

E-SYNAPSE

Empowering business owners with real-time, data-driven insights for smarter, faster decisions in the dynamic e-commerce landscape.

AGENDA

PART 1

Domain, Problem Statement, Why it Matters, Roadmap,
Initial Architecture, Data Streaming, Metrics

PART 2

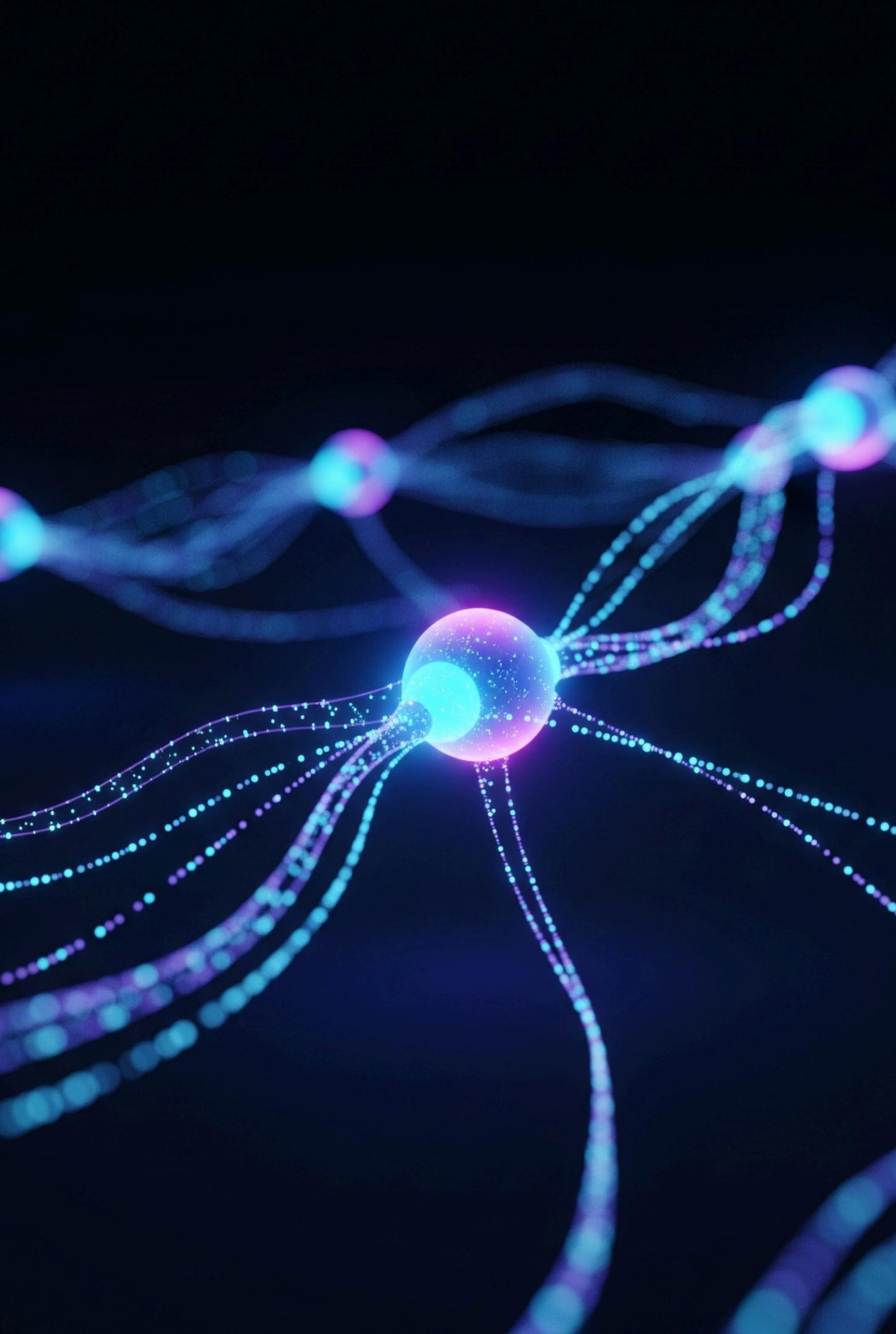
Changes Since Last Time & Synthetic Data

Implementation Steps

Key Insights

Final Answer to The Business Problem

Challenges and Next Steps



E-SYNAPSE

PART 1

Our Arena: The Pace of Modern Commerce

Today's customer decides in microseconds. Our project operates in this dynamic world of Retail and Digital Commerce, where real-time is the only relevant time.

