

Students Performance

Mathilde Cornille | Ludovic Darguy | Louis-Baptiste Rabillon

5 ISS A

Introduction

Dans le cadre du projet de traitement de masse de données, nous avons analysé une base de données recensant, pour 1000 cas, les performances d'étudiants selon plusieurs critères. Ces données proviennent du site <https://www.kaggle.com/datasets> et sont incluses au format csv dans le projet. Nous avons ensuite exploité les données qui nous paraissaient utiles afin de répondre à plusieurs questionnements. Nous mettons donc en relation les performances des étudiants dans différentes matières avec leur environnement familial, leur genre ou encore leur régime alimentaire afin de définir s'il existe un contexte plus favorable à la réussite scolaire.

Avant propos

Afin d'analyser nos résultats nous avons souhaité définir certaines règles en amont. Ayant uniquement des notes dans trois matières distinctes à disposition nous avons choisi d'en faire la moyenne. Nous avons ainsi défini une moyenne générale pour chaque étudiant, ce qui nous permet d'alléger certains graphiques. Nous avons également penser qu'il n'était pas nécessaire, pour certaines études, de connaître les notes exactes mais plutôt le taux de réussite. C'est pourquoi nous avons choisi de définir un seuil de réussite à 50/100. Les étudiants ayant une note supérieure ou égale à ce seuil obtiendront donc "PASS" tandis que les étudiants ayant une note inférieure auront "FAIL".

Il nous semble également important de préciser que nous n'avons pas utilisé les données de la colonne "race/ethnicity" car nous ne comprenons pas comment les personnes ont été attribuées aux différents groupes (A à E), nous ne pouvons donc pas en tirer de conclusions.

Nos études

Etude n°1 : Le pourcentage de réussite dépend-t-il du genre de l'étudiant?

Dans une première étude nous nous sommes questionnés sur l'influence du genre sur les résultats scolaires. Pour cela nous avons d'abord déterminé le nombre d'hommes et de femmes dans l'échantillon. En effet, il est important de se rapporter au nombre total afin de pouvoir analyser les résultats. Le jeu de données recense 518 femmes et 482 hommes, ce qui rend les données relativement comparables.

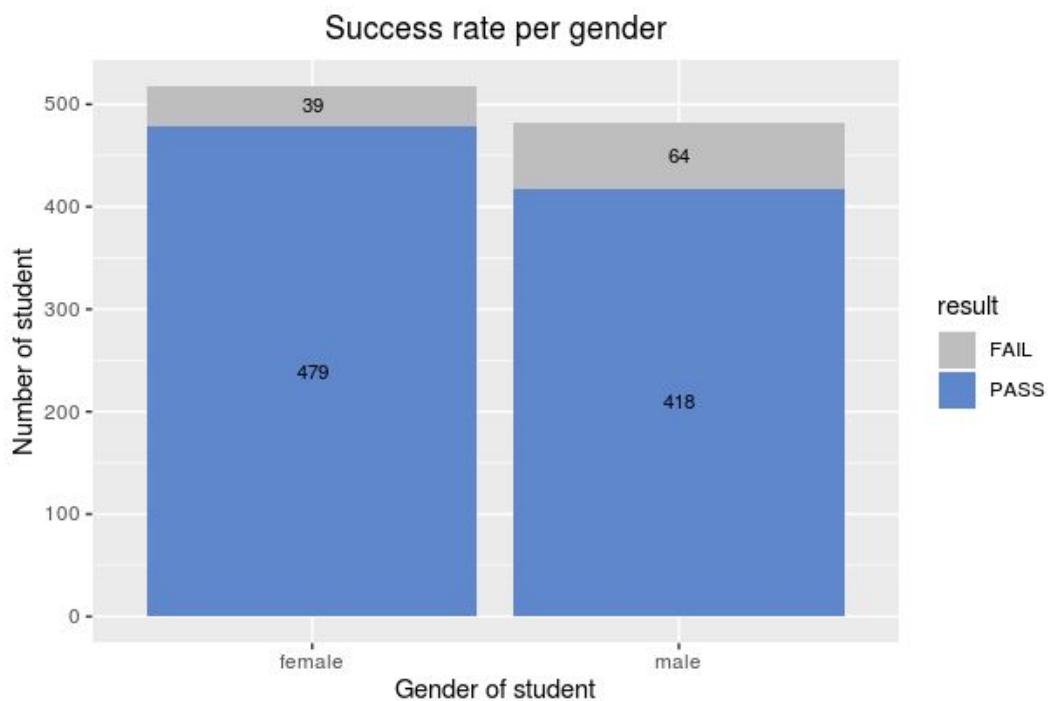


Figure 1 : Résultats selon le genre de l'étudiant.

