

# 01-01

## Introduction au partitionnement de données

---

**NOUS ÉCLAIRON.  
VOUS BRILLEZ.**

---

FORMATION CONTINUE  
ET SERVICES AUX ENTREPRISES



# Sommaire

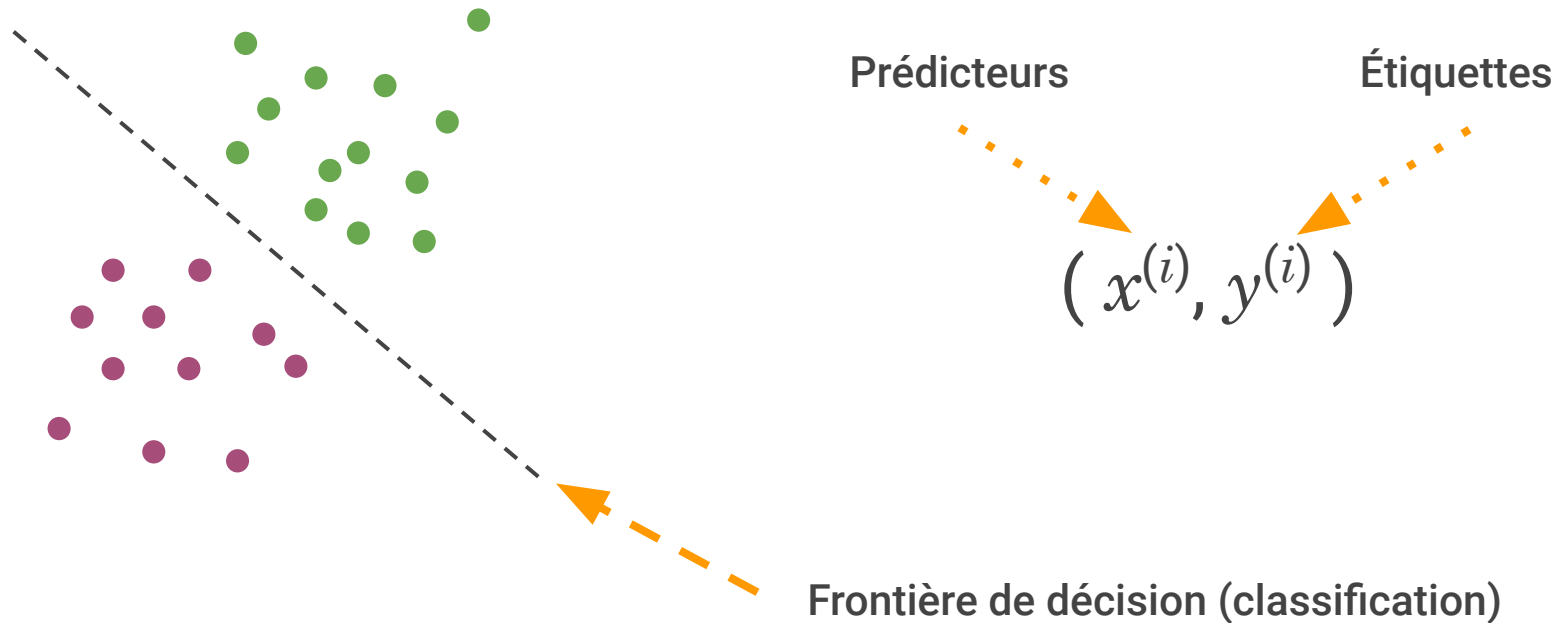
1. Introduction
2. Les principaux types de partitionnement
3. Applications
4. Lecture obligatoire
5. Références

# Sommaire

1. Introduction
2. Les principaux types de partitionnement
3. Applications
4. Lecture obligatoire
5. Références

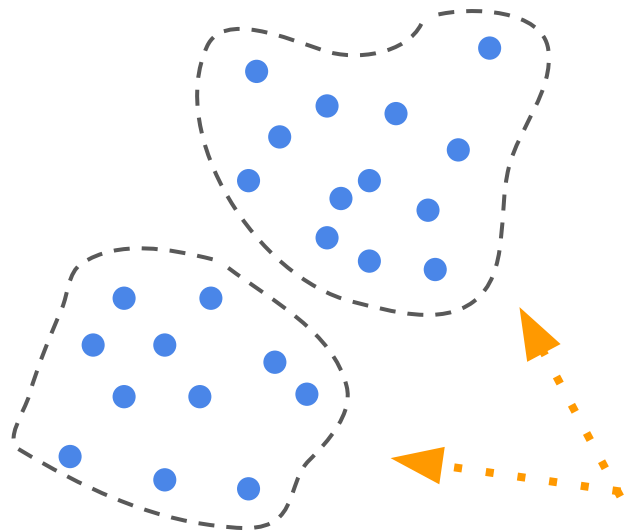
# Introduction

Rappelez-vous, en **apprentissage supervisé** ...



