

Università Politecnica delle Marche

Facoltà di Ingegneria

Corso di Laurea in Ingegneria Informatica e dell'Automazione



Analisi con Qlik sulla vendita di telefono in Indonesia

DOCENTI

Prof. Ursino Domenico

Prof. Luca Virgili

STUDENTI

Mori Nicola

Sospetti Mattia

Zitoli Francesca

Anno accademico 2021-2022

Indice

1	Introduzione	3
1.1	Qlik	3
2	Gestione dei dati	4
2.1	Dataset	4
2.2	ETL	4
2.3	Caricamento dei dati	5
3	Dashboard	6
3.1	Dashboard Generale	6
3.2	Dashboard Prodotto	8
3.3	Dashboard Luogo	10
3.4	Dashboard Analisi Temporale	12

1 Introduzione

1.1 Qlik

Qlik, figura 1, è un software di business intelligence che consente di creare dashboard che descrivono in modo immediato le interazioni che esistono nei dati di un dataset. Il software riesce ad essere di semplice utilizzo con un'interfaccia semplice ed immediata, fornendo ulteriormente consigli su quali dati sono in correlazione tra di loro grazie al machine learning. Da sottolineare la possibilità di utilizzare un cloud con altri colleghi per modificare il lavoro svolto e condividere dati, dashboard, consigli e workspace.



Figura 1: Logo Qlik

2 Gestione dei dati

2.1 Dataset

Il dataset scelto per il progetto di business intelligence riguarda le vendite di telefoni cellulari in Indonesia nel periodo temporale che comprende Marzo e Aprile del 2020. È composto da un unico file .csv con 16 colonne corrispondenti ai campi. Di seguito riportiamo la tabella che riassume la struttura del dataset:

ID	Tipo	Descrizione
date	Timestamp	Data.
sku_name	String	Nome del telefono.
product_id	String	Id associato al telefono.
merchant_code	Number	Id sotto forma di numero come identificatore del mercante.
mercant_badge	String	Tipologia del mercante.
product_type	String	Tipologia telefono.
merchant_city	String	Città del mercante.
cod	Boolean	Contrassegno.
price	Number	Prezzo del telefono (in rupie indonesiane).
stock	Number	Numeri di telefoni in stock.
daily_gmv	Number	Valore totale creato nella giornata.
daily_sold	Number	Vendite giornaliere.
daily_view	Number	Viste giornaliere.
rating	Number	Valore della recensione (0-100).
total_review	Number	Numero totale di recensioni fino alla data registrata.
total_sold	Number	Numero totale di vendite fino alla data registrata.

2.2 ETL

La trasformazione dei dati è stata eseguita attraverso un programma Python poiché abbiamo notato che in alcuni casi gli attributi venivano letti in modo errato poiché nel dataset di base il carattere che separa i campi era la virgola, questo portava dei conflitti con l'attributo `sku_name` il quale a suo volta presentava una virgola al suo interno, ciò comportava un problema in fase di lettura. Lo `sku_name`, quindi, non veniva interpretato come un unico attributo poiché la virgola al suo interno provocava la sua lettura come due

attributi diversi, quindi questo comportava lo scorrimento di tutti i campi successivi ad esso provocando la lettura errata del dataset. Per risolvere tale problema abbiamo modificato le stringhe affinché non presentassero la virgola (sostituita con un semplice spazio, " "), abbiamo cambiato inoltre l'elemento separatore scegliendo il “;”. Applicata questa modifica vi erano 3 tuple rimanenti che contenevano il ";", a quel punto abbiamo deciso di procedere manualmente nel modificarle.

