WEEK 1

Trading vs investing

Financial sector firms are usually designated as either buy side or sell side. Buy side refers to firms whose primary business is investing and advising. This group includes asset managers such as private equity, mutual funds, life insurance companies, pension funds, and quantitative trading firms such as hedge funds. Sell side refers to banks and broker dealers that sell investments and services to asset managers and hedge funds and that also provide market-making services. That means they act as dealers in secondary markets. At investment firms, we find portfolio managers make longer-term decisions about the strategic allocation of assets and somewhat shorter-term decisions about the tactical allocations. Strategic allocation decides how much to invest in each asset category such as equities, bonds, real estate, or commodities. Tactical allocation chooses specific assets to buy or short within each category. The most common investment strategy is to buy and hold assets that had been identified as undervalued based on fundamental analysis.

Portfolio managers' gains and losses are measured relative to a relevant benchmark, portfolio, or index which mirrors their strategic allocation in their portfolio. This performance is measured net of any moves in the benchmark or index. So they can only outperform if their asset choices beat a passive portfolio that mirrors the benchmark.

Hedge funds on the other hand, we find traders, developers, and researchers all working together to identify and implement quantitative strategies. Their goal is to generate a positive return that is independent of overall moves in the market. This hedge return is called Alpha. The term Alpha is also widely used by portfolio managers to refer to their outperformance or return above a benchmark. This excess return seems similar to the Alpha generated by hedge funds, but there's an important difference. Portfolio manager Alpha comes from long asset holdings that are exposed to market, sector, and company risk as opposed to hedge fund Alpha which comes from a hedge strategy that has eliminated or at least attempted to minimize these risks. This is one of the key benefits of investing in a hedge strategy, especially if you're worried there'll be a sell off in the overall market.

Portfolio managers use fundamental analysis to rebalance the allocations in their portfolios. Rebalancing includes changes in the strategic allocation to give undervalued asset categories a heavier weight in the portfolio. It can also include tactical reallocation, where ideally you're selling winning assets that have achieved their full target value and replacing them with undervalued assets that have the potential to help the portfolio achieve returns in excess of its benchmark. At trading firms, traders have a much shorter-term opportunistic focus. The time frame for their investments ranges from a few months to a few milliseconds. A millisecond end of this spectrum dominates over all trading volume. Traders rarely use fundamental analysis as a factor in their decisions as they consider this information to already be baked in to the market price and essentially worthless for generating Alpha. Traders instead look constantly for market behaviors and inefficiencies that will generate high-risk adjusted returns on their trading capital. Although they trade the same assets as portfolio managers, they generally ignore fundamentals and focus on other sources of mispricing.

Buy-side quantitative methods include regression, prediction models, statistical arbitrage, and machine learning which we'll cover later in the course. Sell-side quantitative methods are mostly execution strategies. These are designed to reduce the market impact cost of large orders and also to capture spreads by providing liquidity through market making. Keep in mind though that buy-side firms also employ execution strategies when they execute their trading orders.

QUIZ

Which of the following are a feature of hedge fund alpha ?

- ○ Return in excess of a benchmark
- ○ Risk adjusted return
- ◉ Return once market and sector risk have been minimized

Correct
This is a hedged return alpha and would be a feature of a quant strategy such as matched pairs trading.

Order execution strategies are used by both portfolio managers and quant traders. What is the objective of this type of strategy?

- ☐ Break up large orders into smaller orders to minimize brokerage costs and exchange fees.
- ☑ Minimize market impact costs.

Correct
Market impact costs are are important for large orders as there is usually insufficient liquidity to fill orders at the current bid or ask price. These orders tend to get filled at worse prices as market makers lower their bids and raise their ask prices in response to a large order being shown all at once.

- ☑ Preserve privacy of strategy as long as possible.

Correct
Trading firms want to keep their strategies secret from other market participants. Breaking up orders can "camouflage" a strategy so it is harder for other traders to detect.

QUANT DETAILS

A successful quantitative strategy depends on good quality data which you can analyze statistically to uncover potentially profitable patterns and behaviors.

Algorithmic trading is a type of quant trading that uses pre-specified machine executable instructions to determine the size and timing of trades based on a quantitative model of an asset's price behavior. Over 70 percent of US trading volume is algorithmic. Most of this volume is high-frequency trading.

