



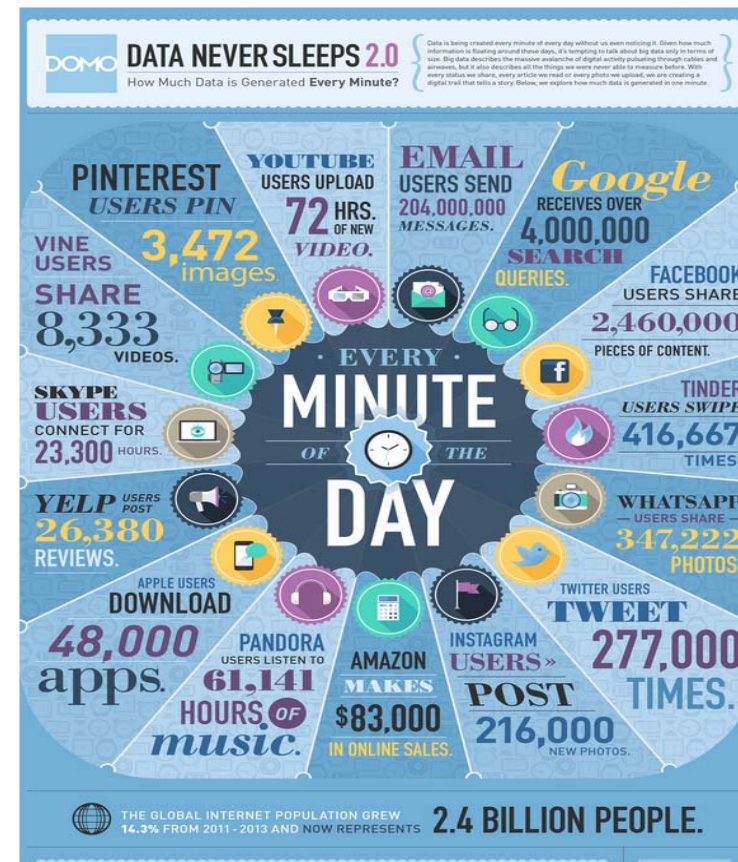
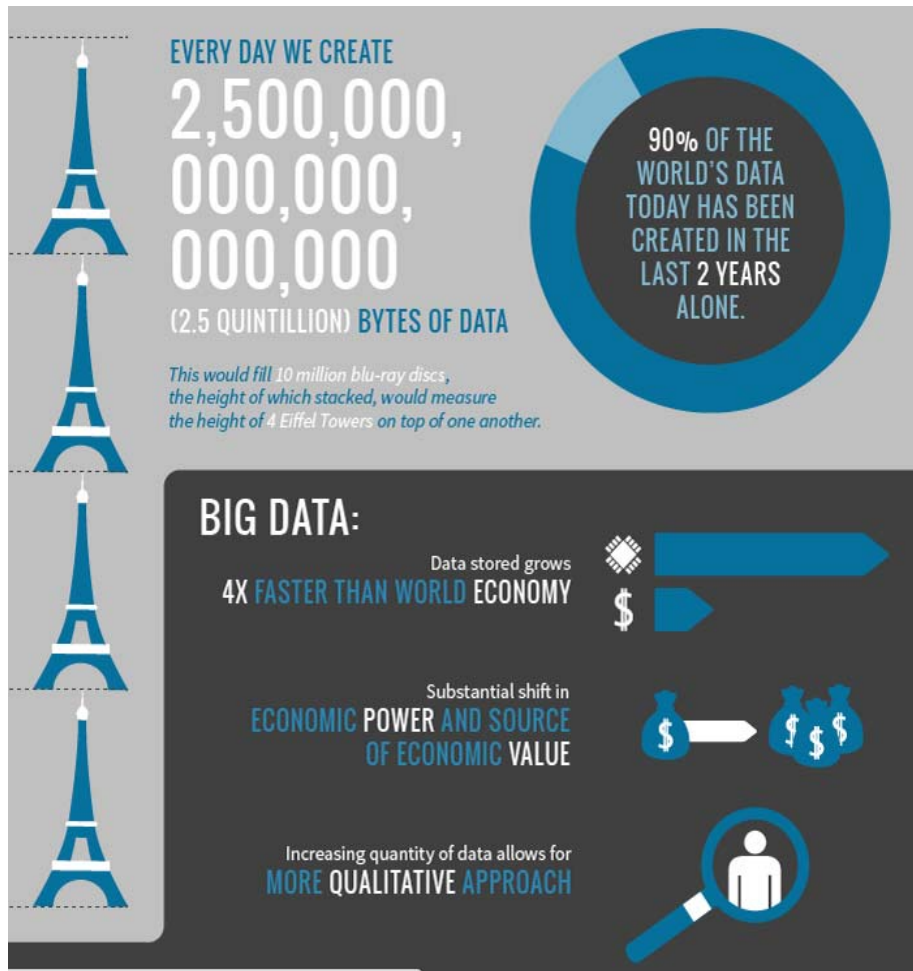
# How does Data Science play out in industry and what skills are we looking for?

