2016 Big Data Maturity Survey

Brought to you by Trifacta & AtScale

TRIFACTA

atscale

Introduction

Now in its second year, the Big Data Maturity Survey 2016 (formerly the "Hadoop Maturity Survey") delivers a wealth of insight into Hadoop as it continues to be the fastest growing force in Big Data.

But how do you separate hype from reality? How do you find out what works and what doesn't? AtScale has partnered with Cloudera, Hortonworks, MapR, Tableau, Trifacta and Cognizant to identify companies that are working with Big Data or about to. We asked them how they got value from it, what tools they are using and the tactics they used to succeed.

With answers from over 2,550 participants worldwide, this survey is the first global and industry's largest study on Big Data Maturity. The study focuses on companies that have deployed Big Data, their best practices and their intentions.

Executive Summary

atscale

Big Data is growing fast

97% will do as much or more with Big Data over the next 3 months.

Big Data Cloud is King

72% of respondents plan on doing Big Data in the Cloud.

Governance is a growing concern

Governance is the fastest growing area of concern year-over-year (21% YOY).

Business Intelligence is #1

75% of respondents say they planning on using BI on Big Data.

A Deep and Wide View of the Market

atscale

More Value

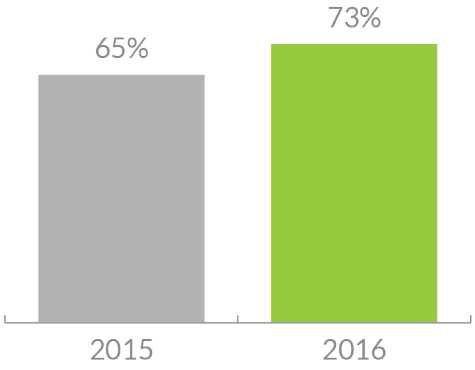
95% of respondents have achieved positive value with Big Data or are anticipating they will.

More People

Over 2,550 answers. 77 countries, over 1400 companies. 76% are using Hadoop today!

More Production

73% of respondents now in production (vs.65% last year).



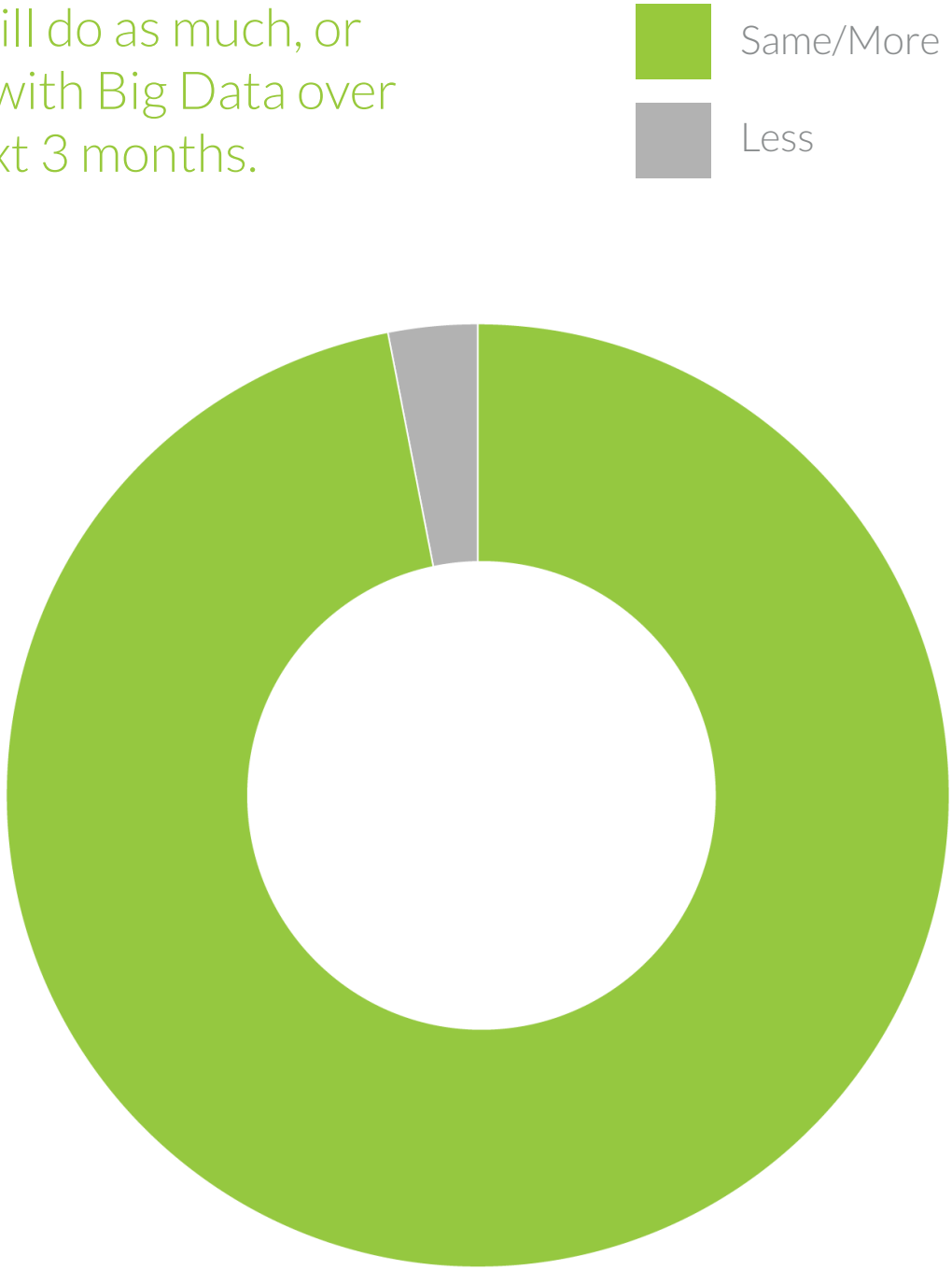
More Mature

1 in 5 Have Over 100 Nodes!

