

```
library(dplyr)
```

```
rladies_global %>%  
  filter(city == 'London')
```



@AnalyticsPanda

Demystifying Data Science Job Hunting

UCL - Structures Lab
June '17



Objective

Your Data Science Transition

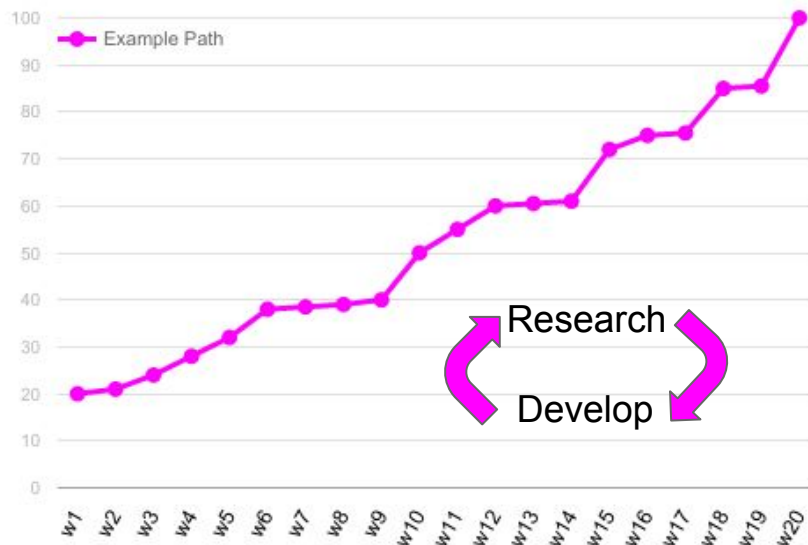


Identify
Goal/s



Upskilling: Succeed or Learn

Data Science Journey over time



Achieved!

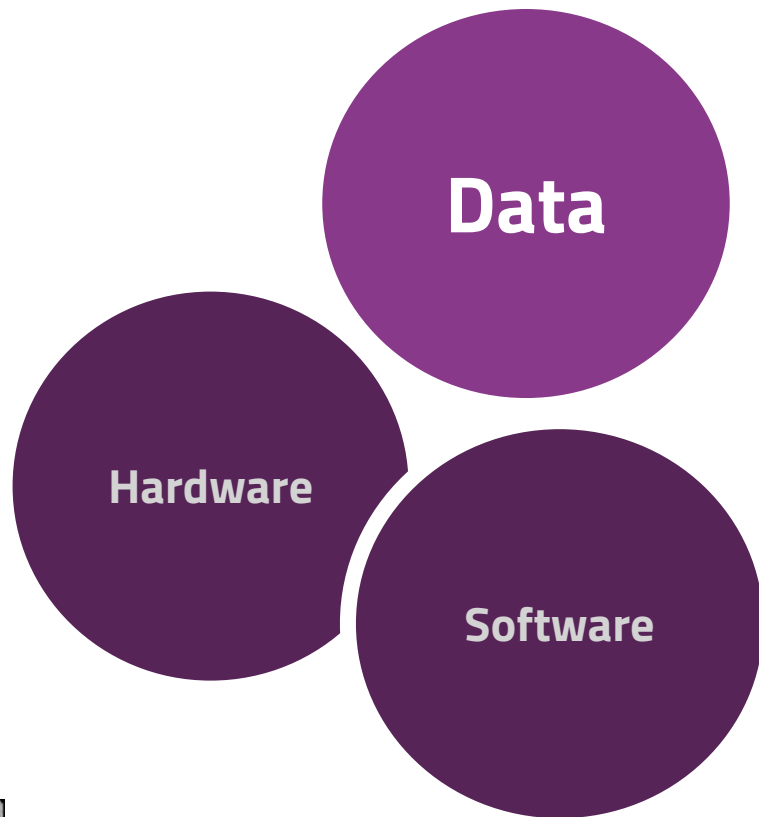




Base Definitions



Tech Industry - not all Apps, Robots & PhDs



P.S you do **not** have to have a PhD in Computer Science to work in Tech!!

“69% of all (coders) tell us they are at least partly self-taught.

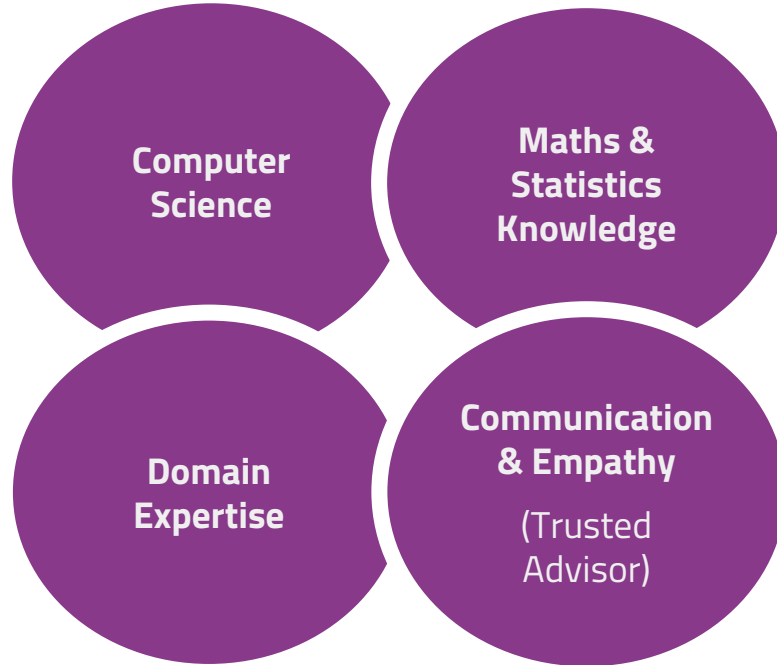
13% of respondents across the globe tell us they are **only** self-taught.”

—**Stack Overflow**
Developer Survey Results 2016
56,033 coders in 173 countries

What is Data Science?

