

[Open in app](#)[Get started](#)

Published in Towards Data Science

This is your **last** free member-only story this month. [Sign up for Medium and get an extra one](#)

Mario Emmanuel [Follow](#)

May 28, 2019 · 16 min read · ✨ · ⏰ Listen

[Save](#)

Backtesting trading strategies: less sorcery and more statistics on your side



[Open in app](#)[Get started](#)

later describe which are the key statistical parameters/figures that will lead us to meaningful conclusions. The impact of these figures on profitability and risk management will be covered in detail.

As it will be shown, simple statistics and metrics can be applied in a semiautomated way to provide quality insight even on discretionary trading, traditionally less prone to be statistically backed up.

Backtesting does not belong exclusively to automated trading

Some analysts already stated that the average holding time of an asset is as low as 22 seconds¹. This figure comes from studies in 2012, which could be already considered outdated.

Such a figure could be misleading² because there are different players operating over different timeframes, and while some of them hold assets for seconds or milliseconds, others hold assets for years.

In any case, it looks reasonable to state that the amount of automatic trading (high-frequency, algorithmic, quantitative or AI based) is growing and that the amount of discretionary trading is decreasing or losing relevance.

This trend can be confirmed both by large investment institutions trading floors³ and by the trading pits⁴, which are now either fully dismantled or operating at marginal levels.

This trend can also be found in retail trading. The approach is usually different from the one used by professional trading firms. The latest tend to rely more on statistics, factoring, regression models, neural networks and strategies based on coverages, pair spreads and network neutrality, while most of the automated trading done by retails is plain directional trading using basic indicator-based strategies which usually perform poorly on different market conditions.

Within this context, it might seem that discretionary trading days are numbered and that retails are doomed to failure. While this is true for most of the people (CFD brokerage



[Open in app](#)[Get started](#)

discretionary trading. Their results might not be replicable for large capitals, but they are good enough to provide a wealthy and lucrative job or even to run small funds or investment clubs.

As in any other occupation, success is determined by a combination of factors which include attitude, aptitudes, training, education, mentoring, experience, learning, commitment, mistakes and time. Most of the retails fail, that is a fact, but it is also a fact that most of them do not follow the necessary path to become a successful trader.

Backtesting is to trading what simulation/training is to other fields

I had, back in my late 20s, a brief opportunity to learn about the Aviation industry when I tried to become an Air Traffic Controller. From that period of my life, I got some takeaways about the safest industry in the world that I still find useful today.

In the Aviation industry, nothing is left to improvisation. Training — including simulation — is an integral part of the job for all roles. It does not matter if you are a pilot, an air traffic controller or a flight assistant. Everyone in the industry follows procedures and receive proper training including simulation to get exposure to both normal and abnormal conditions.

Checklists, cross-checking and procedures are the norms, and there is a correlation between aviation incidents and a miss to follow these procedures and checklists. This approach can be extrapolated to other industries, where there is no such strong focus on training and simulation.

“Who in his right mind would want to pilot a 747 without thousands of hours of flying experience and proper flight simulation time?”

Simulation is a key element in training. It allows exposure to adverse and stress conditions without being hurt and it develops the capabilities required to follow procedures and take



[Open in app](#)[Get started](#)

The benefits of backtesting in discretionary trading

As it came early in the article, backtesting needs to be an integral step of any strategy.

Waiving this point will simply cost money and can even ruin a trading account in a matter of days — or hours if you like leverage—.

While many people associate backtesting with automated trading (where backtesting can be easily programmed and it is, therefore, a natural part of the process), fewer people associate backtesting with technical analysis or price action discretionary trading.

The benefits of backtesting are:

1. It enables to fully understand the methodology/strategy, discovering issues that simply can not be covered by the mere description of the strategy.
2. It brings you down-to-earth regarding expectations and it will clearly show how the market is fully designed to wipe out small player's accounts.
3. It allows you to adapt the strategy to a particular asset/instrument.
4. It helps to determine filtering conditions that might impact profitability.
5. It provides a realistic assessment of our skills and capabilities to properly execute that strategy.
6. It provides real-world stats which will lead to realistic and detailed expectations on profitability, drawdown, risks and required capital.

While this is not rocket science and I am definitively not discovering anything new, backtesting is not always taken seriously. Many times it is waived, replaced with trading in demo accounts, poorly executed or executed without a proper commitment.

What tools and resources do you require to do backtesting?