High-frequency trading is a subset of algo trading, focuses on the rapid execution of short term trading strategies at the millisecond and sub-millisecond timescale. We base or high-frequency trading strategies on quant models and market micro-structure features such as liquidity and latency

QUANT STRATEGIES

Quant strategies are a precise set of rules that generate orders and manage the risk of your current position. Rules can be quite complex and are a product of systematic statistical analysis of an asset's historical price data. You can use this analysis to uncover hidden patterns and behaviors in the market for a particular asset or group of assets such as an index.

Strategies range from low-frequency, at least by trading standards, to high-frequency. You must implement high-frequency trading strategies with software and market interfaces that can handle sub-millisecond order creation and submission. Quant strategies are also a type of forecasting method that attempts to predict the future value of a stock or other instrument where the direction of a price spread between two instruments. These forecasts are based on an observed price response to factors that you have identified as statistically significant in predicting changes in an instrument's price. Mean reversion strategies depend on the stability or stationarity of an asset's price or spread relative to another asset. When the price risk spread gets too high, it becomes more likely that it will decrease, and so revert to its long-term mean or neutral value. When the price risk spread gets too low, it becomes more likely that it will increase, and again revert to its mean.

Correlation is a measure of how well two variables move together over time. Correlation coefficients range from negative one or perfect negative correlation to zero, or no correlation to one which is perfect positive correlation. Positive correlation means that the variables move in tandem in the same direction while negative correlation means that they move in tandem but in opposite directions. When you look for a correlation in a price series, you normalize the prices so that each starts at 100 percent. The change in prices is then the cumulative return on each asset. Notice how the returns in this example track closely, and then diverge by about 10 percent, and then by about 15 percent. These two assets are correlated, but the difference between their means is not stable.

Co-integration tasks do not measure how well two variables move together, but rather whether the difference between their means remains constant. Often variables with high correlation will also be co-integrated and vice versa, but this isn't always the case. In contrast to correlation when you test for co-integration, use prices rather than returns since you're more interested in the trend between the variable's means over time than in the individual price movements.

In momentum or trend-following strategies, you're just buying assets that had been past winners and, selling assets that had been past losers. This is the opposite of a mean reversion strategy where you sell winners and buyback losers. Momentum strategies has been particularly popular over the last five years, and they outperformed the S&P 500 by a wide margin as you can see from this chart, which compares the performance of the iShare momentum ETF to the S&P. We will find possible explanations for this outperformance in both the underreaction and overreaction of prices to new market information. Underreaction means it takes time for the market to fully incorporate new information, and so the positive or negative effect is spread out over a longer time period. Overreaction means that the market tends to feed on its own positive or negative reaction. The market over rewards companies that release good news, and over punishes those that release bad news. You can also argue that it is inherently riskier to continue buying stocks after up-moves or selling them after down-moves. So momentum investors should get an extra return for taking this risk. This argument, however, is undercut by the fact that iShares Momentum ETF has a three-year beta of 0.91 versus 1.0 for the S&P 500. It is weakly supported however by the volatility of the momentum ETF being 12.5 percent versus 12.1 percent for the S&P. I think that the last argument is the strongest. Investors believe that momentum strategies will outperform the market, will cause them to bid up the prices of momentum stocks creating a positive reinforcement cycle.

You can also argue that it is inherently riskier to continue buying stocks after up-moves or selling them after down-moves. So momentum investors should get an extra return for taking this risk. This argument, however, is undercut by the fact that iShares Momentum ETF has a three-year beta of 0.91 versus 1.0 for the S&P 500. It is weakly supported however by the volatility of the momentum ETF being 12.5 percent versus 12.1 percent for the S&P. I think that the last argument is the strongest. Investors believe that momentum strategies will outperform the market, will cause them to bid up the prices of momentum stocks creating a positive reinforcement cycle. Many traders believe that momentum stocks and also quality stocks are overvalued, and that this positive reinforcement cycle will unwind in a sharp sell-off of the stocks favored by these strategies. Market micro-structure is a combination of the physical trading infrastructure used by a trading platform and its participants, the platform's trading rules, and the behaviors, and trading patterns of its participants. The strategies used and abused by high-frequency traders had been described extensively in the media and books such as Flash Boys. High-frequency traders who attempt to capture an arbitrage profit by taking advantage of flaws and anomalies in this infrastructure that exists for a few milliseconds or less. They also seek to identify and exploit patterns in the trading behavior of their major competitors in each market. These patterns are often identified using game theory models. Equally important is their ability to detect and profit from the execution of large orders that have been broken up and hidden using an order execution strategy. Let's say a firm wants to purchase an unusually large quantity of stock. Now use an execution strategy that enters the trades in a way that is supposed to break up and hide the order but which unintentionally signals search strategy. Sophisticated traders have their own algorithms which are able to detect these stealth orders. They're able to use this information to jump in front of the large orders, and profit by driving prices up ahead of a buy order or down ahead of a sell order. Spoofing is where traders attempt to mislead other traders by submitting orders to the market, order book that they don't intend to execute. Say a trading group needs to sell a stock and wants to get a better price. They can enter temporary large-sized buy orders below the current market bid to encourage other bidders to pay higher prices for the stock.