My suggestion:

“Creating Value from Data by applying science/scientific methods”



- General consensus = There is no consensus!
- DJ Patil & Jeff Hammerbacher purport to have first coined the term “Data Scientist” as a job title in 2008
- Hybrid & Interdisciplinary
- P.S Don’t assume an Organisation recruiting for a Data Scientist actually knows what a Data Scientist is.....!

Data Science - Misconceptions

Data Science != Hadoop

Data Science != Big Data

Data Science != Machine Learning

Data Science != Just Programming

Data Science != Successful Data Science



Realities

Ambiguous Profession - no standardisation



- **Rule of Thumb: Assume every Data Science vacancy is Unique**
- Fast Tech Evolution + Differing Maturity of org/industry + Different literacy of the specific people involved in recruitment process + Current Projects + Current Tech Resources
- **Rule of Thumb: Prepare to advise the recruiters how to recruit, Don't expect the Hiring Manager has any better idea about the role than you do!**
- If you've been actively in job hunting, you'll almost certainly know the current market better than those making the hiring decision!!
- **Rule of Thumb: Keep focussing on what YOU can Deliver, and QUANTIFY as much as possible!! | MONEY TALKS**
- Obviously try and tailor what you can offer to the specificities of the exact vacancy, but if they don't appreciate what you offer, this may well be a reflection of their inabilities, NOT yours!
- A good company will be able to make good judgement
- Be open to feedback being valid, but don't assume it is: gender bias is still rife amongst both technical/untechnical



Framing, Framing, Framing - Sell Yourself

- **Rule of Thumb: Know your proposition (what you offer) like the back of your hand!!**
- You should be able to pitch yourself in 15 seconds, or 2 hours | And with confidence
- **Rule of Thumb: You don't need to tick all the job spec boxes, You just need to be the best candidate who applies | Success is RELATIVE, not absolute | Based on your competition**
- Job specs are wish-lists, NOT a representation of reality
- It is well-known that successful candidates will only fulfill maybe 60% of requirements
- **Rule of Thumb: You MUST portray yourself as best as possible | IMPRESS THEM!**
- Your competition are likely already over-inflating their CVs, so if you downplay yourself you've created an even bigger discrepancy
- **Rule of Thumb: Recruiters are time-pressured and making Snap Decisions, so make it easy for them to review you | IMPRESS THEM FAST!**
- The hiring manager may have 30secs to click Yes or No on your application to invite you to interview: you CANNOT afford to be humble, subtle and indirect!



Salary Negotiations - START AS HIGH AS POSSIBLE

- **Rule of Thumb: START HIGH**
- Employers will offer the absolute minimum they can get away with
- You will end up with a better result if you go high and get negotiated down vs go low initially and then try to suggest you meant you wanted more (also shows lack of judgement)
- **After all, if you undervalue yourself, why shouldn't they?**



Please takeaway these points!

TRYING TO BE LIKED & TRYING TO CONFORM WILL GET YOU ABSOLUTELY NOWHERE

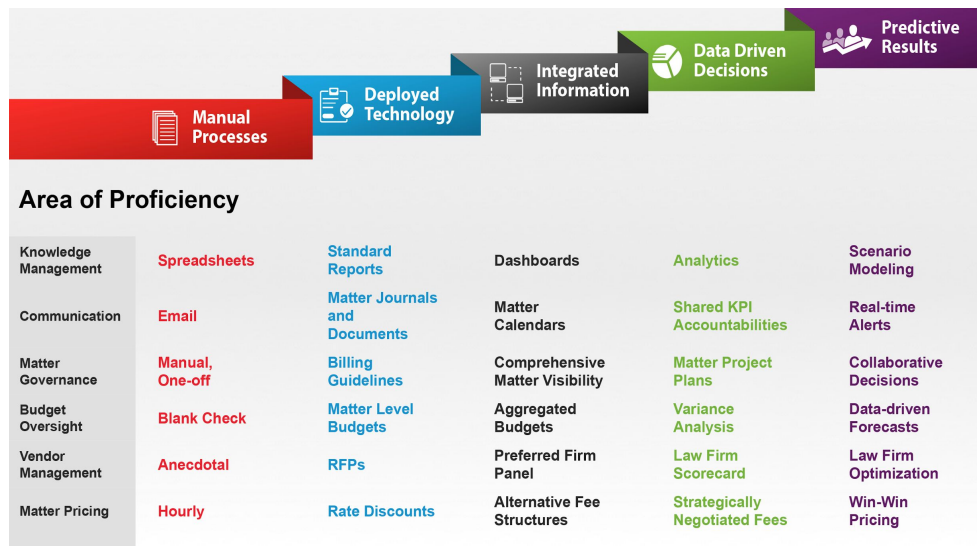
For you to WIN in Tech, you HAVE to play the game - it's your choice to play the game or not, but your likelihood of success will be minimal if you don't

We are not in a position where we can afford to be SELF-DEPRECATING | This is a COMPETITION, and as we all know, our minority identities already put us at a disadvantage - we must be prepared for a challenge!



Maturity

Maturity Model



Level	Structured Processes	Discoverability & Compounding	Analytical Speed & Agility	Breadth & Depth of Impact	Organizational Cohesion
1 Ad Hoc Exploration	Practitioners operate autonomously in a black box	Assets stored locally, emailed around	Limited talent and tools	Ivory tower, no tangible value	Analytics island, purely transactional
2 Repeating, but Limited	Recurring workflows discussed, no enforcement	Assets stored centrally, but lack metadata / permissions	Some tools and talent investment	Static reports in a few business areas	Some collaboration with line managers
3 Defined and Controlled	Formalized process, manually enforced	Assets stored and tagged centrally with metadata and permissions	Rapidly test ideas with novel methods / tools	Results translated into multiple operational workflows	Analytics are key stakeholders in strategic decisions
4 Optimized and Automated	Best practices codified into infrastructure, transparency for all	All asset versions stored/tagged, searchable, reproducible	Cutting-edge tools, comfortable at the analytical frontier	Data products drive org with robust safeguards	Analytics enmeshed in business and proactively anticipates needs



Targeting



Industries/Sectors

In order of DS sophistication

Legacy/Slow

Public Sector

Retail

Media

Marketing

Middling

Consultancies

Marketing

E-commerce

Finance

Cutting/Bleeding Edge

Academia

Major tech players

Tech start-ups

Latest self-reported Salaries for Data Scientists (London, Glassdoor)

