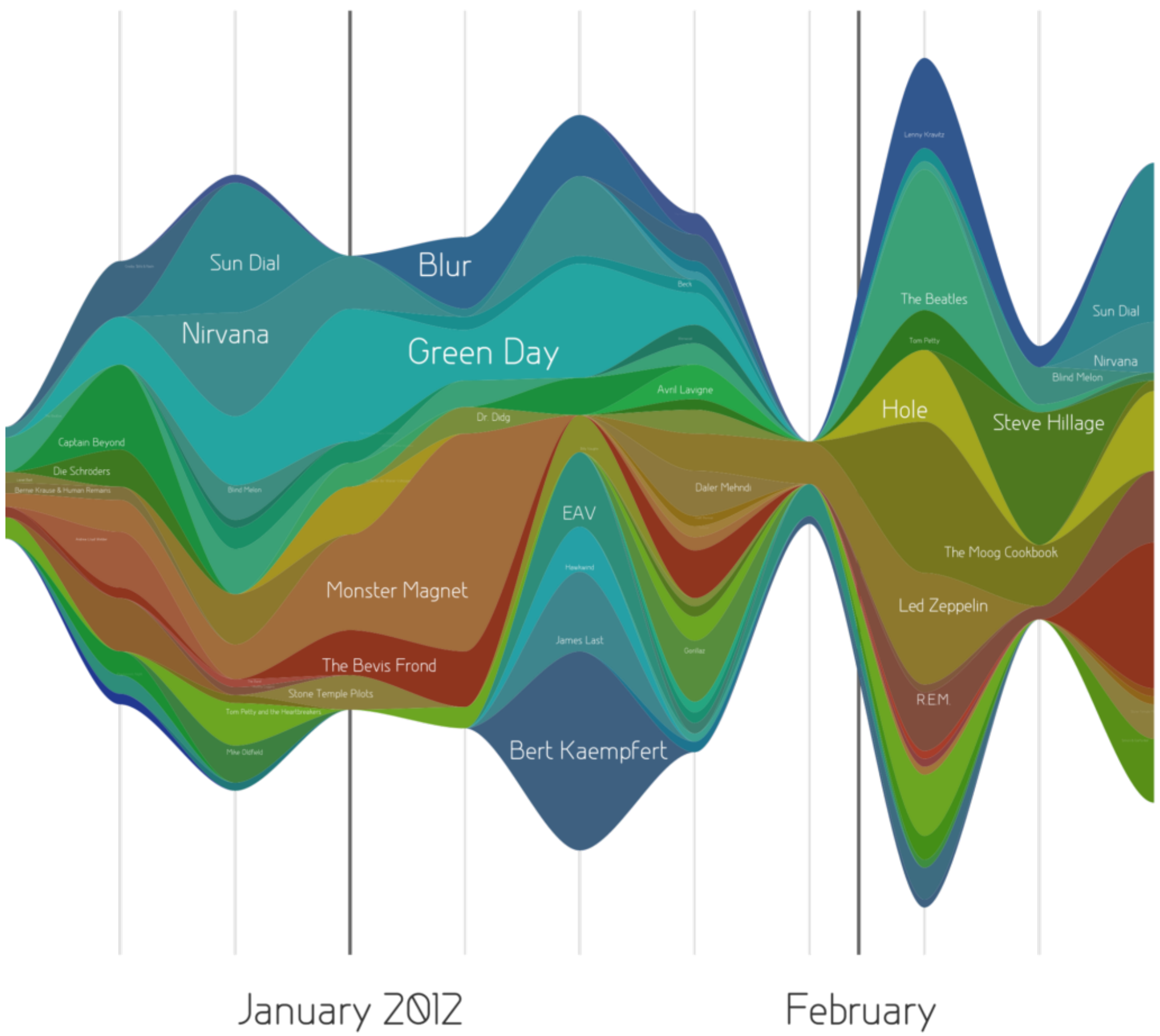
