

A Decade of Surveying Analytic Professionals: 2017 Survey Highlights

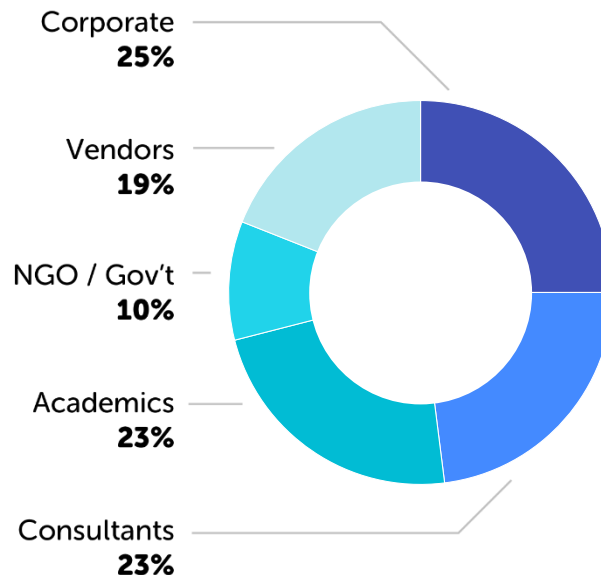
Karl Rexer, PhD
President, Rexer Analytics

www.RexerAnalytics.com

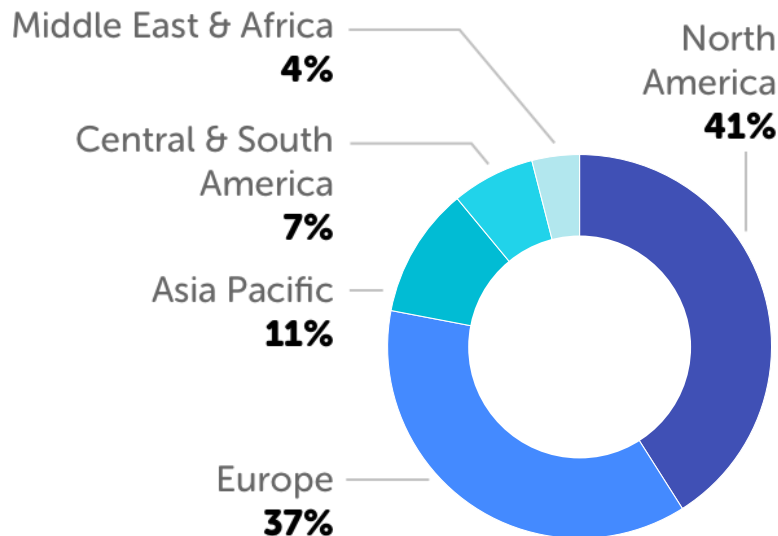
2017 Data Science Survey: Overview



- 8th survey since 2007
- 67 questions
- 10,000+ invitations emailed & promoted by newsgroups, vendors and bloggers
- Respondents: 1,123 analytic professionals from 91 countries
- Data collected in first half of 2017



- 8th survey since 2007
- 67 questions
- 10,000+ invitations emailed & promoted by newsgroups, vendors and bloggers
- Respondents: 1,123 analytic professionals from 91 countries
- Data collected in first half of 2017



Formal Data Science Training is Important



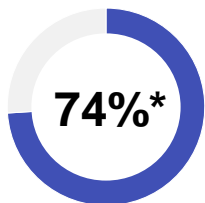
Do you agree or disagree with the following statement? You need to have formal training in data analytics in order to properly model data.

