Data science in the wild: Data science as a career
**data·scientist** *n.* a person employed to analyse and interpret complex digital data, such as the usage statistics of a website, especially in order to assist a business in its decision-making.
data·scientist \textit{n.} job title adopted by statisticians and business analysts to achieve significant pay increases.
the good news
AI will generate **$2.9 trillion** in business value and recover 6.2 billion hours of worker productivity by 2021.

*Gartner Predictions* (via Forbes)
Forrester Predictions 2017
Artificial Intelligence Will Drive The Insights Revolution

AI-driven companies will take $1.2 trillion from competitors by 2020
COMPANIES CAN’T AFFORD TO WAIT

Companies are already reaping massive value from AI.

“By the time a late adopter has done all the necessary preparation, earlier adopters will have taken considerable market share - they'll be able to operate at substantially lower costs with better performance. In short, the winners may take all and late adopters may never catch up.”

Source: Notes from the AI frontier: Modeling the impact of AI on the world economy
Most large organizations recognize the potential of AI and have embarked on a journey to become AI driven.

AI mandates globally are clear

- **80%** AI ADOPTION
  - Large global companies in working to adopt
  - (McKinsey Global Survey 2018)

Investments are substantial

- **55%** FORTUNE 500
  - Have invested $50M+ in AI
  - (HBR 2019)

Pace of investment and change is accelerating

- **90%** Executives reporting increasing urgency to invest at a higher level
  - (Harvard Business Review 2019)

- **$1.9 TRILLION**
  - AI derived business value in 2019
  - (Gartner 2018)

- **68%** Invested in a Chief Data Officer
  - (HBR 2019)

- **$3.9 TRILLION**
  - AI derived business value in 2022
  - (Gartner 2018)
Data Scientists Are Scarce

**Programming**
Get, manipulate & explore data. Build & implement models.

- R
- Python
- Spark
- Hadoop

**Math & Stats**
Foundational statistics. Internals of algorithms. Practical knowledge & experience.

- Logistic Regression
- GLM
- GBM
- Random Forest
- Decision Trees
- Neural Nets
- Deep Learning
- Text Mining
- Feature Engineering
- Blending
- Cross Validation

**Domain Expertise**
Understand the business problem & data

- Products
- Business rules
- Customers
- KPIs
- Regulations
- Accounting

**UK data science salaries**

### PERMANENT - AVERAGE ANNUAL SALARY

<table>
<thead>
<tr>
<th>DATA SCIENCE</th>
<th>Entry Level</th>
<th>Mid-Level</th>
<th>Technical Lead</th>
<th>Director</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Scientist</td>
<td>£45k</td>
<td>£34k</td>
<td>£65k</td>
<td>£60k</td>
</tr>
<tr>
<td>Quantitative Analyst</td>
<td>£55k</td>
<td>£43k</td>
<td>£83k</td>
<td>£66k</td>
</tr>
<tr>
<td>Machine Learning Eng.</td>
<td>£44k</td>
<td>£36k</td>
<td>£67k</td>
<td>£58k</td>
</tr>
</tbody>
</table>

*Annual Base Salary - £ GBP [London | Outside London]*

### CONTRACT - AVERAGE DAY RATES

<table>
<thead>
<tr>
<th>DATA SCIENCE</th>
<th>Entry Level</th>
<th>Mid-Level</th>
<th>Technical Lead</th>
<th>Director</th>
</tr>
</thead>
<tbody>
<tr>
<td>Role type</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Scientist</td>
<td>£650</td>
<td>£600</td>
<td>£750</td>
<td>£700</td>
</tr>
<tr>
<td>Machine Learning Eng.</td>
<td>£575</td>
<td>£550</td>
<td>£675</td>
<td>£625</td>
</tr>
</tbody>
</table>

*Daily Contract rate - £ GBP [London | Outside London*

**Source:** Harnham
US data science salaries

<table>
<thead>
<tr>
<th>DATA SCIENCE</th>
<th>Entry Level</th>
<th>Mid-Level</th>
<th>Technical Lead</th>
<th>Head of/VP</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Scientist</td>
<td>$121k</td>
<td>$105k</td>
<td>$152k</td>
<td>$135k</td>
</tr>
<tr>
<td>Deep Learning &amp; AI</td>
<td>$126k</td>
<td>$131k</td>
<td>$170k</td>
<td>$165k</td>
</tr>
<tr>
<td>M. Learning Engineer</td>
<td>$110k</td>
<td>$96k</td>
<td>$165k</td>
<td>$150k</td>
</tr>
<tr>
<td>Natural Lang. Processing</td>
<td>$133k</td>
<td>$125k</td>
<td>$170k</td>
<td>$160k</td>
</tr>
<tr>
<td>Computer Vision</td>
<td>$120k</td>
<td>$110k</td>
<td>$150k</td>
<td>$140k</td>
</tr>
</tbody>
</table>

Average annual salary

AVERAGE SALARY INCREASE ACHIEVED
15% When accepting a role over the last year.

