# Insight Data Science: Bridging the gap between academia and industry

Josiah Walton

Physics Careers Seminar

UIUC Department of Physics

April 23, 2015

# Outline

Personal Timeline

Data Science

Insight Program

Internships

Tips and Tactics

# Physics to Data Science

ARKANSAS RAZORBACKS

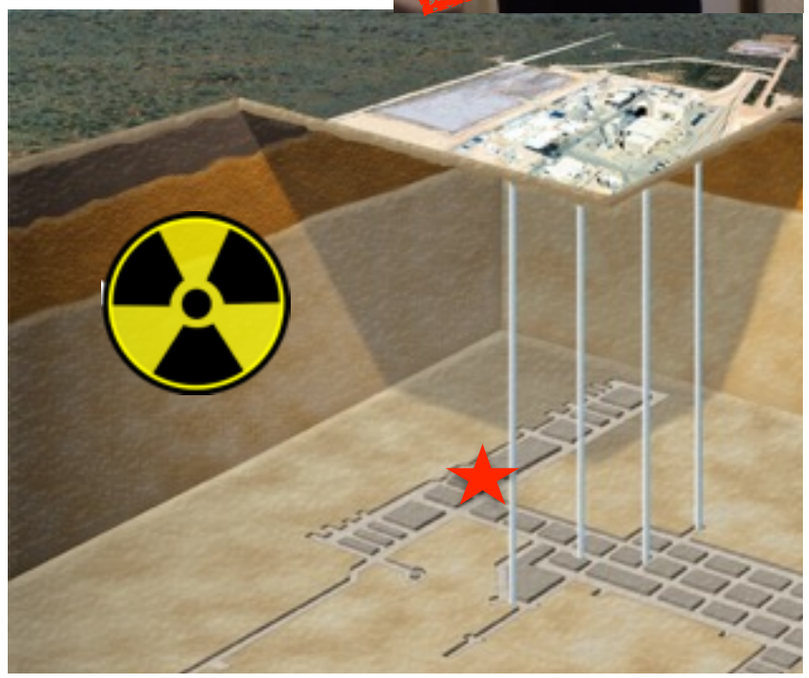
CARNEGIE INSTITUTION FOR SCIENCE

NOAO

THE DARK ENERGY SURVEY

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Enriched Xenon Observatory
EXO
for double beta decay



