BUILDING YOUR DATA SCIENCE CAPABILITIES

@lenguyenthedat

The Asia InsurTech Summit 2018

~ 2010 UNIVERSITY OF OXFORD - MSc. Computer Science

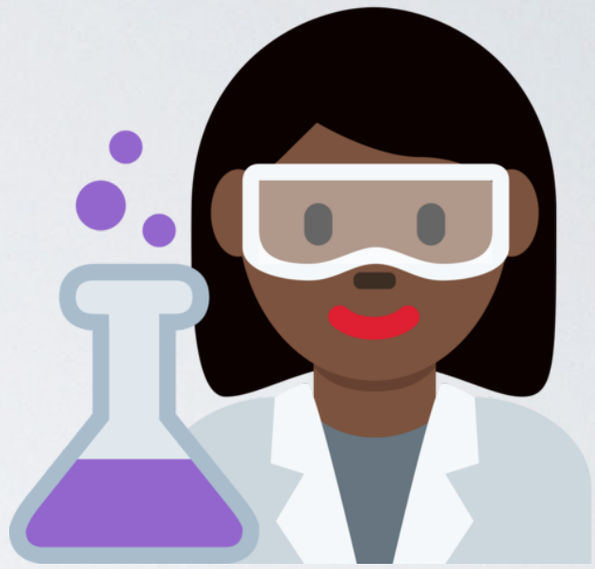
2011-2013 A*STAR Research SINGAPORE

2013-2015 ZALORA

2016-2017 honestbee

2017-NOW UBER

"A unicorn is just a horse with a ***point of view****" - Ron Sexsmith*



"A Data Scientist is someone who possesses the skill sets to leverage data, delivering ***business impact****"*

BUILDING YOUR DATA SCIENCE CAPABILITIES

STEP #1 - HIRING YOUR DATA SCIENTISTS

BUILDING YOUR DATA SCIENCE CAPABILITIES

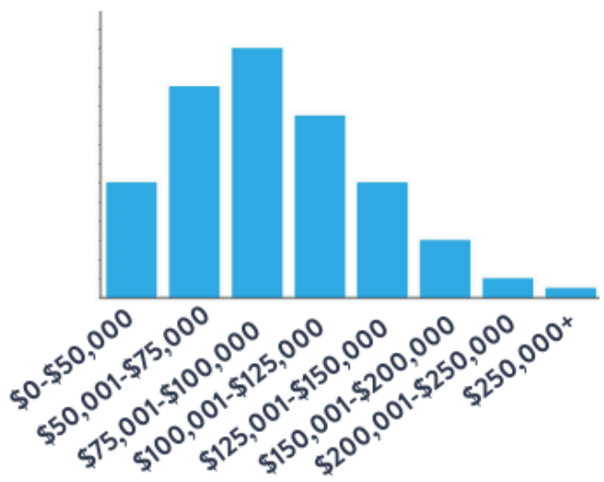
STEP #1 HIRING YOUR DATA SCIENTISTS

BUILDING YOUR DATA SCIENCE CAPABILITIES

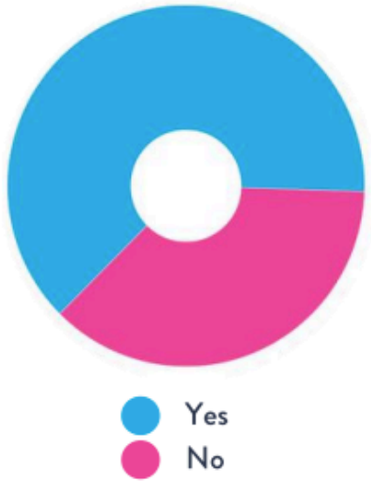
STEP #1 - LET'S TAKE A STEP BACK AND TALK NUMBERS

BUILDING YOUR DATA SCIENCE CAPABILITIES

	EXPERIENCE LEVEL		
	Junior	Mid	Senior
Average Annual Salary	$59,480	$104,626	$136,783



27% of respondents are being paid between $75,000 - $100,000



63% of respondents receive an annual bonus, with the average amount received being $16,858

https://www.techinasia.com/data-scientist-salaries-singapore-revealed

Salary of a data scientist or a data engineer can easily go higher than **$100k**, or in some cases, over **$200k** a year

