

# Chapitre 1: L'analyse de données dans l'industrie



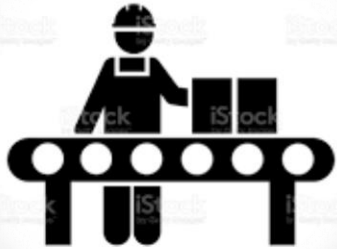
Présentée par:

Dr Imane NEDJAR

# Introduction

Dans l'industrie des grandes quantités de données sont générées de plus en plus. Ces données brute provenant de différents domaines

**Le Contrôle de qualité**



**Conception**

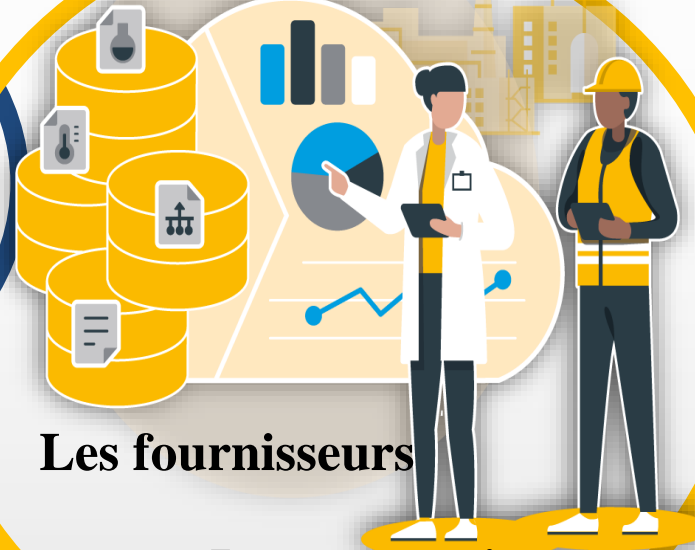
**Le processus d'assemblage**

**Les moteurs**



**Les compresseurs**

**Les convoyeurs**



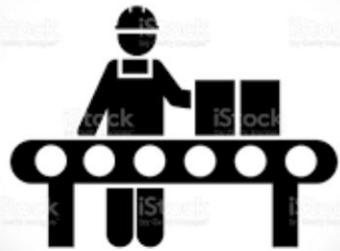
**Les fournisseurs**

**Les partenaires**

# Introduction

Ces données ont rarement été utilisées pour des analyses détaillées, elles étaient utilisées uniquement pour les contrôles et dans le processus techniques de routine d'exécutions de journaux

