



# Neural Data Science with **Python**

## L2 : Time Series

*Michael Graupner*

*SPPIN – Saint-Pères Institute for the Neurosciences*

*Université de Paris, CNRS*

# Examples of time series

?



# Examples of time series: stock prices

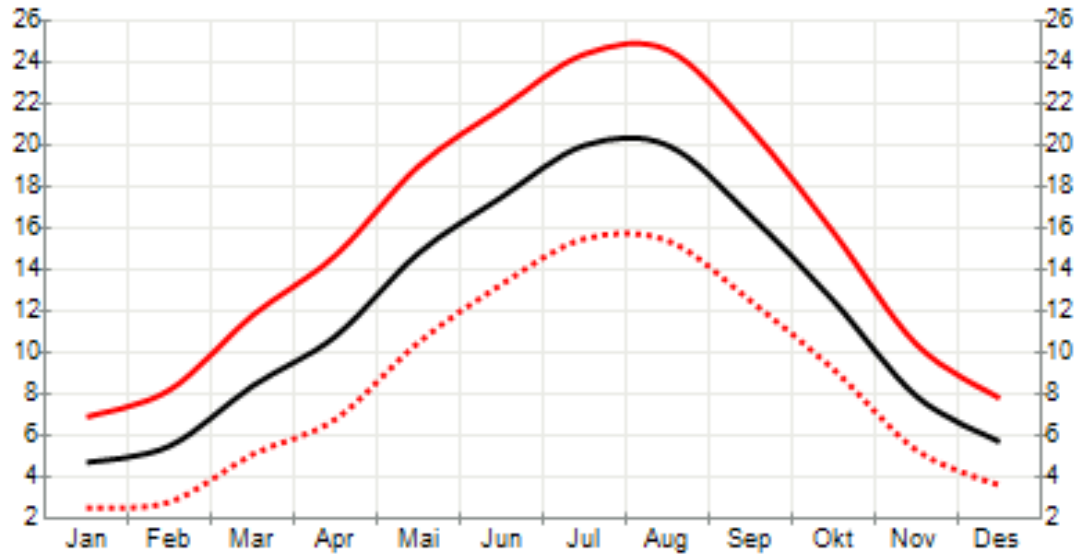
## Apple's shares in 2018

Share price in US Dollars



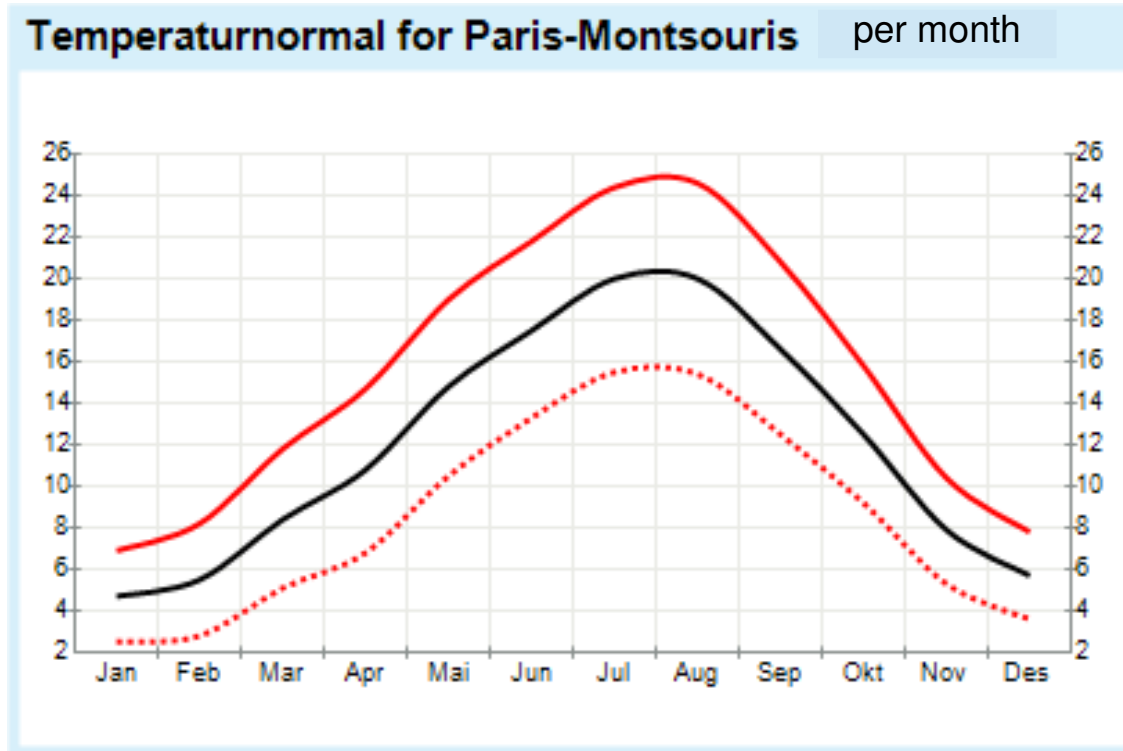
Source: Bloomberg. Last update: 21/11/2018, 8:00am GMT

# Examples of time series



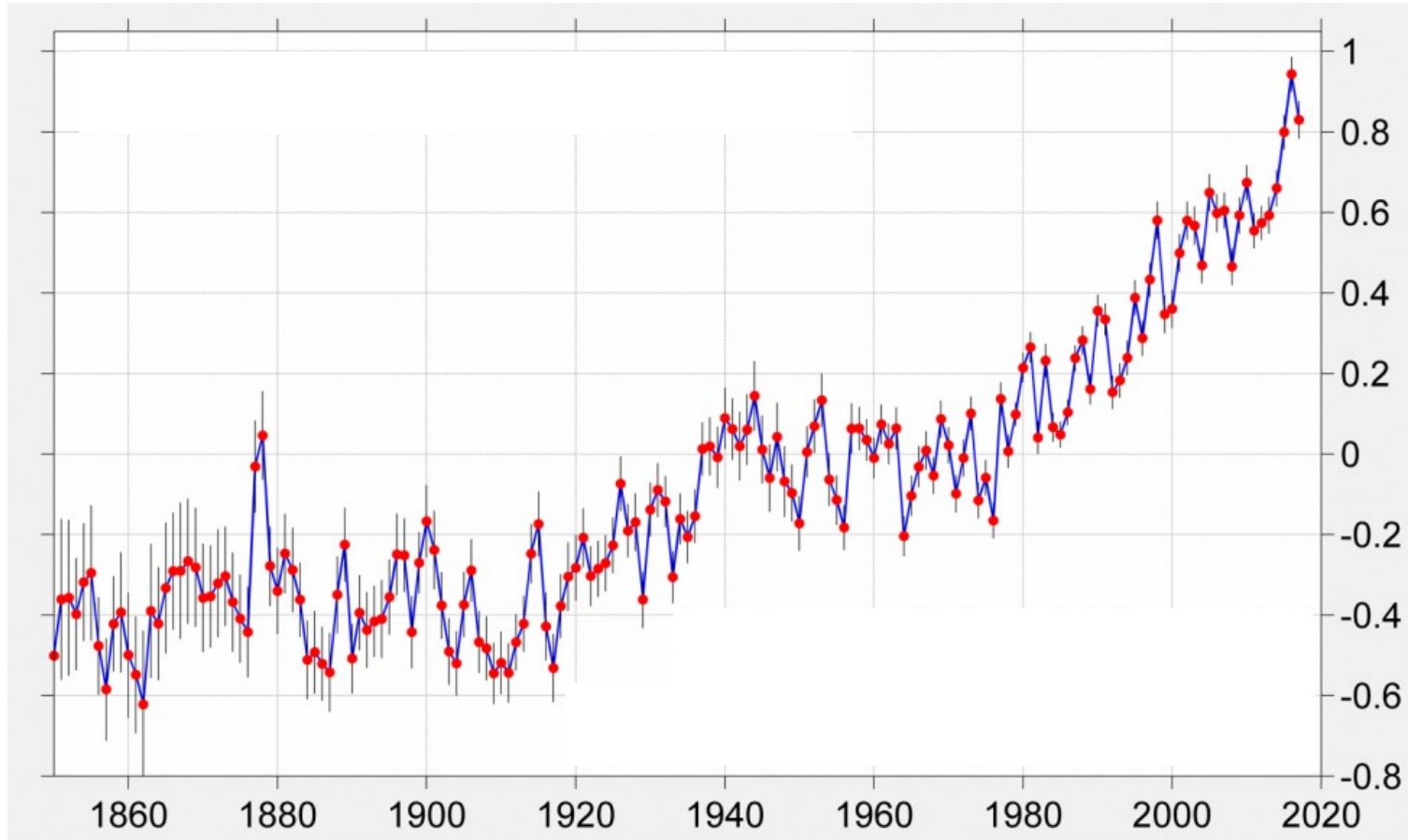
?

