Newton School of Technology

Lecture 1

Unit 1
Time Series

By Sanchit, Rishabh and Shobhit

Newton School of Technology

Join the lecture online on your dashboard.

Newton School of Technology

Let’s start with a minute of silence.

आचार्यात् पादं आधत्ते पादं शिष्यः स्वमेधया ।
पादं सब्रह्मचारिभ्यः पादं कालक्रमेण च ॥

Meaning:

A student acquires knowledge in four equal parts:

- One-fourth from the teacher
- One-fourth through self-reflection and independent thinking
- One-fourth through discussions with peers and fellow learners
- One-fourth over time through personal experience

Recap

Newton School of Technology

- Why Advance ML?
- What is ML?
- Course Objectives
- Roadmap
- Evaluation and Classroom Rules
- Required Readings and References

Agenda

Newton School of Technology

- What is Time Series?
- Components of Time Series Data
- Decomposition of Time Series Data

City- Energy Demand Forecasting

Newton School of Technology

- **Problem Statement**
  - A **power distribution company** must forecast electricity demand at multiple horizons:
    - 15-minute ahead (real-time load balancing)
    - Day-ahead (generation planning)
    - Week-ahead (maintenance scheduling)
  - **Wrong forecasts** lead to **blackouts** or **wasted** generation cost.
- **Dataset Link:** https://data.open-power-system-data.org/time_series/

City- Energy Demand Forecasting

Newton School of Technology

Assumed Data & Features

Assume we have 15-minute interval data with:

- Target load → electricity demand (MW)

Time-based features

- hour
- day_of_week
- is_weekend
- month

Historical features

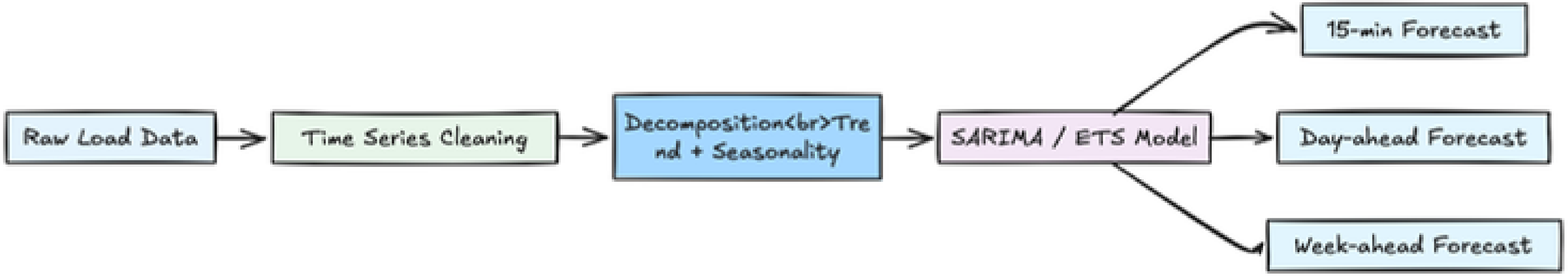
load_t-1 (last 15 min)
load_t-4 (1 hour ago)
load_t-96 (same time yesterday)
rolling_mean_1h
rolling_mean_24h

External features (optional but realistic)

temperature
humidity
holiday_flag

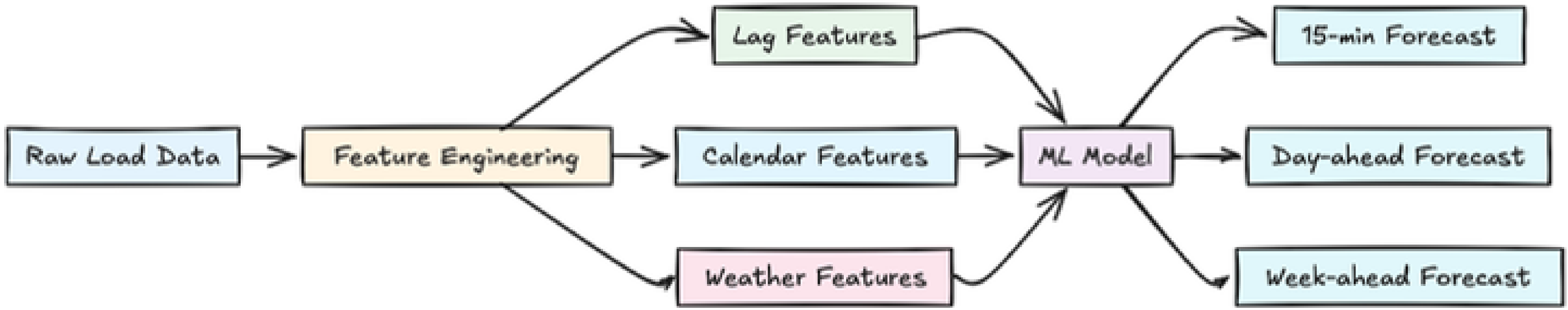
City- Energy Demand Forecasting

Newton School of Technology



City- Energy Demand Forecasting

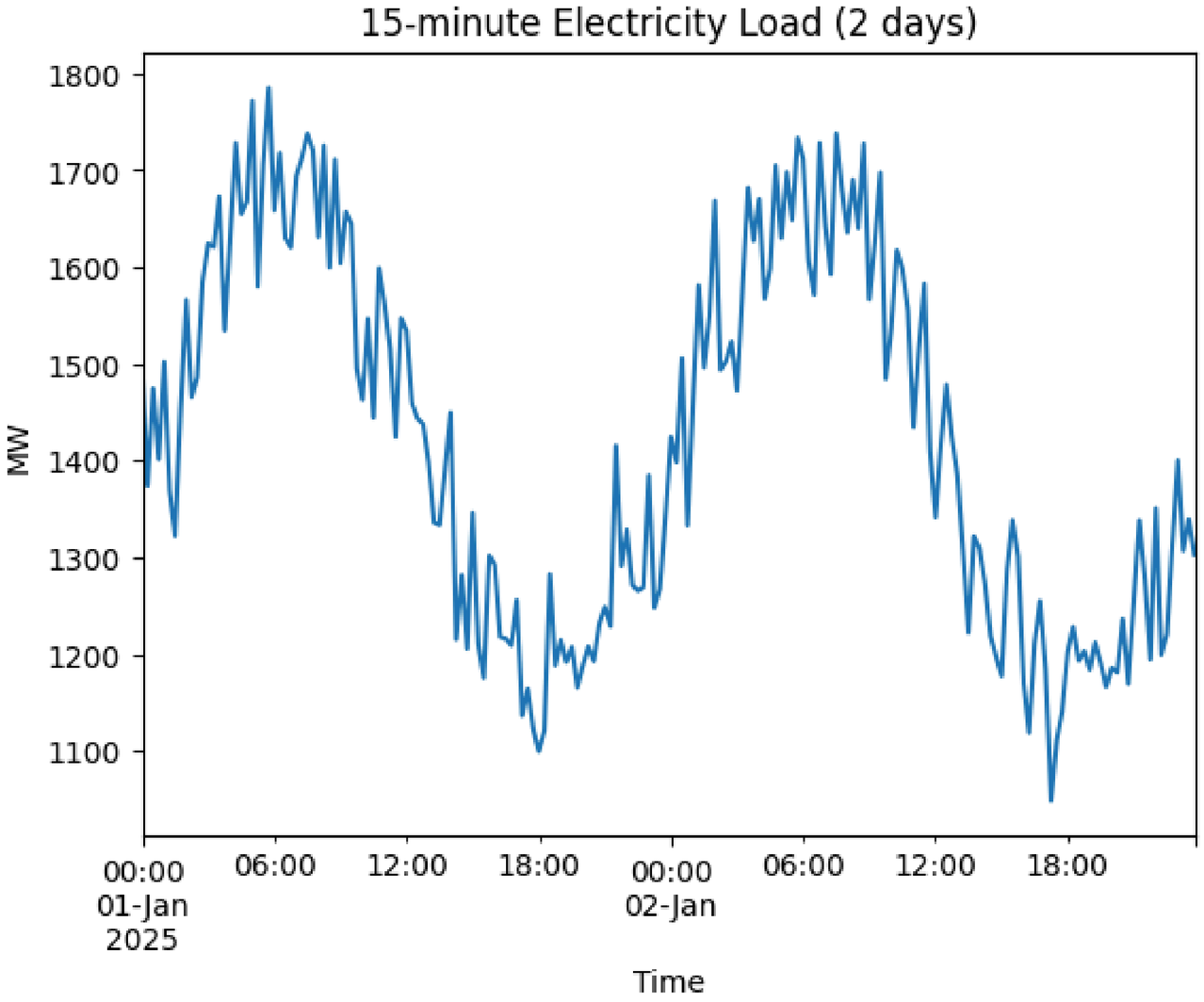
Newton School of Technology



City- Energy Demand Forecasting

Newton School of Technology

15-minute Electricity Load for 2 days

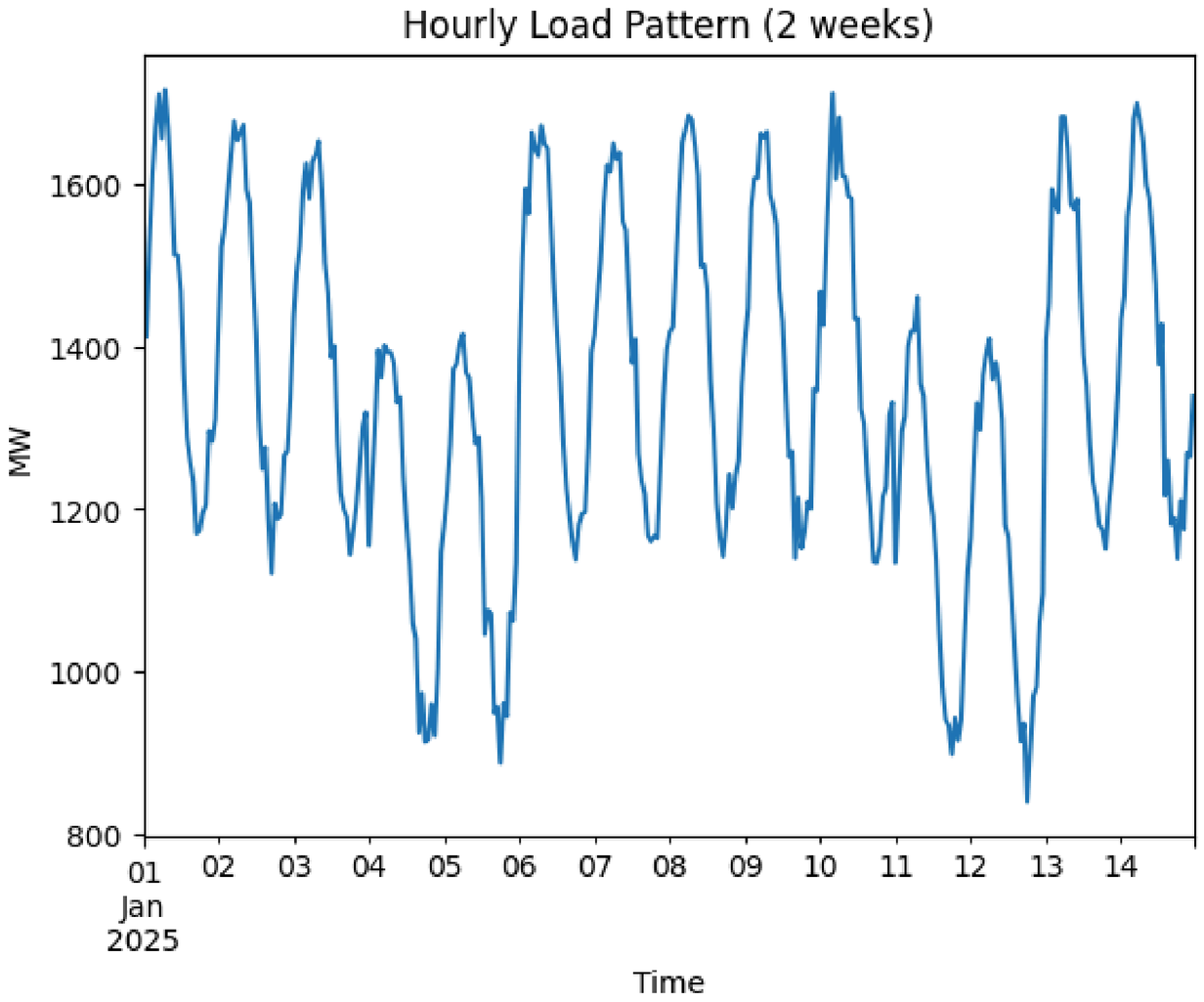


What do you notice as you look from left to right in this plot?

City- Energy Demand Forecasting

Newton School of Technology

Hourly load for 2 weeks

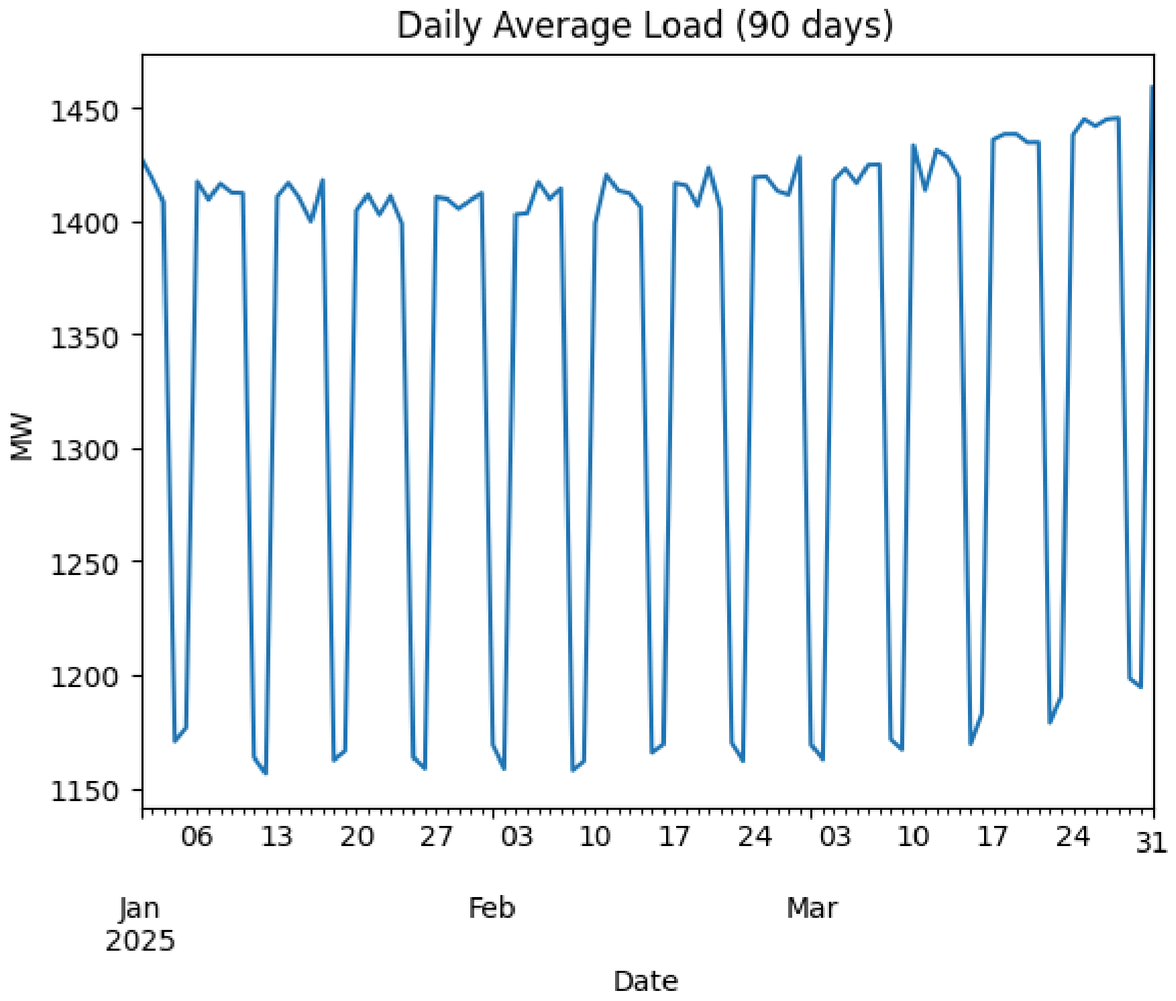


What do you notice as you look from left to right in this plot?

City- Energy Demand Forecasting

Newton School of Technology

Daily average for 90 days



What do you notice as you look from left to right in this plot?

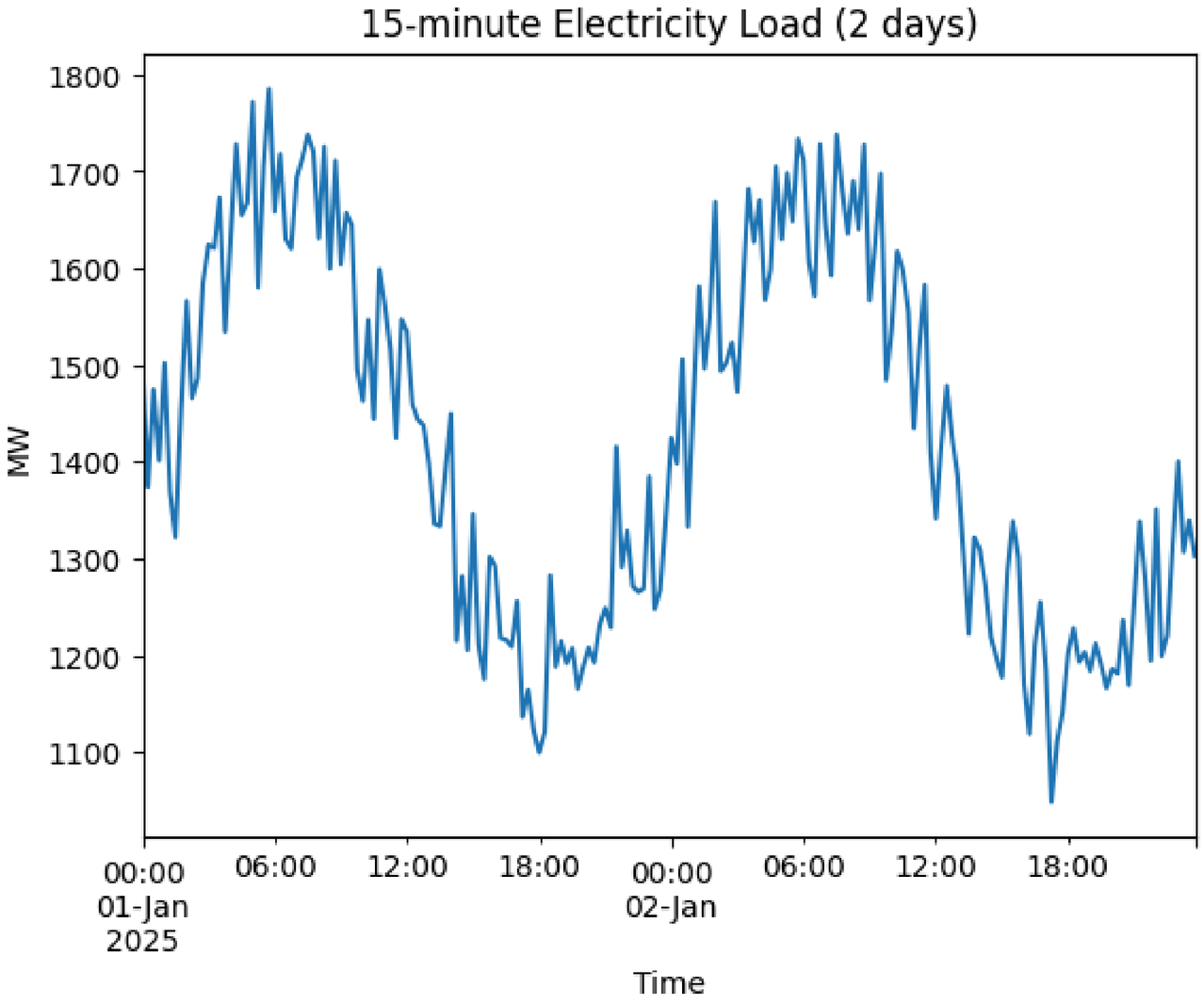
Newton School of Technology

What is the relation between time and electicity load/consumption?

What is the relation between time and electicity load/consumption?

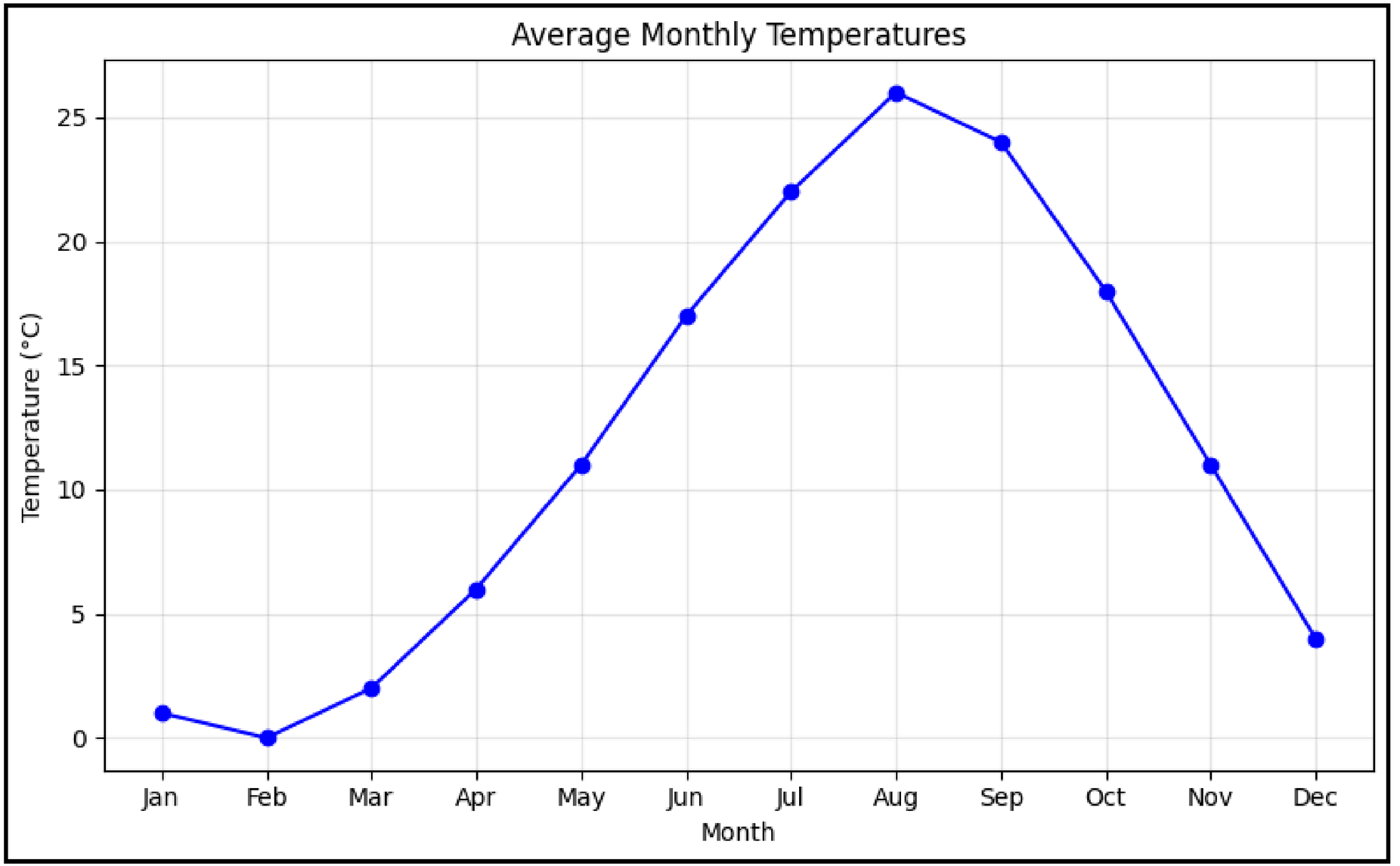
Newton School of Technology

- As time increases, there is a pattern in the consumption.
- The consumption is highest at 6:00 and lowest at 18:00.
- 1 Jan 2025 consumption is very similar to 2nd jan 2025



What about this graph between time and temperature?

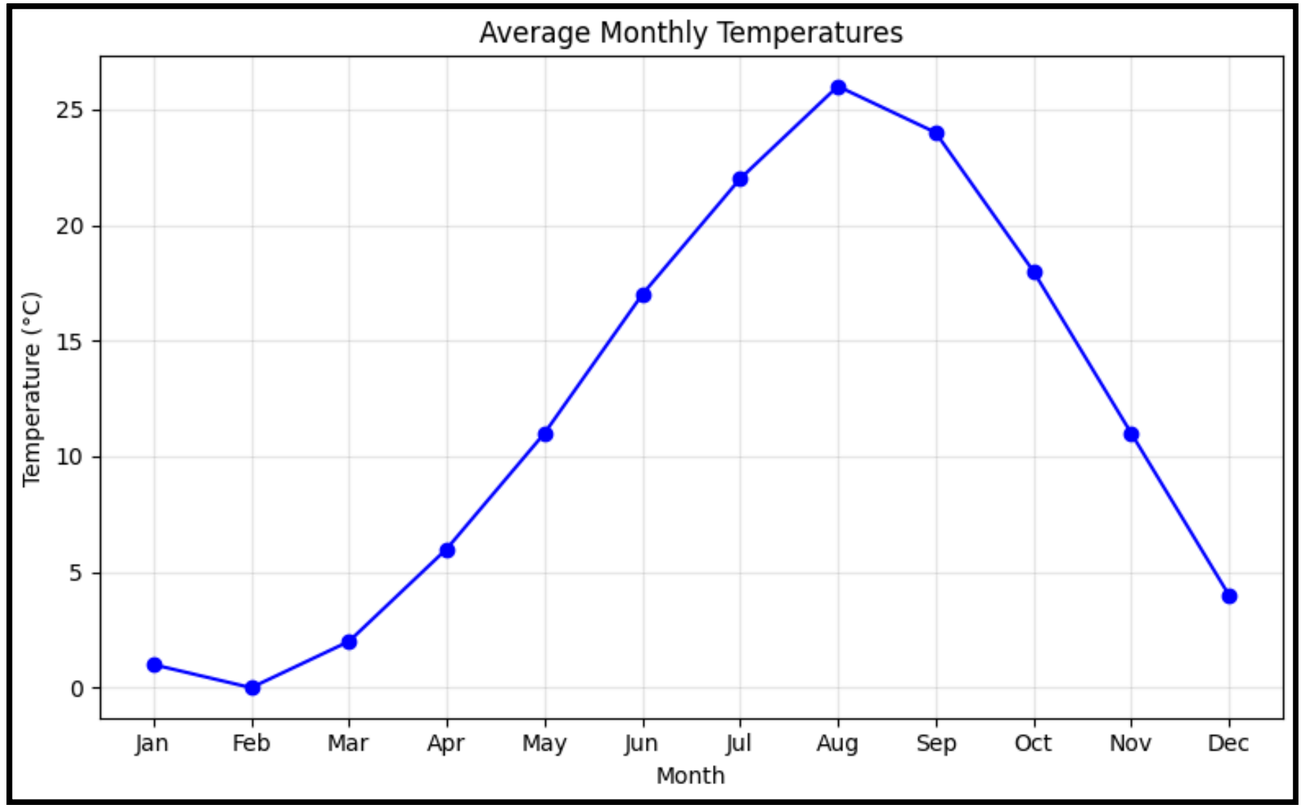
Newton School of Technology



What about this graph between time and temperature?

Newton School of Technology

- As time increases, there is a pattern in the temperature.
- The Temperature is highest in July-Aug and lowest in Feb.
- Jan Temp is very similar to that of Feb.

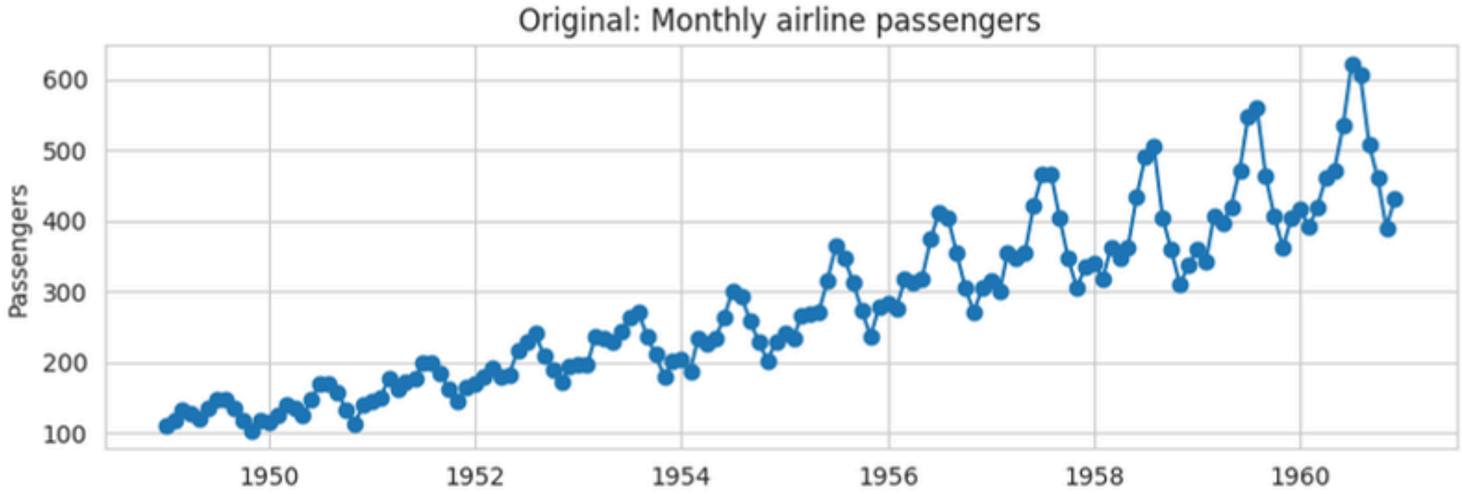


Newton School of Technology

What is Time series?

Time Series Data

Newton School of Technology

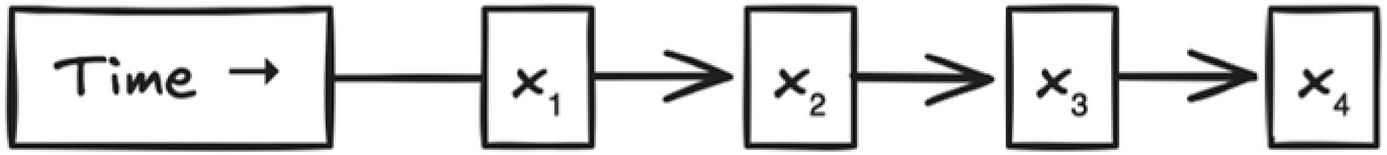


- Time series data is a **sequence of data points** collected or recorded at **specific time intervals.**

- Unlike **standard data**, where the **order of observations** may not matter, in time series, the **temporal ordering is critical.**

Time Series Data

Newton School of Technology

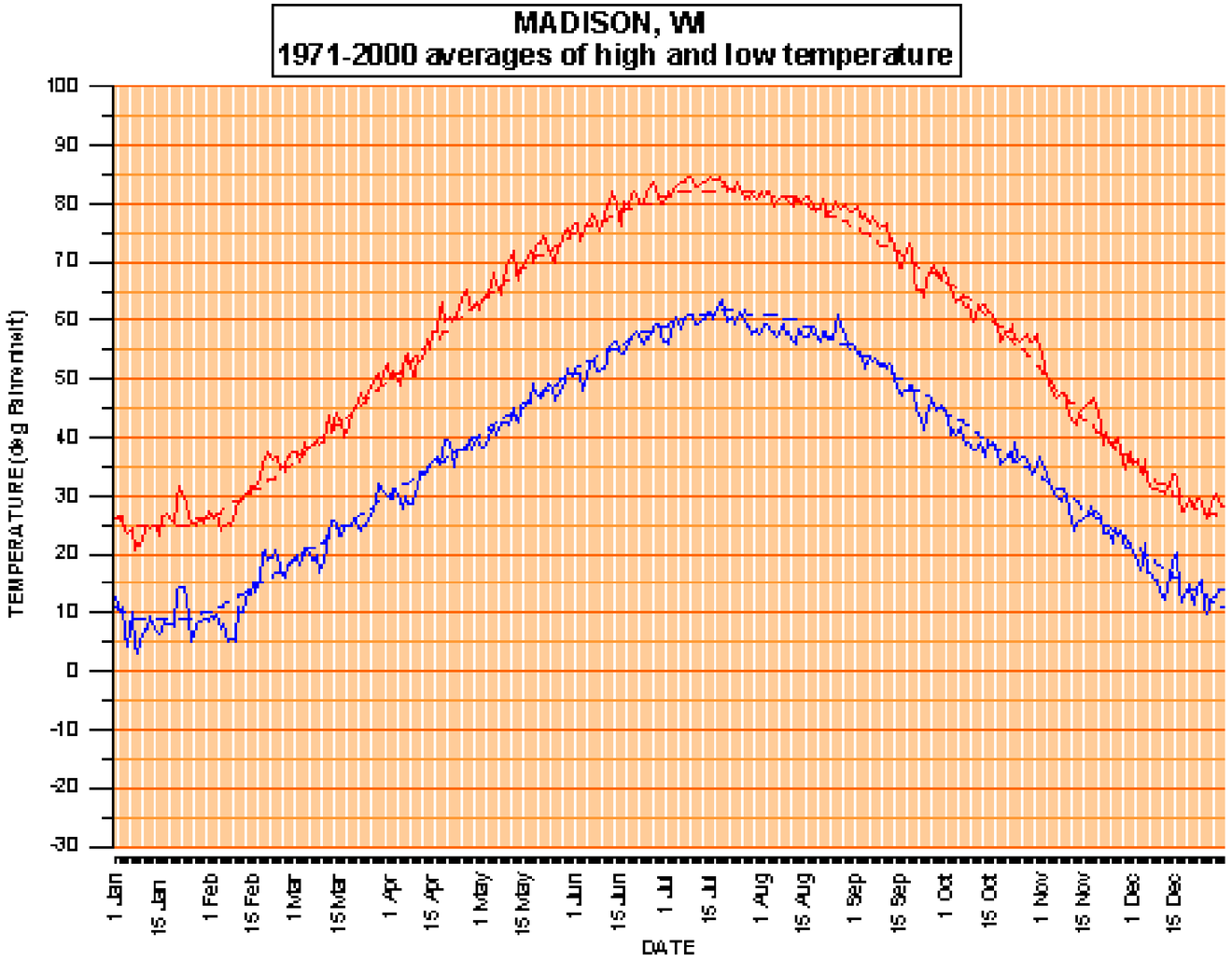


Values are **indexed** by **time**, and the **order** cannot be **changed**.

Examples of Timeseries

Newton School of Technology

Weather Data:
Daily Temperature



Examples of Timeseries

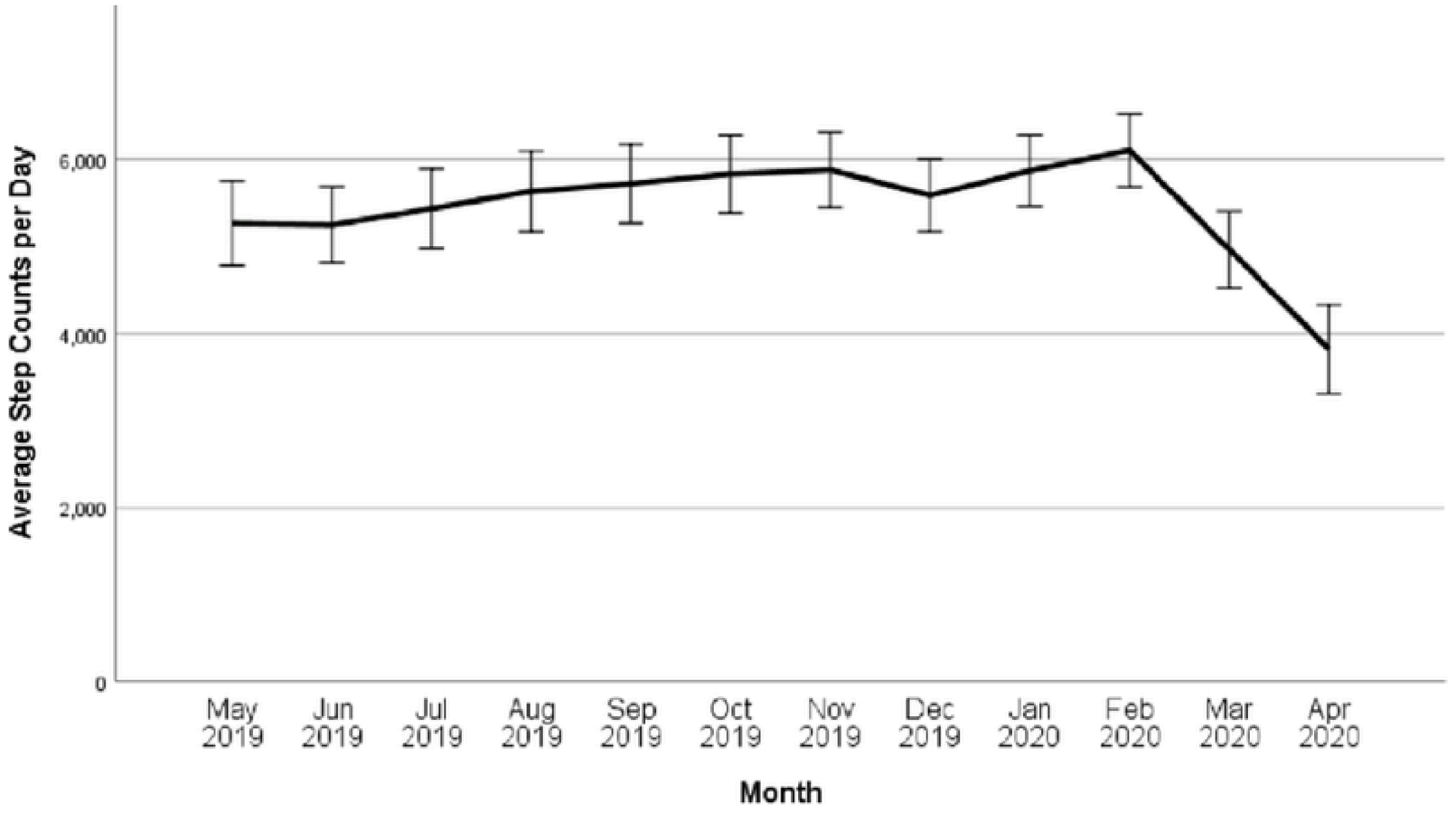
Newton School of Technology

Sales/Revenue Data: Monthly Sales/Revenue for Products

Examples of Timeseries

Newton School of Technology

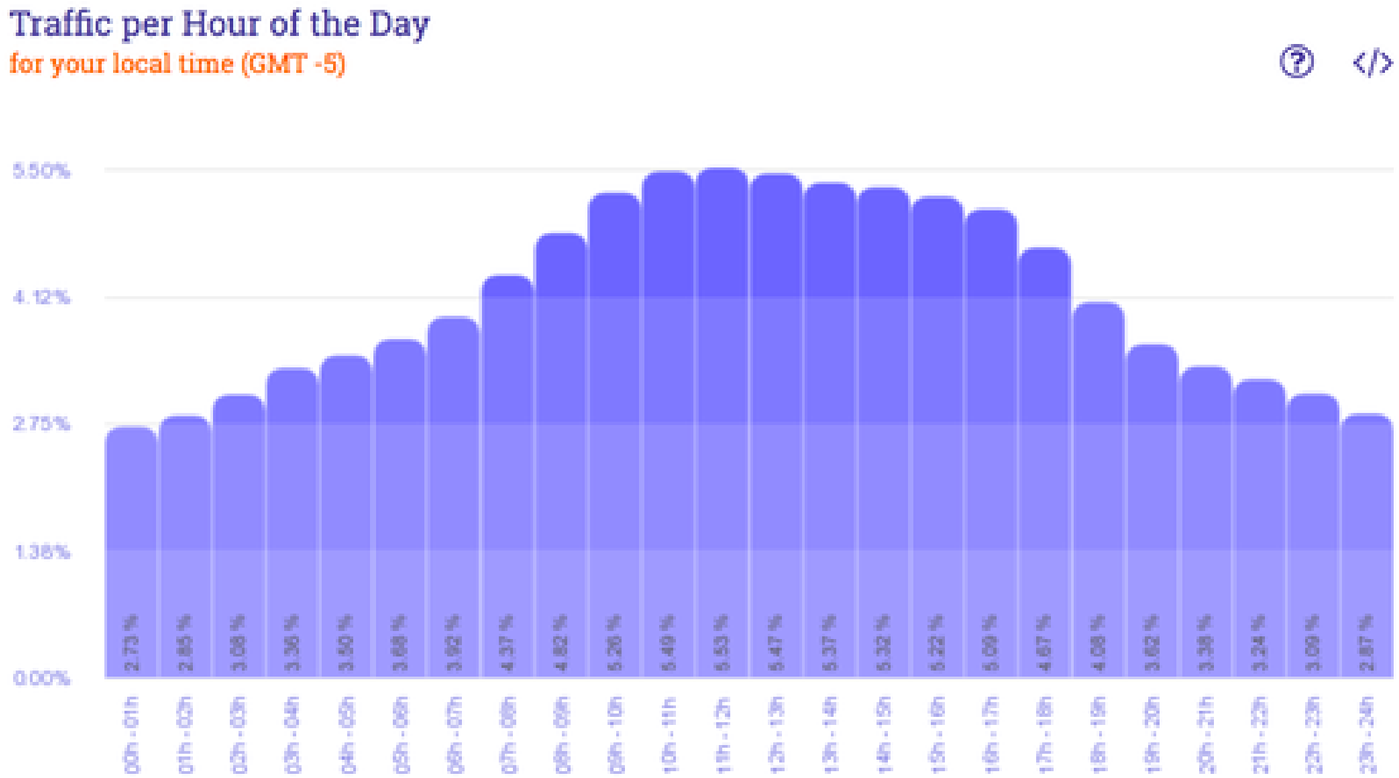
Fitbit Data:
Steps Count per Day



Examples of Timeseries

Newton School of Technology

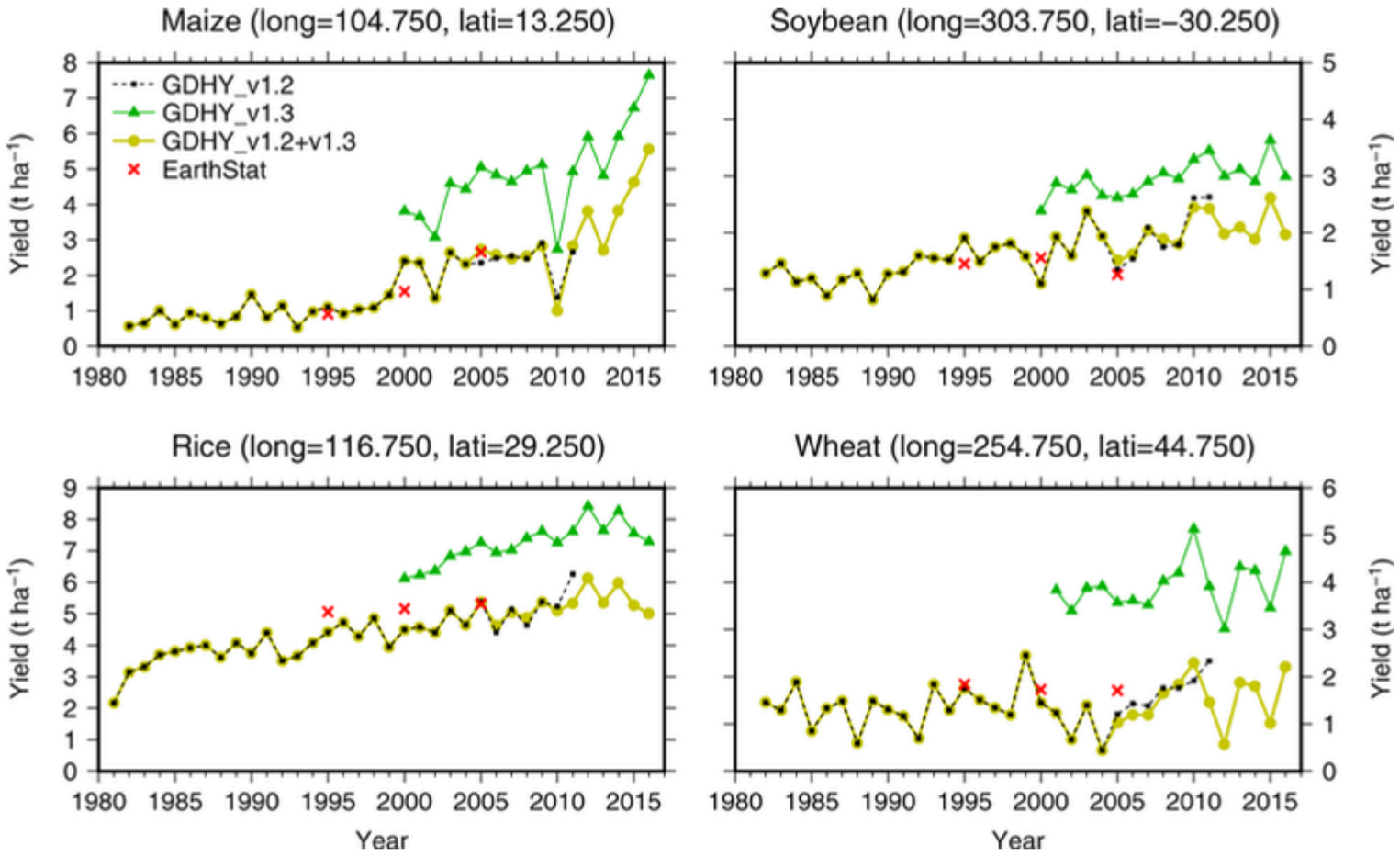
Website Traffic:
Page views per hour/day



