

LA SCELTA DEL CAMPIONE

A cura dell'Area Pianificazione strategica e controllo direzionale di Arpa Emilia-Romagna

Premessa

In questa presentazione viene descritta la procedura scelta per determinare le persone da coinvolgere nell'indagine sul benessere organizzativo, ovvero le persone a cui somministrare il Questionario multidimensionale della salute organizzativa.

Le indagini di tipo statistico

Le indagini di tipo statistico servono a conoscere determinati comportamenti di una popolazione. Le informazioni su tali comportamenti possono essere desunte da:

1. **rilevazioni totali o censuarie**: che prendono in considerazione l'intero universo della popolazione considerata;
2. **rilevazioni campionarie**, quando l'analisi è limitata ad una parte dell'universo considerato, detto campione.

Le rilevazioni totali o censuarie

Le rilevazioni totali o censuarie coinvolgono ogni singola unità del collettivo di riferimento e vengono definite totali proprio perché osservano tutte le unità che appartengono all'universo considerato. Questa tipologia di indagine permette una comprensione effettiva del fenomeno indagato, tuttavia le rilevazioni censuarie comportano un notevole dispendio di energie e di risorse, soprattutto quando l'universo a cui ci si riferisce è molto ampio

Le rilevazioni campionarie

I risultati ottenuti da un'indagine svolta su di un campione, presentano delle differenze rispetto alle vere caratteristiche della popolazione che vogliamo stimare. L'indagine campionaria porta sempre con sé una certa quota di errore, in quanto il campione non potrà mai contenere tutte le caratteristiche dell'universo considerato (per assurdo se anche si intervistassero 9 utenti su 10 di un servizio, non si potrebbe parlare di risultati assoluti, in quanto il comportamento dell'unica persona non considerata potrebbe essere molto difforme dalla media).

Le rilevazioni campionarie

Gli errori insiti nella rilevazione campionaria, definiti errori di campionamento, possono essere contenuti entro limiti, più o meno ristretti, adottando appropriati metodi di inferenza statistica. Queste tecniche consentono di limitare gli errori insiti in una modalità di rilevazione che non analizza tutte le unità dell'universo, facilitando un processo di generalizzazione e il passaggio dal particolare al generale, ossia dal campione alla popolazione.

Le rilevazione in ARPA

In ARPA si è deciso di realizzare una **rilevazione campionaria** in quanto tale procedura, rispetto alle indagini totali, presenta alcuni vantaggi:

1. il campione consente di ottenere informazioni in tempi più brevi,
2. l'indagine campionaria ha un costo ridotto,
3. l'indagine campionaria consente un più accurato controllo del processo di acquisizione dell'informazione.

La tecnica di campionamento in ARPA

Si è deciso di procedere all'individuazione di un **campione probabilistico**, per ridurre al minimo i margini di errore, ed effettuare un **campionamento casuale semplice**

Il campione probabilistico

Il campione probabilistico è quel campione i cui risultati possono essere estesi con un certo livello di fiducia (detto anche livello di confidenza) alla popolazione. La numerosità campionaria (n), ovvero il numero delle persone da intervistare, è la risultante di una formula standard, ampiamente utilizzata nelle rilevazioni quantitative

Il campione probabilistico

Questa la formula applicata nell'ambito dell'indagine in ARPA

$$n = \frac{z_{\frac{\alpha}{2}}^2 N p (1-p)}{e^2 (N-1) + z_{\frac{\alpha}{2}}^2 p (1-p)}$$

In tale formula:

- **N** corrisponde alla numerosità della popolazione (numero dei dipendenti).
- **n** indica la numerosità del campione.
- **p** rappresenta il valore della proporzione nella popolazione. Non avendo conoscenze a priori sulle opinioni dei rispondenti, il calcolo viene svolto, con approccio conservativo, assumendo in questi casi $p = 0,5$.
- **e** esprime il valore massimo accettabile dell'errore.
- **α** la probabilità che l'errore superi il livello massimo accettabile e.
- **$Z_{\alpha/2}$** rappresenta il valore della variabile casuale normale standardizzata Z tale che la probabilità che Z assuma valori maggiori di $Z_{\alpha/2}$ è uguale a $\alpha/2$.

