TIME SERIES ECONOMETRICS

PROJECT TITLE : Forecasting (2024) India's Space Launches: An ARIMA-Based Time Series Analysis.

Introduction

In recent decades, India has increasingly become an important player in space exploration and satellite technology, marked by frequent launches of satellites and other objects into outer space. The Indian Space Research Organisation (ISRO) has been responsible for these advancements, contributing to both national interests and international collaborations. Forecasting the number of objects launched annually is crucial for ISRO's strategic planning, budget allocation, and resource management.

This report applies the ARIMA (Auto Regressive Integrated Moving Average) model, a powerful time series forecasting method, to predict India's future space activities. Specifically, it seeks to provide a forecast for the year 2024 based on historical launch data from 1975 to 2023. This analysis follows the stages of data preprocessing, model selection, diagnostic testing, and forecasting, providing a comprehensive approach to assessing and predicting India's space program trajectory.

Objective

The primary objective of this report is to forecast the number of objects India is likely to launch into outer space in 2024 using historical launch data. This will enable:

1 .Exploration of Historical Patterns: Analyzing trends and any periodic increases or decreases in launch frequencies.

2. Model Development: Identifying and fitting the best ARIMA model based on the data's characteristics.

3. Forecasting and Uncertainty Assessment: Generating a forecast with confidence intervals to account for variability.

4. Strategic Insights: Providing actionable insights for policymakers and space program administrators.

This report leverages the ARIMA model's capacity to accommodate non-stationary time series data, making it ideal for modelling the observed trends in India's space launch data.

Data Overview

Data Source and Description

The dataset, sourced from the United Nations Office for Outer Space Affairs (UNOOSA), consists of annual records of the Number of objects launched by India into outer space. It spans 39 years from 1975 to 2023. The dataset contains two main variables:

• Year: The year corresponding to each record.

• No. of Objects Launched: The number of objects launched by India in each respective year

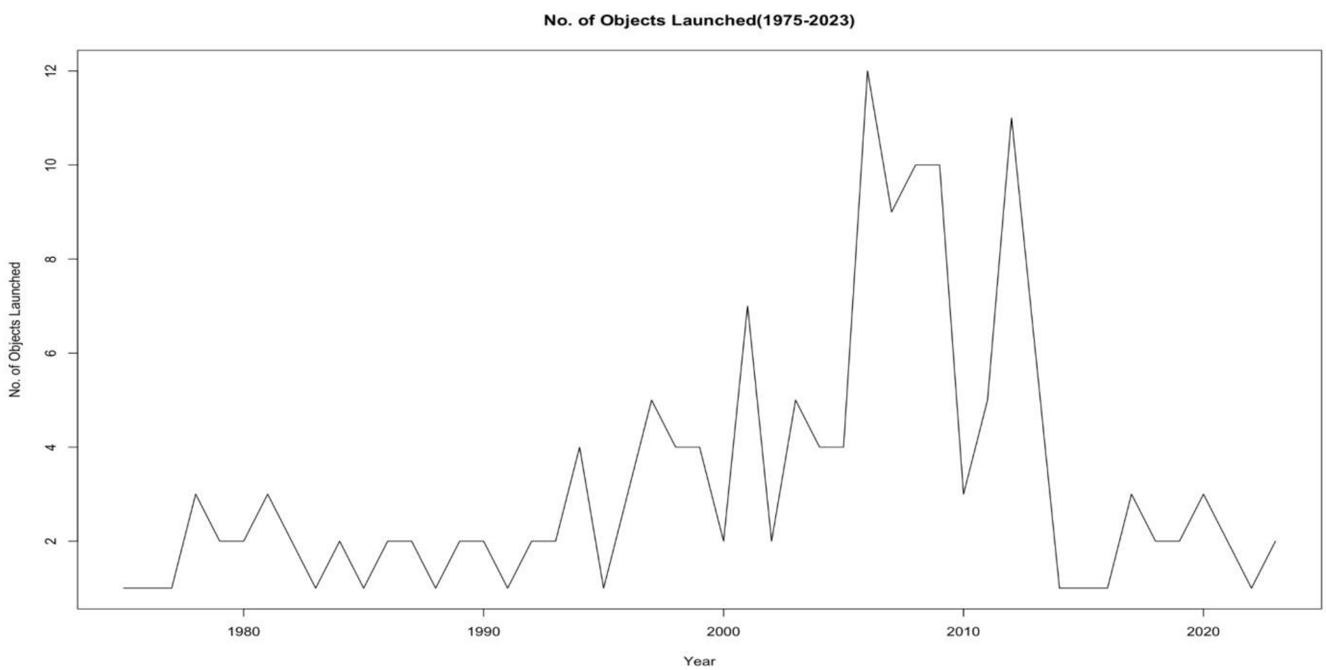
.

