

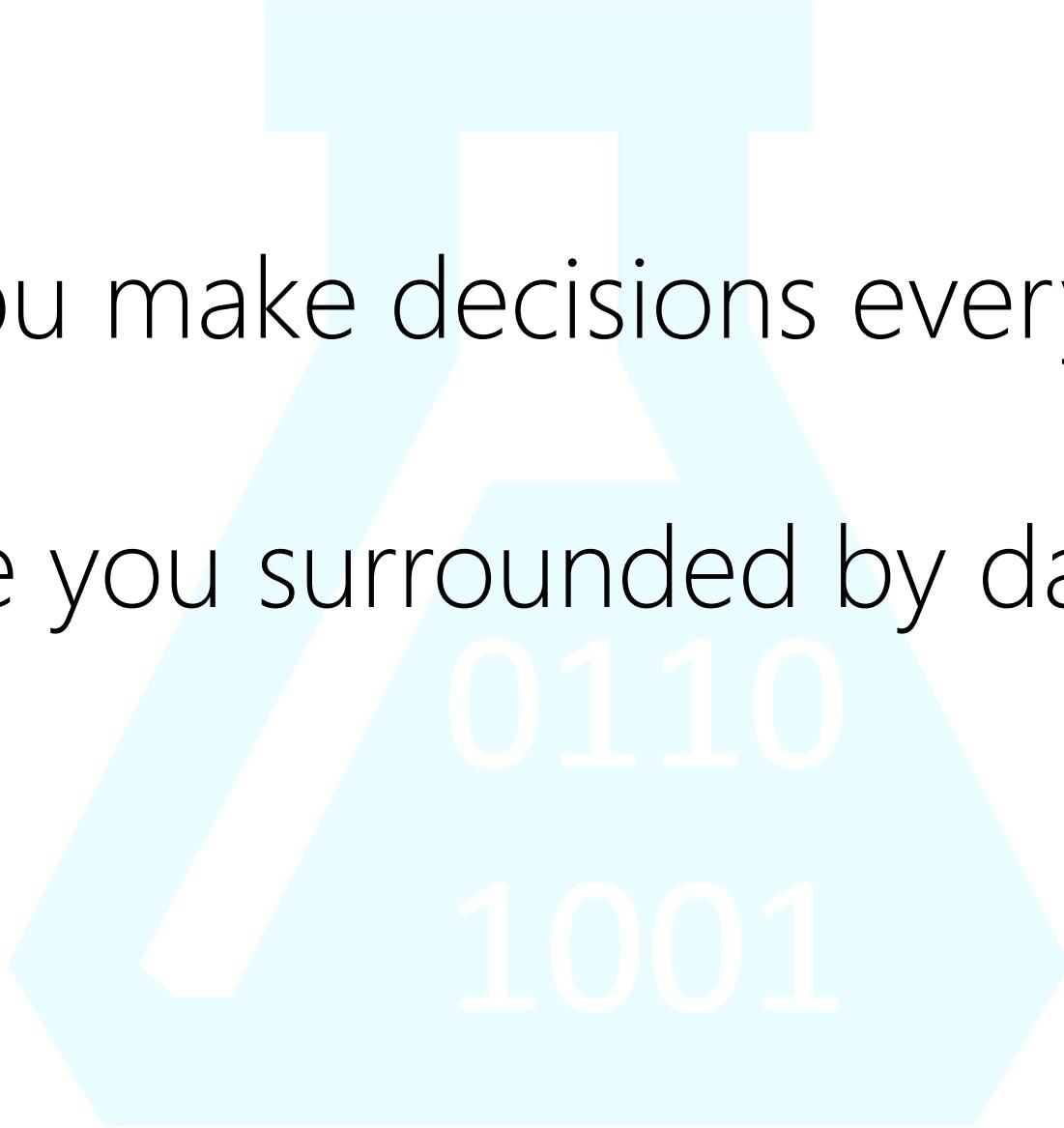
DevSci: Better Software Through Data Science

@MatthewRenze

#MicrosoftUAE



Do you make decisions every day?



Do you make decisions every day?

Are you surrounded by data?



Do you make decisions every day?

Are you surrounded by data?

Are you using data to make decisions?



What is data science?

Why is it important?

How do I get started?







The Economist

FEBRUARY 27TH-MARCH 5TH 2010

Gordon Brown's pitch
What went wrong at RBS
Genetically modified crops blossom
The EU woos Russia
The right to eat cats and dogs

The data deluge

AND HOW TO HANDLE IT: A 14-PAGE SPECIAL REPORT

The New York Times

For Today's Graduate, Just One Word: Statistics

By STEVE LOMK
Published: August 5, 2009

MOUNTAIN VIEW, Calif. — At Harvard, Carrie Grimes majored in anthropology and archaeology and ventured to places like Honduras, where she studied Mayan settlement patterns by mapping where artifacts were found. But she was drawn to what she calls “all the computer and math stuff” that was part of the job.

- [TWITTER](#)
- [LINKEDIN](#)
- [COMMENTS \(58\)](#)
- [SIGN IN TO E-MAIL](#)

Data Scientist: *The Sexiest Job of the 21st Century*

Meet the people who can coax treasure out of messy, unstructured data.

by Thomas H. Davenport
and D.J. Patil

When Jonathan Goldman arrived for work in June 2006 at LinkedIn, the business networking site, the place still felt like a start-up. The company had just under 8 million accounts, and the number was growing quickly as existing members invited their friends and colleagues to join. But users weren't seeking out connections with the people who were already on the site at the rate executives had expected. Something was apparently missing in the social experience. As one LinkedIn manager put it, “It was like arriving at a conference reception and realizing you don't know anyone. So you just stand in the corner sipping your drink—and you probably leave early.”

Job Postings for Data Scientists

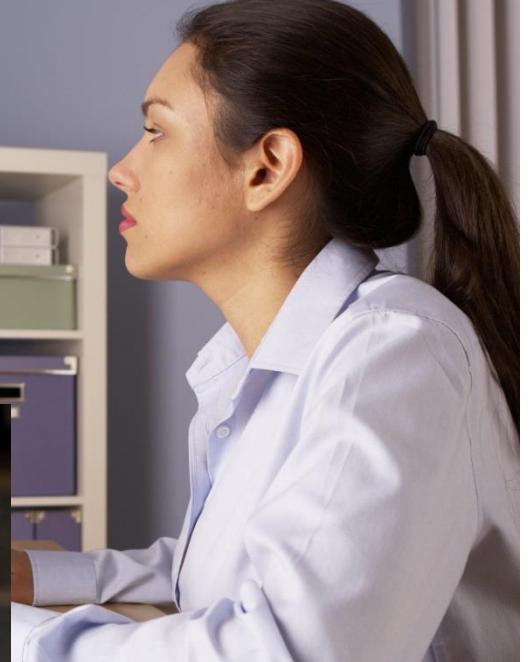
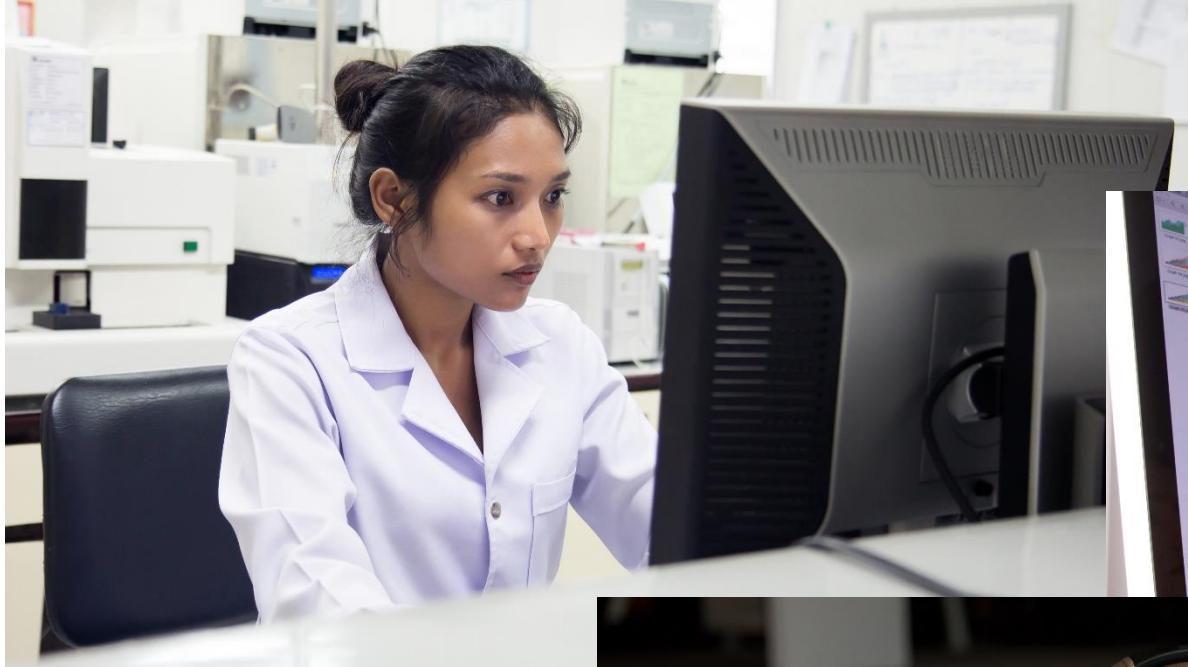


Top-paying Tech Skills

Skill	2016	Change
HANA (High Performance Analytical Application)	\$ 128,958	-3.3%
MapReduce	\$ 125,009	-0.3%
Cloud Foundry	\$ 124,038	n/a
Hbase	\$ 123,934	5.7%
Omnigraffle	\$ 123,782	-1.9%
Cassandra	\$ 123,459	2.2%
Apache Kafka	\$ 122,728	n/a
SOA (Service Oriented Architecture)	\$ 122,094	-1.9%
Ansible	\$ 121,382	n/a
Jetty	\$ 120,978	1.3%
PaaS (Platform as a Service)	\$ 120,403	-4.4%
Elasticsearch	\$ 120,002	n/a
ABAP (Advanced Business Application Programming)	\$ 119,961	0.5%
NoSQL	\$ 119,498	1.3%
CMMI (Capability Maturity Model Integration)	\$ 119,466	-0.6%
Amazon Redshift	\$ 119,197	n/a
Pig	\$ 119,118	-4.2%
Solr	\$ 119,032	0.1%
Cloudera	\$ 118,896	-9.0%
Docker	\$ 118,873	0.2%

Skill	2016	Change
Amazon Route 53	\$ 118,828	n/a
Hadoop	\$ 118,625	-2.5%
Hive	\$ 118,589	-1.3%
Korn Shell	\$ 118,273	1.4%
PMBok (Project Management Body of Knowledge)	\$ 118,233	0.7%
Dynamo DB	\$ 118,119	n/a
Groovy	\$ 117,897	-0.1%
IaaS (Infrastructure as a Service)	\$ 117,422	n/a
JAX-RS (Java API RestFUL Services)	\$ 116,997	n/a
RabbitMQ	\$ 116,909	n/a
JDBC (Java Database Connectivity)	\$ 116,833	2.0%
SOX (Sarbanes Oxley)	\$ 116,743	0.6%
Objective C	\$ 116,667	2.5%
FCoE (Fibre Channel over Ethernet)	\$ 116,145	7.2%
UML (Unified Modeling Language)	\$ 115,285	-3.6%
XSLT (Extensible Stylesheet Language Transformations)	\$ 115,089	3.5%
Redis	\$ 114,922	2.8%
ETL (Extract Transform and Load)	\$ 114,892	2.6%
SDN (Software Defined Network)	\$ 114,739	-2.3%
Informatica	\$ 114,143	1.1%

Source: Dice Salary Survey 2017







About Me

IOWA STATE
UNIVERSITY



About Me

Data Science Consultant

IOWA STATE
UNIVERSITY



About Me

Data Science Consultant

Education

B.S. in Computer Science (ISU)

B.A. in Philosophy (ISU)

IOWA STATE
UNIVERSITY



About Me

Data Science Consultant Education

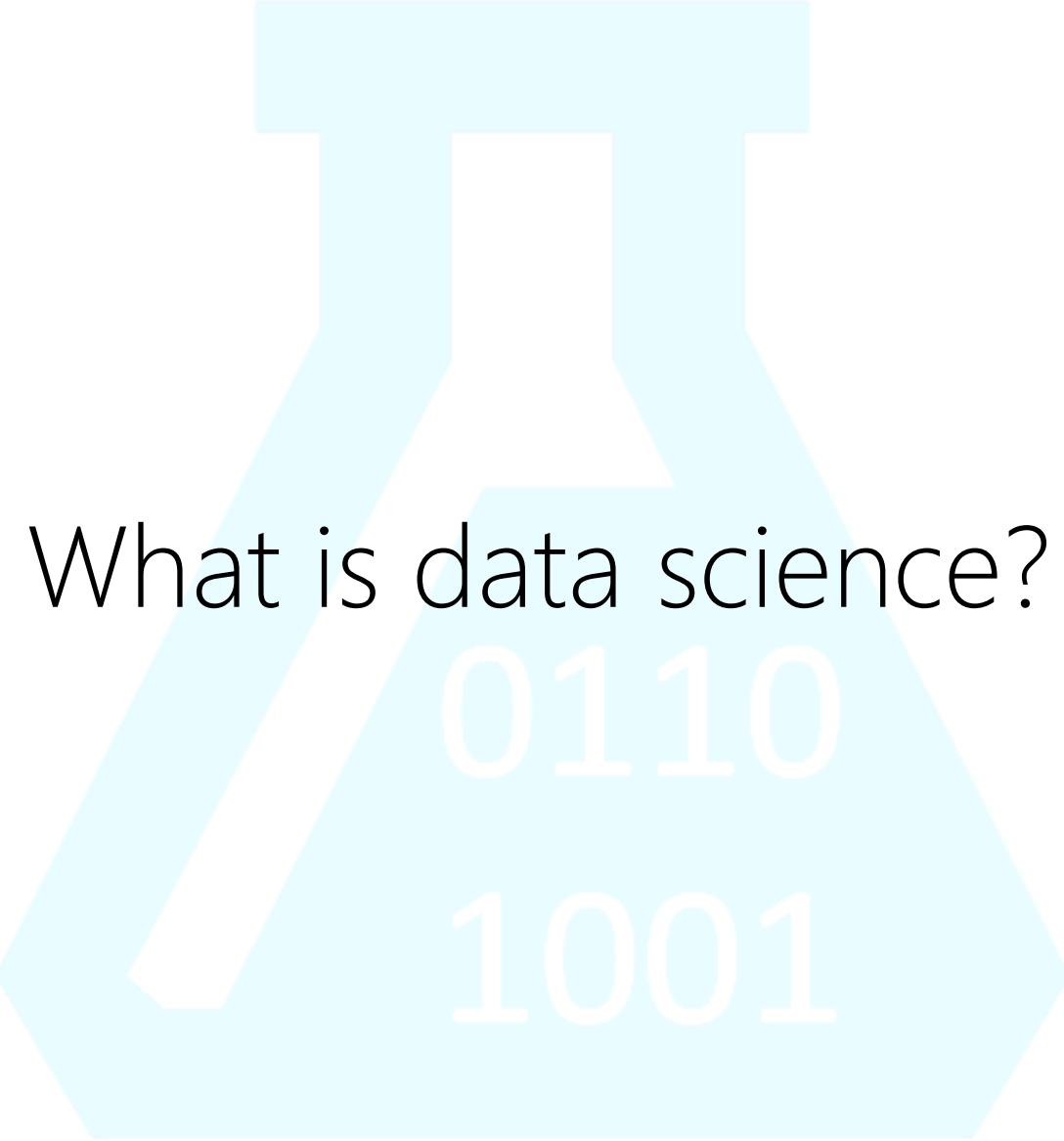
B.S. in Computer Science (ISU)
B.A. in Philosophy (ISU)

Community

Keynote speaker
Pluralsight author
DataCamp author
Microsoft MVP AI
ASPIInsider

IOWA STATE
UNIVERSITY

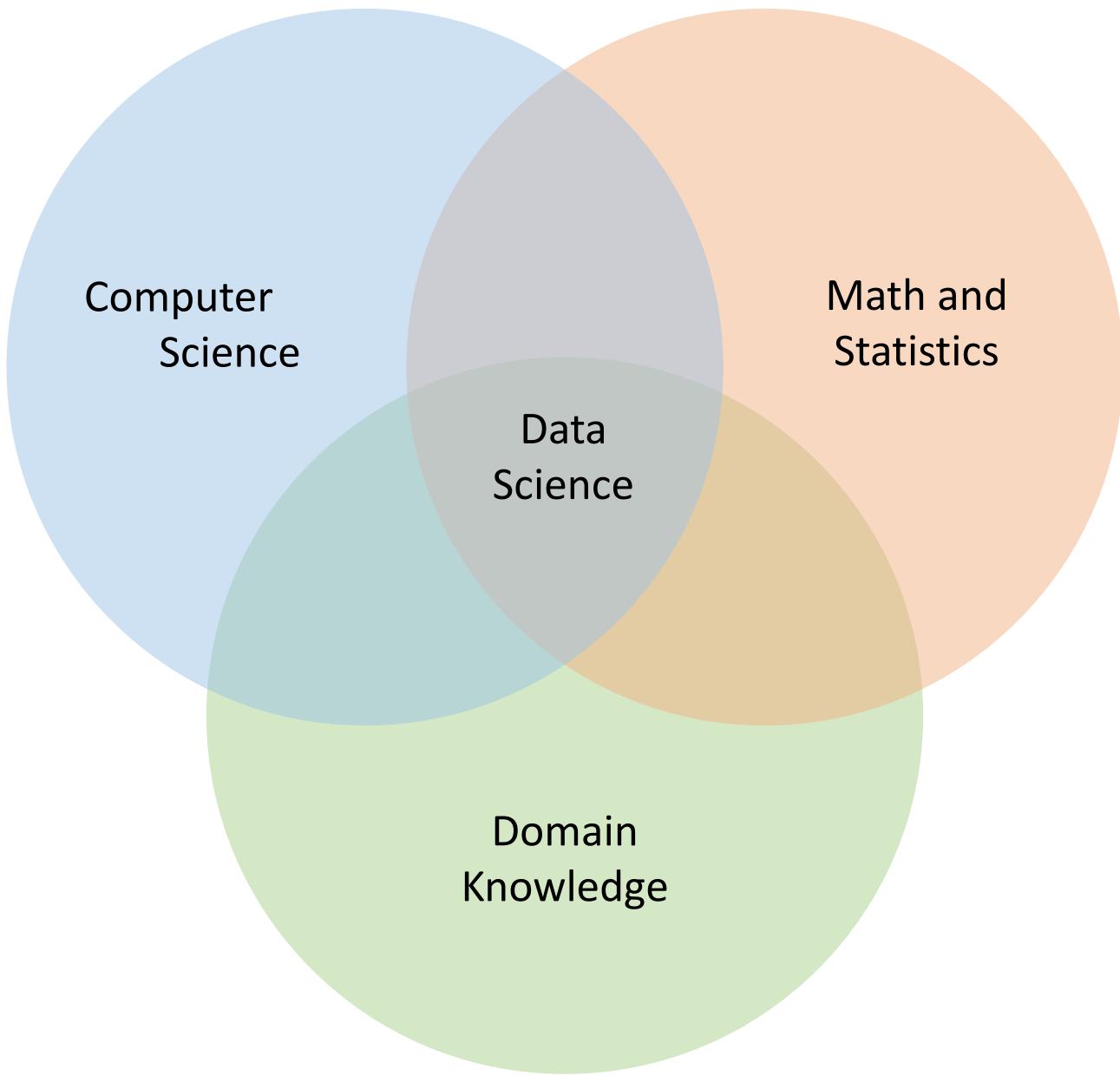


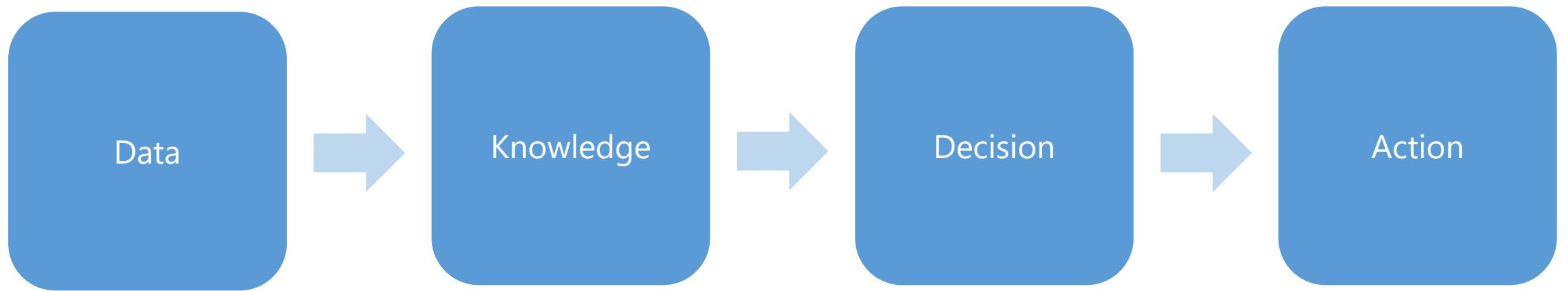


What is data science?

0110

1001





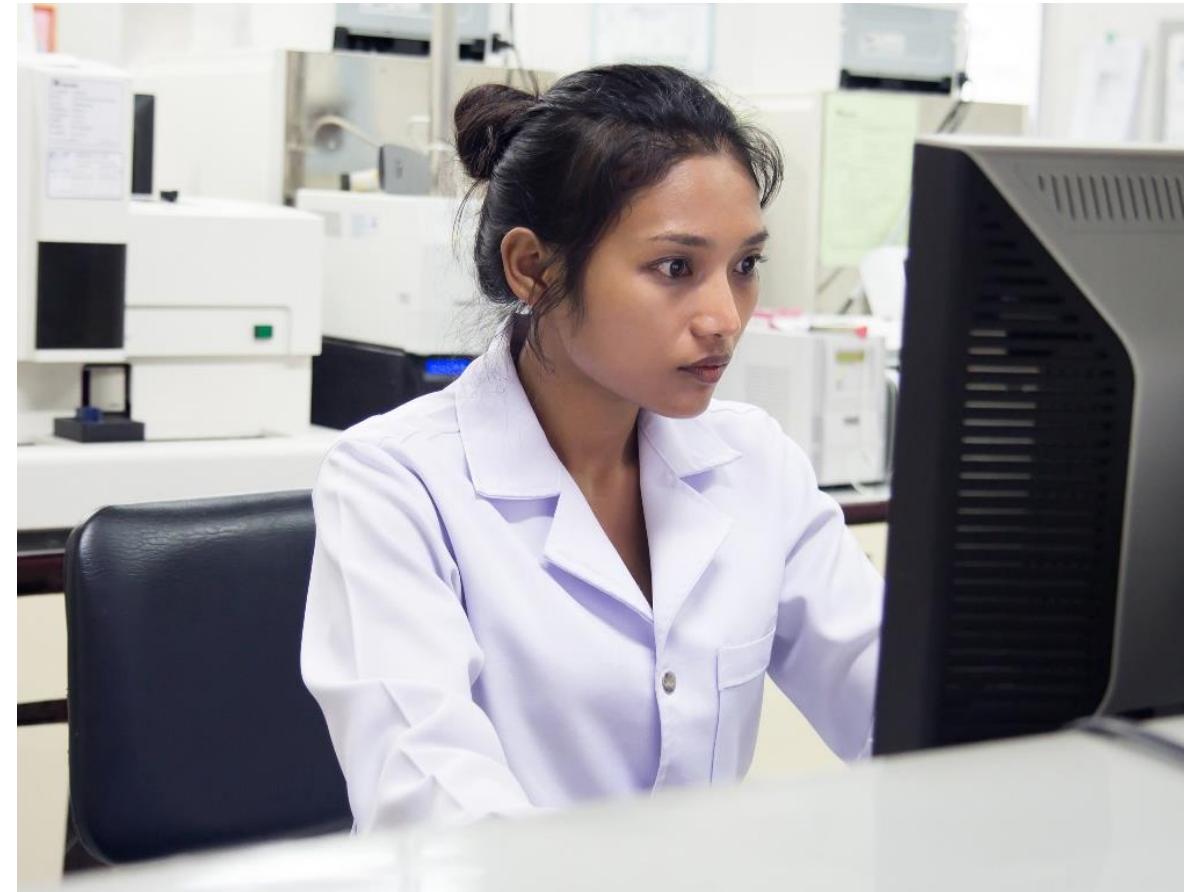
What Is a Data Scientist?

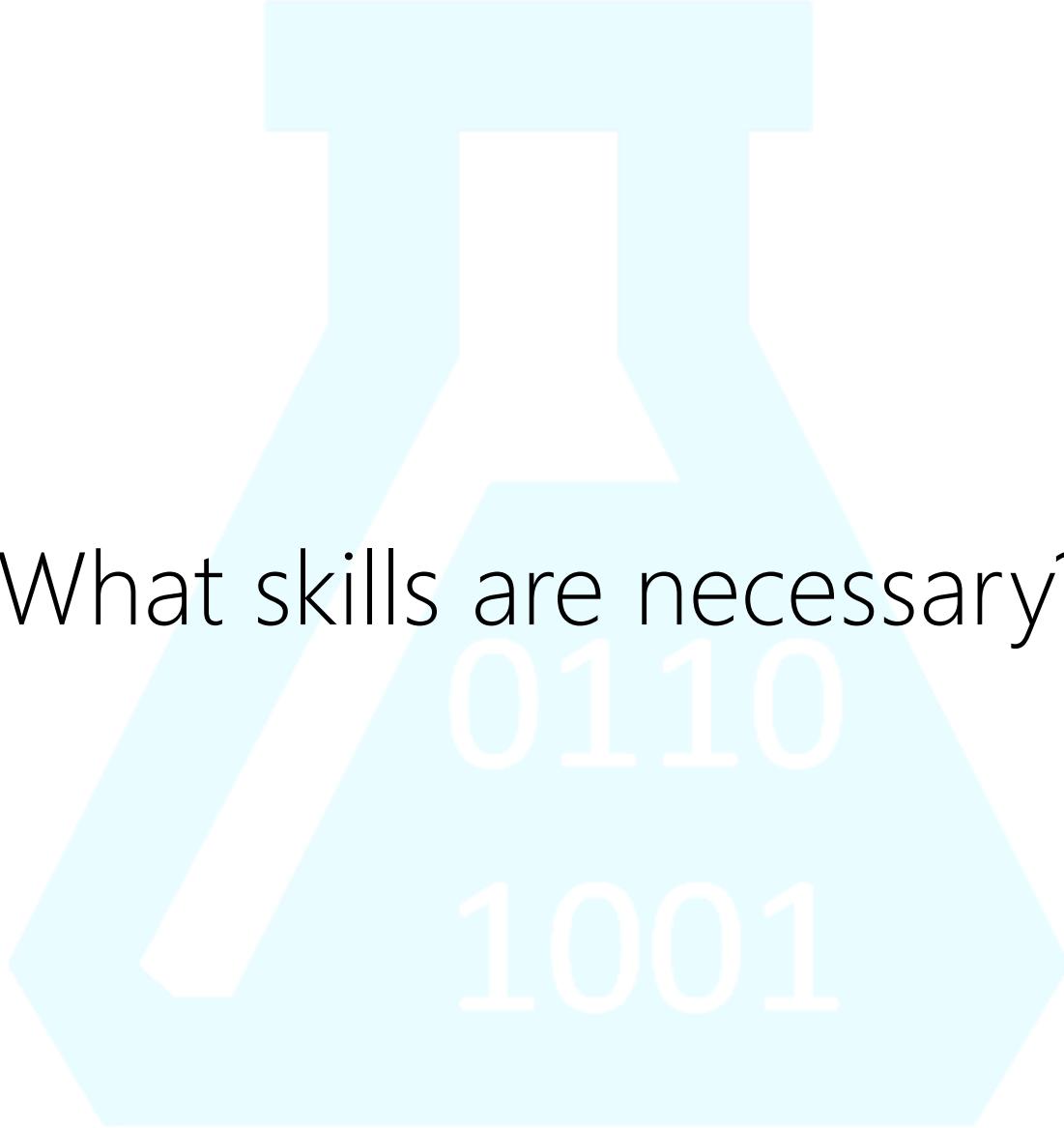
Performs data science

More than a scientist

More than an analyst

More than a developer





What skills are necessary?

