



# Data Science x Artificial Intelligence

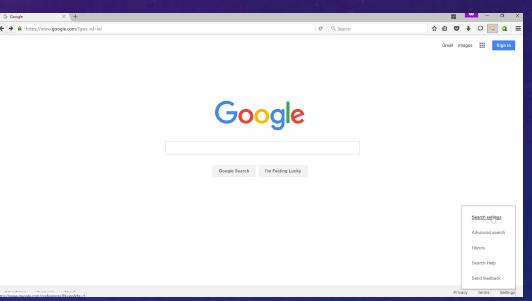


# How my day looks like

7:00  
“Hey Google?  
Turn off my alarm”



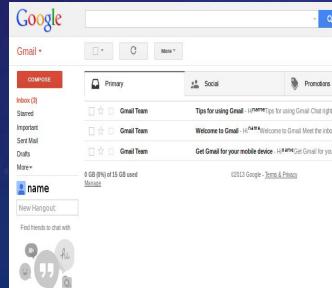
8:15



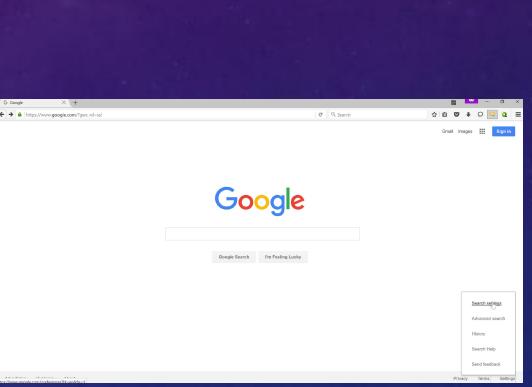
9:30



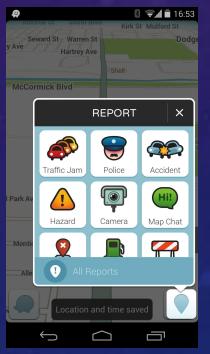
7:30



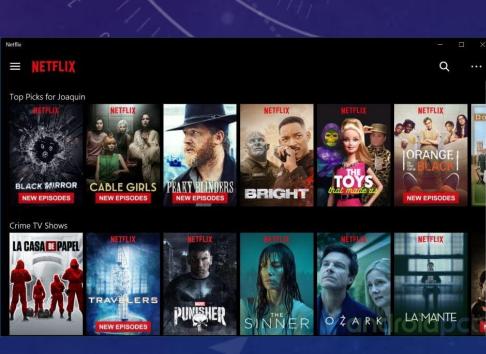
8:15



17:30



20:30



19:30

22:00

# AI is present in almost everything we do

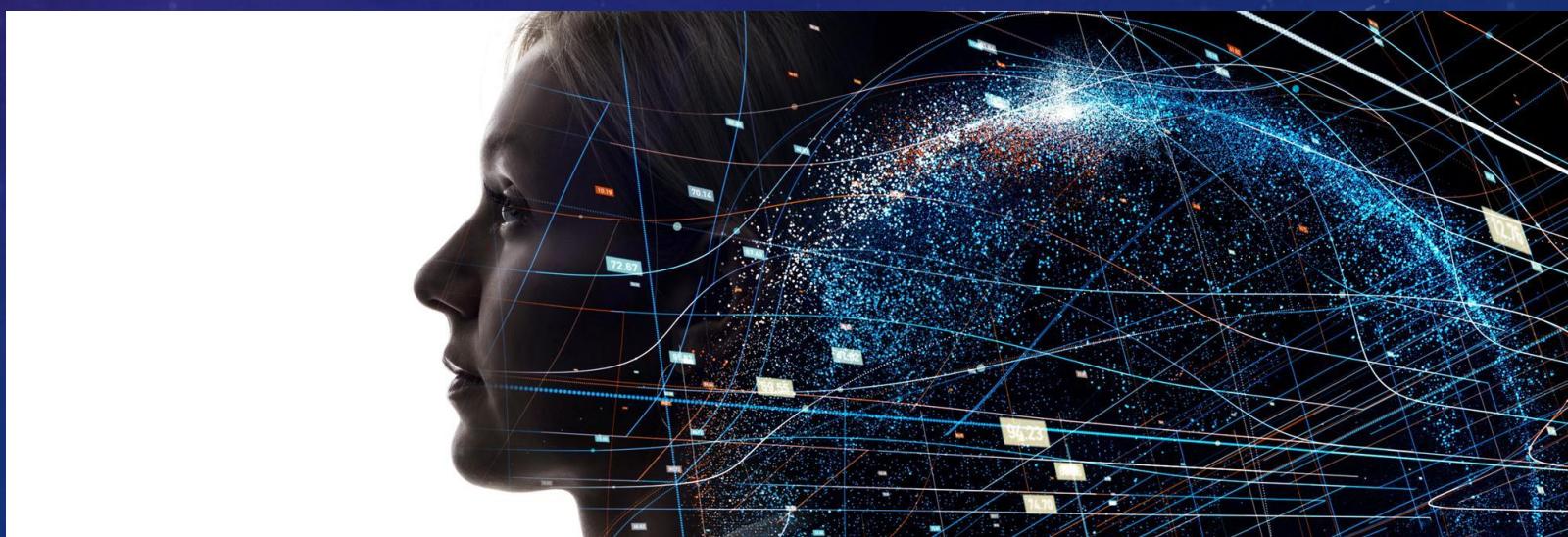
“When algorithms become **deeply embedded** in our daily lives, they have the potential to greatly **influence how we behave.**”

— Mike Walsh,  
Author of *The Algorithmic Leader*

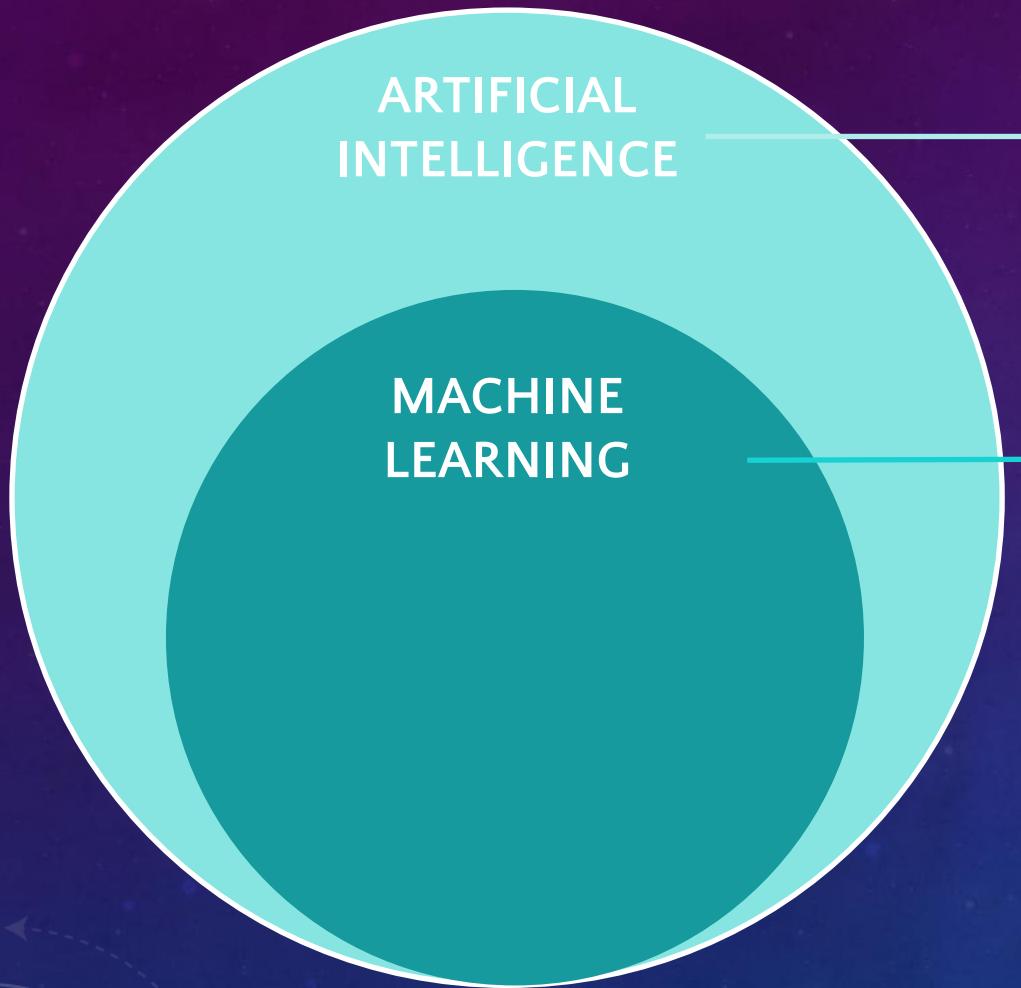


# BY THE END OF THIS SESSION, YOU WOULD BE ABLE TO:

- Differentiate the terminologies: Data Science vs AI vs Machine Learning (ML)
- Learn about the process, tools and skills needed to build and implement data science / machine learning models
- Get an appreciation of how DS & ML are applied in industries, society and in your everyday life



# What is AI? ML? DS?



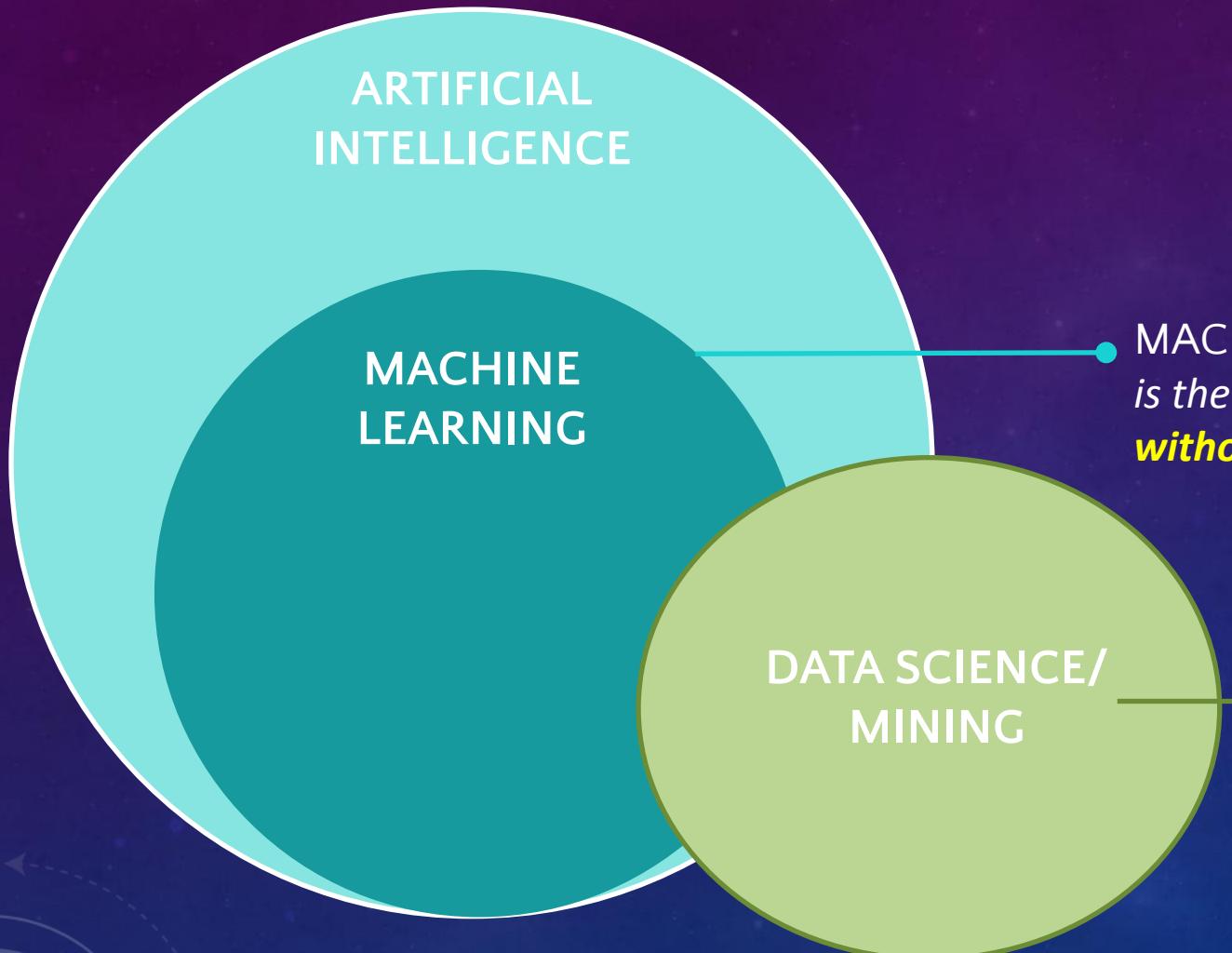
- ARTIFICIAL INTELLIGENCE:  
*Is the general field that enables the capability for systems to emulate / mimic human tasks & behavior.*

