

# Eurostat

## - das statistische Amt der Europäischen Union -

Vorbereitungsmaterial

Betreuerin: Andrea Wiencierz

Autorin: Anna Chulkeper

07.06.2013

# Inhaltsverzeichnis

1. Einleitung	3
2. Historie von Eurostat	4
3. Ziele und Aufgaben	4
4. Veröffentlichungen	5
5. Daten	6
6. Mikrodaten	7
7. Saisonbereinigung	8
7.1 Gründe für die Saisonbereinigung	8
7.2 Die Komponente einer Zeitreihe	9
7.3 Arten der Komponentenverknüpfung	9
7.4 Verwendete Saisonbereinigungsverfahren	9
8. Das Verfahren X12- Arima	10
9. Literaturverzeichnis	12

## 1. Einleitung

Amtliche Statistik spielt in der öffentlichen Verwaltung, in Politik und Wirtschaft, wie auch in der Forschung eine wichtige Rolle. Der Bedarf an qualitativ hochwertigen Statistiken ist groß, um die Entwicklungen in Wirtschaft, Gesellschaft, Umwelt und Kultur zu erfassen und zu beschreiben, z.B. makroökonomische Entwicklungen wie die Inflation, die Beschäftigung und die öffentlichen Finanzen. So können politische Entscheidungen auf einer unparteiischen und sachlichen Grundlage getroffen werden. Amtliche Statistiken sind also für eine moderne Demokratie unverzichtbar.

Eurostat ist das statistische Amt der Europäischen Union und hat seinen Sitz in Luxemburg. Es beschafft auf europäischer Ebene Statistiken. Die Daten bezieht Eurostat hauptsächlich von den nationalen statistischen Ämtern seiner 27 Mitgliedstaaten. Nach vorgegebenen statistischen Methoden und Normen werden diese dann verarbeitet, analysiert und EU-weit veröffentlicht.

Eine der Aufgaben von Eurostat ist es, die Entwicklung ähnlicher statistischer Systeme in Nachbarländern der Europäischen Union zu fördern und zu harmonisieren.

Nicht nur auf nationaler, sondern auch auf europäischer Ebene sind Statistiken für die Entscheidung, Durchführung, Überwachung und Bewertung politischer Maßnahmen unentbehrlich. Ebenso kann die wirtschaftliche und soziale Lage wirklichkeitsgetreu erfasst werden. Die statistischen Daten stehen aber nicht nur der Politik, Wirtschaft und Forschung zur Verfügung, sondern auch den Medien und der Öffentlichkeit.

Durch die europäischen Wirtschaftsindikatoren (WEWI), durch Indikatoren für nachhaltige Entwicklung, Strukturindikatoren und Indikatoren für Beschäftigungs- und Sozialpolitik können europaweite Entwicklungen überwacht werden und der Vergleich zwischen Ländern und Regionen wird ermöglicht.

So können Statistiken viele Fragen beantworten, wie zum Beispiel, ob die Entwicklung der Gesellschaft so war, wie es die Politik versprochen hat. Auch die Entwicklung der Arbeitslosigkeit, des CO<sub>2</sub>-Ausstoß in den letzten 10 Jahren, die Erwerbstätigkeit von Frauen, die Leistungsfähigkeit einer nationalen Wirtschaft im Vergleich zu den anderen EU-Ländern wird erfasst und beschrieben, Eurostat (2012).

## **2. Historie von Eurostat**

Ein Vorgänger von Eurostat wurde 1953 von der Europäischen Gemeinschaft für Kohle und Stahl (Montanunion) gegründet. Über die Zeit wuchs sein Aufgabengebiet stetig, 1958 wurde die Europäische Wirtschaftsgemeinschaft und Generaldirektion (GD) der Europäischen Kommission gegründet.

1959 wurde der Name Eurostat für das bisherige Statistische Amt der Europäischen Gemeinschaft eingeführt.

Eurostat liefert heutzutage eine große Bandbreite an wichtigen Daten, die von den Regierungen, Unternehmen, Bildungseinrichtungen, den Medien und der Öffentlichkeit in vielen Bereichen genutzt werden.

Mit der Entwicklung der EU hat sich aber auch die Rolle von Eurostat verändert. Die Erhebung von Daten für die Wirtschafts- und Währungsunion und der Aufbau statistischer Systeme in den Ländern, die der EU beitreten möchten, hat eine größere Wichtigkeit als noch vor 10 Jahren, (<http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/>).

