

Kurzdokumentation

zum

Gehaltsrechner

Bearbeitungsstand: **21.12.2023**



STATISTIK AUSTRIA
Bundesanstalt Statistik Österreich
Direktion Bevölkerung
Bereich Soziales und Lebensbedingungen
A-1110 Wien, Guglgasse 13
www.statistik.at

Für Rückfragen:
Mag. Serhan Marcel Bilgili
serhan.bilgili@statistik.gv.at
Tel.: +43-1-71128-8284

1 Einleitung / Datenbasis

Im Gehaltsrechner werden für Beschäftigungsverhältnisse, die durch verschiedene Merkmale wie Alter, Beruf, Branche etc. der beschäftigten Person bestimmt sind, jeweils die mittleren Bruttomonatsverdienste (Median) für unselbständig Erwerbstätige mit Wohnsitz und Arbeitsort in Österreich berechnet. Nicht enthalten sind Arbeitnehmer:innen in der Land- und Forstwirtschaft, in privaten Haushalten und bei extraterritorialen Organisationen und Körperschaften. Ebenfalls nicht enthalten sind Lehrlinge sowie Leiharbeitskräfte.

Als Grundlage für die Berechnungen wurde ein Datensatz erstellt, der die gewünschten einkommensbestimmenden Merkmale enthielt und auch weitere Anforderungen hinsichtlich Aktualität und Aktualisierungsmöglichkeit, Qualität sowie die Abdeckung der beschriebenen Grundgesamtheit erfüllte. Dieser Datensatz musste aus mehreren Datenquellen verknüpft werden, da die Berechnungen auf den Ergebnissen statistischer Schätzmodelle beruhen, welche Einkommensinformationen aus den Lohnsteuerdaten und Mikrozensusinformationen wie das Alter, die Qualifikation etc. verwenden. Des Weiteren wurden für eine Vergrößerung der zugrunde liegenden Stichprobe die Datensätze der drei Jahre 2020, 2021 und 2022 gepoolt. Der endgültige Stichprobenumfang umfasst 47.220 Angestellte, 19.551 Arbeiter:innen, 2.984 Beamt:innen, 4.909 Vertragsbedienstete.

Um möglichst aktuelle Ergebnisse zu erhalten, wurden zudem die Einkommenswerte mit branchenspezifischen Informationen auf den Monatswert Juli 2023 des Tariflohnindex¹ fortgeschrieben.

2 Datenquellen

Der kombinierte Datensatz besteht aus mehreren Komponenten (Mikrozensusdaten, Lohnsteuerdaten, Daten des Hauptverbandes der österreichischen Sozialversicherungsträger), denen unterschiedliche Konzepte und Zielsetzungen zugrunde liegen. Besonders herausfordernd war, dass die Lohnsteuerdaten Jahreseinkommen beinhalten, während sich die Mikrozensusinformationen auf eine bestimmte Referenzwoche beziehen. Daher musste eine Verknüpfung der Datenquellen sicherstellen, dass die Einkommensinformationen in geeigneter Weise mit den Mikrozensusinformationen verbunden und umgerechnet werden².

Die Lohnsteuerdaten wurden in Form der Lohnzetteldaten herangezogen. Für jedes Beschäftigungsverhältnis einer Person in einem Jahr muss, unabhängig von der Dauer des Dienstverhältnisses, ein solcher Jahreslohnzettel vom Dienstgeber ausgestellt und an die Finanzbehörden übermittelt werden. Statistik Austria steht dieser Datensatz für statistische Zwecke zur Verfügung.

Im Zusammenhang mit der Umrechnung von Jahreseinkommen auf Monatseinkommen haben die Daten des österreichischen Hauptverbandes der Sozialversicherungsträger (HV) besondere Bedeutung. Aus den HV-Daten wurde vor allem die Dauer der Erwerbsverhältnisse verwendet, um damit aus dem Bruttojahreseinkommen ein Bruttomonatseinkommen berechnen zu können.

