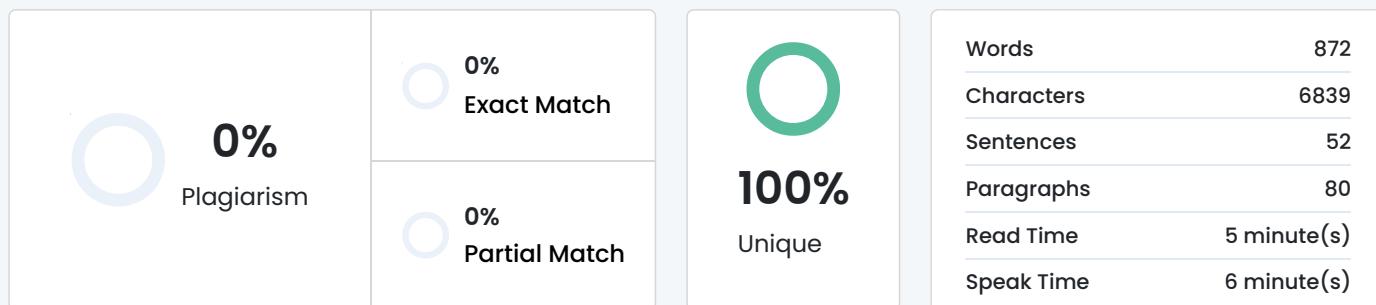


## Plagiarism Scan Report



### Content Checked For Plagiarism

INDIAN INSTITUTE OF TECHNOLOGY JODHPUR  
NH 62 Nagaur Road, Karwar, Jodhpur, Rajasthan

MTech – Data Engineering  
TIME SERIES ANALYSIS

RESERVE BANK of INDIA BALANCE SHEET FORECASTING

Submitted By:

M24DE3011 Aneerban Chowdhury

M24DE3077 Sneha Sawla

M24DE3053 Neha Challa

#### Executive Summary

This report presents an extensive exploration into a time series forecasting project for the weekly balance sheet data of the RBI. The proposed analysis integrates statistical modeling and monetary economics in understanding financial patterns and generating predictions. Through structured preprocessing, in-depth exploratory analysis, correlation evaluation, and forecasting using ARIMA models, this study brings forth how the quantitative modeling of central bank balance sheet movements can be used to support monetary decisions and manage liquidity. We have 2004 to 2025 years of data.

Through analysis, it brings forth meaningful trends in deposits, loans, foreign assets, and currency circulation. The ARIMA-based forecasting part underlines short-run predictability and thus presents data-driven insights for policymakers, financial institutions, and researchers.

## 1. Introduction

### 1.1 Project Overview

The project investigates RBI's weekly balance sheet data to analyze temporal dynamics in India's monetary system. The dataset, sourced from RBI's official Weekly Statistical Supplement, includes high-frequency financial indicators reflecting India's evolving economic environment. The study dissects assets and liabilities, identifying growth patterns, liquidity conditions, and policy implications.

#### Dataset Description

Source: Official RBI Database – Weekly Statistical Supplement (WSS) published by the Reserve Bank of India.

File: RBI\_Liabilities\_and\_Assets.csv

Each row represents the week-ended balance sheet data (usually every Friday).

Category Column Description

Notes A1\_Notes\_in\_Circulation

A2\_Notes\_Held\_in\_Banking\_Department Currency in circulation and held by RBI

Deposits (B) B1\_Deposits\_Central\_Government

B3\_Deposits\_State\_Governments

B4\_Deposits\_Scheduled\_Commercial\_Banks

B5\_Deposits\_Scheduled\_State\_Coop\_Banks Deposits maintained with RBI by various institutions

Liabilities C\_Other\_Liabilities

D\_Total\_Liabilities\_OR\_Assets Overall liabilities and assets on the RBI balance sheet

Assets (D) D1\_Foreign\_Currency\_Assets

D2\_Gold

D3\_Rupee\_Securities Composition of RBI's asset holdings

Loans & Advances (E) E1\_Loans\_Central\_Government, E2\_Loans\_State\_Governments,

E4\_Loans\_Scheduled\_Commercial\_Banks, E5\_Loans\_Scheduled\_State\_Coop\_Banks Loans extended by RBI to various entities

Investments & Others F\_Investments, G\_Other\_Assets Investments and miscellaneous assets

Date Week\_Ended (set as index) Date of the report

### 1.2 Significance of Time Series Analysis

Financial datasets exhibit sequential dependence—past events significantly influence future outcomes.

