## Café vs sport vs calorie vs productivité

Rendus : script python avec les png générés

Datasets :

* Sport\_raw.csv
* Travail\_raw.csv

Relations :

* calories ≈ MET(activite) \* 3.5 \* poids\_kg/200 \* duree\_min + bruit où MET(course)=9.8, MET(velo)=7.5, MET(natation)=8.0. (MET unité physiologique)
* productivite ≈ 45 + 4.5 \* tasses\_cafe + bruit puis rendement décroissant (aplatissement)

Travail à faire

* Pour une activité à la fois (ex. *vélo*), modéliser calories = a \* duree\_min + b
* Pour une activité à la fois (ex. *vélo*), modéliser la version multivariée avec poids\_kg: calories = a\*duree\_min + b\*poids\_kg + c.
* Faire une régression linéaire, récupérer pente.
* Tracer scatter + droite de régression
* Facultatif : Tracer les résidus vs duree\_min ; commenter (linéarité, hétéroscédasticité, outliers restants).
* Interpréter la pente (calories par minute) et comparer aux ordres de grandeur des METs.
* Modéliser productivite = a \* tasses\_cafe + b sur l’intervalle 0–6 tasses.
* Tracer scatter + droite de régression
* Que signifie la pente ? À partir de quel niveau de café l’effet décroît ?
* Répondre aux questions suivantes
  + Les jours où les individus font du sport intense, boivent-ils plus de café le lendemain ?
  + Les individus actifs sont-ils plus productifs ?
  + Trop de café et trop de sport → baisse de productivité ?
  + Y a-t-il un équilibre entre “heures\_travail + duree\_sport” et la productivité ?
  + Y a-t-il une corrélation entre calories et tasses\_cafe ?
  + Les individus les plus sportifs consomment-ils en moyenne plus de café ?
* Facultatif : Vérifier si le sport du jour précédent a un effet sur la productivité du lendemain
* Facultatif : transformer votre application console en api
* Facultatif : docker-compose pour lancer votre projet

Si vous voulez travaillez avec vos données : 2 csv avec un lien entre les deux, une formule à modéliser, scatter +régression linaire + pente, une liste à répondre

Si vos fichiers sont propres : enlever des données pour les rendre réels

**Me rendre également le sujet : thème de vos données, ce que vous calculez et à quelles questions vous répondez avec les fichiers csv des données**