# Microsegmentation for better diversification in the Tunisian Stock Exchange Market (BVMT)

Fayad Ali Banna *        Louay Boukhris†        Ahmed Rebai ‡

September 20, 2021

**Abstract**

Historical financial data, related to companies listed on the Tunisian stock exchange, has been collected from the Tunis Stock Exchange Market (BVMT) during the first 6 months of the year 2021. Based on a statistical approach, several features were extracted from the data. The dimensionality reduction algorithm U-MAP was applied to these variables. BVMT presented in june 2021, a classification of the tunisian stock market paritioned into classes. These classes were introduced after analysing their financial data both on the stock market itself and some internal data. These results were obtained in the first plan by the U-MAP algorithm.

**Keywords:** Unsupervised Learning, Tunisian Stock Market, Machine Learning, Asset Management

## 1 Introduction

When investing money in the stock market, rational investors try to avoid the risk of concentration and therefore diversify their investment over several securities. In this article, we continue the work that was done on diversification techniques in the Tunisian market in the article (1). It is important to mention that our study focuses on financial data from the Tunis Stock Exchange and that several factors must be considered while reading the previous article and this article. Indeed, the Tunisian financial market is a new market, hampered by strict banking regulations that put several obstacles to investors, such as the law n° 1994-25 of 7 February 1994 modifying the law n°1967-51 of 7 December 1967 regulating the banking profession. On the other hand, this market is located in a developing country ravaged by a deep structural economic crisis accentuated by a political crisis for 10 years and by the current Covid-19 crisis since March 2020. Therefore, the astute reader of this article is advised to understand the difficulties encountered during the fine analysis of the data and the interpretation of the results. Since the publication of the last article (2), the Tunis Stock Exchange authorities have published in June 2021 a classification of the main market for companies according to four criteria:

- The market capitalisation of the company is calculated as a daily average of the capitalisations between 1 September and 30 November of each year.

- The IPO price for newly listed companies.

---

*fayad.alibanna@ept.u-carthage.tn : Quant intern @Value Digital Services and engineering student in the last year @ Ecole Polytechnique de Tunisie

†louay.boukhris@value.com.tn : Data scientist @ Value Digital Services

‡ahmed.rebai@value.com.tn : Corresponding author (Senior Data Engineer @ Value Digital Services)

- Regulatory criteria related to transparency in relation to financial reporting.

- Criteria related to an excessive risk likely to durably disrupt the financial stability of a company or to compromise the proper functioning of the Tunisian market.

From now on, the Tunis Stock Exchange is grouped into 3 groups A, B and S:

- Group A contains companies with a market capitalisation of 200 million dinars or more.

- Group B includes issuers with a market capitalisation of less than 200 million dinars.

- Group S contains companies under supervision that are not in compliance with their financial reporting obligations and are affected by events that may have a high-risk exposure to market events.