$100k-500k spend on tools, servers, and licenses a year

BUILDING YOUR DATA SCIENCE CAPABILITIES

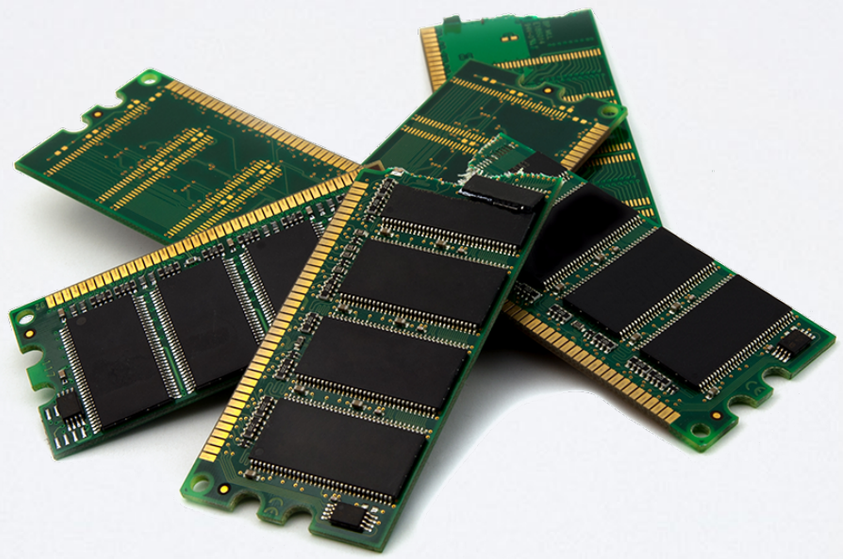
Ballpark figure: ***$1M-2M/year***

BUILDING YOUR DATA SCIENCE CAPABILITIES

QUESTION #1 - DO YOU REALLY NEED DATA SCIENTISTS?

QUESTION #1 - DO YOU REALLY NEED DATA SCIENTISTS?

"My data is too big, Excel wouldn't load anymore! I need a data scientist!"



32GB DDR4 RAM = ***$400***

QUESTION #1 - DO YOU REALLY NEED DATA SCIENTISTS?

Sometimes you could hire an external consultant to solve the actual business problem you have

QUESTION #1 - DO YOU REALLY NEED DATA SCIENTISTS?

"So you want to hire me as a Data Scientist for Intelligent Virtualized Deep Machine Learning Real-time Big Data in the Cloud for Social Networks? Ok, but if you also want Hadoop, increase my salary by 50%."

QUESTION #1 - DO YOU REALLY NEED DATA SCIENTISTS?