0110

1001

Data Science Skills

Programming

Working with data

Descriptive statistics

Data visualization

Data Science Skills

Programming

Working with data

Descriptive statistics

Data visualization

Statistical modeling

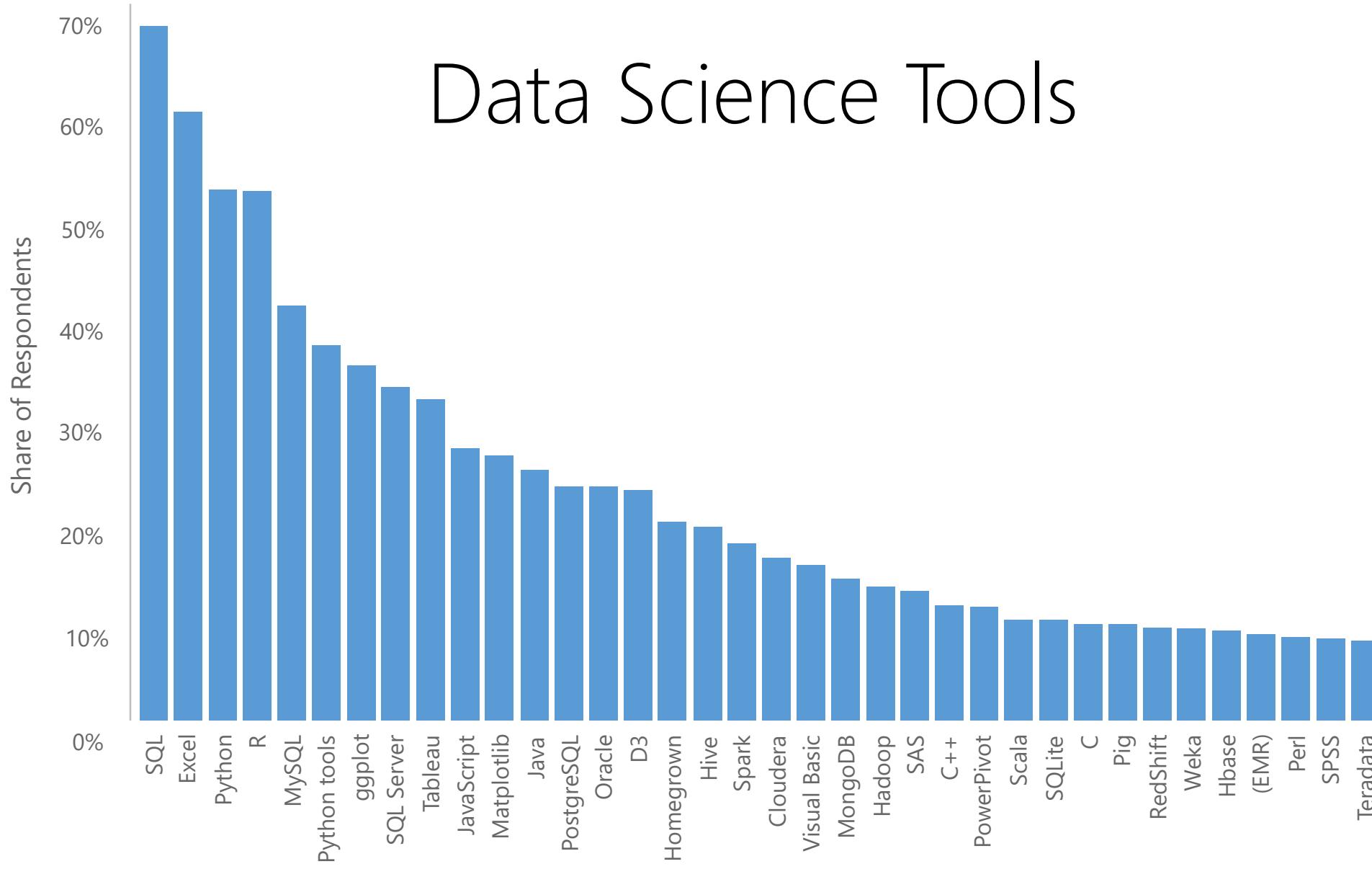
Handling Big Data

Machine learning

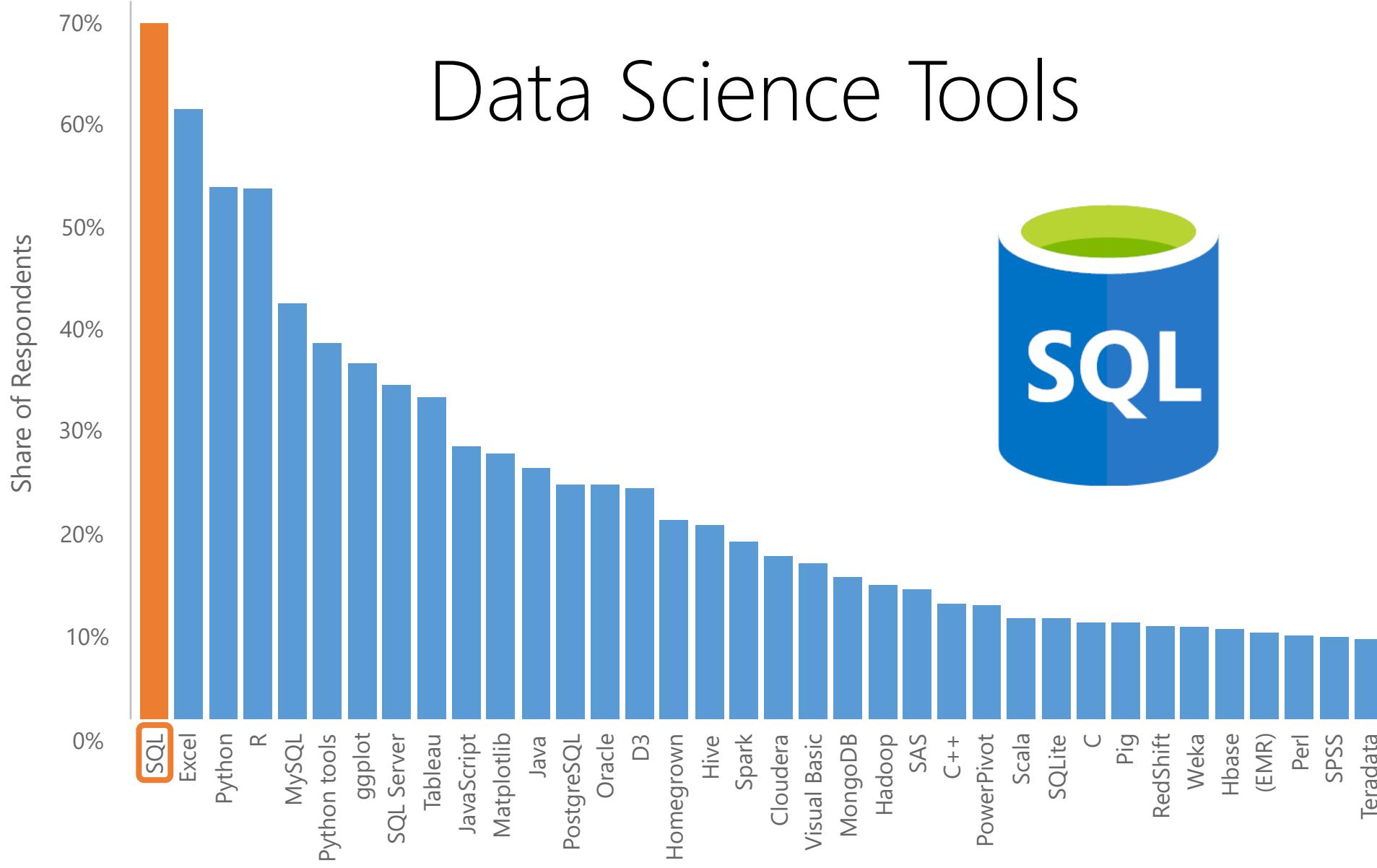
Deploying to production



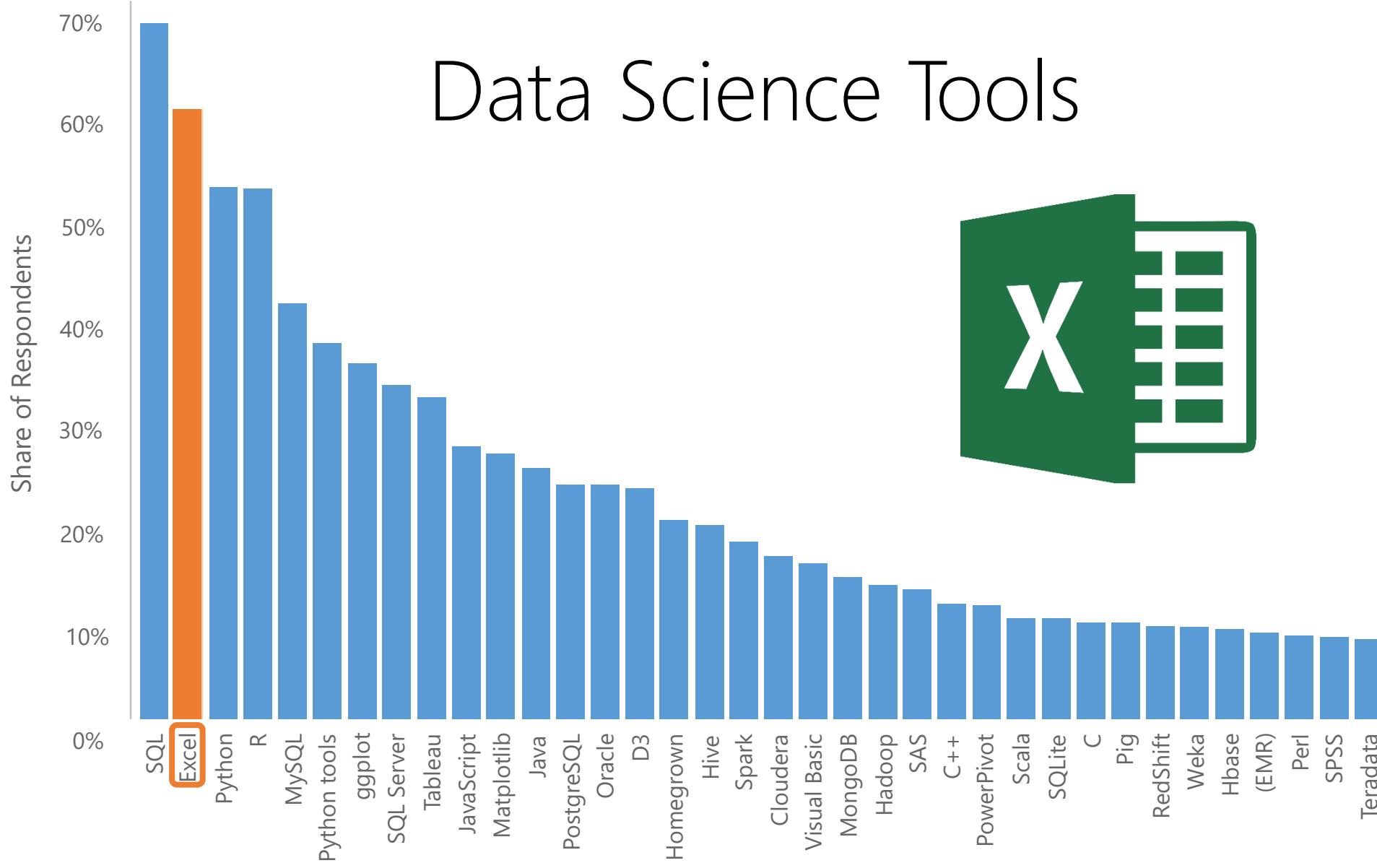
What tools are used?



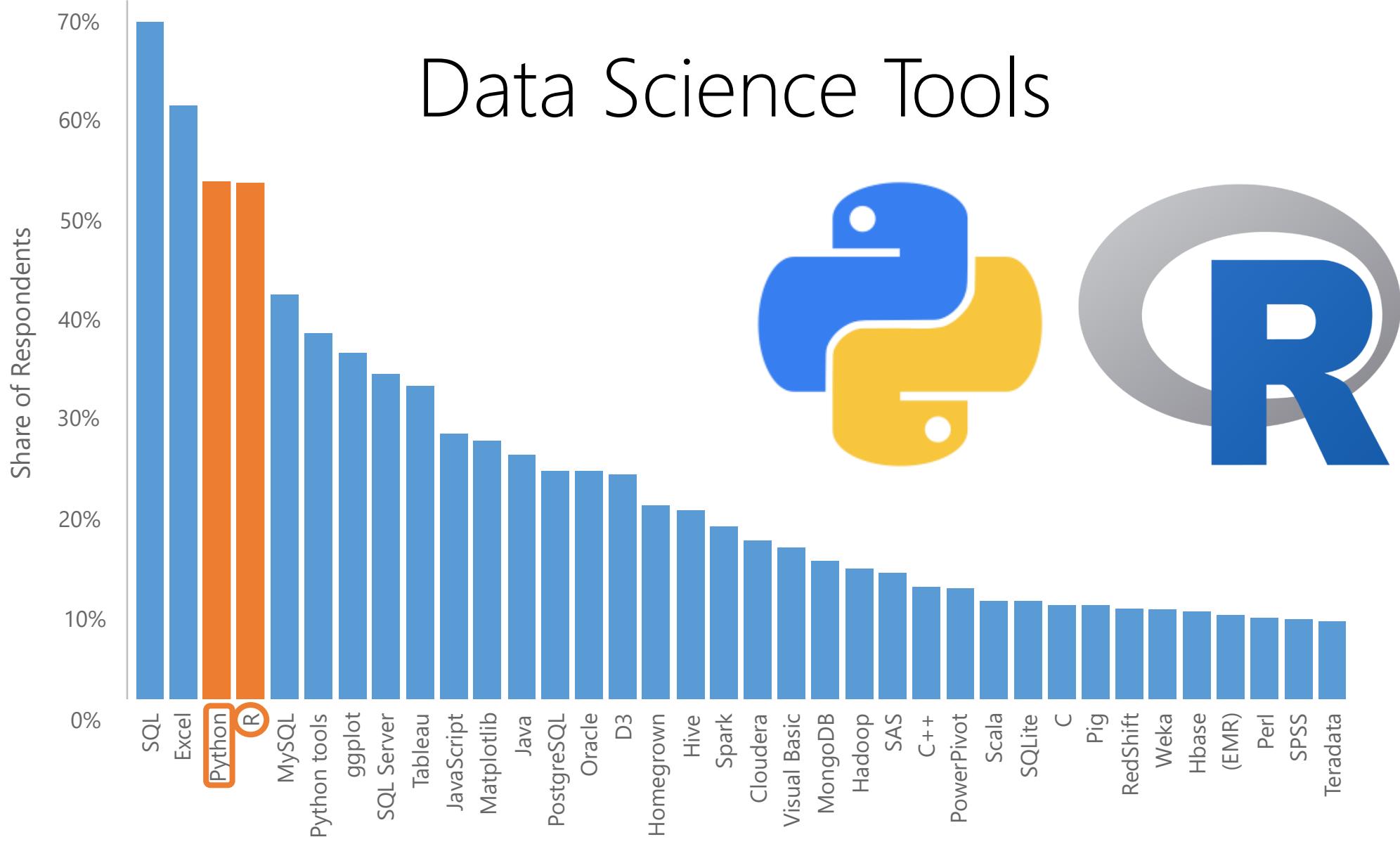
Source: O'Reilly 2015 Data Science Salary Survey



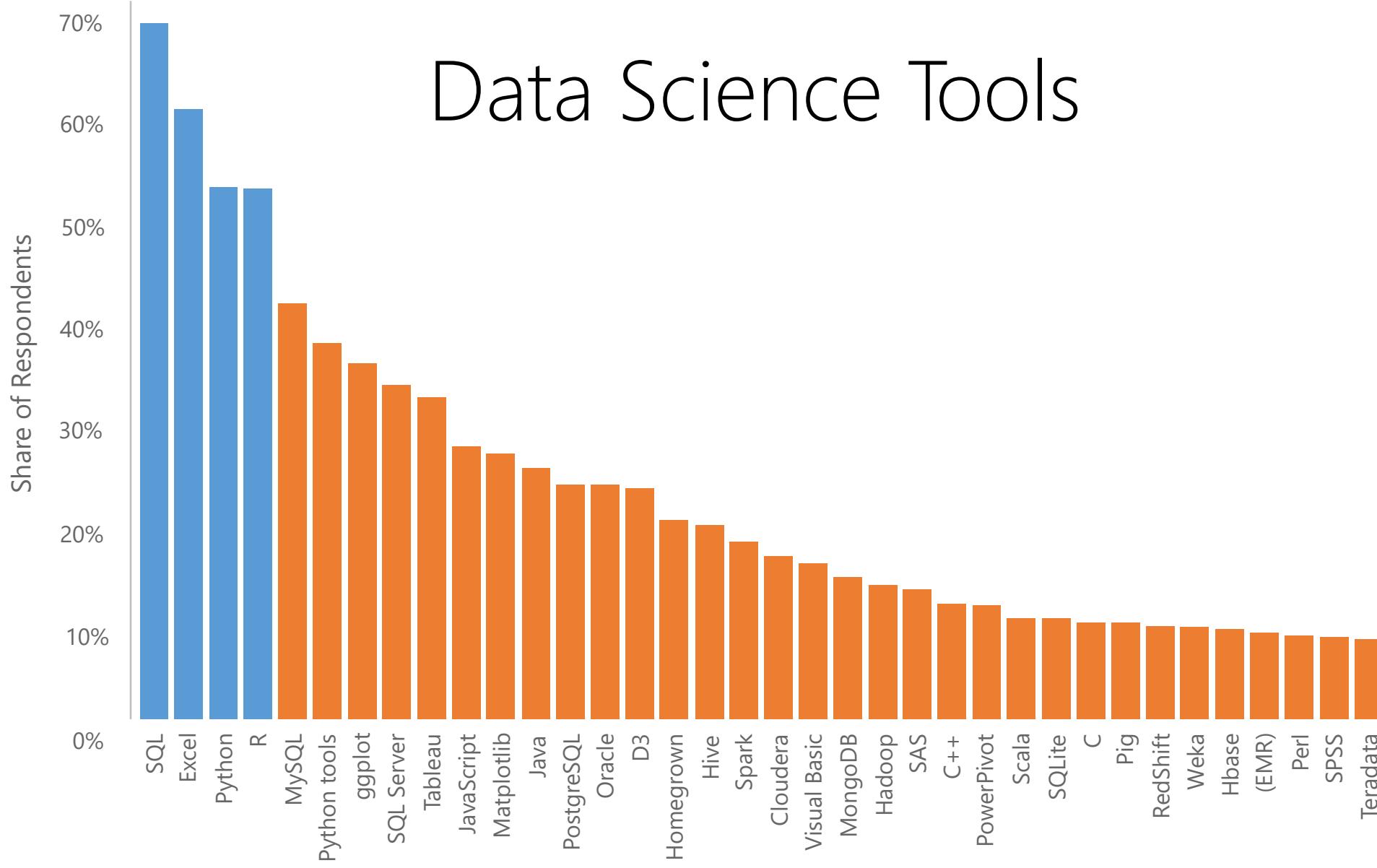
Source: O'Reilly 2015 Data Science Salary Survey



Source: O'Reilly 2015 Data Science Salary Survey



Source: O'Reilly 2015 Data Science Salary Survey



Source: O'Reilly 2015 Data Science Salary Survey



How is data science performed?

The Data Science Process



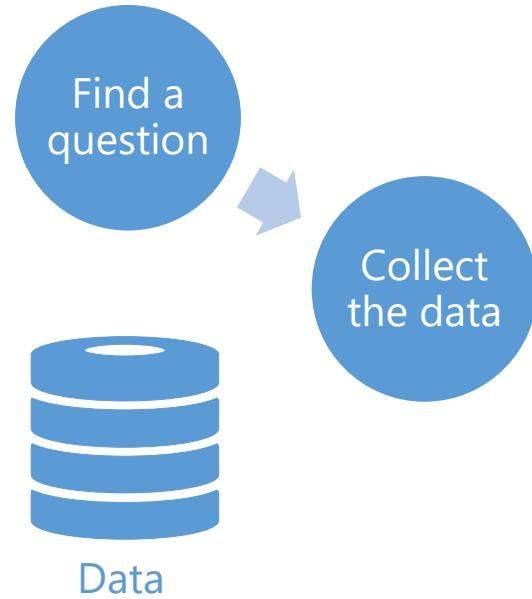
Data

The Data Science Process

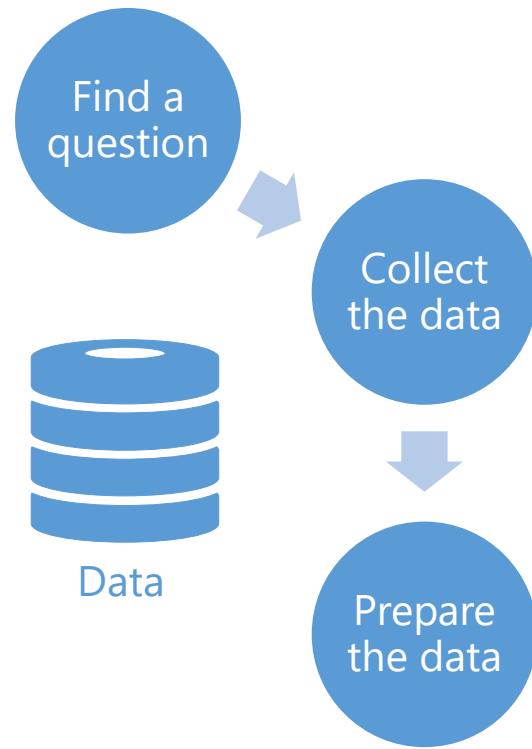


Data

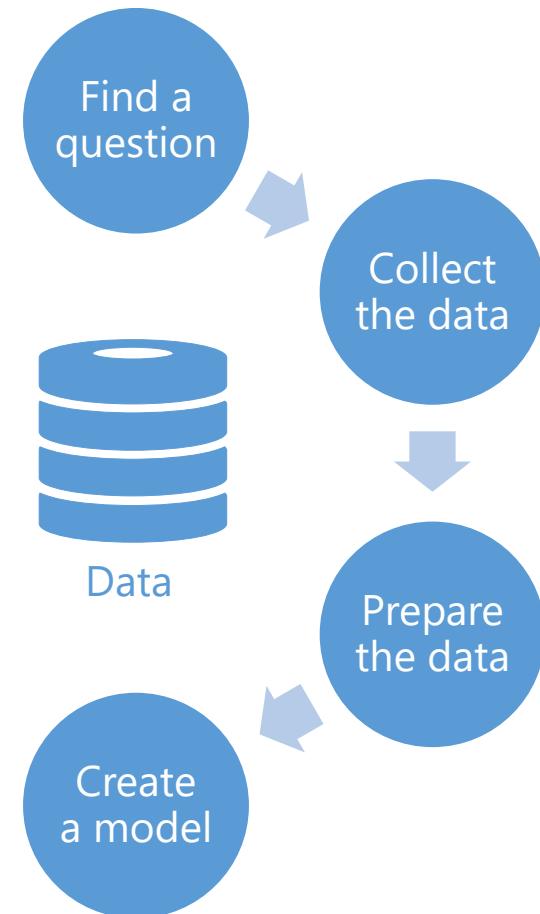
The Data Science Process



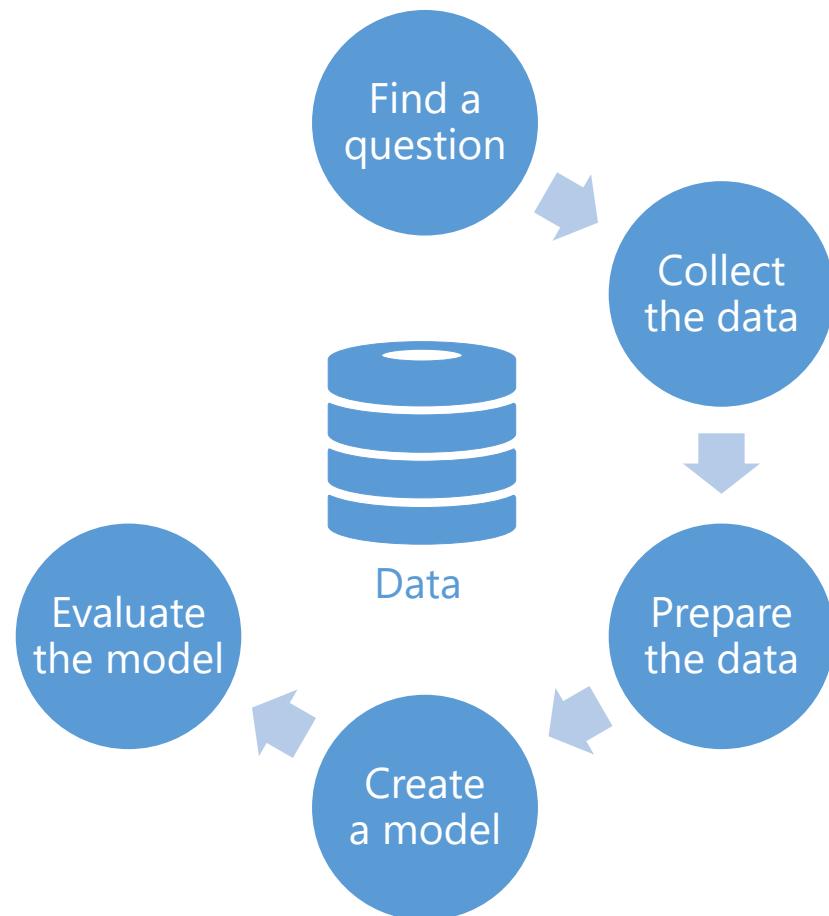
The Data Science Process



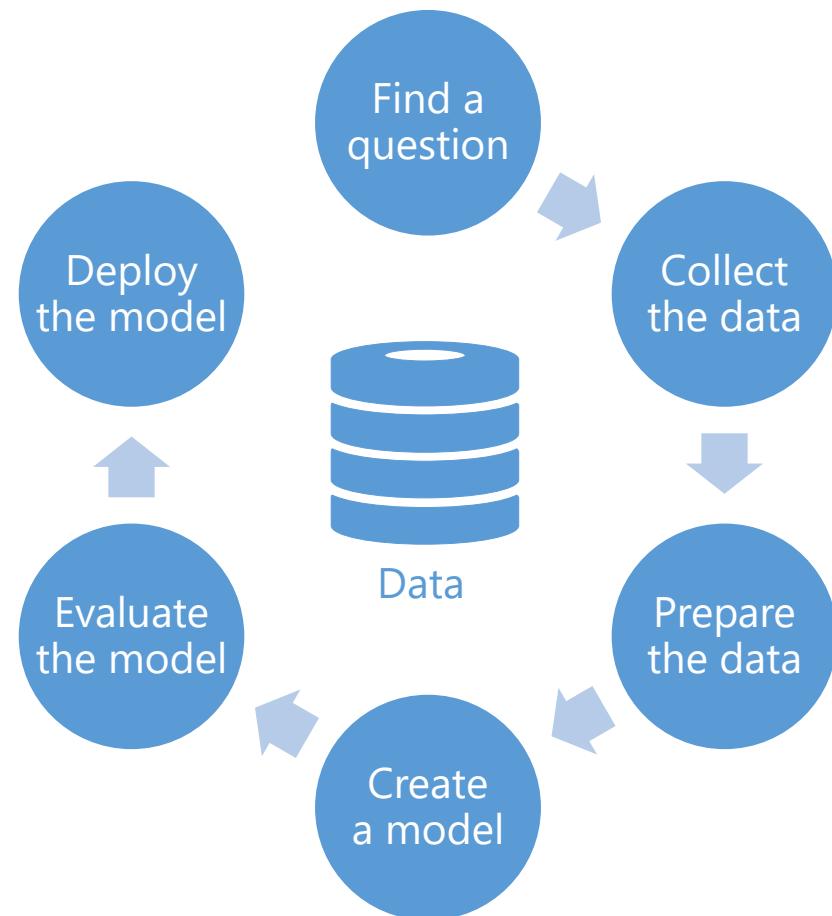
The Data Science Process



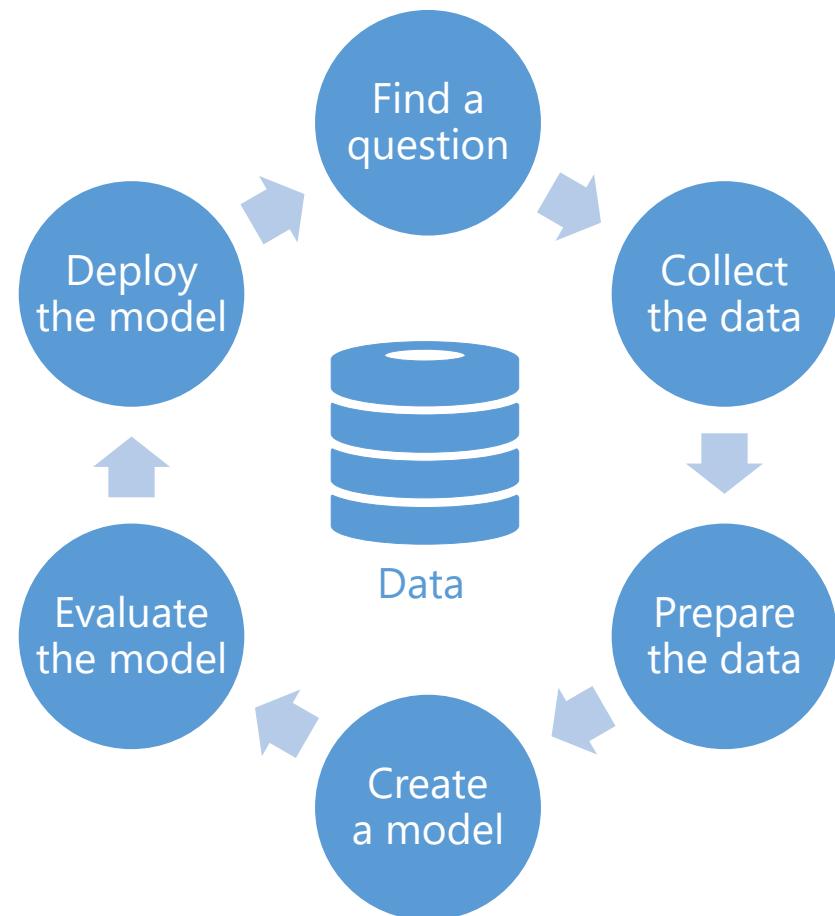
The Data Science Process



The Data Science Process

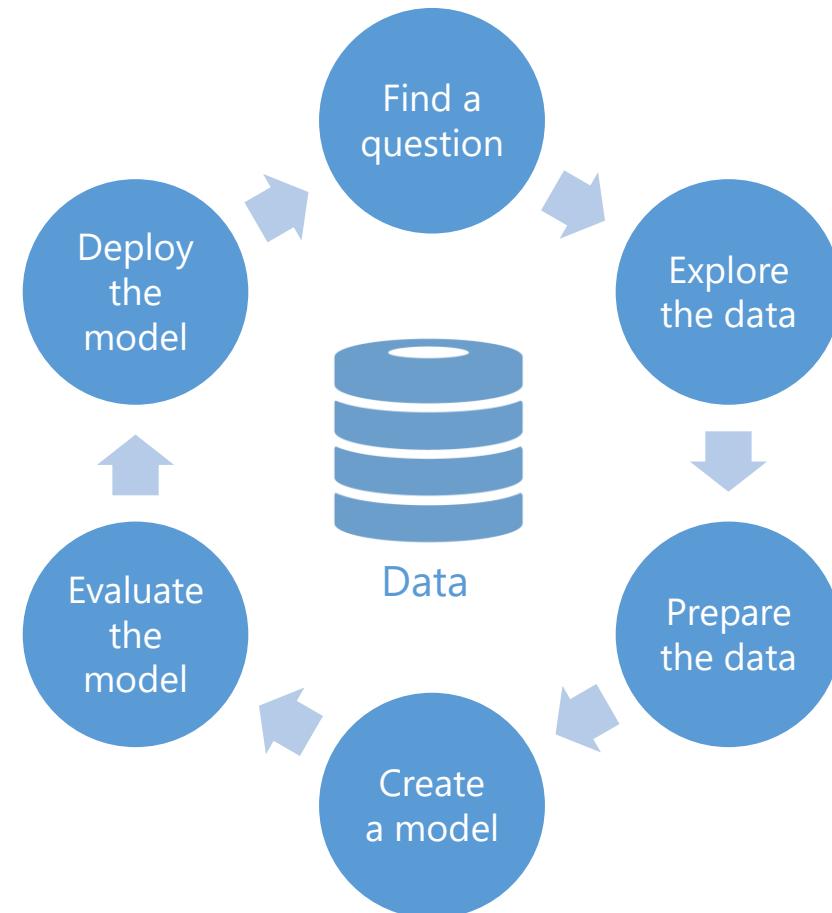


The Data Science Process



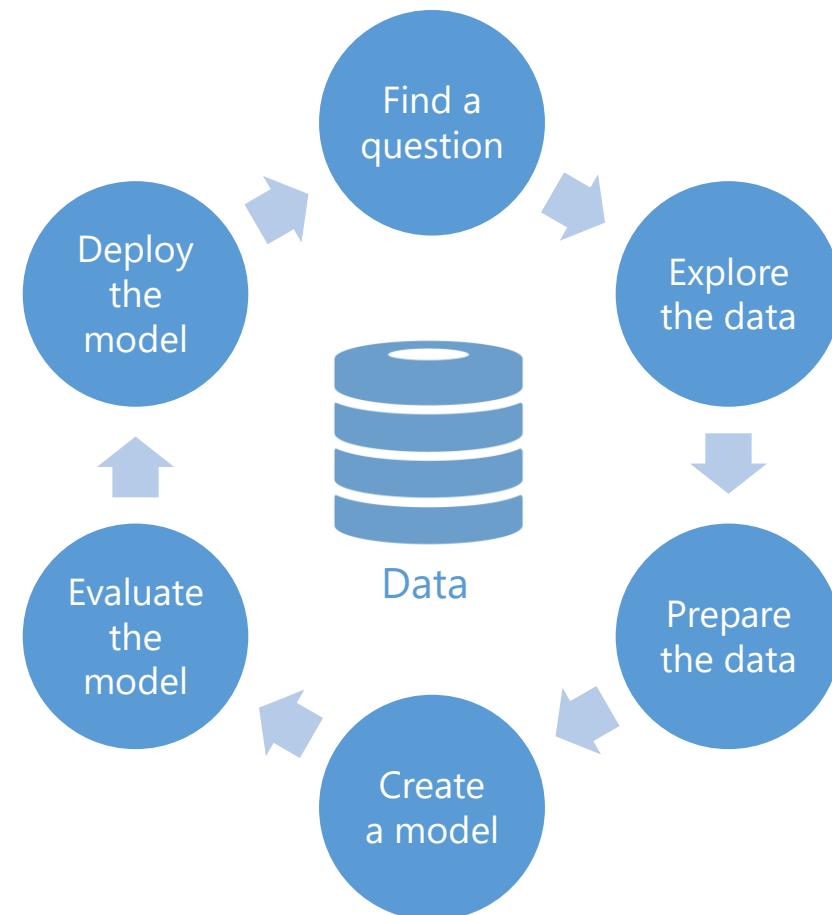
The Data Science Process

Iterative process



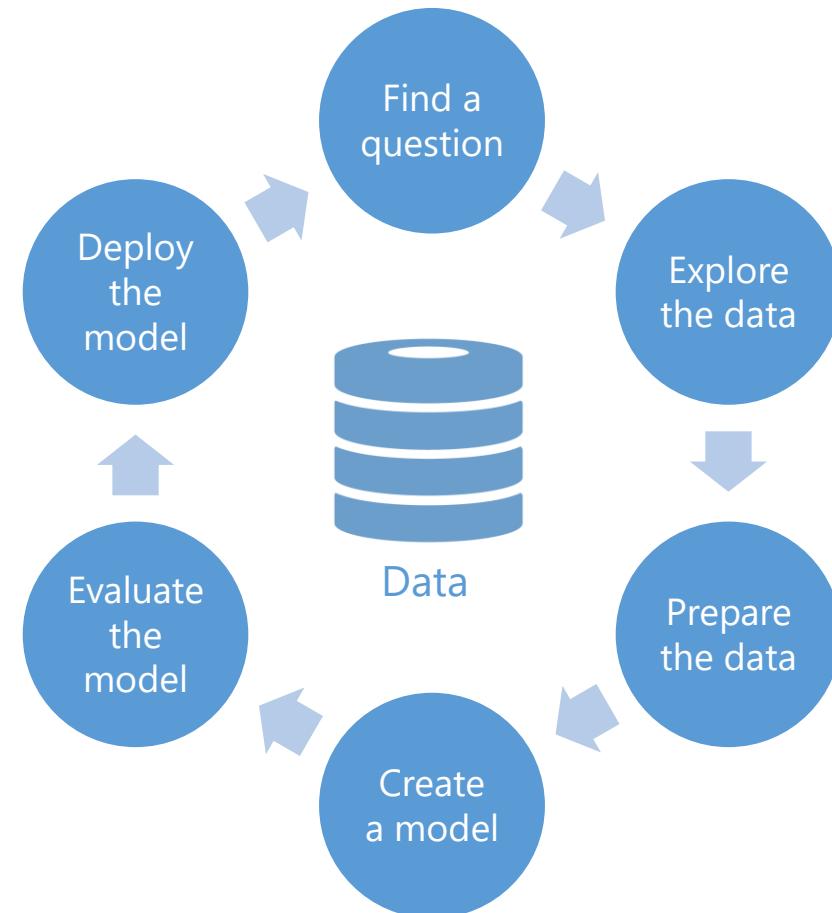
The Data Science Process

Iterative process
Non-sequential



The Data Science Process

Iterative process
Non-sequential
Early termination





Why is data science important?

