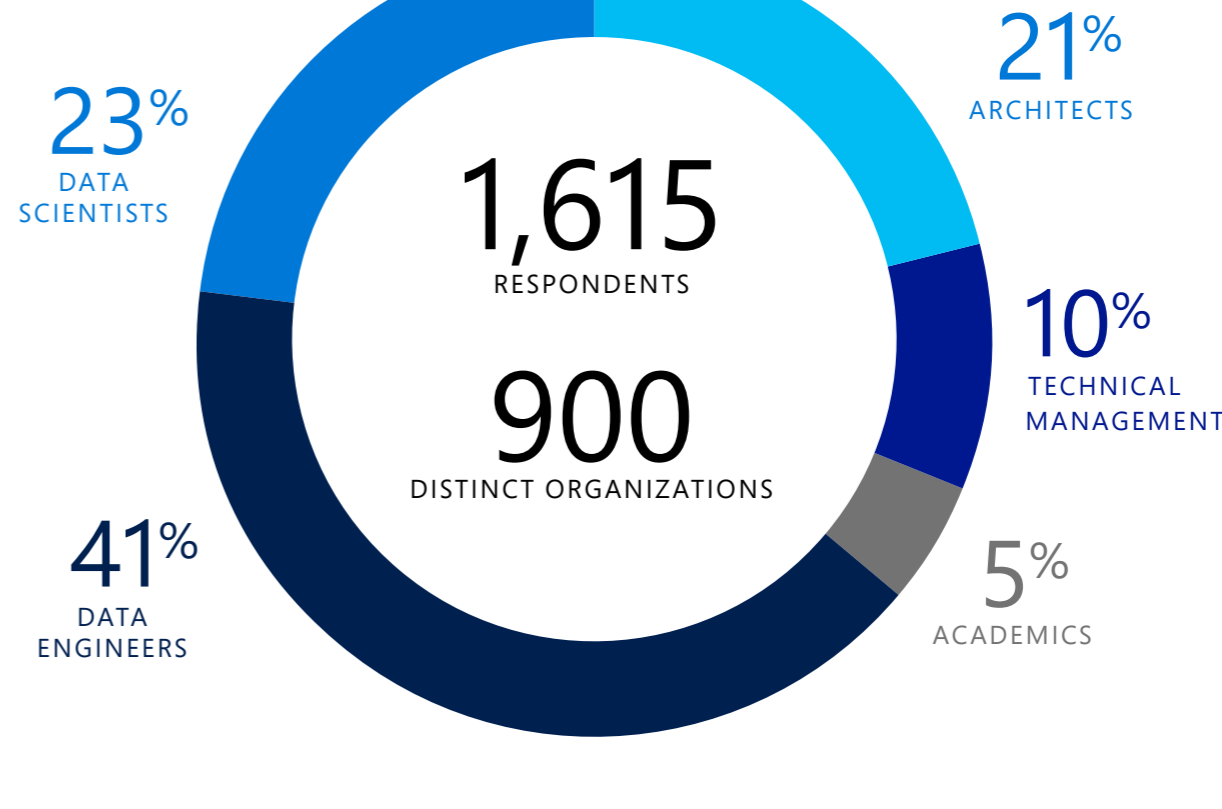


Better insights. Faster innovation.

Apache® Spark™ Survey highlights

The Apache Spark Survey identifies how organizations use Apache Spark.

The results from this survey reflect the answers and opinions of 1,615 respondents representing over 900 organizations. Survey respondents were predominantly Apache Spark users.