The basic tools required are:

1. Quality historical data for the asset/instrument you want to trade



[Open in app](#)[Get started](#)

3. A full definition of the strategy (can be an iterative process as first simulation rounds might lead to fine-tune the strategy or define further conditions).
4. Quality time (several days) to complete the backtesting until at least hundreds (ideally greater than 1000) trades have been done.
5. An extensive analysis of the backtest, revisiting executed trades and trying to identify if there could be any correlated factor or scenario that could minimize losing trades.

All items are self-explanatory, although we will cover a bit more in detail the first two ones in the following paragraphs.

Is historical data an issue?

Data is no longer so hard to find. Of course, if your strategy is aggressive in terms of time frames and you are expecting to take decisions on real-time based on the one-minute candles evolution or your strategy uses tick charts, data will be expensive. But even in those cases some of those strategies can be approximated by 1-minute candles, much less expensive to obtain/purchase.

The important fact of data is that you need to understand that your timeframe and strategy will determine how much data do you need.

If you are testing an intraday strategy, a couple of years, probably with 1-minute candles (even 5-minutes it is a long intraday strategy) would be the minimum. Whatever data length you choose, the following conditions shall be meet:

1. The strategy must be tested in trending and non-trending market conditions. A strategy might work well in a continued bullish market but it will not perform well in lateral or bearish markets.
2. The strategy shall be tested against different market volatilities.
3. You shall plan to test 1000 operations. It can be hundreds too, but the larger the better.





Open in app

Get started

For manual backtesting, I use Forex Tester. It is an inexpensive tool, it has a perpetual licence model and it is extremely simple to use. It supports both tick and minute data.

This is the only manual backtesting tool I have tested, but I am sure that there are other tools perfectly capable of doing the job. The important point is that you need something that replicates the market with actual market historical data.



[Open in app](#)[Get started](#)

Backtesting software Forex Tester with a classic intraday Triple Screen configuration. In this particular setup a 1 minute intraday strategy is being tested using daily, 15 minutes and 1 minute charts of the FDAX symbol. The multiple screen setup mimics a live market setup where you can monitor scenario for different timeframes.

Special attention shall be put to scales (I personally try to fix grid scales so you can develop a better and heuristic understanding of how a specific asset/instrument behaves).

Templates shall be always the same ones and you can use whatever indicators help you, while many price action traders disregard indicators I personally think they can be a good confirmation aid and I tend to incorporate a standard 14-periods RSI and the venerable Connors 2-periods RSI, but this is just a personal preference.

Volume is always relevant if data comes from regulated liquid markets, over the counter volume such as the one found on CFD brokers might be misleading.

Simulation software will allow moving market forwards and backwards, either automatically or by just pressing a key. Switching from one timeframe to another is easy and you can quickly move forward when the market is ranging in an area which is not relevant for your strategy.



[Open in app](#)[Get started](#)

a few days. The same amount of information would require months in regular or demo trading. Not to mention swing or position strategies where trades can last weeks.”

It should be avoided to analyse past static charts. I have seen this many times, people just deploy past charts and validate strategy against the charts. The human brain is extremely good — and creative — at finding patterns, so just finding the patterns on past charts is not enough. An empty right side of the chart is needed to feel the strategy and understand how you perform under those conditions.

Additionally, a common mistake is to avoid analysing simulated trades — to avoid analysing trades in general — . You can learn from profit or loss trades and identify conditions that might lead to generating filters that improve profitability. It is better to try and post-analyse two hundred trades than just test trading one thousand ones. It is also important to correlate and annotate market conditions to each trade, i.e. you are expected to enrich profit/loss tabular data provided by the tool with additional market condition situation. This will allow to further enrich the statistical analysis and further refine the strategy.

Key statistics from backtesting

While discretionary trading has a reputation of being closer to sorcery than to science, there is nothing far from it to be true. Although a certain degree of art is required, it can be considered a more heuristic experience than esotericism.

The scientific part of discretionary trading comes from the rules of engagement (i.e. the trading plan, setups, market profiling, etc.) and from the statistical analysis of our methodology.

