MDB Project History & Transformation Overview

Project Name:

MDB

Client / Stakeholder:

Million Dollar Baby

Overview & Background

The MDB journey began in **June 2022** with a modest start—a single Power BI resource and a growing need for better dashboards and actionable insights. At the time, the existing dashboards weren't delivering what stakeholders needed, and much of the focus was on web scraping to pull relevant data.

As we dug deeper, it became clear that a more unified, scalable approach was required—not just patchwork dashboards or fragmented data sources. What started as a dashboard enhancement project quickly evolved into a comprehensive data transformation initiative.

Business Objective

- Build a robust, unified data infrastructure
- Improve BI effectiveness and dashboard usability
- Enable predictive analytics through machine learning
- Automate manual, repetitive processes
- Enhance decision-making across departments (especially marketing and operations)

Timeline:

- **Start Date:** June 2022 (on schedule)
- **Current Status:** Ongoing (key focus areas now in 2025)

Initial Scope & Early Challenges

Initially, we had only one Power BI resource, while some effort was directed toward web scraping. It quickly became clear that dashboard requirements were vague and fragmented, limiting their usefulness.

To address this, we expanded our scope. We began R&D on possible unified platform solutions and proposed a modern data architecture, including **Azure** and **Databricks**. Though we suggested a cloud-native structure, the client preferred a **unified hybrid platform**, which influenced how we built out the system.

Team Evolution

- Started with 2 team members: 1 Power BI +1 Web Scraping resource
- Scaled to 7+ over time with Data Engineers, ML experts, and BI specialists
- Onboarded senior engineers to manage complex data workloads
- BI resource joined during the second phase to strengthen analytics delivery

Key Milestones

- **Q3 2022:** Data engineering team expanded
- **Q4 2022:** First ML initiative kicked off – Demand Forecasting
- **Q1 2023:** Migration to Sigma Computing began
- **2023 Mid-Year:** ML models operational
- **2025 Plan:** Focused on Savant migration & UCP consolidation; further team scaling with an additional React Js resource

Technology Stack

- **Cloud & Data:** Azure, Databricks, Lakehouse architecture

- **BI & Dashboards:** Power BI, Sigma Computing
- **Languages & Frameworks:** Python, PySpark, SQL, Spark SQL
- **Automation & ML:** Selenium, MLOps, LLMs
- **Platform:** SAVANT (DE), 30+ integrated APIs, 50+ tables

Highlights & Achievements

Machine Learning & AI:

- Developed **demand forecasting models** that directly solved out-of-stock inventory issues
- Built **custom chatbots** using LLMs for customer and internal support
- Launched **sentiment analysis pipelines** to extract and act on customer feedback
- Created a **Price Elasticity of Demand (PED)** model for the marketing team to optimize discount strategies

Data Engineering:

- Introduced a modern **Lakehouse** architecture
- Integrated over **30+ APIs** to reduce manual effort
- Automated **200+ workflows**, streamlining data operations
- Migrated **25% of SAVANT** data in under a month, covering 20+ endpoints and 50+ tables

Business Intelligence:

- Rolled out 20+ high-impact Power BI dashboards in the first quarter
- Successfully migrated all dashboards from Power BI to Sigma
- Leveraged Sigma's UI/UX to make dashboards more intuitive and user-friendly

Key Deliverables

- Unified data architecture
- ML-powered insights (demand forecasting, sentiment analysis, pricing strategy)
- Seamless transition from Power BI to Sigma
- Real-time, automated data pipelines
- Easy-to-use, insightful dashboards for end users

Challenges & How We Solved Them

Challenge	Resolution
Poor dashboard quality	Redesigned from scratch with BI experts
Disconnected data sources	Unified via Azure, Databricks, and Lakehouse
Manual, repetitive data tasks	Automated with workflows and API integrations
Lack of prediction capabilities	Introduced ML models for forecasting & pricing
UI/UX friction in old dashboards	Migrated to Sigma for smoother user experience

2025 Focus: Unified Customer Profile (UCP) & Strategic Expansion

As we move into 2025, a major focus area for the team is building out the **Unified Customer Profile (UCP)**—a strategic initiative that aims to consolidate all customer-related data into a single, accessible, and intelligent profile. This effort is foundational to unlocking deeper personalization, better customer engagement, and smarter decision-making across the organization.

Why UCP?

Over the past two years, the data we've collected has grown exponentially—from transactional logs to support tickets, user feedback, behavioral analytics, and more. However, this data has

often remained siloed across platforms and teams. The UCP initiative is about bringing all of this together in one cohesive structure.

We're not just aggregating customer data—we're enriching it. UCP will integrate inputs from multiple systems:

- Web & app behavior
- Purchase history
- Customer support interactions
- Survey feedback and reviews
- ML-driven sentiment scores
- Loyalty program data
- Demographic and firmographic overlays

This unified view will be made accessible to marketing, sales, support, and strategy teams via dynamic dashboards and real-time APIs.

Our Approach

To make this vision a reality, we're building on the scalable foundation of Azure and Databricks, leveraging:

- **Lakehouse architecture** for centralized, performant data storage
- **Sigma** for user-facing dashboards with deep filters and segmentation
- **ML models** to predict customer intent, churn likelihood, and product affinity
- **LLMs** to auto-generate customer summaries and surface key behavioral insights
- **MLOps pipelines** to automate and continuously improve data enrichment and scoring

Early Wins

Although the UCP is in early phases, we've already:

- Defined over 50 key customer attributes

- Mapped and integrated ~70% of the needed data sources
- Created early prototypes for real-time customer summary cards
- Begun testing lead scoring and product recommendation models with the marketing team

What It Enables

- **Sales** can now prioritize leads using ML-driven intent scoring
- **Marketing** can run hyper-targeted campaigns using real-time behavioral segments
- **Customer Support** gets a complete view of the customer journey, improving service quality
- **Leadership** gets better forecasting based on customer trends and engagement signals

Strategic Vision

The UCP is more than just a data integration project—it's the backbone of a **customer-first, insight-led strategy**. With it, MDB is moving toward a future where every customer interaction is informed, timely, and personalized.

Client Feedback & User Response

The feedback has been overwhelmingly positive. Users found Sigma dashboards significantly easier to navigate, while the marketing team reported actionable gains from the PED model. BI tools are now faster, cleaner, and far more useful.

MDB team went on an trip to Qatar as a token of appreciation from the client in 2024.