Nowcasting

Muhammad Apriandito A.S.

Agenda

1. Introduction to Nowcasting
2. Nowcasting vs. Forecasting
3. Applications of Nowcasting
4. Requirements for Nowcasting
5. Model Evaluation
6. Practical Session

What is Nowcasting?

- **Definition**: Nowcasting is the prediction of the present, the very near future
- **Key Difference**:
  - **Forecasting**: Predicts future events (e.g., weather next week).
  - **Nowcasting**: Focuses on the current or immediate future (e.g., weather right now or in the next few hours).
- **Common Confusion**:
  - *"Forecasting can also predict short-term events, so why nowcasting?"*
  - **Answer**: Nowcasting uses real-time data and high-frequency indicators to provide more accurate and timely insights than traditional forecasting.

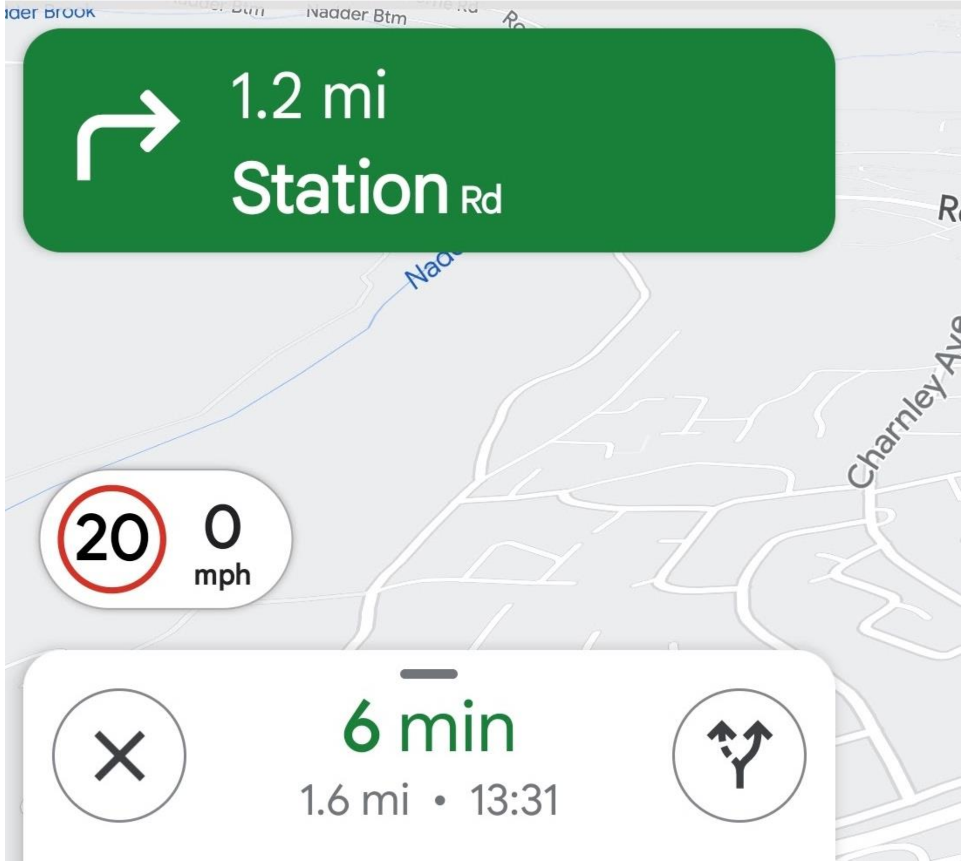
Nowcasting vs. Forecasting

Aspect	**Nowcasting**	**Forecasting**	**Key Point**
Time Horizon	**Present or very near future** (e.g., hours, days).	**Short to long-term future** (e.g., weeks, months, years).	Forecasting can predict short-term, but nowcasting focuses on **real-time or immediate future**.
Data Used	**Real-time, high-frequency data** (e.g., live traffic, stock prices).	**Historical and periodic data** (e.g., monthly GDP, annual reports).	Nowcasting uses **live data** for accuracy, while forecasting relies on **historical trends**.
Purpose	**Immediate decision-making** (e.g., adjust routes on Google Maps now).	**Strategic planning** (e.g., budget for next quarter).	Nowcasting is for **urgent actions**, while forecasting is for **long-term strategies**.
Accuracy	**Highly accurate for the present** due to real-time data.	**Less accurate for the present** due to reliance on lagged data.	Forecasting's short-term predictions are **less precise** than nowcasting's real-time insights.
