

# Master of Technology in Knowledge Engineering

## Data Mining Tools Survey and Demo

Dr. Zhu Fangming  
Institute of Systems Science,  
National University of Singapore.  
E-mail: [isszfm@nus.edu.sg](mailto:isszfm@nus.edu.sg)

© 2018 NUS. The contents contained in this document may not be reproduced in any form or by any means, without the written permission of NUS ISS, other than for the purpose for which it has been supplied.

# Agenda for Day1

---

- Data Mining – What and Why?
- Applications Overview
- Identifying and planning data mining projects
- Data Mining Planning Workshop
- Data Mining Tools Overview/Demo



- About 100 software suites listed!

- Some Popular Commercial Tools

- SAS (Enterprise Guide/Miner)
- IBM SPSS Modeler
- Angoss Knowledge Studio
- Salford Systems (CART, MARS)

- Popular Freeware

- R
- Rattle (R)
- Python
- Weka
- RapidMiner
- KNIME
- Orange

KDNuggets™ Subscribe to KDNuggets News | [Twitter](#) [Facebook](#) [LinkedIn](#) | [Contact](#)

SOFTWARE | News/Blog | Top stories | Opinions | Tutorials | JOBS | Companies | Courses | Datasets | EDUCATION | Certificates | Meetings | Webinars

**To explain or to predict?** > Watch the webcast  
That is the question. Author of *Data Mining for Business Analytics* says understanding the difference is crucial. *jmp* Statistical Discovery • From SAS

To explain or to predict? Author of *Data Mining for Business* says the difference is crucial

**Machine Learning, Data Science, Data Mining, Big Data, Analytics, AI**

[Software](#) (Suites, Text, Visualization) | [Jobs - Industry](#) | [Academic](#) | [Meetings, Conferences](#) | [Companies](#) (Consulting, Products) | [Courses in Big Data, Data Science](#) | [Datasets](#) (APIs/Markets, Gov) | [Data Mining Course](#) | [Gregory Piatetsky](#) | [Education](#) (online, USA, Europe, cert) | [FAQ](#) | [Polls](#) | [Publications](#) (Books) | [Solutions](#) (Fraud, Data Cleaning) | [Webcasts](#) | [Websites](#) (Blogs, Cartoons, Podcasts)

**Most Recent**

- Apple: Apple Media Products Commerce Engineering Manager
- Top tweets, Jan 3-9: A collection of Jupyter notebook...
- Top 10 TED Talks for Data Scientists and Machine Learning
- Regularization in Machine Learning
- How Docker Can Help You Become A More Effective Data Scientist
- KDNuggets 18:n02, Jan 10: Quantum Machine Learning: AI...

Learn everything from Predictive Analytics to autonomous vehicles  
**ODSC EAST, May 1-4, Boston**  
SAVE 65%  
ODSC East, Save 65%

Latest poll results: Artificial General Intelligence (AGI) in less than 50 years, say KDNuggets readers

**Latest**  
[News](#) | [Software](#) | [Tutorials](#)

Subscribe to KDNuggets News [Twitter](#) [Facebook](#) [LinkedIn](#)

**Northwestern**  
DATA SCIENCE  
School of Professional Studies  
Online MS in Data Science

SEE THE BEAUTY BEHIND BIG DATA  
M.S. IN DATA SCIENCE  
REGIS UNIVERSITY  
Regis MS in Data Science

MS IN ANALYTICS at the University of San Francisco  
USF  
Big Data Requires Big Skills  
MS in Analytics at USF - Big Data Requires Big Skills

