Institute and Faculty of Actuaries

Actuary vs Data Scientist

Richard Pugh
Chief Data Scientist, Mango Solutions
President @ R Consortium

Chris Reynolds
Head of Life Solutions Actuarial, PartnerRe

10 November 2015

Disclaimer

The following presentation is for general information, education and discussion purposes only.

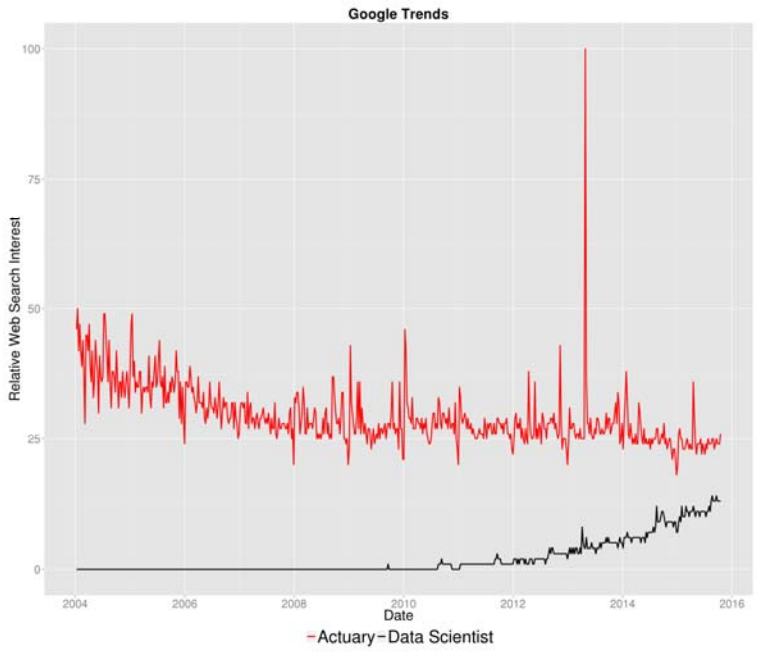
Views or opinions expressed, whether oral or in writing do not necessarily reflect those of PartnerRe nor do they constitute legal or professional advice.

Source: Aljawad / Wikimedia Commons / Public Domain

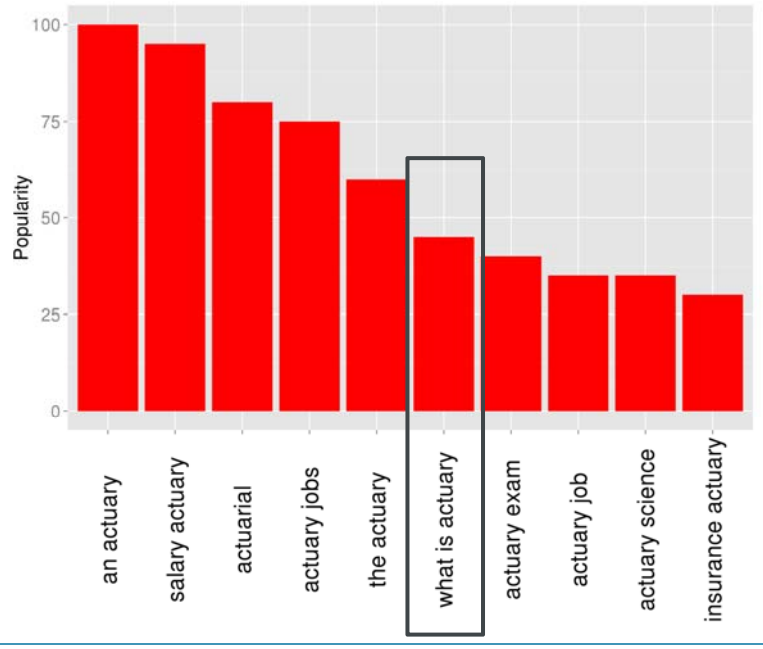
What makes a good actuary?