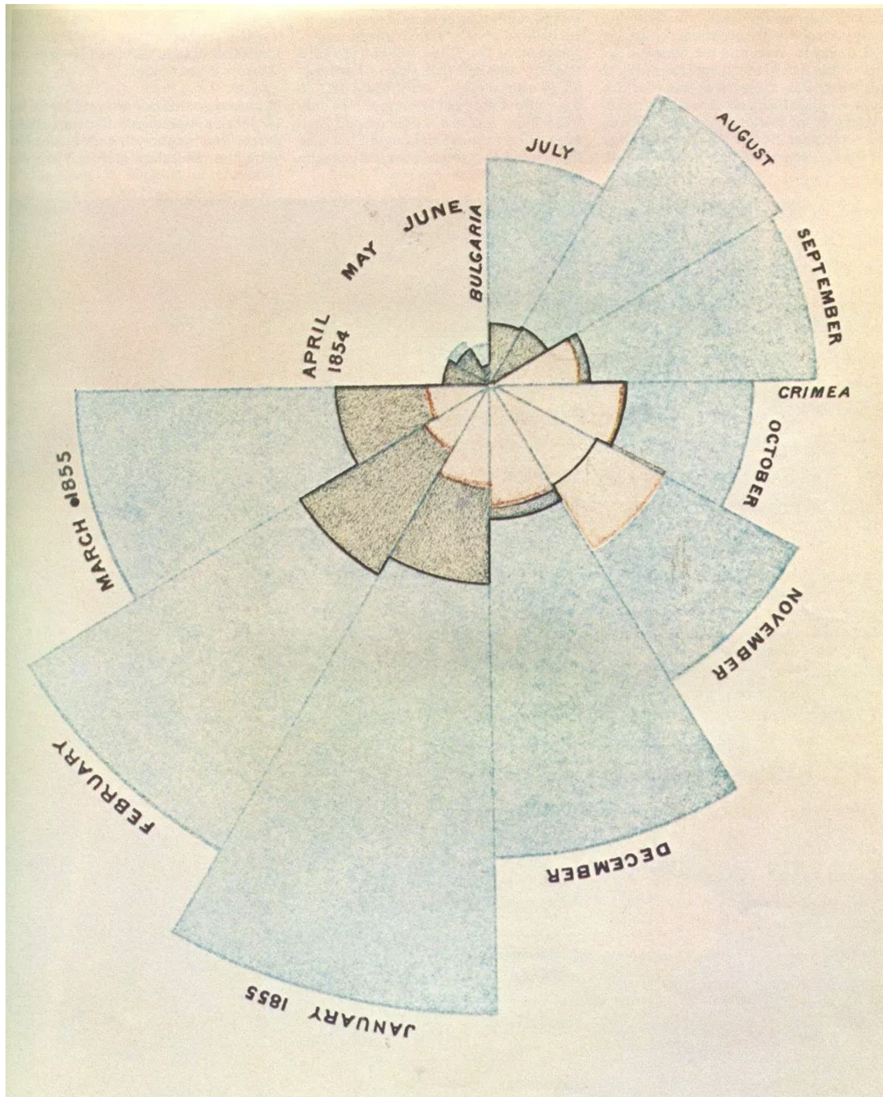
Time Series Forecasting — ARIMA Models

