



RACCOLTA DATI STATISTICI *a cura di Patrizia Grossi*

Premessa: Prime due fasi della Raccolta dati [“Specifica dei fabbisogni” e “Progettazione”]

1. Strumento di Raccolta: il QUESTIONARIO

1.1. Struttura

1.2. Redazione

1.3. Procedure di verifica

2. RACCOLTA DATI

2.1 Creazione della cornice e selezione del campione

2.2 Pianificazione

2.3 Conduzione

2.4 Tecniche di acquisizione



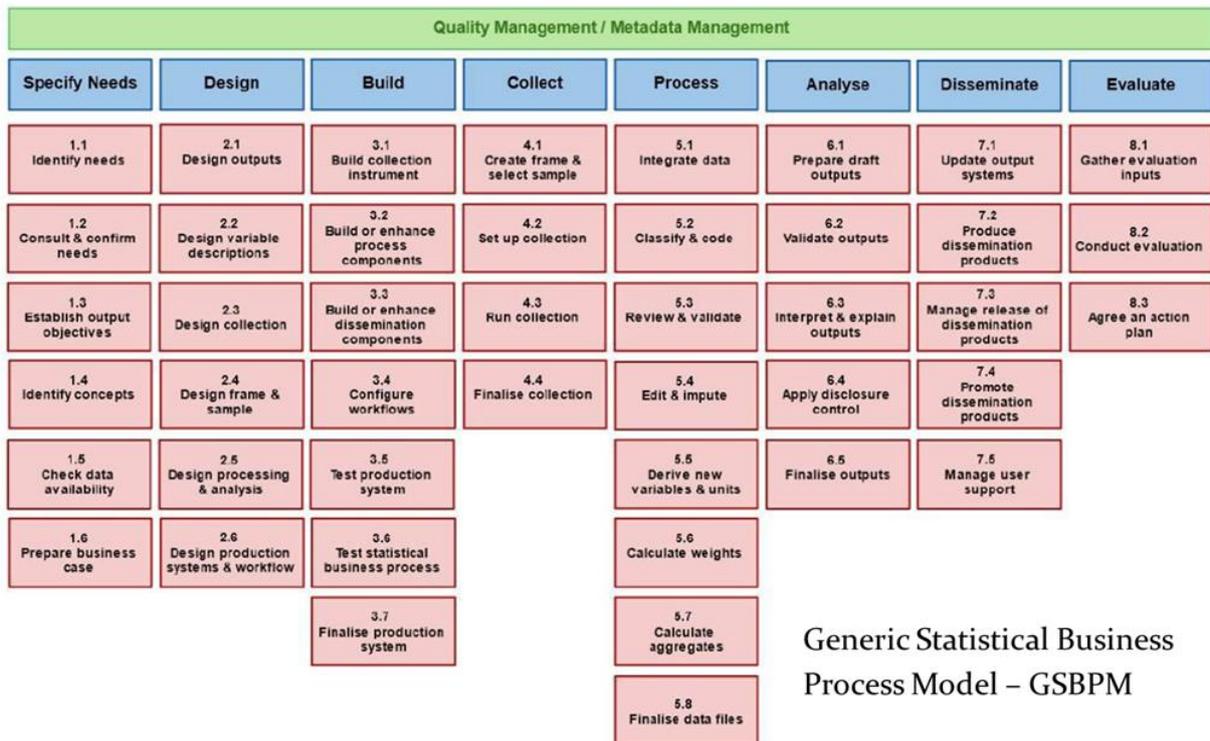
Premessa:

Il dato, in sé, non possiede valore, o meglio lo possiede solo in forma potenziale. Grazie al processo di analisi, elaborazione e presentazione che si passa dal dato all'informazione e grazie al confronto di diverse informazioni, provenienti anche da fonti diverse, si passa poi dall'informazione alla conoscenza, attraverso la quale si può rispondere a domande del tipo "come" e "perché".