We would say the dominant trait among actuaries is an intense curiosity – a desire to go beneath the surface of a problem, find the questions at its heart, and distil them into a very clear set of hypotheses that can be tested

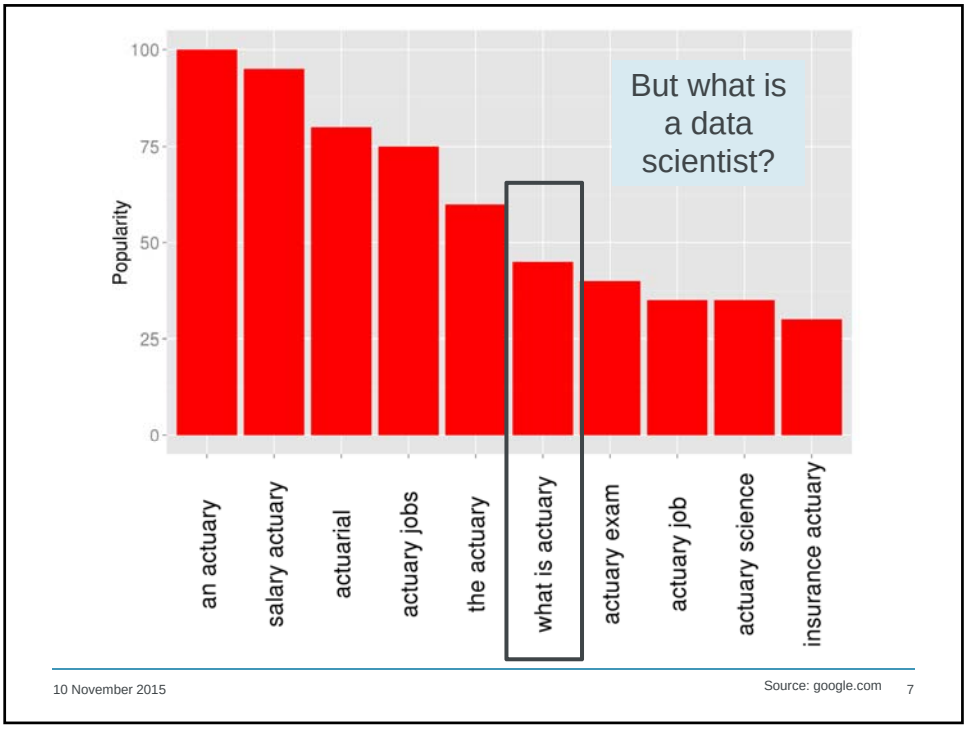
Would you agree?



10 November 2015 Source: google.com 5



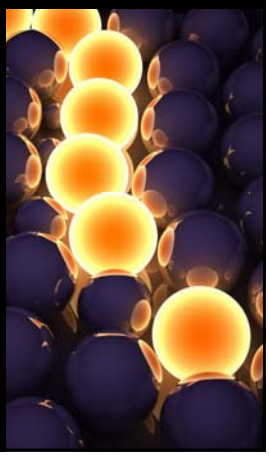
10 November 2015 Source: google.com 6


But what is a data scientist?
Popularity
100
75
50
25
0
an actuary
salary actuary
actuarial
actuary jobs
the actuary
what is actuary
actuary exam
actuary job
actuary science
insurance actuary
10 November 2015
Source: google.com
7


Institute and Faculty of Actuaries
Data Science

Institute and Faculty of Actuaries

What is a "Data Scientist"?



10 November 2015

?

A Data Scientist is a statistician who lives in San Francisco

Josh Wills
@josh_wills

Data Science is Statistics on a Mac.