Visualize the data

- Visualize the time-series data — is there any overall trend? Are there observable seasonal trends?
- Other patterns — e.g. is there “too” much granularity in the data — aggregating sales by day or week may be better than looking at hourly sales data?
- Explore interesting visualization schemes that may be appropriate to the task at hand.

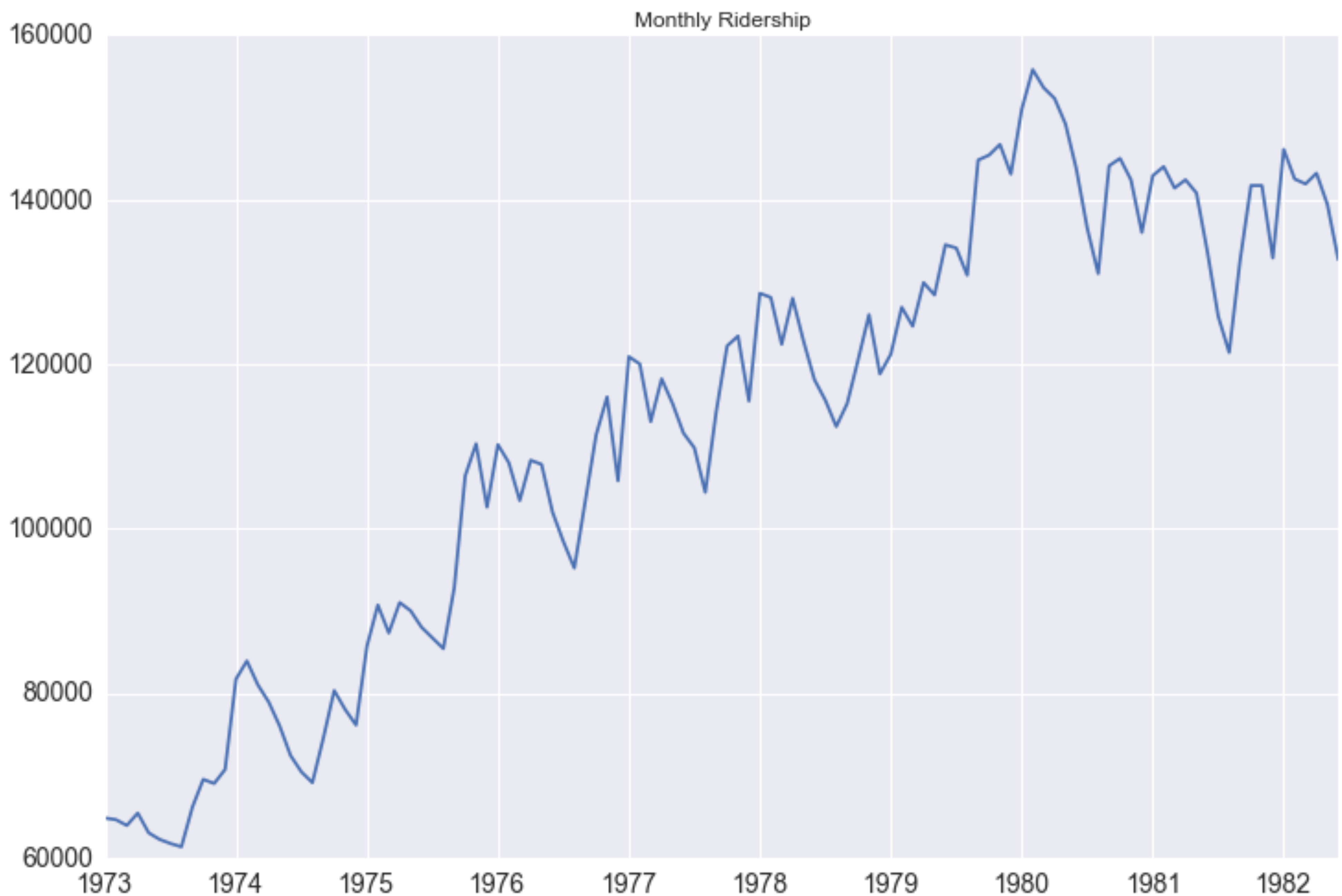


