**Time Series – Report**

Part I:

Question 1:

This series represents the seasonally and working-day adjusted industrial production index for the pharmaceutical industry (NAF rev. 2, division 21) in Metropolitan France, with base 100 in 2021. The index measures monthly production, adjusted to remove seasonal and calendar effects. It’s well-suited for time series modeling, but we must check for stationarity.

Part II:

* **Interpretation of ARIMA Model Results:**

Based on the output from summary(auto\_model) and checkresiduals(auto\_model), here's the breakdown of the results and the steps you can take to interpret and validate the chosen ARIMA model.

* **ARIMA Model Summary:**
* **Model Chosen**: ARIMA(2,0,1)(0,0,2)[12] with non-zero mean

This indicates the following components of the model:

* + **ARIMA(2,0,1)**:
    - **AR(2)**: The model includes two autoregressive (AR) terms, meaning that the current value depends on the previous two observations.
    - **I(0)**: The series is already stationary, so no differencing is needed (d = 0).
    - **MA(1)**: The model includes one moving average (MA) term, which means that the current value is influenced by the previous error term.
  + **Seasonal Component (0,0,2)[12]**:
    - **Seasonal AR(0)** and **Seasonal MA(2)**: The model also accounts for seasonal effects, where it includes two seasonal moving average terms over a 12-period seasonal cycle.
    - **No seasonal differencing** (d = 0 for the seasonal part).
* **Coefficients**: The coefficients for the AR and MA terms are as follows:
  + **AR1**: -0.0636
  + **AR2**: -0.0211
  + **MA1**: -0.6452
  + **SMA1**: -0.1829 (Seasonal MA term)
  + **SMA2**: -0.1119 (Seasonal MA term)
  + **Mean**: 0.0034 (Indicates a slight non-zero mean in the data)

These coefficients suggest the impact of past observations (AR terms) and past errors (MA and seasonal MA terms) on the current value.

* **Model Diagnostics:**

1. **Sigma-squared**: 0.001792  
   This represents the variance of the residuals. A smaller value typically suggests a better fit.
2. **Log-likelihood**: 736.17  
   This value is used to compare models. A higher log-likelihood indicates a better model fit.
3. **AIC/BIC**:
   * **AIC**: -1458.35
   * **BIC**: -1430.05
   * Lower values of AIC and BIC suggest a better-fitting model. The AIC and BIC values here are quite low, indicating that this model fits the data well relative to other potential models.
4. **Training Set Error Measures**:
   * **ME** (Mean Error): 0.0002686 (close to zero, good sign)
   * **RMSE** (Root Mean Squared Error): 0.0420 (a low value indicates good model fit)
   * **MAE** (Mean Absolute Error): 0.0305
   * **MAPE** (Mean Absolute Percentage Error): 143.36% (This value is high, possibly due to high variance in the series, but could also be seasonal effects.)
   * **ACF1**: 0.00098 (close to zero, indicating that there are no significant autocorrelations in the residuals at lag 1)

* **Ljung-Box Test for Residuals:**
* **Ljung-Box test**: Q\* = 34.747, df = 19, p-value = 0.01498
  + The **Ljung-Box test** checks for autocorrelation in the residuals. The null hypothesis of the test is that the residuals are white noise (no autocorrelation).
  + The **p-value** = 0.01498, which is **less than 0.05**, suggesting that **the residuals may still exhibit autocorrelation**, meaning that the model might not fully capture all dependencies in the data.
* **What to Do Next:**

**1. Model Selection:**

* **ARIMA(2,0,1)(0,0,2)[12]** is the model selected by auto.arima(), and while it appears to fit well based on AIC/BIC, there is some residual autocorrelation (as indicated by the Ljung-Box test). This could mean that the model may not have captured all the patterns, particularly seasonal autocorrelation.

**Recommendation**: You might consider re-fitting the model with more seasonal components or further adjusting for potential seasonality or residual autocorrelation.

**2. Model Improvement:**

* **Increase p or q values**: Try testing higher AR or MA terms to capture more dependencies in the residuals, especially after checking ACF/PACF plots again.
* **Seasonal adjustments**: If there’s significant seasonal autocorrelation, you might increase the seasonal AR and MA terms.

**3. Forecasting and Validation:**

* Use this model to make forecasts and validate them with out-of-sample data.
* Perform further diagnostic checks (like residual analysis, out-of-sample validation) to assess the model's validity over time.
* **ARIMA Model Representation:**

The ARIMA model chosen is:

Xt=−0.0636Xt−1−0.0211Xt−2+ϵt−0.6452ϵt−1−0.1829ϵt−12−0.1119ϵt−13+0.0034X\_t = -0.0636 X\_{t-1} - 0.0211 X\_{t-2} + \epsilon\_t - 0.6452 \epsilon\_{t-1} - 0.1829 \epsilon\_{t-12} - 0.1119 \epsilon\_{t-13} + 0.0034

Where:

* XtX\_t is the differenced series,
* ϵt\epsilon\_t is the error term at time tt,
* Seasonal terms are included for 12-period seasonality.

Let me know if you need further help or adjustments!

Une image contenant texte, capture d’écran, Tracé, ligne

Le contenu généré par l’IA peut être incorrect.

Une image contenant diagramme, ligne, Tracé

Le contenu généré par l’IA peut être incorrect.

Une image contenant diagramme, texte, ligne, Parallèle

Le contenu généré par l’IA peut être incorrect.

Une image contenant capture d’écran, diagramme, texte

Le contenu généré par l’IA peut être incorrect.