

Restitution séjour LORIA – décembre 2021

L'analyse de données symboliques repose la théorie des ensembles et les notions :

- d'**individu** : les objets qui sont décrits dans la base
- d'**attribut** : un élément de description élémentaire : une valeur quantitative ou qualitative, un intervalle, une variation. . .
- de **description** : un ensemble d'attributs qui sont utilisés pour décrire les individus
- une fonction d'incidence : la fonction qui associe à un individu une description
- de **concept** : une classe d'individus qui possèdent des attributs communs

Restitution séjour LORIA – décembre 2021

PREMIERS RESULTATS SUR :

- 1) Causal Formal Concept Analysis (Bazin 2021)
- 2) Biclustering :
 - BicPAMS (Madeira & Henriques, <https://web.ist.utl.pt/rmch/bicpams/>)
 - Python (Juniarta 2021)
- 3) Associations

1) Causal Formal Concept Analysis - concept

Studying causality from data involves two related tasks:

- discovering the causal structure (is there a causal relation between these two sets of variables?)
- inferring causal effects (how does the cause causes the effect?)

- v_1 causes v_2 : $\{v_1\} \xrightarrow{c} \{v_2\}$
- v_3 causes v_4 : $\{v_3\} \xrightarrow{c} \{v_4\}$
- v_1 and v_4 together cause v_5 : $\{v_1, v_4\} \xrightarrow{c} \{v_5\}$

1) Causal Formal Concept Analysis – premiers résultats

FCA_Expe_agro_05 - Bloc-notes

Fichier Edition Format Affichage Aide

303.46402645111084 seconds

13 -> 12

11 -> 9

10 -> 13

10 12 13 -> 0 8 4

9 -> 7

8 -> 4 12

7 -> 11

9 11 12 7 -> 0 8 4 13

12 6 -> 0 4 13

6 7 9 11 14 -> 10 13

4 -> 10 12 13

2 12 -> 4 13

1 12 -> 0 8 4 13

0 -> 4 12

0 4 8 10 12 13 -> 9 11 6 7

FCA_Expe_agro_07 - Bloc-notes

Fichier Edition Format Affichage Aide

817.7814698219299 seconds

11 -> 9

10 12 14 -> 13

10 12 13 -> 0 4

9 -> 7

8 13 -> 12 4

8 12 -> 4

8 10 -> 4

7 -> 11

9 11 12 7 -> 4

7 9 10 11 13 14 -> 4

13 6 -> 4

12 6 -> 4

6 7 9 10 11 -> 13

4 -> 12

4 12 -> 10

10 4 12 -> 0 13

2 12 -> 4

1 12 -> 4

0 13 -> 12 4

0 12 -> 4

0 10 -> 4 12

0 8 14 -> 12

0 7 9 11 14 -> 12

0 7 8 9 11 -> 4 12

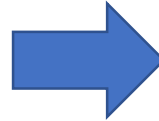
0 8 6 -> 4

0 6 7 9 11 -> 4

0 4 10 12 13 -> 6

0 4 6 10 12 13 -> 9 11 7

0 4 6 7 9 10 11 12 13 -> 8



**0 – 14 = nombre attribué à
chaque colonne du dataset**

0 – 9 = attributs climatiques
10 – 14 = attributs plante

Quelles sont les relations qui
ont vraiment du sens du
point de vue agronomique?

i.e. un facteur climatique doit
causer une variable plante, le
contraire n'est pas possible!