La determinazione della numerosità campionaria

Partendo dalla formula precedentemente descritta e applicandola al contesto lavorativo di ARPA, è stato possibile determinare la numerosità campionaria.

$$n = \frac{1,96^2 * 1.015 * 0,5(1 - 0,5)}{0,04(1.015 - 1) + 1,96^2 * 0,5(1 - 0,5)} = \underline{\underline{337}}$$

Non avendo conoscenze a priori circa il valore della proporzione nella popolazione, si è assunto che $p = 0,5$. In secondo luogo si è fissato un valore soglia di massimo errore accettabile nel calcolo del campione pari al 4% (ne deriva che $e = 0,04$). In questo modo è stata fissata la probabilità che l'errore non superi la soglia del 4%, considerando tale probabilità pari al 95% (quindi $1 - \alpha = 0,95$); infine avvalendosi di un software è stato determinato anche il valore della variabile casuale normale standardizzata, uguale a 1.96

La determinazione della numerosità campionaria

| ARPA | CAMPIONE | % |
|---|-----------|------------------------------------|
| 1.015 | 337 | 33% |
| Tutto il personale (esclusi: CdD e CoCoCo) | $e < 4\%$ | Rappresentativa di tutti i nodi |

Il campionamento casuale semplice (senza ripetizione)

- ✓ Selezione di un campione di numerosità n da una popolazione di N elementi in modo tale che ogni possibile elemento abbia *uguale* probabilità di essere estratto.
- ✓ Analogamente qualsiasi sotto insieme di n elementi ha la stessa probabilità di estrazione.
- ✓ L'assenza di ripetizione sta a significare che un'unità non può entrare a far parte dello stesso campione più di una volta.
- ✓ Riduce al minimo la discrezionalità della scelta.

Il campionamento casuale semplice (senza ripetizione)

Il campionamento casuale si realizza assegnando ad ogni soggetto della popolazione un numero ed estraendo dei numeri in modo casuale, per avere una lista di estrazione creata in modo effettivamente casuale.

