A Scuola di OPENCOESIONE

Conoscere e preparare un'indagine di statistica ufficiale

Rita Lima

Direzione Centrale per la Metodologia e il Disegno dei Processi Statistici

18 Aprile 2023

Istat

CONOSCERE E PREPARARE UN'INDAGINE DI STATISTICA UFFICIALE

Fasi di una indagine statistica, operazioni e principali fonti di errore (non campionario)

- La progettazione dell’indagine
- L’elaborazione dei dati raccolti
- La diffusione dei risultati
- La qualità dei dati

Alcuni esempi

L'INDAGINE STATISTICA E Il CICLO PRODUTTIVO DEI DATI

Per INDAGINE STATISTICA si intende un insieme di attività finalizzate ad approfondire la conoscenza di un fenomeno mediante alcuni passaggi (FASI):

- TRADUZIONE un concetto teorico (CARATTERE DI INTERESSE) in un concetto operativo (CARATTERE MISURABILE)

- RACCOLTA informazioni sul comportamento del singolo componente (UNITÀ STATISTICA) dell'insieme che si vuole indagare (POPOLAZIONE O FENOMENO COLLETTIVO)

Esempio

Fenomeno collettivo	Situazione occupazionale in Sardegna
Popolazione	Residenti in Sardegna in un certo momento
Unità statistiche	Singoli individui
Caratteri	Sesso, età, stato civile, condizione professionale, …
Osservazione	Registrazione delle risposte ad un questionario
Aspetti del fenomeno	• Quota disoccupati • Quota disoccupazione giovanile • Durata della disoccupazione • Relazione tra occupazione, disoccupazione e sesso • …

- TRATTAMENTO delle informazioni raccolte:
  - ELABORAZIONE dei risultati dell'OSSERVAZIONE di uno o più CARATTERI del FENOMENO COLLETTIVO
  - ANALISI dei dati raccolti (QUANTITATIVI e QUALITATIVI) e DIFFUSIONE dei risultati ottenuti

LE STATISTICHE UTILIZZANDO DATI GIÀ RACCOLTI DA ALTRI

Statistiche multifonte
Registri statistici
Big Data

- La produzione di statistiche, da parte di istituti preposti a tale funzione, costituisce un immenso patrimonio informativo a cui è facile attingere grazie anche alle possibilità delle nuove tecnologie

- I dati prodotti dalla FONTE STATISTICA possono essere usate in ANALISI SECONDARIE

SVANTAGGI

Archiviazione di dati con finalità diverse da quelle per cui vengono riutilizzati

- Dati non sono sempre "adeguati all'uso" rispetto ai propri obiettivi
- Problemi di comparabilità dati di diversa fonte
- È necessario conoscere il processo della loro produzione
- Occorre verificare e valutare la loro qualità

VANTAGGI

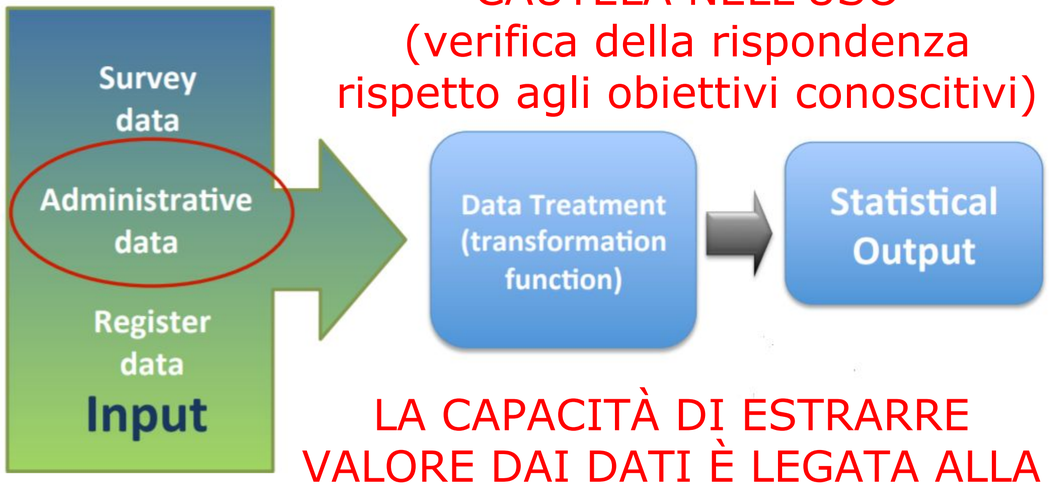
- Riduzione di costi e tempi di produzione
- Disponibilità di un ricco insieme di informazioni raccolte da più fonti
- Onere non eccessivo sui rispondenti
- Efficienza rispetto ai costi

ANALISI SECONDARIE – FONTI AMMINISTRATIVE

Si definiscono fonti amministrative le informazioni raccolte e conservate da istituzioni pubbliche:

- ✔ registrazione di uno stato di fatto
- ✔ riguardanti persone fisiche o giuridiche del territorio di competenza
- ✔ per finalità fiscali, pensionistiche, giuridiche o anagrafiche

esempi: iscritti in anagrafe per nascita, matrimoni, delitti denunciati, separazioni civili,ect.



CAUTELA NELL'USO
(verifica della rispondenza rispetto agli obiettivi conoscitivi)

LA CAPACITÀ DI ESTRARRE VALORE DAI DATI È LEGATA ALLA CAPACITÀ DI INTEGRARE DATI CHE PROVENGONO DA FONTI DIFFERENTI!

VANTAGGI

- Riducono il costo della rilevazione
- Riducono il disturbo statistico

SVANTAGGI

- Concetti, definizioni, classificazioni adottate possono non coincidere
- La qualità con cui le informazioni sono raccolte nei dati amministrativi può non essere sufficiente
- Non sempre è garantita la disponibilità dei dati in tempi utili alle necessità di informazione statistica

LE STATISTICHE UTILIZZANDO DATI GIÀ RACCOLTI DA ALTRI

I DATI SECONDARI – fonti di statistiche ufficiali

Come trovare le informazioni:

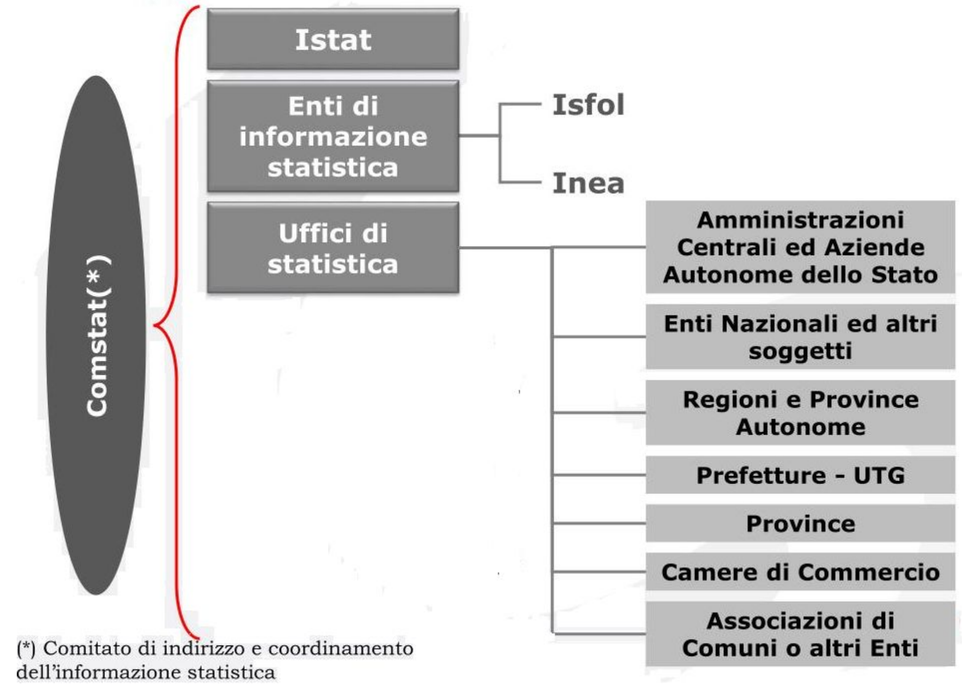
- Cataloghi (es. catalogo SISTAN)
- Uffici specializzati (es. Istat)
- Internet

- Pubblicazioni periodiche
- Banche dati
- Mezzi di informazione di massa

- WHO= Chi è il produttore? È riconosciuto ed autorevole a livello pubblico?
- WHY= Per quali obiettivi viene prodotta l'informazione statistica: per servire il paese e i decisori politici o per obiettivi commerciali?
- WHAT= La produzione statistica è riportata in programmi statistici normati? Quali fenomeni sono investigati? Le statistiche sono confrontabili?
- HOW= Si adottano metodologie solide, classificazioni standard, procedure e strumenti standardizzati, sistemi di controllo della qualità?

