Python for Finance: Course Notes

365√DataScience

When we think of an investment, we have to remember two things – **its upside and its downside.** In other words, we should consider the profit that will be made if everything goes well and the risk of losses if the investment is unsuccessful.

Type of asset	Definition
Government bonds	Government bonds offer an average rate of return of 3%, and historically, there have been very few cases of governments going bankrupt and not repaying what's owed to investors. So, some risk comes with this investment; it isn't risk-free, but the risk is very contained.
Equity shares	Equity shares have a higher rate of return – approximately 6%. However, they are associated with much more frequent fluctuations and price changes, as different variables influence a company's share price.

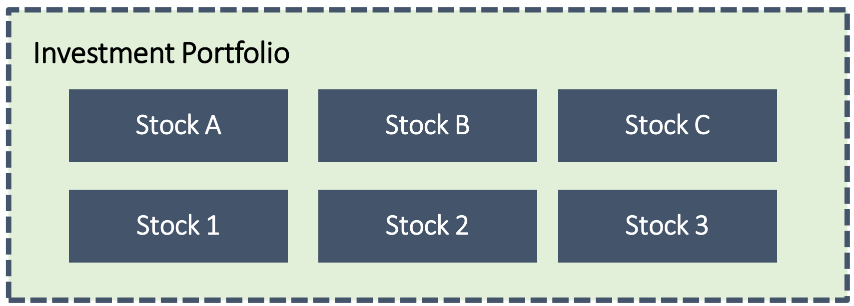
The art of finance isn't about maximizing an investor's returns in a year. It is about **making informed decisions that consider both dimensions, risk and return**, and optimizing the risk-return combination of an investment portfolio.

Calculating rates of return

Term	Definition
Simple returns	(Ending Price – Beginning Price) / (Beginning Price) Preferable when you have to deal with multiple assets over the same timeframe
Log returns	Log (Ending Price / Beginning Price) Preferable when you make calculations about a single asset over time

Portfolios of stocks

Most investors own several stocks, and the set of stocks that an investor owns is called his **investment portfolio.**



Calculating the **rate of return of a portfolio** is an easy and intuitive task. We have the rates of return of individual securities, and we only have to multiply each security's rate of return by the weight it has in the overall portfolio.

$$r_p = w_1 r_1 + w_2 r_2 \ldots + w_n r_n$$

Market Indices

A market index provides an idea about how a given stock market is performing. It represents a large enough sample of the overall number of stocks in that market and can be considered a good enough proxy of the overall development of the market.

Market Index	Description
S&P500	Comprises 500 of the largest listed companies. The diverse constituency of the S&P 500 makes it a true approximation of the US stock market. It is a market-cap-weighted index, so companies are weighted according to their market value
Dow Jones Industrial Average	The Dow Jones industrial average index uses an average of 30 large public stocks traded in the US market
NASDAQ	Most companies that are part of the NASDAQ index are information technology companies, and NASDAQ gives us an idea about the general development of the tech industry

A stock index gives you **a sense of the type of return** you can expect if you invest in a well-diversified portfolio in a given market.

Variability plays an important role in the world of finance. It is the best measure of risk we have. A volatile stock is much more likely to deviate from its historical returns and surprise investors negatively.

$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1}$$

Sample variance

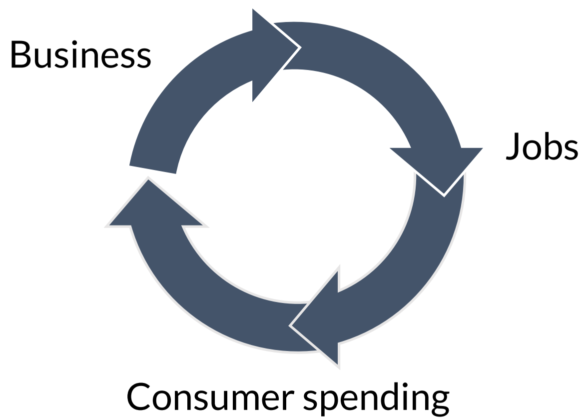
$$s = \sqrt{s^2}$$

Standard deviation

Commonly used statistical measures, such as variance and standard deviation, can help us a great deal when we try to quantify risk associated with the dispersion in the likely outcome. Such dispersion is measured by a security's variance and standard deviation.

The relationship between securities

It is reasonable to expect **the prices of shares in a stock exchange are influenced by common factors.** The most obvious example is the development of the economy. In general, favorable macroeconomic conditions facilitate the business of all companies.



Companies' shares are influenced by the state of the economy. However, different industries are influenced in a different way. Some industries do better than others in times of crisis.

There is a relationship between the prices of different companies, and it is very important to understand what causes this relationship and how to use this measurement to build optimal investment portfolios.

