

# RAPPORT : ANALYSE EXPLORATOIRE DE DONNEES

*Projet : Analyse temporelle et exploratoire du Bitcoin*

Cours : 8PRO408 - Outils de programmation pour la science des données

Chargé de Cours : HN Doukaga

Ayad-Zeddam Malik / Lopez Valentin / Abgarov Artur /  
Razafindratsima Nathan

07/12/2025

## INTRODUCTION

Le Bitcoin est la première cryptomonnaie décentralisée et représente aujourd’hui un actif financier majeur sur les marchés numériques. Son évolution rapide, sa forte volatilité et son adoption croissante en font un sujet d’étude particulièrement intéressant en science des données. L’objectif de ce projet est de réaliser une analyse exploratoire et temporelle des données historiques du Bitcoin, afin de mieux comprendre l’évolution de son prix, la dynamique de son volume d’échange ainsi que les périodes de forte instabilité du marché. Cette étude vise également à poser les bases pour de futures analyses prédictives.

## Description des données

Le jeu de données utilisé contient plus de 7,3 millions de lignes et six variables : *Timestamp*, *Open*, *High*, *Low*, *Close* et *Volume*. Chaque ligne correspond à une observation du marché du Bitcoin à une minute donnée. La période couverte s'étend de janvier 2012 jusqu'à aujourd'hui. Cette granularité élevée permet une analyse fine mais implique également un niveau de bruit important dans les observations. Ensuite les données ont été vérifiées afin d'identifier d'éventuelles valeurs manquantes, doublons ou incohérences temporelles. Cette étape n'a pas révélé d'anomalies majeures, ce qui a permis de travailler directement sur les données brutes. En raison à la fréquence très élevée des observations, plusieurs séries agrégées ont ensuite été construites à différentes échelles temporelles : horaire, journalière et mensuelle. Ces agrégations permettent de lisser partiellement le bruit des données et de faciliter l'analyse des tendances de fond du marché.

## Méthodologie

L’analyse repose sur une approche exploratoire et temporelle. Des statistiques descriptives ont tout d’abord été calculées afin d’étudier les distributions du prix de clôture et du volume d’échange. Des visualisations univariées (histogrammes, courbes de densité, boxplots) ont permis d’identifier l’asymétrie des distributions et la présence de valeurs extrêmes. Une analyse bivariée a ensuite été menée afin d’étudier la relation entre le prix et le volume, à l’aide de nuages de points et d’un coefficient de corrélation. Enfin, une analyse multivariée basée sur une matrice de corrélation et une heatmap a permis de vérifier la cohérence interne des variables de prix. De plus, on a également intégré des outils d’analyse temporelle, tels que les moyennes mobiles et la volatilité

glissante, afin d'examiner l'évolution des tendances et du niveau de risque du Bitcoin au cours du temps.

## Résultats principaux

Les résultats montrent que le prix du Bitcoin présente une distribution fortement asymétrique, avec une majorité de valeurs relativement faibles et quelques pics très élevés correspondant aux périodes de *bull market*. Cette configuration reflète la nature hautement spéculative de cet actif. Le volume d'échange est caractérisé par une grande majorité de faibles valeurs, en raison de la fréquence à la minute, ainsi que par quelques pics très importants correspondant à des périodes d'activité intense du marché. Les boxplots confirment la présence de nombreuses valeurs aberrantes. L'analyse conjointe du prix et du volume met en évidence une corrélation linéaire faible et légèrement négative, indiquant que le volume seul ne permet pas d'expliquer les variations du prix. La volatilité glissante révèle quant à elle des phases de forte instabilité, notamment lors des périodes de hausse rapide ou de chutes brutales du prix. Les moyennes mobiles permettent enfin de visualiser les grandes tendances à court et moyen terme.

## Limite de l'étude

Cette étude repose sur une analyse exploratoire basée sur des statistiques descriptives et des relations linéaires simples. Elle ne permet pas d'identifier de relations causales ni de modéliser précisément l'évolution future du prix du Bitcoin. De plus, l'analyse ne tient pas compte de facteurs externes tels que les données macro-économiques, les annonces réglementaires ou le contexte géopolitique, qui influencent fortement le marché des cryptomonnaies. Enfin, la fréquence élevée des données engendre un bruit important dans les observations.

## Conclusion et perspectives

Cette étude repose sur une analyse exploratoire basée sur des statistiques descriptives et des relations linéaires simples. Elle ne permet pas d'identifier de relations causales ni de modéliser précisément l'évolution future du prix du Bitcoin. De plus, l'analyse ne tient pas compte de facteurs externes tels que les données macro-économiques, les annonces réglementaires ou le contexte géopolitique, qui influencent fortement le marché des cryptomonnaies. Enfin, la fréquence élevée des données engendre un bruit important dans les observations.