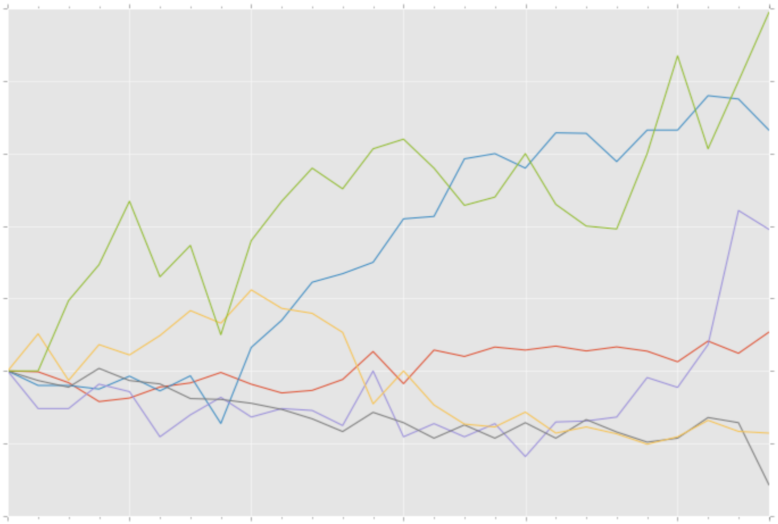
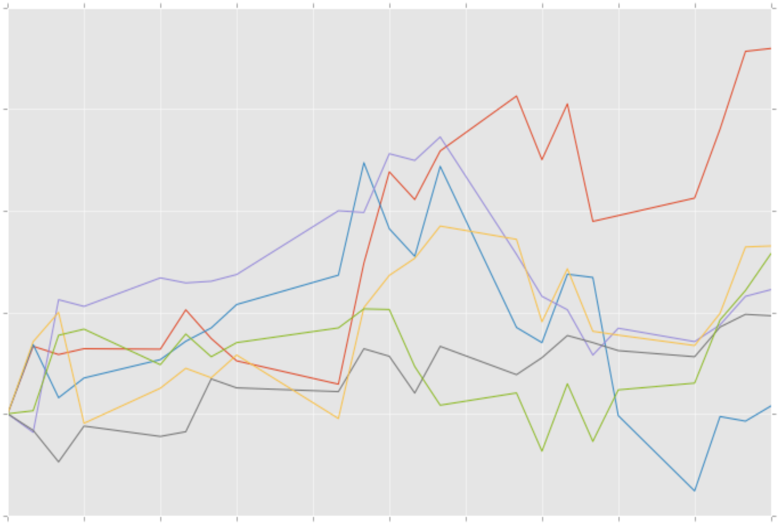
Politics and Finance: Comparing Data

KEVIN SANDERS

BOSTON UNIVERSITY

Motivation





Overview

Motivation

Limitations

- Data (type and amount)