The majority of respondents agree that formal data training is needed to properly model data

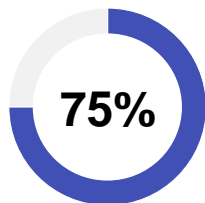
Top problems untrained staff have:

- Poor data preparation
- Mis-interpreting results

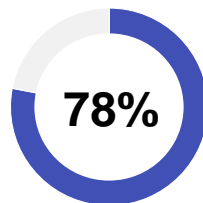
Corporate



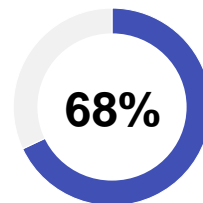
Consultants



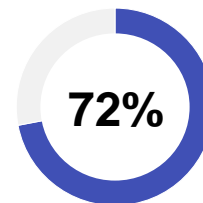
Academics



NGO / Gov't



Vendors



Do-it-yourself Tools Can Cause Problems



Have there ever been difficulties when people at your company outside the data science team have utilized these do-it-yourself tools or services?

One third of respondents have seen difficulties when do-it-yourself tools or services are used*

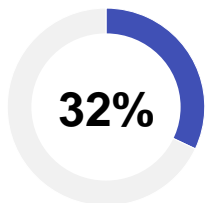
"Typically, they blow up the ship"

"People have run analyses that resulted in completely false conclusions, which were then used by the business"

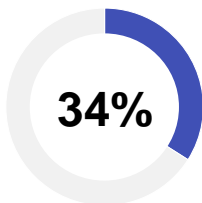
"A large part of my job is cleaning up those problems"

"Its like giving a child a bazooka. Easy to pull the trigger, hard to know what will happen in advance."

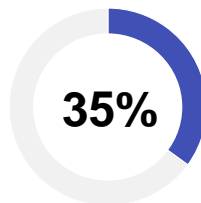
Corporate



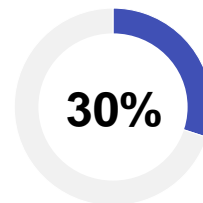
Consultants



NGO / Gov't



Vendors

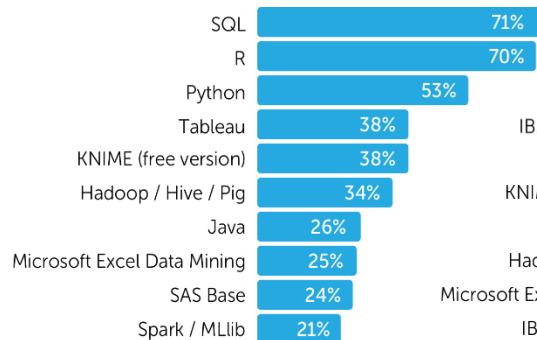


Most Data Scientists use Multiple Tools

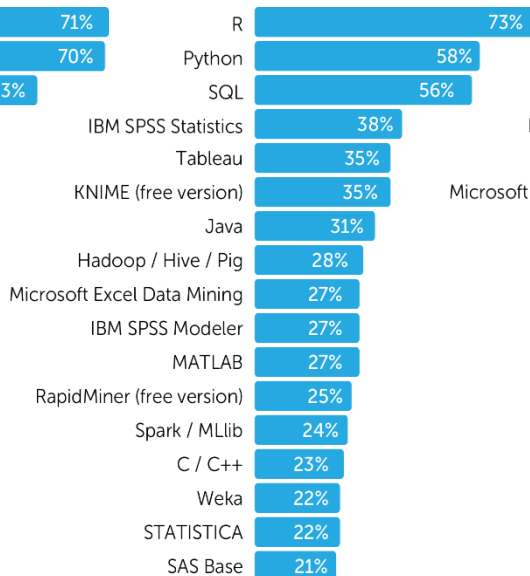


What data science / analytic tools, technologies, and languages did you use in the past year?

