

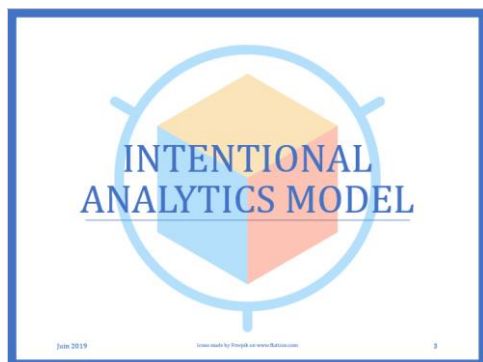
RAPPORT DE STAGE STORYTELLING, OLAP ET CLUSTERING

Antoine Chédin – Université de Tours

antoine.chedin@gmail.com



SOMMAIRE





INTENTIONAL ANALYTICS MODEL

UN NOUVEAU MODÈLE POUR OLAP

Beyond Roll-Up's and Drill-Down's: An Intentional Analytics Model to reinvent OLAP, Panos Vassiliadis, Patrick Marcel, and Stefano Rizzi

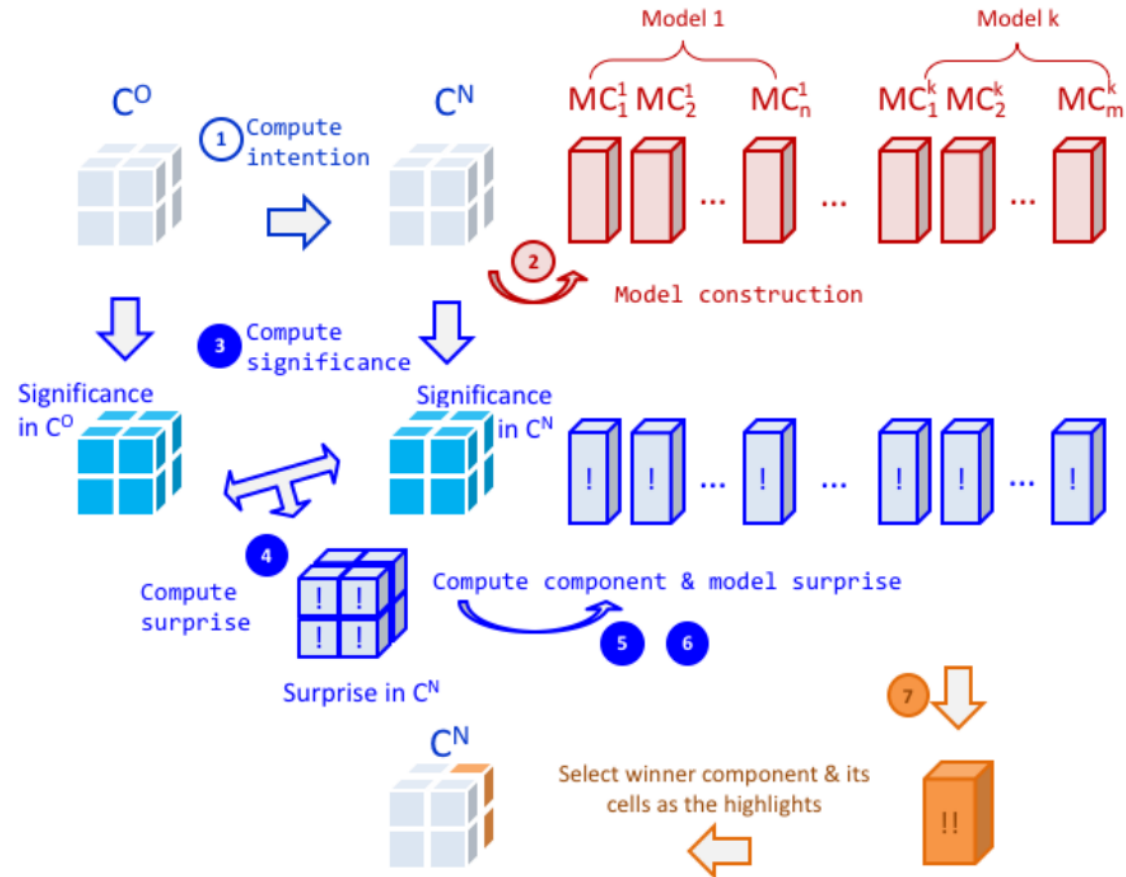
- Drill-Down
- Roll-Up
- Slicing
- Dicing



- Describe
- Assess
- Explain
- Predict
- Suggest



HIGHLIGHT ALGORITHM

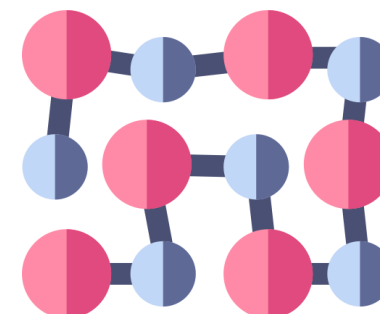


Panos Vassiliadis, Patrick Marcel, and Stefano Rizzi. Beyond roll-up's and drill-down's: An intentional analytics model to reinvent olap.

OBJECTIF: SURLIGNER LES CELLULES SURPRENANTES



PROXIES



Categorie	En métropole	Dans un DOM-TOM	A l'étranger
Personne de référence	74 151,20	7,36	60,22
Conjoint de la personne	46 335,55	4,01	4,17
Enfant de la personne de référence ou de son conjoint	9 787,27		18,25
Autre lien	2 038,53		
Sans Objet	1 043,84		

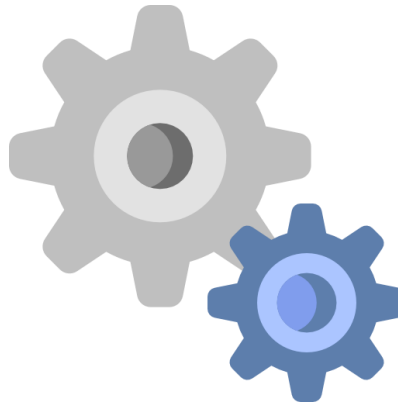
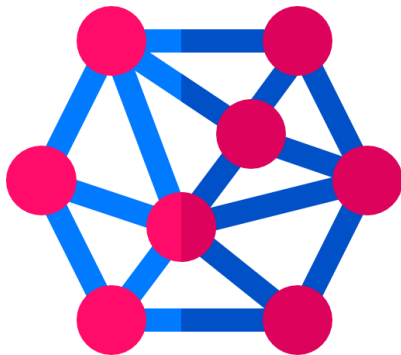