Expectations

- Based on financial data

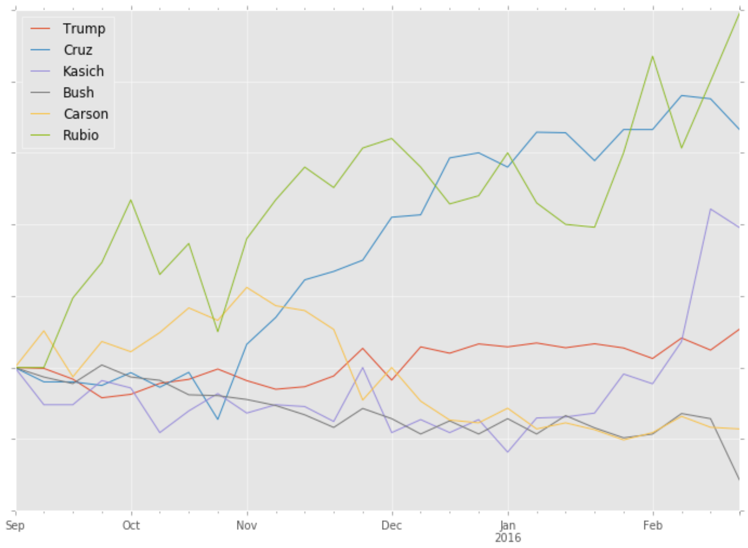
Reality

- Based on political data

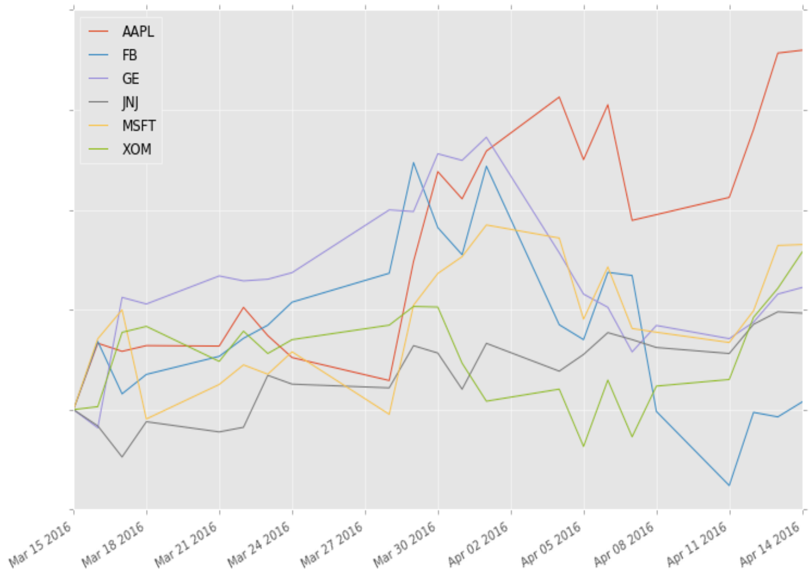
Next Steps

Motivation

2016 GOP PRIMARY



S&P 6



Limitations

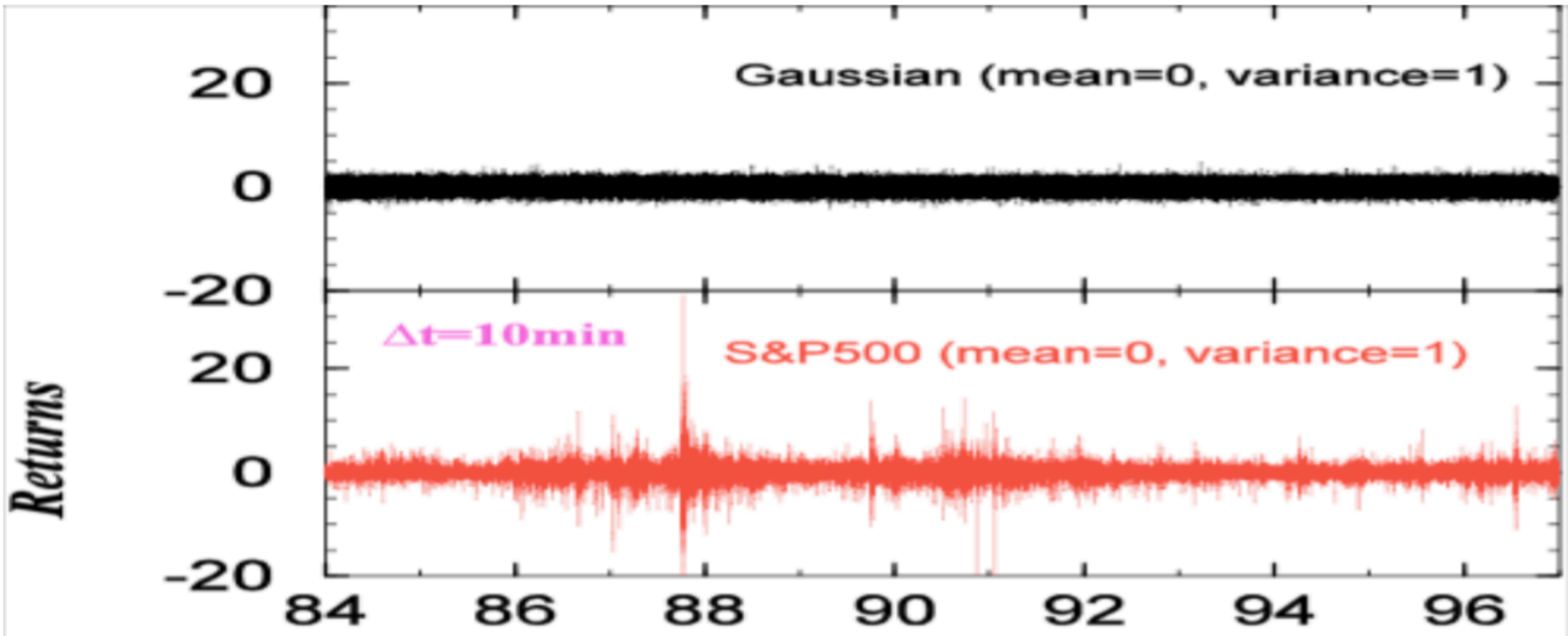
Lack of Data

- State-based results
- Polls daily at best
- Limited electronic data

Polls

- Only estimate actual value
- Different methods (phone vs internet)
- Differing levels of accuracy
- Arbitrage between different polls
- Focused around certain times

S&P500 Normalized Returns



*from Professor Stanley's slides.

Expectations for 2016 GOP Primary Polls

Similarities

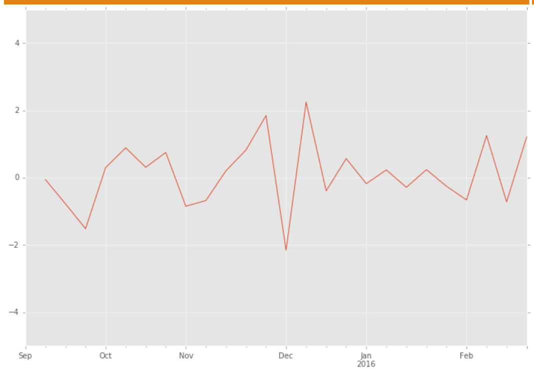
- Both time series data
- Unlikely to be normally distributed
- Probably both biased in some way

Differences

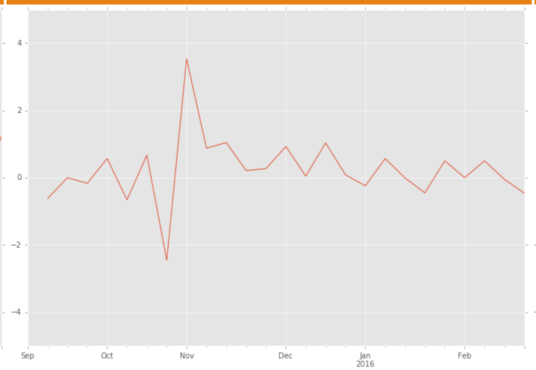
- Vastly different period of measurement
- Polls are limited by 100% (likely affects bias → one goes up the others must go down)
- Can only change by whole percentage points
- Candidates disappear much more frequently