1) Causal Formal Concept Analysis – premiers résultats



FCA_Expe_agro_05 - Bloc-notes

Fichier Edition Format Affichage Aide

303.46402645111084 seconds

Les causalités les plus intéressantes

13 -> 12

Lipids -> PMG

11 -> 9

10 -> 13

Seed_yield -> lipids

10 12 13 -> 0 8 4

9 -> 7

8 -> 4 12

Harvest_25 -> PMG

7 -> 11

P1000_30 -> seed_nb

9 11 12 7 -> 0 8 4 13

12 6 -> 0 4 13

6 7 9 11 14 -> 10 13

4 -> 10 12 13

P600_25 -> seed_yield, PMG, lipids

2 12 -> 4 13

1 12 -> 0 8 4 13

0 -> 4 12

Ptot_25 -> PMG

0 4 8 10 12 13 -> 9 11 6 7

Significat des colonnes

0 = Ptot_25

1 = Ptot_30

2 = P300_25

3 = P300_30

4 = P600_25

5 = P600_30

6 = P1000_25

7 = P1000_30

8 = harvest_25

9 = harvest_30

10 = seed_yield

11 = seed_nb

12 = PMG

13 = lipids

14 = proteins

1) Causal Formal Concept Analysis – premiers résultats

FCA_Expe_agro_07 - Bloc-notes

Fichier Edition Format Affichage Aide

817.7814698219299 seconds

11 -> 9

10 12 14 -> 13

Les causalités les plus intéressantes

Seed_yield, PMG, proteins -> lipids

10 12 13 -> 0 4

9 -> 7

8 13 -> 12 4

Harvest_25, lipids -> PMG

8 12 -> 4

8 10 -> 4

7 -> 11

P1000_30 -> seed_nb

9 11 12 7 -> 4

7 9 10 11 13 14 -> 4

13 6 -> 4

12 6 -> 4

6 7 9 10 11 -> 13

P1000_25, P1000_30, harvest_30, seed_yield, seed_nb -> lipids

4 -> 12

P600_25 -> PMG

4 12 -> 10

P600_25, PMG -> seed_yield

10 4 12 -> 0 13

2 12 -> 4

1 12 -> 4

0 13 -> 12 4

Ptot_25, lipids -> PMG

0 12 -> 4

0 10 -> 4 12

Ptot_25, seed_yield -> PMG

0 8 14 -> 12

0 7 9 11 14 -> 12

0 7 8 9 11 -> 4 12

0 8 6 -> 4

Significat des colonnes

0 = Ptot_25

1 = Ptot_30

2 = P300_25

3 = P300_30

4 = P600_25

5 = P600_30

6 = P1000_25

7 = P1000_30

8 = harvest_25

9 = harvest_30

10 = seed_yield

11 = seed_nb

12 = PMG

13 = lipids

14 = proteins

2) Biclustering - concept

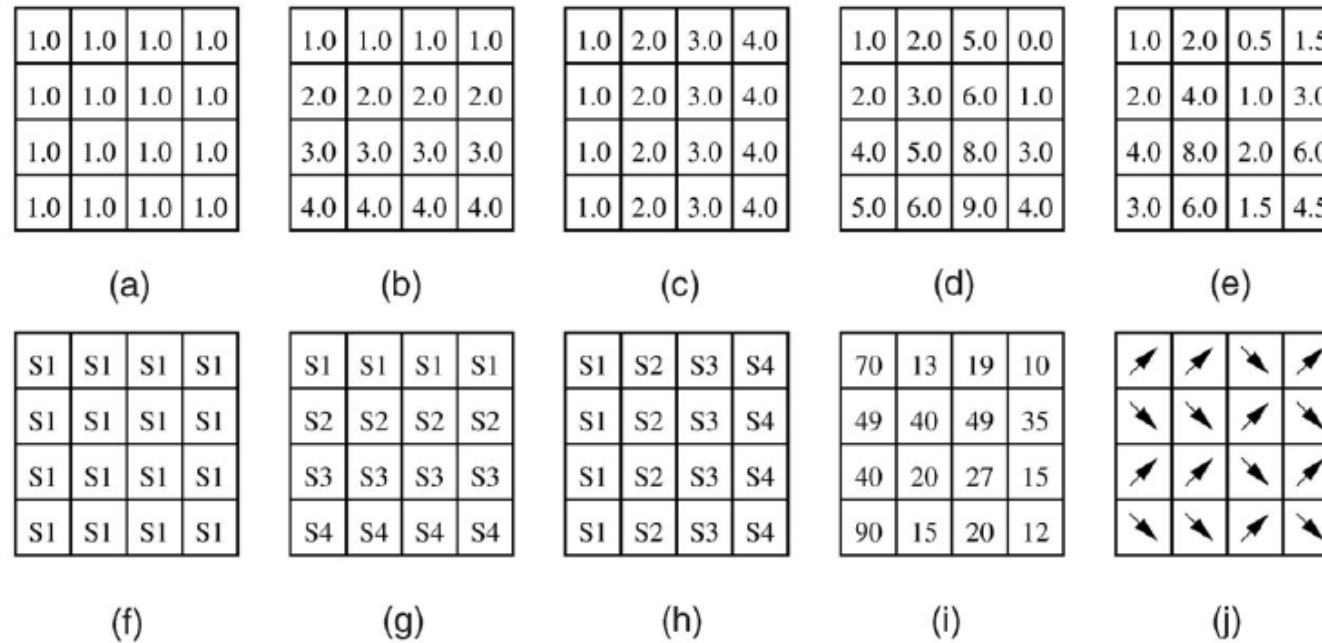
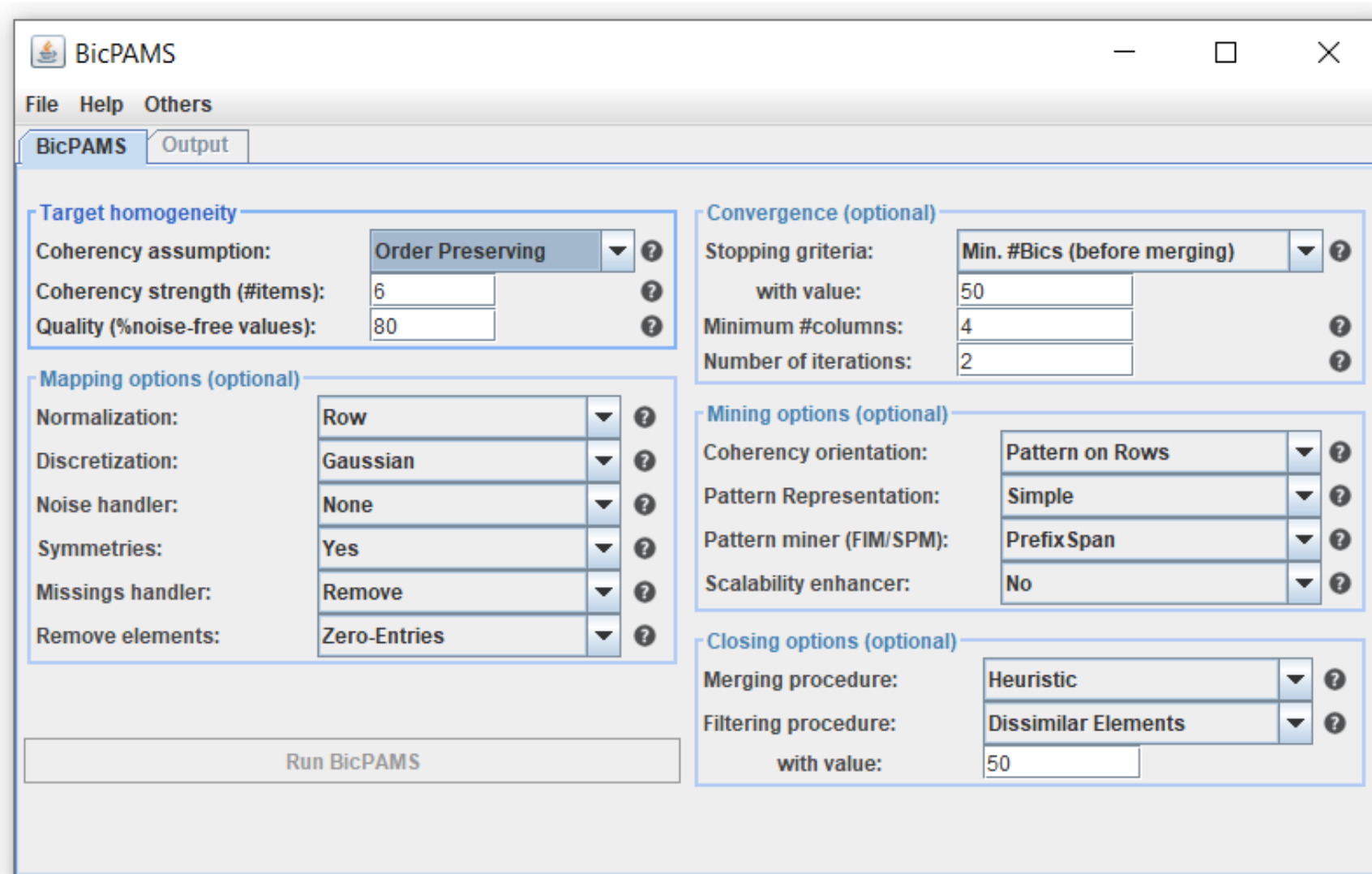


Fig. 1. Examples of different types of biclusters. (a) Constant bicluster, (b) constant rows, (c) constant columns, (d) coherent values (additive model), (e) coherent values (multiplicative model), (f) overall coherent evolution, (g) coherent evolution on the rows, (h) coherent evolution on the columns, (i) coherent evolution on the columns, and (j) coherent sign changes on rows and columns.