Example	Predicting **today's rainfall** using live satellite data.	Predicting **next month's average rainfall** using historical weather patterns.	Even if forecasting predicts short-term, it **lacks the immediacy and precision** of nowcasting.

When and Why Do We Need Nowcasting?

- **Everyday Example**: Google Maps.
  - **Problem**: Predicting traffic conditions in real-time.
  - **Solution**: Nowcasting uses live data (e.g., GPS from cars) to update traffic conditions instantly.

- **Why Nowcasting?**
  - Real-time decision-making.
  - High accuracy for immediate needs.
  - Adaptability to rapidly changing conditions.



Nowcasting for Economics

- **Problem**: Economic indicators (e.g., GDP, unemployment) are often released with a lag.
- **Urgency**:
  - Policymakers need real-time data to make informed decisions.
  - Businesses need timely insights to adjust strategies.
- **Example**: Predicting quarterly GDP before official data is released.

Central Bank Challenges: Delayed Speedometer

- **Analogy**: Imagine driving a car with a **delayed speedometer**.(Updated every 5 second)
  - You’re driving at 100 km/h, but the speedometer shows 80 km/h (from 5 seconds ago).
  - **Problem: You can’t react in real-time to avoid accidents or adjust speed.**
- **Central Bank Challenge**:
  - Economic data (e.g., inflation, GDP) is like a **delayed speedometer**.
  - Decisions (e.g., interest rates) are made on outdated information.

- ***Solution**: Nowcasting acts like a **real-time speedometer**, providing up-to-date economic snapshots for timely policy adjustments.*

What Do We Need for Nowcasting?