We are focusing on D2C brands—businesses that have a direct line to their customers but face intense competition and must leverage every possible advantage.



Real-Time Shopper Behavior

For these brands, the most valuable data is happening now. We will capture live signals of customer intent, including:

- Cart Activity
- Wishlist Popularity
- Product Page Views



Strategic Optimization

By understanding this live behavior, we unlock the ability to fine-tune critical business levers with unprecedented agility:

- Pricing Strategies
- Ad Targeting
- Inventory Management

The Problem: Beyond Basic Reports

Traditional sales reports offer a rear-view mirror view. Business owners need a crystal ball, blending current performance with predictive analytics to stay ahead.



Limited Actionability

Current reporting often lacks the depth to identify underlying issues or hidden opportunities within the product catalog.

Which popular products are secretly hurting my brand?



Slow Decision-Making

Without real-time insights, businesses react rather than anticipate market shifts and customer trends.

How can I spot a new trend before my competitors do?



Disconnected Data

Sales, customer feedback, and behavioral data often reside in silos, preventing a holistic view.

Are my website views and sales figures telling the same story?

Why It Matters: From Data to Decisions

Our E-Synapse platform translates raw data into strategic actions that drive growth, efficiency, and customer loyalty.

Boost Revenue & Marketing ROI

We will move beyond just tracking sales to understanding what drives them. By answering, "How effectively do views convert to sales?", we can optimize marketing spend and pricing for maximum impact.

Protect Brand Reputation

Our platform acts as an early warning system. By answering, "Which popular products suffer from poor customer satisfaction?", we can instantly flag "high-risk" items and prevent long-term brand damage.

Enable Predictive Operations

We will use AI to look into the future. By answering, "What will our sales look like next week?", we can prevent stockouts, reduce overstocking, and optimize inventory across all locations.



Roadmap to Success

Our structured approach ensures a comprehensive and effective implementation of the AI-powered intelligence center.

Planning and Setup 5

- ✓ Brainstorm ideas and define domain, problem statement & why it matters
👤 Jul 7
- ✓ Explore Best Buy APIs, get API Key and run test calls
👤 Jul 8
- ✓ Create Architecture Diagram
👤 Jul 9
- ✓ Define KPIs
👤 Jul 9
- ✓ Design BigQuery Schemas
👤 Tomorrow

Streaming Pipeline 4

- ✓ Build Cloud Function to Pull API Data
👤 Yesterday
- ✓ Stream Events into BigQuery via Pub/Sub
👤 Monday
- ✓ Validate Data in BigQuery
👤 Monday
- ✓ Log Monitoring + Backup Plan
👤 Wednesday

dbt Modelling and Testi... 4

- ✓ Set Up dbt Project & BigQuery Connection
👤 Thursday
- ✓ Build Staging & Intermediate Models
👤 Thursday
- ✓ Implement Metrics in dbt Models
👤 Thursday
- ✓ Add dbt Tests + Automate via Scheduler
👤 Friday

Dashboard and Final Pr... 3

- ✓ Connect BigQuery to Tableau
👤 Jul 19
- ✓ Build Visuals: Trending, Drops, Reviews
👤 Jul 19
- ✓ Polish Dashboard: Filters, UX, Labels
👤 Jul 20

