

# Transitioning to Data Related Careers in Healthcare

July 27, 2020

Ravi Patel  
Program Director, Health Data Science  
Insight

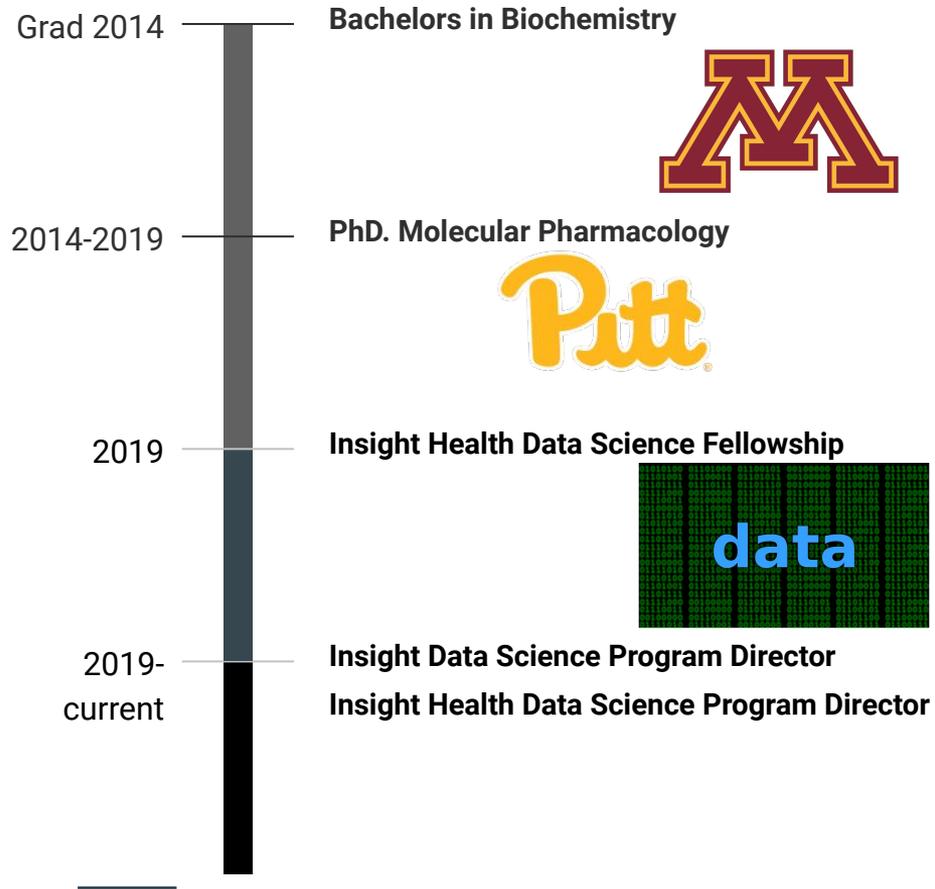




# Ravi Patel

Program Director  
Health Data Science

[linkedin.com/ravikpatel-phd/](https://www.linkedin.com/in/ravikpatel-phd/)  
[ravi@insightdata.com](mailto:ravi@insightdata.com)



# Today we will cover

- Why are careers in data important?
- Health industry trends
- What does a data related career in healthcare look like?
  - Data Scientist at a Healthcare company
  - Computational Biologist
- What does it take to get a job in this field?
- Insight Data Science and Transitions
- Q&A

# Why are careers in data important?

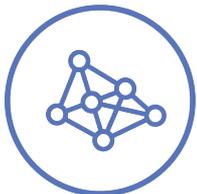
# Why are these careers booming?

- Information Explosion & Big Data
  - People were ready to share information
  - We had means of sharing the information
- Data Driven Decisions
  - Not taking decisions based on “gut”.
  - Better understanding of users.
  - Personalized services: Recommendations.
- Technology Advancements
  - NGS
  - EHR
  - Drug discovery platforms
  - Digital health

What can you work on?

