



Time Series Analysis

Ngo Thai Hoa

Faculty of Data Science and Artificial Intelligence
College of Technology, National Economics University



Learning Outcomes

1. Understand time series data
2. Time series decomposition (Components of time series)
3. Autocorrelation and lagged series
4. Autocorrelation Function – ACF and Correlogram



Table of Contents

- 1. Time series analysis**
- 2. Components of Time Series**
- 3. Time Series Representations**
- 4. Auto-Correlation**
- 5. Advanced Time Series Analysis**
- 6. Time Series Forecasting**



1. Time Series Analysis



Time Series Data

A time series is a sequence of data points collected, recorded, or measured at successive, equally spaced points in time.

Examples: stock prices, daily/monthly temperature, or quarterly/annual business performance data.

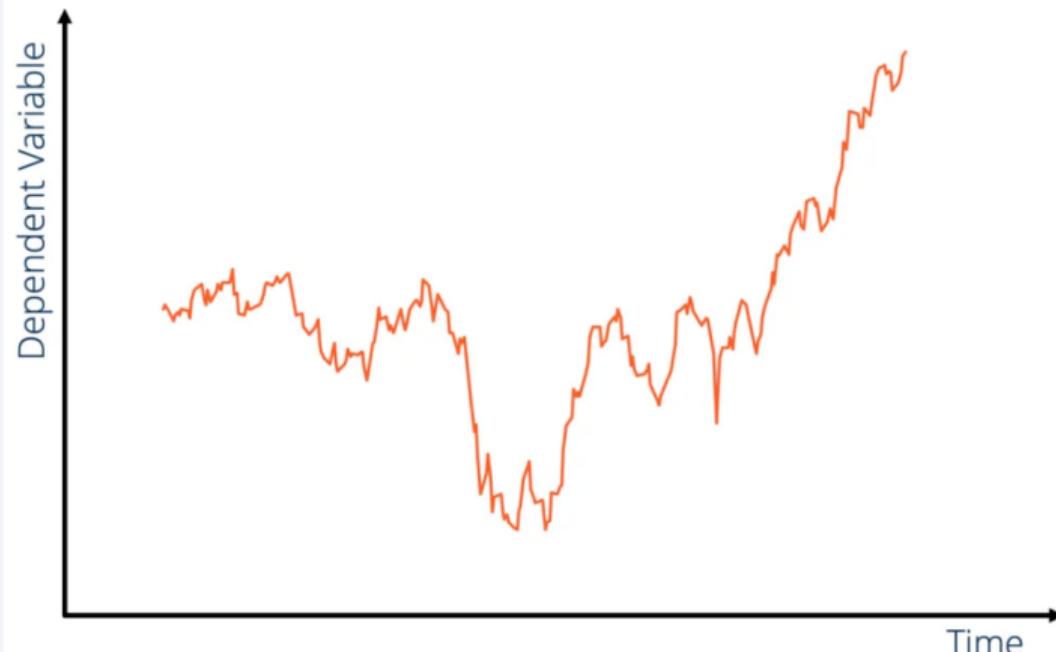
Date	DayOf Week	Time_t	Sandwiches-Asst	Bottled Water	Chips-Asst	Cinnamon Rolls
1/22/2015	Thu	1	15	10	12	9
1/23/2015	Fri	2	0	8	3	1
1/26/2015	Mon	3	5	8	4	4
1/27/2015	Tue	4	11	10	9	6
1/28/2015	Wed	5	4	7	8	3
1/29/2015	Thu	6	7	11	15	2
1/30/2015	Fri	7	0	3	2	0
2/2/2015	Mon	8	6	4	2	0
2/3/2015	Tue	9	20	11	14	3
2/4/2015	Wed	10	5	8	2	5
2/5/2015	Thu	11	26	12	5	12
2/6/2015	Fri	12	1	4	2	7
2/9/2015	Mon	13	8	11	8	4
2/10/2015	Tue	14	20	11	6	7

1.1 Concept of Time Series Data

Each data point represents observations or measurements taken over time.

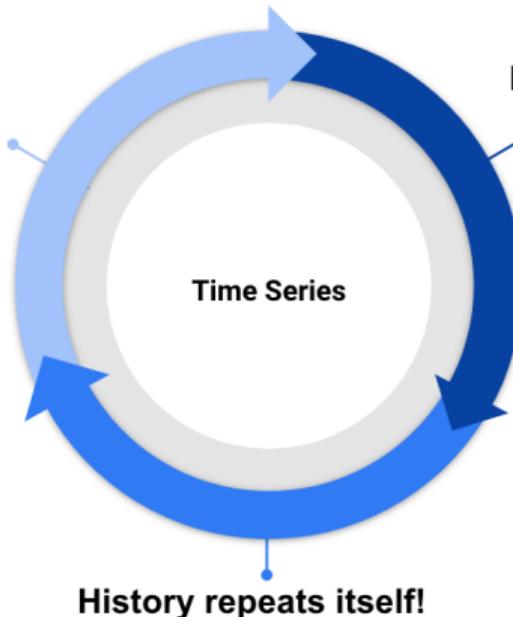
- **x-axis:** Time
- **y-axis:** Dependent variable values

Time-Series Analysis



1.2 Why Time Series Data Is Important

Phổ biến trong nhiều bài toán



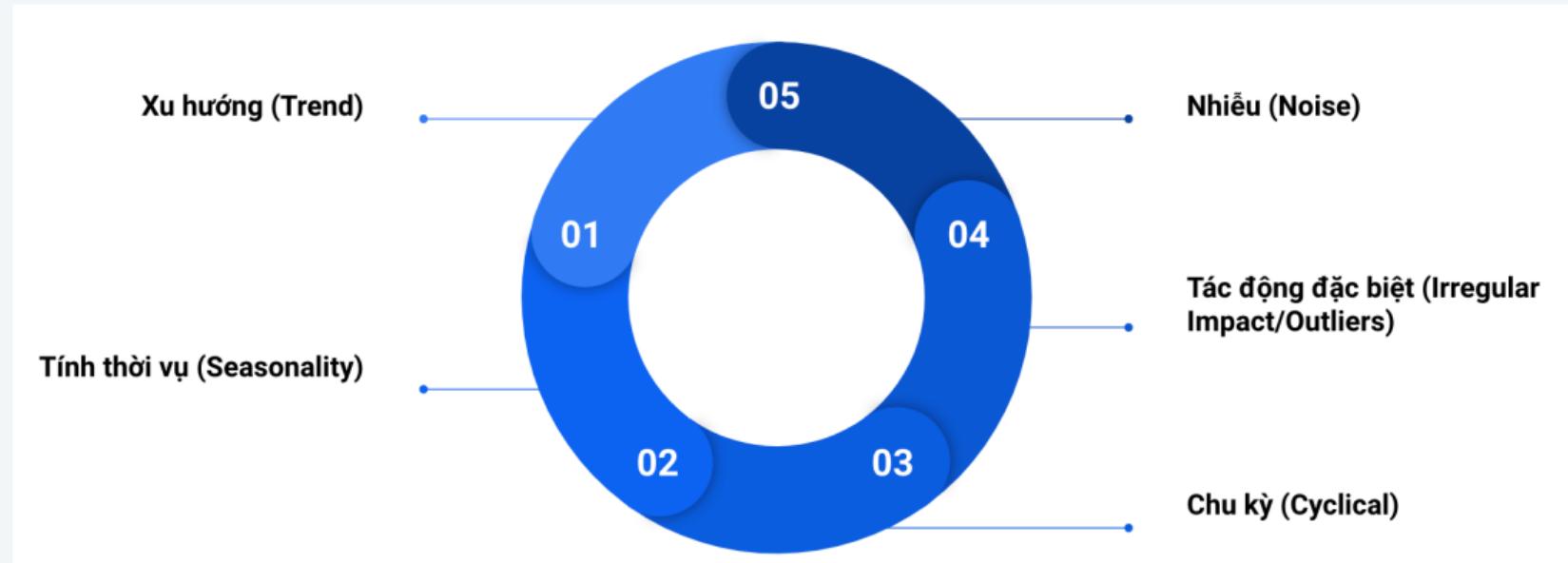
Nhiều đặc điểm khác biệt so với các loại dữ liệu thông thường

History repeats itself!