- MACHINE LEARNING:  
*is the field of AI that gives computers the ability to learn without being explicitly programmed*  
– Arthur Samuel(1959)

*Examples:*

- Computer game of chess
- Online Shopping Site's "You May Like This"
- Hey Google / Siri

# What is AI? ML? DS?



**MACHINE LEARNING:**  
*is the field of AI that gives computers the ability to learn without being explicitly programmed* – Arthur Samuel(1959)

□ **OUTPUT:** *An algorithm usually deployed in a system/program*

**DATA SCIENCE / DATA MINING:**  
*Is a science of extracting knowledge & insights from data*

□ **OUTPUT:** *A study presented as a slide deck or a paper*

# MACHINE LEARNING VS DATA SCIENCE

## HISTORICAL DATA ON HOME PRICES\*

size of house (square feet)	# of bedrooms	# of bathrooms	newly renovated	price (1000\$)
523	1	2	N	115
645	1	3	N	150
708	2	1	N	210
1034	3	3	Y	280
2290	4	4	N	355
2545	4	5	Y	440



308	3	2	Y	???
-----	---	---	---	-----

 INPUTS



308	3	2	Y	???
-----	---	---	---	-----



OUTPUT

## MACHINE LEARNING EXAMPLE: Estimating/Predicting Prices of Houses

- o Input -> Output
- o Input: Size, # BRs, #T&Bs, New (Y/N)
- o Output: Price of the House
- o Deployed in House Price Calculator Site



House Plan 5219: Wyndsong Farm   
This country style home boasts a wide open plan that captures views through a generous outdoor entertaining area. Master bedroom is cozy and private with a large WIC, dual sinks and garden tub. The family room is vaulted with a see-through fireplace which defines the dining room. [Read more >](#)

House Plan Details							
Sq. Ft.	Beds	Baths	Stories	Galleries	Width	Depth	
2,575	3	2.5	2	3	79' 7"	65' 0"	
	1st Floor 2,044	2nd Floor 531	Bonus 853	Garage 870			
<a href="#">Customize this Home Plan</a>						<a href="#">Estimate the Cost to Build</a>	

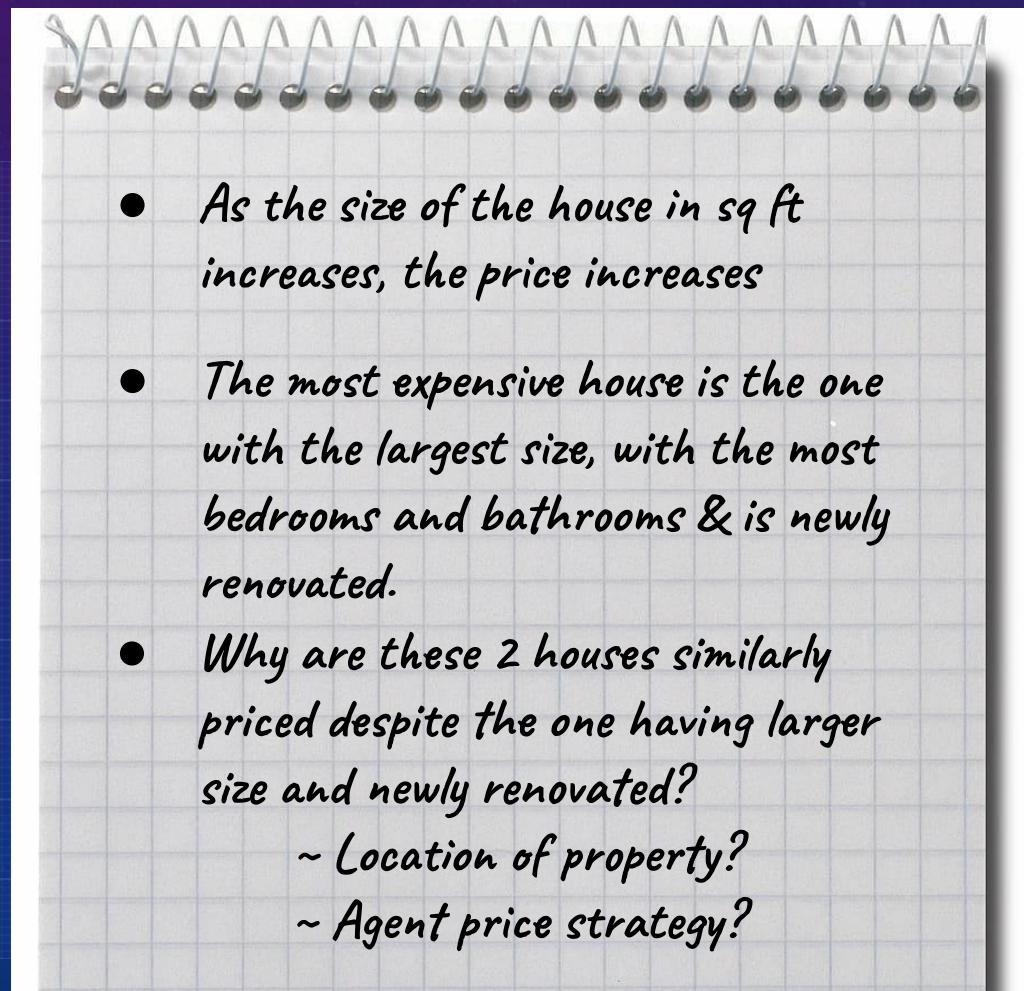
# MACHINE LEARNING VS DATA SCIENCE

HISTORICAL DATA ON HOME PRICES\*

size of house (square feet)	# of bedrooms	# of bathrooms	newly renovated	price (1000\$)
523	1	2	N	115
645	1	3	N	150
708	2	1	N	210
1034	3	3	Y	280
2290	4	4	N	355
2545	4	5	Y	440

## DATA SCIENCE / DATA MINING

- Insights are drawn from the data

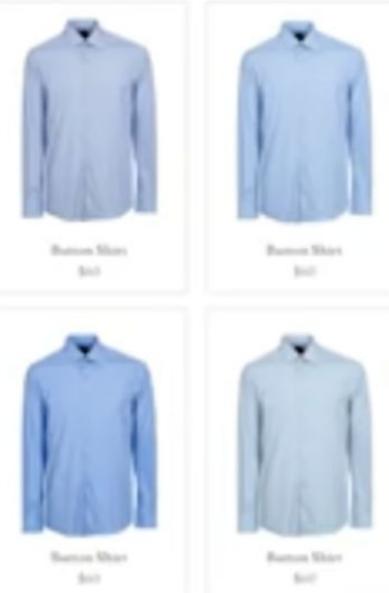
- 
- As the size of the house in sq ft increases, the price increases
  - The most expensive house is the one with the largest size, with the most bedrooms and bathrooms & is newly renovated.
  - Why are these 2 houses similarly priced despite the one having larger size and newly renovated?
    - ~ Location of property?
    - ~ Agent price strategy?

# MACHINE LEARNING OR DATA SCIENCE

## APPLICATION: MARKETING

???

Recommended for you



???

Break into AI

Whether you want to build algorithms or build a company, deeplearning.ai's courses will teach you key concepts and applications of AI.

[Learn More > Deep Learning Specialization](#)

Break into AI

Whether you want to build algorithms or build a company, deeplearning.ai's courses will teach you key concepts and applications of AI.

[Take My Core Learning Specialization](#)

A

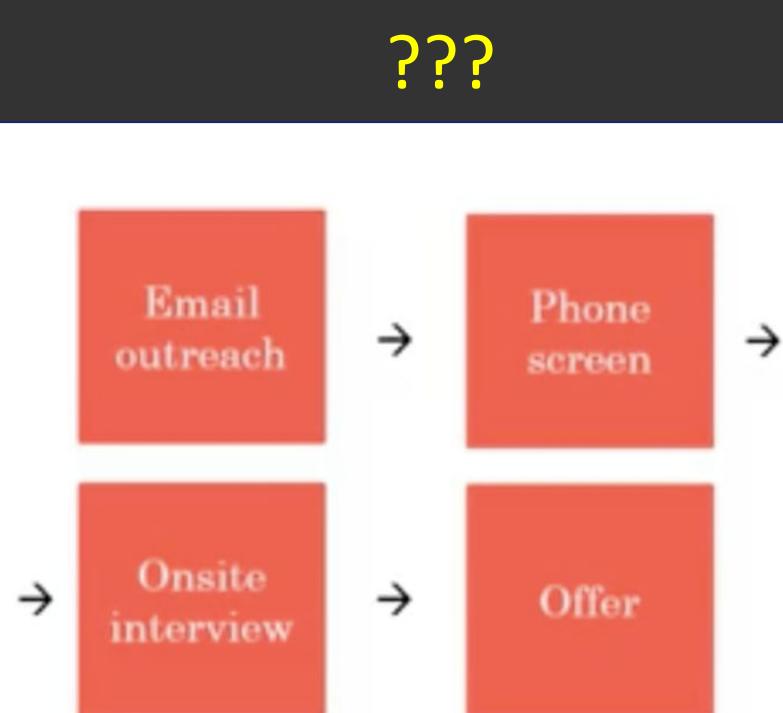
B

Customized product recommendation

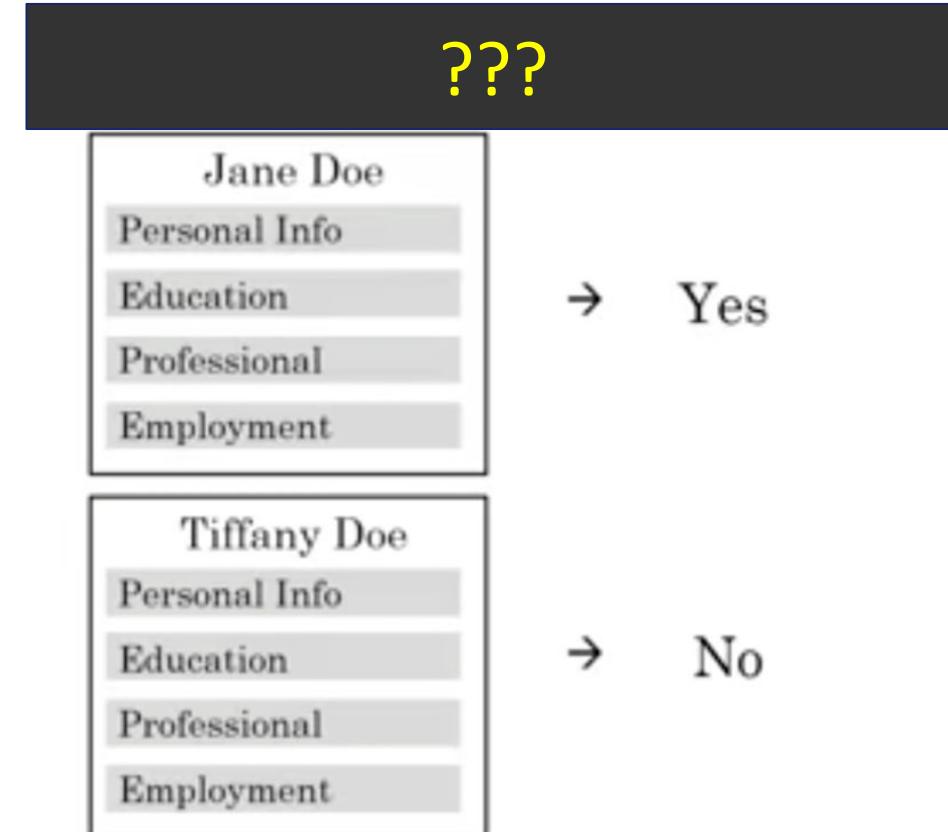
A/B testing

# MACHINE LEARNING OR DATA SCIENCE

## Recruiting



Optimize recruiting funnel



Automated resume screening

# MACHINE LEARNING VS DATA SCIENCE

Agriculture

???



