



**Université MOHAMED V de Rabat
École Nationale Supérieure d'Informatique
et d'Analyse des Systèmes**



PROJET DATA MINING

Analyse des relevés des notes des étudiants de l'ENSIAS

Réalisé par :

ELBRIKI Amal
HAMAOUI Amine
SAIDI Yasmine

Encadré par :

Mme. Houda BENBRAHIM

Remerciement

Nous profitons par le biais de ce rapport pour exprimer nos vifs remerciements à toute personne contribuant de près ou de loin à l'élaboration de ce travail.

Nous tenons à remercier vivement tous nos professeurs. Le Directeur de notre établissement Mr Mohammed HALIM, qui ont contribué à la réalisation de ce modeste projet, qui nous ont encadrés et aidés tout au long de notre parcours. Ainsi nous remercions l'administration de l'École Nationale Supérieure d'Informatique et d'Analyse des Systèmes qui à travers leur programme nous ont fourni des outils de qualité facilitant notre spécialisation.

Un merci bien particulier adressé également à Mme. Houda BENBRAHIM notre professeur de Data mining, pour ces remarques, ses directives, et l'intérêt qu'elle porte à ses étudiants. Nous tenons à lui exprimer nos sincères remerciements pour son suivi et ses orientations.

Table des matières

1	La compréhension métier	7
1.1	Détermination des objectifs stratégiques et opérationnelles	7
1.2	Analyse de la situation actuelle	7
1.3	Détermination des objectifs du Data Mining	7
1.3.1	Clustering	7
1.3.2	K-means	8
2	La compréhension des données	9
2.1	Collecte des données	9
2.2	Description des données	9
2.3	Exploration des données	11
2.3.1	Exploradion à l'aide des tableaux	11
2.3.2	Visualisation du pourcentage de réussite	12
2.3.3	Visualisation du pourcentage des filières	12
2.4	Vérification de la qualité des données	15
3	Préparation des données	17
3.1	Sélection des données	17
3.2	Nettoyage des données	17
3.2.1	Données manquantes	17
3.2.2	Erreurs dans les données	18
3.3	Intégration des données	20
4	Modélisation	22
4.1	Sélection des techniques de modélisation	22
4.2	Génération d'une conception de test	22
4.3	Construire le modèle	22
5	Evaluation et déploiement	28
5.1	Evaluation	28
5.1.1	Qualité du modèle	28
5.1.2	Amélioration	28
5.2	Déploiement	28

Table des figures

1	Principe algorithmique de K-means	8
2	Les données collectées	9
3	Les composants de chaque fichier	9
4	Création des nœuds sources	11
5	Explorer les données à l'aide des tableaux	11
6	Pourcentage de réussite	12
7	Extraire le champ des noms	13
8	Ajouté la colonne de la filière	13
9	Union des nœuds	14
10	Pourcentage des filières	15
11	Qualité des données	16
12	Qualité des données de nœud notes SSI	16
13	Sélection des données	17
14	Insertion d'une valeur manquante	18
15	vérification des données	19
16	Détection d'une valeur erronée	19
17	Remplacer la valeur erronée par la valeur moyenne	20
18	Fusionner les données de la première année et de la deuxième année de la filière Iel . .	21
19	Définir les inputs du modèle	23
20	Génération du modèle	24
21	Temps d'exécution	25
22	Taille des clusters	25
23	Importance des prédicteurs	26
24	Résultat	27
25	Qualité des clusters	28

Liste des tableaux

1	Dictionnaire des données	10
---	------------------------------------	----

Introduction

Dans notre école l'ENSIAS, on travaille plusieurs projets dans des groupes de deux à six personnes de façon collaborative et responsable. Le travail de groupe développe bien sûr les compétences sociales, mais il poursuit aussi l'objectif d'intensifier l'apprentissage disciplinaire. Pour cela, nous devons bien choisir notre groupe. Les enseignants laissent aux élèves la liberté de former leurs groupes. En général, les groupes sont formés en basant sur l'amitié qui relie les élèves et non pas leurs compétences.

Le but de ce projet est d'aider les élèves à former des groupes en exploitant intelligemment leurs relevés des notes de la première et la deuxième année et extraire des groupes de personnes ayant plus ou moins les mêmes notes dans les différentes matières. Donc nous obtenons des groupes homogènes de personnes qui peuvent exploiter leurs points fort communs pour réaliser un bon projet.

Dans ce projet on va suivre la méthode CRISP-DM qui signifie « Cross-industry standard process for data mining ». Chaque chapitre va être une étape de cette méthode. Donc nous aurons cinq chapitres dans ce rapport qui sont : la compréhension métier, la compréhension des données, la préparation des données, la modélisation et enfin l'évaluation et le déploiement.

1 La compréhension métier

Cette première phase consiste à bien comprendre les éléments métiers et la problématique qu'on vise à résoudre. On va commencer par déterminer les objectifs stratégiques et opérationnelles. Puis, on va évaluer la situation actuelle. Finalement, vient la traduction de l'objectif stratégique en concepts de Data mining.

1.1 Détermination des objectifs stratégiques et opérationnelles

Le problème qu'on rencontre dans le travail de groupe dans notre école l'ENSIAS c'est la non homogénéité des groupes. Le seul facteur de la formation de ces groupes c'est la relation entre les membres. Ainsi les outils techniques peuvent ne pas plaire tous les membres d'un groupe. D'où vient l'idée de notre projet, former des groupes homogènes des étudiants ayant les mêmes compétences pour qu'ils puissent travailler sur un projet qui exploite leurs compétences communes. Tous cela, en analysant leurs résultats et chercher les élèves ayant des résultats similaires dans la même filière.

1.2 Analyse de la situation actuelle

Actuellement, nous disposons de quelques relevés de notes de la deuxième année de quelques filières : la filière e-Management et Business Intelligence (eMBI), la filière Génie Logiciel (GL) et les notes de la première année. Tous ces relevés de notes sont de la promotion de l'année 2020. Nous devons encore collecter les relevés de notes d'autres filières : la filière Ingénierie e-Logistique (IeL), la filière Ingénierie des Systèmes Embarqués et Mobiles (ISEM) et la filière Sécurité des Systèmes d'Information (SSI).

Parmi les problèmes que nous avons rencontrés c'est la sensibilité des données. On ne peut pas diffuser telles données personnelles, donc on doit cacher les noms et les prénoms des étudiants pour protéger leurs confidentialités.

1.3 Détermination des objectifs du Data Mining

Puisque l'objectif stratégique est clairement défini, il convient maintenant de le traduire en concepts de Data Mining. On va opter à une technique descriptive car nous sommes en train de présenter une information cachée par le volume des données. Et plus spécifiquement on va travailler un problème de clustering qui sert à regrouper des données non étiquetées présentant des propriétés similaires. Dans notre cas les propriétés sont les notes de chaque module.

1.3.1 Clustering

Le Clustering ou le partitionnement des données est une méthode de classification non supervisée rassemble un ensemble d'algorithmes d'apprentissage pour relever des sous-ensembles des données difficiles à identifier à l'œil nu.

Pour atteindre ce but, cette technique vise à maximiser l'inertie entre les sous-ensembles pour bien les différencier et minimiser celle au sein du même ensemble pour qu'il soit plus homogène. Il existe plusieurs méthodes de clustering : méthodes hiérarchiques, méthodes de partitionnement et méthode

mixtes.

1.3.2 K-means

Nous avons choisis travailler avec l'algorithme de K-means c'est un algorithme de partitionnement qui permet de regrouper les observations de notre data set en K clusters distincts. L'algorithme de K-means est comme suit :

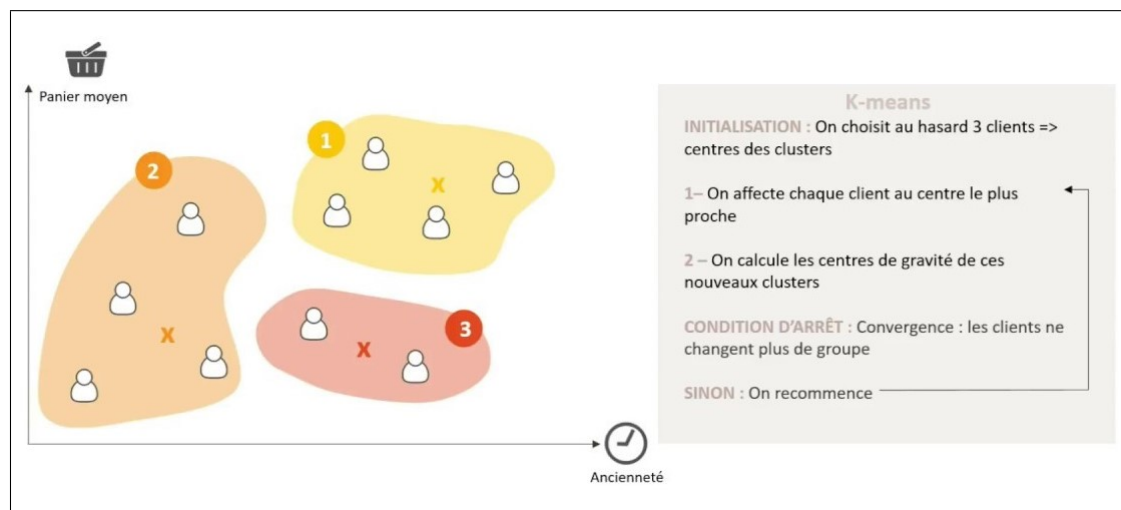


FIGURE 1 – Principe algorithmique de K-means

2 La compréhension des données

La phase de compréhension des données de CRISP-DM implique l'étude des données disponibles pour le Data mining. On doit tout d'abord collecter les données puis faire la description de ces données. La troisième étape est l'exploration des données et comme étape finale vérifier la qualité des données collectées.

2.1 Collecte des données

Nous avons pu collecter les notes de toute la promotion 2020 dans toutes les filière sauf la filière (IWIM) auprès des élèves de cette promotion. Les notes de la première année qui est une année commune à toutes les filières. Puis les notes des élèves de chaque filière : eMBI, GL, IeL, SSI, et ISEM en deuxième année.







Nom	Modifié le	Type	Taille
 Plate forme 2A_eMBI 2018-2019	09/04/2020 22:28	Fichier PDF	243 Ko
 Plate forme 2A_IeL 2018-2019	13/04/2020 13:44	Fichier PNG	114 Ko
 Plate forme 2A_isem 2018-2019	09/04/2020 23:20	Microsoft Excel W...	12 Ko
 Plate forme 2A_SSI 2018-2019	09/04/2020 22:28	Fichier PDF	183 Ko
 Plate forme 1A 2017-2018	09/04/2020 22:29	Fichier PDF	1 375 Ko
 Plate forme 2A_GL 2018-2019	04/04/2020 23:54	Fichier PDF	71 Ko

FIGURE 2 – Les données collectées

Nos fichiers sont de différents formats : des fichiers pdf, un fichier png et un fichier Microsoft Excel.

2.2 Description des données

Chaque fichier contient le nom et le prénom de l'étudiant, le moyen dans chaque semestre et chaque module ainsi la décision finale.

