

Monitoring Trends in Tools for Data Science

Bob Muenchen (Pronounced MINchen)

muenchen@utk.edu

Greve Hall 522

<http://oit.utk.edu/research>

Why Bother?

- ◆ Rapid rate of change
- ◆ Plethora of tools
- ◆ Your career & salary are on the line
- ◆ Only the paranoid ~~survive~~ thrive

Tools I Track

- ◆ Alpine
- ◆ Alteryx
- ◆ Angoss
- ◆ C / C++ / C#
- ◆ BMDP
- ◆ IBM SPSS Statistics
- ◆ IBM SPSS Modeler
- ◆ InfoCentricity Xeno
- ◆ Java
- ◆ JMP
- ◆ KNIME
- ◆ Lavastorm
- ◆ Mathworks' MATLAB
- ◆ Megaputer's PolyAnalyst
- ◆ Minitab
- ◆ NCSS
- ◆ Python
- ◆ R
- ◆ RapidMiner
- ◆ SAS
- ◆ SAS Enterprise Miner
- ◆ Salford Predictive Modeler
- ◆ SAP's KXEN
- ◆ Stata
- ◆ Statistica
- ◆ Systat
- ◆ WEKA / Pentaho

bit.ly/statpop

r4stats.com



[Home](#) [Articles](#) [Blog](#) [Books](#) [Downloads](#) [Examples](#) [Miscellaneous](#) [Workshops](#)

The Popularity of Data Analysis Software

by Robert A. Muenchen

Abstract

This article presents various ways of measuring the popularity or market share of software for advanced analytics software. Such software is also referred to as tools for data science, statistical analysis, machine learning, artificial intelligence, predictive analytics, business analytics, and is also a subset of business intelligence.

Updates: The latest data for Language Popularity Measures was updated 10/17/2015.

I announce the updates to this article on Twitter: <http://twitter.com/BobMuenchen>

Sources of Information

- ◆ Job Advertisements
- ◆ Scholarly articles
- ◆ IT Research Firm Reports – Gartner, Forrester
- ◆ Discussion Forum Activity
- ◆ Surveys of Use
- ◆ Books
- ◆ Blogs
- ◆ Programming Language Popularity Measures
- ◆ Growth in Capability

Beware of Ambiguity, e.g. “SAS”

Senior Analysts (abbreviated SAs)

System Administrators (abbreviated SAs)

SAS Shoes

Samsung Austin Semiconductor

School Age Services

Scandinavian Airlines System

SCSI Attached Storage

Sedation Agitation Scoring

Sexual Assault Services (e.g. “SAS Nurse”)