2) Biclustering – software BicPAMS



The screenshot displays the BicPAMS software window, which includes a menu bar (File, Help, Others) and two tabs: BicPAMS and Output. The main area is divided into several sections for configuring the biclustering process:

- Target homogeneity:**
 - Coherency assumption: Order Preserving (dropdown)
 - Coherency strength (#items): 6 (text input)
 - Quality (%noise-free values): 80 (text input)
- Mapping options (optional):**
 - Normalization: Row (dropdown)
 - Discretization: Gaussian (dropdown)
 - Noise handler: None (dropdown)
 - Symmetries: Yes (dropdown)
 - Missings handler: Remove (dropdown)
 - Remove elements: Zero-Entries (dropdown)
- Convergence (optional):**
 - Stopping criteria: Min. #Bics (before merging) (dropdown)
 - with value: 50 (text input)
 - Minimum #columns: 4 (text input)
 - Number of iterations: 2 (text input)
- Mining options (optional):**
 - Coherency orientation: Pattern on Rows (dropdown)
 - Pattern Representation: Simple (dropdown)
 - Pattern miner (FIM/SPM): PrefixSpan (dropdown)
 - Scalability enhancer: No (dropdown)
- Closing options (optional):**
 - Merging procedure: Heuristic (dropdown)
 - Filtering procedure: Dissimilar Elements (dropdown)
 - with value: 50 (text input)

A large button labeled "Run BicPAMS" is located at the bottom left of the configuration area.

2) Biclustering v1 – Constant columns avec BicPAMS

(5,35) X=[43, 46, 47, 49, 53, 54, 55, 56, 57, 58, 60, 61, 62, 63, 64, 66, 73, 83, 84, 85, 86, 88, 89, 90, 117, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129]
Y=[P300_25, P600_25, P1000_25, P1000_30, proteins]

(4,42) X=[1, 2, 3, 4, 5, 6, 24, 28, 30, 31, 32, 33, 34, 35, 37, 38, 39, 40, 41, 55, 56, 57, 58, 60, 61, 62, 63, 64, 65, 66, 74, 81, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129]
Y=[P300_25, P1000_25, P1000_30, PMG]



indiv	P300_25	P600_25	P1000_25	P1000_30	seed_nb	PMG	proteins
30	0	1	2	0	59359,76	4,42	24,25
31	0	0	2	0	95067,14	4,22	21,71
32	0	0	2	0	96205,64	4,47	20,15
33	0	0	2	0	90141,78	4,39	20,64
34	0	0	2	0	76851,13	4,34	22,12
35	0	0	2	0	81840,22	4,23	20,31
36	0	0	2	0	74565,74	4,56	20,01
37	0	0	2	0	81925,19	4,21	21,82
38	0	0	2	0	66568,32	4,4	22,48
39	0	0	2	0	103089,87	4,38	21,25
40	0	0	2	0	83178,84	4,27	21,29
41	0	0	2	0	84373,37	4,19	22,74
42	0	0	2	0	63619,97	5,17	22,82
43	0	0	2	0	55434,18	5,26	18,5
44	0	0	2	0	58085,01	5,42	20,01
45	0	0	2	0	63651,76	5,34	20,52
46	0	0	2	0	57438,91	5,56	19,29
47	0	0	2	0	63201,48	5,37	19,33
48	0	0	2	0	60579,25	5,39	19,83
49	0	0	2	0	62286,49	4,95	18,87
50	0	0	2	0	64338,34	5,3	19,62
51	0	0	2	0	62343,64	5,37	20,88
52	0	0	2	0	55241,58	5,4	20,69
53	0	0	2	0	62237,66	5,36	19,38
54	0	0	2	0	67559,34	4,57	18,35
55	0	0	2	0	86886,53	4,23	18,39

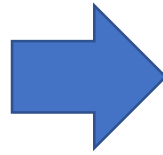
Points d'amélioration

- Valeurs des attributs plante trop contrastants (entre 4 et 100.000)
- P300_30 et P600_30 sont toujours à 0, donc souvent présents dans les biclusters (pas d'intérêt pratique)
- Sur BicPAMS pas assez précis (lipides = 24 et 35 dans le même bicluster...)

2) Biclustering v1 – Constant columns avec Python

11 x 5 bicluster

	P300_25	P300_30	P600_25	P600_30	PMG
0	0	0	1	0	4
1	0	0	1	0	4
2	0	0	1	0	4
3	0	0	1	0	4
4	0	0	1	0	4
5	0	0	1	0	4
109	0	0	1	0	4
23	0	0	1	0	4
27	0	0	1	0	4
29	0	0	1	0	4
110	0	0	1	0	4



11 x 5 bicluster

	P300_25	P300_30	P600_30	harvest_30	lipids
0	0	0	0	2	51
1	0	0	0	2	51
2	0	0	0	2	51
65	0	0	0	2	51
96	0	0	0	2	51
54	0	0	0	2	51
55	0	0	0	2	51
56	0	0	0	2	51
57	0	0	0	2	51
59	0	0	0	2	51
60	0	0	0	2	51

Points d'amélioration

- Valeurs des attributs plante trop contrastants (entre 4 et 100.000)
- P300_30 et P600_30 sont toujours à 0, donc souvent présents dans les biclusters (pas d'intérêt pratique)
- Sur Python c'est trop précis (lipides = 51)

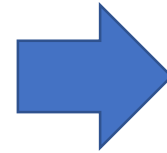
2) Biclustering v2 – Constant columns avec Python

7 x 6 bicluster

	Ptot_30	P300_25	P1000_30	harvest_25	harvest_30	lipids
37	0	0	0	2	0	46.68
40	0	0	0	2	0	46.82
111	0	0	0	2	0	47.1
18	0	0	0	2	0	46.91
19	0	0	0	2	0	46.64
30	0	0	0	2	0	47.16
115	0	0	0	2	0	47.2

4 x 7 bicluster

	Ptot_30	P300_25	P1000_30	harvest_25	harvest_30	PMG	proteins
18	0	0	0	2	0	4.6	21.4
19	0	0	0	2	0	5.06	21.31
118	0	0	0	2	0	4.73	21
111	0	0	0	2	0	4.88	20.8



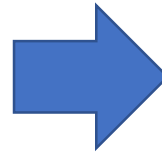
Changements

- Exclusion de P300_30 et P600_30
- Pas d'arrondissement (important pour les valeurs petites comme le PMG)
- Augmentation de l'échelle de variation accepté pour les paramètres plante

