

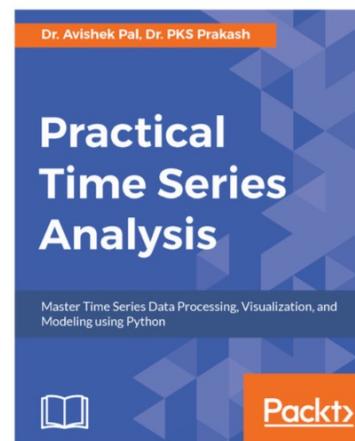
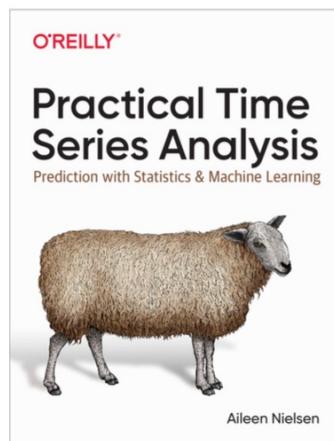
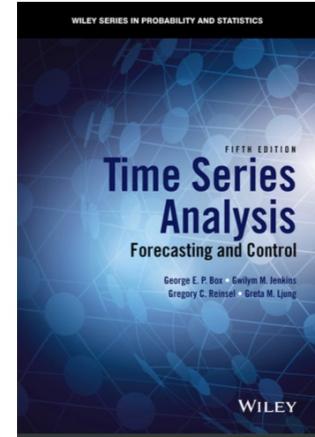
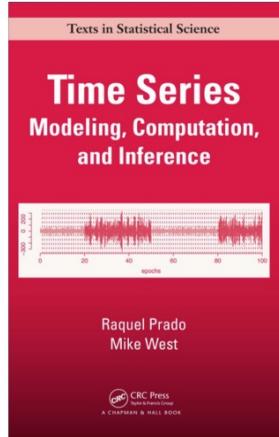
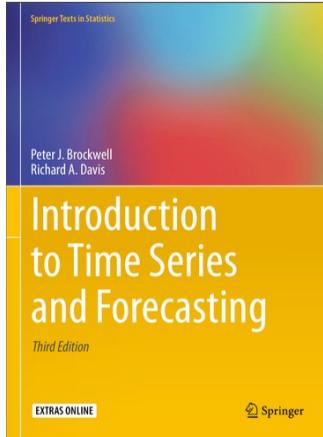
Time Series Analysis

INFO 523 - Lecture 11

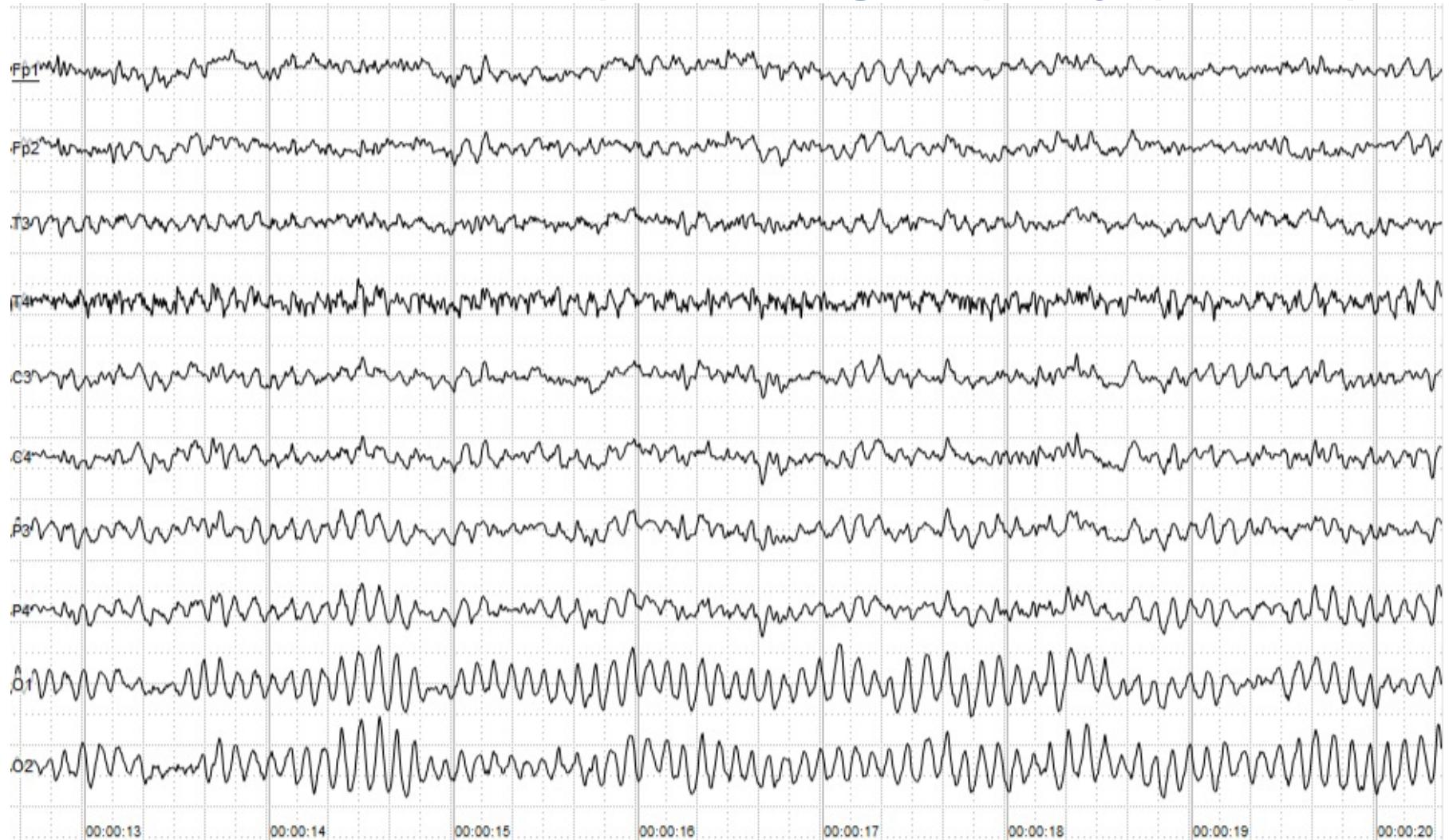
Dr. Greg Chism

Lesson 1: **Understanding Timeseries**

References

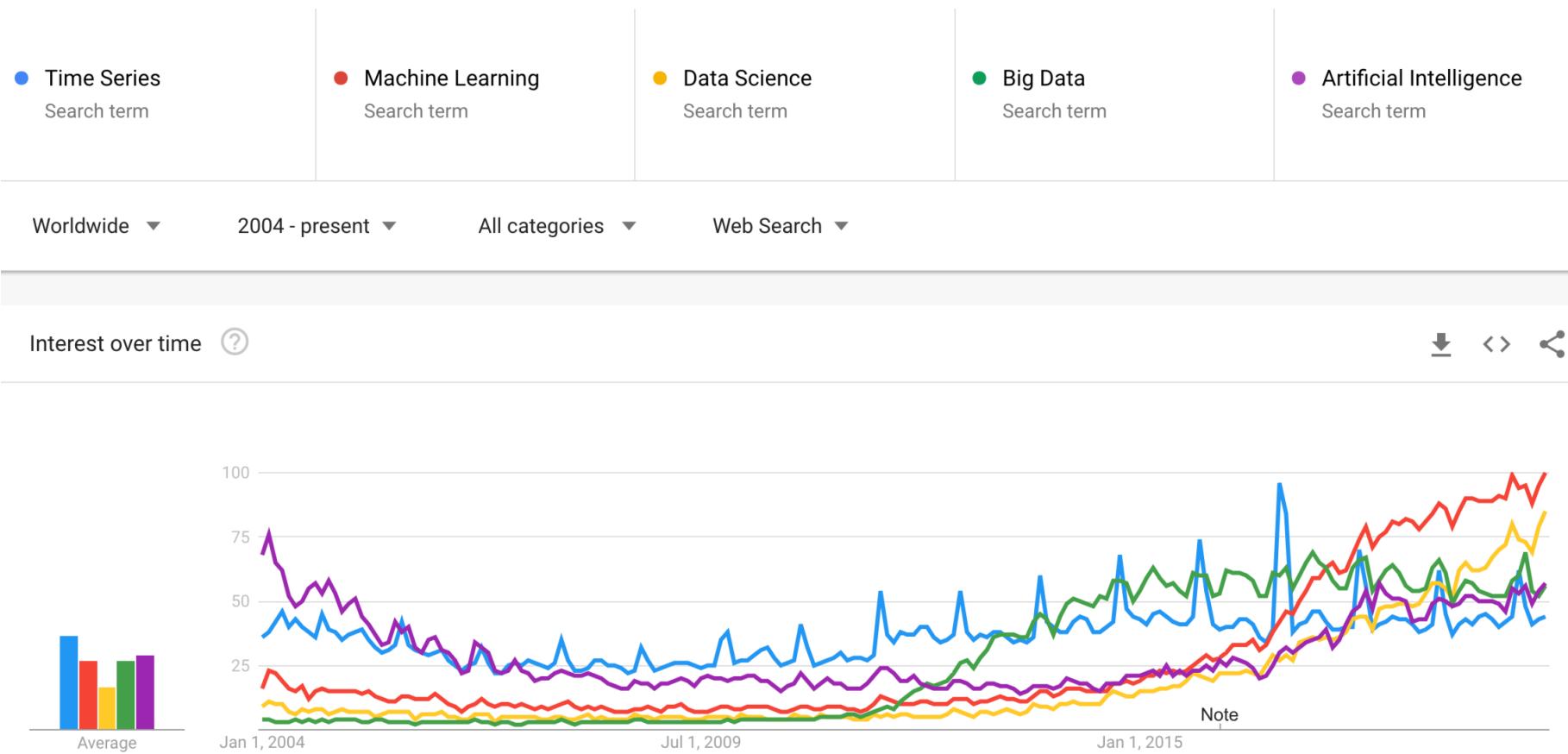


Electroencephalography(EEG)

