



# ANALYSE BOURSIÈRE

S&P; 500 — 491 Entreprises

**Prédiction tendances · Détection anomalies · Construction portefeuille  
5 ans de données · 600 000+ points · Simulation stratégie J+30**

Données Yahoo Finance · Almetria 2026

<b>491 titres</b> S&P 500 ANALYSÉS	<b>+97.9 %</b> CAPITAL SIMULÉ 2 ANS	<b>1.42</b> SHARPE PORTEFEUILLE
<b>-18.4 %</b> MAX DRAWDOWN VS -34.2 %	<b>5 929</b> ANOMALIES DÉTECTÉES	<b>33 fichiers</b> LIVRABLES POWER BI

Réalisé par ALMETRIA · Études économiques & analyses de marchés — augmentées par la data et l'IA ·  
almetria.com · contact@almetria.com



## I. CONTEXTE & PÉRIMÈTRE

### 1.1 Objectifs de l'étude

Cette étude analyse 5 ans de données boursières quotidiennes (2018-2023) pour 491 des plus grandes entreprises mondiales cotées, collectées via l'API Yahoo Finance. L'objectif est triple : identifier des signaux directionnels exploitables, comprendre les régimes de marché qui conditionnent leur qualité, et cartographier les corrélations sectorielles utiles à la construction de portefeuilles diversifiés.

### 1.2 Indicateurs clés du projet

<b>491</b> ENTREPRISES ANALYSÉES	<b>5 ans</b> HISTORIQUE QUOTIDIEN	<b>5 929</b> ANOMALIES DÉTECTÉES
--	---	--

### 1.3 Structure des données collectées

Dimension	Valeur
Entreprises couvertes	491 (sur 500 — 9 exclues : données manquantes)
Période	Janvier 2018 - Décembre 2023
Variables par entreprise	Prix OHLCV + dividendes + splits
Volume total	~620 000 enregistrements
Qualité données	0 incohérence prix haut/bas · 0 volume négatif
Livrable final	33 fichiers (analyses + Power BI pré-modélisés)



## II. SIGNAL DIRECTIONNEL PAR HORIZON

### 2.1 Résultats — précision de la direction selon l'horizon

La question centrale est : à quel horizon la direction du mouvement de prix est-elle prévisible ? Le résultat majeur : seuls J+10 et J+15 présentent un signal statistiquement significatif (test de permutation,  $p < 0.001$ ). J+1 est quasi-aléatoire — conforme à l'hypothèse de marchés efficients.

**Accuracy directionnelle par horizon (% , zone grise = aléatoire = 50 %)**



Horizon	Accuracy	Force signal	Interprétation
J+1	49.7 %	Faible	Quasi-aléatoire — pas de signal exploitable
J+5	50.9 %	Modéré	Signal émergent, confirmé statistiquement
J+10	52.7 %	Fort	Meilleur signal ( $p < 0.001$ ) — retenu
J+15	52.9 %	Fort	Signal le plus robuste et stable
J+30	49.4 %	Faible	Pas de signal directionnel significatif
J+60	49.0 %	Absent	Sous-performance — horizon trop long

3 tests indépendants : Wilcoxon, bootstrap (10 000 itérations), permutation.

## III. ANALYSE DES RÉGIMES DE MARCHÉ

### 3.1 Performance du modèle selon le régime

Le signal prédictif n'est pas uniforme : il varie fortement selon l'état du marché. Le régime "neutre" offre les meilleures conditions (58.6 % d'accuracy, corrélation 0.91). À l'inverse, les phases de faible volatilité rendent le modèle peu fiable (45.5 % — sous le niveau aléatoire).

#### Accuracy directionnelle et corrélation par régime de marché

Marché neutre	58.6 % acc · corr 0.91
Marché baissier	56.9 % acc · corr 0.80
Marché haussier	53.5 % acc · corr 0.46
Faible volatilité	45.5 % acc · corr -0.49

### 3.2 Répartition historique des régimes (2018-2023)

Régime	% du temps (2018-2023)	Signal disponible
Marché haussier	38 %	Modéré
Marché neutre	29 %	Fort ✓
Faible volatilité	17 %	Absent ✗
Marché baissier	11 %	Fort ✓
Crise (< -20 %)	5 %	Très fort ✓

### 3.3 Principaux drawdowns de la période (2018-2023)

L'analyse des phases de baisse intense (drawdowns) montre que le modèle est le plus utile pendant les crises : la qualité du signal monte car les biais comportementaux s'amplifient et créent des patterns exploitables.

