Изображение выглядит как текст, Шрифт, логотип, Цвет электрик

Автоматически созданное описание

**BU7152: Financial Modelling & Analysis**

**Group Members:**

**Álvaro Tapia**

**Niccolò Guerzoni Panichi**

**Ivan Nemtsev**

**Carlos Ramirez**

**Jose Ramon Cobos**

**Date: 07/05/2023**

# Table of Contents

[Table of Contents 2](#_Toc134396730)

[Introduction 3](#_Toc134396731)

[Data preparation 3](#_Toc134396732)

[General Statistics 4](#_Toc134396733)

[Univariate tests 8](#_Toc134396734)

[First technique: Shapiro-Wilk 8](#_Toc134396735)

[Second technique: Kolmogorov-Smirnov 9](#_Toc134396736)

[Third technique: D’Agostino test of skewness 9](#_Toc134396737)

[Fourth technique: Anscombe-Glynn test of kurtosis 9](#_Toc134396738)

[Fifth technique: Bonett-Seier test of kurtosis 10](#_Toc134396739)

[Multivariate tests 10](#_Toc134396740)

[First technique: Pearson’s product moment correlation coefficient t-test 11](#_Toc134396741)

[Second technique: Spearman rank correlation test 11](#_Toc134396742)

[Third technique: Kendall’s tau correlation coefficient 12](#_Toc134396743)

[T-tests 12](#_Toc134396744)

[Causality 13](#_Toc134396745)

[Volatilities 14](#_Toc134396746)

[Conclusion 16](#_Toc134396747)

# Introduction

The Trinity Wealth Management firm (TWM) currently accounts for a staggering 4.5 billion of Euros of assets under management. The firm provides portfolio management services to individual clients with different financial goals. Clients range from those with long-term investment goals to others with higher risk appetite. Roman Matkovsky, one of TWM’s most loyal and long-lasting clients, has the intention to retire in the foreseeable future. For this reason, he has requested to the firm to provide a report reflecting how his portfolio is currently performing in the market. Roman’s portfolio accounts for a varied percentage of asset classes, ranging from crypto currencies to treasury debt securities. Based on Roman’s preference and risk tolerance, the financial analyst team has selected the following assets to run the portfolio:

|  |  |  |
| --- | --- | --- |
| **Portfolio** | | |
| Asset Type | Name | Industry |
| Stock | BlackBerry | Technology |
| Stock | Uber | Technology |
| Stock | Intel | Technology |
| Stock | Walmart | Retail |
| Treasury Security | Treasury | Government Bonds |
| ETF | Franklin U.S. Treasury Bond | Government Bonds |
| Commodity | Gold | Minerals |
| Commodity | Silver | Minerals |
| Crypto currency | Bitcoin | Fintech |
| Crypto currency | Binance coin | Fintech |

The portfolio report details how the assets influence the overall performance of the portfolio. This objective will be achieved through different statistical approaches and data-driven analysis. The analytical team will conduct general statistics to understand the main characteristics of every asset, as well as the distribution. Furthermore, more complex analysis will be performed to calculate if the differences in return are significant as well as the volatilities of different assets. Finally, the correlation between variables and their causalities will be studied in the report.

# Data preparation

The data gathering process started out by downloading and connecting to datasets from the Yahoo Finance website for each asset. The data analysis was conducted on the R Studio software. The characteristics of the datasets rely on a time-series format. The data analysis starts by calculating the logarithmic returns of all 10 assets, and they have been converted into a data frame to simplify the analysis. Here is an example of the daily returns for the silver commodity:

A picture containing line, plot

Description automatically generated

# General Statistics

The mean returns have been calculated. At this stage missing values have been excluded, in order to avoid approximations. In a later stage, missing values will be filled with linear interpolation. The following table reflects the outputs of the mean returns:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Asset Name** | **Mean** | **Trimmed Mean 10%** | **Trimmed Mean 20%** | **Median** |
| BlackBerry | -0.001037237 | -0.001497504 | -0.001634854 | -0.001594059 |
| Uber | -0.0005657535 | -0.001037832 | -0.001338969 | -0.001337041 |
| Intel | -0.0004550487 | -2.842831e-05 | 0.0001610635 | -0.0001901157 |
| Walmart | 0.0002888531 | 0.0004239817 | 0.0005041657 | 0.0004035349 |
| Treasury Notes | -0.0002500413 | -0.0002513966 | -0.0002203835 | -0.0002282137 |
| Franklin T-Bond | 0.0003168652 | 0.0004224397 | 0.0004481281 | 0.0005548623 |
| Gold | 0.0002867146 | 0.0004451224 | 0.0005424217 | 0.0005146292 |
| Silver | 0.0003168652 | 0.0004224397 | 0.0004481281 | 0.0005548623 |
| Bitcoin | 8.798334e-05 | 0.0006606419 | 0.0007935275 | 0.000638879 |
| Binance coin | 0.00182914 | 0.001622646 | 0.001644858 | 0.0009758786 |