Normalized Returns

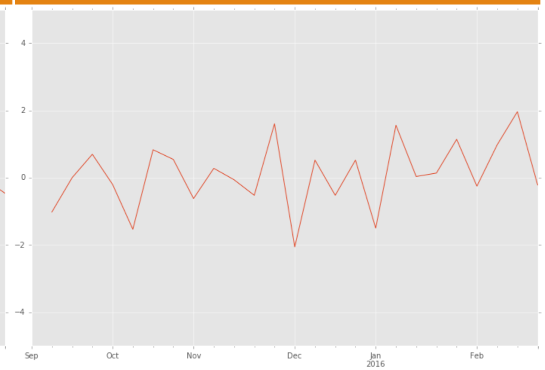
Trump



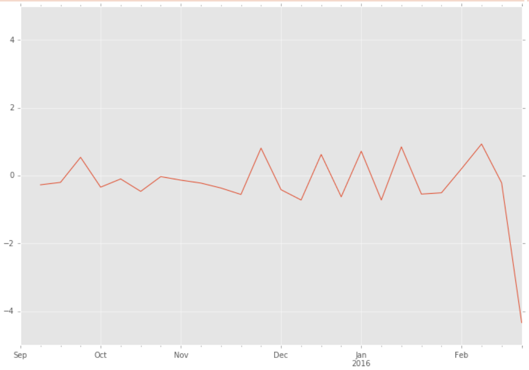
Cruz



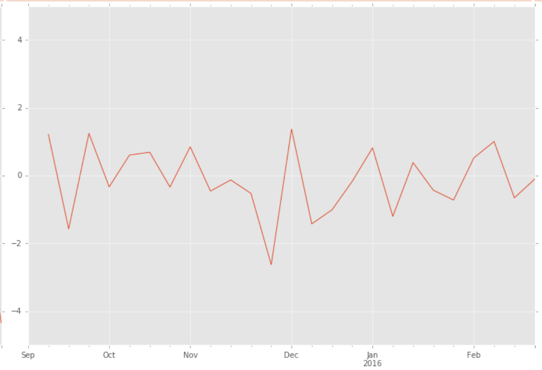
Kasich



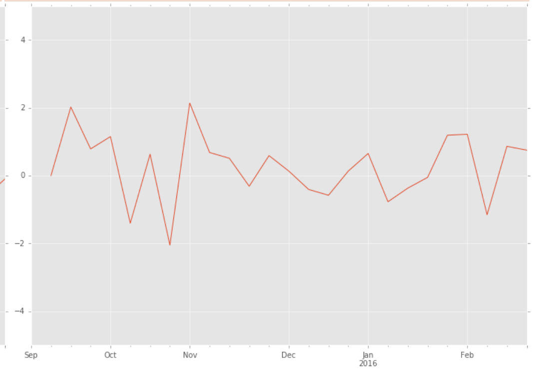
Bush



Carson

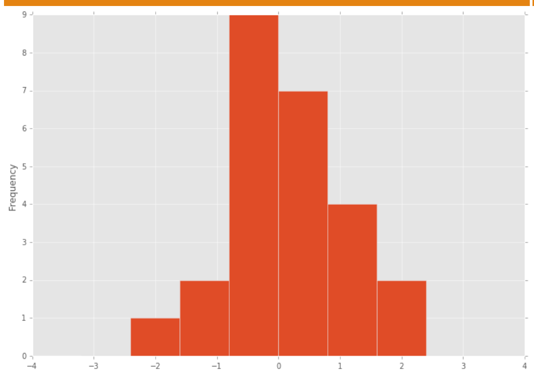


Rubio

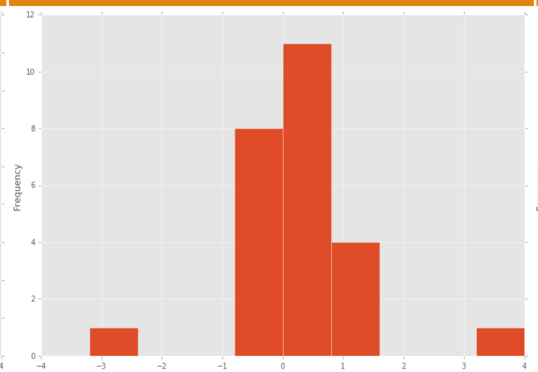


Normalized Returns

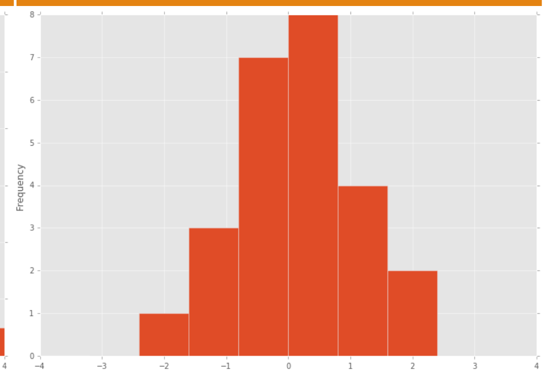
Trump



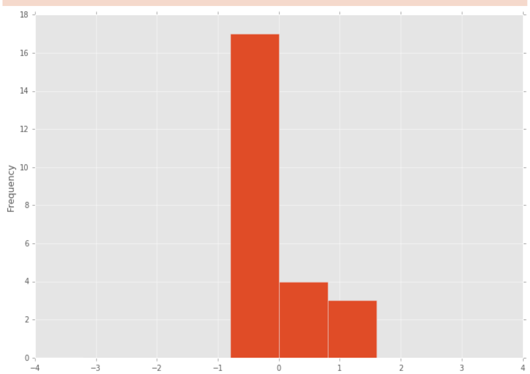
Cruz



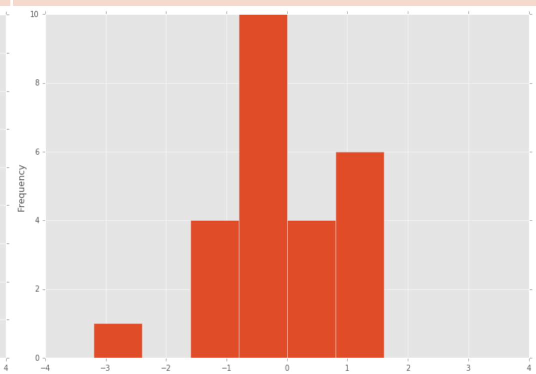
Kasich



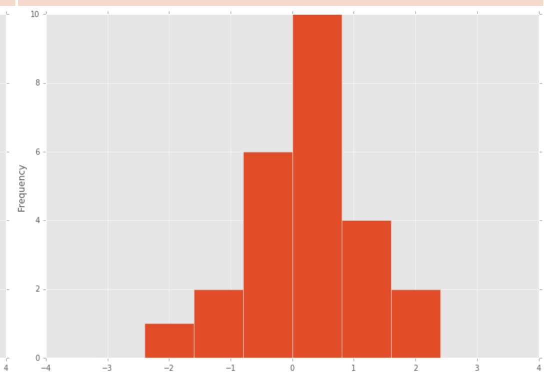
Bush



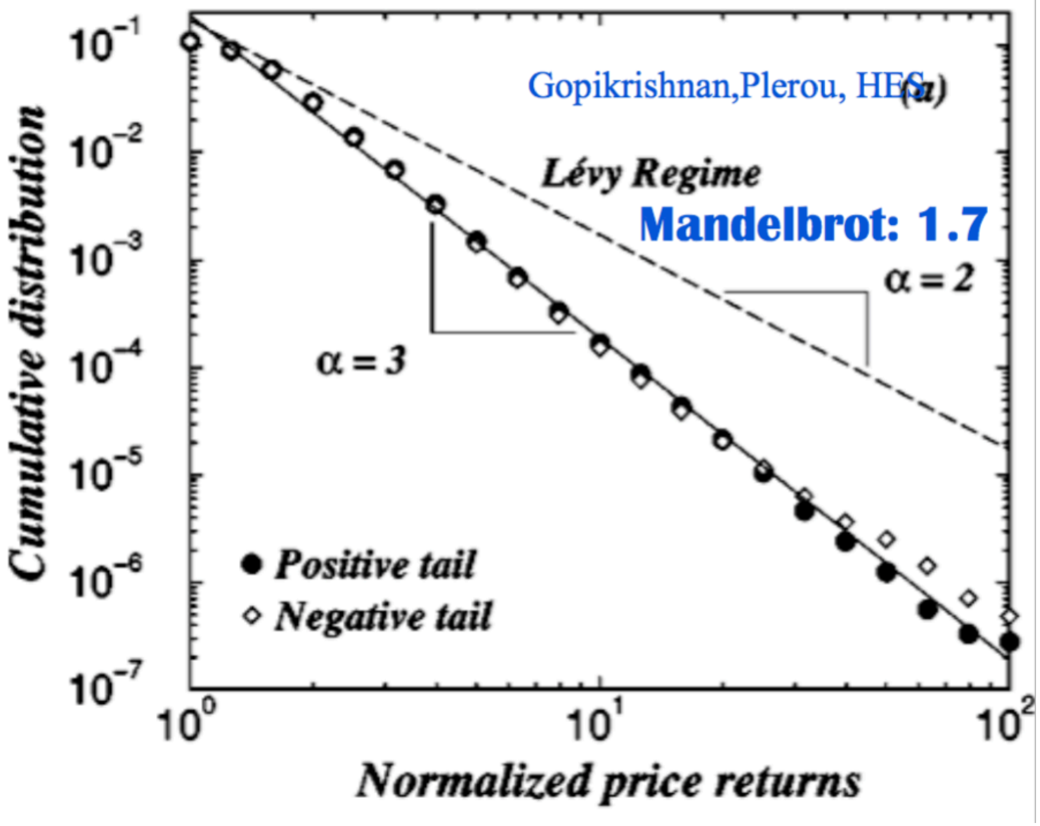
Carson



Rubio



Power Laws



*from Professor Stanley's slides.

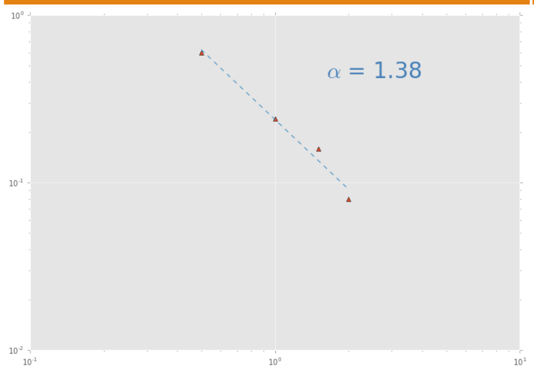
Power Laws

The histograms of returns appear to have a certain drop off away from the mean.

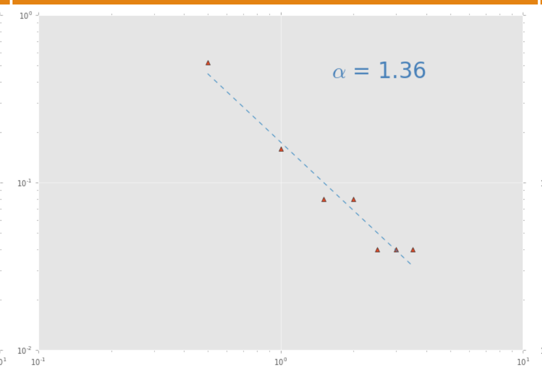
One way to confirm this is to look at the cumulative distribution for each histogram.

