DRP Spr2025

Time Series

By: Celine Sachi and Tracia Pan

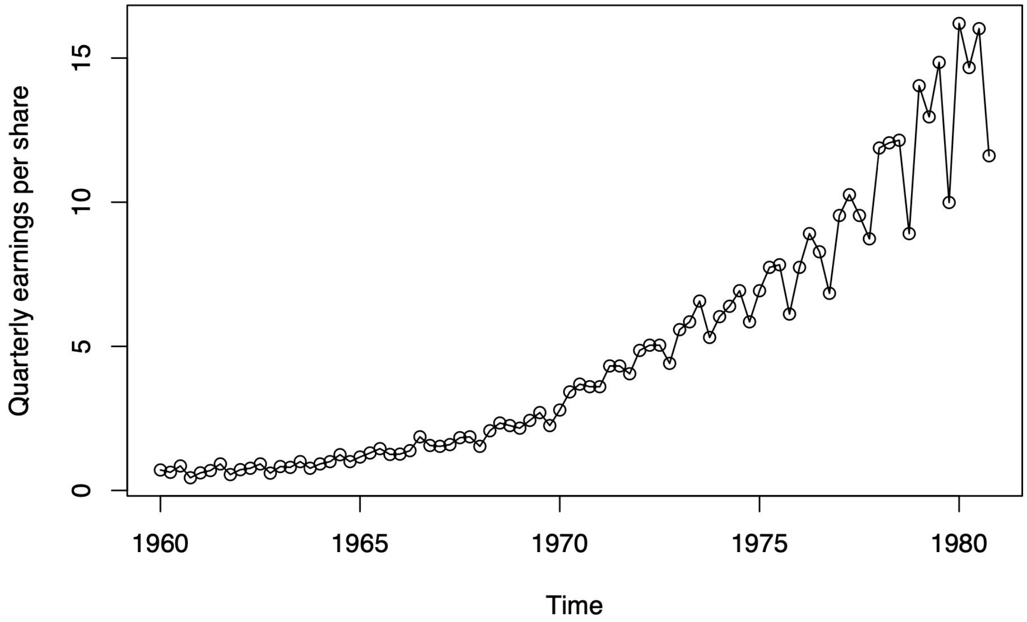
Mentor: Stefan Noel Inzer

Contents

1. What is Time Series?

- Data recorded over successive time intervals
- Not i.i.d

Figure 1: *Johnson & Johnson quarterly earnings per share (from SS).*



02. Characteristic of Time Series

Mean

Definition:

Given a sequence x_t where $t = 1, 2, 3.$
$\mu_t = \mathbb{E}(x_t)$

Variance

Definition:

Given a sequence x_t where $t = 1, 2, 3.$
$\sigma_t^2 = Var(x_t) = \mathbb{E}[(x_t - \mu_t)^2]$

Example:
White noise: a collection of points that are uncorrelated and identically distributed. It has mean of 0 and variance of σ^2 (constant).

Auto-Covariance

Definition:

$$\text{Given a sequence } x_t \text{ where } t = 1, 2, 3.$$
$$\gamma_x(s,t) = \mathrm{Cov}(x_s, x_t)$$
$$\text{Note: } \gamma_x(s,t) = \sigma^2_{x,t}$$

Why is this *useful*?

- Measures linear dependence between variates along the series.
- Used in Model Building (AR, MA, ARIMA)
- Precursor to Auto-correlation (Auto-correlation is the normalized version of auto-covariance which lies between -1 and 1; often easier to interpret).

Cross-Covariance

Definition:

$$\text{Given a sequence } x_t \text{ and } y_t \text{ where } t = 1, 2, 3.$$
$$\gamma_{xy}(s,t) = \mathrm{Cov}(x_s, y_t)$$

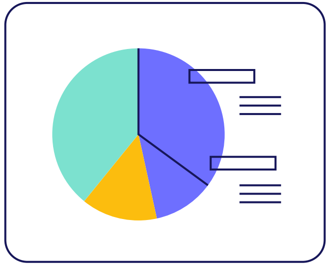
What is the *purpose*?

- Detects Interdependence
- Used in Multivariate Time Series Models
- Precursor to Cross-correlation (cross-correlation is the normalized version of cross-covariance).

