© Journal of the Technical University - Sofia
Plovdiv branch, Bulgaria
"Fundamental Sciences and Applications"
Vol. 32 2026

SQL Database Optimization as a Driver for SME Digital Transformation: A Case Study in the PV Sector †

Elena Gurova [1]

[1] Department of Information Technologies in Industry, Faculty of Computer Systems and Technologies, Technical University of Sofia, 8 Kliment Ohridski blvd., 1000 Sofia, Bulgaria;
* Correspondence: elenakr.gurova@gmail.com

Abstract: Small and Medium-sized Enterprises (SMEs) in the photovoltaic sector face significant challenges in managing and leveraging data for strategic decision-making. This case study examines a comprehensive database optimization initiative within a European PV company, focusing on the digitalization of sales processes through Databricks and Power BI integration. The research documents a practical methodology for data consolidation, cleaning, semantic layer design, and installer segmentation that enabled real-time analytics and evidence-based decision-making. Results demonstrate improvements in data quality (67% to 94% completeness), reporting efficiency (days to real-time), and measurable business outcomes including resource reallocation and customer win-back campaigns.

Keywords: Database Optimization; SME Digital Transformation; Databricks; Power BI; Installer Segmentation; PV Commissioning; Data Integration; Sales Process Digitalization

1. Introduction

The photovoltaic (PV) industry has experienced rapid growth across Europe, creating both opportunities and operational challenges for Small and Medium-sized Enterprises (SMEs) operating within this sector. As these organizations scale from regional to multi-country operations, their data management capabilities often fail to keep pace with business growth. Customer records become fragmented across disparate systems, sales processes diverge by country, and management visibility into operational performance deteriorates—precisely when data-driven insights become most critical for competitive positioning.

Digital transformation offers a pathway to address these challenges [1], yet SMEs face unique constraints that distinguish their transformation journey from that of large enterprises [2]. Limited IT budgets, scarce technical expertise, and the need for rapid return on investment demand practical, scalable approaches rather than enterprise-grade solutions that require extensive implementation timelines and specialized personnel [5].

This paper presents a case study of database optimization within a European PV company, demonstrating how the strategic combination of Databricks data platform capabilities [4] with Microsoft Power Platform tools (Power BI, Power Apps, Power Automate) can enable meaningful digital transformation within SME resource constraints [5]. The implementation focuses on creating live-updated analytical views that integrate data from multiple source systems—including CRM databases, mobile commissioning

Online ISSN 2603-459X

applications, and geographic reference data—to support real-time sales reporting and customer segmentation.

The contribution of this work is threefold: (1) a documented methodology for data integration and semantic layer design appropriate for SME contexts; (2) a practical demonstration of Databricks and Power BI architecture for sales process digitalization; and (3) empirical evidence of business outcomes achieved through structured data optimization, including measurable improvements in reporting efficiency, decision-making quality, and sales team productivity.

The paper builds upon theoretical foundations established in prior work examining SME digitalization frameworks [1], extending from conceptual methodology to concrete implementation within a real-world PV sector organization [8].

1.1 Database Challenges in SMEs

Small and Medium-sized Enterprises face distinct challenges when attempting to leverage data for competitive advantage [2]. Unlike large enterprises with dedicated data engineering teams and established data governance frameworks, SMEs typically operate with limited technical resources, fragmented data sources, and minimal formal data management processes [1]. In the photovoltaic sector, these challenges are amplified by the rapid growth of the industry, geographical dispersion of operations, and the need to track complex customer relationships across multiple dimensions including installation sites, service contracts, and equipment lifecycles.

The problem is not merely one of data volume, but of data fragmentation and accessibility. Sales teams across different European countries often maintain customer information in disparate systems—spreadsheets, local databases, CRM systems with inconsistent implementations, and even paper records. This fragmentation creates several critical business problems:

Inconsistent Reporting: Different countries generate reports using different methodologies, making cross-country comparisons unreliable

Delayed Decision-Making: Manual data consolidation processes can take days or weeks, by which time market conditions may have changed

Limited Visibility: Management lacks real-time insight into customer account status, sales pipeline health, or emerging market trends

Resource Inefficiency: Sales teams spend significant time on data gathering rather than customer engagement

Database optimization in the SME context goes beyond technical performance improvements. It represents a strategic capability that enables standardization, transparency, and data-driven decision-making—fundamental prerequisites for digital transformation [6].

1.2 Research Context and Objectives

This case study is situated within a larger research program examining digital transformation in SMEs through the Microsoft 365 ecosystem [1]. While the broader research explores multiple dimensions of digitalization including collaboration tools, process automation, and change management [12], this article focuses specifically on the data foundation layer—the databases, data architectures, and semantic models that enable downstream business intelligence and analytics capabilities [5].

The research question addressed in this case study is: How can database optimization and structured data architectures enable digital transformation in resource-constrained SME environments, particularly within the photovoltaic sector?

The specific objectives are:

To document a practical methodology for assessing, cleaning, and structuring sales data in a multi-country PV organization

To demonstrate the implementation of a modern data architecture (Databricks + Power BI) suitable for SME constraints and capabilities

To evaluate the business outcomes achieved through database optimization, including improvements in reporting accuracy, decision-making speed, and sales process standardization

To identify lessons learned and scalability considerations applicable to similar SME digitalization initiatives

This case study draws on a real-world implementation project conducted within a European PV company operating across multiple countries [8]. The project aimed to standardize and digitalize sales operations by creating a unified data infrastructure that could support consistent reporting, customer segmentation, and sales performance monitoring [6].

2. Methodology

2.1 Research Approach and Data Collection

This case study employs a participatory action research methodology [10], where the researcher is actively involved in the design and implementation of the database optimization initiative while simultaneously collecting data on the process and outcomes. This approach follows established case study research principles [10] and is particularly appropriate for SME digitalization research, as it enables deep understanding of organizational constraints, technical challenges, and change management dynamics that may not be visible to external observers [11].

Data Collection Methods:

Document Analysis: Review of existing database schemas, data dictionaries, business process documentation, and sales reporting templates across European countries

Stakeholder Interviews: Semi-structured interviews with sales managers, inside sales teams, account owners, and business analysts to understand data needs and current pain points

Technical Assessments: Direct analysis of database structures, data quality metrics, query performance logs, and existing reporting systems

Implementation Logs: Systematic documentation of design decisions, technical configurations, data transformations, and iterative refinements throughout the project

Performance Metrics: Collection of quantitative metrics before and after implementation, including report generation times, data accuracy rates, and user adoption metrics

Research Setting:

The case study organization is a mid-sized European photovoltaic company with operations in multiple countries. The company designs, sells, and services solar energy installations for commercial and residential customers. The sales organization employs approximately 50-75 sales professionals across Europe, organized into country-specific teams with regional coordination.

