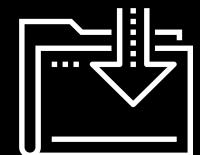
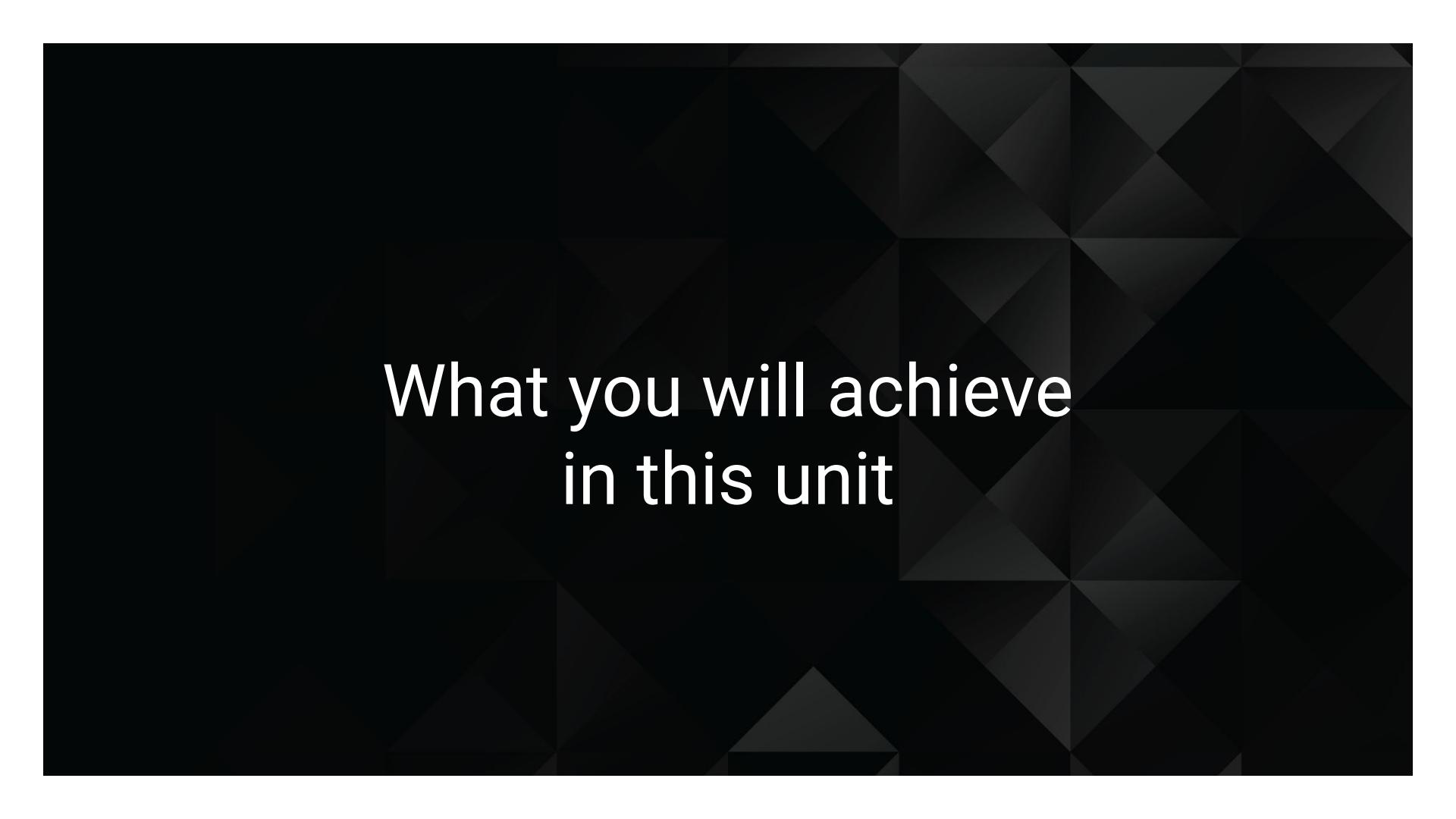


# Unsupervised Learning and The Cloud

FinTech  
Lesson 13.1





What you will achieve  
in this unit

# The Cloud Toolbox

---

You will gain hands-on experience in the following Amazon Web Services (AWS) tools:



**Amazon SageMaker:** To deploy ML models



**Amazon Lex:** To create conversational interfaces.



**Amazon S3:** To store files in the cloud.



**AWS Lambda:** To create serverless applications.

# Conversational Interfaces (Chatbots)

The image shows two separate instances of a conversational interface, likely from a cloud-based platform like Microsoft Bot Framework or similar. Both instances feature a header with a back arrow, the text 'Test bot (Latest)', and a green circular icon with a checkmark and the text 'Ready. Build complete.'

**Left Instance:**

- Header: 'Test bot (Latest)' and 'Ready. Build complete.'
- Middle Text Area: 'You're now ready for testing. Type an utterance below to begin conversation with your chatbot.'
- Message Bubbles:
  - Bot Message (Grey): 'Hello, I can help you to convert dollars to crypto.'
  - User Message (Blue): 'Hey, I would love to assist you with investing for retirement.'
- Buttons at the bottom: 'Clear chat history' and a microphone icon followed by 'Chat with your bot...' with a cursor pointing to it.

**Right Instance:**

- Header: 'Test bot (Latest)' and 'Ready. Build complete.'
- Middle Text Area: 'You're now ready for testing. Type an utterance below to begin conversation with your chatbot.'
- Buttons at the bottom: 'Clear chat history' and a microphone icon followed by 'Chat with your bot...'.

# Congratulations!

---

You already mastered these skills that will empower the cloud applications you will create.