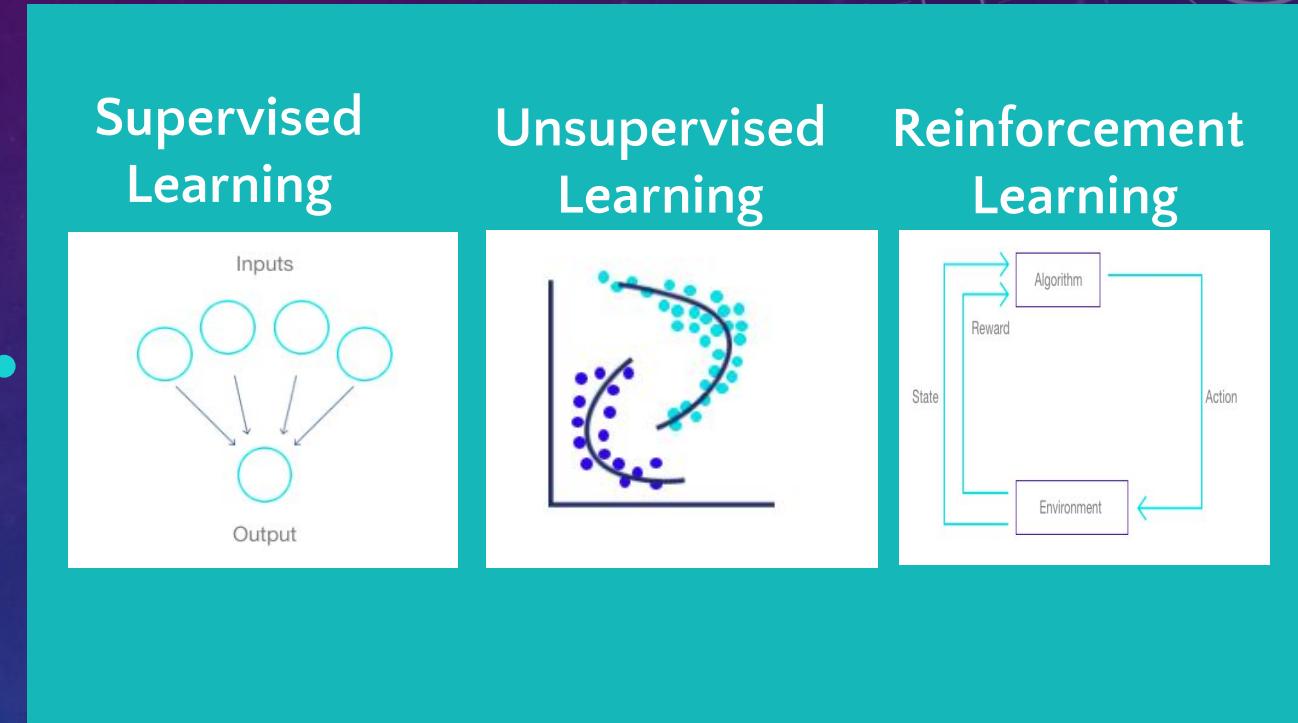
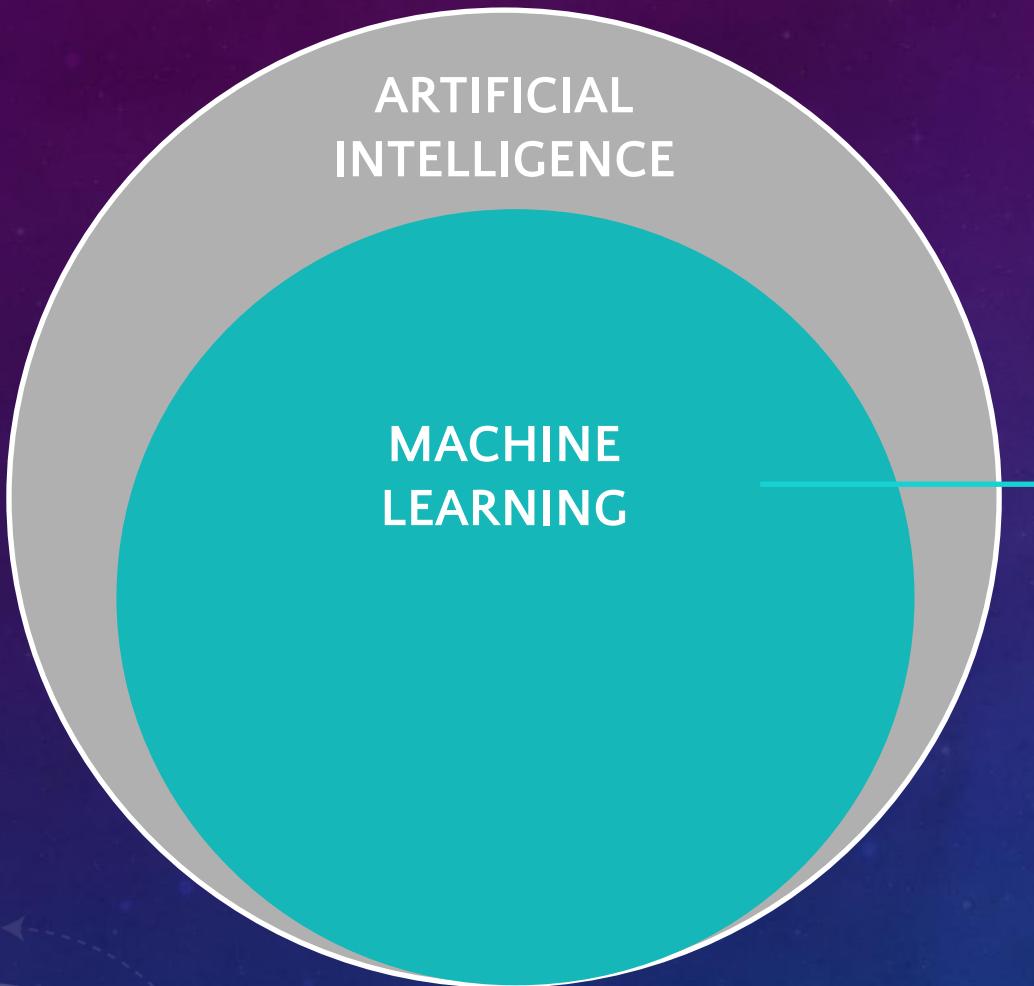
Crop analytics

???



Precision weed killing

# TYPES OF MACHINE LEARNING



# SUPERVISED LEARNING

A teacher shows the girl what a cat looks like..



INPUT

*A machine learning algorithm that uses labeled data to learn the relationship of given inputs to a given output\**



So when the girl sees this furry animal like this one, she now knows that is a cat.

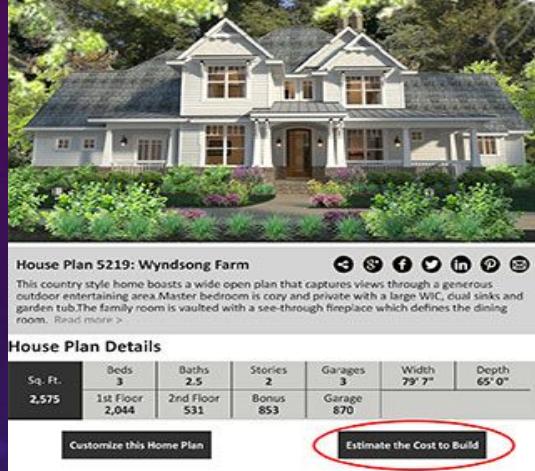


OUTPUT

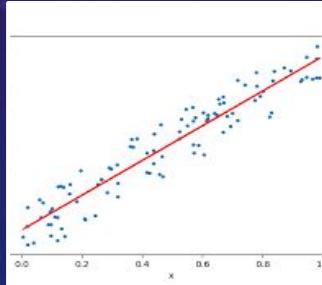
# SUPERVISED LEARNING EXAMPLES

## MODEL TECHNIQUES

Estimate price of a house



### Linear Regression

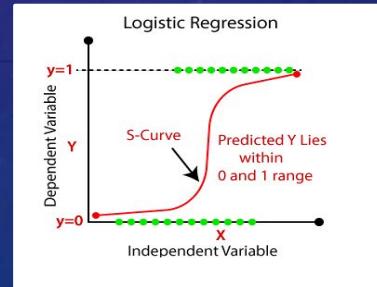


Models relationship between input variables and dependent output (continuous) variables  
OUTPUT =  $a+b (\text{INPUT}_1)..+n(\text{INPUT}_n)$

Predict probability of being a bad credit card user



### Logistic Regression

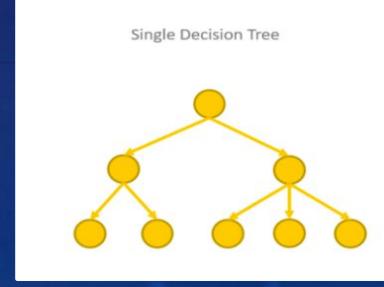


Also a regression model but target is binary & output is probability bet 0-1  
 $\text{Prob} (0,1) = a+b (\text{INPUT}_1)..+n(\text{INPUT}_n)$

Predict food / product preference



### Decision Trees



Model that splits data feature values into branches at decision until a final decision output is made

# UNSUPERVISED LEARNING

*is used if you do not know how to classify the data, and you want the algorithm to find patterns and classify the data for you*



Like how Hogwarts' Sorting Hat works

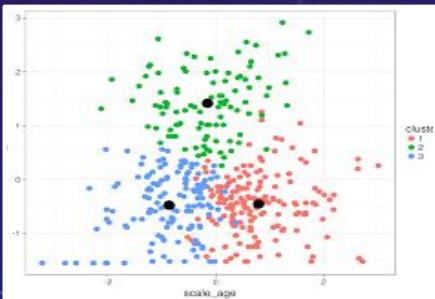


# UNSUPERVISED LEARNING EXAMPLES

Personalized Telco offers based on subscriber's usage and browsing behaviors

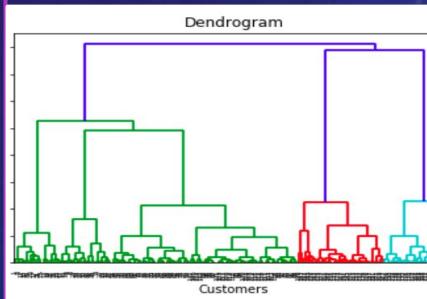


## K-Means Clustering



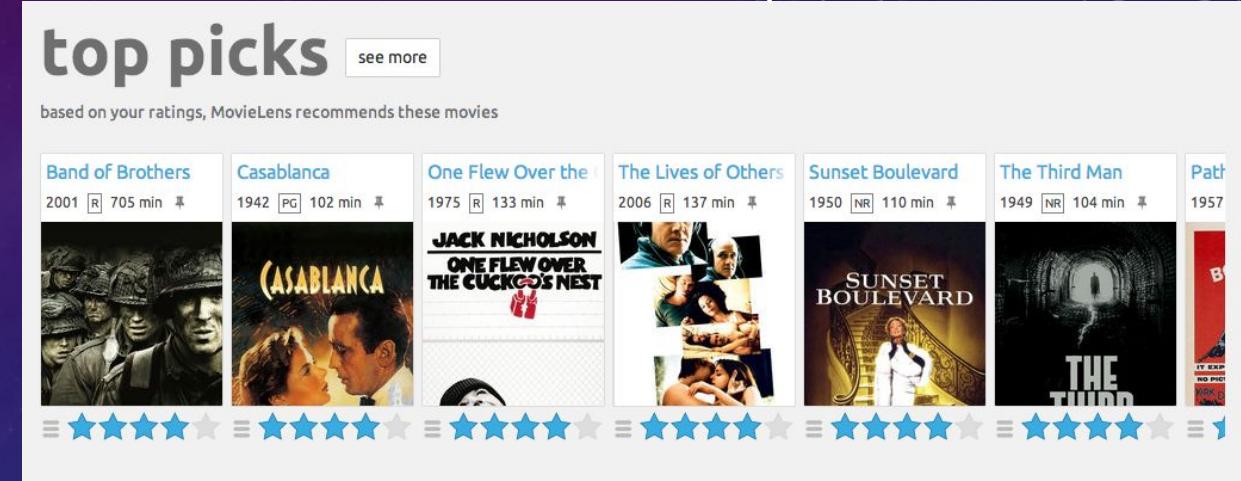
Puts data into a number of groups ( $k$ ) that each contain data with similar characteristics (as determined by the model)