2. Components of Time Series

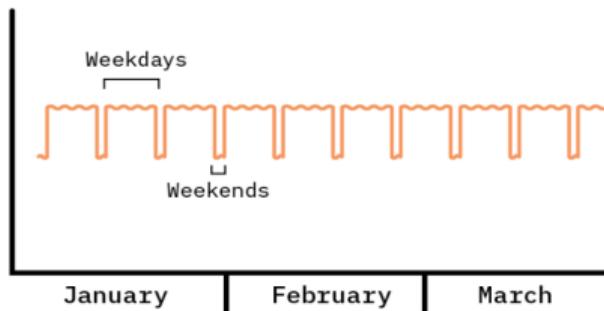
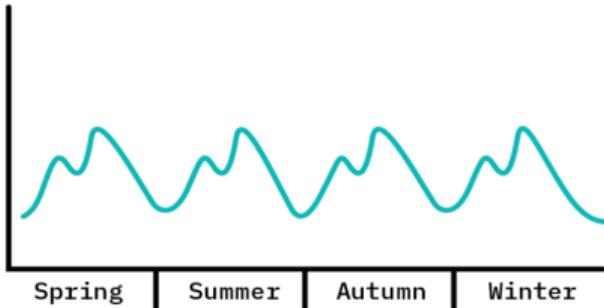
2.1 Components of Time Series



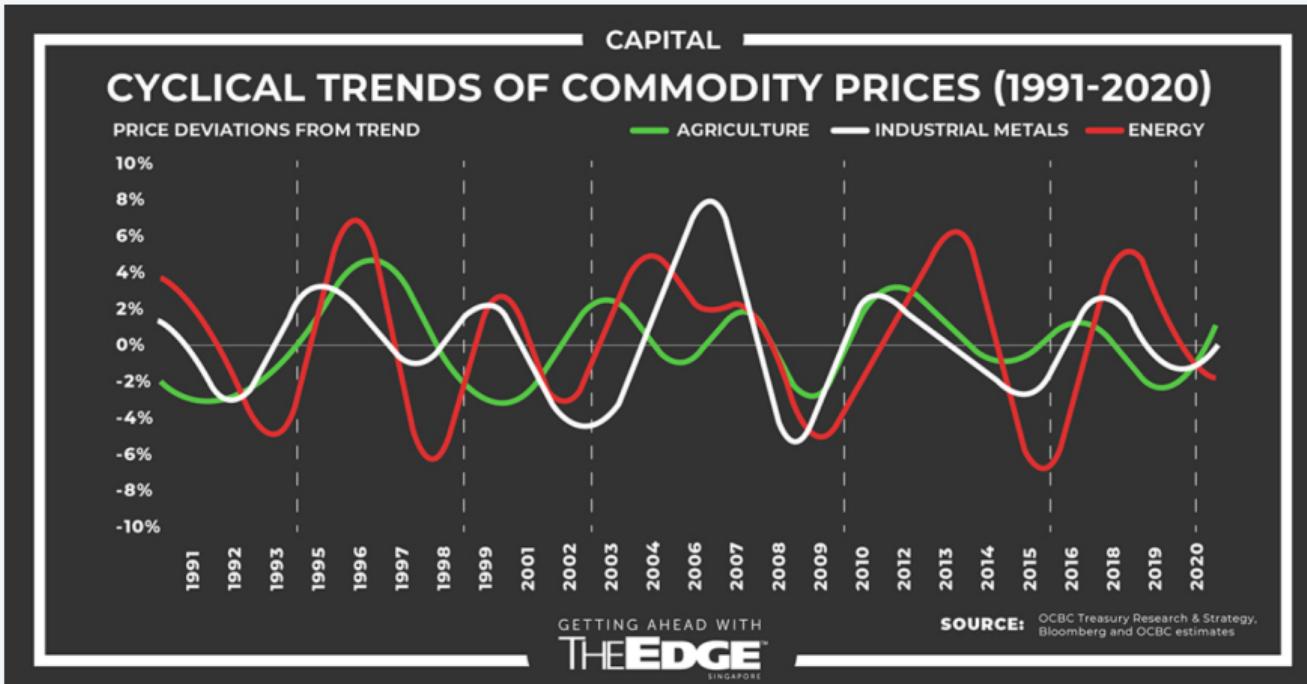
2.2 Components of Time Series - Trend



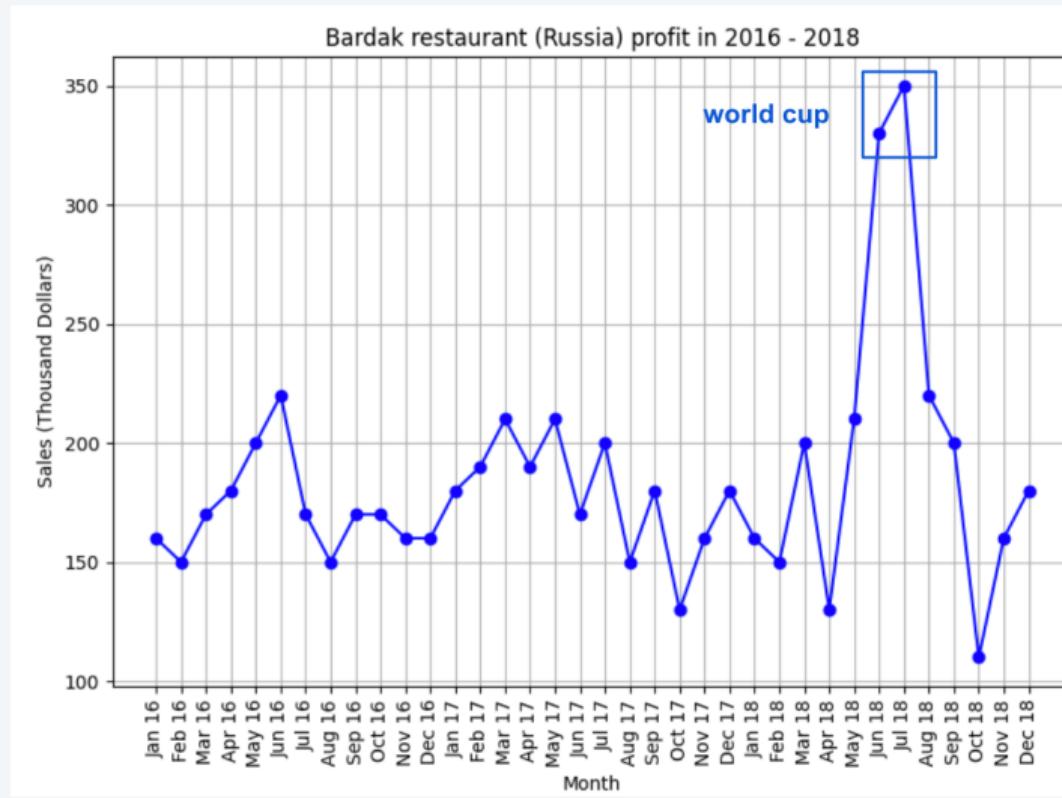
2.3 Components of Time Series – Seasonality



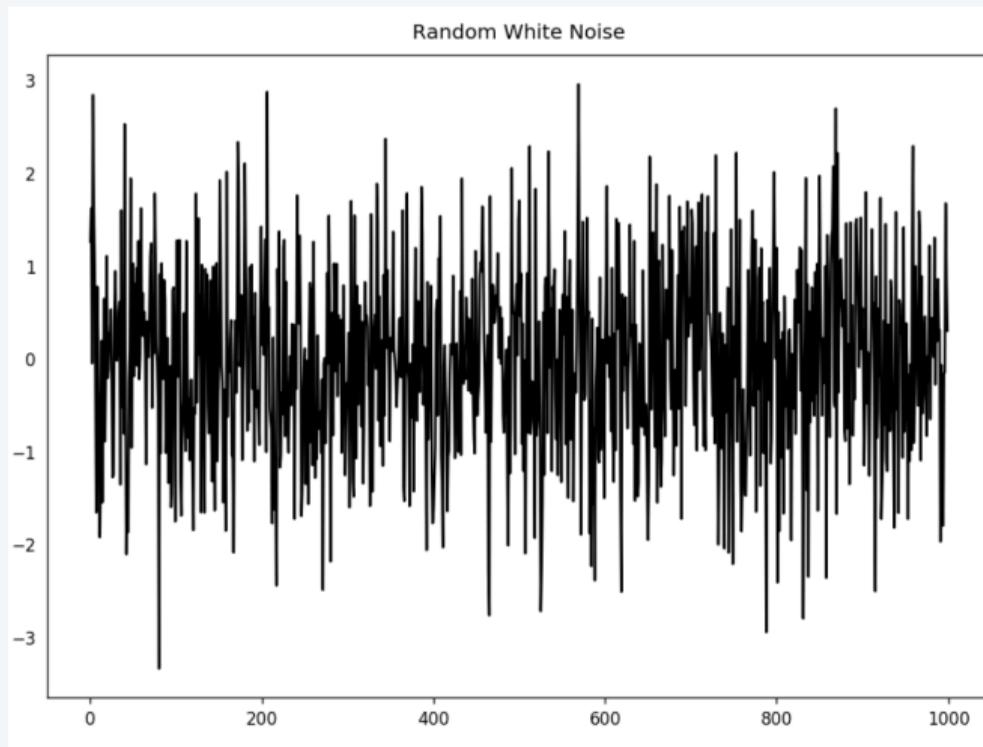
2.4 Components of Time Series - Cyclical