# Different Roles in the data space

## Data Scientist / Data Analyst



- Product Analytics
- Understanding Business
- Understanding Data

## Data Engineer



- Data Pipelining
- Distributed Storage
- Scalability

## Data Product Manager



- Market analysis
- Future of Data products
- Business growth

## ML Engineer / AI Professional



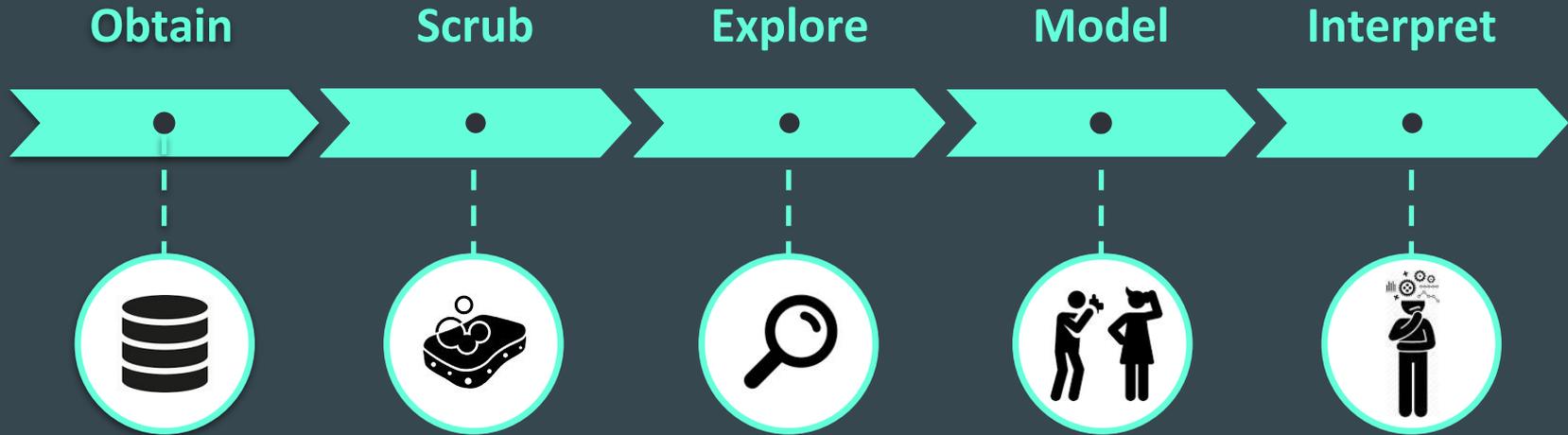
- Designing ML Algorithms
- Productionalizing ML models

## Data Infrastructure Engineer / DevOps



- Designing infrastructure
- Automated deployments
- Handling Dev pipelines

# What do data scientists do?



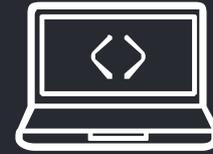
# There are many unique challenges to analyzing *health data*



