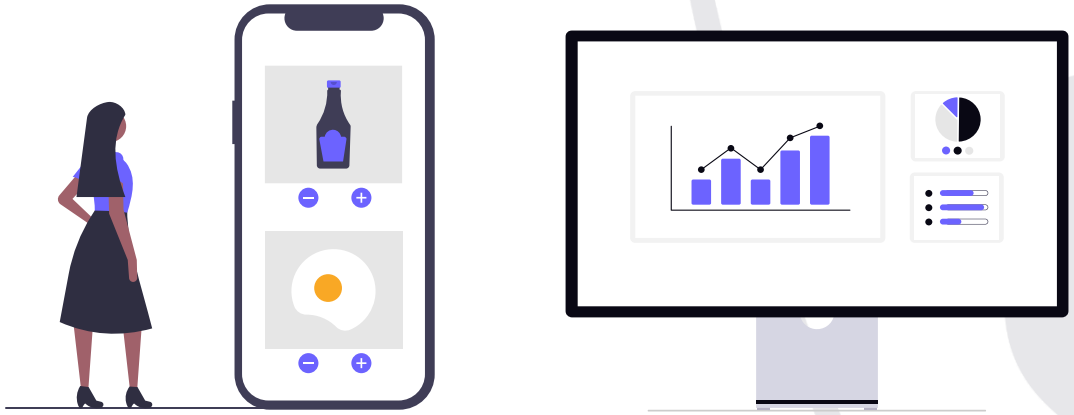
McGill | DESAUTELS

PriceScout

a South African price intelligent platform



Client: GigEfx

Project Owner: Ziyan Chen

June 27th, 2025

Executive Summary

User Pain Points:

- **Emerging Opportunities:** The rapid growth in mobile users and the surge in online shopping have created substantial business opportunities for centralized price comparison platforms. However, there is currently no such comprehensive platform available in the South African market, leaving a significant gap to be filled.
- **Consumer Inconvenience:** Without a unified platform, shoppers must manually browse multiple websites, which is time-consuming and can lead to uninformed purchase decisions.
- **Lack of Comprehensive Historical Data:** There is no single resource for tracking and analyzing historical pricing trends, limiting consumers' ability to assess true value and timing for purchases.

Solution Contents	Descriptions
Database Design	• Entity Relationship Diagram (ERD) for a price intelligent platform • Data Definition Language(DDL) script for database creation • Manual data retrieval from several platforms and inserted to database as a proof of concept
Local Market Research	• Local Market Research on South African retailors with data source prioritization recommendations
Data Retrieval Guide	• Data Retrieval Guide outlines the essential fields mapped from retailor websites to the project's designated database.

Solution Results:

- Purpose-built database structure enables efficient data storage and seamless aggregation, supporting robust future data analysis.
- Recommended data sources, identified through focused market research, ensure reliable, comprehensive, and accessible information gathering.
- Clear data retrieval guide provides sustainable approach for data mapping between database and retailor platforms
- Automation-ready framework allows for scalable solutions and future operational efficiency.

Introduction-Overview

Project Overview

GigEfx is introducing **PriceScout**, an innovative, community-focused price intelligence platform aimed at empowering South African consumers. The core mission is to provide transparent, real-time, and historical retail pricing information for mobile devices across major retailers.

Goal

While PriceScout is in its early development phase, this foundational project is centered on conducting initial market research and establishing a robust database design, effectively laying the groundwork to enable and accelerate future platform development.

Key Objectives:

1. Establish a scalable and secure relational database capable of integrating and managing diverse retail data sources.
2. Identify and evaluate leading South African retail platforms to ensure broad, reliable data coverage for price, availability, and product attributes.
3. Develop clear data mapping and integration guides, enabling efficient future automation and minimizing manual intervention.

Target Community

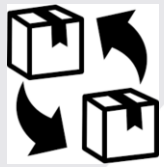
South African mobile users—shoppers, families, and tech enthusiasts—who seek to make smarter, faster, and more informed purchasing decisions in an increasingly digital retail environment.

Introduction-Goal

PriceScout product features:


Price Comparison


Store Information


Product Alternatives


Product Availability


Sales Event


Product Ratings

- Establishes the foundational infrastructure for the PriceScout app through the design of a scalable and well-structured database.
- Ensures all data is systematically organized and readily accessible.
- Determines the specific data requirements for each feature by defining key database fields and structure, ensuring all essential information is captured and appropriately organized for future integration and scalability.