2) Biclustering v3 – Multiplicative avec Python

6 x 3 bicluster

	P600_25	P1000_25	proteins
129	0	8	16.40
132	0	8	16.90
133	0	8	17.00
9	0	11	23.17
10	0	11	22.94
11	0	11	22.99



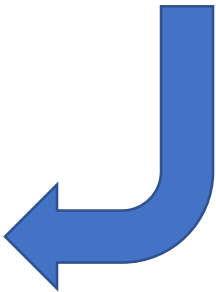
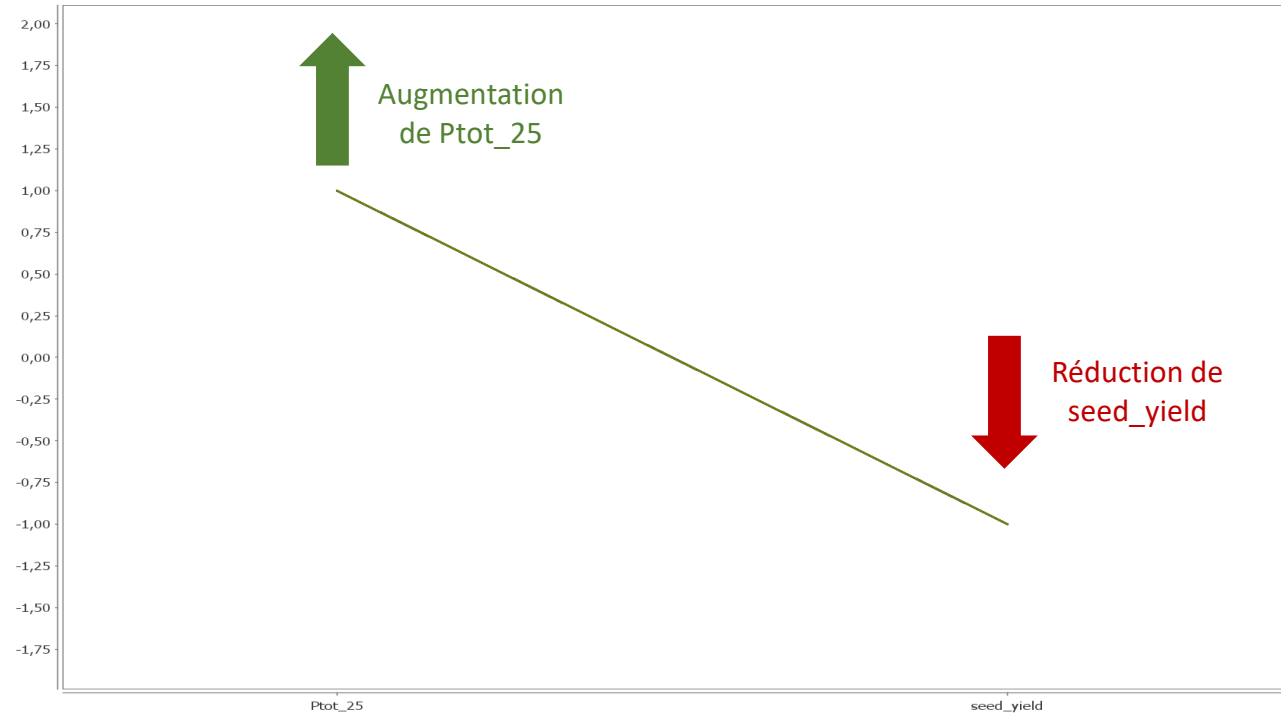
11 x 5 bicluster

	Ptot_30	P600_25	P1000_25	harvest_30	PMG
6	1	0	11	1	3.04
7	1	0	11	1	2.90
8	1	0	11	1	2.51
9	1	0	11	1	3.00
10	1	0	11	1	2.87
11	1	0	11	1	3.03
12	1	0	9	1	2.79
14	1	0	9	1	2.77
15	1	0	9	1	2.61
16	1	0	9	1	2.68
17	1	0	9	1	2.66

Les relations changent dans la même proportion entre les colonnes du bicluster





2) Biclustering v4 – Binary constant avec BicPAMS

	Ptot_25	Ptot_30	P300_25	P600_25	P1000_25	P1000_30	harvest_25	harvest_30	seed_yield	seed_nb	PMG	lipids	proteins
0->6	1.0	-1.0	1.0	-1.0	1.0	1.0	0.0	-1.0	-1.0	-1.0	-1.0	-1.0	1.0
0->7	1.0	-1.0	1.0	-1.0	1.0	1.0	0.0	-1.0	-1.0	-1.0	-1.0	-1.0	1.0
0->8	1.0	-1.0	1.0	-1.0	1.0	1.0	0.0	-1.0	-1.0	1.0	-1.0	-1.0	1.0
0->9	1.0	-1.0	1.0	-1.0	1.0	1.0	0.0	-1.0	-1.0	-1.0	-1.0	-1.0	1.0
0->10	1.0	-1.0	1.0	-1.0	1.0	1.0	0.0	-1.0	-1.0	1.0	-1.0	-1.0	1.0
0->11	1.0	-1.0	1.0	-1.0	1.0	1.0	0.0	-1.0	-1.0	-1.0	-1.0	-1.0	1.0
0->12	1.0	-1.0	1.0	-1.0	1.0	0.0	1.0	-1.0	-1.0	1.0	-1.0	-1.0	1.0
0->13	1.0	-1.0	1.0	-1.0	1.0	0.0	1.0	-1.0	-1.0	1.0	-1.0	-1.0	1.0
0->14	1.0	-1.0	1.0	-1.0	1.0	0.0	1.0	-1.0	-1.0	1.0	-1.0	-1.0	1.0
0->15	1.0	-1.0	1.0	-1.0	1.0	0.0	1.0	-1.0	-1.0	1.0	-1.0	-1.0	1.0



Cette approche permet de voir les tendances de croissance (1), constance (0) et réduction (-1) entre 2 objets (lignes)