Dato → Informazione → Conoscenza

Un'indagine statistica è un processo produttivo complesso che ha come obiettivo la produzione di informazione statistica

L'acquisizione delle informazioni presso le unità rispondenti deve corrispondere alla soddisfazione di un'esigenza conoscitiva, oggetto della ricerca, ed è preceduta da un insieme complesso ed articolato di attività (Fase 1 e 2 del GSBPM) che permettano di raccogliere, per poi misurare i diversi aspetti del fenomeno indagato.



Nella Fase 1 Specifica dei fabbisogni “Specify needs” vengono realizzati i sottoprocessi necessari ad individuare gli obiettivi di indagine e a tradurli in concetti che, da un lato dovranno essere comprensibili e accessibili ai rispondenti, e dall’altro, dovranno essere misurabili e quindi trasformabili nelle variabili statistiche, che saranno disegnate nella fase 2. In questa prima fase è importante verificare l’esistenza dei dati presso altre fonti (ad esempio dati amministrativi) in modo da ridurre l’insieme delle variabili da rilevare con effetti positivi sia sui costi di rilevazione che sul fastidio statistico.

Nella Fase 2 Progettazione “Design” sono descritte le attività di sviluppo e progettazione della ricerca necessaria a produrre risultati statistici,, i concetti, le metodologie, gli strumenti di raccolta e i processi operativi. Le variabili individuate nella Fase 1 permetteranno la progettazione del piano di tabulazione (passaggio utile anche per la fase di diffusione dei risultati) dal quale saranno desumibili anche le variabili derivate (che non saranno rilevate ossia raccolte in fase di acquisizione) e le classificazioni ufficiali che saranno utilizzate. La definizione della variabili statistiche da raccogliere dovrebbe essere fatto tenendo in considerazione i metadati già esistenti per favorire il processo di standardizzazione, nonché il riuso di elementi definiti in altri processi di acquisizione.

Nella Fase 3 Costruzione “Build” vengono realizzati gli strumenti per la raccolta dati. La fase di acquisizione dati avviene attraverso l’utilizzo di una o più tecniche (strumenti), assistite o non dal computer, con o senza il supporto del rilevatore, nonché attraverso strumenti, quali EDI (*Electronic Data Interchange*), XBRL

(*eXtensible Business Reporting Language*), adatti allo scambio di informazioni tra le unità di rilevazione (aziende e/o organi istituzionali produttori di dati sotto forma di fonti amministrative o registri).

Lo strumento di osservazione scelto per la raccolta dei dati può considerarsi come un filtro posto tra l'osservatore e l'oggetto, cambiando il mezzo di osservazione dallo stesso oggetto possono ottenersi dati differenti, ed è per questo che la loro scelta influenza i dati dell'indagine.

I dati possono essere raccolti attraverso:

- misurazione di entità oggettive
- dichiarazioni (questionario)

La costruzione del **QUESTIONARIO** è un momento delicato nel quale viene richiesta competenza tecnica, abilità ed esperienza. Affinché lo strumento si riveli adeguato a garantire la confrontabilità delle informazioni rilevate, si impongono 2 condizioni:

- Le domande siano rivolte a tutti nella stessa forma;
- Le domande abbiano lo stesso significato per tutti i rispondenti.

La **FUNZIONE** del questionario è quella di raccogliere informazioni sulle variabili qualitative e/o quantitative oggetto del fenomeno da indagare, pertanto sarà necessario convertire le generiche esigenze conoscitive in concetti operativi:

- Precisare bene le finalità conoscitive
- Definire aree (argomenti), ambiti e costrutti da indagare
- Elencare le variabili da "misurare"
- Redigere e formulare le domande in forma scritta
- Definire l'ordine di presentazione delle domande
- Prevedere la presenza di domande filtro

In generale il **questionario** conterrà codici identificativi, domande sulle caratteristiche strutturali degli individui, domande specifiche dell'indagine, domande utilizzate per i controlli di qualità. Al di là dei vincoli imposti dalla necessità di utilizzare forme standardizzate per le domande relative alle caratteristiche strutturali e a quelle dei controlli di qualità, esistono molti margini di libertà nella formulazione delle altre domande. La tentazione a cui è necessario resistere è di incrementare a piacere il numero di domande, ogni domanda o batterie di domande (aperte o chiuse) inserite nel questionario deve avere una sua funzione precisa nel corpo dell'indagine, ognuna costituisce un'ipotesi di ricerca.