Calculating Covariance:

$$\sigma_{xy} = \frac{(x - \bar{x}) * (y - \bar{y})}{n - 1}$$

Calculating Correlation:

$$\rho_{xy} = \frac{(x - \bar{x}) * (y - \bar{y})}{\sigma_x \sigma_y} \qquad \rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

Type of correlation	Description
Perfect correlation	Equal to 1. The entire variability of the second variable is explained by the first variable
Negative correlation	It can be perfect negative correlation of -1 or much likelier an imperfect negative correlation of a value between -1 and 0
Neutral correlation	A correlation of 0 between two variables means they are independent from each other

Calculating Portfolio Variance

If a portfolio **contains two stocks**, its risk will be a function of the variances of the two stocks and of the correlation between them.

$$(w_1\sigma_1 + w_2\sigma_2)^2 = w_1^2\sigma_1^2 + 2w_1\sigma_1 w_2\sigma_2\rho_{12} + w_2^2\sigma_2^2$$

$$w_1 \ = weight\ of\ Security\ 1$$
$$w_2 \ = weight\ of\ Security\ 2$$
$$\sigma_1 \ = standard\ deviation\ of\ security\ 1$$
$$\sigma_2 \ = standard\ deviation\ of\ security\ 2$$
$$\rho_{12} = corrleation\ between\ security\ 1\ and\ 2$$

Type of risk	Description
Systematic risk	This is the uncertainty that is characteristic of the entire market. Systematic risk is made of the day to day changes in stock prices and is caused by events that affect all companies
Unsystematic risk	These are company-specific, even industry-specific, risks that can be smoothed out through diversification

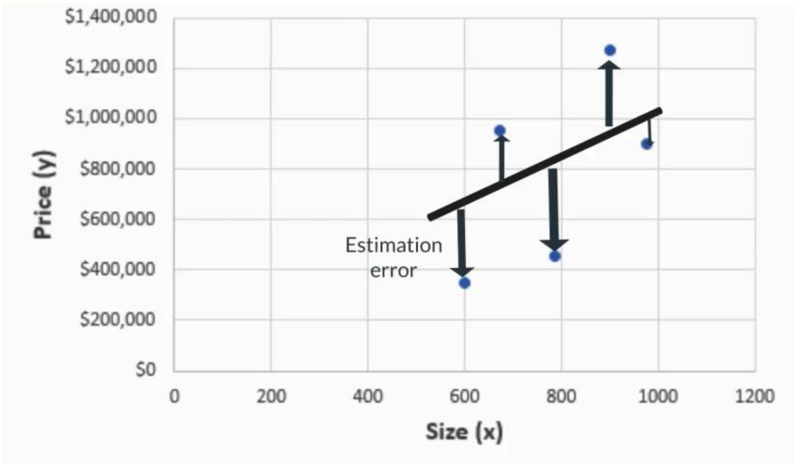
Regression analysis

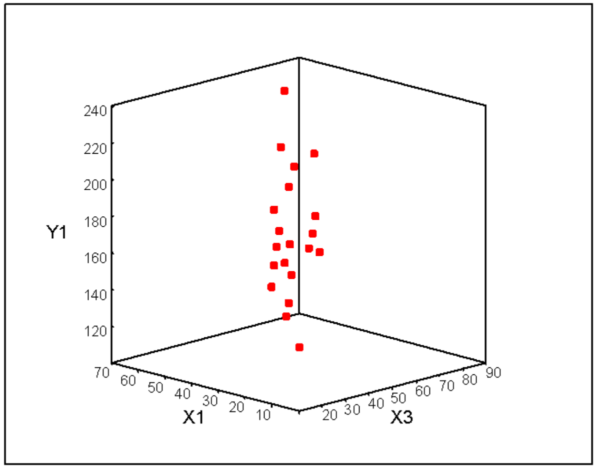
Regression analysis is **one of the most frequently used tools** in the world of finance. It quantifies the relationship between a variable, called dependent variable, and one or more explanatory variables, also called independent variables.

Type of regression	Description
Simple regressions	Yi=β0+β1 X1+εi Regression analysis assumes the existence of a linear relationship between the two variables. One straight line is the best fit and can help us describe the rapport between all the data points we see in the plot
Multivariate regressions	Yi=β0+β1 X1+β2 X2+β3 X3+εi By considering more variables in the regression equation, we'll improve its explanatory power and provide a better idea of the full picture of circumstances that determine the development of the variable we are trying to predict

Regression analysis

In order to calculate regression coefficients, we are estimating a best fitting line minimizing the sum of squared residuals.





Minimizing residuals (the amount of estimation error) in a **simple regression setting**

Minimizing residuals (the amount of estimation error) in a **multivariate regression setting**

Simple regression

Y = 1,000+0.8*x1

Multivariate regression

Y = 600+0.8*x1+0.4*x2+0.3*x3

How good is a regression?