Audience Metrics

2017: Year of Big Data

97% will do as much, or more, with Big Data over the next 3 months.

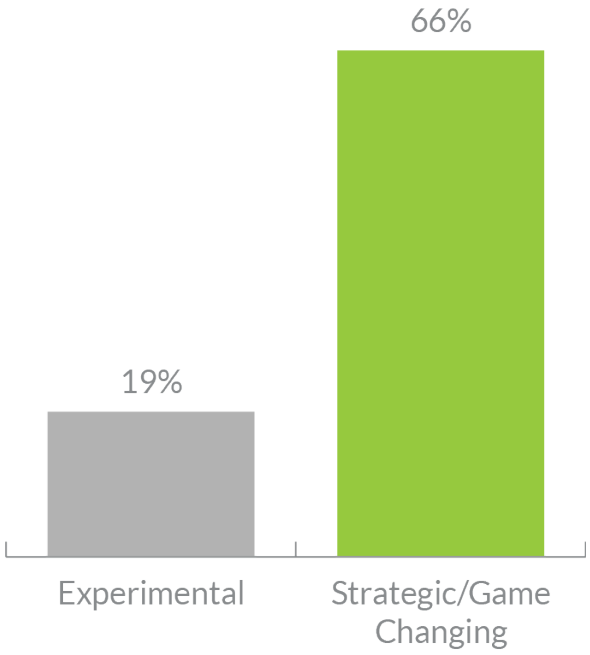


Big Data: No Longer an Experiment

atscale

Big Data is Strategic

Two-thirds (66%) view Big Data as "Strategic" or "Game Changing", while only 19% of respondents consider it experimental.



Big Data is No Longer an Experiment

Dramatically fewer described their usage of Big Data as "Experimental" this year versus last (a 37% year over year drop).

Big Data Value: "Scale-Out" is #1

atscale

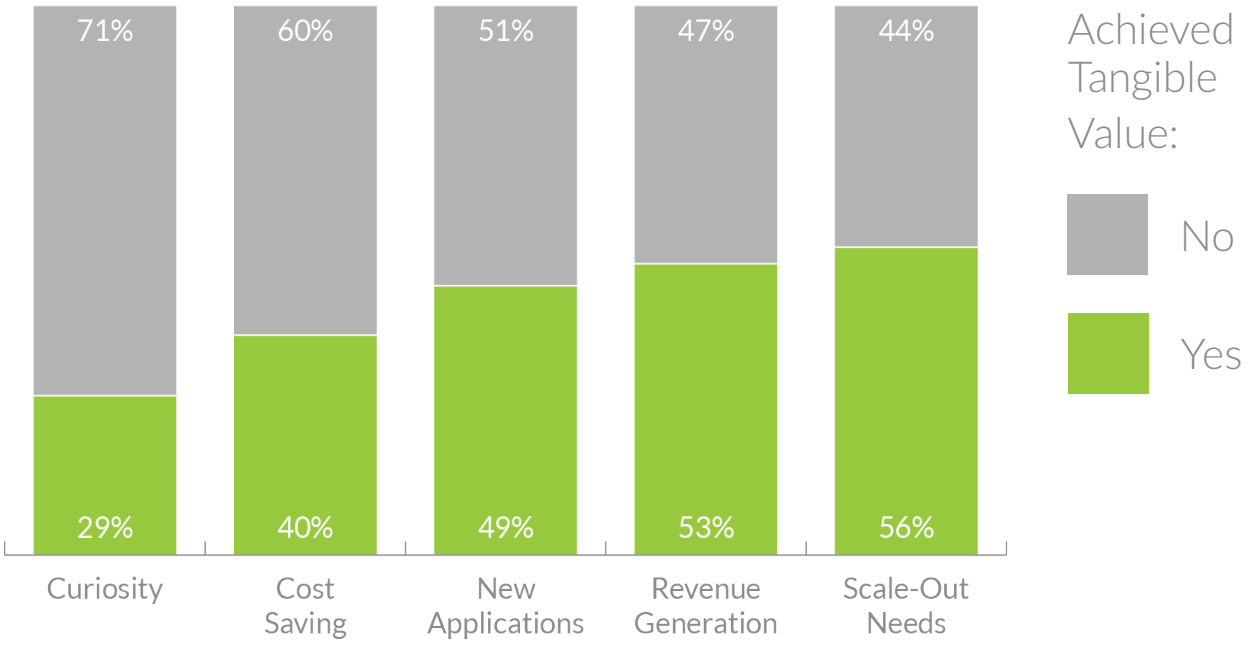
Best way to gain value from Big Data is to solve scale-out needs

Scale-out use cases are 56% more likely to yield tangible value.

Least effective way to gain value is to be motivated by curiosity or cost savings

Companies who use Big Data out of curiosity are least likely to succeed (29%).

