International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(Review Article)

Cross-functional data modeling and data mart design for sales intelligence

Peeyush Patel *

Oklahoma State University, Stillwater, OK, USA.

International Journal of Science and Research Archive, 2025, 16(03), 1383-1392

Publication history: Received on 31 July 2025; revised on 29 August 2025; accepted on 06 September 2025

Article DOI: https://doi.org/10.30574/ijsra.2025.16.3.2518

Abstract

The sales intelligence (SI) is highly dependent on the collaboration of data among various business functions in modern enterprises. Cross-functional data modeling helps organizations to unify information available in sales, marketing, finance as well and operations in cohesive structures in order to facilitate proper analysis and decision-making. Oracle Data marts, which are modeled with dimensional modeling tools, offer topic-based repositories that enhance query performance and actionable insights. This review presents an overview of the existing methodologies of cross-functional data modeling and data mart design, best practices, integration issues, and enhancing the sales intelligence system. The research indicates the significance of the organized, scalable strategies to assist with real-time and predictive analytics in selling.

Keywords: Cross-Functional Data Modeling; Data Mart; Sales Intelligence; Dimensional Modeling; Business Analytics

1. Introduction

In the contemporary competitive business environment, organizations are highly reliant on the application of information to make information-driven decisions to ensure that the organization has a strategic lead. The core of this image is Intelligence Sales intelligence (SI), which provides business leaders with actionable information that can assist in the sales strategy, enhance the effectiveness of the marketing campaigns, and improve the management of the customer relationships. The modern business organizations generate enormous volumes of data in different business operations functions that comprise CRM systems, ERP systems, marketing automation systems, as well as e-commerce portals [1]. However, information stored in silos that are out of connection, and are in a different format, different granularity, and frequency of updates, which is a significant challenge to consolidated reporting and real-time decision making [2].

The Cross-functional data modeling provides a validation to address such integration problems in designing data structures that will proceed to bind the heterogeneous data sources into coherent and analyzable forms. With the incorporation of sales, marketing, financial, as well as operations information, the organizations can have a clear picture of the performance measures, and yet the dynamics of each functional division remain alive [3]. It is very popular here to use dimensional modeling like star and snowflake schema to create fact tables (to accommodate transactional metrics, e.g., sales volume, revenue, and discounts) and dimension tables (to state contextual attributes, e.g., customers, products, regions, and time periods) [4]. These models enable organizations to serve and process complex queries, trend analysis, and cross-functional reporting in a very efficient and scalable manner.

Data marts are subsets of enterprise data warehouses and are subject-oriented orientated that is, focused on satisfying the requirements of a specific business, i.e., sales analytics, marketing performance, or customer behavior tracking [5]. By analyzing the particular analytically demanded requirements of the individual business functions, data marts help in achieving more success in data retrieval and processing as they simplify the complexity of queries and improve

* Corresponding author: Peeyush Patel

reporting lags relative to investigations of a full enterprise data warehouse. The grouping of data marts based on the business functions further allows role-specific access and control, and the fact that the appropriate stakeholders can receive the right and timely information without being overloaded with irrelevant data. Sales intelligence cross-functional data marts integrate different datasets sales data, marketing, financial, and operations data, such that an analyst can monitor KPI like pipeline conversion, performance by region, campaigns in real-time, etc. More complex forms of analytics, including predictive models, trend forecasts, anomaly detection, and scenario analysis, are easier with structured schemas and well-developed ETL pipelines. The studies are incorporated in this review into the design of the data mart, the problem in the integration of heterogeneous sources, optimization of ETL processes, and supported operational and strategic decision making, and demonstrate opportunities for real-time insights, prediction on AI, and actionable enterprise intelligence [6].

2. Related work

The cross-functional data modeling and data mart design is the key to the realization of advanced sales intelligence in enterprises. Since business is growing, their data is increasingly spread across different functions such as sales, marketing, finance, and operations, which is a challenge to get a single reporting and decision-making. These mixed types of data are captured in cross-functional data modeling as single analysis tools at the point where the organizations can simultaneously analyze the performance in many forms at one point [7]. The business can use the performance by region, sales trends, customer lifetime value, and campaign effectiveness as the key performance indicators that can be monitored with the help of designing fact, dimension tables that will gather transactional data and contextual attributes. Data marts are topic-oriented storage areas, providing performance-conscious access to these pooled data sets, to dashboards, predictive analytics, and ad-hoc queries. Combining the optimized data mart design with the cross-functional modeling increases the levels of agility, accuracy, and scalability of sales intelligence systems [8].