Lien avec la personne de reference	Dans la commune de résidence actuelle	Dans un autre département de la région de résidence	Dans une autre commune du département de résidence	Hors de la région de résidence actuelle : en métropole	Hors de la région de résidence actuelle : dans un DOM	Hors de la région de résidence actuelle : dans une COM	A l'étranger
Personne de référence du ménage	26 768,25	8 505,98	36 036,58	2 840,40	7,36		60,22
Conjoint de la personne de référence du ménage	15 989,73	4 479,92	24 971,44	914,45		4,01	4,17
Enfant de la personne de référence du ménage ou de son conjoint	2 962,26	965,31	5 477,40	382,30			18,25
Ami	191,34	102,03	312,36	7,01			
Ascendant	36,10	7,94	24,96	3,86			
Autre parent	437,20	111,05	570,63	49,72			
Domestique ou salarié logé	11,02						
Pensionnaire ou sous-locataire	43,35	7,92	48,17	5,28			
Petit-enfant	18,96	6,97	42,67				
Hors logement ordinaire	647,74	116,77	227,33	52,00			

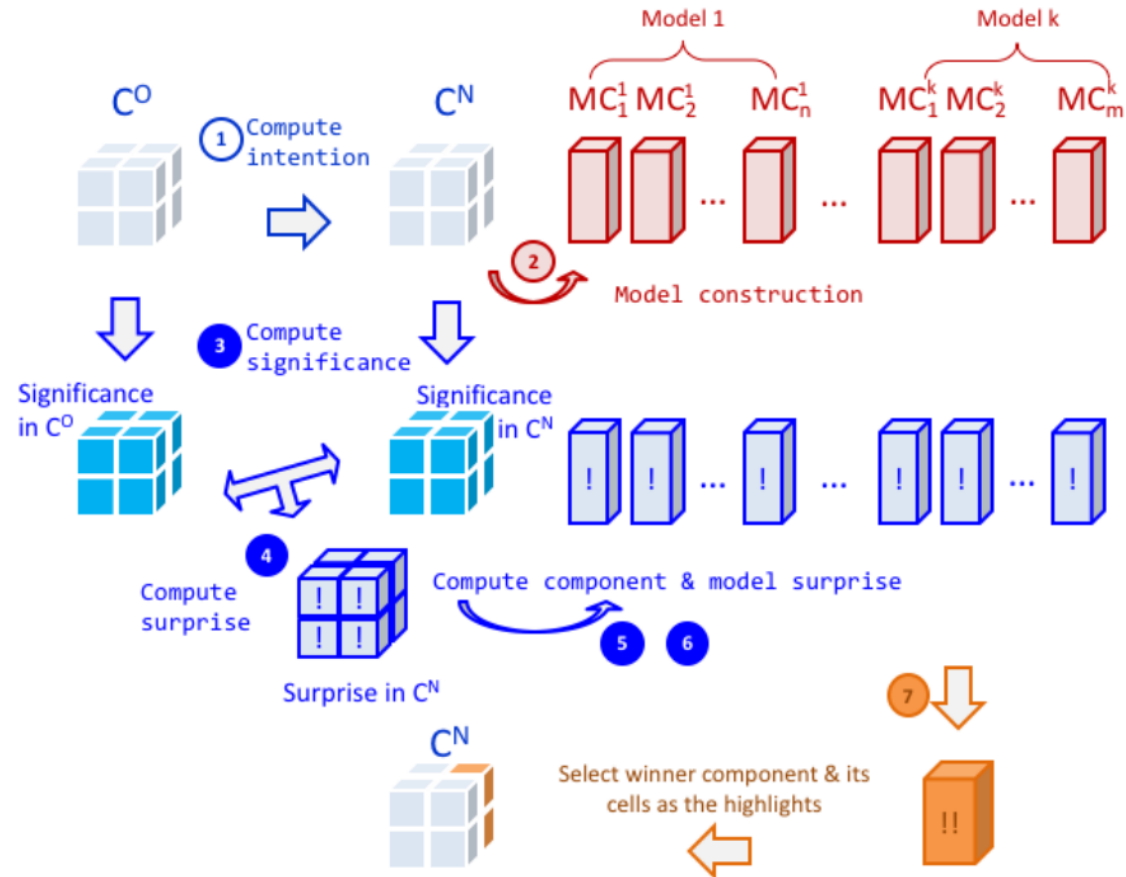
MODÈLES DE SÉLECTION

Name	Input signature	Output signature
Model types for description		
Top-K	(Number of values, Name of measures)	(Rank)
Outlier	(Threshold, Name of measure)	(Outlierness)
Clustering	(Number of clusetes, Name of measure)	(Cluster ₁ , ..., Cluster _n , Representative)
Shrink	(Number of cells, Name of measure)	(Cell ₁ , ..., Cell _n)
Dominating Slice	(Name of measure)	(DomSlice ₁ , ..., DomSlice _n)

Panos Vassiliadis, Patrick Marcel, and Stefano Rizzi. Beyond roll-up's and drill-down's: An intentional analytics model to reinvent olap.



