

Data and analytics in production at scale

Allison Nau

Co-Founder and Director, Test Learn Iterate

Why are we here?

Digital Leaders Outperform Laggards on Three Financial Measures

Based on 2012–2014 data from 344 enterprises listed on U.S. exchanges.

3-YEAR AVERAGE GROSS MARGIN



3-YEAR AVERAGE OPERATING MARGIN



3-YEAR AVERAGE PROFIT MARGIN



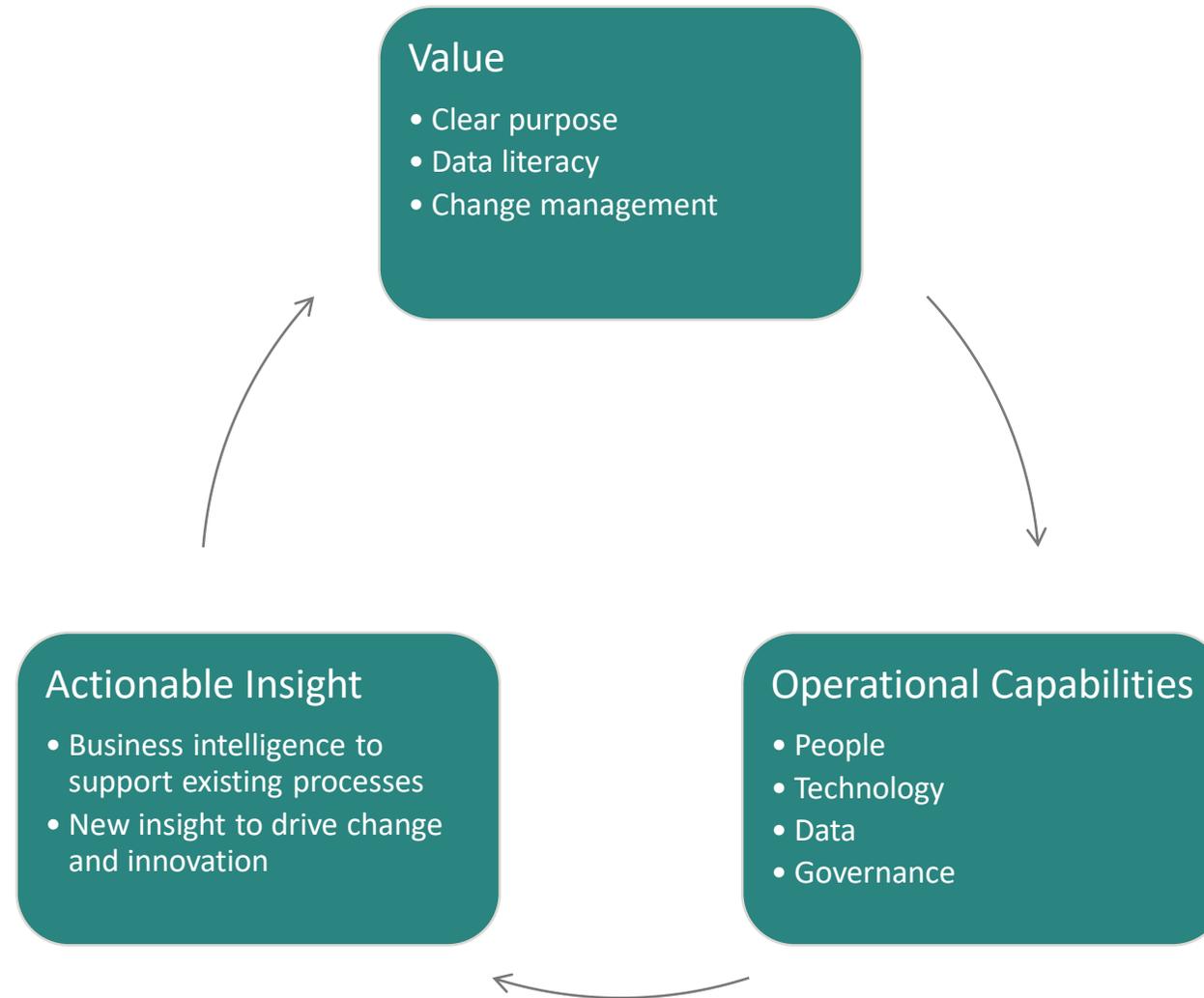
Digital leaders were defined by their ability to manage, analyse, and apply insights from the data collected.

SOURCE S&P CAPITAL IQ, KEYSTONE STRATEGY ANALYSIS

© HBR.ORG

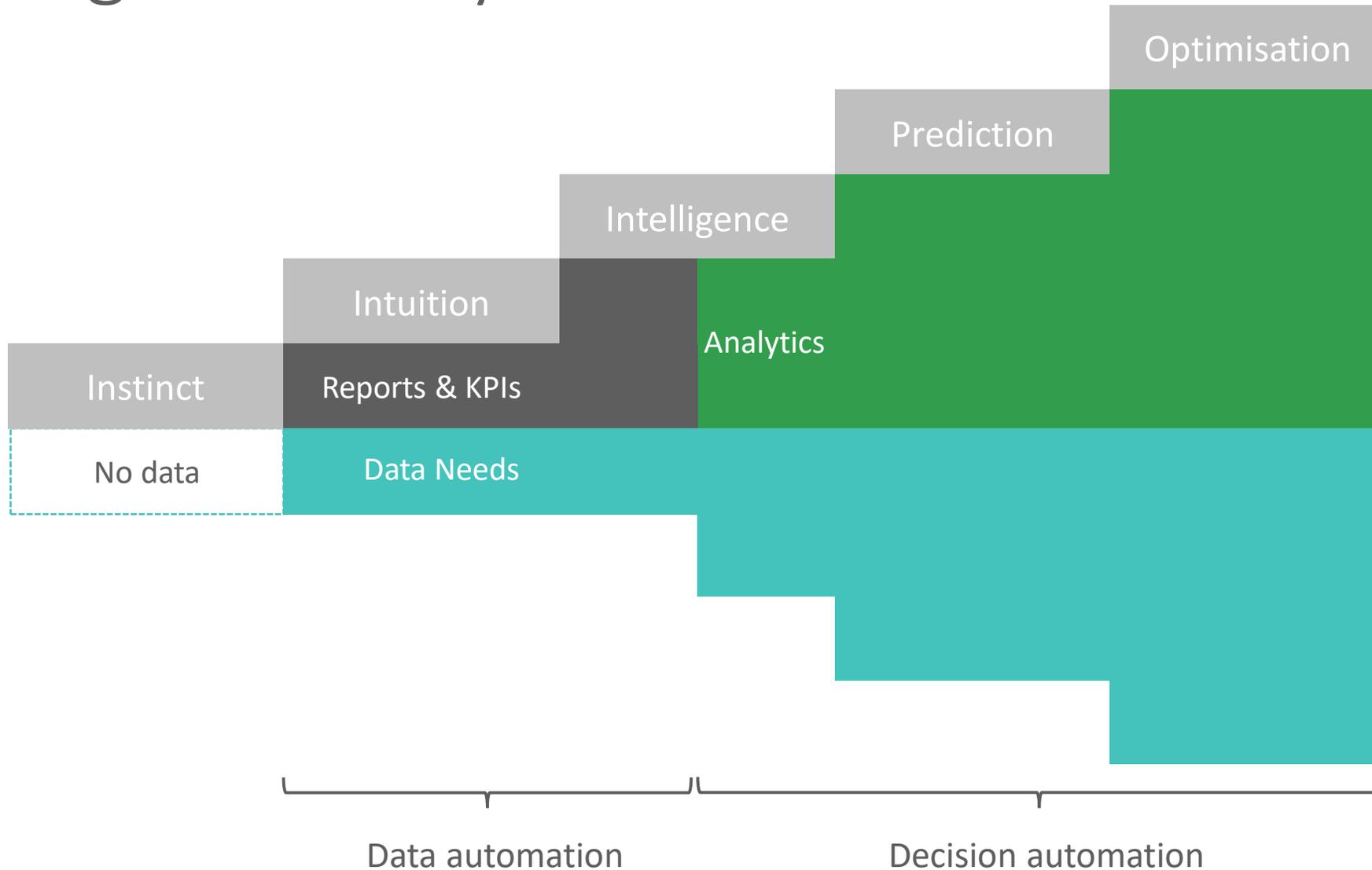
Bock, Robert, Marco Iansiti and Karim R. Lakhani. "What the companies on the right side of the digital business divide have in common", *Harvard Business Review* January 2017. <https://hbr.org/2017/01/what-the-companies-on-the-right-side-of-the-digital-business-divide-have-in-common>

How do organisations become data-driven?