Cumulative Distribution

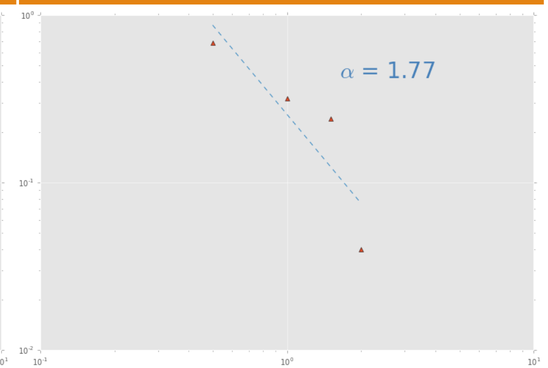
Trump



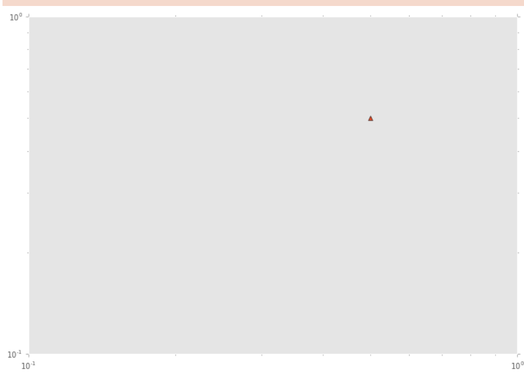
Cruz



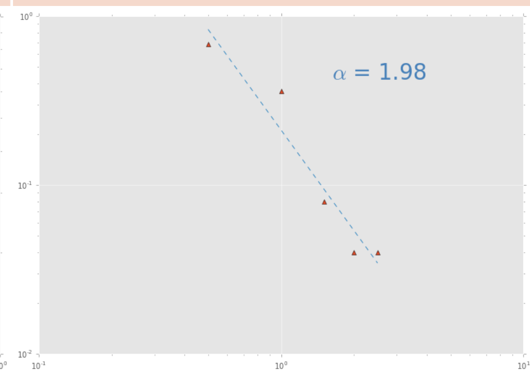
Kasich



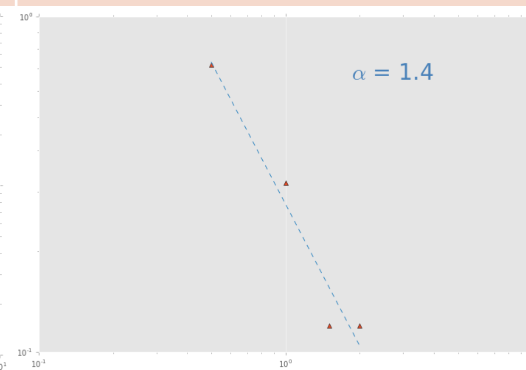
Bush



Carson



Rubio

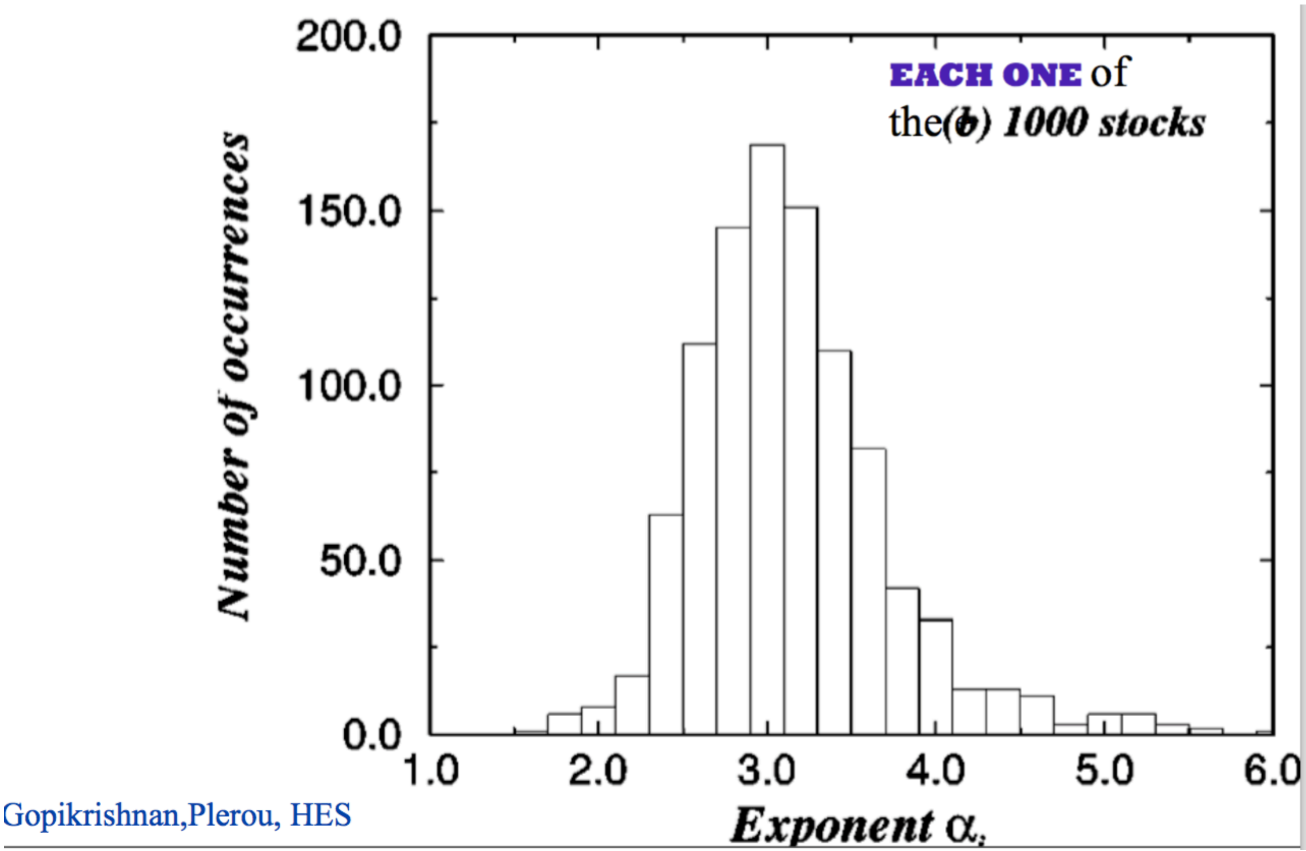


Info from this Power Law

Lower exponents tend to fair better

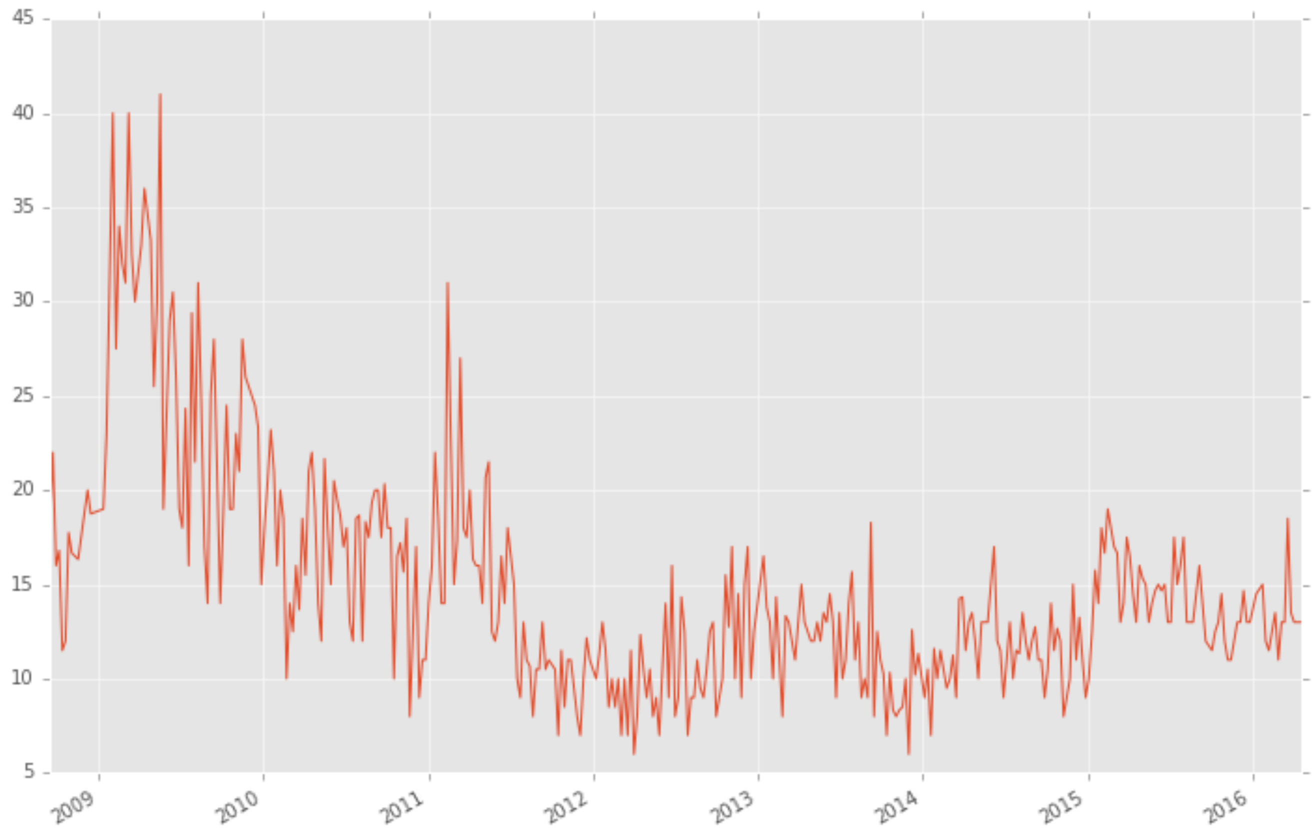