3) Associations - concept

	G \ M	a	b	c	d
1		x			x
2		x		x	
3			x	x	
4			x	x	x

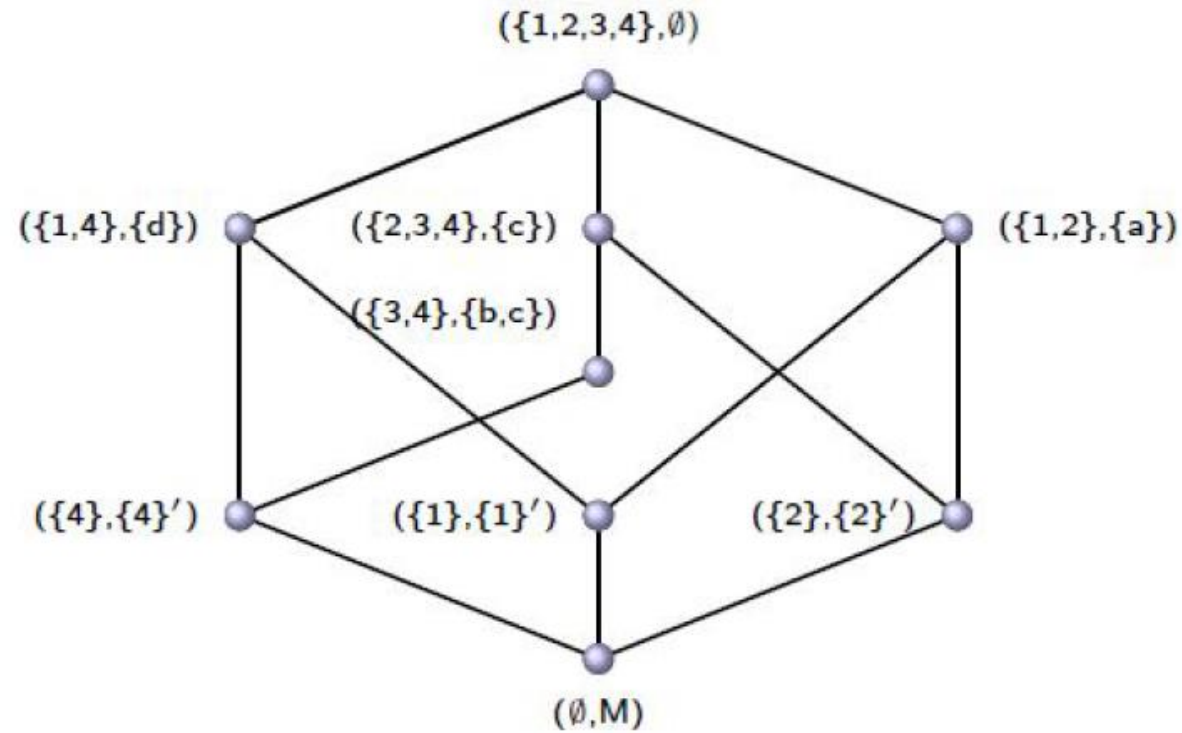
Objects:





- 1 – equilateral triangle
- 2 – rectangle triangle
- 3 – rectangle
- 4 – square

Attributes:

- a** – has 3 vertices
- b** – has 4 vertices
- c** – has a direct angle
- d** – equilateral

3) Associations - exemple



	G \ M	a	b	c	d
1		x			x
2		x		x	
3			x	x	
4			x	x	x

- a – has 3 vertices
- b – has 4 vertices
- c – has a direct angle
- d – equilateral

3) Associations – exemple climat

Exemple de règles d'association sur les variations climatiques¹:

year	Rule	Confidence
2013	$\{\text{rainday} \geq 21.6\} \rightarrow \{\text{humidity} < 82.3, \text{temperature} < 27.62, \text{windspeed} < 1.06\}$	100.0
	$\{\text{rainday} \leq 21.6\} \rightarrow \{\text{humidity} < 82.3, \text{temperature} < 27.62\}$	100.0
	$\{\text{rainday} \leq 21.6\} \rightarrow \{\text{humidity} < 82.3, \text{windspeed} < 1.06\}$	100.0
2014	$\{\text{humidity} < 82.5, \text{temperature} < 27.46, \text{windspeed} < 0.94, \text{rainfall} < 624.0\} \rightarrow \{\text{rainday} \leq 21.4\}$	100.0
	$\{\text{humidity} < 82.5, \text{temperature} < 27.46, \text{windspeed} < 0.94\} \rightarrow \{\text{rainday} \leq 21.4\}$	100.0
	$\{\text{humidity} < 82.5, \text{windspeed} < 0.94, \text{rainfall} < 624.0\} \rightarrow \{\text{rainday} \leq 21.4\}$	100.0
2015	$\{\text{humidity} < 74, \text{temperature} < 28.64\} \rightarrow \{17.0 \leq \text{rainday} < 20.0\}$	100.0
	$\{\text{humidity} < 74, \text{windspeed} < 1.2, \text{rainfall} < 437.96\} \rightarrow \{17.0 \leq \text{rainday} < 20.0\}$	100.0
	$\{\text{humidity} < 74\} \rightarrow \{17.0 \leq \text{rainday} < 20.0\}$	75.0

¹Analyzing Climate Variability in Malaysia Using Association Rule Mining, Rabiatul et al., 2018

3) Associations – organisation du jeu de données

G \ M	m1	m2	m3
g1	1	3	4
g2	2	2	3
g3	4	1	1
g4	2	2	1

Nominal Scaling:

G \ M	m1=1	m1=2	m1=4	m2=1	m2=2	m2=3	m3=1	m3=3	m3=4
g1	x					x			x
g2		x			x			x	
g3			x	x			x		
g4		x			x		x		

Interordinal Scaling:

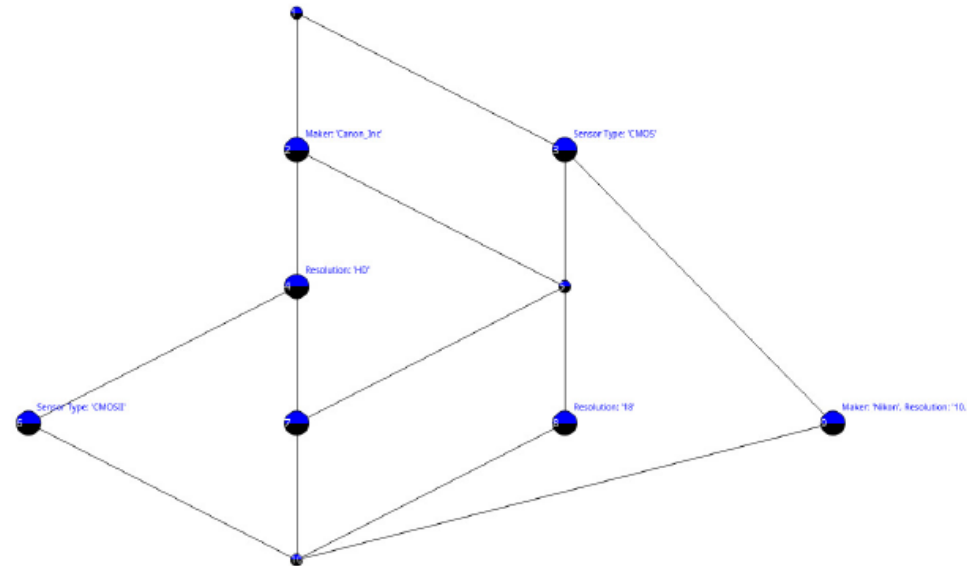
G \ M	m1 \leq 1	m1 \geq 1	m1 \leq 2	m1 \geq 2	m1 \leq 3	m1 \geq 3	m1 \leq 4	m1 \geq 4	m2 \leq 1	m2 \geq 1
g1	x	x	x		x		x			x
g2		x	x	x	x		x			x
g3		x		x		x	x	x	x	x
g4		x	x	x	x		x			x