SKILLS AND QUALIFICATIONS

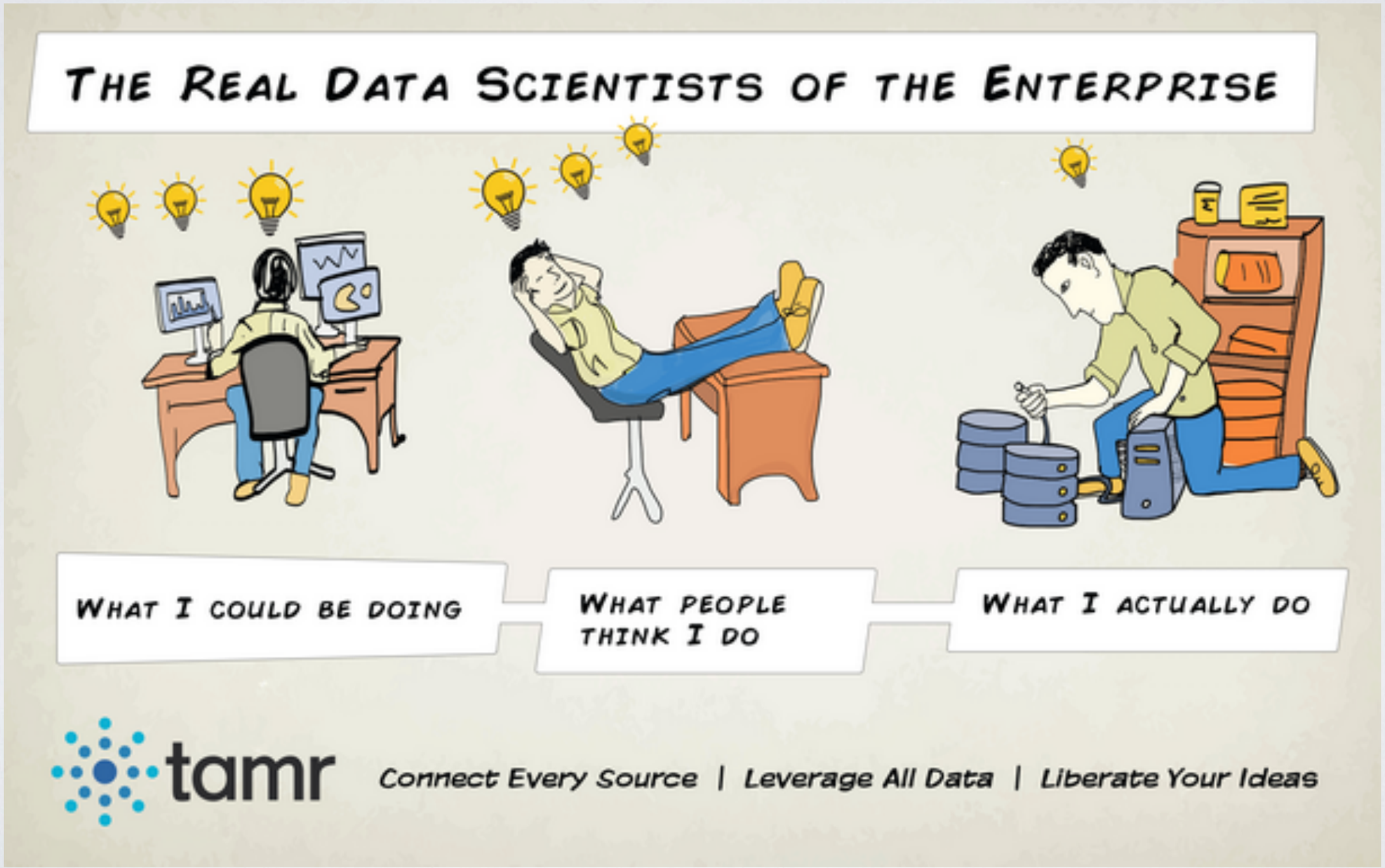
- PhD or equivalent experience in machine learning, computer science or
- logistic regression, gradient boosting, random forests, SVMs, bagging, RNNs, etc.
- hierarchical and multi-label classification models
- NLP, computational linguists, search
- graph algorithms and/or social network analysis
- python or perl
- as weka, R, scikit-learn or
- Hadoop, Hive, Pig etc.

Do you have a clear project, direction, business impact or do you just need some keywords?