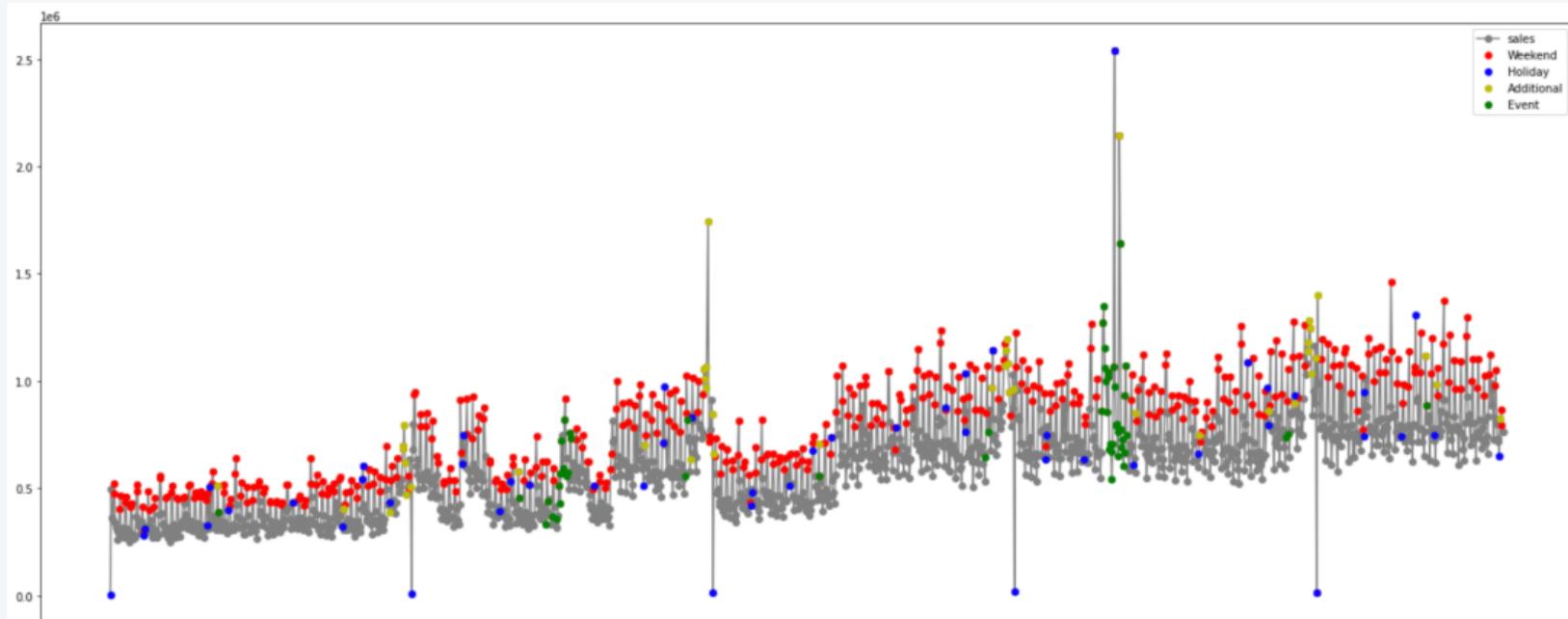
2.5 Components of Time Series - Impact Effect