Examples of Timeseries

Newton School of Technology

Crop Yield Over Years



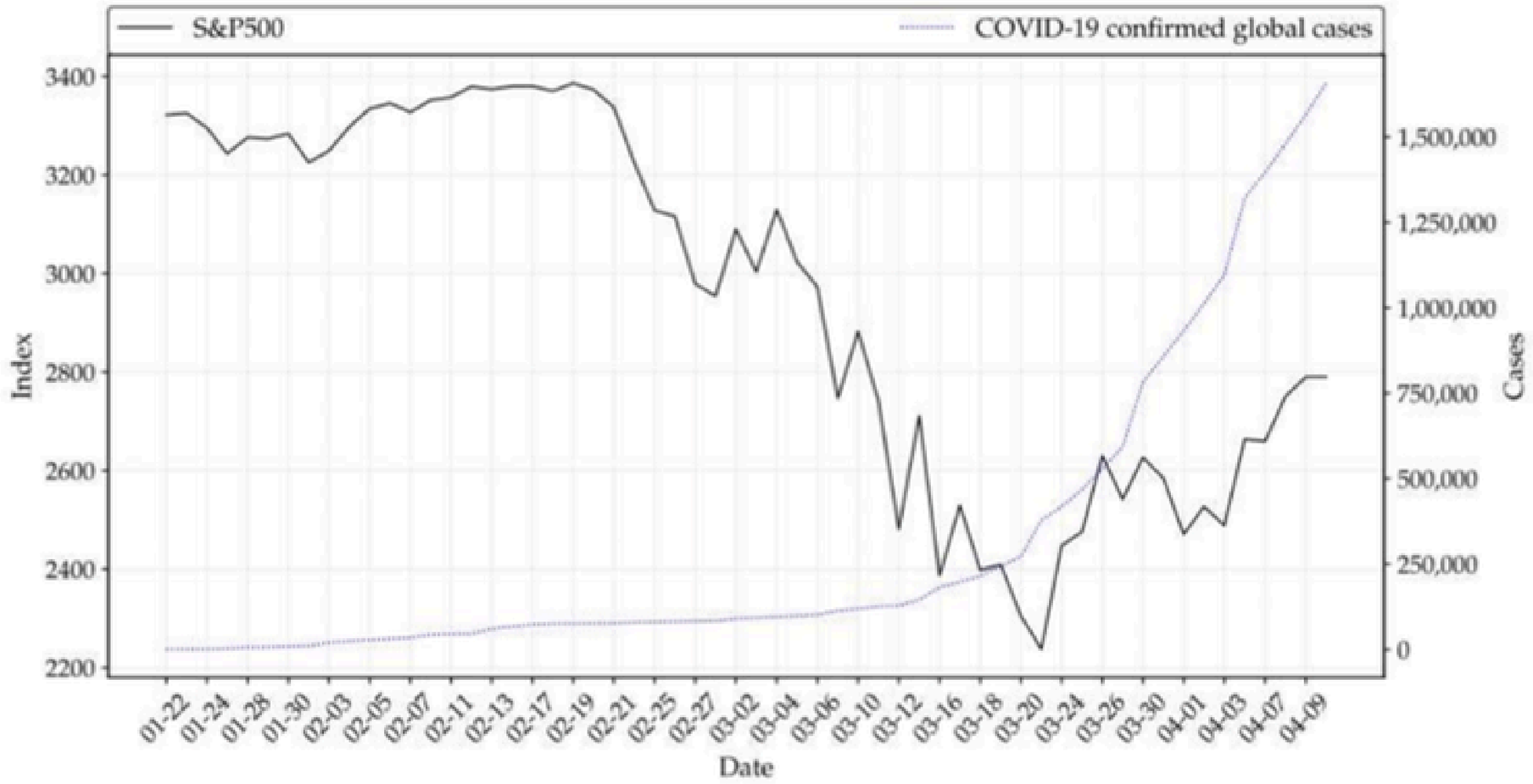
Examples of Timeseries

Newton School of Technology

Stock Price/commodity Price/Bitcoin etc

Examples of Timeseries

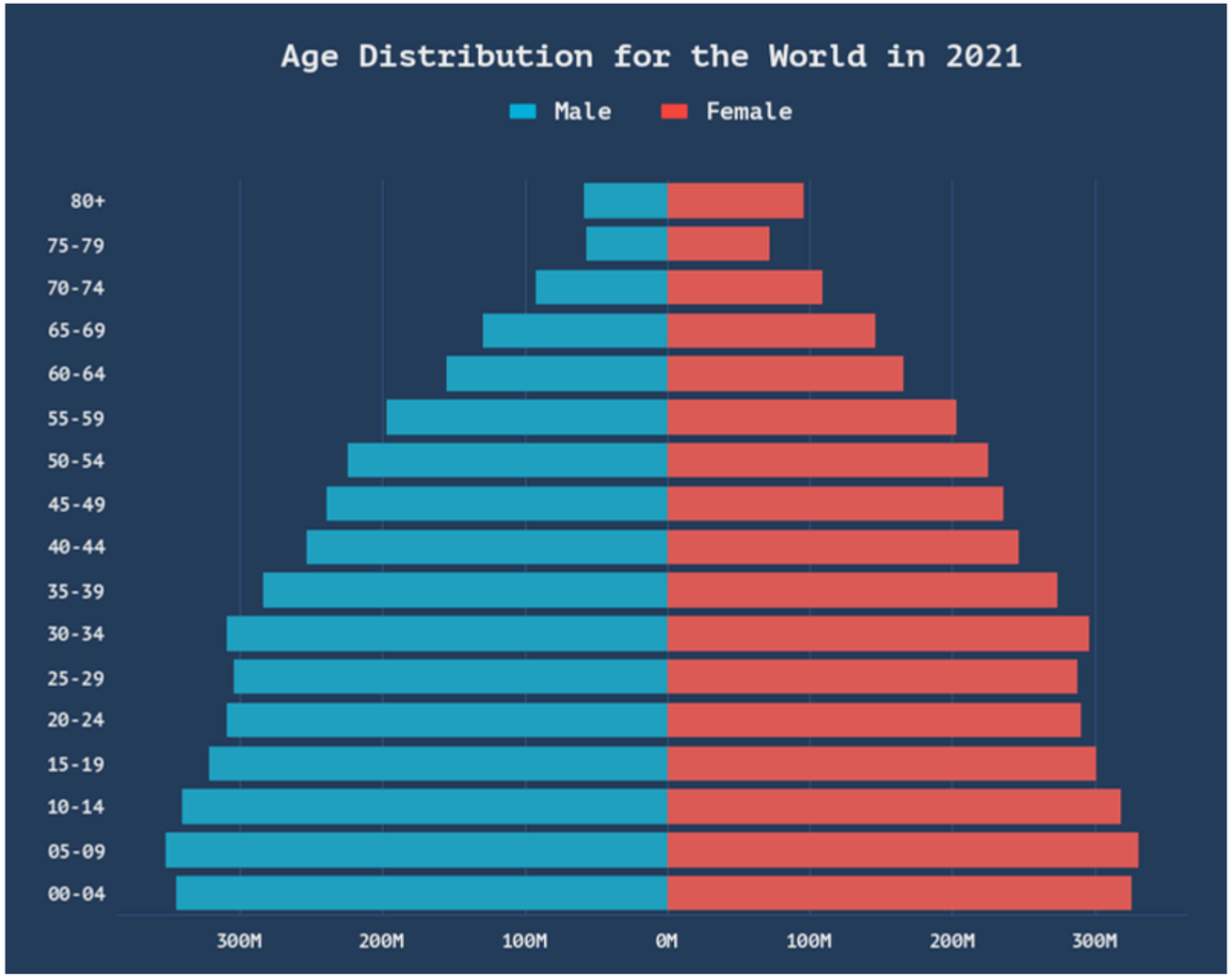
Newton School of Technology



S&P500 vs. COVID-19 confirmed cases (Engelhardt et al., 2020)

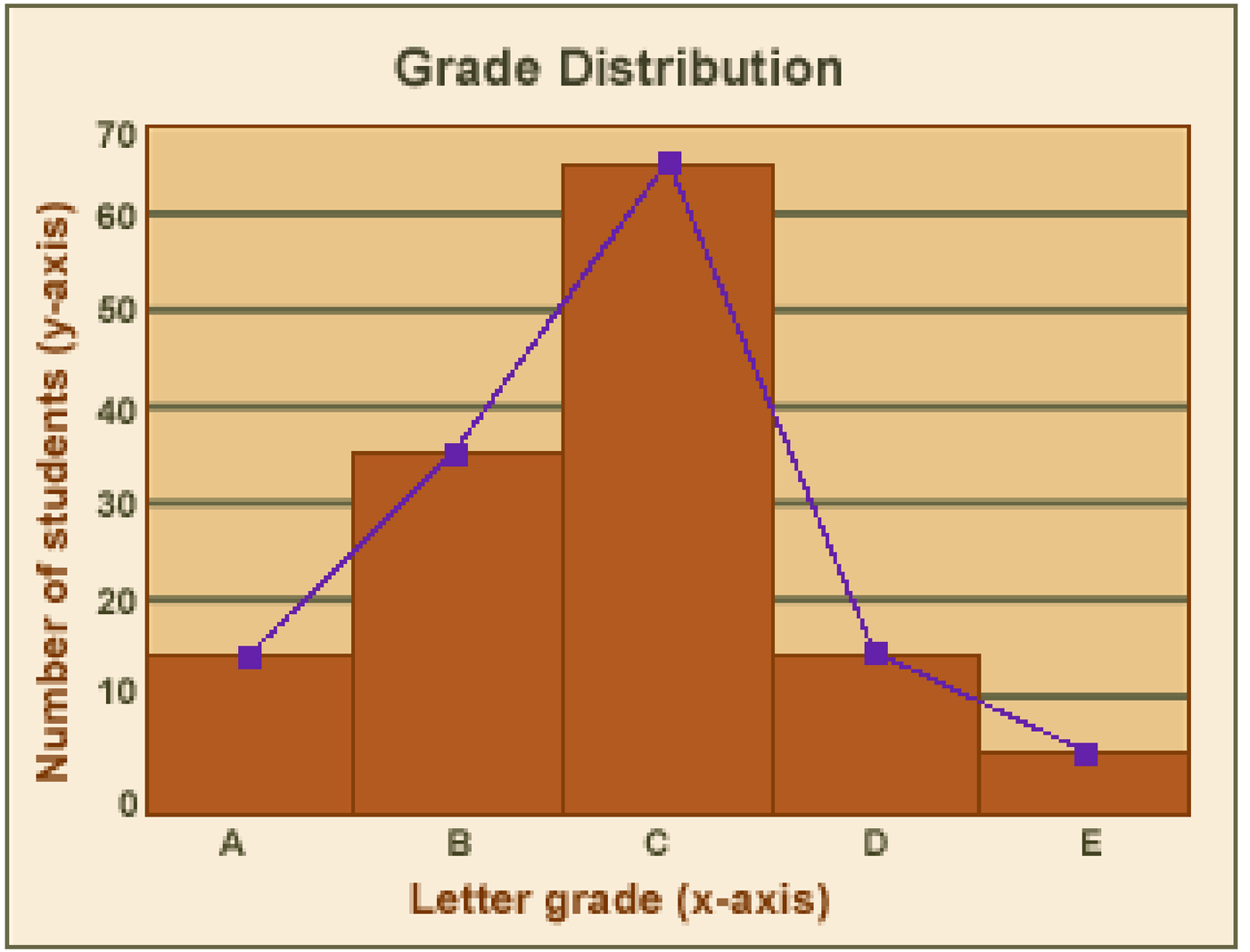
Is this a timeseries?

Newton School of Technology



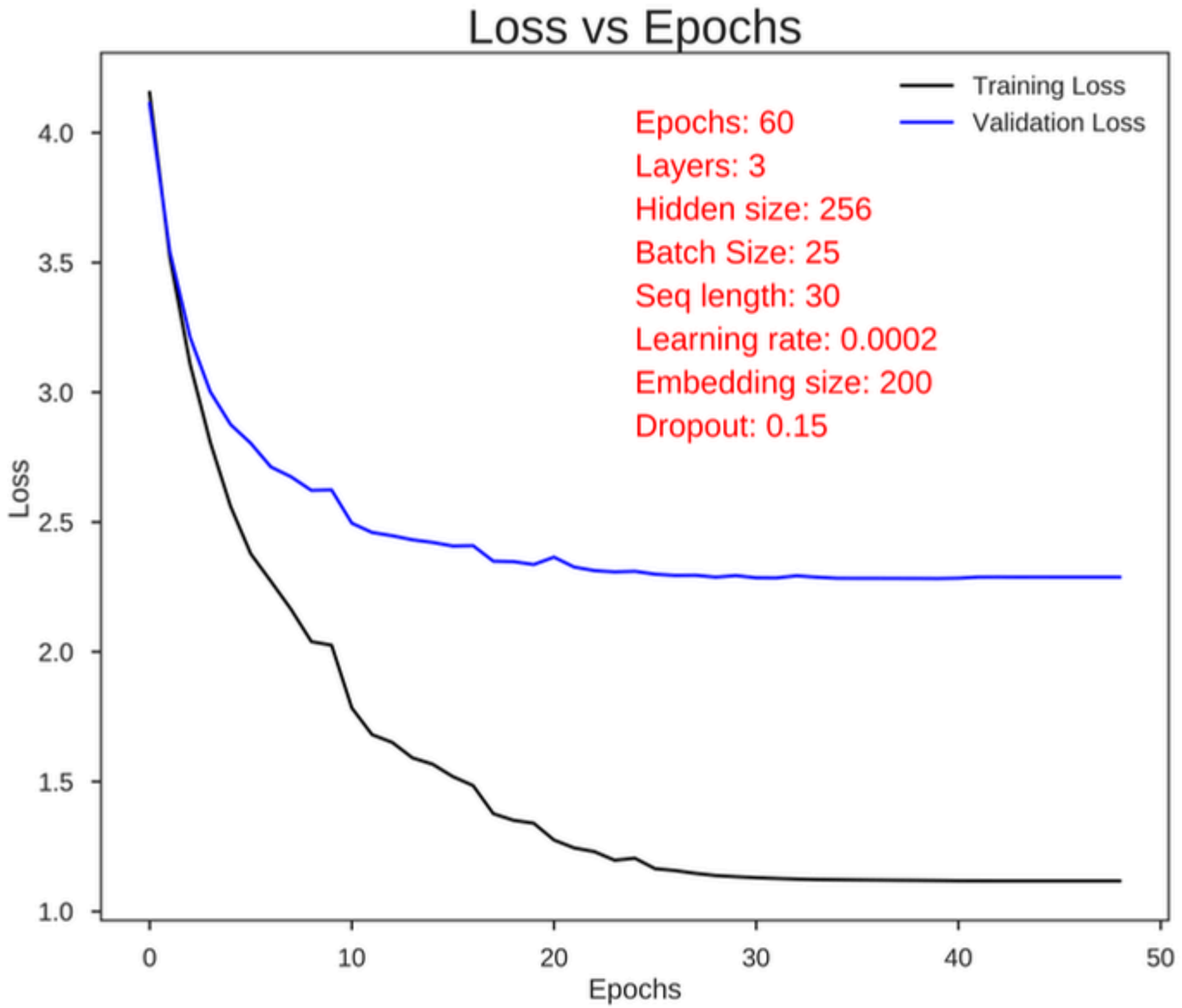
Is this a timeseries?

Newton School of Technology



Is this a timeseries?

Newton School of Technology



Is this a timeseries?

Newton School of Technology

Student_ID	Submission_Time (minutes)	Exam_Score
S01	35	78
S02	10	92
S03	55	60
S04	20	88
S05	45	70
S06	30	80

Problem

Newton School of Technology

A **factory engineer** records the **number of defects found** after **each production** batch inspection.

Batch Number	Defects Found
1	14
2	11
3	9
4	8
5	6
6	5
7	5
8	4

Problem

Newton School of Technology

A **factory engineer** records the **number of defects found** after **each production** batch inspection.

Is this **dataset** a **time series**?

Batch Number	Defects Found
1	14
2	11
3	9
4	8
5	6
6	5
7	5
8	4

Solution

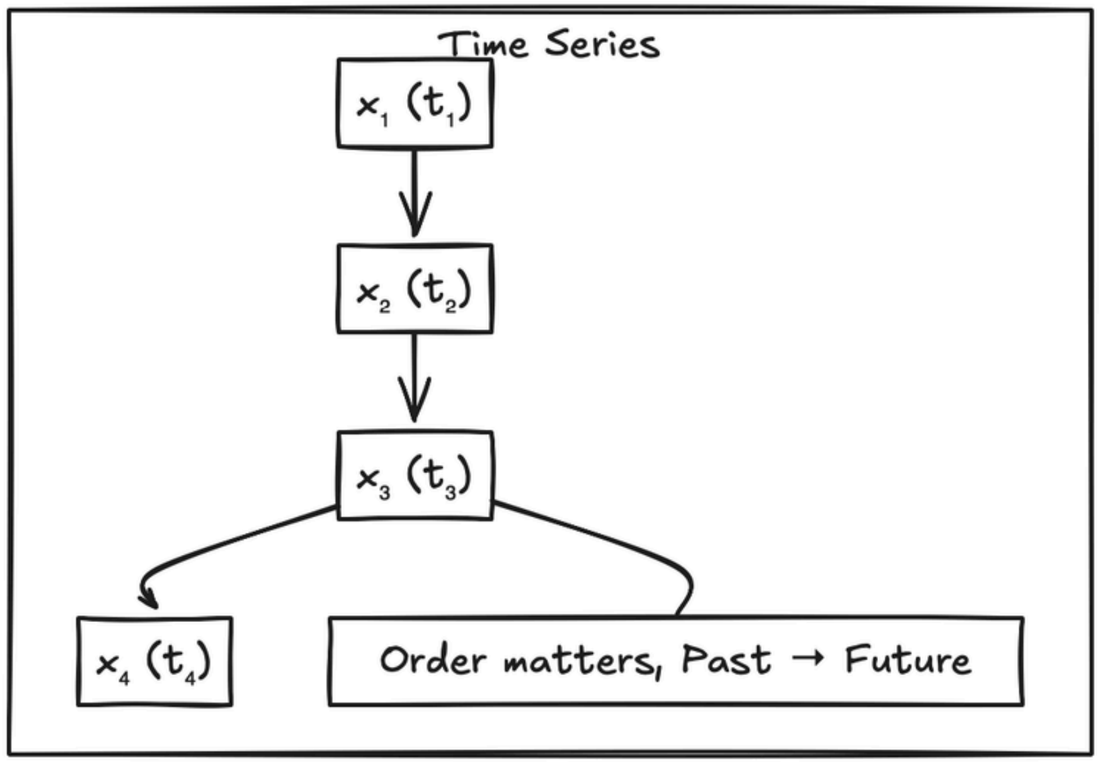
Newton School of Technology

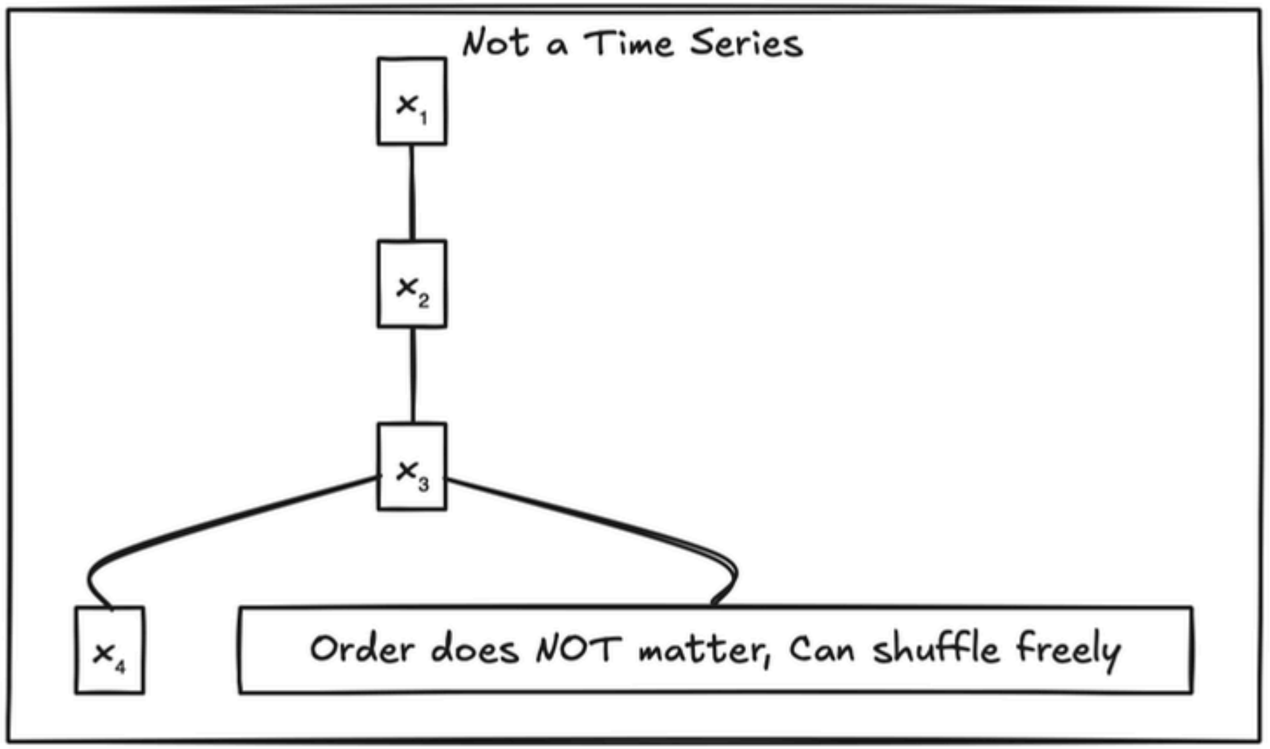
- **Observations are sequential**
- The **order cannot** be **changed**
- **Past production** improvements **affect future defect counts.**
- This is a **time series** data.

Batch Number	Defects Found
1	14
2	11
3	9
4	8
5	6
6	5
7	5
8	4

How to identify if data is timeseries?

Newton School of Technology





Time series: values are indexed by time, and changing order breaks meaning

Non–time series: values may be ordered, but order carries no information

If you can shuffle the x-axis and nothing breaks, it was never a time series.

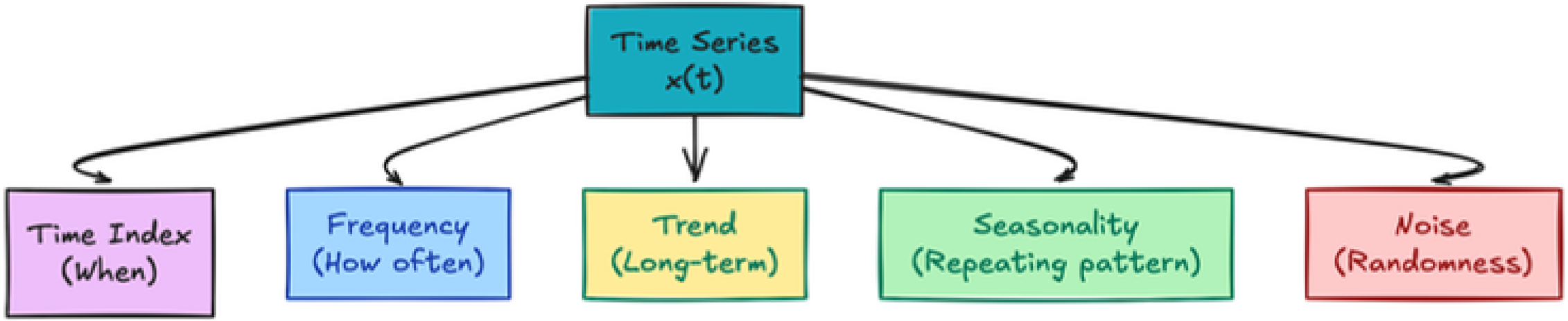
Newton School of Technology

All time series are line graphs, but not all line graphs are time series.

Property	Time Series	Looks Like Time Series
X-axis	Time	Anything ordered
Order matters	✅	❌
Temporal dependency	✅	❌
Forecasting possible	✅	❌
Can shuffle x-axis	❌	✅

Component of Time Series

Newton School of Technology



x(t)=Trend + Seasonality + Noise

Time Index and Frequency (structure),

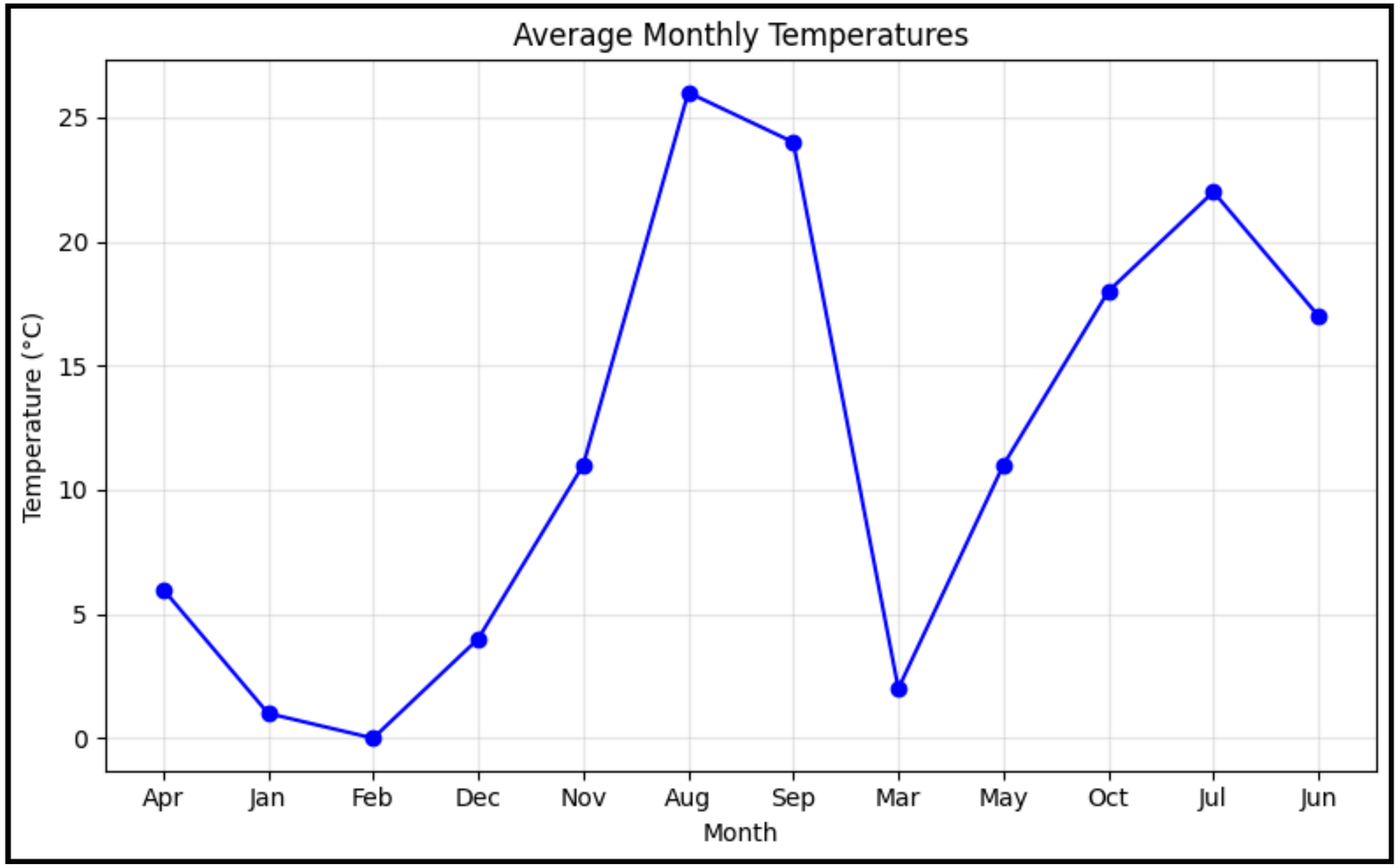
Trend, Seasonality, Noise (behavior)

Newton School of Technology

What will happen if we arrange the data at random?

Newton School of Technology

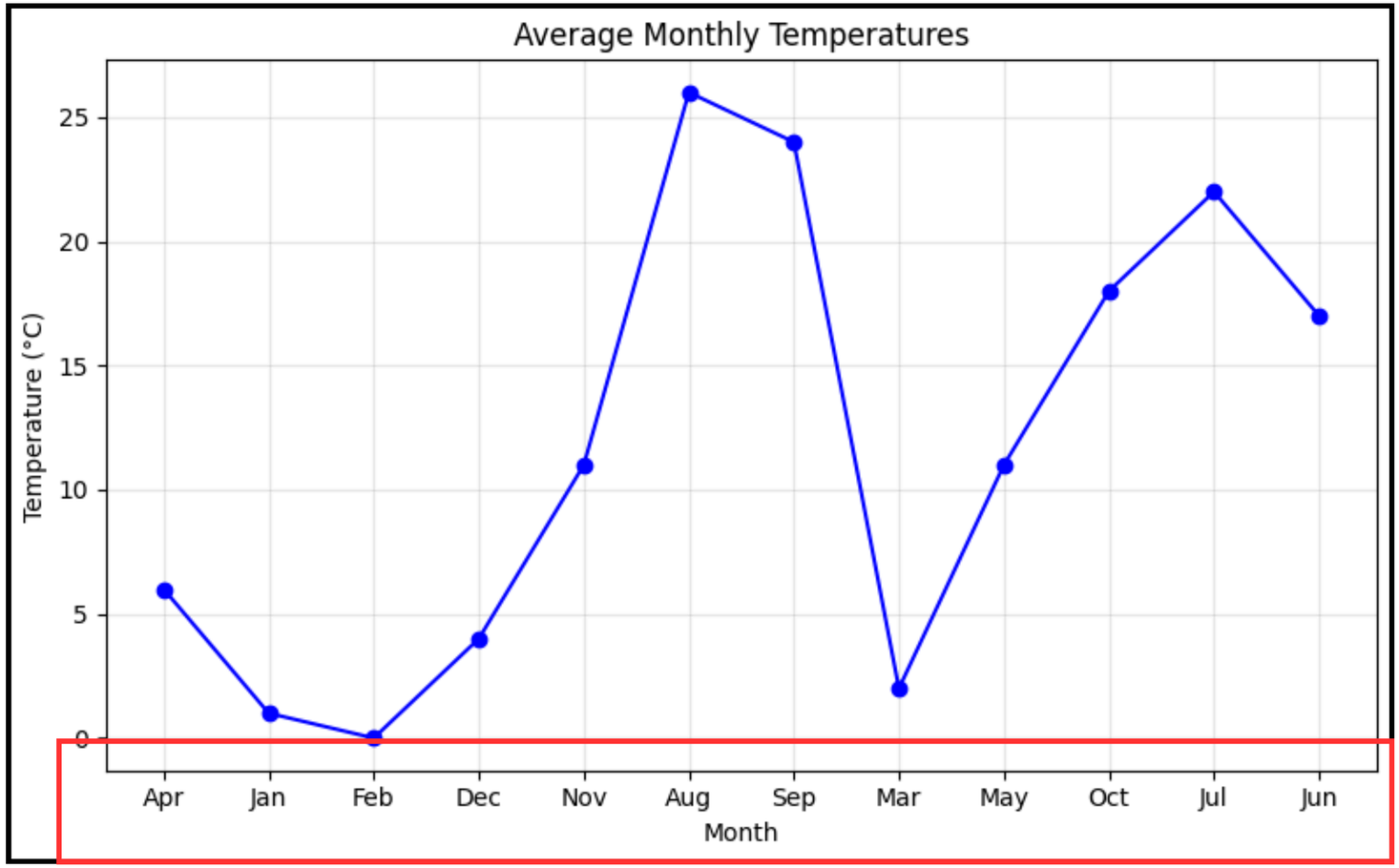
- What will happen if we arrange the data at random?



Time Index

Newton School of Technology

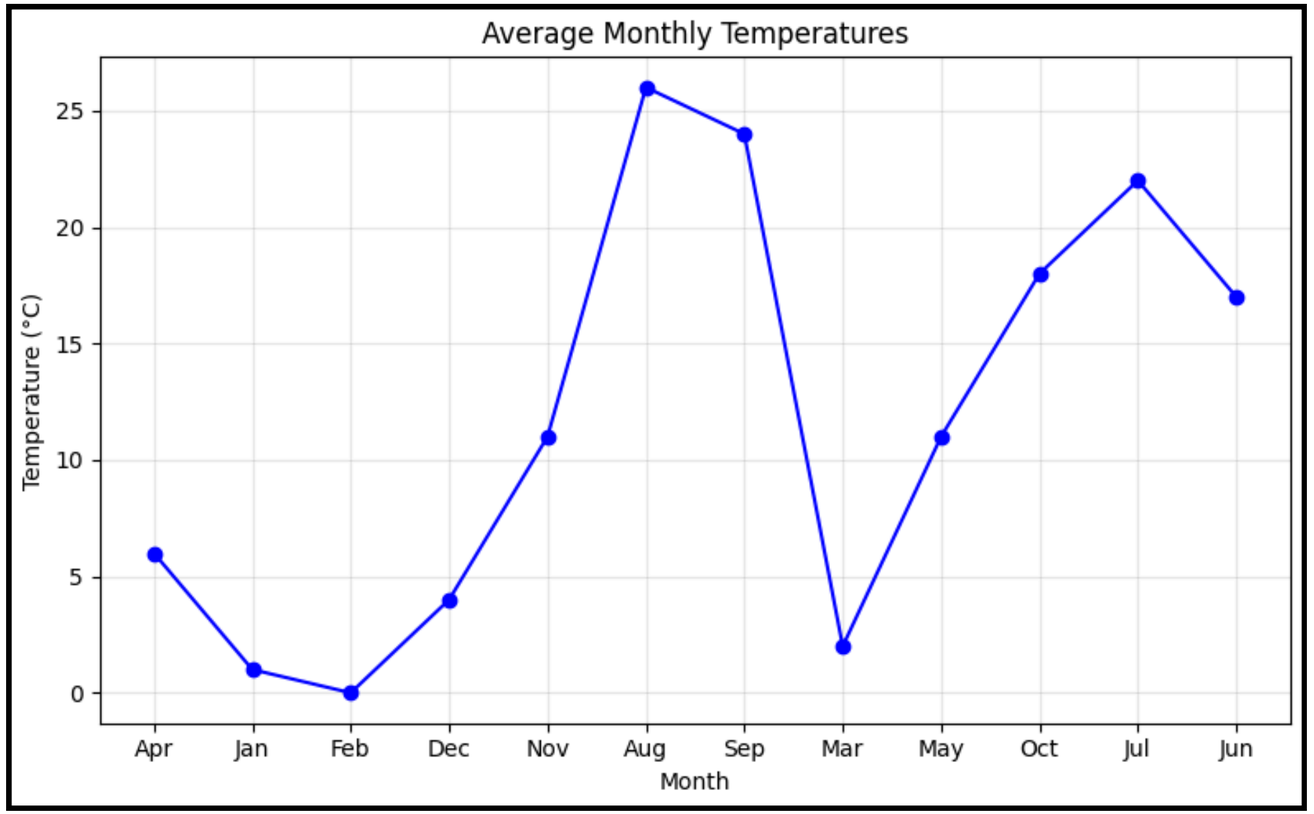
- The time index tells us WHEN each point was recorded.
- Without correct time stamps:
  - order becomes unclear
  - patterns break
  - interpretation becomes unreliable
- This is why shuffling the data earlier destroyed meaning — the time index was violated



Time Index

Newton School of Technology

- Time Index allows us to:
  - detect gaps
  - compare events
  - measure spacing between observations



Newton School of Technology

If we record something more often or less often, how does the pattern change?

Frequency

Newton School of Technology

- Frequency = time gap between observations
- **Examples:**
  - **Yearly (Inflation)**: Focuses on long-term trends; ignores temporary noise.
  - **Monthly (Temperature)**: Reveals seasonal cycles; smooths out daily variance.
  - **Daily (COVID):** Tracks short-term fluctuations and weekly patterns.
  - **Hourly (Traffic)**: Shows intraday peaks and valleys (e.g., rush hour).
  - **Per-second (IoT):** Captures instantaneous, critical state changes.

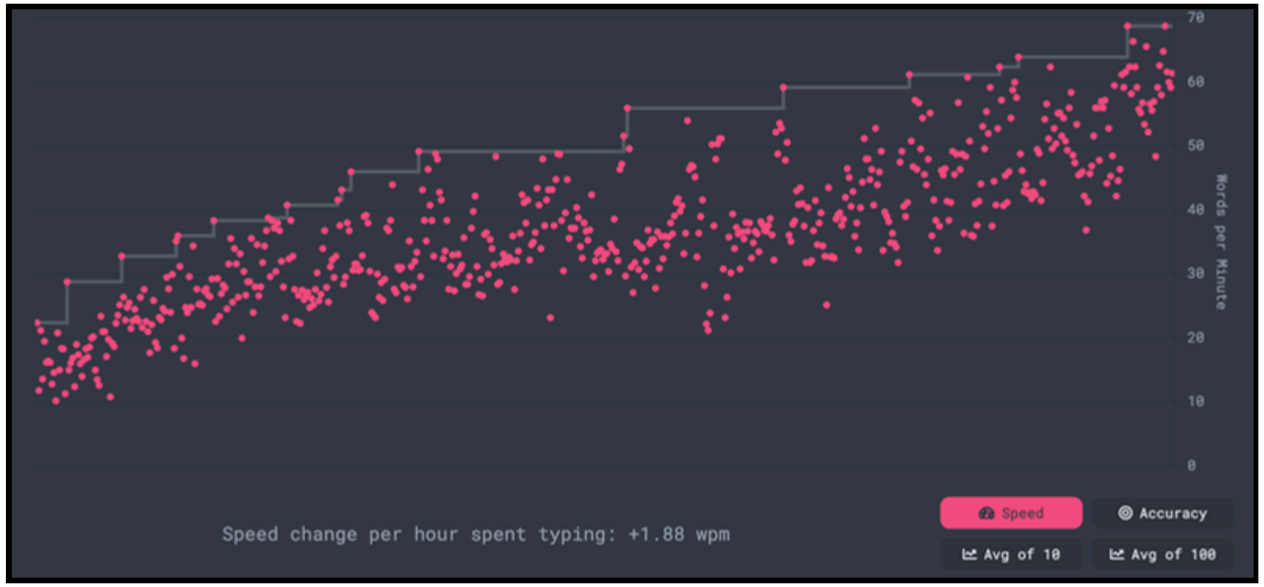
Frequency

Newton School of Technology

- How frequency affects meaning:
  - Higher frequency → more detail, more noise
  - Lower frequency → smoother, but hides patterns

Frequency

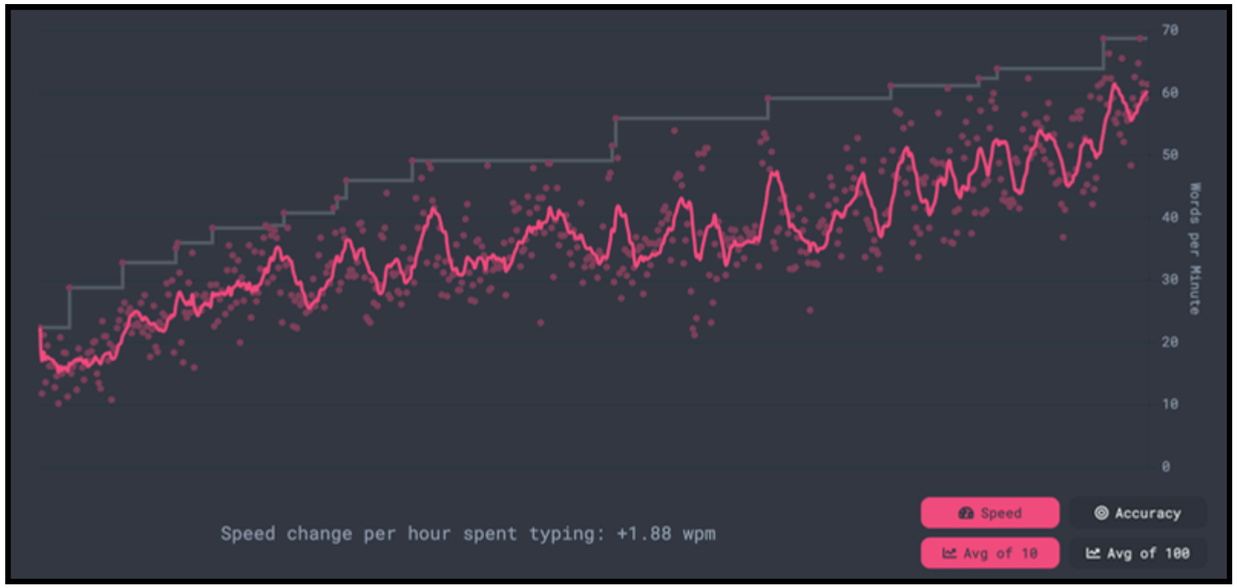
- This graph from Typing Monkey Website shows Shobhit Sir's typing progress.
- The **x-axis** represents **test timestamps**, and the **y-axis** shows **typing speed** in words per minute.
- Below are two buttons: one for the average of the **last 10 tests**, and the other for the average of the **last 100 tests**.



Frequency

Newton School of Technology

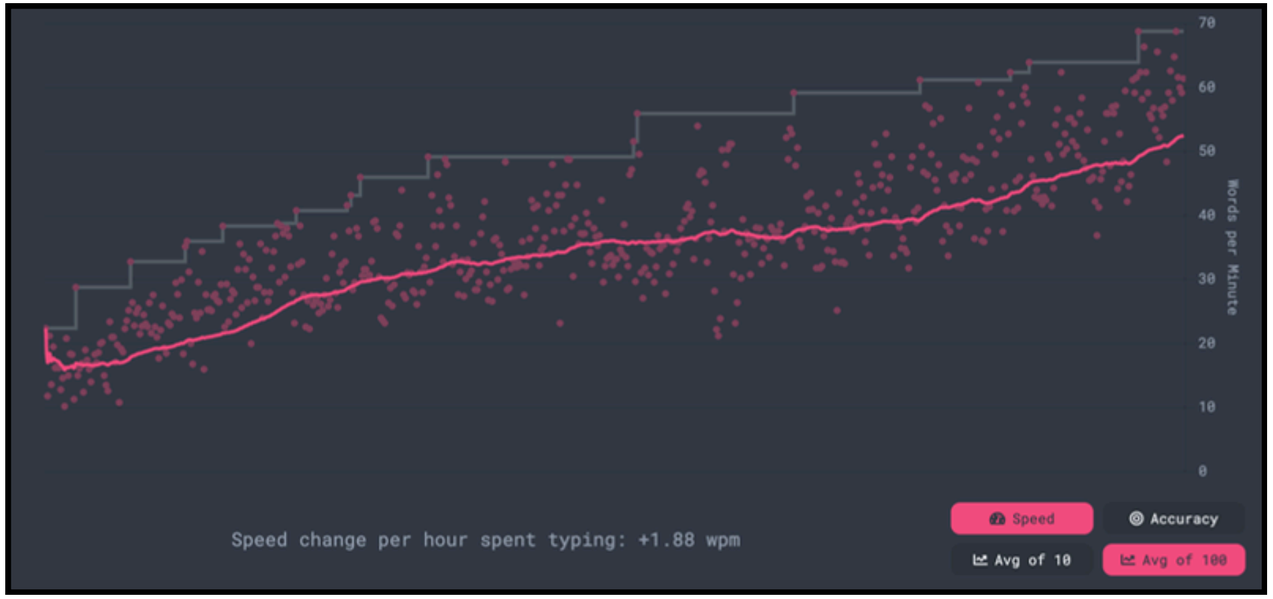
When we click the 'Avg of 10' button, it displays **high-frequency** time series data.



Frequency

Newton School of Technology

When we click the 'Avg of 100' button, it displays **low-frequency** time series data.



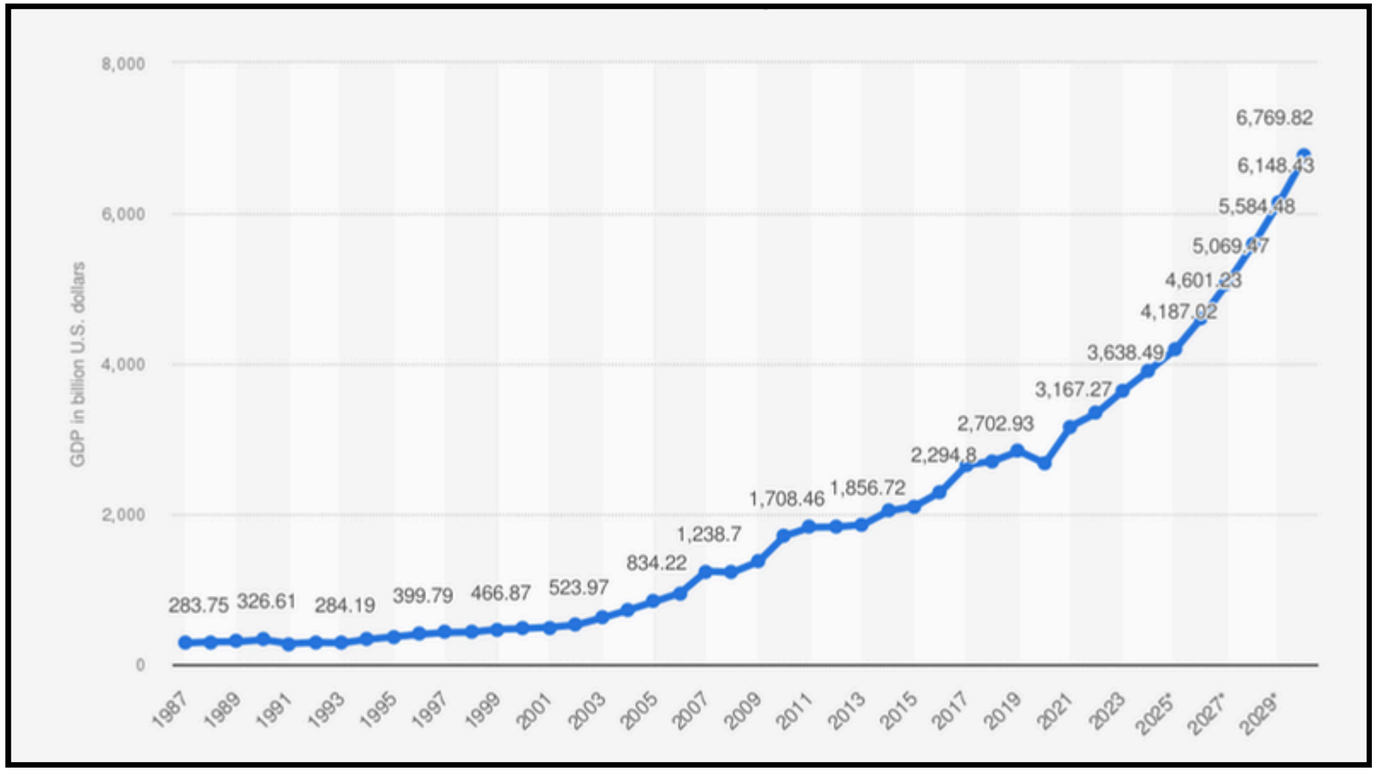
Frequency

Newton School of Technology

Frequency	Best For...	Shows	Hides
Yearly	Policy & History	Decades of trends	Seasons & Specific events
Monthly	Planning & Climate	Seasons (Summer vs Winter)	Daily weather spikes
Daily	News & Reaction	Weekly habits (Weekends)	Hour-by-hour shifts
Hourly	Operations & Logistics	Rush hours & Peaks	Instantaneous glitches
Per-Second	Safety & Machines	Immediate failures	Long-term drifts

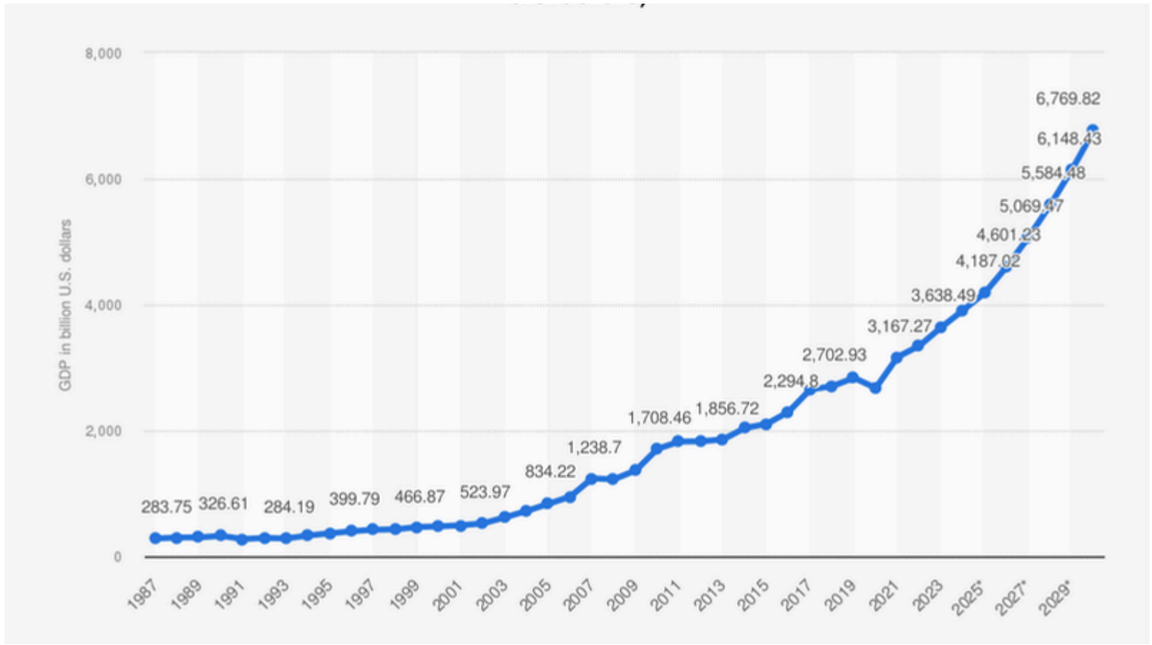
Newton School of Technology

Ignore the small ups and downs… what overall direction do you see?



