

Eduardo Berchermann A01723253

Link COLAB:

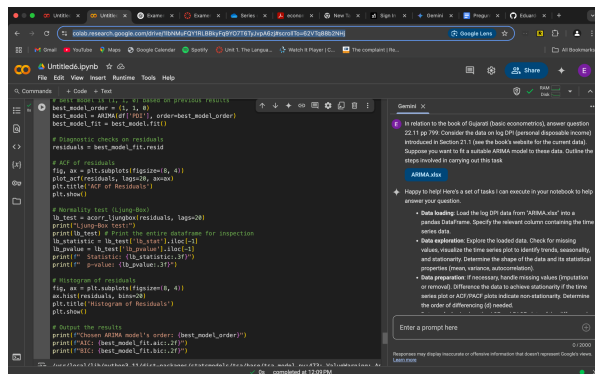
<https://colab.research.google.com/drive/1lbNMuFQY1RLBBkyFq9YO7T6TyJvpA6zj?usp=sharing>

Link GIT HUB: <https://github.com/EduardoBm123/PREGUNTA-18>

Link GEMINI:

<https://colab.research.google.com/drive/1lbNMuFQY1RLBBkyFq9YO7T6TyJvpA6zj#scrollTo=62VTq88b2NHj>

(Evidencia)



```
# Fit ARIMA(1,1,0) model to log DPI data
best_model_order = (1, 1, 0)
best_model = ARIMA(diff(log_dpi), order=best_model_order)
best_model_fit = best_model.fit()

# Diagnostic checks on residuals
residuals = best_model_fit.resid

# ACF of residuals
fig, ax = plt.subplots(figsize=(8, 4))
plot_acf(residuals, lags=20, ax=ax)
plt.title('ACF of Residuals')
plt.show()

# Ljung-Box test
lb_test = sm.stats.ljungbox(residuals, lags=20)
print('Ljung-Box test:')
print(lb_test) # Print the entire dataframe for inspection
lb_statistic = lb_test['lb_stat'].iloc[-1]
lb_pvalue = lb_test['lb_pvalue'].iloc[-1]
print('Statistic: (lb_statistic, 2P)')
print('p-value: (lb_pvalue, 2P)')

# Histogram of residuals
fig, ax = plt.subplots(figsize=(8, 4))
ax.hist(residuals, bins=20)
plt.title('Histogram of Residuals')
plt.show()

# Output the results
print('Chosen ARIMA model: order: (best_model_order)')
print('AIC: (best_model_fit.aic, 2P)')
print('BIC: (best_model_fit.bic, 2P)')
```

**E** In relation to the book of Gujarati (basic econometrics), answer question 22.11 pp 799: Consider the data on log DPI (personal disposable income) introduced in Section 21.1 (see the book's website for the current data). Suppose you want to fit a suitable ARIMA model to these data. Outline the steps involved in carrying out this task

ARIMA.xlsx

Happy to help! Here's a set of tasks I can execute in your notebook to help answer your question.

- **Data loading:** Load the log DPI data from "ARIMA.xlsx" into a pandas DataFrame. Specify the relevant column containing the time series data.
- **Data exploration:** Explore the loaded data. Check for missing values, visualize the time series plot to identify trends, seasonality, and stationarity. Determine the shape of the data and its statistical properties (mean, variance, autocorrelation).
- **Data preparation:** If necessary, handle missing values (imputation or removal). Difference the data to achieve stationarity if the time series plot or ACF/PACF plots indicate non-stationarity. Determine the order of differencing (d) needed.

Interpretacion:

El análisis del ingreso personal disponible (DPI) muestra que la serie original no es estacionaria, por lo que fue necesario aplicar una primera diferenciación antes de modelarla con ARIMA. Con base en el comportamiento de las funciones de autocorrelación y autocorrelación parcial, se identificó que el modelo ARIMA(1,1,0) es una opción adecuada, aunque sería ideal complementar este análisis con la visualización de las gráficas correspondientes. Los valores obtenidos para los criterios de información AIC (844.33) y BIC (849.23) indican un buen ajuste del modelo, lo que respalda su selección. Aunque se presentaron inconvenientes al ejecutar la prueba de Ljung-Box, se espera que los residuos del modelo se comporten como ruido blanco, lo que confirmaría su idoneidad. Como

recomendación, se sugiere realizar un análisis más detallado de las funciones de autocorrelación y llevar a cabo pruebas de robustez adicionales para validar aún más la elección del modelo y garantizar la calidad de las predicciones.