Python



APIs



Time Series



Classification Algorithms

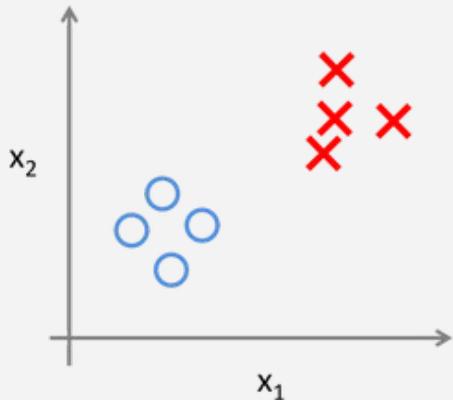


NLP

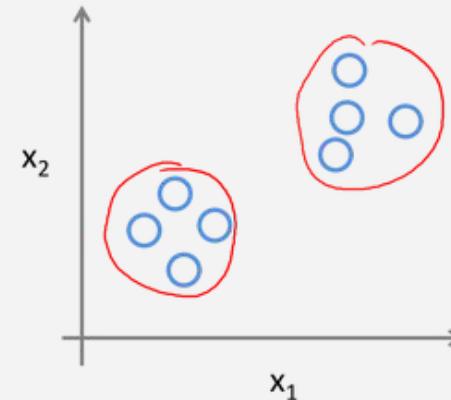
# Unsupervised Learning

# Supervised vs. Unsupervised Learning

## Supervised Learning



## Unsupervised Learning

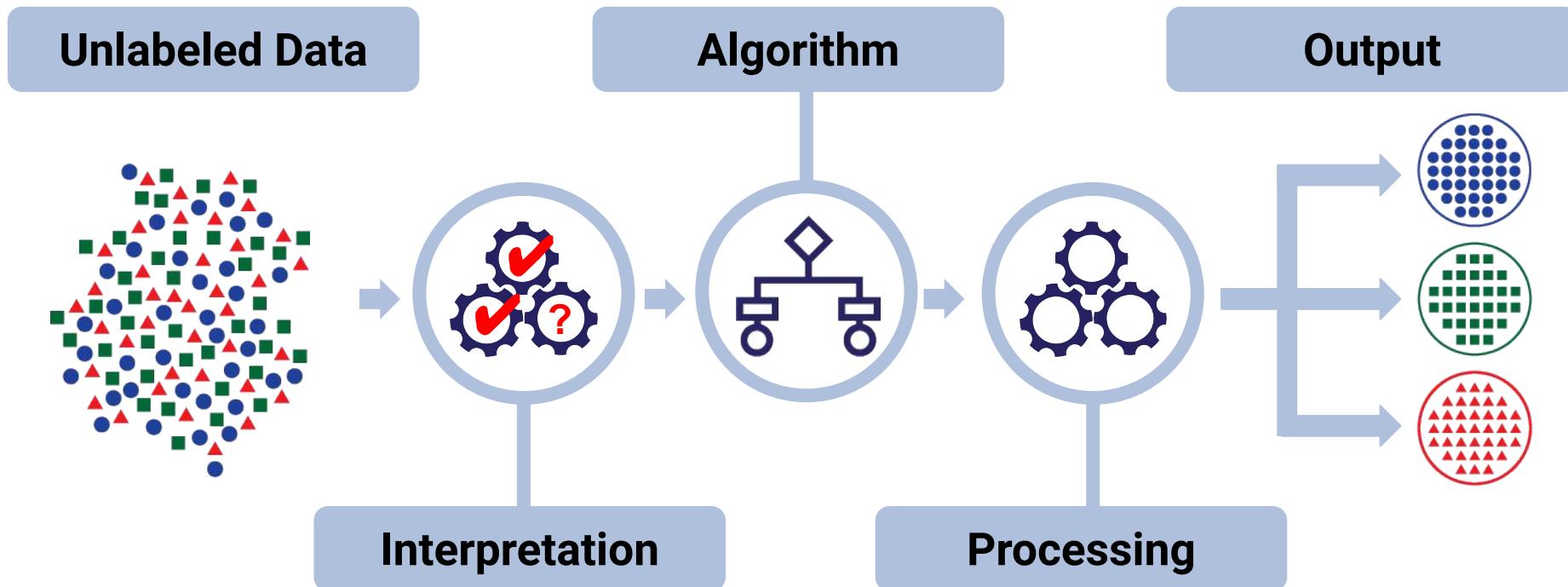


- Input data is labeled.
- Uses training datasets.
- **Goal:** Predict a class or value.

- Input data is unlabeled.
- Uses just input datasets.
- **Goal:** Determine patterns or grouping data.

# Unsupervised Learning at a Glance

**Unsupervised learning** uses unlabeled data to find patterns or group data.





**If unsupervised learning deals with unlabeled data, what kind of questions or business problems do you think we can afford?**

# Unsupervised Learning

We can group customers on a retail chain by shopping habits, so we can send customized offers by e-mail or using a mobile app to increase sales.

**FN** ≡ SUBSCRIBE NEWSLETTERS 

BUSINESS ▸ TECHNOLOGY OCT 3, 2019 11:51AM PT

## This New Technology Integration Helps Footwear Firms Offer More Personalization

BY MADELEINE STREETS





CREDIT: STANISIC VLADIMIR/SHUTTERSTOCK

As consumers begin to demand more curated shopping experiences both on- and offline, retailers are aiming to boost levels of personalization. But to do that well, they first need to know who the customer is.

# Unsupervised Learning

Having thousands of transactions per day on credit card operations, it's hard to identify anomalous or fraudulent transactions.

We can use unsupervised learning to find patterns among transactions data to identify anomalies and potential fraudulent transactions.

The screenshot shows the Finextra website homepage. At the top, there is a navigation bar with links for NEWS, TV, RESEARCH, EVENTS, RESOURCES, COMMUNITY, BLOGS, CAREERS, and options to SIGN UP and log IN. Below the navigation is a search icon. The main header is "/ Regulation & Compliance" with a subtitle "Regulation and Compliance News and Resources". On the left, there is a sidebar titled "News" with a "Sign Up" button. The main content area features a news article titled "HSBC introduces financial crime detection systems" dated 26 September 2019. The article discusses HSBC's deployment of an industry-leading Anti-Money Laundering (AML) system and an automated sanctions checking system. To the right of the article, there are several advertisements: one for HID (On Demand Webinar) Fighting fraud during digital identification using AI and data, another for Temenos (AWS Cloud Series), and a third for Refinitiv (Get a 360° view of your book of business with Refinitiv's next generation Wealth Management solutions). At the bottom, there are logos for ACI, Universal Payments, and IFTech.

# Unsupervised Learning

We can use unsupervised learning to cluster stock data, so we can create investment portfolios according to the resulting groups.

## March of the machines

The stockmarket is now run by computers, algorithms and passive managers

*Such a development raises questions about the function of markets, how companies are governed and financial stability*



Satoshi Kambayashi

Print edition | Briefing >

Oct 5th 2019 | NEW YORK



FIFTY YEARS ago investing was a distinctly human affair. "People would have to take each other out, and dealers would entertain fund managers, and no one would know what the prices were," says Ray Dalio, who worked on the trading floor of the New York Stock Exchange (NYSE) in the early 1970s before founding Bridgewater Associates, now the world's largest hedge fund. Technology was basic. Kenneth Jacobs, the boss of Lazard, an investment bank, remembers using a pocket calculator to analyse figures gleaned from company reports. His older colleagues used slide rules. Even by the 1980s "reading the *Wall Street Journal* on your way into work, a television on the trading floor and a ticker tape" offered a significant information advantage, recalls one investor.



How can we understand  
our customers?

# Supervised Vs. Unsupervised Learning

---

## Supervised Learning Approach

- Is this person a female or male?
- Is this person happy?
- How much is this person going to expend next month?



**Upbeat Millennial**  
purchases Matcha tea drinks  
most often

### Behaviors

Average Purchase per month: \$64.32

### Motivations

The most important factor in purchase is knowing that their tea is sustainably sourced.

# Supervised Vs. Unsupervised Learning

---

## Unsupervised Learning Approach

- How can I create a customized offer to customers?

A smartphone screen displays a promotional offer for a green tea smoothie. The phone has a white back cover with a camera and speaker grille on the left and a power button on the right. The screen shows a green smoothie in a cup topped with whipped cream and a straw, followed by promotional text and a 'Buy Now' button.

Buy one Green Tea Smoothie, **get one free**

**A Great deal WITH great health benefits.** Green tea is known for the remarkable and essential health benefits that a daily cup can bring.

[Buy Now](#)

# Applications of Unsupervised Learning

## Clustering

Allows automatic splitting of the dataset into groups according to similarity.

It can be used for customer segmentation and targeting.



# Applications of Unsupervised Learning

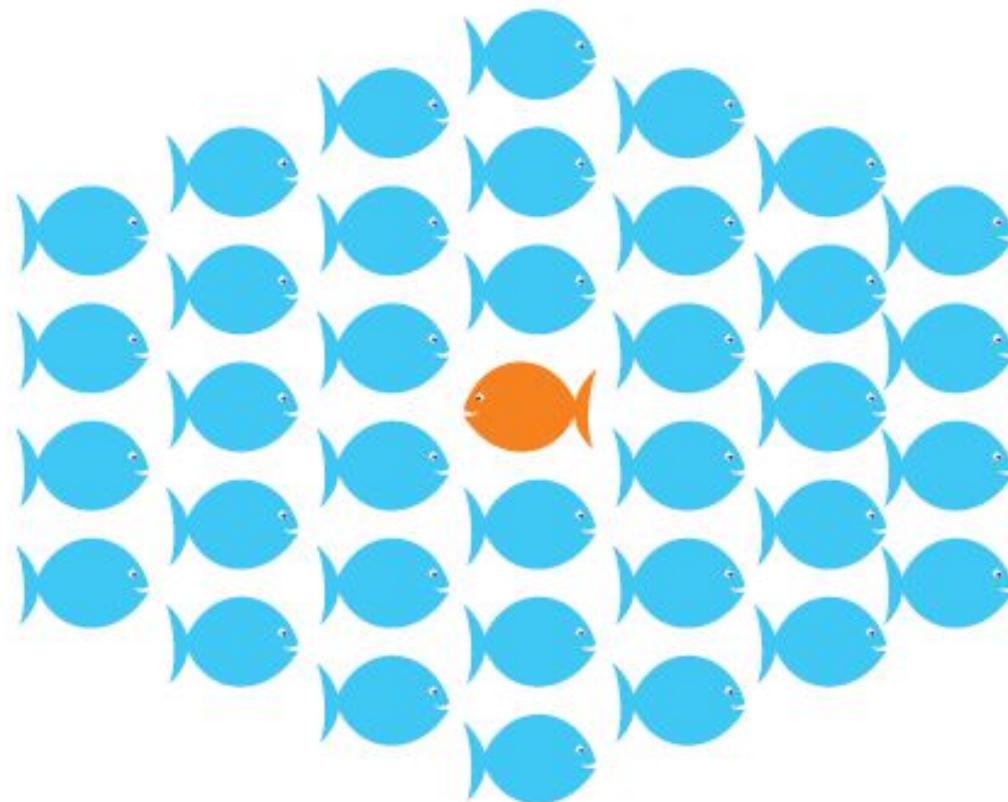
---

## Anomaly Detection

Automatically discovers unusual data points in a dataset.

