

Business Process Automation with VBA and Python

Mr. Eddie Chow / 9 Sep 2025



Table Of Contents

→ Introduction to business process automation

Business Process automation with VBA

Business process automation with Python

Introduction to project management for business process automation

Development and implementation of business process automation

Final Group Presentation



Lecturer Biography

<https://www.linkedin.com/in/eddie-chow-a3860464/>



- Founder of InnoVi, Co-founder & CTO of 4 companies, covering from property, retail, education and ESG
- 19+ Years' programming experience, including 7 Years' IT Management Experience
- 7 years' teaching experience in tertiary education
- 10+ years' coaching or mentoring experience on large project development..
- Development experience on blockchain, computer vision, machine learning/deep learning
- Researcher in Cutting-edge AI Research, particular in area in Computer vision
- AI Tech Speaker, AI Mentor
- Personal Website : eddiecityu.github.io





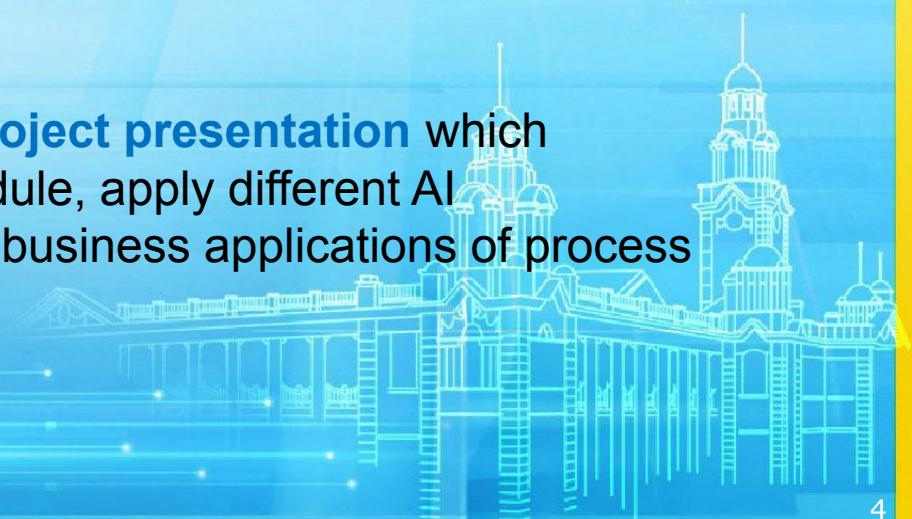
Assessment

- **Continuous Assessment**

In-class discussions and computer assignments which include the usage of different applications & methodologies

- **Final Assessment**

20-mins (3 persons) group project presentation which integrate knowledge in this module, apply different AI methodologies and software to business applications of process automation





What is Process Automation, RPA, IPA?

Overview of the technological building blocks related to process automation

Contemporary tools for process automation

Challenges and opportunities of business process automation

Business implications of process automation



Intended Learning Outcomes

1. describe the contemporary trends of process automation and explain the opportunities and challenges of business process automation;
2. outline key steps in process automation project management and illustrate the significance of each step;
3. apply computational tools to implement process automation;
4. discuss the development of process automation and practical cases for business.

Assessment

Memory Test?

- NO Midterm / Final exams

Essays?

- NO research paper
- NO project report

Practical Work!

- 2 in-class coding exercises
- 1 Group Project



Taking
Exams

Building
Solutions

Lecture 1 - Business Process Automation with VBA and Python

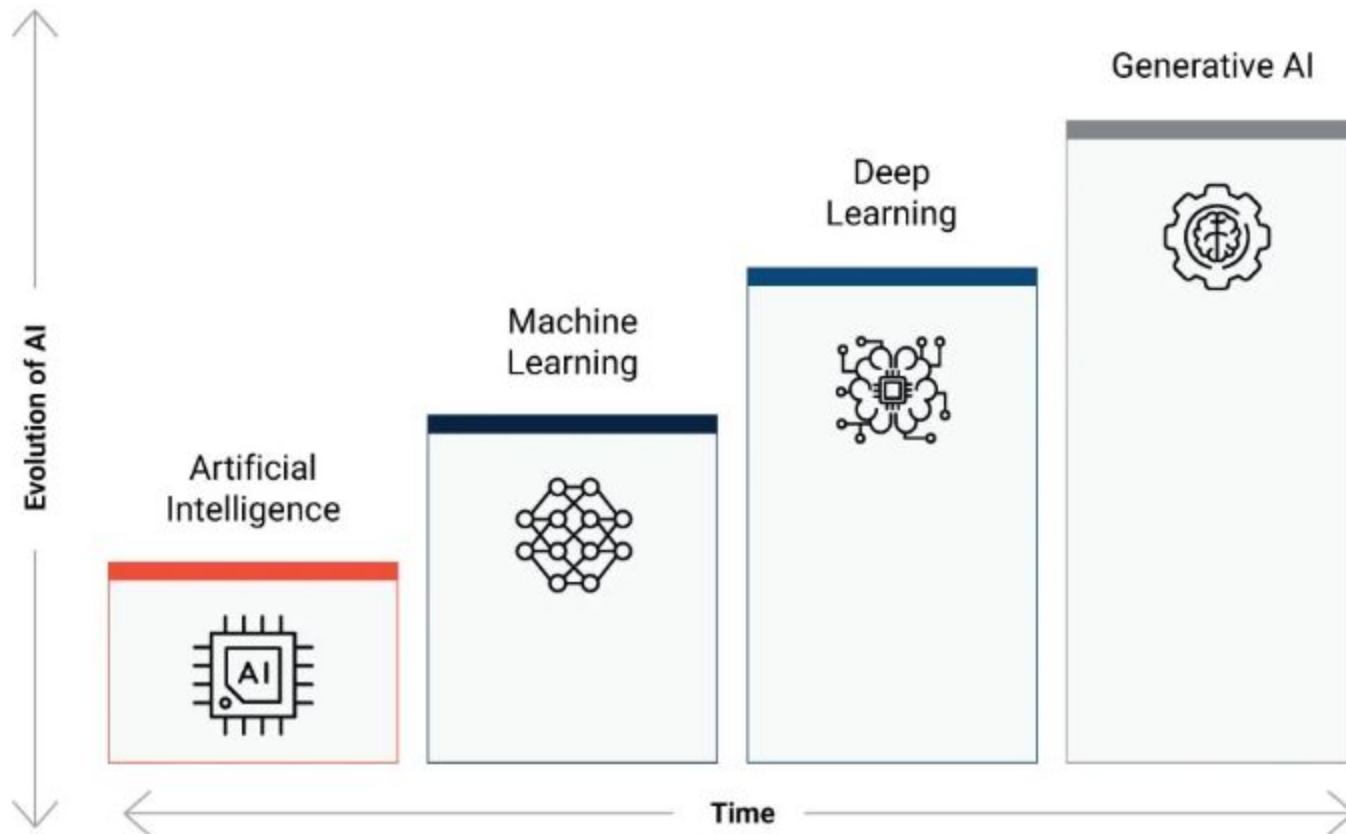


Introductions: Getting To Know You

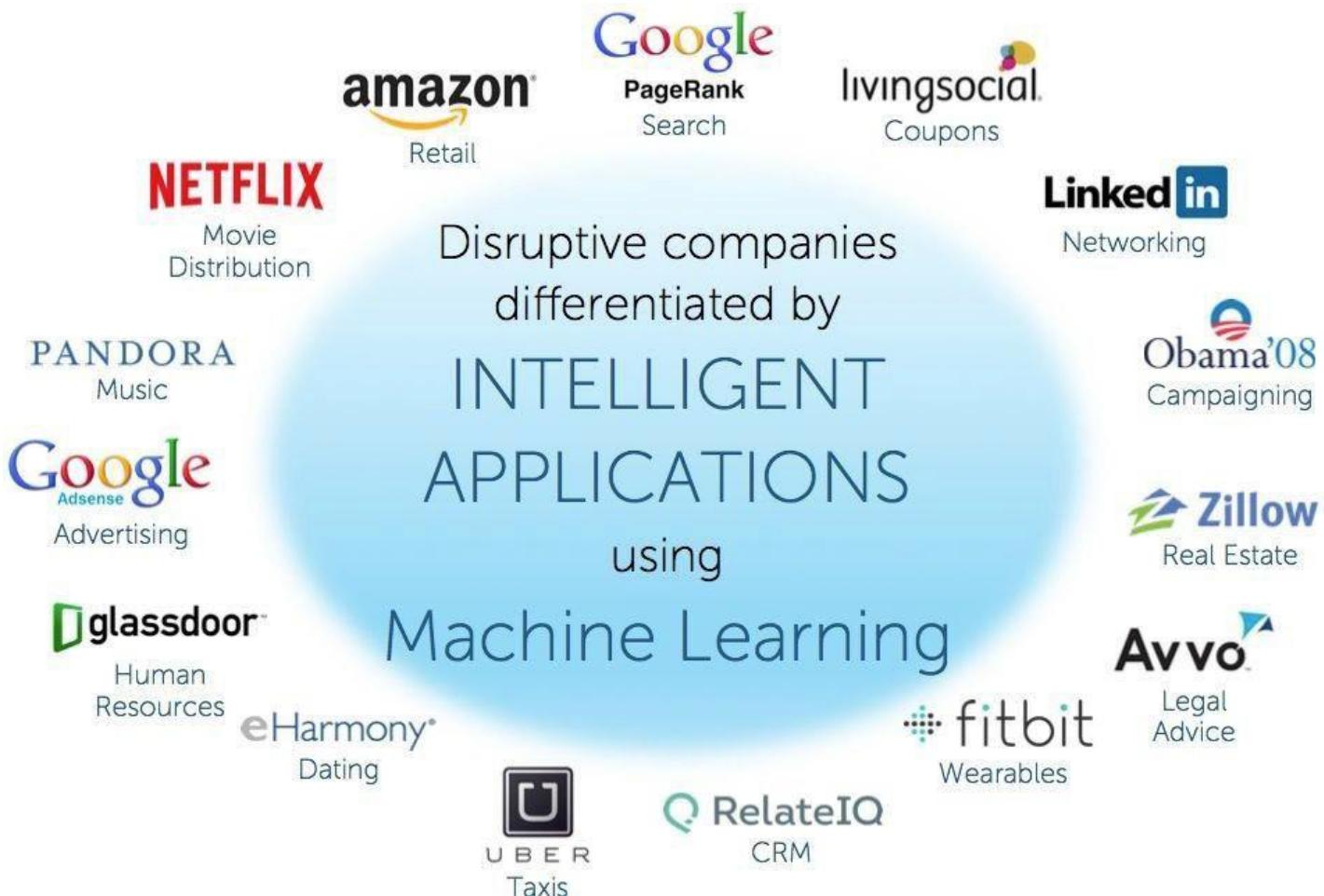
What industry are you working on?
Do you have any coding experience?
Why you want to study this courses?

Forming Groups of 2 – 3 for
Group Project

Overall Artificial Intelligence Journey



The new disruptors



Machine Learning Business Use Case

Face Recognition Algorithm



Cases : Facial recognition for customer analytics



Face Detection

Reliably detect human faces in a photo, along with their coordinates and sizes.



Emotions, Gender, Age

Recognize people's facial expressions and detects primary and secondary emotions. Also detects age and gender

<https://findface.pro/en/solutions/face-recognition-for-customer-analytics.html>

Shifting mindsets of customers: From Products To Relationships



Products
1970s



Customer
Centric



Relationship
Centric Today

These companies applied analytics to stay 1 step ahead of the competition



AI Features in Retail

- Chatbots, Virtual Assistants
- Facial Recognition
- Demographic Segmentation
- Facial Personality Analytics
- AI-enabled Customer Analytics
- Real-time insights
- Personalization
- Inventory Management
- Visual Recognition
- Customer Purchase Prediction
- Trend Prediction
- Purchasing Recommendation
- Ideal Price Point Recommendation
- Predict the trends in Fashion, e-commerce, etc
- Predictive Analysis
- Predictive Personalization

Business Case Studies: Artificial Intelligence

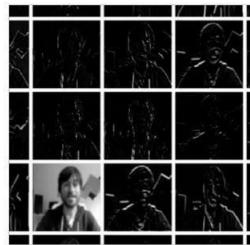
Artificial Intelligence

Common Use Cases:

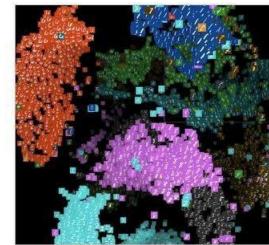
Object Detection
Speech Recognition
Natural Language Processing
Translation between languages
Creativity - Style Transfer
Art Restoration



TEACHABLE MACHINE



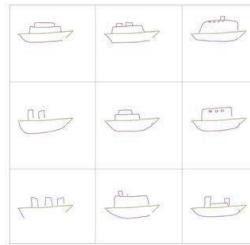
WHAT NEURAL NETWORKS SEE



VISUALIZING HIGH-DIMENSIONAL SPACE



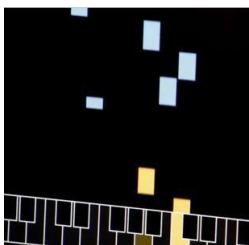
QUICK, DRAW!



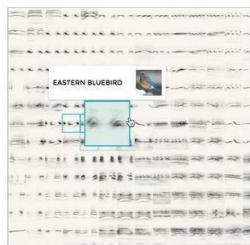
SKETCH-RNN DEMOS



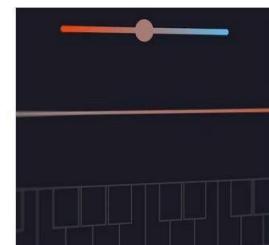
HANDWRITING WITH A NEURAL NET



AI DUET



BIRD SOUNDS



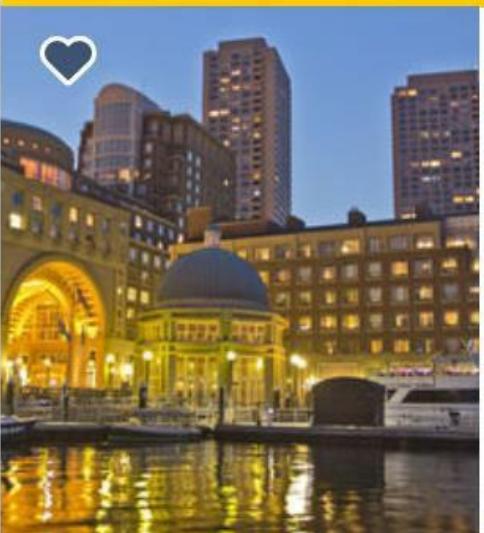
NSYNTH: SOUND MAKER

NLP in real life

- Machine Translation
 - Google Translate translates language from one language to another.
- Text Simplification
 - Rewordify simplifies the meaning of sentences.
- Sentiment Analysis
 - Hater News gives us the sentiment of the user.
- Text Summarization
 - Smmry or Reddit's autotldr gives a summary of sentences.
- Auto-Predict
 - Google Search predicts user search results.
- Auto-Correct
 - Google Keyboard and Grammarly correct words otherwise spelled wrong.
- Speech Recognition
 - Google WebSpeech or Vocalware.
- Question Answering
 - IBM Watson's answers to a query.
- Natural Language Generation
 - Generation of text from image or video data

Machine Learning Business Use Case

Hotel Recommendation Engine



Heart

Boston Harbor Hotel

★★★★★

Downtown Boston

Come Experience our 5 Star Luxury Hotel.

Iconic Waterfront Hotel with a Convenient Location to Everything Boston Has to Offer. Every Room Boasts City or Harbor Views.

1-866-286-0843 • Expedia Rate

Viewed

4.8/5 Exceptional!
(1,132 reviews)

\$693-\$581

nightly price
Sponsored

Get member price

People who looked at the Boston Harbor Hotel also viewed these:



Four Seasons Hotel Boston

★★★★★

\$645



The Ritz-Carlton, Boston

★★★★★

\$595

Introduction to Business Process Automation

What is Business Automation?

From ChatGPT...

JA

What is business automation? asking for a friend :)

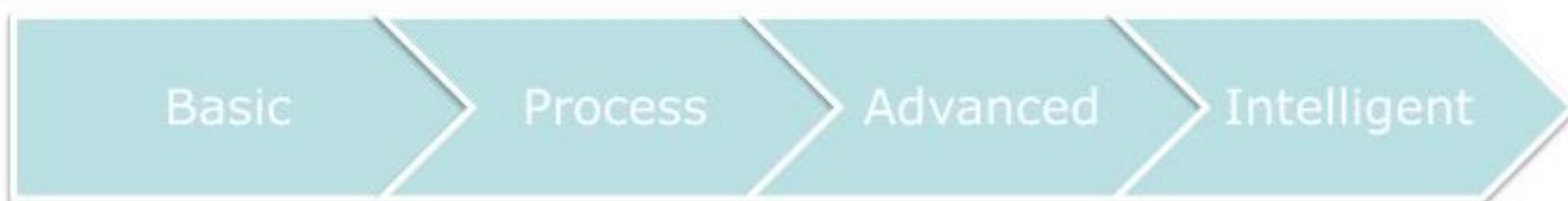


Business automation refers to the process of using technology to automate repetitive or routine tasks and processes within a business, such as data entry, invoicing, and inventory management. It can also involve the integration of various software applications and systems to streamline workflows and improve overall efficiency.

Introduction to Business Process Automation

Types of Business Automation

1. Basic Automation
2. Process Automation
3. Advanced Automation
4. Intelligent Automation



1. *The Basics of Business Automation.* (2021, May 12). The Basics of Business Automation | IBM.
<https://www.ibm.com/cloud/blog/basics-of-business-automation>

Introduction to Business Process Automation

Types of Business Automation – Basic Automation

- **Target →** Simple / Repetitive tasks
- **Why**
 - Eliminate human errors
 - Accelerate pace of transactional work
- **Example →** Robotic Process Automation (RPA)*



* To be covered in the next section

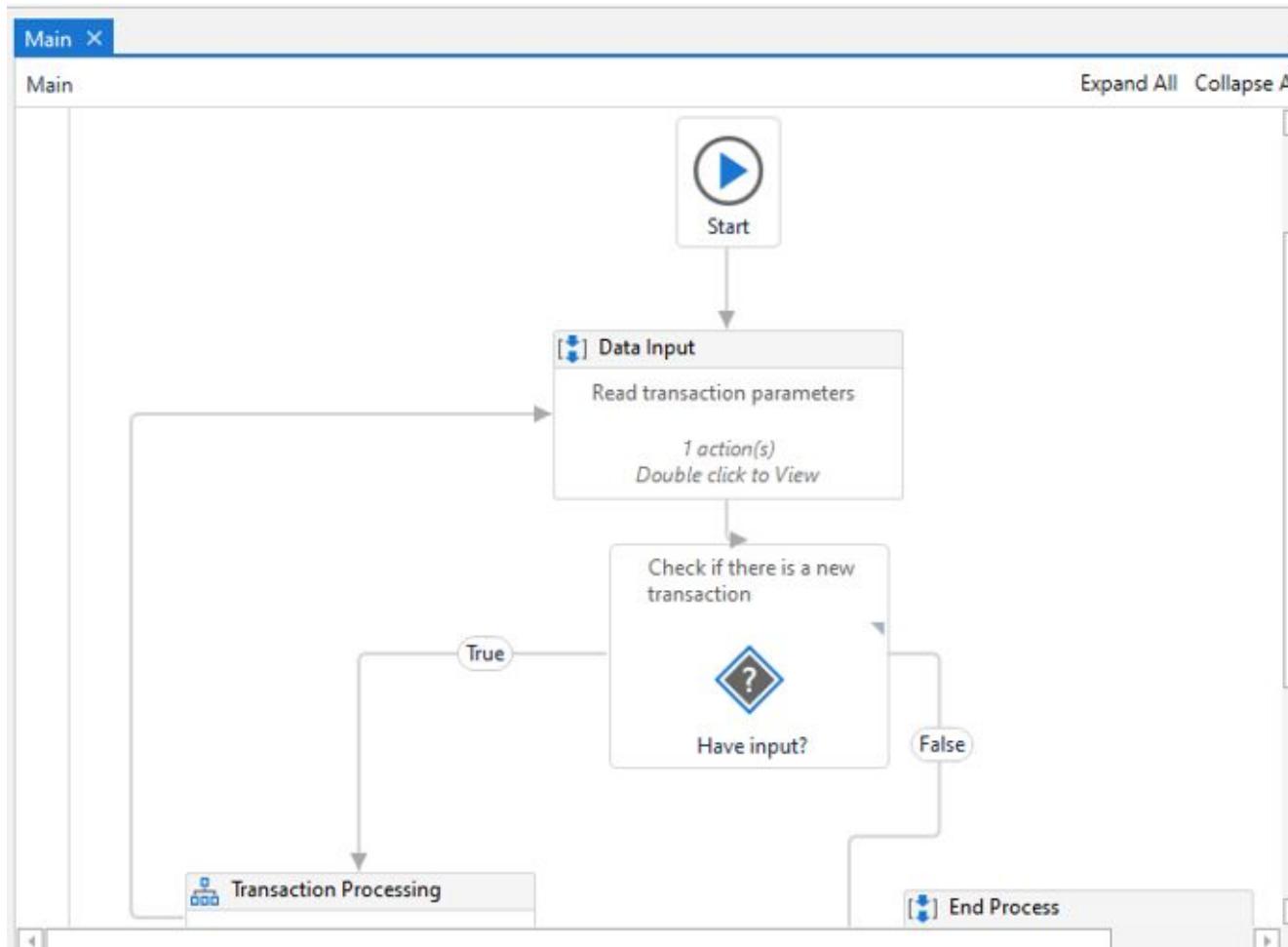
Introduction to Business Process Automation

Types of Business Automation – Process Automation

- **Target →** Process requiring dedicated solutions
- **Why**
 - Increase productivity and efficiency
- **Example →** Process Mining, workflow automation



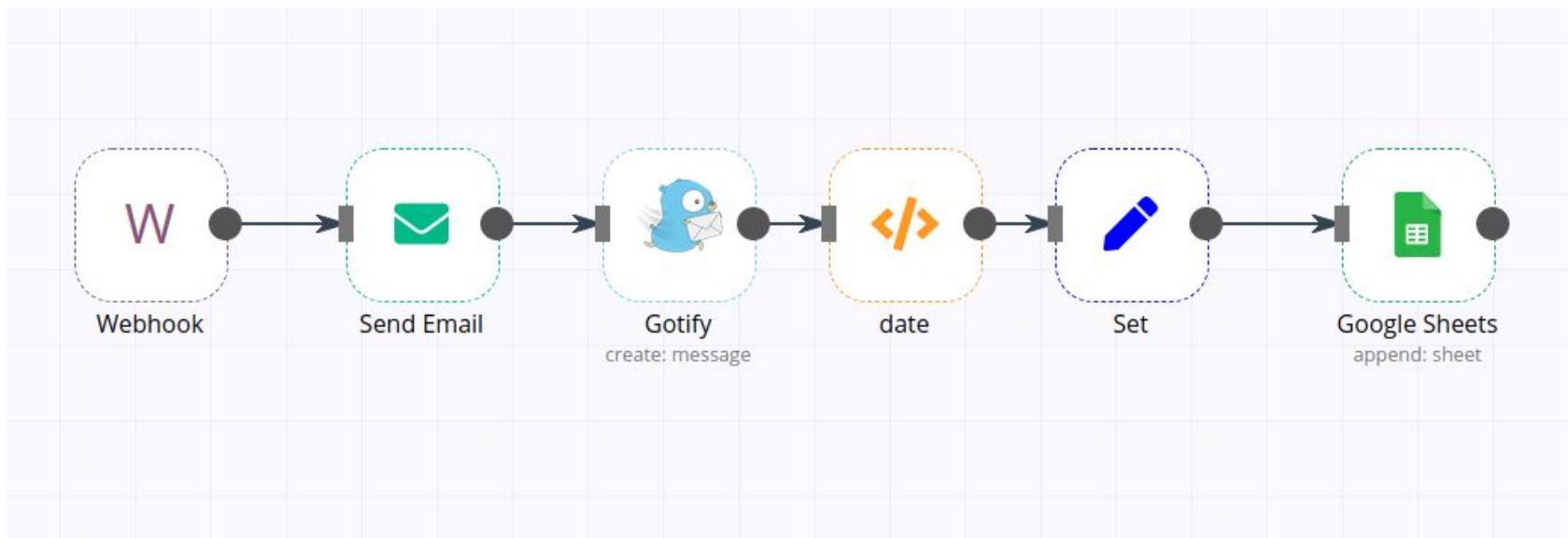
Process Automation Software Example 1 : UiPath