Nov 7th<sup>th</sup> 2019

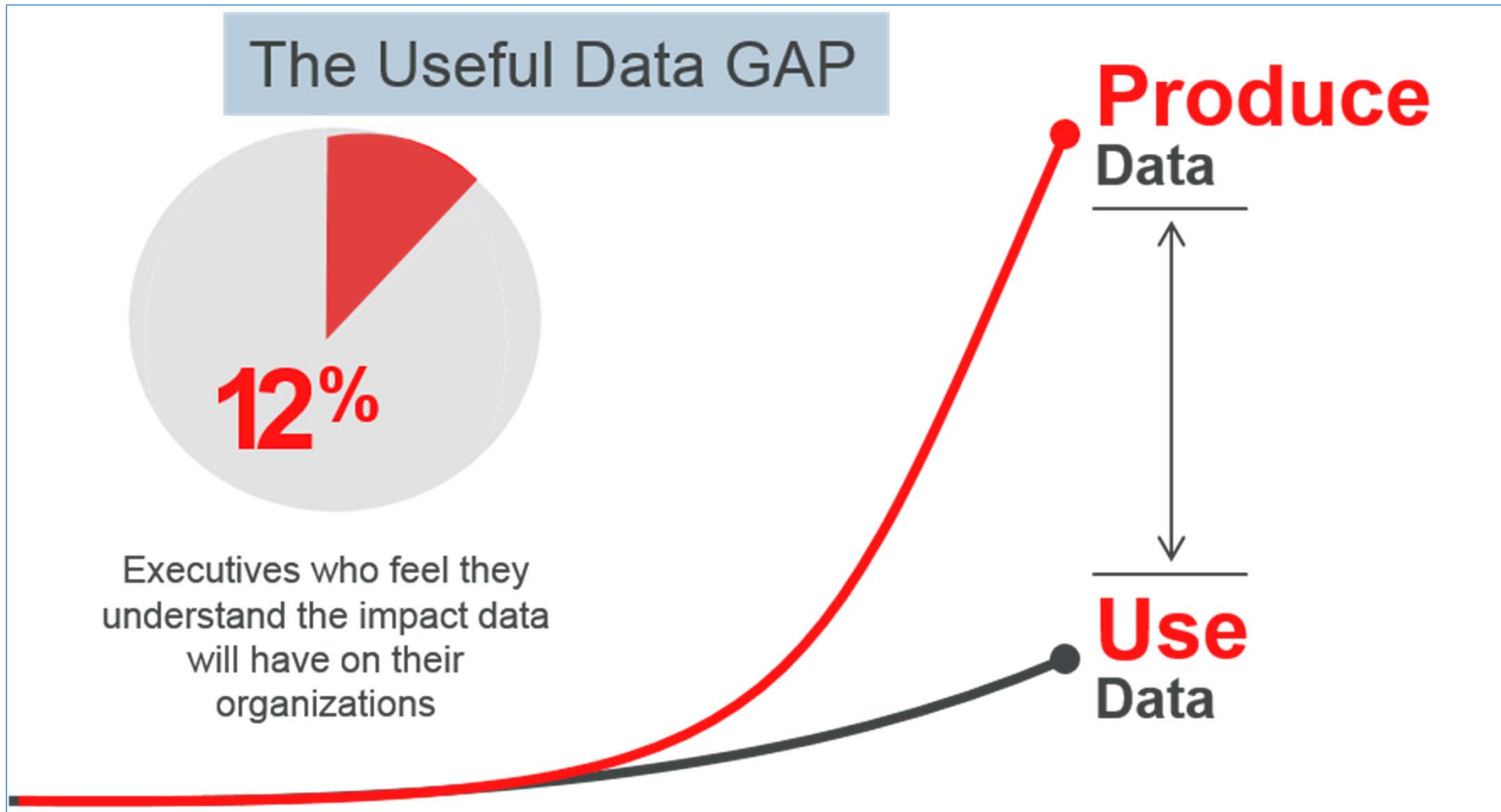
Dr Virginia Wheway

Manager, Data Systems and Strategy, Thiess / Chief Data Officer CIMIC

# More and more data is being collected....



# Most organisations don't fully utilise their data



# Demand for Data Scientists is increasing dramatically....

## IBM Predicts Demand For Data Scientists Will Soar 28% By 2020

<https://www.forbes.com/.../ibm-predicts-demand-for-data-scientists-will-soar-28-by-2020> ▼

May 13, 2017 - Annual demand for the fast-growing new roles of **data scientist**, **data developers**, and **data engineers** will reach nearly 700,000 openings by 2020. By 2020, the number of **jobs** for all US data professionals will increase by 364,000 openings to 2,720,000 according to IBM. **Data Science and Analytics (DSA)** ...

## The Supply And Demand Of Data Scientists: What The Surveys Say

<https://www.forbes.com/.../the-supply-and-demand-of-data-scientists-what-the-surveys-say> ▼

Apr 30, 2015 - Gil Press , Contributor I write about technology, entrepreneurs and innovation. Opinions expressed by Forbes Contributors are their own. Tweet This. The median salary of a junior level **data scientist** is \$91,000 · When changing jobs, **data scientists** see a 16 percent increase in their median base salary.

## Data science and analytics job market predictions for 2017: PwC

<https://www.pwc.com/us/en/library/data-science-and-analytics.html> ▼

To help you think about what this means for your own hiring, we've worked with the Business-Higher Education Forum to examine the landscape of **jobs** requiring **data science** and analytics competencies and skills. We'd like to share our data-driven predictions about what you will need to do to hire the best in 2017.

## Proficiency in Data Science Skills by Job Role | Data Science ...

[www.datascienceassn.org/content/proficiency-data-science-skills-job-role](http://www.datascienceassn.org/content/proficiency-data-science-skills-job-role) ▼

Oct 13, 2016 - Proficiency in **data science** skills can easily be understandable by this chart. Halloween Coupons I hope if we follow the basics of this chart we will come to know about the data professionals who select their **job roles**. Thanks for this amazing chart description to increase our knowledge. Log in or register to ...

