Time Series Analysis of the Industrial Production Index for Food and Beverages in Germany

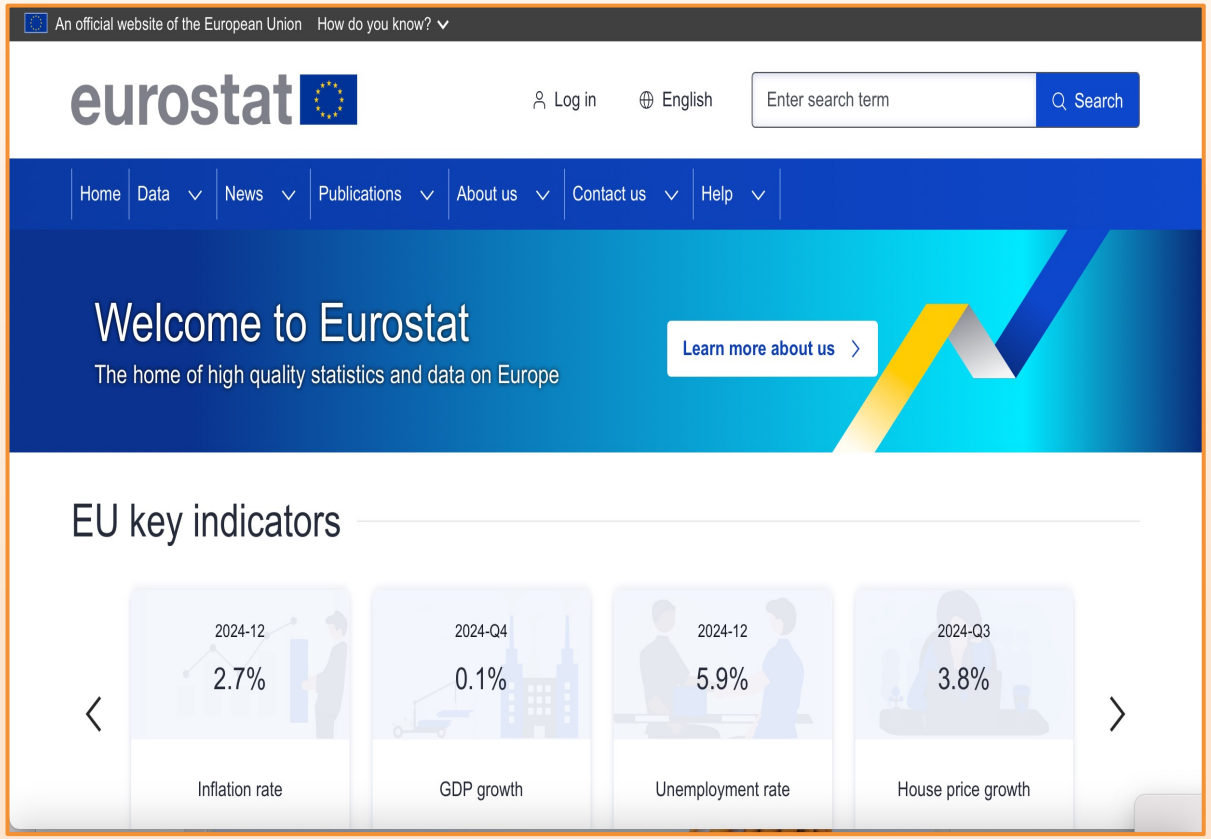
A Data-Driven Approach to Forecasting Trends and Market Dynamics

Conducted by :

Nazila Fazeli: M.Sc. Artificial Intelligence

Eurostat : The Statistical Office of the European Union

- Free & Open Data Access
- High Credibility & Standardization
- Wide Coverage & Diverse Sectors
- Advanced Data Filtering
- Detailed Metadata for Transparency
- Long-Term & Regularly Updated Time Series



https://ec.europa.eu/eurostat/databrowser/product/view/sts_inpr_m?category=sts.sts_ind.sts_ind_prod

Dataset Overview

Dataset Name
Manufacture of Food Products & Beverages

Source
Eurostat - Short-Term Business Statistics

Sector
Production in Industry - Monthly Data

Time Range
Jan 1991 – Oct 2024 (406 Months)

Country
Germany

Data Type
Unadjusted Monthly Production Index

Dataset Overview

What Does This Data Represent?

- ✓ Industrial Production Index (IPI)
- ✓ Relative to base year 2021 (Index = 100)

Why This Dataset?

