

$$F = G \frac{m_1 m_2}{d^2}$$

Using Data science tools and techniques
for creating and maintaining a passive
investment portfolio

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$$

Arthur Valle - TRENDSET

$$\frac{df}{dt} = \lim_{h \rightarrow 0} \frac{f(t+h) - f(t)}{h}$$

Content



WHY INVEST?



WHY PASSIVE
INVESTING?



KEY TERMINOLOGY



EXAMPLE OF
PASSIVE PORTFOLIO



3 STEPS TO
IMPLEMENT THE
PASSIVE STRATEGY



DATA SCIENCE TOOLS
& TECHNIQUES



CONCLUSION

WHY INVEST?

*"YOU, OWNER OF PROFITABLE COMPANIES ALL
OVER THE WORLD"*

- **You won't get rich** working for others - or even for yourself - from 9 am to 6 pm every day.

"FREE TIME IN THE FUTURE"

For a decent retirement, you need to put your money to work for you and wait the time act on it.

WHY PASSIVE INVESTING?

"IF YOU CAN'T BEAT IT, JOIN IT"

- In the long run, **few investors are able to consistently "beat the market"**, with low risk.

"INVESTING SHOULD BE AS BORING AS WATCHING THE GRASS TO GROW"

- Do not waste your time reading balance sheets to find "the next Tesla": **it's the time in the market and periodic contributions that will buy you "free time in the future"**.
- Dedicate your hours for your current source of income (or even new ones), but **put your savings (and time) to work for you!**

Key terminology

Passive Investment:

- Investment strategy that **aims to obtain the same average return as the stock market indexes**, such as S&P 500 or iBOVESPA.
- Implemented via index funds: higher diversification (and consequently lower risk) than picking isolated companies' stocks.
- Requires less effort (and knowledge):
 - it avoids unnecessary *buys* and *sells*
 - In the long run, it tends to have returns similar to “active investing”

ETFs (or index funds):

- ETF (Exchange Traded Fund) is a **financial asset traded on the stock exchange. It replicates the performance of a given index.**
- ETFs are used in the passive investment strategy as they are naturally diversified and have lower management fees.
- Examples of ETFs:
 - Vanguard Total World Stock ETF (VT)
 - SPDR S&P 500 ETF (SPY)
 - It Now S&P 500 Fundo De Indice (SPXI11)
 - Ishares Ibovespa Fundo De Indice (BOVA11)

EXAMPLE OF A SIMPLE PASSIVE PORTFOLIO, WITH JUST 4 ETFs:

25% in **Developed Markets** Equities

- Invest in listed companies from developed markets: US, Europe, Japan, etc

25% in **Emerging Market** Equities

- Invest in listed companies from Emerging Markets: China, Russia, Brazil, Mexico, etc

25% in **REITS Real Estate Investment Trusts**

- Invest in the Real Estate Industry.

25% in **Fixed Income**

- Invest in Government and Corporate Bonds.

3 STEPS TO IMPLEMENT THE PASSIVE STRATEGY:

Step 1: Objectives, profile and portfolio

- Define your **investment goals** (ex: 12% return per year) and **build a portfolio with assets that have the capability** to achieve the goals but that also **fits your profile** as an investor (conservative, moderate, aggressive, etc)

STEP1: DATA SCIENCE TOOLS & TECHNIQUES

- For checking (statistical) capability: *Normality Test, Mean, Standard Deviation, Confidence Intervals, Hypothesis Testing, Control Charts, Sigma level*
- For selecting assets classes and assets and composing the portfolio: *(in addition to the ones above) DOE-Design of Experiments, ANOVA, Back testing, Correlation*

Shapiro-Wilk test, using a right-tailed normal distribution

1. H_0 hypothesis

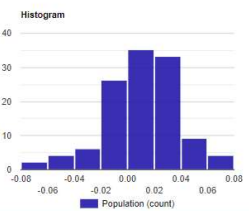
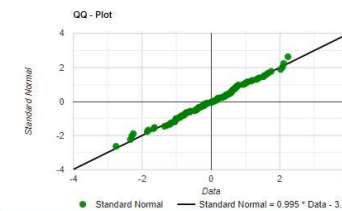
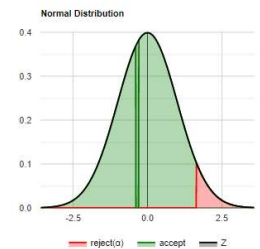
Since $p\text{-value} > \alpha$, we accept the H_0 . It is assumed that the data is normally distributed. In other words, the difference between the data sample and the normal distribution is not big enough to be statistically significant.