Der österreichische Mikrozensus (MZ) ist die zentrale Quelle für die erklärenden Variablen. Jeder Haushalt wird in fünf aufeinanderfolgenden Quartalen jeweils einmal im Quartal befragt, wobei pro Quartal jeweils ein Fünftel der Haushalte ausscheidet und ersetzt wird (Fünftel-Rotation). Pro Quartal werden ca. 20.000 Haushalte erhoben. Da entsprechend den Quartalen jeweils vier Datensätze pro Jahr vorliegen, und eine bestimmte Person somit bis zu viermal vorkommen kann, wurden die Datensätze für die Zwecke des Gehaltsrechners zu Jahresdatensätzen konsolidiert.

¹ Der Jahreswert 2023 des Tariflohnindex (Durchschnitt des Monatswerte Jänner bis Dezember) lag zum Zeitpunkt der Modellherstellung im Dezember 2023 noch nicht vor. Daher wurde als Approximation ein Monatswert aus der Mitte des Jahres gewählt.

² Wesentliche Anregungen dafür stammen aus Baierl, A., Gumprecht, D. und Gumprecht, N. (2011) "Monatliches Nettoeinkommen im Mikrozensus - Konzept", Statistische Nachrichten 7/2011. 596-612, Wien.

3 Einkommensbegriff

Generell gilt, dass die Höhe des Einkommens aus den Lohnsteuerdaten stammt, für die Bestimmung des monatlichen Bruttoeinkommens muss aber zunächst die Bezugsdauer aus den Daten des HV in Tagen bestimmt werden, die sich für das Referenzmonat ergibt, um dann die Einkommenshöhe der relevanten Einheiten zu ermitteln.

Da für den Gehaltsrechner nur regelmäßig bezogene Einkommenskomponenten im Bruttomonatseinkommen enthalten sein sollen, mussten alle anderen Komponenten, wie z.B. Abfertigungen, unregelmäßig ausbezahlte Prämien und Boni abgezogen werden. Aus den gesamten Bruttojahresbezügen kann man mittels folgender Formel das Jahreseinkommen (*JB*), das sich aus den laufenden Monatseinkommen zusammensetzt, berechnen:

$$JB = KZ\ 210 - FESTSAT - KZ\ 220 - NTSONST$$

wobei die Kennzahlen des Lohnzettels folgende Bedeutung haben:

KZ210.....Bruttobezüge gemäß § 25 EStG (ohne § 26³ und ohne Familienbeihilfe)

FESTSAT.....Steuerfreie bzw. mit festen Sätzen versteuerte Bezüge gem. § 67 Abs. 3 bis 8 EStG. Darunter fallen vor allem die gesetzlichen und kollektivvertraglichen Abfertigungen sowie Urlaubsentschädigungen und -abfindungen.

KZ220.....Bezüge gemäß § 67 Abs. 1 und 2 EStG (innerhalb des Jahressechstels), wobei vor allem der 13. und 14. Monatsbezug zu nennen ist; vor Abzug der Sozialversicherungsbeiträge.

NTSONST.....Nach dem Tarif versteuerte sonstige Bezüge gemäß § 67 Abs. 2, 6 und 10 EStG vor Abzug der Sozialversicherungsbeiträge.

4 Datenaufbereitung

Während die Kohärenz und Vollständigkeit in den Mikrozensusdaten und im erstellten Datenbestand der Lohnsteuer-HV-Daten jeweils für sich gewährleistet ist, da die Daten bereits im Vorfeld überprüft und fehlende Werte im Mikrozensus durch geeignete statistische Verfahren imputiert wurden⁴, kann es durch die Zusammenführung dieser Datenkörper zu unplausiblen Kombinationen kommen. Daher musste überprüft werden, ob auch Merkmale aus dem Administrativdatenbestand (in diesem Zusammenhang vor allem das berechnete Monatseinkommen) mit den Mikrozensus-Merkmalen der Personen stimmig sind. Für den Umgang mit unplausiblen und fehlenden Werten musste eine geeignete Vorgehensweise festgelegt werden.