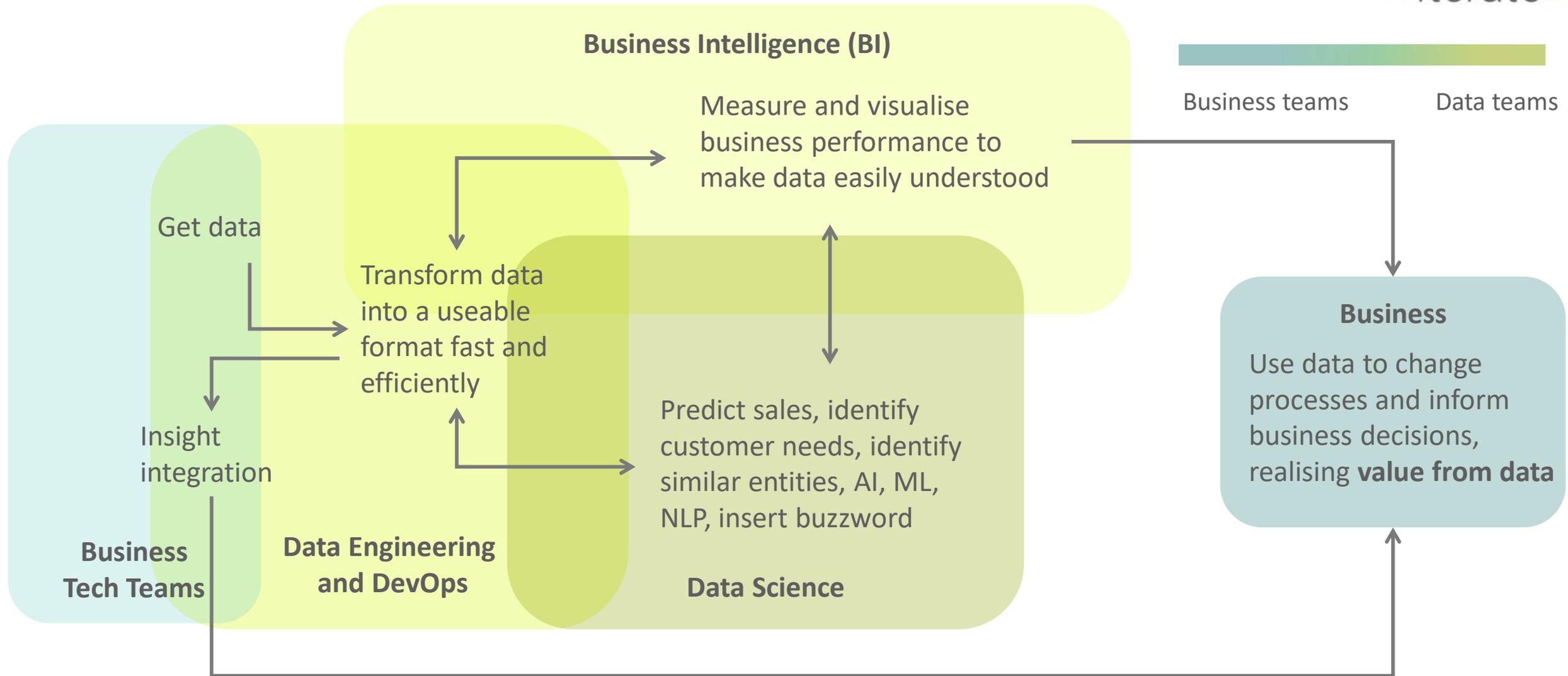


Insight Journey

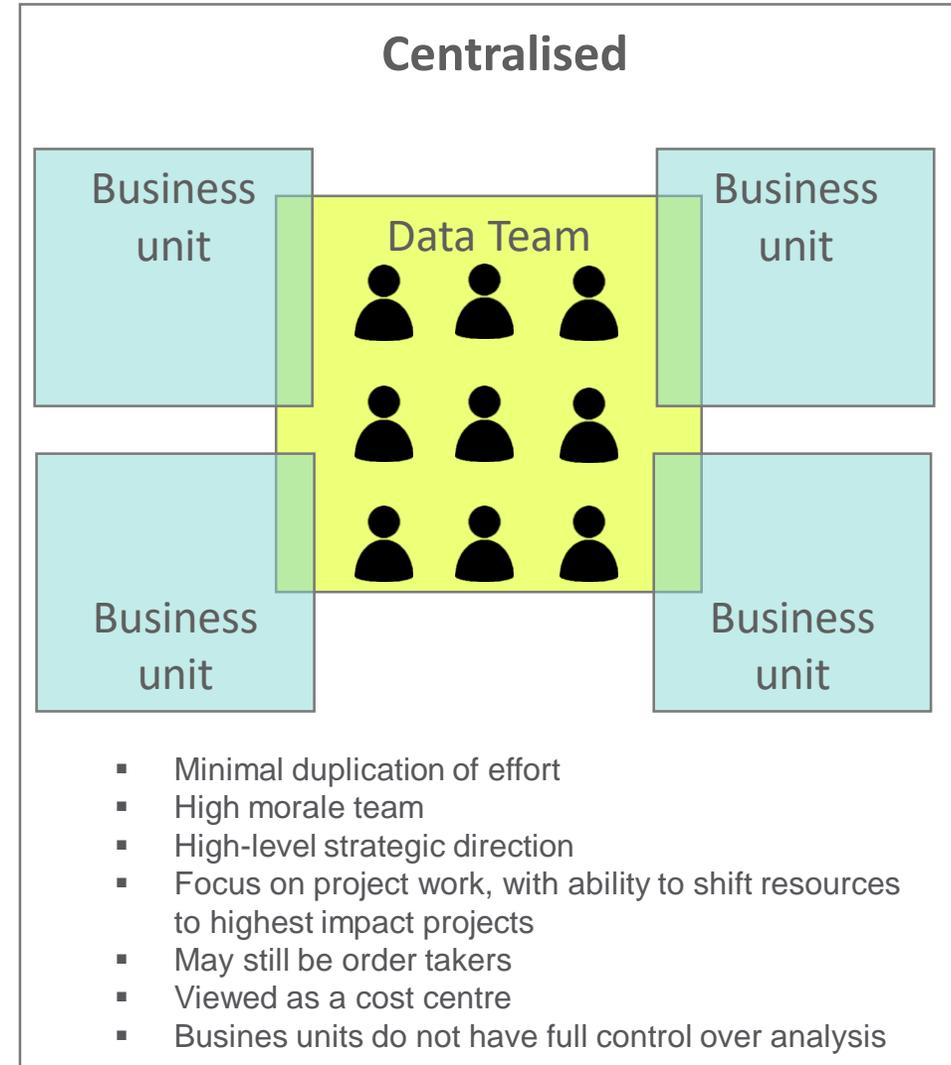
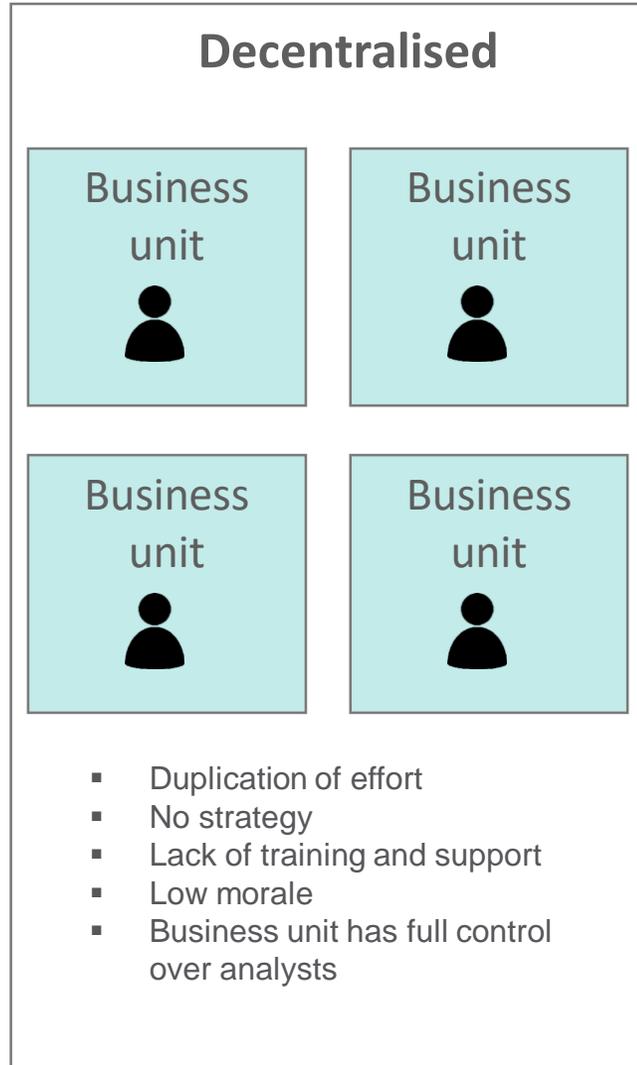
Test
Learn
Iterate



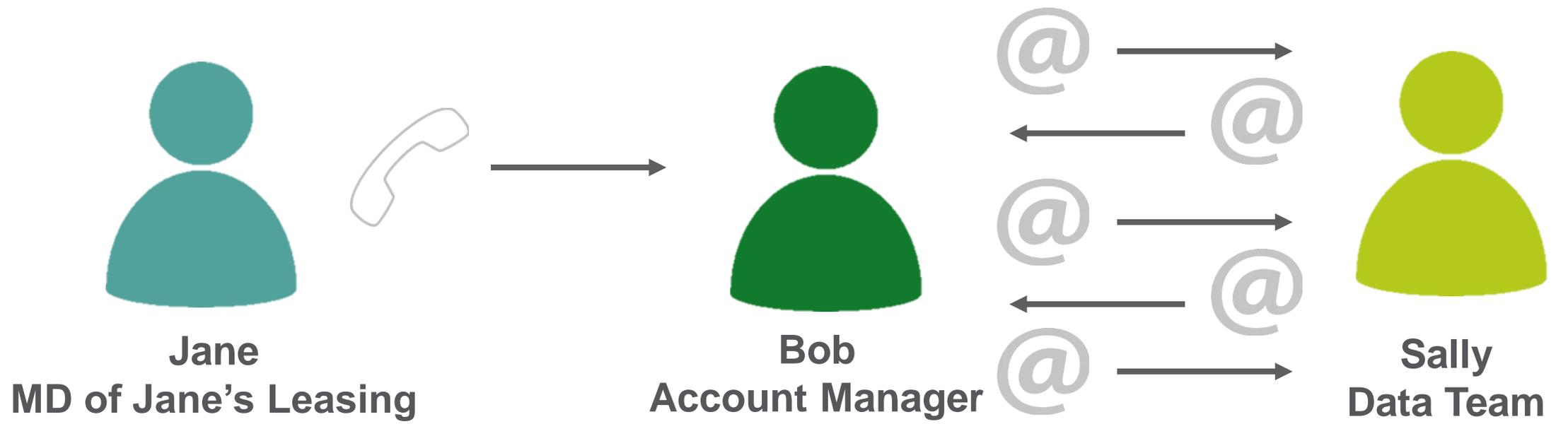
Roles for extracting value from data



Data team models



Test
Learn
Iterate





Jane is not happy with the sales performance. She **wants to know what action can be taken to ensure this doesn't happen in future**