<https://www.uipath.com/rpa/robotic-process-automation>

Process Automation Software Example 2 : n8n

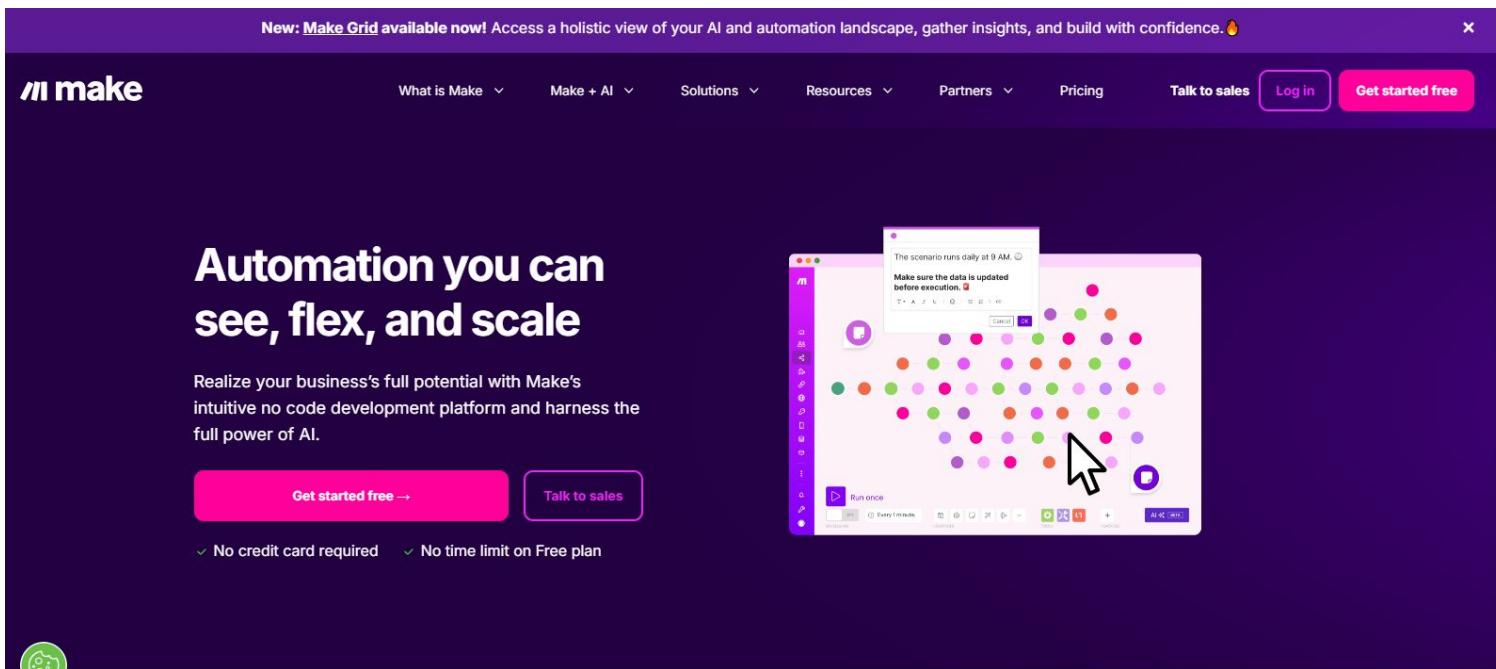
Simple enough to ship in hours, sophisticated enough to scale. n8n lets you automate business processes without limits on your logic.



<https://n8n.io/features/>

Practical Exercise - Process Automation (Make.com)

Go to <https://www.make.com/en>



Click “Get started free” or enter <https://www.make.com/en/login>

Simple Automation using NoCode Platform

Sign in with Google

The screenshot shows the 'My Organization' dashboard. On the left is a purple sidebar with navigation links: 'Organization' (selected), 'Team', 'Scenarios', 'AI Agents (BETA)', 'Templates', 'Connections', 'More', 'Resource Hub', 'What's New', 'Notifications', 'Help', and 'Eddie Chow'. The main area has a header 'My Organization' with tabs: 'ORGANIZATION' (selected), 'TEAMS', 'USERS', 'SUBSCRIPTION', 'PAYMENTS', 'INSTALLED APPS', 'VARIABLES', and 'SCENARIO PROPERTIES'. A red box highlights the 'Create a new scenario' button in the top right corner. Below the header, there's a 'Dashboard' section with a 'Current plan - Free' summary table:

| Price | \$0.00 |
|-------------------|---------|
| Operations/mo | 1,000 |
| Unused operations | 998 |
| Billing interval | Monthly |

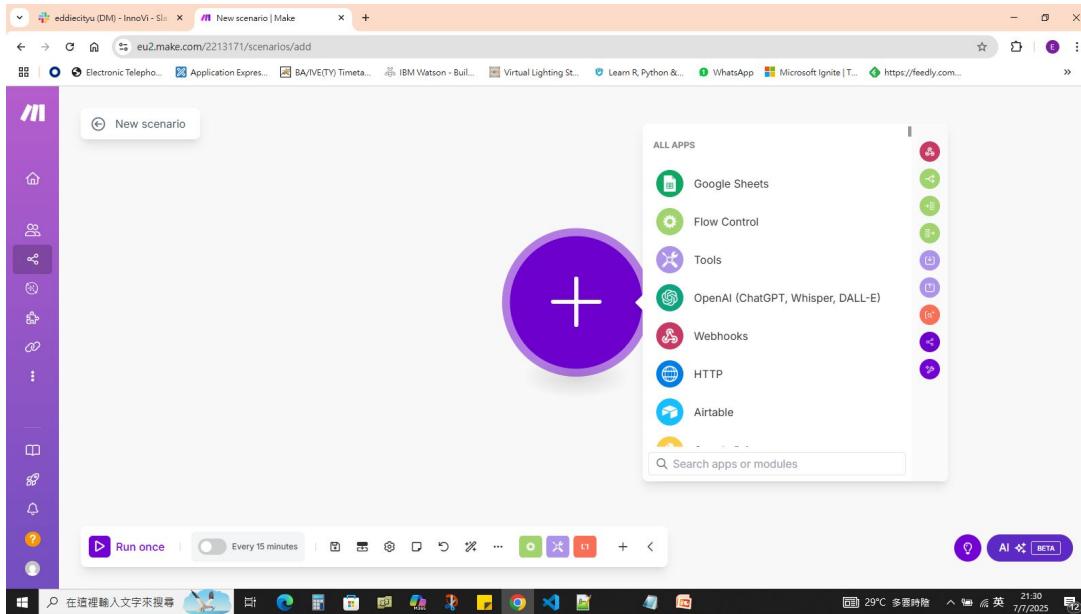
A purple 'Upgrade' button is located below this table. To the right is a chart showing 'Active scenarios' (0/2), 'Operations (~2/1,000 (0% used))', and 'Data transfer (~592.0 B/512.0 MB (0% used))'. A large graph at the bottom shows usage over time from July to June, with a sharp spike in late May.

Below the dashboard are sections for 'Active scenarios' and 'Recommended for you'. A blue 'Search for help' button is in the bottom right. The overall interface is clean with a white background and light gray grid lines.

Click “Create a new scenario”

Simple Automation using NoCode Platform

An interface with module is displayed



Create Module “Weather” -> Click “Get daily weather forecast”

Simple Automation using NoCode Platform

Enter Days “9 days from today” and City “Hong Kong”
-> Click “Save”

Right-Click and select “Run this module only”

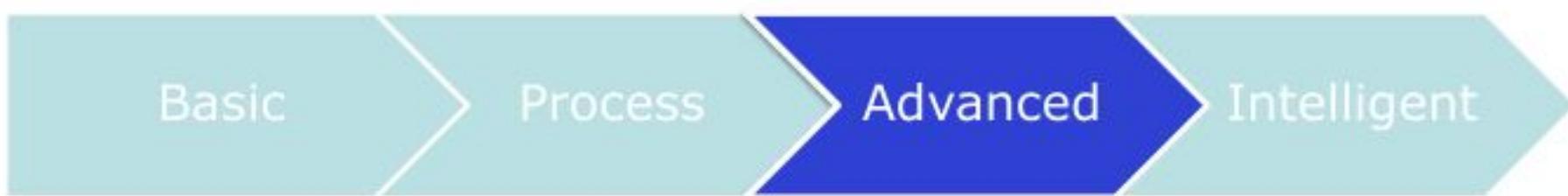
The screenshot shows a NoCode platform interface with a sidebar containing various icons. In the main area, there is a weather forecast module with a sun and cloud icon. A context menu is open over the module, listing options: Settings, Run this module only, Add a module, Add error handler, Rename, Clone, Copy module, Add a note, and Delete module. To the right of the module, there is a preview window titled "Weather" showing a summary of the forecast. The preview includes a large orange circle with a clock and a speech bubble, followed by the word "Weather" and a small green checkmark. Below this, it says "Get daily weather forecast". On the far right, there is a detailed view of the "Weather" module's configuration, showing sections for Initialization, Operation 1, INPUT, and OUTPUT. The INPUT section shows a bundle with City: Hong Kong and Days: (array). The OUTPUT section shows a bundle with Date: 2025年7月16日 12:00, Temperature: (Collection), Atmospheric pressure: 997, Air humidity: 80, Cloudiness: 100, Wind: (Collection), Rain: 7.92, Status: Rain, Code: 500, Description: light rain, and an icon URL. There are also sections for Commit and Finalization.

You can view an output file in JSON format

Introduction to Business Process Automation

Types of Business Automation – Advanced Automation

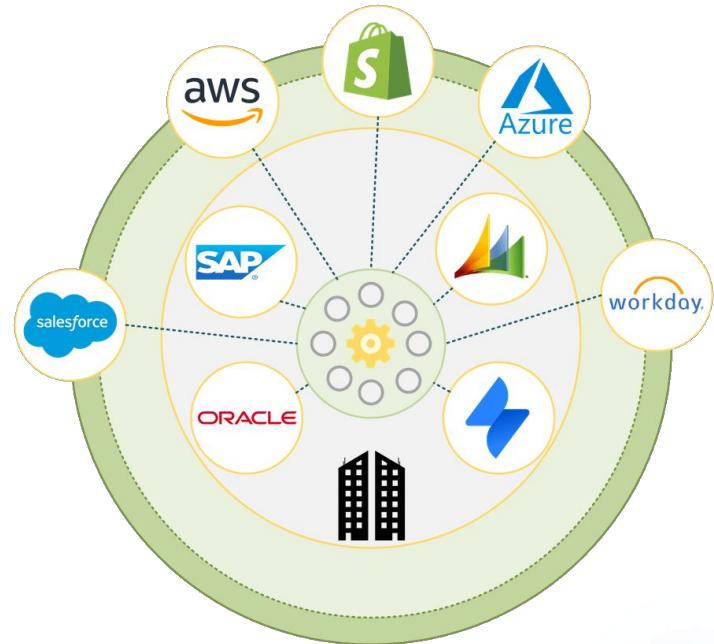
- **Target →** Complex process: Multiple systems + Human
- **Why**
 - Handles unstructured data / methods of access
- **Example →** Machine learning, natural language processing



Advanced Automation : System Integration



<https://fitsmallbusiness.com/crm-integration/>

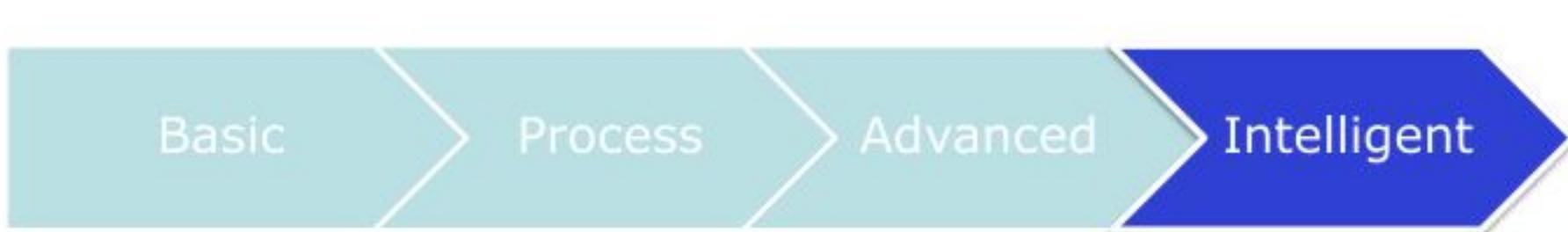


<https://iwconnect.com/making-multiple-systems-work-as-one-through-integration/>

Introduction to Business Process Automation

Types of Business Automation - Intelligent Automation

- **Target →** Processes requiring AI/customized decisions
- **Why**
 - Smarter interactions, personalization
- **Example →** Voice-based interfaces



Introduction to Business Process Automation

Types of Business Automation - Checkpoint

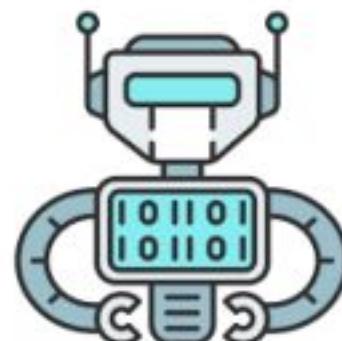
1. Daily email template to enter Market commentary
2. HR time off workflow
3. Online supermarket order processing
4. Voice-automated credit card hotline



Introduction to Business Process Automation

Robotic Process Automation (RPA) - Introduction

- Rule-based software
- Automating high-volume activities
- Free up human workers (for more meaningful tasks)

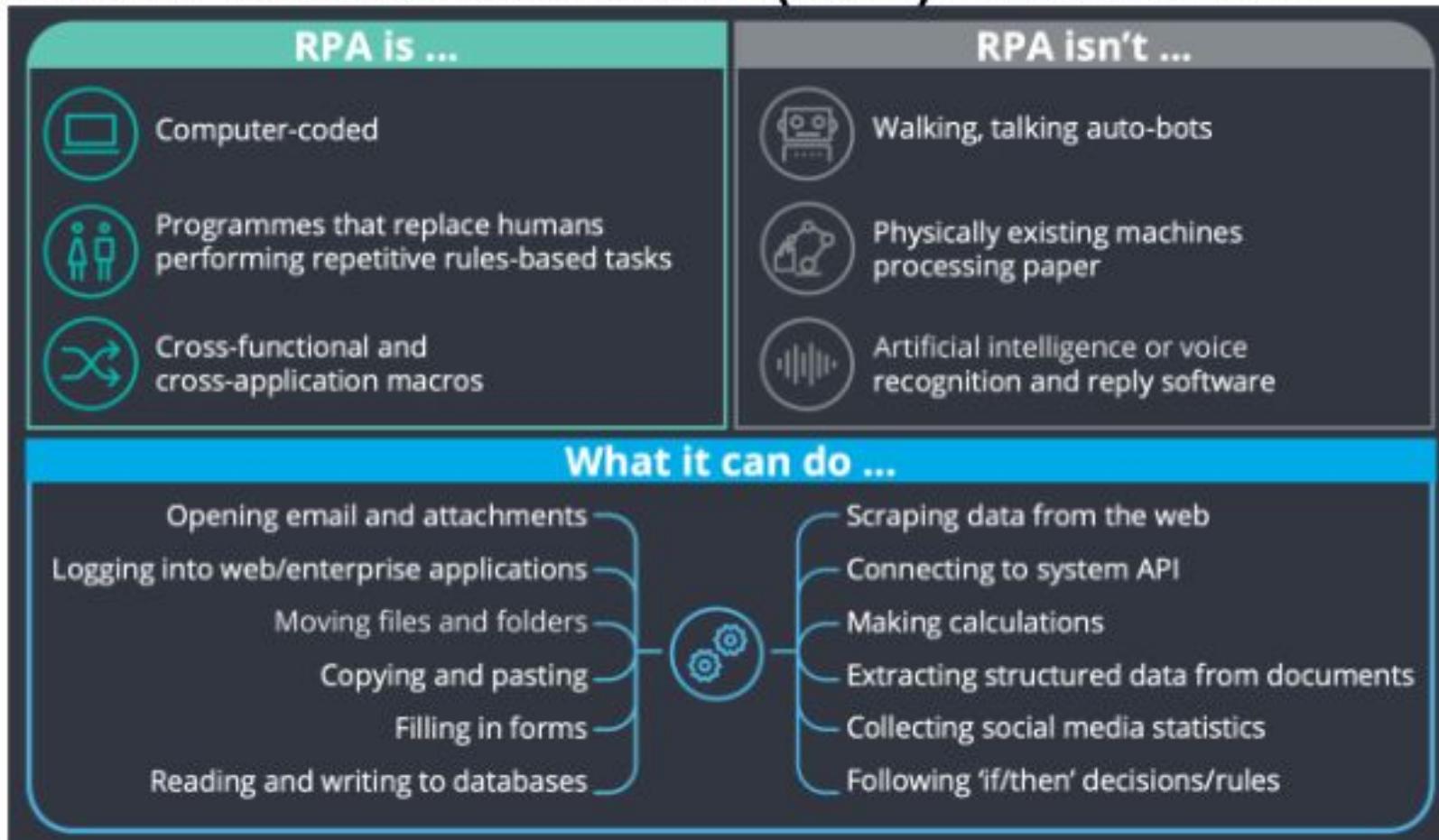


RPA bot

1. *What is Robotic Process Automation (RPA)?* | IBM. What Is Robotic Process Automation (RPA)? | IBM. <https://www.ibm.com/topics/rpa>

Introduction to Business Process Automation

Robotic Process Automation (RPA) - Introduction



1. Deloitte. (2018). *The robots are ready. Are you?*

<https://www2.deloitte.com/content/dam/Deloitte/tr/Documents/technology/deloitte-robots-are-ready.pdf>

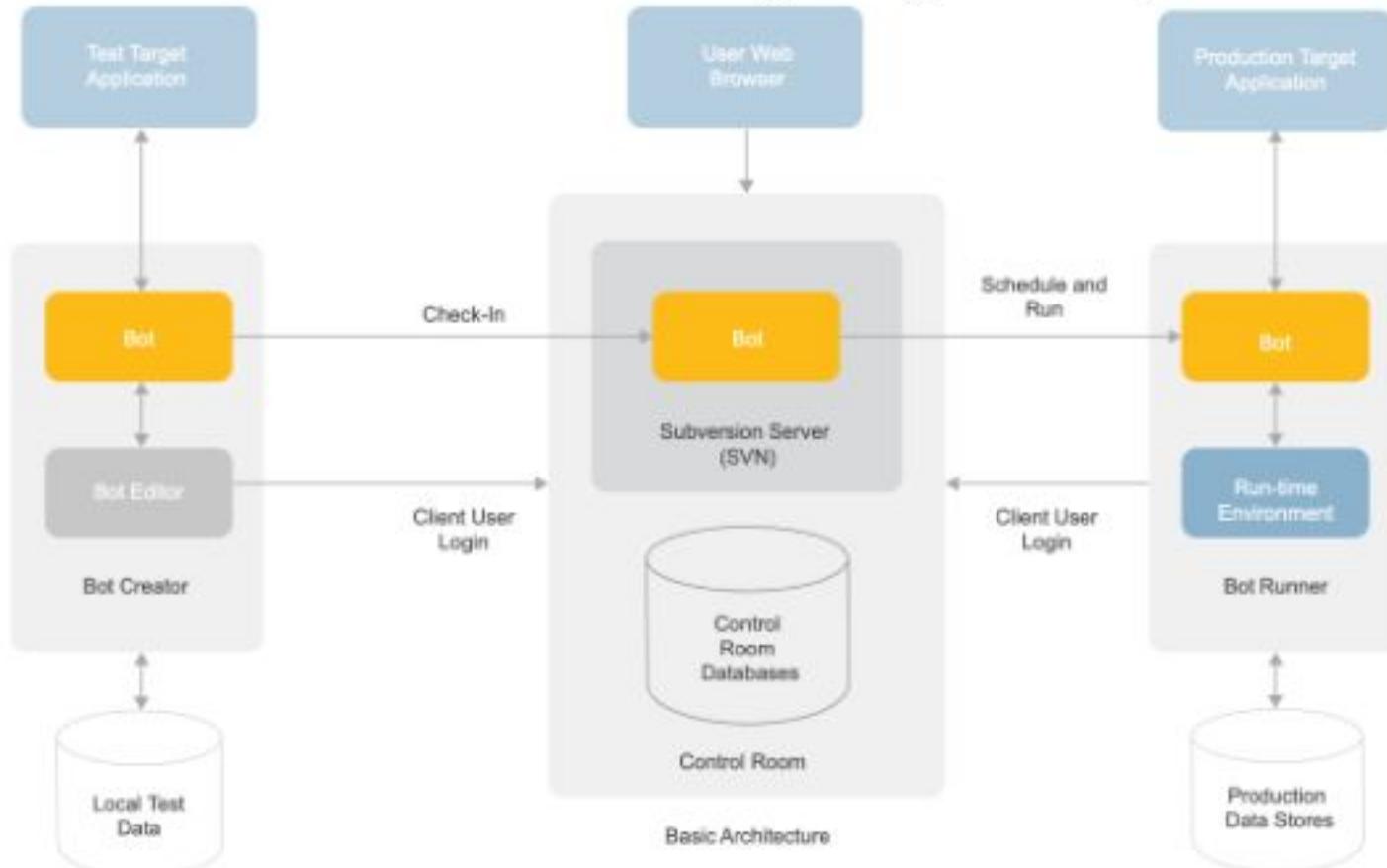
Robotic Process Automation (RPA) - Components

- **Bot modeling**
 - Bot Creator → development/ testing
 - Bot Runner → execution of tested codes in production
- **Control Room**
 - Deployment / Scheduler
 - Audit trails
 - Performance Analytics
 - Role-based access control



Introduction to Business Process Automation

Robotic Process Automation (RPA) - Components



1. Automation Anywhere. *Enterprise-class Security for Robotic Process Automation.*
https://www.automationanywhere.com/sites/default/files/internal-assets/uberflip/security-whitepaper_en.pdf

Introduction to Business Process Automation

Intelligent Process Automation (IPA)

- Using Artificial Intelligence (AI), e.g.,
 - Machine Learning (ML)
 - Natural Language Processing (NLP)
- Enhancing cognitive ability of the automation



Introduction to Business Process Automation

Intelligent Process Automation (IPA)