## Hierarchical Clustering

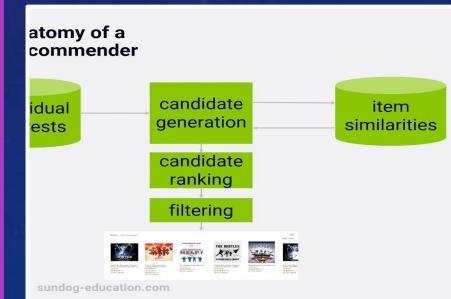


Splits or aggregates clusters along a hierarchical tree to form a classification system

Movie recommendation based on ratings of viewers with similar profiles



## Recommender Systems

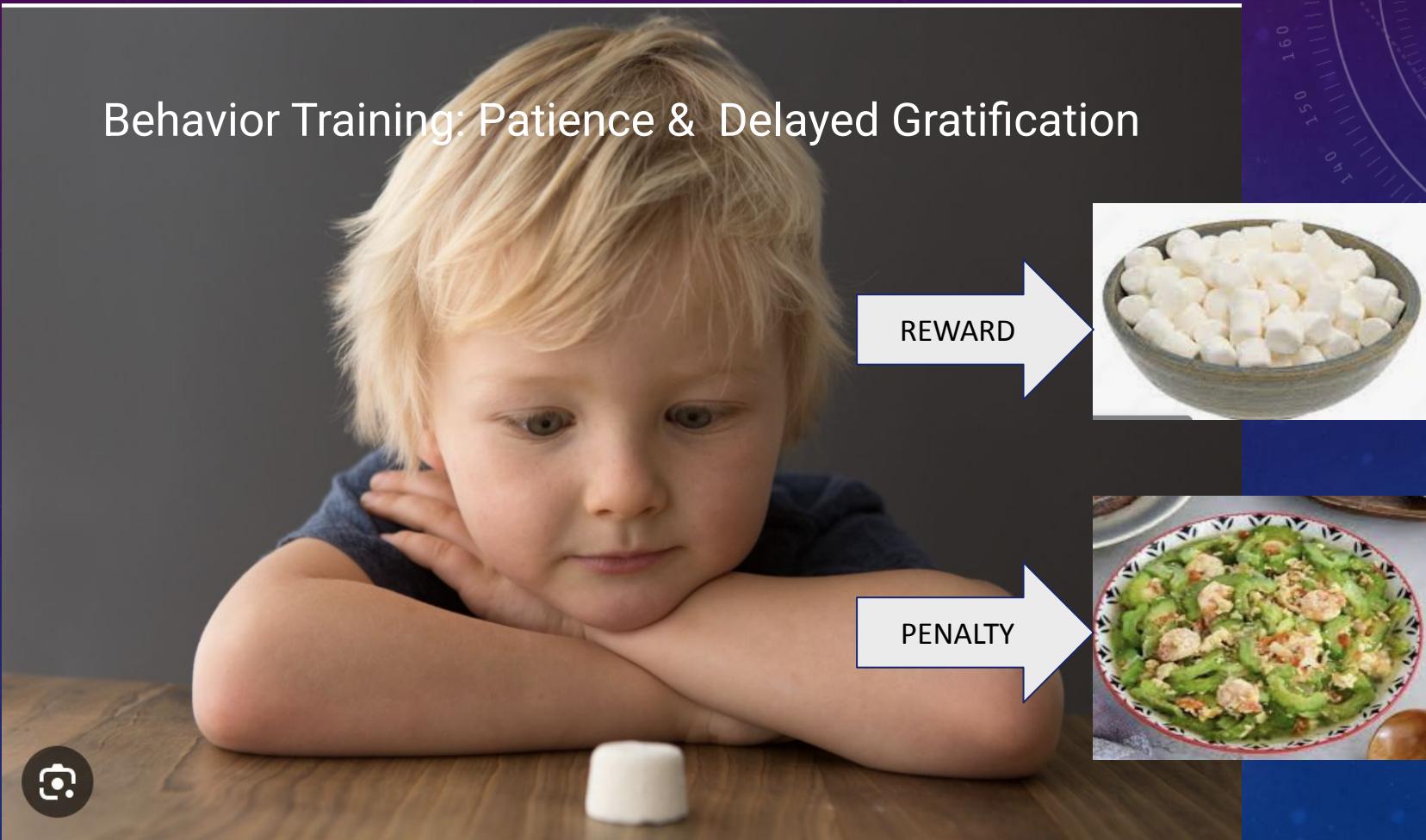


A system that suggests products, services, information to users based on analysis of data.

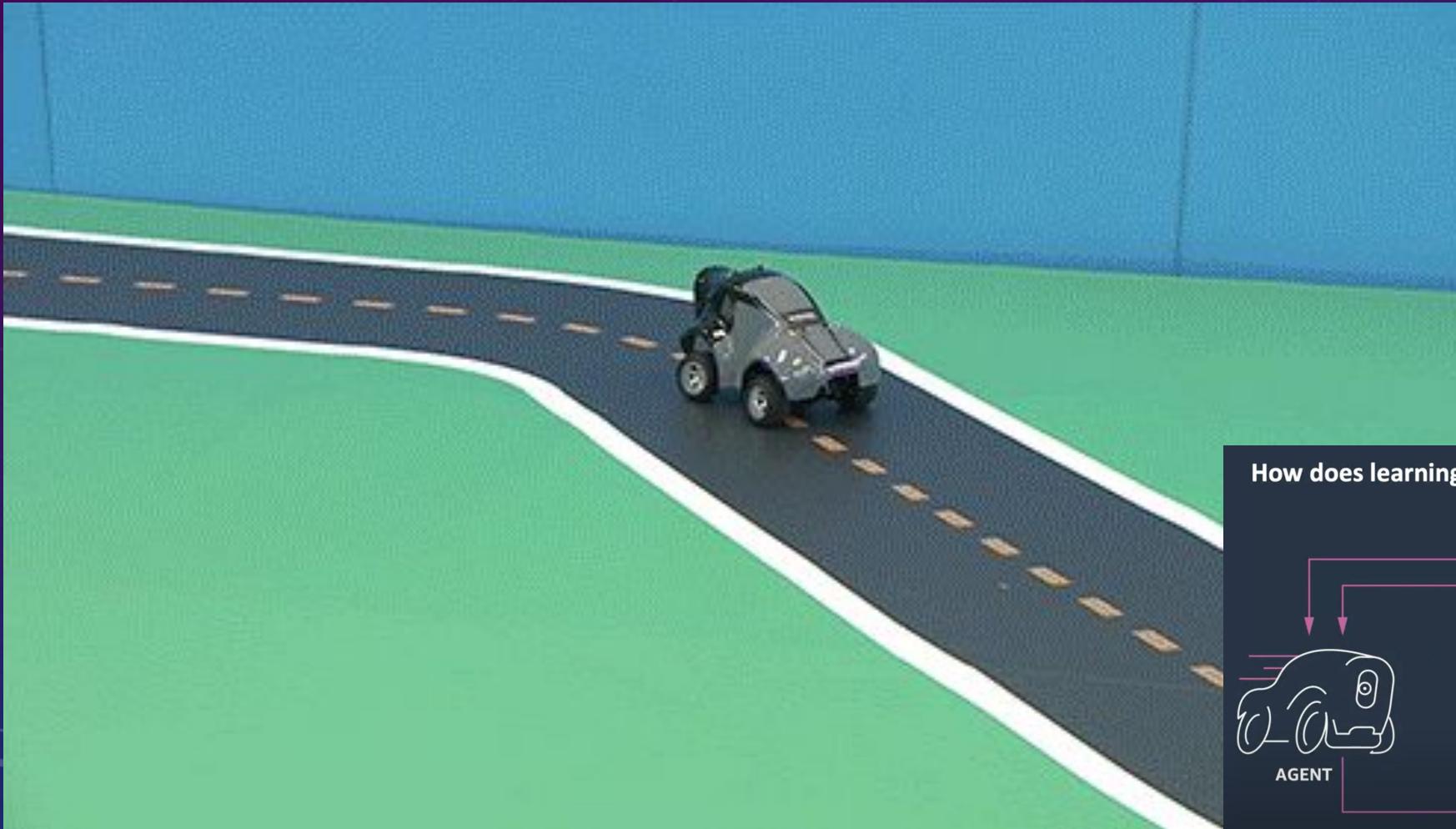
# REINFORCEMENT LEARNING

An algorithm learns to perform a task simply by trying to **maximize rewards** it receives for its actions

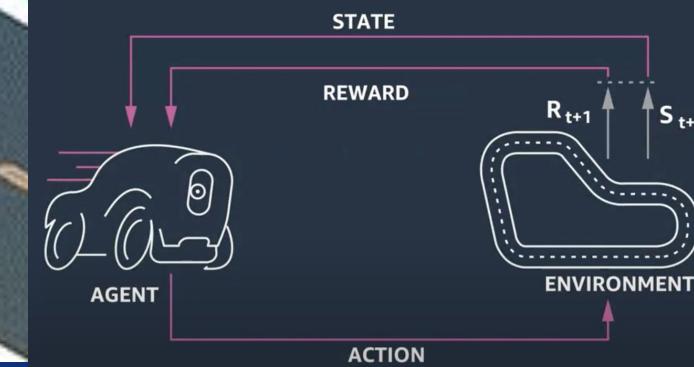
Behavior Training: Patience & Delayed Gratification



# REINFORCEMENT LEARNING: SELF DRIVING CARS



How does learning happen?