QUIZ

We have been talking about correlation and cointegration. Please select the statements below that are correct

- [x] Cointegration is the same as correlation except that cointegration is the correlation of the means of two price series.

This should not be selected
Cointegration indicates that the difference between the means of two price series is stationary (reverts to a long-term mean). A cointegrated series series can also be correlated but this is not required. Also, the mean of a price series is a single number, so the means of two price series cannot be correlated. All we can measure is their difference.

- [ ] Cointegration and correlation both imply that the two price series are stationary.
- [x] Cointegration indicates that the difference between the means of two price series is stationary. Correlation implies that the returns from two prices series are related, either positively or negatively.

Correct
Cointegration indicates that the difference between the means of two price series is stationary (reverts to a long-term mean). A cointegrated series series can also be correlated but this is not required. Correlation indicates that a price series has some relationship with another price series. This relationship can be very strong in the case of perfect positive (+1) or negative (-1) correlation or non-existent in the case of zero correlation.

Quant trading has several advantages over discretionary trading as it is objective and takes the emotion out of strategy and risk management decisions. That said, it also has some challenges (many of which are shared by discretionary trading). Please look at the list below and select the statements that reflect these challenges.

- [ ] Quant strategies are consistently profitable.
- [ ] Quant strategies are optimal for all market conditions.
- [x] Quant strategies must be constantly adapted to new market correlations and volatility.

Correct
Quant strategies are only optimal for the market conditions present in the data that was used to create and test the strategy. If correlations or volatility change, strategies must be modified to have a chance of being profitable.

Which type of arbitrage is used in Pairs Trading

- (•) Statistical Arbitrage
- ( ) Exchange Arbitrage
- ( ) Carry Arbitrage

Correct
This is correct. Pairs trading relies on a statistical relationship, usually correlation, between two assets (stocks).

QUANT TRADING: ADVANTAGES/ DISADVANTAGES

Markets are often driven as much by emotions as by fundamentals. While it is sometimes important to factor in market sentiment, it is usually not a good idea to let your own emotions color your trading decisions. In my experience, trader sentiment swing wildly between greed and fear, blinding them to what is really going on in the market. Greed and its enabler egomania often cause a trader to refuse to take profits and rides huge winners all the way down to huge losses. Fear and its enabler self-doubt often cause traders to liquidate positions due to short-term losses that would have eventually turned into big winners.

Exchange and Statistical Arbitrage

Exchange arbitrage is a form of riskless or pure arbitrage in which the same asset trades on two exchanges with the asking price on one exchange overlapping with the bid price on another exchange. Normally, the bid price is lower than the ask price. In the example you see here, the asking price on Nasdaq is 100 and the bid on NYSE is 100.10 cents, or an Arb profit of 0.10 cents per share. The opportunity to make this type of riskless profit is becomes scarcer as links between exchanges and alternative trading networks gets faster, which reduces the time window in which arbitrage opportunities like these are available. Also, the increase in high-frequency trading firms specialized in low latency strategies or very fast order execution strategies has reduced the profits available from this type of arbitrage.

Carry is the cost of holding and storing an asset, net of any income that the asset pays you. Unlike bonds and stocks, gold only pays a yield if it is leased out, and the rate for a one-year lease on gold is only about 10 basis points right now. The difference between the spot price and the future delivery price is the compensation of the carry trader receives to offset their costs which are funding, storage, and insurance. In the example shown, the total carry is $30 or about two percent of the asset's price. This would be less by 10 basis points if you're leasing the gold out.