Per la costruzione di un buon questionario si identificano 3 stadi:

1. identificazione degli obiettivi e definizione dei concetti

2. redazione
3. verifica

1° stadio: Le domande dovranno essere formulate in modo da suscitare lo stesso stimolo per tutti i rispondenti. Un rispondente che non conosce il significato di una parola non potrà fornire una risposta accurata, ma soltanto approssimativa e il ricercatore non potrà elaborare correttamente i dati.

Due sono gli spetti fondamentali da prendere in considerazione:

- L'argomento che si vuole indagare
- La popolazione di riferimento

Costruire un buon questionario viene considerato come compito specifico del ricercatore, mentre il compito di aiutare il rispondente, fornirgli chiarimenti e soprattutto motivarlo a rispondere, viene considerato come parte della sfera di attività e capacità di un buon intervistatore.

2° stadio: stabilito il tema da affrontare, occorre decidere quanta libertà lasciare all'intervistato nell'articolare la risposta.

2.1 In quest'ottica le **DOMANDE** si distinguono in

1. Domande a risposta libera (domande aperte)

Vantaggi:

- Non condiziona la risposta
- Utile per trattare argomenti delicati

Svantaggi

- Elevato livello di arbitrarietà di chi intervista
- Criticità nell'interpretazione delle risposte

2. Domande a risposta fissa predefinita (domande chiuse)

Vantaggi:

- Elimina errori di scrittura
- Facilita compito del rispondente
- Stimola l'analisi e la riflessione
- Riduce tempi di elaborazione

Svantaggi

- Esclude le alternative di risposta non previste
- Influenza risposte anche in chi non ha opinioni (le prime e le ultime modalità vengono più facilmente memorizzate)

2.2 Non esiste un **LINGUAGGIO** facile o difficile in assoluto, ma alcuni accorgimenti possono essere adottati:

- usare termini semplici
- evitare domande troppo lunghe
- evitare troppe alternative di risposta
- evitare termini tecnici
- evitare definizioni ambigue
- evitare l'utilizzo di parole con significato dispregiativo o elogiativo
- evitare l'utilizzo di locuzioni poco conosciute
- evitare domande composte da più proposizioni e troppe specificazioni
- usare termini precisi che indirizzino verso una corretta interpretazione della domanda

2.3 La **SUCCESSIONE DEI TEMI**. Spesso c'è la necessità di utilizzare molte domande per indagare uno stesso tema ed esistono alcune tecniche per gestire particolari gruppi di domande:

1. La successione **a imbuto**: si passa da domande generali a domande più particolari per dare la possibilità al rispondente di focalizzare l'attenzione sul tema proposto. Facilita la raccolta di opinioni "non meditate".
2. La successione **a imbuto rovesciato**: si fanno prima le domande specifiche e poi quelle più generali. Utile quando si desidera raccogliere opinioni "meditate" su un determinato argomento.

2.4 Le **DOMANDE FILTRO** permettono di saltare uno o più quesiti successivi se sono verificate alcune condizioni. Sono utili quando:

- è necessario indirizzare gruppi particolari di rispondenti verso blocchi di domande specificamente rivolte a loro (es: occupati o non occupati);
- si vuole evitare di porre quesiti fortemente dettagliati quando è inutile (es. non fare domande sulle vacanze svolte nell'anno a coloro che dichiarano di non aver svolto vacanze nell'anno)
- si vogliono evitare condizionamenti nella risposta (es. non è opportuno chiedere opinioni sull'ultimo libro letto nei 12 mesi a chi non ha letto nessun libro nei 12 mesi: si rischierebbe di provocare risposte "forzate" date dal rispondente al solo scopo di non fare "brutta figura")

2.5 La predisposizione delle **SCALE DI VALUTAZIONE** è un altro elemento molto importante da tenere in considerazione nelle risposte chiuse. Il rispondente dovrà biffare la risposta più congeniale tenendo presente le alternative presentate.

