

R Programming Fundamentals

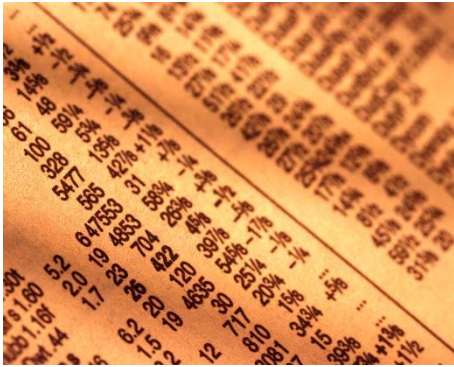
Getting Started

Abhishek Kumar
ItsAbhishekKumar.com
@MeAbhishekKumar



pluralsight
hardcore dev and IT training







Data scientists : The sexiest job of
21st century



— Industry experts

About the Course

R framework & its
capabilities

Core programming
concepts

Basic **data analysis**

Course Objective

Provide Solid
Foundation

Real World
Examples

What & Why ?

Follow Along

Not To Cover
Every Aspect

Take The First
Step

Overview



Implementation of S
language (created by **John
Chambers**)

Developed by **Ross Ihaka**
and **Robert Gentleman**

Developed as a **free**
software environment for
teaching

Maintained by **R**
development core team

Overview



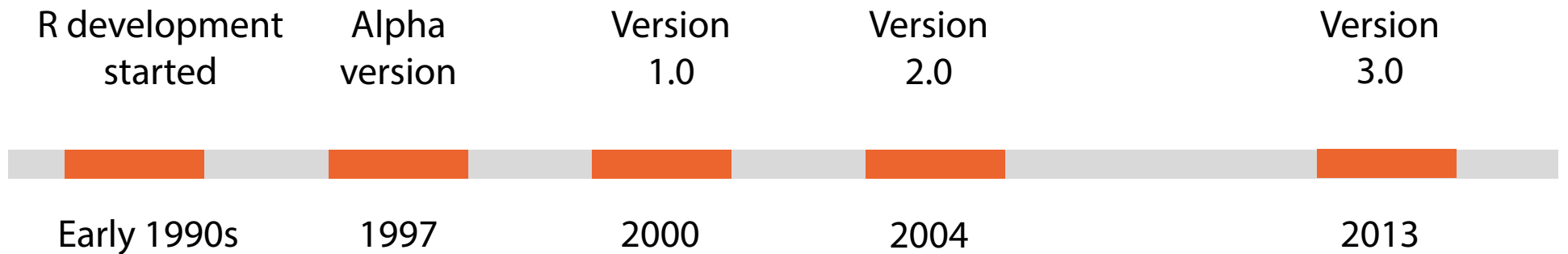
Statistical software package

Comparable to commercial packages such as **SAS, SPSS**

Free interactive programming language and environment

Used in various applications

Timeline



<http://cran.r-project.org/src/base/NEWS>

Why **R**?