NOM	PRENOM	Moy_1 A	Moy_2 A	Moy_3 1	Moy_4 52	Moy_5 3	Moy_6 3	Moy_7 3	Moy_8 3	Moy_9 3	Moy_10 3	Moy_11 3	Moy_12 3	Moy_13 3	Moy_14 3	Moy_15 3	Moy_16 3	Moy_17 3	Moy_18 3	Moy_19 3	Moy_20 3	Moy_21 3	Moy_22 3	Moy_23 3	Moy_24 3	Moy_25 3	Moy_26 3	Moy_27 3	Moy_28 3	Moy_29 3	Moy_30 3	Moy_31 3	Moy_32 3	Moy_33 3	Moy_34 3	Moy_35 3	Moy_36 3	Moy_37 3	Moy_38 3	Moy_39 3	Moy_40 3	Moy_41 3	Moy_42 3	Moy_43 3	Moy_44 3	Moy_45 3	Moy_46 3	Moy_47 3	Moy_48 3	Moy_49 3	Moy_50 3	Moy_51 3	Moy_52 3	Moy_53 3	Moy_54 3	Moy_55 3	Moy_56 3	Moy_57 3	Moy_58 3	Moy_59 3	Moy_60 3	Moy_61 3	Moy_62 3	Moy_63 3	Moy_64 3	Moy_65 3	Moy_66 3	Moy_67 3	Moy_68 3	Moy_69 3	Moy_70 3	Moy_71 3	Moy_72 3	Moy_73 3	Moy_74 3	Moy_75 3	Moy_76 3	Moy_77 3	Moy_78 3	Moy_79 3	Moy_80 3	Moy_81 3	Moy_82 3	Moy_83 3	Moy_84 3	Moy_85 3	Moy_86 3	Moy_87 3	Moy_88 3	Moy_89 3	Moy_90 3	Moy_91 3	Moy_92 3	Moy_93 3	Moy_94 3	Moy_95 3	Moy_96 3	Moy_97 3	Moy_98 3	Moy_99 3	Moy_100 3	Moy_101 3	Moy_102 3	Moy_103 3	Moy_104 3	Moy_105 3	Moy_106 3	Moy_107 3	Moy_108 3	Moy_109 3	Moy_110 3	Moy_111 3	Moy_112 3	Moy_113 3	Moy_114 3	Moy_115 3	Moy_116 3	Moy_117 3	Moy_118 3	Moy_119 3	Moy_120 3	Moy_121 3	Moy_122 3	Moy_123 3	Moy_124 3	Moy_125 3	Moy_126 3	Moy_127 3	Moy_128 3	Moy_129 3	Moy_130 3	Moy_131 3	Moy_132 3	Moy_133 3	Moy_134 3	Moy_135 3	Moy_136 3	Moy_137 3	Moy_138 3	Moy_139 3	Moy_140 3	Moy_141 3	Moy_142 3	Moy_143 3	Moy_144 3	Moy_145 3	Moy_146 3	Moy_147 3	Moy_148 3	Moy_149 3	Moy_150 3	Moy_151 3	Moy_152 3	Moy_153 3	Moy_154 3	Moy_155 3	Moy_156 3	Moy_157 3	Moy_158 3	Moy_159 3	Moy_160 3	Moy_161 3	Moy_162 3	Moy_163 3	Moy_164 3	Moy_165 3	Moy_166 3	Moy_167 3	Moy_168 3	Moy_169 3	Moy_170 3	Moy_171 3	Moy_172 3	Moy_173 3	Moy_174 3	Moy_175 3	Moy_176 3	Moy_177 3	Moy_178 3	Moy_179 3	Moy_180 3	Moy_181 3	Moy_182 3	Moy_183 3	Moy_184 3	Moy_185 3	Moy_186 3	Moy_187 3	Moy_188 3	Moy_189 3	Moy_190 3	Moy_191 3	Moy_192 3	Moy_193 3	Moy_194 3	Moy_195 3	Moy_196 3	Moy_197 3	Moy_198 3	Moy_199 3	Moy_200 3	Moy_201 3	Moy_202 3	Moy_203 3	Moy_204 3	Moy_205 3	Moy_206 3	Moy_207 3	Moy_208 3	Moy_209 3	Moy_210 3	Moy_211 3	Moy_212 3	Moy_213 3	Moy_214 3	Moy_215 3	Moy_216 3	Moy_217 3	Moy_218 3	Moy_219 3	Moy_220 3	Moy_221 3	Moy_222 3	Moy_223 3	Moy_224 3	Moy_225 3	Moy_226 3	Moy_227 3	Moy_228 3	Moy_229 3	Moy_230 3	Moy_231 3	Moy_232 3	Moy_233 3	Moy_234 3	Moy_235 3	Moy_236 3	Moy_237 3	Moy_238 3	Moy_239 3	Moy_240 3	Moy_241 3	Moy_242 3	Moy_243 3	Moy_244 3	Moy_245 3	Moy_246 3	Moy_247 3	Moy_248 3	Moy_249 3	Moy_250 3	Moy_251 3	Moy_252 3	Moy_253 3	Moy_254 3	Moy_255 3	Moy_256 3	Moy_257 3	Moy_258 3	Moy_259 3	Moy_260 3	Moy_261 3	Moy_262 3	Moy_263 3	Moy_264 3	Moy_265 3	Moy_266 3	Moy_267 3	Moy_268 3	Moy_269 3	Moy_270 3	Moy_271 3	Moy_272 3	Moy_273 3	Moy_274 3	Moy_275 3	Moy_276 3	Moy_277 3	Moy_278 3	Moy_279 3	Moy_280 3	Moy_281 3	Moy_282 3	Moy_283 3	Moy_284 3	Moy_285 3	Moy_286 3	Moy_287 3	Moy_288 3	Moy_289 3	Moy_290 3	Moy_291 3	Moy_292 3	Moy_293 3	Moy_294 3	Moy_295 3	Moy_296 3	Moy_297 3	Moy_298 3	Moy_299 3	Moy_300 3	Moy_301 3	Moy_302 3	Moy_303 3	Moy_304 3	Moy_305 3	Moy_306 3	Moy_307 3	Moy_308 3	Moy_309 3	Moy_310 3	Moy_311 3	Moy_312 3	Moy_313 3	Moy_314 3	Moy_315 3	Moy_316 3	Moy_317 3	Moy_318 3	Moy_319 3	Moy_320 3	Moy_321 3	Moy_322 3	Moy_323 3	Moy_324 3	Moy_325 3	Moy_326 3	Moy_327 3	Moy_328 3	Moy_329 3	Moy_330 3	Moy_331 3	Moy_332 3	Moy_333 3	Moy_334 3	Moy_335 3	Moy_336 3	Moy_337 3	Moy_338 3	Moy_339 3	Moy_340 3	Moy_341 3	Moy_342 3	Moy_343 3	Moy_344 3	Moy_345 3	Moy_346 3	Moy_347 3	Moy_348 3	Moy_349 3	Moy_350 3	Moy_351 3	Moy_352 3	Moy_353 3	Moy_354 3	Moy_355 3	Moy_356 3	Moy_357 3	Moy_358 3	Moy_359 3	Moy_360 3	Moy_361 3	Moy_362 3	Moy_363 3	Moy_364 3	Moy_365 3	Moy_366 3	Moy_367 3	Moy_368 3	Moy_369 3	Moy_370 3	Moy_371 3	Moy_372 3	Moy_373 3	Moy_374 3	Moy_375 3	Moy_376 3	Moy_377 3	Moy_378 3	Moy_379 3	Moy_380 3	Moy_381 3	Moy_382 3	Moy_383 3	Moy_384 3	Moy_385 3	Moy_386 3	Moy_387 3	Moy_388 3	Moy_389 3	Moy_390 3	Moy_391 3	Moy_392 3	Moy_393 3	Moy_394 3	Moy_395 3	Moy_396 3	Moy_397 3	Moy_398 3	Moy_399 3	Moy_400 3	Moy_401 3	Moy_402 3	Moy_403 3	Moy_404 3	Moy_405 3	Moy_406 3	Moy_407 3	Moy_408 3	Moy_409 3	Moy_410 3	Moy_411 3	Moy_412 3	Moy_413 3	Moy_414 3	Moy_415 3	Moy_416 3	Moy_417 3	Moy_418 3	Moy_419 3	Moy_420 3	Moy_421 3	Moy_422 3	Moy_423 3	Moy_424 3	Moy_425 3	Moy_426 3	Moy_427 3	Moy_428 3	Moy_429 3	Moy_430 3	Moy_431 3	Moy_432 3	Moy_433 3	Moy_434 3	Moy_435 3	Moy_436 3	Moy_437 3	Moy_438 3	Moy_439 3	Moy_440 3	Moy_441 3	Moy_442 3	Moy_443 3	Moy_444 3	Moy_445 3	Moy_446 3	Moy_447 3	Moy_448 3	Moy_449 3	Moy_450 3	Moy_451 3	Moy_452 3	Moy_453 3	Moy_454 3	Moy_455 3	Moy_456 3	Moy_457 3	Moy_458 3	Moy_459 3	Moy_460 3	Moy_461 3	Moy_462 3	Moy_463 3	Moy_464 3	Moy_465 3	Moy_466 3	Moy_467 3	Moy_468 3	Moy_469 3	Moy_470 3	Moy_471 3	Moy_472 3	Moy_473 3	Moy_474 3	Moy_475 3	Moy_476 3	Moy_477 3	Moy_478 3	Moy_479 3	Moy_480 3	Moy_481 3	Moy_482 3	Moy_483 3	Moy_484 3	Moy_485 3	Moy_486 3	Moy_487 3	Moy_488 3	Moy_489 3	Moy_490 3	Moy_491 3	Moy_492 3	Moy_493 3	Moy_494 3	Moy_495 3	Moy_496 3	Moy_497 3	Moy_498 3	Moy_499 3	Moy_500 3	Moy_501 3	Moy_502 3	Moy_503 3	Moy_504 3	Moy_505 3	Moy_506 3	Moy_507 3	Moy_508 3	Moy_509 3	Moy_510 3	Moy_511 3	Moy_512 3	Moy_513 3	Moy_514 3	Moy_515 3	Moy_516 3	Moy_517 3	Moy_518 3	Moy_519 3	Moy_520 3	Moy_521 3	Moy_522 3	Moy_523 3	Moy_524 3	Moy_525 3	Moy_526 3	Moy_527 3	Moy_528 3	Moy_529 3	Moy_530 3	Moy_531 3	Moy_532 3	Moy_533 3	Moy_534 3	Moy_535 3	Moy_536 3	Moy_537 3	Moy_538 3	Moy_539 3	Moy_540 3	Moy_541 3	Moy_542 3	Moy_543 3	Moy_544 3	Moy_545 3	Moy_546 3	Moy_547 3	Moy_548 3	Moy_549 3	Moy_550 3	Moy_551 3	Moy_552 3	Moy_553 3	Moy_554 3	Moy_555 3	Moy_556 3	Moy_557 3	Moy_558 3	Moy_559 3	Moy_560 3	Moy_561 3	Moy_562 3	Moy_563 3	Moy_564 3	Moy_565 3	Moy_566 3	Moy_567 3	Moy_568 3	Moy_569 3	Moy_570 3	Moy_571 3	Moy_572 3	Moy_573 3	Moy_574 3	Moy_575 3	Moy_576 3	Moy_577 3	Moy_578 3	Moy_579 3	Moy_580 3	Moy_581 3	Moy_582 3	Moy_583 3	Moy_584 3	Moy_585 3	Moy_586 3	Moy_587 3	Moy_588 3	Moy_589 3	Moy_590 3	Moy_591 3	Moy_592 3	Moy_593 3	Moy_594 3	Moy_595 3	Moy_596 3	Moy_597 3	Moy_598 3	Moy_599 3	Moy_600 3	Moy_601 3	Moy_602 3	Moy_603 3	Moy_604 3	Moy_605 3	Moy_606 3	Moy_607 3	Moy_608 3	Moy_609 3	Moy_610 3	Moy_611 3	Moy_612 3	Moy_613 3	Moy_614 3	Moy_615 3	Moy_616 3	Moy_617 3	Moy_618 3	Moy_619 3	Moy_620 3	Moy_621 3	Moy_622 3	Moy_623 3	Moy_624 3	Moy_625 3	Moy_626 3	Moy_627 3	Moy_628 3	Moy_629 3	Moy_630 3	Moy_631 3	Moy_632 3	Moy_633 3	Moy_634 3	Moy_635 3	Moy_636 3	Moy_637 3	Moy_638 3	Moy_639 3	Moy_640 3	Moy_641 3	Moy_642 3	Moy_643 3	Moy_644 3	Moy_645 3	Moy_646 3	Moy_647 3	Moy_648 3	Moy_649 3	Moy_650 3	Moy_651 3	Moy_652 3	Moy_653 3	Moy_654 3	Moy_655 3	Moy_656 3	Moy_657 3	Moy_658 3	Moy_659 3	Moy_660 3	Moy_661 3	Moy_662 3	Moy_663 3	Moy_664 3	Moy_665 3	Moy_666 3	Moy_667 3	Moy_668 3	Moy_669 3	Moy_670 3	Moy_671 3	Moy_672 3	Moy_673 3	Moy_674 3	Moy_675 3	Moy_676 3	Moy_677 3	Moy_678 3	Moy_679 3	Moy_680 3	Moy_681 3	Moy_682 3	Moy_683 3	Moy_684 3	Moy_685 3	Moy_686 3	Moy_687 3	Moy_688 3	Moy_689 3	Moy_690 3	Moy_691 3	Moy_692 3	Moy_693 3	Moy_694 3	Moy_695 3	Moy_696 3	Moy_697 3	Moy_698 3	Moy_699 3	Moy_700 3	Moy_701 3	Moy_702 3	Moy_703 3	Moy_704 3	Moy_705 3	Moy_706 3	Moy_707 3	Moy_708 3	Moy_709 3	Moy_710 3	Moy_711 3	Moy_712 3	Moy_713 3	Moy_714 3	Moy_715 3	Moy_716 3	Moy_717 3	Moy_718 3	Moy_719 3	Moy_720 3	Moy_721 3	Moy_722 3	Moy_723 3	Moy_724 3	Moy_725 3	Moy_726 3	Moy_727 3	Moy_728 3	Moy_729 3	Moy_730 3	Moy_731 3	Moy_732 3	Moy_733 3	
-----	--------	------------	------------	------------	-------------	------------	------------	------------	------------	------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	-------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--------------	--