3) Associations – FCA et Iceberg Lattices

Online tool to perform and visualize FCA results, by Orpailleur team, Loria.

- Binary data or interval patterns
- Multiple algorithms
- Display and manipulate the lattice
- Compute concepts (closed itemsets)
- Compute implications

<https://latviz.loria.fr/>



3) Associations – FCA et Iceberg Lattices

Binary data

Create a context table:

Number of objects: Number of attributes: Create Context

Or choose one example of binary data context:

Or load JSON context file: Aucun fichier sélectionné. ☒ Display context table

Load JSON lattice file (the output of Sofia software): Aucun fichier sélectionné.

Load JSON lattice diagram file (exported from LatViz app): Aucun fichier sélectionné.

	Att 1	Att 2
Obj 1		
Obj 2		
Obj 3		
Obj 4		

Add object Add attribute Delete selected rows Delete selected columns Save context file

Build Concept Lattice (AddIntent) Build Concept Lattice (Ganter) Display Full Concept Lattice Save lattice json file (Sofia output)

Display Concepts By Levels Display AOC Poset Calculate Implications Save Graph Diagram Save Graph Image

☐ Set the filter

☐ Mark new concepts when editing context

☐ Highlight AOC Poset

☐ Show full labelling of lattice

☐ Display objects' name on node ☐ Display attributes' name on node

Type of positioning nodes on graph:

☐ Search on displayed lattice

Context: Number of objects: 4, Number of attributes: 3

Prochaines étapes

1 - Définition du contexte de recherche! Quelles questions veut-on répondre?

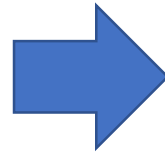
2 - Choix de l'approche qui présente le mieux compromis :

- Pour **Analyse Causal** : travailler sur les relations déjà trouvées avec Alexandre
- Pour **Association** : transformer les données numériques en données symboliques
 - Division du jeu de données en deux catégories (cf profil « cherché » et « pas cherché ») OU
 - Création des intervalles (cf binarisation et/ou formalisation du jeu de données)
- Pour **Bicluster** : décider la meilleure approche (additive, constante, multiplicative?) et travailler sur BicPAMS pour affiner les résultats et/ou contacter Nyoman pour Python
- Pour **autres** : essayer les approches vues chez ACTA (GAM, random forest...)?

Pistes d'amélioration des données d'entrée

Transformation des attributs pour avoir la même dimension numérique

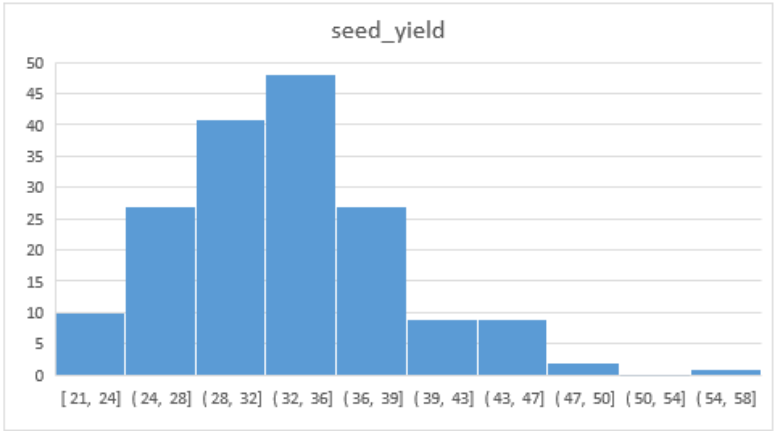
seed_yield	seed_nb	PMG	lipids	proteins
40,4	95386	4,23	51,4	16,4
37,8	84953	4,45	50,8	17,9
41	92816	4,42	51,1	17,7
55,1	140813	3,91	51,8	17,1
39,1	96437	4,05	52,1	16,7
43,4	101070	4,29	51,7	17,4
25,6	84213	3,04	42,29	23,54
27	93070	2,9	41,79	23,83
28,5	113687	2,51	40,64	24,6
26,4	87849	3	41,96	23,17
28,1	97586	2,87	42,01	22,94
25,3	83573	3,03	42,77	22,99
29,3	104919	2,79	43,5	22,69
31,2	108616	2,87	42,27	24,07
34,1	123339	2,77	44,15	22,04



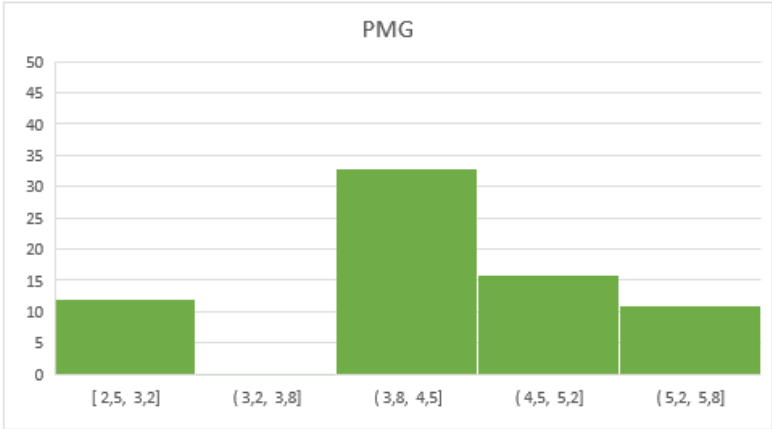
seed_yield	seed_nb/200	PMG*10	lipids	proteins
80800	95386	84600	102800	32800
75600	84953	89000	101600	35800
82000	92816	88400	102200	35400
110200	140813	78200	103600	34200
78200	96437	81000	104200	33400
86800	101070	85800	103400	34800
51200	84213	60800	84580	47080
54000	93070	58000	83580	47660
57000	113687	50200	81280	49200
52800	87849	60000	83920	46340
56200	97586	57400	84020	45880
50600	83573	60600	85540	45980
58600	104919	55800	87000	45380
67400	108616	57400	84540	48140

