



École Nationale Supérieure d'Informatique pour l'Industrie et l'Entreprise

Rapport de stage 1A

« Indicateurs pour l'analyse technique de l'évolution du
prix d'un actif »

Élève : Lina El haddaj

Tuteur académique à l'ENSIIE : Sergio Pulido Nino

Diplôme préparé : Ingénieur ENSIIE (promotion 2025)

le cnam

Entreprise : Le Cnam (Paris, 75)

Maître de stage : Dariush Ghorbanzadeh

Stage de première année du lundi 12 juin 2023 au vendredi 4 août 2023

Remerciements

Je tiens à exprimer ma profonde gratitude envers M. Ghorbanzadeh pour m'avoir accueillie au sein du département Mathématiques et Statistique pendant la période de mon stage au CNAM. Ce stage a été une expérience enrichissante et formatrice qui m'a permis de découvrir de nouveaux horizons et d'approfondir mes connaissances dans le domaine du développement et de l'analyse des données.

M. Ghorbanzadeh a joué un rôle essentiel dans mon apprentissage en me donnant l'opportunité de collaborer étroitement avec lui. Sa patience, son expertise et son dévouement envers l'enseignement m'ont inspiré et m'ont aidé à progresser tout au long de ce stage. J'ai eu la chance de travailler directement avec lui, d'observer son travail pédagogique auprès de ses élèves, et d'apprendre de ses compétences.

Au-delà de l'aspect professionnel de ce stage, M. Ghorbanzadeh m'a également permis de découvrir ma passion pour des sujets tels que l'étude des données. Ses encouragements et son soutien ont été déterminants dans ma prise de conscience de l'importance de ce domaine et de son potentiel pour l'avenir de ma carrière.

Ce stage a été une expérience inoubliable, et j'en suis reconnaissante. Il a renforcé ma passion pour l'apprentissage et l'analyse des données, et je suis impatiente d'appliquer les connaissances et les compétences acquises au cours de ce stage dans ma future carrière.

Merci encore pour cette opportunité exceptionnelle.

Le Conservatoire National des Arts et Métiers

Le Conservatoire National des Arts et Métiers (CNAM) est un grand établissement d'enseignement supérieur et de recherche français placé sous la tutelle du ministère chargé de l'Enseignement Supérieur. Fondé par l'abbé Henri Grégoire à Paris dans le but de « perfectionner l'industrie nationale », il est avec l'École Polytechnique et l'École Normale Supérieure l'une des trois créations de la Révolution française ayant pour but de promouvoir les sciences et les techniques.

The logo consists of the text 'le cnam' in a red, lowercase, sans-serif typeface. The 'l' is tall and thin, while the 'e' is shorter and wider. The 'c' is a simple curve, and 'nam' is in a bold, slightly condensed font.

FIGURE – *Le logo du CNAM*

Liste des sigles et abréviations

Voici les sigles et abréviations utilisés dans ce mémoire (ou rapport, au choix) :

- **CNAM** : Conservatoire National des Arts et Métiers ;
- **MFI** : *Money Flow Index* en anglais, ou "Indice de Flux Monétaire" en français ;
- **RSI** : *Relative Strenght Index* en anglais, ou "Indice de Force Relative" en français ;
- **MACD** : *Moving Average Convergence Divergence* en anglais, ou "Convergence/-Divergence de Moyennes Mobiles" en français ;
- **BA** : symbole financier correspoondant à "The Boeing Company" ;
- **RSE** : Responsabilité sociétale des entreprises.

Sommaire

1	Introduction	1
2	Implémentation des Formulaires de Simulation et de la Communication PHP-Python	2
2.1	Création de Formulaires de Simulation Interactifs	2
2.2	Communication PHP-Python avec argparse	3
3	Simulation et traitement de données boursières	5
4	Analyse des données boursières	8
4.1	Statistiques descriptives	8
4.2	Histogramme des données	9
4.3	Indicateurs pour l'analyse technique de l'évolution du prix d'un actif	12
4.4	Traitement des données	13
5	Conclusion	22
A	Annexe : Développement durable et responsabilité sociétale	23
A.1	Engagement envers le développement durable	23
A.2	Résultats et Mesures d'Impact	23
A.3	Responsabilité sociétale	24
A.4	Environnement de travail	25
B	Code développé	26
B.1	Fichier Librairie_fonctions.py	26
B.2	back.py	28

1 Introduction

L'objectif principal de ce stage consiste à acquérir les compétences fondamentales de la Data Science en vue de leur application sur des ensembles de données provenant du domaine financier. Plus précisément, notre mission consistera à créer des modèles pour comprendre l'évolution de différents indicateurs financiers d'une entreprise donnée sur une période spécifique, en utilisant Python pour le calcul et l'analyse des résultats.

Pour atteindre cet objectif, nous débiterons par la création d'une infrastructure web de base en utilisant HTML et PHP. Cette infrastructure comprendra des pages destinées à l'affichage des résultats, ainsi qu'une page dédiée à la visualisation de l'évolution des indicateurs financiers pour l'entreprise en question. Lorsque nous parlons de "résultats", nous faisons référence aux informations clés extraites de l'ensemble de données, telles que les valeurs maximales, minimales, moyennes, et d'autres statistiques pertinentes, qui seront calculées à l'aide de Python pour analyser l'évolution de ces indicateurs au fil du temps.

Par la suite, nous représenterons la distribution des valeurs des différents indicateurs financiers à l'aide d'histogrammes. De plus, nous générerons des graphiques basés sur les résultats de l'exécution des algorithmes spécialement développés dans le cadre de ce stage, toujours en utilisant Python, afin de visualiser de manière plus détaillée l'évolution de ces indicateurs financiers pour l'entreprise sur la période donnée.

2 Implémentation des Formulaires de Simulation et de la Communication PHP-Python

2.1 Création de Formulaires de Simulation Interactifs

M. Ghorbanzadeh nous a mis à disposition une page d'accueil web complète, comprenant un menu latéral offrant diverses options de lois de probabilité. Lorsqu'un utilisateur sélectionne l'une de ces options, un formulaire de simulation correspondant s'affiche à l'écran. Cette approche a rendu la simulation de différentes lois de probabilité accessible à un large public, facilitant ainsi l'analyse statistique pour des données diverses.

Prenons l'exemple du formulaire lié à la loi normale, bien que le processus soit similaire pour d'autres lois continues. Le formulaire standard que nous avons développé est illustré ci-dessous :

Simulation de la loi normale	
μ	<input type="text" value="5"/>
σ	<input type="text" value="4"/>
taille	<input type="text" value="537"/>
nombre de classes	<input type="text" value="11"/>
fichier sortie	<input type="text" value="Normale_H"/>

Exécuter python

FIGURE 1 – Formulaire

Ce formulaire offre la possibilité de spécifier des paramètres tels que la moyenne μ et l'écart-type σ de la loi normale à simuler, ainsi que la taille de l'échantillon et le nombre de classes modales pour l'histogramme. Le champ "Nom de la figure" permet à l'utilisateur de nommer la sortie générée, afin d'afficher l'histogramme suivant :

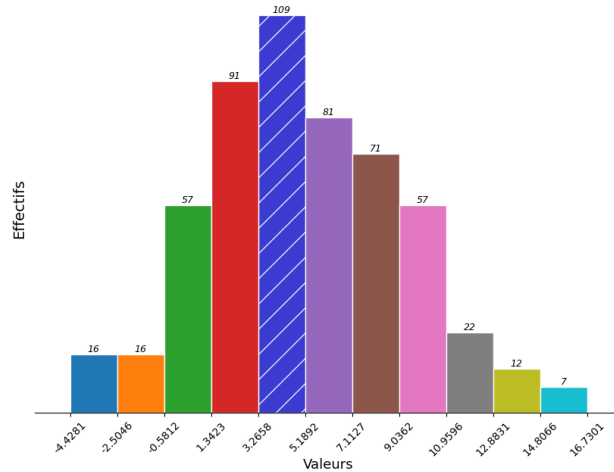


FIGURE 2 – Histogramme renvoyé par le formulaire

2.2 Communication PHP-Python avec argparse

Une fois que l'utilisateur a complété et soumis le formulaire, il est essentiel de récupérer les données saisies pour les utiliser dans la simulation. C'est ici que la communication entre le langage de programmation PHP et Python entre en jeu. Au lieu de transmettre les données via l'URL (méthode "get"), nous utilisons la méthode "post" pour garantir la sécurité des informations.

La première étape consiste à récupérer les paramètres de la distribution et les détails de l'échantillon à partir de la méthode "post" en PHP :

```

1 $pmu = htmlspecialchars($_POST['mu']);
2 $psigma = htmlspecialchars($_POST['sigma']);
3 $nbObs = htmlspecialchars($_POST['taille']);
4 $nbc = htmlspecialchars($_POST['nbclasses']);
5 $fichier_figure = htmlspecialchars($_POST['NomFigure']);

```

Ensuite, nous créons une commande PHP pour exécuter un script Python spécifique en utilisant les arguments récupérés du formulaire. La variable *\$output* recueille les résultats de l'exécution :

```

1 exec("python normale_simulation.py $pmu $psigma $nbObs $nbc
2 $fichier_figure", $output);

```

C'est ici que le module Python argparse entre en jeu. Pour le script Python, nous définissons les arguments que la fonction doit recevoir lors de son exécution :

```

1 parser = argparse.ArgumentParser(description="Ecriture dans un

```

```

2 fichier_externe_en_format_txt")
3 parser.add_argument('mu', type=float, help='parametre_mu\n\'')
4 parser.add_argument('sigma', type=float, help='sigma\n\'p\'')
5 parser.add_argument('taille', type=int, help='taille_de
6 l\'echantillon \n\'')
7 parser.add_argument('k', type=int, help='nombre_de_classes\n\'')
8 parser.add_argument('Nomfigure', type=str, help="fichier_figure")
9 args = parser.parse_args()

```

Une fois que les arguments sont définis, la simulation est effectuée et les résultats sont renvoyés à PHP pour affichage sur la page web.