Code	Signification	Type	Longueur
NOM	Le nom de l'étudiant.	Texte	32
PRENOM	Le prénom de l'étudiant.	Texte	32
Moy_1A	Le moyen de la première année.	Numérique	5
Moy_S1	Le moyen de la première semestre de la première année.	Numérique	5
Moy_S2	Le moyen de la deuxième semestre de la première année.	Numérique	5
NV_1A	Le nombre des modules non validé dans la première année .	Numérique	2
NV_S1	Le nombre des modules non validé dans la première semestre de la première année.	Numérique	1
NV_S2	Le nombre des modules non validé dans la deuxième semestre de la première année.	Numérique	1
Mx.y_MOY_Finale	Le moyen du yème module dans la xème semestre.	Numérique	5
$NV_{Mx.y}$	Le module y de la xème semestre est-il valide ? — 0 : si le module est validé (Mx.y-MOY-Finale ≥ 12). — 1 : si le module n'est pas validé.	Numérique	5
PFA	La note du projet de la fin de l'année.	Numérique	5
Décision	La décision du conseil il prend 6 valeurs : — Félicitation si $Moy_{1A} \geq 15$. — Encouragement si $15 > Moy_{1A} \geq 14$. — Admis si $14 > Moy_{1A} \geq 12$ et $NV_{1A} \leq 4$. — Admis avec indulgence si $12 > Moy_{1A} \geq min_moy$ ou $NV_{1A} > 4$. — Ajourné si $min_moy > Moy_{1A}$ (min_moy est décidé par le conseil). — Réorienté si l'étudiant n'était pas admis pendant deux ans.	Texte	32

TABLE 1 – Dictionnaire des données

2.3 Exploration des données

Pour explorer les données il faut tout d'abord transformer nos fichiers en un format adopté par le Spss modeler. Nous avons trois formats : des fichiers excel adopté par le Spss modeler et des fichiers PDF et une image. Pour les fichiers PDF nous avons pu les convertir en fichiers excel facilement. Par contre la conversion de l'image a été faite manuellement puisque sa qualité était mauvaise.

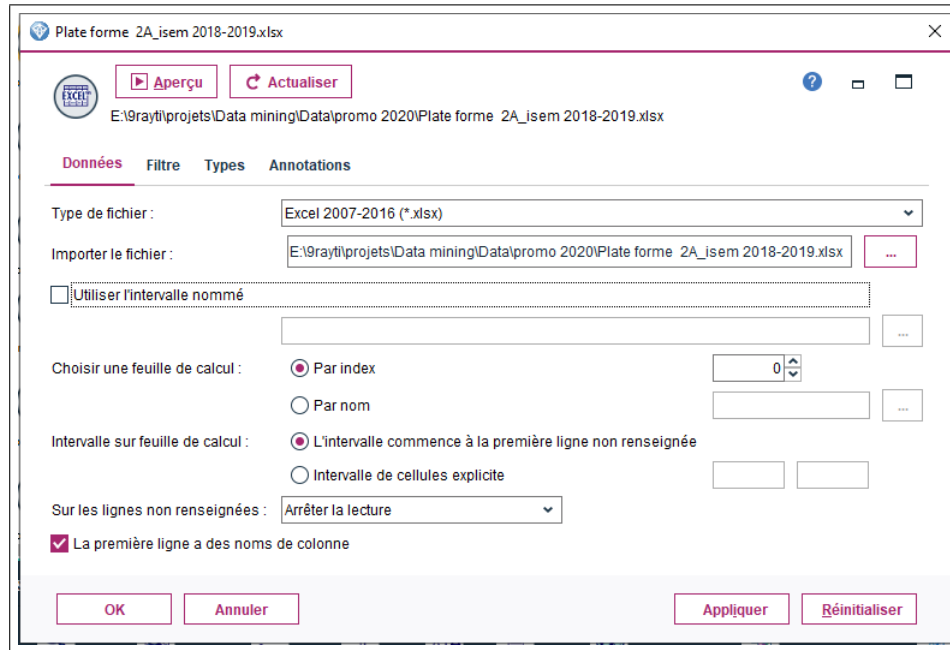


FIGURE 4 – Création des nœuds sources

Nous avons placé des nœuds sources excel et nous avons édité les nœuds en important les fichiers excel qui contiennent les données.

2.3.1 Exploradion à l'aide des tableaux

Moy_S3	NV_S3	M3.1_MOY_Finale	NV_M3.1_Finale	M3.2_MOY_Finale	NV_M3.2_Finale	
10.160	3.000	12.000	0.000	6.000	1.000	
9.710	4.000	12.000	0.000	7.000	1.000	
7.310	8.000	8.080	1.000	4.750	1.000	
9.920	4.000	12.000	0.000	7.250	1.000	
9.710	5.000	10.380	1.000	5.500	1.000	
9.130	6.000	9.780	1.000	5.000	1.000	
10.970	3.000	12.000	0.000	8.000	1.000	
10.160	4.000	10.980	1.000	6.500	1.000	
7.860	8.000	8.600	1.000	6.500	1.000	
8.820	4.000	7.350	1.000	4.500	1.000	
7.910	7.000	12.000	0.000	4.000	1.000	
10.180	5.000	12.000	0.000	9.000	1.000	
8.590	6.000	7.450	1.000	7.250	1.000	
9.540	4.000	12.000	0.000	3.750	1.000	
10.470	3.000	12.000	0.000	8.750	1.000	
9.710	4.000	12.000	0.000	7.000	1.000	
8.800	7.000	10.980	1.000	6.500	1.000	
9.490	5.000	10.700	1.000	6.000	1.000	
9.380	5.000	9.250	1.000	8.000	1.000	
10.050	4.000	12.000	0.000	6.000	1.000	

FIGURE 5 – Explorer les données à l'aide des tableaux

Nous avons connecté un nœud tableau avec chaque nœud source pour visualiser les données.

2.3.2 Visualisation du pourcentage de réussite

Il sera utile de connaître le pourcentage des gens qui ont réussi avec félicitation et le pourcentage des admis et aussi les autres catégories donc nous avons choisi d'explorer cette donnée. D'après le

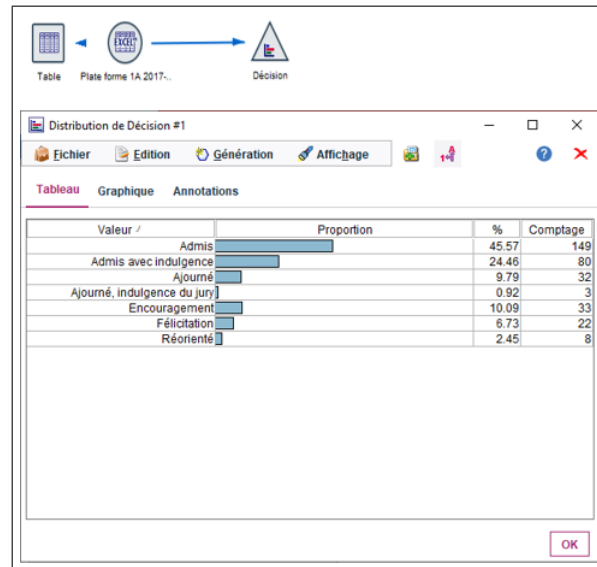


FIGURE 6 – Pourcentage de réussite

résultat du nœud Distribution, la plupart des étudiants sont des admis. Ainsi nous avons constaté que nous avons seulement 22 étudiants qui réussirent avec félicitation, c'est-à-dire ils ont une note supérieure à 15.

2.3.3 Visualisation du pourcentage des filières

Il sera utile aussi de visualiser le pourcentage et le nombre des étudiants dans chaque filière. Pour savoir le nombre des groupes quand on aura. Pour cela nous avons ajouté un champ qui contient la filière de chaque étudiant en utilisant le nœud « calculer ». Mais avant nous allons extraire le champ « NOM », puisque dans le fichier des étudiants « GL » nous disposons seulement des noms des étudiants, avec le nœud « filter ».