### It's useful in:

- identifying fraudulent transactions
- discovering faulty pieces of hardware
- identifying an outlier caused by a human error during data entry



# Customer Segmentation

---

~~This division is based on customers~~  
This is the division of potential customers in a given market into discrete groups.  
~~that division is based on customers~~  
having similarities such as:

- Customer needs  
(e.g. a particular product can satisfy some of them)
- Responses to online marketing channels
- Buying habits  
(e.g. best day for buying, weekly spend)



# Customer Segmentation in Action

Some facts about how customer segmentation is driving revenue in leading companies:

**75% of Netflix viewer activity  
is driven by recommendation**

The screenshot shows the Netflix homepage with a red header. On the left is the Netflix logo. To its right are navigation links: "Watch Instantly", "Just for Kids", "Instant Queue", and "Personalize". A search bar contains the placeholder text "Movies, TV shows, actors, directors, etc.". To the right of the search bar is a user profile icon of a woman with dark hair and red lips. Below the header, a large white section displays a welcome message: "Angela, welcome to your **very own Netflix homepage!** Based on your ratings, we've filled it with personalized suggestions **JUST FOR YOU.** The more you rate, the better we get at giving you suggestions you'll love."

# Customer Segmentation in Action

Some facts about how customer segmentation is driving revenue in leading companies:

**35% of Amazon's sales** are generated through their recommendation engine



**Recommended for You**

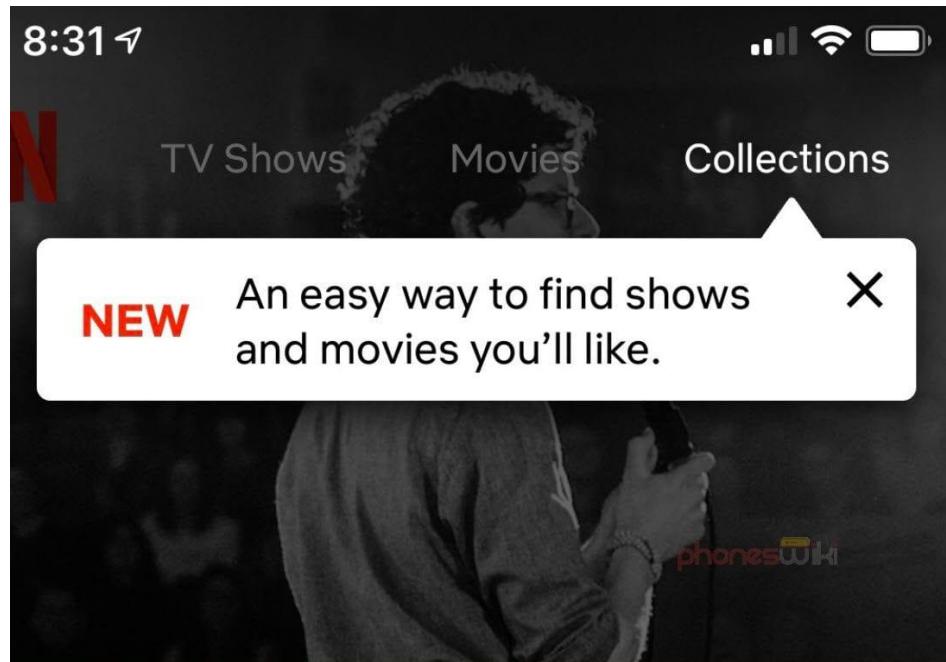
Amazon.com has new recommendations for you based on items you purchased or told us you own.

([Source](#))

# Customer Segmentation in Action

Some facts about how customer segmentation is driving revenue in leading companies:

Netflix's recommendation system saves the company an estimated **\$1 Billion** per year through reduced churn



([Source](#))

# Customer Segmentation in Finance

How a bank could use customer segmentation to improve their products?

= Forbes

EDITOR'S PICK | 508 views | Oct 10, 2019, 11:22am

## The 10 Best Gas Rewards Credit Cards Of 2019

 Jason Steele Contributor  
Advisor Contributor Group ©  
Personal Finance

f Editorial Note: Forbes adheres to strict editorial integrity standards. The content on this page may contain partner offers and/or affiliate links, and Forbes may receive compensation if you click through a link on this page. To the best of our knowledge, all content is accurate as of the date posted, though offers contained herein may no longer be available. See our Advertiser Disclosure for more details.

t in We reviewed dozens of credit cards to find those that pay increased rewards when you fill up your tank. The result is our list of the 10 best gas rewards credit cards.



To figure out the finances of the 2020 candidates, Forbes examined financial disclosure statements, scoured local real estate records and calculated pension benefits.

Lots of companies sell marketing clouds.  
**Listrak delivers results.**

Email | Marketing Automation | CRM  
Predictive Analytics | SMS

# Customer Segmentation in Finance

How can an investment portfolio be improved using customer segmentation?

## How Customer Segmentation Can Help You

Before we talk about the most effective types of customer segmentation, let's take a quick look at how customer segmentation can benefit your brand.

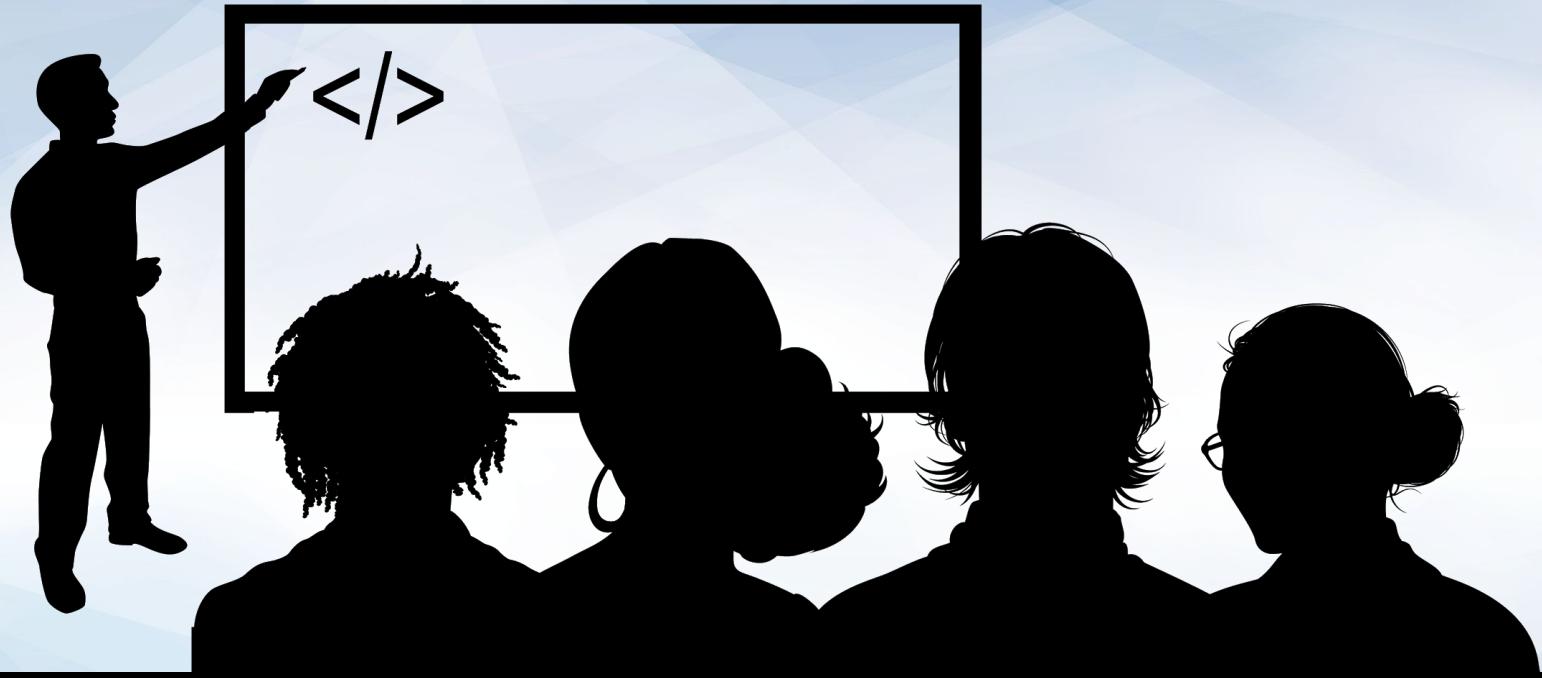
### 1. Develop Effective Retention Strategies

Did you know that [\\$1.6 trillion](#) was the estimated cost incurred by businesses due to lost customers in 2016? And, once customers leave, 68% of them don't ever go back to the same brand again. If that doesn't motivate you to retain your customers, maybe you should also consider that acquiring new customers is [5X as expensive](#).