CHI FA PARTE DELLA STATISTICA UFFICIALE?

I PRODUTTORI DI STATISTICHE UFFICIALI E IL SISTAN



(*) Comitato di indirizzo e coordinamento dell'informazione statistica

SISTAN

Istat
Istituto Nazionale di Statistica

eurostat

ALTRI ORGANISMI E ENTI NAZIONALI E INTERNAZIONALI

Eurostat, OCSE, FMI, Bankitalia, ect.

ALTRI PRODUTTORI DI INFORMAZIONI STATISTICHE

Censis, Università, Centri Studi, Fondazioni, Istituti privati di ricerca e di ricerche di mercato

IL SITO ISTAT: COSA CERCARE?

L'Istituto nazionale di statistica (ISTAT), attivo dal 1926, è il maggiore produttore di dati economici e sociali del Paese e dal 2016 è un ente pubblico di ricerca, rivolto alla produzione e all'analisi di dati.

Provvede a:

- INDIRIZZARE e COORDINARE le attività statistiche degli altri soggetti del SISTAN
- ASSISTERE sugli aspetti tecnici enti e uffici del SISTAN

www.istat.it fonte d'informazione OPEN fruibile gratuitamente e riutilizzabile.

Microdati:

IDQUESTIONARIO	MODELLO	TIPALL	NOCC	NFAM	TITGOD	SUPERF
111111111111	L	1	5	1	4	200

collezioni di dati elementari/individuali

Macrodati:

dati aggregati, riferiti a gruppi omogenei di dati elementari

Popolazione residente al 1° Gennaio 2014 per età, sesso e stato civile - Italia

Eta'	Totale Maschi	Totale Femmine	Maschi + Femmine
100 e più	2.993	14.891	17.884

Metadati:

- descrizioni
- definizioni e informazioni

per interpretare correttamente i dati

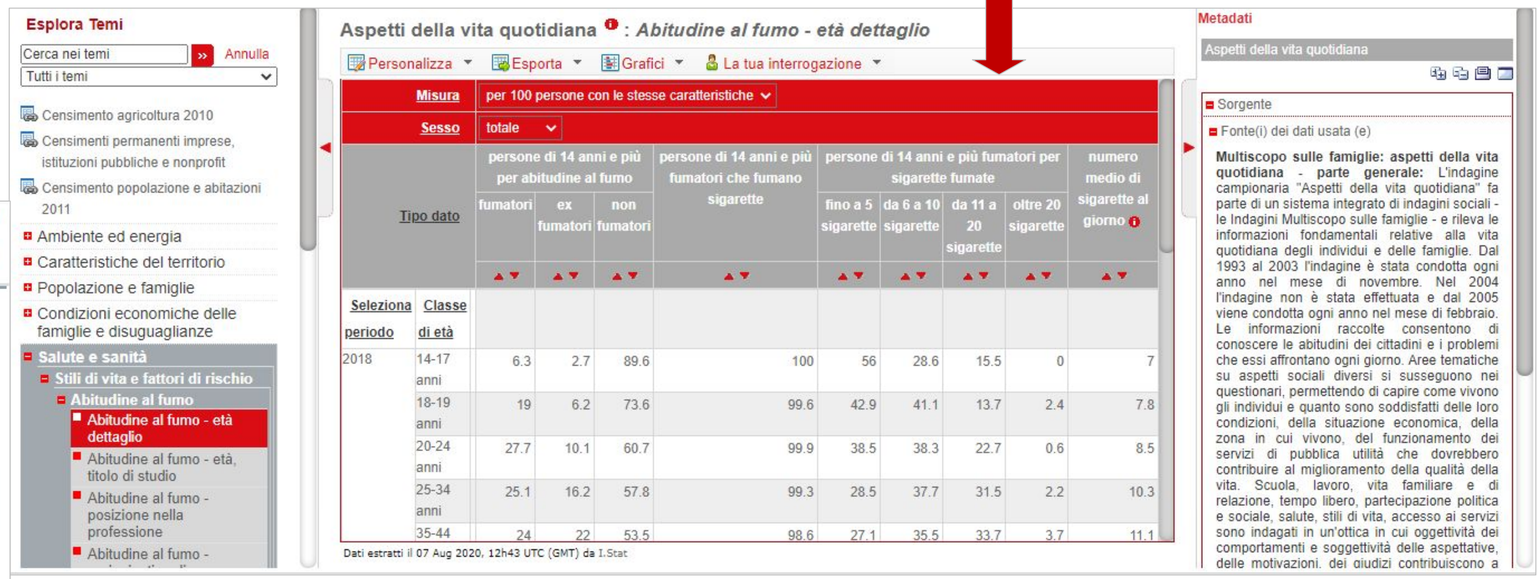
ESEMPIO: TRA I GIOVANI DI 14-19 ANNI CHI FUMA DI PIÙ?

Fonte: Rilevazione Istat Indagine campionaria multiscopo «Aspetti della vita quotidiana»

Collettivo: giovani di 14-19 anni

Estrazione dati: http://dati.istat.it/

Informazioni sul processo statistico:
https://www.istat.it/it/archivio/91926



In Italia nel 2018 fuma:

- ✔ il 6,3% dei ragazzi tra 14 e 17 anni
- ✔ il 19,0% dei ragazzi tra 18 e 19 anni

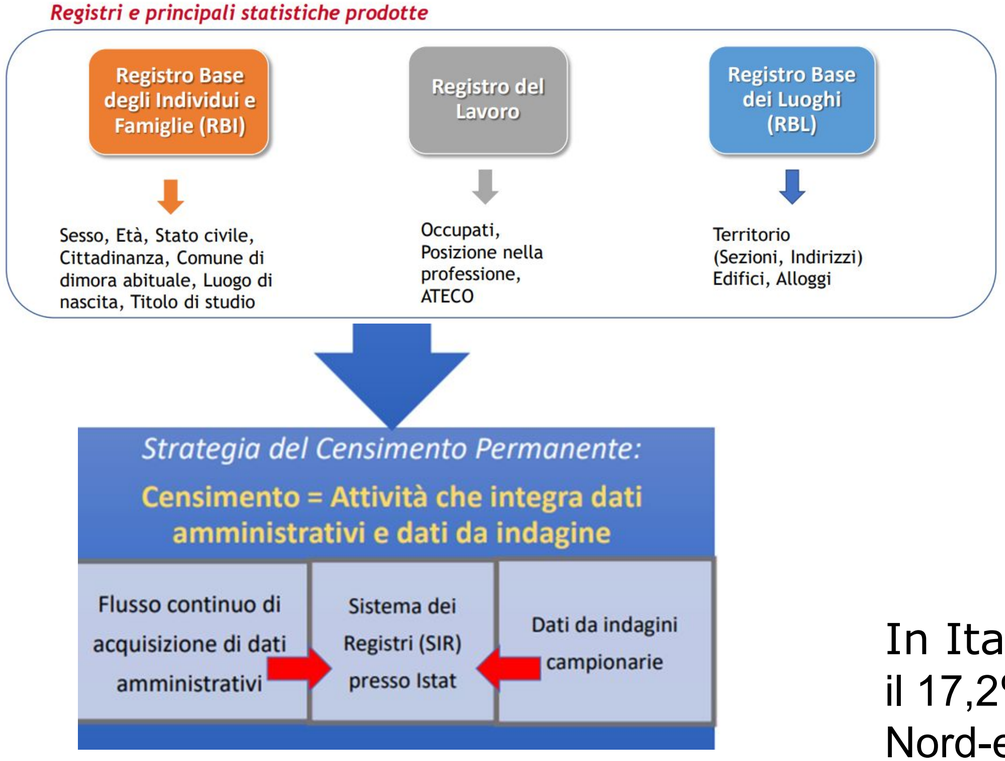
Fonte: Istat, Indagine multiscopo "Aspetti della vita quotidiana"

ESEMPIO: IN QUALE RIPARTIZIONE TERRITORIALE CI SONO PIÙ LAUREATI?

