

Zeitreihen und ihre Merkmale visualisieren

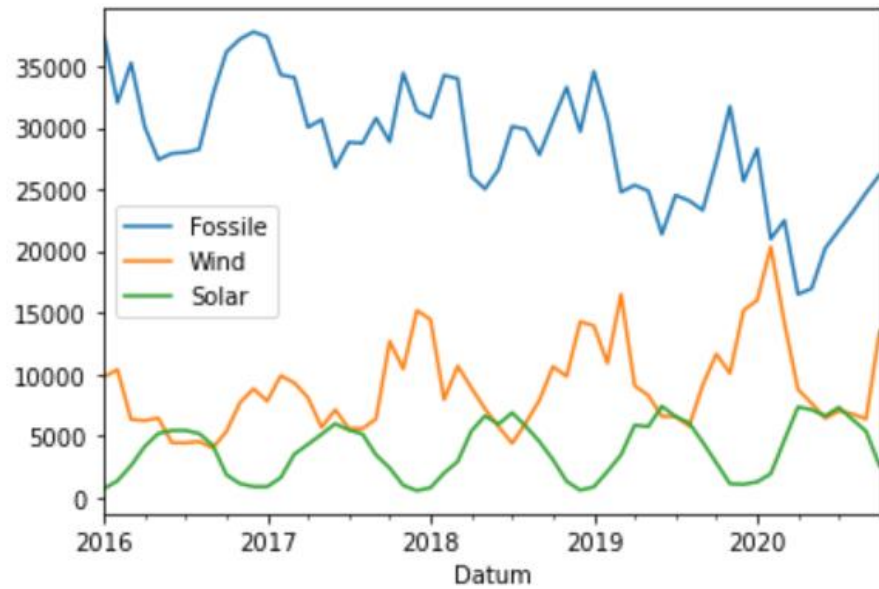
Session 2 (Montag 11:00 – 12:30)



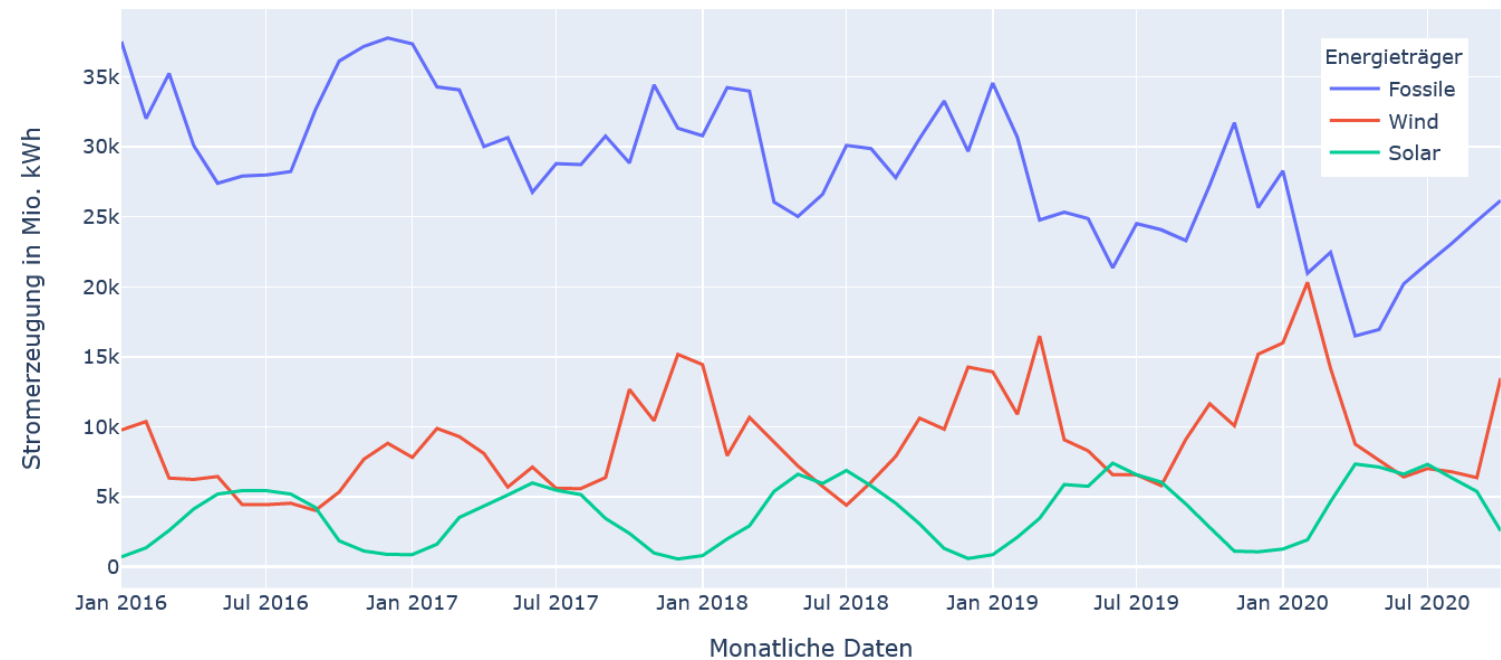
Zeitreihen und ihre Merkmale visualisieren