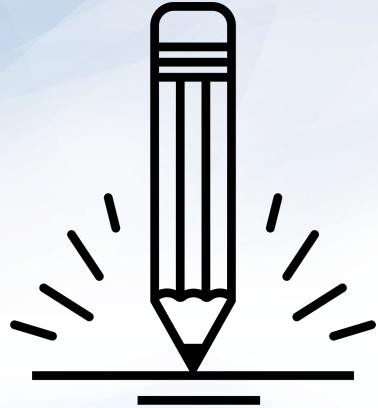
Customer segmentation can help you develop more focused strategies to retain customers. For example, you may want to identify your top-paying customers and create exclusive offers for them. Or you may even want to re-engage people who haven't purchased in a while. The possibilities are endless.

### 2. Provide Superior Customer Experience

[81%](#) of consumers reported getting frustrated working with brands that don't provide a good customer experience. A bigger concern is that 44% of customers who are unhappy with the experience will vent on social media. This can harm your reputation and your



## Instructor Demonstration Data Preparation for Unsupervised Learning



## **Activity:**

### Understanding customers

In this activity, you will perform some data preparation tasks on a dataset that contains data from purchases on an e-commerce website made by 200 customers.

**Suggested Time:**  
20 minutes





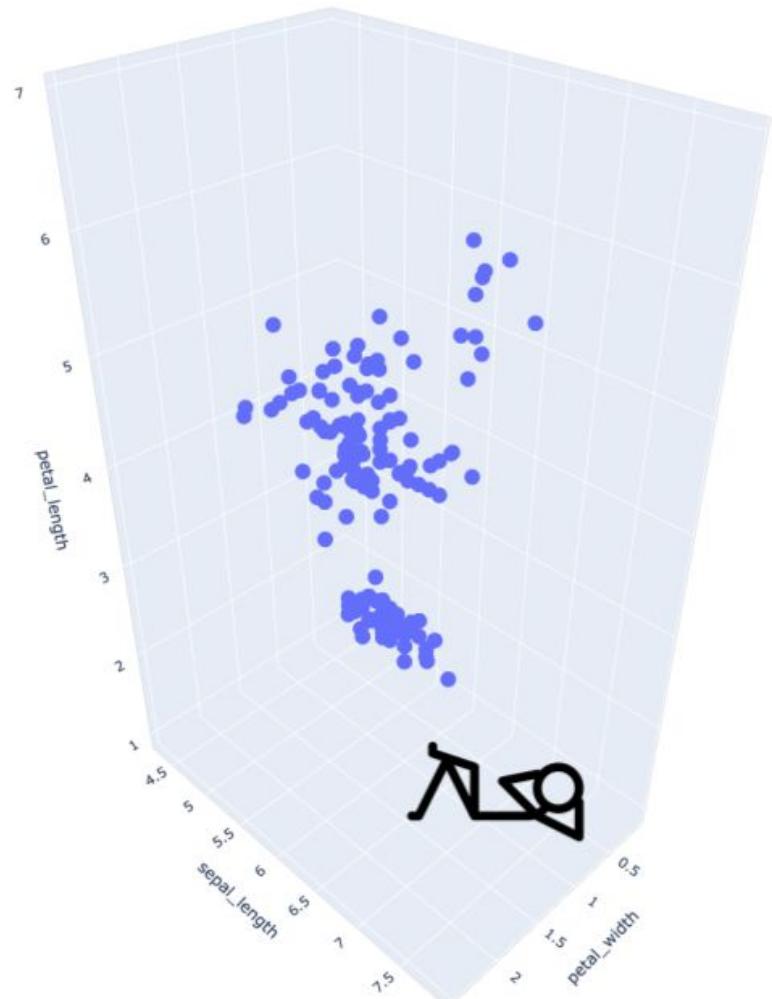
**Time's Up! Let's Review.**

# The K-Means Algorithm

# What K-Means can do for me?

---

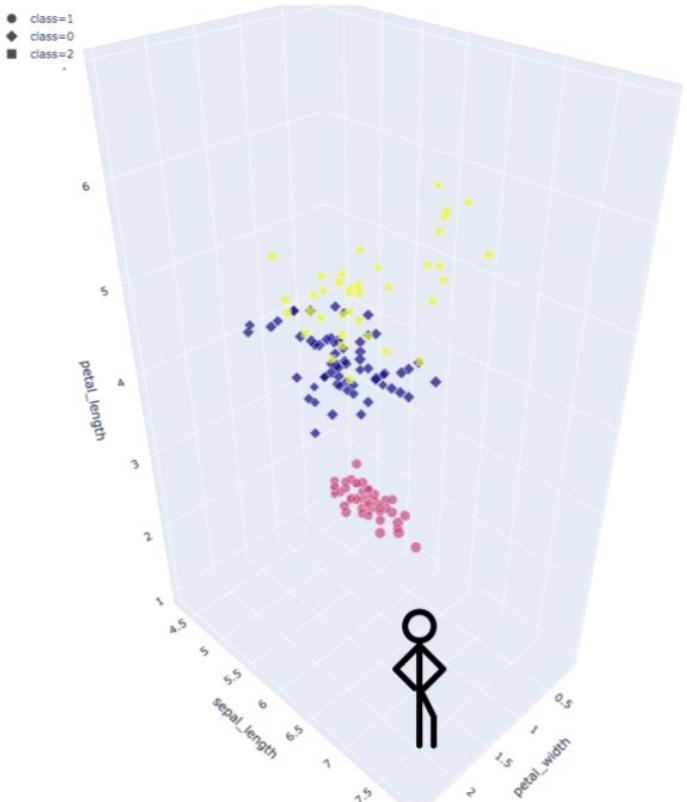
- Imagine that you are in a room full of spheres (data points), and you want to learn more about them, so you start to observe.
- You realized that every sphere represents a flower and that axes represent features of flowers.
- After observing the flowers, you discovered that there are some patterns when you combine the three features.



# What K-Means can do for me?

---

- At this point, K-Means comes to the rescue!
- K-means is an unsupervised learning algorithm used to identify clusters and solve clustering issues.



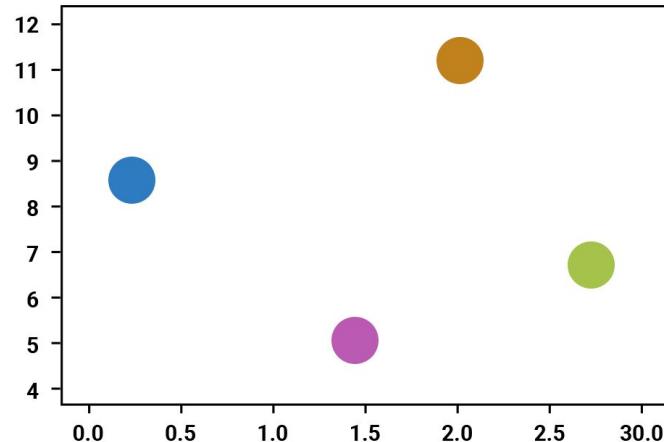
# How K-Means Algorithm Works

---

- K-Means algorithm groups the data into **k clusters**, where belonging to a cluster is based on some similarity or distance measure to a centroid.
- A **centroid** represents a data point that is the arithmetic mean position of all the points on a cluster.

## K-means Clustering

### Initial Clusters



# How K-Means Algorithm Works

## Algorithm at a glance

01

Randomly initialize the  $k$  starting centroids.

