University of Naples Federico II

Political Sciences Department

Methods for fintech and artificial intelligence in Finance

Erasmus+ Blended Intensive Program (BIP)

Relationship between Stablecoins and other cryptocurrencies for crypto asset diversification using PCA

Supervisor:
Professor Maria Iannario

Author:
Azeezat Mosunmade Mustapha

Summer 2024

Contents

1 Introduction

In recent years, Cryptocurrency has emerged as an interesting financial asset due to its sharp price fluctuations, attracting several individuals and investors. Among these digital assets are stablecoins, cryptocurrencies designed to maintain a stable value relative to a fiat currency or Gold, which is crucial in facilitating transactions and providing liquidity (Coinbase, 2024). Due to its low volatility, it has become attractive to investors with low-risk appetite and investors looking to create stability in their portfolios (Enacache, 2024). Despite its major advantage of being less volatile, some stablecoins have deviated in recent years, with prices falling far below their pegged currencies or assets.

This report uses Correlation and Principal component analysis (PCA) to explore the relationship between stablecoins and other top cryptocurrencies. Analyzing percentage returns and identifying the key drivers of variation in the price movements of these assets, The goal is to gain insights into the relationship of these assets that can help make informed risk management and investment decisions in the cryptocurrency market. The details and findings of this analysis are explained in the following sections.

2 Problem statement

The data used in this report is an extract of the daily historical price data for Cryptocurrencies with the top market capitalization from Cryptocompare(CCData) API (Data, 2021), CCData provides institutional-grade digital asset data. The data extracted was the 1-year daily historical price of the top 10 cryptocurrencies and 8 stablecoins by market capitalization supported by CCData from 09-22-2023 to 09-21-2024, making 366 observations. The cryptocurrencies included are Bitcoin(BTC), Ethereum (ETH), BNB, Solana (SOL), XRP, AVAX, Dogecoin (DOGE), Cardano (ADA), Toncoin (TON), and Tron (TRX) while the stablecoins are USDT, USDC, DAI, USTC, PYUSD, TUSD, USDP and GUSD. There was no data for the price of TON for the first 25 days for which the rows were discarded. The price was used to calculate the percentage daily change in price to get the daily return. The total number of observations used in this analysis was 340. The calculated return is then scaled to ensure that PCA captures variance from a common reference point. The main objective is to gain insights into the relationship between these assets, identify the key drivers of variation in the price changes, and investigate whether diversifying across both stablecoins and cryptocurrencies can reduce risk and improve portfolio efficiency.

3 Statistical methods

3.1 Correlation

Correlation measures the strength and direction of the relationship between two quantitative variables. The value is called the correlation coefficient and is represented by r. The value of r ranges from -1 to +1 (Bluman, 2012). This report used the Pearson correlation(r), which measures the linear relationship between two variables. The Pearson correlation coefficient between 2 variables x and y with n number of observations is given below:

$$r = \frac{n(\sum_{i=1}^{n} x_i y_i) - (\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i)}{\sqrt{[n(\sum_{i=1}^{n} x_i^2) - (\sum_{i=1}^{n} x_i)^2][n(\sum_{i=1}^{n} y_i^2) - (\sum_{i=1}^{n} y_i)^2]}}$$

where x_i = values of x variables and y_i = values of y variables. (Bluman, 2012)

A positive r-value shows that there is a positive linear relationship between the two variables. That is, as one variable increases, the other also increases. A negative r-value suggests that as one variable increases, the other reduces. 0 r-value indicates that there is no relationship between the variables. The closer r is to 1 or -1, the stronger the relationship between the measured variables.

3.2 Principal component analysis

Principal component analysis is a dimension reduction method that transforms a data set with correlated variables into a set of uncorrelated principal components by capturing the variation in the data. Principal components(PCs) of a data set is a sequence of best linear approximation of the variables in the data set, which captures the most important patterns or structure in the data (James et al., 2021, p. 375). The first PC captures the highest variation in the data followed by the second and the last PC captures the least variation.

