# Seasonality and cyclical patterns overview

1. Terminology

* Period (T): the time taken for one repetition of a cycle
* Frequency (1/T): Number of occurrences of a repeating event over a unit of time
* E.g. For an hourly time series, the daily period is 24 hours and the frequency is 1/24 per hour

1. Seasonality

* A pattern or effect that repeats with a **fixed frequency** over time
* Typically related to the calendar or time of day

1. Cyclical patterns

* A pattern that repeats **without a fixed frequency** over time

A graph showing a number of lynx

Description automatically generated

1. Features to capture seasonality and cyclical patterns

* Seasonality
* Lag features
* Calendar features (datetime features)
* Seasonal dummies (seasonal indicators)
* Fourier features (sine and cosine)
* Cyclical patterns
* Lag features

Seasonal lags

1. Lag features can help capture seasonality

* Lag of the seasonal period
* E.g. monthly data, yearly seasonality -> feature
* If we know the seasonal period then we can use that directly
* If we don’t know the seasonal period then we can use the following to help pick a lag
* Domain knowledge
* Plots
* ACF and PACF

1. E.g. Air passengers

* Features
* Trend features:
* Lag of 1 month:
* Model:

A graph of a forecast

Description automatically generated with medium confidence

* With Lag of 12 months:

A graph showing the growth of a period of time

Description automatically generated with medium confidence

1. Pros and cons

* Pros: easy to create
* Cons:
* Creates missing data at the start of the time series – big problem for long seasonality with high frequency time series
* Does not reliably capture seasonality
* Lag features ignore the calendar. If seasonality is driven by the calendar then use datetime features or seasonal dummies
* Creates a strong dependence on what happened exactly one seasonal period ago. So outliers or other unexpected behavior impacts predictions. Can cause instable forecasts

1. Summary

* Lag features can be used to capture seasonality
* The lag of the seasonal period is used
* Simple to create but come with many cons
* Better to combine seasonal lags with other features or use other methods to capture seasonality

# Date and time features for seasonality

1. Datetime features

* Seasonality is often driven by the calendar date and time
* E.g.
* Traffic patterns – daily seasonality related to hour of the day
* Retail sales – weekly seasonality related to day of the week
* Air passenger – yearly seasonality related to month or week of the year
* Extracting features from the date and time can therefore help capture seasonality

1. Implementation

* Sktime:

A white background with black text

Description automatically generated

A screenshot of a computer

Description automatically generated

# Why linear models struggle with date and time features

1. E.g. electricity demand
2. Why

* When computing these features in most packages we receive numeric features
* Most of these datetime variables are cyclical. The numeric representation does not capture this
* Cyclical variables often have non-linear relationship with the target
* Tree-based models can model the non-linear relationship between the target and features
* Linear models are constrained to fit a linear relationship between the target and features
* Additional feature engineering can help linear models better use date & time variables (e.g., sine transformation, treat date time features as a categorical variable and use one hot encoding / target encoding)

# Summary

1. What is seasonality

* A pattern or effect that repeats with a fixed frequency over time
* Typically related to the calendar or time of day

1. Features to capture seasonality

* Lag features
* Calendar features
* Seasonal dummies
* Fourier features