Rapport Livrable 3

0. Les données - Problématique

1. Présentation des donnés

Le fichier Vue.csv contient plusieurs séries statistiques sur l’ensemble des moyennes des étudiants inscrits à l’IUT.

* 1ère valeur : Le code postal de l’établissement
* 2ème valeur : La différence de la moyenne des étudiants entre le premier et le deuxième semestre
* 3ème valeur : La différence de la moyenne des étudiants entre le premier et le deuxième semestre
* 4ème valeur : La différence de la moyenne des étudiants entre le premier et le deuxième semestre
* 5ème valeur : La valeur de la moyenne des étudiant au 4ème semestre

1. Problématique

En utilisant ces données nous allons essayer de répondre à la problématique suivante :

*L’emplacement géographique des IUT influe-t-il sur l’évolution des moyennes des étudiant au cours de l’année ?*

1. Import des données mise en forme

1. Importer des données en python

On importe notre vue sur python sous forme de DataFrame avec la commande :

VueDF = pd.read\_csv(“Vue.csv”)

1. Mise en forme

Suppression des cases vides dans notre vue et on transforme notre DataFrame en Array :

VueDF = VueDF.dropna()

VueAR = VueDF.to\_numpy()

1. Normalisation

On ne garde que les colonnes de notre tableau qui contiennent des donn´ees num´eriques, on peut alors centrer-réduire ces donn´ees :

2. Représentation graphique

on mettra des screens de diagramme de moyenne par semestre pour le code postal d’un établissement

“observation sur le diagramme”

même chose mais pour un autre établissement

“observation sur le diagramme”