Specialist Alarm Services

Student Access to Science

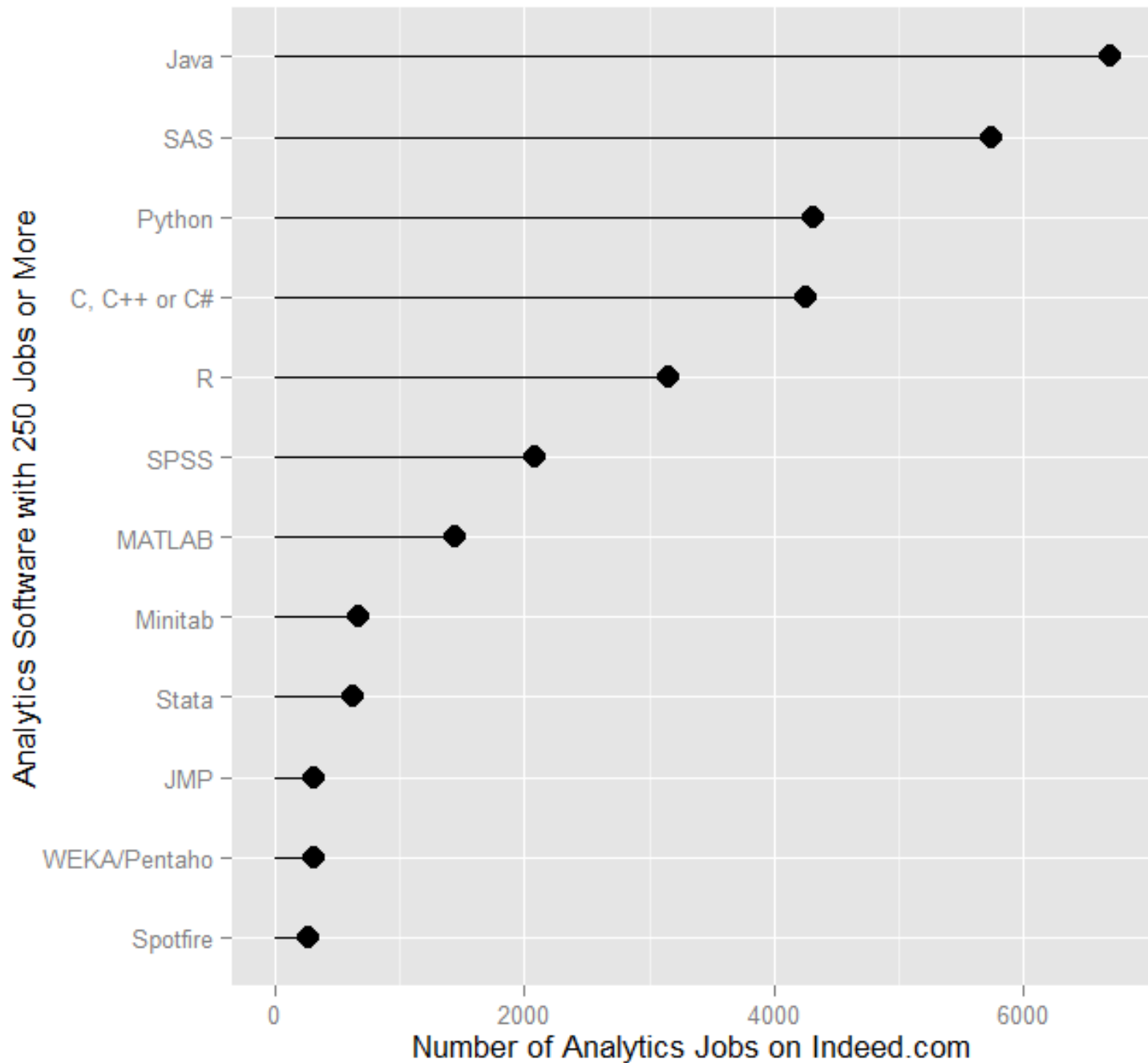
Surgical Admissions Suite

Synthetic Aperture Sonar

Focus Your Search

Search Terms	Number of Jobs
1 Analytics (not well focused)	65,136
2 Survey (not well focused)	47,342
3 Statistical (not well focused)	42,890
4 Statistics (not well focused)	37,783
5 Business intelligence (too much reporting)	19,805
6 Analyze data (not well focused)	13,339
7 Big data *	10,378
8 Statistical analysis *	9,719
9 Data mining *	7,776
10 Data analytics *	6,209
11 Machine learning *	3,658
12 Quantitative analysis *	3,365
13 Research associate (too vague)	3,022
14 Business analytics *	2,867
15 Statistical software *	2,102