Auto-covariance: within **one** series → looks at *internal memory.*

Cross-covariance: between **two** series → looks at *how they relate across time.*

Stationarity



Strong → Weak
Weak ↛ Strong (general)
Weak → Strong (Gaussian)

Definition

Important property that determines whether a time series behave consistently over time.

This is *useful* because:

- Model Assumptions
- Stationary process is more predictable
- Simplified Analysis

Strong Stationarity

$$(x_{t_1}, x_{t_2}, \ldots, x_{t_k}) \stackrel{d}{=} (x_{t_1+\ell}, x_{t_2+\ell}, \ldots, x_{t_k+\ell}), \quad \text{for all } k \geq 1, \text{ all } t_1, \ldots, t_k, \text{ and all } \ell$$

Any collection of variates along the series has the same joint distribution after we shift the time indices forward or backward.

Note: rare useful in practice.

Weak Stationarity

$$\mu_{x,t} = \mu, \quad \text{for all } t$$

$$\gamma_x(s,t) = \gamma_x(s+\ell, t+\ell), \quad \text{for all } s, t, \ell$$

Mean and Variance are constant.

Auto-covariance depends only on lag.

Note: common in modeling (ARIMA)

3. Decomposing Time Series

Time series model:

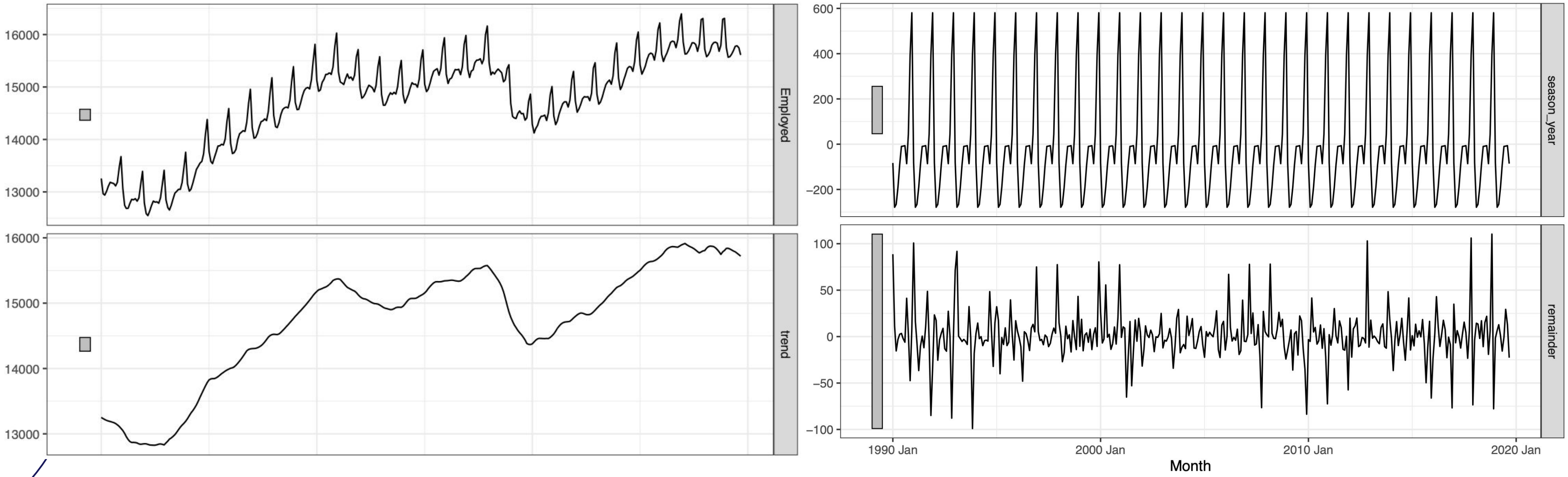
$$\overbrace{x_t = \underbrace{\theta_t}_{u_t + s_t} + w_t}$$