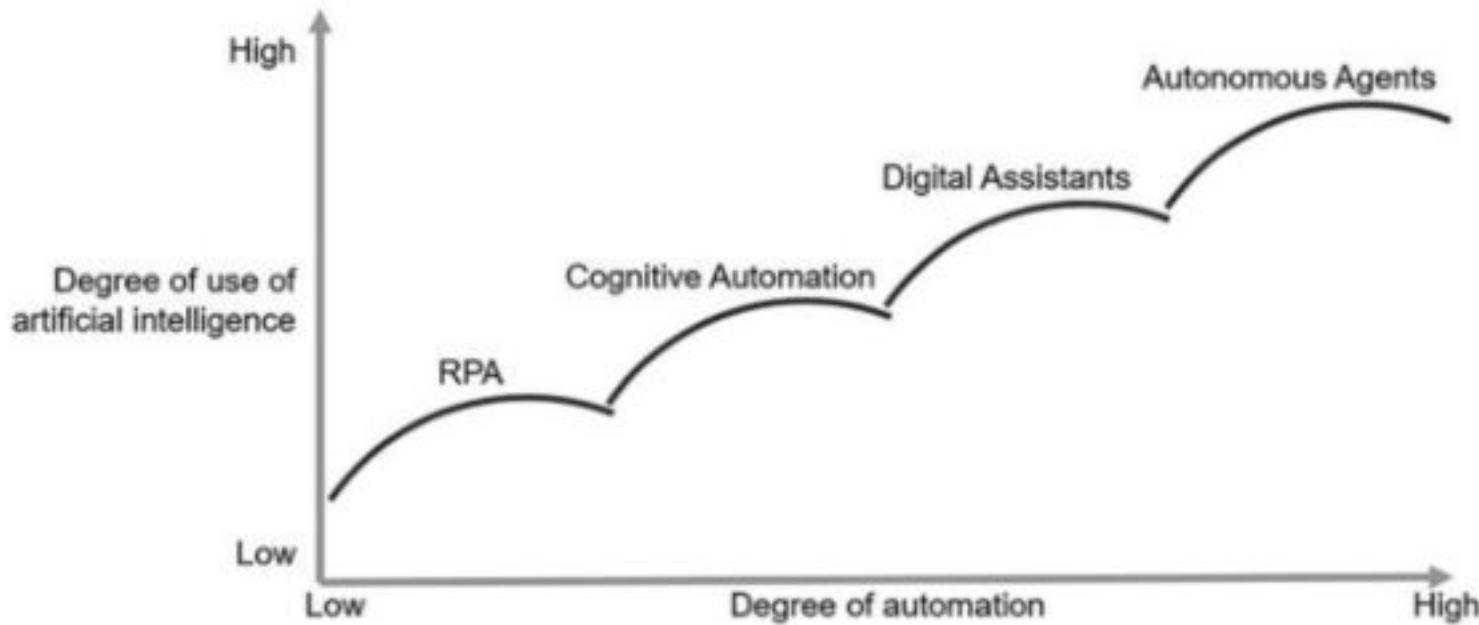
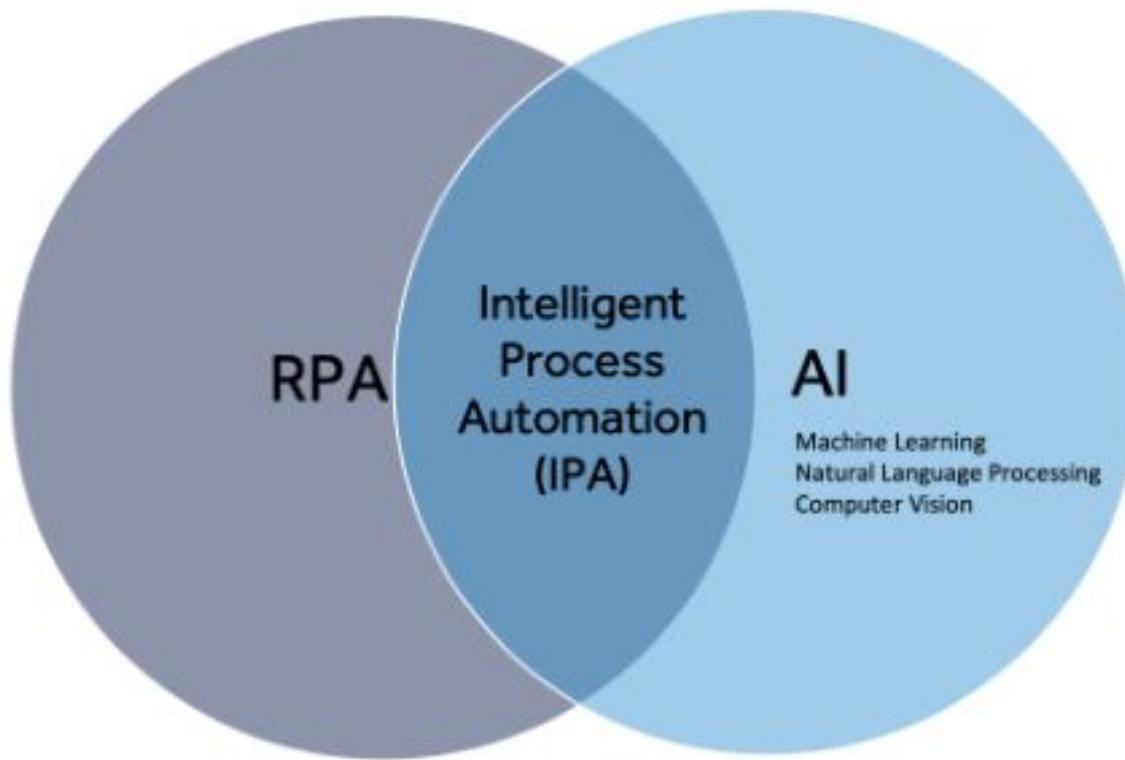


Fig. 2.2 Classification of RPA according to the degree of automation and the use of artificial intelligence (based on Ostrowicz 2018, p. 4)

1. Smeets, M., Erhard, R., & Kaußler, T. (2021, July 30). *Robotic Process Automation (RPA) in the Financial Sector: Technology - Implementation - Success for Decision Makers and Users*. Springer. <https://doi.org/10.1007/978-3-658-32974-7>

Introduction to Business Process Automation

Intelligent Process Automation (IPA)



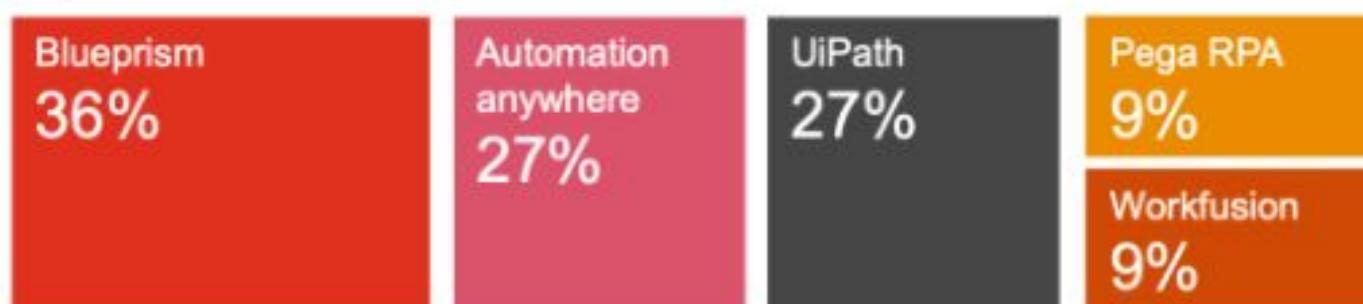
1. Kosmopoulos, C. (2021, May 11). *What is Intelligent Process Automation (IPA)?* | Blueprint. What Is Intelligent Process Automation (IPA)? | Blueprint. <https://www.blueprintsys.com/blog/rpa/what-is-intelligent-process-automation-ipa>

Introduction to Business Process Automation

Tools for Automation - Overview

Which tool is being used ?

As some companies use more than one tool, the percentages in this graphic total to more than 100%

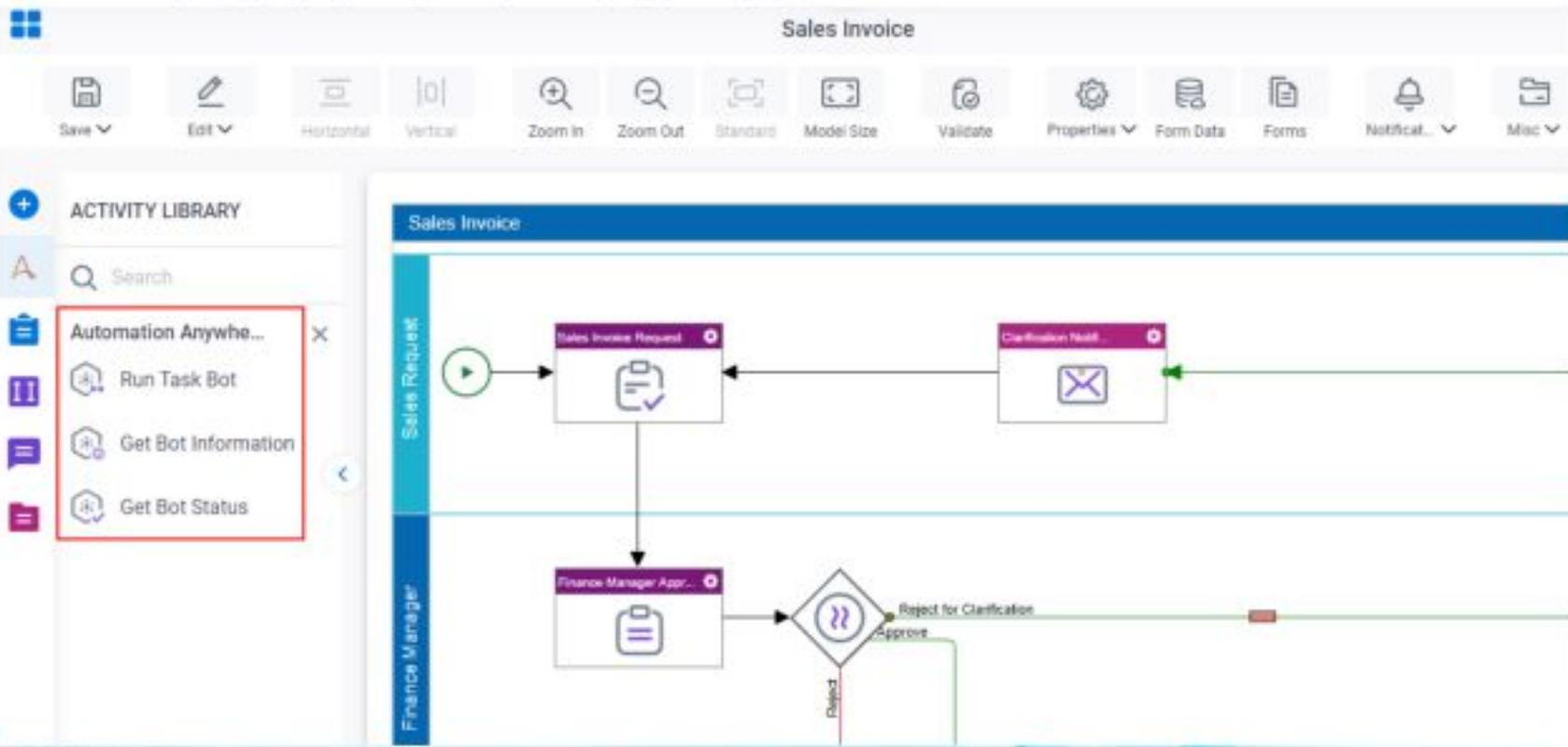


1. 2019 actuarial robotic process automation (RPA) survey report. PwC.
<https://www.pwc.com/gx/en/industries/financial-services/publications/pwc-2019-actuarial-robotic-process-automation-survey-report.html>

Introduction to Business Process Automation

Tools for Automation - Types

- **Low-code platforms**
 - Automation Anywhere, XCEPTOR, Microsoft Power Automate



Introduction to Business Process Automation

Tools for Automation - Types

- In-house Solutions

- Excel VBA, Java, Python, Web

Microsoft®
Excel VBA



Most in-demand functional areas

Compliance
Operational Risk
Fund Accounting
Corporate Accounting
Actuarial
Finance Operations
IT Audit

In-demand technical skills and experience

IFRS 17 Reporting and Accounting Policy
RBC and Capital Reporting
IT skills (Python, C+ BASIC, C, C++, COBOL, Java, R, Tableau, SQL, VBA, and Macro)
Operational Risk, Information and IT Risk, Risk Transformation, Risk Analytics, Internal Control Framework



1. Robert Half. 2024 Salary Guide HK.
<https://content.roberthalfonline.com/SG24/SG24-PDF/2024-Salary-Guide-HK.pdf>

Introduction to Business Process Automation

Tools for Automation – Getting a taste

- **Visualping** (<https://visualping.io/>)
 - Websites monitoring



<https://hkuspace.hku.hk/prog/cert-for-module-business-process-automation-with-vba>

Certificate for Module (Business Process Automation with VBA ...)

(2) Mr Jackie Liu. Mr. Jackie Liu is veteran quantitative strategist currently working in one of the top global investment banks. With over 10 years of ...

 hku.hk
<https://hkuspace.hku.hk/prog/cert-for-module-business-process-automation-with-vba>

Certificate for Module (Business Process Automation with VBA ...)

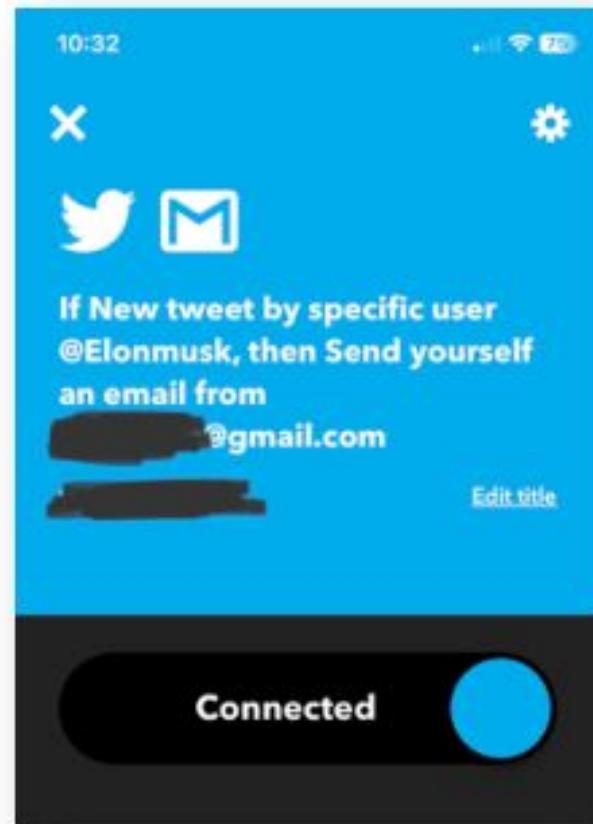
(2) Mr Jackie Liu. Mr. Jackie Liu is veteran quantitative strategist currently working in one of the top global investment banks. With over 10 years of ...

Introduction to Business Process Automation

Tools for Automation – Getting a taste

- **IFTTT (<https://ifttt.com/>)**

- Web + App interfaces
- Service-based triggers and actions



Introduction to Business Process Automation

Tools for Automation – Considerations

- Cost

- One-time cost

- Initial effort of adaptation
 - Implementation and rollout

- Running cost

- License, e.g., Power Automate: \$Millions/year¹
 - Infrastructure, e.g., Cloud storage
 - Training, e.g., Automation Anywhere, 10+ hours²
 - Support, e.g., bot re-configuration

| Cost | | |
|------------|-------|--------|
| Production | Gold | Maint. |
| 60 🎁 | 240 🎁 | 1 🎁 |

1. RPA Pricing: Comparison of Leading RPA Vendors' Fees in 2023. <https://research.aimultiple.com/rpa-pricing/>

2. RPA Course Learning Trails | Automation Anywhere University. <https://university.automationanywhere.com/training/rpa-learning-trails/>

Introduction to Business Process Automation

Tools for Automation – Considerations

- **Flexibility**
 - Specialized vs Generalized, e.g., XCEPTOR → Tax
- **Scalability**
 - Desktop vs Enterprise, e.g., Excel VBA
 - Version Control, e.g., Use of repository
- **Time to delivery**
 - “Perfect” solution does not exist



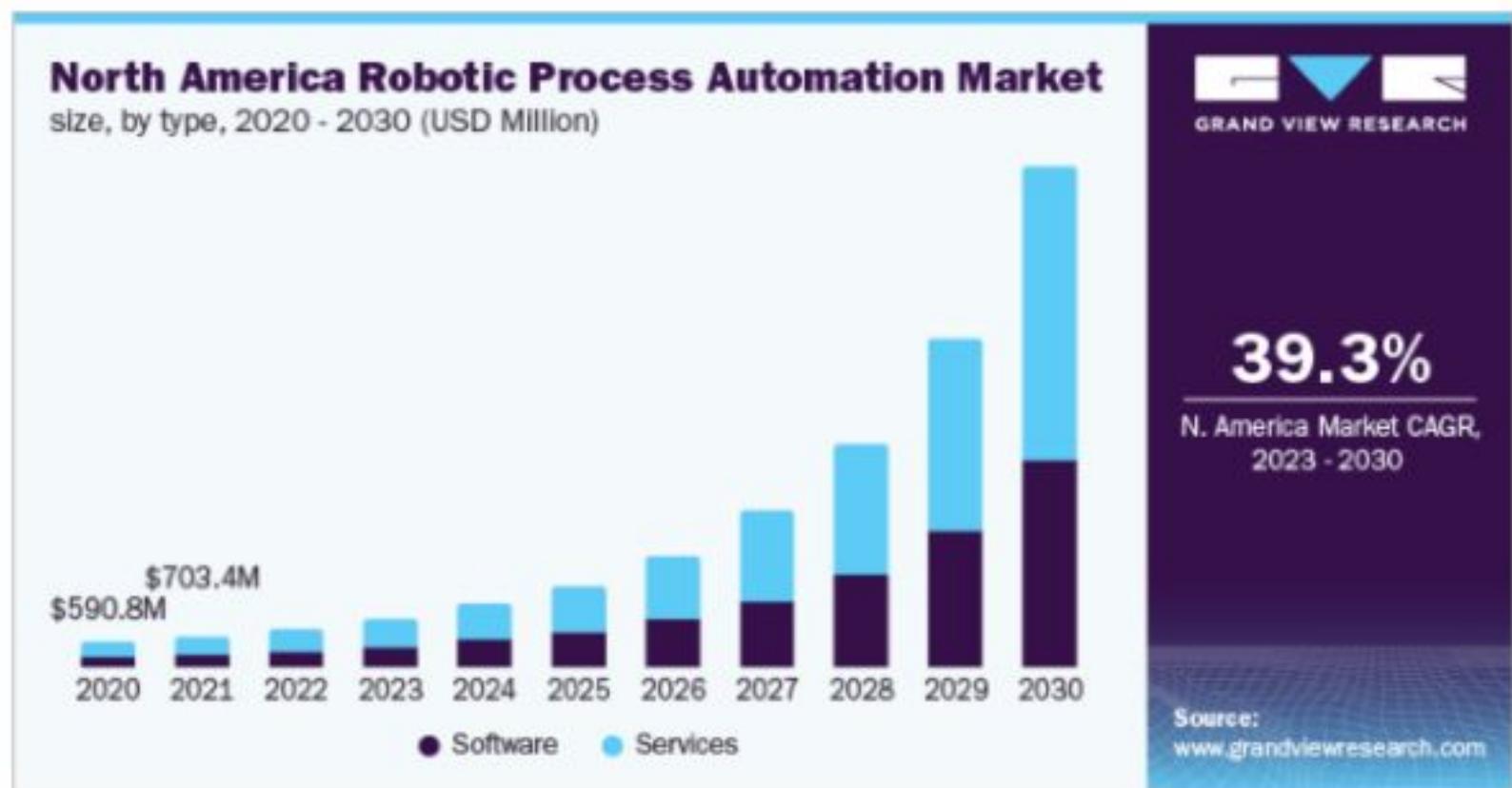
GitHub



1. Smeets, M., Erhard, R., & Kaußler, T. (2021, July 30). *Robotic Process Automation (RPA) in the Financial Sector: Technology - Implementation - Success for Decision Makers and Users*. Springer. <https://doi.org/10.1007/978-3-658-32974-7>

Introduction to Business Process Automation

Opportunities of BPA – The Money



1. *Robotic Process Automation Market Size & Share Report 2030.* Robotic Process Automation Market Size & Share Report 2030.
<https://www.grandviewresearch.com/industry-analysis/robotic-process-automation-rpa-market>

Introduction to Business Process Automation

Opportunities of BPA – Sectors

| Industries | % of RPA Solutions |
|---|--------------------|
| Banking, Financial Services & Insurance | 51 % |
| Business Process Outsourcing (BPO) | 14 % |
| Manufacturing – Consumer Packaged Goods (CPG) | 7 % |
| Professional, Legal & Accountancy Services | 7 % |
| Retail Trade | 7 % |
| Technology (IT, Internet, SAAS) | 7 % |
| Utilities | 7 % |

1. Top 15 RPA Use Cases & Examples in Banking in 2023. <https://research.aimultiple.com/banking-rpa/>

Opportunities of BPA – Sectors

- **Healthcare**
 - Monitoring, drug dispensing
- **Retail**
 - Auto-checkout, smart shelf, inventory management
- **Marketing**
 - Lead generation
 - social media management
- **Finance**
 - Fraud detection, Credit management, Regtech



Introduction to Business Process Automation

Opportunities of BPA – Regtech

- Follow the money

 CNN

Wells Fargo ordered to pay \$3.7 billion for 'illegal activity' including unjust foreclosures and vehicle repossession

The CFPB said the more than \$2 billion in customer refunds Wells Fargo has been ordered to pay includes more than \$1.3 billion to consumers hurt...

20 Dec 2022



 Spiceworks

SEC Penalizes Major Wall Street Firms \$1.97B For Using Unauthorized Messaging Apps

If there are allegations of wrongdoing or misconduct, we must be ... Morgan Stanley recently agreed to a \$35 million fine by the SEC for...