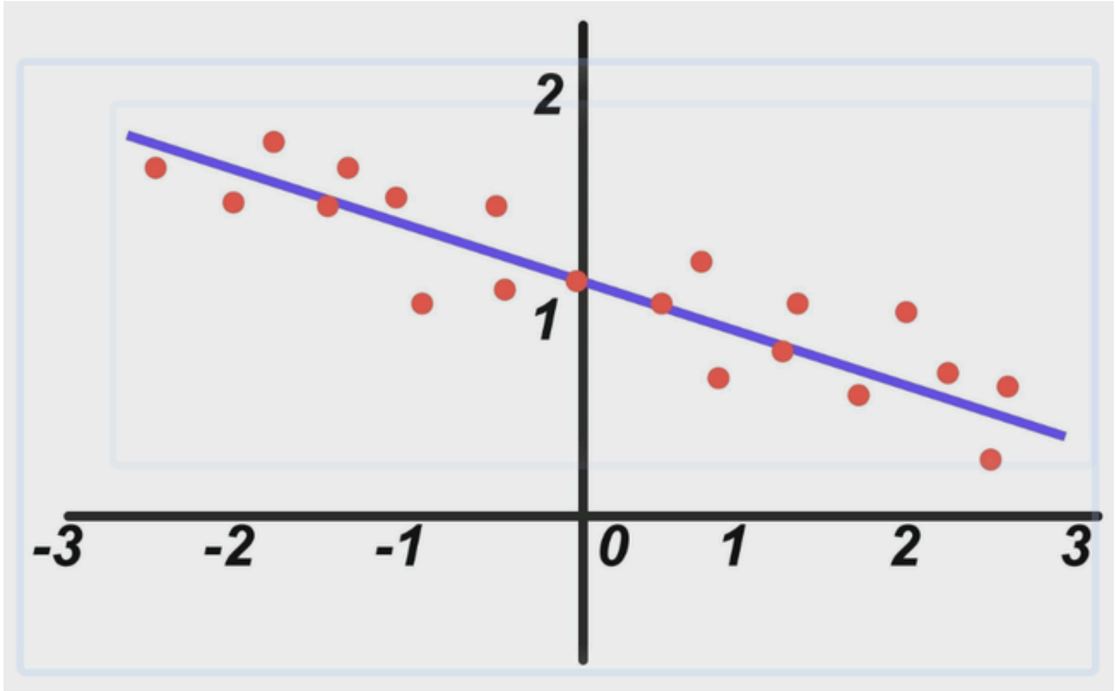
Trend — The long-term movement

- **Trend** = the general direction over long time periods
- **Could be**:
  - Increasing
  - Decreasing
  - Flattening
  - **The “big picture” hidden under local wiggles**



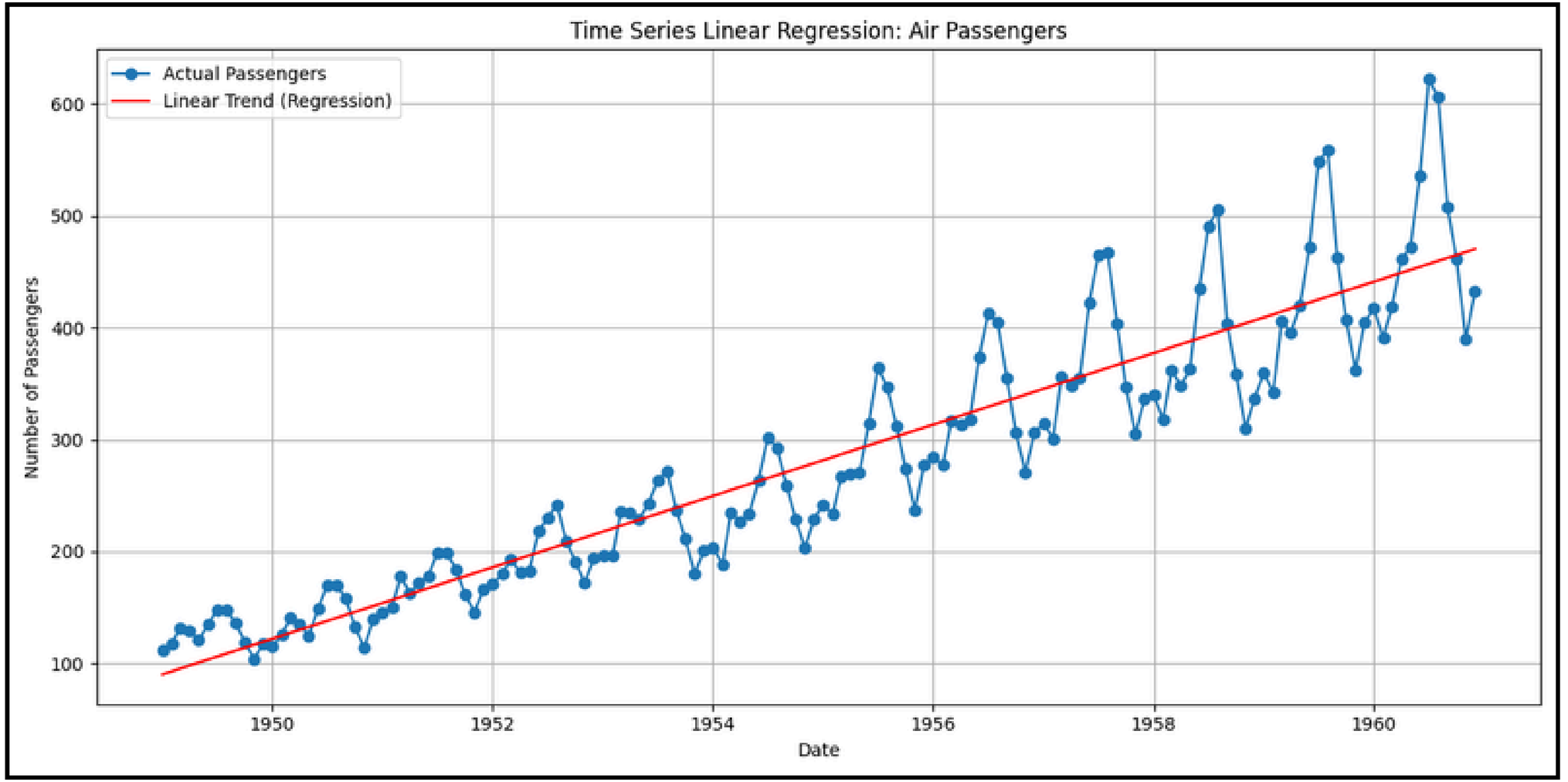
Linear Trends

- **Constant Rate:** The value increases or decreases by a fixed, unchanging amount for every unit of time (e.g., saving exactly $100 every month).
- **Straight Line:** When plotted on a graph, the data follows a straight diagonal path; it never curves, accelerates, or bends.
- **Predictability:** Because the speed of change is stable, it is the simplest pattern to forecast—you just extend the straight line.



Real life example of Linear Trend

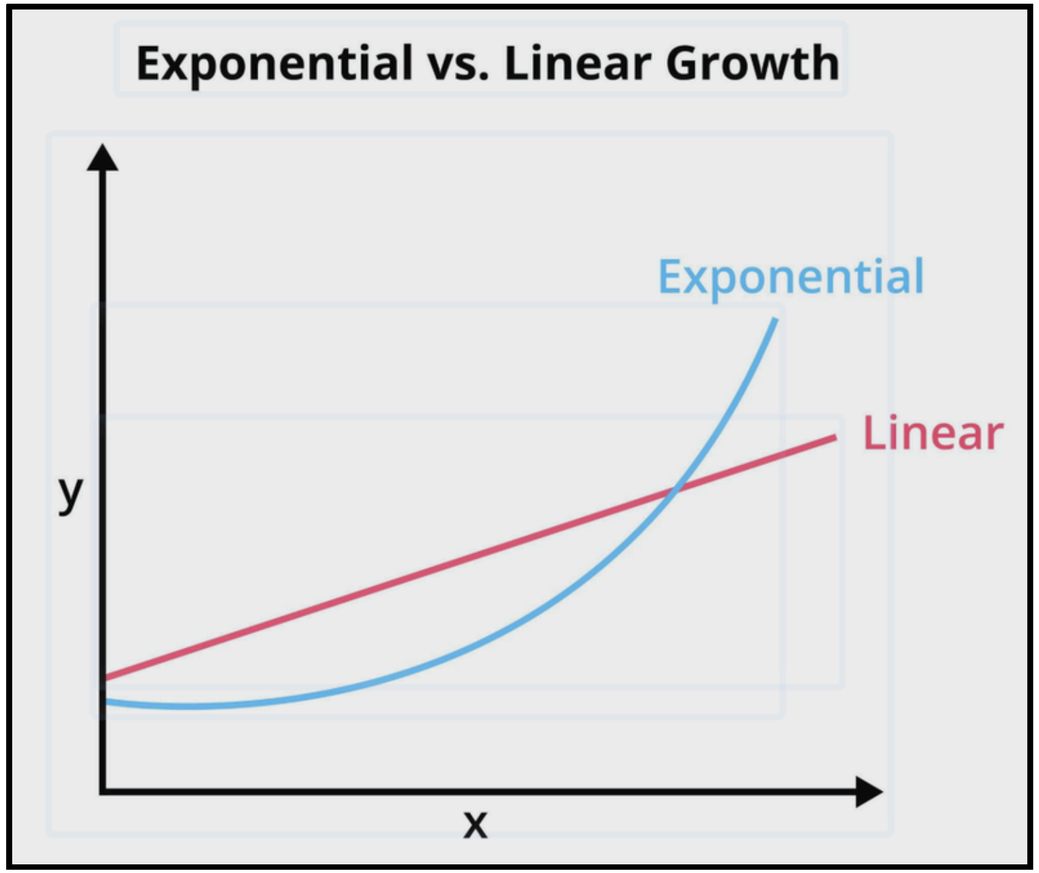
Newton School of Technology



Non-Linear Trends

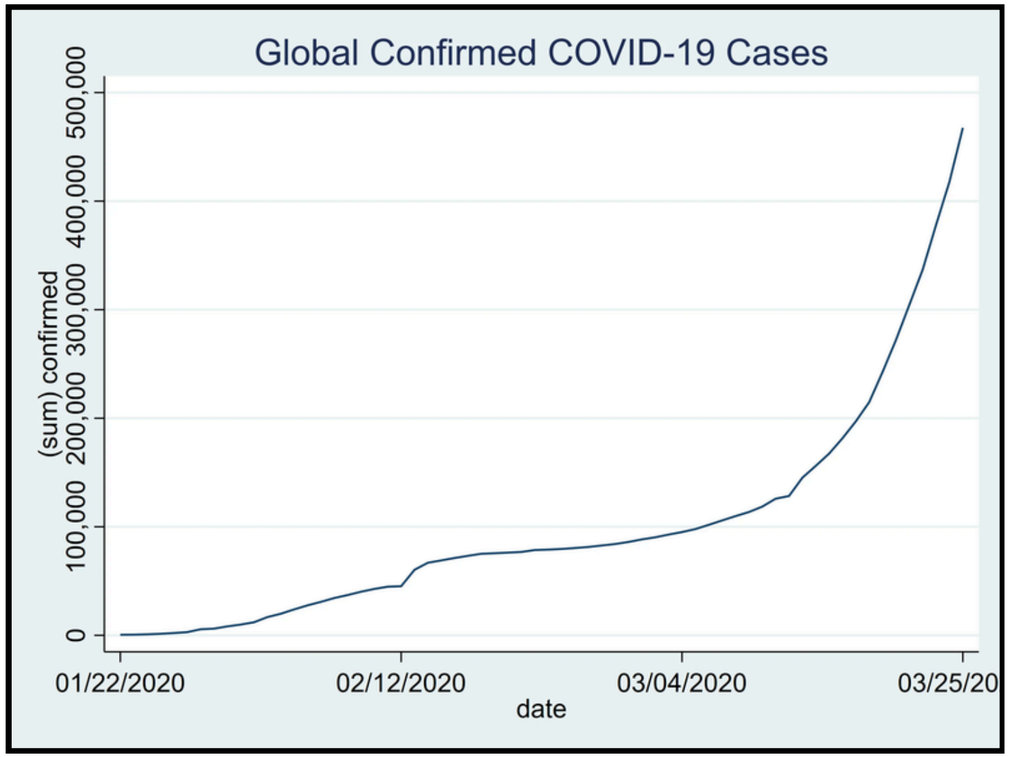
Newton School of Technology

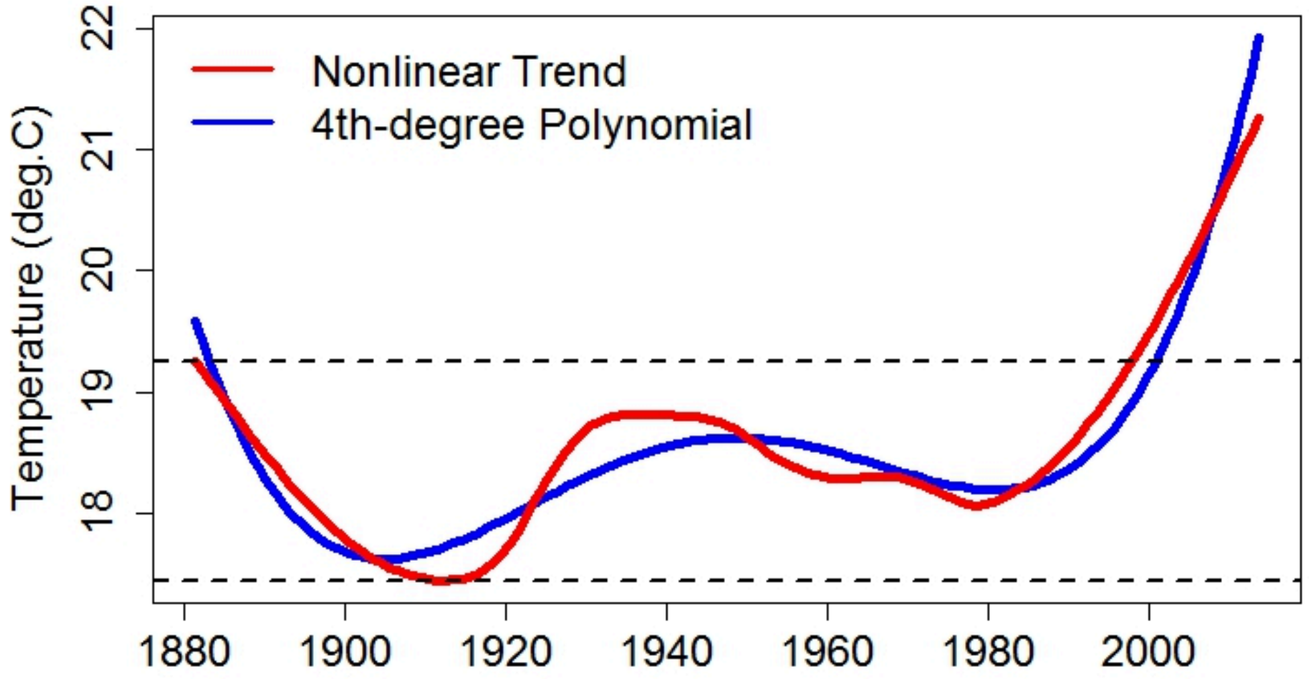
- **Variable Rate:** The speed of change is not constant; it accelerates (speeds up like a virus spreading) or decelerates (slows down like a car braking).
- **Curved Line:** When plotted, the graph bends, arcs, or waves; it does not follow a straight diagonal path.
- **Complex Behavior:** These patterns often show explosive growth (exponential) or diminishing returns (plateaus), making them harder to predict than linear trends.
- Patterns can be Exponential, Quadratic, etc.



Real life example of Non-Linear Trends

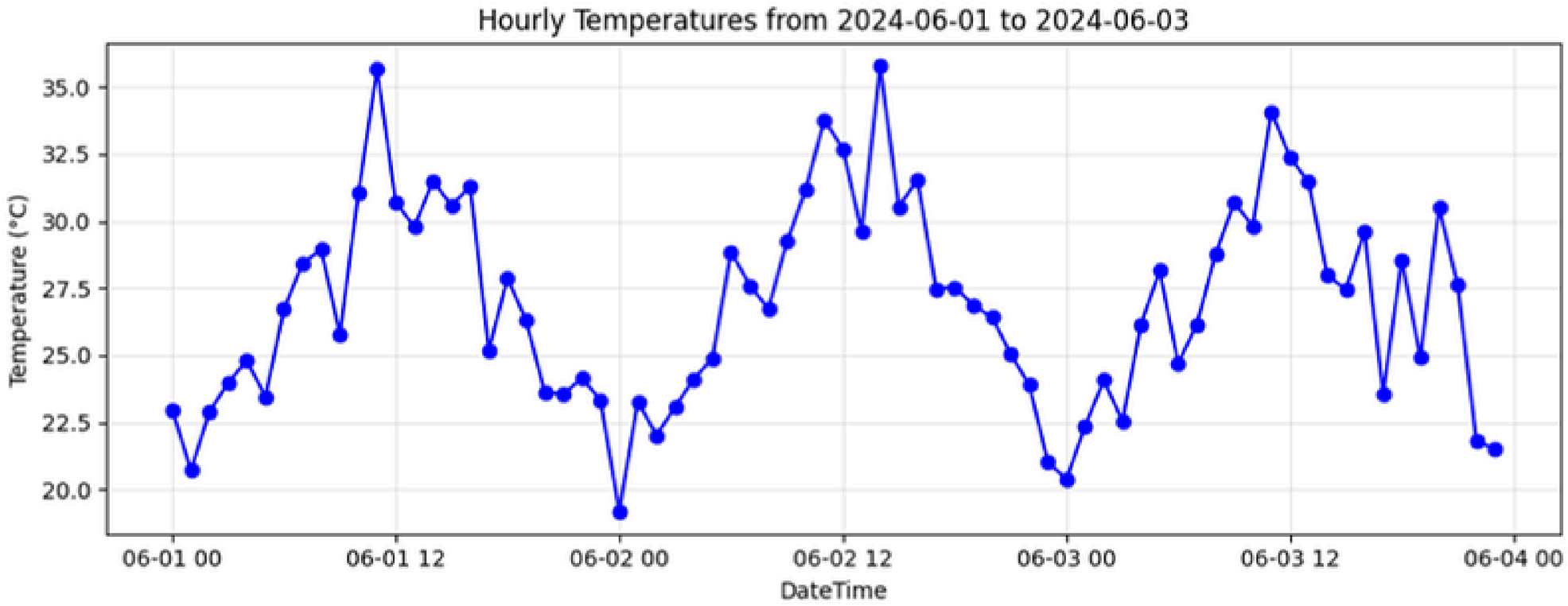
Newton School of Technology





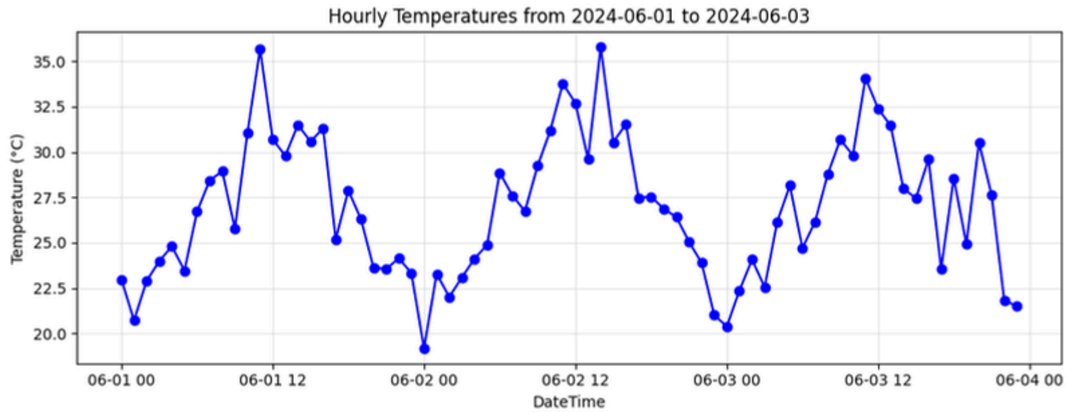
Do you see anything repeating regularly?

Newton School of Technology



Seasonality

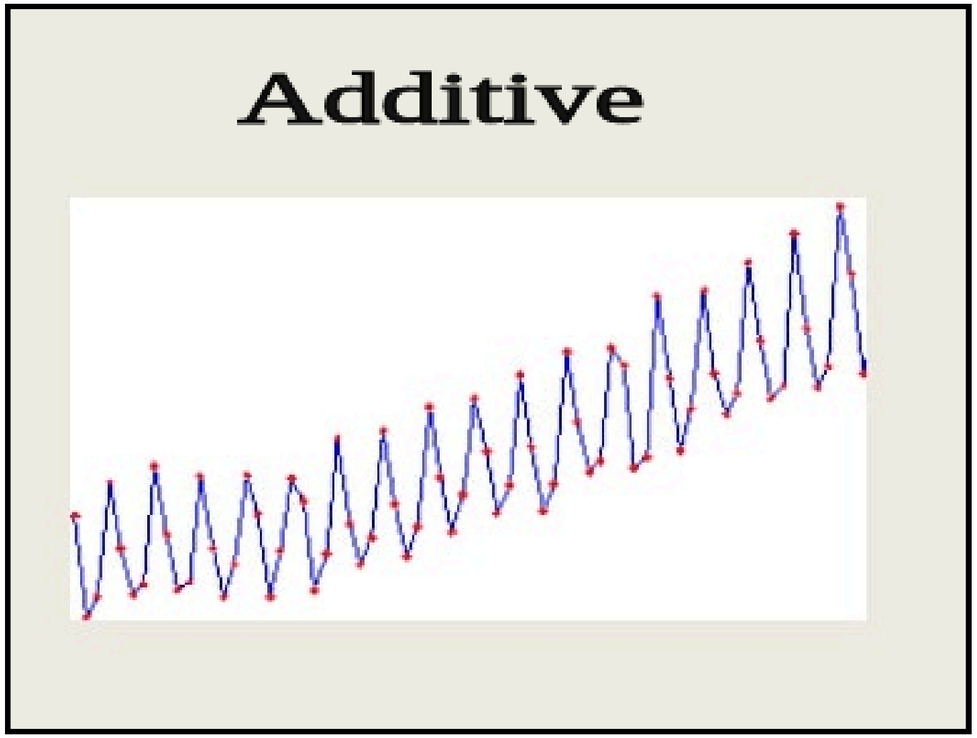
- **Some patterns repeat at fixed intervals:**
  - daily
  - weekly
  - monthly
  - yearly
- **Examples:**
  - Weekend dips in traffic
  - Festival season spikes
  - Summer heat, winter cold



Additive Seasonality

Newton School of Technology

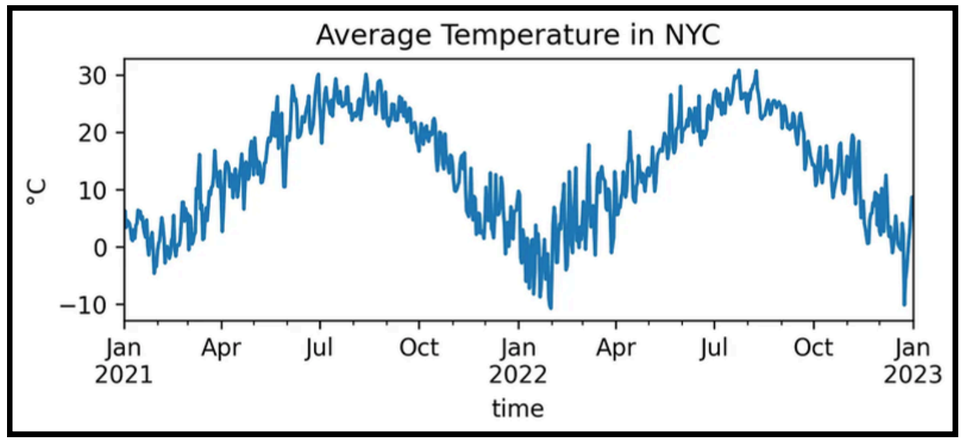
Additive seasonality means that a time series has regular seasonal changes that stay about the same size over time.



Additive Seasonality

Newton School of Technology

- **Example**: **Daily Temperature**
  - Summer is always warmer than winter
  - The difference between summer and winter is fairly stable each year
  - Even if the average temperature changes slightly, the seasonal swing remains similar

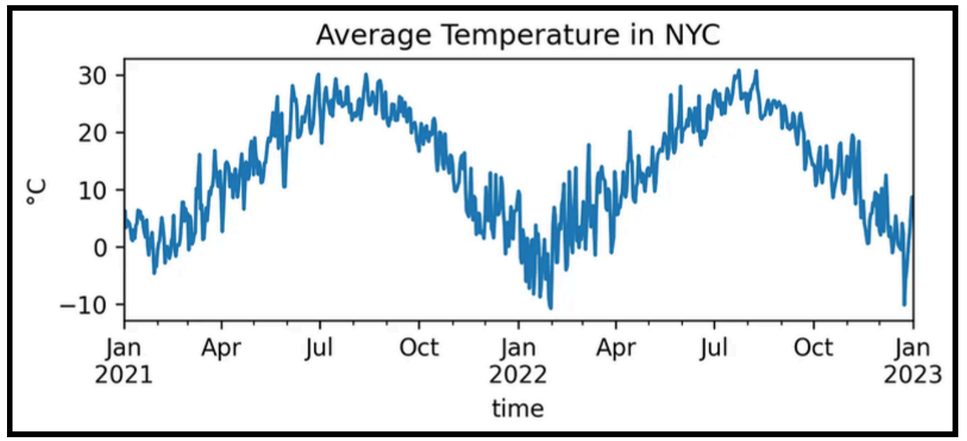


Additive Seasonality

Newton School of Technology

How It Looks Visually

- Seasonal peaks and troughs have similar height
- Fluctuations do not grow as the trend increases



When You Typically See Additive Seasonality

Newton School of Technology

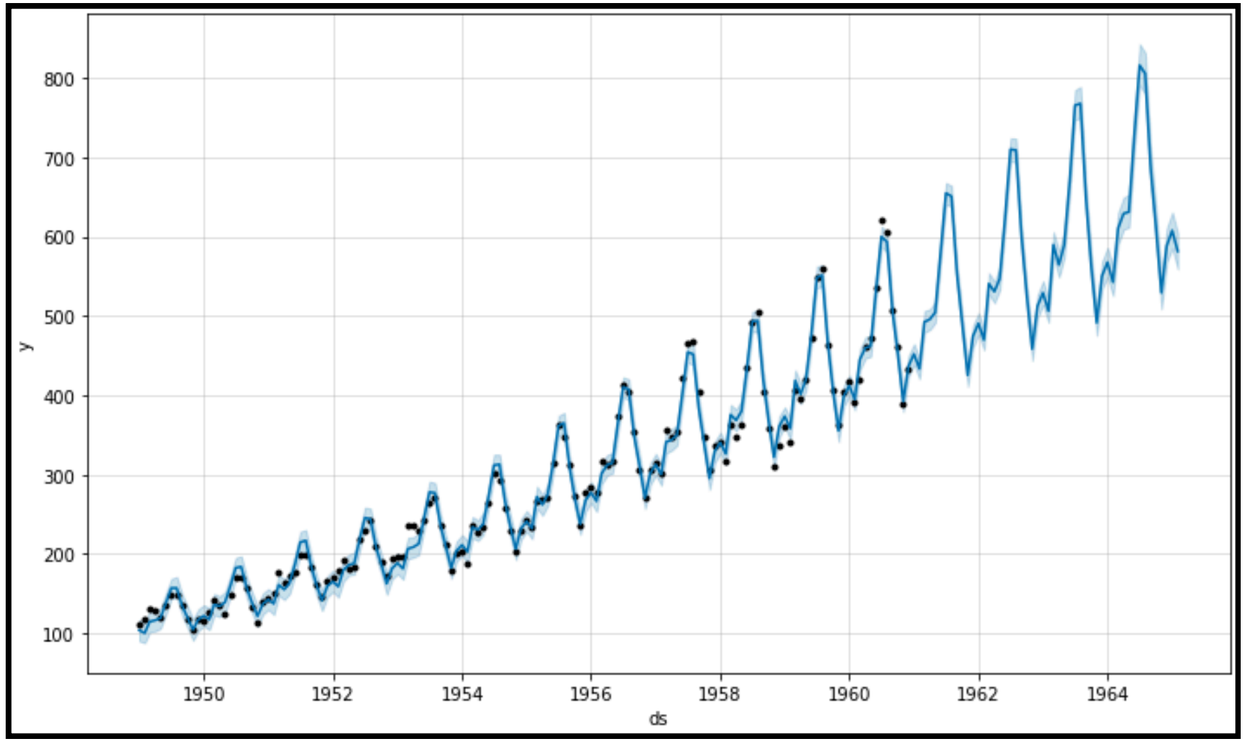
- Temperature data
- Electricity usage in a stable population
- Website traffic with fixed user base
- Manufacturing output with capacity limits

Multiplicative Seasonality

Newton School of Technology

Multiplicative seasonality means that seasonal changes depend on how big the data already is.

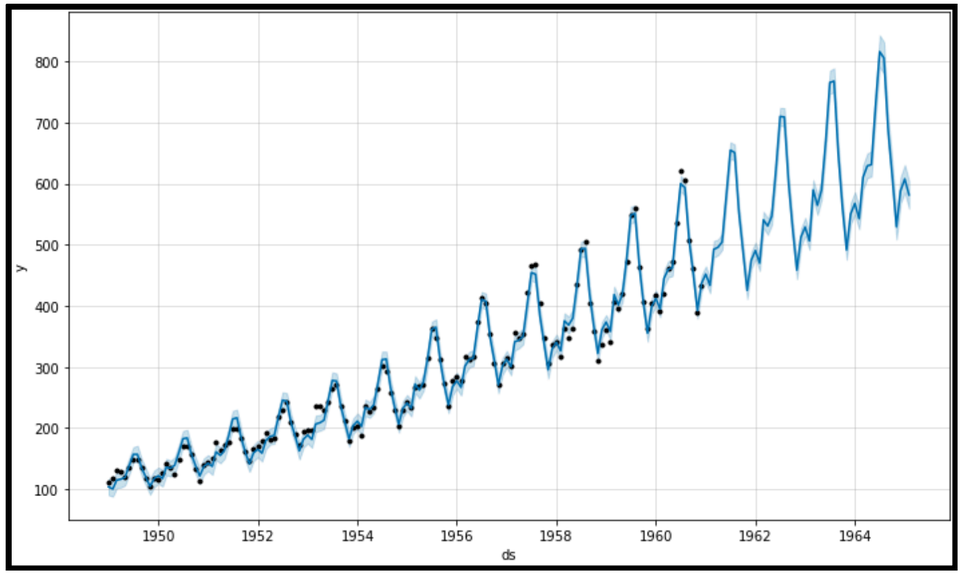
The season changes the value by a percentage, not a fixed amount.



Multiplicative Seasonality

Newton School of Technology

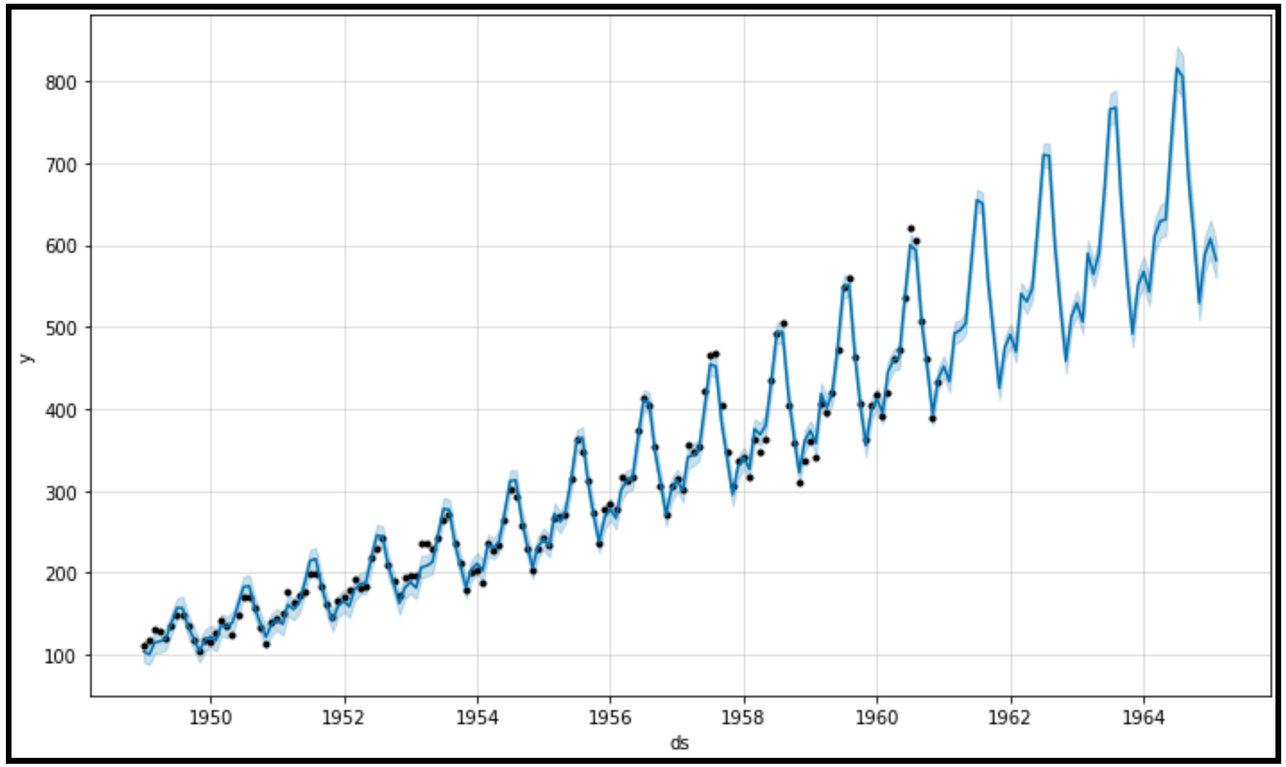
- **Example**: A growing business sells more each year
  - Holiday sales increase in proportion to total sales
  - The seasonal spike becomes larger over time
  - Example:
  - Year 1 December sales: +200 units
  - Year 5 December sales: +2,000 units
  - The seasonality scales with growth.



Multiplicative Seasonality

Newton School of Technology

- **How It Looks Visually**
  - Seasonal fluctuations fan out over time
  - Peaks and troughs grow as the trend increases



When You Typically See Multiplicative Seasonality

Newton School of Technology

- Retail sales
- Revenue and profits
- E-commerce traffic
- Population-driven demand

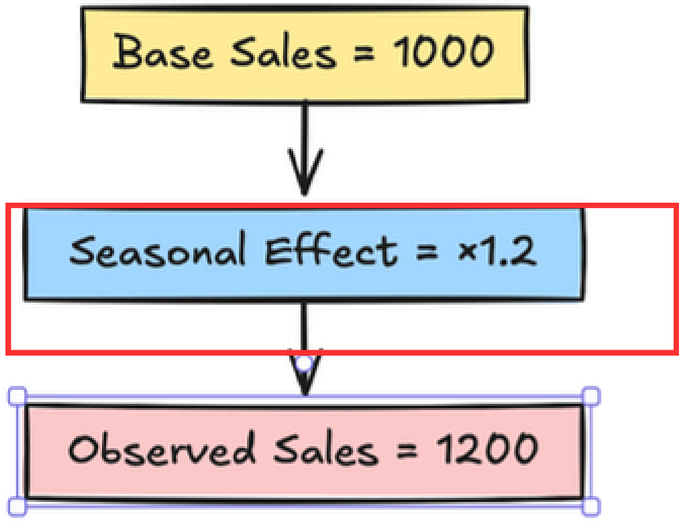
Additive vs Multiplicative Seasonality

Newton School of Technology

Additive Seasonality

Year 1

Multiplicative Seasonality



If next year base becomes 2000 — what happens?

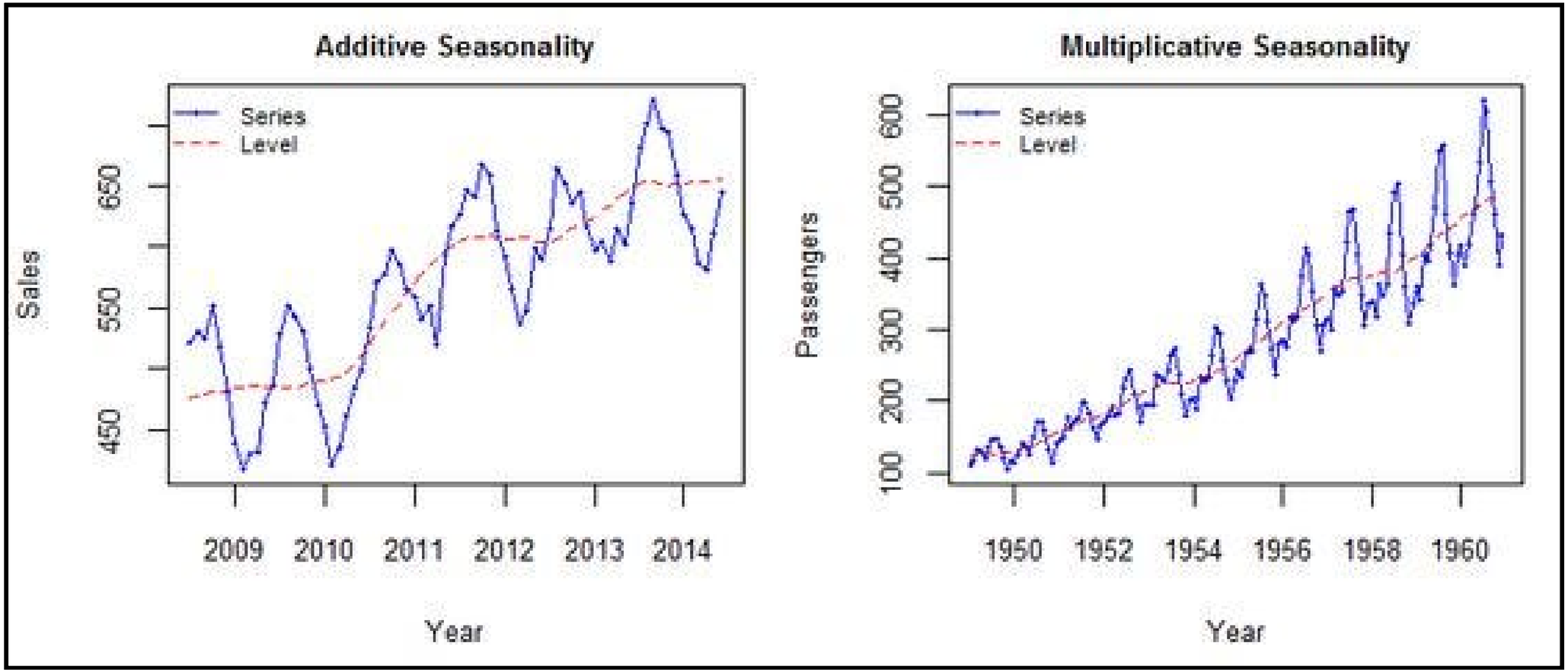
- Additive → 2000 + 200 = 2200
- Multiplicative → 2000 × 1.2 = 2400

Same amount → Additive

Same percentage → Multiplicative

Additive Seasonality vs Multiplicative Seasonality

Newton School of Technology



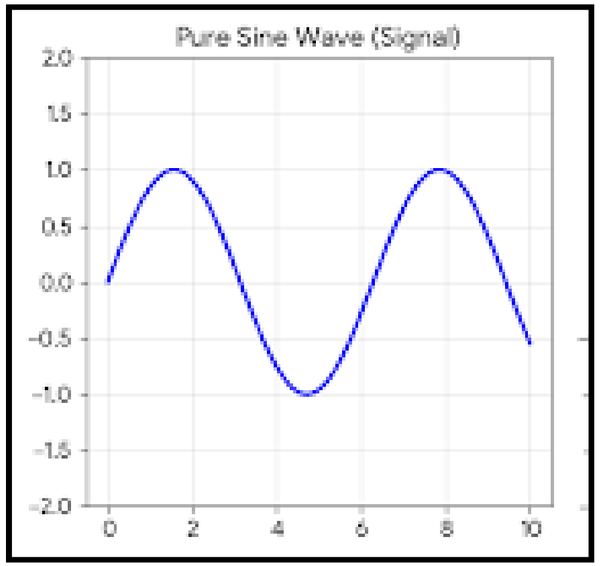
Noise

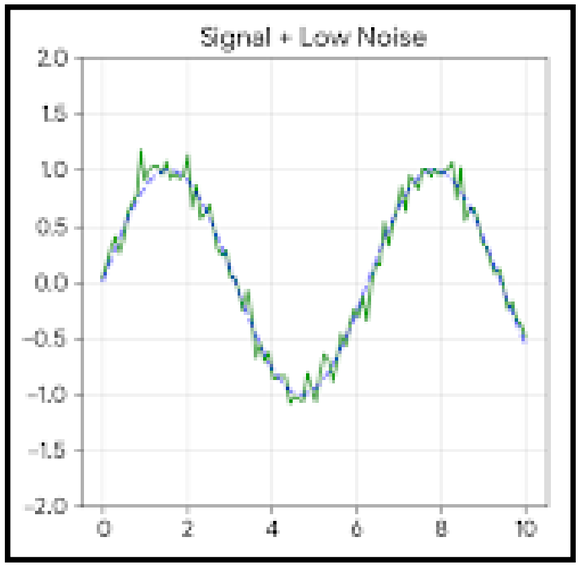
Noise is the **random, irregular fluctuation** in data that obscures the true underlying pattern and **cannot be predicted**.

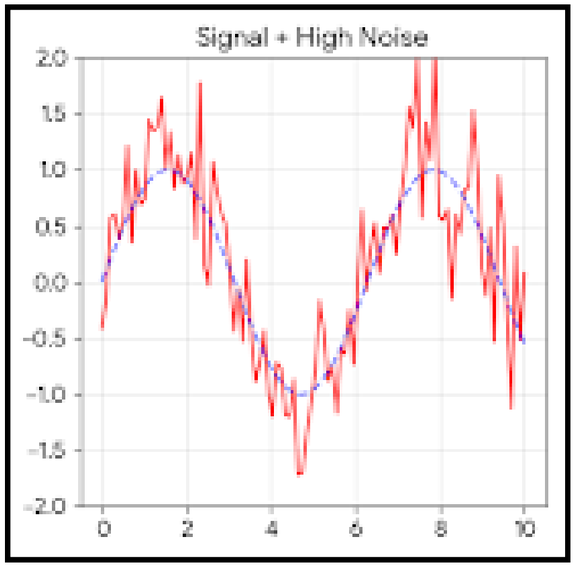
- **Unpredictable**: It is the irregular fluctuation in data that has no pattern or specific cause (like a measurement error or a sudden, random event).
- **"Jittery" Look**: On a graph, it appears as erratic spikes, static, or "fuzziness" scattered around the main trend line, making the plot look messy.
- **Masks the Actual Data:** Unlike trends or seasonality, noise contains no useful information; it is the "static" that makes it hard to see the real pattern.

Noise

- **Noise includes:**
  - measurement errors
  - unexpected events
  - anomalies
  - randomness





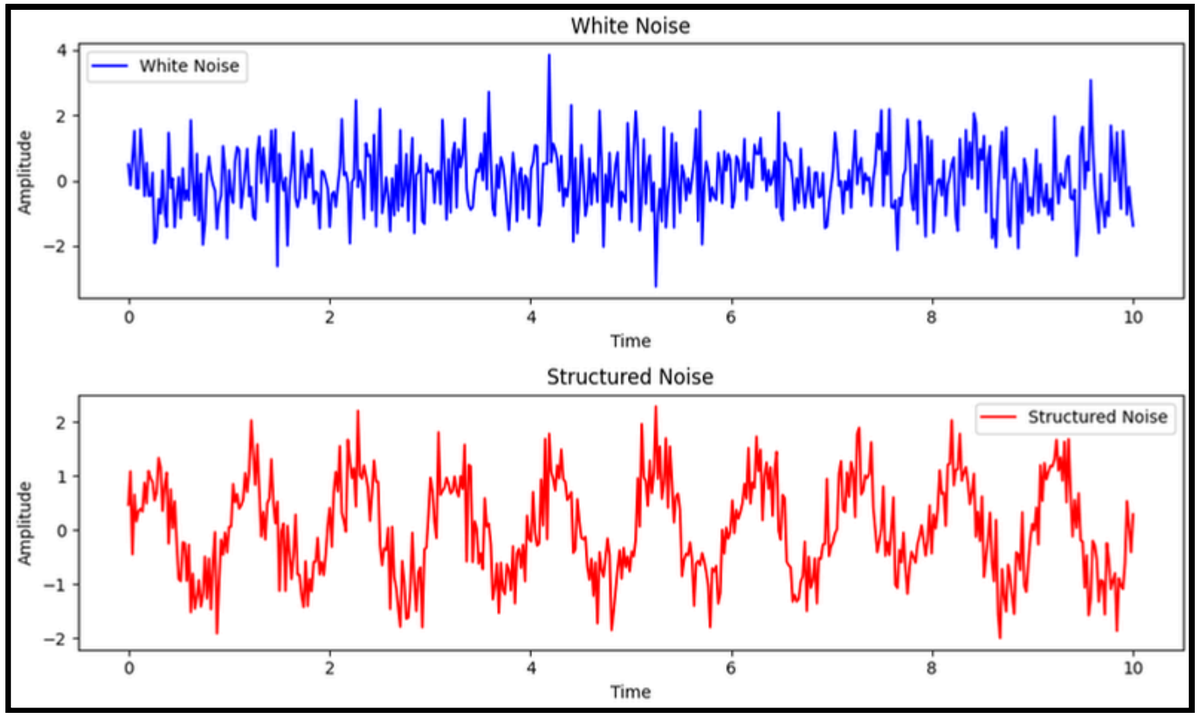


Noise

Newton School of Technology

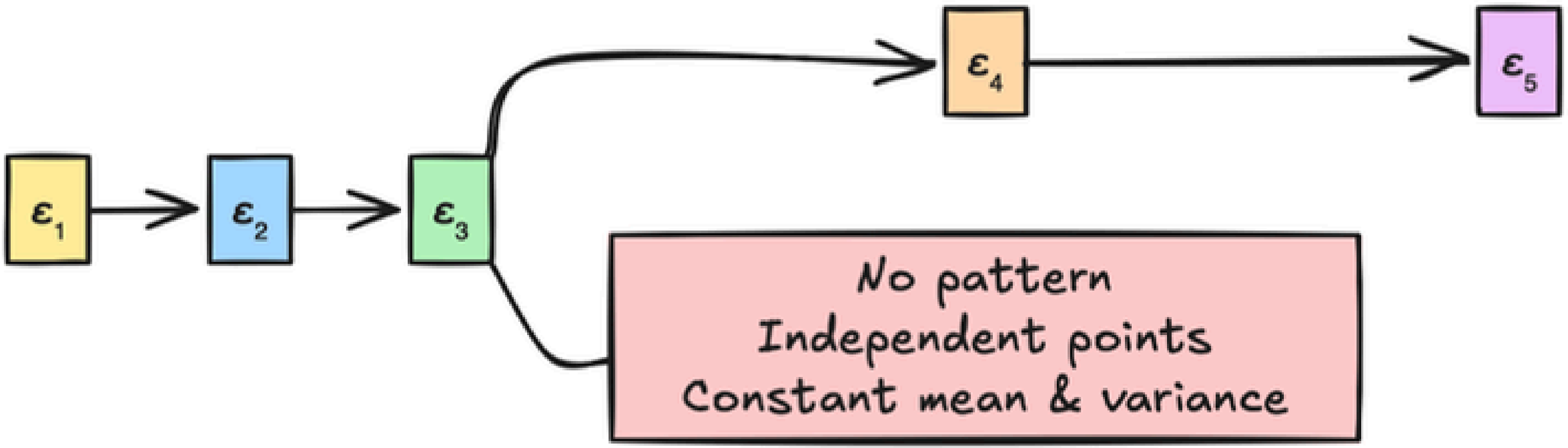
Two types:

- **White Noise:**
  - Random and unpredictable without any patterns or trends.
  - Each data point is independent of others.
  - Constant mean and variance.
- **Structured Noise:**
  - Shows some pattern or structure (could be cycles, trends, or repeating patterns) is present.
  - Mean and Variance may be variable.