Signal: θ_t

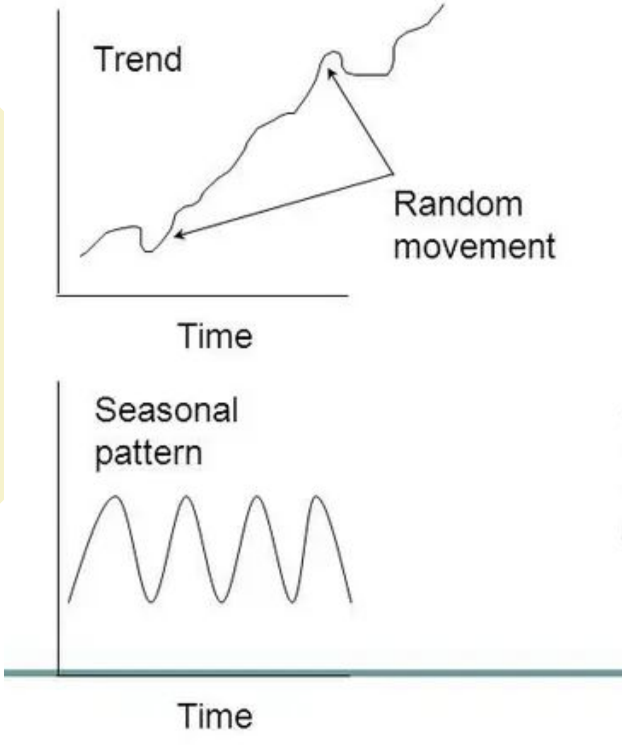
Noise: w_t

$$\theta_t = \underbrace{u_t}_{\text{Trend}} + \underbrace{s_t}_{\text{Seasonal}}$$

3. Decomposing Time Series Example



04. Seasonality



- **Definition**: regular, repeating patterns or cycles in time series that occur at fixed intervals due to seasonality factors (time of year, month, week, day, etc)

- **Seasonality vs Trend:**
  - Trend: long-term upward/downward movement
  - Seasonality: short-term, cyclic, periodic variation

- **How to detect Seasonality?**
  - Fourier Decomposition
  - Discrete Fourier Transform
  - Seasonal-Trend Decomposition

Detecting Seasonality Methods

1. Fourier Decomposition (Continuous)

- A method for breaking down a complex time series into a sum of simple **sine and cosine waves** with different frequencies, amplitudes, and phases.
- Definition:

$$c_{tj} = \cos(2\pi j/n \cdot t), \quad t = 1, \ldots, n$$
$$s_{tj} = \sin(2\pi j/n \cdot t), \quad t = 1, \ldots, n$$

$$x_t \approx a_0 + \sum_{j=1}^{p} (a_j c_{tj} + b_j s_{tj}), \quad t = 1, \ldots, n$$

