<https://www.etftrends.com/crypto-channel/the-race-to-the-first-bitcoin-etf/>

Arma model

Duration lag

Arima

Arma

sarima

Kaggle references

<https://www.kaggle.com/lp187q/etf-directional-prediction-sensitivity-analysis>

<https://www.kaggle.com/general/272226>

<https://www.kaggle.com/andreshg/timeseries-analysis-a-complete-guide>

<https://www.kaggle.com/general/272429>

<https://www.kaggle.com/freespirit08/time-series-for-beginners-with-arima>

<https://www.kaggle.com/thebrownviking20/everything-you-can-do-with-a-time-series>

<https://www.kaggle.com/anshuls235/time-series-forecasting-eda-fe-modelling>

<https://www.kaggle.com/biphili/time-series-data-analysis-stock-price-code-12>

<https://www.kaggle.com/arashnic/time-series-forecasting-with-yahoo-stock-price>

Reading notes of <https://www.kaggle.com/freespirit08/time-series-for-beginners-with-arima>

**4. What are the components of Time Series?**

A. There are 4 components:

a) Trend - Upward & downward movement of the data with time over a large period of time. Eq: Appreciation of Dollar vs rupee.

b) Seasonality - seasonal variances. Eq: Ice cream sales increases in Summer only

c) Noise or Irregularity - Spikes & troughs at random intervals

d) Cyclicity - behavior that repeats itself after large interval of time, like months, years etc.

**5. What is Stationarity?**

A. Before applying any statistical model on a Time Series, the series has to be staionary, which means that, over different time periods,

a) It should have constant mean.

b) It should have constant variance or standard deviation.

c) Auto-covariance should not depend on time.

Trend & Seasonality are two reasons why a Time Series is not stationaru & hence need to be corrected.

解决方法：remove trend, detrend (很多方法) ， how to make stationarity of time series

平稳时间序列。差分。

**7. Tests to check if a series is stationary or not**

A. There are 2 ways to check for Stationarity of a TS:

a) Rolling Statistics - Plot the moving avg or moving standard deviation to see if it varies with time. Its a visual technique.

b) ADCF Test - Augmented Dickey–Fuller test is used to gives us various values that can help in identifying stationarity. The Null hypothesis says that a TS is non-stationary. It comprises of a **Test Statistics** & some **critical values** for some confidence levels. If the Test statistics is less than the critical values, we can reject the null hypothesis & say that the series is stationary. THE ADCF test also gives us a **p-value**. Acc to the null hypothesis, lower values of p is better.

**8. What is ARIMA model?**

A. ARIMA(Auto Regressive Integrated Moving Average) is a combination of 2 models AR(Auto Regressive) & MA(Moving Average). It has 3 hyperparameters - P(auto regressive lags),d(order of differentiation),Q(moving avg.) which respectively comes from the AR, I & MA components. The AR part is correlation between prev & current time periods. To smooth out the noise, the MA part is used. The I part binds together the AR & MA parts.

**9. How to find value of P & Q for ARIMA ?**

A. We need to take help of ACF(Auto Correlation Function) & PACF(Partial Auto Correlation Function) plots. ACF & PACF graphs are used to find value of P & Q for ARIMA. We need to check, for which value in x-axis, graph line drops to 0 in y-axis for 1st time.

From PACF(at y=0), get P

From ACF(at y=0), get Q

**10. What Is ADCF test?**

A. In statistics and econometrics计量经济学, an augmented Dickey–Fuller test (ADF) tests the null hypothesis that a unit root is present in a time series sample. The alternative hypothesis is different depending on which version of the test is used, but is usually stationarity or trend-stationarity. It is an augmented version of the Dickey–Fuller test for a larger and more complicated set of time series models.

The augmented Dickey–Fuller (ADF) statistic, used in the test, is a negative number. The more negative it is, the stronger the rejection of the hypothesis that there is a unit root at some level of confidence.

p value(0<=p<=1) should be as low as possible. Critical values at different confidence intervals should be close to the Test statistics value.