



* Part of this slide is modified from a slide of Prof.Natawut



Introduction to Data Science

2190513 Data Science (ICE) (2024/1)

Assoc. Prof. Peerapon Vateekul, Ph.D.
Department of Computer Engineering,
Faculty of Engineering, Chulalongkorn University
Peerapon.v@chula.ac.th
www.cp.eng.chula.ac.th/~peerapon/



Outline

- Introduction
 - Data is important
 - Data Science Definition by Dr.Virote
 - Data Science Definition by Aj.Natawut
- Big Data
- Data Science Process & Data Science Trend





Introduction



Data is important (in 2017)

The Economist Topics ▾ Current edition More ▾

Regulating the internet giants

The world's most valuable resource is no longer oil, but data

The data economy demands a new approach to antitrust rules



David Parkins

Print edition | Leaders >
May 6th 2017

Twitter Facebook LinkedIn Email Print

- Alphabet (Google's parent company), Amazon, Apple, Facebook and Microsoft
- \$25bn in net profit in the first quarter of 2017
- Amazon captures half of all dollars spent online in America.
- Google and Facebook accounted for almost all the revenue growth in digital advertising in America last year



Data is important (in 2018)! (cont.)

The New Oil

Jennifer Presley Executive Editor, E&P Magazine Hart Energy Thursday, November 1, 2018 - 6:40am



With a number of successful projects under its collective belt, the oil and gas industry is proving Big Data is more than just a buzzword. (Source: Makhnach_S/Shutterstock.com; Design by Felicia Hammons)

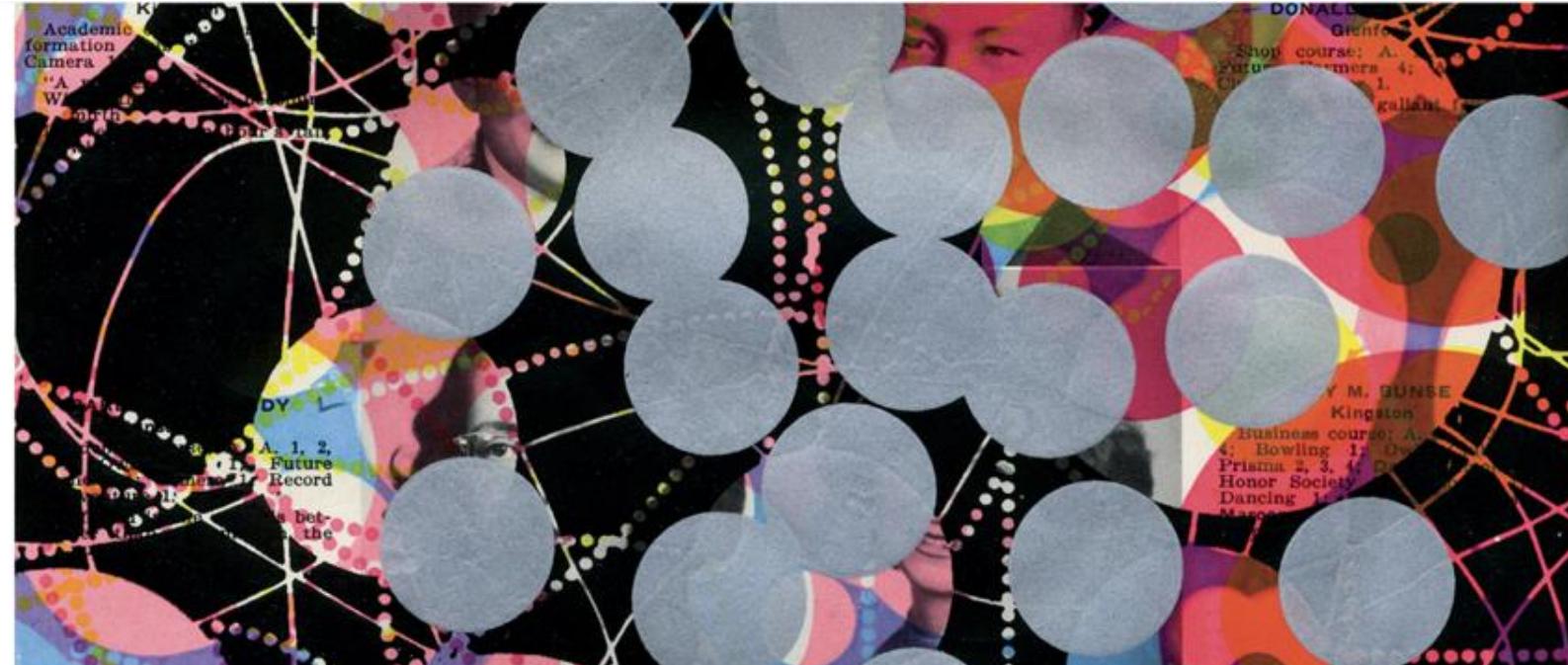
<https://www.epmag.com/new-oil-1720651>

Data Science
(AI,ML,DM)
+
Big Data



Who analyzes these data!

Harvard
Business
Review



ARTWORK: TAMAR COHEN, ANDREW J. BUBOLTZ, 2011, SILK SCREEN
ON A PAGE FROM A HIGH SCHOOL YEARBOOK, 8.5" X 12"

DATA

Data Scientist: The Sexiest Job of the 21st Century

by Thomas H. Davenport and D.J. Patil

FROM THE OCTOBER 2012 ISSUE

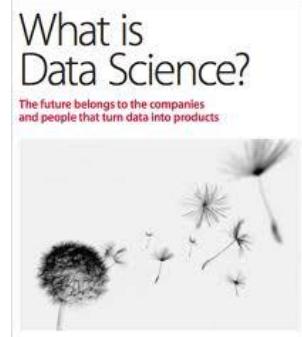
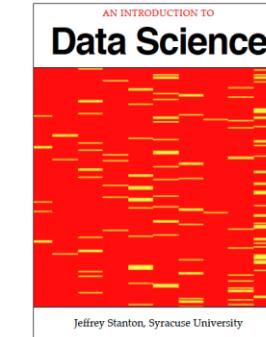
WHAT TO READ NEXT



Competing on Analytics



What is Data Science?



■ Data

- Facts and statistics collected for reference or analysis



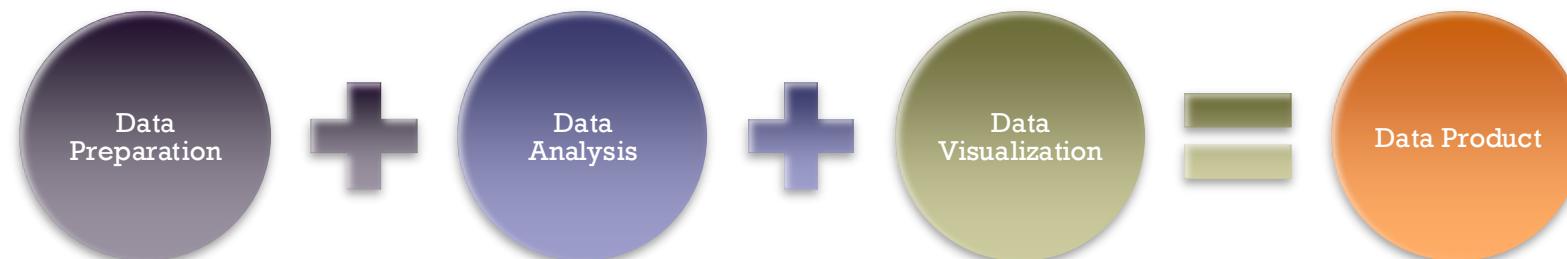
■ Science

- A systematic study through observation and experiment



■ Data Science

- The scientific exploration of data to extract meaning or insight,
- and the construction of software to utilize such insight in a business context.





What is Data Science? (cont.)

1. Transform data into **valuable insights**
2. Transform data into **data products**
3. Transform data into **interesting stories**



Ta Virot Chiraphadhanakul
Data Scientist, Facebook

Code Mania 2 (01), Jan-2015



1) Transform data into valuable insights

How Obama Raised \$60 Million by Running a Simple Experiment



By Dan Siroker
November 29, 2010

A/B Testing Case Studies, Multivariate Testing, Political Campaign

The image displays two versions of the Obama '08 campaign website side-by-side, illustrating a design experiment. Both versions feature the same overall layout: a blue header with the 'OBAMA '08' logo, a main section with a photo of Barack Obama, and a footer with a 'JOIN THE MOVEMENT' call-to-action.

The primary difference between the two versions is the position of the 'SIGN UP' button:

- Left Version:** The 'SIGN UP' button is located at the bottom right of the main content area, below the photo of the Obamas. An arrow points from the word 'Button' to this red button.
- Right Version:** The 'SIGN UP' button has been moved to the top right of the main content area, above the photo of the Obamas. An arrow points from the word 'Button' to this red button.

Both versions include fields for 'Email Address' and 'Zip Code' and a 'CONTINUE TO WEBSITE' link at the bottom.



1) Transform data into valuable insights (cont.)



BUSINESS

Amazon introduces next major job killer to face Americans

By James Covert, Linda Massarella and Bruce Golding

December 5, 2016 | 9:59pm | Updated



The Amazon Go storefront
Amazon

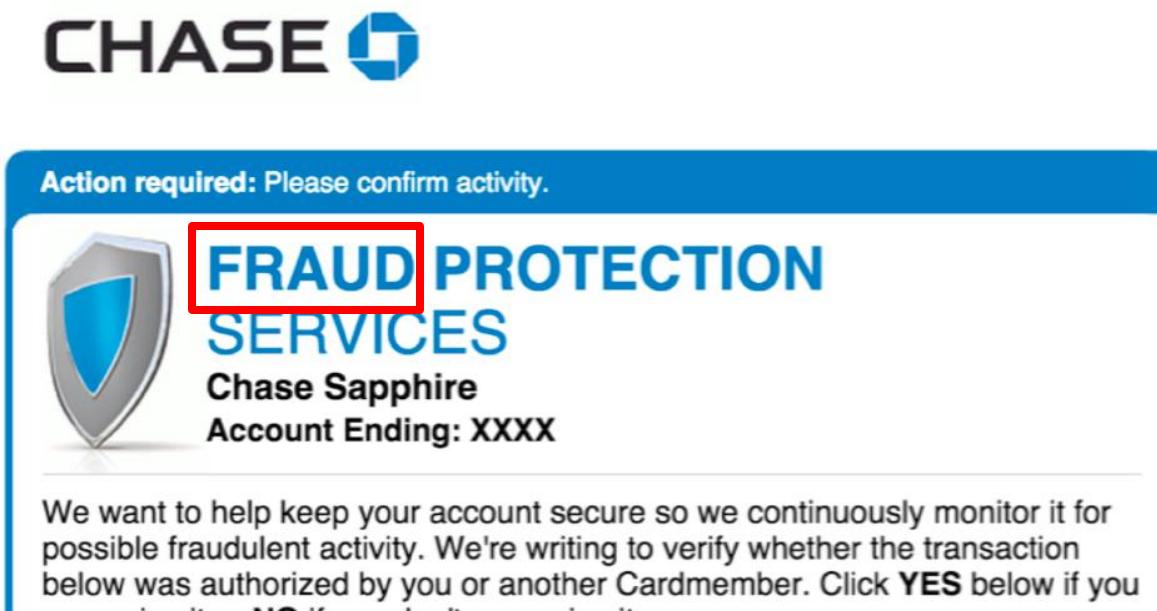
Ta Virot Chiraphadhanakul
Data Scientist, Facebook
Code Mania 2, Jan-2015



Amazon's new supermarket will blow your mind — and cost America jobs
Why employment may be optional in the near future
Amazon's futuristic grocery store could spell employment doom
Amazon introduces next major job killer to face Americans
Amazon's latest idea could kill off jobs forever



2) Transform data into data products



The screenshot shows the Microsoft Outlook interface. The top navigation bar includes "Outlook", a search bar, and various action buttons like "New message", "Empty folder", "Mark all as read", and "Undo". The left sidebar lists "Favorites" and "Folders": "Inbox" (45 items), "Junk Email" (128 items, highlighted with a red box), "Drafts", "Sent Items", "Deleted Items", "Archive", "Conversation Hist...", and "Notes". The main pane displays the "Junk Email" folder content for "Today". It shows several spam emails: "Work At Home Opportunities" (1:47 PM), "Client service NETFLIX SURVEY" (1:40 PM), "Thank You Costco Re: Costco Has a Surprise Fo..." (12:01 PM), and "Client service - Are you a friend of Amazo..." (8:43 AM).

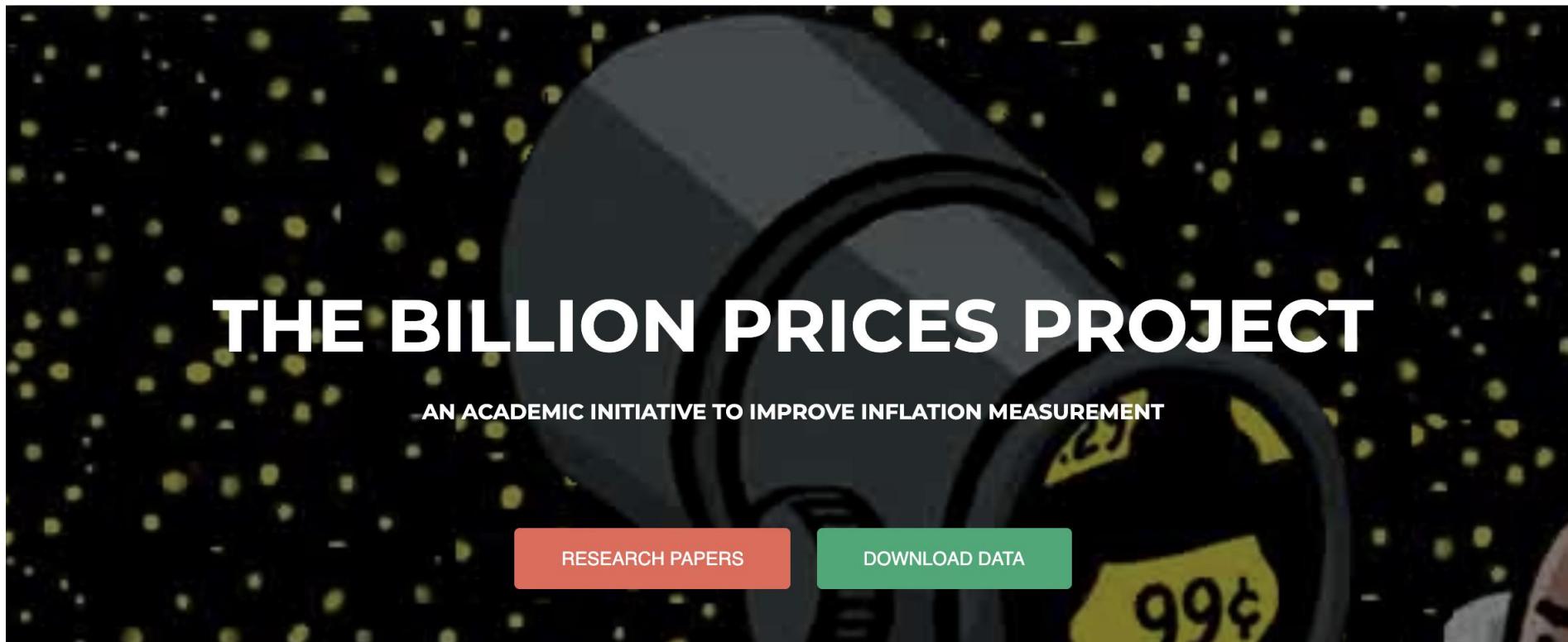


3) Transform data into interesting stories

Consumer Price Index (CPI) - Inflation

The Billion Prices Project

Home Our Public Data Our Research News



<http://www.thebillionpricesproject.com/>

A promotional poster for Tesco Lotus Express featuring large red text 'TESCO' and 'Lotus express' on the left, and large red text 'ลด กู้ เดือน' (Promotion Month) on the right. The poster is filled with various product offers, each with a yellow starburst containing a price and a promotional message like 'ลด 30%' or 'ลด 50%'. Products shown include oils, soft drinks, tea, coffee, snacks, personal care items, and household goods.

13

กิ๊บส์ ได้รับการยอมรับเป็น Trusted Brand
ซึ่งเป็นสุดยอดแบรนด์ที่พูนรักโภคให้ความเชื่อมั่นมากที่สุด
ต่อเนื่องกันถึง 5 ปี สาขาดูแลรักษารักษา จัดการผลิตภัณฑ์ได้ดีที่สุด

30 ล. - 19 ก.พ. 2560

Product	Price
Red Hot	159 บ.
แมกโนเลีย ชีสเค้ก 10 ชิ้น	12 บ.
แมกโนเลีย ชีสเค้ก 25 ชิ้น	22 บ.
Tops น้ำผึ้ง 100% 1 ล.	59 บ.
แมกโนเลีย ชีสเค้ก 10 ชิ้น	119 บ.
แมกโนเลีย ชีสเค้ก 25 ชิ้น	129 บ.
น้ำยาล้างจาน 400 มล.	135 บ.
น้ำยาล้างจาน 400 มล.	135 บ.
น้ำยาล้างจาน 400 มล.	49 บ.
น้ำยาล้างจาน 400 มล.	219 บ.
น้ำยาล้างจาน 400 มล.	129 บ.
น้ำยาล้างจาน 400 มล.	21 บ.
น้ำยาล้างจาน 200 มล.	95 บ.
น้ำยาล้างจาน 200 มล.	99 บ.
น้ำยาล้างจาน 300 มล.	42 บ.
น้ำยาล้างจาน 300 มล.	59 บ.
น้ำยาล้างจาน 300 มล.	63 บ.
น้ำยาล้างจาน 300 มล.	85 บ.
น้ำยาล้างจาน 300 มล.	30 บ.
น้ำยาล้างจาน 300 มล.	55 บ.
น้ำยาล้างจาน 100 มล.	123 บ.
น้ำยาล้างจาน 100 มล.	129 บ.
น้ำยาล้างจาน 100 มล.	65 บ.
น้ำยาล้างจาน 100 มล.	30 บ.
น้ำยาล้างจาน 100 มล.	79 บ.

