Introduction to Time Series

Paul Sangrey

University of Pennsylvania
paul@sangrey.io

August 13, 2018

What is Time Series?

- Statistics starts considering independently, identically distributed data.
- We have some underlying true distribution and bunch of observations from it — X.
- We can easily estimate the population (true) mean μ using the sample mean $\bar{x}_t := \frac{1}{T}\sum_{t=1}^{T} x_t$
- **Sometimes order matters.**
  - Did Trump copy the term “fake news” from the mainstream media, or vice-versa?
  - Does it make sense to think of William Randolph Hearst, Walter Cronkite, and Stephen Colbert as doing the same thing.
    - We cannot take means if the data are not from the same distribution.

Why Time Series for Communications?

News is Dynamic!

- Something is only news if it is new.
- We must figure out what's new.
- Order matters.

Examples:

1. What is "Fake News"? Who started talking about it first?
2. Did the 2017 Congressional Baseball Shooting get as much coverage as we expected?
3. What about the Russian bots?
   - Are the bots just the natural consequence of AI and online media or was the 2016 Election special?
   - Did they even matter? If they hadn't posted would someone have taken their place?

Decomposition Time Series

- A time series has 4 components.
  1. Trend
  2. Season
  3. Cycle (Predictable but not persistent components.)
  4. Noise

Plan for the Talk

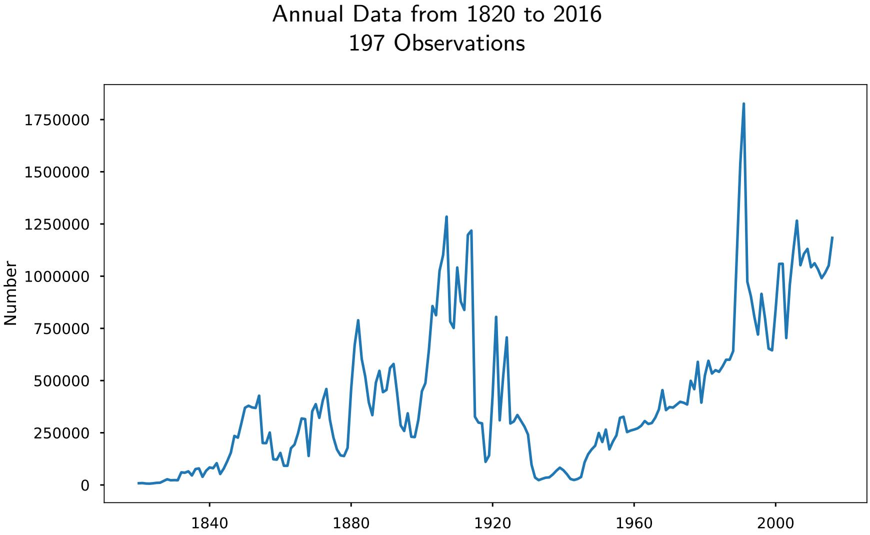
Goal:
Learn how to grasp the dynamics of a process.

(Exploratory Data Analysis for Time Series!)

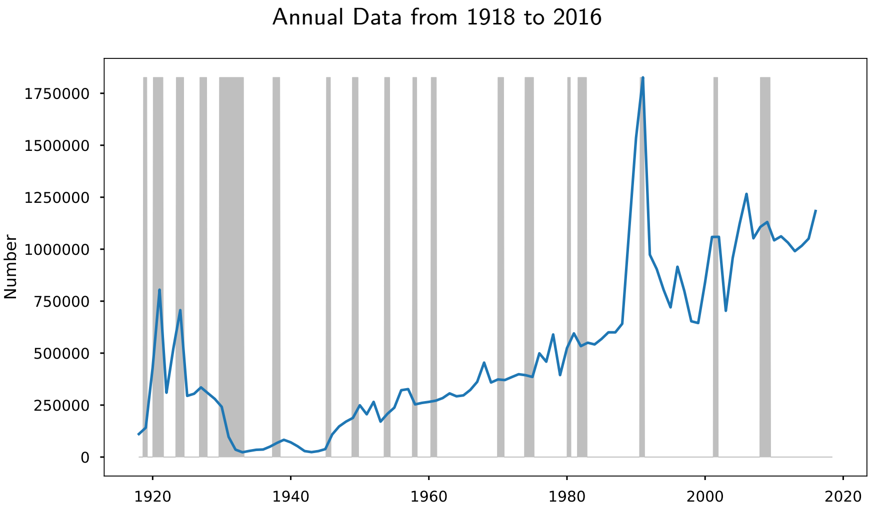
1. Consider Legal Immigration to the U.S.
   - Figure out how to separate the trend, predictable movements (cycle), and noise.
   - Discuss when you can view data from different periods as being from the same distribution and what to do about it if you can't.