0110

1001





4% higher productivity

6% higher profits















Family life



Medical



Hospital



Residence

Car



Factory



Travel







Why is data science important to ME?

Two Main Approaches

Build
intelligent
software

Improve
development
practices

Two Main Approaches



Build
intelligent
software



Improve
development
practices



Ask a question about the data on this dashboard

Total Stores

104

This Year's Sales

\$22.05M

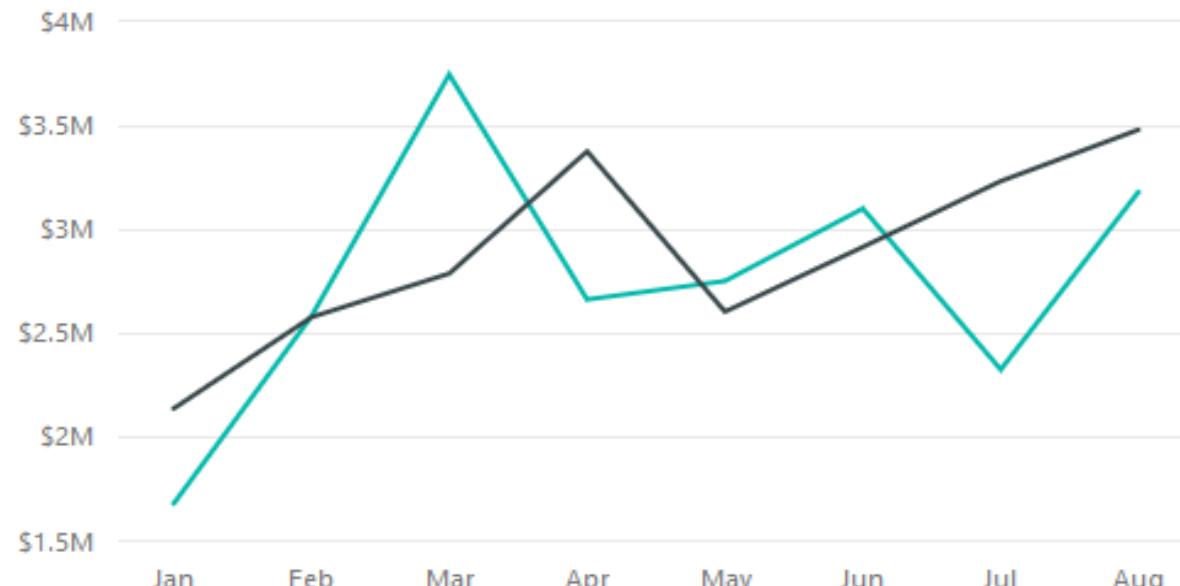
This Year's Sales
BY CHAIN

New Stores Opened This Year

10

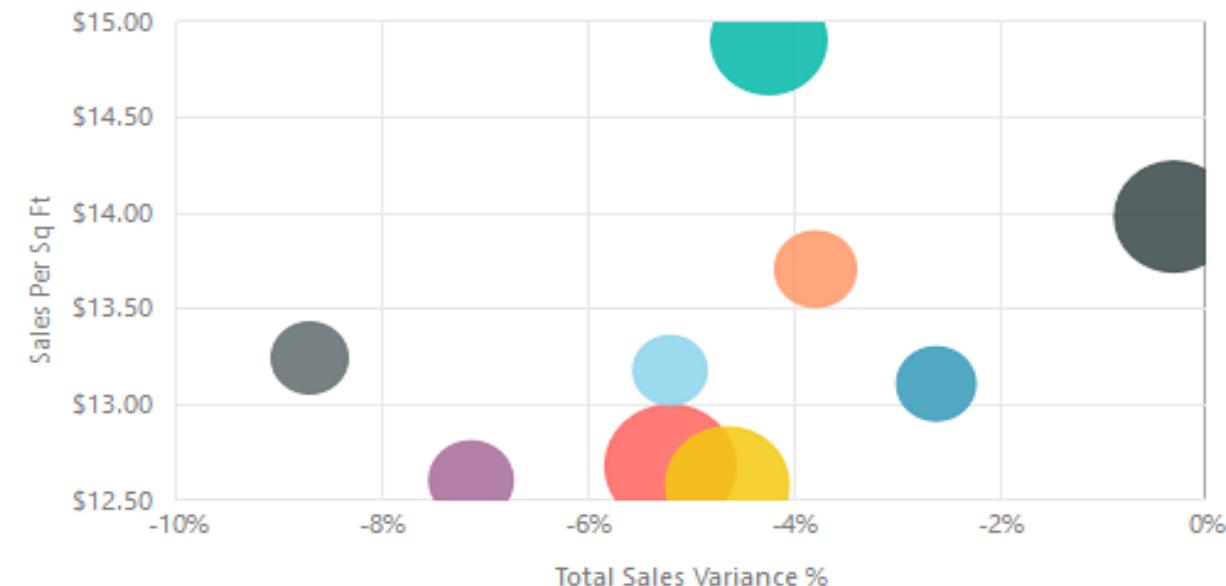
This Year's Sales, Last Year's Sales

BY FISCAL MONTH

● This Year Sales ● Last Year Sales

Sales Per Sq Ft, Total Sales Variance %, This Year's Sales

BY DISTRICT

District ● FD - 01 ● FD - 02 ● FD - 03 ● FD - 04 ● LI - 01 ● LI - 02 ● LI - 03



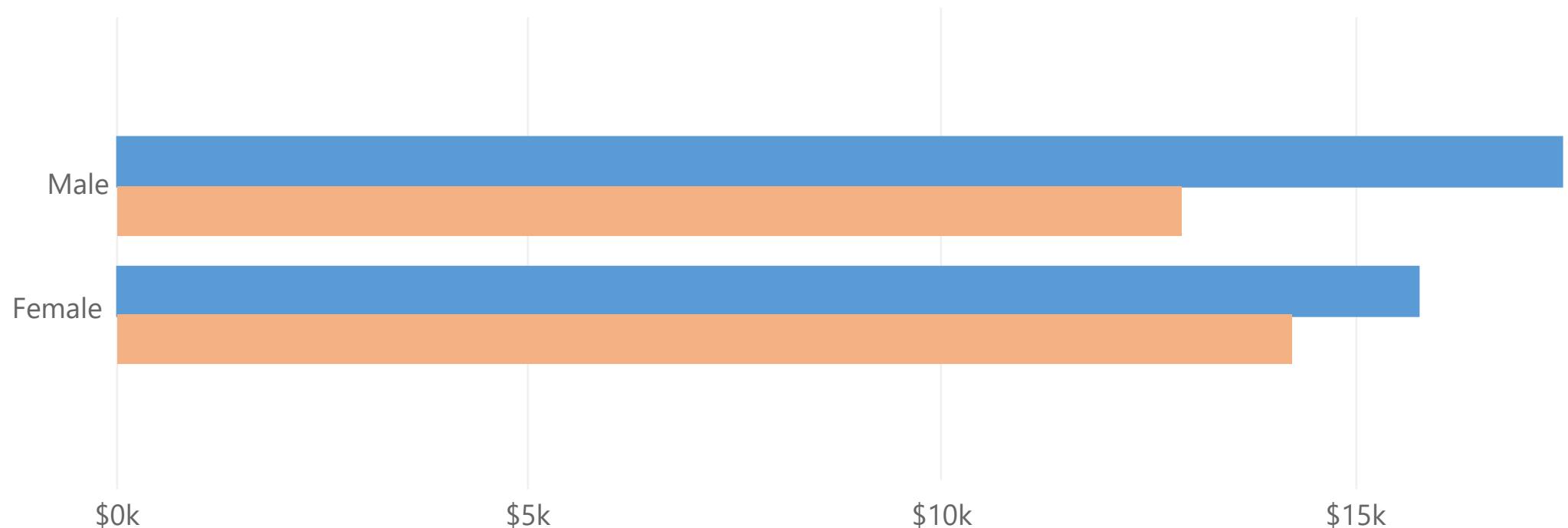
Internet Sales

Show me sales by gender and marital status.

Displaying sum of sales by gender and marital status

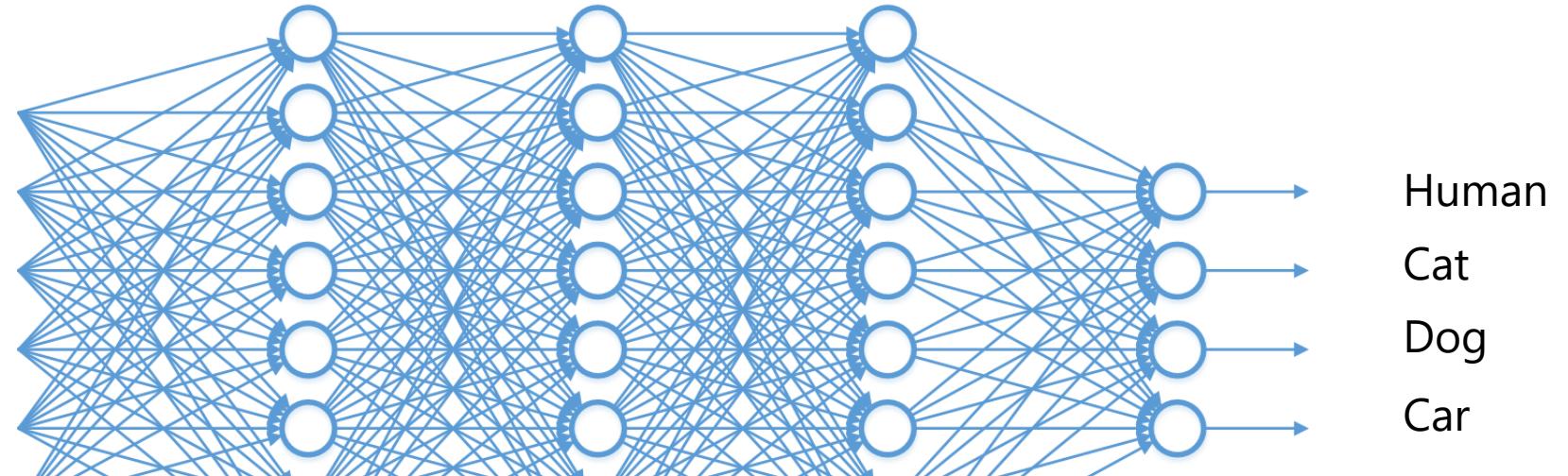
Marital Status:

- Married
- Single

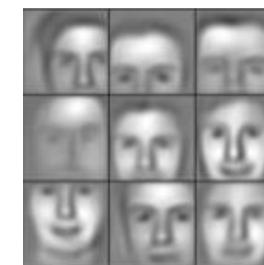
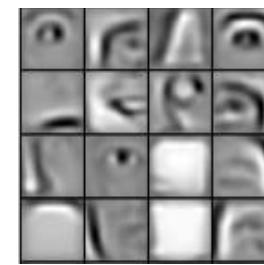
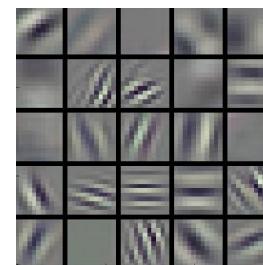


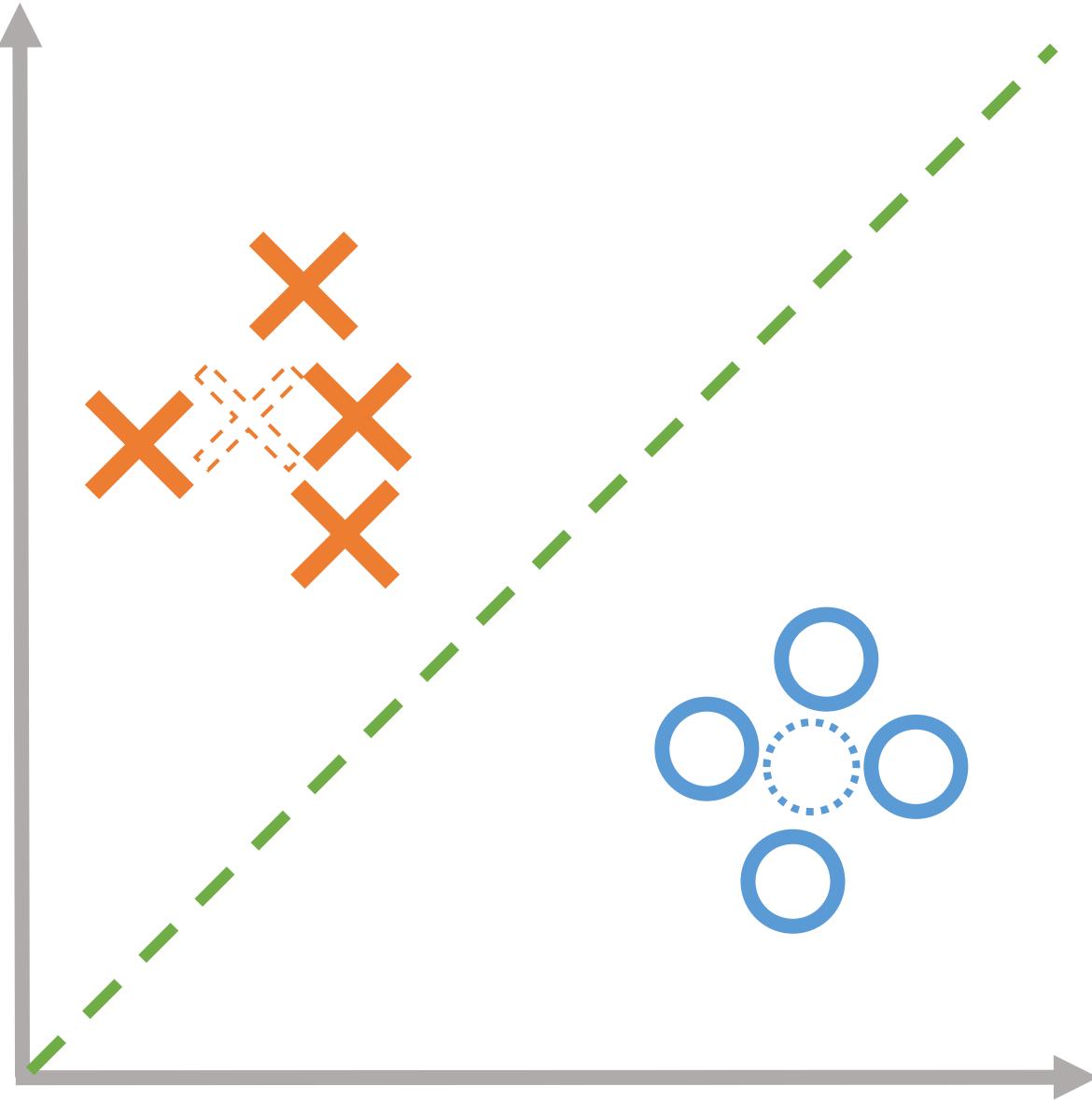
"Show me sales by
gender and marital
status."

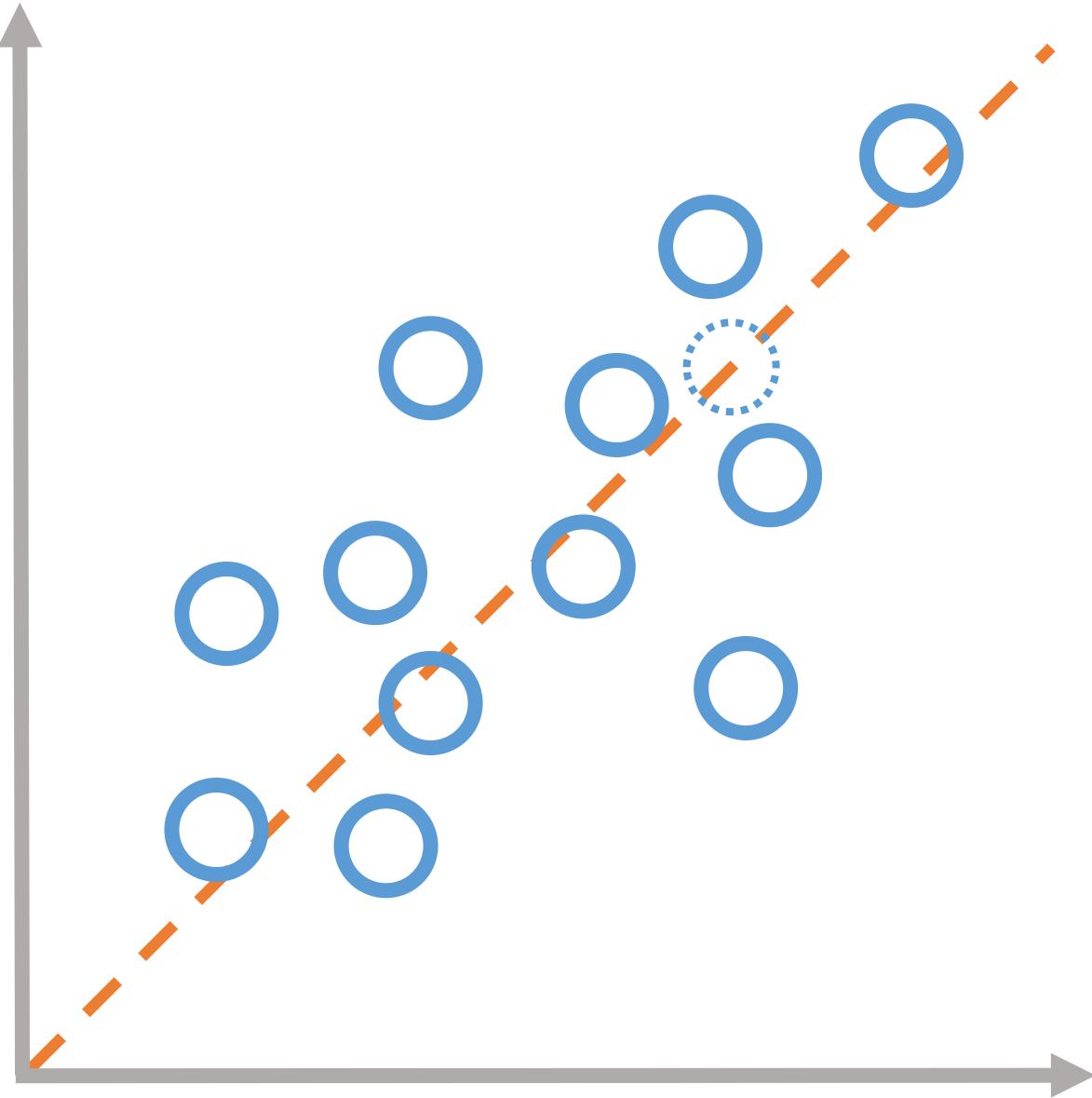
Machine Learning

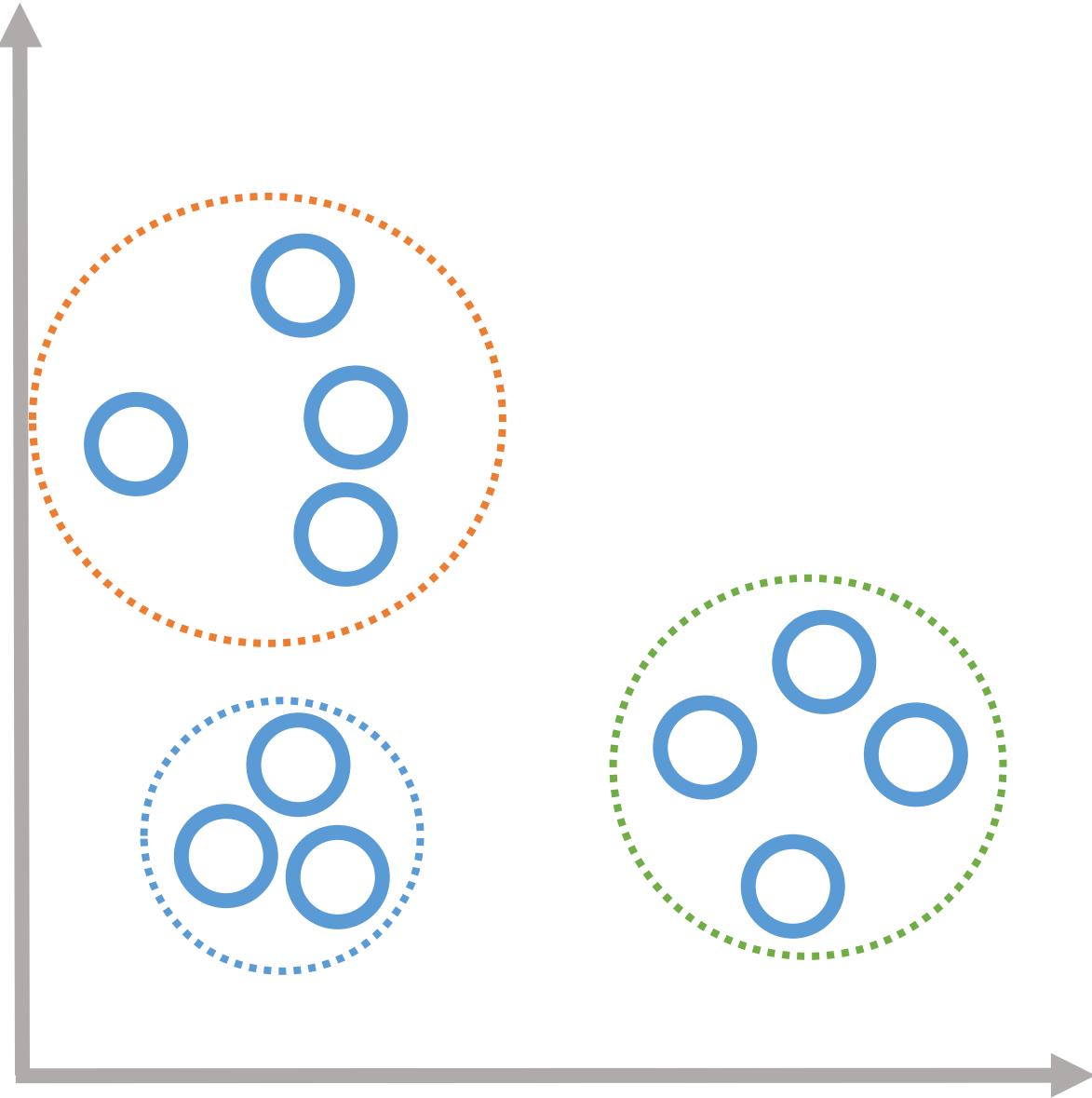


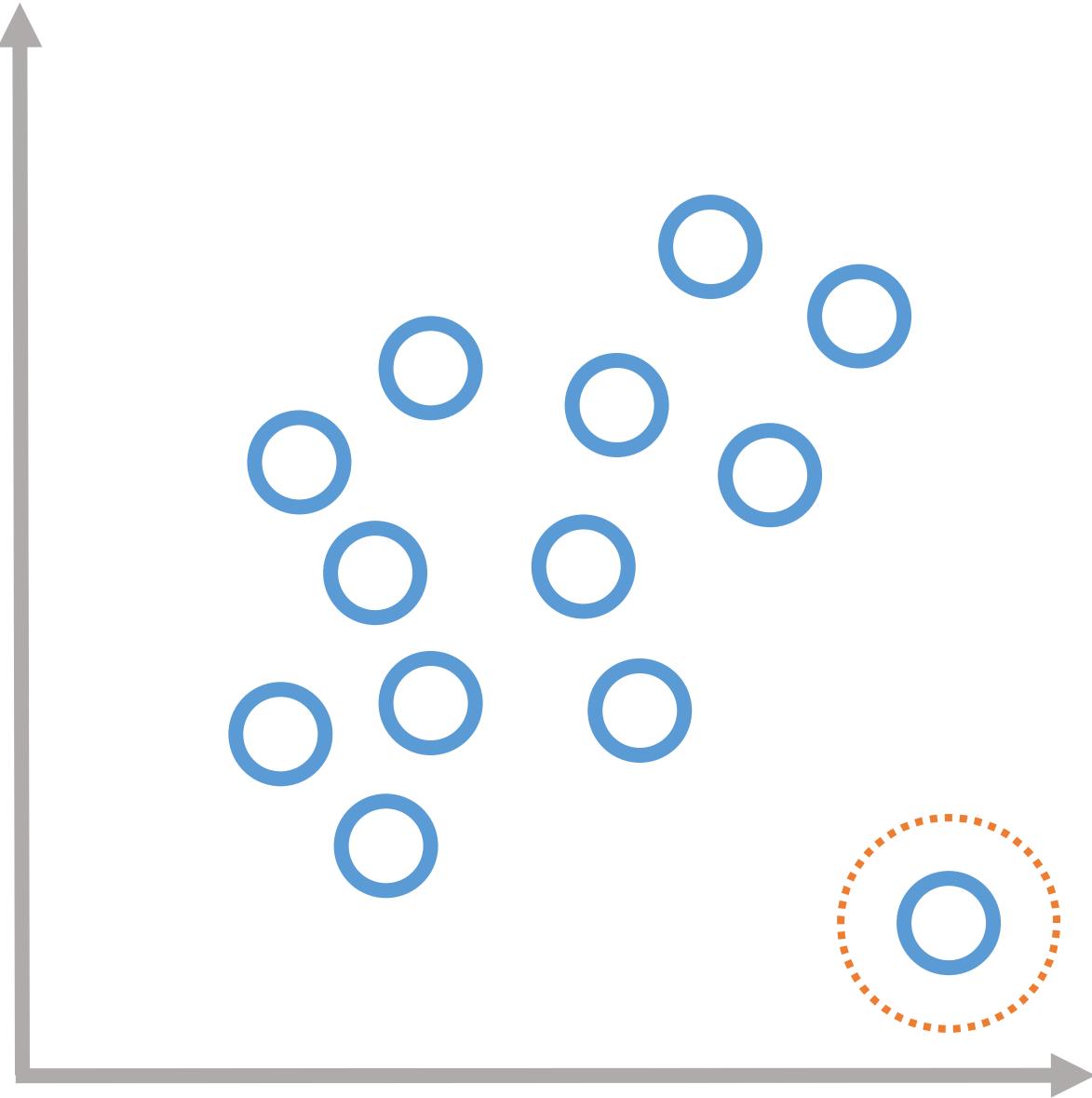
Human
Cat
Dog
Car



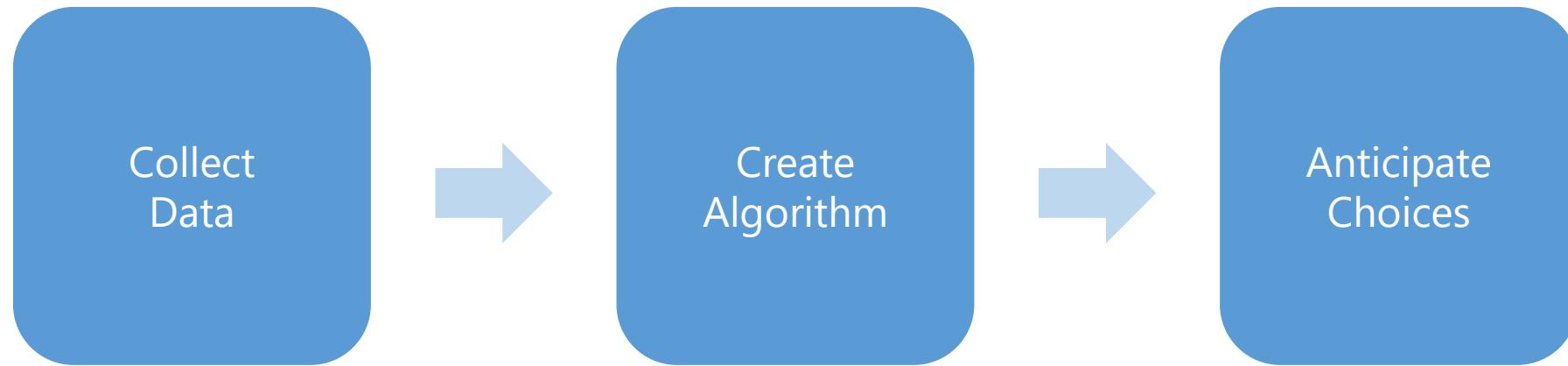








Anticipatory Design



...