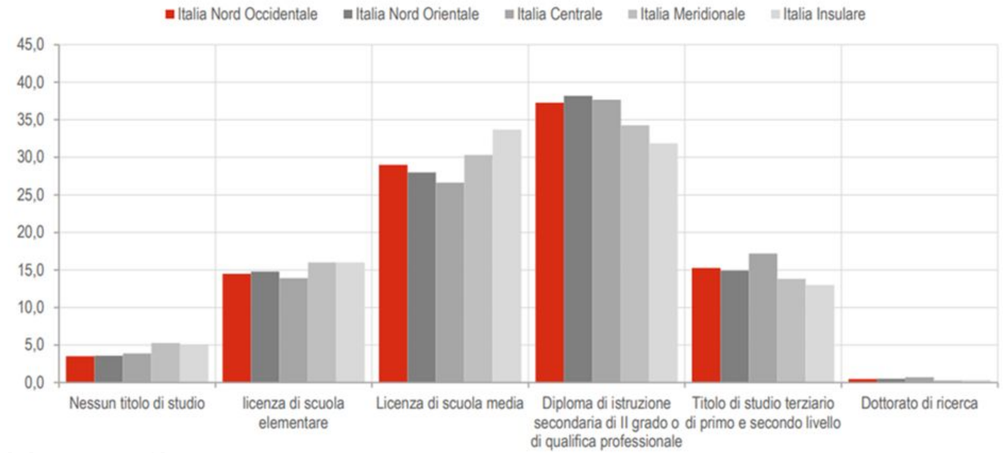
Fonte: Rilevazione campionaria Istat «Il censimento permanente»

Collettivo: popolazione di 9 anni e più per grado di istruzione e ripartizione

Informazioni sul processo statistico: dati: ttps://www4.istat.it/it/censimenti-permanenti/popolazione-e-abitazioni



Estrazione dati: http://dati-censimentipermanenti.istat.it/



In Italia nel 2021 i laureati sono:
il 17,2% al Centro, il 15,3% al Nord-ovest, il 14,9% al Nord-est, il 13,8% nel Meridione e il 13% nelle Isole.

Fonte: Istat, Censimento e dinamica demografica , 2021

ESEMPIO: QUALE E' IL FATTURATO MEDIO DELLE IMPRESE ITALIANE A CONTROLLO ESTERO?

Fonte: Rilevazione multifonte Istat « Struttura e competitività delle imprese multinazionali»

Collettivo: imprese e unità locali attive italiane a controllo estero nei settori ATECO da B a N e da P a S

Informazioni sul processo statistico: https://www.istat.it/it/archivio/263634

ELENCO DELLE FONTI INFORMATIVE IMPIEGATE PER LA PRODUZIONE DELLE STATISTICHE *INWARD* FATS

Fonti informative	Caratteri e variabili economiche impiegati per le stime
A. Panel delle imprese a controllo estero integrato con i risultati della rilevazione censuaria sulle imprese a controllo estero in Italia	Nazionalità estera del controllante ultimo e scambi con l'estero di merci e servizi (nel complesso e per la componente *intra-firm)*
B. Archivio dei gruppi di imprese	Struttura del gruppo e nazionalità del vertice, lista di imprese a controllo diretto estero.
C. Archivio statistico delle imprese attive in Italia (ASIA)	Caratteri anagrafici dell'impresa (codice di attività economica, numero di addetti, ecc), elenco delle imprese operanti in Italia con forma giuridica estera.
D. Principali giornali, pubblicazioni specialistiche, database commerciali, a livello nazionale e internazionale	Monitoraggio delle principali imprese a controllo estero in relazione ad eventi di entrata o uscita dalla popolazione di imprese a controllo estero nonché di cambiamento di nazionalità estera.
E. Rilevazione sul sistema dei conti delle imprese (SCI)	Fatturato, valore aggiunto al costo dei fattori, valore della produzione, costi per il personale, acquisti totali di beni e servizi, acquisti di beni e servizi per la rivendita senza trasformazione, investimenti fissi lordi
F. Dati Frame-SBS integrati con i risultati della rilevazione sulle piccole e medie imprese e sull'esercizio di arti e professioni (PMI)	
G. Rilevazione sulla Ricerca e sviluppo nelle imprese (RS1)	Spesa per Ricerca e sviluppo
H. Rilevazioni sul commercio estero (COE)	Esportazioni e importazioni di merci

Estrazione dati:
https://www.istat.it/it/imprese?dati/

PRINCIPALI INDICATORI ECONOMICI DELLE IMPRESE A CONTROLLO ESTERO RESIDENTI IN ITALIA PER MACROSETTORE. Anno 2019

	IMPRESE	ADDETTI	FATTURATO (mln di euro)	DIMENSIONE MEDIA
	Imprese a controllo estero residenti in Italia			
Industria	4.548	525.255	246.096	115,5
Servizi	11.231	985.935	378.071	87,8
TOTALE	15.779	1.511.190	624.167	95,8

Fonte: Istat, Struttura e competitività delle imprese multinazionali, 2019

ESEMPIO QUALE E' L'OPINIONE COLLETTIVA SULL'ECONOMIA DEL PAESE?

Fonte: Statistica sperimentale Istat «Il Social Mood on Economy Index»
Collettivo: circa 26.000 tweet giornalieri in lingua italiana di contenuto economico

Informazioni sul processo statistico: https://www.istat.it/it/archivio/21958534

POPOLAZIONE E FAMIGLIE | SOCIETÀ E ISTITUZIONI | ISTRUZIONE E LAVORO | ECONOMIA | AMBIENTE E TERRITORIO | CERCA NEL SITO | Statistiche A-Z | Glossario

[ENGLISH]

STATISTICA SPERIMENTALE

SOCIAL MOOD ON ECONOMY INDEX

Una misura del sentiment italiano sull'economia basata sui dati di Twitter

L'Istat aggiorna il Social Mood on Economy Index (SMEI), un indice sperimentale reso disponibile per la prima volta a ottobre 2018. L'indice fornisce misure giornaliere del sentiment italiano sull'economia, derivate da campioni di tweet pubblici in lingua italiana, catturati in streaming.

La procedura di produzione dell'indice seleziona ed elabora unicamente tweet il cui testo contenga almeno una parola appartenente ad un insieme filtro predisposto da esperti di dominio. La procedura analizza, con tecniche di sentiment analysis, una media di circa 26.000 tweet al giorno.

Nella sezione download è riportata la serie storica giornaliera dell'indice dal 10 febbraio 2016.