- ✓ Key Economic Sector
- ✓ Real-World Impact
- ✓ Long-Term Data Availability
- ✓ Indexed Data for Comparisons
- ✓ Public & Free Access

Main Objective

To Analyze and Forecast the Industrial Production Index (IPI) for the food & beverage sector in Germany using time series modeling

Key Objectives:

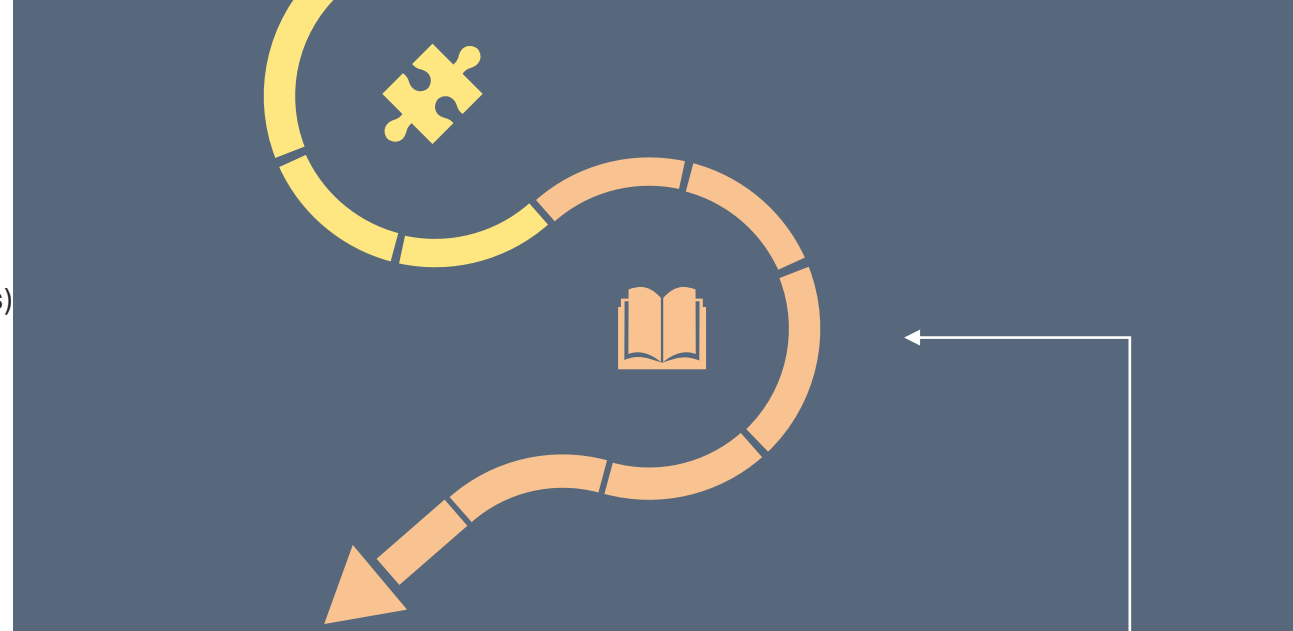
01 Data Preparation & Exploration

02 Model Selection & Training

03 Forecasting & Evaluation

Data Pre-processing

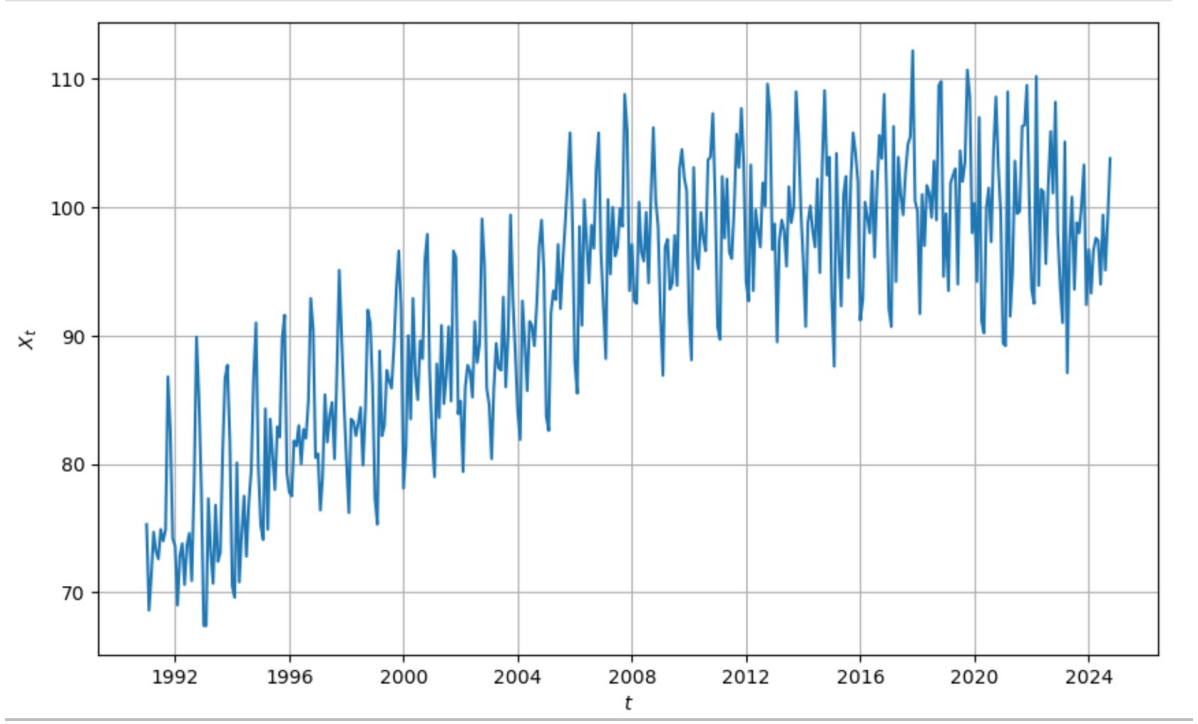
- **Selected Relevant Columns:** TIME_PERIOD & OBS_VALUE
- **Checked for missing values:** None found (406 complete records)
- **Fixed Numerical Formatting:** Replaced commas with dots
- **Converting to Time Series**