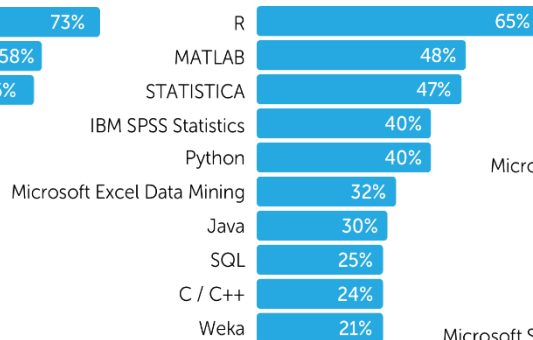
Corporate



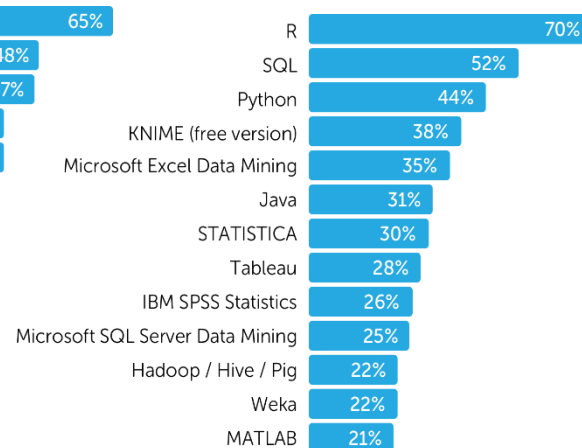
Consultants



Academics



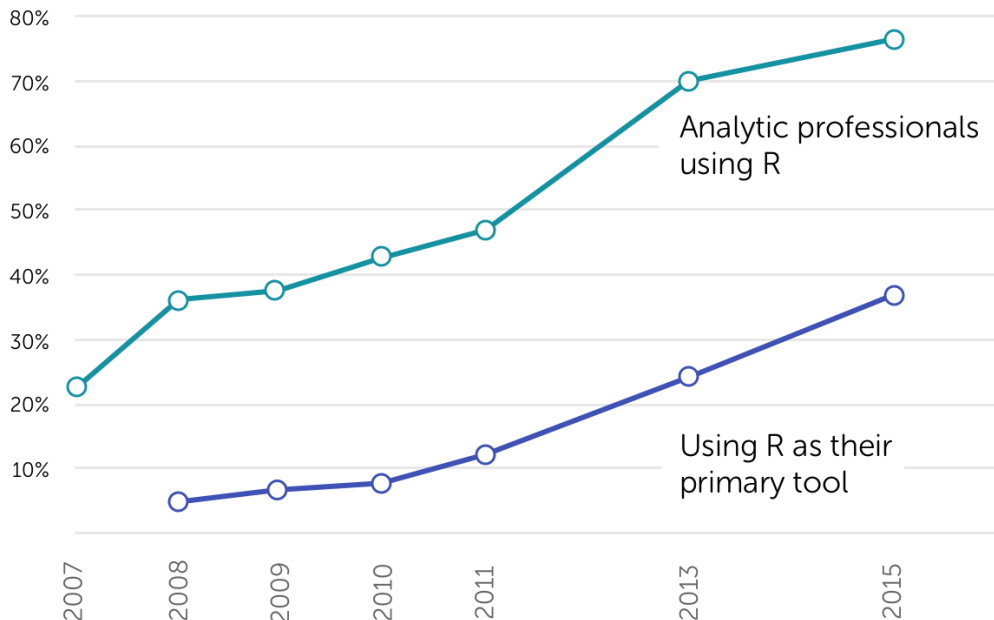
NGO / Gov't



A Decade of Dramatic Growth in the Use of R



The biggest change in tool adoption we've seen over the past decade has been the dramatic growth in the use of R