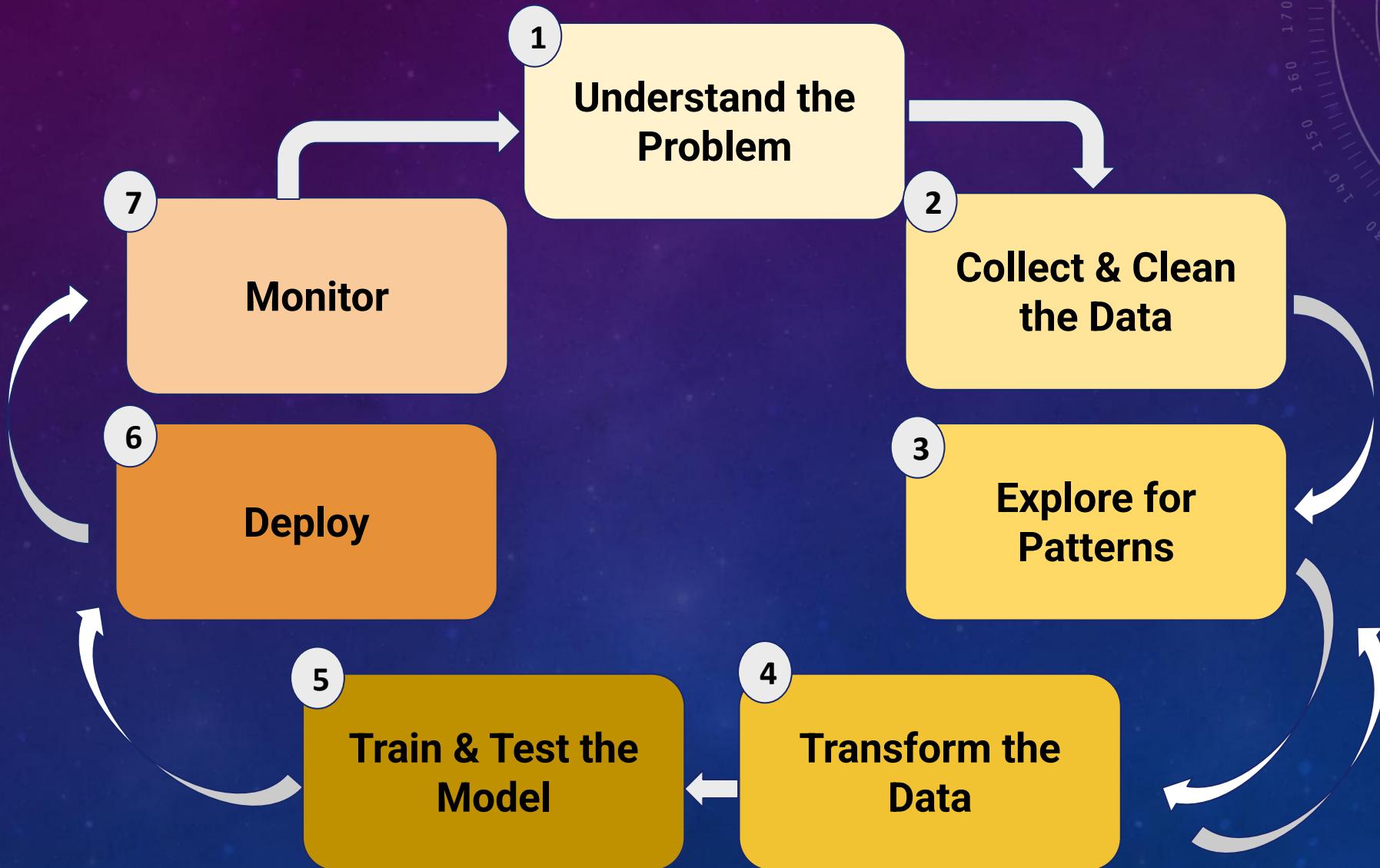
VALUE FUNCTION



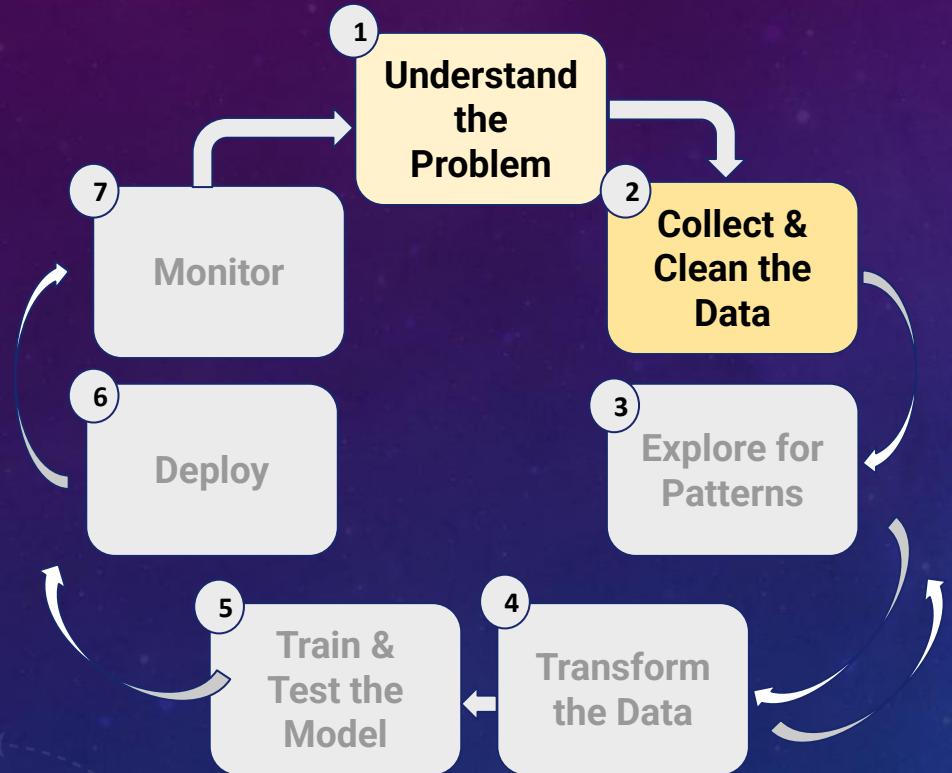
POLICY FUNCTION



# MACHINE LEARNING MODEL LIFECYCLE



# MACHINE LEARNING MODEL LIFECYCLE

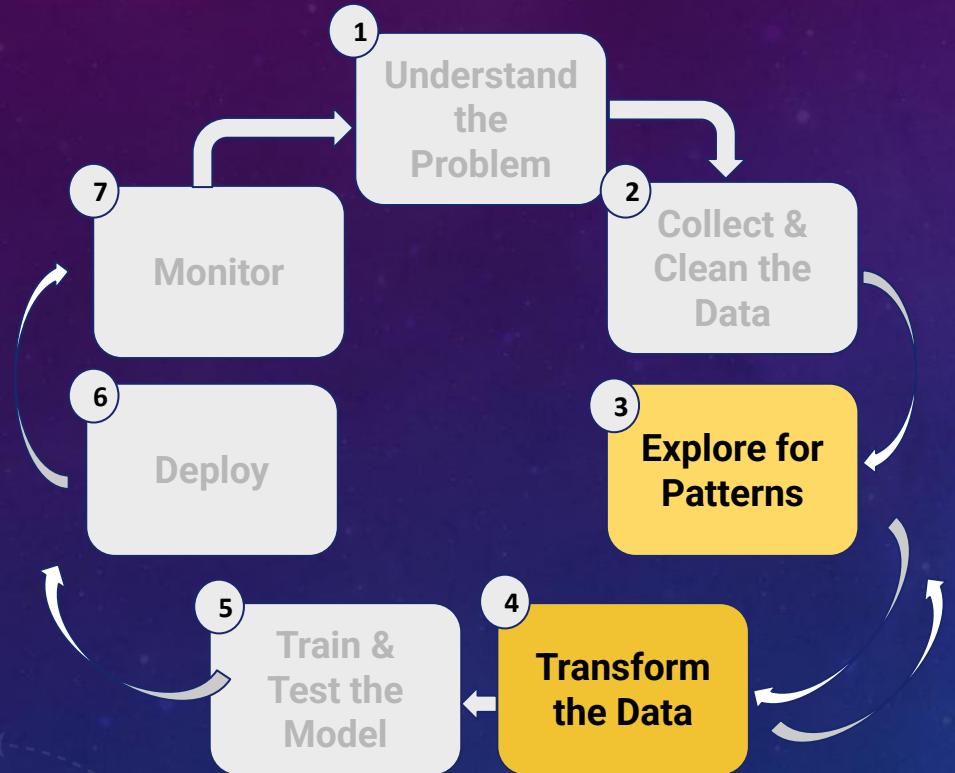


**PROBLEM:**  
*Predict the demand for Ride-hailing cars  
to optimize service for commuters*

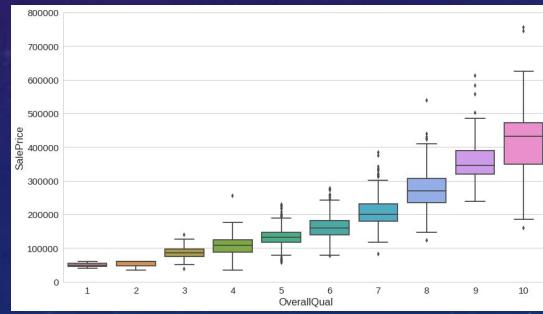
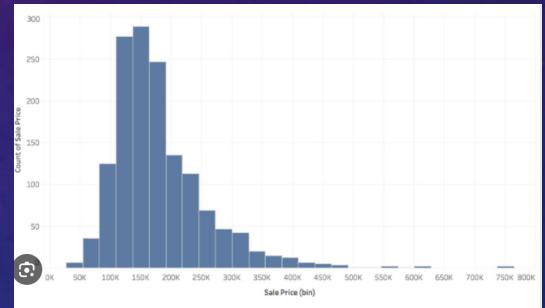
## DATA COLLECTION:

- *Historical trip data*
- *Location*
- *Time of Day*
- *Weather Conditions*
- *Events*
- ....

# MACHINE LEARNING MODEL LIFECYCLE



## DATA EXPLORATION & VISUALIZATION

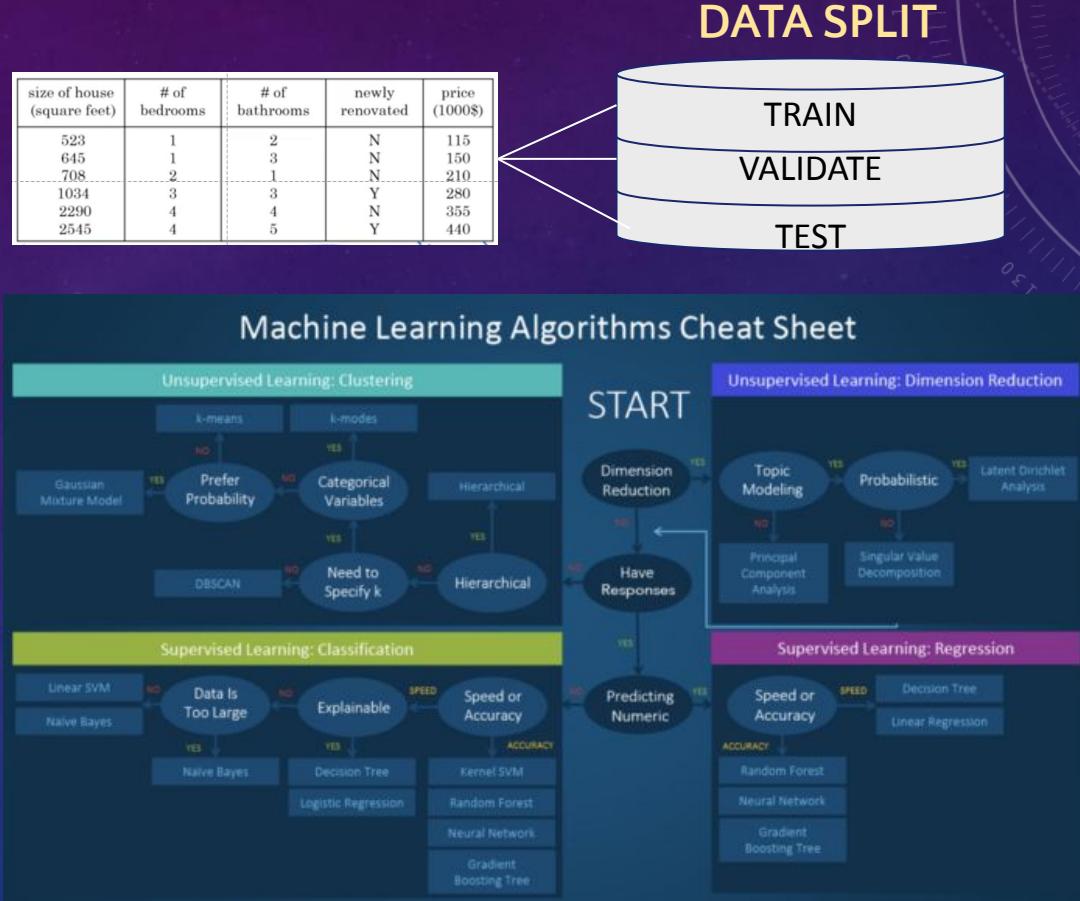
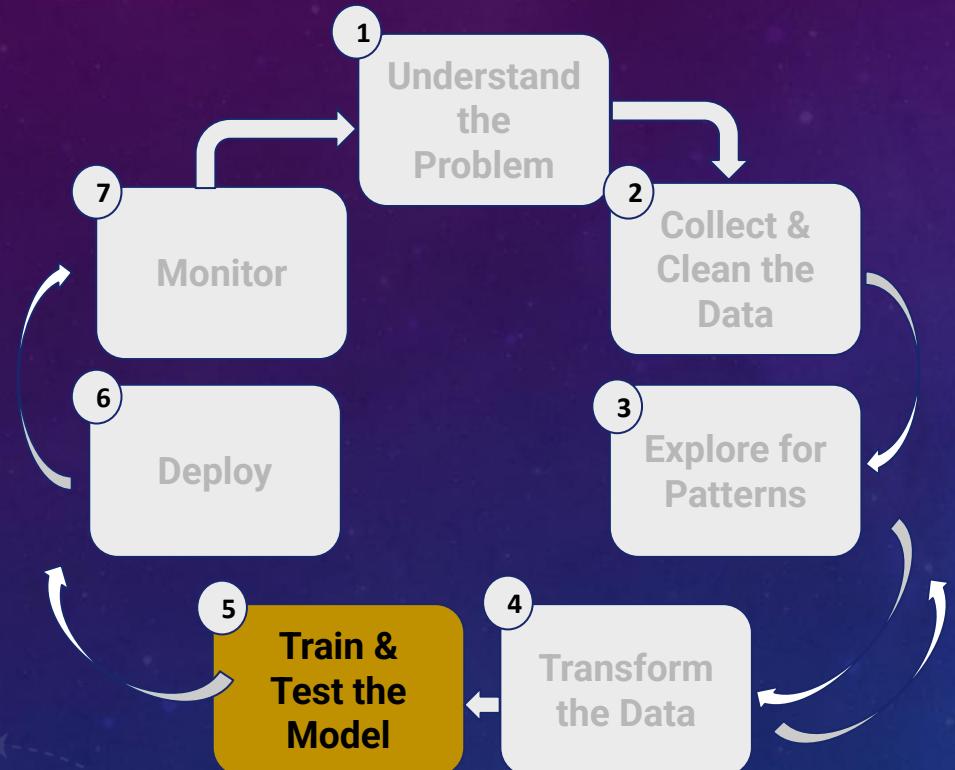


