



*Data Science*

# *Preparing your business for data science*

A checklist for success

@stefflocke  
@lockedata



*Preparedness is  
about people,  
processes, and  
tools!*

# Agenda



*Data Science*

## Preparing your business for data science

A checklist for success

@stefflocke  
@lockedata

## Data science

What is data science and what's involved at a high level?

Locke Data

## People

What are the people factors you need to evaluate?

Locke Data

## Processes

What are the process factors you need to evaluate?

Locke Data

## Tools

What tools do you need to succeed?

Locke Data

## Data scientists

Locke Data

## Readiness

- People
- Processes
- Tools



Locke Data

*Steph Locke*

Locke Data

Data Science &  
DataOps focus

Microsoft MVP

Author

@stefflocke

@lockedata



# Abstract

- Data platform maturity
- Use of data products within the business
- Skillsets of existing business intelligence and other analytical teams
- Analytical teams processes and capabilities
- IT and analytical teams alignment to business goals
- Recruitment, induction, and professional development processes
- Legal, ethical, and regulatory considerations

Starting out on a project without adequate planning, risks wasted time and money when you hit unexpected roadblocks. Additionally, putting a data science project into production without sufficient testing, monitoring, and due diligence around legal obligations, can expose you to substantial problems.

Armed with the checklist, there'll be fewer "unknown unknowns" that could derail your project or cause extra cost. Let's get planning!

# *Data science*

What is data science and what's involved at a high level?



# *Data science is solving problems by building models of how things work from data*



90% of people who bought  
paddling pools went on to buy  
sandals, swimsuits, or suntan  
lotion

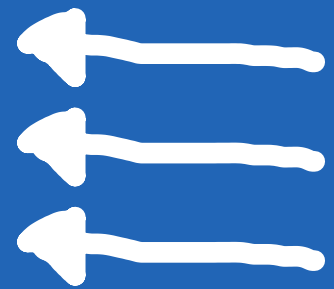


Hey, you bought a  
pool! Do you need  
these items too?



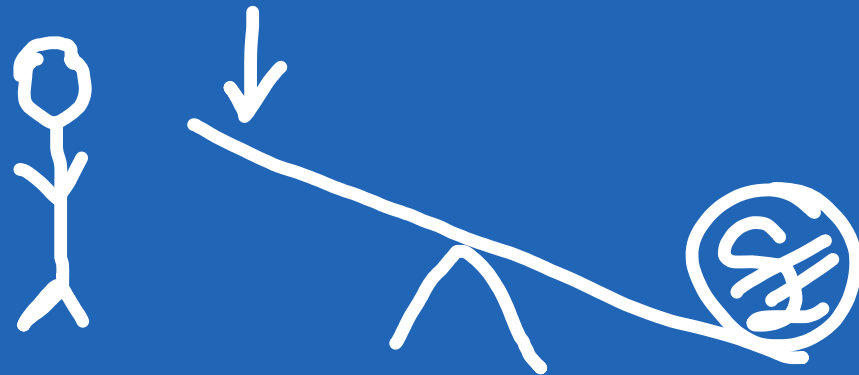
# *The problems data science should solve are business problems*