Jobs



Job Trends @ Indeed.com

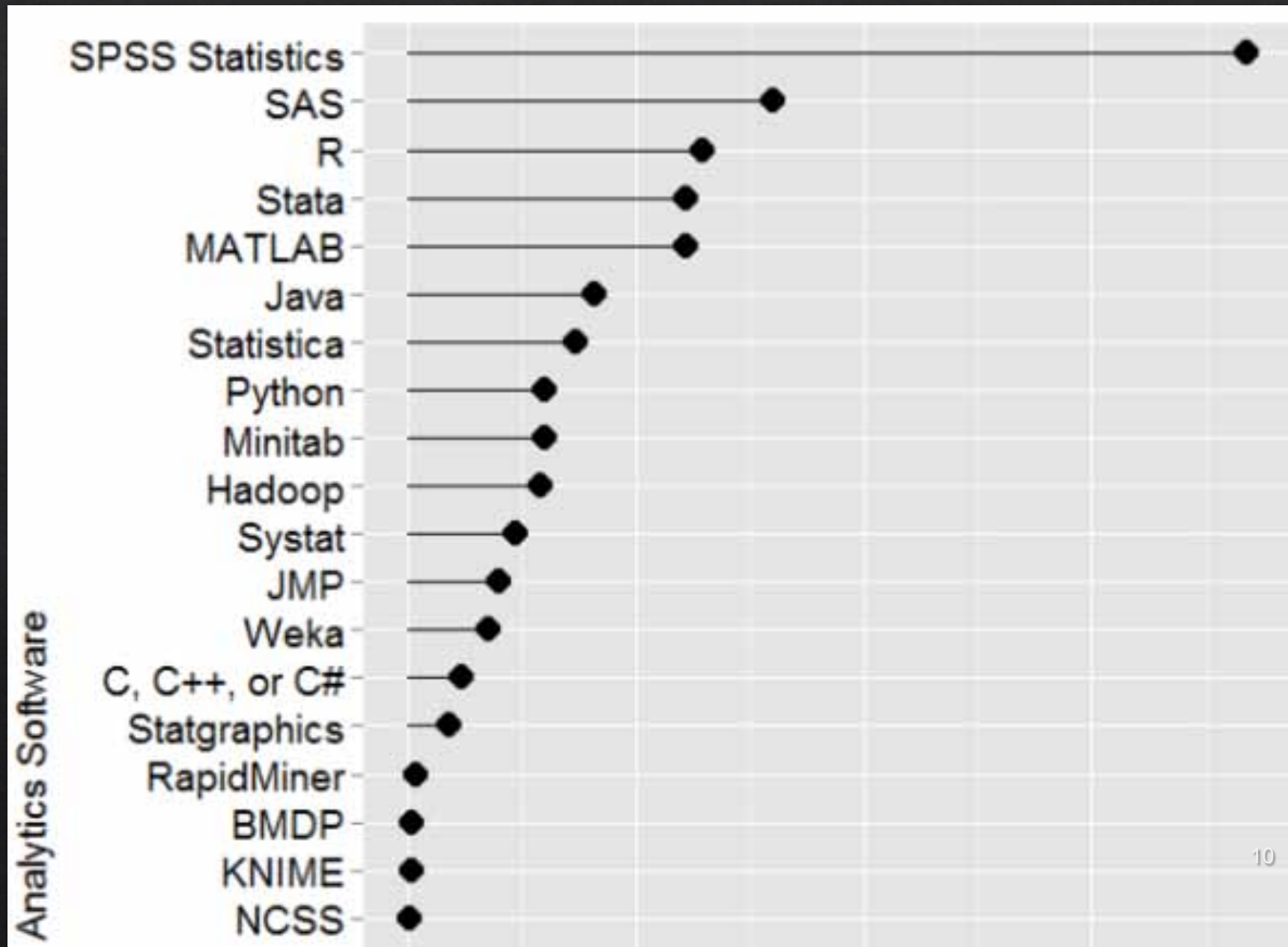
Job Trends from Indeed.com

— R "R D" "A R" "H R" "R N" Toys kids " R Walgreen" Walmart "HVAC R" "R Bard" and (

