

Rapport livrable n°3 : Modèle

Problématique

- (a) On se demande si l'établissement et la filière d'origine des étudiant.e.s de DUT Info a une influence sur leur réussite dans les différents modules du S1. On souhaite regrouper les établissements+filière d'origine par groupes selon les profils de réussite des étudiant.e.s dans les différents modules du S1.
- (b) Par exemple, y a-t-il des établissements+filières d'origine qui permettent aux étudiant.e.s de bien réussir en BD et en Maths Discrètes mais moins en Economie, tandis que d'autres permettent de réussir surtout en Systèmes et en Communication ?
- (c) L'intérêt d'une telle observation pourrait être de créer des cours de remise à niveau à la rentrée du S1 en fonction de l'établissement+filière d'origine.

Création d'une Vue appropriée

A partir de la Base de Données des étudiant.e.s de DUT Info, on crée une Vue sous forme d'un tableau .csv.

Dans ce tableau, chaque ligne correspond à un couple (établissement d'origine, filière d'origine). Les colonnes correspondent chacune à un module du 1e semestre du DUT Info.

Dans les cases du tableau, on a les moyennes sur l'ensemble de tou.te.s les étudiant.e.s issu.e.s d'un même (établissement, filière), pour chaque module du S1.

Le choix de cette vue permet d'avoir un aperçu de la réussite des étudiant.e.s issu.e.s d'un même établissement+filière dans les différentes matières du S1.

Voici un aperçu de la Vue créée :

libelle_etablissement	serie	moy_m1101	moy_m1102	moy_m1103	moy_m1104	moy_m1105	moy_m1106	moy_m1201	moy_m1202	moy_m1203	moy_m1204	moy_m1205	moy_m1206	moy_m1207
Atelier de Sévres - A*	o	11.25000000	12.67000000	11.50000000	12.88000000	14.42000000	18.75000000	17.00000000	14.00000000	12.42000000	12.33000000	12.80000000	12.00000000	14.00000000
CFA CMA 35	BTM	15.90000000	16.36000000	19.15000000	16.44000000	16.13000000	18.00000000	16.35000000	8.00000000	12.83000000	13.17000000	16.27000000	12.88000000	16.75000000
ENI de Brest	1ère année	11.08000000	11.21000000	14.75000000	13.53000000	13.21000000	17.00000000	18.00000000	15.20000000	13.00000000	13.00000000	14.13000000	16.50000000	10.25000000
ENI de Brest	Cycle prép	9.09000000	11.41000000	9.00000000	7.88000000	14.17000000	16.00000000	15.25000000	16.40000000	14.00000000	16.42000000	17.37000000	17.13000000	15.25000000
ENI de Brest	formation i	8.80000000	15.41000000	11.70000000	11.03000000	15.58000000	19.00000000	18.00000000	16.40000000	11.00000000	12.00000000	16.37000000	15.13000000	15.00000000
ENI de Brest	Préparation	13.07000000	13.43000000	12.85000000	4.90000000	14.33000000	13.00000000	12.50000000	16.40000000	12.33000000	11.75000000	11.87000000	17.00000000	12.00000000

Utilisation de la méthode des k-moyennes

On choisit de répartir les couples (établissement d'origine, filière d'origine) en 4 clusters :

- (a) On commence par importer dans Python sous forme de DataFrame les données de notre Vue :

Puis on extrait les données numériques de ce DataFrame sous la forme d'un tableau de type `numpy.array`. Ce qui nous permet ensuite de centre-réduire nos données :

Ensuite on exécute la fonction `KMeans` sur nos données. On choisit ici de faire 4 clusters :

On crée ensuite une fonction qui renvoie l'ensemble des couples (établissement d'origine, filière d'origine) compris dans un même cluster :

- (b) Grâce à la commande `res.cluster_centers_` on obtient les valeurs suivantes pour les centres des 4 clusters :