En utilisant cette méthode, nous avons pu développer des formulaires interactifs pour différentes lois de probabilité et établir une communication fluide entre PHP et Python grâce à l'utilisation d'argparse pour le traitement des arguments. Cette approche nous permet d'étendre facilement ces fonctionnalités à d'autres lois de probabilité en créant des formulaires similaires et en adaptant les scripts Python correspondants.

En résumé, cette section a décrit le processus de création de formulaires interactifs pour la simulation de lois de probabilité, ainsi que la communication fluide entre PHP et Python grâce à argparse pour le traitement des arguments. Cette solution nous a permis de fournir un outil complet et flexible pour l'analyse statistique des données, en facilitant l'ajout de nouvelles distributions au fur et à mesure de nos besoins.

3 Simulation et traitement de données boursières

M. Ghorbanzadeh m'a envoyé une série de documents et vidéos au sujet d'argparse, un module qui permet de faire la passerelle entre Python et PHP comme illustré dans la partie précédente.

L'objectif principal de cette passerelle est d'afficher les données provenant de Yahoo Finance sur la page web, après les avoir traitées.

Au départ, il était prévu d'utiliser argparse pour gérer les arguments passés à un script Python depuis PHP.

Cependant, après une analyse plus approfondie, nous avons finalement opté pour une approche différente, en utilisant la liste `sys.argv`.

Nous allons expliquer les spécificités de ces deux méthodes dans la suite pour mieux comprendre pourquoi ce choix a été fait.

Introduction à argparse

argparse est une bibliothèque Python qui facilite la gestion des arguments passés à un script en ligne de commande. Lorsque des scripts qui doivent être exécutés en ligne de commande sont créés, argparse permet de spécifier comment les utilisateurs doivent fournir des informations à ce script. Il automatise la conversion d'arguments en valeurs Python, gère les erreurs d'entrée utilisateur et fournit une aide conviviale.

Introduction à `sys.argv`

`sys.argv` représente un élément essentiel de Python pour la gestion des arguments passés à un script via la ligne de commande.

Lorsque vous concevez des scripts destinés à une interaction en ligne de commande, `sys.argv` devient un outil incontournable. Il s'agit d'une liste de chaînes de caractères contenant au minimum le nom du script lui-même, suivi des arguments fournis par l'utilisateur au moment de l'exécution du script.

Contrairement à argparse, qui automatise la gestion structurée des arguments, `sys.argv` nécessite une manipulation plus manuelle. Cependant, il offre une flexibilité totale pour la manière dont vous traitez ces arguments, ce qui peut s'avérer particulièrement avantageux lorsque des besoins spécifiques de traitement des données passées via la ligne de commande se présentent.

En utilisant `sys.argv`, vous conservez un contrôle direct sur la récupération, la conversion et la manipulation des arguments de votre script, ce qui en fait un choix approprié lorsque la personnalisation approfondie est requise.

Intégration des Scripts Python et PHP avec sys.argv

Dans le cadre de mon projet de stage, j'ai été confronté à la nécessité de créer une solution où des scripts Python (.py) et des scripts PHP (.php) interagissent avec une page web via un formulaire.

L'un des défis majeurs était de trouver une manière efficace de communiquer des données entre ces deux langages de programmation tout en préservant la performance de l'ensemble du système.

Avantages de sys.argv dans ce Contexte

L'une des solutions que j'ai adoptées était d'utiliser sys.argv pour gérer les arguments passés aux scripts Python. Cette approche s'est avérée avantageuse pour plusieurs raisons spécifiques à notre contexte :

- **Intégration Simplifiée :** Le fait que les données provenant de la page web soient généralement envoyées sous forme de paramètres POST ou GET au serveur web simplifie l'intégration. Les scripts PHP peuvent récupérer ces données et les transmettre aux scripts Python en tant qu'arguments via la ligne de commande. sys.argv offre une manière simple et directe de récupérer ces arguments dans les scripts Python.
- **Performance Améliorée :** Une des contraintes majeures de notre projet était la performance. Le traitement des données depuis la page web et leur passage au script Python étaient des opérations intensives en termes de temps. Dans ce contexte, sys.argv offrait l'avantage de réduire la surcharge liée à l'analyse des arguments, contribuant ainsi à accélérer l'exécution globale du système.
- **Simplicité de Configuration :** La configuration de sys.argv pour récupérer des arguments de ligne de commande est simple et directe. Cette simplicité est particulièrement utile lorsque vous avez déjà un flux de données pré-défini provenant des scripts PHP et devant être acheminé vers les scripts Python via la ligne de commande.

Limitations de sys.argv

Il est important de noter que sys.argv présente également certaines limitations :

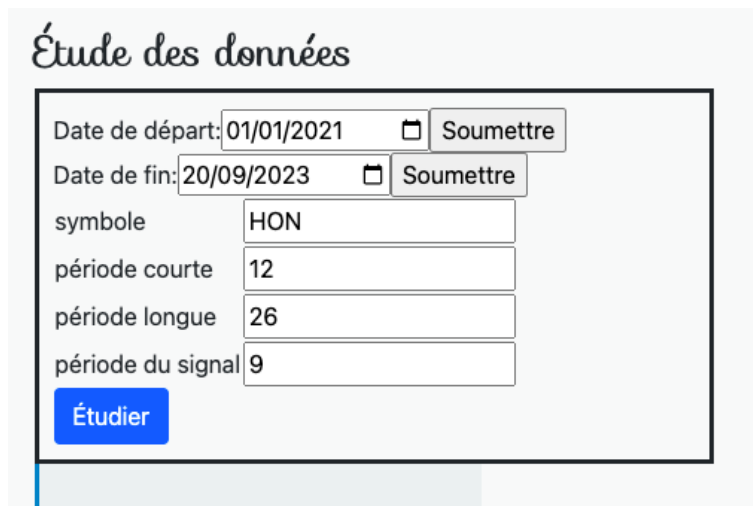
- **Gestion de la Conversion de Type :** Avec sys.argv, vous devez gérer la conversion de type des arguments vous-même. Cela signifie que si vous attendez des entiers, des flottants ou d'autres types spécifiques, vous devrez effectuer ces conversions manuellement.
- **Validation Personnalisée :** Vous devez mettre en place des mécanismes de validation personnalisée pour vos arguments. Cette tâche peut devenir complexe si

vous avez de nombreux arguments avec des contraintes spécifiques.

- **Documentation Séparée :** `sys.argv` ne fournit pas de documentation automatique pour les utilisateurs. Vous devrez créer une documentation distincte pour expliquer comment utiliser les arguments de vos scripts.

En conclusion, l'utilisation de `sys.argv` s'est avérée être une solution appropriée pour notre projet, où nous devons intégrer des scripts Python et PHP tout en préservant la performance globale du système. Cependant, il est essentiel d'évaluer cette approche en fonction des besoins spécifiques de votre projet, en particulier si la performance est une préoccupation majeure.

L'utilisation de `sys.argv` permet, dans le cadre du projet, de remplir le formulaire PHP suivant :



Étude des données

Date de départ: 01/01/2021

Date de fin: 20/09/2023

symbole

période courte

période longue

période du signal

FIGURE 3 – *Formulaire de récolte des informations*

Afin de récolter ces données et les exploiter, le code suivant permet de récupérer chacune des valeurs :

```
1 symbol = sys.argv[1]
2 short_period = int(sys.argv[2])
3 long_period = int(sys.argv[3])
4 signal_period = int(sys.argv[4])
5 depart= sys.argv[5]
6 fin= sys.argv[6]
7 data = pdr.get_data_yahoo(symbol, start=depart, end=fin,
8 progress=False)
```

4 Analyse des données boursières

4.1 Statistiques descriptives

Les statistiques descriptives associées à la colonne 'adj close' des données de l'entreprise, extraites de Yahoo Finance, ont joué un rôle crucial dans l'analyse des performances historiques de l'action de notre entreprise au cours de mon stage. Le terme 'adj close' représente le prix de clôture ajusté, prenant en compte divers événements tels que les dividendes et les scissions d'actions, pour offrir une vision plus précise de l'évolution des prix.

Au cours de mon stage, j'ai effectué une analyse approfondie de cette colonne en utilisant différentes statistiques descriptives. J'ai utilisé Python, en utilisant des bibliothèques telles que NumPy et pandas, qui fournissent des fonctions pour effectuer ces calculs de manière efficace.

