

# AmazonDB – Progetto BD2

Autore: Vincenzo Tarantino, Bartolomeo Mazzeo

## 1. Introduzione

Il progetto sviluppato è un sistema per la gestione, visualizzazione e analisi dei prodotti venduti su Amazon India e delle relative recensioni da parte degli utenti.

L'obiettivo è creare un'infrastruttura scalabile, flessibile e modulare per l'inserimento, la consultazione e la modifica delle informazioni sui prodotti e sui feedback degli acquirenti.

Il sistema si basa su tecnologie NoSQL (in particolare MongoDB) ed è pensato per supportare un'interfaccia web sviluppata usando Flask.

## 2. Individuazione dei dataset e preprocessing

Per il progetto è stato utilizzato un dataset pubblico disponibile su kaggle al link: in formato CSV contenente informazioni su prodotti venduti su Amazon appartenenti a diverse categorie merceologiche.

Il dataset presenta un'unica tabella con le seguenti colonne:

- **product\_id**  
Identificatore univoco del prodotto su Amazon.  
→ *Singolo valore per riga.*
- **product\_name**  
Nome completo del prodotto.  
→ *Singolo valore per riga.*
- **category**  
Categoria/e del prodotto espressa come stringa  
→ *Può contenere più categorie come stringa unica.*

- **discounted\_price**  
Prezzo scontato del prodotto (stringa con simbolo ₹ (rupia)).  
→ *Singolo valore per riga.*
- **actual\_price**  
Prezzo originale del prodotto prima dello sconto.  
→ *Singolo valore per riga.*
- **discount\_percentage**  
Percentuale di sconto sul prodotto.  
→ *Singolo valore per riga.*
- **rating**  
Valutazione media del prodotto.  
→ *Singolo valore per riga.*
- **rating\_count**  
Numero totale di recensioni ricevute dal prodotto.  
→ *Singolo valore per riga.*
- **about\_product**  
Descrizione dettagliata del prodotto.  
→ *Singolo valore per riga.*
- **user\_id**  
Identificatori degli utenti che hanno recensito il prodotto, separati da virgola.  
→ *Contiene più valori per riga.*
- **user\_name**  
Nomi degli utenti autori delle recensioni, separati da virgola.  
→ *Contiene più valori per riga.*
- **review\_id**  
Identificatori univoci delle recensioni, separati da virgola.  
→ *Contiene più valori per riga.*

- **review\_title**  
Titoli delle recensioni, separati da virgola.  
→ *Contiene più valori per riga.*
- **review\_content**  
Testi completi delle recensioni, separati da virgola.  
→ *Contiene più valori per riga.*
- **img\_link**  
URL all'immagine del prodotto.  
→ *Singolo valore per riga.*
- **product\_link**  
URL alla pagina Amazon del prodotto.  
→ *Singolo valore per riga.*

In fase di preprocessing, è stato effettuato il parsing delle colonne contenenti valori multipli (user\_id, user\_name, review\_id, review\_title, review\_content) per separare ogni recensione in una riga distinta, mantenendo il collegamento con il relativo prodotto.

### 3. Architettura Back-End

#### 3.1 Database NoSQL: MongoDB

Il database è stato modellato in MongoDB, sfruttando la flessibilità del modello a documenti.

Sono state definite quattro collezioni principali:

- Products: contiene i dati commerciali dei prodotti;
- Reviews: contiene le recensioni associate a ciascun prodotto, collegate tramite product\_id;
- Users: contiene le informazioni degli utenti recensori (user\_id, user\_name) e consente di individuare l'utente che ha fatto una

determinata recensione per un determinato prodotto.

- Categories: contiene le categorie disponibili al momento dell'inserimento di un nuovo prodotto.

I riferimenti tra collezioni sono gestiti tramite foreign key logiche che aiutano ad effettuare le JOIN.

