Delivering Value with Big Data and Analytics

Fern Halper

VP and Senior Research Director for Advanced Analytics

TDWI

November 30, 2016

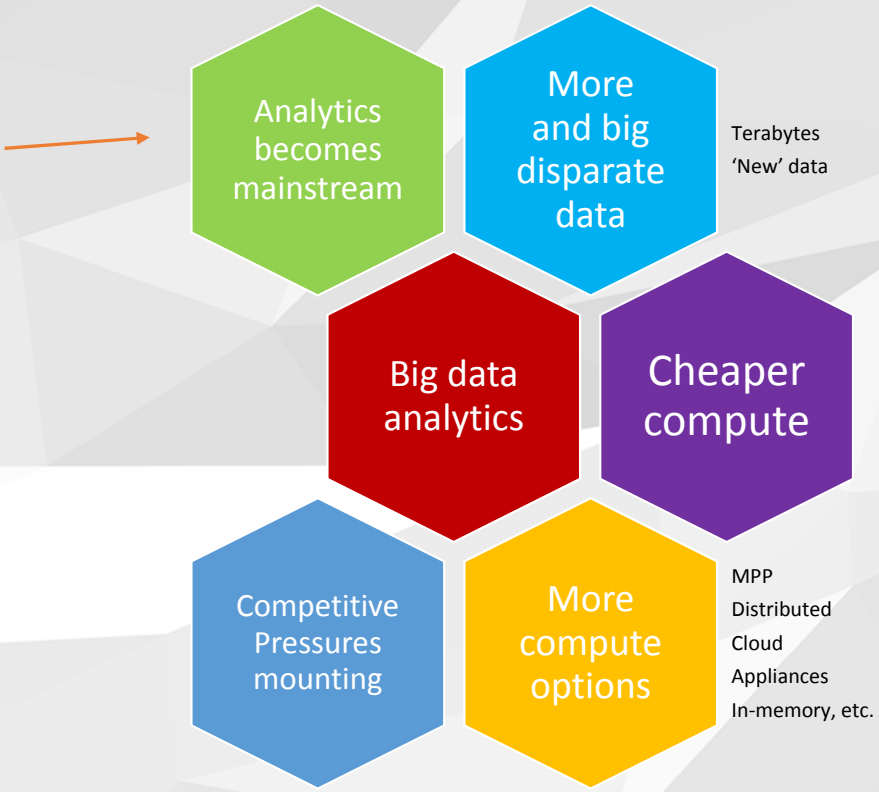
tdwi
Transforming Data With Intelligence™

My goals

- Paint a picture of what is happening in the world of big data and advanced analytics
  - Market trends and drivers
  - Adoption
  - Use cases for big data/advanced analytics
- Open source and Spark as an environment for advanced analytics
- Opinions of open source

BIG DATA TRENDS AND DRIVERS

Big data ecosystem trends



Drivers for big data analytics

60%
More accurate insights

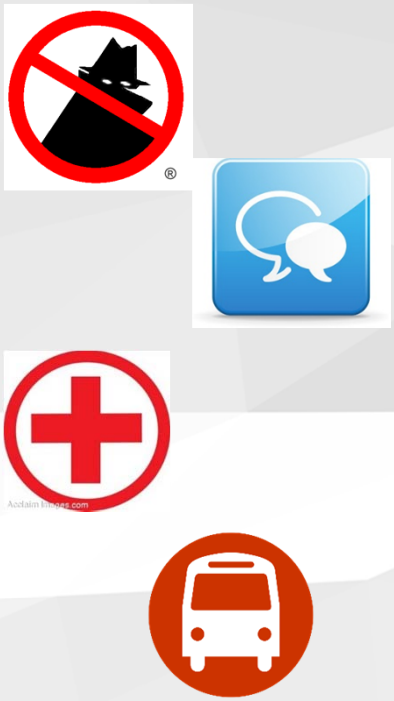
48%
Understand customers

40%
Improve business processes

Advanced analytics approaches

- Predictive analytics including machine learning
  - Systems that can learn from data to identify patterns and predict future results with minimal human intervention
- Natural language processing
  - Analyzing, understanding, and generating languages to ultimately enable interfacing with systems using human language
- Streaming analytics
  - Analyzing data in motion, often in real time
- Graph analytics
  - Models relationships between people, entities in a network. Many-many
- Open source options too!