02

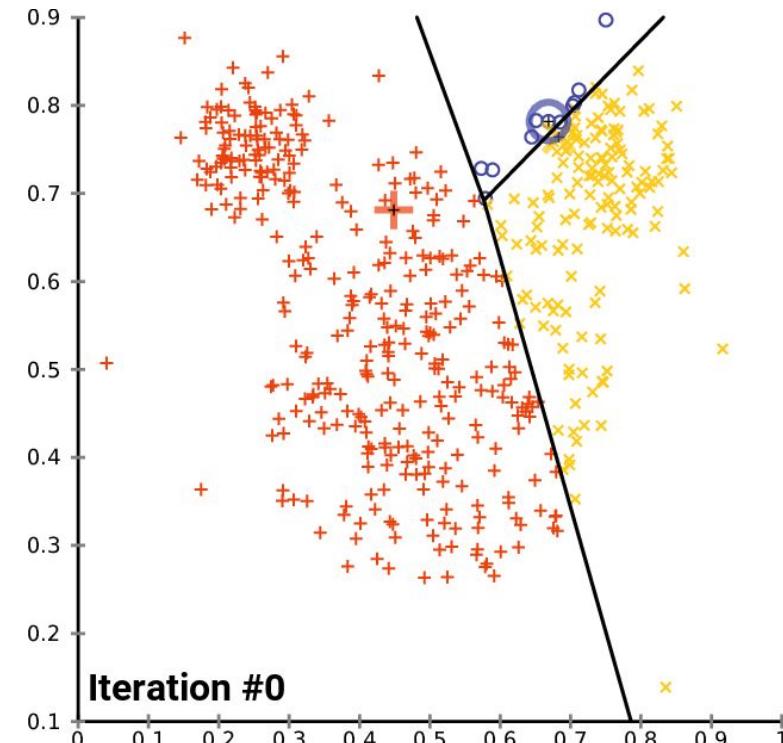
Each data point is assigned to its nearest centroid.

03

The centroids are recomputed as the mean of the data points assigned to the respective cluster.

04

Repeat steps 1 through 3 until the stopping criteria is triggered.

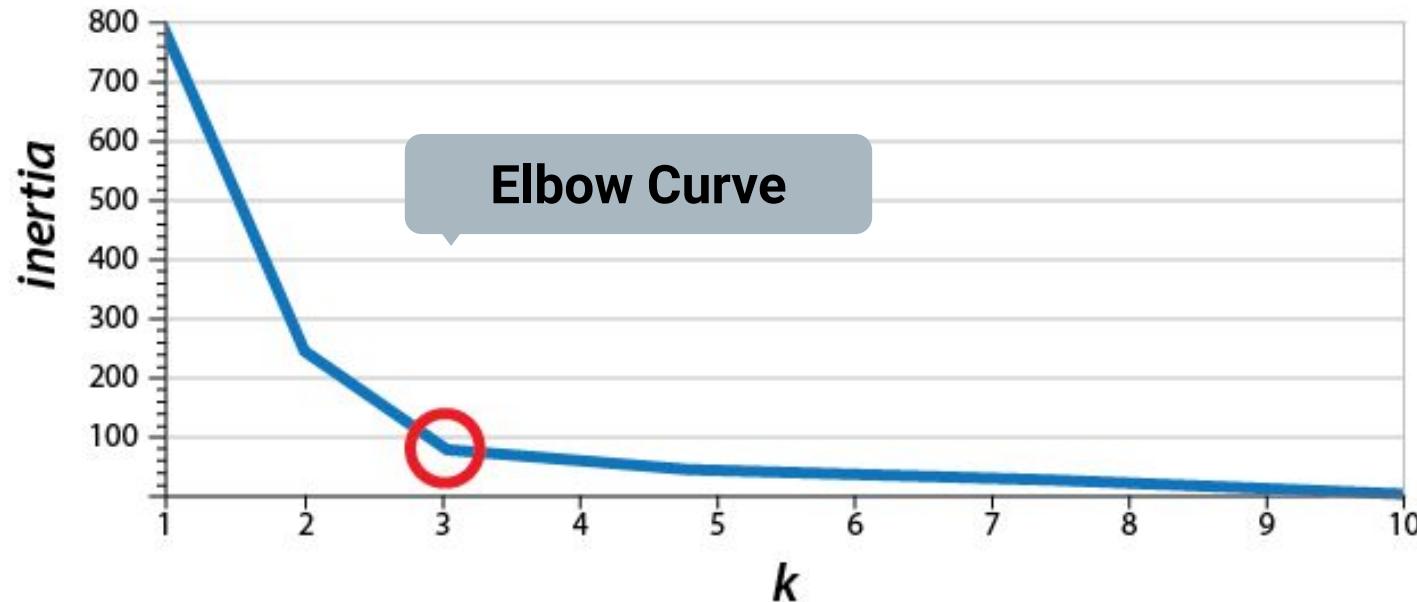


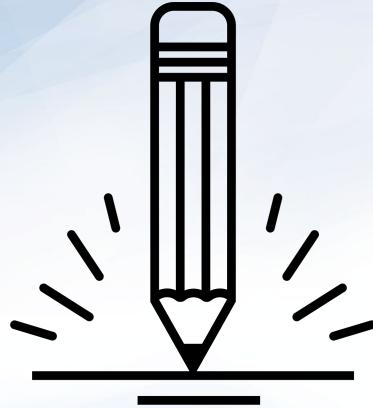
# What is the Best Number for $k$ ?

---

This is done using an **elbow curve**, where the  $x$  axis is the  $K$ -value and the  $y$  axis is some *objective function*.

A common objective function is the *inertia*.





## **Activity:** Customer segments for e-commerce

In this activity, you will identify the best number of clusters on a customers clustering scenario using the Elbow Curve, next you will use K-Means to cluster data and make conclusions about the obtained results.

**Suggested Time:**  
20 minutes

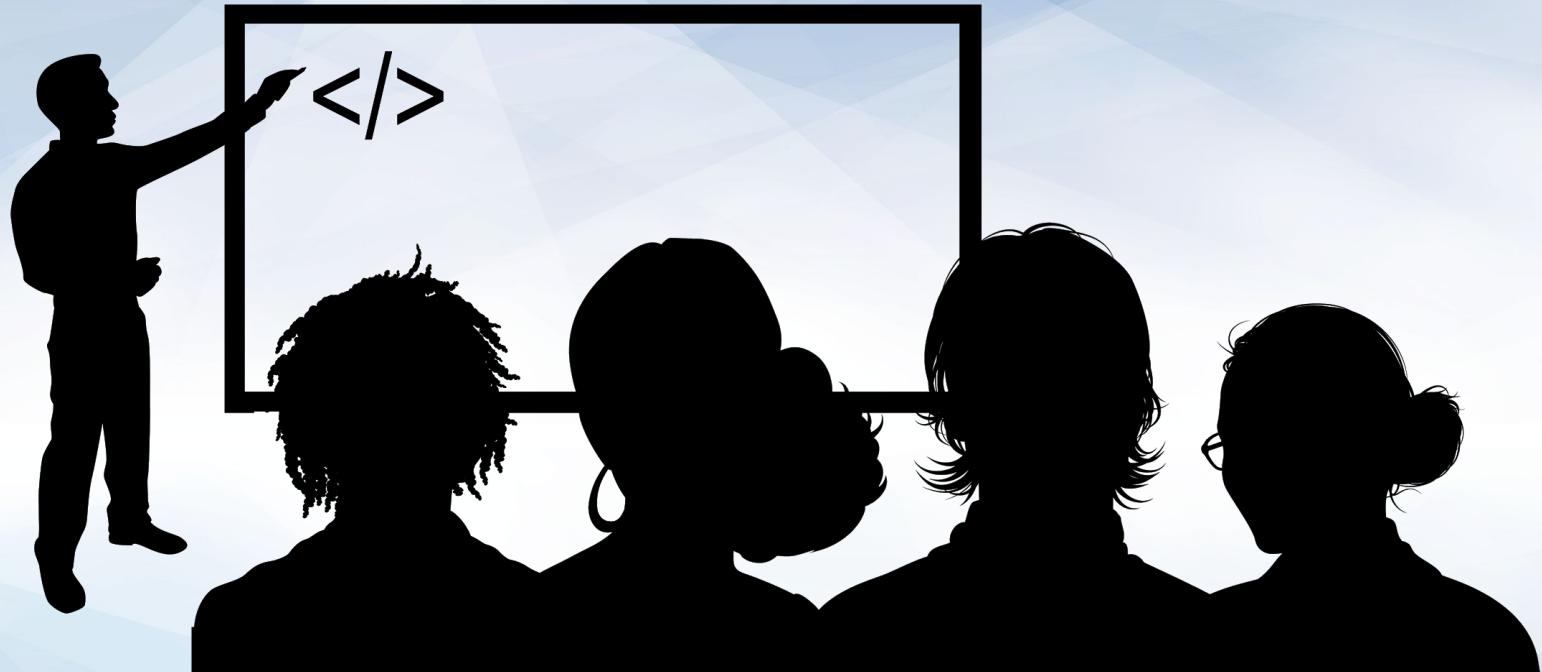




**Time's Up! Let's Review.**

*Break*





Instructor Demonstration  
Speeding up ML algorithms with PCA



## Activity: PCA in Action

In this activity, you will use PCA to reduce the dimensions of the consumers shopping dataset.