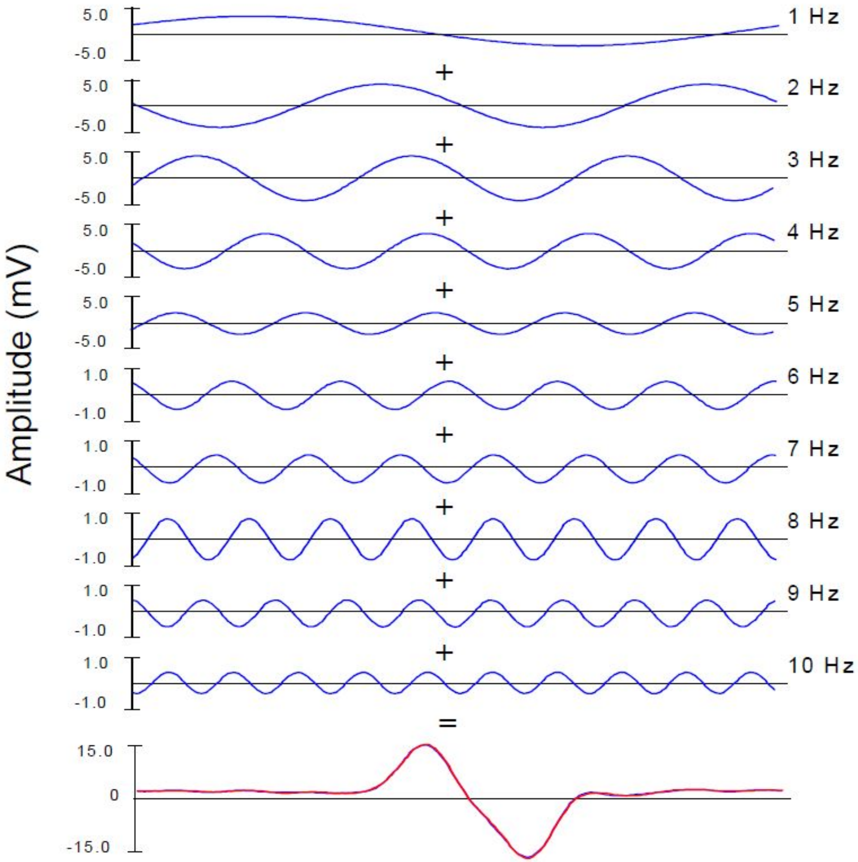
Note: set p = (n – 1)/2

- **Why is this useful?**
  - *Uncover hidden cycles* by expressing the series in terms of its frequency components.
  - *Turns complicated repeating patterns into a mix of simple wave shapes* (smooth up-and-down curves)
  - Helps us build tools to *remove noise, focus on important parts, or study signals* more easily

In theory, FD is very slow: takes about O(n^2) for n points.

A faster method: “Fast Fourier Transform” : O(n log n)

Example of FD



2. Discrete Fourier Transform (DFT)

- A method used to convert a **discrete time series (a list of numbers)** into a **set of frequencies** that represent it.
- Definition:

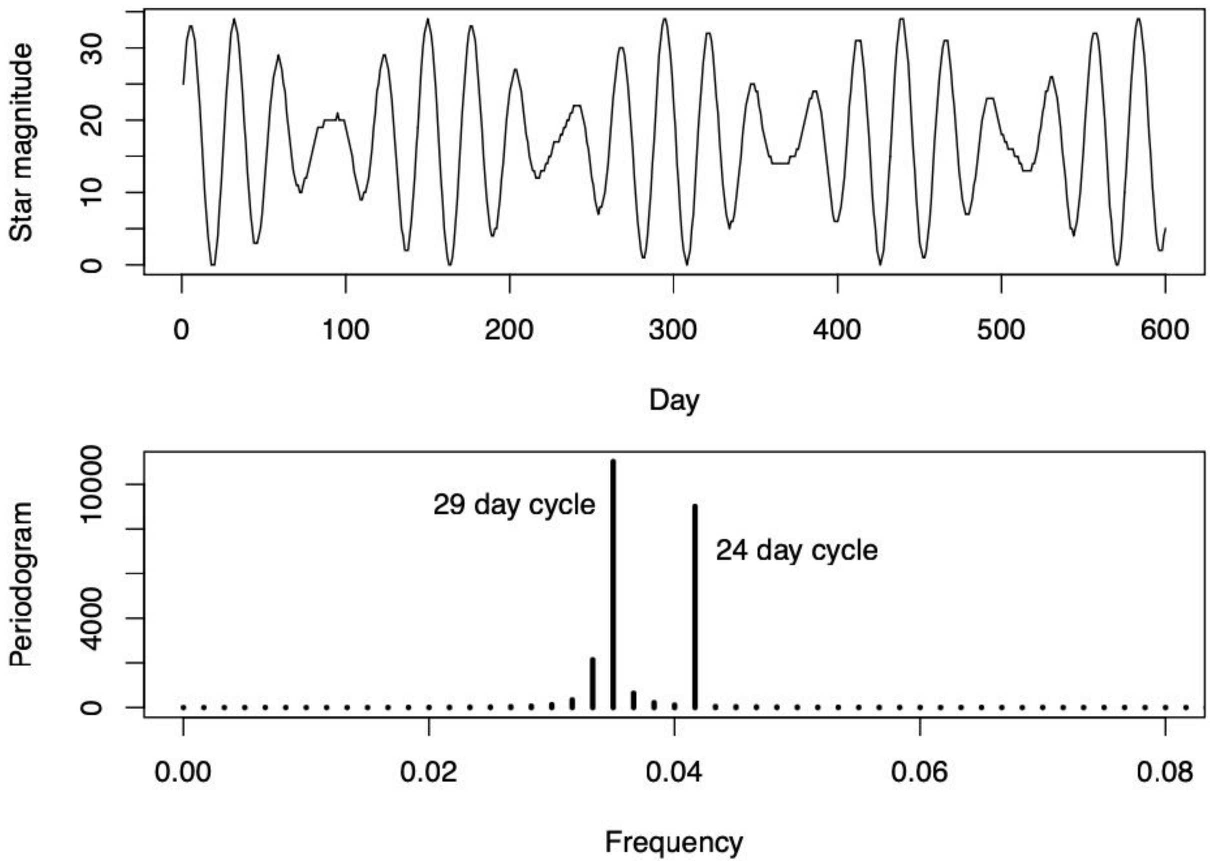
For a discrete time series $x_0, x_1, \ldots, x_{N-1}$, the DFT is defined as:

$$X_k = \sum_{n=0}^{N-1} x_n \cdot e^{-2\pi i \frac{kn}{N}}, \quad k = 0, 1, \ldots, N-1$$

- X_k: complex-valued DFT coefficient at frequency k
- N: number of time samples
- The **inverse DFT** reconstructs the time series from the frequency domain

- **Why is this useful?**
- Converts a time series into the **frequency domain**, showing how much of each frequency is present.
- To *detect cycles, analyze seasonality and filter out noise*
- Can also be computed using *Fast Fourier Transform.*

Example of DFT



3. Seasonal-Trend (ST) Decomposition

- A statistical method to *detect the presence of seasonality* in time series by separating into 3 components:
  Time Series = Trend + Seasonality + Noise

- **Types of Decomposition**:

- *Additive:* Use when seasonal variation is roughly **constant over time**.

$$y_t = T_t + S_t + e_t$$

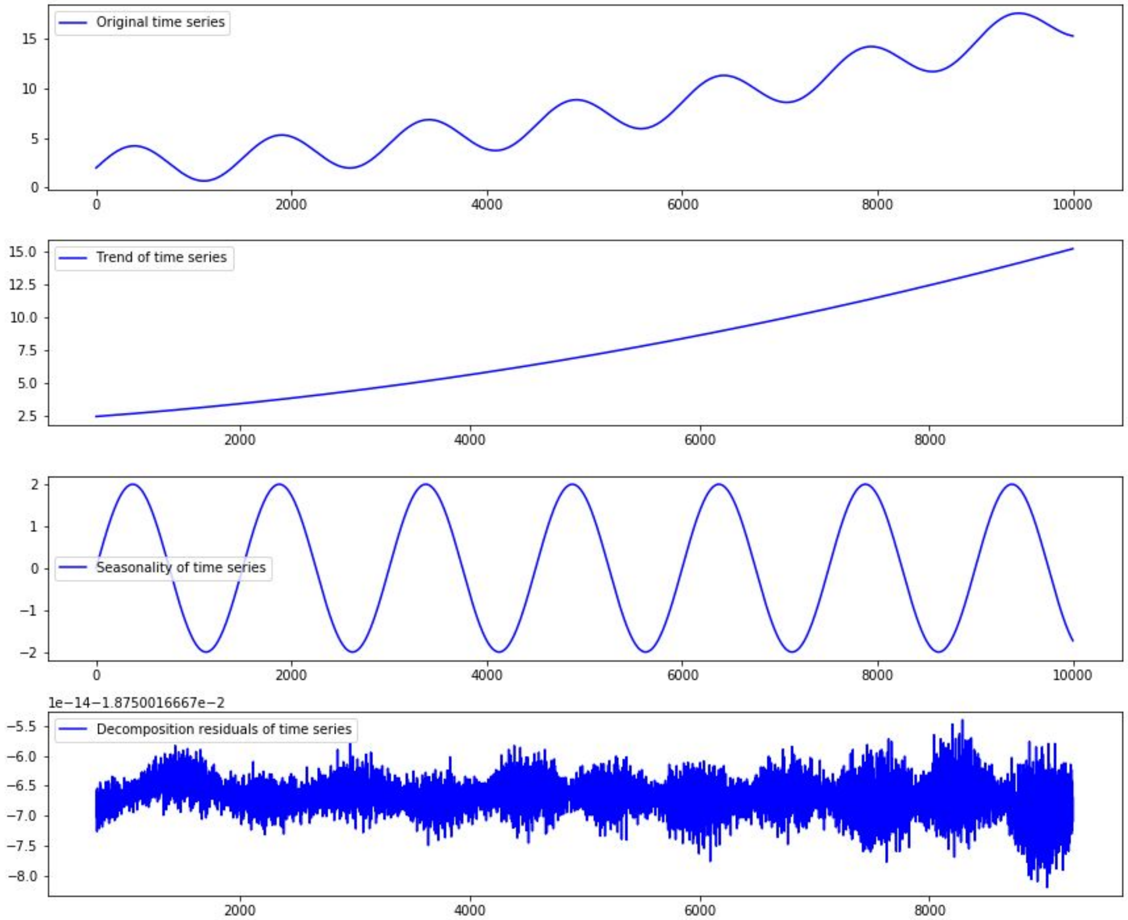
- *Multiplicative:* Use when seasonal variation **changes with trend level** (e.g., growing sales with bigger seasonal peaks).