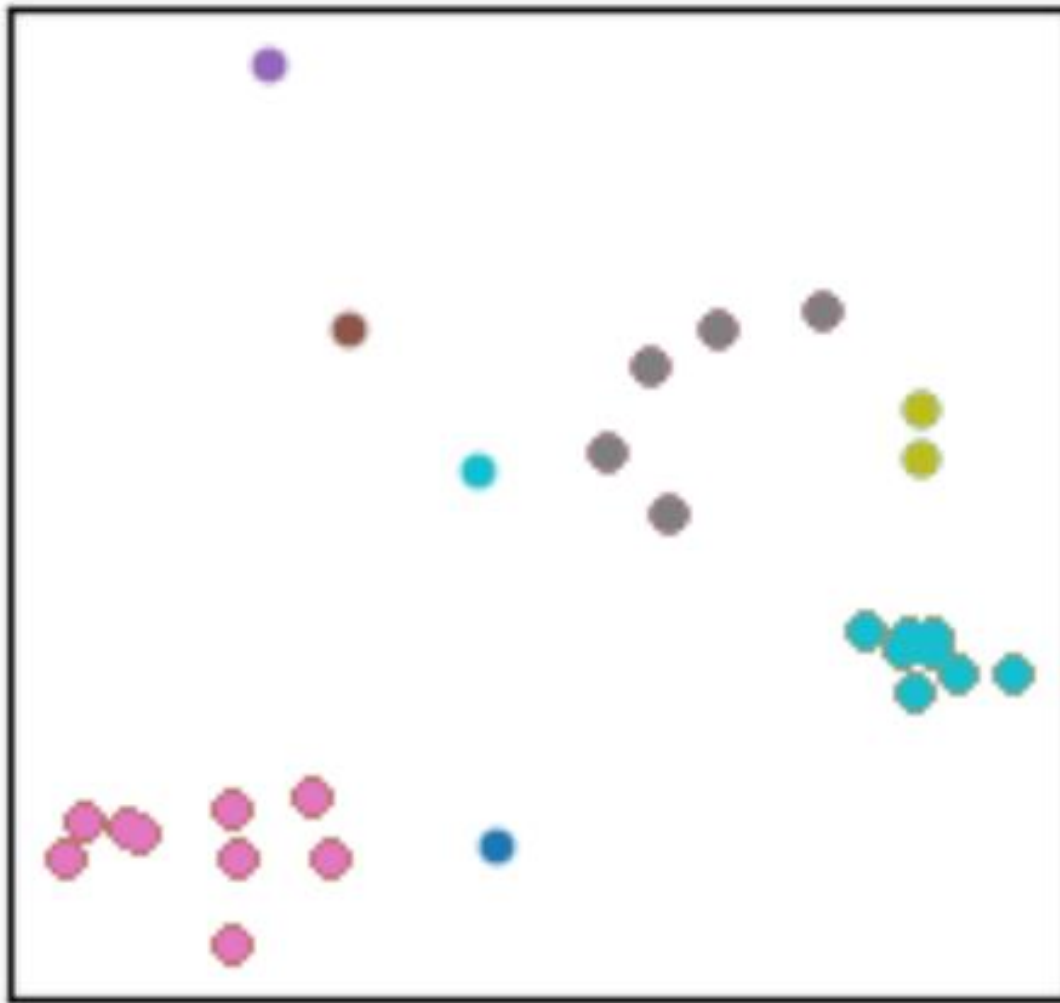
HIGHLIGHT ALGORITHM



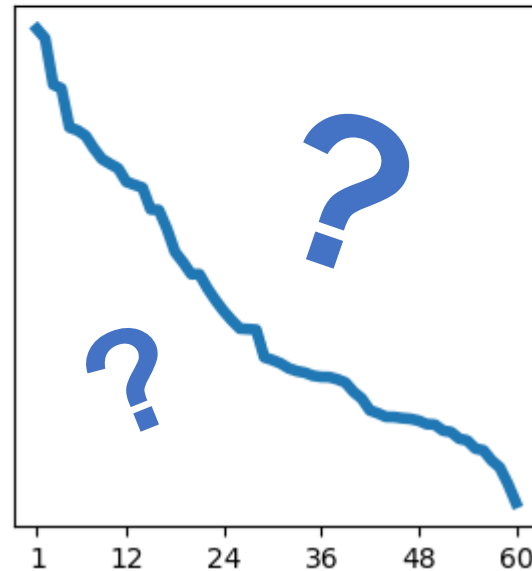
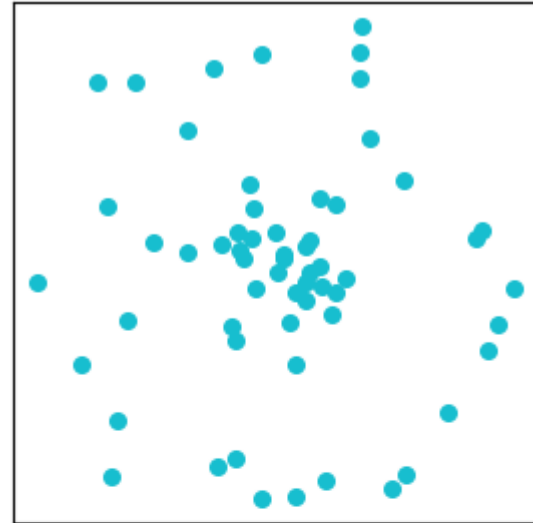
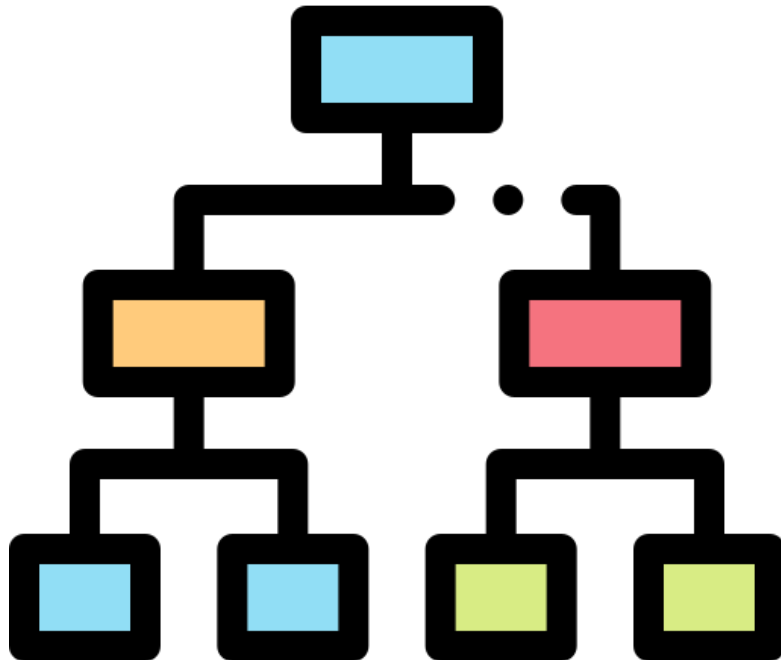
Panos Vassiliadis, Patrick Marcel, and Stefano Rizzi. Beyond roll-up's and drill-down's: An intentional analytics model to reinvent olap.

