

Advancing digital marketing solutions with an improved data warehouse

Business objective

Process Optimization

Sector

Media & Advertising

Technical services

Context & objectives

A leader in digital marketing solutions needed a data warehouse due to growing data volume and complexity. Data from different sources in Tableau for blending resulted in slowly reacting dashboards, thus curbing commercial and operational workflows.

Our client knew it was time to leverage a cloud data management infrastructure and reap benefits like cost optimization, less maintenance, cutting-edge technology, and ease of development. They contacted Agilytic to implement a new solution.

Approach

We set out to implement a proof-of-concept for a data pipeline and a data warehouse hosted in the cloud.

First, we defined the scope and the business requirements. We identified short and long-term use cases via client stakeholder interviews to help our client redesign data flows.

Next, we performed a market study to investigate the best cloud service provider suited to the client's needs. Moving forward with Amazon Web

Services (AWS), we identified the ideal cloud integration (development, test, production) and iteration of the analytical stack to ensure the infrastructure suited our client.

ETL development with AWS

Afterward, we performed a cloud infrastructure set-up (partly IaC) and ETL development for our data warehouse. There were two data sources for the use cases we worked on: a PostgreSQL database hosted in the AWS cloud and Google BigQuery containing analytics data. The client had experience with the cloud and was proficient with some good DevOps practices like Infrastructure as Code (IaC). Since we decided on the vendor, the only decision for us to make was about the ETL tool. Two options considered were: AWS Glue or Lambda functions written in Python triggered by Step. Tempted by the data catalog functionality, we chose Glue. From a technical point of view, the project consisted mainly of translating SQL queries into Spark. We gave two training sessions to introduce Glue and ETL concepts.

Before we completed the data warehouse project, we worked out the business logic to offer final fixes for the client. Additionally, we created a data catalog with a defined lineage and data governance guidelines for more straightforward long-term management.

Results

We developed an ETL pipeline and a robust, flexible data warehouse based on trusted cloud infrastructure in two months.

Business departments can access the most up-to-date data to make quick and impactful decisions with the solution. Moreover, in the case of a simple hesitation, they can quickly understand data by referencing the provided data catalog.

Newly hired developers will easily be able to add features or adapt changes requested by business users, given the scalability of this solution.

About Agilytic

Since 2015, Agilytic helps innovative leaders solve their biggest challenges through the smarter use of data. With over 150 successful projects to date, we have perfected a pragmatic approach to putting data at the service of business goals, be they commercial, operational, financial, or human. Reach out today for a quick introduction, we'd love to hear from you.