Die Plausibilitätsprüfung unterteilte sich in eine interne und eine externe Prüfung und konzentrierte sich auf die Höhe des Bruttomonatseinkommens in Abhängigkeit von verschiedenen anderen Variablen. Da das Bruttomonatseinkommen die abhängige Variable ist, die im Gehaltsrechner beschrieben wird, musste sichergestellt werden, dass keine unplausiblen Werte, die z.B. durch Datenfehler entstehen, in die Modellbildung eingehen. Eine deskriptive Analyse der Einkommenswerte ergab, dass an den Rändern der Verteilung zum Teil unplausible Extremwerte auftraten. Diese unplausiblen Werte können unterschiedliche Ursachen haben, wobei hier auch die Verknüpfung zweier Datenquellen (Mikrozensus und Lohnsteuerdaten) bedacht werden sollte. Ausreißer können beim gewählten Berechnungsmodell relativ großen Einfluss auf die Ergebnisse haben. Es wurde daher entschieden, für jede der vier beruflichen Stellungen die obersten und untersten 0,5% der Einkommen aus der Berechnung auszuschließen.

Fehlende Werte ergaben sich, weil nicht jeder der unselbständig erwerbstätigen Personen im Mikrozensus auch ein Einkommen aus den Administrativdaten zugewiesen werden konnte. In diesen Fällen wurde mit Hilfe eines geeigneten Verfahrens (Hot-Deck) imputierte Werte eingesetzt. Während die Imputation

³ Darin finden sich Vergütungen für Dienstreisen, Ersatz für Arbeitsbekleidung u.ä.

⁴ Vgl. dazu Moser, W. (2005). Das Datenmanagement im neuen Mikrozensus – eine Prozessbeschreibung. Austrian Journal of Statistics, S. 327-343.

beim Mikrozensus-Datensatz, der für die Verknüpfung mit den Lohnsteuer- und Hauptverbandsdaten verwendet wurde, bereits im Vorfeld isoliert geschah, mussten fehlende bzw. unplausible Einkommenswerte im kombinierten Datensatz noch imputiert werden.

5 Statistisches Modell

Die Schätzung der Einkommen erfolgt durch Regression der logarithmierten Einkommen auf erklärende Variablen.⁵

In Abhängigkeit des Arbeitsverhältnisses kommen folgende erklärende Variablen zur Anwendung: Alter (linear und quadratisch), Dauer der Betriebszugehörigkeit (linear und quadratisch), Bundesland des Arbeitsorts, Zahl der Beschäftigten an der Arbeitsstätte, Branche, berufliche Tätigkeit, Berufsgruppe, höchste abgeschlossene Ausbildung, wöchentliche Normalarbeitszeit.

Die erklärte Varianz (R^2) des Modells beträgt 74% für Angestellte, 73% für Arbeiter:innen, 58% für Beamt:innen und 67% für Vertragsbedienstete.

Da die Einkommen nur für eine Stichprobe aller unselbständig Erwerbstätigen vorliegen, sind die Ergebnisse des Gehaltsrechners mit einer gewissen statistischen Unsicherheit behaftet, die durch Konfidenzintervalle illustriert wird. Die Berechnung des Unsicherheitsbereichs erfolgt mittels Bootstrapgewichten des Mikrozensus.⁶

6 Ergebnisdarstellung

Die Darstellung der Ergebnisse erfolgt separat für Angestellte, Arbeiter:innen, Beamt:innen sowie Vertragsbedienstete.

Nutzer:innen wählen zunächst eines der vier Arbeitsverhältnisse aus und geben dann, je nach Arbeitsverhältnis, bestimmte Merkmale an, die Einfluss auf die Höhe des Einkommens haben (z.B. Alter, Branche und Bildungsabschluss).

Ausgewiesen werden das geschätzte (Median) monatliche Bruttoeinkommen in Euro, der Unsicherheitsbereich dieser Schätzung (95%-Konfidenzintervall) und der prozentuelle Einkommensnachteil von Frauen gegenüber Männern.

⁵ Fahrmeir, L., Kneib, T. und Lang, S. (2007): Regression. Modelle, Methoden und Anwendungen. Springer-Verlag Berlin-Heidelberg-New York.

⁶ Meraner, A., Gumprecht, D. und Kowarik, A. (2016): Weighting Procedure of the Austrian Microcensus using Administrative Data. Austrian Journal of Statistics, September 2016, Volume 45, 3-14.