DESCRIPTIVE CLUSTERING

Le clustering, c'est pas automatique



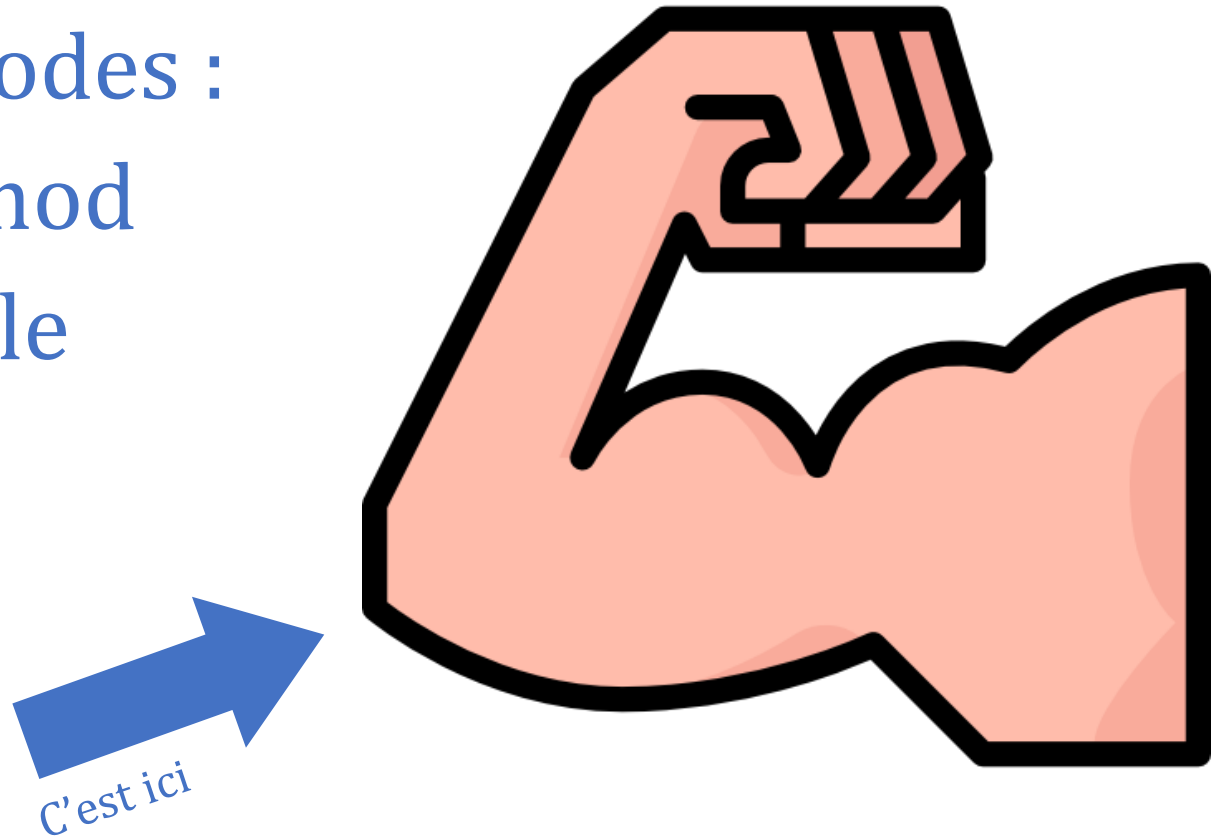
COMMENT ÇA MARCHE ?