1. Data
2. Model (Predictive Algorithm)

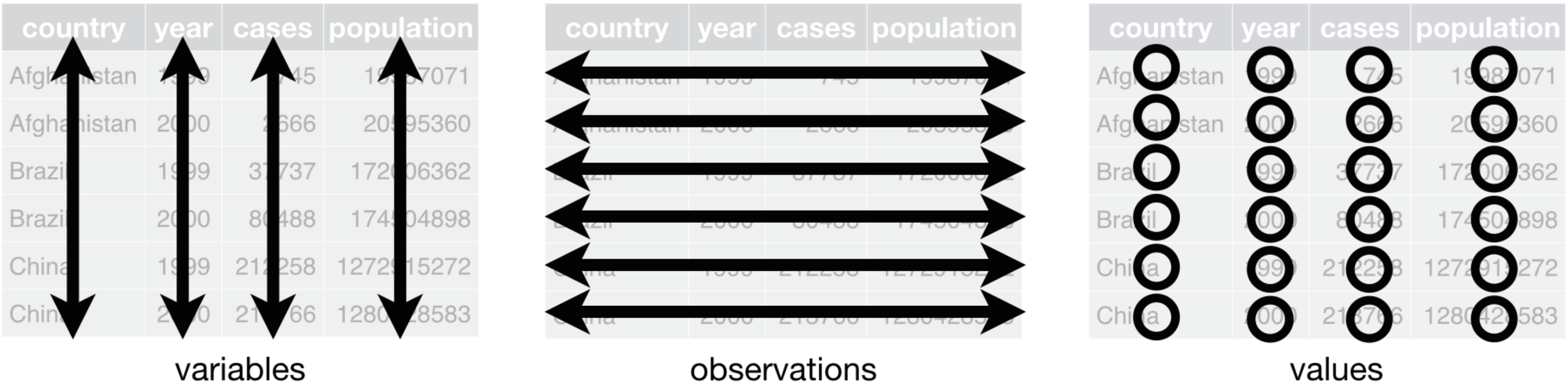
Data for Nowcasting

Criteria	Description	Example for Economic Nowcasting
Timeliness	Data must be available in real-time or with minimal delay.	Daily stock market data, weekly unemployment claims.
Granularity	High-frequency data is preferred for accurate predictions.	Monthly retail sales data vs. quarterly GDP reports.
Relevance	Data should directly relate to the economic indicator being predicted.	Industrial production data for GDP nowcasting.
Quality	Data must be accurate, complete, and free from errors.	Cleaned and validated consumer price index (CPI) data.
Consistency	Data collection methods should remain consistent over time.	Uniformly collected labor market statistics.
Frequency	Higher data frequency improves nowcasting accuracy.	Weekly jobless claims vs. monthly employment reports.

Format: Tidy Data

Tidy data is a structured format for organizing datasets where:

1. **Each variable is a column** (e.g., GDP, inflation).
2. **Each observation is a row** (e.g., a specific month or country).
3. **Each value is a single cell** (e.g., GDP for January 2025).