BO BUY 1 GET 1 FREE

1 แถม 1



The Billion Prices Project: Using Online Prices for Measurement and Research *

Alberto Cavallo

MIT and NBER

Roberto Rigobon

MIT and NBER

This Version: April 8, 2016

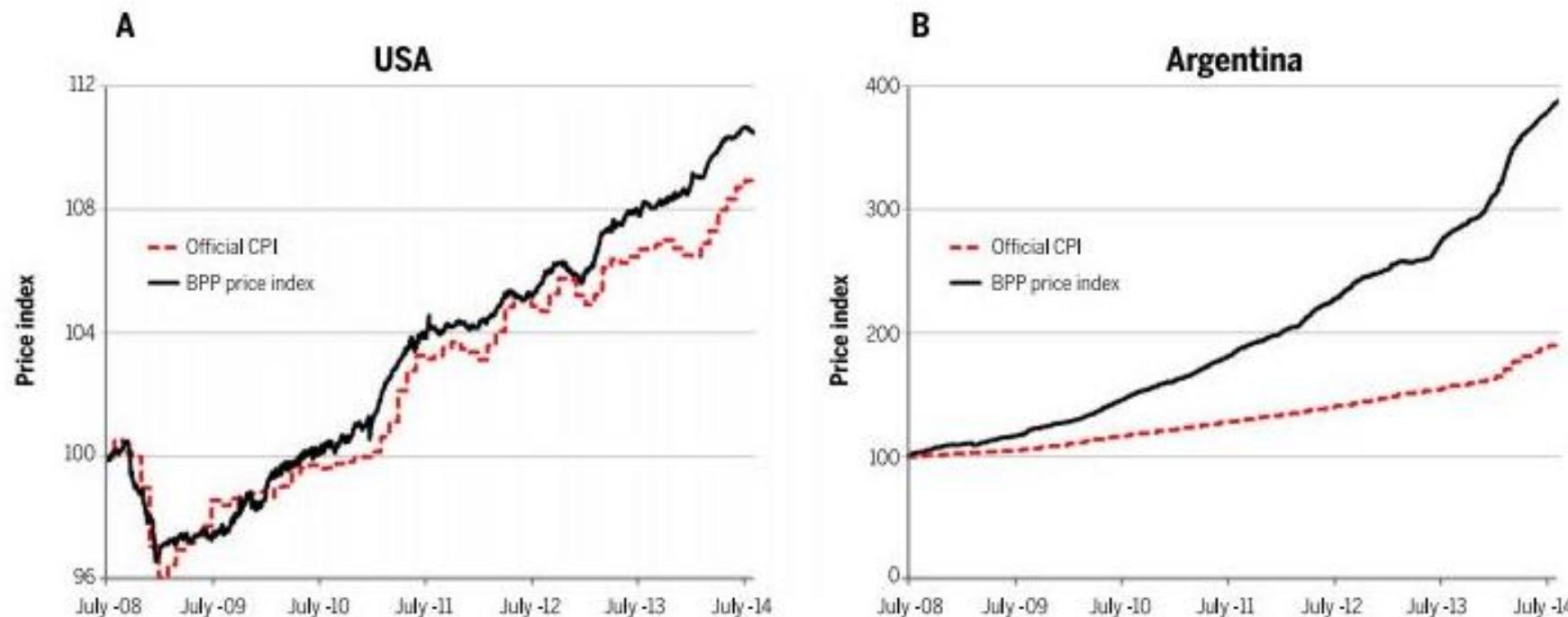
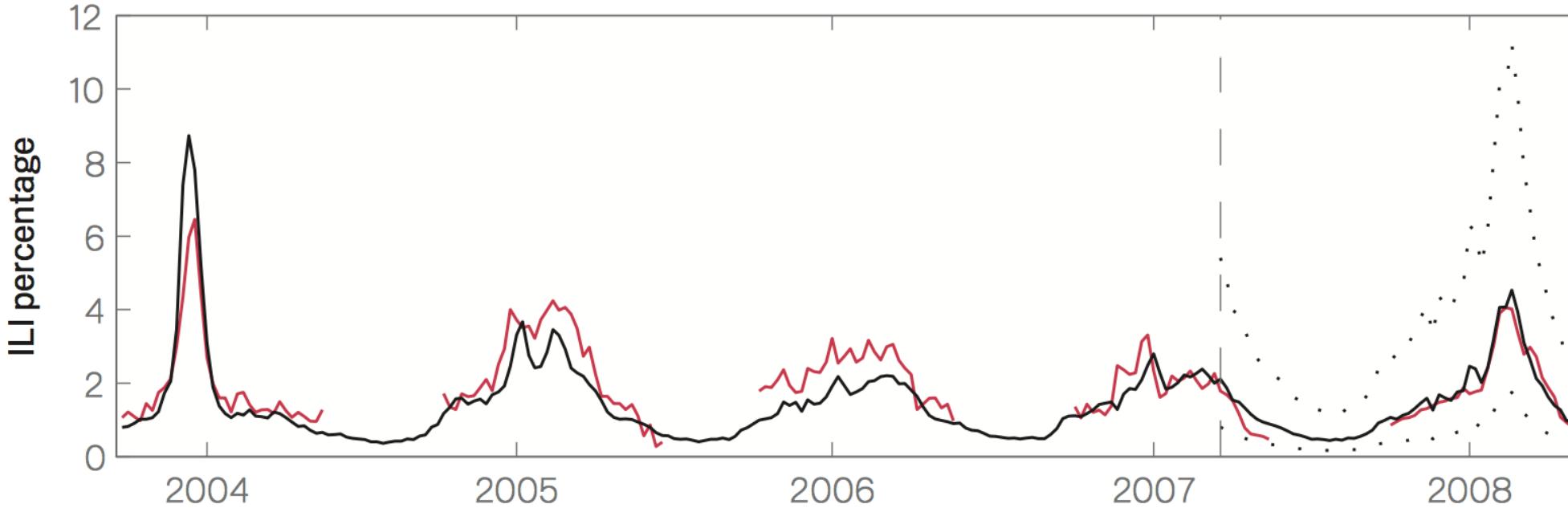


Fig. 2. BPP price index. Dashed red lines show the monthly series for the CPI in the United States (A) and Argentina (B), as published by the formal government statistics agencies. Solid black lines show the daily price index series, the "State Street's PriceStats Series" produced by the BPP, which uses scraped Internet data on thousands of retail items. All indices are normalized to 100 as of July 2008. In the US chart, the two series track

each other quite closely, although the BPP index is available in real time and at a more granular level (daily instead of monthly). In the plot for Argentina, the indices diverge considerably, with the BPP index growing at about twice the rate of the official CPI. [Updated version of figure 5 in (18), provided courtesy of Alberto Cavallo and Roberto Rigobon, principal investigators of the BPP]



Google Flu Trend



Ginsberg, Jeremy; Mohebbi, Matthew H.; Patel, Rajan S.; Brammer, Lynnette;
Smolinski, Mark S.; Brilliant, Larry (19 February 2009). "Detecting influenza
epidemics using search engine query data". *Nature*. **457** (7232): 1012–1014.



What are they using data science for?

1. Measurement
2. Insights
3. Data Products





1) Measurement

- To make a decision based on data
- Aka. benchmarking
- Turning qualitative information into quantitative values
 - Usually called metrics or indicators
- Direct and indirect measurement

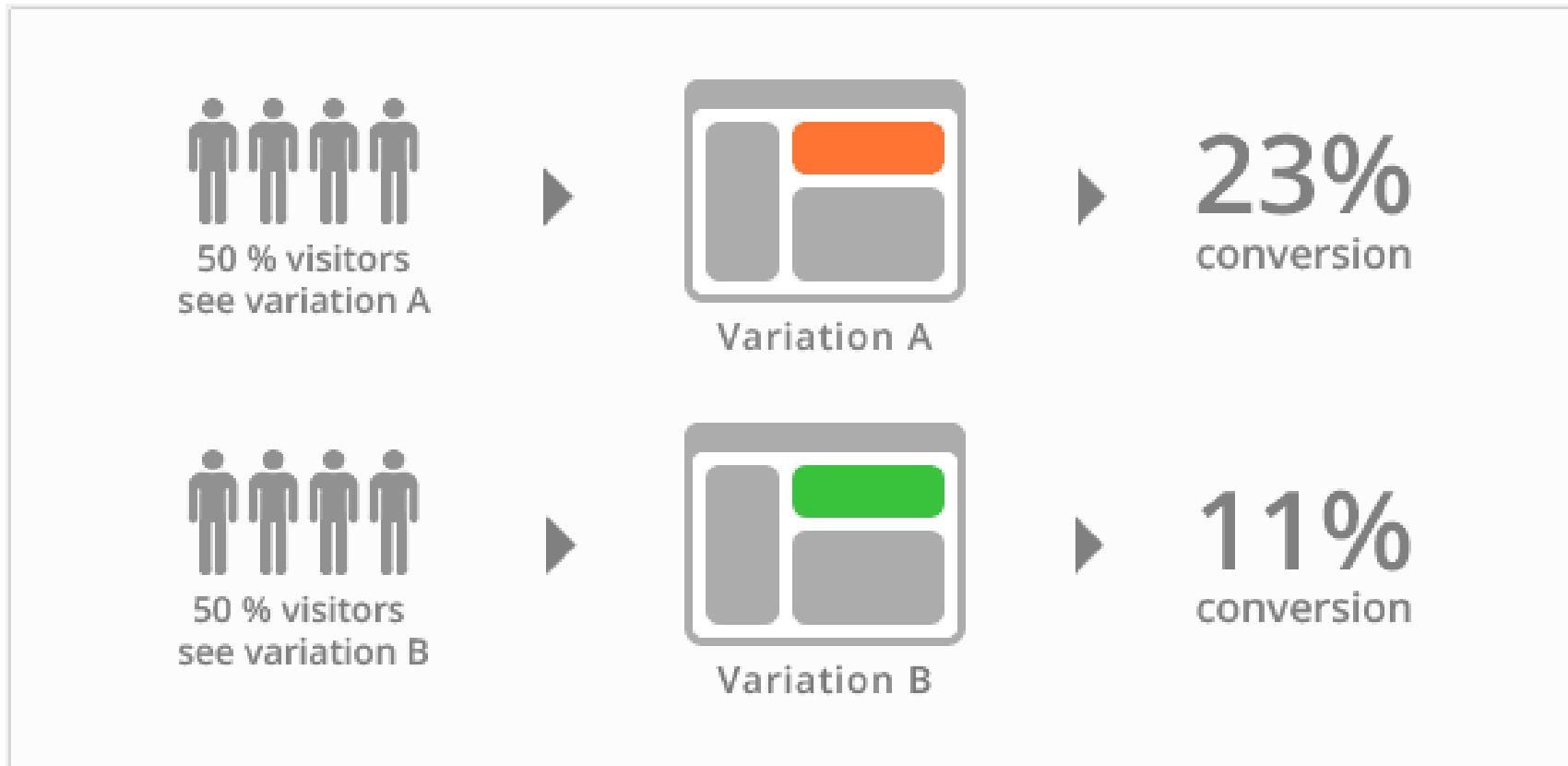


Why do we need to measure?

- Comparison between alternatives (**make a selection**)
 - Choosing which notebook to buy
- Comparison after **improvement** or tuning
 - Should I add memory to my notebook?
- **A/B Testing** (split testing)
 - Let the actual users decide their preferences
 - Very popular for UI design



A/B Testing



Source: <https://vwo.com/ab-testing/>



Example: SimCity

1. Remove product banner: SimCity sees 43% more conversions without hero banner image

20

Control

The screenshot shows the official SimCity website homepage. A large hero banner at the top left encourages users to "PRE-ORDER AND GET \$20 OFF YOUR NEXT PURCHASE". Below this, there are two main product offerings: "SIMCITY™" priced at \$59.99 and "SIMCITY™ DIGITAL DELUXE EDITION" priced at \$79.99. Both are offered as PC Download or PC Physical options. A call-to-action button labeled "BUY NOW" is present for each. To the right, a box titled "DIGITAL DELUXE EDITION INCLUDES" lists four items: HEROES AND VILLAINS SET, FRENCH CITY SET, GERMAN CITY SET, and BRITISH CITY SET. At the bottom, a "Key Features" section is visible.

Variation

The screenshot shows the same SimCity website as the control, but the hero banner for pre-ordering has been removed. Instead, there are two separate product cards for "SIMCITY™" and "SIMCITY™ DIGITAL DELUXE EDITION". Each card includes its price (\$59.99 or \$79.99), download/physical options, and a "BUY NOW" button. To the right of the "SIMCITY™" card, a blue circular callout contains the text "43% increase in checkouts". Below the products, a "Key Features" section is displayed, which includes a "WHAT IS SIMCITY?" section and a "DEPTH OF SIMULATION" section.



Example: SmartWool

3. Use a well-defined grid layout for your online shopping experience: Uniform product page images increase ARPV 17% for SmartWool

Control

Variation

A blue circle contains the text: 17% increase in average revenue per visitor.

The Control version shows a grid layout with 12 products per row. The Variation version shows a grid layout with 16 products per row, which is highlighted by a red border. A blue circle on the Variation side contains the text "17% increase in average revenue per visitor".



2) Insights

<https://blogs.scientificamerican.com/guest-blog/9-bizarre-and-surprising-insights-from-data-science/>

- **Good understanding of user behavior** can lead to new product development or improvements of the existing products

- Walmart -- Pop-Tarts before a hurricane
 - Prehurricane, Strawberry Pop- Tart sales increased about sevenfold

- Financial startup -- Typing with proper capitalization indicates creditworthiness
 - Online loan applicants who complete the application form with the correct case are more dependable debtors

- Starbucks use customer purchase information from My Starbucks Mobile Apps to figure out new products