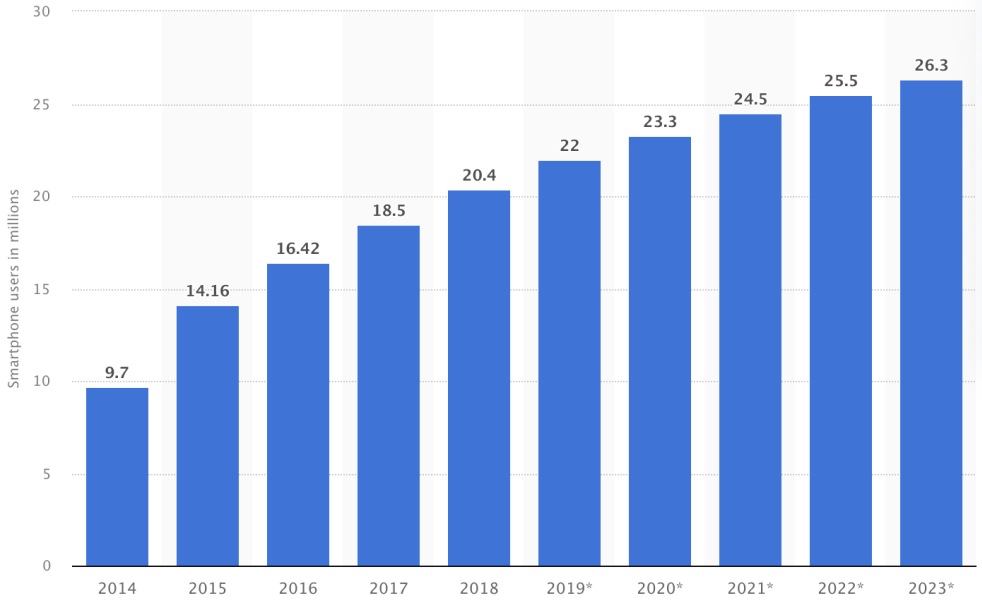
Introduction-Goal

- Conducts targeted local market research to select trustworthy and high-value data sources for ongoing and future data integration.
- Develops a practical data retrieval guide that ensures efficient and systematic integration of external data into the platform.
- Lays the groundwork for future automation and scalable development, enabling continuous improvements and advanced features for end users.

Introduction-Impact

Market Opportunity

- **Continued growth of mobile users in South Africa**: The number of smartphone users in South Africa was forecast to continuously increase between 2024 and 2029. The smartphone user base is estimated to reach 25.83 million users and therefore a new peak in 2029.
- **Hybrid Shopping pattern:** 80% of shoppers use a mix of online and in-store shopping, with only 13% exclusively online and 7% solely in-store. Customers often research products online before buying in-store.
- **Limited API availability among major retailers:** The absence of public-facing APIs on major supermarket websites makes it difficult to efficiently aggregate pricing and product information.



**Number of smartphone users in South Africa from 2014 to 2023 (in millions)

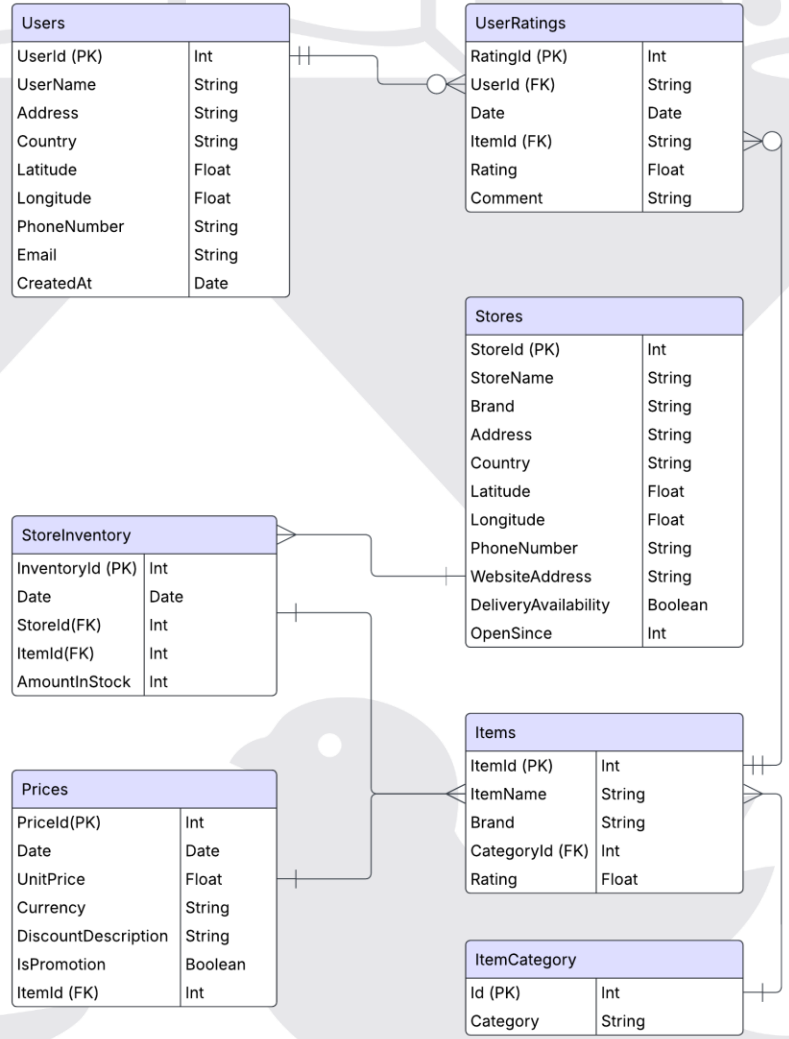
Source:
1. Statista 2025
2. SA Retail and E-commerce Market: Thriving in the Post-Covid Era