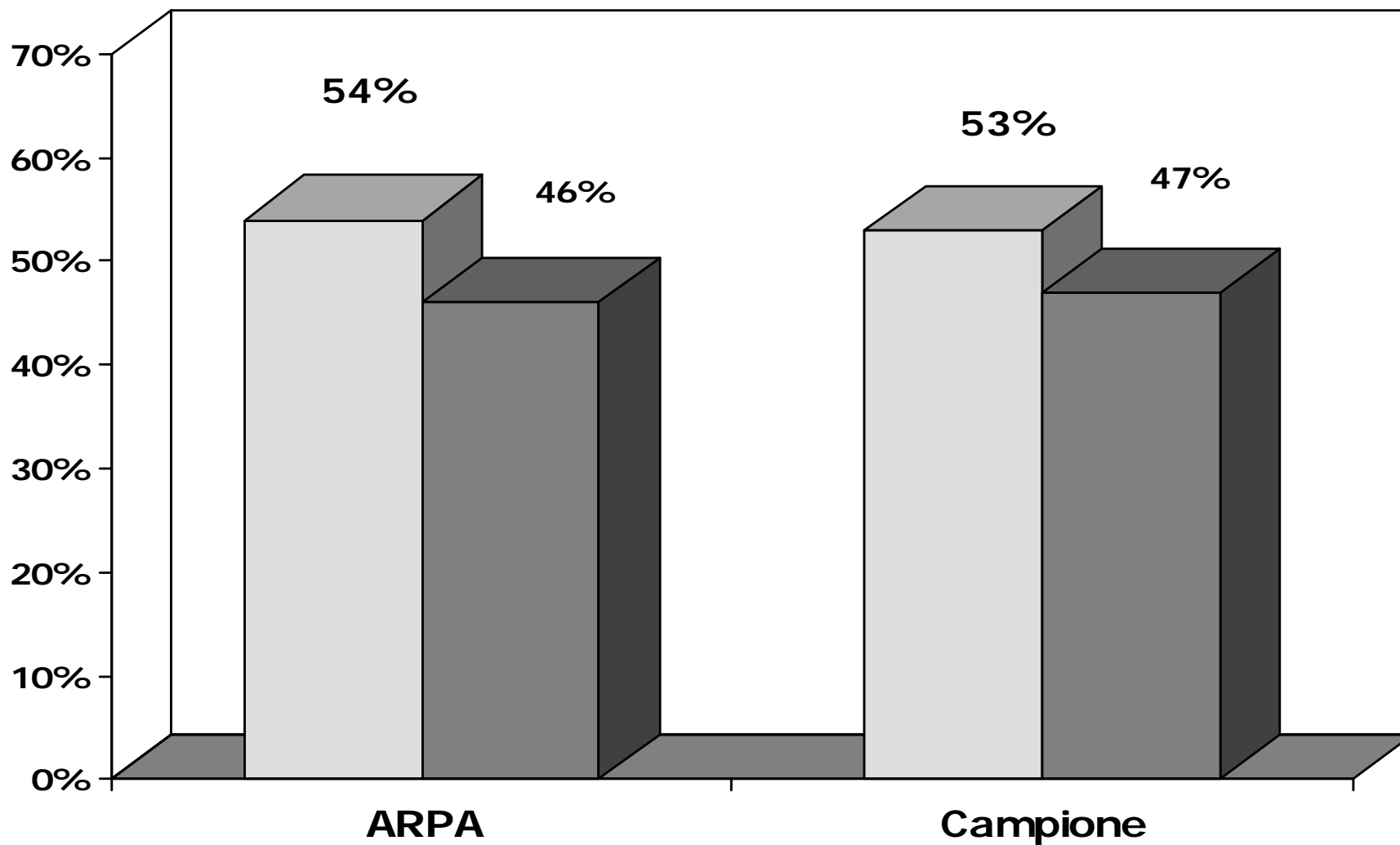
Il campionamento casuale in ARPA

Da un insieme comprendente 1.015 numeri, ne sono stati estratti casualmente 337. Prendendo poi la lista degli occupati in ARPA, ordinata in ordine alfabetico, i numeri estratti sono stati associati all'elenco degli occupati, individuando così le 337 persone da intervistare.