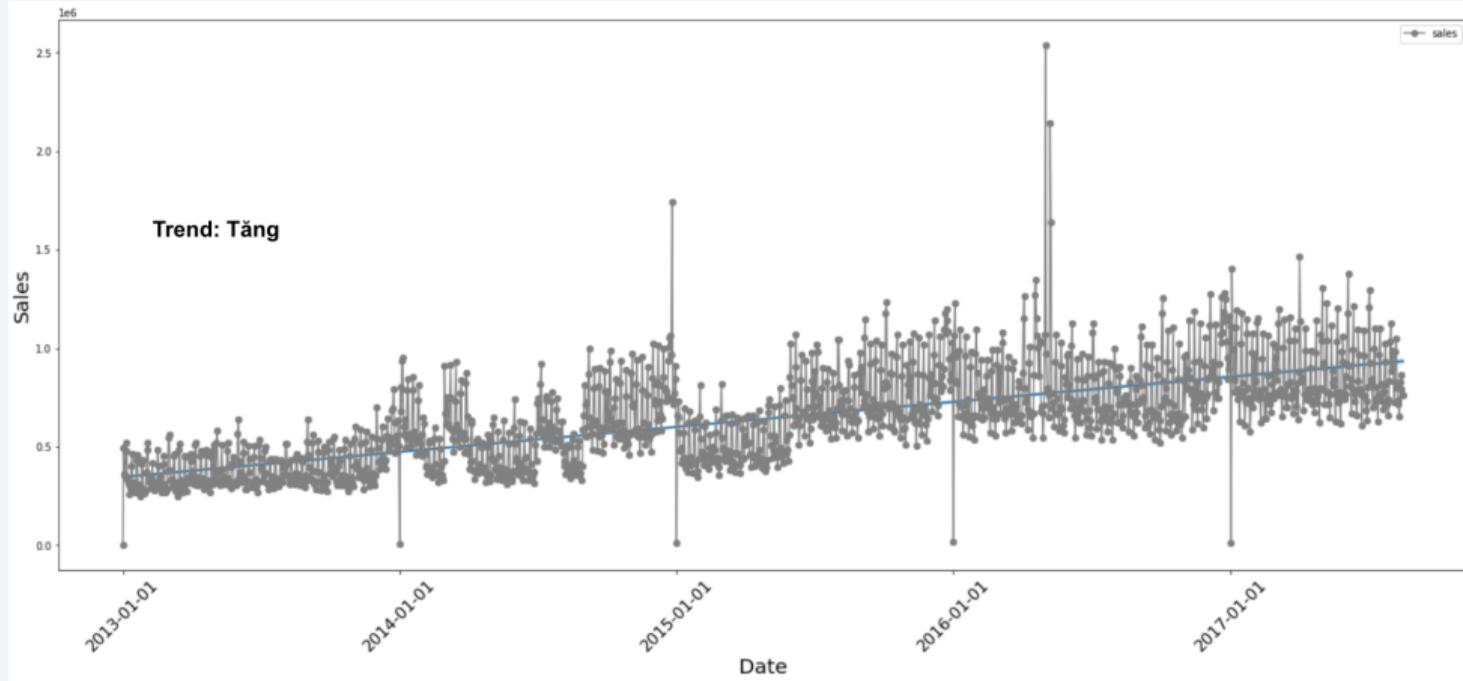
2.6 Components of Time Series – Noise



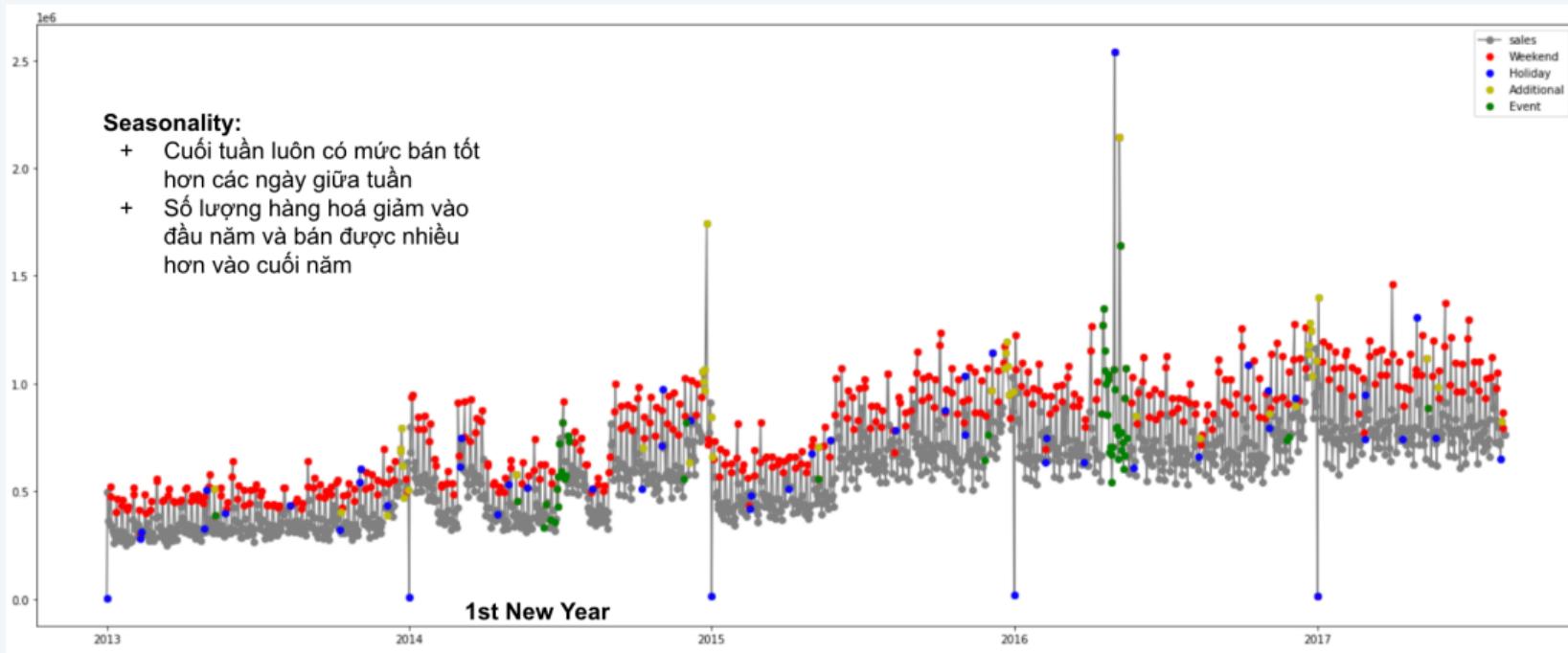
Ex: Sales Analysis Over Time



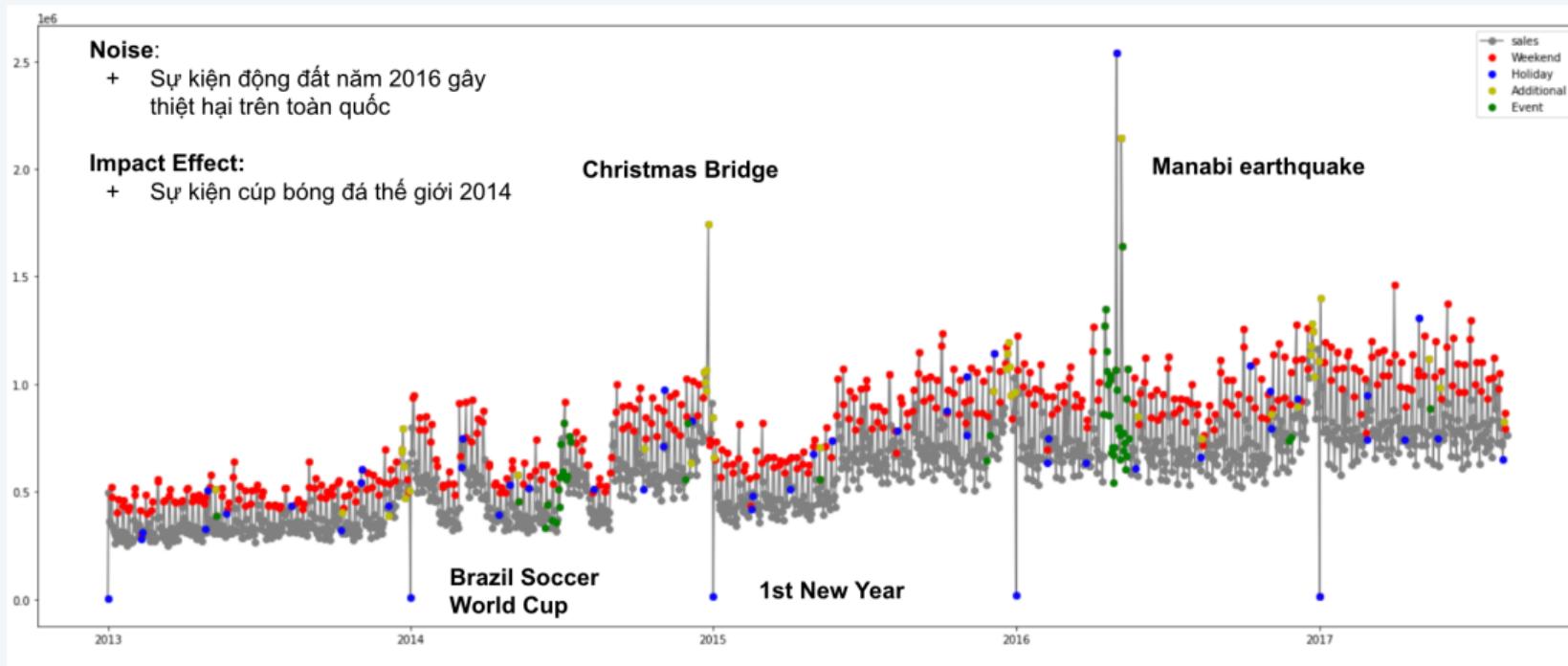
Sales Analysis Over Time



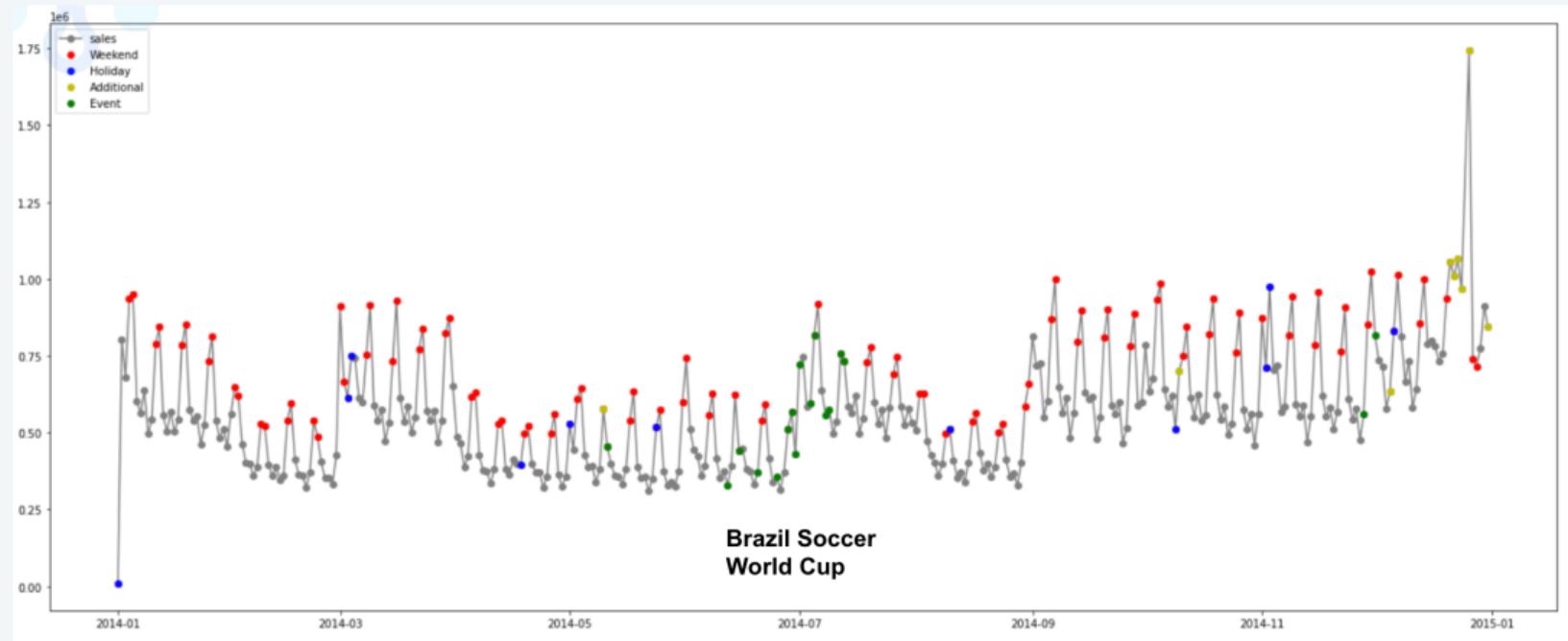
Sales Analysis Over Time



Sales Analysis Over Time



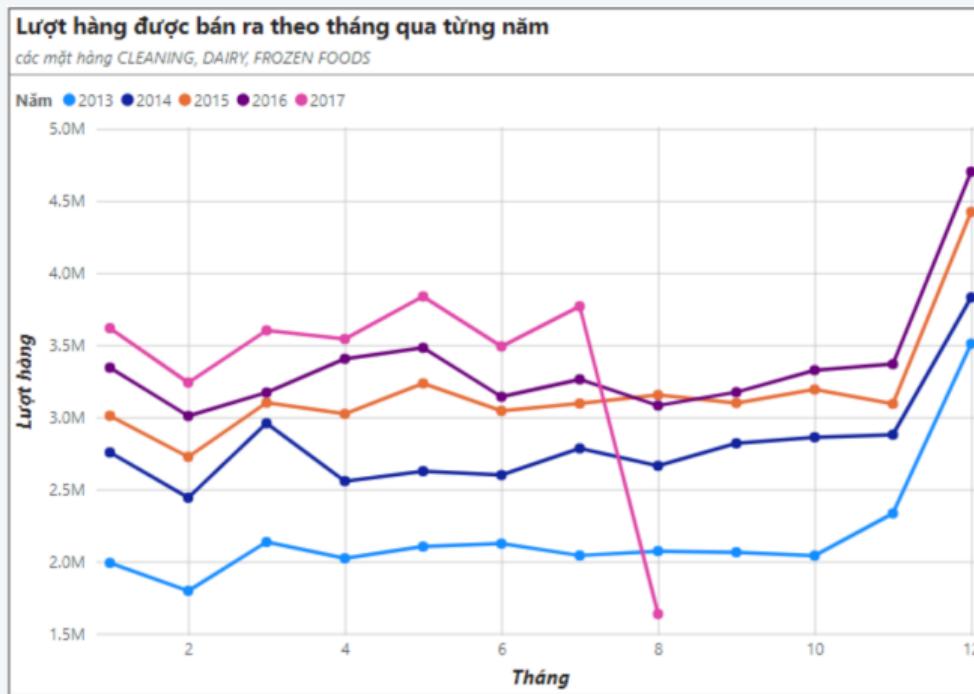
Sales Analysis Over Time



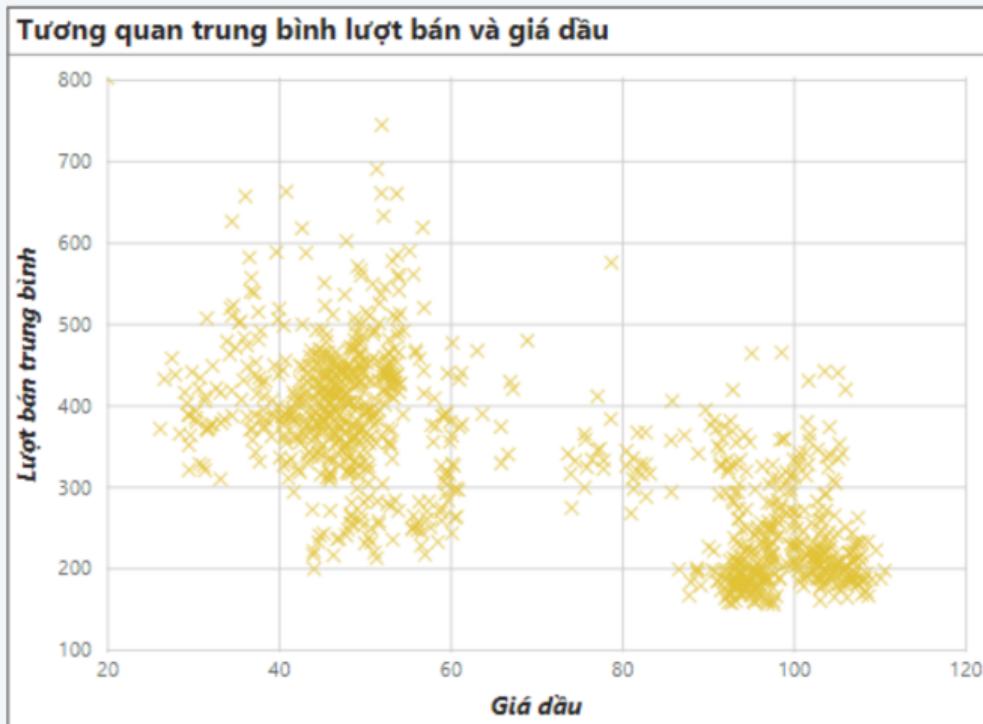


3. Time Series Representations

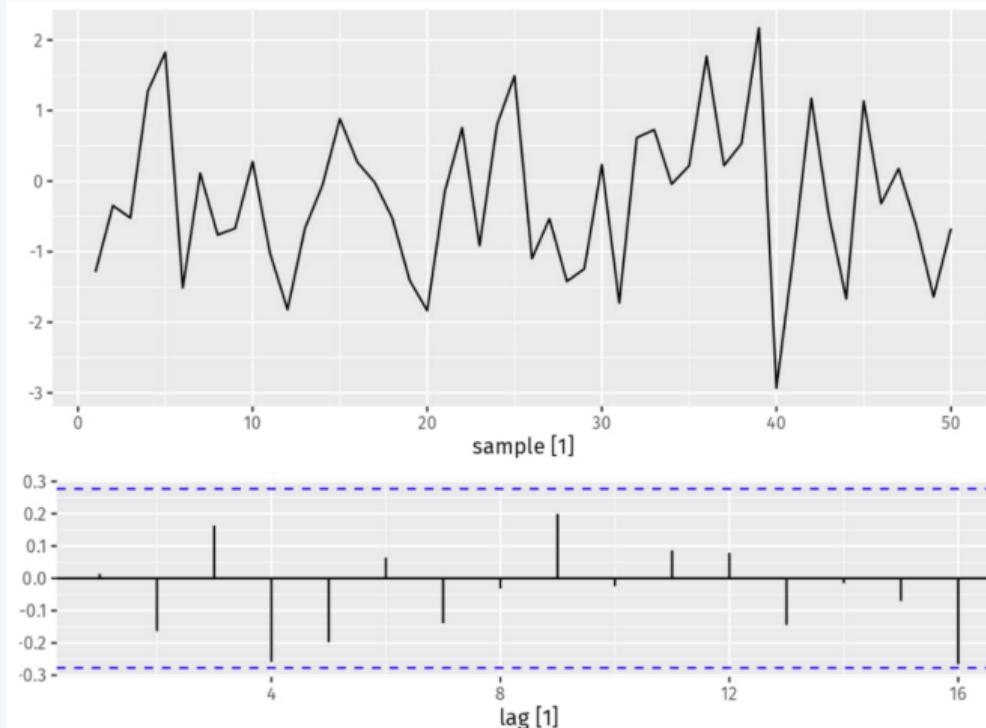
Time Series Representations - Line Graph