AVERAGE BONUS
20% 65% of Data Science professionals received a bonus over the last year.

HAVE FLEXIBLE WORKING OPTIONS
68% Including flexible hours and the ability to work from home, the highest percentage in the industry.

NUMBER OF FEMALE PROFESSIONALS
16% This is the lowest in the industry, and a decrease from 18% last year.

Source: Harnham
but
85% of big data projects fail

(Gartner, 2017)
4% of executives say their business has successfully implemented AI

(PwC, 2018)
Few have been able to scale AI in a meaningful way...

Most organizations are in the early days of their AI journey and have not yet been able to achieve enterprise-wide scale...

**30% PILOTING**
AI in the business
(Mckinsey Global Survey 2018)

**52% NO CLEAR KPIs**
Instead focused on piloting and learning
(EY/MIT Technology Review Insights 2018)

**50% IMPLEMENTED**
AI in standard business process
(Mckinsey Global Survey 2018)

**80% SINGLE BU**
Focused on adopting; not yet enterprise-wide
(Mckinsey Global Survey 2018)

**20%** Of companies adopt the core practices to enable AI at scale
(Mckinsey Insights)

**85%** Of big data projects fail. “[T]he problem isn't technology. [It’s] org and cultural change”
(Gartner Nov 2017)

**93%** Executives identify people & process issues as key obstacle
(Harvard Business Review 2019)

...Scaling AI demands new ways to engage business experts with technology
(BCG Henderson Institute)

...and many are encountering significant challenges
...as a result of key challenges (2/2)

**Recognition**
Spotting and prioritizing the highest potential AI opportunities

**Alignment**
Framing AI opportunities in a compelling manner that secures sponsorship and executive focus from the business

**Delivery**
Effectively managing AI project delivery

**Measurement**
Defining valid pilot strategies and measuring business outcomes

**People**
Meeting business AI demand with existing DS staff

**Change**
Identifying and implementing the needed change management (communication, training, role-modeling, KPI changes)

**Pace**
Delivering DS projects quickly; efficiently preparing data and evaluating models

**Quality**
Executing projects with consistently high quality in a rapidly evolving field

**Deployment**
Efficiently and quickly putting models into production

**Technical**

**Monitoring**
Managing (multiple) in-production models
GETTING FROM DATA TO VALUE WITH AI HAS RISKS

Anywhere Along the Journey from Data to Value your AI Project Can Stall and Fail!

Data Prep

Feature Engineering

Model Creation and Evaluation

Model Ops and Deployment

Model Management and Monitoring

DATA

VALUE

005 SILOED TEAMS
Average number of teams that an AI project passes through before getting to the business consumer

007 DIFFERENT TOOLS
Average number of Data Science tools in organizations today to perform AI

012 DIFFERENT LANGUAGES
Average number of languages that can be used
Data Scientists

People who understand your business

Your employees
the idea
no more skunkworks
It’s been a long time coming. But I know a change is gonna come.

Sam Cooke
Generalist data scientist
Industry specialist
Deep specialist
Analytics developer
Data engineer

ROI VS DIFFICULTY PLOT

Complexity

Impact

Needs
AI-Enabling
Architecture
The global demand for machine learning & AI solutions greatly exceeds the production capacity of all the data scientists in the world, and this gap is growing exponentially.
Machines must learn to learn on their own.
Simple maths dictates that the vast majority of AI applications will be developed automatically by machines.
Simple maths dictates that the vast majority of AI applications will be developed automatically by machines.

“Automation-First”
Data Scientists: 10-100x increased productivity for the world’s few real applied data scientists

Democratize AI: Enable 1000x more AI Producers: business analysts, software developers, data engineers, RPA devs
black boxes?
Not necessarily
Humans have to pass job interviews. AIs should, too.
| How well did you perform at your job?  
What were your exam results like? | How accurate is the model?  
How well does it perform on unseen data? |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>How would you deal with a particular task?</td>
<td>What steps does the model take?</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>What information do you typically need to make a decision?</td>
<td>Which inputs/variables were important?</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>What typically influences the decisions you make?</td>
<td>What patterns did the algorithm find?</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>--------------------------------</td>
</tr>
</tbody>
</table>
| Here are some case studies. Tell me the decisions you would make, and why you would make that decision. | Why did this data point have that value?  
Which inputs determined this output? |
|--------------------------------|--------------------------------|
| Show me your references. | Document how you work.  
And show me your academic citations. |
How well does your AI answer these questions?