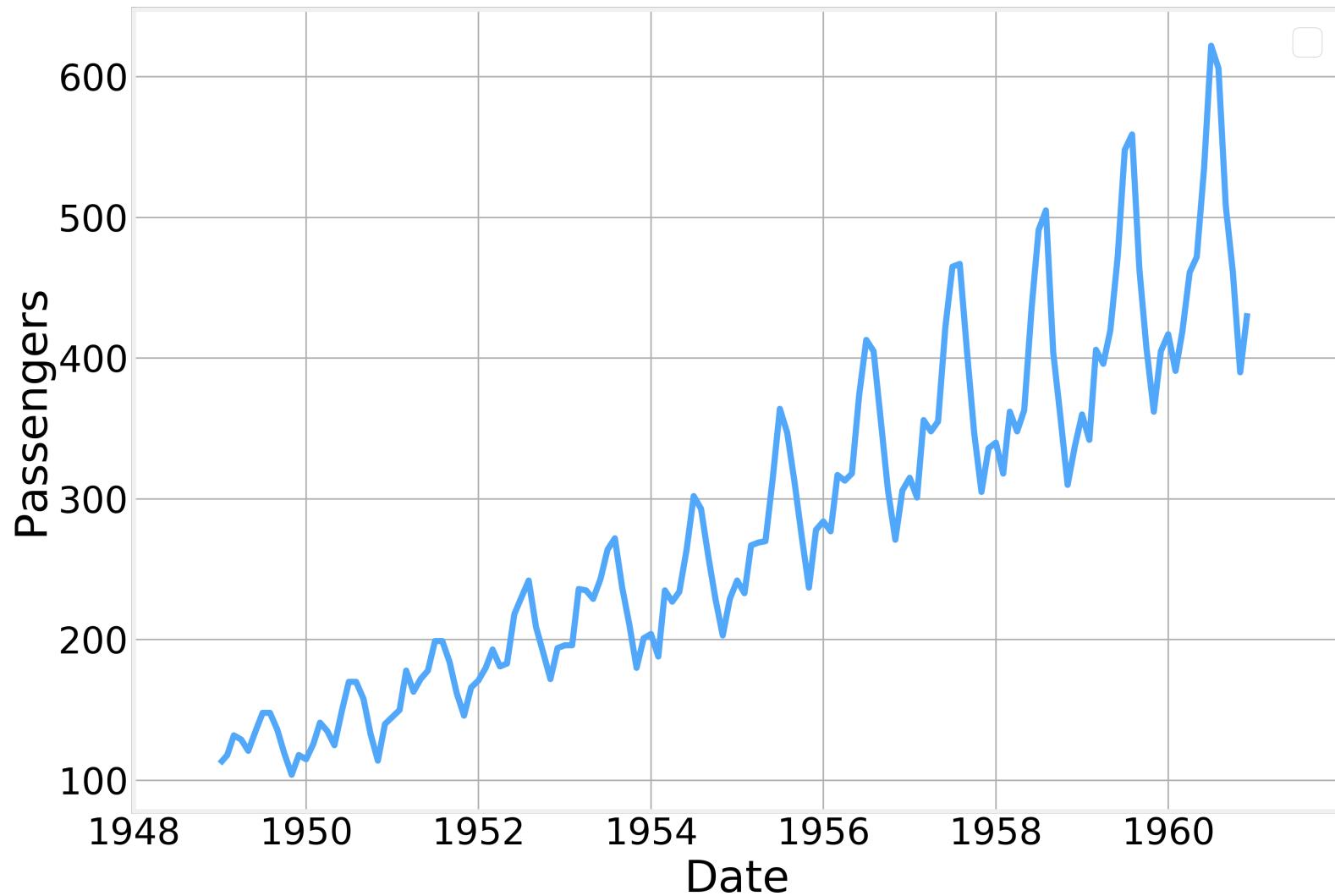


<https://en.wikipedia.org/wiki/Electroencephalography>

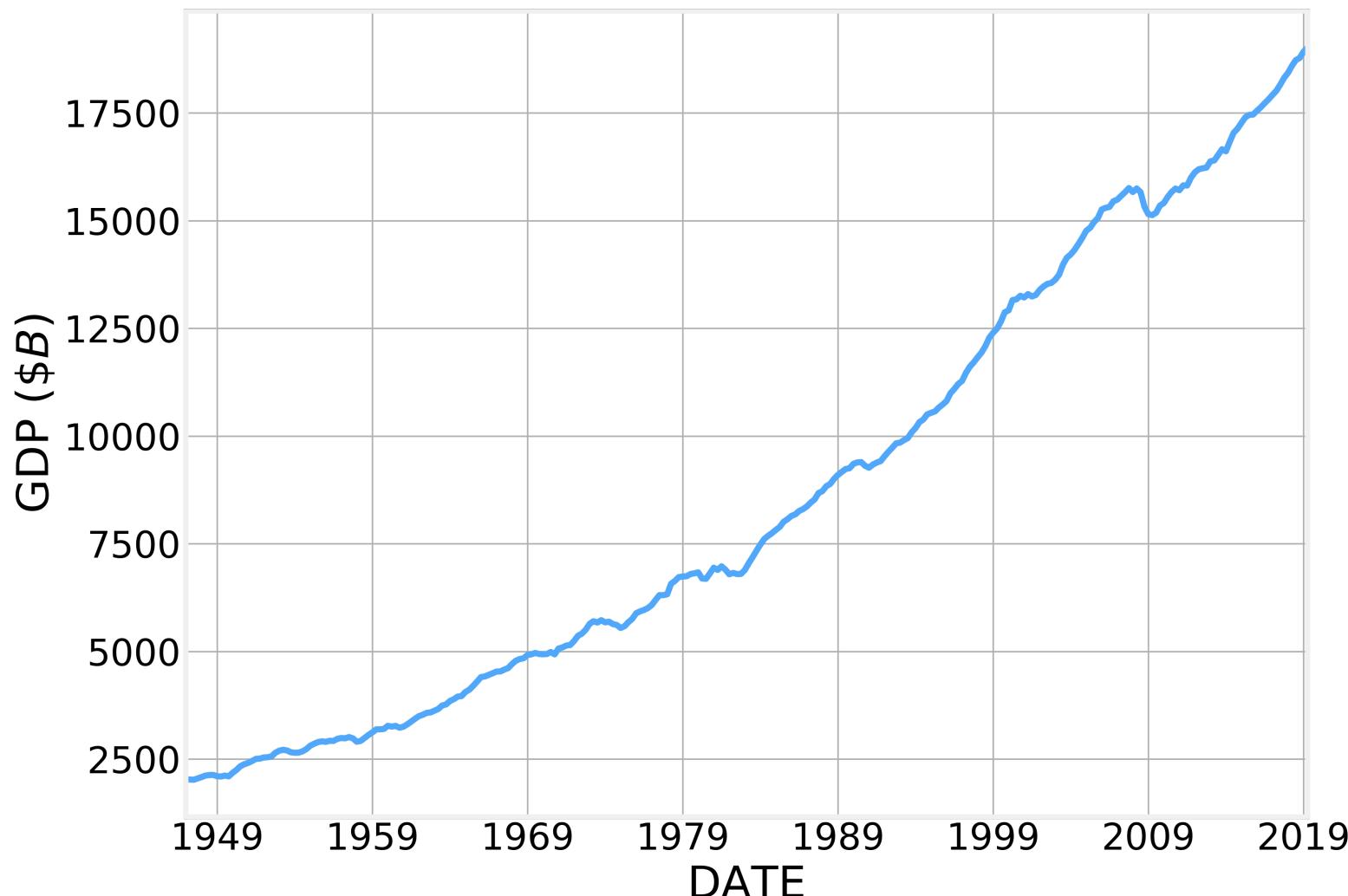
Public Interest – Search Trends



Airline Passengers

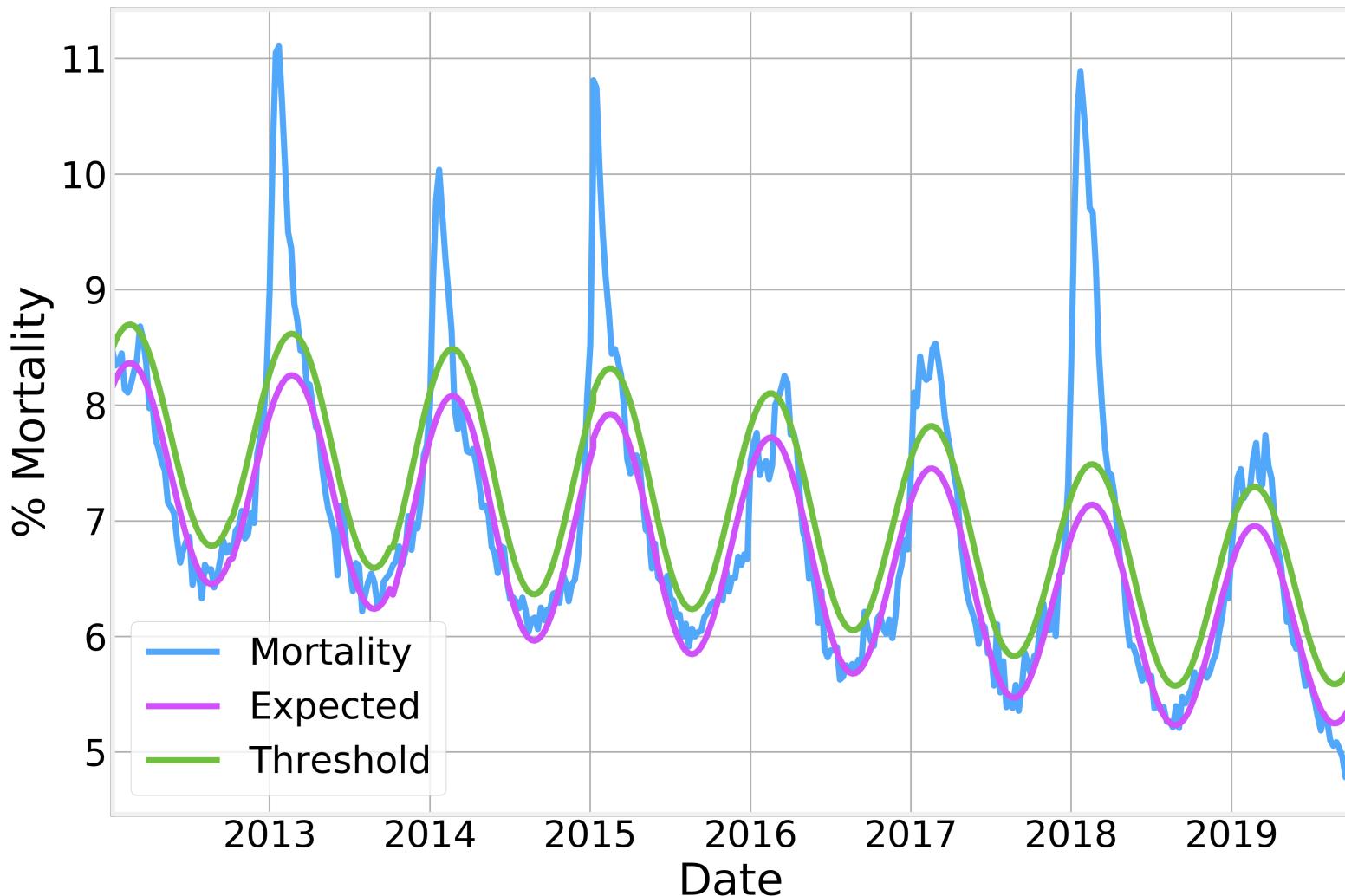


Gross Domestic Product (GDP)