Pistes d'amélioration des données d'entrée

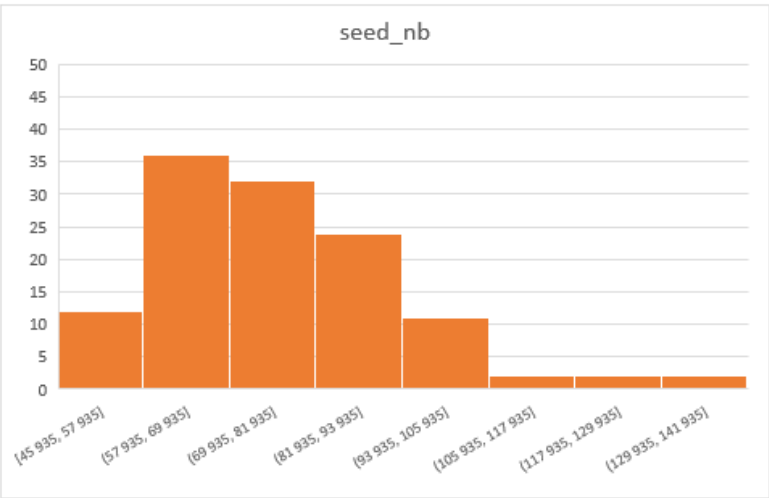
Creation des intervalles d'intérêt avec les données Rapsodyn



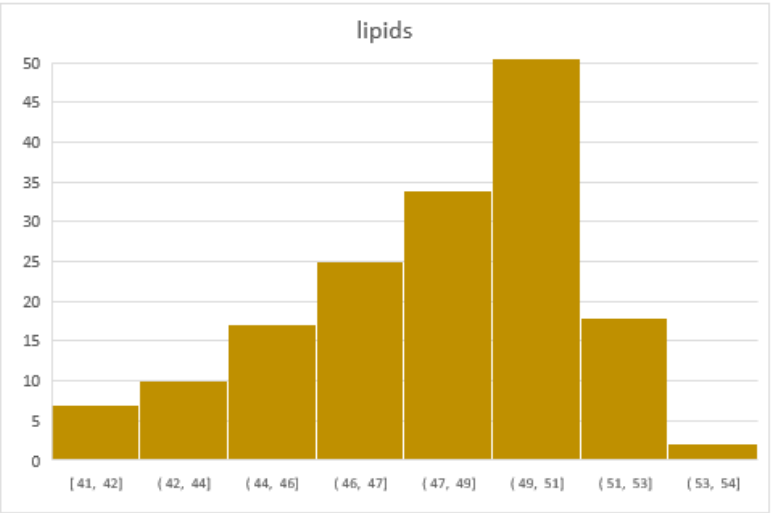
Intervalles	seed_yield
yield_A	<24
yield_B	24-28
yield_C	28-32
yield_D	32-36
yield_E	36-39
yield_F	39-43
yield_G	>43



Intervalles	PMG
pmg_A	<3,8
pmg_B	3,8-4,5
pmg_C	4,5-5,2
pmg_D	>5,2



Intervalles	seed_nb
nb_A	<47000
nb_B	47000-57000
nb_C	57000-67000
nb_D	67000-77000
nb_E	77000-87000
nb_F	87000-97000
nb_G	>97000



Intervalles	lipids
lip_A	<42
lip_B	42-44
lip_C	44-46
lip_D	46-48
lip_E	48-50
lip_F	50-52
lip_G	>52

Pistes d'amélioration des données d'entrée

Séparation du jeu de données en deux classes : profil « performant/recherché » et « moins performant »

- Pour les attributs climatiques, ce n'est pas possible ni intéressant pour le moment.
- Pour les attributs plante*, cela me semble pertinent de chercher :

seed_yield > 30
seed_nb > 55.000
PMG > 4
lipides > 43
protéines > 19

- On pourrait dire que les objets avec des résultats inférieurs à ces chiffres sont "moins intéressants".

* Le profil « performant/recherché » a été défini par rapport à des sources bibliographiques (Yara, Terres Univia) mais aussi en prenant en considération la distribution du jeu de données Rapsodyn. Voir <https://www.yara.fr/fertilisation/blog/construction-rendement-colza/>

Rappel approche

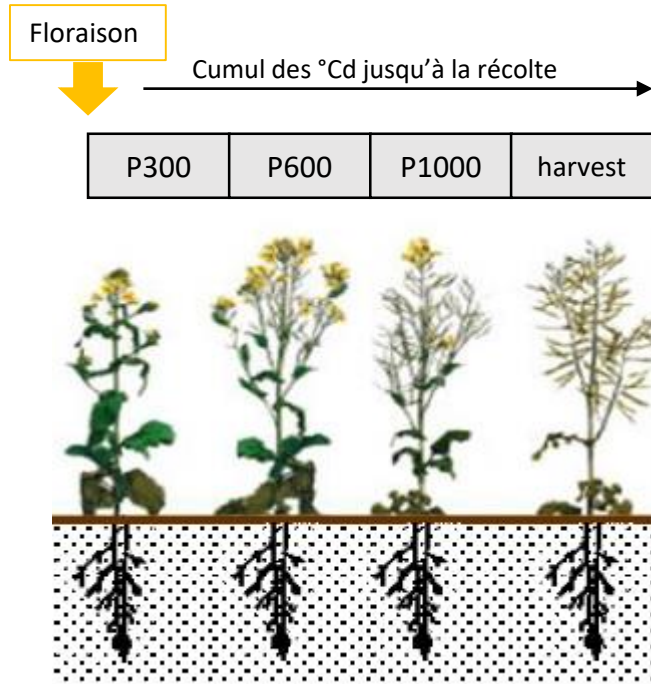
- Rapeseed crops in France, between 2011-2020
- Characterisation of the trials : ITK
- Environmental factors : locality, year, daily max/min/aver temperature, radiation, rainfall
- Plant and initial state variables : genotype, sowing/flowering/harvest dates, N supply
- Final performance variables : grain yield and components (seed number, PMG), oil and protein content
- Pre-sorting already carried out to discard irrelevant data sets (other stresses, diseases, missing data...)
- **Total of 27 combinations harvest year x local, with 174 individuals**