Search



renze1577



Browse

Radio

YOUR LIBRARY

Your Daily Mix

Recently Played

Songs

Albums

Artists

Stations

Local Files

Videos

Podcasts

PLAYLISTS

Discover Weekly

Closer

One Man Puppet Show

A Puppet Named Julio

This Will Destroy You

Explosions in the Sky

Mastodon - Call Me

New Playlist

Your Daily Mixes

Play the music you love, without the effort. Packed with your favorites and new discoveries.



Daily Mix 1

Bonobo, The Album Leaf, Thievery Corporation and more

MADE FOR RENZE1577



Daily Mix 2

Epic45, Sigur Rós, mún and more

MADE FOR RENZE1577



Your Daily Mix 3

Joey Fehrenbach, The Album Leaf, The Abbasi Brothers and more

MADE FOR RENZE1577

See what your friends
are playing

FIND FRIENDS

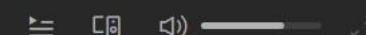
Julia Eger
Heartbeats
José González
The Indie Mix

Ben Khan
Night Changes
One Direction
One Direction

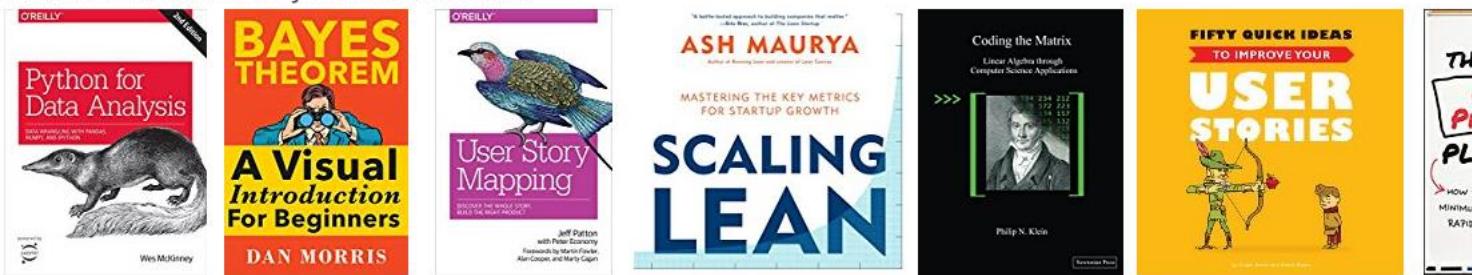
Sean Aquilina 12m
Flashing Lights
Kanye West
Graduation

Maddie Stocker 25m
Fake Plastic Trees
Radiohead
Radiohead Essential...

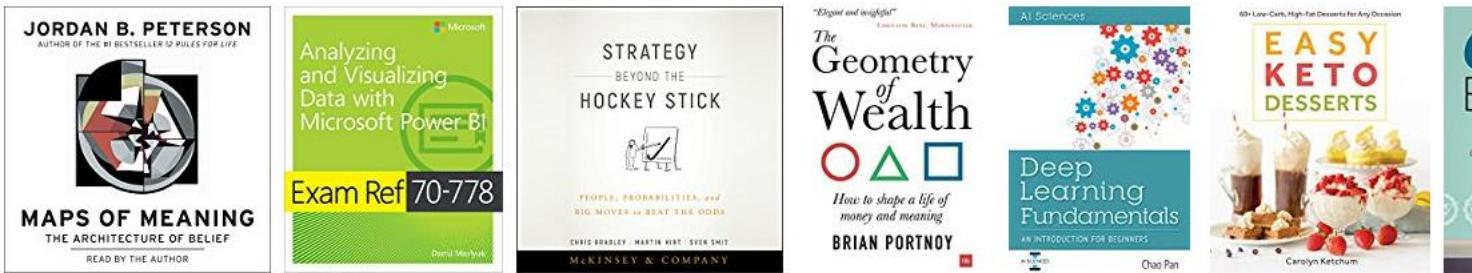
Laura Stephenson 1h



Recommendations for you in Kindle Store



New for you [See more](#)



New & exclusive products
[prime day launches](#)



Shop innovative tech
[exclusives](#)



Buy it again



Amazon
Gift Cards



Millions of items,
no expiration.

> Shop now

Advertisement

Tap Below to Reorder

Your Dash Buttons

[Learn more](#)

TAP HERE for info

TAP HERE to buy



[See all your Dash Buttons](#)

Prime Members
read **free**

[prime reading](#)





SEARCH

MENU

- Suspendisse
- Velit Sit Amet
- Pellentesque
- Voluptat
- Pretium Nulla
- Eget Turpis

PROJECTS

- Nunc venenatis
- Maecenas
- Curabitur

Dashboard > Ornare Orci Velit

Consectetur Adipiscing

Donec sollicitudin, ipsum non

\$1,310

Nunc et metus magna



Aenean Porta

Donec sollicitudin, ipsum non

\$24,815

Nunc et metus magna



Aenean Porta

Donec sollicitudin, ipsum non

94.2%

Nunc et metus magna



Aenean Porta

Donec sollicitudin, ipsum non

Nunc et metus magna

Integer

Nam ultrices

Phasellus vulputate

Vivamus

< NULLAM SAPIEN >

Mon	Tue	Wed	Thu	Fri	Sat	Sun
30	31	1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31	1	2

Nullam Ac Leo Vitae

5000

4000

3000

2000

1000

JAN

FEB

1,283

Cursus Id Arcu Eget

3,155

Cursus Id Arcu Eget

MAR

APRIL

MAY

JUN

JULY

Lorem Ipsum

Nunc faucibus dignissim nibh

Cras Pharetra

Aliquam

Nunc faucibus dignissim nibh

Donec orci nulla

Phasellus Interdum

Dignissim nibh

872

Lobortis non nisi quis

650

Bibendum

Nunc Leo Tortor

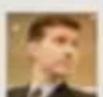
Nunc faucibus dignissim nibh

Felis ligula rutrum orci 2016

Phasellus interdum ligula eget turpis

Sed id egestas 2016

Wednesday September 16, 2015



Show Feed

Refresh

Customer Page

As of 9/16/2015 5:29 PM. Displaying data as Gerald Frazer.

Current Quarter Sales Predictions



Predicted Risk by Team

Sort: By Highest Value

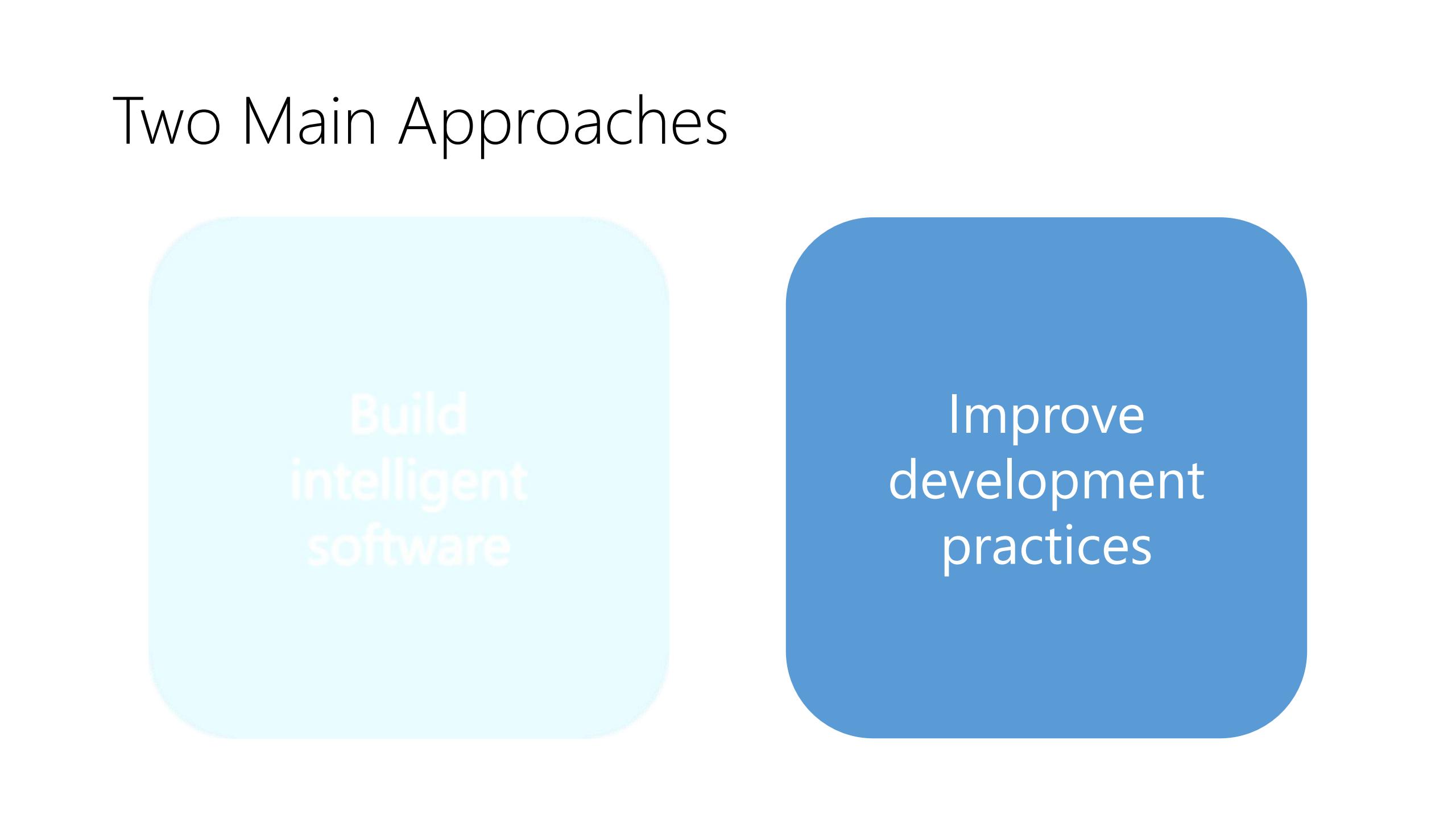


Changes Over 7 Days



Powered by
Microsoft Cortana Analytics Suite

Two Main Approaches

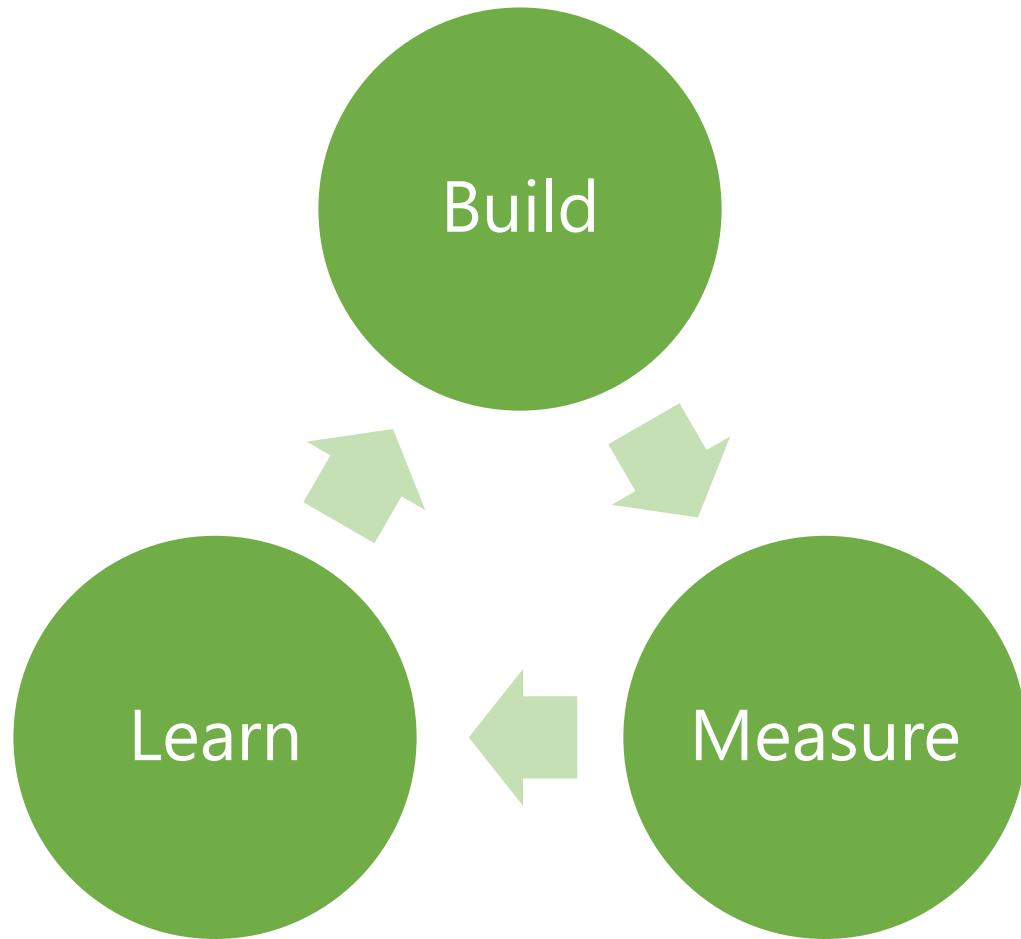


The diagram consists of a central title 'Two Main Approaches' above two rounded rectangular boxes. The left box is light blue and contains the text 'Build intelligent software'. The right box is dark blue and contains the text 'Improve development practices'.

