



RESEARCH DATA CENTER

REX - Remote Execution / Sistema di elaborazione a distanza

Indagine sulle imprese industriali e dei servizi

Descrizione degli archivi



Novembre 2022

Indice

Indice.....	2
Descrizione degli archivi.....	3
1. Indagine sulle imprese industriali e dei servizi	3
1.1. Indicazioni generali	3
1.2. Aggiornamento periodico degli archivi	5
1.3. Variabili presenti sui questionari ma non reperibili negli archivi.....	5
1.4. Variabili presenti negli archivi ma non incluse nei questionari	5
1.4.1.Variabili di classificazione rispetto all'anno di rilevazione.....	5
1.4.2.Variabili di classificazione rispetto al settore di attività economica.....	6
1.4.3.Variabili di classificazione rispetto alla classe dimensionale	7
1.4.4.Variabili di classificazione rispetto all'area geografica	8
1.4.5.Variabili relative al disegno campionario e al sistema di ponderazione	8
1.4.6.Variabili di classificazione della quota del fatturato esportato.....	9
1.4.7.Variabili i cui livelli sono disponibili a prezzi correnti e costanti	9
1.4.8.Variabili imputate e relative indicatrici	10

Descrizione degli archivi

1. Indagine sulle imprese industriali e dei servizi

1.1. Indicazioni generali

Sono disponibili le informazioni sulle imprese per il periodo **1984-2014**. Le informazioni riferite alle varie indagini sono accodate in un unico archivio e sono separabili attraverso la variabile **annoril** che indica l'anno oggetto di rilevazione. La rilevazione viene effettuata nei primi mesi dell'anno successivo.

Ogni impresa ha un codice identificativo (la variabile **ident**) che consente l'aggancio delle informazioni nel caso in cui sia stata oggetto di rilevazione in più anni. Questo codice, generato in modo casuale, serve esclusivamente per le analisi longitudinali, non può essere tabulato, è del tutto incorrelato con le variabili identificative delle imprese e viene periodicamente rinnovato. La coppia (**annoril**, **ident**) identifica le informazioni riguardanti una data impresa in un dato anno.

Tavola 1

STRUTTURA DELL'ARCHIVIO DEI DATI

ANNORIL	IDENT	V24	V200	V209
1984	1	40	10	140
1985	2	190	100	2000
...
...
2020	1	35	20	500
2021	2	240	100	7000

La tavola 1 mostra la struttura logica degli archivi a puro titolo esemplificativo: l'utente non può accedere in alcun modo ai dati individuali né è abilitato a generare tavole simili.

Gli archivi contengono alcune variabili relative al riporto all'universo delle stime campionarie. L'adozione del peso campionario consente di allineare la struttura del campione a quella dell'universo secondo le variabili di stratificazione¹; se ne raccomanda l'uso nelle analisi per ottenere stime non distorte che riportino alla popolazione di riferimento.

Per le variabili che sono fornite dalle imprese sotto forma di rapporto, o comunque prive di fattore di scala, si dovrebbero in generale adottare pesi che tengano conto anche della dimensione del fenomeno. Gli archivi contengono anche alcune variabili relative alle classificazioni adottate per la stratificazione secondo il disegno campionario. Va tenuto presente che l'area geografica è considerata nel disegno dell'indagine come variabile di post-stratificazione. Per l'elenco delle variabili di ponderazione, stratificazione e post-stratificazione si veda il paragrafo 1.4.

¹ Poiché la numerosità della popolazione è nota con ritardo, talvolta anche di uno o due anni, i pesi sono provvisoriamente calcolati adoperando le numerosità della popolazione negli strati degli anni più recenti. I pesi sono poi ricalcolati quando i dati della popolazione divengono disponibili. I piccoli scostamenti tra l'effettiva numerosità della popolazione di riferimento e la somma dei pesi sono dovuti al processo di post-stratificazione.

Le informazioni monetarie provenienti dal questionario sono espresse a **prezzi correnti e in migliaia di euro** (ad eccezione delle variabili relative alle **retribuzioni**, espresse in **euro**), anche per quelle riferite agli anni in cui non era entrato in vigore l'euro.