HIPAA + PHI

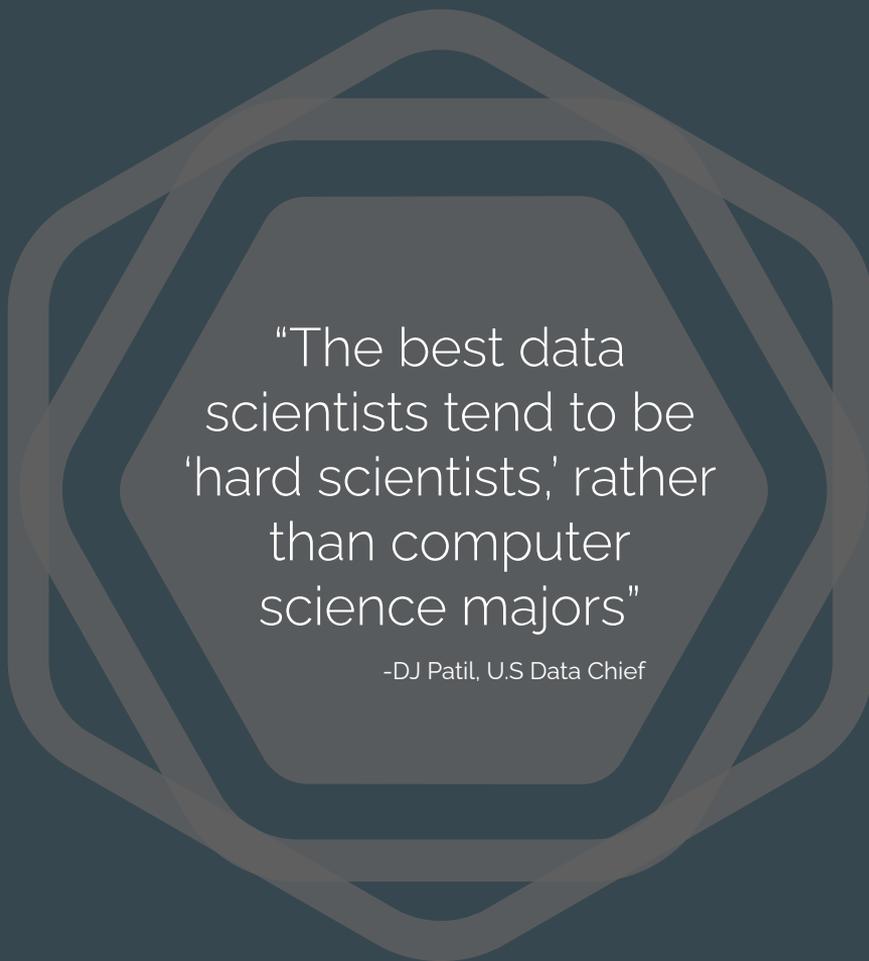


Socioeconomic  
inequalities



Health tech  
literacy

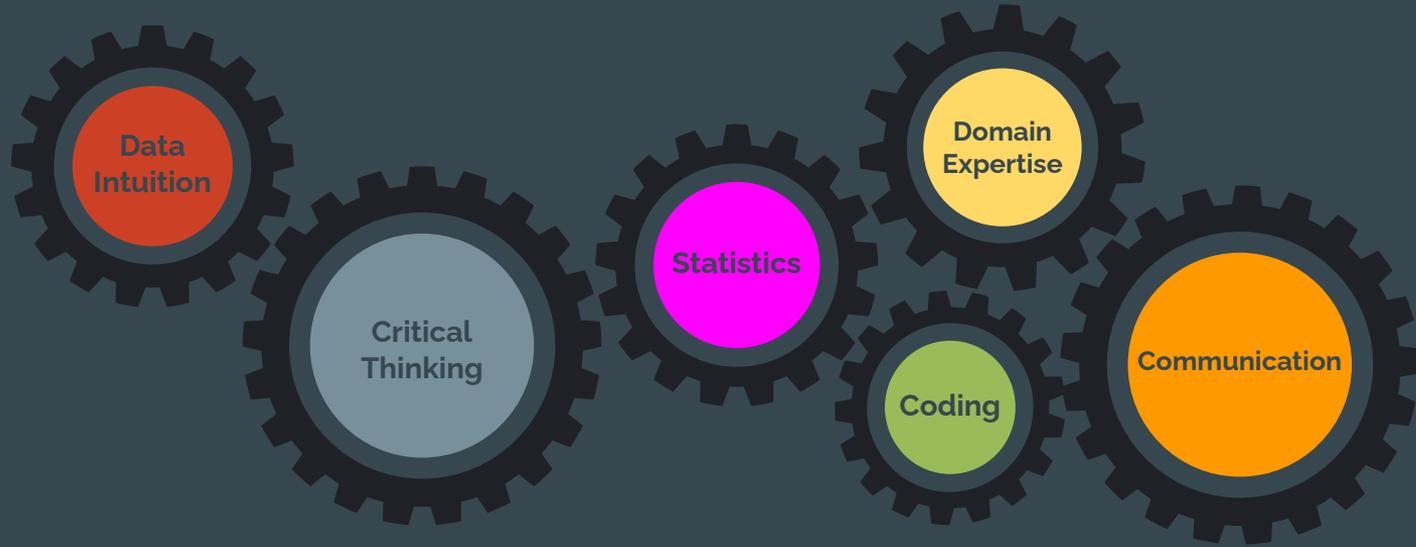
Health  
researchers are  
primed to address  
these critical data  
challenges



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

-DJ Patil, U.S. Data Chief

# Health researchers make great data scientists



What does health  
data science look like  
in practice?

