

THE SAN FRANCISCO GIANTS WIN BIG WITH MODERN BUSINESS ANALYTICS

A world-class cloud data stack increases time-to-insight by 50 percent



ABOUT THE SAN FRANCISCO GIANTS

Location: San Francisco, CA

Founded: 1883

Industry: Sports and Entertainment

For sports fans, the San Francisco Giants need no introduction. The team is one of the oldest franchises in baseball and boasts 8 World Series Championships and 66 representatives in the Major League Baseball Hall of Fame. Off the field, Forbes Magazine ranks the Giants as the fifth most valuable team in the league at an estimated \$3.4 billion.

SAN FRANCISCO GIANTS DATA STACK:



A QUEST FOR EXCELLENCE, ON AND OFF THE FIELD

Every season, the San Francisco Giants take the field and play hard in front of thousands of avid fans. But the quest for excellence isn't limited to the baseball diamond. In the front office, the Giants' sales and marketing teams work hard to understand what makes fans tick to increase both engagement and ticket sales. That requires two things: A 360 view of fan activity, and a lot of data. The Giants had the fans and the data; putting it all together to gain insight was the challenge.

SILOED, UNTRUSTWORTHY DATA AND SLOW ANALYTICS

Similar to most sports and entertainment companies, the Giants wanted to better understand their customers, ticketing sales, partner campaign success, and venue data to drive better outcomes, increase revenue, and provide the best fan experience possible. But their existing data and analytics solution presented several challenges. First of all, all of that data was siloed within separate data sources. As a result, the data was inaccessible to people who needed it and the company lacked a single, centralized, source of truth. Plus, the data that did exist was untrustworthy, containing duplicate records that yielded incomplete or incorrect information. With these three barriers, reporting lagged and the team was hesitant to even start analysis.

A MODERN PLATFORM AND A "MASTER KEY" TO MANAGE FAN DATA

Partnering with Data Clymer, the Giants set out to democratize fan data, clean it up, and start to build that important 360 view. Together, the Giants and Data Clymer teams implemented a modern cloud data platform with three critical components.

snowflake : **WELL-SUITED AS A CLOUD DATA PLATFORM**

In addition to being a world-class data cloud platform, Snowflake was uniquely suited as an enterprise data platform for the Giants. With native support of JSON files, the data team could store results from Melissa, the master data management platform, as a variant data type directly in Snowflake. Then they could easily flatten the data into a structured format for further transformation processing.

MATILLION : **ELT BUILT FOR SNOWFLAKE**

As a cloud-native ELT tool, Matillion was the product of choice to bring data from multiple siloed sources into Snowflake and transform it into an analytics-ready format. Matillion is designed to work with Snowflake, enabling the team to:

- Leverage the power of Snowflake to handle large volumes of data quickly and at scale
- Dedicate Snowflake compute power to Matillion jobs so the team can run Matillion jobs and analytical queries at the same time
- Create custom metadata-driven ELT applications by creating configuration tables in Snowflake. This allows the team to extract a new table from a source without touching the Matillion code; they just need to add a new entry into a configuration table.

melissa® : KEEPER OF THE “MASTER KEY”

Data Clymer and the Giants chose [Melissa](#) to work with Matillion as the master data management (MDM) platform. With MDM, the teams were able to address duplicate data and better track changes.

Melissa enabled the teams to clean email addresses and mailing addresses leveraging access to the US Postal Service address database (among other libraries). The Giants could identify if an address is business or residential and attach a latitude and longitude. Then, the data is sent through another Melissa service that applies rules that start with a fan's first and last name as an anchor, then match across email address, mailing address, and phone number against the entire fan database of millions of records. Melissa then returns a master key for all matches.

The result? A “mapping table” that contains all the original fan contact data and the master key, providing the ability to trace back from the master key to the raw keys. This mapping table data is then rolled up into the master contact, account, and person tables. The Giants also track historical changes to the master tables to understand what changes may have occurred over time with the master keys. For those familiar with data warehousing best practices, Data Clymer used a slowly changing dimension type 2 design pattern. That means that any change to the data in a master row triggers a new row to be inserted and the old row to be marked as historical.

FASTER ANALYTICS, MORE ENGAGING OFFERS, AND FAN BASE GROWTH

This new modern data and analytics stack transformed how the Giants work with data. Silos are gone, and data now lives in a central location in the cloud where it is accessible to the departments that need it.

DUPLICATE RECORDS ARE ALSO A THING OF THE PAST:



The Giants reduced their fan record base by 15 percent with deduplication. **Where complex analytics projects once seemed out of reach because they took too much effort, the data team can now analyze data 50 percent faster than before.** Insight gets into the hands of the business faster, and analytics projects bring more value than ever before.

Within weeks of implementing the new system, the Giants had a better understanding of their customers including deeper insight into:

- **Ticketing** – Usage, sales, and packaging efficiencies
- **Stadium data** – Total entrances, seat data, and concession sales
- **Forecasting** – Marketing, ticket pricing, and partner campaigns
- **Ad-hoc game-day analytics** – Up-to-date data to support custom queries





All of this insight has added up to a deeper understanding of customer behavior and the ability to use that information to target customers with the deals and packages they really want.

“We used to base (customer insight) on TDC account number, which overestimated new customers being attracted through offers,” said Rocky Koplik, Vice President of Analytics at the Giants organization. “Now, we can see when a new customer truly enters our database. We can better assess the types of offers that draw their interest and follow their customer path forward.”

RESULTS

- **15 percent reduction** in fan datasets after deduplication
- **50 percent faster** time to insight with improved analytics
- Better understanding of customers within weeks of implementation

ABOUT DATA CLYMER

Data Clymer is a next-gen data consultancy enabling clients with trusted, full-stack cloud data and analytic solutions that drive positive business results through data accessibility and actionable insights. Our experienced team of consultants are certified data engineering and analytics experts and provide a highly-personalized level of service to enable data-driven companies like Peet’s Coffee, Thirty Madison, the Big Ten Conference, and the Las Vegas Raiders to unlock the value of data.

If you want to create a modern data stack for your organization, contact us at sales@dataclymer.com.