MS IN APPLIED DATA SCIENCE  
ENROLLING NOW FOR FEBRUARY 26th  
100% ONLINE  
BAY PATH UNIVERSITY

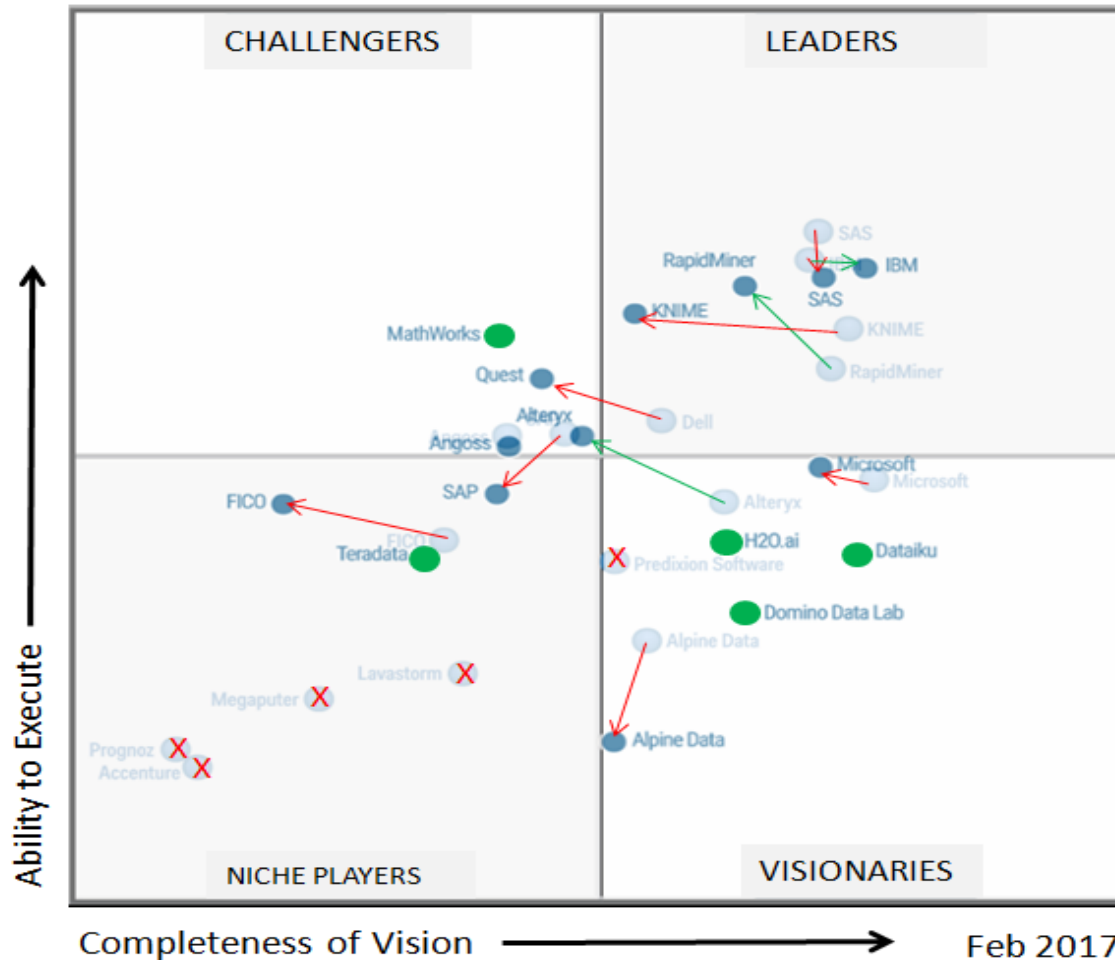
# Gartner Research : Advanced Analytics Platforms