Big Data Borat
@BigDataBorat

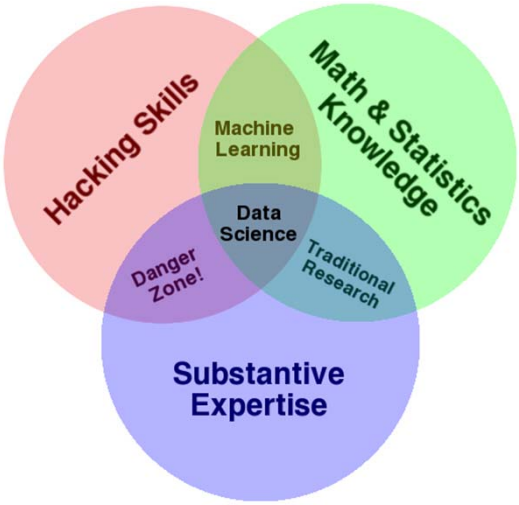
A Data Scientist is a device for turning coffee and data into better decisions

Likelihood T. Prior
@CjBayesian
Corey Chivers, Penn Medicine

Josh Wills
@josh_wills

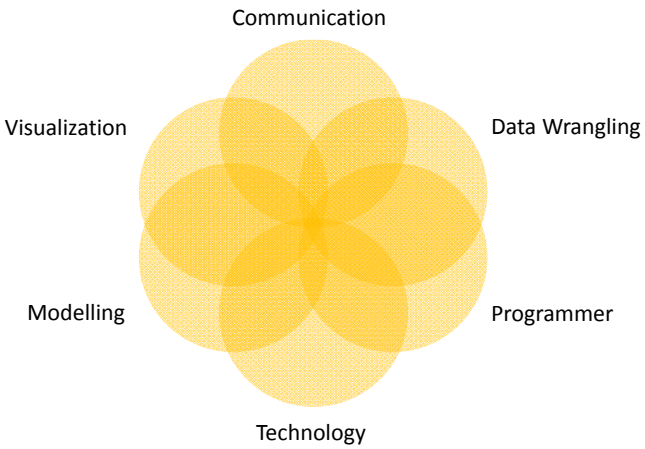
Data Scientist (n.): Person who is better at statistics than any software engineer and better at software engineering than any statistician.

Drew Conway
@drewconway



The View from Mango …

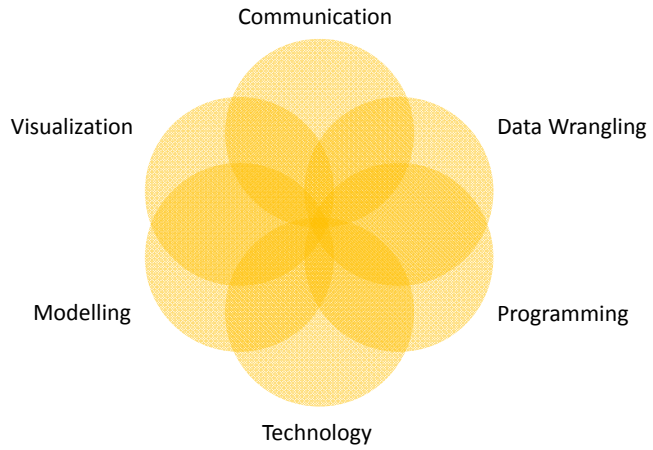
Data Science Skills



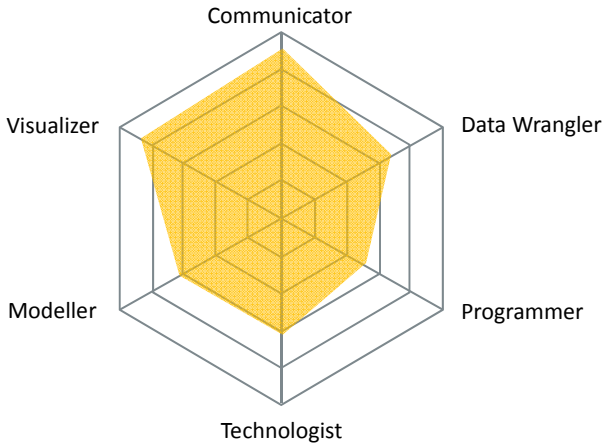
Skills Required

Skill	Statistician	Data Scientist
Communication	Communicate results	Interact with business to explore challenges
Visualization	Spot relationships in data	Clearly communicate insight
Data Wrangling	Manage modelling data	Integrate analytic workflows with data
Modelling	Analyse data	Understand how “solvable” business challenges are
Programming	Do Analysis	Embed analytics in deployable apps
Technology	Hardly needed	Understand technical aspects of solution – look for new ways to meet challenges

