Capital asset pricing model (CAPM)

CAPM shows a linear relationship between risk and expected return and that is used in the pricing of risky securities. The security market line of the CAPM is defined as

$$\boxed{\text{SML: } R_{it} - R_{ft} = a_i + \beta_i (R_{Mt} - R_{ft}) + \varepsilon_{it}}$$

The alpha coefficient (alpha) is the constant (intercept) in the security market line of the CAPM. The alpha coefficient indicates how an investment has performed after accounting for the risk it involved. If markets are efficient then alpha = 0. If **alpha < 0** the security has **earned too little for its risk**. If **alpha > 0** the investment has a **risk in excess of the return** for the given risk.

Fama-French model (FFM) (Link: [refer])

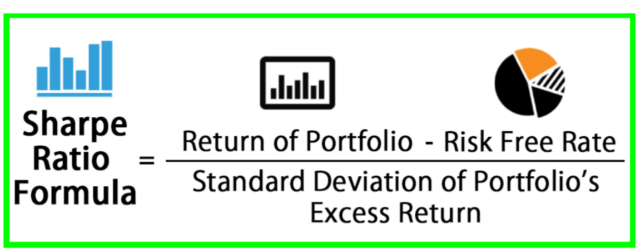
Arbitrage Pricing Theory (APT) (Link: [refer])
APT states that the return of a financial asset is a linear function of a number of factors such as macroeconomic factors (e.g., the market rate of interest, economic growth, inflation) and market indices.
The APT produces a rate of return that can be checked against actual return to see if the asset is correctly priced.

Yield curve (term structure of interest rates)
Yield curve helps to understand the structure of interest rates over the time period.

The expected one-year returns are all at time t. Assume the investor bases his/her expectation on the current one-year bond, $r_{1,t}$.The estimating equation for the n –year bond (long-term) is $r_{n,t} = a + b \ \ r_{1,t}$ [estimating equation for yield curve].

Note that the theory (if it panned out) suggests that $b = 1$. This can be tested using the t-test from OLS.

Sharpe ratio
The Sharpe ratio measures investment risk. When comparing two assets, the one with a higher Sharpe ratio provides a superior return for the same risk.
In general, Sharpe Ratio = Avg daily returns / Std of daily returns



	X	Y
R_p	12%	20%
R_f	5%	5%
σ_p	0.04	0.15

Sharpe Ratio for X is calculated using below formula

Sharpe Ratio = $(R_p - R_f) / \sigma_p$

Sharpe Ratio Formula	=(B7-B8)/B9
Sharpe Ratio for X	**1.75**

Sharpe Ratio above 3 is good.

VAR (Value-at-Risk)

Value at Risk (VaR) is a key concept in portfolio risk management. It uses the past observed distribution of portfolio returns to estimate what your future losses might be at different likelihood levels.

Measure of risk of loss for a particular asset or portfolio. Meaning estimates the amount exposed to risk by hypothesis testing. Using the confidence interval concept.

Example:
Amount of dollars invested: $10 million Standard deviation: 1.99% or 0.0199
The VaR at the 95% confidence level is 1.645 x 0.0199 or 0.032736.
The portfolio has a market value of £10 million, so the VaR of the portfolio is
0.032736 x 10,000,000 = $327,360
Z = 2.33 for a 99 % confidence level considering the normal distribution.
Z = 1.64 for a 95 % confidence level considering the normal distribution.

ADF - Augmented Dickey-Fuller test - (Link: [refer]())

The method helps to determine that there is a time dependent factor involved in the series using hypothesis testing. Determine if the hypothesis is correct or not based on p value.

```
import numpy as np
import pandas as pd
import statsmodels.tsa.stattools
#the following is a list of returns
a=[1,2,-1,2,-3,0,.5,.8,-.21,2,-1,2,-3,0,.5,.8,-2,2,-1,2,-3,0,.5,.8,-.21,2,-1,2,-3,0,.5,.8,-.2]
statsmodels.tsa.stattools.adfuller(a, maxlag=None)
Out[10]:
(-1.7034844119211205,0.42931810315407193,7,25,{'1%': -3.7238633119999998, '10%': -2.6328003999999998, '5%': -2.98648896},
 49.908550173095492)
```

```
# Import the Time Series library
import statsmodels.tsa.stattools as ts

# Import fix_yahoo-finance module
import fix_yahoo_finance as yf

# Download Google data from 1/1/2000 to 1/1/2019
goog=yf.download("GOOG", start="2000-01-01", end="2019-01-01")
# Output the results of the Augmented Dickey-Fuller test for Google
# with a lag order value of 1
ts.adfuller(goog['Adj Close'], 1)
```

