Time Series Analysis

Lecture 5

Vector Autoregressive (VAR) Models

datascience@berkeley

Regression With Multiple Trending Time Series: An Introduction

Introduction

- Classical linear regression models assume that the stochastic errors are uncorrelated.

- In the context of time series analysis, however, this assumption is often violated.

- More troublesome is that two independent time series could appear related to each other and have very high "correlation" when they are in fact independent of each other.

- We will study examples of this phenomenon and introduce statistical tests relevant for this situation.

- After discussing concepts such as spurious correlation, **cointegration**, and testing for **unit roots**, we will discuss **vector autoregressive models** that can be used to model multiple time series.

An Example:
Electricity and Chocolate Production

- These are two series that we have worked with before and are provided by the book.
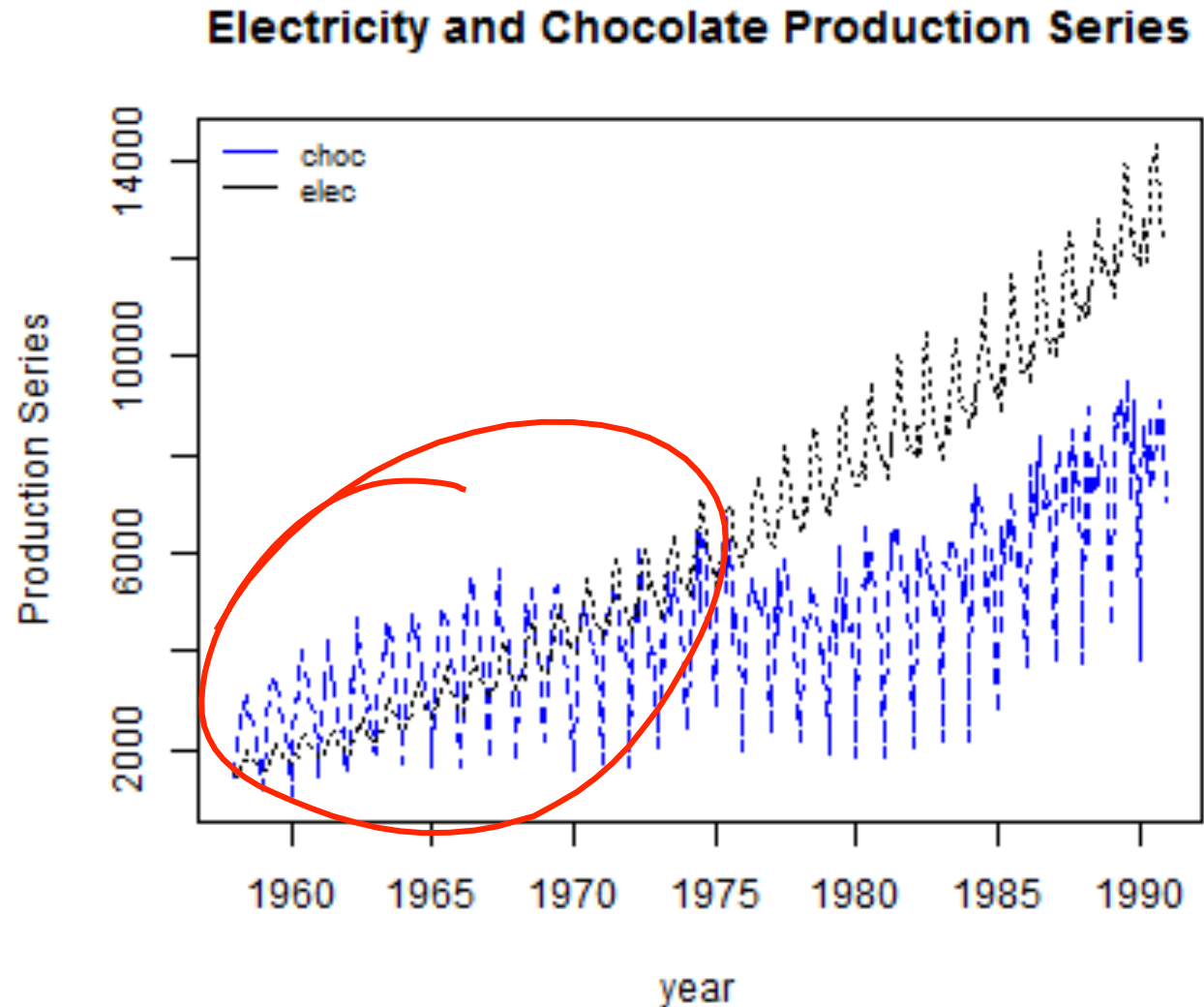
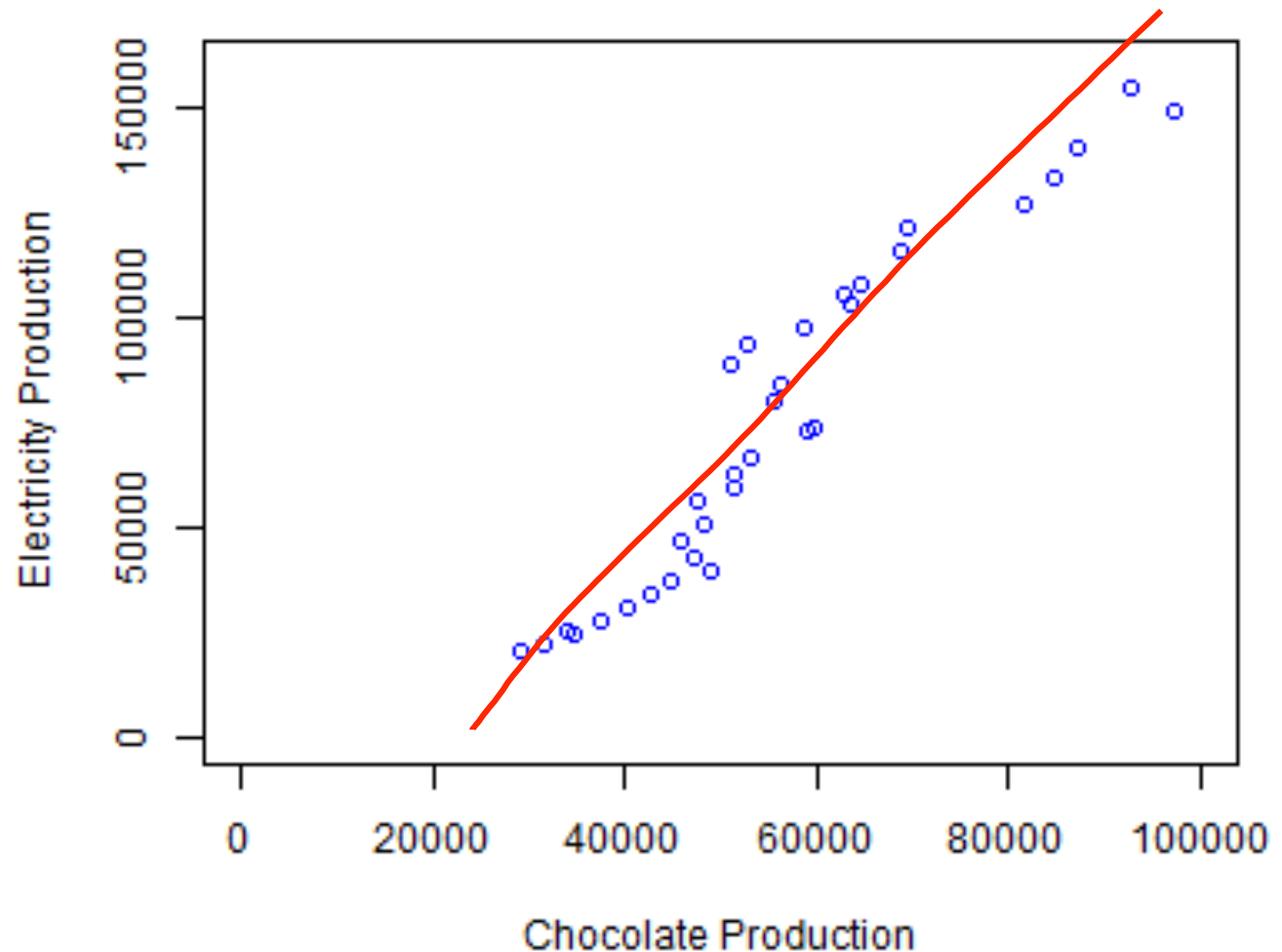# An Example:
Electricity and Chocolate Production

- Aggregate the series into an annual series.

- According to the scatter plot, the correlation with the two series seems to be very high.

- In fact, the correlation is **0.96!**

Annual Electricity Production vs. Annual Chocolate Produc

An Example:
Electricity and Chocolate Production

- It may be very tempting to fit a regression from one trending series on another and report high R^2. Don't!

- Fitting a regression of one variable as a linear function of the other, with added random variation, can often lead to *spurious* regression.

- The term *spurious regression* is also used when underlying stochastic trends in both series happen to be "coincident."