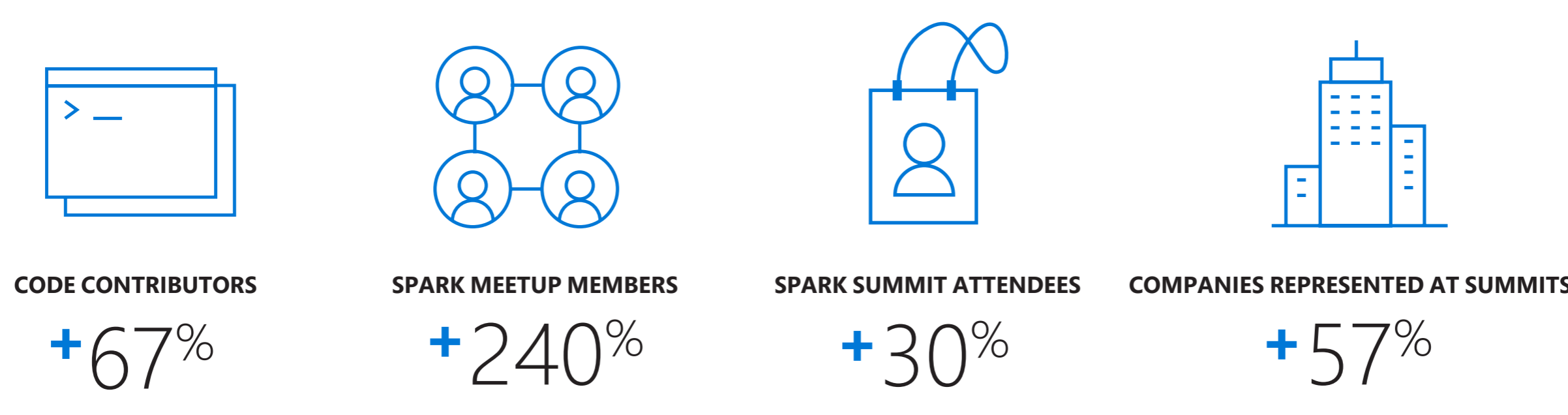


The survey results suggest that Spark's growth continues across various industries, as more people in various functional roles use it to build increasingly sophisticated data solutions. Mixing and matching multiple Spark components, different users build various types of products in production.