Stream graphs



Polar Plots

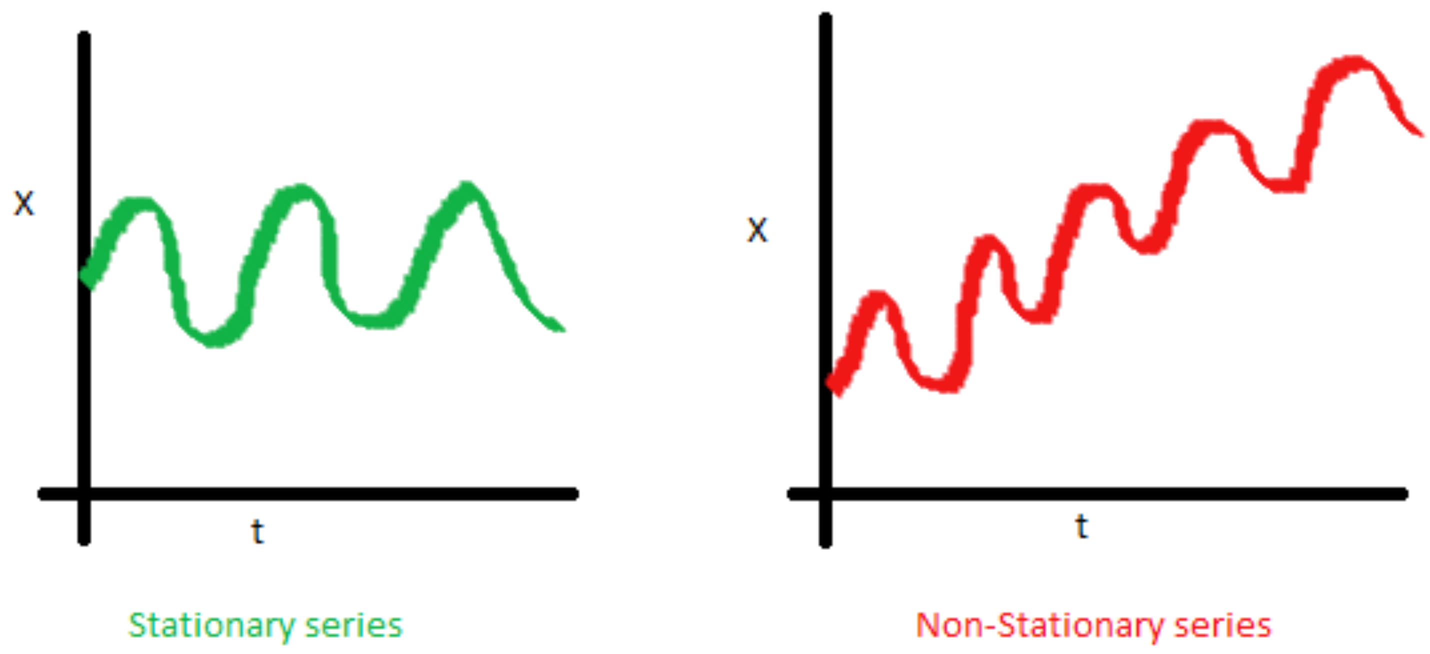
Visualize the data



Seasonality and Trend

Stationarizing the data

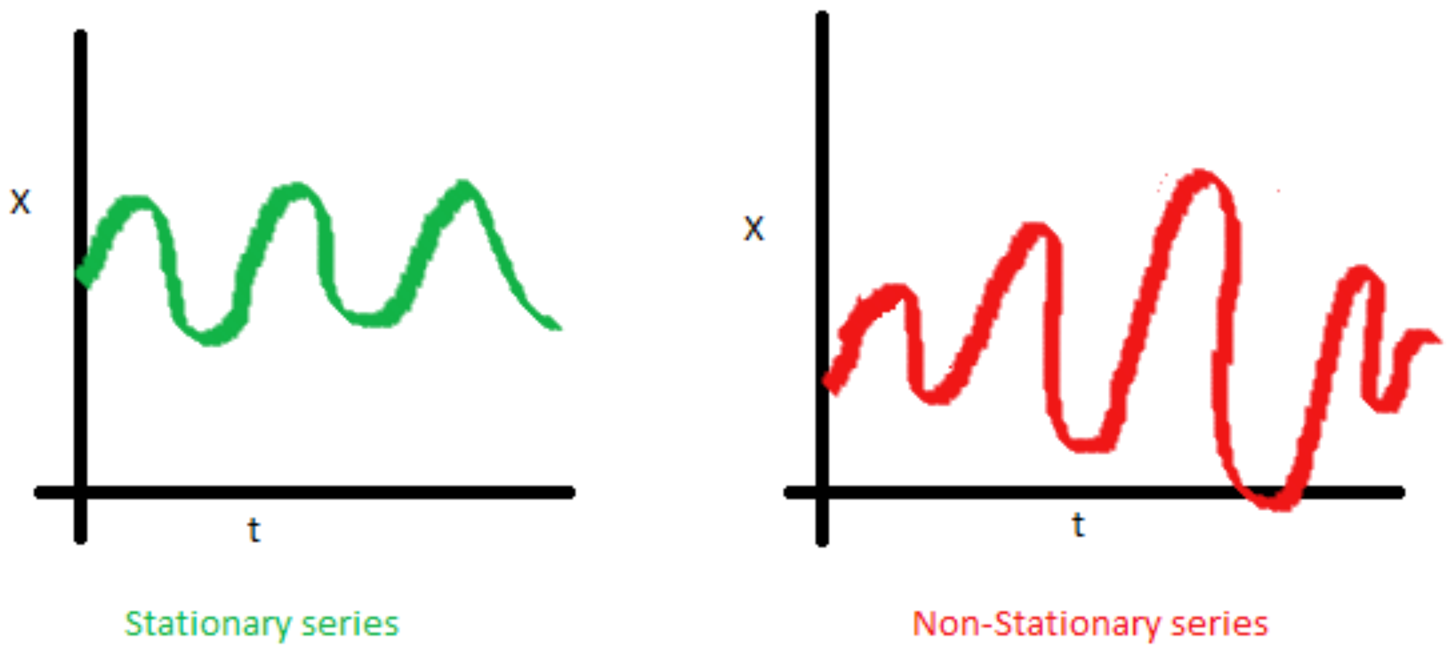
- The first step is to stationarize the time-series data.
- What is stationarity?



i. Mean is not a function of time

Stationarizing the data

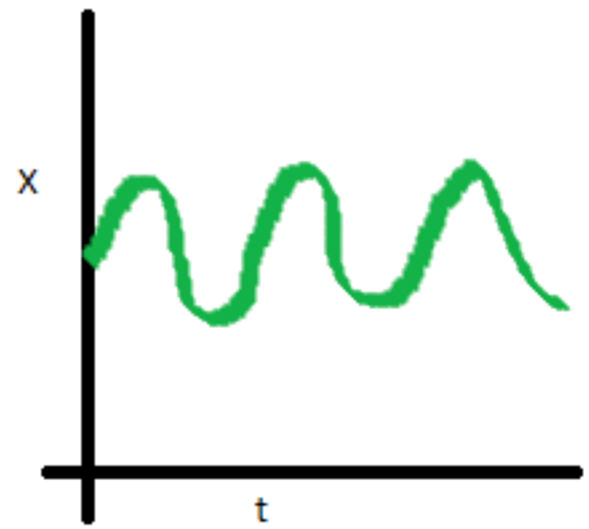
- What is stationarity?