28 Sept 2022



Introduction to Business Process Automation

Opportunities of BPA – Regtech

- Follow the money



1. KPMG. (2019, June). *There's a revolution coming.*

<https://assets.kpmg.com/content/dam/kpmg/cn/pdf/en/2019/06/embracing-the-challenge-of-the-new-regtech-era.pdf>

Introduction to Business Process Automation

Opportunities of BPA – Regtech

- Common themes



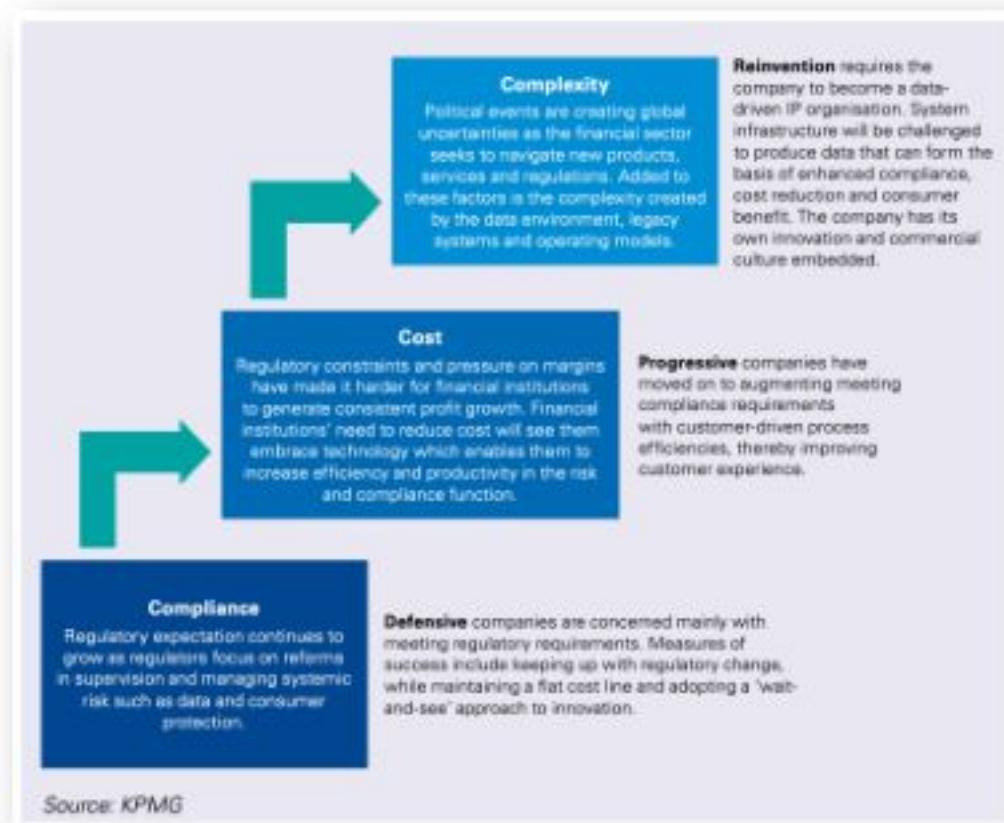
1. KPMG. (2022, November). <https://assets.kpmg.com/content/dam/kpmg/uk/pdf/2022/11/innovate-finance-regtech-industry-and-adoption.pdf>

Introduction to Business Process Automation

Opportunities of BPA – Regtech

Adoption phases:

1. Defensive
2. Progressive
3. Reinvention



1. KPMG. (2019, June). *There's a revolution coming.* <https://assets.kpmg.com/content/dam/kpmg/cn/pdf/en/2019/06/embracing-the-challenge-of-the-new-regtech-era.pdf>

Introduction to Business Process Automation

Opportunities of BPA – Regtech

- HKMA 3-Year Roadmap
 - RPA: automating rule-based processes
 - Network Analysis: Company holdings and exposures

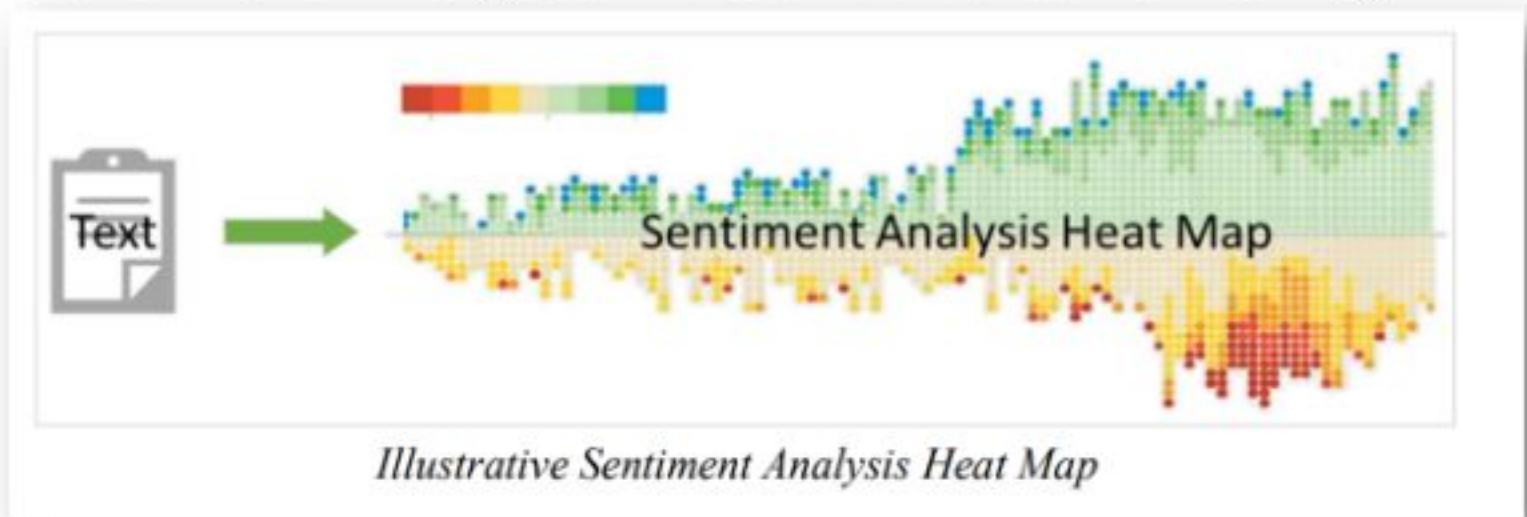


1. HKMA. (2021, June). *Regtech Watch Issue No. 7*.
<https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2021/20210617e1a1.pdf>

Introduction to Business Process Automation

Opportunities of BPA – Regtech

- **HKMA 3-Year Roadmap**
 - **Speech-to-text (STT)**: Audio to time-stamped text
 - **Sentiment Analysis**: Realtime news monitoring

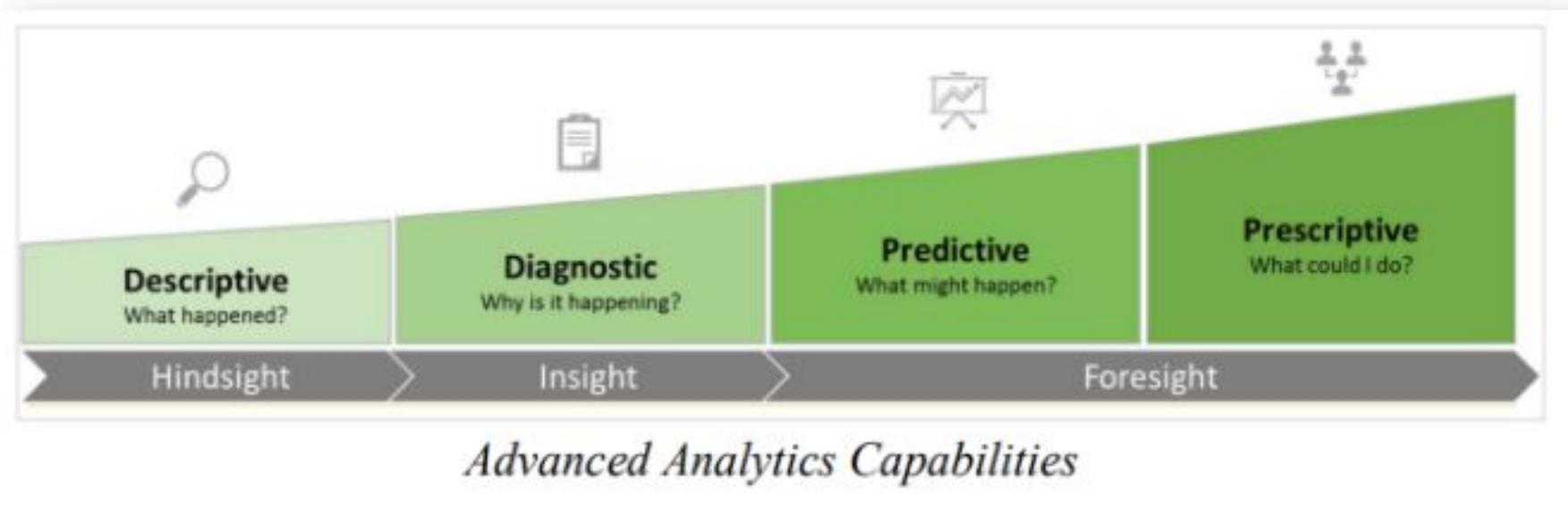


1. HKMA. (2021, June). *Regtech Watch Issue No. 7*.
<https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2021/20210617e1a1.pdf>

Introduction to Business Process Automation

Opportunities of BPA – Regtech

- Analytics enabled by BPA



1. HKMA. (2021, June). *Regtech Watch Issue No. 7*.
<https://www.hkma.gov.hk/media/eng/doc/key-information/guidelines-and-circular/2021/20210617e1a1.pdf>

Introduction to Business Process Automation

Challenges of BPA – Insight from the banks

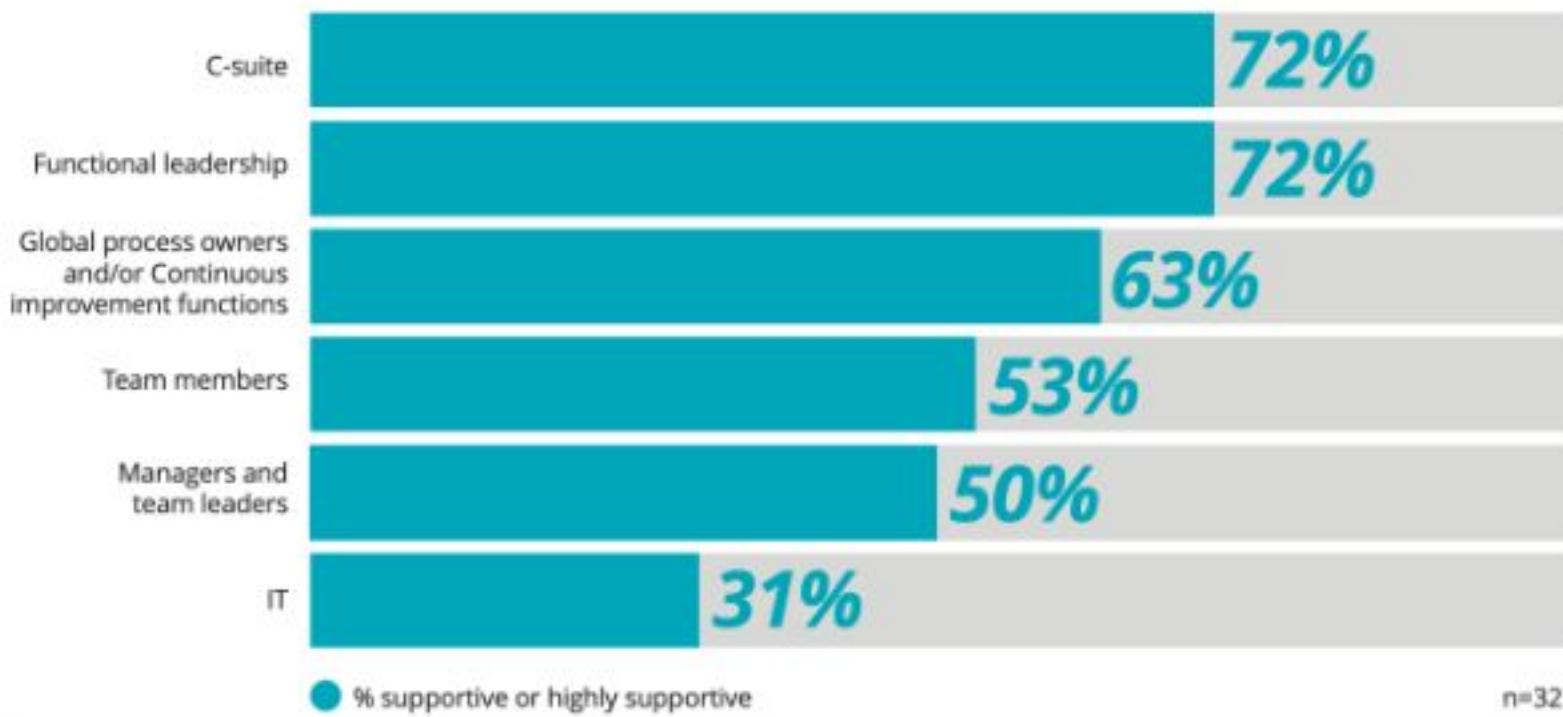


1. KPMG & HKMA. (2020). *Transforming Risk Management and Compliance: Harnessing the Power of Regtech.* <https://www.hkma.gov.hk/media/chi/doc/key-information/press-release/2020/20201102c3a1.pdf>

Introduction to Business Process Automation

Challenges of BPA – Worries from the workers?

Figure 5: How supportive of the RPA implementation were your stakeholder groups?

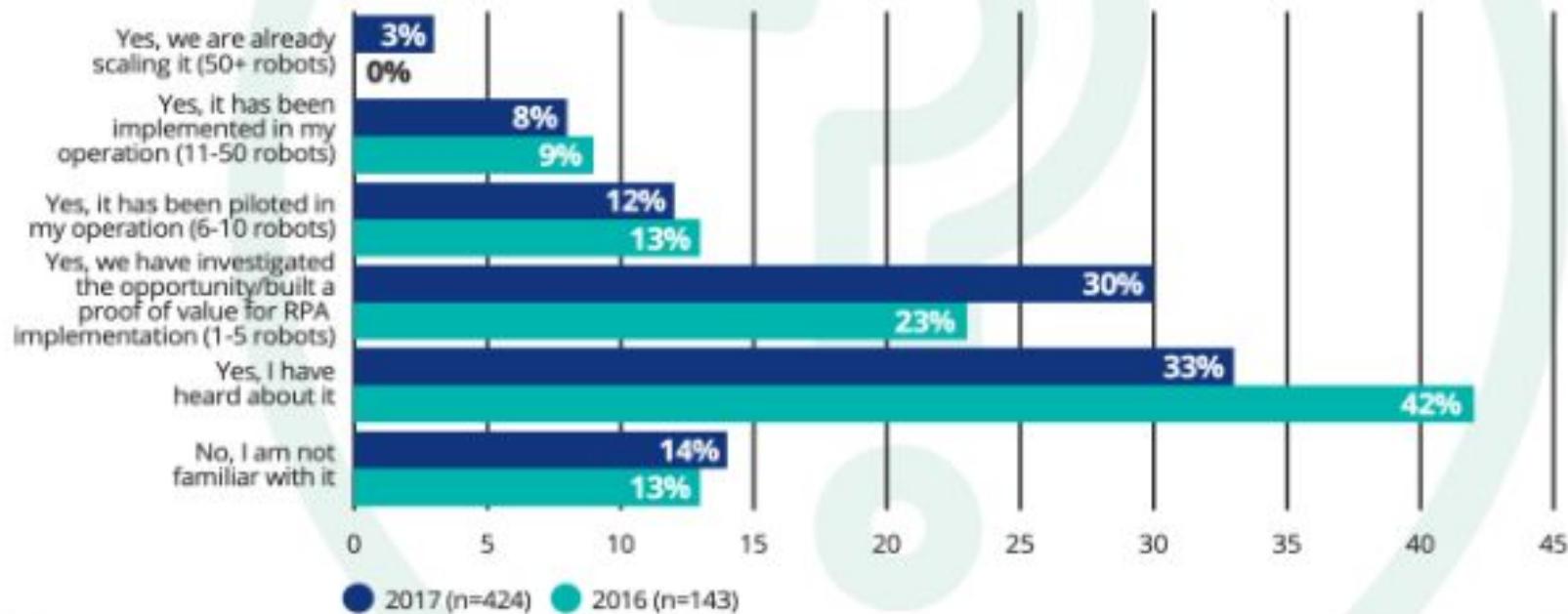


1. Deloitte. (2018). *The robots are ready. Are you?*
<https://www2.deloitte.com/content/dam/Deloitte/tr/Documents/technology/deloitte-robots-are-ready.pdf>

Introduction to Business Process Automation

Challenges of BPA – Hesitant to scale up

Figure 2: Are you familiar with RPA?



1. Deloitte. (2018). *The robots are ready. Are you?*
<https://www2.deloitte.com/content/dam/Deloitte/tr/Documents/technology/deloitte-robots-are-ready.pdf>

Challenges of BPA

- **Shortage**
 - Budget / Resource/ Talent
- **Concerns**
 - Risk of rushing / failing
 - Moral / Structural unemployment

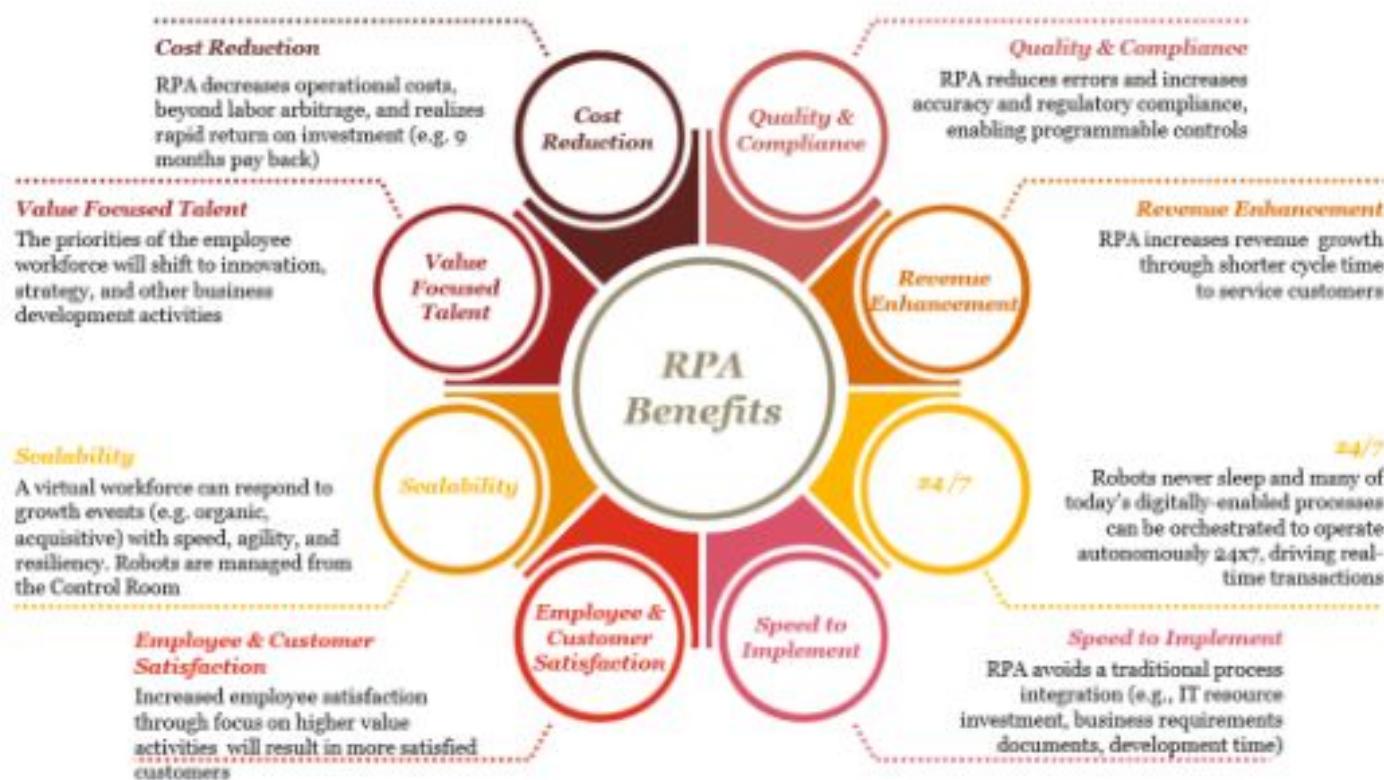


* Success depends on planning
(More on project management later)

Introduction to Business Process Automation

Implications of BPA – The Good

- Cost, Efficiency, Customer Experience



1. PwC Robotics Process Automation Solutions. PwC.

<https://www.pwchk.com/en/services/entrepreneurial-and-private-business/new-technology-digitalisation-and-transformation/process-automation-solutions.html>

Introduction to Business Process Automation

Implications of BPA – The Bad

- Automation complacency / dependency
- Risk of Technical Issues / Downtime

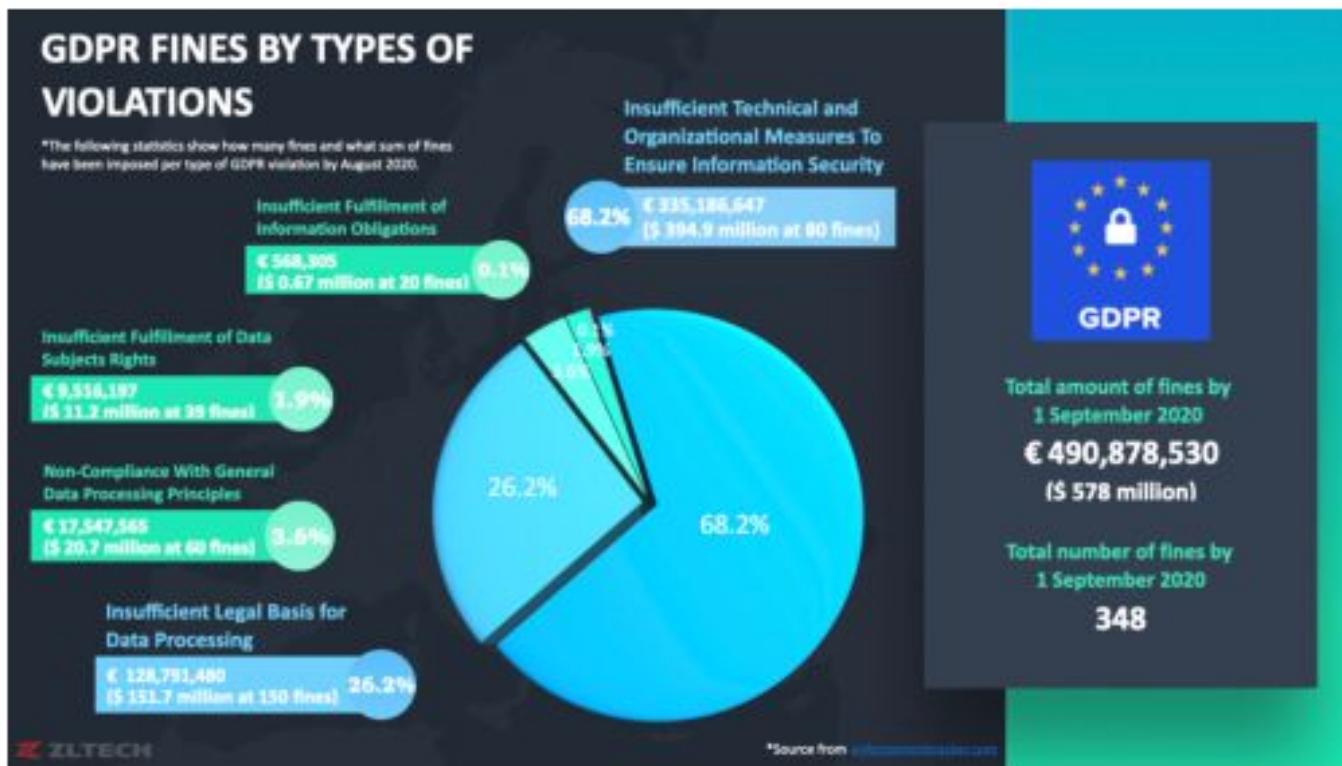


Image: Civil Aviation Safety Authority

Introduction to Business Process Automation

Implications of BPA – The Ugly

- Data Collection and Analysis¹
- Vulnerability to Cyberattacks

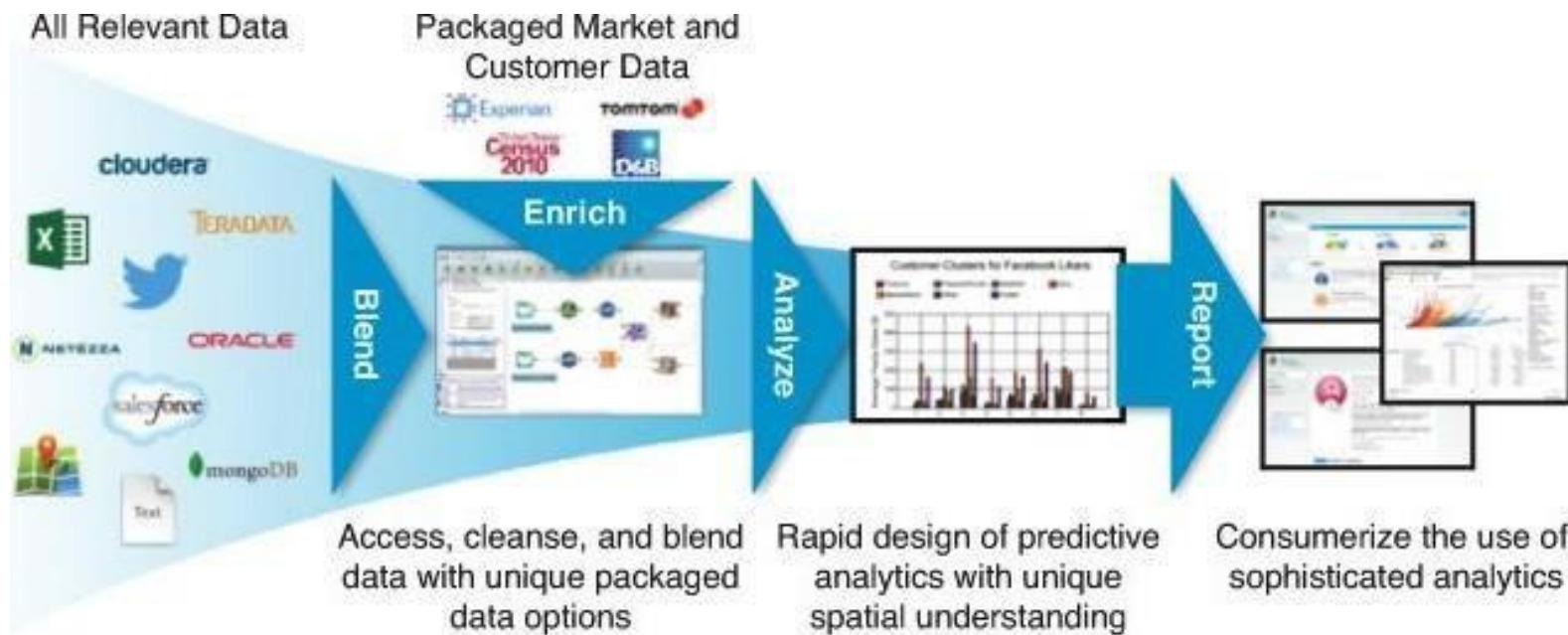


1. Chen, B. (2020, September 9). GDPR Recap: 28 Nations, 348 Fines, Half a Billion Euros. ZL Tech.
<https://www.zltech.com/blog/gdpr-recap-28-nations-343-fines-half-a-billion-euros/>

Modern BI Process - Blend, Enrich, Analyze, and Report

Predictive Analytics: By applying machine learning algorithms to historical project data, organizations can forecast outcomes more accurately and mitigate risks in project delivery

Business Intelligence: Automating the preparation of data for analytics tools allows organizations to gain actionable insights more rapidly, enhancing their decision-making capabilities.