Si possono utilizzare:

- scale numeriche, presentando al rispondente un ventaglio di punteggi da attribuire al fenomeno (da 0 a 5 oppure da 0 a 10)

- scale verbali, presentando al rispondente un ventaglio di giudizi espressi con delle frasi da un minimo di due a un massimo di sei) riportiamo alcuni esempi

1. si 2. no	1. per niente 2. poco 3. abbastanza 4. molto 5. non so	1. in assoluto disaccordo 2. in disaccordo 3. in accordo 4. completamente in accordo 5. non so	1. insufficiente 2. sufficiente 3. discreto 4. buono 5. ottimo 6. eccellente
----------------	--	--	---

2.6 Le MODALITÀ DI SOMMINISTRAZIONE del questionario sono diverse e possono essere utilizzate a seconda degli obiettivi e dell'estensione della ricerca, della popolazione di riferimento, dei tempi e dei costi della rilevazione.

Esistono vantaggi e svantaggi nell'utilizzo di ciascuna tecnica che vanno attentamente prese in esame: vediamo alcune

1. Intervista diretta (faccia a faccia)
2. Intervista telefonica
3. Questionario auto compilato

1. L'intervista diretta viene condotta da un rilevatore che legge le domande e le opzioni di risposta nell'esatto ordine e con lo stesso linguaggio adottati nel questionario, riportandovi quindi le risposte così come sono fornite dal rispondente.

Vantaggi

- Si presta meglio ad alcuni disegni di indagine (es. censimenti)
- Maggiore possibilità di contattare e convincere il rispondente a collaborare
- Il rispondente è identificato con certezza
- Possibilità di istruire il rispondente sul significato delle domande e sul modo corretto di fornire le risposte
- Flessibilità negli strumenti utilizzabili (audiovisivi, tecniche particolari)
- Interviste di maggiore durata

2. L'intervista telefonica viene condotta al telefono da un intervistatore che legge le domande e le opzioni di risposta nell'esatto ordine e con lo stesso linguaggio adottati nel questionario, riportandovi quindi le risposte così come sono fornite dal rispondente (supporto cartaceo o CATI)

Vantaggi

- Costi minori rispetto all'intervista faccia a faccia
- Tempestività della raccolta dati
- Non è richiesta una capillare organizzazione sul territorio

- Maggiore possibilità di controllo dell'operato dei rilevatori
 - Possibilità di contatto anche per le persone che non si trovano in casa in orari "canonici"
 - Mancando la presenza fisica dell'intervistatore: bassi rischi di condizionamento e maggiore possibilità di porre quesiti delicati
- Il rispondente riceve il questionario a mezzo posta o corriere e provvede a **compilarlo** nelle parti ad esso spettanti e a rispedirlo indietro o eventualmente a riconsegnarlo ad un addetto che lo ritira a domicilio

Vantaggi

- Bassi costi di realizzazione; è richiesta un'organizzazione minore
- Bassi rischi di condizionamento
- Modalità molto adatta per porre quesiti delicati
- Disponibilità di tempo per reperire eventuale documentazione necessaria alla compilazione
- È possibile sottoporre al rispondente liste più lunghe di categorie di risposta

Svantaggi

- Tempi molto lunghi di raccolta dei dati
- Impossibilità di identificare con certezza il rispondente
- Autoselezione dei rispondenti (svantaggio molto grave)
- Minore capacità di ottenere la partecipazione all'indagine (il tema deve essere coinvolgente; non è possibile un'opera di convincimento diretto)
- È quasi impossibile aiutare i rispondenti nella comprensione delle domande e nella compilazione del questionario (sono molto importanti: la chiarezza delle domande, la grafica, la presenza di istruzioni chiare)
- Spesso, i questionari autocompilati danno luogo a risposte logicamente incongruenti, e questo è molto difficile da correggere ex-post

3° stadio: Le **PROCEDURE DI VERIFICA** costituiscono un supporto irrinunciabile per un corretto giudizio di validità del questionario, pertanto occorrerà:

- Svolgere una somministrazione pilota, per testare preventivamente il questionario (pre-test)
- Effettuare una revisione del questionario prima della somministrazione
- Valutare bene quale sarà la durata della somministrazione (non dovrebbe superare la mezz'ora)
- Decidere le modalità di somministrazione.