Suggested Time:  
20 minutes





**Time's Up! Let's Review.**

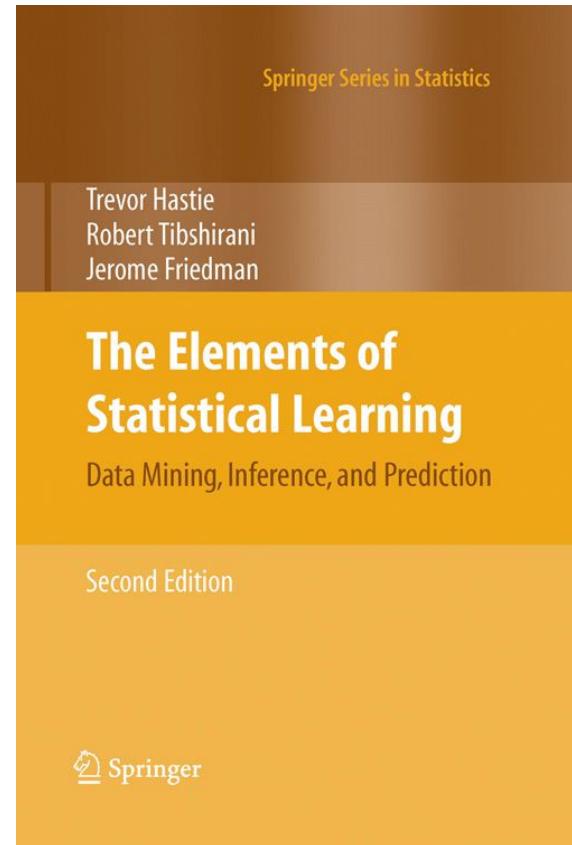
# Do you want to learn more?

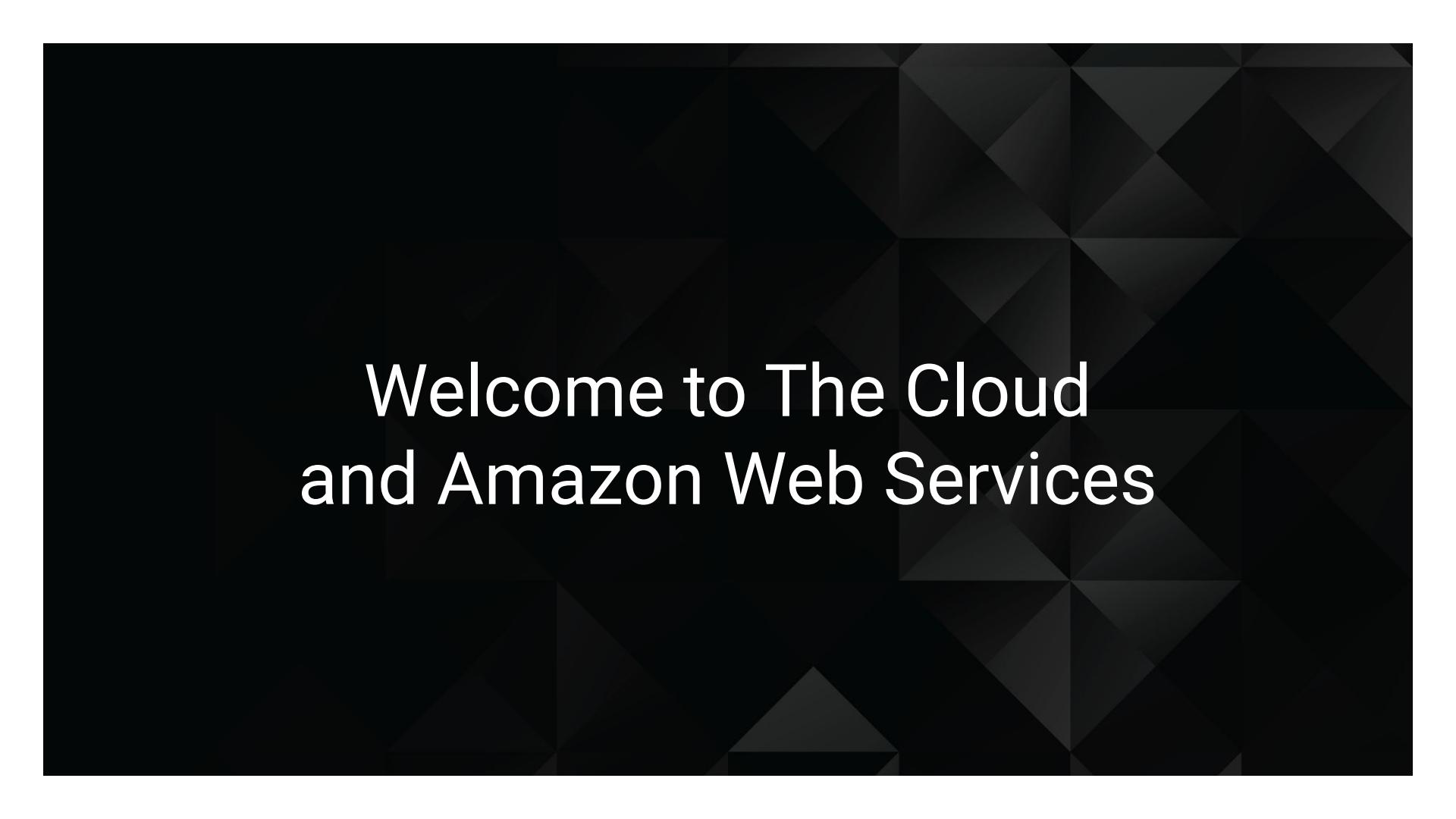
---

**Focus on**  
Chapter 14, Unsupervised Learning

This book is freely shared by authors at  
[stanford.edu](http://stanford.edu)

*The Elements of Statistical Learning: Data Mining, Inference, and Prediction*  
Second Edition  
by Trevor Hastie, Robert Tibshirani, Jerome Friedman  
(Springer 2009)



The background of the slide features a dark, abstract geometric pattern composed of numerous small, dark gray triangles of varying sizes, creating a subtle, organic texture.

# Welcome to The Cloud and Amazon Web Services

# Hello Cloud!



What do you think  
the cloud is?

# Hello Cloud!

---

The Cloud is a term used in the internet world to refer to different services, such as:

Data Storage

e.g.

- Google Drive
- DropBox
- Microsoft OneDrive

Computing Power

e.g.

