

# Saisonbereinigung – gleitender Durchschnitt

als grundlegendes Verfahren zur Glättung von Zeitreihen

---

„... die Zahl der arbeitslosen Menschen [ist] von Dezember auf Januar um 268.000 auf 3.032.000 gestiegen...“<sup>1</sup>

Lässt der Anstieg der Arbeitslosenzahlen von Dezember 2014 auf Januar 2015 den Schluss zu, dass im betrachteten Zeitraum die Probleme am Arbeitsmarkt drastisch zugenommen haben?

## Lernziele:

Im vorliegenden Arbeitsblatt wird am Beispiel der Entwicklung der Arbeitslosenzahlen das Saisonbereinigungsverfahren gleitender Durchschnitt erörtert. Ziel ist es, dass die grundlegende Funktionsweise von Saisonbereinigungsverfahren verstanden wird. Der Leser ist abschließend in der Lage, Zeitreihen mit Hilfe des gleitenden Durchschnitts von saisonalen Effekten zu bereinigen.

<sup>1</sup> Presse Info 005 vom 29.01.2015. URL im Internet:

<https://www.arbeitsagentur.de/web/content/DE/Presse/Presseinformationen/ArbeitsundAusbildungsMarkt/Detail/index.htm?dfContentId=L6019022DSTBAI721188>.

### Komponentenmodell :

Das Komponentenmodell besagt, das sich die Zeitreihe in verschiedene Komponenten zerlegen lässt.

Komponenten sind:

- Trend,
- Konjunktur-,
- Saison- und
- Rest- bzw. irreguläre Komponente

Die Trend und die Konjunkturkomponente werden aus definitorischen Abgrenzungsproblemen als glatte Komponente zusammengefasst.

Das **Berliner Verfahren** (BV) wurde Ende der 1960er Jahre von der Technischen Universität Berlin und dem Deutschen Institut für Wirtschaftsforschung (DIW) entwickelt.

Das Verfahren **Census X-12-ARIMA** (Autoregressive Integrated Moving Average) ist vom U.S. Bureau of the Census entwickelt worden.

Die Statistik der Bundesagentur für Arbeit verwendet zur Bereinigung von saisonalen Einflüssen das **Census X-12-ARIMA** Verfahren.

## Warum ist eine Saisonbereinigung notwendig?

Arbeitsmarktzahlen, aber auch viele andere wirtschaftliche Größen sind stark von saisonalen Effekten beeinflusst. Diese regelmäßig auftretenden Schwankungen überlagern die trendmäßige und konjunkturbedingte Entwicklung. Bei der Interpretation von Zeitreihen müssen die saisonalen Schwankungen berücksichtigt werden, um keine falschen Schlüsse aus der Veränderung auf die trendmäßige oder konjunkturelle Entwicklung zu ziehen. Bei Arbeitsmarktdaten haben alljährliche Effekte eine starke Wirkung auf Beschäftigung und Arbeitslosigkeit. Damit diese Zahlen sinnvoll interpretiert werden können, werden Saisonbereinigungsverfahren angewandt, um die Ursprungsdaten von saisonalen Einflüssen zu bereinigen.

## Was sind Saisonbereinigungsverfahren und wie funktionieren diese?

Saisonbereinigungsverfahren sind Verfahrensweisen zur Identifizierung saisonaler Einflüsse und deren Eliminierung aus den Ursprungswerten. Die zu bereinigende Zeitreihe muss durch ein mathematisches Modell beschrieben sein.

Ein derartiges Modell kann mit Hilfe von Komponentenzerlegungsverfahren in einzelne Komponenten zerlegt werden. Die einzelnen Komponenten wiederum werden mittels mathematischer Verfahren geschätzt. Saisonbereinigte Werte entstehen dadurch, dass die Ursprungswerte einer Zeitreihe von der Saisonkomponente bereinigt werden. Gängige Bereinigungsverfahren, die auf einer Komponentenzerlegung basieren, sind das Berliner Verfahren und das Verfahren Census X-12-ARIMA. Diese Verfahren sind sehr komplex. Zur Veranschaulichung der Grundprinzipien der Saisonbereinigung soll hier ein elementares Bereinigungsverfahren, der gleitende Durchschnitt, vorgestellt werden.