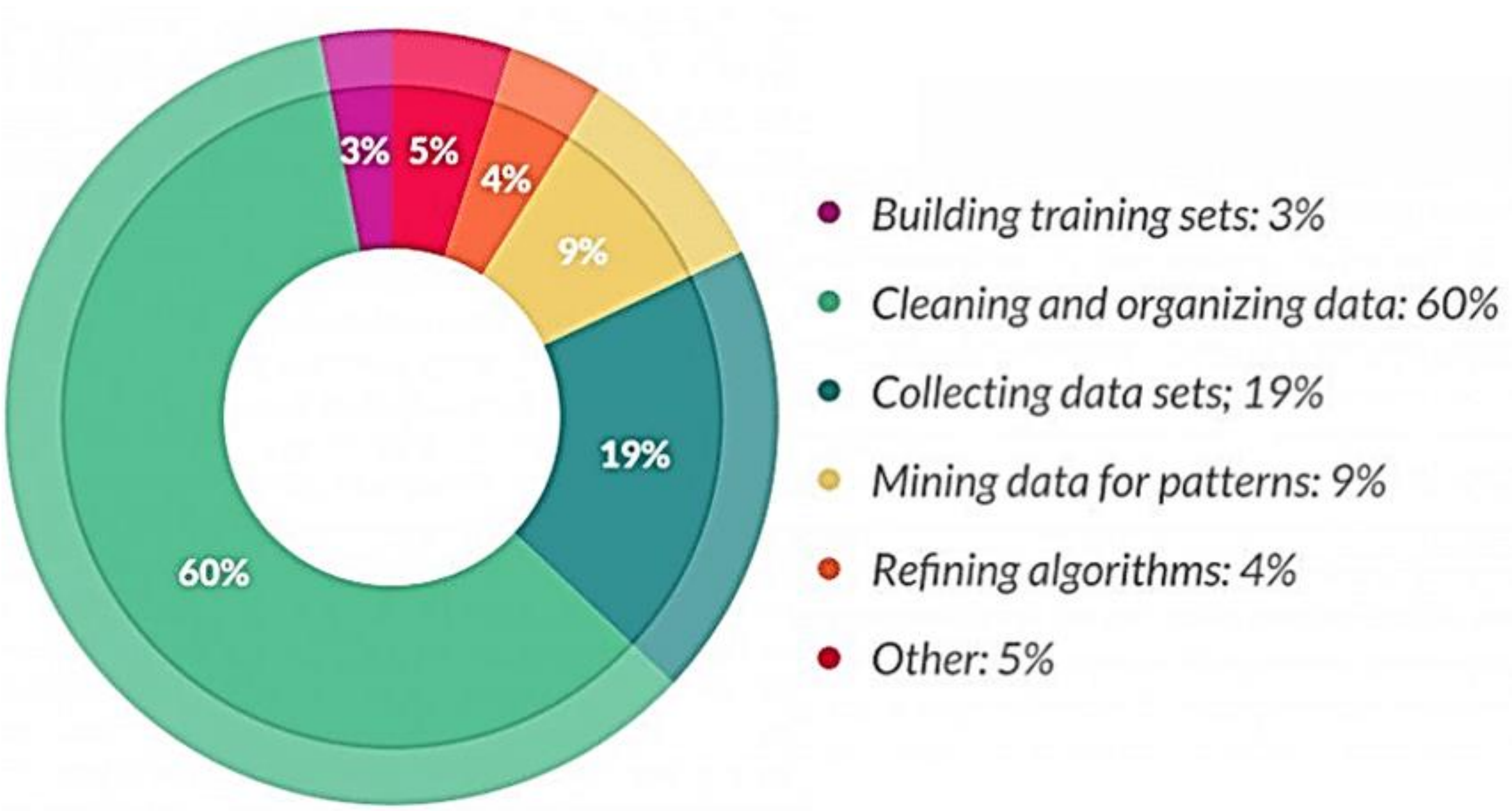


Problem

	Inflasi (2018=100) Menurut Kelompok dan Sub Kelompok 01 Makanan, Minuman dan Tembakau												
	Makanan, Minuman dan Tembakau												
													2023
90 Kota Inflasi (2018=100)	Januari	Februari	Maret	April	Mei	Juni	Juli	Agustus	September	Oktober	November	Desember	Tahunan
KOTA MEULABOH	1.67	0.79	-2.39	-2.2	1.32	1.25	0.82	-0.62	1.06	0.66	0.37	-2.01	-
KOTA BANDA ACEH	1.64	1.8	-1.23	0.15	0.37	0.75	0.87	-1.29	0.1	-1.01	1.39	-0.08	-
KOTA LHOKSEUMAWE	1.56	1.16	-1.29	-1.44	1.4	0.07	-0.43	-0.33	0.53	-0.67	0.57	0.27	-
KOTA SIBOLGA	2.94	0.62	-2.15	-1.34	0.37	1.64	0.81	-0.03	1.05	-0.55	-0.63	-0.01	-
KOTA PEMATANG SIANTAR	2.17	-0.71	-1.04	-0.81	1.02	1.39	0.02	-1.31	0.27	-0.52	0.83	1.02	-
KOTA MEDAN	2.23	-0.83	-1.16	-0.99	0.81	1.06	0.21	-0.25	0.53	-0.4	1.26	1.33	-
KOTA PADANGSIDIMPUAN	2.62	0.91	-0.62	-1.19	1.02	0.89	1.18	-0.9	1.18	-0.44	-0.36	0.27	-
KOTA GUNUNGSITOLI	3.21	-1.23	-2.1	-1.35	0.27	-0.09	1.53	0.22	0.66	-1.03	1.74	0.85	-
KOTA PADANG	1.93	0.78	-0.77	-0.59	0.51	0.75	0.21	0.24	0.11	-0.3	1.9	0.34	-
KOTA BUKITTINGGI	1.11	0.14	-0.57	-0.49	0.86	-0.06	0.51	-0.16	1.28	-0.56	1.47	-0.3	-
TEMBILAHAN	1.28	0.07	-0.48	0.48	0.68	0.34	1.11	-1.47	0.19	-0.47	1.23	0.08	-
KOTA PEKANBARU	1.69	0.38	-0.72	-0.88	0.3	0.75	0.89	-0.12	0.83	-0.31	2.13	0.3	-
KOTA DUMAI	2.19	-0.35	-0.19	-0.95	0.61	0.58	0.89	-0.52	0.96	-1.03	2.12	-0.21	-
BUNGO	2.17	0.51	-0.58	-0.13	0.99	0.37	0.79	-1.04	0.79	-0.15	2.45	0.81	-
