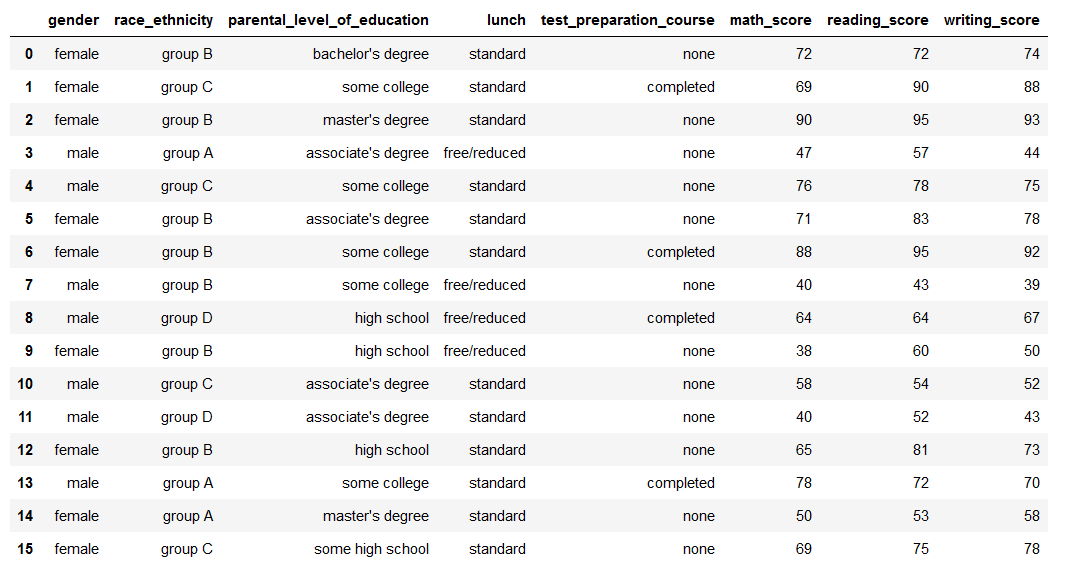
**TP FINAL**

Cours**: Système d’Information Décisionnel**

**1- Lien de téléchargement du dataset : https://www.kaggle.com/**

Bref aperçu du dataset



**2- Explication des données récupérées**

Les données récupérées montrent les performances des élèves (résultats aux tests) par des variables telles que le sexe, l'origine ethnique, le niveau d'éducation des parents, le repas de midi et le cours de préparation aux tests.

Cet ensemble de données comprend les notes obtenues par les élèves dans différentes matières.

- Description des données

Gender : sexe des étudiants -> (homme/femme)

Race/ethnicité : ethnicité des étudiants -> (Groupe A, B, C, D, E)

Niveau d'éducation des parents : niveau d'éducation final des parents -> (bachelor's degree, some college, master's degree, associate's degree, - high school)

Déjeuner : déjeuner avant le test (standard ou gratuit/réduit)

Cours de préparation à l'examen : suivi ou non avant l'examen

Notes des élèves par matière : Score en maths - Score en lecture - Score en écriture

**3- Quelques pistes d’analyse de la donnée**

- Analyser l’impact du niveau d’éducation des parents (parental\_level\_of\_education) par les résultats obtenus (Score). Il s’agit de voir en quoi le niveau d’éducation des parents peut impacter la performance des élèves.

- Croiser les variables « test\_preparation\_course (Cours de préparation à l'examen) et résultats obtenus ». L’objectif est de voir l’impact des cours de préparation ou non avant l’examen sur la performance des élèves afin de recommander cela aux parents si il le faut.

- Faire une analyse de la performance des élèves en fonction du type de déjeuner

-Analyser la performance des élèves suivant le sexe. L’objectif est de décider du type de préparation à faire à chaque groupe d’élève

**4- Mise en place d’un ETL en python qui permettre de charger les données dans une base de données MySQL**

- Code python commenté : Voir fichier main\_study.py sur github

- Aperçu de la base de données MySQL :