$$y_t = T_t \times S_t \times e_t$$

- **Why is this useful?**

- Understanding patterns
- Improve forecasting
- Improve interpretability

Example of ST Decomposition



04. Smoothers

- Eliminates noise and irregularities to reveal underlying trend and patterns in the data
- Linear Filters:

$$\hat{\theta}_i = \sum_{j=-k}^{k} a_j y_{i-j}, \quad i = 1, \ldots, n$$

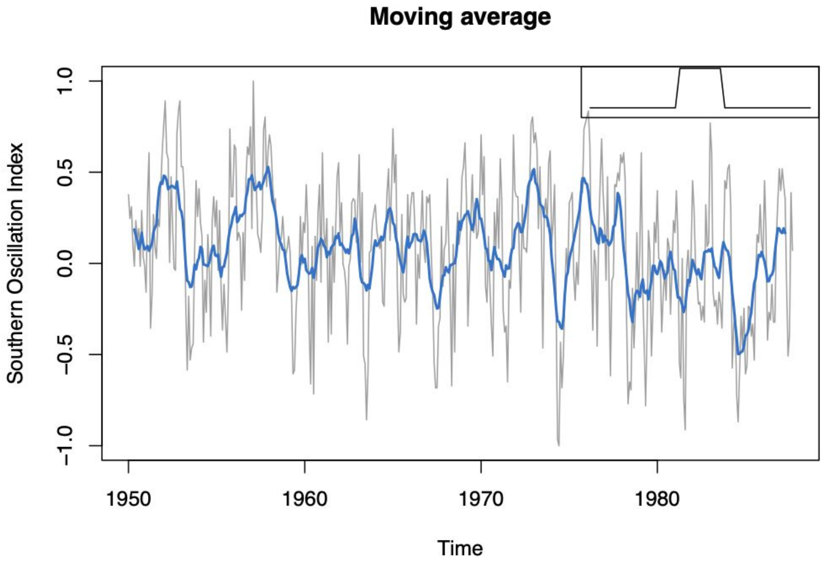
- Penalized Least Squares:

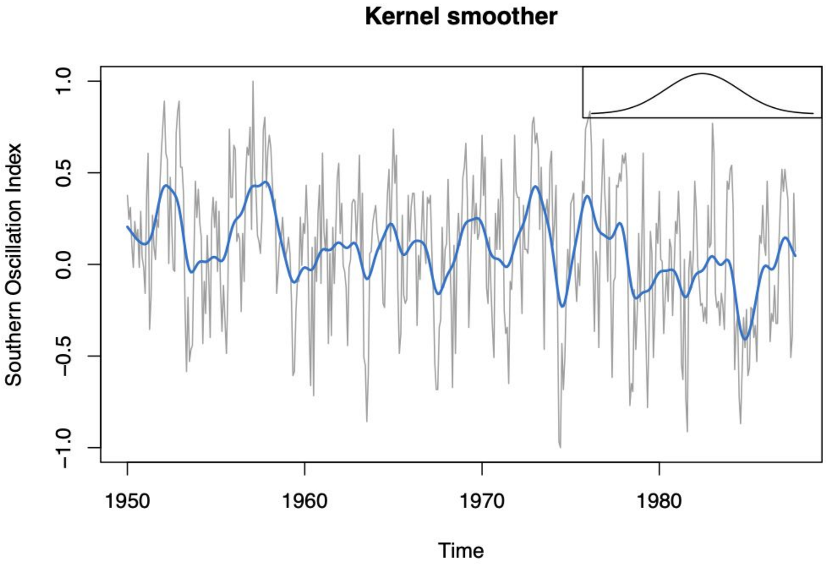
$$\min_{\theta \in \mathbb{R}^n} \sum_{i=1}^{n} (y_i - \theta_i)^2 + \lambda P(\theta)$$