Fermilab

Discovery CHANNEL

INSIGHT

# Why data science?

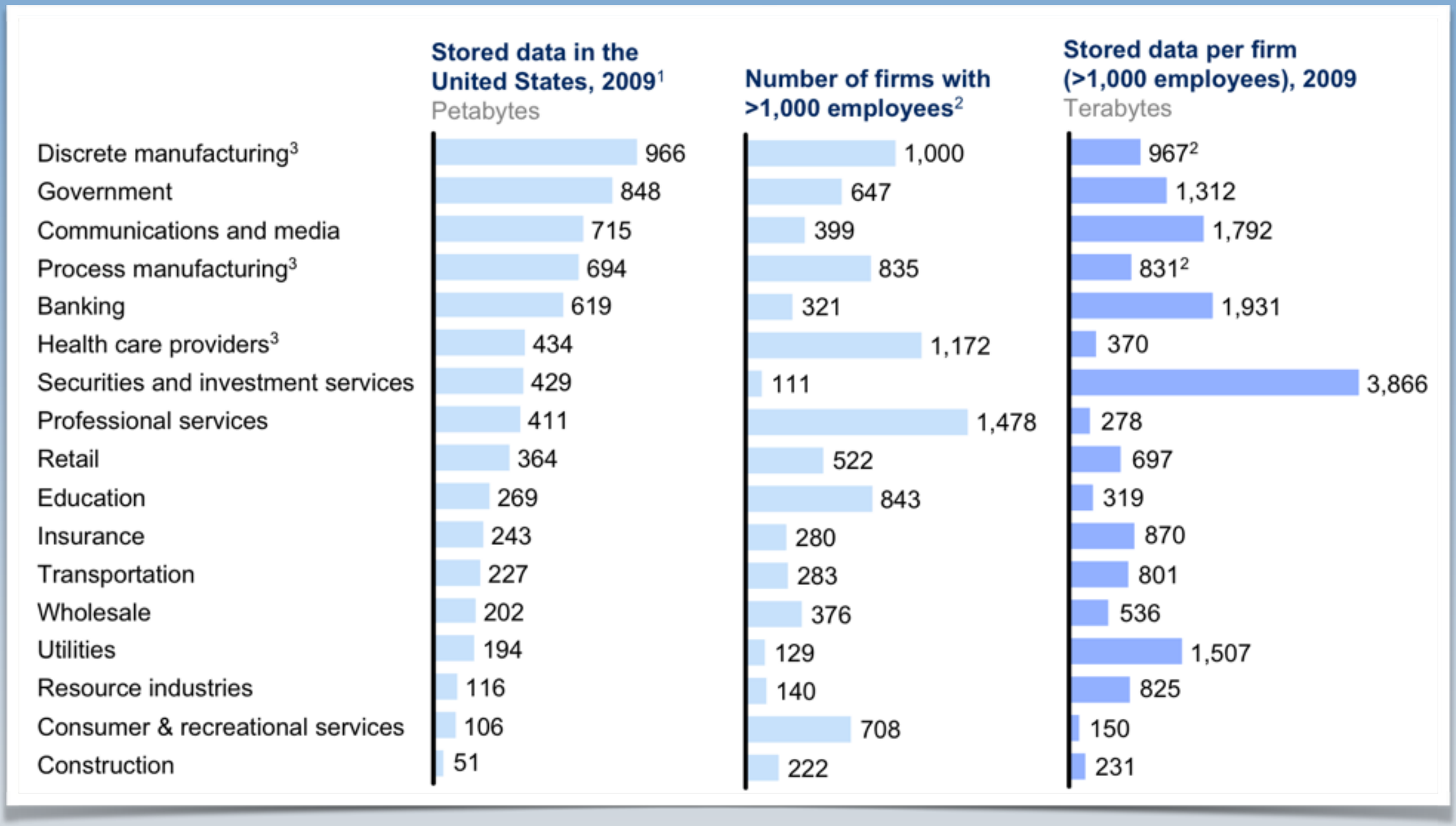
- Chance to positively impact people's lives
- Variety: learn many new skills and areas
- Pace of work
- Location
- Compensation
- Many opportunities