Nachfolgend wird ein Komponentenmodell mit additiver Verknüpfung unterstellt (siehe Arbeitsblatt Zeitreihenanalyse).

$$x_t \text{ für } t = 1, \dots, n:$$

$$x_t = T_t + S_t + R_t$$

mit:

$x_t$	Ursprungswert
$T_t$	Wert der glatten Komponente (Trend und Konjunkturkomponente)
$S_t$	Wert der Saisonkomponente
$R_t$	Wert der Restkomponente

Beim gleitenden Durchschnitt wird die Saison- und Restkomponente ausgeschaltet, indem die glatte Komponente jeweils als Durchschnittswert über den Zeitraum der Periodenlänge dargestellt wird ( $\hat{T}_t$ ). Je größer der Zeitraum bzw. je mehr Werte in die Durchschnittsberechnung eingehen, desto stärker wird die Zeitreihe geglättet. Der einzubeziehende Zeitraum ist durch die Periodenlänge der unterstellten Saisonschwankung bestimmt. Unterstellt wird bei dieser Methode, eine streng periodische Saisonfigur und dass sich die Saisonkomponenten einer Periode auf Null summieren.

### Gleitender Durchschnitt ungerader Ordnung ( $2k + 1$ ):

Der Durchschnittswert wird aus einer ungeraden Anzahl von Werten ermittelt und als arithmetisches Mittel aus  $x_t$ , den  $k$  vorausgehenden Werten  $x_{t-k}, x_{t-k+1}, \dots, x_{t-1}$  und den  $k$  nachfolgenden Werten  $x_{t+1}, x_{t+2}, \dots, x_{t+k}$  berechnet:

$$\hat{T}_t = \frac{1}{2k + 1} (x_{t-k} + x_{t-k+1} + \dots + x_t + \dots + x_{t+k})$$

$$= \frac{1}{2k + 1} \sum_{i=t-k}^{t+k} x_i$$

$$\hat{T}_t \text{ für } t = k + 1, \dots, n - k$$

mit:

$x_t$  Ursprungswert  
 $\hat{T}_t$  Wert der geschätzten glatten Komponente

**Beispiel:** Wird unterstellt, dass bei einer Zeitreihe aus Monatsdaten saisonale Schwankungen im Quartalsrhythmus auftreten, beträgt die Periodenlänge 3 Monate ( $k = 1$ ). Für den gleitenden Durchschnitt kann der saisonbereinigte Wert wie folgt berechnet werden:

$$\hat{T}_t = \frac{1}{3}(x_{t-1} + x_t + x_{t+1}) = \frac{1}{3} \sum_{i=t-1}^{t+1} x_i$$

### Gleitender Durchschnitt gerader Ordnung (2 k):

Der gleitende Durchschnitt kann auch für eine gerade Anzahl von Werten ermittelt werden. Der berechnete Wert würde dann zwischen zwei Zeitpunkten liegen. Um diesen ungewünschten Effekt zu vermeiden, wird ein zweiter Durchschnitt ermittelt, so dass das arithmetische Mittel aus beiden Durchschnittswerten dem Wert zum gewünschten Zeitpunkt entspricht. Das arithmetische Mittel wird aus  $x_t$ , den  $k$  vorausgehenden Werten und den  $k$  nachfolgenden Werten bestimmt:

$$\hat{T}_t = \frac{1}{2} \left[ \frac{1}{2k} (x_{t-k} + \dots + x_t + \dots + x_{t+k-1}) + \frac{1}{2k} (x_{t-k+1} + \dots + x_t + \dots + x_{t+k}) \right]$$

$$\hat{T}_t = \frac{1}{2k} \left( \frac{1}{2} x_{t-k} + x_{t-k+1} + \dots + x_t + \dots + x_{t+k-1} + \frac{1}{2} x_{t+k} \right)$$

$$= \frac{1}{2k} \left[ \frac{1}{2} x_{t-k} + \sum_{i=t-k+1}^{t+k-1} x_i + \frac{1}{2} x_{t+k} \right]$$

$$\hat{T}_t \text{ für } t = k + 1, \dots, n - k$$

**Beispiel:** Die Saisonschwankungen bei Arbeitsmarktzahlen treten im Jahresrhythmus auf. Liegen der Zeitreihe Monatswerte zu Grunde, ist eine Periodenlänge von 12 Monaten zu berücksichtigen ( $k = 6$ ). Der gleitende Jahresdurchschnitt berechnet sich wie folgt:

$$\hat{T}_t = \frac{1}{2} \left[ \frac{1}{12} (x_{t-6} + \dots + x_t + \dots + x_{t+5}) + \frac{1}{12} (x_{t-5} + \dots + x_t + \dots + x_{t+6}) \right]$$

$$\hat{T}_t = \frac{1}{12} \left( \frac{1}{2} x_{t-6} + x_{t-5} + \dots + x_t + \dots + x_{t+5} + \frac{1}{2} x_{t+6} \right)$$

$$= \frac{1}{12} \left[ \frac{1}{2} x_{t-6} + \sum_{i=t-5}^{t+5} x_i + \frac{1}{2} x_{t+6} \right]$$