2.2 Database Structure Analysis Framework

The initial assessment phase employed a multi-dimensional framework to understand the current state of sales data management [3]:

Dimension 1: Data Source Inventory

Identification of all systems containing sales-relevant data (CRM systems, spreadsheets, local databases, ERP systems)

Mapping of data ownership and update responsibilities

Assessment of data redundancy and conflict resolution processes

Dimension 2: Data Quality Assessment

Completeness: Percentage of required fields populated

Accuracy: Validation of data against known ground truth sources

Consistency: Cross-system comparison of shared entities (customer names, addresses, contact information)

Timeliness: Lag between business events and data updates

Dimension 3: Business Process Mapping

Documentation of how sales data is currently created, updated, and consumed

Identification of manual data entry points and handoff processes

Analysis of reporting workflows and decision-making cycles

Dimension 4: User Needs Analysis

Segmentation of sales roles and their specific data requirements

Identification of key questions that data should answer (e.g., "Which accounts are at risk?" "Where should we focus new business development?")

Understanding of preferred data consumption formats (dashboards, reports, alerts)

2.3 Optimization Techniques and Technical Architecture

The optimization methodology centered on creating a "single source of truth" data architecture while accommodating the distributed nature of sales operations [3].

Technical Solution Architecture:

The chosen architecture separates concerns between data processing/storage (backend) and data consumption (frontend):

Backend: Databricks Data Platform

Purpose: Centralized data warehouse for all sales-related data from multiple source systems

Capabilities:

Data ingestion from heterogeneous sources (SQL databases, Excel files, API connections)

Data cleaning and transformation using SQL and Python

Creation of semantic layers and business logic definitions

Maintenance of live, continuously updated tables

Rationale: Databricks provides enterprise-grade data capabilities with the flexibility needed for evolving SME requirements, without requiring on-premises infrastructure [4]

The architecture integrates three primary data sources: (1) CREdge, a CRM system providing customer and installer account master data with near real-time sync; (2) SetApp Server, a mobile commissioning application capturing PV installation and commissioning records in real-time from field installers; and (3) GEO Mapping reference tables linking postal codes to sales areas and assigned sales managers, updated quarterly. The detailed configuration of each source is described in Section 3.

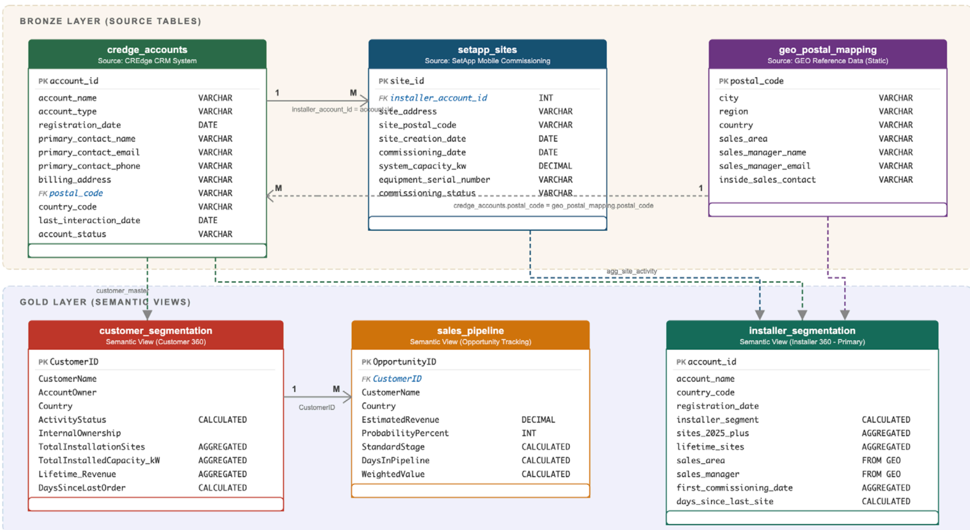

Figure 1 presents the relational model of the database architecture, illustrating the entities, their attributes, primary and foreign key relationships, and the data flow from source tables through the medallion architecture (Bronze, Silver, Gold layers) to the semantic views consumed by Power BI dashboards embedded in Microsoft Teams.
This medallion architecture (Bronze → Silver → Gold) enables progressive data quality improvement while maintaining traceability back to source systems [4]. The frontend layer uses Power BI dashboards with real-time Databricks connections, providing interactive visualization accessible to non-technical sales professionals within Microsoft Teams.

A key architectural decision was establishing clear separation of responsibilities: sales teams interact exclusively with Power BI dashboards, the project lead manages backend data architecture in Databricks, and the IT department provides infrastructure support. This separation enables rapid iteration on data models without disrupting end-user workflows.

2.4 Customer Segmentation Strategy

A core component of the optimization methodology was developing a coherent customer segmentation framework that could drive both data structuring decisions and sales process standardization [6].

Primary Segmentation: Customer Activity Status

Customers were categorized based on their current relationship with the organization:

Lost Accounts: Previously active customers who have not placed orders or engaged with the company within a defined period (e.g., 18-24 months)

Business Objective: Win-back campaigns, churn analysis, competitive intelligence

Data Requirements: Historical revenue, last interaction date, reasons for disengagement, competitive movements

New Accounts: Prospects and recently acquired customers (e.g., first order within last 12 months)

Business Objective: Onboarding optimization, early lifecycle engagement, expansion opportunity identification

Data Requirements: Lead source, acquisition cost, initial purchase details, onboarding status

Active Accounts: Current customers with ongoing relationships

Business Objective: Revenue growth, service quality, upsell/cross-sell opportunities

Data Requirements: Installation sites, equipment inventory, service history, revenue trends, satisfaction metrics

Secondary Segmentation: Internal Ownership

Within each activity category, accounts were further classified by internal ownership structure:

Account Owner: Primary relationship manager responsible for strategic account planning

Inside Sales Manager: Coordinates remote sales activities, qualification, and pipeline development

Sales Manager: Oversees regional or product-specific sales teams

This dual-layered segmentation enabled both strategic analysis (focusing on customer lifecycle stages) and operational reporting (tracking individual and team performance).

Tertiary Segmentation: Installer Commissioning Activity (Primary Use Case)

For the PV sector specifically, a critical segmentation dimension emerged around installer commissioning behavior [8]. Installers represent B2B partners who purchase equipment and install PV systems at end-customer sites. The installer segmentation framework categorizes accounts based on the relationship between account registration date and first commissioning activity:

The framework defines four categories: (1) New Account 2025 – installers who registered and began commissioning in the current strategic year, representing organic market growth; (2) Conversion from Old Account – installers registered before 2025 with new commissioning activity, indicating reactivation or market shift; (3) Dormant Account – installers registered before 2025 with historical but no current-year commissioning, signaling potential competitive displacement; and (4) Pending First Commissioning – recently registered installers with no commissioning activity yet, requiring activation support. The complete SQL implementation and data logic for this segmentation is provided in Appendix A.

This commissioning-based segmentation proved particularly valuable for the European sales organization because it directly measures market activity rather than merely tracking CRM interactions or revenue [9]. An installer with active commissioning is demonstrably engaged in the market, regardless of their direct purchasing pattern with the organization.