2. P-value

p-value is **0.636628**, hence, if we would reject H_0 , the chance of type1 error (rejecting a correct H_0) would be too high: 0.6366 (63.66%). The larger the p-value, the more it supports H_0 .

3. The statistics

W is **0.990919**. It is in the 95% critical value accepted range: [0.9779 ; 1.0000]



Sample size (n):	117
Average (\bar{x}):	0.0104894
Median:	0.011455171
Sample Standard Deviation (S):	0.0249749
Sum of Squares:	0.0723545
b:	0.267764
Skewness:	-0.182217
Skewness Shape:	Potentially Symmetrical (pval=1.585)
Excess kurtosis:	0.0576749
Tails Shape:	Potentially Mesokurtic , normal like tails (pval=0.897)
P-value:	0.636628
Outliers:	-0.162790698, -0.061967027, 0.083333333

WHAT ARE EXAMPLES OF LONG-TERM FINANCIAL GOALS?

Example 1: **Preservation of capital**

- for example, a return target of around 1.5% per year.

Example 2: **Protection against inflation and/or devaluation of the currency**

- a return target of 3 to 5% per year.

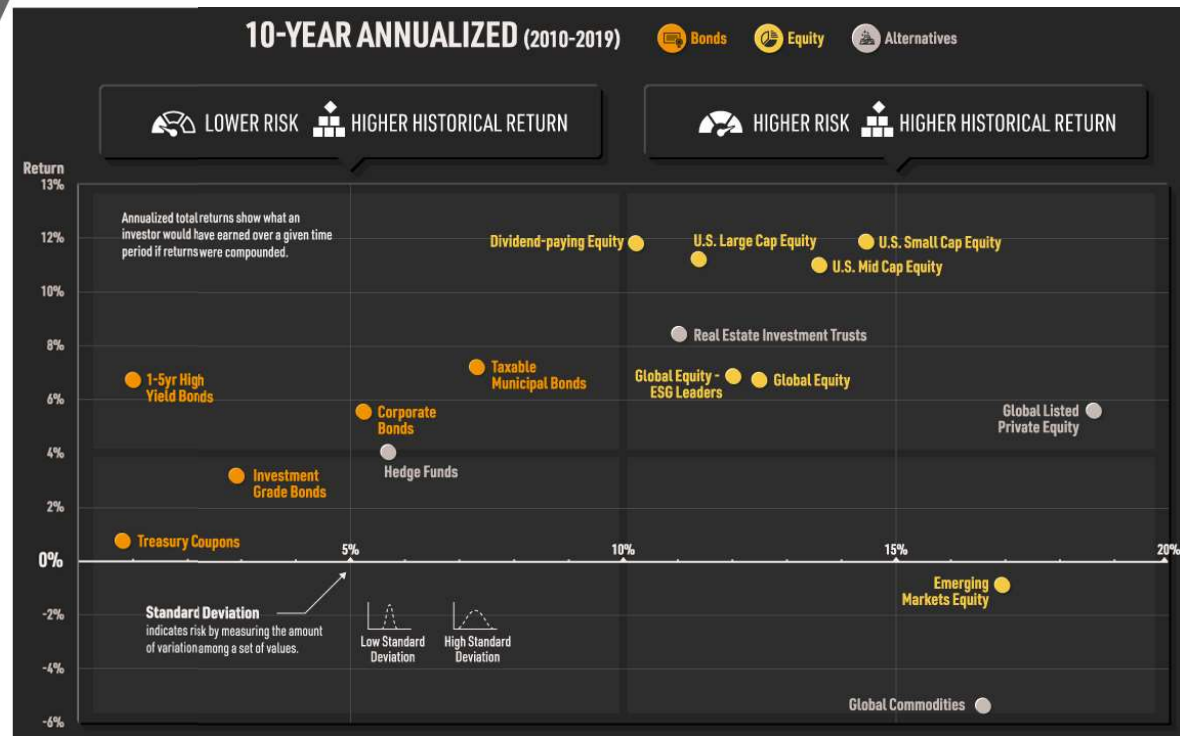
Example 3: **Passive income**

- a target of 0.40% per month, similar to the returns of real estate rentals.

