

# Careers in Data

February 25, 2020



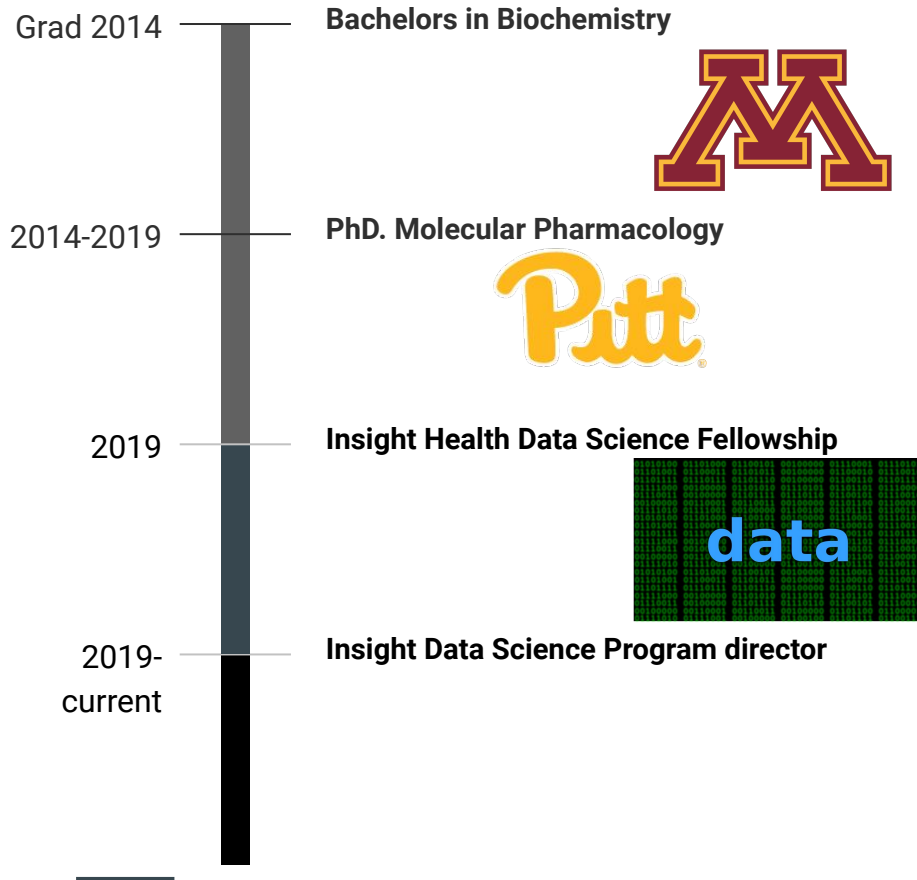
University of California  
San Francisco



# Ravi Patel

Program Director  
Health Data Science

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



# Today we will cover

- Why are careers in data important?
- What are the different roles and responsibilities?
- Case Study: Uber Eats: Providing personalized food recommendations
- What does it take to get a job in this field?
- Insight Data Science and Transitions

# Why are careers in data important?

# Why are these careers booming?

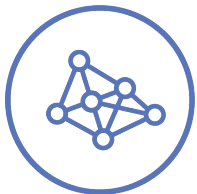
- Information Explosion & Big Data
  - People were ready to share information
  - We had means of sharing the information
- Data Driven Decisions
  - Not taking decisions based on “gut”.
  - Better understanding of users.
  - Personalized services: Recommendations.
- Technology Advancements
  - NGS
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

What can you work on?