- Amazon Web Services
- Microsoft Azure
- Google Cloud

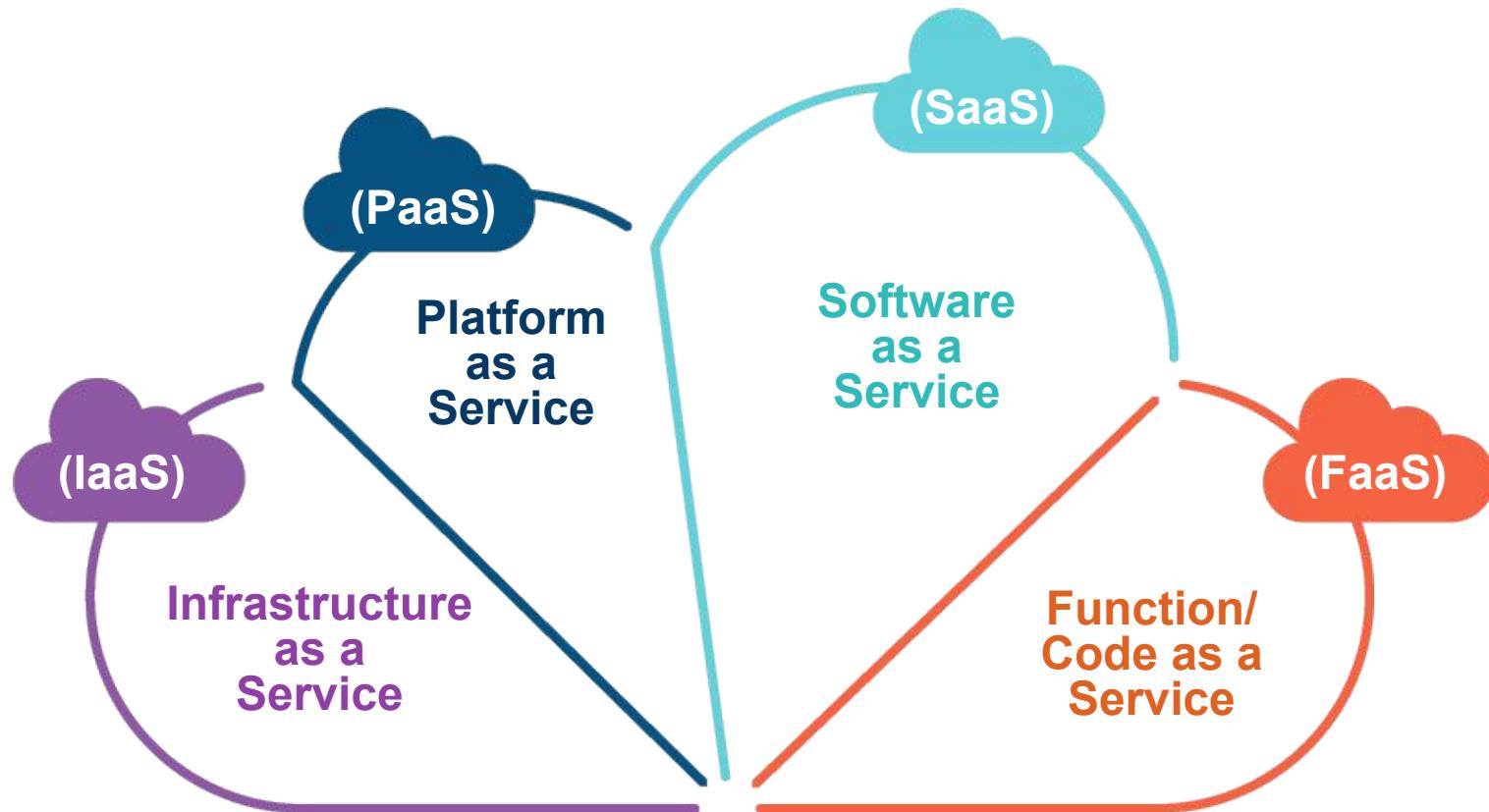
Collaboration Platforms

e.g.

- G-Suite
- Microsoft Office 365
- Zoho

# Models of Cloud Services

---



# Disadvantages of the Cloud

# Disadvantages of the Cloud

Having to trust a  
3rd party provider  
with your data.

THE WALL STREET JOURNAL.  
English Edition | October 10, 2019 | Print Edition | Video

Home World U.S. Politics Economy Business Tech Markets Opinion Life & Arts Real Estate WSJ. Magazine

TECH

## Capital One Breach Casts Shadow Over Cloud Security

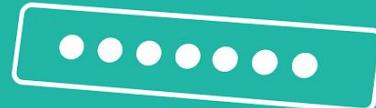
Massive data exposure highlights sustained risk from poor information-protection practices



# Disadvantages of the Cloud

---

You probably have to give up  
some privacy and confidentiality.



# Disadvantages of the Cloud

---

You depend on the SLAs of a provider for issues to be addressed

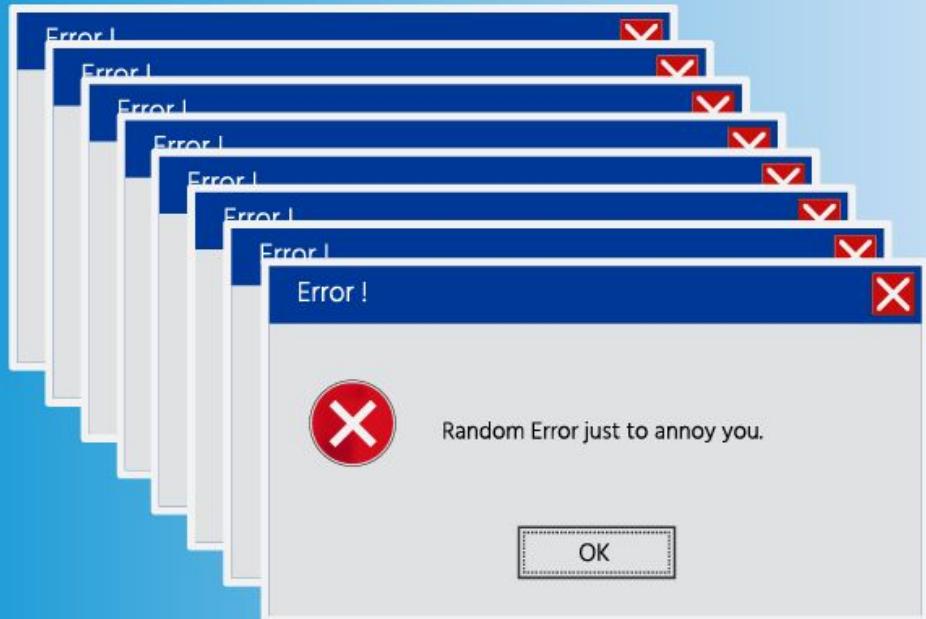
The image shows a portion of a laptop keyboard at the top, with keys like 'ctrl', 'fn', and various arrow keys visible. Below the keyboard, a hand holds an orange and yellow striped pen over a white document. The document has 'Service Level Agreement' printed in large, bold, black letters at the top. In the bottom left corner of the document, the words 'Agreement Overview' are partially visible. The background is a light color, and the overall composition suggests a professional or technical environment.

Service Level Agreement

Agreement Overview

# Disadvantages of the Cloud

Incidents such as hardware failure  
are inevitable.



# Disadvantages of the Cloud

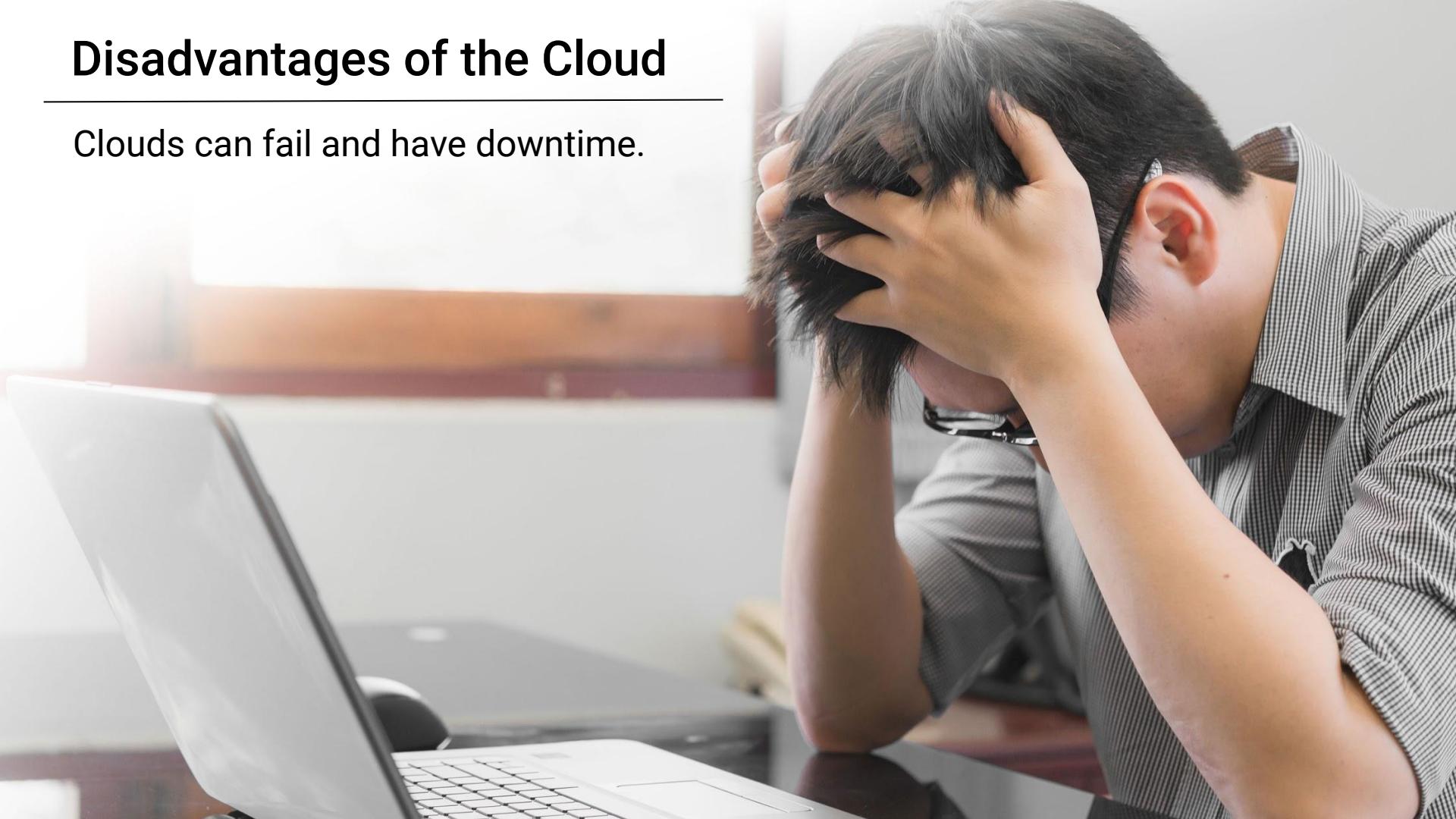
Hardware might be shared with other tenants,  
"bad neighbors" can affect your resources