Example 4: **Monetizing capital**

- an average return expectation from 8 to 12% per year.

KNOW THE ASSET CLASSES' EXPECTED RETURNS (AND RISKS)



<https://advisor.visualcapitalist.com/asset-class-risk-and-return/>

HAVE A GLOBALLY DIVERSIFIED PORTFOLIO (WITH LOW CORRELATED ASSETS)

Figure 6. Trailing 12-month return differential between U.S. and non-U.S. stocks



Figure 1. Historical mix of global equity market capitalization

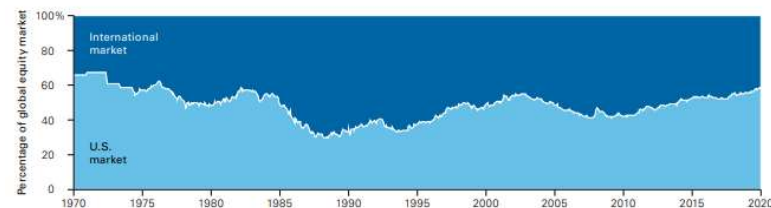


Figure 4. Historically, correlations have risen, meaning less impact from global diversification

12-month and 10-year rolling correlations between U.S. and international stocks



SOURCE:

Vanguard

Global equity investing:
The benefits of diversification
and sizing your allocation

Vanguard Research

April 2021

Figure 1: Relative sizes of world stock markets, end-1899 (left) versus start-2021 (right)