Initial Data Visualization

A time series plot of the data reveals fluctuations in launch numbers. Notable increases in launch frequency align with advancements in India's space program, reflecting ISRO's expanding capabilities and the nation's growing ambitions in space technology. Recent years indicate a steady rise, suggesting potential for further growth in launch activity.

```
# Plot the data to check for trend and stationarity
plot(launches_ts, main = "No. of Objects Launched(1975-2023)",
     xlab = "Year", ylab = "No. of Objects Launched")
```

The initial plot provides a visual baseline for observing trends and variability in the data, laying the foundation for further analysis.



Stationarity Check

Time series forecasting requires data to be stationary, meaning its statistical properties should not change over time. To assess stationarity, we applied the Augmented Dickey-Fuller (ADF) test, a statistical test that determines if a unit root (a sign of non-stationarity) is present.

• **Test Result: The ADF test returned a p-value of 0.6825, indicating non-stationarity at a 5% significance level**.

Differencing

To address non-stationarity, first-order differencing was applied to the data. This transformation involves subtracting each value from the preceding one, which can stabilize the mean and remove trends.

• **Post-Differencing ADF Test Result: Re-testing on the differenced series yielded a p-value of 0.01, confirming stationarity.**

```
# Check stationarity using the Augmented Dickey-Fuller Test
adf_test <- adf.test(launches_ts)
print(adf_test)

# If non-stationary (p-value > 0.05), difference the series and check
again
if(adf_test$p.value > 0.05) {
  launches_ts_diff <- diff(launches_ts)

  # Suppress warning and run the test again
  adf_test_diff <- suppressWarnings(adf.test(launches_ts_diff))

  # Print results and check if p-value is below 0.05 (indicating
stationarity)
  print(adf_test_diff)
}
```

```
OUTPUT:
```

```


Augmented Dickey-Fuller Test

data: launches_ts
Dickey-Fuller = -1.7294, Lag order = 3, p-value = 0.6825
alternative hypothesis: stationary
```

After first differencing,

```
Augmented Dickey-Fuller Test

data:  launches_ts_diff
Dickey-Fuller = -5.1208, Lag order = 3, p-value = 0.01
alternative hypothesis: stationary
```

ARIMA Model Selection

Model Identification

The auto.arima() function in R, which automatically selects the optimal ARIMA model by minimizing the Akaike Information Criterion (AIC), identified ARIMA(0,1,1) as the best-fit model. This model structure suggests:

• p = 0: No autoregressive terms.
• d = 1: One differencing operation to ensure stationarity.
• q = 1: One moving average term, which captures residual patterns in the data.