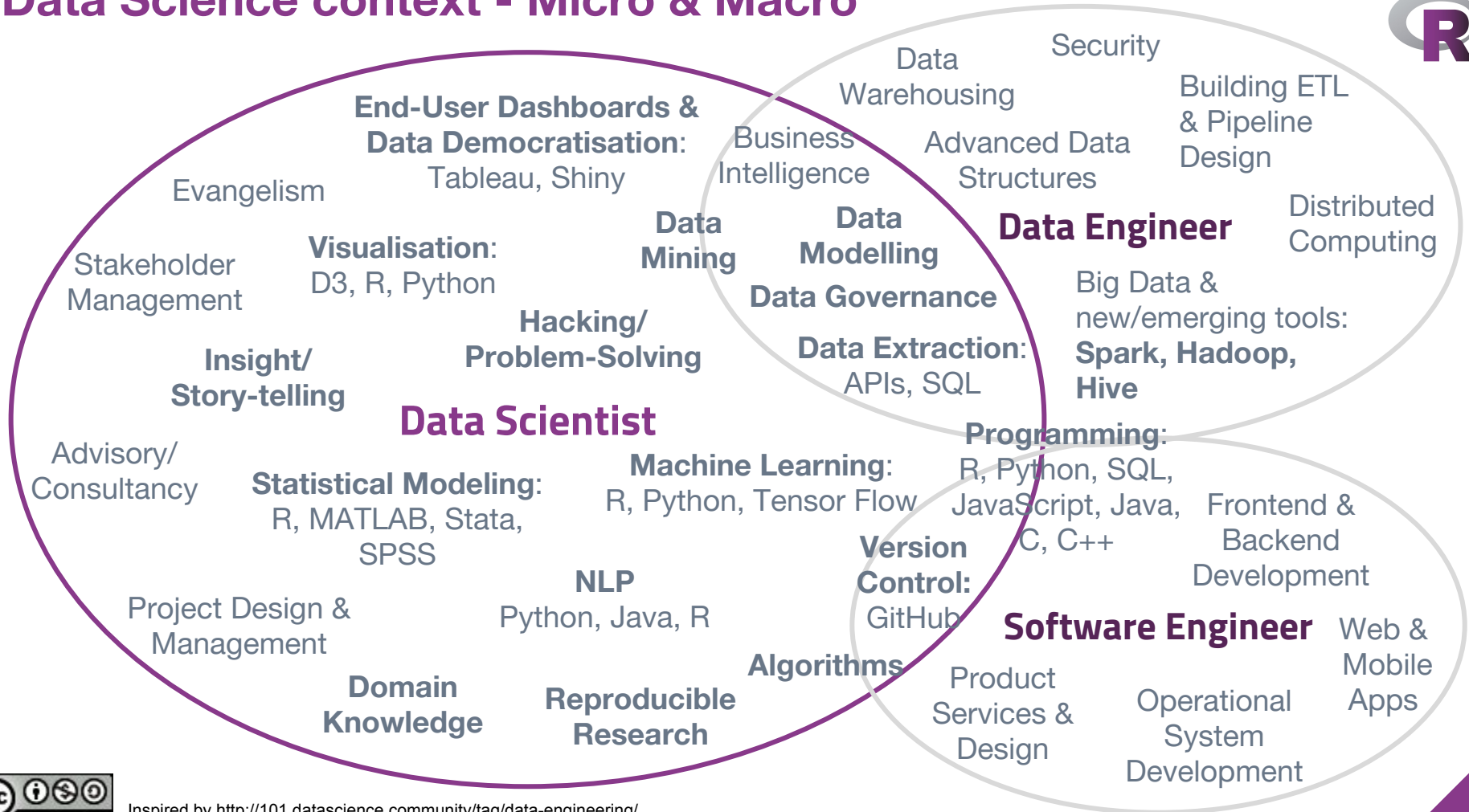
Data Scientist in London Salaries

Job Title	Location	Salary
Ocado Data Scientist	London, England (UK)	£38,172
Tesco Data Scientist	London, England (UK)	£45,296
Unilever Data Scientist	London, England (UK)	£36,878
Dunnhumby Data Scientist	London, England (UK)	£70,432
Imagini Holdings Data Scientist	London, England (UK)	£52,744
Virgin Media Data Scientist	London, England (UK)	£47,424
First Derivatives Data Scientist	London, England (UK)	£20,104
King Senior Data Scientist	London, England (UK)	£63,333



