# Summary/Review

## Time Series

Time Series is a sequence of data points organized in time order.
The sequence captures data at equally spaced points in time. Data that is not collected regularly at equally spaced points is not considered time series.

## Time Series Motivation

For most forecasting exercises, standard regression approaches do not work for Time Series models, mostly because:
- The data is correlated over time
- The data is often non-stationary, which is hard to model using regressions
- You need a lot of data for a forecast

## Forecasting Problems

These are the two types of forecasting problems. Consider that the vast majority of applications employ univariate models, harder to combine variables when using time series data.

1. **Univariate**

Think of single data series containing of:
- Continuous data, binary data, or categorical data
- Multiple unrelated series
- Conditional series

2. **Panel or Multivariate**

Think of multiple related series identifying groups such as customer types, department or channel, or geographic joint estimation across series

## Time Series Applications

Time series data is common across many industries. For example:
- Finance: stock prices, asset prices, macroeconomic factors
- E-Commerce: page views, new users, searches
- Business: transactions, revenue, inventory levels

Time series methods are used to:
- Understand the processes driving observed data
- Fit models to monitor or forecast a process
- Understand what influences future results of various series
- Anticipate events that require management intervention

## Time Series Components

A time series can be decomposed into several components:
**Trend** – long term direction
**Seasonality** – periodic behavior
**Residual** – irregular fluctuations

Generally, models perform better if we can first remove known sources of variation such as trend and seasonality. The main motivation for doing decomposition is to improve model performance. Usually we try to identify the known sources and remove them, leaving resulting series (residuals) that we can fit against time series models

# Decomposition Models

These are the main models to decompose Time Series components:
## – Additive Decomposition Model

Additive models assume the observed time series is the sum of its components.
**i.e. Observation = Trend + Seasonality + Residual**
These models are used when the magnitudes of the seasonal and residual values are independent of trend.

## – Multiplicative Decomposition Model

**Multiplicative** models assume the observed time series is the product of its components.
**i.e. Observation = Trend * Seasonality * Residual**
A multiplicative model can be transformed to an additive by applying a log transformation:
**log(Time*Seasonality*Residual) = log(Time) + log(Seasonality) + log(Residual)**
These models are used if the magnitudes of the seasonal and residual values fluctuate with trend.

## – Pseudo-additive Decomposition Model

Pseudo-additive models combine elements of the additive and multiplicative models.
They can be useful when:
Time series values are close to or equal to zero.
We expect features related to a multiplicative model.
A division by zero needs to be solved in the form: **$O_t = T_t + T_t(S_t – 1) + T_t(R_t – 1) = T_t(S_t + R_t – 1)$**
Decomposition of time series allows us to remove deterministic components, which would otherwise complicate modeling.
After removing these components, the main focus is to model the residual.

# Other Methods

These are some other approaches of time series decomposition:
- Exponential smoothing
- Locally Estimated Scatterplot Smoothing (LOESS)
- Frequency-based methods