# Introduction

Maintenant, en **apprentissage non supervisé**



Prédicteurs

Pas d'étiquettes !

$(x^{(i)})$

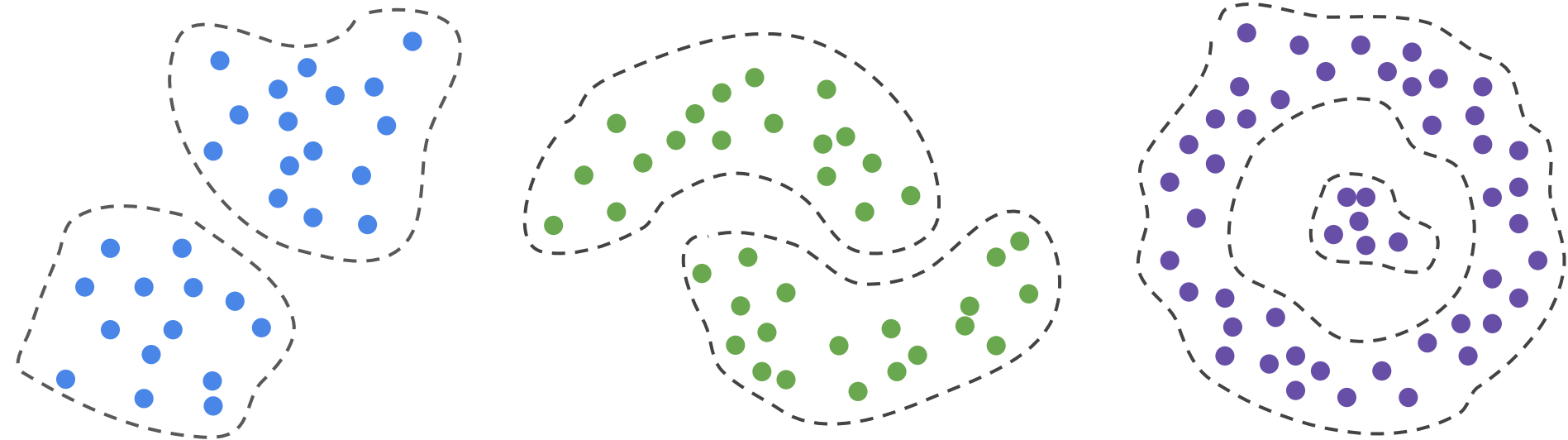
)

Partitions (clusters)

On réalise alors un **partitionnement des données**  
(**data clustering**) ou classification non supervisée

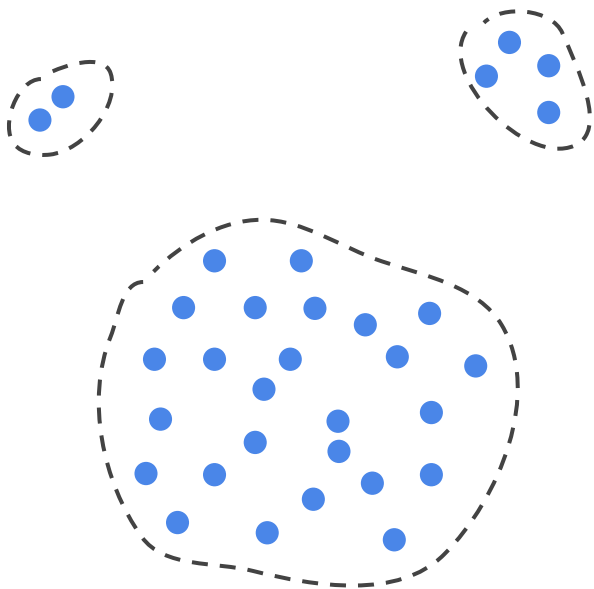
# Introduction

Les clusters peuvent avoir différentes morphologies

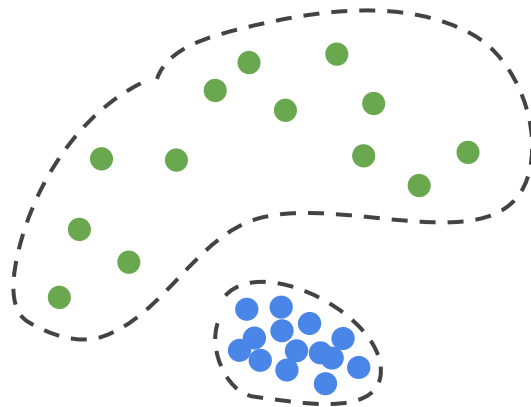


# Introduction

Ils peuvent être de différentes tailles ...