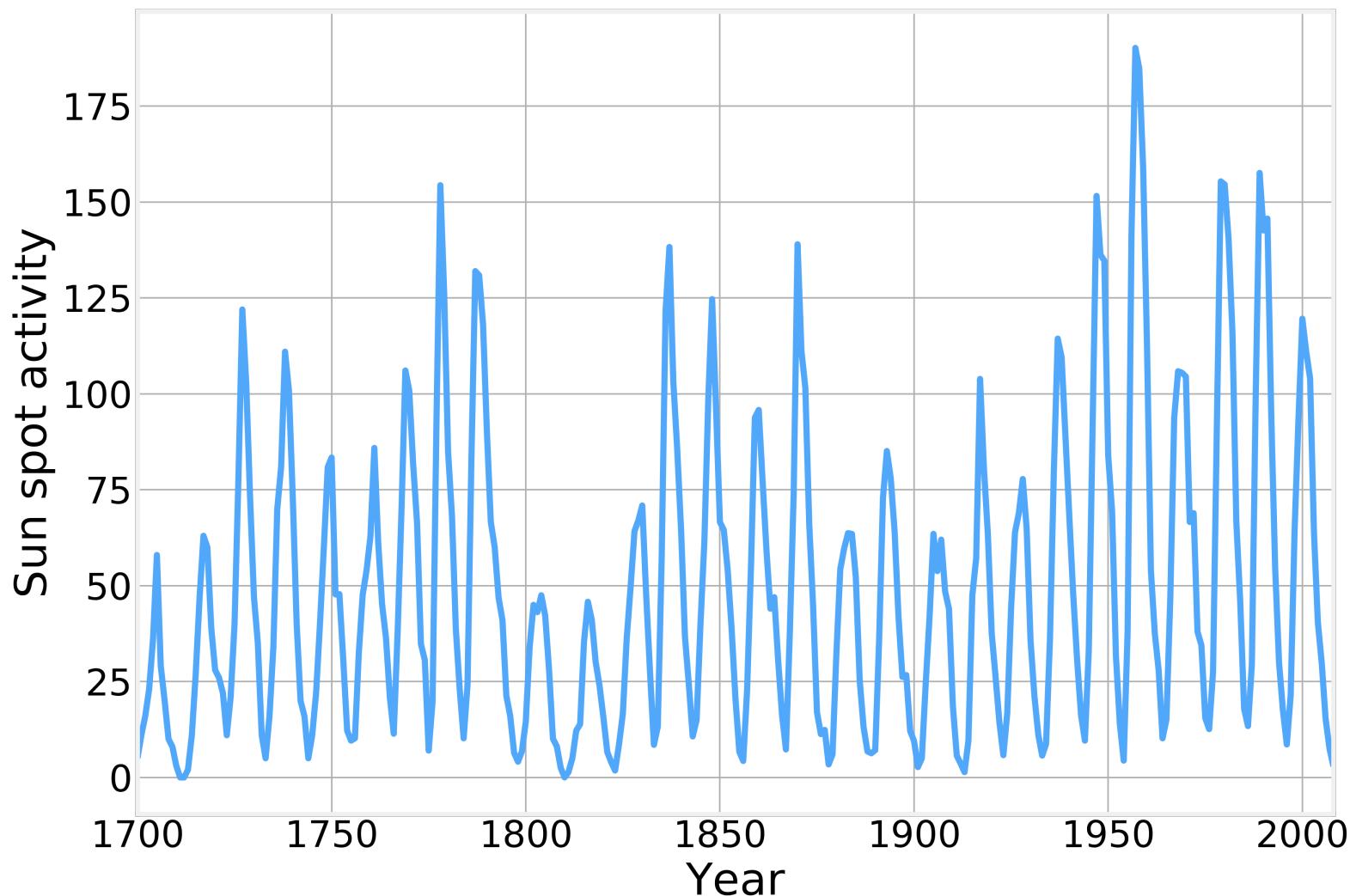


<https://fred.stlouisfed.org/series/GDPC1>

Influenza



Sunspot Activity



Stock Market - DJIA



<https://fred.stlouisfed.org/series/DJIA>

Time Series

Time Series

- A set of values measured sequentially in time

Time Series

- A set of values measured sequentially in time
- Values are typically (but not always) measured at equal intervals, x_1, x_2, x_3, x_3 , etc...

Time Series

- A set of values measured sequentially in time
- Values are typically (but not always) measured at equal intervals, x_1, x_2, x_3, x_4 , etc...
- Values can be:
 - continuous
 - discrete or symbolic (words)

Time Series

- A set of values measured sequentially in time
- Values are typically (but not always) measured at equal intervals, x_1, x_2, x_3, x_3 , etc...
- Values can be:
 - continuous
 - discrete or symbolic (words)
- Associated with empirical observation of time varying phenomena:

Time Series

- A set of values measured sequentially in time
- Values are typically (but not always) measured at equal intervals, x_1, x_2, x_3, x_3 , etc...
- Values can be:
 - continuous
 - discrete or symbolic (words)
- Associated with empirical observation of time varying phenomena:
 - Stock market prices (day, hour, minute, tick, etc...)
 - Temperatures (day, minute, second, etc...)
 - Number of patients (week, month, etc...)
 - GDP (quarter, year, etc...)

Time Series

- A set of values measured sequentially in time
- Values are typically (but not always) measured at equal intervals, x_1, x_2, x_3, x_3 , etc...
- Values can be:
 - continuous
 - discrete or symbolic (words)
- Associated with empirical observation of time varying phenomena:
 - Stock market prices (day, hour, minute, tick, etc...)
 - Temperatures (day, minute, second, etc...)
 - Number of patients (week, month, etc...)
 - GDP (quarter, year, etc...)
- Forecasting requires predicting future values based on past behavior

Time Series

- The value of the time series at time t is given by x_t

Time Series

- The value of the time series at time t is given by x_t
- The values at a given lag l are given by x_{t-l}

Time Series

- The value of the time series at time t is given by x_t
- The values at a given lag l are given by x_{t-l}
- The mean of the overall signal is μ and the corresponding running value is μ_t^w

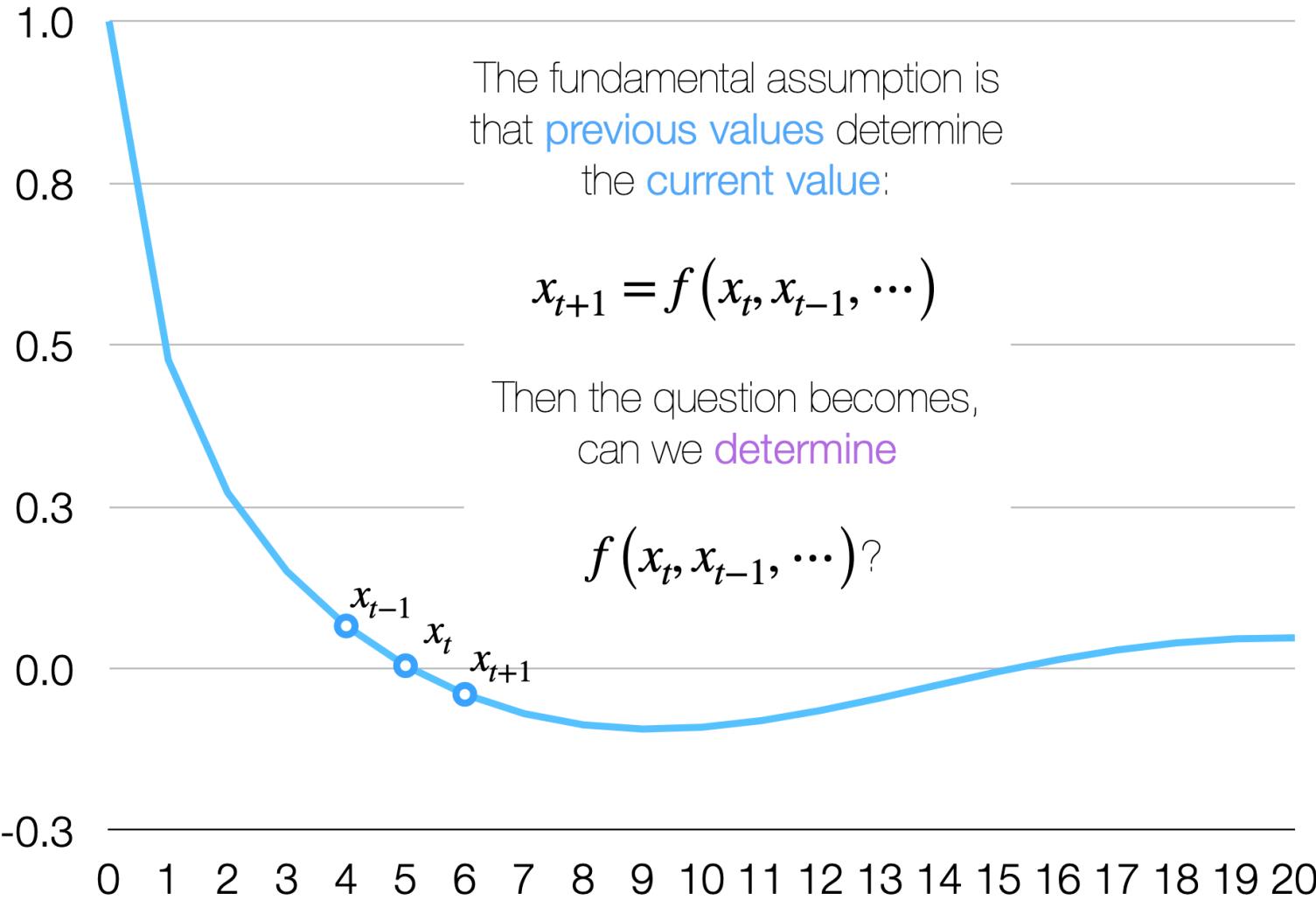
Time Series

- The value of the time series at time t is given by x_t
- The values at a given lag l are given by x_{t-l}
- The mean of the overall signal is μ and the corresponding running value is μ_t^w
- The variance of the overall signal is " and the corresponding running value is σ_t^w

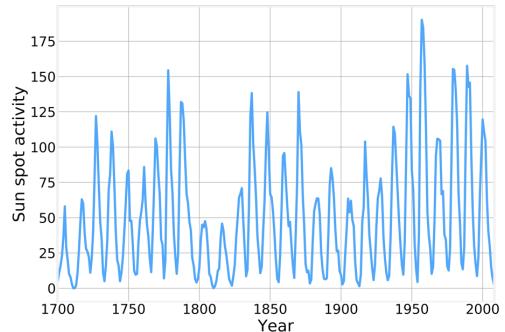
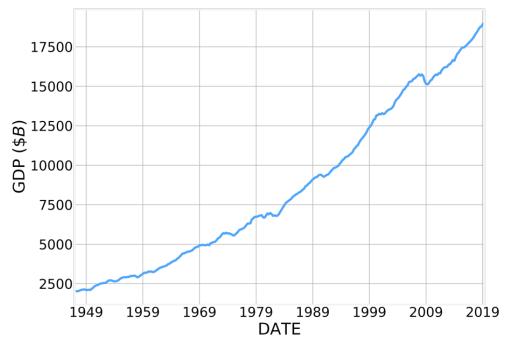
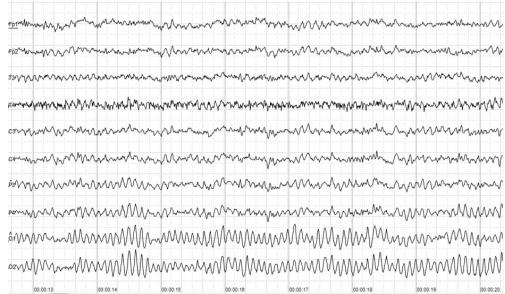
Time Series

- The value of the time series at time t is given by x_t
- The values at a given lag l are given by x_{t-l}
- The mean of the overall signal is μ and the corresponding running value is μ_t^w
- The variance of the overall signal is " and the corresponding running value is σ_t^w
- Running values are calculated over a window of width w

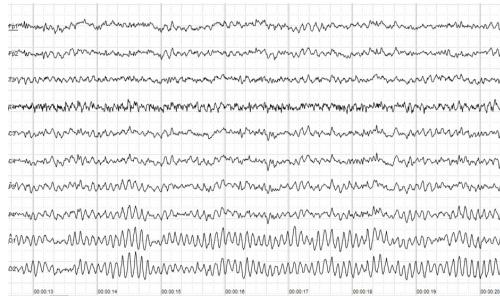
Time Series Analysis



Three fundamental behaviors

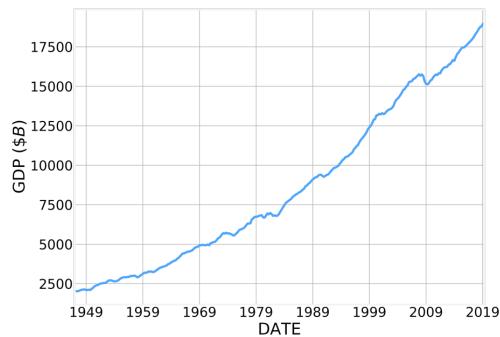


Three fundamental behaviors



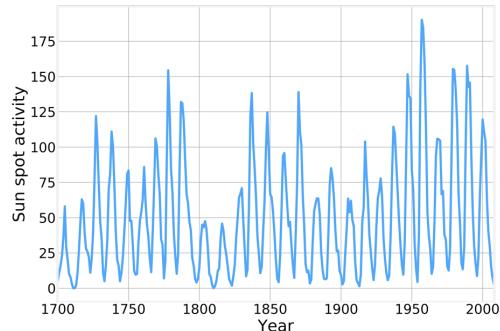
Stationarity

$$\langle x_t \rangle \approx \text{constant}$$



Trend

$$\langle x_t \rangle \approx ct$$



Seasonality

$$X_{t+T} \approx X_t$$

Stationarity

- A time series is said to be stationary if its basic statistical properties are **independent of time**

