Review, 4-21

- How to turn an unnormalized posterior into a normalized posterior
- What is Bayesian Inference?
- Typical definition of a posterior
- Predictive Distribution

Bayesian Vs. Frequentist

Frequentist

- Limiting relative frequencies => probability is an observed property
- Parameters fixed and unknown => no need for probability of parameter
- Procedures for long-run frequencies (e.g. 95% CI)

Bayesian Vs. Frequentist

Bayesian

- Probability is degree of belief
  => can derive probability of many things
- Can estimate probability of parameters
- Can draw inferences about parameter
  probability distribution, point estimates, intervals

Frequentist

- Limiting relative frequencies => probability is an observed property
- Parameters fixed and unknown => no need for probability of parameter
- Procedures for long-run frequencies (e.g. 95% CI)

Bayesian Vs. Frequentist

Pro Bayes:

- Estimating distributions => uncertainty built in
- No need to choose model; always “admissible”
- Automatic regularization

Con:

- Need to assume prior (even if nothing can obviously work)
- Approximate solutions: tend to be a little less accurate for simple classification / regression problems

Bayesian Vs. Frequentist

Pro Bayes:

- Estimating distributions => uncertainty built in
- No need to choose model; always “admissible”
- Automatic regularization

There is at least one situation where

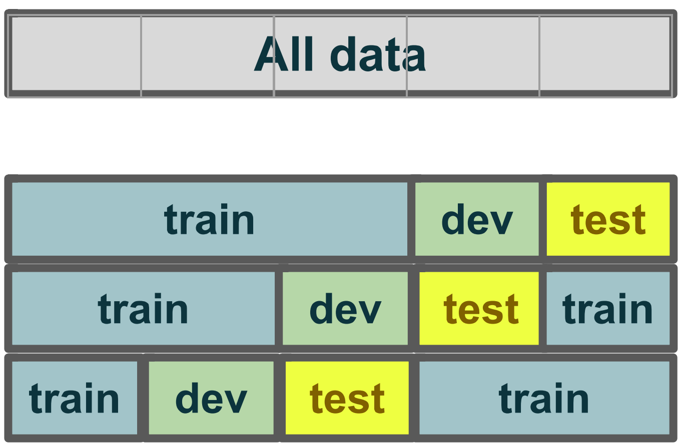
Con:

- Need to assume prior (even if nothing can obviously work)
- Approximate solutions: tend to be a little less accurate for simple classification / regression problems

