



Città di Viareggio

PIANO STRUTTURALE DELLA CITTÀ DI VIAREGGIO

QUESTIONARI

Nota metodologica



PRINCIPI ATTIVI
atelier progettuale

Soggetto incaricato di curare il percorso partecipativo
Impostazione metodologica, conduzione del confronto, reporting degli esiti

Premessa

Per gli abitanti, i fruitori e i lavoratori di Viareggio è stato possibile far sentire da subito la propria voce, rispondendo alle domande di un **QUESTIONARIO dedicato a esplorare gli elementi di forza e debolezza del territorio, le sfide e le opportunità da cogliere, le criticità e i cambiamenti da affrontare, ma anche raccomandazioni e idee per migliorare insieme la città**. Per assicurare un'analisi inclusiva e dettagliata, sono stati predisposti quattro diversi format di questionario, ognuno destinato a un particolare segmento di comunità: il Questionario Under 14 per residenti giovanissimi, il Questionario Under 30 per giovani residenti e frequentatori, il Questionario Over 30 per residenti adulti, e il Questionario Imprese. La raccolta delle risposte si è svolta online dal 20 febbraio al 31 maggio 2024, utilizzando la **piattaforma SURVIO** (Sicurezza 2048-bit SSL, GDPR conforme).

I risultati ottenuti dai questionari hanno evidenziato temi ricorrenti ed elementi di interesse per la comunità, delineando un possibile punto di vista collettivo su come i rispondenti vedono, vivono, vogliono Viareggio. **Tali contenuti saranno oggetto di approfondimento durante le successive attività informative e consultive** (confronto pubblico).

Analisi dei risultati

ANALISI QUANTITATIVA

Per valutare l'attendibilità dei risultati ottenuti tramite il questionario, sono impiegati due parametri statistici specifici, il margine di errore e il livello di confidenza, che permettono di interpretare la rappresentatività dei dati raccolti rispetto alla popolazione totale.

Il **margine di errore** è una stima che indica entro quale intervallo percentuale i risultati di un sondaggio possono variare per riflettere la vera opinione della popolazione totale. Esso dipende principalmente dalla dimensione del campione e dalla variabilità delle risposte. Un margine di errore più piccolo indica che i risultati del sondaggio sono più vicini alla realtà effettiva della popolazione intera, risultando più precisi.

Il **livello di confidenza** indica quanto frequentemente il margine di errore calcolato cattura il vero valore nella popolazione se il sondaggio venisse ripetuto molte volte. Comunemente usato al 95%, questo parametro riflette la nostra sicurezza che, ripetendo lo stesso sondaggio, i risultati si posizioneranno entro il margine di errore il 95% delle volte.

La relazione tra il numero di rispondenti e questi parametri è diretta: aumentando la percentuale di rispondenti, diminuisce il margine di errore e si incrementa la precisione dei dati.

Questa corrispondenza è spiegata di seguito per diverse fasce di risposta (basato sul metodo di calcolo della dimensione del campione interno alle piattaforme di sondaggistica specializzate fra le più utilizzate: Qualtrics, Survio, Survey Monkey)

- **Percentuale di risposta fra 1-5% della popolazione** (parametri: livello di confidenza 95% margine di errore fra 2-3%) > **valore dei risultati "ESPLORATIVO"**: i dati raccolti potrebbero non riflettere l'intera comunità e dovrebbero essere interpretati con cautela. Tali risultati sono utili principalmente per identificare questioni ricorrenti da approfondire con ulteriori indagini o come punto di partenza per discussioni più ampie.
- **Percentuale di risposta fra 5-10% della popolazione** (parametri: livello di confidenza 95% margine di errore fra 1-2%) > **valore dei risultati "INDICATIVO"**: i dati raccolti forniscono una visione affidabile delle tendenze e delle opinioni prevalenti, ma ancora con alcuni limiti in termini di completa rappresentatività. Tali risultati sono utili per influenzare decisioni di media importanza e per formulare indirizzi iniziali.
- **Percentuale di risposta fra 10-25% della popolazione** (parametri: livello di confidenza 95% margine di errore fra 1-2%) > **valore dei risultati "RAPPRESENTATIVO"**: i dati raccolti sono considerati affidabili per decisioni di ampio raggio. Tali risultati forniscono una solida base che può guidare decisioni significative e influenzare politiche sostanziali, riflettendo un'ampia accettazione e partecipazione da parte della comunità.
- **Percentuale di risposta fra > 25% della popolazione** (parametri: livello di confidenza 95% margine di errore fra 0-1%) > **valore dei risultati "CONCLUSIVO"**: i dati raccolti sono estremamente affidabili e riflettono in modo accurato le opinioni e le preferenze dell'intera comunità. Tali risultati offrono una base decisiva per le politiche e le decisioni strategiche, garantendo che le azioni intraprese siano fortemente supportate da un consenso diffuso e ben rappresentato.