# The health and biotech sectors are growing and generating huge amounts of data

-  Health Management
-  Therapeutics
-  Consumer Genomics
-  Health Insurance
-  Digital Health
-  Biotechnology





**15**  
THERAPIES



**MULTIPLE OPTIONS**

**Varied mechanisms,**  
efficacy, side effects and  
tolerability

**TOP 5**  
DRUG SPEND



**HIGH COST**

**\$60-80,000** per year,  
**over \$14B** spend, with  
unguided use

**950K**  
PATIENTS



**STRIKES EARLY**

**Late 20's average**  
**onset,** affects 3:1 women  
to men

**\$2-4M**  
LIFETIME COST



**CHRONIC DISEASE**

**Long term costs** include  
medical, drug, ancillaries  
and disability

**50%**  
DISABILITY



**PROGRESSIVE**

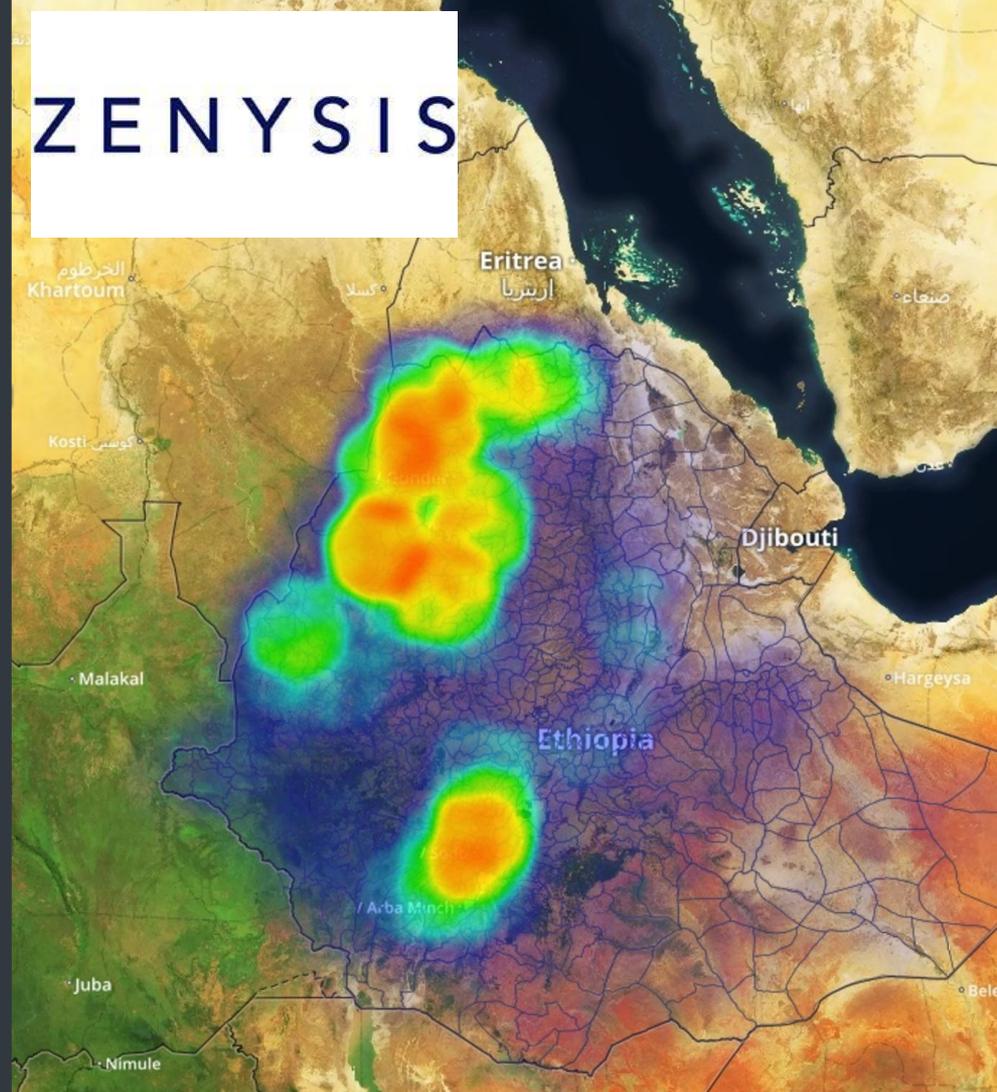
Symptom burden and  
co-morbidities  
**impact employment**