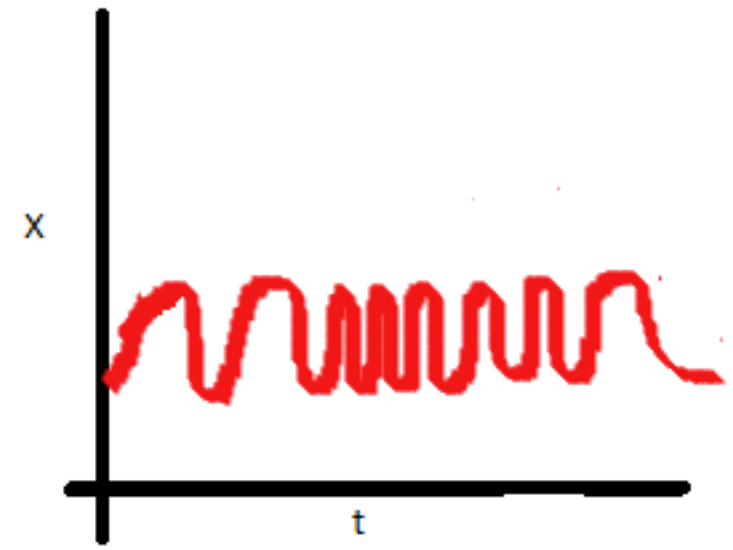
ii. Variance is not a function of time (no heteroskedasticity)

Stationarizing the data

- What is stationarity?



Stationary series



Non-Stationary series

iii. Autocorrelation is invariant to time

Stationarizing the data

- How to stationarize the data? — First or second differencing, logarithmic transform, seasonal difference.

- This is analagous to applying suitable transformations in order to satisfy the assumptions of a linear regression model (residuals should be normally i.i.d and homoscedastic).

- How to check for stationarity? — Visualization and the Dickey-Fuller Test*

- Also, remove seasonalities.

- A stationarized time series has no trend, a constant variance and a constant auto-correlation.

---

*Read up what it is.

ARIMA Model

Differencing to achieve stationarity:

If d=0: $y_t = Y_t$

If d=1: $y_t = Y_t - Y_{t-1}$

If d=2: $y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2}$

Autoregressive Integrated Moving Average Model:

$$\hat{y}_t = \mu + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \ldots - \theta_q e_{t-q}$$

ARIMA models

$$\hat{y}_t = \mu + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \ldots - \theta_q e_{t-q}$$

- A nonseasonal ARIMA model is classified as an "ARIMA(p,d,q)" model, where:

  p is the number of autoregressive terms

  d is the number of nonseasonal differences needed for stationarity, and

  q is the number of lagged forecast errors in the prediction equation.

Forecast for *y* at time *t* = constant + weighted sum of the last *p* values of *y* + weighted sum of the last *q* forecast errors

The forecast for Y is then obtained by undoing the transformations performed to achieve stationarity.

Estimation of ARIMA

$$\boxed{\hat{y}_t = \mu + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \ldots - \theta_q e_{t-q}}$$

- The model parameters can be estimated using a maximum likelihood estimation procedure.

- Assume that the time series is correctly specified with model order (p,d,q) known.

- Possible to derive analytically the MLE estimators for an AR(1) model.

- Alternate estimation procedure: Yule-Walker method.

Building an ARIMA model

$$\boxed{\hat{y}_t = \mu + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \ldots - \theta_q e_{t-q}}$$

ARIMA model

- We have three decisions to make —

  a. How many orders to difference (d)?
  b. How many AR terms to incorporate (p)?
  c. How many MA terms to incorporate (q)?

“Rules” for differencing

$$\hat{y}_t = \mu + \phi_1 y_{t-1} + \ldots + \phi_p y_{t-p} - \theta_1 e_{t-1} - \ldots - \theta_q e_{t-q}$$

- **Rule 1:** If the series has positive autocorrelations out to a high number of lags, then it probably needs a higher order of differencing.

- **Rule 2:** If the lag-1 autocorrelation is zero or negative, or the autocorrelations are all small and patternless, then the series does not need a higher order of differencing. If the lag-1 autocorrelation is -0.5 or more negative, the series may be overdifferenced.

- **Rule 3:** The optimal order of differencing is often the order of differencing at which the standard deviation is lowest.

- **Rule 4:** A model with no orders of differencing assumes that the original series is stationary (mean-reverting). A model with one order of differencing assumes that the original series has a constant average trend (e.g. a random walk or SES-type model, with or without growth).

- **Rule 5:** A model with no orders of differencing normally includes a constant term (which allows for a non-zero mean value). A model with two orders of total differencing normally does not include a constant term. In a model with one order of total differencing, a constant term should be included if the series has a non-zero average trend.

Model Specification

- The next step is to determine the tuning parameters of the model, p and q.
- We do not want to incorporate too many terms — overfitting.
- We choose these parameters by looking at the autocorrelation and partial autocorrelation graphs and some qualitative judgement*.
- The partial-autocorrelation between y at time t and y at lag k is the coefficient of the k^{th} lag in a regression that includes all lower lags.
- Thus, it is the correlation that is not explained by lower lags.