	TIME_PERIOD	OBS_VALUE
1	1991-01	75,3
2	1991-02	68,6
3	1991-03	71,2
4	1991-04	74,7
5	1991-05	73,2
6	1991-06	72,6
7	1991-07	74,9
8	1991-08	74
9	1991-09	74,9
10	1991-10	86,8
11	1991-11	82,6
12	1991-12	74,2
13	1992-01	73,6
14	1992-02	69
15	1992-03	72,8
16	1992-04	73,8
17	1992-05	70,6
18	1992-06	73,7
19	1992-07	74,6
20	1992-08	70,9

DATAFLOW	LAST_UPDATE	freq	indic_bt	unit	geo	TIME_PERIOD	OBS_VALUE	OBS_FLAG
All	All	All	All	All	All	All	All	All
ESTAT:STS_INPR_M(1.0)	09/01/25 11:00:00	Monthly	Production (volume)	Index, 2021=100	Germany	1991-01	75.30	
ESTAT:STS_INPR_M(1.0)	09/01/25 11:00:00	Monthly	Production (volume)	Index, 2021=100	Germany	1991-02	68.60	
ESTAT:STS_INPR_M(1.0)	09/01/25 11:00:00	Monthly	Production (volume)	Index, 2021=100	Germany	1991-03	71.20	
ESTAT:STS_INPR_M(1.0)	09/01/25 11:00:00	Monthly	Production (volume)	Index, 2021=100	Germany	1991-04	74.70	
ESTAT:STS_INPR_M(1.0)	09/01/25 11:00:00	Monthly	Production (volume)	Index, 2021=100	Germany	1991-05	73.20	
ESTAT:STS_INPR_M(1.0)	09/01/25 11:00:00	Monthly	Production (volume)	Index, 2021=100	Germany	1991-06	72.60	
ESTAT:STS_INPR_M(1.0)	09/01/25 11:00:00	Monthly	Production (volume)	Index, 2021=100	Germany	1991-07	74.90	
ESTAT:STS_INPR_M(1.0)	09/01/25	Monthly	Production	Index,	Germany	1991-08	74.00	