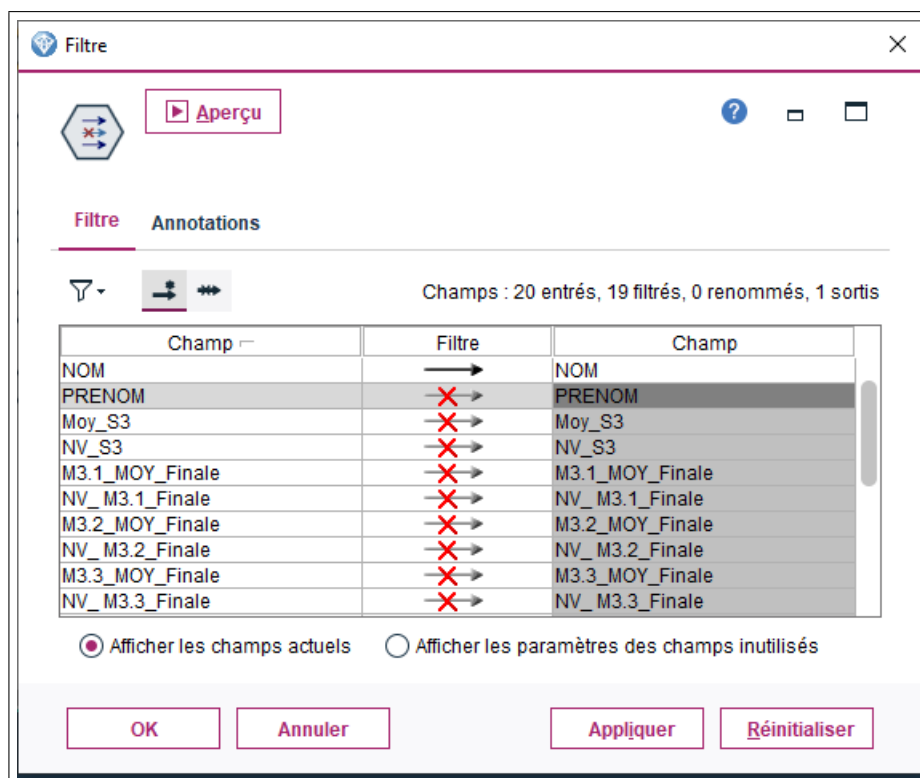


FIGURE 7 – Extraire le champ des noms

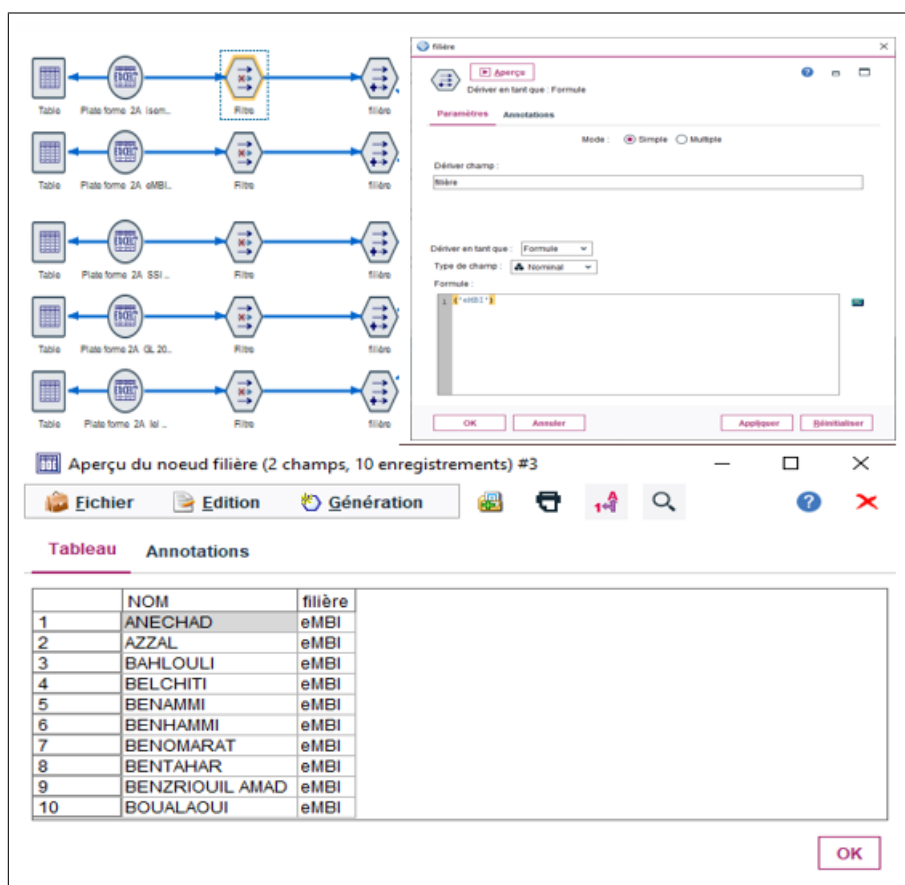


FIGURE 8 – Ajouté la colonne de la filière

Puis on va faire une union entre les nœuds résultants et le nœud qui contient les notes de la première année avec le nœud « ajouter ».

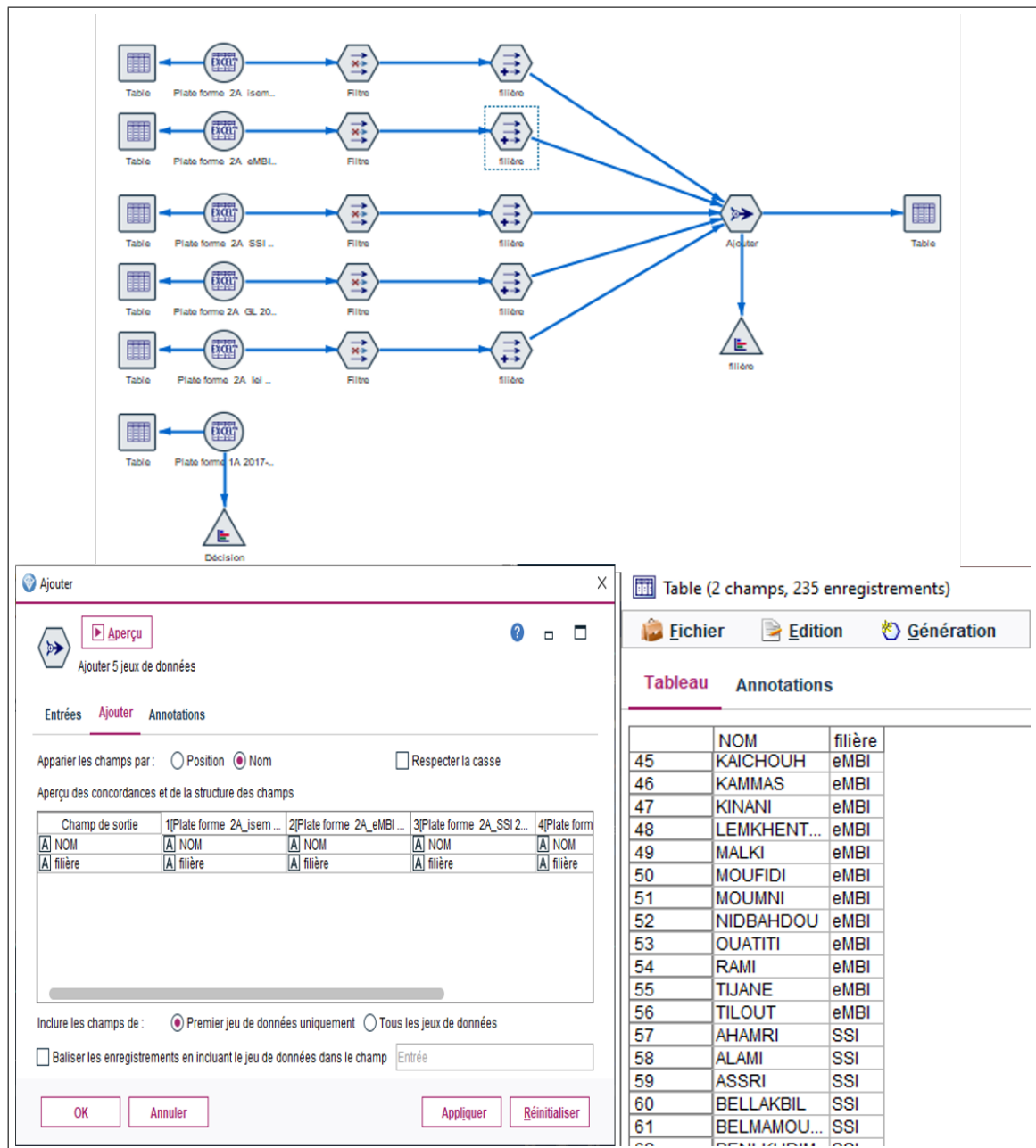


FIGURE 9 – Union des nœuds

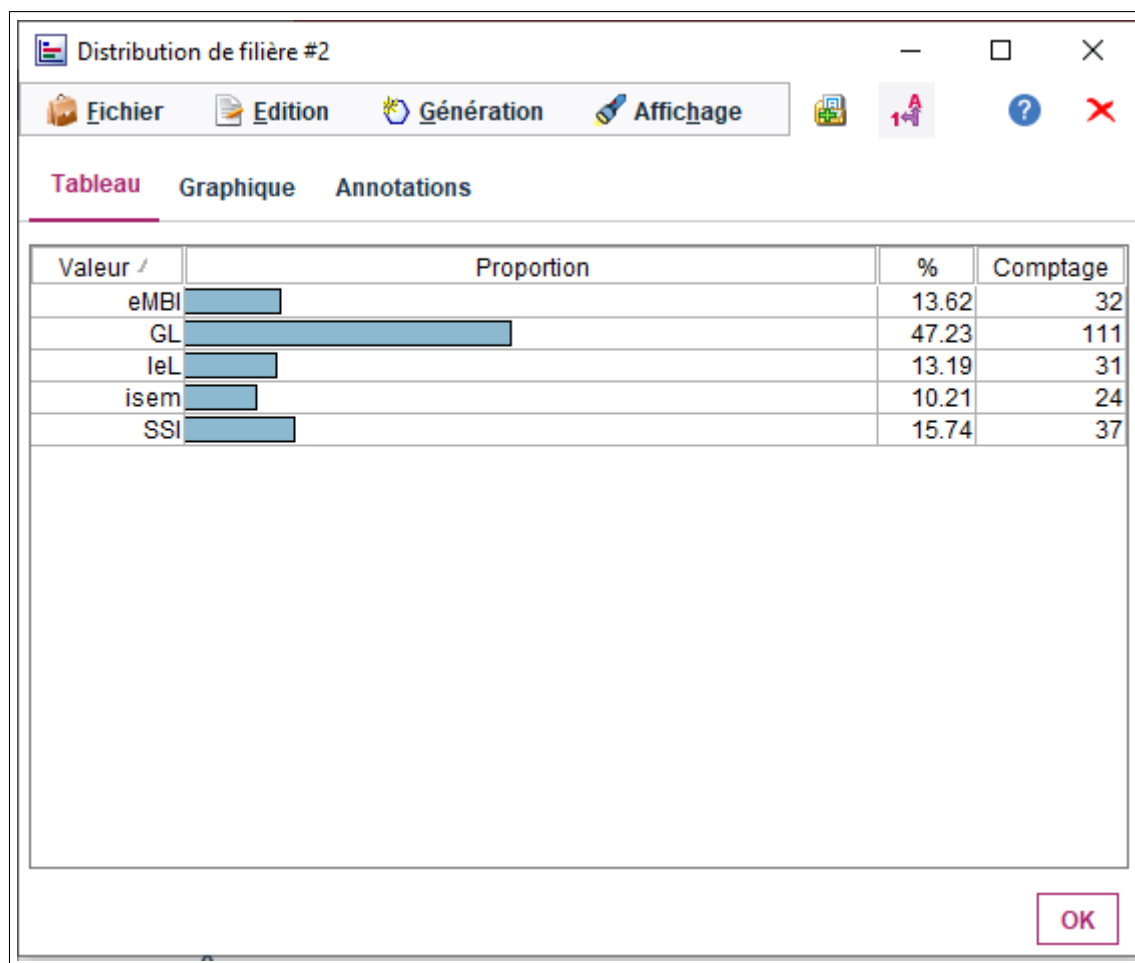


FIGURE 10 – Pourcentage des filières

Nous avons remarqué que l'effectif des étudiants en « GL » est le plus grand. Par contre l'effectif en « Isem » est le plus petit. Ainsi on peut constater que les filières « eMBI », « Isem » contient un nombre pair des élèves donc on n'aura que des binômes. Par contre dans les autres filières, il y a un nombre impair des étudiants donc il y aura des monômes.

2.4 Vérification de la qualité des données

Pour la qualité des données tous les fichiers ont des champs complets sauf le fichiers qui contient les notes de la filière SSI.