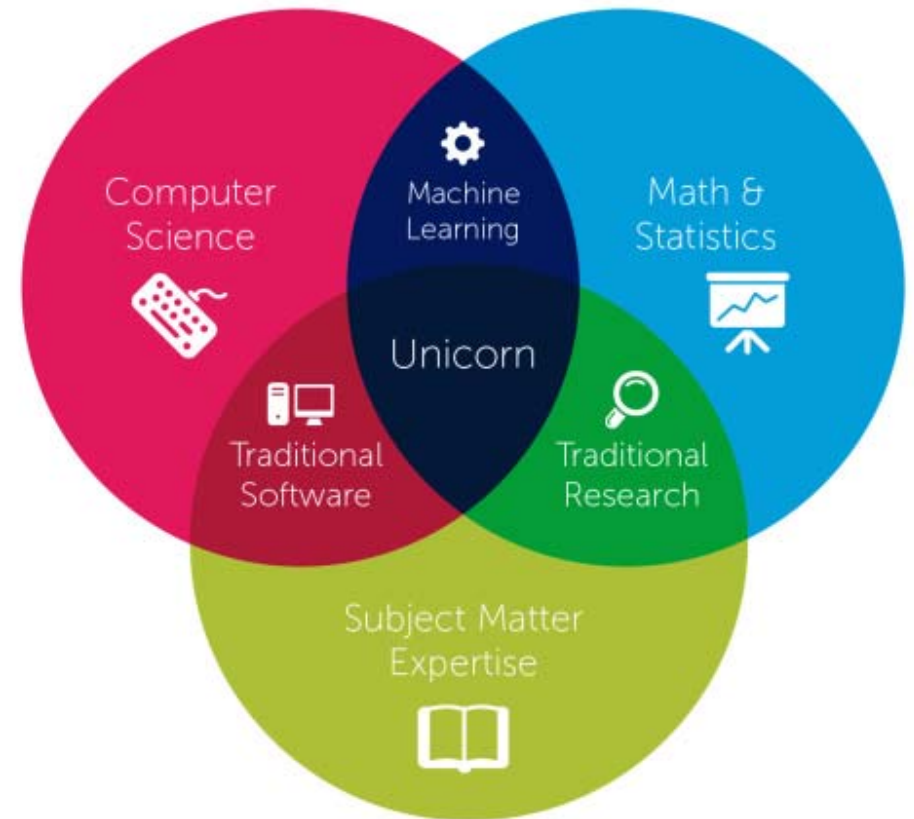
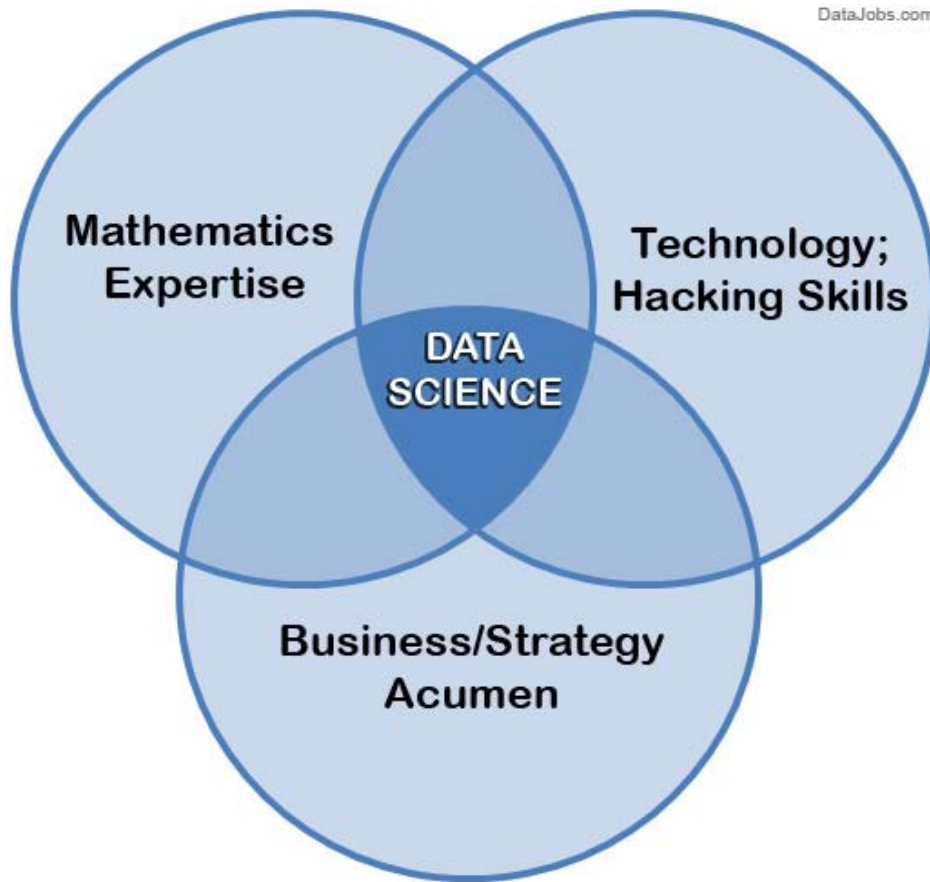
## The growth of data science over the last two years: 300% - Data ...

<https://www.datasciencecentral.com/.../the-growth-of-data-science-in-the-last-two-year...> ▼

Nov 8, 2014 - Other indicators of **growth** include keyword trends for **data science** related keywords (number of searches on Google), and number of **job openings** (there's data available on Indeed, but it looks wrong). Finally, figure 1 offers the most realistic picture: **growth rate** is constant. LinkedIn and Twitter charts ...

# What is Data Science?

DataJobs.com







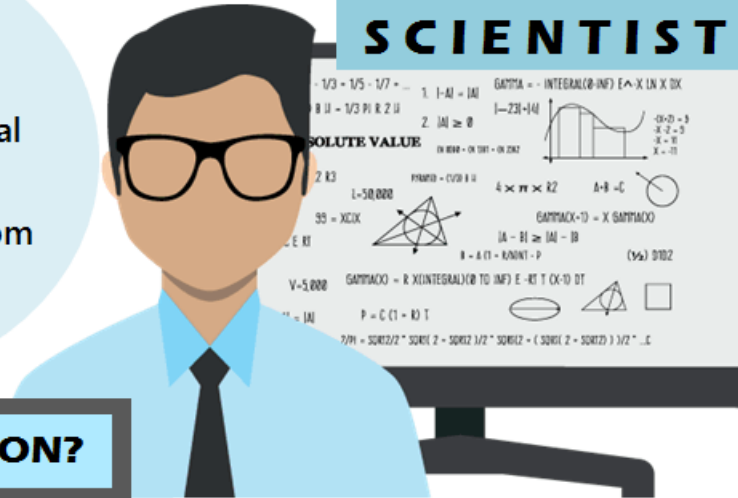
## WHO AM I?

I am a part analyst & part artist. I use my analytical and technical abilities to extract meaning / insights from massive data sets.

# D A T A



## SCIENTIST



## WHAT DO I DO?

1. I cleanse existing raw data & build models to predict future data.
2. I go beyond merely collecting and reporting data, to look at data from multiple angles & give meaning to it.
3. I identify the correct business problem(s) & offer solutions (via visualizations, reports or blogs) by best applying the data.

## WHAT DO I RELY ON?

1. Analytics
2. Predictive Models
3. Statistical Analysis & Modeling
4. Data Mining
5. Sentiment Analysis
6. What-if Analysis

## THE PROCESS I FOLLOW

Define  
Problem

Structure  
Data

Use  
Programming  
Language

## WHAT DO I EARN?

After oil & gas geologists, mine is the 2<sup>nd</sup> highest paid job in the world!

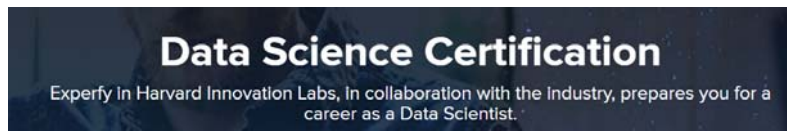
## HOW DO I HELP ORGANIZATIONS TODAY?

- Increase data accuracy
- Develop strategies
- Improve operational efficiency
- Reduce costs
- Mitigate risks
- Offer personalized products/services