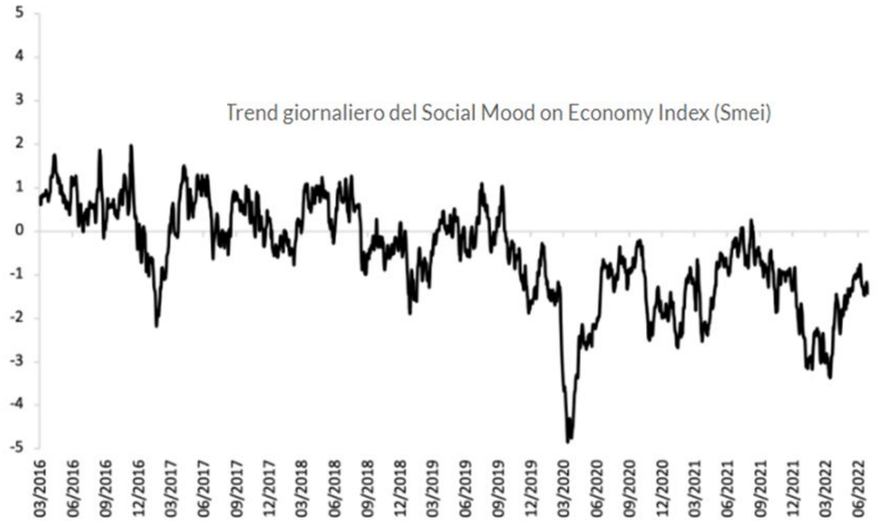
Visto il rapido aggiornamento delle tecniche di sentiment analysis e la natura sperimentale dell'indice, la procedura di calcolo originale è stata parzialmente revisionata e, con il rilascio del quarto trimestre del 2021, la serie storica è stata interamente ricalcolata.
Le principali innovazioni introdotte e le differenze con la precedente versione sono descritte in un paragrafo dedicato della nota metodologica disponibile nella sezione download.

PERIODO DI RIFERIMENTO: **ANNI 2016-2022**
DATA DI PUBBLICAZIONE: **08 FEBBRAIO 2023**

DOWNLOAD
SERIE STORICA GIORNALIERA (xlsx 137 kb)
NOTA METODOLOGICA (pdf 874 kb)
SMEI GUERRA IN UCRAINA (pdf 446 kb)

STATISTICHE SPERIMENTALI

Estrazione dati: https://www.istat.it/it/statistiche-sperimentali/sperimentazioni-su-big-data



Fonte: elaborazioni proprie su dati Istat

EUROSTAT https://ec.europa.eu/eurostat/home

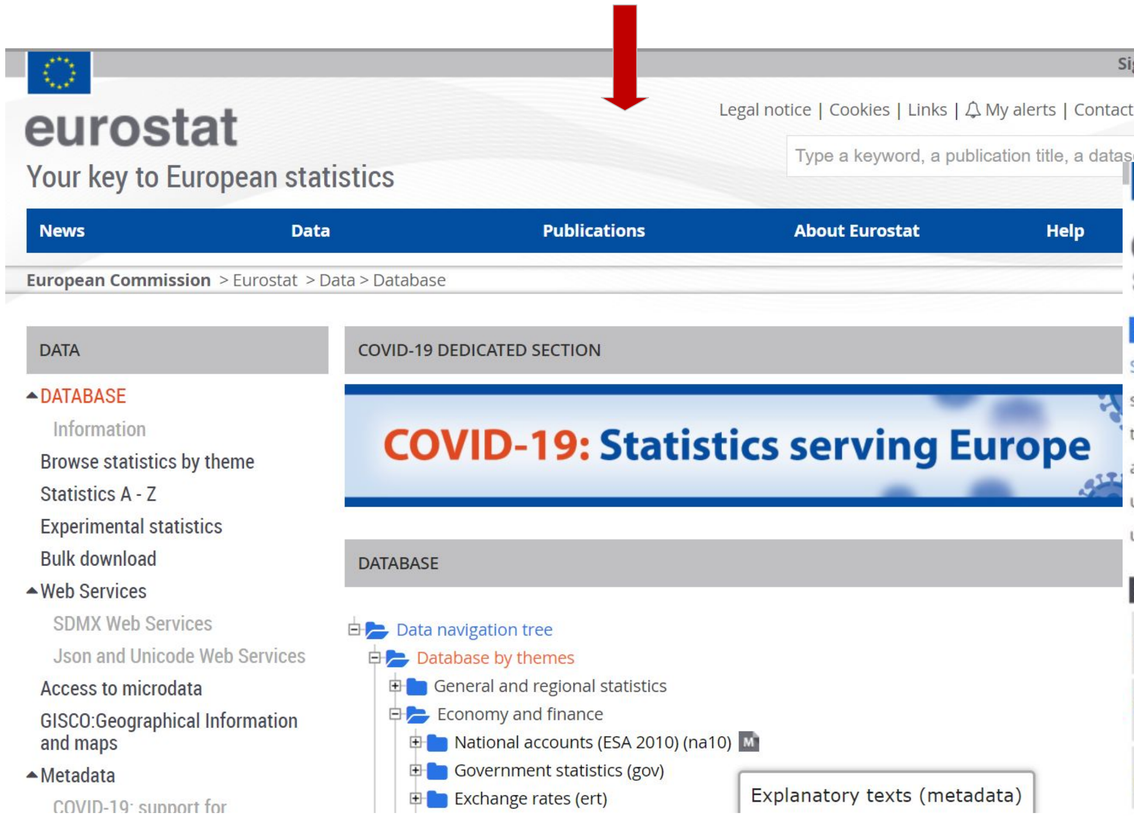
L'Ufficio statistico dell'Unione europea è una direzione generale della Commissione europea che raccoglie ed elabora dati provenienti dagli Stati membri dell'Unione europea a fini statistici, promuovendo il processo di armonizzazione della metodologia statistica tra gli Stati stessi.

La sua missione è quella di fornire all'Unione europea un servizio informativo statistico di elevata qualità, con dati comparabili tra Paesi e regioni Membri (lingue diverse, forme differenti di organizzazione amministrativa, nomenclature specifiche, ecc) in stretta cooperazione con gli INS.

Tra le sue attività principali rientra la definizione di dati macroeconomici che sostengono le decisioni della Banca centrale europea relative alla definizione delle politiche monetarie per l'euro e dati/classificazioni regionali come ad esempio la nomenclatura delle unità territoriali statistiche (o NUTS) che ha sostenuto la definizione delle politiche regionali europee e dei fondi strutturali.

IL SITO DI EUROSTAT FONTE DI STATISTICHE EUROPEE ARMONIZZATE

https://ec.europa.eu/eurostat/data/database fonte da cui estrarre tabelle standard con grafici e mappe, e creare tavole multidimensionali in diversi formati con il software Data Explorer.



IL SITO DI EUROSTAT DELLE STATISTICHE SPERIMENTALI

https://ec.europa.eu/eurostat/web/experimental-statistics/overview/ess fonte da cui estrarre dati e metodi per rispondere meglio alle nuove esigenze in modo tempestivo.

Ad esempio statistiche sulle visite ai siti patrimonio mondiale dell'UNESCO mediante l'uso di Wikipedia come nuova fonte per misurare non solo la popolarità dei siti, ma anche il «consumo culturale».

Poiché queste statistiche non hanno raggiunto la piena maturità in termini di armonizzazione, copertura o metodologia, esse sono sempre contrassegnate da un logo chiaramente visibile e corredate di note metodologiche dettagliate.

LA PRODUZIONE DELLE STATISTICHE: DALLA PROGETTAZIONE ALLA DIFFUSIONE

I DATI PRIMARI – RILEVAZIONI AD HOC

Definire lo scopo della ricerca, gli obiettivi conoscitivi che si vogliono raggiungere e le unità statistiche oggetto di indagine calibrandoli in modo da garantire la qualità dei dati rilevati.