INSIGHT

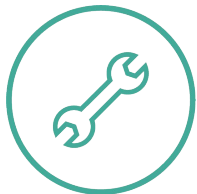
# Different Roles in the data space

## Data Scientist / Data Analyst



- Product Analytics
- Understanding Business
- Understanding Data

## Data Engineer



- Data Pipelining
- Distributed Storage
- Scalability

## Data Product Manager



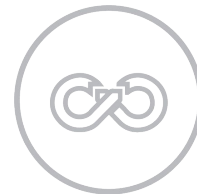
- Market analysis
- Future of Data products
- Business growth

## ML Engineer / AI Professional



- Designing ML Algorithms
- Productionalizing ML models

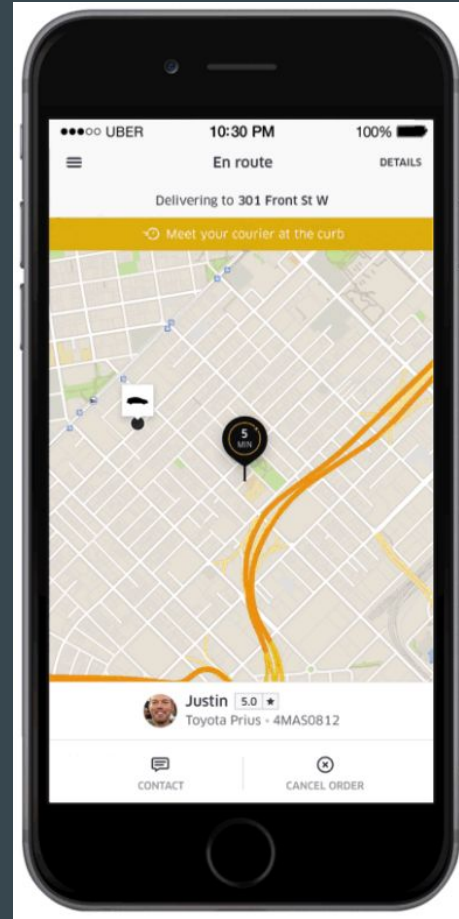
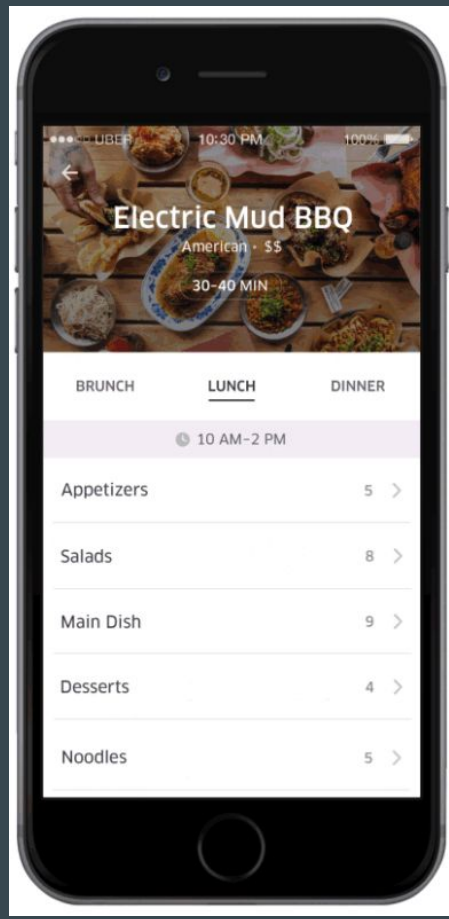
## Data Infrastructure Engineer / DevOps