Many use cases for big data and analytics



- Fraud and risk analysis
  - portfolio and investment risk and credit card fraud detection
- Customer-related analytics
  - churn and retention analysis, up-sell, next best offer, customer sentiment, customer loyalty, recommendation engines, customer journey
- Patient analytics
  - proactive patient health engagement, population health analysis, patient readmission
- Asset management
  - Predictive maintenance, asset optimization

Ex 1: Next Best Offer enabled by machine learning

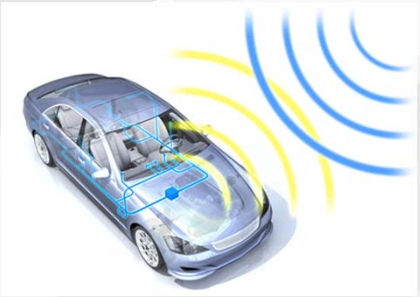
- **Goal:**
  - provide customers with the offer (e.g. a product, promotion, etc.) that is most relevant to them
- **Analyze: Machine learning**
  - A learning model is trained based on historical behavior of how customers with similar characteristics responded to an offer

Ex. 2: VOC enabled by NLP

- **Goal:**
  - Understand customer opinions and preferences
  - understand what customers think of products and what they like and don't like about them.
- **Analyze: NLP**
  - analyze emails, tweets, reviews, etc. using NLP to understand what the people are talking about and their sentiment

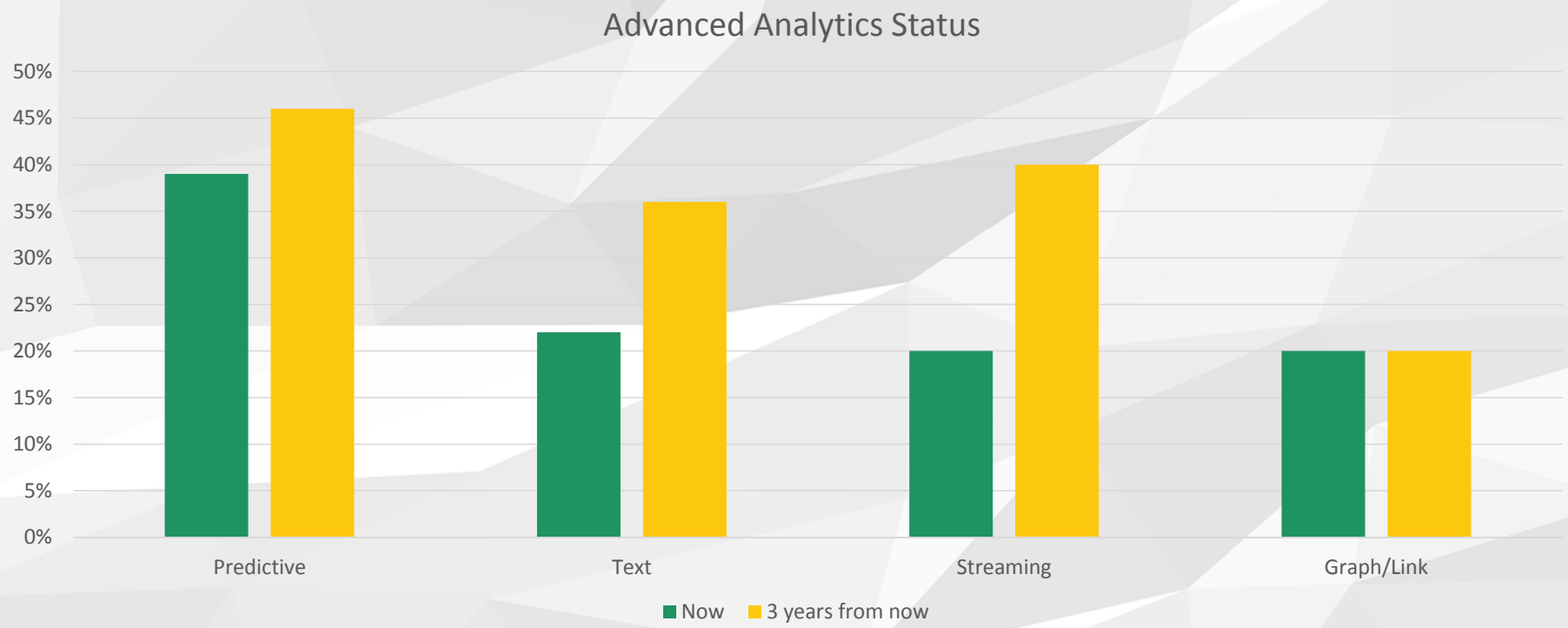
Ex. 3: Predicting insurance risk enabled by predictive and streaming analytics