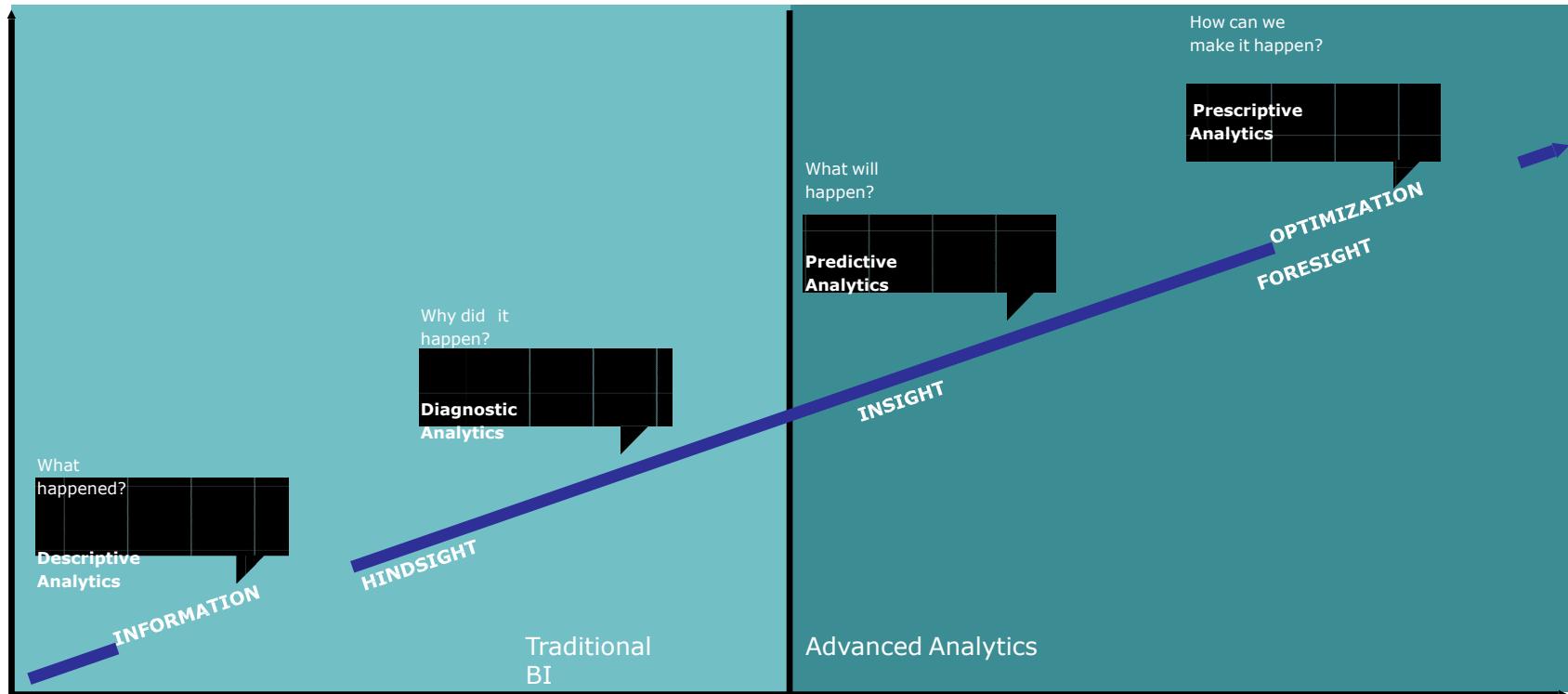


What Is Data Science?

From Wikipedia:

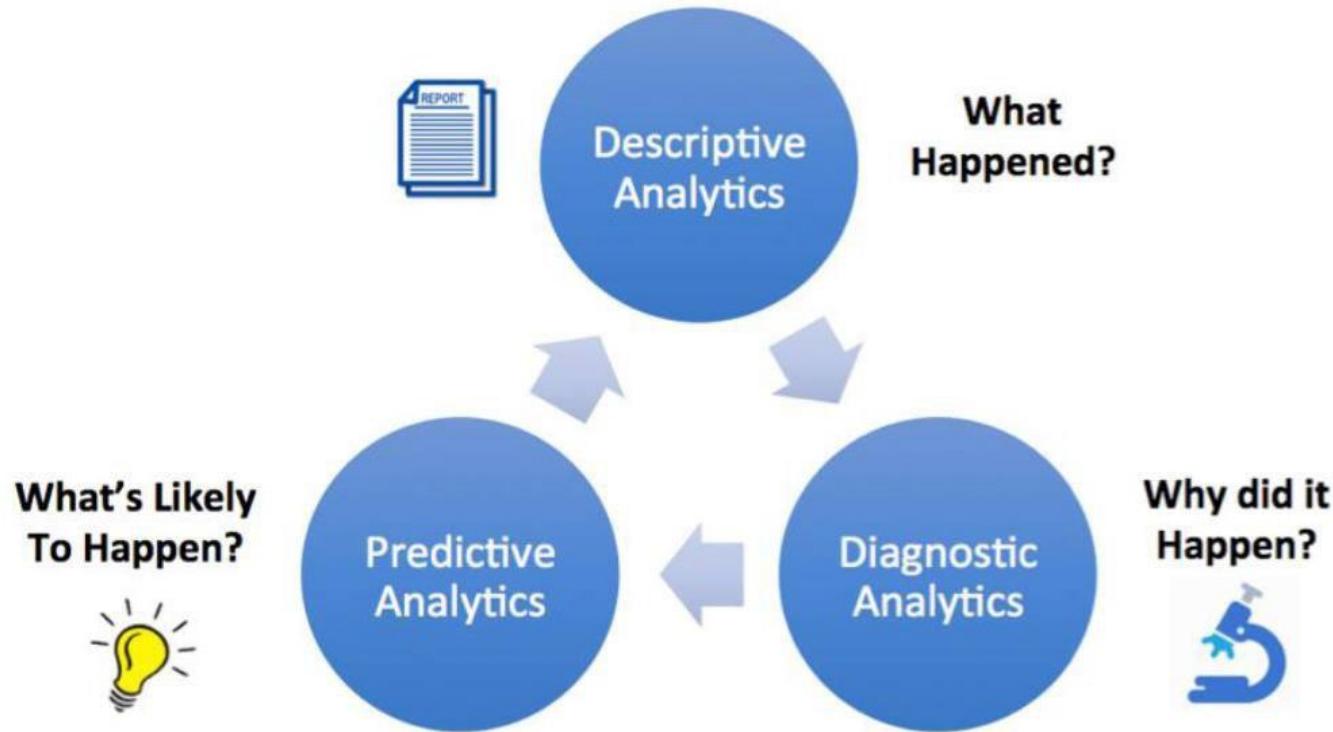
Data science, also known as data-driven science, is an interdisciplinary field about scientific processes and systems to extract knowledge or insights from data in various forms, either structured or unstructured, which is a continuation of some of the data analysis fields such as statistics, machine learning, data mining ,and predictive analytics similar to Knowledge Discovery in Databases (KDD).

How Is Business Intelligence Different From Data Science?



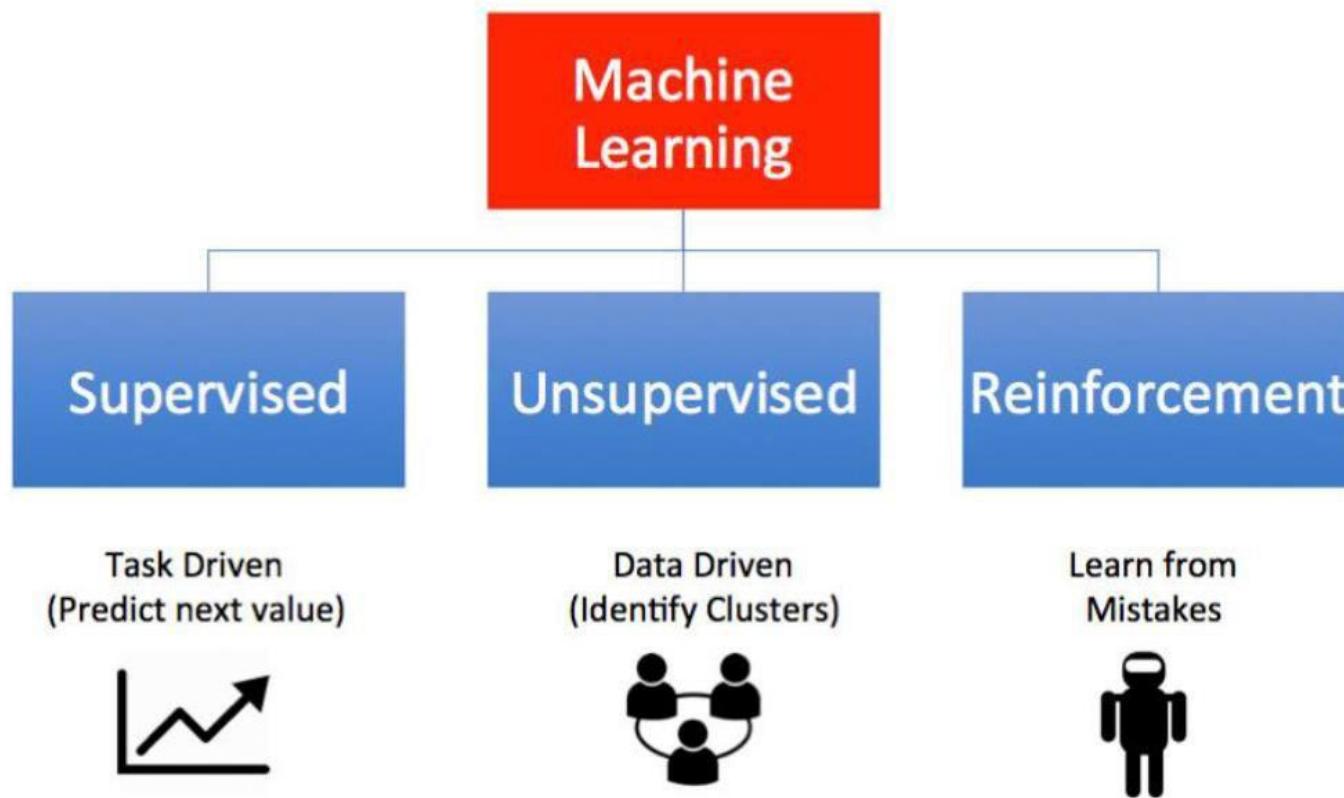
What is Business Analytics & Machine Learning?

Types of Analytics

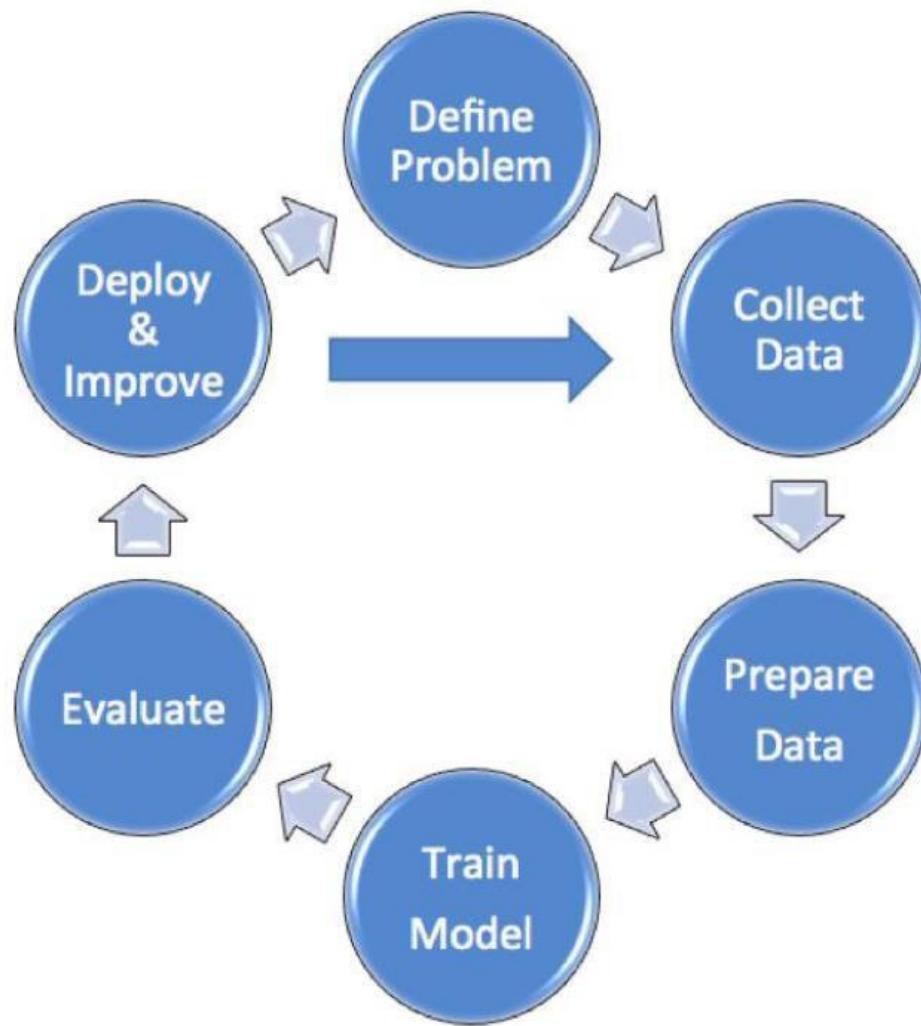


What is Business Analytics & Machine Learning?

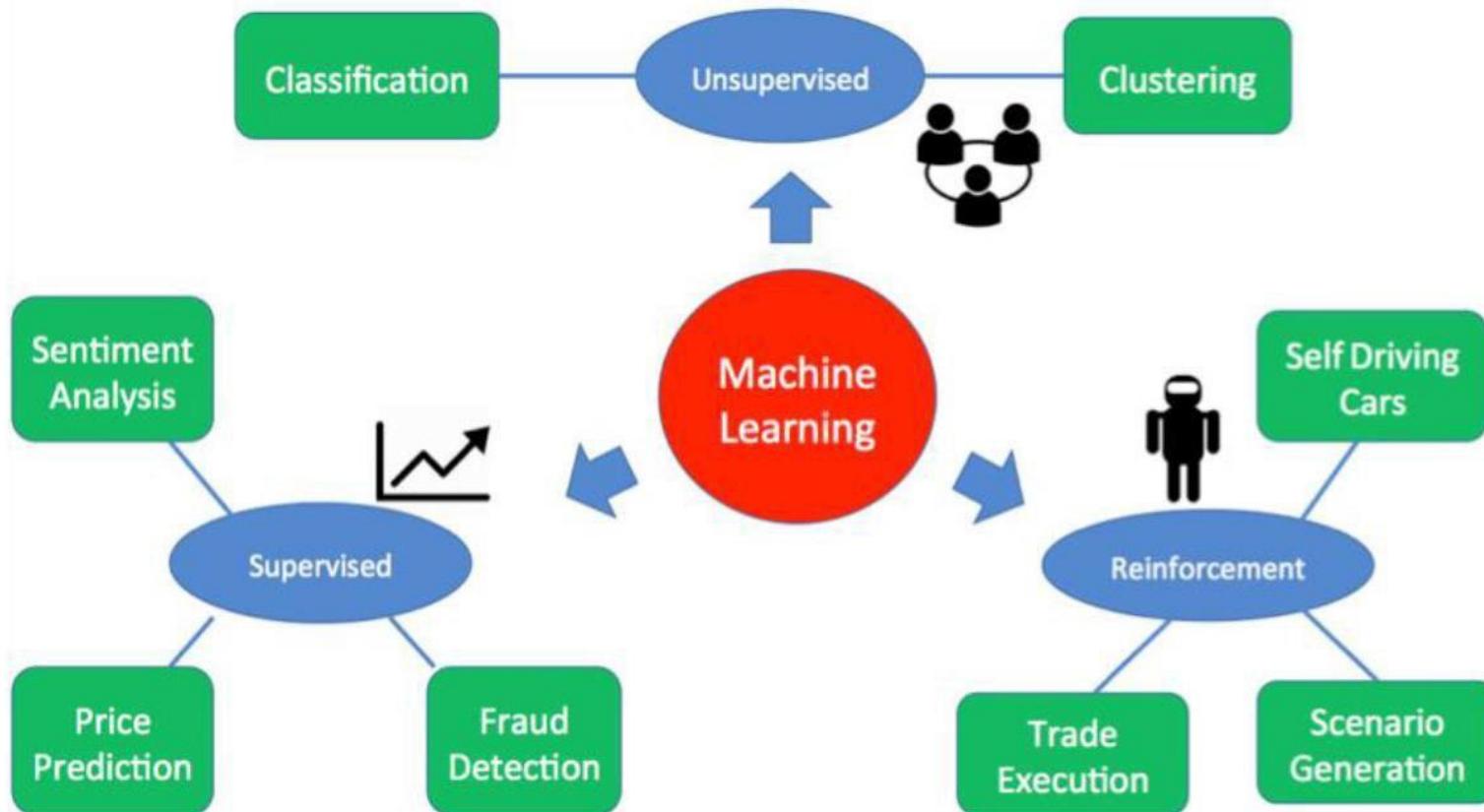
Types of Machine Learning



What is Business Analytics & Machine Learning?



What is Business Analytics & Machine Learning?



How To Do Data Mining: Machine Learning Techniques

Essential Concepts of Machine Learning

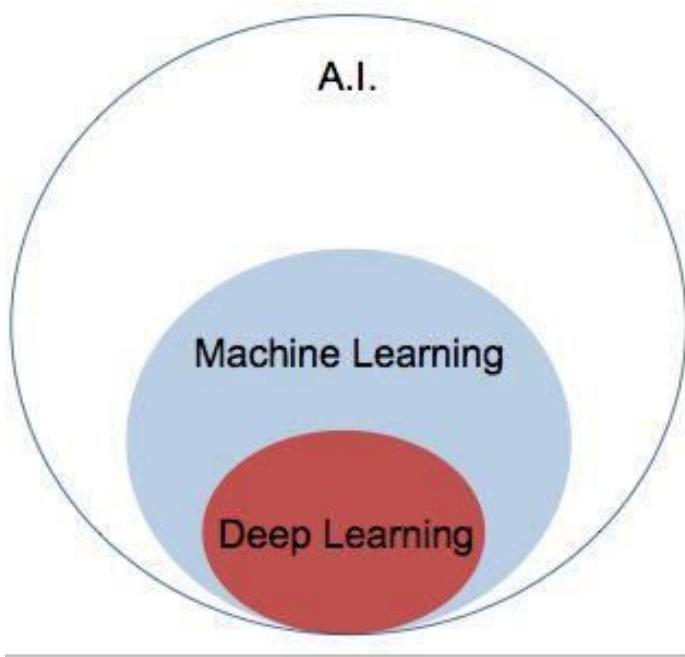
Artificial Intelligence

Machine Learning

Deep Learning

Source: <https://www.becomingadatascientist.com/2017/07/17/introductory-machine-learning-terminology-with-food/>

Clearing The Confusion About Data Science And Big Data



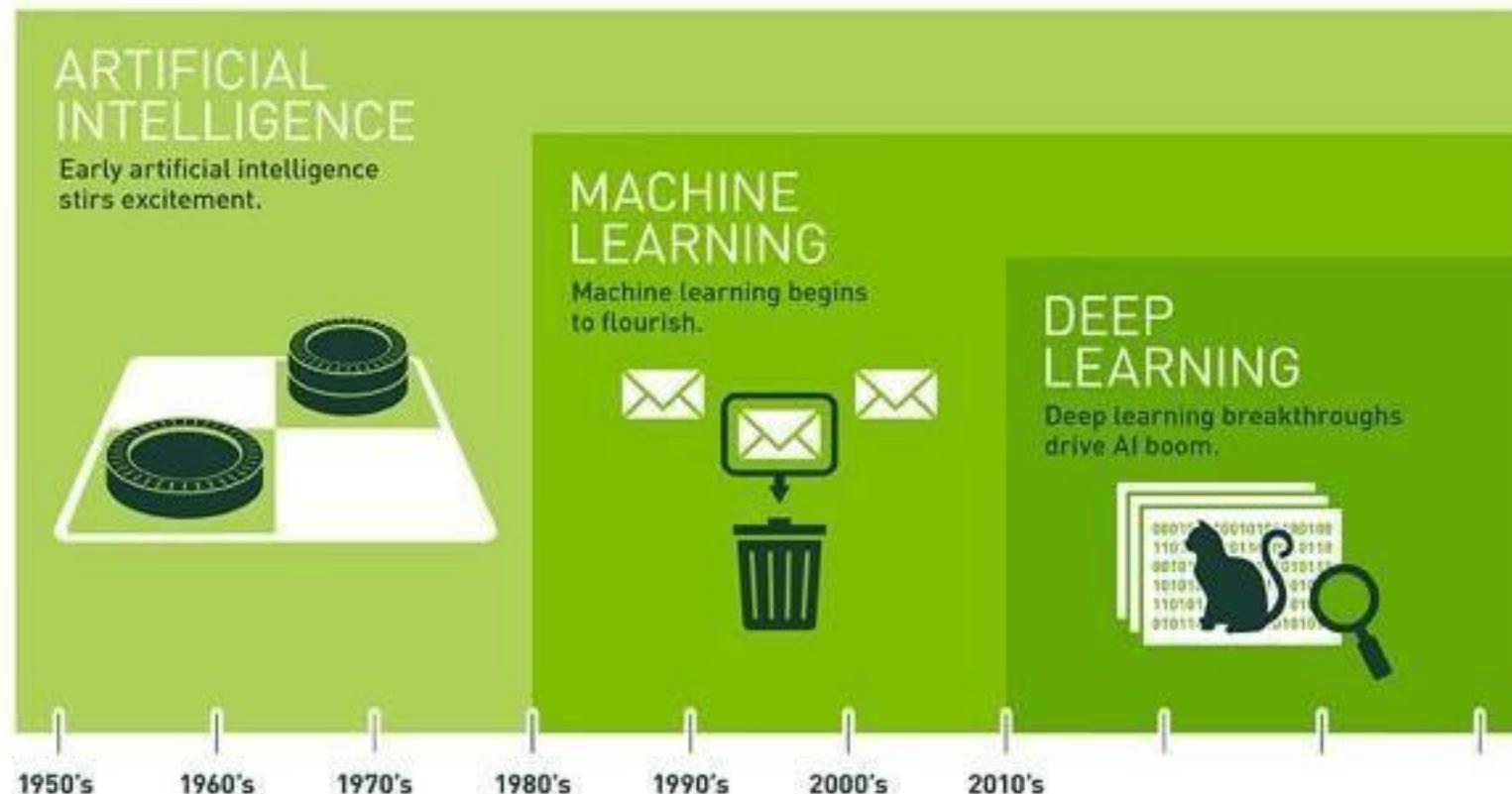
Definition of AI: a branch of computer science trying to get computers to exhibit intelligence.