BlackBerry: This investment has shown consistently negative mean and trimmed mean returns, indicating a lack of profitability. The majority of returns have been below zero, as evidenced by the negative median return. Although the asset has demonstrated relative stability with a small range of returns, it has not yielded positive outcomes.

Uber: Similar to BlackBerry, Uber has experienced negative average returns based on both mean and trimmed mean calculations. However, the median return is slightly higher, indicating occasional positive returns. Uber is characterized by a larger range of returns, suggesting higher volatility compared to BlackBerry. Despite occasional positive returns, the central tendency measures still indicate an overall lack of profitability.

Intel: This investment has displayed slightly positive mean and trimmed mean returns, suggesting some positive average returns. However, the negative median return implies that the majority of returns have been below zero. Although the asset has exhibited relative stability with a small range of returns, the mixed central tendency measures indicate inconsistent profitability.

Walmart: The mean and trimmed mean returns for Walmart have been positive, indicating favourable average returns. The positive median return further supports the notion that the majority of returns have been above zero. With a relatively small range of returns, Walmart has demonstrated stability, and the positive central tendency measures highlight its profitability.

Treasury Notes: This investment has consistently shown negative central tendency measures, indicating negative average returns. Despite the asset's relative stability with a small range of returns, the negative central tendency measures imply that it has not been a profitable investment.

Franklin T-Bond: The central tendency measures for Franklin T-Bond have consistently been positive, signifying positive average returns. With a relatively small range of returns, this investment has displayed stability, and the positive central tendency measures confirm its profitability.

Gold: The central tendency measures for gold have consistently been positive, indicating positive average returns. With a relatively small range of returns, the asset has demonstrated stability, and the positive central tendency measures suggest its profitability.

Silver: Silver has shown positive central tendency measures, signifying positive average returns. Similar to gold, silver has displayed stability with a relatively small range of returns. The positive central tendency measures further indicate its profitability.

Bitcoin: This investment has positive central tendency measures, indicating positive average returns. However, it is important to note that the range of returns is significantly larger compared to other assets, highlighting its high volatility. Despite the positive central tendency measures, investors should be cautious due to the associated risk and volatility.

Binance coin: Among all the assets, Binance coin has exhibited the highest central tendency measures, indicating the highest average returns. However, the range of returns surpasses even that of Bitcoin, indicating extreme volatility. While the high central tendency measures suggest profitability, investors must exercise caution and be aware of the significant risk and volatility associated with this asset. The variability of the different assets has been estimated using varying measures.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Asset Name** | **Mean Abs. Deviation** | **Variance** | **Standard Deviation** | **Median Abs. Deviation** |
| BlackBerry | 0.02623463 | 0.001789854 | 0.04230667 | 0.02657133 |
| Uber | 0.02649867 | 0.001401978 | 0.037443 | 0.03009158 |
| Intel | 0.01604557 | 0.0005774313 | 0.0240298 | 0.01680423 |
| Walmart | 0.009583272 | 0.0002202657 | 0.01484135 | 0.01005302 |
| Treasury Notes | 0.002871305 | 0.003247868 | 0.05699007 | 0.03586497 |
| Franklin T-Bond | 0.01322967 | 0.00152231 | 0.03901679 | 0.02431101 |
| Gold | 0.006704784 | 9.257072e-05 | 0.009621368 | 0.006717626 |
| Silver | 0.01322967 | 0.0003854474 | 0.01963281 | 0.01303763 |
| Bitcoin | 0.03629784 | 1.594921e-05 | 0.003993646 | 0.003007252 |
| Binance coin | 0.02594925 | 0.0003854474 | 0.01963281 | 0.01303763 |

With the data previously obtained we can divide assets into four different types depending on the accuracy with which we can predict how they will perform.