La fase di **RACCOLTA dei DATI ["Collect"]**, raccoglie tutte le informazioni necessarie ad analizzare il fenomeno (dati e metadati), utilizzando diverse modalità (comprese estrazioni da registri e database statistici e amministrativi), queste informazioni

vengono poi caricate nell'ambiente appropriato per l'ulteriore elaborazione e analisi.

La fase "Raccolta Dati" è suddivisa in quattro sotto processi:

1. Creazione della cornice e selezione del campione da intervistare
2. Pianificazione della raccolta dati
3. Conduzione della raccolta dati
4. Acquisizione dei dati

Il primo sottoprocesso identifica e specifica la popolazione di interesse, definisce il campione (e, ove necessario, il registro da cui è derivato) e determina i criteri e la metodologia di campionamento più appropriati (che potrebbe includere l'enumerazione completa). Le fonti comuni per un quadro di campionamento sono i registri amministrativi e statistici, i censimenti e informazioni provenienti da altre indagini campionarie.

Il secondo sottoprocesso garantisce che le persone e la tecnologia siano pronti a raccogliere dati e metadati, in tutte le modalità previste. Si svolge in un periodo di tempo prestabilito, in quanto include la pianificazione e formazione del personale addetto alla raccolta dati, la preparazione degli strumenti di raccolta (stampa di questionari, pre-riempimento con dati esistenti, caricamento di questionari e dati sui computer, ecc.) oltretutto la garanzia di trasparenza e riservatezza dei dati da raccogliere.

Il terzo sottoprocesso identifica gli strumenti utilizzati per raccogliere le informazioni, include la richiesta di microdati grezzi o aggregati prodotti all'origine, nonché tutti i metadati associati, include l'inserimento manuale dei dati o la gestione delle attività di rilevazione sul campo, a seconda dell'origine e della modalità di raccolta. Quando la raccolta raggiunge i suoi obiettivi, viene chiusa la fase e prodotto un report.

Le **TECNICHE DI ACQUISIZIONE** assistite dal computer giocano un ruolo sempre più importante nella fase di Raccolta dei dati, si distinguono in **CADI** (*Computer Assisted Data Inputing*), **CAPI** (*Computer Assisted Personal Interviewing*), **CATI** (*Computer Assisted Telephone Interviewing*), **CAWI** (*Computer Assisted Web Interviewing*). La caratteristica più significativa di queste tecniche consiste nel fatto di permettere l'inserimento già in fase di raccolta dati di tutti quei controlli tipici delle successive fasi di controllo e correzione, inibendo di fatto l'acquisizione del dato errato. Si differenzia la tecnica **CADI** che, usata nell'ambito di rilevazione tramite modelli cartacei, consiste di fatto in un'acquisizione controllata, dove i controlli sono inseriti solo al fine di ridurre quelli di registrazione o anche come supporto alla fase di revisione.

Un'altra peculiarità di queste tecniche (ad eccezione delle **CADI**) è quella di consentire la personalizzazione della formulazione dei quesiti in funzione delle caratteristiche del rispondente (nome, sesso) o di risposte fornite a precedenti quesiti del questionario stesso o di informazioni già disponibili, rendendo così l'intervista più colloquiale e facilitando la disponibilità del rispondente a collaborare. La tipologia e complessità dei controlli inseriti in fase di acquisizione è notevolmente differente tra **CATI** e **CAPI** da un lato e **CAWI** dall'altro:

Per **CATI** e **CAPI**, la presenza di un rilevatore formato sui contenuti dell'indagine permette sia di inserire una quantità di controlli decisamente più elevata di quanto sia possibile effettuare con il **CAWI**, dove il questionario è auto-compilato dal rispondente, è necessario non solo contenere il numero di controlli inseriti, ma anche di gestirli con degli avvertimenti che segnalano l'incongruenza senza l'obbligo di sanarla per procedere con l'intervista.