Time Series Representations - Scatter Plot



Correlogram (ACF & PACF)





4. Auto-Correlation

4. Auto-Correlation

Components of auto-correlation:

1. Correlation
2. Lagged Series



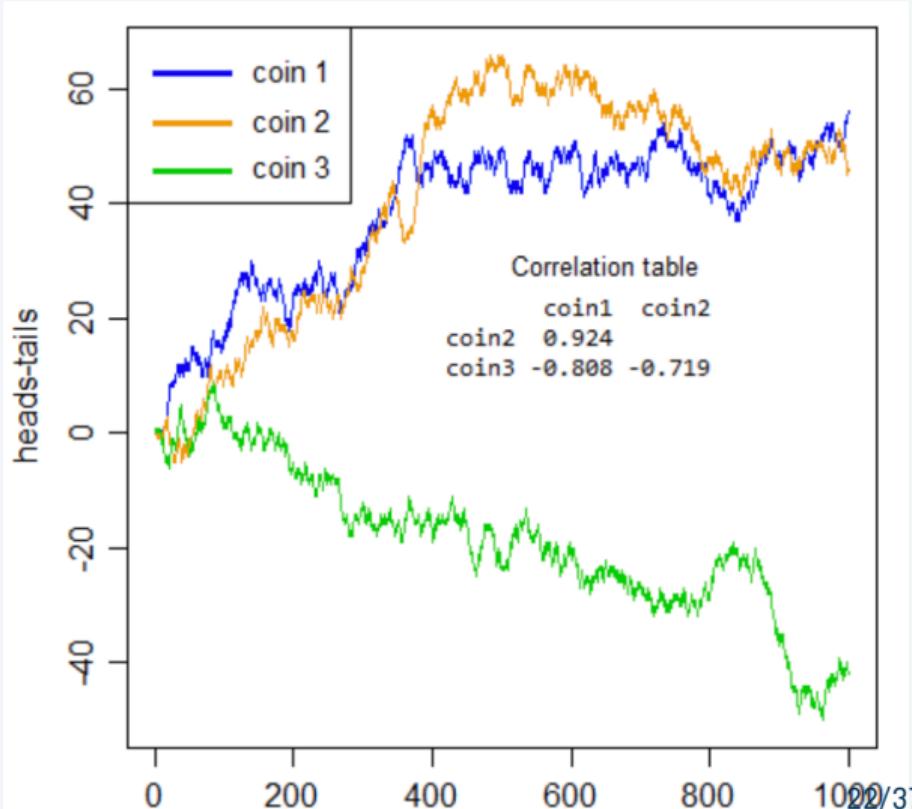
*“Không ai tắm hai lần
trên cùng một dòng sông”*
- Heraclitus



4.1 Correlation

Correlation: a measure of linear relationship between two variables, ranging from -1 to 1:

- **1:** Perfect positive correlation.
- **0:** No linear correlation.
- **-1:** Perfect negative correlation.