3. Case Study: European Sales Process Digitalization

This section presents the practical implementation of the methodology within a European PV company operating across five countries: Germany, France, Spain, Bulgaria, and Poland. The organization, classified as a medium-sized enterprise with approximately 200 employees, specializes in photovoltaic system components and installation services, maintaining relationships with thousands of installer accounts across its geographic footprint.

The digitalization initiative emerged from strategic recognition that fragmented data systems were constraining growth and limiting management's ability to make evidence-based decisions [6]. Despite successful country-level operations, the organization lacked unified visibility into customer relationships, sales pipeline health, and installer commissioning activity—information critical for resource allocation, market prioritization, and competitive positioning.

The case study documents the journey from initial state assessment through technical implementation to measurable business outcomes, with particular emphasis on the installer segmentation capability that became the flagship use case for demonstrating Databricks-to-Power-BI-to-Teams integration.

3.1 Project Overview and Objectives

The project emerged from a strategic initiative to standardize and digitalize sales operations across the European organization. The executive sponsor framed the challenge as: "We have successful sales teams in each country, but we operate as separate entities rather than a unified European organization. We lack visibility into what's working where, and we can't easily share best practices or reallocate resources based on data."

Core Project Objectives:

Identify: Pinpoint key business processes within the sales organization that should operate consistently across countries

Align: Standardize the expected outcomes and performance metrics for these processes across all European markets

Deliver: Implement a unified digital infrastructure for sales and marketing departments that enables consistent data capture, reporting, and analysis

The project scope intentionally started narrow—focusing on customer account tracking and sales activity reporting—with the explicit design goal of creating a scalable platform that could expand to additional processes over time [11].

Project Constraints:

Timeline: 6-month initial implementation for core reporting capabilities

Budget: Limited to existing Microsoft 365 and Databricks licenses; minimal incremental spending

Technical Resources: One dedicated project lead with data engineering capabilities; part-time support from IT for infrastructure

User Change Capacity: Sales teams fully occupied with revenue targets; minimal tolerance for disruptions to daily workflows

3.2 Initial State Assessment

The assessment phase revealed substantial variation in how different country operations managed sales data, consistent with patterns observed in SME digitalization literature [2]:

Data Landscape Findings:

Germany: Using a customized CRM system with extensive historical data, but limited integration with other systems

France: Combination of Excel spreadsheets and a local Access database, with strong informal knowledge among long-tenure sales staff

Spain: Recently implemented a cloud CRM, but adoption was inconsistent with many reps maintaining parallel Excel files

Bulgaria: Minimal formal systems; primary reliance on email communications and personal notes

Poland: Shared Excel workbooks on SharePoint, frequently encountering version control issues

Data Quality Issues Identified:

Customer Master Data Inconsistency:

Same customer appearing with different names across countries (e.g., "ABC Solar GmbH" vs. "ABC Solar" vs. "ABC-Solar Germany")

Missing or incomplete contact information (approximately 30% of accounts lacking phone numbers or email addresses)

Duplicate account records within single countries due to lack of pre-entry validation

Site and Installation Tracking:

No standardized approach to linking customer accounts to physical installation sites

Critical information (site address, system capacity, installation date) captured inconsistently or not at all

Inability to answer basic questions like "How many active installations do we have?" or "Which customer sites are due for service?"

Sales Activity Tracking:

No consistent definition of sales pipeline stages across countries

Revenue forecasting based on individual spreadsheets with varying assumptions and methodologies

Historical win/loss data largely undocumented, preventing pattern analysis

Business Impact:

The fragmented data landscape created tangible business problems:

European sales director unable to produce consolidated pipeline reports for executive leadership

Marketing campaigns executed without understanding total customer base or segmentation

 Online ISSN 2603-459X

Resource allocation decisions (hiring, territory planning) based on incomplete or anecdotal information

Customer service issues when installations were managed from different countries than original sale

3.3 Data Cleaning and Standardization Process

The data cleaning effort focused on creating a reliable foundation for the semantic layer, prioritizing customer account and installation site entities [3].

Phase 1: Customer Master Data Harmonization

Approach:

Data Extraction: Export customer records from all identified source systems into staging tables in Databricks

Fuzzy Matching: Use string similarity algorithms to identify potential duplicate records across and within systems

Algorithms: Levenshtein distance, Jaro-Winkler similarity for company names

Threshold: Similarity score >85% flagged for manual review

Manual Review Workflow: Present matched pairs to country sales managers for validation and merge decisions

Golden Record Creation: Establish master customer records with standardized fields including legal company name, primary and billing addresses, primary contact details (name, phone, email), customer-since date, assigned account owner, and activity status (Active/Lost/New).

Key challenges included sales managers lacking time for data review (addressed by prioritizing top 20% of customers by revenue), disagreements between countries about authoritative customer information (resolved by establishing a governance rule based on most recent direct interaction), and merging historical data from disparate schemas (solved through standardized target schemas with country-specific transformation scripts).

Phase 2: Installation Site Tracking

The "First Commissioning Report" became the pilot implementation for structured site tracking.

The "First Commissioning Report" became the pilot implementation for structured site tracking, enabling equipment inventory management, preventive maintenance scheduling, service contract renewal tracking, and geographic analysis of the installed base.

Data Model:

```
Customer (1) ----< (many) Installation Site
- CustomerID (Primary Key)          - SiteID (Primary Key)
- CustomerName                      - CustomerID (Foreign Key)
- AccountOwner                      - SiteAddress
- ActivityStatus                    - InstallationDate
```

- ... - SystemCapacity (kW)

```
- EquipmentSerialNumbers
- LastServiceDate
- NextServiceDue
- CommissioningStatus
```

- ...

Commissioning statuses were defined as: Planned (contract signed, not yet started), In Progress (installation underway), Commissioned (installed and operational), and Decommissioned (removed or end-of-life).

Data collection involved initial bulk upload from various sources, validation by regional service teams, and ongoing updates through automated ERP feeds, manual entry forms for field technicians, and quarterly reconciliation.

3.4 Semantic Layer and KPI Creation

The semantic layer in Databricks serves as the "business logic" layer that translates raw, cleaned data into business-meaningful metrics and dimensions [4].

Core Semantic Layer Tables:

1. Customer Segmentation View: Combines customer master data with calculated fields for segmentation. The view joins three tables (customer_master, site_summary, revenue_summary) and uses CASE logic to classify accounts as New Account (registered within 12 months), Lost Account (no orders in 18+ months), or Active Account. It also aggregates installation site counts, installed capacity, and lifetime revenue metrics. The complete SQL is provided in Appendix A.2.

2. Sales Pipeline View: Standardizes opportunity tracking across countries by mapping local pipeline stage names to a unified six-stage framework (Qualify, Develop, Propose, Negotiate, Won, Lost) through CASE logic. The view also calculates DaysInPipeline and WeightedValue (estimated revenue weighted by probability), joining opportunity data with customer segmentation for cross-dimensional analysis. The complete SQL is provided in Appendix A.3.

