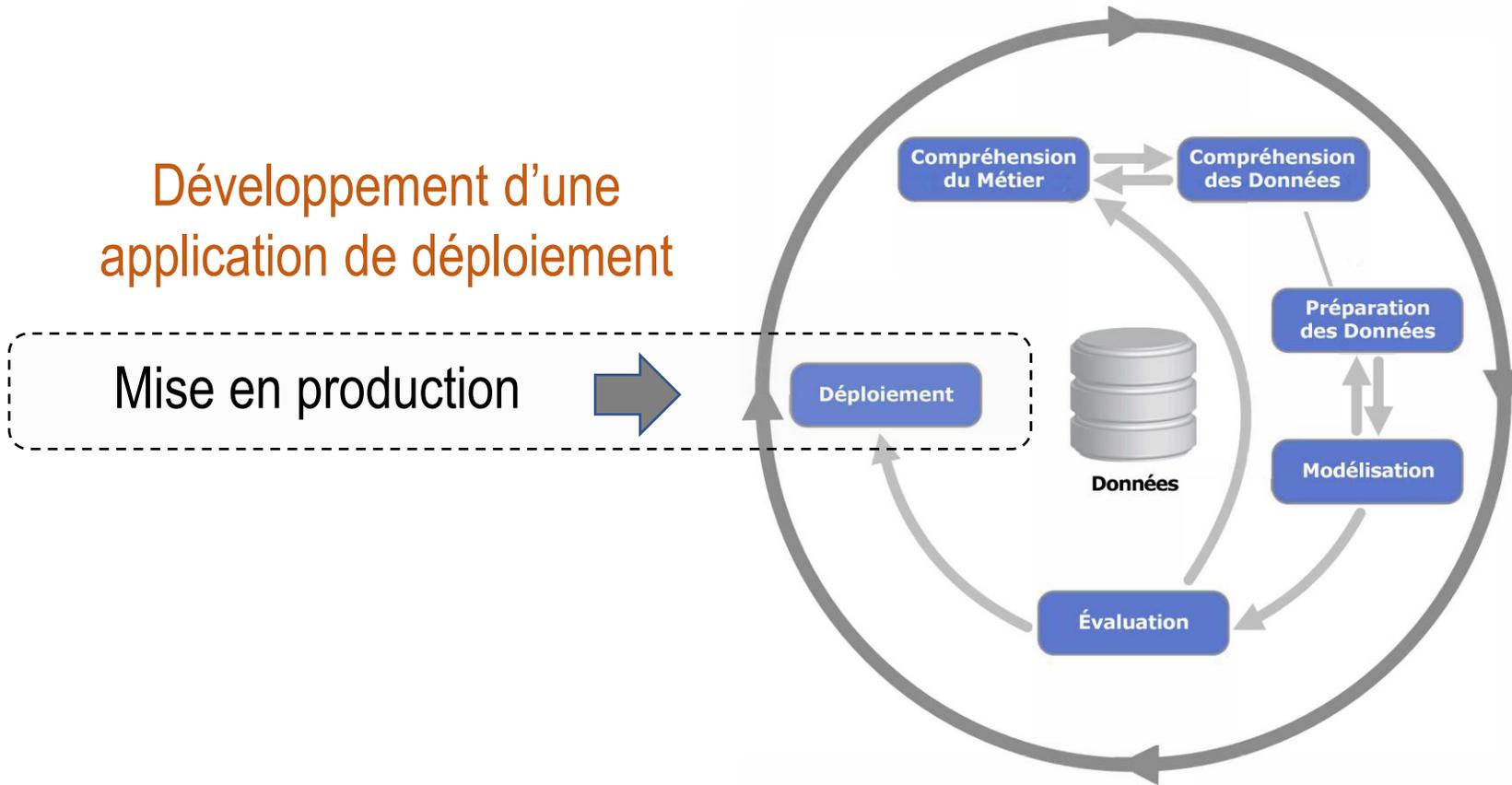


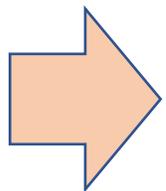
# CRISP-DM

[https://fr.wikipedia.org/wiki/Cross\\_Industry\\_Standard\\_Process\\_for\\_Data\\_Mining](https://fr.wikipedia.org/wiki/Cross_Industry_Standard_Process_for_Data_Mining)