Per diversi fenomeni, quali ad esempio occupazione, investimenti e fatturato, l'indagine rileva gli ammontari riferiti a più anni. Ad esempio, nella rilevazione sul 2021, è possibile riscontrare l'informazione dell'occupazione media nel 2021, nel 2020 e quella prevista per il 2022 (anno nel corso del quale l'intervista si svolge). Questo offre la possibilità di effettuare calcoli di variazioni annuali, senza la necessità di ricorrere all'utilizzo congiunto di più anni di indagine. Tale scelta, utile per la stabilizzazione della stima dei tassi di variazione, comporta, a livello di micro dato, che sia possibile osservare le revisioni dei valori riferiti ad anni diversi nel corso delle successive rilevazioni, come nel seguente esempio sull'occupazione media (variabili **v15**, **v24**, **v611m**) riferito ad una singola impresa:

anno indagine di riferimento	riferimento temporale del dato				
	2018	2019	2020	2021	2022
2019	115	120	125		
2020		120	125	80	
2021			121	100	79

corrispondente al seguente contenuto del database:

anno indagine di riferimento	impresa	occupazione media anno t-1	occupazione media anno t	occupazione media anno t+1 (prev.)
annoril	ident	v15	v24	v611m
2019	999	115	120	125
2020	999	120	125	80
2021	999	121	100	79

Nel corso dell'intervista, all'impresa rispondente vengono proposti, per memoria, i valori forniti l'anno precedente, tuttavia l'impresa è libera di rivedere i valori già forniti; nessuna imputazione è operata sulla base di tale confronto.

Negli archivi, limitatamente ai livelli di investimenti e fatturato, sono presenti variabili espresse sia a prezzi correnti che a prezzi costanti. I prezzi costanti sono riferiti sempre all'anno di riferimento più recente disponibile. Per l'elenco delle variabili interessate si veda il paragrafo 1.4.7.

Per alcune variabili, i dati mancanti sono imputati mediante apposita procedura; per consentire al ricercatore di tenere traccia di questo intervento, l'archivio contiene anche delle variabili flag (indicatrici). Per l'elenco delle variabili coinvolte si veda il paragrafo 1.4.8.

Nei questionari sono riportati i nomi delle variabili e i relativi codici di risposta per le variabili di tipo qualitativo. Per le variabili di tipo no/sì la decodifica è sempre 0=no; 1=sì.

A partire dall'indagine sul 2010, alcune variabili monografiche, specificamente indicate nel database, sono state rilevate su metà del campione. Inoltre, le sezioni dei questionari che le contengono sono chiaramente contrassegnate dalla lettera "A" o "B", per indicare se esse siano rivolte alla prima o alla seconda metà del campione.

La suddivisione del campione totale in due sotto-campioni è effettuata in base a un meccanismo casuale, tale da mantenere la rappresentatività delle due metà del campione originario rispetto alla popolazione di riferimento.

1.2. Aggiornamento periodico degli archivi

Gli archivi vengono creati anno dopo anno sulla base delle risposte fornite dalle imprese nel corso delle interviste. Le risposte sono sottoposte a un processo di controllo qualità che precede la formazione del dataset completo.

Le stime pubblicate sui Supplementi al Bollettino Statistico fanno riferimento all'archivio come disponibile al momento delle elaborazioni e all'universo delle imprese più recente disponibile.

Limitati scostamenti, tra l'archivio utilizzato per le stime del Supplemento e quello messo a disposizione per le elaborazioni a distanza, possono essere dovuti sia a successive revisioni delle risposte fornite dalle imprese, sia a revisioni nel sistema di pesi, dovuto all'aggiornamento della popolazione di riferimento.

Anno dopo anno, il dataset con i dati aggiornati al nuovo anno di riferimento viene messo a disposizione degli utenti del sistema nel corso dell'estate.

Agli utenti viene notificata via e-mail la disponibilità di un nuovo dataset. Mentre la revisione del sistema di ponderazione ha carattere sistematico, eventuali rettifiche operate sulle risposte degli anni precedenti sono da considerarsi del tutto episodiche.

1.3. Variabili presenti sui questionari ma non reperibili negli archivi

Per motivi di riservatezza, non sono disponibili le variabili del questionario collegate all'identificazione dell'impresa rispondente, tra le quali: Codice Fiscale, Ragione sociale, Filiale di rilevazione, Gruppo di appartenenza. Tali variabili corrispondono ai campi presenti all'inizio di ciascun questionario.

Non sono inoltre rese disponibili le risposte di tipo "Altro specificare".

1.4. Variabili presenti negli archivi ma non incluse nei questionari

Oltre alle variabili presenti nel questionario, l'archivio contiene alcune variabili derivate utili per le elaborazioni e le stime.

1.4.1. Variabili di classificazione rispetto all'anno di rilevazione

annoril: indica l'anno di riferimento dell'indagine (ad esempio: 2007 per l'indagine svolta tra gennaio e aprile 2008).