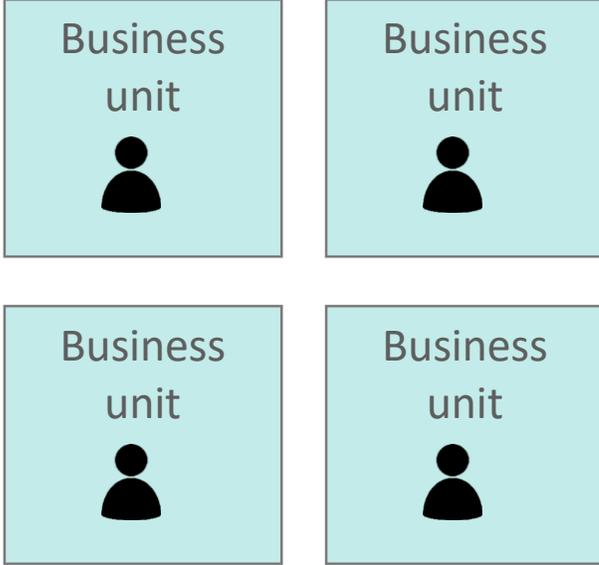
Bob is anxious about losing a customer. He **wants to tell a credible story of why things aren't going well and what actions can be taken to address it** in order to keep the customer.



Sally is frustrated because she's feels like she's wasting time on analysis that isn't used. She **wants to understand what problem Jane has, so she can solve it and craft the data story for Bob to share with Jane.**

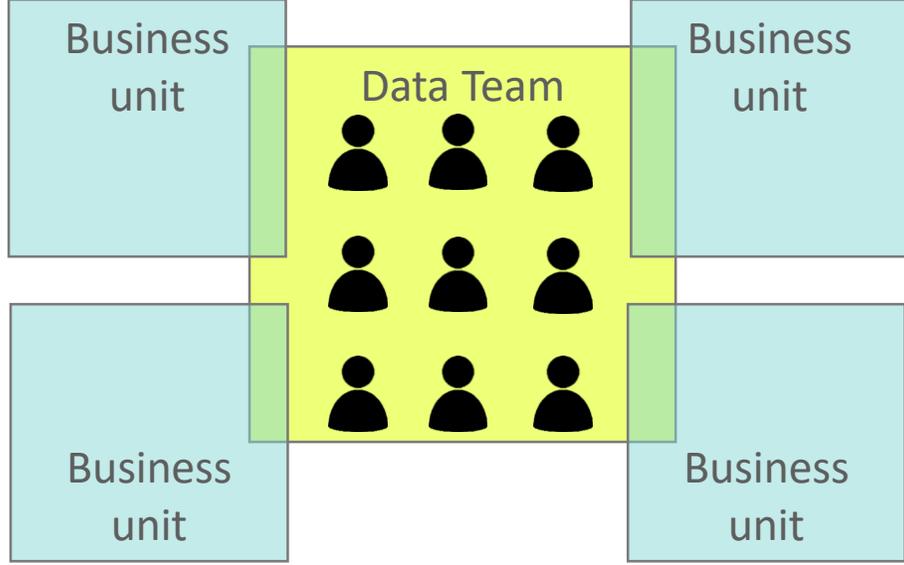
A third team model

Decentralised



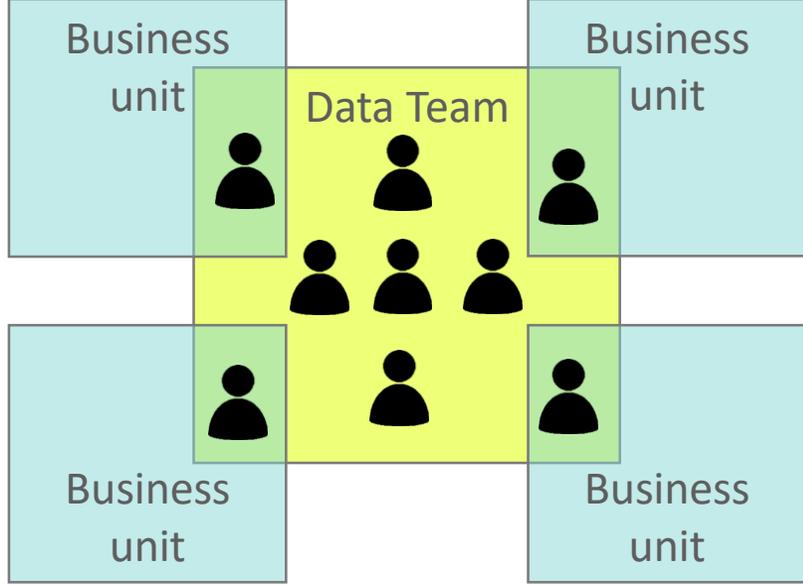
- Duplication of effort
- No strategy
- Lack of training and support
- Low morale
- Business unit has full autonomy over analysts

Centralised



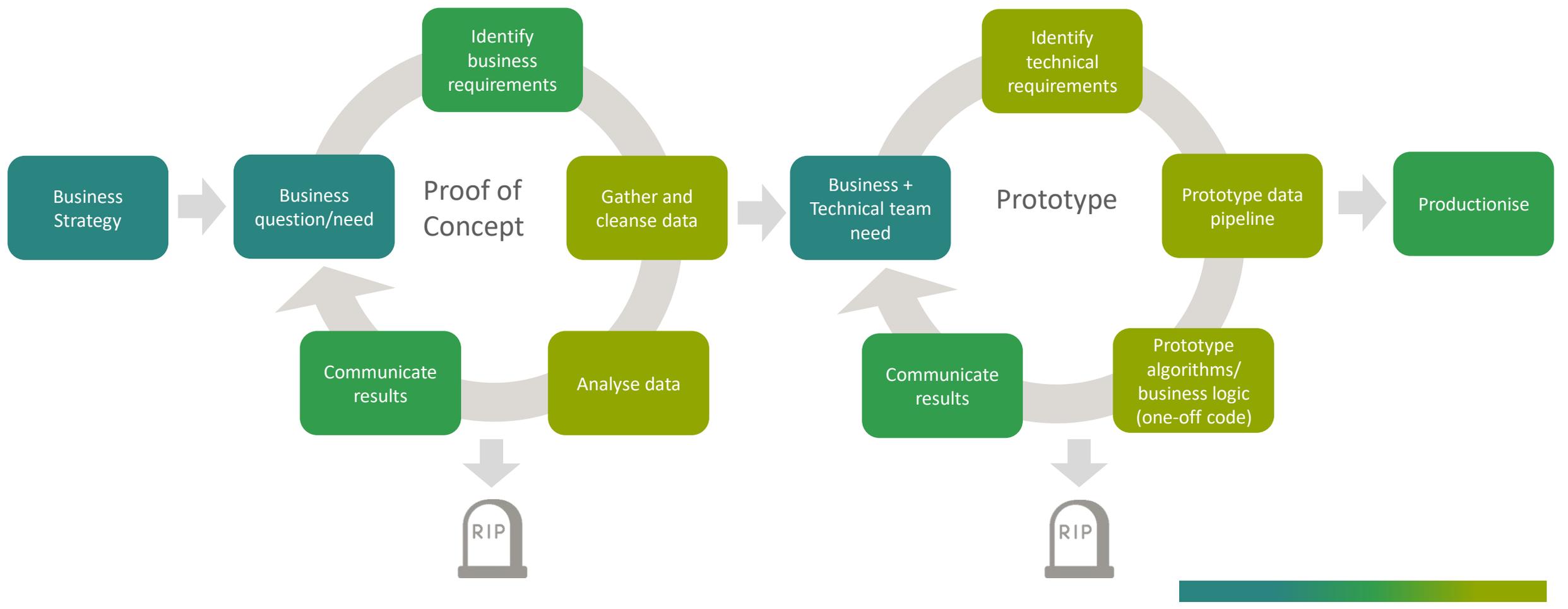
- Minimal duplication of effort
- High morale team
- High-level strategic direction
- Focus on project work, with ability to shift resources to highest impact projects
- May still be order takers
- Viewed as a cost centre