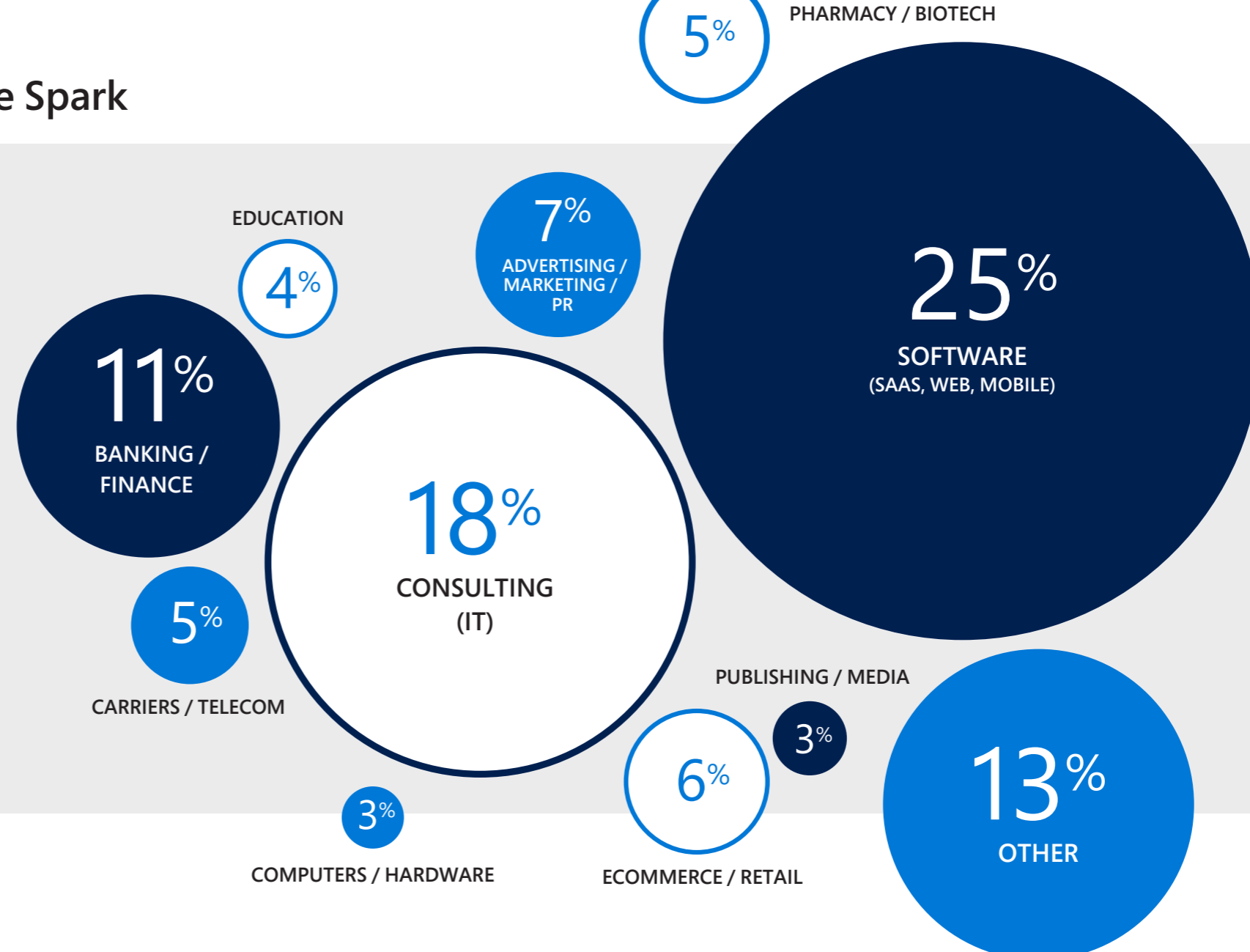
The results indicate that Spark has moved well beyond the early-adopter phase at high-tech companies and is now mainstream in large data-driven enterprises. And with the rise of public cloud computing, the survey findings reflect that users show an increased affinity toward using Spark in the public cloud.