2. Baker, Bloom, and Davis (2016) measure people's uncertainty about economic policy changes by aggregating over 12,000 newspaper articles into a single measure.
   - Published in one of the best economics journals and already has almost 2000 citations.
   - This uncertainty has a weakly cycle, (i.e. seasonality), we'll extend the type of analysis to handle this case as well.

U.S. Legal Immigration

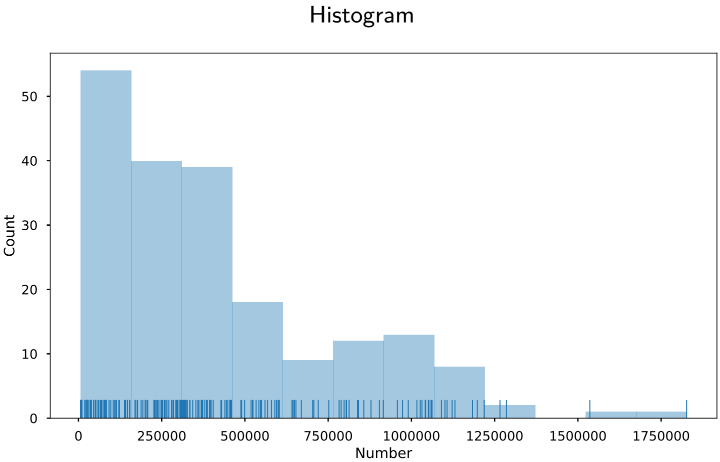


U.S. Legal Immigration Since WWI



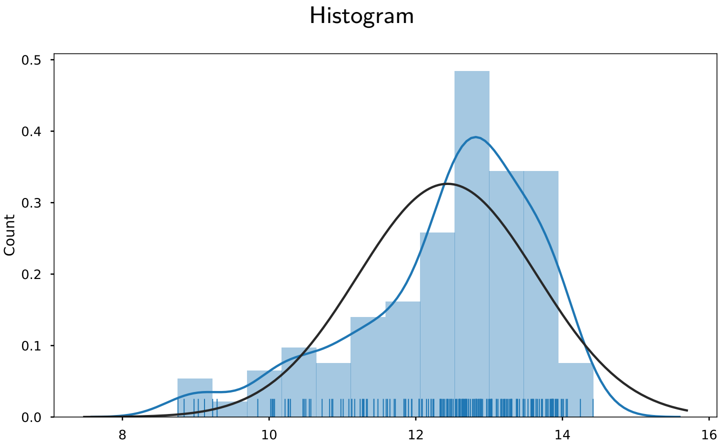
U.S. Legal Immigration

Count	197
Mean	4.20×10^5
Standard Deviation	3.58×10^5
Skewness	1.09
Excess Kurtosis	0.76