# Disadvantages of the Cloud

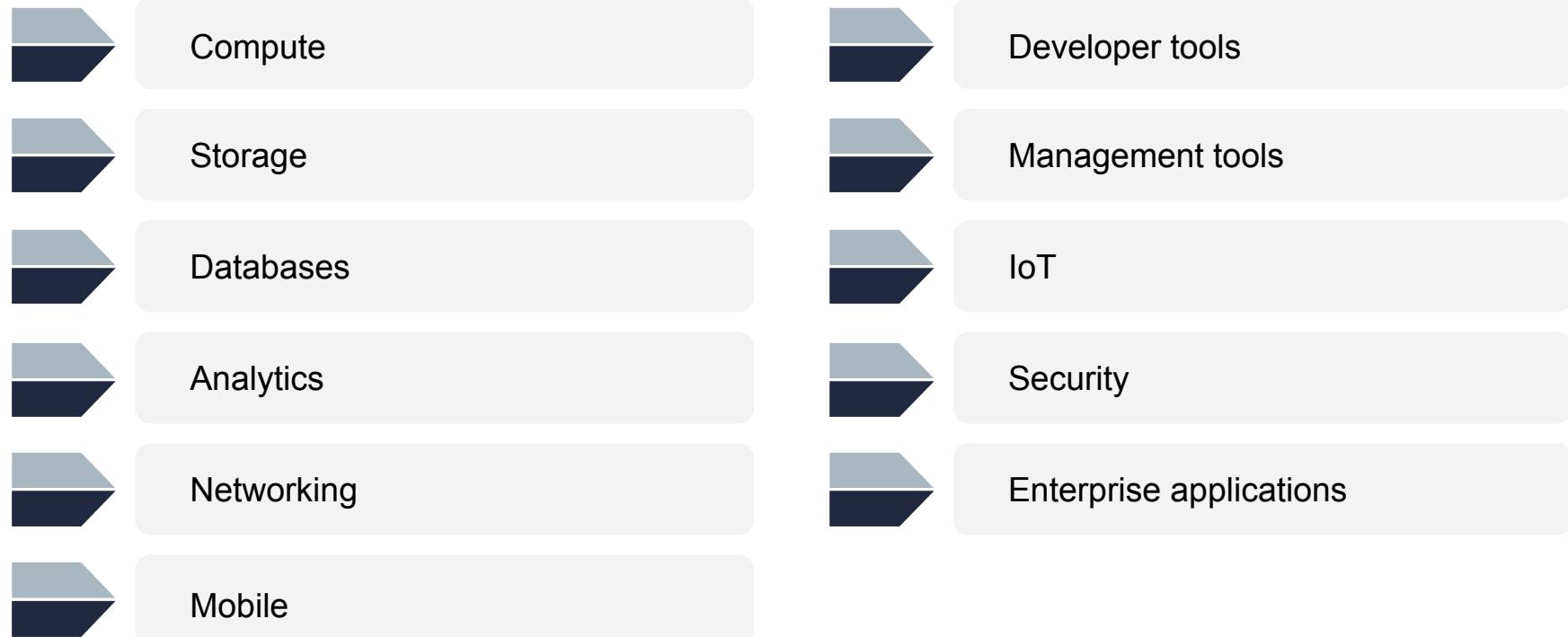
---

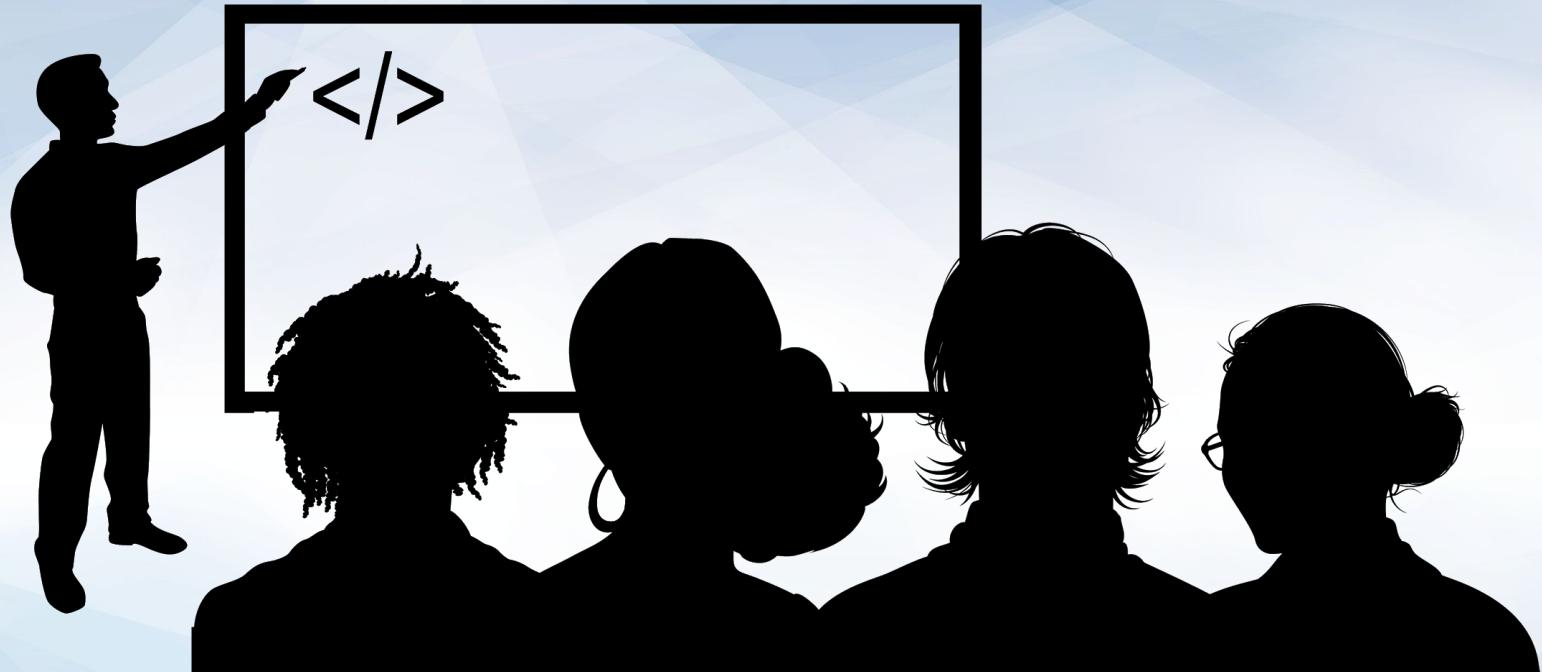
Clouds can fail and have downtime.



# Computing Power: Amazon Web Services

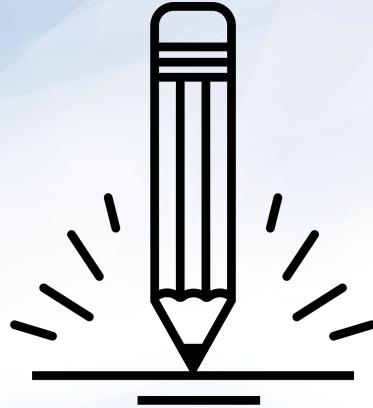
 comprises more than 140 services across these categories:





# Instructor Demonstration

## Intro to Amazon SageMaker



# Activity: Explore SageMaker

In this activity you will explore SageMaker.

Suggested Time:  
5 minutes



Questions?