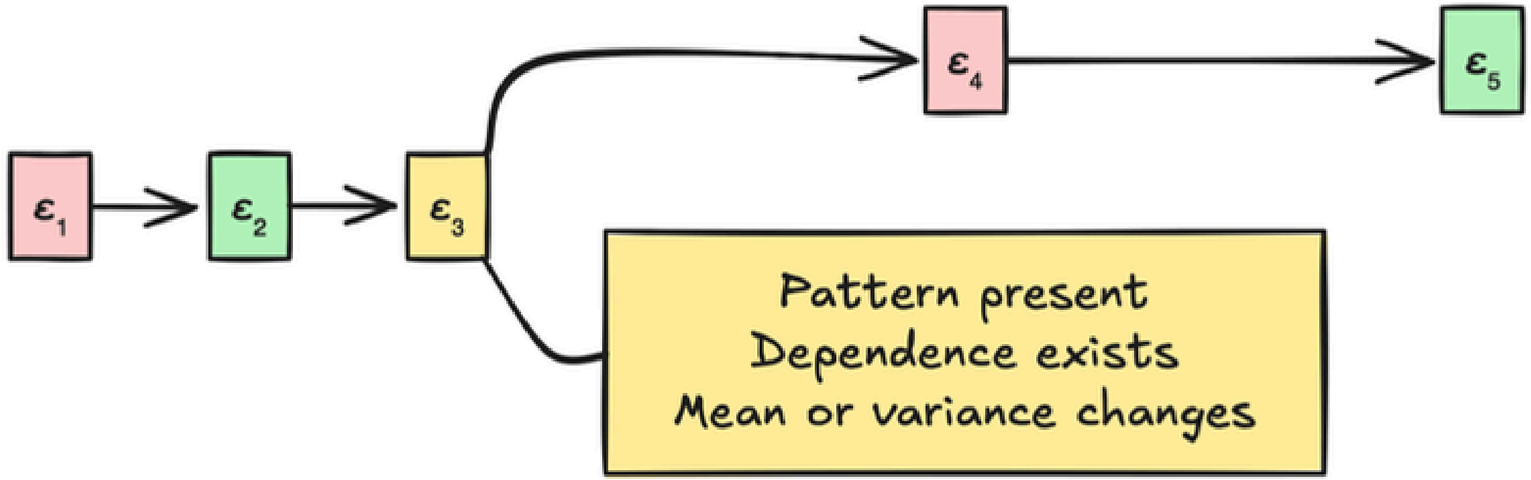
White Noise

Newton School of Technology



Each point is random, unrelated to the previous one, and comes from the same distribution.

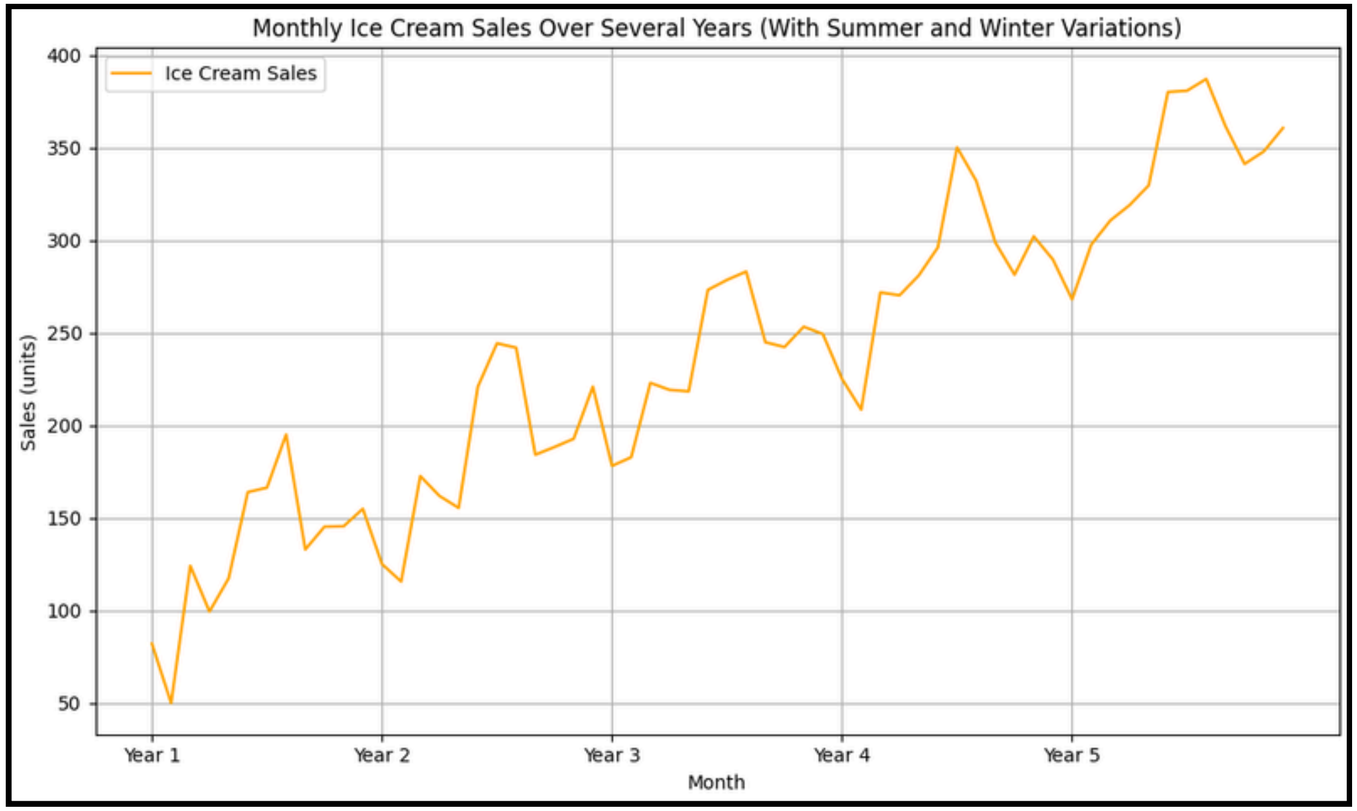
Structured Noise



It looks noisy, but there is hidden structure over time.

Problem

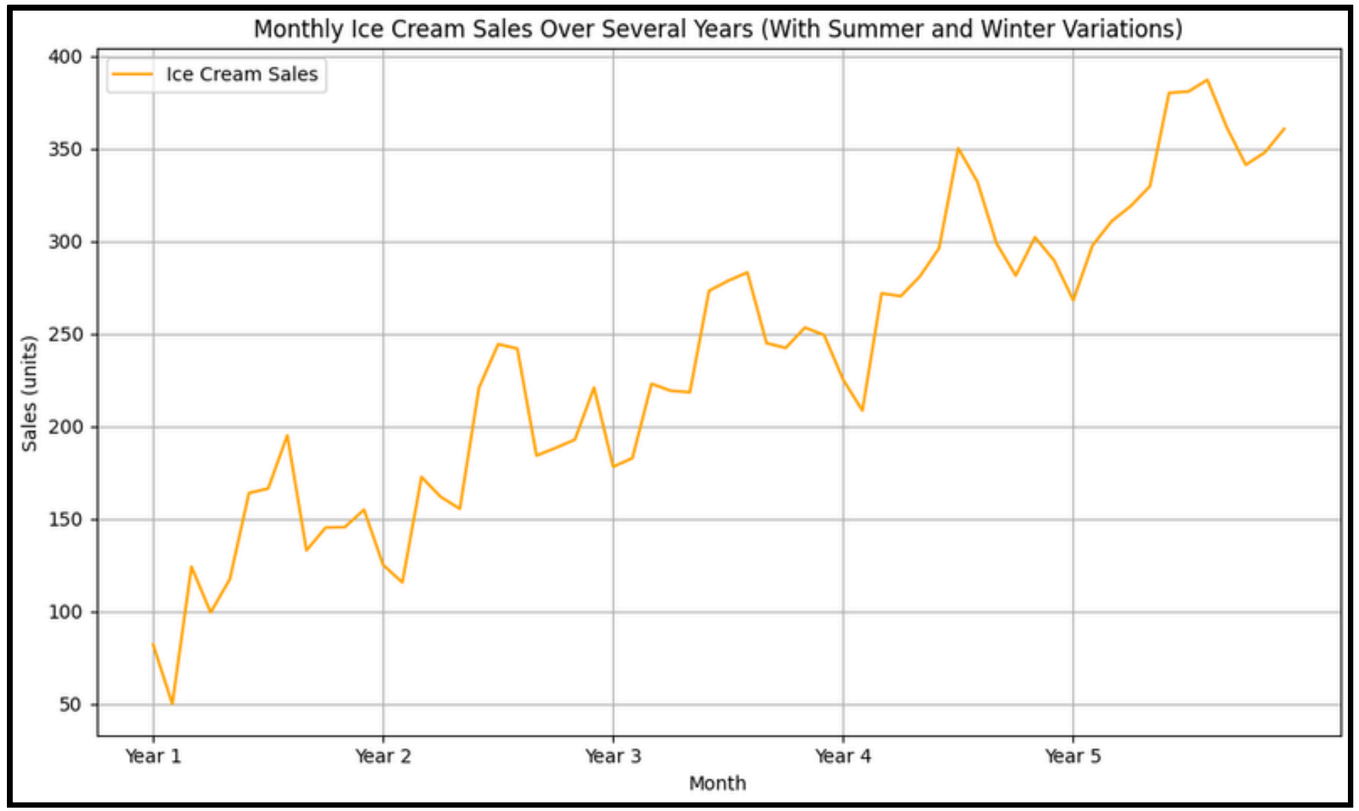
- You are given a line chart showing **monthly ice cream sales** over several years.
- You notice that sales are generally **increasing**, **peak every summer**, and **dip every winter.**
- What **different patterns** do you think are **mixed** in this **single line of data**?



Pattern in the Problems

Newton School of Technology

- Sales seem to be going up over time.
- There’s a repeating pattern every year.
- Some months look unusually high or low.



Decomposition

Newton School of Technology

- Time series decomposition is the process of **breaking a time series into simpler components**, each representing a different **type of pattern** in the **data**.
- Instead of analyzing one complicated line, we separate it into:
  - **Trend** – long-term movement
  - **Seasonality** – repeating patterns
  - **Residuals** – leftover randomness

Types of Decomposition

Newton School of Technology

Additive Decomposition

- The time series is modeled as:

Time Series = Trend + Seasonality + Noise

- Seasonal effect is added to the trend (a fixed amount).
- Used when seasonal variations are constant and do not change with the overall level of the data.

Types of Decomposition

Newton School of Technology

Multiplicative Decomposition:

- The time series is modeled as:

$$\textbf{Time Series} = \textbf{Trend} \times \textbf{Seasonality} \times \textbf{Noise}$$

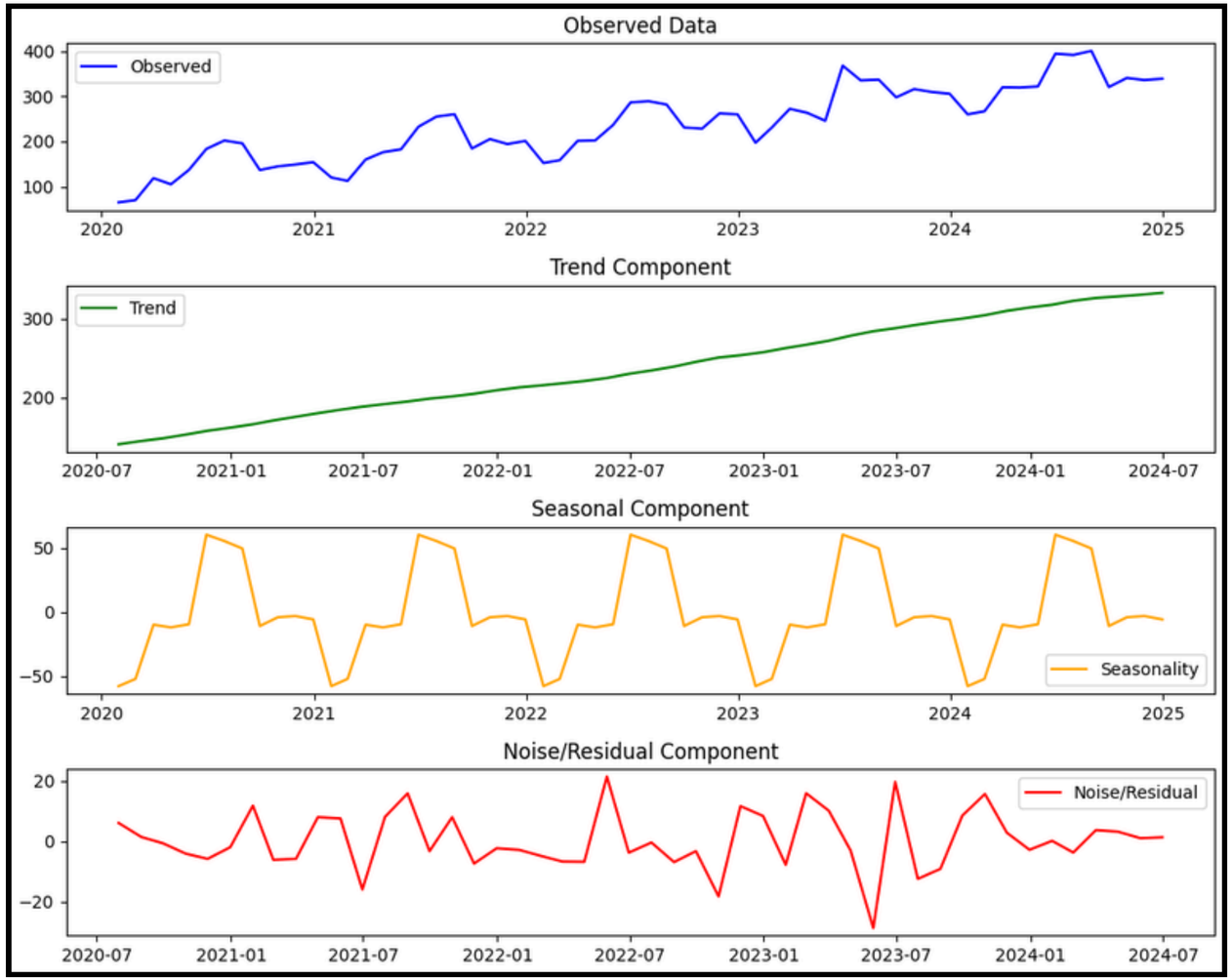
- Seasonal effect is multiplied by the trend.
- Used when seasonal variations are proportional to the trend and increase/decrease as the overall data level changes.

Decomposition Plot

Time series decomposition is usually visualized as a vertical stack of **4** graphs sharing the same time axis.

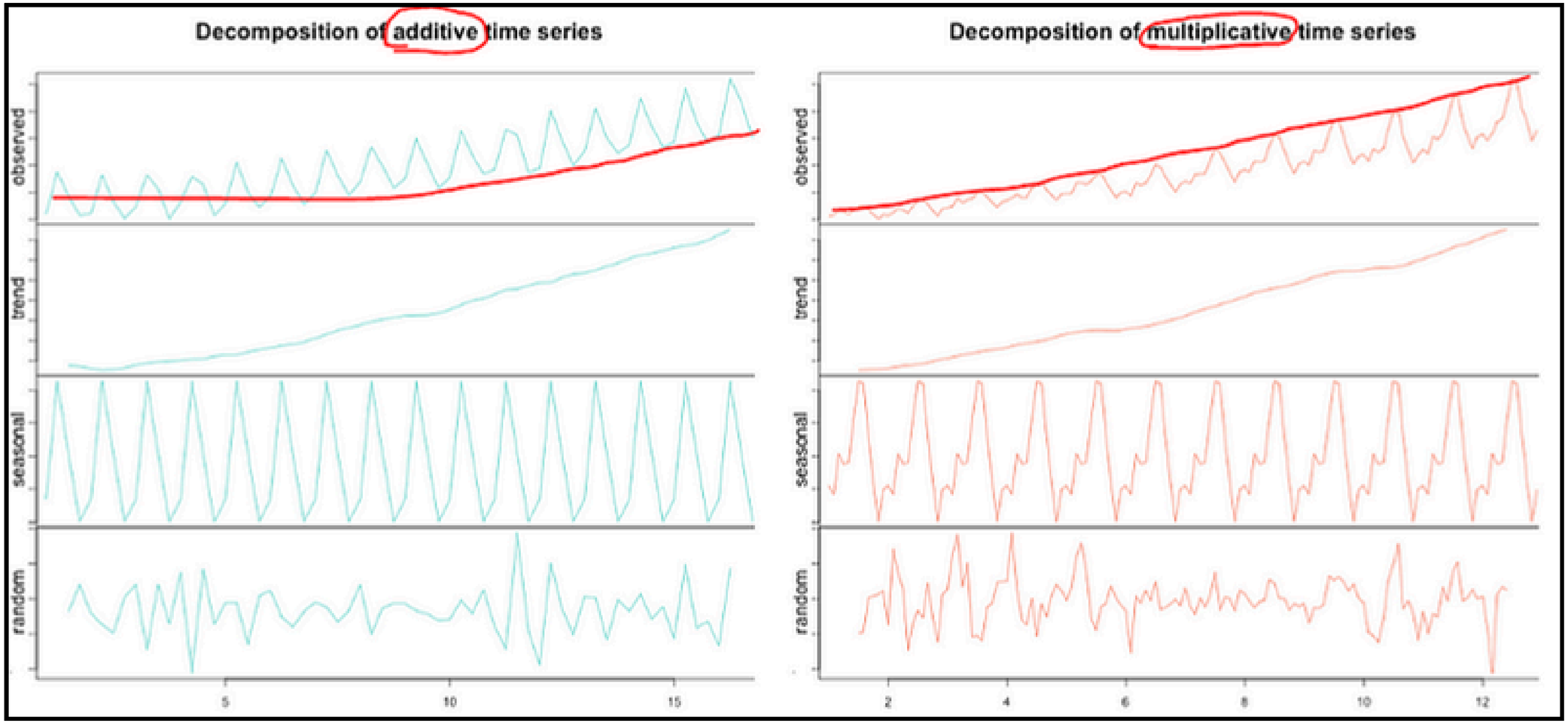
It visually unzips the complex raw data into simple parts.

- Observed
- Trend
- Seasonality
- Residual/Noise



Additive vs Multiplicative Decomposition

Newton School of Technology



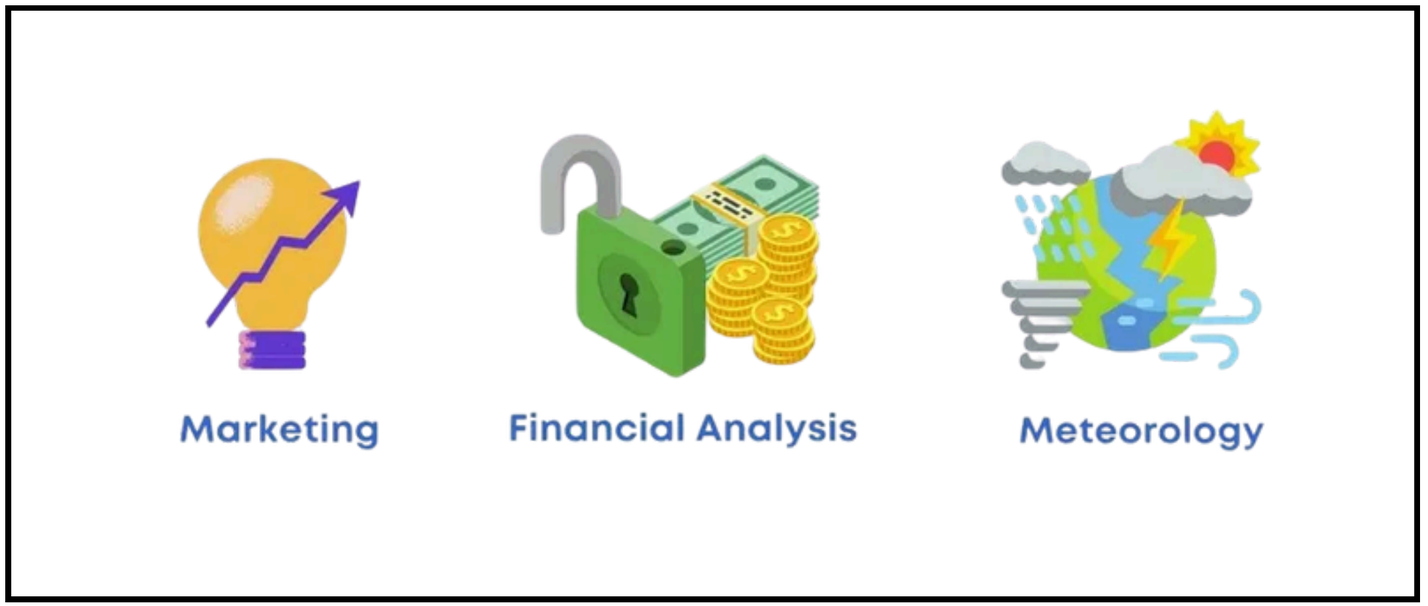
Newton School of Technology

In what other fields are temporal relationships important?

Field where time relation are important

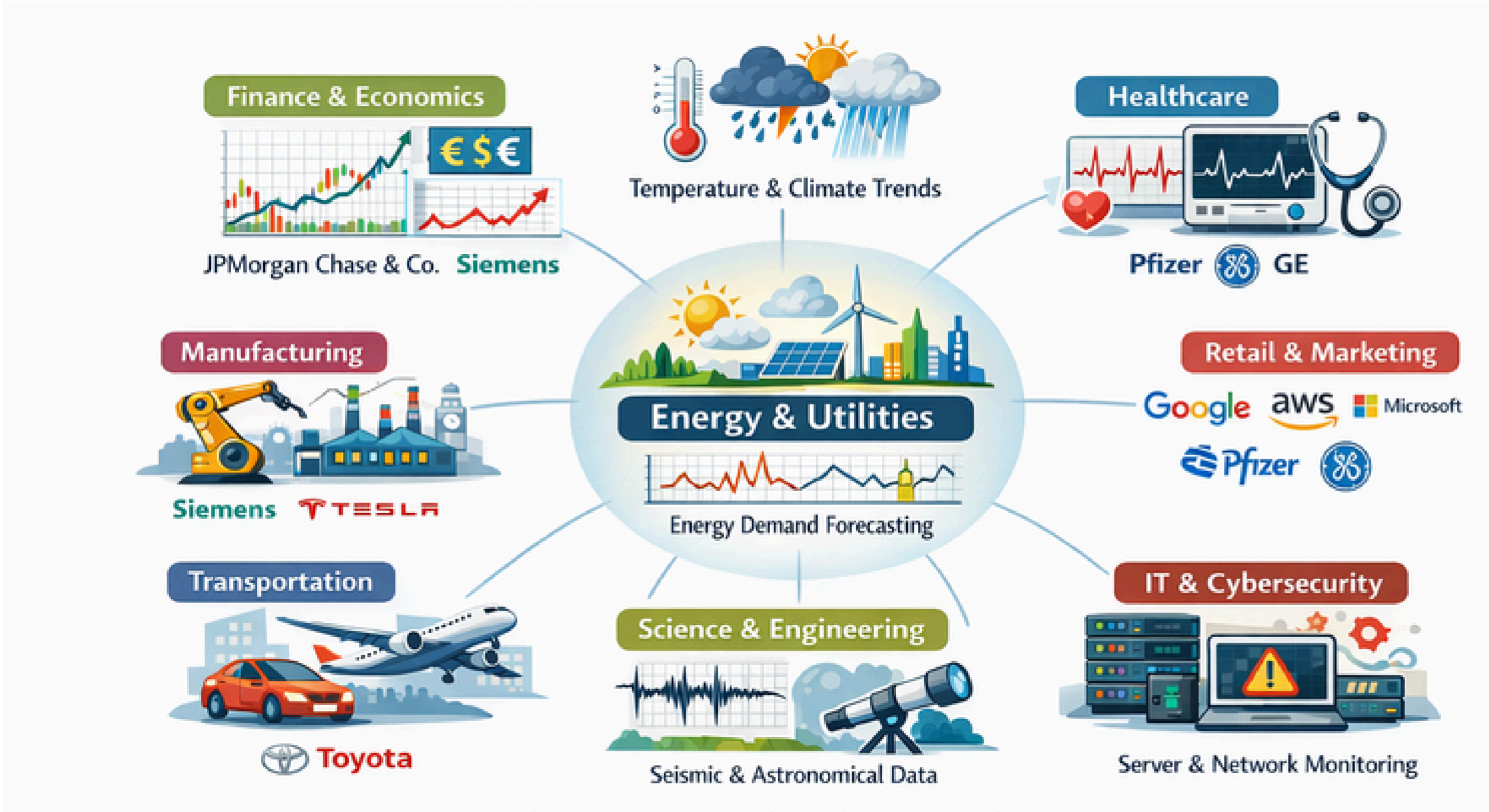
Newton School of Technology

- Stock price prediction.
- Marketing
- Inflation rates over decades
- Music



Field where time relation are important

Newton School of Technology



Summary

Newton School of Technology

- Time series → Data collected over time, where order matters
- Time index → When each observation occurs (date, hour, timestamp)
- Frequency → How often observations are recorded (hourly, daily, monthly)
- Trend → Long-term direction or movement in the data
- Seasonality → Regular, repeating patterns at fixed intervals
- Noise → Random variation not explained by trend or seasonality
- Decomposition → Separating a time series into trend, seasonality, and noise (often visual)

Time series = signal (trend + seasonality) + noise, indexed by time

Case Study/Project

Newton School of Technology

Can we teach an **algorithm** to **predict** a **stock crash** if we **tell** it how **fast** a **virus** is **spreading**?

Actual Volatility vs. COVID-19 Pandemic

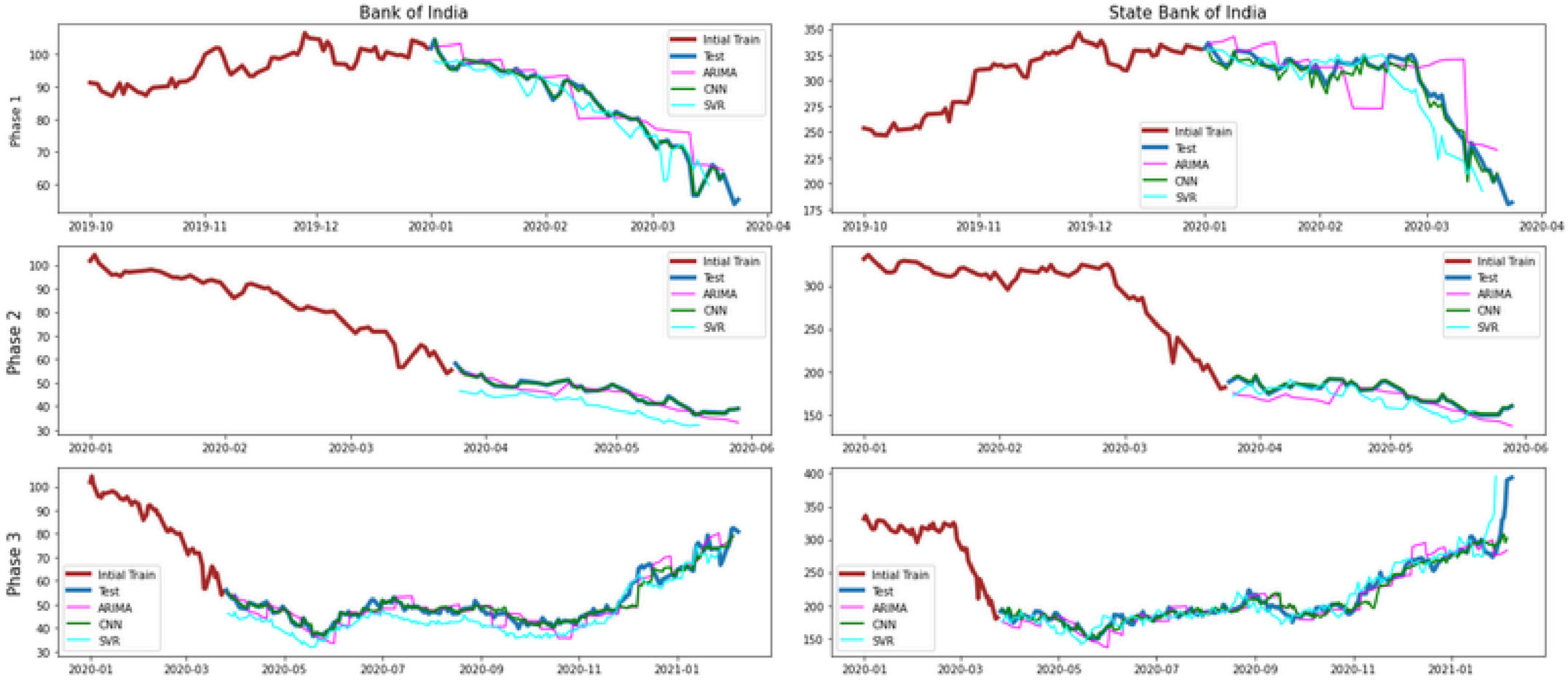
Newton School of Technology

Intervals	P	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
2020-01-02 to 2020-02-12	P1	3.19	40.85	6.85	2.59	7.74	2.20	92.65	1.74	0.50	1.75	9.20	8.74	4.13	13.37	6.51	178.63	45.79	44.02	38.36	19.41	10.92	40.33	14.18	3.00	10.22	82.04	13.43
2020-02-13 to 2020-03-20	P2	15.69	47.95	13.74	17.05	38.85	4.41	170.68	3.95	0.72	16.87	15.15	44.82	9.22	31.23	20.09	584.43	269.43	187.01	187.61	86.52	60.07	125.48	25.49	3.33	18.42	120.19	20.78
2020-03-26 to 2020-05-13	P3	6.31	40.98	4.85	4.41	11.92	0.52	100.35	3.34	0.66	2.58	7.91	8.86	4.27	3.82	5.83	351.02	185.77	153.86	85.05	26.55	33.84	77.69	15.74	3.28	51.21	196.93	78.31
2020-03-14 to 2020-05-29	P4	2.86	9.66	2.20	3.78	23.21	0.48	33.06	1.14	0.16	0.72	2.10	7.46	2.54	4.23	4.27	251.33	74.96	94.47	38.88	18.18	11.19	38.79	4.28	1.14	10.11	33.44	28.05
2020-03-27 to 2020-05-14	P5	6.15	37.11	4.62	4.61	11.08	0.53	97.93	3.23	0.67	2.57	7.66	8.92	3.91	71.75	3.74	351.35	182.52	159.78	86.91	26.41	34.99	71.05	14.42	2.94	48.59	189.36	3.88
2020-05-15 to 2020-06-26	P6	5.80	15.01	4.27	9.30	22.79	3.22	44.47	6.77	0.14	5.36	4.04	12.99	4.54	22.02	4.34	331.90	101.63	125.05	54.10	21.58	22.59	104.06	18.27	7.61	17.27	50.18	11.90
2020-06-29 to 2020-08-07	P7	3.28	17.64	2.22	1.82	8.52	0.68	114.35	2.73	0.31	1.11	5.61	4.76	2.07	30.71	1.68	215.84	75.17	86.50	60.91	87.49	53.29	101.20	24.39	6.89	18.87	173.57	2.30
2020-08-10 to 2020-09-18	P8	3.12	10.68	2.83	7.53	12.68	1.12	45.74	3.79	0.55	7.28	4.88	8.56	1.47	20.66	2.05	209.35	65.05	116.52	79.27	21.38	31.20	61.21	12.41	2.83	12.59	94.86	3.47
2020-09-21 to 2020-11-02	P9	2.06	14.19	1.86	2.82	6.69	0.54	135.87	1.49	0.29	4.95	2.66	7.08	1.13	17.76	1.53	225.86	56.65	149.04	142.13	53.92	25.25	57.21	7.37	2.72	18.21	70.33	2.79
2020-11-03 to 2020-12-15	P10	12.94	28.02	9.41	11.76	16.47	2.46	133.21	4.58	0.29	8.83	8.47	21.02	7.55	17.67	18.27	359.42	144.46	92.20	54.20	26.66	19.11	54.39	13.42	11.05	29.40	191.82	14.03
2020-12-15 to 2021-01-28	P11	7.11	30.97	4.34	3.02	7.32	1.01	101.74	1.39	0.24	7.45	5.71	13.83	6.07	20.48	4.93	263.78	232.92	110.43	178.56	66.06	49.85	30.43	8.37	6.00	16.11	111.69	8.34
2021-01-29 to 2021-02-05	P12	4.33	10.36	4.59	8.44	17.93	0.95	118.42	1.94	0.22	9.62	10.37	17.17	5.14	9.91	6.69	222.65	74.97	181.01	80.18	31.15	27.62	35.28	11.02	6.81	21.94	162.20	62.47
2021-03-10 to 2021-04-27	P13	9.13	8.92	5.37	10.64	44.02	6.93	72.45	8.56	1.19	10.10	11.35	32.85	5.82	7.87	7.30	243.34	172.15	80.03	49.41	30.80	16.83	64.82	18.03	15.17	19.20	98.40	24.95
2021-04-28 to 2021-06-09	P14	5.94	18.30	3.10	6.54	8.31	0.76	231.62	3.94	0.75	12.07	13.96	5.42	2.18	2.79	1.52	201.70	57.40	42.16	57.91	48.37	13.03	24.00	10.03	6.84	7.12	176.47	9.28
2021-06-10 to 2021-07-22	P15	3.07	11.81	2.14	3.52	14.08	1.00	399.39	3.57	2.98	8.23	9.11	10.99	2.83	4.36	3.17	175.25	58.42	53.64	212.21	48.77	66.46	49.28	29.21	16.74	36.23	106.97	19.91
2021-07-23 to 2021-09-03	P16	5.83	24.29	15.77	26.20	43.82	1.81	243.41	1.33	20.78	4.28	9.07	18.66	4.07	12.51	2.99	305.69	77.42	44.46	85.45	26.76	44.75	23.71	9.37	6.02	27.65	148.79	23.84
2021-09-06 to 2021-10-19	P17	2.52	12.70	1.70	6.19	9.47	1.02	112.27	1.15	6.52	8.00	3.26	7.02	2.72	4.01	0.56	105.10	44.51	59.26	63.36	46.04	13.16	10.13	4.99	2.99	6.44	59.23	6.32
2021-10-20 to 2021-10-22	P18	7.72	11.54	4.02	11.94	26.07	1.71	584.11	3.31	80.66	14.98	11.21	21.05	6.11	13.80	4.11	313.28	202.89	100.10	56.56	33.42	6.11	41.27	15.30	7.64	24.62	177.91	21.28
2021-10-26 to 2021-12-08	P19	2.74	6.75	5.04	9.90	11.80	1.70	229.87	2.88	35.07	11.66	9.51	17.65	5.93	10.46	2.11	86.77	89.81	67.43	30.51	38.81	5.93	52.69	3.35	8.29	13.42	77.65	11.83
2021-12-09 to 2021-12-23	P20	6.63	33.31	11.33	14.12	17.68	7.74	233.53	2.25	17.68	13.22	9.69	26.71	9.85	21.94	2.55	454.45	119.85	118.39	91.23	77.26	83.07	21.20	7.99	10.64	27.19	233.11	21.94
2021-12-28 to 2022-02-08	P21	2.65	11.10	2.55	5.07	8.47	2.18	79.13	3.29	27.30	19.56	12.21	14.36	4.77	11.68	1.48	144.50	48.17	46.62	81.98	21.16	12.07	33.71	5.89	9.10	17.92	81.40	11.68
2022-02-09 to 2022-02-23	P22	7.75	16.66	5.37	21.53	41.71	5.07	127.68	2.54	19.29	38.23	4.82	22.06	7.61	12.93	1.76	273.50	181.72	62.49	78.44	62.72	22.90	28.38	20.00	12.83	34.95	138.73	38.25
2022-03-03 to 2022-04-22	P23	2.58	7.34	1.30	3.42	9.23	1.32	126.39	1.51	6.86	3.96	2.69	5.72	1.02	1.98	0.39	200.55	35.38	20.65	63.77	24.98	13.25	32.05	1.31	5.66	9.02	49.60	22.39

- Banking sector experienced moderate levels of volatility during all events
- Each sector and stock tend to fall during the second phase of every crisis but begin to move upwards during the last phase

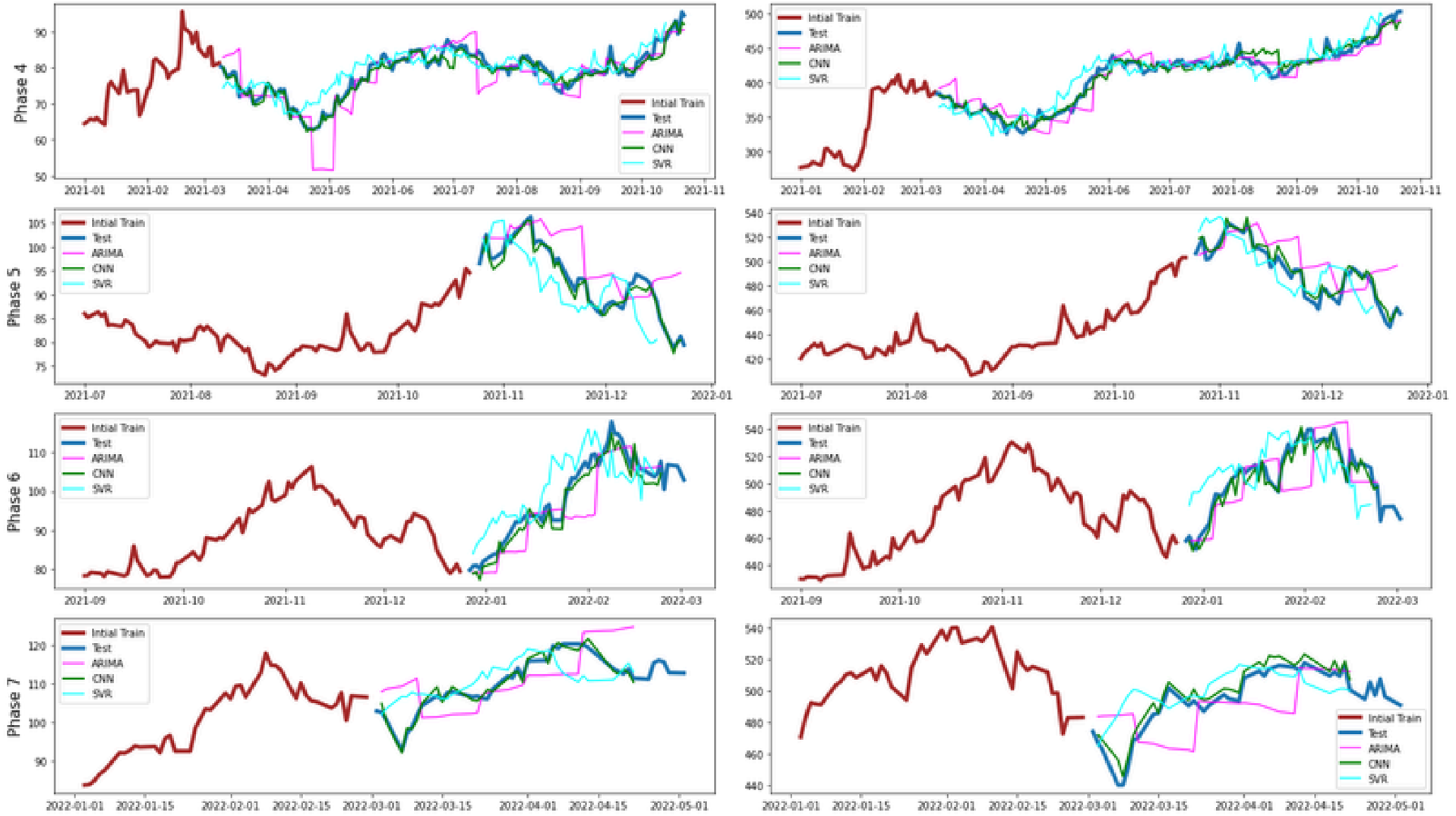
Actual Volatility vs. COVID-19 Pandemic

Newton School of Technology



Actual Volatility vs. COVID-19 Pandemic

Newton School of Technology



Newton School of Technology

Please fill the feedback form.

Newton School of Technology

Thank You

Newton School of Technology

Lecture 2

Unit 1
Time Series

By Sanchit, Rishabh and Shobhit

Newton School of Technology

Join the lecture online on your dashboard.

Newton School of Technology

Let’s start with a minute of silence.

आचार्यात् पादं आधत्ते पादं शिष्यः स्वमेधया ।
पादं सब्रह्मचारिभ्यः पादं कालक्रमेण च ॥

Newton School of Technology

Meaning:

A student acquires knowledge in four equal parts:

- One-fourth from the teacher
- One-fourth through self-reflection and independent thinking
- One-fourth through discussions with peers and fellow learners
- One-fourth over time through personal experience

Recap

Newton School of Technology

- **Time Series**
  - Time series: Data points recorded over time (in time order).
  - What is NOT a time series: Data without a time order (e.g., customer ages, marks list, product prices without dates).
- **Core Concepts / Components**
  - Time Index: The timestamp/date showing when each value occurs.
  - Frequency: How often values are recorded (e.g., hourly, daily, weekly).
  - Trend: The long-term direction of the series (upward/downward/stable).
  - Seasonality: Repeating patterns at fixed intervals (daily/weekly/yearly cycles).
  - Noise: Random variation that cannot be explained by pattern/trend.
- **Decomposition:** Breaking the time series into parts (Trend + Seasonality + Noise) to visualize and understand it better.

Agenda

- Stationarity vs. Non-Stationarity
- Differencing Transformations
- ACF (Autocorrelation Function)
- PACF (Partial Autocorrelation Function)

Newton School of Technology

A Curious Question Before We Begin...

Newton School of Technology

- **Case 1:** Imagine you measure the **daily temperature** in your **city** for a **year**.
- **Case 2:** Now imagine you also measure the **daily temperature inside** your **fridge** for a **year**.

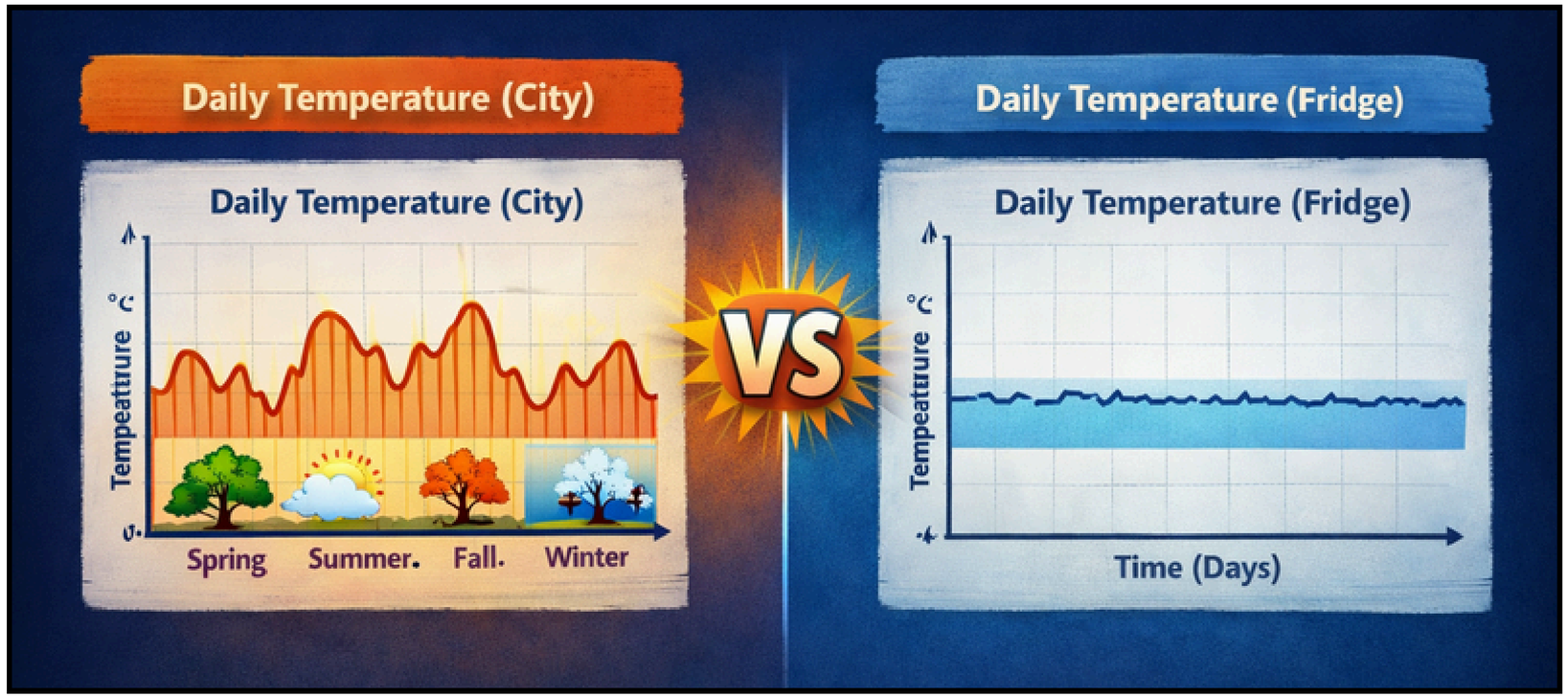
Which one would be easier to predict using past data, and why?

A Curious Question Before We Begin...