Apache Spark's growth continues

Spark remains the most active open source project in Big Data. Today, there are over 1,000 Spark contributors, compared to 600 from 250+ organizations. With such large numbers of contributors and organizations investing in Spark's future development, it is evident that Spark has engaged a community of developers globally.

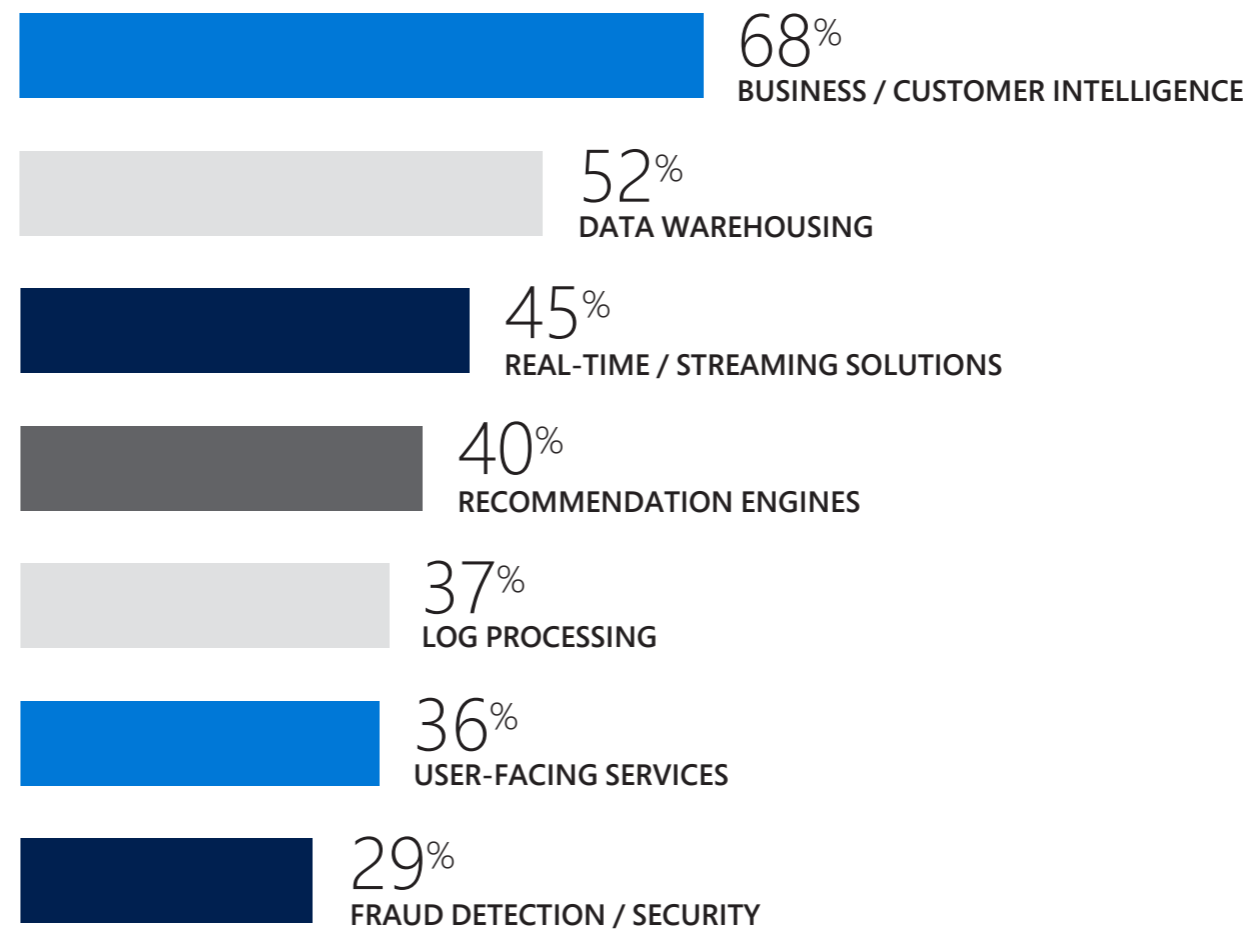


Key industries using Apache Spark

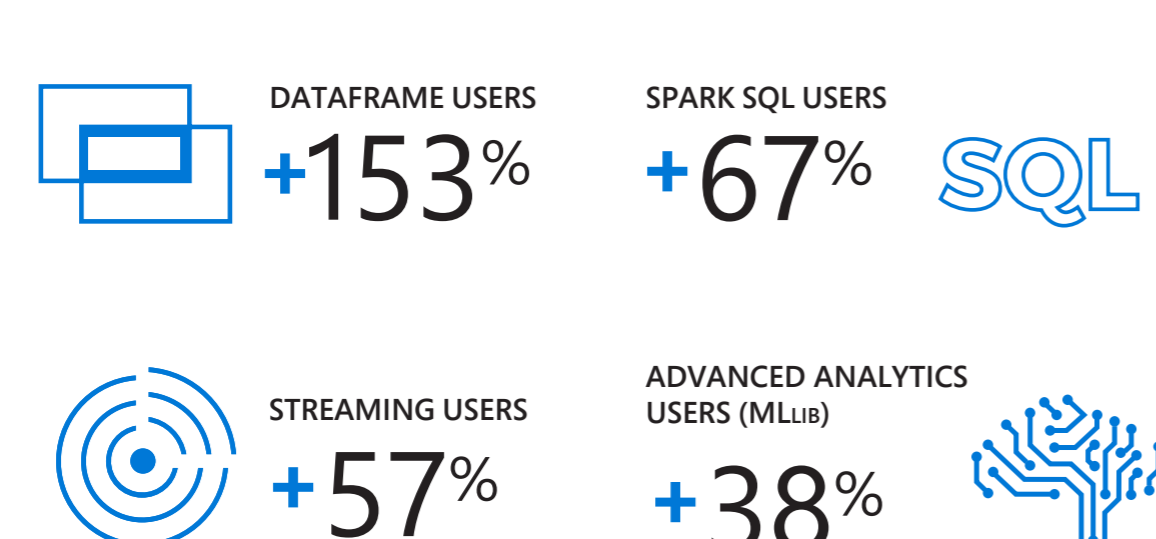


Apache Spark is being used to create sophisticated products within the same organization

Respondents were allowed to select more than one product.