2.3 Caricamento dei dati

Nel momento in cui abbiamo caricato il dataset in Qlik non abbiamo avuto bisogno di legare alcuna tabella con un'altra visto che di base il dataset era composto da un unico file .csv. Abbiamo notato alcune incongruenze in alcuni campi nel dataset:

- Il campo **stock** contiene valori molto variabili, in cui in alcuni casi sono di base 999999, probabilmente un valore di default che alcuni negozi hanno inserito nel dataset, mentre altri negozi come valore hanno inserito 0. Avere un milione circa di telefoni in stock ci sembra eccessivo, quindi date queste incongruenze, presenti in larga misura, con cui sono stati inseriti i dati in questo attributo abbiamo deciso di non utilizzarlo nelle nostre dashboard.
- Inoltre sono presenti alcuni dati mancanti nel periodo indicato, ovvero in alcuni casi non sono presenti informazioni in alcuni giorni, questo compromette ad esempio quando si effettua la sommatoria del campo *daily_sold* (ovvero l'attributo che esprime le vendite nella singola giornata) per tutti i giorni indicati il valore ottenuto non corrisponde con il campo *total_sold*.
- Il prezzo è stato mantenuto in rupie indonesiane nelle nostre valutazioni nonostante vi era l'idea di convertirlo in euro, poiché per analizzare i risultati ottenuti abbiamo ritenuto essere di migliore lettura avere di fronte la valuta locale.

3 Dashboard

Abbiamo deciso di suddividere il lavoro in diverse *dashboard*, ognuna delle quali approfondisce un determinato argomento:

- **Dashboard Generale:** presenta delle informazioni generali o che verranno approfondite successivamente. Ha lo scopo di mostrare le prime generali informazioni relative al Dataset, così da avere una visione generale di quello che si affronterà poi.
- **Dashboard Prodotto:** approfondisce informazioni legate al prodotto, il focus si sposta quindi ad esempio sui prezzi, sulle vendite e sulle informazioni legate ai prodotti.
- **Dashboard Luogo:** vengono analizzate maggiormente le informazioni relative al luogo, come ad esempio sono disposti i negozi, dove si verificano vendite più alte, etc..
- **Dashboard Analisi Temporale:** viene approfondita l'evoluzione temporale dei dati, come ad esempio vendite o prezzi.

3.1 Dashboard Generale

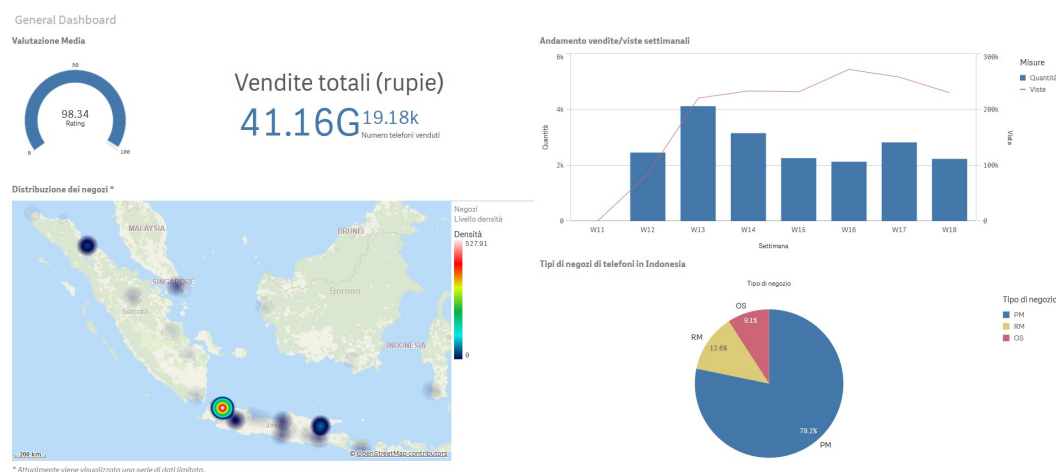


Figura 2: Dashboard Generale

La dashboard generale, Figura 2, ci fornisce delle informazioni generali sulla situazione in Indonesia in merito alla vendita di telefoni cellulari ed ha lo scopo di introdurre i temi che poi verranno approfonditi nelle dashboard successive.

Il grafico in alto a sinistra mostra quale è il rating medio dei telefoni nella nazione dove si può notare che è del 98,34 su 100, il che ci suggerisce che non è un buon indicatore per valutare le vendite in base al rating, in quanto non è un discriminante che aiuta nell'analisi dei dati.

Il grafico a fianco invece mostra il dato sulle vendite totali, ovvero 19.18k di telefoni e il denaro prodotto da tali vendite ovvero 41.16G (rupie) .