- Designing infrastructure
- Automated deployments
- Handling Dev pipelines

# Case Study: Uber Eats



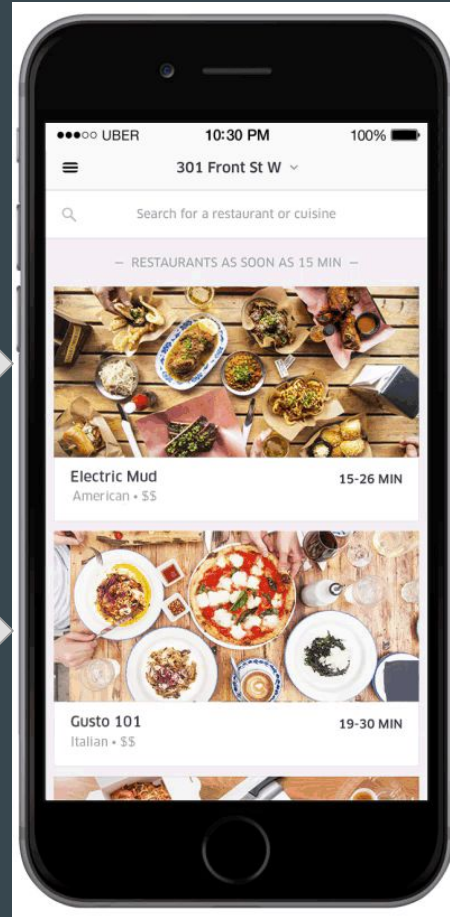


Copyrights Uber Eats and Wired

INSIGHT

Use case: Provide personalized food choices

Update the options that gets shown on the app based on user preferences.



# Use case: Provide personalized food choices

- Need to capture user specific data
  - Past orders by the user
  - Famous restaurants in user's location
  - User search history, clicks??
  - Nearby Orders
- Data Product managers:
  - Identify the need for the product.
  - Work with others to define the development lifecycle.
  - Define user-testing, AB testing frameworks for the product.
- Data Scientists:
  - Analyze the User data to find out what features are important.
  - Run experiments, build out models.
  - Retrain existing models with new data.
- Data Engineer:
  - Define a pipeline to get data to scientists/analysts.
  - Automate process for new data.
  - Build a scalable data pipeline.
- DevOps engineers:
  - Maintain infrastructure for data pipelines.
  - Define development best practices.
  - Build continuous integration pipelines to auto deploy code, test them.
  - Define integration tests for the code deployed.
- Machine Learning Engineer / AI professionals:
  - Build a deep learning model to get super accurate prediction of user behaviour
  - Monitor quality of recommendation

How do you get these jobs?

# 1. Knowledge

- Understanding Concepts related to each role
- Knowing the fundamentals
  - Stats
  - Computer Science
  - Math
  - Programming
  - Analytical Thinking
- Knowing the tools
  - Pandas, Numpy, Scikit learn

## 2. Skills

- What is the industry tackling
- 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

# 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
- 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)
  - Over the phone coding
  - 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 Data Science