### 3.2 Server-Side Logic: Python e Flask

La logica di gestione dei dati è implementata in Python in maniera combinata con il framework Flask, tipico per applicazioni web leggere. Flask consente di gestire le rotte HTTP in modo semplice e flessibile, facilitando la creazione di un'interfaccia RESTful per l'accesso e la manipolazione dei dati memorizzati su MongoDB.

Nel dettaglio:

Il file **app.py** rappresenta il cuore dell'applicazione Flask.

In questo modulo sono definite tutte le rotte dell'applicazione, ciascuna associata a una funzionalità specifica come:

- creare, leggere, aggiornare o cancellare un prodotto o una recensione;
- effettuare ricerche complesse sui prodotti combinando filtri come prezzo, categoria e rating;
- interagire con gli utenti, ad esempio per associare recensioni agli utenti registrati.

Ogni rotta gestisce la logica di ricezione e validazione dei dati provenienti dal front-end, invoca le funzioni di CRUD e restituisce le pagine HTML renderizzate.

Il file **crud.py** contiene tutte le funzioni di interazione diretta con il database MongoDB.

Queste funzioni sono suddivise logicamente per collection (prodotti e recensioni,) e si occupano di:

- Creare nuovi documenti nelle rispettive collezioni.
- Leggere e filtrare i dati secondo le specifiche esigenze dell'utente.
- Aggiornare campi dei documenti esistenti.
- Cancellare documenti in base agli identificatori.

In questo modo, il file `crud.py` svolge il ruolo di modulo di accesso ai dati, incapsulando tutte le query e le operazioni di database, mentre `app.py` si concentra sulla gestione delle rotte e della logica di business.

#### 4. Architettura Front-End

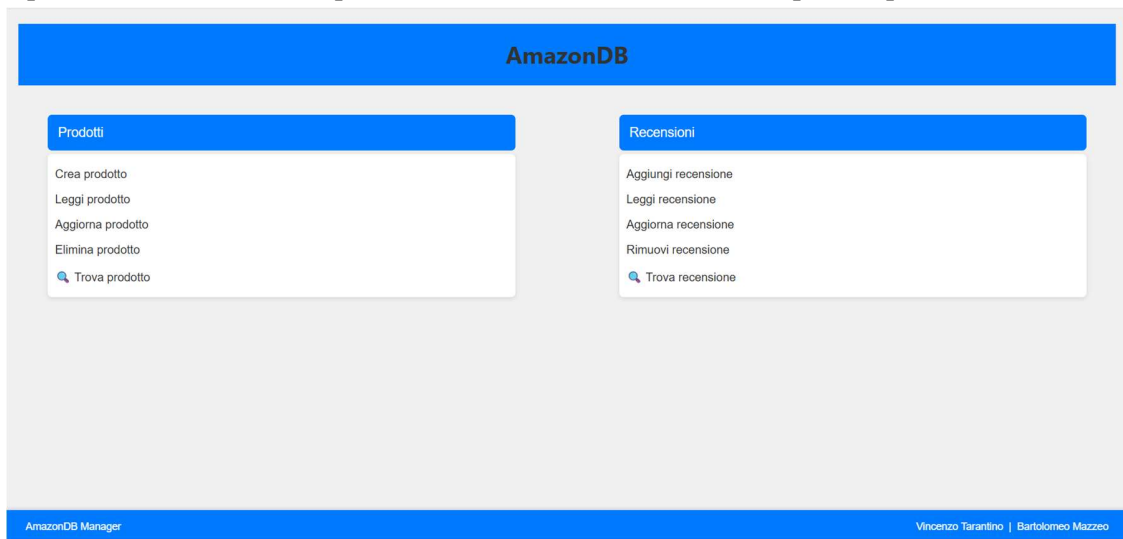
L'interfaccia utente è stata sviluppata come una **Web App interattiva** basata su Flask per il rendering server-side delle pagine HTML, arricchite con fogli di stile CSS personalizzati per garantire un'esperienza utente semplice, pulita e coerente.