Primary Big Data Use Cases



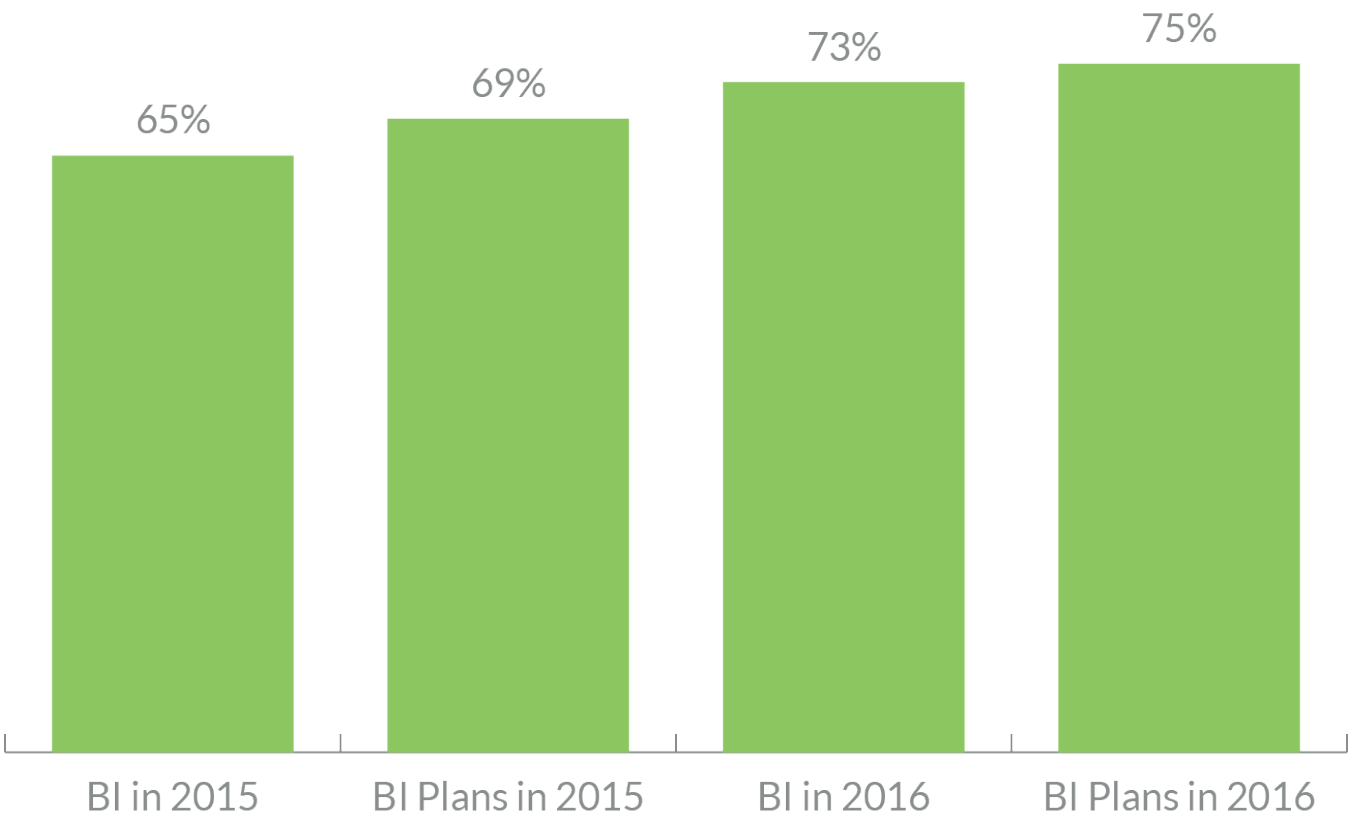
Business Intelligence

BI: #1 Use Case for Big Data

atscale

In 2015, BI overtook ETL as the #1 workload planned for Big Data.

This trend continues in 2016. Business Intelligence is #1 workload for Big Data with 75% of respondents planning on using BI on Big Data.