Xôi Yến



How can we estimate the amount of sticky rice that can be sold each day in order to prepare the right amount of rice and ingredients?

Based on historical data:

- Quantity of sticky rice sold yesterday, the day before, or over the past 7 days
- Weather conditions, holidays, and special events



4.2 Lagged Series

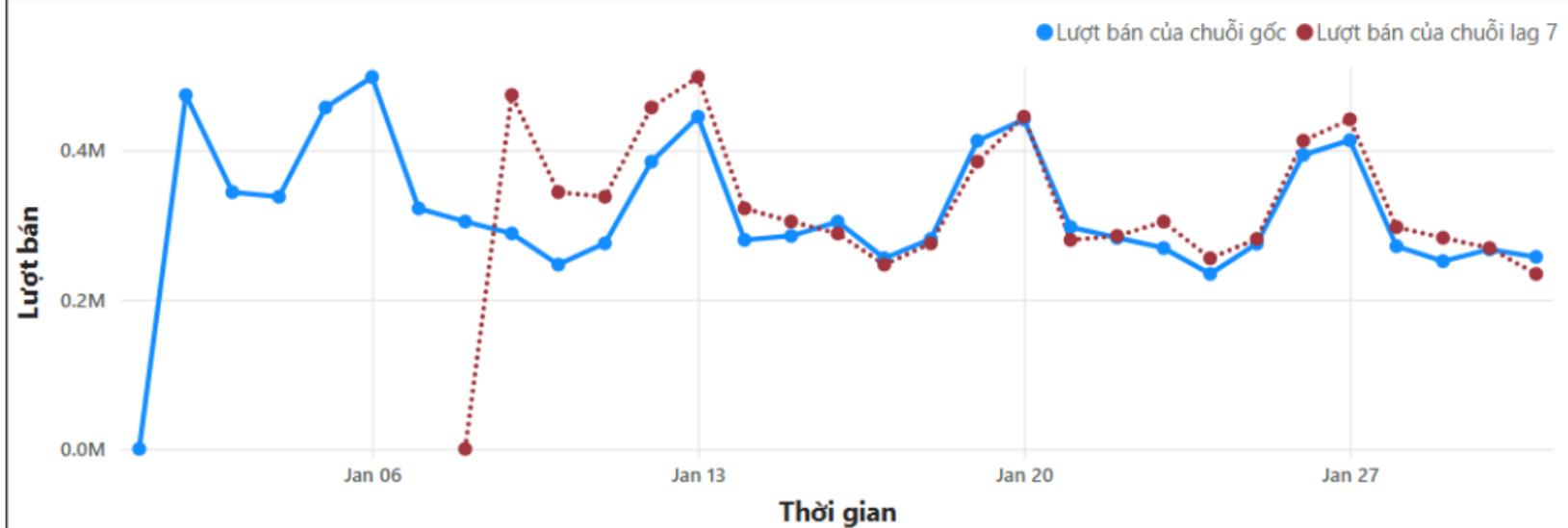
Lagged Series $y(t-k)$: A time series created by shifting the original series $y(t)$ backward by k values. The value k is called the lag.

Date	Value	Value _{t-1}	Value _{t-2}
1/1/2017	200	NA	NA
1/2/2017	220	200	NA
1/3/2017	215	220	200
1/4/2017	230	215	220
1/5/2017	235	230	215
1/6/2017	225	235	230
1/7/2017	220	225	235
1/8/2017	225	220	225
1/9/2017	240	225	220
1/10/2017	245	240	225

4.2 Lagged Series

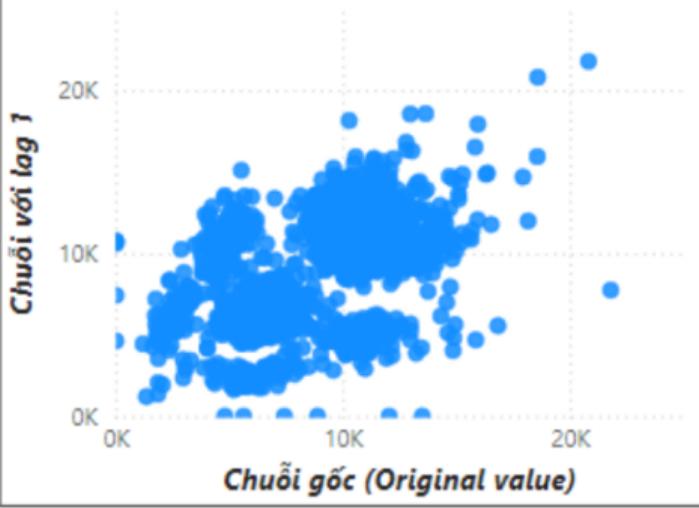
Tổng lượt bán của chuỗi gốc và chuỗi có lag = 7