| N° | Compartiment A | | N° | Compartiment B | | N° | Compartiment S |
|---|---|---|---|---|---|---|---|
| 1 | SFBT | | 1 | SOTUVER | | 1 | TUNISAIR |
| 2 | POULINA GP HOLDING | | 2 | WIFACK INT BANK | | 2 | UADH |
| 3 | BIAT | | 3 | TPR | | 3 | SOPAT |
| 4 | BT | | 4 | ARTES | | 4 | TAWASOL GP HOLDING |
| 5 | ATTIJARI BANK | | 5 | TUNIS RE | | 5 | SERVICOM |
| 6 | SAH | | 6 | BH ASSURANCE | | 6 | CELLCOM |
| 7 | DELICE HOLDING | | 7 | CITY CARS | | 7 | OFFICEPLAST |
| 8 | AMEN BANK | | 8 | SOTIPAPIER | | 8 | ELECTROSTAR |
| 9 | ONE TECH HOLDING | | 9 | MONOPRIX | | 9 | GIF-FILTER |
| 10 | BNA | | 10 | ICF | | 10 | SIPHAT |
| 11 | CARTHAGE CEMENT | | 11 | AIR LIQUDE TSIE | | 11 | AMS |
| 12 | STB | | 12 | TUNISIE LEASING ET FACTORING | | 12 | STIP |
| 13 | UIB | | 13 | TELNET HOLDING | | 13 | AETECH |
| 14 | UBCI | | 14 | SOTEMAIL | | 14 | MIP |
| 15 | BH BANK | | 15 | ADWYA | | | |
| 16 | ASTREE | | 16 | CIL | | | |
| 17 | ENNAKL AUTOMOBILES | | 17 | LAND OR | | | |
| 18 | UNIMED | | 18 | ALKIMIA | | | |
| 19 | ATB | | 19 | CIMENTS DE BIZERTE | | | |
| 20 | STAR | | 20 | ATL | | | |
| 21 | MAGASIN GENERAL | | 21 | ASSAD | | | |
| 22 | ASSURANCES MAGHREBIA | | 22 | BEST LEASE | | | |
| 23 | SPDIT - SICAF | | 23 | SOTRAPIL | | | |
| 24 | EURO-CYCLES | | 24 | SIAME | | | |
| | | | 25 | PLACEMENT TSIE-SICAF | | | |
| | | | 26 | SIMPAR | | | |
| | | | 27 | SOTUMAG | | | |
| | | | 28 | SOMOCER | | | |
| | | | 29 | HANNIBAL LEASE | | | |
| | | | 30 | CEREALIS | | | |
| | | | 31 | MPBS | | | |
| | | | 32 | SITS | | | |
| | | | 33 | ATTIJARI LEASING | | | |
| | | | 34 | SOTETEL | | | |
| | | | 35 | SANIMED | | | |
| | | | 36 | NEW BODY LINE | | | |
| | | | 37 | ATELIER MEUBLE INT | | | |
| | | | 38 | ESSOUKNA | | | |
| | | | 39 | BH LEASING | | | |
| | | | 40 | BTE (ADP) | | | |
| | | | 41 | TUNINVEST-SICAR | | | |

Figure 1: The figure shows the lists of companies listed on the Tunis Stock Exchange and classified according to the above-mentioned criteria according to the BVMT communiqué of June 2021.

In order to calibrate our analysis, we will start with a verification of our algorithms by showing that the data collected and processed by us are able to detect real behaviours recorded by the trade specialists. In the first part of this article, we propose to reproduce this segmentation using fully public data.

---

[0]Nowadays trading carries a high level of risk, and may not be suitable for all investors. Since we are not economists, arbitrageurs, market analysts, market bulls, brokers, or even daily traders, we disown any responsibility for any errors or misunderstandings caused by the given strategies presented in this paper. That said, we are a team of data scientists with an affinity for market finance who aim to democratize the application of mathematical theories and machine learning models to the Tunisian financial market.

# 2 Microsegmentation of the Tunisian market

## 2.1 Concept of micro-segmentation

The practice of micro-segmentation consists of dividing the groups (clusters) obtained after a segmentation. This allows a better understanding of the data. This practice is widespread in marketing, digital marketing and networking. There is a difference between segmentation and micro-segmentation. Segmentation allows you to characterise the data and generate characteristics that describe the groups of companies.

## 2.2 Data collection and extraction of explanatory variables

For our issue we collected one types of data. The structured data corresponds to the quotations of the Tunisian stock market.This digital data contains the following indicators

- **Stock market index:** A stock market index is an important indicator that determines the performance of a stock market. Composed of a group of shares, it represents a market, a sector of activity. It shows the trend of the economy and of the largest companies in the index. In Tunisia, there are two stock market indices: Tunindex which represents all listed Tunisian companies, and Tunindex20 which represents the twenty most liquid listed companies with the largest market capitalization.

- **Asset price:** The price of an asset at any point in time(daily basis)is the price at which the greatest number of securities can be traded at that point in time.

- **Opening price:** The opening price of an asset is the first price displayed at the beginning of a trading day. Generally calculated using the fixing technique, it is calculated by comparing buy and sell orders to extract an appropriate equilibrium price in the event that the maximum number of assets are traded.