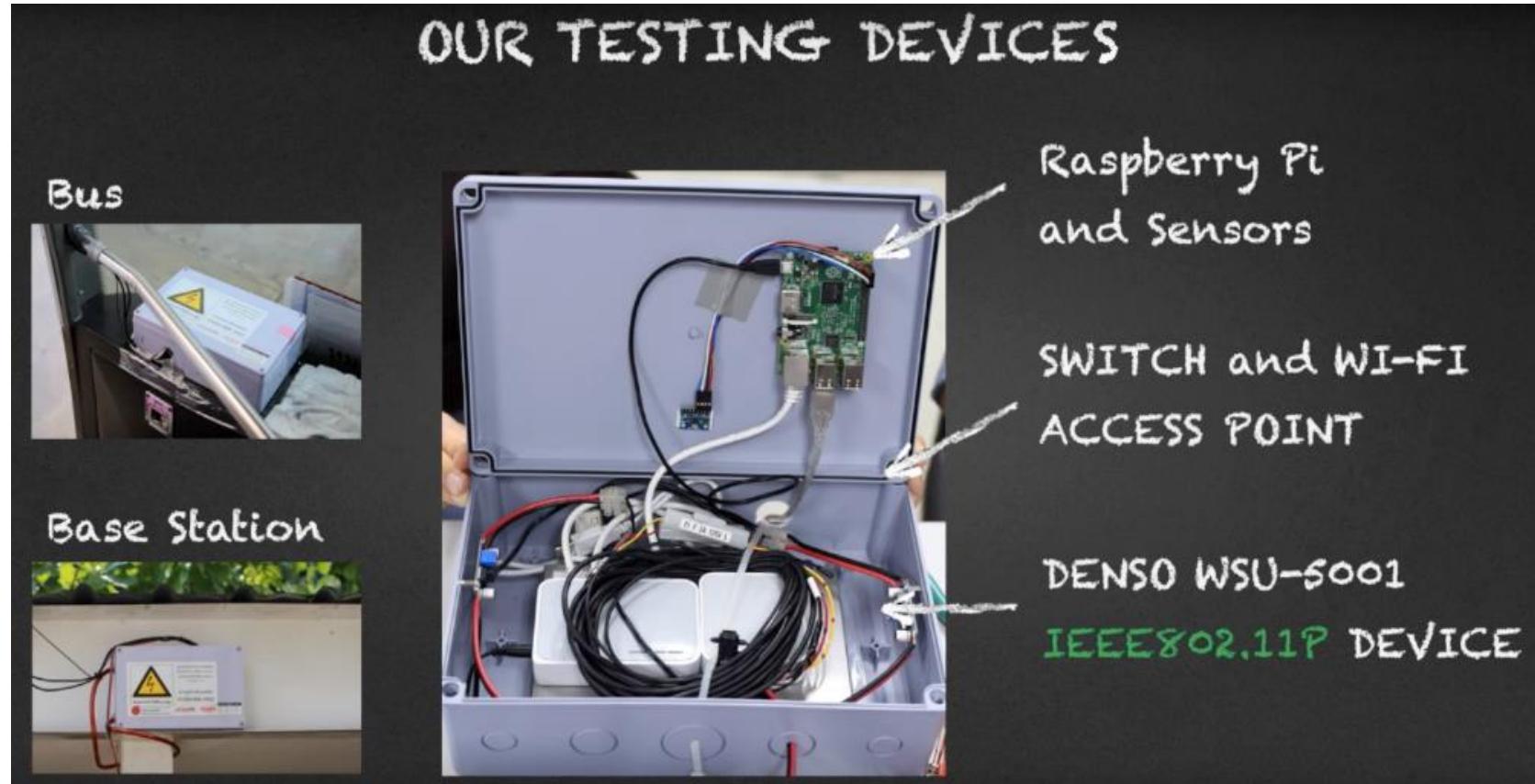


Example: Tracing Traffic





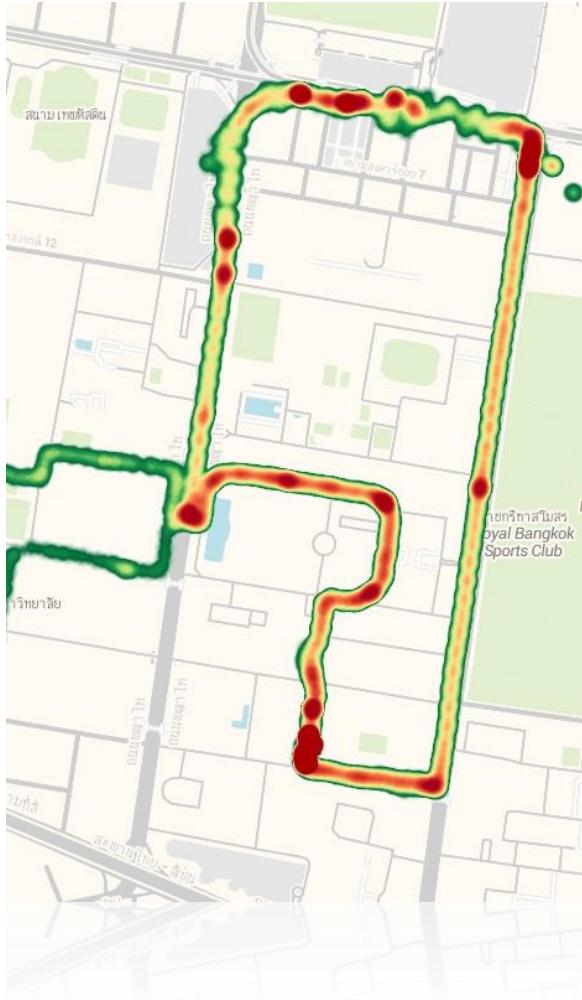
Example: Tracing Traffic



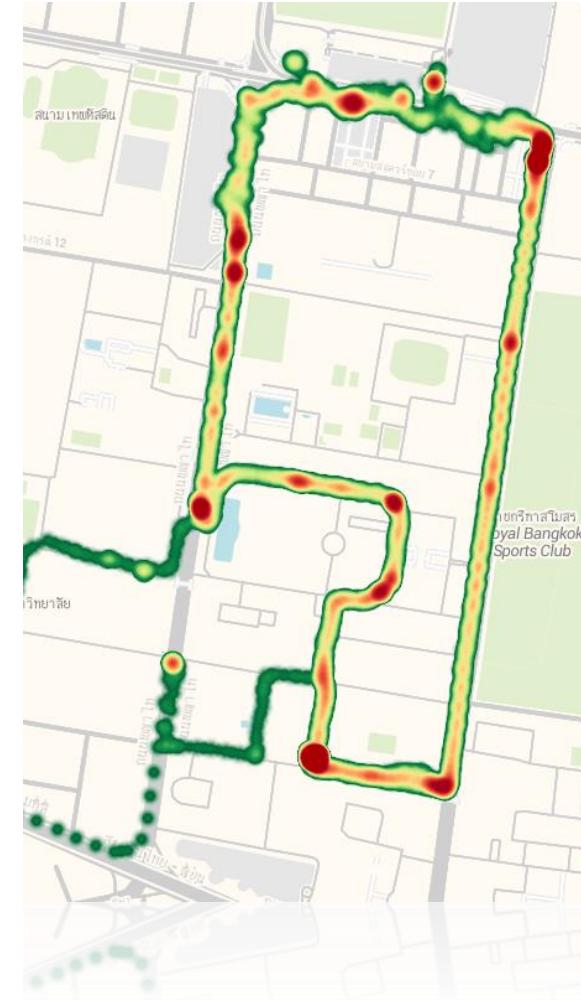


GPS Average Speed

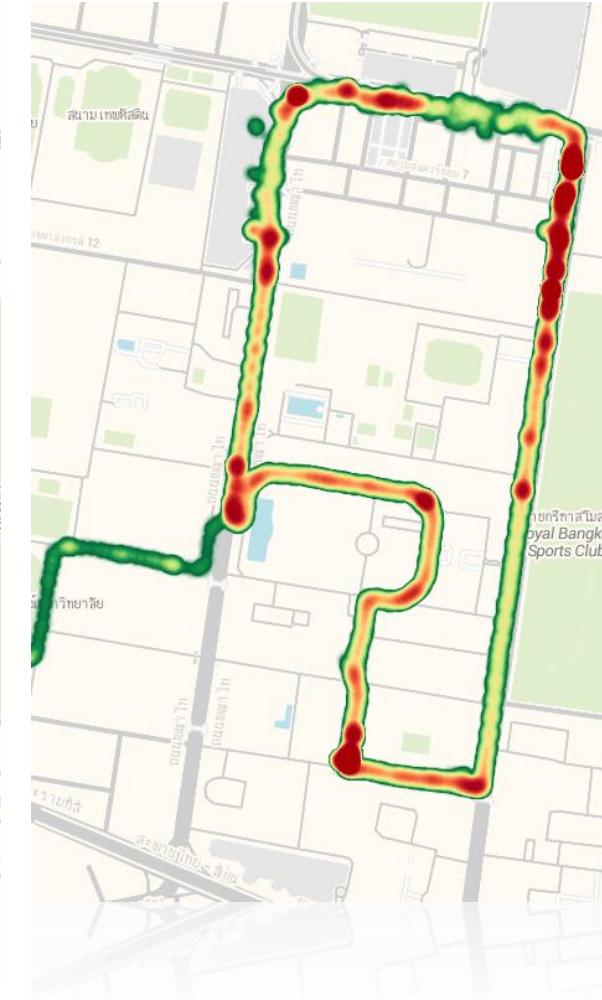
6:00-10:00



10:00-15:00



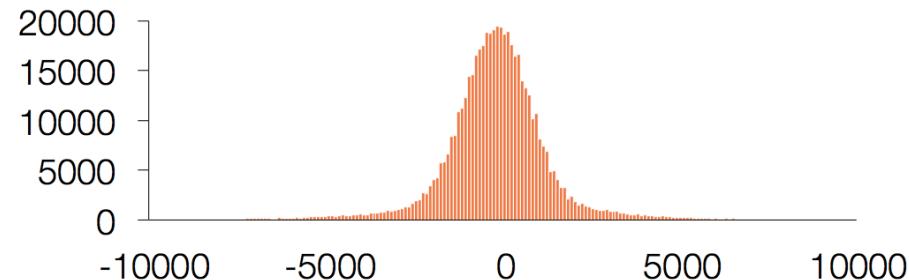
15:00-18:00



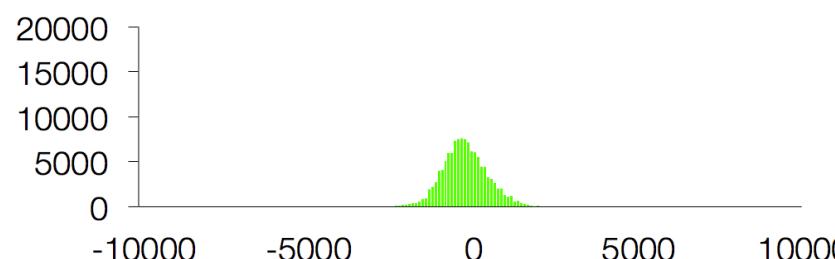
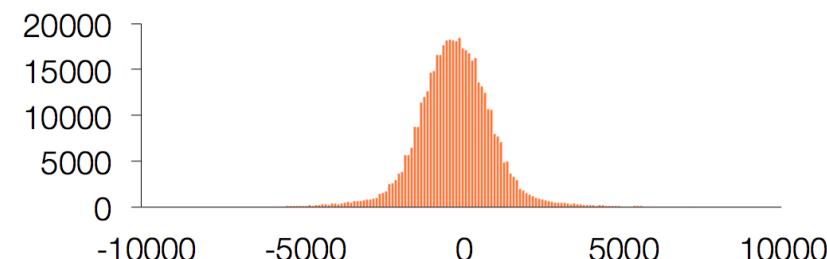
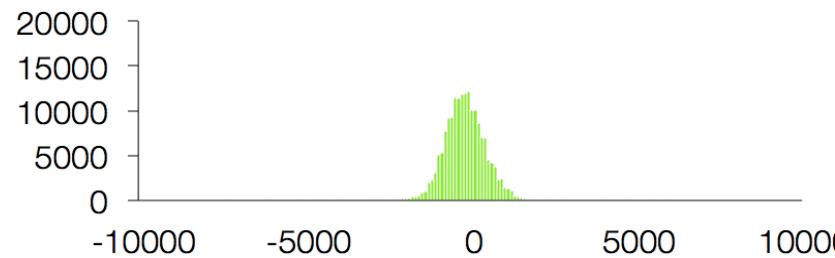
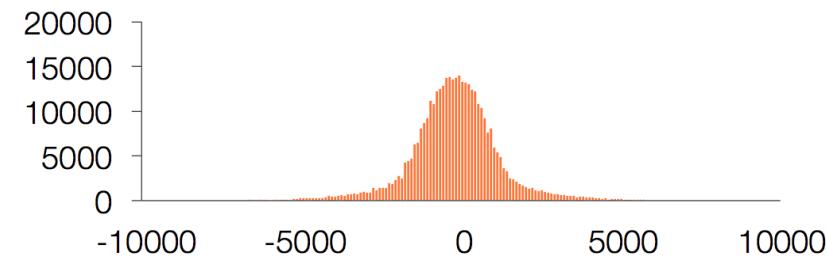
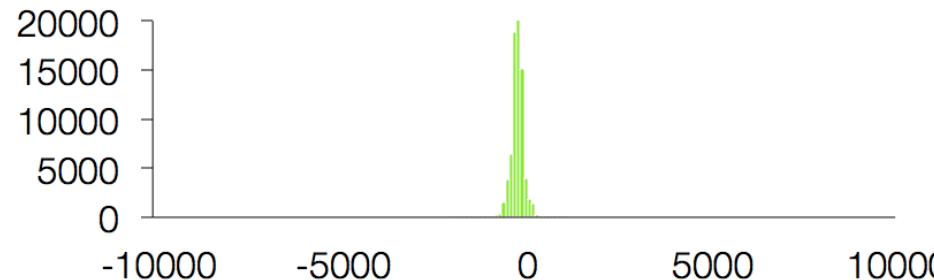


Bus Drivers' Behaviors

Bus A



Bus B





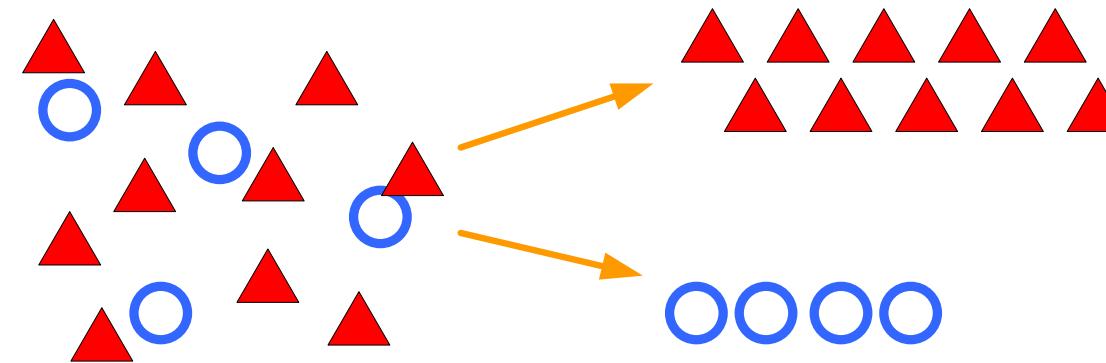
3) Data Products

- An application or system that uses data to provide “**intelligent**” products or services, which create more data that can be further used
- **Machine learning** plays an important role in building great data products



Machine Learning Classification

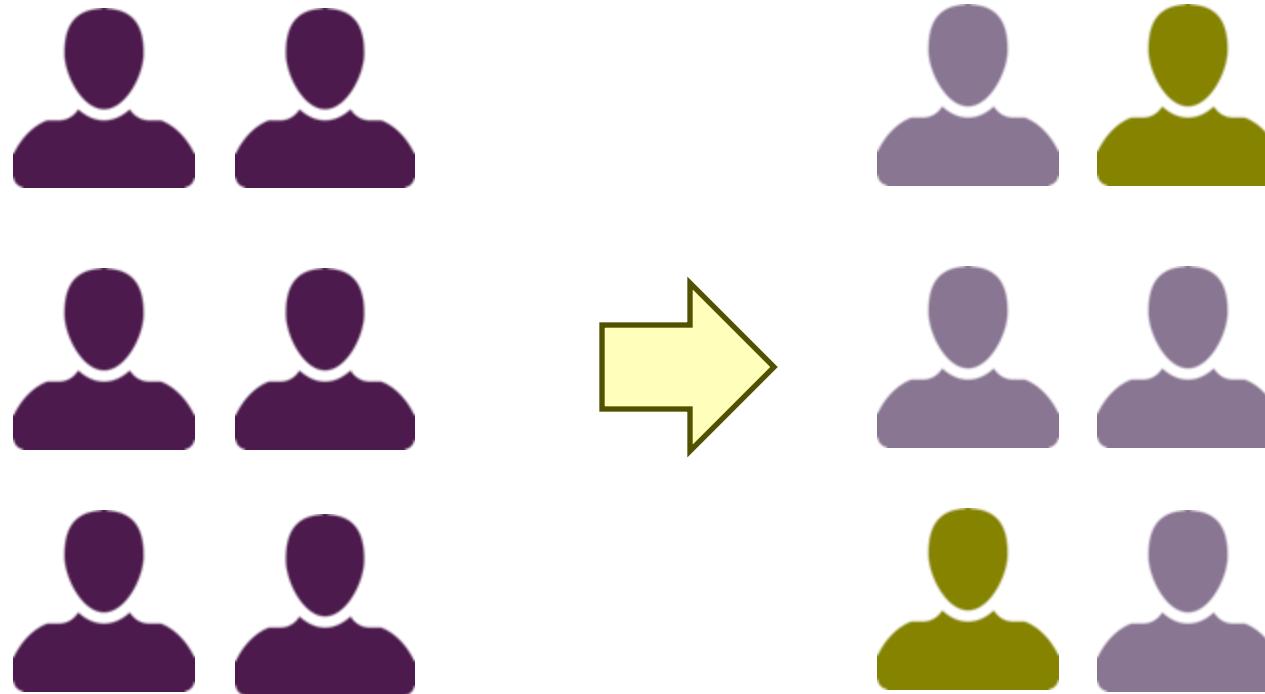
- Identify to which set of categories a new observation belong

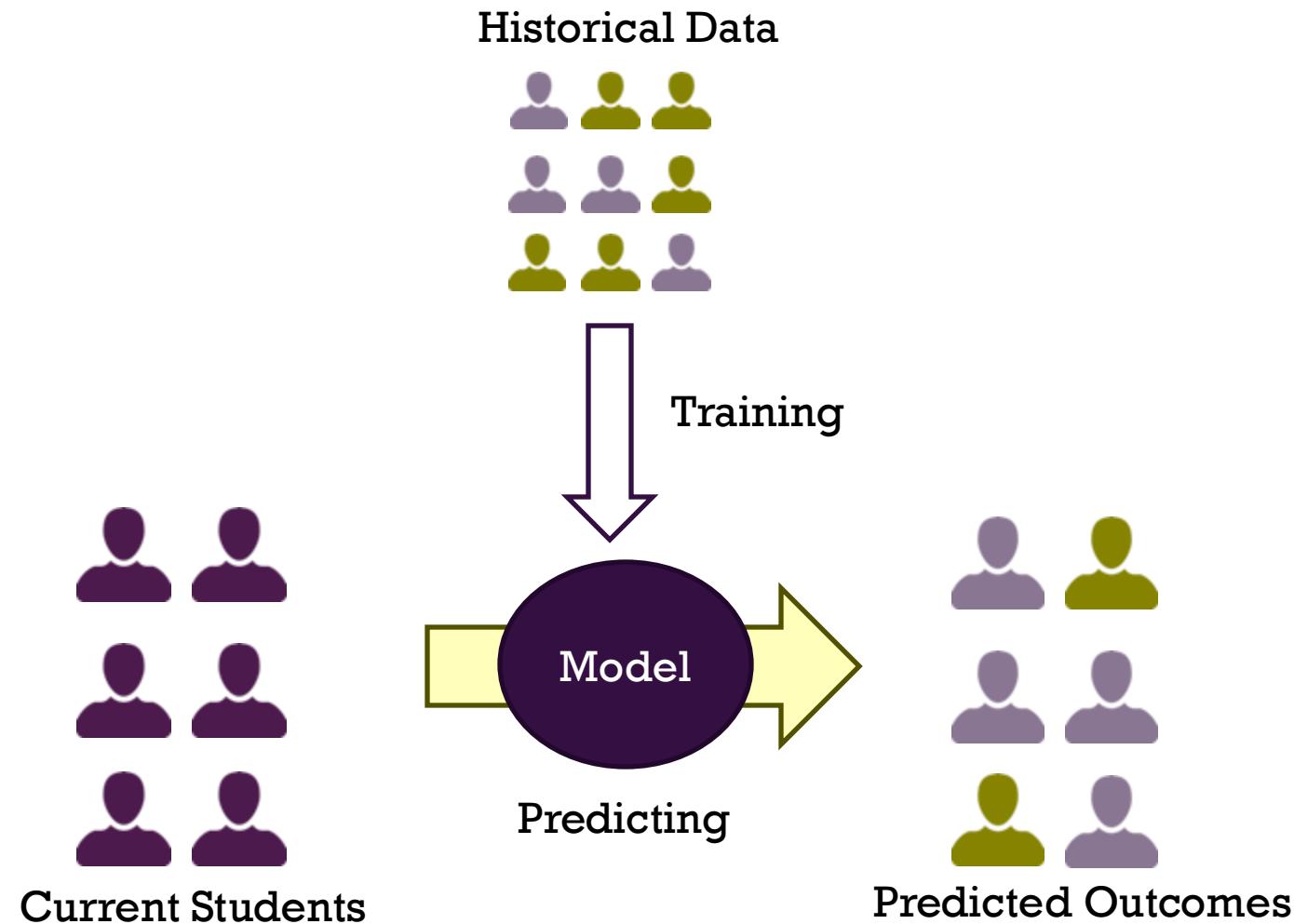


- Example: spam filtering, customer churn prediction, complaint classification



Example: Students Grade Prediction





$$\frac{OS \times Data\ Struct \times Prog}{9} > 7$$

Example: Amazon Recommendation

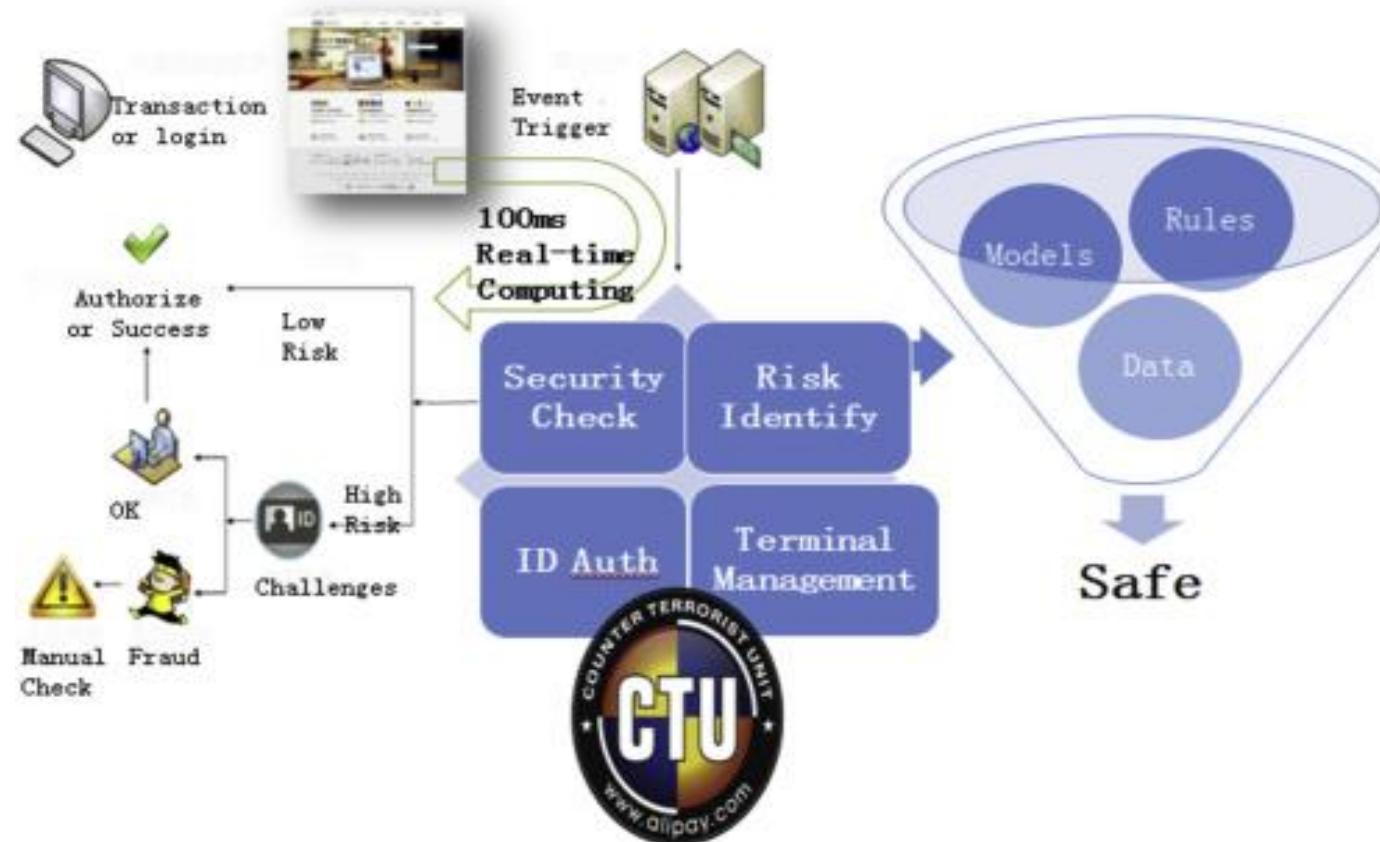
- Amazon sells 480M products (485k new products per day)
- Use recommendation systems to bring products to customers
- Analyze data from 300M customers
 - Purchase history
 - Reviews / Ratings
 - Search history
 - Views

The screenshot shows the top navigation bar of the Amazon website, including 'Browsing History', 'Natawut's Amazon.com', 'Today's Deals', and 'Gift Cards'. Below the header, it says 'Natawut's Amazon' and 'You could be seeing useful stuff here! Sign in to get your order status, balances and more...'. The main content area displays recommended products:

- Computer & Technology Books**: 92 ITEMS. Includes a book cover for 'Hadoop Application Architectures' by Mark Grover, Ted Malaska, Jonathan Seidman & Gwen Shapira.
- Science & Math Books**: 51 ITEMS. Includes a book cover for 'Storytelling with Data' by Cole Nussbaumer Knaflic.
- Storytelling**: A Guide on How to Tell a Story, With Storytelling Techniques and Storytelling Secrets. Includes a book cover for 'Own the Room' by Robert J. Knell.
- Introductory Machine Learning**: Fully revised and expanded version. Includes a book cover for 'Machine Learning for Absolute Beginners' by Robert J. Knell.