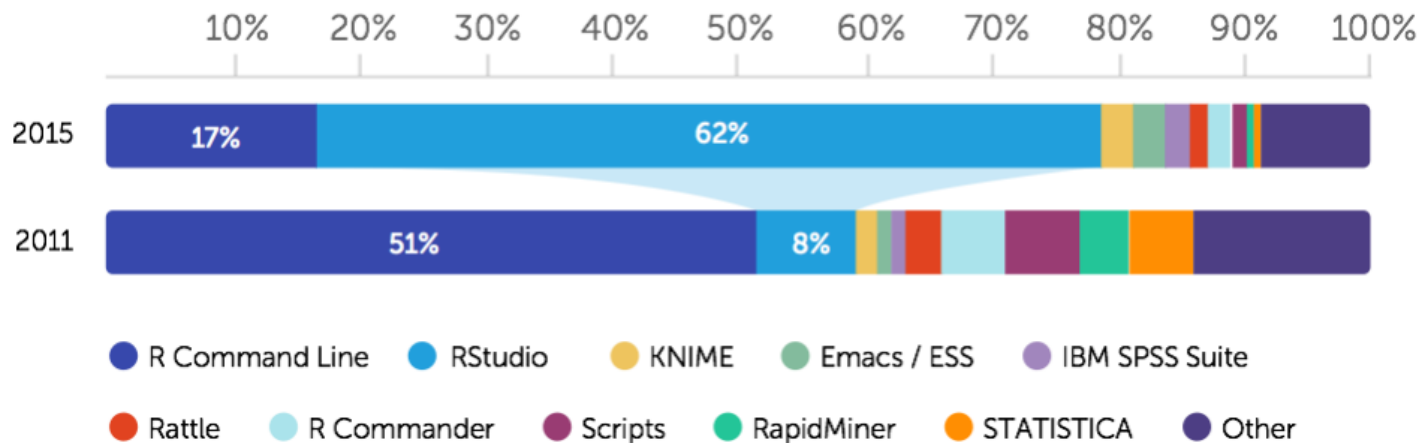


RStudio has Become the Dominant R Interface



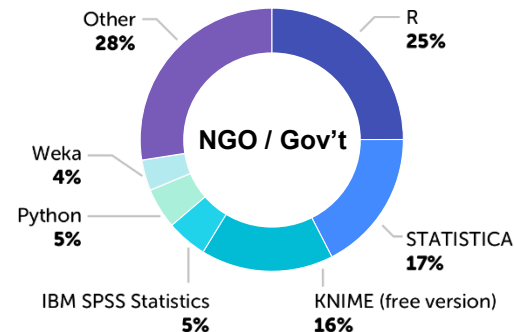
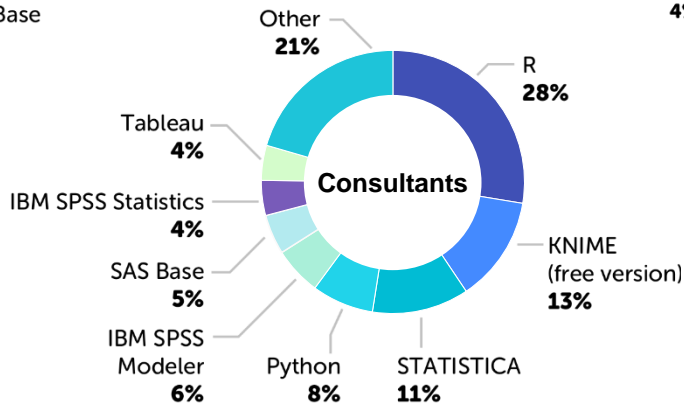
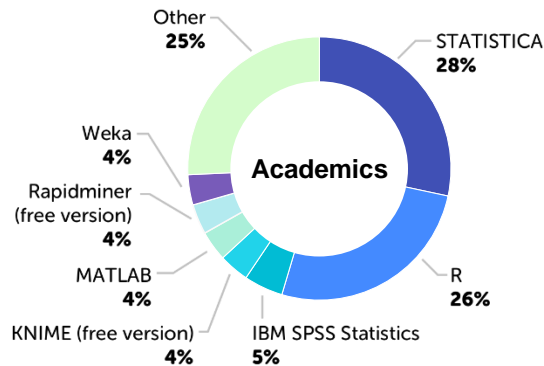
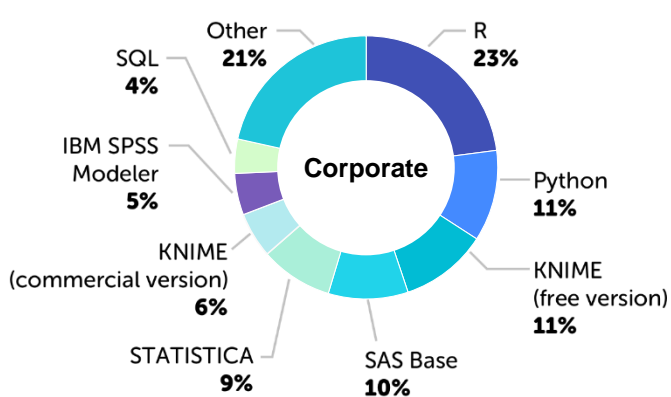
What is your primary interface to R?