## Transition to Careers in Data

Smart  
individuals  
with the right  
skills



High impact  
Teams



7 Weeks

\$0

Full time

Need based  
scholarship

Professional  
environment

INSIGHT

# The Fellowship



Knowledge

Skills

Network

# Meeting Data Teams from mentor companies



INSIGHT

# Collaboration



INSIGHT

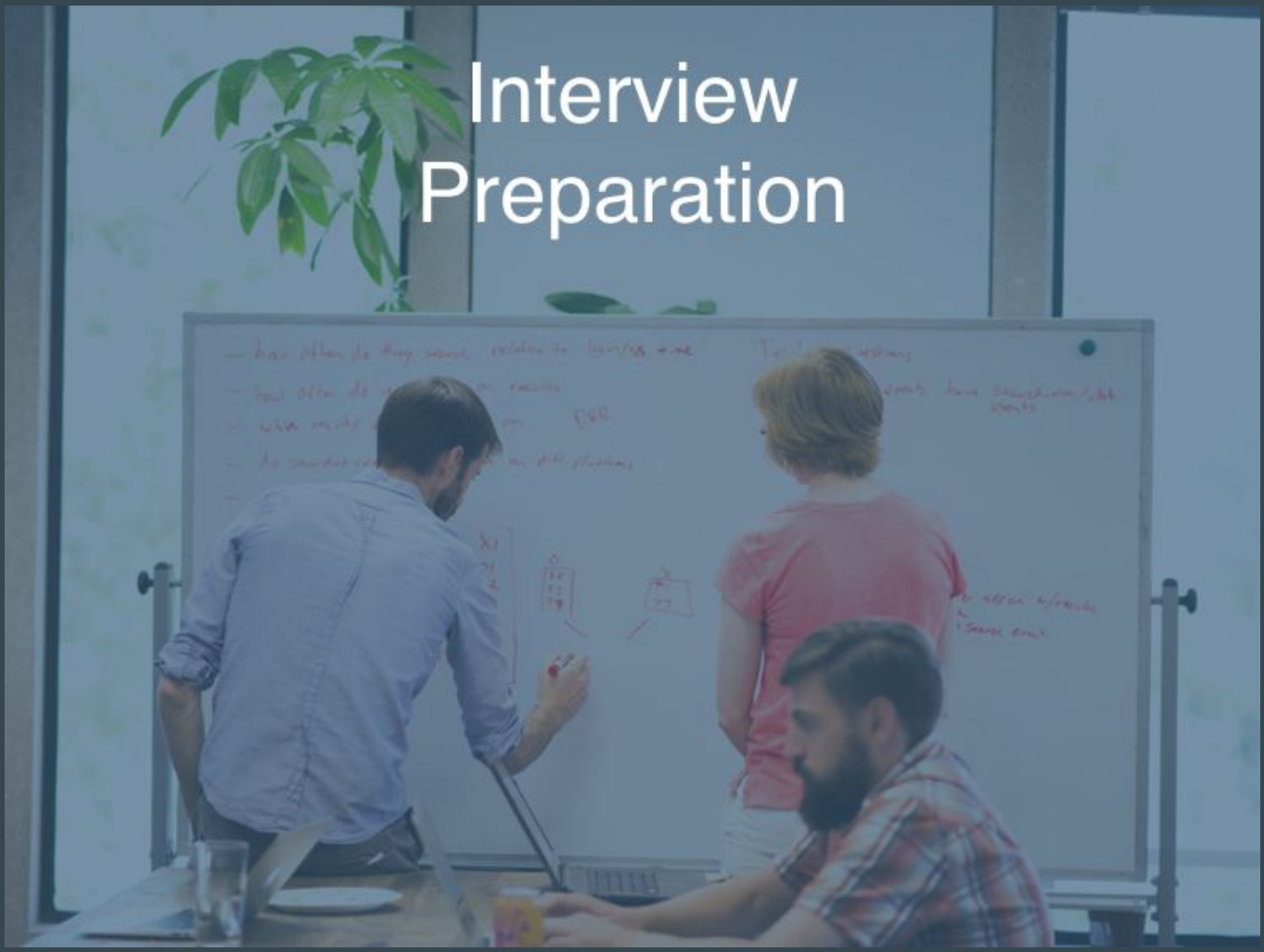


# Alumni Mentoring



INSIGHT

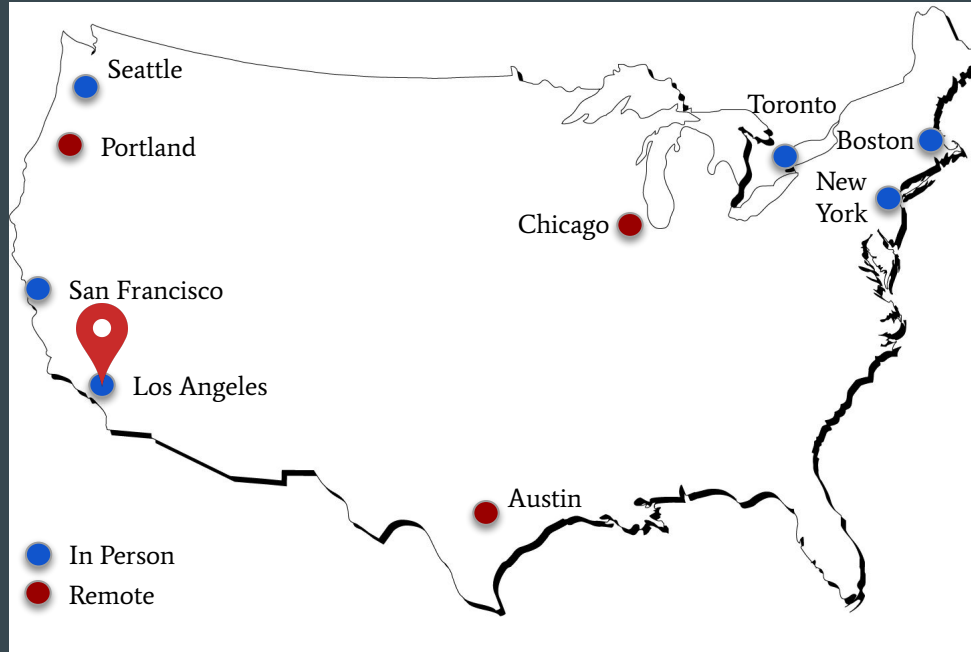
# Interview Preparation



# Program Timeline

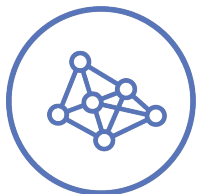


# Where are we?



# Insight Offerings - Which one to pick

## Data Science Program



- PhD in quantitative fields.
- Have worked in analysing data.
- Good problem solving skills

## Data Engineering Program



- Engineering background.
- Worked on and maintained building engineering systems.
- Java/Python

## Health Data Science Program



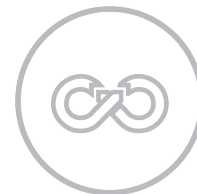
- Postdoctoral researcher, medical doctors
- Interested in genome sequences, clinical trials.

## Artificial Intelligence Program



- Engineering background.
- Have worked on training and deploying ML or NN.

## DevOps Engineering Program



- Systems admin and Linux background.
- Problem solver critical thinker.
- Can understand containerized sys.

# New Programs - More focused domains

## Security Engineering Program



- Designing security measures