- **Closing price:** The Closing Price of an Asset is the last price fixed for that Asset on the day of listing. It is considered to be the reference price for tax purposes and is one of the key pieces of information that enables investors to value their investments on the stock exchange.

- **Volume:** The volume represents the number of securities traded over a certain period of time (one trading day in our case).

The figure below shows the correlation between the 4 prices of a stock.

From this important financial time series, the other explanatory variables (or features) are constructed:

- **Skewness**: Skewness refers to a distortion or asymmetry that deviates from the symmetrical bell curve, or normal distribution, in a set of data. It is assumed to be bent whether the curve is moved to the left or to the right. As a representation of the degree to which a given distribution differs from a normal distribution, skewness can be quantified.

- **Kurtosis**: Kurtosis, like skewness, is a statistical measure used to characterize distribution. Although skewness distinguishes extreme values in one tail from the other tail, in either tail, kurtosis measures extreme values. Large kurtosis distributions exhibit tail data exceeding the normal distribution tails. Low kurtosis distributions exhibit tail data that is usually less extreme than the normal distribution tails.

- **Beta**: This coefficient measures the volatility of a security in relation to market fluctuations and therefore measures non-diversifiable risk, which makes it complementary to diversification methods aimed at minimizing diversifiable risk.
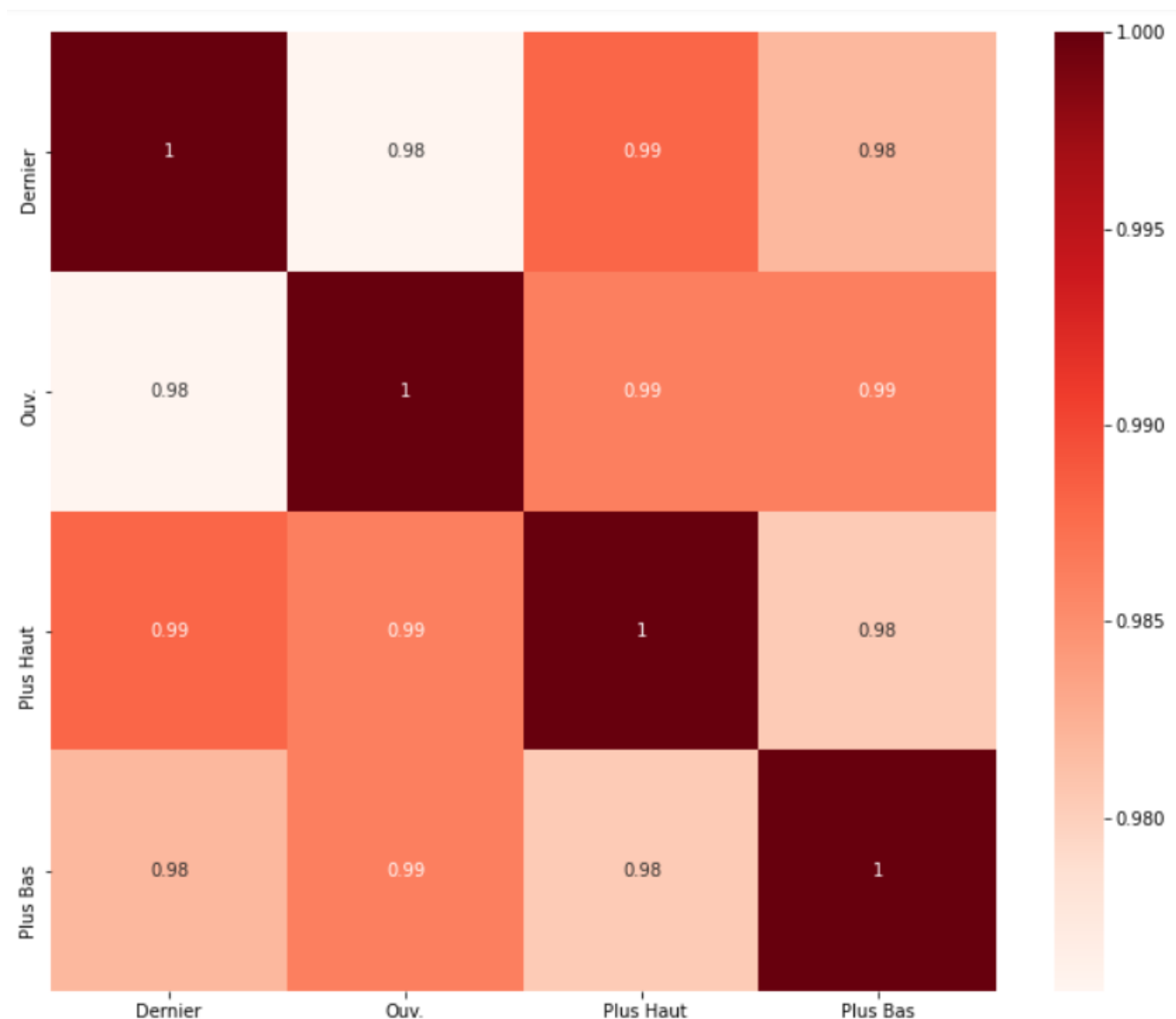
3