KOTA JAMBI	3.07	-0.66	-0.36	-0.37	0.25	1.47	0.92	-1.45	0.96	1.19	2.21	0.53	-
KOTA PALEMBANG	1.32	-0.08	0.36	0.14	0.57	1.24	0.72	-0.61	1.07	0.41	1.74	0.25	-
KOTA LUBUKLINGGAU	1.06	-0.07	0.18	0.62	0.72	0.64	0.68	-0.45	0.92	0.39	0.9	0.31	-
KOTA BENGKULU	1.11	-0.03	0.34	0.27	0.62	0.63	0.91	-0.75	0.57	0.78	0.93	0.7	-
KOTA BANDAR LAMPUNG	2.41	0.37	0.1	0.38	0.18	0.52	-0.17	0.89	0.77	0.58	3.02	-0.02	-
KOTA METRO	2.65	0.62	-0.61	0.61	0.8	0.11	0.61	-0.4	0.62	1.29	2.81	0	-
TANJUNG PANDAN	4.84	-0.93	1.22	-0.77	2.95	0.47	-0.18	-0.48	3.16	-1.74	-0.7	-1.47	-
KOTA PANGKAL PINANG	0.62	-0.38	0.54	0.8	0.33	1.67	-0.75	-0.18	1.81	-0.22	-0.41	0.48	-
KOTA BATAM	-0.71	2.3	-1.35	-0.24	0.3	1.24	0.2	0.52	-0.15	0.44	2.03	0.92	-
KOTA TANJUNG PINANG	-0.25	1.6	-0.69	-0.66	-0.57	1.33	0.51	0.11	0.54	-0.03	2.16	0.07	-
DKI JAKARTA	0.94	0.57	0.6	0.52	0.13	0.2	0.29	-0.25	0.55	0.04	1.08	1.55	-
KOTA BOGOR	1.53	1.3	0.17	0.48	0.74	0.38	0.49	-0.64	-0.21	0.12	2.04	0.5	-
KOTA SUKABUMI	0.92	0.6	0.11	0.92	0.68	0.53	0.19	-0.09	0.3	0.11	0.92	0.43	-
KOTA BANDUNG	0.87	0.89	0.57	0.56	0.37	0.59	0.78	-0.06	0.06	-0.19	0.81	1.04	-
KOTA CIREBON	0.56	1.1	0.49	0.66	0.77	0.78	0.24	-0.54	0.41	-0.01	0.6	0.46	-
KOTA BEKASI	0.6	0.78	0.49	0.99	0.29	0.2	-0.08	0.02	-0.14	0.36	1.96	1.51	-
KOTA DEPOK	1.46	1.13	0.08	0.39	0.3	0.05	0.42	-0.5	-0.31	0.22	1.06	0.94	-

Source: *https://www.bps.go.id/id/statistics-table/2/MTg5MCMy/inflasi--2018-100--menurut-kelompok-dan-sub-kelompok-01-makanan--minuman-dan-tembakau.html*