Free and open source

Run on Windows/ Mac/ Linux

Great extensibility

Active community

**Integration with other
language and packages**



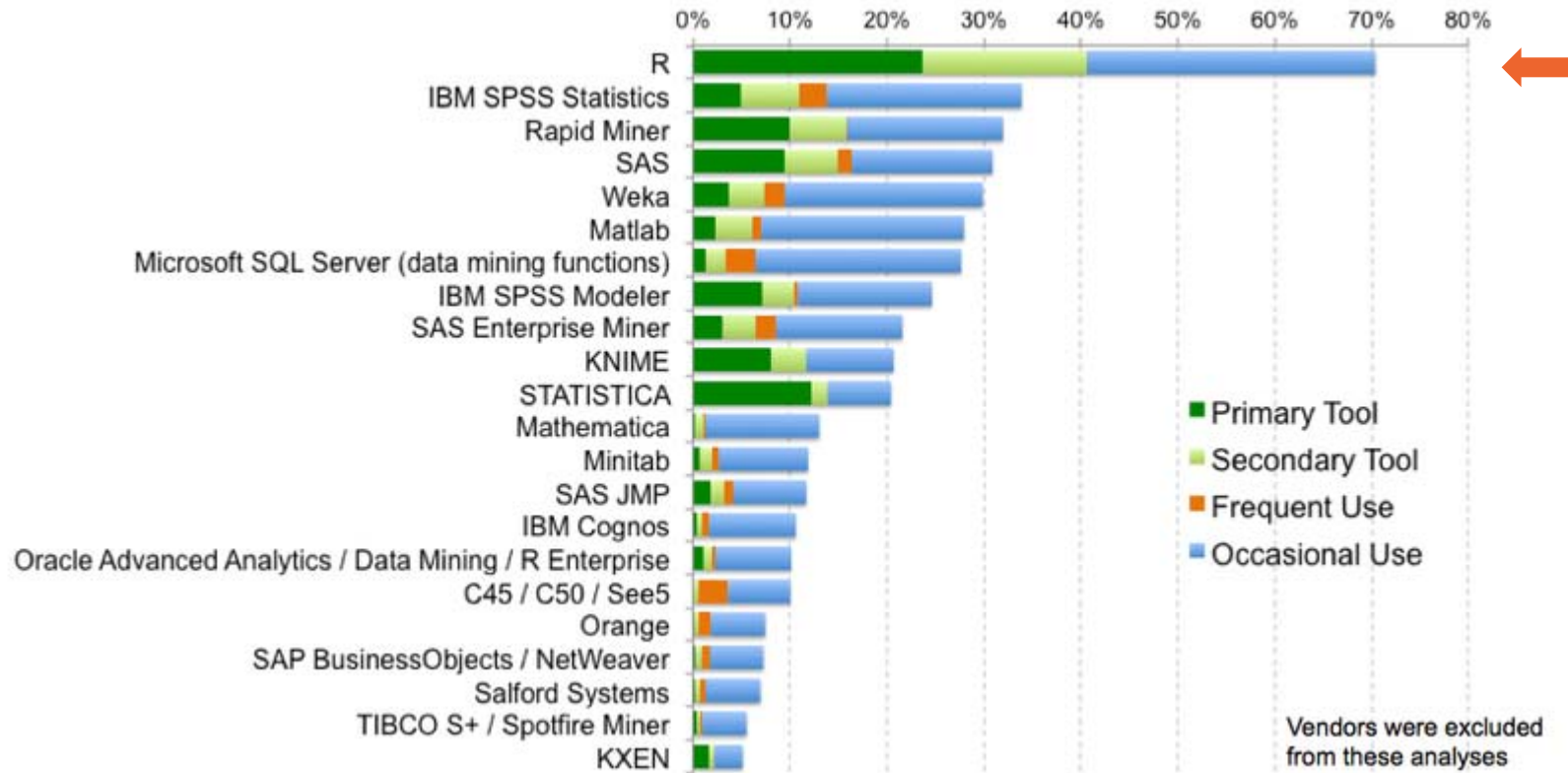
Users

Statistician

Data analyst

Data miner

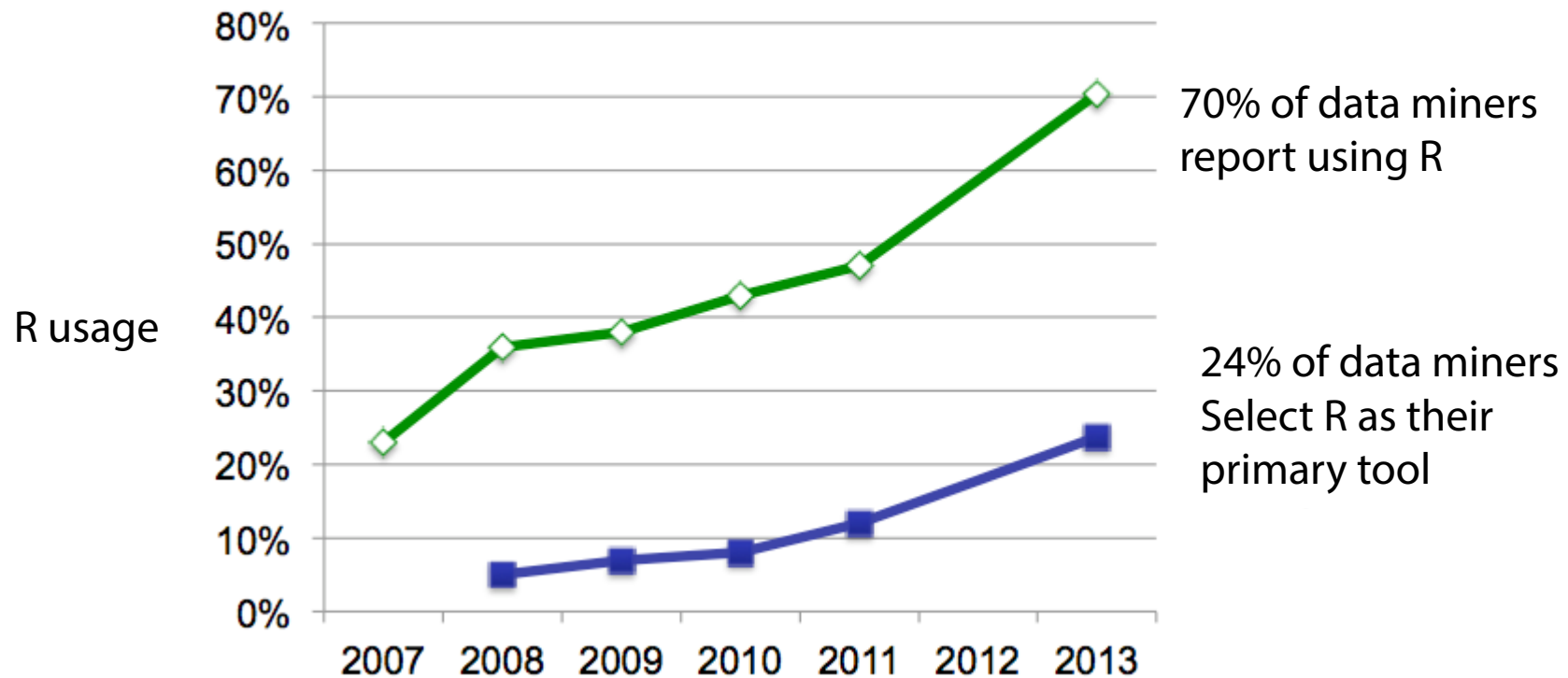
R Popularity



Rexter Analytics data miner survey results in 2013

Source : <http://blog.revolutionanalytics.com/2013/10/r-usage-skyrocketing-rexer-poll.html>

R Popularity



Rexter Analytics data miner survey results in 2013

Source : <http://blog.revolutionanalytics.com/2013/10/r-usage-skyrocketing-rexter-poll.html>

Competitions

kaggle

Customer Solutions

Competitions

Community ▾

Sign Up

Login

We're the global leader in solving
business challenges through
predictive analytics.