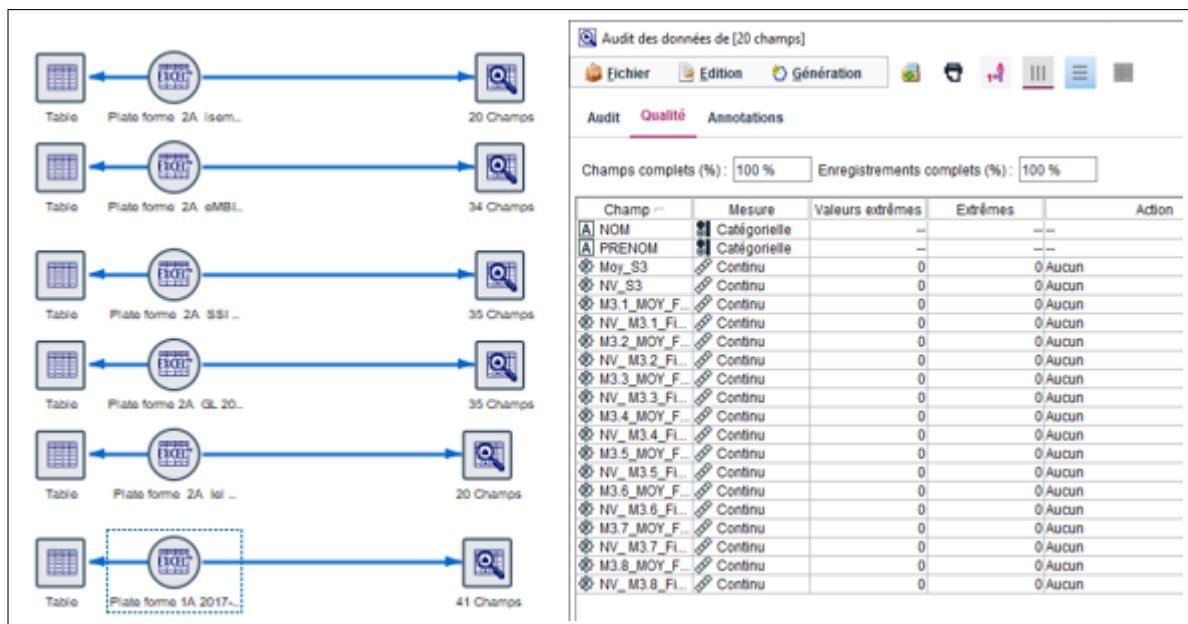


FIGURE 11 – Qualité des données

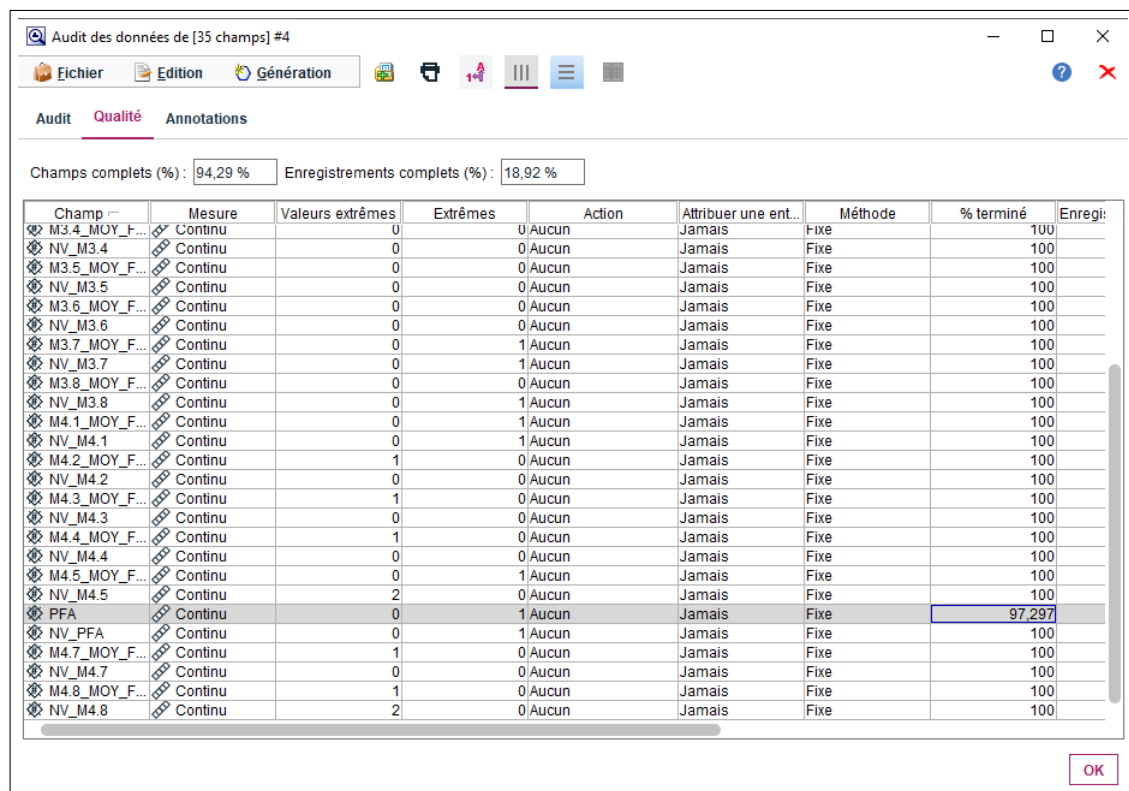


FIGURE 12 – Qualité des données de nœud notes SSI

Il y' a un manque des valeurs dans le champ status (30 valeurs manquantes) ainsi dans le champ PFA il y a une seule valeur manquante.

3 Préparation des données

La préparation des données est l'un des aspects les plus importante et les plus coûteux en temps du Data Mining. Dans cette section, on va voir la sélection des données, le nettoyage des donnés et l'intégration des données.

3.1 Sélection des données

Après avoir effectué la collecte initiale de données on doit maintenant choisir les données pertinentes pour nos objectifs de Data Mining.

Les données que nous avons besoin sont : les noms et les prénoms des étudiants et leurs notes.

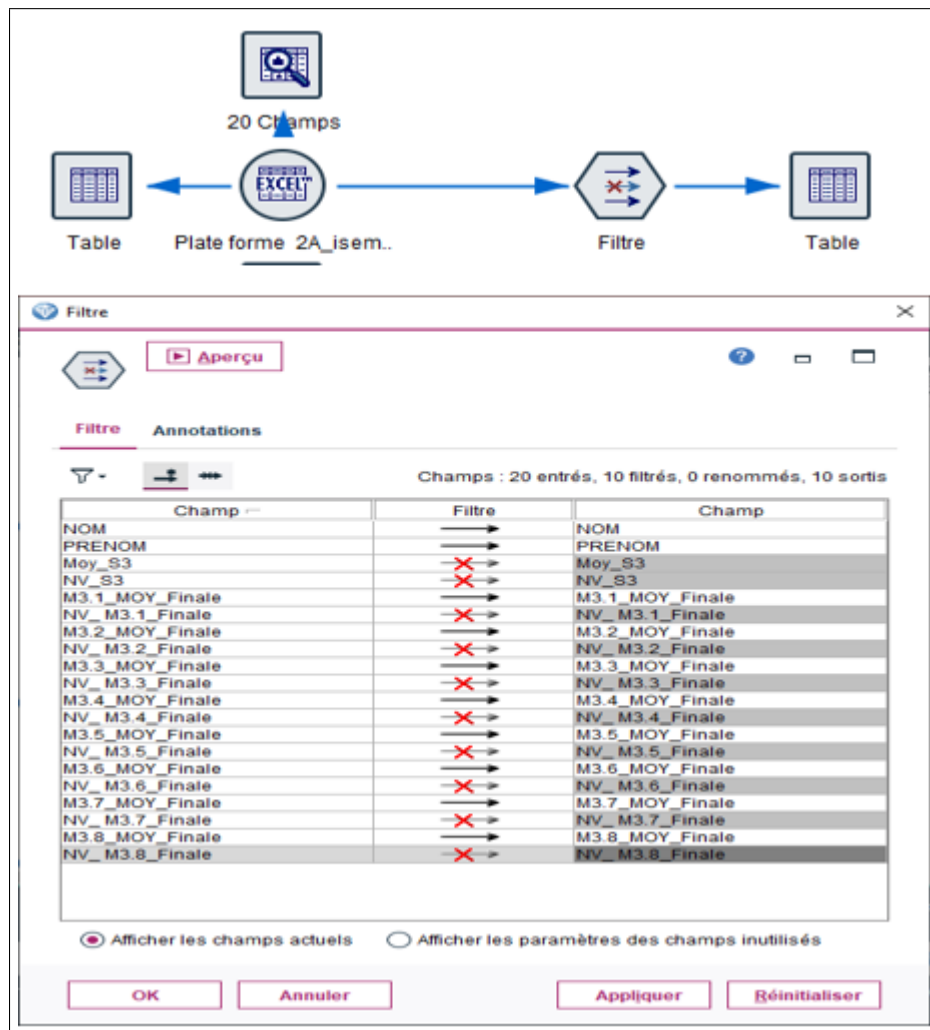


FIGURE 13 – Sélection des données

Pour tous les fichiers, nous avons sélectionné les noms, les prénoms et les notes des étudiants dans les huit modules et nous avons éliminé les autres attributs en utilisant le nœud « Filter ».

3.2 Nettoyage des données

3.2.1 Données manquantes

Comme nous avons déjà montré dans le chapitre précédant il y a une seule valeur manquante. C'est la note du projet de la fin d'année (PFA) d'un étudiant dans la filière SSI. On va donc la

remplacer avec la valeur moyenne.

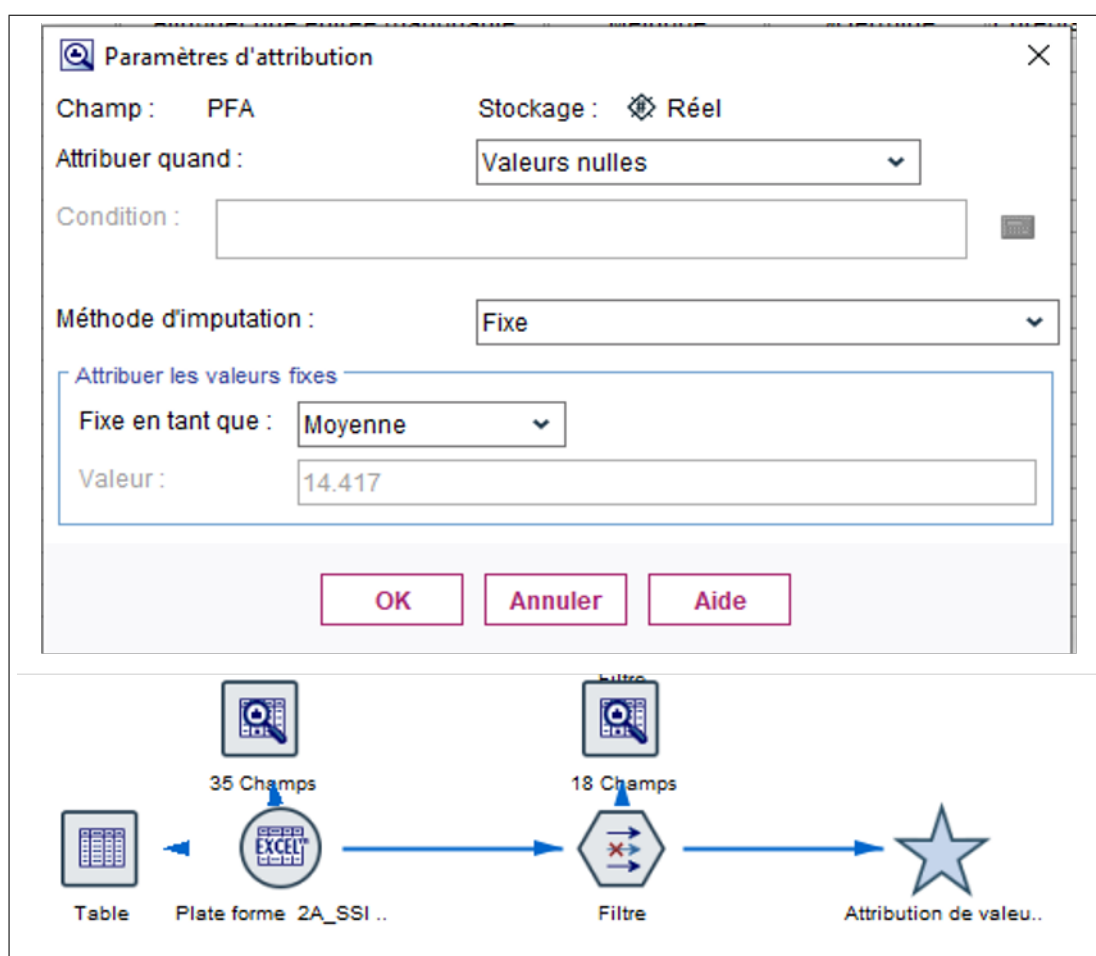


FIGURE 14 – Insertion d'une valeur manquante

3.2.2 Erreurs dans les données

Nous avons vérifié manuellement les valeurs de telle manière que la valeur d'une note soit incluse entre 0 et 20.