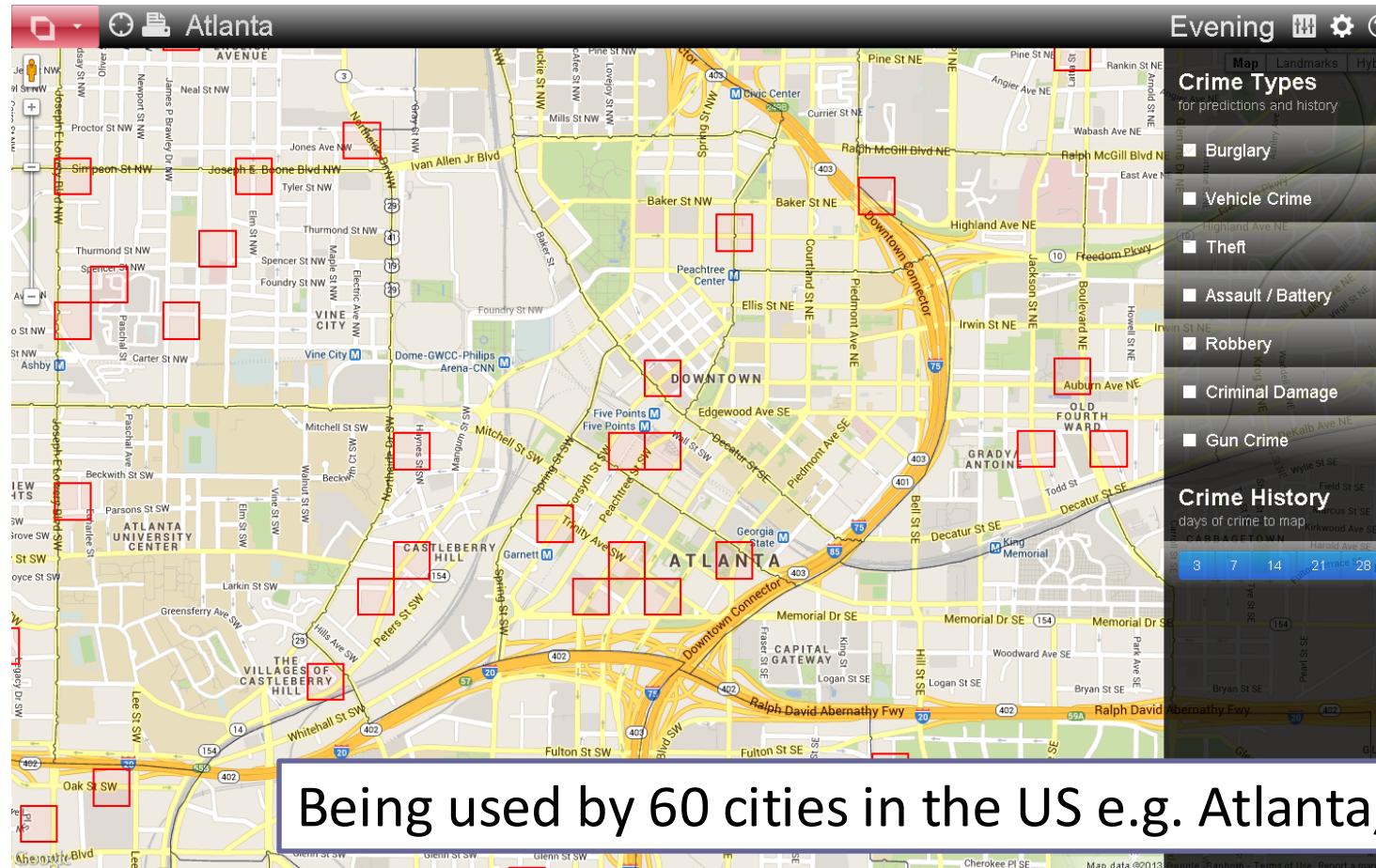
Case study: Alibaba Fraud Detection



Source: <http://www.sciencedirect.com/science/article/pii/S2405918815000021>



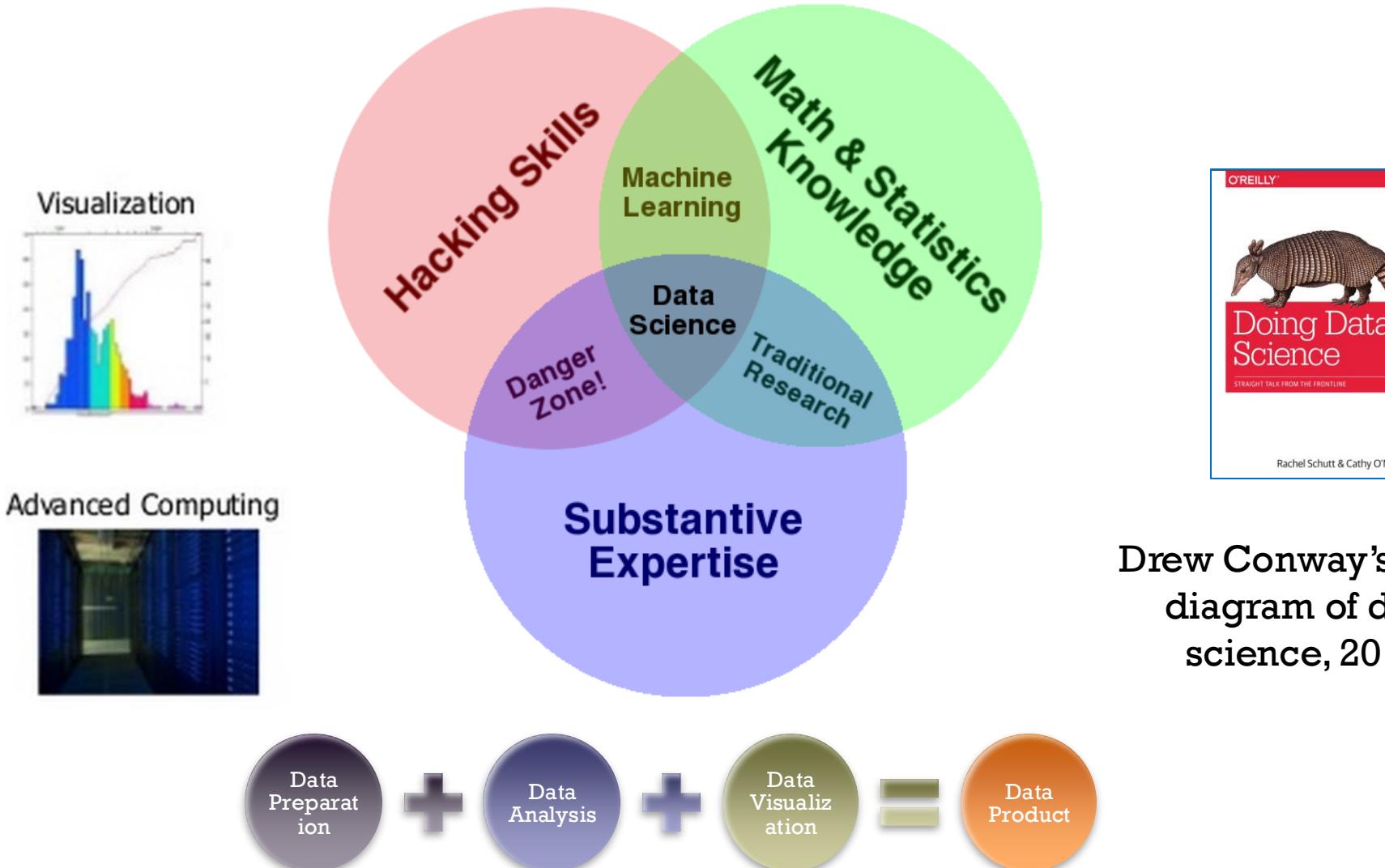
Case study: Predictive Policing



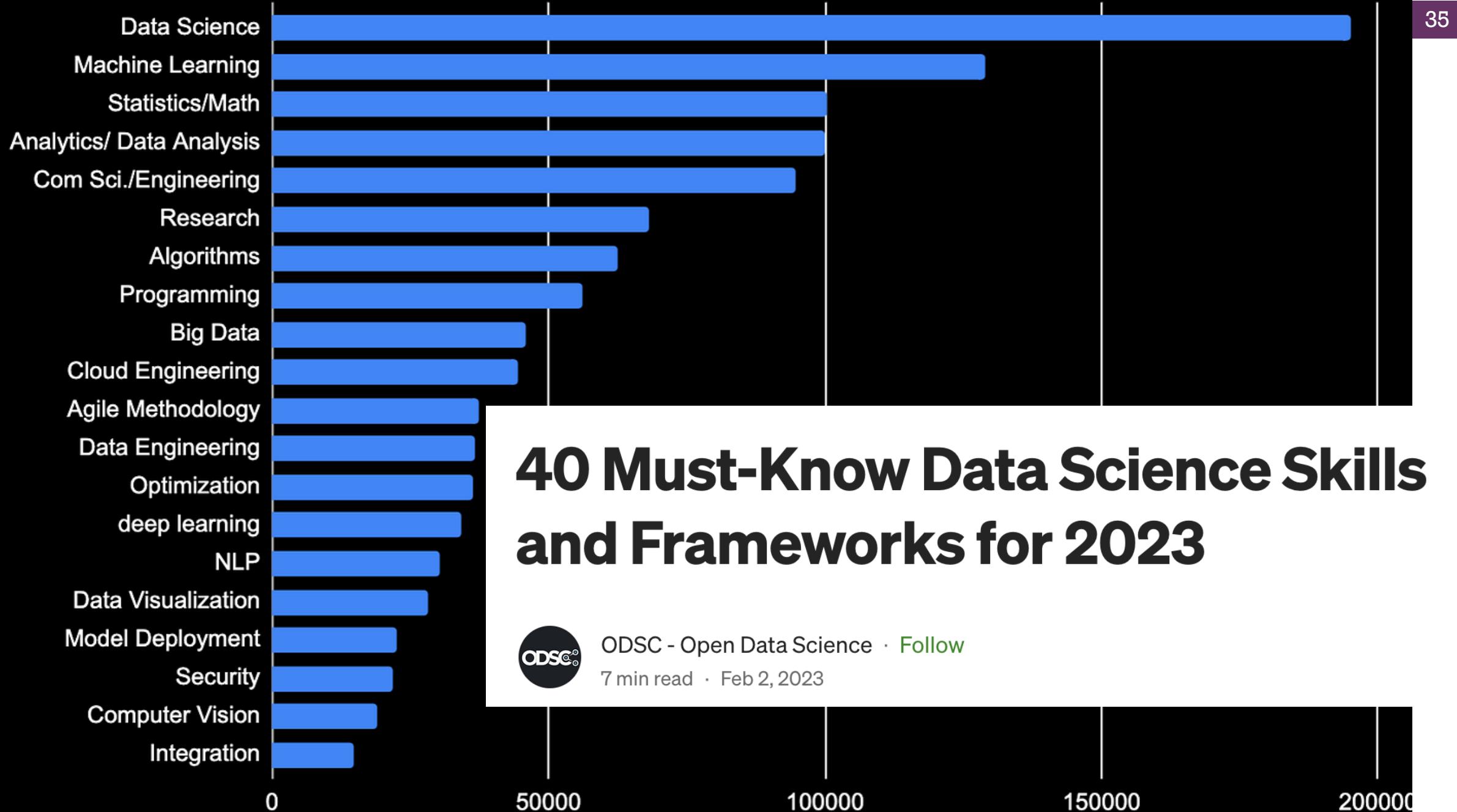
Source: <http://www.forbes.com/sites/ellenhuet/2015/02/11/predpol-predictive-policing>



Drew Conway's Data Science Venn diagram (Skills)



Drew Conway's Venn diagram of data science, 2010



40 Must-Know Data Science Skills and Frameworks for 2023



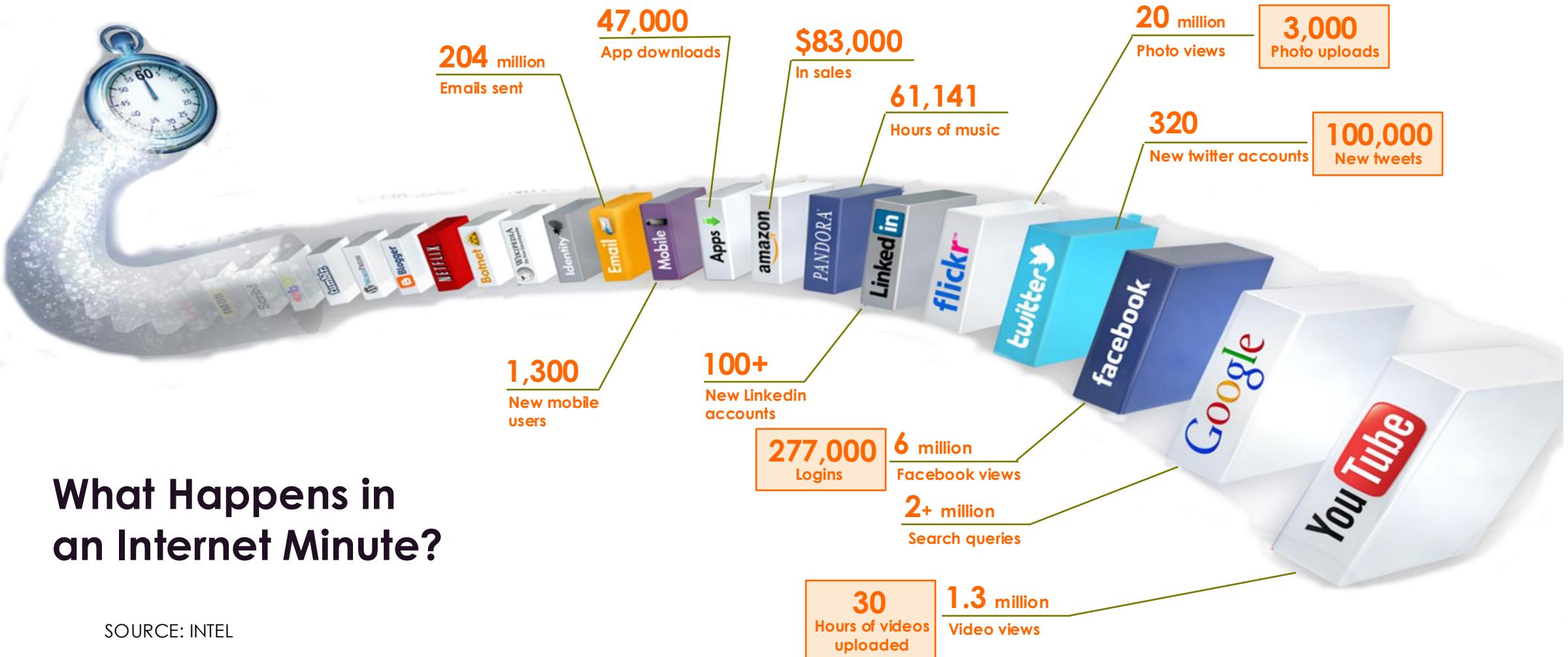
ODSC - Open Data Science · Follow

7 min read · Feb 2, 2023



Big Data

Big Data Explosion



40 ZETTABYTES

[43 TRILLION GIGABYTES] of data will be created by 2020, an increase of 300 times from 2005

6 BILLION PEOPLE
have cell phones

Volume SCALE OF DATA

2005

2020

It's estimated that
2.5 QUINTILLION BYTES
[2.3 TRILLION GIGABYTES] of data are created each dayMost companies in the U.S. have at least
100 TERABYTES
[100,000 GIGABYTES] of data stored

The FOUR V's of Big Data

From traffic patterns and music downloads to web history and medical records, data is recorded, stored, and analyzed to enable the technology and services that the world relies on every day. But what exactly is big data, and how can these massive amounts of data be used?

As a leader in the sector, IBM data scientists break big data into four dimensions: **Volume**, **Velocity**, **Variety** and **Veracity**.

