Time series decomposition I

EC 361–001

Prof. Santetti
Spring 2024

Materials

Required readings:

- `Hyndman & Athanasopoulos, ch. 3`
  - Sections 3.1—3.4.

Motivation

Motivation

Given that a time series may exhibit several different **features**, these can be **split** into different components, each representing an underlying pattern category.

Recall the three main **features**:

- Trend;
- Seasonality;
- Cyclical component.

The usual approach is to consider the **trend and cycle** components *together*; the **seasonal** features; and a third component containing *anything else* that the other two do not comprise, known as the **remainder**.

Motivation

This week, we will learn how to **decompose** a time series into these features.

This way, we **improve** our understanding of the time series at hand.

First, though, some **adjustments** may be necessary to make our jobs easier.

Data transformations

Data transformations

Among the *several* different data **adjustment/transformation** techniques that exist in Statistics, (macro)economists use mainly *three*:

1. *Per capita* adjustments;
2. *Inflation* (real) adjustments;
3. *Logarithmic* transformations.

Data transformations

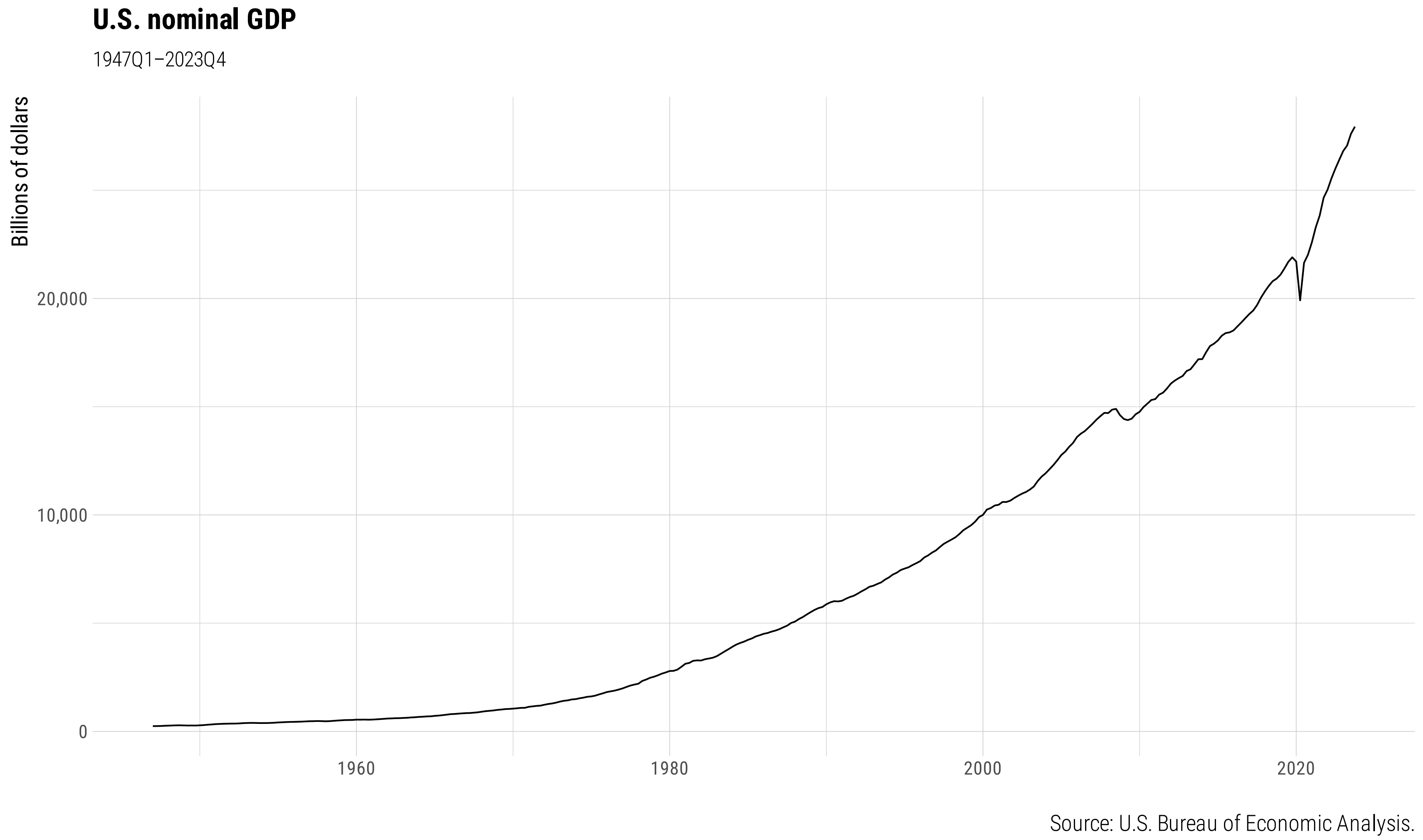
- ***Per capita*** adjustments:

Many times, we are interested in economic measures *relative to some populational reference*.

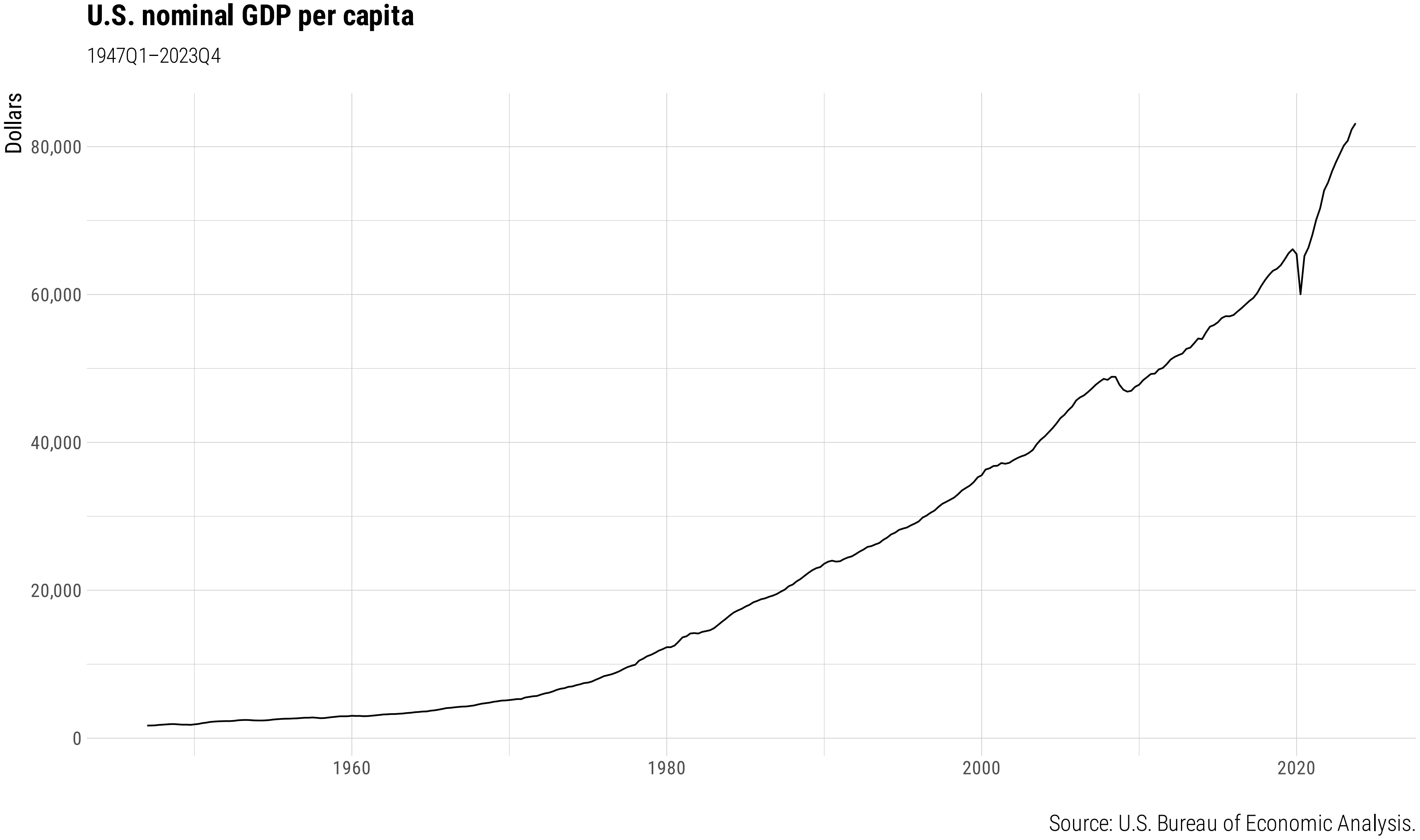
For instance, it is not uncommon to analyze a country's **Gross Domestic Product** (GDP) relative to its population size

