

# Microsoft Azure Partner Solution Case Study



**Partner:** Kernel Analytics

**Website:** [www.kernel-analytics.com](http://www.kernel-analytics.com)

**Customer size:** 250 employees

**Country:** Spain (Catalonia)

**Industry:** Airlines/Aviation

## Partner profile

Combining mathematics, statistics, computer science, econometrics, and business expertise, Barcelona-based software house Kernel Analytics works to help customers realize the value of their data.

## Software and services

- Microsoft Azure Resource Manager
- Cloudera Enterprise including:
  - HDFS
  - MapReduce
  - Sqoop data import tool
  - Hive
  - Impala

## Open source technologies

- Apache Hadoop
- Programming languages:
  - R
  - Python
- Operating systems:
  - Linux (CentOS)

For more information about other Microsoft customer successes, please visit: [www.customers.microsoft.com](http://www.customers.microsoft.com)

Boutique Spanish analytics experts help a major airline create a Big Data platform to gain a competitive edge

*“The Microsoft-Cloudera combination is proving to be a highly workable foundation for data science that’s propelling our customers’ business forward.”*

Miquel Camprodon, Team Lead, Kernel Analytics

In an increasingly competitive environment, one of Europe’s premier airlines is using its accumulated online data to better personalize the service it can offer its customers. Barcelona’s Kernel Analytics, a highly specialized interdisciplinary team, has built a turnkey Big Data analytics platform based on Apache Hadoop Cloudera Enterprise on Microsoft Azure – a combination offering true power, reliability and scalability.

## A thirst for data-driven transformation

Airline users always have a choice. In today’s hyper-competitive market, the opportunity lies in making that choice as seductive as possible by offering compelling, relevant marketing efforts to both regular and casual customers.

The only way one major European air carrier could see to do that was to make better use of the large repository of online data it had accumulated around the habits and spending preferences of its frequent

fliers. The business objective was to retain customer interest by ensuring highly personalized communication and marketing deals. This can only really be done with a deep understanding of past behavior. The best way of doing that: analytics. The problem – while it had a massive amount of the data, it simply didn’t have big data in-house skills or the pre-requisite technology.

So, rather than simply outsourcing the analytics opportunity, its leadership was determined to develop a significant internal resource to enable in-depth data exploitation. To meet these challenges, the

airline approached Barcelona and Madrid-based software house, Kernel Analytics, which has built significant expertise around the application of analytics to handle a variety of client issues – from pricing and promotion optimization to fraud detection.

Kernel Analytics accepted the challenge, says Team Lead Miquel Camprodon. “Our cross-industry experience has enabled us to create a wide portfolio of services and solutions to complex business problems. And we’re happy to go the extra mile to deliver the best results for our clients. We decided to see how we could help this client design the best on-ramp to big data analytics.”

## Banking on Microsoft Azure and analytics

After a short but intensive period of on-site research, the team determined that the best solution was to deploy a customized internal analytical platform with the power to handle big data.

And the best technology for that platform? The highly regarded Cloudera Enterprise, optimized to run on the Azure cloud platform. Why? “Simple,” says Emiliano Carluccio, Science Lead at Kernel Analytics. “This combination is at the leading edge of all real big data work in the world.”

The team decided that the combination of Azure and Cloudera offered this customer the strength of two proven enterprise platforms not just for current needs but for the long term as well. “The client decided that its best option was to leverage the advanced analytics tools that Microsoft is deploying on Azure,” explains Camprodon. “And for us, the big bet on analytics that Microsoft is making with Azure was soon evident to both us and the client.”

Once they had decided on the technologies to use, the Kernel Analytics team started on the deployment of an engine that the customer could use to better understand its customers. First, the Kernel team created a full production Hadoop cluster, then performed an extensive migration of the airline’s existing data and a full extract, transform, and load (ETL) process from an old development environment to the new Cloudera-Azure platform.

The new Cloudera-Azure big data platform features three “master” and three “worker” nodes. The master nodes are comprised of DS12 machines offering 28 GB of RAM and the worker nodes are DS13 machines with 56 GB of RAM – all running on Linux (CentOS). This cluster is already performing intensive data loads at night and is the data hub for the realization of the client’s many planned data analysis projects and studies. The Cloudera Enterprise distribution being used employs the High Availability feature set.

At the same time, the Kernel Analytics team deployed a high-performance server based on R and Python that will host the dense statistical models that the client needs to analyze its massive data sets.

## Open source on Azure

The hard work of Kernel Analytics has resulted in a fully functional proprietary analytical technology that manages high volumes of internal data. The customer’s earliest goals for the system include approaches to better customer segmentation and ideas for ways to deliver personalized recommendations today’s airline users crave. Meanwhile, Kernel Analytics is already thinking beyond these immediate business needs to see how it

can help with fraud detection and, eventually, real-time, fully personalized ticket pricing.

“Now, our customer has the basis for a solid data strategy at a very low entry cost and is already working with us on multiple other projects,” notes Camprodon. “The Microsoft-Cloudera combination is proving to be a highly workable foundation for data science that’s propelling this customer’s business forward.”

Based on its experience with this and other deployments, Kernel Analytics has high praise for the Cloudera-Microsoft pairing. “We are now confident that the solution can scale seamlessly to at least two orders of magnitude without any major intervention or need for architecture modification,” says Camprodon. “We found that Azure was a very cost effective solution, allowing us to get exactly what we needed – when we needed it – without incurring huge extra costs for the analytics team/department,” adds Carluccio.

Without a doubt, a major feature of this successful project is the dovetailing of leading open source tools with Microsoft technology. For Camprodon, this is par for the course. “Many of our projects are developed using open source tools, from operating systems to programming languages to statistical tools. We’ve also been a Microsoft shop for many years, from the operating system level, to the Office suite, to development environments. We have now joined the Azure program, planning to certify our first Kernel Analytics professionals in this key enterprise cloud technology within the next three months, as we see a huge potential for this winning combination,” he predicts.