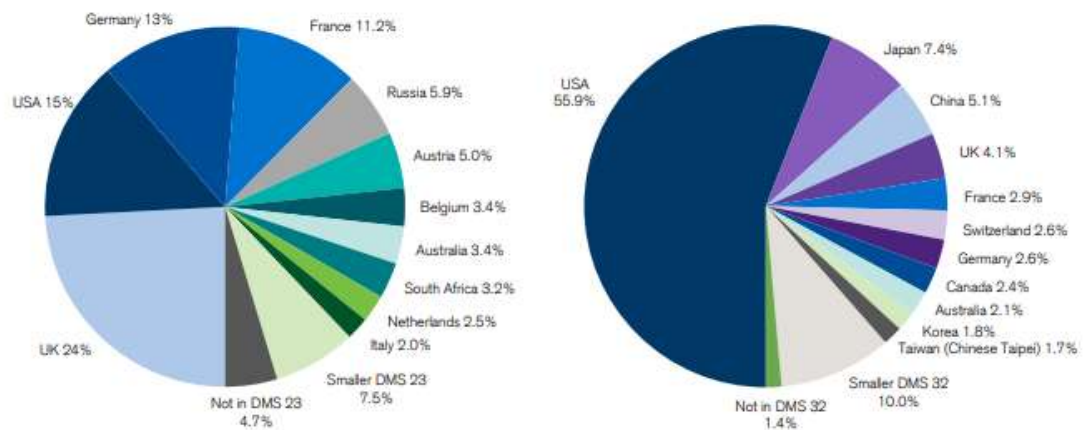


Figure 19: EM versus DM weightings, 1980-2020

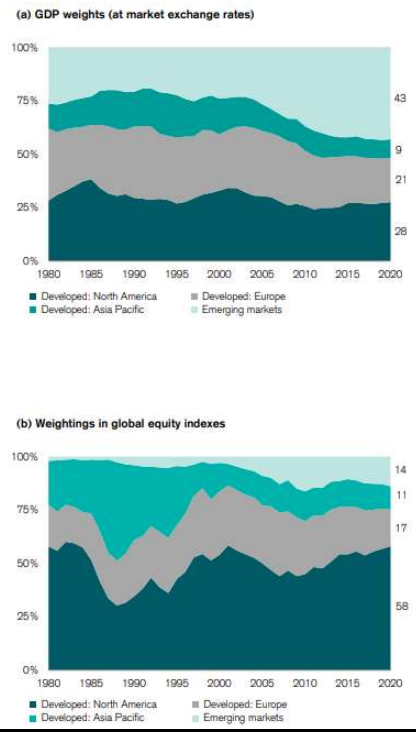
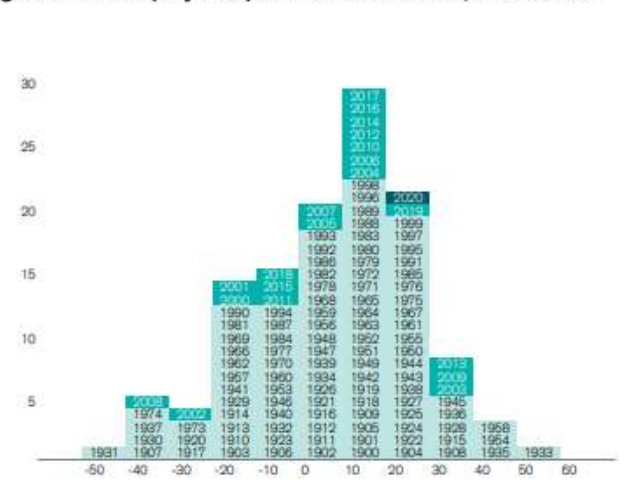


Figure 10: US equity risk premium versus bills, 1900-2020

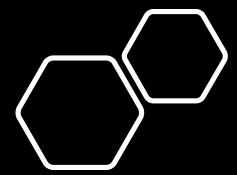


SOURCE:

March 2021 CREDIT SUISSE

Research Institute
Credit Suisse Global Investment Returns
Yearbook 2021 Summary Edition

ASPECTS TO CONSIDER:
A) LOCAL OR GLOBAL?



ASPECTS TO CONSIDER: B) DIVERSIFICATION AMONG ASSET CLASSES

Asset Class Returns

2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
EM 39.8%	HG Bnd 5.2%	EM 79.0%	REIT 28.0%	REIT 8.3%	REIT 19.7%	Sm Cap 38.8%	REIT 28.0%	REIT 2.8%	Sm Cap 21.3%	EM 37.8%	Cash 1.8%	Lg Cap 31.5%	Sm Cap 20.0%
Int'l Stk 11.6%	Cash 1.7%	HY Bnd 57.5%	Sm Cap 26.9%	HG Bnd 7.8%	EM 18.6%	Lg Cap 32.4%	Lg Cap 13.7%	Lg Cap 1.4%	HY Bnd 17.5%	Int'l 25.6%	HG Bnd 0.0%	REIT 28.7%	EM 18.7%
AA 7.6%	AA -22.4%	Int'l Stk 32.5%	EM 19.2%	HY Bnd 4.4%	Int'l Stk 17.9%	Int'l Stk 23.3%	AA 6.9%	HG Bnd 0.6%	Lg Cap 12.0%	Lg Cap 21.8%	HY Bnd -2.3%	Sm Cap 25.5%	Lg Cap 18.4%
HG Bnd 7.0%	HY Bnd -26.4%	REIT 28.0%	HY Bnd 15.2%	Lg Cap 2.1%	Sm Cap 16.4%	AA 11.5%	HG Bnd 6.0%	Cash 0.0%	EM 11.6%	Sm Cap 14.7%	REIT -4.0%	Int'l Stk 22.7%	AA 9.8%
Lg Cap 5.5%	Sm Cap -33.8%	Sm Cap 27.2%	Lg Cap 15.1%	AA 0.3%	Lg Cap 16.0%	HY Bnd 7.4%	Sm Cap 4.9%	Int'l Stk -0.4%	REIT 8.6%	AA 14.6%	Lg Cap -4.4%	AA 18.9%	Int'l Stk 8.3%
Cash 4.4%	Lg Cap -37.0%	Lg Cap 26.5%	AA 13.5%	Cash 0.1%	HY Bnd 15.6%	REIT 2.9%	HY Bnd 2.5%	AA -1.3%	AA 7.2%	REIT 8.7%	AA -5.6%	EM 18.9%	HY Bnd 7.5%
HY Bnd 2.2%	REIT -37.7%	AA 24.6%	Int'l Stk 8.2%	Sm Cap -4.2%	AA 12.2%	Cash 0.1%	Cash 0.0%	Sm Cap -4.4%	HG Bnd 2.7%	HY Bnd 7.5%	Sm Cap -11.0%	HY Bnd 14.4%	HG Bnd 6.1%
Sm Cap -1.6%	Int'l Stk -43.1%	HG Bnd 5.9%	HG Bnd 6.5%	Int'l Stk -11.7%	HG Bnd 4.2%	HG Bnd -2.0%	EM -1.8%	HY Bnd -4.6%	Int'l Stk 1.5%	HG Bnd 3.5%	Int'l Stk -13.4%	HG Bnd 8.7%	Cash 0.6%
REIT -15.7%	EM -53.2%	Cash 0.1%	Cash 0.1%	EM -18.2%	Cash 0.1%	EM -2.3%	Int'l Stk -4.5%	EM -14.6%	Cash 0.3%	Cash 0.8%	EM -14.3%	Cash 2.2%	REIT -5.1%