tại tháng 1/2013

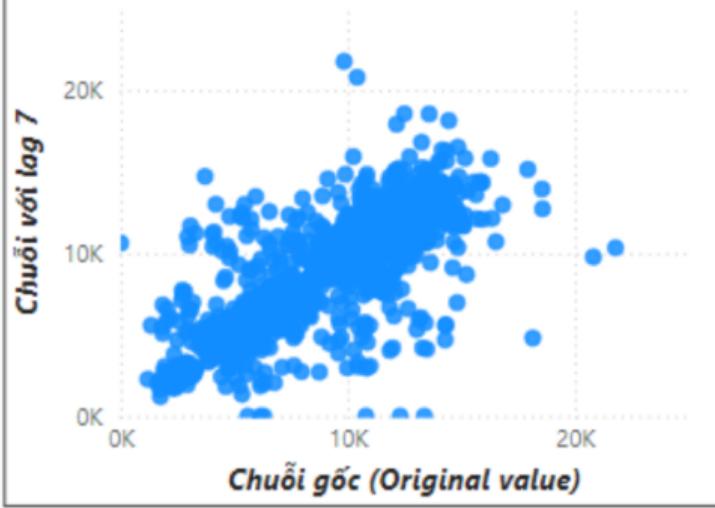


4.2 Lagged Series

Biểu đồ phân tán với lag = 1

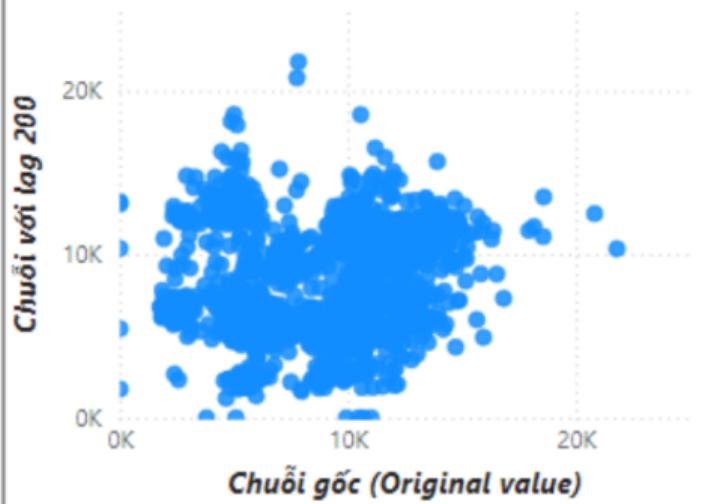


Biểu đồ phân tán với lag = 7

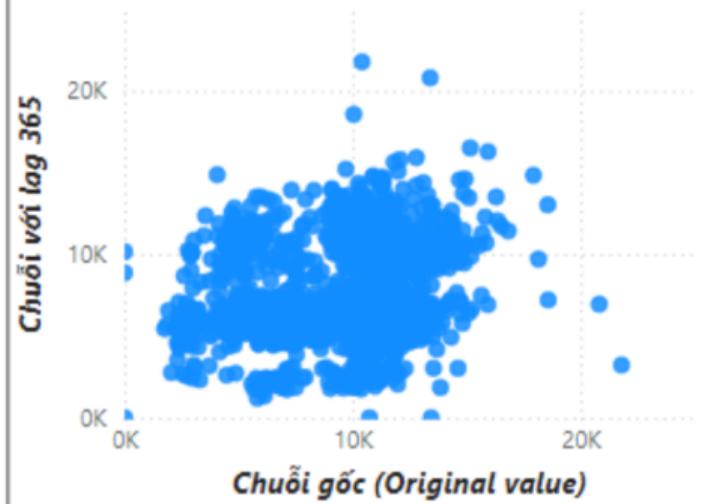


4.2 Lagged Series

Biểu đồ phân tán với lag = 200



Biểu đồ phân tán với lag = 365

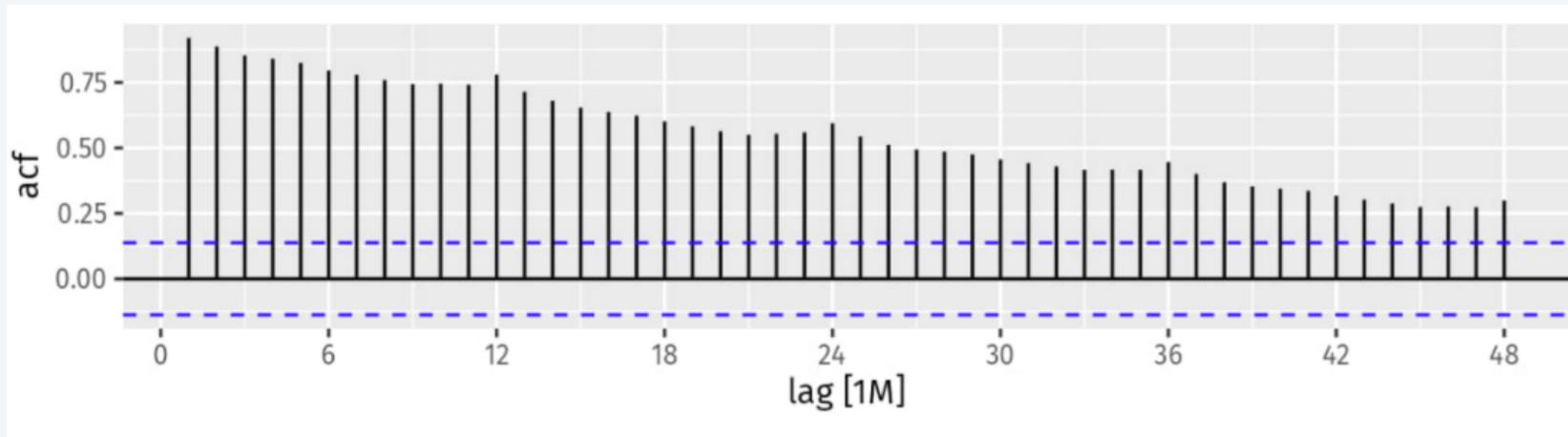


4.3 Auto-Correlation

Auto-Correlation: The correlation between an original time series and its own lagged series.

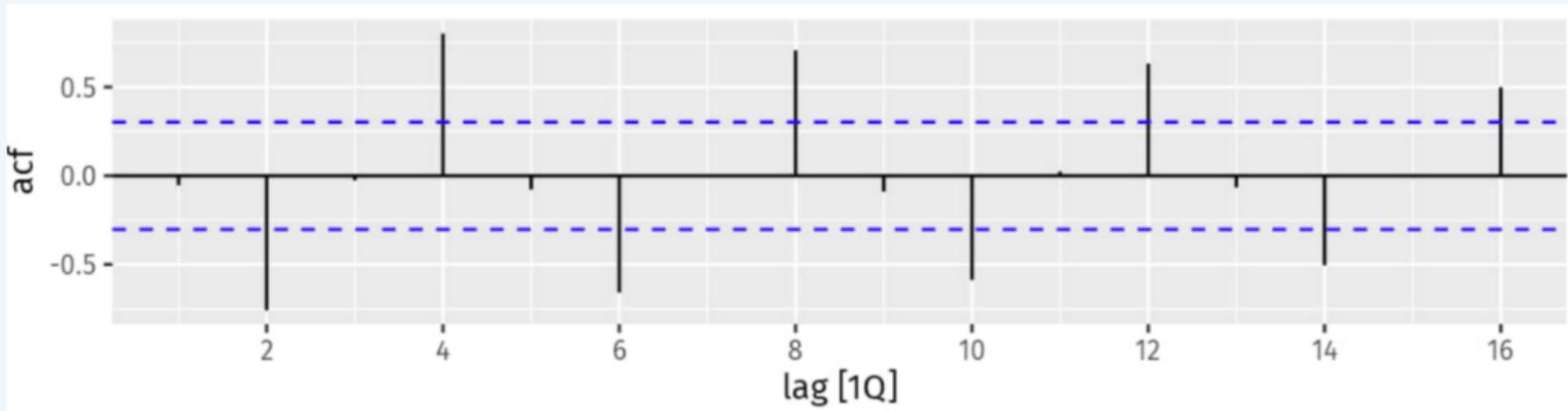
Quarter	Sales (\$000)	Lagged Sales (1)	Lagged Sales (2)	Lagged Sales (3)	Lagged Sales (4)	Lagged Sales (5)
1995Q1	52.04					
1995Q2	59.42	52.04				
1995Q3	55.66	59.42	52.04			
1995Q4	53.86	55.66	59.42	52.04		
1996Q1	61.29	53.86	55.66	59.42	52.04	
1996Q2	75.28	61.29	53.86	55.66	59.42	52.04
1996Q3	73.74	75.28	61.29	53.86	55.66	59.42
1996Q4	68.59	73.74	75.28	61.29	53.86	55.66
1997Q1	71.19	68.59	73.74	75.28	61.29	53.86
1997Q2	85.52	71.19	68.59	73.74	75.28	61.29
1997Q3	76.50	85.52	71.19	68.59	73.74	75.28
1997Q4	73.18	76.50	85.52	71.19	68.59	73.74
1998Q1	72.14	73.18	76.50	85.52	71.19	68.59
1998Q2	84.79	72.14	73.18	76.50	85.52	71.19
1998Q3	77.85	84.79	72.14	73.18	76.50	85.52
1998Q4	75.84	77.85	84.79	72.14	73.18	76.50
1999Q1	88.75	75.84	77.85	84.79	72.14	73.18
1999Q2	98.54	88.75	75.84	77.85	84.79	72.14
1999Q3	94.68	98.54	88.75	75.84	77.85	84.79
1999Q4	90.41	94.68	98.54	88.75	75.84	77.85
2000Q1	87.65	90.41	94.68	98.54	88.75	75.84
2000Q2	98.90	87.65	90.41	94.68	98.54	88.75
2000Q3	95.62	98.90	87.65	90.41	94.68	98.54
2000Q4	91.58	95.62	98.90	87.65	90.41	94.68
Correlation		0.85	0.72	0.75	0.86	0.67

4.3 Auto-Correlation



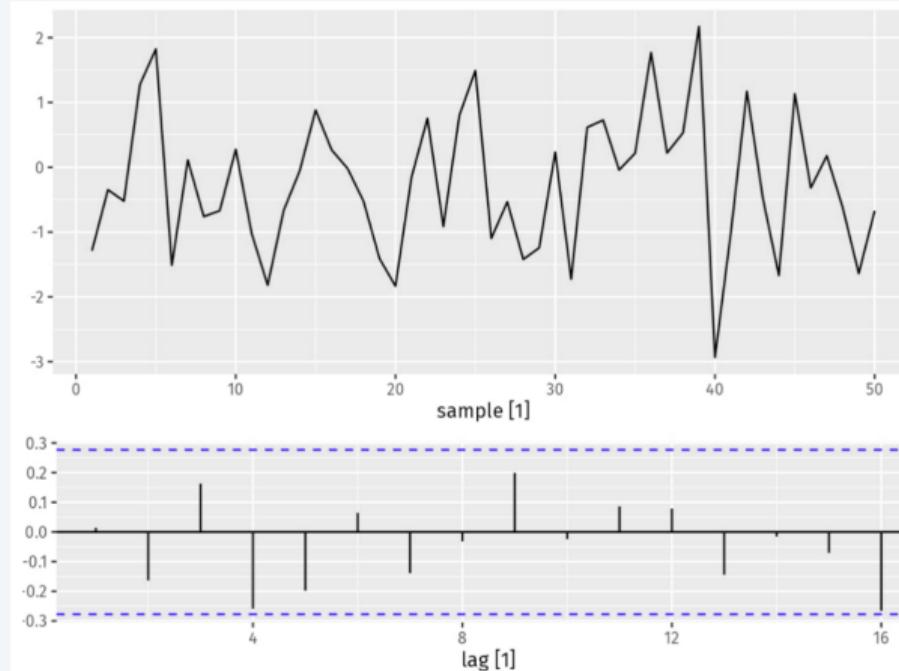
For a time series with a trend, auto-correlation is typically large and positive at small lag values (small k), and gradually decreases as the lag increases.

4.3 Auto-Correlation



For data with seasonality, auto-correlation is usually high at lag values corresponding to the seasonal lag.