Database Design

The relational database is structured to efficiently store all necessary data to support key product features. An Entity-Relationship Diagram (ERD) was created to illustrate the relationships between different entities.

Data normalization: For example, by storing price records in a distinct Prices table, separate from product item data, the schema ensures that updates to product attributes do not overwrite or compromise historical pricing records. This approach preserves referential integrity and enables accurate longitudinal analysis of price trends.

Table Name	Features
Users & Ratings	enable tracking of user feedback and activity, supporting future features such as product reviews and platform engagement metrics. ✓ Product Ratings
Stores & StoreInventory	capture details about retailer locations and their availablestock, giving users the flexibility to select stores based on location and inventory. Store Information ✓ Product Availability ✓ Store Information
Items & ItemCategory	define the characteristics of each product (such asname, brand, and category), enabling efficient data aggregation and analysis. ✓ Product Alternatives
Prices	maintains up-to-date/historic information on prices, discounts, and promotions for each item, supporting accurate price comparison features. ✓ Price Comparison ✓ Sales Event



Local Market Research

Overview:

This local market research evaluates **top online retail platforms** within South Africa to identify the most reliable sources of mobile device pricing and product data. The goal is to build a robust foundation for accurate price comparisons, manual data collection, and future automation.

Evaluation Methodology:

The selection of key retail platforms was guided by the following criteria:

- **Market Share & Relevance:**
  Focus on platforms with extensive market coverage and high consumer engagement to ensure broad product and pricing representation.
- **Data Accessibility & Structure:**
  Prioritize platforms with well-organized, accessible, and consistently formatted product and price data, facilitating both manual and future automated data integration.
- **Reliability & Update Frequency:**
  Select sources based on the regularity and accuracy of their price and product updates, ensuring up-to-date and dependable information.
- **Reputation & Consumer Trust:**
  Consider the credibility and reputation of each platform to strengthen confidence in data quality and long-term solution adoption.

Local Market Research

Platform	Key Focus	Online Presence	Physical Stores	Key Strengths
Takealot.com	Everything/E-comm	Yes	No (online only)	Broadest selection, fast nationwide delivery
Checkers	Groceries, household, drinks, clothing, electronics, more	Yes (catalog only)	Yes	Wide physical reach, diverse selection, weekly specials
Pick n Pay	Groceries/general	Yes	Yes, 100s nationwide	Wide footprint, strong value proposition, hybrid shopping
Makro	Bulk shopping, electronics, appliances	Yes	Yes, 20+ warehouse clubs	Wholesale prices, large-item diversity, Walmart-backed logistics
Woolworths	Premium groceries, fashion, lifestyle	Yes	Yes, 100s nationwide	High quality standards, organic products, premium branding

Local Market Research

Platform	Data completeness	Data update frequency
Takealot.com	High – As a dedicated e-commerce platform, the website serves as the primary and most complete data source for products, pricing, and availability.	High – Information is frequently updated to reflect current stock, prices, and promotions in real time.
Checkers	Medium – Does not provide real-time inventory information.	Not specified/Unclear
Pick n Pay	High – Full product listings available online, clear categorization including electronics	Daily updates (pricing and promotions refreshed frequently; inventory synced reliably)
Makro	High – Detailed product descriptions, specifications, and images available; well-structured webpages	Daily updates (product data and pricing refreshed regularly to support bulk purchase model)
Woolworths	High – Rich product data with well-organized structure, especially for food and fashion items	Frequent updates (emphasis on freshness and stock accuracy for premium goods)

Strategic Prioritization of Data Sources

1. Pick n Pay
High data completeness and clear structure make it the top priority for foundational and diverse data collection.