04. Smoothers: Linear Filters

Moving Average: calculates a series of averages from a specified number of consecutive data points in a time series

Kernel Smoothing: creates smoother-looking estimates by using a smoother weight sequence (larger k)

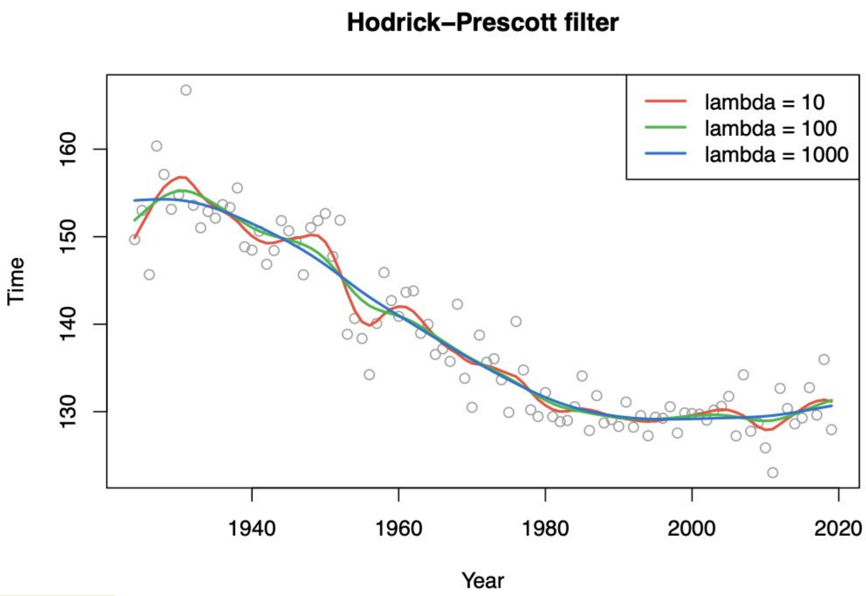


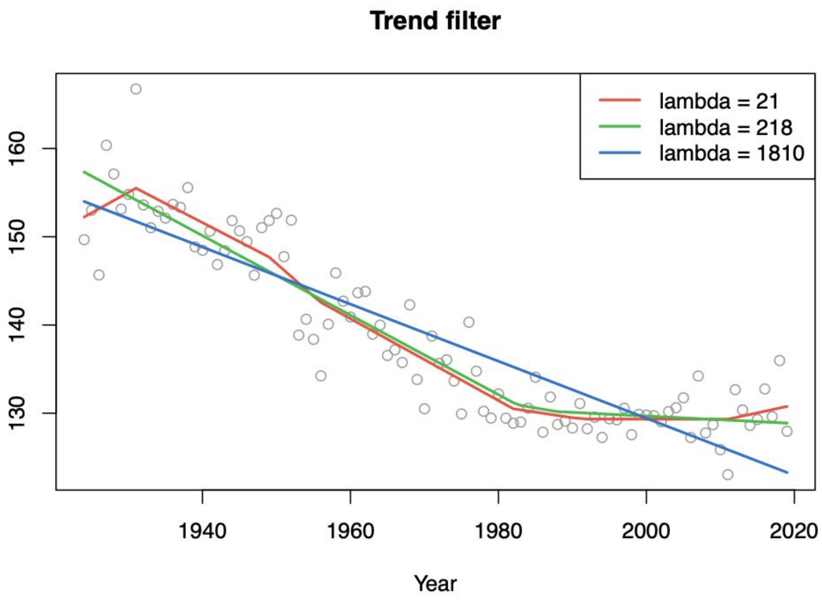


04. Smoothers: Penalized Least Squares

Hodrick-Prescott Filter: decompose a time series into a smooth trend component and a cyclical component

Trend Filter: smooths out short-term fluctuations, noise, and seasonal variations





Thanks!

Any Questions?