Initial Data Analysis & Stationarity Check Style



Plotted time series

Observed upward trend & frequency variations

Checked stationarity with KPSS Test

Confirmed non-stationarity

Trend in mean detected

Transformation needed before modeling

Data Transformation & Stationarity

01 **First differencing**

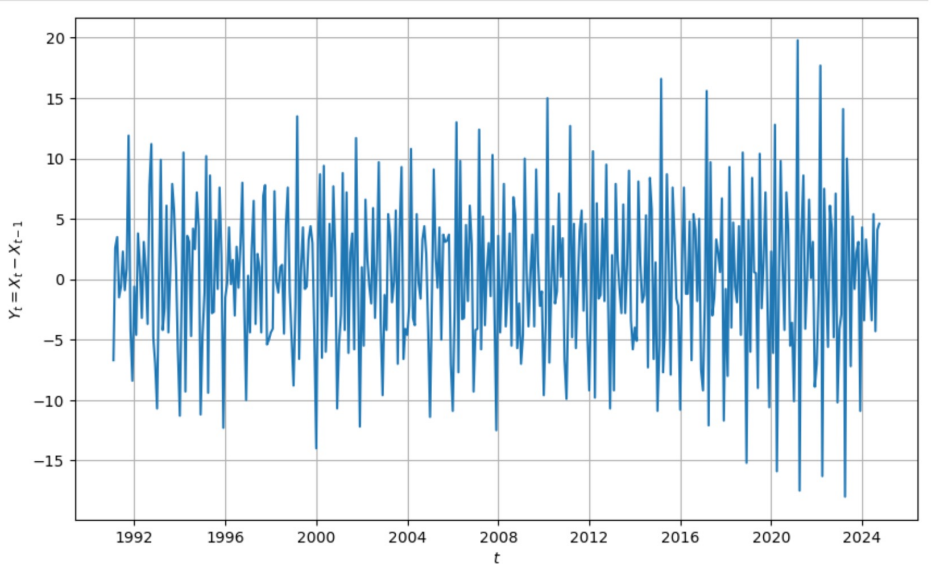
Removed trend in mean

02 **KPSS Test after differencing**

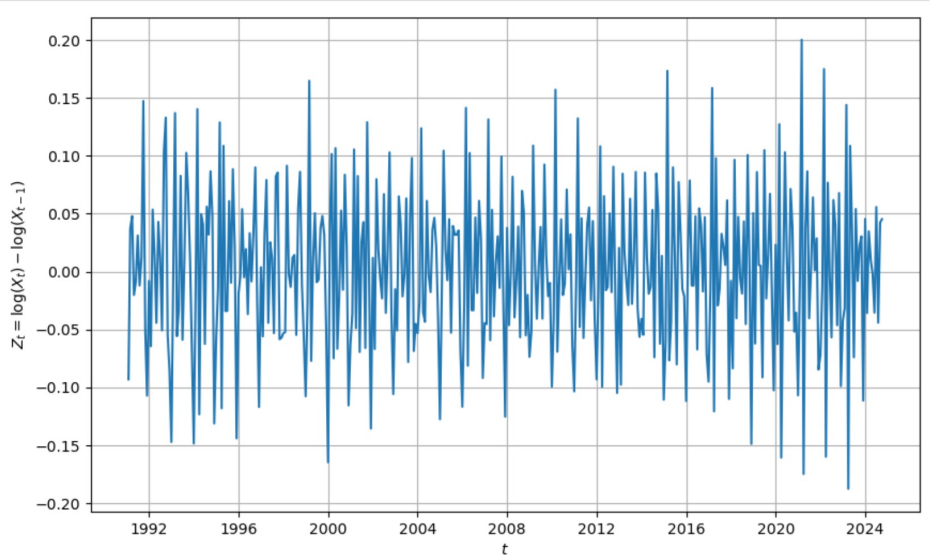