Développement d'une application de déploiement

Mise en production

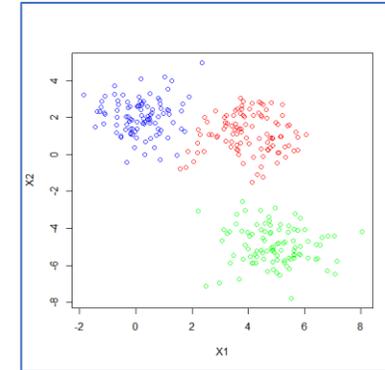
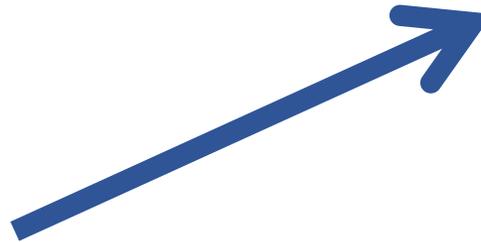
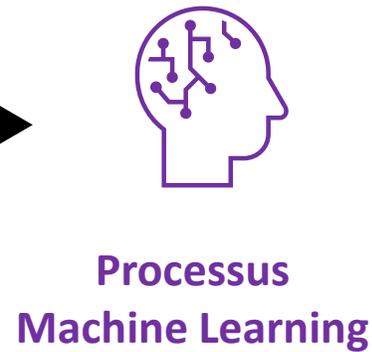
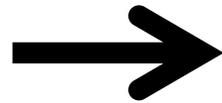
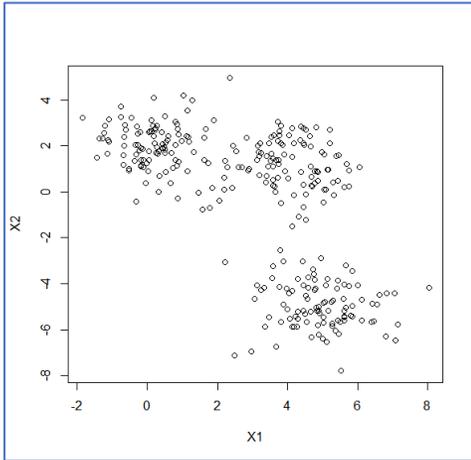


Que l'on retrouve de manière plus prégnante aujourd'hui avec la notion « **Workflow Machine Learning** », de « **cycle de vie** » d'un projet ML qui évolue par « **itérations** »

<https://ml-ops.org/content/end-to-end-ml-workflow>

# Clustering : modélisation, interprétation, exportation du « modèle »

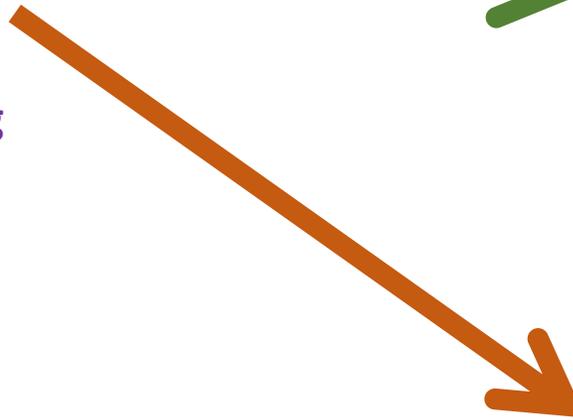
Données décrites par un ensemble de variables (actives)