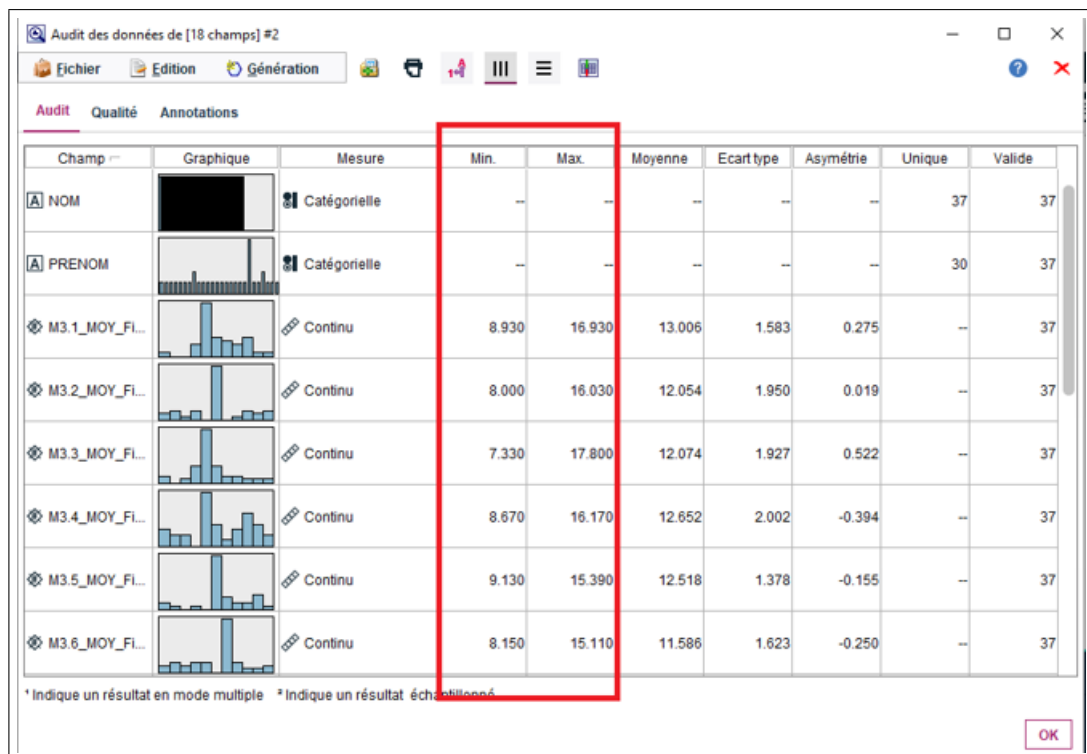


FIGURE 15 – vérification des données

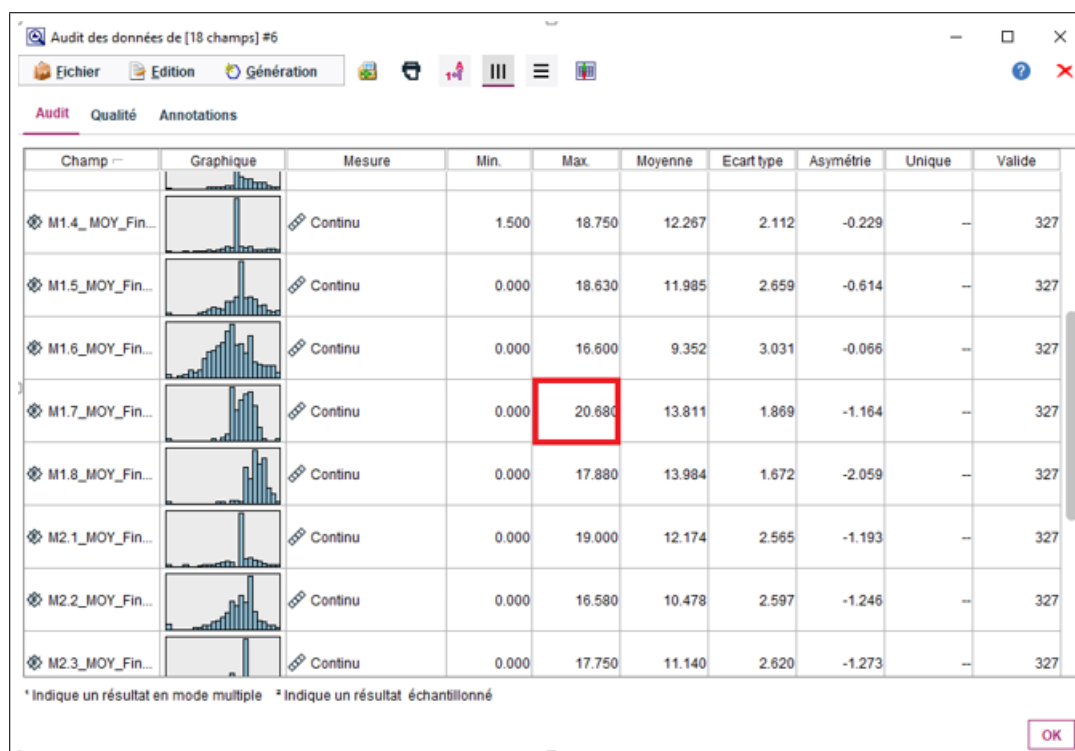


FIGURE 16 – Détection d'une valeur erronée

Nous avons trouvé une valeur anormale au niveau du module 1.7 dans les notes de la première année, la note est 20.68 ce qui est impossible.

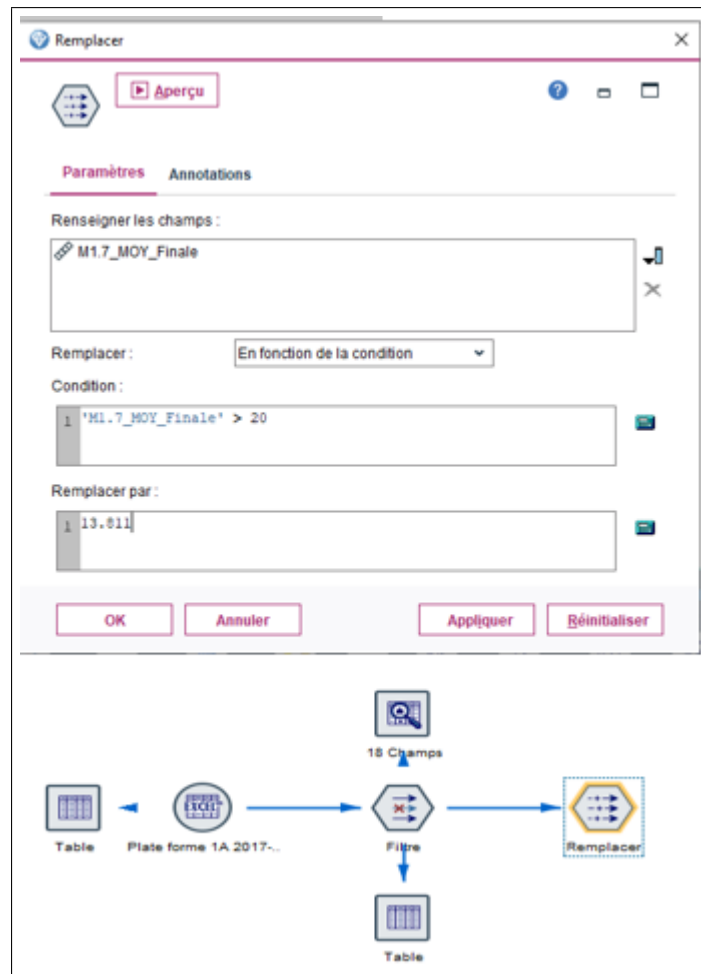


FIGURE 17 – Remplacer la valeur erronée par la valeur moyenne

On va remplacer cette valeur par la moyenne qui est 13.811.

3.3 Intégration des données

Il existe deux méthodes principales pour l'intégration des données : la fusion et l'ajout. Dans notre cas on va faire la fusion des données. Nous avons fusionné les ensembles de données de la première année avec les données de la deuxième année de chaque filière.

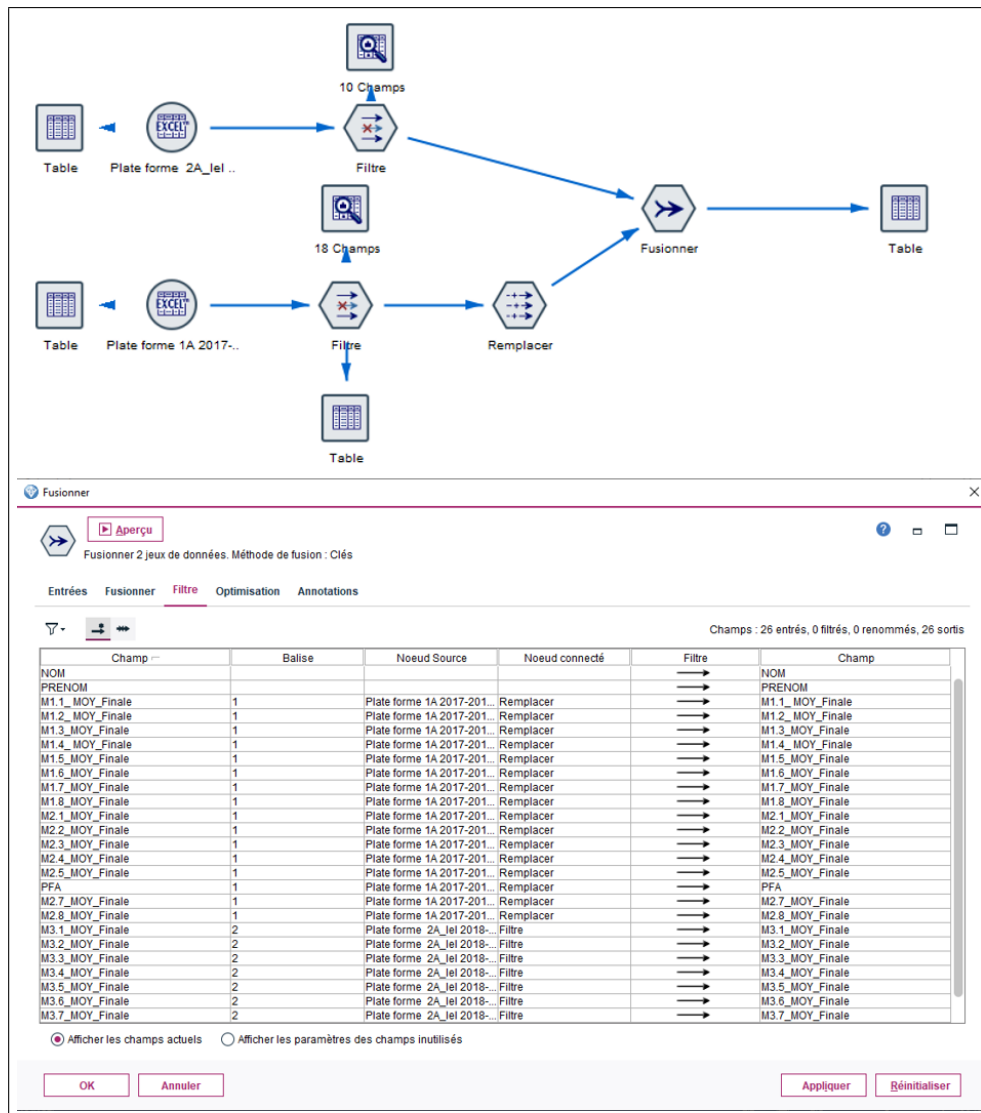


FIGURE 18 – Fusionner les données de la première année et de la deuxième année de la filière Iel

Par exemple dans la figure 18 on a la fusion des deux ensembles des données : les données de la première année et les données de la deuxième année des étudiants de la filière Iel.

4 Modélisation

La quatrième étape est la modélisation qui est au cœur de tout projet d'apprentissage automatique. Cette étape est responsable des résultats qui devraient satisfaire ou aider à atteindre les objectifs du projet.

Bien que ce soit la partie glamour du projet, c'est aussi la plus courte dans le temps, car si tout ce qui précède est fait correctement, il y a peu à ajuster. Dans le cas où les résultats sont améliorables, la méthodologie est définie pour revenir à la préparation des données et améliorer les données disponibles.

Dans ce chapitre, on va voir la sélection des techniques de modélisation, la génération d'une conception de test et la construction du modèle.

4.1 Sélection des techniques de modélisation

Il y a plusieurs types de modélisation. Le choix du modèle le plus adéquat sera généralement basé sur les critères suivants :

- Les types de données disponibles pour l'exploration :

Les champs intéressants sont numérique.