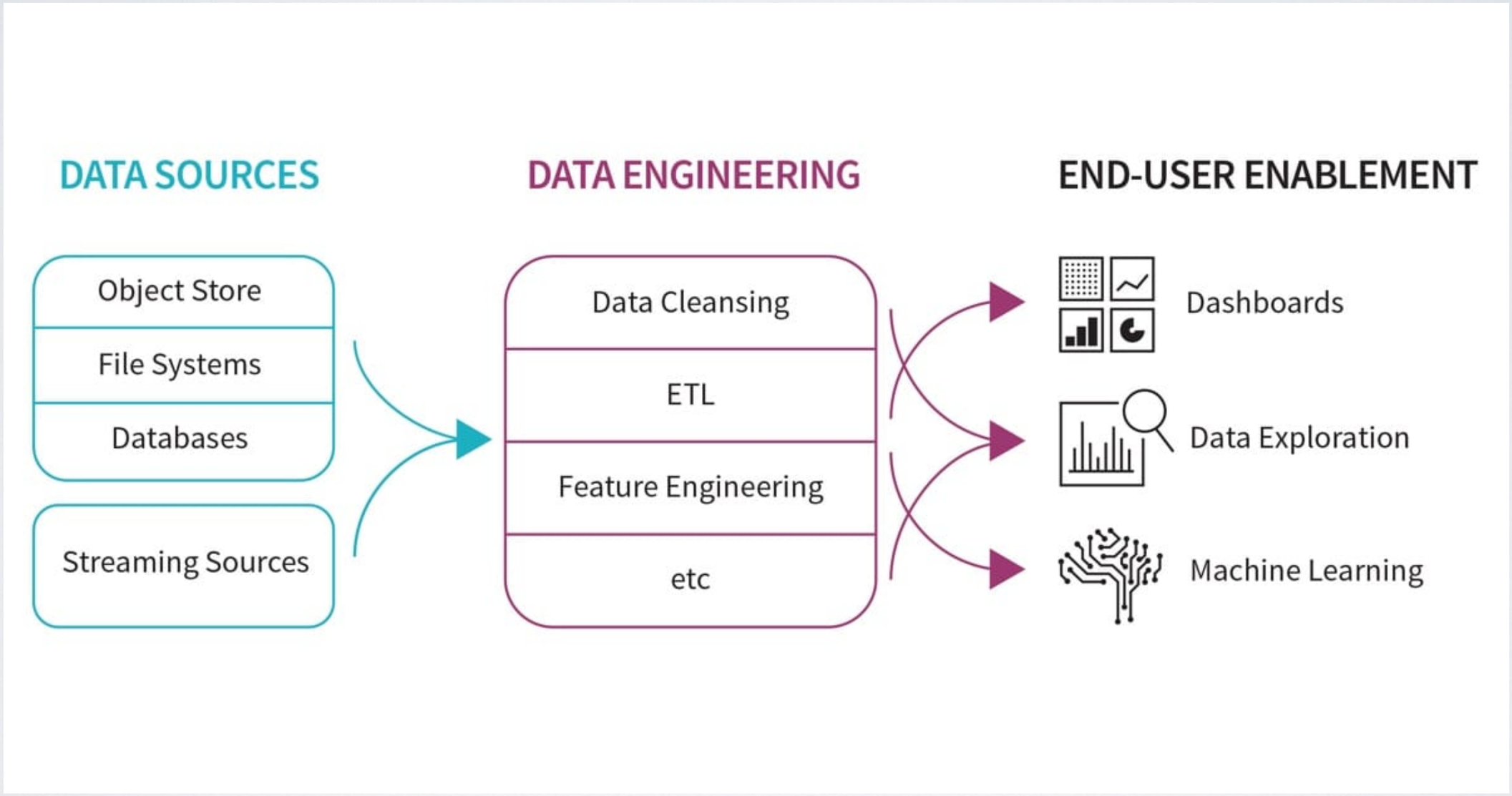
BUILDING YOUR DATA SCIENCE CAPABILITIES

QUESTION #2 - DO YOU HAVE YOUR DATA INFRASTRUCTURE READY?

QUESTION #2 - DO YOU HAVE YOUR DATA INFRASTRUCTURE READY?



QUESTION #2 - DO YOU HAVE YOUR DATA INFRASTRUCTURE READY?