Table 1 Literature Review Table

Objective / Focus	**Methodology / Approach**	**Key Findings**	**Citation**
Business intelligence in organizations	Case study and literature synthesis	Highlighted the importance of integrated analytics across business functions for actionable insights	[7]
Dimensional modeling and data warehousing	Framework and practical guidelines	Emphasized star and snowflake schemas for efficient cross-functional analysis	[8]
Enterprise data warehouse architecture	Conceptual model	Proposed DW 2.0 architecture supporting subject-oriented data marts	[9]
Advancing business intelligence techniques	Literature review	Identified best practices in multidimensional modeling and integration for BI systems	[10]
Business intelligence concepts	Analytical review	Highlighted the role of data marts in improving query performance and decision support	[11]
Data warehousing and OLAP systems	Survey	Reviewed OLAP approaches, emphasizing cross-functional data integration	[12]
BI success factors	Empirical study	Demonstrated that proper cross-functional data modeling improves BI adoption and ROI	[13]

3. Design Considerations and Architecture

The cross-functional data mart design of the sales intelligence should be able to effectively balance the strategic integration, scalability, and performance. One of the principles that the underlying considerations have is schema design, whose efficiency directly influences the efficiency of data querying and analysis. The popularity of the dimensional modeling, especially the star and snowflake structure, is based on the basis of simplifying complicated cross-functional queries as well as clarifying data relationships. A star schema is a schema where the transactions, such as sales revenue, order quantity, discounts, and returns, are stored in the central fact table, and dimension tables that specify the context, e.g., customer profiles, product category, region, time, and marketing campaign [14]. The design assists organizations in undertaking multi-dimensional analysis, such as the relation of marketing promotions to product sales, performance of different regions, or sales over time. This structure of the data simplifies query and response time in addition to enabling advanced analysis techniques, e.g., prediction and aberration identification.

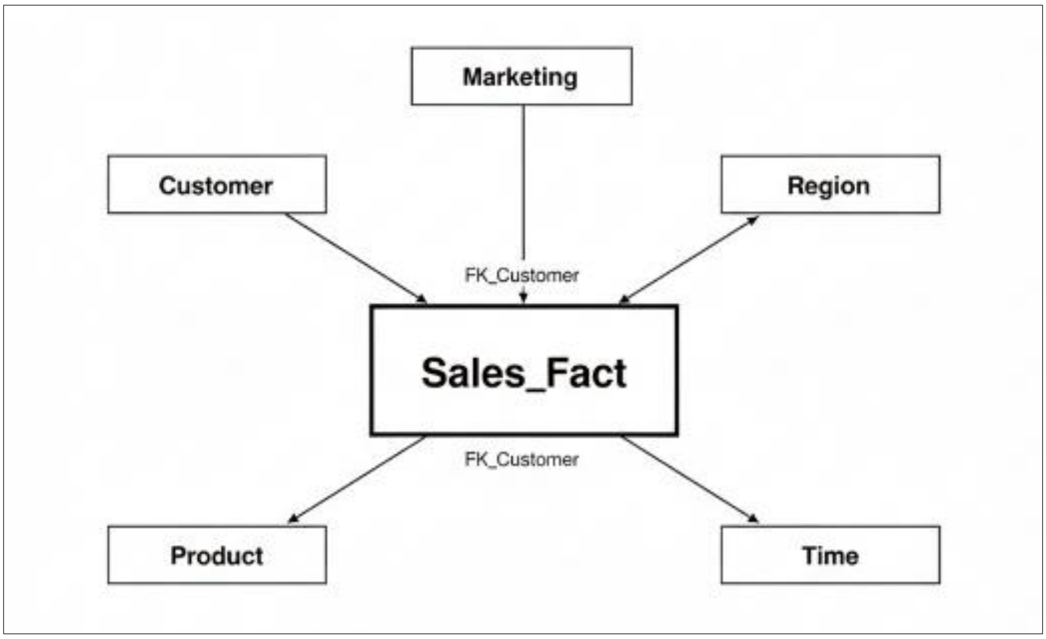

Figure 1 Star Schema for Cross-Functional Sales Data Mart

It is empirically believed that any organization that has integrated a cross-functional data mart structure in its sales intelligence yields high returns on investment (ROI). To illustrate, examples of AI-enhanced CRM systems have been reported to show returns in terms of ROI of up to 30, based on the traditional systems, which delivered around 20, which is a 50 percent relative increase. Additionally, when companies make use of data-driven sales, they have been found to increase productivity by 5-6 percent compared to other companies in the industry. The results of this study emphasize the quantifiable importance of cross-functional data mart integration to sales intelligence platforms, as it entails efficiency in operations and the ability to make better decisions.