Confirmed stationarity

03 **Log transformation applied**

Stabilized variance



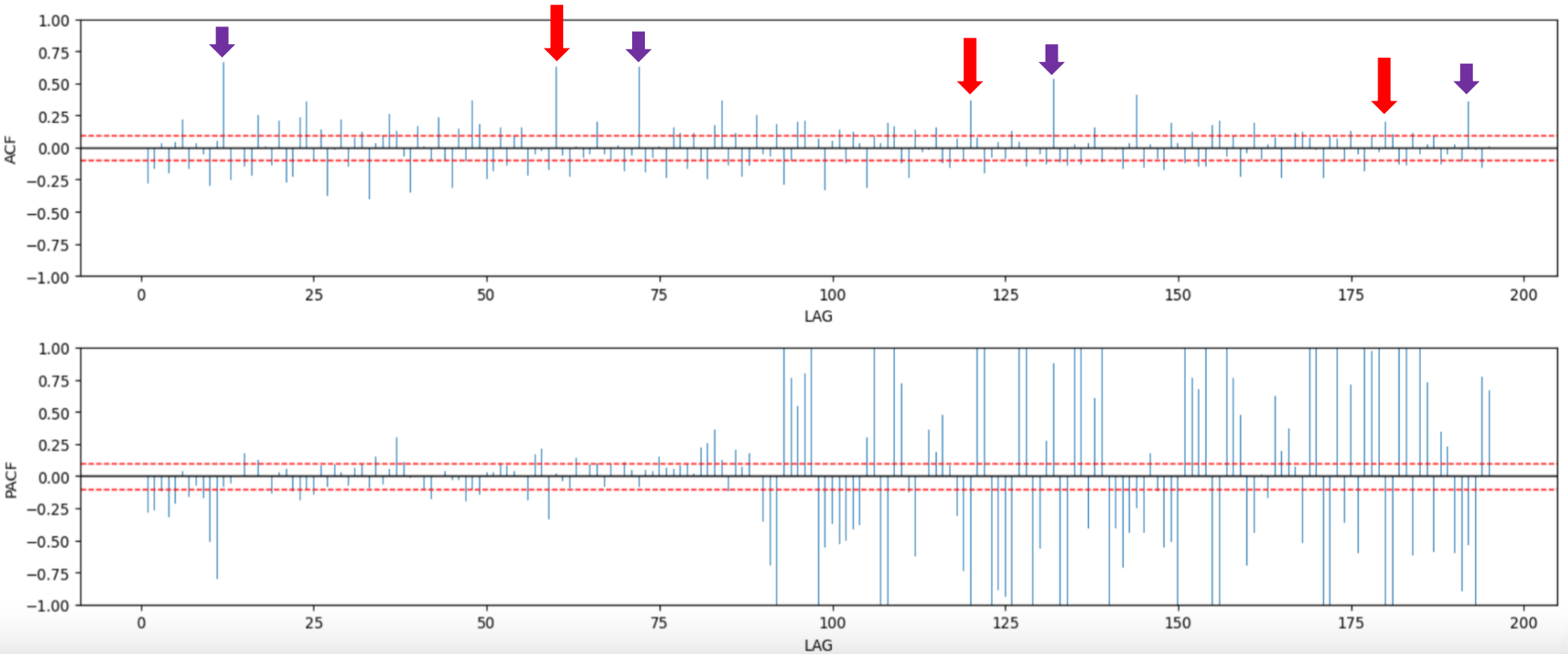
KPSS Test p-value: 0.935

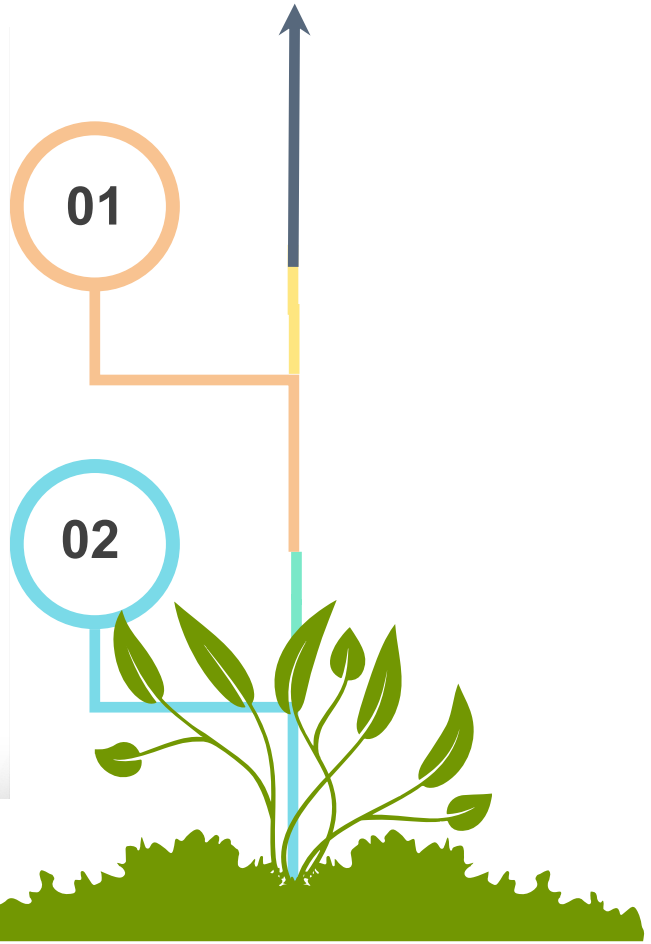


KPSS Test p-value: 0.908

Identifying Seasonality & Seasonal Parameters

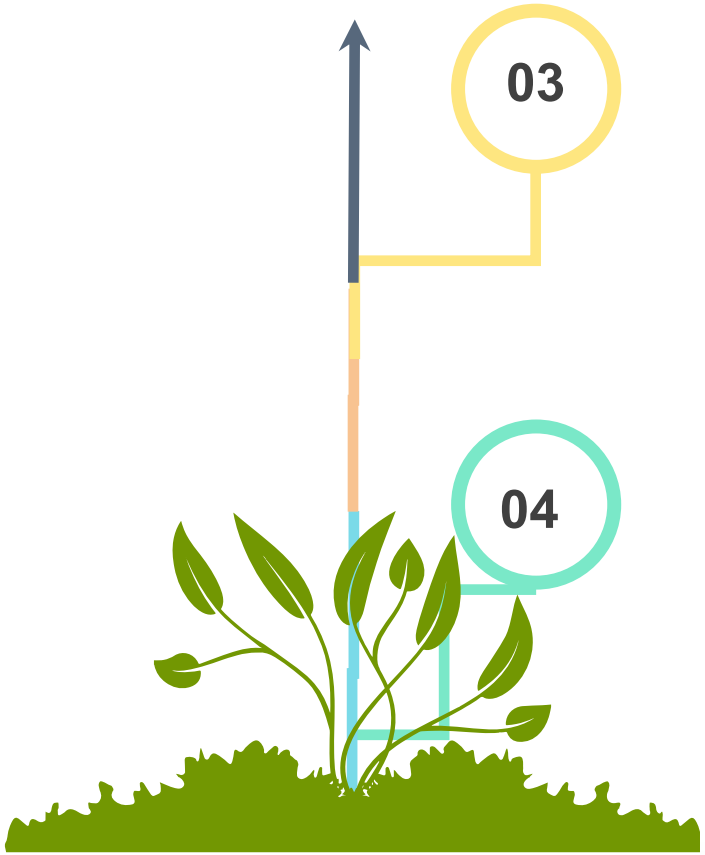
- **ACF & PACF Plotted after first differencing**
- **Detected seasonality every 60 lags (~5 years)**
- **SARIMA Seasonal parameters determined: (P, D, Q, S)**