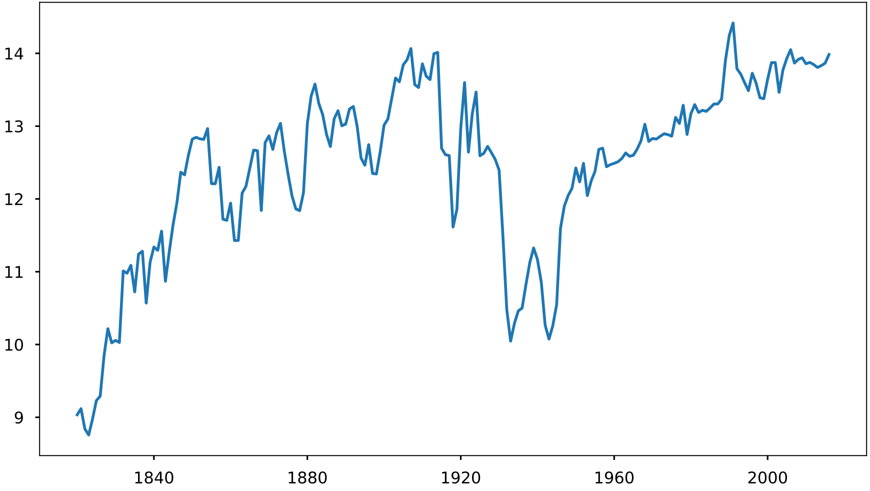
Log Immigration

Count	197
Mean	12.43
Standard Deviation	1.23
Skewness	−0.99
Excess Kurtosis	0.53



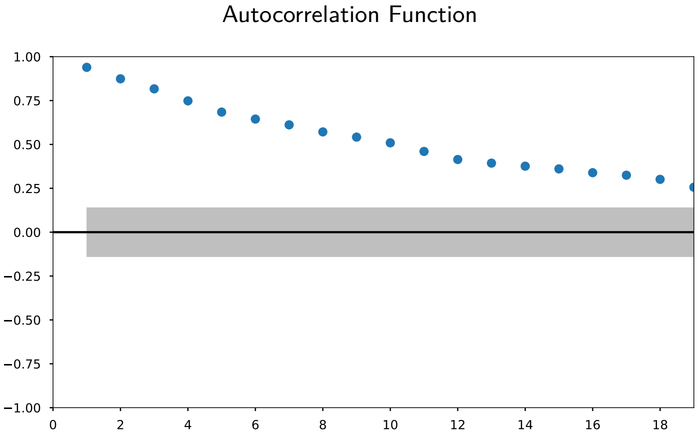
Standardized Log Immigration

Annual Data from 1918 to 2016



Autocorrelation

$$\text{Autocorrelation}(s) = \mathbb{C}\text{orr}(x_t, x_{t-s}) = \frac{\mathbb{C}\text{ov}(x_t, x_{t-s})}{\sqrt{\mathbb{V}\text{ar}(x_t)}\sqrt{\mathbb{V}\text{ar}(x_{t-s})}}$$



What is a Trend?

- When can we take means / variances, run OLS, etc?
- Simple answer: When the data are not trending.
- Intuitively, we need to the data to come from the some distribution in 1820, 1900, and 2010.
  - Statisticians & econometricians call this stationarity.
- If the data follow a random walk with drift, the data aren't stationary!

$$x_t = x_{t-1} + \eta_t$$

- The data do not revert some long run mean.

$$x_t = (x_{t-2} + \eta_{t-1}) + \eta_t = (x_{t-3} + \eta_{t-2}) + \eta_{t-1} + \eta_t = x_{t-h} + \sum_{j=0}^{h-1} \eta_{t-j}$$

