Time series and forecasting

Filippo Biscarini

(Biostatistician, bioinformatician and quantitative geneticist) **CNR**, Milan (Italy)

The lure of **forecasting**

- Haruspex: forecasters would foretell the future based on the distribution of maggots in a rotten sheep's liver

- We progressed a little from that …

Bronze model of a sheep's liver found with Etruscan inscriptions [from Wikipedia.org]

The lure of **forecasting**

- I think there is a world market for maybe five computers (1943)

- Computers in the future may weigh no more than 1.5 tons (1949)

- There is no reason anyone would want a computer in their home (1977)

- Millennium bug: the world will go kablooey at midnight on January 1, 2000

- The horse is here to stay but the automobile is only a temporary novelty (1903)

- If excessive smoking actually plays a role in the production of lung cancer, it seems to be a minor one (1954)

Forecasting: what?

- you may have caught a hint that forecasting is about <u>predicting the future</u>

- forecast well ahead in the future (investments, policies) or in a few seconds range (computer scheduler)

- perfect forecasting: deterministic processes (e.g. time at sunrise on any day; Easter sunday)

- impossible forecasting: e.g. lottery numbers

Forecasting: what?

Forecasting ability depends on:

1. how well we understand the factors behind the phenomen to forecast
2. how much data are available
3. whether the forecasts can affect the thing we are trying to forecast (e.g. self-fulfilling prophecies)

- short-term weather forecasts
- short-term stock prices forecasts (up or down? Toss a coin!)

Forecasting: how

- usually account for changing environment (remember, it's about the future!)
- we assumed that the way in which the environment is changing will continue into the future (e.g. volatile business env will remain volatile; cyclic climate will remain cyclic; etc.)
- many different methods: from simple and naive forecasting methods, to neural networks and systems of equations

Forecasting: time series data

We talked about predicting the future, more specifically, in forecasting we are dealing with **time-series data**
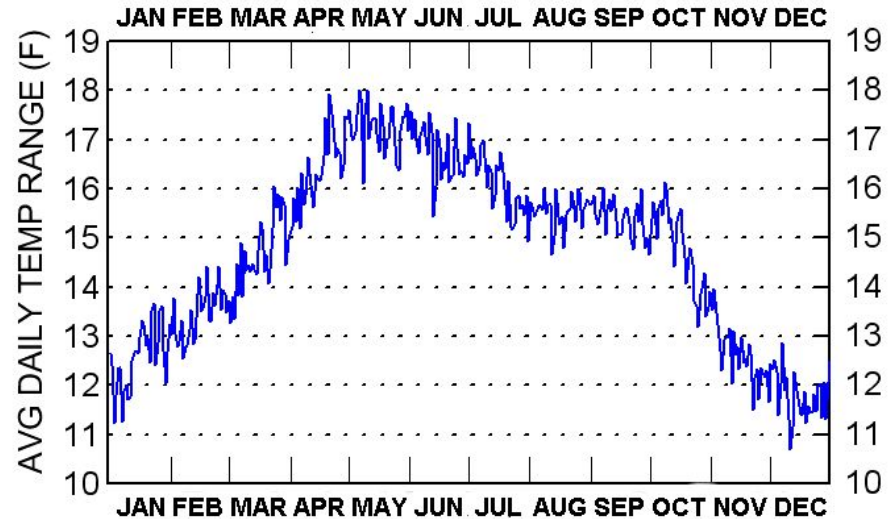
- stock prices

Forecasting: time series data

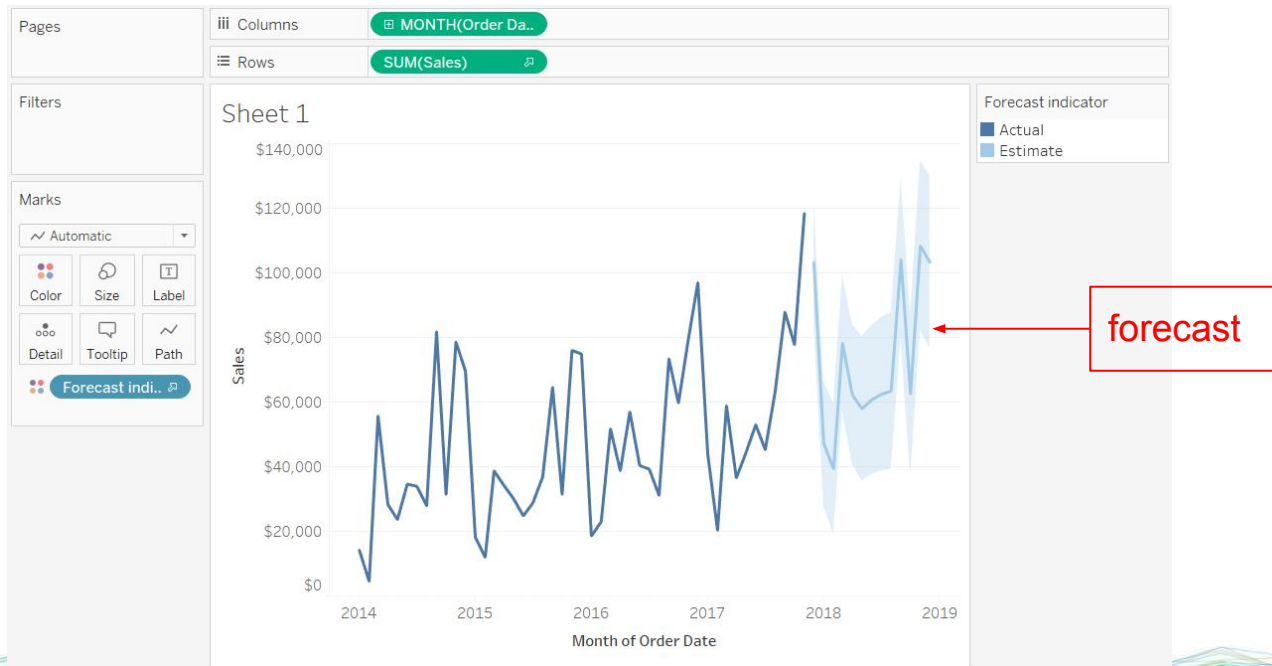We talked about predicting the future, more specifically, in forecasting we are dealing with **time-series data**