# Digital Therapeutics

- **Focus:** Chronic diseases
- **Product:** Digital prescription apps + tests
- **Data:** Survey data + text entries + app engagement + Clinical tests
- **Projects:**
  - User adherence to prescriptions
  - Classification of depression severity
  - Signal processing used for screening patients and monitoring disease

# Public Health Management

- **Focus:** Developing countries and non-profit organizations
- **Product:** Develop software platforms for harnessing PHI
- **Data:** Longitudinal PHI
- **Projects:**
  - Anomaly detection for malaria outbreaks
  - Time series analyses for resource allocations
  - Outbreak investigations





## WOMEN'S HEALTH



**100X fewer  
false positives**

The false positive rate of noninvasive prenatal screening (NIPS) is 100x lower than serum screening in patients <35 years old.<sup>3</sup>



**86% of affected  
pregnancies no longer  
missed**

86% of affected pregnancies detected by expanded carrier screening (ECS) are missed when screening for cystic fibrosis and spinal muscular atrophy alone.<sup>1</sup>



**50% of hereditary  
cancers are non  
BRCA1/2**

At least 50% of hereditary cancers are caused by genes other than BRCA1 and BRCA2.<sup>4</sup>



**76% of couples  
took action**

76% of couples found by ECS to be at risk for severe or profound conditions pursued alternative reproductive actions such as prenatal diagnosis or IVF with preimplantation genetic diagnosis.<sup>2</sup>

# Diagnostics/Therapeutics

- **Focus:** Molecular biology techniques to diagnose a disease
- **Product:** Diagnostic test and interpretation of result
- **Data:** NGS + Clinical Assay + Clinical tests
- **Projects:**
  - Diagnose a genetic disease
  - Work with high dimensional data
  - Build pipelines to support new product features

