Models Used

ARIMA (Autoregressive Integrated Moving Average)

- Uses 3 parameters p, d, and q to parametrize the 3 major aspects of a time series - seasonality, trend, and noise.
- P is the number of autoregressive terms.
- D is the number of nonseasonal differences needed for stationarity.
- Q is the number of lagged forecast errors.

SARIMA (Seasonal ARIMA)

- Uses an additional parameter m in the parametrization to take into consideration the number of time steps for a single period

SARIMAX

- SARIMA with an additional argument X, which allows exogenous variables to be added.

Understanding the Models:

- Time series are simply stochastic processes where the index set T in $\{X_t : t \in T\}$ is time.
- An **autoregressive process** of order 1 is an i.i.d. $N\left(0, \tau^2\right)$ process with $\{Z_n : n \in \mathbb{Z}\}$ and $X_n = \alpha X_{n-1} + Z_n$ where X_{n-1} is independent of Z_n

Stationarity:

- For $T \subset \mathbb{R}^k$, a stochastic process is said to be **weakly stationary** if

  1. its mean function $\mu(t)$ is constant in t and
  2. its autocovariance function $\sigma(s,t) = \kappa(s-t)$ for some $\kappa : \mathbb{R}^k \to \mathbb{R}^1$.

  Note: κ is a positive semidefinite function, i.e. it must satisfy $\kappa(0) \geq 0, \kappa(t) = \kappa(-t)$, and $\sum_{i=1}^{n}\sum_{i=1}^{n} x_i x_j \kappa\left(t_i - t_j\right) \geq 0$ for all $\{t_1, \ldots, t_n\} \subset T, x = \left(x_1, \ldots, x_n\right)^T \in \mathbb{R}^n$

- A **strictly stationary** process has the property that $\left(X_{t_1+h}, \ldots, X_{t_n+h}\right) \sim \left(X_{t_1}, \ldots, X_{t_n}\right)$ where h is such that $\{t_1 + h, \ldots, t_n + h\} \subset T$

Note: A weakly stationary Gaussian process is always strictly stationary, since $\sigma\left(t_i+h, t_j+h\right)=\kappa\left(t_i-t_j\right)=\sigma\left(t_i, t_j\right)$ (similar to how covariance = 0 implies statistical independence with joint normality)

Autocorrelation:

- Suppose $\{(t, X_t): t \in T\}$ is a stochastic process such that $E\left(X_t^2\right)<\infty$ for all $t \in T$. Then provided $\sigma(t, t)>0 \; \forall t \in T$, the **autocorrelation function** $\rho: T \times T \rightarrow \mathbb{R}^1$ is defined as $\rho(s, t)=\dfrac{\sigma(s, t)}{\sqrt{\sigma(s, s)} \sqrt{\sigma(t, t)}}$
- In words, autocorrelation is the correlation of a signal with a delayed copy of itself as a function of delay (Wikipedia definition). Informally, it is the similarity between observations as a function of the time lag between them.