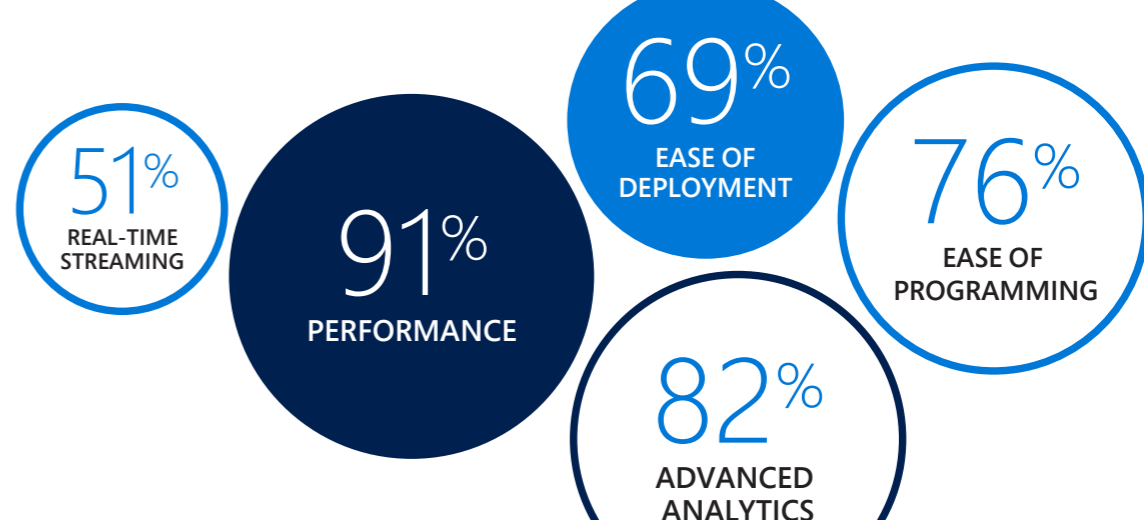


Apache Spark's fastest growing areas



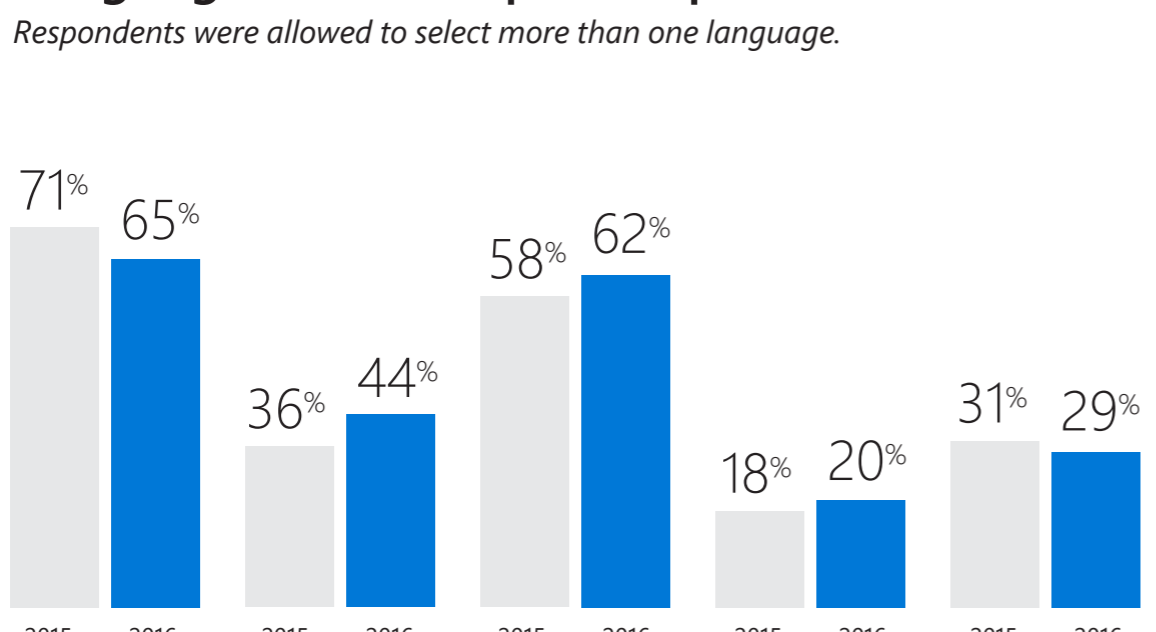
Features users consider important

Respondents were allowed to select more than one feature.