## DATA TRANSFORMATION aka DATA WRANGLING aka FEATURE ENGINEERING

- Riders Ratio = # of Riders / # Population
- Trip Classification = Accepted, Cancelled..



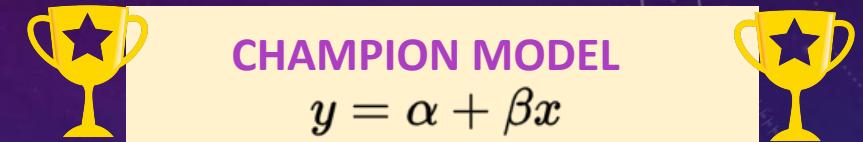
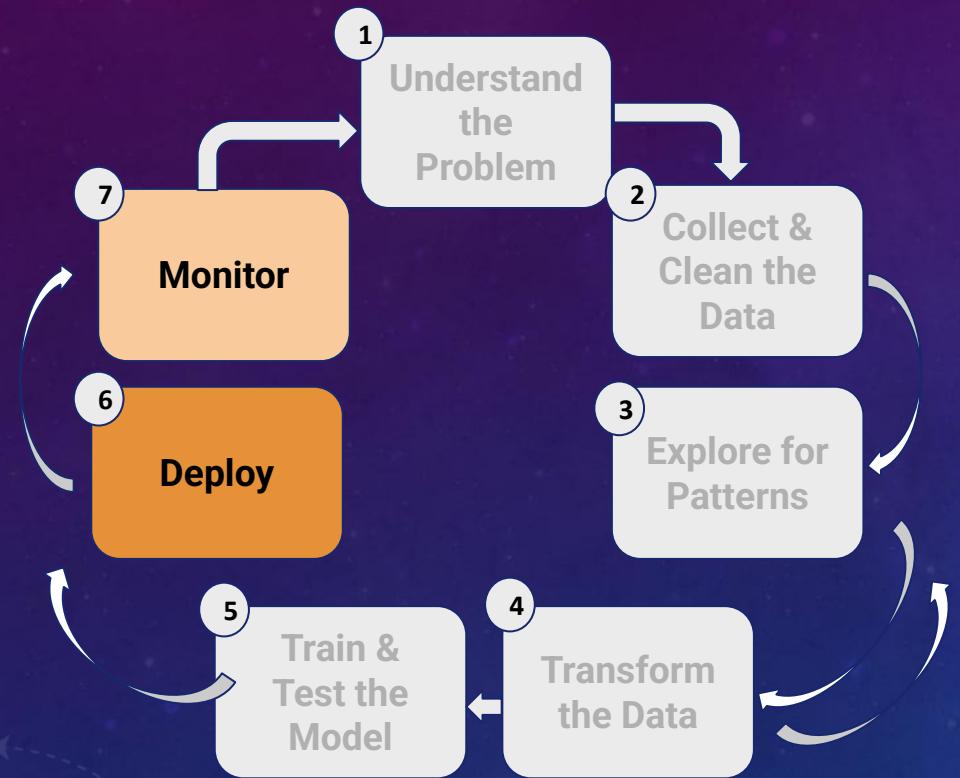
# MACHINE LEARNING MODEL LIFECYCLE



**CHAMPION MODEL**  
 $y = \alpha + \beta x$



# MACHINE LEARNING MODEL LIFECYCLE



OPTIMIZER ALGO EMBEDDED IN APP

# DS & ML Applications

- Predictive Traffic Congestion & Vehicle Flow
- P2P Route Optimization
- Road Safety Risk Classification
- Parking Space Demand Allocation
- Mobility Insight Monetization



# SKILLS & TOOLS



In more than two-thirds of our use cases, artificial intelligence (AI) can improve performance beyond that provided by other analytics techniques.

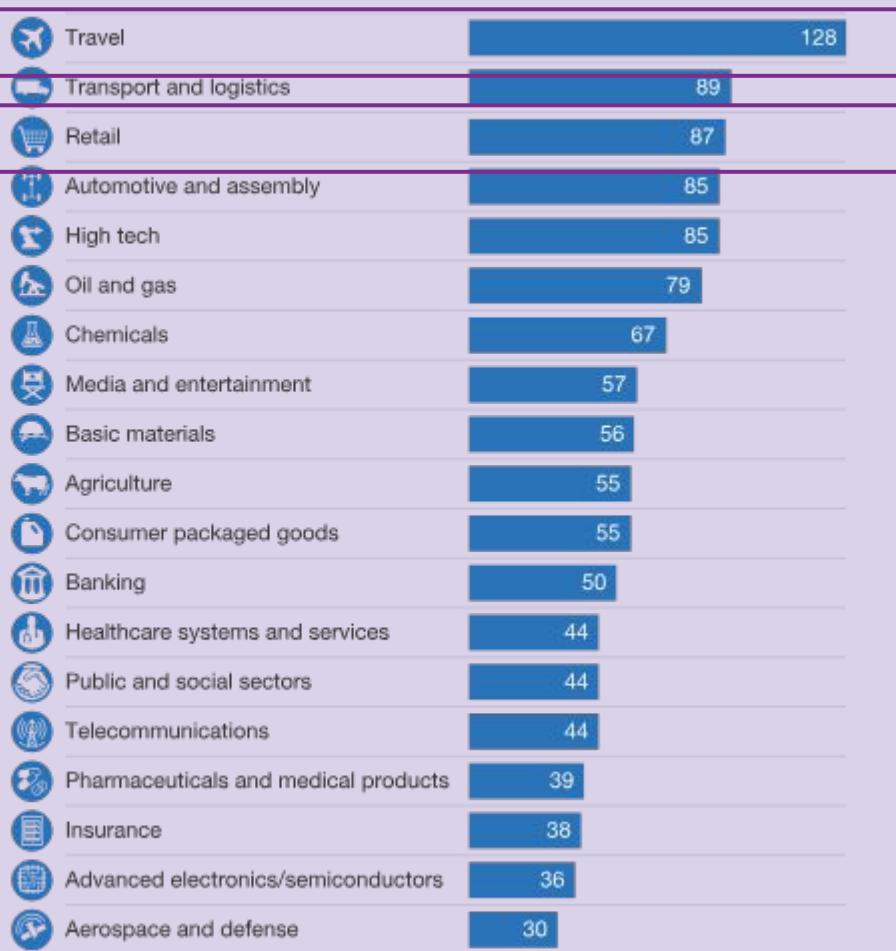
Breakdown of use cases by applicable techniques, %

Full value can be captured using non-AI techniques

AI necessary to capture value ("greenfield")

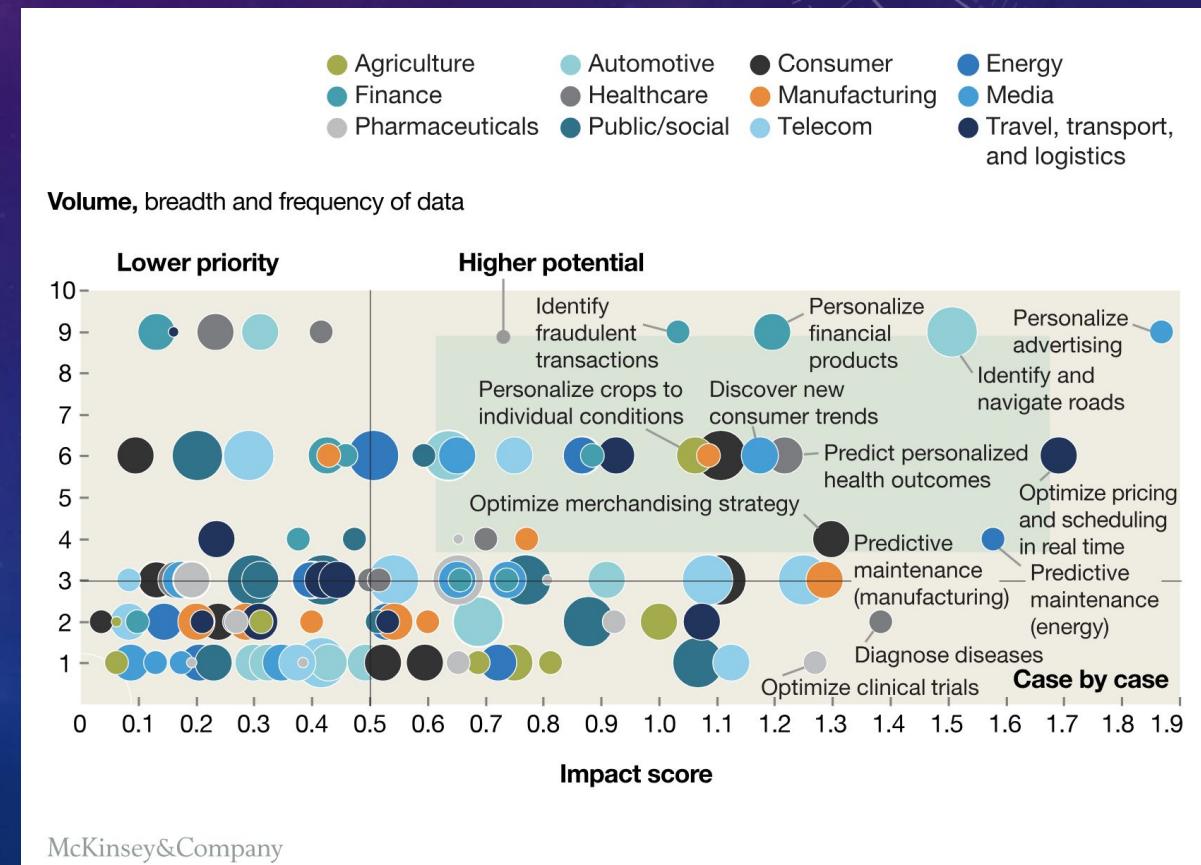
AI can improve performance over that provided by other analytics techniques

