

SCHEDA DI CASO

RISULTATO ATTESO 3 - Provvedere all'analisi ed all'elaborazione dei dati, gestendo banche dati e curando la tracciabilità degli stessi in archivi



1 - ESTRAZIONE E CLASSIFICAZIONE

Grado di complessità 3

1.3 POST-CODIFICA DEI DATI

Nel caso di interviste con domande aperte, dopo aver preso visione dello spoglio preliminare di un campione delle risposte, redigere un codice (es. una lista dei più frequenti tipi di risposte) e assegnare la risposta data da ciascun intervistato alla voce di codice che ne interpreta meglio il significato

Grado di complessità 2

1.2 CONTROLLO E PULIZIA DEI DATI RACCOLTI (CLEANING)

Controllare la validità dei dati dal punto di vista formale (es. la coerenza logica tra informazioni diverse), segnalando e gestendo eventuali errori o incoerenze

1.2 CLASSIFICAZIONE DEI DATI

Inserire i dati in archivi e banche dati secondo regole univoche (es. suddivisione per regioni, per intensità della risposta ecc.)

Grado di complessità 1

1.1 IMMISSIONE DEI DATI (DATA ENTRY)

Immettere le stringhe di informazioni, eventualmente corrispondenti alle risposte degli intervistati, nel database di riferimento per l'indagine, in modo manuale mediante digitazione su tastiera o prevedendo un inserimento automatico tramite esportazione da altri database (es. tramite wordprocessor, fogli di calcolo, ecc.)



2 - ANALISI DATI

Grado di complessità 3

2.3 UTILIZZO DI TECNICHE DI STATISTICA MULTIVARIATA E DI ALTRE TECNCHE AVANZATE

Applicare tecniche di analisi multivariata dei dati (analisi fattoriale, cluster analysis, Analisi delle



Corrispondenze, ecc.)

Grado di complessità 2

2.2 UTILIZZO DI TECNICHE DI STATISTICA BIVARIATA

Applicare tecniche di analisi bivariata per l'elaborazione dei dati (tavole di contingenza, analisi della varianza, regressione lineare, ecc.)

Grado di complessità 1

2.1 UTILIZZO DI TECNICHE DI STATISTICA MONOVARIATA

Applicare tecniche di analisi descrittiva, mediante strumenti standard di tipo grafico (dotplot, tabella, istogramma di frequenza, frequenza comulata, boxplot, probability plot, ecc.) o indici di sintesi (indici di posizione o tendenza centrale, indici di variabilità o dispersione, indice di asimmetria, ecc.)



3 - GESTIONE BANCHE DATI E ARCHIVI

Grado di complessità 3

3.3 AGGIORNAMENTO DEL DATABASE

Organizzare la modalità per mantenere aggiornata la banca dati e le informazioni in essa contenute: nel caso di banche dati, reperire le informazioni recenti a esse relative (es. interrogazione database Istat), nel caso di utilizzo di interviste, armonizzare i dati del database con i risultati delle survey più recenti

Grado di complessità 2

3.2 COSTRUZIONE DEL DATABASE E DEGLI ARCHIVI

Strutturare il set informativo utile alla lettura univoca del database da parte di altri soggetti attraverso la redazione di apposite guide (es. manuali o stringhe esplicative, foglio "leggimi" di Excel, ecc.)

Grado di complessità 1

3.1 UTILIZZO DEL DATABASE

Creare strumenti automatici per l'interrogazione della banca dati (es. query di access o tool di graficazione con opzioni di default), anche attraverso la collaborazione dell'ufficio IT aziendale per



l'implementazione del database nel sistema informatico aziendale



SCHEDA RISORSE A SUPPORTO DELLA VALUTAZIONE DEL RISULTATO ATTESO 3



RISORSE FISICHE ED INFORMATIVE TIPICHE (IN INPUT E/O PROCESS ALLE ATTIVITÀ)

- Norme in materia di protezione della privacy e protezione dati personali (GDPR)
- Software di statistica
- Database per l'archivio e storage dei dati
- Strumenti automatici per l'interrogazione delle banche dati (query access; tool di gratificazione con opzioni di default)