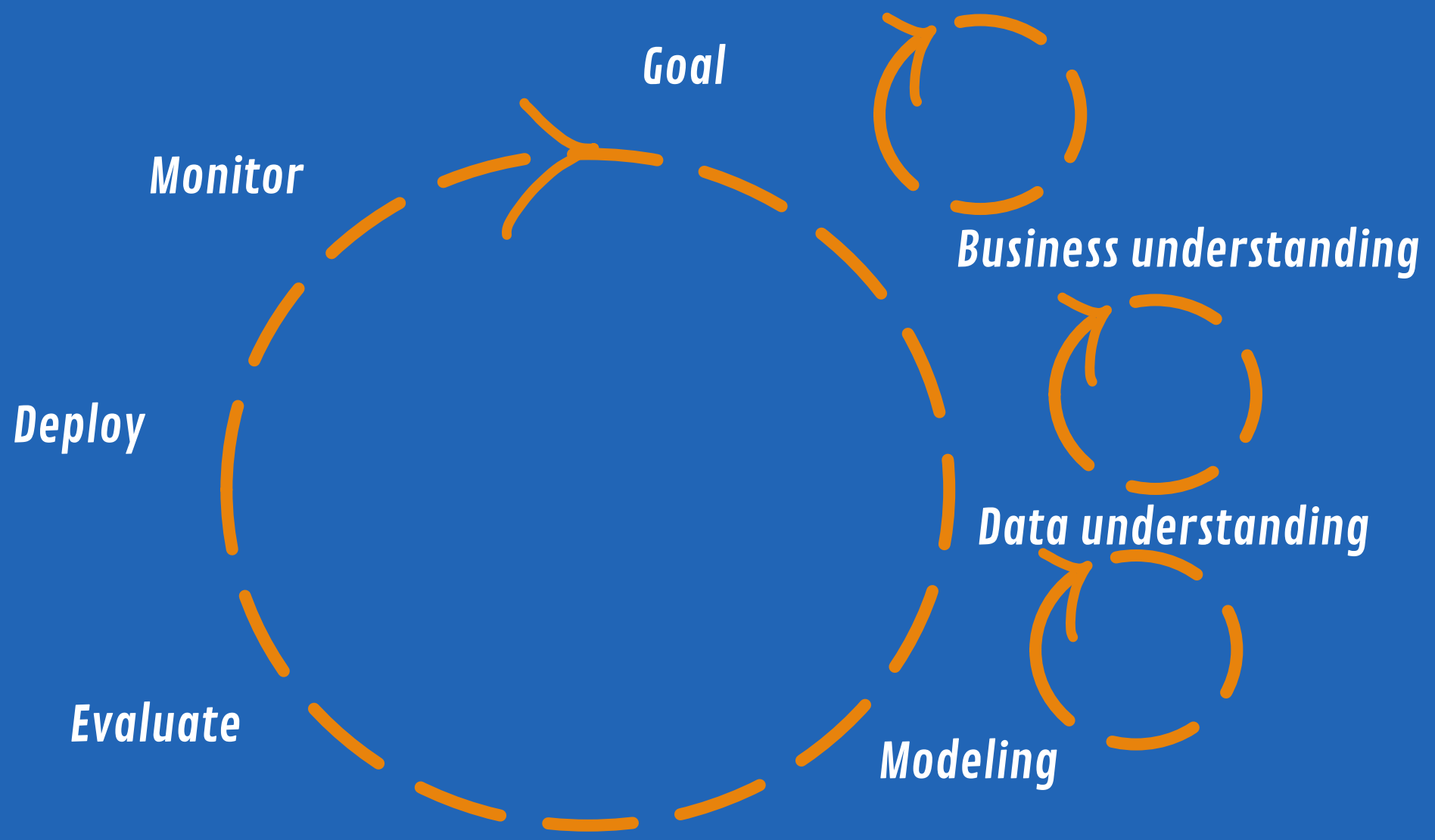
Increase market share  
Increase profits  
Decrease overheads



Improve marketing?  
Improve upselling?  
Reduce waste?

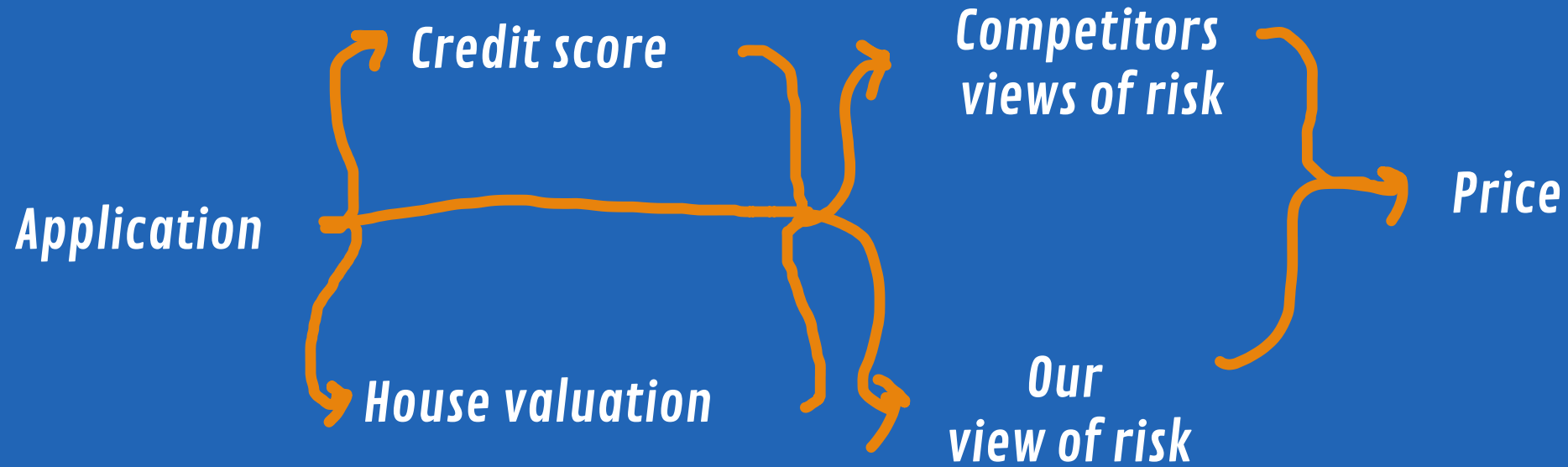
*Data scientists aim to change behaviours in order to achieve the business goal*