Data Science Skills

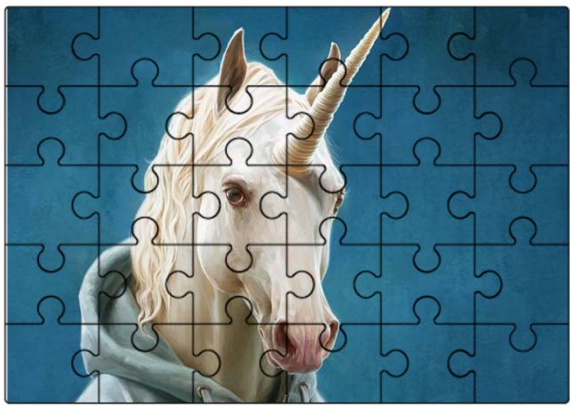


Data Science Radar



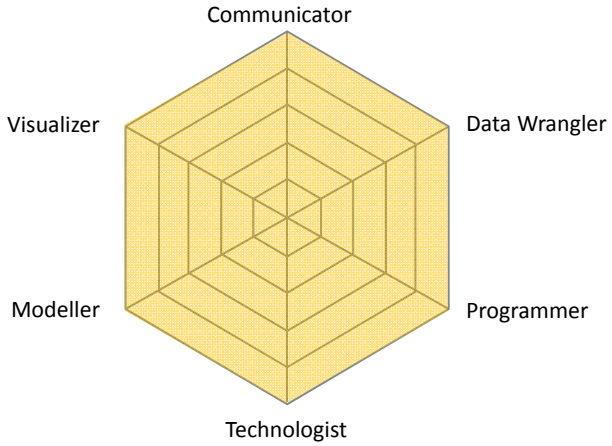
Data Science Radar

- Used at Mango, and by our clients to:
  - Understand the skills in the existing team
  - Understand training needs
  - Supporting the recruitment process
  - Build project teams
  - Promote discussions about Data Science best practices



Building the Unicorn

Who is the Unicorn?



The Search for the Unicorn

- Mango has been providing Data Science services since 2002
- We interview ~4 prospective Data Scientists each week
- The Unicorn. Doesn't. Exist

Building the Unicorn

- We can “build” the unicorn with teams of Data Scientists
- This allows us to represent all necessary skills for a project
- Then we need to ensure they can perform Analytics as a Team

The Data Science Remit

Back in 1998 …


Google!
BETA
Search the web using Google!
Google Search
I'm feeling lucky
Special Searches
Stanford Search
Linux Search
Help!
About Google!
Company Info
Google! Logos
Get Google!
updates monthly:
your e-mail
Subscribe
Archive
Copyright ©1998 Google Inc.

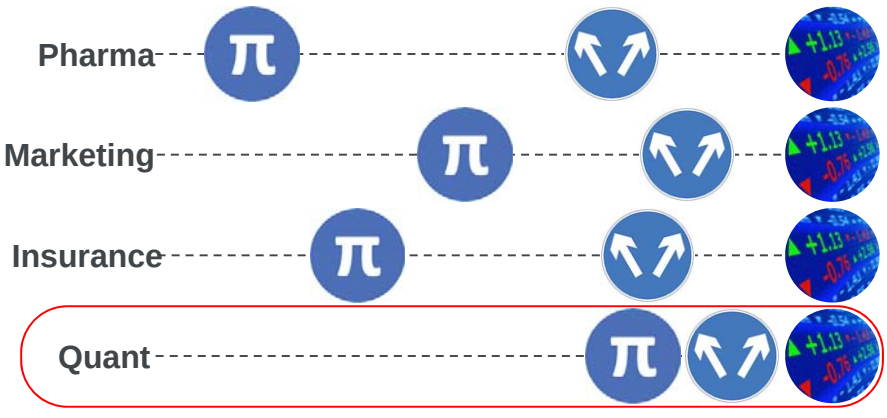

ARMAGEDDON

Missing from my Stats Degree …

- Domain Knowledge
- How be a Programmer
- How to create a good graph
- How to communicate results