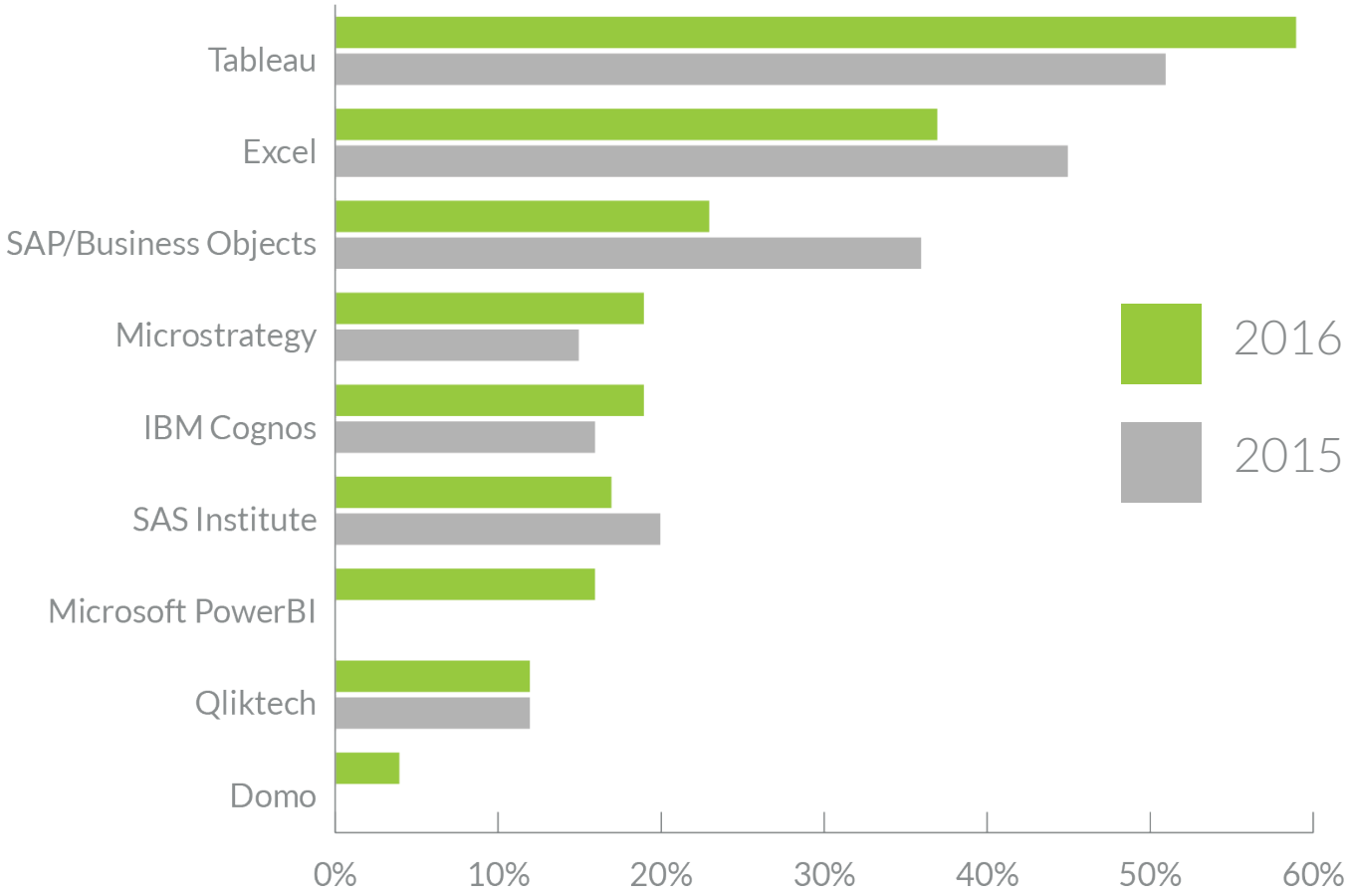
Multi-Dimensional & Visual on the Rise

atscale

Tableau continues to be the #1 BI Tool for Big Data.

Year-over-year, visual and multidimensional vendors have increased their share (MicroStrategy, Cognos) over Tabular models like Microsoft Excel.

BI Tools on Big Data



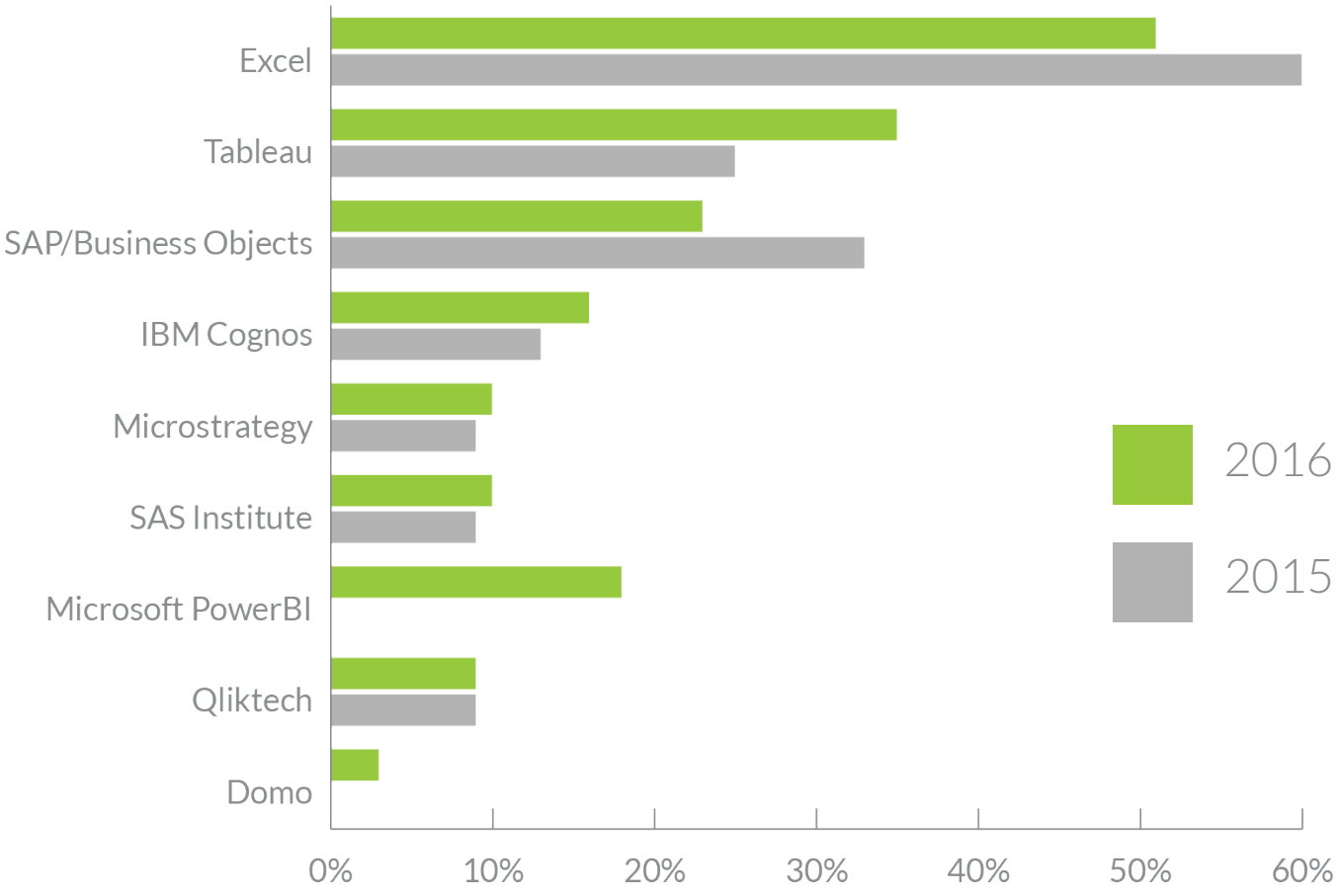
Microsoft Leads on Small Data

atscale

Microsoft Excel maintains the #1 place for BI Tool for Small Data.

New entrant PowerBI is helping Microsoft stay on top.

59% are planning on using Microsoft software (Excel, PowerBI or both).