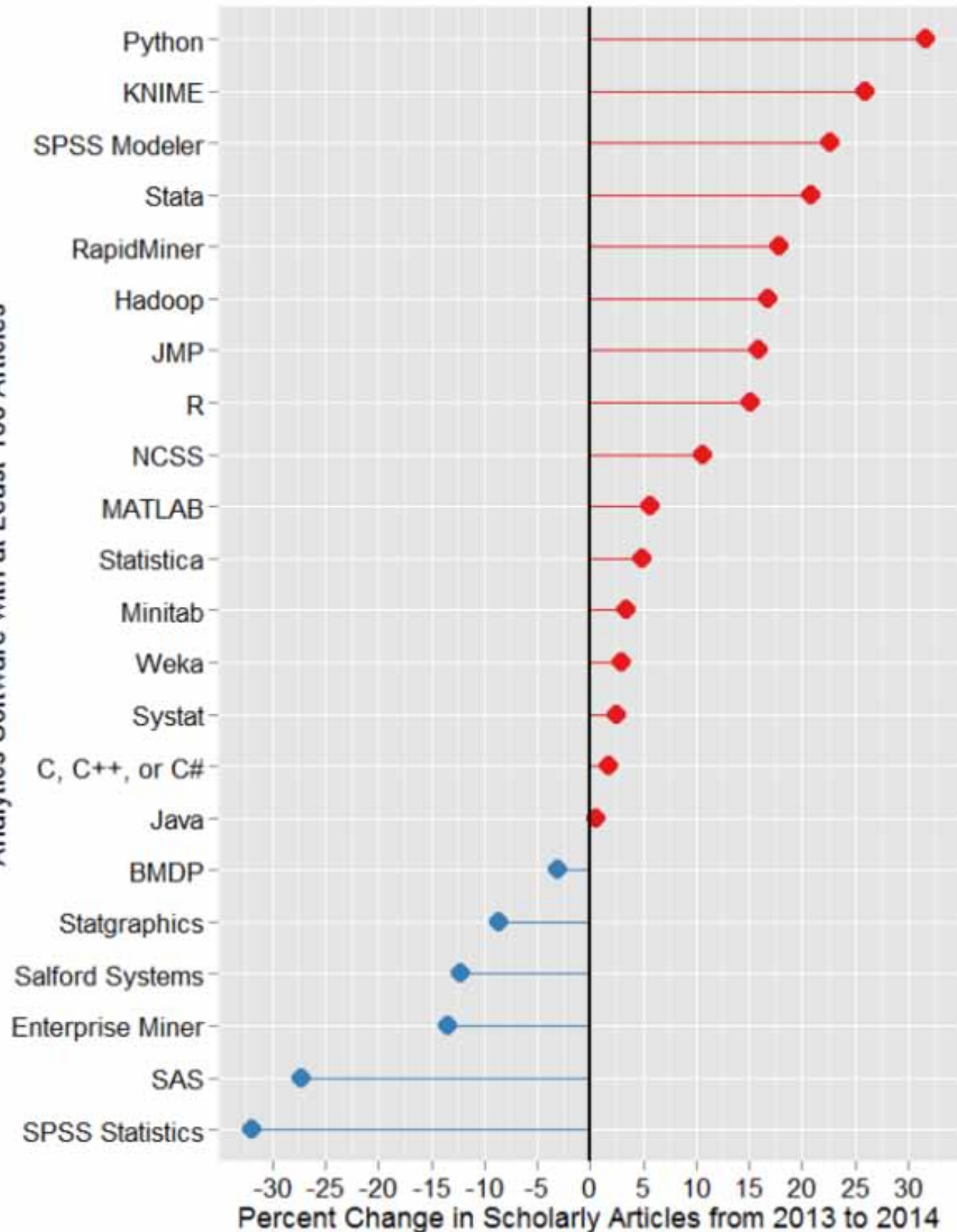
— SPSS and ("big data" or "statistical analysis" or "data mining" or "data analytics" or "machir



