

Forward Deployed Engineer

Overview

Merchandising analysts often need to retrieve insights tied to:

- Regional performance differences
- Category-level revenue contribution
- Channel-specific sales patterns (Online vs. Offline)

For example, an analyst may need to answer:

- “How did *Beverages* perform in the *Sub-Saharan Africa* region during Q2?”
- “Which item types generated the highest revenue in *Online* channels last year?”
- “What were the sales trends for *Cosmetics* across regions in 2013–2015?”

Since the dataset is large and analysts frequently slice by multiple dimensions (region × item × channel × time), manually filtering this data is slow. Build a retrieval system that lets a user express a natural-language query and returns an explanation grounded in the dataset.

Requirements

- UI allowing free-text queries referencing:
 - Region (e.g., “Europe”, “Middle East and North Africa”)
 - Item type
 - Sales channel
 - Date ranges / quarters / years
- Backend must:
 - Identify the relevant subset of rows
 - Aggregate metrics such as units sold, revenue, average selling price
 - Support multi-facet questions (e.g., item × region × time window)

Deliverables

- Demo of the system and executable codebase
- High level overview of the solution

Dataset

<https://www.kaggle.com/datasets/kyanyoga/sample-sales-data>