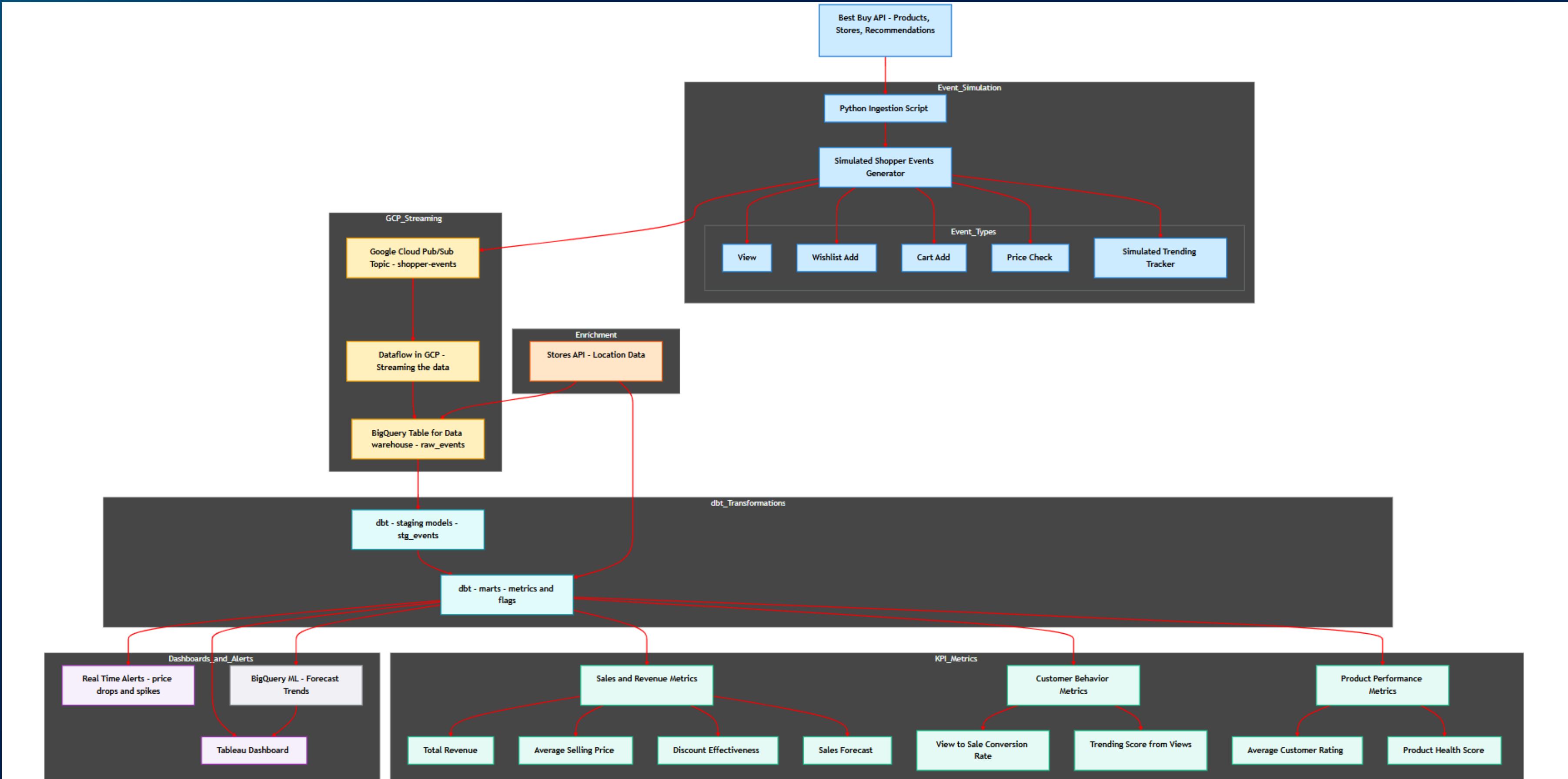
Presentations and Dead... 5

- ✓ Create Slides, Assign Presenter Roles and Rehearse for In-Class Business Presentation 1
👤 Today
- ✓ July 12th: In-Class Business Presentation 1
👤 Today
- ✓ July 25th: Submit Peer Review
👤 Jul 25
- ✓ Create Slides, Assign Presenter Roles and Rehearse for In-Class Business Presentation 2
👤 Jul 26
- ✓ July 26th: In-Class Business Presentation 2
👤 Jul 26