L'UI è composta da diverse pagine dedicate alle principali operazioni CRUD (Create, Read, Update, Delete) sui prodotti e sulle recensioni, oltre a funzionalità di ricerca avanzata:

##### **Homepage (home.html)**

La pagina principale del sito, contiene due aree Prodotti e Recensioni che se cliccate mostrano un menù che porta alle pagine per effettuare le

operazioni CRUD sui prodotti e recensioni, oltre a quelle per la ricerca.



## Creazione Prodotto (create\_product.html)

Permette di inserire un nuovo prodotto specificando:

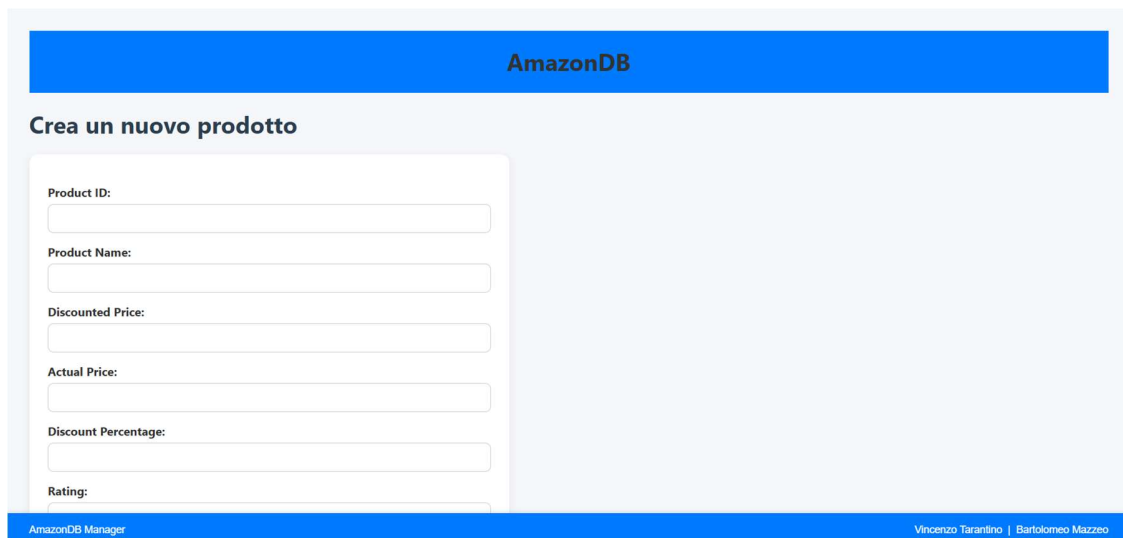
- ID prodotto
- nome
- prezzo non scontato
- percentuale di sconto
- rating iniziale
- descrizione
- link immagine e link alla pagina prodotto
- categorie selezionabili da una lista dinamica

Un controllo lato server assicura che tutti i campi obbligatori siano compilati, che siano rispettati i formati e che le categorie selezionate

esistano effettivamente nel database(vengono mostrate solo le categorie già esistenti nel DB).

Il numero di recensioni di un prodotto appena creato è di default 0, viene aggiornato dinamicamente quando si crea una nuova recensione per quel prodotto.

Il prezzo scontato viene calcolato applicando lo sconto sul prezzo iniziale.



The screenshot displays the 'AmazonDB Manager' web application. At the top, there is a blue header with the text 'AmazonDB'. Below the header, the main content area is titled 'Crea un nuovo prodotto'. On the left side of this area, there is a form with several input fields: 'Product ID:', 'Product Name:', 'Discounted Price:', 'Actual Price:', 'Discount Percentage:', and 'Rating:'. Each field has a corresponding text input box. The form is set against a light blue background. At the bottom of the page, there is a blue footer bar containing the text 'AmazonDB Manager' on the left and 'Vincenzo Taranino | Bartolomeo Mazzeo' on the right.