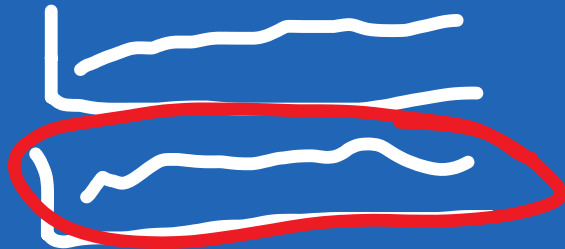
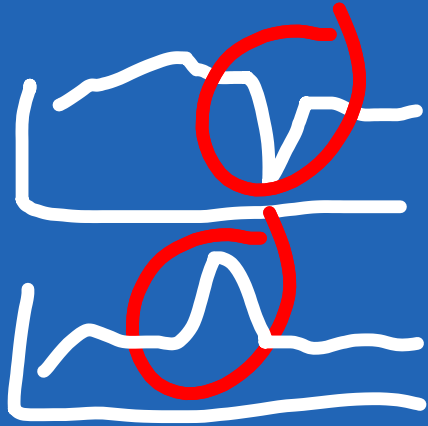
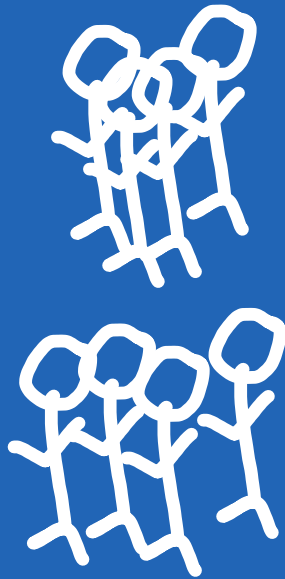