**Le Contrôle de qualité**



**Conception**

**Le processus d'assemblage**

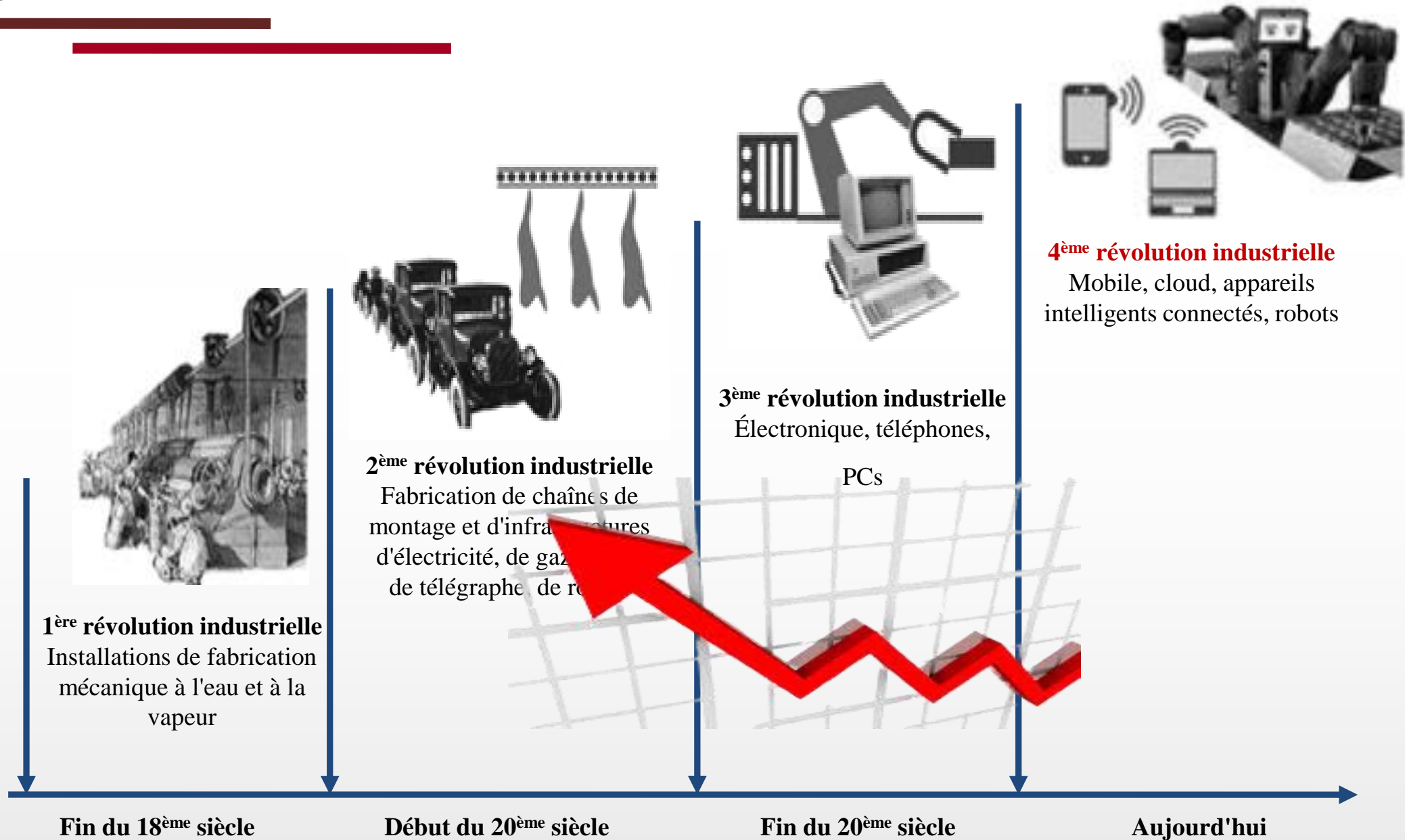
**Les moteurs**



**Les fournisseurs**

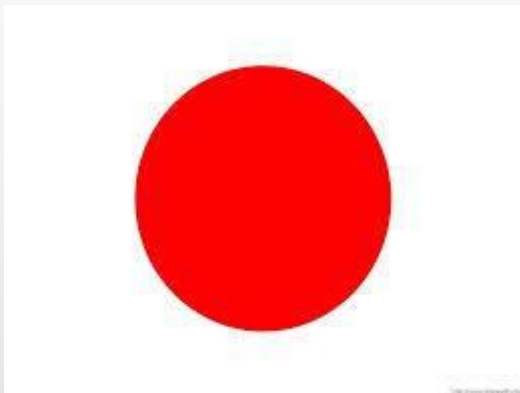
**Les partenaires**

# Introduction



# Introduction

## Usines de fabrication de produits intelligents



# Introduction

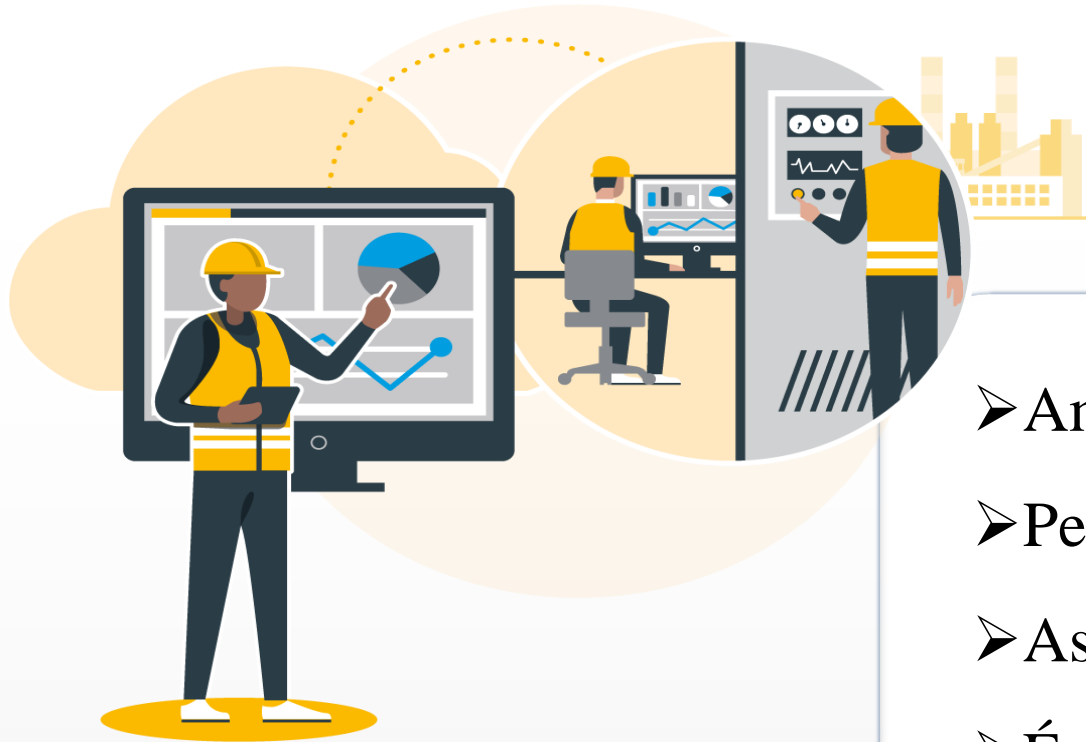
Avec ce large flux de données, la prise de conscience de l'importance de l'extraction d'informations à partir de données a joué un rôle de premier plan dans l'industrie