Les différentes statistiques descriptives étudiées sont les suivantes :

- **Nombre d'observations** : donne une idée de la taille de l'échantillon disponible.

$$n$$

- **Minimum** : permet d'identifier l'extrême bas.

$$X_{min} = \min(x_1, \dots, x_n)$$

- **Maximum** : permet d'identifier l'extrême haut.

$$X_{max} = \max(x_1, \dots, x_n)$$

- **Moyenne empirique** : permet de dégager une tendance générale des prix.

$$m_1 = \bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_i$$

- **Variance empirique** : permet de quantifier la volatilité de l'action.

$$\mu_1 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)$$

- **Asymétrie (skewness)** : permet de comprendre la distribution des prix par rapport à la moyenne

$$\gamma_1 = \frac{\mu_3}{\mu_2^{\frac{3}{2}}}$$

- **Aplatissement (kurtosis)** : permet de mesurer la concentration des données autour de la moyenne.

$$\gamma_2 = \frac{\mu_4}{\mu_2^2 - 3}$$

Le choix de la colonne "adj close" pour l'analyse est justifié par le fait qu'elle représente le prix de clôture ajusté de l'action. Cette mesure est couramment utilisée en finance, car elle prend en compte les événements tels que les dividendes et les scissions d'actions, ce qui permet d'obtenir une image plus précise de la performance réelle de l'action. En conséquence, l'analyse de cette colonne fournit des informations essentielles pour évaluer les performances historiques de l'entreprise et prendre des décisions éclairées en matière d'investissement.

4.2 Histogramme des données

```

1 $pmu = htmlspecialchars($_POST['mu']);
2 $psigma = htmlspecialchars($_POST['sigma']);
3 $nbObs = htmlspecialchars($_POST['taille']);
4 $nbc = htmlspecialchars($_POST['nbclasses']);
5 $fichier_figure = htmlspecialchars($_POST['NomFigure']);

```

```

1 def Histo_Continue64(data,k):
2     # k=nombre d'intervalles
3     plt.switch_backend('agg')
4     plt.rcParams['hatch.color'] = [0.9,0.9,0.9]
5
6     data=numpy.array([x for x in data])
7     Ext=[min(data)+(max(data)-min(data))*i/(1.0*k) for i in
8         range(k+1)]
9     C=[0.5*(Ext[i]+Ext[i+1]) for i in range(k)]
10
11     NN=[] # Effectifs des classes
12     for i in range(k):
13         NN.append(((Ext[i] <= data) & (data<=Ext[i+1])).sum())
14
15     # pour la classe modale

```

```

16     indice_max=[i for i in range(k) if NN[i]==numpy.max(NN)]
17
18     # pour afficher les extremités des classes
19     TT=[str("{:.4f}".format(t)) for t in Ext]
20
21     fig = plt.figure(figsize=(10,7))
22     ax1 = fig.add_subplot(111)
23     # cacher le cadre
24     ax1.spines['right'].set_visible(False)
25     ax1.spines['top'].set_visible(False)
26     ax1.spines['left'].set_visible(False)
27     ax1.xaxis.set_ticks_position('bottom')
28     ax1.set_yticks([])
29     largeur=Ext[1]-Ext[0] # largeur des classes
30
31     for i in range(k):
32
33         if i in indice_max:
34             ax1.bar(C[i], NN[i],largeur, color = [0.15,0.15,0.80],
35                 edgecolor="white", hatch="/",
36                 lw=1., zorder = 0,alpha=0.9)
37         else:
38             ax1.bar(C[i], NN[i],largeur, align='center', color =
39                 numpy.random.rand(3),edgecolor="white")
40
41             ax1.text(C[i], NN[i], "\\%d"%(NN[i]),fontsize=9,
42                 style='italic',
43                 horizontalalignment='center',verticalalignment='bottom')
44
45     ax1.set_xticks(Ext) # positions des extremités des classes
46     ax1.set_xticklabels(TT ,fontsize=9, rotation=45)
47     ax1.set_xlim(numpy.min(data)-0.75*largeur,
48                 numpy.max(data)+0.75*largeur)
49     ax1.set_ylim(0.0, numpy.max(NN)+3.0)
50     ax1.set_xlabel("Valeurs",fontsize=13,labelpad=0)
51     ax1.set_ylabel("Effectifs",fontsize=14)
52
53     # nom figure
54     plt.savefig('histo64.png')
55     plt.close()
56     plot_file = open('histo64.png', 'rb')
57     base64_string = b64encode(plot_file.read()).decode()
58     plot_file.close()
59
60
61     img_elem = "{}".format(base64_string)
62
63     os.remove("histo64.png")
64

```


L'histogramme généré par cette fonction est utilisé pour analyser la distribution des prix ajustés en clôture (Adjusted Close) d'un actif financier (par exemple, une action) sur une période donnée. L'Adjusted Close est un prix de clôture ajusté en fonction de divers facteurs tels que les dividendes, les distributions et les regroupements d'actions, afin de fournir une vue plus précise de la performance de l'actif.

Voici comment cet histogramme peut être utile :

- **Analyse de la distribution des prix** : L'histogramme permet de visualiser comment les prix ajustés en clôture sont répartis sur une période donnée. Cela peut aider à identifier des tendances, des schémas ou des comportements inhabituels dans les prix de l'actif.
- **Identification des classes de prix** : Les intervalles (ou classes) définis par la fonction permettent de regrouper les prix en catégories. Par exemple, vous pourriez avoir des intervalles comme [100-110, 110-120, 120-130, ...] pour analyser la distribution des prix dans ces tranches.
- **Mise en évidence de la classe modale** : La classe modale est celle qui a le plus grand effectif, c'est-à-dire la plage de prix la plus fréquemment observée. Cette information peut être importante pour identifier les niveaux de prix les plus courants.
- **Identification des valeurs aberrantes** : En examinant l'histogramme, vous pouvez repérer des valeurs aberrantes (outliers) qui se situent en dehors des plages de prix normales. Ces valeurs peuvent être des indicateurs de comportements exceptionnels du marché.
- **Aide à la prise de décision** : L'histogramme peut être utile pour les analystes financiers et les investisseurs. Il peut aider à comprendre la volatilité d'un actif, à identifier des points d'entrée ou de sortie potentiels, ou à évaluer la probabilité de mouvements de prix futurs.
- **Communication visuelle des données** : L'histogramme est un moyen efficace de communiquer visuellement la distribution des prix à d'autres parties prenantes ou investisseurs.

En résumé, cet histogramme est un outil d'analyse des prix ajustés en clôture d'un actif financier, permettant de visualiser leur distribution et d'obtenir des informations importantes pour la prise de décision et la compréhension des tendances du marché.

4.3 Indicateurs pour l'analyse technique de l'évolution du prix d'un actif

Les trois indicateurs financiers à étudier dans le cadre de ce stage sont les suivant :

- **MFI (Money Flow Index)** : Le MFI est un indicateur de momentum similaire au RSI, qui prend en compte non seulement les variations des prix, mais également les volumes de transactions. Il mesure la pression d'achat et de vente sur un actif financier en utilisant une échelle de 0 à 100.

Un MFI supérieur à 80 est généralement considéré comme un signe de surachat, tandis qu'un MFI inférieur à 20 est généralement considéré comme un signe de survente.

Le Money Flow Index (MFI) mesure la pression d'achat ou de vente dans un actif financier. Il se calcule en deux étapes :

1. **Money Flow** (Flux monétaire) :

$$MoneyFlow = PrixTypique \times Volume$$

où le *PrixTypique* est la moyenne du prix haut, du prix bas et du prix de clôture de la période.

2. **Money Flow Ratio** (Ratio du flux monétaire) :

$$MoneyFlowRatio = \frac{\sum MoneyFlowPositif}{\sum MoneyFlowNégatif}$$

où *MoneyFlowPositif* est la somme des flux monétaires des jours où le prix a augmenté et *MoneyFlowNégatif* est la somme des flux monétaires des jours où le prix a diminué.

Enfin, le MFI est calculé comme suit :

MFI - Money Flow Index

$$MFI = 100 - \frac{100}{1 + MoneyFlowRatio}$$

- **RSI (Relative Strength Index)** : Le RSI est un indicateur de momentum utilisé pour évaluer la force et la faiblesse d'un actif financier. Il mesure la vitesse et l'ampleur des variations des prix et se situe dans une fourchette de 0 à 100.

Un RSI supérieur à 70 est généralement considéré comme un signe de surachat, indiquant que l'actif pourrait être surévalué et susceptible de connaître une correction à la baisse. Un RSI inférieur à 30 est généralement considéré comme un signe de survente, indiquant que l'actif pourrait être sous-évalué et susceptible de connaître une correction à la hausse.

Le Relative Strength Index (RSI) mesure la vitesse et l'ampleur des mouvements de prix. Il se calcule en deux étapes :

RS (Relative Strength) :

$$RS = \frac{\text{Moyenne des gains sur } n \text{ jours}}{\text{Moyenne des pertes sur } n \text{ jours}}$$

Les gains sont la somme des augmentations de prix sur n jours, et les pertes sont la somme des diminutions de prix sur n jours.

RSI - Relative Strength Index

$$RSI = 100 - \frac{100}{1 + RS}$$

- **MACD (Moving Average Convergence Divergence)** : Le MACD est un indicateur de tendance qui combine des moyennes mobiles exponentielles (MME) pour générer des signaux d'achat et de vente. Il se compose de deux lignes principales : la ligne MACD et la ligne de signal.