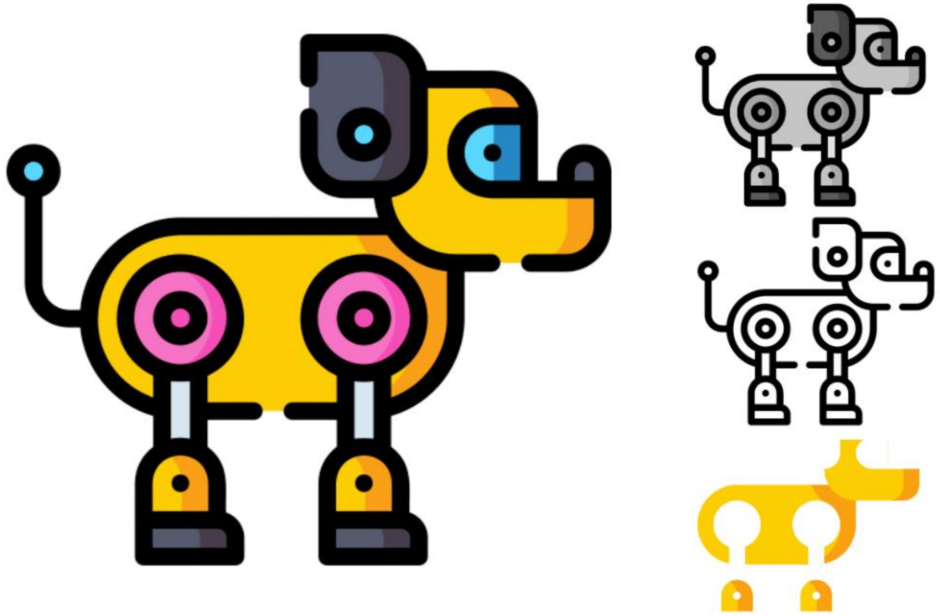
Where do we spend our time?



Source: *https://www.researchgate.net/publication/335577003_Data_preparation_and_preprocessing_for_broadcast_systems_monitoring_in_PHM_framework*

Feature Engineering

Feature engineering is the process of transforming raw data into meaningful features (variables) that improve the performance of prediction models. It involves creating, selecting, or modifying features to better represent the underlying problem, making it easier for the model to learn patterns and make accurate predictions.



Feature Engineering

Technique	Advantages	When to Use (Purpose)
Lag Features	Captures temporal dependencies by using past values as features.	When historical data is crucial for predicting future values (e.g., stock prices, demand forecasting).
Rolling Statistics	Computes moving averages, standard deviations, etc., to smooth data.	When trends or patterns over a time window are important (e.g., demand forecasting, economic indicators).
Difference Features	Calculates changes between consecutive time steps (e.g., Δ value).	When the rate of change is more important than absolute values (e.g., GDP growth, inflation rates).
Scaling/Normalization	Rescales features to a standard range (e.g., 0 to 1 or z-score).	When features have different scales and need to be comparable (e.g., multi-variable data).
Encoding Categorical Variables	Converts categorical data into numerical format (e.g., one-hot encoding).	When categorical data (e.g., regions, product types) is part of the dataset.

Model (Predictive Algorithm)

Model (Predictive Algorithm)

Category	Strengths	Weaknesses
Statistical Models	- Well-established and interpretable. - Handles time series data well. - Suitable for linear relationships.	- Struggles with non-linear relationships. - Requires strict assumptions (e.g., stationarity). - Limited scalability with big data.
Machine Learning Models	- Handles non-linear relationships well. - Scalable with large datasets. - Flexible and adaptable to various data types.	- Requires large amounts of data for training. - Less interpretable (black-box nature). - Prone to overfitting without proper tuning.
Hybrid Models	- Combines strengths of statistical and ML models. - Improved accuracy and robustness. - Can handle both linear and non-linear patterns.	- Complex to implement and tune. - Computationally expensive. - Requires expertise in both statistical and ML methods.
Big Data Approaches	- Utilizes real-time and alternative data sources. - Captures unconventional insights (e.g., sentiment, mobility). - Enhances predictive power with diverse data.	- Data quality and noise can be issues. - Requires advanced preprocessing. - Computationally intensive.
Econometric Models	- Designed for economic data and relationships. - Handles cointegration and error correction. - Provides theoretical grounding.	- Requires strong econometric knowledge. - Less flexible with non-traditional data. - Can be complex to implement.