Verifica rappresentatività campione

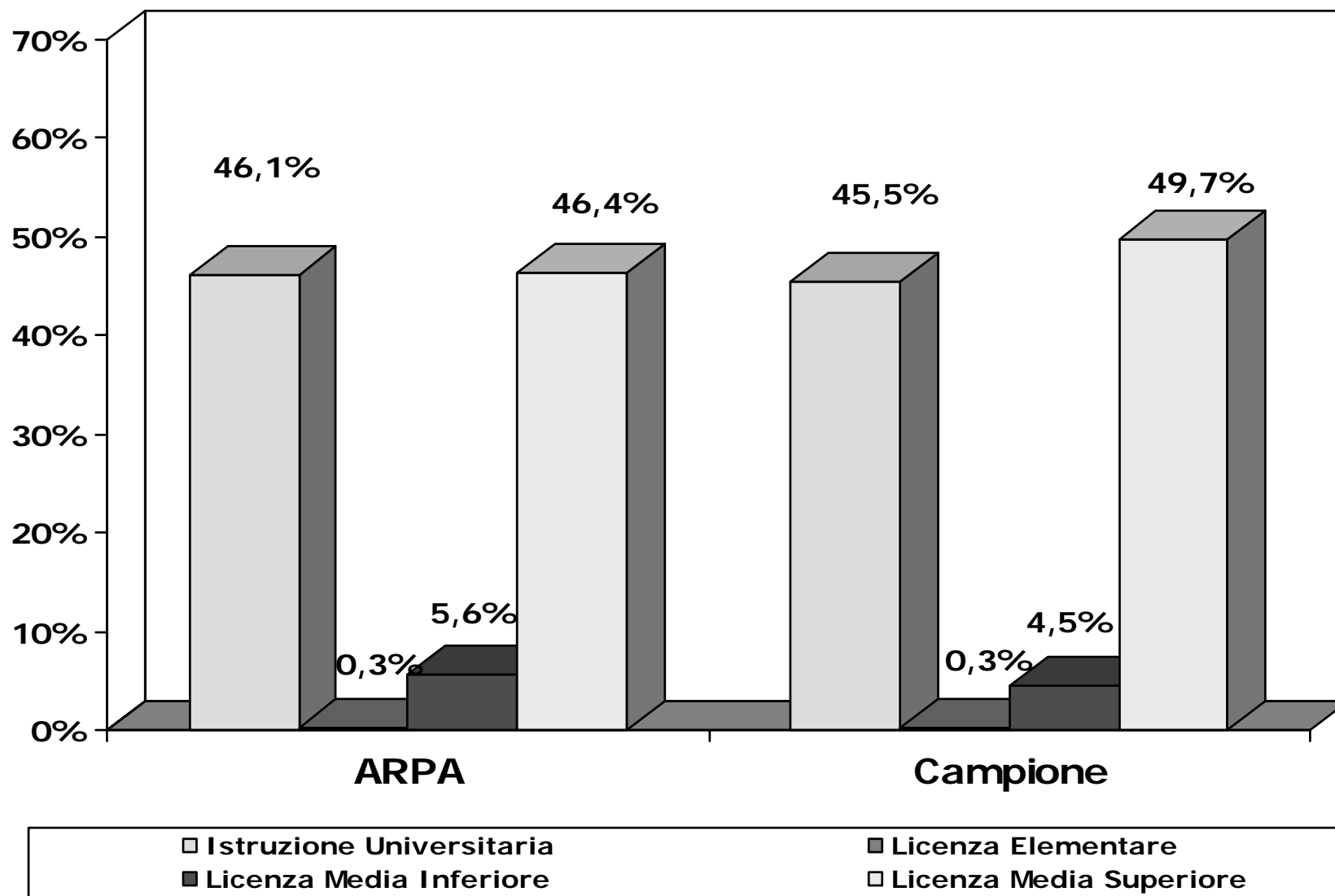
Una volta estratti i nominativi e definito il campione, le principali caratteristiche anagrafiche sono state confrontate con l'universo della popolazione lavorativa, in modo da verificare la congruità della scelta. Di seguito vengono riportati i principali confronti.