Revisiting N-Fold Cross-Validation

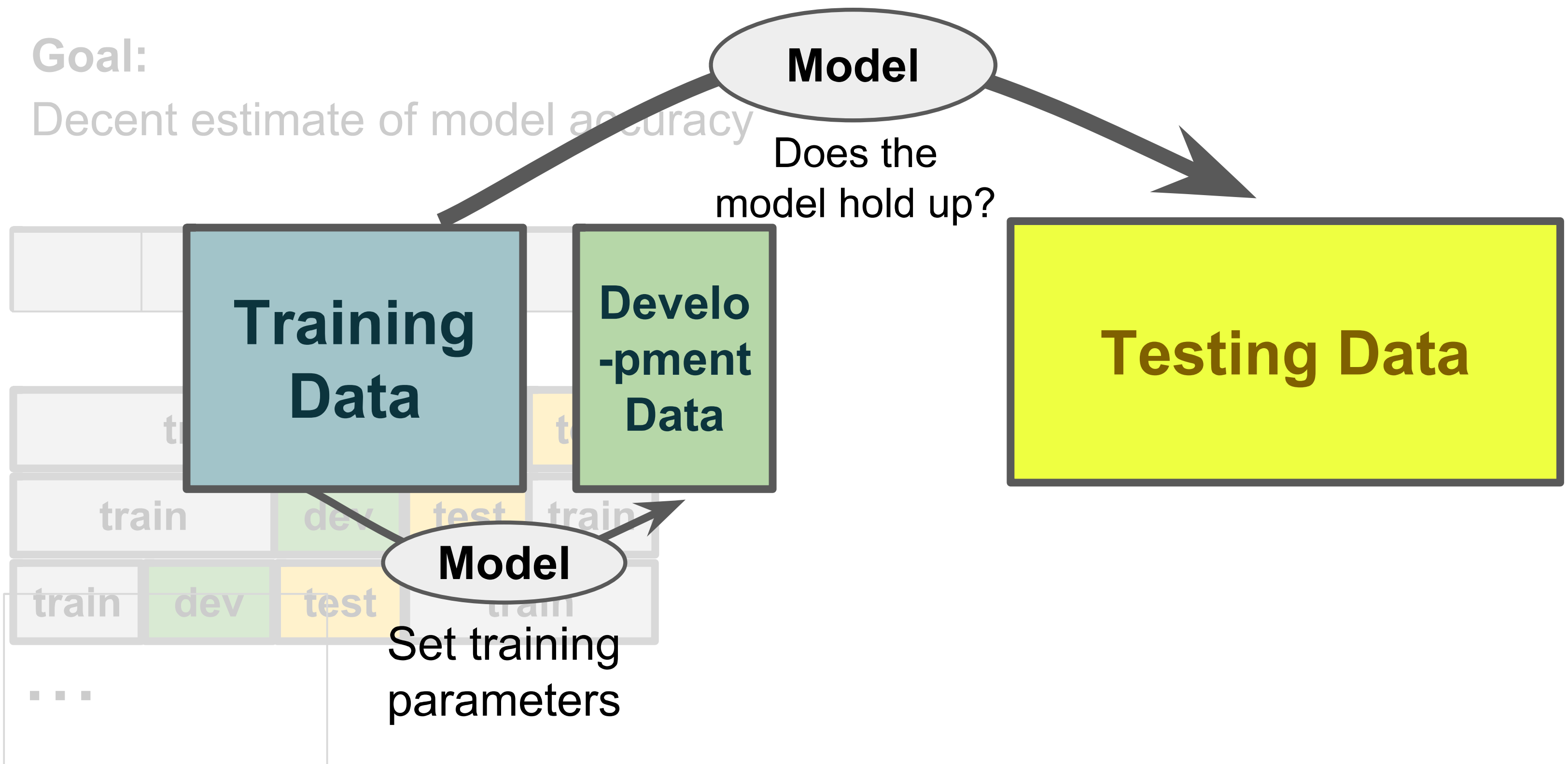
Goal:

Decent estimate of model accuracy



...

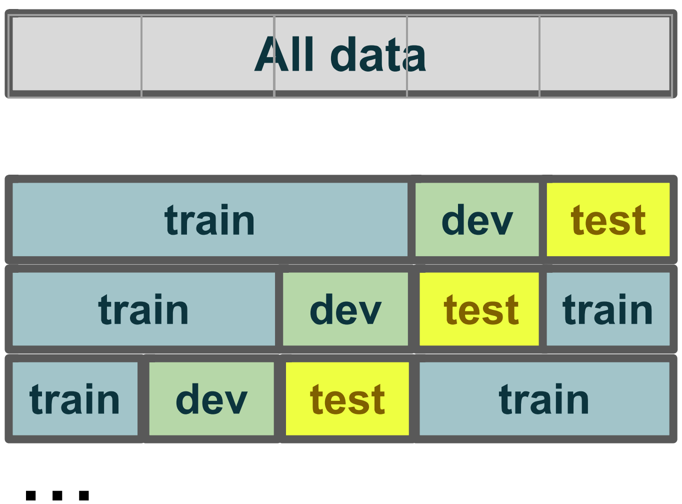
Revisiting N-Fold Cross-Validation



Revisiting N-Fold Cross-Validation

Goal:

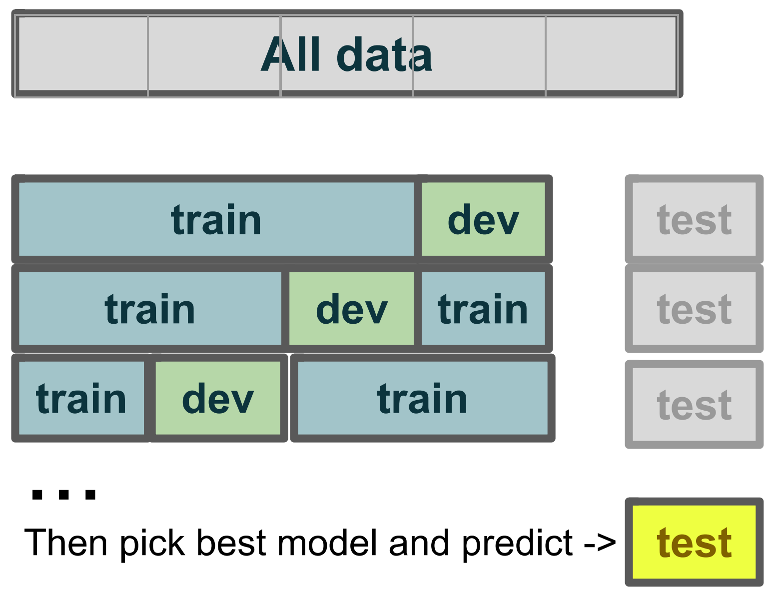
Decent estimate of model accuracy



Revisiting N-Fold Cross-Validation

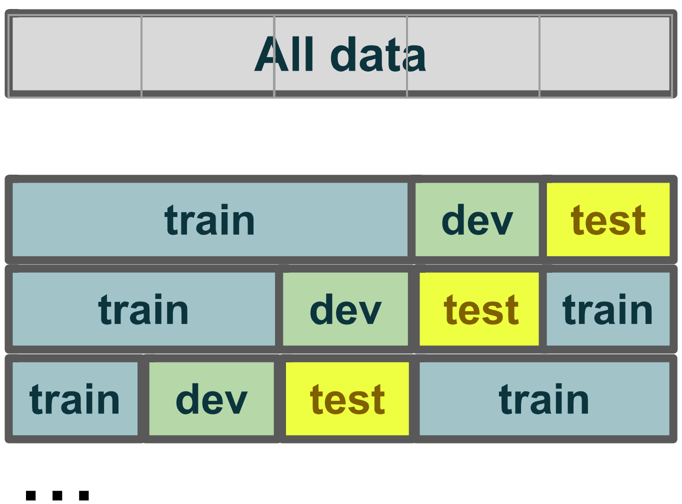
Goal:

Select a super-reliable penalty (alpha)
(this is overkill)