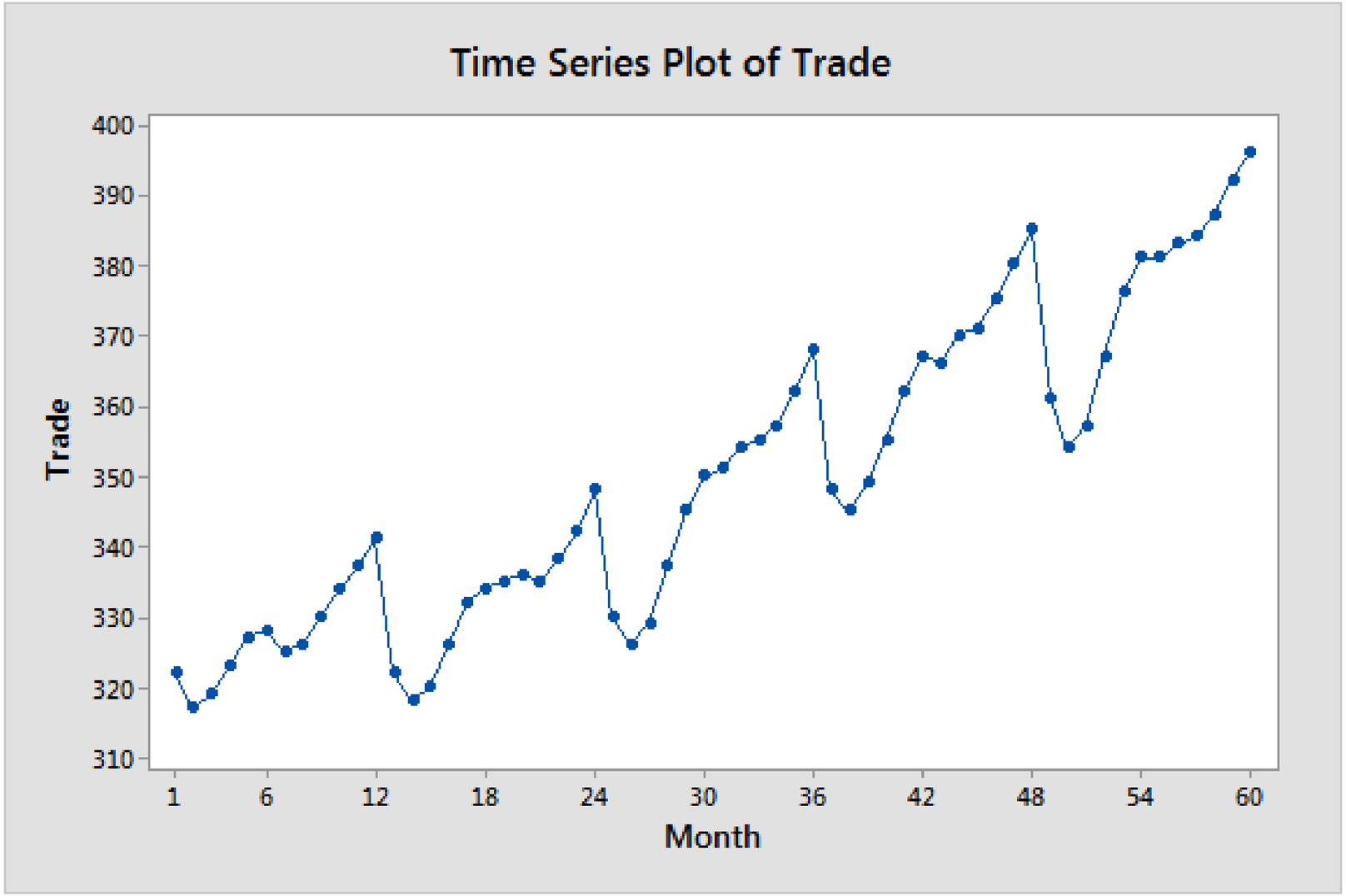
Newton School of Technology

- **Case 1:** Imagine you measure the **daily temperature** in your **city** for a **year**.
- **Case 2:** Now imagine you also measure the **daily temperature inside** your **fridge** for a **year**.

Which one would be easier to predict using past data, and why?



- **Fridge temperature**: fluctuates, but around a constant level.
- **City temperature**: changes across different months, day and night.

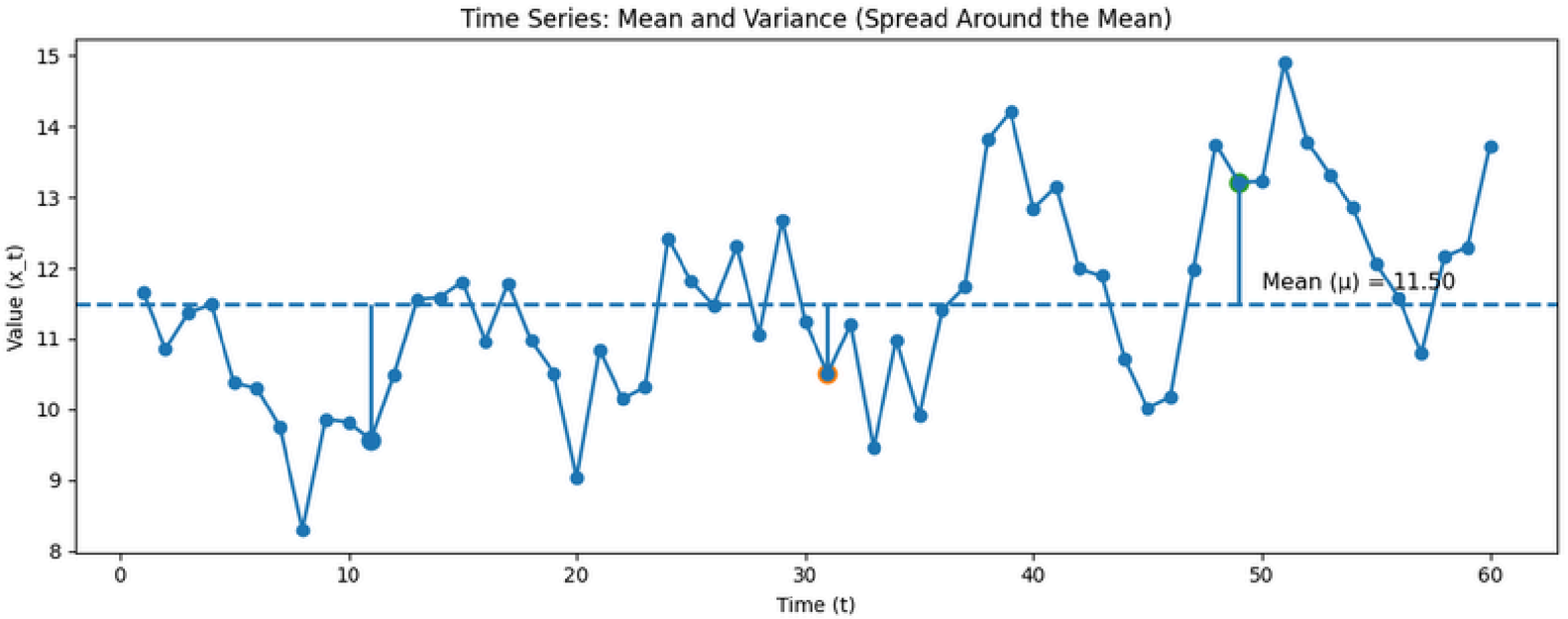


Newton School of Technology

What is mean, variance and correlation in timeseries?

Mean and Variance in Time Series

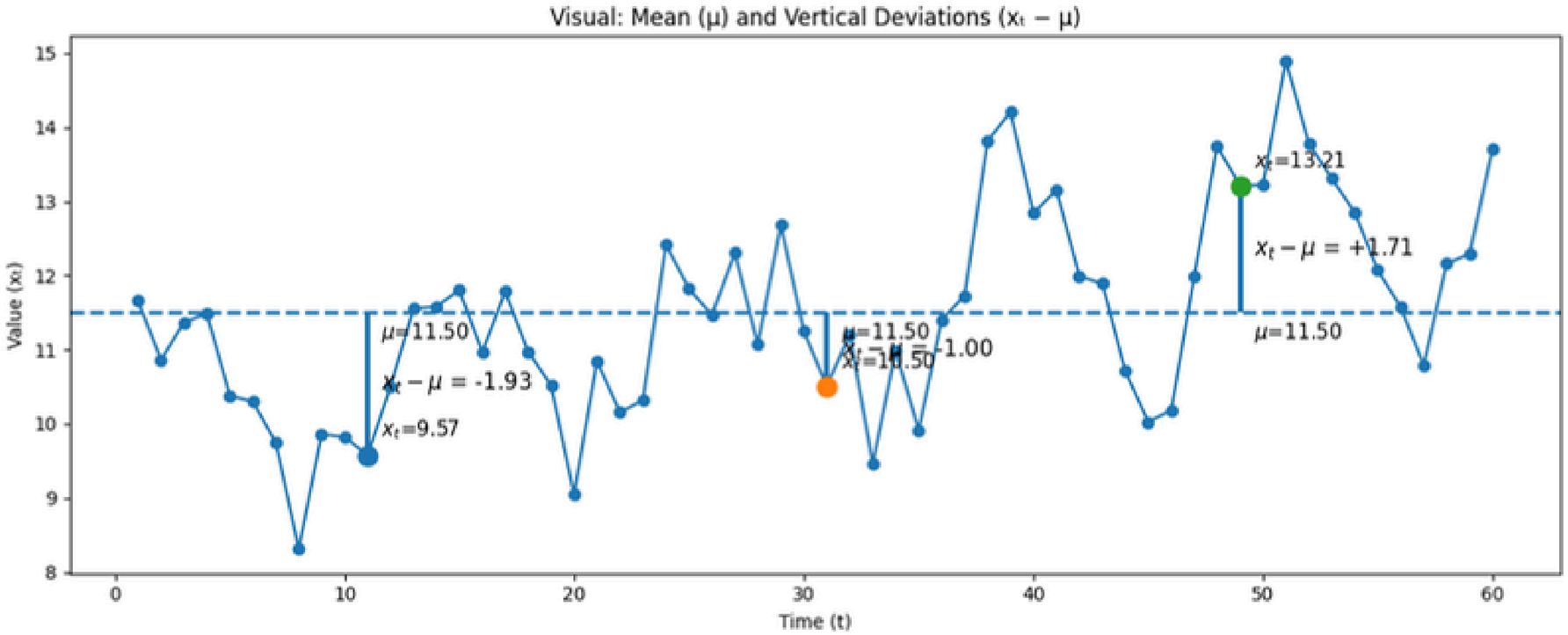
Newton School of Technology



- Dashed horizontal line = Mean (μ)
- Vertical lines from mean to points = deviation from mean (these deviations create variance)

Mean and Variance in Time Series

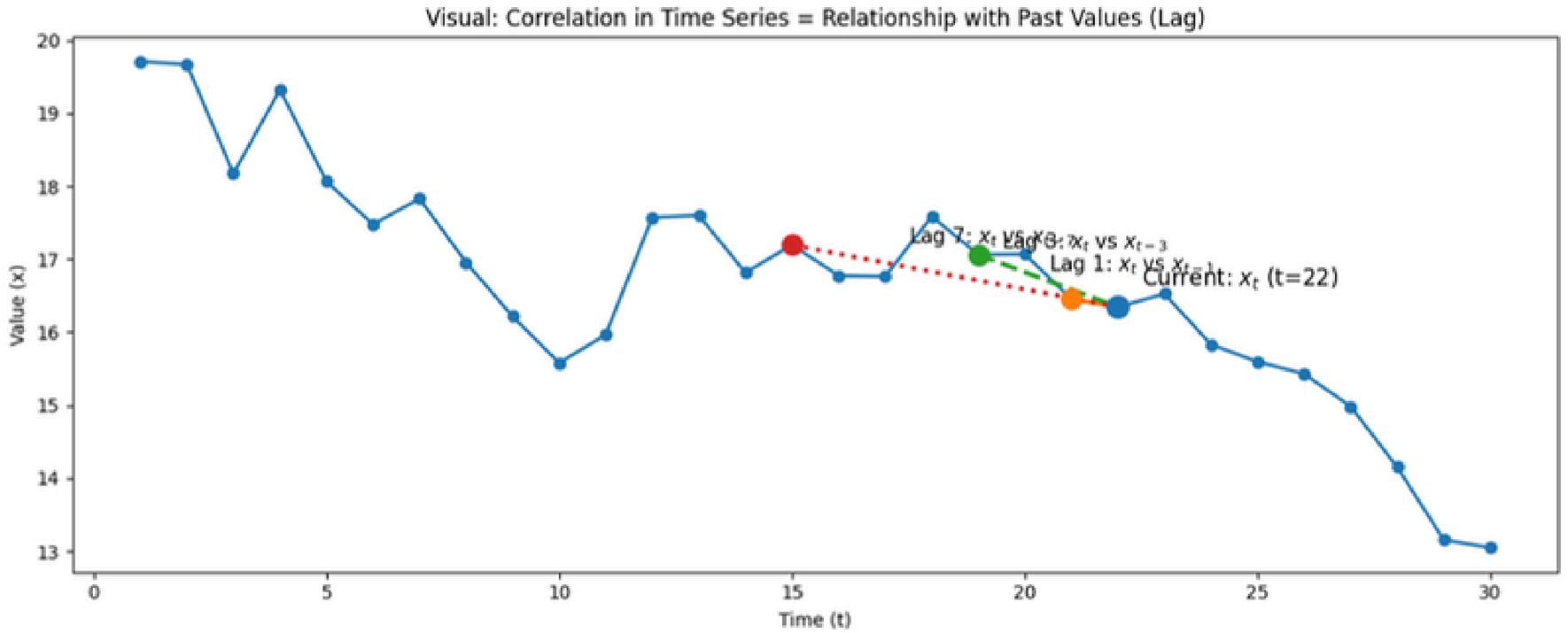
Newton School of Technology



- Mean line **(μ)** = “average level” of the series
- A point **x(t)** = actual value at time t
- Vertical line length = **x(t)-μ**
- This is called **deviation from mean.**

Correlation in Time Series

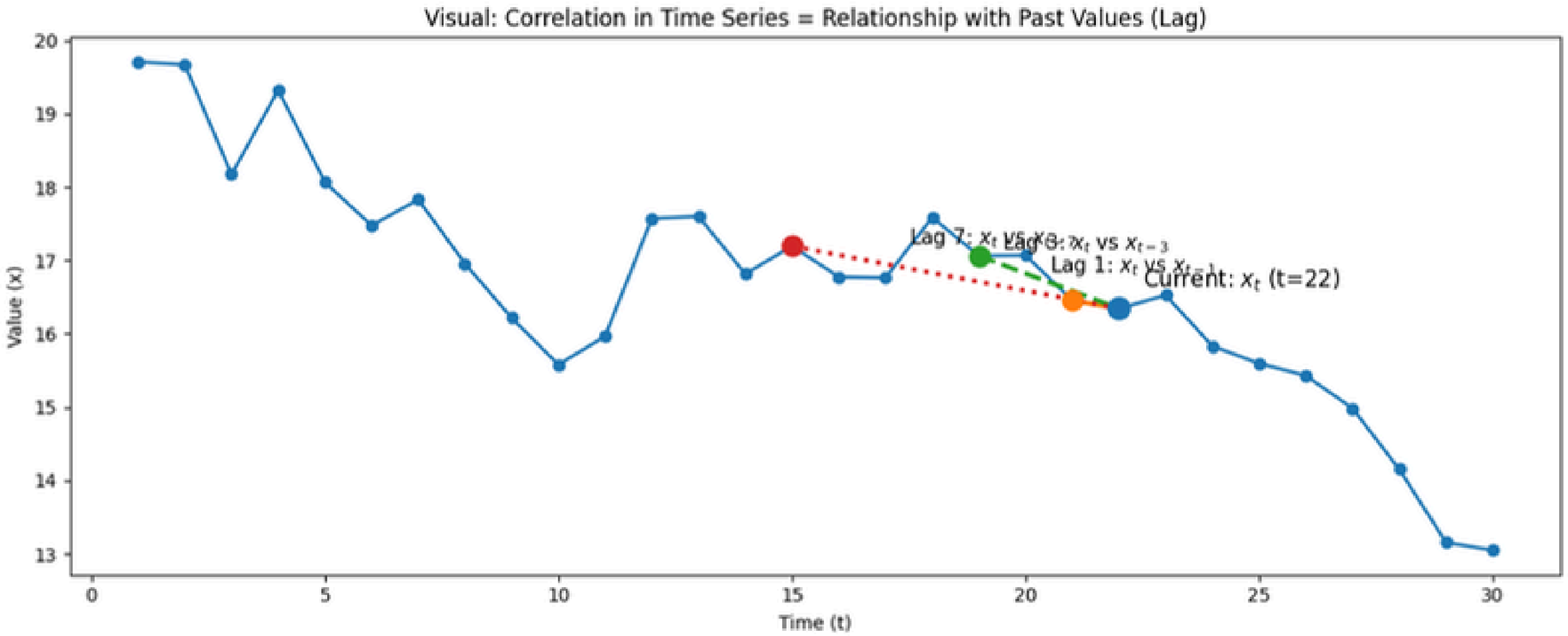
Newton School of Technology



- What is correlation? How strongly **two values** move together.
- In time series, we don't compare two different **variables,** we compare the same **variable** at **two different times.**

Correlation in Time Series

Newton School of Technology

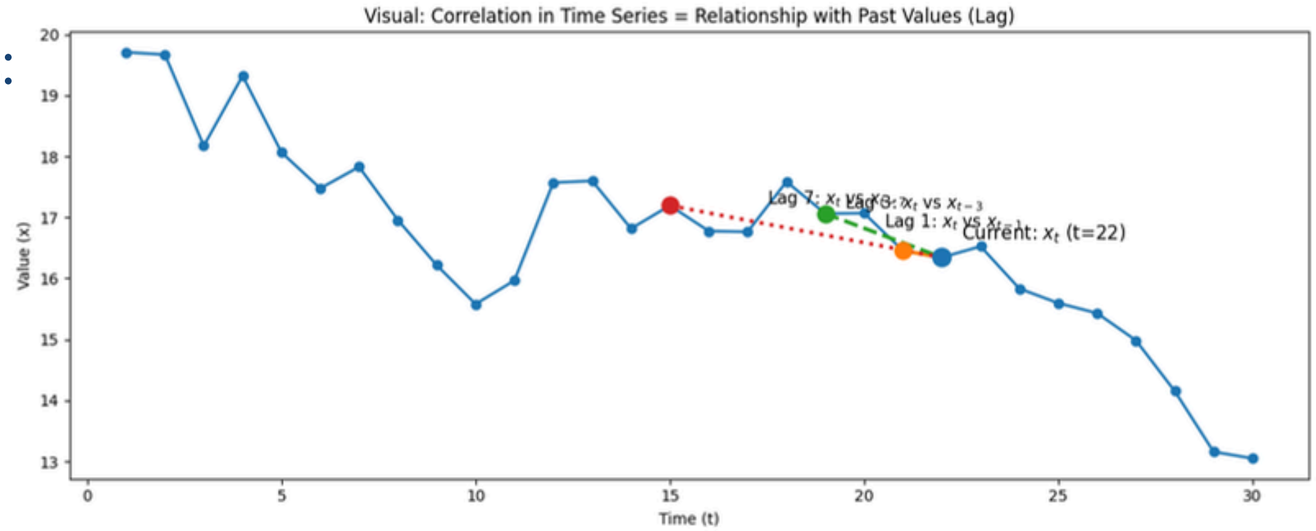


- How to calculate correlation in timeseries?
- **We will cover this in depth in upcoming slides.**

Correlation in Time Series

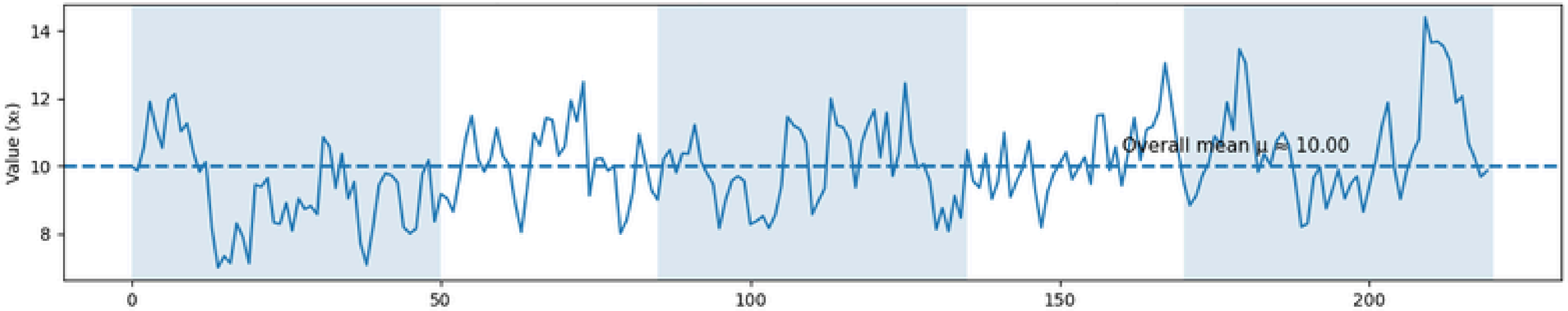
Newton School of Technology

- How to calculate correlation in timeseries? with **Lags**
- The **blue dot** is the **current value $x_{(t)}$**
- The connected dots show past values:
  - **Lag 1:** compare $x_{(t)}$ with $x_{(t-1)}$
  - **Lag 3:** compare $x_{(t)}$ with $x_{(t-3)}$
  - **Lag 7:** compare $x_{(t)}$ with $x_{(t-7)}$
  - **Lag K:** compare $x_{(t)}$ with $x_{(t-k)}$



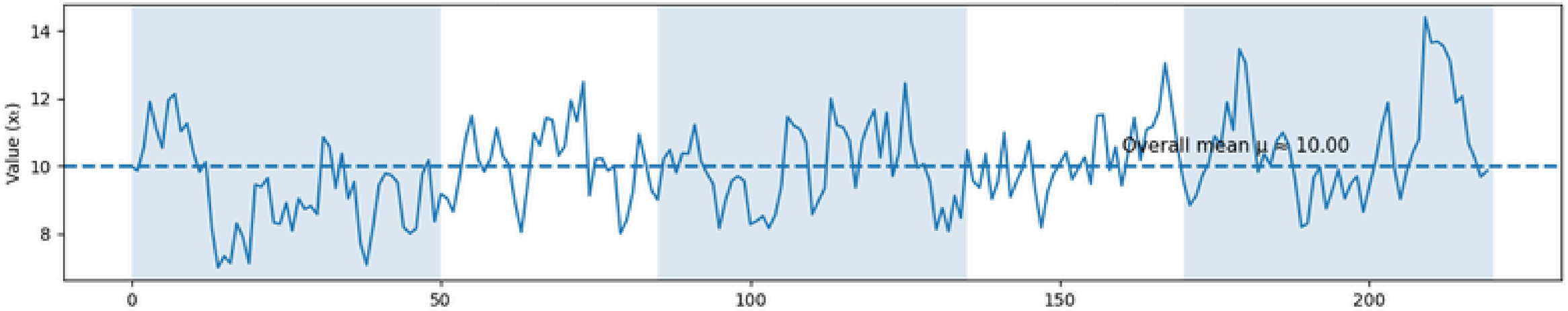
If points that are connected (lagged points) are usually close / move similarly, the correlation is high → past helps prediction.

Newton School of Technology



What is your **observeration** about this **timeseries**?

Newton School of Technology



What is your **observeration** about this **timeseries specfically around mean and variance?**?

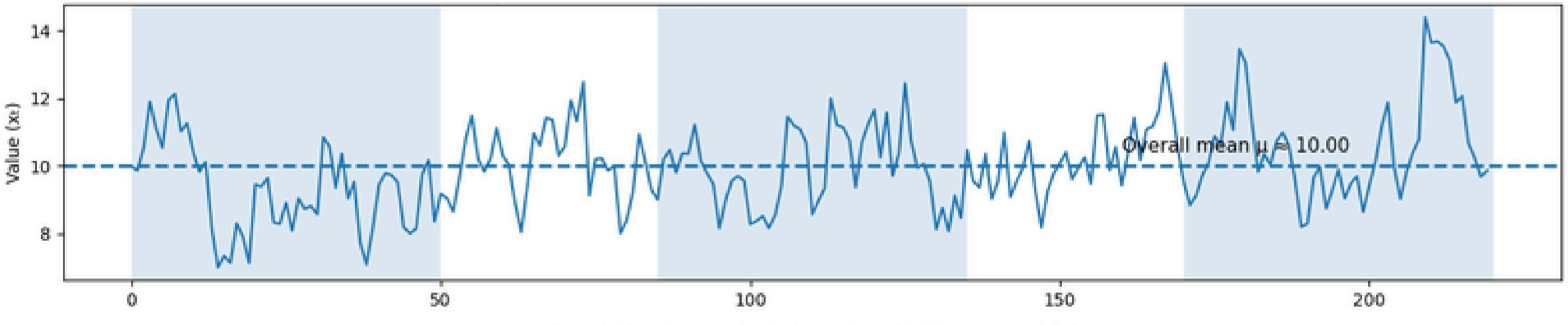
Stationarity

Newton School of Technology

- A time series is stationary if its **behavior** does not **change over time**.
- Which means if its **major statistical properties (Mean, variance, etc.)** remain constant over time.
- A time series is stationary when:
  - **Mean** is constant (average level stays same)
  - **Variance** is constant (spread/volatility stays same)
  - **Correlation** depends only on **lag**

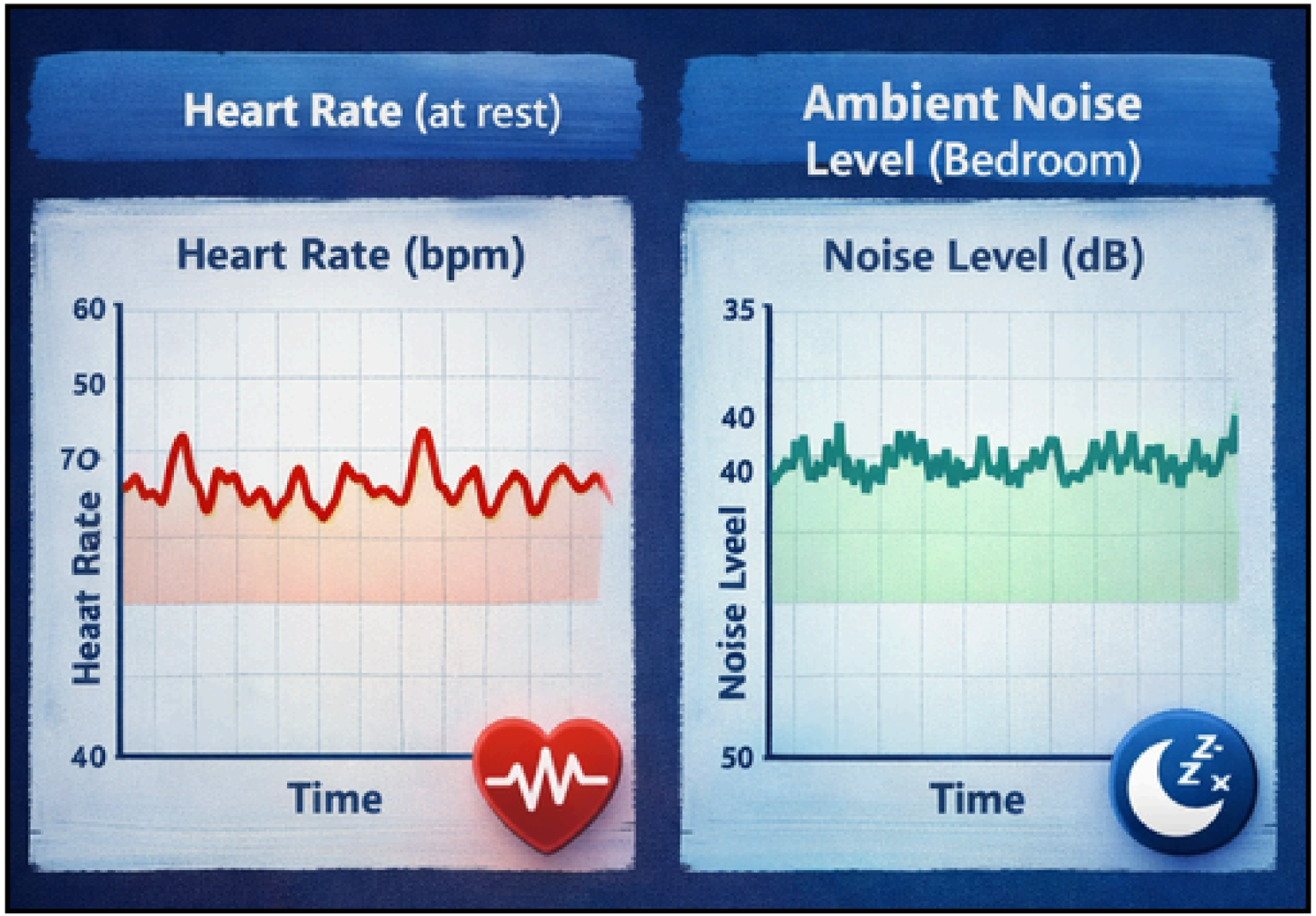
Is this time series Stationary?

Newton School of Technology



Real life examples of Stationary Time Series Data

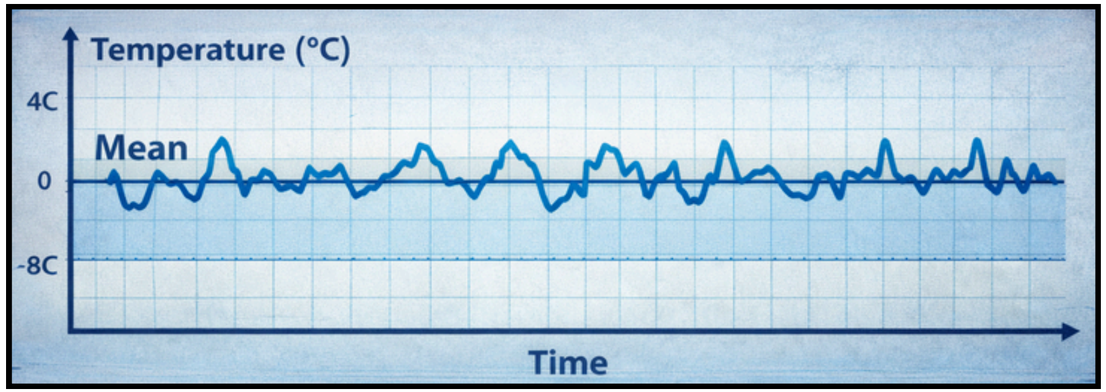
Newton School of Technology



Mean Stability

Newton School of Technology

A stationary series should “hover” around the **same average** level at all times this is called as **Mean Stability**.



Mean Stability

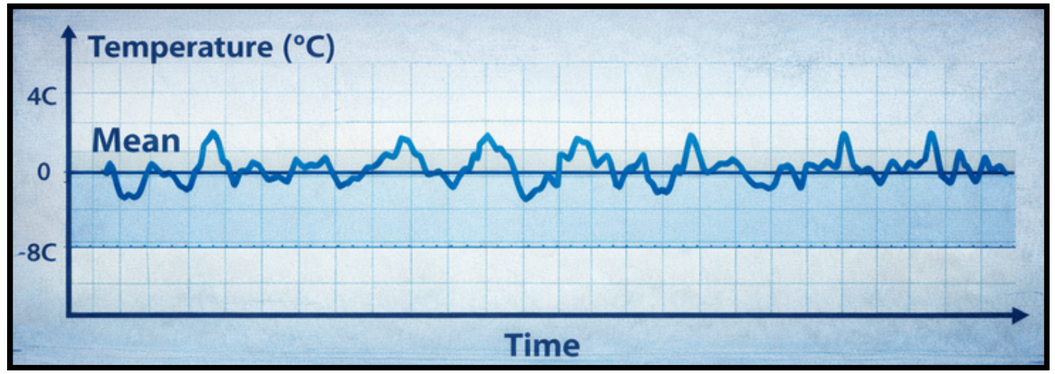
Newton School of Technology

Think about a fridge:

- It is set to a fixed cold temperature.
- Even if it fluctuates a little when you open the door,
- It always returns to roughly the same level.

The average temperature stays the same every day.

This means the **fridge temperature** has a **mean stability.**



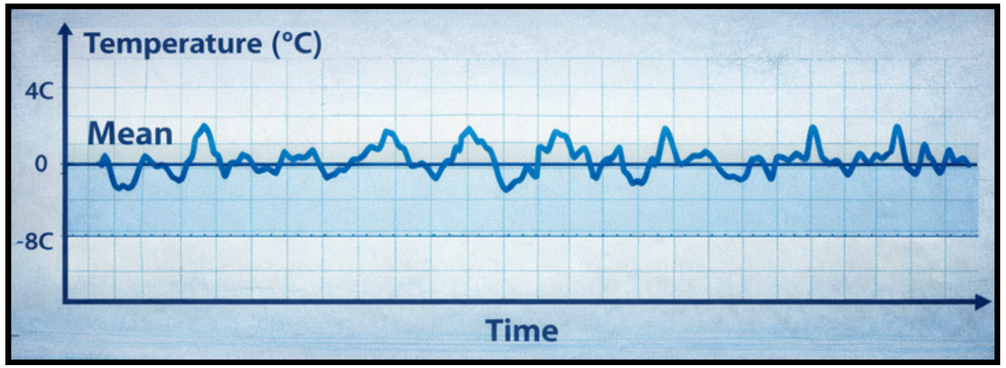
Variance Stability (Constant Variance)

Newton School of Technology

A stationary series keeps its spread or variance **consistent.**

The variance stays the same every day.

This means the fridge temperature also has **variance stability.**

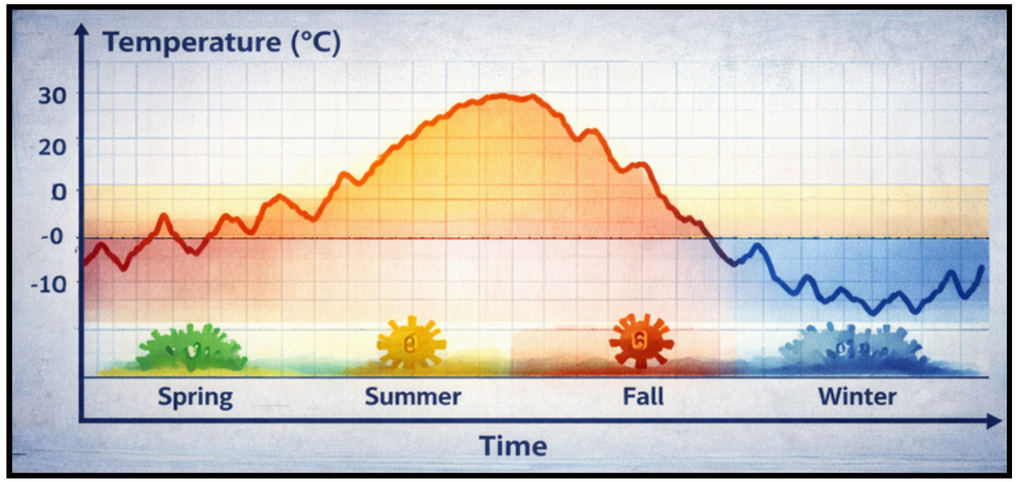


Examples of Non-Mean Stability

Newton School of Technology

Now look at outdoor temperature:

- In summer, the average temperature is higher.
- In winter, the average temperature is lower.
- The typical temperature changes through seasons.



The average does **not stay the same**.

The **outdoor temperature** does not **have mean stability.**

Example of Non-Variance Stability

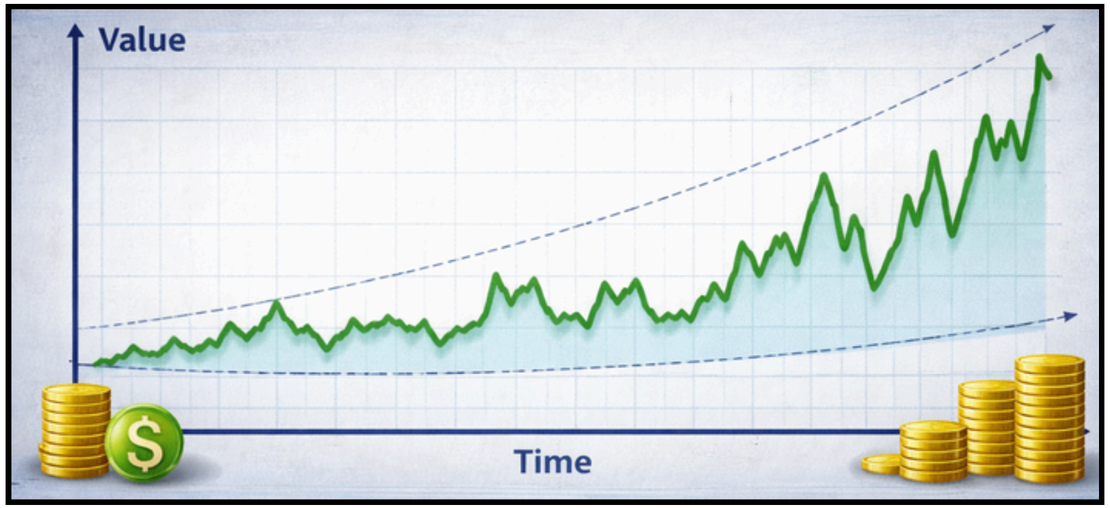
Newton School of Technology

Non-variance stability patterns include:

- increasing volatility
- sudden jumps
- variance that expands over time

Example:

Financial markets often show this: small fluctuations early, large swings later.

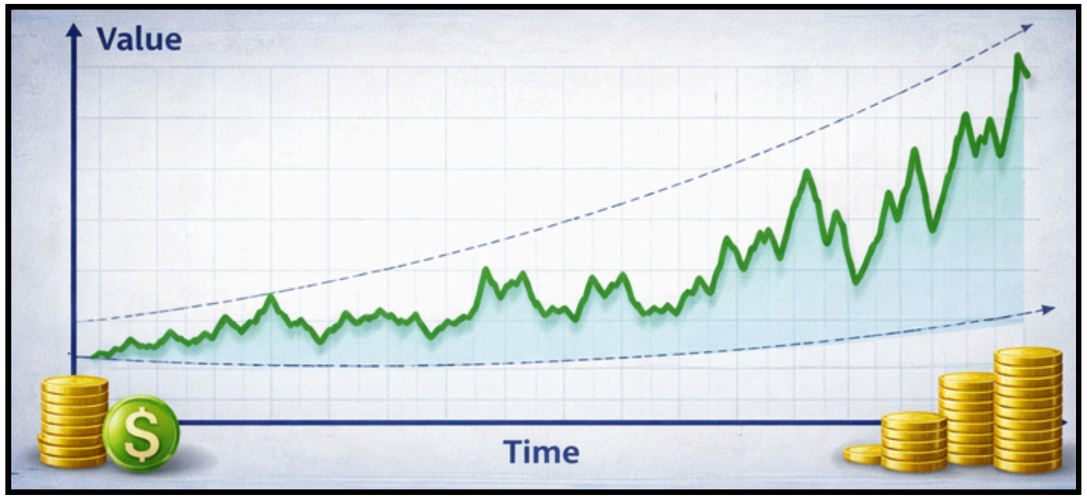


Non Stationary Data

Newton School of Technology

Non-stationary data patterns include:

- upward trend
- downward trend
- structural breaks
- seasonality

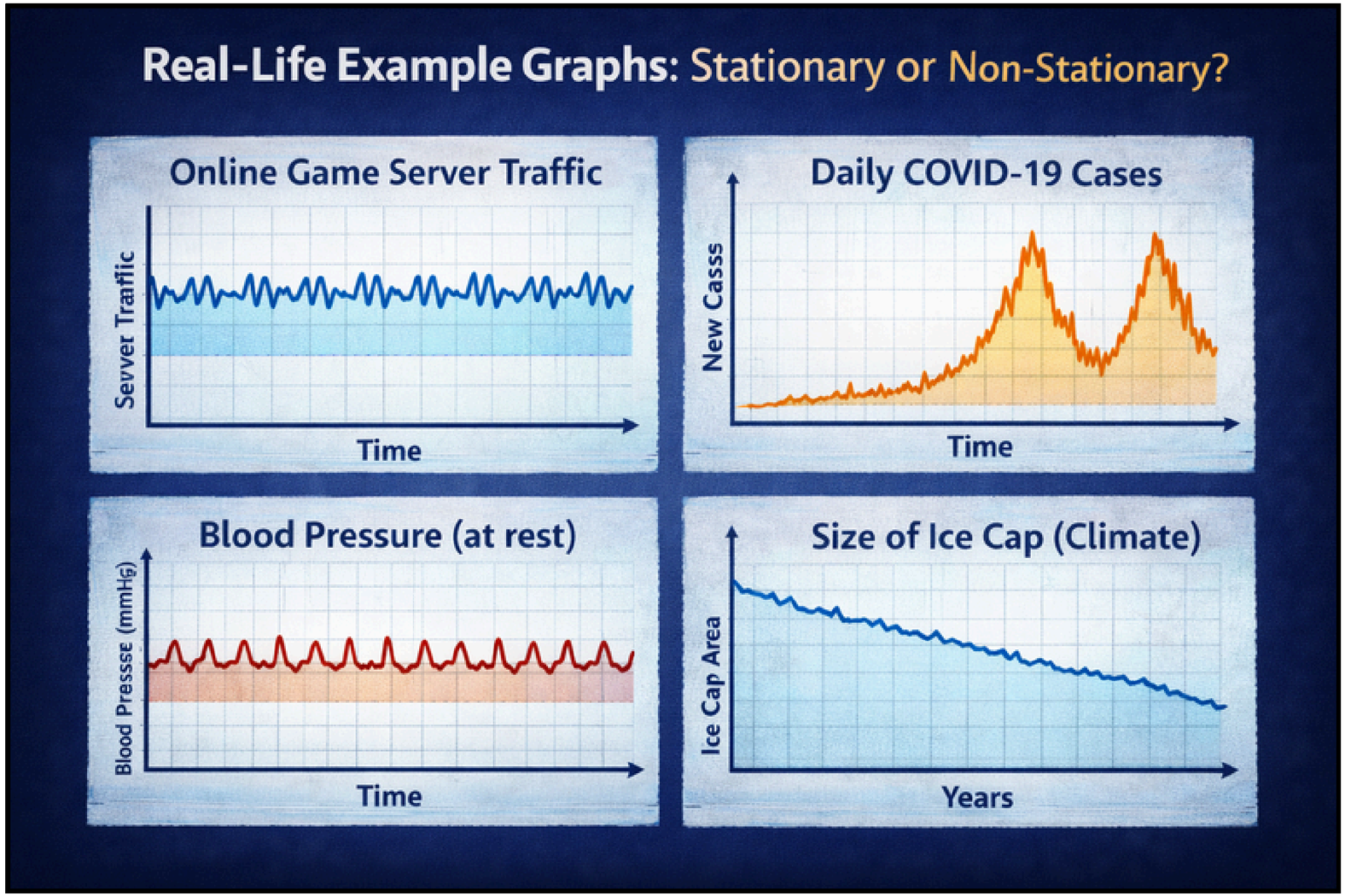


Example:

Monthly sales are steadily rising → the mean is changing.

Identify the following graphs as Stationary or Non-Stationary

Newton School of Technology



Why Stationarity Matters?

Newton School of Technology

Few Reasons why Stationarity matters in time series data:

- Past helps predict the future
- Consistent behavior
- Reliable predictions
- Simpler analysis

Stationary vs Non-Stationary Data

Newton School of Technology

Aspect	Stationary	Non-Stationary
Mean	Constant	Time-dependent
Variance	Constant	Changes
Modeling	Easier	Harder
Forecasting	Stable	Risky
Real-world realism	Lower	Higher
Statistical tests	Valid	Often invalid

Newton School of Technology

Problem Statement

If a restaurant received **X Zomato orders today**, can we predict **tomorrow's orders** using just that **single number**?

Or do we need to know what happened **yesterday**, the **day before**, or **last week**?

Newton School of Technology

- **Food orders usually don’t behave randomly. There are predictable effects:**
  - If **Monday** is busy → **Tuesday** is generally similar.
  - **Weekend** orders are consistently higher.
  - **Rainy** days often **increase orders** for 2–3 days.
  - A **discount** running today affects **orders** tomorrow.
- **Today’s order volume is related to yesterday’s order volume.**

Day	Orders	Contextual Factor
Monday	240	Regular weekday
Tuesday	255	Exam season begins
Wednesday	265	Exam season continues
Thursday	260	Light rain
Friday	275	Rain + pre-weekend

Lag Dependence (How Far Back the Relationship Goes)

Newton School of Technology

- **Lag means:** how many time-steps back we look.
  - So if today is Day T, then lag tells us which previous day/value we are using.
- If we are looking at today's value:
  - **Lag 1 → yesterday**
  - **Lag 2 → day before yesterday**
  - **Lag 7 → same day last week**
  - **Lag 30 → same day last month**
- **Basically: Lag = past value used to understand today.**

Day	Orders	Orders (Lag 1)
Monday	240	
Tuesday	255	240
Wednesday	265	255
Thursday	260	265
Friday	275	260

Lag Dependence (How Far Back the Relationship Goes)

Newton School of Technology

Let the time series be: y_t

Where:

- t= Current time
- y_t = value at time t

Then, **Lag-k** value is: y_{t-k}
meaning the value at time k units before time t i.e. value at time t-k.

- **Think of a lag as a conversation between two time points.**
  - “What happened today?”
  - “What happened one step before?”
- If today and yesterday are similar → strong lag-1 relationship.
- If today and **7 days** ago are **similar → strong lag-7 relationship**.