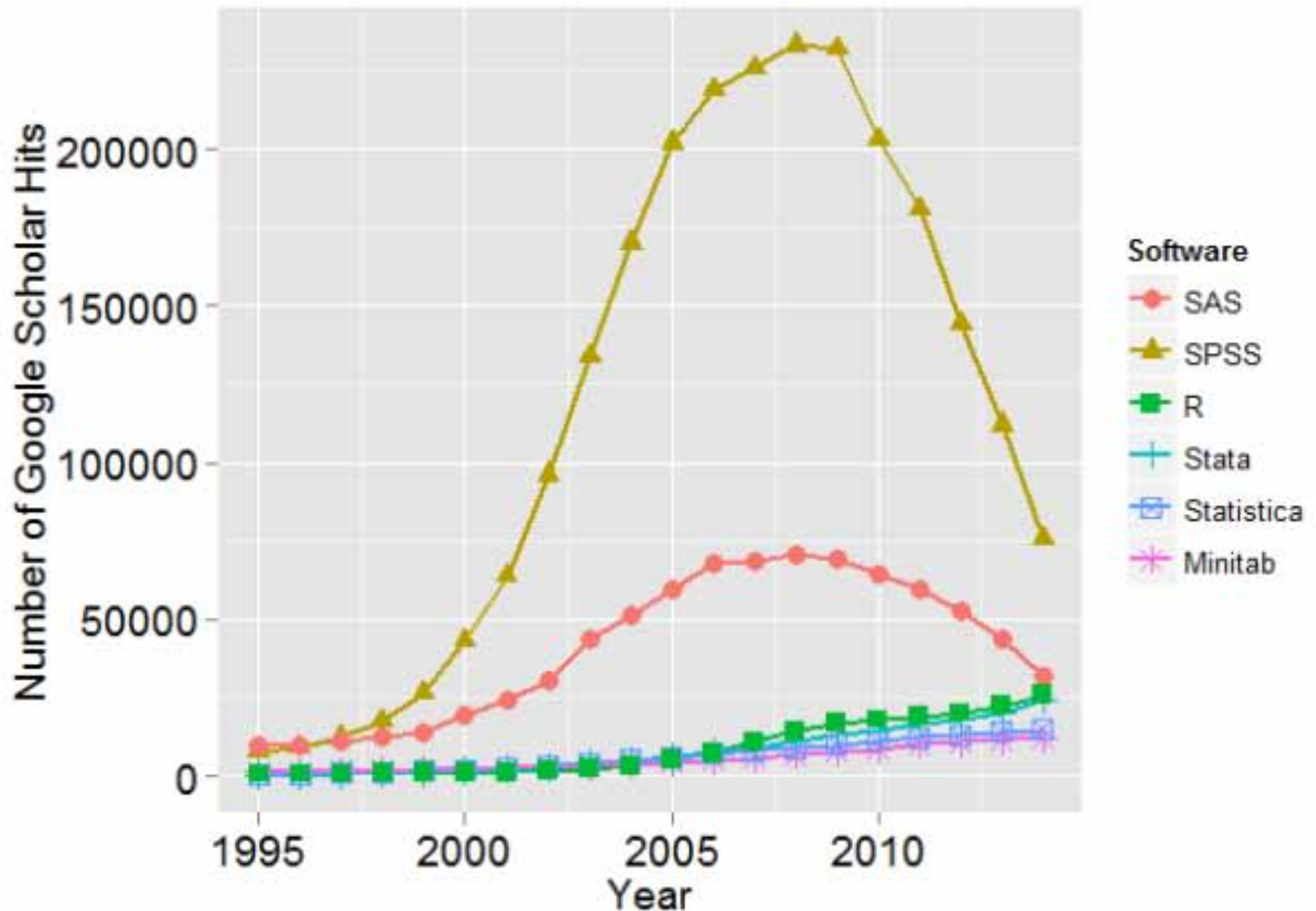
Scholar.Google.Com



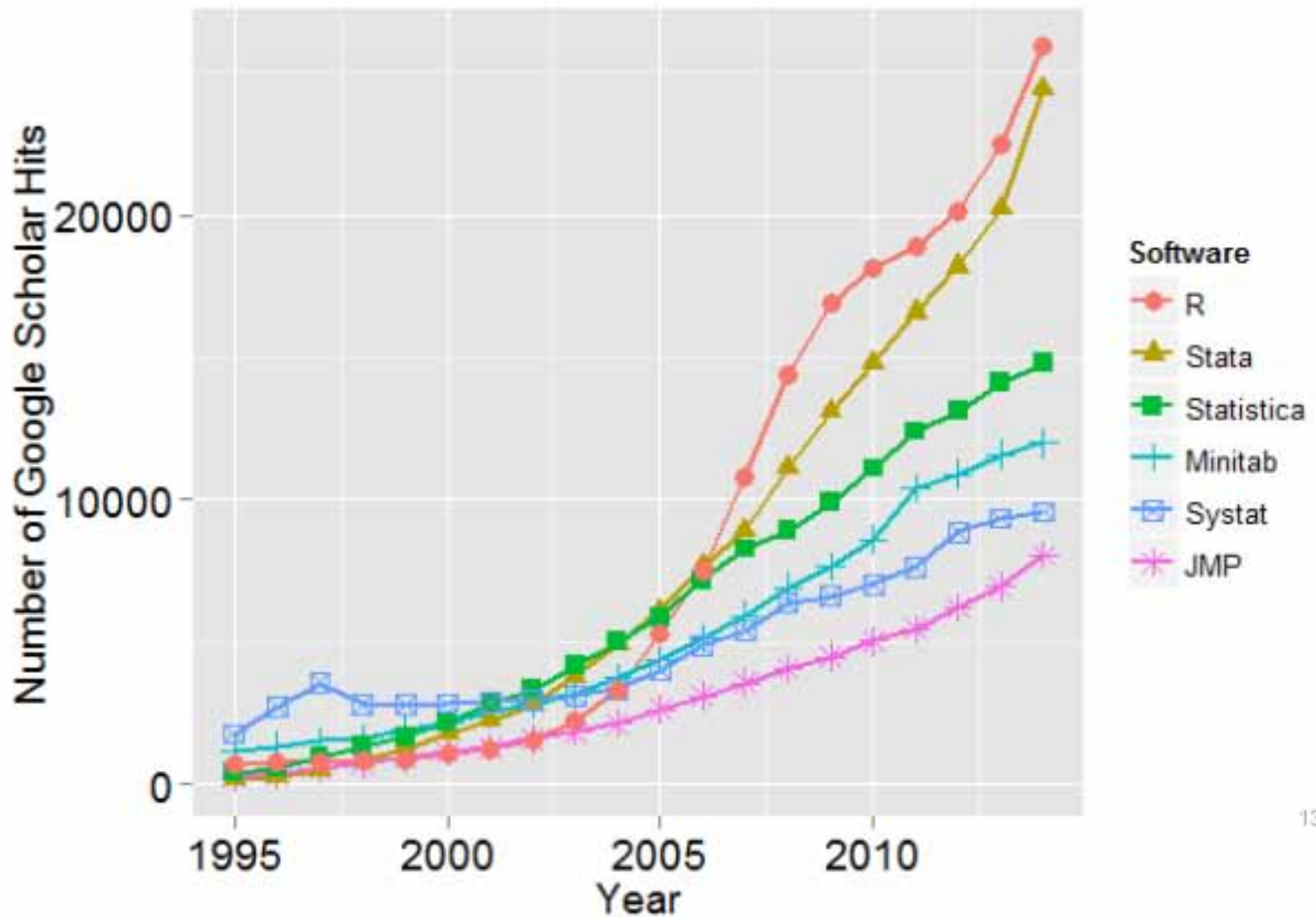
Analytics Software with at Least 100 Articles