PROGETTAZIONE (1)

- WHAT= Fissare gli obiettivi e individuare il fenomeno collettivo di interesse dell'indagine (nota che più ampio l'arco degli argomenti trattati, maggiori le complessità da affrontare sul piano concettuale, statistico ed operativo);
- WHO= Scegliere/definire la popolazione (campo di osservazione) le unità statistiche e i caratteri o variabili statistiche da indagare (es. qualitative ordinali o sconnesse, quantitative discrete o continue);
- HOW= Definire se interessa descrivere un fenomeno nella sua componente statica o dinamica e se si vuole confrontare i risultati con informazioni relative ad altre realtà territoriali;
- WHEN - WHERE= Delimitare il periodo di riferimento e l'area geografica a cui ci si riferisce (limiti spazio-temporali).

PROGETTAZIONE (2)

I DATI PRIMARI – RILEVAZIONI AD HOC

Come si vuole indagare? Indagine totale o campionaria? Con quale tecniche e strumenti di rilevazione?

- Definizione del disegno di indagine ossia specificazione del tipo di indagine per la raccolta delle informazioni e la produzione delle relative statistiche
- Individuazione della popolazione e della lista delle unità statistica
  - rilevazione totale o esaustiva (censimento) si ha la conoscenza esatta del fenomeno studiato: viene rilevata tutta la popolazione di interesse.
  - rilevazione parziale (campionaria) si perviene a una stima (più o meno precisa) del fenomeno studiato: viene osservata una parte o campione della popolazione di interesse.

    Le fonti dell'errore sono innumerevoli e non tutte sono prevedibili ed eliminabili

    - Definizione del piano di campionamento (lista, schema, errori)

TIPI DI RILEVAZIONE A CONFRONTO

INDAGINE TOTALE

VANTAGGI

- misura reale (non affetta da errore campionario) della popolazione
- informazioni rilevate base di riferimento per studi successivi
- elevato dettaglio di analisi

SVANTAGGI

- tempi lunghi per rilascio delle informazioni
- costi elevati sia in termini di risorse che economici
- errore non campionario

INDAGINE CAMPIONARIA

VANTAGGI

- riduzione dei costi
- maggiore rapidità
- maggiore accuratezza (<dimensioni)
- assenza di alternative (popolazioni illimitate)

SVANTAGGI

- errore campionario
- discriminazioni all'interno della popolazione
- eccessiva dimensione del campione in caso di eventi rari

PROGETTAZIONE (3)

- Promuovere la partecipazione alla rilevazione delle unità di rilevazione (famiglia, impresa,…)
- Scelta della tecnica di indagine (contatto) più idonea a raccogliere le informazioni
- Formulazione del questionario e pre - test
  - Concettualizzazione (relazione e gerarchie), individuazione aree e sotto-aree, contenuti, formulazione e sequenza domande
- Rilevazione su campo
  - Formazione dei rilevatori
  - Predisposizione di meccanismi di controllo

Sistema di controlli dell'errore affiancati alla fase: invio di lettera preavviso ai rispondenti; istituzione di un numero verde per chiarimenti; selezione e formazione rilevatori; analisi completezza e ridondanza delle liste utilizzate; monitoraggio mancate risposte, indagini ad hoc su non rispondenti, ect.

Fasi	Operazioni	Fonti errore	Tipo errore
Progettazione	scelta unità, variabili	modello concettuale	rilevanza teorica
	redazione questionario	struttura del vocabolario, quesiti, codifica	errori di misura
	diffusione dei dati		rilevanza effet-tiva trasparenza

TIPI DI TECNICHE DI RILEVAZIONE A CONFRONTO

INTERVISTATORE TECNOLOGIA	CON INTERVISTATORE	SENZA INTERVISTATORE
CON COMPUTER	CATI, CAPI	CASI, CAWI
SENZA COMPUTER	PAPI	POSTALI

La scelta è connessa a:

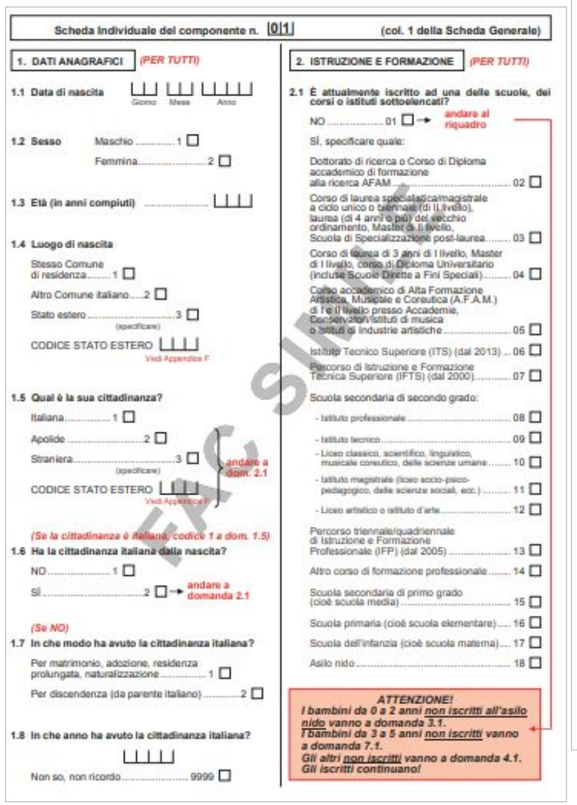
- il fenomeno indagato
- gli archivi di base
- la strategia di campionamento
- i costi e i tempi attesi
- l'organizzazione del personale sul campo

Spesso sono usati due o più modi di raccolta dei dati per: risparmiare, aumentare il tasso di risposta, ridurre gli errori di misura, ect.

ESEMPIO: Rilevazione Istat: «Aspetti della vita quotidiana»

dati anagrafici

consumo di tabacco



Scheda Individuale del componente n. 01 (col. 1 della Scheda Generale)

1. DATI ANAGRAFICI (PER TUTTI)

1.1 Data di nascita — Giorno Mese Anno

1.2 Sesso — Maschio 1 — Femmina 2

1.3 Età (in anni compiuti)

1.4 Luogo di nascita
Stesso Comune di residenza 1
Altro Comune italiano 2
Stato estero 3 (specificare)
CODICE STATO ESTERO (Vedi Appendice F)

1.5 Qual è la sua cittadinanza?
Italiana 1
Apolide 2
Straniera 3 (specificare) → andare a dom. 2.1
CODICE STATO ESTERO (Vedi Appendice F)

(Se la cittadinanza è italiana, codice 1 a dom. 1.5)

1.6 Ha la cittadinanza italiana dalla nascita?
NO 1
SÌ 2 → andare a domanda 2.1

(Se NO)

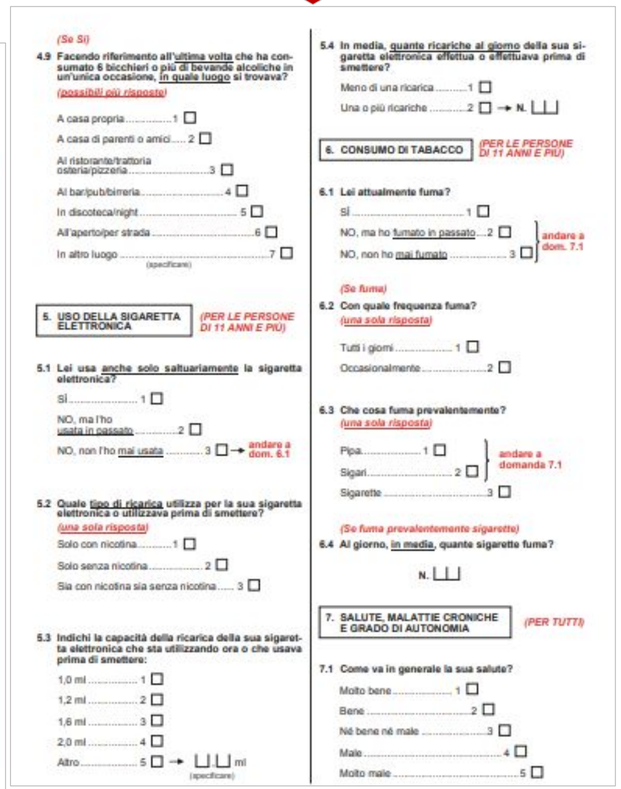
1.7 In che modo ha avuto la cittadinanza italiana?
Per matrimonio, adozione, residenza prolungata, naturalizzazione 1
Per discendenza (da parente italiano) 2