CONFRONTO DATI ARPA E CAMPIONE SELEZIONATO CASUALMENTE - RAPPRESENTATIVITA' DI GENERE

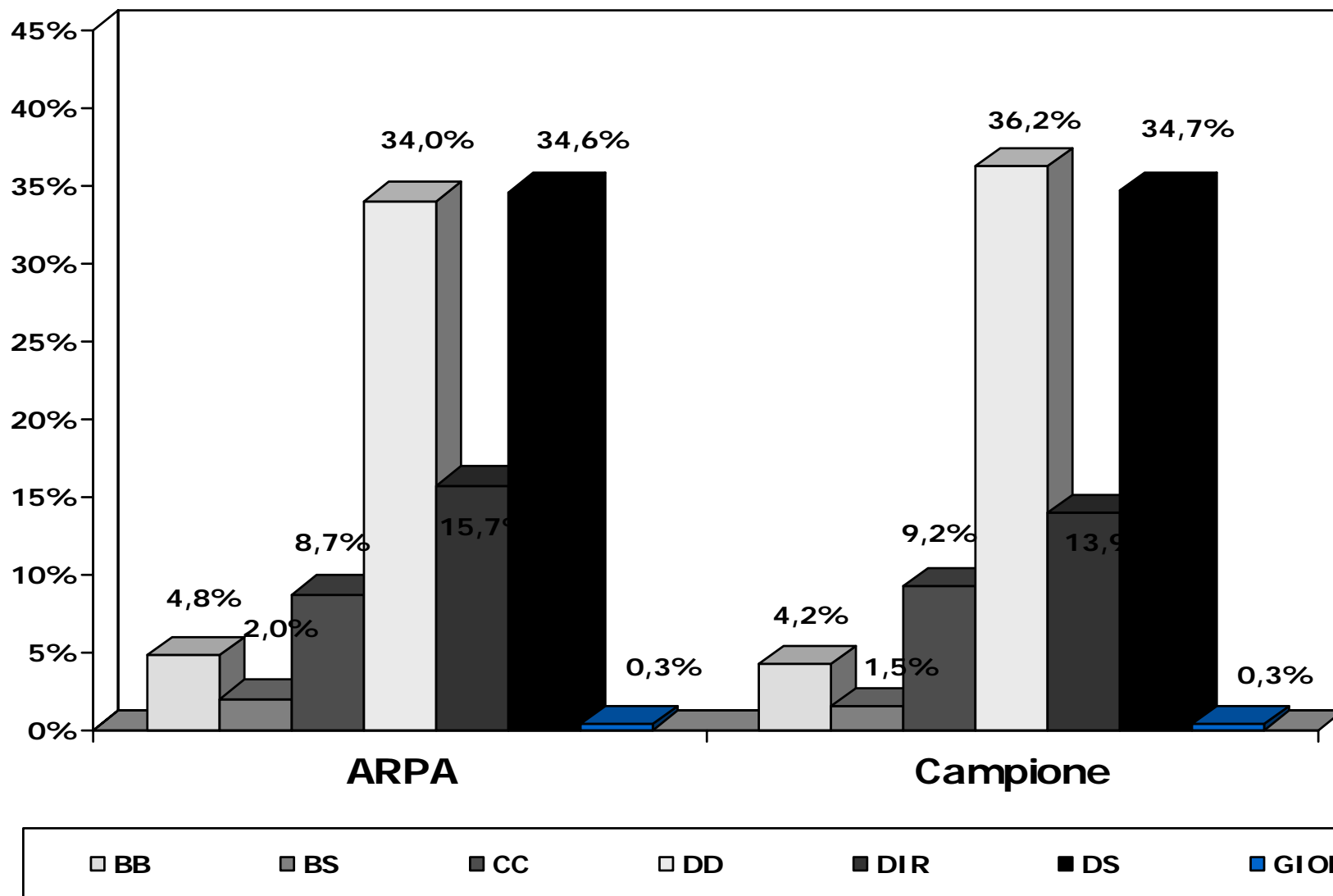


□ Donne ■ Uomini

CONFRONTO DATI ARPA E CAMPIONE SELEZIONATO CASUALMENTE – TITOLO DI STUDIO



CONFRONTO DATI ARPA E CAMPIONE SELEZIONATO CASUALMENTE – CATEGORIA



| NODI | ARPA | | CAMPIONE | |
|----------------------|------|-----------------|----------|-----------------|
| | % | Valore Assoluto | % | Valore Assoluto |
| DIR.GEN. | 0% | 1 | 0% | 0 |
| DIR.AMM. | 4% | 38 | 3% | 10 |
| DIR.TEC. | 6% | 56 | 5% | 17 |
| SAIPC | 2% | 21 | 1% | 5 |
| SGI:SQE | 1% | 8 | 1% | 3 |
| SSI | 1% | 11 | 1% | 3 |
| SOFE | 1% | 8 | 1% | 3 |
| Sez. BOLOGNA | 13% | 137 | 16% | 53 |
| Sez. FERRARA | 9% | 90 | 9% | 31 |
| Sez. FORLI '- CESENA | 8% | 78 | 7% | 24 |
| Sez. MODENA | 7% | 74 | 7% | 22 |
| Sez. PARMA | 7% | 67 | 7% | 22 |
| Sez. PIACENZA | 8% | 85 | 9% | 30 |
| Sez. RAVENNA | 9% | 96 | 11% | 36 |
| Sez. REGGIO EMILIA | 10% | 100 | 10% | 34 |
| Sez. RIMINI | 5% | 55 | 5% | 16 |
| SOD | 1% | 10 | 1% | 3 |
| SIMC | 8% | 80 | 7% | 25 |
| TOTALE | | 1.015 | | 337 |

INDIVIDUAZIONE DEI PARTECIPANTI ALLA RILEVAZIONE IN PRESENZA

1. Perché per la somministrazione del questionario in presenza è stata coinvolta solo una parte del personale?

Nelle indagini di tipo statistico, per conoscere determinati comportamenti di una popolazione, le informazioni possono essere desunte da:

- rilevazioni totali o censuarie: che prendono in considerazione l'intero universo della popolazione considerata (p.es. i censimenti);
- rilevazioni campionarie, quando l'analisi è limitata ad una parte dell'universo considerato, detto campione.

Le rilevazioni censuarie comportano un notevole dispendio di energie e di risorse, soprattutto quando l'universo a cui ci si riferisce è molto ampio. Per tali ragioni, in ARPA, si è preferito ricorrere a un'indagine campionaria. Per loro stessa natura, le indagini campionarie producono delle informazioni affette da errori, tuttavia presentano, rispetto alle indagini totali, alcuni vantaggi:

- il campione consente di ottenere informazioni in tempi più brevi,
- l'indagine campionaria ha un costo ridotto,
- l'indagine campionaria consente un più accurato controllo del processo di acquisizione dell'informazione.