## **Lettura Prodotto (read\_product.html)**

Consente di cercare un prodotto selezionandolo tra tutti quelli nel DB e visualizzare tutte le informazioni relative, incluse le categorie di appartenenza e i dati commerciali.

AmazonDB

### Leggi Prodotto

Seleziona un prodotto dal menu a tendina:

-- Seleziona un prodotto --

ID Prodotto:

Cerca

AmazonDB ManagerVincenzo Tarantino | Bartolomeo Mazzeo

## Aggiornamento Prodotto (update\_product.html)

Permette di modificare uno o più campi di un prodotto esistente, comprese le categorie.

Il modulo è dinamico: vengono aggiornati solo i campi compilati nel form, lasciando inalterati gli altri.

AmazonDB

### Aggiorna Prodotto

Seleziona un prodotto dall'elenco:

-- Seleziona un prodotto --

ID Prodotto da aggiornare:

Campi da aggiornare (lascia vuoto se non vuoi cambiarli):

Nome:

Prezzo reale (actual price):

es. 1000 o ₹1000

Sconto (%):

AmazonDB ManagerVincenzo Tarantino | Bartolomeo Mazzeo

## Elimina Prodotti (delete\_products.html)



## Elimina un prodotto selezionato

The screenshot shows a web application titled 'AmazonDB Manager'. The main heading is 'Elimina Prodotto' in red. Below it, a instruction says 'Seleziona un prodotto dall'elenco:'. There is a dropdown menu with the placeholder text '-- Seleziona un prodotto --'. Below the dropdown, there is a form with a label 'ID Prodotto da eliminare:' and an input field. A red button labeled 'Elimina' is positioned below the input field. The footer of the page contains the text 'AmazonDB Manager' on the left and 'Vincenzo Tarantino | Bartolomeo Mazzeo' on the right.

## Ricerca Avanzata Prodotti (search\_products.html)

Offre un'interfaccia per effettuare ricerche complesse combinando diversi filtri:

- categoria (dropdown)
- range di prezzo minimo e massimo
- rating minimo

La pagina mostra l'elenco dei prodotti trovati, aggiornato dinamicamente in base ai criteri di ricerca impostati.

The screenshot shows the 'AmazonDB Manager' interface. At the top is a blue header with the text 'AmazonDB'. Below the header, the main section is titled 'Ricerca Prodotti'. On the left side of this section, there is a white form with the following fields: 'Categoria:' with a dropdown menu showing '-- Seleziona una categoria --'; 'Rating minimo:' with a text input field; 'Prezzo minimo:' with a text input field; 'Prezzo massimo:' with a text input field; and a green 'Cerca' button at the bottom. The footer of the interface is a blue bar containing 'AmazonDB Manager' on the left and 'Vincenzo Tarantino | Bartolomeo Mazzeo' on the right.

## Aggiunta Recensione (create\_review.html)

Permette di inserire una nuova recensione selezionando:

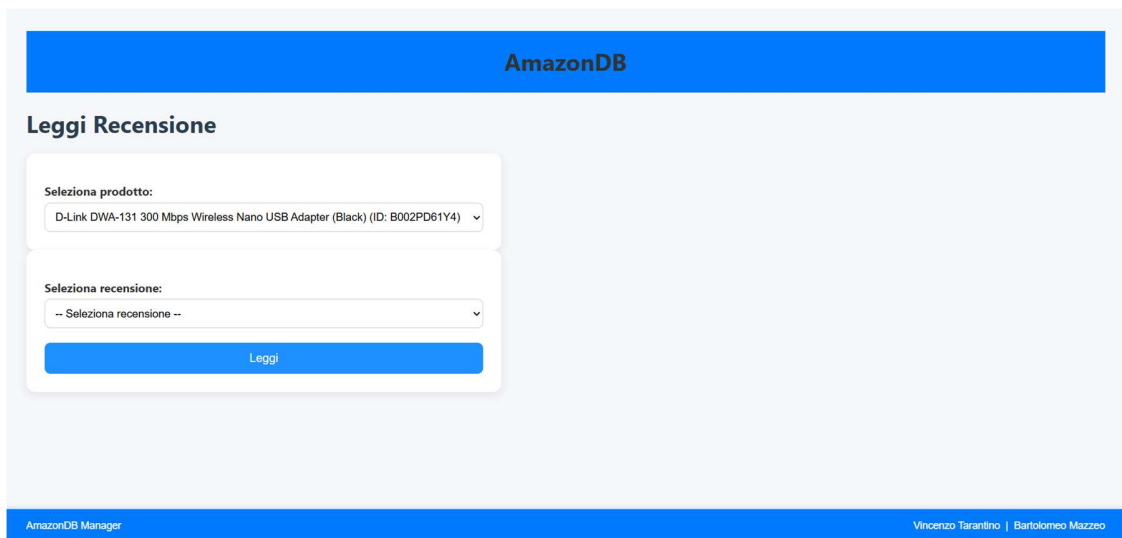
- il prodotto (tramite dropdown popolato dinamicamente)
- l'utente autore della recensione
- titolo e contenuto della recensione

In automatico la recensione viene associata al prodotto e all'utente selezionati.

The screenshot shows the 'AmazonDB Manager' interface. At the top is a blue header with the text 'AmazonDB'. Below the header, the main section is titled 'Aggiungi Recensione'. On the left side of this section, there is a white form with the following fields: 'ID Recensione:' with a text input field; 'Prodotto:' with a dropdown menu showing '-- Seleziona un prodotto --'; 'Utente:' with a dropdown menu showing '-- Seleziona un utente --'; 'Titolo Recensione:' with a text input field; and 'Contenuto Recensione:' with a text input field. At the bottom of the form is a blue 'Aggiungi' button. The footer of the interface is a blue bar containing 'AmazonDB Manager' on the left and 'Vincenzo Tarantino | Bartolomeo Mazzeo' on the right.

## Lettura Recensione (read\_review.html)

Consente di visualizzare una recensione specifica cercando il prodotto a cui è associata la recensione, selezionando la recensione interessata tra tutte quelle disponibili per il prodotto scelto (se presenti recensioni per quel prodotto) e mostrando tutte le informazioni relative alla recensione (contenuto, autore, prodotto collegato).



The screenshot displays the 'AmazonDB Manager' application. At the top, a blue header bar contains the text 'AmazonDB'. Below this, the main section is titled 'Leggi Recensione'. It features a form with two dropdown menus. The first dropdown, labeled 'Seleziona prodotto:', has a selected value: 'D-Link DWA-131 300 Mbps Wireless Nano USB Adapter (Black) (ID: B002PD61Y4)'. The second dropdown, labeled 'Seleziona recensione:', has a selected value: '-- Seleziona recensione --'. Below these dropdowns is a blue button labeled 'Leggi'. At the bottom of the interface, a blue footer bar contains the text 'AmazonDB Manager' on the left and 'Vincenzo Tarantino | Bartolomeo Mazzeo' on the right.

## Aggiornamento Recensione (update\_review.html)

Offre la possibilità di modificare il titolo o il contenuto di una recensione esistente.

Anche in questo caso, solo i campi compilati vengono aggiornati, mentre gli altri restano invariati.

AmazonDB

### Aggiorna Recensione

Seleziona prodotto:  
D-Link DWA-131 300 Mbps Wireless Nano USB Adapter (Black) (ID: B002PD61Y4) ▾

Seleziona recensione da aggiornare:  
-- Seleziona recensione -- ▾

Nuovo Titolo:

Nuovo Contenuto:

Aggiorna

AmazonDB ManagerVincenzo Tarantino | Bartolomeo Mazzeo

## Eliminazione Recensione (delete\_review.html)

Fornisce un form essenziale per eliminare una recensione specificando il suo ID.