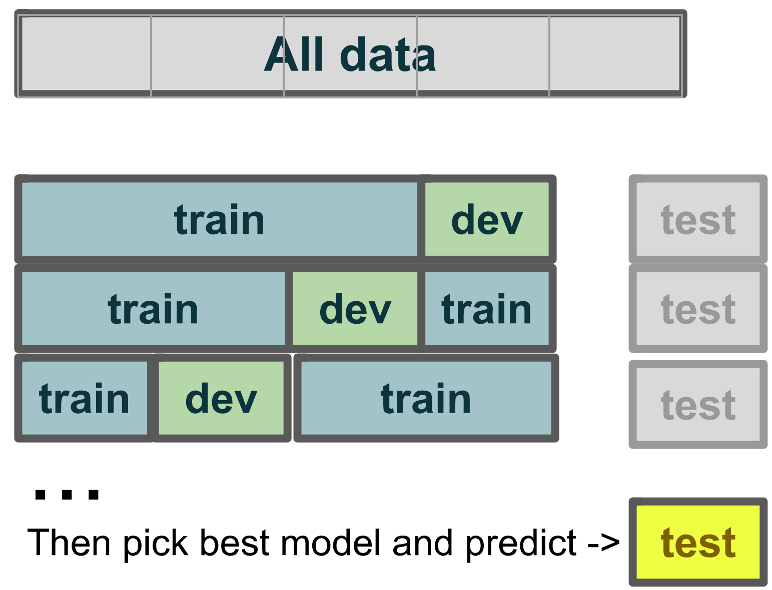
Revisiting N-Fold Cross-Validation

Goal:
Decent estimate of model accuracy



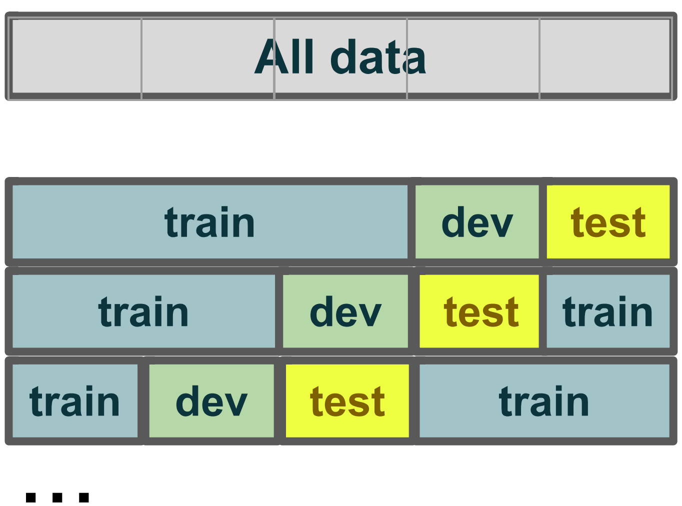
≠

Goal:
Select a super-reliable penalty (alpha)
(this is overkill)

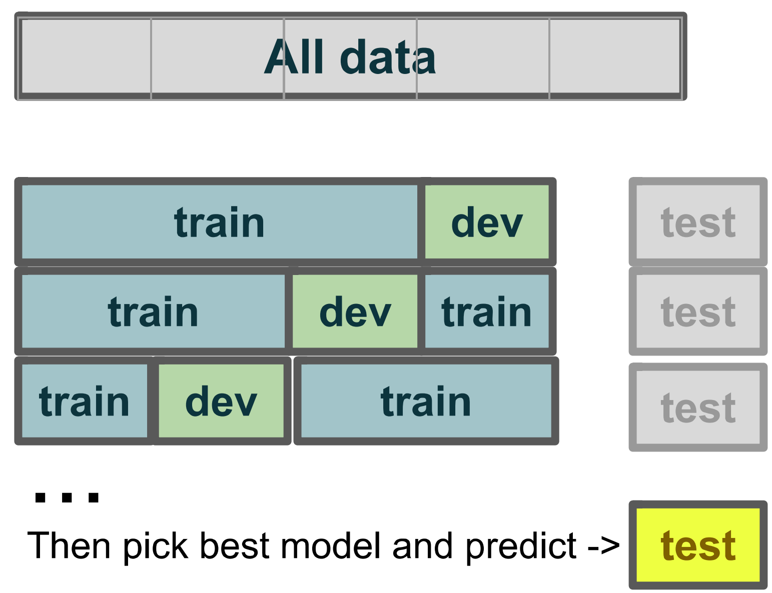


Revisiting N-Fold Cross-Validation

Goal:
Decent estimate of model accuracy



Goal:
Select a super-reliable penalty (alpha)
(this is overkill)



Example: Assignment 3

Introduction Time Series Analysis

Goal: Understanding temporal patterns of data (or real world events)

Common tasks:

- **Trend Analysis:** Extrapolate patterns over time (typically descriptive).
- **Forecasting:** Predicting a future event (predictive).
  (contrasts with “cross-sectional” prediction -- predicting a different group)

Introduction to Causal Inference (Revisited)

X causes Y as opposed to X is associated with Y

Changing X will change the distribution of Y.