Stationarity

- A time series is said to be stationary if its basic statistical properties are **independent of time**
- In particular:
 - **Mean** - Average value stays constant

Stationarity

- A time series is said to be stationary if its basic statistical properties are **independent of time**
- In particular:
 - **Mean** - Average value stays constant
 - **Variance** – The width of the curve is bounded

Stationarity

- A time series is said to be stationary if its basic statistical properties are **independent of time**
- In particular:
 - **Mean** - Average value stays constant
 - **Variance** – The width of the curve is bounded
 - **Covariance** – Correlation between points is independent of time

Stationarity

- A time series is said to be stationary if its basic statistical properties are **independent of time**
- In particular:
 - **Mean** - Average value stays constant
 - **Variance** – The width of the curve is bounded
 - **Covariance** – Correlation between points is independent of time
- Stationary processes are **easier** to analyze

Stationarity

- A time series is said to be stationary if its basic statistical properties are **independent of time**
- In particular:
 - **Mean** - Average value stays constant
 - **Variance** – The width of the curve is bounded
 - **Covariance** – Correlation between points is independent of time
- Stationary processes are **easier** to analyze
- Many time series analysis algorithms assume the time series to be stationary

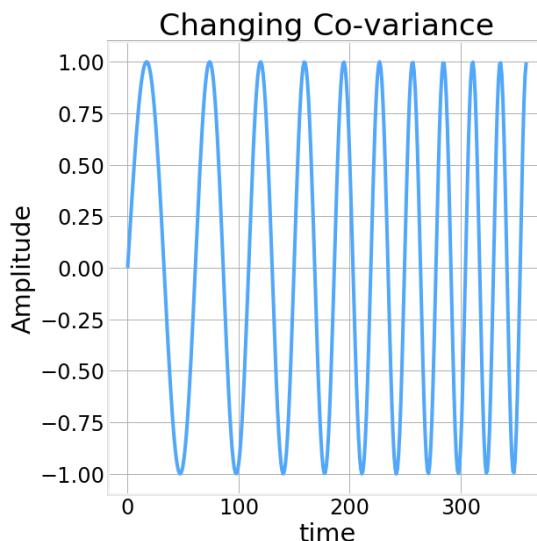
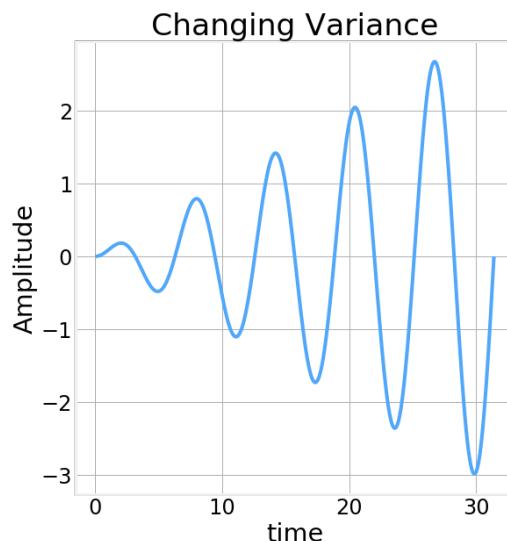
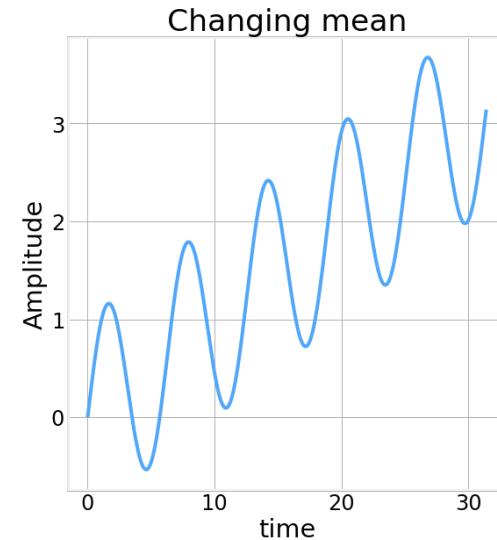
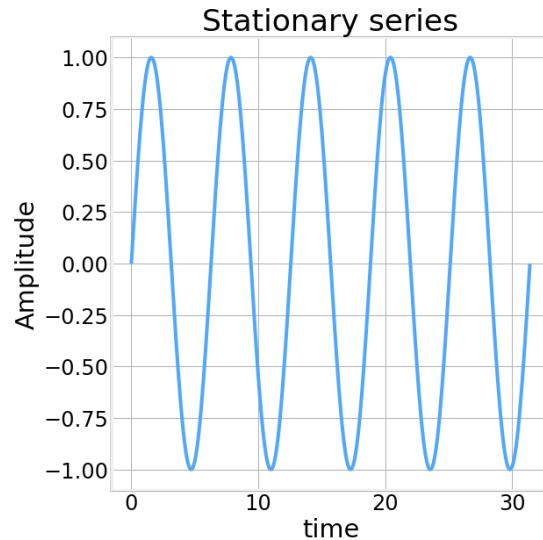
Stationarity

- A time series is said to be stationary if its basic statistical properties are **independent of time**
- In particular:
 - **Mean** - Average value stays constant
 - **Variance** – The width of the curve is bounded
 - **Covariance** – Correlation between points is independent of time
- Stationary processes are **easier** to analyze
- Many time series analysis algorithms assume the time series to be stationary
- Several rigorous tests for stationarity have been developed such as the **(Augmented) Dickey-Fuller** and **Hurst Exponent**

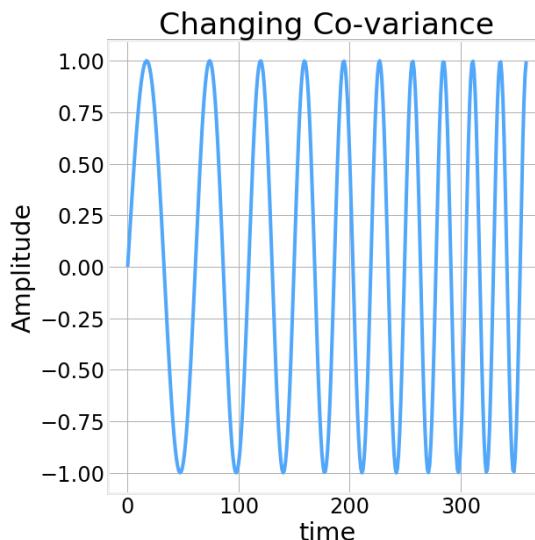
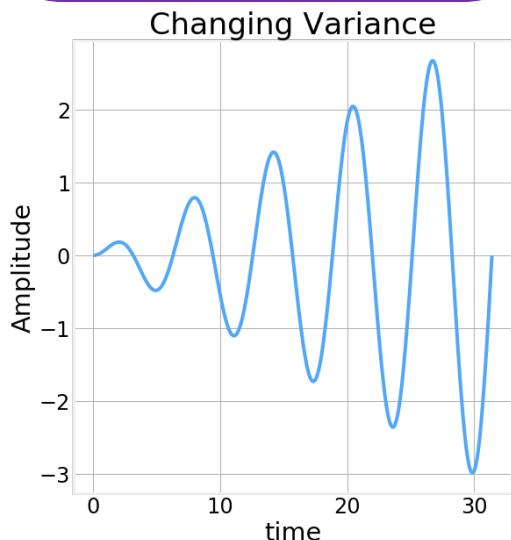
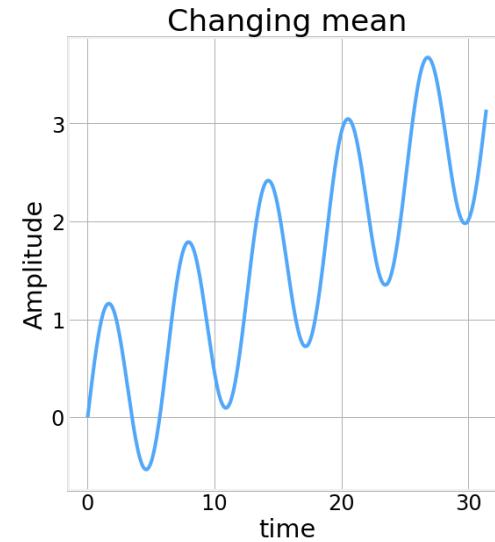
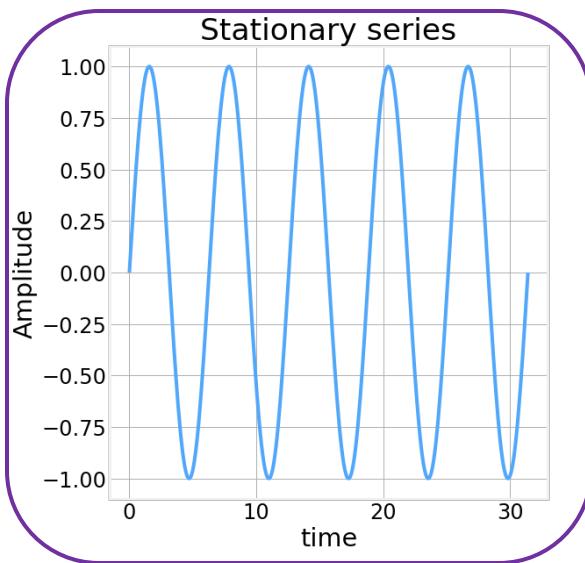
Stationarity

- A time series is said to be stationary if its basic statistical properties are **independent of time**
- In particular:
 - **Mean** - Average value stays constant
 - **Variance** – The width of the curve is bounded
 - **Covariance** – Correlation between points is independent of time
- Stationary processes are **easier** to analyze
- Many time series analysis algorithms assume the time series to be stationary
- Several rigorous tests for stationarity have been developed such as the **(Augmented) Dickey-Fuller** and **Hurst Exponent**
- Typically, the first step of any analysis is to transform the series to make it stationary

Stationarity



Stationarity



Code – Data Exploration

Trend

- Many time series have a clear trend or tendency:

Trend

- Many time series have a clear trend or tendency:
 - Stock market indices tend to go up over time

Trend

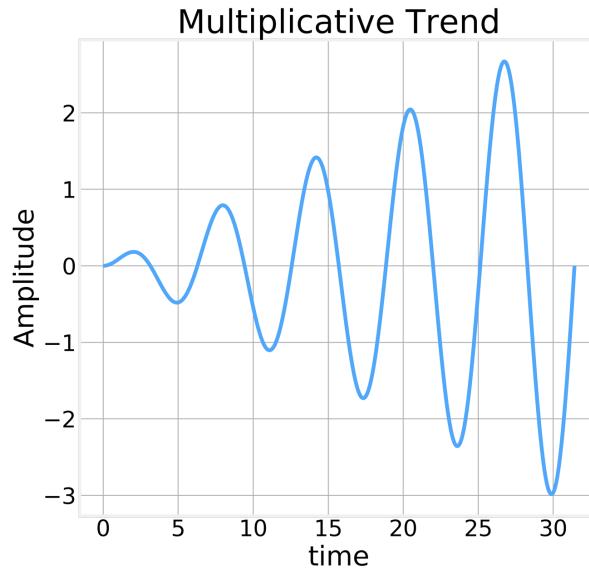
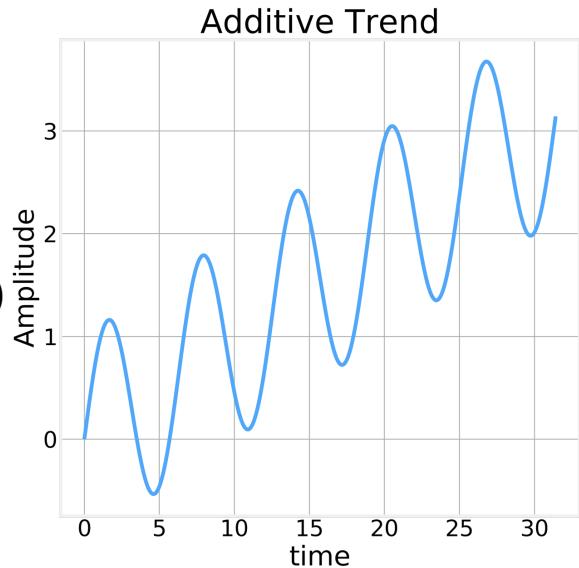
- Many time series have a clear trend or tendency:
 - Stock market indices tend to go up over time
 - Number of cases of preventable diseases tends to go down over time