	0	1	2	3	4	5	6	7	8	9	10	11	12
0	-0.206702	-0.417781	-0.356228	-0.348537	-0.101972	0.21819	-0.31216	0.158792	0.865035	0.587422	1.08869	0.505227	0.140611
1	1.02642	1.10951	1.13302	1.01947	0.701052	0.178046	1.12138	0.762751	0.420648	0.511886	0.356771	0.595491	0.443959
2	-1.36165	-1.31181	-1.44919	-1.11059	-1.20619	-0.0281194	-0.765564	-0.396705	-0.0250025	-0.0912512	-0.257687	-0.581082	0.0391068
3	-0.0589981	-0.0394838	-0.0334573	-0.0782995	0.0288396	-0.168667	-0.251274	-0.347736	-0.542145	-0.466807	-0.521159	-0.332073	-0.302436

(c) On voit que le Cluster 0 correspond à des étudiant.e.s ayant des notes un peu plus basses que la moyenne dans les matières informatiques (les colonnes 0 à 4) et en Maths Discrètes (colonne 6) mais au-dessus de la moyenne dans les matières non informatiques et particulièrement en Expression Communication (colonne 10).

Le Cluster 1 correspond à des étudiant.e.s nettement au-dessus de la moyenne dans les matières informatiques et en Maths Discrètes, et un peu au-dessus de la moyenne dans les matières non informatiques.

Le Cluster 2 correspond à des étudiant.e.s en difficulté dans les matières informatiques et les Maths Discrètes, et un peu en-dessous de la moyenne dans les matières non informatiques.

Le Cluster 3 correspond à des étudiant.e.s légèrement en-dessous de la moyenne dans les matières informatiques, et plus nettement en-dessous de la moyenne dans les matières non informatiques.

(d) Le cluster 0 contient les couples (établissement d'origine, filière) suivants :

ENI de Brest	Cycle préparatoire ENIB Ecole Nationale d'Ingénieurs...			
ENI de Brest	Préparation intégrée pour l'ENI			
ENI de Brest	Scientifique cycle préparatoire			
Facultés libres de l'ouest - UCO - Bretagne Nord	Licence - Sciences humaines et sociales			
LGT SAINT-FRANCOIS D'ASSISE	S			
Lycée Alcide D'Orbigny	S			
Lycée Auguste Brizeux	S			
Lycée Chaptal	STI2D			
Lycée Curie-Corot	S			
Lycée De L'Elorn	S			
Lycée de l'Harteloire	S			
Lycée du Leon	S			
Lycée Jean Macé	S			
Lycée La Mennais	S			
Lycée Notre Dame Du Voeu	S			
Lycée Rene Cassin	S			
Lycée Sacre Coeur	STI2D			
Lycée Ste Elisabeth-St Blaise	S			
Lycée St Magloire	S			
Lycée St Pierre	S			
Université de Rennes 1	DU - Intégrateur	d'applications	mobiles	et in...

Le cluster 1 :

Atelier de Sèvres - Atelier Préparatoire aux Ecoles Supérieures d'Art	o
CFA CMA 35	BTM
ENI de Brest	1ere année ecole nationale d'ingénieurs de Brest
ENI de Brest	formation ingénieur généraliste intégrée
INSA Lyon	Ingénieur Mention PREMIER CYCLE 1 ^{re} Année
INSA Rennes	Scientifique Sciences techniques pour l'ingénieur première année
Lycée Amiral Ronarc'h	S
Lycée Bertrand D'Argentre	S
Lycée de Kerneuzec	S
Lycée DE L'IROISE	S
Lycée Emile Littré	S
Lycée Estran Charles De Foucauld	S
Lycée Eugene Freyssinet	S
Lycée Felix Le Dantec	S
Lycée Francois Rabelais	Classe préparatoire scientifique
Lycée Frédéric OZANAM	S
Lycée Henri Avril	S
Lycée Jacques Cartier	S
Lycée Jean-Paul II	S
Lycée La Pérouse - Kerichen	Classe préparatoire scientifique
Lycée Le Likès	S
Lycée Leonard de Vinci	S
Lycée Maupertuis	S
Lycée Notre Dame De Campostal	S
Lycée Saint-Charles	S
Lycée Saint Joseph - Bossuet	S
Lycée St Esprit	S
Pôle Chateaubriand - Joliot-Curie	Classe préparatoire scientifique
Université de Rennes 2	Licence - Arts-lettres-langues
Université Paris 2 Panthéon - Assas	Licence - Droit-économie-gestion