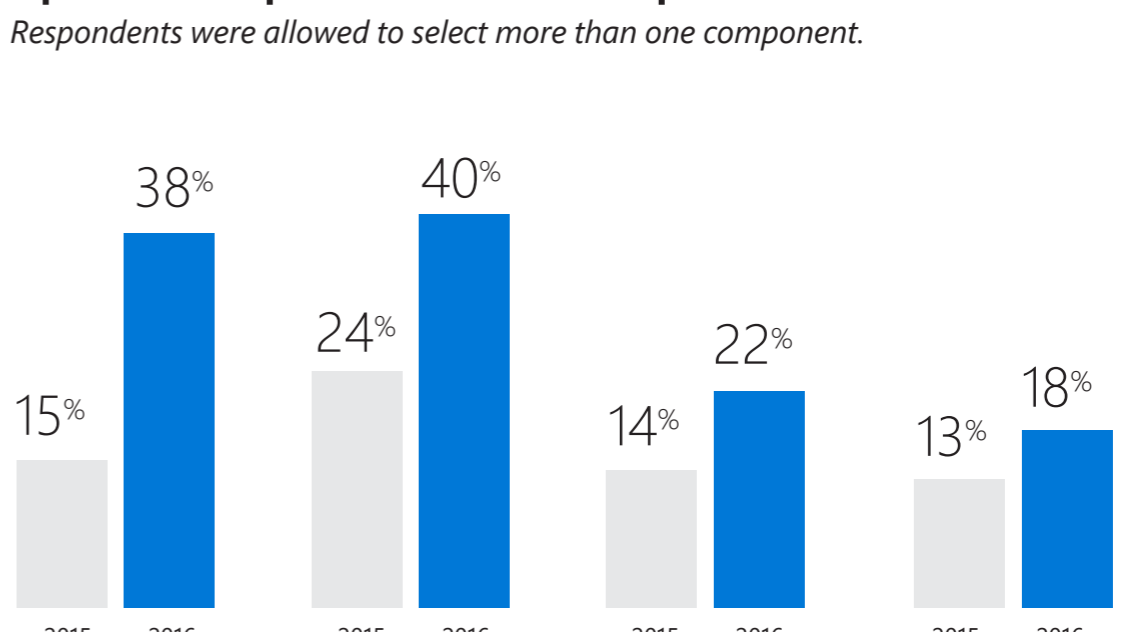
Languages used in Apache Spark

Respondents were allowed to select more than one language.



Spark components used in production

Respondents were allowed to select more than one component.