1.4.2. Variabili di classificazione rispetto al settore di attività economica

Fino al 2009 le seguenti variabili erano l'aggregazione di alcune sottosezioni ATECO 2002 ISTAT, specificate in tabella.

nome della variabile	valori assumibili
settore11	SS1 = Industrie alimentari, bevande e tabacco (DA) SS2 = Industrie tessili, dell'abbigliamento, pelli cuoio e calzature (DB, DC) SS3 = Fabbricazione di coke, industria chimica, gomma e plastica (DF, DG, DH) SS4 = Industria della lavorazione dei minerali non metalliferi (DI) SS5 = Industria metalmeccanica (DJ, DK, DL, DM) SS6 = Altre industrie manifatturiere (DD, DE, DN) SS7 = Altre industrie in senso stretto (CA, CB, E) SS8 = Commercio ingrosso e dettaglio (G) SS9 = Alberghi e ristoranti (H) SS10= Trasporti e comunicazioni (I) SS11= Attività immobiliari, informatica, etc.(K)
indag3	1 = INDUSTRIA MANIFATTURIERA - sezione ATECO 2002 ISTAT D 2 = INDUSTRIA ESTRATTIVO - ENERGETICA - sezioni ATECO 2002 ISTAT C ed E 3 = SERVIZI - sezione ATECO 2002 ISTAT G+I+H+K
indagine	1 = INDUSTRIA IN S.S. - sezione ATECO 2002 ISTAT C+D+E 2 = SERVIZI - sezione ATECO 2007 ISTAT G+I+H+K
Costruzioni	Indagine sulle costruzioni - sezione ATECO 2007 ISTAT F



Dal 2010 le variabili di classificazione basate sul settore di attività economica sono costruite in base ad aggregazioni di Divisioni della classificazione ATECO 2007 (prime due cifre) nel modo seguente.

nome della variabile	valori assumibili
settor11	SS1 = Industrie alimentari, bevande e tabacco (10,11,12) SS2 = Industrie tessili, dell'abbigliamento, pelli cuoio e calzature (13,14,15) SS3 = Fabbricazione di coke, industria chimica, gomma e plastica (19,20,21,22) SS4 = Industria della lavorazione dei minerali non metalliferi (23) SS5 = Industria metalmeccanica (24,25,26,27,28,29,30,33) SS6 = Altre industrie manifatturiere (16,17,18,31,32) SS7 = Altre industrie in senso stretto (05,06,07,08,09,35,36,37,38,39) SS8 = Commercio ingrosso e dettaglio, riparazione di autoveicoli e motocicli (45,46,47) SS9 = Servizi di alloggio e ristorazione (55,56) SS10 = Trasporto e magazzinaggio, servizi di informazione e comunicazione (49,50,51,52,53,58,59,60,61,62,63) SS11 = Attività immobiliari, professionali, scientifiche e tecniche, amministrative e di servizi di supporto (68,69,70,71,72,73,74,75,77,78,79,80,81,82)
indag3	1 = INDUSTRIA MANIFATTURIERA - sezione ATECO 2007 ISTAT C 2 = INDUSTRIA ESTRATTIVO - ENERGETICA - sezioni ATECO 2007 ISTAT B, D ed E 3 = SERVIZI - sezione ATECO 2007 ISTAT G+I+H+J+L+M+N
indagine	1 = INDUSTRIA IN S.S. - sezione ATECO 2007 ISTAT B+C+D+E 2 = SERVIZI - sezione ATECO 2007 ISTAT G+I+H+J+L+M+N
Costruzioni	Indagine sulle costruzioni – sezione ATECO 2007 ISTAT F

1.4.3. Variabili di classificazione rispetto alla classe dimensionale

Fino all'anno di riferimento 2003 la classe dimensionale è riferita al numero di addetti a fine anno; dal 2004 in poi al numero di addetti in media d'anno².

² Questa discontinuità è dovuta al fatto che uno dei criteri di inclusione delle imprese nel campione è la loro dimensione in termine di numero di addetti (tav. 1). Tale criterio è stato modificato nel senso indicato a partire dall'anno di riferimento 2004.

nome della variabile	valori assumibili
cldimet	0 = 20 - 49 addetti 1 = 50 - 99 2 = 100 - 199 3 = 200 - 499 4 = 500 - 999 5 = 1.000 addetti e oltre
cc2	1 = 20 - 49 addetti 2 = 50 addetti e oltre

1.4.4. Variabili di classificazione rispetto all'area geografica

Per motivi di riservatezza non sono disponibili le classificazioni relative alle singole regioni e province, ma solo quelle relative alle macroaree geografiche.