facebook

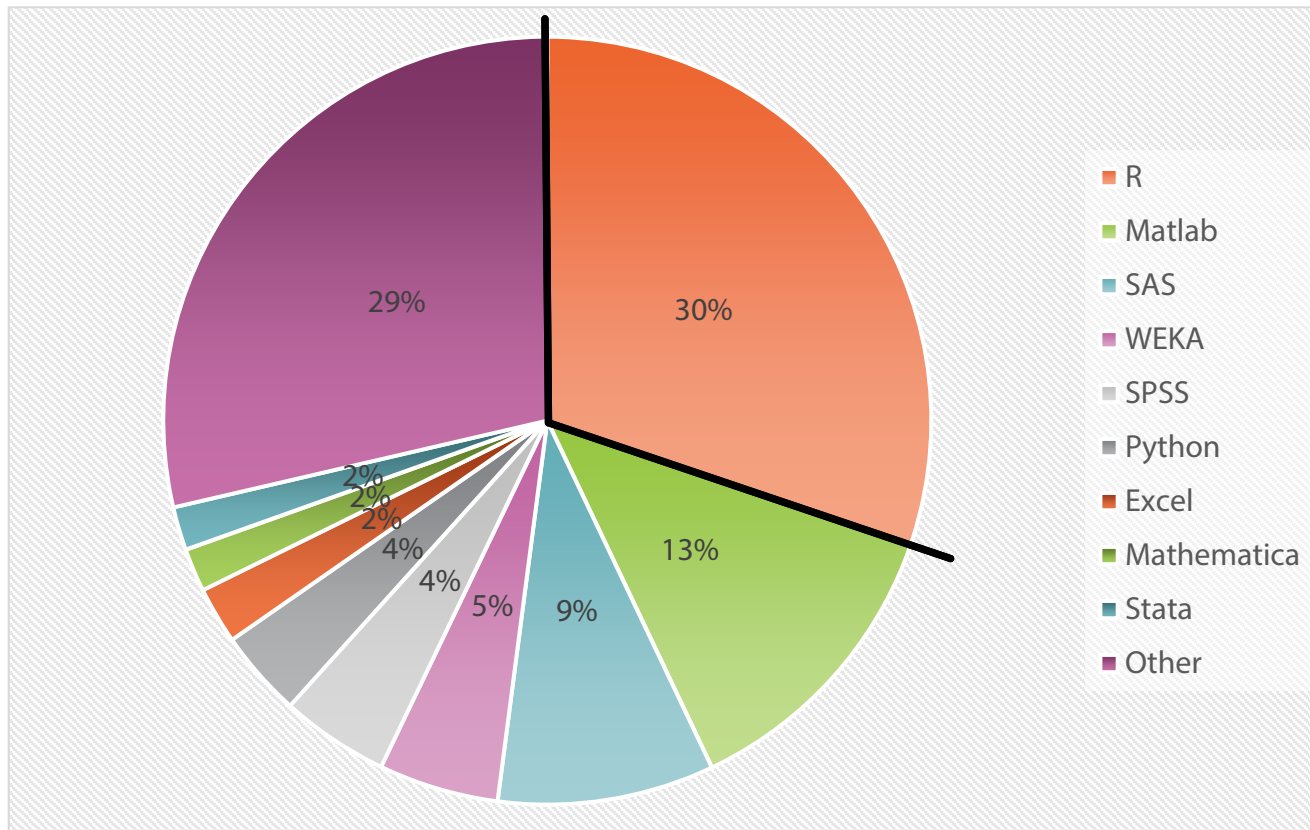


Compete as a data scientist for fortune, fame and fun »



FOCUS INDUSTRY The Energy industry uses our
expertise in machine learning and big data to
drive high-stakes decisions. [Find out more](#)

Kaggle Competition



Source : Apr 21, 2011 presentation by Kaggle CEO Anthony Goldbloom

Tasks

Statistics

Optimization

Machine
Learning

Spatial

Visualization

High
Performance
Computing

Application Areas

Finance

Natural
Language
Processing

Social Science

Genetics

Market
Research

Big Data

Why to Learn **R** Programming?

GUI applications

R programming

Why to Learn **R Programming**?

Perform **specific**
tasks

Repeatable steps

Understand existing
implementation

Why to Learn R Programming?