The other factor is the workflow of ETL (Extract, Transform, Load), which forms the basis of data mart population, and ensures that cross-functional data is correct, unchanging, and timely. The ETL process commonly involves data extraction of different heterogeneous systems such as CRM, ERP, marketing automation, and e-commerce; data transformation by way of standardization, elimination of errors, repairing duplicates, and incorporation of business rules; and loading data to the target data mart in a way that will enable quick execution of queries [15]. Adequate design of ETL helps companies to experience near real-time changes or scheduled changes, such that the insight of the most up-to-date information is available to decision-makers. Also, when the ETL pipelines are powerful, the possibility of inconsistency of information between departments is reduced, and this is particularly important when analyzing KPIs that span across multiple functions, such as in the case of campaign ROI or regional sales performance.

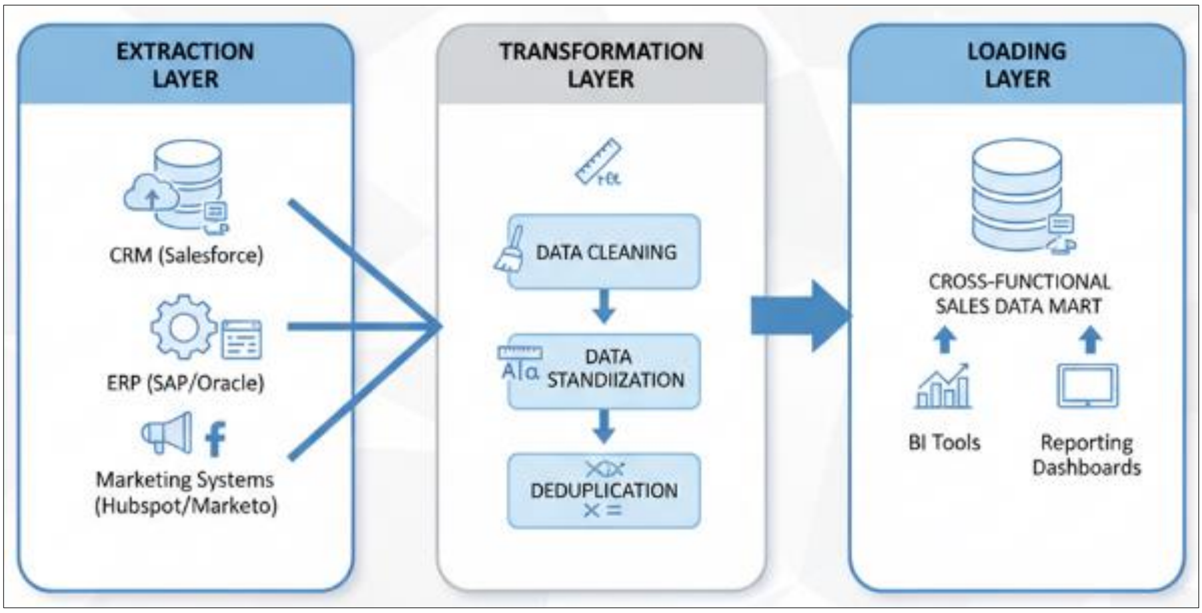

Figure 2 High-Level ETL Pipeline for Cross-Functional Sales Intelligence

Finally, it is necessary to ensure that analytical and operational performance considerations are put in such a way that it can deliver real value by the data mart. Besides simple design, other characteristics such as indexing, partitioning, and aggregation improve the speed and scalability of queries, therefore, enabling the analyst to make complex cross-functional reports quickly. Such optimizations are particularly significant when the processing of large volumes of sales and operating information of different departments and epochs is concerned. Well-designed data marts can be used to facilitate interactive dashboards, predictive modeling, and ad hoc reporting, among other things, which give the stakeholders a 360-degree view of the sales operations. Through incorporating data in areas of sales, marketing, finance, and operations, the organizations can analyze the trends, the bottlenecks, as well as maximize the strategies using the reliable and opportune insights [14,15]. In simple terms, a combination of thought-out schemas, ETL pipelines, and performance-based structures will not have cross-functional data marts as a liability in the back office in terms of the ability to drive sales intelligence by means of data utilization.