Die geschätzte glatte Komponente  $\hat{T}_t$  entspricht dem um die Saison- und Restkomponente bereinigten Ursprungswert:

$$\hat{T}_t = x_t - S_t - R_t$$

### Saisonfigur:

Die periodisch wiederkehrenden saisonalen Schwankungen werden als Saisonfigur bezeichnet. Bei Arbeitslosenzahlen ist ein alljährlich wiederkehrendes Muster zu erkennen. So sorgen beispielsweise die Wintermonate in den witterungsabhängigen Wirtschaftszweigen für unfreiwillige Arbeitspausen und damit für einen Anstieg der Arbeitslosenzahlen.

### Glättung:

Glätten ist eine Methode zur Bereinigung von Datenreihen. Ziel dieser Methode ist es, die wichtigen Muster der Datenreihe aufzuzeigen und von Störgrößen und irregulären Komponenten zu befreien.

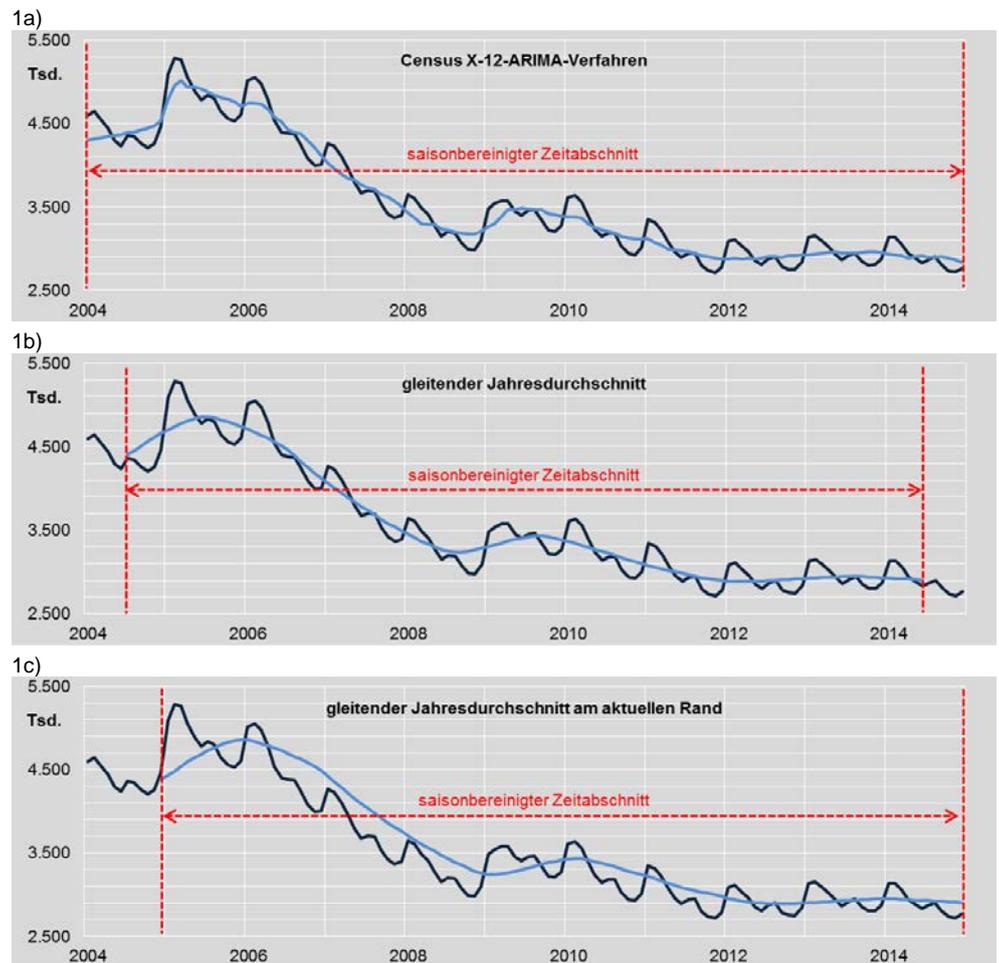
Ein Glättungsverfahren ist der gleitende Durchschnitt, bei dem die Bereinigung von Datenreihen basierend auf der Bildung von arithmetischen Mittelwerten erfolgt.