4.3 Auto-Correlation

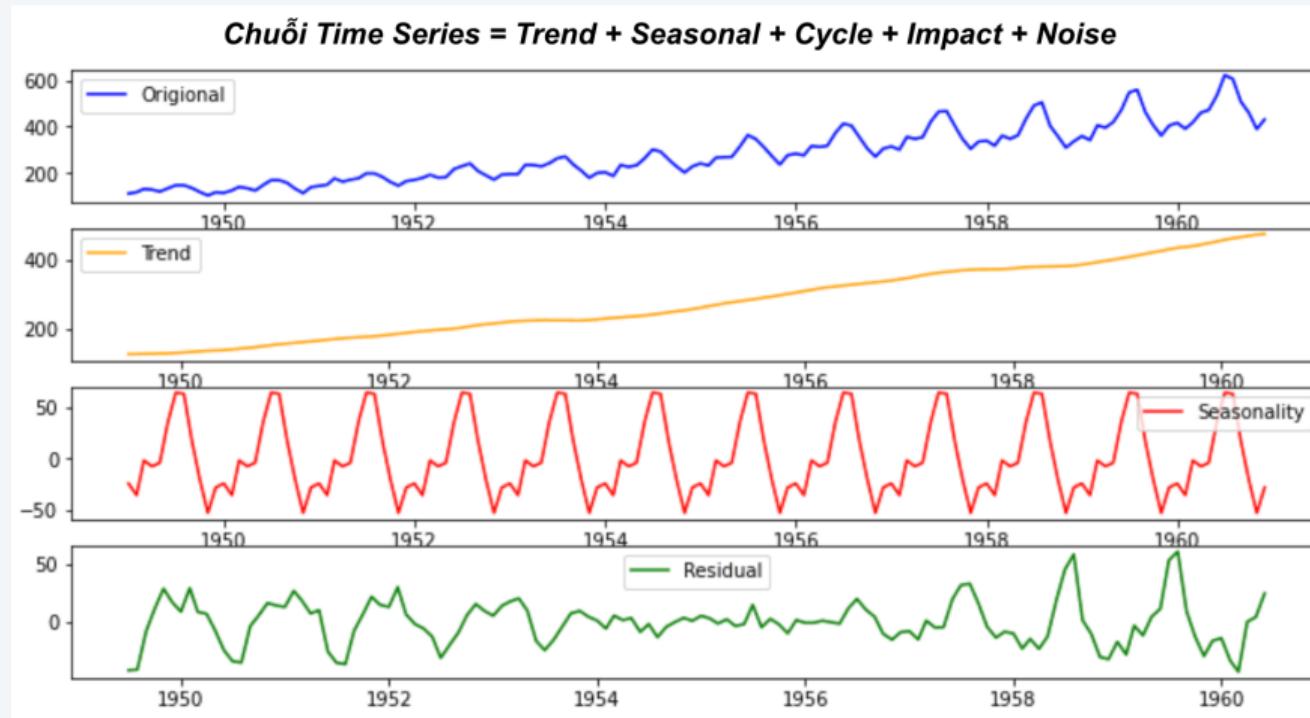


Time series that do not exhibit significant auto-correlation beyond a certain threshold are often considered noise.



5. Advanced Time Series Analysis

5.1 Time Series Decomposition

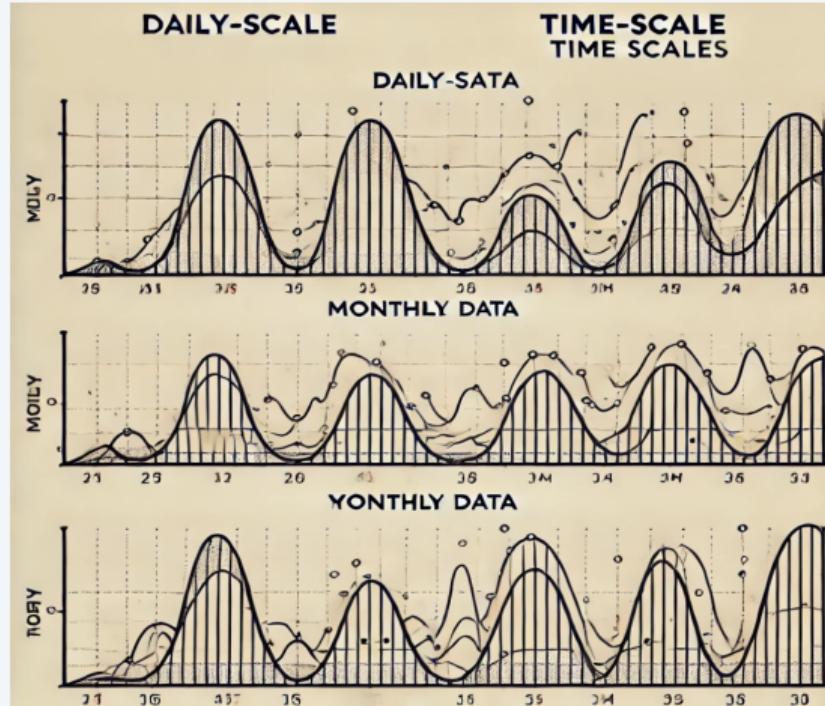


5.2 Time-scale Analysis

Time-scale Analysis:

- **Definition:** Analyzing time series across multiple time scales (short-term, long-term) to understand behaviors and variations at different periods.
- **Significance:** Helps capture data characteristics at different time horizons, enabling the detection of cycles, long-term trends, or short-term fluctuations.

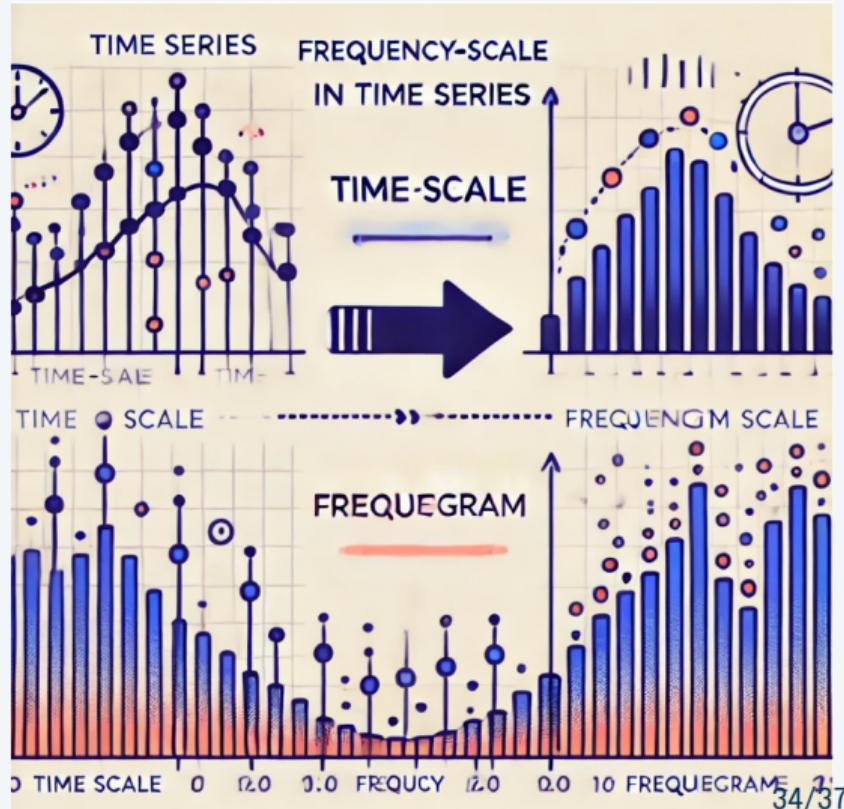
Examples: Monthly revenue data, daily oil prices, annual customer counts



5.3 Frequency-based Analysis

Frequency-based Analysis:

- **Definition:** Analyzing data in the frequency domain to identify repeating cycles and hidden patterns in time series. Data are transformed from the time domain to the frequency domain for easier analysis.
- **Significance:** Enables the detection of periodicities that may not be obvious when directly observing the time series.
- **Examples:** Using **Fourier Transform** or **Wavelet Transform** to identify cycles in weather or financial data.





6. Time Series Forecasting

6. Time Series Forecasting

Suppose we are given the Statistics exam scores of class AI66A.

Consider the following question:

- If a new student joins class AI66A and takes this exam, what score would you expect them to get?

Two fundamental properties of a dataset:

- **Centrality** (mean)
- **Dispersion** (variance)





Time Series Modeling: An Overview

Time series forecasting extends this idea by modeling how these quantities change over time.

In particular, they focus on:

- How the **mean** evolves over time
- How the **variance** (uncertainty) changes
- Recurrent **patterns** such as trend and seasonality
- Temporal dependencies across observations

- **AR / MA / ARIMA / SARIMA**

- Model mean dynamics
- Assume simple variance structure

- **GARCH family**

- Model time-varying variance

- **Deep Learning (RNN, LSTM, Transformer)**

- Learn complex patterns automatically
- Capture nonlinear dynamics



Thank you for your attention!

Ngo Thai Hoa

Faculty of Data Science and Artificial Intelligence
College of Technology, National Economics University