- Zeitreihen plotten, Plots anpassen
- Autokorrelation, partielle Autokorrelation, gleitende Mittel berechnen und visualisieren
- Komponenten von Zeitreihen (Trend, Saisonalität) berechnen und visualisieren

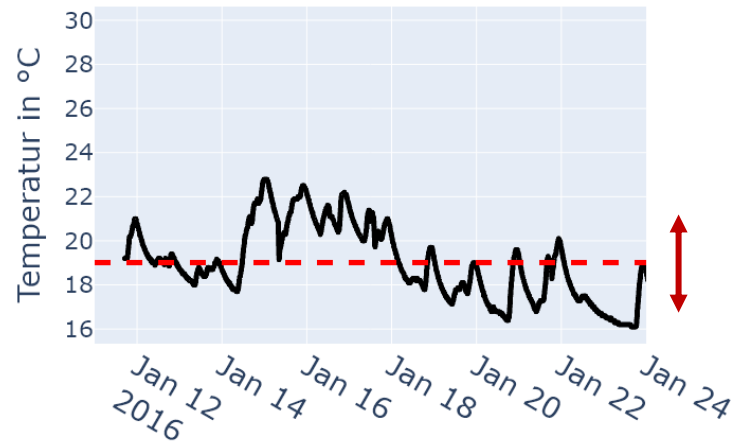
Zeitreihen visualisieren



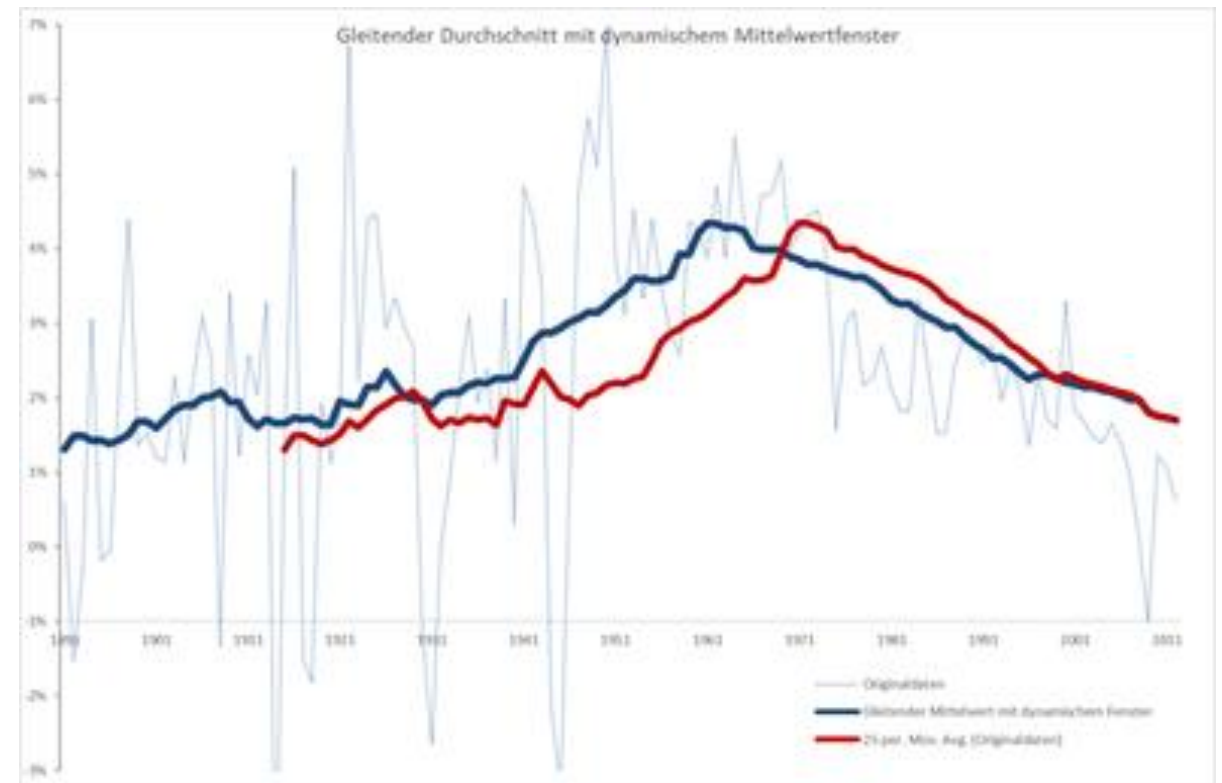
Stromerzeugung in Deutschland nach Energieträger