In basso a sinistra abbiamo un grafico che mostra invece una mappa che mostra la distribuzione dei negozi in Indonesia; qui si può subito notare un exploit in corrispondenza della zona di Jakarta, evidente il centro in rosso, dove c'è un'alta densità di negozi.

In alto a destra possiamo notare la porzione della dashboard che rappresenta con l'istogramma l'andamento delle vendite settimanalmente, mentre per quanto riguarda le viste l'andamento è rappresentato dalla linea arancione. Un aspetto interessante che questo grafico mostra è il fatto che nonostante con l'avanzare delle settimane il numero di viste sui prodotti aumenta la quantità delle vendite diminuisce.

In basso a destra abbiamo un grafico a torta che mostra in percentuale i tipi di negozi in Indonesia e vediamo che la maggior parte è di negozi fisici, rispetto a quelli online e retail. Qui possiamo notare come siano dominanti i negozi fisici che forniscono il servizio di telefoni cellulari.

3.2 Dashboard Prodotto

Nella dashboard legata al prodotto, Figura 3 abbiamo focalizzato l'attenzione sul Prodotto e sui dati legati ad esso.

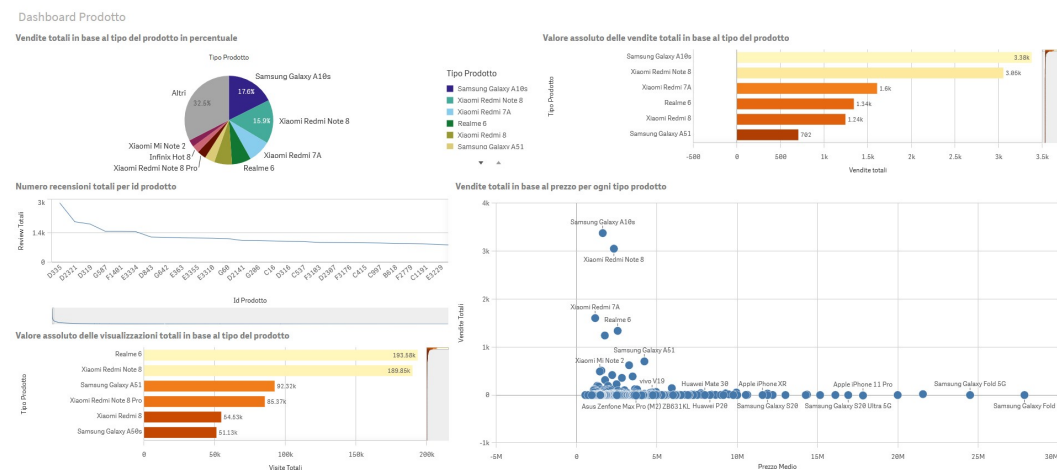


Figura 3: Dashboard Prodotto

In alto a sinistra abbiamo un grafico a torta che mostra in percentuale quali sono i telefoni più venduti, di cui possiamo notare che il Samsung Galaxy A10s insieme allo Xiaomi Redmi Note 8 occupano insieme più del 30% delle vendite, seguiti dallo Xiamo Redmi 7A e dal Realme 6. La combinazione di questi 4 telefoni occupa circa il 50% delle vendite sul mercato indonesiano il che ci suggerisce che non c'è una forte differenziazione tra i clienti.

A destra abbiamo un grafico a barre che mostra invece il numero quantitativo delle vendite totali del periodo considerato dal dataset per ogni tipo di prodotto, dove possiamo notare che il Samsung Galaxy A10s ha totalizzato 3.38k di vendite e lo Xiaomi Redmi Note 8 invece 3.05k di vendite, lo Xiami Redmi 7A ha registrato 1.6k di vendite ed infine, il Realme 6 1.34k, con un totale di 9.37k di vendite come già detto quasi il 50% delle vendite complessive.

A metà dashboard a sinistra abbiamo su un asse l'ID dei prodotti mentre sull'altro il numero di Review Totali. Nel nostro caso possiamo notare che il D335 è il più recensito e corrisponde al modello ASUS Zenfone con circa 3k recensioni, il secondo e il terzo invece sono un unico telefono ovvero lo Xiaomi Redmi Note 7 considerati diversamente per via del colore, il quarto invece è lo Xiaomi Redmi Note 4. Da questo particolare grafico abbiamo intuito che valutare i prodotti in base al numero di recensioni non è ottimale in quanto i valori sono probabilmente influenzati dalla longevità del prodotto;

l'alto valore è dato dall'accumulo di recensioni anche prima del periodo da noi analizzato. I telefoni più venduti quindi non corrispondono allo stesso tempo a quelli più recensiti.