1.8 In che anno ha avuto la cittadinanza italiana?
Non so, non ricordo 9999

2. ISTRUZIONE E FORMAZIONE (PER TUTTI)

2.1 È attualmente iscritto ad una delle scuole, dei corsi o istituti sottoelencati?
NO 01 → andare al riquadro
SÌ, specificare quale:
Dottorato di ricerca o Corso di Diploma accademico di formazione alla ricerca AFAM 02
Corso di laurea specialistica/magistrale a ciclo unico o biennale (di II livello), laurea (di 4 anni o più del vecchio ordinamento, Master di II livello, Scuola di Specializzazione post-laurea 03
Corso di laurea di 3 anni di I livello, Master di I livello, corso di Diploma Universitario (incluse Scuole Dirette a Fini Speciali) 04
Corso accademico di Alta Formazione Artistica, Musicale e Coreutica (A.F.A.M.) di I e II livello presso Accademie, Conservatori/Istituti di musica o Istituti di industrie artistiche 05
Istituto Tecnico Superiore (ITS) (dal 2013) 06
Percorso di Istruzione e Formazione Tecnica Superiore (IFTS) (dal 2000) 07
Scuola secondaria di secondo grado:
- Istituto professionale 08
- Istituto tecnico 09
- Liceo classico, scientifico, linguistico, musicale coreutico, delle scienze umane 10
- Istituto magistrale (liceo socio-psico-pedagogico, delle scienze sociali, ecc.) 11
- Liceo artistico o istituto d'arte 12
Percorso triennale/quadriennale di Istruzione e Formazione Professionale (IFP) (dal 2005) 13
Altro corso di formazione professionale 14
Scuola secondaria di primo grado (cioè scuola media) 15
Scuola primaria (cioè scuola elementare) 16
Scuola dell'infanzia (cioè scuola materna) 17
Asilo nido 18

ATTENZIONE!
I bambini da 0 a 2 anni non iscritti all'asilo nido vanno a domanda 3.1.
I bambini da 3 a 5 anni non iscritti vanno a domanda 7.1.
Gli altri non iscritti vanno a domanda 4.1.
Gli iscritti continuano!



(Se Sì)

4.9 Facendo riferimento all'ultima volta che ha consumato 6 bicchieri o più di bevande alcoliche in un'unica occasione, in quale luogo si trovava? (possibili più risposte)
A casa propria 1
A casa di parenti o amici 2
Al ristorante/trattoria osteria/pizzeria 3
Al bar/pub/birreria 4
In discoteca/night 5
All'aperto/per strada 6
In altro luogo (specificare) 7

5. USO DELLA SIGARETTA ELETTRONICA (PER LE PERSONE DI 11 ANNI E PIÙ)

5.1 Lei usa anche solo saltuariamente la sigaretta elettronica?
SÌ 1
NO, ma l'ho usata in passato 2
NO, non l'ho mai usata 3 → andare a dom. 6.1

5.2 Quale tipo di ricarica utilizza per la sua sigaretta elettronica o utilizzava prima di smettere? (una sola risposta)
Solo con nicotina 1
Solo senza nicotina 2
Sia con nicotina sia senza nicotina 3

5.3 Indichi la capacità della ricarica della sua sigaretta elettronica che sta utilizzando ora o che usava prima di smettere:
1,0 ml 1
1,2 ml 2
1,6 ml 3
2,0 ml 4
Altro 5 (specificare) ml

5.4 In media, quante ricariche al giorno della sua sigaretta elettronica effettua o effettuava prima di smettere?
Meno di una ricarica 1
Una o più ricariche 2 → N.

6. CONSUMO DI TABACCO (PER LE PERSONE DI 11 ANNI E PIÙ)

6.1 Lei attualmente fuma?
SÌ 1
NO, ma ho fumato in passato 2
NO, non ho mai fumato 3 → andare a dom. 7.1

(Se fuma)

6.2 Con quale frequenza fuma? (una sola risposta)
Tutti i giorni 1
Occasionalmente 2

6.3 Che cosa fuma prevalentemente? (una sola risposta)
Pipa 1
Sigari 2 → andare a domanda 7.1
Sigarette 3

(Se fuma prevalentemente sigarette)

6.4 Al giorno, in media, quante sigarette fuma?
N.

7. SALUTE, MALATTIE CRONICHE E GRADO DI AUTONOMIA (PER TUTTI)

7.1 Come va in generale la sua salute?
Molto bene 1
Bene 2
Né bene né male 3
Male 4
Molto male 5

- popolazione:
  i giovani italiani tra i 14 e i 19 anni
- campione:
  i giovani tra i 14 e i 19 anni appartenenti alle circa 25.000 famiglie campione
- tecnica di rilevazione:
  tecnica mista, che si avvale di un questionario online auto compilato (CAWI) o di una intervista diretta con questionario sia elettronico che cartaceo, somministrato da un intervistatore (CAPI-PAPI)
- questionario:
  articolato in più parti: una scheda generale con informazioni socio-demografiche (età, sesso, stato civile, titolo di studio, ecc.) per i componenti famiglia; un questionario familiare con informazioni generali sulla famiglia; un questionario individuale

RILEVAZIONE E REGISTRAZIONE DEI DATI

- Raccogliere, senza influenzare il rispondente, le informazioni utilizzando le diverse modalità
- Registrare i dati riportando le risposte dei questionari (strumento di osservazione) su supporto informatico:
  - in alcuni casi lettura ottica dei questionari
  - ricodifica quesiti aperti in codici standard (es. ATECO)
  - non è necessaria quando la rilevazione è in CATI, CAPI, CAWI

Sistema di controlli dell'errore affiancati alla fase: buona impressione per contatti futuri, uso programmi per registrazione controllata dei dati; applicazione tecniche di identificazione automatica di incoerenze nei dati (es. un professionista con la sola licenza elementare); correzioni dati con valori accettabili mediante re intervista; calcolo indicatore di qualità (tassi di risposta), ect.

Fasi	Operazioni	Fonti errore	Tipo errore
Rilevazione sul campo	formazione elenchi di unità	liste supervisori rilevatori	errori di misura (selez. da mancata copertura)
	raccolta dati	supervisori rilevatori rispondenti	errori di misura (mancate risposte, incongruenze)
	registrazione	operatori	errori di misura

ELABORAZIONE DEI RISULTATI

Nella fase di elaborazione si applicano strumenti propri dell'analisi statistica (metodi inferenziali-induttivi) al fine di ottenere una sintesi e una descrizione dei dati sotto forma di:

- tabelle
- grafici
- calcolo di indici e rapporti sintetici

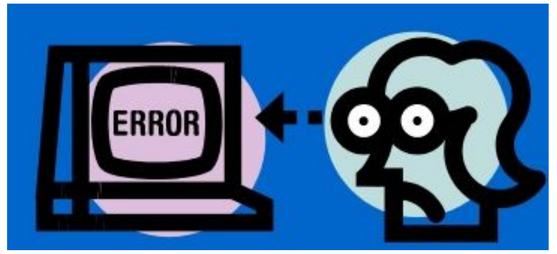
La scelta del metodo da usare dipende dal tipo di fenomeno osservato (quantitativo o qualitativo).

Sistema di controlli dell'errore affiancati alla fase: analisi descrittive e esplorative per individuare incongruenze (valori anomali ed errori;) ricerca di relazioni; analisi confermative o verifica di ipotesi; indicatori di qualità (varianza), ect.