Skills

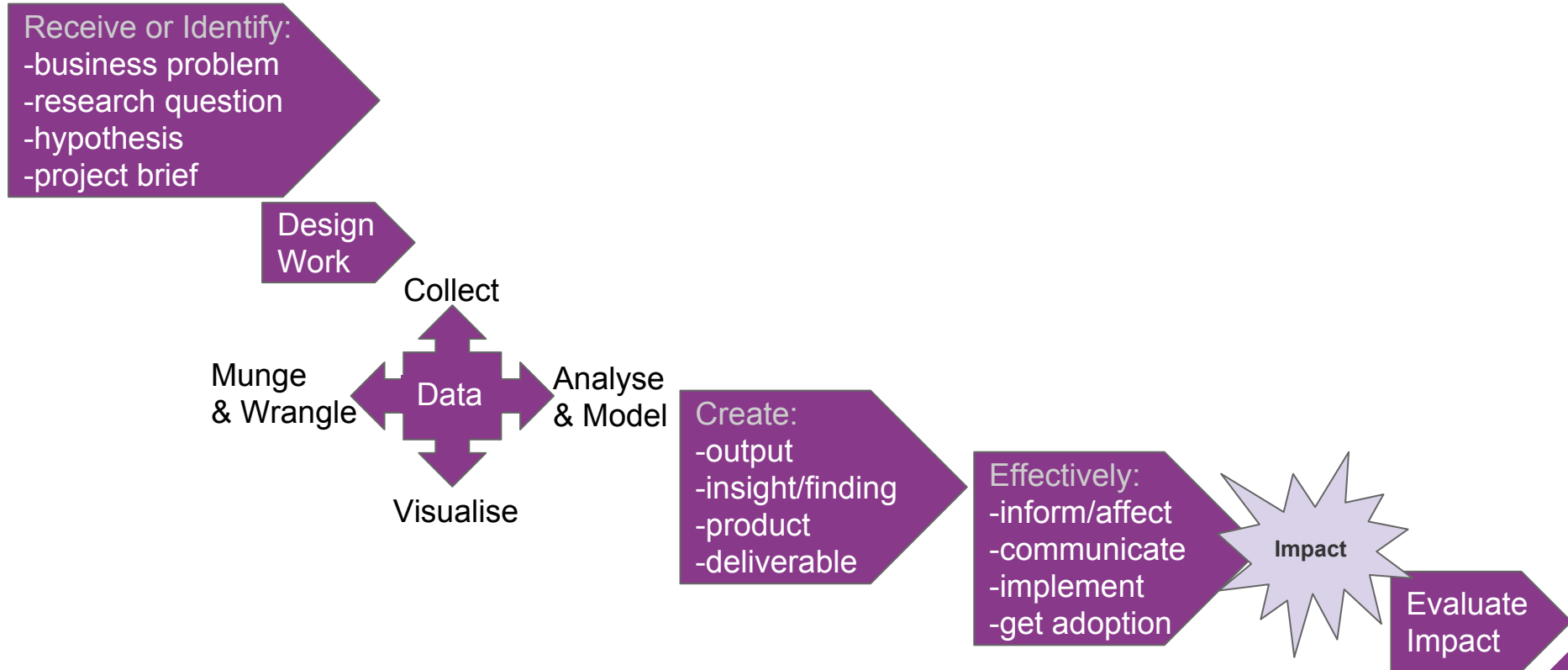
Data Science context - Micro & Macro





Workflow

Data Science - Workflow



Data Science - Workflow

Receive or Identify:

- business problem
- research question
- hypothesis
- project brief

Design
Work

Collect

Munge
& Wrangle

Data

Analyse
& Model

Visualise

Create:

- output
- insight/finding
- product
- deliverable

Effectively:

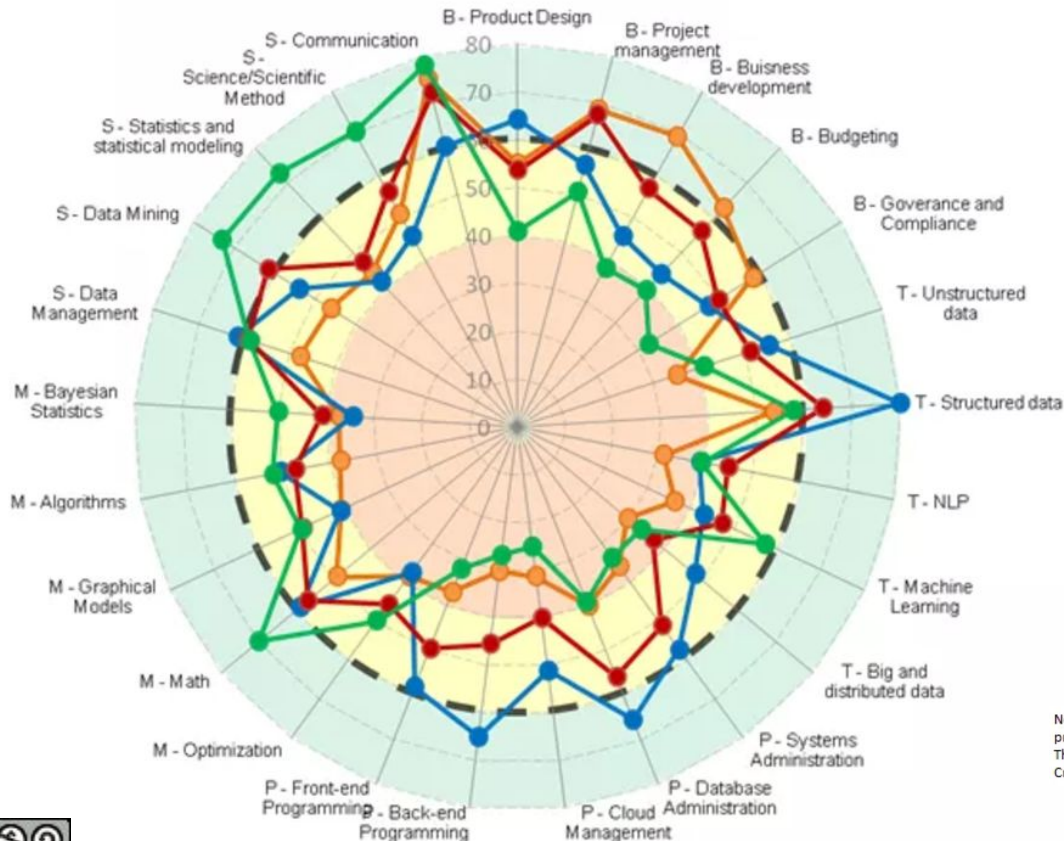
- inform/affect
- communicate
- implement
- get adoption

Impact

Evaluate
Impact

Computational data skills are
Necessary but **not** Sufficient!!

