**Time Series Algorithm Implementation in Project**

## **How to Analyze Time Series?**

To perform the [time series analysis](https://www.analyticsvidhya.com/blog/2022/05/a-comprehensive-guide-to-time-series-analysis-and-forecasting/), we have to follow the following steps:

* Collecting the data and cleaning it
* Preparing Visualization with respect to time vs key feature
* Observing the stationarity of the series

## **Components of Time Series Analysis**

components of [Time Series Analysi](https://www.analyticsvidhya.com/blog/2018/09/multivariate-time-series-guide-forecasting-modeling-python-codes/)s:

* **Trend**: In which there is no fixed interval and any divergence within the given dataset is a continuous timeline. The trend would be Negative or Positive or Null Trend
* **Seasonality**: In which regular or fixed interval shifts within the dataset in a continuous timeline. Would be bell curve or saw tooth
* **Cyclical**: In which there is no fixed interval, uncertainty in movement and its pattern
* **Irregularity**: Unexpected situations/events/scenarios and spikes in a short time span.

Let’s see the types of moving averages:

* Simple Moving Average (SMA),
* Cumulative Moving Average (CMA)
* Exponential Moving Average (EMA)

Simple Moving Average (SMA)

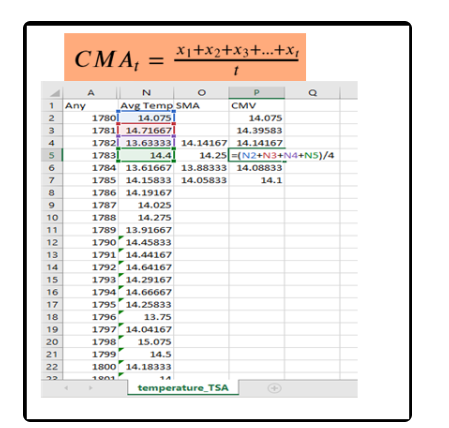
The Simple Moving Average (SMA) calculates the unweighted mean of the previous M or N points. We prefer selecting sliding window data points based on the amount of smoothing, as increasing the value of M or N improves smoothing but reduces accuracy.

#### Cumulative Moving Average (CMA)

The CMA is the unweighted mean of past values till the current time.

#### Cumulative Moving Average (CMA)

The CMA is the unweighted mean of past values till the current time.



**STEP1:IMPORTING LIBRARIES AND READING THE DATA**

**import pandas as pd**

**from matplotlib import pyplot as plt**

**from statsmodels.graphics.tsaplots import plot\_acf**

**df\_temperature = pd.read\_csv(r"C:\Users\GT-499\Desktop\Excelr\Dr.Gns resentation\PRESENTATION 3-5\42\_Time series\GlobalTemperatures.csv")**

**df\_temperature.head()**

**STEP2: EDA**

**df\_temperature.info()**

**Step3: SMA over a period of 10 and 20 years**

**df\_temperature['SMA\_10'] = df\_temperature.LandAverageTemperature.rolling(10, min\_periods=1).mean()**

**df\_temperature['SMA\_20'] = df\_temperature.LandAverageTemperature.rolling(20, min\_periods=1).mean()**

**colors = ['green', 'red', 'orange']**

**# Line plot**

**df\_temperature.plot(color=colors, linewidth=3, figsize=(12,6))**

**plt.xticks(fontsize=14)**

**plt.yticks(fontsize=14)**

**plt.legend(labels =['LandAverageTemperature', '10-years SMA', '20-years SMA'], fontsize=14)**

**plt.title('The yearly average air temperature in city', fontsize=20)**

**plt.xlabel('Year', fontsize=16)**

**plt.ylabel('Temperature [°C]', fontsize=16)**

#### STEP : Cumulative Moving Average (CMA)

**# CMA Air temperature**

**df\_temperature['CMA'] = df\_temperature.LandAverageTemperature.expanding().mean()**

**# green -Avg Air Temp and Orange -CMA**

**colors = ['green', 'orange']**

**# line plot**

**df\_temperature[['LandAverageTemperature', 'CMA']].plot(color=colors, linewidth=3, figsize=(12,6))**

**plt.xticks(fontsize=14)**

**plt.yticks(fontsize=14)**

**plt.legend(labels =['Average Air Temperature', 'CMA'], fontsize=14)**

**plt.title('The yearly average air temperature in city', fontsize=20)**

**plt.xlabel('Year', fontsize=16)**

**plt.ylabel('Temperature [°C]', fontsize=16)**

**Note:**

**from statsmodels.graphics.tsaplots import plot\_acf**

**ACF is an autocorrelation function**

Plot the autocorrelation function

Plots lags on the horizontal and the correlations on vertical axis.

What is rolling() in Python?

The rolling() method in Pandas is used to perform rolling window calculations on sequential data. A rolling window is a fixed-size interval or subset of data that moves sequentially through a larger dataset.

 Minimum number of observations in window required to have a value; otherwise, result is np.nan . For a window that is specified by an offset, min\_periods will default to 1. For a window that is specified by an integer, min\_periods will default to the size of the window.

What does PLT legend() do?

The plt.legend().Legend() A legend is used to describe elements for a particular area of a graph. Python has a function called legend() which is used to place a legend on the axis.

The expanding() method is one of the window methods of pandas and it Provides expanding transformations. And it returns a window subclassed for the particular operation

**>>>>>>>>>>>>end of code…………………………………….**

**Dataset**

Global surface temperature data sets are an essential resource for monitoring and understanding climate variability and climate change. The most commonly used data sets combine historical observations of near surface air temperatures at land stations with global data sets of sea surface temperatures (SST) obtained from a changing mix of ship-based and measurements. While the concept of these data sets is fairly simple, their construction is challenging due to difficulties in obtaining data, documenting and accounting for changes in instrumentation and observing practices, addressing changes in station location and local land use, understanding random measurement errors, and deciding where and how to infill missing data in space and time.