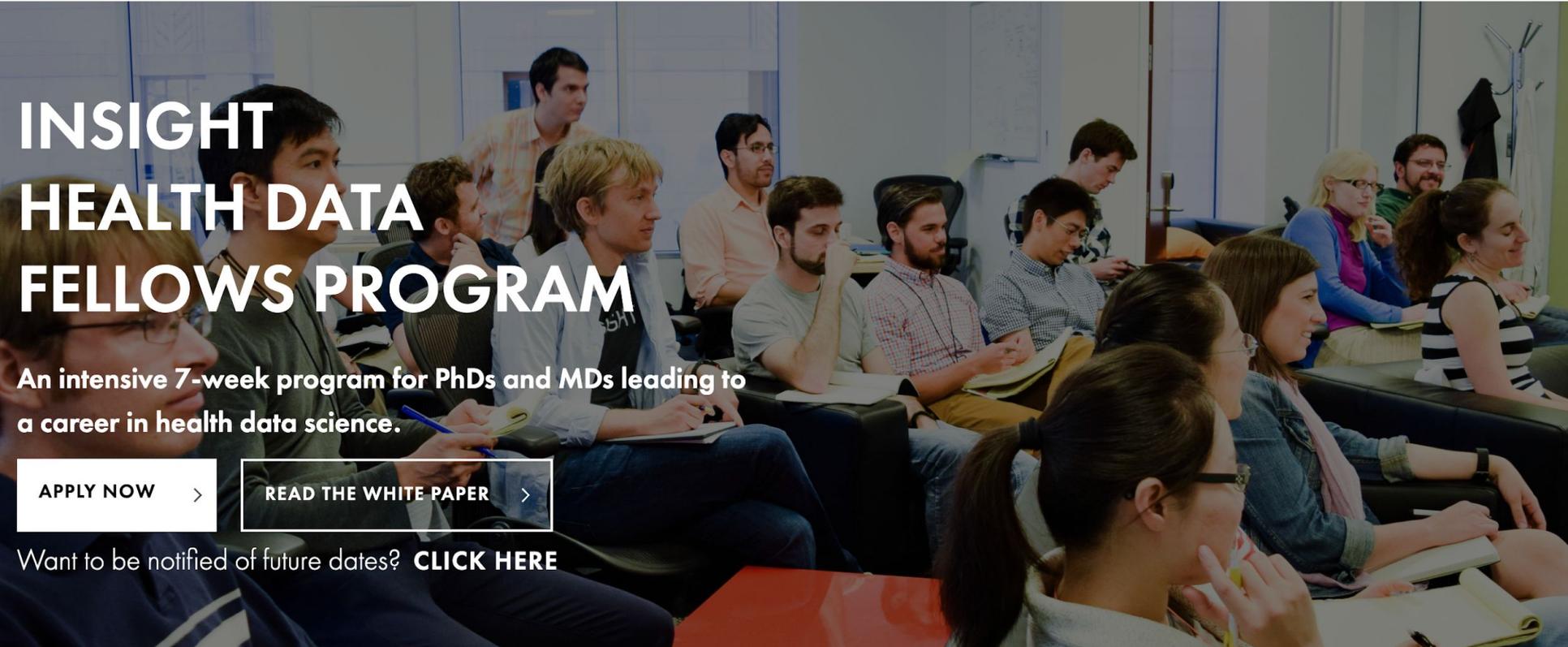
# AI and Healthcare

- **Focus:** Using deep learning to solve challenging healthcare problems
- **Product:** Develop software
- **Data:** Drug Screening + Clinical Data + Medical imaging
- **Projects:**
  - Develop an in silico drug screening method
  - Diagnose disease
  - Work with medical imaging



**BlackThorn**<sup>TM</sup>  
THERAPEUTICS





# INSIGHT HEALTH DATA FELLOWS PROGRAM

An intensive 7-week program for PhDs and MDs leading to a career in health data science.

[APPLY NOW](#) >

[READ THE WHITE PAPER](#) >

Want to be notified of future dates? [CLICK HERE](#)

Start by  
asking  
lots of  
questions

What is data science,  
computational biology 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?

How do you get these jobs?

# 1. Knowledge/Skills

- Understanding Concepts related to each role
- Knowing the fundamentals
  - Stats, Computer Science, Math, Programming, Analytical Thinking, domain knowledge
- Knowing the tools
  - Pandas, Numpy, Scikit learn
- Understand how to apply the knowledge to industry problems
  - Reading blogs
  - Attending meetups
  - Knowing to ask the right questions?
- Applying the relevant tools / techniques to relevant problems
  - Understanding trade-offs
  - Doing Projects

## 2. Build Evidence

- Your PhD project
  - Can you add a “DS-like” component
- Doing small non-academic projects
  - What is the industry tackling?
  - Can you do something unrelated to your research
- Post your code on Github
- Write a blog post
- Can people outside your field understand what you did?

# 3. Network

- Know the right people to talk to
  - Hiring Managers
  - Technical Recruiters
- Know how to pitch yourself
  - Tailor a pitch for each role you apply to.
  - Use career fairs effectively, don't just drop in your resume.
- Meet the right people
  - Meetups, Hackathons, Fellowships/ Internships/ Bootcamps, Code for America, Insight
- Know to have technical conversations (Knowledge about the field will help here)

# Interview Process (Internship & Full-time)

- **Preparing a good resumé to get a foot in the door!**
- A quick phone screen to get you started
- Take home assignment
  - Timed Coding assignment (HackerRank, Leetcode)
  - Technical assessment over video conference or phone
  - Take home Data Challenges
- **Onsite**
  - 4-6 hours
  - Mix of Behavioural, Technical and data oriented.
  - Speak to multiple team members
  - Topics: Designing data platforms, Whiteboarding coding questions, Technical conversations

**Most importantly,  
Be open to rejection  
and  
Learn from your mistakes**

# Insight

## Transition to Careers in Data

# Summary of Insight

Full-time program

Scholarships available

Collaboration and mentoring

Interview Prep

Meet teams hiring data scientists, computational biologists, bioinformaticians, etc.



Join the community! Learn the domain!  
Build your network! Get a job!

Questions?  
Careers in Data, making the transition, or  
Insight?

[Ravi@InsightDataScience.com](mailto:Ravi@InsightDataScience.com)

INSIGHT