Depending on the industry and organization, big data encompasses information from multiple internal and external sources such as transactions, social media, enterprise content, sensors and mobile devices. Companies can leverage data to adapt their products and services to better meet customer needs, optimize operations and infrastructure, and find new sources of revenue.

By 2015
4.4 MILLION IT JOBS
will be created globally to support big data, with 1.9 million in the United States.



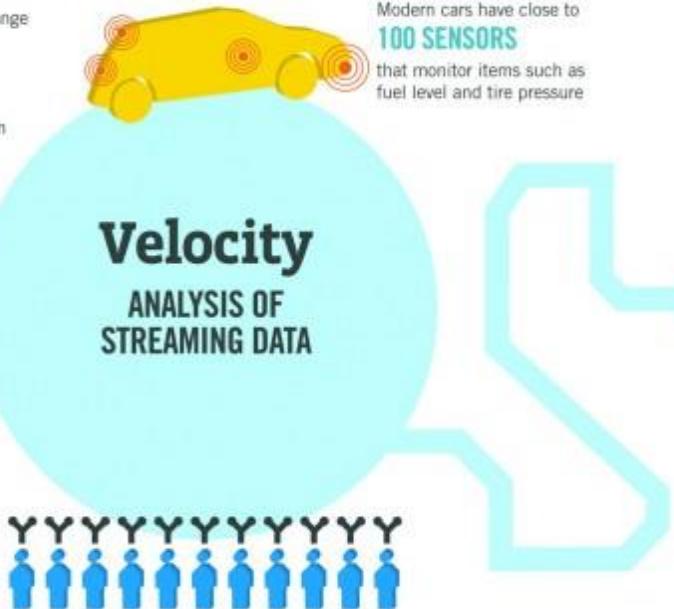
The New York Stock Exchange captures
data at a rate of

1 TB OF TRADE INFORMATION
during each trading session


Velocity ANALYSIS OF STREAMING DATA

2005

2020

Modern cars have close to
100 SENSORS
that monitor items such as fuel level and tire pressure

By 2016, it is projected there will be

18.9 BILLION NETWORK CONNECTIONS

– almost 2.5 connections per person on earth



As of 2011, the global size of data in healthcare was estimated to be

150 EXABYTES

[161 BILLION GIGABYTES]


30 BILLION PIECES OF CONTENT

are shared on Facebook every month



Variety DIFFERENT FORMS OF DATA



By 2014, it's anticipated there will be
420 MILLION WEARABLE, WIRELESS HEALTH MONITORS

4 BILLION+ HOURS OF VIDEO

are watched on YouTube each month

**400 MILLION TWEETS**

are sent per day by about 200 million monthly active users



1 IN 3 BUSINESS LEADERS

don't trust the information they use to make decisions



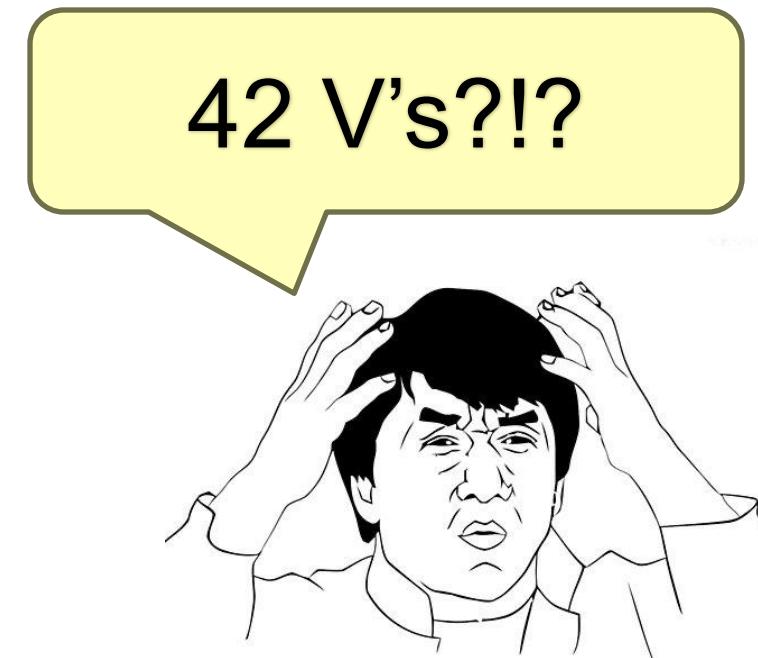
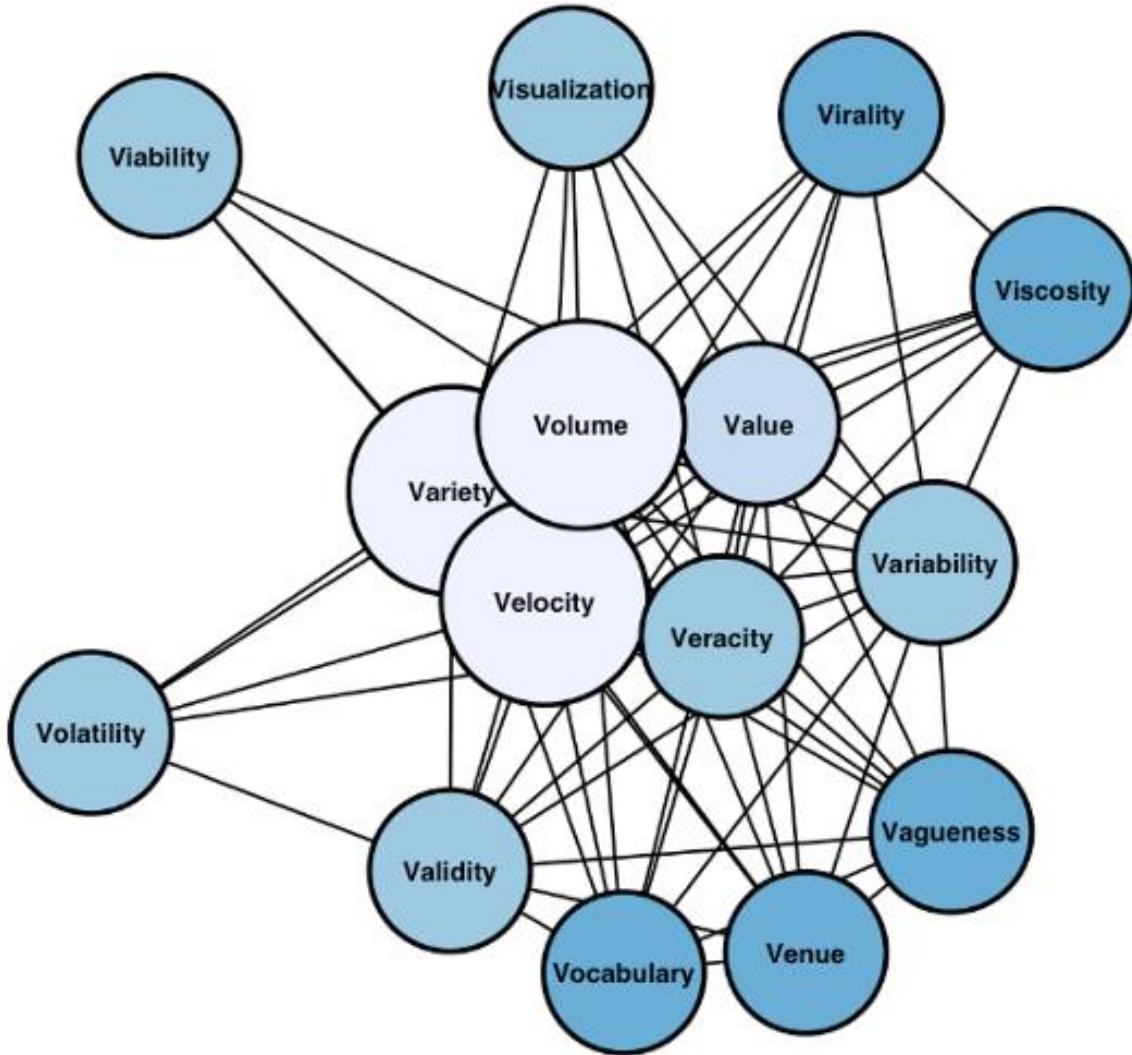
in one survey were unsure of how much of their data was inaccurate

Veracity UNCERTAINTY OF DATA

Poor data quality costs the US economy around

\$3.1 TRILLION A YEAR

Now 42 V's of Big Data



Big Data Driver: Internal + External Data





A Better Way to Know

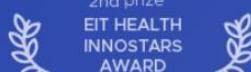
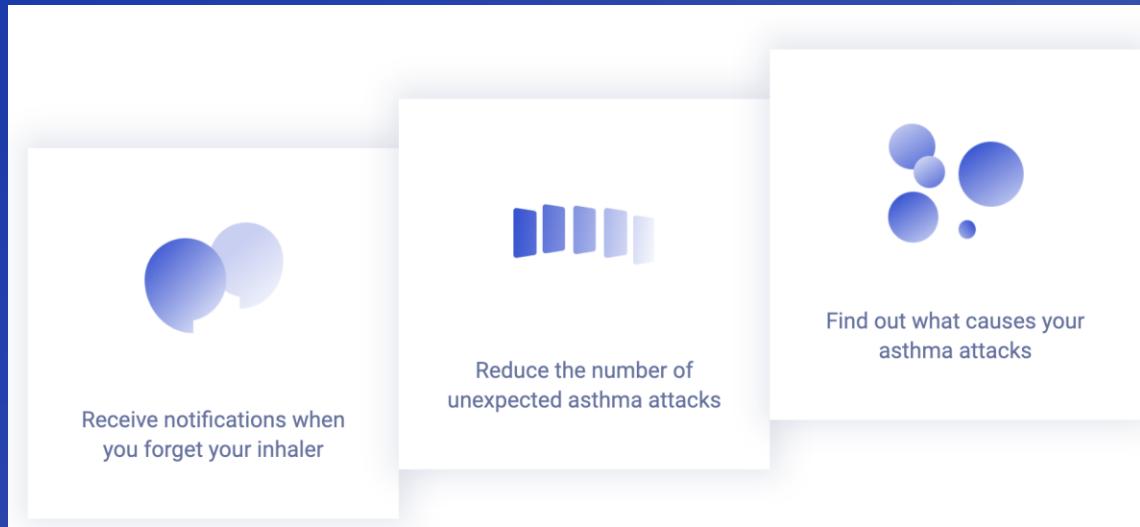
We understand what it's like to hover over the crib at night. That's why we invented the Owlet Smart Sock. It's a better way to check on your baby and smarter way to know they're okay.

Every Beat. Every Breath.

KNOWING IS BEST

Control your asthma with FindAir smart inhaler

Full Control. Less asthma attacks. Better life.



2nd prize
EIT HEALTH
INNOSTARS
AWARD
2018



1st prize
UPC DIGITAL
IMAGINATION
CHALLENGE
2017



Best Pitch
SMART
HEALTH
BUDAPEST
2017



1st prize
INNOLABS
DIGITAL HEALTH
HACKATHON
2017



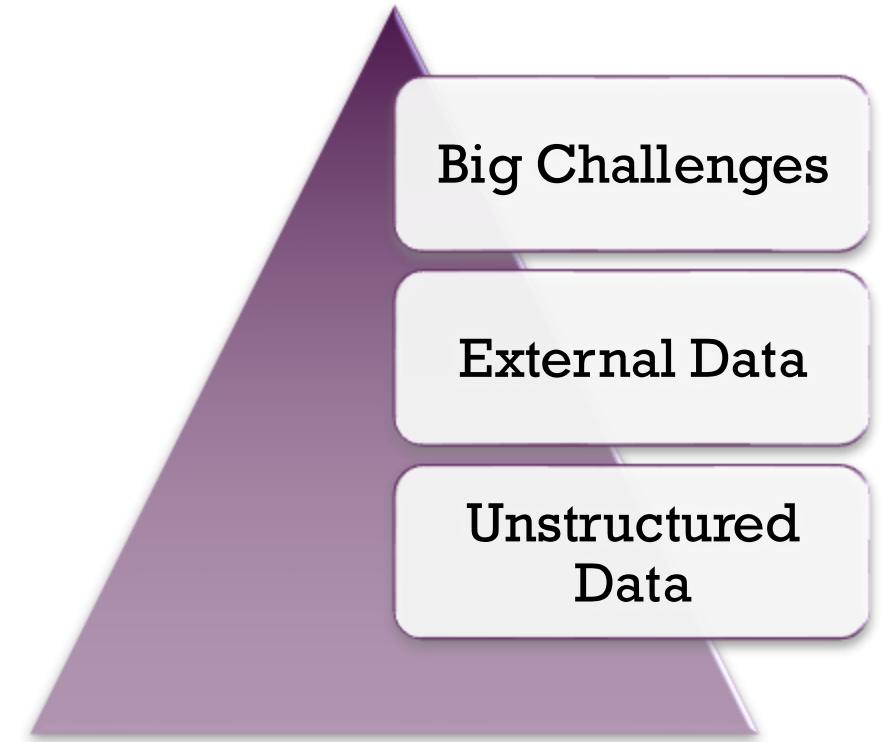
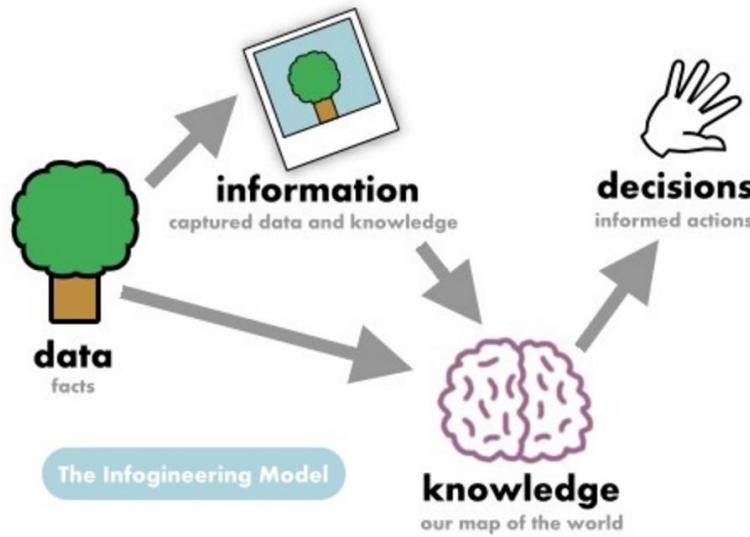
1st prize
INNOLABS
DIGITAL HEALTH
HACKATHON
2017



2018

Big Data Analytics

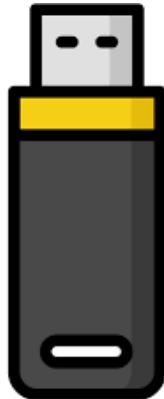
- It is a process of examining **Big Data** to uncover useful information and knowledge.
- More data means better decision!



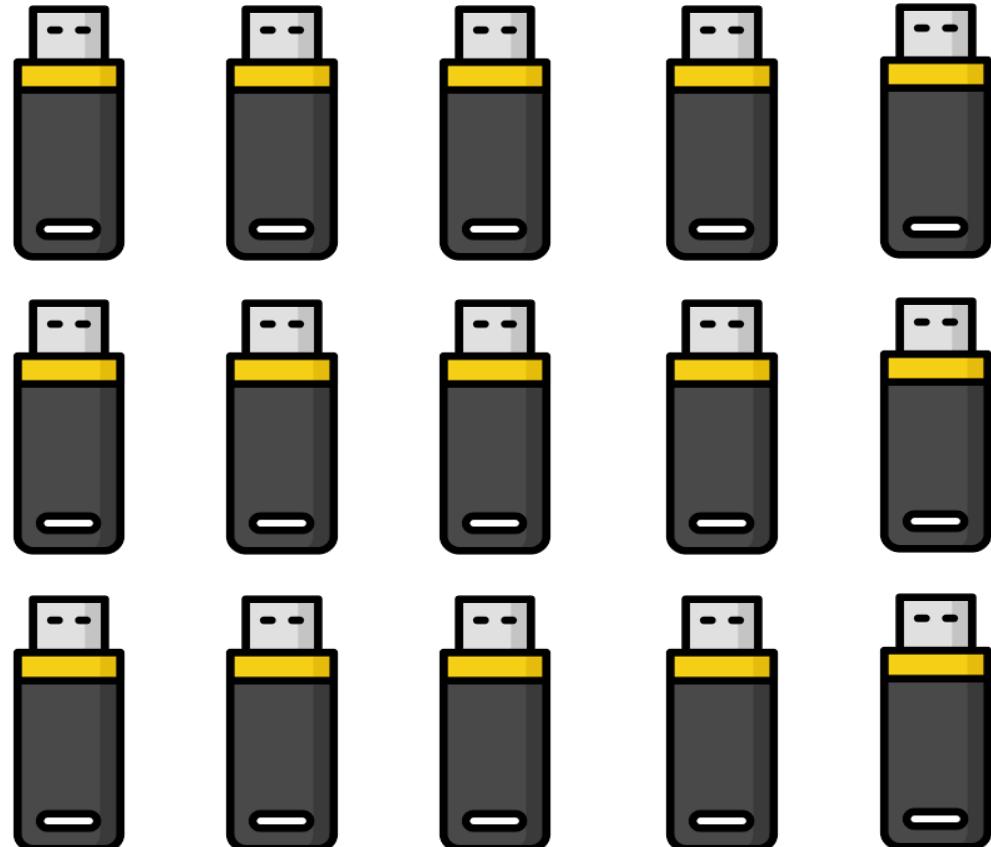
Big Data Challenges

Same tasks, but much more difficult!