Most people using R utilize RStudio as their primary interface to R.
In the 2017 survey responses, this grew to 69% (we just don't have the graphic updated yet)





What one data science tool, technology or language do you use most frequently?



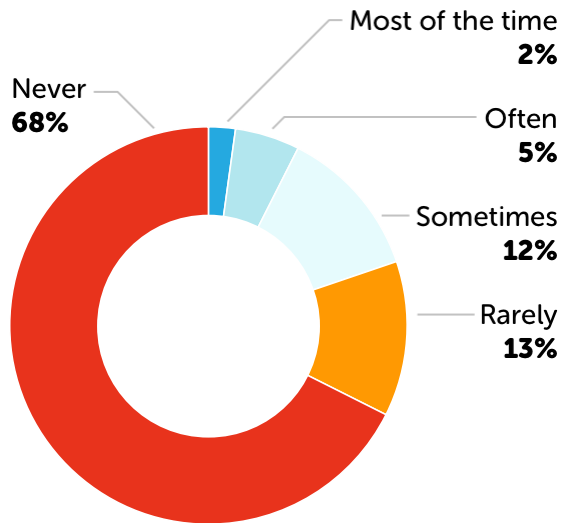
Deep Learning: It's Use is Increasing



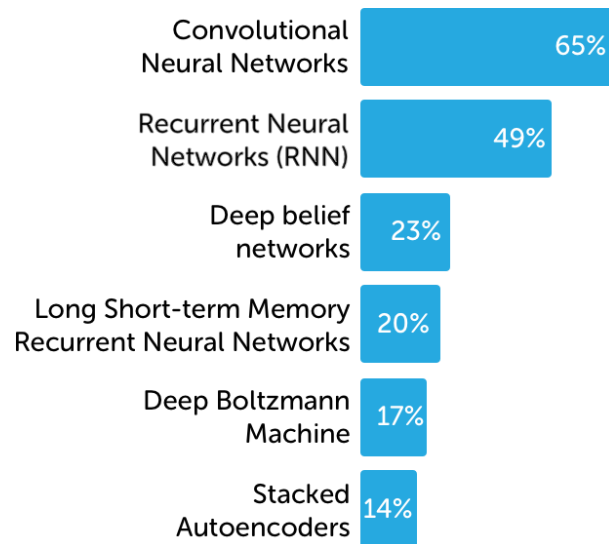
There is a growth in the use of Deep Learning. But it's still only being used by a small proportion of analytic professionals.



What algorithms/analytic methods do you use?:
Deep Learning



Which Deep Learning algorithms have you used?



Selected Success Stories

- Medical imaging segmentation
- Analysis of histopathological images
- Colorization of black and white images
- Classification of 20,000+ species using crowd-sourced images
- Handwriting interpretation and voice recognition
- Closed vocabulary verbal intent determination
- Impact of weather on airport arrival and departure rates
- Chip defect identification in semiconductor industry
- Prediction of negative prices in energy market
- Estimation of the daily demand on ATM knowing historical demands, stock and recharge times