Stat Arb is the most common type of risk Arb strategy and mean reversion is a more formal way of saying buy low, sell high. In this example, you want to trade a stock whose average price has been 100. The issuer is a mature company with limited growth opportunities and you assume the share price will be more or less stable, or more formally, you can model it as a stationary process. This is an unusual assumption as most stocks are assumed to have an upward price script due to either profit growth or retaining earnings. The variation of the stock has been five percent over the previous 100 days. So you expect it to vary by the same amount in the future.

You create trading rules so that when the stock increases to five percent above its mean, you sell or short it. When it falls to five percent below its mean, you buy it.

Using these rules, you could earn a total profit of 20 based on the buy and sell points marked on the graph. So far so good. But what happens is volatility is greater than expected. What happens if you short at 105 and then the price goes to 110? You sell more stock at 110 or do you lock in your loss by covering the short at 110? Similar problem. If you buy the stock for 95 and then the price drops to 90, giving you a five percent loss. Do you buy more at 90 or do you lock in your loss by selling it 90? This is the unavoidable challenge of Stat Arb strategies. If you set your buy and sell levels too close to the current market, you will likely make a transaction, but you're going to have a low level of confidence that it will end up being profitable. If you set your buy and sell levels too far from the

current price, you're less likely to take a loss, but it's also less likely that the market is going to reach those levels. So you may never make a trade.

Designing a profitable strategy means finding the balance between these extremes and also setting stop-loss orders to preserve your trading capital. Another common Arb strategy is pairs trading, where you identified two stocks with similar characteristics and go long one and short the other. With this example, we assume the stocks are trading at the same price and you were able to finance your purchase of stock A with the proceeds from shorting stock B. In real life, you would have to borrow stock B and you'd have to post part of the proceeds as collateral, but we're just going to assume you're able to finance the purchase completely with stock B. The main benefit of this strategy is that since A and B trade in the same market, belong to the same sector of the same industry, you are hedging the industry and sector risks. So any return that you get will be based only on the strength of the choice you make of which stocks to pair.

Your decision about which stock to buy and which stock to sell could be based on either a fundamental evaluation analysis, or on relative historical returns, or correlations, or some combination of all three. You might determine that the fundamental value of A is $10 more than B despite them having the same market price. Or you might have observed that B is outperformed A by 10 percent historically, and going forward, you expect the price of A to catch up to B because you expect their returns to be the same in the long-term. So you buy A for 100 and you sell B for the same price. You are now long the spread between A and B. Almost immediately, A share price increases to 110 while B stays at 100. You sell A and buy B back for a quick $10 profit, and then wait for A and B to be mispriced again which happens when A and B both fall to 95, and you get longer spread again by buying A and shorting B. This time, A's price drifts and ends up back at 95. Meanwhile, B has rallied to 105 causing you to lose $10 on your short. If you panic and issue a stop-loss order to sell A and buyback B, you will lock in a loss of $10. If instead you decide to write out the trade, you will see your loss increase to $15 when A stays at 95 and B rallies to 110. It would be hard not to cover the trade now as you've lost 15 percent of your capital and you will get pressure from your funders. If you're able to talk them into letting you keep the trade on, your 15 loss eventually transforms into an unrealized seven dollar profit when A rallies to 107 and B falls to 100. By the end of the period, stock A is trading at 83 and stock B at 103 and you have an unrealized loss of $20 on the trade and an overall loss of $10 for the year. Your funders are not happy and you are wishing you issue that stop-loss order or even better had taken that unrealized profit of seven dollars and ended the year up 17.

Correlations between stock pairs are notoriously unstable and can move from positive to negative, and then back again depending on what time period you're looking at. This makes pair trading a fairly risky proposition even though it's considered fully hedge except for the company or spread risk. Other trade such as pairs of bonds with different maturities and commodities paired with their byproducts have more stable correlations and are potentially better candidates for pairs trading, which is essentially just trading a spread between the price of two related assets.

INDEX ARBITRAGE

In this section, we will explore one of the most popular Arb strategies: Index Arbitrage. The original exchange traded funds or ETFs were created to mirror the most heavily traded indexes: the S&P 500, Nasdaq 100, and the Dow 30. As volume of these ETFs increased and passive investing became more popular, the number of ETFs increased to more than 6,000 worldwide by 2018 of which 3,500 were equity-based. ETFs currently manage over $5 trillion of assets and are traded actively by both institutional and retail traders. Index Arbitrage is a high-frequency strategy. Indexes are composed of multiple stocks or bonds each of which is traded separately or as part of an index ETF. Traders need to know the current bid, offer, and size of orders for each index component, and also the weighting of each component in each index in order to know exactly what price they can buy and sell the index in the market. They then compare that with the best bid and offer for the index ETF and see if there's an arbitrage opportunity.