The main goal of the principal component analysis is to find a matrix A that transforms the initial vector of variables X into a new one such that $Y = AX$ where A is the matrix with the eigenvectors or loading of the covariance matrix of X (James et al., 2021, p. 376).

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ a_{p1} & \cdots & \cdots & a_{pp} \end{pmatrix}$$

a_{ij} indicates much the j_{th} variable contributes to the i_{th} principal component

p is the number of variables

All analysis was done with Python programming language (Python Software Foundation.), version 3.9.18.

4 Statistical analysis

4.1 Descriptive analysis

To explore the linear relationship between the returns of the Cryptocurrencies and Stablecoins, Correlation was used, the heatmap in Figure 1 shows the relationships between these variables with the strength and the direction of the relationships indicated with colours. The colour red indicates a positive relationship, while the colour blue indicates a negative relationship. The chart shows that there are positive relationships between the cryptocurrencies that are not stable coins with values ranging from 0.80 to 0.22, this means that as one price increases, the other also increases. The relationship between the Stablecoins are mostly weak and in opposite directions.

Considering the relationships between the Cryptocurrencies that are Stablecoins and not Stablecoins, their relationships are weak. The stablecoin USTC has the highest positive correlation value of 0.25 with ADA, AVA, ETH and SOL.

4.2 Principal Component Analysis

To analyze the relationship between these crypto assets, PCA is used to decompose the data into principal components and understand the variation in the return data.

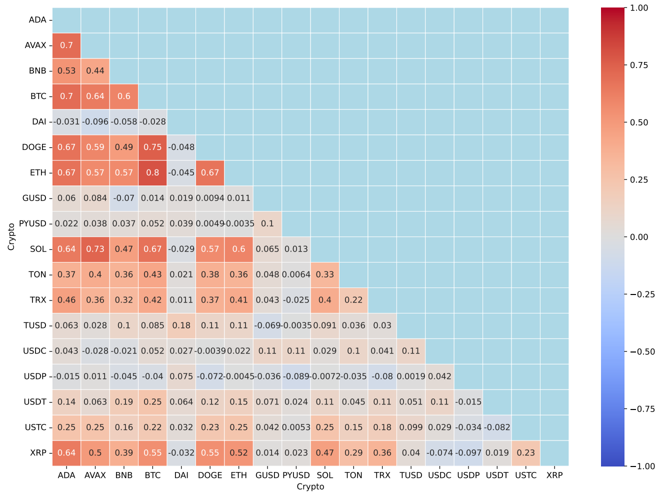

Figure 1: Heat map of the correlations between the Cryptocurrencies and the Stablecoins

Figure 2 shows the bar chart for the variation percentage explained by each principal component.