Reactive Analytics

The Role of Analytics



Data Science
=
Team of Multi-Skilled Individuals
+
Proactive Analytics

Institute and Faculty of Actuaries

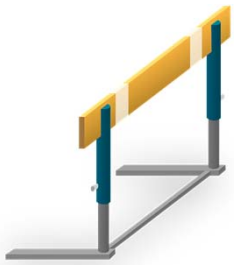
DS Case Study

Mortgage Asset Pricing

Project Overview

- Client was Investment firm
- (How much) should they bid for large mortgage portfolio
- ~120,000 Mortgages from a single EU Country
- Guide price ~€7bn

Challenges

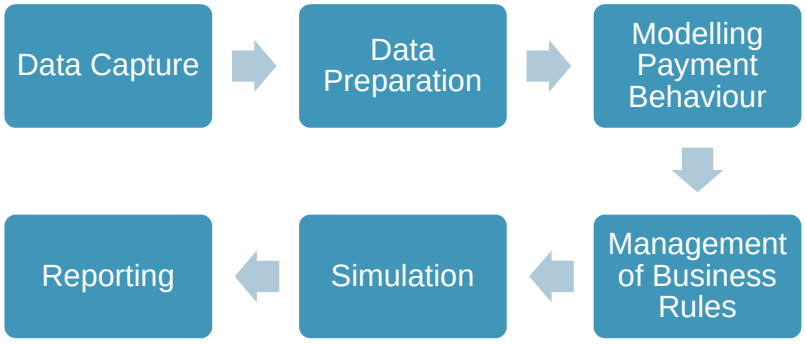


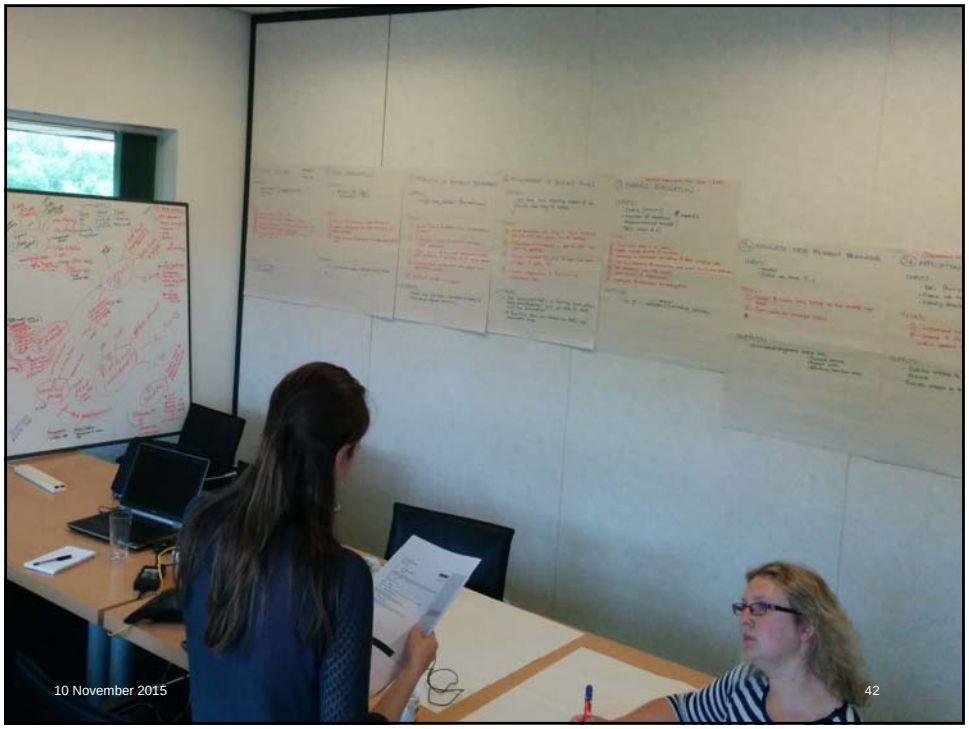
- Data held in multiple Excel Spreadsheets and was messy
- Lots of compute: 10,000 iterations x ~120,000 assets x 50 years x 8 sets of business rules
- Business Rules to be applied held as formulae in Spreadsheets
- Client had attempted for 6 months, now there were 6 days to the deadline!