Asset Class - Index	Annual	Best	Worst
Large Cap Stocks - S&P 500 Index	9.88%	32.4%	-37.0%
Small Cap Stocks - Russell 2000 Index	8.91%	38.8%	-33.8%
International Developed Stocks - MSCI EAFE Index	4.97%	32.5%	-43.1%
EM Stocks - MSCI Emerging Markets Index	6.95%	79.0%	-53.2%
REITs - FTSE NAREIT All Equity Index	7.15%	35.1%	-37.7%
High Grade Bonds - Bloomberg Barclays U.S. Agg Index	4.40%	8.7%	-2.0%
High Yield Bonds - ICE BofA US High Yield Index	7.44%	57.5%	-26.4%
Cash - S&P U.S. Treasury Bill 0-3 Mth Index	1.11%	4.7%	0.0%
Asset Allocation Portfolio*	7.02%	24.6%	-22.4%

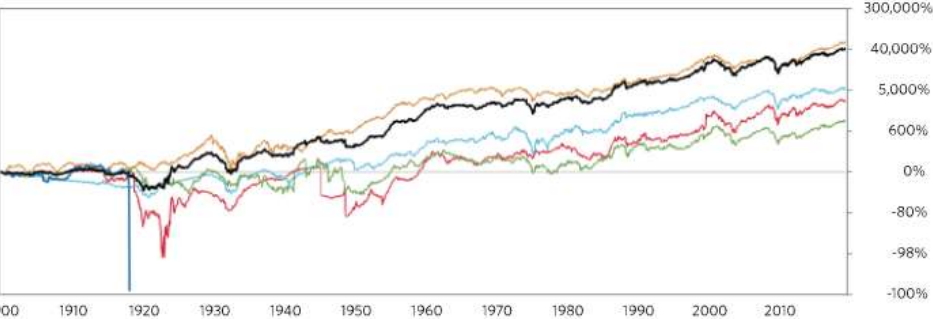
Past performance does not guarantee future returns. The historical performance in market trends across several asset classes over the past fifteen years. Return annual returns (reinvestment of all distributions) and does not include fees and investments you choose should reflect your financial goals and risk tolerance. Ft to a financial professional. All data are as of 6/30/21.

*Asset Allocation Portfolio is 15% large cap stocks, 15% international stocks, 10% stocks, 10% emerging market stocks, 10% REITs, 40% high-grade bonds, and ann

Source: <https://novelinvestor.com/asset-class-returns/>

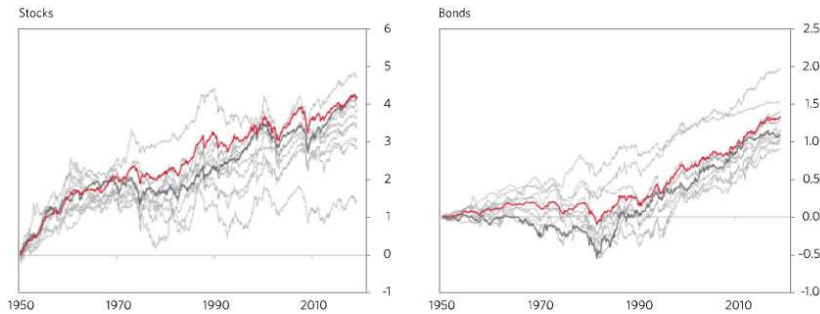
Equity Market Cumulative Excess Returns Since 1900 (In Scale)