Statisticians have come up with a tool that's easy to understand. **It is called R^2 .**
To understand R^2 , we need to think of the total variability of the data. TSS provides a sense of the variability of the data. The formula for the variance of x is the same, but N – 1 is omitted.

$$s^2 = \frac{\sum(x-\bar{x})^2}{n-1}$$

$$\text{TSS} = \sum(x-\bar{x})^2$$

The total variability of the data can also be decomposed as:

$$\text{TSS} = SSE + SSR,$$

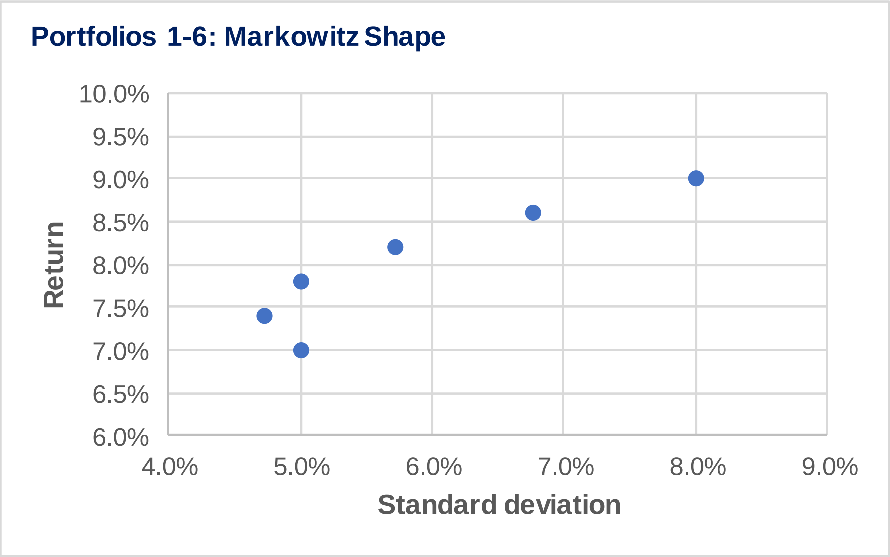
where SSE is the sum of the squares that were explained, and SSR is the sum of the squared residuals (unexplained)

$$R^2 = 1 - \frac{SSR}{SST} \qquad R^2 \epsilon [0; 1]$$

Typically, R square is looked at as a percentage value, and it can range from 0% to 100%. **The higher it is, the greater the explanatory power of the regression model** (the lower the weight of unexplained squares, the better the model).

Efficient frontier

Markowitz proved **the existence of an efficient set of portfolios that optimize investors' return for the amount of risk they are willing to accept.**
One of the most important highlights of his work is that investments in multiple securities shouldn't be analyzed separately, but should be considered in a portfolio, and financiers must understand how different securities in a portfolio interact with each other.



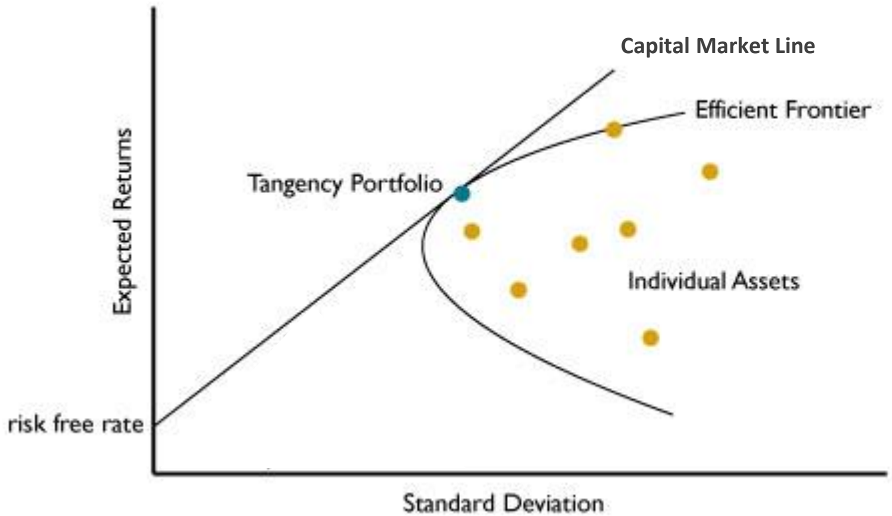
Markowitz **suggested that through the combination of securities with low correlation**, investors can optimize their returns without assuming additional risk.

The Capital Asset Pricing Model

William Sharpe's Capital Asset Pricing Model contains:

CAPM concepts	Description
Market portfolio	A bundle of all possible investments in the world (both bonds and stocks), and the risk-return combination of this portfolio is superior to the one of any other portfolio
Risk-free asset	An investment with no risk (zero standard deviation). It has a positive rate of return, but zero risk associated with it
Beta coefficient	It helps us quantify the relationship between a security and the overall market portfolio
Capital Market Line	The line that connects the risk-free rate and is tangent to the efficient frontier is called the Capital Market Line. The point where the Capital Market Line is tangent to the efficient frontier is the market portfolio

The Capital Asset Pricing Model



- The Tangency portfolio is the Market portfolio
- Individual assets (portfolios of assets) that are not lying on the efficient frontier are inefficient
- The risk-free rate has a standard deviation of 0
- The tangency portfolio lies on the Capital Market Line
- Once we introduce a risk-free asset, all investors would be interested in buying only two assets – the risk-free asset and the Market portfolio (the Tangency portfolio)

The Capital Asset Pricing Model

The CAPM formula:

$$r_i = r_f + \beta_{im}(r_{m-}r_f)$$

A security's expected return is equal to the return of the risk-free asset plus beta times the expected return of the market minus the return of the risk-free asset.