- **Goal:**
  - Understand risk associated with drivers to better price premiums
- **Analyze: predictive and streaming**
  - Analyze telematics data to understand driving patterns including excessive breaking, etc.
  - Build models based on observed patterns



tdwi
Transforming Data With Intelligence™

Some analytics becoming mainstream



(source: TDWI Best Practices Report, Next Generation Analytics, 2015)

ADD OPEN SOURCE TO MIX

Why open source?

- Source code freely available
- A community of innovation
- A community of support, but not formal support
- Analytics
  - R
  - Spark

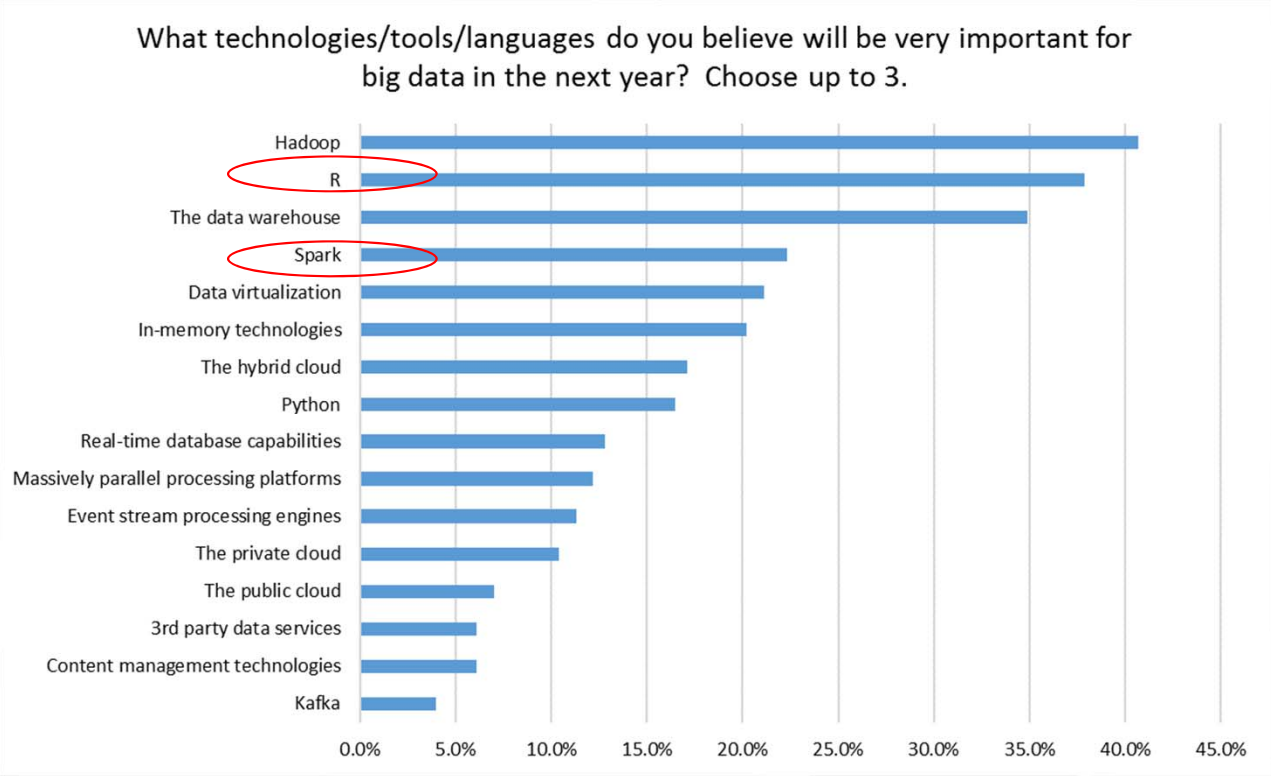