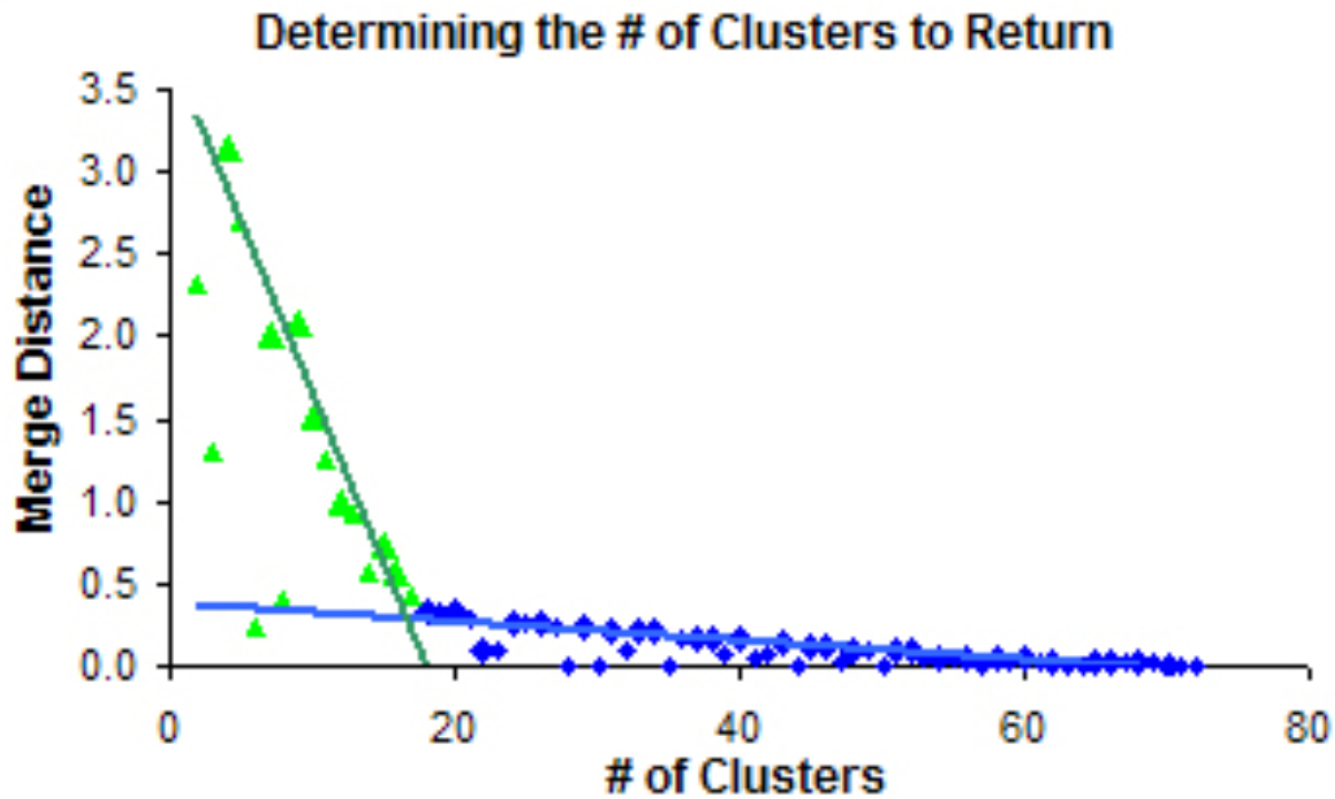
DÉTECTION DU COUDE

2 méthodes :