Trend

- Many time series have a clear trend or tendency:
 - Stock market indices tend to go up over time
 - Number of cases of preventable diseases tends to go down over time
- Trends can be additive or multiplicative:

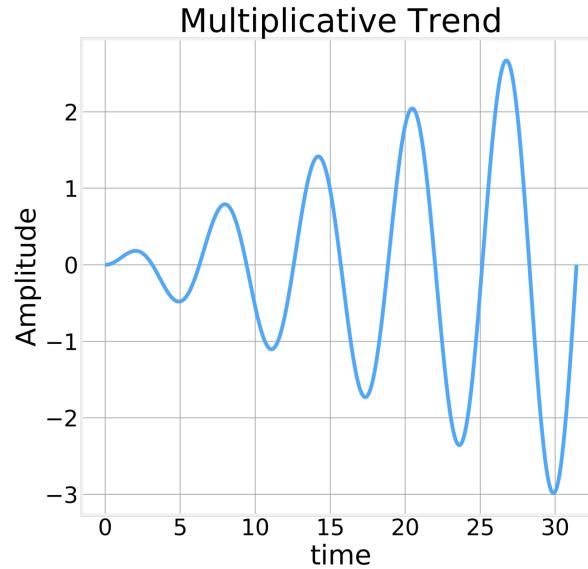
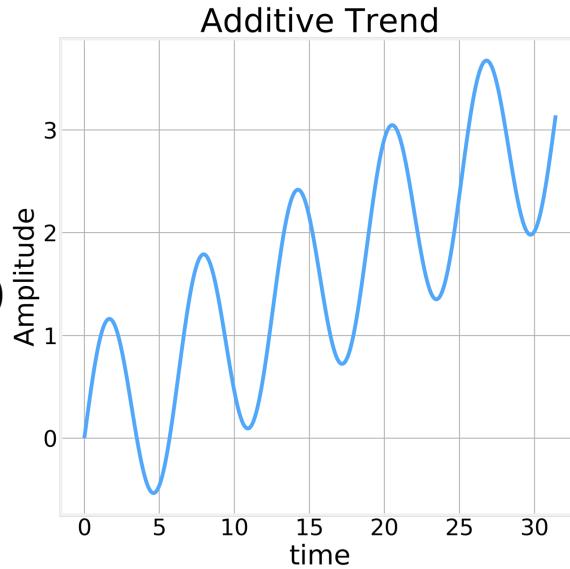
Trend

- Many time series have a clear trend or tendency:
 - Stock market indices tend to go up over time
 - Number of cases of preventable diseases tends to go down over time
- Trends can be additive or multiplicative:



Trend

- Many time series have a clear trend or tendency:
 - Stock market indices tend to go up over time
 - Number of cases of preventable diseases tends to go down over time
- Trends can be additive or multiplicative:



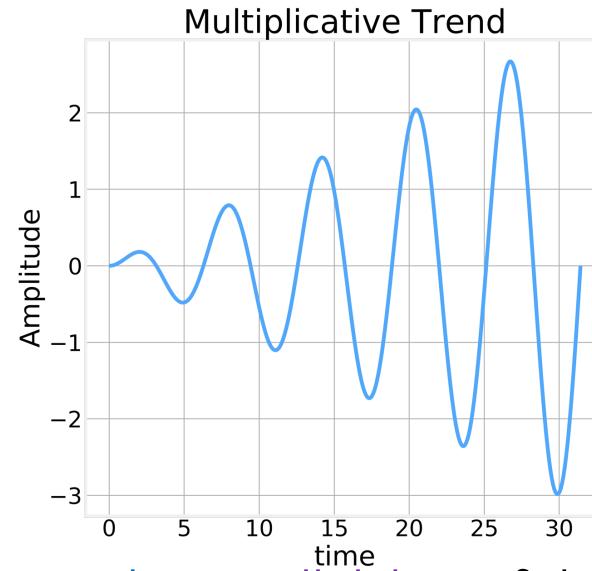
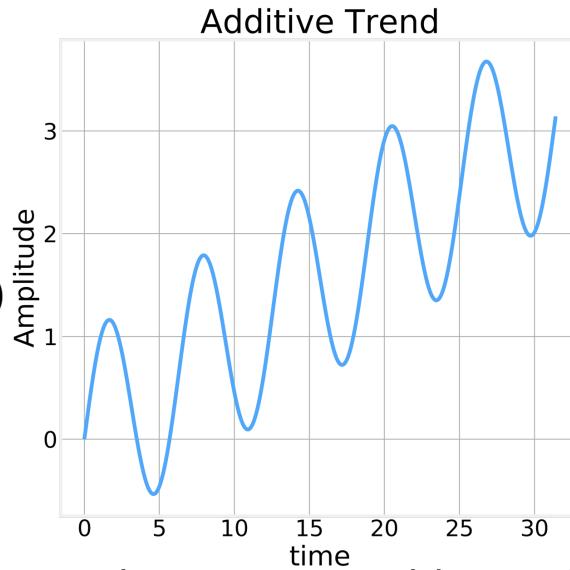
$$x = \frac{t}{10} + \sin(t)$$

$$x = \frac{t}{10} \sin(t)$$

- Trends can be removed by subtraction or division of the correct values

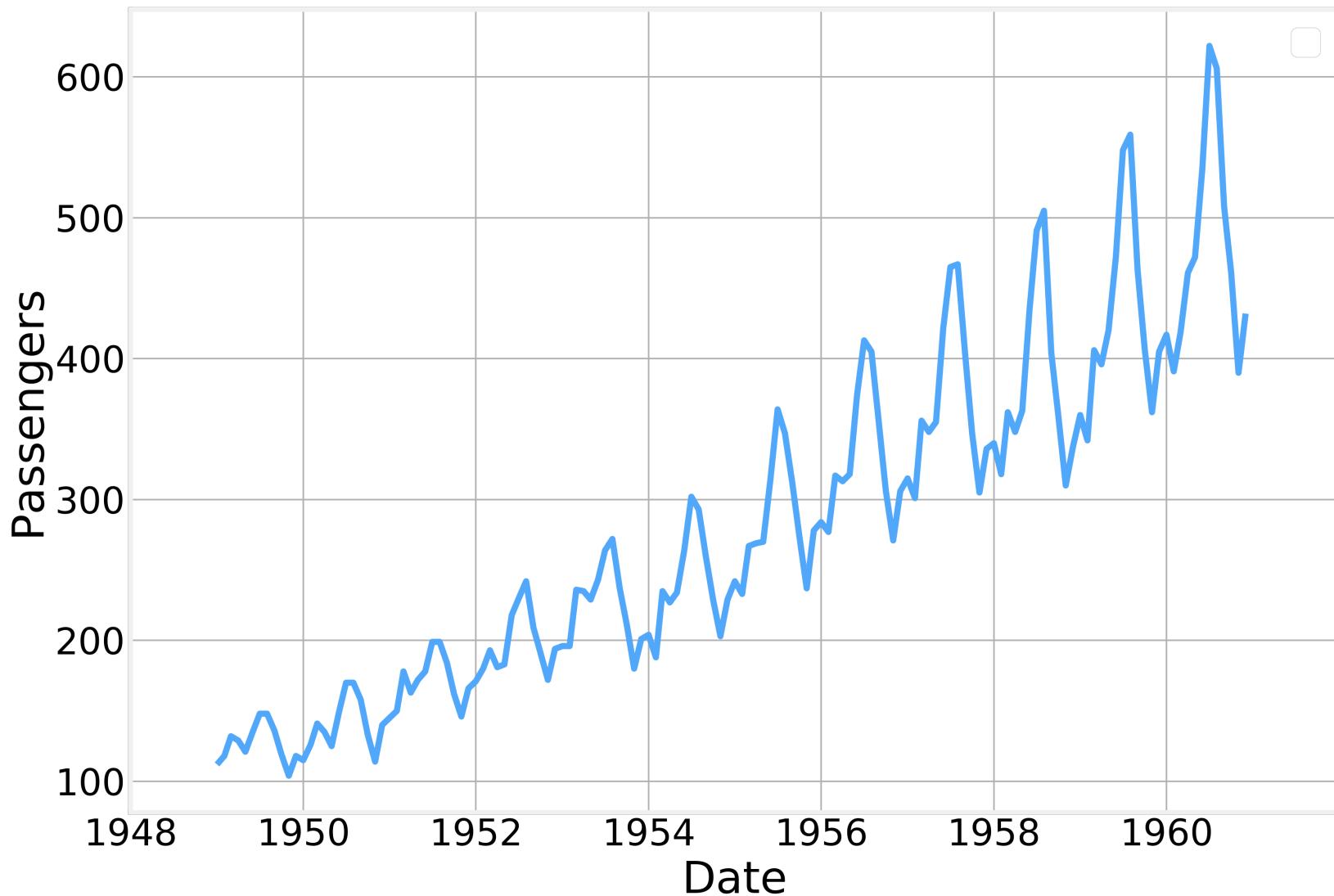
Trend

- Many time series have a clear trend or tendency:
 - Stock market indices tend to go up over time
 - Number of cases of preventable diseases tends to go down over time
- Trends can be additive or multiplicative:

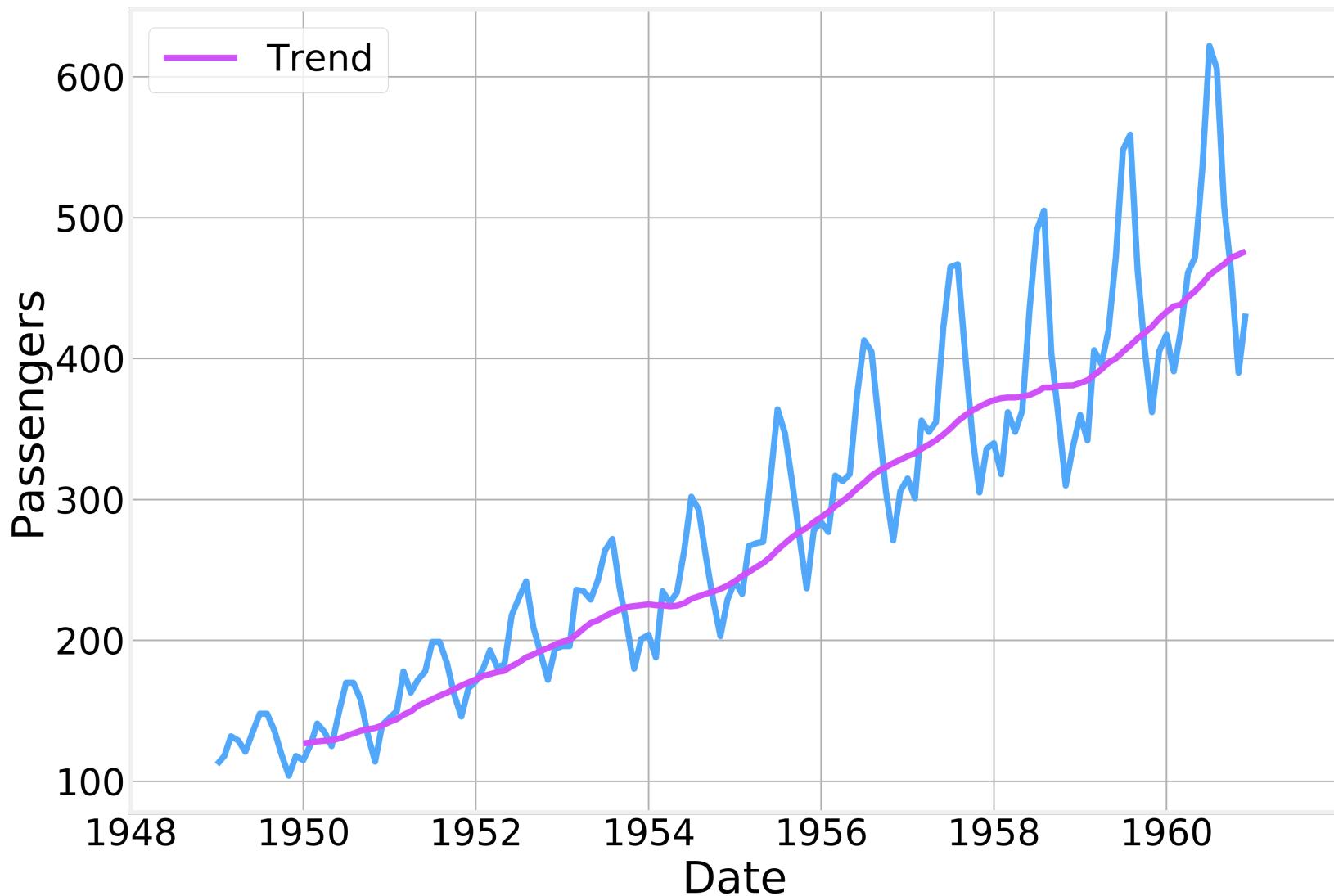


- Trends can be removed by subtraction or division of the correct values
- One way to determine the trend is to find the running average