**Saisonbereinigter Zeitabschnitt:**

Bei den Verfahren des gleitenden Durchschnitts bzw. Jahresdurchschnitts können nicht zu allen Ursprungswerten saisonbereinigte Werte ermittelt werden. In Abhängigkeit der Länge der periodischen Saisonschwankungen ist für die  $k$  ersten und  $k$  letzten bzw. am aktuellen Rand für die  $k$  ersten Ursprungswerte keine Durchschnittswert-Berechnung möglich.

Mit Hilfe von Prognosen werden beim **Berliner Verfahren** und bei **Census X-12-ARIMA** fehlende Ursprungswerte geschätzt, um die Saisonbereinigung über die gesamte Zeitreihe zu ermöglichen.

**Abbildung 1:** Arbeitslosenzahlen von 2004 bis 2014  
Ursprungswerte (dunkelblau) und saisonbereinigte Werte (hellblau) nach drei Bereinigungsverfahren



**Gleitender Durchschnitt am aktuellen Rand ( $k$ ):**

Bei der Interpretation von Entwicklungen werden auch am aktuellen Rand saisonale Einflüsse untersucht. Jedoch stehen für die Berechnung mittels gleitenden Durchschnitts noch keine dem Betrachtungszeitpunkt nachfolgenden (zukünftige) Werte zur Verfügung. Die glatte Komponente  $\hat{T}_t$  wird für alle Betrachtungszeitpunkte aus  $x_t$  und den  $k - 1$  vorausgehenden Werten geschätzt:

$$\hat{T}_t = \frac{1}{k} (x_{t-k+1} + x_{t-k+2} + x_{t-k+3} + \dots + x_t)$$

$$= \frac{1}{k} \left[ \sum_{i=t-k+1}^t x_i \right]$$

$$\hat{T}_t \text{ für } t = k, \dots, n$$

**Beispiel:** Bei monatlichen Arbeitslosenzahlen beträgt die Periodenlänge 12 Monaten ( $k = 12$ ). Der gleitende Jahresdurchschnitt am aktuellen Rand wird wie folgt berechnet:

$$\hat{T}_t = \frac{1}{12} (x_{t-11} + x_{t-10} + \dots + x_t)$$

$$= \frac{1}{12} \left[ \sum_{i=t-11}^t x_i \right]$$

Der gleitende Durchschnitt ist eine sehr einfache Methode zur Eliminierung saisonaler Effekte aus den Ursprungswerten. Für fast jeden Zeitpunkt kann die glatte Komponente  $\hat{T}_t$  als Mittelwert vorausgehender und nachfolgender oder nur aus vorangegangenen Beobachtungen geschätzt werden. Die Anzahl der zu berücksichtigten Ursprungswerte bestimmt sich aus der Periodenlänge der Saisonschwankung. Das Prinzip, der Zerlegung der Ursprungswerte einer Zeitreihe in einzelne Komponenten findet grundlegend auch bei sehr komplexen Verfahren der Saisonbereinigung, wie dem Berliner Verfahren oder dem Verfahren Census X-12-ARIMA Anwendung.

### Übungen:

Arbeitslosenzahlen für die Übungsaufgaben finden Sie hier:

[http://statistik.arbeitsagentur.de/nn\\_31892/SiteGlobals/Forms/Rubrikensuche/Rubrikensuche\\_Form.html?view=processForm&resourceId=210368&input\\_=&pageLocale%0b=de&topicId=17590&year\\_month=aktuell&year\\_month.GROUP=1&search=Suchen](http://statistik.arbeitsagentur.de/nn_31892/SiteGlobals/Forms/Rubrikensuche/Rubrikensuche_Form.html?view=processForm&resourceId=210368&input_=&pageLocale%0b=de&topicId=17590&year_month=aktuell&year_month.GROUP=1&search=Suchen)

- 1) Bereinigen Sie eine Zeitreihe um saisonale Effekte: Bestimmen Sie auf Basis der Methode *gleitender Durchschnitt gerader Ordnung* für jeden Ursprungswert zwischen Januar 2000 und Dezember 2013 die glatte Komponente  $\hat{T}_t$ . Tragen Sie die bereinigte Reihe und die Ursprungswerte in eine Grafik ein.
- 2) Unterstellen Sie, dass die Restkomponente Null ist. Berechnen Sie nun für jeden Beobachtungszeitpunkt (Jan. 2000 bis Dez. 2013) die Saisonkomponente und tragen Sie diese in die obige Grafik ein.
- 3) Wenden Sie nun die Methode *gleitender Durchschnitt am aktuellen Rand* auf die Ursprungswerte an. Vergleichen Sie den Graph der bereinigten Reihe mit dem aus Übung 1 und diskutieren Sie die Unterschiede und worauf diese zurückzuführen sind.