Evaluation

Evaluating a nowcasting model is essential to:

1. **Measure Accuracy**: Ensure the model's predictions are close to actual values.
2. **Assess Reliability**: Determine if the model can be trusted for real-time decision-making.
3. **Identify Improvements**: Highlight areas where the model underperforms and needs refinement.
4. **Compare Models**: Decide which model performs best for the specific nowcasting task.

Evaluation Metric

Metric	Explanation	How to Interpret
MAE	Mean Absolute Error - Average absolute difference between predicted and actual values.	Lower MAE = Better accuracy. Represents average error in the same units as the data.
MSE	Mean Squared Error - Average squared difference between predicted and actual values.	Lower MSE = Better accuracy. Sensitive to large errors (penalizes outliers more heavily).
RMSE	Root Mean Squared Error - Square root of MSE.	Lower RMSE = Better accuracy. Easier to interpret than MSE (same units as the data).
MAPE	Mean Absolute Percentage Error - Average percentage difference between predicted and actual values.	Lower MAPE = Better accuracy. Useful for understanding error relative to the actual values (in %).
R^2	R-Squared - Proportion of variance in the target variable explained by the model.	Closer to 1 = Better fit. Indicates how well the model captures the variability in the data.

Splitting

Splitting data into training and testing sets is essential to evaluate how well a model performs on unseen data.

Imagine you're studying for an exam:

- *Training Data = Your study materials (textbooks, notes, practice problems).*
- *Testing Data = The actual exam questions.*

Splitting

Splitting Method	Explanation	When to Use
Holdout Split	Data is divided into one part for training and one part for testing, without considering time order.	For initial exploration or when the data does not have a strong time sequence.
Sliding Window (Fixed Window)	Uses a fixed-size training window that slides forward for each testing step.	When older data is considered less relevant than newer data (e.g., stock market prediction).
Expanding Window (Forward Chaining)	Training data gradually expands to include all previous data.	When all historical data is considered relevant for prediction.
Walk-Forward Validation	Similar to expanding window, but predicts one step ahead each time. Training data grows over time, while testing is always for the next time step.	For nowcasting or real-time prediction systems.

Splitting

1. Holdout Split (70-30)

	t1	t2	t3	t4	t5	t6	t7	t8	t9	t10	t11	t12
Split	Train	Train	Train	Train	Train	Train	Train	Train	Test	Test	Test	Test

2. Sliding Window (Window Size: 6)

	t1	t2	t3	t4	t5	t6	t7	t8	t9	t10	t11	t12
Fold 1	Train	Train	Train	Train	Train	Train	Test	Test				
Fold 2			Train	Train	Train	Train	Train	Train	Test	Test		
Fold 3					Train	Train	Train	Train	Train	Train	Test	Test

3. Expanding Window

	t1	t2	t3	t4	t5	t6	t7	t8	t9	t10	t11	t12				
Fold 1	Train	Train	Train	Train	Test	Test	Test	Test								
Fold 2	Train	Train	Train	Train	Train	Train	Train	Train	Test	Test	Test	Test				
Fold 3	Train	Train	Train	Train	Train	Train	Train	Train	Train	Train	Train	Train	Test	Test	Test	Test

4. Walk-Forward Validation

	t1	t2	t3	t4	t5	t6	t7	t8	t9	t10	t11	t12
Step 1	Train	Train	Train	Test								
Step 2	Train	Train	Train	Train	Test							
Step 3	Train	Train	Train	Train	Train	Test						
Step 4	Train	Train	Train	Train	Train	Train	Test					

Let’s Practice

Google Collab Link:
https://colab.research.google.com/drive/1rvyj6qA45MtzZevXQ3CLP2MZTiPFb0wy?usp=sharing

Key Takeaways

- Nowcasting is about **real-time predictions** using **live data**.
- It's crucial for **immediate decision-making** in fields like economics and traffic management.
- **Data quality** and **model selection** are critical for accurate nowcasting.
- **Evaluation** and **data splitting** ensure the model performs well on unseen data.

Thank You

See You Again!