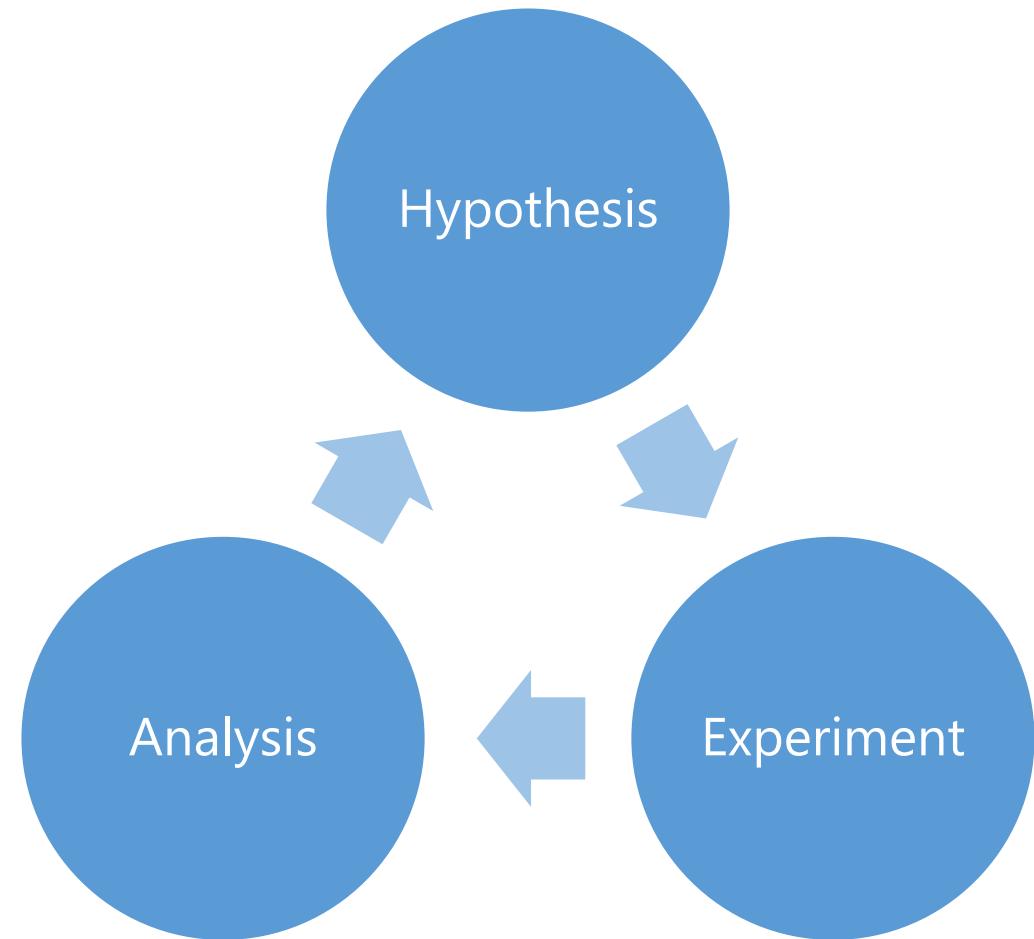
Build
intelligent
software

Improve
development
practices

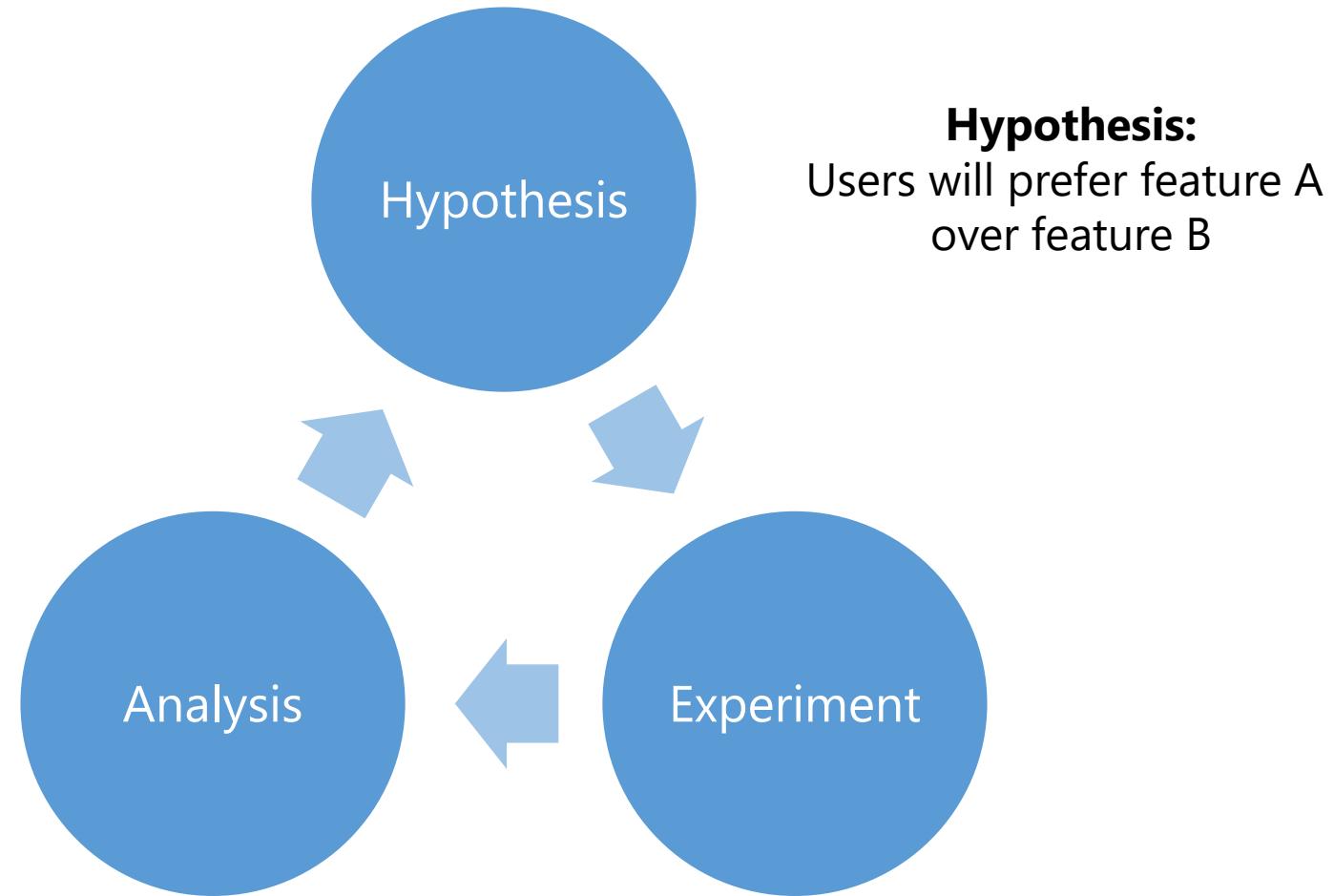
Data-Driven Decision Making



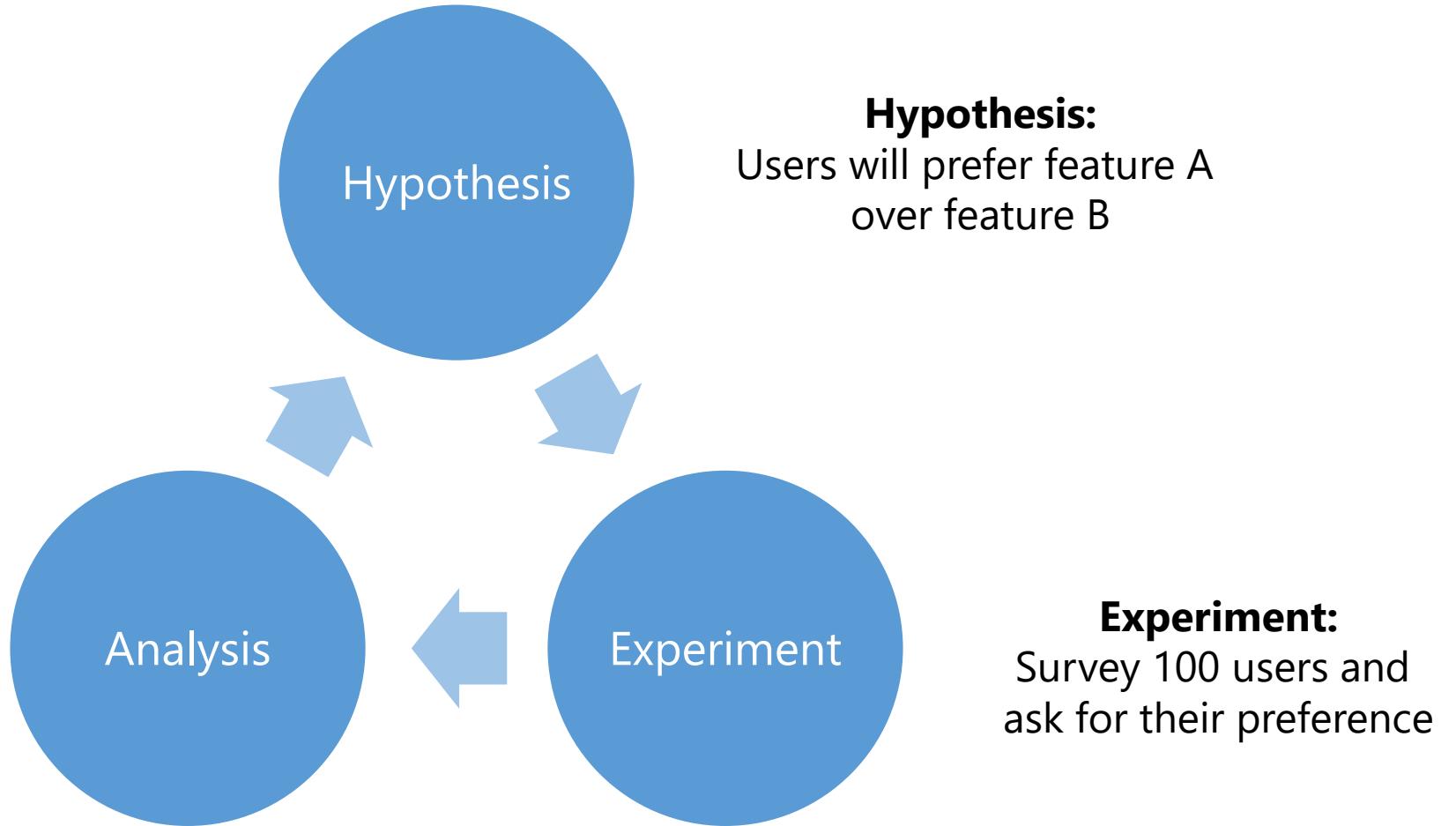
Hypothesis-Driven Development



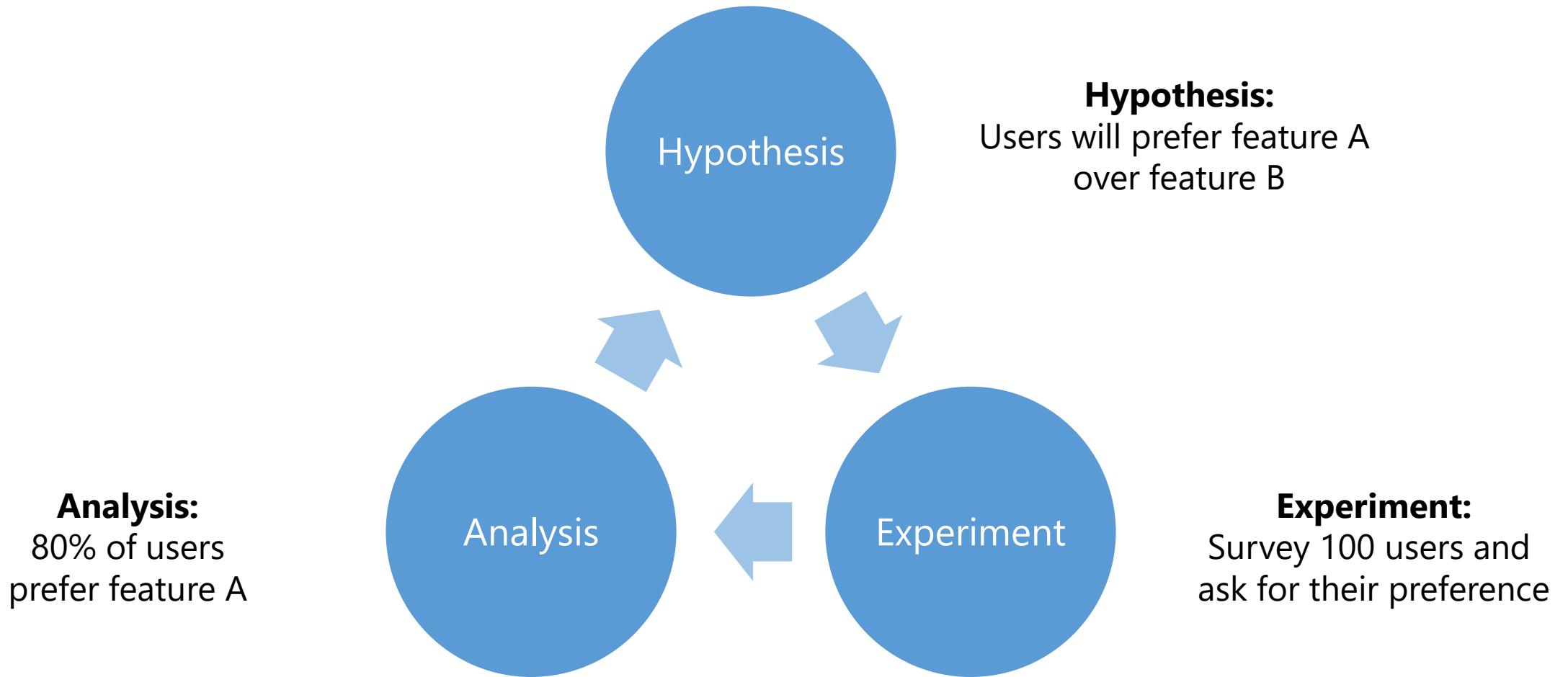
Hypothesis-Driven Development



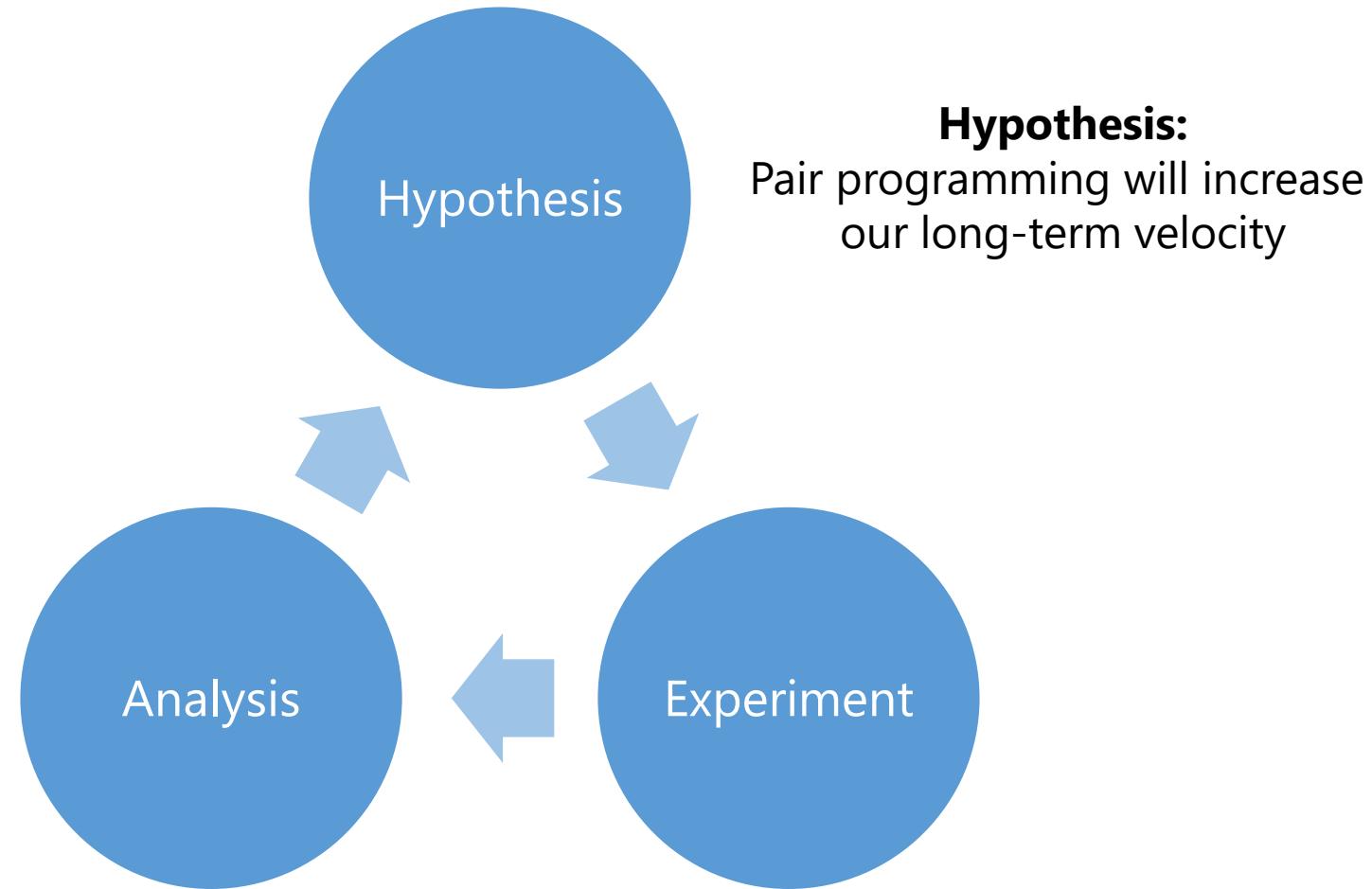
Hypothesis-Driven Development



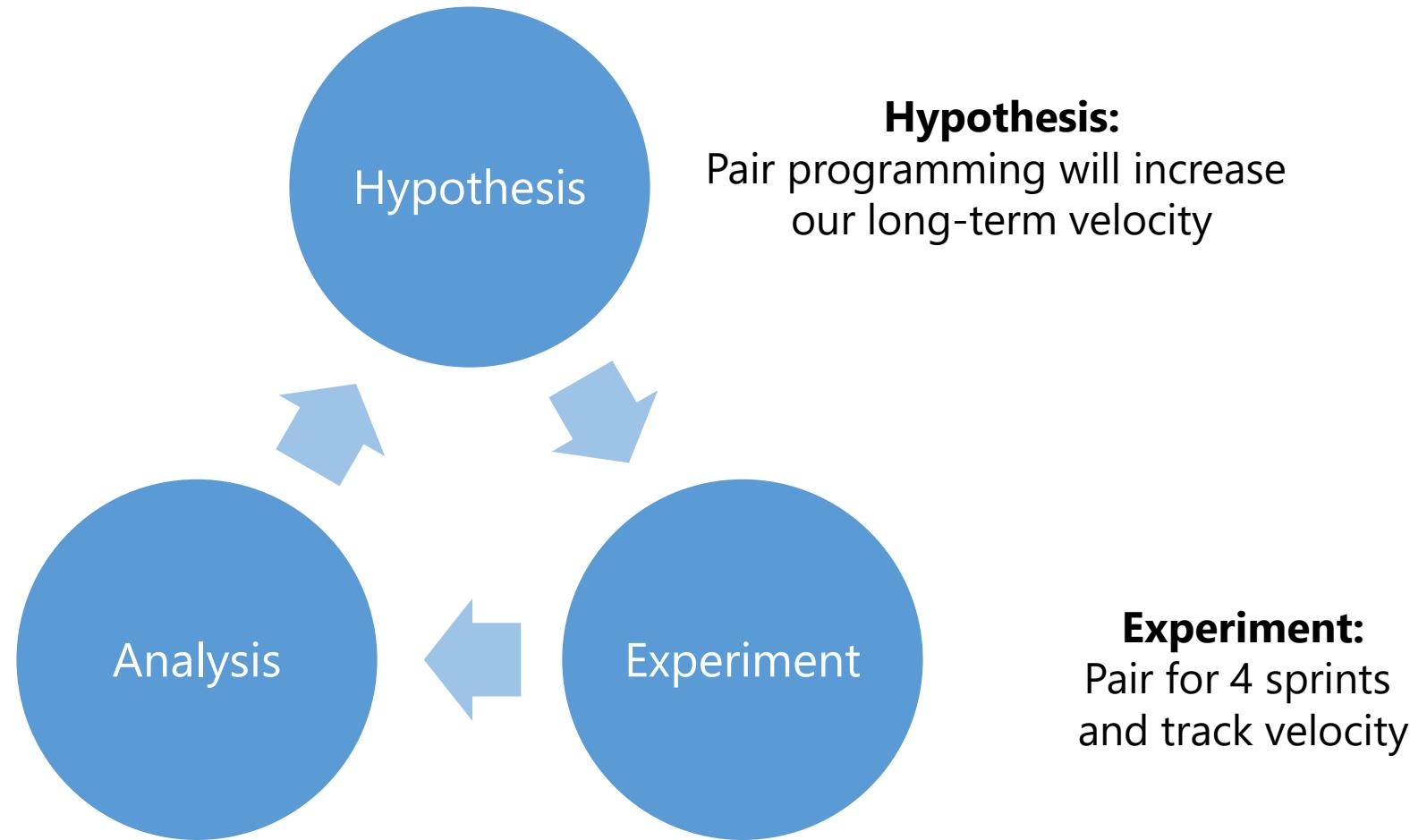
Hypothesis-Driven Development



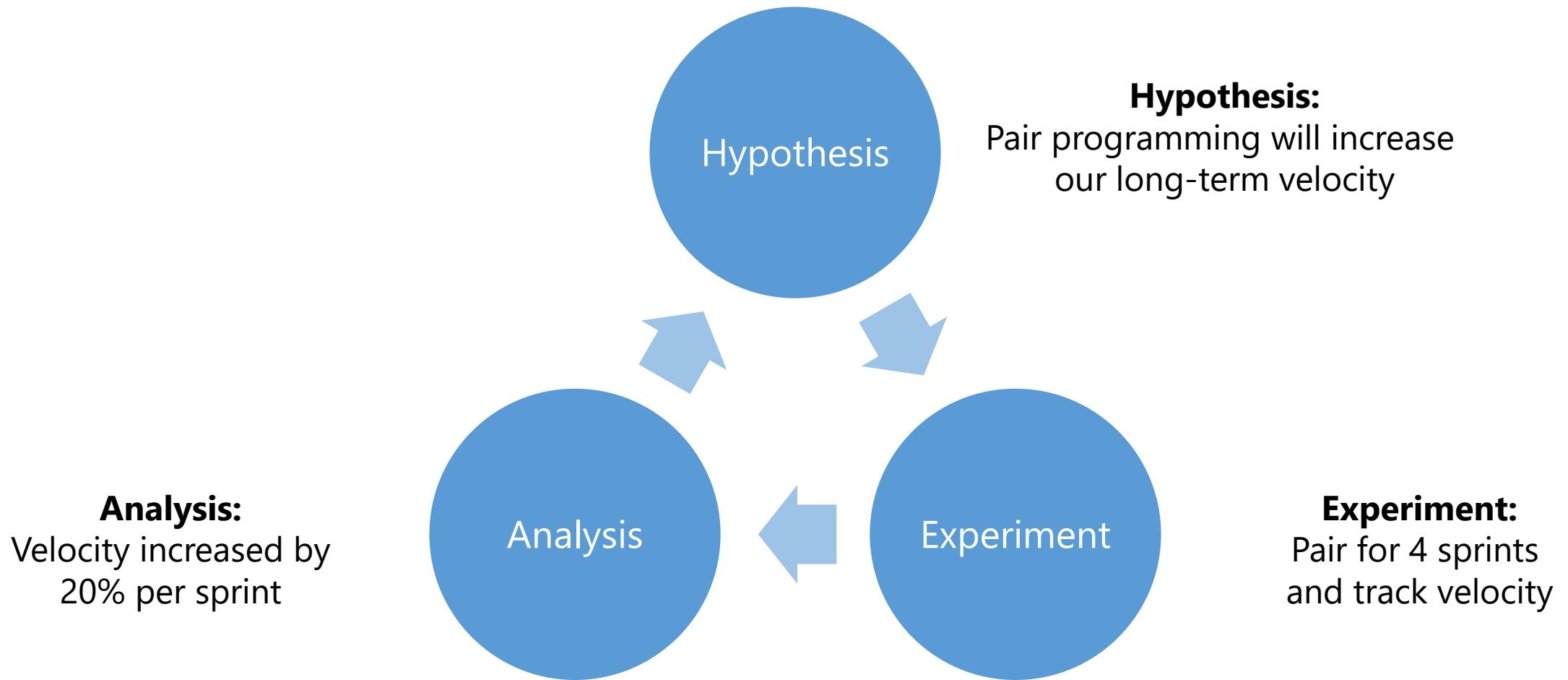
Hypothesis-Driven Development



Hypothesis-Driven Development



Hypothesis-Driven Development



Hypothesis Stories

<Hypothesis>

We assume that <hypothesis>

Will result in <outcome>

We will have succeeded when <measurable result>

Hypothesis Stories

Pair Programming Hypothesis

We assume that pair programming

Will result in higher long-term velocity

We will have succeeded when we have seen a 10%
or greater increase in velocity after 4 sprints.

Hypothesis Stories

"Kill Bill" Hypothesis

We assume that murdering Bill in his sleep

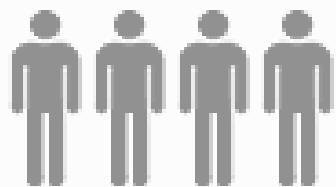
Will result in our codebase not sucking so bad

We will have succeeded when Bill is gone forever
and our code doesn't feel like the 7th level of hell

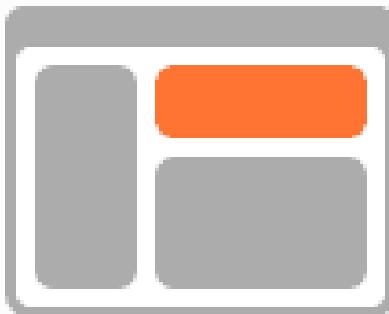
A/B Testing



A/B Testing



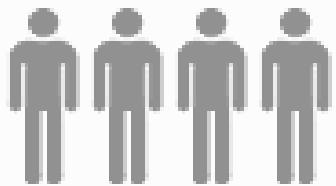
50 % visitors
see variation A



Variation A



23%
conversion



50 % visitors
see variation B

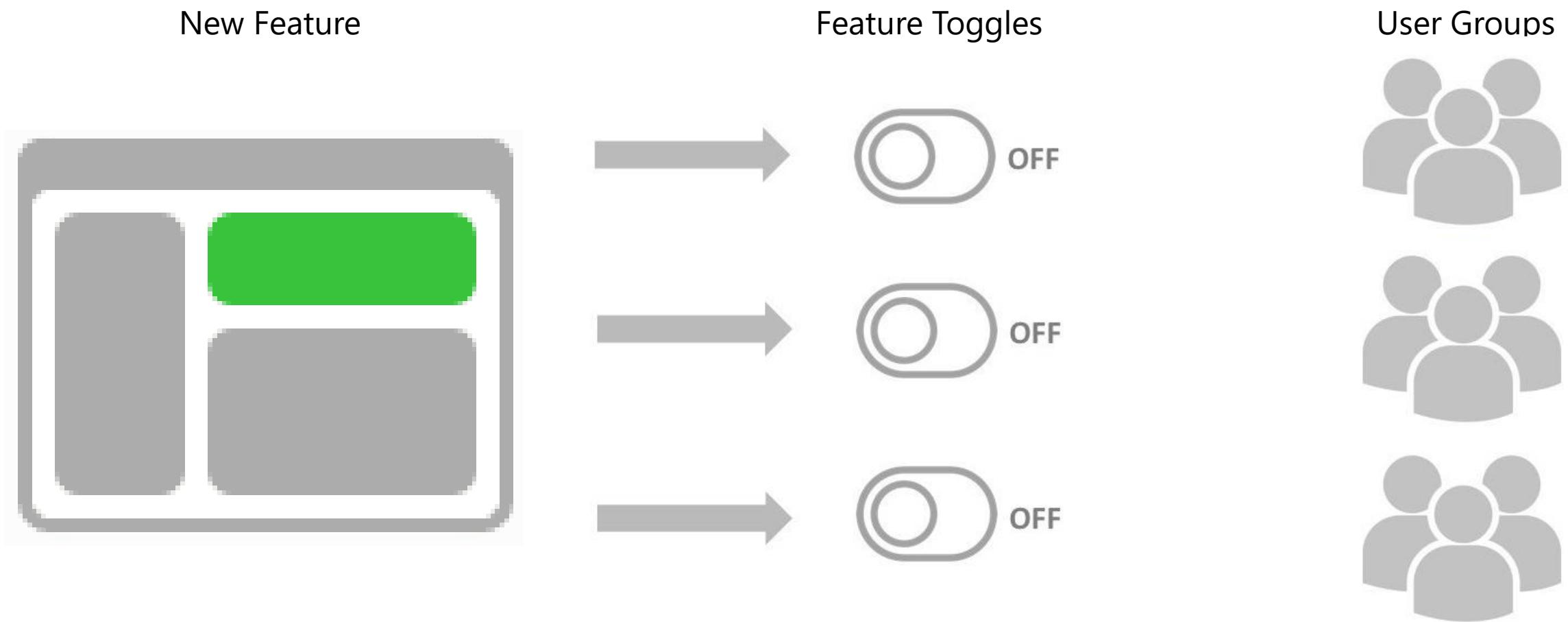


Variation B

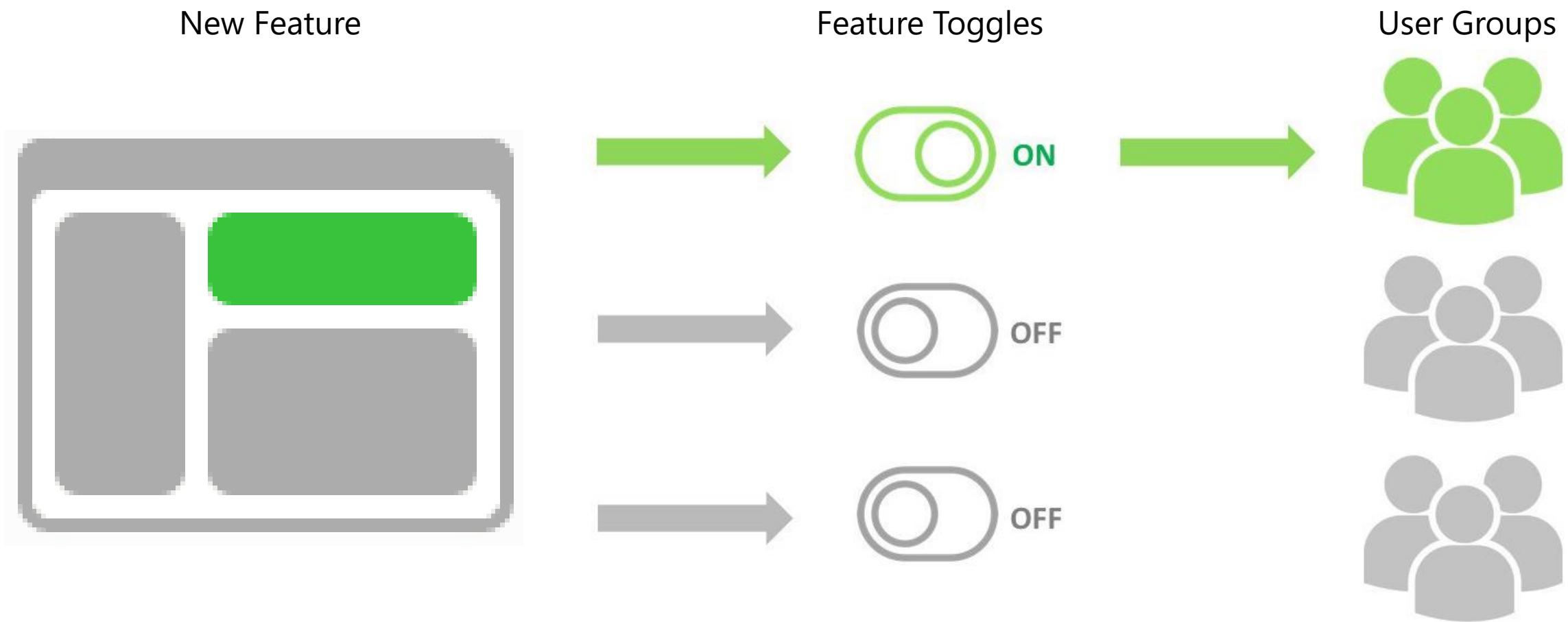


11%
conversion

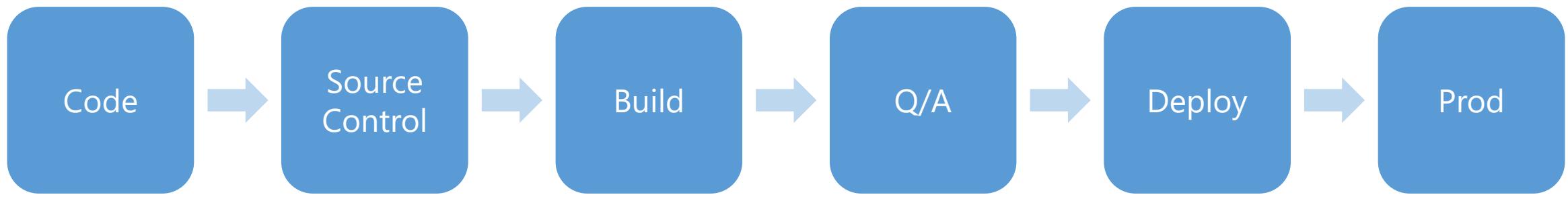
Feature Toggles



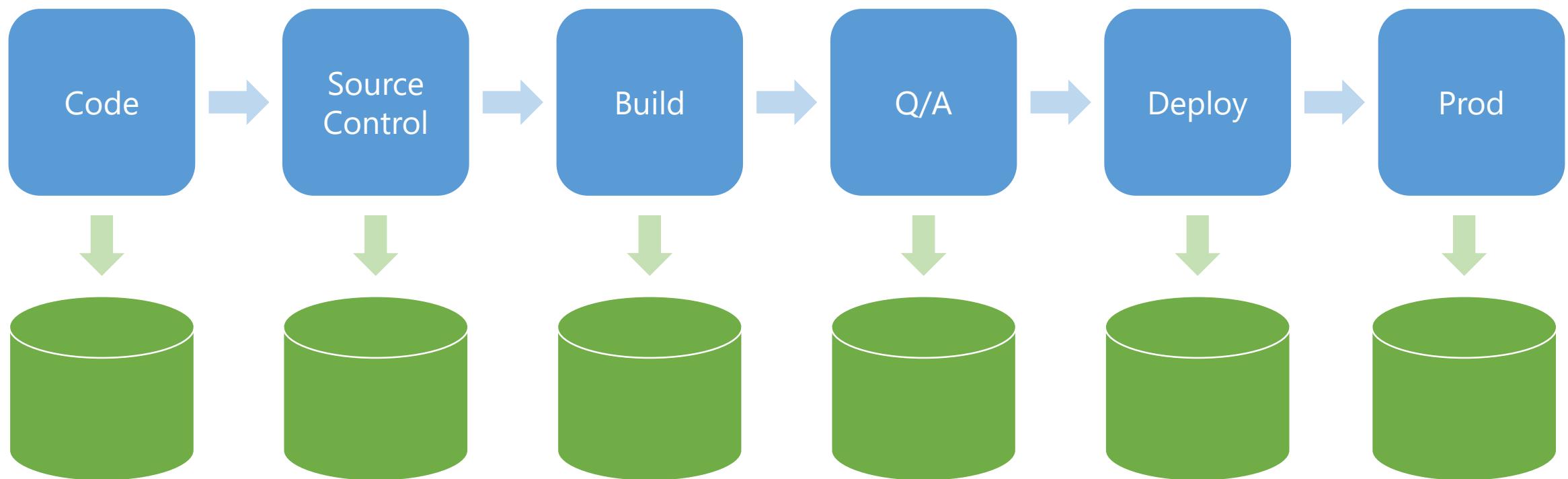
Feature Toggles



DevOps Pipeline



DevOps Pipeline



Code Quality Metrics

Source: NDepend

Code Metrics View



Error List

0 Errors

Description

Context-Sensitive Help

Show description of the method

In this Metric View, each rectangle represents a **method**. The area of a rectangle is proportional to **# lines of code (LOC)** of the corresponding method.

The color of a rectangle depends on its **method** value for the metric **Percentage Coverage**. Rectangles without such value remain gray.

Treemap helps see patterns that would be hard to spot with other ways

- Introduction
- Too Big - Too Complex
- View Code Coverage
- View Code Structure
- Pinpoint Where Are Flaws

Source: NDepend

Solution Explorer

Search Solution Explorer (Ctrl+\$)

Solution 'NDepend.All.v6' (27 projects)

- App
 - NDepend.Analysis
 - NDepend.API
 - NDepend.Console
 - NDepend.Core
 - NDepend.Platform.DotNet
 - NDepend.ReflectorAddIn
 - NDepend.UI
 - NDepend.VisualStudioExtension
 - NDepend.VisualStudioExtension.DotNet
 - NDepend.VisualStudioExtension.Impl
 - NDepend.VisualStudioExtension.Install
 - VisualNDepend
- Integration
- Test

Properties Code Ana... Solution... Find Resu...

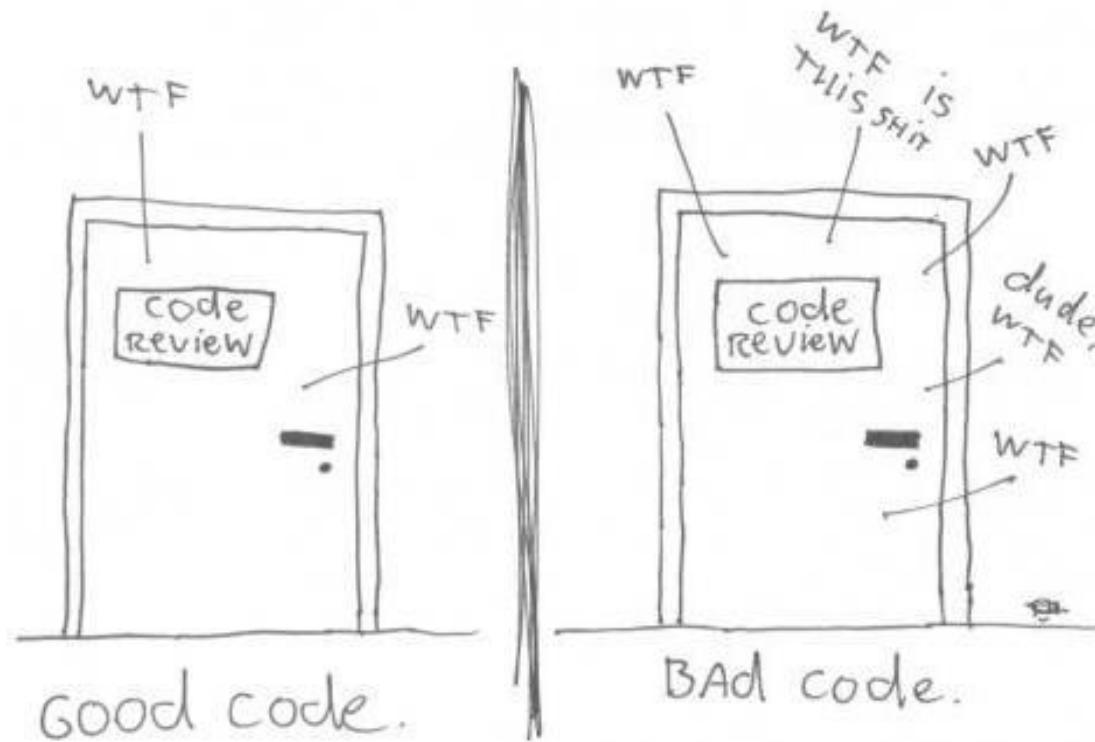


How maintainable is our code?

How complex is our code?

How should we organize our code?

The ONLY VALID MEASUREMENT OF CODE QUALITY: WTFs/MINUTE



(c) 2008 Focus Shift

Source Control Metrics



Search or jump to...

Pull requests Issues Marketplace Explore



matthewrenze / matthewrenze.com

[Unwatch](#) 3[Unstar](#) 3

Fork 2

Code

Issues 0

Pull requests 0

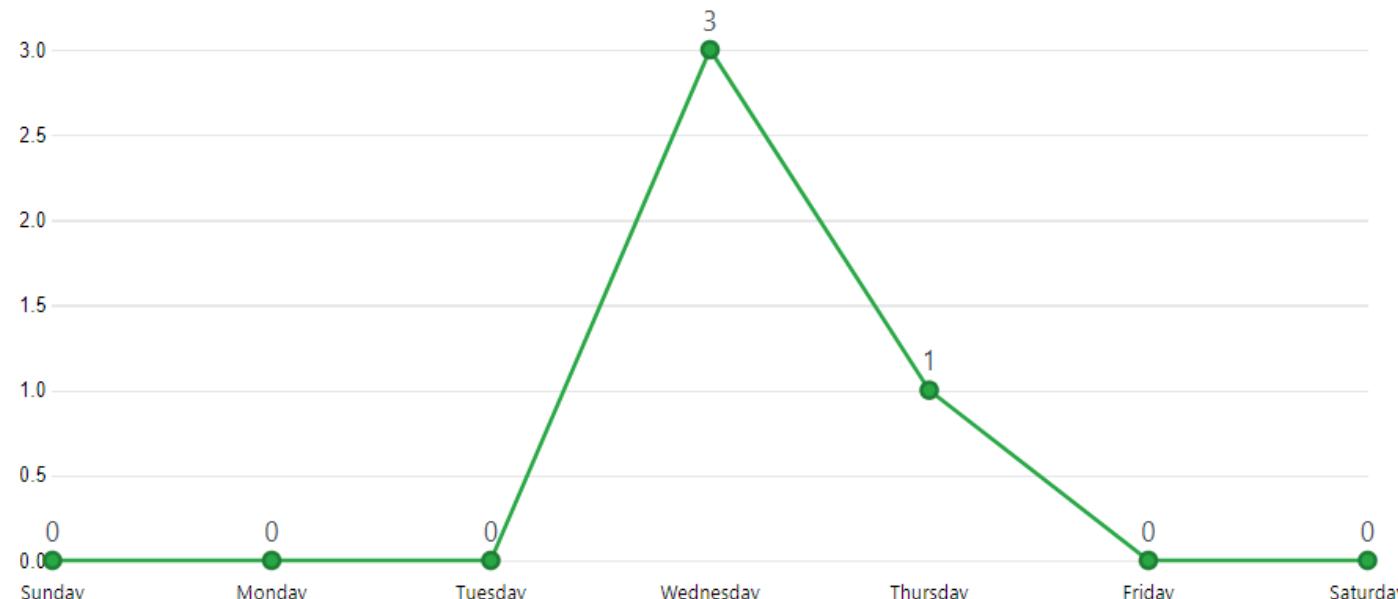
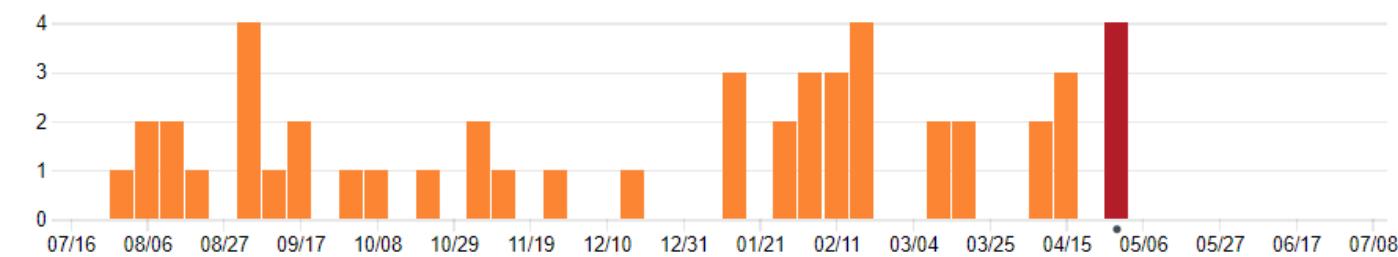
Projects 0

Wiki

Insights

Settings

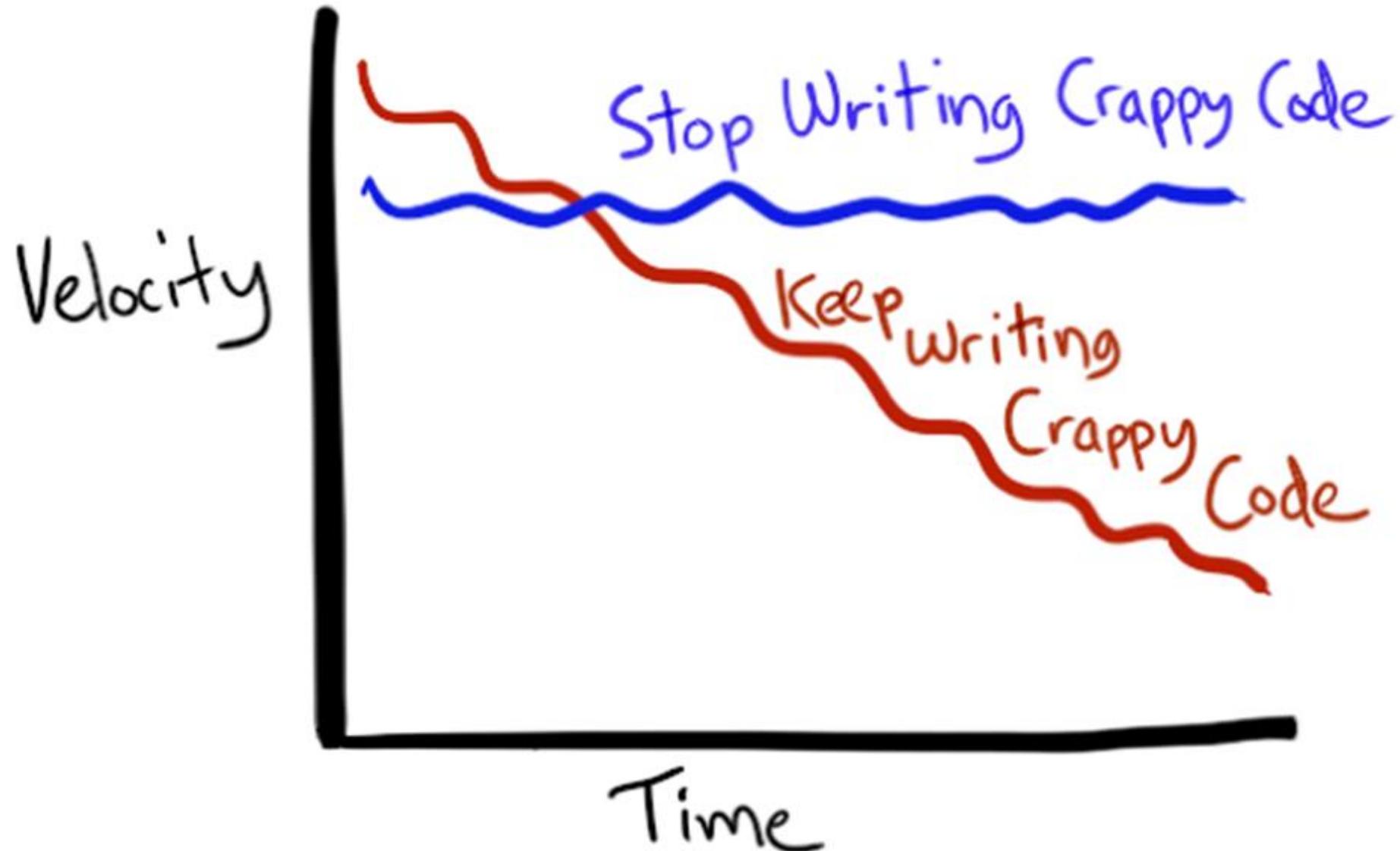
- Pulse
- Contributors
- Community
- Traffic
- Commits
- Code frequency
- Dependency graph
- Network
- Forks



How often are we checking in code?

Which files are modified most frequently?

How is our code evolving over time?



Source: <http://talkingincode.com/wp-content/uploads/2014/12/crapcode.png>

Build Metrics

Build 86

Build

- ✓ Get sources
- ✓ Build solution ***.sln
- ✓ Build TestPkg
- ✓ Test TestPkg
- ✓ Publish Test Results *.xml
- ✓ Copy Publish Artifact: Database_Pac...