Trend



Trend



Seasonality

Seasonality

- Many of the phenomena we might be interested in vary in time in a **cyclical** or **seasonal** fashion

Seasonality

- Many of the phenomena we might be interested in vary in time in a **cyclical** or **seasonal** fashion
 - Ice-cream sales peak in the **summer** and drop in the **winter**

Seasonality

- Many of the phenomena we might be interested in vary in time in a **cyclical** or **seasonal** fashion
 - Ice-cream sales peak in the **summer** and drop in the **winter**
 - Number of cell phone calls made is larger during the day than at night

Seasonality

- Many of the phenomena we might be interested in vary in time in a **cyclical** or **seasonal** fashion
 - Ice-cream sales peak in the **summer** and drop in the **winter**
 - Number of cell phone calls made is larger during the day than at night
 - Many types of crime are more frequent at **night** than during the **day**

Seasonality

- Many of the phenomena we might be interested in vary in time in a **cyclical** or **seasonal** fashion
 - Ice-cream sales peak in the **summer** and drop in the **winter**
 - Number of cell phone calls made is larger during the day than at night
 - Many types of crime are more frequent at **night** than during the **day**
 - Visits to museums are more frequent in the **weekend** than in **weekdays**

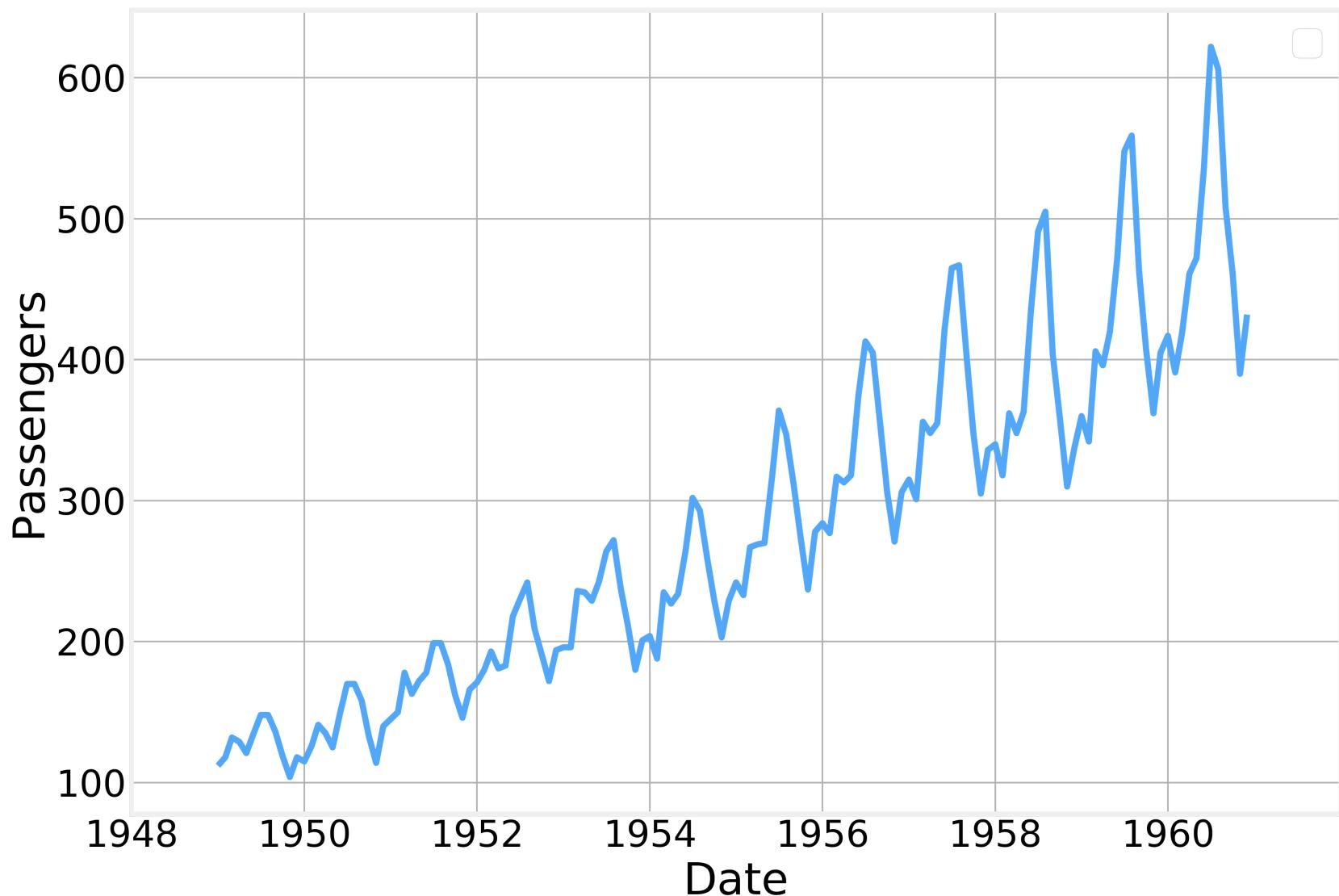
Seasonality

- Many of the phenomena we might be interested in vary in time in a **cyclical** or **seasonal** fashion
 - Ice-cream sales peak in the **summer** and drop in the **winter**
 - Number of cell phone calls made is larger during the day than at night
 - Many types of crime are more frequent at **night** than during the **day**
 - Visits to museums are more frequent in the **weekend** than in **weekdays**
 - The stock market grows during **bull** periods and shrinks during **bear** periods

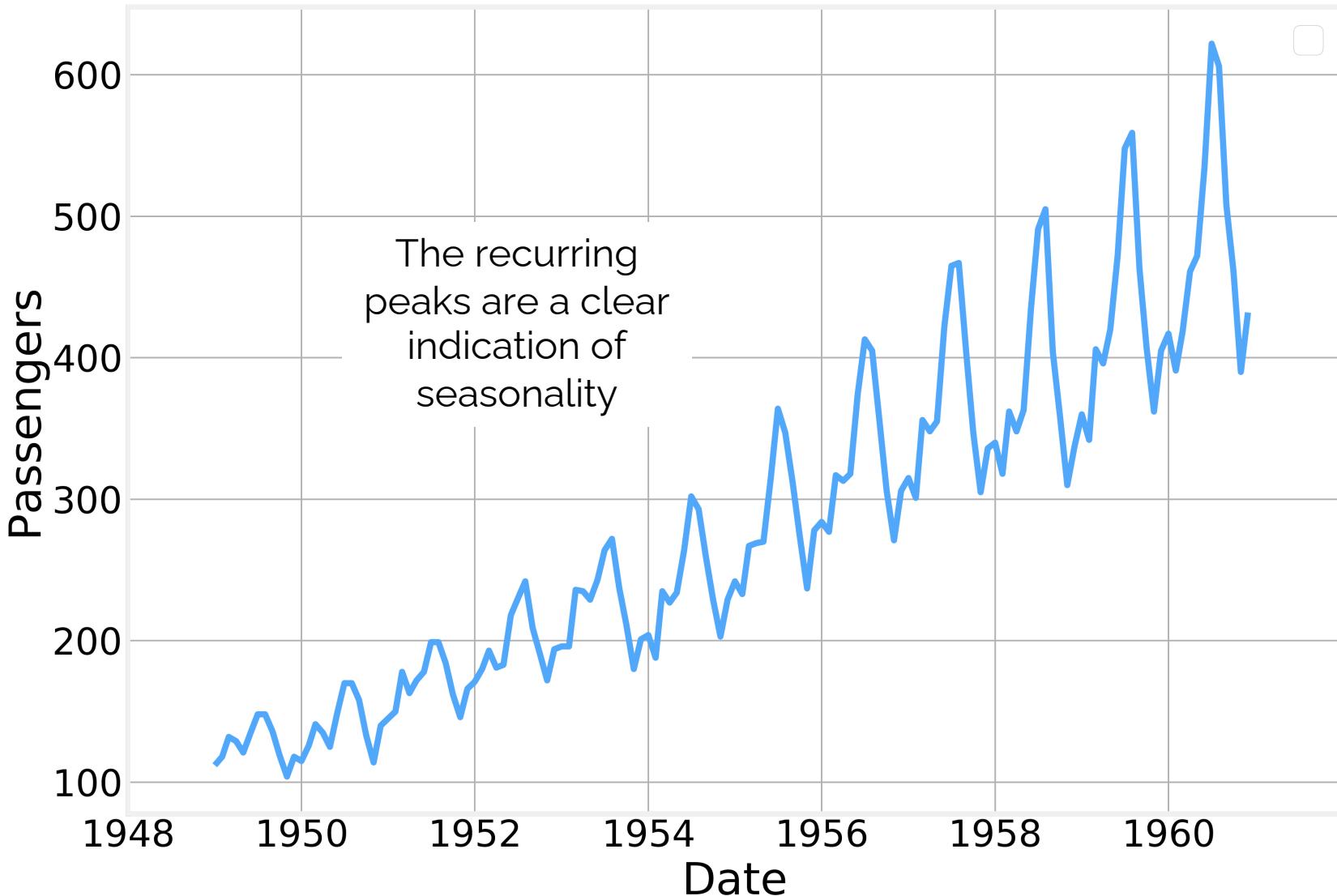
Seasonality

- Many of the phenomena we might be interested in vary in time in a **cyclical** or **seasonal** fashion
 - Ice-cream sales peak in the **summer** and drop in the **winter**
 - Number of cell phone calls made is larger during the day than at night
 - Many types of crime are more frequent at **night** than during the **day**
 - Visits to museums are more frequent in the **weekend** than in **weekdays**
 - The stock market grows during **bull** periods and shrinks during **bear** periods
- Understanding the seasonality of a time series provides important information about its long-term behavior and is extremely useful in predicting future values

