GENERAL ASSEMBLY

INTRODUCTION TO TIMESERIES MODELS

Christoph Rahmede, Data Science Immersive

AGENDA

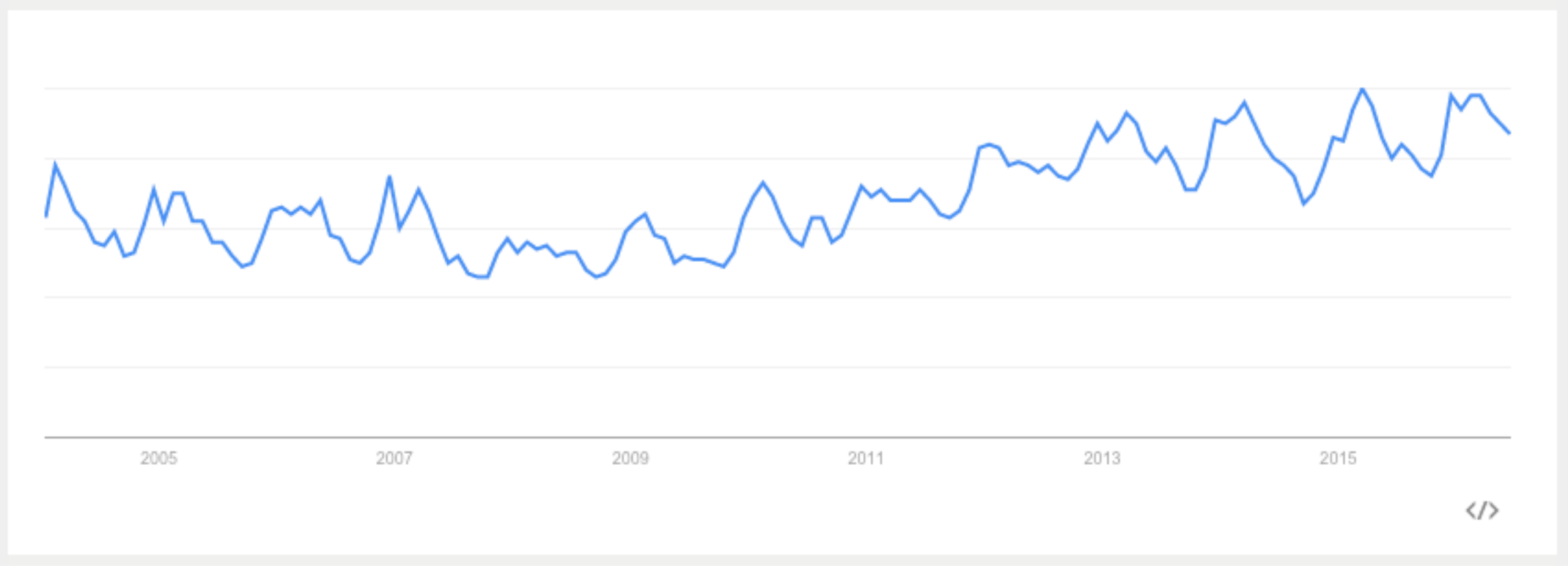
- What are time series
- EDA for time series
- Time series models
- Assumptions for time series models
- Model evaluation and forecasting

WHAT ARE TIME SERIES?

- A sequence of observations in chronological order
- We will mostly focus on
    - univariate time series (only one variable)
    - of continuous variables
    - over regular time intervals