Lag Dependence

Newton School of Technology

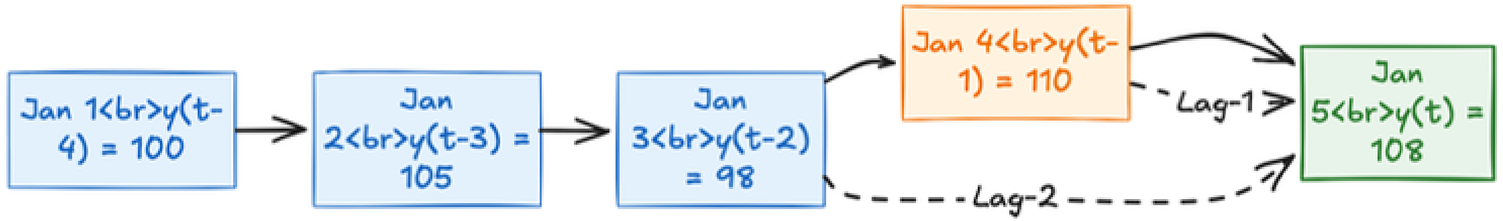
Day	Orders	Lag 1	Lag 2	Lag 7	Comments
Mon (Week 1)	220	—	—	—	**Start**
Tue (Week 1)	**230**	**220**	—	—	**Slight weekday rise**
Wed (Week 1)	235	230	220	—	
Thu (Week 1)	240	235	230	—	
Fri (Week 1)	**300**	240	**235**	—	**Weekend spike**
Sat (Week 1)	320	300	240	—	
Sun (Week 1)	310	320	300	—	
Mon (Week 2)	**225**	310	320	**220**	**Weekly pattern**
Tue (Week 2)	235	225	310	230	

Newton School of Technology

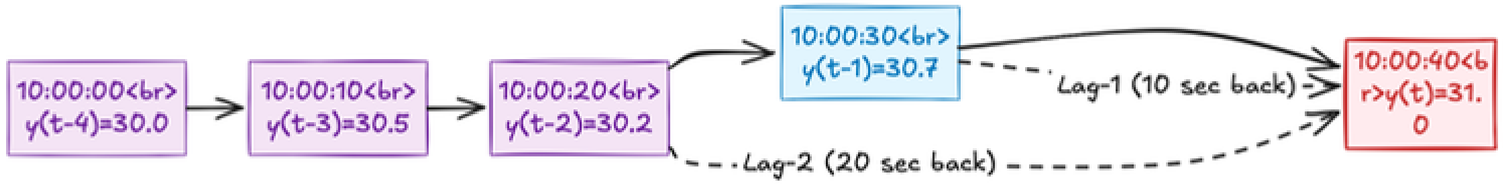
Lag = “how many steps back” (step size depends on granularity)

Data Type	1 step =	Lag 1 means
Second-wise	1 second	1 second ago
Minute-wise	1 minute	1 minute ago
Hourly data	1 hour	1 hour ago
Daily data	1 day	1 day ago

Lag for Daily Data (Lag-1, Lag-2)



Lag for Time Data (every 10 seconds)



Newton School of Technology

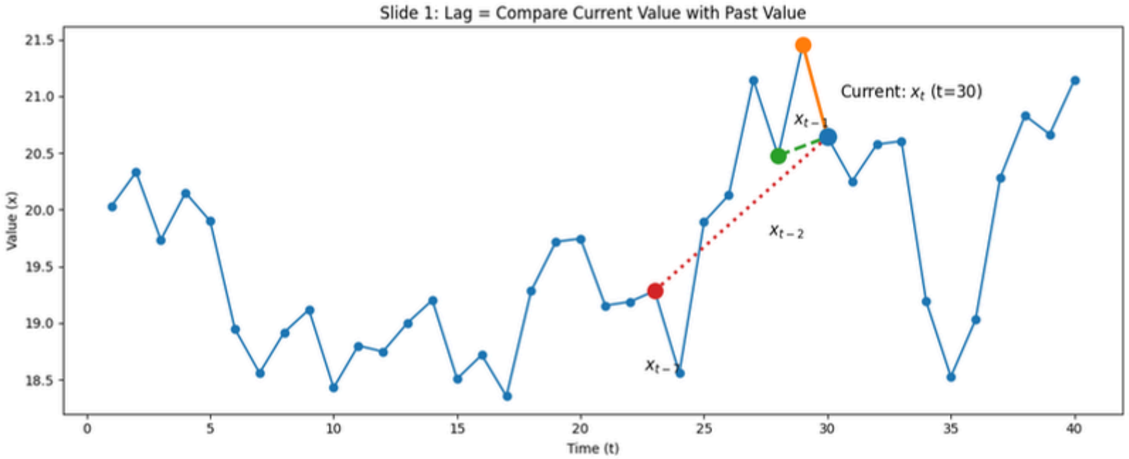
How do we choose useful lags?

Use lags based on **domain pattern**:

- **Daily series:**
  - Lag-1 (yesterday effect)
  - Lag-7 (weekly seasonality)
- **Hourly series:**
  - Lag-1 (last hour)
  - Lag-24 (same hour yesterday)
- **15-min series:**
  - Lag-1 (last 15-min)
  - Lag-96 (same time yesterday)

Practice rule:

Pick lags where **real-world cycles exist.**



Problem Statement

Newton School of Technology

Is this data stationary?

Time (t)	Value
1	5
2	7
3	9
4	11
5	13
6	15
7	17

Problem Statement

Newton School of Technology

This data is **not stationary**, as there is a clear upward trend in the data with respect to time.

Time (t)	Value
1	5
2	7
3	9
4	11
5	13
6	15
7	17

Problem Statement

Newton School of Technology

How would you convert this data into stationary data?

Time (t)	Value
1	5
2	7
3	9
4	11
5	13
6	15
7	17

Problem Statement

Newton School of Technology

We will apply the **transformation** in the data

Time (t)	Value
1	5
2	7
3	9
4	11
5	13
6	15
7	17

Difference Transformation

Newton School of Technology

We take the difference of the two lags to make the data stationary.

$$\Delta X_t = X_t - X_{t-1}$$

X_t : **Current Value**

X_{t-1} : **Previous Value (Lag 1)**

ΔX_t : **Change in Value**

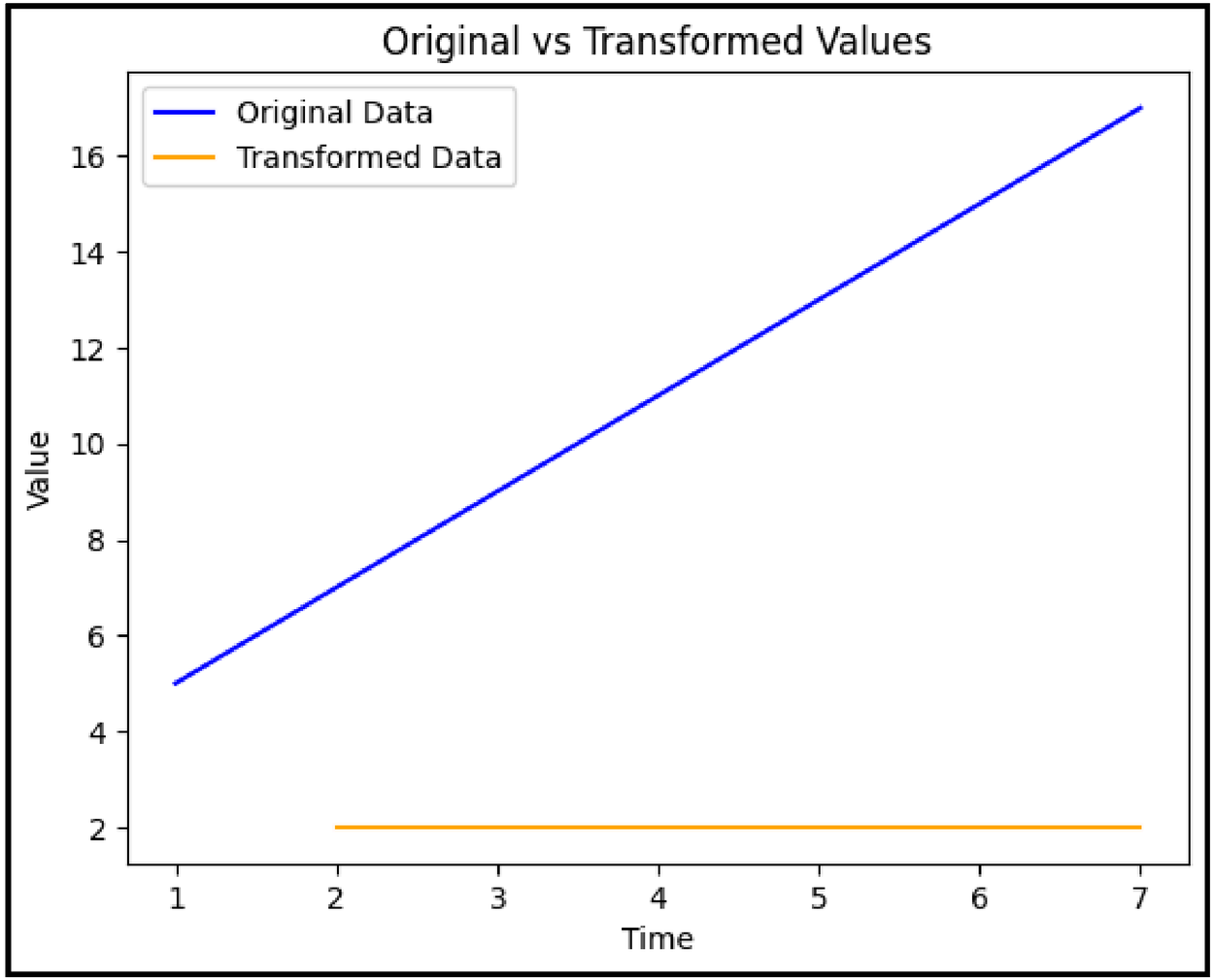
This Produces a stable mean and variance.

Time (t)	Value	Lag 1	Transformed Values (Value - Lag 1)
1	5		
2	7	5	**2**
3	9	7	**2**
4	11	9	**2**
5	13	11	**2**
6	15	13	**2**
7	17	15	**2**

Difference Transformation

Newton School of Technology

Plot of the previous tabular dataset



Newton School of Technology

Can you analyze this data?

Time (t)	Sales
1	2
2	10
3	4
4	20
5	6
6	30
7	8
8	40

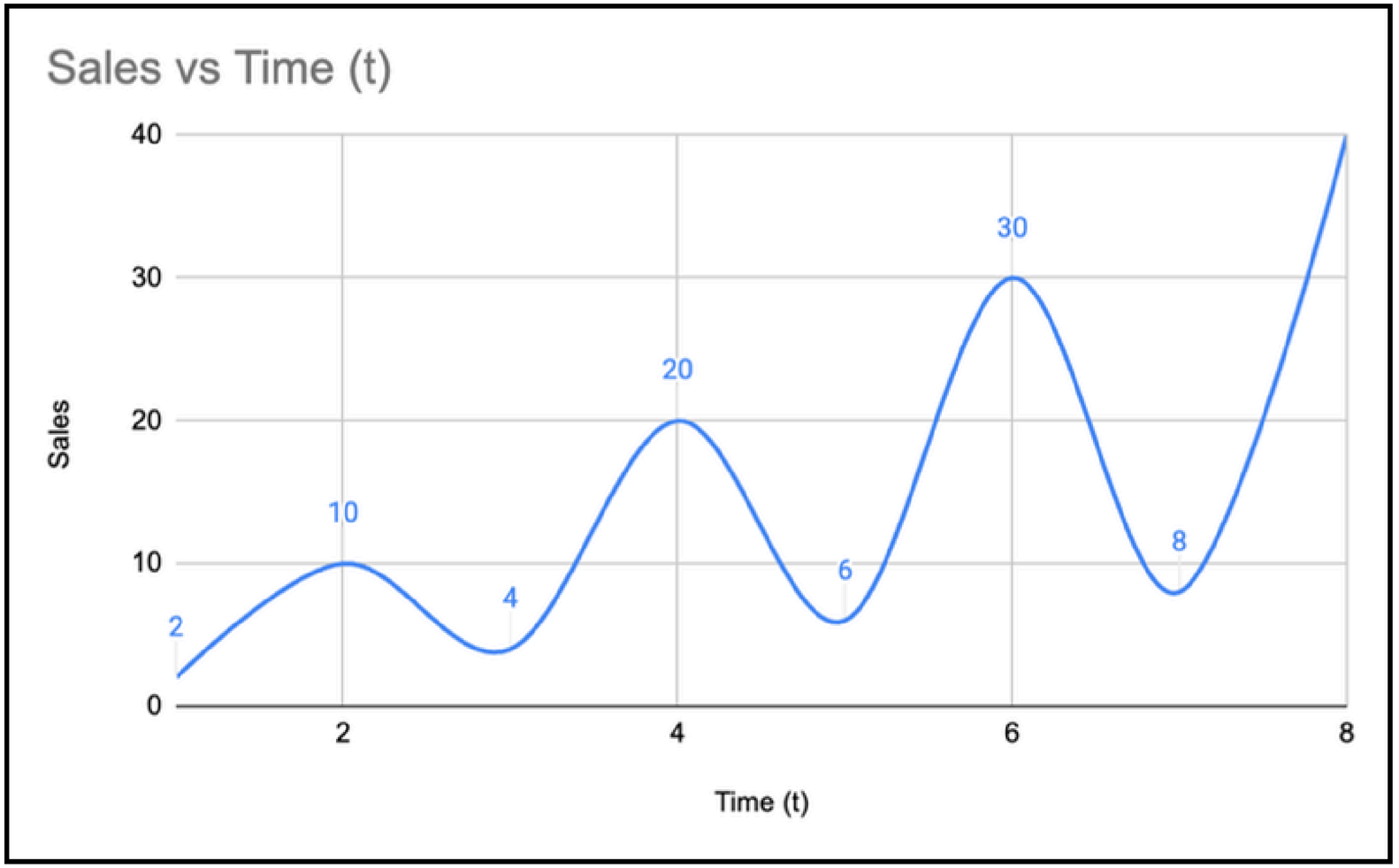
Newton School of Technology

Can difference transformation be applied on it?

Time (t)	Sales
1	2
2	10
3	4
4	20
5	6
6	30
7	8
8	40

Graph of Sales vs Time

Newton School of Technology



Time (t)	Sales
1	2
2	10
3	4
4	20
5	6
6	30
7	8
8	40

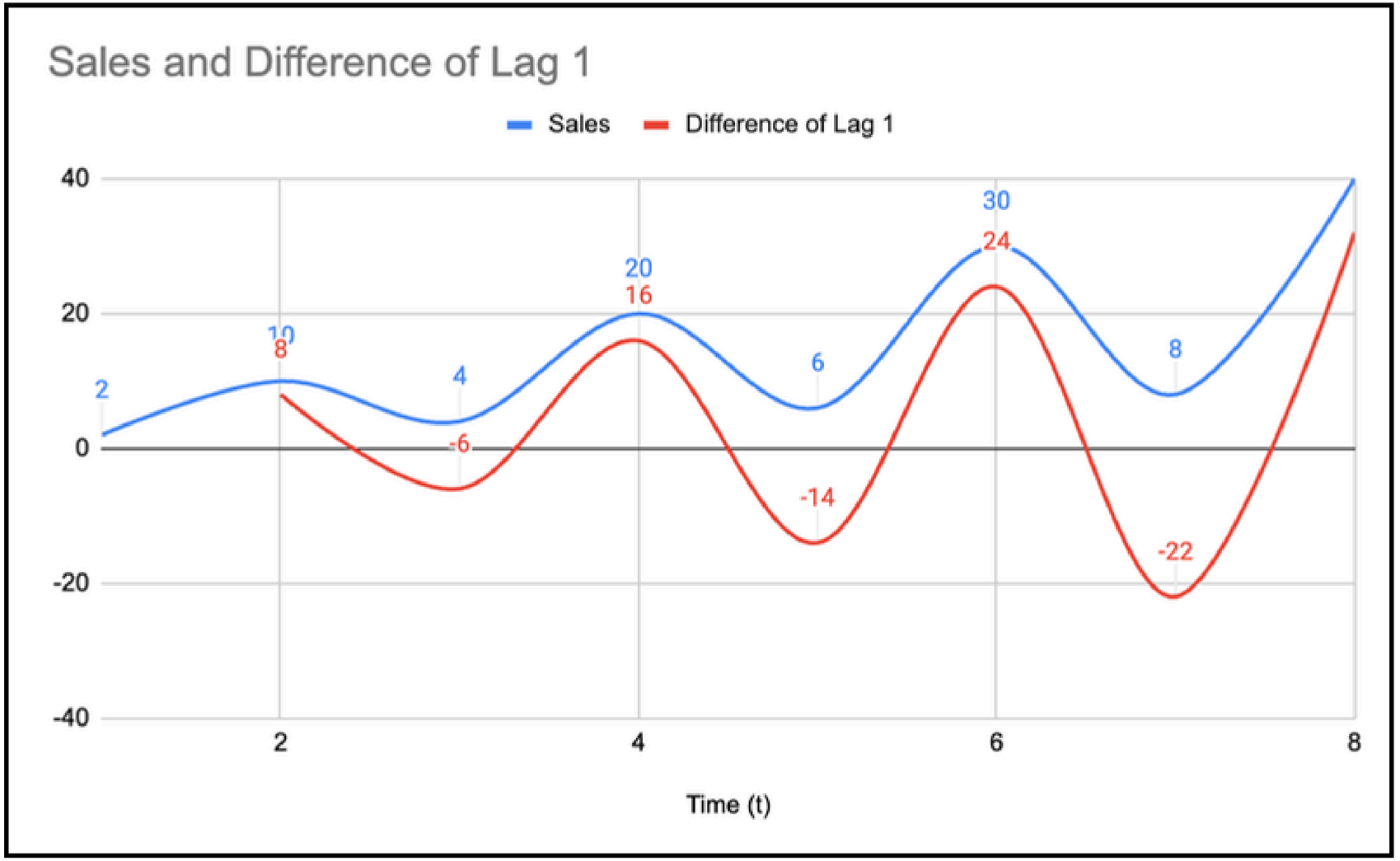
Non Stationary Data

Applying difference transformation on it.

Newton School of Technology

Time (t)	Sales	Lag 1	Difference (Sales - Lag 1)
1	2		
2	10	2	8
3	4	10	-6
4	20	4	16
5	6	20	-14
6	30	6	24
7	8	30	-22
8	40	8	32

Applying difference transformation on it.



Applying difference transformation on it.

Why Differencing Fails

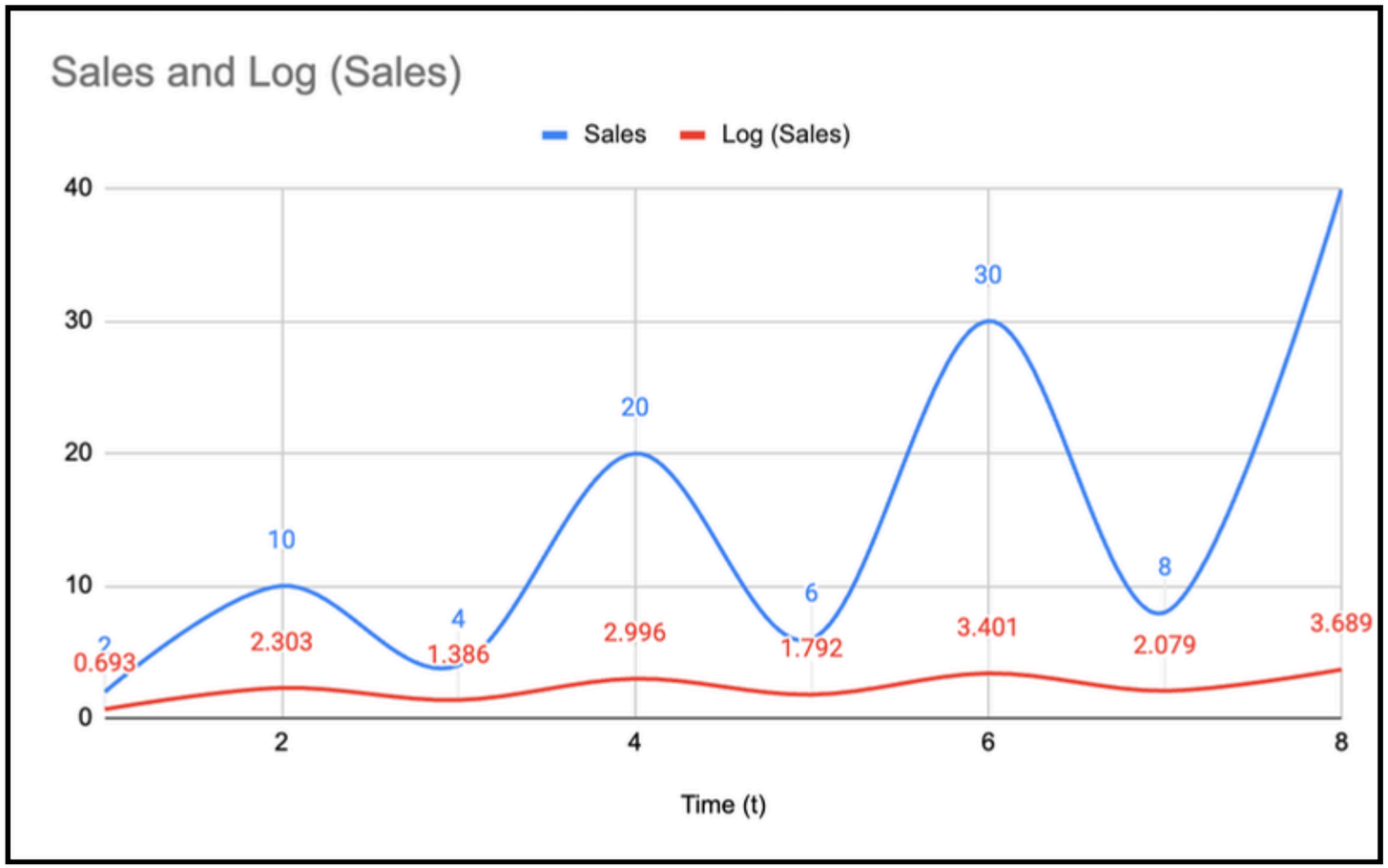
- Differences become larger as values grow
- Same pattern (jump up, fall down) but magnitude keeps increasing
- Variability is not stable
- Large values dominate the series

Key Insight: The problem is not only the change over time, but also that the change depends on how big the value already is

Time (t)	Sales	Lag 1	Difference (Sales - Lag 1)
1	2		
2	10	2	8
3	4	10	-6
4	20	4	16
5	6	20	-14
6	30	6	24
7	8	30	-22
8	40	8	32

Log Transformation

Newton School of Technology



Time (t)	Sales	Transformed Values (Log (Sales))
1	2	0.693
2	10	2.303
3	4	1.386
4	20	2.996
5	6	1.792
6	30	3.401
7	8	2.079
8	40	3.689

Log Transformation

Newton School of Technology

- Reduces **large fluctuations** by **compressing big** values.
- Stabilizes **variance**, especially when **ups and downs** grow over time.
- Makes **exponential growth** look more **linear** and easier to model.
- Helps **reveal underlying patterns** that were hidden by **big swings**.

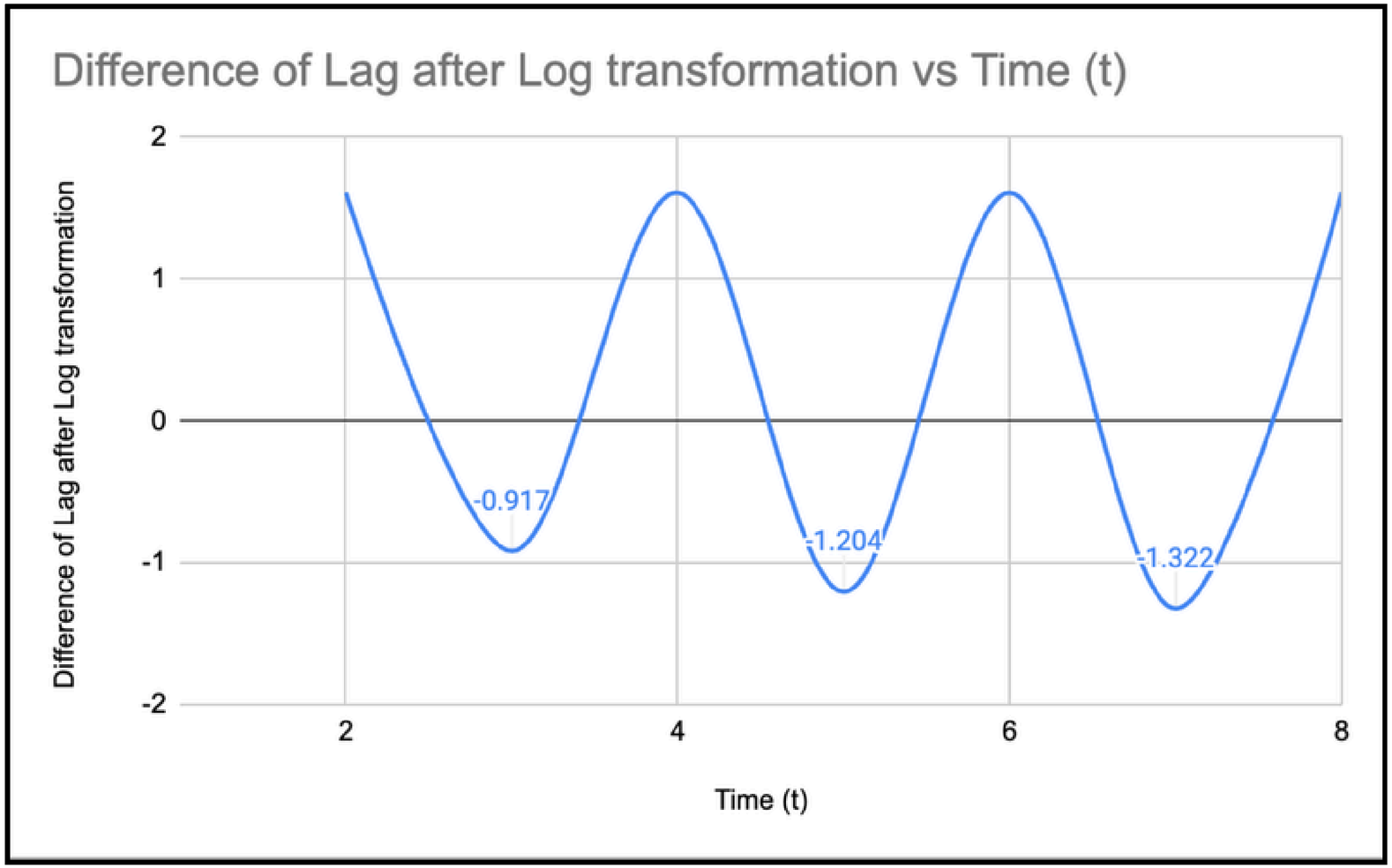
$$Y_t = \log(X_t)$$

Time (t)	Sales	Transformed Values (Log(Sales))
1	2	0.693
2	10	2.303
3	4	1.386
4	20	2.996
5	6	1.792
6	30	3.401
7	8	2.079
8	40	3.689

Lag Difference after Log Transformation

Newton School of Technology

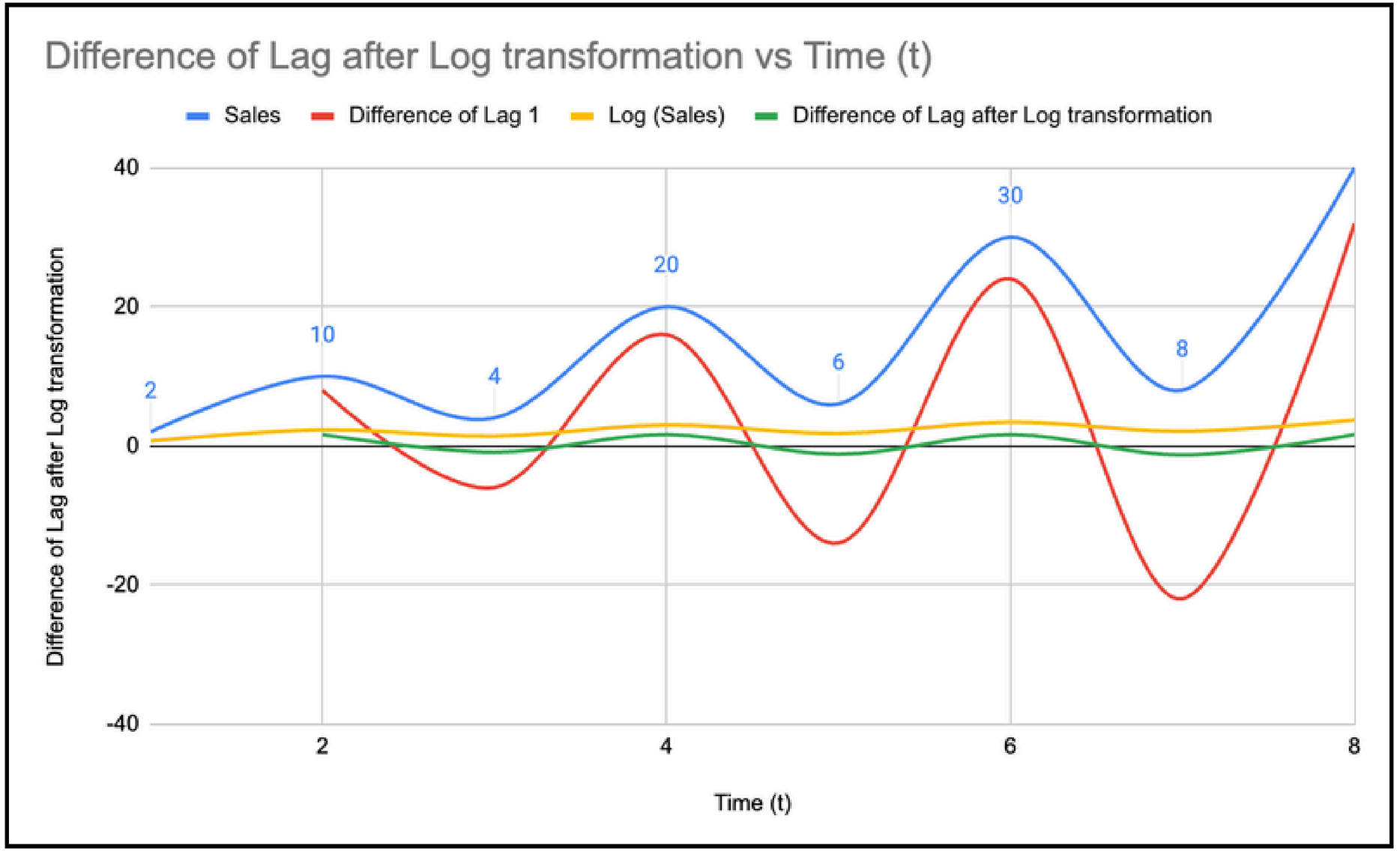
Time (t)	Log(Sales)	Lag 1	Difference
1	0.693		
2	2.303	0.693	1.61
3	1.386	2.303	-0.917
4	2.996	1.386	1.61
5	1.792	2.996	-1.204
6	3.401	1.792	1.609
7	2.079	3.401	-1.322
8	3.689	2.079	1.61



Stationary Data

Comparison of all of the graphs in one image

Newton School of Technology



Autocorrelation

Newton School of Technology

- **What is Autocorrelation?**
  - Autocorrelation is a mathematical concept to check the dependency of today's values with past values.
  - In simple terms: to check whether the past influences the present
- **How does it help?**
  - Detect patterns + seasonality (weekly cycles)
  - Build better forecasts
  - Avoid assuming the data is random

Mathematical formula for Autocorrelation

Newton School of Technology

Autocorrelation at lag k:

$$\rho(k) = Corr(x_t, x_{t-k})$$

Expanded form:

$$\rho(k) = \frac{\sum_{t=k+1}^{N}(x_t - \mu)(x_{t-k} - \mu)}{\sum_{t=1}^{N}(x_t - \mu)^2}$$

Where:

- N = number of observations
- μ = mean of time series
- k = lag

Newton School of Technology

Please fill the feedback form.

Newton School of Technology

Thank You

Newton School of Technology

Join the lecture online on your dashboard.

Newton School of Technology

Let’s start with a minute of silence.

आचार्यात् पादं आधत्ते पादं शिष्यः स्वमेधया ।
पादं सब्रह्मचारिभ्यः पादं कालक्रमेण च ॥

Newton School of Technology

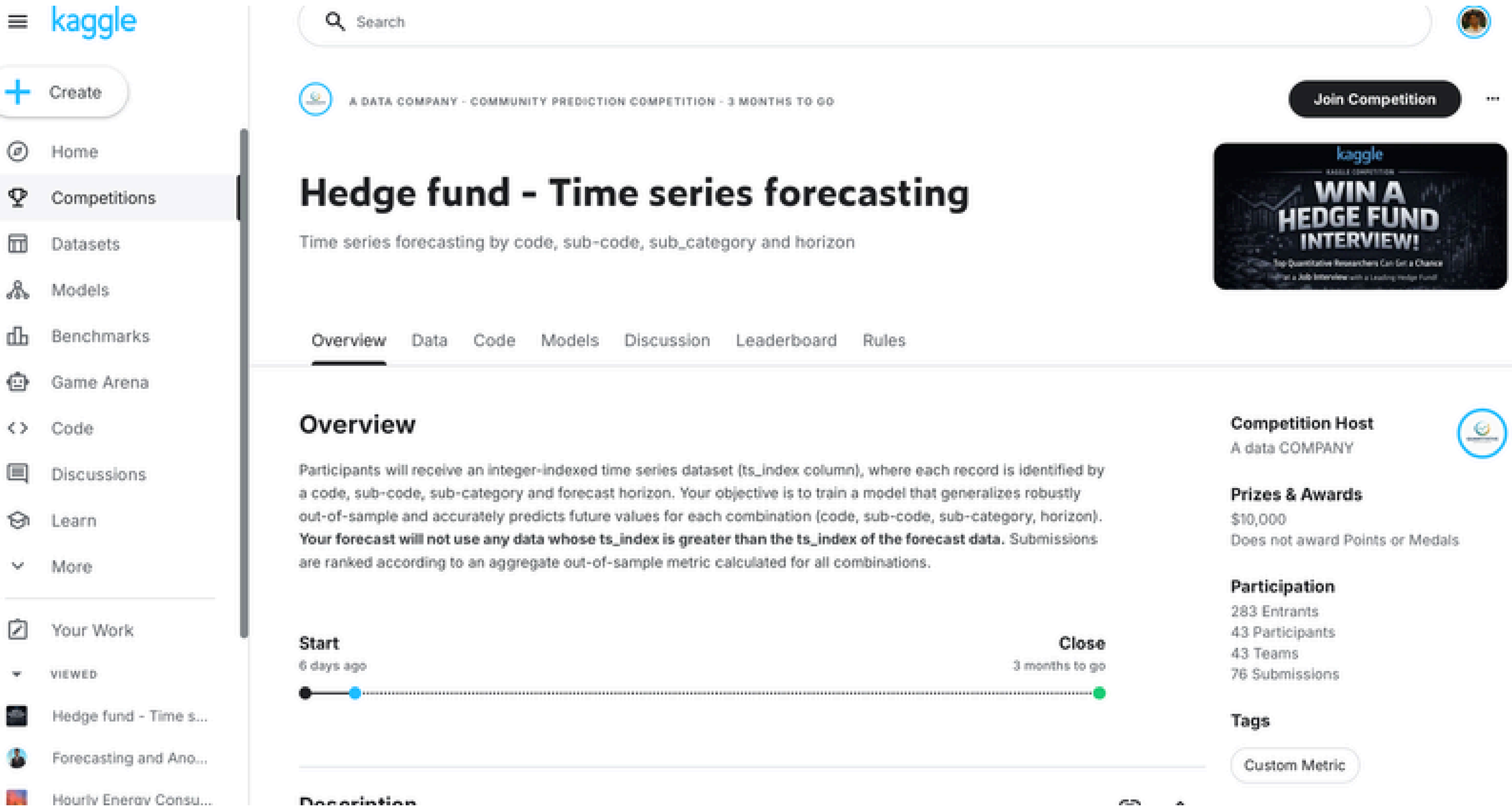
Meaning:

A student acquires knowledge in four equal parts:

- One-fourth from the teacher
- One-fourth through self-reflection and independent thinking
- One-fourth through discussions with peers and fellow learners
- One-fourth over time through personal experience

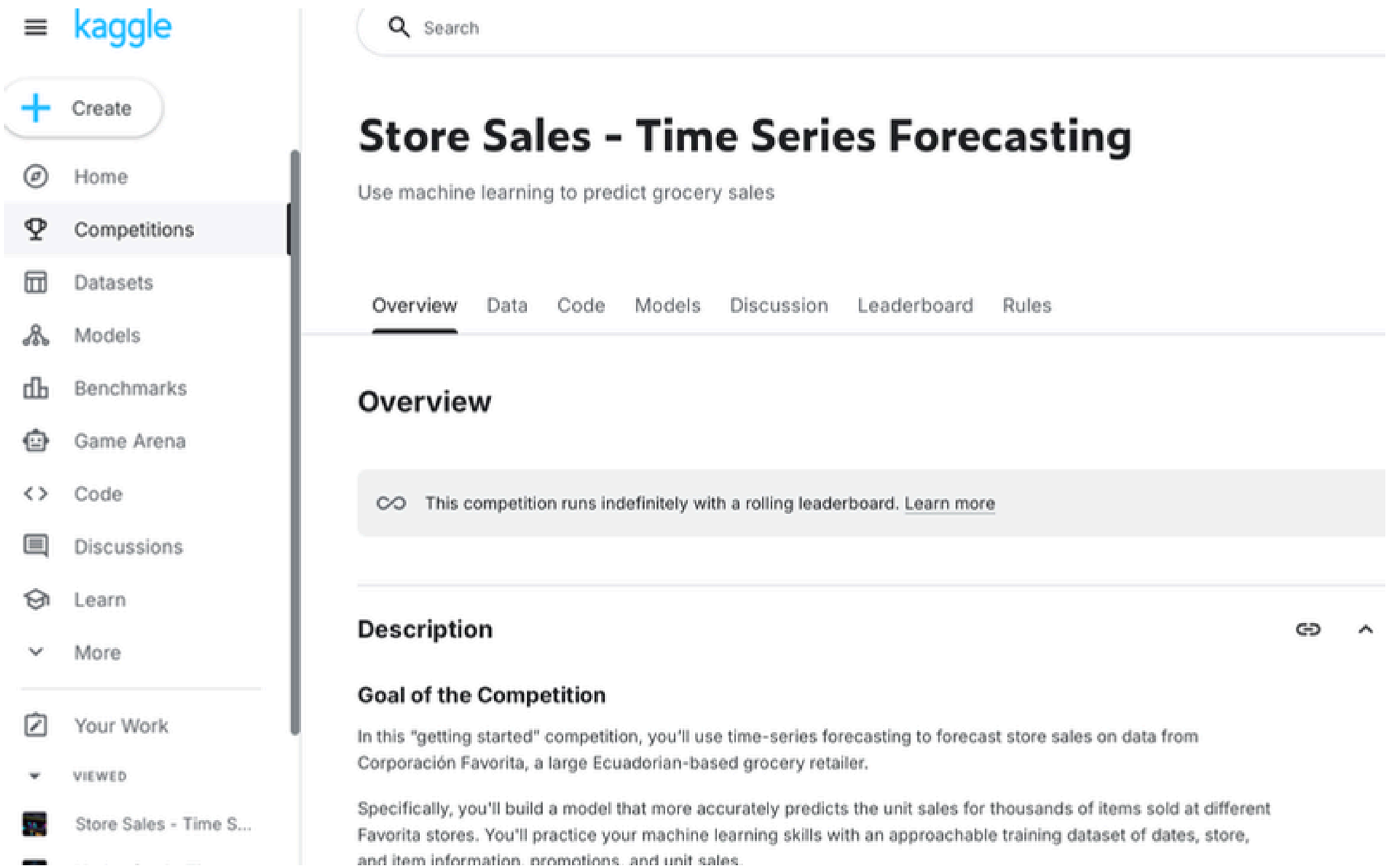
Projects for this Unit

Newton School of Technology



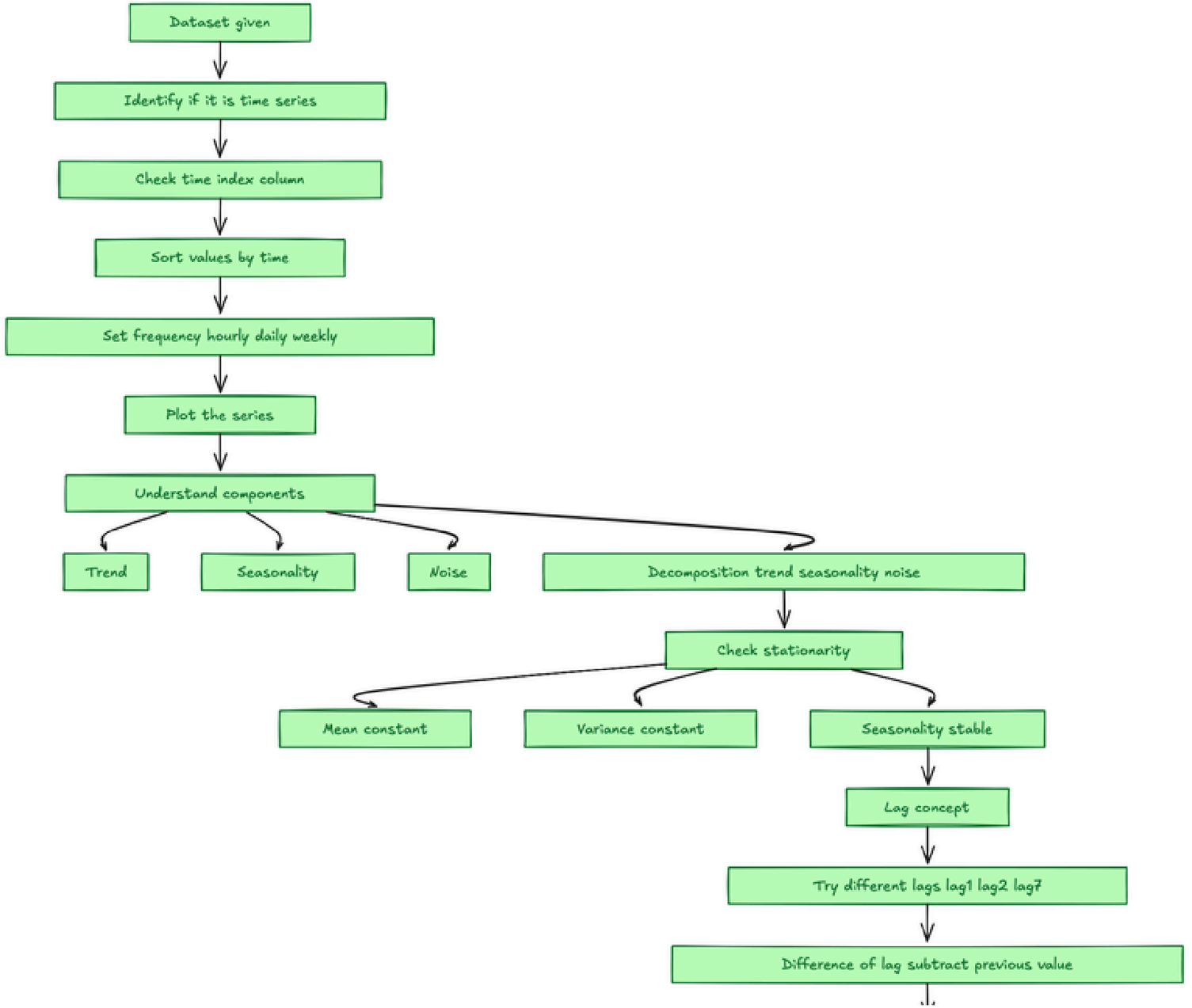
Projects for this Unit

Newton School of Technology

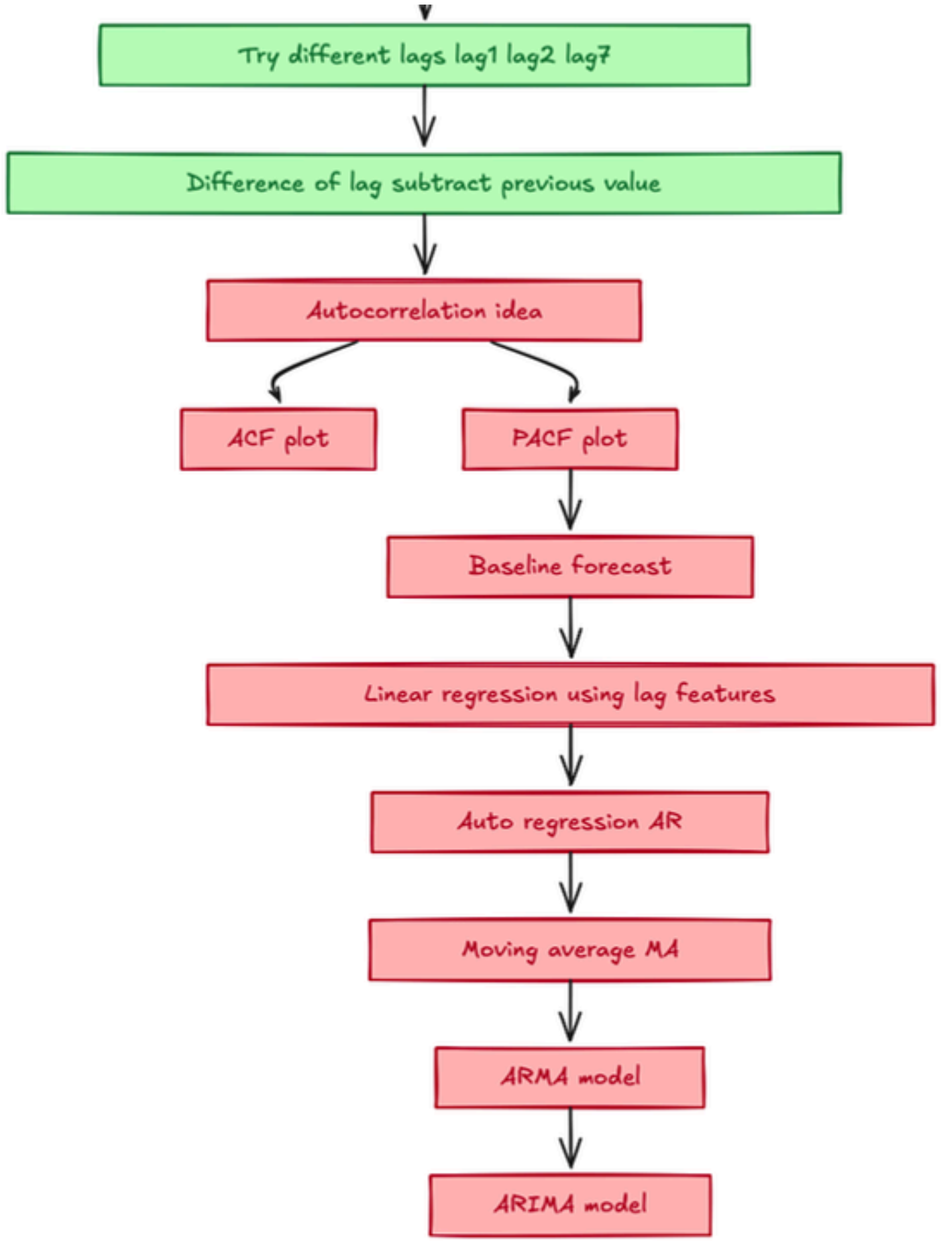


Recap

Newton School of Technology

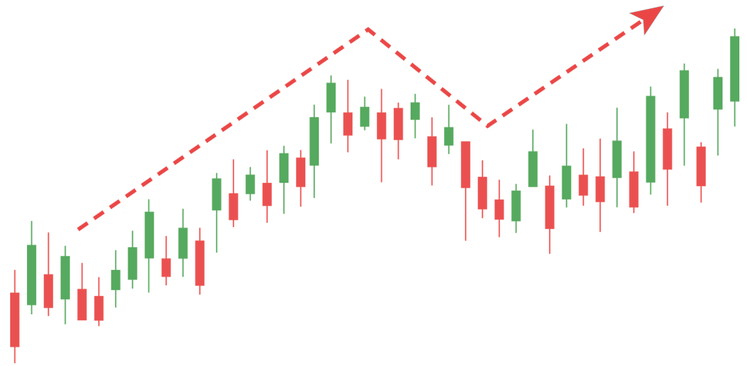


Agenda

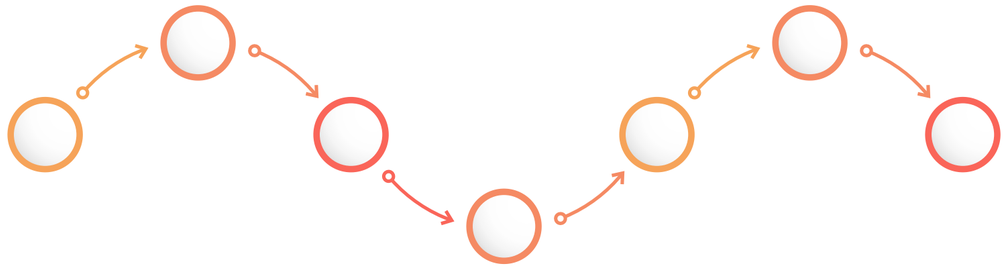


Newton School of Technology

Newton School of Technology



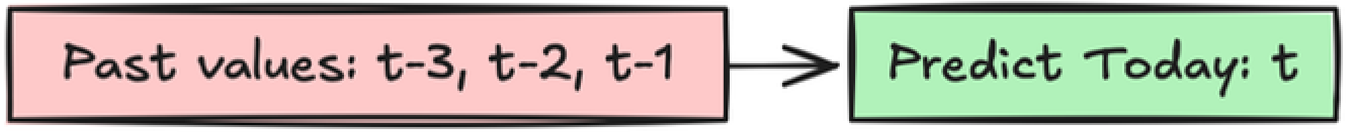
In the previous lecture, we learned how **present** value **depends** on the **past** data and why the **order of time** is **important**. What strategy will you use to predict the present values?