---

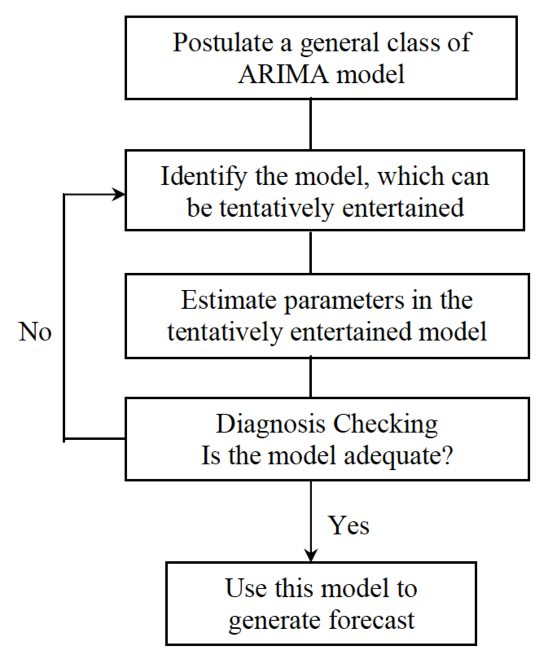
* http://people.duke.edu/~rnau/411arim3.htm

Model Specification

If the ACF plot cuts off sharply at lag k but the PACF plot decays more gradually, then set *q=k* and *p=0*. This is a so-called "*MA(q) signature*."

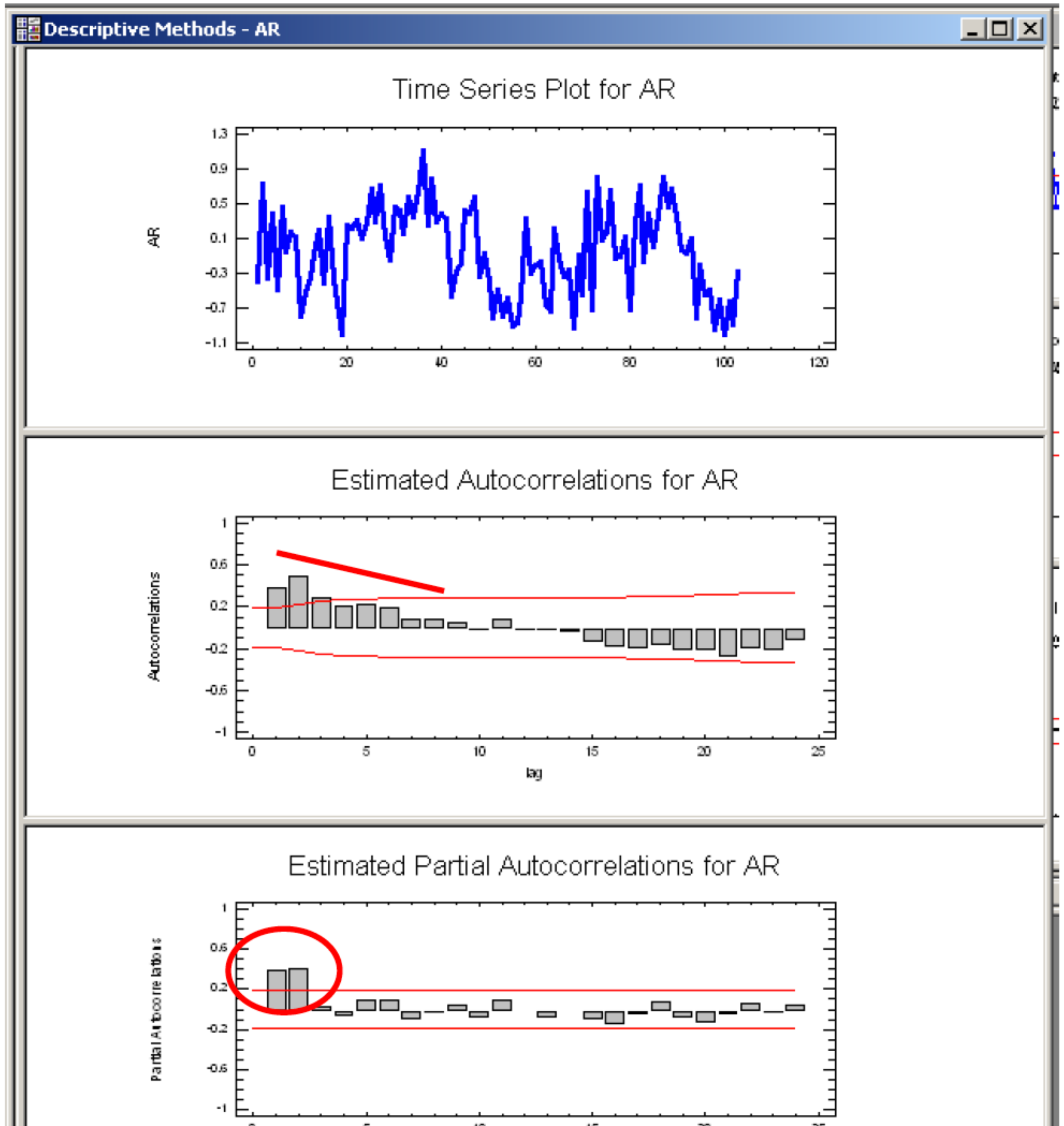
If the PACF plot cuts off sharply at lag *k* while the ACF plot decays more gradually, then set *p=k* and *q*=0. This is a so-called "*AR(p) signature.*"