Candidates with very low support will likely have very different exponents (when taking into consideration the polls only change by whole percentage points)

Primary Polls and Stock Returns center around a particular α

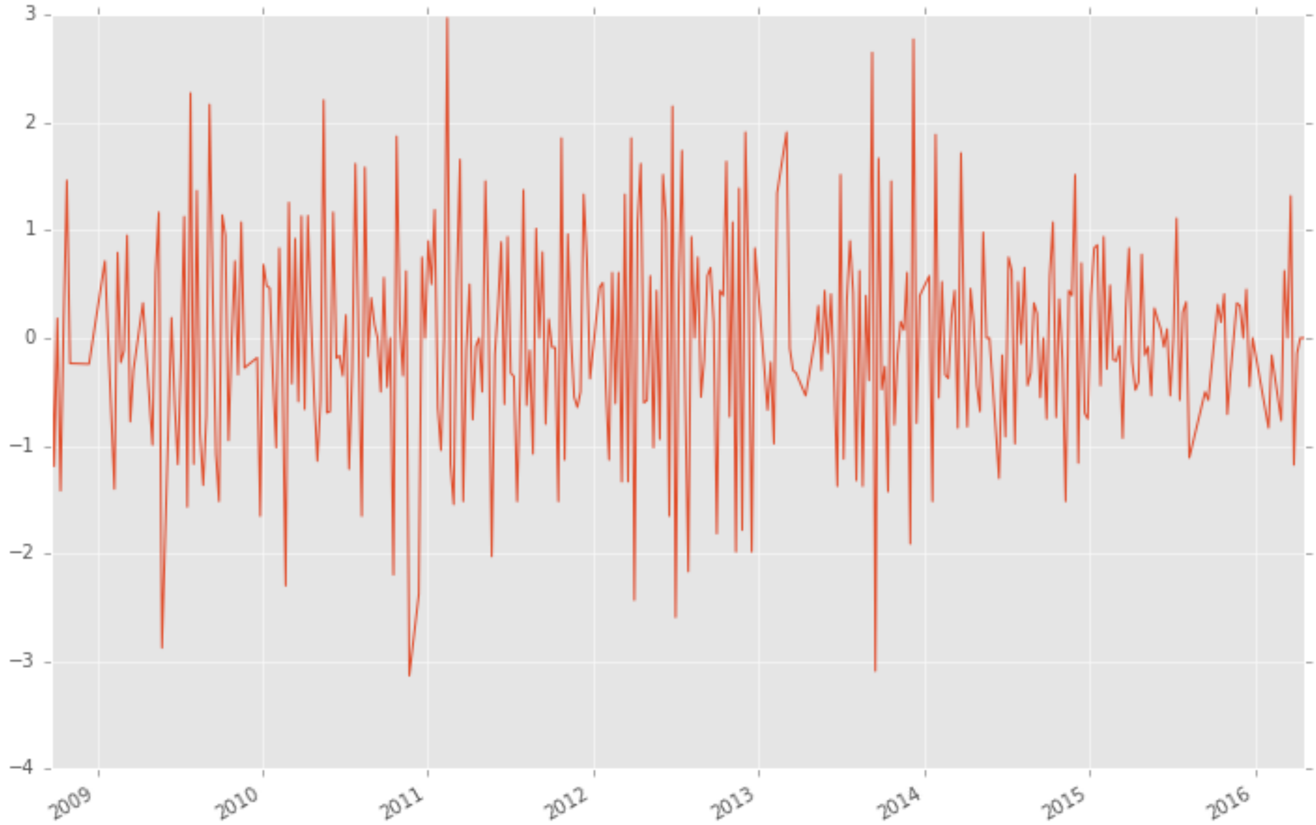


*from Professor Stanley's slides.

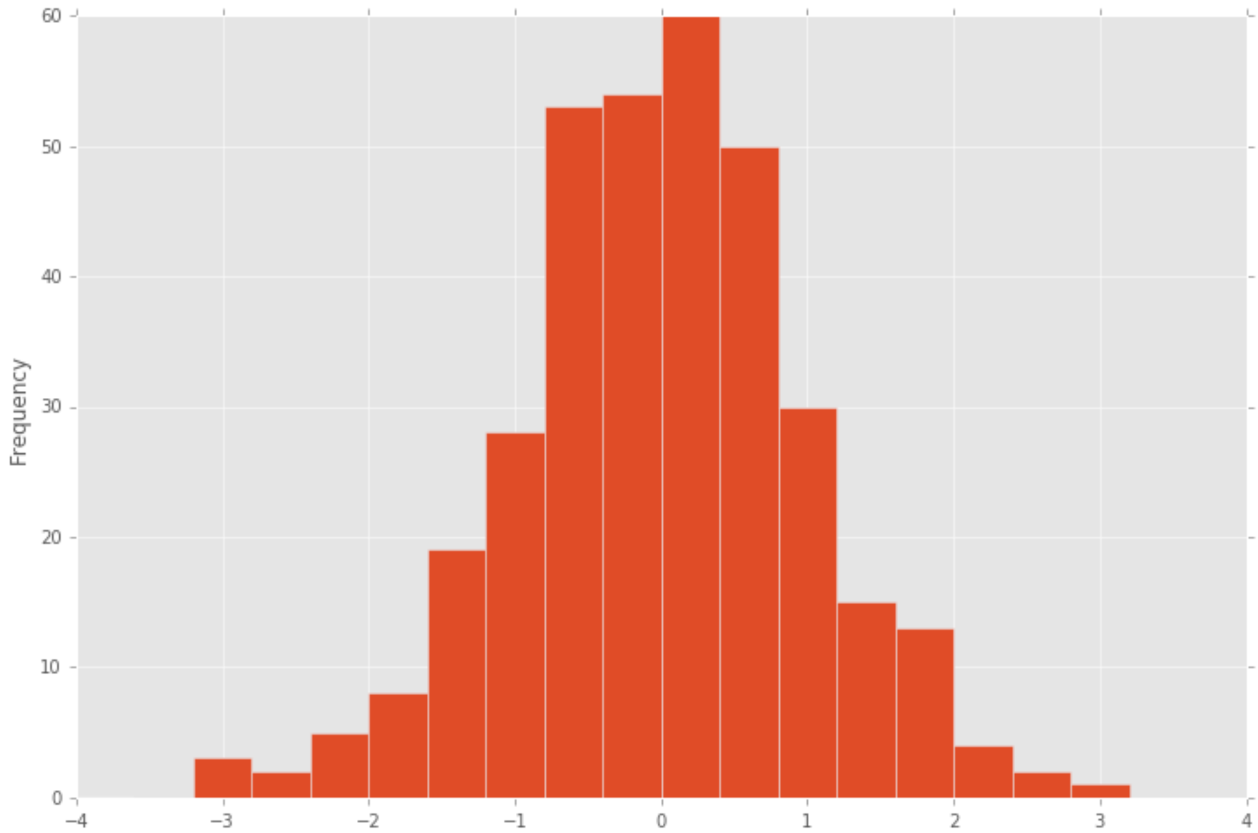
What about a longer time period?