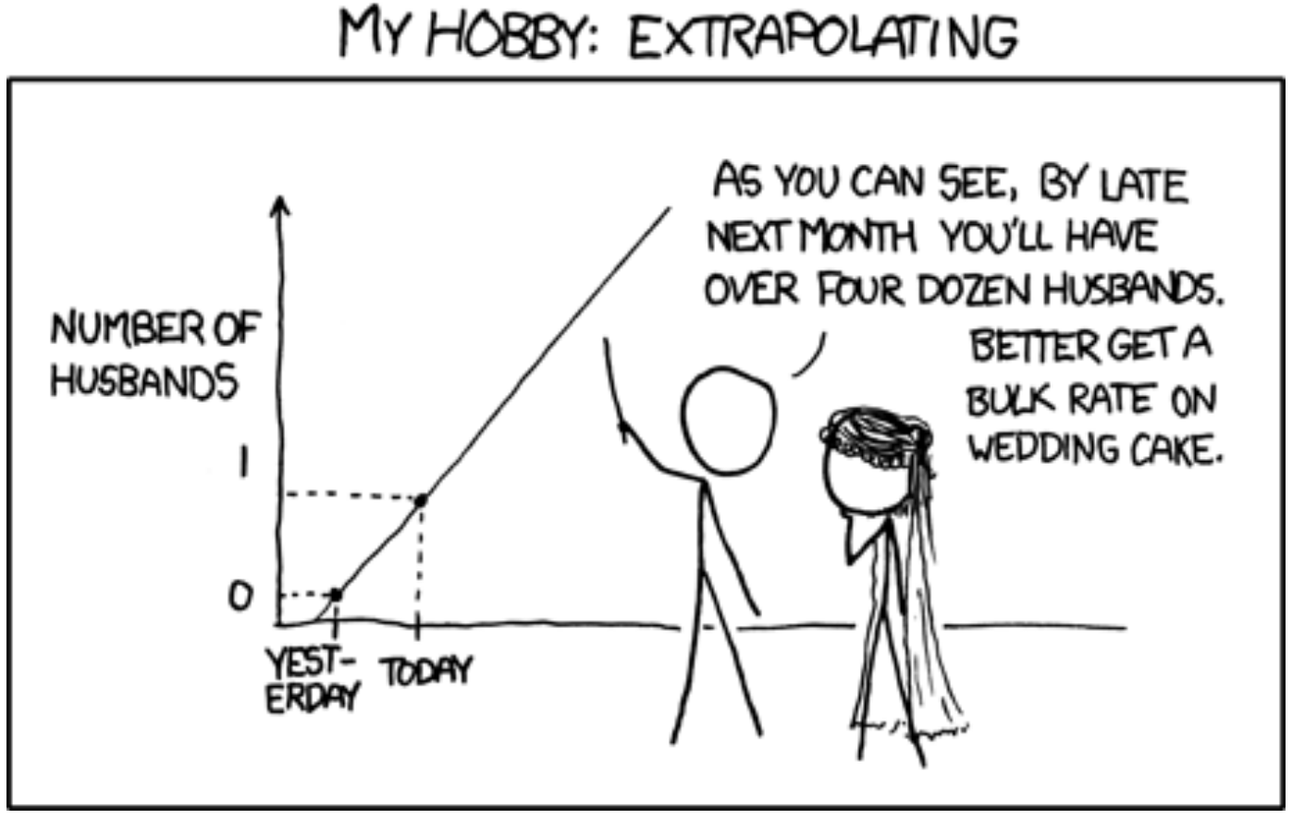
VISUALISE THE DATA/EDA

- Is there a clear trend in the data?
- Are there regular periods/seasonality we observe?
- Does time seem explanatory?



VISUALISE THE DATA/EDA

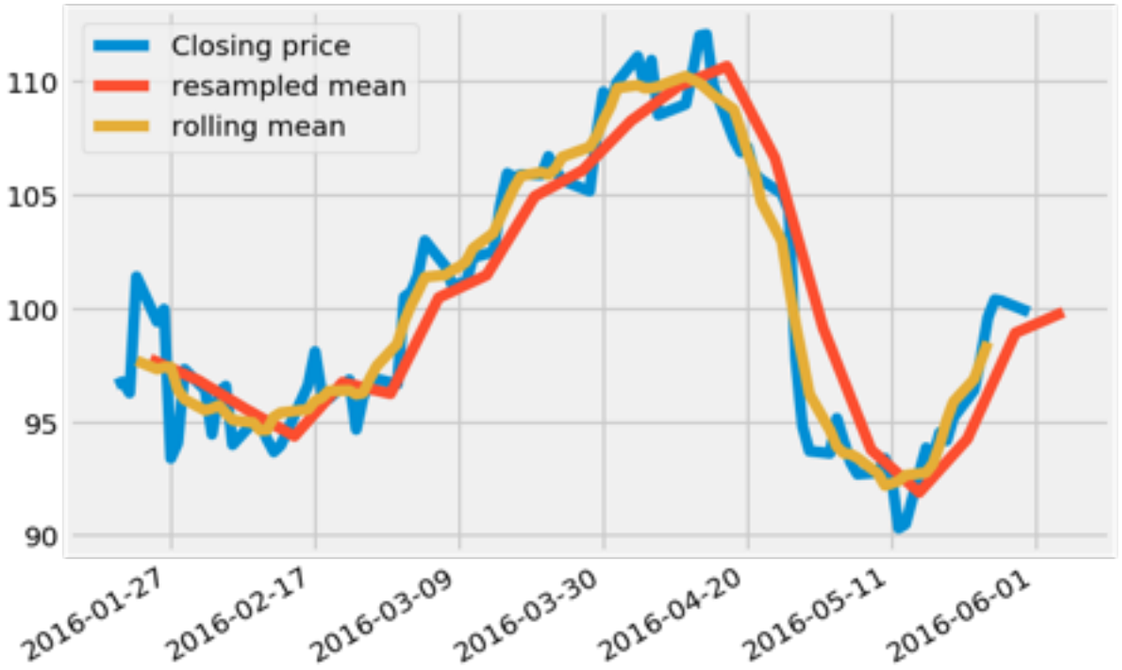
- Is there a clear trend in the data?
- Are there regular periods/seasonality we observe?
- Does time seem explanatory?