Big Data in the Cloud

Cloud: 72% of Big Data will be in the Cloud

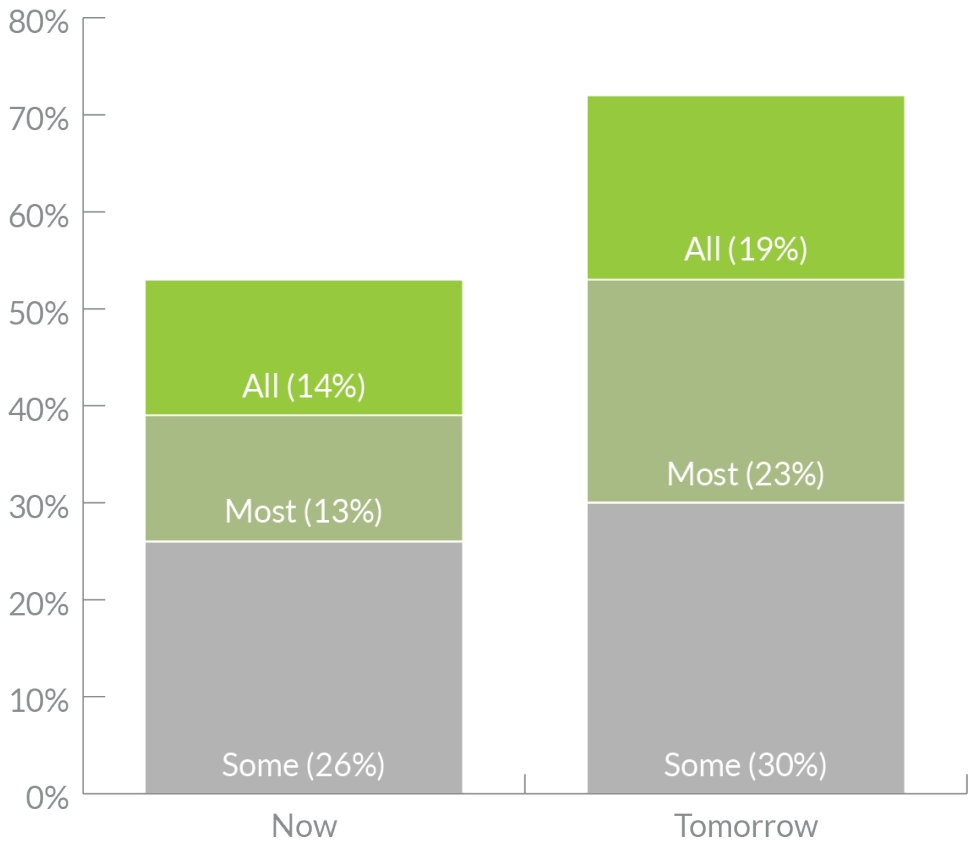
atscale

Now

- 53% are using Cloud for their Big Data deployment
- 14% have ALL of their Big Data in the Cloud
- 27% have more than half of their Big Data in the Cloud