OctoFX Database build / Build 86

[Queue new build...](#) [Download all logs as zip](#)

Build Succeeded



Build 86

Ran for 61 seconds (Default), completed 19.9 hours ago

[Summary](#) [Timeline](#) [Artifacts](#) [Tests](#) [Database deployment preview](#)

Build details

Definition	OctoFX Database build (edit)
Source branch	refs/heads/master
Source version	Commit 223852
Requested by	John Walley
Queued	08 March 2016 14:29:15
Started	08 March 2016 14:29:17
Finished	08 March 2016 14:30:19

Associated changes

No changes associated with this build.

Tags

[Add...](#)

Test Results

Total tests	Passed (222)	Failed tests	Pass percentage	Run duration
	222 (+148)	■ Passed (222) ■ Failed (0) ■ Others (0)	100% (+0%)	30s 120ms (+19s 341ms)

[Detailed report >](#)

Code Coverage

No build code coverage data available.

Associated work items

No associated work items found for this build

Database deployment preview

Migrations pending

1. Migrations\0002_20160301-2020_John.Walley.sql
2. Migrations\0003_20160304-1024_John.Walley.sql
3. Migrations\0004_20160304-1333_John.Walley.sql

This list is based on [.\SQL2014].[OctoFX-Test] at 2016-03-08 14:29:29.

The actual list of migrations to be deployed may vary depending on the state of the target environment.

Table changes

- [» tSQLt.TestResult](#)
- [» tSQLt.TestMessage](#)
- [» tSQLt.Run_LastExecution](#)
- [» tSQLt.Private_EqualsTableSchema_Actual](#)
- [» tSQLt.Private_RenamedObjectLog](#)
- [» tSQLt.Private_NullCellTable](#)
- [» tSQLt.Private_Configurations](#)
- [» tSQLt.Private_EqualsTableSchema_Expected](#)
- [» tSQLt.CaptureOutputLog](#)
- [» tSQLt.Private_ExpectException](#)
- [» dbo.Account](#)

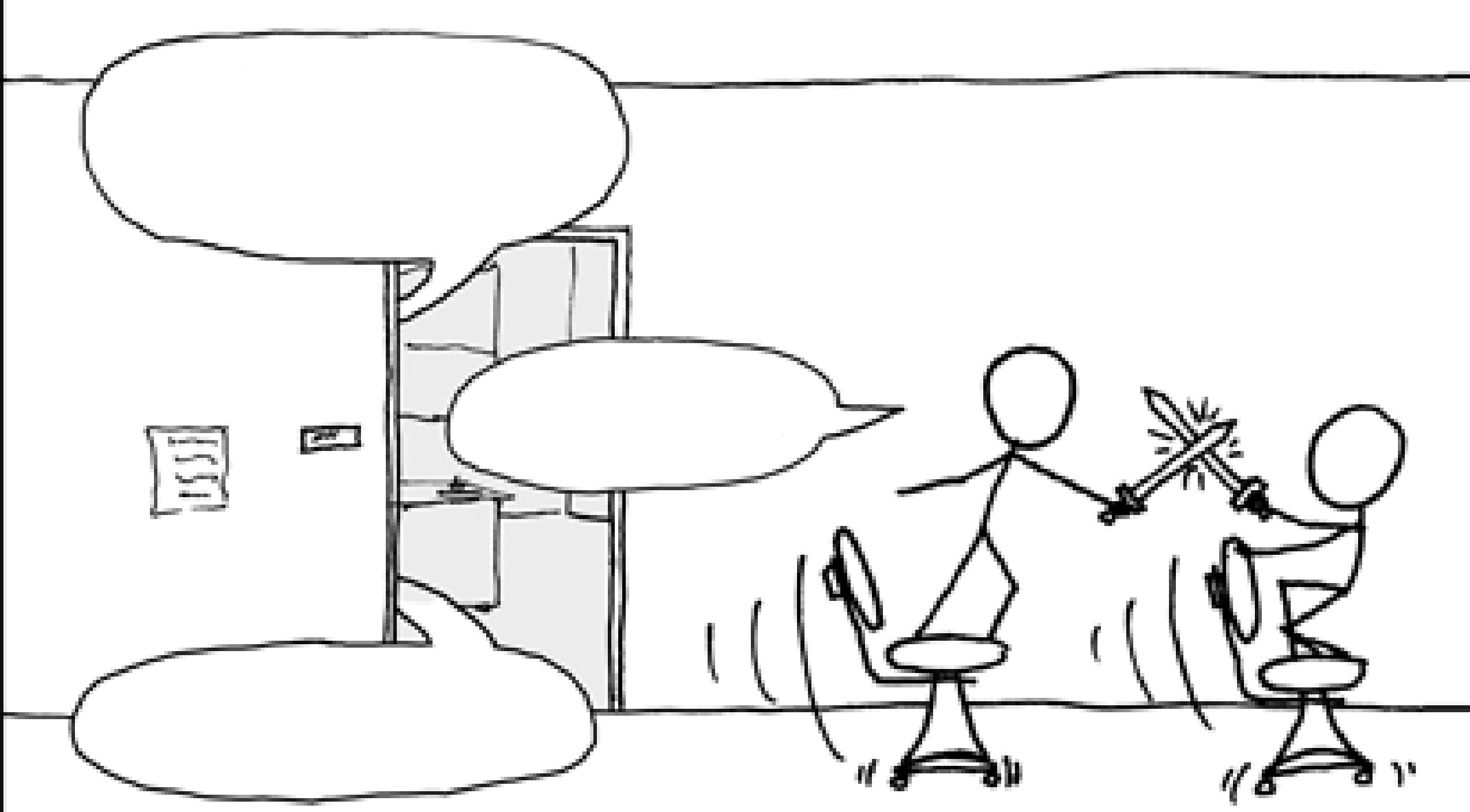
Other object changes

- [» \(Assembly\) dbo.tSQLtCLR](#)
- [» \(Function\) tSQLt.Private_GetForeignKeyParColumns](#)
- [» \(Function\) tSQLt.TestCaseSummary](#)
- [» \(Function\) tSQLt.Private_SqlVersion](#)
- [» \(Function\) tSQLt.Private_SqlVariantFormatter](#)
- [» \(Function\) tSQLt.Private_ScriptIndex](#)
- [» \(Function\) tSQLt.Private_ResolveSchemaName](#)
- [» \(Function\) tSQLt.Private_ResolveObjectName](#)
- [» \(Function\) tSQLt.Private_DatabaseName](#)

Are there warnings to look into?

How long are our builds taking?

Would faster machines save money?



Q/A Metrics

Type to search

► ✓ Domain.Tests	(536 tests)	Success
► ✓ Infrastructure.Tests	(5 tests)	Success
► ✓ Persistence.Tests	(132 tests)	Success
► ✓ Presentation.Tests	(423 tests)	Success
▲ ✓ Application.Tests	(369 tests)	Success
▲ ✓ DataExplorer.Application.Tests	(369 tests)	Success
► ✓ Application	(24 tests)	Success
► ✓ Clipboard	(4 tests)	Success
► ✓ Columns	(3 tests)	Success
► ✓ Core	(36 tests)	Success
► ✓ FilterTrees	(8 tests)	Success
► ✓ Filters	(6 tests)	Success
► ✓ Importers.CsvFile	(23 tests)	Success
► ✓ Layouts	(91 tests)	Success
► ✓ Legends	(41 tests)	Success
► ✓ Projects.Commands	(17 tests)	Success
► ✓ Rows	(6 tests)	Success
► ✓ Views.ScatterPlots	(107 tests)	Success
► ✓ Web	(3 tests)	Success
▲ ✗ DataExplorer.Specs	(5 tests)	Failed: 3 tests failed
▲ ✗ DataExplorer.Specs	(5 tests)	Failed: 3 tests failed
▲ ✗ Application	(1 test)	Failed: 1 test failed
▲ ✗ ExitApplicationFeature	(1 test)	Failed: One or more child tests had errors: 1 test failed
✗ ExitTheApplication		Failed: System.IO.DirectoryNotFoundException : Could no
✗ Importers	(1 test)	Failed: 1 test failed
✗ Projects	(2 tests)	Failed: 1 test failed
▲ ✗ ScatterPlots	(1 test)	Inconclusive
▲ ✗ ViewScatterPlotFeature	(1 test)	Inconclusive
✗ SinglePointAt00		Inconclusive: One or more step definitions are not implem

ExitTheApplication failed

Exit the Application

```
System.IO.DirectoryNotFoundException : Could not find a part of the
path 'C:\Users\Matthew\AppData\Local\Data Explorer\Logs
\2018-07-11.txt'.
at System.IO.__Error.WinIOError(Int32 errorCode, String
maybeFullPath)
at System.IO.FileStream.Init(String path, FileMode mode, FileAccess
access, Int32 rights, Boolean useRights, FileShare share, Int32
bufferSize, FileOptions options, SECURITY_ATTRIBUTES secAttrs,
String msgPath, Boolean bFromProxy, Boolean useLongPath, Boolean
checkHost)
at System.IO.FileStream..ctor(String path, FileMode mode,
FileAccess access, FileShare share, Int32 bufferSize, FileOptions
options, String msgPath, Boolean bFromProxy, Boolean
useLongPath, Boolean checkHost)
at System.IO.StreamWriter.CreateFile(string path, Boolean append,
Boolean checkHost)
at System.IO.StreamWriter..ctor(String path, Boolean append,
Encoding encoding, Int32 bufferSize, Boolean checkHost)
at System.IO.StreamWriter..ctor(String path, Boolean append,
Encoding encoding)
at System.IO.File.InternalAppendAllText(String path, String
contents, Encoding encoding)
at System.IO.File.AppendAllText(String path, String contents)
at DataExplorer.Infrastructure.Logs.LogFile.WriteLine(String filePath,
String text) in C:\Users\Matthew\Dropbox\Professional\Software
\Data Explorer\Source\DataExplorer\Infrastructure\Logs
\LogFile.cs:line 23
at DataExplorer.Application.Core.Logs.LogProvider.WriteLine(DateTime
dateTime, Int32 threadId, LogLevel level, String message) in C:
\Users\Matthew\Dropbox\Professional\Software\Data Explorer\Source
```

How much of our code is covered by tests?

How long does it take to run our tests?

Are we actually benefiting from TDD?



Deployment Metrics

Projects



Artifact Example

[CREATE RELEASE](#)[Overview](#)[Process](#)[Variables](#)[Triggers](#)[Channels](#)[Releases](#)[Settings](#)

Release 0.0.1

Deploy Artifact Example release 0.0.1 to Local

⋮

TASK SUMMARY

TASK LOG

Task Progress

This task started 5 minutes ago and ran for 8 seconds

Deploy Artifact Example release 0.0.1 to Local

Step 1: Get host file from web server

Web01-hosts.txt

Step 2: Get host file from sql server

Sql01-hosts.txt

Artifacts

Web01-hosts.txt

Sql01-hosts.txt

Task History

When	Who	What
5 minutes ago	system	Deploy to Local (#2) succeeded for Artifact Example release 0.0.1 to Local

5 minutes ago	system	Deploy to Local (#2) started for Artifact Example release 0.0.1 to Local
---------------	--------	--

6 minutes ago	admin	Deployed Artifact Example release 0.0.1 to Local (#2)
---------------	-------	---

Did any of our deployments fail?

Why did our deployment fail?

Should we use continuous deployment?

:(
:(

Your PC ran into a problem and needs to restart. We're just
collecting some error info, and then we'll restart for you.

0% complete



For more information about this issue and possible
fixes, visit <http://windows.com/stopcode>

If you call a support person, give them this info:
Stop code: MANUALLY INITIATED CRASH

Software Telemetry

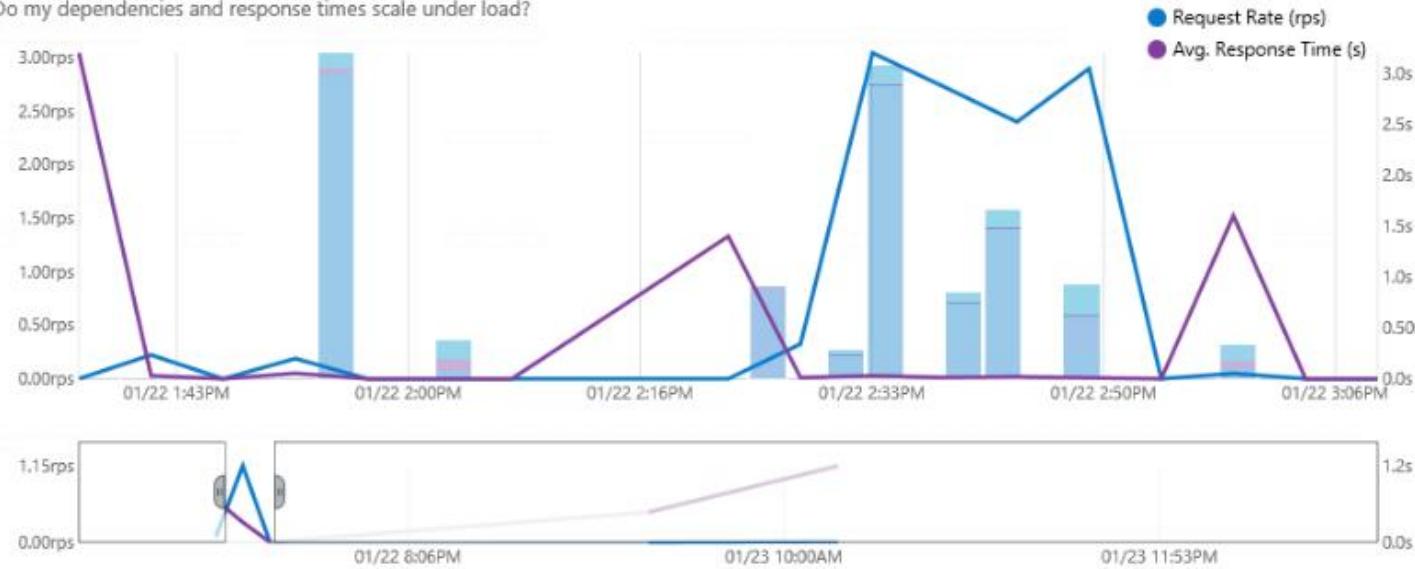
Server Performance

FF-Intranet-Test - Server Performance

Selected date range: Last 48 hours

Response Time and Load vs. Dependencies

Do my dependencies and response times scale under load?



Resource

- (localdb)\v11.0\FabrikamFiber-Express
- (localdb)\v11.0\master
- https://dc.services.visualstudio.com/

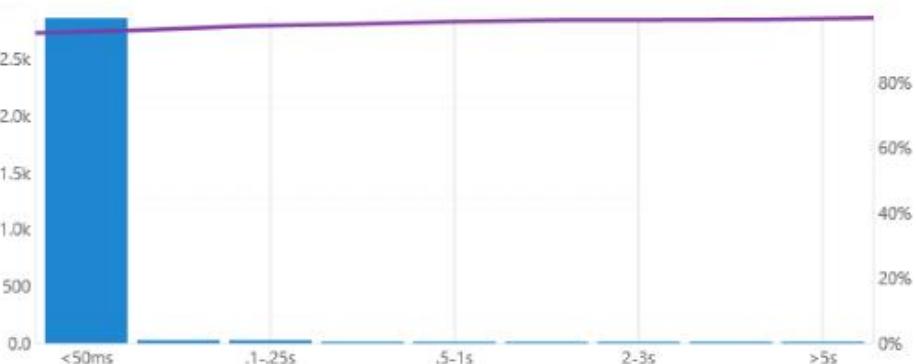
Type	Calls	Weight	Avg	Req/Sec
Sql	2.58k	86.72%	20ms	0.3932
Sql	18.0	2.1%	68ms	0.0027
HTTP	30.0	11.18%	219ms	0.0045

Top 10 slowest requests by issue count

Description	Avg	Max	Count	Last Date
MVC page : Home.Index [28,270 ms] slow at FabrikamFiber.DAL.Data.ServiceTicketRepository.AllIncluding() [...	26.51s	28.23s	7	2014/01/22 02:13:03 PM
MVC View rendering: Index [11,461 ms] slow at MVC View rendering: Index [11,461 ms]	10.09s	11.46s	7	2014/01/22 01:32:04 PM
MVC page : Home.Index [15,514 ms] slow at MVC page : Home.Index [15,514 ms]	15.51s	15.51s	1	2014/01/22 02:50:00 PM

Response time distribution

How many requests have high latency?



Exceptions Rate

How many exceptions are thrown per second?



CPU

How much CPU does my service consume?



Network

What is the network I/O rate for my service?



Memory In Use

How much memory does my service use?



Average Instance Count

How many instances does my service utilize?

Which features are being used?

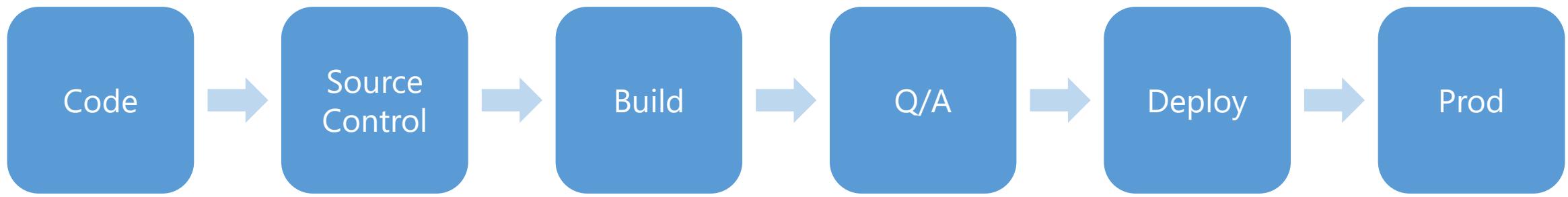
How long to complete a task?

Is someone hacking into our system?

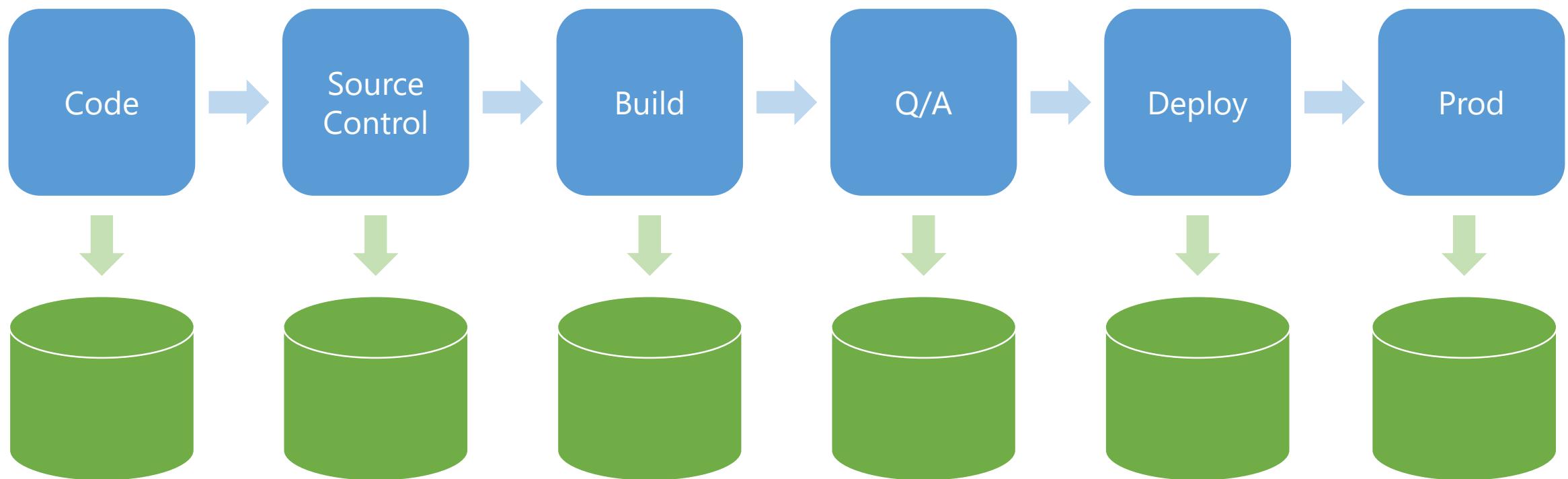


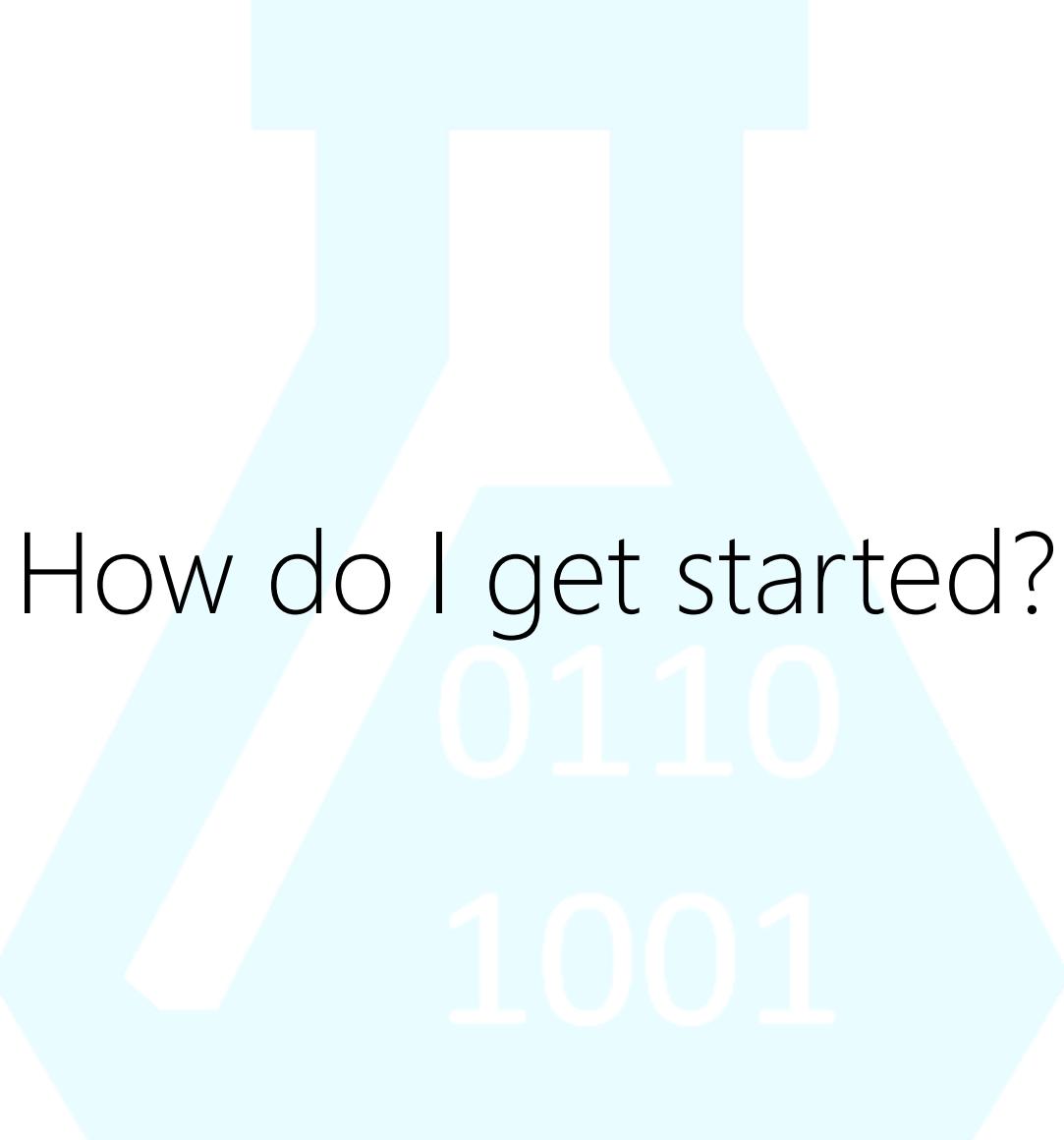
HACKERMAN

DevOps Pipeline



DevOps Pipeline

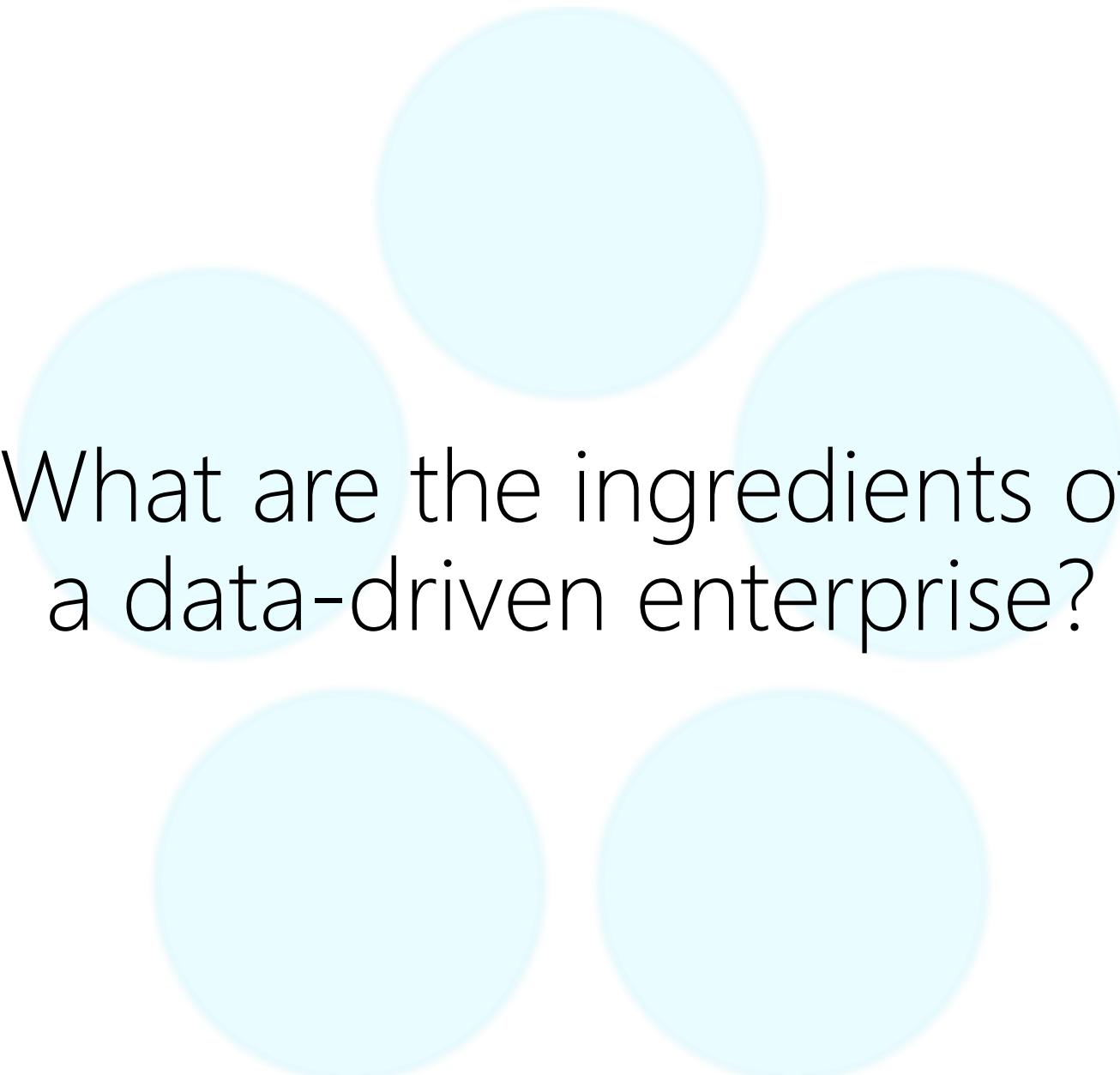




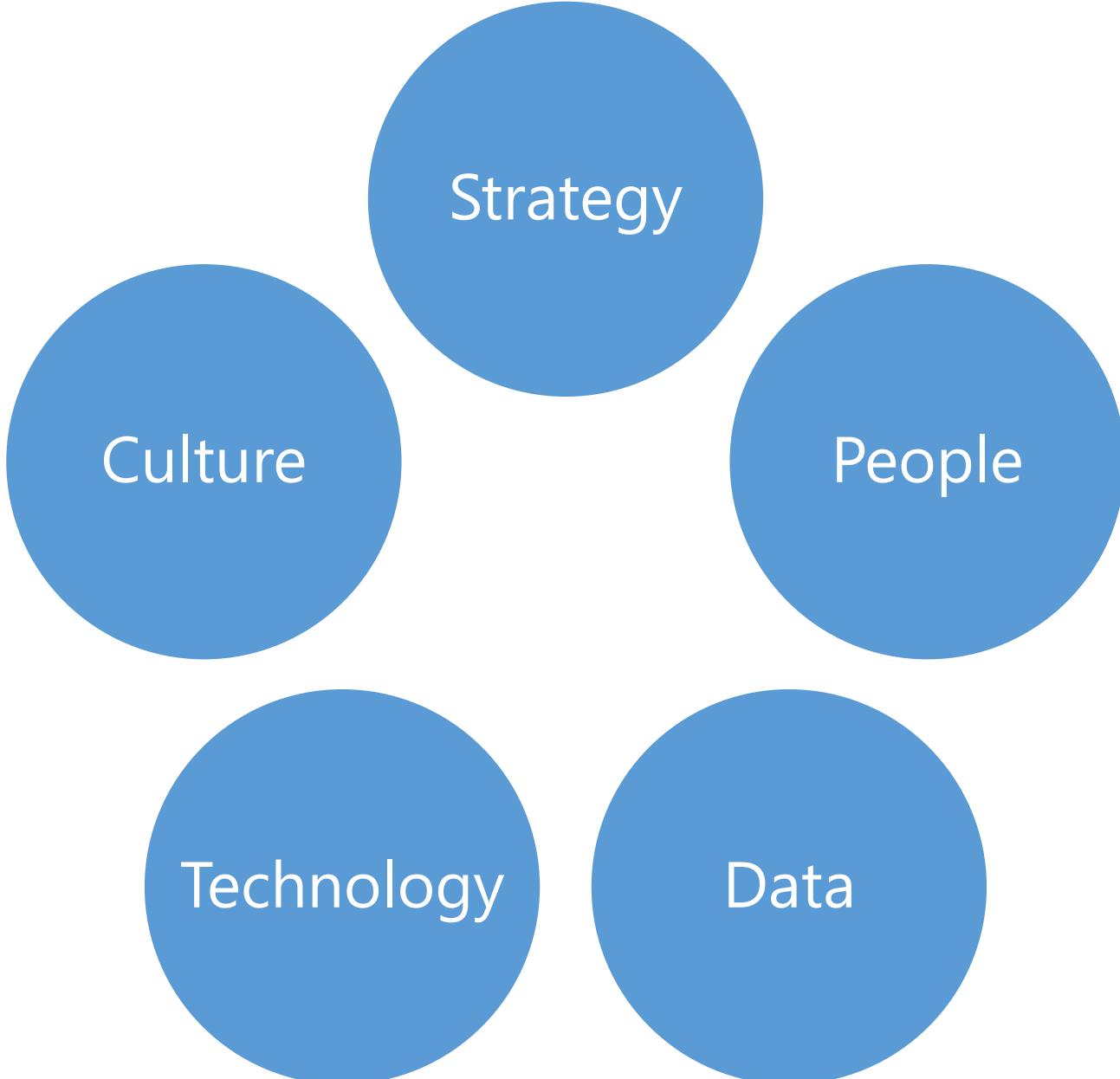
How do I get started?