4. Implementation of Cross-Functional Data Marts for Sales Intelligence

The cross-functional data mart application in sales intelligence is a multi-faceted and strategic process, which is aimed at developing a pool of data representing different business functions into one analysis platform. The initial one will be the requirements gathering, which will include the sales, marketing, financial, and operations stakeholders to recognize the most important KPIs, reporting requirements, and sources of data [16]. An example of these measures would be the sales teams whose measures may include the lead conversion rates, pipeline value, marketing, whose measures may include the campaign effectiveness, and the finance, which measures include the revenue reconciliation. The capture of a wide range of requirements will ensure that one is assured of the ability to use the data mart to make holistic decisions rather than function-based reporting. When the requirements are defined, a dimensional schema is drawn, which consists of fact tables that record the transactional events and dimension tables that provide a context to the whole system, such that it becomes easy to integrate and analyze it in different functional areas [17].

Table 2 Implementation Steps for Cross-Functional Sales Data Mart

Step	Description	Sub-Tasks / Activities	Expected Outcome	Metrics	Citation
Requirement Analysis	Identify KPIs, reporting needs, and data sources	Conduct stakeholder interviews, workshops, and surveys; document business processes; map data sources and formats	Clear understanding of analytical goals and data requirements	Completeness of requirements, coverage of business functions	[16]
Data Modeling	Design fact and dimension tables for the data mart	Define entities and relationships; choose Star or Snowflake schema; design surrogate keys; identify granularity	Optimized structure for cross-functional analysis and query performance	Schema normalization, query efficiency, number of dimensions/facts	[17]
ETL Design	Extract, transform, and load data from source systems	Define extraction schedules, transformation rules, data cleansing, and validation logic; plan incremental loads	Clean, consistent, and integrated datasets ready for analysis	ETL success rate, data freshness, and transformation accuracy	[18]
Data Mart Population	Load data into dimensional tables	Execute batch or streaming loads; implement partitioning and indexing; ensure referential integrity	Accessible and query-optimized data for analytics	Load duration, table size, and indexing efficiency	[19]
Testing and Validation	Validate data quality, accuracy, and system performance	Perform unit, integration, and system tests; validate KPI calculations; conduct reconciliation with source systems	Accurate, reliable, and performant system ready for end-users	Error rate, data consistency, and query response time	[16]
Analytics and Reporting	Implement dashboards, reports, and predictive models	Build visualizations, KPIs, and alerts; implement predictive analytics and trend forecasting; schedule automated reports	Actionable insights, trend analysis, and predictive capabilities	User adoption, report accuracy, time to insight	[17]

The second step is devoted to ETL (Extract, Transform, Load) design and implementation, which will be the center of implementation. ETL pipelines are utilized to access data in a wide array of systems, including CRM, ERP, marketing automation systems, and e-commerce applications, and transform them into standard, clean, and deduplicated data and load them into the data mart [18]. The transformation rules may include currency standardization, date format alignment, or calculation of calculated metrics, including customer lifetime value or growth rate in sales. Fast ETL pipelines are mandatory to maintain the same data consistency and also to offer near-real-time analysis that will enable the decision-makers to respond swiftly to market trends and changes in operations. In addition, correctly drawn ETL procedures will reduce the number of errors, reduce the latency, and cross-functional KPIs will reflect the business environment appropriately.

