

Business Intelligence per Big Data

Looker Studio

1. Login

Connettiti a [Google Looker Studio](https://lookerstudio.google.com), accedi con il tuo account Google o crea un nuovo account Google gratuito.

- <https://lookerstudio.google.com>

2. Report di benvenuto

Clicca su "Tutorial Report" e segui il tutorial.

- <https://lookerstudio.google.com/reporting/0B5FF6JBKbNJxOWItcWo2SVVVeGc>
- Impara le basi dello strumento Looker Studio copiando il "Welcome report" e seguendo le istruzioni passo-passo fornite.
- La pagina 8, "Track report usage with Google Analytics", può essere saltata.

3. Nuovo report sulle recensioni Airbnb Boston

Per creare un nuovo report da zero, è necessario identificare una fonte di dati. A tale scopo, una porzione del [dataset Kaggle delle recensioni Airbnb a Boston](#) è stata caricata in un [Google Sheets condiviso](#) da utilizzare come fonte dati per Looker Studio.

- Il Google Sheets, con circa 10.000 recensioni, è disponibile all'indirizzo: <https://docs.google.com/spreadsheets/d/1a2c9vCMFFfDXmhjoEoX2EwS2IYTbqE4WfZY72TXW9co/edit#gid=285360760>
- Dedica del tempo per comprendere i dati leggendo la loro descrizione su Kaggle e osservando la tabella su Google Sheets.
- La tabella della fonte dati è stata creata unendo le tabelle originali "Listings" e "Reviews" fornite da Kaggle, ed esportando le prime 10.000 righe unite ordinate per "listing_id" crescente.

Fonti di dati

Le fonti di dati hanno due tipi di campi: dimensioni e metriche.

- **Dimensione:** è una categoria di dati.
- **Metrica:** è un numero che quantifica qualcosa in quella categoria.
- Un report Looker Studio consente di visualizzare tali dimensioni e metriche in grafici e tabelle.
- Nelle proprietà delle fonti dati e del report di Looker Studio, le dimensioni appaiono come chip **verdi**, mentre le metriche appaiono come chip **blu**.

Creare un nuovo report

- Vai alla homepage di Looker Studio.
- Clicca su "Crea" → "Report" (vuoto).
- Rinomina il "Report senza titolo" con un nome a tua scelta cliccando sul nome stesso.

- Crea una nuova fonte dati cliccando sul pulsante blu in basso a destra, oppure seleziona la fonte dati Airbnb se è già presente nel pannello di destra.

Connessione alla fonte dati (Fogli Google)

Nella finestra "**Aggiungi dati al report**" che si aprirà automaticamente, seleziona il connettore "**Fogli Google**" (Di Google — "Connettiti a Fogli Google").

Dopo aver selezionato Fogli Google, completa la configurazione:

- Nella schermata successiva, scegli l'opzione "URL" nel pannello di sinistra.
- Incolla l'URL del Google Sheet Airbnb nel campo apposito:
<https://docs.google.com/spreadsheets/d/1a2c9vCMFFfDXmhjoEoX2EwS2lYTbqE4WfZY72TXW9co/edit>
- Seleziona il foglio di lavoro "Reviews Query DW" nel menu a tendina del foglio.
- Verifica che l'opzione "Usa la prima riga come intestazioni" sia spuntata.
- Clicca su "**Aggiungi**" per completare la connessione alla fonte dati.

4. Dimensioni, metriche e trasformazioni

- Verifica il **tipo** e l'**aggregazione** di ogni campo e che tutti i campi siano correttamente interpretati come dimensione o metrica.
- **Crea nuovi campi utili** (dimensioni o metriche) da quelli esistenti sfruttando le formule (clicca sui segnaposto "+" e "fx"). Per dettagli: <https://support.google.com/looker-studio/answer/6299685?hl=it>
 - `LENGTH(comments)` → per contare il numero di caratteri del campo commento
 - `CONCAT(latitude, CONCAT(', ', longitude))` → per generare un campo (lat, long) utile per le mappe; imposta prima "Tipo=Testo" per latitudine e longitudine
 - `price / square_feet` → prezzo medio per piede quadrato (prova con metri quadri: 1 piede = 0,3048 m)
 - `MONTH(Date)` → estrae il mese dalla data completa, es. 12
 - `YEAR(Date)` → estrae l'anno dalla data completa, es. 2017
 - `CONCAT(YEAR(Date), MONTH(Date))` → mese completo, es. 201712
- Dopo aver creato i nuovi campi e aggiornato quelli esistenti, clicca su "Aggiungi al report".

5. Analisi dei dati

Analizza i dati costruendo le seguenti visualizzazioni. Quindi, esplora e crea nuove visualizzazioni per trovare autonomamente insight interessanti.