2. Woolworths
Offers premium and well-structured product data, particularly valued for quality. Provide insights for mid-to-high-end consumer insights.

3. Makro
Ranked third due to its primary focus on wholesale rather than general retail, and smaller physical presence.

4. Takealot.com
Lower prioritization given its online-only nature and less direct integration with physical retail presence.

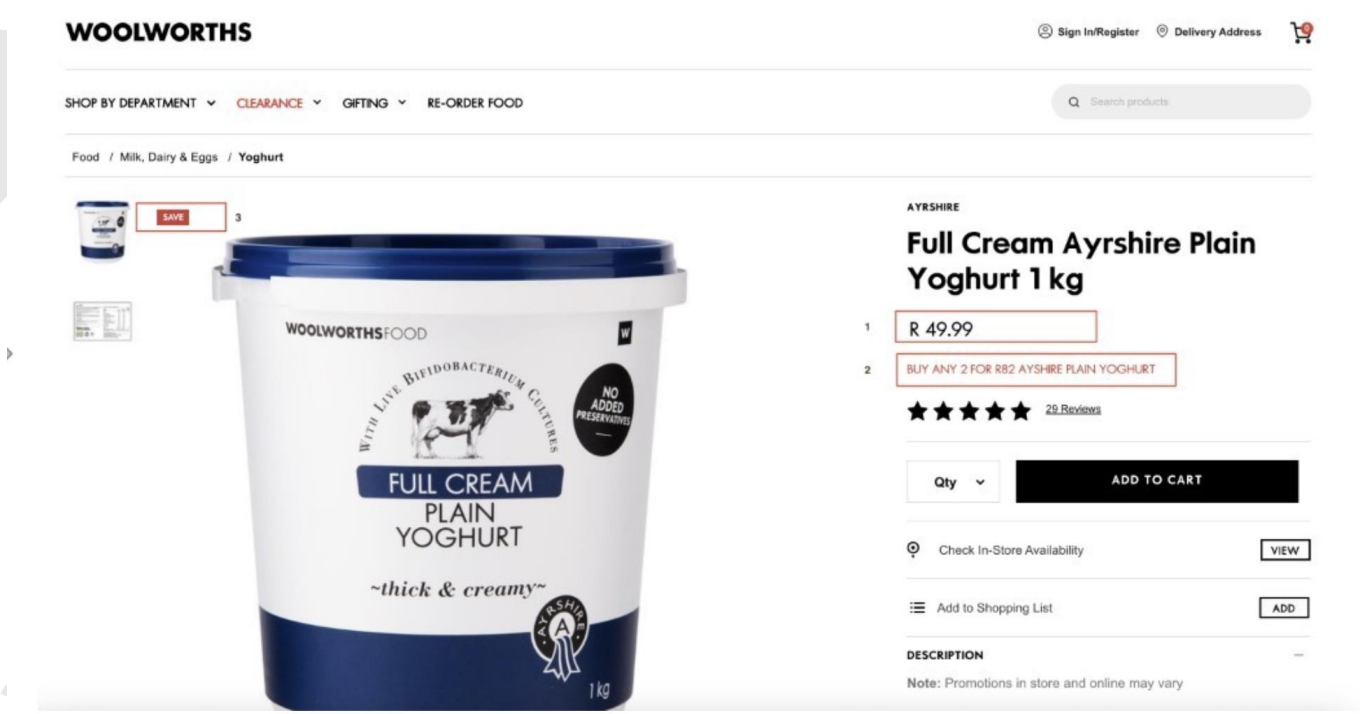
5. Checkers
Currently least prioritized due to limited visibility of local product inventory.

Data Retrieval Guide

The data retrieval guide outlines the essential fields mapped from retailer websites to the project's designated database, including precise instructions for navigating to each relevant page.

Designed for clarity and ease of use, the guide helps users quickly identify and extract key information needed for the database, streamlining future automation or web-scraping efforts.

By focusing on data mapping rather than rigid, site-specific scripts, the approach promotes long-term sustainability and reduces maintenance demands as retailor website designs change.



1.Prices.UnitPrice, Prices.Currency
2.Prices.DiscountDescription
3.Prices.IsPromotion

Database Prototype

The initial plan was to deploy a relational database on an AWS server provided by the client; however, technical challenges—including user access restrictions and IP whitelisting issues—prevented successful setup.