Model Summary

The ARIMA(0,1,1) model yielded:

• Log-Likelihood: -108.55
• AIC: 221.1
• BIC: 224.84
• MA(1) Coefficient: -0.5524 (Standard Error: 0.1266)

These metrics indicate the model’s fit, with lower AIC and BIC values suggesting an effective balance between model complexity and goodness of fit.

```
# Determine ARIMA model order (p, d, q) using auto.arima()
model <- auto.arima(launches_ts, seasonal = FALSE, stepwise = FALSE,
approximation = FALSE)
summary(model)
```

```
OUTPUT:
```

```
Series: launches_ts
ARIMA(0,1,1)

Coefficients:
          ma1
      -0.5524
s.e.   0.1266

sigma^2 = 5.466:  log likelihood = -108.55
AIC=221.1   AICc=221.37   BIC=224.84

Training set error measures:
                     ME      RMSE       MAE       MPE     MAPE
MASE       ACF1
Training set 0.03694212 2.289668 1.474651 -35.19711 64.53581
0.8528101 0.04311669
```

Diagnostic Checks

Residual Analysis

Residuals from the ARIMA(0,1,1) model were analyzed to ensure that they exhibit randomness, as expected if the model has captured all significant patterns in the data. The residual plot showed no discernible pattern, suggesting the residuals are white noise.

Ljung-Box Test

The Ljung-Box test was conducted to confirm that the residuals were uncorrelated. With a p-value of 0.6955, the test supports the hypothesis that the residuals are uncorrelated, further validating the model.

```
# Diagnostic checks
# Plot residuals and ACF of residuals
checkresiduals(model)
```

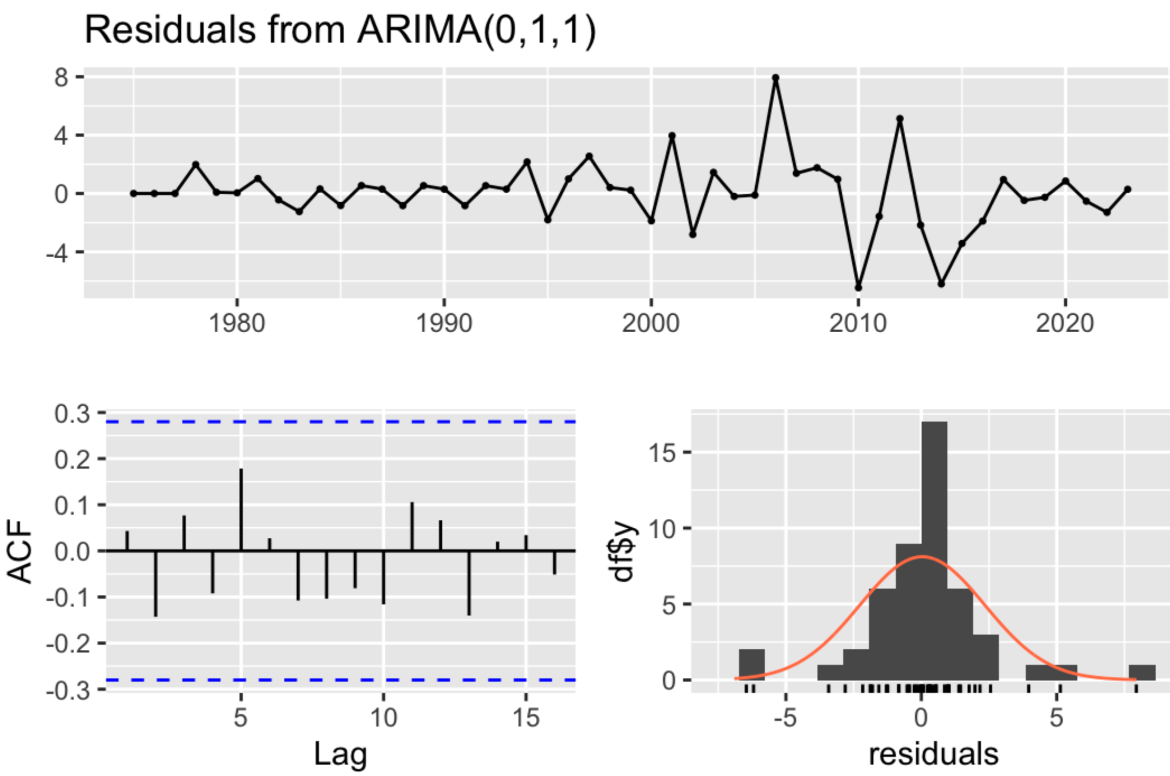
```
OUTPUT:
```

```
Ljung-Box test
```

```
data:  Residuals from ARIMA(0,1,1)
Q* = 6.4373, df = 9, p-value = 0.6955

Model df: 1.   Total lags used: 10
```

```
PLOT:
```