Le Cluster 2 :

ENI de Brest	Scientifique ingénieur
I.U.T d'Evreux	DUT - Production
Lycée ASSOMPTION	S
Lycée Brest - Rive Droite - Javouhey	S
Lycée De La Fontaine Des Eaux	STI2D
Lycée De L'Elorn	E5
Lycée François René de Chateaubriand	S
Lycée Jean-Baptiste Colbert	STI2D
Lycée Jeanne D'Arc	STI2D
Lycée Kerraoul	S
Lycée Kerraoul	ST2S
Lycée La Pérouse - Kerichen	S
Lycée Marcel Gambier	S
Lycée Notre Dame De Kerbertrand	S
Lycée St Louis	S
Lycée St-Sébastien	S
Lycée Victor et Hélène BASCH	S

Et le Cluster 3 :

INSA Rennes	Scientifique prépa STPI (Sciences et Techniques Pour l'Ingénieur
Institution St Malo-Providence	S
I.U.T de Lannion	DUT - Production
LPO SAINT AUBIN LA SALLE	S
Lycée Alain René Lesage	S
Lycée Anita Conti	S
Lycée Auguste Pavie	S
Lycée Benjamin Franklin	S
Lycée Bréquigny	S
Lycée Bréquigny	STI2D
Lycée Chaptal	S
Lycée De La Fontaine Des Eaux	S
Lycée Emile Zola	STI2D
Lycée Ernest Renan	S
Lycée Felix Le Dantec	STI2D
Lycée Francois Rabelais	S
Lycée Frédéric OZANAM	STI2D
Lycée Fulgence Bienvenue	S
Lycée Giacante De Casabianca	S
Lycée Henri Bergson	S
Lycée International Victor Hugo	S
Lycée Jean-Baptiste Colbert	S
Lycée Jean De Lattre de Tassigny	S
Lycée Jean De Lattre de Tassigny	STI2D
Lycée Jean Guehenno	S
Lycée Jean Mace	S
Lycée Jean Moulin	S
Lycée Jeanne D'Arc	S
Lycée Jean Perrin	S
Lycée Jean XXIII	S
Lycée Joliot-Curie	S
Lycée Joseph Loth	S
Lycée Jules Ferry	STI2D
Lycée Julliot De La Morandiere	S
Lycée La Croix Rouge	S
Lycée La Mennais	STI2D
Lycée Le Likès	STI2D
Lycée Marcel Callo	STI2D
Lycée Marcelin Berthelot	S
Lycée Marcellin Berthelot	S
Lycée Maupertuis	STI2D
Lycée N-D DU KREISKER	S
Lycée Pierre Mendes France	S
Lycée professionnel Felix Le Dantec	P
Lycée Rene Laennec	S
Lycée Ste Therese	S
Lycée St Joseph	S
Lycée St Joseph - LaSalle Lorient	S
Lycée St-Sébastien	STI2D
Lycée Tristan Corbiere	S
Lycée Vauban	S
Lycée Vauban	STI2D
Lycée Yves Thepot	S
Lycée Yves Thepot	STI2D
Université de Rennes 1	Licence - Sciences - technologies - santé
nan	nan

Propositions d'actions

D'après les résultats obtenus avec la méthode des k-moyennes, il pourrait être intéressant de

- proposer des mises à niveaux ou/et du soutien en informatique et en maths aux étudiant.e.s issu.e.s des établissements et filières présents dans le Cluster 0 et le Cluster 2
- proposer du soutien dans les matières non informatiques aux étudiants.e.s issu.e.s des établissements et filières présents dans le Cluster 3
- on pourrait aussi envisager que les étudiant.e.s issu.e.s du Cluster 0 puissent aider les étudiant.e.s issu.e.s du Cluster 3 en informatique et maths et réciproquement dans les autres matières.