Open source R

- Been around for several decades
- A language and environment for statistical analysis
- Includes data handling and storage facilities, a large set of tools for data analysis (including machine learning and NLP), tools for graphical analysis, and its programming environment
- Vendors now supporting it

Spark

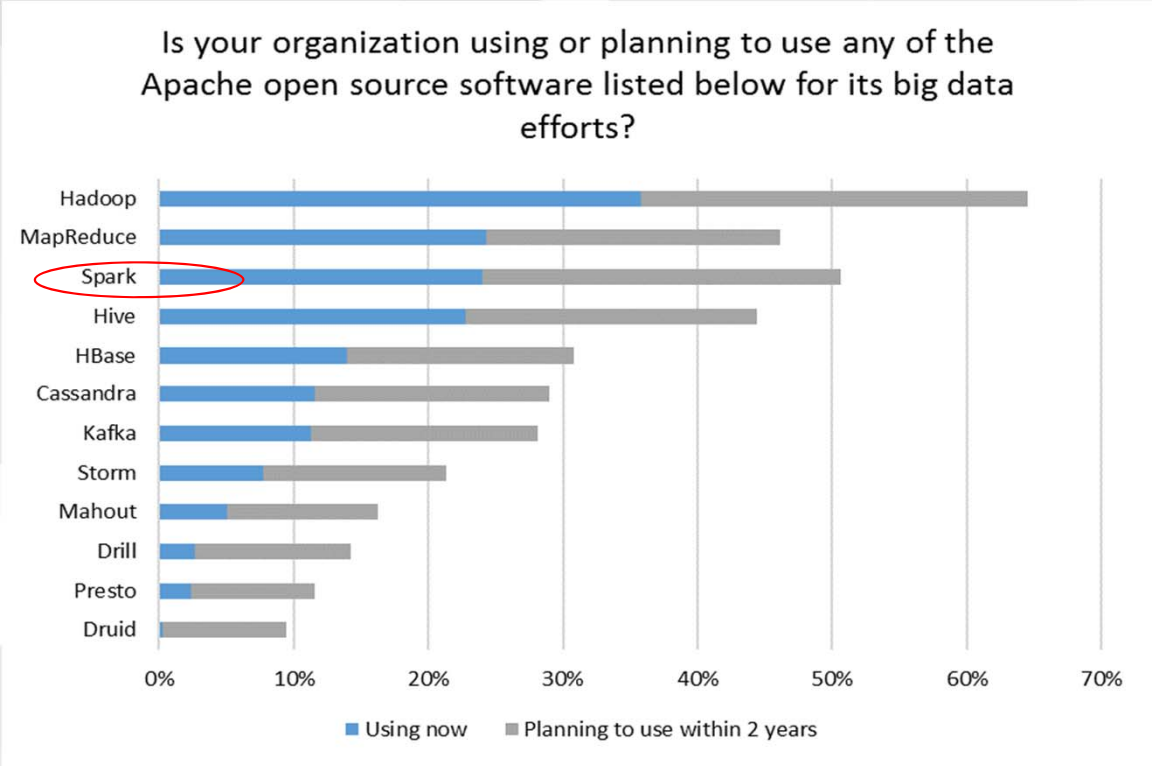
An open source big data processing framework that provides processing capabilities for multiple kinds of big data. Spark also offers analytics libraries, including a machine learning library

Increased interest in open source



(TDWI BPR on Data Science, 2016. n=338)