- Machine Learning is a range of techniques for computers to perform cognitive functions
- Deep Learning (a type of Machine Learning that uses layers of neural networks) has been the most accurate and productive technique in AI research

Frank Chen of a 16z

Big Data Tools and Algorithm Exploration

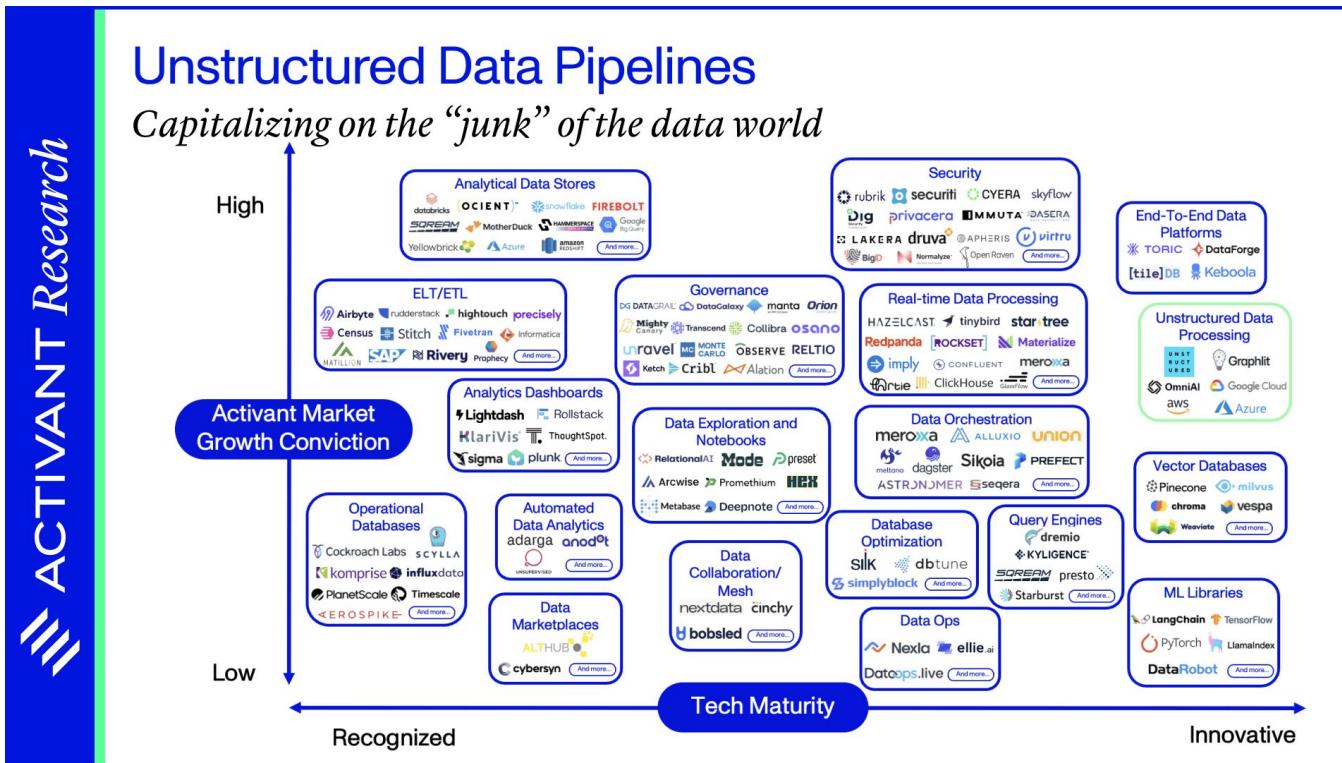
Difference Between AI, Machine Learning & Deep Learning



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

IPA – Data Science

IPA's integration with data science is particularly impactful in several areas:
Unstructured Data Processing: IPA can convert unstructured data into structured formats suitable for analysis, enabling businesses to leverage insights from previously untapped data source.



Python – Origin

- Released in 1991 (same as Visual Basic)
- Free and open-source



imgflip.com

Business Process Automation with Python

Motivation – Soft Limits

- Protective measures not matching modern days

| Feature | Maximum limit |
|---|----------------------------------|
| Total number of rows and columns on a worksheet | 1,048,576 rows by 16,384 columns |
| Column width | 255 characters |
| Row height | 409 points |
| Page breaks | 1,026 horizontal and vertical |

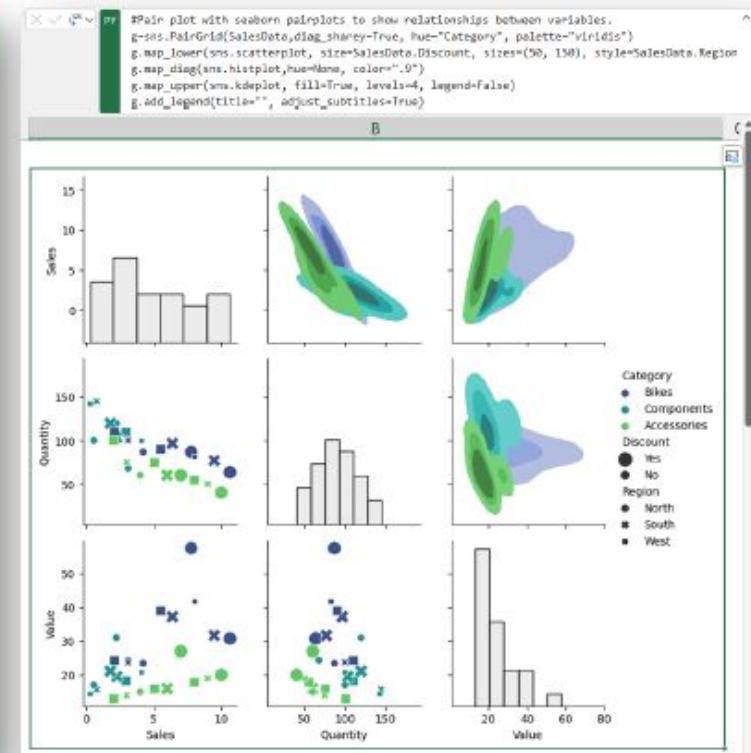
1. Microsoft. Excel Specifications and limits. <https://support.microsoft.com/en-us/office/excel-specifications-and-limits-1672b34d-7043-467e-8e27-269d656771c3>

Business Process Automation with Python

Motivation – Extensibility

- Python in Excel¹

The screenshot shows a Microsoft Excel spreadsheet titled "Python in Excel.xlsx". In the formula bar, the code "#Announcing Python in Excel!" is visible. Below it, the code `DataFrame=xl("A1:B10", headers=True)` and `DataFrame.groupby('Category').agg('mean')` is displayed. The spreadsheet contains a table of data with columns "Category" and "\$". To the right of the table, there are two data visualizations: a bar chart titled "Image" and a scatter plot titled "Image". The bar chart shows values for Components (\$20), Bikes (\$17), Accessories (\$9), Bikes (\$9), Clothing (\$8), Accessories (\$4), Clothing (\$4), Components (\$3), and Components (\$1). The scatter plot shows the same data points with different colors.



1. Announcing Python in Excel: Combining the power of Python and the flexibility of Excel. TECHCOMMUNITY.MICROSOFT.COM.
<https://techcommunity.microsoft.com/t5/excel-blog/announcing-python-in-excel-combining-the-power-of-python-and-the/ba-p/3893439>

Business Process Automation with Python

Motivation – Jobs

- Python → 1 of the top languages for work

Top Programming Languages 2024

Click a button to see a differently weighted ranking



1. *The Top Programming Languages 2024.* (2024 August 22). IEEE Spectrum.
<https://spectrum.ieee.org/top-programming-languages-2024>

Business Process Automation with Python

Motivation – Jobs

- Python → Recently added as part of CFA exams

Level II

- Python Programming Fundamentals**

A fundamentals course to demonstrate the basics of Python and how to use Jupyter Notebook for developing, presenting, and sharing data science projects related to finance. (if not taken at Level I)

- Analyst Skills**

Focuses on the skills equity and credit analysts need using insights gained from hundreds of successful analysts.

- Python, Data Science & AI**

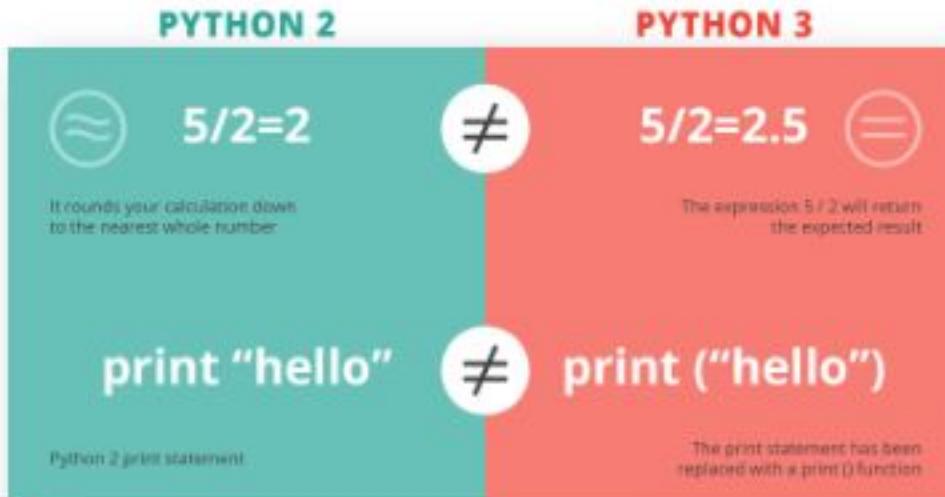
Introduces candidates to machine learning, artificial intelligence, and data science to understand financial statements, reporting, and analysis using Python.

1. *Practical Skills Module | CFA Program Evolution.* Practical Skills Module | CFA Program Evolution.
<https://evolve.cfainstitute.org/practical-skills-modules.html>

Business Process Automation with Python

Python runtime

- The language and basic runtime environment
- Just like Apple iOS, it keeps updating (e.g., Python 3.12.2)
 - Python 2 was no longer supported from 2020¹



1. Sunsetting Python 2. Python.org. <https://www.python.org/doc/sunset-python-2/>

Business Process Automation with Python

Installation – Python Runtime

- Official website
 - <https://www.python.org/downloads/>
- Installing in Windows → Tick “Add python.exe to PATH”

The screenshot shows the Python Downloads page for macOS. At the top, there's a navigation bar with links for About, Downloads, Documentation, and Community. Below that, a large button says "Download the latest version for macOS". Underneath this button is a yellow call-to-action button labeled "Download Python 3.11.3". To the right of the main content area, there's a sidebar with two main sections: "Install Now" and "Customize installation". The "Install Now" section shows the default installation path: "C:\Users\Jacki\AppData\Local\Programs\Python\Python312". It also includes a note that it "Includes IDLE, pip and documentation" and "Creates shortcuts and file associations". The "Customize installation" section allows users to "Choose location and features". Two checkboxes are present: one for "Use admin privileges when installing py.exe" (unchecked) and one for "Add python.exe to PATH" (checked). A blue arrow points to the checked "Add python.exe to PATH" checkbox.

python™

About Downloads Documentation Community

Download the latest version for macOS

Download Python 3.11.3

→ Install Now
C:\Users\Jacki\AppData\Local\Programs\Python\Python312

Includes IDLE, pip and documentation
Creates shortcuts and file associations

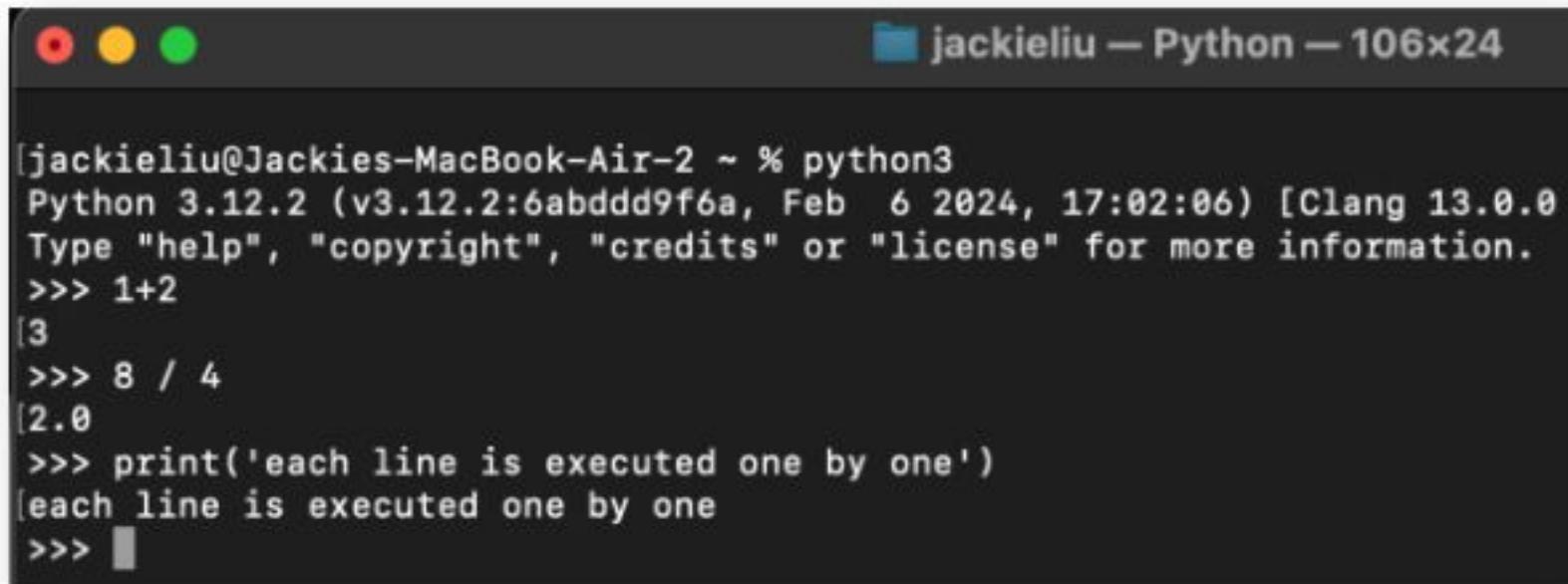
→ Customize installation
Choose location and features

Use admin privileges when installing py.exe

Add python.exe to PATH

How to run Python – In the old days

- Python interpreter
 - This is the core of Python
 - Input a line of codes → Press Enter → Run



```
[jackieliu@Jackies-MacBook-Air-2 ~ % python3
Python 3.12.2 (v3.12.2:6abddd9f6a, Feb  6 2024, 17:02:06) [Clang 13.0.0
Type "help", "copyright", "credits" or "license" for more information.
>>> 1+2
[3
>>> 8 / 4
[2.0
>>> print('each line is executed one by one')
each line is executed one by one
>>> ]
```

Business Process Automation with Python

How to run Python – Running as a file

- Python interpreter + Text editor
 - Write codes in a text editor (e.g., Notepad, Vim, Sublime Text)
 - Save codes as a “.py” file
 - Run all the lines in one go

The screenshot displays two windows side-by-side. On the left is a text editor window titled "test.py". The code inside is:

```
print(f'1 + 2 = {1 + 2}')
```

On the right is a "Command Prompt" window showing the output of running the script:

```
C:\Users\Jacki\Documents>python test.py
1 + 2 = 3

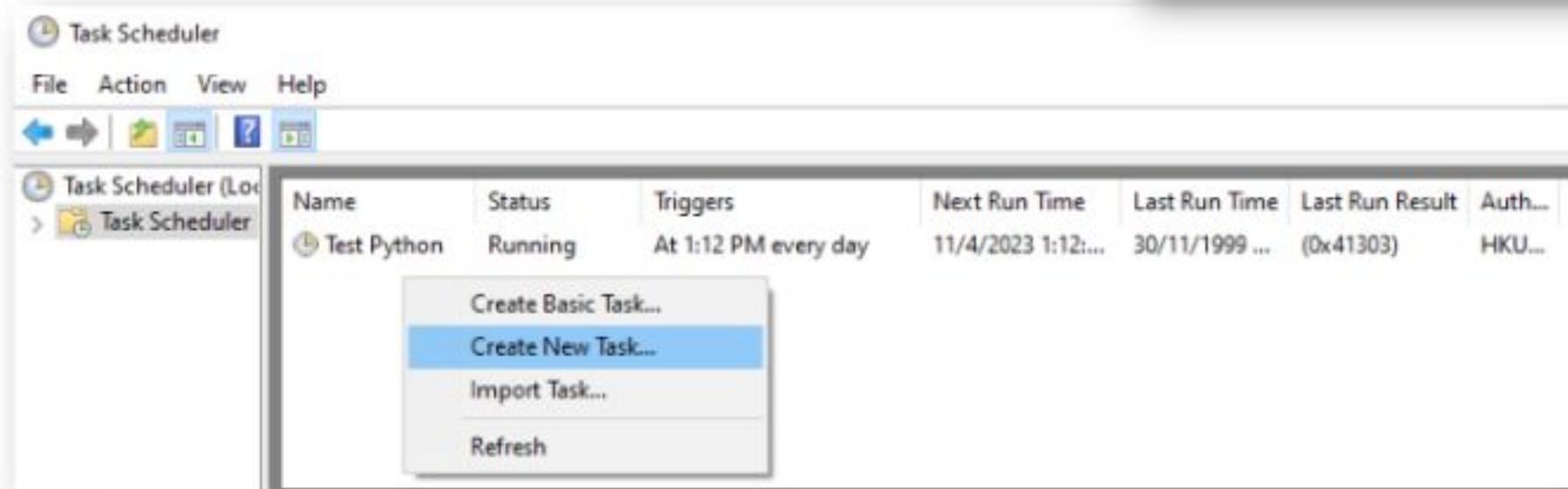
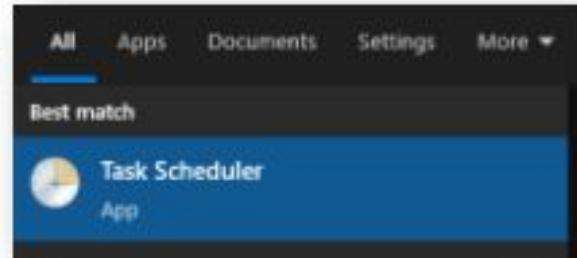
C:\Users\Jacki\Documents>
```

A yellow dashed circle highlights the "test.py" file tab in the text editor, and another yellow dashed circle highlights the command "python test.py" in the Command Prompt window.

Business Process Automation with Python

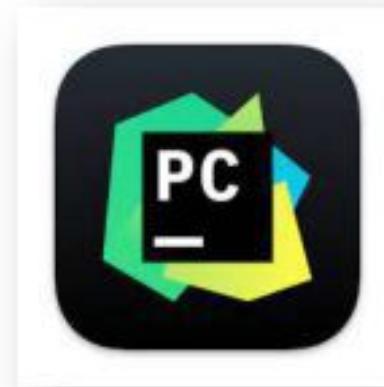
How to run Python – Running as a file (bonus)

- Python interpreter + Text editor + Task Scheduler
 - Running a Python script daily/ hourly
 - e.g., Task Scheduler/ Cron / Autosys etc
 - Create a task pointing to the .py file



How to run Python – IDE

- Integrated Development Environment (IDE)
 - A smart editor tailored for writing codes
 - e.g., Syntax highlighting, variable tracking, debugging mode
 - May support useful extensions/plugins (e.g., GitHub Copilot)
 - Common IDEs
 - e.g., Spyder, PyCharm, Visual Studio Code



Business Process Automation with Python

Example - PyCharm

The screenshot shows the PyCharm IDE interface with a Jupyter notebook open. The notebook contains Python code for data analysis, specifically focusing on the Ames Housing dataset. The code includes importing pandas, matplotlib.pyplot, and numpy, handling missing values, performing feature engineering by creating an 'AreaCategory' column, printing group statistics, performing statistical analysis with .describe(), and finally creating a scatter plot of SalePrice vs Br-Liv-Area. The scatter plot is color-coded by AreaCategory (small, medium, large) and shows a positive correlation.

```
import pandas as pd
import matplotlib.pyplot as plt
import numpy as np

df = pd.read_csv("/Users/Stanislav.Garkusha/Downloads/Shad_Python_81.2/Ames_dataset/AmesHousing.csv")

# Handle Missing Values
# Use apply function to apply a specific function across each column of the DataFrame
df = df.apply(lambda x: x.fillna(x.mean()) if x.dtype.kind in 'biufc' else x.fillna(x.mode()[0]))

# Feature Engineering
df["AreaCategory"] = pd.cut(df["Br-Liv-Area"], bins=[0, 1000, 2000, df["Br-Liv-Area"].max()], labels=["small", "medium", "large"], include_lowest=True)

print(df.groupby("AreaCategory")[
      "SalePrice"].mean()) # printing mean sales price for small, medium, and large living areas

# Statistical Analysis
print(df.describe()) # prints descriptive statistics of all numerical columns

# Data Visualization
fig, ax = plt.subplots()
ax.scatter(df["Br-Liv-Area"], df["SalePrice"], alpha=0.5)
ax.set_title('Scatter plot of Br-Liv-Area vs SalePrice')
ax.set_xlabel('Br-Liv-Area')
ax.set_ylabel('SalePrice')

# Scatter plot instead of boxplot
fig, ax = plt.subplots()
area_categories = ['small', 'medium', 'large']
for category in area_categories:
    ax.scatter(df[df['AreaCategory'] == category]['Br-Liv-Area'], df[df['AreaCategory'] == category]['SalePrice'])
```

Scatter plot of SalePrice grouped by AreaCategory

SalePrice

Br-Liv-Area

Data View

| Order | PID | MS SubClass | MS |
|-------|------|-------------|---------|
| 1553 | 1554 | 9102510... | 20 |
| 2903 | 2904 | 923125... | 20 |
| 942 | 943 | 9111030... | 50 |
| 727 | 728 | 9024771... | 30 |
| 726 | 727 | 9024771... | 30 |
| 1557 | 1558 | 9112260... | 30 |
| | | | C (all) |

How to run Python – Jupyter Notebook

- Google Colaboratory (Colab)
 - Jupyter Notebook powered by Google Cloud ¹
 - <https://colab.research.google.com/>
 - Benefits ²
 - “Zero configuration required”
 - Easy sharing
 - Free access: CPU & GPU



1. Google Colab FAQ. <https://research.google.com/colaboratory/faq.html>

2. Google Colaboratory. <https://colab.research.google.com>

Cloud Software Introduction

Colaboratory, or Colab for short, is a Google Research product, which allows developers to write and execute Python code through their browser. Google Colab is an excellent tool for deep learning tasks. It is a hosted Jupyter notebook that requires no setup and has an excellent free version, which gives free access to Google computing resources such as GPUs and TPUs.

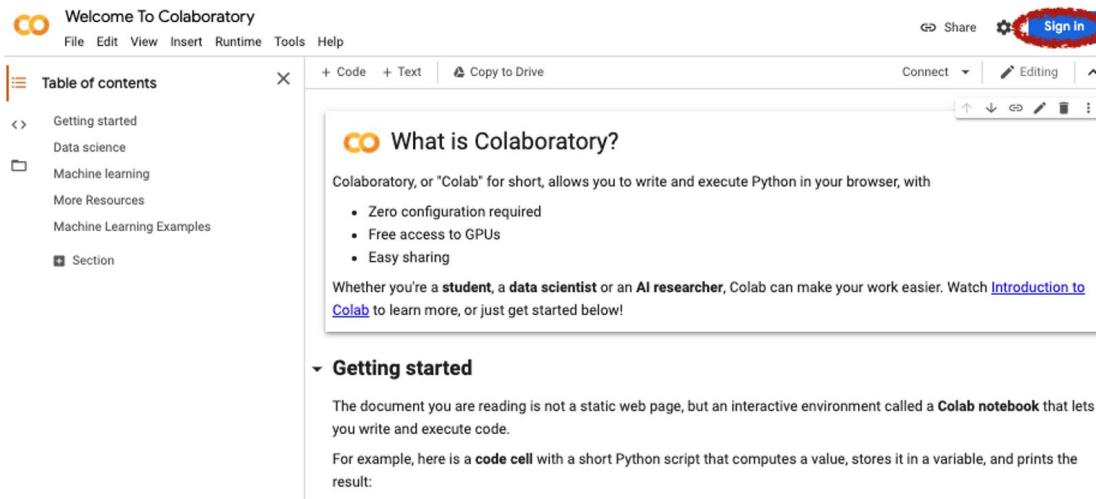
Why Should I Use Google Colab?