- i.e., GDP *per capita*.

Data transformations



Data transformations



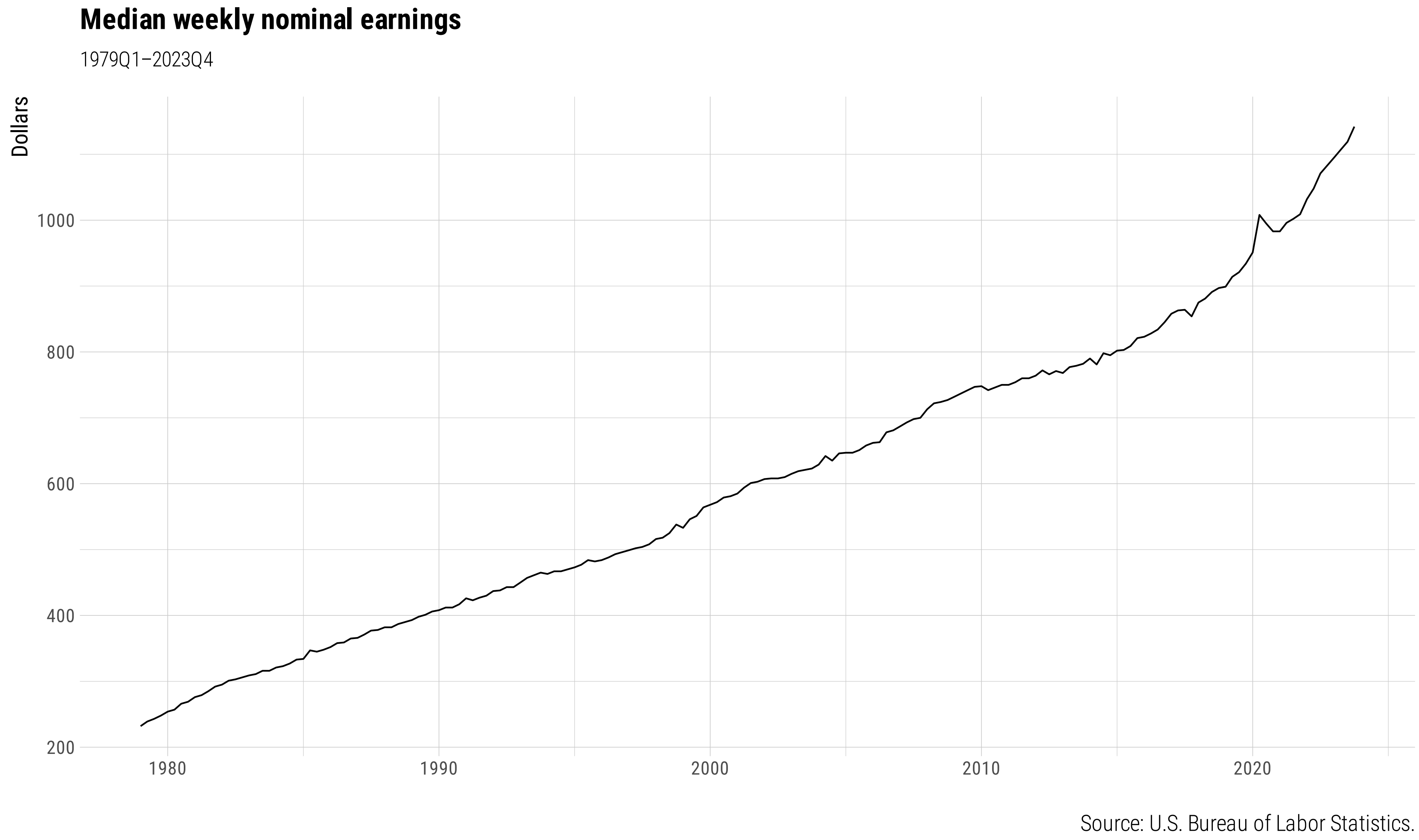
Data transformations

- **Inflation (real)** adjustments:

As economists, we know that **nominal** measures may be misleading, since changes in **prices** may *distort* a statistic of interest.

Therefore, many times we need to use **real** adjustments to our data.