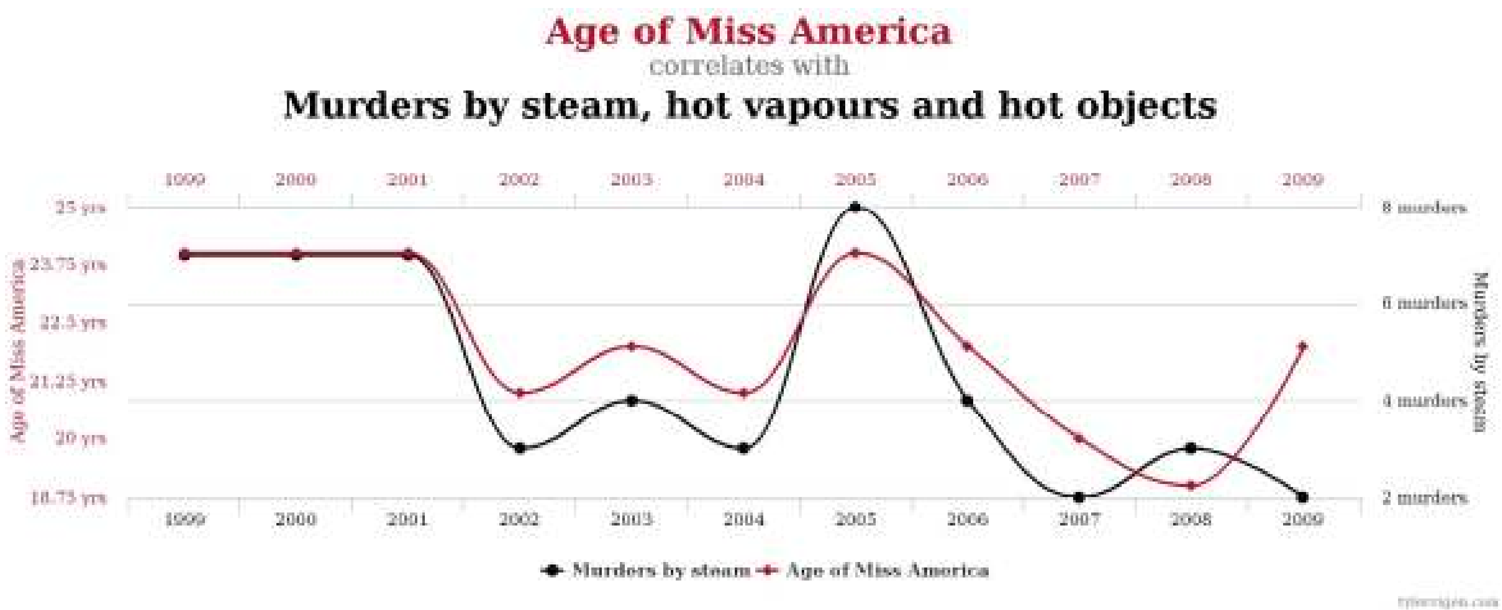
X causes Y $\not\leftrightarrow$ Y causes X

Spurious Correlations

Extremely common in time-series analysis.

Spurious Correlations

Extremely common in time-series analysis.



Introduction to Causal Inference (Revisited)

X causes Y as opposed to X is associated with Y

Changing X will change the distribution of Y.

X causes Y $\not\leftrightarrow$ Y causes X

$$P(Y = 1|X = 1) - P(Y = 1|X = 0)$$

Counterfactual Model: Exposed or Not Exposed: X = 1 or 0

$$Y = \begin{cases} C_0 & \text{if } X = 0 \\ C_1 & \text{if } X = 1 \end{cases}$$

Causal Odds Ratio: $$\frac{\left(\frac{P(C_1=1)}{P(C_1=0)}\right)}{\left(\frac{P(C_0=1)}{P(C_0=0)}\right)}$$

Simpson's "Paradox"

	Y=1	Y=0	Y=1	Y=0
X=1	.15	.225	.1	.025
X=0	.0375	.0875	.2625	.1125
	Z = men		Z = women	

http://vudlab.com/simpsons/

Autocorrelation

“(a.k.a. Serial correlation).”

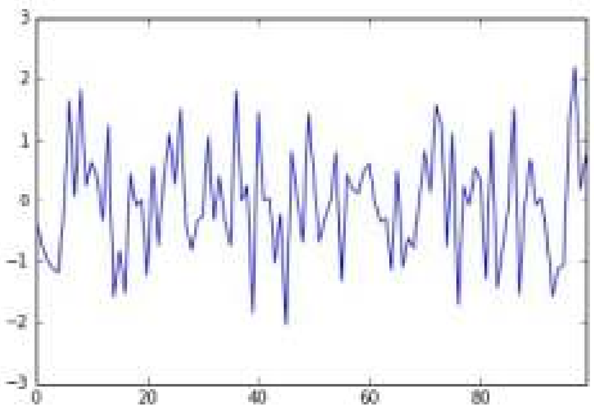
Quantifying the strength of a temporal pattern in serial data.

