Dupli Checker

Date: 16-11-2025

Plagiarism Scan Report

0% Plagiarism	
0% Exact Match	
0% Partial Match	
100% Unique	

Words	998
Characters	6960
Sentences	40
Paragraphs	91
Read Time	5 minute(s)
Speak Time	7 minute(s)

Content Checked For Plagiarism

4. Exploratory Time Series Analysis
4.1 Univariate Time Series Visualization
The project uses line plots to display individual time series, which is the very first step in understanding temporal patterns. The critical aspects to be investigated include:

- Trend Analysis: This involves the determination of long-term directional movements from visual inspection and moving averages. Trends in financial data may be:
- Upward: Reflecting an increase in reserves, deposits, or circulation.
- Downward: Suggesting contraction or policy-driven reductions
- Stationarity: nonconstant signal around a constant mean

Moving Averages: The application of rolling windows (e.g., 4-week or 12-week moving averages) serves several purposes:

- Smooths short-run volatility
- Reveals underlying trends
- Used as a simple benchmark for forecasts
- Helps in providing changes in the series.

4.2 Multivariate Time Series Relationships
The project investigates ties between paired variables, which is very important to grasp the connectedness of financial systems:
Correlation Analysis:

- B1 vs E1: Central Government deposits in comparison to loans shows the liquidity management between the government and the central bank
- B3 vs E2: State Government patterns depict fiscal dynamics at the sub-national level
- B4 vs E4: Commercial bank correlations point to monetary transmission mechanisms
- B5 vs E5: Cooperative sector trends signal rural and agricultural credit cycles

These ties frequently demonstrate:

- Positive Correlation: Corresponding movements of Variables (e.g., when deposits increase, available funds for loans also increase)
- Negative Correlation: Movement inversely (i.e., when loans are extended, deposits may decrease)
- Lagged Relationships: One time series influencing another with a delay

4.3 Comparative Time Series Analysis
Asset Composition Analysis (D1, D2, D3): Comparing Foreign Currency Assets, Gold, and Rupee Securities over time reveals the RBI's reserve diversification strategy. This multivariate comparison shows:

- Relative proportions changing over time
- Policy shifts in asset allocation
- Response to external economic conditions

Currency Circulation (A1 vs A2): Analyzing notes in circulation versus notes held by the banking department provides insights into:

- Money supply dynamics
- Seasonal patterns in cash demand
- Impact of demonetization or digital payment adoption

5. Time Series Forecasting: ARIMA Methodology

5.1 ARIMA Model Foundations

The project uses an ARIMA model, which is among the most common univariate time series forecasting methods. The orders of the main parameters in ARIMA are three: (p, d, q).

AutoRegressive (AR) Component - p: The AR component models the relationship between an observation and a predetermined number of lagged observations. An AR(p) model is defined as

$y_t=\varphi_1 y_{(t-1)}+\varphi_2 y_{(t-2)}+.+\varphi_p y_{(t-p)}+\varepsilon_t$

This reflects the fact that current values are in a sense determined by recent past values. In financial data, this would imply momentum and persistence in monetary aggregates.

Integrated (I) Component - d: The integration parameter represents the number of differencing operations required to make the series stationary. Stationarity is important since it guarantees:

- Constant mean over time
- Constant variance
- Constant autocovariance structure

Most financial time series are non-stationarity because of the presence of trends or structural changes. Differencing transforms the data:

- First difference: $y't=y_t-y(t-1)$
- Second difference: $y''_t = y't - y'(t-1)$
- For the RBI balance sheet data, differencing once (d=1) usually is enough to make it stationary.

Moving Average (MA) Component - q: The MA component models the dependency between an observation and residual errors from previous predictions:

$y_t = \mu + \varepsilon_t + \theta_1\varepsilon_{(t-1)} + \theta_2\varepsilon_{(t-2)} +. + \theta_q\varepsilon_{(t-q)}$

The moving average captures short-term irregularities and shock effects which persist over several periods.

5.2 Model Identification and Selection

The project uses a systematic approach to identify optimal ARIMA parameters:

Stationarity Testing:

- Augmented Dickey-Fuller (ADF) Test: Tests the null hypothesis that a unit root is present (non-stationary)

- KPSS Test: Tests the null hypothesis that the series is stationary
- Visual Inspection: Plotting the series and checking for constant mean and variance

ACF and PACF Analysis:

- Autocorrelation Function (ACF): Measures correlation between observations at different lags, helping identify the MA order (q)
- Partial Autocorrelation Function (PACF): Measures correlation between observations while controlling for intermediate lags, helping identify the AR order (p)

Automated Selection: The project leverages pmdarima (Python's implementation of R's auto.arima), which:

- Tests multiple combinations of (p, d, q)

- Uses information criteria (AIC, BIC) for model comparison
- Selects the model with the best balance between fit and complexity

Akaike Information Criterion (AIC): AIC = 2k - 2ln(L)

Where k is the number of parameters and L is the likelihood. Lower AIC indicates better model performance with appropriate complexity penalty.

5.3 Model Estimation and Fitting

The ARIMA model is fitted to the historical data using the library statsmodels. The estimation involves:

Maximum Likelihood Estimation: The determination of parameter values that maximize the probability of the given data being observed.

Residual Analysis: Residuals (errors) should be:

- Normally distributed
- Show no autocorrelation (white noise)
- Have constant variance (homoscedastic)
- Poor residual behavior is indicative of model misspecification, suggesting alternative parameterization or model structure.

5.4 Forecasting Process

The project generates 12-step-ahead forecasts (12 weeks into the future). The forecasting process involves:

Point Forecasts: Best estimates of future values based on the fitted model

Confidence Intervals: Typically 95% intervals that represent the uncertainty around predictions. These intervals widen as the forecast horizon increases, reflecting:

- Accumulated prediction errors
- Uncertainty propagation
- Model limitations

Recursive vs Direct Forecasting: ARIMA uses recursive (or iterative) forecasting, where:

- Forecast for t+1 uses actual data up to t
- Forecast for t+2 uses actual data up to t and forecast for t+1
- This continues, with each forecast depending on previous forecasts

5.5 Forecast Evaluation

Although not explicitly detailed in the repository documentation, proper time series forecasting includes:

Train-Test Split: Using historical data for training and withholding recent data for validation

Error Metrics:

- MAE (Mean Absolute Error): Average magnitude of errors
- RMSE (Root Mean Square Error): Penalizes larger errors more heavily
- MAPE (Mean Absolute Percentage Error): Scale-independent metric useful for comparing across series

Forecast Accuracy Assessment: Comparing predicted values against actual values in the test set

Matched Source

Congratulations !

No Plagiarism Found