# *Examples*

# Loan pricing



# Risky behaviour online



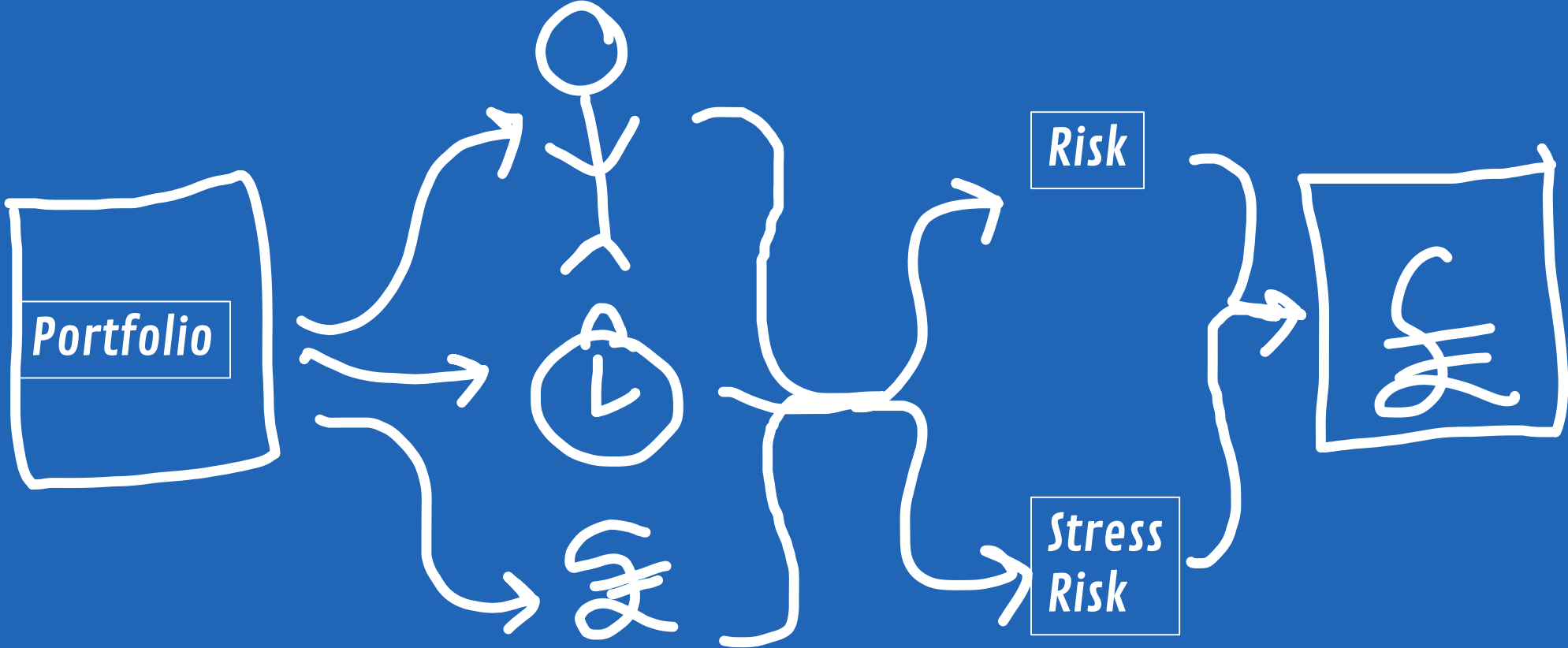
Productivity app  
traffic dropped  
off

Big spike in  
traffic to  
malware by Jim



Bob's web traffic  
doesn't look like  
everyone else's in  
the dev team.