Data transformations



Data transformations

Data transformations

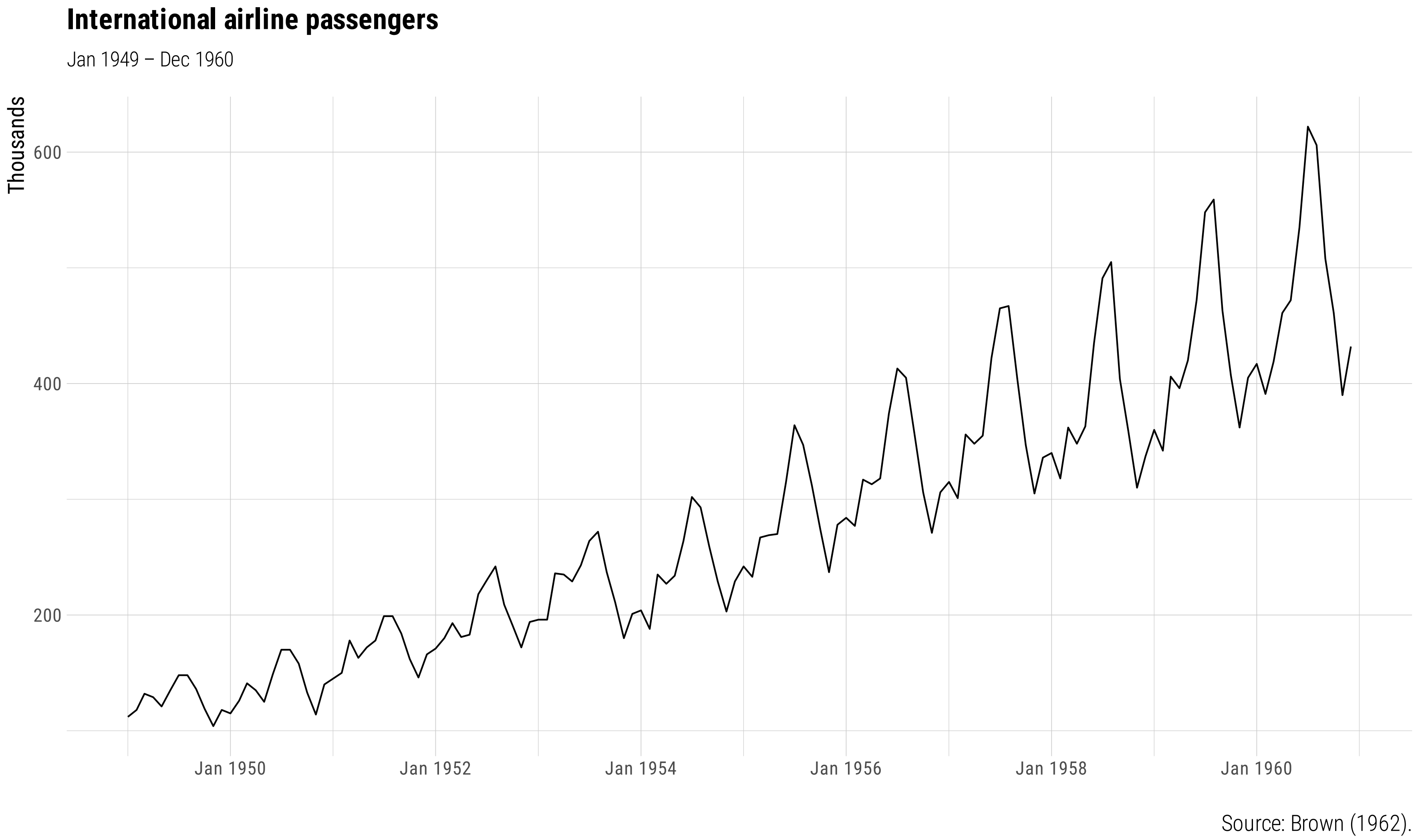
- **Logarithmic** transformations:

In case the **variance** of our data changes at different levels of the series, a **logarithmic transformation** may be useful.

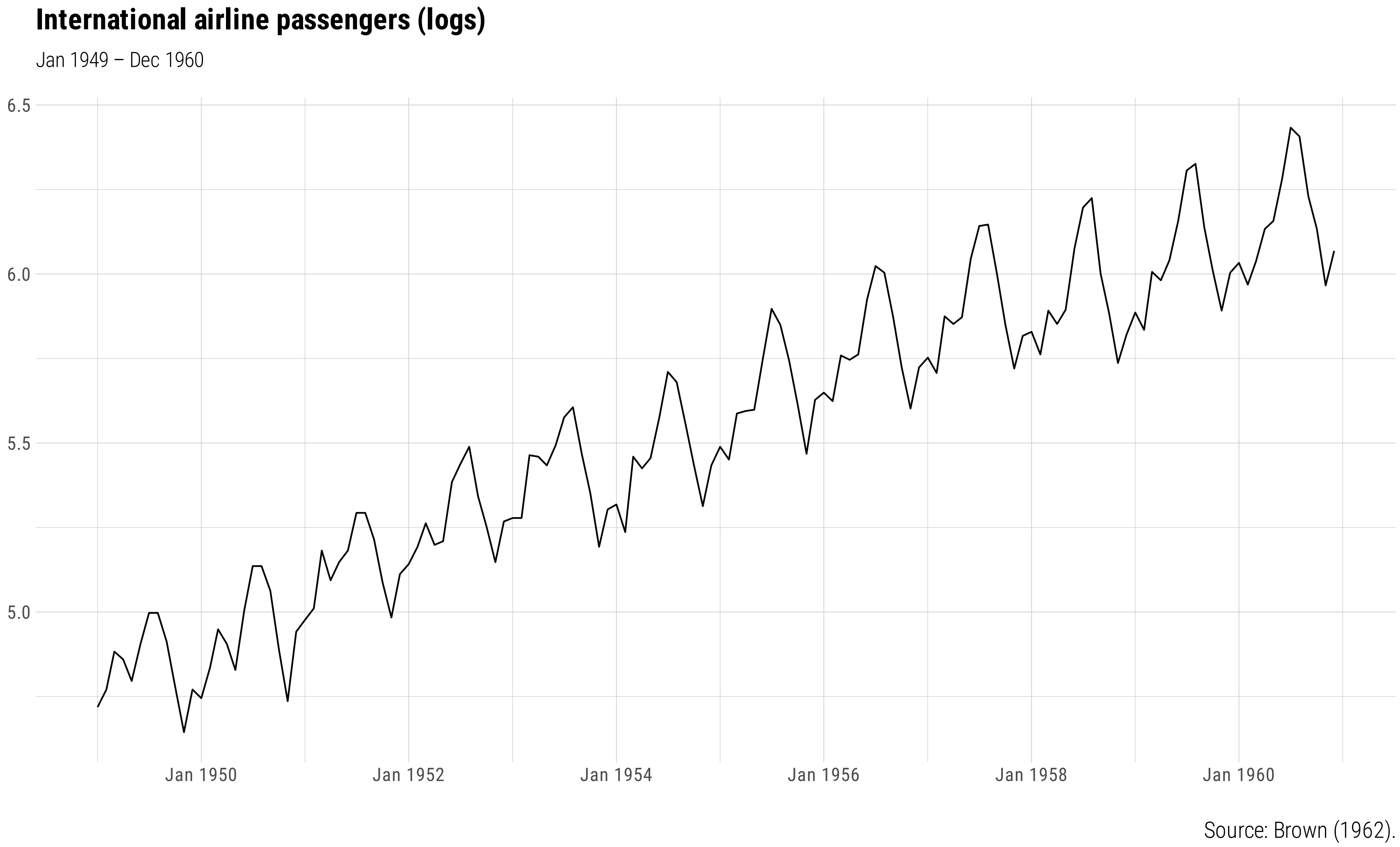
In addition, using **natural logarithms** (base e) are *directly interpretable*:

- Changes in a *log* value are relative (percent) changes on the original scale.