Federated



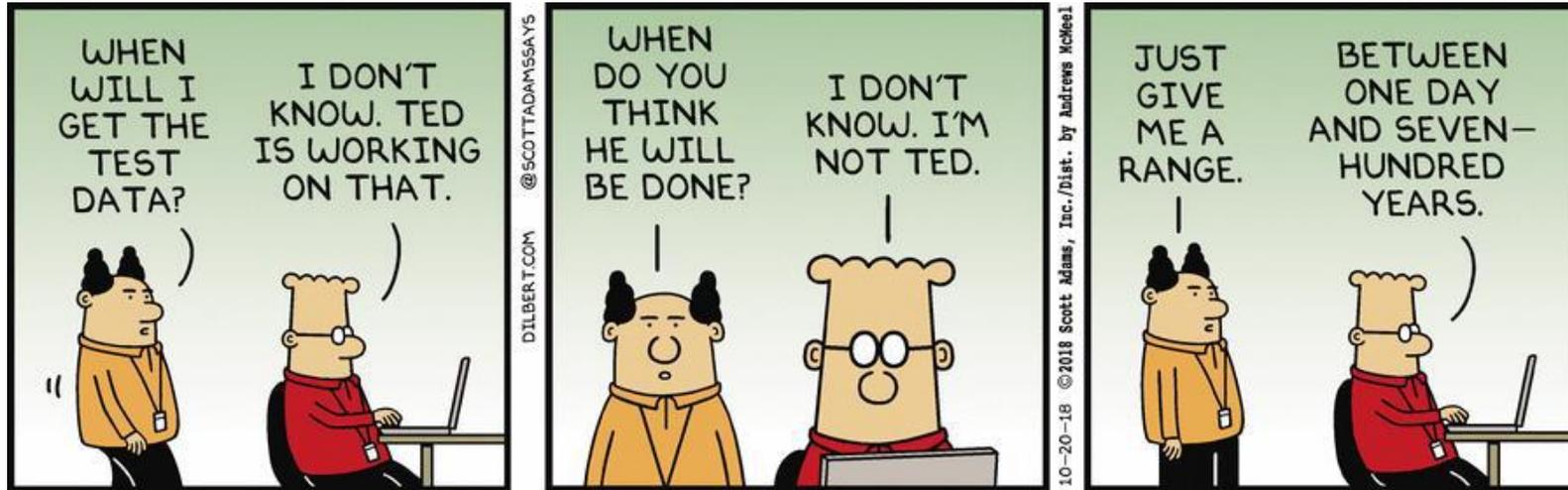
- Federated analysts supported by central team of data scientists and data engineers
- Analysts sit on management team of function
- Focus on identifying and delivering solutions to maximize value for business area
- Greater visibility to leadership on value team produces
- Level of data maturity required

An iterative approach to development



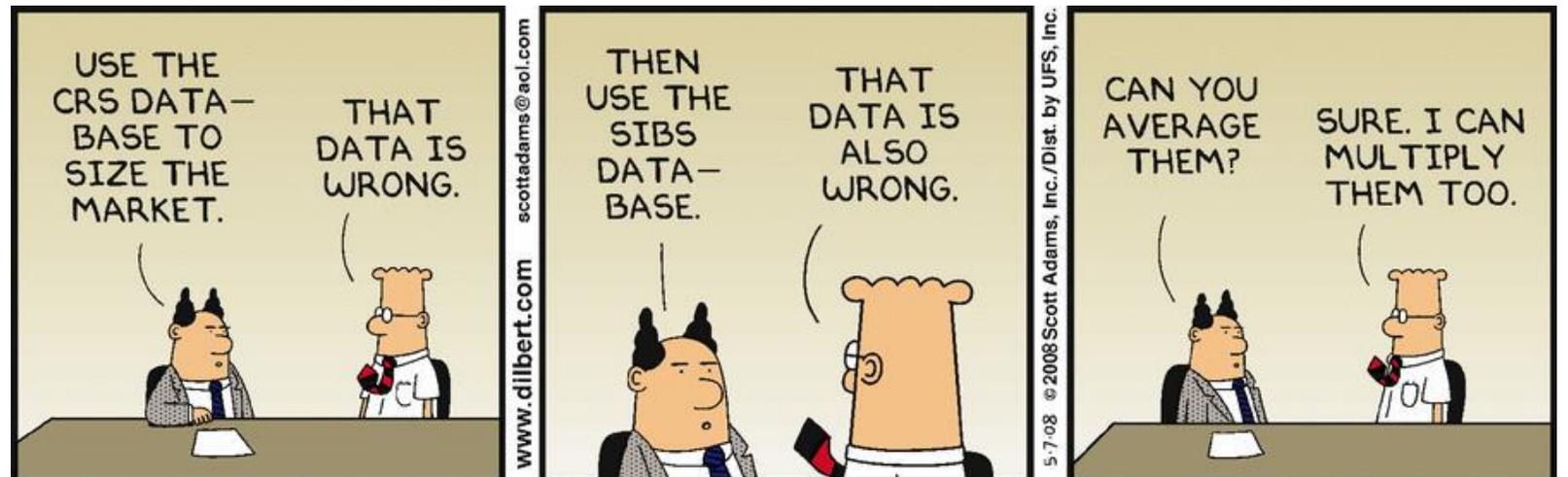
Collaborative, iterative approach helps demystify the data development process for business stakeholders

Test
Learn
Iterate



Gathering data

Cleansing data



It also allows data teams to work on more meaningful work

Test
Learn
Iterate



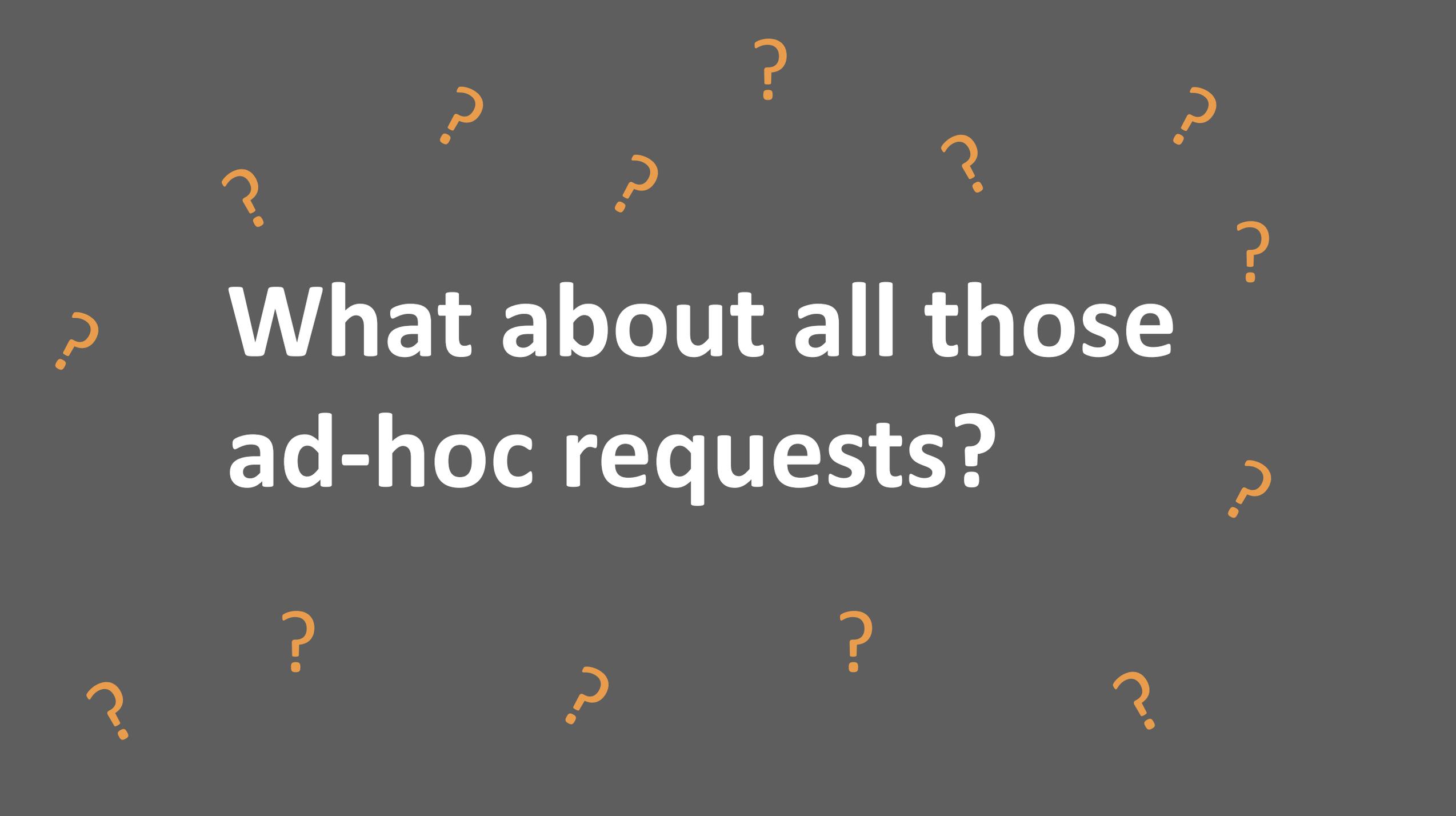
Jane is happy because she is informed with the data she needs at the time that she can make a decision to influence the outcome - and feels more in control



Bob is relieved that he has something new to tell the customer that is a differentiator from the competition, and will reduce the number of customer complaints he has to address



