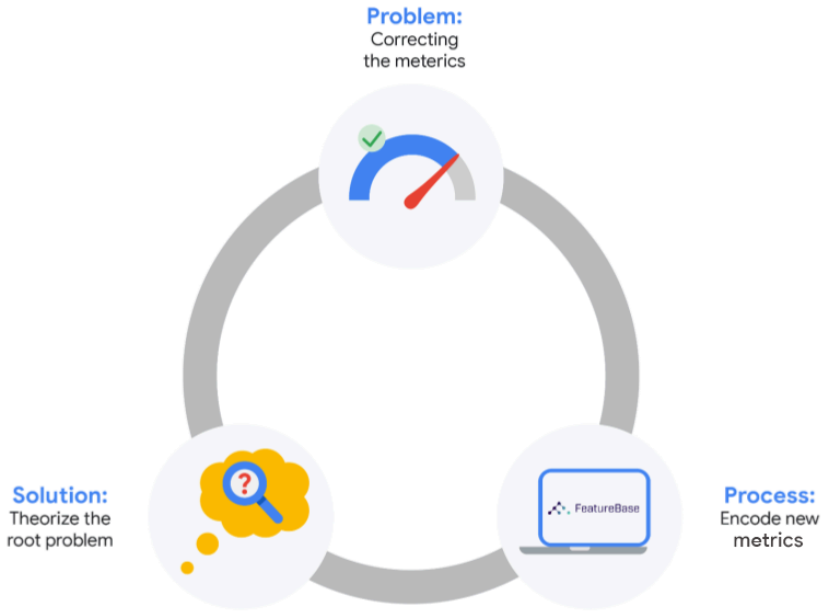
Reading: Case study: FeatureBase, Part 2: Alternative solutions to pipeline systems

This case study with FeatureBase will focus on the Analyze stage of the BI process, where you examine relationships in the data, draw conclusions, make predictions, and drive informed decision-making. This follows an earlier case study, where you explored the Capture phase of FeatureBase's project. In a follow-up case study, you'll learn about how FeatureBase addressed the Monitor stages of this project to solve their business problem. Like the previous FeatureBase case study, you'll consider the problem, process, and solution for this stage of the project.



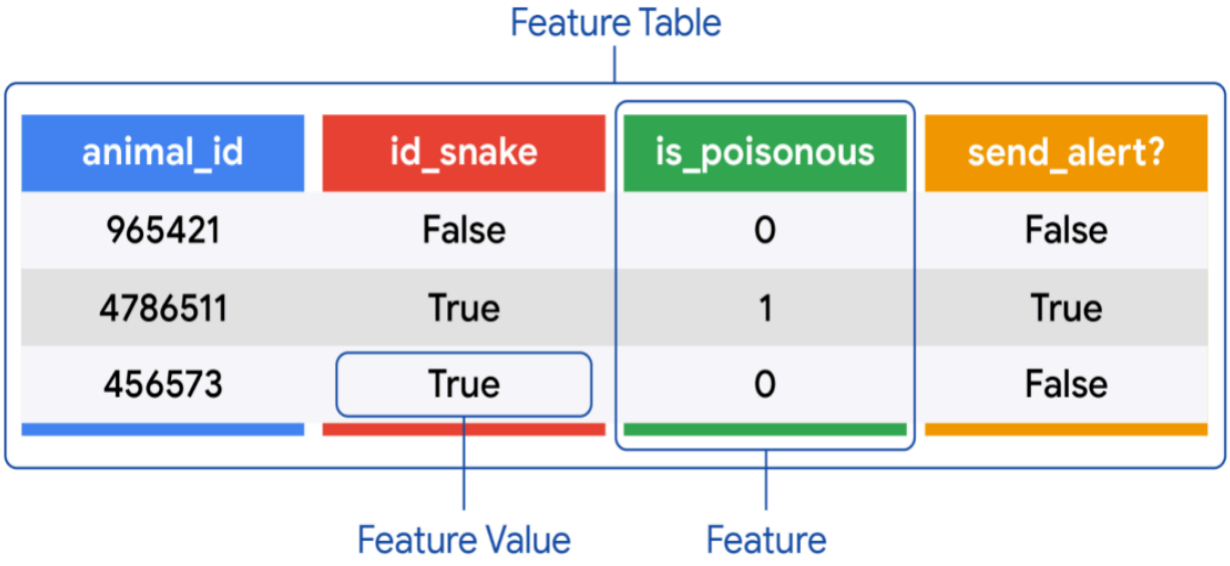
In a previous reading, you were introduced to FeatureBase, the OLAP database company. For a quick refresher, you can review **part one of this case study**. Their core technology, FeatureBase, is the first OLAP database built entirely on bitmaps that power real-time analytics and machine learning applications by simultaneously executing low latency, high throughput, and highly concurrent workloads. Last time, you learned about a business problem the FeatureBase team was facing: they realized that customers were falling off during the sales cycle, but that their data collection didn't have the necessary measurements to investigate when and why this was happening. The first step to addressing this issue was collaborating across sales, marketing, and leadership teams to determine what data they needed to understand when customers were falling off. Then, they could use that insight to investigate and address those issues for future sales. In this reading, you're going to focus on the database tools FeatureBase uses to collect data for monitoring and reporting.