# No shortage in tools, and many are open source



# Other offerings and trends



## The 9 Best Free Online Big Data And Data Science Courses - Forbes

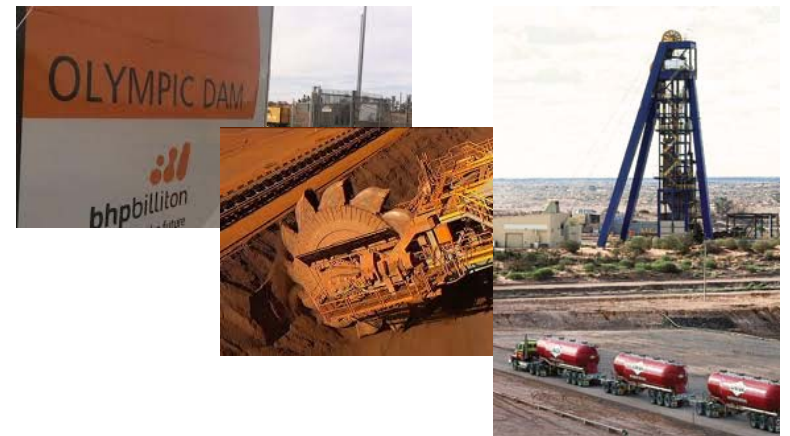
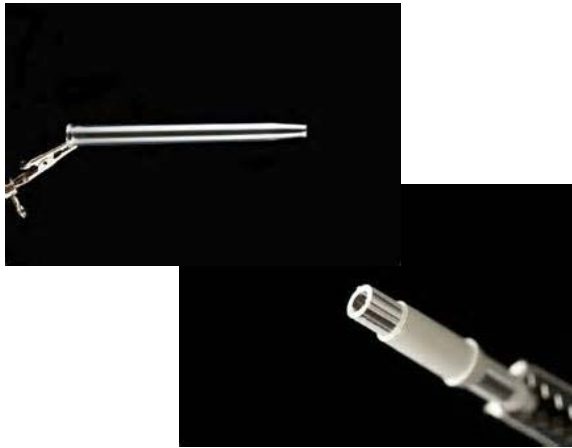
<https://www.forbes.com/.../the-9-best-free-online-big-data-and-data-science-courses/> ▼

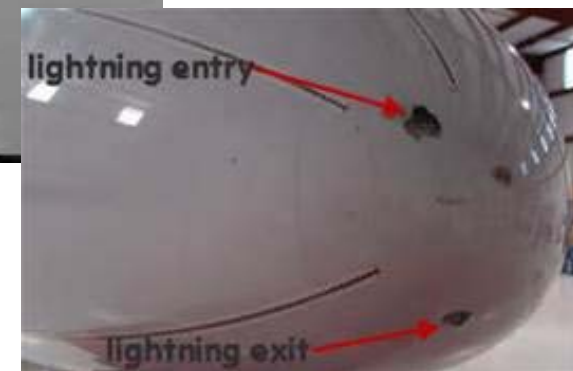
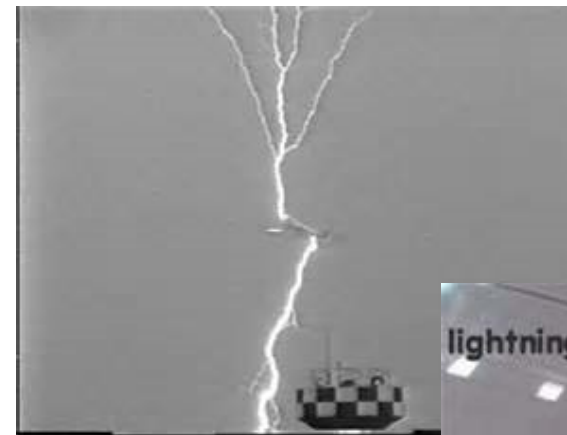
Jun 6, 2017 - The ability to turn data into insights is one of the most thought-after skills anyone could have in today's big data world. A great way of finding your way into a **data science** or analytics career is through fee **online courses**. Here I look at my favorite free **data science online courses**.



Industry problems I've  
helped solve as a Data  
Scientist









**THIESS**





# Examples



## WEEKLY VERIFIED STORAGE AND SUPPLY REPORT

Thursday 16 June 2016

WaterNSW publishes weekly bulk water storage and supply reports to provide information about how much rainfall has been recorded, how full the dams are, and how much bulk water we've supplied to our customers.

### Catchment rainfall

During the week the following rainfall (approximately) was recorded in the catchments:

WARRAGAMBA	UPPER NEPEAN	WORONORA	SHOALHAVEN	BLUE MOUNTAINS
Negligible	Negligible	Negligible	Negligible	Negligible

### Water supply

At 8 am Thursday, 16 June 2016, the inflow rate to Warragamba Dam was 1,450 megalitres per day and to Tailowa Dam was 2,510 megalitres per day.

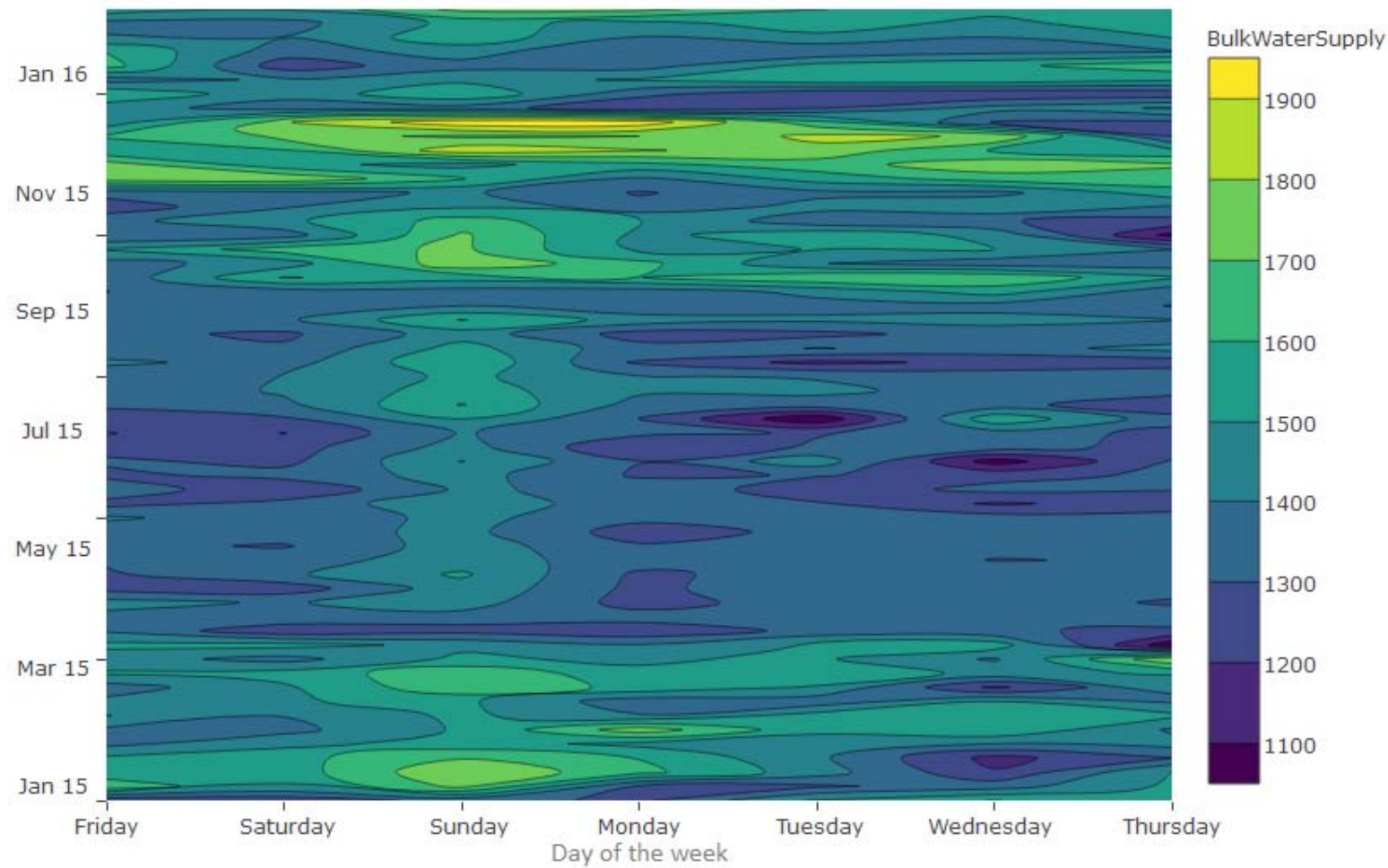
Total available storage as at 8 am Thursday, 16 June 2016 was **97.4%**, an increase of 0.4% since last Thursday.