## **3. Ziele und Aufgaben**

Wie schon in der Einleitung beschrieben, ist die wichtigste Aufgabe Eurostats die Verarbeitung und Veröffentlichung vergleichbarer statistischer Daten auf EU-weiter Ebene. Es sollen gemeinsame statistische Begriffe, Methoden, Strukturen und technische Normen für alle Mitgliedsstaaten gelten.

Die Daten, die Eurostat von den Statistikbehörden der Mitgliedstaaten erhält, werden geprüft und analysiert. Es wird gewährleistet, dass sie vergleichbar sind, d.h. nach einer einheitlichen Methodik erstellt wurden. Außer Eurostat gibt es keine Behörde, die auf europäischer Ebene statistische Daten offiziell verarbeitet.

Ein Beispiel ist die Arbeitslosigkeit in der EU. Um vergleichbare Daten zu haben, müssen die Arbeitslosen in Frankreich oder Spanien nach demselben Verfahren erfasst werden wie in Finnland oder Deutschland.

Damit dies klappt, wird eine gemeinsame Methodik für die betreffenden Gebiete erstellt oder es

werden zum Beispiel in den Erhebungsbögen die gleichen Fragen gestellt. Anhand der so ermittelten Daten lassen sich die Arbeitslosenquoten der einzelnen Länder miteinander empirisch vergleichen.

Die Wirtschafts- und Währungsunion (WWU) hat als gemeinsame Währung den Euro und so können auch einzelne Mitgliedstaaten aus der WWU mit anderen Ländern mit Währungen wie dem US-Dollar und dem Yen verglichen werden.

Vorteile der Währungsunion sind, dass Eurostat dadurch Wirtschaftsindikatoren für das gesamte EU-Gebiet entwickeln und veröffentlichen kann,

(<http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/>).

#### **4. Veröffentlichungen**

Eurostat hat eine große Anzahl an regelmäßigen Veröffentlichungen, die beiden wichtigsten möchte ich kurz näher vorstellen.

Das *Eurostat Jahrbuch der Regionen* liefert ein detailliertes Bild über viele statistische Themen in den Regionen seiner Mitgliedstaaten (EU) sowie den Regionen der EFTA (Europäische Freihandelsassoziation) und der Kandidatenländer. Das Jahrbuch liefert statistische Informationen in Karten, Abbildungen und Tabellen und beschreibt die wichtigsten Ergebnisse, die Datenquellen und den politischen Kontext. Die regionalen Indikatoren sind in 11 Themenbereiche aufgeteilt, zum Beispiel Wirtschaft, Bevölkerung, Gesundheit, Bildung und Arbeitsmarkt. Außerdem gibt es drei spezielle Kapitel, die Städte und Küstenregionen näher betrachten und Methoden zur Trennung von ländlichen und städtischen Gebieten beinhalten.

Eine weitere wichtige Veröffentlichung ist der *Statistische Atlas*. Hierbei handelt es sich um eine interaktive Anwendung zum Betrachten von statistischen und topographischen Karten. Es können geografische Basisinformationen, wie zum Beispiel die Grenzen der NUTS-Regionen oder die Urban-Audit-Städte, mit Informationen aus statistischen Karten kombiniert werden.

Einzelne Kartendetails können auf unterschiedliche Art hervorgehoben werden:

- Herein- und Herauszoomen auf Europa und Regionen in äußerster Randlage
- Ändern der Transparenz-Einstellungen der Kartenebenen
- Abrufen des Codes, der geographischen Bezeichnung und der statistischen Daten einer bestimmten NUTS-Region oder Urban-Audit-Stadt.

Im *Statistischen Atlas* sind alle Karten aus dem *Eurostat Jahrbuch der Regionen* nach Themen und Kapiteln sortiert enthalten.

Weithin umfasst das Veröffentlichungsprogramm von Eurostat folgende Sammlungen:

- Diverse Pressemitteilungen, die aktuelle Informationen über Euro-Indikatoren und über soziale, ökonomische, regionale, landwirtschaftliche oder ökologische Themen liefern.
- Pocketbooks (Taschenbücher) als kostenlose Veröffentlichungen mit einer Auswahl an wesentlichen Daten über ein spezifisches Thema.
- *Statistik kurz gefasst* mit aktuellen Daten und weiteren Informationen über die Ergebnisse von Erhebungen, Studien und statistischen Analysen.
- *Methodologies and working papers* (Methodologien und Arbeitspapiere) als technische Veröffentlichungen für statistische Experten, die auf einem speziellen Gebiet forschen und arbeiten.
- Weitere Broschüren und Kataloge.