There are several reasons to opt to use Google Colab instead of a plain Jupyter Notebook instance:

- Pre-Installed Libraries
- Saved on the Cloud
- Collaboration
- Free GPU and TPU Use

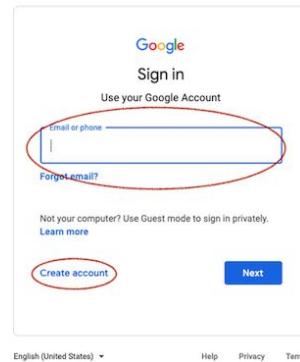
Google Colab Setup

1. Visit the [Google Colab](http://colab.research.google.com/) (<http://colab.research.google.com/>) page, which will direct you to the [Google Colaboratory Welcome Page](#).
2. Click the **Sign in** button on the right top.



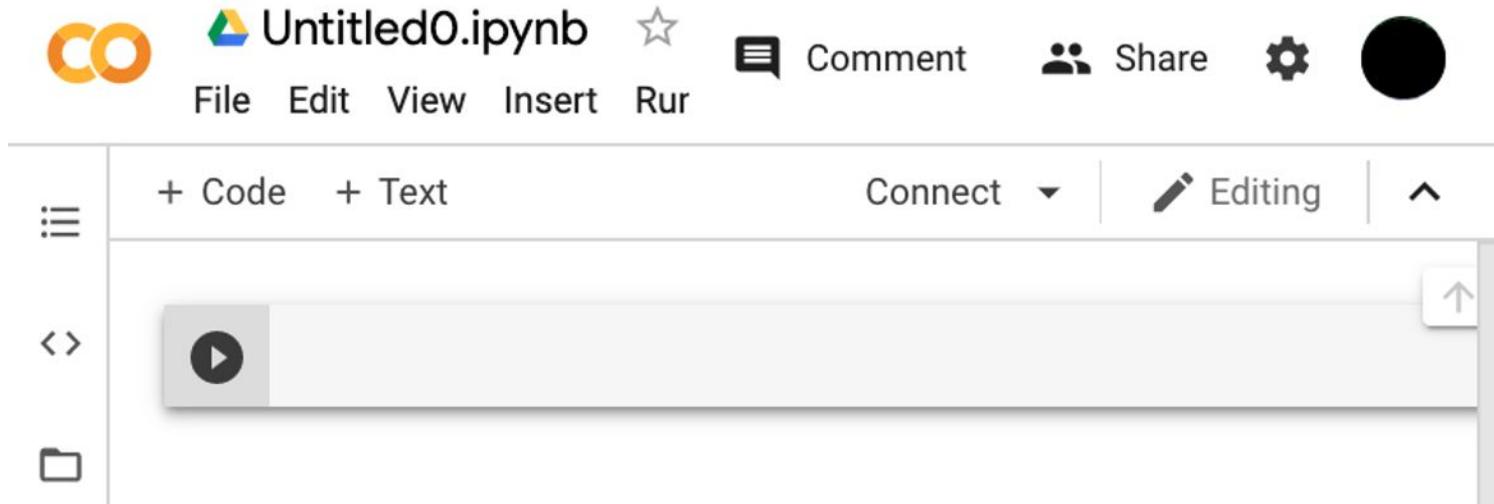
The screenshot shows the 'Welcome To Colaboratory' page. At the top, there's a navigation bar with 'File', 'Edit', 'View', 'Insert', 'Runtime', 'Tools', and 'Help'. On the far right of the bar is a 'Sign in' button, which is circled in red. Below the bar, there's a toolbar with icons for 'Code', 'Text', 'Copy to Drive', 'Connect', 'Editing', and other document-related functions. To the left, there's a 'Table of contents' sidebar with sections like 'Getting started', 'Data science', 'Machine learning', 'More Resources', 'Machine Learning Examples', and 'Section'. The main content area has a heading 'What is Colaboratory?' followed by a description and a bulleted list: 'Zero configuration required', 'Free access to GPUs', and 'Easy sharing'. It also mentions that Colab can make work easier for students, data scientists, and AI researchers. Below this, there's a section titled 'Getting started' with a description of what a Colab notebook is and an example of a code cell.

3. Sign in with your Gmail account.



Google Colab Setup

4. As soon as you complete the sign-in process, you are ready to use Google Colab.
5. You may easily create a new Colab notebook on this page by clicking *File> New notebook*.



How to run Python – Summary

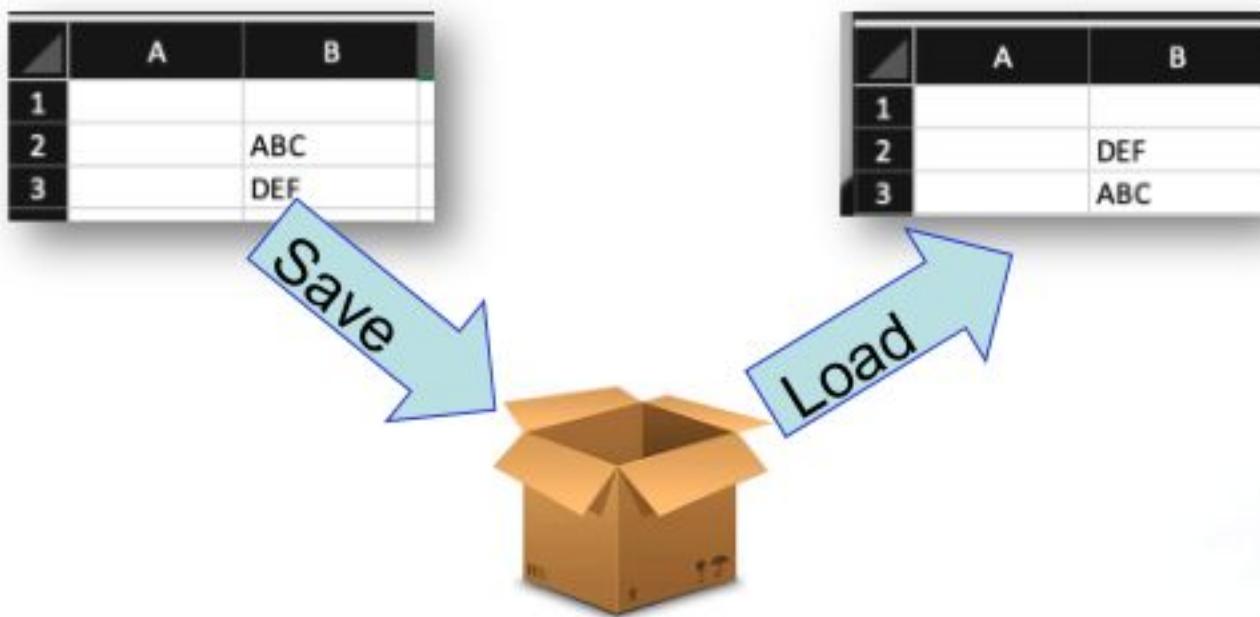
- Python interpreter
- Python (.py) files with
 - Text editors
 - IDEs
- Jupyter Notebook running
 - Locally
 - Remotely



Business Process Automation with Python

Variables

- From the previous lecture
 - Variables → Storage of values
 - Values → carry data types



Variables

- Declaration
 - In VBA, declarations (“dim”) are recommended
 - Dim year As **Integer**
 - year = **2024**
- In Python, variables are created during **assignment**
 - year = **2024**
- To assign without an actual value
 - year = **None**



Business Process Automation with Python

Value Check

- VBA
 - Debug.Print()

```
Sub HighlightCell()
    If Range("B1").Value > 50 Then
        Range("B1").Interior.Color = vbYellow
    Else
        Range("B1").Interior.Color = vbGreen
    End If
    Debug.Print (Range("B1"))
End Sub
```

40

- Python
 - print()
 - Last line in a block

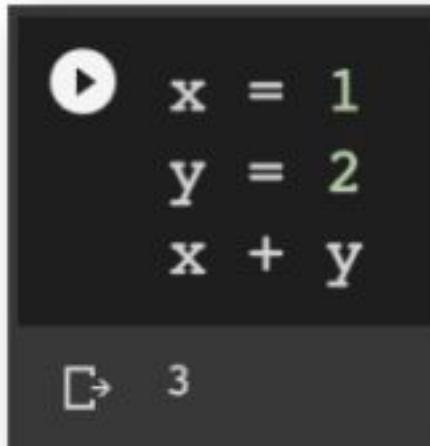
```
▶ print(1 + 2)          # This gives 3
  print(1 + 2 + 3)      # This gives 6
□ 3
  6
```

```
▶ 1 + 2          # This is not shown
  1 + 2 + 3      # This is shown
□ 6
```

Business Process Automation with Python

Primitive Data Types - Numeric

- Int

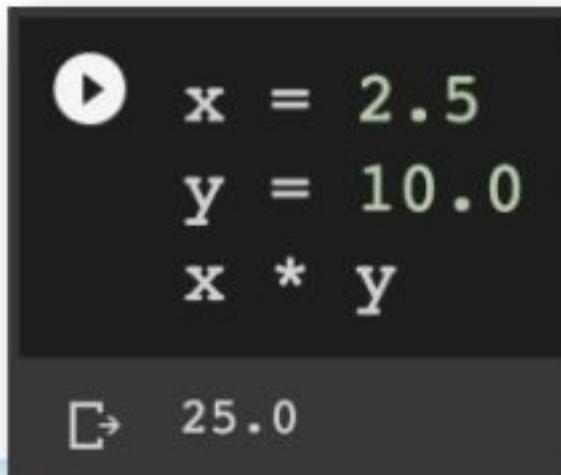


```
x = 1  
y = 2  
x + y
```

[Out]: 3

A screenshot of a Jupyter Notebook cell. The cell contains three lines of code: 'x = 1', 'y = 2', and 'x + y'. Below the code, a dark grey bar indicates the output, which is '3'.

- Float



```
x = 2.5  
y = 10.0  
x * y
```

[Out]: 25.0

A screenshot of a Jupyter Notebook cell. The cell contains three lines of code: 'x = 2.5', 'y = 10.0', and 'x * y'. Below the code, a dark grey bar indicates the output, which is '25.0'.

Primitive Data Types - Numeric

- Complex (For scientific calculations)

THE QUADRATIC FORMULA

© CHILIMATH.COM

If $ax^2 + bx + c = 0$ but $a \neq 0$

then

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

DISCRIMINANT

- $b^2 - 4ac > 0$ two real solutions
- $b^2 - 4ac = 0$ one real solutions
- $b^2 - 4ac < 0$ zero real solutions

▶ $x = 3 + 4j$
 $y = 2 + 2j$
 $x - y$
⇒ $(1+2j)$

$i = \sqrt{-1}$

Primitive Data Types - Numeric

- Common Functions
 - `round()` / `pow()` / `abs()`

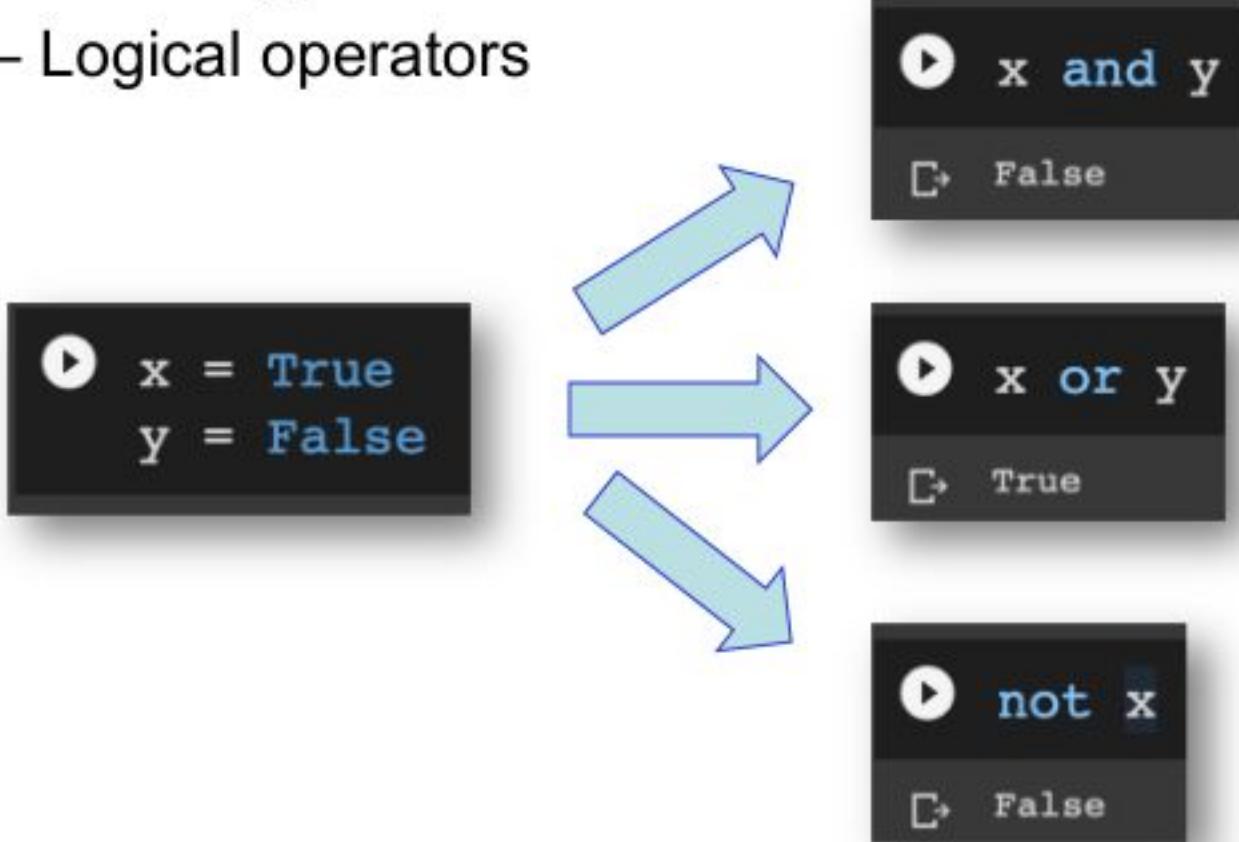
```
▶ print(round(9876.54321, 2))      # Round to 2 decimal places
    print(round(9876.54321, 0))      # Round to nearest integer
    print(pow(2, 3))                  # 2 * 2 * 2 = 8
    print(abs(-5))                   # Negative numbers become positive
```

```
▷ 9876.54
  9877.0
  8
  5
```

Business Process Automation with Python

Primitive Data Types - Boolean

- Bool – Logical operators



Business Process Automation with Python

Primitive Data Types - Boolean

- Bool – comparators
 - Generates True/ False

| Operator | Name | Example |
|--------------------|--------------------------|--------------------------|
| <code>==</code> | Equal | <code>5 == 5</code> |
| <code>!=</code> | Not equal | <code>26 != 3</code> |
| <code>></code> | Greater than | <code>100 > 67</code> |
| <code><</code> | Less than | <code>89 < 216</code> |
| <code>>=</code> | Greater than or equal to | <code>90 >= 54</code> |
| <code><=</code> | Less than or equal to | <code>23 <= 77</code> |

Primitive Data Types - Text

- String
 - Representation

```
▶ x = 'This is a string'                      # Single quote
    y = "This is also a string"                 # Double quote
    z = '''This is a very long string
spanning across more than 1 line'''           # Multi-line

    print(x)
    print(y)
    print(z)

⇒ This is a string
    This is also a string
    This is a very long string
        spanning across more than 1 line
```

Primitive Data Types - Text

- String
 - Concatenation → “+” → glue 2 strings together

```
▶ x = 'This is a string'  
    y = "This is also a string"
```

```
▶ print(x + ' and ' + y) # Adding 2 strings together  
◀ This is a string and This is also a string
```

Business Process Automation with Python

Primitive Data Types - Text

- String
 - Concatenation → Number vs Text



```
lobster_price = 936
lobster_text = str(lobster_price)
print(2 * lobster_price) # 1872
print(2 * lobster_text) # 936936
```

1872
936936

```
▶ print(10 * '=')
print('WELCOME')
print(10 * '=')
```

```
=====
WELCOME
=====
```

- 星島日報. (2023, May 6). 荃灣中菜館驚現「天價龍蝦」兩隻竟索價90萬元 酒樓親解釋. Singtaousa.com; 星島日報.
<https://www.singtaousa.com/2023-05-06/%e8%bd%83%e7%81%a3%e4%b8%ad%e8%8f%9c%e9%a4%a8%e9%a9%9a%e7%8f%be%e3%80%8c%e5%a4%a9%e5%83%b9%e9%be%8d%e8%9d%a6%e3%80%8d-%e5%85%a9%e9%9a%bb%e7%ab%9f%e7%b4%a2%e5%83%b990%e8%90%ac%e5%85%83-%e9%85%92/4488305>

Primitive Data Types - String formatting

- F-String (Python 3.6 or newer)
 - A smart way to format strings

```
▶ x = 4
  y = 8.8
  f'x is {x} and y is {y}; y divided by x is {y / x}'
⇒ 'x is 4 and y is 8.8; y divided by x is 2.2'
```

1. *Input and Output.* Python Documentation. <https://docs.python.org/3/tutorial/inputoutput.html>

Primitive Data Types - String formatting

- F-String (Python 3.6 or newer)
 - A smart way to format strings (advanced)

```
▶ z = 1234.56789
    print(z)
    print(f'{z:.2f}')
    print(f'{z:,.2f}')

⇨ 1234.56789
    1234.57
    1,234.57
```

1. *Input and Output.* Python Documentation. <https://docs.python.org/3/tutorial/inputoutput.html>

Primitive Data Types - String

- Common Functions
 - upper() / lower() / replace()

```
▶ # Common functions
    test_string = 'Hello Hong Kong!'
    print(test_string.upper())
    print(test_string.lower())
    print(test_string.replace('Hello', 'Bello'))
```



```
⇨ HELLO HONG KONG!
    hello hong kong!
    Bello Hong Kong!
```

Primitive Data Types - Conversion

- **VBA**
 - CInt() / CDbl() / Format()
- **Python**
 - int() / float() / str() / bool()

```
▶ int(4.5)
□ 4
```



```
▶ float(5)
□ 5.0
```



```
▶ str(5.55)
□ '5.55'
```

⚠️

```
▶ bool('Some values')
□ True
```



```
▶ bool('')
□ False
```

Methods

- VBA
 - Sub-routine: reusable code blocks
 - Function: sub-routine which gives an output
- Python
 - Methods: Use “`return`” if output is required

VBA

```
Function AddNumbers(x As Double, y As Double) As Double
    AddNumbers = x + y
End Function
```

```
?AddNumbers(1,2)
3
```

Python

```
def add_numbers(x, y):
    return x + y
```

```
def add_numbers(x, y):
    return x + y

print(add_numbers(1,2))
```

Business Process Automation with Python

Practice 1 – Market Capitalization

- Compute market capitalization of 2 companies
- Show the difference between the two
- Format the result up to 2 decimal places



- **Starting notebook:**

https://github.com/innoviai/IPA_Courses-Sept2025-/blob/main/practice1_simple_calc_202409.ipynb

Data Processing & Analysis

Decision Trees: Ensemble Linear Regression, Decision Tree

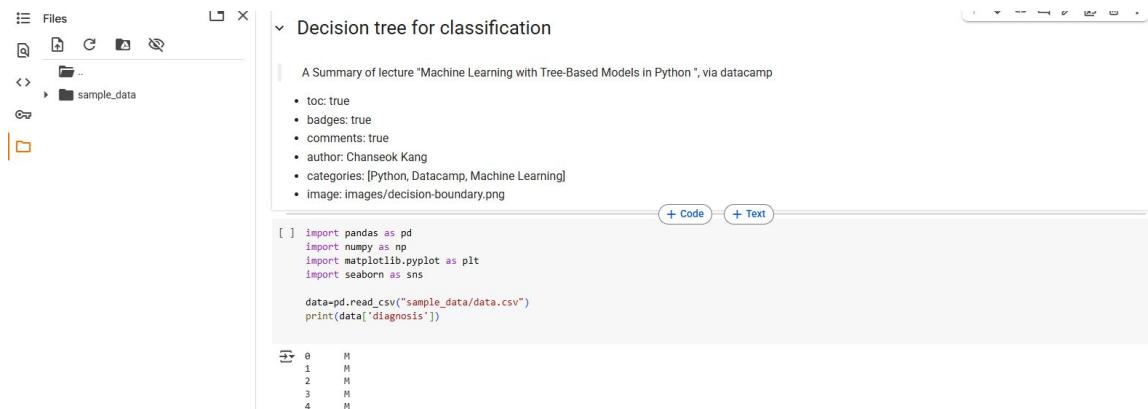
Practice 2 – Run machine learning model

Colab Lab Materials :

<https://colab.research.google.com/drive/1T59haRQLd3KSL6rSa4CKFXG2AyRflcIP>

Data :

https://github.com/innoviai/IPA_Courses-July2025-/blob/main/data.csv



The screenshot shows the Google Colab interface. On the left is a sidebar with 'Files' and a folder named 'sample_data'. The main area has a title 'Decision tree for classification' and a summary from DataCamp. Below it is a code cell containing Python code to read a CSV file and print its first column. At the bottom, there's a preview of the data showing five rows: 0, M; 1, M; 2, M; 3, M; 4, M.