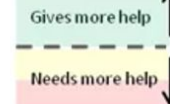
Data Science - The Unicorn = Group of Horses



Data Roles

- Business Management (e.g., leader, business person, entrepreneur)
- Developer (e.g., developer, engineer)
- Creative (e.g., Jack of all trades, artist, hacker)
- Researcher (e.g., researcher, scientist, statistician)

Skill Levels



- Don't know (0)
- Fundamental Knowledge (20)
- Novice (40)
- Intermediate (60)
- Advanced (80)
- Expert (100)

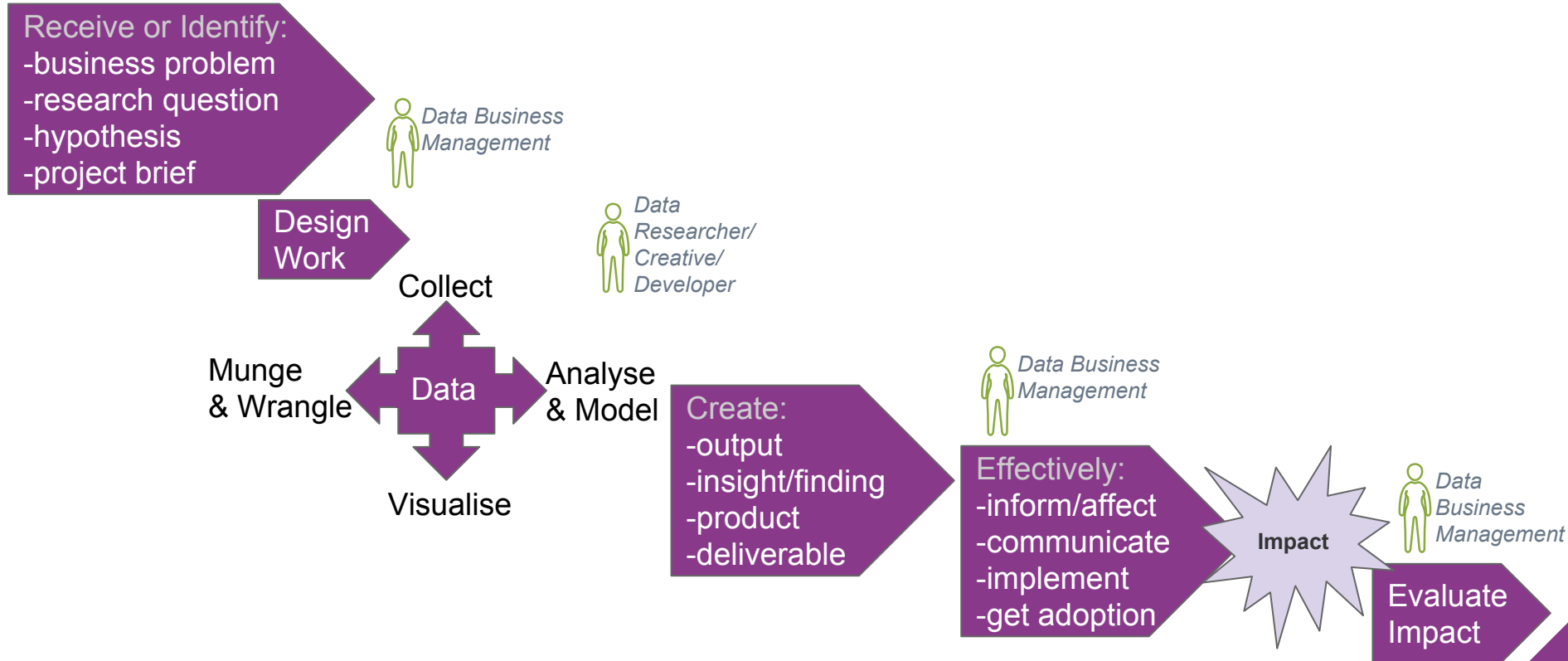
Note: Data are based on responses from 490 data professionals. Data professionals were asked to rate their proficiency across 25 skills using a scale from 0 (don't know) to 100 (expert). This graph is based on respondents who selected only one primary job role. Business (n = 65); Developer (n = 47); Creative (n = 25); Researcher (n = 101)



Data Science - The Unicorn = Group of Horses

- **Leader, Business person, Entrepreneur | Data Business Management** are the product and profit-focused data scientists. They're leaders, managers, and entrepreneurs, but with a technical bent. A common educational path is an engineering degree paired with an MBA.
- **Developer, Engineer | Data Developers** are focused on writing software to do analytic, statistical, and machine learning tasks, often in production environments. They often have computer science degrees, and often work with so-called "big data".
- **Generalist, Hacker | Data Creatives** are eclectic jacks-of-all-trades, able to work with a broad range of data and tools. They may think of themselves as artists or hackers, and excel at visualization and open source technologies.
- **Researcher, Scientist, Statistician | Data Researchers** apply their scientific training, and the tools and techniques they learned in academia, to organizational data. They may have PhDs, and their creative applications of mathematical tools yields valuable insights and products.