Analisi (1) – Lunghezza commenti vs punteggio per tipo di proprietà

Obiettivo: confrontare il trend della lunghezza media dei commenti (numero di caratteri) rispetto al "review_scores_rating" medio, raggruppati per "property_type", ordinati per lunghezza commenti decrescente, con filtro per data.

Passo 1 – Inserire un grafico combinato (barre + linea)

- Nel menu in alto clicca su **Inserisci** → **Grafico** → **Grafico a barre**.
- Disegna l'area del grafico sulla tela del report tenendo premuto il mouse e trascinando.

- Il pannello delle proprietà si aprirà automaticamente sulla destra.

Passo 2 – Configurare i dati del grafico

- Nel pannello di destra, sezione **Dati**, verifica che la fonte dati selezionata sia "AirBnB reviews in Boston by Kaggle".
- Nel campo **Dimensione** clicca su "Aggiungi dimensione" e seleziona **property_type**.
- Nel campo **Metrica** clicca su "Aggiungi metrica" e seleziona **comment_length**. Looker Studio userà automaticamente l'aggregazione Media (AVG).
- Aggiungi una seconda metrica: clicca di nuovo su "Aggiungi metrica" e seleziona **review_scores_rating** (aggregazione: Media).

Passo 3 – Trasformare il grafico in combinato (barre + linea)

- Nel pannello di destra, clicca sulla scheda **Stile**.
- Nella sezione **Serie**, individua la serie "review_scores_rating".
- Clicca sul menu a tendina accanto al nome della serie e seleziona **Tipo: Linea**. La serie "comment_length" resterà come barra.
- Spunta l'opzione **"Asse destro"** per la serie "review_scores_rating", così i due assi Y saranno indipendenti.

Passo 4 – Ordinare per lunghezza commenti decrescente

- Torna alla scheda **Dati** nel pannello di destra.
- Nella sezione **Ordina**, clicca su "Aggiungi campo di ordinamento" e seleziona **comment_length**.
- Imposta la direzione su **Decrescente** (freccia verso il basso).

Passo 5 – Aggiungere un filtro per intervallo di date

- Nel menu in alto clicca su **Inserisci** → **Intervallo di date**.
- Posiziona il controllo sopra o accanto al grafico.
- Seleziona il controllo appena inserito, poi nel pannello di destra verifica che sia collegato alla stessa fonte dati del grafico.
- Torna al grafico, clicca su di esso e nel pannello **Dati** sezione **Filtri predefiniti**, verifica che il controllo data sia applicato (Looker Studio lo collega automaticamente per fonte dati).

Risultato atteso: Un grafico con le barre rosse (lunghezza media commenti) e una linea (punteggio medio), con le proprietà ordinate dalla più "verbosa" a sinistra. Le barche con commenti più lunghi non sempre corrispondono ai punteggi più alti.

Analisi (2) – Listing recensiti per room_type e mese

Obiettivo: confrontare l'andamento mensile del numero di "listing_id" distinti recensiti, separati per tipo di stanza ("room_type"), con filtri per data e tipo di superhost.

Passo 1 – Inserire un grafico a linee

- Nel menu in alto clicca su **Inserisci** → **Grafico** → **Grafico a linee**.
- Disegna l'area del grafico sulla tela.

Passo 2 – Configurare i dati

- Nel pannello **Dati**, imposta la **Dimensione** su **month** (il campo calcolato MONTH(Date) creato in precedenza).
- Imposta la **Dimensione di suddivisione** su **room_type**. In questo modo ogni tipo di stanza sarà una linea separata.
- Nella sezione **Metrica**, clicca su "Aggiungi metrica" e seleziona **listing_id**.
- Clicca sul nome della metrica "listing_id" nel pannello per aprire le opzioni e cambia l'aggregazione da **Somma** a **Conteggio distinto** (Count Distinct).

Passo 3 – Aggiungere i filtri interattivi

- Inserisci un controllo **Intervallo di date**: menu **Inserisci** → **Intervallo di date**. Posizionalo sopra il grafico.
- Inserisci un controllo **Elenco a discesa** per il superhost: menu **Inserisci** → **Elenco a discesa**.
- Seleziona il controllo a discesa appena creato; nel pannello **Dati**, imposta il **Campo di controllo** su **host_is_superuser**.

Risultato atteso: Tre linee (Entire home/apt, Private room, Shared room) che mostrano l'andamento mensile. Gli "Entire home/apt" risultano generalmente i più recensiti. Filtrando per `superhost=true` la distribuzione cambia visibilmente.

Analisi (3) – Numero di recensori distinti per posizione geografica

Obiettivo: visualizzare su mappa quante recensioni distinte ha ricevuto ogni alloggio, identificato dalla sua posizione geografica (lat, long).