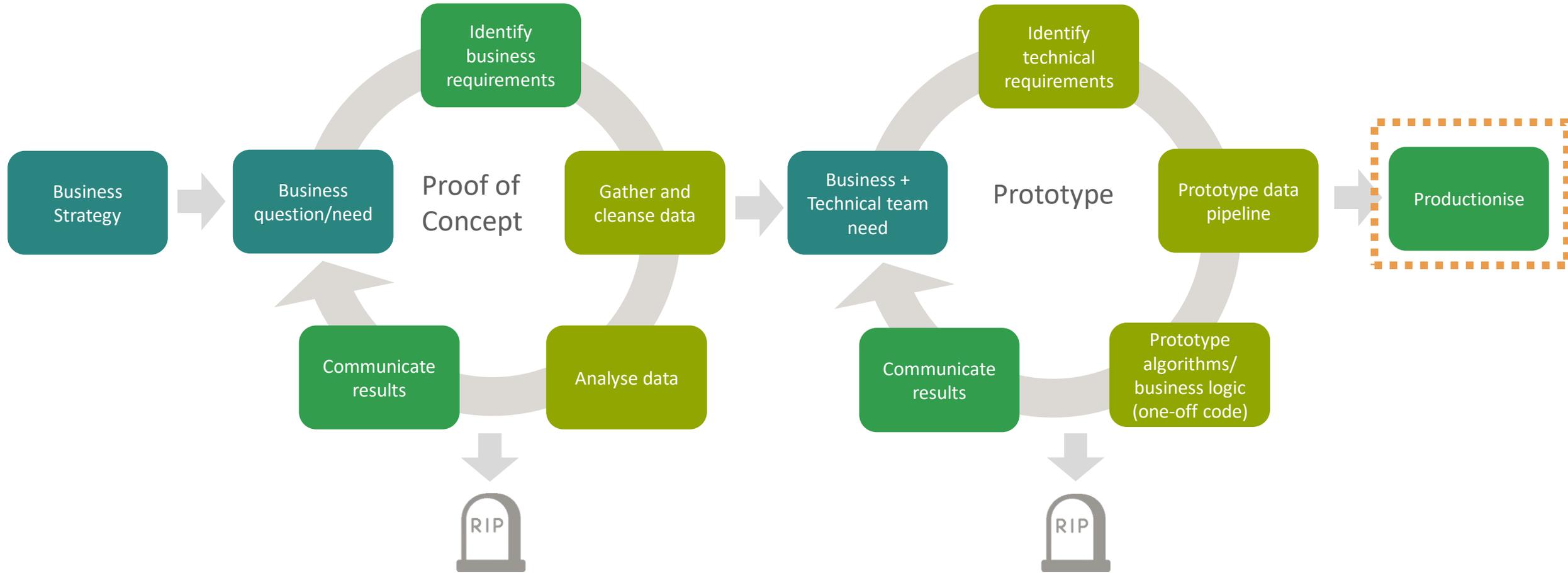
Sally identifies that Bob's concerns are similar across the sales team. Brainstorming with data colleagues, Sally realises a better solution would be to inform customers of the likelihood their cars will sell before an auction- when the customer can take action to adjust the reserve price.



**What about all those
ad-hoc requests?**

How do you put things into production?

Test
Learn
Iterate



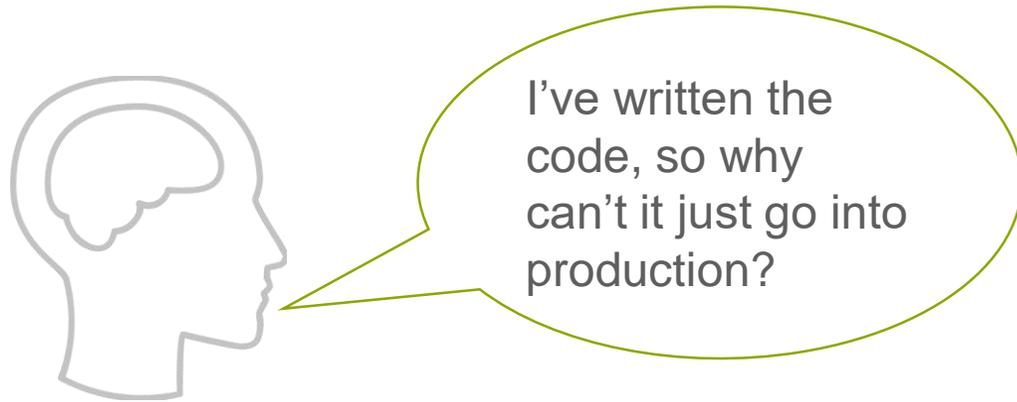
Realising value from data

Test
Learn
Iterate



Data team members care about different things

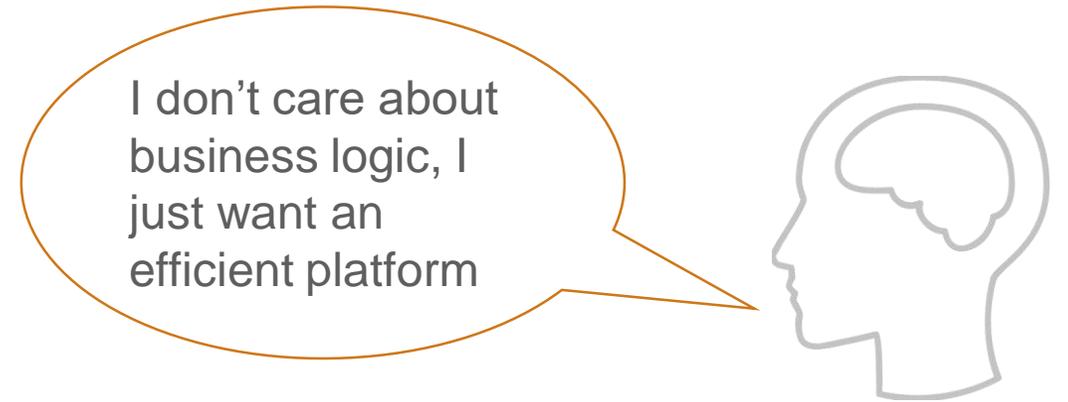
Analysts and Data Scientists



Analysts write business logic in one big messy chunk of code

- This is the fastest path to “the answer” given they are experimenting on opaque requirements
- Optimal path to execute the business logic remains obscure (and out of scope)

Data and DevOps Engineers

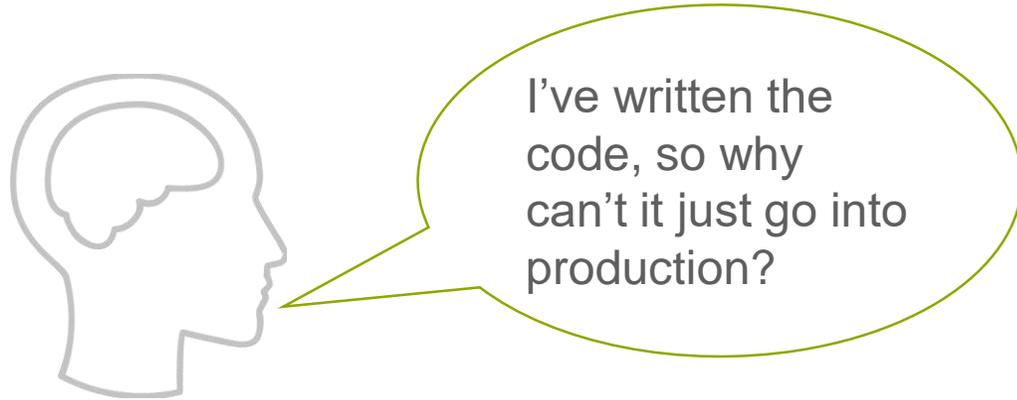


Developers want the code for each step in the flow of logic to be independent

- This allows isolation and safe (ie fully tested) improvement of less-efficient steps in the flow
- Optimal path to breaking down “big messy chunk of code” not clear without seeing business logic

Who should productionise analysts' code?

Analysts and Data Scientists

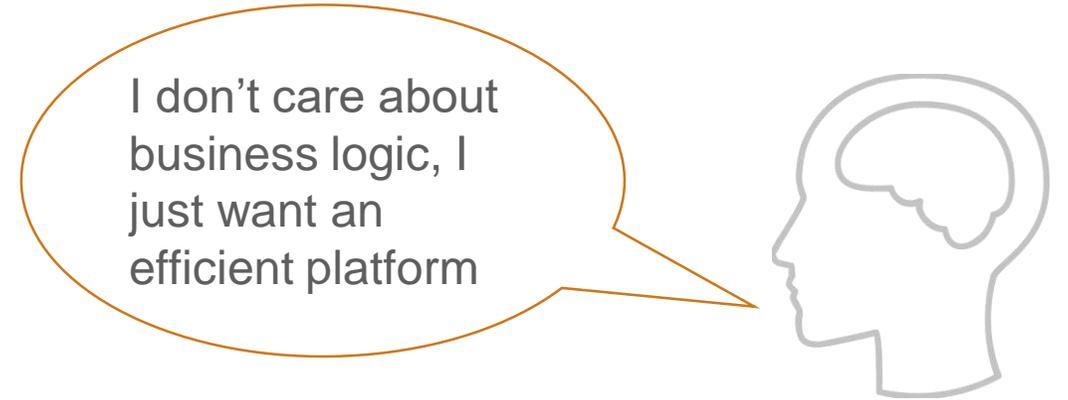


I've written the code, so why can't it just go into production?