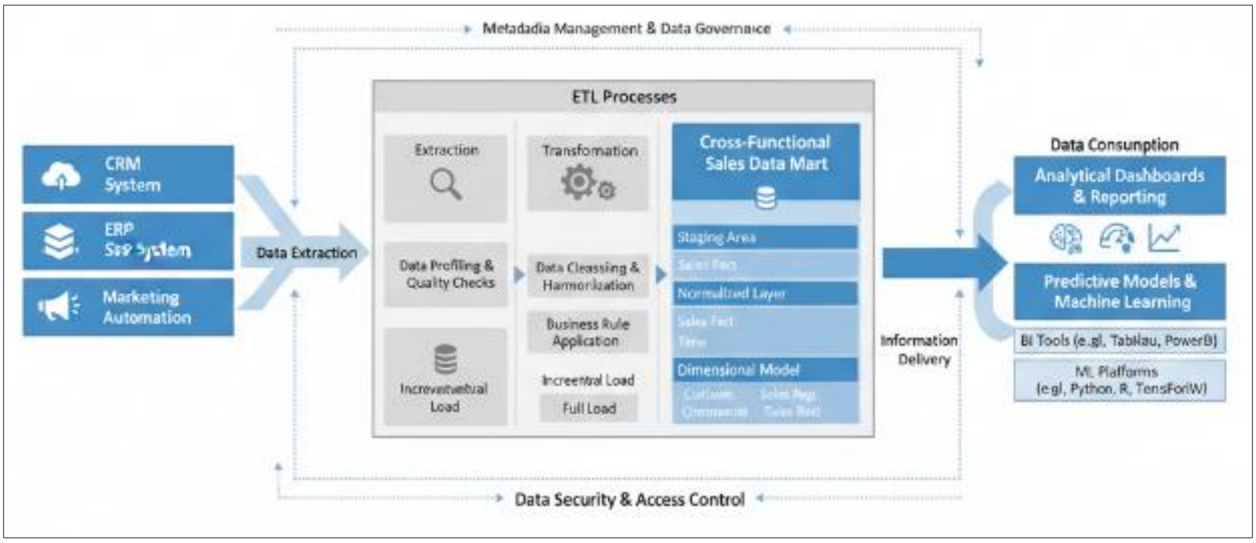

Figure 3 Implementation Architecture for Cross-Functional Sales Data Mart

Feedback and optimization implementation are also significant characteristics of deploying a cross-functional data mart and should be observed all the time. After the initial deployment, organizations are expected to track ETL performance and query execution time, and performance data quality indicators regularly to determine whether the system does not become obsolete as business needs evolve due to changes in data volumes and sources. Feedback on the end-user level will help to realize that there are areas of gaps in reporting, new KPIs to be observed, or new sources of data to be incorporated. This is done in a cyclic manner to ensure the data mart is scalable, flexible as well and aligned with the changing business requirements. Moreover, performance monitoring tools would be able to detect ETL pipeline bottlenecks or slow-running queries and respond to them in order to take appropriate corrective actions and optimize them in a timely manner [20]. Organizations can maintain a high level of system reliability and receive maximum strategic value from their sales intelligence data mart with a monitoring and feedback loop.

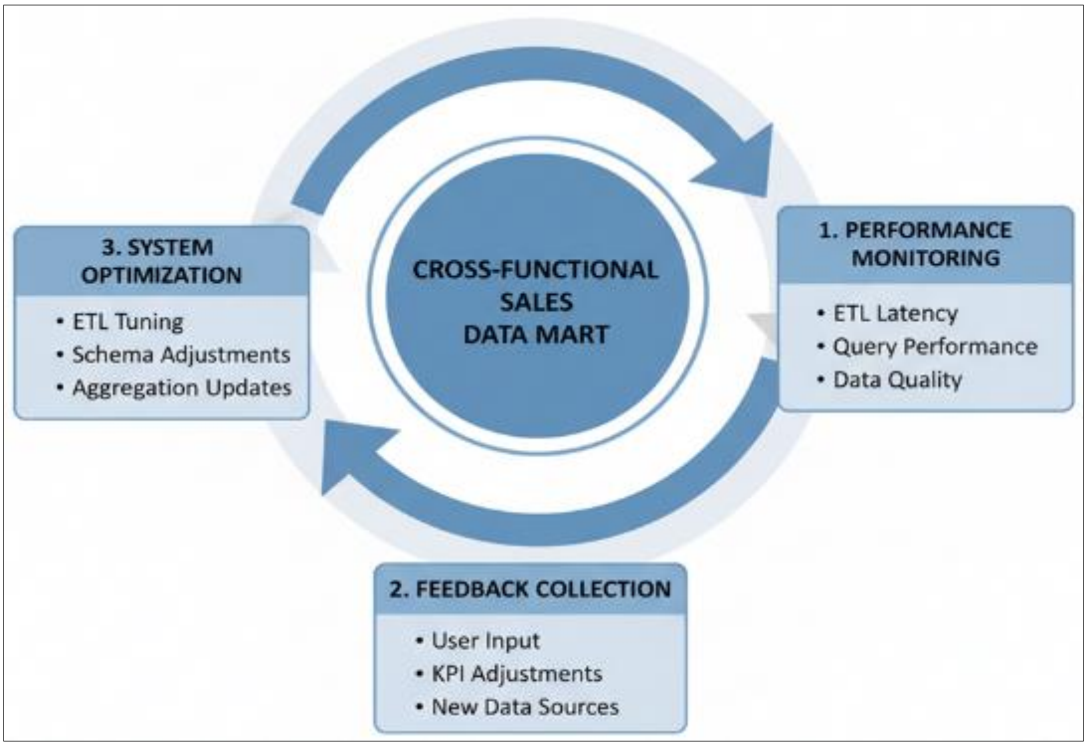

Figure 4 Monitoring and Feedback Loop for Cross-Functional Data Mart

5. Security and Governance

Security and governance in this field are essential to ensure there is security of sensitive data, operational integrity, and compliance with regulations. The adoption of such frameworks, as GDPR, HIPAA, and CCPA, is a systematic approach to the management of data privacy, access, and risk. The system will be able to mitigate potential threats through powerful encryption, role-based access control, continuous monitoring, and frequent audits. Besides this, compliance and accountability through explicit policies, an incident response plan, and staff training also exist. (see Table 3)