Forecasting for 2024

The final ARIMA model was used to generate a forecast for the number of objects India is expected to launch in 2024. The forecast includes point estimates and confidence intervals to account for uncertainty.

Forecast Summary

• Point Forecast for 2024: 1.84 objects

• 80% Confidence Interval: [-1.15, 4.84]

• 95% Confidence Interval: [-2.74, 6.42]

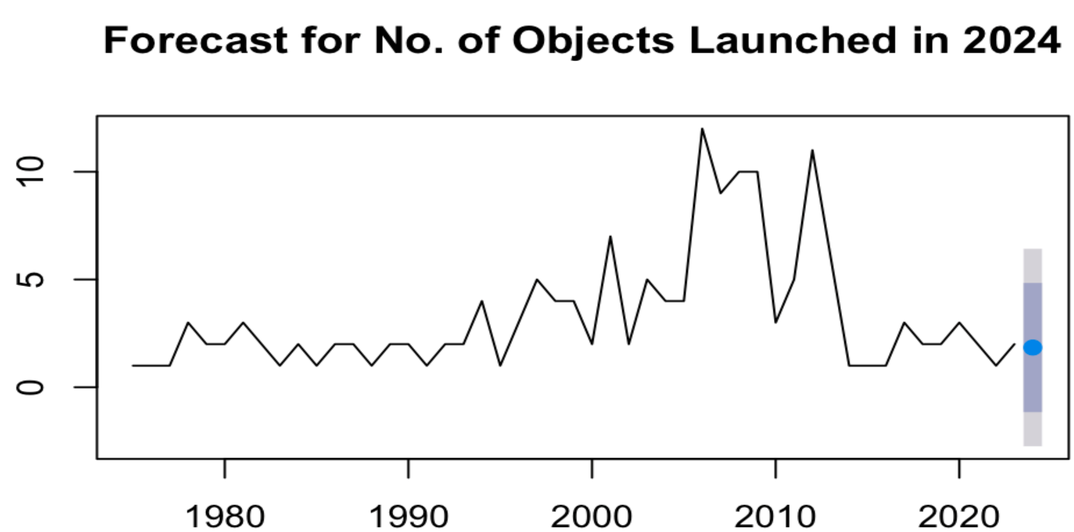
The forecast value suggests a launch count close to recent historical averages, though the wide confidence intervals indicate significant variability.

```
# Forecast for 2024
forecast_2024 <- forecast(model, h = 1)
print(forecast_2024)
plot(forecast_2024, main = "Forecast for No. of Objects Launched in
2024")
```

```
OUTPUT
```

```
Point Forecast     Lo 80     Hi 80      Lo 95     Hi 95
2024       1.841589 -1.15452 4.837698 -2.740565 6.423742
```

```
PLOT:
```



diagnostics confirming its adequacy.

• Forecast Implications: The point forecast for 2024 aligns with recent historical activity, although the large confidence intervals highlight potential variability in future launch activity.

• Strategic Insight: ISRO and policymakers can leverage these insights for resource allocation and strategic planning, considering the model's forecasted value and its uncertainty bounds.

This analysis highlights the usefulness of ARIMA modeling in strategic planning for government and space agencies. Future research could consider multivariate models incorporating additional factors, such as technological advancements and budget changes, to improve forecast accuracy.