HOW ARE WE GOING TO USE THE DATA?

- Look at the individual data points recorded at different moments of time
- Consider how quantities change over time
- Aggregate the data (for example instead of looking at hourly data we consider the daily or monthly aggregates)
- Transform the data before building the model, e.g. taking the logarithm
- Model the differences between subsequent observations

AGGREGATED TIME SERIES

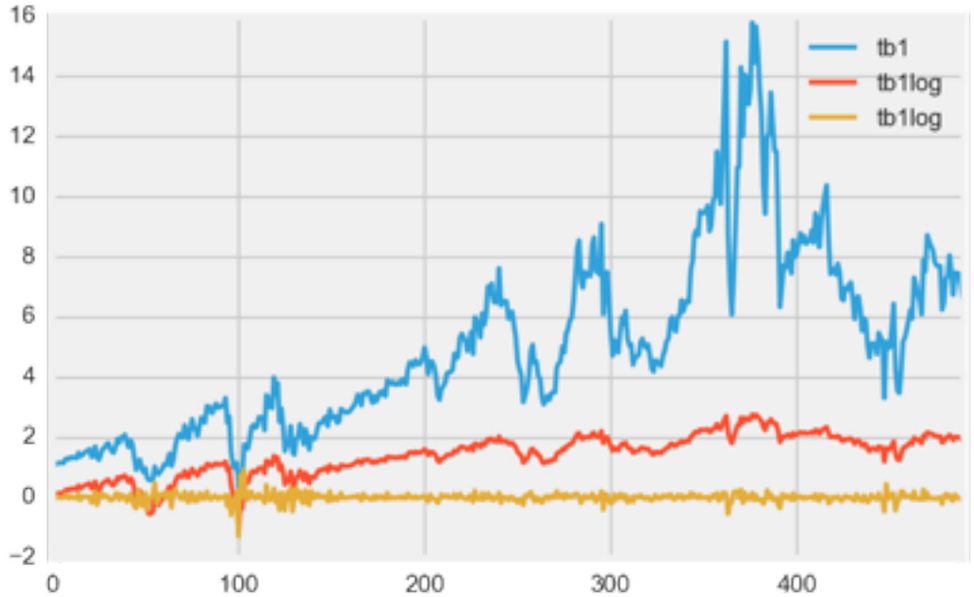


Values, weekly resampled mean values and weekly rolling mean values of Apple stock daily closing prices

TRANSFORMING AND DIFFERENCING

Below, American T-bills data are displayed
(one-month T-bill rate in percent annual rate from R's Ecdat package (Mishkin))

- Blue: returns
- Red: log returns
- Yellow: differenced log returns

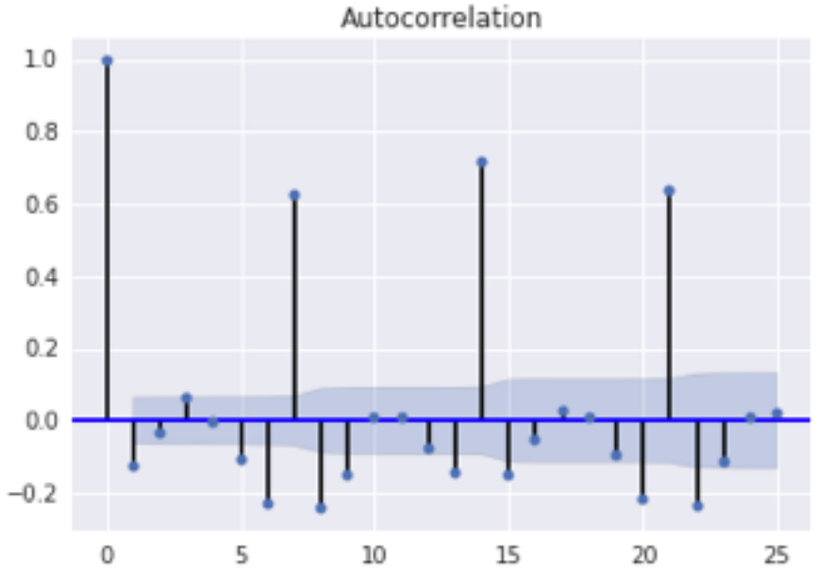


AUTOCORRELATION

- In regression models, we look at correlations between outcome and predictor variables
- For time series, we will look at correlations between observations at different points in time
- Doing so will inform us about which time series model to use

AUTOCORRELATION

$$\mathrm{acf}(k) = \frac{\sum_{t=k+1}^{n} \left(x_t - \bar{x} \right) \left(x_{t-k} - \bar{x} \right)}{\sum_{t=1}^{n} \left(x_t - \bar{x} \right)^2}$$