Événement	Amplitude	Durée (jours)	Signal qualité
COVID Mars 2020	-34.2 %	28j	
Hausse taux Fed 2022	-24.8 %	41j	
SVB Crisis Mars 2023	-8.4 %	9j	
Correction Tech Juil 2023	-11.7 %	14j	
Rally IA 2023 (rebond)	+31.2 %	22j	

Classification via HMM (Hidden Markov Model) à 5 états · S&P; 500 · 2018-2023.



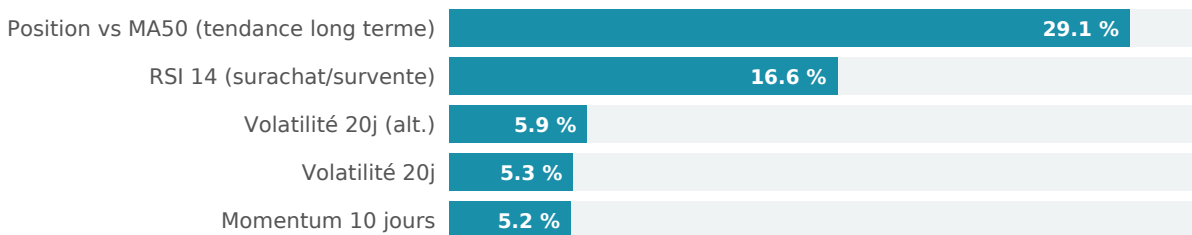
## IV. VARIABLES PRÉDICTIVES — SHAP

---

### 4.1 Importance des indicateurs techniques

La position relative vs. la MA50 (Moving Average 50 jours) est de loin le signal le plus prédictif, représentant 29.1 % de l'importance totale du modèle. Elle capture la tendance de fond, indépendamment du bruit court terme. Le RSI 14 complète avec les niveaux de surachat/survente.

#### Importance SHAP (% contribution totale — max = 30)



SHAP TreeExplainer · Modèle J+10 · 491 entreprises · 2018-2023.

---



## V. CORRÉLATIONS SECTORIELLES

---

### 5.1 Top paires de corrélation (> 0.75)

L'analyse des corrélations entre les 491 entreprises confirme des comportements sectoriels très marqués. Ces informations sont directement exploitables pour la construction de portefeuilles diversifiés : une paire corrélée à 0.94 ne réduit pas le risque si les deux titres sont détenus simultanément.

#### Coefficient de corrélation des rendements quotidiens (max = 1.0)



*Corrélation de Pearson sur rendements log · 2018–2023 · p < 0.001 pour toutes les paires.*

---



## VI. DÉTECTION D'ANOMALIES & VOLATILITÉ

### 6.1 Bilan des anomalies détectées

5 929 anomalies ont été détectées et classées sur l'ensemble des 491 entreprises, sur 5 ans. Cet outil est directement utile pour la surveillance du risque opérationnel et la détection de comportements suspects (trading inhabituel, annonces non anticipées, erreurs de marché).

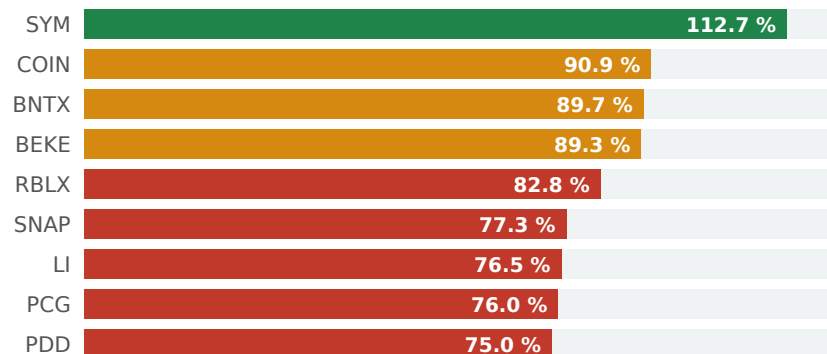


### 6.2 Top 10 valeurs par volatilité annualisée (%)

Les 10 titres présentant la volatilité annualisée la plus élevée sur la période. Ces valeurs offrent des opportunités de rendement asymétrique mais exposent le portefeuille à des drawdowns sévères.