# Examples of time series: temperature profiles



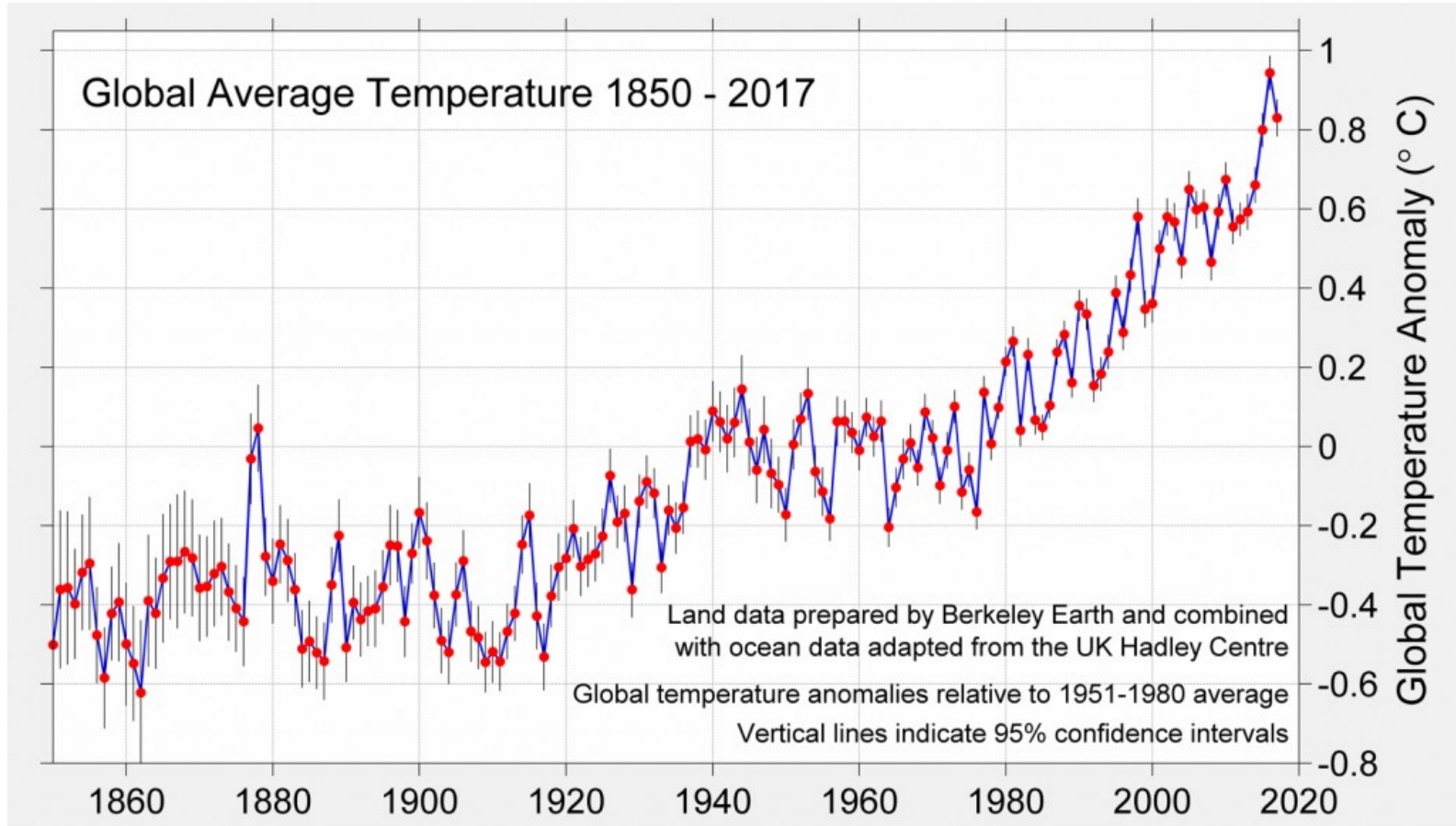
- Max temperature: Average max daily (24h) temperature per month
- Minimum temperature: Average minimum daily (24h) temperature per month
- Average temperature: Average daily (24h) temperature per month
- The temperature normals are measured in the period 1961–1990.

# Examples of time series



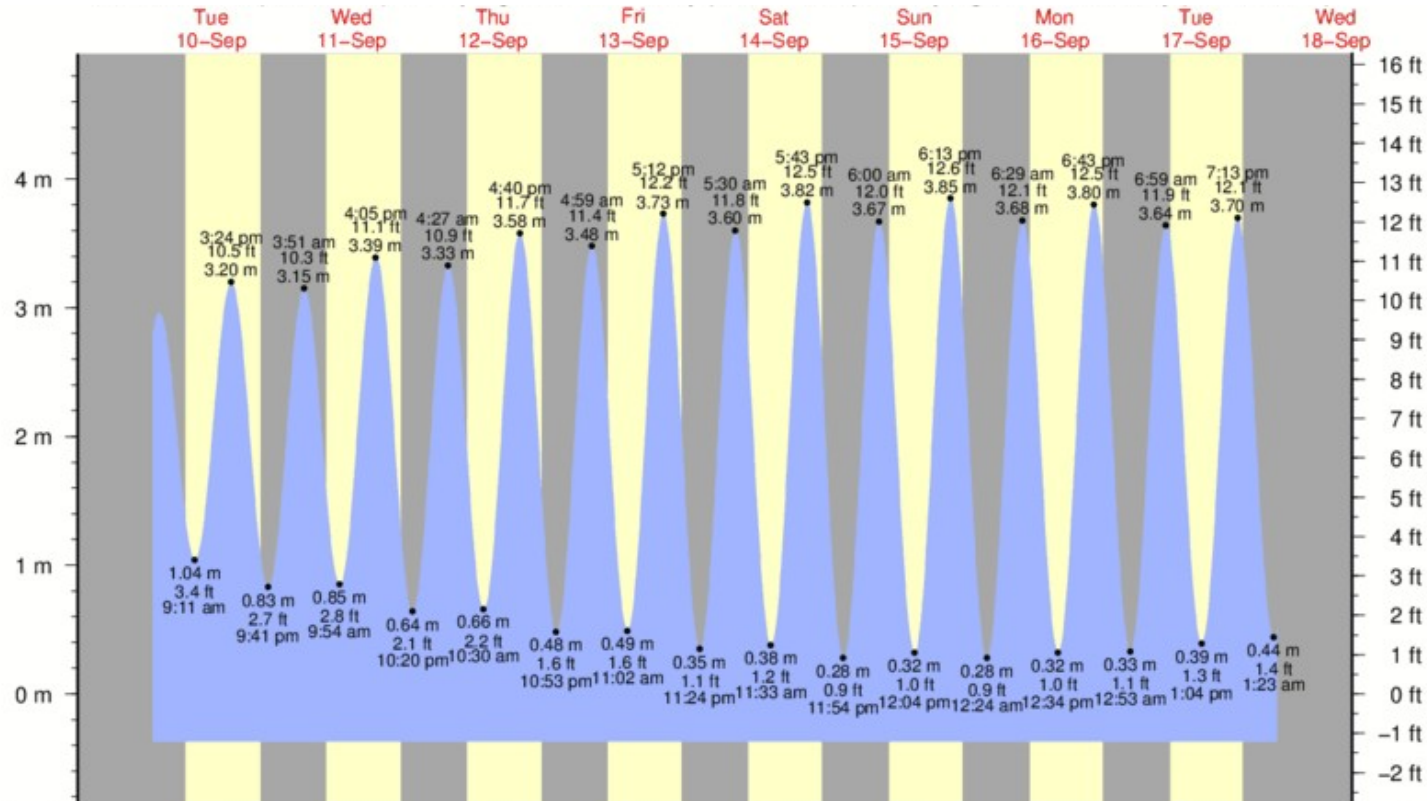
?

# Examples of time series: temperature profiles



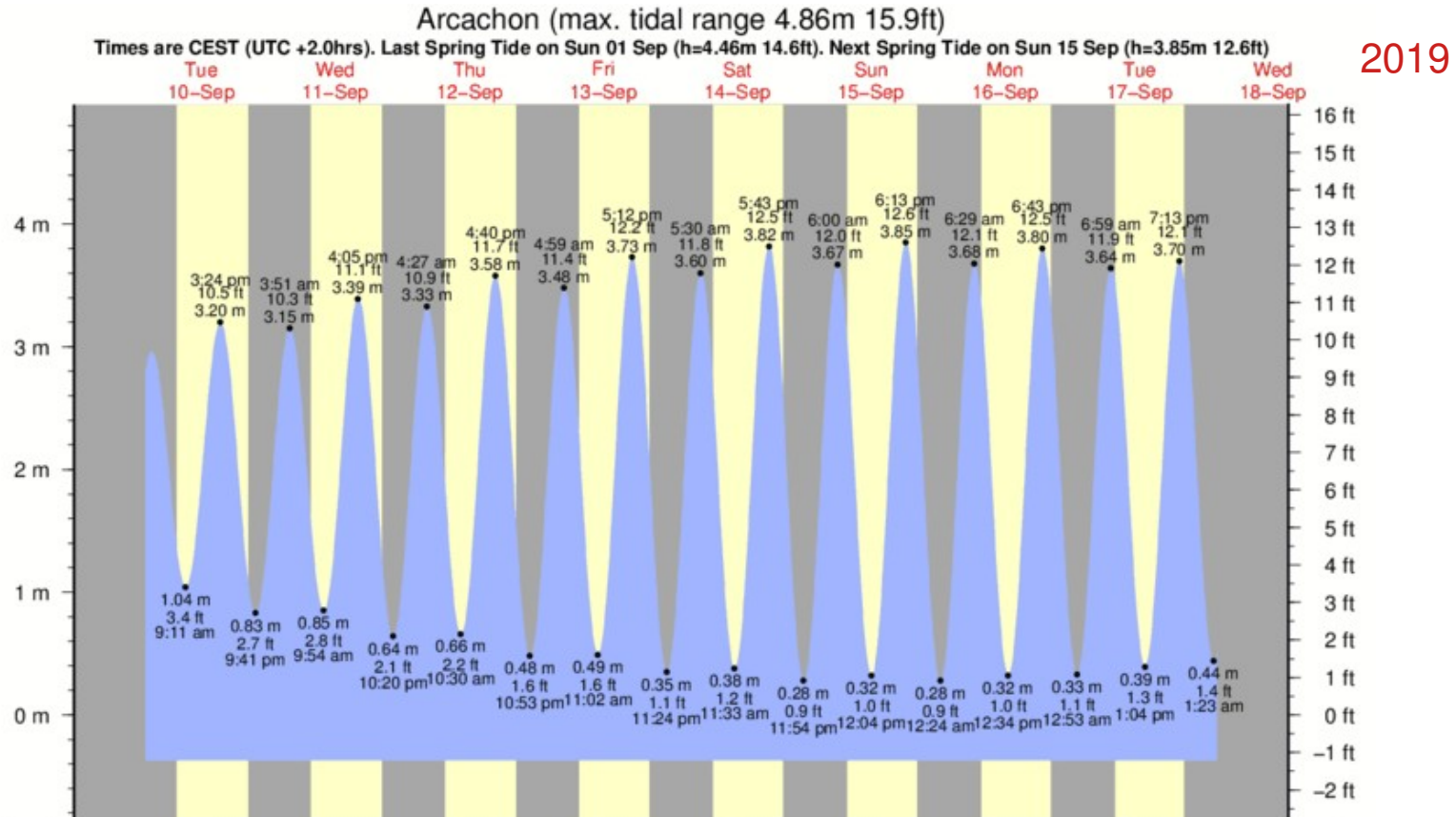
# Examples of time series

?