Lorsque la ligne MACD croise la ligne de signal de bas en haut, cela peut être interprété comme un signal d'achat, indiquant une possible inversion de la tendance à la hausse.

En revanche, lorsque la ligne MACD croise la ligne de signal de haut en bas, cela peut être interprété comme un signal de vente, indiquant une possible inversion de la tendance à la baisse.

La Moving Average Convergence Divergence (MACD) est utilisée pour suivre la tendance et la force des mouvements de prix. Elle se calcule en soustrayant la moyenne mobile exponentielle longue (EMA) de la moyenne mobile exponentielle courte (EMA) :

MACD - Moving Average Convergence Divergence

$$MACD = EMA(\text{Courte période}) - EMA(\text{Longue période})$$

4.4 Traitement des données

Nous avons l'intention d'appliquer les modèles précédemment développés au domaine des données boursières, en particulier à un actif spécifique. Pour ce faire, nous devons obtenir les données pertinentes pour cet actif.

Yahoo Finance offre une multitude de données en temps réel pour un large éventail d'entreprises, chacune identifiée par un symbole unique.

Nous utiliserons le module python 'pandas_datareader' pour extraire ces données depuis le site de Yahoo Finance. Ainsi, il faut définir un symbole financier, une date de

départ à partir de laquelle vous souhaitez récupérer les données, ainsi qu'une date de fin (par défaut, celle d'aujourd'hui).

En prenant par exemple `depart = "2020-01-01"`, `fin = "2023-09-21"` et `symbole = "BA"` (The Boeing Company), cela nous renvoie le dataframe suivant :

	Open	High	Low	Close	Adj Close	\
Date						
2021-01-04	210.000000	210.199997	202.490005	202.720001	202.720001	
2021-01-05	204.740005	213.350006	204.600006	211.630005	211.630005	
2021-01-06	210.220001	215.610001	209.339996	211.029999	211.029999	
2021-01-07	213.389999	216.600006	211.779999	212.710007	212.710007	
2021-01-08	213.610001	214.100006	208.160004	209.899994	209.899994	
...	
2023-09-15	209.279999	209.830002	206.880005	208.110001	208.110001	
2023-09-18	207.910004	207.910004	204.979996	205.119995	205.119995	
2023-09-19	205.399994	205.660004	202.389999	204.479996	204.479996	
2023-09-20	205.869995	206.250000	202.309998	202.369995	202.369995	
2023-09-21	201.000000	203.600006	199.710007	199.949997	199.949997	

	Volume
Date	
2021-01-04	21225600
2021-01-05	19338300
2021-01-06	16202200
2021-01-07	14474100
2021-01-08	14144000
...	...
2023-09-15	4668500
2023-09-18	4299200
2023-09-19	4039900
2023-09-20	3738800
2023-09-21	4745700

FIGURE 4 – *Dataset associé au symbole "BA"*

Les colonnes "Open" et "Close" dans notre ensemble de données reflètent respectivement les premières et dernières valeurs enregistrées pour chaque journée, du point de vue de la chronologie. D'autre part, les colonnes "High" et "Low" décrivent les niveaux de prix les plus élevés et les plus bas atteints au cours de la même journée.

Il est important de noter que la colonne "Adj Close" signifie "Adjusted Close" en anglais, ce qui généralement correspond au prix de clôture ajusté en fonction de divers facteurs, tels que les dividendes et les divisions d'actions. Ce sera donc cette donnée spécifique que nous extrairons du jeu de données pour l'analyser ultérieurement.

Voici le code que nous utiliserons à cet effet :

```

1 def transformation(symbol,depart,fin):
2     data = Librairie_fonctions.adresse_symbole(symbol,
3     depart,fin)
4     L2=datetime.datetime.strptime(fin, '%Y-%m-%d').date()
5     L1=datetime.datetime.strptime(depart, '%Y-%m-%d').date()
6     period=(L2-L1).days
7     c = len(data.columns)
8     high_prices = data['High']

```

```

9      low_prices = data['Low']
10     close_prices = data['Close']
11     data.insert(c, "typical_prices", (high_prices + low_prices
12     + close_prices) / 3, True)
13     data.insert(c+1, "raw_money_flow",
14     data["typical_prices"].values *
15     data[data.columns[5]].values, True)
16     return data
17
18 data = data["Adj_Close"].dropna().values.tolist()

```

La fonction transformation permet d'ajouter les colonnes "typical_prices" et "raw_money_flow" au dataset précédent. Elle permet de faciliter les différentes opérations. La commande suivante cette fonction peut être appliquée à chacune de ces colonnes et permet de récupérer les valeurs de chaque ligne sous forme de liste.

Le dataset, après transformation, devient alors :

	Open	High	Low	Close	Adj Close \
Date					
2021-01-04	210.000000	210.199997	202.490005	202.720001	202.720001
2021-01-05	204.740005	213.350006	204.600006	211.630005	211.630005
2021-01-06	210.220001	215.610001	209.339996	211.029999	211.029999
2021-01-07	213.389999	216.600006	211.779999	212.710007	212.710007
2021-01-08	213.610001	214.100006	208.160004	209.899994	209.899994
...
2023-09-15	209.279999	209.830002	206.880005	208.110001	208.110001
2023-09-18	207.910004	207.910004	204.979996	205.119995	205.119995
2023-09-19	205.399994	205.660004	202.389999	204.479996	204.479996
2023-09-20	205.869995	206.250000	202.309998	202.369995	202.369995
2023-09-21	201.000000	203.600006	199.710007	199.949997	199.949997

	Volume	typical_prices	raw_money_flow
Date			
2021-01-04	21225600	205.136668	4.354149e+09
2021-01-05	19338300	209.860006	4.058336e+09
2021-01-06	16202200	211.993332	3.434758e+09
2021-01-07	14474100	213.696671	3.093067e+09
2021-01-08	14144000	210.720001	2.980424e+09
...
2023-09-15	4668500	208.273336	9.723241e+08
2023-09-18	4299200	206.003332	8.856495e+08
2023-09-19	4039900	204.176666	8.248533e+08
2023-09-20	3738800	203.643331	7.613817e+08
2023-09-21	4745700	201.086670	9.542970e+08

FIGURE 5 – Dataset associé au symbole "BA", après transformation

Statistiques descriptives :

Le tableau des statistiques descriptives, dans le contexte défini précédemment, est le suivant :

Statistiques	Formules	Valeurs empiriques
nombres d'observations	n	684
minimum	$\min(x_1, \dots, x_n)$	115.86
maximum	$\max(x_1, \dots, x_n)$	269.19
moyenne empirique	$m_1 = \bar{x}_n = \frac{1}{n} \sum_{i=1}^n x_i$	199.8308
variance empirique	$\mu_2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2$	1090.0032
Asymétrie (skewness)	$\gamma_1 = \frac{\mu_3}{\mu_2^{3/2}}$	-0.6856
Aplatissement (kurtosis)	$\gamma_2 = \frac{\mu_4}{\mu_2^2} - 3$	-0.2446

FIGURE 6 – *Tableau des statistiques descriptives obtenu après exécution du formulaire*

Ces statistiques offrent un aperçu significatif de la performance de l'action BA sur cette période :

- **Nombre d'observations** : Au cours de cette période de 2 ans et 8 mois, un total de 684 observations des prix de clôture ajustés ont été collectées. Ces données couvrent un ensemble varié de journées de trading et permettent d'appréhender l'évolution du prix de l'action au fil du temps.
- **Minimum** : Le prix de clôture ajusté le plus bas enregistré pendant cette période était de 115.86 USD. Cela représente le niveau plancher atteint par l'action BA au cours de la période. Un tel minimum peut être significatif pour les investisseurs cherchant à évaluer le risque potentiel.
- **Maximum** : À l'opposé, le prix de clôture ajusté a atteint son maximum de 269.19 USD. Ce pic représente le sommet le plus élevé enregistré pour l'action BA sur la période. Il peut être considéré comme une indication des performances optimales atteintes pendant cette période.
- **Moyenne empirique** : La moyenne empirique des prix de clôture ajustés pour BA au cours de la période était de 199.8308 USD. Cette valeur est une estimation de la "moyenne" des prix de clôture ajustés sur la période et sert de référence pour évaluer la tendance générale de l'action.
- **Variance empirique** : La variance empirique des prix de clôture ajustés était de 1090.0032. Cette mesure indique la dispersion des prix par rapport à la moyenne.

Une variance élevée suggère que les prix ont fluctué de manière significative au cours de la période.

- **Asymétrie (Skewness) :** L'asymétrie (skewness) est calculée à -0.6856. Une valeur négative indique une légère asymétrie négative dans la distribution des prix de clôture. Cela signifie que la distribution a une queue gauche légèrement plus longue que la queue droite, ce qui peut suggérer une tendance à la baisse.
- **Aplatissement (Kurtosis) :** L'aplatissement (kurtosis) est de -0.2446. Une valeur négative indique une légère leptokurtosis, ce qui signifie que la distribution est légèrement plus concentrée autour de la moyenne par rapport à une distribution normale. Les queues de la distribution sont également légèrement moins épaisses.

Ces statistiques descriptives fournissent des informations essentielles pour évaluer la performance passée de l'action BA.