Il bottone per la cancellazione è evidenziato con uno stile che richiama il rosso, a indicare la natura distruttiva dell'operazione.

AmazonDB

### Rimuovi Recensione

Seleziona prodotto:  
D-Link DWA-131 300 Mbps Wireless Nano USB Adap ▾

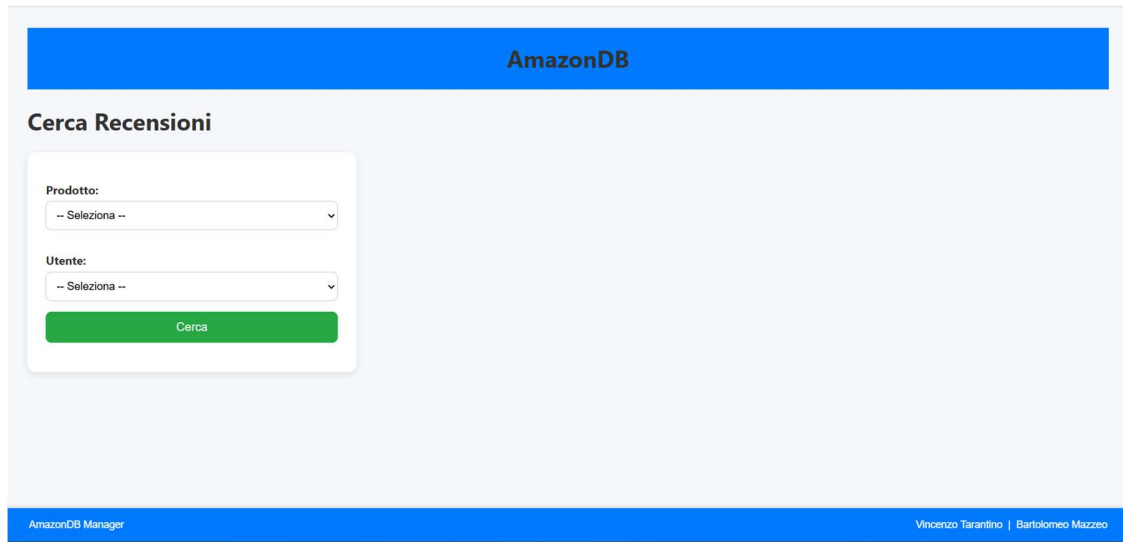
Seleziona recensione da rimuovere:  
-- Seleziona recensione -- ▾

Rimuovi

AmazonDB ManagerVincenzo Tarantino | Bartolomeo Mazzeo

## Ricerca dettagliata Recensione (search\_review.html)

Consente di specificare un prodotto e/o un utente per ricercare tutte le recensioni per un determinato prodotto, tutte le recensioni fatte da un determinato utente e tutte le recensioni fatte da un determinato utente per uno specifico prodotto



The screenshot shows a web application titled "AmazonDB Manager". At the top, there is a blue header bar with the text "AmazonDB" in white. Below the header, the main content area has a light blue background. On the left side of this area, there is a white box titled "Cerca Recensioni". Inside this box, there are two dropdown menus: the first is labeled "Prodotto:" and the second is labeled "Utente:". Both dropdown menus currently show "-- Seleziona --" with a small downward arrow on the right. Below these two dropdowns is a green button with the text "Cerca" in white. At the bottom of the page, there is a blue footer bar. On the left side of the footer, it says "AmazonDB Manager" in white. On the right side of the footer, it says "Vincenzo Tarantino | Bartolomeo Mazzeo" in white.

## Intestazioni e Footer

Tutte le pagine condividono header e footer che contengono rispettivamente il link per tornare alla home e branding dell'applicazione, per un'esperienza coerente e continua durante l'uso