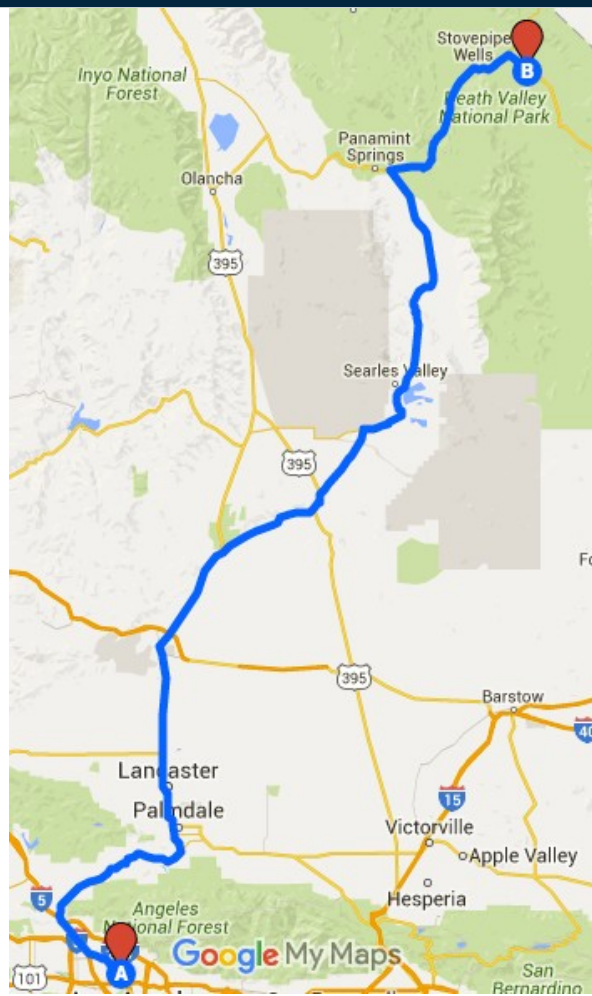




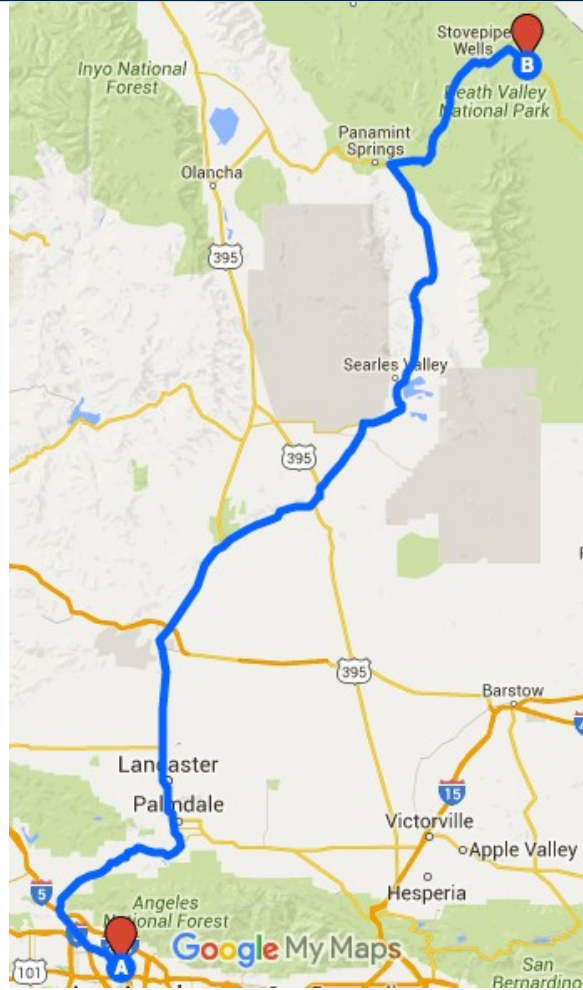
# Examples of time series: ocean tides



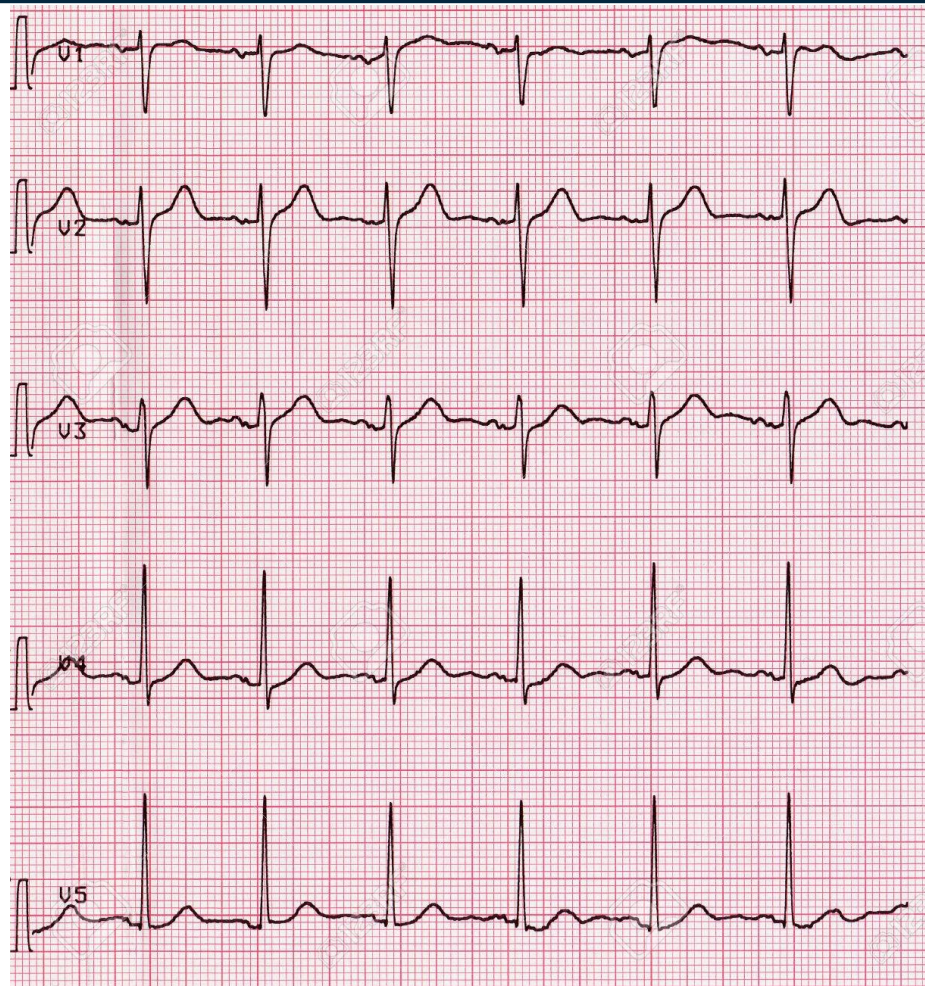
?



# Examples of time series: location profiles

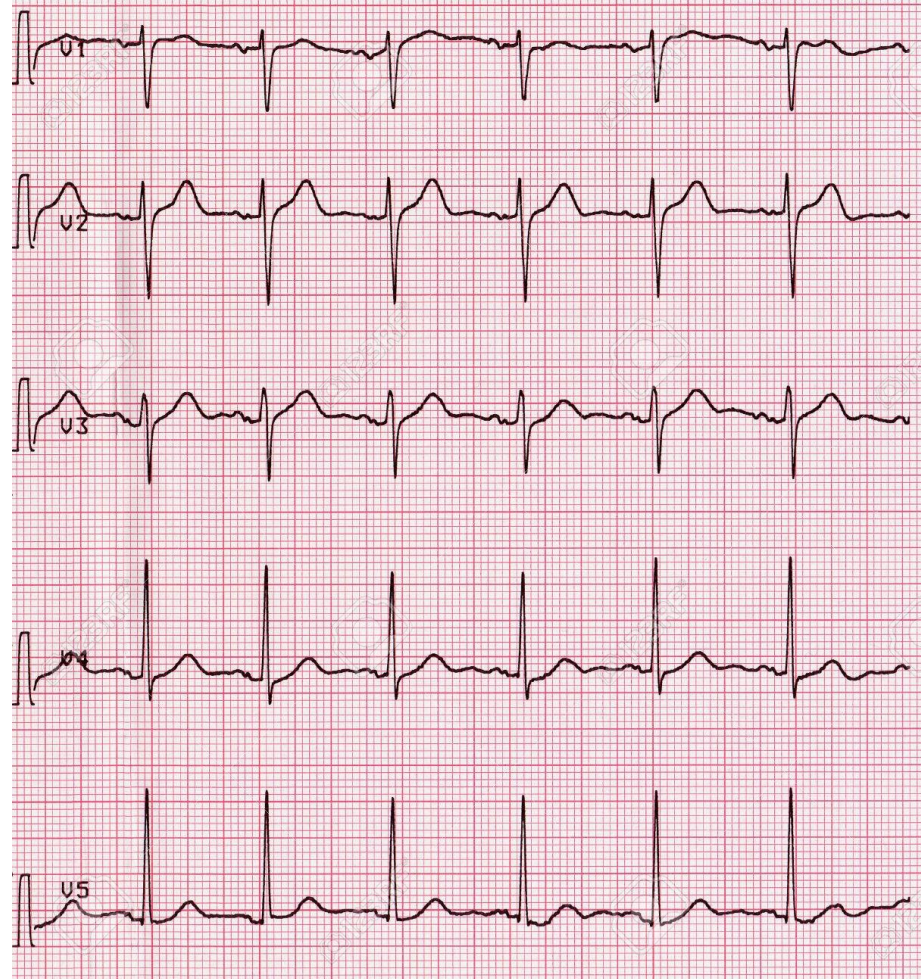


?