Conclusion : Nous observons un taux de réussite moyen plus élevé chez les femmes avec $479/518 = 92.5\%$ de validations contre un taux de $418/482 = 86.7\%$ chez les hommes.

Etude n°2 : Le régime alimentaire suivi par l'étudiant influence-t-il ses résultats?

Afin de réaliser cette étude nous avons décidé de tracer deux diagrammes circulaires. Le premier représente le ratio de "PASS" et "FAIL" chez les personnes suivant un régime alimentaire particulier. Le second représente le même ratio, chez les personnes suivant un régime alimentaire standard.

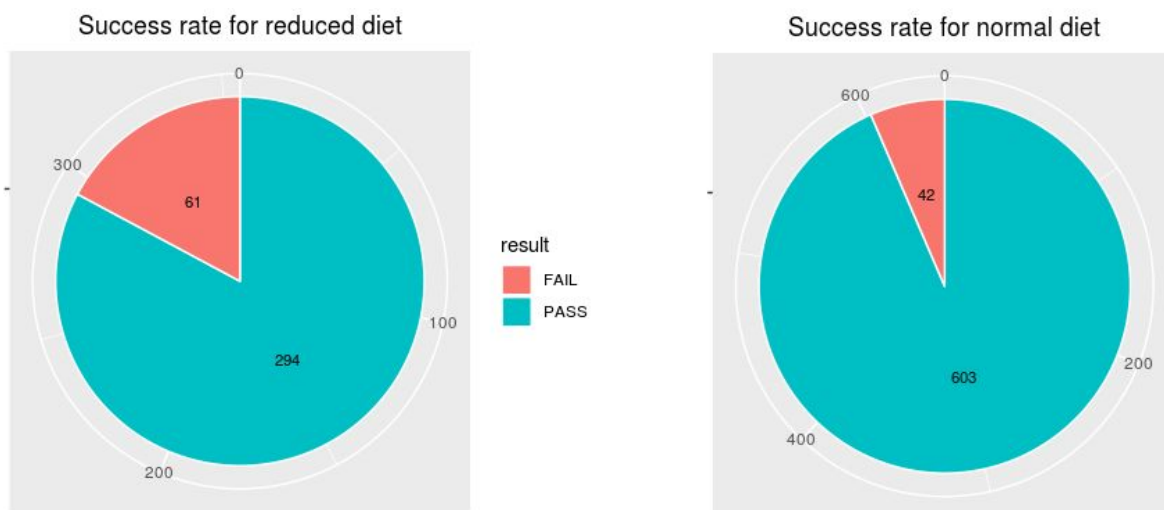


Figure 2 : Influence d'un régime alimentaire adapté sur les résultats scolaires.

Conclusion : Il est intéressant de noter une différence considérable entre les deux observations. En effet, régime alimentaire et performances ne semble que très peu liés, pourtant nous observons une forte distinction entre les deux catégories étudiées.

Etude n°3 : Un étudiant est-il destiné à réussir davantage dans une matière selon son genre?