3. Installer Segmentation View (Primary Implementation): The cornerstone of the live reporting system, this view integrates three distinct source systems – CREdge CRM (account data), SetApp Server (commissioning records), and GEO Mapping (sales territory assignment) – into a unified analytical view. It segments installers into four categories based on registration date and commissioning activity as described in Section 2.4, with aggregated metrics including site counts, capacity, and temporal indicators. The complete SQL implementation is provided in Appendix A.1.

Integration with Power BI and Teams:

This view is configured as a live-updated table in Databricks, with direct connection to Power BI. The resulting dashboard is embedded in Microsoft Teams, allowing sales managers to:

Filter installers by segment (New 2025, Conversion, Dormant)

Drill down by sales area or individual sales manager

Track week-over-week changes in segment distribution

Export targeted lists for outreach campaigns

The Teams integration eliminates the need for sales staff to access separate BI tools—the dashboard appears directly within their daily workflow environment [7].

Key Performance Indicators (KPIs):

The semantic layer enables calculation of standardized KPIs across three categories: Customer Health Metrics (account growth rate, churn rate, account concentration, average account value), Sales Performance Metrics (pipeline velocity, win rate, average deal size, sales cycle by segment), and Operational Metrics (data completeness score, site commissioning time, service coverage).

Rather than pre-calculating KPIs in stored tables, the semantic layer uses views and live calculations. This ensures KPIs always reflect the latest data, allows business logic changes without data migration, and enables Power BI to leverage server-side query optimization. For the scale of this SME (tens of thousands of records), live calculation proved performant.

4. Results

The database optimization initiative produced measurable outcomes across three dimensions: data quality improvements [3], reporting efficiency gains [5], and business decision-making enhancements [6]. This section presents quantitative metrics where available, supplemented by qualitative observations from stakeholder interviews conducted during the project evaluation phase.

Importantly, results are reported at two levels: immediate technical outcomes (e.g., data completeness, query performance) and downstream business outcomes (e.g., time savings, campaign effectiveness). While technical metrics were captured systematically through automated monitoring, business outcomes rely on self-reported assessments from sales leadership, introducing potential bias that should be considered when interpreting findings.

4.1 Before/After Metrics

The implementation produced measurable improvements across multiple dimensions, consistent with BI system success factors identified in prior research [5]:

Data Quality Improvements:

Table 1. Data quality improvements before and after implementation.

Metric	Before Implementation	After Implementation	Change
Customer record completeness (required fields)	67%	94%	+27 pp
Duplicate customer records identified	Not measured	847 duplicates found and merged	N/A
Installation sites tracked in system	~40% of actual sites	~95% of actual sites	+55 pp
Countries with standardized reporting	0 of 5	5 of 5	100%

Reporting Efficiency Improvements:

Table 2. Reporting efficiency improvements before and after implementation.

Report Type	Time to Generate (Before)	Time to Generate (After)	Time Saved
European consolidated pipeline	3-4 days (manual aggregation)	Real-time (instant refresh)	~3.5 days
Customer account status by segment	Not feasible (data unavailable)	Real-time	N/A
Installation site inventory	1-2 weeks (field team survey)	Real-time	~10 days
Sales performance by owner	2-3 days per country	Real-time, all countries	~12 days total

Business Decision-Making Improvements:

Weekly Sales Reviews: Previously conducted at country level only, with no European-wide view. Now conducted with unified dashboard showing:

Pipeline by country, owner, and segment

Week-over-week changes in pipeline value and opportunity count

Flagged accounts (new large opportunities, stalled deals, at-risk customers)

Result: European Sales Director reports 40% reduction in meeting time due to shared real-time data; more time spent on strategy discussion vs. data clarification

Customer Segmentation Initiatives:

Lost Account Win-Back Campaign: Marketing team identified 230 lost accounts with >€50K historical lifetime value. Targeted outreach campaign recovered 34 accounts in first quarter (15% win-back rate)

New Account Onboarding: Standardized "90-day onboarding journey" implemented for all new accounts, tracked through Power BI dashboard. Early indication of improved retention.

Active Account Growth: Account managers now have visibility into installation site details, enabling proactive service contract renewals and expansion conversations

Resource Allocation:

Identified Germany had disproportionate pipeline relative to account owner capacity; reallocated inside sales support to German team

Discovered Spain had high number of stalled opportunities in "Proposal Sent" stage; implemented additional sales training on negotiation

4.2 Performance Improvements

Technical Performance:

Query Response Times: Power BI reports load in 2-5 seconds for typical user queries (10,000-50,000 row result sets)

Data Freshness: Near real-time updates (5-minute latency from source system changes to Power BI availability)

System Reliability: 99.7% uptime over 6-month measurement period; planned maintenance windows communicated in advance

Scalability Headroom: Current Databricks cluster utilization ~30% of provisioned capacity; room for significant data growth or additional use cases

User Adoption Metrics:

Power BI Report Access: 87% of intended users (sales team members and managers) accessed reports at least weekly within 3 months of launch

Training Completion: 92% of sales team completed 1-hour Power BI navigation training

User Satisfaction: Post-launch survey (n=42 respondents):

81% agreed or strongly agreed: "The new reports provide information I need for my role"

76% agreed or strongly agreed: "The reports are easy to use"

68% agreed or strongly agreed: "The reports have improved my effectiveness"

Common positive feedback: real-time access, cross-country visibility, consistent definitions

Common improvement requests: mobile access (addressed in subsequent release), additional filtering options

4.3 Decision-Making Enhancements

Strategic Planning Impact:

The unified data platform enabled European leadership to make evidence-based decisions that were previously based on limited or anecdotal information [6]:

Market Prioritization:

Data Insight: Bulgaria showing highest win rate (64%) but smallest pipeline value due to limited sales resources

Decision: Approved two additional sales hires in Bulgaria; within 6 months, pipeline value increased 180%

Product-Market Fit:

Data Insight: Analysis of lost opportunities revealed that certain product configurations were consistently losing to competitors in specific countries

Decision: Product team adjusted standard offerings and pricing for those markets; subsequent quarter showed improved win rates

Service Model Optimization:

Data Insight: Installation site tracking revealed that 40% of commissioned sites had no documented service activity in past 24 months

Decision: Launched proactive service campaign; converted 25% to paid service contracts, improving recurring revenue stream

Operational Efficiency Impact:

Sales managers reported concrete changes to daily and weekly workflows:

Account Prioritization: "Before, I would essentially work my accounts alphabetically or based on whoever emailed me most recently. Now I start my week reviewing the dashboard to see which accounts have the most significant pipeline changes or upcoming service renewals. I'm working smarter, not just harder." (Account Owner, Germany)

Pipeline Management: "The pipeline view with standardized stages helps me coach my team. When I see deals stuck in 'Proposal Sent' for more than 30 days, I know to ask questions. That visibility didn't exist before." (Sales Manager, France)

Customer Service: "When a customer calls about their installation, I can immediately see all their sites, equipment details, and service history. I don't have to put them on hold to dig through files or call the service team. It makes us look professional and organized." (Inside Sales Manager, Spain)

Data-Driven Culture Shift:

Perhaps the most significant qualitative outcome was the observable shift toward data-driven decision-making culture [6]:

Sales meetings now routinely start with dashboard review rather than subjective status reports

Disagreements about market conditions or competitive dynamics increasingly referenced with data queries rather than opinions

Sales teams began requesting additional data points and KPIs, demonstrating engagement and value recognition

Cross-country collaboration improved as teams could see and learn from each other's metrics

5. Discussion

The implementation outcomes described in Section 4 warrant interpretation within the broader context of SME digital transformation research [1]. This section examines the lessons learned, considers scalability and generalizability of the approach, and acknowledges limitations that qualify the conclusions drawn from this single-case study.

