazione o a sottopopolazioni specifiche.

Notare che nel caso si richieda il rilascio di output pesato, lo stesso deve essere presentato anche in versione non pesata per consentirne la valutazione.

#### NOTE

Riportare ogni altra informazione si ritenesse utile ad una corretta interpretazione dei file di output.

Il richiedente: \_\_\_\_\_

Data: \_\_\_/\_\_\_/\_\_\_

**N.B.:** la descrizione dell'output deve essere sufficiente a comprenderlo; non è consentito il riferimento ad altre fonti (quali, ad esempio, i file di sintassi utilizzati).