- Building secure applications.

## Decentralized Consensus Program



- Blockchain technology
- Smart contract management
- Decentralized architectures

facebook

LinkedIn

okcupid

schireson:

meetup

Zillow

ACTIVISION

jet

VIACOM

DataKind

greenhouse

Square

Capital One

reddit

NBC

STITCH FIX

airbnb



fitbit

ZocDoc

verizon

VECTRA

twitch

climacell

OSCAR

wayfair

Pinterest

JAWBONE

BLACKROCK

Microsoft

TWO SIGMA



JPMorgan

salesforce

amazon

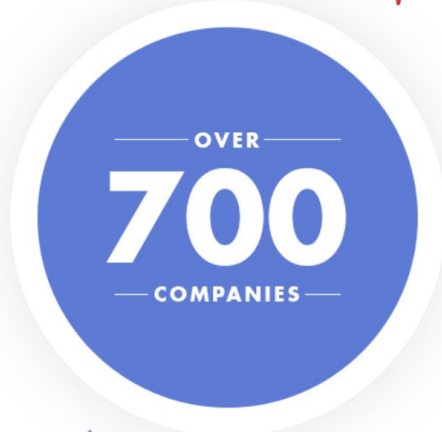
enigma

SAMBA TV

nielsen

intuit

KHANACADEMY



# UCSF and Insight Alum

Melissa Runfeldt - Senior Data Scientist (Highmark Health)

Dominic Tong- Data Scientist (InsightRX)

Natalie Korn- Data Scientist (AppAnnie)

**Peter Chisnell**- Scientist/Data Scientist (23andMe)

Jasmine Hughes- Data Scientist (InsightRX)

Yiran Chen- Data Scientist (Abbot)

**Chung Yan Cheung** - Data Scientist (Rally Health)

Rebecca Clarkson- Software Engineer (Apple)

**Claire Tang** - Data Scientist (SambaTV)

Keshi Jordan- Data Scientist (Rally Health)

Jessie Lee- Data Scientist (Lyra Health)

Emily Anderson- Data Scientist (Clover Health)

John Nicoludis - Data Scientist (Invitae)





All applications due 3/16!



<http://bit.ly/UCSF-Spring>

INSIGHT

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)