Careers in Data Science

Yan Kou | Insight Data Science

Baylor Medical College | 25 June 2018
THE GRANT CYCLE

HOW IT'S SUPPOSED TO WORK:

WRITE GRANT → GET $ → DO RESEARCH → PUBLISH RESULTS

(REPEAT)

HOW IT REALLY WORKS:

DO RESEARCH → GET RESULTS BUT DON'T PUBLISH THEM YET. CALL THEM "PRELIMINARY RESULTS" → WRITE GRANT TO DO WHAT YOU ALREADY DID → GET $ → OK, NOW YOU CAN PUBLISH RESULTS

USE $ TO PAY FOR AN UNRELATED NEW PROJECT

WWW.PHDCOMICS.COM
A 7-week Fellowship program designed to bridge the gap between academics and Data Science
There are LOTS of opportunities out there for you.

However, career transitions are hard.

Often, the hardest part is identifying what it is you even want to do.
Start by asking lots of questions.
What is data science and where do I fit in?

Which industries am I interested in?

What companies/institutes are at the intersection of those two things?

How can I best prepare?
“Data Science” is a very broad term.

- Analytics
- Data Products
- Machine Learning
- Engineering
“Data Science” is a very broad term.

Analytics, experimentation, visualization, and data stories

“Data Scientists” job trend from indeed.com
“Data Science” is a very broad term.
“Data Science” is a very broad term.
“Data Science” is a very broad term.
"Data Science" is a very broad term.
“Data Science” is a very broad term.
"Data Science" cuts across an enormous range of industries

**Health**
- Memorial Sloan Kettering
- Biogen, ZocDoc, CVS

**Media**
- The New York Times
- MTV, Netflix
- Nielsen

**Social**
- Facebook
- Twitter
- LinkedIn

**Advertising**
- Sailthru
- Tapad
- Dstillery

**Energy**
- Bright Power
- EnergyHub

**Travel**
- AirBnB
- TripAdvisor

**Retail**
- Etsy
- Macy’s
- Birchbox

**Apparel**
- Rent the Runway
- StitchFix
- Bonobos

**Finance**
- Bloomberg
- Capital One Labs
- AmEx

**Government/Non-Profit**
- The City of NYC, The Census Bureau
- DataKind
Who are Data Scientists?

Job of an academic scientist:

• Collect and clean data

• Use programming and statistics knowledge to discriminate between signal and noise

• Convey results to the scientific community
Who are Data Scientists?

Job of a data scientist in industry:

• Collect and clean data

• Use programming and statistics knowledge to discriminate between signal and noise

• Convey results to the team/company/investors

• Make data-informed decisions that directly impact the product and ultimately the business
Who are Data Scientists?

PhDs in quantitative fields:

• Have data intuition
• Know the right questions to ask
• Have statistics knowledge
• Have strong coding skills
• Communicate clearly
Tips and Tactics

**Connect:** Join LinkedIn. Read Twitter. Subscribe to DS weekly. Go to Meet-ups or watch online. Find out if nearby universities have a center for data science. Look into civic data groups: DataKind, Code for America. Where are alumni from your group now? Talk with them!

**Talk to people:** Learn terminology, ask about what tools people use at places you’re interested in working (what should you learn and demonstrate?), practice describing your research. Practice interviewing.

**Build evidence:** Do small non-academic projects to show your interest, post code on github, write a blog post. Write a one-page resume. Show it to people outside your field — can they understand what you did?

**Apply for Insight!** We help with all of these things. :)
Why Insight?

• **Connection**: meet 20-30 hiring managers in 4 weeks

• **Community**: 1600+ alumni at 400+ companies

• **Continued career support**: career growth and more

• **It's fun!**
Start now, don’t apologize.
What you know

What they know
What you know  What they know
To learn more about Insight, visit www.insightdatascience.com

PROGRAMS

DATA SCIENCE FOR PHDS

DATA ENGINEERING FOR ENGINEERS

HEALTH DATA SCIENCE FOR GOOD MINDS (AND PHDS)