Risk-free: The minimum amount of compensation an investor would expect from an investment

Beta coefficient: How risky is the specific asset with respect to the market

Market Risk Premium: The amount of compensation an investor would expect when buying the Market portfolio

The Sharpe ratio formula:

$$\text{Sharpe Ratio} = \frac{r_{i} - r_{f}}{\sigma_i}$$

Rational investors want to maximize their rate of return and minimize the risk of their investment, so they need a measure of risk-adjusted return, a tool that would allow them to compare different securities, as they will be interested in investing in the ones that will provide the highest return for a given amount of risk.

This is how William Sharpe came up with his famous Sharpe ratio. It is a great way to make a proper comparison between stocks and portfolios and decide which one is preferable in terms of risk and return.

What is alpha?

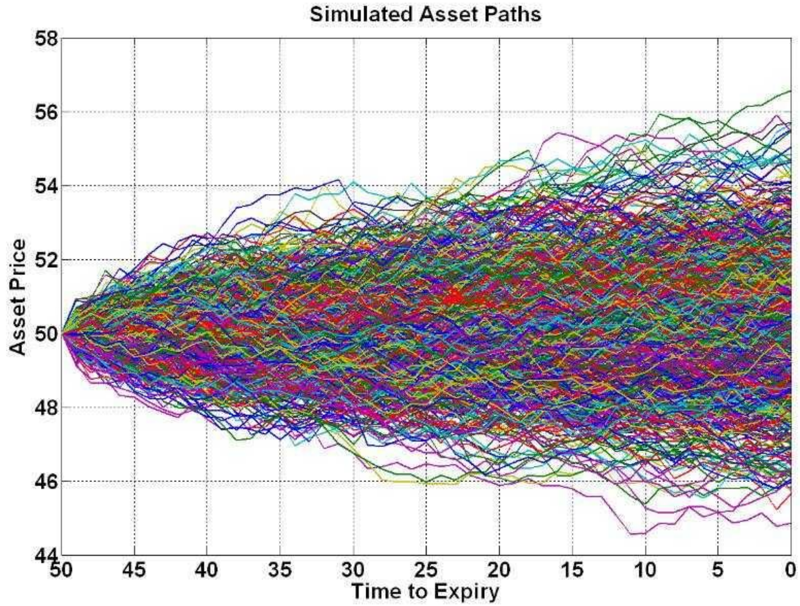
$$r_i = \alpha + r_f + \beta_{im}(r_{m} - r_f)$$

In the world of finance and investments, alpha is often thought of as a measure of how good (or poor) the performance of a fund manager has been. The standard CAPM setting assumes efficient financial market and an alpha of 0. Given that beta multiplied by the equity risk premium gives us the compensation for risk that's been taken with the investment, alpha shows us how much return we get without bearing extra risk. A great portfolio manager, someone who outperforms the market, can achieve a high alpha. And conversely, a poor investment professional may even obtain a negative alpha, meaning he underperformed the market.

Types of Investment Strategies

Type of correlation	Description
Passive Investing	Consists in buying a portfolio of assets and holding it in the long-run regardless of short-term macroeconomic developments
Active Investing	Frequent trading based on expectations regarding macroeconomic and company-specific developments,
Arbitrage Trading	Find pricing discrepancies on the market and exploit these discrepancies in order to make a profit without assuming risk
Value Investing	Invest in specific companies, hoping they will outperform their peers

Monte Carlo Simulations



When we run a Monte Carlo simulation, we are interested in observing the different possible realizations of a future event. What happens in real life is just one of the possible outcomes of any event.

And that's where a Monte Carlo simulation comes in handy. We can use past data, something we already know, to create a simulation – a new set of fictional but sensible data. These realizations are generated by observing the distribution of the historical data and calculating its mean and variance.

Such information is valuable, as it allows us to consider a good proxy of the probability of different outcomes and can help us make an informed decision.

Monte Carlo in a Corporate Finance setting

Type of correlation	Description
Revenues	Current revenues = Previous revenues * (1 + growth rate) Growth rate is the unknown variable. We can simulate its development if we know its distribution, mean, and standard deviation. This would allow us to obtain multiple simulations about the development of revenues
Cost of goods sold	Cost of goods sold = Percentage of revenues; For each “revenue path” Cogs can be simulated as a percentage of the observed amount of revenues. All we have to do is simulate the percentage as a random variable with a known distribution, mean, and standard deviation
Gross profit	Revenues – Cost of goods sold = Gross profit

$$Price\ Today = Price\ Yesterday\ * e^{r}$$

The price of a share today is equal to the price of the same share yesterday multiplied by the log return of the share (r). Remember that:

$$e^{\ln(x)} = x$$

So, above we have written:

$$Price\ Today = Price\ Yesterday\ * e^{\ln(\frac{Price\ Today}{Price\ Yesterday})}$$