Approach

- Set up collaborative environment
- Create team of 6 Data Scientists + 2 Business Analysts
- Split workflow into steps, then assign steps to members of the team
- Clearly define inputs and outputs to each step

The Process

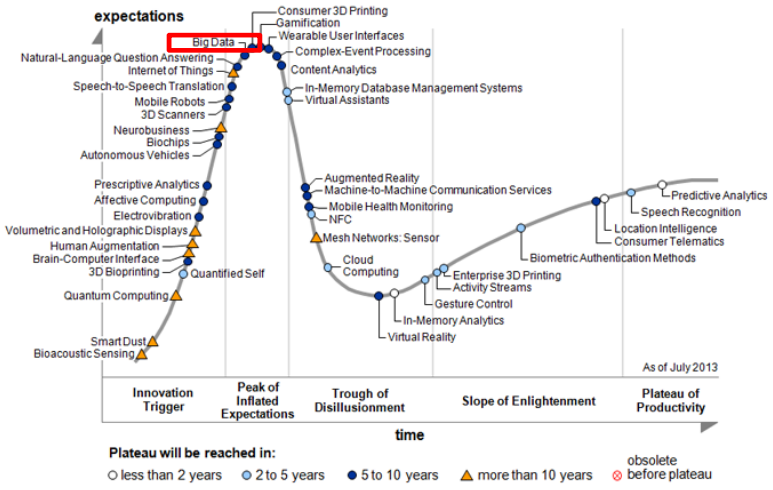

Data Capture
Data Preparation
Modelling Payment Behaviour
Management of Business Rules
Simulation
Reporting


10 November 2015
42

Skills Used by Data Scientists

- Collaboration
  - Version control
  - Package development
  - Continuous Integration
- Database Creation
- SQL Processing
- R Data Manipulation
- Modelling & Simulation
  - Logistic Regression (R)
  - Markov State Switching (R)
  - GBR Trees (Python)
- Scale using AWS
- Image Mgment with Docker
- Reporting using markdown

Big Data – Big Hype?

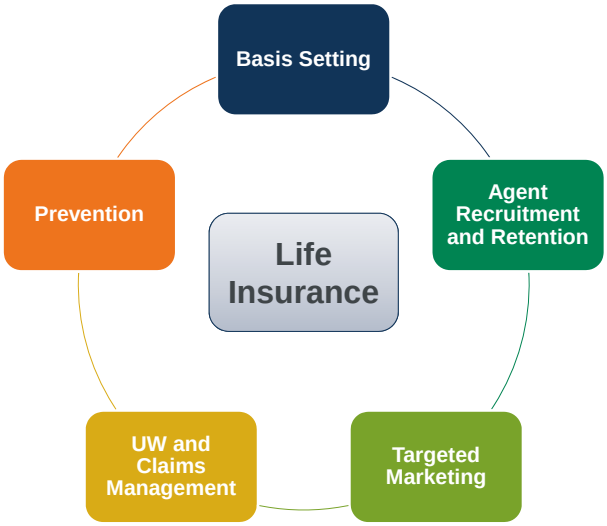


Source: Gartner (August 2013)