Data transformations



Data transformations



Time series components

Time series components

Time series data can be decomposed into its **trend-cycle** (T_t), **seasonal** (S_t), and **remainder** (R_t) components in *two* ways:

- **Additive**: $T_t + S_t + R_t$
- **Multiplicative**: $T_t \times S_t \times R_t$

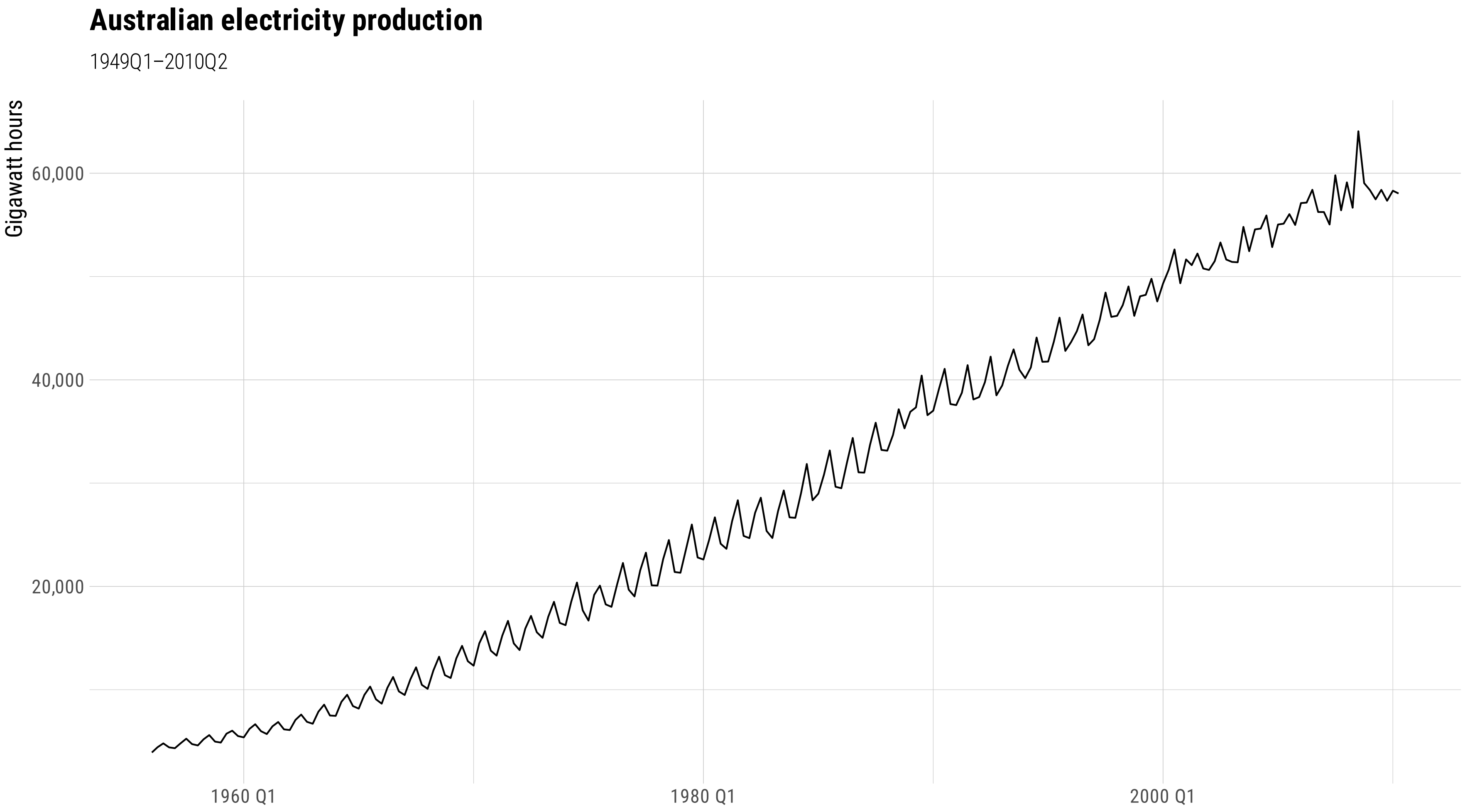
In case the magnitude of the seasonal fluctuations **does not** vary with the level of the time series, the **additive** decomposition is appropriate.

However, in case the variance of the time series is not constant, the **multiplicative** method is a better choice.

In our case, we will stick with the **additive method**:

- Recall that if $y_t = T_t \times S_t \times R_t \rightarrow log\ y_t = log\ T_t + log\ S_t + log\ R_t$