Gartner 2017 Magic Quadrant for Advanced Analytics Platforms



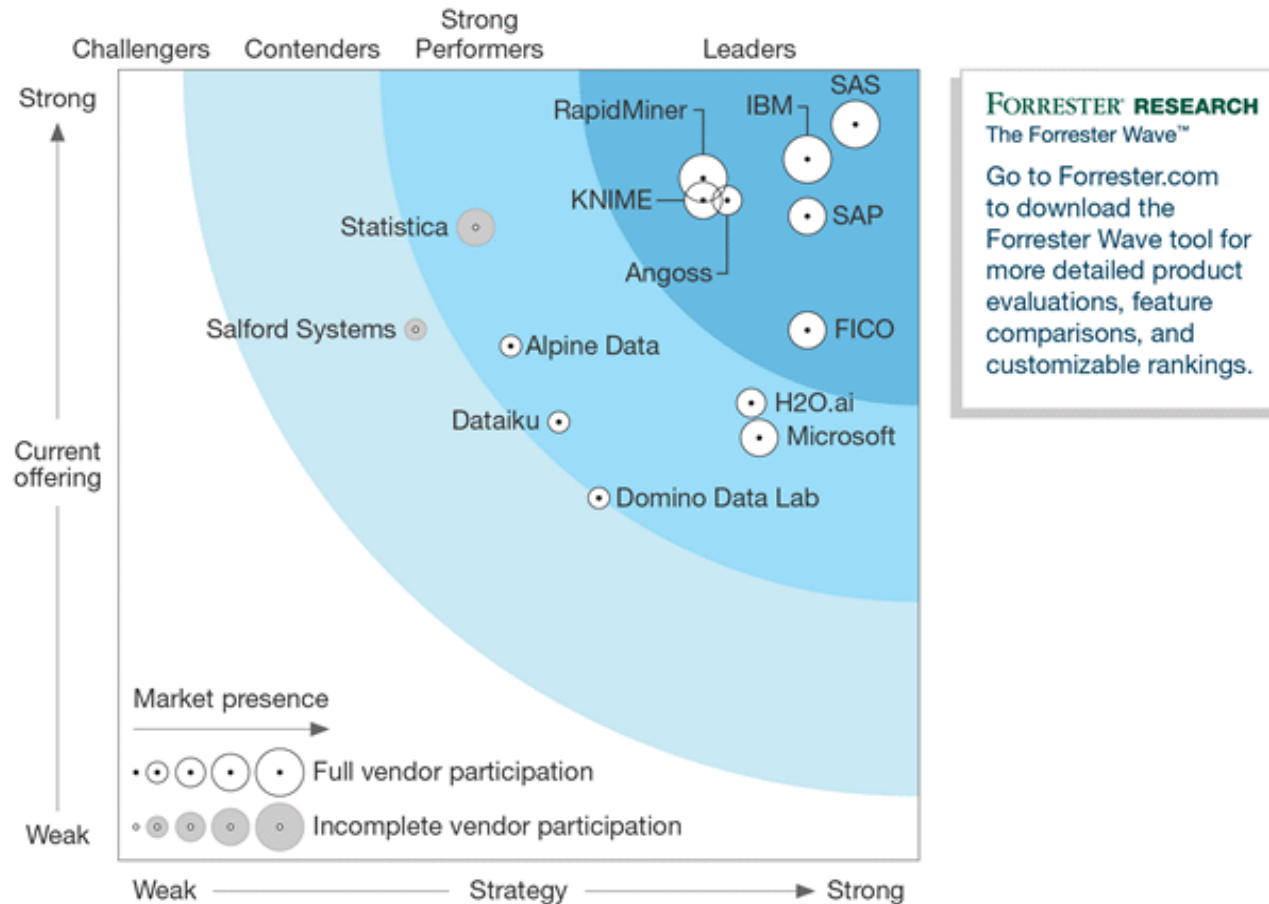
# Gartner Research : Advanced Analytics Platforms

Gartner Magic Quadrants for Advanced Analytics Platforms compared, 2016 vs 2017



# Forrester Research: Predictive Analytics And Machine Learning

## The Forrester Wave™: Predictive Analytics And Machine Learning Solutions, Q1 2017




129452

Source: Forrester Research, Inc. Unauthorized reproduction, citation, or distribution prohibited.






# The Tools Surveyed (Forrester)


Vendor	Product name
Alpine Data	Chorus
Angoss	KnowledgeSTUDIO
Dataiku	Dataiku DSS
Domino Data Lab	Domino
FICO	FICO Decision Management Suite
H2O.ai	H2O
IBM	IBM SPSS Modeler
KNIME	KNIME Analytics Platform
Microsoft	Cortana Intelligence Suite, Microsoft R Server
RapidMiner	RapidMiner Platform
Salford Systems	The SPM Salford Predictive Modeler software suite
SAP	SAP BusinessObjects Predictive Analytics, SAP HANA SPS
SAS	SAS Analytics Suite
Statistica	Statistica



# DATA SCIENCE WARS

 python

R and Python are waging war:  
while both programming languages are gaining prominence  
in the data analytics community, they are fighting  
to become data scientists' language of choice.  
Which side are you taking?



## Introducing The Opponents

Current Version

3.1.3  
March 2015

3.4.3 / 2.7.9  
February 2015/ December 2014

<http://blog.datacamp.com/r-or-python-for-data-analysis>