TARGET	N° RISPONDENTI	POPOLAZIONE DI RIFERIMENTO Dato ISTAT al 1/01/2023	PERCENTUALE DI RISPOSTA
RAGAZZI UNDER 14 età 8-13	297	2.917	10,2% (valore rappresentativo)
GIOVANI UNDER 30 abitanti età 14-29	250	8.824	2,8% (valore esplorativo)
GIOVANI UNDER 30 fruitori età 14-29	166	-	-
ADULTI OVER 30 età 30-99+	999	46.077	2,2% (valore esplorativo)
IMPRESE	38	-	
TOTALE RISPONDENTI	1750		

ANALISI QUALITATIVA

Per elaborare la sintesi delle risposte alle domande aperte, è stata condotta un'analisi **semantica bifase**, articolata nelle seguenti fasi:

- **FASE INIZIALE** - è stata realizzata una classificazione semantica di parole e concetti, impiegando tecniche di text mining supportate dall'uso di due modelli AI, GPT-4 Turbo e GPT-4o (Large Language Model) su tre differenti applicazioni (ChatGPT, Claude, Perplexity). È stato così possibile identificare temi ricorrenti, facilitando un'aggregazione coerente e organizzata dei contenuti dalle risposte.
- **FASE AVANZATA** - una revisione *manuale* (intervento umano) ha permesso di verificare e affinare i risultati ottenuti nella fase iniziale, assicurando che i temi e i concetti emersi riflettano accuratamente il contenuto e il senso delle risposte fornite.

L'output dell'**analisi semantica bifase** consiste in una **sintesi integrata** basata su criteri di ricorrenza e rilevanza delle risposte relative a una specifica domanda. A seconda dei contenuti, la sintesi integrata ha assunto la forma dei seguenti **output: definizioni indicative, slogan evocativi, elenchi organizzati, cluster tematici, set di parole, set di concetti chiave**.

Gli output dell'analisi semantica sono stati editati nel seguente modo:

- **formato di testo normale** - questo formato rappresenta testi che sono il risultato dell'aggregazione e della rielaborazione di contenuti affini identificati nelle risposte;
- **formato di testo corsivo** - utilizzato per i testi che derivano unicamente dall'aggregazione dei contenuti, con una revisione limitata alla sintassi dei testi originali;
- **testo virgolettato** - questo formato è adottato per i contenuti che sono stati estratti e riproposti esattamente come appaiono nelle risposte originali, mantenendo l'autenticità del linguaggio e delle espressioni utilizzate dai rispondenti.

Per alcuni output è stato possibile e/o significativo mettere in evidenza la **percentuale di ricorrenza** dei temi, cioè la frequenza con cui uno specifico sotto-insieme, di contenuti semanticamente affini, è stato identificato all'interno dell'insieme generale di risposte ad una o più domande, arricchendo la comprensione della rilevanza di ciascun tema all'interno del corpus di risposte analizzato. La **ricorrenza** è stata classificata secondo la seguente scala:

- **predominante** – percentuale di ricorrenza superiore al 50%;
- **elevata** – percentuale di ricorrenza compresa fra il 35% e il 50%;
- **alta** – percentuale di ricorrenza compresa fra il 25% e il 35%;
- **media** – percentuale di ricorrenza compresa fra il 15% e il 25%;
- **moderata** – percentuale di ricorrenza inferiori al 15%.