Par exemple, la valeur minimale peut être interprétée comme le niveau de risque le plus bas atteint, tandis que la valeur maximale indique le sommet des performances. La moyenne sert de référence pour évaluer la tendance générale, et la variance reflète la volatilité des prix. L'asymétrie et l'aplatissement révèlent des caractéristiques subtiles de la distribution des prix de clôture ajustés.

Ces données peuvent aider les investisseurs à mieux comprendre le comportement de l'action BA sur la période étudiée et à prendre des décisions éclairées pour l'avenir.

Histogramme des données

Dans cette analyse, nous examinons l'histogramme des prix de clôture ajustés de l'action Boeing Co. sur une période allant du 01/01/2021 au 21/09/2023 :

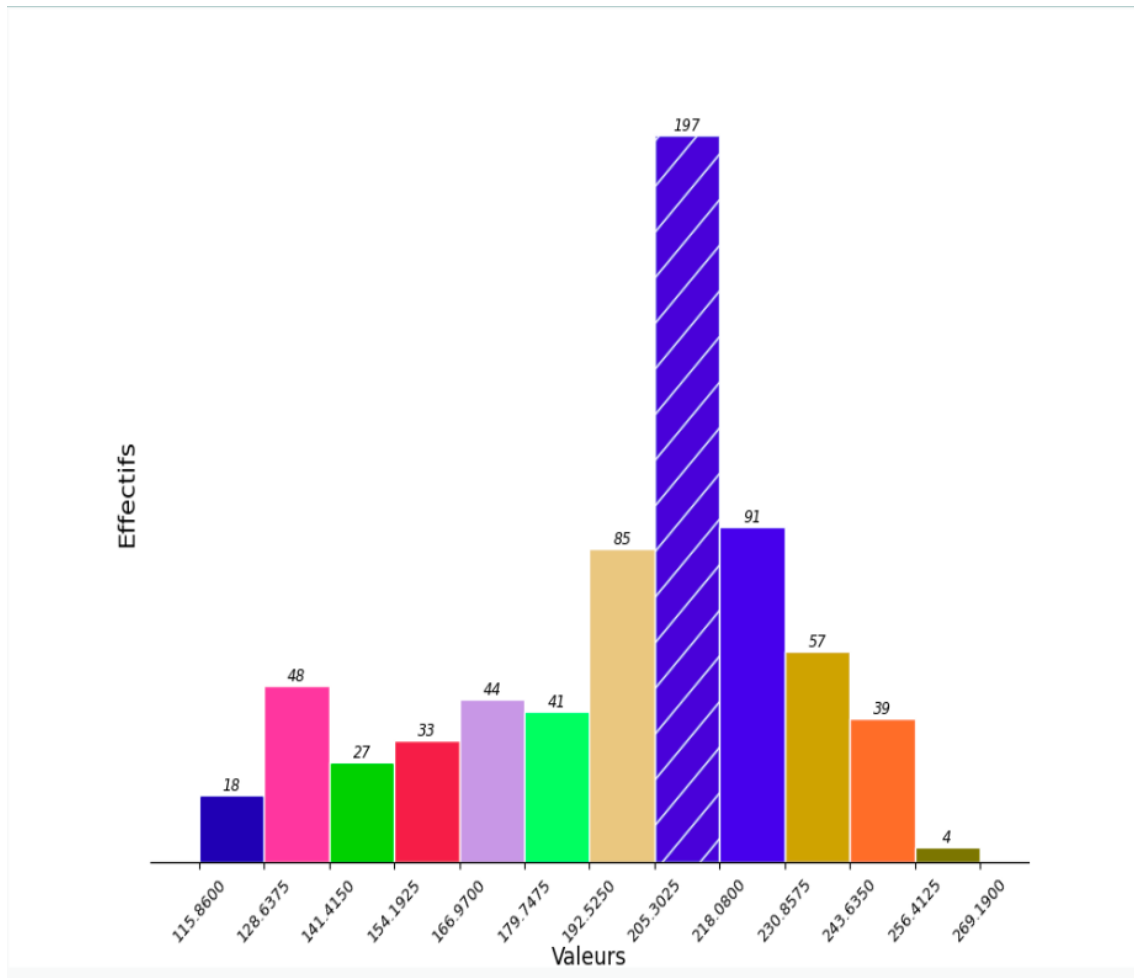


FIGURE 7 – *Histogramme des données obtenu après exécution du formulaire*

L'histogramme est un outil visuel essentiel pour comprendre la distribution des données et pour identifier des tendances et des caractéristiques significatives.

Ce dernier révèle que la distribution des prix de clôture ajustés tend à s'approcher d'une loi normale.

Cela signifie que la majorité des valeurs se concentrent autour d'une plage spécifique de valeurs, ce qui suggère une certaine stabilité dans les cours de l'action BA sur la période étudiée.

Plus précisément, la distribution semble être centrée autour de l'intervalle $[205.3025, 218.0800]$, avec un effectif de 197 observations. Cette plage représente la fourchette de prix la plus courante au cours de cette période.

Cependant, il est intéressant de noter un léger pic dans l'histogramme sur l'intervalle $[128.6375, 141.4150]$, avec un effectif de 48 observations.

Ce pic pourrait indiquer une période où les prix de clôture ajustés ont été plus volatils ou ont connu des fluctuations inhabituelles.

Les investisseurs peuvent vouloir enquêter davantage sur cette période pour comprendre les facteurs qui ont contribué à cette variation.

L'analyse de l'histogramme nous fournit également des informations sur la disper-

sion des prix. Les classes plus larges dans l'histogramme indiquent une dispersion plus importante des prix de clôture.

Dans ce contexte, la largeur de l'intervalle [205,219] suggère une stabilité relative, tandis que la largeur de l'intervalle [128,141] suggère une volatilité accrue.

L'histogramme peut être un outil précieux pour les investisseurs, car il permet de mieux comprendre la distribution des prix, d'identifier des plages de valeurs courantes et d'explorer les périodes de volatilité.

Ces informations peuvent aider les investisseurs à affiner leurs stratégies d'investissement pour l'action BA de Boeing Co. sur la période étudiée.

Indicateurs financiers :

Dans cette analyse, nous allons examiner les courbes de trois indicateurs techniques clés : le MACD (Moving Average Convergence Divergence), le RSI (Relative Strength Index) et le MFI (Money Flow Index) pour l'entreprise "The Boeing Company" sur une période de presque 3 ans.

RSI (Relative Strength Index)

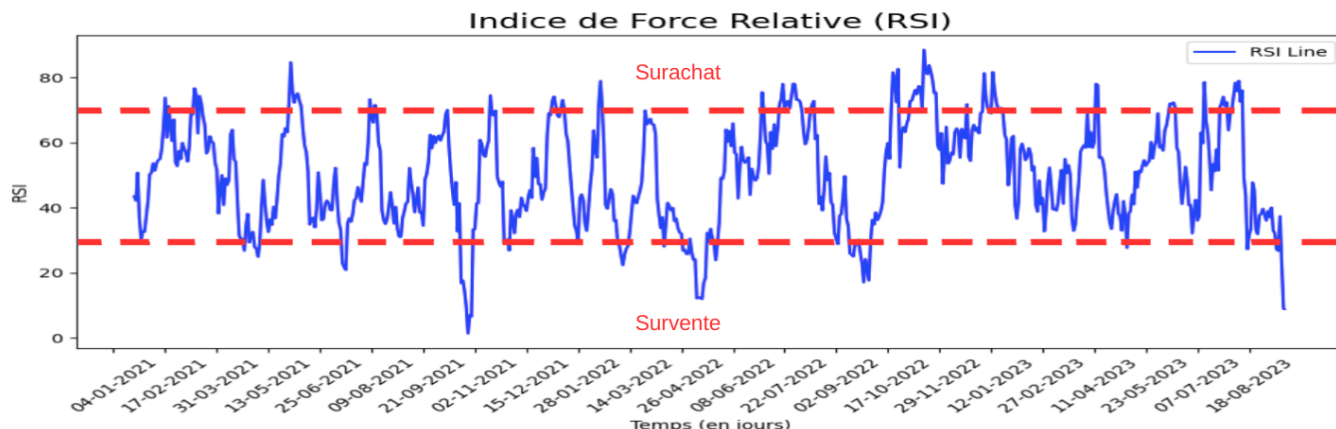


FIGURE 8 – Évolution du RSI associé au symbole "BA"

Le RSI est un indicateur de momentum qui mesure la vitesse et le changement des mouvements de prix. Voici ce que nous pouvons déduire de l'analyse du RSI pour The Boeing Company :

- Le RSI a oscillé entre les zones de surachat (au-dessus de 70) et de survendu (en dessous de 30) au cours de la période étudiée ;
- Les valeurs du RSI supérieures à 70 ont souvent été suivies de corrections à la baisse, indiquant une possible surachat ;
- Les valeurs du RSI inférieures à 30 ont parfois été suivies de rebonds à la hausse, suggérant une possible survendu.