+ Add task

Initial High-Level Architecture Diagram

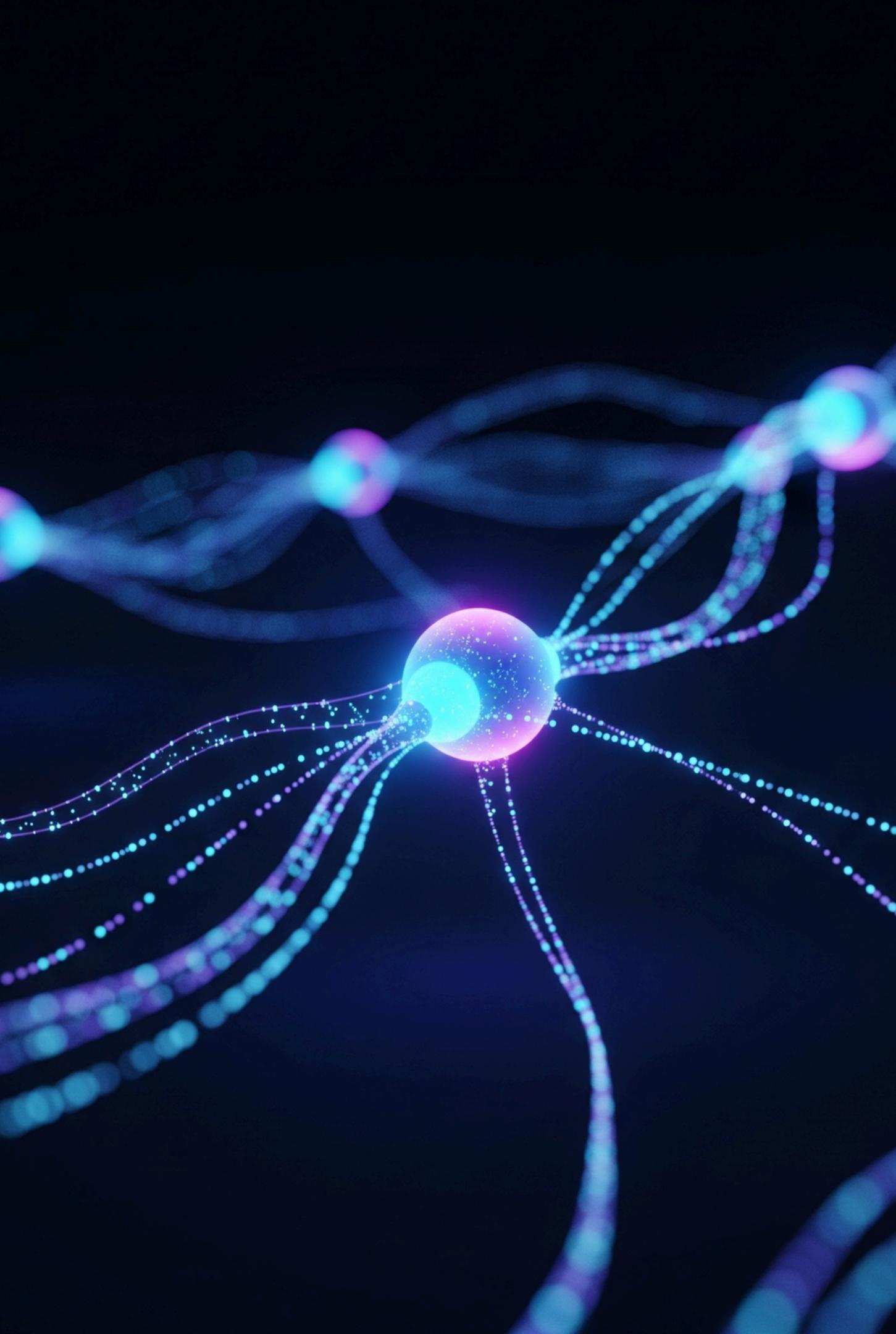
An end-to-end streaming data platform on Google Cloud, designed for scalability and real-time processing.



Key Metrics: Sales & Financial KPIs

Understanding the financial pulse of the business through precise calculations.

KPI / Metric	Required Attributes & Source API(s)	KPI / Metric	Required Attributes & Source API(s)
Total Revenue	salePrice from Products API + sales transaction counts.	Average Customer Rating	customerReviewAverage & customerReviewCount from Products API.
Average Selling Price (ASP)	salePrice from Products API.	Product Health Score	salePrice, customerReviewAverage, customerReviewCount (Products API); time-stamped sku values (Recommendations API).
Discount Effectiveness Ratio	salePrice, regularPrice, onSale status from Products API.	"View-to-Sale" Conversion Rate	Time-stamped sku values (Recommendations API - "Most Viewed" endpoint) & sales data.
Sales Forecast	Historical sales data (collected over time) and the BigQuery ML forecasting model.	Trending Score	Time-stamped sku values (Recommendations API - "Trending" endpoint).



E-SYNAPSE

PART 2

Changes Since Last Presentation

Data Source Change:

Instead of using the BestBuy API, we created synthetic data because API access was not available. This approach was designed to mimic actual user behavior including Add-to-Cart, Purchase and Return scenarios, historical data, varying product popularity, seasonality, etc.