2. Come è stato definito il numero di persone cui somministrare il questionario?

Come ricordato in precedenza, l'indagine campionaria porta sempre con sé una certa quota di errore, in quanto il campione non potrà mai contenere tutte le caratteristiche dell'universo considerato (per assurdo se anche si intervistassero 9 utenti su 10 di un servizio, non si potrebbe parlare di risultati assoluti, in quanto il comportamento dell'unica persona non considerata potrebbe essere molto difforme dalla media). Tale errore (definito *errore di campionamento*) può essere, adottando appropriati metodi di *inferenza statistica*, contenuto entro limiti più o meno ristretti. Queste tecniche, utilizzate anche per la determinazione del campione in ARPA, consentono di limitare gli errori, facilitando un processo di generalizzazione e il passaggio dal particolare al generale, ossia dal campione alla popolazione.

Grazie a quest'approccio statistico, partendo dall'universo della popolazione lavorativa in ARPA, pari a **1.015** unità¹, è stato possibile individuare il numero di persone da intervistare (337).

3. Come sono stati selezionati i partecipanti?

Da un insieme comprendente 1.015 numeri, ne sono stati estratti casualmente 337. Prendendo poi la lista degli occupati in ARPA, ordinata in ordine alfabetico, i numeri estratti sono stati associati all'elenco degli occupati, individuando così le 337 persone da intervistare. Questa tecnica di individuazione del campione, definita *campionamento casuale semplice*, è stata ritenuta la più idonea, in quanto riduce al minimo la discrezionalità nell'individuazione dei partecipanti all'indagine.

4. Il campione è rappresentativo?

Sì. Alcuni dati di verifica hanno dimostrato che il campione è rappresentativo dell'intera popolazione. Se si prendono, infatti, come riferimento alcune caratteristiche anagrafiche della popolazione lavorativa (genere, titolo di studio, nodo o struttura di appartenenza, categoria contrattuale), il campione presenta le stesse caratteristiche.

5. Che succede se qualcuno delle persone scelte non può partecipare alla somministrazione del questionario?

Per ogni partecipante sono stati, sempre con il metodo casuale, individuati dei possibili sostituti. Tuttavia per garantire i criteri di rappresentatività del campione è importante che le persone scelte partecipino alla giornata di somministrazione.

¹ L'universo comprende gli occupati a tempo indeterminato, tempo determinato e lavoratori comandati presso l'amministrazione, sono esclusi il Comitato di Direzione e i lavoratori Co.Co.Co.

| Rappresentatività rispondenti a questionari I.B.O. | | | |
|--|-----------------|--------------------|----------------|
| Dati Anagrafici | ARPA Universo | QUESTIONARI I.B.O. | |
| | Totale 1.015 | In Presenza 335 | On-line 209 |
| Genere | | | |
| M | 46% | 45% | 46% |
| F | 54% | 55% | 54% |
| Categoria | | | |
| DS | 28% | 34% | 28% |
| B/BS | 7% | 5% | 8% |
| C | 9% | 8% | 9% |
| D | 33% | 34% | 35% |
| D/DS tit. di PO | 7% | 6% | 7% |
| Dirigente | 16% | 13% | 13% |
| Età | | | |
| 18-30 | 2% | 2% | 2% |
| 31-40 | 22% | 23% | 27% |
| 41-50 | 43% | 45% | 44% |
| 51-60 | 30% | 26% | 26% |
| Oltre 60 | 3% | 3% | 1% |
| Tit.Studio | | | |
| Licenza Scuola Elem. | 0% | 0% | 0% |
| Licenza Media Inf. | 6% | 2% | 3% |
| Scuola Media Sup. | 46% | 49% | 45% |
| Laurea/Spec.post-laurea | 46% | 49% | 51% |
| Anni lav. Arpa | | | |
| 0-5 anni | 17% | 10% | 14% |
| 6-10 anni | 23% | 26% | 28% |
| Oltre 10 anni | 60% | 65% | 58% |

Esposizione dati percentuali con arrotondamento all'unità