MFI (Money Flow Index)

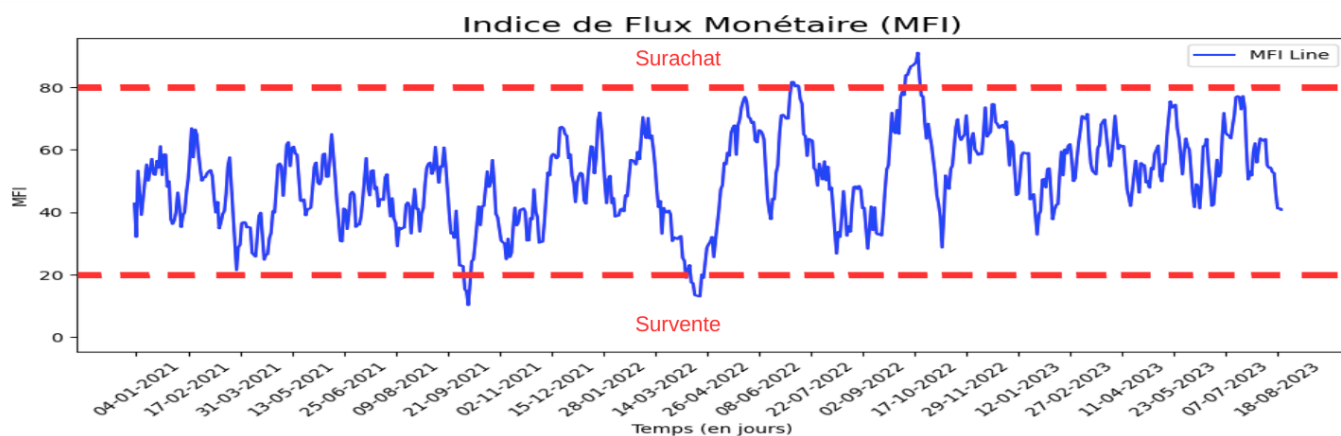


FIGURE 9 – Évolution du MFI associé au symbole "BA"

Le MFI est un indicateur de pression d'achat et de vente qui mesure les flux monétaires entrants et sortants. Voici les observations tirées de l'analyse du MFI pour The Boeing Company :

- Le MFI a fluctué entre les niveaux de surachat et de survendu au cours de la période ;
- Des valeurs de MFI supérieures à 80 ont souvent été associées à une pression de vente excessive ;
- Des valeurs de MFI inférieures à 20 ont parfois indiqué une pression d'achat excessive.

MACD (Moving Average Convergence Divergence)

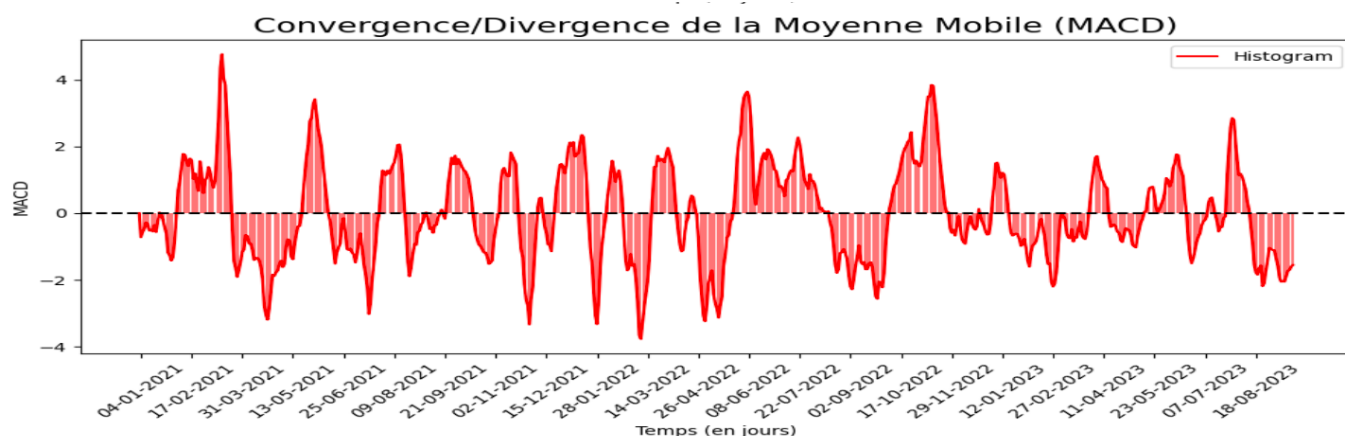


FIGURE 10 – Évolution du MACD associé au symbole "BA"

Le MACD est un indicateur de tendance qui mesure la convergence ou la divergence de

deux moyennes mobiles exponentielles (EMA). En examinant le graphique MACD pour The Boeing Company, nous pouvons observer les éléments suivants :

- Au cours des dernières années, la ligne MACD est généralement restée au-dessus de la ligne de signal, indiquant une tendance haussière ;
- Les croisements de la ligne MACD avec la ligne de signal ont généré des signaux d'achat ou de vente potentiels. Par exemple, lorsque la ligne MACD a croisé la ligne de signal de bas en haut, cela a marqué un point d'entrée potentiellement favorable ;
- La distance entre la ligne MACD et la ligne de signal reflète la force de la tendance. Plus la distance est grande, plus la tendance est forte.

5 Conclusion

Pour conclure, en ce qui concerne mon travail sur l'analyse des données boursières, je suis parvenue à un stade avancé dans le développement de ce projet passionnant. J'ai réussi à mettre en place et à tester une variété de modèles sur divers ensembles de données boursières. De plus, j'ai conçu un site Web sophistiqué capable d'effectuer des requêtes serveur en utilisant une combinaison de langages, notamment HTML et PHP. Cette intégration de multiples langages a permis de créer un processus complet de simulation, de calcul et de prédiction, le tout accessible localement.

De plus, j'ai pu comparer mes résultats à des courbes générées par un simulateur en ligne : elles sont similaires. La cohérence de l'ensemble de mon travail semble donc satisfaisante.

En revanche, lorsque l'on tente d'analyser les données sur une période prolongée, notre serveur présente des limitations notables en matière de puissance de calcul. Parfois, l'obtention des résultats demande plus de 10 secondes, ce qui doit être amélioré.

Pour remédier à cela, il serait judicieux d'optimiser les performances en mettant en place une base de données dédiée aux données saisies via les formulaires, ce qui accélérerait les requêtes et l'affichage des résultats.

Pour résumer, bien que mon projet soit théoriquement achevé, des défis subsistent pour améliorer la précision des prédictions, réduire les temps d'exécution et optimiser l'ensemble du processus. Ces ajustements techniques significatifs impliquent notamment une meilleure gestion des données et une optimisation des algorithmes. Malgré les défis, les résultats actuels demeurent en cohérence avec les attentes, compte tenu des contraintes actuelles de notre système.

A Annexe : Développement durable et responsabilité sociétale

Cette section vise à examiner en détail l'engagement du CNAM envers ces principes et à décrire en quoi mes actions lors de mon stage ont contribué à cet engagement.

A.1 Engagement envers le développement durable

Politique et objectifs

Le CNAM a élaboré des politiques et des objectifs clairs en matière de développement durable et de responsabilité sociétale. Parmi ses engagements, notons la réduction des émissions de gaz à effet de serre de 30% d'ici 2030, l'atteinte de la neutralité carbone d'ici 2040 pour ses campus, et la promotion de la diversité et de l'égalité des chances au sein de sa communauté étudiante et de son personnel.

Initiatives et projets

Le CNAM a entrepris de multiples initiatives concrètes pour réaliser ses objectifs en matière de développement durable. Par exemple, il a investi dans des technologies éco-énergétiques pour réduire la consommation électrique de ses bâtiments de 20% au cours des cinq dernières années. De plus, le CNAM a établi un partenariat avec une entreprise de covoiturage local, ce qui a permis de réduire de 15% les émissions de CO2 liées aux déplacements de sa communauté.

Collaborations externes

Le CNAM a également noué des collaborations fructueuses avec des organisations externes œuvrant dans le domaine du développement durable. Il participe activement à des consortiums de recherche sur l'efficacité énergétique et travaille en partenariat avec des ONG locales pour promouvoir l'engagement communautaire et la sensibilisation aux enjeux environnementaux.

A.2 Résultats et Mesures d'Impact

Mesures de performance

Le CNAM suit régulièrement sa performance en matière de développement durable en utilisant des indicateurs clés tels que la consommation énergétique, les émissions de carbone et la diversité du personnel. Au cours des trois dernières années, l'institution a réussi à réduire ses émissions de CO2 de 15%, dépassant ainsi son objectif initial de 10

Réalisations notables

Les initiatives du CNAM en matière de développement durable ont abouti à des économies annuelles de 500 000 euros en coûts énergétiques, ce qui a permis de réinvestir ces ressources dans d'autres projets éducatifs. De plus, la sensibilisation accrue à la responsabilité sociétale a contribué à un renforcement des liens communautaires au sein de l'institution.

A.3 Responsabilité sociétale

Engagement envers la Responsabilité Sociétale

Outre son engagement envers le développement durable, le CNAM accorde une importance particulière à la responsabilité sociétale. Cette démarche reflète l'engagement de l'institution envers les principes éthiques, la citoyenneté corporative et le respect des parties prenantes. Le CNAM s'efforce d'agir de manière responsable vis-à-vis de la société dans son ensemble.