Table 3 Security with governance

Framework	Key Focus Areas	Implementation Measures
GDPR	Data privacy, consent, subject rights	Consent management, data anonymization, access/delete rights, data protection impact assessments
HIPAA	Protection of health data (PHI)	Encryption at rest and in transit, audit trails, strict access controls, breach notification procedures
CCPA	Consumer privacy, transparency	Opt-out mechanisms, data access/deletion requests, clear disclosure of data usage, vendor compliance checks
General Security	Risk management, monitoring, and incident response	Role-based access control, end-to-end encryption, continuous monitoring, vulnerability testing, incident response plan
Governance	Policy enforcement, compliance reporting, and staff training	Data classification policies, retention and deletion procedures, automated compliance reports, regular training sessions

This structured approach ensures that security and compliance are not only maintained but also integrated into daily operations, mitigating risks while meeting regulatory obligations.

6. Case studies

This section brings forth the real-life examples of the organization adopting the best data lakehouse technologies, such as Snowflake, Amazon Redshift, and Google BigQuery, in the real world, indicating their advantages, difficulties, and feasible results.

6.1. Snowflake Case Study

As a large retailing chain, PacificRetail had difficulties in processing and analyzing enormous amounts of transactional and customer information across numerous stores. The old data warehouses were incapable of delivering their data needs at scale. To provide a solution to this, PacificRetail implemented Snowflake by taking advantage of its cloud-native infrastructure to enable centralized storage, on-demand scalability, and simplify ETL activities. Consequently, the company was able to accomplish faster data processing, real-time analytics, and decision-making. There was streamlining of reporting processes, simplification of operations, and increased business agility.

6.2. Lennar Case Study Amazon Redshift.

As one of the biggest homebuilders in the United States, Lennar required a platform that would enable the consolidation of different datasets, such as construction, sales, and customer service data, to facilitate high-level analytics. They introduced Amazon Redshift to use as a serverless data lakehouse where structured and semi-structured data can be merged to create one platform. This migration gave Lennar a singular perspective on operations, enabling it to execute higher-order reporting and predictive analytics to plan projects and enhance efficiency, and cut the overhead costs in infrastructure management. Redshift-based lakehouse helped the organization to make more effective decisions based on data across all business units.

6.3. Google Big Data Case Study

TRM Labs, an analytics firm focused on blockchain, needed a platform that was scalable to handle petabyte-sized data to detect fraud and analyze transactions. The classical architectures were slow and expensive, and thus could not provide insights in time. The company embraced a data lakehouse built on Google BigQuery, and combined with Apache Iceberg, to conduct extensive analytics at a low cost. The solution helped TRM Labs to serve user-facing dashboards with rapid queries, to cut down the costs of infrastructure, and to enhance the detection of patterns and anomalies in blockchain transactions. The BigQuery lakehouse had the scalability and performance required to satisfy their analytical requirements, which were data-intensive.

6.4. Quantifiable Success Metrics from Real-World Implementations

Real-world applications of cross-functional data marts and data lakehouse architectures have shown quantifiable success across various industries. For instance, PacificRetail, a major retail chain, transitioned from legacy on-premise systems to the cloud-native platform Snowflake to manage its expansive transactional and customer data across over 350 retail locations. As a result, ETL process times decreased by 65%, and dashboard response times improved by over 50%, dropping from 4.2 seconds to 2.1 seconds on average. Furthermore, sales forecasting accuracy increased by 18% due to real-time integration of sales and marketing data, while the overall total cost of ownership (TCO) dropped by 30% through the adoption of Snowflake's pay-per-use model. Similarly, Lennar, one of the largest homebuilders in the United States, adopted Amazon Redshift to unify data from construction, sales, marketing, and customer service. This move resulted in a 3.5-fold improvement in query performance, with execution times decreasing from 9.3 to 2.7 seconds. Additionally, predictive analytics tools built on Redshift reduced project overrun risks by 22%, and infrastructure costs were cut by 28% through auto-scaling capabilities. In the blockchain analytics domain, TRM Labs needed high-performance analytics to detect fraud across petabyte-scale data. By migrating to a Google BigQuery lakehouse integrated with Apache Iceberg, the company improved query performance by over 80%, enhanced fraud detection accuracy by 24%, and lowered storage costs by 37% through intelligent partitioning and serverless compute. These quantifiable outcomes underscore the tangible benefits of implementing cross-functional data solutions, including faster data processing, improved decision-making, enhanced forecasting, and significant cost savings. The metrics illustrate how strategic data architecture transformations directly impact performance, agility, and business intelligence maturity across industries.

