

Data Science in Industry

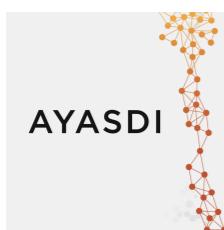
Johan Grahnen, PhD

2022-08-05

Summary

- Data science is decision support for money
- The job involves a lot of different tasks, and is very different from what you do now
- Company size is a very important variable
- Interviewing is difficult but persistence pays off

Why I am qualified to tell you about this



- PhD, U. Wyoming (2007 - 2012)
- Postdoc, Georgia Tech (2012 - 2013)
- Principal Data Scientist, Ayasdi (2013-2015)
- Data Scientist, Auransa (2015 - 2018)
- Senior Data Scientist, Microsoft (2018 - 2020)
- Senior Data Scientist, Lucid Motors (2020 - 2022)
- Staff Data Scientist, Angi (2022 -)

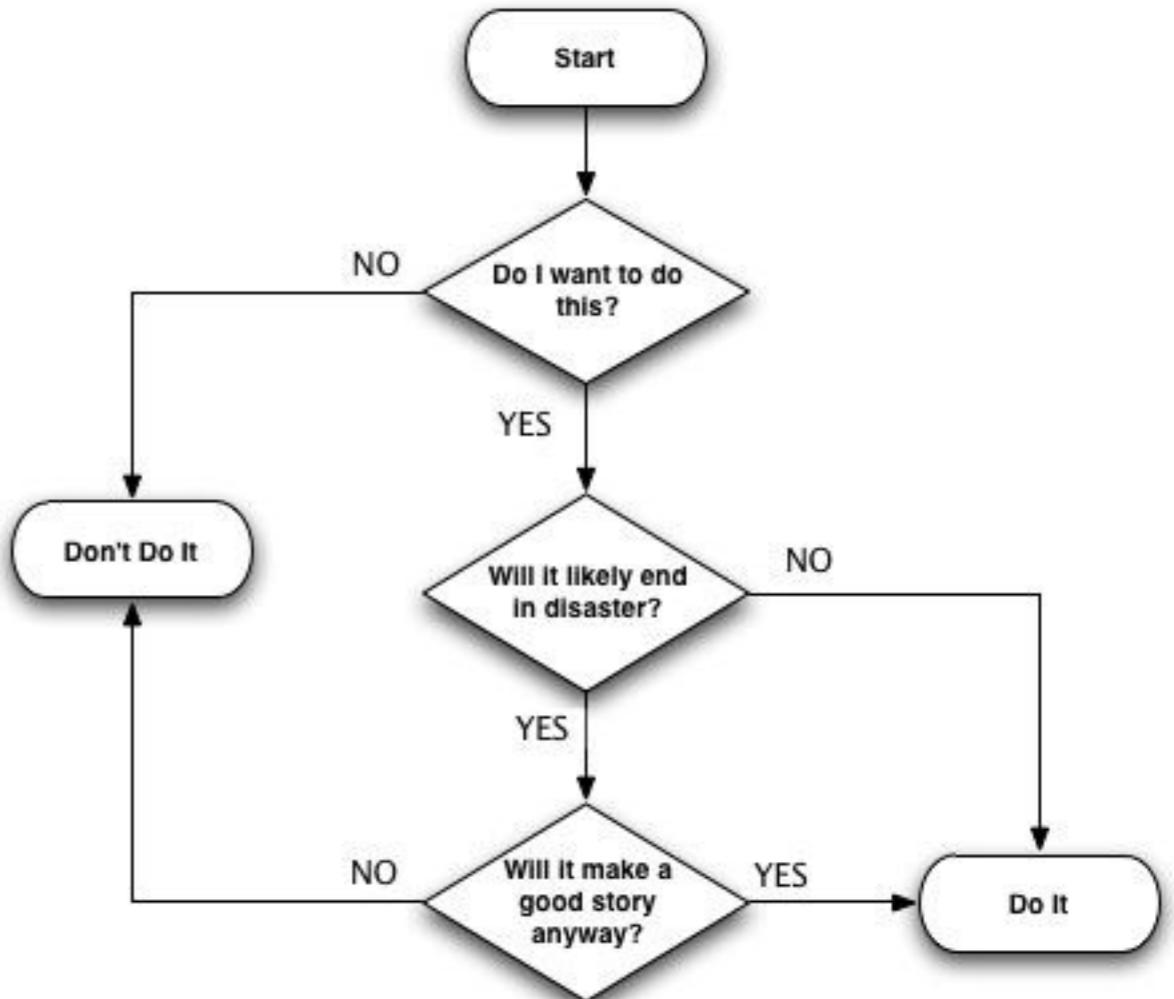
Other things you might want to do

- Be like Tom
- Ecology consultancy
- Government
- Non-profit
- etc.



What's the point of data science?

- Support decision making using data to make more money
- Assist internal decisions for the company
 - Ex: UI changes
- Help customers make decisions
 - Ex: Netflix recommendations
- Part of automated decision system
 - Ex: Stock trading algos



Types of data science work

- **Analytics:** cmp “business intelligence specialist”
- **Product development:** cmp “software engineer”
- **Tool/systems development:** cmp “ML/data engineer”
- **General R&D:** cmp “researcher”



What do I do in a week?

- Explore, massage data: 20%
 - Software creation, maintenance: 20%
 - Statistical analysis, modeling, ML/DL: 10%
 - Understand business problems: 10%
 - Prepare for stakeholder meetings: 10%
 - Coordination overhead (meetings, etc.): 20%-100%
 - Team bonding: 10%
 - Learning, skills development: 0%-5%
-
- The diagram illustrates the distribution of time spent on various activities. It features four main groups of tasks, each enclosed in a curly brace and labeled with its purpose and percentage contribution to the total week:
- Technical work: 50%** (Explore, massage data; Software creation, maintenance; Statistical analysis, modeling, ML/DL)
 - Business-specific: 20%** (Understand business problems; Prepare for stakeholder meetings)
 - Working in a team: 30%** (Coordination overhead; Team bonding)
 - Self-indulgence: almost zero** (Learning, skills development)

Academia vs industry

- Long cycle times (years)
- Novelty/rigor rewarded
- Important to find new problems
- Centered on individual
- More freedom
- Tenure
- Do it *now* (hours-months)
- \$\$\$ creation/savings rewarded
- Important to remove existing pain points
- Centered on team, company
- More money
- Large job market

Small vs medium vs large companies

Small (<500)

- Fast pace, high chaos
- Many hats, learn fast
- Pay low, upside high

Medium (<5000)

- Sometimes fast, often slow
- Potentially many areas
- Pay highly variable

Large (>5000)

- Slow pace, predictable
- Deep focus on one thing
- Pay can be *very* high

Interviewing for DS jobs

You will likely be assessed on

- Programming: Python, R, SQL, data manipulation
- Machine/deep learning: algo how, why, pro/con/pitfalls
- Statistics: probability, distributions, hypothesis testing
- Previous experience: your stories
- Soft skills: playing well with others, independent thought, company-specific traits, “cultural fit”

Getting a foot in the door

- Network inbound
- Network outbound
- Recruiting inbound (LinkedIn, etc.)
- Placement agencies, bootcamps
- Cold calls/e-mails/applications

BigCo vs SmallCo

Negotiating compensation



Questions?