

=====

EXPLORATORY DATA ANALYSIS – SUMMARY REPORT

=====



Time Period: 1946-01-01 to 2025-11-17



Number of Indicators: 5



Total Observations: 26167

INDICATOR SUMMARY:



unemployment:

Mean: 5.67

Std Dev: 1.71

Min: 2.50

Max: 14.80

Missing: 25235 (96.4%)

Stationarity: Stationary (p=0.0018)

Outliers: 8



cpi:

Mean: 123.42

Std Dev: 88.82

Min: 21.48

Max: 324.37

Missing: 25222 (96.4%)

Stationarity: Non-Stationary (p=1.0000)

Outliers: 0



gdp:

Mean: 7600.85

Std Dev: 7972.38

Min: 243.16

Max: 30485.73

Missing: 25853 (98.8%)

Stationarity: Non-Stationary (p=1.0000)

Outliers: 0



dff:

Mean: 4.61

Std Dev: 3.56

Min: 0.04

Max: 22.36

Missing: 94 (0.4%)

Stationarity: Stationary (p=0.0490)

Outliers: 840

◆ consumer:
Mean: 84.90
Std Dev: 13.22
Min: 50.00
Max: 112.00
Missing: 25502 (97.5%)
Stationarity: Non-Stationary (p=0.0991)
Outliers: 0

=====

KEY FINDINGS:

=====

- ✓ Data Quality:
 - 101906 missing values require attention
- ✓ Correlations:
 - 1 pairs with strong correlation ($|r| > 0.7$)
- ✓ Stationarity:
 - 2/5 indicators are stationary
 - 3 require differencing for modeling

=====

NEXT STEPS:

=====

1. Handle missing values (forward fill, interpolation)
 2. Apply differencing to non-stationary series
 3. Feature engineering (lags, rolling statistics)
 4. Build forecasting models (ARIMA, Prophet, LSTM)
 5. Evaluate model performance
- =====