Data Sources for Life Insurers

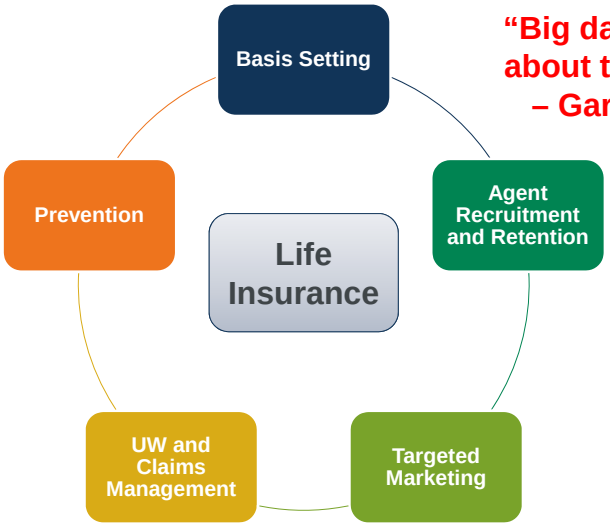
Traditional Data Sources	New Data Sources
Application Data	Search Engines
Inforce Experience Data	Social Media
CMI	Wearables
National Statistics (ONS, HES)	Genome Sequencing
Geodemographic Profilers	Subject Access Requests
Economic Data	??
Credit Ratings	

Life Insurance – Big Data Applications



10 November 2015 47

Life Insurance – Big Data Applications



"Big data is not about the data" – Gary King

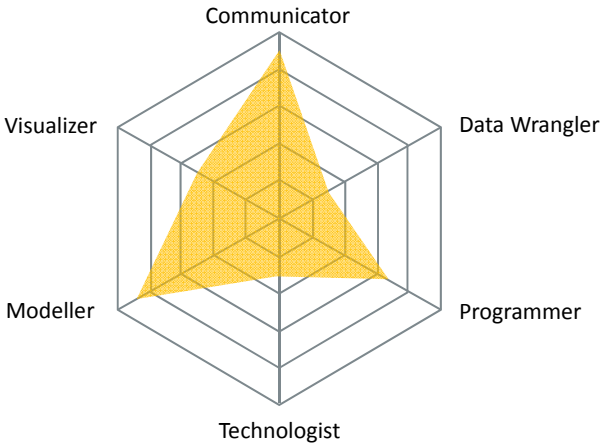
10 November 2015 48

Are Actuaries the Data Scientists of Insurance?

- ✚ Core statistical training
- ✚ Analytical thinking and quantitative skills
- ✚ Ongoing training
- ✚ Communicating insight to the business
- ✚ Domain Knowledge

- ▬ Limited skills to subdue unstructured data
- ▬ Unlikely to be better programmers
- ▬ Lacking the visual representation ability
- ▬ Actuarial models yet to reach “Big Data”
- ▬ Confirmation Bias

Data Science Radar



Actuaries – What's the SOA doing?

Actuaries in Advanced Business Analytics

E-learning

Launched "Application of Statistical Techniques" module

Instruction regarding use of R

Least Squares, GLMs, Cluster Analysis, Credibility using GLMs

Seminars

3 day seminar in ABA

Using R

Perform basic data manipulations

Graphically explore data

Interpret and critically examine model output

Learning Strategy

Broaden ABA education – task force established

New curriculum will include topics in predictive analytics

Post qualification products and services – first in predictive analytics

Actuaries – What's the UK Profession Doing?

CT3

Summarise a set of data

Describe the level and variability

Explain symmetry and skewness

CA2

Analyse data

Develop model with clear documentation

Analysis of model output

Interpretation of results

Communication of approach and results

Education Review

???

Completing the Jigsaw

Source: Fotolia

Questions

Comments

The views expressed in this [publication/presentation] are those of invited contributors and not necessarily those of the IFoA. The IFoA do not endorse any of the views stated, nor any claims or representations made in this [publication/presentation] and accept no responsibility or liability to any person for loss or damage suffered as a consequence of their placing reliance upon any view, claim or representation made in this [publication/presentation].

The information and expressions of opinion contained in this publication are not intended to be a comprehensive study, nor to provide actuarial advice or advice of any nature and should not be treated as a substitute for specific advice concerning individual situations. On no account may any part of this [publication/presentation] be reproduced without the written permission of the IFoA [*or authors, in the case of non-IFoA research*].