

Dossier Introduction à la Data Science : 1^{ère} étape

Noms et prénoms des membres du trinôme :

- ADROVER Killian
- LAVAUD Lisa
- LAVESNES Tao

Question traitée dans le dossier :

Problématique : *Quels sont les principaux déterminants des performances en mathématiques des élèves ?*

Modèle à estimer :

- Variable à expliquer :

math_score : score obtenu par l'élève en mathématiques (variable quantitative continue).

- Principales variables explicatives :

- Variables quantitatives :

reading_score : score en lecture de l'étudiant.

writing_score : score en écriture de l'étudiant.

- Variables qualitatives :

parental_level_of_education : Niveau d'étude des parents :

lunch : Type de repas de l'étudiant

test_preparation_course : L'étudiant a t-il suivi une préparation supplémentaire ?

Modèle utilisé : $\text{math_score}_i = \beta_0 + \beta_1 \cdot \text{reading_score}_i + \beta_2 \cdot \text{writing_score}_i$

+ $\beta_3 \cdot \text{parental_education}_i + \beta_4 \cdot \text{lunch}_i$

+ $\beta_5 \cdot \text{test_prep}_i + \varepsilon_i$

- Nombre d'observations dans la base de données : la base contient 1000 observations.

- Lien vers la base de données :

<https://www.kaggle.com/datasets/spscientist/students-performance-in-exams?resource=download>

Saisie d'écran de la base de données (10 premières lignes) :

(également trouvable dans le GitHub :
[data-science-student-performance/outputs/tables/first_10_rows.csv](https://github.com/justmarkham/data-science-student-performance/blob/main/outputs/tables/first_10_rows.csv))

| gender | race/ ethnicity | parental level of education | lunch | test preparation course | math score | reading score | writing score |
|--------|--------------------|-----------------------------------|------------------|-------------------------------|---------------|------------------|------------------|
| female | group B | bachelor's degree | standard | none | 72 | 72 | 74 |
| female | group C | some college | standard | completed | 69 | 90 | 88 |
| female | group B | master's degree | standard | none | 90 | 95 | 93 |
| male | group A | associate's degree | free/ reduced | none | 47 | 57 | 44 |
| male | group C | some college | standard | none | 76 | 78 | 75 |
| female | group B | associate's degree | standard | none | 71 | 83 | 78 |
| female | group B | some college | standard | completed | 88 | 95 | 92 |
| male | group B | some college | free/ reduced | none | 40 | 43 | 39 |
| male | group D | high school | free/ reduced | completed | 64 | 64 | 67 |
| female | group B | high school | free/ reduced | none | 38 | 60 | 50 |