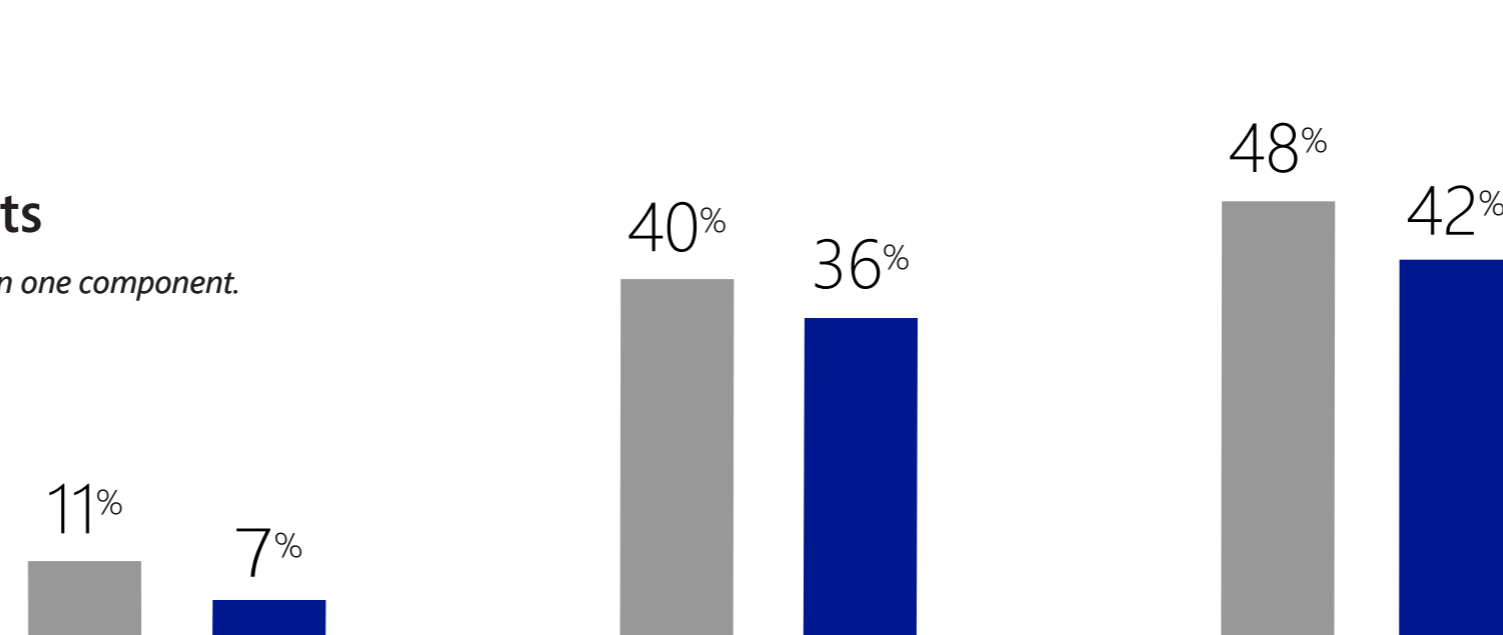


Apache Spark's growth and adoption continues as users, industries, development environments, disciplines, and programming languages embrace its ease of use and programming, its unified computer engine, and its performance to solve complex data problems at scale.

Apache Spark in the cloud is growing

Percentage decrease in on-premises deployments

Respondents were allowed to select more than one component.

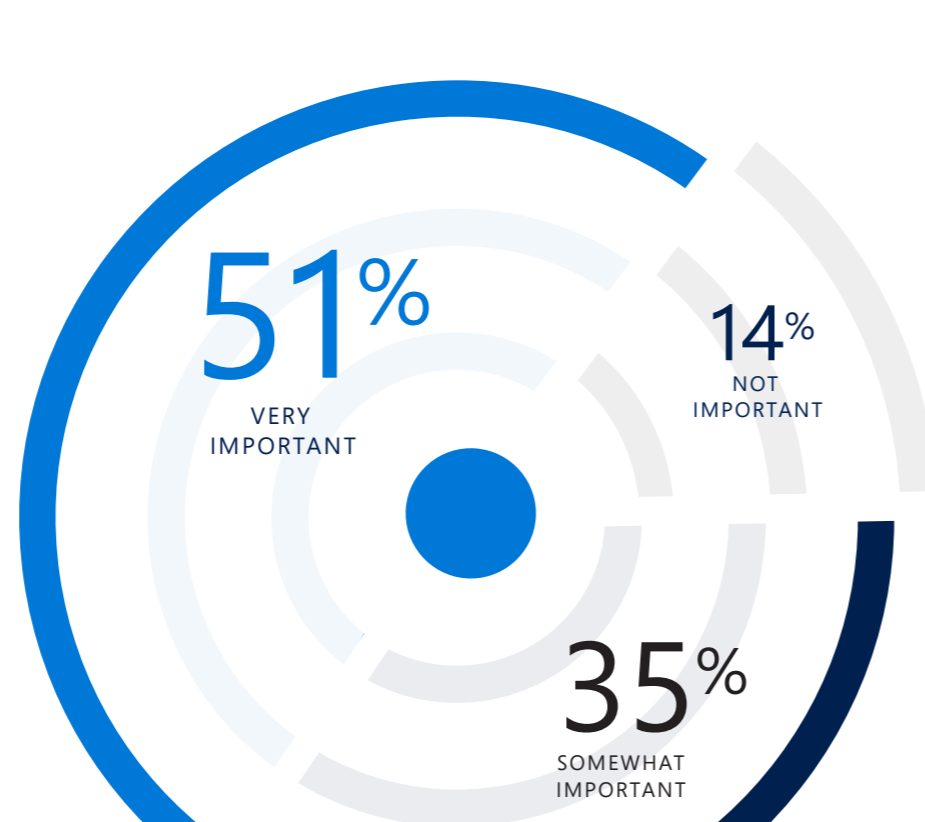


The rise of cloud computing is rapid, inexorable, and causing a huge upheaval in the tech industry. We observe this trend reflecting in Spark, as many respondents have moved to cloud, reaping its many benefits. Not only do cloud deployments have lower deployment costs and fewer management headaches, they have higher and proven performance benefits.