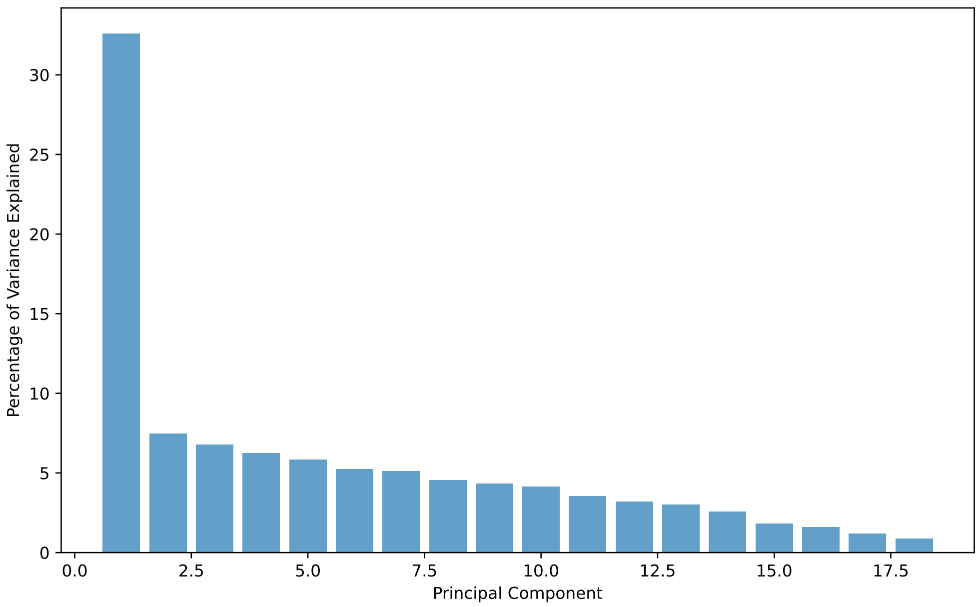

Figure 2: Percentage variance explained by each component

Figure 2 above shows that the first component explains 33% of the variation in the data while the second and third explain only 7% while the other components indicate lower values. This means that only the first three PCAs are important for explaining the relationships in the data.

To visualise how much these Cryptocurrencies contribute to each principal component(PC), a heat map is used to visualize the eigenvectors. Figure 3 shows the heatmap of the weight of each cryptocurrency against the 18 principal components. Colors indicate the intensity of the weight of each cryptocurrency, strong positive red indicates that the variable has a high positive contribution to the corresponding principal component, and strong negative blue has a high negative contribution to the PC. In contrast, light colours or values close to zero indicate minimal contribution to the PC.

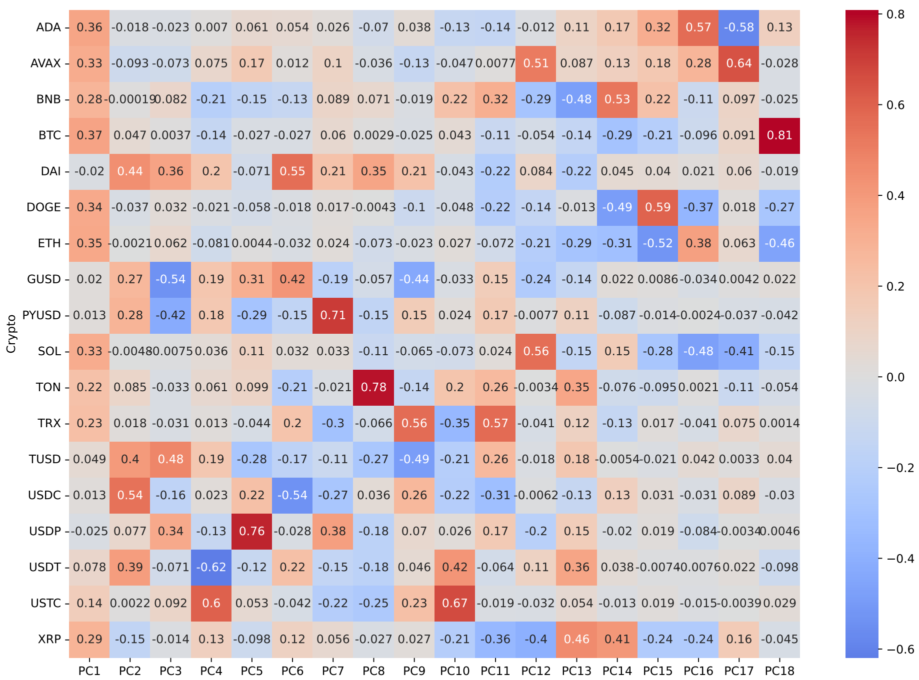

Figure 3: Principal component eigenvalues for Cryptocurrencies and Stablecoins

From figure 3, the first principal component(PC1), which contributes 33% of the total information, shows that the cryptocurrencies that are not stable coins have the highest positive weights, with Bitcoin (BTC) having weight of 0.37 followed by ADA with the weight of 0.36 and the least being 0.22 from TON. This shows that cryptocurrencies that are not Stablecoins move in the same direction and drive most of the variations in PC1. Stablecoins have negative values and are close to zero indicating that they contribute less to the PC1, and move in opposite directions with the return of other cryptocurrencies. PC2 and PC3 from 2 contribute 7% each to the total variation in the data.