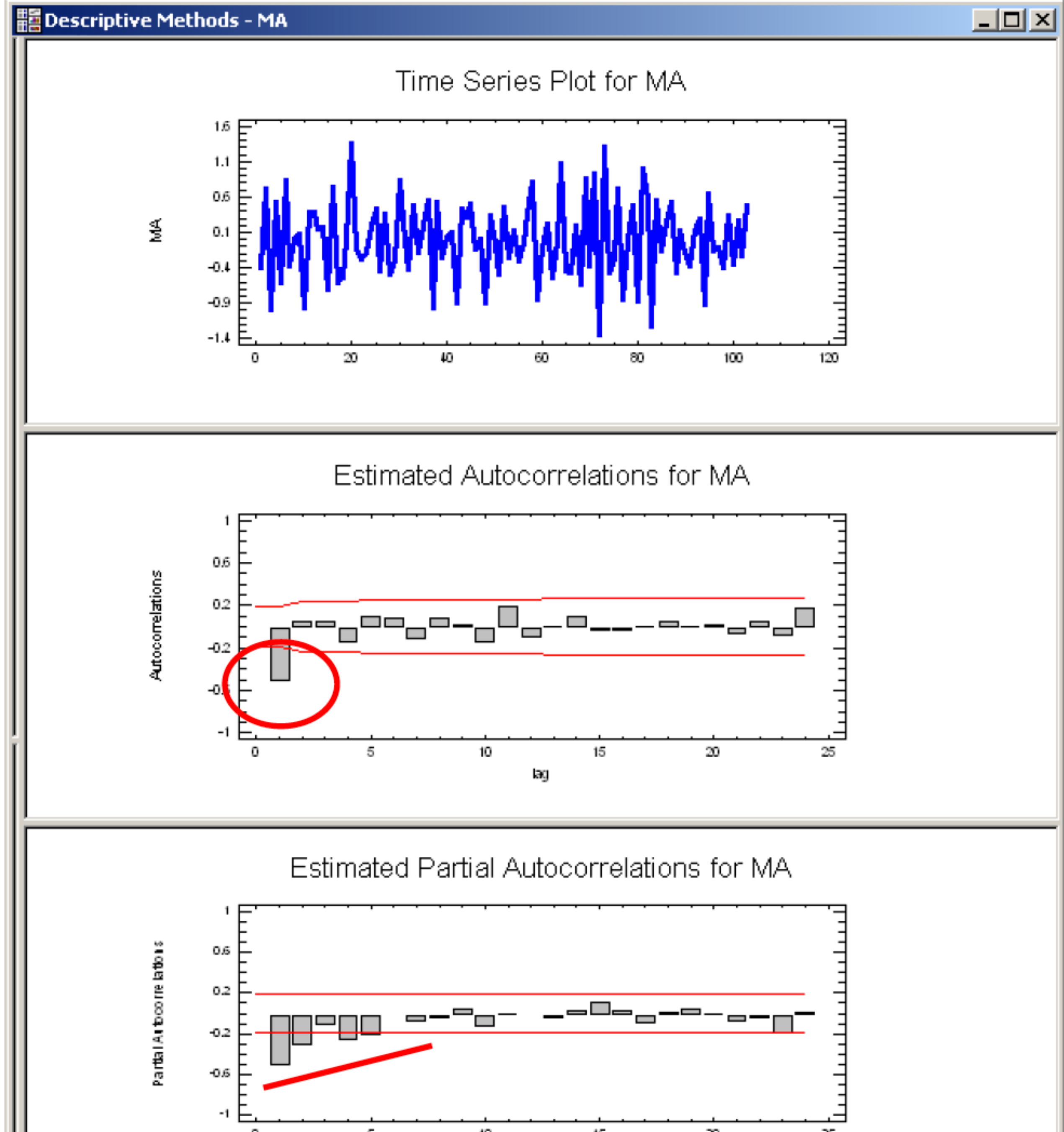
Use the AIC criterion to select between candidate ARIMA models: AIC = $2(p+q) - 2\ln(L)$