- L-method
- Kneedle



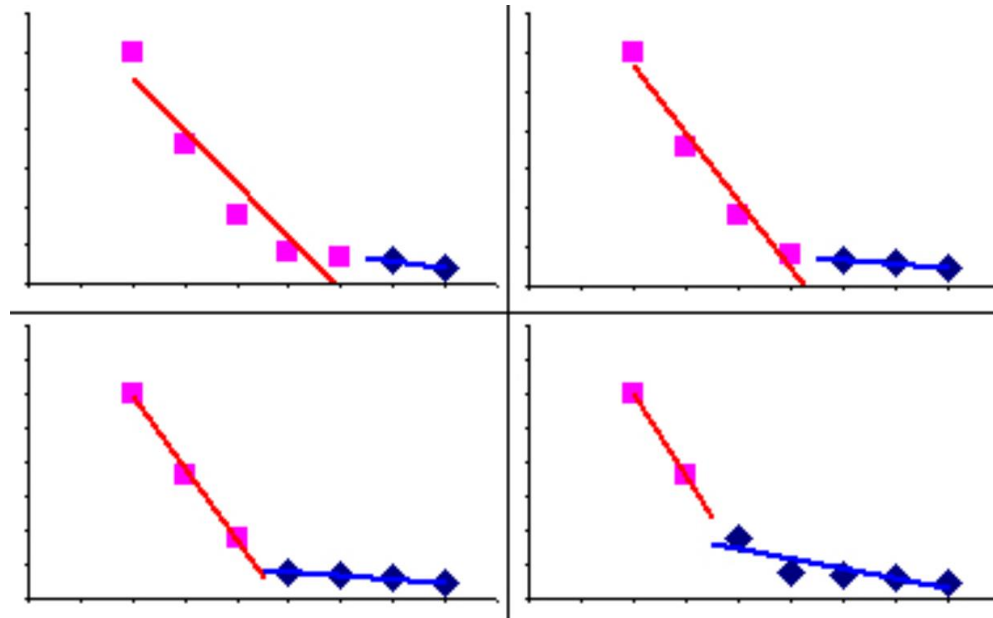
L-METHOD



S. Salvador and P. Chan. Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms

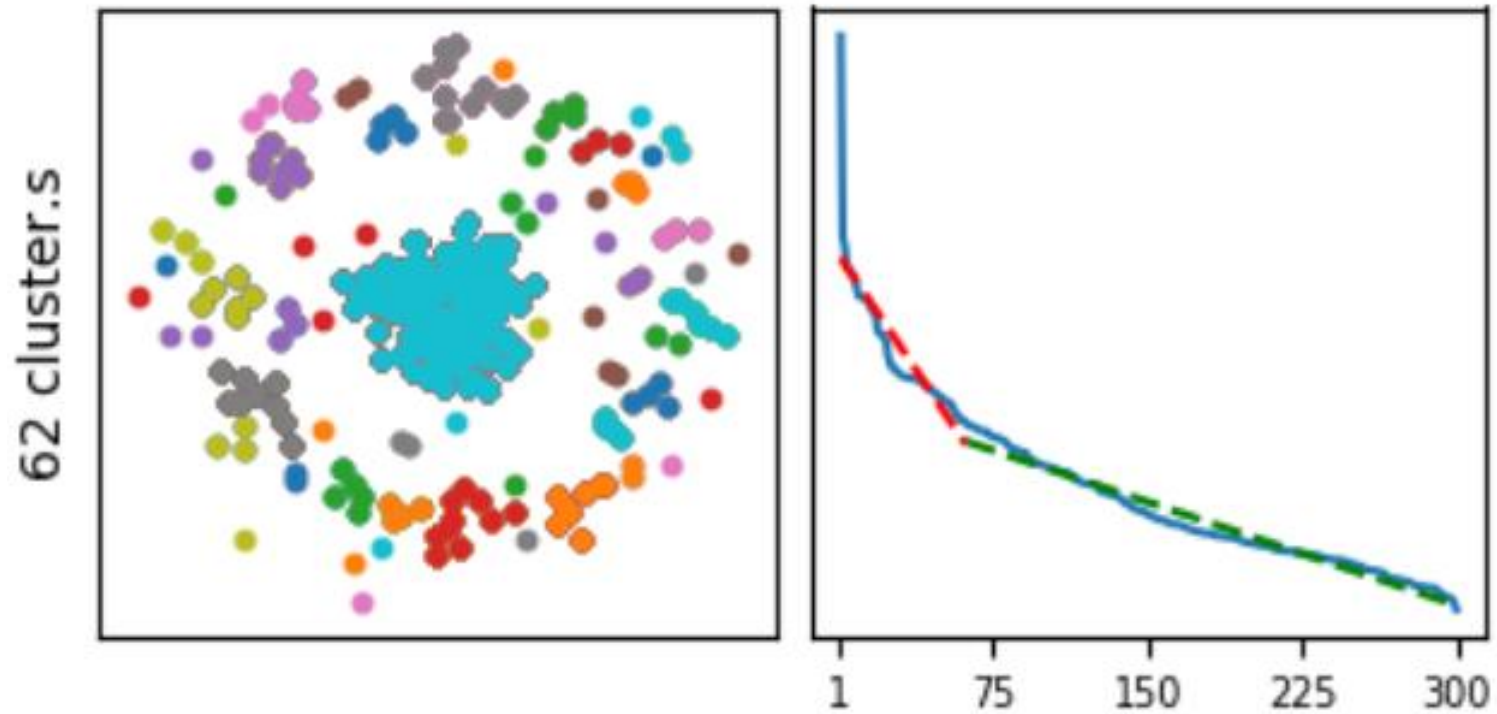
L-METHOD

- Essayer toutes les combinaisons possibles
- Garder celle avec la meilleur erreur quadratique