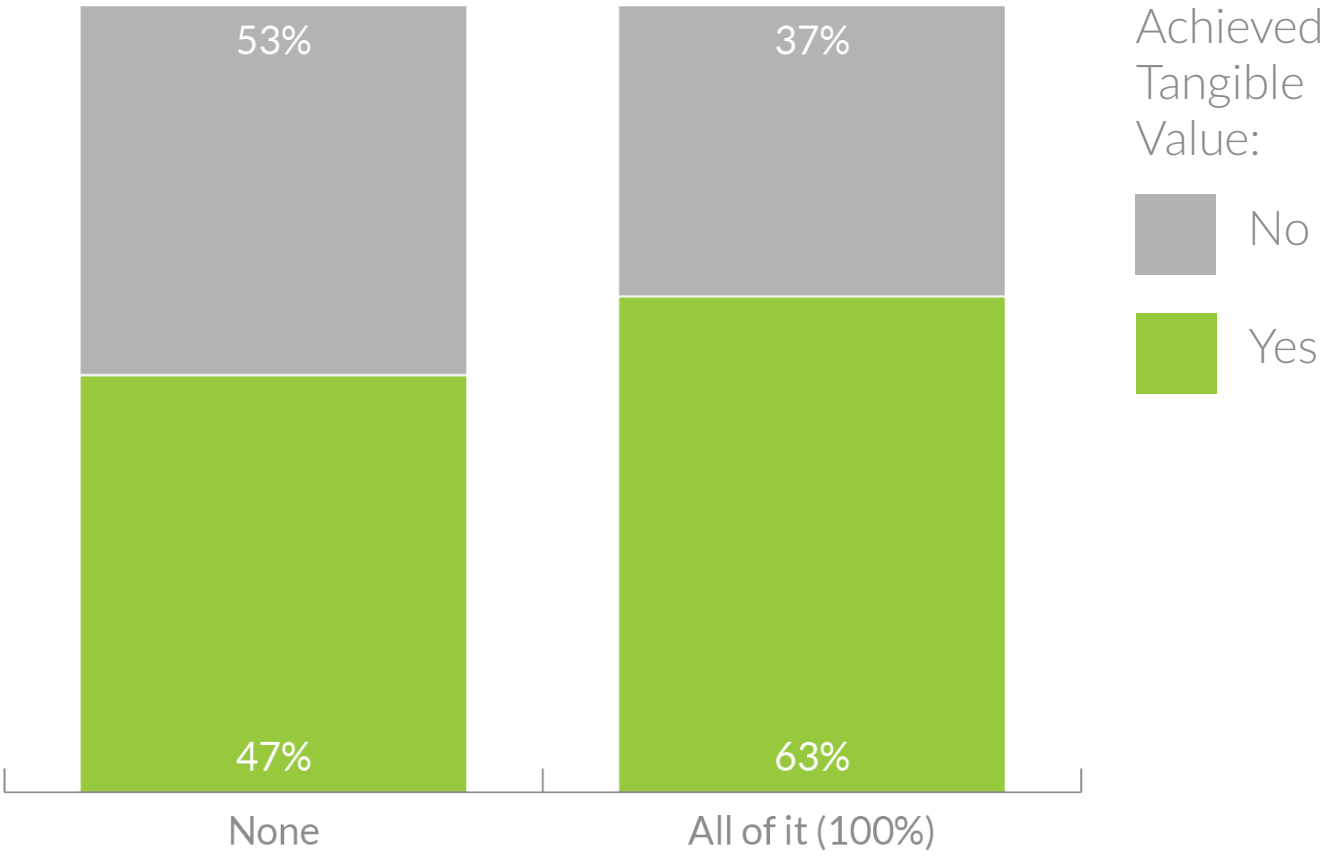
Tomorrow

- 72% plan on doing Big Data in the Cloud
- 42% of them plan on having more than half of their Big Data in the Cloud



All Cloud: 34% More Success

Respondents who have deployed all of their Big Data in the Cloud are 34% more likely to achieve tangible value, compared to those that have none.



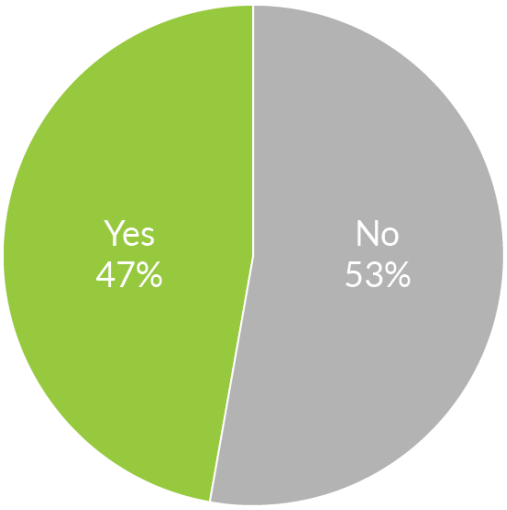
Governance & Self-Service

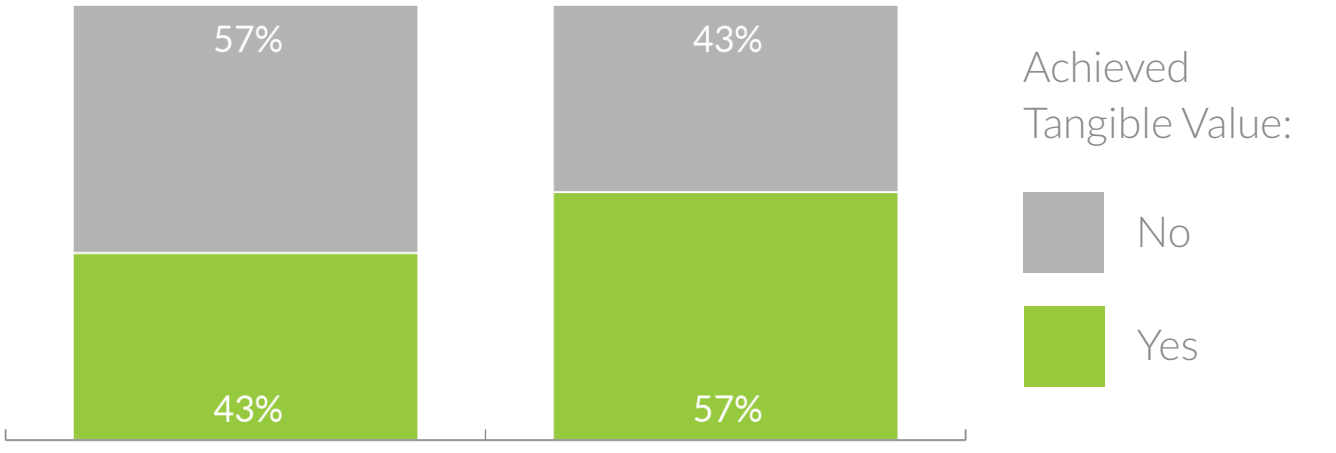
The State of Self-Service Access to Big Data

atscale

Self-Service Access to Hadoop

Year over year, self-service access on Big Data grew by 15%, however most companies (53%) still suffer from a lack of self-service access to Big Data.