— DEU — GBR — USA — FRA — RUS — Equal-Weight



Cumulative Excess Returns (In)

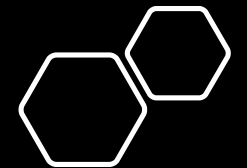
— Individual Countries — USA — Equal-Weight



1980s	1990s	2000s	2010s
Sweden 503%	Switzerland 231%	Norway 48%	United States 182%
Korea 354%	United States 217%	Brazil 45%	New Zealand 149%
Japan 310%	Sweden 190%	Canada 42%	Sweden 146%
Spain 188%	France 177%	Australia 36%	Japan 105%
Equal Weight 185%	United Kingdom 110%	Korea 22%	Germany 99%
Germany 179%	Spain 98%	Spain 17%	Switzerland 97%
United Kingdom 173%	Germany 92%	Equal Weight 1%	France 92%
Italy 169%	Australia 59%	New Zealand -3%	United Kingdom 83%
France 158%	Equal Weight 53%	Switzerland -4%	Norway 78%
Switzerland 96%	Canada 52%	Sweden -13%	Equal Weight 74%
United States 96%	Italy 40%	Taiwan -23%	Taiwan 55%
Australia 39%	Norway 2%	United Kingdom -25%	Canada 54%
Norway 23%	New Zealand -6%	United States -27%	Australia 41%
Canada -4%	Japan -47%	France -32%	Korea 27%
	Taiwan -49%	Italy -35%	Italy 20%
	Korea -56%	Germany -36%	Spain 11%
		Japan -41%	Brazil -26%
Avg. Correl. 46%	Avg. Correl. 50%	Avg. Correl. 74%	Avg. Correl. 65%
Best - Worst 507%	Best - Worst 296%	Best - Worst 89%	Best - Worst 209%

Source: <https://www.bridgewater.com/research-and-insights/geographic-diversification-can-be-a-lifesaver-yet-most-portfolios-are-highly-geographically-concentrated>

ASPECTS TO CONSIDER:
C) EQUAL-WEIGHTED OR CAP-WEIGHTED PORTFOLIO?



3 STEPS TO IMPLEMENT THE PASSIVE STRATEGY?

Step 2: Stay invested

- **Contribute periodically** to your portfolio (**and re-balance it whenever needed**).

STEP2: DATA SCIENCE TOOLS & TECHNIQUES

- For Predictions: *Confidence Intervals, Regression, Time Series, MonteCarlo Simulation, Machine Learning*



<https://graphite-note.com/>

Results - APA Style

M: 95% CI [0.0059, 0.015]
SD: 95% CI [0.022, 0.029]

Mean confidence interval: [0.005916280526, 0.01506253749].

Alternatively: 0.01048940901 ± 0.004573128482

Margin of Error (MOE): 0.004573128482.

Standard Error (S.E): 0.002308930928.

Since the population's σ is not known, the formula uses the **T distribution** with $n-1$ degrees of freedom.

If you would calculate the confidence interval over an infinite number of samples with a sample size of **117**, 95% of the calculated confidence intervals will contain the mean's true value.

$$\bar{x} \pm T_{1-\alpha/2}(df) * \frac{S}{\sqrt{n}}$$

$$\bar{x} \pm T_{1-0.05/2}(116) * \frac{0.02497490656}{\sqrt{117}}$$

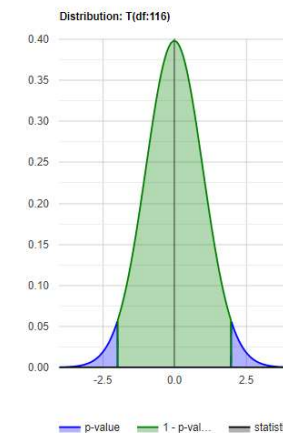
$$\bar{x} \pm T_{0.975}(116) * \frac{0.02497490656}{\sqrt{117}}$$

$$0.01048940901 \pm 1.980625937 * \frac{0.02497490656}{\sqrt{117}}$$

$$0.01048940901 \pm 1.980625937 * 0.002308930928$$

Since $T_{\alpha/2} = -T_{1-\alpha/2}$, you may use $T_{\alpha/2}$ instead of $T_{1-\alpha/2}$