- Les objectifs de data mining :

Notre problème est un problème de classification non supervisé. Nous avons plusieurs méthodes de clustering : méthode hiérarchiques, méthodes de partitionnement, méthodes mixtes et analyse floue. Nous avons choisi de travailler avec l'algorithme de K-means : c'est une méthode de partitionnement. Il permet de regrouper en K clusters distincts les observations du data set.

Nous avons choisi K-means pour ces raisons :

- Simple et robuste.

- On définit nous même le nombre de clusters.

- Efficace : $O(n.t.k)$ avec n : le nombre des objets, t : nombre des itérations et k : le nombre des clusters.

4.2 Génération d'une conception de test

La dernière étape avant la création du modèle consiste à considérer la façon dont les résultats du modèle seront testés. Puisqu'on est dans le cas d'un modèle non supervisé le critère de « qualité d'ajustement » de notre modèle sera la taille des clusters. En effet nous voulons minimiser les groupes avec une seule personne. Ainsi il sera comme critère la facilité d'interprétation et le temps de traitement nécessaire.

4.3 Construire le modèle

Nous avons commencer par ajouter le noeud « typer » pour les définir les données entrants du modèle.

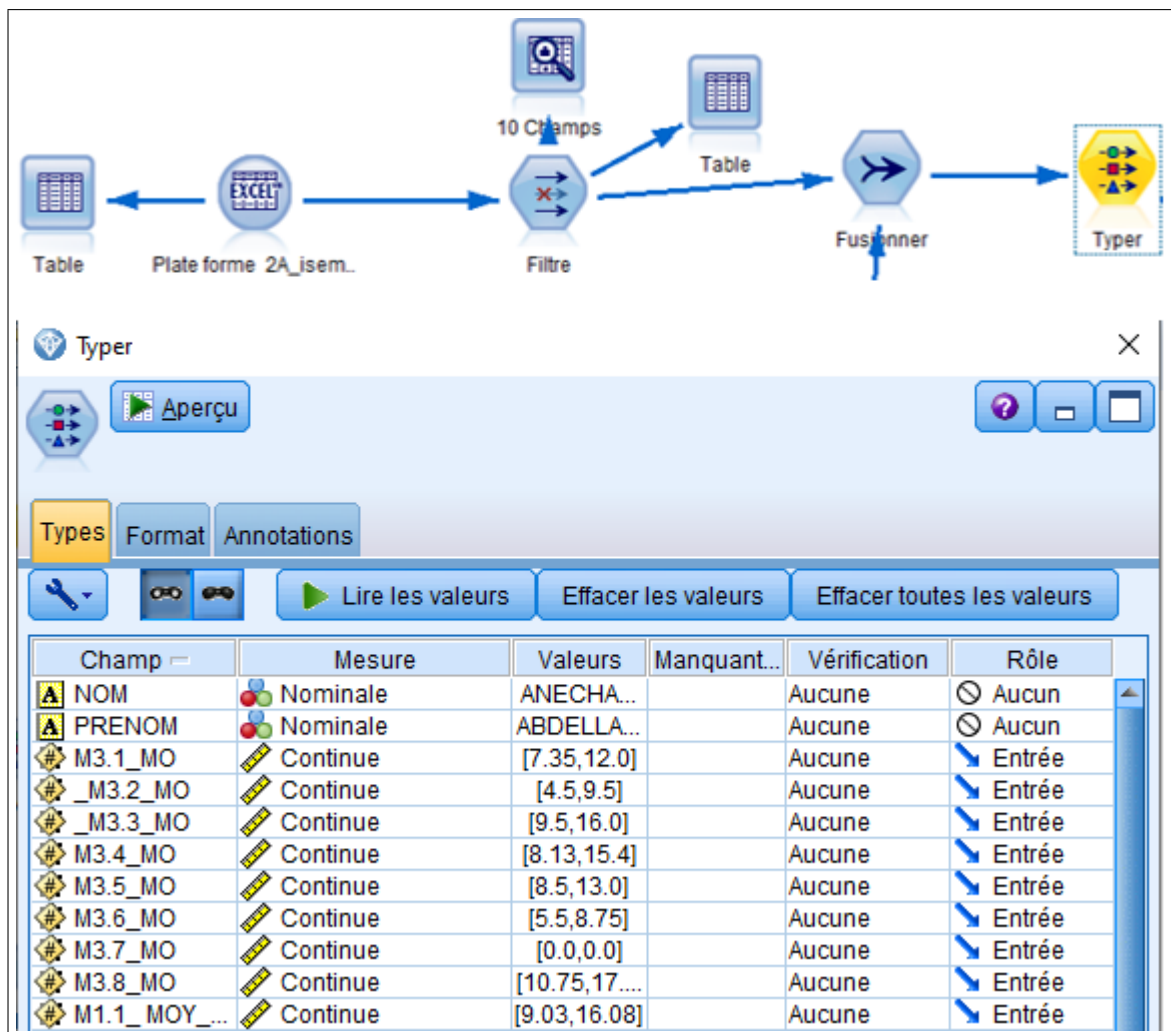


FIGURE 19 – Définir les inputs du modèle

Nous avons défini tous les champs comme des champs « entrée » sauf nom et prénom que nous avons défini leur rôle comme « aucun ».

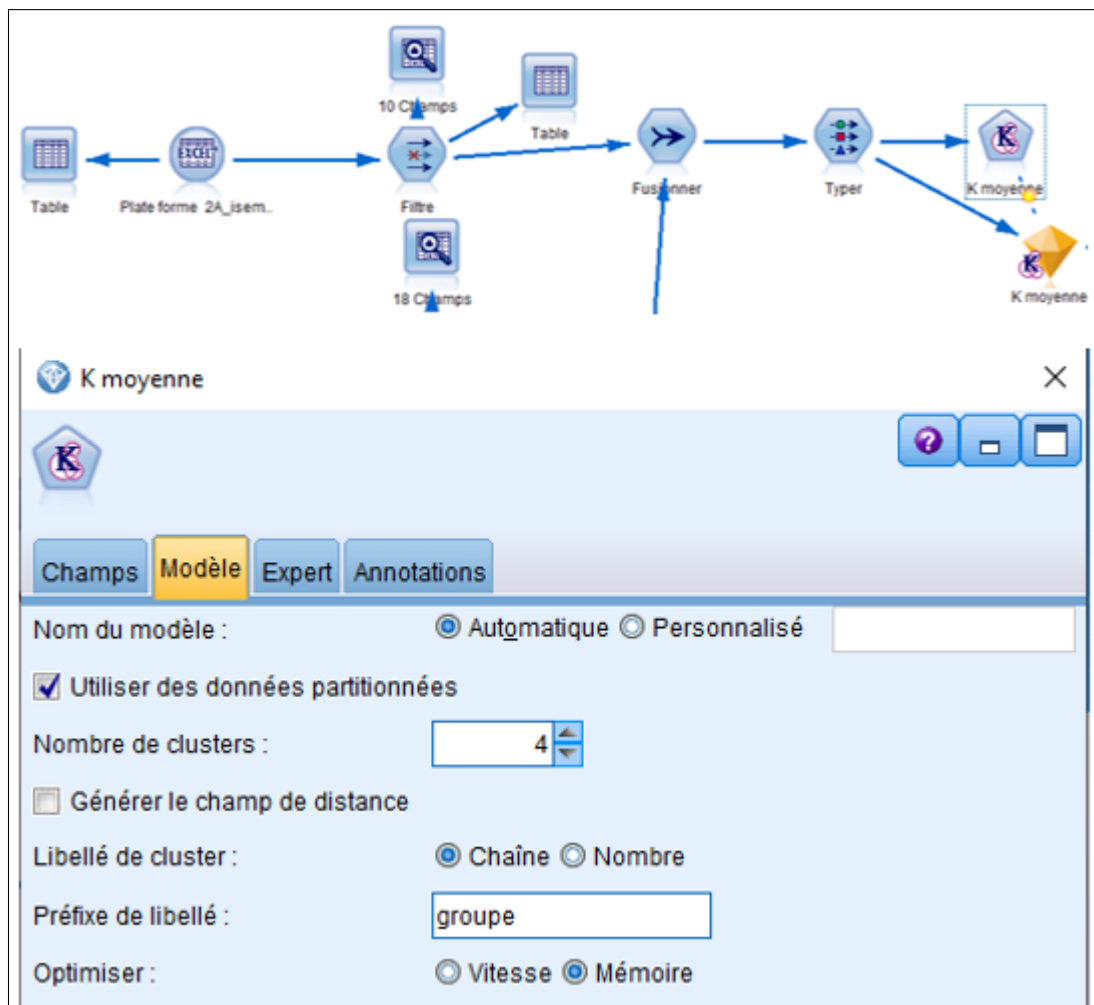


FIGURE 20 – Génération du modèle

Nous avons fixé le nombre de clusters pour la filière isem sur 4. Nous aurons donc 4 groupes des étudiants chaque groupe est constitué des étudiants ayant des notes similaires.

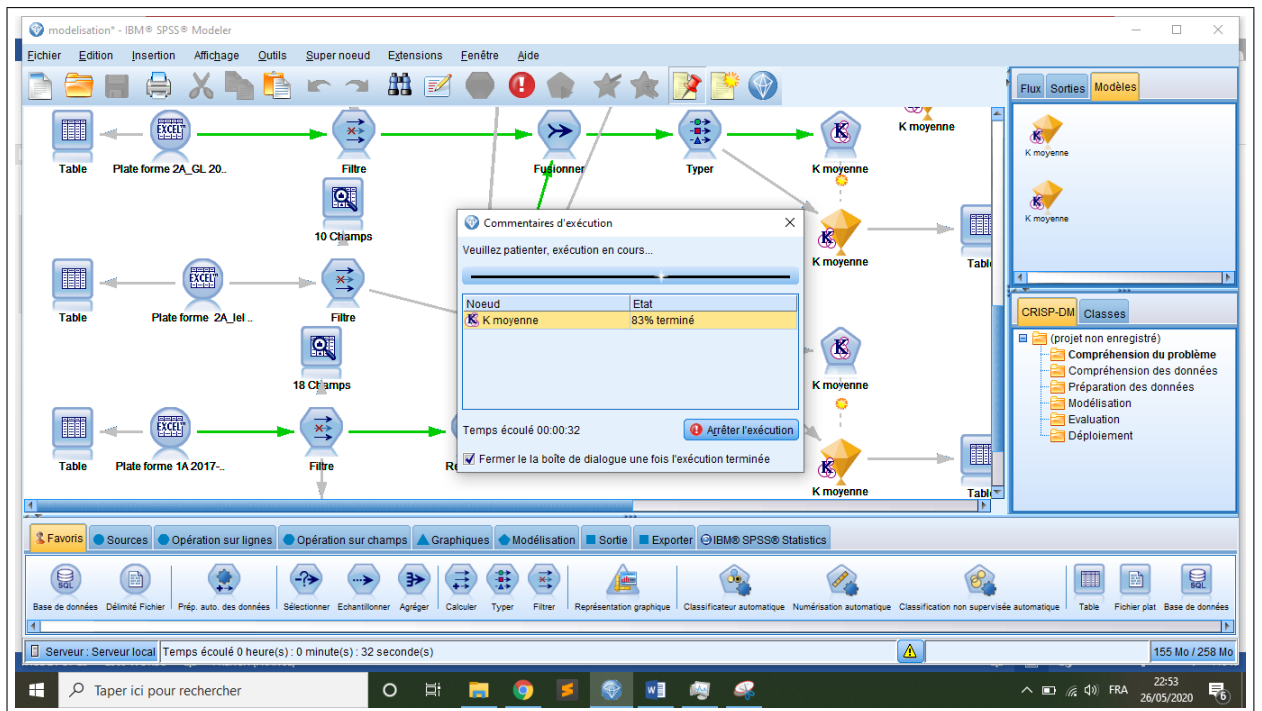


FIGURE 21 – Temps d'exécution

Pour le temps d'exécution, il est raisonnable par exemple dans le cas de la filière GL le temps d'exécution est 33 secondes