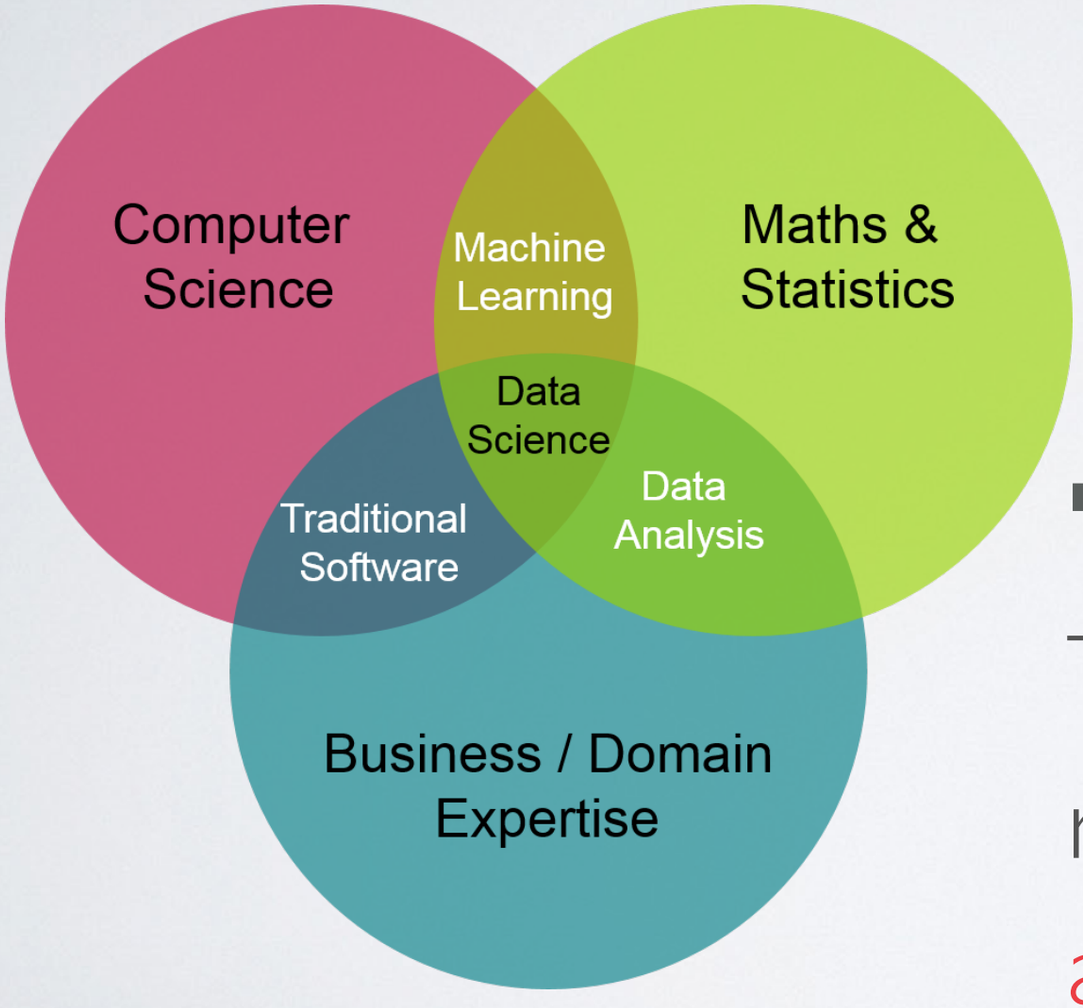
Data Engineers enable Data Scientists to do their jobs more effectively.

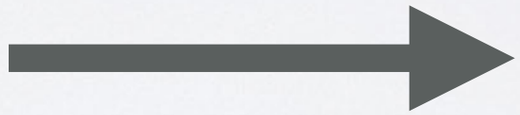
BUILDING YOUR DATA SCIENCE CAPABILITIES

HIRING YOUR DATA SCIENTISTS

when there is clearly a shortage of "unicorns"

THE TRADITIONAL APPROACH TO DATA SCIENCE



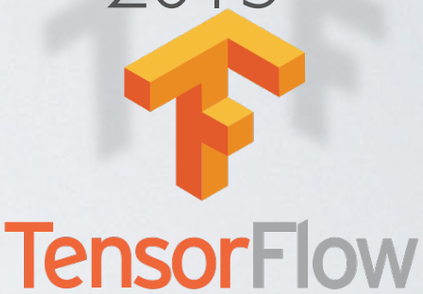


Business Impact

The traditional data scientist requires years of training done ahead of time

WHAT CHANGES IN THE LAST FEW YEARS?

2010 2011 2014 2014 2014 2015



Modern Data Science toolkits

2010 scikit learn

2011 R

2014 jupyter

2014 dmlc XGBoost

2014 Spark MLlib

2015 TensorFlow

Practical and real-world experience + ***Modern Data Science toolkits*** + *Continuous learning ability*

=

Education can then start with the practical knowledge at the same time, or continue after one has already made **business impact**

HIRING YOUR DATA SCIENTISTS

Try not to look for people who ticks all the boxes

Focus on demonstration of real business impact and ability to learn new things

TAKEAWAYS

Your data science capabilities do not start from hiring data scientists

The shortage of unicorn data scientists is real, and how we could emphasize on real-world experience and learning ability instead

HOW
WHY
WHEN
WHERE
WHERE
WHY
WHERE
WHO
WHO
WHEN
WHAT
WHAT
WHY
WHAT
WHEN
WHAT
WHEN
WHO
WHERE
WHY
HOW

THANK YOU!