For example, you could buy shares of all 100 components of the Nasdaq 100 in exactly the right proportions and then sell shares of the QQQ ETF. This may seem like a lot of effort for a small profit, but it's the main business of lead market-makers for the QQQ who are constantly converting mirror portfolios into QQQ shares and then converting QQQ shares into their component's stocks. Index Arbitrage also has high-infrastructure costs due to its need for real-time market data feeds, high-speed computers, and co-location with exchange servers. Co-location reduces the time between when your order execution program receives market data and the exchange receives your trading orders. This is called reducing data and execution latency.

QUIZ

Index arbitrage trading makes a profit by

- ○ Using exchange arbitrage to trade component stocks of an index that have bid prices that are higher than ask prices on two different exchanges.
- ○ Using cointegration between two indexes to decide when to buy one index and sell the other short.
- ◉ Buying the component stocks and selling the index ETF or future when the component stocks are cheaper than the index ETF or future.

Correct
This is the foundation of index arbitrage which also includes selling the component stocks and buying the index ETF or future when the component stocks are more expensive than the index ETF or future.

STATISTICAL ARBITRAGE OPPORTUNITIES AND CHALLENGES

Stat Arb is very competitive and it's hard for new entrants to compete in areas like index arbitrage especially in major markets. Fortunately, there are lots of opportunities in smaller markets and less liquid stocks and other assets. These markets are less efficient, and so they offer more opportunities to trade profitably but on a smaller scale. Stat Arb profits are eaten away by a range of costs including trading, clearing, and exchange fees. These are especially heavy for high-frequency traders. There are also risk-based charges which factor in the cost of trading capital being tied up in collateral and margin for derivatives positions in short sales. Added to this are the cost to borrow assets for short sales, especially when there's very high demand and limited supply as is the case when several hedge funds decide to short the same stock of a closely held company.

Think Beyond Meat or Lyft. This is referred to as the asset being on special, where the stock lender charges you as much as 60 percent to borrow the stock. This means you put up collateral of $100 and receive $40 back at the end of the year. Obviously, this is not a good strategy for more than a few weeks. Lastly, there's a market impact cost of trading. Lower frequency arbitrators usually sell at the bid price and they buy at the ask price. So they lose about half of that bid-ask spread on each transaction. This is paying for liquidity or being a liquidity taker rather than a liquidity provider such as a market-making firm. Arb trading firms consume lots of data. Initially, a firm must purchase and cleanse historical data that it will use as the input to models and strategies. Market feeds vary in price depending on the level of detail needed by the firm. Daily open, high-low, and closed data is essentially free. High frequency data is more costly, but is essential for most quantitative strategies.

High-frequency traders also need to have full visibility into each asset's order book and tick by tick data on the timing and size of transactions. This type of feed can be quite expensive, especially if a firm trades in several markets at the same time. Remember my earlier observation that the most promising areas for traders were smaller, less liquid stocks and markets? This comes with a big caveat. In order to profit from these stocks, you need to actually be able to trade them. The reality is that liquidity is heavily concentrated in the largest cap stocks in the largest markets. This is part of why these markets offer fewer mispricings and other inefficiencies. When you design a trading strategy, you need to take into account how much trading actually takes place in the stock and what size of bid-ask spreads you will face if you decide to transact. These hurdles can make an otherwise profitable strategy. You're unexecutable or unprofitable.

Another challenge is it's sometimes hard to borrow certain stocks. Recall our discussion of stocks being on special. In this case, you may not be able to find a stock lender when you want to shorten them. Also trading in stock may be halted due to pending news or simply because the market is not open. This can cause you to lose a lot of potential profit from a strategy. In our earlier example of spread trading of pair stocks, we had to wait a long time before the spread reached a level where we can sell it, and then another long time before it reached a level where we can buy it back again. During these waiting periods, correlations can change dramatically due to new information or events in the market. The longer your time frame, the more sensitive your trade is to these risks. This type of risk is generally not an issue for high-frequency traders who focus on ultra-short time frames. Instead, they have to worry about covering high trading costs with frequent but usually small profits.