nome della variabile	valori assumibili
areag4:	1 = Nord ovest 2 = Nord est 3 = Centro 4 = Sud e Isole
areag2:	1 = Nord, Centro 2 = Sud e Isole

1.4.5. Variabili relative al disegno campionario e al sistema di ponderazione

- strato:** Formato dalle 68 combinazioni di **settor11** e **cldimet**. Si tenga presente che le imprese con almeno 5000 addetti hanno peso unitario e sono considerate, separatamente per industria e servizi, facenti parte di due strati a parte (rispettivamente 67 e 68).
- poststrato:** Formato dalle 48 combinazioni di **areag4**, **cc2** e una riaggregazione dei settori di attività economica in 6 gruppi: 1) indag3=1; 2) indag3=2; 3) settor11=SS8; 4) settor11=SS9; 5) settor11=SS10; 6) settor11=SS11.
- peso:** peso di espansione all'universo: a livello di strato e post-strato, la somma dei pesi uguaglia la numerosità della popolazione di riferimento, separatamente per ciascun anno senza tener conto della dimensione longitudinale del campione.
- pesoadd:** Peso campionario (di espansione): a livello di strato e poststrato, la somma dei pesi equivale alla numerosità degli addetti della popolazione di riferimento separatamente per ciascun anno e non tiene conto della dimensione longitudinale del campione (disponibile dal 2007).
- pesoa:** Equivalente a **peso** per le imprese appartenenti alla prima metà del campione (presente dal 2010, da usare per i quesiti monografici rilevati sulla prima metà del campione).

- pesoadda:** Equivalente a **pesoadd** per le imprese appartenenti alla prima metà del campione (presente dal 2010, da usare per i quesiti monografici rilevati sulla prima metà del campione).
- pesob:** Equivalente a **peso** per le imprese appartenenti alla seconda metà del campione (presente dal 2010, da usare per i quesiti monografici rilevati sulla seconda metà del campione).
- pesoaddb:** Equivalente a **pesoadd** per le imprese appartenenti alla seconda metà del campione (presente dal 2010, da usare per i quesiti monografici rilevati sulla seconda metà del campione).
- pesorisc:** La variabile **pesorisc** è ottenuta come prodotto di **peso** e un opportuno fattore di scala, in modo che essa, anno per anno, sommi alla numerosità del campione.
- popstr:** Numerosità della popolazione a livello di strato.
- poppopst:** Numerosità della popolazione a livello di post-strato.

1.4.6. Variabili di classificazione della quota del fatturato esportato

nome della variabile	valori assumibili
a6:	0= impresa non esportatrice 1 = meno di 1/3 di fatturato esportato 2 = tra 1/3 e 2/3 di fatturato esportato 3 = oltre 2/3 di fatturato esportato
qexp:	1 = meno di 1/3 di fatturato esportato o nessuna esportazione 2 = tra 1/3 e 2/3 di fatturato esportato 3 = oltre 2/3 di fatturato esportato

1.4.7. Variabili i cui livelli sono disponibili a prezzi correnti e costanti

Negli archivi, limitatamente ai livelli di investimenti e fatturato, sono presenti variabili espresse sia a prezzi correnti che a prezzi costanti. I prezzi costanti sono riferiti sia all'anno di riferimento più recente disponibile sia all'anno di riferimento di ogni singola indagine. Questi ultimi consentono di calcolare variazioni a prezzi costanti che non variano con l'aggiunta negli archivi di nuove edizioni dell'indagine. I deflatori per il calcolo dei prezzi costanti sono ricavati secondo una metodologia che utilizza – a livello di medie di sottosezione e classe dimensionale - i deflatori forniti dalle stesse imprese. Si rimanda alla Nota metodologica di ciascuno dei Supplementi al Bollettino Statistico – Indagini campionarie – dedicati all'indagine e reperibili sul sito internet della Banca d'Italia.

a prezzi correnti	a prezzi costanti riferiti all'anno più recente	a prezzi costanti riferiti all'anno di riferimento dell'indagine	descrizione
v200	v200cos^(a)	v200k^(a)	Investimenti fissi t-1
v202	v202cos	v202k	Investimenti fissi t
v203	v203cos	v203k	Investimenti fissi t+1
v209	v209cos	v209k	Fatturato t-1
v210	v210cos	v210k	Fatturato t
v437	v437cos	v437k	Fatturato t+1
v211	v211cos	v211k	Fatturato esportato t-1
v212	v212cos	v212k	Fatturato esportato t
v438	v438cos	v438k	Fatturato esportato t+1
v810	v810cos	v810k	Investimenti immateriali t-1
v811	v811cos	v811k	Investimenti immateriali t
v812	v812cos	v812k	Investimenti immateriali t+1

(a) Disponibile dal 1985.