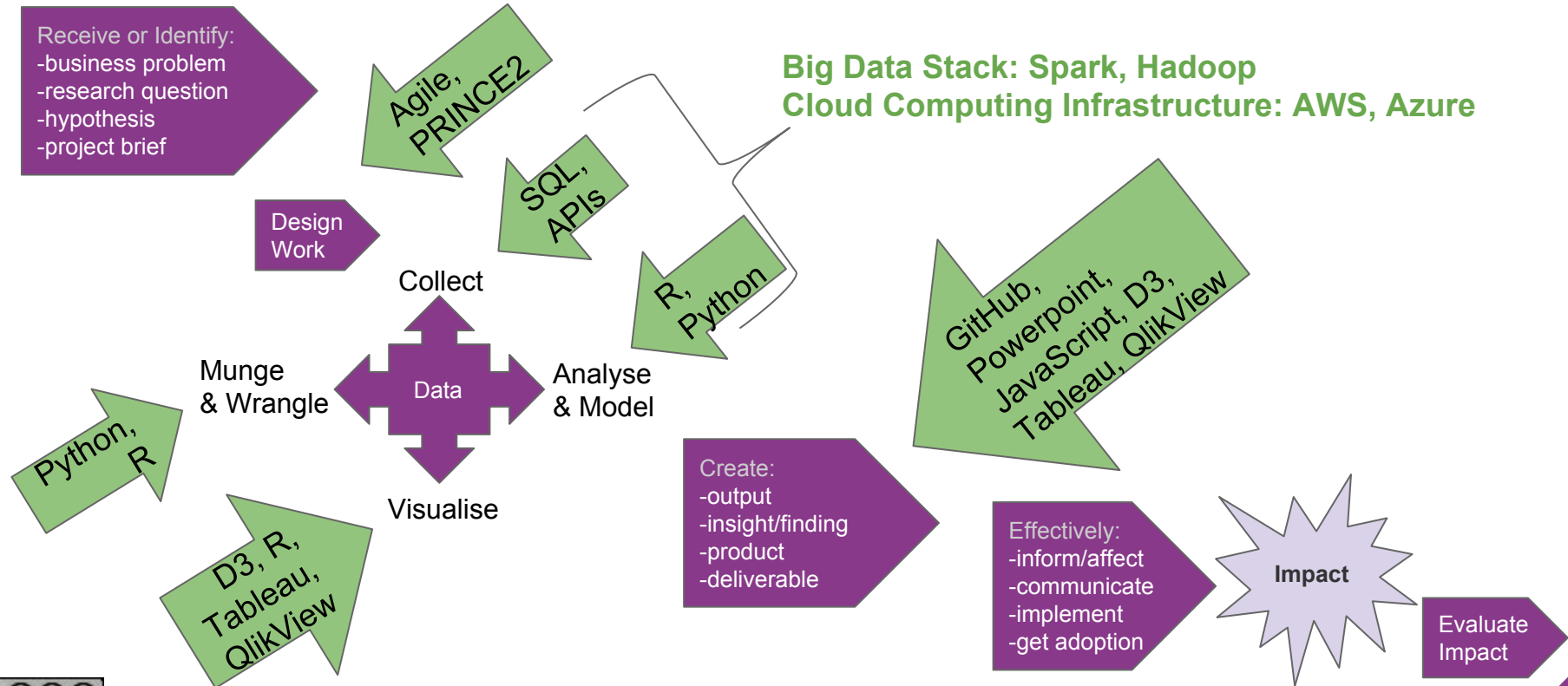
Data Science Workflow





Prepare

Data Science Workflow - where tools fit



To Do - Interview Questions

Core

- Walk me through your CV
- Why do you want this job
- Why are you suitable for this job
- Why are you looking to move from your current role/situation
- What are your salary expectations

Possible

- Greatest (data science-related) achievement
- Explain any gaps in your CV
- What are your weaknesses

Vacancy-Specific

- Are you aware of the domain-specific challenges relating to <eCommerce/Finance/Gov/Health etc>?

Things to Ask

- What are the next steps of the process?
- What is the provenance/background to this vacancy?
- Where does this role sit in the organisation structure, e.g. under Chief Tech Officer, Chief Marketing Officer?
- Why did you (to interviewer) take this job/What have you enjoyed about this job?
- What is the organisation's vision/strategy/position w.r.t Data Science/Tech, e.g. investing, exploring?

To Do

Fundamentals

Useful

If Possible

LinkedIn



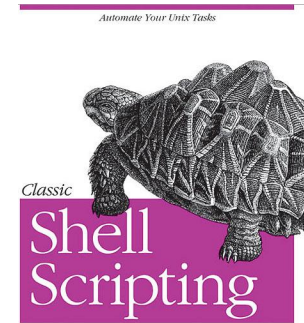
L^AT_EX



kaggle



codewars





Example hand-picked live Data Science roles

(at time of writing)



Data Scientist
Entry-Level

<https://www.linkedin.com/jobs/view/361665375/>



Data Analyst
Associate (Junior) |
Excel & SQL

<https://www.linkedin.com/jobs/view/364310099/>



Junior Data Scientist
Entry-Level
SQL & R

Reasonable
Requirements

<https://www.linkedin.com/jobs/view/340146959/>



British
RedCross

Data Scientist
Mid-Senior

Reasonable Requirements

£34,136 - £47,450 + London weighting
<https://www.linkedin.com/jobs/view/322581997/>



Data Scientist
Mid-Senior

Reasonable Requirements

<https://www.linkedin.com/jobs/view/297108433/>



Junior/Graduate Data Scientist
Excel & SQL (Access)

Easy Requirements

<https://www.linkedin.com/jobs/view/357961672/>



Insight Analyst - Data Science
Entry-Level

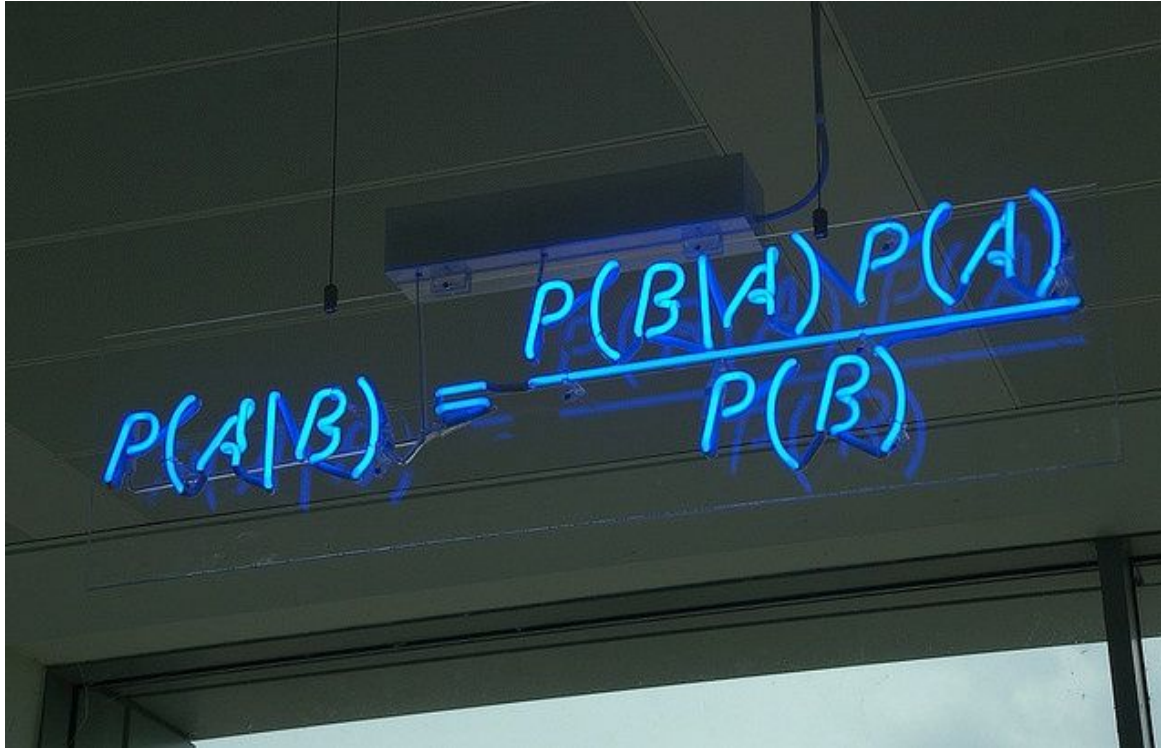
Reasonable Requirements

<https://www.linkedin.com/jobs/view/304630939/>



Let's Go!