Extend existing
implementation

Create Software for
data analysis

Aspire to be a data
scientist

Editors

- **Through base installation**

- Windows : R-Gui
- Mac : R.app
- Linux : No specific editor

- **Other editors**

- Emacs : <http://ess.r-project.org/>
- Vim-R : http://www.vim.org/scripts/script.php?script_id=2628
- NPPToR : <http://sourceforge.net/projects/npptor/>
- TINN-R : <http://www.sciviews.org/Tinn-R/>

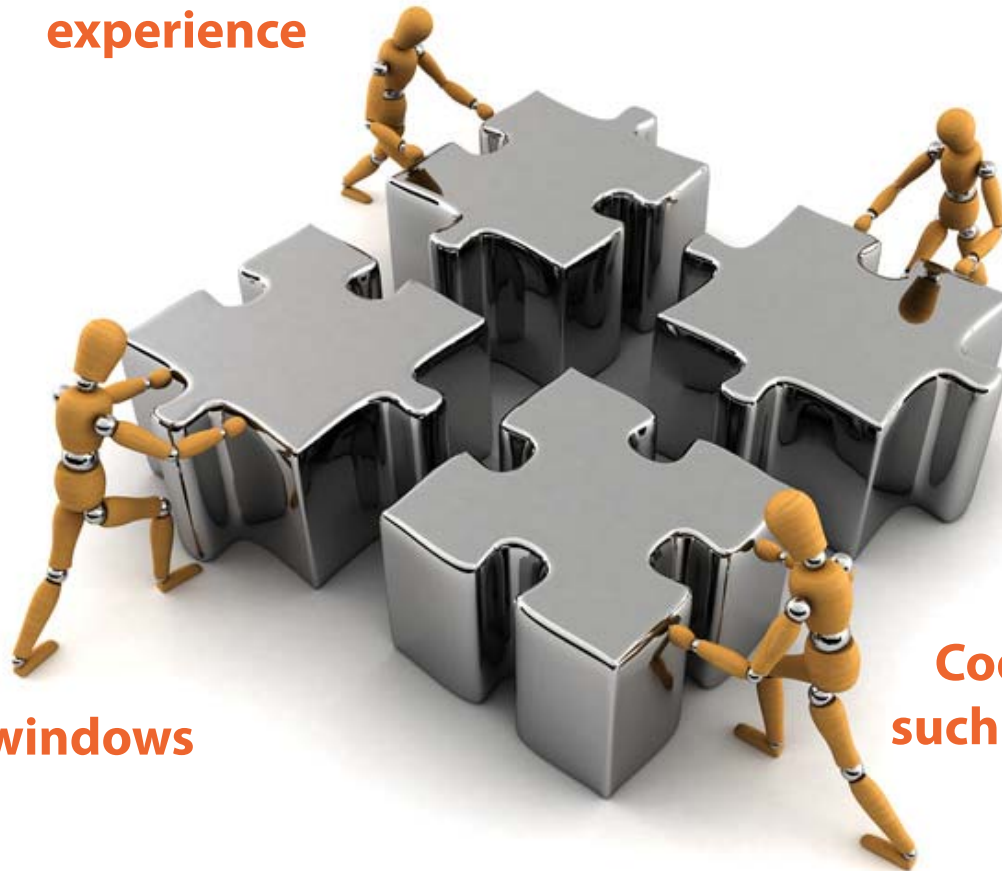
Why **IDE**?

**Better code writing
experience**

Easy debugging

Multiple windows

**Code management
such as version control**



IDE for R

Eclipse

- Used by Java developers (<http://bit.ly/1kRtU3e>)
- StatET plugin (<http://bit.ly/OH1G03>)

Architect

- R oriented variant of Eclipse
- <http://bit.ly/1mLYogP>

Revolution-R

- Enterprise edition
- <http://bit.ly/1gkJzZo>

Live-R

- R in web application
- <http://bit.ly/PZN38V>

RStudio

- R oriented IDE
- <http://bit.ly/NBHT0U>

Course Structure



Getting help
(Module **2**)



Core R programming
(Module **3-8**)



Working with data
(Module **9-10**)