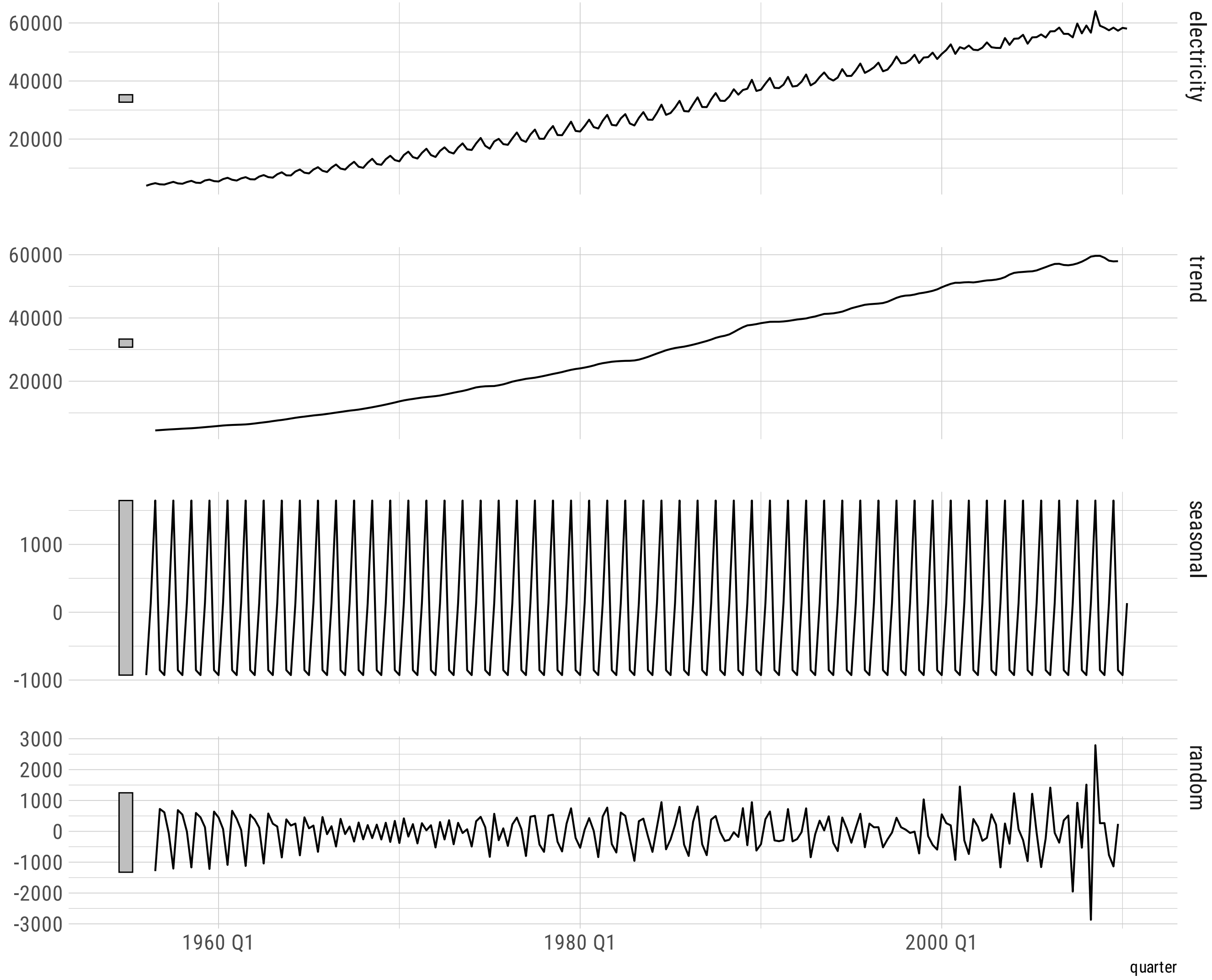
Time series components



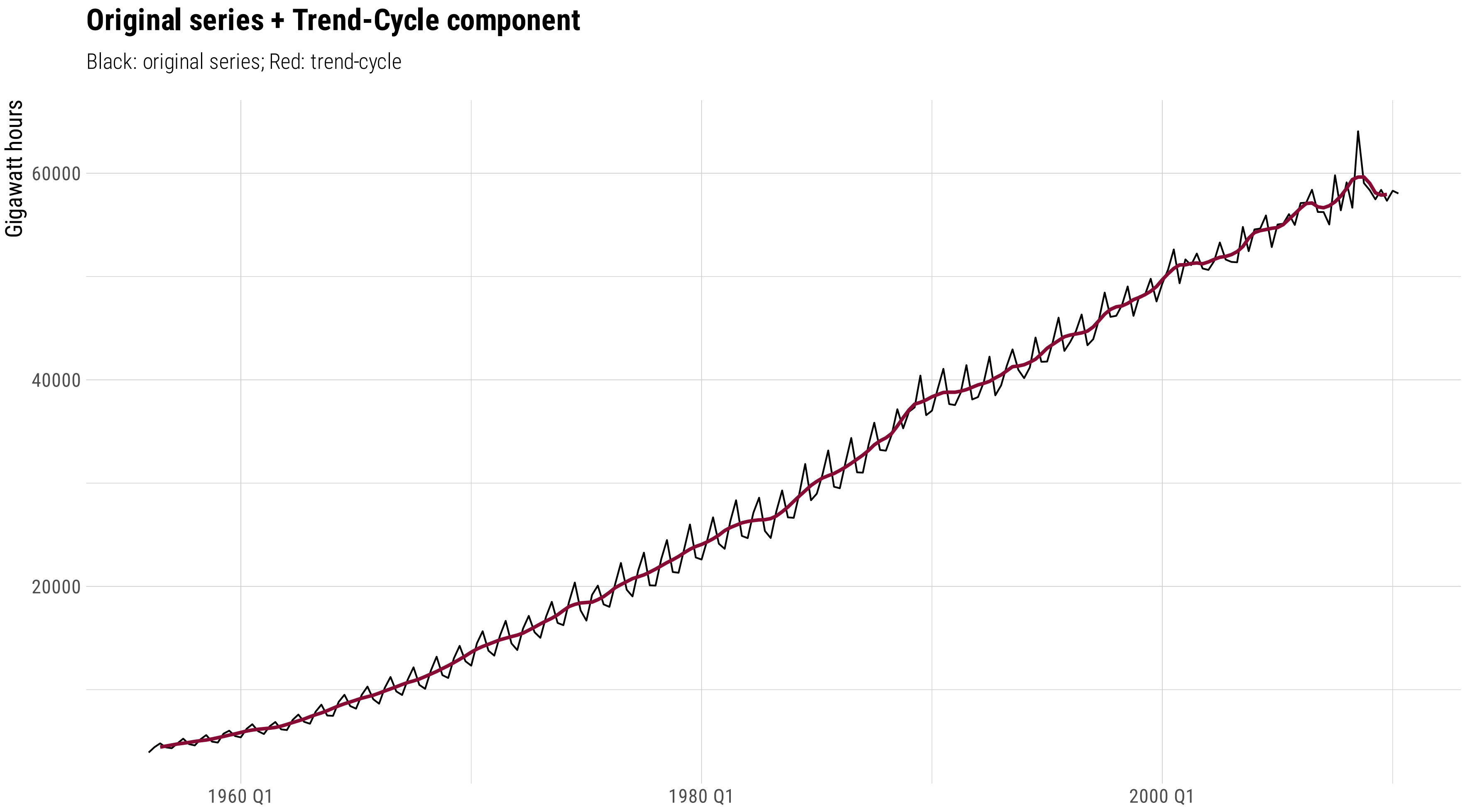
Source: Hyndman & Athanasopoulos (2021)