Statistics Packages Across Time



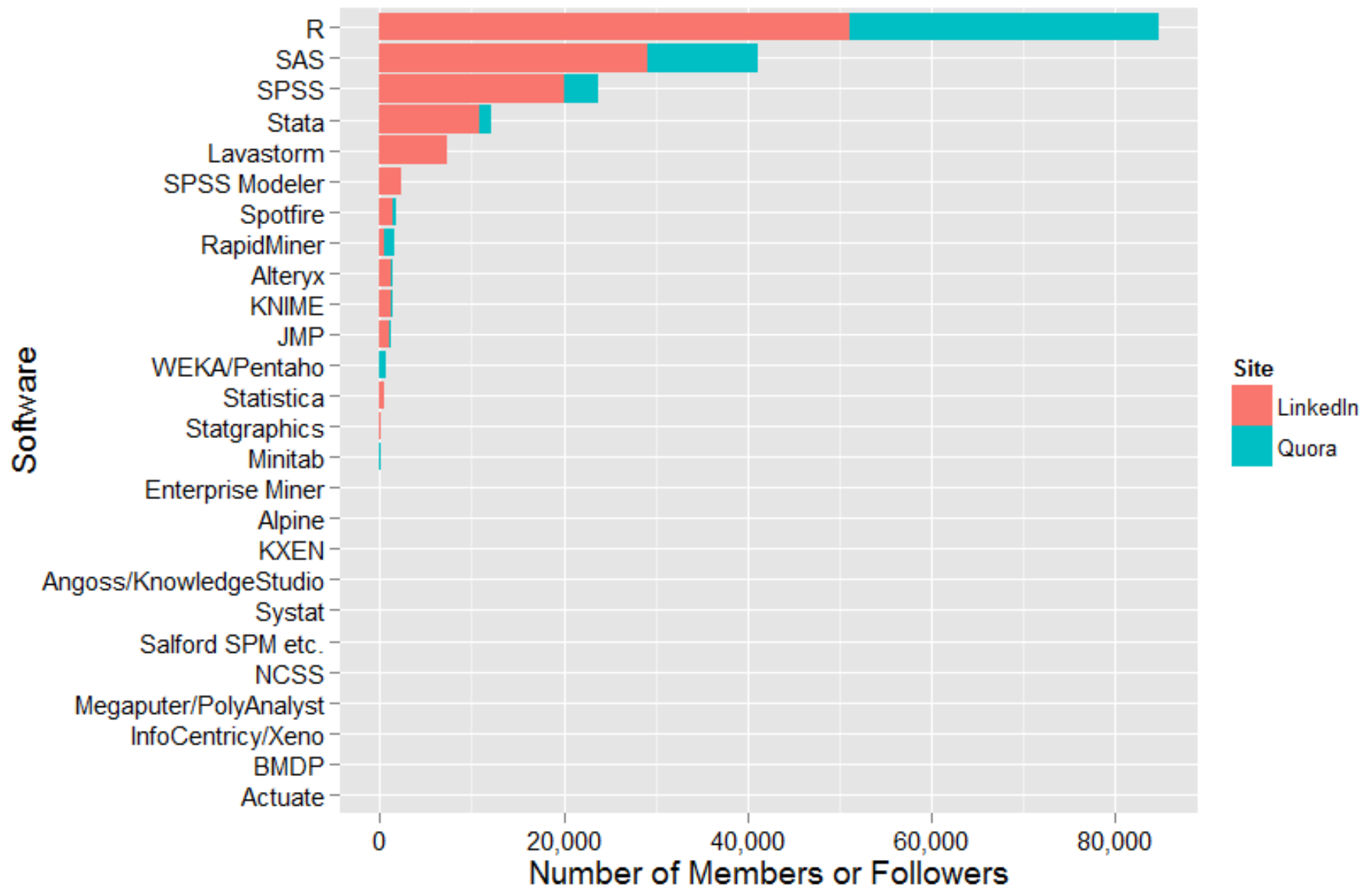
SPSS & SAS Removed



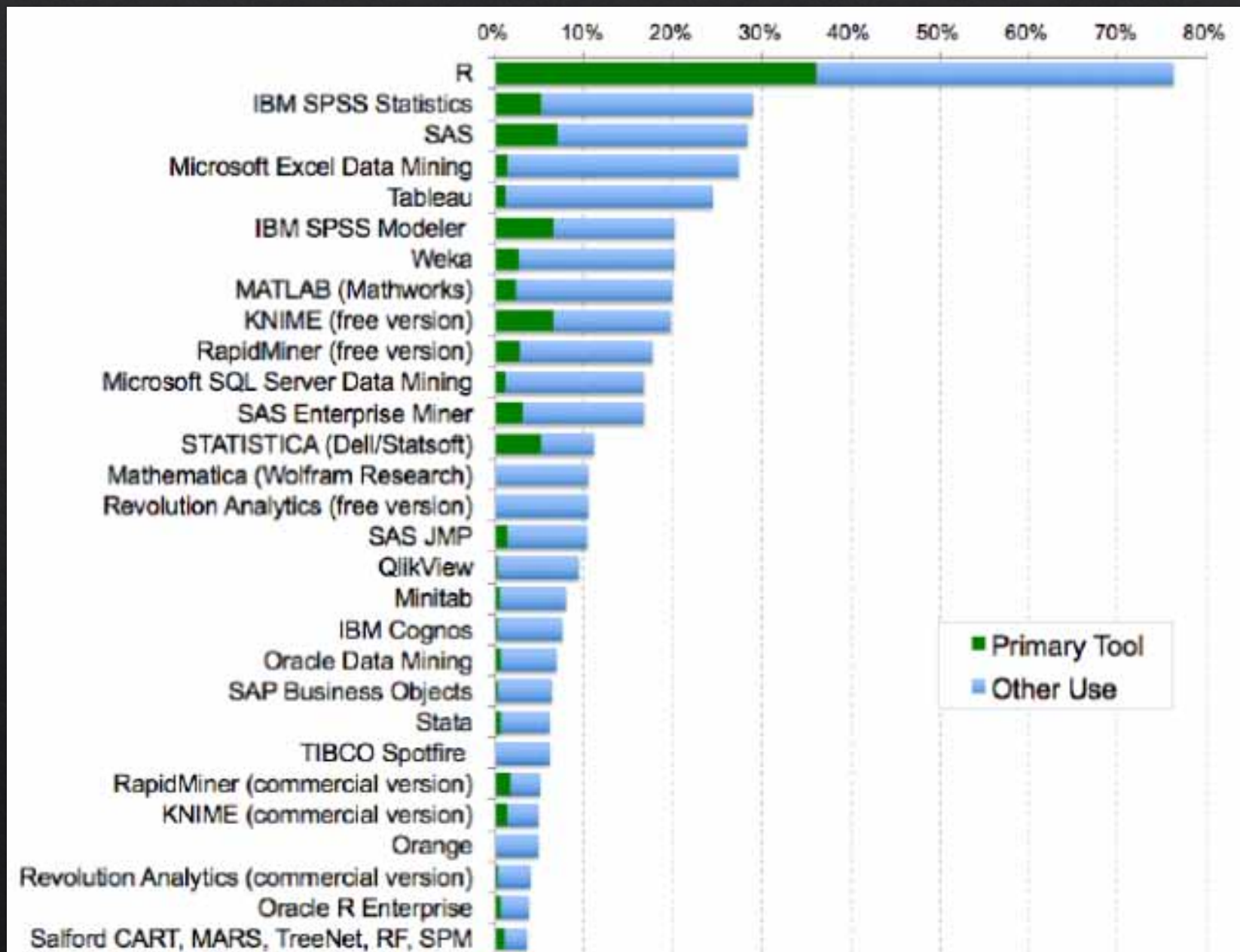
IT Research Firms: Gartner



Discussion Forum Membership



Surveys of Use: Rexer Analytics