DATE	MAXIMUM SYDNEY TEMPERATURE	SYDNEY RAINFALL	BULK WATER SUPPLY	AVAILABLE STORAGE AT 8AM	% STORAGE AT 8AM
Fri 10.6.16	22°C	0.2 mm	1 352 ML	2 508 315 ML	97.2%
Sat 11.6.16	21°C	Nil	1 408 ML	2 512 662 ML	97.3%
Sun 12.6.16	20°C	Nil	1 509 ML	2 513 439 ML	97.4%
Mon 13.6.16	18°C	Nil	1 283 ML	2 513 977 ML	97.4%
Tue 14.6.16	20°C	Nil	1 337 ML	2 514 966 ML	97.4%
Wed 15.6.16	21°C	Nil	1 410 ML	2 514 367 ML	97.4%
Thu 16.6.16	21°C	Nil	1 387 ML	2 513 673 ML	97.4%

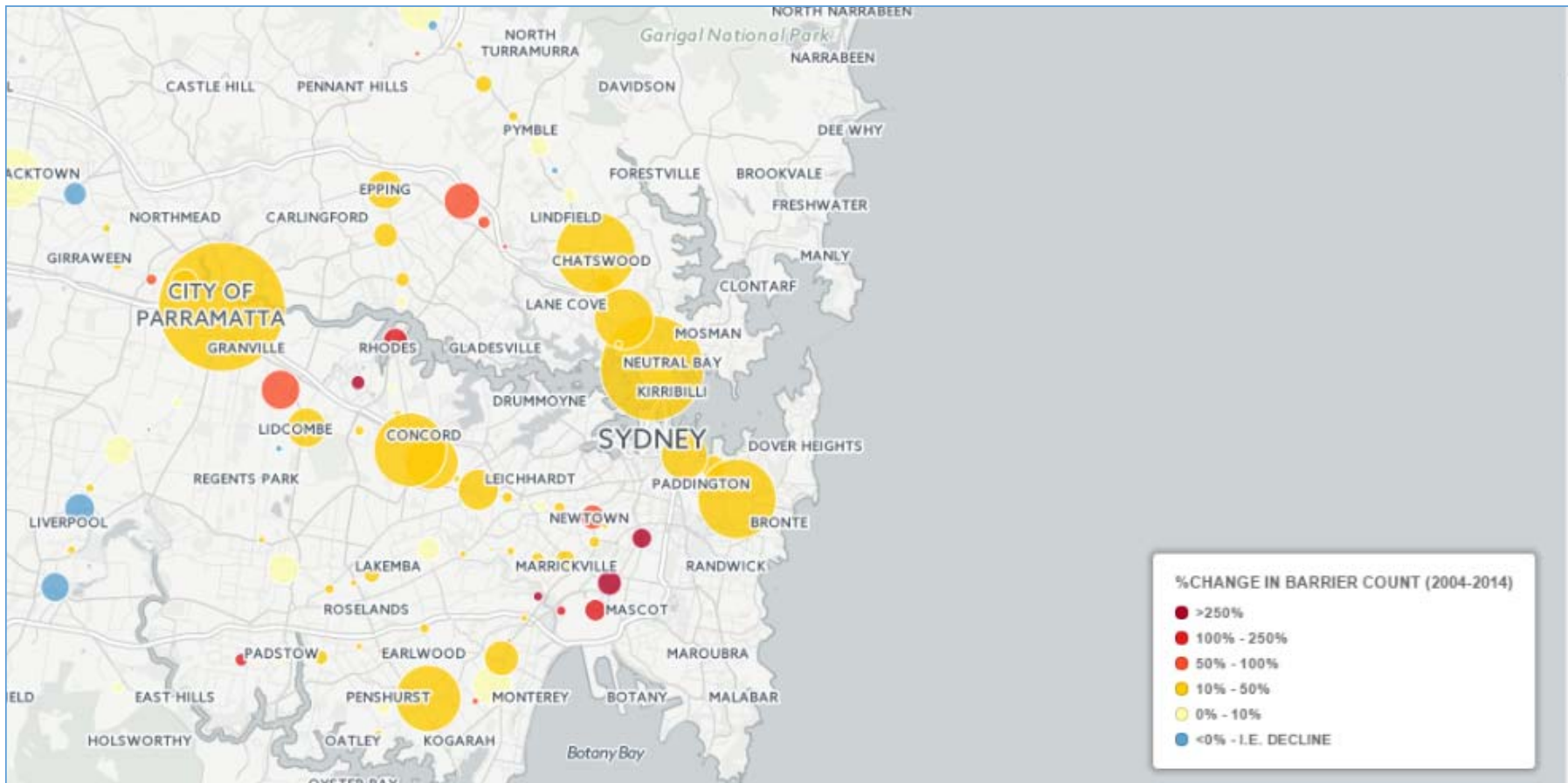
Note: Maximum Sydney temperature, rainfall and water supply figures refer to the 24 hours to 8 am.

No more tables....