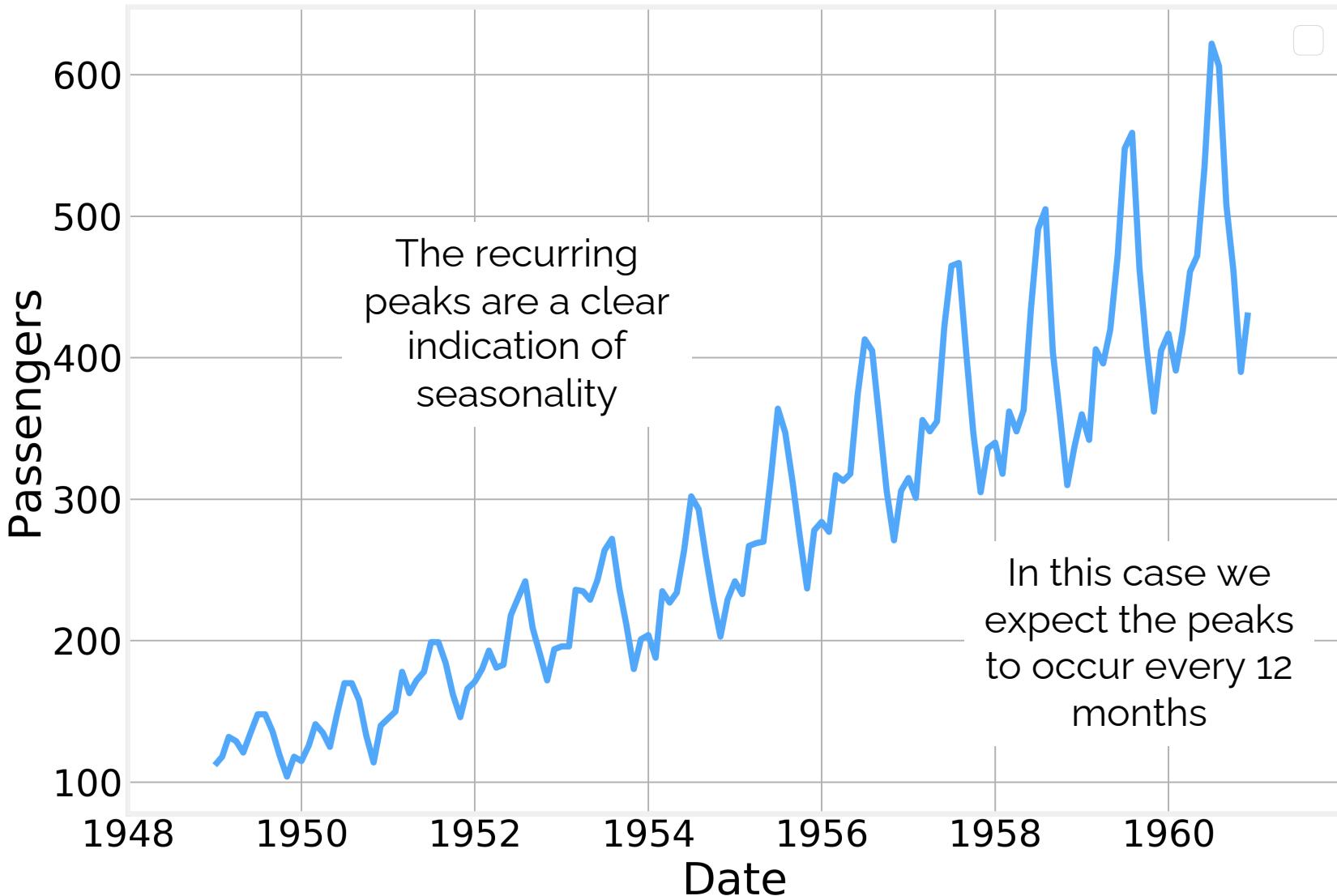
Seasonality



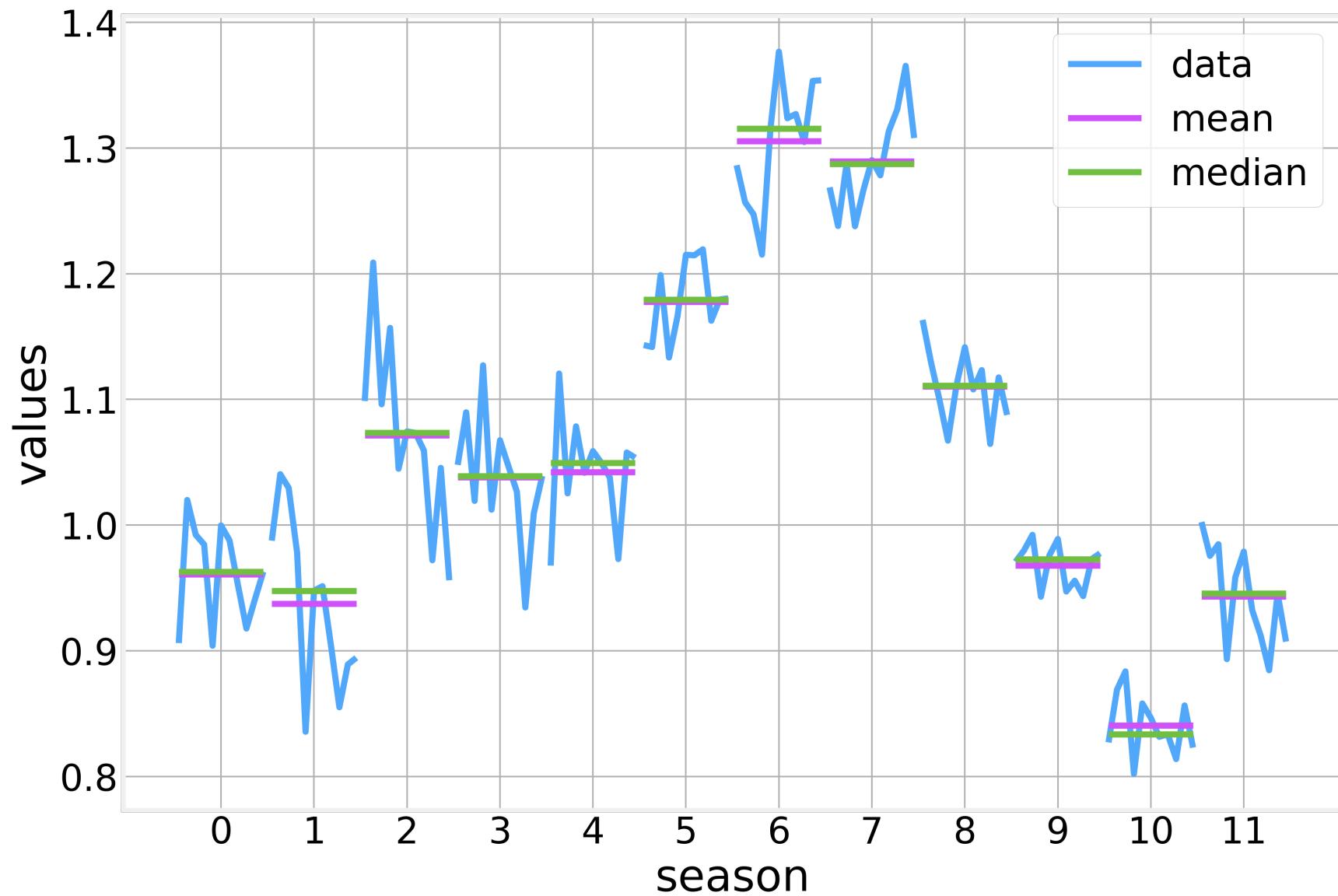
Seasonality



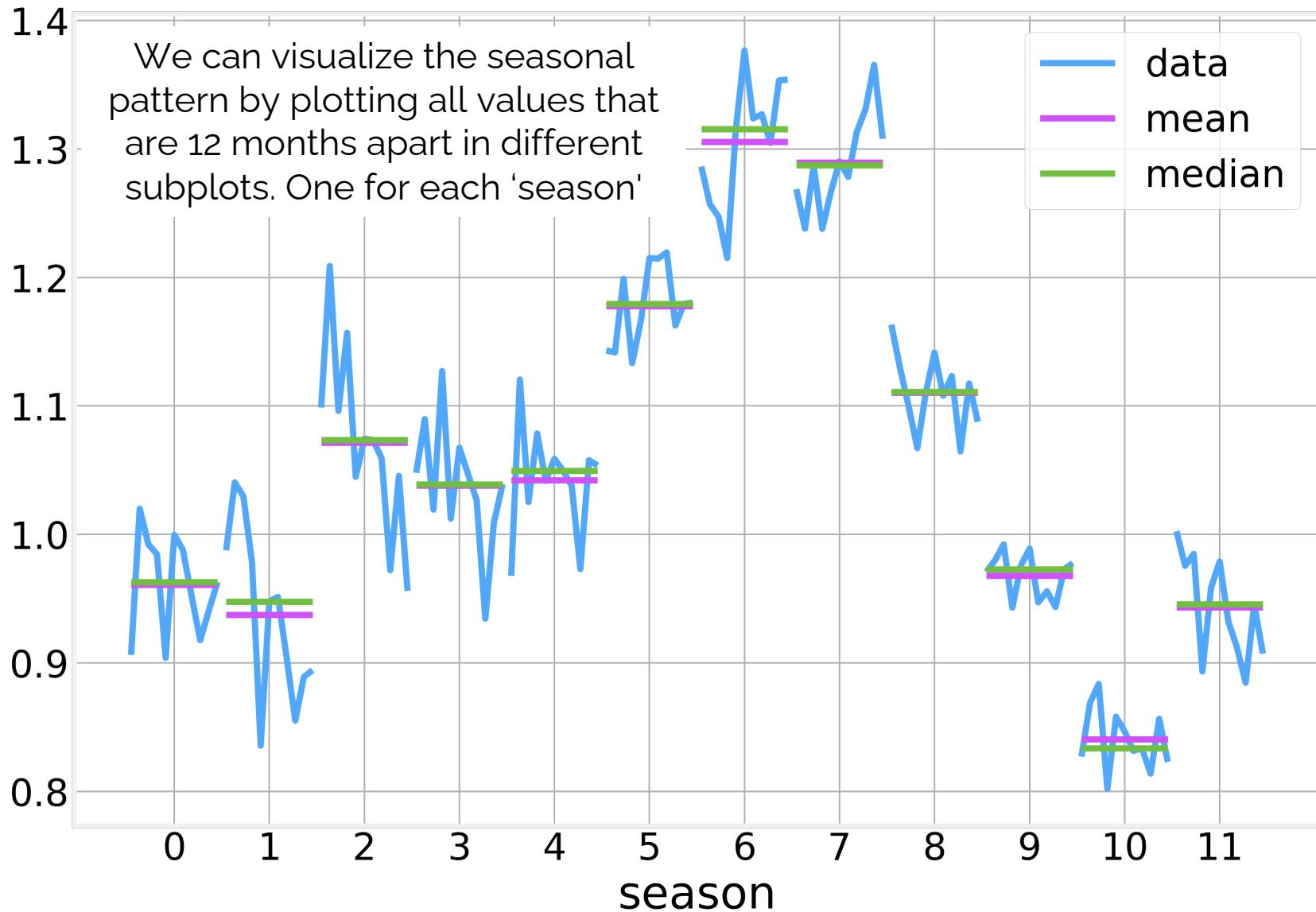
Seasonality



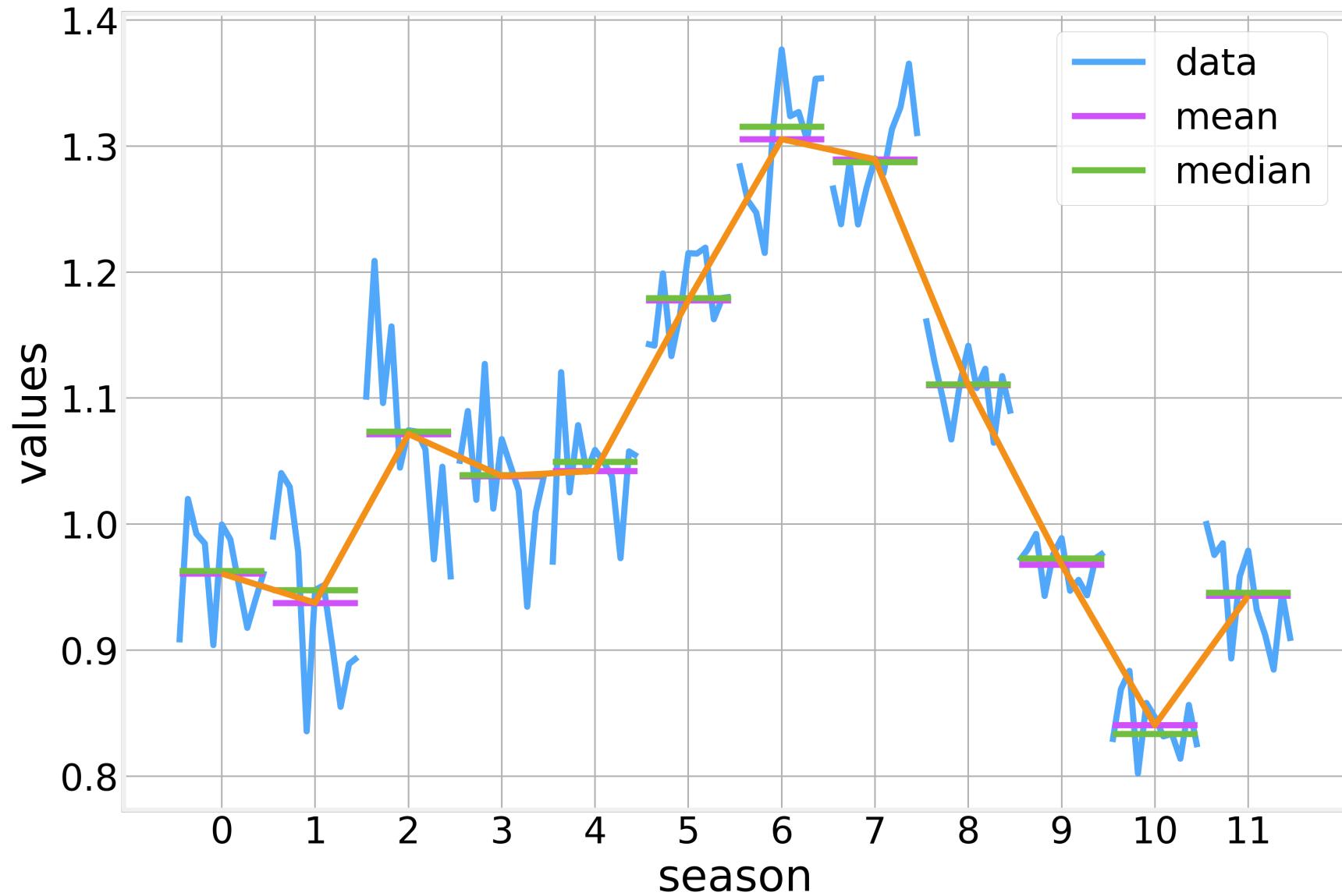
Seasonality



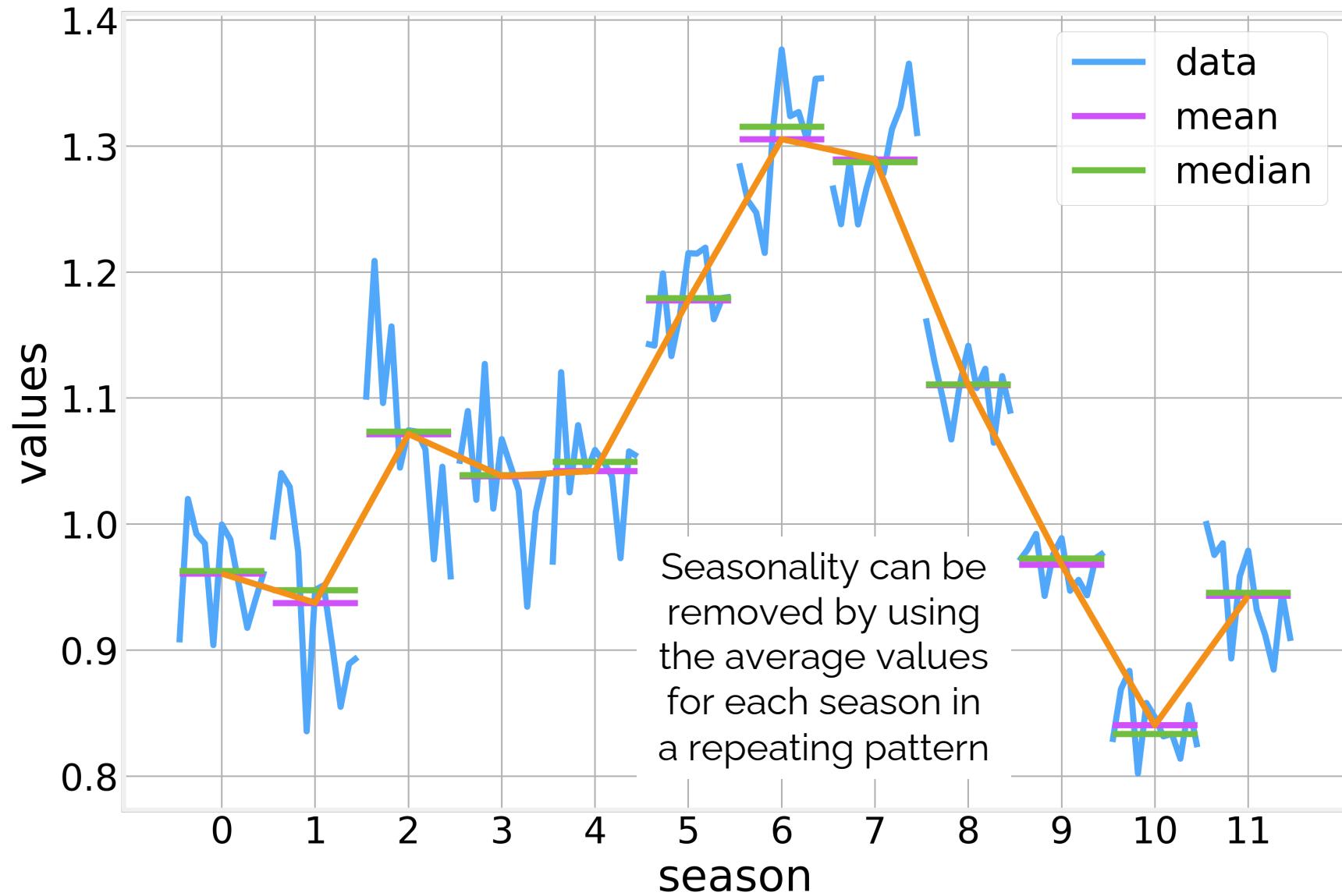
Seasonality



Seasonality



Seasonality



Time Series Decomposition

- A time series can be decomposed into **three components**:

Time Series Decomposition

- A time series can be decomposed into **three components**:
 - Trend, T_t
 - Seasonality, S_t
 - Residuals, R_t

Time Series Decomposition

- A time series can be decomposed into **three components**:
 - Trend, T_t
 - Seasonality, S_t
 - Residuals, R_t
- Decompositions can be

Time Series Decomposition

- A time series can be decomposed into **three components**:
 - Trend, T_t
 - Seasonality, S_t
 - Residuals, R_t
- Decompositions can be
 - additive - $x_t = T_t + S_t + R_t$
