

ISEN

ALL IS DIGITAL!

OUEST

TP n°6 - Bases de données - Python

D'un fichier texte à une base de données

benoit.lardeux@isen-ouest.yncrea.fr

Objectifs

- Créer une base de données à partir d'un fichier texte

1 Description du fichier

Nous allons dans ce TP créer une base de données utilisée pour calculer la probabilité d'admission d'étudiants postulant à un programme de Master. La probabilité est déterminée en fonction de caractéristiques en lien avec le parcours universitaire de chacun des candidats.

A partir du fichier texte intitulé *admissions.txt*, l'objectif est de construire et remplir une base de données *Postgresql*. La première ligne du fichier *admissions.txt* comporte les titres des différentes propriétés associées aux caractéristiques des candidats. Celles-ci sont expliquées ci-dessous. Dans ce fichier texte, le séparateur des données est la virgule.

- *Serial No.* : identifiant de l'élève
- *GRE score* (max 340) : note à un test général d'admission en études supérieures
- *TOEFL score* (max 120) : note à un test d'anglais
- *University rating* (max 5) : classement de l'université
- *SOP* (max 5) : impact des lettres de recommandations
- *ML* (max 5) : impact de la lettre de motivation
- *CGPA* (max 10) : moyenne cumulative pondérée en Licence
- *Research* (0= pas d'expérience ou 1= expérience) : expérience de recherche
- *Admission* (0= pas admis ou 1= admis) : admission en master

2 Installation de l'environnement de développement

Installation de *python3*

- Taper l'adresse <https://www.python.org/downloads/>
- Télécharger la dernière version de Python
- Installer Python
- N'oublier pas de cocher la case "Add Python x.x to PATH"

Inscription à *JetBrains*

En tant qu'étudiant ISEN et en utilisant votre adresse email YNCREA ISEN Ouest (prenom.nom@isen-ouest.yncrea.fr), vous pouvez télécharger gratuitement les produits JetBrains comme PyCharm.

Pour ce faire, enregistrer vous sur le site :

<https://www.jetbrains.com/shop/eform/students>

Un mail d'activation de la licence va vous être envoyés.

Installation de *JetBrain Toolbox*

Télécharger puis installer JetBrain Toolbox via le site :
<https://www.jetbrains.com/toolbox/>

Installation de *PyCharm Edu*

Dans Jetbrain Toolbox, chercher PyCharm Edu et cliquer sur "Install"

Premier programme : *Hello World!*

- File>NewProject...
- Créer le projet dans un dossier *tp1* par exemple
- Une fois le projet créé, pour créer un fichier python *tp1.py*, clic droit sur le projet (sous forme de dossier) puis "new" puis "python file". Remarquer que l'extension va être ajoutée automatiquement.
- Taper `print("Hello World!")`
- Clic droit sur le code et choisir "Run 'tp1'". La sortie du programme va être affichée dans l'onglet "4 :Run" en bas de la fenêtre
- Cliquer sur "tp1" en haut à droite à côté de l'icône *play* puis choisir "Save tp1 configuration"

3 Importer la base de données

- a) A partir de l'ENT, téléchargez le fichier intitulé *admissions.txt*
- b) Créez un nouveau projet Python
- c) Développez une fonction permettant de lire le fichier *admissions.txt*
- d) Développez la fonction écrivant le fichier *creation_admissions.sql* qui permet de créer la table "*Profil*" dans PostgreSQL. Prenez soin de définir les types appropriés aux propriétés de la table
- e) Développez la fonction qui écrit le fichier *populer_admissions.sql* permettant de remplir la table "*Profil*" avec les données
- f) Dans PostgreSQL, en tant que super-utilisateur *postgres*, créez un nouveau rôle appelé *admiadmin* qui n'est pas super-utilisateur mais qui peut hériter d'autres rôles et peut se connecter à une base de données en entrant un mot de passe
- g) Créez la base de données *admissions* dont le détenteur (owner) est "*admiadmin*"
- h) Donnez à *admiadmin* tous les privilèges sur toutes les tables de la base de données *admissions*
- i) Créez la table dans la base de données *admission* à partir du fichier *creation_admissions.sql*
- j) Remplissez la table grâce à *populer_admissions.sql*
- k) Connectez vous sur la base de données *admissions* en tant que *admiadmin*

4 Requêtes SQL de manipulation de données

- a) Sélectionner l'ensemble des candidats ayant un *GRE score* supérieur à 320, un *TOEFL score* supérieur à 110 et un *University rating* supérieur à 3
- b) Quel est le taux d'admission pour le sous-groupe sélectionné en a) ?