One of the most important things you need to keep in mind as you develop strategies is that when you trade a correlation or a spread, you can also change the market you're trading in. Prices will change immediately and other traders may withdraw liquidity if they sense a spread is being traded more heavily in one direction. They may also be able to figure out your strategy even if your trades are anonymous by noticing changes in price patterns in a asymmetries in the order book. Ultimately, if your trades are large enough or persistent enough, they will eventually eliminate the profitability of the correlation or inefficiency. This is especially true when other terms start trading the same strategy.

BACKTESTING

When I first started trading, I thought trading was all about coming up with clever strategies and then, just to be sure, testing them on historical data. This is the complete opposite of how strategies are actually developed. First, you collect and cleanse your data. Then you use the development back tester to identify potentially profitable correlations and behaviors in the data. After you rank these from most promising to least, you validate the best ones using an implementation backtester. This is a long and somewhat tedious process, but gives you a much higher probability of making money once the strategy goes live.

First, we will explore how developers validated quantitative strategy that has been created using a historical dataset. A classic backtest starts with a time series of historical asset prices that is divided into two series: a training series A and a testing series B. You train your model using data from the training window, and then you apply that model on data in the testing window. This allows you to compute the performance of your model with data that was not used in its creation. Backtesting is part of an iterative process in which you first use a series of price data to fit a mathematical model of predictable price patterns and behaviors. You then use this model of patterns and behaviors to design a quantitative strategy that will allow you to initiate orders to buy and sell the asset and also to set stop loss orders to manage position risk.

You optimize a strategy to maximize the profit within your risk tolerance, from trading the price series in period A. It wouldn't make sense for you to validate a strategy with the same data you use to create it. So you reserve part of your historical data series to test your strategy out of sample. Backtesting at its simplest is just running your strategy on the price data from period B, and seeing if you can replicate the performance of your strategy achieved in period A. If performance in the backtest is negative or extremely weak, you abandon the strategy and start over. If it is positive but does not meet your minimum risk return requirements, you might be tempted to tweak your strategy and retest it on period B data. Although tweaking seems reasonable, you run the risk of overfitting your model to the data. If the strategy's performance in period B meets or even better exceeds period A results, then you are ready to either go live with your strategy, which is a bit risky, or more commonly paper trade the strategy for a while with live data and then gradually commit trading capital if it continues to generate Alpha.

A backtest is a simulation of the performance you would have achieved if you had traded your strategy during a particular historical period. It includes trading and risk management rules to simulate a live trading environment. You can then compare those performance to that of other strategies implemented on the same asset. Last week, backtest give you an estimate of the capital you will need, risk involved, and transaction costs you could incur if you decide to live trade a

strategy. Another type of task that we commonly use for time series models that have been built using machine learning methods is the sliding window backtest. In a sliding window test, the whole dataset is divided into a series of adjoining pairs of training and testing windows. You train a model using a training window and then apply that model on the adjoining testing window to compute performance through the first run. For the next run, you slide the training window to the new set of training records and repeat this process until you've used all of the training windows. With this technique, you can calculate an average performance metric across the entire dataset. Performance metrics you derived through sliding window validation are generally more robust than split validation techniques that we discussed earlier. Notice how the training window or the testing window moved together across the entire time series of data. This allows your strategy to be modeled and tested on almost the entire dataset rather than just half of it.

BACKTESTING DESIGN

When you design and implement a backtest, it is important that you have a clear idea of what you want the backtest to accomplish.Your goal at the initial development stage is to set up a strategy pipeline and then filter out any strategies that don't meet your minimum performance criteria. Backtesting provides an efficient filtration mechanism so that you can eliminate week strategies that have a low risk adjusted return, or that have large anticipated drawdowns. You can use Sharpe ratios and Calmar ratios to rank your strategies, but keep in mind that risk and drawdown may be higher when move to a live trading strategy. Backtesting also allows you to test your model and fine-tune without risking trading capital. It helps you to identify microstructure issues early on, such as transactions costs, liquidity, and latency and data in order execution. Backtesting provides a framework in which you can optimize your strategy by modifying the quantities or values of the parameters associated with that strategy, and then recalculating its performance. Unfortunately, by optimizing your strategy, you run the risk of introducing overfitting bias into your model

Now, we will talk about some of the traps that many quantitative strategy developers fall into when backtesting their models. Models depend on correlations, which can vary or even reverse given enough time. Backtesting takes the correlation in the test data series as given. At a minimum, you should measure the value and variability of critical correlations over your entire data set. Data can be incomplete or contained fictitious outliers. You need to carefully cleanse your data before using it in a backtest. You also need to decide if the frequency of data that you have available is sufficient to test your strategy. For long-term strategies, open, high, low and close data may be sufficient. For high-frequency strategies, you need tick by tick data and visibility into the order book.