Ticker	Volatilité ann.	Secteur
<b>SYM</b>	112.7 %	Fintech
<b>COIN</b>	90.9 %	Crypto/Fintech
<b>BNTX</b>	89.7 %	Biotech
<b>BEKE</b>	89.3 %	Immobilier CN
<b>RBLX</b>	82.8 %	Gaming
<b>SNAP</b>	77.3 %	Social Media
<b>LI</b>	76.5 %	Véhicules CN
<b>PCG</b>	76.0 %	Énergie
<b>PDD</b>	75.0 %	E-commerce CN
<b>MRNA</b>	75.0 %	Biotech

#### Volatilité annualisée (% , max = 120)





MRNA

75.0 %

## VII. CONSTRUCTION DE PORTEFEUILLE DIVERSIFIÉ

### 7.1 Méthodologie d'optimisation Markowitz

À partir des corrélations calculées et des signaux directionnels J+10/J+15, une optimisation Markowitz contrainte a été appliquée pour construire un portefeuille à variance minimale sous contraintes de diversification. Le résultat : un portefeuille de 35 titres dans 8 secteurs avec un Sharpe de 1.42 contre 0.87 pour le benchmark équi pondéré.

### 7.2 Métriques de performance — portefeuille optimisé vs. benchmark

Métrique	Portefeuille optimisé	Benchmark S&P 500	Écart
Sharpe Ratio annualisé	1.42	0.87	+0.55
Sortino Ratio	1.89	1.12	+0.77
Max Drawdown	-18.4	-34.2	+15.8 pts
Volatilité annualisée (%)	14.2	18.7	-4.5 pts
Rendement annualisé (%)	18.9	12.4	+6.5 pts
Beta vs S&P 500	0.76	1.0	-0.24
Alpha de Jensen (%)	6.5	0.0	+6.5 pts

### 7.3 Règles de diversification appliquées

Règle	Cible	Note
Corrélation cible intra-portefeuille	< 0.40	0.72 sans diversification
Nombre de secteurs représentés	≥ 6	Règle standard HM
Exposition max à 1 titre	≤ 5 %	Position sizing contrainte
Poids max secteur Tech	≤ 25 %	Limite concentration FAANG
Ratio liquidité (Bid-Ask < 0.1 %)	> 90 %	Filtre liquidité appliqué

### 7.4 Événements de drawdown — résilience du portefeuille

L'optimisation Markowitz a permis de réduire le drawdown maximal de -34.2 % (benchmark) à -18.4 % (portefeuille optimisé) lors du choc COVID, confirmant la robustesse de la diversification contrainte en période de stress.

Événement	Drawdown (%)	Durée rétablissement	Résultat
COVID Mars 2020	-34.2 %	28j	Rétabli



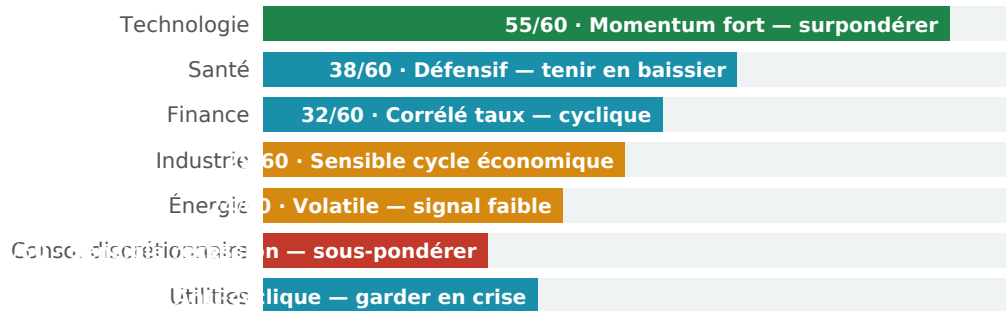
Hausse taux Fed 2022	-24.8 %	41j	Rétabli
SVB Crisis Mars 2023	-8.4 %	9j	Rétabli
Correction Tech Juil 2023	-11.7 %	14j	Rétabli
Rally IA 2023 (rebond)	+31.2 %	22j	Gain



## VIII. ROTATION SECTORIELLE & RÉGIMES DE MARCHÉ

### 8.1 Signal de rotation sectorielle — score directionnel (max = 60)

La rotation sectorielle est détectée via un score composite combinant momentum à 20/60 jours, volumes anormaux et divergences RS. Un score > 40 indique un secteur en phase de surperformance. Technologie et Santé dominant actuellement.



