A Tale of a Data-Driven Culture

Gloria Lau
VP of Data, Timeful (Acquired by Google)

Keynote @ Spark Summit, 2015

Acquired Scientist
Google
May 2015 – Present (2 months)

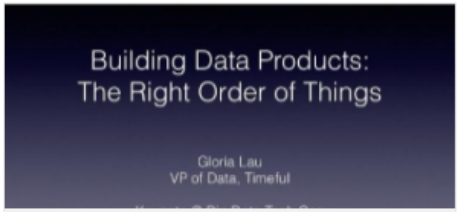
Consulting Associate Professor
Stanford University
August 2011 – Present (3 years 11 months)

My research interest is in legal informatics and e-rulemaking.

Advisor
Various startups
2014 – Present (1 year)

VP of Data
Timeful (acquired by Google)
September 2014 – May 2015 (9 months)

Time is our most precious resource but we seldom manage it like one. Can we be smarter in managing our time? How do we quantify productivity with data, and how do we change behavior given what we know from data?



Keynote at Big Data Tech Con SF 2014

SXSW Interactive Panel - How Beautiful Tool…

Senior Manager, Data Science
LinkedIn
February 2011 – August 2014 (3 years 7 months)

http://www.linkedin.com/in/gloriatlau/
@gloriatlau

This talk is

- a hindsight on why Spark matters to non-engineers too

Every company aspires to be more data-driven

- What is involved?
  - Analytics
  - Deep dives
  - Explorations
- on *production* data *real time* with *speed*

Common implementation

- Hire more data scientists, turn them into SQL monkeys

Problem

- *everyone* should have be able to dig into data
- **long** innovation cycle
- monkeys *do not* last

Truly data-driven

- Say **NO** to job descriptions
  - easier said than done

How to throw out job descriptions

Friction = Bad
Less friction = Less bad

- Most in your company **want** answers from data
- Most in your company *are* capable of finding answers from data, *given* the right tool
- Your goal is to find that tool that minimizes friction to data

Biggest friction to data

- Lack of speed
- Imagine your favorite search engine's SLA is… 10 seconds

Experiment

- Kill the SQL monkey by turning *everyone* into their own SQL monkeys
  - Give PMs, designers and *everyone* access to Databricks.
  - Fail fast. Learn from examples. Visualize. Inspect. Fail again. Iterate.
  - You have to trust your team to be capable. Give them the tool and they shall learn to be data-driven.

Experiment Results

- **Everyone** spends time with data, forms their own insights, and no longer needs a crutch to explore
- Data scientists* focus on building data products
  - feature engineering etc *also* benefits from speed
- Much ***shorter*** innovation cycle.

* disclaimer: very liberal definition throughout

What did we learn

- Every company aspires to be data-driven. You hire SQL monkeys. Wrong.
- Democratize thou data by removing friction. Speed should be top on your list.
- #SQLMonkeyEveryone