2MB



200TB



Big Data Solution



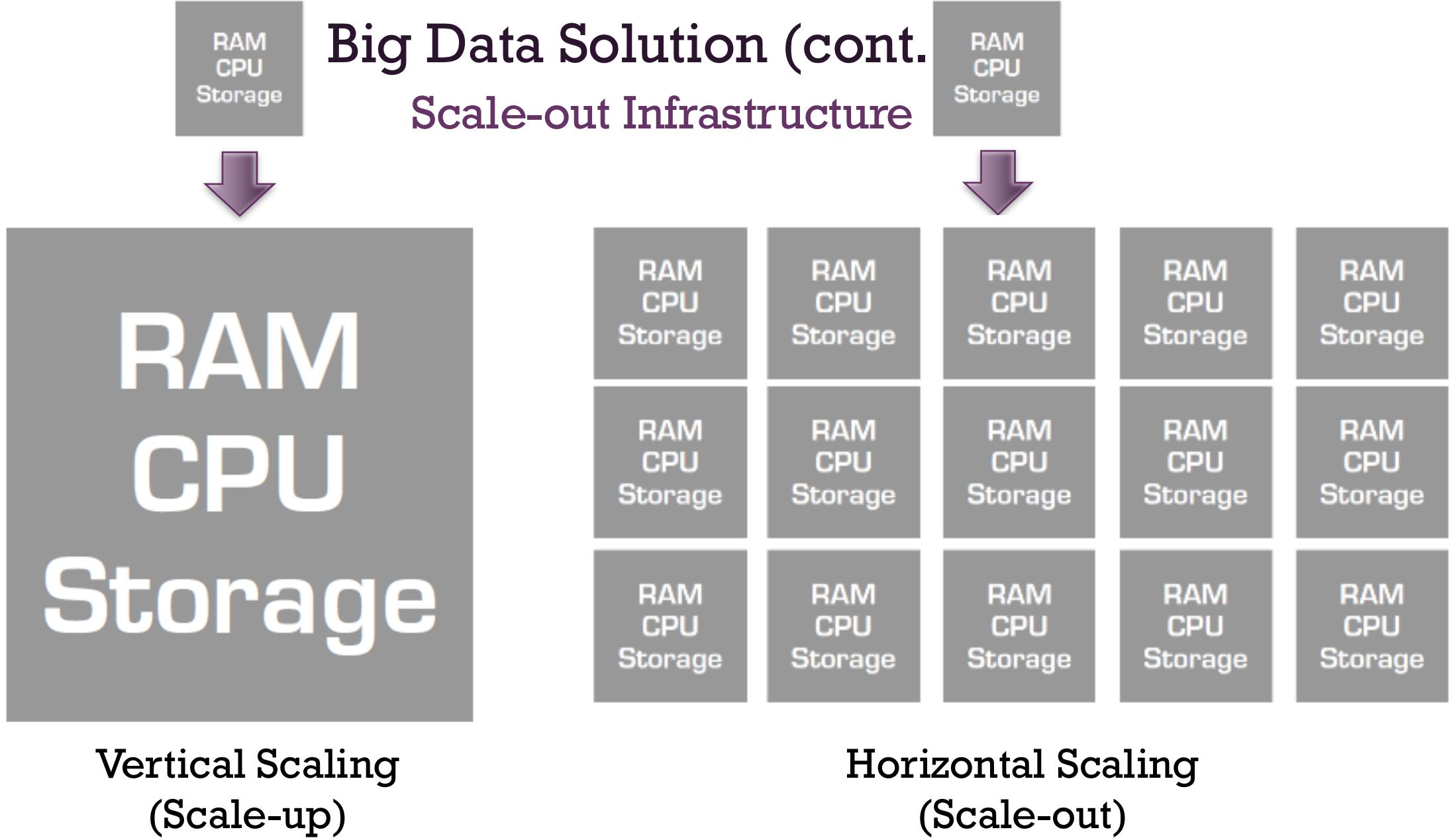
INFRASTRUCTURE



ALGORITHM

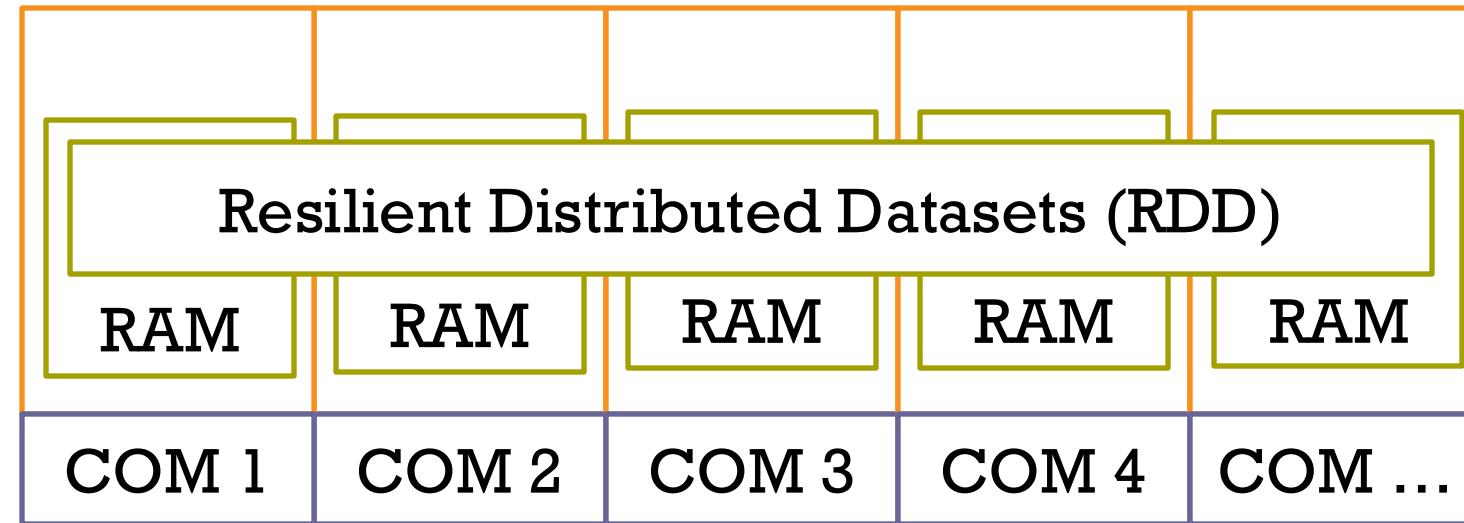
Big Data Solution (cont.)

Scale-out Infrastructure

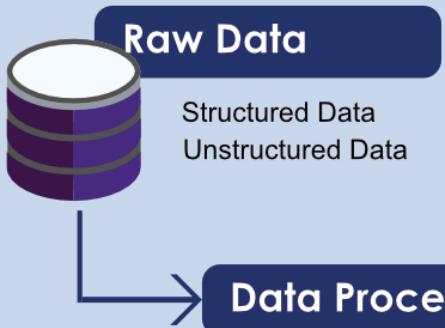


Big Data Solution (cont.)

In-memory & Distributed Computing



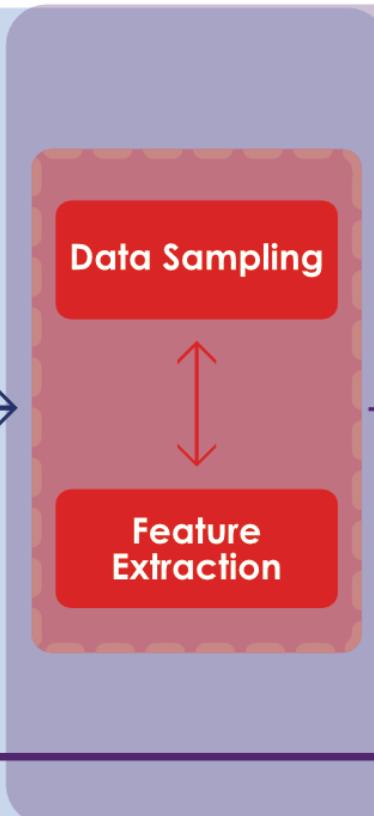
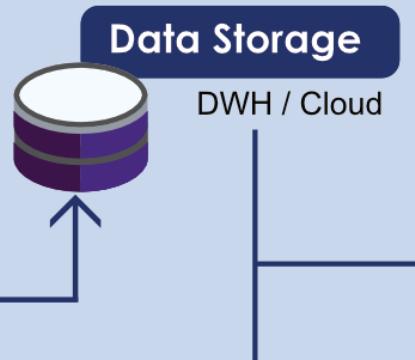
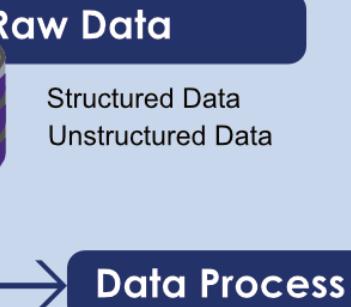
DATA ENGINEER



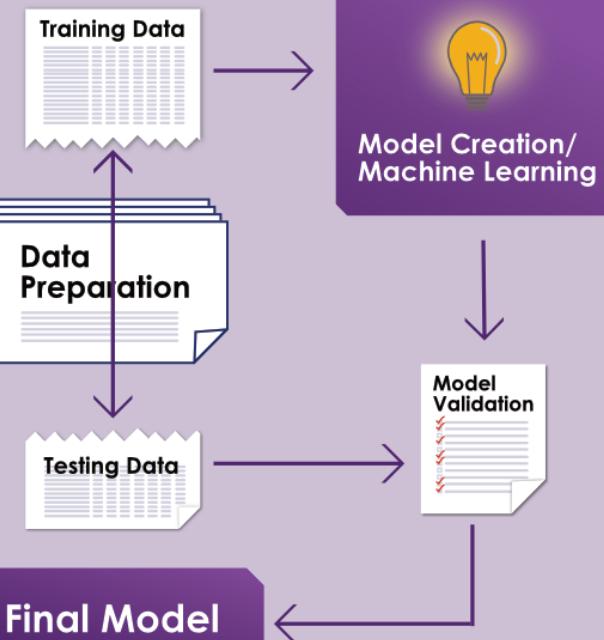
SQL
NoSQL
Python

Hadoop
Spark

- Clustering
- Cleaning
- Structuring
- Streaming
- Extract ,Transform and Load (ETL)



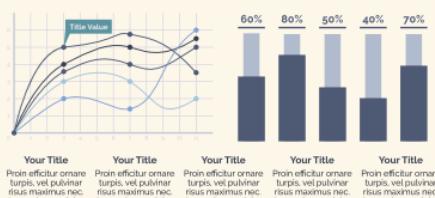
DATA SCIENTIST



Final Model

DATA ANALYST

Data Visualization
Bi
Application
Report
Action



BigData Ecosystem





<https://blog.datath.com/data-engineer-guide/>



Top Chef Thailand ตอนสุดท้าย ที่ผู้เข้าแข่งขันต้องช่วยกันทำงานเป็นทีม – ขอบคุณรูปจาก one31

Data Engineer ก็เหมือนกับผู้ช่วยเชฟ มีหน้าที่จัดเตรียมข้อมูลจากแหล่งต่าง ๆ มารวมกันไว้ในจุดเดียว โดยต้องทำให้ข้อมูลมีความถูกต้อง และดูและระบบว่าทำงานได้ไม่เกิดปัญหาอะไร (ในชีวิตจริงนี่ต้องให้เราวางแผนมาดีแค่ไหน เจอข้อมูลเยอะ ๆ วันดีคืนดีก็กล่อมได้ครับ T_T)