**Identification  
des « clusters »**

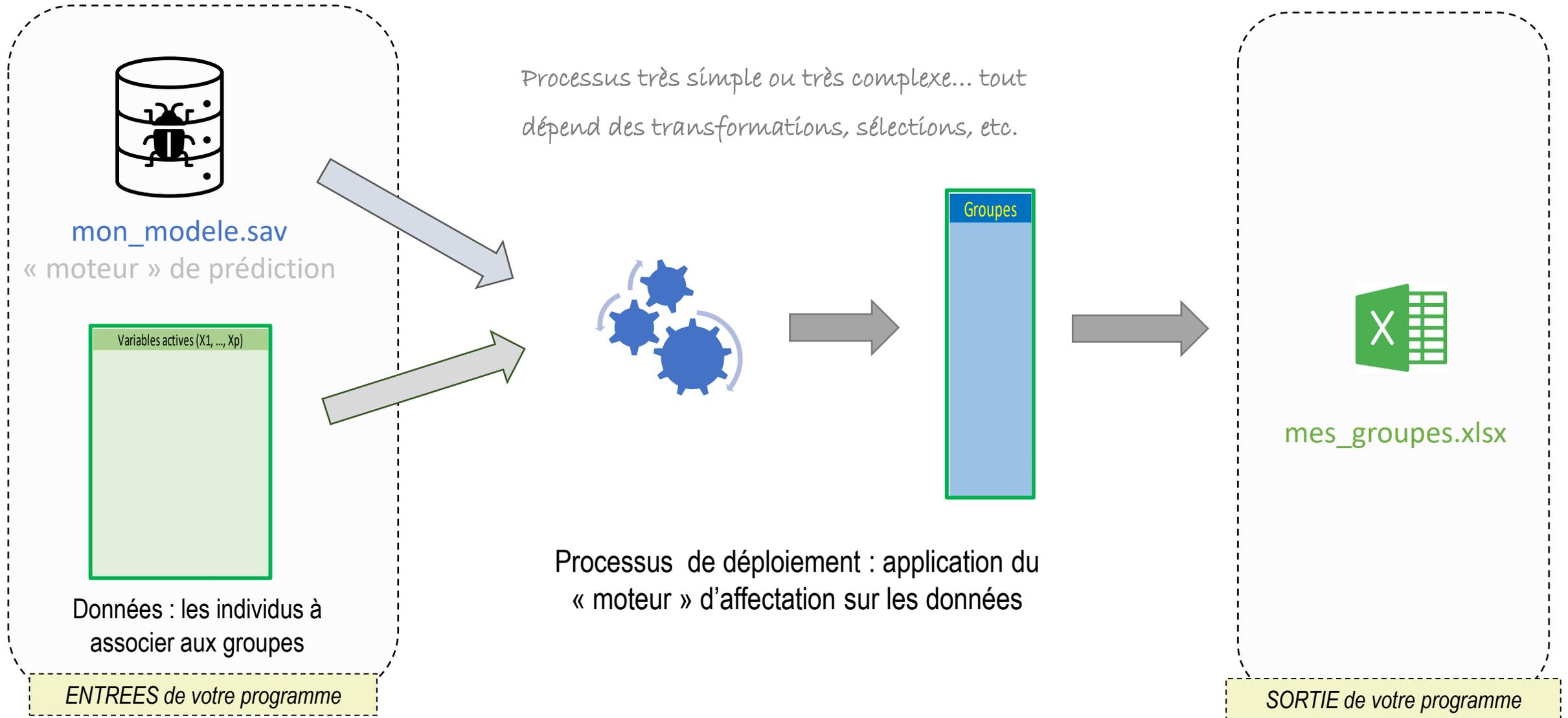


**Interprétation  
des « clusters »**



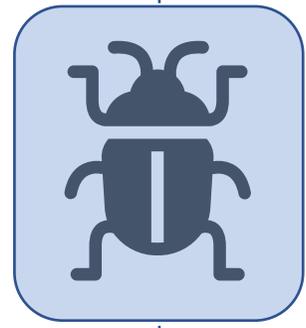
**« Programme »  
(application) de  
déploiement**

# Programme de déploiement

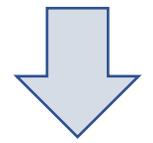


« Modèle » prédictif :  
enchaînement d'opérations

Pipeline

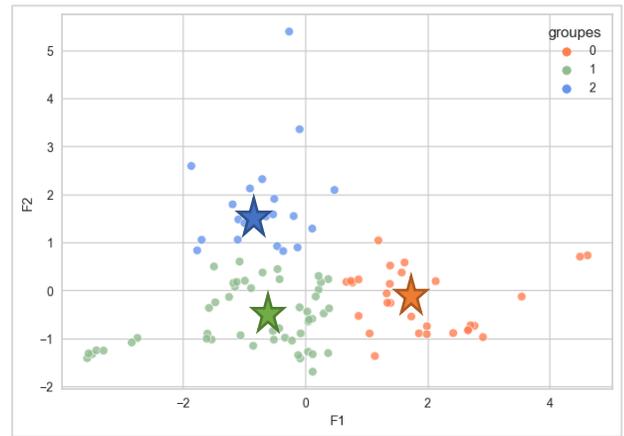


$(X_1, X_2, \dots, X_p)$

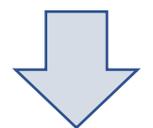


Analyse des correspondances multiples  
Réduction de dimension

$(Z_{(1)}, Z_{(2)}, \dots, Z_{(q)})$   $q \ll p$



Algorithme des K-Means



« Modèle » d'affectation