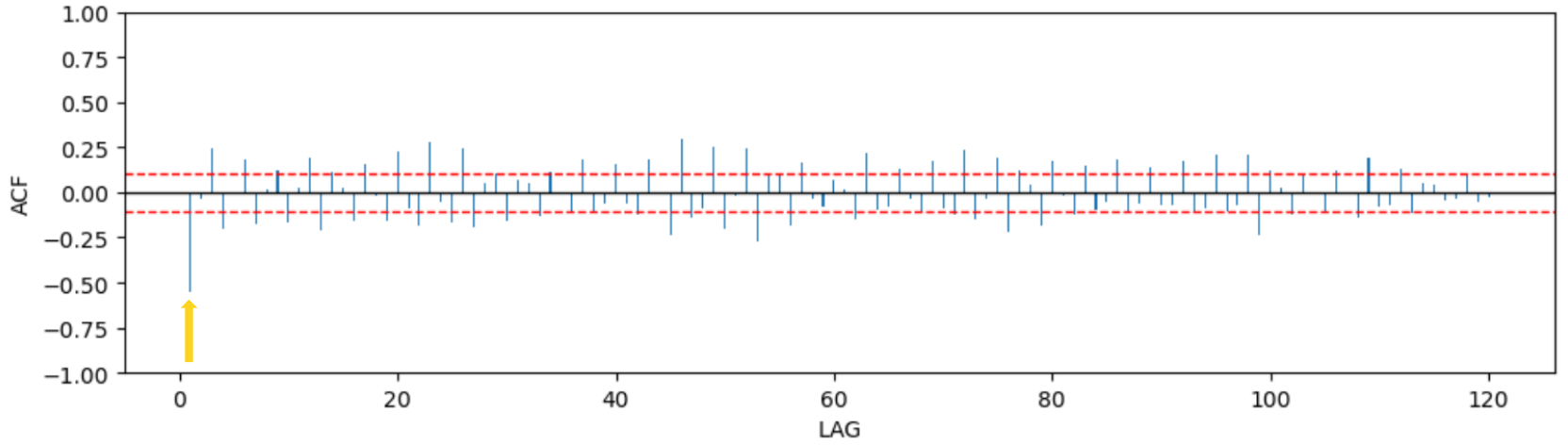


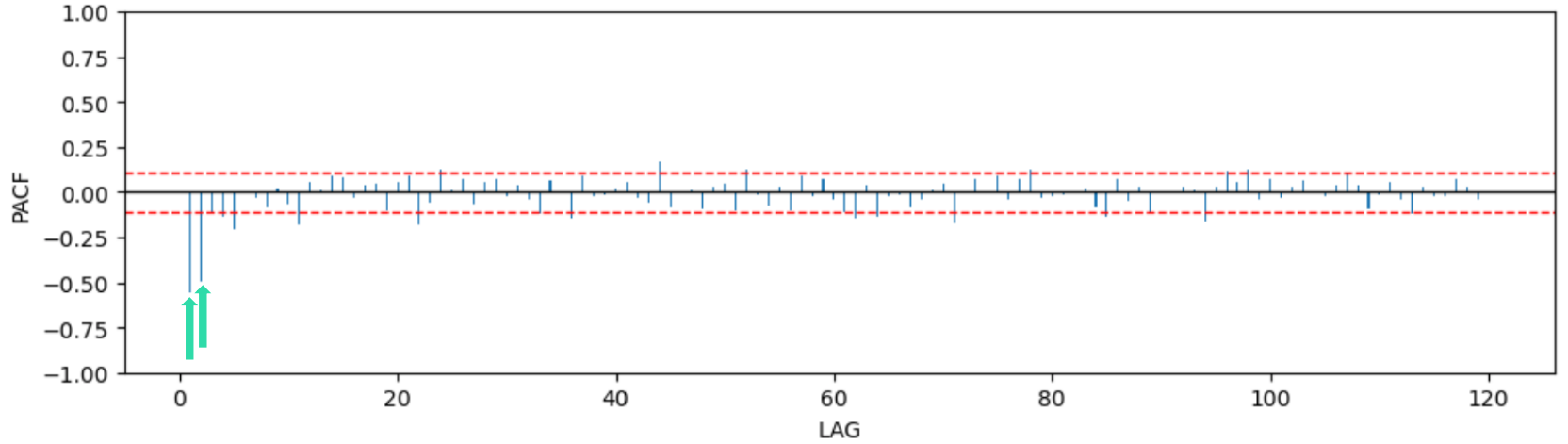


SARIMA Model Non-Seasonal Parameter Estimation

- **ACF & PACF Plotted after removing seasonal lags**
- **Determined Non-Seasonal SARIMA parameters: (p, d, q)**



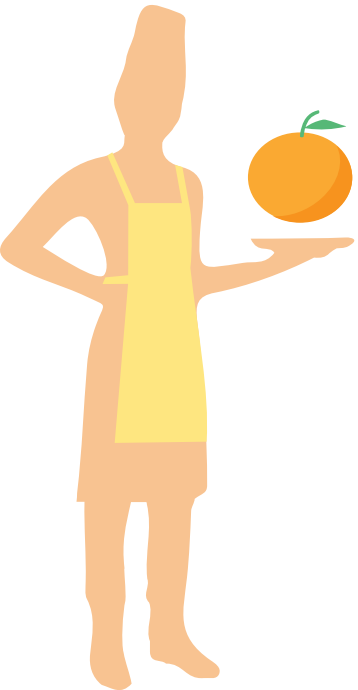




Model Selection

Suggested Models:

1. SARIMA(1, 1, 1)(1, 1, 1) _ {60}
2. SARIMA(1, 1, 1)(1, 1, 0) _ {60}
3. SARIMA(1, 1, 1)(0, 1, 0) _ {60}
4. SARIMA(2, 1, 1)(1, 1, 1) _ {60}
5. SARIMA(2, 1, 1)(1, 1, 0) _ {60}
6. SARIMA(2, 1, 1)(0, 1, 0) _ {60}
7. SARIMA(2, 1, 2)(0, 1, 0) _ {60}