Data Scientist + ML Engineer



Data Scientist

VS

Data Engineer



VS



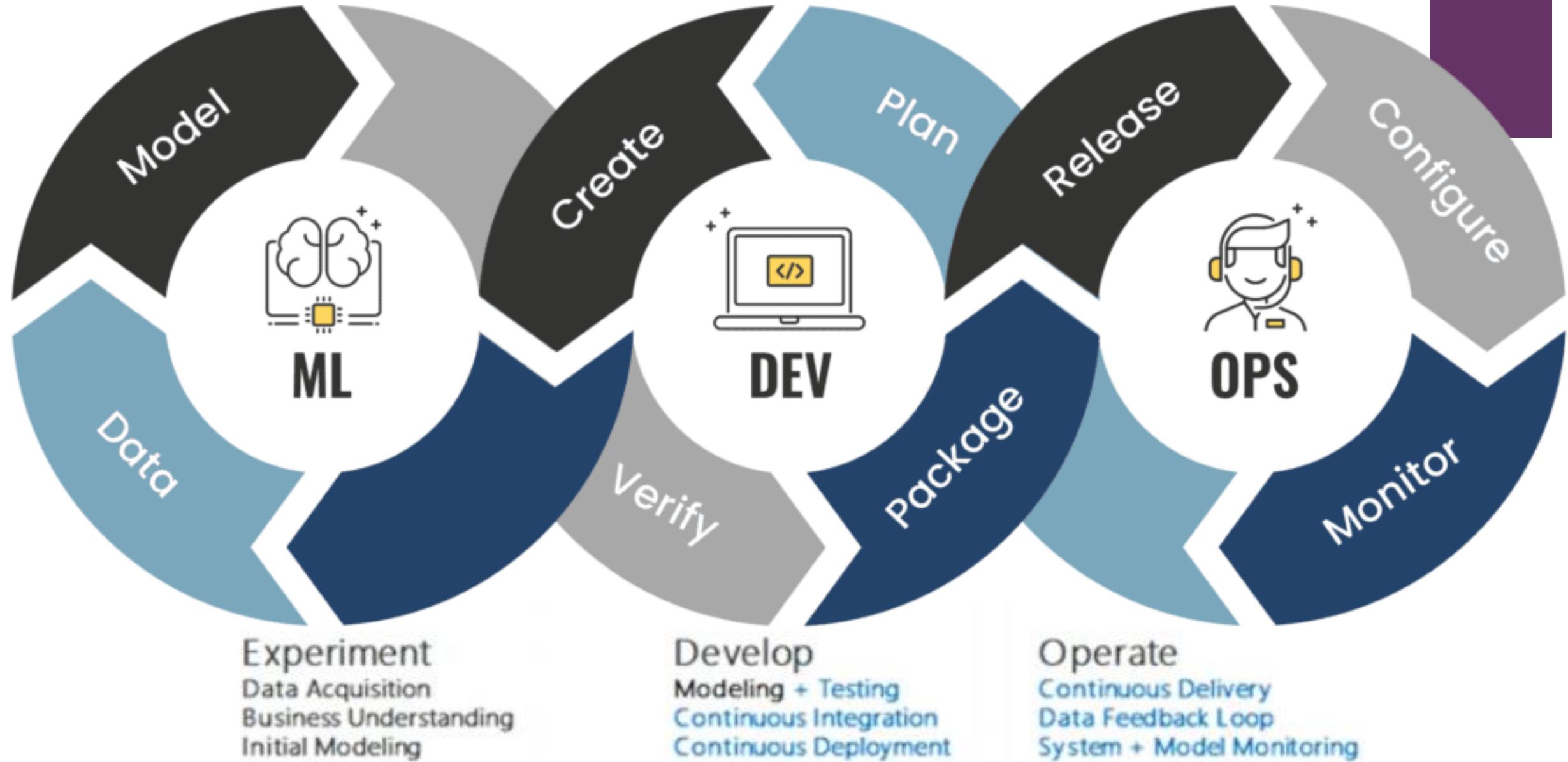
ML Engineer

VS



MLOps Engineer

+ MLOps = ML + DEV + OPS





Data Science Process

+

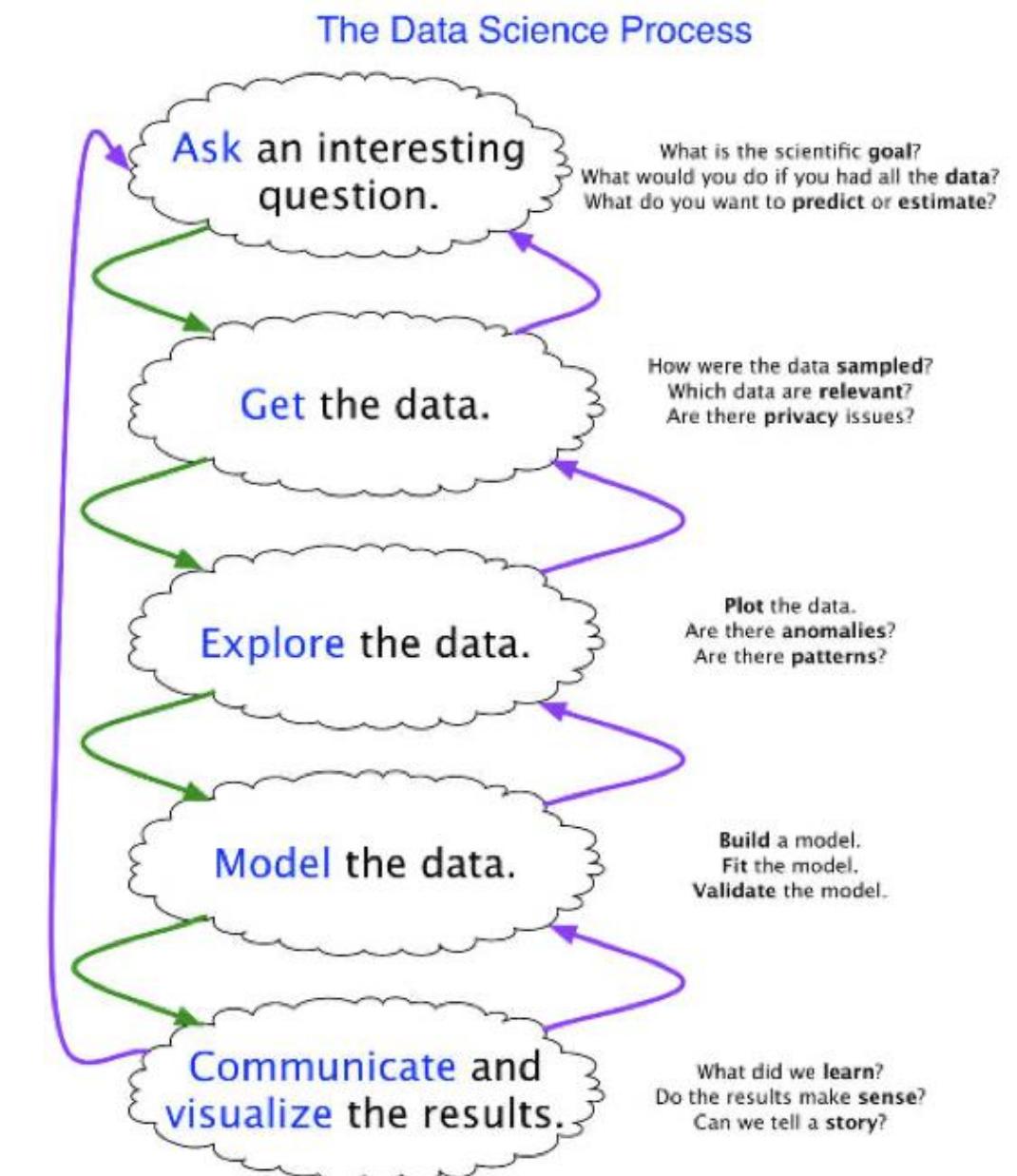
Data Science Process

Dr.Virote

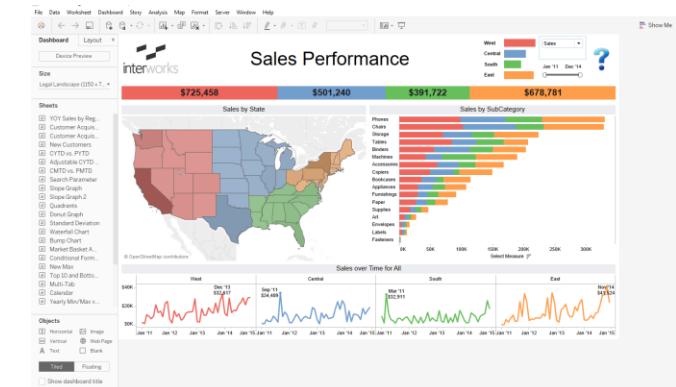
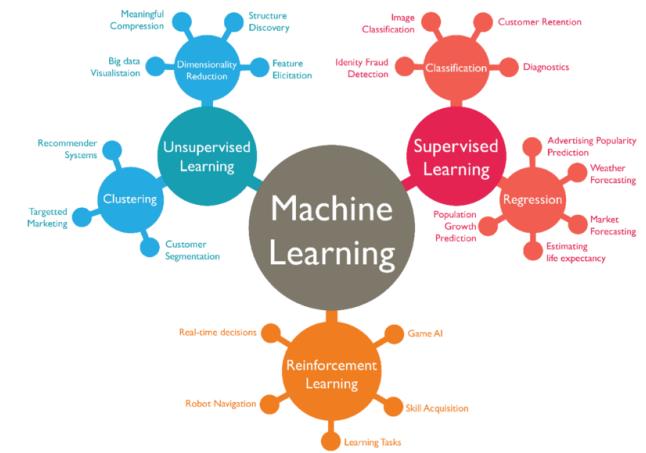
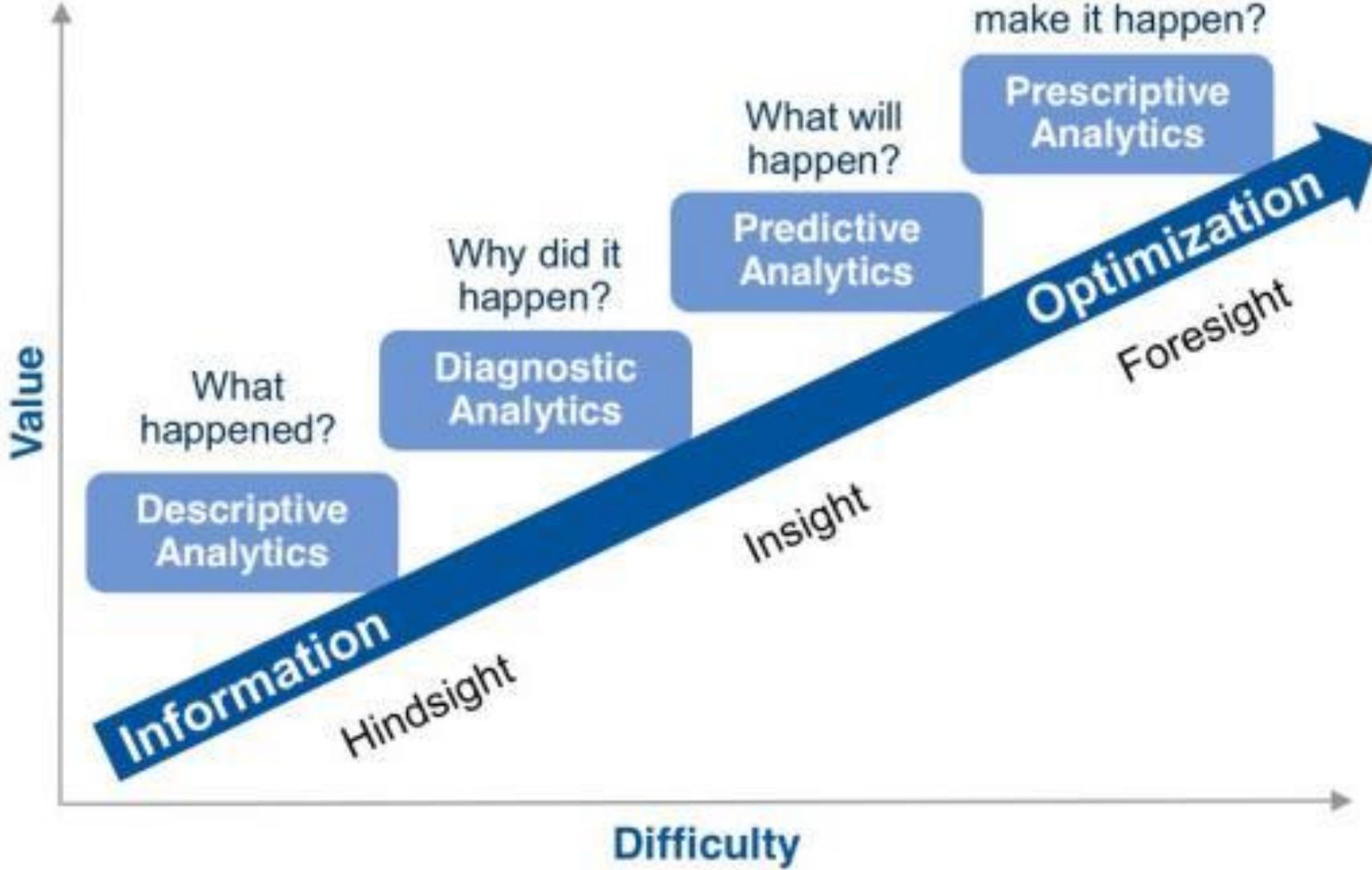
1. Transform data into **valuable insights**
2. Transform data into **data products**
3. Transform data into **interesting stories**

Aj.Natawut

1. Measurement (**decision**)
2. Insights (**knowledge**)
3. Data Products (**Innovation, Intelligent**)



Data Analytics (Data Science)



BIG DATA





Types of Data Science Projects

Valuable insights

- Data visualization
- Analytical skills & storytelling
 - Infographic



Advanced analytics

- AI/Machine Learning/Deep Learning
 - Prediction, Forecasting, Clustering, etc.



Ask a question about your data

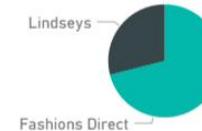
Total Stores
NEW & EXISTING STORES

104

This Year's Sales
NEW & EXISTING STORES

\$22M

This Year's Sales
BY CHAIN



New Stores, New Stores Target
YEAR TO DATE

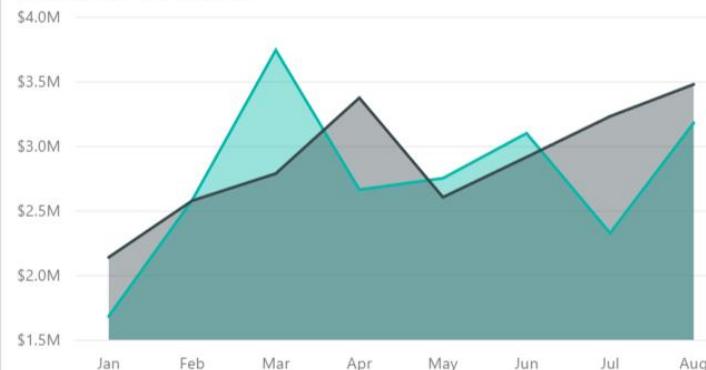


This Year's Sales
NEW STORES ONLY

\$2M

This Year's Sales, Last Year's Sales
BY FISCAL MONTH

This Year Sales ● Last Year Sales



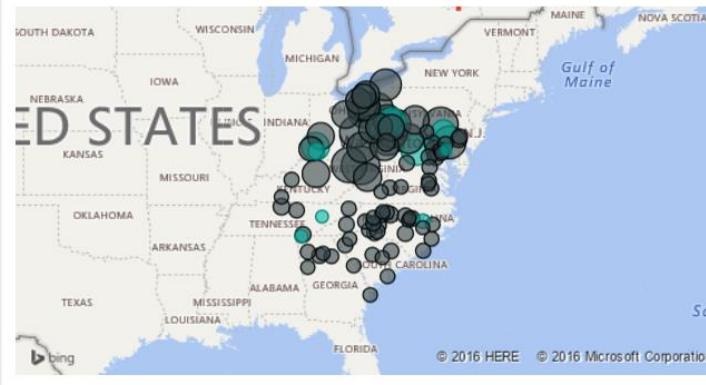
Total Sales Variance %, Sales Per Sq Ft, This Year's Sales
BY DISTRICT



New Stores
NEW STORES ONLY

10

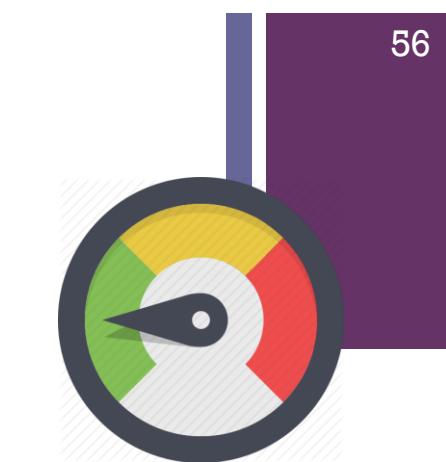
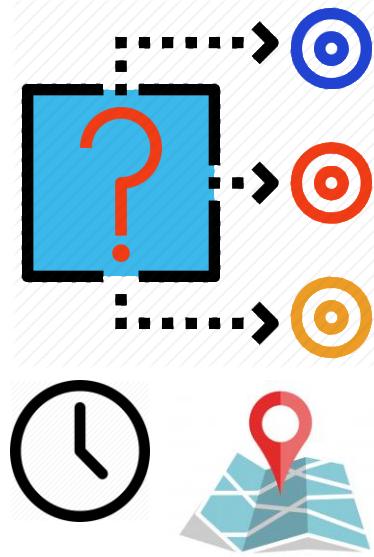
This Year's Sales
BY POSTAL CODE, STORE TYPE



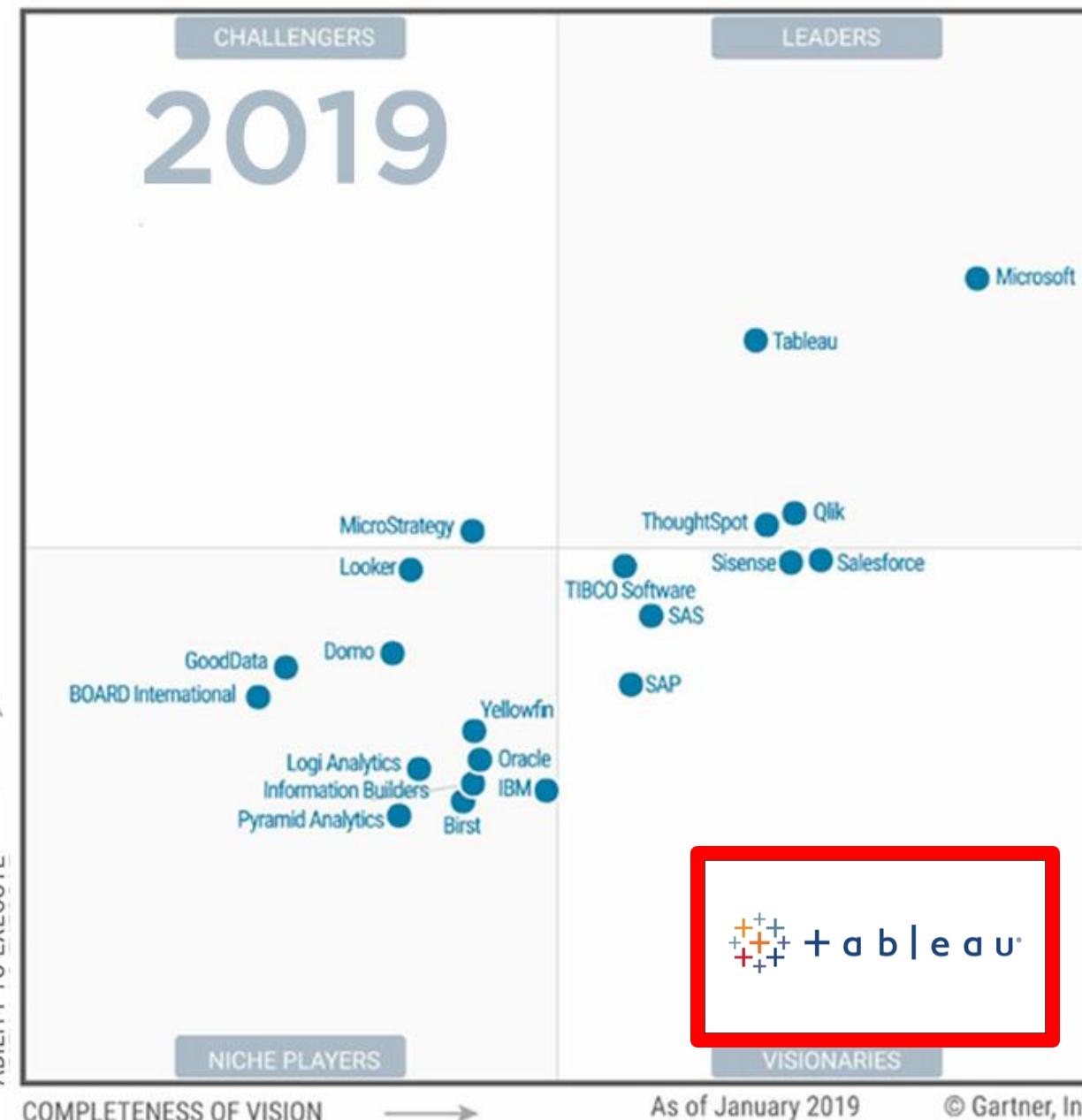
This Year's Sales
BY CITY, CHAIN



Sales Per Sq Ft
BY NAME



Magic Quadrant for Analytics and Business Intelligence Platforms.



Source: Gartner (Feb 2019 and 2020)



PRODUCT INNOVATOR



CHALLENGER

<https://dataforest.ai/blog/best-business-intelligence-tool-of-2023-top-16-bi-tools-by-dataforest>

LEADER
Microsoft Power B Tableau

SAP Analytics Cloud SAP MicroStrategy

Tableau CRM Tableau Oracle Analytics Cloud

Qlik Sense Qlik Zoho Analytics

Cognos Analytics IBM Analytics Amazon QuickSight

Dundas BI

Domo

Sisense

TIBCO Spotfire

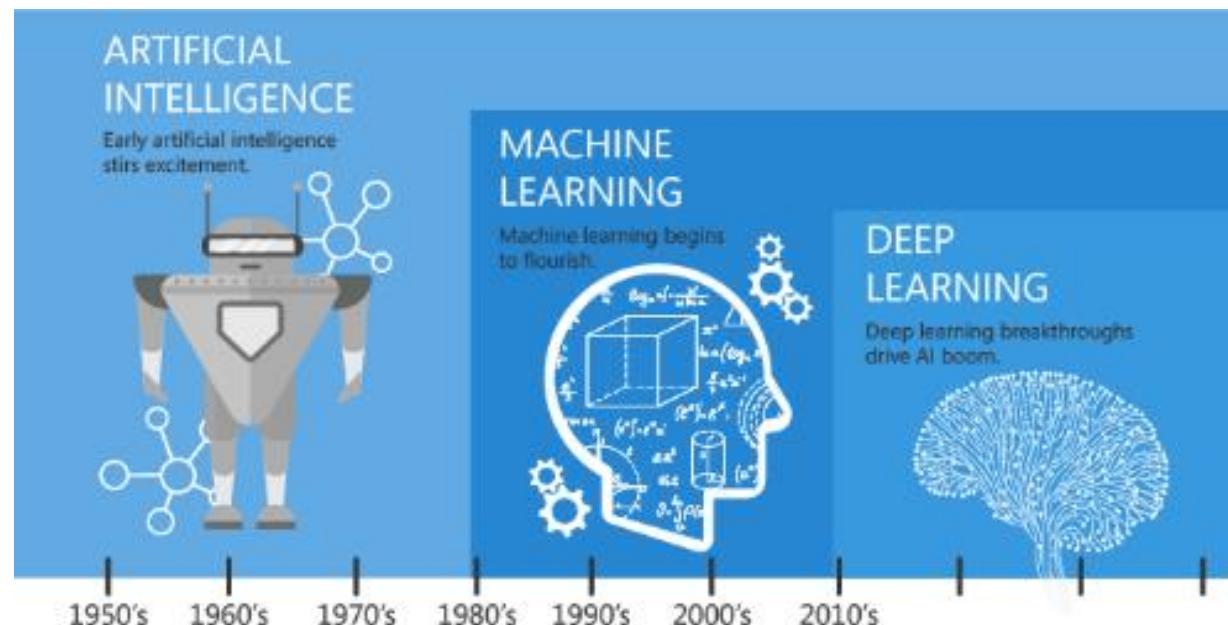
SAS Business Intelligence

SERVICE STAR

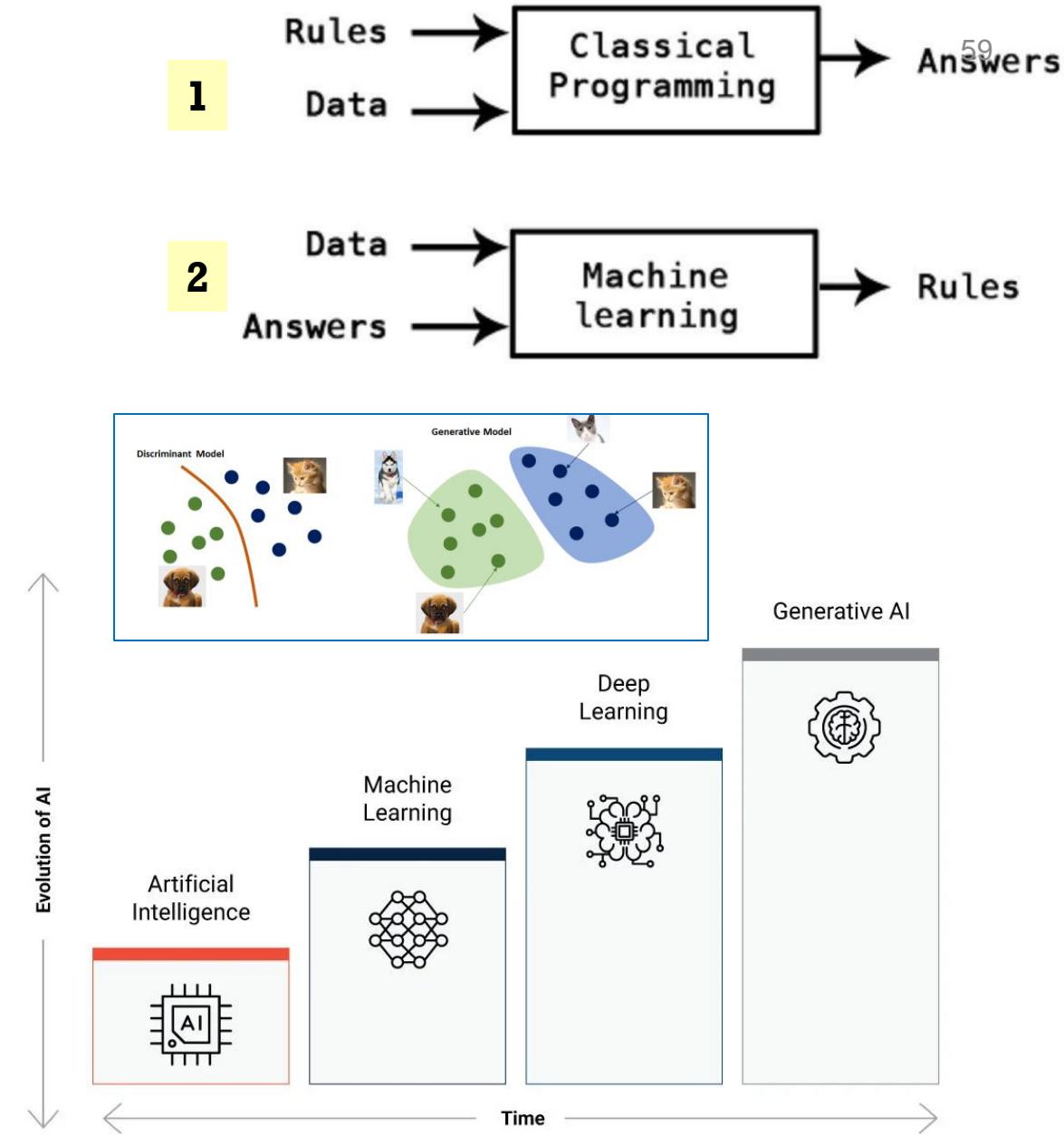


AI - Automation

- 1) Rule-based AI
- 2) Machine Learning (ML)