Question

Newton School of Technology

Can you think of a method to predict the value of **today's data point** using **previous data points**.



Question: Can you predict Sales at 6th time stamp?

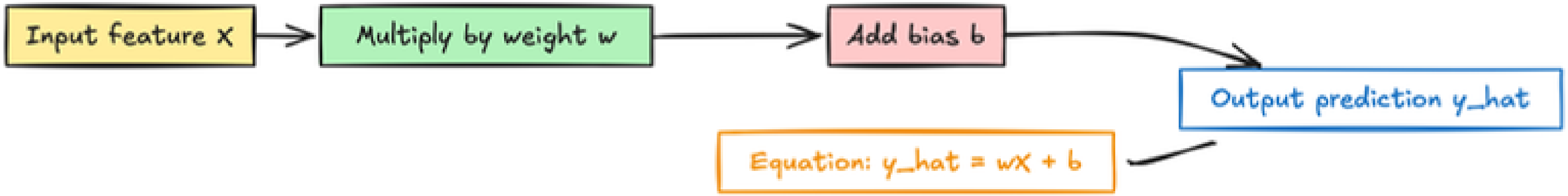
Newton School of Technology

Day	Sales
1	120
2	135
3	128
4	150
5	160
6	???

Question

Newton School of Technology

Can you relate it with linear regression?



$$Y = wX + b$$

Question

Newton School of Technology

Day	Sales
1	120
2	135
3	128
4	150
5	160
6	???

What is X and Y in this?

Question

Newton School of Technology

Day	Sales
1	120
2	135
3	128
4	150
5	160
6	???

Can we consider Day as X?

Question

Day	Sales
1	120
2	135
3	128
4	150
5	160
6	???

Newton School of Technology

Can we consider Values of Sales from Day 1 to Day 6 as X?

Question

Newton School of Technology

Day	Sales	Sales (Lag 1)
1	120	
2	135	120
3	128	135
4	150	128
5	160	150
6	???	

Lag 1 - Can we consider Values of Sales from Day 1 to Day 6 as X?

SOLUTION: Linear Regression using Lags

Newton School of Technology

White noise

$$y_t = c + \phi y_{t-1} + \epsilon$$

Current Target Value

A constant

Influencing Factor

Value of target 1 day ago

We can use the lags and develop a relation of the target variable with the specific past values (lags).

Newton School of Technology

Can we **incorporate** the **importance** of all the **other lags**?

Question

Newton School of Technology

Day	Sales	Sales (Lag 1)	Sales (Lag 2)
1	120		
2	135	120	
3	128	135	120
4	150	128	135
5	160	150	128
6	???		

Can Lag 1 and Lag 2 Act as X_1 and X_2 ?

Equation for prediction (other lags)

Newton School of Technology

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t$$

y_t : (Value at time t)

c : (Constant/Intercept)

ϕ_i : (Coefficient for lag i)

ϵ_t : (White noise/Error term)

Equation for prediction (other lags)

Newton School of Technology

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t$$

- How can we **determine the values of lag coefficients** and decide which **lags** to include in our model?
- Including **all lags** could lead to unnecessary complexity.
- Therefore, it's **crucial** to **select** only the most **relevant lags** that **significantly influence** the time series, ensuring the model remains both efficient and interpretable.

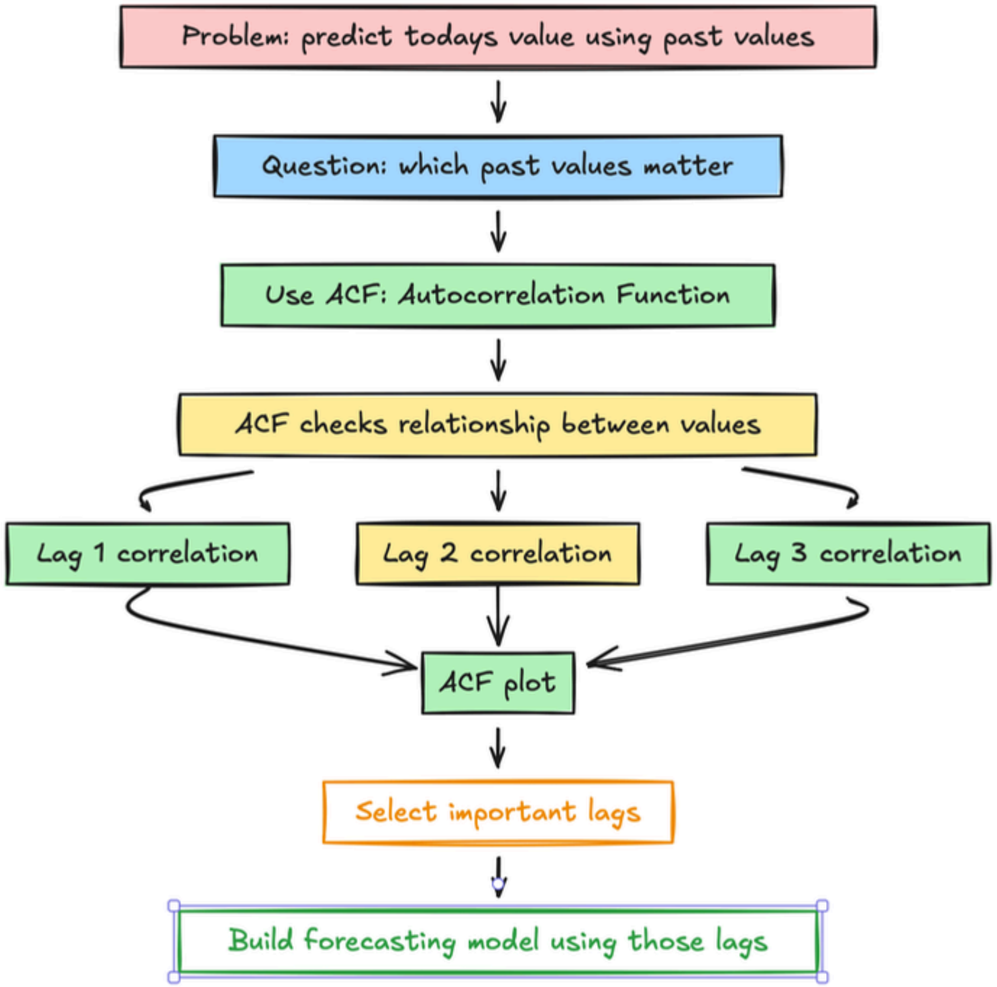
Newton School of Technology

How to solve this problem?

Solution

Newton School of Technology

We use Autocorrelation Function to solve this problem.



Autocorrelation Function (ACF)

Newton School of Technology

- The ACF is a **collection of autocorrelations** across **many lags**, shown together.
- ACF answers:
  - How strong is **lag 1**?
  - How strong is **lag 2**?
  - How strong is **lag 3**?
  - How strong is **lag k**?
- The **ACF** is usually displayed as a **plot of autocorrelations** at **different lags**.
- Think of it like this:
  - **Autocorrelation** = one value
  - **ACF** = list of autocorrelations for multiple lags

Range of ACF

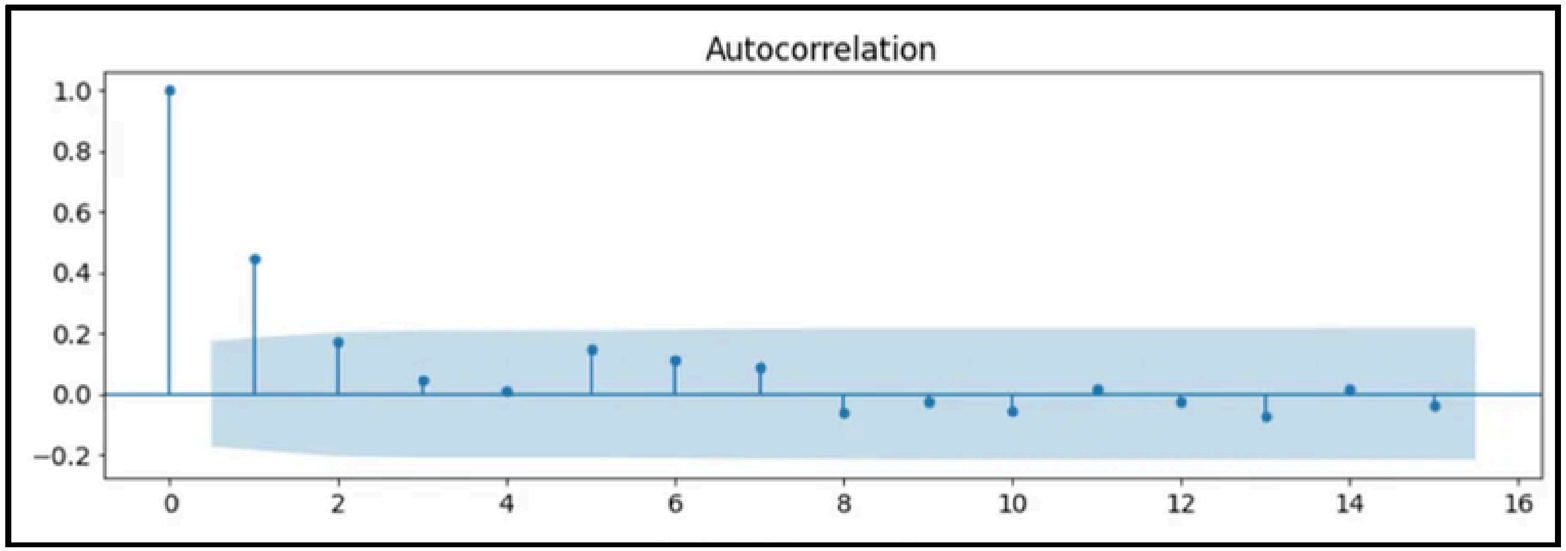
Newton School of Technology

Because ACF is literally a **correlation coefficient**, and correlation is always bounded:

$$-1 \leq \rho(k) \leq 1$$

Interpretation:

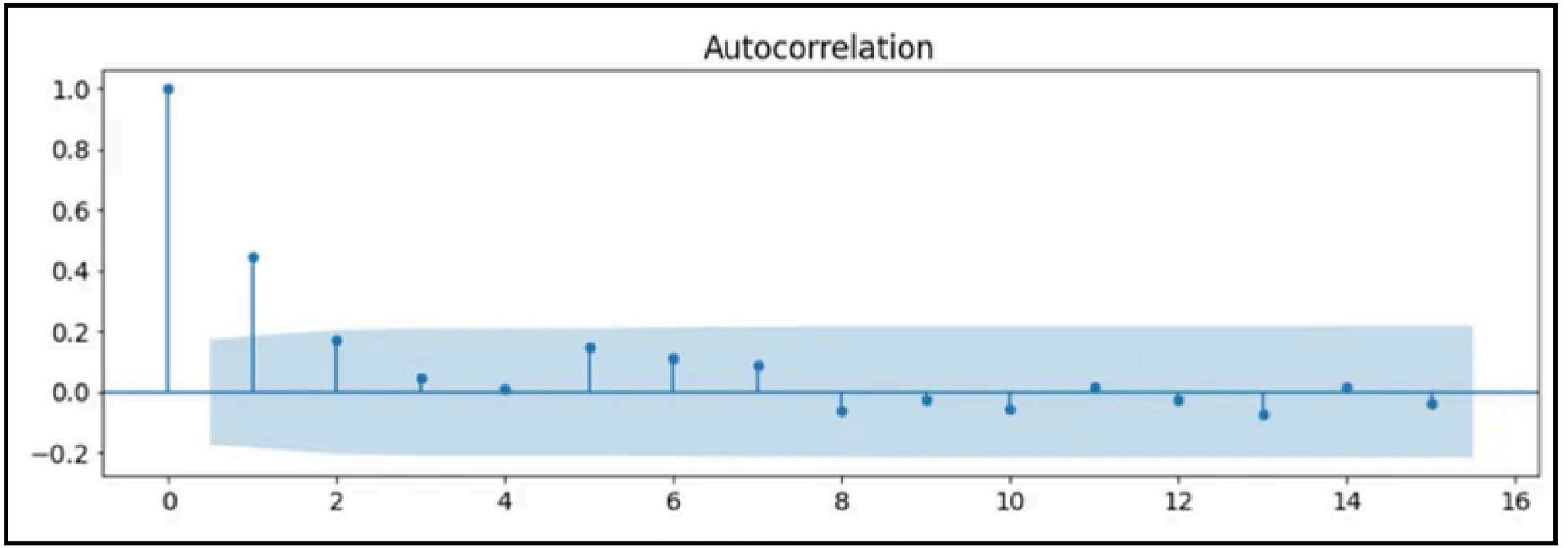
- $\rho(k) \approx 1$: strong positive relation (moves together)
- $\rho(k) \approx -1$: strong negative relation (moves opposite)
- $\rho(k) \approx 0$: no relation



Autocorrelation Function Plot

Newton School of Technology

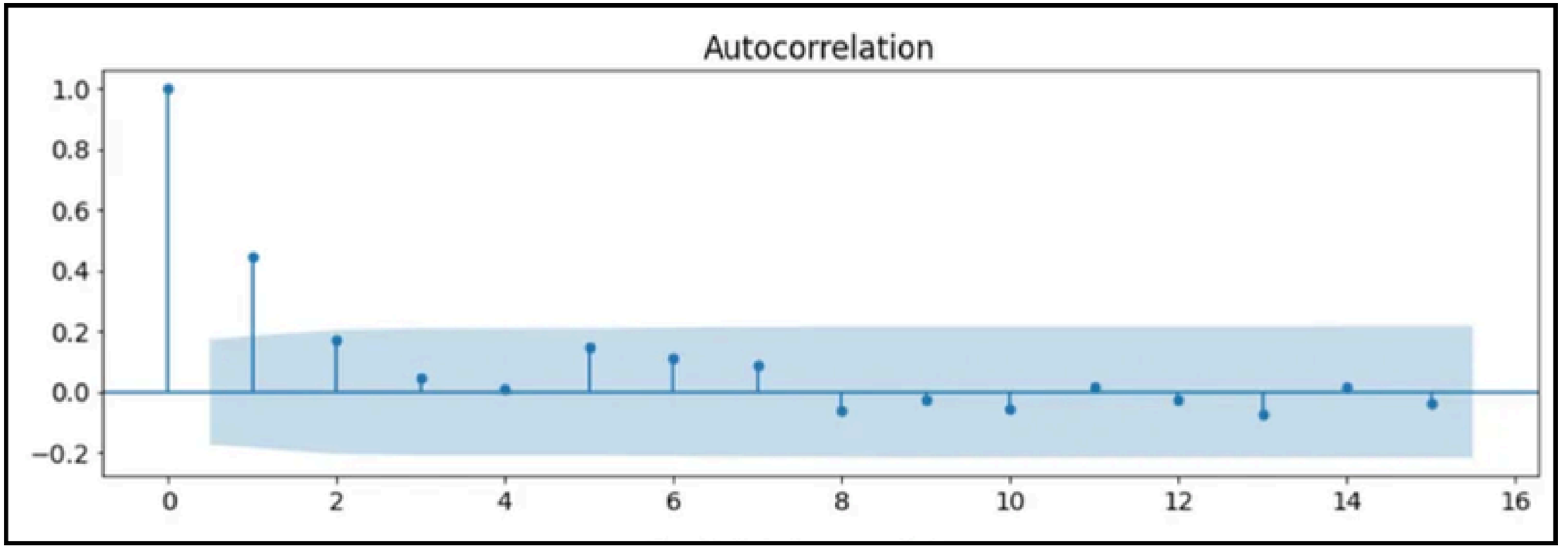
- Each vertical spike represents the strength of correlation at that lag.
- The shaded **blue** region represents the **confidence band**:
  - Bars inside the box → NOT statistically significant
  - Bars outside the band → statistically significant



What is confidence band?

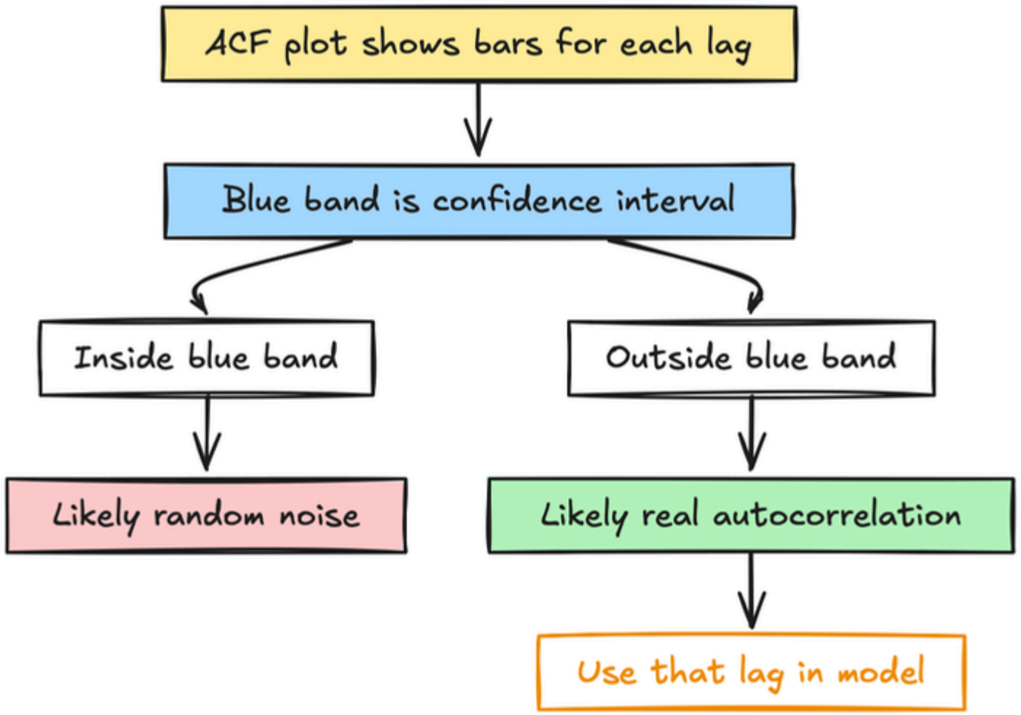
Newton School of Technology

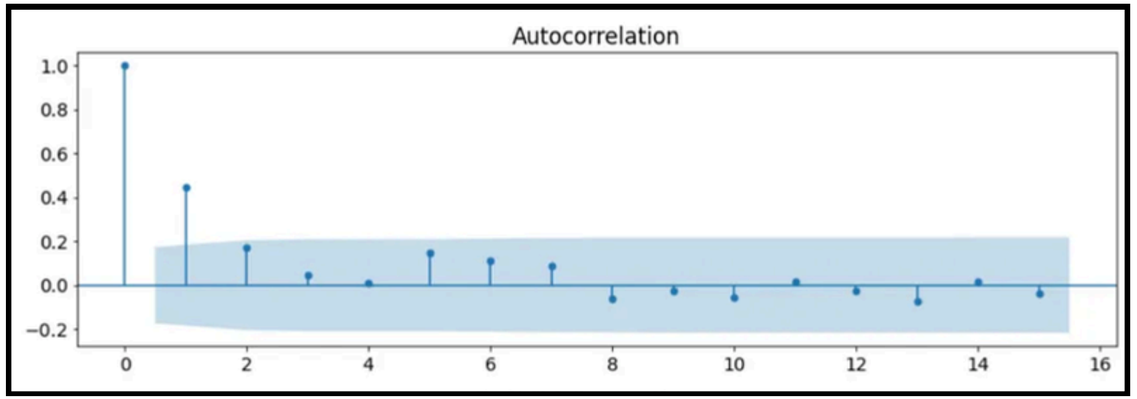
- **So the plot shows a blue shaded region = range of correlations that are “normal by chance”.**
- If a spike goes outside the blue band, it is likely a real signal (statistically significant autocorrelation).
- Most ACF plots use an approximate 95% confidence interval



What is confidence band?

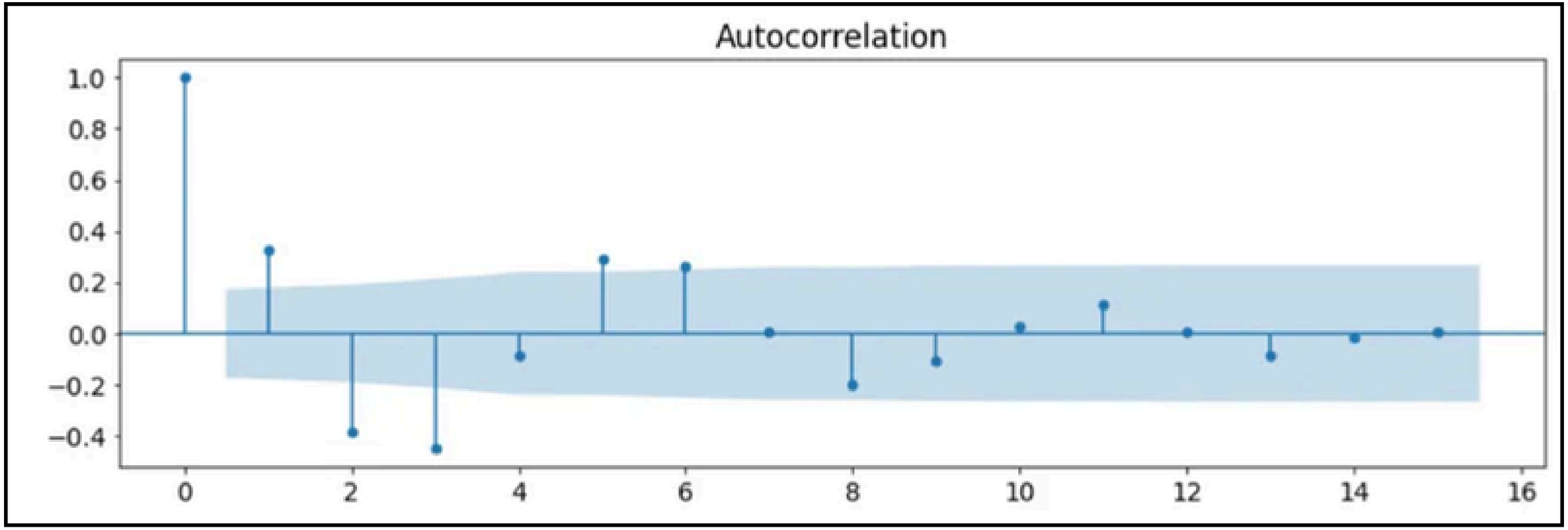
Newton School of Technology





Problem Statement

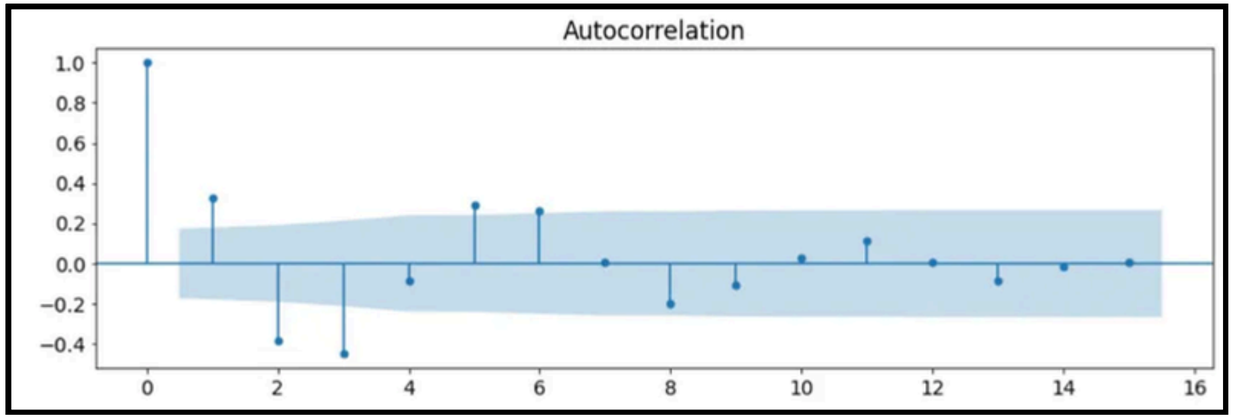
What does the negative bar in the following plots mean ?



Autocorrelation Function Plot

Newton School of Technology

- A **negative autocorrelation** at **lag 2** means:
- The value **today tends to move** in the **opposite direction** from the value from two time steps ago.
- In simpler terms:
  - If the value 2 days ago was high,
    - **Today tends to be lower**.
  - If the value 2 days ago was low,
    - **Today tends to be higher**.
- This indicates a **reversal effect** or **correction pattern** after two time steps.



Autocorrelation Function Plot

Newton School of Technology

Most ACF plots use an approximate 95% confidence interval:

$$CI \approx \pm \frac{1.96}{\sqrt{N}}$$

Where:

- N = number of observations in the time series
- 1.96 = z value for 95% confidence

So:

$$\text{Blue band upper} = +\frac{1.96}{\sqrt{N}}$$

$$\text{Blue band lower} = -\frac{1.96}{\sqrt{N}}$$

SOLUTION: Linear Regression using Lags

Newton School of Technology

White noise

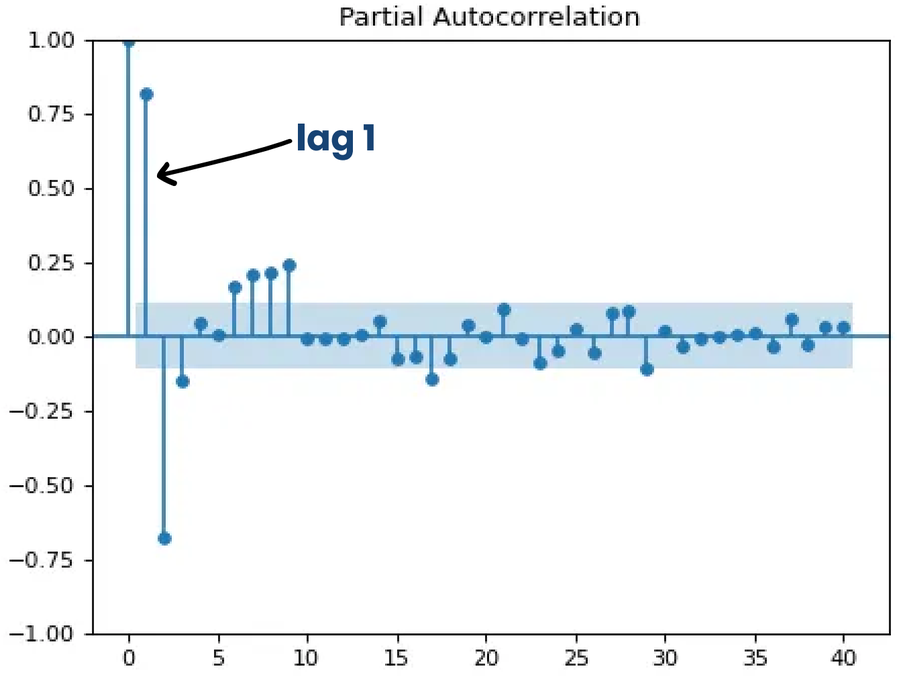
$$y_t = c + \phi y_{t-1} + \epsilon$$

Current Target Value

A constant

Influencing Factor

Value of target 1 day ago



$\lg 1 = 0.8$

$\lg 2 = 0.75$

$\lg 3 = 0.9$

$Ci = 0.79$

Partial Autocorrelation (PACF)

Newton School of Technology

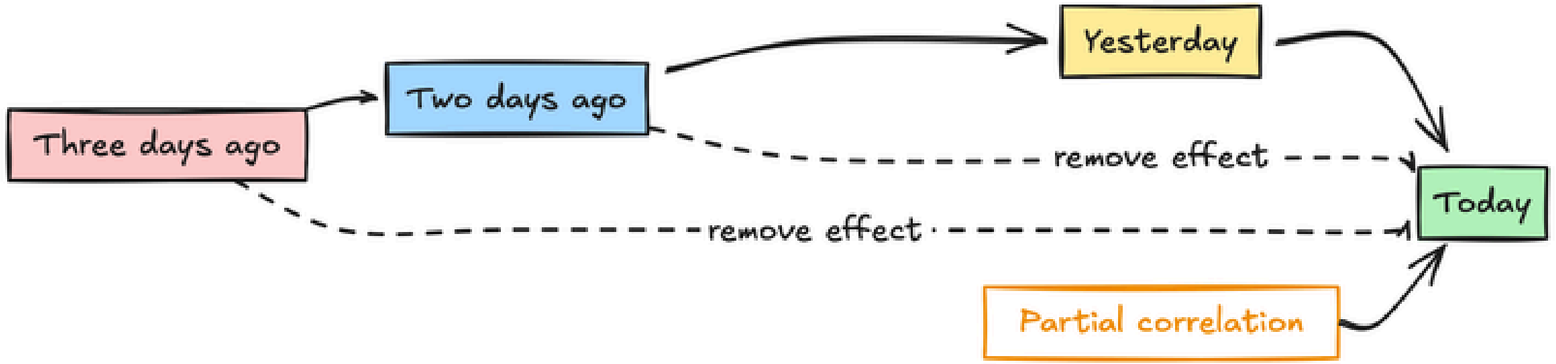
- Imagine you're predicting today's Zomato orders:
  - **Yesterday's orders** influence **today** → **lag 1**
  - The day before yesterday also influences today,
  - But part of that influence is already explained by lag 1.
- PACF removes all "**shortcut**" effects.
- **ACF says:**
  - "How related are Day 3 and Day 1 overall?"
- **PACF says:**
  - "How related are Day 3 and Day 1 if I ignore what happened on Day 2?"

ACF Vs PACF

Newton School of Technology

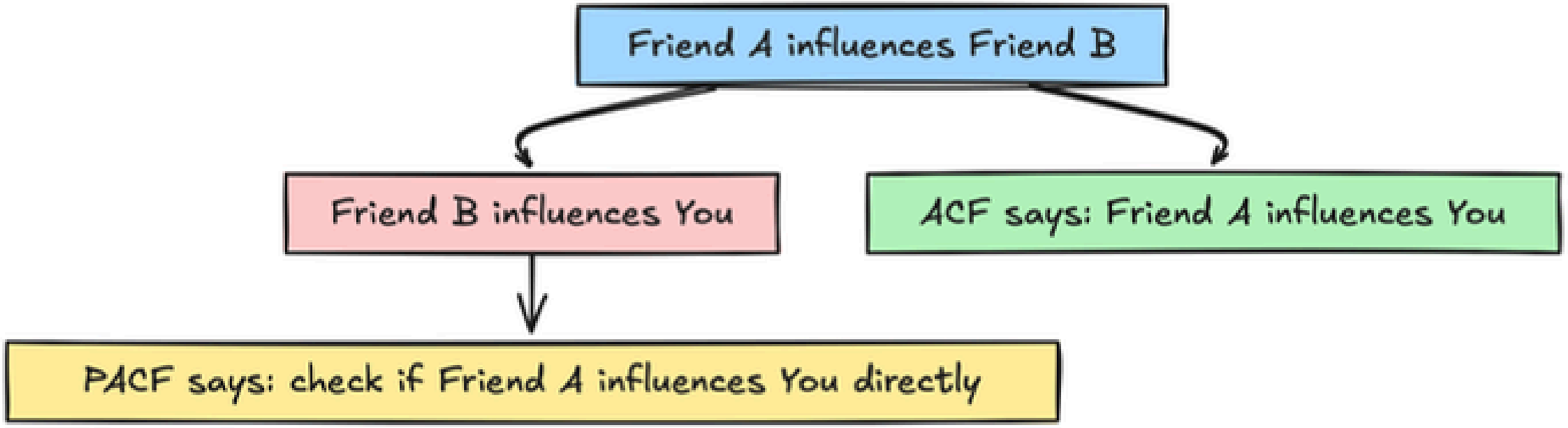
ACF shows total influence

PACF shows direct influence



ACF Vs PACF

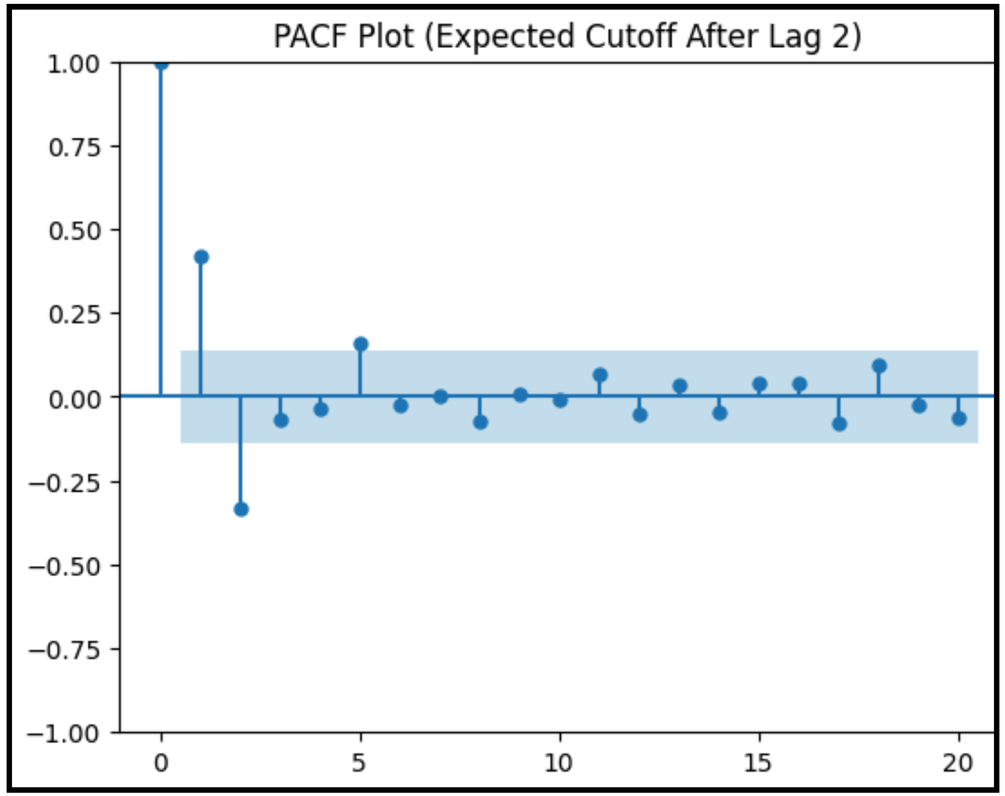
Newton School of Technology



Interpretation of PACF Plot

Newton School of Technology

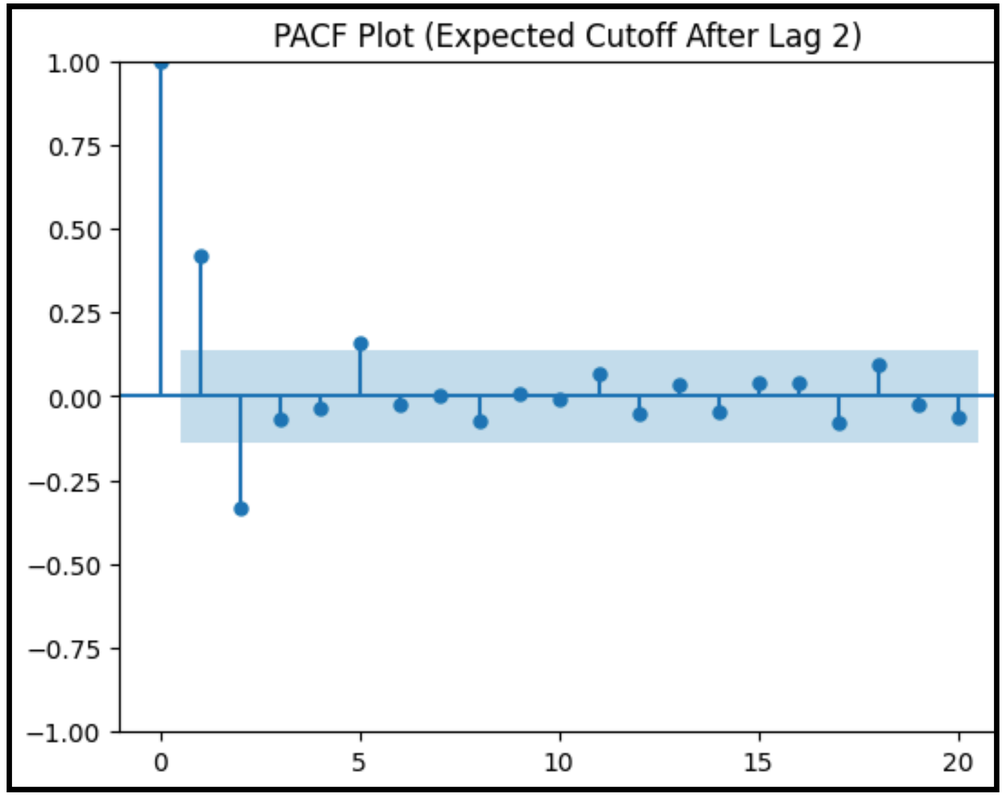
You are analyzing a **daily sales time series** that has already been made **stationary** (trend and seasonality removed). You compute the Partial Autocorrelation Function (PACF) and obtain the following plot:



Interpretation of PACF Plot

Newton School of Technology

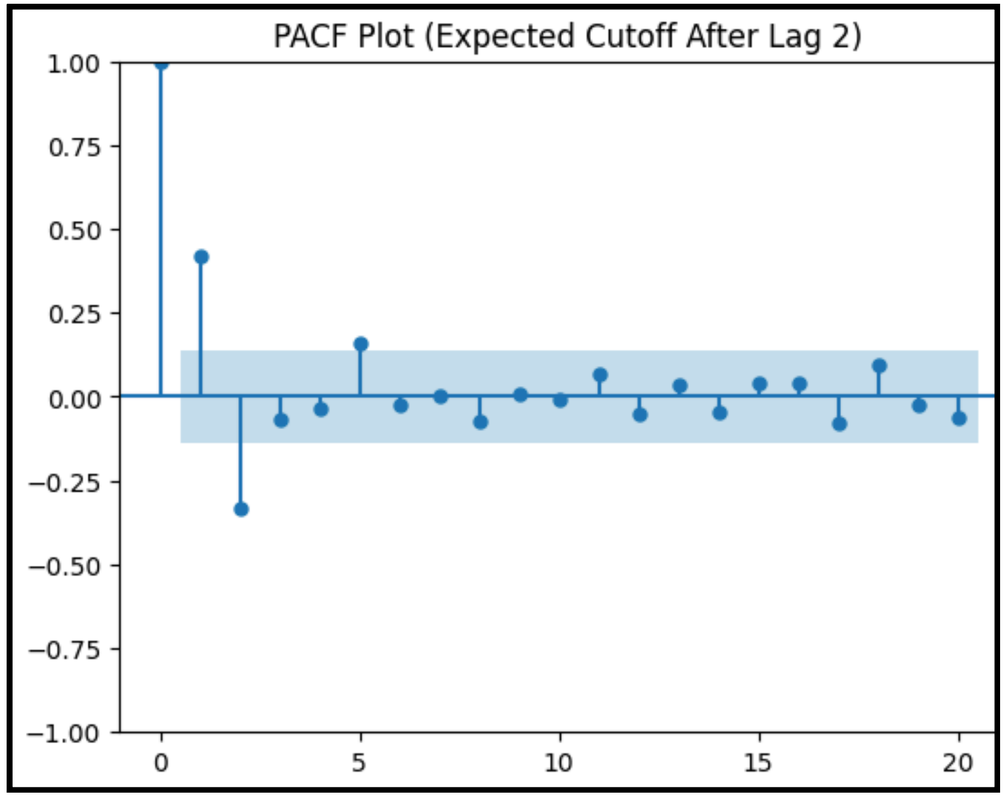
What does this **PACF plot suggest** about **how current sales** depend on **past sales**?



Interpretation of PACF Plot

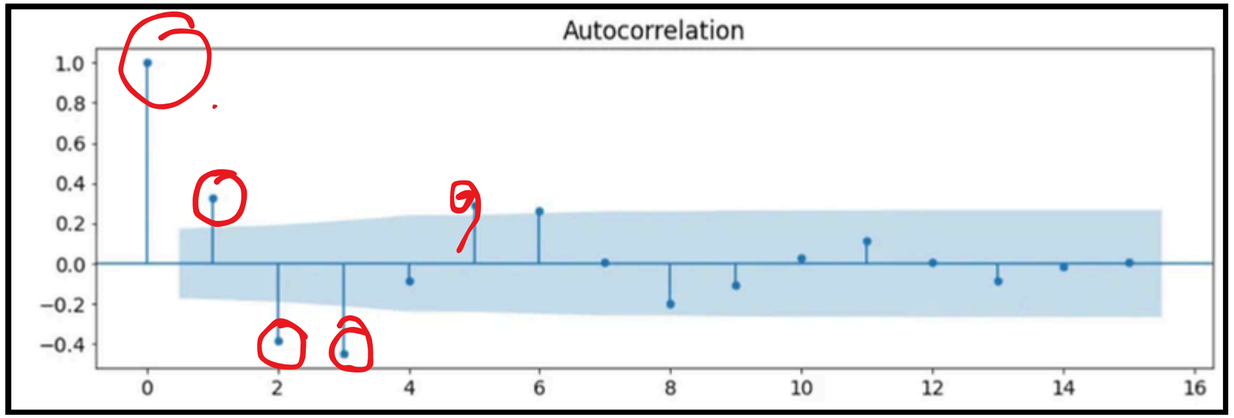
Newton School of Technology

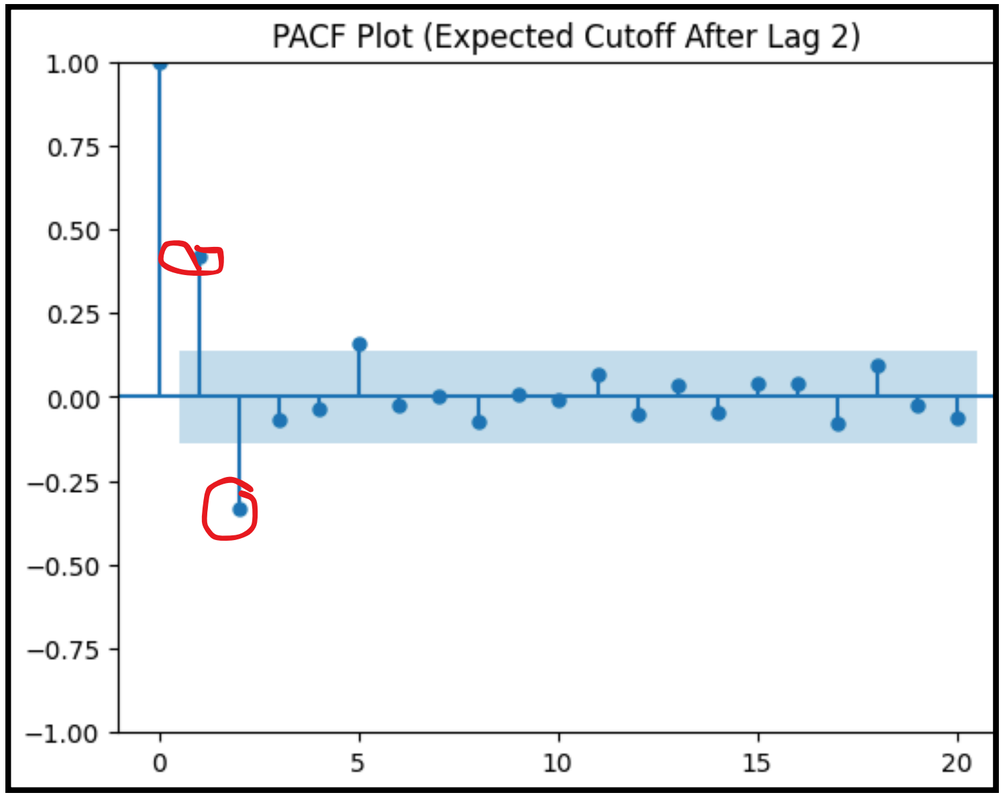
- **Lag 1:** Extreme positive spike
  - Yesterday has a strong direct influence on today.
- **Lag 2**: Strong negative spike
  - A negative PACF at lag 2 tells you:
  - Day-before-yesterday has a strong inverse effect on today, once we remove yesterday's influence.
- **Lags 3 afterwards**: All small, inside the confidence band
  - Once you go past lag 2, no more significant spikes
  - Everything is basically noise
  - No long-range "direct" relationships



Interpretation of PACF Plot

Newton School of Technology





Newton School of Technology

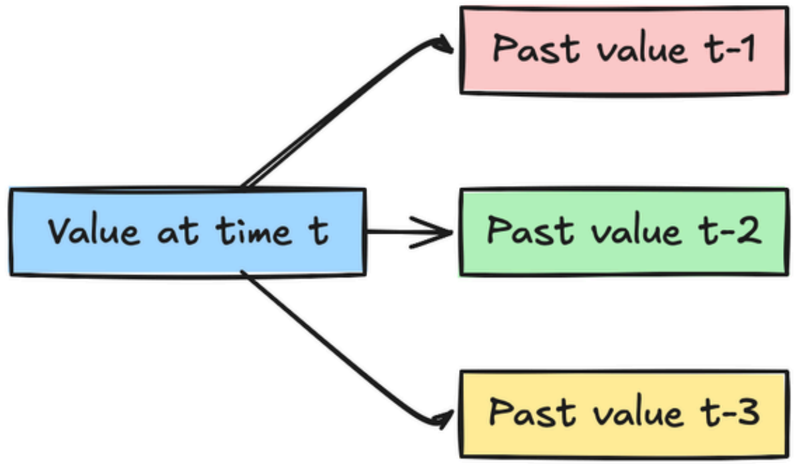
What We just **discovered** is

Autoregression

Auto Regression

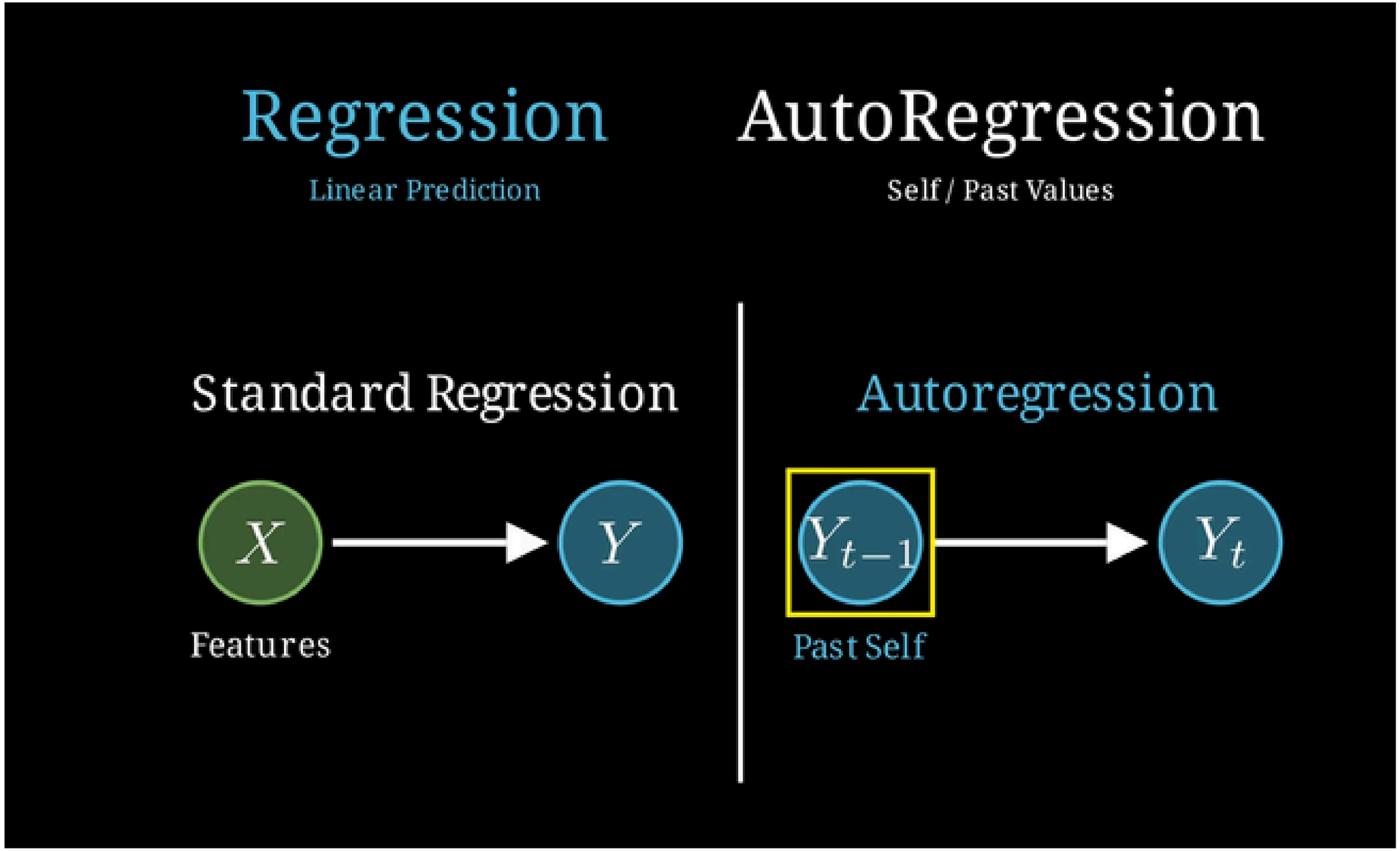
Newton School of Technology

- **Auto** (Self): implies the variable is **predicted** by itself (specifically, its own past values).
- **Regression**: a **statistical technique** for predicting a **target value** from a **combination of inputs**.
- Unlike **standard regression** (which uses other features like x1, x2 to predict y), Autoregression **uses past values** as the features.