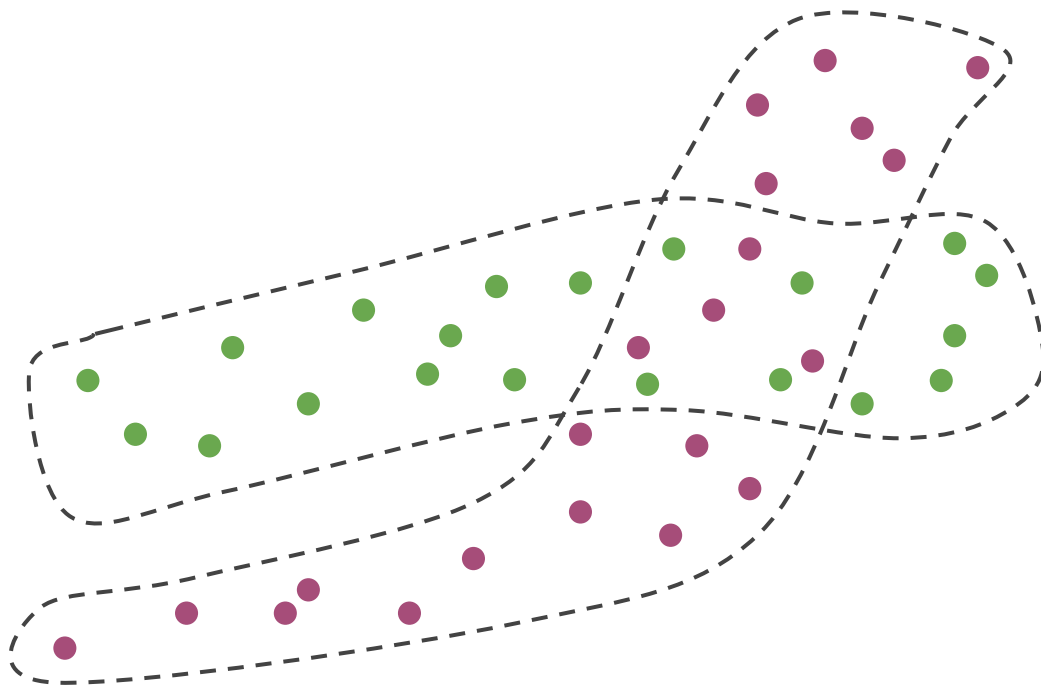


... et de différentes densités



# Introduction

Dans certains cas, il peut même y avoir chevauchement





# Sommaire

1. Introduction
2. Les principaux types de partitionnement
3. Applications
4. Lecture obligatoire
5. Références

# Les principaux types de partitionnement

- Partitionnement basé sur
  - les centroïdes (K-moyennes, CURE, ...)
  - la connectivité (hiérarchique, ...)
  - la distribution (BFR, ...)
  - la densité (DBSCAN, OPTICS, ...)
  - les grilles
- Et d'autres ....

# Les principaux types de partitionnement

- Partitionnement basé sur
  - les centroïdes (**K-moyennes**, CURE, ...) ➡
  - la connectivité (**hiérarchique**, ...) ➡
  - la distribution (BFR, ...)
  - la densité (**DBSCAN**, OPTICS, ...) ➡
  - les grilles
- Et d'autres ....

# Sommaire

1. Introduction
2. Les principaux types de partitionnement
3. Applications
4. Lecture obligatoire
5. Références

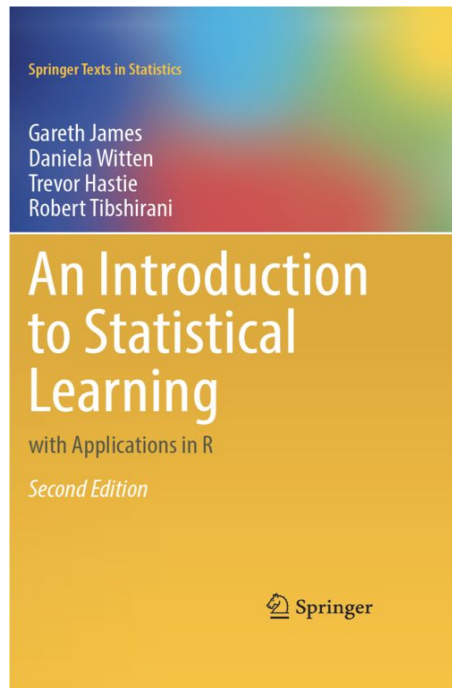
# Applications du partitionnement

- Optimisation de portfolio
- Centres de données
- Astrophysique
- Cybersécurité
- Marketing
- Génétique
- Télédétection
- ...

# Sommaire

1. Introduction
2. Les principaux types de partitionnement
3. Applications
4. Lecture obligatoire
5. Références

# Lecture obligatoire



- Introduction to Statistical Learning with Applications in R  
Second edition (2021)

**Lire le chapitre 12.4 Clustering Methods**  
**Pages 516-532**

# Sommaire

1. Introduction
2. Les principaux types de partitionnement
3. Applications
4. Lecture obligatoire
5. Références



# Références

[1] CS229: Machine Learning - Stanford University

[2] Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani, "Introduction to Statistical Learning with Applications in R - Second edition"

[3] [Popular Unsupervised Clustering Algorithms \(Kaggle\)](#)