# Update: I've accepted a data scientist position at Biogen Idec

# The Rise of Data Science

# Big Data

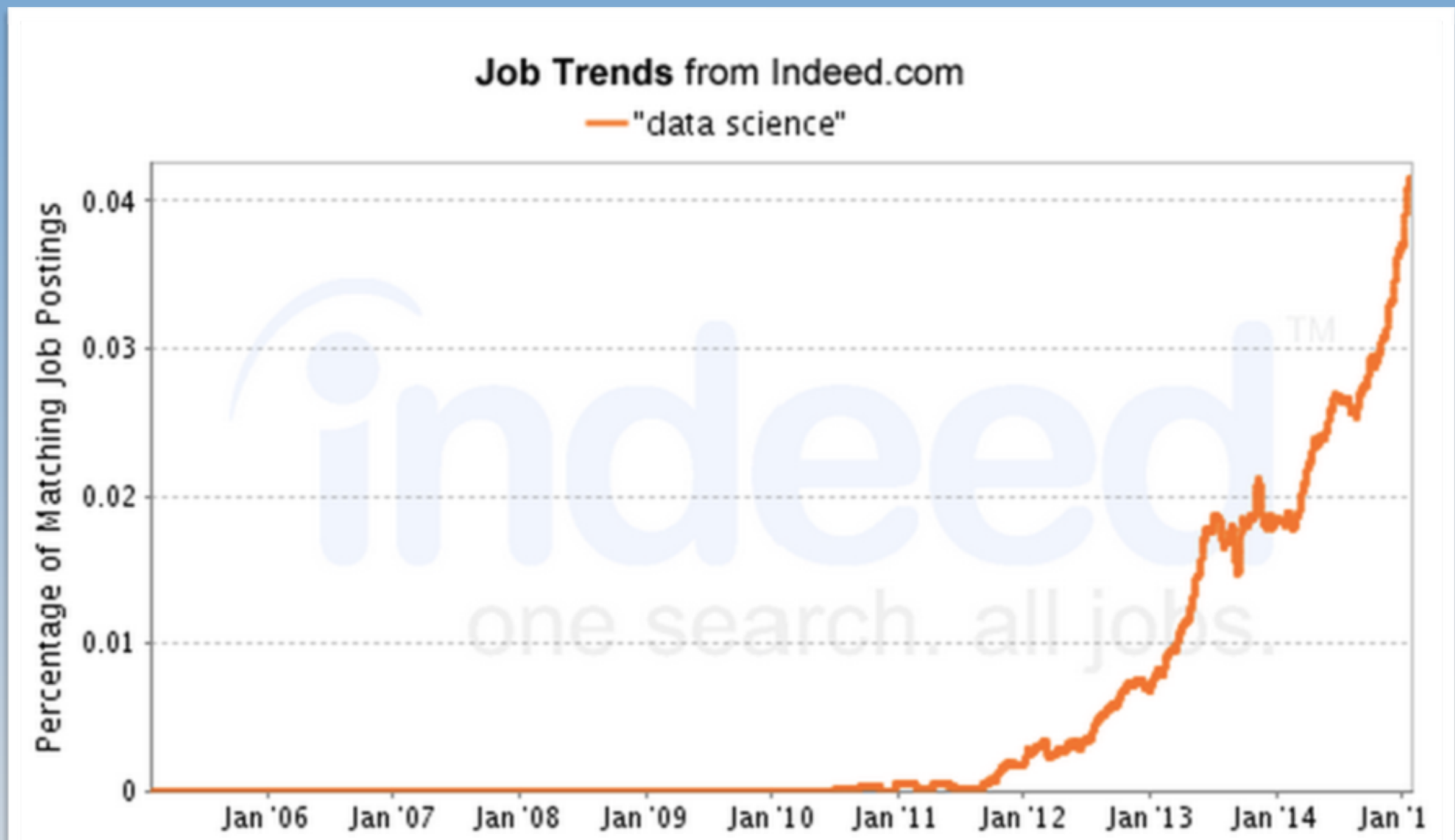
# Big Data




SOURCE: IDC; US Bureau of Labor Statistics; McKinsey Global Institute analysis 2011

# Estimated demand by 2018

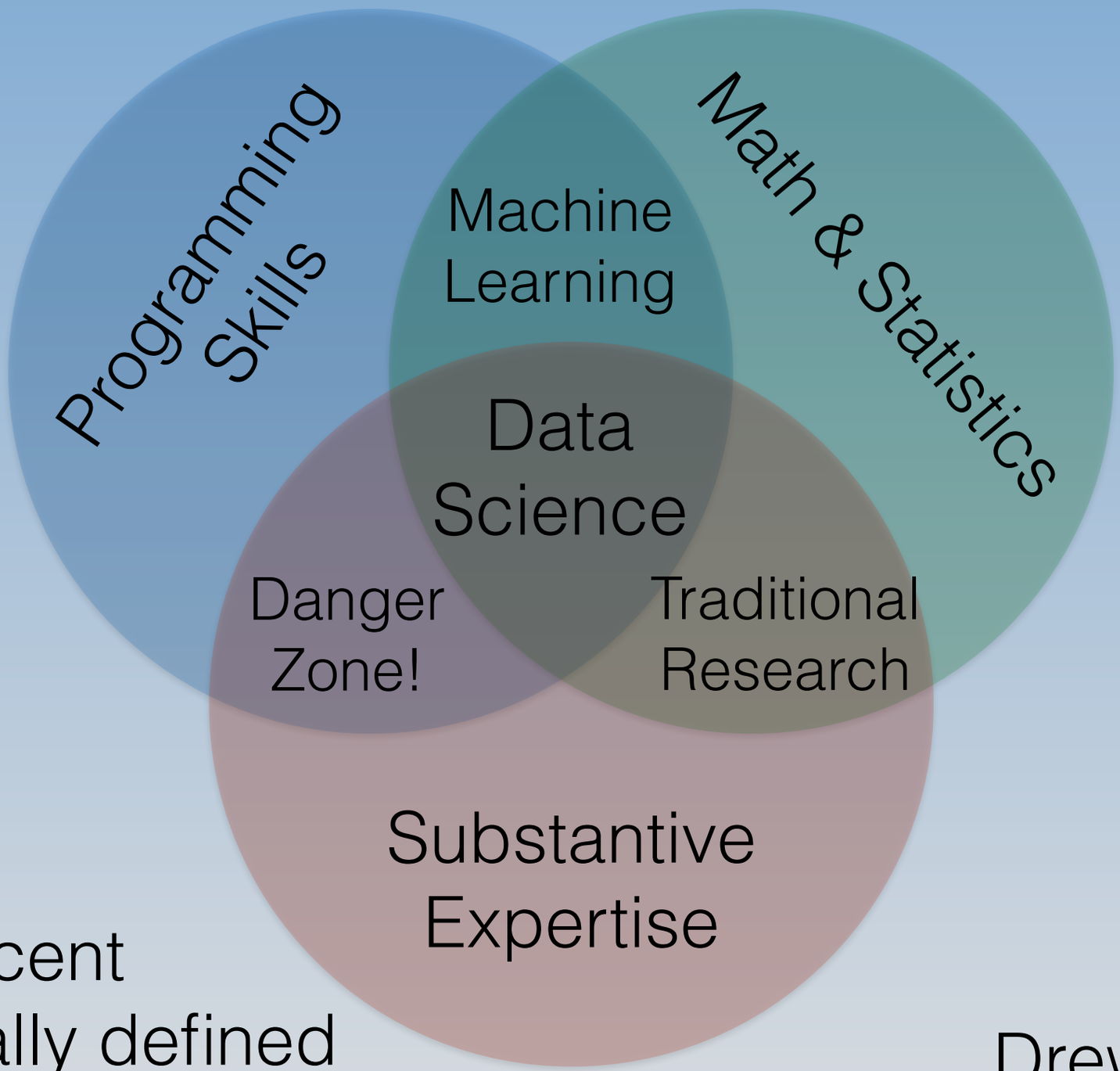
SOURCE: IDC; US Bureau of Labor Statistics; McKinsey Global Institute analysis 2011