Gleitender Mittelwert



**Mittelwert hängt
von der Zeit ab!**

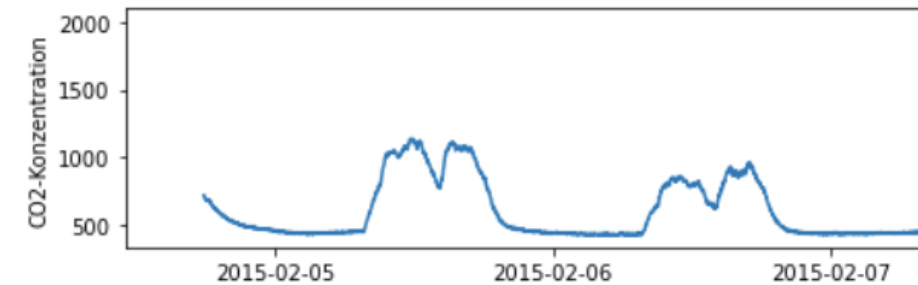


Autokorrelation

Beispiel

- (A) Ich messe heute Mittag die CO²-Konzentration in 72 Büros in München (**keine Zeitreihe**)
- (B) Ich messe 72 Stunden lang die CO²-Konzentration in meinem München (**Zeitreihe**)

- Erinnerung: Bei **Zeitreihen**:
- Beobachtung zu Zeitpunkt t : *könnte etwas zu tun haben* mit Beobachtungen zu anderem Zeitpunkt $t + h$
- Deshalb: **Kovarianz** der Zeitreihe mit sich selbst (zeitverzögert) wichtig!



Stochastischer Prozess $X_1, X_2, X_3, \dots, X_n$

Daten $x_1, x_2, x_3, \dots, x_n$

Kovarianz

$$\text{Cov}(X_{t+h}, X_t) = \mathbb{E}[(X_{t+h} - \mu_{t+h})(X_t - \mu_t)] = \gamma_t(h)$$

Stichprobenkovarianz

$$\hat{\gamma}(h) = \frac{1}{n} \sum_{t=1}^{n-|h|} (x_{t+|h|} - \bar{x})(x_t - \bar{x})$$

Die **Autokorrelationsfunktion ACF** (engl. Auto Correlation Function) ist definiert als

$$\rho_t(h) = \frac{\gamma_t(h)}{\gamma_t(0)}$$

ACF Plot

- Die ACF-Darstellung ist ein Balkendiagramm der Korrelationskoeffizienten zwischen einer Zeitreihe und ihren verzögerten Werten.
- Einfach ausgedrückt: ACF erklärt, wie der aktuelle Wert einer gegebenen Zeitreihe mit den vergangenen Werten (1 Einheit Vergangenheit, 2 Einheiten Vergangenheit, ..., n Einheiten Vergangenheit) korreliert ist.
- In der ACF-Darstellung drückt die y-Achse den Korrelationskoeffizienten aus, während die x-Achse die Anzahl der Verzögerungen angibt.