Weiterführende Produkte zu Zeitreihen und saisonbereinigten Arbeitsmarktdaten finden Sie auf den Seiten der Statistik der Bundesagentur für Arbeit:

<http://statistik.arbeitsagentur.de/>

### Zeitreihen:

[http://statistik.arbeitsagentur.de/nn\\_31892/SiteGlobals/Forms/Rubrikensuche/Rubrikensuche\\_Form.html?view=processForm&resourceId=210368&input\\_=&pageLocale=de&topicId=17590&year\\_month=aktuell&year\\_month.GROUP=1&search=Suchen](http://statistik.arbeitsagentur.de/nn_31892/SiteGlobals/Forms/Rubrikensuche/Rubrikensuche_Form.html?view=processForm&resourceId=210368&input_=&pageLocale=de&topicId=17590&year_month=aktuell&year_month.GROUP=1&search=Suchen)

### Saisonbereinigte Zeitreihen:

[http://statistik.arbeitsagentur.de/nn\\_32022/SiteGlobals/Forms/Rubrikensuche/Rubrikensuche\\_Form.html?view=processForm&resourceId=210368&input\\_=&pageLocale=de&topicId=734738&year\\_month=aktuell&year\\_month.GROUP=1&search=Suchen](http://statistik.arbeitsagentur.de/nn_32022/SiteGlobals/Forms/Rubrikensuche/Rubrikensuche_Form.html?view=processForm&resourceId=210368&input_=&pageLocale=de&topicId=734738&year_month=aktuell&year_month.GROUP=1&search=Suchen)

### Analytikreports:

<http://statistik.arbeitsagentur.de/Navigation/Statistik/Statistische-Analysen/Analytikreports/Zentral/Monatliche-Analytikreports/Fuehindikatoren-Arbeitsmarkt-nav.html>

### Monatsberichte:

<http://statistik.arbeitsagentur.de/Navigation/Statistik/Arbeitsmarktberichte/Monatsbericht-Arbeits-Ausbildungsmarkt-Deutschland/Monatsbericht-Arbeits-Ausbildungsmarkt-Deutschland-Nav.html>

### Methodenbericht:

<http://statistik.arbeitsagentur.de/Statistischer-Content/Grundlagen/Methodenberichte/Arbeitsmarktstatistik/Generische-Publikationen/Methodenbericht-Saisonbereinigung.pdf>

### Literatur:

MOSLER, K. und SCHMID, F. (2003). Beschreibende Statistik und Wirtschaftsstatistik. Springer Verlag, Heidelberg, 4. Auflage.

SCHLITGEN, R. und STREITBERG, B. (1994). Zeitreihenanalyse. Oldenbourg Verlag, München, 5. Auflage.

SPETH, H.-T. (2004). Methodenberichte. Komponentenerlegung und Saisonbereinigung ökonomischer Zeitreihen mit dem Verfahren BV4.1. Statistisches Bundesamt, Wiesbaden.

## Impressum

**Herausgeber:** Bundesagentur für Arbeit  
Statistik

**Erstellungsdatum:** Februar 2015

Statistik-Service Ost

Tel.: 030/555599-7373

Fax: 030/555599-7375

E-Mail: [Statistik-Service-Ost@arbeitsagentur.de](mailto:Statistik-Service-Ost@arbeitsagentur.de)

## Weiterführende statistische Informationen:

Internet: <http://statistik.arbeitsagentur.de>

### Statistik-Service Nordost

Tel.: 0511/919-3455

Fax: 0511/919-4103456

E-Mail: [Statistik-Service-Nordost@arbeitsagentur.de](mailto:Statistik-Service-Nordost@arbeitsagentur.de)

### Statistik-Service Ost

Tel.: 030/555599-7373

Fax: 030/555599-7375

E-Mail: [Statistik-Service-Ost@arbeitsagentur.de](mailto:Statistik-Service-Ost@arbeitsagentur.de)

### Statistik-Service Südost

Tel.: 0911/179-8001

Fax: 0911/179-908001

E-Mail: [Statistik-Service-Suedost@arbeitsagentur.de](mailto:Statistik-Service-Suedost@arbeitsagentur.de)

### Statistik-Service Südwest

Tel.: 069/6670-601

Fax: 069/6670-910307

E-Mail: [Statistik-Service-Suedwest@arbeitsagentur.de](mailto:Statistik-Service-Suedwest@arbeitsagentur.de)

### Statistik-Service West

Tel.: 0211/4306-331

Fax: 0211/4306-470

E-Mail: [Statistik-Service-West@arbeitsagentur.de](mailto:Statistik-Service-West@arbeitsagentur.de)