The discussion is organized around three themes: practical lessons applicable to similar SME digitalization initiatives; considerations for scaling the approach to larger datasets, additional business processes, or different industry contexts; and methodological limitations that constrain the strength of causal claims about the relationship between database optimization and business outcomes.

5.1 Lessons Learned

Lesson 1: Start with Business Questions, Not Technical Capabilities

The project's success stemmed from a disciplined focus on specific business questions that sales leadership needed answered:

"Which accounts are at risk of churning?"

"Where should we focus new business development efforts?"

"Are we tracking toward quarterly revenue targets?"

Early in the project, there was temptation to build comprehensive data models capturing every conceivable data point. Resisting this temptation and instead focusing on delivering answers to priority questions enabled faster time-to-value and maintained stakeholder engagement. Additional data points were added iteratively based on demonstrated need.

Lesson 2: Data Cleaning is 70% of the Effort (and Worth It)

The ratio of time spent on data cleaning and harmonization versus building the semantic layer and Power BI reports was approximately 70:30. This aligns with data quality research findings [3] and industry observations, but it's worth emphasizing for SME contexts where there may be pressure to "just get something working quickly."

The customer master data harmonization—identifying duplicates, standardizing names, filling gaps in contact information—was tedious and time-consuming. However, every subsequent analytics use case benefited from this foundational work. The alternative (building reports on dirty data) would have produced metrics that stakeholders didn't trust, undermining the entire initiative.

Lesson 3: The Semantic Layer is Where Business Value Lives

The semantic layer—where customer segmentation logic, KPI calculations, and business rules are encoded—proved to be the most valuable and reusable component of the architecture.

When the business decided to refine the definition of "Lost Account" from "no orders in 18 months" to "no orders in 24 months" (based on analysis showing customer buying cycles), this change was implemented in a single view definition in Databricks. All downstream reports automatically reflected the new logic without any Power BI modifications.

This separation of concerns (business logic in semantic layer, visualization in Power BI) proved critical for maintainability and governance.

Lesson 4: User Adoption Requires More Than Good Technology

Despite building a technically sound solution, user adoption was not automatic [7]. Several deliberate change management actions proved essential [12]:

Executive Sponsorship: The European Sales Director's visible commitment—attending training sessions, using the dashboard in meetings, asking questions based on the data—signaled organizational importance

Role-Based Training: Rather than generic "here's how Power BI works" training, sessions were tailored to specific roles (Account Owners, Inside Sales, Sales Managers) showing them exactly how to answer their specific questions

Quick Wins: Prioritizing use cases that would deliver immediate value (e.g., the lost account win-back campaign) created momentum and belief in the system's value

Feedback Loops: Establishing regular "office hours" where users could ask questions and request enhancements maintained engagement and improved the solution

Lesson 5: Choose Technology That Fits SME Constraints

The decision to use Databricks and Power BI was driven by several SME-specific considerations [2]:

Cloud-Based: No on-premises infrastructure to manage; IT team's limited resources not consumed by database administration

Existing Licenses: Power BI was already licensed through Microsoft 365 E5; Databricks community edition was sufficient initially, with affordable paid tiers for scaling

Skills Availability: SQL-based approach meant project lead could implement without rare/expensive data engineering skills; Power BI's low-code interface enabled eventual self-service

Integration: Both tools integrate naturally with existing Microsoft ecosystem (SharePoint, Teams, Excel)

Alternative approaches (building custom applications, implementing open-source solutions like PostgreSQL + Tableau, or enterprise data warehouses) would have required capabilities and resources beyond typical SME reach.

5.2 Scalability Considerations

Technical Scalability:

The current database (~500K rows) represents less than 1% of Databricks capacity, with room for 10-50x growth without architecture changes. Power BI supports hundreds of concurrent users, well above current usage (~50 weekly). The semantic layer approach has already enabled rapid expansion: the marketing team built campaign dashboards reusing customer segmentation views, finance connected revenue tables for forecasting, and the service team began developing predictive maintenance models.

Organizational Scalability:

Expanding to other business functions would require formalizing the currently informal governance framework, developing additional data engineering skills (through training, contractors, or low-code tools), and tailoring change management approaches to functions with different data readiness levels [6].

Geographic Scalability:

Expanding beyond Europe would require addressing data sovereignty requirements (GDPR compliance, regional Databricks instances), language localization for Power BI reports, and timezone considerations for real-time metrics.

5.3 Limitations

Study Limitations:

Single-Case Design: This case study examines one organization's implementation [10]. Findings may not generalize to all SMEs or all industries. The PV sector has specific characteristics (project-based sales, long sales cycles, site-based service model) that influenced design decisions [8].

Limited Timeframe: Results reported cover 6-month implementation and initial months of operation. Longer-term outcomes (multi-year impact on revenue growth, customer retention, etc.) are not yet measurable.

Researcher Role: The participatory action research approach provides deep insight but introduces potential bias. The researcher's involvement in implementation may influence how challenges and successes are perceived and reported.

Quantitative Metrics: While several quantitative metrics are reported (data quality scores, time savings, user adoption rates), some key outcomes (decision-making quality, strategic planning improvements) rely on qualitative assessment and stakeholder testimony.

Implementation Limitations:

Data Quality Ceiling: Despite significant improvements, data quality is not perfect. Approximately 5-10% of customer records still have incomplete information, particularly for older, inactive accounts. Ongoing data quality initiatives continue.

Source System Dependencies: The architecture depends on reliable data extraction from source systems (ERP, CRM, etc.). When source systems have issues or undergo changes, downstream impacts occur. This dependency creates fragility.

Technical Debt: Some aspects of the implementation (particularly country-specific transformation scripts) were built pragmatically rather than ideally. As the system matures, refactoring will be needed to maintain technical health.

Mobile Access: Initial Power BI reports were optimized for desktop viewing. Mobile experience (for field sales teams) is suboptimal, though Power BI mobile app provides basic access. Dedicated mobile design would improve usability.