Sotto a questo grafico poi possiamo individuare un istogramma dove nella colonna abbiamo il tipo del prodotto e come lunghezza della colonna le visite totali, notiamo un gran numero di visite per il modello Realme6 che però risulta essere il quarto più venduto, mentre in seconda posizione abbiamo lo Xiaomi Redmi Note 8 che notiamo invece essere il secondo più venduto, mentre il Samsung Galaxy A10s nonostante abbia poche visualizzazioni ha un alto numero di vendite, probabilmente grazie alla fama del brand. Quindi possiamo notare che ad un alto valore di visualizzazioni effettivamente c'è una maggiore vendita di prodotti, questo però non si verifica per un prodotto come il Samsung Galaxy A10s che grazie al marchio consolidato nel mercato riesce ad ottenere un alto numero di vendite nonostante le poche visualizzazioni.

L'ultimo grafico in basso a destra mostra in base al prezzo di ogni tipo di prodotto le vendite effettuate (prezzo in rupie sull'asse delle x, vendite totali sulla y). Notiamo che il prezzo medio per i telefoni più venduti è tra i 2 e 4 milioni di rupie indonesiane, mentre i telefoni di 'lusso' risultano avere una quantità di vendite irrilevanti, questo dà un'idea di quale sia il potere di acquisto del cliente medio indonesiano che si aggira in un valore compreso tra i 120€ e i 250€.

3.3 Dashboard Luogo

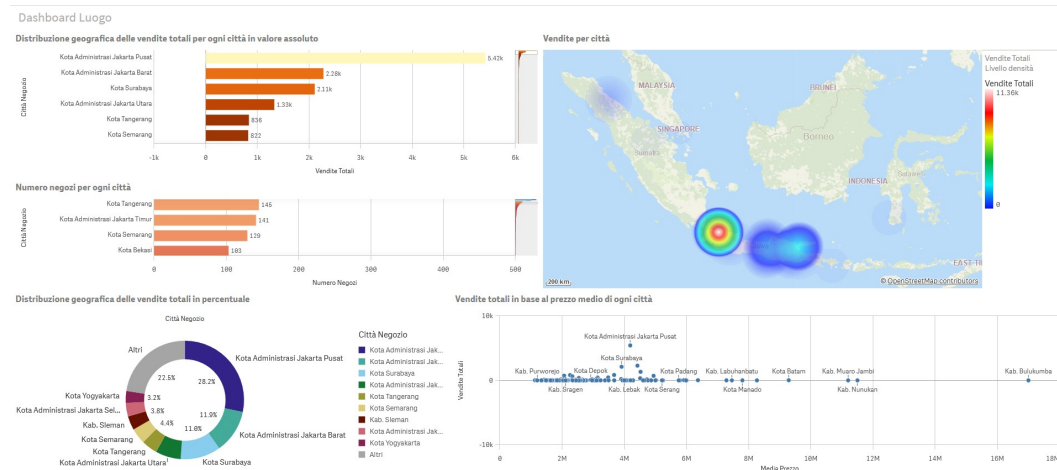


Figura 4: Dashboard Luogo

La dashboard del luogo, Figura 4 mostra le analisi che abbiamo svolto focalizzandoci sulla localizzazione. Il grafico a barre in alto a sinistra mostra come si distribuiscono le vendite per ogni città, in particolar modo il valore assoluto delle vendite; come si può notare Kota Administrasi Jakarta Pusat (Zona centrale di Giacarta) domina le vendite con 5.42k di vendite. Segue invece Kota Administrasi Jakarta Barat (Zona occidentale di Giacarta) con 2.28k, ovvero. Anche al quarto posto notiamo una zona di Jakarta in particolare Utara (Zona Nord di Giacarta). Come si notava nella mappa corrispondente alla Dashboard generale Jakarta è il fulcro delle vendite di smartphone in Indonesia.