In *Statistics Explained*, gibt Eurostat weitere Information, es handelt sich hierbei um ein benutzerfreundliches Online-Veröffentlichungssystem im Wiki-Format, (<http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/>).

## 5. Daten

Die Eurostat Datenbank umfasst zahlreiche Datensätze. Es besteht die Möglichkeit, einzelne Datensätze oder die gesamte Datenbank unter dem Bereich „Statistik“ vollständig herunterzuladen.

Die wichtigsten Datensätze sind nach folgenden Themenbereichen geordnet:

- Allgemein- und Regionalstatistiken
- Wirtschaft und Finanzen
- Bevölkerung und soziale Bedingungen
- Industrie, Handel und Dienstleistungen
- Landwirtschaft, Forstwirtschaft und Fischerei
- Außenhandel
- Verkehr

- Umwelt und Energie
- Wissenschaft und Technologie

Zu jedem dieser Themen gibt es einige Unterthemen (beispielsweise beim Thema Bevölkerung und soziale Bedingungen die Unterthemen Bevölkerung, Gesundheit, Bildung und Weiterbildung, Arbeitsmarkt, Lebensbedingungen und Sozialleistungen, Kriminalität und Strafverfolgung sowie Kultur). Diese Unterthemen sind als Hyperlinks dargestellt, über die man zu einem speziellen Bereich gelangt, in dem man in der Regel Informationen zu Daten (Haupttabellen und Datenbanken), Veröffentlichungen, Rechtsetzung, Methodik und weitere Hintergrundinformationen findet.

Außerdem bietet Eurostat ein Geografisches Informationssystem der Europäischen Kommission an. Dieses Informationssystem, kurz GIS, ist ebenfalls in Einzelbereiche aufgeteilt:

- GIS (Geografisches Informationssystem), es umfasst Hardware, Software und Daten zur Erfassung, Verarbeitung, Analyse und Wiedergabe von geografisch referenzierten Informationen aller Art.
- GISCO (Geografisches Informationssystem der Kommission), es dient der Verwaltung der geografischen Referenzdatenbank der Kommission und der Verbreitung ihrer Inhalte. Es produziert Karten- und Raumanalysen, dient der Georeferenzierung und bietet den Nutzern Unterstützung,

[\(http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/\)](http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/).

## 6. Mikrodaten

Bei Mikrodaten handelt es sich um vertrauliche Daten, die Informationen über einzelne statistische Einheiten enthalten und anonymisiert sind. Aufgrund der derzeitigen Rechtssituation ist der Zugang zu diesen Daten nur für wissenschaftliche Zwecke genehmigt, aber kostenlos.

Folgende Mikrodatenbestände gibt es:

- ECHP (Haushaltspanel der Europäischen Union)
- AKE (Arbeitskräfteerhebung)
- CIS (Innovationserhebung der Union)
- AES (Erhebung über die Erwachsenenbildung)

- EU-SILC (Statistik der Europäischen Union über Einkommen und Lebensbedingungen)
- VSE (Verdienststrukturerhebung)
- FSS (Erhebung über die Struktur der landwirtschaftlichen Betriebe)

Am Beispiel des Haushaltspanel der Europäischen Union (ECHP) möchte ich erklären, wie Eurostat Mikrodaten erstellt:

Es handelt sich hierbei um eine Panelerhebung, bei der im jährlichen Abstand eine Haushalts- und Personenstichprobe zu ihren Lebensbedingungen befragt wird. Im Fragenkatalog sind viele verschiedene Themen enthalten, zum Beispiel detaillierte Angaben zum Einkommen, die finanzielle Lage generell, die Erwerbsbeteiligung, die Wohnsituation, die sozialen Beziehungen, Angaben zum Gesundheitszustand und weitere soziodemografische Angaben zum Befragten. Das ECHP wurde in acht Wellen über einen Zeitraum von insgesamt acht Jahren (1994-2001) durchgeführt, (<http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/>).

## **7. Saisonbereinigung**

Die Saisonbereinigung hat die Aufgabe, saisonbedingte Auswirkungen zu schätzen und diese aus den Zeitreihen zu entfernen, damit etwaige saisonale Schwankungen eliminiert werden. Dies gewährleistet eine Vergleichbarkeit der Daten über längere Zeitabschnitte hinweg, da eine Störung der Daten durch ungewollte Saisoneffekte aus den Daten entfernt wurde.