Unlike PC1, Cryptocurrencies that are not stable coins have lower weights and contribute less to this component. In contrast, most Stablecoins have high positive weights except USTC and USDP, which weigh almost zero. This shows that PC2 captures a factor that differentiates these Stablecoins from the other cryptocurrencies even though the contribution to the total variation is low. The other principal components from PC4 to PC11 hardly differentiate the Stablecoins from the other Cryptocurrencies by the weight of the eigenvalues.

5 Summary

This report focuses on the relationship between the top 18 cryptocurrencies by market capitalization using Principal Component Analysis (PCA) to make informed decisions on portfolio diversification. The data for this analysis was extracted from the Cryptocompare (CCData) API, which provides institutional-grade digital asset data. The data was a 1-year daily historical price of these cryptocurrencies. The historical price was used to calculate the percentage of daily return on investment, and these returns were then standardized to ensure that all assets were on the same scale. The total number of cryptocurrencies considered in this report was 18, and the number of observations was 340. This report aims to examine the relationships and identify the key drivers of variation in the price movements of these assets and distinguish between the cryptocurrencies that are Stablecoins and highly volatile cryptocurrencies.

Correlation and heatmap were used to examine and visualize the relationship between the returns of these crypto assets. Cryptocurrencies that are not Stablecoins are more positively correlated with each other, indicating that the prices of these assets are going in the same direction. In contrast, Stablecoins are either negatively correlated to other cryptocurrencies and other Stablecoins, or their correlations are close to zero. Principal component analysis was used to capture the complex relationship between these assets and identify the key drivers of market variability. The PCA results reveal that the first principal component (PC1) explains a significant portion of the variance (33%); this component was able to differentiate between these two sets of cryptocurrencies, with Stablecoins contributing less to the variation in this component, while the others have higher positive contributions to the variations in the component.

The second and third components (PC2 and PC3) explain smaller variations and were able to capture fewer differences between the assets. These components may represent several factors but contribute far less to the overall market changes than PC1. This analysis shows that stablecoins and Cryptocurrencies show different price behaviors, with stablecoins providing potential diversification benefits due to their lower volatility and weaker Correlation with the broader market trends captured by PC1. For further exploration of the data, the assets for which each stablecoin is pegged can be considered, and further analysis, such as cluster analysis, can be done.

Bibliography

Allan G. Bluman. *Elementary statistics : a step by step approach.* The McGraw-Hill Companies, Inc., 1221 Avenue of the Americas, New York, NY 10020, eightht edition, 2012.

Coinbase. What is a stablecoin? 2024. URL: https://www.coinbase.com/en-de/learn/crypto-basics/what-is-a-stablecoin. (accessed 21 September 2024).

Crptocompare Data. Daily pair ohlcv. 2021. URL: https://developers.cryptocompare.com/documentation/legacy/Historical/dataHistoday (visited on 21st of september 2024).

Camelia Enacache. Stablecoin investment in 2024: Smooth sailing or risky venture? 2024. URL: https://tangem.com/en/blog/post/stablecoin-investment/ Last Reviewed: 11/04/2024. (accessed 21 September 2024).

Gareth James, Daniela Witten, Trevor Hastie, and Robert Tibshirani. *An Introduction to Statistical Learning with Applications in R.* Springer New York, 2 edition, 2021. doi: /10.1007/978-1-0716-1418-1.

Python Software Foundation. *Python Language Reference.* URL `https://www.python.org/`.