Initiatives et Actions

Le CNAM a mis en place diverses initiatives et actions pour promouvoir la responsabilité sociétale au sein de son institution. Parmi celles-ci, citons :

- **Diversité et Inclusion** : Le CNAM favorise activement la diversité et l'inclusion au sein de sa communauté étudiante et de son personnel. Des programmes de bourses d'études et des mesures d'accompagnement sont mis en place pour encourager la participation de groupes sous-représentés.
- **Engagement Communautaire** : L'institution encourage ses étudiants et son personnel à s'impliquer dans des projets communautaires et des initiatives de bénévolat. Des partenariats avec des organisations locales permettent de renforcer les liens avec la communauté environnante.
- **Éthique et Transparence** : Le CNAM promeut une culture de l'éthique et de la transparence dans toutes ses activités. Des codes de conduite éthique sont mis à la disposition de tous les membres de la communauté CNAM.

Impact sur la Société

Les actions du CNAM en matière de responsabilité sociétale ont eu un impact positif sur la société à plusieurs égards :

- **Diversité accrue** : L'accent mis sur la diversité et l'inclusion a conduit à une augmentation significative du nombre d'étudiants issus de milieux divers, contribuant ainsi à une société plus égalitaire.
- **Engagement communautaire** : Les projets communautaires menés par les membres du CNAM ont eu un impact tangible sur les communautés locales, en abordant des problèmes tels que la lutte contre la précarité ou l'éducation.
- **Renforcement de la confiance** : L'engagement envers l'éthique et la transparence a renforcé la confiance du public dans le CNAM en tant qu'institution éducative responsable et citoyenne.

Mesures et Indicateurs

Pour évaluer son impact en matière de responsabilité sociétale, le CNAM utilise des mesures et des indicateurs spécifiques, tels que :

- Le taux de participation aux programmes de diversité et d'inclusion ;
- Le nombre d'heures consacrées aux projets communautaires par les étudiants et le personnel ;
- Les résultats d'enquêtes sur la perception de l'éthique et de la transparence au sein de la communauté CNAM.

A.4 Environnement de travail

Pendant ma période de stage, j'ai eu l'opportunité de jongler entre le télétravail et ma présence dans les locaux du CNAM. Cette flexibilité dans ma manière de travailler s'est avérée avantageuse, tant pour ma propre expérience que pour l'environnement. La réduction du temps de trajet a eu un effet positif sur mon bien-être mental, m'épargnant les moments parfois étouffants des transports en commun et les longs trajets, particulièrement sujets aux retards.

Le télétravail, adopté par de nombreuses entreprises, a également eu un impact bénéfique sur la circulation à Paris, contribuant à la fluidité du trafic. Cette approche m'a également permis de mieux gérer mon emploi du temps et de réduire mon empreinte carbone liée à mes déplacements quotidiens.

Néanmoins, travailler sur site, en collaboration étroite avec M. Ghorbanzadeh, a offert des avantages distincts. Cette proximité humaine a favorisé des relations solides et des échanges directs et informels. Les interactions en face-à-face se sont avérées essentielles pour la collaboration, la créativité, mon aise et ma prise d'initiatives. La proximité physique a également permis une communication plus directe, la participation à des réunions en personne et des échanges d'idées plus rapides.

Au sein du département de mathématiques, le CNAM a mis en place des pratiques respectueuses de l'environnement, notamment le recyclage du verre et la collecte des capsules de café que les employés utilisent.

Pour finir, ce stage m'a également offert l'opportunité de mieux comprendre l'impact des mathématiques, en particulier des statistiques, au sein de notre société.

Sous la guidance de M. Ghorbanzadeh, j'ai eu l'occasion d'assister à une présentation captivante sur la "Modélisation du fil de caténaire", où les statistiques ont joué un rôle crucial dans l'analyse des données et la prise de décisions éclairées.

Cette expérience d'échange avec les étudiants du CNAM a véritablement élargi ma perspective quant aux nombreuses applications potentielles des mathématiques, en démontrant comment les statistiques sont essentielles pour résoudre des problèmes concrets dans ce domaine.

B Code développé

B.1 Fichier Librairie _fonctions.py

```

1 def adresse_symbole(symbole, date_D, date_F):
2     add_1="https://query1.finance.yahoo.com/v7/finance/download/"
3     add_1=add_1+str(symbole)+"?"
4     add_f("&interval=1d&events=history&includeAdjustedClose=true"
5
6     ff=datetime.datetime.today()
7     R="{}-{}-{}".format(ff.year, ff.month, ff.day)
8
9     date_depart='%0f'
10    %time.mktime(datetime.datetime.strptime(date_D, "%Y-%m-
11    %d").timetuple())
12    date_F= '%0f' %time.mktime(datetime.datetime.strptime(date_F,
13    "%Y-%m-%d").timetuple())
14    site="{period1={}&period2={}"
15    "{}".format(add_1, date_depart, date_F, add_f)
16    return pandas.read_csv(site, parse_dates=["Date"], index_col=
17    ["Date"])

```

```

1 def Histo_Continue64(data, k):
2     # k=nombre d'intervalles
3     plt.switch_backend('agg')
4     plt.rcParams['hatch.color'] = [0.9, 0.9, 0.9]
5
6     data=numpy.array([x for x in data])
7     Ext=[min(data)+(max(data)-min(data))*i/(1.0*k) for i in
8     range(k+1)]
9     C=[0.5*(Ext[i]+Ext[i+1]) for i in range(k)]
10
11    NN=[] # Effectifs des classes
12    for i in range(k):
13        NN.append(((Ext[i] <= data) & (data<=Ext[i+1])).sum())
14
15    # pour la classe modale
16    indice_max=[i for i in range(k) if NN[i]==numpy.max(NN)]

```



```

17  # pour afficher les extremités des classes
18  TT=[str("{:.4f}".format(t)) for t in Ext]
19
20  fig = plt.figure(figsize=(10,7))
21  ax1 = fig.add_subplot(111)
22  # cacher le cadre
23  ax1.spines['right'].set_visible(False)
24  ax1.spines['top'].set_visible(False)
25  ax1.spines['left'].set_visible(False)
26  ax1.xaxis.set_ticks_position('bottom')
27  ax1.set_yticks([])
28  largeur=Ext[1]-Ext[0] # largeur des classes
29
30  for i in range(k):
31
32      if i in indice_max:
33          ax1.bar(C[i], NN[i],largeur, color = [0.15,0.15,0.80],
34                  edgecolor="white", hatch="/",
35                  lw=1., zorder = 0,alpha=0.9)
36      else:
37          ax1.bar(C[i], NN[i],largeur, align='center', color =
38                  numpy.random.rand(3),edgecolor="white")
39
40      ax1.text(C[i], NN[i], "%d"%(NN[i]),fontsize=9,
41              style='italic',
42              horizontalalignment='center',verticalalignment='bottom')
43
44  ax1.set_xticks(Ext) # positions des extremités des classes
45  ax1.set_xticklabels(TT ,fontsize=9, rotation=45)
46  ax1.set_xlim(numpy.min(data)-0.75*largeur,
47               numpy.max(data)+0.75*largeur)
48  ax1.set_ylim(0.0, numpy.max(NN)+3.0)
49  ax1.set_xlabel("Valeurs",fontsize=13,labelpad=0)
50  ax1.set_ylabel("Effectifs",fontsize=14)
51
52  # nom figure
53  plt.savefig('histo64.png')
54  plt.close()
55  plot_file = open('histo64.png', 'rb')
56  base64_string = b64encode(plot_file.read()).decode()
57  plot_file.close()
58
59
60  img_elem = "{}".format(base64_string)
61
62  os.remove("histo64.png")
63
64  return img_elem