# WaterNSW Bulk Water Supply Jan15 – Feb 16

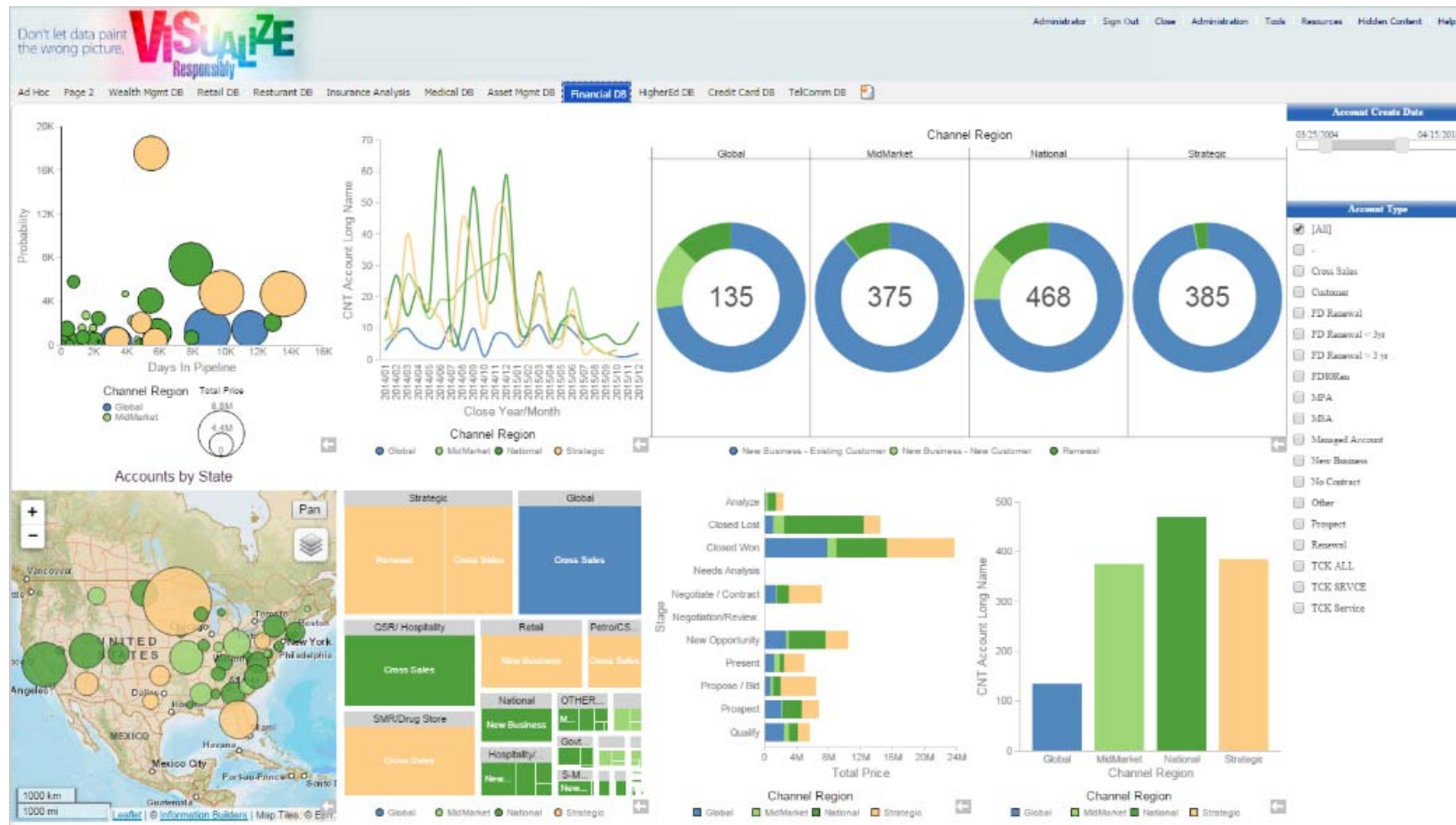


# Interactive maps example - NSW station barrier counts



[https://vwheaway.cartodb.com/viz/f14bbef8-cfbd-11e5-bc6e-0ecd1babdde5/public\\_map](https://vwheaway.cartodb.com/viz/f14bbef8-cfbd-11e5-bc6e-0ecd1babdde5/public_map)

# High demand for production of self-service dashboards





# Sydney's Road Network Performance **This Week**

This week

Last week

By location

## Cost of congestion this week

**\$234,180**



Last week \$232,597 **1%**

## Average percentage of speed limit



This week  
**74.1%**



Last week  
**74.8%**



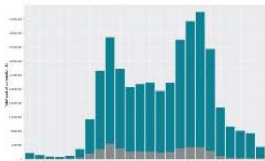
**1%**

## Best & worst performing TMCs



Average percentage  
of speed limit (6-9am)  
25 100

## Total cost of congestion (Pacific Hwy)



■ Heavy vehicles ■ Light vehicles



## WEEKDAY WORST performing road segments



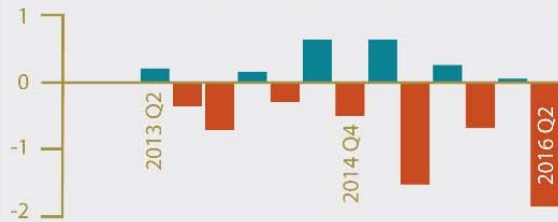
## WEEKEND WORST performing road segments



Road	Segment
Missenden Rd	Southbound before Carillon Ave
M5 East Fwy	Westbound after Princes Hwy Crossroad
Park St	Westbound between Pitt St & George St
Parramatta Rd	Westbound between Luke Ave & Shaftesbury Rd
High St Penrith	Westbound after Henry St