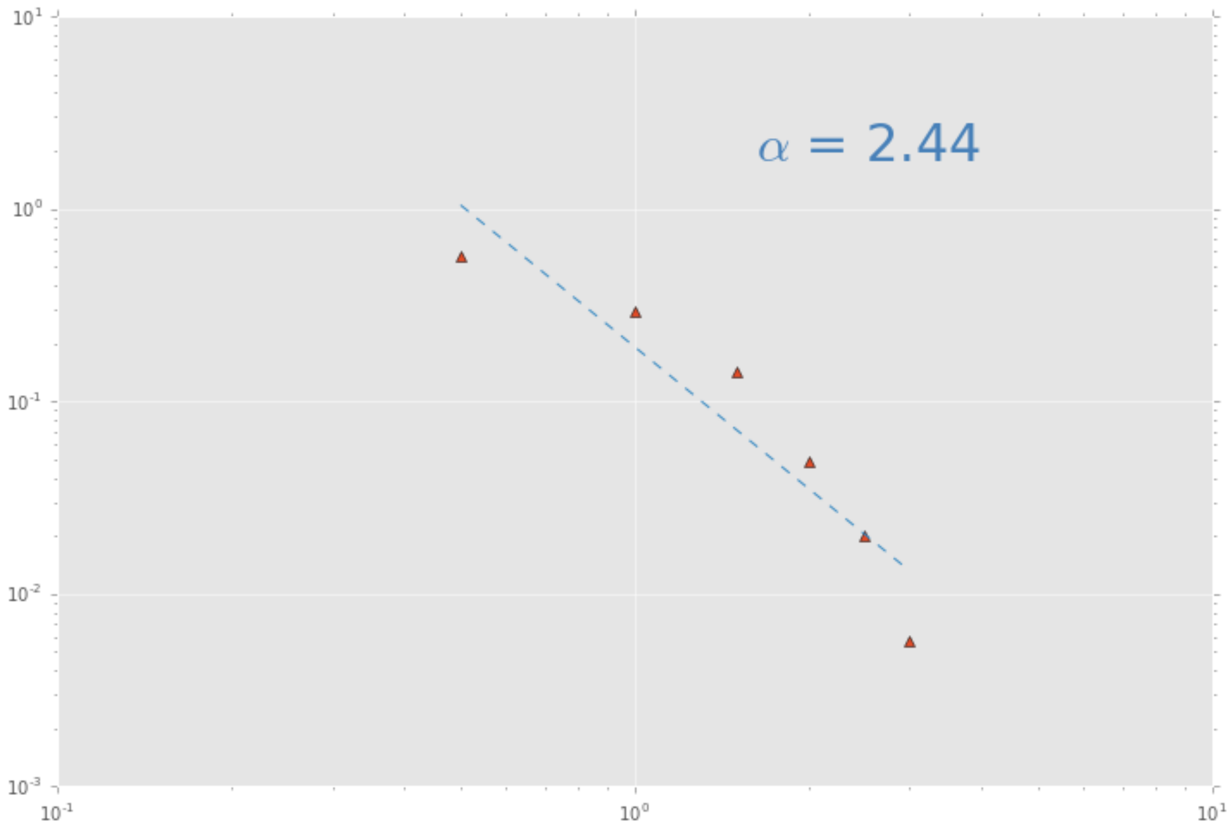
Congressional Approval normalized ‘Returns’



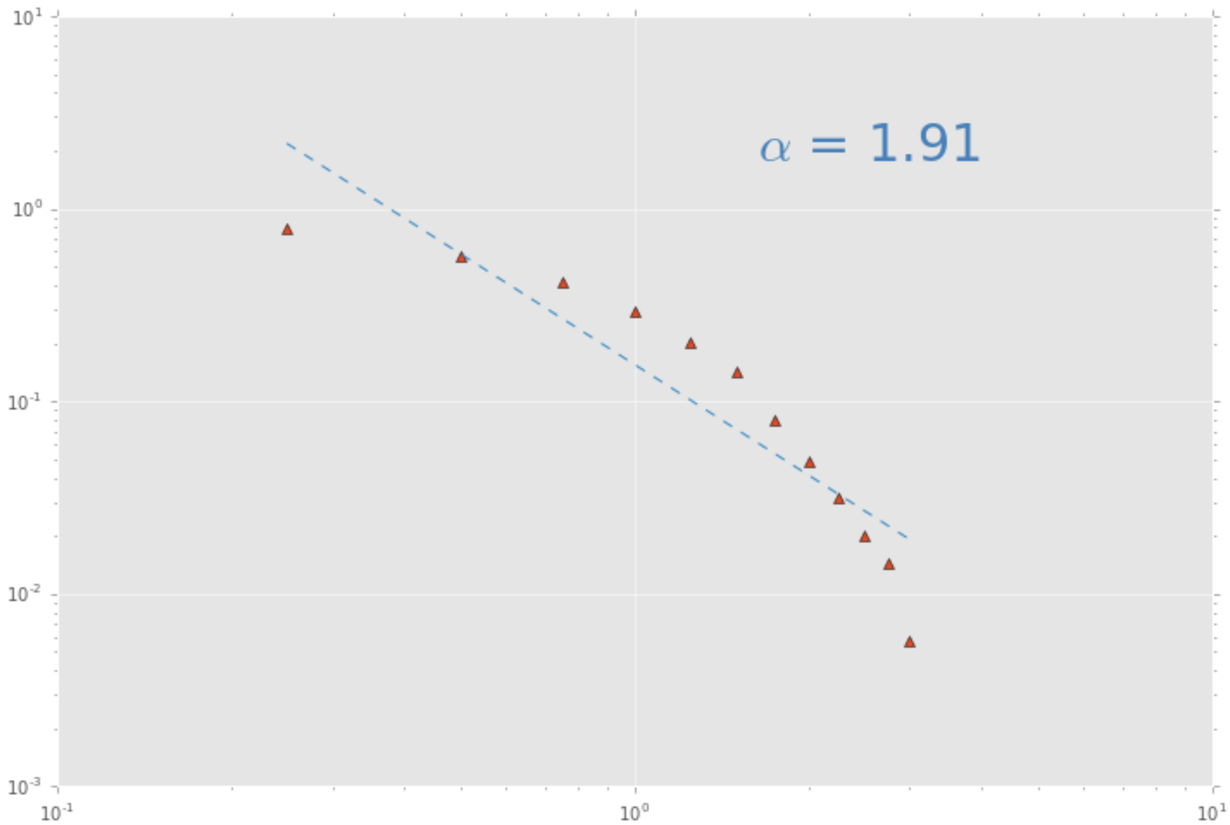
Congressional Approval normalized ‘Returns’