Popular Apache open source technologies

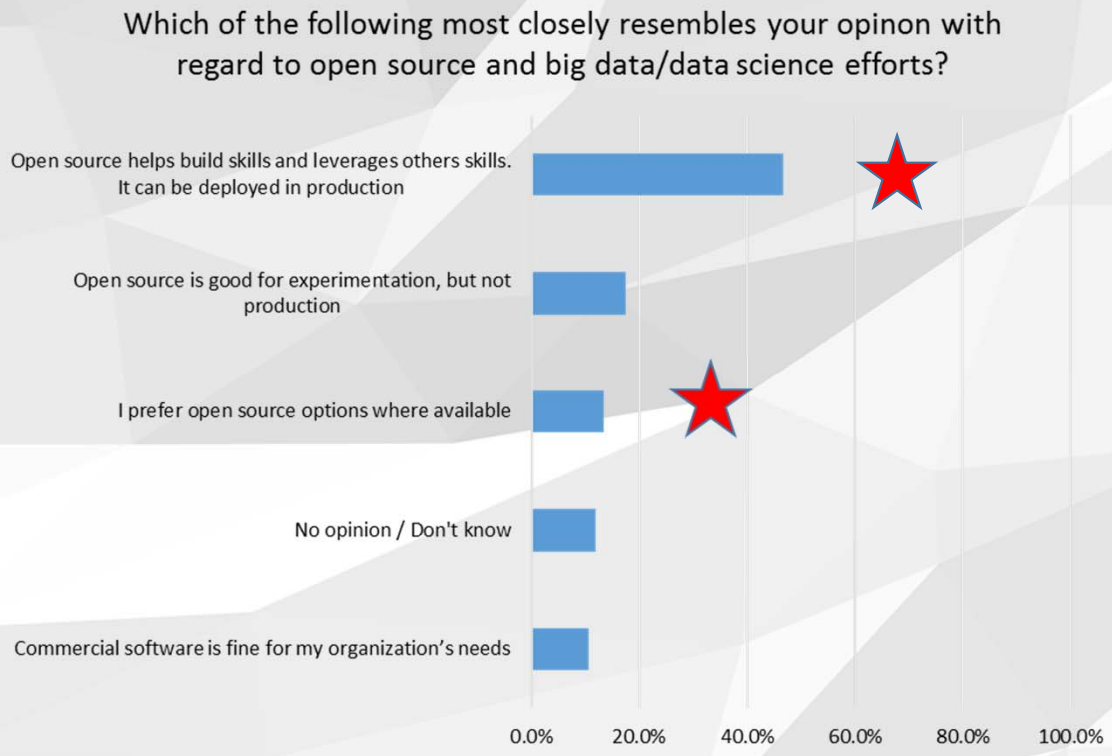


(TDWI BPR on Data Science, 2016. n=338)

What's so great about Spark?

- **Processes data in memory- fast**
- **A unified engine with support for SQL queries (ANSI standard), streaming data, machine learning and graph processing**
- **Makes it easy for development – includes** filter, join and group-by
- Broadly compatible – HIVE, ODBC/JDBC, Etc.**- Flexible** deployment too (IBM, others)

Open source opinions



(TDWI BPR on Data Science, 2016. n=338)

Recommendations for value

- Analyze disparate data types
- Use more advanced analytics
  - Get the training needed!
- Consider open source
- Leverage what you already have in place

Learn more! 2016 Best Practices Report From TDWI on
Big Data and Data Science

- The report discusses best practices for data science
- Download the free report in a PDF file at: www.tdwi.org in December.
- TDWI Leadership Summit in Vegas focuses on this topic, too

tdwi
Transforming Data With Intelligence™

Learn more!


< TDWI HOME
tdwi
CONFERENCES
Agenda
Pricing
Travel & Accommodations
Exhibit & Partners
Why TDWI?
REGISTER
TDWI LAS VEGAS
DATA INSIGHT FOR THE ENTERPRISE
View Agenda
CAESARS PALACE
FEBRUARY 12–17, 2017
WELCOME
TO Fabulous
LAS VEGAS
NEVADA