First, we have BlackBerry, Uber and Intel, all of which have higher mean absolute deviation, variance and standard deviation values than the other assets. This shows that the returns are more dispersed with respect to the mean and that it is more difficult to predict what the outcome will be. The only difference between these assets is the median absolute deviation, which in the case of BlackBerry and Intel is similar to the mean absolute deviation, so there are no significant outliers in the data set. In the case of Uber, this value is higher than the mean, indicating that the dataset has a skewed distribution or contains extreme values or outliers. This makes it even less reliable than the rest when we have to choose between them.

In the second group we have the assets of Walmart, Treasury Notes and Franklin T-Bond. Unlike the previous ones, they have a lower mean value, variance and standard deviation, so they can be considered as assets that are easier to predict. But in all three cases they have a median absolute deviation higher than the mean deviation so that outliers can be found in all three cases. Care should be taken with these outliers as they affect the rest of the analysis.

Finally, we have two completely different groups of assets, commodities and cryptocurrencies. In the former, the mean absolute deviation is relatively low, as is the variance and standard deviation. Therefore, we can consider these assets as the easiest to predict and they have a median equal to the mean, so there are no outliers. The latter have a much larger mean deviation, but the variance and standard deviation are low. This can occur when the data are highly clustered around the mean, with little variation between data points. In this case, the median is smaller than the mean, meaning that the data set has a symmetric or normal distribution with relatively small dispersion around the mean.

In order to have an overlook of returns, different quantiles have been calculated. Boxplots have been created to visualize the results. Here is an example for silver:

Another method of visualization is the histogram since it shows different way of plotting the distributions. Here is a histogram example for silver:

A picture containing line, screenshot, diagram, text

Description automatically generated

A picture containing diagram, line, technical drawing, screenshot

Description automatically generated

Then a Kernel density estimation was performed to estimate the probability density function of each asset. Accordingly, the estimated density function was plotted. Here is an example with the Silver commodity again:

A picture containing line, plot, diagram, text

Description automatically generated

Assets’ distributions have been compared to the expected distribution to the variables if it were normally distributed thanks to the Quantile plot. Here is an example with silver:

A picture containing line, plot, diagram, text

Description automatically generated

From this graph it is clear that the data points are not normally distributed by their distance from the slope. This outcome could mean that the returns have heavier tails or are more skewed than a normal distribution. Our team has taken this information to further investigate of underlying factors that might be producing the outliers.

# Univariate tests

The report further studies different techniques to understand the univariate test for each series. These tests are useful for TWM’s interest in analysing if the asset returns follow a particular distribution, thus determined the presence of outliers.

## First technique: Shapiro-Wilk

This tests if the dataset comes from a normal distribution. Departures from normality are taken into account within this technique. The null hypothesis of this test is that the data is normally distributed, and the alternative hypothesis is that it is not normally distributed. The tests statistic W measures the agreement of the dataset compared to the expected values in a scenario with normality.

The following code resembles the results from this test of the silver commodity:

A picture containing text, font, screenshot

Description automatically generated

In this case, the p-value is less than 2.2e-16 (which is essentially zero), which means that we can reject the null hypothesis and conclude that the data is not normally distributed. Since the W value is 0.94011, it can be suggested that the data is not perfectly normal, but close.

## Second technique: Kolmogorov-Smirnov

This technique determines whether a sample of data comes from a specified distribution or not. Differently to the Shapiro technique, which focuses only on normality. It calculates the maximum distance of the t-distribution. Example with the code of the silver commodity:

A picture containing text, font, screenshot

Description automatically generated

Since the p-value is lower than the critical value of 0.05, we reject the null hypothesis.

## Third technique: D’Agostino test of skewness

This technique determines the difference of skewness from the dataset compared to the one of a normal distributed one. Example with the code of the silver commodity:

A picture containing text, font, screenshot

Description automatically generated

The skewness value of -0.53434 indicates that the data are slightly left-skewed. The p-value of 2.942e-13 is less than the significance level of 0.05, so we reject the null hypothesis and conclude that the data are not normally distributed.

## Fourth technique: Anscombe-Glynn test of kurtosis

The Anscombe-Glynn tests if a sample data has a kurtosis that is significantly different from a normal distributed. This test assumes that sample datasets are normally distributed. Example with the code of the silver commodity:

A picture containing text, screenshot, font

Description automatically generated

The p-value shows a result lower than 0.05, this reflects enough evidence to reject the null hypothesis as the kurtosis is different than the critical value 3.