- stock prices
- monthly temperatures
- etc.

NY CENTRAL PARK AVG DAILY TEMPERATURE RANGES
- 1876-2015 PERIOD of RECORD -

Forecasting: time series data

We talked about predicting the future, more specifically, in forecasting we are dealing with **time-series data**

forecast

Forecasting: time series data

In time series forecasting we use only information on the variable to be forecast, and make no attempt to discover the factors (predictors) that affect its behaviour

- naive forecasting
- exponential smoothing
- ARIMA/SARIMA
- one-step-ahead forecast

Forecasting: basic steps

1. Problem definition: what needs to be forecast and how
2. Gathering data: historical data, old/recent data
3. EDA: are there consistent patterns? Is there a significant trend? Is seasonality important? Is there evidence of the presence of cycles? Are there outliers in the data that need to be explained? Autocorrelation?
4. Choosing and fitting models: it's common to compare two or more models. Each model is based on a set of assumptions (explicit and implicit) and may involve parameters to be estimated
5. Evaluating the forecasting model: the performance of the model must always be measured, with cross-validation/test sets, and then also after the data for the forecast period have become available (once the future has happened)

Forecasting: final remarks

Total international visitors to Australia
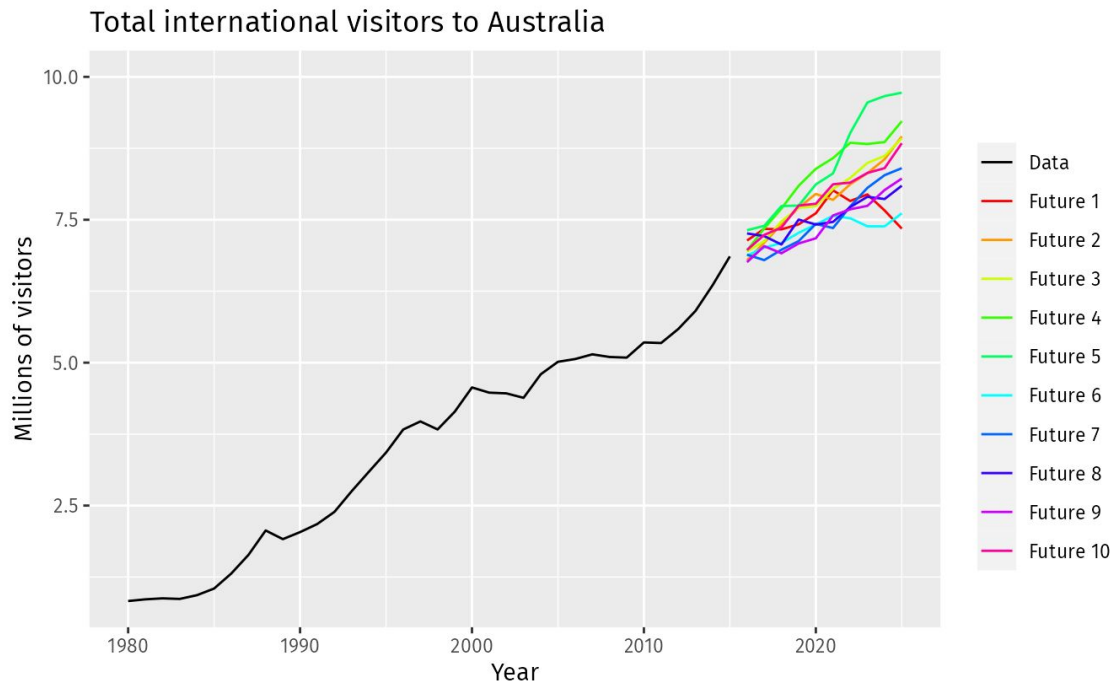
- The quantity to forecast is unknown → random variable that can take up many possible values → many possible futures

- Futures diverge with the forecasting range

- Average or middle scenario +/- prediction interval (e.g. 95% interval → will include the actual future value with probability 0.95)