



Departement  
Informatique  
&  
Méthodes  
Quantitatives

## **Mini Projet Aide à la décision Avec Python**

AU : 2023/2024  
Fénière : 2<sup>ème</sup> MP-IEF  
Responsable :  
Dr.Fadoua BOUAFIF

### ➤ **Contexte du projet**

Vous travaillez pour une entreprise de conseil financier qui souhaite optimiser les investissements pour ses clients. Votre mission implique l'analyse de données financières, la détection de tendances et la formulation de recommandations en matière d'investissement.

### ➤ **Objectif du projet**

Le projet vise à utiliser les connaissances acquises en programmation Python, notamment la manipulation de données avec Pandas, la visualisation graphique avec Matplotlib et Seaborn, et les concepts d'aide à la décision pour analyser les données financières, prendre des décisions basées sur ces analyses, et présenter les résultats de manière visuelle.

### ➤ **Organisation**

1. Groupe de deux étudiants maximum (ceux (ou celles) qui souhaitent travailler seul(e)s peuvent le faire).
2. Tous les étudiants doivent créer un compte sur **GitHub** (<https://github.com/>)
3. Vos codes, données et rapports doivent être chargés sur **GitHub** puis vous allez me fournir le lien d'accès à vos projets sur Google Classroom.  
(<https://classroom.google.com/c/NTQxNzg1NTgwODE2?cjc=vxiuyz4>)
4. Rapport en Google doc, Jupyter-Nootbook ou LaTeX.
5. Date de rendu : **05 Janvier 2024 à 23h59mn**
6. Date de présentation : **12 et 13 Janvier 2024**

### ➤ **Travail demandé**

#### **A. Phase1 : Pratique**

##### **1. Manipulation de données :**

- a. Charger un ensemble de données financières pertinentes à partir d'une source externe (Yahoo Finance, Quandl, etc.).
- b. Effectuer des opérations de base sur les données en utilisant Pandas (filtrage, tri, groupby, etc.).

## 2. Analyse et Visualisation :

- a. Calculer des statistiques descriptives sur les données avec Numpty (moyennes, médianes, écarts types, etc.)
- b. Créer des graphiques pertinents à l'aide de Matplotlib et Seaborn pour illustrer les tendances du marché, les performances d'actifs, etc.

## 3. Prise de Décision :

- a. Mettre en œuvre des algorithmes simples d'aide à la décision (par exemple, recommandation d'investissement basée sur des critères prédéfinis).
- b. Justifier les décisions prises en se basant sur l'analyse des données.

## B. Phase2 : Théorique

Rédiger un rapport qui inclure une :

1. Présentation descriptive des données utilisées
2. Présentation des résultats de l'analyse et des recommandations
3. Interprétation des visualisations graphiques

## ➤ Ressources de données

### 1. **Données Financières :**

- Yahoo Finance (utilisation de la bibliothèque **yfinance** pour extraire les données). (<https://finance.yahoo.com/quote/CSV/history/>)

### 2. **Données Économiques :**

- Banque Mondiale (utilisation de la bibliothèque **wbdata**).([https://datacatalog.worldbank.org/search?q=&sort=last\\_updated\\_date%20des c](https://datacatalog.worldbank.org/search?q=&sort=last_updated_date%20des c))
- Fonds Monétaire International (FMI) pour les données économiques mondiales. (<https://data.imf.org/?sk=388dfa60-1d26-4ade-b505-a05a558d9a42>)

### 2. **Données du Marché Boursier :**

- **Kaggle** propose plusieurs ensembles de données liés aux marchés financiers. (<https://www.kaggle.com/datasets/qks1lver/financial-data-of-4400-public-companies>)

### 3. **Autres :**

- Data Word : <https://data.world/datasets/finance>

## ➤ **Evaluation**

L'évaluation porte sur :

- La présentation : Qualité, clarté, organisation, animation, etc.
- Le rapport : clarté, organisation, etc.
- Le programme : exécution, commentaires, etc.

**N.B :**

**1. Le contenu produit doit être original, et la duplication de travaux équivalents sera sujet à des sanctions.**

**2. Remplir le fichier Goole sheet avant le 15 décembre 2023 Lien :**

[https://docs.google.com/spreadsheets/d/1VMe\\_vL0TL6g93bgiiyZxmyCo9tcf0iSPtOVeoibRJwU/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1VMe_vL0TL6g93bgiiyZxmyCo9tcf0iSPtOVeoibRJwU/edit?usp=sharing)

**Bon Courage**

○

Les différents algorithmes de machine learning :

1. Linear regression
2. Logistic Regression
3. Support Vector Machine (SVM)
4. Naïve Bayes
5. Anomaly Detection
6. Decision Trees
7. Neural Networks
8. K-Means
9. Gradient Descent
10.     Algorithme KNN (K- Nearest Neighbors)
11.     Algorithme Random Forest
- 12.

Les algorithmes de classification peuvent être divisés en deux catégories principales :

Modèles linéaires :

- Régression logistique
- Support Vector Machines – SVM

Modèles non linéaires :

- K-Nearest Neighbors (K plus proche voisins)
- Kernel SVM
- Naïve Bayes (réseaux bayésiens)
- Classification par arbre de décision
- Classification Random Forest

Lien pour dataset

➤ **UCI Machine learning Repository:** <https://archive.ics.uci.edu/ml/index.php>

• Google Scholar • UCI Machine Learning Repository • Kaggle • Sci-Hub • Google Colab