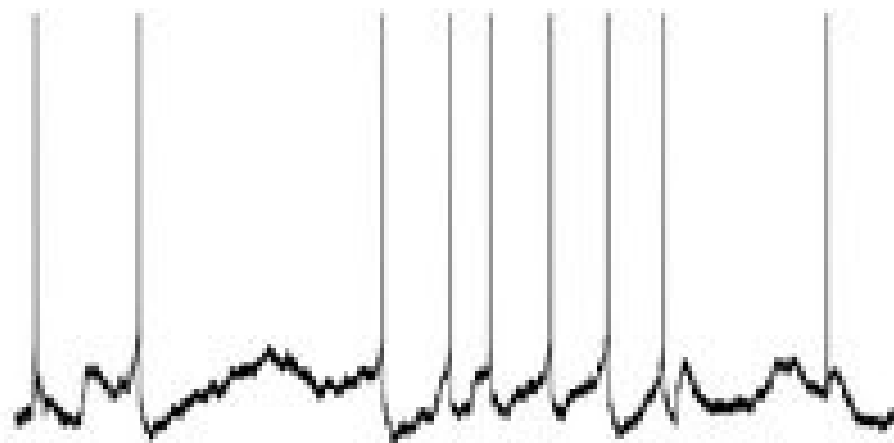


# Time series in biology: electrocardiogram

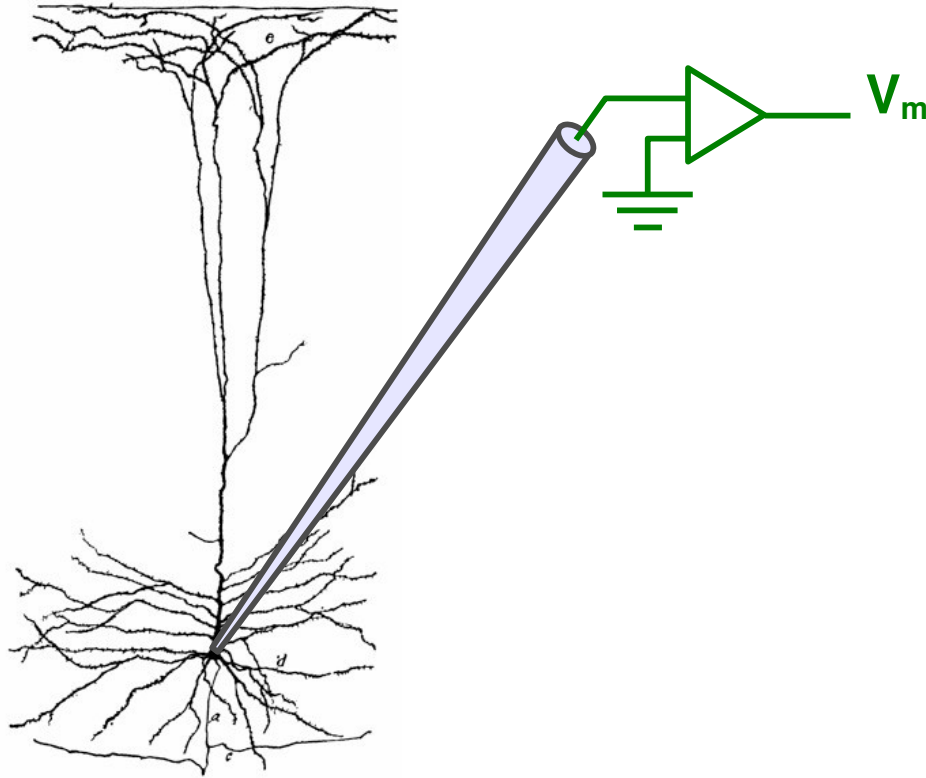


?

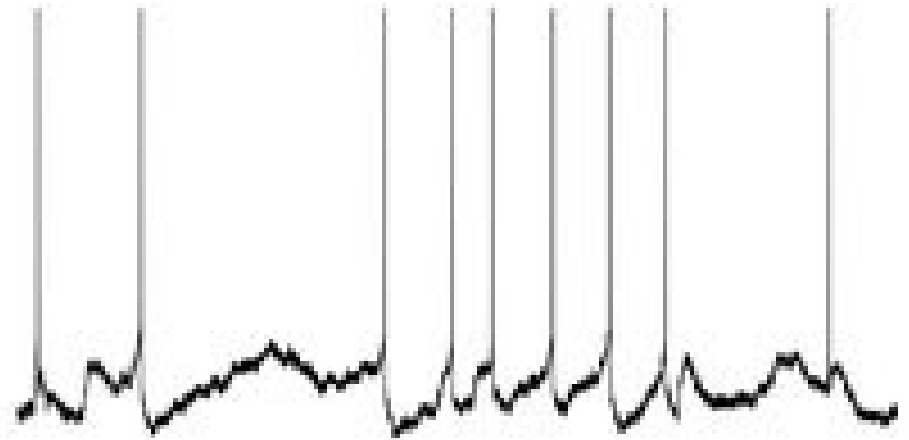
10 mV  
100 ms



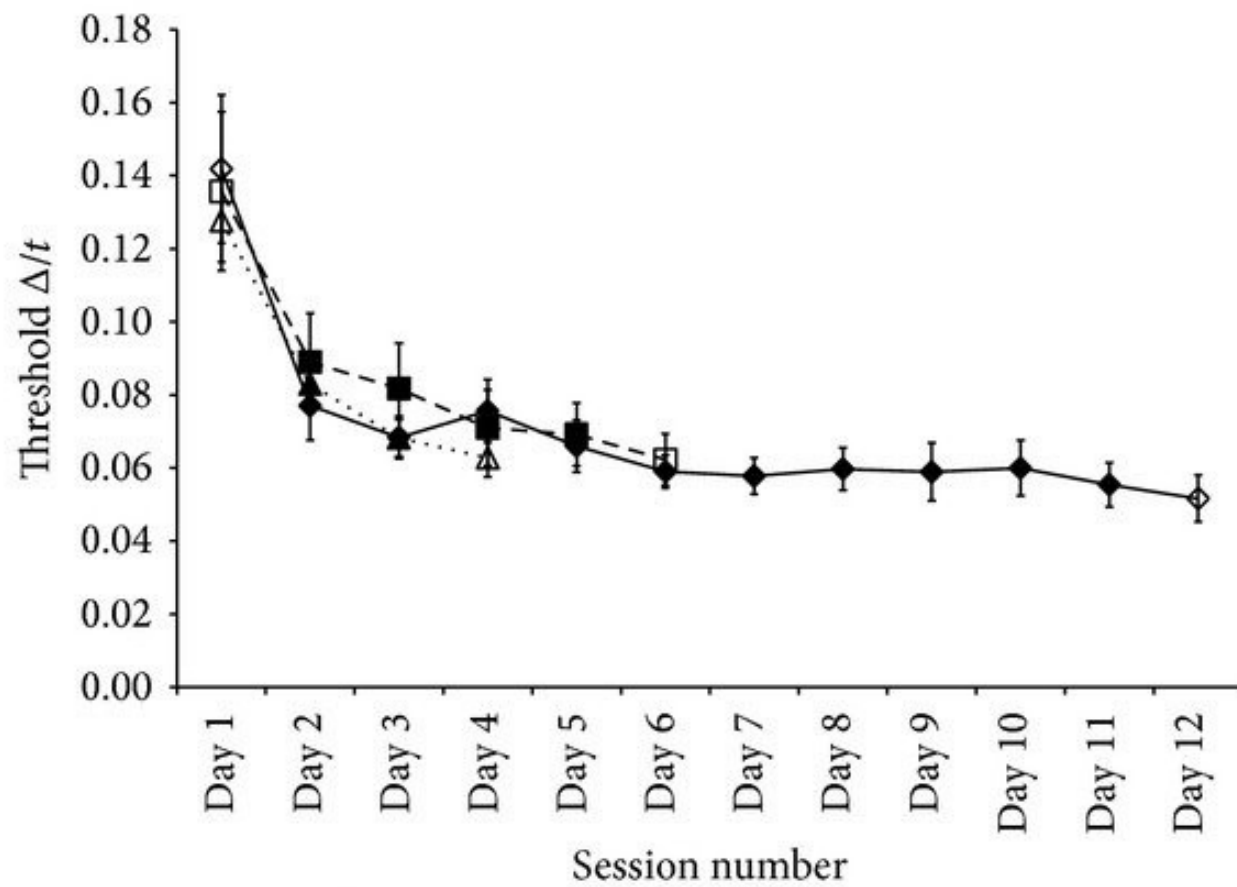
# Time series in neuroscience: membrane potential



10 mV  
100 ms

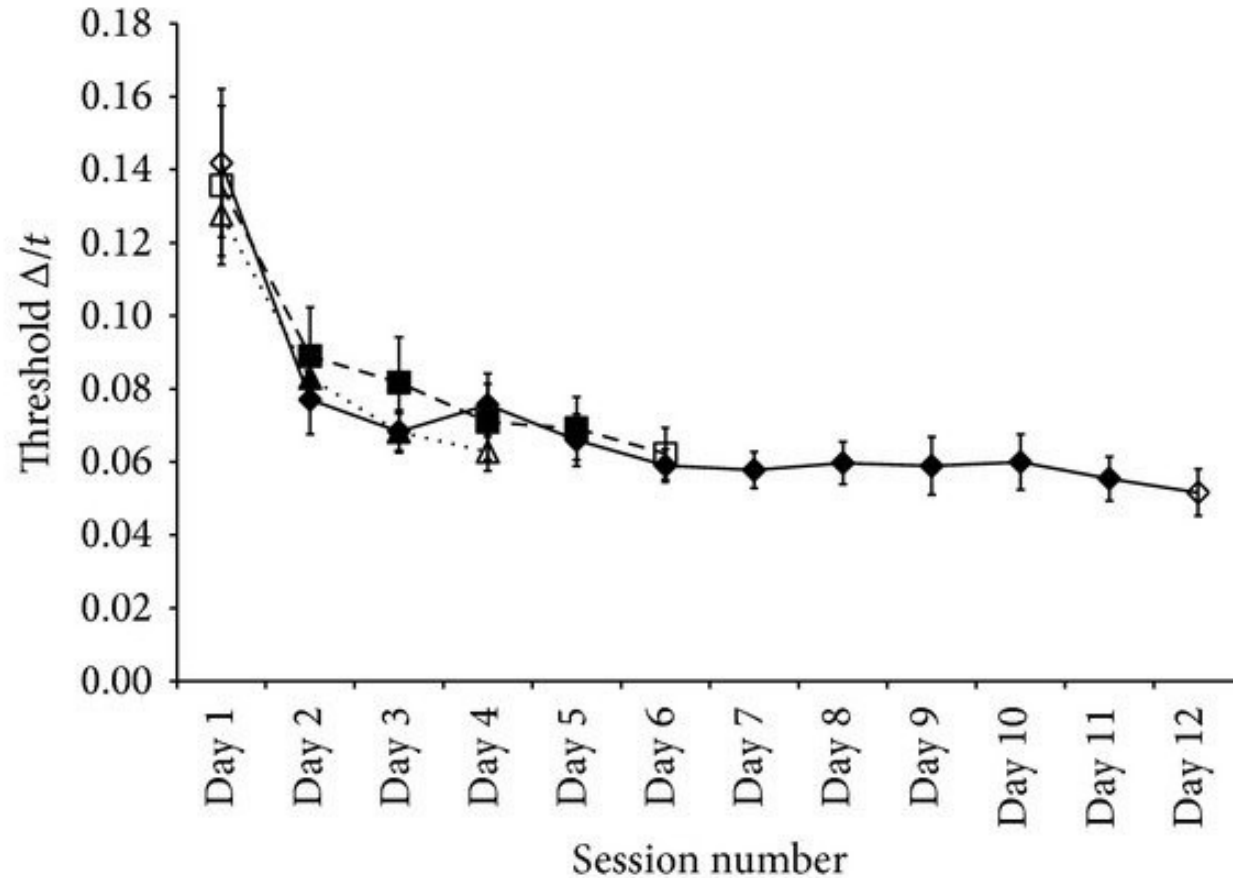


?