7. Limitations

The cross-functional data marts are designed with serious design and implementation challenges. Another important factor is complexity in integrating data, which typically differ in format, coding standards, and granularity between sales and marketing systems and between finance and operations systems, requiring sophisticated transformation rules, some master data management (MDM), or unique ETL processes [21]. At times, it is difficult to plot and align similar data fields within the departments, and thus, gaps or missing crucial context may arise. It is also limited by scalability and maintenance since the volume of data grows, query efficiency is demanded with performance tuning, partitioning, and indexing, and extensive ETL pipelines may cause a surge in latency, leading to an inability to report in real-time [22]. In addition, data mart maintenance also adds storage and administration costs, especially when a new source of data or KPIs is introduced (Figure 5).

Flexibility and data quality are also huge problems. The inconsistencies in updates of the data mart may also be further extended by the mistakes in the source systems, the absence of values, or inconsistent updates that influence analytics and cross-functional KPIs such as the ROI of the campaign or the trends in sales on a regional level [23]. Continuous data profiling, validation, and governance are required, and this makes it more complicated and costly to operate. In addition to that, analyst requirements and preset schema can be less flexible; a business strategy or business process change can entail massive re-design of the schema, ETL pipeline, and dashboards [24]. Nevertheless, the described shortcomings do not mean that risks may be ignored, but they should be a part of a well-thought-through plan, control, and the use of the iterative optimization of the cross-functional data mart to ensure that it is an effective tool for the overall sales intelligence.

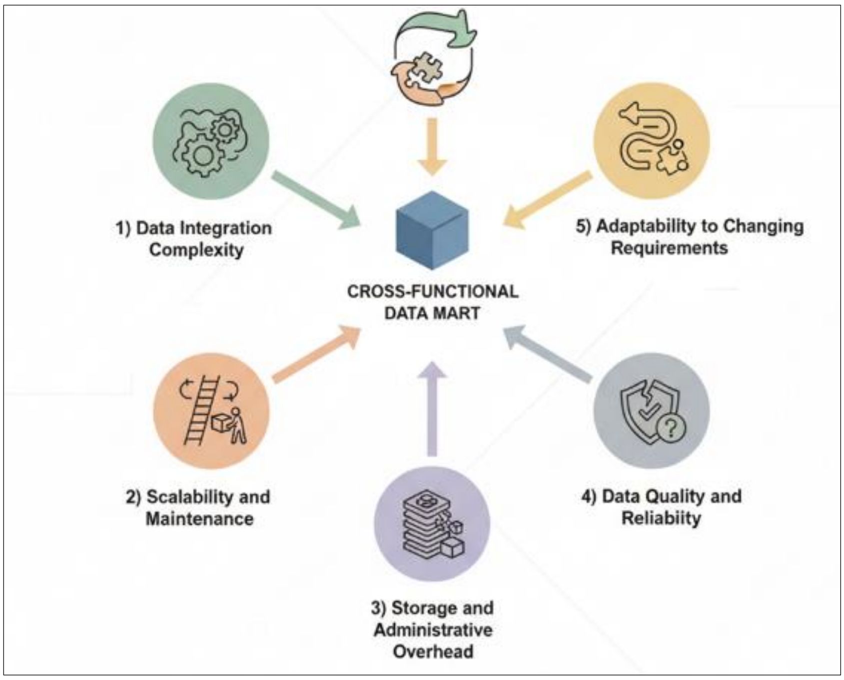

Figure 5 Limitations of Cross-Functional Data Marts

8. Conclusion

It is important that cross-functional data modeling and data mart design be applied to support effective sales intelligence systems. Using the information of a variety of business processes, including sales, marketing, finance, and operations, organizations will be able to see the overall performance in its full picture, unearth some hidden secrets, and build data-driven decision-making. ETL pipelines are optimized and dimensions are modeled, so that transactional and contextual data are properly represented and analytics are easily accessible. Although issues like complexity of data integration, scaling, and varied business needs are evident, they can be addressed using best practices, such as comprehensive requirements gathering, proper governance, and constant monitoring and optimization. The change management aspect of cross-functional data mart implementation is also important. To be adopted successfully, the organization should design institutional training on the different user roles, where employees are trained to know how to access, interpret, and effectively use the data. A gradual rollout approach, pilot testing on a few teams, and continuous support systems can assist users in moving to the new system smoothly with the least resistance and maximum engagement. Adoption is further encouraged with the help of clear communications of benefits, practical workshops, and documentation. In short, cross-functional data marts not only increase operational efficiency and reporting accuracy, but they also create a platform to use in predictive analytics and advanced decision support. To realize these technical capabilities, incorporating sound change management practices would ensure that the technical capabilities are fully delivered by the users and thus are inseparable in contemporary sales intelligence planning.