Job Trends from Indeed.com
"data science"
Percentage of Matching Job Postings
0.04
0.03
0.02
0.01
0
Jan '06
Jan '07
Jan '08
Jan '09
Jan '10
Jan '11
Jan '12
Jan '13
Jan '14
Jan '1
indeed
one search. all jobs.

# What is Data Science?

"Nobody knows what it means, but it's provocative"

# What is Data Science?




Field still nascent
Role continually defined
Varies from company to company

Drew Conway

# What do data scientists do?

Leverage critical thinking and analytical expertise to ask piercing questions

Extract “actionable insights” with real business value from complex data sets

# Many Questions in Data Science

- **Khan Academy:** Millions of math problems are attempted online each day.
  - When is someone proficient?
  - What is the best problem order?
- **The New York Times and Wall Street Journal:**
  - What is the right payment model for journalism?
  - How do we prevent customer churn?
- **LinkedIn:** Does this LinkedIn account correspond to a real person?
- **Netflix:** What new shows should we make?
- **Everywhere:** What are some efficient ways in which the data can be stored, accessed, and streamed?

# Where do data scientists come from?

“The best data scientists tend to be ‘hard scientists’, rather than computer science majors.”

- DJ Patil

- Chief Data Scientist at U.S. OSTP
- Former head of data science at LinkedIn
- Insight mentor
- PhD mathematician
- Coined “data science”