Classical decomposition

electricity = trend + seasonal + random

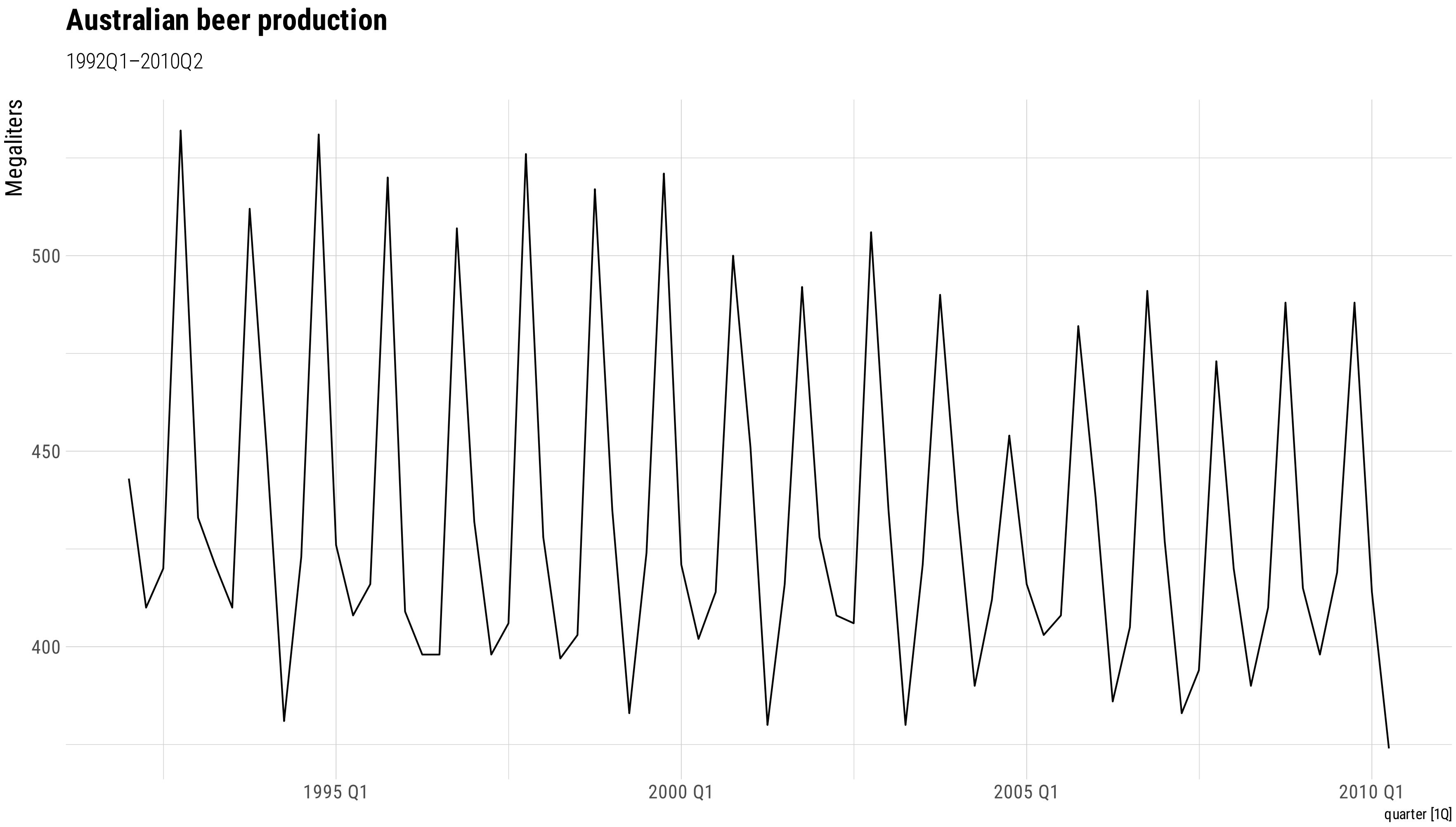


Time series components



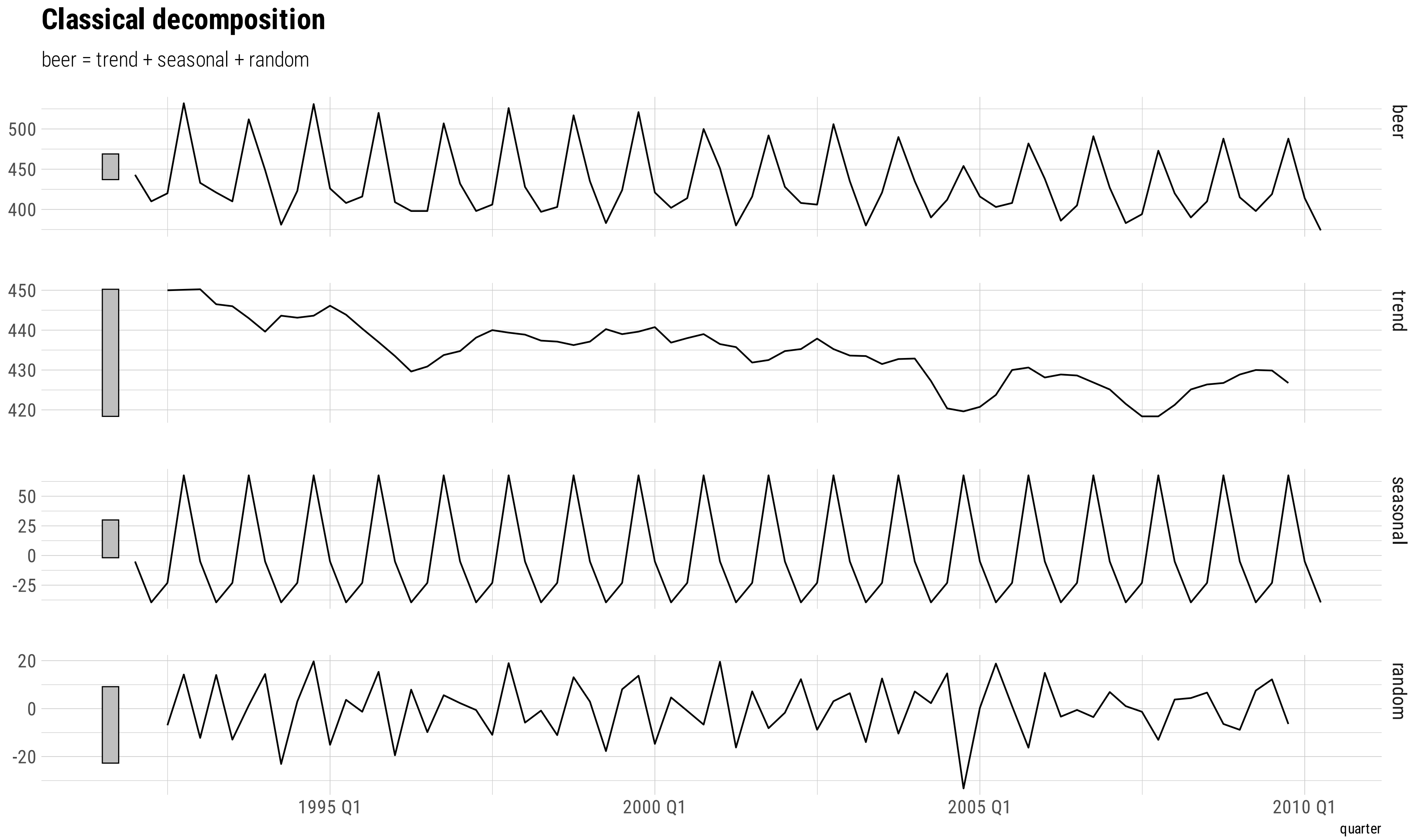
Source: Hyndman & Athanasopoulos (2021).

Time series components



Source: Hyndman and Athanasopoulos (2021).

Time series components



Time series components

If the **seasonal** component is removed from the original data, the resulting values are the ***seasonally adjusted*** data.

In case of an additive decomposition, seasonally adjusted data are simply

$$y_t - S_t = T_t + R_t$$

Seasonally adjusted data are especially **useful** to understand the variations in the data due to the underlying *state of the economy*, rather than due to seasonal factors.