Feature-oriented databases

Throughout this program, you have been learning about different database technologies that use pipeline systems to ingest, transform, and deliver data to target databases. This setup is fairly common across many different organizations, and you will often work with pipelines as a BI professional. However, there are also other types of database systems that use different technology to make data accessible and useful to users.

FeatureBase is one example of an alternative solution to traditional databases– it is built on top of bitmaps, which is a format that stores data as raw features. To build predicting models, the AI depends on the feature-oriented database (and not the other way around) to find patterns which can be used to guide decision making. These models are fed measurable data points, or features, that help it learn how to parse the data more effectively over time. Feature-oriented databases provide an alternate approach to data prep by automating the feature extraction as the first step. The feature-oriented approach enables real-time analytics and AI initiatives because the data or "features" are already in a model-ready data format that is instantly accessible and reusable across the organization, without the need to re-copy or re-pre-process.

Here's an example of this system: imagine you were trying to predict a type of animal based on traits. Humans would have a list of traits they might think about: wings, snouts, or the number of legs, for example. But these traits aren't actionable like that– but if they're converted into features like "has_wings" or "has_4_legs" with yes or no values coded as 1 or 0, they can be fed into a model and processed more quickly. This is why FeatureBase is built on bitmaps; the arrays of 1s and 0s are easily actionable by machines and models.



Fine-tuning features

In the last reading, you followed along as FeatureBase leadership considered how to approach their ultimate question: Why are customers dropping off during the sales cycle?

At that point in this project's cycle, the team didn't have metrics built into their system to find out when customers weren't completing the sales process, which was key to investigating why customers were dropping off. Having determined what their system was lacking, the team encoded these new features into the collection process– they recreated their original sales funnel with new attributes about customers at every stage of the sales cycle. These new features were fed into their database model, which began training to identify patterns, allowing them to immediately draw insights from their pool of data.

One of the attributes they added to the data collection process tracked exactly when customers were dropping off. They discovered that most customers who didn't complete the sale dropped off during the technical validation stage. This is the point at which the FeatureBase team would set up FeatureBase within the customer's system so they can try the product for themselves. The team realized that this was the critical stage they needed to investigate more. They theorized that customers weren't leaving the technical validation stage confident about their ability to adopt this new technology. They also wondered if customers were having concerns about the reliability and stability of the product since FeatureBase would replace their current, still-functioning database system. To understand this better, they would need to explore the customer data gathered at this stage in their dashboard.

The next step

As a BI professional, you might find yourself working with a variety of database technologies connected with pipeline systems like more traditional row based technologies and newer alternatives such as FeatureBase. Understanding your tools and how they operate will help you focus on what's most important–empowering your team with access to the answers they need. As they continued investigating their problem, the FeatureBase team found that most customers who did not complete the sales process were falling off in the technical validation stage.

This is the point at which FeatureBase was being implemented in the customer's data environment to determine if it was actually functional for them. This is how the FeatureBase team can showcase FeatureBase's utility and provide proof that it is a workable solution for a customer's needs.

The team theorized that, while many customers loved the capabilities of the service, they had some anxieties about how adoptable it would be for their organization and how reliable it would be. But before the team can confirm this theory or make any decisions based on this potential answer, they would need to be able to explore dashboard reports and investigate their metrics deeply. Which is what you will focus on when you return to this case study next time!