Let's Test This

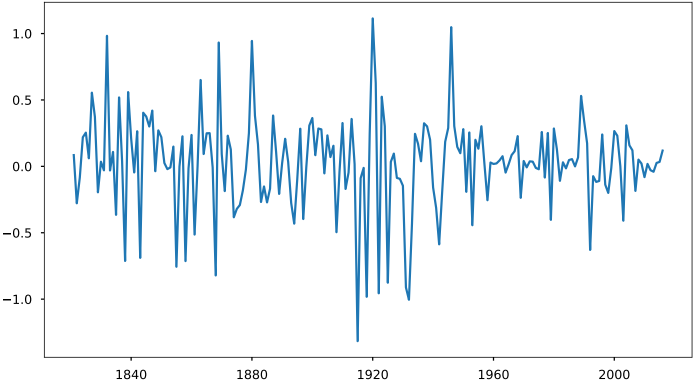
- Run OLS: $x_t = \beta_0 + \beta x_{t-1} + \eta_t$.
- $\beta = \begin{pmatrix} 12.86 & 0.94 \end{pmatrix}$.
- We cannot use Student's t-test to check for stationarity when the process might have infinite variance.
- Use the Augmented Dickey-Fuller (ADF) Test.
  - Standard Programming languages have commands to do this that report the p-values.
  - p-value: $8.03\,\% > 5\,\%$
  - We cannot reject a random-walk.
- Can we reject mean-reversion in favor a unit-root?
- Use the Kwiatkowski-Phillips-Schmidt-Schin (KPSS) test.
  - p-value: $2.66\,\% < 5\,\%$.
  - Yes, we can reject!
  - Immigration has a trend, (or at least it is not stationary).

How Do We Fix Stationarity?

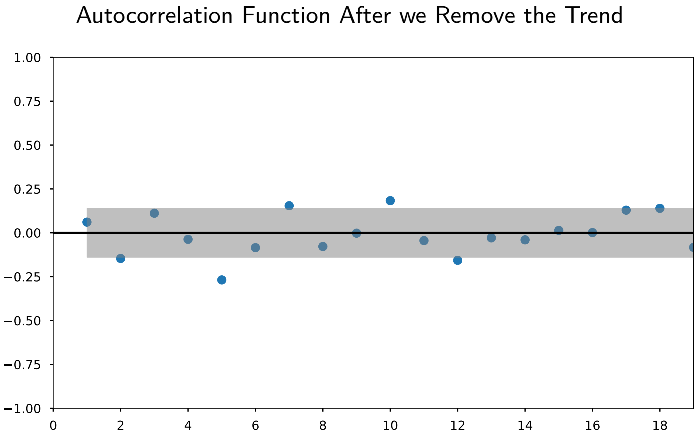
- We difference the data.
- Recall $x_t = x_{t-1} + \eta_t \implies x_t - x_{t-1} = \eta_t$.
- Define $\tilde{x}_t = x_t - x_{t-1}$.

- The data have a much less obvious pattern.
- A pattern is a signal.
- We want to find signals.
- When we're done, we'll have noise.

Differenced Log Immigration ($\approx$ percent-change)