0110

1001



What are the ingredients of
a data-driven enterprise?



Strategy

Culture

People

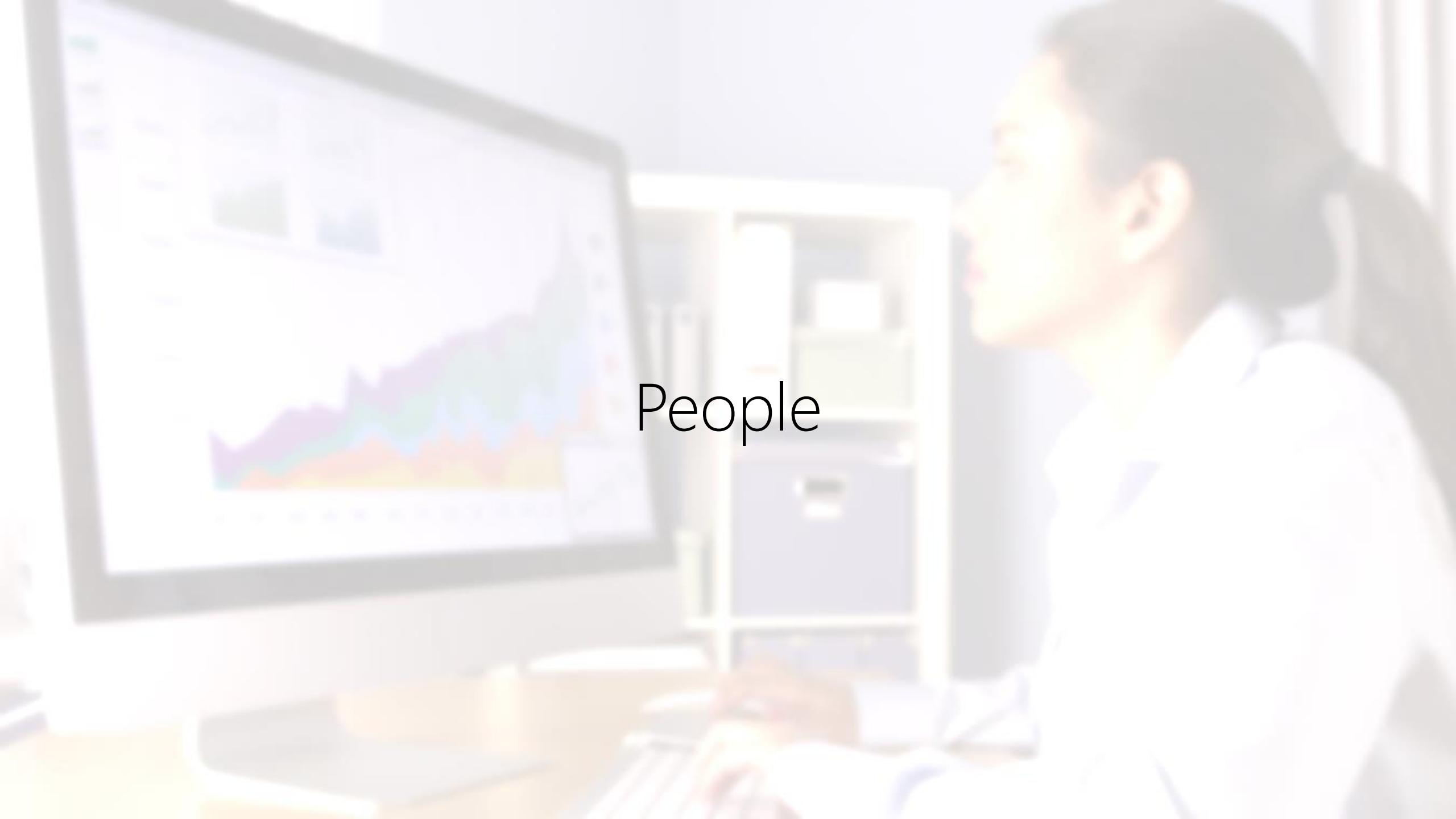
Technology

Data



Strategy



A man in a dark suit and white shirt is seated at a desk, looking intently at a large computer monitor. The monitor displays a complex, multi-colored bar chart with several peaks and troughs. The background is a blurred office environment.

People



Data



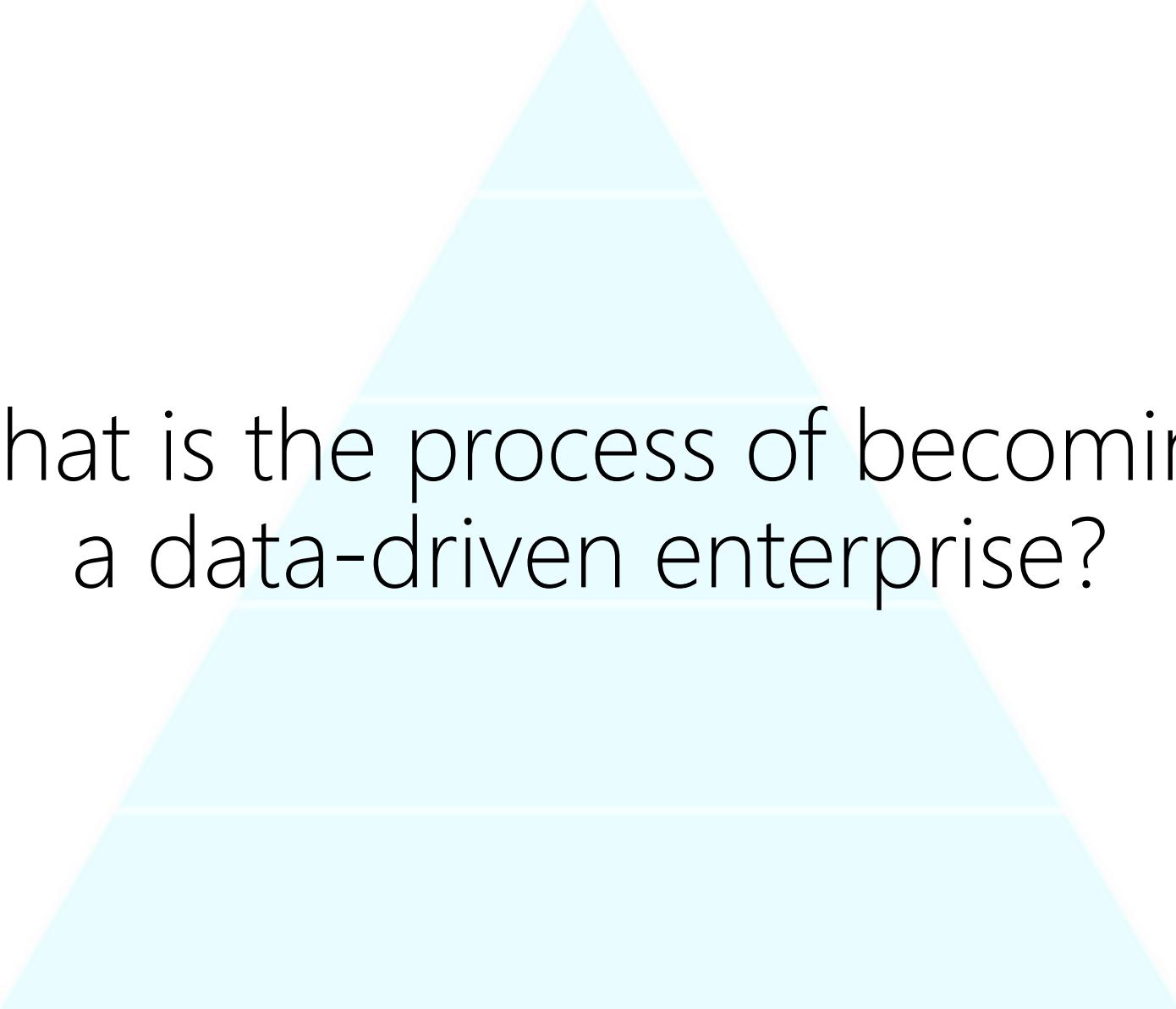
Technology



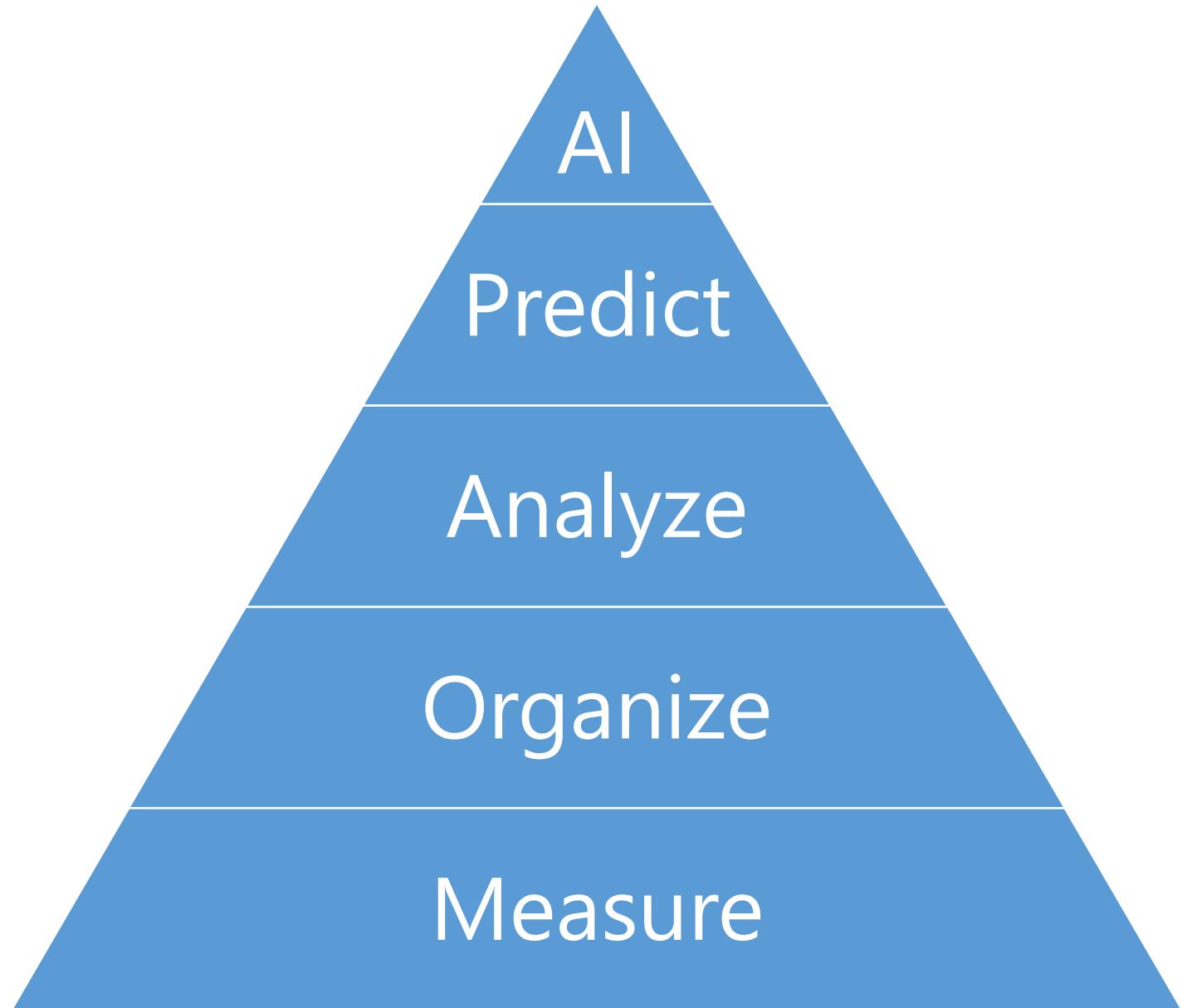
A blurred background image of a person sitting at a desk, working on a computer. There are multiple monitors visible, and the person appears to be looking at one of them. The overall scene is out of focus.

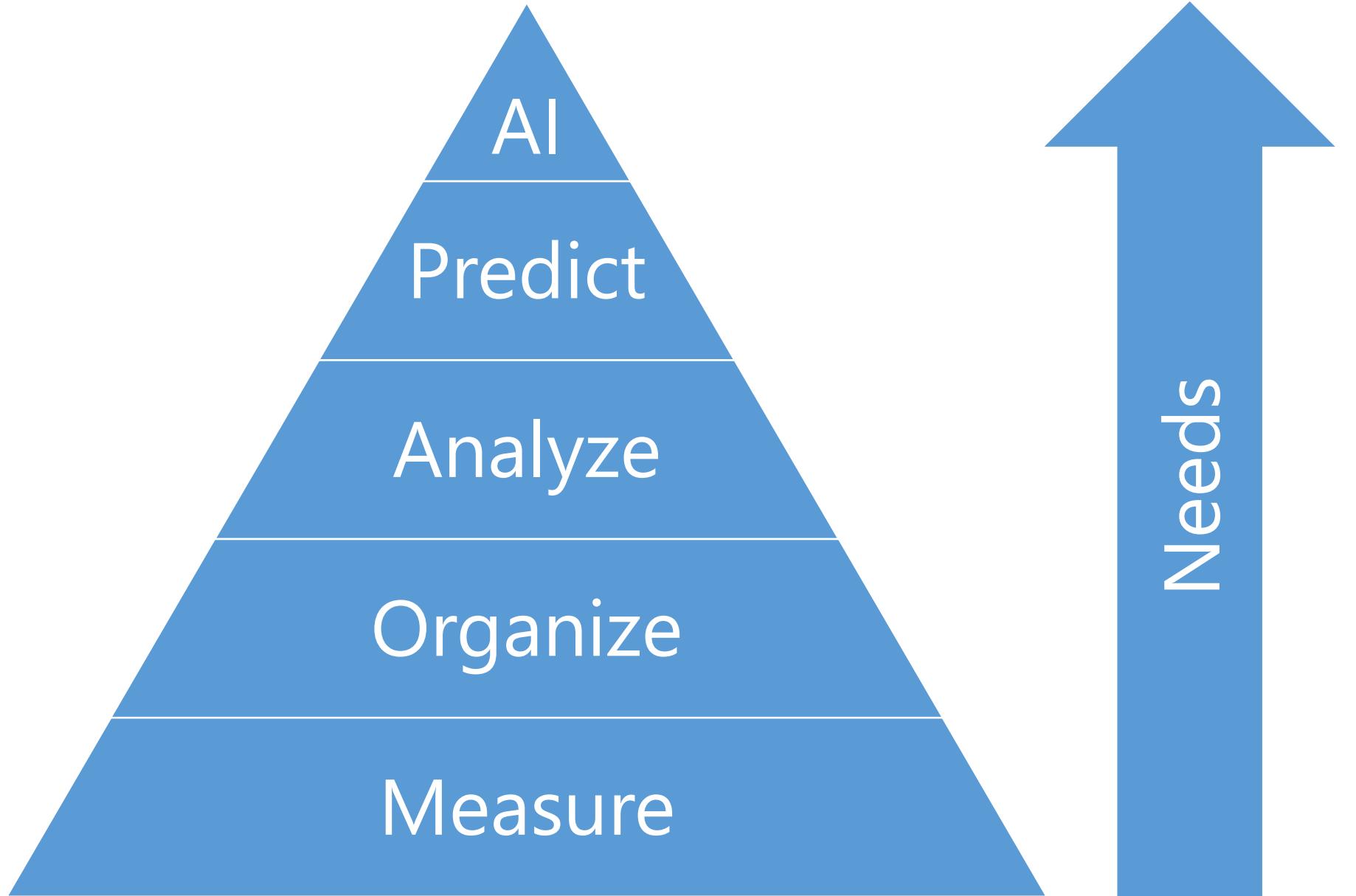
Culture





What is the process of becoming
a data-driven enterprise?