Benefits of Synthetic Data:



Full Control over Data Volume & Frequency



No API Rate Limits or Outages



Easy to Simulate Returns and Stockouts



Ensures Privacy & Avoids Using Real Customer Data

Architecture Adjusted:

Pipeline components were updated to accommodate the synthetic dataset (high-level flow remains the same).

A Deeper Dive into the Synthetic Data

Built using these two Python Scripts:

- **Product Catalog:** 500 products, 18 categories, realistic prices & reviews (Python script).
- **Event Stream:** Simulated user actions (views, carts, purchases, returns) → streamed via Pub/Sub to BigQuery.

Behaviors that were Simulated:



Multi-Item
Browsing



Multiple Add
To Carts



Repeat
Customers



Seasonal
Traffic



Product
Performance



Stockouts

Implementation Steps

1

Ingestion: A Python script generates a live event stream that is ingested via Pub/Sub into a raw BigQuery table.

2

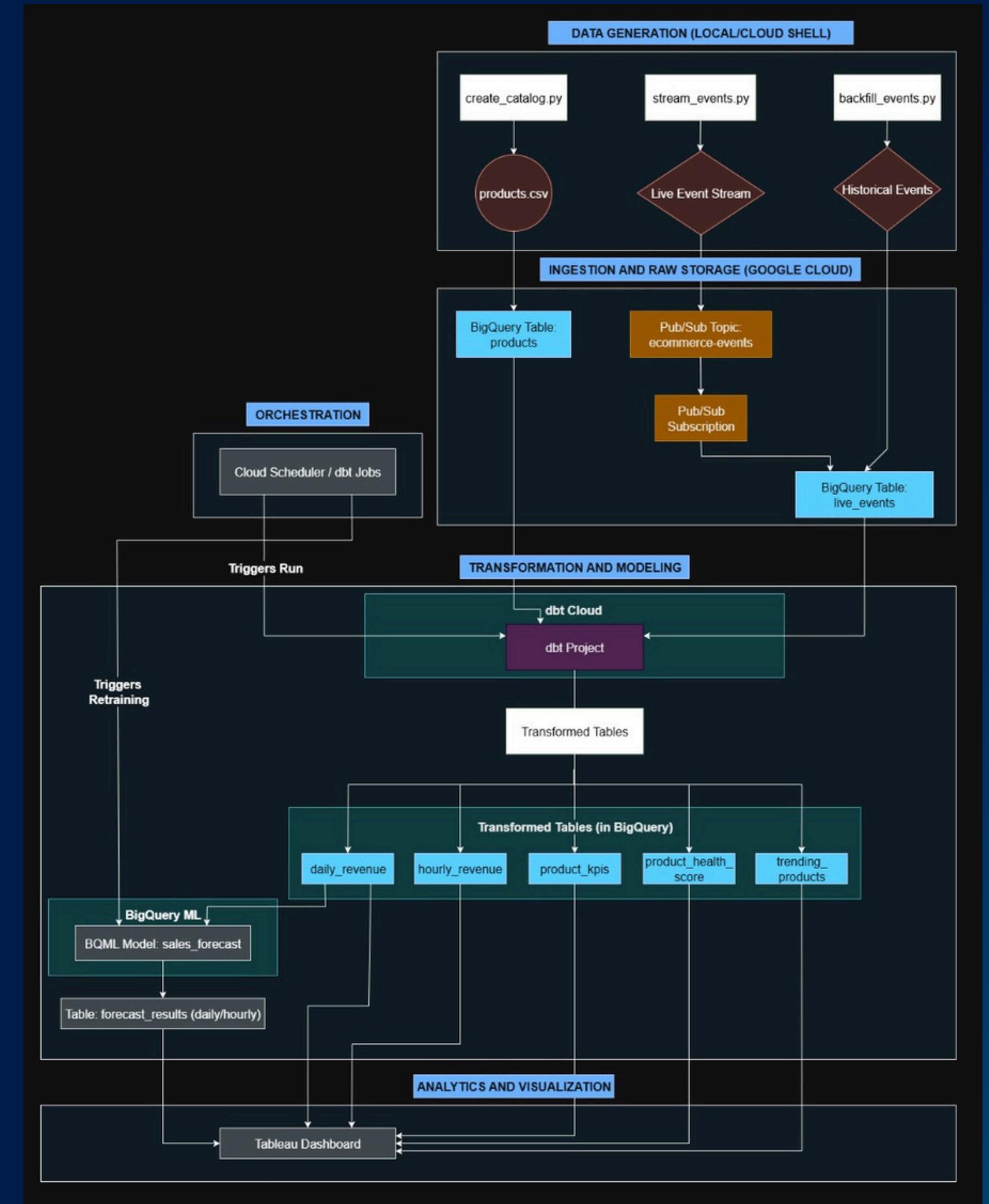
Transformation: A scheduled dbt Cloud job runs hourly, transforming the raw data into clean, aggregated tables (e.g., daily_revenue, product_kpis).