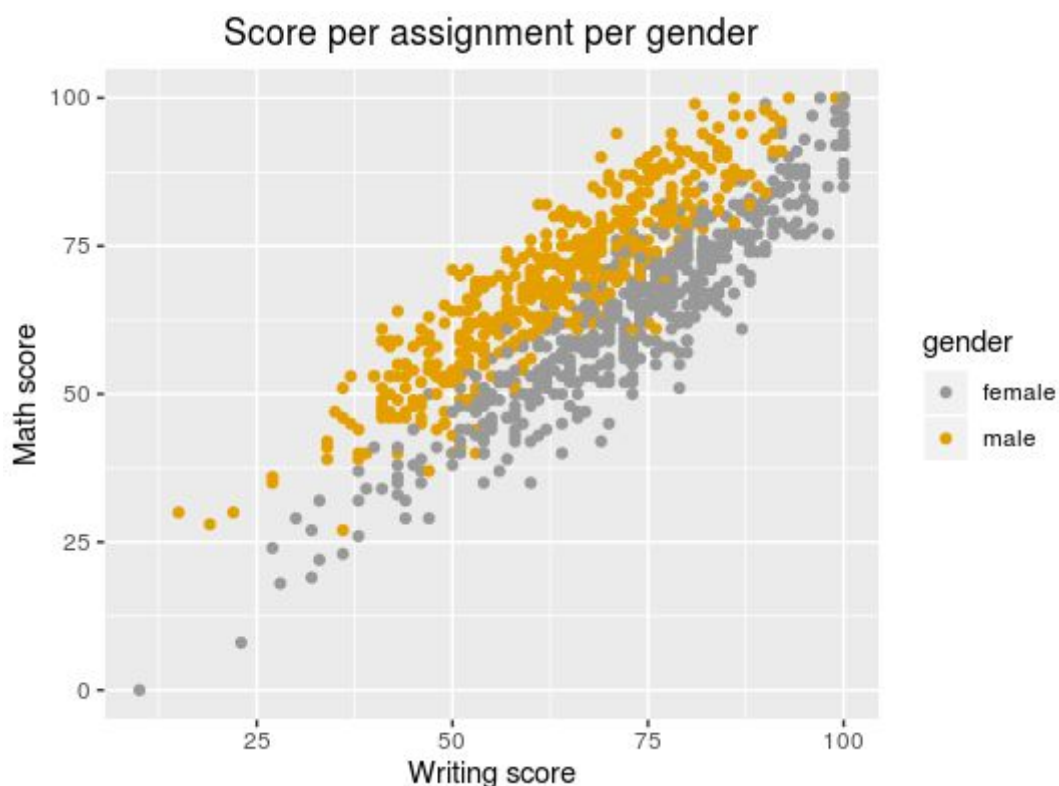


Figure 3 : Notes obtenues en maths et en écriture selon le genre

Conclusion : Nous pouvons conclure à partir de la figure 3 que les hommes sont généralement meilleurs en maths et les femmes en écriture. À noter également que plus de femmes atteignent la note maximale de 100 dans les deux matières.

Etude n°4 : La réussite des étudiants est-elle influencée par le parcours scolaire de leurs parents?

À travers cette quatrième étude, nous avons voulu savoir si le niveau d'études atteint par les parents pouvait influencer significativement les résultats des enfants. Pour cela nous avons extrait les étudiants de chaque catégorie parentale (du lycée jusqu'à un niveau master) et nous avons calculé le ratio de réussite de chaque catégorie.