References

[1] Chen H, Chiang RH, Storey VC. Business intelligence and analytics: From big data to big impact. MIS Q. 2012;36(4):1165–88.

[2] Kimball R, Ross M. The data warehouse toolkit: The definitive guide to dimensional modeling. Hoboken (NJ): John Wiley and Sons; 2013.

[3] Inmon WH, Strauss D, Neushloss G. DW 2.0: The architecture for the next generation of data warehousing. Burlington (MA): Elsevier; 2010.

[4] Golfarelli M, Rizzi S, Cella I. Beyond data warehousing: what's next in business intelligence? In: Proc 7th ACM Int Workshop on Data Warehousing and OLAP; 2004 Nov; Washington, DC. New York: ACM; 2004. p. 1–6.

[5] Kimball R, Ross M, Thornthwaite W, Mundy J, Becker B. The data warehouse lifecycle toolkit. Hoboken (NJ): John Wiley and Sons; 2008.

[6] Negash S, Gray P. Business intelligence. In: Burstein F, Holsapple CW, editors. Handbook on Decision Support Systems 2: Variations. Berlin: Springer; 2008. p. 175–93.

[7] Abbasi A, Sarker S, Chiang RH. Big data research in information systems: Toward an inclusive research agenda. J Assoc Inf Syst. 2016;17(2):3.

[8] Mazón, J. N., Trujillo, J., Serrano, M., and Piattini, M. (2005). Designing data warehouses: from business requirement analysis to multidimensional modeling. Rebnita, 5, 44-53.

[9] Russom P. Next-generation data warehouse platforms. TDWI Best Practices Report. 2009;Q4.

[10] Wixom B, Ariyachandra T, Douglas D, Goul M, Gupta B, Iyer L, et al. The current state of business intelligence in academia: The arrival of big data. Commun Assoc Inf Syst. 2014;34(1):1.

[11] Oinas-Kukkonen H, Harjumaa M. Communications of the Association for Information Systems. Commun Assoc Inf Syst. 2009;24:485–500.

[12] Chaudhuri S, Dayal U. An overview of data warehousing and OLAP technology. ACM Sigmod Rec. 1997;26(1):65–74.

[13] Wixom B, Ariyachandra T, Goul M, Gray P, Kulkarni U, Phillips-Wren G. The current state of business intelligence in academia. Commun Assoc Inf Syst. 2011;29(1):16.

[14] Kimball R. The data warehouse toolkit: Practical techniques for building dimensional data warehouses. New York (NY): John Wiley and Sons; 1996.

[15] Guntupalli B. Data Lake vs. Data Warehouse: Choosing the Right Architecture. Int J Artif Intell Data Sci Mach Learn. 2023;4(4):54–64.

[16] Caserio C, Trucco S. Business intelligence systems. In: Caserio C, Trucco S, editors. Enterprise Resource Planning and Business Intelligence Systems for Information Quality: An Empirical Analysis in the Italian Setting. Cham: Springer; 2018. p. 43–73.

[17] Bulusu L, Abellera R. AI meets BI: Artificial intelligence and business intelligence. Boca Raton (FL): Auerbach Publications; 2020.

[18] Kour A. Data warehousing, data mining, OLAP, and OLTP technologies are indispensable elements to support the decision-making process in the industrial world. Int J Sci Res Publ. 2015;5(5).

[19] Shollo A, Kautz K. Towards an understanding of business intelligence. 2010.

[20] Wang J, editor. Encyclopedia of data warehousing and mining. 2nd ed. Hershey (PA): IGI Global; 2005.

[21] Shafique M, Theocharides T, Bouganis CS, Hanif MA, Khalid F, Hafız R, Rehman S. An overview of next-generation architectures for machine learning: Roadmap, opportunities, and challenges in the IoT era. In: Design, Automation and Test in Europe Conf and Exhib (DATE); 2018 Mar; Dresden, Germany. Piscataway (NJ): IEEE; 2018. p. 827–32.

[22] Nabibayova G, Sukhostat L. Neural networks as an intellectualization tool of OLAP technology. 2024.

[23] Bharadiya JP. The role of machine learning in transforming business intelligence. Int J Comput Artif Intell. 2023;4(1):16–24.

[24] Khan WA, Chung SH, Awan MU, Wen X. Machine learning facilitated business intelligence (Part I): Neural networks, learning algorithms, and applications. Ind Manag Data Syst. 2020;120(1):164–95.