3

Predictive Analytics: BigQuery ML models are automatically retrained daily on the transformed dbt data to generate sales forecasts.

4

Visualization: Tableau connects to the clean dbt tables and BQML forecasts to provide an always up-to-date business dashboard.



New Architecture Diagram

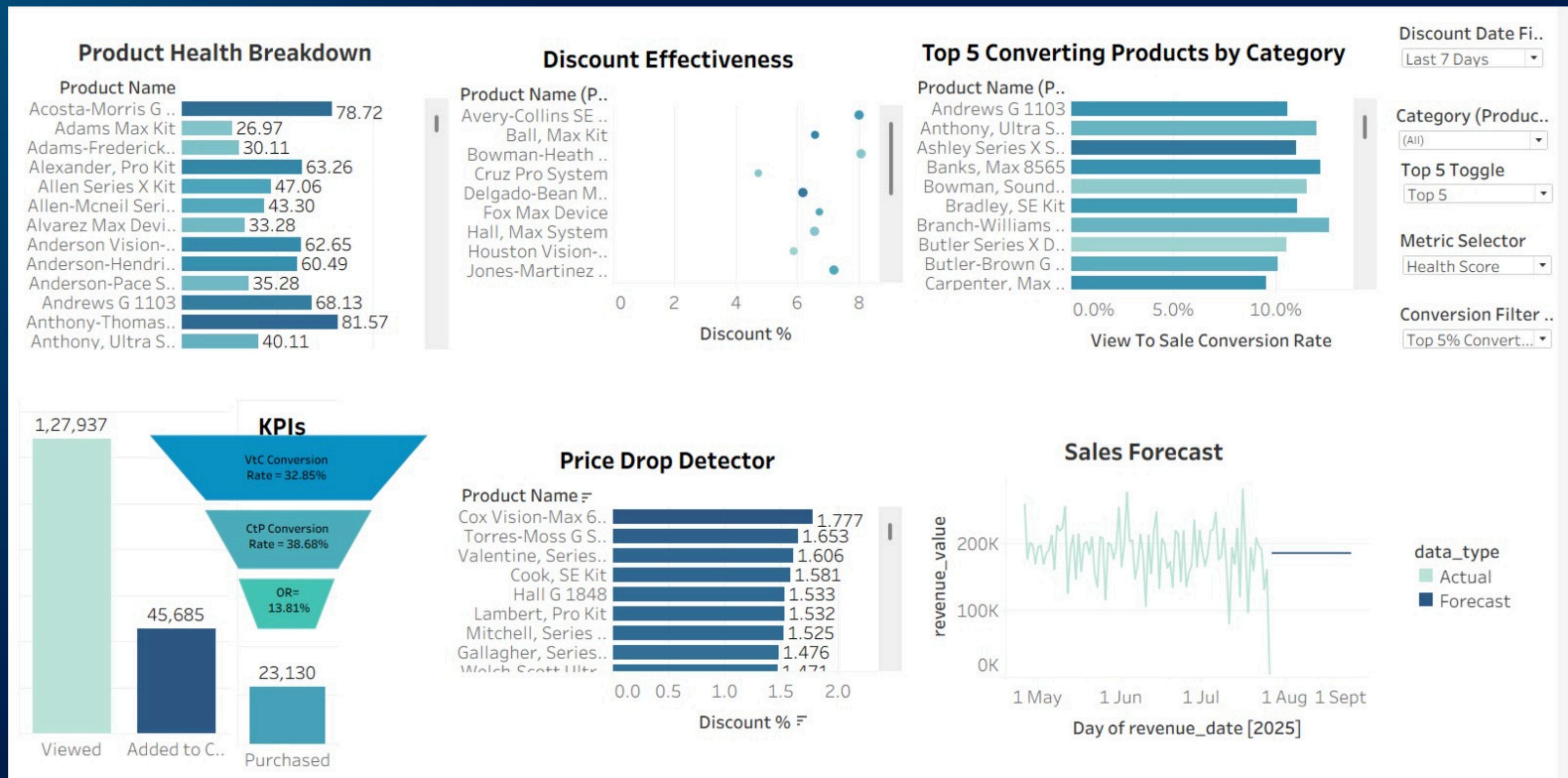
Demo

The screenshot shows the Google Cloud BigQuery interface with the following details:

- Project:** E-commerce Project
- Search Bar:** Search (/) for resources, docs, products, and more
- Explorer:** Shows the `ecommerce` dataset with two models:
 - `sales_forecast_30day` (selected)
 - `sales_forecast_7day`
- Model Details:** `sales_forecast_30day`
 - Model type:** ARIMA Plus
 - Model ID:** e-commerce-project-467004.ecommerce.sales_forecast_30day
 - Date created:** Jul 26, 2025, 12:00:16 AM UTC-7
 - Model location:** US
- Model Details (continued):**
 - Date modified:** Jul 26, 2025, 12:00:17 AM UTC-7
 - Model expiration:** Never
 - Description:** -
- Job history:** Refreshed
- CLOUD SHELL:**
 - Terminal (e-commerce-project-467004)
 - Output:

```
--- PUBLISHING NEW SESSION: session-106476e4-1b46-48d3-b223-146024ca0ed2 (1 events) ---
--- PUBLISHING NEW SESSION: session-7e88d3aa-5dd-4687-9186-556d42d860a8 (9 events) ---
--- PUBLISHING NEW SESSION: session-306ac915-ec68-4eaa-950b-4eaeaa7d29d6 (7 events) ---
--- PUBLISHING NEW SESSION: session-25274398-7584-45e0-a98f-7b1ddff77e44 (3 events) ---
--- PUBLISHING NEW SESSION: session-6157eff0-41b2-4757-ac19-758b6e387545 (6 events) ---
--- PUBLISHING NEW SESSION: session-a4d74520-2f80-472f-9dde-60a26043ca41 (13 events) ---
--- PUBLISHING NEW SESSION: session-68afb19f-5b84-4d7c-a835-6de26398da28 (3 events) ---
--- PUBLISHING NEW SESSION: session-1464ele9-f5de-4d0e-ae61-c9e95f731de1 (5 events) ---
--- PUBLISHING NEW SESSION: session-cf69893f-cb8d-4524-aa49-2a82fdb2259 (8 events) ---
--- PUBLISHING NEW SESSION: session-54575086-b9e8-4154-b4d1-6fc73a264d9d (3 events) ---
--- PUBLISHING NEW SESSION: session-db988354-721f-45b6-b75e-8b447e4d8178 (1 events) ---
--- PUBLISHING NEW SESSION: session-dad076bd-1443-4064-ab3d-9172e77a6295 (4 events) ---
--- PUBLISHING NEW SESSION: session-89dac879-de75-4acd-aabe-d9c2962dacce (13 events) ---
^C
Stream stopped by user.
```

Key Insights



Final Answer to the Business Problem

To sum it all up, here are the answers to the business questions:

1

How effectively do views convert to sales?

About 32.8% of views lead to cart additions, but only 13.6% result in purchases, showing significant drop-offs after adding to cart.

2

Which popular products suffer from poor satisfaction?

Several high-selling items have ratings below 4, indicating they drive sales but risk damaging brand reputation.

3

What will our sales look like next week?

Using historical patterns and current trends, moderate discounts (5–7%) and highly rated products are expected to drive steady sales growth if maintained.



Challenges Faced

- **API Limitation:** No BestBuy API access → switch to synthetic data.
- **Data Realism Effort:** Designing realistic behavior (returns, seasonal trends, cart abandonment) was time-intensive.
- **Pipeline Adjustments:** Architecture had to be revised for the new source.
- **Time Constraint:** Limited time meant prioritizing core pipeline and dashboard features over building advanced models, richer synthetic behaviors, and robust quality checks.

Next Steps

- **Richer Data Patterns:** Enhance synthetic data with more complex user journeys and seasonal patterns.
- **Conversion Optimization:** A/B test pricing and promotions to reduce cart abandonment.
- **Enhanced Analytics:** Add new KPIs to the dashboard for revenue impact and cohort behavior.
- **API Integration:** Connect real APIs when available to validate findings.

THANK YOU