Which is equal to:

$$Price\ Today = \cancel{Price\ Yesterday}\ * \frac{Price\ Today}{\cancel{Price\ Yesterday}}$$

$$Price\ Today = Price\ Today$$

$$Price\ Today = Price\ Yesterday\ * e^{r}$$

We know yesterday's stock price. We don't know r, as it is a random variable. Brownian motion is a concept that would allow us to model such randomness. The formula we can use is made of two components:

1 Drift

$$\mu - \frac{1}{2}\sigma^2$$

2 Random component

$$\sigma * Z(Rand(0;1))$$

$$Price\ Today = Price\ Yesterday\ * e^{\mu-\frac{1}{2}\sigma^2+\sigma*Z(Rand(0;1))}$$

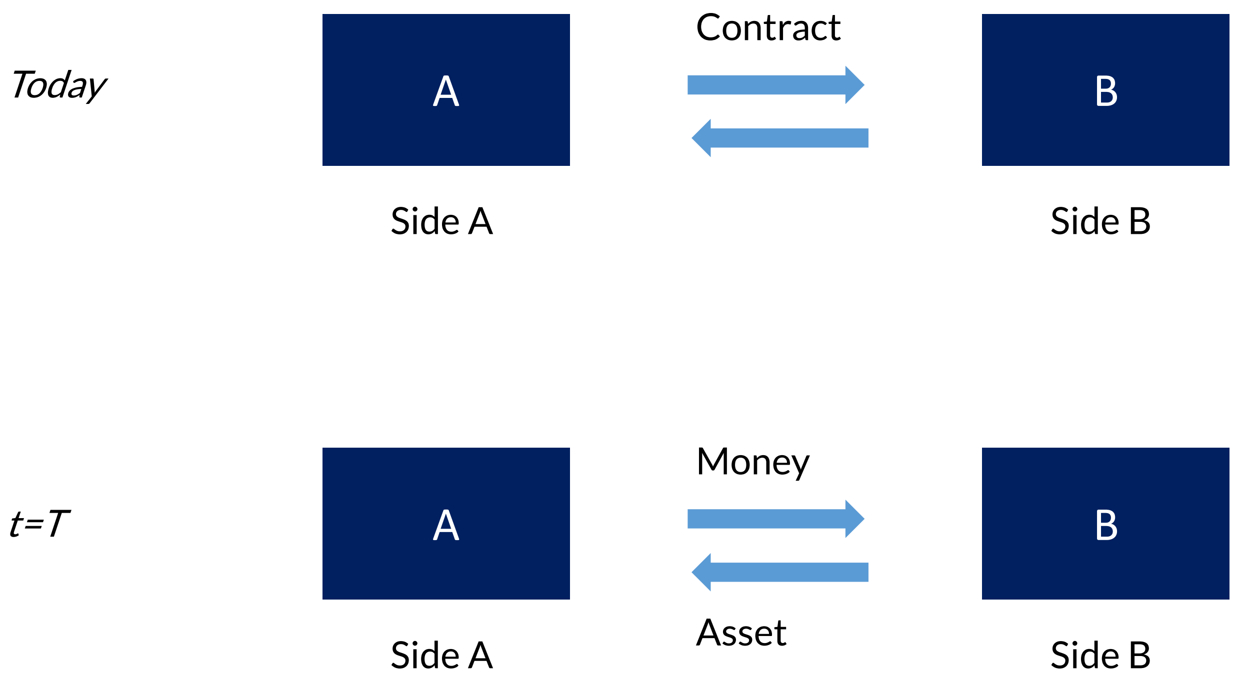
Derivative Instruments

Originally, derivatives served as a hedging instrument. Companies interested in buying these contracts were mostly concerned about protecting their investment. However, with time, financial institutions introduced a great deal of innovation to the scene, the so-called financial engineering was applied, and new types of derivatives appeared.

Type of correlation	Description
Forwards	A forward contract is used when two parties agree that one party will sell to the other an underlying asset at a future point of time. The price of the asset is agreed beforehand.
Futures	Futures are highly standardized forward contracts typically stipulated in a marketplace. The difference between futures and forwards is the level of standardization and the participation of a clearing house – the transaction goes through the marketplace, and the counterparties do not know each other.
Swaps	Swap contracts are derivative instruments in which two parties agree to exchange cash flows, based on an underlying asset at a future point of time. The underlying asset can be an interest rate, a stock price, a bond price, a commodity price, and so on.
Options	An option contract enables its owner to buy or sell an underlying asset at a price, also known as strike price. The owner of the option contract may buy or sell the asset at the given price, but he may also decide not to do it if the asset's price isn't advantageous.