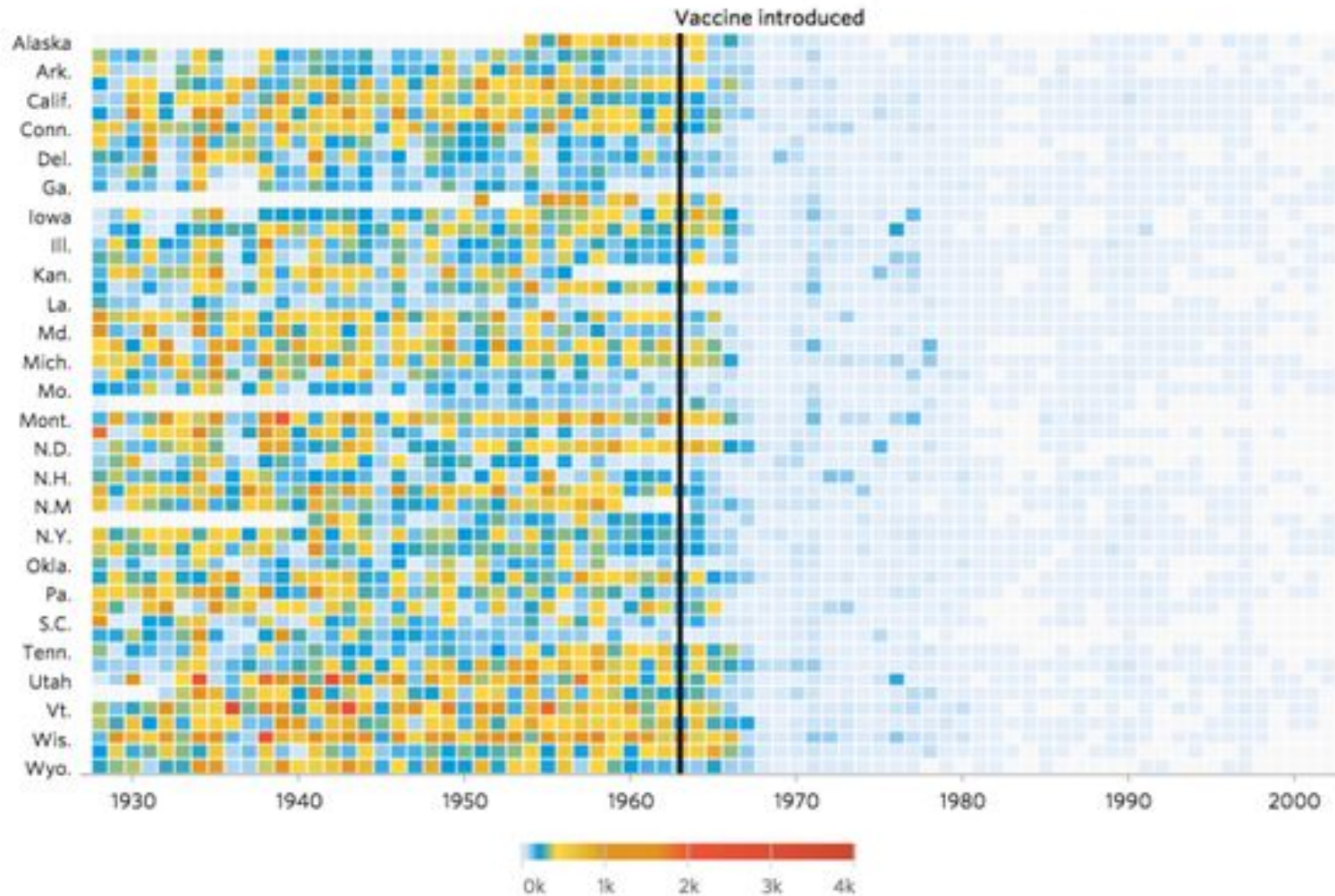
Road	Segment
Illawarra Rd Marrickville	South-Westbound at Sydenham Rd
Market St	Kent St
King St Newtown	South-Westbound
Military Rd	Avenue Rd
King St	After Missenden Rd

## Quarterly % change in average speed



# Winner of international Data Viz competition

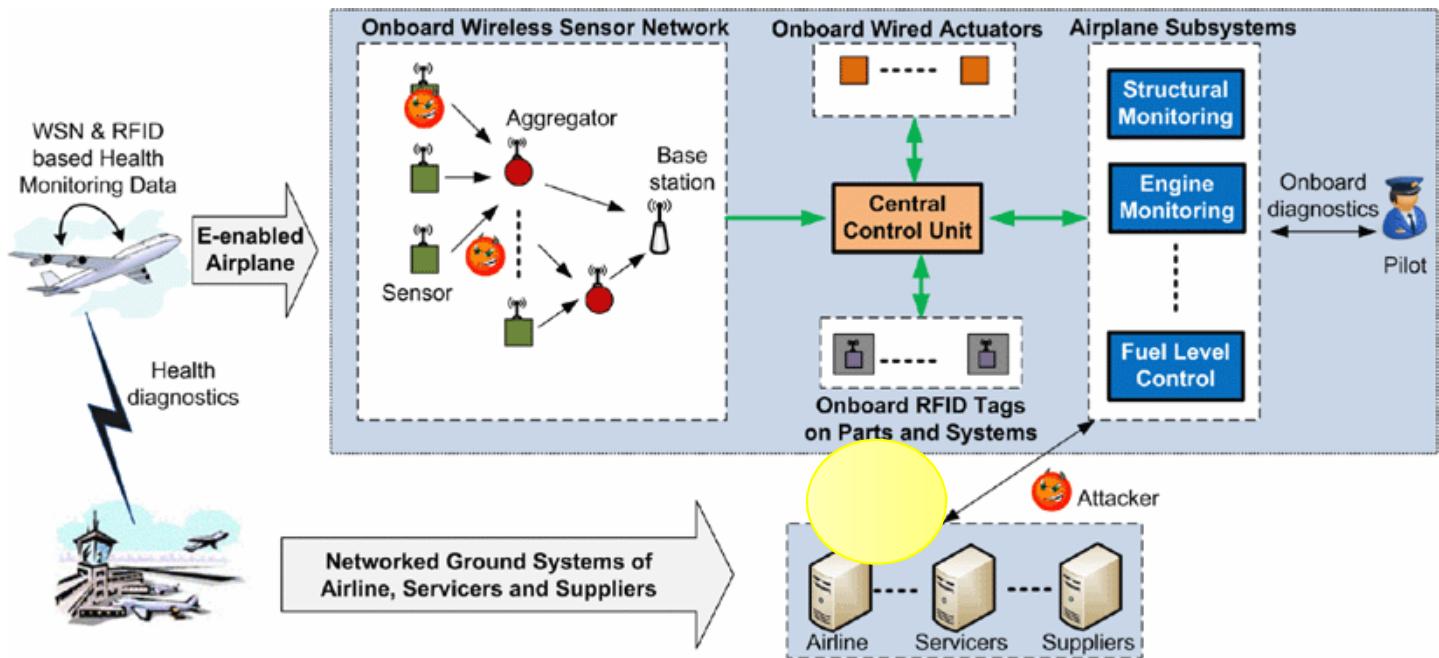
## Measles

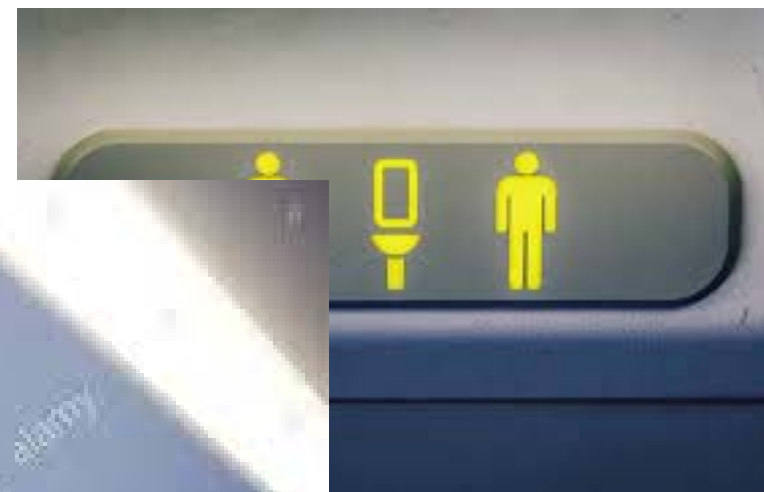


# Predictive Maintenance – Data Mining













# Building a Data Science team in industry

# WHY build an internal Data Science team?

---

## More “Reactive”

1. There are increasing regulatory or tender requirements about the collection and reporting of timely, accurate information. More data requirements needed to even be able to respond to a tender