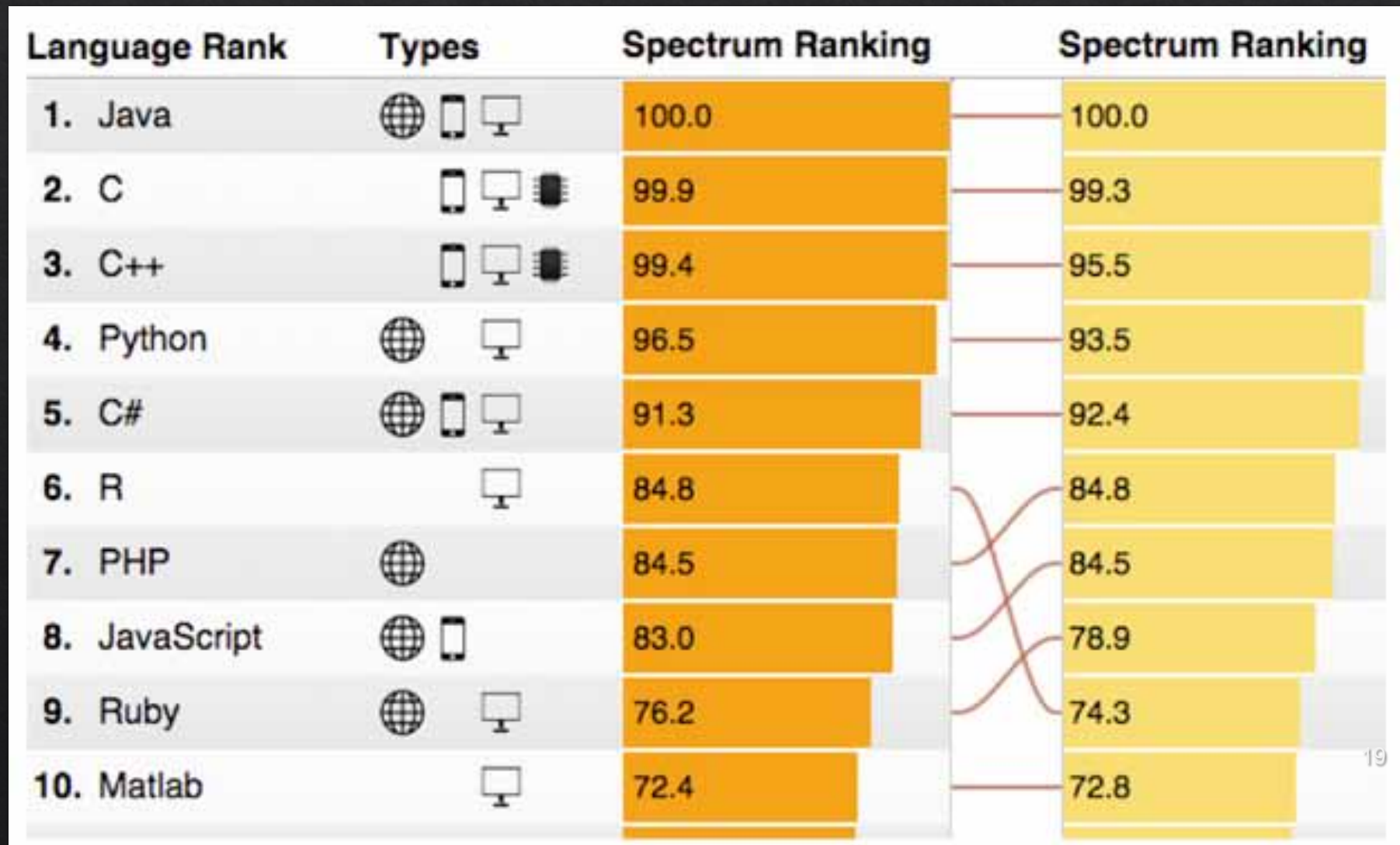
Books With Titles Containing X

<u>Software</u>	<u>Number of Books</u>
SAS	576
SPSS Statistics	339
R	240
JMP	97
Hadoop	89
Stata	62
Minitab	33
Enterprise Miner	32
All Others	<8 (often 1 or 0)