SARIMAX Results

Dep. Variable:	OBS_VALUE	No. Observations:	394
Model:	ARIMA(2, 1, 1)x(0, 1, [], 60)	Log Likelihood	-820.325
Date:	Mon, 03 Feb 2025	AIC	1648.651
Time:	19:19:55	BIC	1663.883
Sample:	01-01-1991	HQIC	1654.725
	- 10-01-2023		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	-0.3518	0.066	-5.332	0.000	-0.481	-0.222
ar.L2	-0.1985	0.069	-2.869	0.004	-0.334	-0.063
ma.L1	-0.6745	0.061	-11.062	0.000	-0.794	-0.555
sigma2	8.0443	0.663	12.128	0.000	6.744	9.344

Ljung-Box (L1) (Q):	0.25	Jarque-Bera (JB):	0.14
Prob(Q):	0.61	Prob(JB):	0.93
Heteroskedasticity (H):	1.57	Skew:	0.05
Prob(H) (two-sided):	0.02	Kurtosis:	2.99

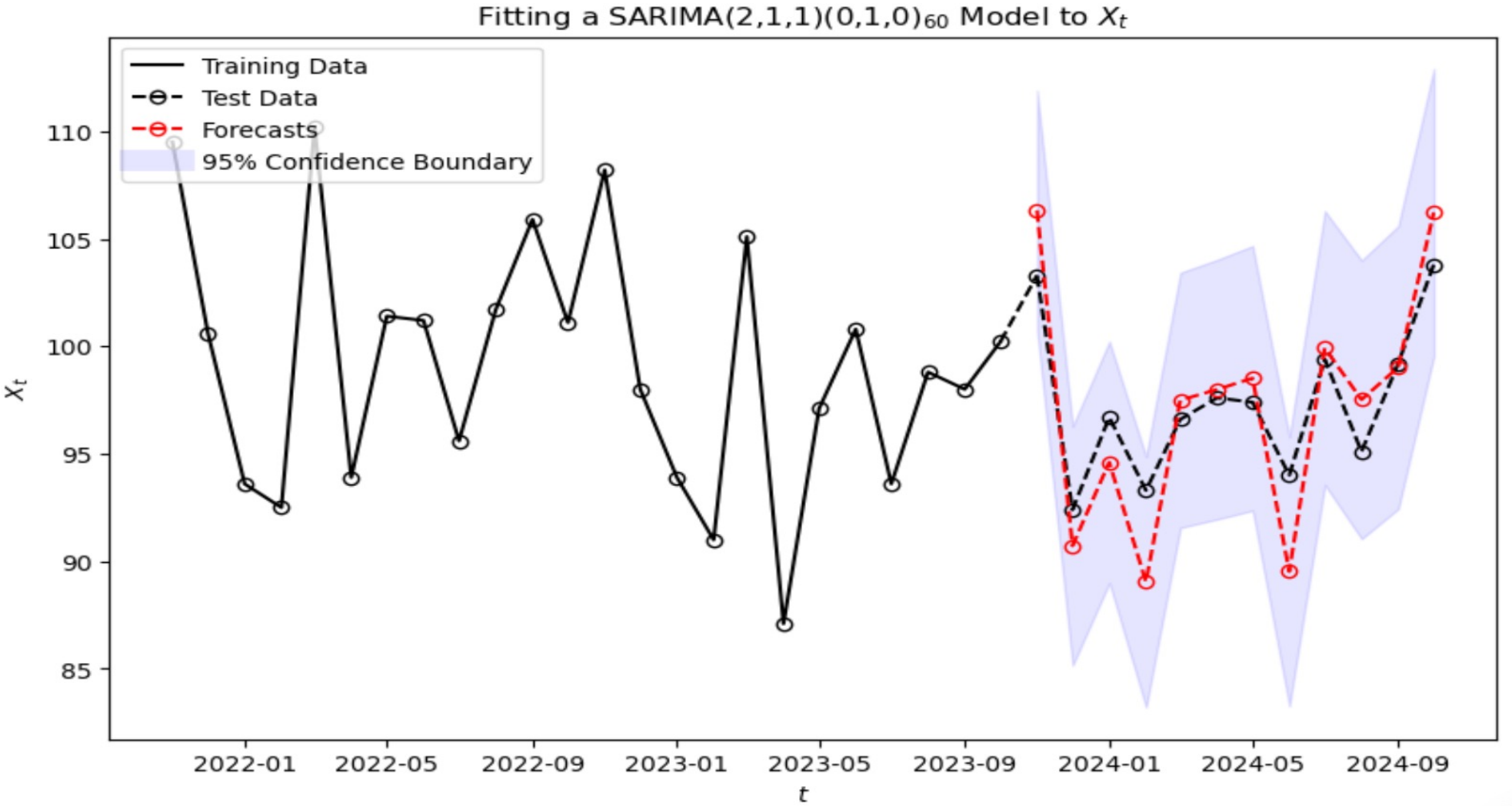
Evaluation Criteria:

- ✓ **p-values for significance**
- ✓ **AIC, BIC scores**
- ✓ **Ljung-Box test for residuals**

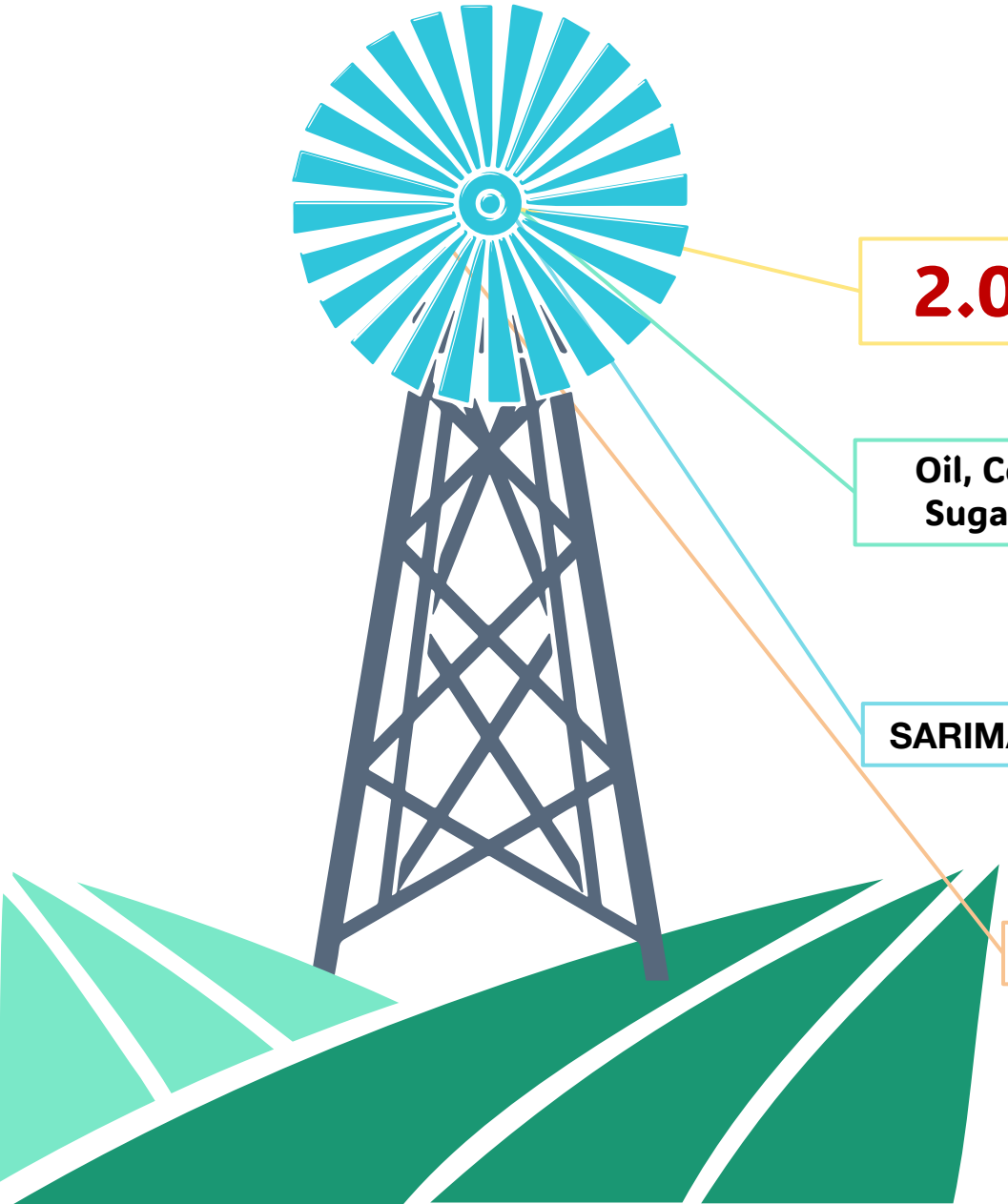
Forecasting Using Selected Model

1. **Using SARIMA (2,1,1)(0,1,0)_{60}:**

   - ✓ **Forecasted for the next 12 months**
   - ✓ **Evaluated the model performance using MAPE (2.02%)**



Model Evaluation & External Factors



2.02%

MAPE (Mean Absolute Percentage Error)

Oil, Cereals, Sugar GDP

Tested external variables

SARIMA (2,1,1) (0,1,0,60)

SARIMA model automatically dropped exogenous variables

better variables?

Further improvements might be possible with better variables

Conclusion & Future Work

SARIMA model successfully forecasted next 12 months.

MAPE showed high accuracy (2.02%).

External factors did not improve predictions

Future improvements:

- Investigate the identified 12-month seasonal pattern observed after every 60-month seasonality.
- exploring other exogenous factors.
- Test advanced machine learning models for improved forecasting.

THANK YOU

Group2

“You are already naked;

There is no reason not to follow your heart.”

— Steve Jobs