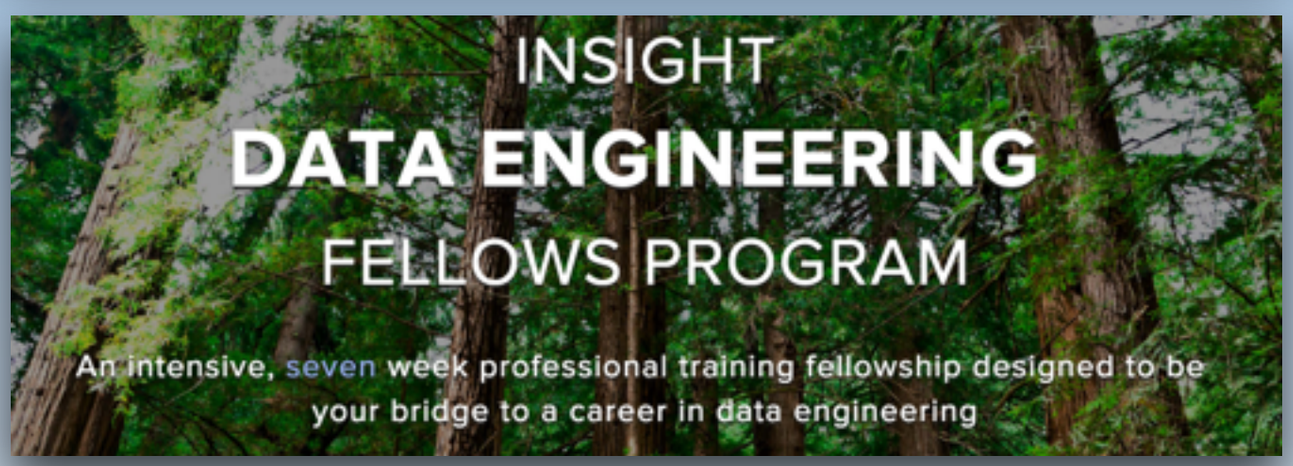
INSIGHT

Founded 2012

NYC

Palo Alto, CA

insightdatascience.com




Palo Alto, CA

insightdataengineering.com

Boston, MA

insighthealthdata.com

# Applying to Insight

- **Initial Stage** - Formal application:
  - CV
  - Experience with programming, machine learning, etc.
  - Why you'd like to participate
- **Final stage:** Interview
- **Competitive:** ~5% acceptance (growing though!)

Upcoming Sessions

Data Science
SILICON VALLEY & NYC
Sept. 2015 (Deadline: June 29, 2015)

Data Engineering
SILICON VALLEY
Sept. 2015 (Deadline: June 29, 2015)

Health Data Science
BOSTON
July 2015 (Deadline: May 4, 2015)

INSIGHT

# Program Timeline

## 7 Weeks

## 4-8 Weeks

Week 1 — Intro Week

Weeks 2-4 — Data Project

Weeks 5-7 — Company Visits

Job Interviews

# The Bridge

You
(~90% of skills)

Insight
(last ~10% of skills)

Data Science Career

1. Tools
2. Context
3. Pace
4. Priorities

Data
Project

Personal Portfolio

Machine Learning
New languages/tools
Short time scale
Concrete Deliverable
Acceptable Performance

# Outcomes

- Individual data product, effectively walks you into interviews
- Extensive job application and **mock interview** preparation
- Outstanding professional network (cohort and leading data scientists)
- Interest from **many** recruiters
- **100% job placement** (usually within 2 months)
- Starting salary:

$110,000 $120,000 $150,000

200+
**Insight Fellows**
working as
**Data Scientists**
and
**Data Engineers**

# Insight Companies

NOVARTIS

Memorial Sloan-Kettering Cancer Center
*The Best Cancer Care. Anywhere.*

NBC

NETFLIX

in

Microsoft

airbnb

Palantir

Square

MUSIC TELEVISION

KHAN ACADEMY

yelp

Spotify

The New York Times

+Many Startups…

# Doing an internship while in grad school

# Difficult

- Loss of research time
- Potentially delayed graduation
- Incurred debt

# Rewarding

- Experience
- Networking

# Important Questions

- Applicability of new skills?
- Teaching/research responsibilities?
- Pay?
- Work while away?
- Time spent away?
- Current lease/renting situation?

Ideally light at end of the tunnel

# It’s your life…

# Don’t Fret!

You don’t have to get into an internship like Insight to find an awesome career in data science!

# What you do possess are a very particular set of skills…

- Analytical mindset
- Innate curiosity and passion
- Communication skills
- Team work mentality
- Adaptability
- Technical expertise (math, stats, complex/messy data. etc.)
- …

# Tips and Tactics

- Start early!
- Learn open source software (python, R, etc.)
- Build knowledge of machine learning, statistics, comp sci fundamentals, etc., using library, free blogs, coursera, youtube, etc…
- Accumulate evidence of your skills (learn by doing!):
  - Develop real projects
  - Post work on Github
  - Write personal blog
- Connect (network!): **LinkedIn**, Twitter, meetups
- Informational interviews: Vocabulary, substantive expertise, tools
- Learn to sell yourself and accomplishments!

# Data Science Job Interviews

- Interview prep: practice, practice, practice (mock interviews)

- **Be confident and charismatic!** Hiring process is much about human nature as intellectual ability

- Typical interview route (assuming your resume makes it):

  1. Phone screen (HR and or technical questions)

  *2. Data challenge*

  3. On-site presentation and meetings

# Ending thoughts

- Get in touch with me if you're interested in Insight or want to talk about data science in general!
- Connect with me on LinkedIn!
- Thanks for listening!