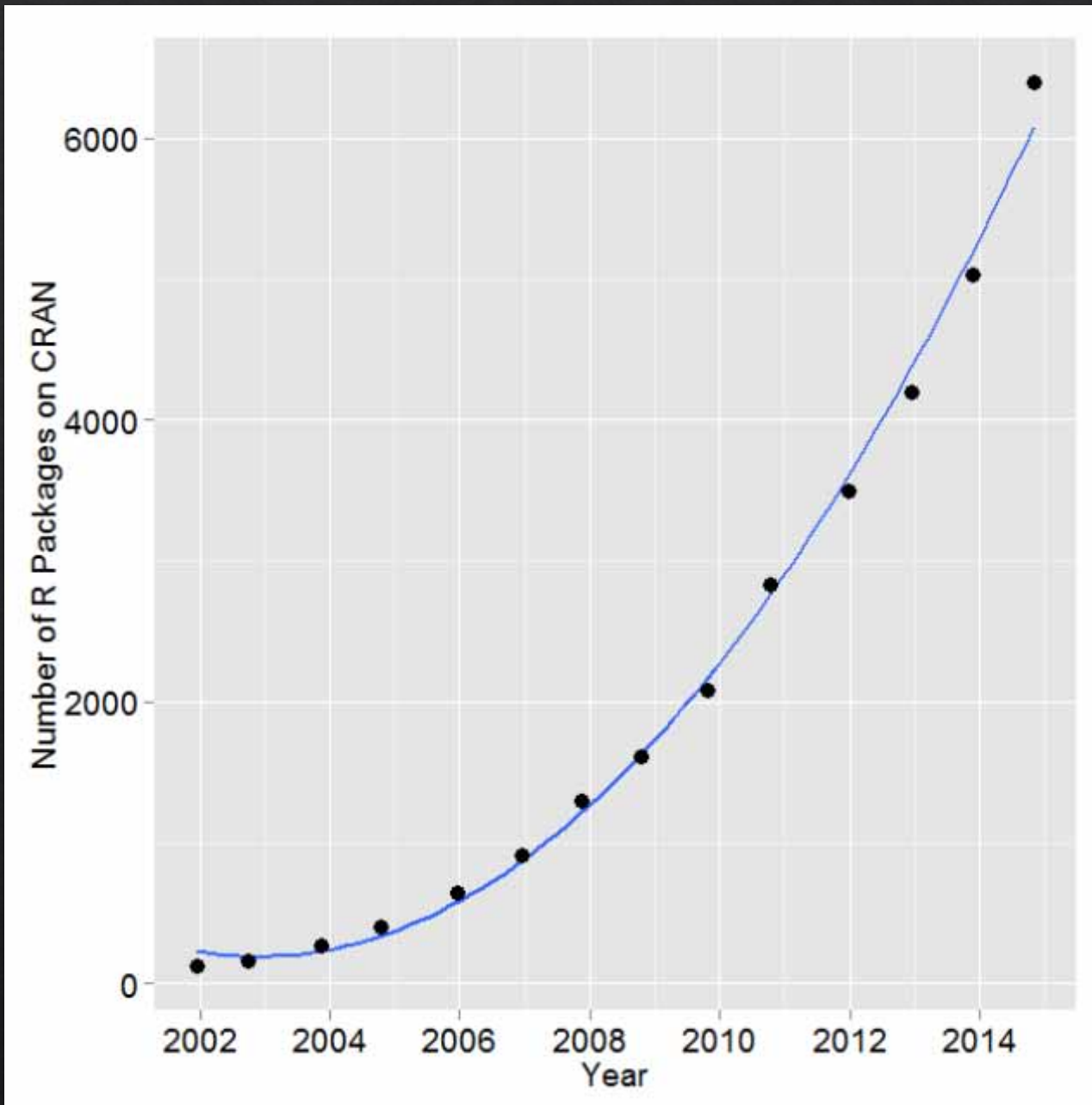
Blogs

<u>Software</u>	<u>Number of Blogs</u>		<u>Source</u>
	<u>2015</u>	<u>2014</u>	
R	573	550	R-Bloggers.com
Python	78	60	SciPy.org
SAS	39	40	PROC-X.com
Stata	7	11	http://www.stata.com/...

Programming Popularity Sites: IEEE Spectrum



Growth in Capability



Recommendations

- ◆ Track the tools of your field
- ◆ Don't view tools with religious fervor
- ◆ Assume all your current tools will be replaced
- ◆ Learn new ones *before* it becomes mandatory
- ◆ Become a change agent for your organization
- ◆ If you champion a tool,
champion its replacement too

Questions?