|  |
| --- |
| Power BI Report Development Workflow  Dalla modellazione alla visualizzazione |
| Vladyslav Bendovskyy <https://sites.google.com/view/bvlad/home> |

## Olist Dataset 2016-2018



### **MODELLAZIONE DATI**

Dopo un’iniziale esplorazione delle tabelle, sono state selezionate le seguenti:

* **olist\_orders\_dataset**
* **olist\_order\_items\_dataset**
* **olist\_customers\_dataset** (rinominata **customers**)
* **olist\_order\_review\_dataset** (rinominata **reviews**)
* **olist\_product\_dataset**
* **product\_category\_name\_translation**

Dopo aver effettuato attività di pulizia dei dati — tra cui l'utilizzo di funzioni come Trim, la rimozione di duplicati, la verifica dei formati e l’eliminazione di colonne non significative ai fini analitici — sono state implementate le seguenti modifiche:

* In **olist\_orders\_dataset** è stata aggiunta la colonna calcolata *delivery\_time*, che misura i giorni di consegna a partire dalla data dell’ordine.
* Nella tabella **customers**, si è notato che la colonna *customer\_id* si riferisce a ciascuna transazione, mentre *customer\_unique\_id* al cliente effettivo. Inoltre, *customer\_state* è stata sostituita con i nomi delle regioni per esteso (basati su dati da Wikipedia), per evitare confusioni con omonimie internazionali.
* Le tabelle **olist\_orders** e **olist\_order\_items** sono state unite in un’unica tabella chiamata **orders**. In questa, la colonna *delivery\_time* è stata rimpiazzata da *delivery\_time\_grouped*, che raggruppa i valori in tre classi: <7, 7 to 10, >10. Questa scelta facilita l’utilizzo dei filtri. Il merge è stato effettuato per semplificare lo schema delle relazioni, trattandosi di un dataset statico.
* Nella tabella **reviews**, i duplicati su *order\_id* sono stati rimossi in quanto legati a modifiche successive a una stessa recensione.
* Infine, è stato effettuato un merge tra **olist\_product\_dataset** e **product\_category\_name\_translation**, ottenendo la tabella **products\_&\_categories**, contenente gli ID prodotto e il nome della categoria in inglese.

Immagine che contiene testo, schermata, Carattere, logo

Il contenuto generato dall'IA potrebbe non essere corretto.

### **DAX E MISURE**

* **tabella calendar** : create per una gestione corretta delle analisi temporali

**orders:**

* **\_order\_purchase\_date :** ottenuta tramite DATEVALUE a causa di problematiche sulla colonna originale order\_purchase\_timestamp
* **\_tot\_orders** : numero di ordini, calcolato con DISTINCTCOUNT(order\_id)
* **\_tot\_orders\_PY** : calcolata con SAMEPERIODLASTYEAR e condizionato da HASONEVALUE affinché sia visibile nei visual soltanto quando è selzionato un anno in modo da evitare aggregazioni insensate (se non si filtra per un anno non ha senso comparare i trend con “l’anno passato”)
* **\_tot\_revenue** : calcolato con SUMX su price e freight\_value
* **\_tot\_revenue\_PY :** stessa formula di tot\_orders\_PY
* **\_var\_order :** creata con DIVIDE (([\_tot\_orders]-[\_tot\_orders\_PY]) , orders[\_tot\_orders\_PY], "--")
* **\_var\_revenue** : come \_var\_order

**customers**:

* **\_max\_region\_orders** : misura creata per ottenere la regione con più ordini CALCULATE(FIRSTNONBLANK('customers'[customers\_state], 1),

                     TOPN(1, SUMMARIZE(customers, 'customers'[customers\_state], "Total Orders", [\_tot\_orders]), [\_tot\_orders], DESC))

* **\_max\_region\_revenue** : come sopra
* **\_max\_region\_review** : come sopra

**reviews**:

* **\_review** : misura create con DISTINCTCOUNT per comodità di calcolo in \_max\_region\_review

Immagine che contiene testo, schermata, Carattere, logo

Il contenuto generato dall'IA potrebbe non essere corretto.

### **DETTAGLI SUI VISUAL**

La palette cromatica è stata definita in coerenza ai codici colore del sito web aziendale e dell’immagine del logo. La lingua inglese è stata una scelta dettata dal fatto che si tratta di un’azienda brasiliana.

I filtri di ogni pagina sono posizionati sulla sinistra della canvas.

In basso a destra si trovano i pulsanti di navigazione tra le pagine.

L’aggiunta di una scheda relativa alla regione permette anche di visualizzare il nome di quest’ultima se si filtra attraverso la mappa.

Dove opportuno, sono state aggiunte *descrizioni comando*. Per esempio, nei grafici mostra valori percentuali come la variazione rispetto l’anno precedente oppure sulle icone per descrivere le loro azioni.

* Nelle prime due pagine (**Orders** e **Revenue**) i grafici mostrano misure aggregate per mese per poter visualizzare eventuali trend stagionali. Se si filtra per anno, verrà aggiunto al grafico l’andamento dell’anno precedente a quello selezionato.
* In **Revenue**, sono stati esclusi gli ordini con status canceled o unavailable, non apportando ricavi effettivi.
* La pagina **Product** introduce una navigazione tramite icone, che rimandano ad analisi mirate su ordini e profitti. È possibile filtrare le categorie attraverso i grafici ad anello.
* Nella pagina **Rating** vengono rimosse le schede relative ai totali e alla variazione annua per fare spazio ad un grafico a linea che analizza l’andamento nel tempo della media dei voti di recensione.
* Infine, nella pagina **Insight**, cliccando sull’icona del robot è possibile interagire con la funzione Q&A, con domande suggerite e sinonimi già predisposti. La pagina inizialmente si presenta con una guida all’utente e ,una volta selezionata una delle icone, si passa all’analisi relativa.

Immagine che contiene testo, schermata, Carattere, logo

Il contenuto generato dall'IA potrebbe non essere corretto.