ACF for the sales data of the German drug store chain Rossmann

MODELLING TIME SERIES

- We can develop a number of simple models
- These models are similar to regression models
- We will model a point in the time series as a function of previous points
- The particular way we do this will give us some model coefficients

$$Y_t = f(Y_{t-1}, Y_{t-2}, \ldots, Y_{t-n}) + \epsilon_t$$

TIME SERIES MODEL SELECTION

- What kind of model we should choose can be guided by looking at the autocorrelations
- We can evaluate our fit using MSE and R2
- We have to balance between making an accurate fit and model complexity
- Evaluation metrics used are Akaike and Bayesian Information Criterion (AIC/BIC)
- Once having fitted a model we can predict future data points

TIME SERIES MODELS

- Autoregressive models (AR)
- Moving average models (MA)
- Autoregressive integrated moving average models (ARIMA)
- Seasonal ARIMA models (SARIMA)
- Generalised Autoregressive models with conditional heteroscedasticity (GARCH)

OVERVIEW: TIME SERIES FORECASTING STEP-BY-STEP

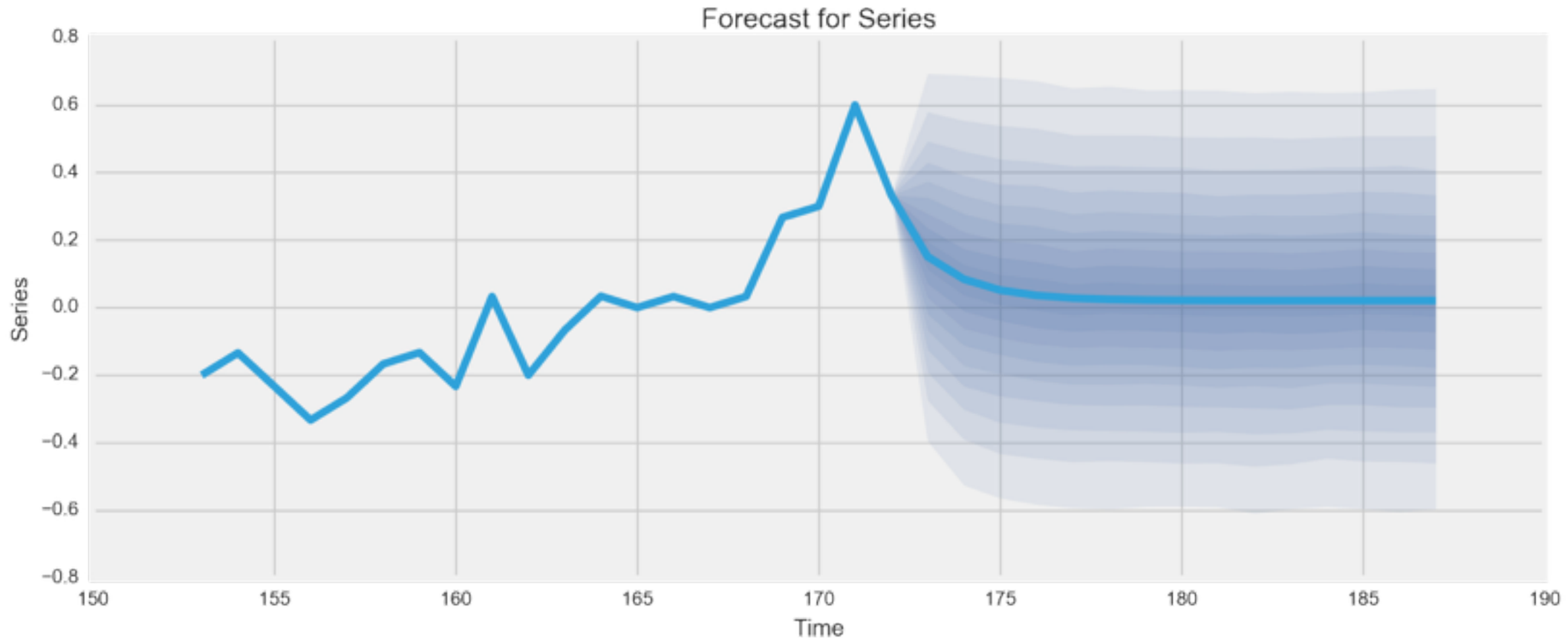
1. Visualise the time series
2. Make the time series stationary
3. Plot ACF/PACF to seek optimal parameters
4. Build the ARIMA model
5. Predict the future

OVERVIEW: TIME SERIES FORECASTING STEP-BY-STEP

1. Visualise the time series
2. Make the time series stationary
3. Plot ACF/PACF to seek optimal parameters
4. Build the ARIMA model
5. Predict the future
6. Profit*

*Lecture sold separately

FORECASTING WITH OUR MODEL



CODE EXAMPLES

- To the repo…