Data Sleek
Streamline Your Data into Business Value

CASE STUDY

Modernizing Analytics at Johns Hopkins University: dbt and Dimensional Modeling for Development and Alumni Relations

Modernizing Development and Alumni Relations Analytics with dbt and Dimensional Modeling

Johns Hopkins University's Development and Alumni Relations (DAR) team supports fundraising, campaign management, and constituent engagement across a highly complex institutional environment. As reporting needs grew, DAR required a modern analytics foundation that could deliver trusted metrics, support future self-service BI, and scale with evolving data demands.

Challenge:

DAR's primary data source was Salesforce, configured for operational workflows rather than analytics. While effective for gift entry and donor management, the source data posed major challenges for reporting and decision-making:

- Complex, denormalized data structures with overlapping concepts
- Ambiguous grains across payments, gifts, credits, and campaigns
- Unstable natural keys and heavy reliance on Salesforce-generated IDs
- Inconsistent metric definitions across teams and dashboards
- Limited trust in analytics outputs due to reconciliation issues

Without a governed analytics layer, reports risked double-counting gifts, misattributing credit, and producing conflicting totals, an especially serious issue in fundraising and donor reporting. At an institutional level, these inconsistencies undermined leadership confidence in reported outcomes and introduced unnecessary risk into campaign performance tracking and donor accountability.

Solution

The objective was to establish a trusted, scalable analytics foundation that delivered consistent metrics and supported future self-service BI through governed transformations and clear data grains.

1. Layered Analytics Architecture

We enforced a clear, scalable modeling structure using dbt:

- Staging models to standardize raw Salesforce data
- Dimension tables representing stable business entities (constituents, campaigns, designations, dates)
- Fact tables capturing transactional events at explicitly documented grains

All business logic was centralized in SQL transformations and never embedded in BI tools.

2. Dimensional Modeling and Key Management

To resolve grain ambiguity and unstable identifiers:

- Deterministic surrogate keys were generated for all dimensions
- Natural keys were retained for traceability but excluded from joins
- Every fact table clearly documented its unit of analysis
- Relationships between facts and dimensions were enforced with automated tests

This approach supported multiple analytical perspectives, payment-level, credit-level, and campaign-level, without conflating them.

3. Automated Data Quality and Governance

Data quality was embedded directly into the pipeline:

- not_null and unique tests on primary keys
- Relationship tests between facts and dimensions
- Layer-specific YAML files for maintainable testing at scale

As Salesforce schemas changed, dbt tests surfaced issues early, preventing them from reaching dashboards.

4. Business-Friendly Documentation

To bridge the gap between technical and non-technical stakeholders:

- dbt documentation was written in plain business language
- Column descriptions explained the meaning, usage, and limitations
- Dimensional ERDs were generated from the analytics model itself

Documentation became a shared source of truth across teams.

Results

1. Technical Outcomes

- A modular, maintainable dbt project aligned with industry best practices
- Simplified SQL and reduced logic duplication in downstream reporting
- Faster development through reusable, conforming dimensions
- Improved change management via version control and automated testing

2. Business Outcomes

- Consistent fundraising metrics across dashboards and reports
- Clear, shared definitions for gifts, credits, campaigns, and constituents
- Reduced reconciliation effort and fewer ad-hoc clarification requests
- Greater confidence and trust in analytics outputs
- Reduced report delivery time from 24+ hours to under 1 hour, a 24× improvement in reporting speed

3. Organizational Impact

- Analytics assets treated as products, not one-off reports
- Governance embedded directly in the data pipeline
- Analytics teams shifted from report builders to data stewards

This foundation positioned DAR for sustainable growth in analytics and future self-service BI.

Technology Stack

- **Analytics Engineering:** dbt Core / dbt Cloud
- **Data Warehouse:** Azure Synapse Analytics (Dedicated SQL Pool)
- **Source Systems:** Salesforce (Development and Alumni Relations)
- **BI and Reporting: Power BI, Tableau**
- **Governance Artifacts:** Dimensional models, ERDs, data dictionaries, automated tests

Why This Matters for Higher Education

Higher education institutions face unique analytics challenges: legacy CRMs, complex stakeholder needs, and high expectations for accuracy and auditability. This engagement demonstrates how dbt and dimensional modeling can:

- Translate operational systems into analytics-ready data products
- Support advancement, alumni engagement, and institutional reporting
- Scale governance without slowing innovation

Rather than requiring a full system replacement, this approach modernizes analytics incrementally and pragmatically, a critical advantage in higher-education environments.

About Data-Sleek

Data-Sleek® is a consulting agency and a leading provider of comprehensive data management solutions. We specialize in data transformation, data migration, data integration, data architecture, data warehousing, and database consulting. Our commitment to client satisfaction empowers organizations of all sizes to unlock the full potential of their data assets. Data-Sleek is a trusted partner of Fortune 500 companies, dedicated to innovating and revolutionizing the way businesses operate.