Statistics must be collected for backtesting, demo and live trading stages. Contrary to





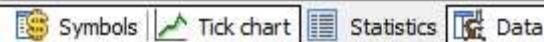
Open in app

Get started

Statistics



Parameter	Value
Time:	
Days processed	405
Months processed	13.32
Trades:	
Total trades	71
Profit trades	49
Loss trades	22
Profit trades cons.	6
Loss trades cons.	3
Trades / day	0.18
Trades / month	5
Profit trades / month	4
Loss trades / month	2
Max profit trade	56.83
Max loss trade	142.87
Income:	
Net profit	342.34
Gross profit	1031.02
Gross loss	688.68
Profit / month	25.71
Average profit	21.04
Average loss	31.30
Max drawdown	208.99
Profit factor	1.50
Return, %	6.85
Other statistics:	
Max lot used	0.01
Restoration factor	1.64
Reliability factor	0.12
Profit probability, %	69
Loss probability, %	31



Backtesting Statistics

The above figure shows the statistics provided by the tool. The whole trade details can be exported to Excel to further analyse the results, but even this basic set will provide enough



[Open in app](#)[Get started](#)

enough sampling data for a meaningful analysis of the strategy). The information is anyway enough to cover the key statistical figures and how they shall be interpreted.

Total Trades / Profit Trades / Loss Trades (relates to profitability)

This is the obvious count for the number of trades executed detailing the win and loss ones. the following information can be extrapolated from this basic count:

1. How many trades can you expect on average from the strategy, and based on the capital and risk management applied this will lead to the net profit that can be expected from this strategy. If our strategy is a swing/position which gives you 15 trades per year and you are risking 100€ per trade, you know you will not make a living out of that particular strategy. Also, the figure of 15 trades per year must immediately lead you to the conclusion that it needs to be backtested for years.
2. Profit/Loss ratios (also included in the statistical report). In this case, we are winning 69% of the times and losing 31% of the times. Note that there are profitable strategies with winning ratios of less than 50%, this data, combined with the average profit for the winning trades and the average loss for the losing trades introduces the concept of Average Profitability Per Trade (APPT)⁵, which is actually the cornerstone to determine if a strategy is profitable or not.

This is the formula for average profitability per trade, which basically states how much money is made out of each trade on average. A positive number will reflect a profitable strategy, a negative number will reflect a non-profitable strategy:

$$\text{APPT} = (\text{Profit \%} \cdot \text{Average Win}) - (\text{Loss \%} \cdot \text{Average Loss})$$

Profit/Loss Consecutive Trades (relates to risk management and psychology of trading)

Depending on the strategy followed you will face longer or shorter sequences of winning or losing streaks. Some strategies fail often than others incurring on small losses that are



[Open in app](#)[Get started](#)

The relevance of knowing in advance how many losing trades you can expect from a strategy is relevant to determine what to expect. Your account shall be properly sized to withstand not only this number of loosing streak but a larger one.

While this is just an indicative factor (simulation does not ensure future reliability) it is a good indicator on what to expect on reality. In this specific test, we had a maximum of 3 losing trades and a maximum of 6 winning trades. This can be taken as a reference so you shall be prepared to at least twice these figures.

The expected consecutive profit/loss trades will have also an impact on your psychology. Withstanding a long losing streak might be a hard experience, especially if you are stressing the amount of capital you are risking per trade. Knowing in advance what you might expect shall help you to cope with the turbulence and it would also help to detect when something truly deviates from the original plan so you can pause operations before it is too late.

Max profit trade / Min profit trade (relates to risk management)

Maximum profit per trade and minimum profit per trade can be misleading. Actually while preparing a strategy, it is good advice to prepare it for a given trade size. In this particular example, symbol stats (leverage and lot size) has been arranged so the profit and loss figures are given in FDAX points. I.E. we have selected a lot size so each point (win or lose) represents 1€. This allows later sizing operations to fit our risk. This means that you can later upgrade or downgrade the size based on your trading account and risk profile.

Based on the stats we notice that the maximum losing trade lost 142.87 DAX points.

This means that if you are accepting a 1% risk per trade (of your trading account alocated for that particular strategy) — which is a common risk assumption in trading — , you would probably properly sized at 1€/point for a 10.000€ trading account, as the worst trade made you lose 1.4%.



[Open in app](#)[Get started](#)

would also need to incorporate the current exchange rate of the pairs against the base currency of your trading account. “*Way too much hassle, don’t you think so?*”.

My personal view is that you do not need to be too strict with the figures; staying around an acceptable losing figure is easier to operate and allows you to operate fixed lots. Keeping your losses narrowed down is the important part of risk management. You can review sizing on a quarterly basis or simply based on equity curve mid or long term evolution. It is easier.