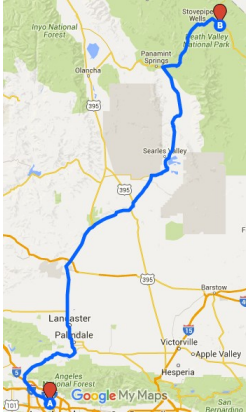




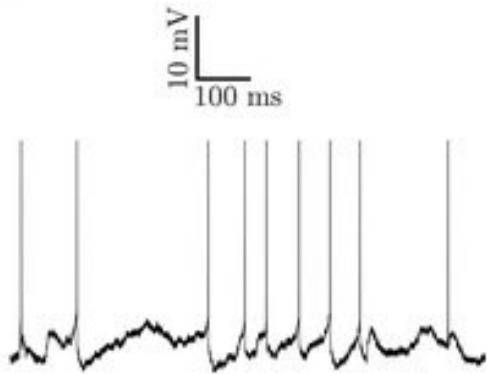
# Time series in neuroscience: learning curves



# Time series: definition



Anything that is observed or measured at many points in time forms a time series.



# Time series: properties

time	value (°)
Jan	3.3
Feb	4.2
March	7.8
April	10.8
May	14.3
June	17.5
July	19.4

] e.g. interval  
1 month

⋮

- list of pairs : time point and data point of specific unit
- listed in time order (ascending time)
- entries are separated by specific intervals (years, months, seconds, ...)

# Interval between data-points

equally spaced points in time

- interval determines frequency of measurement as  $1/\text{interval}$

time (month)	value
Jan	3.3
Feb	4.2
March	7.8
April	10.8
May	14.3
June	17.5
July	19.4

⋮

all intervals  
= 1 month

irregular spaced points in time

time (hh:m)	value
13:20	45.4
13:22	40.1
13:30	38.3
13:35	37.4
13:43	36.1
14:01	35.9
14:08	36.0

⋮

2 min  
8 min  
5 min  
8 min  
18 min  
7 min

# Data can refer to time point, interval, elapsed time

## Apple's shares in 2018

Share price in US Dollars

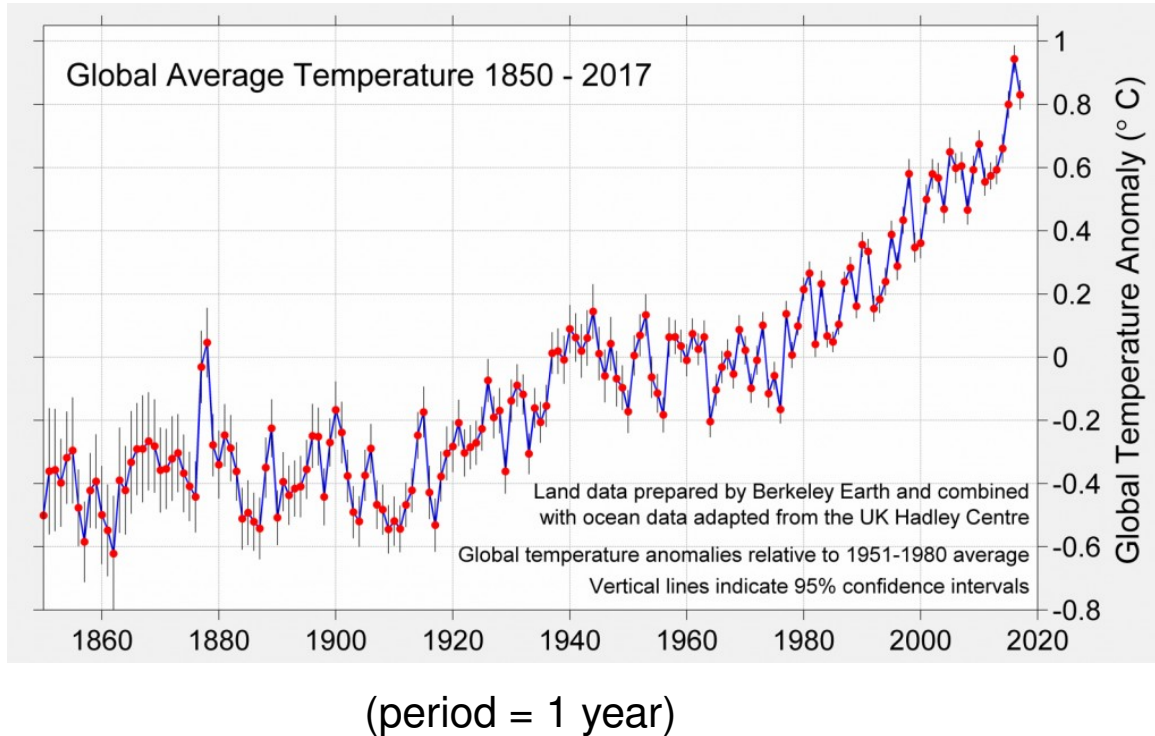


Source: Bloomberg. Last update: 21/11/2018, 8:00am GMT



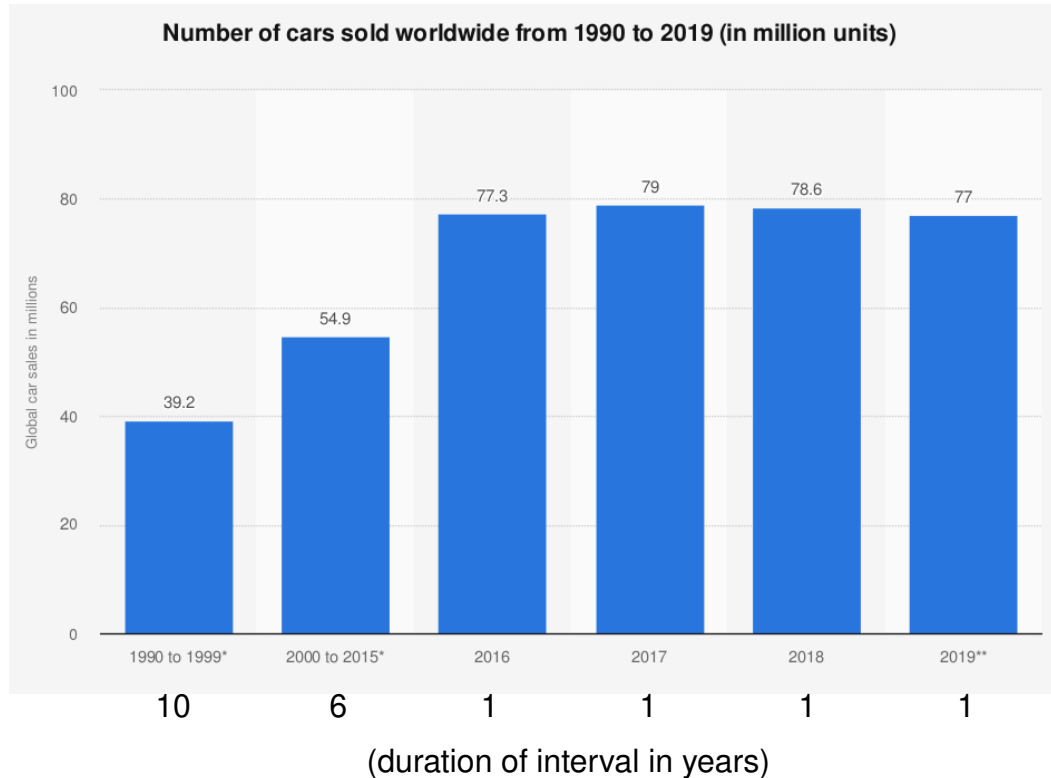
- timestamps – specific instants in time (e.g. every day at 4pm)
- fixed periods : e.g. a month, a year (data represents often average during that period; can be given with further statistics, e.g. standard deviation)
- intervals : indicated start and end of timestamp (general case of fixed periods)
- elapsed time relative to particular start time (often the case for neural data; start is the beginning of a recording)

# Data can refer to time point, interval, elapsed time



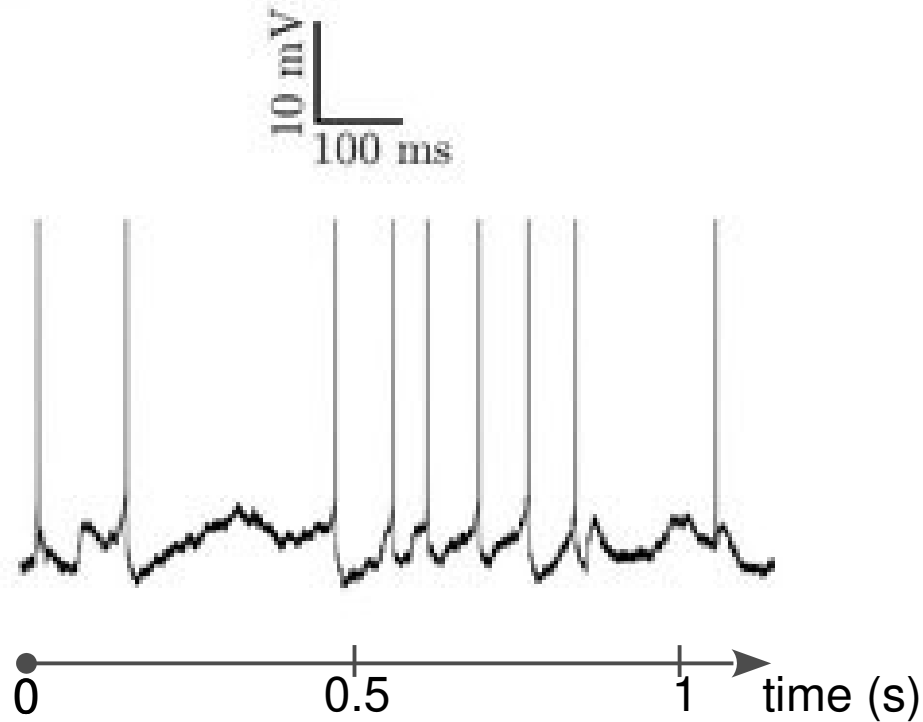
- timestamps – specific instants in time (e.g. every day at 4pm)
- fixed periods : e.g. a month, a year (data represents often sum or average during that period; can be given with further statistics, e.g. standard deviation)
- intervals : indicated start and end of timestamp (general case of fixed periods)
- elapsed time relative to particular start time (often the case for neural data; start is the beginning of a recording)