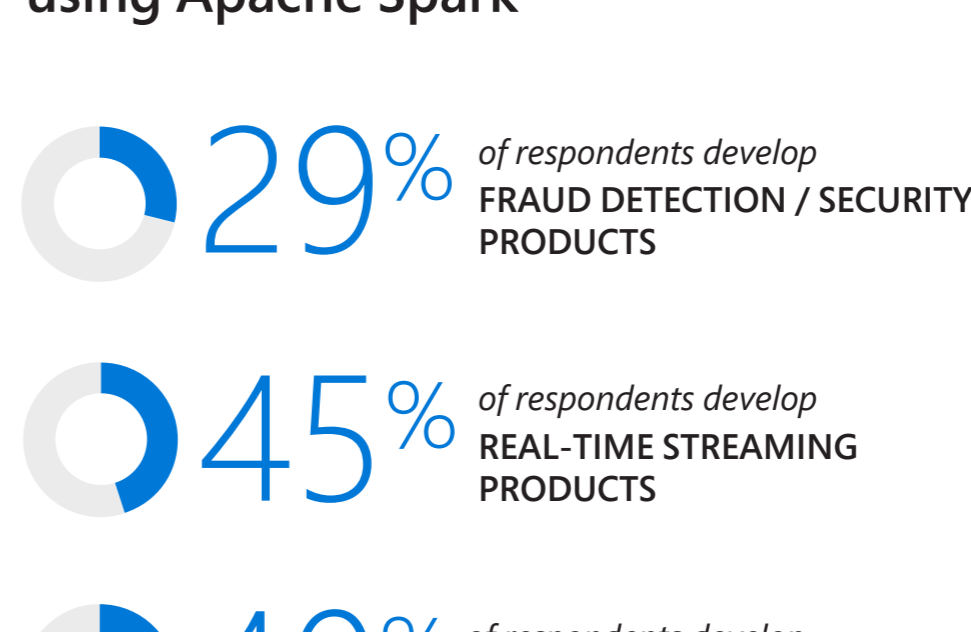
Apache Spark streaming and machine learning surge in usage

Apache Spark Streaming is growing. Since its release, Spark Streaming has become one of the most widely used distributed streaming engines. Interest in developing real-time applications and advanced analytics is on the rise.

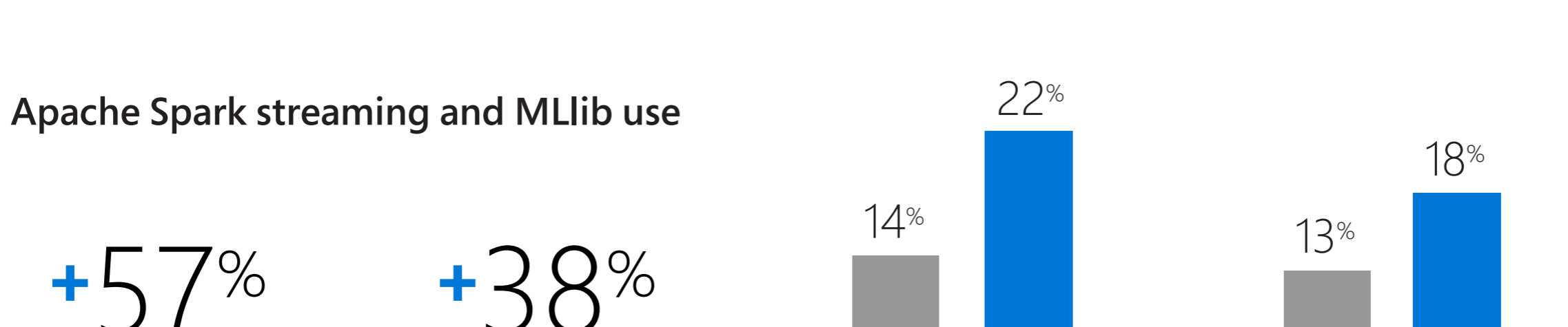
Importance of Apache Spark streaming to users



Types of products developed using Apache Spark



Apache Spark streaming and MLlib use



The numbers show that Spark Streaming is the preferred streaming engine for organizations building real-time, end-to-end streaming solutions, from evaluation to production.

By surveying our community, we're able to gather essential insights about how our users are leveraging the features of Apache Spark. To learn more on how you can accelerate innovation with the most advanced Apache Spark-based platform, visit us at azure.com/databricks

Azure Databricks

Data science, engineering, and business come together like never before with Microsoft Azure Databricks, the most advanced Apache Spark platform. With a high-performance processing engine that's optimized for Azure, you're able to improve and scale your analytics on a global scale—saving valuable time and money, while driving new insights and innovation for your organization. Microsoft Azure Databricks offers an intelligent, end-to-end solution for all your data and analytics challenges.



Copyright © 2017 Microsoft, Inc. All rights reserved. This infographic is for informational purposes only. Microsoft makes no warranties, express or implied, with respect to the information presented here.