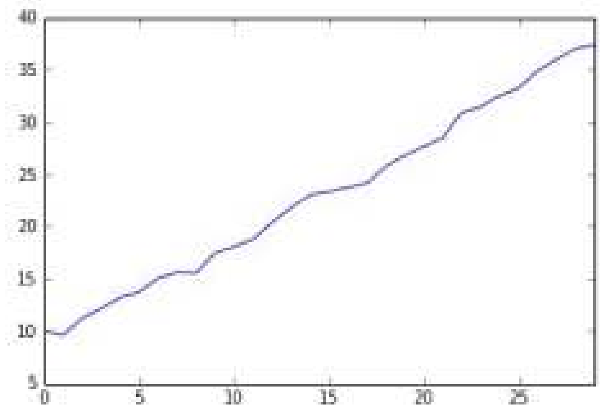
Requirements:

- Assume regular measurement (hourly, daily, monthly...etc..)

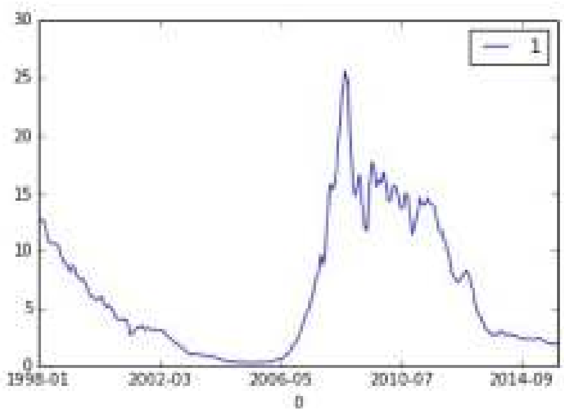
Autocorrelation

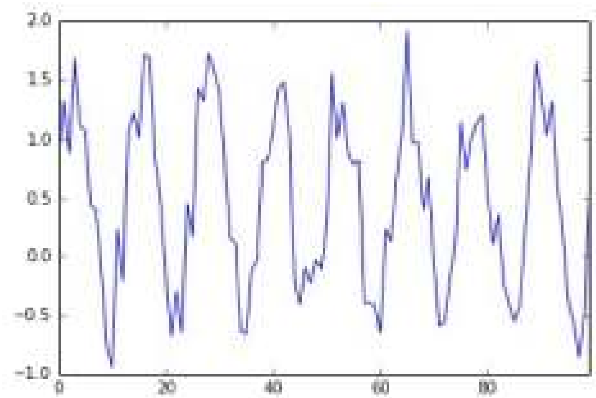
Quantifying the strength of a **temporal pattern** in serial data.





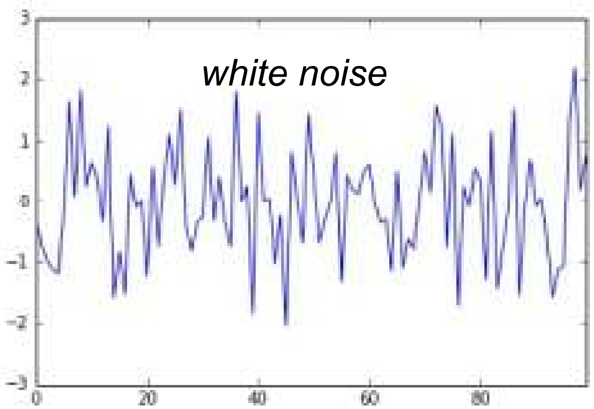
Which have temporal patterns?