## Fifth technique: Bonett-Seier test of kurtosis

Similarly to the fourth technique, the Bonnett-Seier tests the if a data sample shows a kurtosis significantly different from a normal distributed one. The difference is that this technique does not assume the data is normally distributed.

A screen shot of a computer

Description automatically generated with medium confidence

Since the p-value is less than 0.05 the null hypothesis is rejected again for this test. The data is different from a normal distribution in terms of its kurtosis.

The outcomes of the silver example were the same for the other 9 assets within each univariate test technique. It is important to check whether the assets' returns are normally distributed because many statistical methods and financial models rely on the assumption of normality. For example, portfolio optimization and risk management models often assume normality of returns. If the returns are not normally distributed, these models may produce inaccurate or misleading results. Additionally, some hypothesis tests, such as the t-test and ANOVA, assume normality, and violations of this assumption can lead to incorrect conclusions. Therefore, it is important to consider alternative methods for analysing and modelling the returns from these assets, such as using non-parametric methods or other distributional assumptions.

# Multivariate tests

Following the outcomes from the univariate section, TWM’s analytics team further the portfolio’s study by conducting different multivariate tests. Multivariate techniques are useful to compare two different assets. The analysis compares patterns and correlations, this is important to understand how two assets relate by performing in different market conditions. Multivariate tests help understand the relationship regarding volatility and returns of multiple assets, making easier to discover hidden relationships compared to the analysis of an isolated asset. In the following multivariate techniques, the financial analysis team has picked two different types of assets that perform in different markets. The compared assets are the silver commodity and the U.S. treasury note.

## First technique: Pearson’s product moment correlation coefficient t-test

This method analyses the linear relationship or degree of correlation between two assets. This coefficient assessment helps understand the degree of diversification in the overall portfolio. The following chart represents the outcomes from of this method by comparing the assets: silver (commodity) and Treasury Notes (government bond).

A screenshot of a test

Description automatically generated with low confidence

The chart shows a correlation coefficient of 0.1772002, which means there is a weak positive relationship between the two assets. A p-value of 6.07e-06 is lower than the significance level of 0.05, suggesting a strong null hypothesis. Another indicator is the outcome of 3.035e-0 for the alternative null hypothesis, which is greater than 0. There is a positive relationship between both silver and T-note assets by their statistically significant positive correlation.

## Second technique: Spearman rank correlation test

This technique focuses on the monotonic relationship of two assets. Like any multivariate test it analyses the degree of correlation between two variables, in this case assets. The difference with the Pearson test is that this one does not take into account the linear relationship of the comparison. Rather, it ranks the variables and then calculates as a normal Pearson test on the ranks. As any non-parametric statistical measure, this test does not assume the disruption to be normal.

A picture containing text, font, screenshot, white

Description automatically generated

The above graph contemplates the same two assets from the previous example. The output shows that the indicator *rho* is 0.167085 between both variables. Having a p-value lower than 0.05 the null hypothesis is rejected. Nevertheless, the strength of the correlation is weak by a coefficient close to 0.

## Third technique: Kendall’s tau correlation coefficient

This technique is useful to analyse the ordinal relationship between two variables. Alike the previous two techniques this one is used to analyse the degree of correlation of two assets, and similar to the Spearman rank technique it is a non-parametric measure. Kendall calculates by comparing the number of concordant and discordant pairs. This method is better deployed with non-normally distributed returns of an asset.

A picture containing text, font, receipt, screenshot

Description automatically generated

The silver example from the above chart represents the outcomes using this technique. The *tau* value is 0.1131939, which means there is a weak positive correlation between the two assets. The p-value is 1.731e-05 being less than the significance level of 0.005, thus showing a correlation coefficient statistically significant. Again, through this method the null hypothesis is rejected and assume that correlation significance is present in both assets.

# T-tests

The financial analysis team continued with the statistical evaluation of the overall portfolio by conducting a t-test. The test evaluated the difference in the return means between assets which are similar in both asset-type and industry (Silver-Gold, Bitcoin-Binance, T-bonds-T-notes, Walmart-Uber, Intel-Blackberry). This analysis is performed because a statistical significance in the mean of the returns between very similar assets would mean that the asset with the lowest return has to be sold out of the portfolio.