Understanding these dependencies helps forecast macroeconomic indicators. The project identifies four major characteristics of time series data:

- Temporal Dependence: Each observation relies on preceding values, representing continuity in economic behavior.
- Trend Components: These depict long-term growth or decline influenced by fiscal and monetary policies.
- Seasonality: Cyclical patterns emerge due to recurring phenomena like fiscal year-end adjustments or festival-related spending.
- Stochastic Variations: Unpredictable changes caused by shocks such as policy interventions, global crises, or liquidity injections.

Studying these components allows researchers to construct models that explain historical financial behavior and predict future developments.

#### 1. Deposits vs Loans Relationships

- Correlation and trend comparison between respective entities (B1–E1, B3–E2, etc.)
- Insights into liquidity movements between governments and the banking system

#### 2. Foreign Reserves Composition

- Comparative analysis of D1, D2, D3

→ Shows RBI's balance between foreign currency assets, gold, and rupee securities

#### 3. Currency Circulation Trends

- A1 vs A2 illustrates how much currency is in active circulation vs retained by the RBI

#### 4. Forecast Visualization

- 12-week projections for key variables
- Trend continuation and anomaly detection

## 2. Time Series Data Structure and Components

### 2.1 Dataset Overview

The dataset is a multivariate time series, containing weekly records of financial components that define India's central banking structure. Key variables include:

- Currency Components (A-series): Notes in circulation and notes held in the banking department.
- Deposit Components (B-series): Deposits by central and state governments, scheduled commercial banks, and cooperative banks.
- Asset Components (D-series): Foreign currency assets, gold holdings, and rupee securities that make up RBI's investment and reserve portfolio.
- Loan Components (E-series): Loans issued to central and state governments, commercial banks, and cooperative institutions.
- Other Financial Metrics: Investments, total liabilities, and total assets—providing a complete view of RBI's financial position.

### 2.2 Temporal Indexing

Time indexing ensures that all records are chronologically ordered using the Week\_Ended field. The conversion to Python's DatetimeIndex structure (`pd.to_datetime()`) facilitates:

- Temporal Ordering: Preserves data sequence integrity.
- Frequency Recognition: Detects weekly frequency for accurate analysis.
- Rolling Calculations: Enables moving averages and trend analysis.
- Model Compatibility: Many forecasting models, such as ARIMA, require datetime-indexed data.

## 3. Time Series Data Preprocessing Techniques

### 3.1 Missing Value Treatment

Time series data often contains gaps due to holidays, system failures, or reporting delays. The project addresses missing values through several strategies:

**Forward Fill (FFILL):** Propagates the last valid observation forward, appropriate when values change slowly and the gap is small. This assumes that the most recent observation is the best estimate for missing periods.

**Backward Fill (BFILL):** Uses the next valid observation to fill gaps, suitable for preliminary data cleaning.

**Interpolation:** For time series data, linear or polynomial interpolation can estimate missing values based on surrounding observations, maintaining the continuity of trends.

The choice of method depends on the nature of missing data and the underlying assumption about how values evolve over time.

### 3.2 Data Formatting and Standardization

Financial data often requires cleaning of monetary formats, including:

- Removing currency symbols and separators
- Converting string representations to numeric types
- Handling different scaling factors (thousands, millions, crores)
- Ensuring consistent decimal representations

This preprocessing is essential because statistical models require numeric input in standardized formats.

## Matched Source

Congratulations !

**No Plagiarism Found**