## Les techniques d'analyses de données



Consiste à extraire des informations utiles contenu dans des données brutes et les transférer vers des connaissances efficaces pour améliorer la compréhension et l'aides à la décision.





- Améliorer la fabrication
- Personnalisez la conception des produits
- Assurer une meilleure assurance qualité
- Évaluer tout risque potentiel



## Analyses de données

```
graph TD; A[Analyses de données] --> B[Les modèles statistiques]; A --> C[Les modèles classification]; A --> D[Les modèles factorielles];
```

### *Les modèles statistiques*

Sont utilisés pour nettoyer les données au début par l'élimination des valeurs aberrantes, et aussi de visualiser les données, afin de construire l'ensemble initial d'exemples.

### *Les modèles classification*

Construisent des règles et des modèles prédictifs pour synthétiser et structurer l'information contenue dans des données.

### *Les modèles factorielles*

Cherchent à réduire le nombre de variables en les résumant par un petit nombre de composantes synthétiques en utilisant essentiellement des outils de l'algèbre linéaire.

### Analyses de données

```
graph TD; A[Analyses de données] --> B[L\'analytique descriptif]; A --> C[L\'analytique prédictif]; A --> D[L\'analytique prescriptif];
```

#### *L\'analytique descriptif*

Fournit des informations sur ce qui a été réalisé et permet donc de comprendre ce qui s'est produit

#### *L\'analytique prédictif*

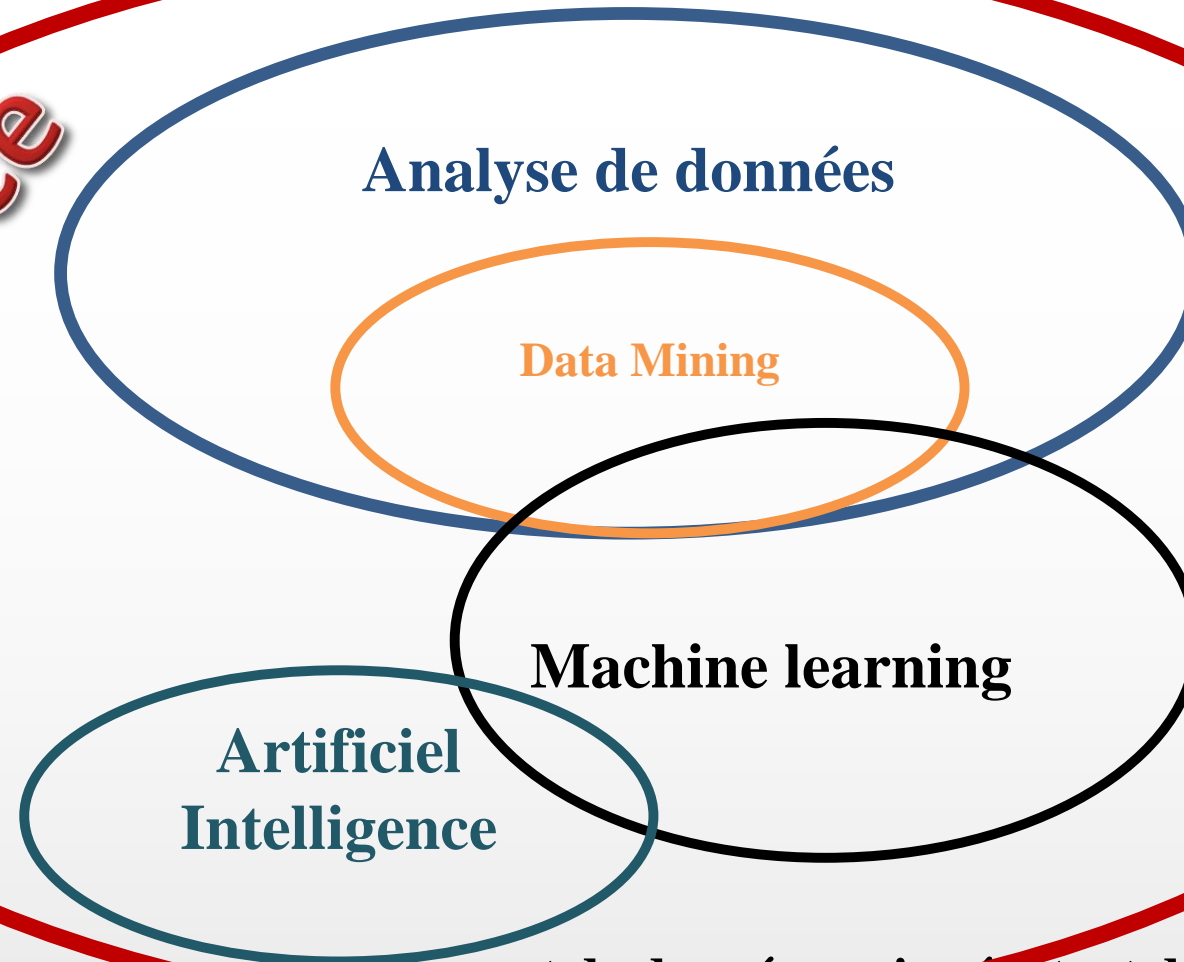
A pour objectif de fournir des modèles afin de prévoir ce qui pourrait se réaliser