Rappel approche

First step : sorting the data, creating the climate indicators and organizing the datasets

Search for days with **Tmax>25°C** and **Tmax>30°C**
in the pre-defined phenological intervals

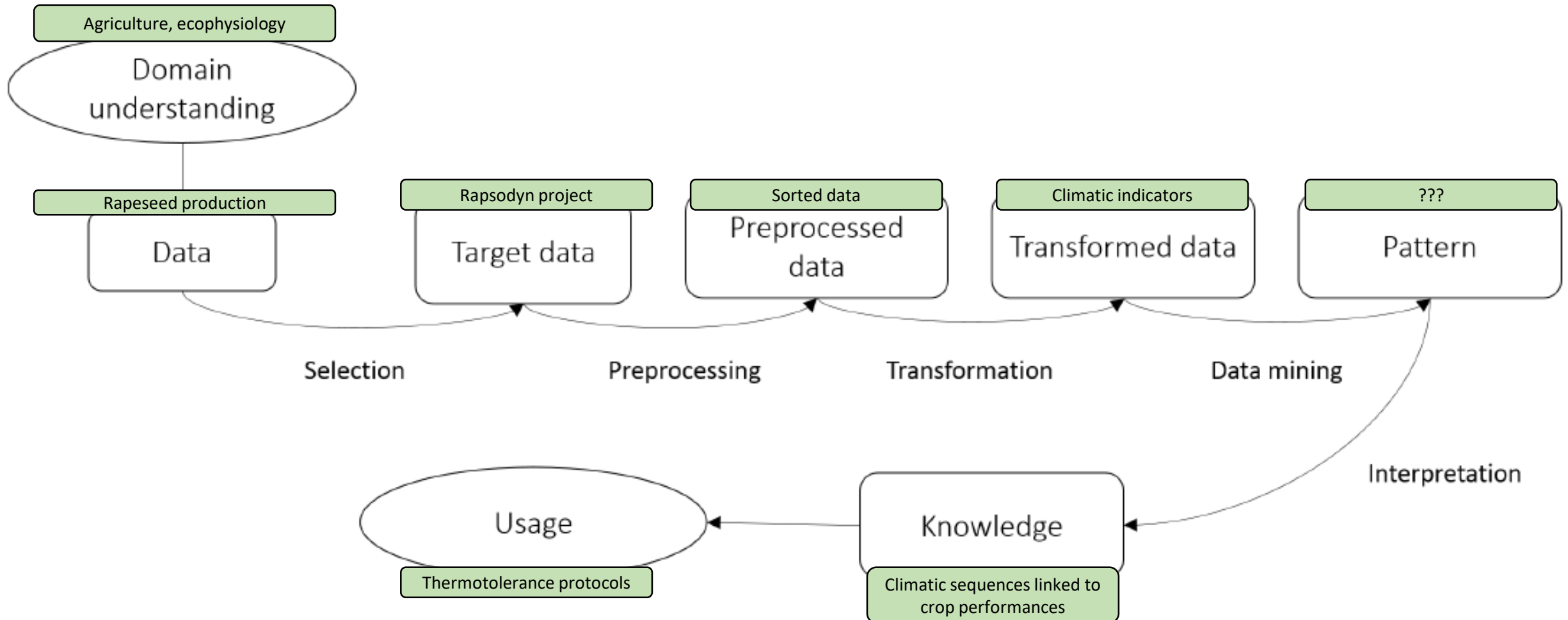


174 rows
(objects)

17 columns (attributes) = 2 qualitatives + 10 climatic + 5 plant performances

indiv	ref	local	genotype	Ptot_25	Ptot_30	P300_25	P300_30	P600_25	P600_30	P1000_25
1	LR11	LR	Aviso	6	2	0	0	1	0	2
2	LR11	LR	Aviso	6	2	0	0	1	0	2
3	LR11	LR	Aviso	6	2	0	0	1	0	2
4	LR11	LR	Montego	6	2	0	0	1	0	2
5	LR11	LR	Montego	6	2	0	0	1	0	2
6	LR11	LR	Montego	6	2	0	0	1	0	2
7	MD11	MD	Aviso	17	1	4	0	0	0	11
8	MD11	MD	Aviso	17	1	4	0	0	0	11
9	MD11	MD	Aviso	17	1	4	0	0	0	11
10	MD11	MD	Aviso	17	1	4	0	0	0	11
11	MD11	MD	Aviso	17	1	4	0	0	0	11
12	MD11	MD	Aviso	17	1	4	0	0	0	11
13	MD11	MD	Montego	17	1	4	0	0	0	9
14	MD11	MD	Montego	17	1	4	0	0	0	9
15	MD11	MD	Montego	17	1	4	0	0	0	9
16	MD11	MD	Montego	17	1	4	0	0	0	9
17	MD11	MD	Montego	17	1	4	0	0	0	9
18	MD11	MD	Montego	17	1	4	0	0	0	9
19	LR12	LR	Aviso	4	0	0	0	1	0	2
20	LR12	LR	Aviso	4	0	0	0	1	0	2
21	LR12	LR	Aviso	4	0	0	0	1	0	2
22	LR12	LR	Aviso	4	0	0	0	1	0	2
23	LR12	LR	Aviso	4	0	0	0	1	0	2
24	LR12	LR	Aviso	4	0	0	0	1	0	2
25	LR12	LR	Aviso	4	0	0	0	1	0	2


Rappel approche



Rappel approche

Illustration de l'approche **bi-clustering** basée sur l'association spécifique entre la **séquence de température** enregistrée et les **caractéristiques mesurées** des performances des plantes

Plant feature T°C sequence	Seed Yield	Seed Oil content	Seed Protein content	Seed Storage capacity	...	Feature n
T°C seq 1	SY1	SO1	SP1	SSc1	...	Fn1
T°C seq 2	SY2	SO2	SP2	SSc2	...	Fn2
T°C seq 3	SY3	SO3	SP3	SSc3	...	Fn3
T°C seq 4	SY4	SO4	SP4	SSc4	...	Fn4
T°C seq 5	SY5	SO5	SP5	SSc5	...	Fn5
...
T°C seq n	SYn	SO _n	SP _n	SSc _n	...	Fn _n

- 
- **Bi-cluster** ({T°C seq 2, T°C seq 3}, {SO}) → effets spécifiques de la séquence thermique sur une caractéristique spécifique des plantes (SO)
 - **Bi-cluster** ({T°C seq 2, T°C seq 5}, {SSc, Feature n}) → effets spécifiques de deux séquences thermiques sur deux caractéristiques des plantes (SSc and Feature n)