Advanced Analytics Gap: Current implementation focuses on descriptive and diagnostic analytics (what happened, why it happened). Predictive analytics (what will happen) and prescriptive analytics (what should we do) are limited. Machine learning and AI capabilities exist in the technology stack but are not yet leveraged.

6. Conclusion

This case study has examined database optimization as a foundational driver for digital transformation in an SME context [1], demonstrating a practical implementation within the European photovoltaic sector [8]. The findings support the proposition that structured data architecture—specifically, the combination of Databricks for data integration and semantic layer design [4] with Power BI for visualization and Teams for end-user access—can deliver meaningful business value within the resource constraints typical of medium-sized enterprises [2].

The installer segmentation capability, which integrates CRM data, mobile commissioning records, and geographic reference information into a live-updated analytical view, exemplifies how targeted data solutions can address specific business questions while establishing reusable infrastructure for broader digitalization initiatives.

6.1 Summary of Findings

This case study demonstrates that database optimization serves as a foundational driver for digital transformation in SME environments [1], particularly in data-intensive industries like the photovoltaic sector [8]. The implementation of a Databricks-based data architecture connected to Power BI visualization tools enabled a European PV organization to:

Standardize fragmented data across five countries into a single source of truth

Improve data quality from 67% to 94% completeness on critical customer attributes

Accelerate reporting from days/weeks to real-time for key sales metrics

Enable evidence-based decision-making that produced measurable business outcomes (resource reallocation, market prioritization, product adjustments)

Build scalable foundation for additional analytics use cases across marketing, finance, and service functions

The technical approach—separating data processing (Databricks backend) from data consumption (Power BI frontend) [4] and investing heavily in semantic layer design—proved well-suited to SME constraints of limited technical resources, budget limitations, and the need for rapid value delivery [2].

Equally important were non-technical factors: executive sponsorship, role-based training, quick-win prioritization, and ongoing stakeholder engagement [7]. Database optimization is ultimately an organizational transformation challenge, not merely a technical one [12].

6.2 Practical Recommendations

Based on the lessons learned, the following practical recommendations are offered for SME leaders considering similar database optimization initiatives [2]: (1) Start with business value by identifying 5-10 specific questions data should answer and designing the architecture to address them; (2) Invest disproportionately in data quality, budgeting 60-70% of the project timeline for cleaning and harmonization; (3) Build the semantic layer deliberately, encoding business logic separately from visualization tools to enable reuse and governance; (4) Choose SME-appropriate technology by evaluating total cost of ownership and leveraging existing ecosystem investments such as Microsoft 365; (5) Plan for change management from day one [7], including executive sponsorship, role-specific training, and quick wins; (6) Design for scalability but implement incrementally, proving value with focused use cases before expanding; and (7) Measure and communicate outcomes through both quantitative metrics and qualitative feedback to maintain organizational momentum.

6.3 Contribution to Digital Transformation Research

This case study contributes empirical evidence on database optimization as a foundation for digital transformation, addressing a gap in literature that predominantly focuses on customer-facing innovations or automation technologies. It documents SME-specific implementation patterns – cloud-based architecture, semantic layer separation, incremental implementation, and role-based adoption strategies – that are transferable to similar digitalization initiatives [2].

Additionally, the research adds to the limited empirical literature on digital transformation in the renewable energy sector [8], particularly in PV installation and service businesses, and demonstrates practical application of the Microsoft 365 ecosystem in combination with cloud data platforms [5].

6.4 Future Research Directions

This case study opens several avenues for future research [10]:

Longitudinal Study: Track the organization over 2-3 years to measure long-term impacts on revenue growth, customer retention, operational efficiency, and competitive positioning. Does early database optimization investment compound over time?

Comparative Analysis: Conduct similar case studies in other PV sector SMEs with different technology choices (e.g., open-source stack, traditional data warehouses, SaaS analytics platforms) to compare outcomes and identify optimal approaches for different SME profiles.

Scalability Validation: Follow the organization as it expands database optimization to additional functions (finance, operations, HR). What new challenges emerge? How does governance evolve? Where are limits of the architectural approach?

Advanced Analytics Progression: Study the organization's journey from descriptive analytics (current state) toward predictive and prescriptive analytics. What capabilities and skills are required for SMEs to leverage machine learning and AI effectively?

Change Management Deep Dive: Conduct detailed analysis of user adoption patterns, resistance factors, and cultural change indicators [7]. What differentiates high-adoption vs. low-adoption user segments? How does data literacy evolve over time?

Cross-Industry Validation: Replicate the implementation pattern in SMEs outside the PV sector (manufacturing, professional services, healthcare, etc.) to test generalizability of findings and identify industry-specific variations.

References

[1] L. Li, F. Su, W. Zhang, and J.-Y. Mao, "Digital transformation by SME entrepreneurs: A capability perspective," Information Systems Journal, vol. 28, no. 6, pp. 1129-1157, 2018. doi: 10.1111/isj.12153

[2] S. Mittal, M. A. Khan, D. Romero, and T. Wuest, "A critical review of smart manufacturing & Industry 4.0 maturity models: Implications for small and medium-sized enterprises (SMEs)," Journal of Manufacturing Systems, vol. 49, pp. 194-214, 2018. doi: 10.1016/j.jmsy.2018.10.005

[3] R. Y. Wang and D. M. Strong, "Beyond accuracy: What data quality means to data consumers," Journal of Management Information Systems, vol. 12, no. 4, pp. 5-33, 1996. doi: 10.1080/07421222.1996.11518099

[4] M. Zaharia, A. Ghodsi, R. Xin, and M. Armbrust, "Lakehouse: A new generation of open platforms that unify data warehousing and advanced analytics," in Proc. Conference on Innovative Data Systems Research (CIDR), Virtual Event, Jan. 2021. [Online]. Available: http://cidrdb.org/cidr2021/papers/cidr2021_paper17.pdf

[5] B. Puklavec, T. Oliveira, and A. Popovic, "Understanding the determinants of business intelligence system adoption stages: An empirical study of SMEs," Industrial Management & Data Systems, vol. 118, no. 1, pp. 236-261, 2018. doi: 10.1108/IMDS-05-2017-0170

[6] A. Popovič, R. Hackney, P. S. Coelho, and J. Jaklič, "Towards business intelligence systems success: Effects of maturity and culture on analytical decision making," Decision Support Systems, vol. 54, no. 1, pp. 729-739, 2012. doi: 10.1016/j.dss.2012.08.017

[7] V. Venkatesh, M. G. Morris, G. B. Davis, and F. D. Davis, "User acceptance of information technology: Toward a unified view," MIS Quarterly, vol. 27, no. 3, pp. 425-478, Sep. 2003. doi: 10.2307/30036540

[8] D. Horváth and R. Z. Szabó, "Evolution of photovoltaic business models: Overcoming the main barriers of distributed energy deployment," Renewable and Sustainable Energy Reviews, vol. 90, pp. 623-635, 2018. doi: 10.1016/j.rser.2018.03.101

[9] P. D. Rigo, J. C. M. Siluk, D. P. Lacerda, G. Rediske, and C. B. Rosa, "The competitiveness factors of photovoltaic installers companies and a BSC model," Solar Energy, vol. 234, pp. 225-238, 2022. doi: 10.1016/j.solener.2022.01.067

[10] R. K. Yin, Case Study Research: Design and Methods, 5th ed. Thousand Oaks, CA, USA: Sage, 2014.