The different return returns for every pair of assets have been merged on a comparative table and converted into a data frame. After creating the table, the t-test was performed.

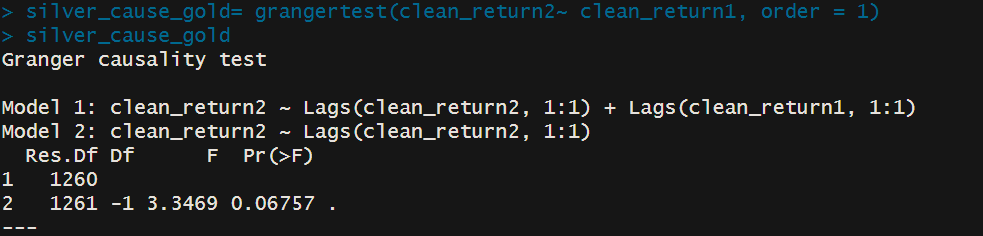
The t-tests results from all assets reflected a p-value below the critical value of 0.05. Therefore, there are not significant differences between the pair of assets, reason why the current portfolio should remain unchanged.

# Causality

Granger causality is a statistical tool that can be used to determine whether one asset's past values can be used to predict another asset's future values. This can provide insight into the causal relationships between assets and help in making informed investment decisions. If one asset Granger causes another, it means that the first asset's past values contain information that can help predict the future values of the second asset. This information can be useful, as it can potentially allow to predict the future movements of an asset based on the movements of another asset. Therefore, studying Granger causality between different financial assets can provide valuable insights for investors looking to make informed investment decisions.

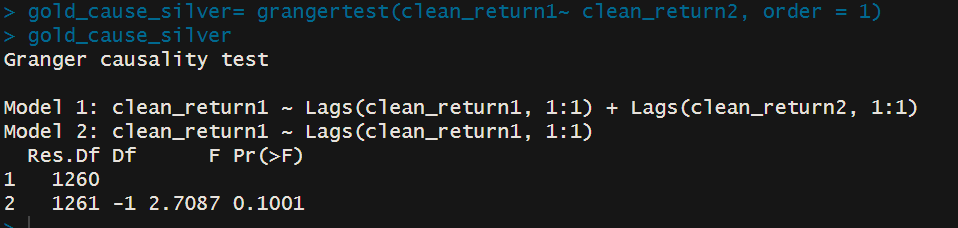
If it is found that some of the assets in a portfolio are Granger caused by others, it would be prudent to consider reducing or eliminating investments in the Granger caused assets. This is because the Granger causality relationship suggests that the causing asset contains valuable information that can be used to predict the behaviour of the caused asset. Therefore, holding the caused asset may expose an investor to additional risk, as the value of the caused asset is likely to be influenced by the fluctuations of the causing asset. On the other hand, if there is no Granger causality relationship between assets, it may be beneficial to include both assets in a diversified portfolio, as the movements of one asset are unlikely to affect the behaviour of the other.

Here there are some examples about assets of Roman’s portfolio that could cause others.



The null hypothesis is that silver does not Granger cause gold. The result indicates that the p-value is 0.06757, which is greater than the commonly used significance level of 0.05. Therefore, we do not reject the null hypothesis, and we conclude that there is no evidence of Granger causality from gold to silver at lag 1.

Since the granger test is asymmetric, the other direction of causality is tested.



Even in this case, we conclude that there is no evidence of Granger causality from silver to gold at lag 1.

# Volatilities

Volatility is a fundamental measurement for the portfolio’s performance and relevancy for the client. The analysis is performed with the ARIMA-GARCH model, which is a popular time-series model used to estimate the volatility of financial assets. It combines two models: the ARIMA model for the conditional mean and the GARCH model for the conditional variance.

The ARIMA-GARCH model estimates several parameters for each asset, which can provide information about the asset's behaviour and potential risk.