As an alternative, the prototype was developed on an **Azure SQL database**.

Manual data extraction from several platforms was carried out as a proof of concept.

A comprehensive DDL script will be supplied, allowing the client to easily replicate and deploy the database within their own environment.

```
1  select p.Date, p.UnitPrice,p.DiscountDescription,i.ItemName,i.Brand,i.Rating from dbo.Prices p
2  left join dbo.Items i on p.ItemId =i.ItemId
3  where i.ItemName like '%Jam%'
```

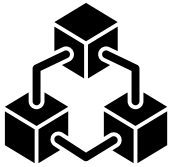
Results | Messages

Search to filter items...

Date	UnitPrice	DiscountDescription	ItemName	Brand	Rating
2024-06-25	239.95		Smooth Apricot Jam ...	Rhodes	4.7
2024-06-25	99.99	In-Store Deal: R84.99	Strawberry Jam 370 g	Bonne Maman	4.8
2024-06-25	44.99		Superfine Apricot Jam...	Rhodes	

The database is structured to enable efficient data aggregation and analysis. As shown in the screenshot, users can quickly query all pricing information for jam-related products and easily identify any available deals or promotions.

Sustainability Analysis



Technical Sustainability

Scalable and Flexible Database Design:
The database architecture separates core entities such as Items and Prices, ensuring each dataset is independently managed. This decoupling allows for easier data updates, streamlined maintenance, and supports seamless expansion as new features or data sources are introduced.

Maintainable Data Integration Approach:
Instead of relying on automated web scraping—which requires continuous code maintenance and frequent script updates when retailer websites change—the project provides a robust data mapping guide. This guide ensures that new data can be efficiently and accurately integrated, regardless of changes in website structure.

Future-Proofing Through API Readiness:
The system is designed to easily adopt public retailer APIs as they become available, providing opportunities for more stable, efficient, and scalable data integration in the future.

Financial Sustainability

Cost Control:
Minimize infrastructure costs by starting with cloud platforms that scale with usage (e.g., AWS free or low-cost tiers).

Build an Early User Base and Trust:
Establish credibility by providing consistently accurate, unbiased, and easily accessible price data. Offer free core features at launch, laying the groundwork for long-term user loyalty and future monetization opportunities.

Future Revenue Models:
Consider monetization through premium features, retailer partnerships for promoted listings, affiliate links, or data analytics services.

Monetization Strategies in Practice

Premium Features
Honey (by PayPal): Offers a free browser extension for coupon and price comparisons but has premium “Honey Gold” rewards and exclusive deals for registered/pay-to-use members.

Retailer Partnerships / Promoted Listings
Google Shopping: Allows retailers to pay for promoted or sponsored listings, giving their products higher visibility in search results.

Affiliate Links
PriceRunner: Provides price comparison, and earns commission when users click through and buy from partner retailers via affiliate links.

Data Analytics Services
SimilarWeb: Offers free website analytics but charges businesses for detailed traffic, competitor, and trend data.

Challenges and Limitations

Inconsistent Data Formats:

Product information and pricing structures differ substantially between retailors, adding complexity to data mapping and standardization. For instance, product categories are not uniform across platforms, requiring strategic decisions on how to align them.

Technical Constraints:

Encountered several technical hurdles, such as access denied errors when configuring the Aurora database on AWX server hosted by client, which delayed implementation. At the end, a prototype database was built on Azure.

Incomplete Data Coverage:

Some retailer platforms lack real-time inventory or comprehensive promotion explanations, limiting the accuracy of trend analyses. For example, Woolworths only provides general inventory levels rather than precise stock counts.

Sample Constraints:

The study's scope is limited to a subset of major retailors, which may not fully represent the entirety of South Africa's mobile device retail landscape.

Future Steps

Automated data retrieval through web-scraping
Implement robust web-scraping solutions for continuous and efficient price and product data collection from key retailers.

ETL pipeline
Develop an Extract, Transform, Load (ETL) pipeline to automate data cleaning, validation, and integration into the database.

Database scaling
Enhance database infrastructure to support increasing data volume, ensuring performance, reliability, and scalability as the platform grows.

API development
Create internal and external APIs for seamless data access, integration, and future mobile or web app expansion.

Advanced analytics and reporting
Introduce analytical tools and dashboards to generate insights on pricing trends, historical comparisons, and customer behavior.

Continuous market research
Monitor changes in the retail landscape and update data sources regularly to ensure ongoing accuracy and relevance of the platform.

McGill | DESAUTELS