[11] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," MIS Quarterly, vol. 28, no. 1, pp. 75-105, Mar. 2004. doi: 10.2307/25148625

[12] A. Hanelt, R. Bohnsack, D. Marz, and C. A. Marante, "A systematic review of the literature on digital transformation: Insights and implications for strategy and organizational change," Journal of Management Studies, vol. 58, no. 5, pp. 1159-1197, 2021. doi: 10.1111/joms.12639

Author Contributions: Conceptualization, E.G.; methodology, E.G.; software, E.G.; validation, E.G.; formal analysis, E.G.; investigation, E.G.; resources, E.G.; data curation, E.G.; writing—original draft preparation, E.G.; writing—review and editing, E.G.; visualization, E.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data supporting the reported results are not publicly available due to confidentiality constraints of the case study organization. Anonymized summary statistics are available upon reasonable request.

Conflicts of Interest: The author declares no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

SME	Small and Medium-sized Enterprise
PV	Photovoltaic
CRM	Customer Relationship Management
KPI	Key Performance Indicator
BI	Business Intelligence
SQL	Structured Query Language
ERP	Enterprise Resource Planning

Appendix A

Appendix A.1. Installer Segmentation View – Complete Databricks SQL

This appendix provides the complete SQL implementation for the installer segmentation view discussed in Section 3.4. The code is designed for Databricks SQL and demonstrates the integration of three source systems: CREdge CRM, SetApp mobile commissioning server, and GEO mapping reference data.

```
-- ============================================================
-- INSTALLER SEGMENTATION VIEW
-- Databricks SQL Implementation
-- Version: 1.0
-- Purpose: Segment installers based on 2025 commissioning activity
-- ============================================================
-- Step 1: Create staging view for SetApp commissioning data
CREATE OR REPLACE VIEW staging_setapp_activity AS
SELECT
site_id,
installer_account_id,
site_address,
site_postal_code,
site_creation_date,
commissioning_date,
system_capacity_kw,
equipment_serial_number,
commissioning_status,
YEAR(site_creation_date) AS site_creation_year
FROM bronze_setapp.sites
WHERE commissioning_status IN ('Commissioned', 'In Progress', 'Planned');
-- Step 2: Create staging view for CREdge account data
CREATE OR REPLACE VIEW staging_credge_accounts AS
SELECT
account_id,
account_name,
account_type,
registration_date,
YEAR(registration_date) AS registration_year,
primary_contact_name,
primary_contact_email,
primary_contact_phone,
billing_address,
postal_code,
country_code,
last_interaction_date,
account_status
FROM bronze_credge.accounts
WHERE account_type = 'Installer'
AND account_status != 'Deleted';
-- Step 3: Create GEO mapping reference view
CREATE OR REPLACE VIEW ref_geo_mapping AS
SELECT
postal_code,
city,
region,
country,
```

```
sales_area,
sales_manager_name,
sales_manager_email,
inside_sales_contact
FROM reference_data.geo_postal_mapping;
-- Step 4: Aggregate site activity per installer
CREATE OR REPLACE VIEW agg_installer_site_activity AS
SELECT
installer_account_id,
-- 2025+ Activity Metrics
COUNT(CASE WHEN site_creation_year >= 2025 THEN 1 END) AS sites_created_2025_plus,
SUM(CASE WHEN site_creation_year >= 2025 THEN system_capacity_kw ELSE 0 END) AS
capacity_kw_2025_plus,
-- Lifetime Activity Metrics
COUNT(*) AS total_sites,
SUM(system_capacity_kw) AS total_capacity_kw,
COUNT(CASE WHEN commissioning_status = 'Commissioned' THEN 1 END) AS commis-
sioned_sites,
-- Temporal Metrics
MIN(site_creation_date) AS first_site_date,
MAX(site_creation_date) AS latest_site_date,
MIN(CASE WHEN site_creation_year >= 2025 THEN site_creation_date END) AS
first_2025_site_date,
-- Days since last activity
DATEDIFF(CURRENT_DATE, MAX(site_creation_date)) AS days_since_last_site
FROM staging_setapp_activity
GROUP BY installer_account_id;
-- Step 5: Main Installer Segmentation View
CREATE OR REPLACE VIEW gold_installer_segmentation AS
SELECT
-- Account Identifiers
ca.account_id,
ca.account_name,
ca.country_code,
-- Registration Information
ca.registration_date,
ca.registration_year,
-- =====================================================
-- CORE SEGMENTATION LOGIC
-- =====================================================
CASE
-- New Account 2025: Registered in 2025+ AND has 2025+ commissioning activity
WHEN ca.registration_year >= 2025
AND COALESCE(sa.sites_created_2025_plus, 0) > 0
THEN 'New Account 2025'
-- Conversion from Old Account: Registered before 2025 BUT has 2025+ activity
WHEN ca.registration_year < 2025
AND COALESCE(sa.sites_created_2025_plus, 0) > 0
THEN 'Conversion from Old Account'
-- Dormant: Registered before 2025, has historical activity, but none in 2025+
WHEN ca.registration_year < 2025
```

```
AND COALESCE(sa.total_sites, 0) > 0
AND COALESCE(sa.sites_created_2025_plus, 0) = 0
THEN 'Dormant Account'
-- New Registration Pending: Registered in 2025+ but no commissioning yet
WHEN ca.registration_year >= 2025
AND COALESCE(sa.sites_created_2025_plus, 0) = 0
THEN 'New Registration - Pending First Commissioning'
-- Legacy Inactive: Old account with no activity ever recorded
ELSE 'Legacy Inactive'
END AS installer_segment,
-- =====================================================
-- ACTIVITY METRICS
-- =====================================================
COALESCE(sa.sites_created_2025_plus, 0) AS sites_2025_plus,
COALESCE(sa.capacity_kw_2025_plus, 0) AS capacity_kw_2025_plus,
COALESCE(sa.total_sites, 0) AS lifetime_sites,
COALESCE(sa.total_capacity_kw, 0) AS lifetime_capacity_kw,
COALESCE(sa.commissioned_sites, 0) AS commissioned_sites,
sa.first_site_date,
sa.latest_site_date,
sa.first_2025_site_date,
sa.days_since_last_site,
-- =====================================================
-- SALES ASSIGNMENT (from GEO mapping)
-- =====================================================
gm.sales_area,
gm.sales_manager_name,
gm.sales_manager_email,
gm.inside_sales_contact,
gm.region,
-- =====================================================
-- CONTACT INFORMATION
-- =====================================================
ca.primary_contact_name,
ca.primary_contact_email,
ca.primary_contact_phone,
ca.billing_address,
-- =====================================================
-- METADATA
-- =====================================================
ca.last_interaction_date,
ca.account_status,
CURRENT_TIMESTAMP AS view_refresh_timestamp
FROM staging_credge_accounts ca
LEFT JOIN agg_installer_site_activity sa
ON ca.account_id = sa.installer_account_id
LEFT JOIN ref_geo_mapping gm
ON ca.postal_code = gm.postal_code;
-- ============================================================
-- SUMMARY STATISTICS VIEW (for dashboard KPIs)
-- ============================================================
```

```
CREATE OR REPLACE VIEW gold_installer_segment_summary AS
SELECT
installer_segment,
sales_area,
country_code,
COUNT(*) AS installer_count,
SUM(sites_2025_plus) AS total_sites_2025_plus,
SUM(capacity_kw_2025_plus) AS total_capacity_2025_plus,
AVG(lifetime_sites) AS avg_lifetime_sites,
AVG(days_since_last_site) AS avg_days_since_activity
FROM gold_installer_segmentation
GROUP BY installer_segment, sales_area, country_code;
```