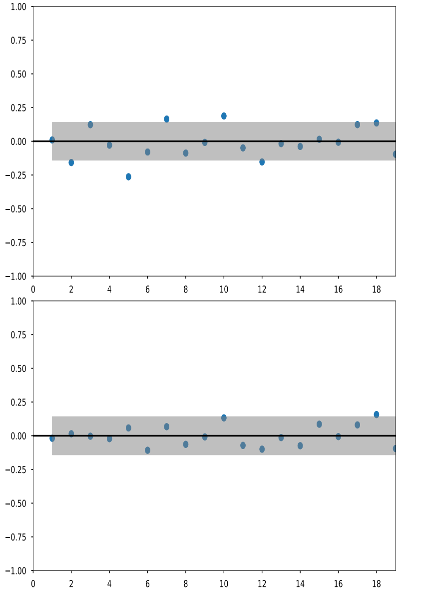
Are Any Patterns Left



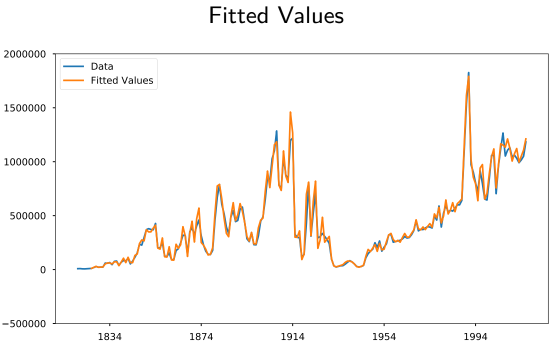
Autoregressive (AR) Models

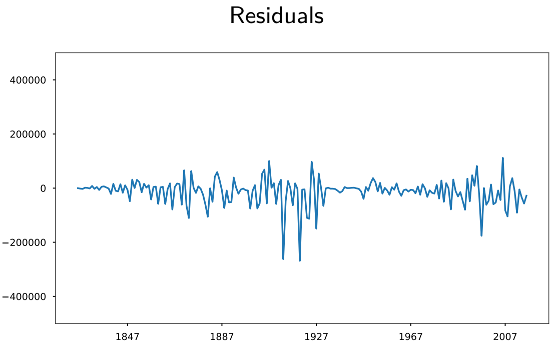
$$\tilde{x}_t = \beta_0 + \beta_1 \tilde{x}_{t-1} + \beta_2 \tilde{x}_{t-2} + \cdots + \beta_k \tilde{x}_{t-k} + \eta_t$$

β_0	β_1	β_2	β_3	β_4	β_5	σ
0.03	0.07					0.35
[0.93]	[0.85]					
0.03	0.08	−0.14				0.35
[1.13]	[0.99]	[−2.10]				
0.03	0.09	−0.16	0.13			0.35
[1.01]	[1.25]	[−2.24]	[1.86]			
0.03	0.08	−0.14	0.10	−0.06	−0.23	0.34
[1.26]	[1.14]	[−1.94]	[1.46]	[−0.83]	[−3.21]	



In-Sample Forecasts



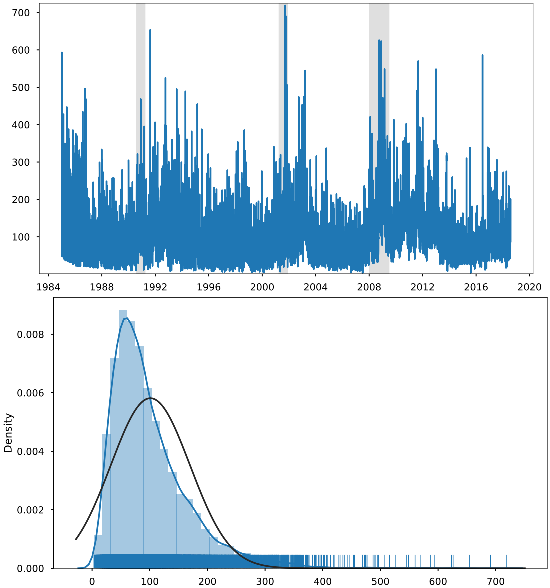


Data: Economic Policy Uncertainty Index

- Measures people's beliefs about economic activity using their words.
- Uses a database of 10 leading U.S. newspapers
- Measures the frequency of the following trio of terms:
    1. economic, economy
    2. Congress, deficit, Federal Reserve, legislation, regulation, White House
    3. uncertain, uncertainty
- As mentioned above, it was published two years ago and already has almost 2000 citations on Google Scholar.

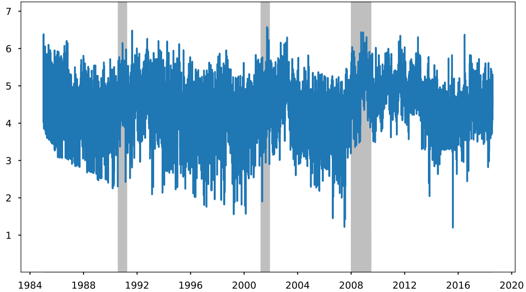
Economic Policy Uncertainty Index

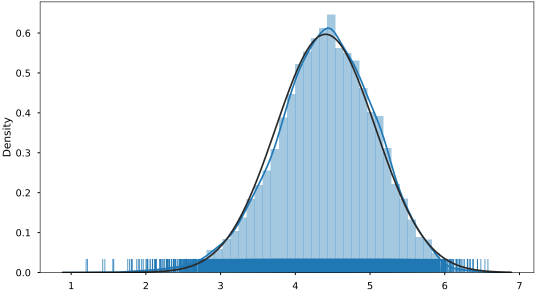
Count	12,266
Mean	101.08
Standard Deviation	68.62
Skewness	1.85
Excess Kurtosis	5.90