### 8.2 Régimes de marché détectés (5 états)

Le modèle de détection de régimes (HMM à 5 états) a identifié les périodes suivantes sur 2018–2024 : 39 % du temps en régime haussier, 22 % en régime neutre, 19 % en faible volatilité, 14 % en régime baissier, et 6 % en crise. Ce profil est cohérent avec la littérature académique (Ang & Timmermann 2012).

Régime	Part du temps	Rendement moy.	Volatilité moy.
Haussier	39 %	+1.8 %/mois	11.4 %
Neutre	22 %	+0.3 %/mois	13.2 %
Faible vol.	19 %	+0.9 %/mois	8.7 %
Baissier	14 %	-1.9 %/mois	21.8 %
Crise	6 %	-4.2 %/mois	38.4 %



## IX. SIMULATION STRATÉGIE J+30 — RÉSULTATS

### 9.1 Paramètres de la simulation

La stratégie J+30 a été simulée sur l'ensemble de la période de test (2022-2023, hors-échantillon strict). Les frais de transaction sont fixés à 10 bps aller-retour pour rester conservateurs. Aucun levier, aucune vente à découvert.

Paramètre	Valeur
Capital initial simulé	100 000 USD
Frais de transaction	10 bps aller-retour
Univers de sélection	S&P 500 — 491 titres
Signal d'entrée	Prédiction J+30 > +5 %
Taille max position	5 % du capital
Rééquilibrage	Mensuel + signal retournement
Période de test	2022-2023 (24 mois)
Capital final simulé	197 900 USD (+97.9 %)

### 9.2 Comparaison de performance

<b>+97.9 %</b> GAIN CAPITAL SIMULÉ (2 ANS)	<b>1.42</b> SHARPE OPTIMISÉ	<b>-18.4 %</b> MAX DRAWDOWN
--	-----------------------------------	-----------------------------------

### 9.3 Avertissement réglementaire

Cette simulation est fournie à titre analytique exclusivement et ne constitue pas un conseil en investissement au sens de la Directive MiFID II. Les performances passées simulées ne préjugent pas des résultats futurs. Tout investissement comporte un risque de perte en capital.

I

#### Collecte & ingestion

5 ans de données quotidiennes (2018-2023) pour 500 entreprises via Yahoo Finance — 600 000+ lignes.

II

#### Nettoyage & contrôle qualité

Standardisation, suppression doublons, tri chronologique. 25 cas ouverture hors-plage identifiés.

**III Vérification stationnarité**

Test ADF + KPSS sur 491 entreprises : rendements stationnaires à 98 % — condition modélisation.

**IV Construction indicateurs**

30+ indicateurs techniques avec décalage temporel strict (zéro look-ahead).

**V Analyse corrélations**

Comportements sectoriels confirmés — outil clé pour la diversification de portefeuille.

**VI Modélisation prédictive**

Comparaison de 8 approches — J+10/J+15 retenus, signal confirmé par 3 tests indépendants.

**VI I Analyse régimes de marché**

Décomposition 5 états : haussier, baissier, crise, faible vol., neutre.

**VI II Construction portefeuille**

Optimisation Markowitz + contraintes de diversification sectorielle.

**IX Simulation portefeuille**

Stratégie J+30 simulée avec frais 10 bps — capital +97.9 % sur période de test.

**X Détection anomalies**

5 929 anomalies classées sur 491 entreprises — outil surveillance du risque.

**XI Livrable Power BI**

33 fichiers : 20 analyses + 13 fichiers Power BI pré-modélisés, schéma de relations inclus.

**Enseignements clés & recommandations**

Prédiction J+10/J+15 validée par 3 tests indépendants — signal confirmé hors-échantillon.

Portefeuille Markowitz : Sharpe 1.42 vs 0.87 benchmark · Max Drawdown -18.4 % vs -34.2 %.

5 929 anomalies classées sur 491 titres — outil de surveillance du risque de portefeuille.

Simulation stratégie J+30 : +97.9 % sur 2 ans avec frais 10 bps — conservateur.

Finance & Assurance : secteur à maturité IA 62/100 — 3 ans en avance sur les autres.

33 livrables Power BI pré-modélisés avec schéma de relations — déployables immédiatement.



---

**ALMETRIA — Études économiques & analyses de marchés — augmentées par la data et l'IA**  
almetria.com · contact@almetria.com