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

data=pd.read_csv("sample_data/data.csv")
print(data['diagnosis'])
```

| | diagnosis |
|---|-----------|
| 0 | M |
| 1 | M |
| 2 | M |
| 3 | M |
| 4 | M |

How to Fit a Decision Tree Model using Scikit-Learn

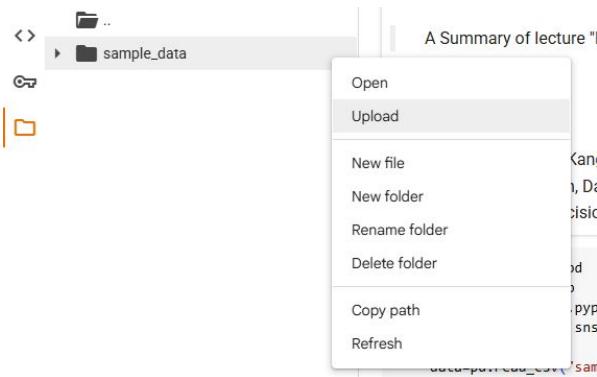
<https://www.kdnuggets.com/2020/04/visualizing-decision-trees-python.html>

Data Processing & Analysis

Decision Trees: Ensemble Linear Regression, Decision Tree

Practice 2 – Run machine learning model

1. Open Colab Lab Materials at
<https://colab.research.google.com/drive/1T59haRQLd3KSL6rSa4CKFXG2AyRflcIP>
2. Click File -> Save a copy in Drive
3. Download the data at
https://github.com/innoviai/IPA_Courses-July2025/blob/main/data.csv and save in C:\Users\user\data.csv
4. In colab, right click the sample files and click “Upload” ->select the data.csv

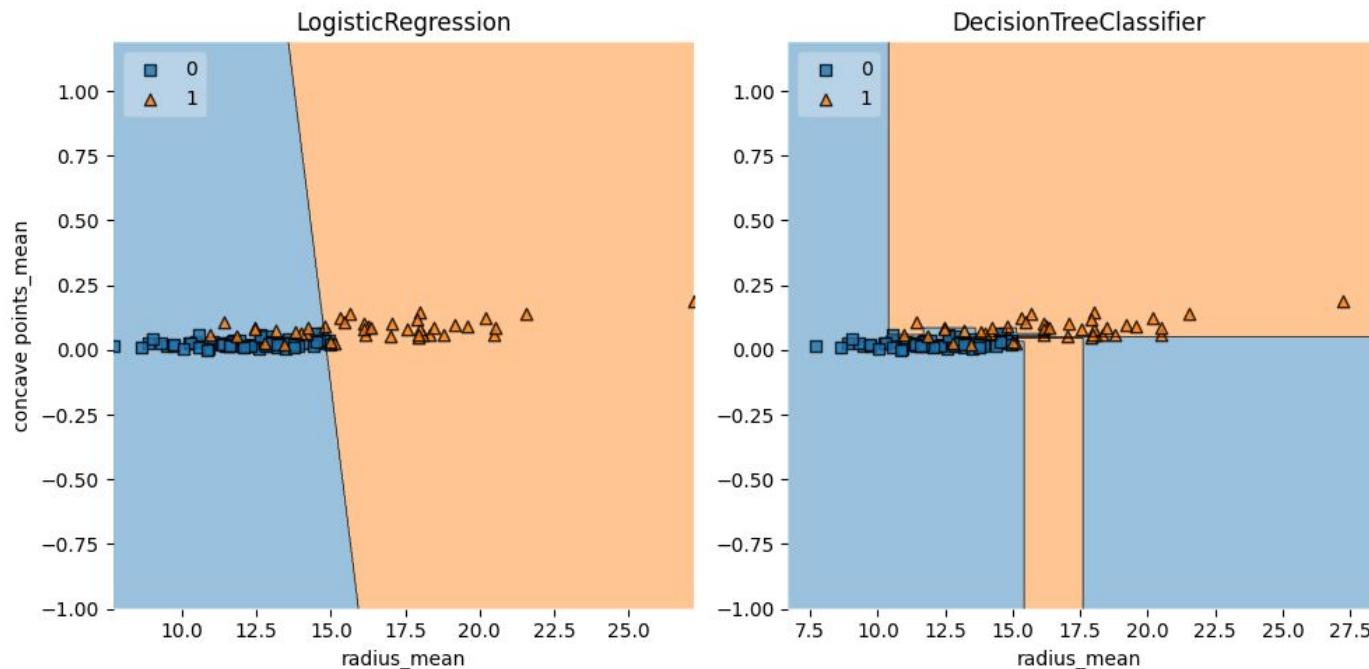


Data Processing & Analysis

Decision Trees: Ensemble Linear Regression, Decision Tree

Practice 2 – Run machine learning model

1. `X = wbc[['radius_mean', 'concave points_mean']]`
2. `y = wbc['diagnosis']`
3. `y = y.map({'M':1, 'B':0})`
4. `X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=1)`



How To Do Data Mining: Machine Learning Techniques

Artificial Intelligence

Artificial Intelligence (AI) is the science of making things smart. Can be defined as:

“Human intelligence exhibited by machines”

A broad term for getting computers to perform human tasks. The scope of AI is disputed and constantly changing over time. Let's go deeper...

How To Do Data Mining: Machine Learning Techniques

Machine Learning

An approach to achieve artificial intelligence through ‘software’ systems that can **learn** from **experience** to find **patterns** in a **set of data**.

Write a computer program
with **explicit rules** to follow

```
if email contains V!agra  
  then mark is-spam;  
if email contains ...  
if email contains ...
```

Write a computer program
to **learn from examples**

```
try to classify some emails;  
change self to reduce errors;  
repeat;
```

Traditional Programming

Machine Learning Programs

How To Do Data Mining: Machine Learning Techniques

Gathering Data



Inspiring Your Future

上環/中環/金鐘 寶泰大廈 [室內相]
售 \$530 萬 月供 \$13,711 按揭計算機»

單位詳情

成交紀錄

屋苑資訊

附近學校



| | |
|------|---------------------------------|
| 樓盤 | 上環/中環/金鐘 寶泰大廈 |
| 樓盤編號 | LCX458-34676 |
| 地址 | 水坑口街9號 <small>中原地圖»</small> |
| 價格 | 售 \$530 萬 |
| 實用面積 | 357呎 (\$14,845/呎) |
| 建築面積 | 519呎 (\$10,211/呎) |
| 實用率 | 69% |
| 樓層 | 低層 (B室) |
| 樓齡 | 33年 |
| 單位類型 | 分層單位 |
| 輟學特色 | <small>小學: 11» 中學: 中西區»</small> |

Business Education @ HKUSPACE

Source: <http://www.becomingdatascientist.com/2017/07/17/introductory-machine-learning-terminology-with-food/>

How To Do Data Mining: Machine Learning Techniques

Features / Attributes

Are used to train a Machine Learning algorithm

Facts and Features

| | | | | | |
|--|-----------------------|--|---------------------|--|-------------------|
| | Type Single Family | | Year Built 1989 | | Heating Other |
| | Cooling No Data | | Parking 2 spaces | | Lot 6,300 sqft |

INTERIOR FEATURES

| | |
|-----------------|------------------------|
| Bedrooms | Flooring |
| Beds: 6 | Floor size: 3,138 sqft |



Using a Home Valuation Prediction Model
as an example

How To Do Data Mining: Machine Learning Techniques

Features / Attributes

Are used to train a Machine Learning algorithm

Sales Price

Saleable

Area

Number of Bedrooms

Year built

(Refurbished) Schools
in the area

Transportation (MTR, buses)

Address: 270-280 QUEEN'S ROAD
WEST
Price per sqft.: 17,841
Qty transaction: 2
Year: 2017



Using a Home Valuation Prediction Model as an example

How To Do Data Mining: Machine Learning Techniques

Machine Learning

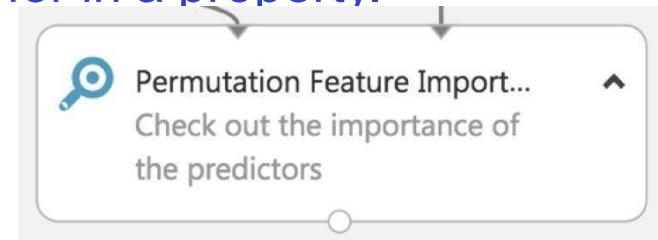
| Feature | Score |
|----------------|----------|
| | 1. |
| SaleableArea | 0.858545 |
| Eff_Build_Ages | 0.166877 |
| Bus_Route | 0.076519 |
| Parks | 0.019011 |

Choosing a ‘useful’ or **significant** feature (predictors) **makes** or **breaks** a data science project.

Domain experts, property agents will tell you what property buyers value in a home.

Data scientists would **NOT** have a clue what home buyers look for in a property.

But not all is lost!



Data Processing & Analysis - What Is Machine Learning?

Supervised Learning

Used as an advanced form of *predictive modeling*

Each observation must be labeled with a "correct answer"

Only then can you build a predictive model because you must tell the algorithm what's "correct" while training it (hence, "supervising" it)

Regression is the task for modeling continuous target variables

Classification is the task for modeling categorical (a.k.a. "class") target variables



Unsupervised Learning

Used either as a form of automated data analysis or automated signal extraction

Unlabeled data has no predetermined "correct answer"

You'll allow the algorithm to directly learn patterns from the data (without "supervision")

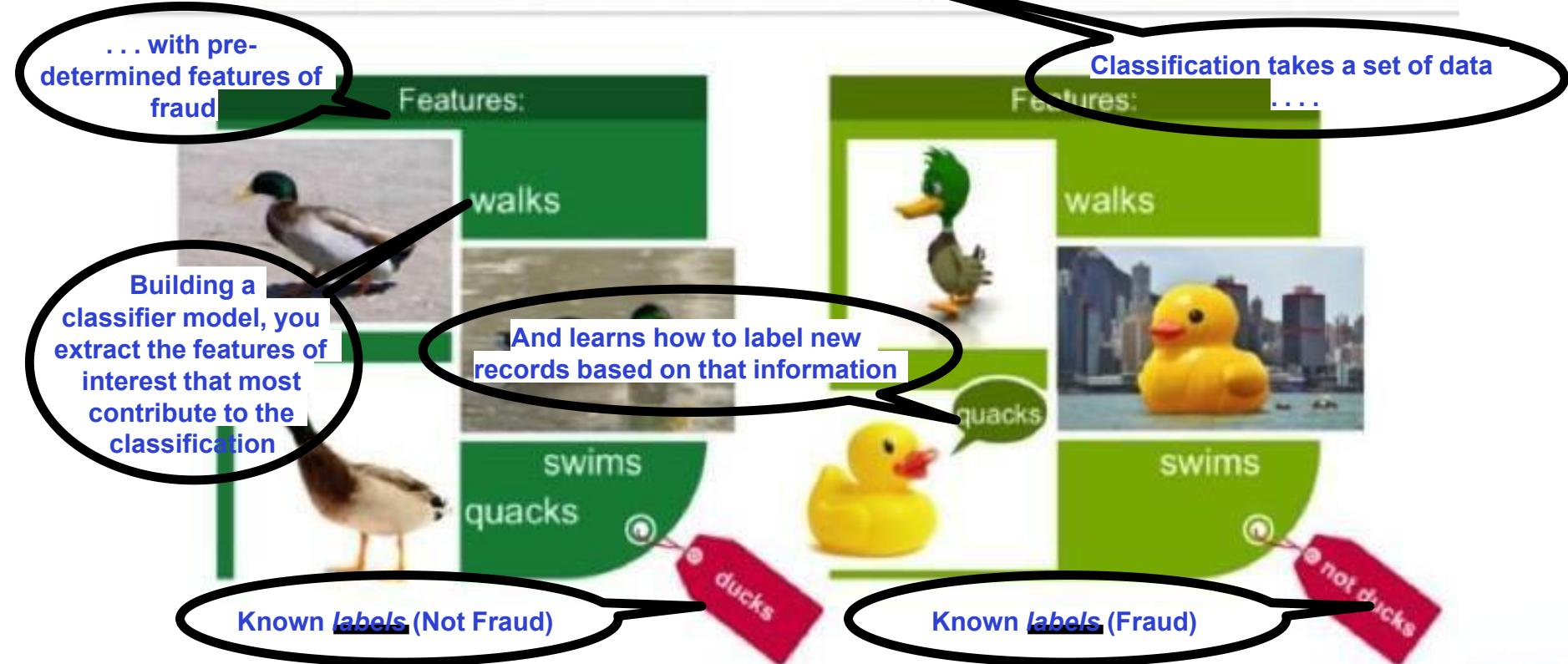
Clustering is the most common unsupervised learning task, and it's for *finding groups within your data*



Data Processing & Analysis

Supervised Learning: Classification For Debit Card Fraud

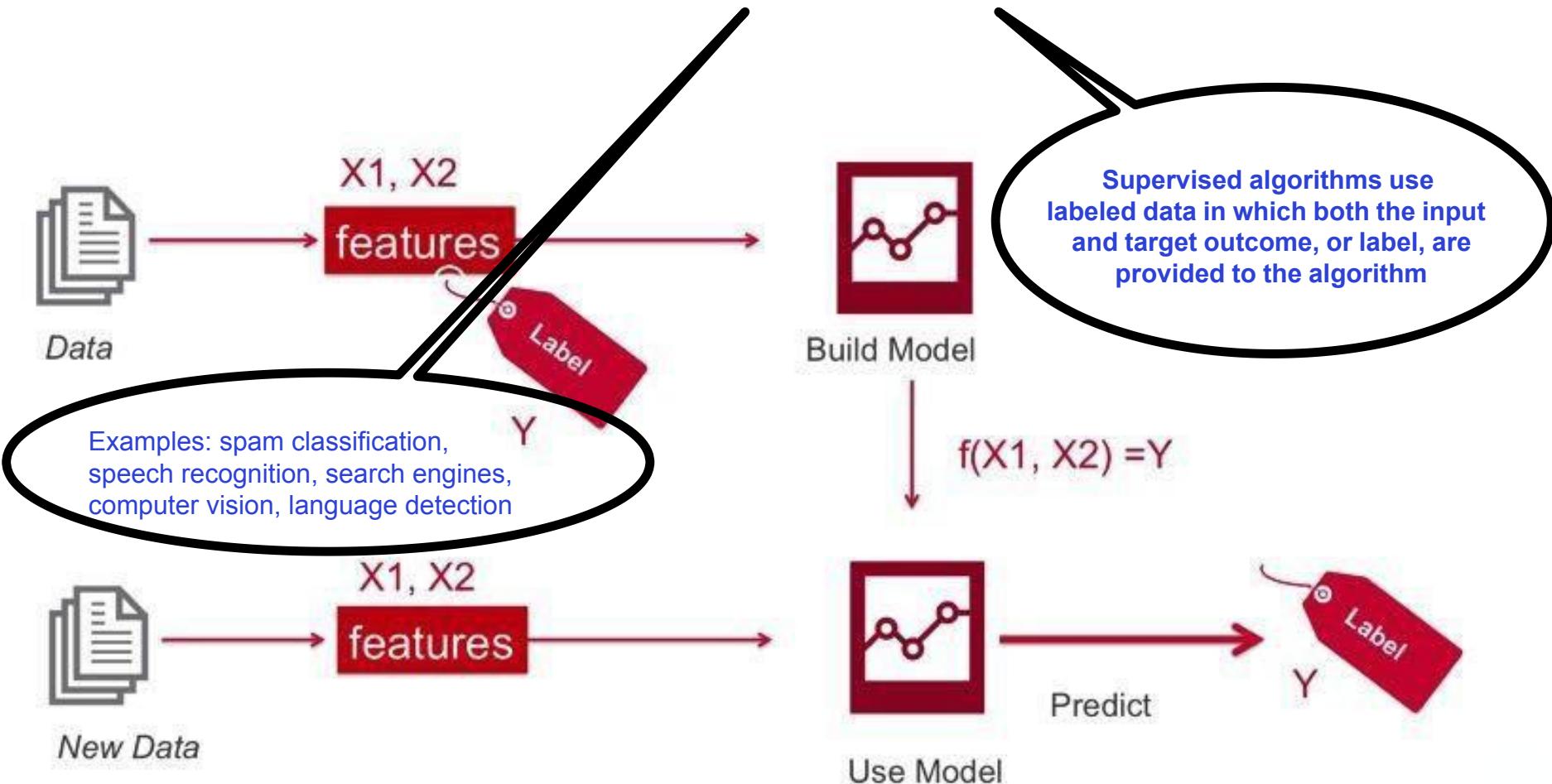
If it Walks/Swims/Quacks Like a Duck . . . Then It Must Be a Duck



Source:
<https://dzone.com/articles/demystifying-ai-machine-learning-and-deep-learning>

Data Processing & Analysis

Machine Learning: Supervised Learning



Source:
<https://dzone.com/articles/demystifying-ai-machine-learning-and-deep-learning>

120

Data Processing & Analysis

Machine Learning: Decision Trees

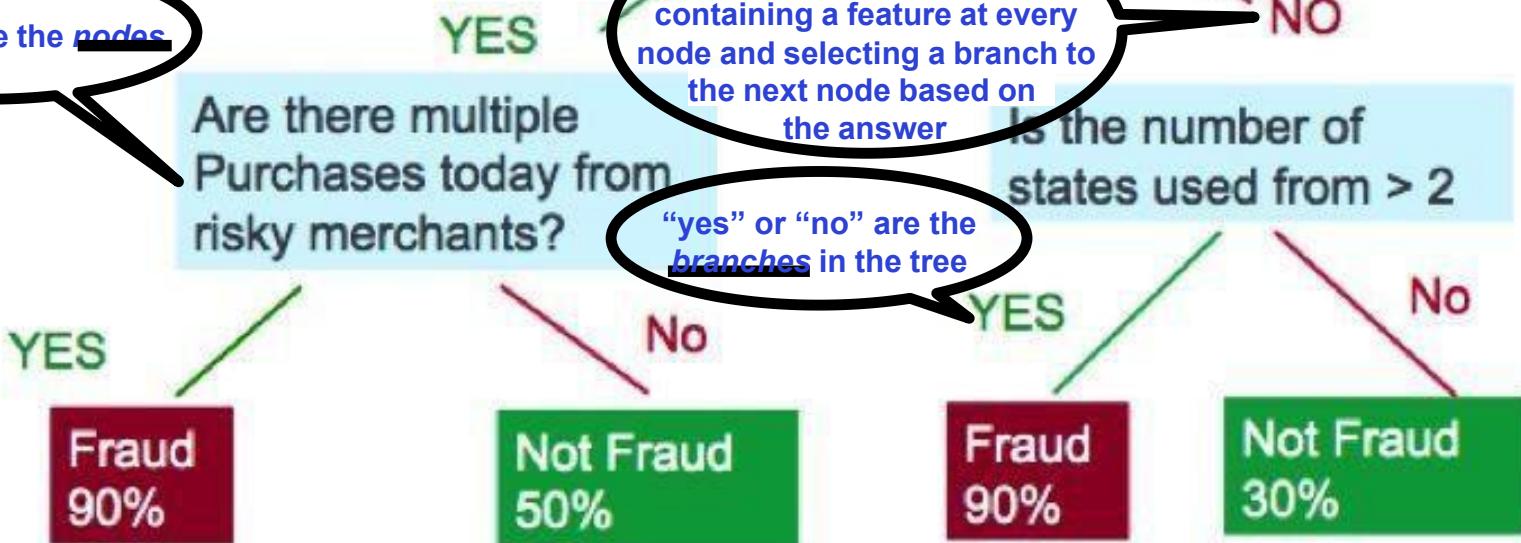
decision tree for predicting debit card fraud

Decision trees create a model that predicts the class or label based on several input features

questions are the nodes

Is the amount spent in 24 hours > average

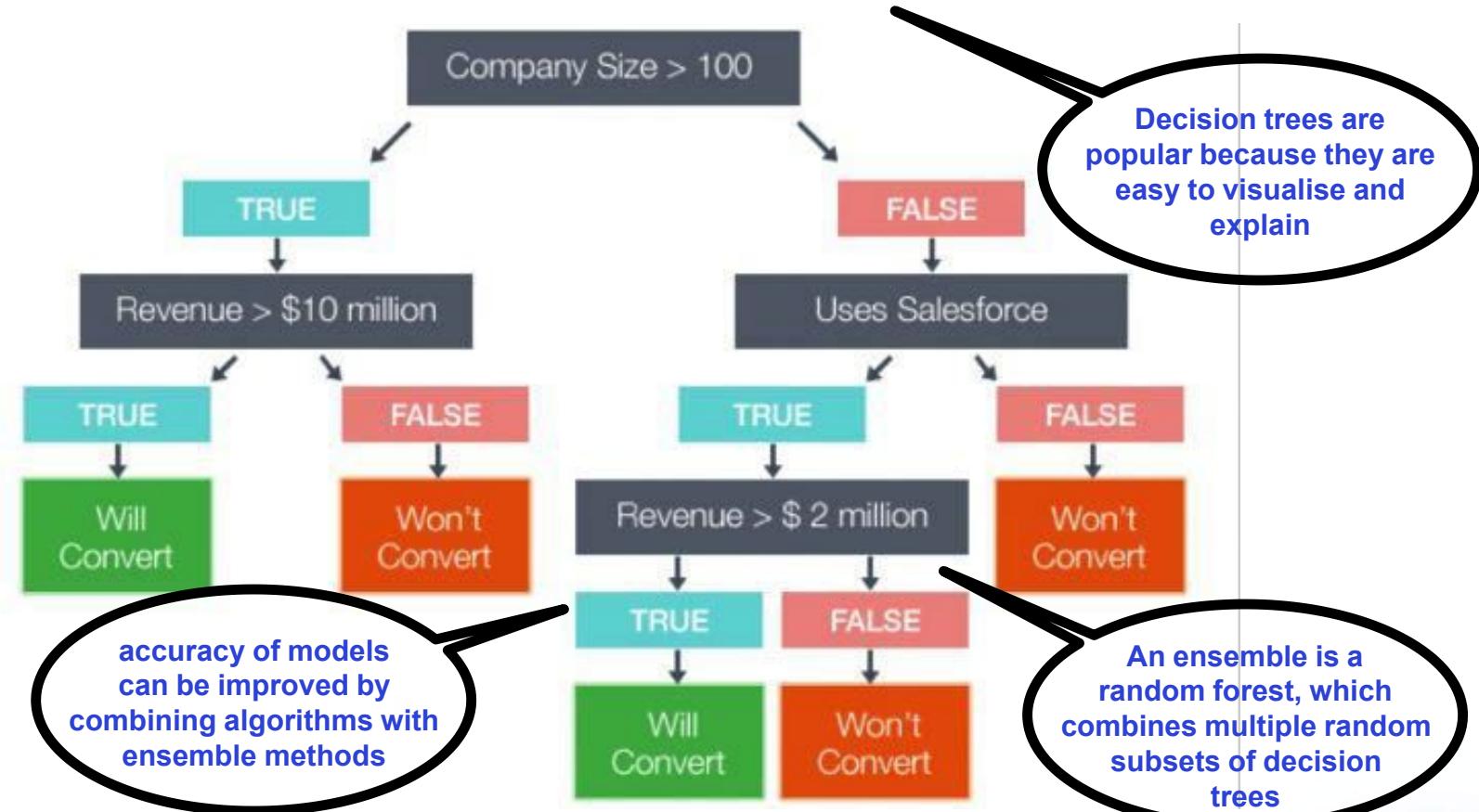
work by evaluating a question containing a feature at every node and selecting a branch to the next node based on the answer



"yes" or "no" are the branches in the tree

Data Processing & Analysis

Decision Trees: Ensemble Random Forests



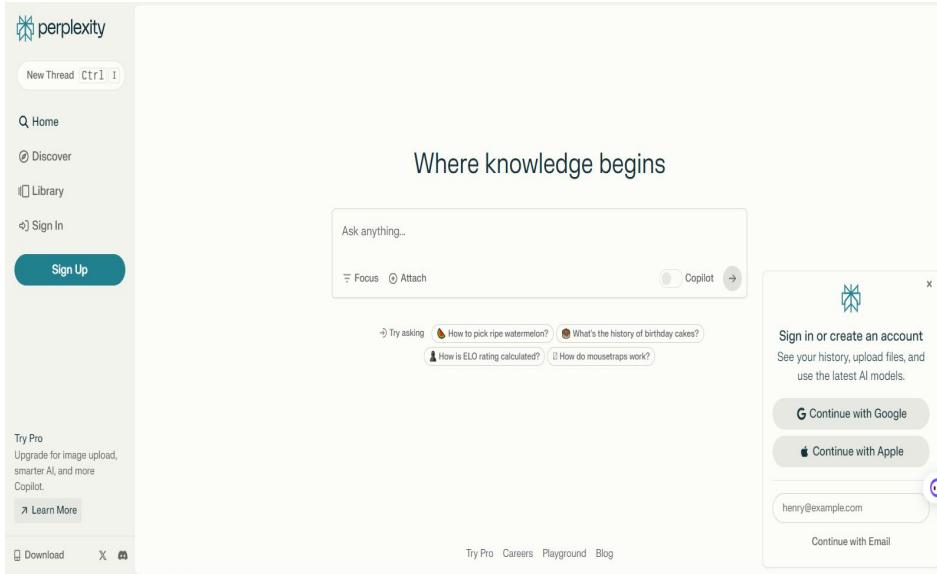
Source: <http://dataeconomy.com/2015/05/predictive-machine-learning-behind-the-scenes-at-flipkart-and-predictions-for-the-future-of-martech/>

AI Applications

- You could be next Steve Jobs
but please check what could be reused first!
- You can revisit this after you mastered Python 😊
- <https://www.kaggle.com/models>



Generative AI Tools Introduction – Perplexity



```
python
def compound_interest(principal, rate, time, n):
    """
    Calculate compound interest.

    principal: initial amount (float)
    rate: annual interest rate in decimal (e.g., 5% as 0.05)
    time: time in years (float)
    n: number of compounding periods per year (int)

    Returns:
        tuple containing (final amount, compound interest)
    """
    amount = principal * (1 + rate / n) ** (n * time)
    interest = amount - principal
    return amount, interest

# Example usage:
principal = 1000 # $1000
rate = 0.05 # 5% interest rate
time = 10 # 10 years
n = 12 # compounded monthly
```

```
# Example usage:
principal = 1000 # $1000
rate = 0.05 # 5% interest rate
time = 10 # 10 years
n = 12 # compounded monthly

final_amount, interest = compound_interest(principal, rate, time, n)
print(f"Final Amount: ${final_amount:.2f}")
print(f"Compound Interest: ${interest:.2f}")
```

Output:

```
text
Final Amount: $1647.01
Compound Interest: $647.01
```

- Enter <https://www.perplexity.ai/>
- Input prompt "write python compound interest calculations"



HKUSPACE
香港大學專業進修學院
HKU School of Professional and Continuing Education

THANK YOU