I dati riguardanti la centralità di Jakarta sono anche mostrati dal numero di negozi che vendono telefoni in questa zona dove nel grafico al centro-sinistra Kota Administrasi Jakarta Barat con 470 negozi, e poi 375 con Kota Administrasi Jakarta Pusat.

Il grafico ad anello in basso a sinistra mostra in percentuale come le vendite si distribuiscono nelle varie zone/città dell'Indonesia, dove in volume Pusat occupa il 28.2%, dove insieme al 11.9% di Kota Administrasi Jakarta Barat e Kota Surabaya con l'11% costituiscono circa il 51% delle vendite complessive in Indonesia.

Infine in basso a destra abbiamo confrontato le vendite con il prezzo medio per ogni città, cercando di trovare una correlazione tra il prezzo d'acquisto e il volume dei telefoni acquistati. Notiamo il prezzo medio degli acquisti è sui 4G per il quartiere Pusat, dove al tempo stesso ci sono più vendite,

quindi suggerisce che la zona centrale di Giacarta sia quella più benestante a livello economico dato che il prezzo medio è nel limite superiore del range che avevamo stabilito precedentemente nella dashboard del prodotto.

3.4 Dashboard Analisi Temporale

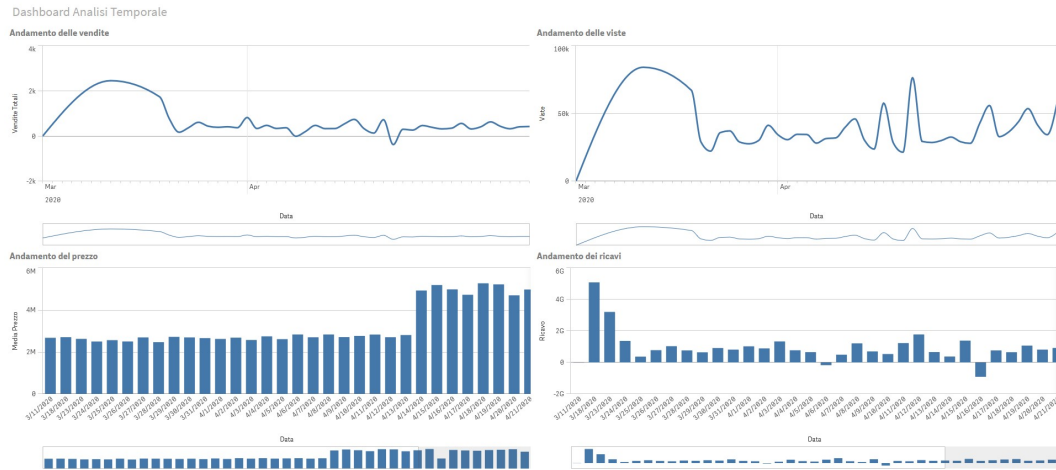


Figura 5: Dashboard Analisi Temporale

Infine abbiamo svolto un'analisi in base al tempo che comprende il periodo tra l'11 Marzo e il 30 Aprile 2020, Figura 5.

In alto a sinistra abbiamo le vendite giornaliere di telefoni in Indonesia, come si può notare non c'è un vero e proprio trend, si nota un picco a Marzo, mentre ad Aprile un dato particolare sono le vendite negative, probabilmente dovute alla restituzione di telefoni durante il periodo del Covid, o in seguito ad un terremoto avvenuto durante quel periodo.

In alto a destra invece abbiamo il grafico delle visualizzazioni giornaliere, anche qui abbiamo un trend irregolare dove notiamo un picco di visual a metà Marzo e a metà Aprile.

Invece in basso a sinistra abbiamo deciso di analizzare la media dei prezzi, dove in particolar modo c'è un aumento durante il periodo di metà aprile, osservando l'andamento della moneta abbiamo visto che c'è stata una svalutazione e quindi l'aumento dei prezzi che poi ha anche portato nello stesso periodo il volume di vendite giornaliere negativo.

Infine in basso a destra l'andamento dei ricavi segue in particolar modo l'andamento delle vendite quindi è probabile che il prezzo medio delle vendita sia costante, ovvero che le vendite avvengono sempre per gli stessi prezzi che abbiamo anche analizzato nella Dashboard Prodotto.