# BASEL II

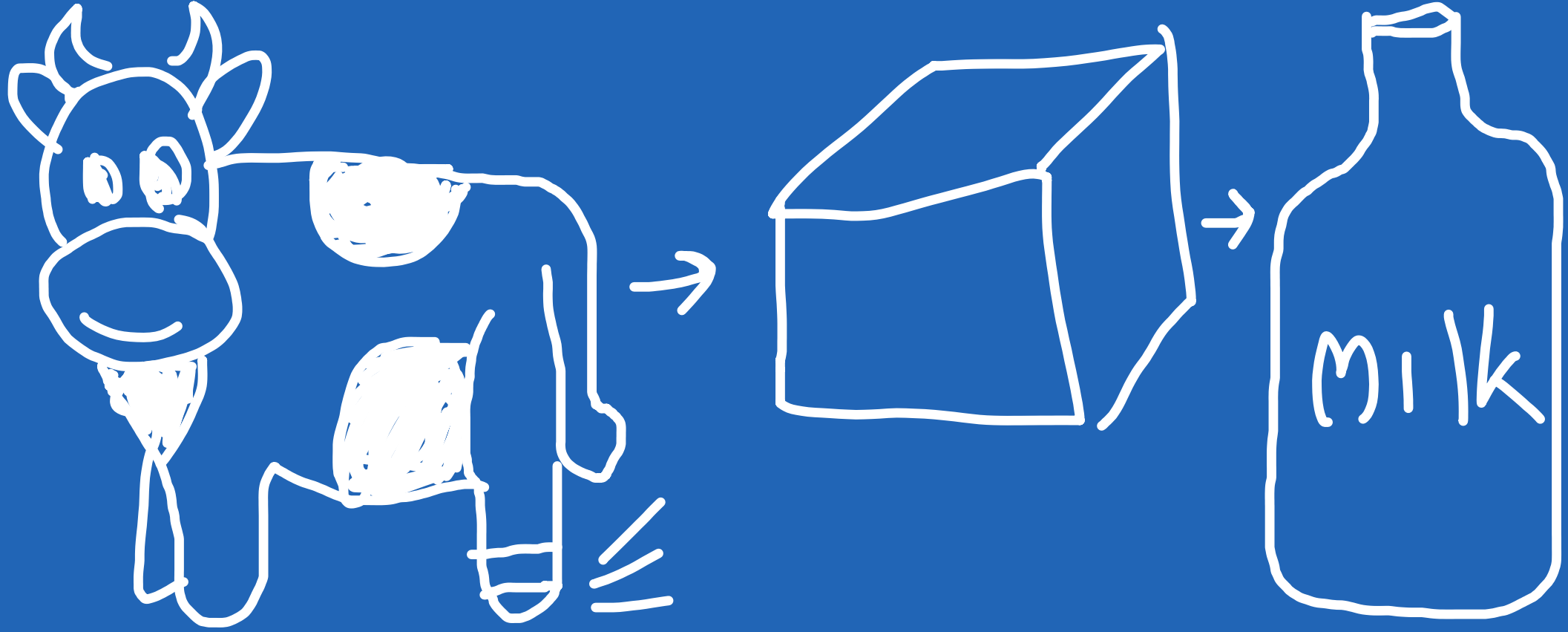


# *Skype's Babel Fish*





IoC



# *People*

What are the people factors you need to evaluate?

# Areas

- Skillsets of existing business intelligence and other analytical teams
- Recruitment, induction, and professional development processes
- Legal, ethical, and regulatory considerations



# *Processes*

What are the process factors you need to evaluate?

# Areas

- Analytical teams processes and capabilities
- IT and analytical teams alignment to business goals
- Legal, ethical, and regulatory considerations



# *Tools*

What tools do you need to succeed?