ARTIFICIAL INTELLIGENCE FOR SCIENTISTS AND ENGINEERS

LOCATIONS

SILICON VALLEY

NEW YORK

BOSTON

SEATTLE

REMOTE

TORONTO
Thank you for your time!

yan@insightdatascience.com

@insightdata

Upcoming  session: September 2018
Upcoming  deadline: July 9th, 2018
Challenges academics face when making the transition to Industry:

**Tools**
- ROOT
- IDL
- MATLAB
- PostgreSQL
- Python
- Spark

**Pace**
- Results published once a year
- Analysis for scientific community
- Results needed weekly, MVP
- Results need to impact bottom line

**Product**
- Tools
- Pace
- Product
- Vocabulary
Thinking about DS?

How do you prepare for a career in data science?

Undergrads / early graduate students:

• Great time to start thinking about using open source tools
  • eg: build software tool for lab in python
  • Start using GitHub to store your projects
  • Organize public outreach events, practice explaining your research to non-experts
Senior graduate students:

• Build a data product as a side project
  • Shows off your initiative, technical ability, product sense
  • Again…use industry standard tools like python and SQL!
  • Working with distributed computing already? Become familiar with frameworks like Apache Spark, Hadoop, Flink…etc
• Start a blog
How to find out more

• Go to data science meetup groups and conferences

• Sign up for newsletters, read blogs
  • eg: Data Science Weekly, Hacker News, etc

blog.insightdatascience.com
Data Engineers are specialized software engineers that enable others to answer questions on datasets within latency constraints

-Nathan Marz
Inventor of Apache Storm and the Lambda Architecture
Author of “Big Data” and Insight Advisor
E.g. No one wants to wait 1 minute to get a response for a Google search: “ASU Sun Devils game”
Collect

Process

Access

Efficiently
Insight Fellows are Data Scientists and Data Engineers at:

facebook  LinkedIn  okcupid  PREMISEx  Gartner  Bloomberg  ACTIVISION  SQUARESPACE
BLACKROCK  BIRCHBOX  greenhouse  Square  CapitalOne  reddit
NBC  STITCH FIX  airbnb  Twitter  UBER  ZocDoc
YAHOO!  VECTRA  twitch  AVANT  amazon.com  McKesson
Pinterest  JAWBONE  Oscar  Microsoft  News Corp  mtv
JPMorgan  salesforceIQ  DOWJONES  Palantir  SAMBA TV  AXON VIBE  INTUIT  KHANACADEMY

Silicon Valley ● New York ● Boston ● Los Angeles ● Seattle + many others…
Who are Insight Fellows?

They come from many different backgrounds!

**INSIGHT BY THE NUMBERS**

The Alumni Split

- **42%** Physical Sciences
- **25%** Biosciences
- **10%** Engineering
- **9%** Mathematics
- **7%** Social Sciences
- **3%** Earth Sciences and Energy
- **3%** Computer Science
- **1%** Other
INSIGHT IN A NUTSHELL

- INTRO WEEK
  - WEEK 1
- DATA PROJECT
  - WEEKS 2-4
- COMPANY VISITS
  - WEEKS 5-7
- JOB INTERVIEWS
Data Science: Silicon Valley, New York City and Remote January, May, September each year!

Data Engineering: Silicon Valley & New York City January, May, September each year!

Health Data Science: Silicon Valley & Boston January, May, September each year!

Artificial Intelligence: Silicon Valley in March 2017 New York in July 2017

7 Weeks $0 Full time need based scholarship
Top Industry Mentors

DJ Patil
Chief Data Scientist, White House
Co-coiner of the term Data Scientist

Hilary Mason
Founder & CEO of Fast Forward Labs

Drew Conway
Founder & CEO of Alluvium
Author of Conway diagram

Monica Rogati
Data Science Advisor
One of the earliest Data Scientists at LinkedIn
Top Industry Mentors

Nathan Marz
Inventor of Apache Storm & Lambda Architecture

Lucian Lita
Head of Data, Intuit

Brian Sullivan
Director of Streaming Analytics, Netflix

Jeff Magnusson
Director of Data Platform, Stitchfix