* The mu parameter is the mean of the asset's return series. It represents the expected return of the asset over time.
* The AR(1) parameter (ar1) is the coefficient of the first lag of the asset's return series in the ARIMA model. It indicates the degree of autocorrelation in the series.
* The omega parameter is the constant term in the GARCH model that captures the long-run volatility of the asset.
* The alpha1 parameter is the coefficient of the first lag of the asset's conditional variance in the GARCH model. It represents the persistence of shocks to the variance of the asset's returns.
* The beta1 parameter is the coefficient of the first lag of the asset's conditional variance in the ARCH model. It represents the decay rate of the impact of past shocks on the asset's current volatility.
* The loglikelihood is a measure of how well the model fits the data. A higher value of the loglikelihood indicates a better fit.
* The AIC (Akaike Information Criterion) is a measure of the relative quality of the model, taking into account the trade-off between the goodness of fit and the complexity of the model. A lower AIC value indicates a better model fit with a balance between the goodness of fit and the number of parameters used.

In addition, the coefficient which are statistically significant (p-value < 0.05) are coloured in blue.

A screenshot of a computer screen

Description automatically generated with low confidence

Gold: The estimated volatility for gold is relatively low compared to other assets, with an omega value of 0.000006. This suggests that gold has low volatility and is relatively stable compared to other assets.

Intel: The estimated volatility for Intel is moderate, with an omega value of 0.000140. The alpha1 value of 0.183661 suggests that the asset's volatility is affected by past shocks, while the beta1 value of 0.565803 indicates that the asset's volatility is positively related to past volatility.

10Y T-notes: The estimated volatility for 10-year Treasury notes is also relatively low, with an omega value of 0.000000. The alpha1 value of 0.121496 suggests that the asset's volatility is affected by past shocks, while the beta1 value of 0.866696 indicates that the asset's volatility is positively related to past volatility.

BlackBerry: The estimated volatility for BlackBerry is high, with an omega value of 0.000148. The alpha1 value of 0.128645 suggests that the asset's volatility is affected by past shocks, while the beta1 value of 0.773265 indicates that the asset's volatility is positively related to past volatility.

Walmart: The estimated volatility for Walmart is moderate, with an omega value of 0.000020. The alpha1 value of 0.285217 suggests that the asset's volatility is affected by past shocks, while the beta1 value of 0.646879 indicates that the asset's volatility is positively related to past volatility.

Uber: The estimated volatility for Uber is relatively high, with an omega value of 0.000098. The alpha1 value of 0.132786 suggests that the asset's volatility is affected by past shocks, while the beta1 value of 0.788258 indicates that the asset's volatility is positively related to past volatility.

Silver: The estimated volatility for silver is relatively low, with an omega value of 0.000004. This suggests that silver has low volatility and is relatively stable compared to other assets.

Binance Coin: The estimated volatility for Binance Coin is high, with an omega value of 0.000089. The alpha1 value of 0.167722 suggests that the asset's volatility is affected by past shocks, while the beta1 value of 0.826324 indicates that the asset's volatility is positively related to past volatility.

Bitcoin: The estimated volatility for Bitcoin is high, with an omega value of 0.000095. The alpha1 value of 0.188084 suggests that the asset's volatility is affected by past shocks, while the beta1 value of 0.795744 indicates that the asset's volatility is positively related to past volatility.

ETF Franklin: The estimated volatility for the ETF Franklin is relatively low, with an omega value of 0.000000. The alpha1 value of 0.029695 suggests that the asset's volatility is affected by past shocks, while the beta1 value of 0.966516 indicates that the asset's volatility is positively related to past volatility.

# Conclusion

The general statistical measures revealed that Binance coin has the highest mean, which means this asset has experienced an increase in its returns over the years. Whereas Blackberry showed the lowest mean, consequently a decrease of value over time. T-notes showed the highest variance, meaning its returns were more spread out from the mean. Bitcoin showed the lowest variance. The t-test analysis showed lower p-values for all 10 assets, no significance differences between assets in the portfolio. Commodity assets showed the lowest volatilities, while stocks like BlackBerry and Uber showed the highest. Meaning, stocks’ returns changes faster over a period of time. Results from the different univariate tests reflected p-values less than the critical value, resulting in rejection of the 10 null hypothesis of all assets. The data points of the returns did not show normality in their distributions. Regarding the multivariate tests, all assets were compared by their correlation into five pairs of two. All showed a positive correlation coefficient, and t-statistics showing coefficients different from zero. The pair of silver and t-notes showed a significantly higher correlation, whereas the pair of silver and Walmart stock showed the lowest positive correlation. Regarding causality was compared with the commodities gold and silver. Gold was not Granger caused by silver, since the p-value was bigger than the significance level of 0.05. There are the main points of Roman’s diversified portfolio.