Note that while trading leveraged derivative instruments such as futures or CFDs, you need to think more in notional value than in actual trading account. You will be required to allocate much less money than the notional value of your investment, but you shall still understand that the leverage counts both for losses and profits. My personal preference is to think always in terms of notional values. Having started my operations on stocks that is the natural way I size operations, but I have met many people who have dealt only with leveraged instruments, and they tend to not fully understand the concept of notional value⁶.

It shall be also noted that the maximum percentage or risk per trade really depends on each individual. There are many ways of determining this value and it is not just related to your trading account but to your available capital and wealth. There are people with less capital but a strong wealth so their capital has relatively small relevance and people with more capital but a weak wealth so preserving their capital is paramount. Your risk profile also plays a significative role.

Blindly using the X % rule would lead to unbalanced risks depending on your specific situation.

The important point with risk management is to have something in place and to clearly understand what it means for your risk profile, wealth and personal situation. Which specific risk management shall be applied depends on every individual.



[Open in app](#)[Get started](#)

Bear in mind that despite most training courses and books stress on the need to keep average win/average loss positive ratios, this is not truly needed as what is relevant is the APPT sign. It is especially important to understand that many profitable short-term intraday and scalping strategies deal with negative profit/loss ratios. This might be related to the noise found in such smaller timeframes, which makes more advisable to deal with short-term momentum and higher trade % win/loss ratios than with positive average win/average loss ratios.

Max drawdown (relates to risk management and psychology)

Max drawdown simulates the maximum erosion experienced by your trading account during the backtested period. While this value ensures nothing, it can give you an idea on what to expect.

You must be definitively comfortable with experiencing such loss without stopping the process and it must fit your risk management policy. Some strategies — such as mean reversion strategies operating mid-caps, just to mention one — might experience extremely severe drawdowns, so it is important to understand your risk profile to feel comfortable.

Profit factor (relates to profitability and optimization)

Profit factor defines how much do you risk and how much do you get. It basically relates gross profits against gross loss. A ratio of 1.5 — like the one presented in this example — means that you are risking 1€ to get 0.50€ of profit. The higher the ratio the better.

Profit factor is often used to further optimize a strategy. A profit factor closer to one shall put you on alert.

Return % (relates to profitability)

The return % is basically how much return is received from the investment. It might be a bit misleading as it is calculated based on your backtested trading account, which might be oversized (or undersized).



[Open in app](#)[Get started](#)

particular example, one month of backtesting is not enough to conclude on the profitability of the strategy.

Max lot factor

If you are operating a variable size strategy, the max lot factor will tell you the maximum size employed. This strongly depends on how you configured the leverage levels of the instruments in the backtesting tool.

In this example lots and leverage well adjusted so each operation deals with 1€/point, hence the value is not meaningful. It would help if you operate variable lots in trades and you want to control risk by determining lot size (in CFDs you end up *feeling natural* what lot size is too much for your trading account).

Restoration factor and Reliability factor (relates to risk management and optimization)

The Restoration factor is calculated with the following formula:

$$\text{Restoration Factor} = (\text{Gross profit} - \text{Gross loss}) / \text{Max drawdown}$$

It is basically a figure of merit which quantifies the relationship between the gross profit against the maximum drawdown experienced. The higher the restoration factor the better.

The reliability factor relates the average monthly profit against the maximum drawdown:

$$\text{Reliability factor} = \text{Profit per month} / \text{Max drawdown}$$

We could try to simplify the statistical implication of both figures stating that they try to quantify how choppy the equity curve will be. As max drawdown is estimative they are just indicators they could be used to optimize parameters and get more friendly equity



[Open in app](#)[Get started](#)

While we will not cover how to enrich your tabular data, it has been already mentioned in the article that it is always a good idea to enrich data with information on the trade.

The ways to enrich the data is not constrained and it requires both experience and experimentation. If you are using moving averages or indicators you might like to include the status of such averages or indicators when the trade took place.