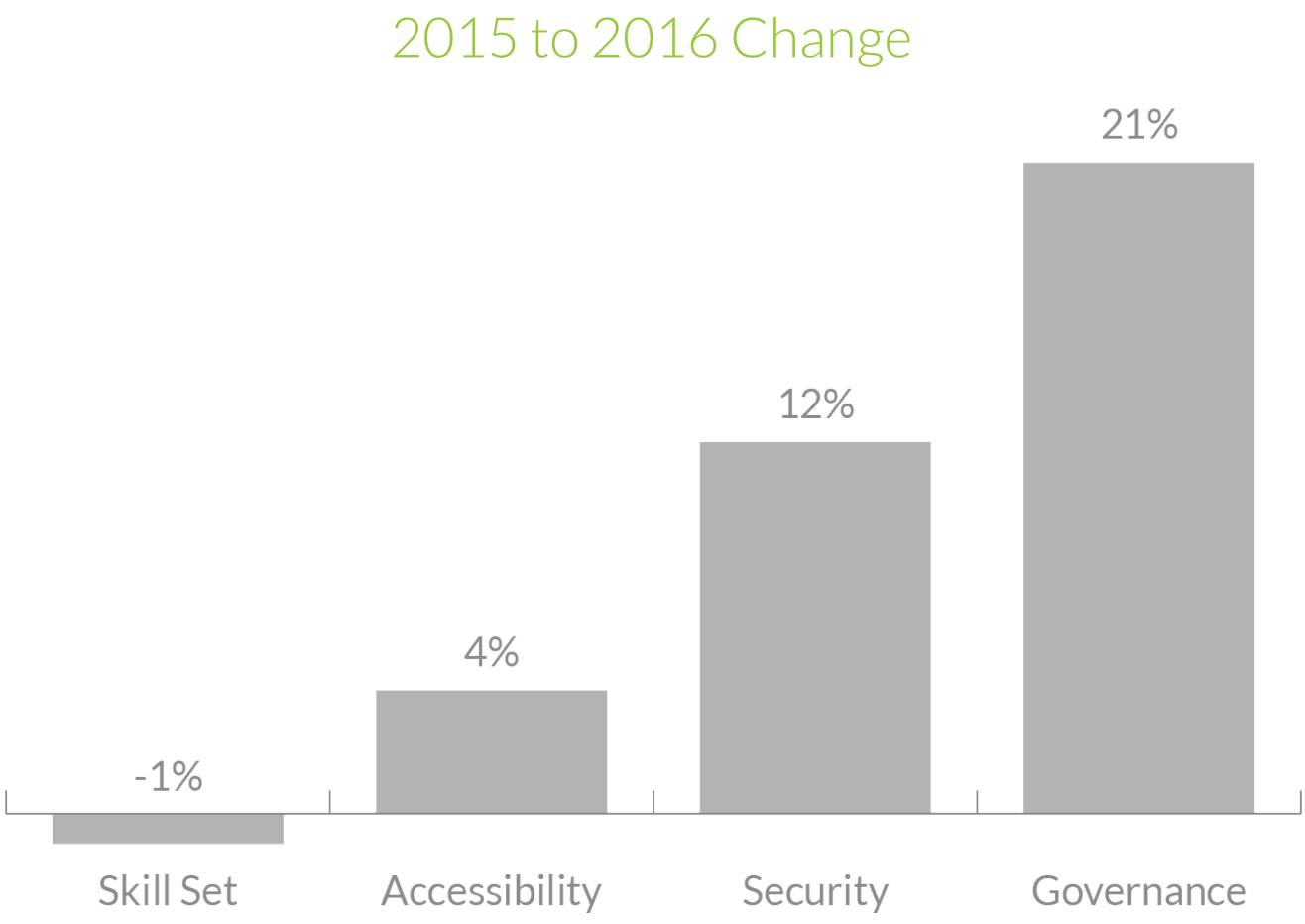
Do Business Units Have Access to Hadoop?

Self-service access is on of the best ways to drive value on Big Data. Companies who have self-service are 32% more likely to succeed.

Growing Concern: Governed Self-Service

While skillset continues to be the top barrier to providing self-service access, fewer respondents worry about it more in 2016.

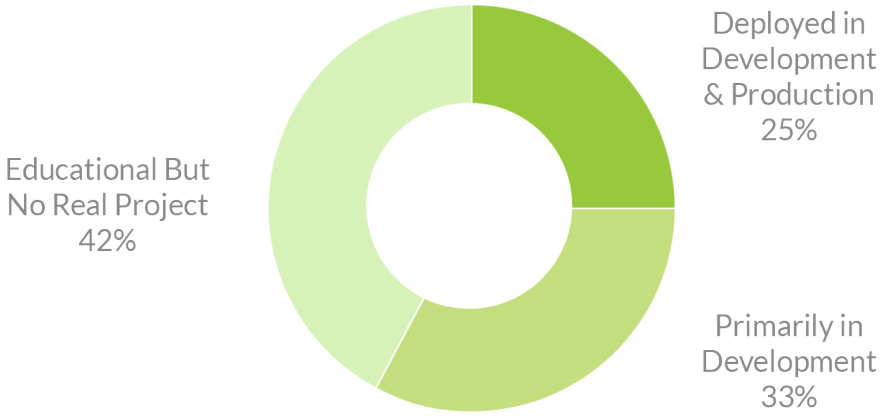
Accessibility, security and governance have become the fastest growing areas of concern year-over-year, with governance growing most at 21%.