Box and Jenkins

* http://people.duke.edu/~rnau/411arim3.htm



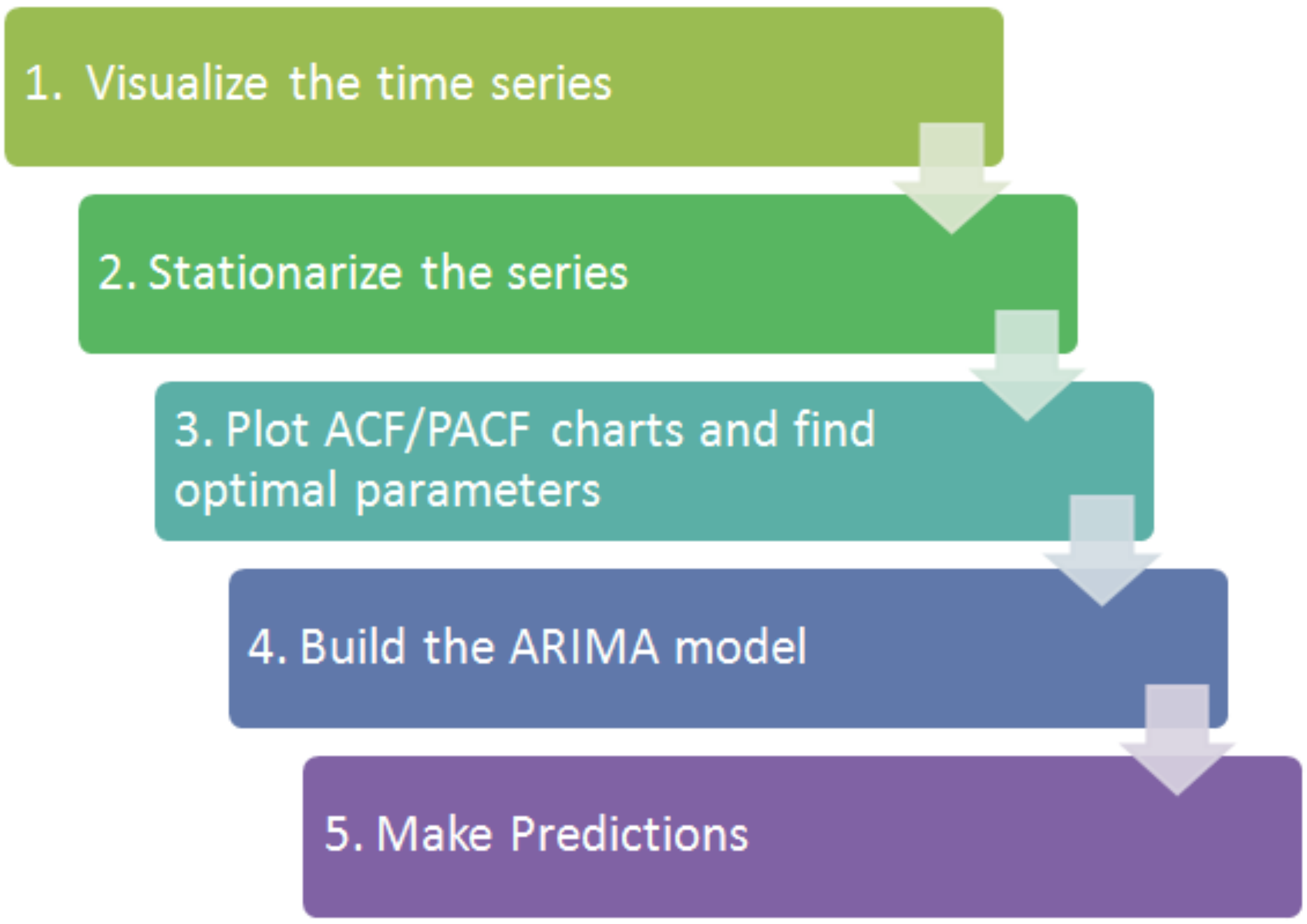


AR(2) signature: spikes at lags 1 and 2 in PACF (the first of which is usually positive), and gradual decay in ACF (usually from above)

MA(1) signature: 1 spike in ACF (which is usually negative) and gradual decay in PACF (usually from below)

* http://people.duke.edu/~rnau/411arim3.htm

ARIMA Models



An ARIMA exercise

- Download dataset from here* (Portland bus ridership data) or you can use the Translink dataset (preferred)

- Build a forecasting model (using ARIMA)†.

---

* https://datamarket.com/data/set/22w6/portland-oregon-average-monthly-bus-ridership-100-january-1973-through-june-1982-n114

* https://github.com/seanabu/seanabu.github.io/blob/master/Seasonal_ARIMA_model_Portland_transit.ipynb

References

1. Notes on Nonseasonal ARIMA models. Robert Nau. Fuqua School of Business, Duke University

2. Time Series Analysis: Forecasting and Control. Box, Jenkins and Reinsel.

3. Estimation of ARIMA Models. Florian Pelgrin, EDHEC Business School.