Va osservato che con l'espressione "investimenti immateriali", secondo il SEC95, si intende la spesa per software, basi di dati e prospezioni minerarie. Sono esclusi invece i brevetti e i marchi, che nel SEC95 rientrano invece tra i consumi intermedi.

1.4.8. Variabili imputate e relative indicatrici

Per alcune variabili fondamentali, i dati mancanti sono imputati mediante apposita procedura; per consentire al ricercatore di tenere traccia di questo intervento, l'archivio contiene anche delle variabili *flag* (indicatrici), indicate con la lettera **f** seguita dal nome della variabile (ad esempio il contenuto della variabile **fv200** indica l'eventuale imputazione della variabile **v200**). Le variabili *flag* hanno valore 1 o sono vuote.

Il valore 1 contrassegna i dati imputati, che quindi sono riconoscibili e possono, volendo, essere esclusi dalle elaborazioni. La tavola 2 contiene l'elenco dei flag delle variabili soggette a imputazione.

Tavola 2

LISTA DEI FLAG DI IMPUTAZIONE

nome variabile	sezione del questionario	nome variabile	sezione del questionario
fser1	1-occupazione	fv27⁽¹⁾	1-occupazione
fser2	1-occupazione	fv27a	1-occupazione
fv15	1-occupazione	fv28⁽¹⁾	1-occupazione
fv16	1-occupazione	fv29⁽¹⁾	1-occupazione
fv17⁽¹⁾	1-occupazione	fv30⁽¹⁾	1-occupazione
fv18⁽¹⁾	1-occupazione	fv31	1-occupazione
fv18a	1-occupazione	fv32	1-occupazione
fv19⁽¹⁾	1-occupazione	fv31t	1-occupazione
fv20	1-occupazione	fv31f	1-occupazione
fv21⁽¹⁾	1-occupazione	fv31ft	1-occupazione
fv207	1-occupazione	fv803f	1-occupazione
fv207new	1-occupazione	fv611	1-occupazione
fv22	1-occupazione	fv611m	1-occupazione
fv23	1-occupazione	fv800	1-occupazione
fv24	1-occupazione	fv801	1-occupazione
fv25	1-occupazione	fv802	1-occupazione
fv26⁽¹⁾	1-occupazione	fv803	1-occupazione
		fv804	1-occupazione

nome variabile	sezione del questionario
fv805	1-occupazione
fv806	1-occupazione
fv807	1-occupazione
fv808	1-occupazione
fv809	1-occupazione
fv200	2-investimenti
fv201	2-investimenti
fv202	2-investimenti
fv203	2-investimenti
fv633	2-investimenti
fv634	2-investimenti
fv635	2-investimenti
fv636	2-investimenti
fv637	2-investimenti
fv638	2-investimenti
fv639	2-investimenti
fv640	2-investimenti
fv641	2-investimenti
fv642	2-investimenti
fv643	2-investimenti
fv644	2-investimenti
fv645	2-investimenti
fv646	2-investimenti
fv647	2-investimenti
fv648	2-investimenti
fv810	2-investimenti
fv811	2-investimenti
fv812	2-investimenti
fv816	2-investimenti
fv817	2-investimenti
fv818	2-investimenti
fv819	2-investimenti
fv820	2-investimenti

nome variabile	sezione del questionario
fv821	2-investimenti
fv822	2-investimenti
fv823	2-investimenti
fv824	2-investimenti
fv825	2-investimenti
fv826	2-investimenti
fv827	2-investimenti
fv828	2-investimenti
fv829	2-investimenti
fv830	2-investimenti
fv831	2-investimenti
fv451p	2-investimenti
fv451an	2-investimenti
fv814an	2-investimenti
fv217	3-capacità produttiva tecnica
fv219	3-capacità produttiva tecnica
fv220	3-capacità produttiva tecnica
fv441	3-capacità produttiva tecnica
fser5	4-fatturato
fser6	4-fatturato
fser7	4-fatturato
fv209	4-fatturato
fv210	4-fatturato
fv211	4-fatturato
fv212	4-fatturato
fv437	4-fatturato
fv438	4-fatturato
ffi12	5-finanziamento dell'impresa
ffi14	5-finanziamento dell'impresa
ffi16	5-finanziamento dell'impresa

(1) Flag relativi a variabili rilevate sino al 1995.