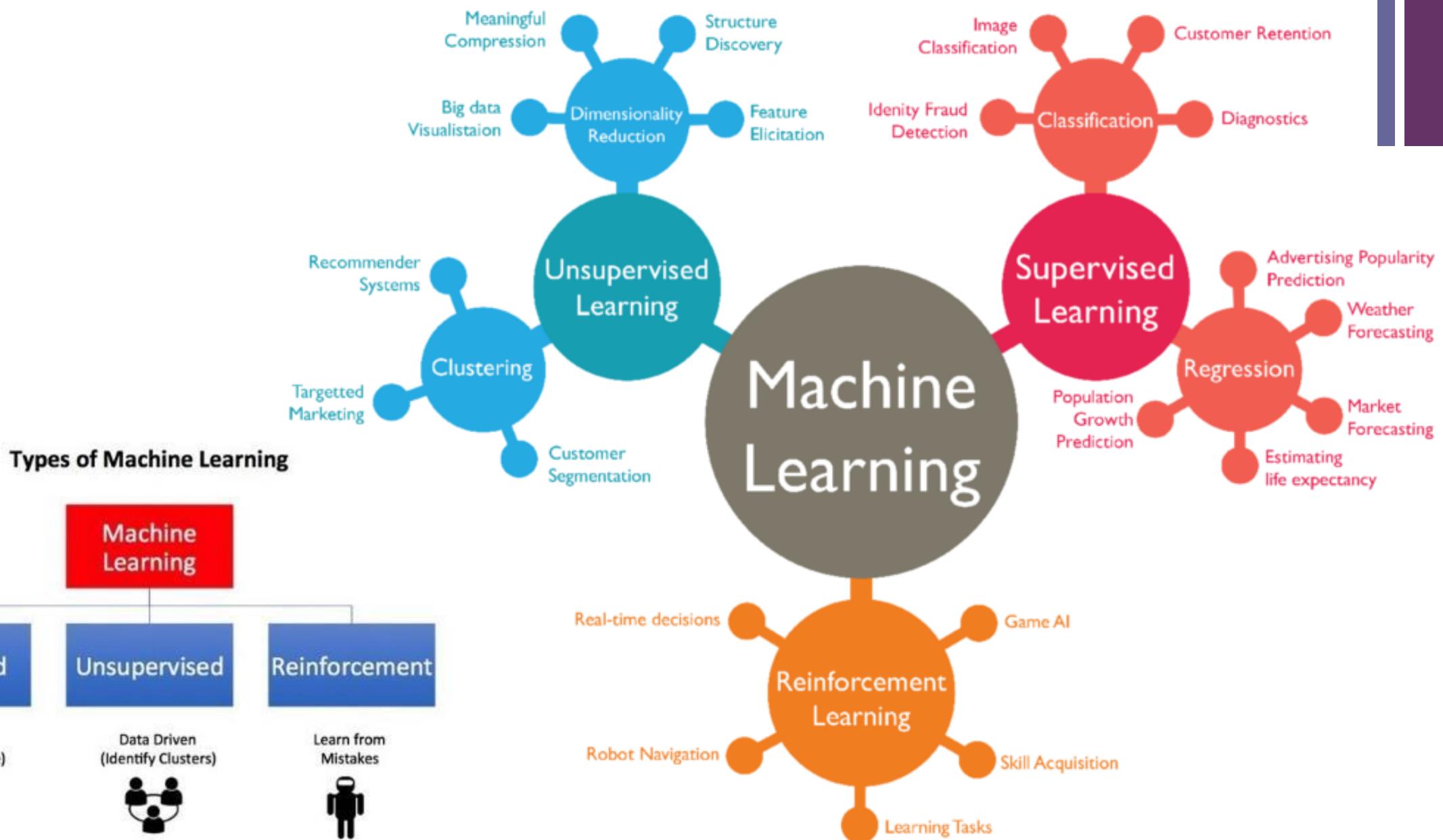
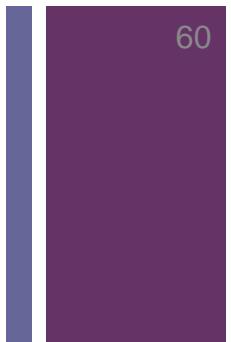
Since an early flush of optimism in the 1950's, smaller subsets of artificial intelligence - first machine learning, then deep learning, a subset of machine learning - have created ever larger disruptions.



<https://mc.ai/machine-learning-basics-artificial-intelligence-machine-learning-and-deep-learning/>

+ Machine Learning (ML)

60



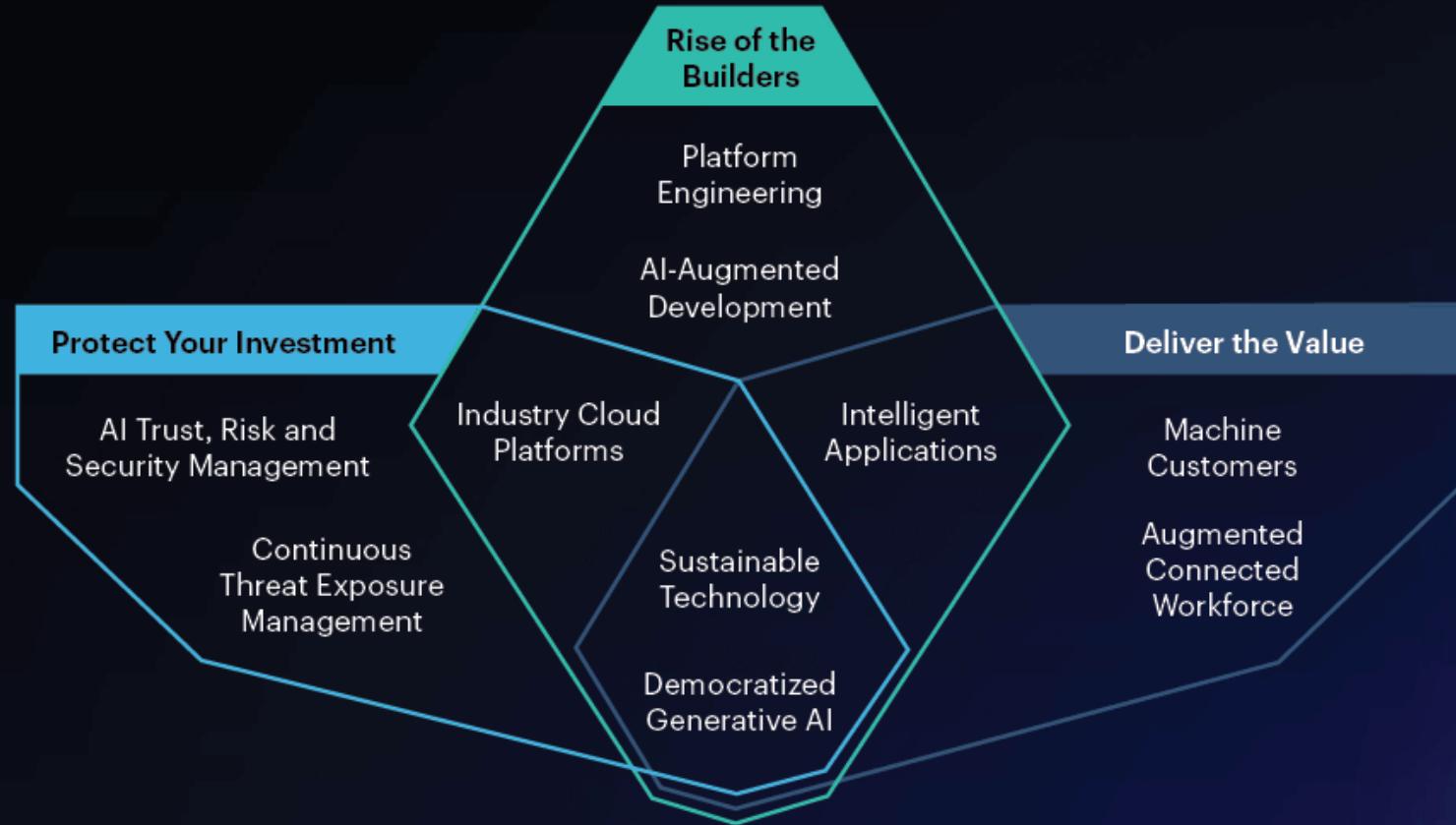
Top Strategic Technology Trends 2024

Top Strategic Technology Trends **2024**

- | | | |
|---|--|--|
| 1 AI Trust, Risk and Security Management | 2 Continuous Threat Exposure Management | 3 Sustainable Technology |
| 4 Platform Engineering | 5 AI-Augmented Development | 6 Industry Cloud Platforms |
| 7 Intelligent Applications | 8 Democratized Generative AI | 9 Augmented Connected Workforce |
| 10 Machine Customers | | |

Source: Gartner
© 2023 Gartner, Inc. and/or its affiliates. All rights reserved. CM_GTS_2080051

Gartner



Source: Gartner
© 2023 Gartner, Inc. and/or its affiliates. All rights reserved. CM_GTS_2080051

Gartner

Data Trend in 2024 (cont.)

- AI (AI everywhere & Gen AI) is the key component.
- Knowledge without action (Platform Engineering) is meaningless.
- Cloud technology is a modern infrastructure.



Gartner Magic Quadrant



Vit Niennattrakul, Ph.D.

Figure 1: Magic Quadrant for Cloud Infrastructure and Platform Services



Figure 1: Magic Quadrant for Cloud AI Developer Services



Categories of AWS services



Vit Niennattrakul, Ph.D.



Analytics



Application
Integration



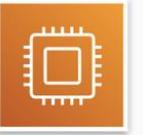
AR and VR



Blockchain



Business
Applications



Compute



Cost
Management



Customer
Engagement



Database



Developer Tools



End User
Computing



Game Tech



Internet
of Things



Machine
Learning



Management and
Governance



Media Services



Migration and
Transfer



Mobile



Networking and
Content Delivery
network



Robotics



Satellite



Security, Identity, and
Compliance



Storage

This document is protected under the copyright laws of Thailand and contains information that is proprietary and confidential to DailiTech, which shall not be disclosed outside the recipient's company or duplicated, used, or disclosed in whole or in part by the recipient for any purpose other than to the purpose of this document. Any other use or disclosures in whole or in part of this information without the express written permission of DailiTech are prohibited.



Vit Niennattrakul, Ph.D.



Amazon Forecast



Amazon Comprehend



Amazon SageMaker



Amazon Rekognition

Amazon Lex
aws

Notebooks: data exploration, experimentation



AWS Academy Service

AWS Academy Learner Lab

- Amazon API Gateway
- AWS App Mesh
- Application Auto Scaling
- AWS AppSync
- Amazon Athena
- Amazon Aurora
- AWS Backup
- AWS Certificate Manager (ACM)
- AWS Batch
- AWS Cloud9
- AWS CloudFormation
- Amazon CloudFront
- Amazon CloudSearch
- AWS CloudShell
- AWS CloudTrail
- Amazon CloudWatch
- AWS CodeCommit
- AWS CodeDeploy
- Amazon CodeWhisperer
- AWS Config
- AWS Systems Manager (SSM)
- Amazon Textract
- AWS Cost and Usage Report
- AWS Cost Explorer
- AWS Data Pipeline
- AWS DeepComposer
- AWS DeepLens
- AWS DeepRacer
- AWS Directory Service
- Amazon EC2 Auto Scaling
- AWS Elastic Beanstalk
- Amazon Elastic Block Store (EBS)
- Amazon Elastic Container Registry (ECR)
- Amazon Elastic Container Service (ECS)
- Amazon Elastic File System (EFS)
- Amazon Elastic Inference
- Amazon Elastic Kubernetes Service (EKS)
- Elastic Load Balancing (ELB)
- Amazon Elastic MapReduce (EMR)
- Amazon ElastiCache
- Amazon EventBridge
- AWS Fargate
- Amazon Timestream
- AWS Trusted Advisor
- Amazon Forecast
- AWS Glue
- AWS Glue DataBrew
- Amazon GuardDuty
- AWS Health
- AWS Identity and Access Management (IAM)
- AWS IAM Access Analyzer
- Amazon Inspector
- AWS IoT 1-Click
- AWS IoT Analytics
- AWS IoT Core
- AWS IoT Greengrass
- Amazon Kendra
- AWS Key Management Service (KMS)
- Amazon Kinesis
- Amazon Lex
- Amazon Machine Learning (Amazon ML)
- AWS Marketplace Subscriptions
- AWS Mobile Hub
- Amazon Neptune
- Amazon Virtual Private Cloud (Amazon VPC)
- AWS WAF - Web Application Firewall
- AWS OpsWorks
- Amazon Personalize
- Amazon QuickSight
- Amazon Redshift
- Amazon Relational Database Service (RDS)
- AWS Resource Groups & Tag Editor
- AWS RoboMaker
- Amazon Route 53
- AWS Secrets Manager
- AWS Security Hub
- AWS Security Token Service (STS)
- AWS Serverless Application Repository (SAR)
- AWS Service Catalog
- Amazon Simple Notification Service (SNS)
- Amazon Simple Queue Service (SQS)
- Amazon Simple Storage Service (S3)
- Amazon Simple Storage Service Glacier (S3 Glacier)
- Amazon Simple Workflow Service (SWF)
- AWS Step Functions
- AWS Storage Gateway
- AWS Well-Architected Tool
- AWS X-Ray

AWS Academy Lab Project - Cloud Data Pipeline Builder

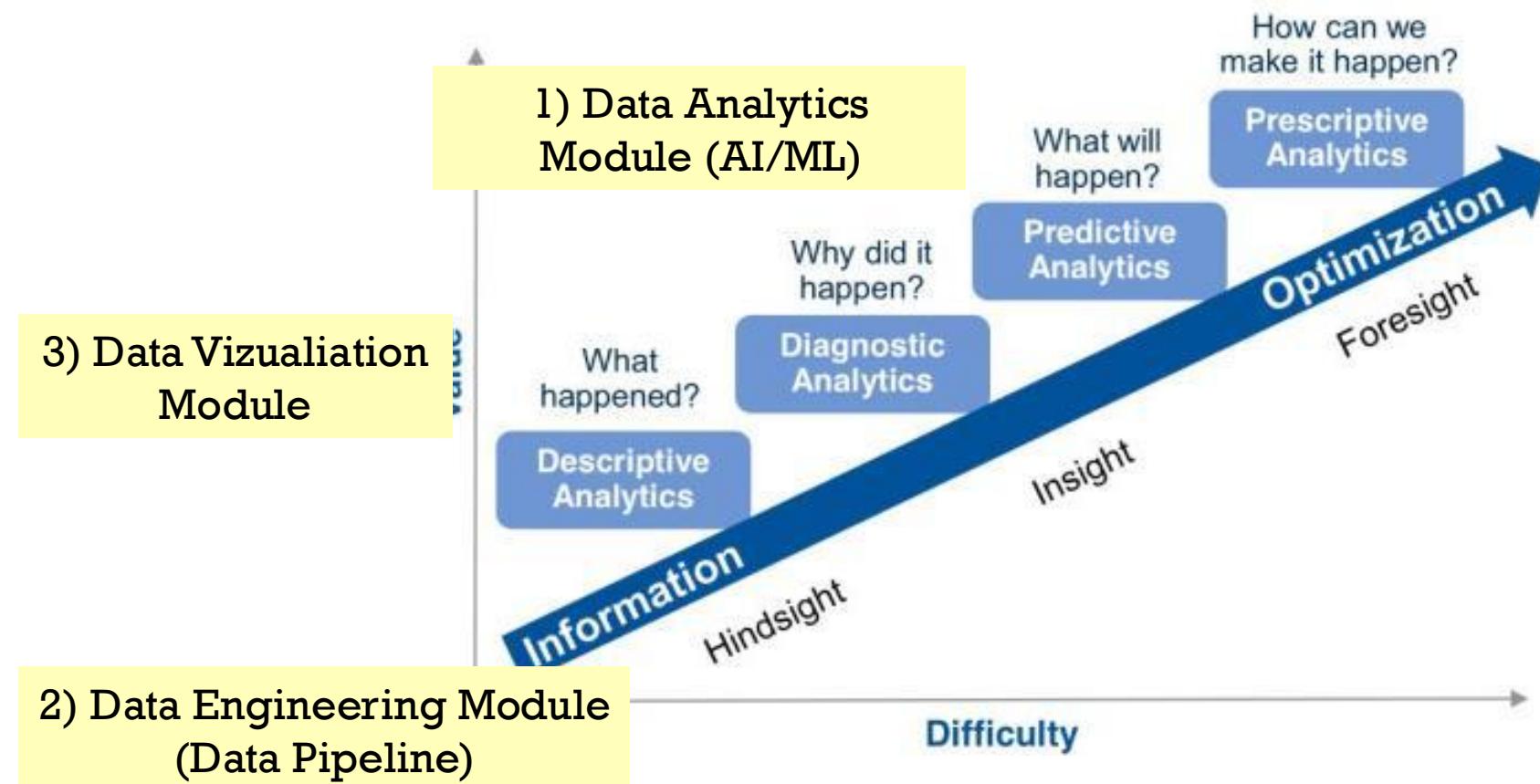
- Amazon Managed Streaming for Apache Kafka (Amazon MSK)

Both Learner Lab & Lab Project - Cloud Data Pipeline Builder

- Amazon SageMaker
- Amazon Elastic Compute Cloud (EC2)
- Amazon DynamoDB
- AWS Lambda
- Amazon Kinesis Video Streams
- Amazon Rekognition

Conclusion

4) Cloud technology





Any questions? ☺