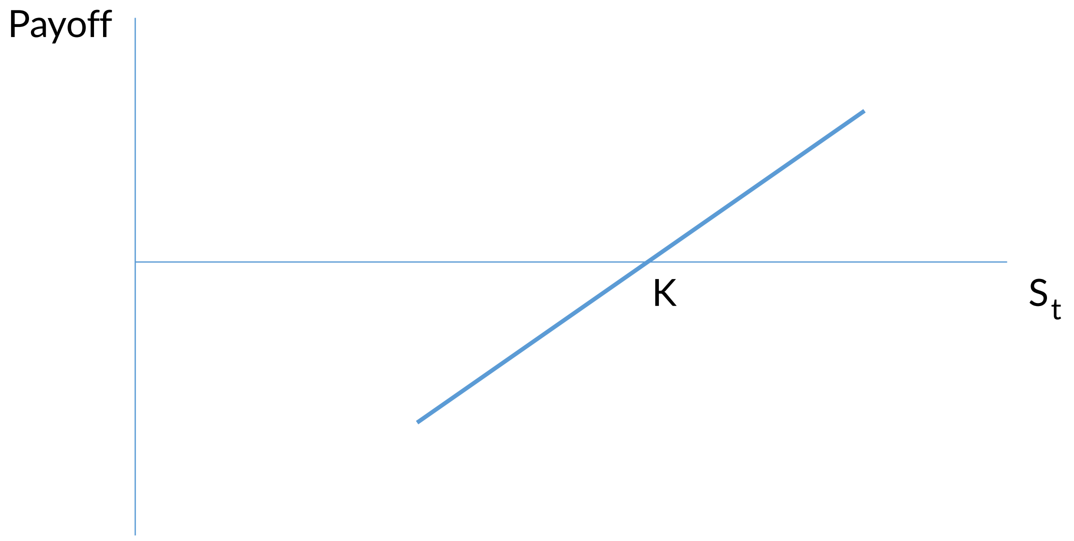
Forward / Future contracts

The two parties enter into an agreement to buy/sell an asset at time T.



Forward / Future contracts

A forward/future contract has the following payoff:



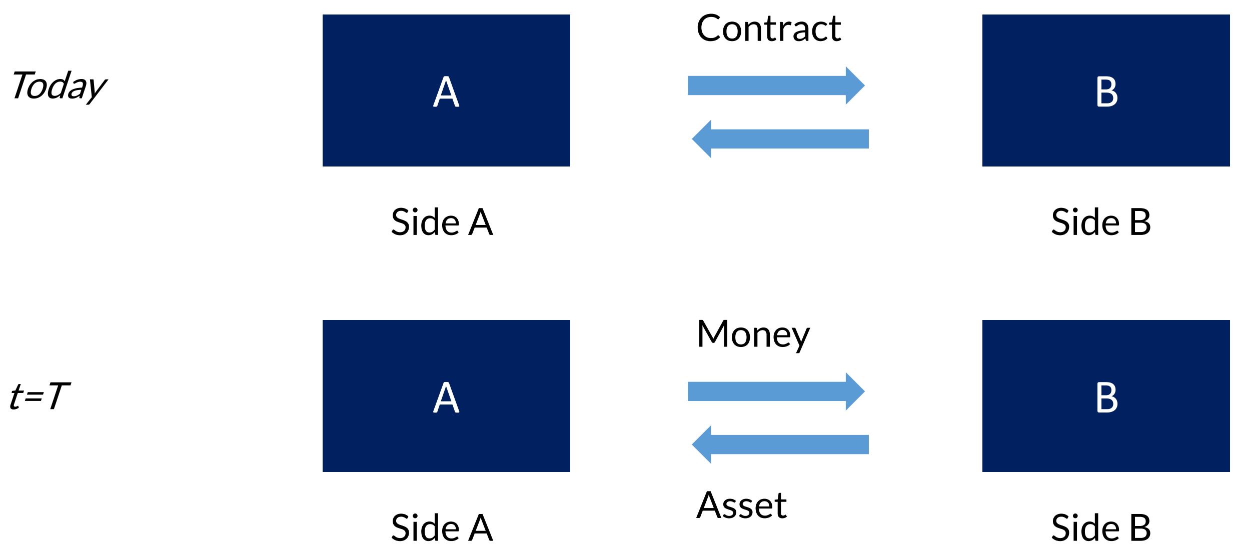
The payoff of a forward/future contract is a function of the agreed price when the contract is signed (K), and the price of the asset at time t

Option contracts

There are two main types of options.

Call Options – the holder has the right to buy an asset at an agreed strike price.

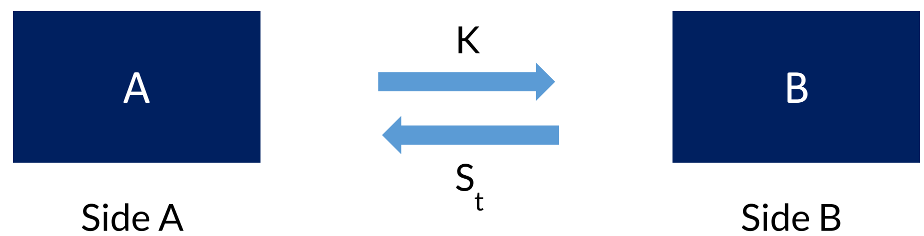
Put Options – the holder has the right to sell an asset at an agreed strike price.



