

**Q1: Based on the summary function (applied to your model object), please explain what are the AR lag coefficients included in the model?**

* From the nonseasonal ARIMA, we only have a MA(2) applied to the second difference of the Netflix data. From the nonseasonal ARIMA, we are not using any AR lags.
* From the seasonal ARIMA, we have P = 2. **What does that mean in terms of the autoregressive lags used in our model?**
  + Y\_t ~ Yt-1+ Yt\_2 if p = 2 (since I am using lowercase p, I am referring to the nonseasonal model). **Not the case based on the above output.**
  + Yt ~ Yt-4­ + Yt-8 (Generally speaking, the seasonal ARs included will be multiple of the frequency, m, which is in the square brackets in the output).
  + **After class addition: So let us say that p = 2, and P = 2. In this case, we will be regressing Yt** on Yt-1 + Yt-2 + Yt-4 + Yt-8 .

**Q2: What does the seasonal ARIMA, D, parameter value indicate?**

* From Last Class, we had D ≠ 0 (I think it was 1). Today, we have D = 0. **What does that mean?** The seasonal data/factors are stationary. Once you take the second difference of the data, no additional differences are needed for the seasonal part.

**Q3: What did we learn from the non-seasonal ARIMA parameters?**

* p = 0, which means that we do not have an autoregressive part.
* d = 2, two differences are needed to bring the data to stationarity.
* q = 2, this means that the error correlations are lagged by 1, and 2 quarters (MA part).

**Q4: In terms of the predictive performance, do we have a reasonable model?**

* The ME is relatively small compared to our data, which indicates that we do not have a biased model (i.e., we are not consistently over predicting or under predicting).
* From the MAE, we are on average 4.6 away from the true value (which given the values we had when we printed the Netflix data 🡪 this is pretty good).
* The MAPE : 3.7% (our average error is about 3.7%)
* MASE: We are having ~ 90% improvement over the naïve forecast.

**Overall, we have a reasonable model. You may need to improve upon it if the 3.7% error is high for your application in practice.**

**Q5: In terms of the assumptions behind the ARIMA model, do we have a reasonable model?**

We do have a reasonable model based on the Ljung-Box test and the plot from the checkresiduals().