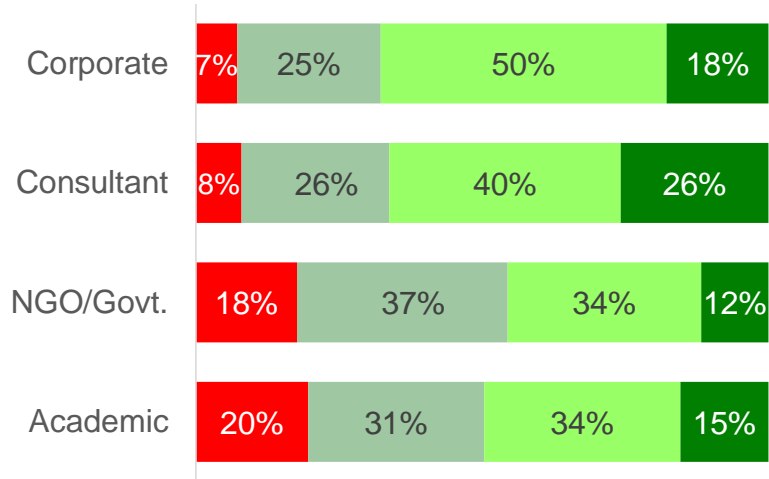


If you have had a great experience using Deep Learning on a project, please provide the details of that project

Employment: It's a Good Time to be a Data Scientist



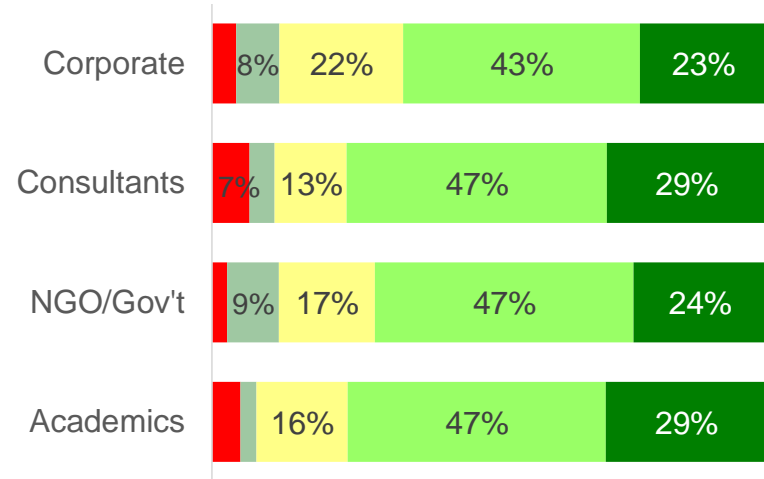
If you left your current role now, how easy would it be for you to find a job fairly equivalent to your current one?



■ Very difficult ■ Somewhat difficult ■ Somewhat easy ■ Very easy



What is your current level of job satisfaction?