At time T, the owner of the option decides whether to exercise it or not:

Scenario 1: Exercise

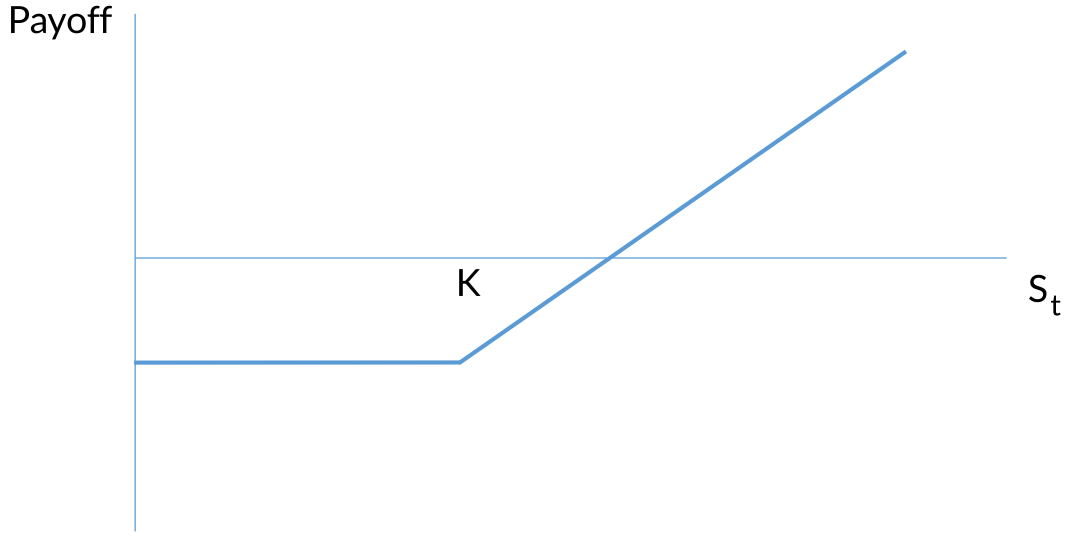
t=T



Scenario 2: Don't exercise

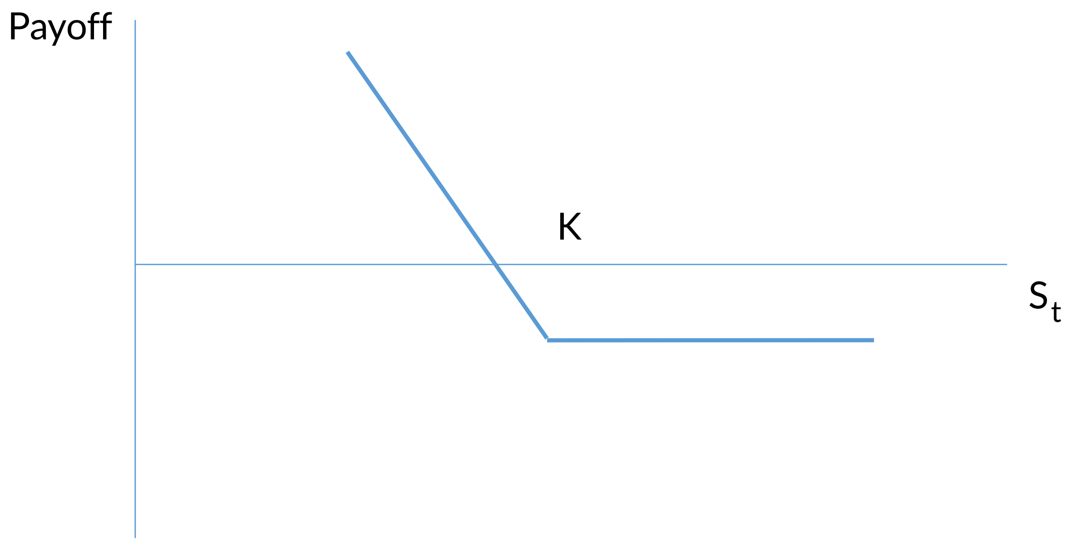
t=T

A call option contract has the following payoff:



The payoff of an option is a function of the agreed strike price when the contract is signed (K), and the price of the asset at the time of maturity of the option (St). In addition, there are two types of options – European and American. European options can be exercised only at maturity, while American options can be exercised at any time and are hence more valuable.