In alcuni output, è stata implementata anche una valutazione del **sentiment**, classificando il sentimento espresso come **positivo, negativo, neutro, neutro con flessioni positive e neutre con flessioni negative**. Questa distinzione ha permesso di aggiungere una ulteriore dimensione analitica che riflette la tonalità emotiva con cui i contenuti sono stati espressi, arricchendo la valutazione dei contributi raccolti.

Parallelamente all'analisi semantica bifase, è stata condotta un'**analisi interpretativa**: un approccio supplementare dedicato all'esame critico e trasversale di tutte le risposte aperte alle domande, al fine di estrarre ulteriori **insight qualitativi**, esplorando i significati impliciti nei dati testuali. L'analisi interpretativa ha consentito di ampliare la comprensione di percezioni, esperienze e visioni espresse dai rispondenti, arricchendo significativamente i risultati ottenuti dall'analisi semantica.

L'**output** dell'analisi interpretativa consiste in una serie di **riflessioni sinottiche, tematicamente organizzate**, che valorizzano il punto di vista dei rispondenti, sottolineando convergenze, divergenze, narrazioni e prospettive comuni.

Gli output dell'analisi interpretativa sono stati editati nel seguente modo:

- le riflessioni sono inserite all'interno di **box colorati** denominati "**Insight**";
- ogni **riflessione tematica** è **titolata ed organizzata in paragrafi** (ove funzionale);
- sono **evidenziate in grassetto le parole-chiave** estratte direttamente dalle risposte e sulle quali si incardina l'interpretazione dei dati testuali, così da sottolineare l'importanza di specifici contenuti all'interno del contesto analitico complessivo.

TARGET	N° DOMANDE totali	N° DOMANDE a risposta aperta	N° CARATTERI spazi esclusi	N° DI PAROLE raccolte/analizzate
RAGAZZI UNDER 14 età 8-13	23	10 (43,5%)	38.046	7.688
GIOVANI UNDER 30 abitanti età 14-29	29	9 (31%)	61.641	11.422
GIOVANI UNDER 30 fruitori età 14-29	12	6 (50,0%)	14.164	2.614
ADULTI OVER 30 età 30-99+	30	6 (20,0%)	258.931	47.061
IMPRESE	28	8 (28,5%)	22.768	3.827

APPUNTI

- Dal punto di vista dell'**analisi quantitativa**, il volume di risposte raccolte (dati numerici) ha permesso una trattazione dei dati di primo livello, dunque senza l'applicazione di filtri combinati. Ciò significa che non è stata effettuata una selezione simultanea di multiple opzioni di risposta per realizzare una profilazione dettagliata su aspetti specifici e circoscritti. Questo approccio ha limitato l'esame delle intersezioni complesse tra diverse variabili, ma ha comunque assicurato un'analisi adeguata e funzionale, consentendo l'emersione di temi e questioni rilevanti da valorizzare e approfondire nel percorso partecipativo.
- Dal punto di vista dell'**analisi qualitativa**, il volume di risposte raccolte (dati testuali) è stato superiore rispetto ad analoghe occasioni di indagine: i questionari sono stati articolati in domande a risposta chiusa e domande a risposta aperta, quest'ultime presenti in percentuali significative (da un minimo di 20% ad un massimo di 50%). Le domande a risposta aperta richiedono un tempo maggiore di compilazione e tendono ad essere trattate con superficialità o bypassando l'obbligatorietà con l'inserimento di lettere casuali, spaziature o generiche espressioni (es. "boh", "non so", "vedi sopra"). Solitamente su 100 rispondenti solitamente solo 20-30 rispondenti su 100 si soffermano sulla compilazione delle domande a risposta aperte. In questo caso la percentuale oscilla fra il 50-60% nel caso dei target giovani (under 30) e adulti (over 30), per giungere al 70% nel caso del target ragazzi (under 14), fino al 90% per il target imprese