Data Science is awesome!!!



A photograph of a blue neon sign mounted on a dark ceiling. The sign displays the formula for Bayes' theorem: $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$. The sign is illuminated with a bright blue light, and the background is dark.

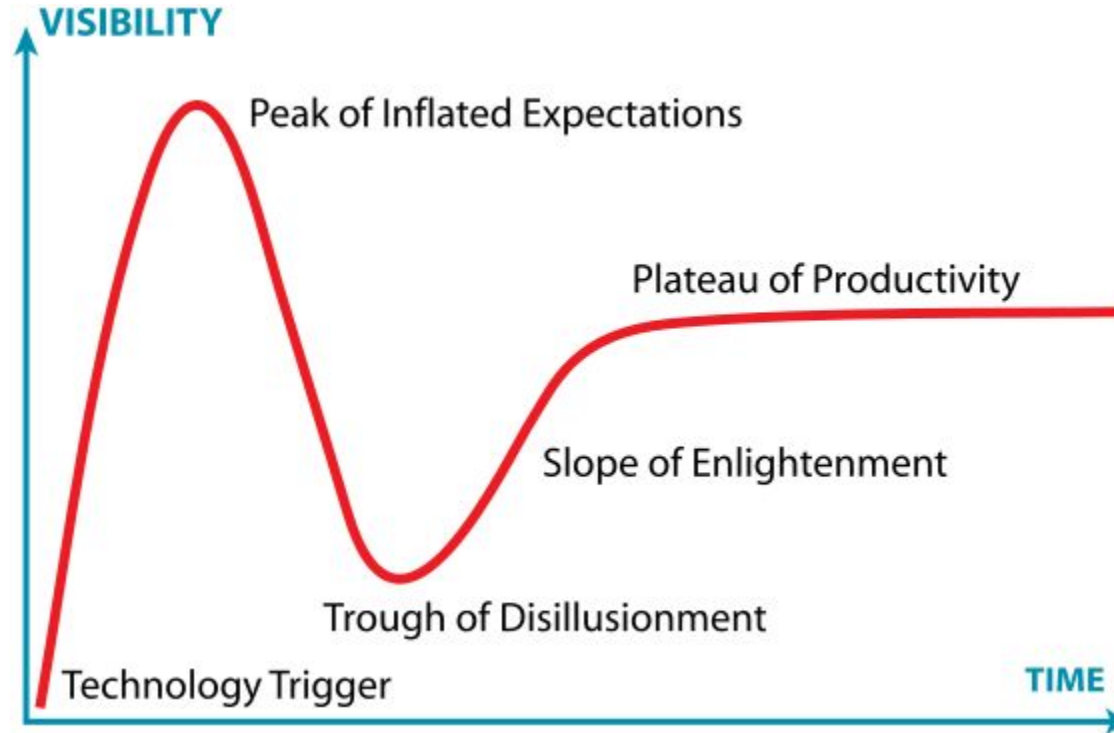


Appendix



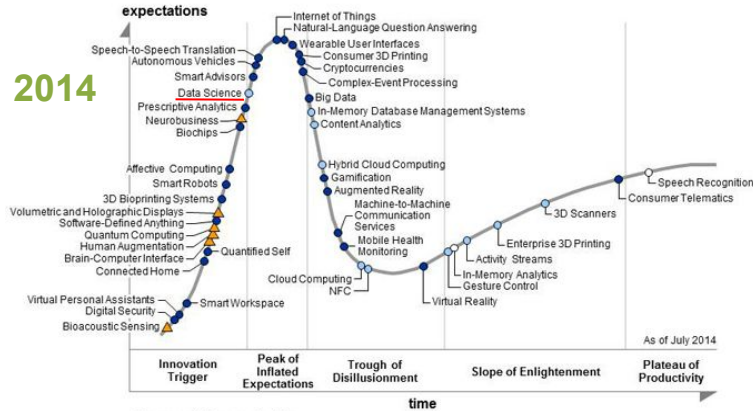
Hype

Hype Cycle

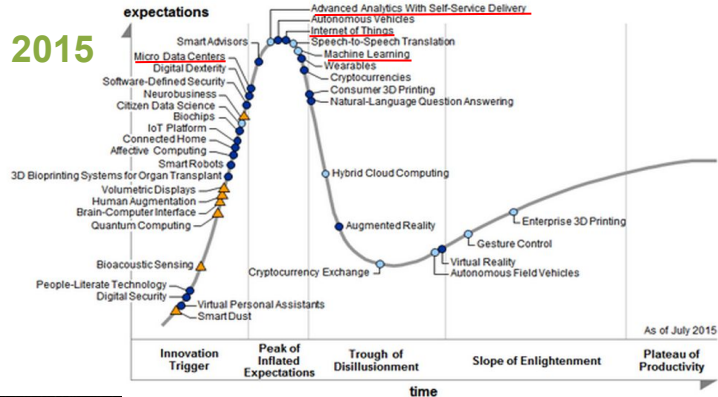


Hype Cycle - Silver Bullets?