Passo 1 – Inserire una mappa a bolle (Bubble Map)

- Nel menu in alto clicca su **Inserisci** → **Grafico** → **Mappa della bolla** (Bubble map).
- Disegna l'area della mappa sulla tela. Looker Studio mostrerà automaticamente la regione geografica in base ai dati.

Passo 2 – Configurare i dati della mappa

- Nel pannello **Dati**, imposta la **Dimensione geografica** sul campo calcolato **latlong** (il campo CONCAT creato in precedenza, di tipo Latitudine/Longitudine).
- Nella sezione **Metrica della bolla**, seleziona **reviewer_id** e cambia l'aggregazione in **Conteggio distinto**.

Passo 3 – Personalizzare la visualizzazione

- Clicca sulla scheda **Stile** nel pannello di destra.
- Regola il **colore delle bolle** e le **dimensioni minima/massima** per rendere le differenze ben visibili.
- La mappa si centererà automaticamente su Boston, Massachusetts, USA.

Risultato atteso: Una mappa del Massachusetts con bolle di dimensioni diverse concentrate nell'area urbana di Boston. Le bolle più grandi indicano gli alloggi con più recensori distinti.

Analisi (4) – Punteggio medio per tipo di proprietà e anno

Obiettivo: creare una tabella che mostri, per ogni coppia (property_type, year), il valore medio di review_scores_value, ordinata per property_type crescente e poi per punteggio medio decrescente, escludendo i valori nulli.

Passo 1 – Inserire una tabella

- Nel menu in alto clicca su **Inserisci** → **Tabella**.
- Disegna l'area della tabella sulla tela.

Passo 2 – Configurare dimensioni e metriche

- Nel pannello **Dati**, aggiungi come prima **Dimensione** il campo **property_type**.
- Aggiungi come seconda Dimensione il campo **year** (il campo calcolato YEAR(Date)).
- Aggiungi come **Metrica** il campo **review_scores_value** con aggregazione **Media**.

Passo 3 – Configurare l'ordinamento

- Nella sezione **Ordina** del pannello Dati, aggiungi come primo criterio **property_type** con direzione **Crescente**.
- Aggiungi come secondo criterio di ordinamento **review_scores_value** con direzione **Decrescente**.

Passo 4 – Escludere i valori nulli di property_type

- Nella sezione **Filtro** del pannello Dati, clicca su **"Aggiungi un filtro"**.
- Clicca su **"Crea un filtro"** nella finestra che si apre.
- Imposta la condizione: campo **property_type** → operatore **"Non è nullo"** (Is not null) → salva il filtro.

Risultato atteso: Una tabella con 44 righe (property_type × year), raggruppata alfabeticamente per tipo di proprietà e, al suo interno, ordinata per punteggio decrescente. Gli Apartment del 2009 risultano tra i più apprezzati.

Analisi (5) – Recensioni con punteggio massimo per anno e tipo di stanza

Obiettivo: visualizzare il conteggio delle recensioni con review_scores_value = 10 per ogni anno e room_type, confrontandolo con il numero di listing distinti recensiti.

Passo 1 – Creare un campo calcolato per le recensioni top

- Vai alla fonte dati (menu **Risorse** → **Gestisci le origini dati aggiuntive** → clicca su **Modifica** accanto al tuo dataset).
- Clicca su **"Aggiungi un campo"** (icona "+").
- Assegna il nome **is_top_score** e inserisci la formula:

```
CASE WHEN review_scores_value = 10 THEN 1 ELSE 0 END
```
- Imposta il tipo su **Numero** e l'aggregazione su **Somma**. Salva e torna al report.

Passo 2 – Inserire un grafico a barre raggruppate

- Nel menu in alto clicca su **Inserisci** → **Grafico** → **Grafico a barre**.
- Disegna l'area del grafico.

Passo 3 – Configurare i dati

- Nel pannello **Dati**, imposta la **Dimensione** su **year**.
- Imposta la **Dimensione di suddivisione** su **room_type** (per avere le barre raggruppate per tipo di stanza).
- Aggiungi come **Metrica** il campo **is_top_score** (Somma = conteggio recensioni con voto 10).

Passo 4 – Aggiungere il confronto con listing distinti

- Aggiungi una seconda **Metrica**: seleziona **listing_id** con aggregazione **Conteggio distinto**.
- Nella scheda **Stile**, imposta questa seconda metrica come **Linea** sull'asse destro, così da sovrapporla alle barre senza confondere le scale.

Risultato atteso: Un grafico a barre raggruppate per anno (dal 2009 al 2016) con le tre tipologie di stanza affiancate. Il 2016 mostra il maggior numero di recensioni top. La linea sovrapposta dei listing distinti aiuta a capire se la crescita è proporzionale all'aumento degli alloggi recensiti.