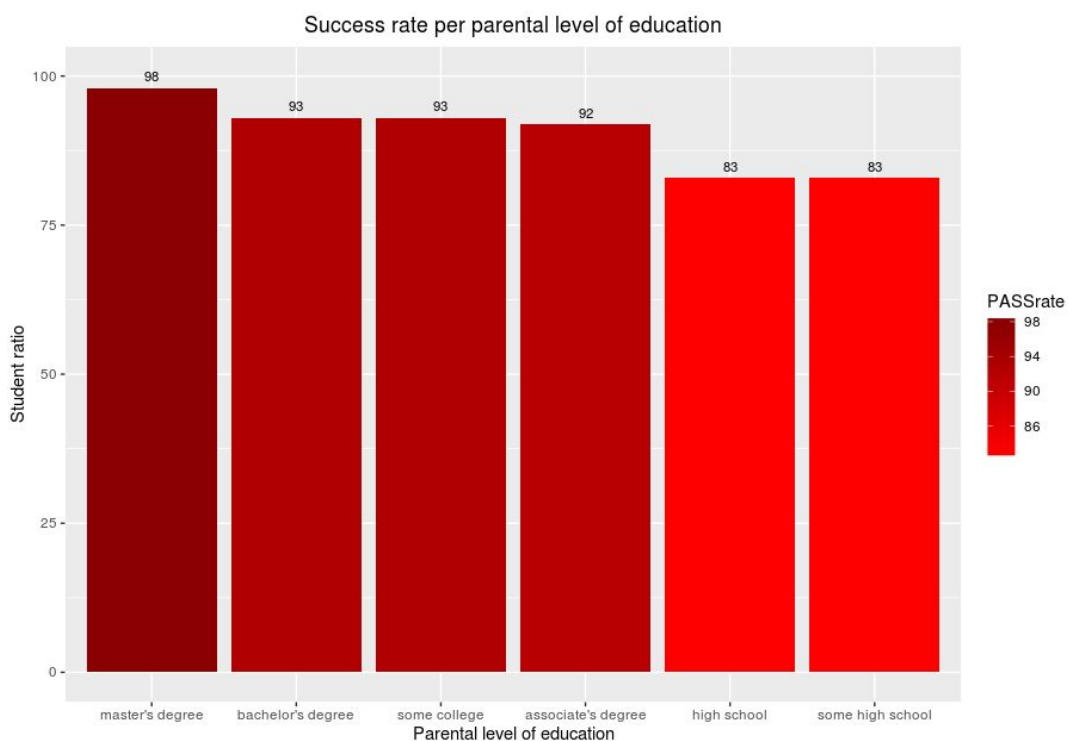


Figure 4 : Résultats des étudiants selon le niveau d'éducation de leurs parents.

Conclusion : Comme nous pouvons le voir sur la figure 4, les résultats des étudiants sont en moyenne proportionnels au niveau d'éducation de leurs parents.

Etude n°5 : Quelle est l'influence de la participation au cours préparatoire sur les résultats?

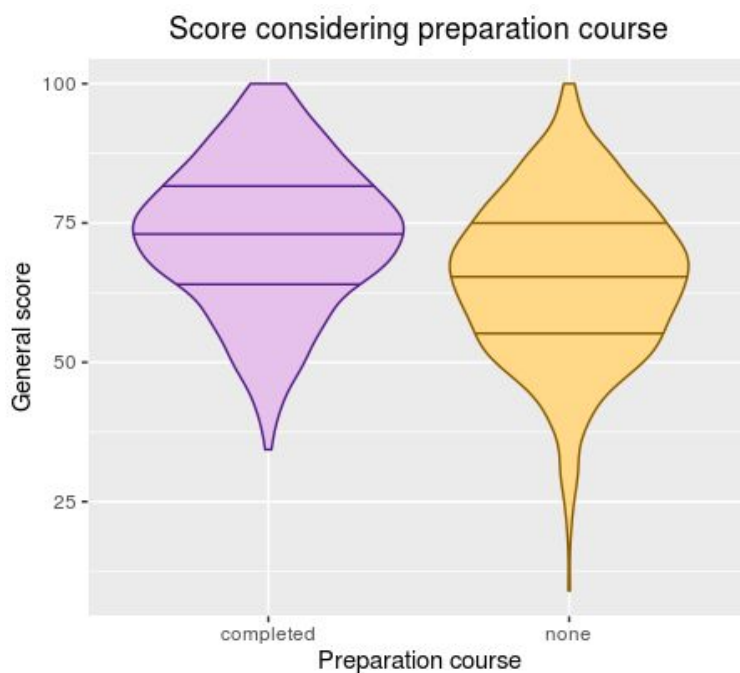


Figure 5 : Influence des cours préparatoires sur les résultats

Conclusion : À partir de la figure 5, nous pouvons remarquer que les notes des personnes ayant assisté à des cours préparatoires ne sont jamais inférieures à 35, contrairement à celles des personnes n'y ayant pas assisté qui atteignent 5. De plus, la valeur du 1er quartile des personnes ayant assisté aux cours préparatoires est quasiment identique à la valeur du 2ème quartile de ceux n'y ayant pas participé. On en conclut donc que participer aux cours préparatoires permettent généralement d'obtenir de meilleurs résultats.

Conclusion

Pour chacune de ces études Il est nécessaire de ramener les résultats obtenus aux échantillons étudiés. Ici notamment, n'ayant pas le même nombre d'hommes et de femmes, il faut en tenir compte dans nos conclusions.

À travers ces 5 études nous avons pu obtenir une esquisse du profil le plus propice à la réussite scolaire. En effet, il semblerait que, pour les données que nous avons, une personne de sexe féminin, qui suit un régime alimentaire standard, dont les parents ont un niveau d'éducation approchant le master et suivant les cours préparatoires serait susceptible d'obtenir les meilleurs résultats. Cette étude, focalisée sur l'écriture, la lecture et les mathématiques n'est pas représentative de l'ensemble des compétences scolaires. Il s'agit cependant d'un premier outil capable de mettre en évidence certains facteurs de performance.