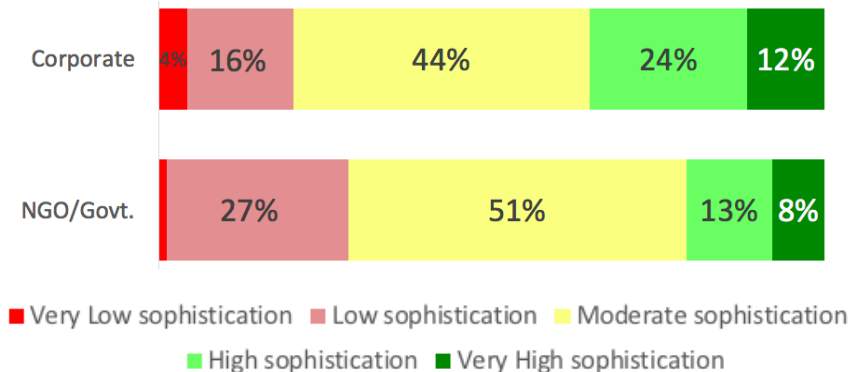


■ Very Unsatisfied ■ Unsatisfied ■ Neutral ■ Satisfied ■ Very Satisfied

Analytic Sophistication: Opportunity for Improvement

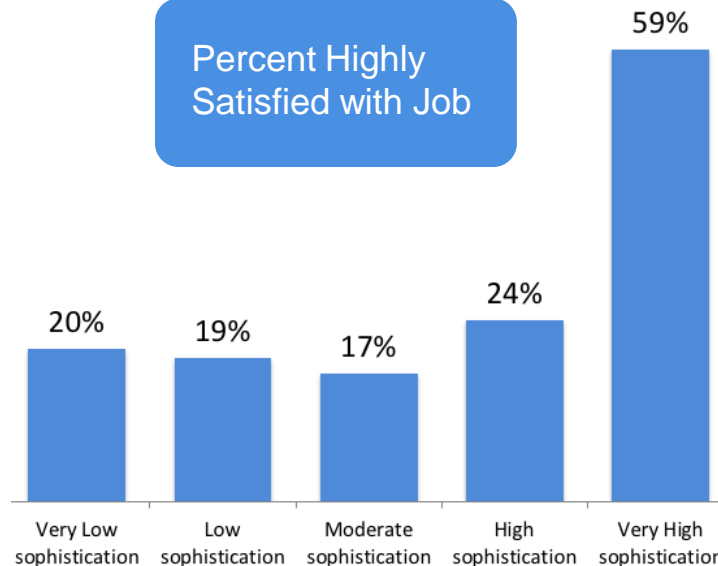


In general, with what degree of sophistication does your company / organization approach analytic problems?



Only a minority of data scientists feel that their firms have high analytic sophistication

Percent Highly Satisfied with Job

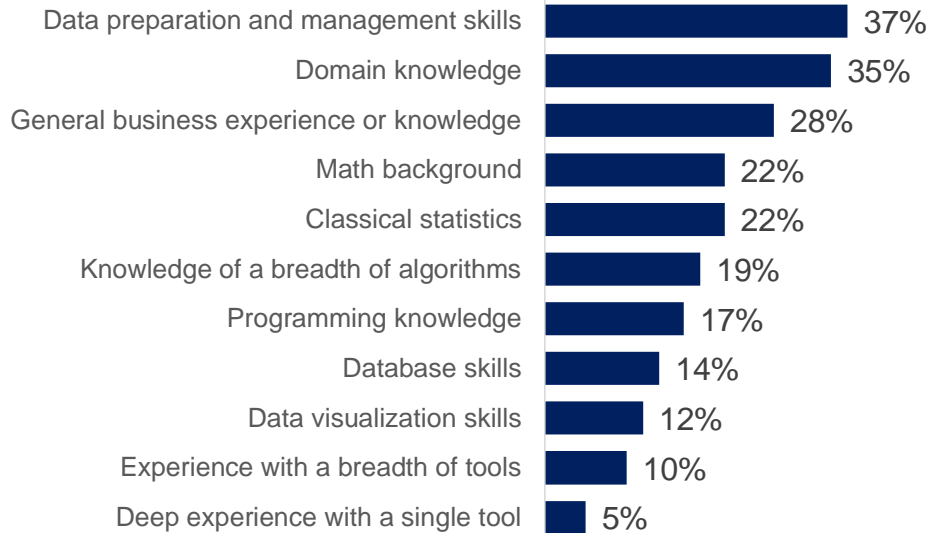


Interpretation: When a company is judged to have very high analytic sophistication, 59% of analytic staff report being very satisfied with their jobs

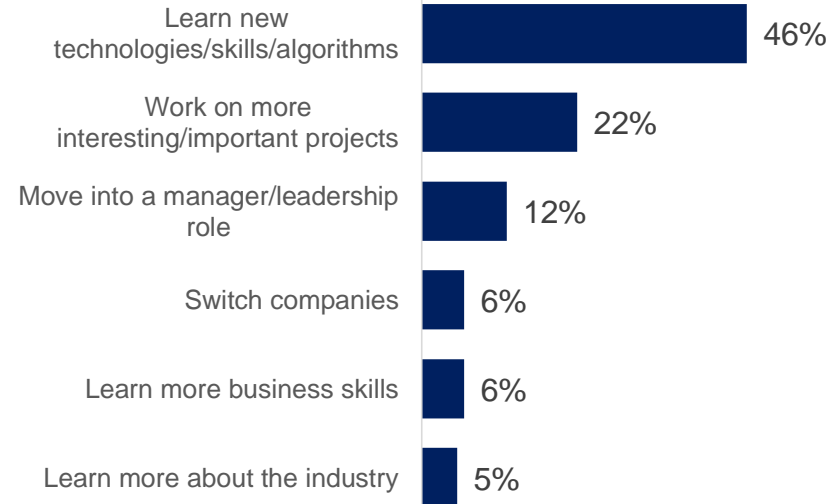
Data Science Skills Needed: Varied and Growing