Fase	Operazioni	Fonti errore	Tipo errore
Elaborazione dati	elaborazione	programmi base statistica	errori di calcolo rilevanza effettiva
	diffusione		tempestività

REVISIONE

La revisione consiste nel:



- valutare la congruità delle informazioni raccolte con le finalità dell'indagine;
- individuare le fonti di errore più rilevanti analizzando i dati e correggendo gli errori;
- predisporre modifiche al processo di produzione in modo da ridurre gli effetti degli errori in successive occasioni di indagine;
- validare i dati corretti (se la qualità dei dati è buona (es. tempestività ai fini della diffusione) dell'informazione agli utenti.

Sistema di controlli dell'errore affiancati alla fase: procedure manuali, automatiche e/o interattive per controlli logico/formali (regole di compatibilità) su campo di variazione variabili, relazioni fra variabili e norme di compilazione dei questionari, ect.

Fase	Operazioni	Fonti errore	Tipo errore
Revisione	registrazione	operatori	errori di misura
	revisione	revisori programmi automatici	errori di misura (possibile selezione)

DIFFUSIONE

È la fase conclusiva della rilevazione, grazie alla quale si rendono disponibili (accessibili e comprensibili) i dati raccolti (le statistiche e l'analisi statistica).

Fa percepire all' «esterno» la specificità e l'utilità del lavoro svolto (promozione).

Aspetti nella strategia di diffusione:

- ❑ adeguata offerta di informazioni
- ❑ massima facilità di accesso ai dati e ai metadati
- ❑ tempo ridotto tra raccolta, elaborazione e diffusione
- ❑ forme efficienti di diffusione e analisi del grado di utilizzo delle statistiche
  - ▪ caratteristiche degli utenti
  - ▪ canali utilizzati
- ❑ valutare se la qualità dei dati è adatta alla diffusione
- ❑ valutare la necessità di modificare il processo di produzione

Hans Rosling - 200 Countries, 200 Years, 4 Minutes - The Joy of Stats
https://www.youtube.com/watch?v=jbkSRLYSojo

ESEMPIO: Rilevazione Istat Indagine Viaggi e vacanze - Diffusione

10 febbraio 2020

statistiche report

Istat Istituto Nazionale di Statistica

VIAGGI E VACANZE IN ITALIA E ALL'ESTERO | ANNO 2019

Rallenta il turismo dei residenti: -8,8% i viaggi, -5,0% i pernottamenti

I viaggi dei residenti in Italia nel 2019 sono 71 milioni e 883 mila (411 milioni e 155 mila pernottamenti) con una flessione sull'anno precedente che interrompe la ripresa iniziata nel 2016. In calo sia le vacanze (-8,4%) sia i viaggi di lavoro (-12%).

In estate, il 37,8% della popolazione fa almeno una vacanza.

Il 76,2% dei viaggi ha come destinazione una località italiana (-12,8% sul 2018), il 23,8% è diretto all'estero.

89%
La quota di viaggi per motivi di vacanza.
Rappresentano il 93,4% dei pernottamenti.

5,7 notti
Durata media dei viaggi

63%
Percentuale dei viaggi estivi occasione di partecipazione ad attività culturali

Nota metodologica

Riferimenti normativi

La rilevazione di informazioni riguardanti il turismo è prevista dal Programma statistico nazionale, che raccoglie l'insieme delle rilevazioni statistiche necessarie al Paese. Inoltre, essa viene svolta in conformità alle definizioni concettuali e metodologiche espresse dal Regolamento per le Statistiche del Turismo 692/2011, che ha sostituito la precedente Direttiva 95/57/CE.

Obiettivi conoscitivi e quadro di riferimento

"Viaggi e vacanze" è un focus inserito nell'intervista finale dell'indagine sulle Spese delle famiglie a partire dal 2014, e consente di rilevare informazioni sui movimenti turistici dei residenti in Italia. Tali informazioni erano rilevate precedentemente dall'indagine trimestrale Viaggi, vacanze e vita quotidiana, condotta dal 1997 al 2013.

Il focus ha la finalità di ottenere informazioni sui movimenti turistici della popolazione (domanda turistica). Le stime prodotte riguardano il numero di turisti, viaggi, pernottamenti in viaggio e escursioni sul territorio nazionale o all'estero.

Il quadro normativo della rilevazione ha come riferimento il Regolamento per le Statistiche del Turismo 692/2011, nell'ambito del *framework* concettuale e metodologico delle International Recommendations for Tourism Statistics 2008 (IRTS 2008). Il turismo è definito come l'insieme delle attività e dei servizi riguardanti le persone che si spostano al di fuori del loro "ambiente abituale" per vacanza o per motivi di lavoro. Rientrano pertanto nei flussi turistici tutti gli spostamenti non abituali, con pernottamento (viaggi) o senza (escursioni). L'individuazione dell'ambiente abituale di una persona permette di distinguere correttamente il fenomeno turistico dalla mobilità, che non rientra nel campo di osservazione della domanda turistica.

Ad esempio, i viaggi e le escursioni abituali, quelli cioè effettuati settimanalmente nella stessa località, diversa dal luogo in cui si vive, sono comunque assimilabili all'ambiente abituale e non rientrano nei flussi turistici; si presuppone, infatti, che tali spostamenti siano riconducibili alla vita quotidiana e alle abitudini dell'individuo. Sono altresì esclusi dalla definizione di "turista" le persone che si spostano giornalmente o settimanalmente per lavoro, per studio o per motivi personali, quando cioè lo spostamento rientra nell'ambito di attività di *routine*.

I viaggi turistici (non abituali) sono classificati, secondo gli standard internazionali, distinguendo i viaggi per motivi di lavoro da quelli per motivi di vacanza e le vacanze 'brevi' (da 1 a 3 notti) da quelle 'lunghe' (più di 3 notti). Tra le vacanze rientrano i viaggi per svago, piacere, relax, per visitare parei religiosi.

Maggiori informazioni sono disponibili sul sito dell'Istat all'indirizzo: h

Fonti di dati

La fonte informativa è rappresentata dall'indagine sulle Spese delle f e vacanze". L'indagine è campionaria e continua (è svolta tutti i mesi d su base trimestrale, è a due stadi di cui il primo è stratificato: le u secondo stadio sono le famiglie.

Classificazioni

Nella rilevazione sono utilizzate le classificazioni territoriali Istat di Comuni, Province e Regioni, le classificazioni Istat degli Stati Esteri e *Nomenclature of Territorial Units for Statistics* – NUTS, la classificazione dell'attività economica Ateco 2007 (Nace Rev.2), la classificazione ISCED dei titoli di studio. Per alcune caratteristiche del viaggio, tra cui tipo di alloggio, motivo e tipo di destinazione, si utilizzano le classificazioni dei metadati di Eurostat, consultabili all'indirizzo: eurostat's metadata server-ramon.

Diffusione

Tra febbraio e marzo di ogni anno la Statistica Report "Viaggi e vacanze in Italia e all'estero" diffonde le stime provvisorie riferite all'anno precedente.