Economic Policy Uncertainty Index

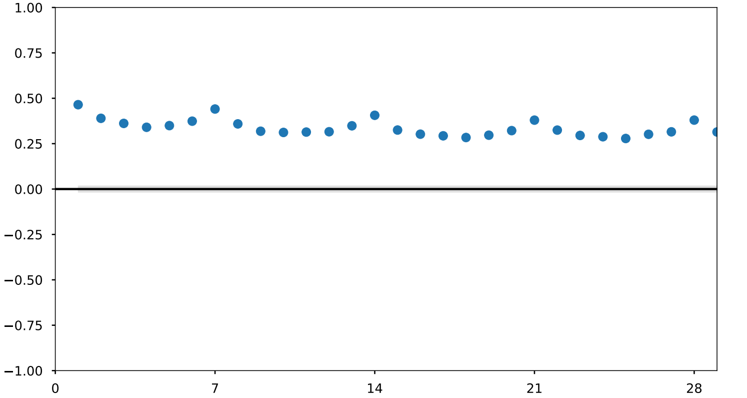
Daily Data	12,266 days
Mean	4.41
Standard Deviation	0.67
Skewness	-0.31
Excess Kurtosis	0.35





Economic Policy Uncertainty Autocorrelation Function

- Clearly not a random walk.
- Still persistent.
- Weekly peaks



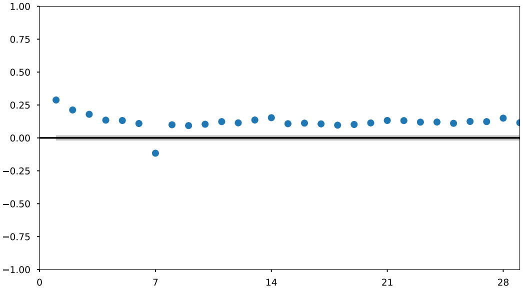
Seasonality

– Can we regress x_t on x_{t-7}?

$$x_t = \beta_0 + \beta_w x_{t-7} + \eta_t$$

β_0	β_1	σ
2.36	0.44	0.36
[73.72]	[47.88]	

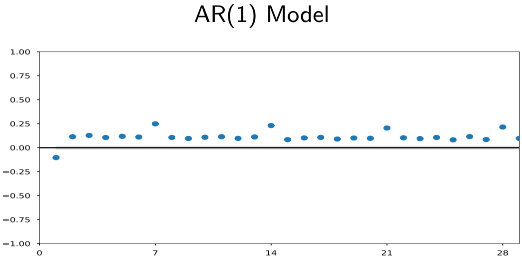
Autocorrelation Plot

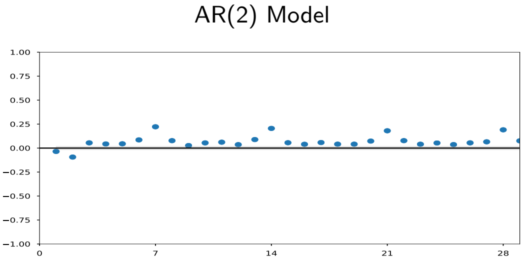


Autoregressive Models

$$x_t = \beta_0 + \beta_1 x_{t-1} + \cdots \beta_k x_{t-k} + \eta_t$$

β_0	β_1	β_2	σ
4.41	0.46		0.59
[441.37]	[58.11]		
1.84	0.36	0.22	0.33
[46.41]	[44.10]	[25.86]	





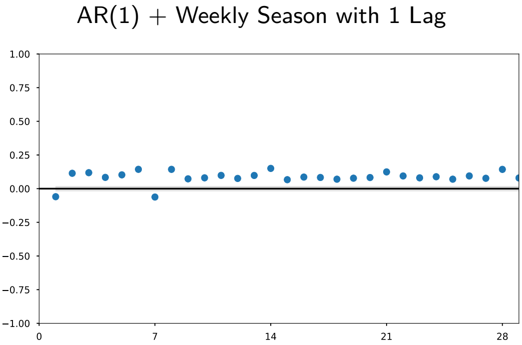
Can We Combine the Autoregressive and Seasonality Models?