# Areas

- Data platform maturity
- Use of data products within the business



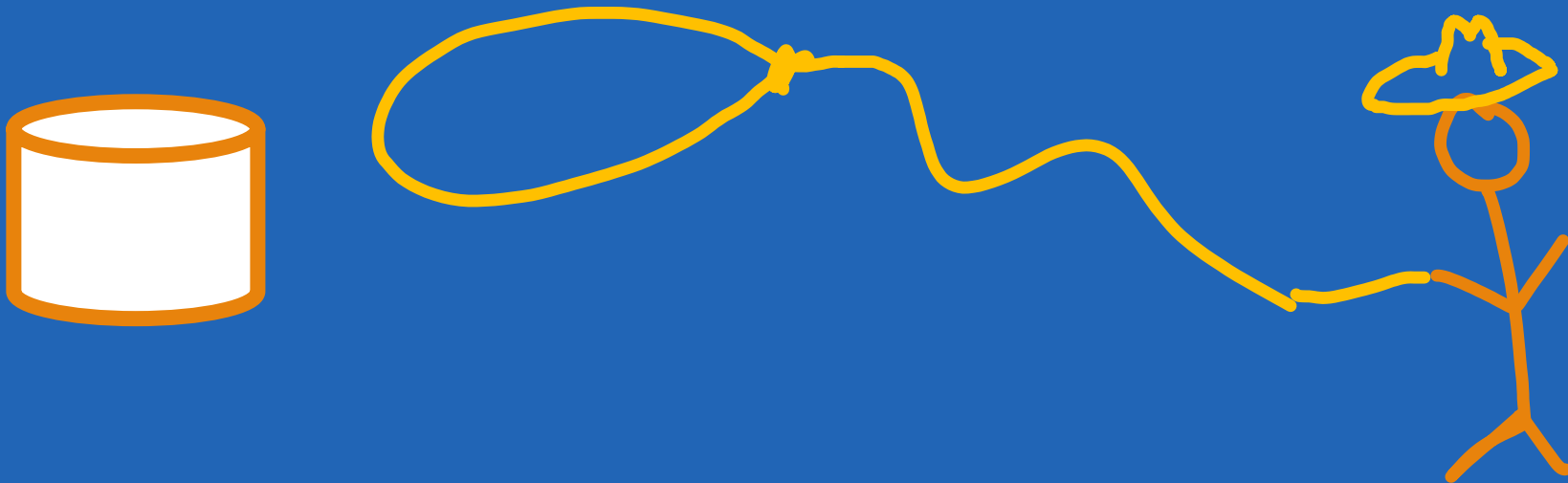
***Data scientists***



*A data scientist is a team*



# *Wrangling skills*



# *Modeling skills*



# Communication skills



# *Coding skills*

SQL

Python

Scala

C#



***Advice***

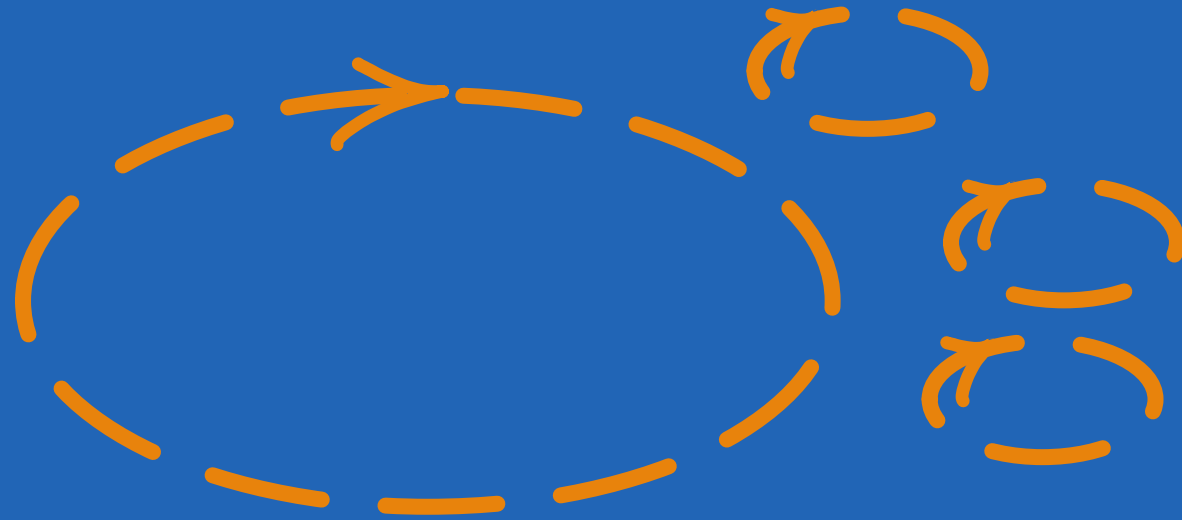
*Start big but not too big*

Impact

OK	Bad	Bad
Great	OK	Bad
Bad	Great	OK

Importance

# *The C-suite needs to understand the process*





# *Do projects in a range of silos*

*Sales* ✓

*Customer Services* ✓

*Finance* ✓

*Logistics* ✓

*Ops* ✓

*Marketing* ✓

*Get something good in production before  
worrying about great*

75% accuracy  
Time taken: 1  
week



78% accuracy  
Time taken: 6  
weeks

# *You can't have good data science with bad data*

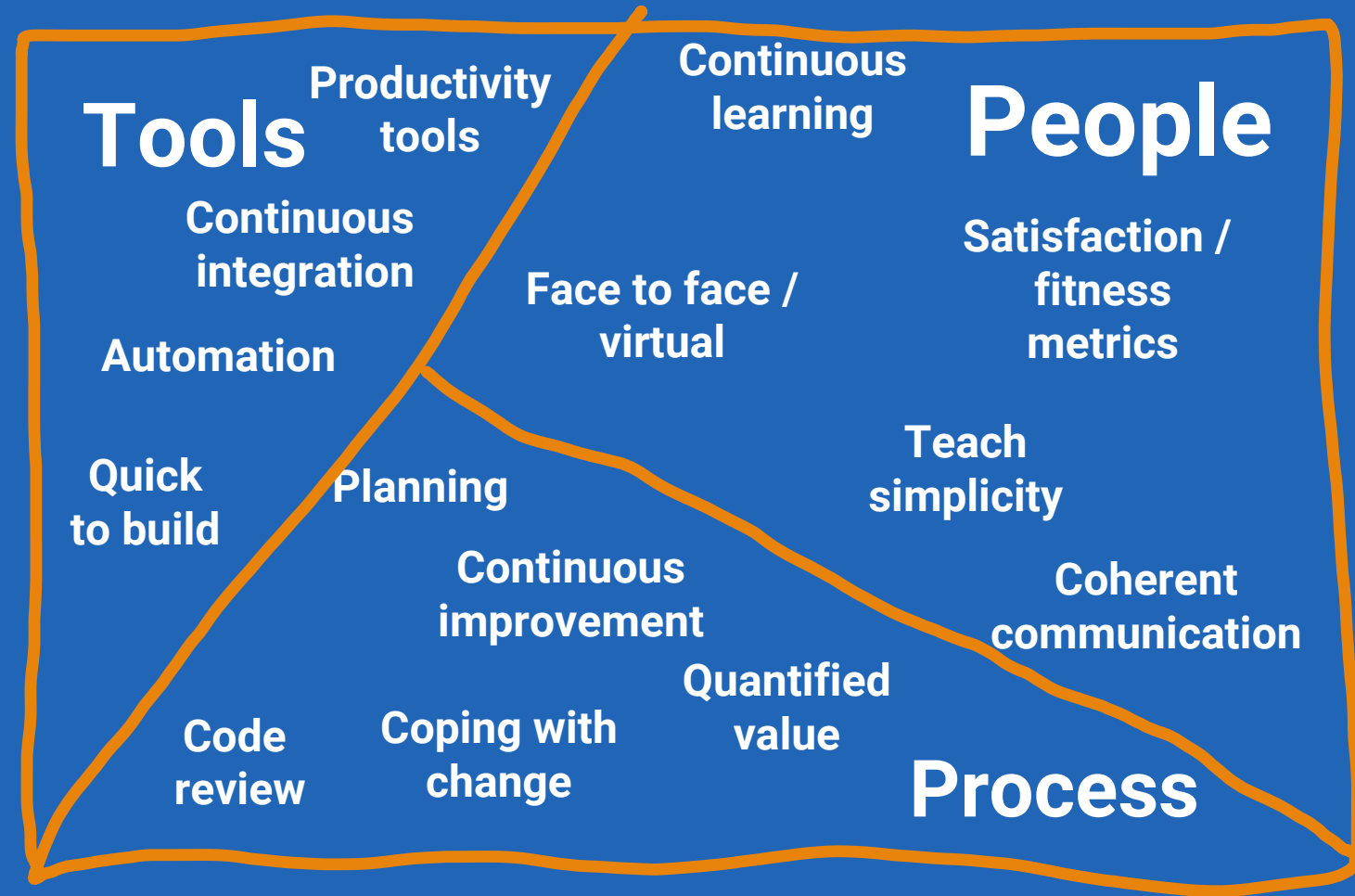


Hey, you  
turned on  