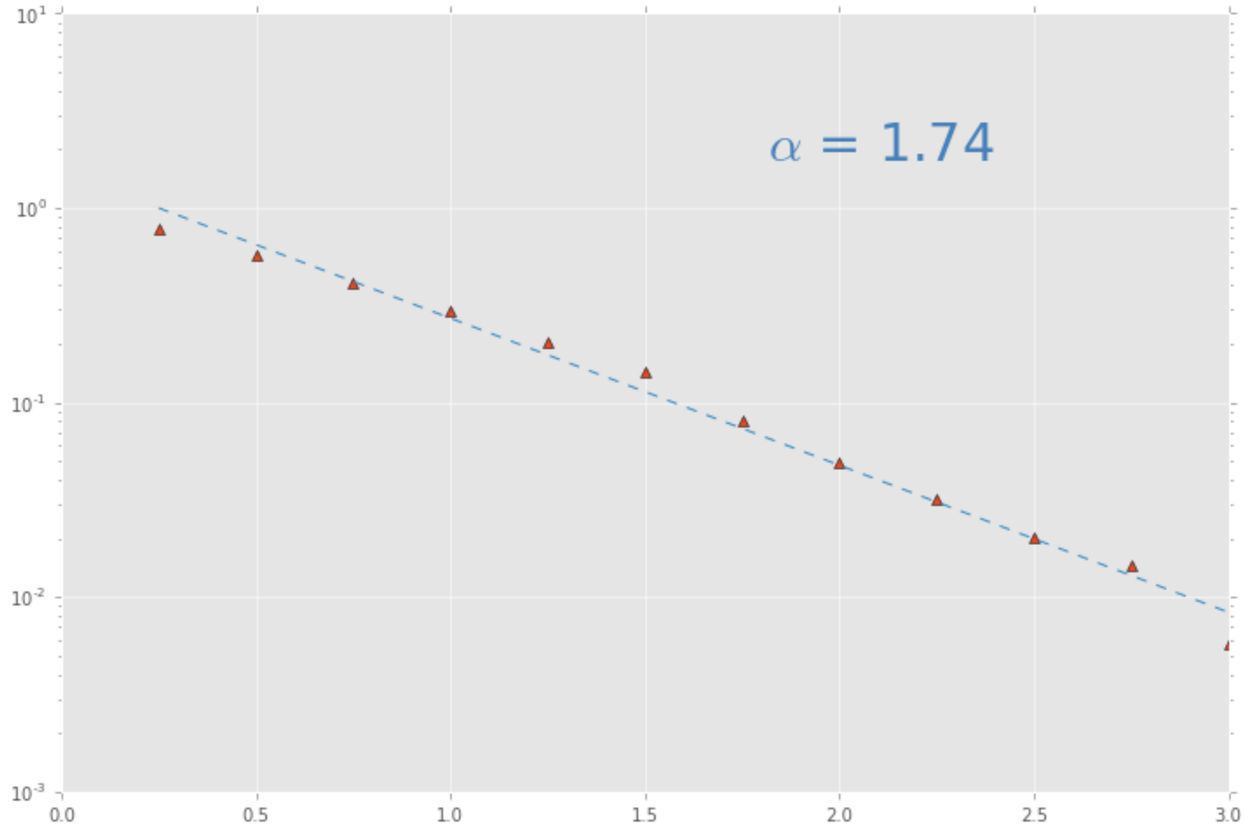
Congressional Approval Cumulative Distribution



Congressional Approval Cumulative Distribution (more bins)

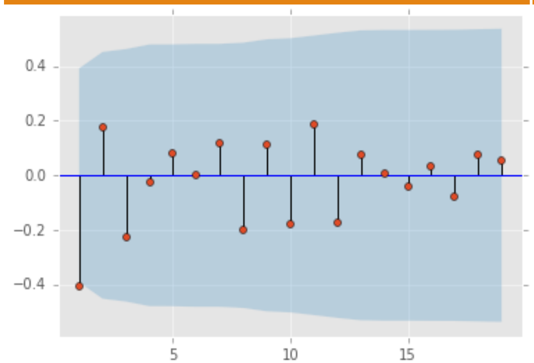


Must be Log Normal!

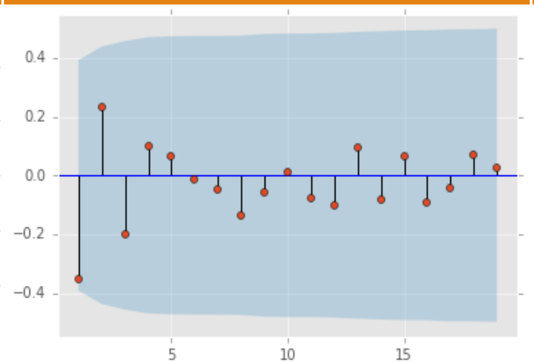


Autocorrelation Function (Returns)