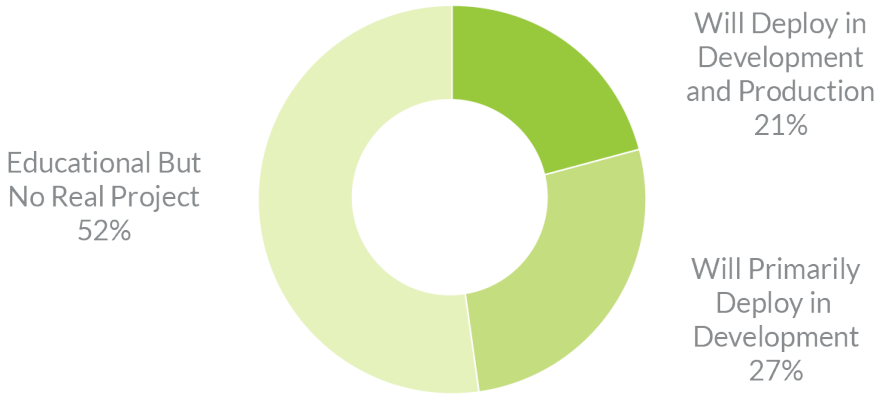
Spark & Spark SQL

Spark: Education & Development

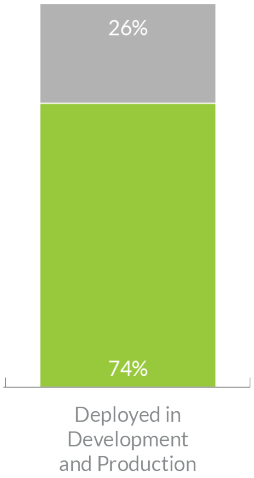
25% Have Spark in Production Today



21% Plan to Use Spark in Production



85% More Likely to Achieve Value When Deployed



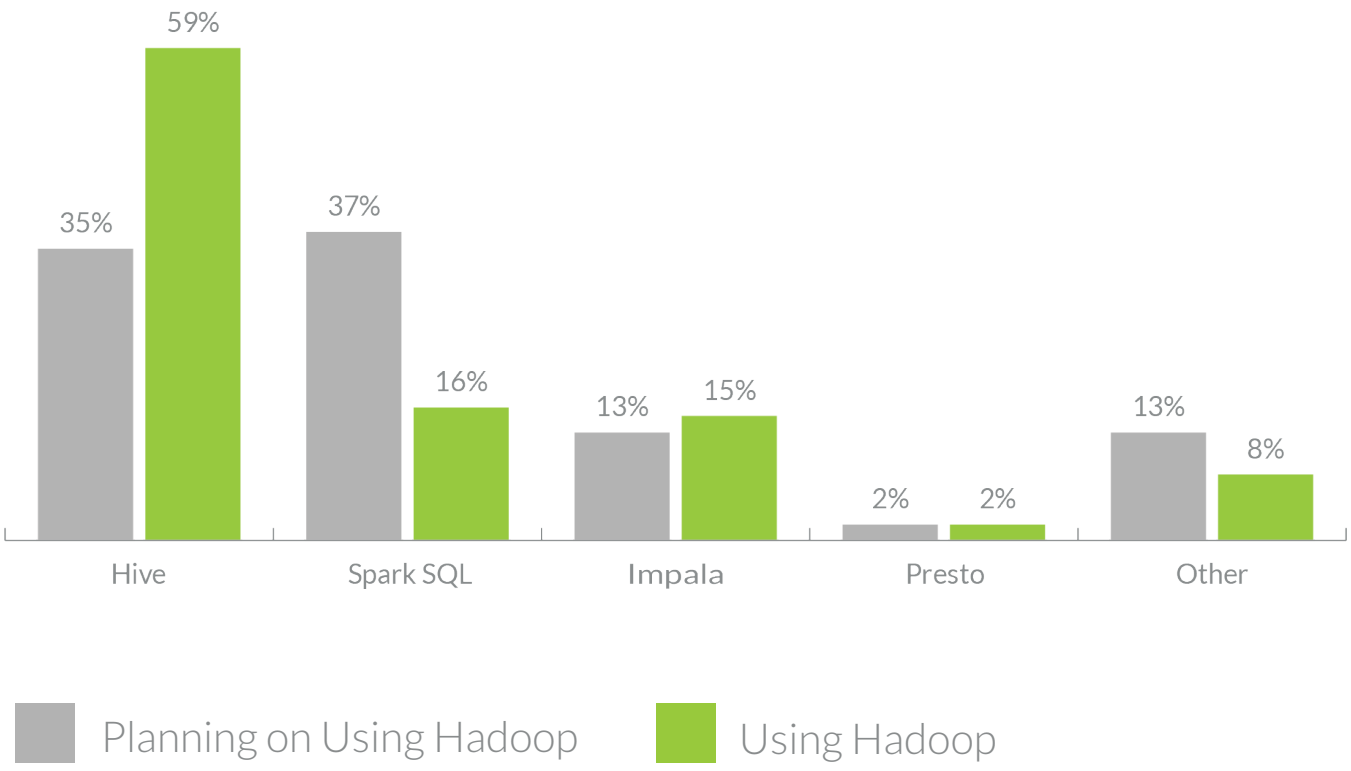
Achieved Tangible Value:

No Yes

Spark SQL: The Excitement is Real

atscale

As indicated in the latest SQL-on-Hadoop performance benchmarks, one engine doesn't fit all. Enterprises will continue to rely on more than one engine to accommodate their needs for reliability, concurrency and fast query response time on small and big data.



Enterprise Business & IT

The CIO's Future: A Hybrid Data Portfolio

atscale

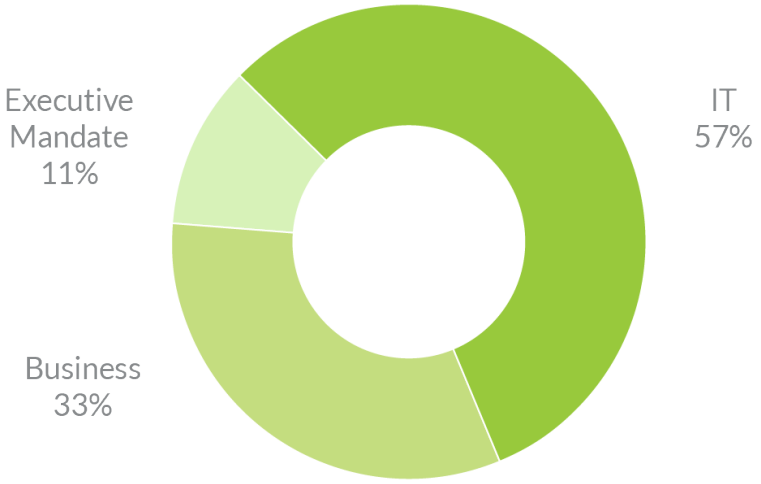
82% are looking to Big Data for EDW Optimization rather than a straight replacement.

The future of the CIO Data Portfolio will combine traditional data warehousing solutions (e.g. Teradata), contemporary data platforms (e.g. Hadoop) with Big Data services (e.g. Google BigQuery, Amazon Redshift and Microsoft Azure).

Executives Drive more Value for Big Data

atscale

Who Was the Primary Driver Behind Big Data at Your company?



IT is the primary driver behind Big Data's adoption (57%) but when executives drive Big Data, companies are 19% more likely to achieve tangible value.

Drivers Behind Big Data Adoption