

Bootcamp IA — Atelier Guidé

Département IA & Ingénierie des Données
Institut Supérieur d'Informatique (ISI Dakar)

Samedi 20 Septembre 2025
14h15 – 16h00

Objectif de l'atelier

Cet atelier vise à mettre en pratique les notions vues en matinée. Les participants apprendront à :

- Construire un modèle simple de **régression linéaire**.
- Visualiser la relation entre une variable explicative et une variable cible.
- Interpréter les résultats de manière claire et concise.

Contexte

Nous disposons d'un dataset fictif contenant des informations sur 100 étudiants :

- **Âge** (en années),
- **Heures de révision** avant l'évaluation,
- **Score final** obtenu au test d'IA.

Mission : prédire le score final en fonction des heures de révision.

Organisation des participants

- Les 70 participants sont répartis en **14 groupes de 5**.
- Chaque groupe dispose d'un poste avec Python (Anaconda ou Google Colab).
- Temps imparti : **1h15**.
- Chaque groupe rend : un code + un graphe + une interprétation.

Étapes à suivre

1. Charger le dataset `etudiants.csv`.
2. Visualiser la relation entre `heures_revision` et `score` (scatter plot).
3. Créer et entraîner un modèle de **régression linéaire**.
4. Évaluer la qualité du modèle (R^2).
5. Prédire le score d'un étudiant hypothétique ayant 5h de révision.

Code de départ

```
1 import pandas as pd
2 import matplotlib.pyplot as plt
3 from sklearn.linear_model import LinearRegression
4
5 # 1. Charger les données
6 df = pd.read_csv("etudiants.csv")
7
8 # 2. Variables explicatives et cible
9 X = df[["heures_revision"]] # entrée
10 y = df["score"]             # sortie
11
12 # 3. Visualisation
13 plt.scatter(X, y)
14 plt.xlabel("Heures de révision")
15 plt.ylabel("Score final")
16 plt.title("Relation heures de révision vs score")
17 plt.show()
```

Ajustement du modèle

```
1 # 4. Création et entraînement du modèle
2 model = LinearRegression()
3 model.fit(X, y)
4
5 # 5. Résultats
6 print("Coefficient :", model.coef_)
7 print("Intercept   :", model.intercept_)
8 print("Score R2     :", model.score(X, y))
9
10 # 6. Prédiction
11 print("Score prévu pour 5h :", model.predict([[5]]))
```

Consignes pour les groupes

Chaque groupe doit :

- Fournir le code Python de la régression linéaire.
- Générer un graphe **scatter** + **droite de régression**.
- Rédiger une phrase d'interprétation, par exemple : *“Chaque heure de révision supplémentaire augmente le score de 3 points en moyenne.”*

Restitution

- Chaque groupe présentera rapidement son graphe et son interprétation.
- Les résultats seront comparés en plénière.
- Une synthèse sera donnée par l'enseignant avec mise en avant des bonnes pratiques.

À retenir

[regular

Points clés]

- La régression linéaire est une première étape vers le Machine Learning prédictif.
- Visualisation + modélisation = compréhension rapide des données.
- Le travail en groupe favorise l'apprentissage collaboratif.