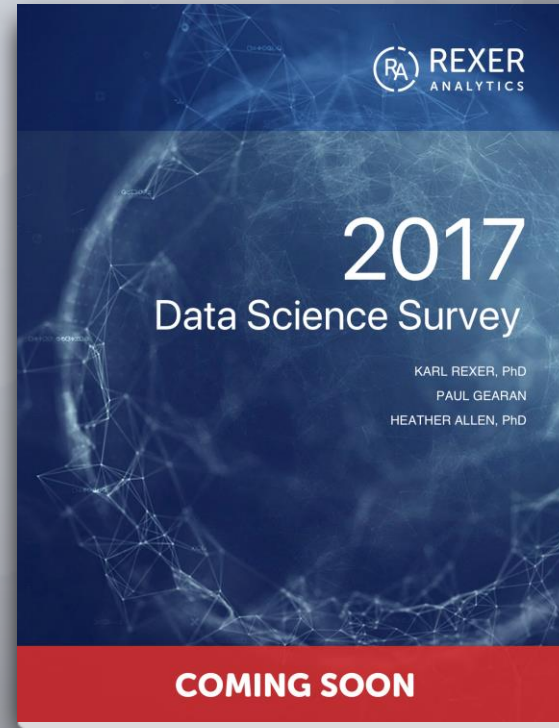
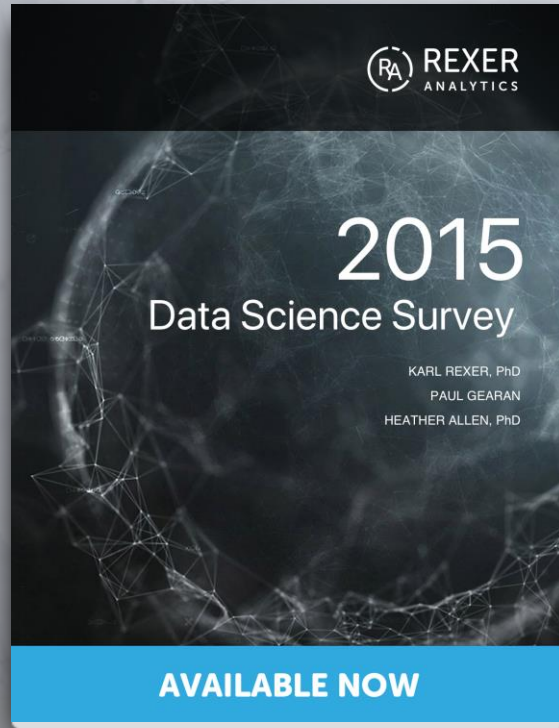


Please rank the following in terms of which are the most important attributes/background/skills for a data miner/data science professional to possess?



What is the most important step you would like to take to advance your career?





www.RexerAnalytics.com

Data Science Survey

Rexer Analytics designed and began this research in 2007. The eight surveys from 2007 through 2017 examine the analytic behaviors, views and preferences of data analytic professionals. We use a small set of consistent questions to enable the tracking of trends across the years. We also incorporate new questions into each survey to explore emerging topics. We appreciate the many people who have submitted survey questions – many have been included in the surveys. Please send us your ideas and feedback about this research by using the form on our [CONTACT](#) page or emailing us at DataScienceSurvey@RexerAnalytics.com. The survey summary reports can be downloaded using the links below.

2017 Survey	▼
2015 Survey	▼
2013 Survey	▼
2011 Survey + “Best Practices” Verbatim Responses	▼
2010 Survey + “Best Practices” Verbatim Responses	▼
2009 Survey	▼
2008 Survey	▼
2007 Survey	▼



For additional information, please contact
Karl Rexer, PhD
President
krexer@RexerAnalytics
617-233-8185

www.RexerAnalytics.com