Figure 2: Correlation between the 4 prices of a given stock.

Figure 3 shows the correlation between the variables created from the closing price. These variables are going to be clustered using kmeans and getting the new vectors by Umap.

## 2.3   Dimension reduction using the U-MAP algorithm

UMAP (Uniform Manifold Approximation and Projection) is a new dimension reduction technique based on manifold learning. The theoretical underpinning for UMAP is based on Riemannian geometry and algebraic topology. UMAP adopts the idea of tSNE in general, but adds a number of enhancements such as a second cost function and the lack of normalization of high- and low-dimensional probabilities.

In our case after applying Umap on our data, and having Umap initializer return two components for easier visualisation we get the information. At a first glance it might seem not clear but if we divide the companies by these hand drawn borders we easily reproduce the results shown in figure 1. It is a confirmation of the classification published by the Tunis Stock Exchange authorities on Friday 28 June 2021
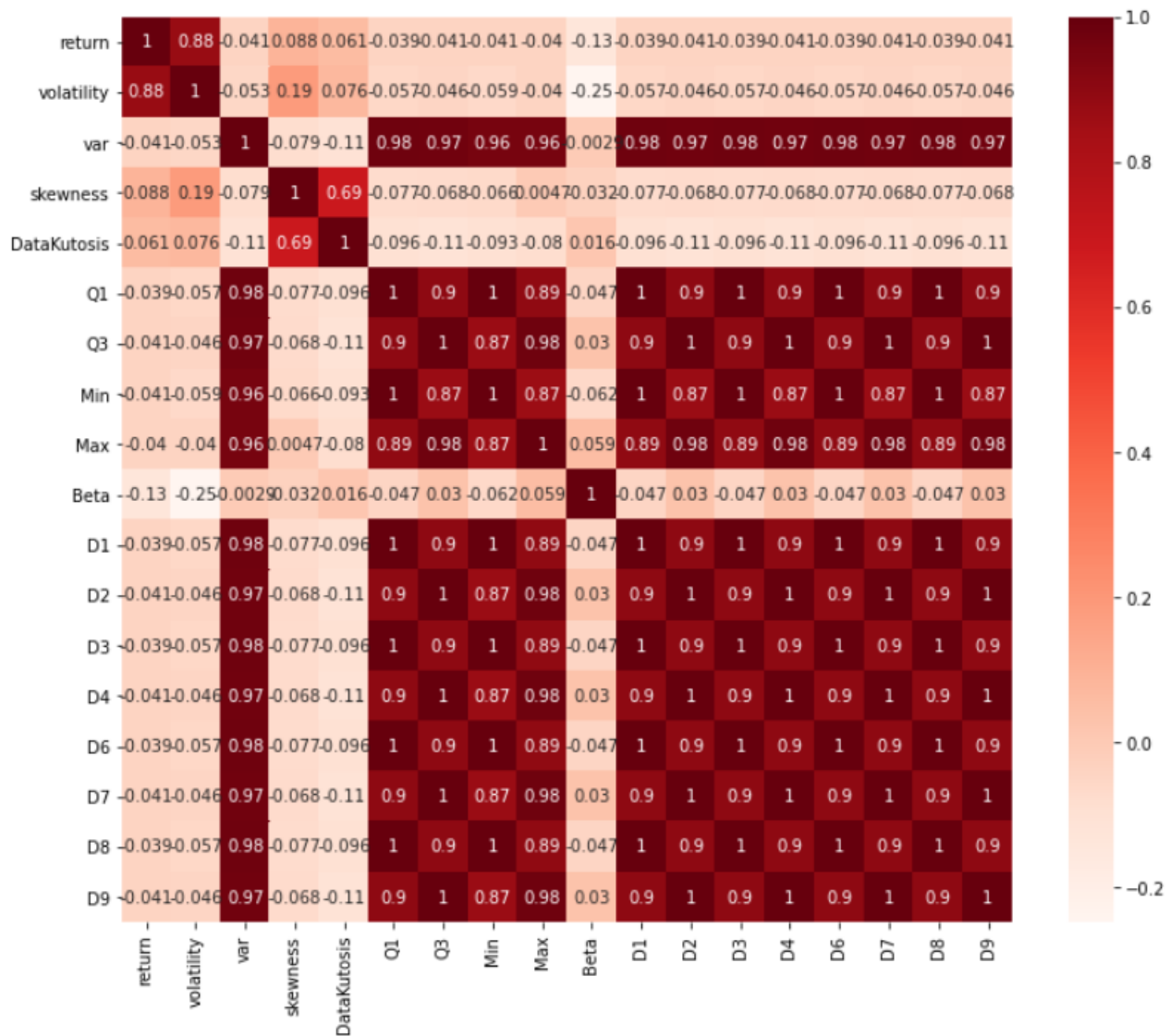
Figure 3: The figure shows the correlation between the variables created from the closing price.