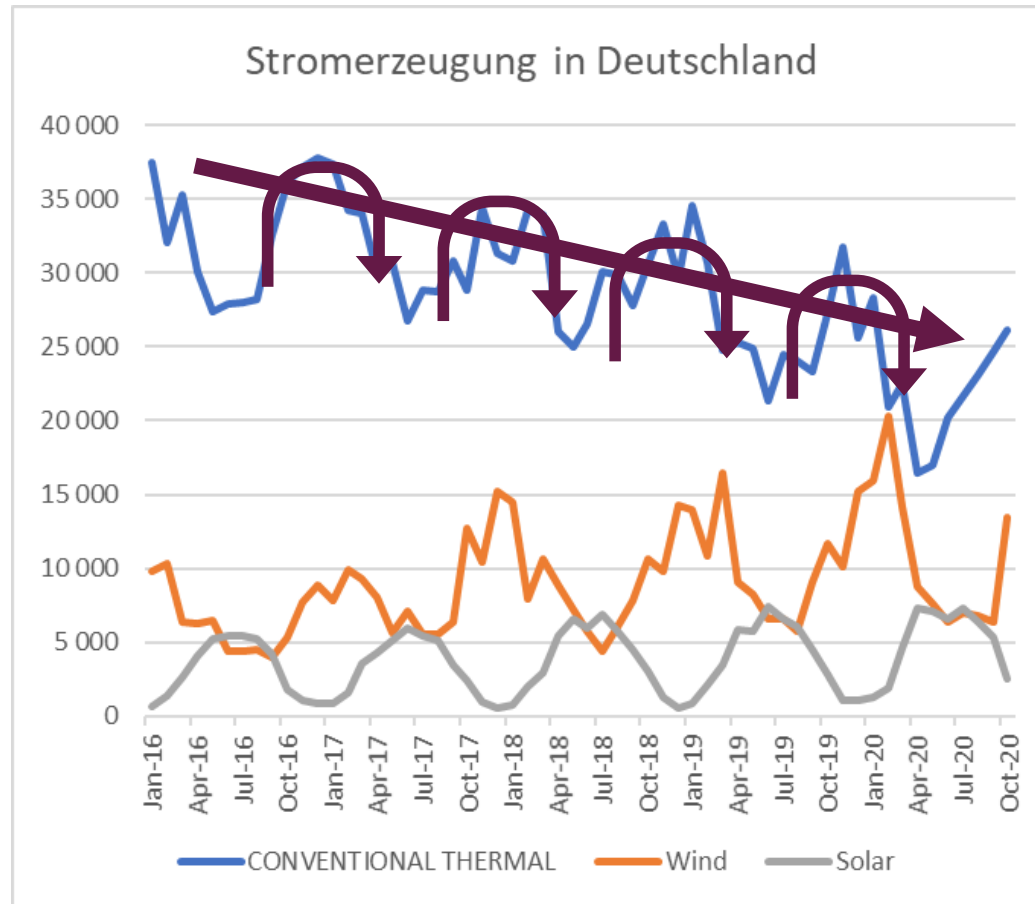


PACF Plot

- PACF ist die partielle Autokorrelationsfunktion, die die partielle Korrelation zwischen der Reihe und den Lags selbst erklärt.
- Vereinfacht ausgedrückt lässt sich die PACF durch eine lineare Regression erklären, bei der wir $y(t)$ aus $y(t-1)$, $y(t-2)$ und $y(t-3)$ vorhersagen [2].
- Bei PACF werden die "Teile" von $y(t)$ und $y(t-3)$ korreliert, die nicht durch $y(t-1)$ und $y(t-2)$ vorhergesagt werden.



Komponenten von Zeitreihen

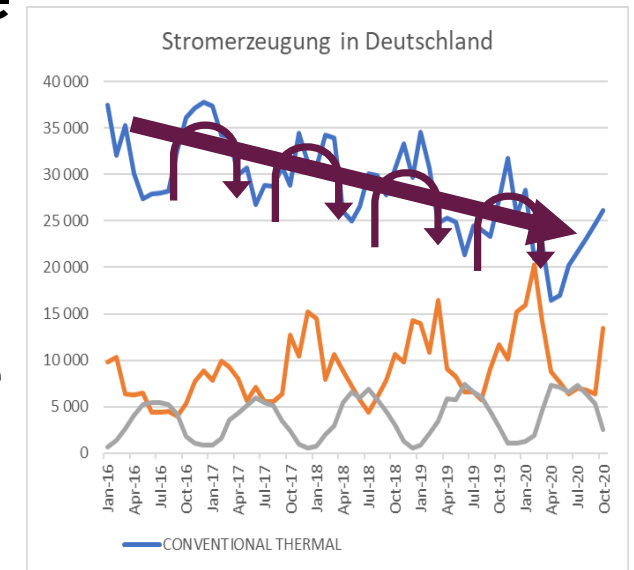


Beispiel

Sie beschreiben die Entwicklung der Stromproduktion aus fossilen Energieträgern in Deutschland ab 2016

Komponenten von Zeitreihen

- **Trend:** Ein Trend liegt vor, wenn es eine langfristige Zunahme oder Abnahme der Daten gibt. Er muss nicht linear sein.
- **Saisonalität:** Ein saisonales Muster tritt auf, wenn eine Zeitreihe von saisonalen Faktoren wie der Jahreszeit oder dem Wochentag beeinflusst wird. Die Saisonalität hat immer eine feste und bekannte Häufigkeit.
- **Additive Zerlegung:** $x_t = T_t + S_t + R_t$
- **Multiplikative Zerlegung** $x_t = T_t \times S_t \times R_t$
- Wo T_t die Trend(-Zyklus)-Komponente, S_t die saisonale Komponente, R_t die Restkomponente



CODING

Stromerzeugung in Deutschland nach Energieträger

- Visualisieren
- Autokorrelation
- Gleitende Mittel
- Komponentenzerlegung

Ab ins Jupyter
Notebook