You may calculate T using the [distribution calculator](#) with distribution: 'T',DF:116



Standard deviation confidence interval: [0.0221331023, 0.02866059369]

Variance confidence interval: [0.0004898742175, 0.0008214296308]

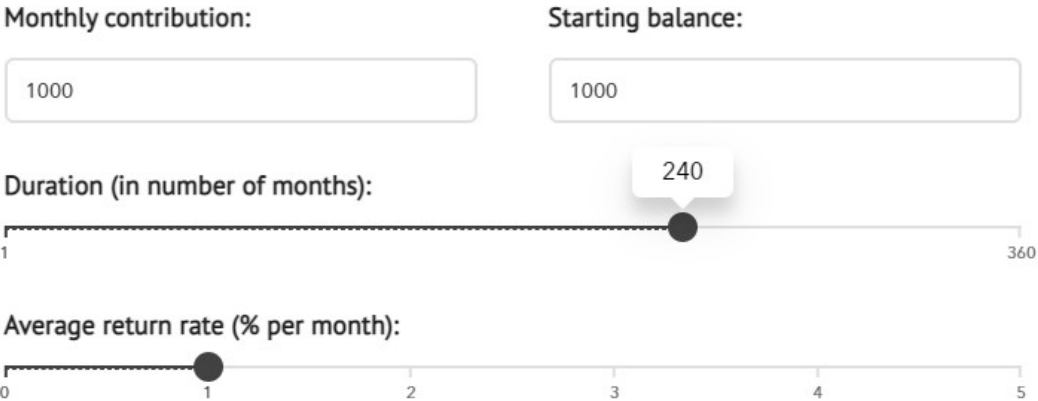
<https://www.statskingdom.com/confidence-interval-calculator.html>



Source: <https://tools.winton.com/thefuture/>

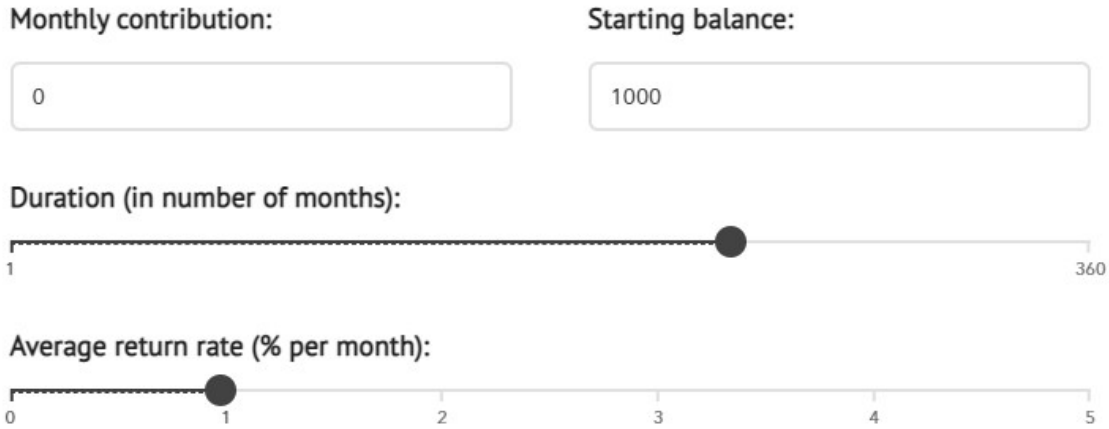
$$Future\ value = Starting\ Value \times (1 + rate)^{time} + \frac{contribution \times ((1 + rate)^{time} - 1)}{rate}$$

With contributions:



PS: Calculations made using the future value with regular contributions formula:
 $FV = StartingBalance \times (1 + rate)^{time} + Contribution \times [(1 + rate)^{time} - 1] / rate$

Without contributions:



PS: Calculations made using the future value with regular contributions formula:
 $FV = StartingBalance \times (1 + rate)^{time} + Contribution \times [(1 + rate)^{time} - 1] / rate$

Forecasted end balance:
1,000,147.92 \$

Forecasted end balance:
10,892.55 \$

3 STEPS TO IMPLEMENT THE PASSIVE STRATEGY?

Step 3: Time and opportunities

- Do not panic about crisis or negative news: **take advantage of opportunities and market conditions/cycles.**
- **Let the time (ie. compound interest!)** to act, protecting and monetizing your capital in the long run.