Potential incremental value from AI over other analytics techniques, %



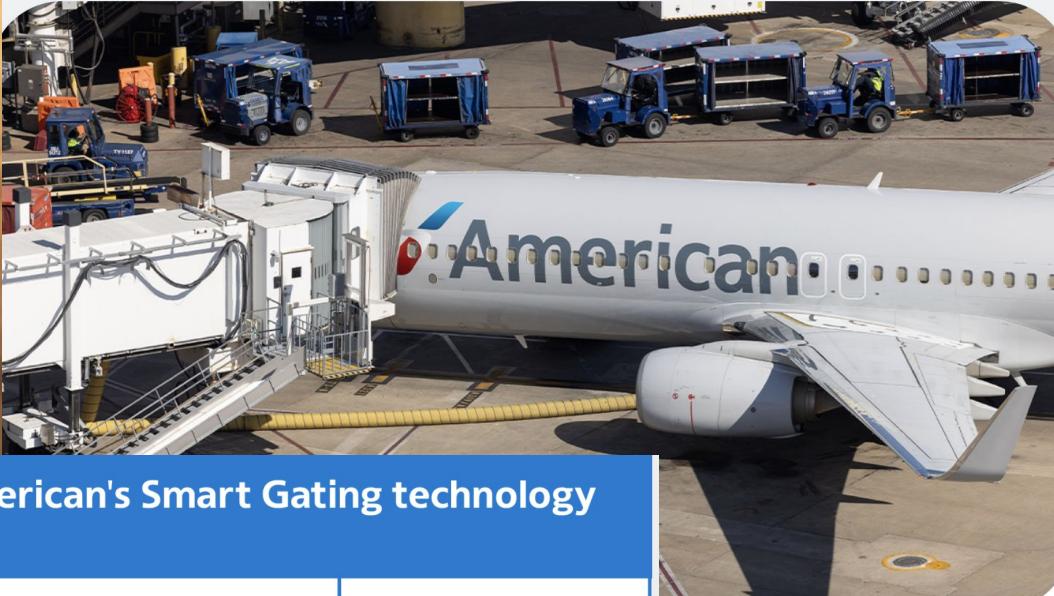
McKinsey&Company | Source: McKinsey Global Institute analysis

# DS & ML have broad potential across industries & use cases



McKinsey&Company

# Application in Travel



## American's Smart Gating technology



Shortens taxi times by 20%, or two minutes per flight, at DFW.



Shortens total taxi time across the system by 17 hours per day.



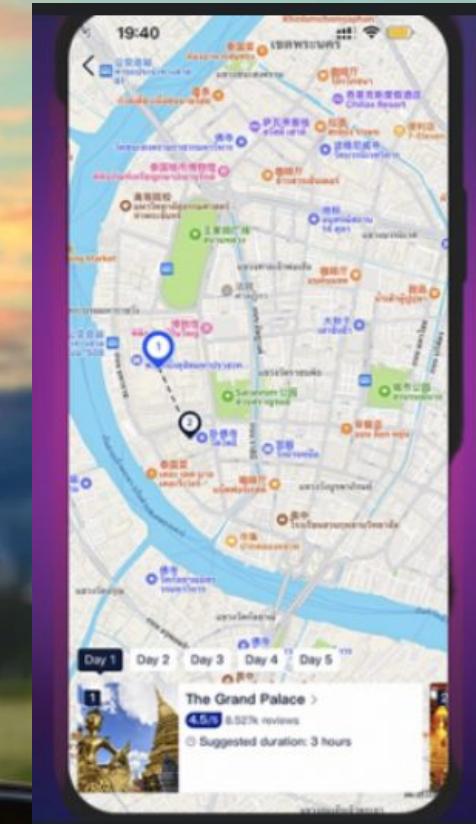
Saves an estimated 1.4 million gallons of jet fuel each year.



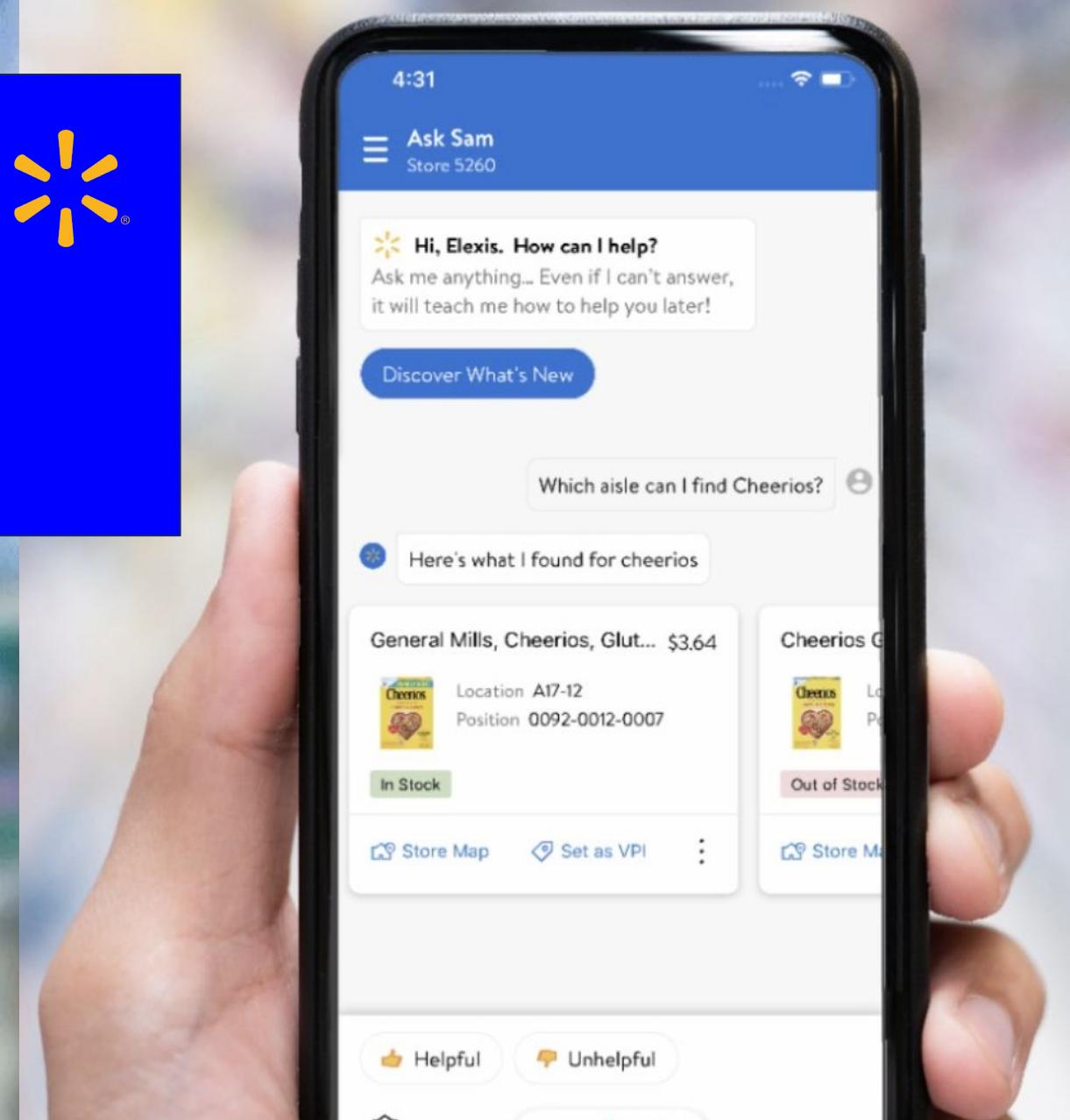
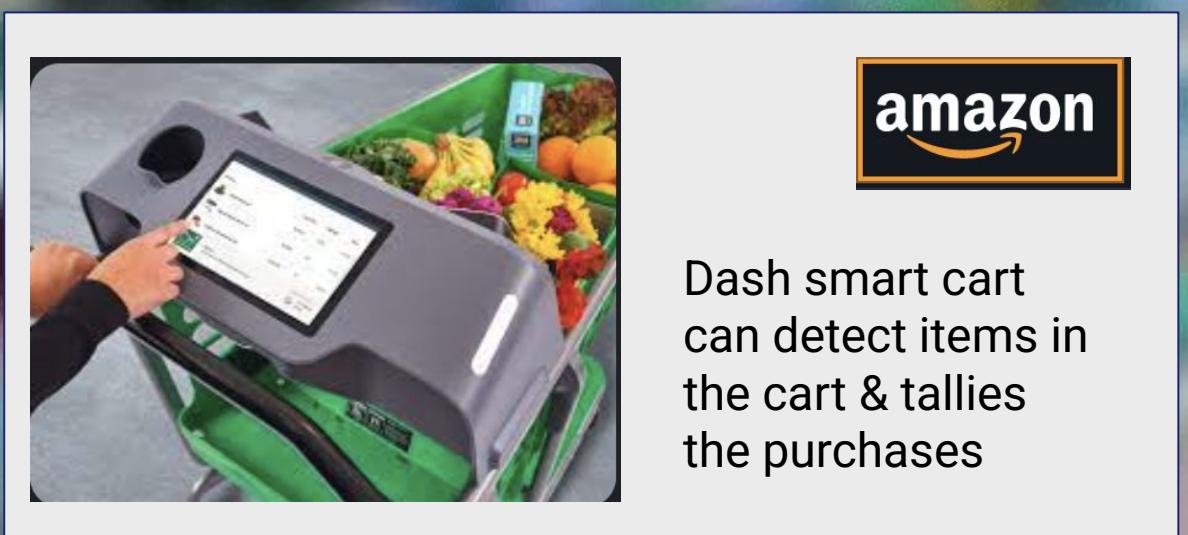
Reduces carbon dioxide emissions by more than 13,000 metric tons annually.



AI Travel assistant with multi-lingual voice & text input.



# Application in Retail industry

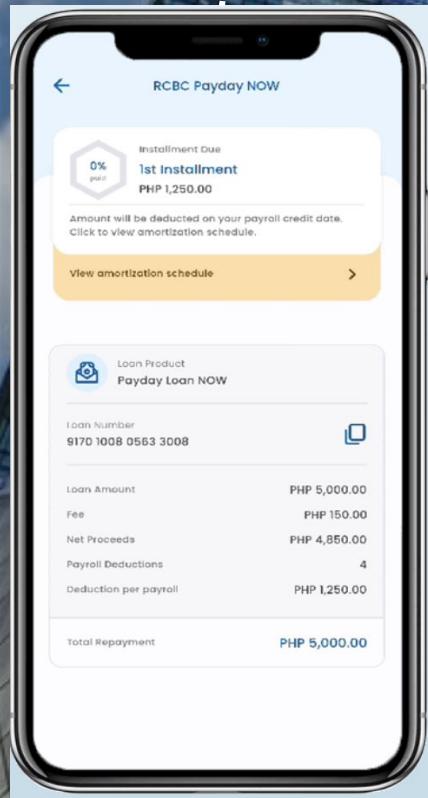


# RCBC uses data science & over 50 ML models as strategic tools



## CREDIT SCORING MODELS

Assess if someone can pay a loan



## PERSONALIZED EMAIL OFFERS

Based on segmented profiles

The image contains two separate promotional banners. The top banner is for the "RCBC Credit Cards Hexagon Club" and features three women smiling and toasting with wine glasses. A central callout box says "Claim your welcome gift of up to P10,000 worth of Rewards today!". The bottom banner is for the "RCBC Hexagon Club" and promotes a "BUY 1 TAKE 1" offer for a birthday treat at The Coffee Bean &amp; Tea Leaf, featuring a coffee drink and two credit cards.