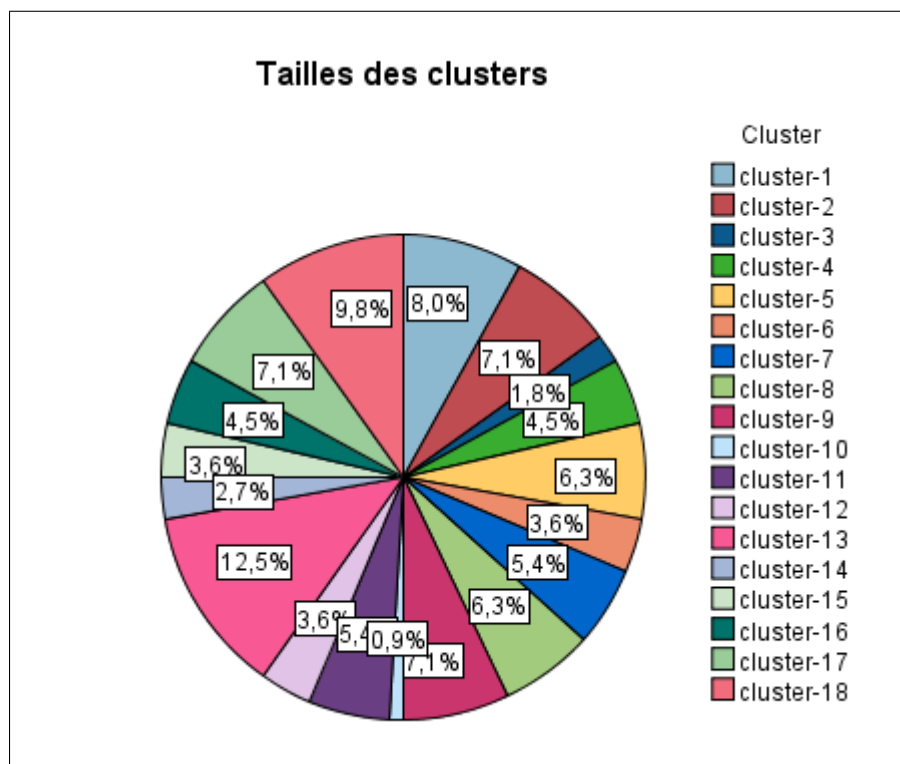


FIGURE 22 – Taille des clusters

Nous avons obtenu des clusters avec différents tailles. On a un seul cluster de taille 1. Donc nous avons atteint notre objectif.



FIGURE 23 – Importance des prédicteurs

Dans le cas des étudiants de la filière Génie Logiciel (GL) le prédicteur le plus important est la note du module 1.2.

M1.1...	M1.2...	M1.3...	M1.4...	M1.5...	M1.6...	M1.7...	M1.8...	M2.1...	M2.2...	M2.3...	M2.4...	M2.5...	PFA	M2.7...	M2.8...	\$KM-groupes GL
15.250	16.770	15.500	18.200	15.380	13.450	16.240	15.500	13.500	14.150	13.700	15.800	18.600	15....	12.000	15.380	groupe-1
12.000	16.070	12.000	12.000	12.000	8.750	14.310	16.130	13.750	13.830	10.330	16.100	19.300	16....	12.000	16.380	groupe-1
11.250	13.180	12.000	12.000	10.380	8.050	12.450	14.500	6.750	9.830	12.000	13.600	15.670	16....	8.250	12.000	groupe-15
15.080	15.500	17.500	17.680	14.500	14.350	16.540	15.130	16.250	12.000	11.750	15.800	17.220	18....	15.380	14.000	groupe-11
12.000	12.000	12.000	12.000	12.000	12.600	12.810	15.630	12.000	9.500	9.830	12.700	15.410	14....	13.500	15.500	groupe-4
12.000	17.750	12.000	12.000	11.130	9.900	12.000	14.500	12.500	11.170	13.080	14.100	12.000	16....	12.000	12.500	groupe-8
12.000	12.980	15.750	12.780	14.250	10.250	15.600	15.130	12.250	10.830	11.750	14.400	17.670	15....	12.250	13.500	groupe-9
9.950	14.820	13.000	13.950	12.000	12.000	16.480	16.880	14.250	10.330	14.540	14.700	17.760	18....	14.500	16.250	groupe-9
12.000	12.000	12.000	12.000	8.500	6.200	12.000	13.750	12.000	10.170	11.500	12.000	13.180	16....	12.000	12.500	groupe-15
12.000	14.050	12.000	11.400	10.630	11.250	15.100	15.250	12.000	10.500	12.000	12.050	17.300	16....	15.000	14.630	groupe-9
15.650	17.000	17.500	17.380	16.130	15.350	15.940	16.880	15.500	12.000	14.250	17.800	19.580	17....	13.630	16.250	groupe-11
12.200	14.800	12.000	12.000	11.630	15.650	14.460	14.500	12.000	12.000	7.330	12.000	17.480	14....	12.250	15.130	groupe-4
13.500	14.370	15.250	16.230	15.750	12.000	15.000	14.630	13.250	12.580	9.000	13.200	13.950	13....	12.380	14.630	groupe-17
12.000	14.680	15.500	12.000	14.630	10.350	15.750	13.880	9.750	11.500	10.460	12.000	17.040	15....	13.130	12.000	groupe-17
13.750	16.770	16.250	14.180	18.000	15.900	15.350	17.000	12.500	13.500	15.480	16.600	19.160	14....	14.630	18.130	groupe-1
12.000	12.500	12.000	12.000	12.000	10.250	12.000	14.000	15.250	9.000	12.000	16.100	12.850	13....	13.630	12.500	groupe-13
13.680	12.890	12.000	12.000	12.000	10.600	12.350	15.380	12.000	8.170	6.760	12.000	15.640	13....	12.500	15.250	groupe-13
14.480	15.770	17.000	15.630	17.750	13.600	16.200	16.630	12.000	14.580	11.650	15.600	18.880	16....	16.000	16.880	groupe-14
13.950	12.000	12.000	12.000	10.250	5.600	12.000	14.750	12.000	12.000	12.000	12.000	12.000	14....	12.000	12.000	groupe-18
11.450	12.840	16.500	17.630	13.380	7.800	16.050	14.250	13.500	12.830	12.000	12.100	18.700	17....	13.000	14.000	groupe-9
10.730	12.000	13.750	12.000	12.000	9.350	12.550	12.000	14.750	12.000	12.000	12.100	13.210	13....	12.000	14.250	groupe-13
12.000	13.320	12.500	12.000	9.000	6.200	12.150	12.000	13.250	10.580	13.730	11.400	14.700	18....	13.880	14.000	groupe-7
15.780	17.980	12.500	16.600	13.130	11.050	16.400	14.000	17.750	10.330	15.140	18.300	19.160	15....	12.000	14.750	groupe-11
10.750	13.800	11.000	12.000	12.000	11.200	12.880	14.000	9.000	9.920	10.250	10.100	12.180	12....	12.000	12.250	groupe-2
12.000	15.020	12.000	10.150	12.000	8.550	12.880	14.250	10.000	9.830	12.000	12.000	12.950	12....	12.000	13.750	groupe-4
12.130	16.590	16.750	14.280	12.000	12.800	17.010	14.250	16.750	15.750	12.000	16.200	17.440	17....	16.750	14.500	groupe-1
15.180	17.320	16.750	12.780	18.630	13.350	13.540	12.500	13.500	12.500	13.000	18.800	16.090	13....	12.750	12.500	groupe-17
12.530	15.950	14.000	12.000	8.130	10.650	14.200	13.500	12.000	12.000	10.050	12.000	13.570	17....	12.000	15.250	groupe-5
15.080	16.700	14.750	13.230	15.630	9.950	17.100	15.500	13.500	12.000	12.010	12.600	16.710	15....	16.130	16.000	groupe-1
8.500	13.520	14.750	12.000	12.000	5.900	14.400	15.000	14.500	9.420	9.650	12.000	14.430	17....	12.000	15.000	groupe-18
15.300	16.070	16.500	14.500	16.000	11.000	13.690	13.000	12.000	15.670	15.080	17.700	16.410	17....	12.250	12.500	groupe-11
14.000	15.680	12.000	13.580	12.880	12.050	14.730	15.000	10.500	11.670	14.080	13.200	16.060	16....	10.500	15.000	groupe-17
11.050	15.270	16.250	12.000	12.750	10.450	14.300	14.000	13.000	12.000	11.500	12.700	12.000	15....	12.000	14.500	groupe-13
12.000	12.520	12.000	14.200	9.880	8.600	15.980	16.000	12.000	10.830	12.000	12.000	13.350	16....	12.000	16.500	groupe-18
13.000	14.820	11.500	12.700	11.500	9.450	11.820	12.000	12.000	12.000	12.000	12.000	15.810	13....	10.130	13.000	groupe-8

FIGURE 24 – Résultat

5 Evaluation et déploiement

5.1 Evaluation

5.1.1 Qualité du modèle

Pour mesurer la qualité de notre modèle, nous prenons en considération la qualité des clusters. Dans notre cas, elle est correcte (Silhouette moyenne = 0.3).

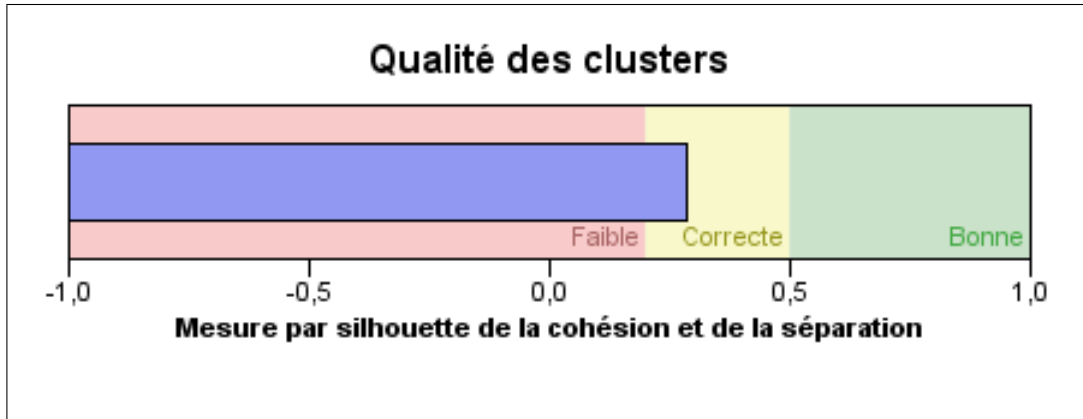


FIGURE 25 – Qualité des clusters

5.1.2 Amélioration

Pour améliorer notre projet, il sera préférable d'obtenir des cluster de la même taille. Pour que les étudiant auront la même liberté de choisir leurs groupes de travail.

5.2 Déploiement

Le déploiement est le processus consistant à utiliser les nouvelles connaissances pour apporter des améliorations au sein de l'établissement.

Notre projet va aider les professeurs de l'ENSIAS à former des groupes homogènes des étudiants pour qu'ils puissent exploiter leurs compétences.

Conclusion

A travers ce projet, nous avons pu en effet consolider les connaissances acquises et d'enrichir Notre expérience en matière de data mining. A travers le sujet choisi « Analyse des relevés de notes des étudiants de l'ENSIAS », on constate que le data mining peut être appliqué dans tous les domaines et s'avère utile pour tous les décideurs.

Pour atteindre notre objectif nous avons mis en place la démarche CRISP-DM (Cross Industry Standard Process for Data Mining). Il s'agit d'un modèle de processus de data mining qui décrit une approche communément utilisée par les experts en data mining pour résoudre les problèmes qui se posent à eux.

Références

1. **Cours Data Mining : Concepts et Techniques**, de Madame Houda Benbrahim
2. **Documentation de IBM SPSS Modeler**, disponible sur : https://www.ibm.com/support/knowledgecenter/SS3RA7_sub/modeler_kc_subscription/clementine/knowledge_center/product_landing_subscription.html
3. **Overleaf** : Editeur de LaTeX en ligne, disponible sur : <https://fr.overleaf.com/>
4. **Documentation de Overleaf**, disponible sur : overleaf.com/learn