2014



2015



2016



Years to mainstream adoption:

○ less than 2 years ● 2 to 5 years ● 5 to 10 years ▲ more than 10 years ✖ obsolete



Pitfalls

Data Science Pitfalls - Noise

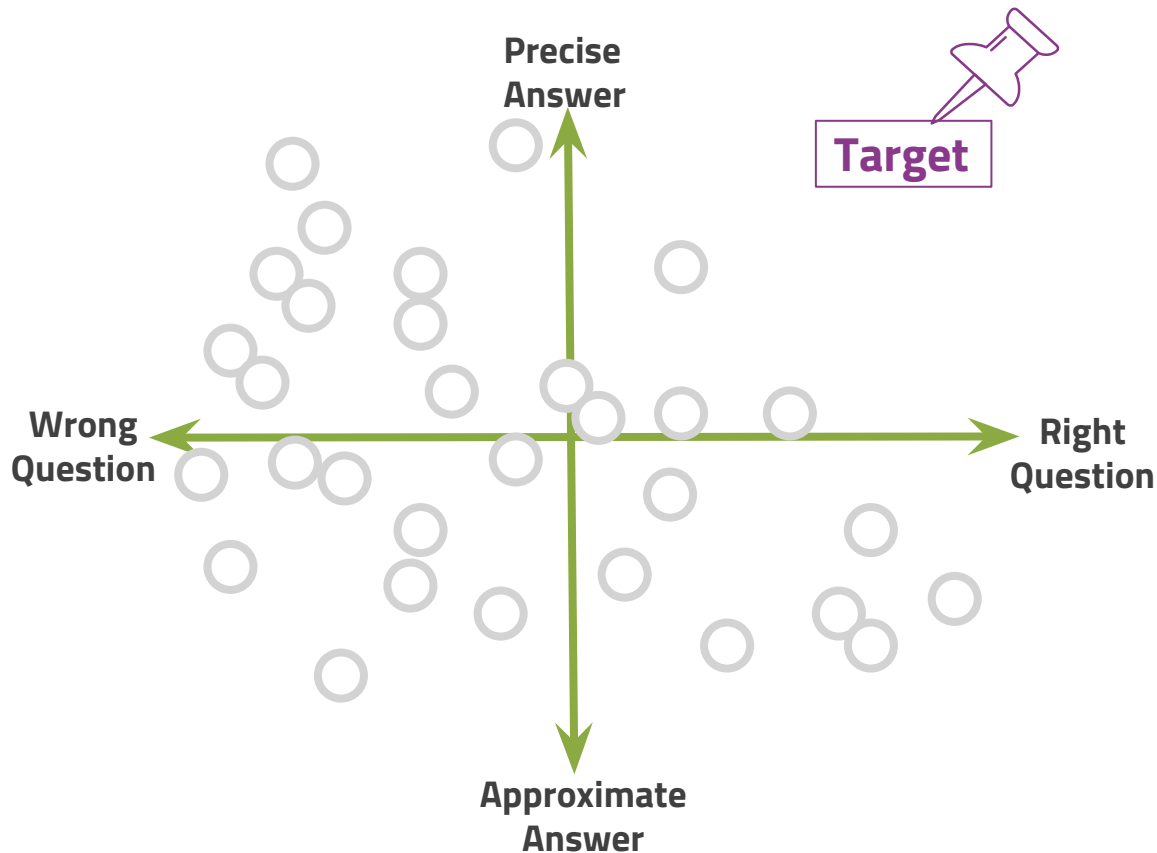


Data Science Pitfalls - Wrong Question

"Far better an **approximate answer to the right question**, which is often vague, than an exact answer to the wrong question, which can always be made precise."

—*John W. Tukey*

The Future of Data Analysis (July 1961)



Data Science Pitfalls - Blind Leading the Blind?



Data Scientist

The Co-op

Manchester, England

£60,000 to £80,000 p.a

We're looking to recruit an experienced Data Scientist to join our team, and unearth information from our data assets that'll help Co-op make smarter decisions for our members. You'll work as an **internal data consultant** with people from right across the Co-op family to find simple solutions to complex problems. And if you can bring us the **combination of data expertise & great people skills** we need, we can offer you a unique opportunity to help build our data capabilities from the ground up. To be successful as a Data Scientist at Co-op, you'll need to be **an established data expert**. Statistical models will **be your speciality**, but in addition to tried and tested technical ability, it's important you've got the skills to put these techniques to good use in our business.

What you'll do:

- Understand Co-op objectives, and **create algorithms** that deliver positive impacts for members
- Be capable and confident in tackling the most complex data-driven and analytical problems
- **Work across several simultaneous projects** with partners from various Co-op businesses
- **Take on different roles on different projects (expert adviser/statistical guru/lead developer)**
- **Write code** when you need to in order to tackle and troubleshoot the most complex tasks
- Make sure new problem solving approaches comply with existing Co-op rules and values
- Plan projects over a 12 month horizon, and **manage budgets using Agile principles**
- Contribute to the **coaching & mentoring** of a pool of less experienced colleagues
- Design solutions that create outputs which are **accessible to non-technical audiences**
- Be a great **advocate for Data Science, promoting our services** to Co-op colleagues

Data Science Pitfalls - Blind Leading the Blind?



Data Scientist
Telegraph Media Group
London, GB

Requirements:

- **Robust experience in SQL** (Hive QL or any other flavour) with databases spanning tens of millions of records
- Solid fundamental understanding of statistics, and experience with statistical analysis using **R or Python** from university degree or previous role
- Be **knowledgeable about the digital ecosystem**
- You are a **communicative person that values building strong relationships** with colleagues and stakeholders and have the **ability to explain complex topics in simple terms**
- A **self-starter** who's comfortable working autonomously
- **Solid commercial skills and business awareness**

To be successful in this role, you will need:

- Outstanding logical thinking, a **very strong background in a quantitative discipline such as mathematics, physics, engineering, operations research**, data science or similar, and a scientific approach to problem solving
- Demonstrate **ability to manage stakeholder, show diplomacy and be an effective influencer**
- Good technical skills (at least some **practical experience with statistical/pattern recognition/machine learning methods**)
- A **sharp focus on getting results** and insights using the most practical and fastest approach (**even if it's not the most challenging or interesting for you**)
- Ability to **perform in a high pressure, fast paced environment** and be comfortable with a high level of ambiguity
- Ability to **communicate complex concepts** concisely and clearly, including **to senior level executives**
- **High energy and a can do, no-excuses attitude**
- Curiosity and **passion for making an impact**
- **Good interpersonal skills**

Behavioural Competencies:

- Bright, driven, love challenges and **want to change the editorial world**