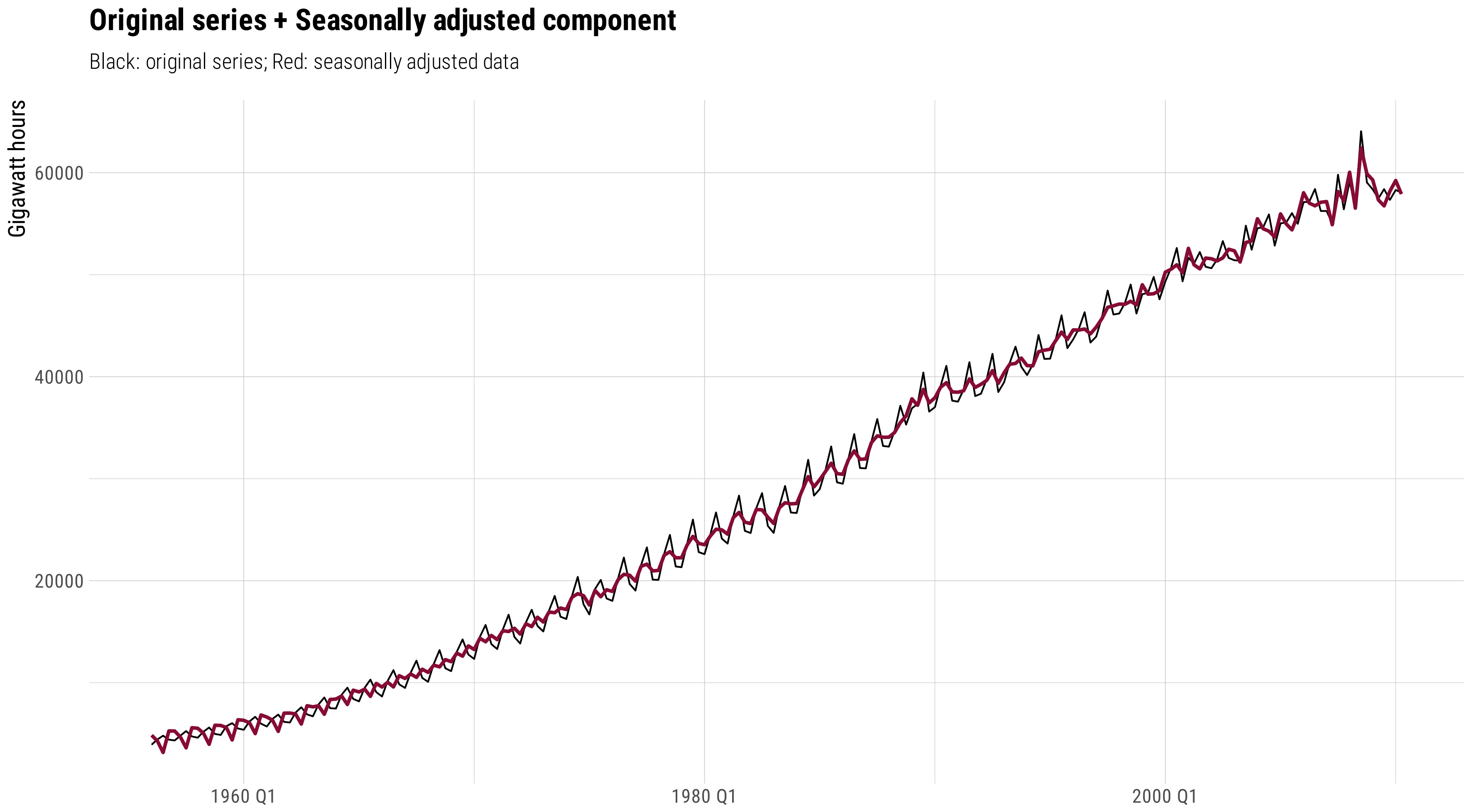
Recall that seasonally adjusted data contain the *trend-cycle* **and** the *remainder.*

- Therefore, seasonally adjusted data have more **noise** than just the trend.

Time series components

`A look at seasonally adjusted data`

Time series components



Source: Hyndman & Athanasopoulos (2021).

Classical decomposition

Classical decomposition

There are **several** methods one can use to decompose a time series into its main *features*.

The *starting point* is the so-called **classical decomposition**.

Regarding seasonality, classical decomposition assumes that the *seasonal* component is **constant** from year to year.

Concentrating on the **additive** decomposition case, the first step is to estimate the **trend-cycle** component.

- This is done by using a **moving-average** approach.

Classical decomposition

A **moving average** of order *m* is defined by

$$m\text{-order MA} = \frac{1}{m}\sum_{j=-k}^{k} y_{t+j} \qquad \text{where } m = 2k+1$$

The **intuition** behind moving averages is that, since *neighboring* observations are likely to be close in value, a more dynamic averaging measure will eliminate some of the **randomness** in the data.

Classical decomposition

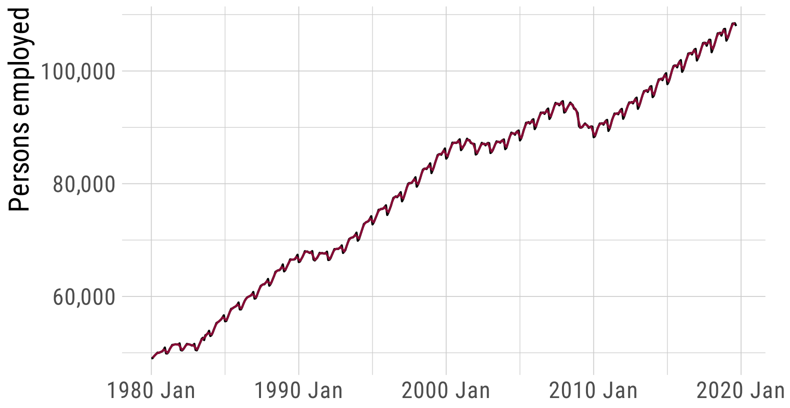
$$\frac{1}{m}\sum_{j=-k}^{k} y_{t+j} \qquad \text{where } m = 2k+1$$

```
us_private_service
```

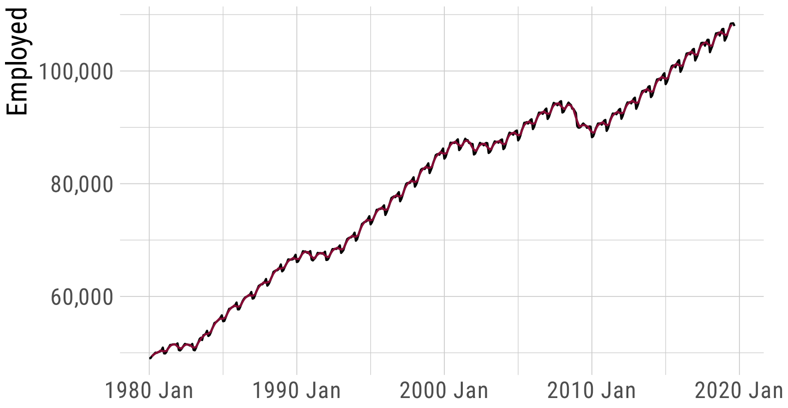
```
#> # A tsibble: 477 x 7 [1M]
#>       Month Title                         Employed    ma3    ma5    ma9   ma15
#>       <mth> <chr>                            <dbl>  <dbl>  <dbl>  <dbl>  <dbl>
#>  1 1980 Jan Private Service-Providing    49040    NA     NA     NA     NA
#>  2 1980 Feb Private Service-Providing    49041 49141.    NA     NA     NA
#>  3 1980 Mar Private Service-Providing    49343 49322. 49358     NA     NA
#>  4 1980 Apr Private Service-Providing    49581 49570. 49557     NA     NA
#>  5 1980 May Private Service-Providing    49785 49800. 49741. 49671     NA
#>  6 1980 Jun Private Service-Providing    50035 49927. 49889. 49807.    NA
#>  7 1980 Jul Private Service-Providing    49962 50026  50007. 49969.    NA
#>  8 1980 Aug Private Service-Providing    50081 50071. 50103  50146. 49910.
#>  9 1980 Sep Private Service-Providing    50171 50173. 50196. 50179. 50019.
#> 10 1980 Oct Private Service-Providing    50266 50312  50391. 50191. 50150.
#> # ℹ 467 more rows
```