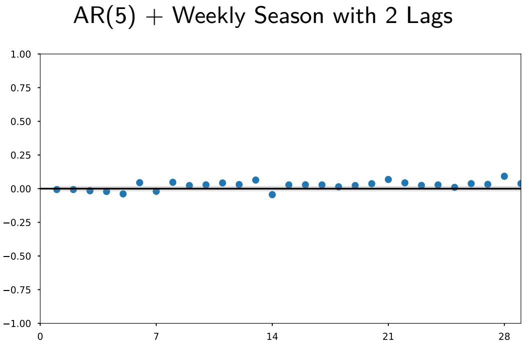
- Yes!
- Intuition:
  1. Apply seasonality model to the data.
  2. Apply autoregressive model to those residuals.
- Not just adding a weekly lag to the autoregressive model. We must worry about interactions.
- Standard statistical programs can do this for us, we just have to tell them to.

Autoregressive + Season Model

β_0	β_1	β_2	β_3	β_4	β_5	β_{w1}	β_{w2}	σ
2.01	0.35					0.30		0.32
[31.49]	[43.24]					[37.18]		
1.70	0.30	0.18				0.26		0.31
[49.49]	[35.94]	[20.85]				[31.70]		
1.17	0.25	0.12	0.09	0.05	0.09	0.18	0.17	0.30
[31.49]	[29.68]	[13.02]	[9.81]	[6.21]	[9.88]	[22.07]	[20.33]	