Bei Prognosen und bei kurzfristigen Analysen für Politik haben saisonbereinigte Daten eine große Bedeutung. Eurostat deckt alle wichtigen Saison- und Kalenderbereinigungsverfahren ab, im Folgenden möchte ich dieses Verfahren näher beschreiben.

### **7.1 Gründe für die Saisonbereinigung**

Es gibt verschiedene Gründe, die eine Saisonbereinigung notwendig machen:

Einerseits natürliche Gründe, wie zum Beispiel alle regelmäßig wiederkehrenden Ereignisse, die mit dem Ablauf des Jahres zusammenhängen. Dann institutionelle Gründe, zum Beispiel die allgemeine Ferienzeit, Ausstellungs- und Messetermine oder den Schlussverkauf. Und außerdem kalenderbedingte Gründe, wie die unterschiedliche Anzahl von Tagen, vor allem Arbeits-, Sonn- und Feiertagen im Monat, Siegmann (1993).

## 7.2 Die Komponenten einer Zeitreihen

Eine Zeitreihe besteht aus folgenden Komponenten, Siegmann (1993):

- $Y(t) = f(T(t), C(t), S(t), I(t))$  für  $t = 1, \dots, n$   
 $Y(t)$  Originalwert,  
 $T(t)$  Trendkomponente,  
 $C(t)$  Zyklische Komponente,  
 $S(t)$  Saisonkomponente,  
 $I(t)$  Irreguläre Komponente,  
 $t$  Zeit.
- Die beiden Komponenten  $T(t)$  und  $C(t)$  kann man auch zu der glatten Komponente  $G(t)$  zusammenfassen.

## 7.3 Die Arten der Komponentenverknüpfung

Es gibt drei verschiedene Modelle der Komponentenverknüpfung, Danckwerts (1971):

1. Additives Modell  $Y(t) = G(t) + S(t) + I(t)$   
Die Komponenten sind hier unabhängig.
2. Multiplikatives Modell  $Y(t) = G(t) * S(t) * I(t)$   
Es herrscht Proportionalität zwischen den Komponenten.
3. Gemischtes Modell: Additiv-Multiplikatives Modell  $Y(t) = G(t) * S(t) + I(t)$

## 7.4 Verwendete Saisonbereinigungsverfahren

Es gibt verschiedene Arten von Saisonbereinigungsverfahren, Siegmann (1993):

- Berliner Verfahren
- X-ARIMA
- Tramo / Seats
- Census Verfahren

Für die Anwendung der Verfahren X12-ARIMA und TRAMO/SEATS wird das von Eurostat zur Verfügung gestellte Software-Paket DEMETRA eingesetzt.

## 8. Das Verfahren X12-ARIMA

Das Verfahren X12-ARIMA basiert auf seinem Vorgänger X11-ARIMA mit geringfügigen Unterschieden und es kann in drei Hauptteile aufgeteilt werden, Stier (1980):

1. Vorläufige Schätzung von Saisonfaktoren und erste Extremwertbereinigung.
2. Verbesserte Schätzung der glatten Komponenten mit den vorläufigen Saisonfaktoren und zweite Extremwertbereinigung.
3. Endgültige Saisonbereinigung / endgültige glatte Komponente.

Im ersten Teil wird die sogenannte „arbeitstägliche“ Bereinigung durchgeführt, um den Einfluss unterschiedlicher Anzahlen von Arbeitstagen eines Monats zu eliminieren:

$A_{ij}$  Anzahl der Arbeitstage im Monat  $j$  des  $i$ -ten Jahres,

$\bar{A}_j$  Durchschnittliche Anzahl von Arbeitstagen im Monat  $j$  in allen  $N$  Jahren.

$x_{ij} = Y_{ij} \frac{A_{ij}}{\bar{A}_j}$  Arbeitstäglich bereinigte Reihe, mit  $i=1, \dots, N$  und  $j=1, \dots, 12$ .

Auf die bereinigte Reihe wird dann ein gleitender 12-Monatsdurchschnitt  $D$  angewendet, als Approximation der glatten Komponenten:  $D^{12}(x_{ij}) \simeq G_{ij}$ .