Account History													
Ticket	Symbol	Type	Lots	Open Time	Open Price	S/L	T/P	Close Time	Close Price	Comment	Swap	Points	Profit
55	fdax	buy	0.01	2015.01.29 09:14	10623.77	10591.40	10649.89	2015.01.29 09:43	10649.89		0.00	2612	26.12 ^
56	fdax	buy	0.01	2015.01.29 13:11	10662.38	10614.68	10694.18	2015.01.29 13:32	10694.18		0.00	3180	31.80
57	fdax	sell	0.01	2015.01.30 09:49	10754.98	10776.24	10740.80	2015.01.30 10:01	10740.80		0.00	1418	14.18
58	fdax	sell stop	0.01	2015.01.30 11:23	10718.18	10746.04	10699.78	2015.01.30 11:23	10748.00		0.00	0	0.00
59	fdax	sell	0.01	2015.01.30 19:39	10721.37	10765.60	10692.76	2015.01.30 19:55	10692.76		0.00	2861	28.61
60	fdax	buy	0.01	2015.02.02 07:53	10704.55	10693.67	10712.55	2015.02.02 07:59	10712.55		0.00	800	8.00
61	fdax	buy	0.01	2015.02.02 10:11	10709.09	10660.54	10744.39	2015.02.02 10:48	10744.39		0.00	3530	35.30
62	fdax	sell	0.01	2015.02.02 15:52	10732.30	10760.50	10713.49	2015.02.02 16:08	10760.50		0.00	-2820	-28.20
63	fdax	sell stop	0.01	2015.02.02 20:29	10823.36	10852.07	10804.38	2015.02.02 20:29	10870.00		0.00	0	0.00
64	fdax	buy	0.01	2015.02.03 07:41	10860.62	10849.21	10869.46	2015.02.03 07:46	10849.21		0.00	-1141	-11.41
65	fdax	buy	0.01	2015.02.03 13:55	10896.85	10851.59	10928.69	2015.02.03 14:41	10926.69		0.00	2984	29.84
66	fdax	sell	0.01	2015.02.03 19:11	10915.54	10944.57	10896.18	2015.02.04 07:46	10896.18		0.00	1936	19.36
67	fdax	sell	0.01	2015.02.04 15:13	10889.09	10934.90	10858.12	2015.02.04 15:30	10858.12		0.00	3097	30.97
68	fdax	sell	0.01	2015.02.05 08:19	10865.65	10905.53	10838.43	2015.02.05 08:47	10905.53		0.00	-3988	-39.88
69	fdax	buy	0.01	2015.02.05 15:03	10846.47	10799.87	10878.88	2015.02.05 15:26	10878.88		0.00	3221	32.21
70	fdax	sell	0.01	2015.02.05 15:54	10893.07	10927.01	10870.05	2015.02.05 20:59	10901.01		0.00	-794	-7.94

Backtesting software will provide you basic information that can — and shall — be enriched separately in an spreadsheet analysis

Further detailed analysis might be obtained based on trade lengths, direction (buy/sell) or period of the day. The objective is to find correlations between certain parameters and profit/losing trades.

I remember a set of statistics done for my forex trades (in that case they were statistics of live trades) where I learnt that I was more profitable buying than selling. While those findings must be taken carefully (it might be related to market conditions) they might also reveal biases in your operative. When properly analysed, filters or further sizing criteria can be applied.

The last but not least recommendation is to analyse the trades. This is something that is not always done (it is boring and requires a lot of time) but it might lead to finding operational mistakes and could help to further improve both strategy and execution.



[Open in app](#)[Get started](#)

The relevance of backtesting has been stressed and it has been shown how a basic understanding of the key statistical figures and metrics can benefit and improve the profitability.

The key figures and formulas to be applied to determine their impact on profitability and risk management have been highlighted and a final comment on how further improvements based on extended data could be achieved has been provided.

[1] <https://www.telegraph.co.uk/finance/personalfinance/investing/9021946/How-long-does-the-average-share-holding-last-Just-22-seconds.html>

[2] <https://www.marottaonmoney.com/what-is-the-average-holding-period-for-a-stock/>

[3] <https://nypost.com/2017/04/19/ubs-has-officially-ditched-its-massive-trading-floor/>

[4] <https://www.ft.com/content/4d221b22-3dfb-11e6-8716-a4a71e8140b0>

[5] https://www.investopedia.com/articles/forex/07/profit_loss.asp

[6] <https://www.cmegroup.com/education/courses/introduction-to-futures/about-contract-notional-value.html>

[Sign up for The Variable](#)



[Open in app](#)[Get started](#)

By signing up, you will create a Medium account if you don't already have one. Review our [Privacy Policy](#) for more information about our privacy practices.

 [Get this newsletter](#)[About](#) [Help](#) [Terms](#) [Privacy](#)

Get the Medium app