If analysts are forced to productionise their own code:

- At best they try hard and make a mess of the “robustness” part
- At worst they are too frightened and never begin

Data and DevOps Engineers



I don't care about business logic, I just want an efficient platform

If developers are forced to productionise analysts' code:

- At best they try hard and make a mess of the “business logic” part
- At worst they are too bored and never begin

Solution: Waimak framework

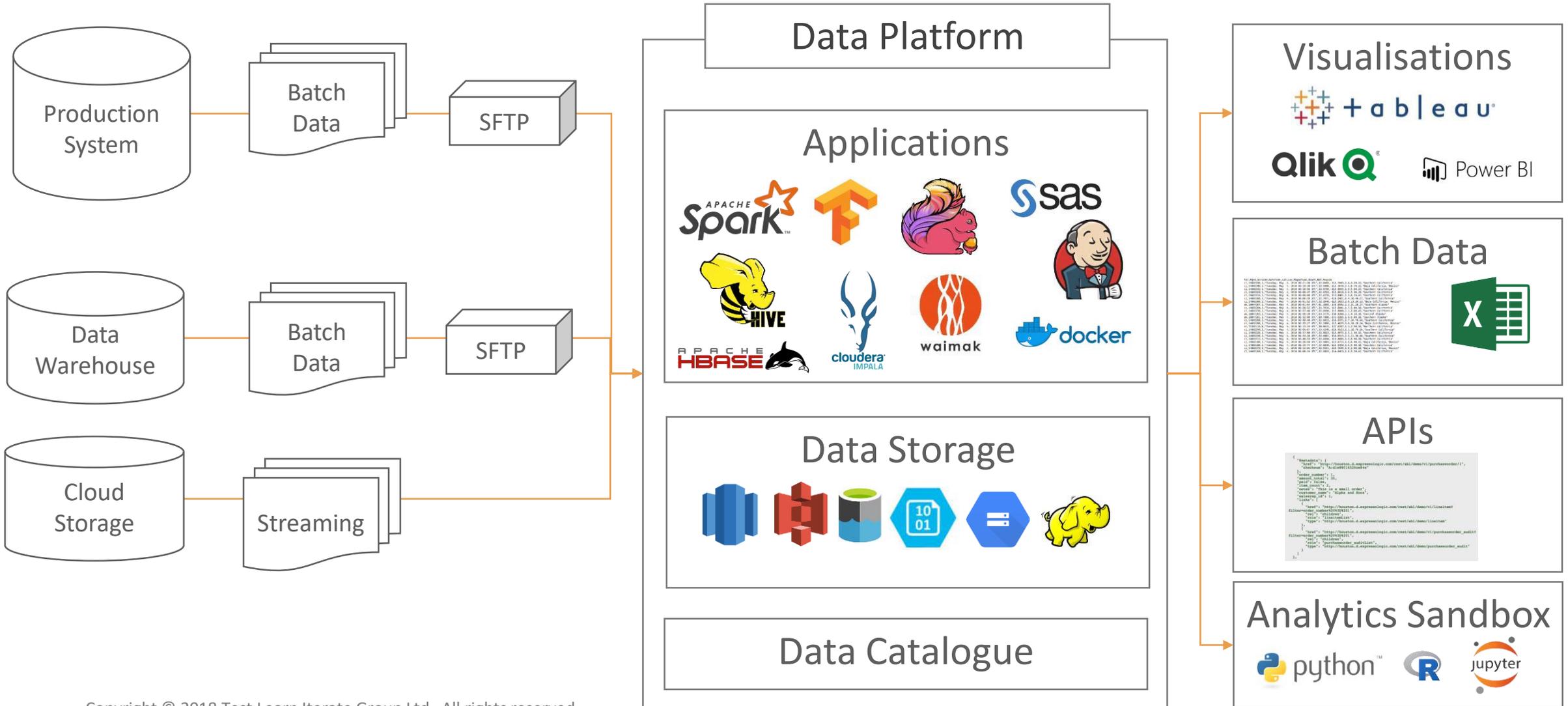
Test
Learn
Iterate



Available at <https://github.com/CoxAutomotiveDataSolutions/waimak>

Pattern for technical implementation

Test
Learn
Iterate

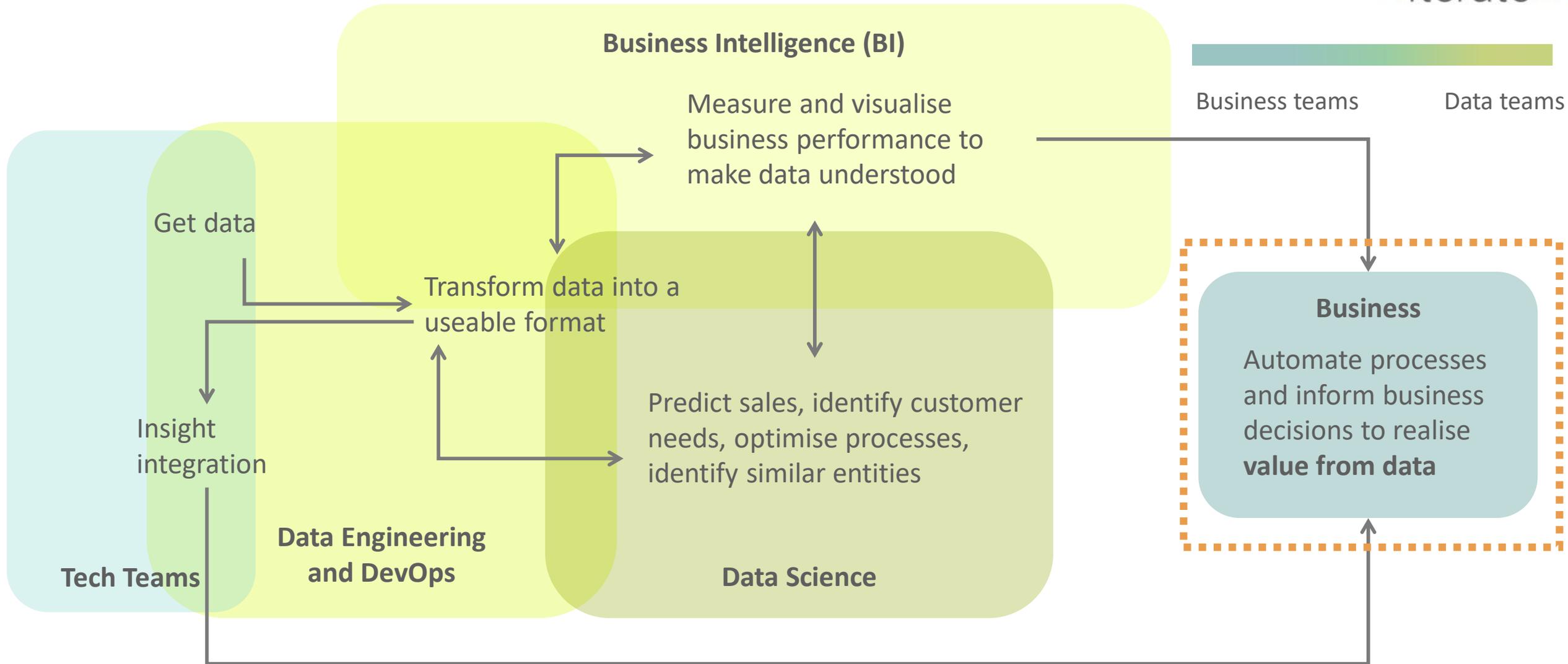


The business doesn't care
about the elegance of your
machine learning algorithm
or the efficiency of your Spark
code*

*If you work for a tech, data, or analytics company with a highly technical board/executive team, this might not be the case. But that is the exception.

How does the business realise value?

Test
Learn
Iterate

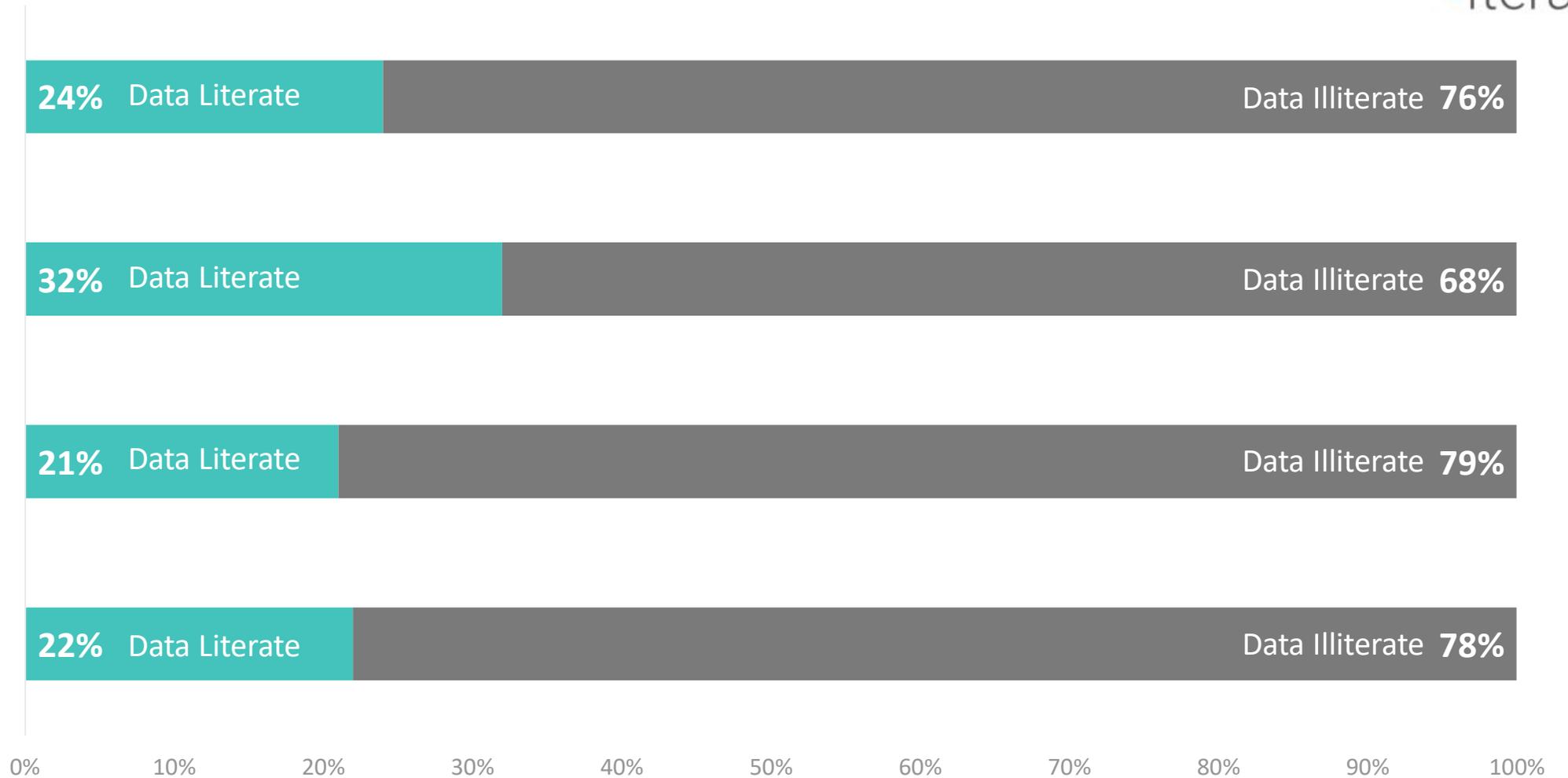


Through the people

Test
Learn
Iterate



Data literacy is a problem



• *"How to Drive Data Literacy in the Enterprise", Qlik, 2018 <https://thedataliteracyproject.org/learn> Data Literacy is defined as "fully confident in their ability to read, work with, analyze, and argue with data." Survey included 7,377 business decision-makers (junior managers and above). Respondents were from across Europe, Asia and the US.