the lights.  
Should I  
start the  
lawnmower?

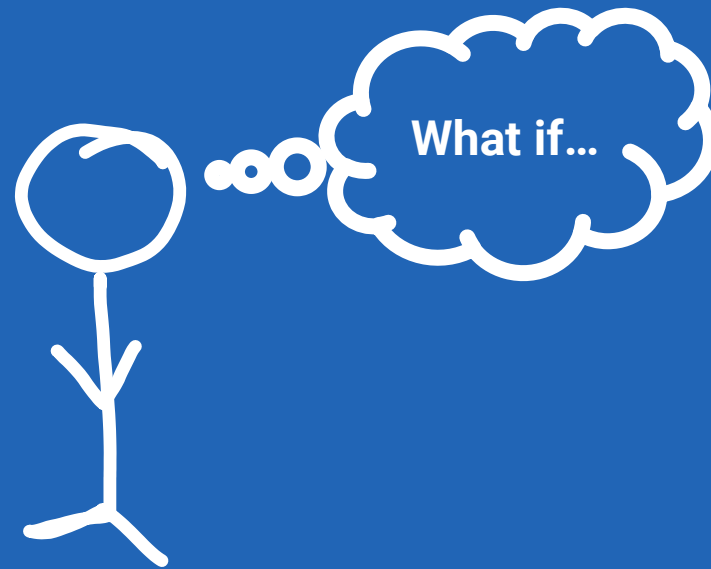
*You don't can't have fast data science  
without DataOps*



# DataOps = Agility




# *Research time required*



# Readiness

- People
- Processes
- Tools



A cartoon character with short, bright red hair, large green eyes, and a wide, open-mouthed smile. The character is wearing a white lab coat over a light blue shirt. The character is positioned on the left side of the image, with a white speech bubble extending from their mouth towards the right.

*Sign up at  
[bit.ly/lockesubscribe](https://bit.ly/lockesubscribe)  
and I'll send you the  
checklist*



Psst ... follow me on  
twitter @stefflocke  
@lockedata and don't  
forget to grab stickers  
and business cards on  
your way out!

***Thank you!***

Thank you to the organisers, sponsors, and volunteers too!