Damit kann man die sogenannten S-I-Quotienten errechnen:  $R_{ij} := \frac{X_{ij}}{D^{12}(X_{ij})} \simeq S_{ij} I_{ij}$

Diese werden nun extremwertbereinigt, indem auf jeden Monat  $j$  ein dreifacher und dreigliedriger Durchschnitt angewendet wird, der im S-I-Quotienten das Element  $I_{ij}$  eliminieren soll, damit für die daraus resultierenden Quotienten  $R'_{ij}$  folgendes gilt:  $R'_{ij} \simeq S_{ij}$ .

Mit der Standardabweichung wird ein Konfidenzintervall  $R'_{ij} \pm 2\sigma_j$  um  $R'_{ij}$  gebildet und eine Extremwertbereinigung nach folgender Regel vorgenommen:

Wenn  $R_{ij} > R'_{ij} + 2\sigma_j$  oder  $R_{ij} < R'_{ij} - 2\sigma_j$  ist, dann wird  $R_{ij}$  durch einen Wert ersetzt, der aus dem arithmetischen Mittel seiner beiden Nachbarwerte gebildet wird.

Durch diese Extremwertbereinigung ergeben sich sog. „vorläufige“ saisonale Faktoren:

$R_{ij}^V \simeq S_{ij} I_{ij}$ , die eine bessere Approximation von  $S_{ij} I_{ij}$  als  $R_{ij}$  darstellen sollen. Diese vorläufige Faktoren werden nun wiederum geglättet durch einen dreifachen dreigliedrigen Durchschnitt. Das ergibt neue saisonale Faktoren  $R_{ij}^{VN} \simeq S_{ij}$ . Mit ihnen wird eine vorläufige

Saisonbereinigung von  $X_{ij}$  vorgenommen:  $V_{ij} := \frac{X_{ij}}{R_{ij}^{vN}} \simeq G_{ij} I_{ij}$

Im zweiten Teil wird eine neue glatte Komponente geschätzt, diesmal mittels der 15-Punkte-Formel von Spencer mit den Gewichten: -0.009, -0.019, -0.016, 0.009, 0.006, 0.144.... -0.009. Damit soll eine verbesserte Approximation von  $G_{ij}$  erzielt werden:  $D^{SP}(V_{ij}) \simeq G_{ij}$ . Mit Hilfe von

$D^{SP}(V_{ij})$  ergeben sich neue S-I-Quotienten:  $NR_{ij} := \frac{X_{ij}}{D_{ij}^{SP}} \simeq S_{ij} I_{ij}$  die ebenfalls eine bessere

Approximation von  $S_{ij} I_{ij}$  sein sollen als dies die  $R_{ij}$  waren.

Die neuen saisonalen Faktoren werden wieder extremwertbereinigt nach demselben Verfahren wie im ersten Teil: Zuerst die Glättung, dann die Berechnung ihrer Standardabweichung und schließlich die Ausschaltung von Extremwerten mit Hilfe des obigen Konfidenzintervalls. Dann werden die extremwertbereinigten  $NR_{ij}$  wieder geglättet, um die irreguläre Komponente  $I_{ij}$  auszuschalten.

Im dritten und letzten Teil ergibt sich die endgültig saisonbereinigte Reihe durch  $\frac{X_{ij}}{NR_{ij}} \simeq G_{ij} I_{ij}$ .

Die endgültige glatte Komponente ergibt sich mittels des obigen Spencer- Durchschnittes, der auf

die endgültig saisonbereinigte Reihe angewendet wird:  $D^{SP} \frac{X_{ij}}{NR_{ij}} \simeq G_{ij} = T_{ij} C_{ij}$ .

Zur Beurteilung der „Güte“ des Verfahrens wurden verschiedene deskriptive Tests entwickelt (Stier 1980).

Dieses Verfahren berücksichtigt insbesondere externe Informationen bei der Berechnung, da verschiedene Gewichtungsfunktionen benutzt werden und unterschiedlich lange gleitende Durchschnitte zur Anwendung kommen (Siegmann 1993).

## 9. Literaturverzeichnis

1. Danckwerts, R.-F.: Moderne Methoden der Saisonbereinigung, Hamburg, 1971
2. Europa in Zahlen, Eurostat Jahrbuch 2012
3. Siegmann Siegmann, F., Pinnekamp, H.-J.: EDV-gestützte Kurzfristprognosen mit Hilfe modifizierter Verfahren der Saisonbereinigung, Münster, Hamburg 1993
4. Stier, W.: Verfahren zur Analyse saisonaler Schwankungen in ökonomischen Zeitreihen, Berlin, Heidelberg, New York 1980
5. <http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/>