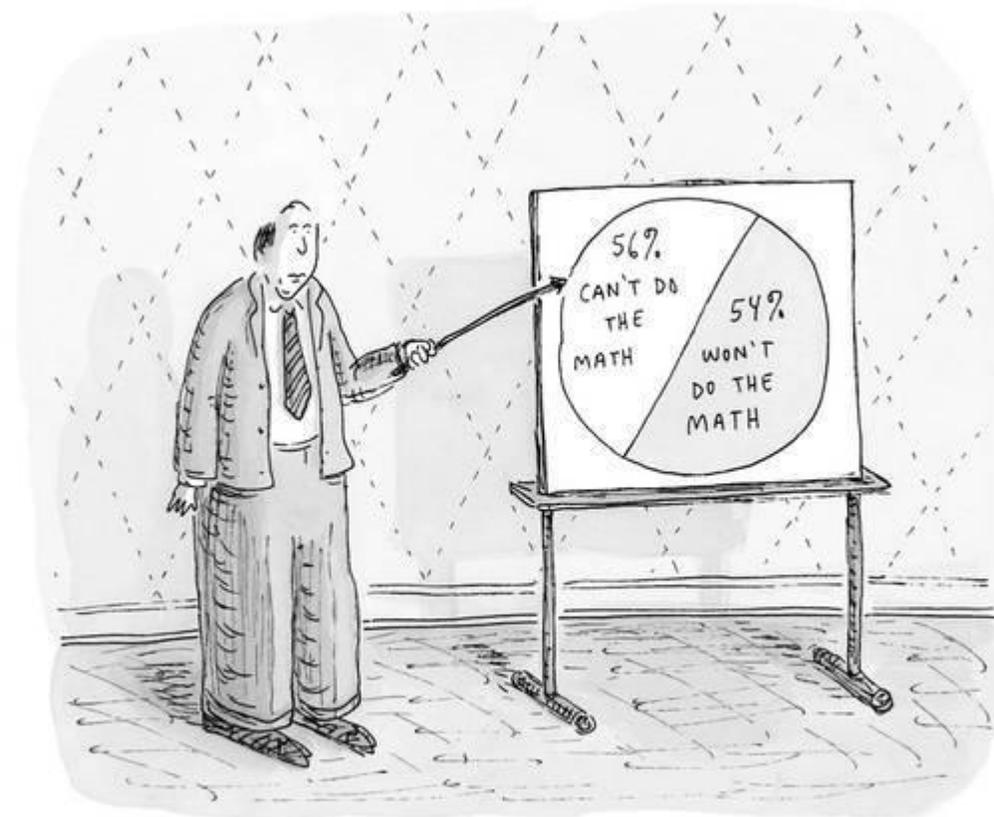
**<https://www.nationalnumeracy.org.uk/what-issue> Data Literacy is defined as GCSE maths level C or higher.

With illiteracy viewed as a badge of honour...

Test
Learn
Iterate



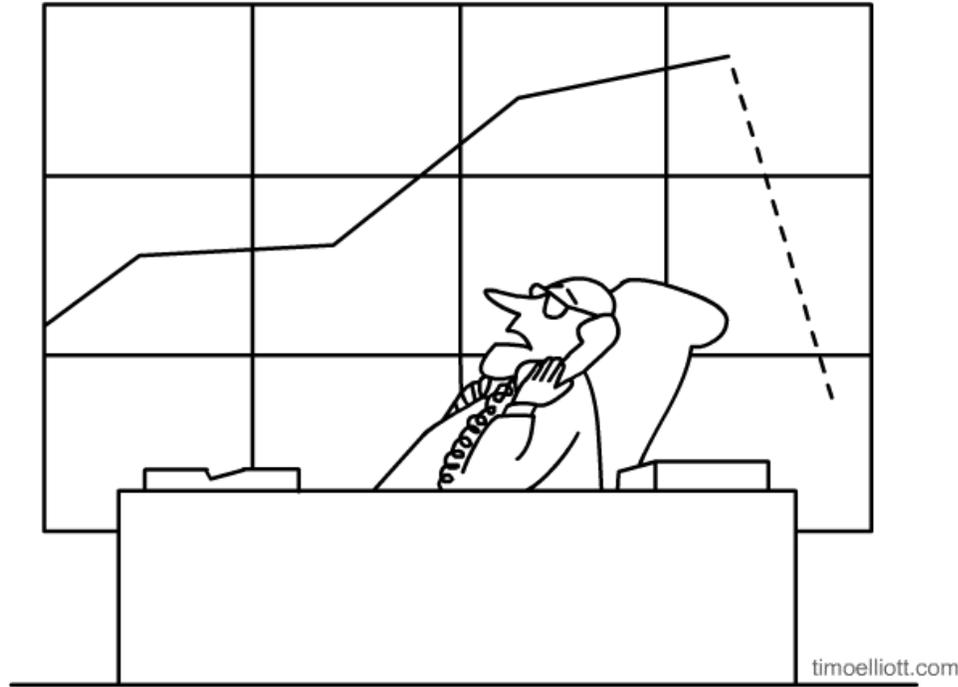
IT'S WEIRD HOW PROUD PEOPLE ARE OF NOT LEARNING MATH WHEN THE SAME ARGUMENTS APPLY TO LEARNING TO PLAY MUSIC, COOK, OR SPEAK A FOREIGN LANGUAGE.



R. Clw

...people ignore the data...

Test
Learn
Iterate



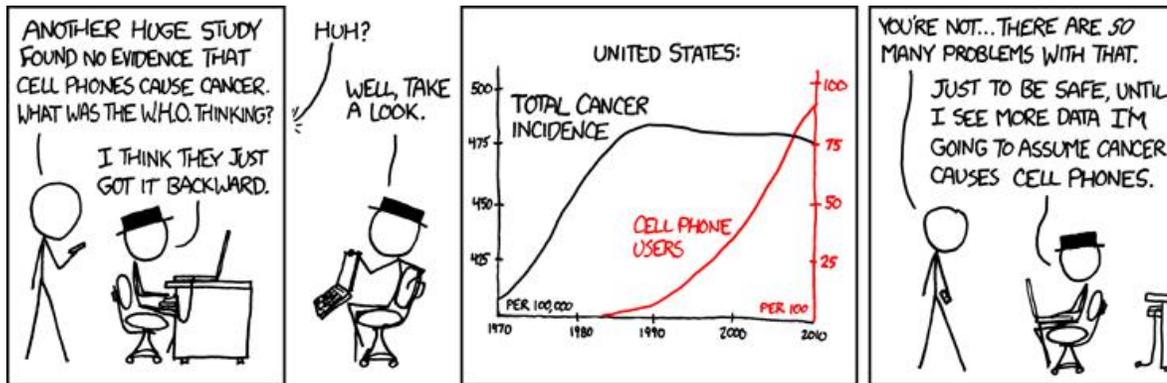
*"Yes, I have made a strategic decision.
I've decided to ignore the bad news..."*



*"You may well have data, Smithers,
but I have strong opinions,
and I pay your wages"*