#### *L\'analytique prescriptif*

Aide à choisir la meilleure solution entre plusieurs actions possibles pour orienter ce qui va se réaliser

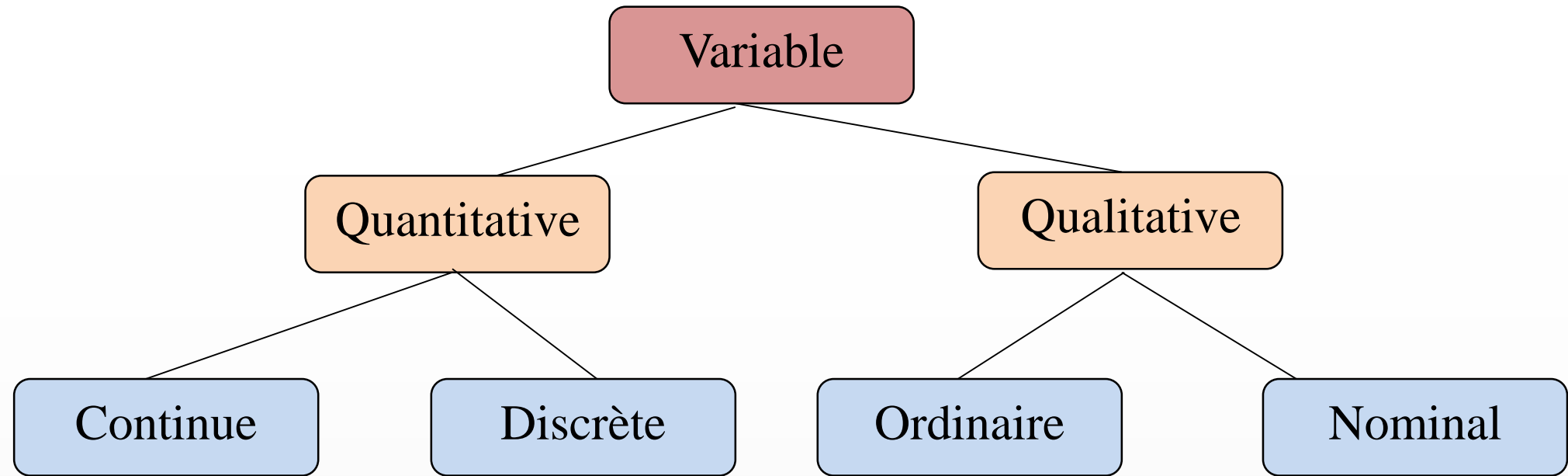
Domaine d'études qui donne aux ordinateurs la capacité d'apprendre sans être explicitement programmés

Consiste à explorer (ou fouiller) les données. Il permet d'établir des associations et relations entre les données (on parle de patterns) qui sont cachées ou non évidentes, très souvent réparties sur plusieurs bases de données relationnelles



Consacrée au développement de systèmes de traitement de données qui exécutent des fonctions normalement associées à l'intelligence humaine, telles que le raisonnement, l'apprentissage et l'auto-amélioration de données