# 3  Conclusion

This article wants to show that it is possible to apply machine learning algorithms in the Tunisian market and that these algorithms can detect the real behavior of companies, then we want to rely on micro segmentation to achieve a better diversification.

Finally, this rapid communication constitutes the beginning of a study in progress within the "Advanced Analytics team of Value Digital Services", on the application of trading strategies on the Tunisian market where we will try to answer fundamental questions such as Is machine learning effective in a financial context? Technical analysis or fundamental analysis? Can econometric models surpass financial machine learning algorithms?
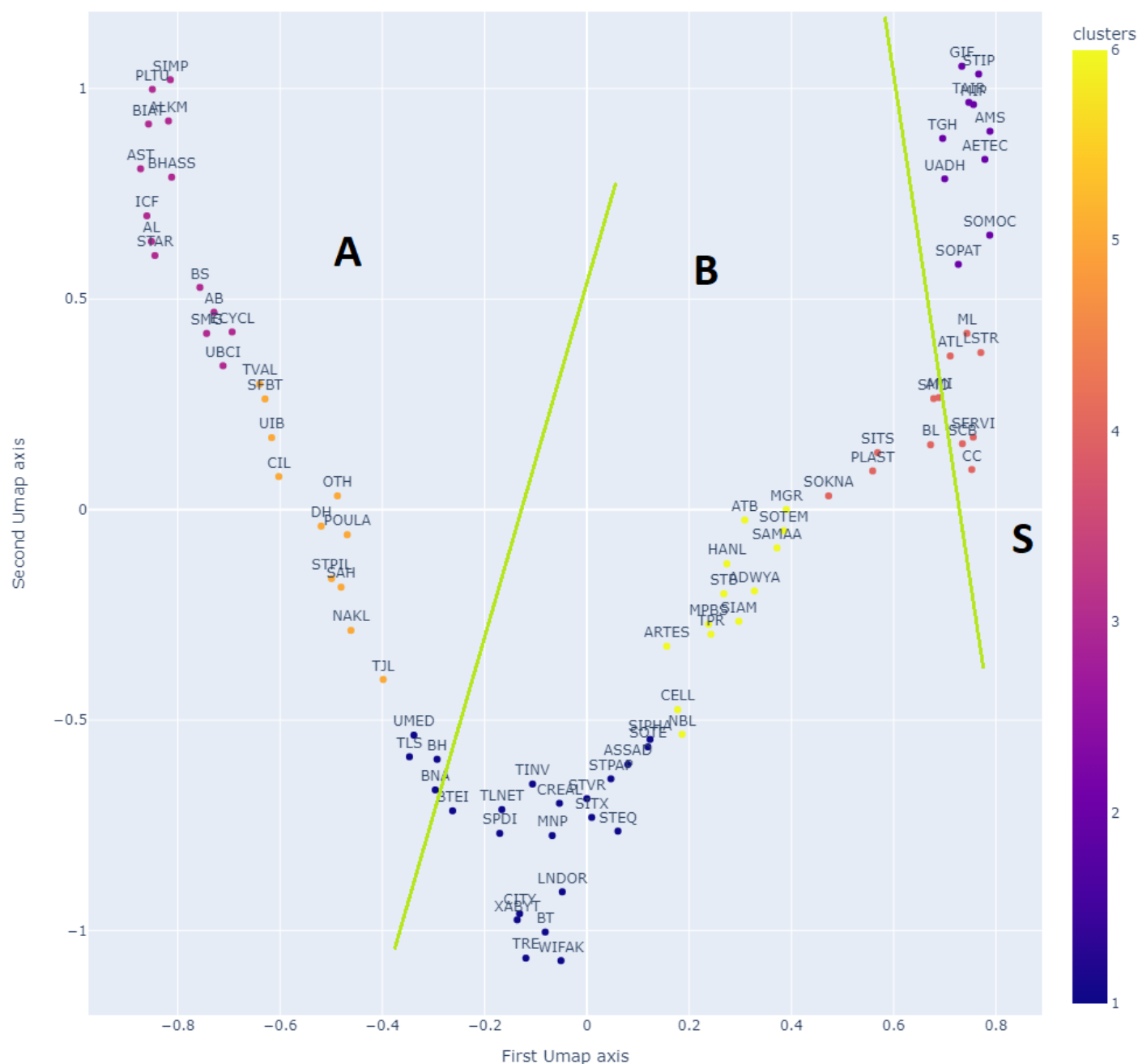
Figure 4: The figure shows the projection of the stocks on the plan obtained by the U-MAP algorithm.The figure shows the projection of the stocks on the plan obtained by the U-MAP algorithm with a hand drawn borders. It is clear that the U-MAP algorithm can reproduce the results shown in the figure number 1.

# References

[1] Rebai, Ahmed and Boukhris, Louay and Lotfi, Ncib and Ben Lasmar, Mohamed Anis, Unsupervised Learning Diversification Applied on the Tunisian Stock Market Before and During the

COVID-19 Crisis (June 5, 2021).

[2] http://www.bvmt.com.tn/sites/default/files/bulletin/pdf/bull20210625.pdf

[3] Micro Segmentation: Today's Success Formulae Nitin Singh Sikarwar and Dinesh Verma International Journal of Operations Management and Services.Volume 2, Number 1 (2012).

[4] Masashi Sugiyama, Motoaki Kawanabe, Machine Learning in Non-Stationary Environments: Introduction to Covariate Shift Adaptation - Adaptive Computation and Machine Learning series (Hardback). .

[5] Daniel C. Funk , (2002),"Consumer-Based Marketing: The Use of Micro-Segmentation Strategies for Understanding Sport Consumption", International Journal of Sports Marketing and Sponsorship, Vol. 4 Iss 3 pp. 39 - 64.

[6] Micro-Segmentation and Personalization in Information Systems in the Financial Service Industry May 2015.

[7] Procédure de micro segmentation du marché industriel : application au service téléphonique interurbain EMMANUEL J. CHÉRON.

[8] Haugen, R. A., Baker, N. L. (1991). The efficient market inefficiency of capitalization–weighted stock portfolios. The Journal of Portfolio Management, 17(3), 35–40. doi:10.3905/jpm.1991.409335