9. You monetises your data and insights (companies that have successfully done this are Boeing via Airplane Health Management, GE, CAT (to some extent), Supermarkets, Banks, social media companies)

## Capabilities required in a “balanced” team

Data Engineer / Data Warehouse

Data Visualisation Specialist

Statistician

Mathematician

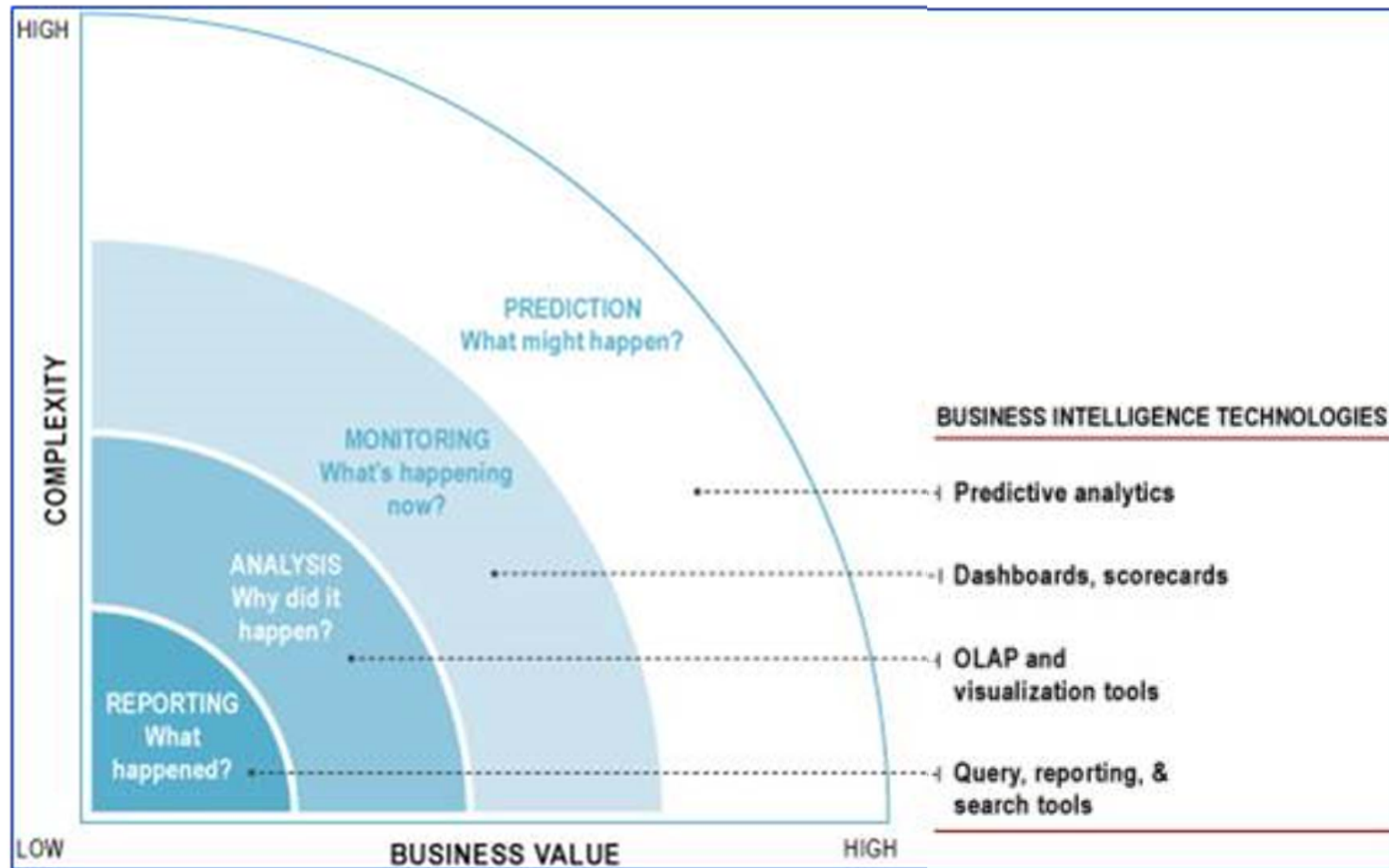
Machine Learning /AI

Software Engineer

Experience as a Consultant

Experience as a Communicator

Experience in business context





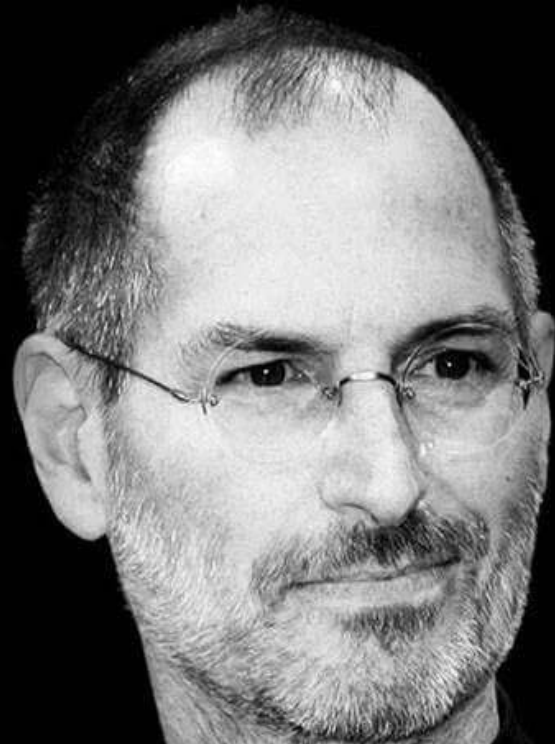
People who describe culture and values and how people behave - I've heard people refer to it as 'the soft stuff'-they often underestimate its importance. The soft stuff actually is the hard stuff.

— *Carly Fiorina* —

AZ QUOTES

Carly Fiorina – previous CEO of HP





**“It doesn't make  
sense to hire smart  
people and then tell  
them what to do;  
we hire smart people  
so they can tell us  
what to do.”**

**- Steve Jobs**

co-founder of Apple computer, co-creator  
of iTunes and the iPhone.

fb/david avocado wolfe

If your job is to tell me how  
to do my job, you should  
at least know how to do  
my job.



someecards  
user card

# Themes

- Big data + combining multiple data sources
- Unstructured data in forms of image, video, text
- Data Engineering, ETL , data 'munging'
- More use of open source tools such as R, python
- Clients wanting 'instant gratification'
- Self-service tools and dashboards
- 80% solution is OK
- Need a combination of skills – not just stats/math's but also Comp Sci and consulting skills

Questions?