# Data can refer to time point, interval, elapsed time



- timestamps – specific instants in time (e.g. every day at 4pm)
- fixed periods : e.g. a month, a year (data represents often average during that period; can be given with further statistics, e.g. standard deviation)
- intervals : with indicated start and end of timestamp (general case of fixed periods)
- elapsed time relative to particular start time (often the case for neural data; start is the beginning of a recording)

# Data can refer to time point, interval, elapsed time

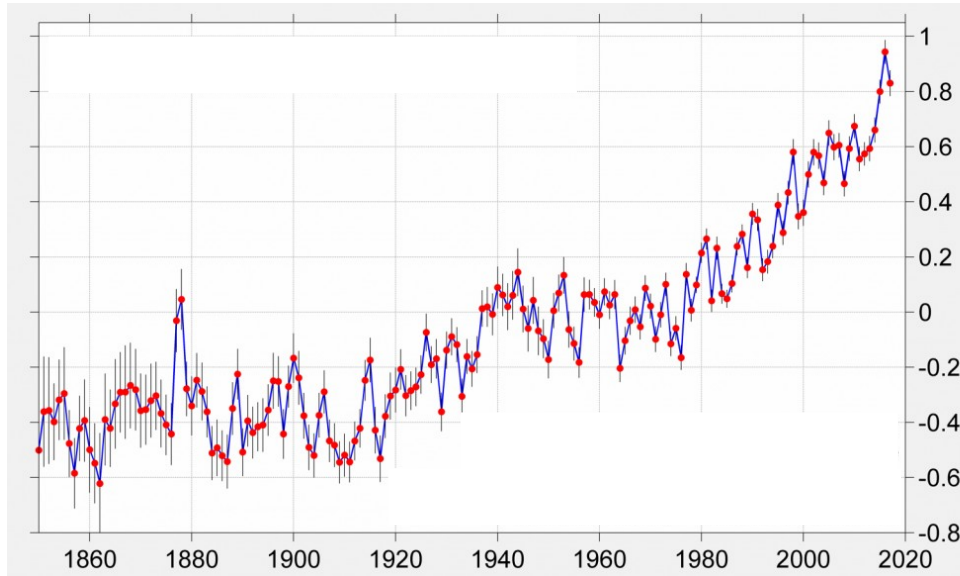


- timestamps – specific instants in time (e.g. every day at 4pm)
- fixed periods : e.g. a month, a year (data represents often average during that period; can be given with further statistics, e.g. standard deviation)
- intervals : indicated start and end of timestamp (general case of fixed periods)
- elapsed time relative to particular start time (often the case for neural data; start is the beginning of a recording)



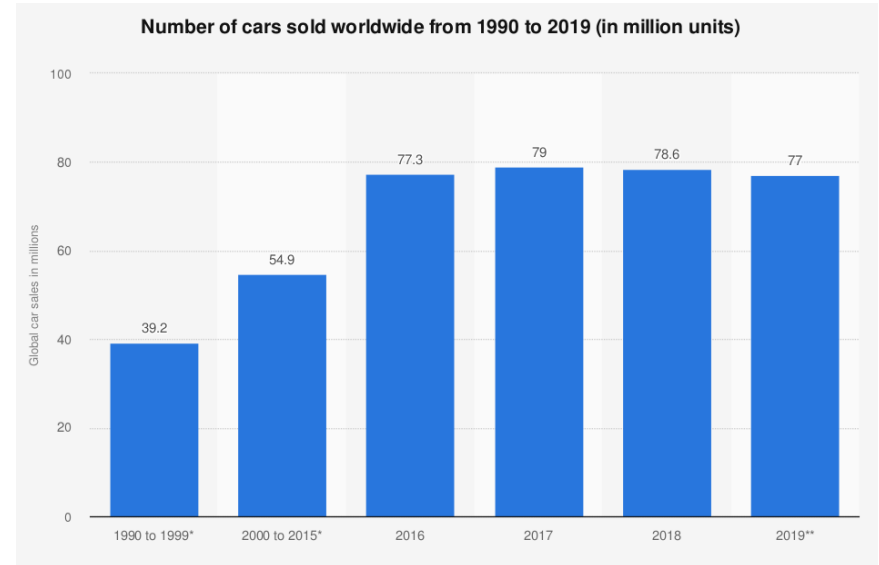
# Common visualizations

line chart



- data-points connected by line
- points themselves can be shown or not

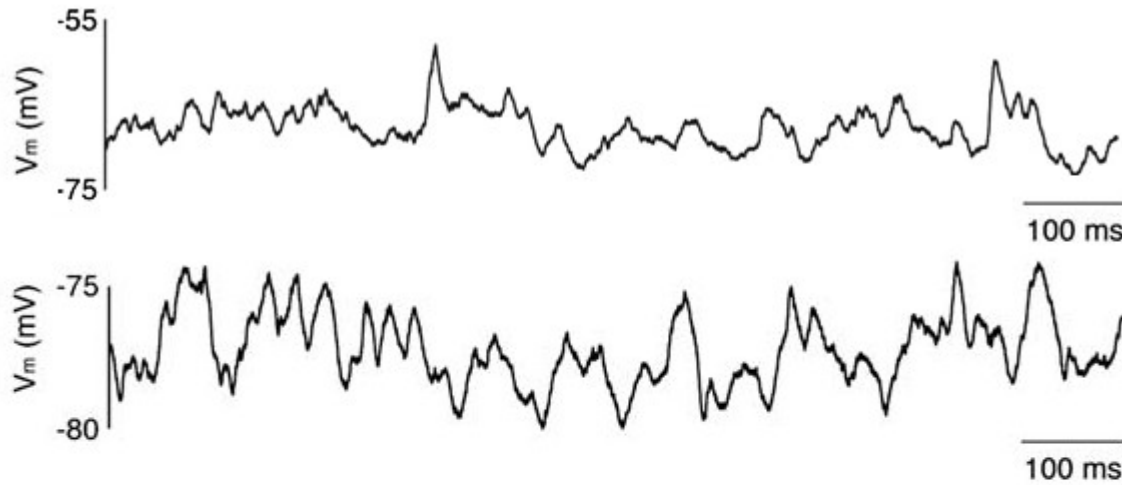
bar graph



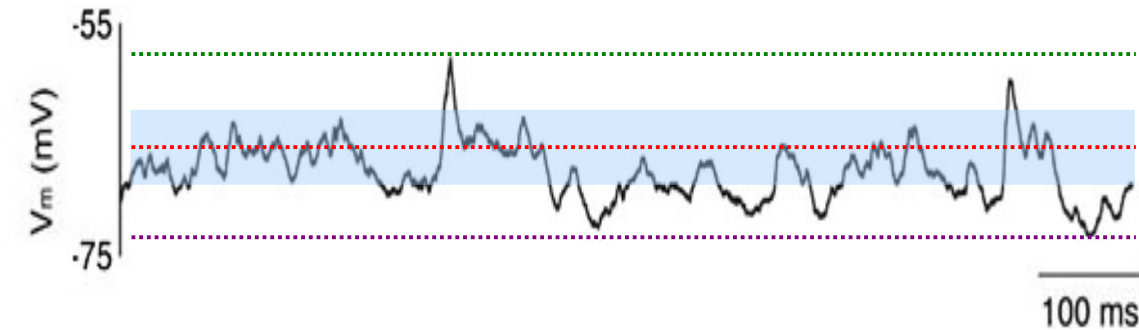
- value shown as height of bar
- limited to displaying few data-points

# Extracting meaningful information from time-series

I want to quantify and compare two membrane potential recordings, which information would be useful?



# Basics statistics I : max, min, mean, SD



- maximum/minimum
- average (arithmetic mean) :
  - sum of all elements divide by total number of elements

$$AM = \frac{1}{n} \sum_{i=1}^n a_i = \frac{a_1 + a_2 + \dots + a_n}{n}$$

- standard deviation (SD):
  - measures variation/dispersion in data-set

$$s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2},$$

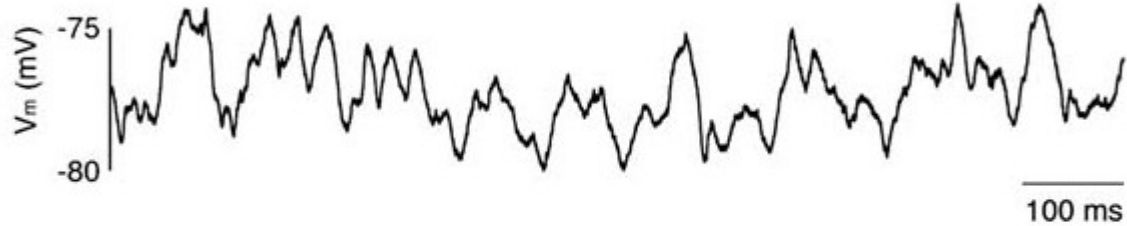
$\bar{x}$  ... mean value

# Comparing data-sets

Recording 1 (R1)



Recording 2 (R2)



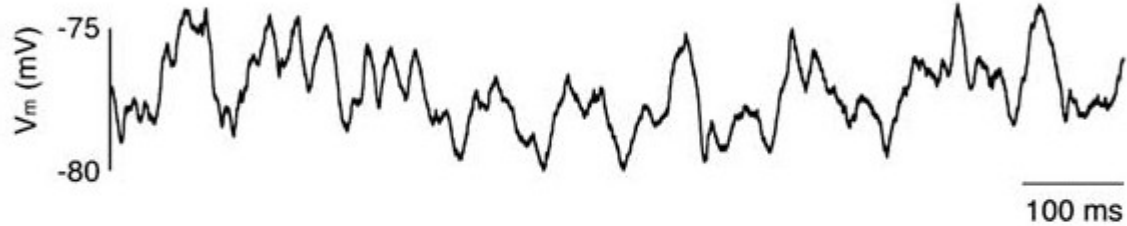
	Comparison
max.	?
min.	
mean	
SD	

# Comparing data-sets

Recording 1 (R1)