Stationarity

A time series is considered stationary if the values never drift far away from their initial value. **Most of the financial time series are not stationary;** however, after applying the first difference (x(t) – x(t-1)), they become stationary.

Classic econometric models require time series to be stationary

A variable that has a finite mean and variance is considered covariance stationary.

Stationary time series variables tend to return to their mean after a period of time (mean reversion), and their deviations from the mean tend to be constant.

Types of non-stationary processes

1. pure random walk
2. random walk with drift
3. deterministic trend
4. random walk with drift and deterministic trend

Auto-correlation

In finance, serial correlation is used by technical analysts to determine how well the past price of a security predicts the future price

Durbin-Watson statistics: **Helps to determine if correlation exists or not.**

d= 2(1 − r), where r is the sample autocorrelation of the residuals.

d = 2, it indicates no auto-correlation

d > 2 , negative auto-correlation

d < 2, Positive auto-correlation

Breusch–Pagan test (Link: refer)
Test to check heteroskedasticity in residuals.
P value helps to determine it.

Use of log in finance:
It converts non-linear sophisticated relations into linear and easier to model connections between the variables. Trend measured in ln units ≈ percentage growth.

Autoregressive and Moving Average
The **Autoregressive (AR)** model is a popular tool for **predicting short-term evolutions** of assets, returns or spread.

Current value is influenced on the immediately preceding value at AR(1) and by the previous two values at AR(2)

AR(1) model for a series with T observations

$$y(T) = \alpha + \beta * y(T-1).$$

An **MA** process shows that the current values of the time series **depend on the previous** (unobserved) random shocks.

ARIMA : Autoregressive integrated moving average model.
ARIMA (p,d,q)
P - order of AR, d- order of stationarity, q - order of MA

In ARMA the order of integration of time series is missing.

AR process, an MA process or an ARMA process, which is best fit?
ACF and PACF

Type of model	Typical pattern of ACF	Typical pattern of PACF
AR(p)	Decays exponentially or with a damped sine wave pattern or both	Significant spikes through lags p
MA(q)	Significant spikes through lags q	Declines exponentially
ARMA(p, q)	Exponential decay	Exponential decay

ARMA modeling requires practice as choosing the right p and q is not an easy task
Quant analysts combine the information provided by ACF and PACF with the one provided by information criteria, such as the **Akaike Information Criterion**, Bayesian Information Criterion and Hannan-Quinn Information Criterion
Goodness-of-fit measures

R^2 and the average of the residual sum of squares.

EPAT
Arch and all
Kalman filter
Kalman filter in algorithmic trading - WQU
GARCH and multivariate GARCH
COUPLA Github

Vector Auto-Regressive Modeling (VAR)
Univariate ARMA modeling approach to **multivariate modeling** in the stationary time series context. This approach is called **VAR**.

Trading Ideas

Pair Trading

Trading correlated stocks (pairs trade or pair trading) means **trading a pair of stocks that are highly correlated**. This means the two stocks move in the same direction in most of the situations. The trading **opportunity appears when the two stocks move temporarily in opposite directions**

Example:
Coca-Cola (KO) and Pepsi (PEP)

The spread between two correlated stocks is easier to predict. The spread can be a stationary time series, and therefore a trader can use an Autoregressive Moving Average (ARMA) model to predict the spread.

The Dividend Discount Model DDM, also known as Gordon Growth Model
It is a quantitative **approach for valuing a stock price**. According to DDM, the price of a stock is the sum of future dividends discounted back to their present value.
Drawback: Need dividends info past and future