STEP3: DATA SCIENCE TOOLS & TECHNIQUES

- For recognizing opportunities: *Hypothesis testing, Control Charts, RSI-Relative Strength Index*

Method: Z-score

How to calculate the standard deviation?

Average = **0.00905**.

S = **0.0308**.

Lower = Average - k*S = $0.00905 - 3 * 0.0308 = -0.08326924627490881$.

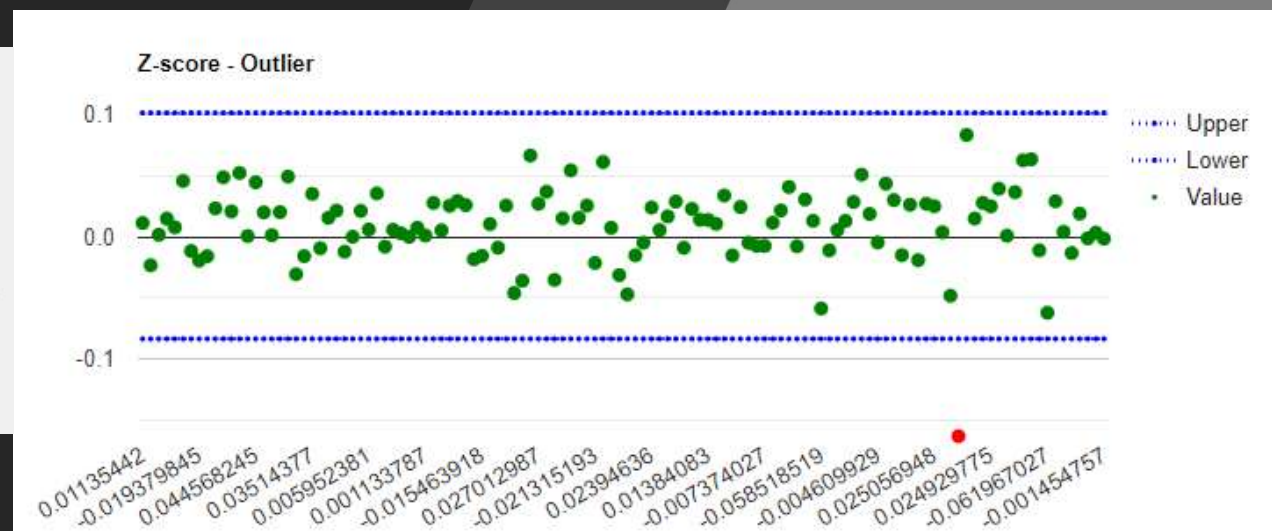
Upper = Average + k*S = $0.00905 + 3 * 0.0308 = 0.10136652064157547$.

Sample size (n) = **120**.

Outlier count: **1**.

Outliers: -0.162790698.

<https://www.statskingdom.com/outlier-calculator.html>



Conclusion

- Investing is a fascinating topic.
- Active (supported by Industry) x Passive (supported by Academia) Strategies is an interesting debate.
- Other disciplines, such as Data Science, can also be part of the investor's toolbox (in addition to Macro Economics, Accounting, Company Valuation, etc).

Arthur Valle, PhD

PhD in Production and System Engineering:
Process Mining

Six Sigma Black Belt

23+ years of experience in IT & Business

DIY Investor (since 2016)

Principal Academic Staff Member at Waikato
Institute of Technology, New Zealand

Founder of **TRENDSET**
(www.trendsetconsulting.com/en)

Disclaimer

I am not an authorised financial advisor.

This is not an investment recommendation.

Past performance does not guarantee future return.

www.trendsetconsulting.com

arthur@trendset.com.br