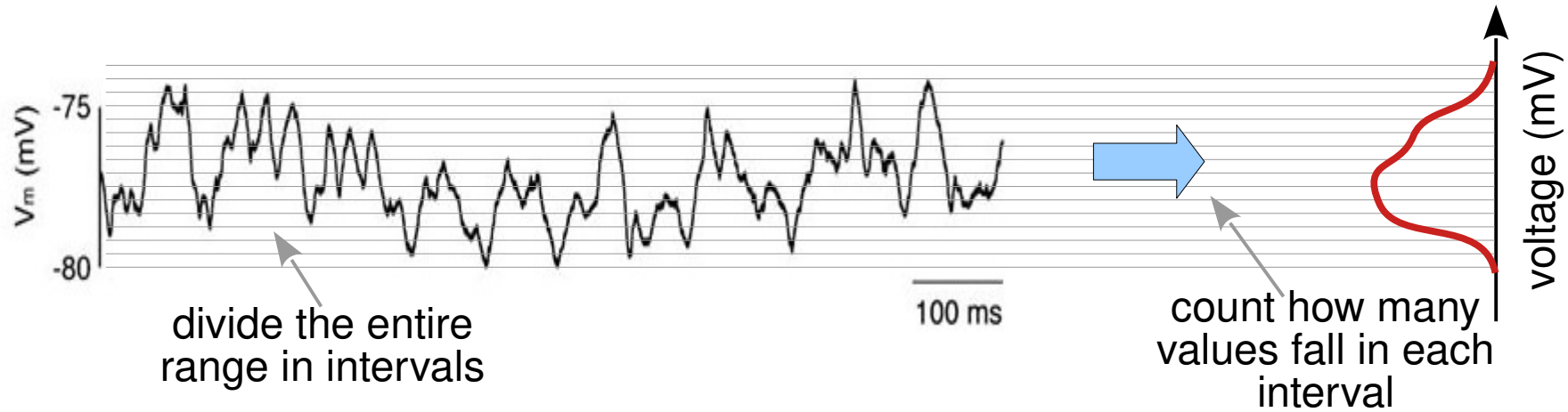


Recording 2 (R2)

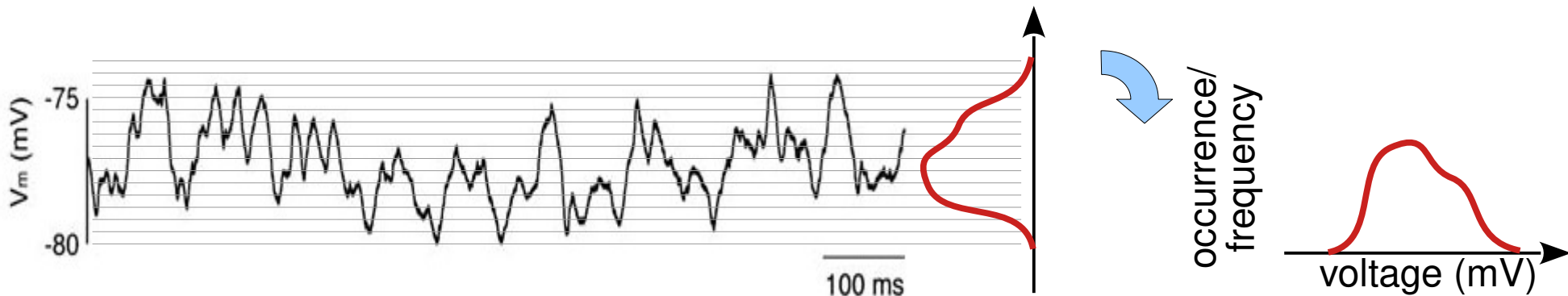


	Comparison
max.	R1>R2
min.	R1>R2
mean	R1>R2
SD	R1<R2

# Histogram – representation of data distribution



# Histogram – shapes

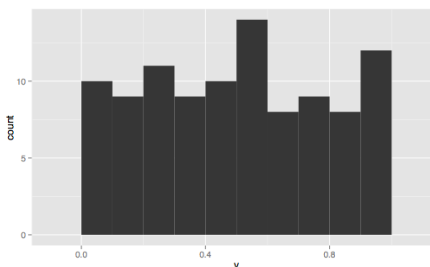
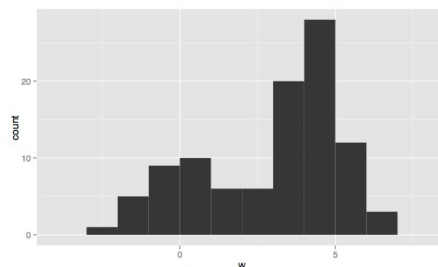
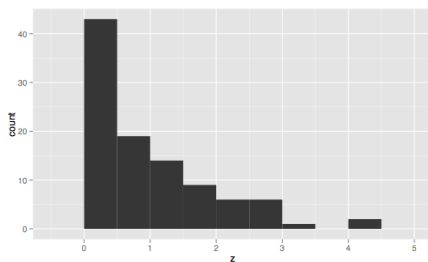
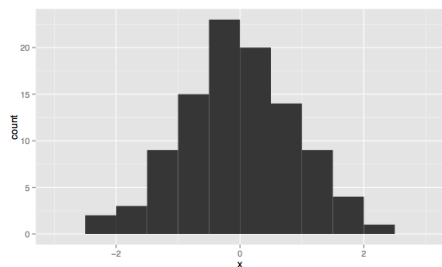


symmetric,  
unimodal

skewed right

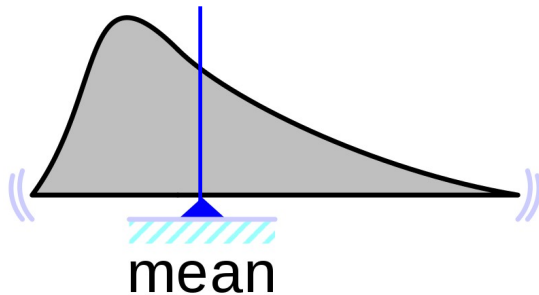
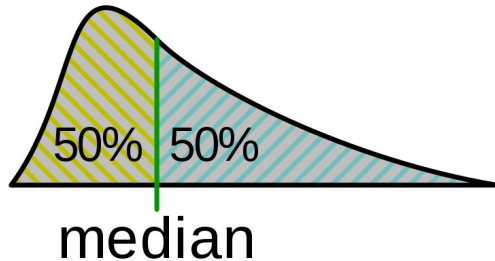
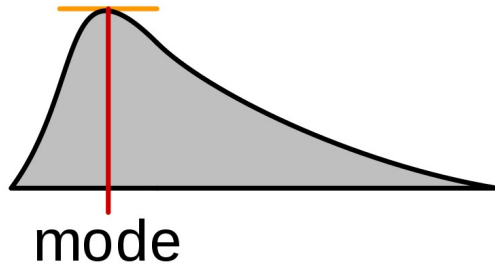
bimodal

symmetric



- good idea to plot histogram with several bin widths to learn more about the data

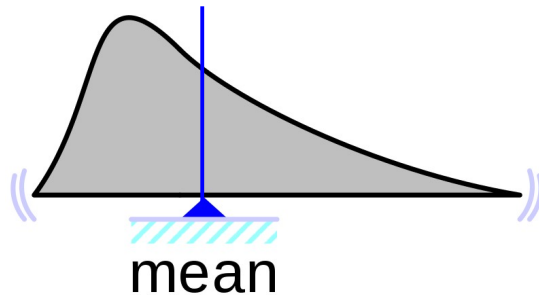
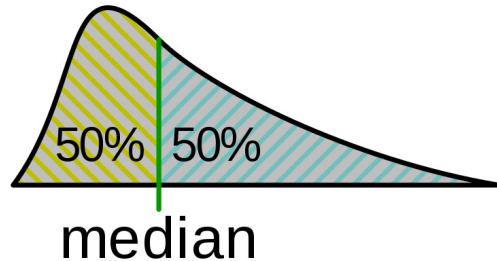
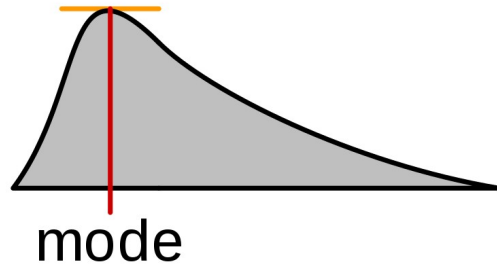
# Basics statistics II : median, percentile



- **mode :**
  - most frequent data point
- **median :**
  - value separating higher half from the lower half of a data-set
  - comparison with mean quantifies skewness of data
- **percentile :**
  - indicating the value below which a given percentile of data-points fall
  - e.g. the median is the 50<sup>th</sup> percentile