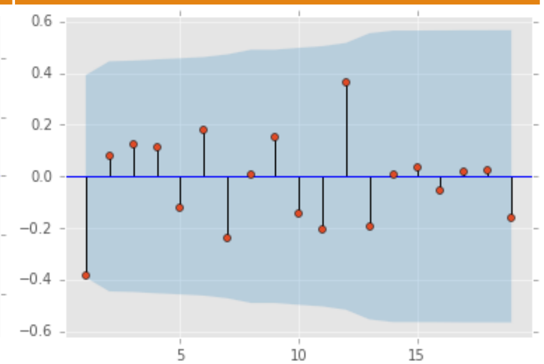
Trump



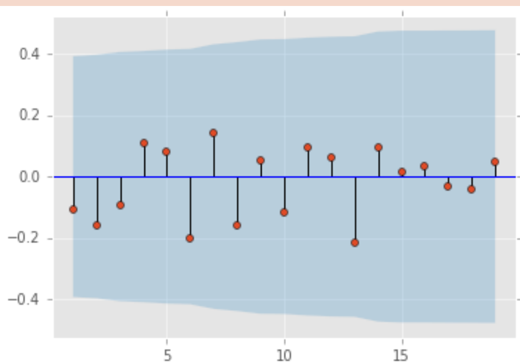
Cruz



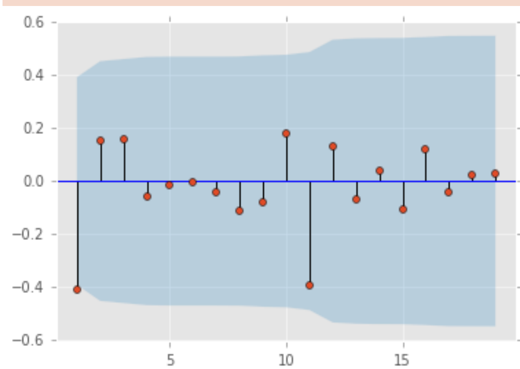
Kasich



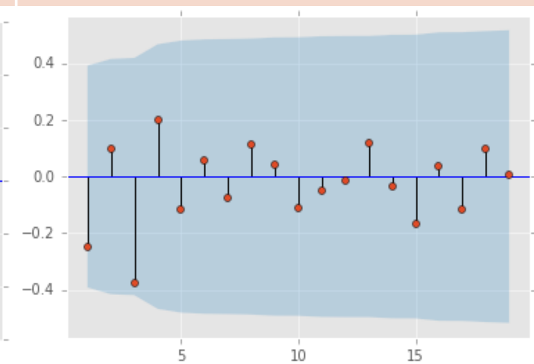
Bush



Carson

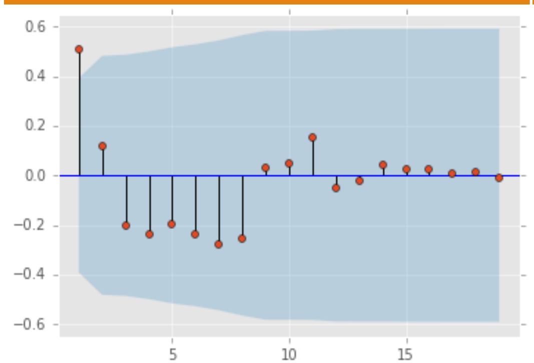


Rubio

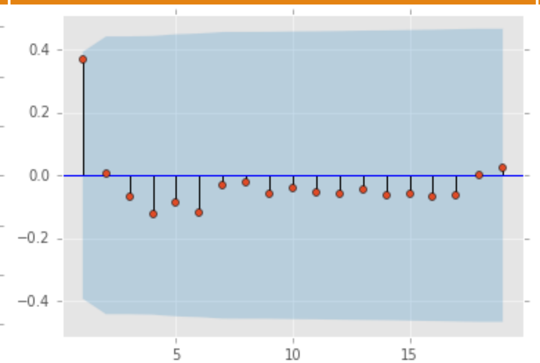


Autocorrelation Function (Volatility)

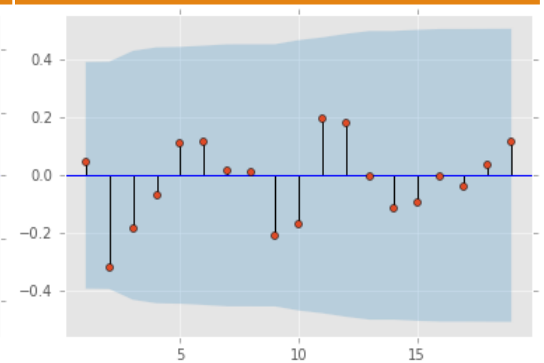
Trump



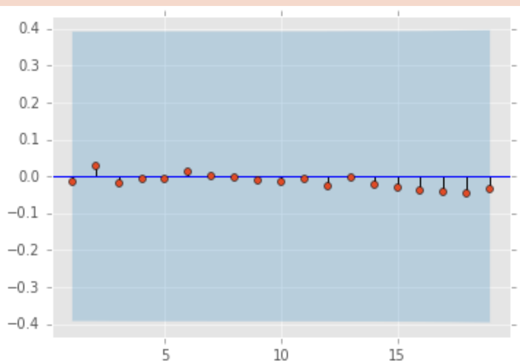
Cruz



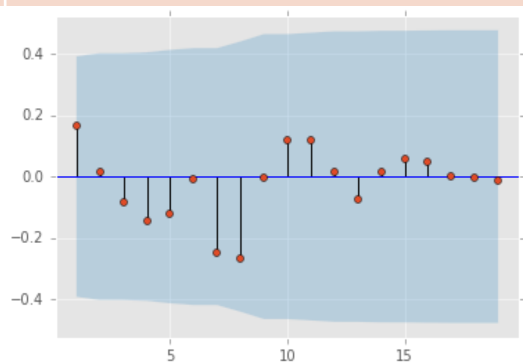
Kasich



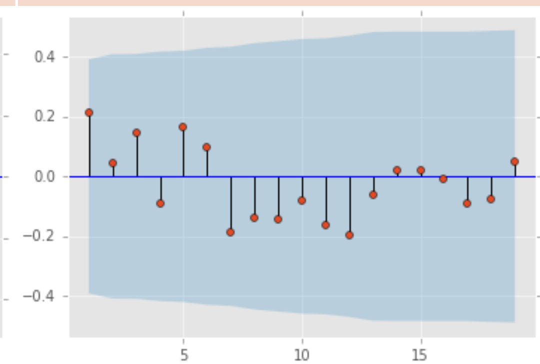
Bush



Carson



Rubio



Time Correlations

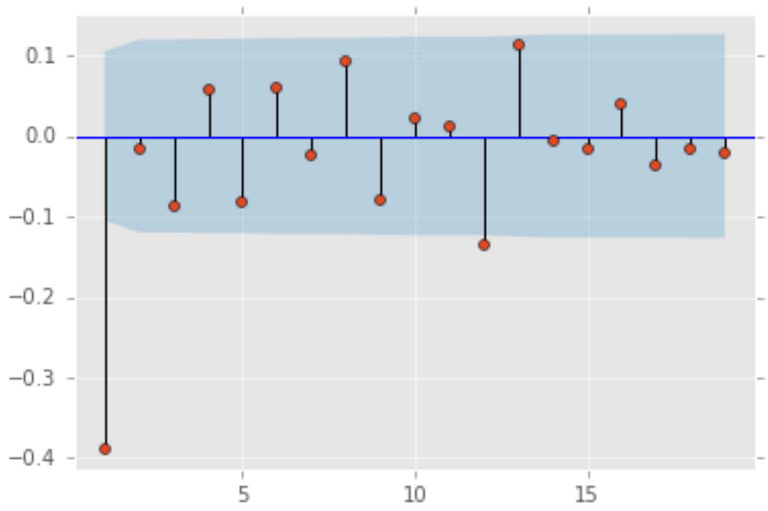
Political data appears uncorrelated for both returns and volatility!

This differs from stock data (short term return correlations but much longer term volatility clustering)

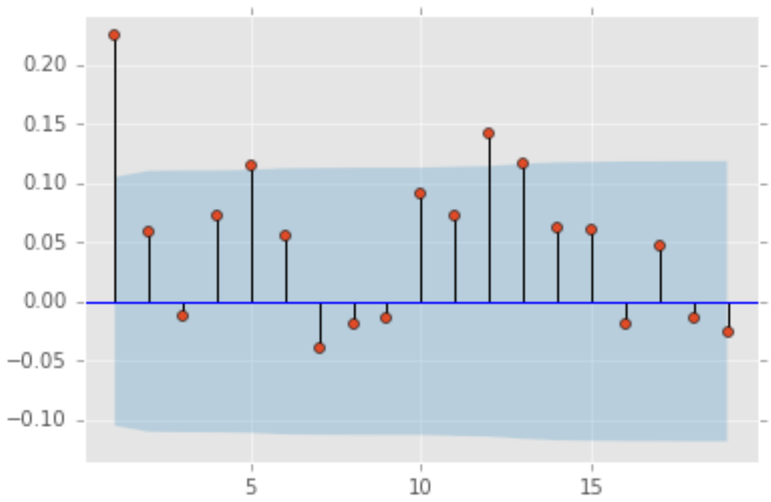
Might be a factor of short time period for data?

Congressional Approval

RETURN



VOLATILITY



Next Steps

Time correlations in long term data like approval ratings

Correlations (probably anticorrelations) between opposing candidates

Which polls are more accurate

Who is going to win

How long does it take for candidates to drop out

References

Professor Stanley's presentations.

Data from Huffington Post