TECNICHE TIPICHE DI REALIZZAZIONE/CONDUZIONE DELLE ATTIVITÀ

- Tecniche di codifica delle domande aperte, trattamento dei dati qualitativi, trascrizione delle registrazioni delle interviste
- Strumenti di classificazione dati
- Metodi e tecniche di controllo qualità (data cleaning)
- Tecniche di immissioni dati (data entry)
- Metodi e tecniche di analisi statistica multivariata e avanzata (analisi fattoriale, cluster analysis, analisi delle corrispondenze)
- Metodi e tecniche di analisi statistica bivariata (tavole di contingenza, analisi della varianza, regressione lineare)
- Metodi e tecniche di analisi monovariata di tipo grafico (dotplot, tabella, istogramma di frequenza, frequenza comulata, boxplot, probability plot) e numerico (indici di posizione, indici di dispersione, indici di asimmetria)
- Metodi e tecniche di definizione di query su basi dati
- Metodi e tecniche per costruzione, aggiornamento e utilizzo di database



OUTPUT TIPICI DELLE ATTIVITÀ

- Dati decodificati
- Data Cleaning effettuata
- Inserimento dati in archivi e banche dati
- Analisi statistiche dei dati effettuate
- Informazioni per aggiornamento database reperite
- Set informativo per la costruzione del database definito
- Dati dei database e dati dei survey armonizzati
- Strumenti automatici per l'utilizzo del database definiti



INDICAZIONI A SUPPORTO DELLA SCELTA DEL METODO VALUTATIVO E DELLA

PREDISPOSIZIONE DELLE PROVE



ESTENSIONE SUGGERITA DI VARIETÀ PRESTAZIONALE

- 1. L'insieme delle tipologie di rilevazione di dati e relativi strumenti
- 2. L'insieme delle tipologie di software di elaborazione statistica
- 3. L'insieme delle procedure di estrazione e classificazione dei dati
- 4. L'insieme delle procedure per la costruzione, utilizzo e aggiornamento del database
- 5. L'insieme delle regole di controllo di qualità dell'indagine
- **6.** L'insieme delle regole di classificazione, analisi ed elaborazione statistica dei dati raccolti (monovariata, bivariata, multivariata/avanzata)
- 7. Un set di strumenti di rilevazione e di relativi dati raccolti

DISEGNO TIPO DELLA VALUTAZIONE

- 1. Prova prestazionale: per almeno una tipologia di rilevazione ed una di strumenti, sulla base di uno strumento software e del set dato, rappresentazione del processo di estrazione, classificazione, analisi ed elaborazione dei dati, nonché di gestione del relativo database
- **2.** Colloquio tecnico relativo al processo di raccolta, controllo, elaborazione statistica dei dati per una tipologia di set di dati non oggetto di prova prestazionale



FONTI

Aledo, J.A., Gamez, J.A., & Molina, D. (2013). Tackling the rank aggregation problem with evolutionary algorithms. Applied Mathematics and Computation, vol.222

Asquini, G., Definizione di una procedura di codifica delle domande aperte basata sui modelli delle indagini internazionali, ECPS Journal, ottobre 2014, pag. 463-483

Cedrola, E., Appunti sulle Ricerche di Mercato, I.S.U. Università Cattolica, Milano, 2001

Corbetta, P. (1999). Metodologia e tecniche della ricerca sociale, Il Mulino, 1999, p. 1-640

Istat (2020), Analisi delle corrispondenze multiple

Istat, L'approccio trasversale alla formazione delle reti di rilevazione, 2019, p. 1-74

Johnson R.A. e Wichern D.W (2007). Applied Multivariate Statistical Analysis. Six Edition. Pearson Education International, Cap. 1-2, 8, 12

Picarelli, A., Tecniche delle ricerche di mercato, Rocco Curto, Napoli, 1994, pag. 79

Tasciotti, C., Professione Statistico: attualità e prospettive, In Rivista Quadrimestrale per la diffusione della cultura statistica, anno III / 2014 – Numero Speciale

UNI 11621-2:2017, Attività professionali non regolamentate - Profili professionali per l'ICT - Parte 2: Profili professionali di "seconda generazione", Febbraio 2017

Zani S. E Cerioli A. (2007). Analisi dei dati e data mining per le decisioni aziendali, Milano: Giuffrè, cap. 1,2,5,6,7,8,9

SITOGRAFIA

Atlante del Lavoro

http://matematica-old.unibocconi.it/statistica/SONDAGGI.htm#rilevazione

http://qualitapa.gov.it/sitoarcheologico/relazioni-con-i-cittadini/utilizzare-gli-strumenti/banca-dati/index.html