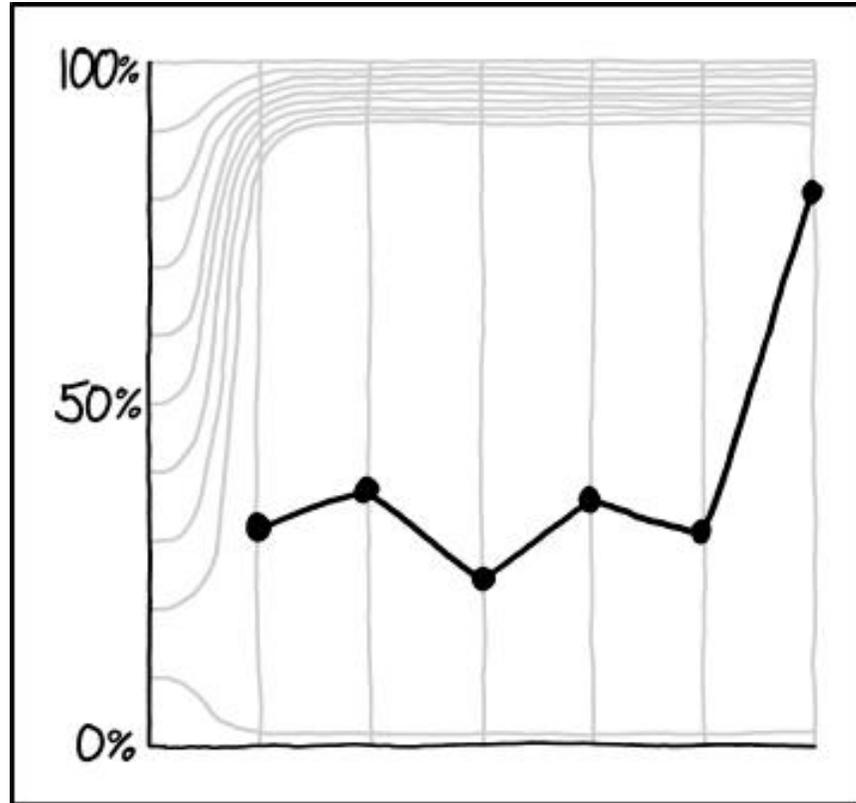
...or misinterpret it

Test
Learn
Iterate

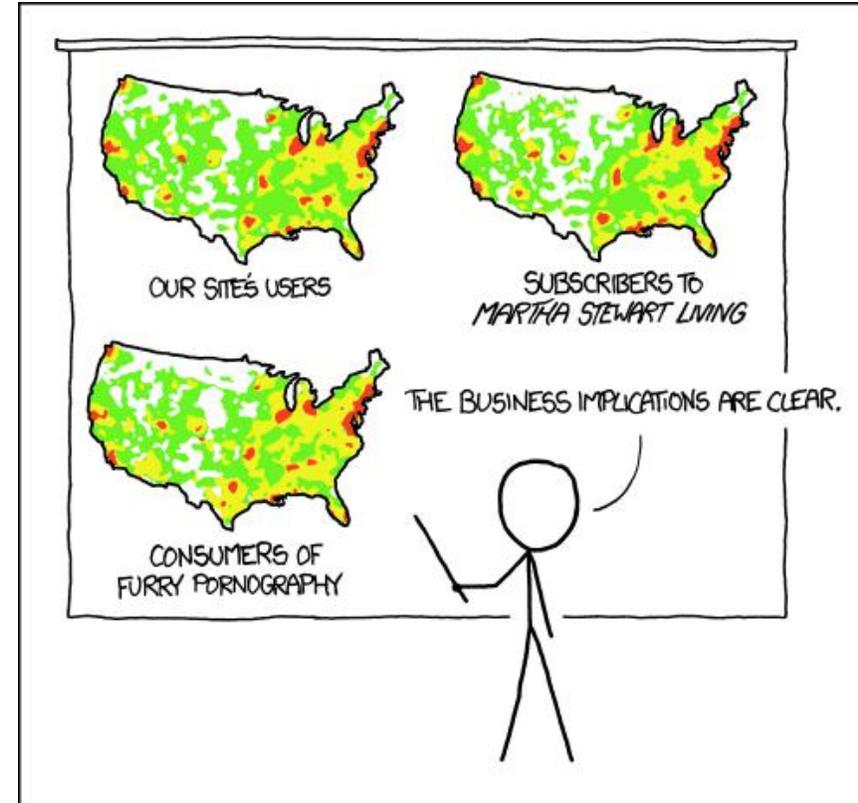


© marketoonist.com

Data teams may communicate misleading results...

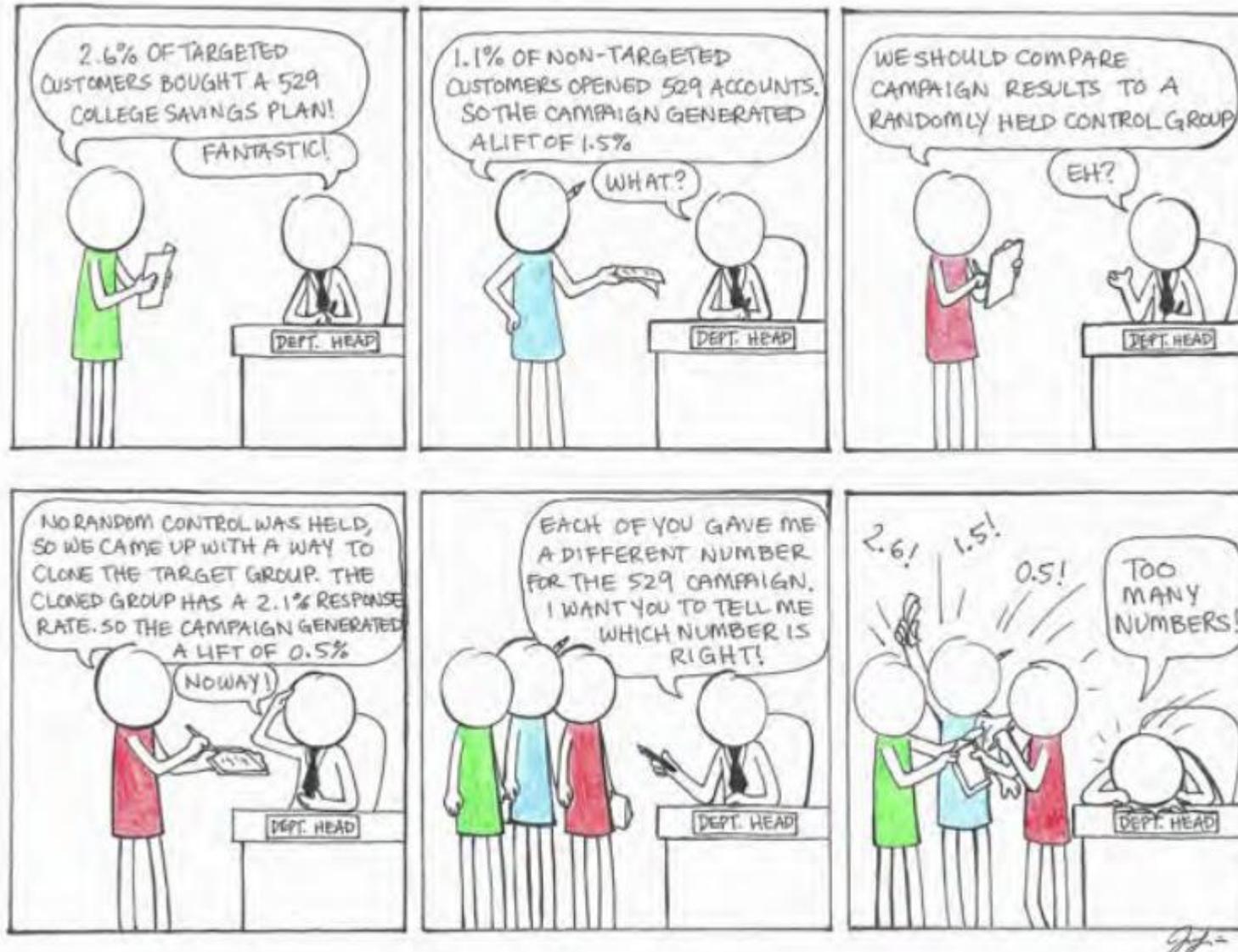


PEOPLE HAVE WISED UP TO THE "CAREFULLY CHOSEN Y-AXIS RANGE" TRICK, SO WE MISLEADING GRAPH MAKERS HAVE HAD TO GET CREATIVE.



PET PEEVE #208:
GEOGRAPHIC PROFILE MAPS WHICH ARE
BASICALLY JUST POPULATION MAPS

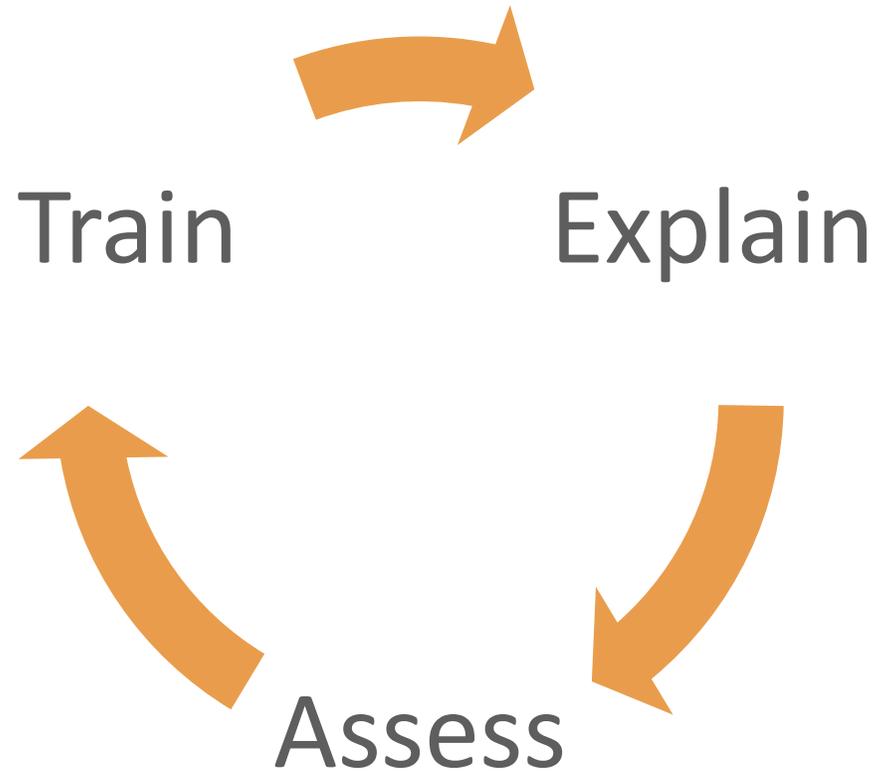
...or overwhelm with too many results



78% of business
decision makers are
willing to improve their
data literacy

How do you tackle this?

Test
Learn
Iterate



This isn't a "once and done" type of training; in order to change a culture training needs to occur frequently and be reinforced from both the top down and bottom up.

Questions/comments

Resources

- The Data Literacy Project <https://thedataliteracyproject.org/>
- National Numeracy <https://www.nationalnumeracy.org.uk/>

Allison Nau

Director, Test Learn Iterate

allison@testlearniterate.com



@allisonmnau



<https://uk.linkedin.com/in/allisonnau>