Le stime definitive sono consultabili, a partire dal mese di luglio, nel datawarehouse dell'Istituto I.Stat, sotto il tema:

Definizione della popolazione, della tipologia di studio, della tecnica di indagine, …

ESEMPIO: Diffusione

HOME > INFORMAZIONI E SERVIZI > PER I GIORNALISTI > APPUNTAMENTI > CALENDARIO DIFFUSIONI ED EVENTI

[ENGLISH

CALENDARIO DELLE DIFFUSIONI E DEGLI EVENTI

DATI ANALISI E PRODOTTI

METODI E STRUMENTI

INFORMAZIONI E SERVIZI

PER GLI UTENTI
- Sportelli sul territorio
- European data support
- Biblioteca
- Archivio storico

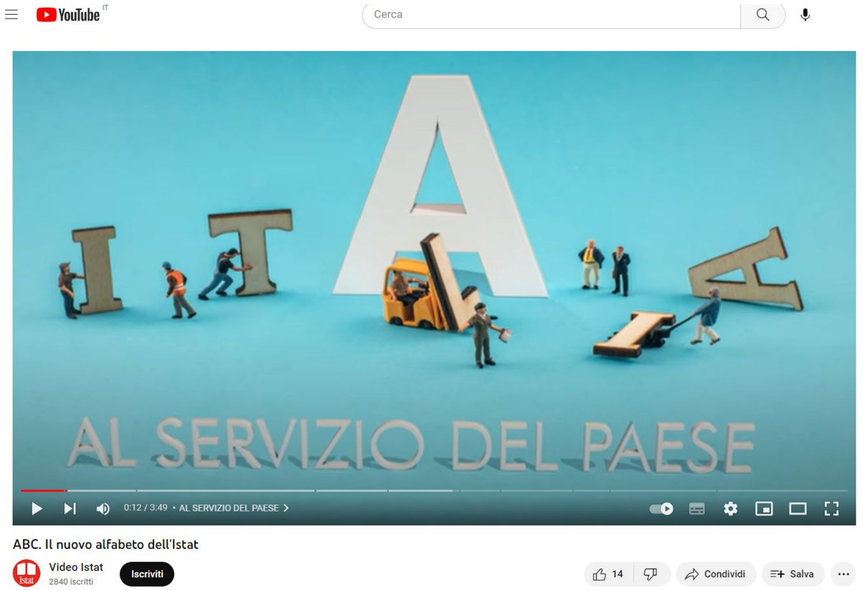
PER I GIORNALISTI
- Appuntamenti
- - Calendario diffusioni ed

Il calendario delle diffusioni e degli eventi, già completo dei comunicati stampa i cui **rilasci sono fissati per l'intero anno**, viene aggiornato con cadenza settimanale con le date di tutte le diffusioni alla stampa, della pubblicazione dei prodotti editoriali, degli eventi e degli aggiornamenti delle banche dati.

Calendario annuale dei comunicati stampa 2022 (pdf)
Calendario annuale dei comunicati stampa 2023 (pdf)

17 – 23 APRILE 2023

Lunedì 17 Aprile 2023



https://www.youtube.com/watch?v=w0-BhhYqN2A

RIASSUMENDO (1)

Disponibilità dati ?

SI

DATI secondari - fonti

FASI DELL'INDAGINE

- ✖ Progettazione
  - ✓ Definizione degli obiettivi e della popolazione di riferimento
- ✖ Registrazione dei dati
  - ✓ Individuazione della/e fonti statistiche e dei dati adeguati all'obiettivo definito
- ✖ Analisi dei dati: elaborazione e/o interpretazione
  - ✓ Trasformazione dei dati: aggregazioni territoriali o per altre categorie, calcolo di indici, calcolo di rapporti statistici ecc..; analisi dei dati; rappresentazioni tabellari e/o grafiche….
- ✖ Validazione e diffusione
  - ✓ Verifiche di coerenza esterna con altri dati, pubblicazione dei risultati

NO

DATI PRIMARI – indagine ad hoc

FASI DELL'INDAGINE

- ✖ Progettazione
  - ✓ Definizione degli obiettivi
    - ➢ Qual è l'esigenza informativa che si vuole soddisfare?
  - ✓ Definizione della popolazione di riferimento o collettivo statistico oggetto dell'indagine
    - ➢ Su cosa si vuole indagare?
    - ➢ Quale è il riferimento temporale?
  - ✓ Definizione del disegno di indagine
    - ➢ Come si vuole indagare? Indagine totale o parziale?
    - ➢ Con quali tecniche e strumenti di rilevazione?
- ✖ Rilevazione dei dati
  - ✓ Liste della popolazione, individuazione del campione, formazione rilevatori, raccolta delle informazioni, registrazione dei dati
- ✖ Analisi dei dati: elaborazione e interpretazione
  - ✓ Verifica della congruenza dei dati, aggregazioni territoriali o per categorie, calcolo di indici, analisi dei dati, rappresentazione tabellare e/o grafica
- ✖ Validazione e diffusione
  - ✓ Verifiche di coerenza esterna con altri dati, pubblicazione delle statistiche

RIASSUMENDO (2)

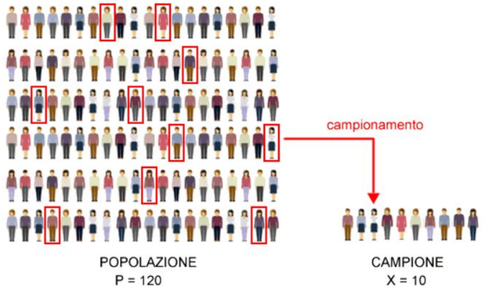
MOLTEPLICI FONTI DI ERRORE

Valore osservato (variabile) = Valore vero (concetto) + Errore

Errore non campionario

* errori di lista delle unità da osservare
* errori nel questionario
* errori dovuti alla tecnica di indagine
* errori dei rilevatori
* errori dovuti ai rispondenti
* mancate risposte (totali e parziali)

Errore campionario



Qual è l'aspetto caratterizzante dell'ufficialità?

Il rispetto
- di regole metodologiche condivise,
- di principi etico-professionali,
- dell'autonomia scientifica.

L'obiettivo finale di generare VALORE PUBBLICO e impatti migliorativi per la collettività non può prescindere da un'attenzione costante verso la QUALITÀ.

CONCLUSIONI (1)

LE QUESTIONI APERTE E IL VALORE DEI DATI:

- Un mondo di dati e di produttori di STATISTICHE NON TUTTE UGUALI per esigenze conoscitive sempre più diverse.

- L'informare come «dare senso alla realtà» richiede quantomeno informazioni CORRETTE

  - AUTOREVOLEZZA DELLA FONTE (credenziali e reputazione di imparzialità, indipendenza scientifica, impianto organizzativo pubblico)
  - QUALITÀ DEL DATO PRODOTTO (affidabilità, pertinenza, trasparenza, riservatezza, tempestività etc.) tanto maggiore è la qualità tanto minore è l'errore connesso in ogni fase

Criteri di qualità	Descrizione
Pertinenza	Il grado in cui le statistiche rispondono alle esigenze attuali e potenziali degli utilizzatori.
Accuratezza	Il grado di corrispondenza fra le stime e i valori veri ignoti.
Comparabilità	La misurazione dell'impatto delle differenze tra i concetti statistici applicati, gli strumenti e le procedure di misurazione quando le statistiche si comparano per aree geografiche, ambiti settoriali o periodi di tempo.
Coerenza	L'idoneità dei dati a essere attendibilmente combinati in modi diversi e a vari scopi.
Tempestività	L'intervallo di tempo intercorrente fra il momento della diffusione dell'informazione e l'evento o il fenomeno da essa descritto.
Puntualità	L'intervallo di tempo intercorrente fra la data di diffusione dei dati e il termine previsto per la loro diffusione.
Accessibilità e chiarezza	Le modalità e le condizioni alle quali gli utilizzatori possono acquisire, utilizzare e interpretare i dati.

Fonte: Corte dei conti europea, sulla base dell'articolo 12 del regolamento (CE) n. 223/2009.

CONCLUSIONI (2)

- DATA REVOLUTION: Democratizzazione della statistica e digitalizzazione dell'informazione

  - I dati - asset strategico per conoscere la realtà
  - Nuove fonti, nuovi metodi e tecniche e nuovi produttori di statistiche che "danno i numeri"
  - Crescente richiesta di dati (accessibilità, tempestività, granularità, ect.)

Data Science – Baba Brinkman Music Video
https://www.youtube.com/watch?v=uHGlCi9jOWY

Grazie per l’attenzione!

Rita lima
lima@istat.it
www.istat.it