caser seguros | **Success story**

More efficient Data Analysis in less time

pue
Make your data shine

About Caser

Caser is a Spanish insurance group with over 75 years of experience, offering a wide range of solutions in all sectors: Home, Automobile, Health, Life, among others. Its activity is characterized by a strong customer orientation, quality of service, and the professionalism of its distribution lines. Additionally, Caser operates in other markets such as elderly care, hospitals, maintenance and assistance services, or financial advice. Caser is part of the Helvetia Group, which currently has a staff of more than 11,600 people.

Caser has committed to a Big Data platform with Cloudera to enhance its Data Centric model.

PUE has contributed by providing technological solutions that allow for more efficient data analysis in less time.

> The challenge for us is to start developing a platform that not only holds and delivers data but does so in a much faster and more agile way, making data as accessible as possible across the business. Data is a critical asset, especially in a financial business like ours. Making it easier to access and add new information that can be used in any context in the simplest way possible is a challenge, but it is one that we are pursuing.

Hugo González
Head of the Information Structures and Advanced Analytics area | Caser

The challenge

For Caser, data analytics is essential for making accurate decisions. They had a Data Centric model based on a Data Warehouse.

However, decisions could not always be made with up-to-date data due to its unavailability, necessitating reduced processing and query times to have real data at the right moment.

They also required a consolidated and reasoned data view and improvements in key areas: scalability and security in massive processes, new capabilities in data governance, traceability, lineage, cataloging, access controls, and auditing for correct data management.

The solution

A three-year strategy with a roadmap, whose first steps have already been executed, was defined. Firstly, the design and implementation of a Data Lake in Cloudera as part of the staging area and as a central piece of the company's Data Warehouse, significantly unifying and optimizing data processing and storage.

Replacing the relational database not only allows users to access and exploit data quickly but having the data in one place has enabled its use in reporting and analytics, as well as integration into production applications.

Future steps in the roadmap involve consolidating the Big Data platform as the central piece for data storage and processing for analytics, thereby increasing the development of use cases and functionalities offered by Cloudera and promoting the use of Machine Learning in the near future.

Technologies applied

Big Data Platform
Specifically:
- **Impala**
- **Kudu**
- **Spark**

Global results

Superior performance in **execution times flow**
- Execution times in complete Workflows reduced by 85% over initial times:
- Writing times on migrated tables in process (ODBC vs FTP+Spark) improved with an 80% reduction over initial times.

Improvement in **data quality**
- Detailed control of data input and output in validations corrected errors in migrated processes during the project.

Refinement of **final volume control**
- Detailed control of data input and output following complete data traceability allowed validation of load volumes with respect to production processes, addressing issues in the validation cycle.

Review of **developments in PWC**
- Migrating PowerCenter's own developments improved the logic applied in conjunction with Caser developers.

Increase in **process scalability**
- More cluster resources achieving an expansion in processing power when needed.

Better **data historification**
- More cluster resources achieving an expansion in processing power when needed.

Hugo González, *Head of the Information Structures and Advanced Analytics area at Caser, highlighted PUE's experience in Big Data but also the commitment and collaboration throughout the project*

pue　**caser** seguros

Hugo González
Head of the Information Structures and Advanced Analytics area en Caser

Adrián Buenaventura
Presentador
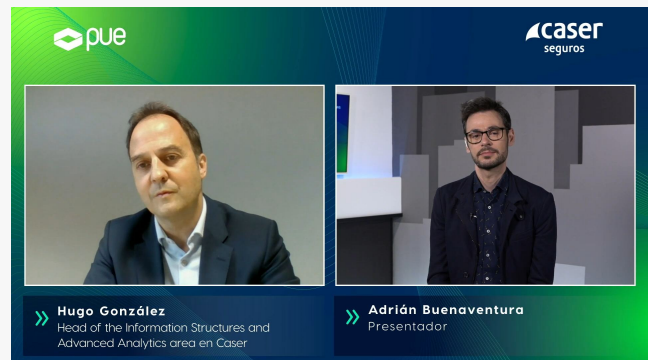
See the interview

PUE services

PUE collaborated by providing technological solutions throughout the project, designing and implementing a data architecture capable of improving Data Lake load performance to allow for more efficient data analysis in less time:

- Conversion of traditional programming languages like PL/SQL to Spark processes with Scala. This has resulted in reducing the time for data analysis processes from hours to minutes, gaining agility and providing an almost real-time business diagnosis.

- Quick and efficient integration of the ETL tool with the Data Lake.
- Configuration of clusters and security policies.
- Integration of PowerCenter processes to the Cloudera Platform.

Any data problem?
Your trusted partner for critical, complex and large-scale data projects

Contact us: sales@pue.es