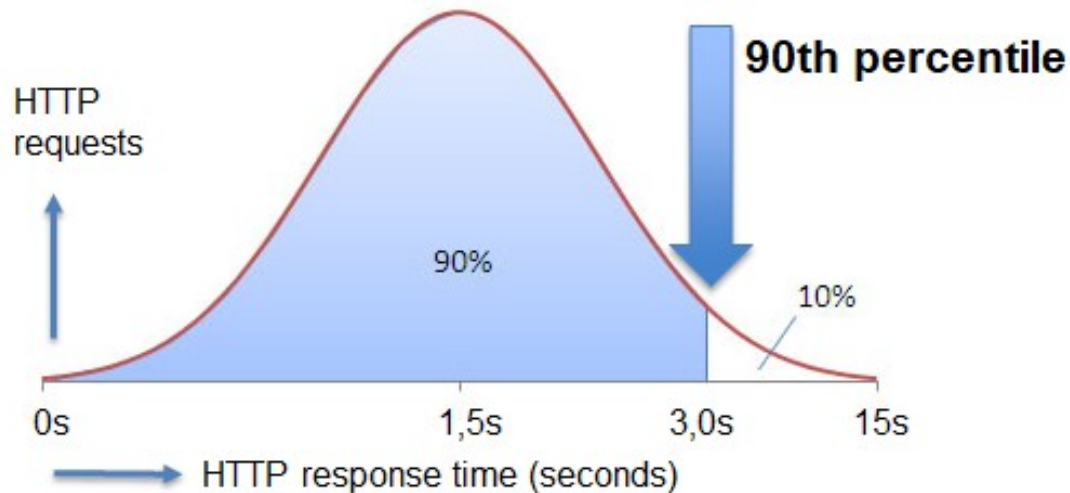


# Basics statistics II : median, percentile



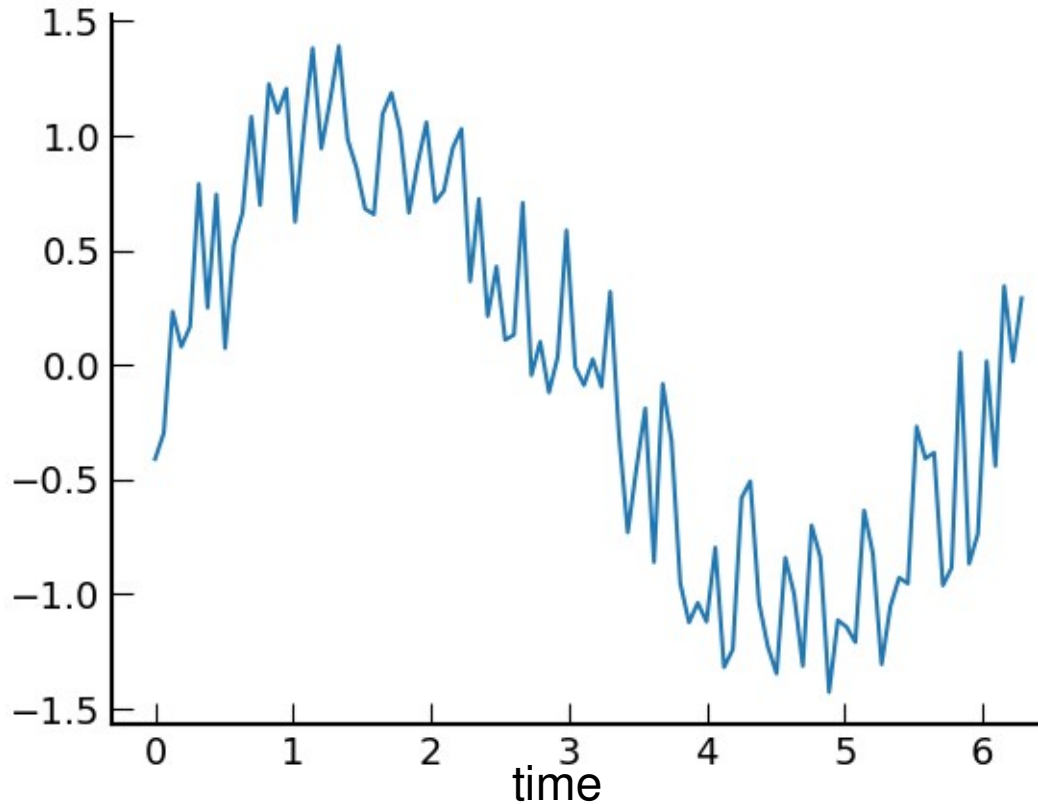
- **mode** :
  - most frequent data point
- **median** :
  - value separating higher half from the lower half of a data-set
  - comparison with mean quantifies skewness of data
- **percentile** :
  - indicating the value below which a given percentile of data-points fall
  - e.g. the median is the 50<sup>th</sup> percentile

# Basics statistics II : median, percentile



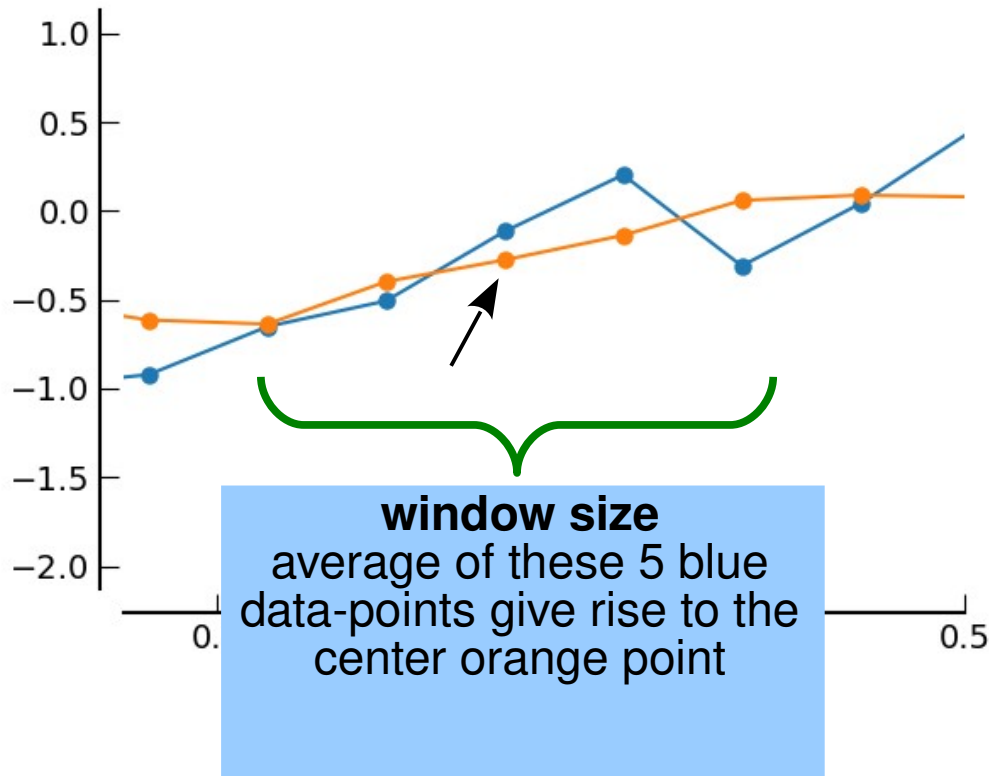
- **mode :**
  - most frequent data point
- **median :**
  - value separating higher half from the lower half of a data-set
  - comparison with mean quantifies skewness of data
- **percentile :**
  - indicating the value below which a given percentile of data-points fall
  - e.g. the median is the 50<sup>th</sup> percentile

# Noisy data



How to reduce noise while preserving characteristics (such as dynamics) of the data ?

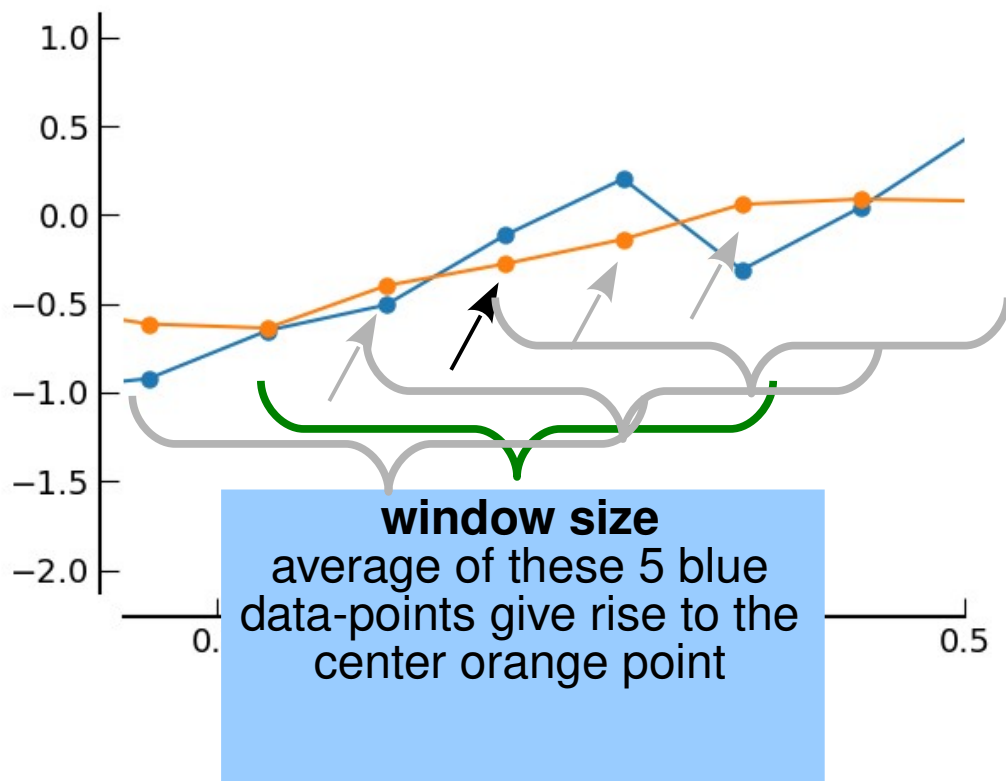
# Moving average



**Moving Average** (or moving mean, rolling mean) : analyze data by creating series of averages of different subsets of the full data

- often used to smooth out short-term fluctuations (an example of low-pass filter)
- mathematically, moving average is a convolution of the data with a flat, normalized kernel

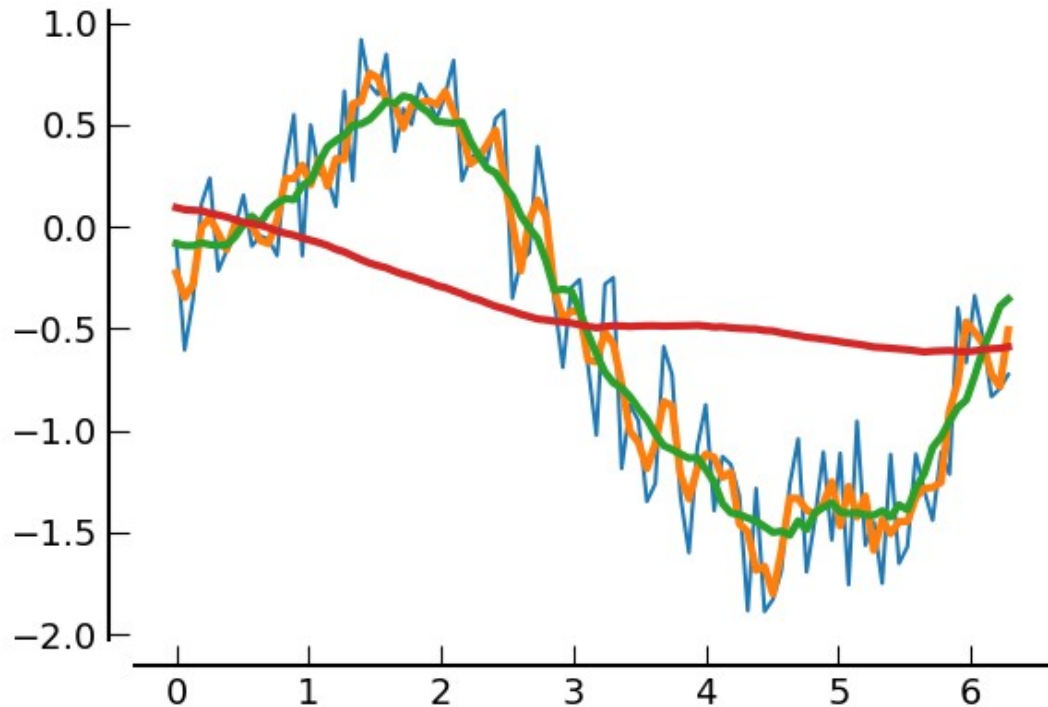
# Moving average



**Moving Average** (or moving mean, rolling mean) : analyze data by creating series of averages of different subsets of the full data

- often used to smooth out short-term fluctuations (an example of low-pass filter)
- mathematically, moving average is a convolution of the data with a flat, normalized kernel

# Moving average: window size



window size = 3  
window size = 11  
window size = 99

## Window size

Pick the smallest window size where the signal starts to flatten out, without affecting dynamics of interest.