2025

4Geeks Academy: data science cohort 12

DAY 26: TIME SERIES

TODO

TIME SERIES DATA	Times in Python and Pandas, forecasting
TIME SERIES PROJECT	Work on Alternative time series project (Time series module), plan to finish over the weekend
CLUSTERING PROJECT	Submit K-means Project Tutorial (Unsupervised Learning Module), if you haven't already

TOPICS

01 TIME DATA: PYTHON & PANDAS

02 FORECASTING

TIME DATA: PYTHON & PANDAS

WHAT

- **Unix timestamp**: number of second since epoch (Jan 1st, 1970 midnight UTC)
- **Python** `time` module: functions for working with time and system clock
- **Python** `datetime()` class: stores & manipulates date and time information
- **Pandas** `Timestamp()` class: optimized replacement for datetime (compatible)

WHY

- Simplifies year, month, day, etc. -> stored as attributes
- Packages time data with common time methods (Object-Oriented Programming)
- Handling strings vs numbers - Aug. 22 is a Friday this year, but not next year
- Other weird time-specific stuff: time zones, AM/PM vs 24 hour clock, etc.

HOW

Converting strings to times (data ingest)

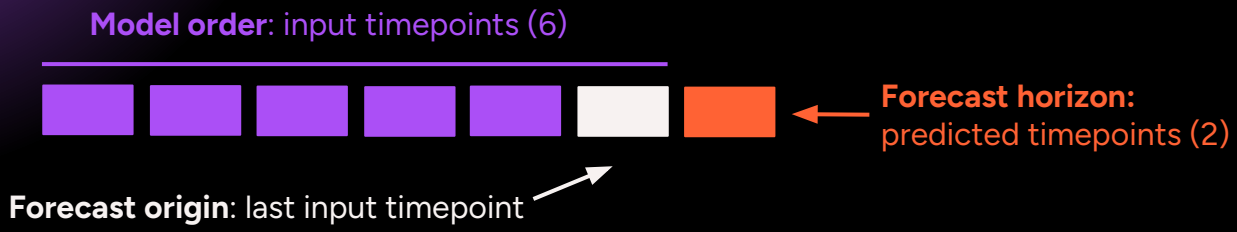
- **Python** `strptime()`: datetime method -> converts string to datetime object
- **Pandas** `pd.to_datetime()`: function to convert DataFrames/Series to timestamp

Both take an input format specifier - EX: `2025-08-22` is `%Y-%m-%d` formatted

FORECASTING

WHAT

Predicting future observation(s) based on past observation(s)



HOW

Using 'standard' ML models

- 'Features': model order
- 'Label': forecast horizon

Using sequence specific models

- **ARIMA**: autoregressive integrated moving average - great for univariate time series (future stock price based on past)
- **Neural networks**: RNN, LSTM, GRU (weather forecast)