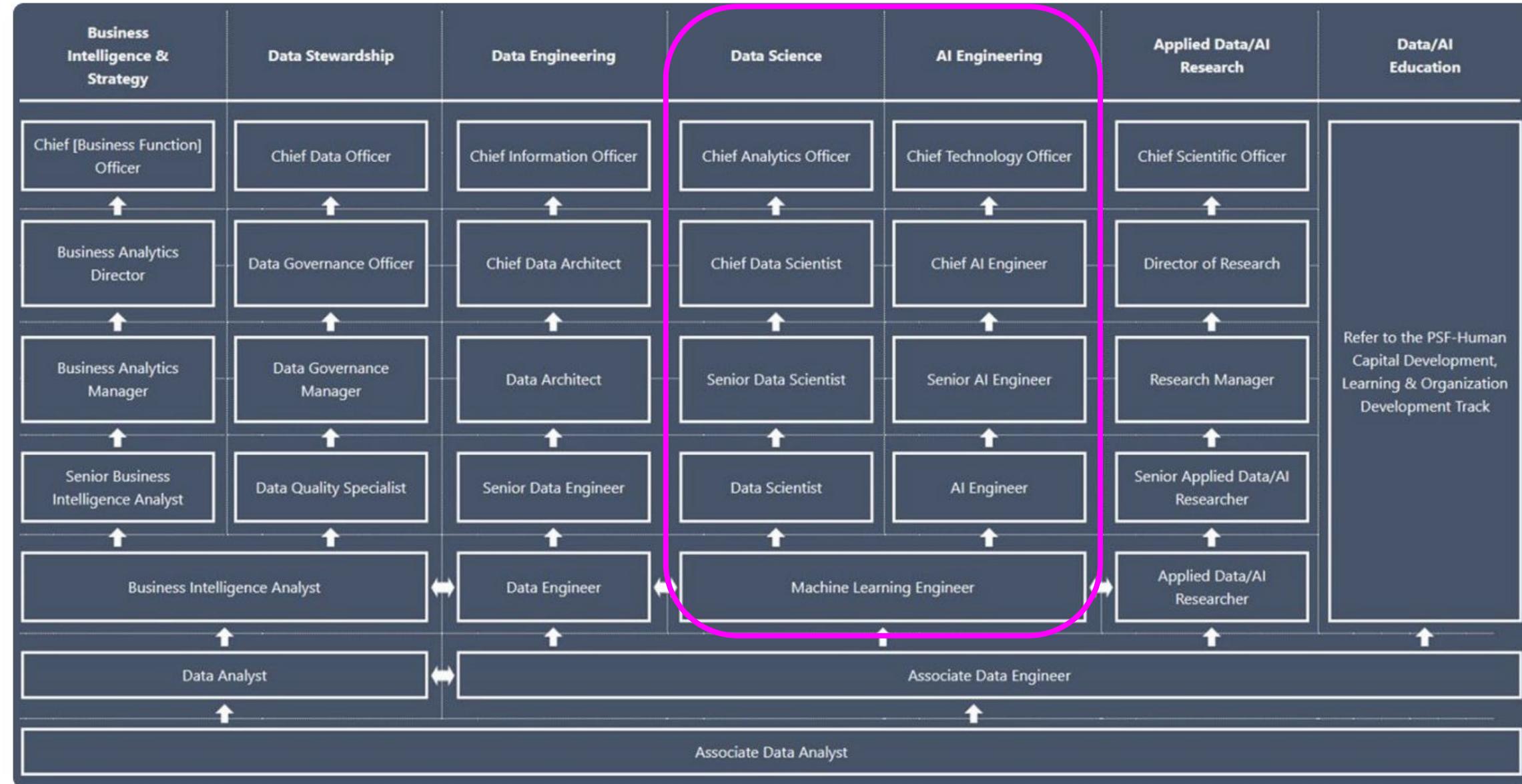
## DETECT FRAUDULENT PATTERNS

By analyzing patterns

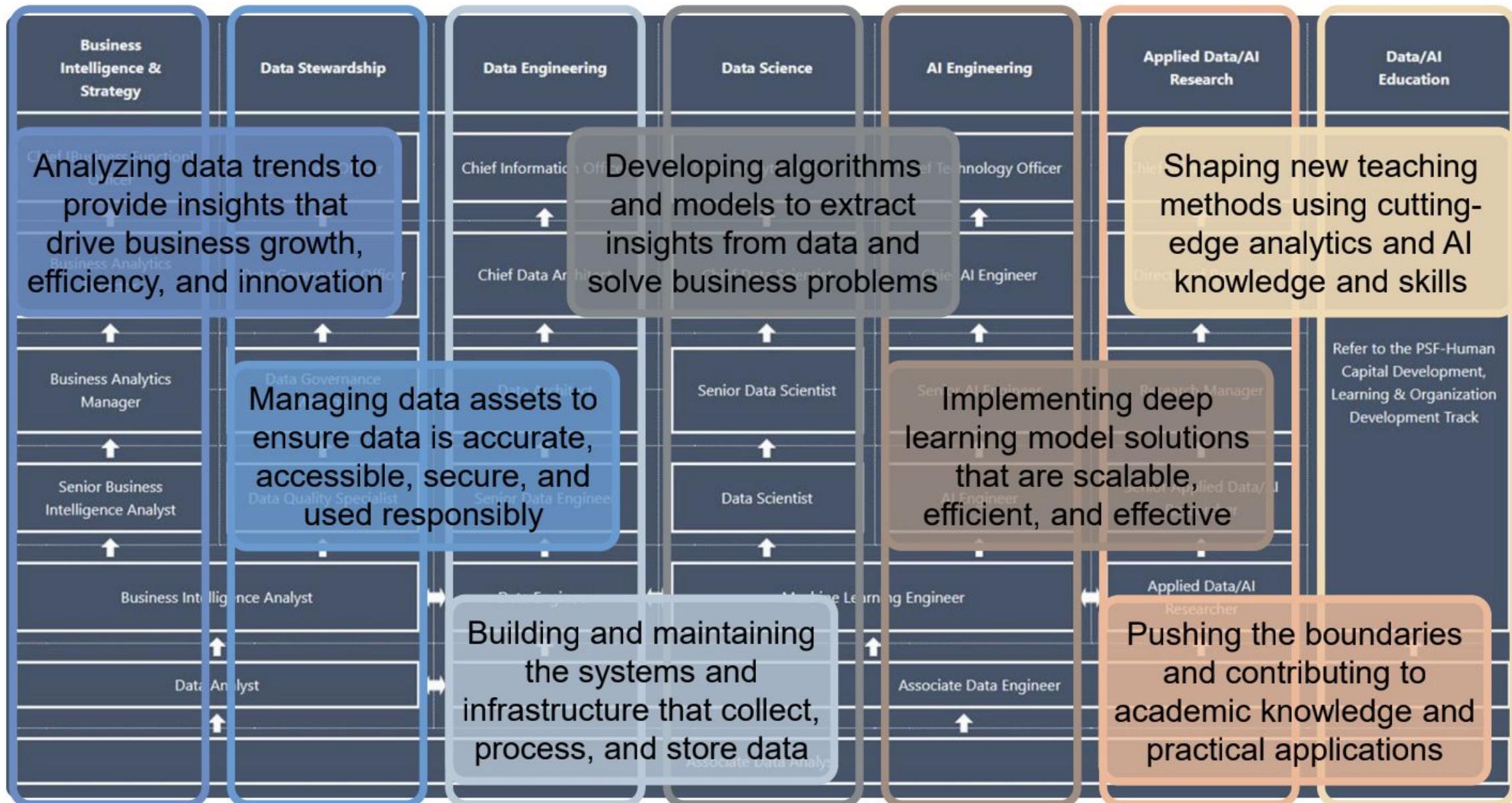




# THE PH SKILLS FRAMEWORK FOR ANALYTICS & AI



# CAREER MAP



## Data Scientist

### Job Description

The Data Scientist plays a crucial role in the development of advanced data analytics techniques and solutions, encompassing the entire process from design and prototyping to testing. The role involves extracting and integrating data from various sources, exploring complex datasets to unearth incremental business value, and creating advanced statistical models tailored for specific business use cases. An integral part of the role is conducting thorough testing of these models, interpreting the findings, and evaluating model performance for scalability and deployment. The Data Scientist is also responsible for developing clear and compelling communication materials to facilitate stakeholder understanding and buy-in. This role demands strong analytical skills to identify and solve complex business problems, proficiency in statistics, scripting, and programming languages relevant to the organisation, and familiarity with software platforms for deploying solutions. The Data Scientist is passionate about analyzing data, displays intellectual curiosity, possesses strong critical thinking skills, and has the ability to narrate and present data findings effectively to influence stakeholders and promote a data-driven approach to resolving business issues.

### Critical Work Functions & Key Tasks

#### *Present data driven business value of data science models*

- Contribute to the creation of leading-edge resources, including playbooks, guides, blog posts, videos, etc.
- Develop compelling, logically structured presentations including story-telling of research and/or analytics findings to secure stakeholder commitment
- Create reports and deliverables based on insights derived from the model results

#### *Build and assess models*

- Document modelling techniques used and assumptions made against test outcomes
- Conduct testing on final model in real-time business conditions prior to deployment
- Enable end user capability to use data science products effectively
- Initiate autonomous monitoring to scale human oversight
- Scale and deploy models in real-time business conditions for end user consumption

#### *Manage data preparation and modelling*

- Interpret and evaluate model performance for scaling and deployment
- Select the best model based on pre-defined evaluation criteria
- Perform model comparison to draw inferences on variable importance
- Develop multiple models and algorithms suitable for the use case
- Analyse the ways in which datasets may be biased and address this in safety measures and deployment strategies

### Functional Skills and Competencies

Skill	Proficiency
Agile Software Development	3
Applications Development	4
Applications Integration	4
Artificial Intelligence Ethics and Governance	3
Business Needs Analysis	4
Change Management	3
Cloud Computing	4
Computational Modelling	4
Computer Vision Technology	4
Configuration Tracking	3
Continuous Integration and Continuous Deployment	4
Cyber and Data Breach Incident Management	3
Data Analytics	3
Data Engineering	3
Data Ethics	4
Data Visualisation	3
Design Thinking Practice	4
Emerging Technology Synthesis	3

# Philippines Skills Framework for Analytics & AI

## FUNCTIONAL SKILLS & COMPETENCIES

### Analytical Thinking

- Systems Thinking
- Business and Project Management**
- Business Agility
- Business Continuity
- Business Environment Analysis
- Business Innovation
- Business Requirements Mapping
- Business Risk Management
- Change Management
- Crisis Management
- Disaster Recovery Management
- Emerging Technology Synthesis
- Manpower Planning
- Portfolio Management
- Process Improvement and Optimization
- Product Management
- Project Feasibility Assessment
- Project Management
- Strategy Planning
- Sustainability Management

**Business Development**

- Business Negotiation
- Data Analytics
- Networking

**Business Finance**

- Budgeting

**Design and Architecture**

- Data Design
- Design Thinking Practice
- Enterprise Architecture
- Organizational Design
- Security Architecture
- Software Design
- Solution Architecture
- Systems Design

**Development and Implementation**

- Agile Software Development
- Applications Development
- Applications Integration
- Cloud Computing
- Computational Modelling
- Computer Vision Technology
- Configuration Tracking
- Continuous Integration and Continuous Deployment
- Data Engineering
- Data Visualization and Storytelling
- Intelligent Reasoning
- Pattern Recognition Systems

**Business Finance**

- Research
- Self-Learning Systems

**Design and Architecture**

- Software Configuration
- Software Testing
- System Integration
- Test Planning
- Text Analytics and Processing

**General Management**

- Business Performance Management

**Governance and Compliance**

- Audit and Compliance
- Cyber Risk Management
- Data Ethics
- Data Governance
- Data Protection Management
- Data Sharing
- IT Governance
- IT Standards

**Operations and User Support**

- Cyber and Data Breach Incident Management
- Data Migration
- Database Administration
- Performance Management
- Problem Management

**People Development**

- Learning and Development
- People and Performance Management

**Project Management**

- Business Needs Analysis

**Risk Management, Governance and Regulatory Compliance**

- Artificial Intelligence Ethics and Governance

**Sales and Marketing**

- Technical Sales Support

**Stakeholder and Contract Management**

- Partnership Management
- Stakeholder Management

**Strategy and Architecture**

- Business Process Re-engineering
- Quality Standards

**Strategy Planning and Implementation**

- Data Strategy
- Infrastructure Strategy
- IT Strategy
- Organizational Analysis
- Strategy Implementation



[bit.ly/psf-aai](http://bit.ly/psf-aai)

## ENABLING SKILLS & COMPETENCIES

**INTERACTING WITH OTHERS**

Building Inclusivity

**THINKING CRITICALLY**

Creative Thinking

Decision Making



Communication



Problem Solving

Sense Making



Developing People



Transdisciplinary Thinking

**STAYING RELEVANT**

Adaptability

Digital Fluency



Global Perspective

Learning Agility



Self Management



# THANK YOU