Slippage is the difference between the last market bid or offer in the price you actually buy or sell at. Your model needs to factor in getting somewhat worse prices, especially in volatile markets when using market orders. You also need to model realistic market impact costs. It's not realistic to assume that you'll be transacting at the midpoint between the bid and the ask price. This is especially critical in markets with limited or variable liquidity. It is easy to allow biases to weaken your backtesting model and strategy.

One bias that we've already mentioned is overfitting your model to a limited data set. This makes your model perfect for the backtest data, but also very likely to have completely different performance when tested in live market.Look-ahead bias sneaks into your model whenever you inadvertently include data outside of your testing window. This can happen when you tweak factors

and correlations to reflect more recent data. Survivorship bias occurs when you use data sets that don't include the stocks of companies that have either gone out of business or been merged into another company. This is when you are backtesting your strategy on a number of different stocks that have all survived to the present, and so represent a positively skewed sample. Drawdown tolerance bias is a common feature of discretionary trading but is mostly ignored for quantitative trading. When you create a backtest that covers a period of five years or more, it's easy to look at an upward trending equity or equity index price curve. Calculate the return, Sharpe ratio and drawdown characteristics and be completely comfortable with the risk you're taking. The strategy might have a maximum drawdown of 25% and a maximum drawdown duration of three months. This would not be unusual for a momentum strategy and it's easy to convince yourself that you can tolerate these losses because the return potential is so good. In practice, a drawdown of 25% extended over three months would raise some serious doubts in your mind about the strategy. Your funders would probably intervene and pressure you to end the strategy and take a loss.

Now we look at the difference between backtest you create in the development phase and those that you used to retest strategies that are candidates for full implementation. You design development backtesters to help create many preliminary models. So strategies with varying sets of parameters can be tested quickly and either discarded or qualified for a more thorough implementation backtests. They are used to identify statistical relationships and market behaviors that have the potential to be exploited in a quant strategy. One big weakness is they tend to give inflated performance, and so you don't consider the full cost of strategies. You can use our pandas to code development testers and it's generally simpler to code these, than decoded implementation backtester.

Implementation backtesters are more difficult to code because they simulate all aspects of a strategy going live. Implementation backtesters require a data interface to handle both historic and live market data, a quant strategy that generates signals to a portfolio manager that then creates orders and measures performance. And then an order execution handler that sends orders and receives confirmations.

QUIZ

Smaller markets offer more opportunities for statistical arbitrage but also have some unique challenges such as

- [x] Lower liquidity

Correct
Yes, volume traded in smaller markets is much lower and bid-ask spreads are much wider.

- [ ] Small bid-ask price spreads
- [x] Low trade frequency

Correct
Yes, low trade frequency, sometimes fewer than once per day, can make trading these markets difficult.

The advantage of using a sliding window backtest over a split window backtest is that

- [x] You are able to use more of the available data to train your model

Correct
Yes, in a sliding window back test you are able to use almost all of your data set to both train **and** test your model

- [x] You are able to use more of the available data to test your model

Correct
Yes, in a sliding window back test you are able to use almost all of your data set to both train **and** test your model

When designing your trading model it is more realistic to assume that you will

- ( ) Buy at the bid price and sell at the ask price
- ( ) Buy and sell at the midpoint between the bid and the ask price.
- (•) Buy at the ask price and sell at the bid price.

Correct
Yes, quantitative traders should realistically expect to sell at the bid price (or worse for large market orders) and buy at the ask price (or worse for large market orders).

Please identify some of the biases that can weaken a backtest below.

- [x] Look-ahead bias

Correct
Yes, this is where your trading strategy inadvertently uses future data (either in sample or out of sample) to improve its performance.

- [x] Look-back bias

This should not be selected
No, looking back is not a bias as you must use historical data to train and test your model.

- [x] Optimization bias

Correct
Yes, optimization bias is where you keep adding variables to maximize performance to your training data but end up with an over-specified model that doesn't perform well on testing data or in live trading.