Notes on Implementation:

Bronze/Gold Layer Convention: The code follows Databricks medallion architecture conventions, with source data in "bronze" schemas and curated views in "gold" schemas.

Null Handling: COALESCE functions handle installers with no commissioning activity, ensuring they appear in the segmentation rather than being excluded by joins.

Parameterization: The year 2025 is hardcoded for clarity; production implementations may parameterize this value for flexibility across reporting periods.

Refresh Strategy: Views are used rather than materialized tables to ensure real-time data; for larger datasets, consider Delta Live Tables with scheduled refresh.

Appendix A.2. Customer Segmentation View – Databricks SQL

The customer segmentation view combines customer master data with calculated fields for activity-based segmentation.

```
CREATE OR REPLACE VIEW customer_segmentation AS
SELECT
    c.CustomerID,
    c.CustomerName,
    c.AccountOwner,
    c.Country,
    -- Activity Status Logic
    CASE
        WHEN c.CustomerSinceDate > DATEADD(month, -12, CURRENT_DATE)
THEN 'New Account'
        WHEN c.LastOrderDate < DATEADD(month, -18, CURRENT_DATE)
THEN 'Lost Account'
        ELSE 'Active Account'
    END AS ActivityStatus,
    -- Ownership Classification
    c.AssignedRole AS InternalOwnership,
    -- Aggregated Metrics
    s.TotalInstallationSites,
    s.TotalInstalledCapacity_kW,
    r.Lifetime_Revenue,
    r.Last12Month_Revenue,
    r.DaysSinceLastOrder
FROM customer_master c
    LEFT JOIN site_summary s ON c.CustomerID = s.CustomerID
    LEFT JOIN revenue_summary r ON c.CustomerID = r.CustomerID;
```

Appendix A.3. Sales Pipeline View – Databricks SQL

The sales pipeline view standardizes opportunity tracking across countries with a unified stage framework.

```
CREATE OR REPLACE VIEW sales_pipeline AS
SELECT
    o.OpportunityID,
    o.CustomerID,
    c.CustomerName,
    c.Country,
    o.EstimatedRevenue,
    o.ProbabilityPercent,
    o.EstimatedCloseDate,
    -- Standardized Stage Definitions
    CASE o.LocalStage
        WHEN 'Qualification' THEN '1-Qualify'
        WHEN 'Needs Analysis' THEN '2-Develop'
        WHEN 'Proposal Sent' THEN '3-Propose'
        WHEN 'Negotiation' THEN '4-Negotiate'
        WHEN 'Closed Won' THEN '5-Won'
        WHEN 'Closed Lost' THEN '6-Lost'
        ELSE '0-Unknown'
    END AS StandardStage,
    -- Calculated Fields
    DATEDIFF(day, o.CreateDate, CURRENT_DATE) AS DaysInPipeline,
    o.EstimatedRevenue * (o.ProbabilityPercent / 100.0) AS WeightedValue
FROM opportunities o
    JOIN customer_segmentation c ON o.CustomerID = c.CustomerID
WHERE o.StandardStage NOT IN ('5-Won', '6-Lost')
    OR o.LastModifiedDate > DATEADD(month, -3, CURRENT_DATE);
```

References

1. Li, L.; Su, F.; Zhang, W.; Mao, J.-Y. Digital transformation by SME entrepreneurs: A capability perspective. Inf. Syst. J. 2018, 28, 1129–1157.
2. Mittal, S.; Khan, M.A.; Romero, D.; Wuest, T. A critical review of smart manufacturing & Industry 4.0 maturity models: Implications for small and medium-sized enterprises (SMEs). J. Manuf. Syst. 2018, 49, 194–214.
3. Wang, R.Y.; Strong, D.M. Beyond accuracy: What data quality means to data consumers. J. Manag. Inf. Syst. 1996, 12, 5–33.
4. Zaharia, M.; Ghodsi, A.; Xin, R.; Armbrust, M. Lakehouse: A new generation of open platforms that unify data warehousing and advanced analytics. In Proceedings of the Conference on Innovative Data Systems Research (CIDR), Virtual Event, January 2021.
5. Puklavec, B.; Oliveira, T.; Popovic, A. Understanding the determinants of business intelligence system adoption stages: An empirical study of SMEs. Ind. Manag. Data Syst. 2018, 118, 236–261.
6. Popovič, A.; Hackney, R.; Coelho, P.S.; Jaklič, J. Towards business intelligence systems success: Effects of maturity and culture on analytical decision making. Decis. Support Syst. 2012, 54, 729–739.
7. Venkatesh, V.; Morris, M.G.; Davis, G.B.; Davis, F.D. User acceptance of information technology: Toward a unified view. MIS Q. 2003, 27, 425–478.
8. Horváth, D.; Szabó, R.Z. Evolution of photovoltaic business models: Overcoming the main barriers of distributed energy deployment. Renew. Sustain. Energy Rev. 2018, 90, 623–635.
9. Rigo, P.D.; Siluk, J.C.M.; Lacerda, D.P.; Rediske, G.; Rosa, C.B. The competitiveness factors of photovoltaic installers companies and a BSC model. Sol. Energy 2022, 234, 225–238.
10. Yin, R.K. Case Study Research: Design and Methods, 5th ed.; Sage: Thousand Oaks, CA, USA, 2014.
11. Hevner, A.R.; March, S.T.; Park, J.; Ram, S. Design science in information systems research. MIS Q. 2004, 28, 75–105.

12. Hanelt, A.; Bohnsack, R.; Marz, D.; Marante, C.A. A systematic review of the literature on digital transformation: Insights and implications for strategy and organizational change. J. Manag. Stud. 2021, 58, 1159–1197.