Regression Vs AutoRegression

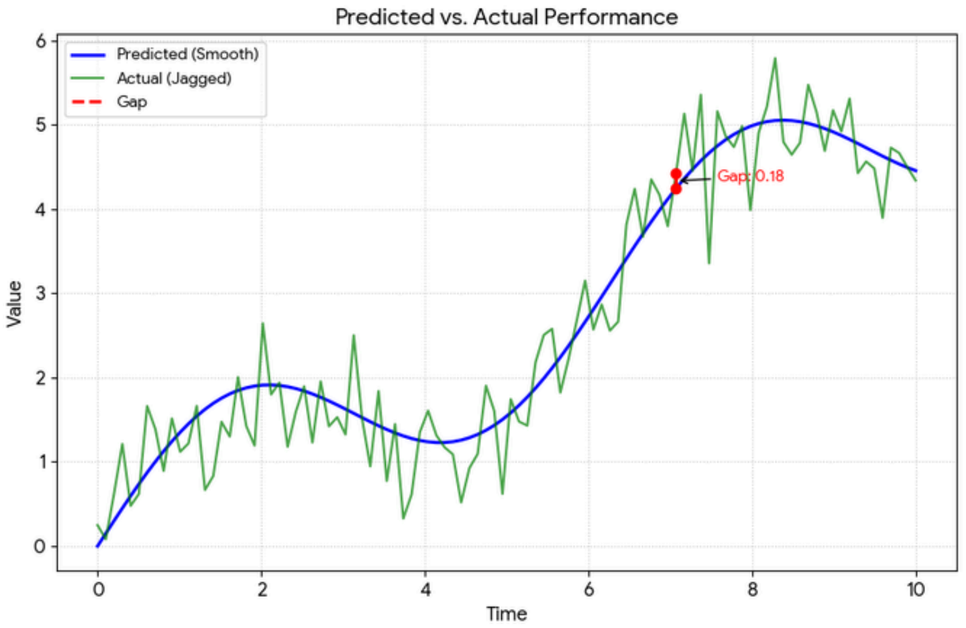
Newton School of Technology



Shock

Newton School of Technology

- **Prediction** is **rarely perfect**.
- Imagine you **predicted** the temperature would be **30°C**, but it was **actually 35°C**.
- That **+5°C difference** is the **Error**, also called the **Residual**, **Innovation**, or **Shock**.
- It represents new information or **unexpected events** that our **past-value model** couldn't see.



Newton School of Technology

If a "**shock**" happened **yesterday** (**e.g., a sudden market crash or a breaking news story**), does its **impact disappear instantly**?

Newton School of Technology

No. The "**ripple effect**" of that **shock** is often **felt today**.

Newton School of Technology

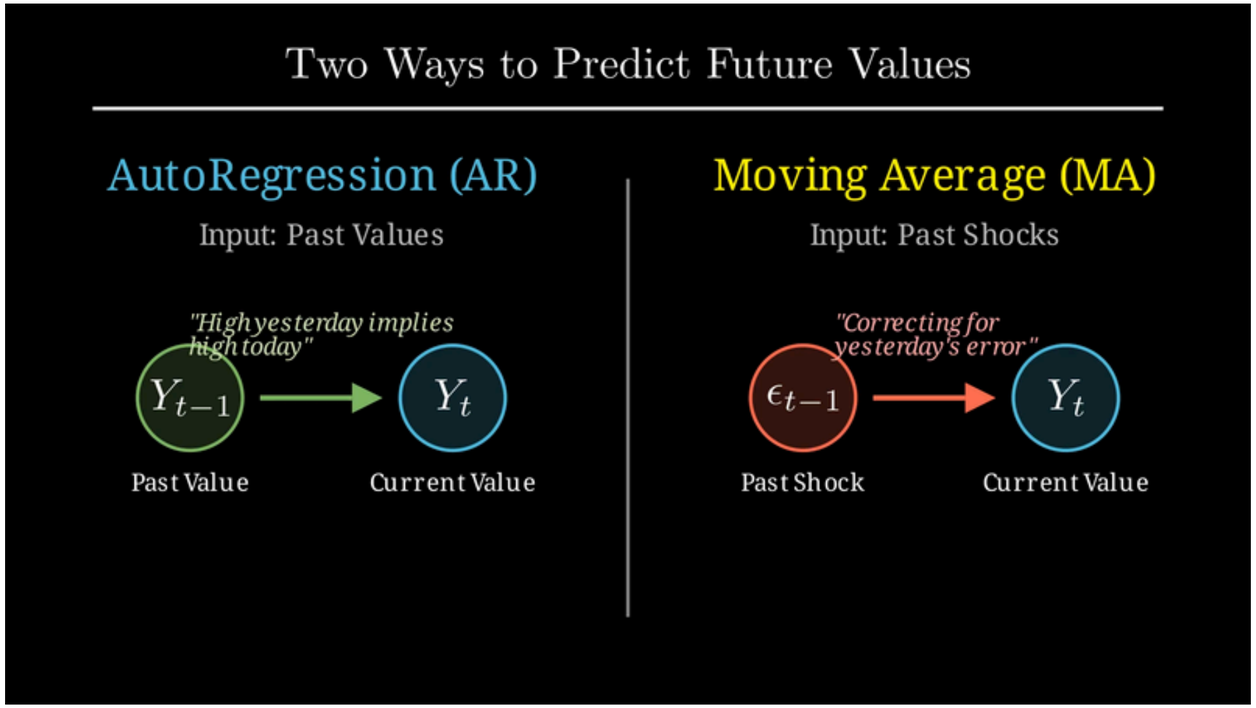
Can we use these Past Errors to improve our Current Prediction?

Moving Average (MA)

Newton School of Technology

- **Moving Average (MA)** model defines the current value of a time series as a **linear combination** of the **series mean** and **past white noise error terms** (shocks).

$$y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}$$



Moving Average

Newton School of Technology

- Coefficient (How much of yesterday's shock persists today?)

$$y_t = \mu + \epsilon_t + \theta_1 \epsilon_{t-1} + \ldots + \theta_q \epsilon_{t-q}$$

Current Target Value

The average

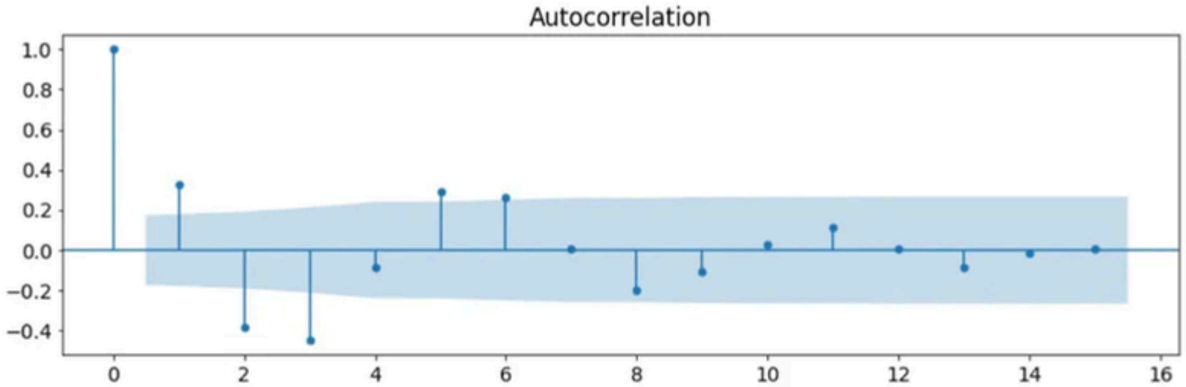
Current Noise

Past Shock (The error we made at t-1)

Moving Average

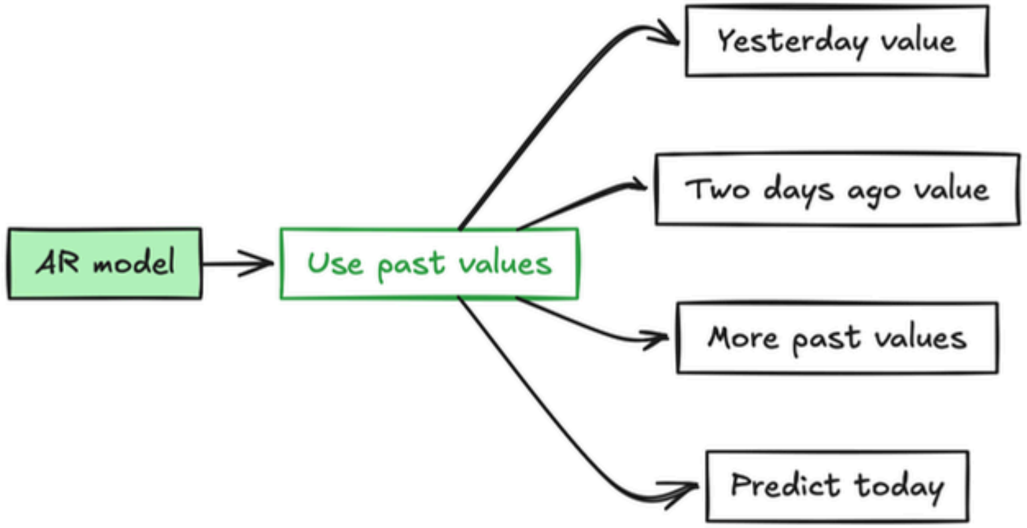
Newton School of Technology

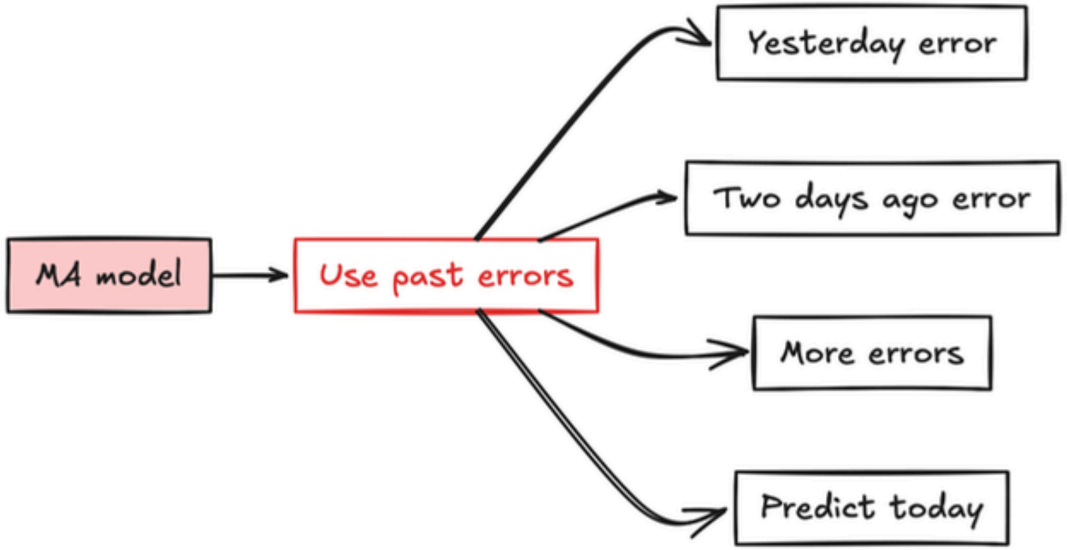
- To find which **past shocks matter**, we look at the **ACF Plot**.
- **Rule**: If the **ACF shows** a **significant spike** at **Lag 1** and then cuts off, it suggests an **MA model**.
- Key Difference:
  - **AR Model - Check PACF**
  - **MA Model - Check ACF**



AR vs MA

Newton School of Technology





Newton School of Technology

Both **AR and MA** assume the data is **Stationary**
(The **mean and variance stay constant**).
Most data has **Trends or Seasonality**. If the mean is constantly **changing**,
our **previous formulas break**.

Integration

Newton School of Technology

- Integration is the process that removes non-stationarity.
- A time series is said to be **integrated** of order d—written as **I(d)** - if:
  - You must difference it d times
  - Before it **becomes stationary** (constant mean, variance, autocovariance)
- EXAMPLE:
  - A series that becomes stationary after one differencing is I(1).
  - One that needs two differencing operations is I(2).

Differencing

Newton School of Technology

Day (t)	Sales X(t)	Lag1 X(t−1)	Diff1 = X(t) −X(t−1)	Diff2 (Difference of Diff 1)	Diff3 (Differnce of Diff 2)
1	10	–	–	–	–
2	20	10	10	–	–
3	40	20	20	10	–
4	70	40	30	10	**0**
5	110	70	40	10	**0**

- **Order d = smallest differencing level that makes the differenced series stationary (stable).**
- Diff1 increases (10, 20, 30, 40)
- Diff2 becomes constant (10, 10, 10)
- Diff3 becomes 0 (0, 0)
- **Order is 2**

Differencing

Newton School of Technology

t	X(t)	Lag2 = X(t-2)	Diff1 = X(t)-X(t-1)		Diff2 = Diff of Diff 1 X(t)-2X(t-
1	10				
2	25		15		
3	40	10	15	15	0
4	60	25	20	15	5
5	85	40	25	20	5

How to detect the needed level of integration?

Newton School of Technology

- **Visual inspection**
  - Does the series have a trend? → Likely I(1)
- **Statistical tests**
  - ADF (Augmented Dickey–Fuller) test
  - KPSS test
  - PP test

ARIMA

Newton School of Technology

- We have learned three separate concepts:
  - **AR (AutoRegressive):** Predicting future values using past values.
  - **I (Integrated):** Making data stationary by differencing.
  - **MA (Moving Average): P**redicting future values using past errors (shocks).

- **ARIMA** stands for **AutoRegressive Integrated Moving Average.**

The ARIMA Equation

Newton School of Technology

- **Differencing**: First, we calculate $y_t^{'}$ (the differenced data) to ensure stationarity.
- Equation: Then, we predict $y_t^{'}$ using both past values and past errors

$$y'_t = c + \underbrace{\phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p}}_{\text{AR terms}} + \underbrace{\theta_1 \epsilon_{t-1} + \cdots + \theta_q \epsilon_{t-q}}_{\text{MA terms}} + \epsilon_t$$

Three Components of ARIMA

Newton School of Technology

- An ARIMA model is defined by three parameters: ARIMA(p, d, q).
  - **p (AR order)**: The number of lag observations included in the model (lag order).
    - **Question**: How many past days do I look at?
  - **d (Integration order)**: The number of times that the raw observations are differenced.
    - **Question**: How many times did I subtract to make the mean stable?
  - **q (MA order)**: The size of the moving average window (order of moving average).
    - Question: How many past error shocks do I include?

Newton School of Technology

Please fill the feedback form.

Newton School of Technology

Thank You

Newton School of Technology

Join the lecture online on your dashboard.

Newton School of Technology

Let’s start with a minute of silence.

आचार्यात् पादं आधत्ते पादं शिष्यः स्वमेधया ।
पादं सब्रह्मचारिभ्यः पादं कालक्रमेण च ॥

Newton School of Technology

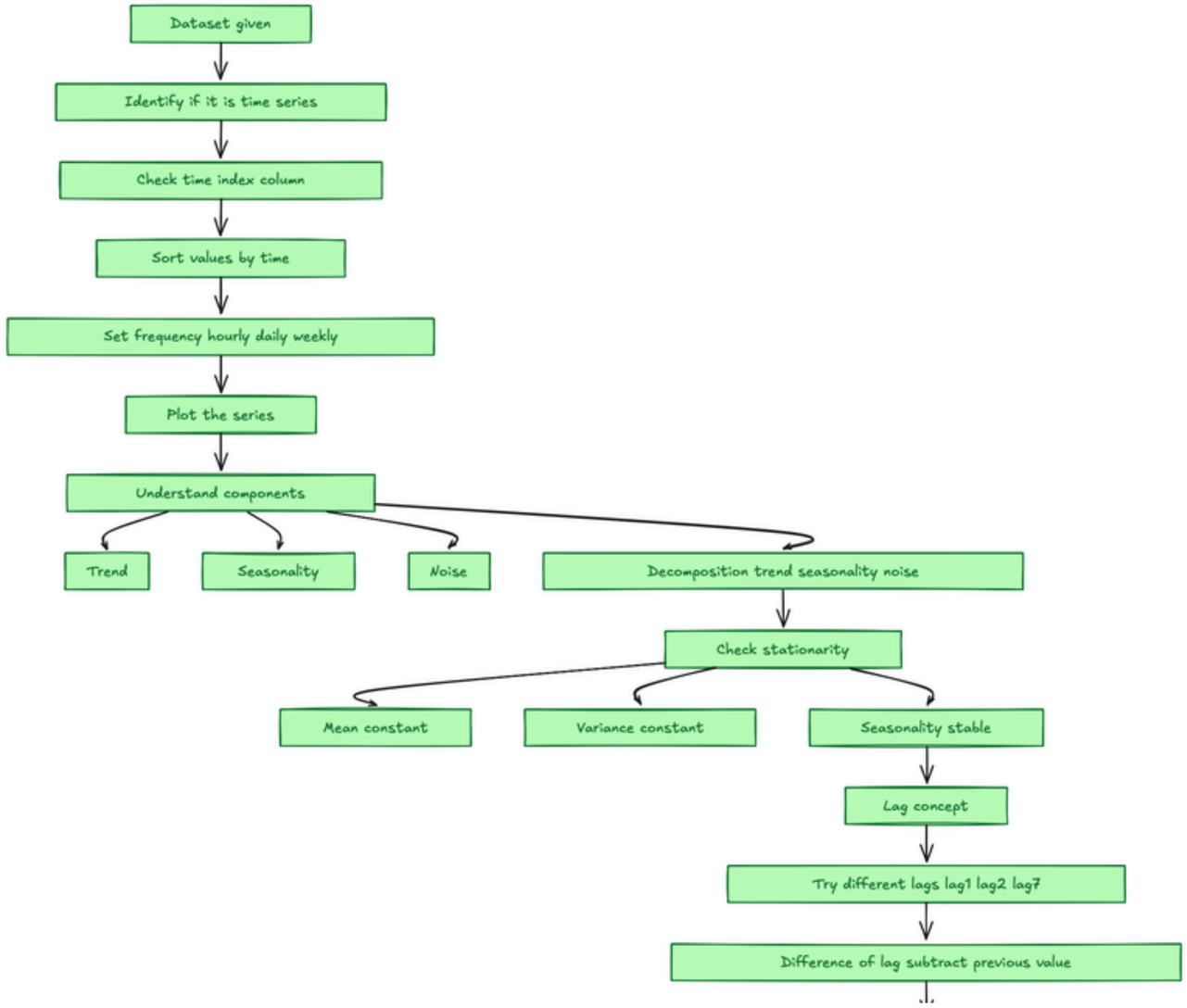
Meaning:

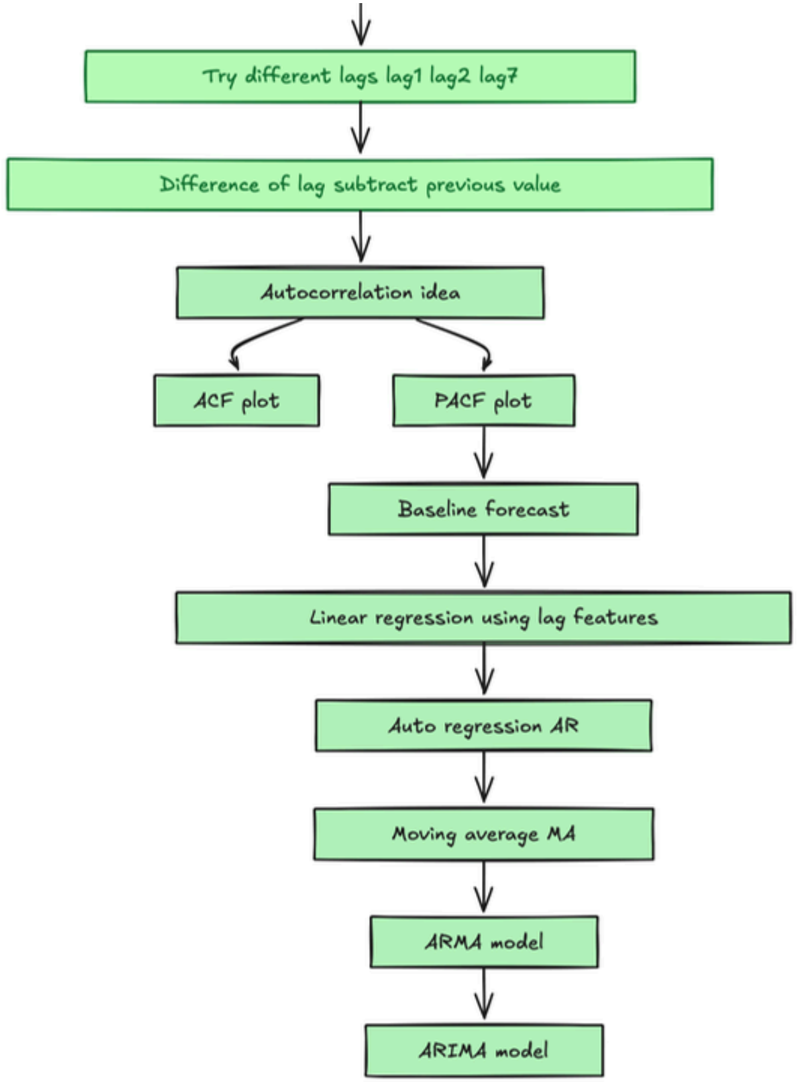
A student acquires knowledge in four equal parts:

- One-fourth from the teacher
- One-fourth through self-reflection and independent thinking
- One-fourth through discussions with peers and fellow learners
- One-fourth over time through personal experience

Recap

Newton School of Technology





Agenda

- ANNOVA
- Problems on ARIMA
- SARIMA
- Case Study

Newton School of Technology

Where do p, d, and q come from?

Newton School of Technology

- **d (differencing order) is chosen by repeatedly differencing the series until the mean becomes stable**

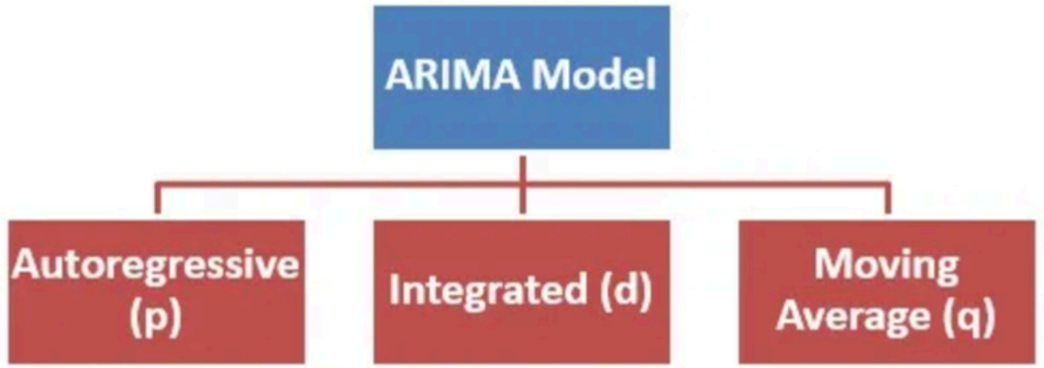
→ checked using the time plot, rolling mean/std, and ADF test

- **p (AR order) is chosen from the PACF plot**

→ count the number of significant spikes before PACF cuts off

- **q (MA order) is chosen from the ACF plot**

→ count the number of significant spikes before ACF cuts off



AR(p): depends on past values ⇒ use PACF (direct value impact)
MA(q): depends on past shocks/errors ⇒ use ACF (shock memory)

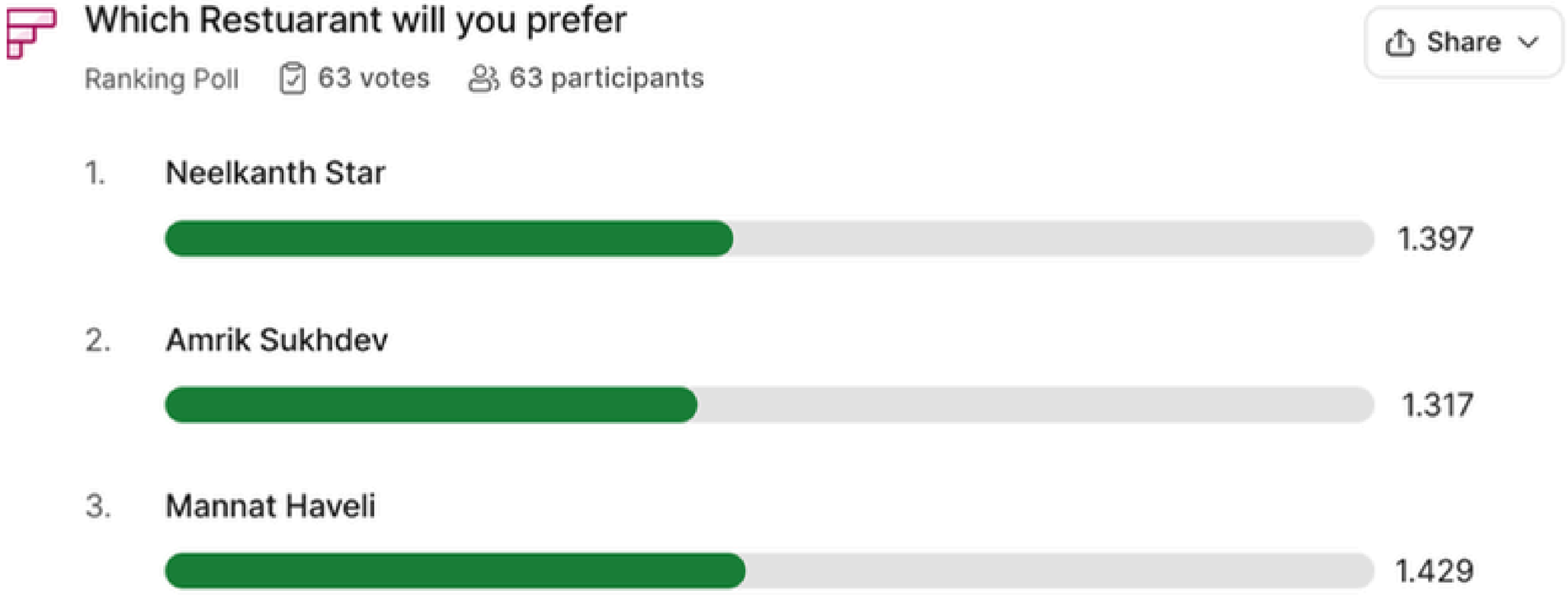
Which Dhaba will you prefer?

Newton School of Technology

1449790

Which Dhaba will you prefer?

Newton School of Technology



Problem

Newton School of Technology

Are these restauratns **really different**?
Or are these **small differences** just because **people's mood** changes daily?

Cafe	Average Rating
Neelkanth Star	7.1
Amrik Sukhdev	7.4
Mannat Haveli	7.2

Problem

Newton School of Technology

Are these restauratns **really different**?

- At first glance, Amrik Sukhdev looks best because it has the **highest average**.
- But think carefully:
  - Your mood changes every day
  - Sometimes paratha tastes better, sometimes worse
  - The difference between 7.1 and 7.4 is minimal.
- That difference might exist even if all dabha's are equally good.

Cafe	Average Rating
Neelkanth Star	7.1
Amrik Sukhdev	7.4
Mannat Haveli	7.2

"Is this difference larger than what random variation normally produces?"

Analysis of Variance(ANOVA)

Newton School of Technology

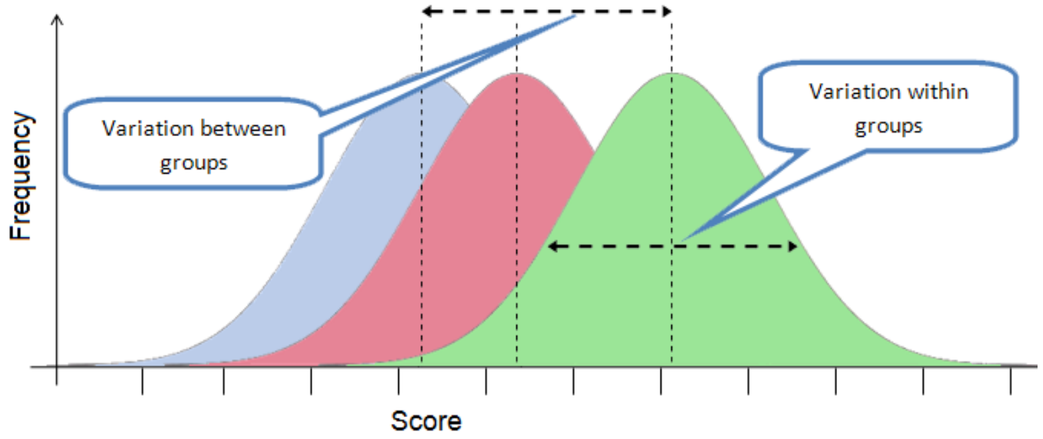
- ANOVA gives a scientific answer to:
  - Are the differences between **groups** larger than **normal random variation**?
- Use ANOVA when:
  - You **compare 3 or more groups**
  - You **measure numbers (ratings, marks, sales)**
  - You want one **decision** instead of many **guesses**.

Cafe	Average Rating
Neelkanth Star	7.1
Amrik Sukhdev	7.4
Mannat Haveli	7.2

ANOVA

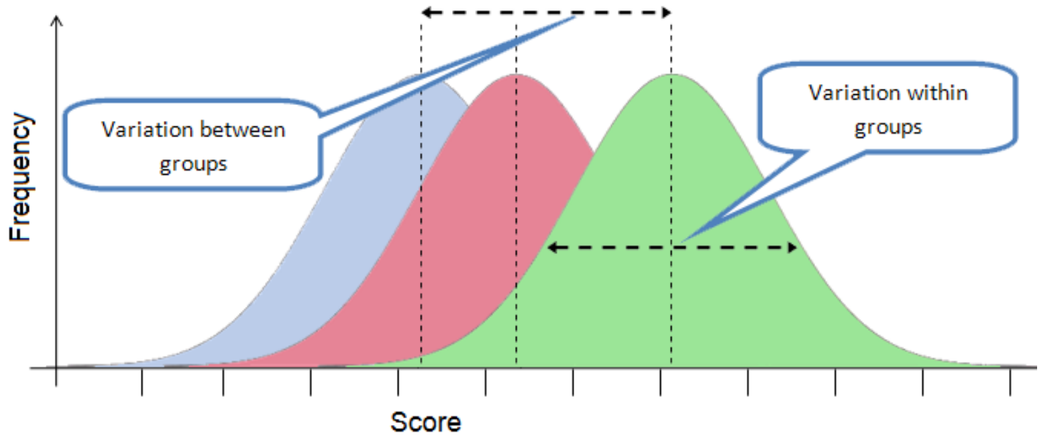
- ANOVA compares two types of variation:
  - **Within-group variation**
    - How much ratings change inside the same Dhabha
  - **Between-group variation**
    - How much Dhabha averages differ from each other

- If **Dhabhas** are truly different:
  - Between-group variation will be much larger



What is “variation”?

- How much do values spread out from an average?
- We measure variation using the sum of squares (SS).



$$SS_{Total} = \sum_{i=1}^{N} (x_i - \bar{x})^2$$

Where:

- x_i = each observation
- $\bar{x}$ = grand mean (mean of all observations)
- N = total number of observations

How spread out is data from the overall average?

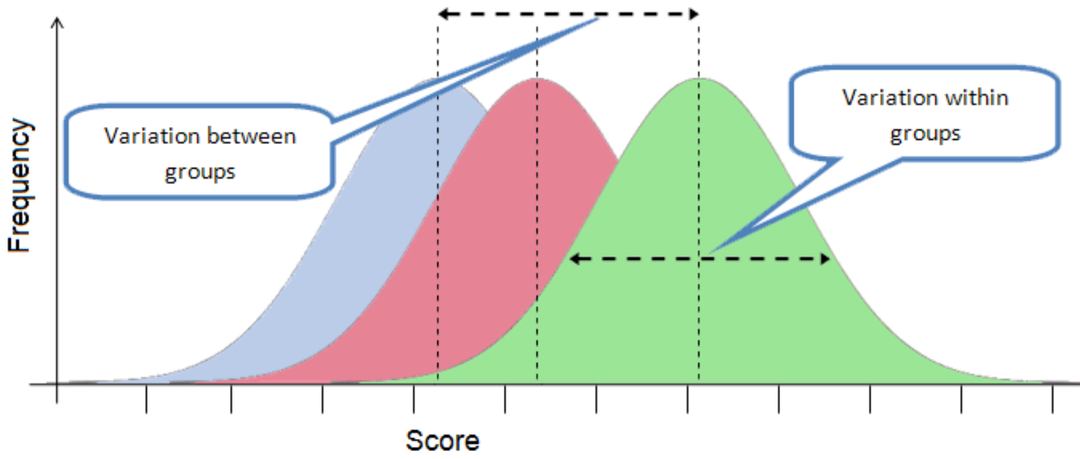
Between-Group Variation

Newton School of Technology

- It calculates the **distance between each group's** individual **mean** and the **overall mean** of the entire dataset (the "Grand Mean").
- The between-group variation is denoted by :

$$SS_{between} = \sum n_i(\bar{x}_i - \bar{x})^2$$

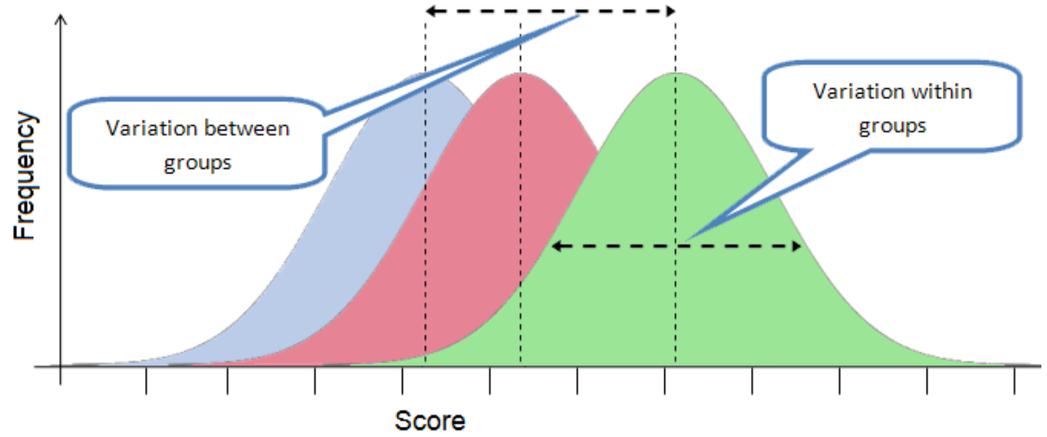
- n_i = size of group i
- $\bar{x}_i$ = mean of group i
- $\bar{x}$ = overall mean



Within-Group Variation

- It tells us how much values vary within the same group.
- It is represented as

$$SS_{within} = \sum\sum(x_{ij} - \bar{x}_i)^2$$



The ANOVA F-Statistic Formula

Newton School of Technology

$$F = \frac{\text{Variance Between Groups}}{\text{Variance Within Groups}}$$

$$F = \frac{MS_{between}}{MS_{within}}$$

- Higher F ratio values indicate the variation between groups is larger than the individual variation of groups.
- In such cases, it is more likely that the mean of the groups are different.

The ANOVA F-Statistic Formula

Newton School of Technology

$$MS_{between} = \frac{SS_{between}}{df_{between}}$$

Degrees of freedom:

$$df_{between} = k - 1$$

(k = number of groups)

The ANOVA F-Statistic Formula

Newton School of Technology

$$MS_{within} = \frac{SS_{within}}{df_{within}}$$

Degrees of freedom:

$$df_{within} = N - k$$

(N = total observations)

Why ANNOVA?

Newton School of Technology

A time series can be seen as:

$$y_t = (\text{level}) + (\text{seasonality}) + (\text{trend}) + (\text{noise})$$

ANOVA helps answer:

Is there evidence that the level differs across groups like:

- Month (Jan vs Feb vs …)
- Day-of-week (Mon vs Tue vs …)
- Pre vs Post intervention
- Festival vs non-festival periods
- Before policy change vs after
- So it becomes a mean comparison tool for time-structured categories.

Annova is used

- To detect seasonality
- To test impact of an intervention / change point (A/B in time)
- To compare multiple time-based regimes

How to Use ANNOVA

Newton School of Technology

How to use ANOVA correctly in time-series workflows

ANOVA on residuals

Common approach:

1. Fit time-series structure first (trend/ARIMA)
2. Take residuals (should be uncorrelated)
3. **Apply ANOVA on residuals grouped by season/month/etc.**

ANOVA is useful in time series when you want to test whether the mean level changes across time-based groups or regimes (seasonality, interventions), but you must account for autocorrelation to avoid misleading significance.

Error

- Predictions without error measurement are meaningless.
- It is essential to recognize our mistakes.
- Always when:
  - You make predictions
  - Decisions depend on correctness.

Error

Newton School of Technology

Situation	Risk
Only average error is low	Rare disasters ignored
Only percentage error used	Breaks when values ≈ 0
Symmetric error assumed	Underprediction may be costlier

Mean Absolute Error (MAE)

Newton School of Technology

- It tells us, on average, **how far off** we are.
- **Interpretation**
  - **Same unit** as data
  - Easy to **explain to non-technical stakeholders**
- **Strengths**
  - Robust to **outliers**
  - **Intuitive**
- **Limitations**
  - Treats all **errors equally**
  - Does not emphasize **big mistakes**

$$MAE = \frac{1}{n}\sum |y - \hat{y}|$$

Mean Squared Error (MSE)

Newton School of Technology

- It tells us how costly our mistakes are if big errors matter more.
- **Interpretation**
  - **Squaring** increases the **penalty** for **large errors**
  - Used heavily in **optimization**
- **Strengths**
  - Strong penalty for big errors
- **Limitations**
  - Hard to interpret (squared units)
  - Hard to interpret (squared units)

$$MSE = \frac{1}{n}\sum(y - \hat{y})^2$$

Root Mean Squared Error (RMSE)

Newton School of Technology

- It tells us how large errors are when big mistakes matter
- **Interpretation**
  - Same **unit** as data
  - Highlights volatility in prediction quality
- **Strengths**
  - Penalizes big errors
  - Interpretable scale
- **Limitations**
  - sensitive to outliers

$$RMSE = \sqrt{MSE}$$

Problem

Newton School of Technology

- Suppose we are **forecasting** daily **electricity demand** (in units) for 5 days.
- Question - Compute the **Mean Absolute error** for this data.

Day	Actual Demand	Predicted Demand
1	100	98
2	120	125
3	130	128
4	150	140
5	170	180

Solution

Newton School of Technology

Day	Actual Value	Predicted Value	Error = Actual - Predicted
1	100	98	2
2	120	125	−5
3	130	128	2
4	150	140	10
5	170	180	−10

$$MAE = \frac{|2| + |-5| + |2| + |10| + |-10|}{5} = 5.8$$

Comparing 2 Models

Newton School of Technology

- Suppose we have actual sales for 5 days and predictions from two models:
- Model A
- Model B

Which Model is better?

Day	Actual Sales	Model A Prediction	Model B Prediction
1	100	95	100
2	120	118	110
3	130	125	140
4	150	145	120
5	170	165	200

Comparing 2 Models

Newton School of Technology

$$MAE_A = \frac{5 + 2 + 5 + 5 + 5}{5} = 4.4$$

$$MAE_B = \frac{0 + 10 + 10 + 30 + 30}{5} = 16$$

- Model A is much more reliable on average.

Day	Actual Sales	Model A Prediction	Model B Prediction
1	100	95	100
2	120	118	110
3	130	125	140
4	150	145	120
5	170	165	200

Comparing 2 Models

Newton School of Technology

$$RMSE_A = \sqrt{\frac{25 + 4 + 25 + 25 + 25}{5}} = 4.456$$

$$RMSE_B = \sqrt{\frac{0 + 100 + 100 + 900 + 900}{5}} = 20$$

- Model B has dangerously large errors, even if sometimes perfect.

Day	Actual Sales	Model A Prediction	Model B Prediction
1	100	95	100
2	120	118	110
3	130	125	140
4	150	145	120
5	170	165	200

Comparing 2 Models

Newton School of Technology

- Model A behaves predictably.
- Model B behaves erratically.
- Model A will be selected because
  - Lower average error
  - No catastrophic failures
  - Easier to plan inventory and staffing

Day	Actual Sales	Model A Prediction	Model B Prediction
1	100	95	100
2	120	118	110
3	130	125	140
4	150	145	120
5	170	165	200

Seasonal ARIMA (SARIMA)

Newton School of Technology

- Regular ARIMA learns:
  - **Trend**
  - Short-term dependence
- But it **cannot remember** repeating **cycles** like:
  - Weekly routines
  - Monthly patterns
  - Yearly seasons
- **Seasonal ARIMA** adds **memory** of the same time in **previous cycles**.

SARIMAX Model

Newton School of Technology

SARIMAX = SARIMA + eXogenous variables

- Used when data has trend, seasonality, and external factors
- **It models:**
  - Past values (AR)
  - Past errors (MA)
  - Differencing (I)
  - Seasonal patterns (S)
  - External variables (X)
- **Best suited when:**
  - Data shows seasonality (daily, monthly, yearly)
  - External factors influence the series
  - Simple ARIMA is not sufficient

Other time series models

Newton School of Technology

ARIMAX Model

- Used when the time series is influenced by external factors
- **Uses:**
  - Past values of the series
  - Past errors
  - Other related variables (X)
- Best suited when:
  - External variables affect the data
  - Better accuracy than ARIMA is needed
  - Future X values are available

VAR Model (Vector Autoregression)

Newton School of Technology

- **VAR models multiple time series together**
  - Each variable is a function of its own past values and the past values of all other variables
- All variables are treated as endogenous (no strict input/output)

- **Best suited when:**
  - Several time series influence each other
  - Variables move together over time
  - System-level forecasting is required

From ARIMA to LSTM

Newton School of Technology

- In time series, we always try to use the past to predict the future
- AR, MA, ARIMA do this using a fixed number of past days and errors

But in real data, we often don't know:

- how far back to look
- which past patterns really matter

LSTM is a deep learning model that learns this automatically

- It keeps a memory of the past and decides:
  - what to remember
  - what to forget
  - what is important for prediction

Problem to solve

Newton School of Technology

We are given the following time series data for variable X_t and the corresponding white noise/error terms ϵ_t:

Time t	X_t	ϵ_t (White Noise/Error Term)
1	5	0.5
2	6	1.0
3	7	-0.3
4	?	0.4

We need to predict X_4 using the following models:

Newton School of Technology

1. Autoregression (AR(1)) model:

The AR(1) model is given by:

$$X_t = c + \phi_1 X_{t-1} + \epsilon_t$$

Where:

- c is the intercept term,
- ϕ_1 is the autoregressive parameter,
- ϵ_t is the white noise error term for time t.

2. Moving Average (MA(1)) model:

The MA(1) model is given by:

$$X_t = c + \theta_1 \epsilon_{t-1} + \epsilon_t$$

Where:

- c is the intercept term,
- θ_1 is the moving average parameter,
- ϵ_{t-1} is the previous error term.

3. ARIMA(1,1,1) model:

The ARIMA(1,1,1) model incorporates autoregression, differencing, and moving average. For this example, we assume $\phi_1 = 0$. The model is given by:

$$\Delta X_t = c + \theta_1 \epsilon_{t-1} + \epsilon_t$$

Where:

$$\Delta X_t = X_t - X_{t-1} \quad \text{is the differenced series.}$$

Task:

- Calculate the intercept term c along with the coefficients ϕ_1 and θ_1 for each model.
- Using the models, predict the value of X_4.

Problem

Consider the following bivariate time series data:

t	X_t	Y_t
1	2	1
2	4	2
3	5	3
4	8	6
5	X_5	Y_5

Assume that the variable X_t follows a **first-order Vector Autoregressive structure (VAR(1))** with a non-zero intercept and zero error terms.

Model Specification

The model for X_t is given by:

$$X_t = C + \phi_1 X_{t-1} + \phi_2 Y_{t-1}, \quad t \geq 2,$$

where:

- C is the intercept term,
- ϕ_1 and ϕ_2 are autoregressive coefficients,
- the error term is assumed to be zero for simplicity.

Task

1. Use the given observations to estimate the parameters C, ϕ_1, and ϕ_2.
2. Using the estimated model, compute the value of X_5.

Case Study

Newton School of Technology

🧪 Urban Air Quality & Public Health Forecasting

🇺🇸 US Government Time Series Case Study

🧑‍🔬 Your Role: Government Data Scientist

You have just joined a **US Government Public Health & Climate Analytics Unit**.

Over the past few years, hospitals across major US cities have reported a steady rise in:

- Heat stroke cases
- Breathing disorders
- Asthma attacks
- Cardiac emergencies

Government officials suspect that **weather conditions and air quality are major hidden drivers** behind these public health problems.

To investigate this, the government has provided you with a nationwide dataset called:

> 📁 **Urban Air Quality and Health Impact Dataset**

It contains daily environmental and weather records along with a computed **Health_Risk_Score**.

Newton School of Technology

Please fill the feedback form.

Newton School of Technology

Thank You