Private services: Number of employed persons, Jan 1980 – Sep 2019

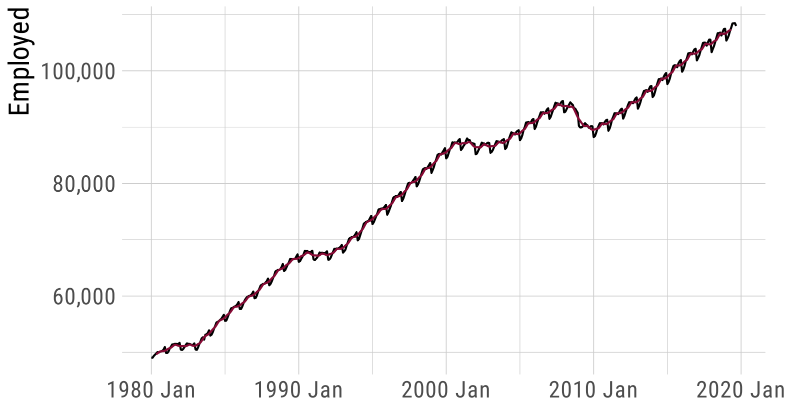
Black: Original series; Red: 3-month moving average



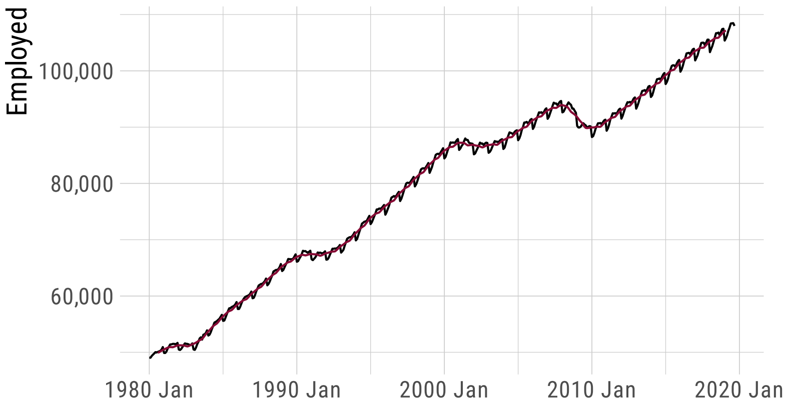
Red: 5-month moving average



Red: 9-month moving average



Red: 15-month moving average



Classical decomposition

After the trend-cycle component is obtained, we may obtain a **detrended** series:

$$y_t - \hat{T}_t$$

The **seasonal** component is calculated by taking the **average** of the *detrended* values for each season.

- Each *quarter* for quarterly data, each *month* for monthly data, and so on.

Finally, the **remainder** results from subtracting the seasonal and trend-cycle components from the original data:

$$\hat{R}_t = y_t - \hat{T}_t - \hat{S}_t$$

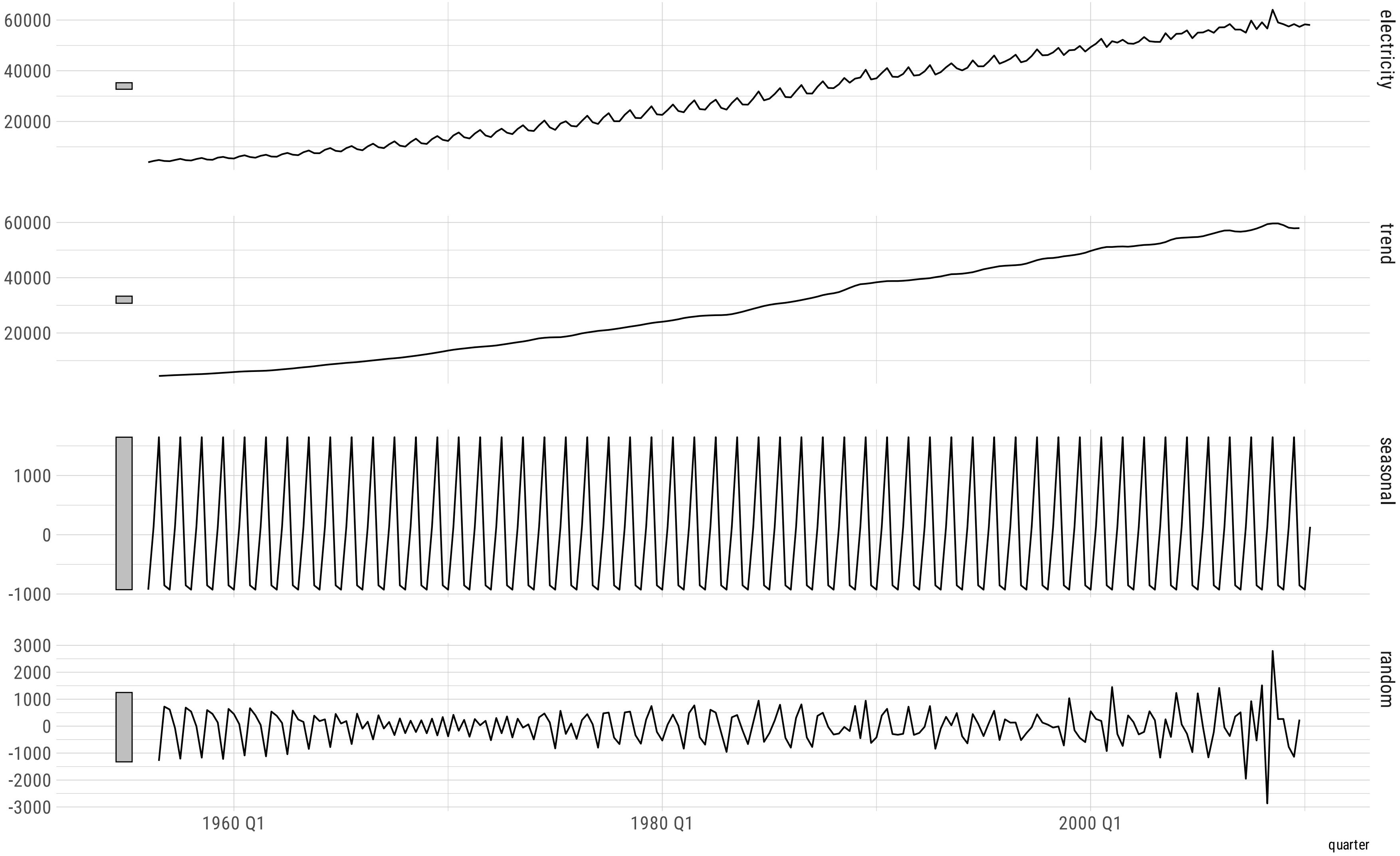
Classical decomposition

In summary:

1. For a time series with seasonal period m:
    - If m is *even*, obtain the trend-cycle component by using a $2 \times m$ moving average;
    - If m is *odd*, obtain the trend-cycle component by using a an *m-order* moving average.
2. Calculate the *detrended* series: $y_t - T_t$;
3. For the *seasonal* component, average the detrended values for each season. If needed, adjust so they add up to *zero*.
4. The *remainder* component is calculated by subtracting the estimated seasonal and trend-cycle components: $R_t = y_t - T_t - S_t$

Classical decomposition

electricity = trend + seasonal + random



Classical decomposition

Some **drawbacks** of classical decomposition:

- Since it uses moving averages, the trend-cycle is **unavailable** for the first few and last few observations;
- The trend-cycle estimate tends to **over-smooth** rapid rises and falls in the data;
- Classical decomposition methods assume that the seasonal component **repeats** from year to year;
- Not robust for **outliers**.

Next time: Time series decomposition II