 - multiplicative - $x_t = Tt \cdot St \cdot Rt$

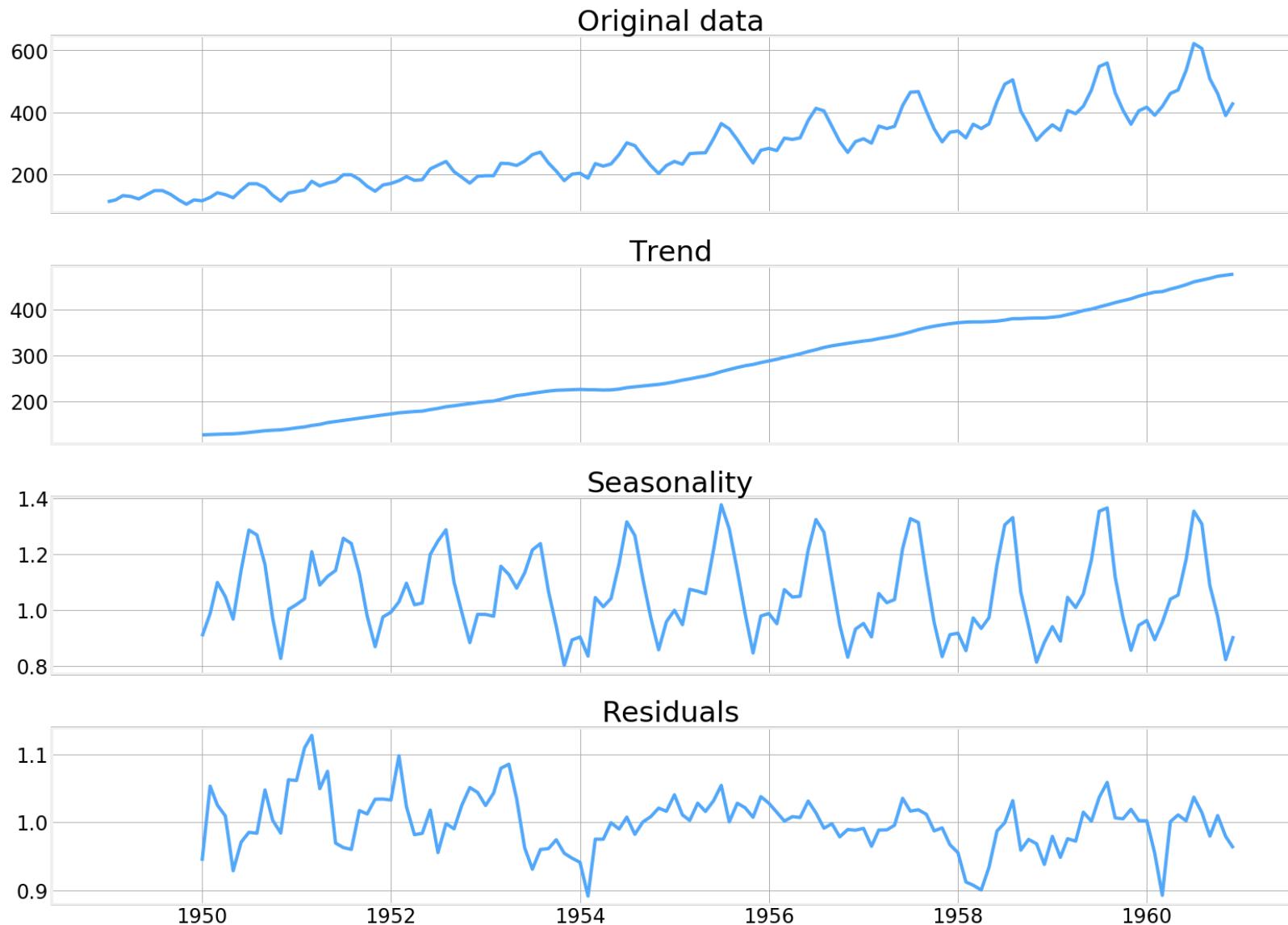
Time Series Decomposition

- A time series can be decomposed into **three components**:
 - Trend, T_t
 - Seasonality, S_t
 - Residuals, R_t
- Decompositions can be
 - additive - $x_t = T_t + S_t + R_t$
 - multiplicative - $x_t = Tt \cdot St \cdot Rt$
- The residuals are simply what is left of the original signal after we **remove the trend and the seasonality**

Time Series Decomposition

- A time series can be decomposed into **three components**:
 - Trend, T_t
 - Seasonality, S_t
 - Residuals, R_t
- Decompositions can be
 - additive - $x_t = T_t + S_t + R_t$
 - multiplicative - $x_t = Tt \cdot St \cdot Rt$
- The residuals are simply what is left of the original signal after we **remove the trend and the seasonality**
- Residuals are typically **stationary**

Time Series Decomposition



Code – Decomposition