Reading: Case study: Deloitte - Optimizing outdated database systems

In this part of the course, you have been learning about the importance of database optimization. This basically means maximizing the speed and efficiency with which data is retrieved in order to ensure high levels of database performance. Part of a BI professional's job is examining resource-use and identifying better data sources and structures. In this case study, you will have an opportunity to explore an example of how the BI team at Deloitte handled optimization when they discovered their current database was difficult for users to query.

Deloitte.

Company background

Deloitte collaborates with independent firms to provide audit and assurance, consulting, risk and financial advisory, risk management, tax, and related services to select clients. Deloitte's brand vision is to be a standard of excellence within the field and to uphold their brand values as they develop leading strategies and cutting edge tools for clients to facilitate their business. These values include integrity, providing outstanding value to clients, commitment to community, and strength from cultural diversity.

The challenge

Because of the size of the company and their ever-evolving data needs, the database grew and changed to match current problems without time to consider long-term performance. Because of this, the database eventually grew into a collection of unmanaged tables without clear joins or consistent formatting. This made it difficult to query the data and transform it into information that could effectively guide decision making.

The need for optimization appeared gradually as the team had to continually track data and had to repeatedly test and prove the validity of the data. With a newly optimized database, the data could more easily be understood, trusted, and used to make effective business decisions.

Primarily, this database contained marketing and financial data that would ideally be used to connect marketing campaigns and sales leads to evaluate which campaigns were successful. But because of the current state of the database, there was no clear way to tie successes back to specific marketing campaigns and evaluate their financial performance. The biggest challenge to this initiative was programming the external data sources to feed data directly into the new database, rather than into

the previous tables that were scheduled to be deprecated. Additionally, the database design needed to account for tables that represented the lifecycle of the data and designed with joins that could easily and logically support different data inquiries and requests.

The approach

Because of the scale of the project and the specific needs of the organization, the BI team decided to design their own database system that they could implement across the entire organization. That way, the architecture of the database would really capture their data needs and connect tables thoughtfully so they were easier to query and use.

For example, the team wanted to be able to easily connect the initial estimate of a marketing campaign's financial success with its ending value and how well internal processes were able to predict the success of a campaign. Increases from the initial estimate were good, but if estimates were frequently much higher than actual outcomes, it could indicate an issue with the tooling used to develop those estimates. But in the database's current state, there were dozens of tables across accounting groups that were creating access issues that were preventing these insights from being made. Also, the different accounting groups had a lot of overlap that the team hoped to more thoughtfully structure for more long-term use.

To achieve these goals, the team strategized the architecture, developed checkpoints to determine if the required functionality could be developed and eventually scaled up, and created an iterative system wherein regular updates to the database system could be made to continue refining it moving forward.

In order to consider the database optimization project a success, the BI team wanted to address the following issues:

- Were the necessary tables and columns consolidated in a more useful way?
- Did the new schema and keys address the needs of analyst teams?
- Which tables were being queried repeatedly and were they accessible and logical?
- What sample queries could promote confidence in the system for users?

A variety of partners and stakeholders had to be involved in the optimization project because so many users across the organization would be affected. The database administrators and engineers working with the BI team were particularly key for this project because they led the design and creation of the database, mapped the life cycle of the data as it matured and changed over time and used that as a framework to construct a logical data-flow design.

These engineers then conducted interviews with various stakeholders to understand the business requirements for teams across the entire organization, trained a team of analysts on the new system, and deprecated the old tables that weren't working.

The results

Deloitte's BI team recognized that, while the database had been continually updated to address evolving business needs, it had grown harder to manage over time. In order to promote greater database performance and ensure their database could meet their needs, the BI team collaborated with database engineers and administrators to design a custom database architecture that thoughtfully addressed the business needs of the organization. For example, the new database structure helped build connections between tables tracking marketing campaigns over time and their successes, including revenue data and regional locations.

This database optimization effort had a lot of benefits. The greatest benefit was the organization's ability to trust their data–the analyst team didn't have to spend as much time validating the data before use because the tables were now organized and joined in more logical ways. The new architecture also promoted simpler queries. Instead of having to write hundreds of lines of complicated code to return simple answers, the new database was optimized for simpler, shorter queries that took less time to run.

This provided benefits for teams across the organization:

- The marketing team was able to get better feedback on the value created by specific campaigns.
- The sales team could access specific information about their regions and territories, giving them insights about possible weaknesses and opportunities for expansion.
- The strategy team was able to bridge the gap between the marketing and sales teams, and ultimately create actionable OKRs (Objectives and Key Results) for the future.

However, as you have been learning, database optimization is an iterative process. The BI team's work didn't stop once they implemented the new database design. They also designated a team to oversee data governance to ensure the quality of the data and prevent the same problems from happening again. This way, the data remains organized and also this team can continue refining the developed databases based on evolving business needs.

Conclusion

The databases where your organization stores their data are a key part of the BI processes and tools you create–if the database isn't performing well, it will affect your entire organization and make it more difficult to provide stakeholders with the data they need to make intelligent business decisions. Optimizing your database promotes high performance and a better user experience for everyone on your team.