**Tout un ensemble de méthodes mathématiques et informatiques**



Le nombre de valeurs possibles d'une variable est très grand, alors on peut la considérer comme continue. Sinon, on la considère comme **discrète poids, la taille, la distance.**

**Nombre de pièces, nombre des employés**

Ses modalités peuvent être ordonnées.  
**Exemple échelle de satisfaction client : Très satisfaisant, satisfaisant, peu satisfaisant**

On ne peut pas les classer,  
**Exemple: couleur des produit, sexe des clients, type de produit.**

<b>Objectif</b>	<b>Variables quantitative</b>	<b>Variables qualitative/mixtes</b>
Repérer et visualiser les corrélations multiples entre variables et/ou les ressemblances entre individus	Analyse en composantes principales (ACP)	Analyse factorielle des correspondances (AFC) et Analyse factorielle des Correspondances Multiples , (AFCM)
Réaliser une typologie des individus	Méthodes de classification	AFC ou AFCM et classification
Caractériser de groupes d'individus à l'aide de variables	Analyse factorielle discriminante (AFD)	Analyse factorielle discriminante (AFD)

Les données se présentent généralement sous la forme d'un tableau

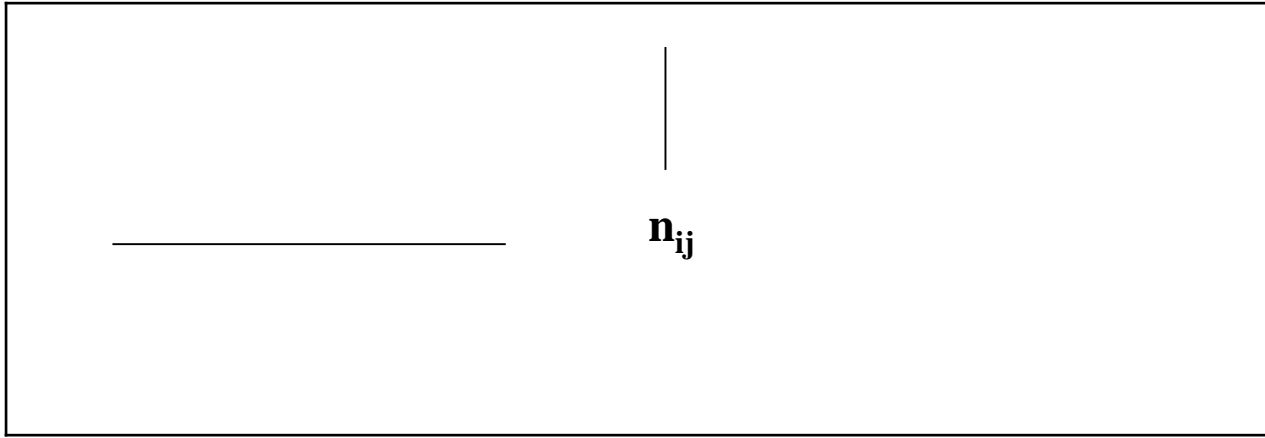
- Les lignes correspondent à des individus ou objets
- Les colonnes correspondent à des attributs ou caractéristiques.

### Le tableau de variables quantitatives

		Variables		
		$X_1$	$X_j$	$X_n$
individus	1			
	.....			
	i			
	.....			
	N			

La méthode d'analyse factorielle qui permet de traiter ce tableau porte le nom d'analyse en composantes principales : ACP

### Le tableau de contingence

		Modalités de Y		
		1	j	n
Modalités de X	1			
	.....			
	i			
	.....			
	N			
		<b>M</b>		

**La méthode d'analyse factorielle qui permet de traiter ce tableau porte le nom d'analyse factorielle des correspondances (AFC)**

## Les tableaux d'enquêtes ou de sondages

Les réponses des  $N$  individus aux questions sont codées en affectant un chiffre aux diverses possibilités de réponses. Le tableau codé est alors transformé en un tableau disjonctif complet. C'est-à-dire un tableau qui ne présente que des 0 ou des 1.

### Questionnaire

Individus	Sexe	Nationalité	Couleur yeux
	H	A	Noir
	F	E	Bleu
	F	A	Marron
	H	A	Noir

### Tableau disjonctif

Homme	Femme	Algerien	Etranger	Yeux bleu	Yeux Noir	Yeux marron
1	0	1	0	0	1	0
0	1	0	1	1	0	0
0	1	1	0	0	0	1
1	0	1	0	0	1	0

**On applique à ce tableau une AFC et la méthode s'appelle analyse factorielle des correspondances multiples (AFCM).**