1. Feature Engineering

Time series Component:

* Trend 🡪 increasing or decreasing in overall value
* Seasonality 🡪 repeating pattern in a certain cycle
* Level 🡪 average value in the series
* Noise 🡪 random variation in Series

Data Stationarity:

* Constant mean
* Constant variance
* Autocovariance does not depend on time

How to test Stationary?

* Augmented Dicky Fuller Test (ADF)

How does the test work?

* the null hypothesis is the time series possesses a unit root and is non-stationary. So, if the P-Value in ADF test is less than the significance level (0.05), you reject the null hypothesis. (We have to reject the null hypothesis to make it stationary)

How to make our data stationary: (there are 5 ways)

1. Transformation 🡪 Log, Square, Root
2. Smoothing 🡪 weekly avg, monthly avg, rolling avg
3. Differencing 🡪 First Order Differencing
4. Polynomial Fitting 🡪 fitting regression model (not use in our project)
5. Decomposition 🡪 Detrending and Deseasonalizing, okay this have longer explanation

So, every time series data contains trend, seasonality, level, noise. We use seasonal\_decompose() function it’ll decompose our data and return us the seasonal, trend, and residual (level + noise). Since residual is our data that doesn’t have the trend and seasonal anymore, so residual most likely a stationary data and we can use that one.

Extra: decompose model have additive and multiplicative model, if using additive method etherPrice = resid+trend+seasonality, is using multiplicative then etherPrice = resid\*trend\*seasonality. The one that we apply on our project is the additive one.