TOP 8
TRENDS FOR 2016

BIG DATA

+†+ +tableau®

The year 2015 was an important one in the world of big data. What used to be hype became the norm as more businesses realized that data, in all forms and sizes, is critical to making the best possible decisions. In 2016, we'll see continued growth of systems that support non-relational or unstructured forms of data as well as massive volumes of data. These systems will evolve and mature to operate well inside of enterprise IT systems and standards. This will enable both business users and data scientists to fully realize the value of big data.

Each year at Tableau, we start a conversation about what's happening in the industry. The discussion drives our list of the top big-data trends for the following year. These are our predictions for 2016.

#datatrends16

The NoSQL Takeover

We noted the increasing adoption of NoSQL technologies, which are commonly associated with unstructured data, in last year's version of Trends in Big Data. Going forward, the shift to NoSQL databases becoming a leading piece of the Enterprise IT Landscape becomes clear as the benefits of schema-less database concepts become more pronounced. Nothing shows the picture more starkly than looking at Gartner's Magic Quadrant for Operational Database Management Systems which in the past was dominated by Oracle, IBM, Microsoft and SAP. In contrast, in the most recent Magic Quadrant, we see the NoSQL companies, including MongoDB, DataStax, Redis Labs, MarkLogic and Amazon Web Services (with DynamoDB), outnumbering the traditional database vendors in Gartner's Leaders quadrant of the report.

Additional Reading:

Magic Quadrant for Operational Database Management Systems

+a b | e a u

Apache Spark lights up Big Data

Apache Spark has moved from a being a component of the Hadoop ecosystem to the Big Data platform of choice for a number of enterprises. Spark provides dramatically increased data processing speed compared to Hadoop and is now the largest big data open source project, according to Spark originator and Databricks co-founder, Matei Zaharia. We see more and more compelling enterprise use cases around Spark, such as at Goldman Sachs where Spark has become the "lingua franca" of big data analytics.

Additional Reading:

Databricks Application Spotlight: Tableau Software

+ableau®

Hadoop projects mature!

Enterprises continue their move from Hadoop Proof of Concepts to Production

In a recent survey of 2,200 Hadoop customers, only 3% of respondents anticipate they will be doing less with Hadoop in the next 12 months. 76% of those who already use Hadoop plan on doing more within the next 3 months and finally, almost half of the companies that haven't deployed Hadoop say they will within the next 12 months. The same survey also found Tableau to be the leading BI tool for companies using or planning to use Hadoop, as well as those furthest along in Hadoop maturity.

Additional Reading:

AtScale's Hadoop Maturity Survey highlights Big Data's relentless growth

+ableau

Big Data grows up:
Hadoop adds to enterprise standards

As further evidence to the growing trend of Hadoop becoming a core part of the enterprise IT landscape, we'll see investment grow in the components surrounding enterprise systems such as security. Apache Sentry project provides a system for enforcing fine-grained, role based authorization to data and metadata stored on a Hadoop cluster. These are the types of capabilities that customers expect from their enterprise-grade RDBMS platforms and are now coming to the forefront of the emerging big data technologies, thus eliminating one more barrier to enterprise adoption.

Additional Reading:

Information Week - Cloudera Brings Role-Based Security To Hadoop

Big Data gets fast:

Options expand to add speed to Hadoop

With Hadoop gaining more traction in the enterprise, we see a growing demand from end users for the same fast data exploration capabilities they've come to expect from traditional data warehouses. To meet that end user demand, we see growing adoption of technologies such as Cloudera Impala, AtScale, Actian Vector and Jethro Data that enable the business user's old friend, the OLAP cube, for Hadoop — further blurring the lines behind the "traditional" BI concepts and the world of "Big Data".

Additional Reading:

5 Best Practices for Tableau & Hadoop

tableau

The number of options for *preparing* end users to discover all forms of data grows.

Self-service data preparation tools are exploding in popularity. This is in part due to the shift toward business-user-generated data discovery tools such as Tableau that reduce time to analyze data. Business users also want to be able to reduce the time and complexity of preparing data for analysis, something that is especially important in the world of big data when dealing with a variety of data types and formats. We've seen a host of innovation in this space from companies focused on end user data preparation for Big Data such as Alteryx, Trifacta, Paxata and Lavastorm while even seeing long established ETL leaders such as Informatica with their Rev product make heavy investments here.

Additional Reading:

Alteryx, Trifacta, Paxata, Lavastorm, Informatica

MPP Data Warehouse growth is heating up… in the cloud!

The "death" of the data warehouse has been overhyped for some time now, but it's no secret that growth in this segment of the market has been slowing. But we now see a major shift in the application of this technology to the cloud where Amazon led the way with an on-demand cloud data warehouse in Redshift. Redshift was AWS's fastest growing service but it now has competition from Google with BigQuery, offerings from long time data warehouse power players such as Microsoft (with Azure SQL Data Warehouse) and Teradata along with new start-ups such as Snowflake, winner of Strata + Hadoop World 2015 Startup Showcase, also gaining adoption in this space. Analysts cite 90% of companies who have adopted Hadoop will also keep their data warehouses and with these new cloud offerings, those customers can dynamically scale up or down the amount of storage and compute resources in the data warehouse relative to the larger amounts of information stored in their Hadoop data lake.

Additional Reading:

Cloud Data Warehouse Race Heats Up

The buzzwords converge!

IoT, Cloud and Big Data come together.

The technology is still in its early days, but the data from devices in the Internet of Things will become one of the "killer apps" for the cloud and a driver of petabyte scale data explosion. For this reason, we see leading cloud and data companies such as Google, Amazon Web Services and Microsoft bringing Internet of Things services to life where the data can move seamlessly to their cloud based analytics engines.

Additional Reading:

All the Things: Data Visualization in a World of Connected Devices

Tableau offers a revolutionary new approach
to business intelligence that allows you to quickly
connect, visualize and share your data with
a seamless experience from the PC to the iPad.

Go to www.tableau.com to learn more.

+ tableau®