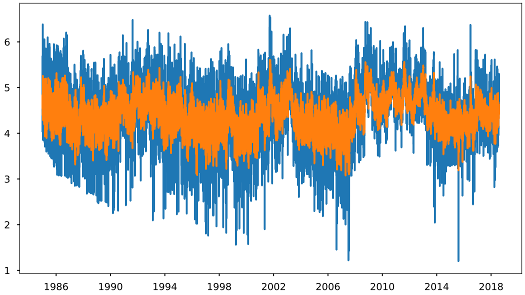
Autocorrelation Plots



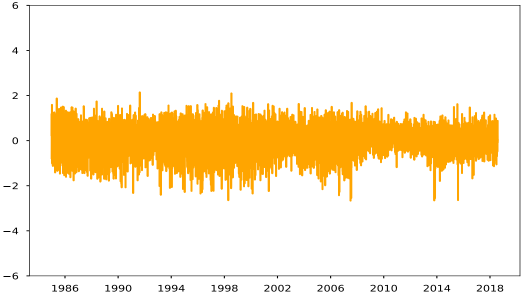


Results

Fitted Values



Residuals

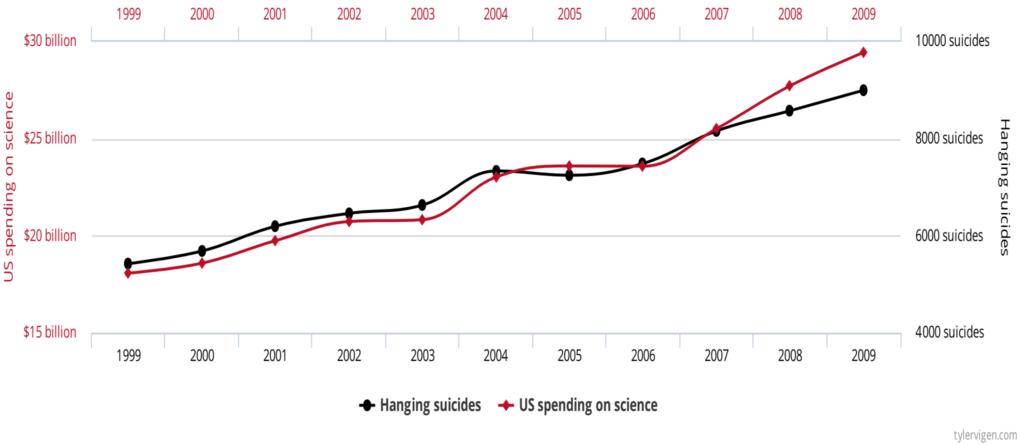


Is it Signal or Noise

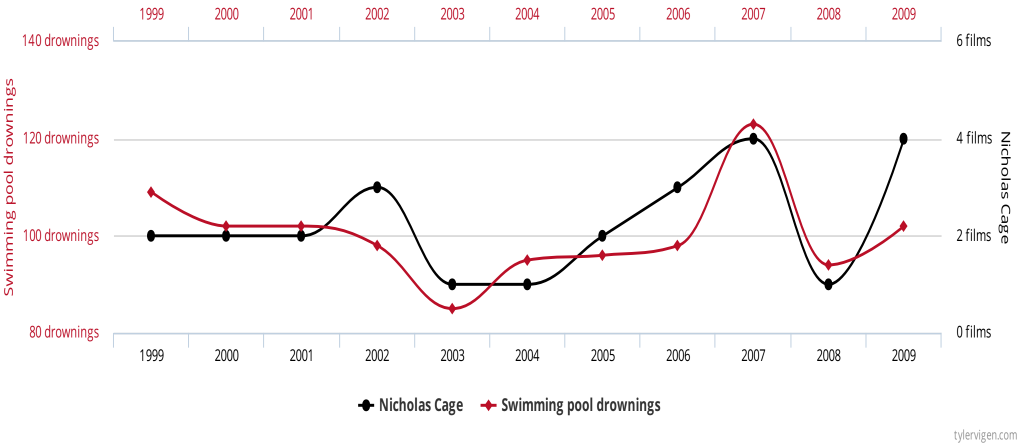
$$\mathbb{R}^2 = 1 - \frac{\sum_{t=1}^{T} \hat{\eta}_t^2}{\sum_{t=1}^{T}(x_t - \bar{x})^2}$$

- OLS picks $\hat{\beta}$ to make $\sum_{t=1}^{T} \hat{\eta}_t^2$ as small as possible.
- As we add more parameters, this gets easier and easier.
- We will start fitting noise eventually!

Example : U.S. Suicide Rate



Example : Drownings

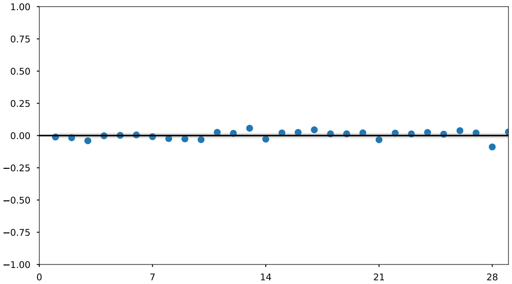


AIC and BIC

1. AIC
   - Akaike Information Criterion
   - $AIC = -2\sum_{t=1}^{T} u_t^2 + 2(\#\mathrm{Params})$ (under Normality)
2. BIC
   - Bayesian Information Criterion
   - Also called Schwarz Information Criterion (SIC)
   - $BIC = -2\sum_{t=1}^{T} u_t^2 + \log(T)(\#\mathrm{Params})$ (under Normality)
   - Useful if you think the true model is one of the ones under consideration.
   - Picks more parsimonious models than AIC.

Selected Model

- BIC Chooses 6 AR lags and 4 seasonal lags.
- AIC Chooses 10 AR lags and 4 seasonal lags



Conclusion

1. Time Series analysis is useful because time matters.
   - We often care what came first or if something acted as we would expect.
2. Time Series can be decomposed into Trend + Seasonal + Cycle + Noise
3. The goal is to figure out what the trend, seasonal and cycle are.
4. We must be careful to ensure the distribution is not changing too much over time. (Stationarity).
5. We can use standard statistical packages, graphs, and common sense to do this well.