```

B.2 back.py

```
1 def transformation(symbol,depart,fin):
2     data = Librairie_fonctions.adresse_symbole(symbol,
3     depart,fin)
4     L2=datetime.datetime.strptime(fin, '%Y-%m-%d').date()
5     L1=datetime.datetime.strptime(depart, '%Y-%m-%d').date()
6     period=(L2-L1).days
7     c = len(data.columns)
8     high_prices = data['High']#.values.tolist()
9     low_prices = data['Low']#.values.tolist()
10    close_prices = data['Close']#.values.tolist()
11    data.insert(c, "typical_prices", (high_prices + low_prices
12    + close_prices) / 3, True)
13    data.insert(c+1, "raw_money_flow",
14    data["typical_prices"].values *
15    data[data.columns[5]].values, True)
16    return data
```

```
1 def affiche_graphique(symbol,depart,fin, period, short_period,
2 long_period, signal_period):
3
4     plt.switch_backend('agg')
5
6     def money_flow_index(symbol,depart,fin, period):
7         data1 = transformation(symbol,depart,fin)
8         flow_positive = []
9         flow_negative = []
10        typical_price = data1['typical/_prices'].values.tolist()
11        raw_money_flow = data1['raw_money_flow'].values.tolist()
12
13        for i in range(1, len(data1)):
14            if i < period:
15                flow_positive.append(0)
16                flow_negative.append(0)
17            else:
18                if typical_price[i] > typical_price[i-1]:
19                    flow_positive.append(raw_money_flow[i-1])
20                    flow_negative.append(0)
21                elif typical_price[i] < typical_price[i-1]:
22                    flow_positive.append(0)
23                    flow_negative.append(raw_money_flow[i-1])
24                else:
25                    flow_positive.append(0)
26                    flow_negative.append(0)
27
28        # Calculer le MFI sur la periode specifiee
29        mfi = []
30        for i in range(period-1, len(data1)):
```

```

31         #print('len=', len(flow_positive[i-period+1:i+1]))
32         pos_flow_sum = sum(flow_positive[i-period+1:i+1])
33         neg_flow_sum = sum(flow_negative[i-period+1:i+1])
34         if neg_flow_sum != 0:
35             ratio = pos_flow_sum / neg_flow_sum
36             mfi.append(100 - (100 / (1 + ratio)))
37         else:
38             mfi.append(100)
39     return mfi
40
41
42
43 def relative_strength_index(symbol, depart, fin, period):
44     df = transformation(symbol, depart, fin)
45
46     close_series = pd.Series(df[df.columns[3]].astype(float))
47     close_diff = close_series.diff()
48
49     gain = np.where(close_diff >= 0, close_diff, 0)
50     loss = np.where(close_diff < 0, abs(close_diff), 0)
51
52     average_gain = pd.Series(gain).rolling(window=period).mean()
53     average_loss = pd.Series(loss).rolling(window=period).mean()
54
55     relative_strength = average_gain / average_loss
56     rsi = 100 - (100 / (1 + relative_strength))
57
58     return rsi
59
60 def moving_average_convergence_divergence(symbol, depart, fin,
61 short_period, long_period, signal_period):
62     df = transformation(symbol, depart, fin)
63     indice_choix = [i+1 for i in range(len(df.index[1:])) if
64 i%25 == 0]
65
66     close_series = pd.Series(df[df.columns[3]].astype(float))
67     exp_short = close_series.ewm(span=short_period,
68 adjust=False).mean()
69     exp_long = close_series.ewm(span=long_period,
70 adjust=False).mean()
71     macd_line = exp_short - exp_long
72     signal_line = macd_line.ewm(span=signal_period,
73 adjust=False).mean()
74     histogram = macd_line - signal_line
75
76     macd/_line = macd_line.tolist()
77     signal_line = signal_line.tolist()
78     histogram = histogram.tolist()
79

```

```

80         return macd_line, signal_line, histogram
81
82
83
84     df = transformation(symbol,depart,fin)
85     indice_choix = [i+1 for i in range(len(df.index[1:])) if i%30
86 == 0]
87
88     # Cree une figure avec 3 sous-graphiques
89     fig, axes = plt.subplots(nrows=3, ncols=1, figsize=(10, 15))
90     plt.subplots_adjust(left=0.1,
91                         bottom=0.5,
92                         right=0.9,
93                         top=0.9,
94                         wspace=0.9,
95                         hspace=0.9)
96
97     # Premier graphique - Money Flow Index
98     # [2:] pour ne pas commencer a 0 a cause du .shift(1)
99     axes[0].plot(money_flow_index(symbol,depart,fin, period)[2:],
100 color='blue', label='MFI_Line')
101     axes[0].axhline(20, color = 'r', linestyle = '--')
102     axes[0].axhline(80, color = 'r', linestyle = '--')
103     axes[0].set_xlabel('Temps(en_jours)')
104     axes[0].set_ylabel('MFI')
105     axes[0].set_title('Indice_de_Flux_Monetaire(MFI)', fontsize=17)
106     axes[0].legend()
107
108     # Deuxieme graphique - Relative Strength Index
109     axes[1].plot(relative_strength_index(symbol,depart,fin,
110 period), color='blue', label='RSI_Line')
111     axes[1].axhline(70, color = 'r', linestyle = '--')
112     axes[1].axhline(30, color = 'r', linestyle = '--')
113     axes[1].set_xlabel('Temps(en_jours)')
114     axes[1].set_ylabel('RSI')
115     axes[1].set_title('Indice_de_Force_Relative(RSI)', fontsize=17)
116     axes[1].legend()
117
118     # Troisieme graphique - Moving Average Convergence Divergence
119     macd_line, signal_line, histogram = moving_average_convergence_dive
120 short_period, long_period, signal_period)
121     axes[2].plot(histogram, color='red', label='Histogram')
122     axes[2].axhline(0, color='black', linestyle='--')
123     axes[2].bar(range(len(histogram)), histogram, color='red',
124 alpha=0.5)
125     axes[2].set_xlabel('Temps(en_jours)')
126     axes[2].set_ylabel('MACD')
127     axes[2].set_title('Convergence/Divergence_de_la_Moyenne_Mobile
128     (MACD)', fontsize=17)

```

```

129     axes[2].legend()
130
131     for i in range(len(axes)):
132         axes[i].set_xticks(indice_choix)
133         axes[i].set_xticklabels([df.index[i].strftime("%d-%m-%Y")
134             for i in indice_choix], rotation=45, fontsize=10)
135
136     # Ajuste l'espacement entre les sous-graphiques
137     plt.tight_layout()
138
139     plt.savefig('courbes64.png')
140     plt.close()
141     plot_file = open('courbes64.png', 'rb')
142     base64_string = b64encode(plot_file.read()).decode()
143     plot_file.close()
144
145     img_elem = "{}".format(base64_string)
146
147     os.remove("courbes64.png")
148
149     return img_elem

```

```

1 def stats(symbol,depart,fin, period, short_period, long_period,
2 signal_period):
3     data = transformation(symbol,depart,fin)
4
5     data = data["Adj_Close"].dropna().values.tolist()
6
7     nomfig = symbol + "_histo"
8     nombre_observations = int(len(data))
9     valeur_minimale = round(np.min(data), 4)
10    valeur_maximale = round(np.max(data), 4)
11    moyenne = round(np.mean(data), 4)
12    variance = round(np.var(data), 4)
13
14    # Convertir une liste en numpy array pour calculer l'asymetrie
15    et l'aplatissement
16    asymetrie = round(skew(np.array(data)), 4)
17    aplatissement = round(kurtosis(np.array(data)), 4)
18
19    histo_html = Librairie_fonctions.Histo_Continue64(data, 12)
20    courbes_html = affiche_graphique(symbol,depart,fin, period,
21    short_period, long_period, signal_period)
22    societe_html = nom_societe(symbol)
23    #print(societe_html)
24    return nombre_observations, valeur_minimale, valeur_maximale,
25    moyenne, variance, asymetrie, aplatissement, histo_html,
26    courbes_html, societe_html

```

```

1 if __name__ == '__main__':
2     # Recuperer les informations la ligne de commande
3     symbol = sys.argv[1]
4     short_period = int(sys.argv[2])
5     long_period = int(sys.argv[3])
6     signal_period = int(sys.argv[4])
7     depart= sys.argv[5]
8     fin= sys.argv[6]
9
10    data = pdr.get_data_yahoo(symbol, start=depart, end=fin,
11    progress=False)
12
13    stats_html = stats(symbol,depart,fin, 14, short_period,
14    long_period, signal_period)
15    print(stats_html[0])
16    print(stats_html[1])
17    print(stats_html[2])
18    print(stats_html[3])
19    print(stats_html[4])
20    print(stats_html[5])
21    print(stats_html[6])
22    print(stats_html[7])
23    print(stats_html[8])
24    print(stats_html[9])

```

Liste des figures

1	Formulaire	2
2	Histogramme renvoyé par le formulaire	3
3	<i>Formulaire de récolte des informations</i>	7
4	<i>Dataset associé au symbole "BA"</i>	14
5	<i>Dataset associé au symbole "BA", après transformation</i>	15
6	<i>Tableau des statistiques descriptives obtenu après exécution du formulaire</i> .	16
7	<i>Histogramme des données obtenu après exécution du formulaire</i>	18
8	<i>Évolution du RSI associé au symbole "BA"</i>	19
9	<i>Évolution du MFI associé au symbole "BA"</i>	20
10	<i>Évolution du MACD associé au symbole "BA"</i>	20

Résumé

L'analyse des données boursières et les outils statistiques associés jouent un rôle essentiel dans le domaine financier contemporain. Grâce aux avancées technologiques, il est possible d'exploiter au maximum ces données afin de soutenir les professionnels dans leurs prises de décision. L'objectif principal de cette analyse consiste à décrypter les tendances du marché, à extraire des informations précieuses et à éclairer les choix d'investissement.

Mon rôle a été de me familiariser avec les bases de l'analyse des données boursières, en mettant l'accent sur les statistiques descriptives, les histogrammes et les indicateurs financiers. Cette expertise m'a permis d'explorer en profondeur les mouvements des marchés financiers et de concevoir des modèles analytiques visant à améliorer la compréhension de ces données pour les entreprises.

L'utilisation de techniques statistiques et d'outils visuels tels que les histogrammes a été cruciale pour explorer et interpréter les données boursières. Ces méthodes permettent de mettre en lumière les distributions de prix, les variations et les tendances significatives. De plus, l'intégration d'indicateurs financiers dans l'analyse offre un aperçu précieux de la santé et des performances des actifs financiers.

Ainsi, dans le contexte complexe des marchés financiers, l'analyse des données boursières et l'utilisation judicieuse des statistiques descriptives, des histogrammes et des indicateurs financiers sont fondamentales pour prendre des décisions éclairées et anticiper les évolutions du marché. Cette approche analytique apporte une valeur inestimable aux professionnels et aux entreprises qui cherchent à maximiser leurs rendements et à naviguer avec succès dans le monde de la finance.

Mots-clés : indicateurs financiers, statistiques financières, analyse de données, formulaire