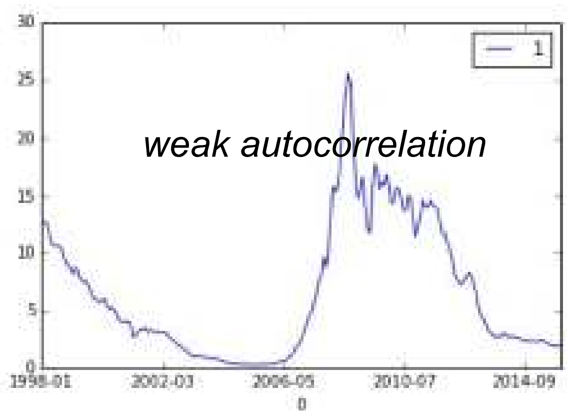




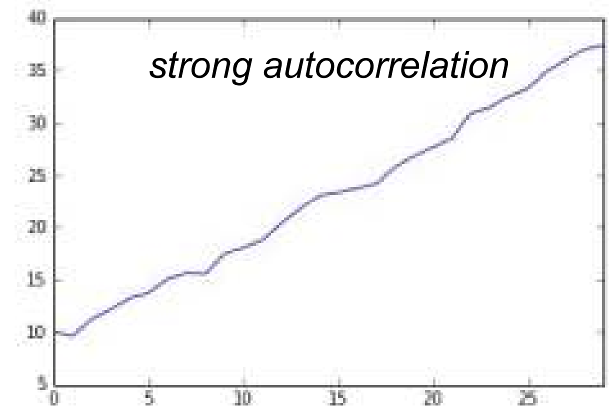
Autocorrelation

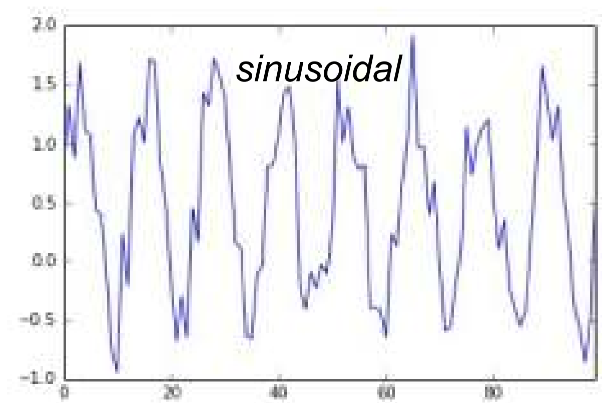
Quantifying the strength of a **temporal pattern** in serial data.





Which have temporal patterns?





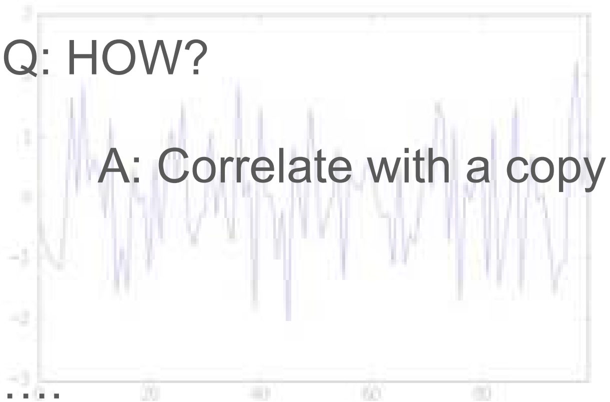
Autocorrelation

Quantifying the strength of a **temporal pattern** in serial data.

Q: HOW?

Autocorrelation

Quantifying the strength of a **temporal pattern** in serial data.



Q: HOW?

A: Correlate with a copy of self, shifted slightly.

....

Autocorrelation

Quantifying the strength of a **temporal pattern** in serial data.

Q: HOW?

A: Correlate with a copy of self, shifted slightly.

```
Y = [3, 4, 4, 5, 6, 7, 7, 8]

correlate(Y[0:7], Y[1:8])  #lag=1

correlate(Y[0:-2], Y[2:8])  #lag=2

....
```

Autocorrelation

Quantifying the strength of a **temporal pattern** in serial data.

Q: HOW?

A: Correlate with a copy of self, shifted slightly.

```
Y = [3, 4, 4, 5, 6, 7, 7, 8]

correlate(Y[0:7], Y[1:8])  #lag=1

correlate(Y[0:-2], Y[2:8])  #lag=2

....
```