

Demystifying Data Science

There's no one way to
become a data scientist.



METIS

```
/* Hi, I'm Greg. */
```



Greg Reda

@gjreda

Sprout Social

Data Science Lead



- Hi
- If that image looks familiar, you've probably read some pandas tutorials
- Here to talk about my path: general analytics or business intelligence to data science
- Education first though

/* Background */

- Planned on being lawyer; saw the light
- Economics & Information Technology
- Loved econometrics; wasn't great at it
- Never planned on using IT degree



- Planned on being a lawyer from 6th grade to Jr in college
- Dad tried talking me out of it
- Studied Econ + IT
- No intention of using IT education -- double majored because it was easy
- Bad at math, loved econometrics, was OK at it
- Wanted to go be Research Analyst, then the market crapped out
- Soooo ... PwC

PwC

- Forensic Technology
- Lots of Databases, SQL, & VBA
- Kind of data engineer: lots of ETL w/ Microsoft stack

Extra: Learning Python, R, + a lot of Twitter



- Got job offer as economy was getting bad
- Sounded cool: tech + analytical skills for legal stuff
- Kind of “data analyst/engineer” stuff -- Lots of database and ETL work
- Got bored: Answer to every project was Oracle + SQL + VBA.
- Started learning Python, R, ggplot, etc.
- Sooo much Data Nerd Twitter - learned so much

GrubHub

- First data hire: Data Analyst
- Start-up: learn as you go
- See something that can be better?
Make it happen.

Extra: Started blogging, digging
into machine learning



- Optimized for learning
- Wanted to work in tech
- First data hire -- was able to shape role a bit -- no prior base
- Start-up!

- Take on as much responsibility as I wanted
- See something that can be improved? Go do it
- Started doing random data projects and blogging about it
- Probably the best professional decision I've made over the years

- Datascope: People! Sports!
- Freelance: Clients, eek!
- Sprout: Four people, new team, new infrastructure, new function

Extra: “Process”, “Systems”



Datascope

- amazing people + more learning!
- NBA team & Deadspin! Huge sports fan

Freelance

- could I strike it on my own? Freedom!!
- short-term view -- not me
- wanted to think long-term and deeply about

things

Spout

- building a data science team
- Lots of thinking about “process” and “systems” to enable teams & org

Things I believe



- Close with some things I believe about data science
- Or at least about *becoming* a data scientist

1

`/* You don't have to be great
at programming */`

(But you have to know a little).



- OK at programming -- not “great”
- Need to know how to do it, but it’s a tool
- Usually not working in huge codebases
- Simple is better - scripts or whatever to get job done

/ You don't have to be great
at math or statistics */*

(But you have to know a little).



- OK at math/stats -- not “great”
- Think in probability distributions
- How can you nullify your work?
 - Try to break things
- Not solving systems of linear equations -- someone's already written that
- Simple is still better
 - don't immediately reach for a model!

3

/ You must be great at
learning and curious */*

This is more important than everything else.



- Tech and DS move fast - constantly evolving
- Jump over hurdles, there will be many
- Problems are hard, curveballs will be thrown
- Data sucks - things are never *just* right

Data Science Demystified

Using **data**, **code** (*that someone else has already written*), and **math** (*that someone else has already done*) to:

- Answer questions
- Solve problems
- Draw a line through things

Learn and be curious.



- Data - gotta have that
- Code - soooooo many libraries that make things easier for you
- Math - encapsulated within all the libraries
- Answer questions, solve problems, and do great things
- Nod to Brian Lange -- “Draw a line through things” talk is great
- Learn and be curious -- above all else!