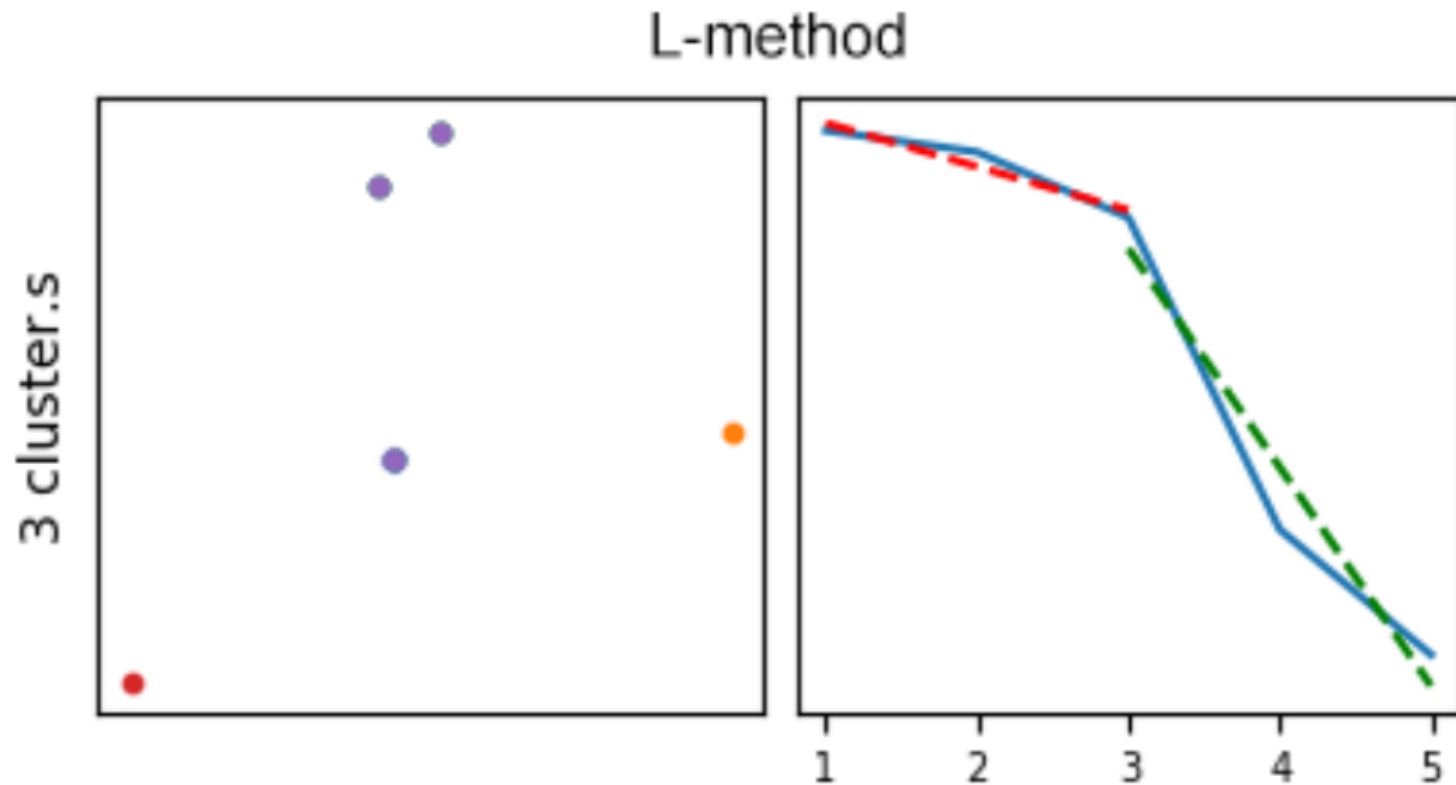


S. Salvador and P. Chan. Determining the number of clusters/segments in hierarchical clustering/segmentation algorithms

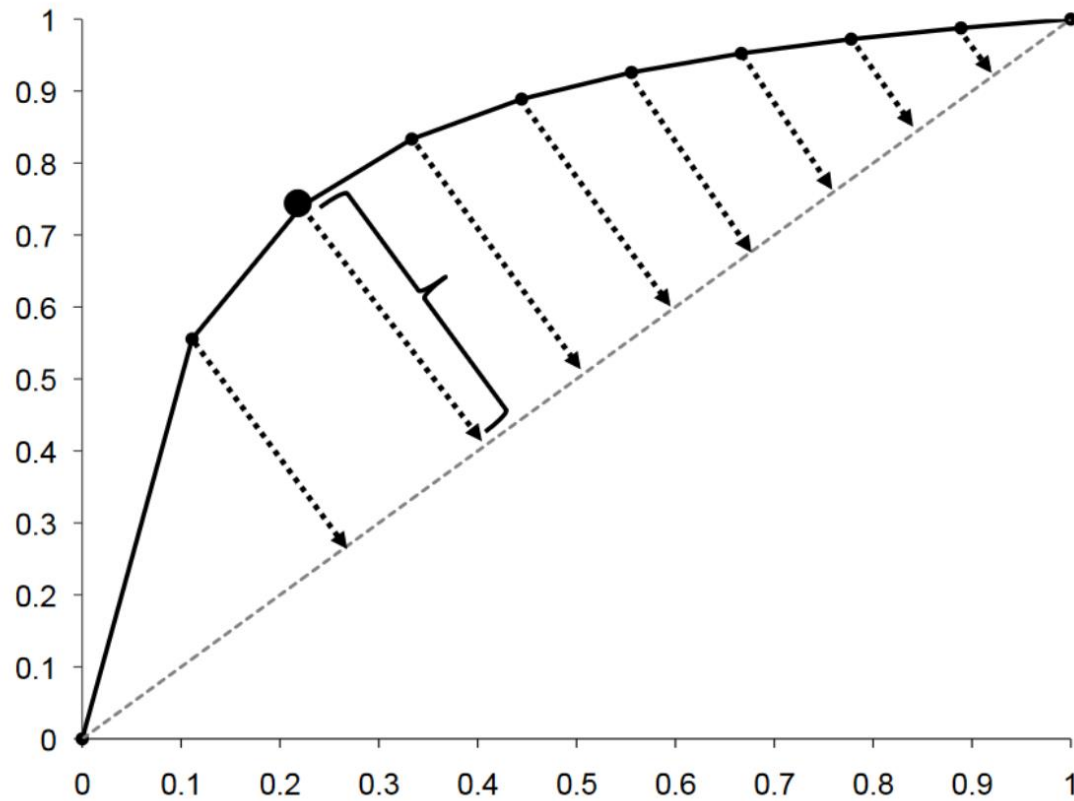
L-METHOD - LIMITATIONS



L-METHOD - LIMITATIONS

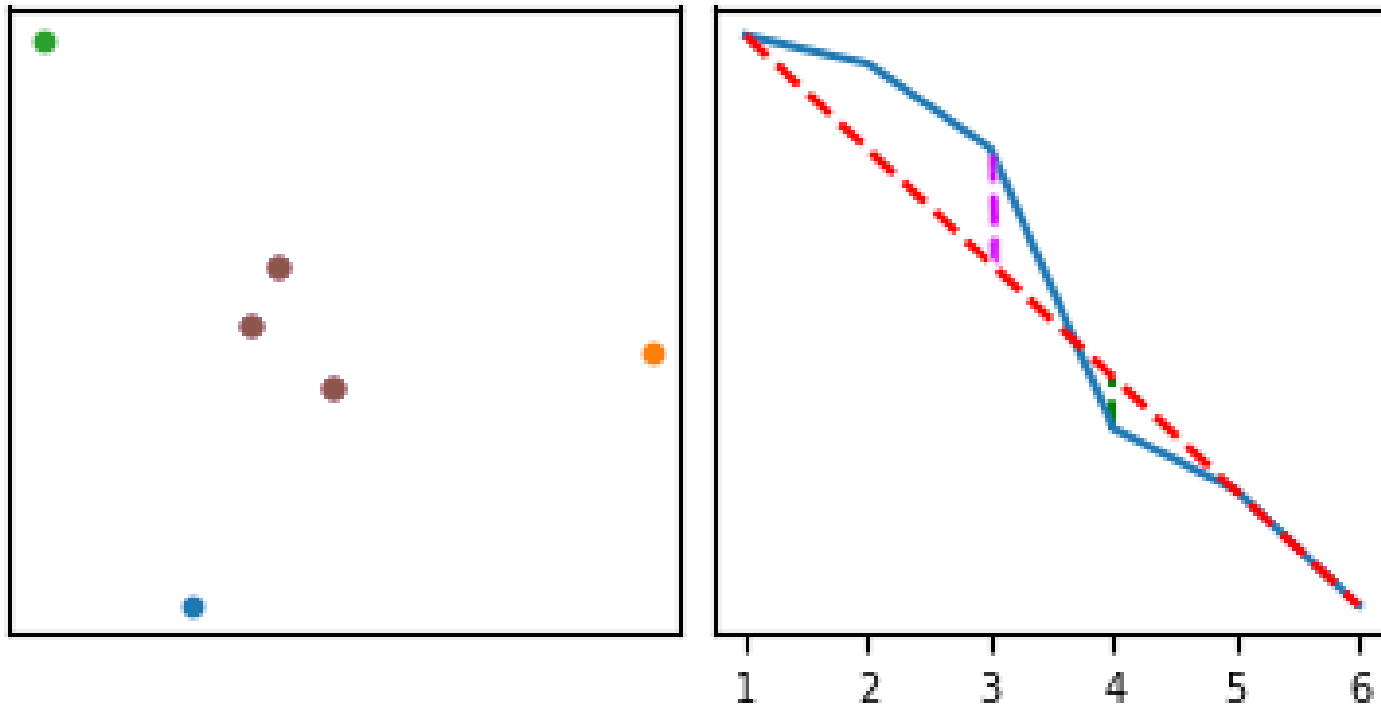


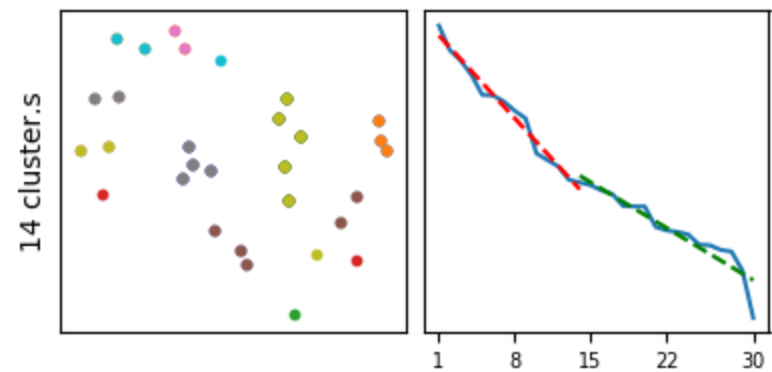
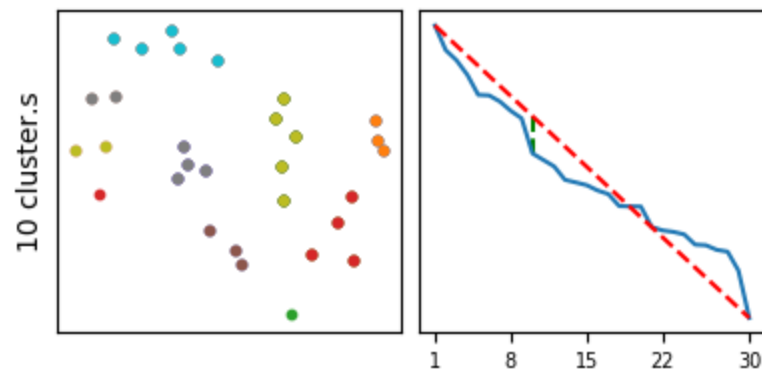
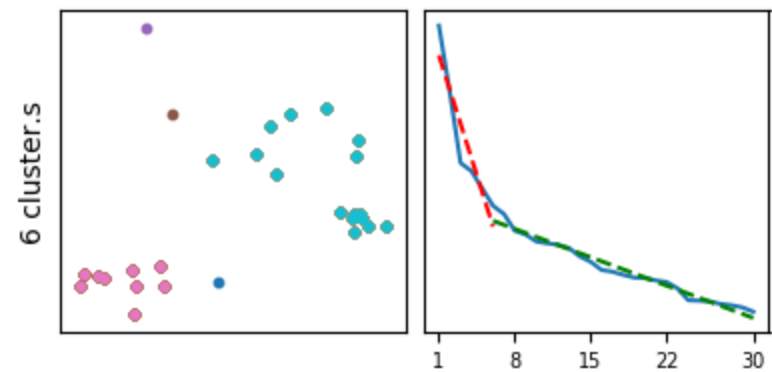
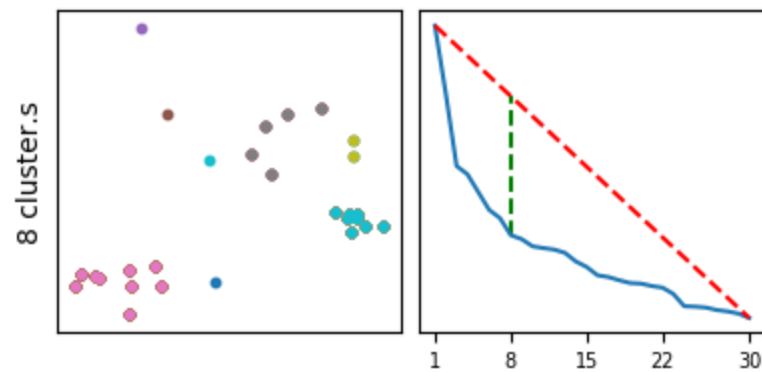
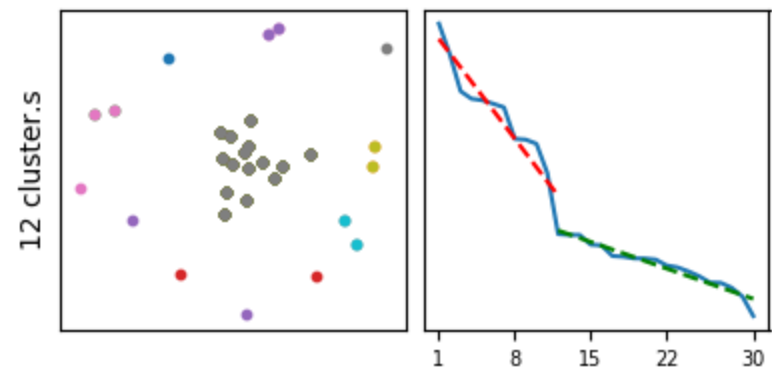
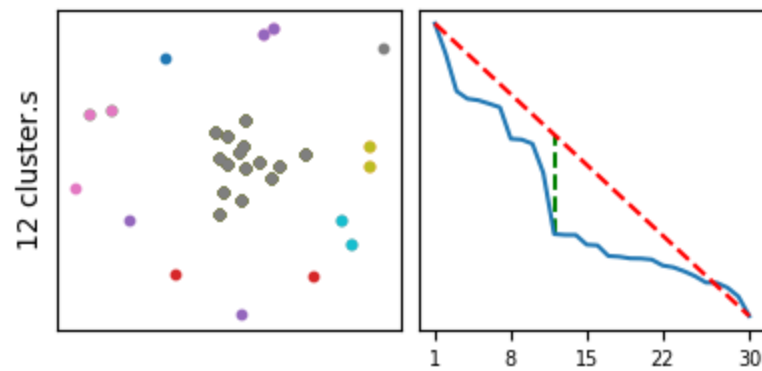
KNEEDLE



V. Satopaa, J. Albrecht, D. Irwin, and B. Raghavan. Finding a "kneedle" in a haystack: Detecting knee points in system behavior

KNEEDLE – LIMITATION FIX







CASTOR

Automatic Storytelling for OLAP

[Local Demo](#)
[Connect](#)

Choose a log file

session_demo.json

[Submit](#)

1

[Data](#)
[Significance](#)
[Surprise](#)
☐ Show clusters

###	Temps complet	Temps partiel
	Nombre total d'individus	Nombre total d'individus
Hommes	65760.66	4039.44
Femmes	46907.81	16742.49

2

[Data](#)
[Significance](#)
[Surprise](#)
☐ Show clusters

###		Temps complet	Temps partiel
Nombre total d'individus	Hommes	0.446	0.336
	Femmes	0.225	0.116



MERCI POUR VOTRE ATTENTION

Antoine Chédin – Université de Tours
antoine.chedin@gmail.com