1. Measure



Measure

1. Measure

Transactions

Logging

Digitization



Measure

1. Measure

Transactions

Logging

Digitization

Telemetry

Experiments

External data



Measure

2. Organize



Organize



Measure

2. Organize

Transform

Clean

Store



Organize



Measure

2. Organize

Transform

Clean

Store

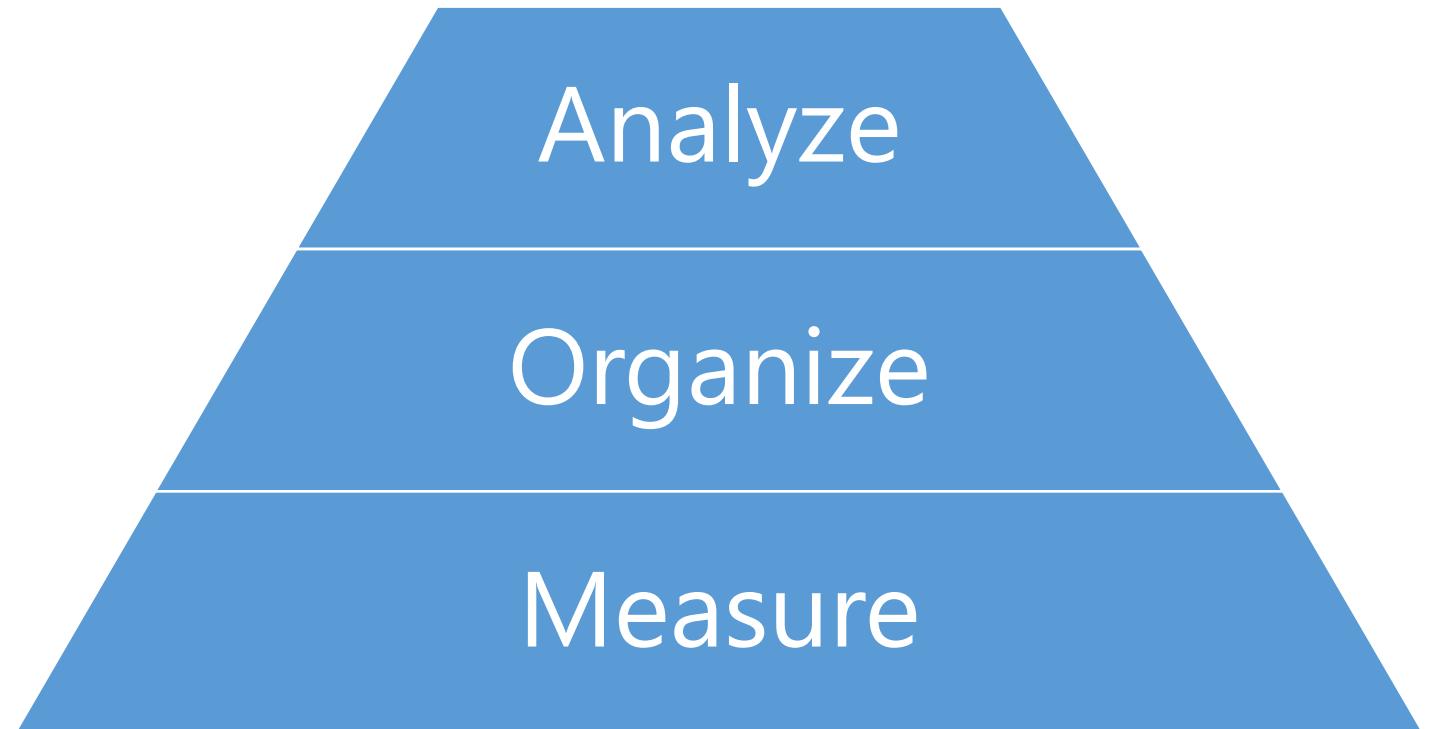
Data ETL

Data Warehouse

Data Lake



3. Analyze

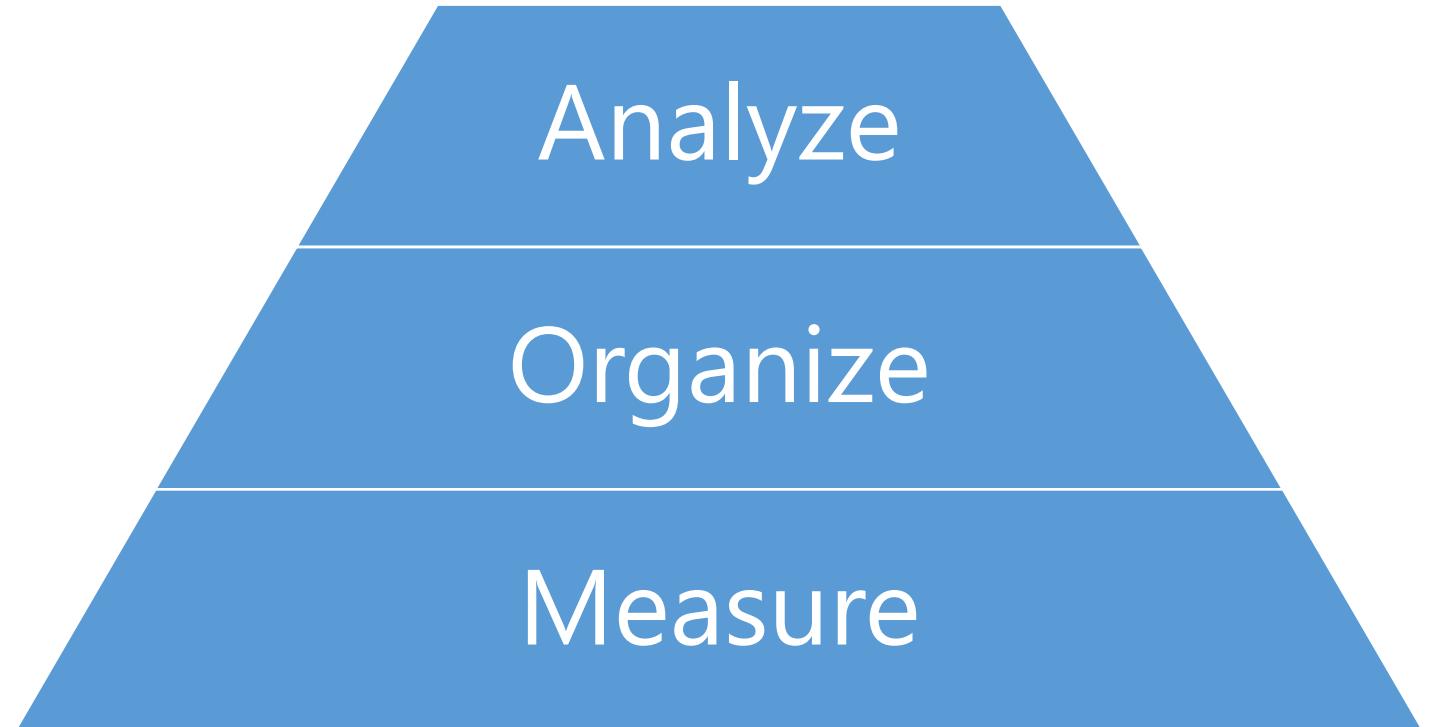


3. Analyze

Reports

Dashboards

KPI monitors



3. Analyze

Reports

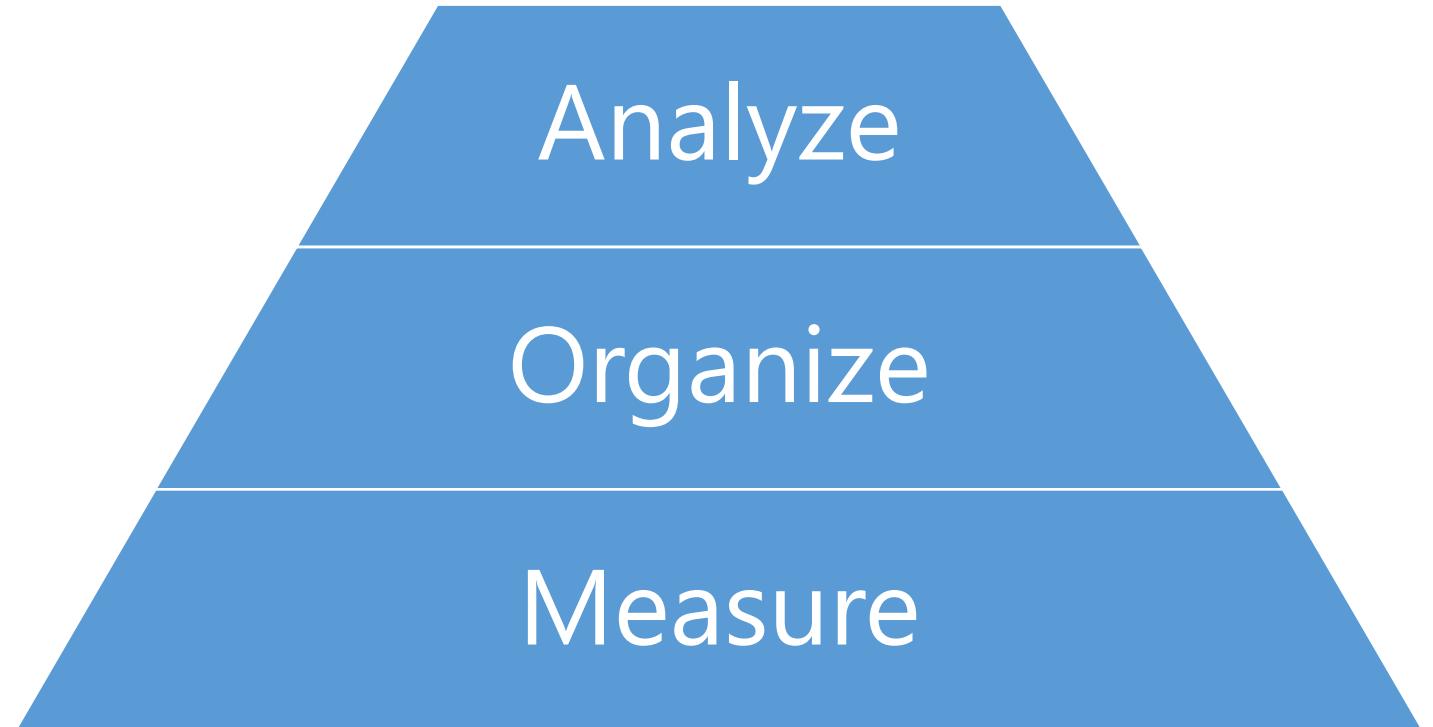
Dashboards

KPI monitors

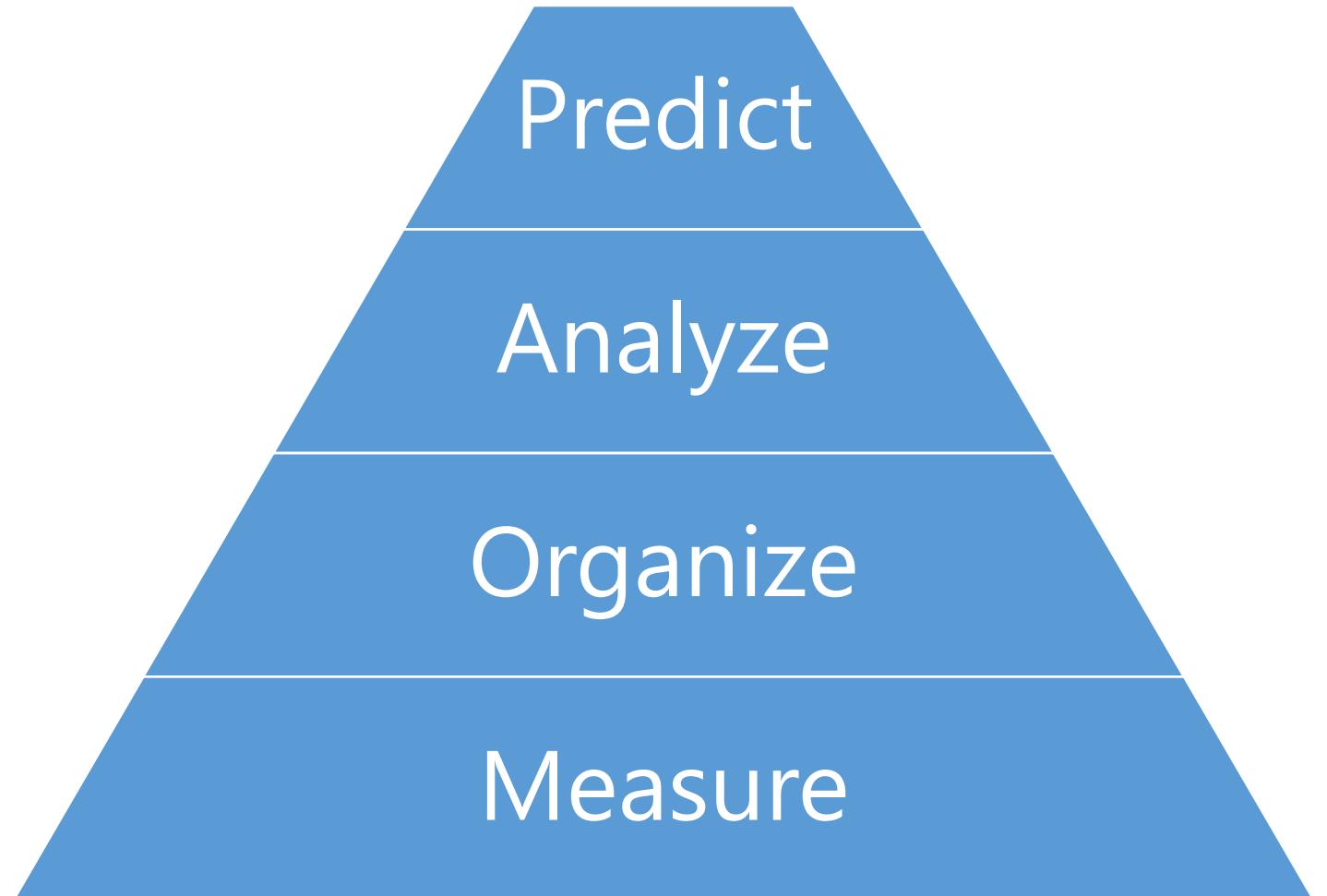
Data mining

Descriptive analytics

Diagnostic analytics

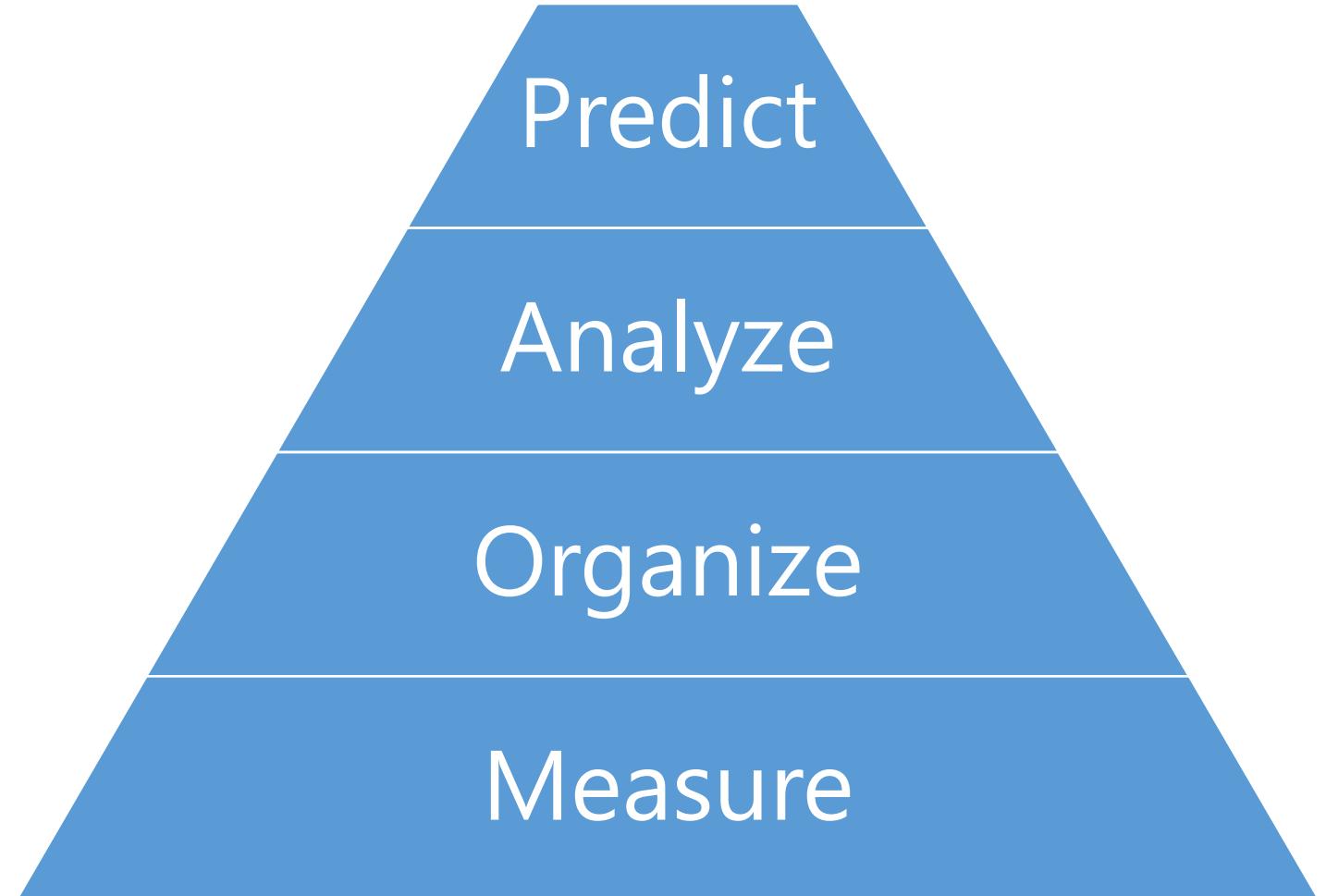


4. Predict

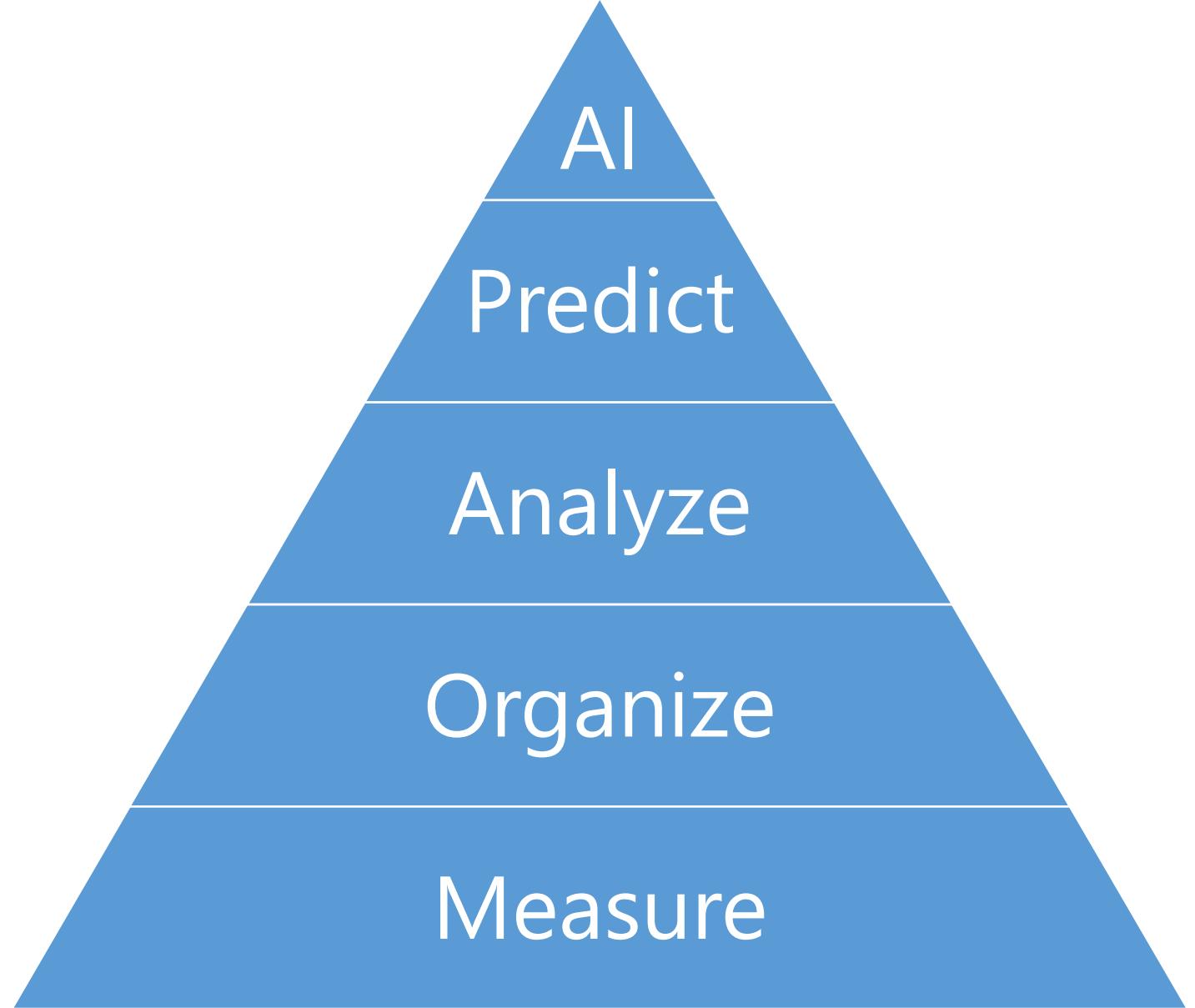


4. Predict

Predictive analytics
Prescriptive analytics
Machine learning

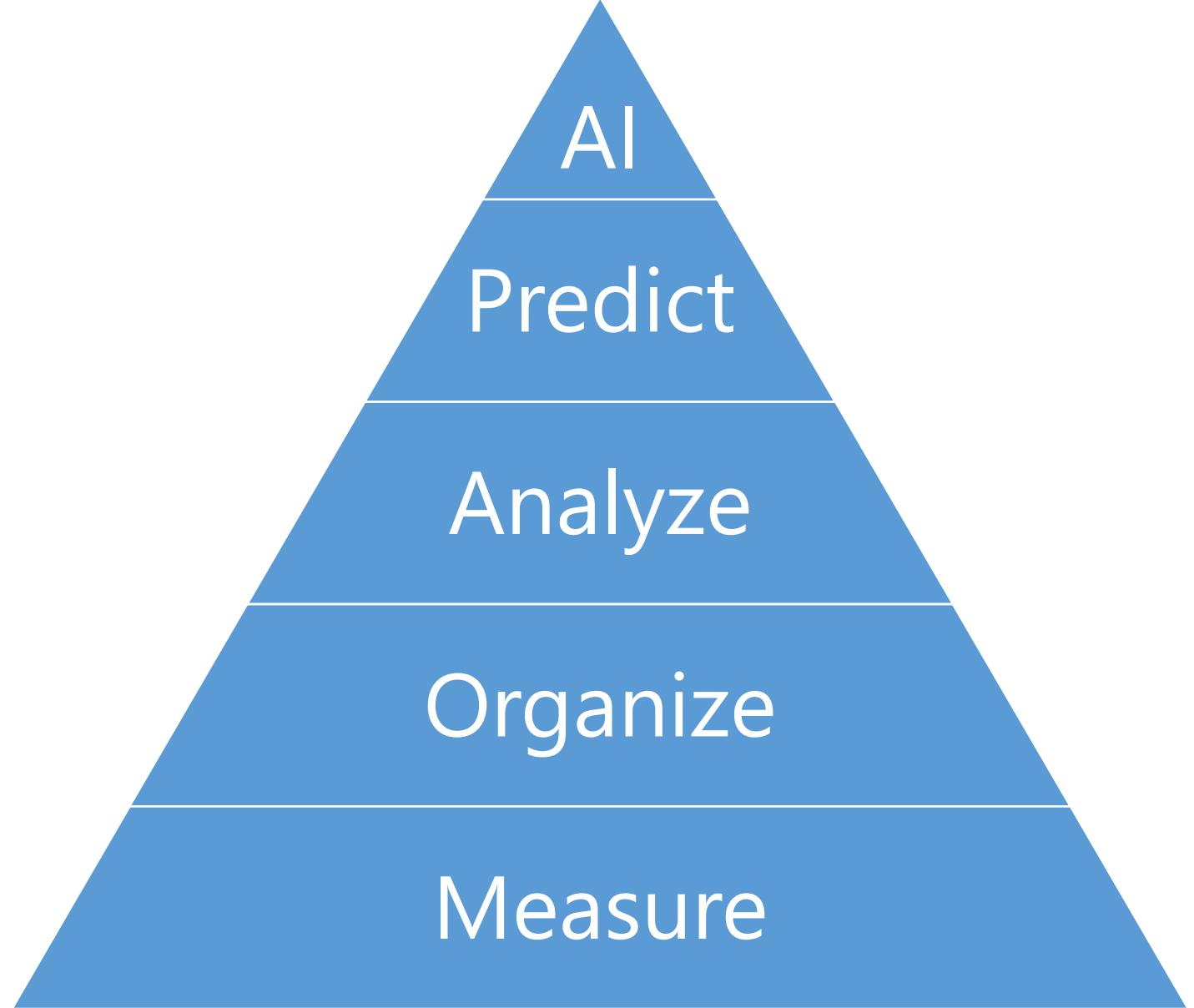


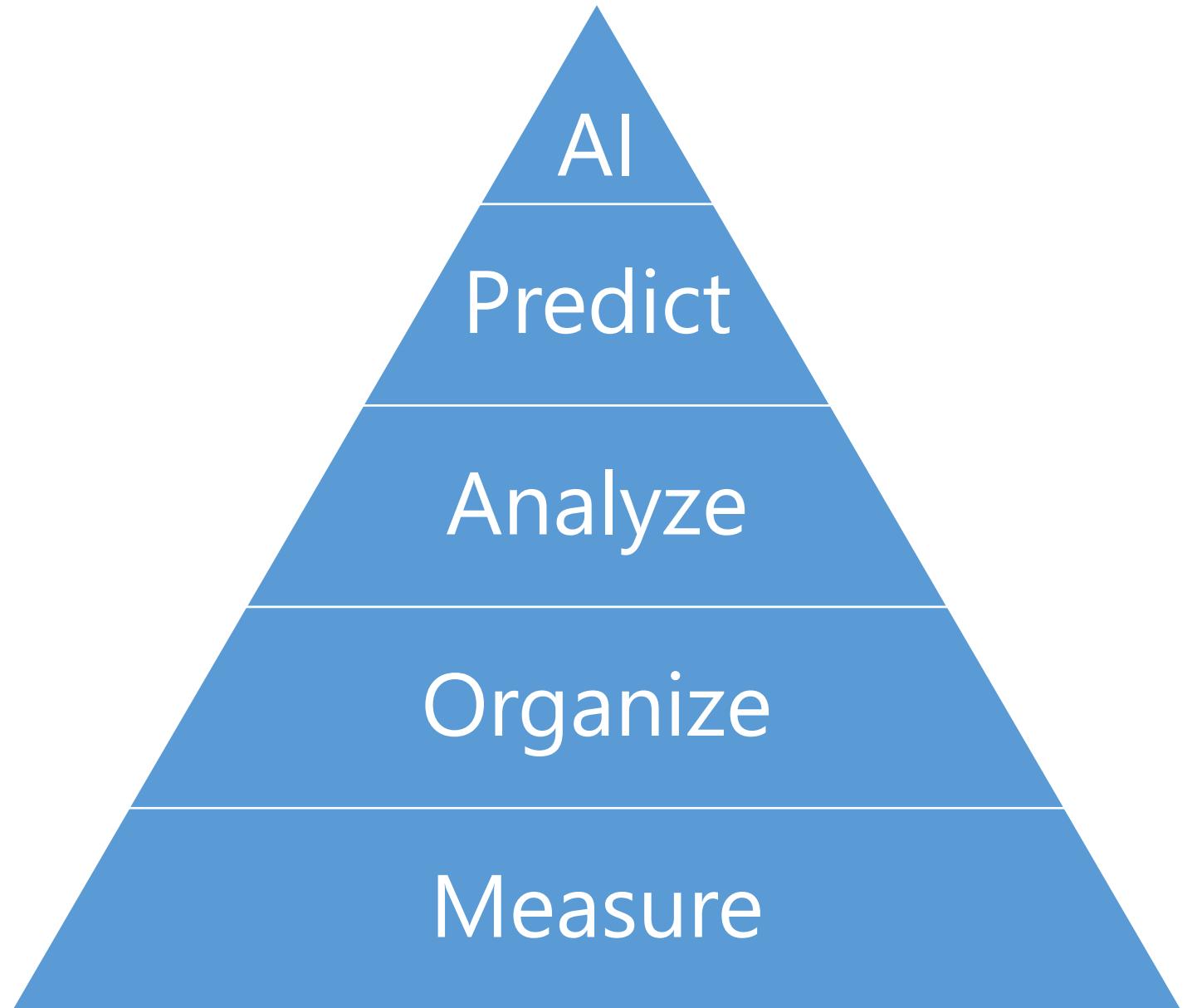
5. Automate

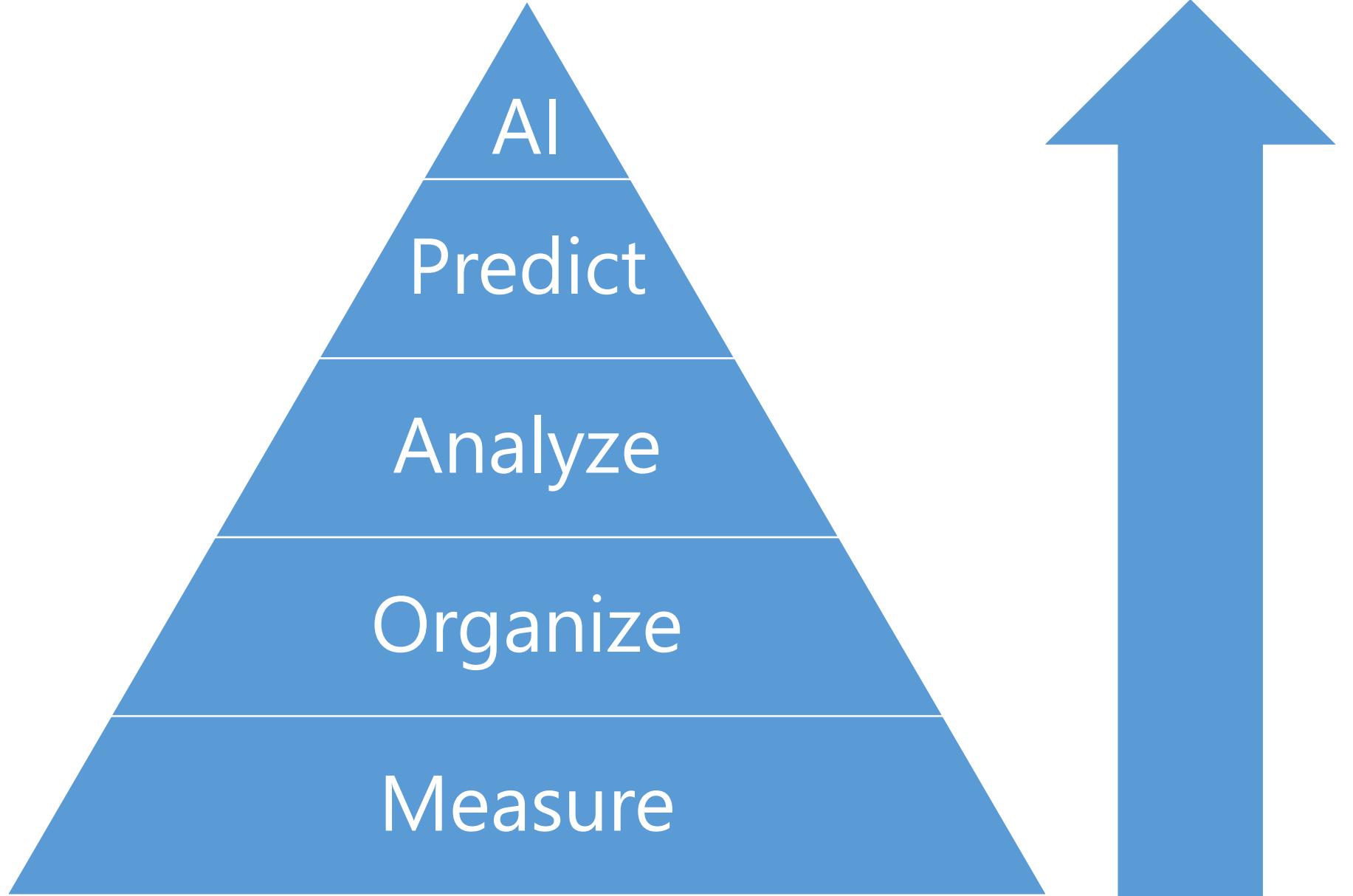


5. Automate

Artificial intelligence
Reinforcement learning
Deep learning







Advice for Success

- Get buy-in from leadership
- Focus on low-hanging fruit
- Don't silo data science teams
- Democratize your data

Advice for Success

Get buy-in from leadership
Focus on low-hanging fruit
Don't silo data science teams
Democratize your data

Embrace smart failure
Focus on feedback
Embed data collection
Avoid the Observer Effect



Where to Go Next?

0110

1001

Where to Go Next

Data Camp: <https://www.datacamp.com>

Pluralsight: <https://www.pluralsight.com>

Coursera: <https://www.coursera.org>

Pluralsight Courses

Data Science: The Big Picture

Data Science with R

Exploratory Data Analysis with R

Data Visualization with R (3-part)

Deep Learning: The Big Picture



<https://www.pluralsight.com/authors/matthew-renze>

News

2017-08-25 - Invitation to Speak at Devoxx Morocco

Very excited to announce that I've been invited to give a keynote in Casablanca at [Devoxx Morocco](#) in November. My keynote presentation will be on [Artificial Intelligence](#).



2017-08-16 - Invitation to Speak at Microsoft Ignite

I've been invited to speak at [Microsoft Ignite](#) in Orlando, Florida in September. This will be my first time speaking at Ignite. Talks will include both Data Science and Machine Learning with R.



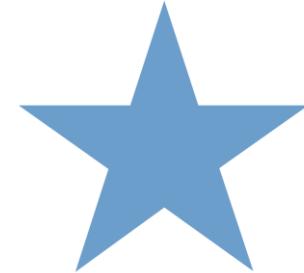
2017-08-14 - Dev on Fire Interview



Matthew is a data science consultant, author for [Pluralsight](#), international public speaker, a [Microsoft MVP](#), [ASPIInsider](#), and open-source software contributor.

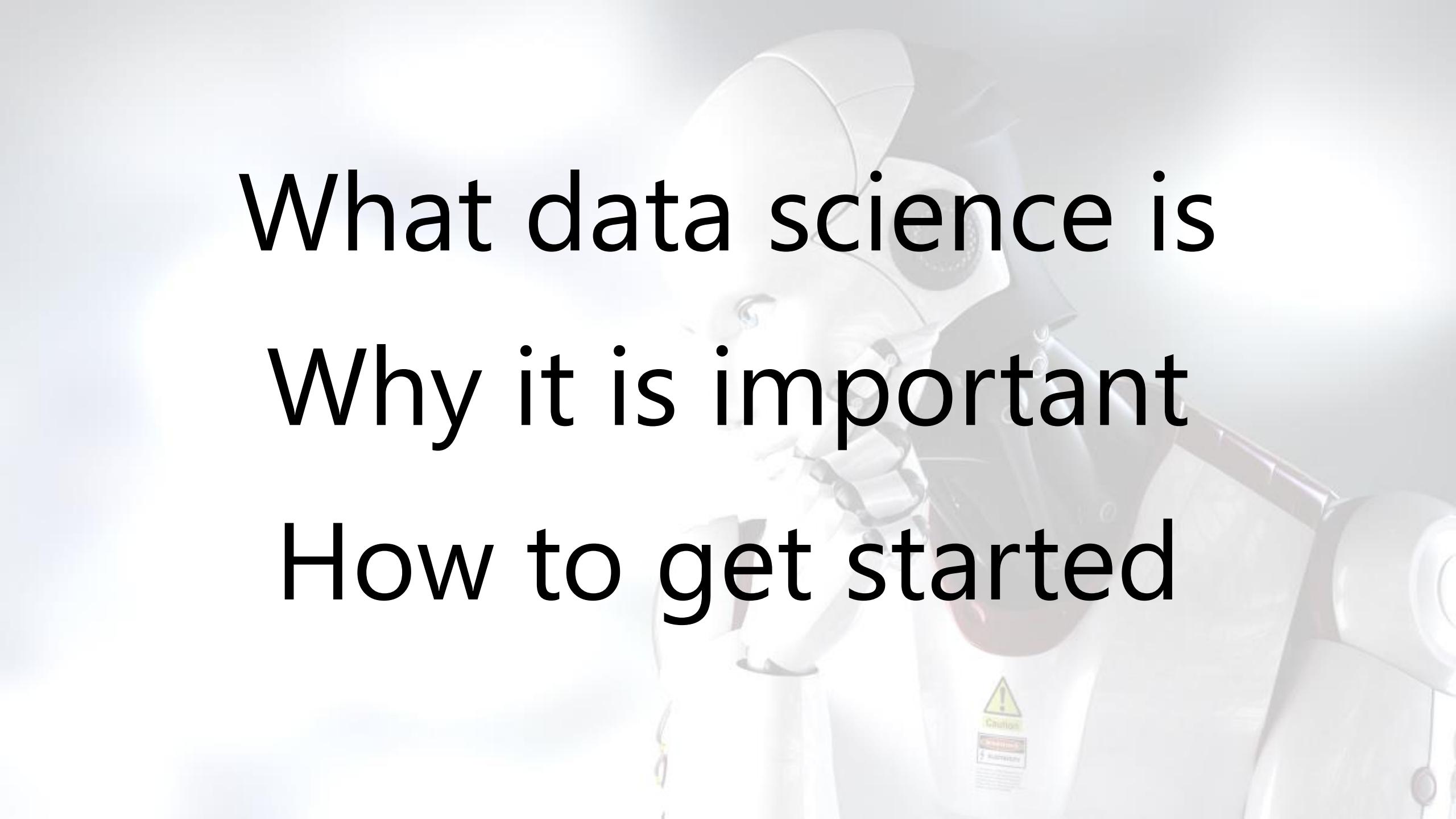
Feedback

Very important to me!
What did you like?
What could I improve?



Conclusion

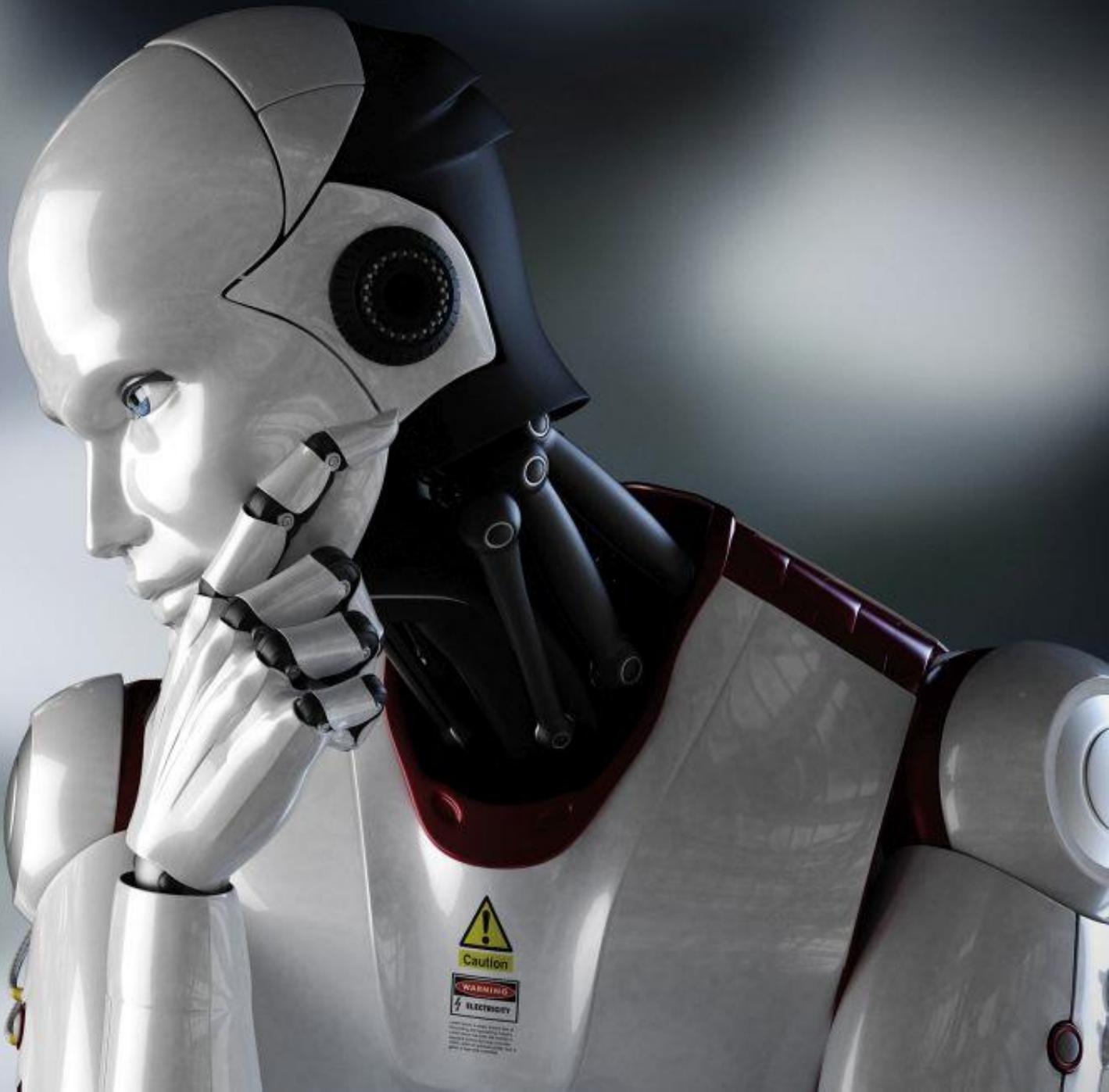


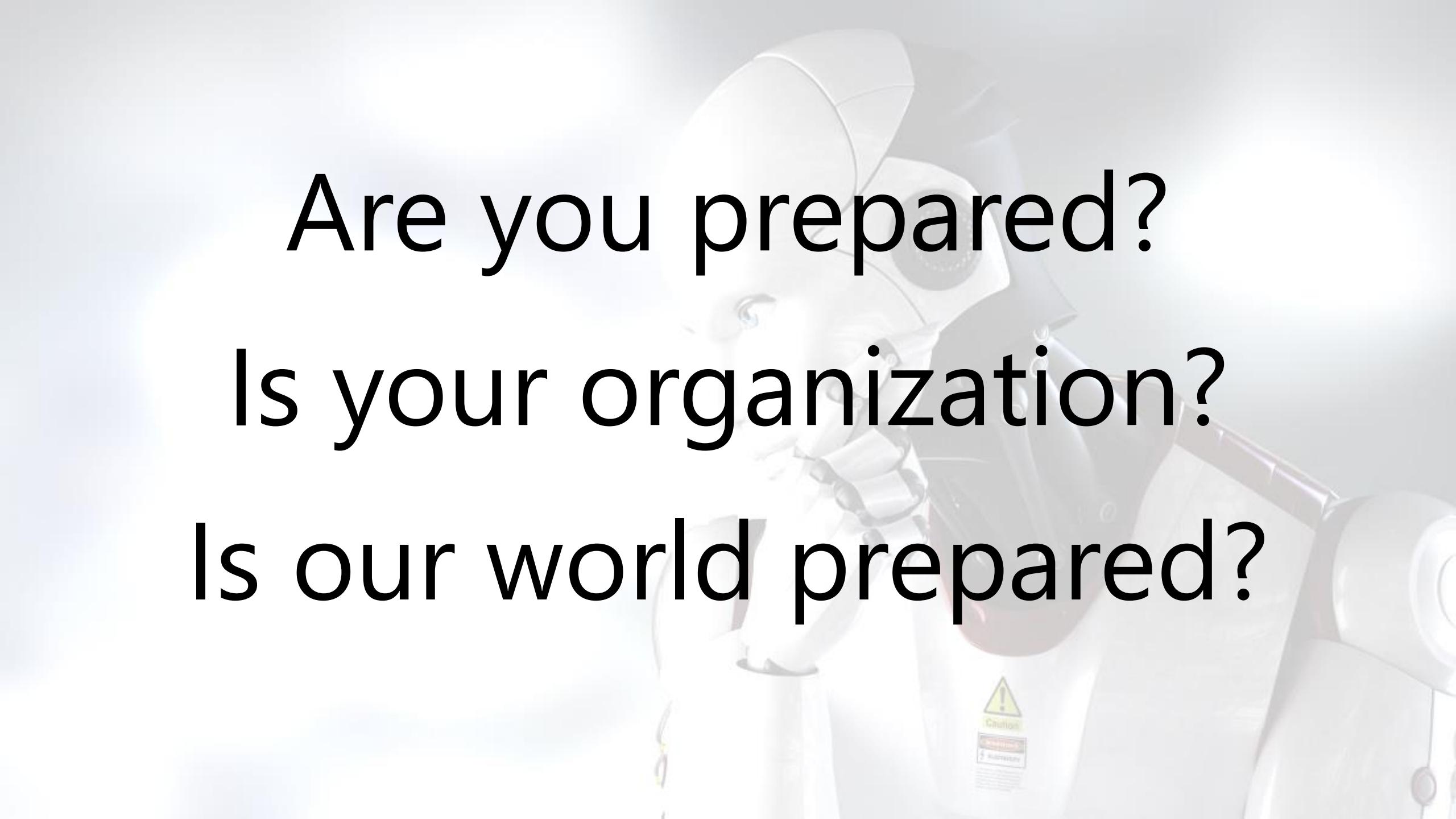
A person wearing a white protective suit and a respirator mask is working on a large industrial machine. The machine has various pipes, valves, and a prominent red circular component. The person's hands are visible, working on a tool or part of the machine. The background is slightly blurred, emphasizing the worker and the equipment.

What data science is

Why it is important

How to get started





Are you prepared?

Is your organization?

Is our world prepared?



Thank You!

Matthew Renze
Data Science Consultant
Renze Consulting

Twitter: [@matthewrenze](https://twitter.com/matthewrenze)
Email: info@matthewrenze.com
Website: www.matthewrenze.com

