**À propos du jeu de données**

Le jeu de données **Education and Career Success** explore comment les réalisations académiques des étudiants, leurs compétences et leurs activités extrascolaires se traduisent en résultats professionnels concrets.  
En combinant des indicateurs éducatifs (GPA, scores de tests, Field\_of\_Study) avec des mesures d’expérience pratique (Internships, Projects, Certifications) et des indicateurs de soft skills (Networking, Communication), ce jeu de données offre une vue d’ensemble des facteurs qui favorisent la réussite professionnelle en début de carrière.

**Principales caractéristiques**

* **Student Demographics**
  + **Student\_ID** : Identifiant unique
  + Age, Gender
* **Academic Performance**
  + **High\_School\_GPA** (2.0–4.0)
  + **SAT\_Score** (900–1600)
  + **University\_GPA** (2.0–4.0)
  + **Field\_of\_Study** (ex. : Computer Science, Medicine)
* **Skills & Extracurriculars**
  + **Internships\_Completed** (0–4)
  + **Projects\_Completed** (0–9)
  + **Certifications** (0–5)
  + **Soft\_Skills\_Score** (1–10)
  + **Networking\_Score** (1–10)
* **Career Outcomes**
  + **Job\_Offers** (0–5)
  + **Starting\_Salary** (USD 25,000–150,000)
  + **Career\_Satisfaction** (1–10)
  + **Years\_to\_Promotion** (1–5 ans)
  + **Current\_Job\_Level** (Entry, Mid, Senior, Executive)
  + **Work\_Life\_Balance** (1–10)
  + **Entrepreneurship** (Yes/No)

**Applications potentielles**

* **Modélisation prédictive**

Créer des classificateurs ou des régressions pour prédire le nombre d’offres d’emploi, le salaire, le délai de promotion ou la satisfaction professionnelle.

* **Analyse d’importance des variables**

Identifier quels facteurs académiques et expérientiels influencent le plus la réussite professionnelle grâce à des techniques d’interprétabilité de modèles (ex. : SHAP).

* **Segmentation et profilage**

Regrouper les étudiants en profils distincts selon leur préparation à la carrière et adapter les conseils ou interventions.

* **Perspectives éducatives**

Aider les conseillers en orientation et les institutions à cibler les domaines d’amélioration — comme le networking ou l’expérience projet — pour accroître l’employabilité des étudiants.

* **Tableaux de bord interactifs**

Développer des outils visuels (ex. : avec Streamlit ou Plotly) permettant d’explorer l’impact des variables démographiques, académiques et de compétences sur les résultats professionnels.

**À propos de ce fichier**

Ce fichier CSV contient **400 enregistrements anonymisés** d’étudiants, reliant les performances académiques, les compétences et les réalisations extrascolaires aux résultats professionnels de début de carrière. Il peut être utilisé pour explorer, modéliser et visualiser comment l’éducation et l’expérience pratique influencent les offres d’emploi, les salaires, les promotions et la satisfaction.

* **Student\_ID**  
  Code alphanumérique unique attribué à chaque étudiant pour assurer la confidentialité tout en permettant le lien entre enregistrements.
* **Age**  
  Âge de l’étudiant, en années entières, au moment de l’obtention du diplôme ou de la collecte des données (plage : 18–30).
* **Gender**  
  Genre auto-déclaré : “Male”, “Female” ou “Other”.
* **High\_School\_GPA**  
  Moyenne générale cumulative au lycée, sur une échelle de 2.0 à 4.0.
* **SAT\_Score**  
  Score total à l’examen d’entrée à l’université SAT, compris entre 900 et 1600.
* **University\_GPA**  
  Moyenne générale cumulative en licence, sur une échelle de 2.0 à 4.0.
* **Field\_of\_Study**  
  Discipline ou spécialité universitaire (ex. : “Computer Science”, “Medicine”, “Business”).
* **Internships\_Completed**  
  Nombre de stages formels effectués pendant les études universitaires (0–4).
* **Projects\_Completed**  
  Nombre de projets académiques ou personnels significatifs (ex. : projets de fin d’études, recherche, applications) réalisés (0–9).
* **Certifications**  
  Nombre de certifications reconnues par l’industrie obtenues en dehors des cours formels (ex. : AWS, PMP) (0–5).
* **Soft\_Skills\_Score**  
  Évaluation par un enseignant ou un pair des compétences en communication, travail en équipe et leadership, sur une échelle de 1 à 10.
* **Networking\_Score**  
  Évaluation quantitative de l’activité de réseautage professionnel (ex. : connexions LinkedIn, événements suivis) sur une échelle de 1 à 10.
* **Job\_Offers**  
  Nombre total d’offres d’emploi à temps plein reçues dans les six mois suivant l’obtention du diplôme (0–5).
* **Starting\_Salary**  
  Salaire annuel de base en USD pour le premier poste à temps plein (plage : 25,000–150,000).
* **Career\_Satisfaction**  
  Satisfaction auto-déclarée vis-à-vis du premier emploi, sur une échelle de 1 (très insatisfait) à 10 (très satisfait).
* **Years\_to\_Promotion**  
  Nombre d’années écoulées entre l’embauche et la promotion au niveau de carrière suivant (1–5).
* **Current\_Job\_Level**  
  Stade de carrière au moment de la collecte des données : “Entry”, “Mid”, “Senior” ou “Executive”.
* **Work\_Life\_Balance**  
  Évaluation personnelle de l’équilibre entre vie professionnelle et vie personnelle, sur une échelle de 1 (faible) à 10 (excellent).
* **Entrepreneurship**  
  Indique si la personne a fondé ou cofondé une entreprise (“Yes” ou “No”).