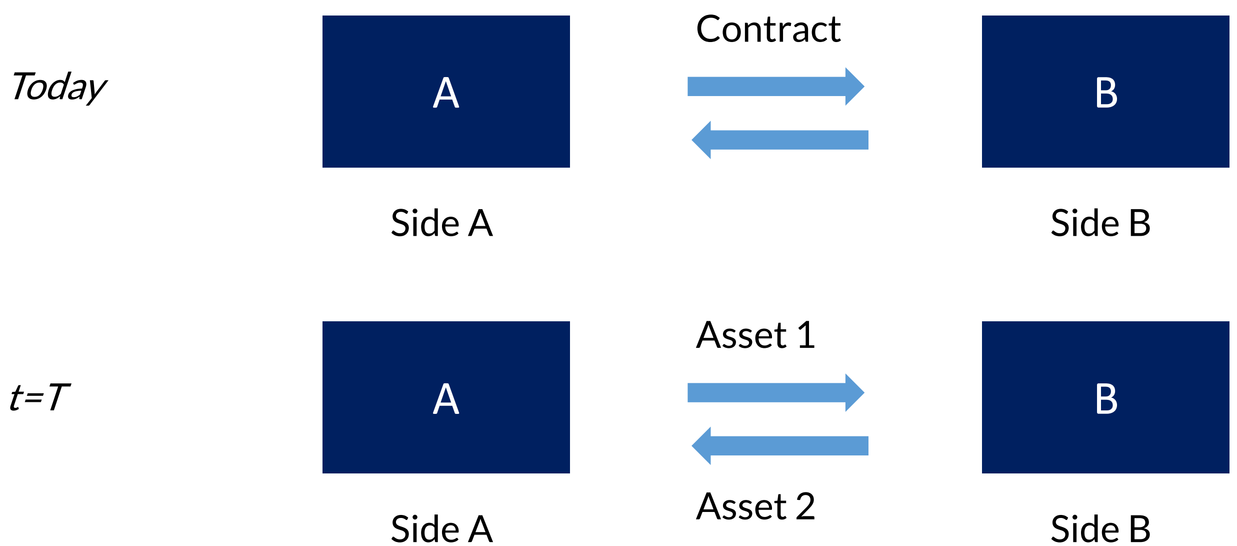
A put option contract has the following payoff:



The payoff of an option is a function of the agreed strike price when the contract is signed (K), and the price of the asset at the time of maturity of the option (St). In addition, there are two types of options – European and American. European options can be exercised only at maturity, while American options can be exercised at any time and are hence more valuable.

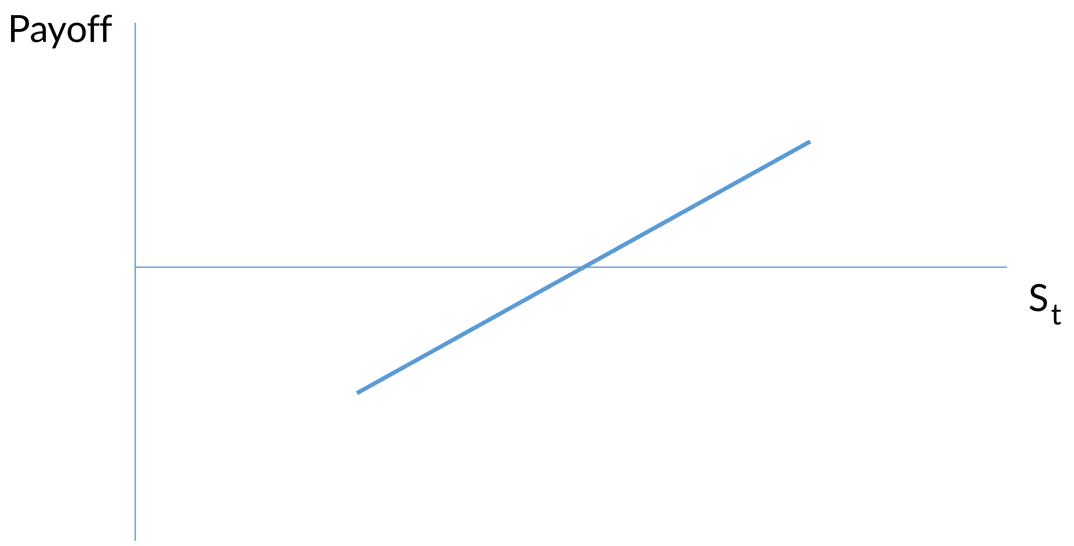
Swap contracts

In a swap contract, the two parties agree to exchange cash flows based on an underlying asset.



<u>Cash flow = Price of Asset 1 – Price of Asset 2</u>

A swap contract has the following payoff:



The payoff of a swap is a function of the price of the underlying asset.

The Black Scholes formula is one of the most widely used tools for derivatives pricing. It can be written in the following way:

$$C(S,t) = N(d_1)S - \mathrm{N}(d_2)Ke^{-r(T-t)}$$
$$d_1 = \frac{1}{s\sqrt{(T-t)}}[\ln\left(\frac{S}{K}\right) + \left(r + \frac{s^2}{2}\right)(\mathrm{T} - \mathrm{t})]$$
$$d_2 = d_1 - s\sqrt{T-t}$$

Type of correlation	Description
S	The stock's current market price
K	The strike price at which the option can be exercised; if we exercise the option, we can buy the stock at the strike price K
T	The option's time until expiration
r	Risk-free rate

The Black Scholes formula is one of the most widely used tools for derivatives pricing. It can be written in the following way:

$$C(S,t) = N(d_1)\mathrm{S} - \mathrm{N}(d_2)Ke^{-r(T-t)}$$

$$d_1 = \frac{1}{s\sqrt{(T-t)}}[\ln\left(\frac{S}{K}\right) + \left(r + \frac{s^2}{2}\right)(\mathrm{T}-\mathrm{t})]$$

$$d_2 = d_1 - s\sqrt{T-t}$$

Type of correlation	Description
s	The standard deviation of the underlying asset
N	Normal distribution