

Univariate Zeitreihenanalyse zur Prognose der Lebensumstände von Senioren in der Bundesrepublik Deutschland

Okhtay Wahid Far

^aBerliner Hochschule für Technik, s90838@bht-berlin.de,

Abstract

In diesem Paper werden Daten über die ökonomische und psychische Verfassung der Senioren in der Bundesrepublik Deutschland verwendet und beleuchtet. Kurzum stellt sich die Frage, ob eine Verelendung der Senioren in der Zukunft vorab prognostiziert und stichhaltig belegt werden kann. Hierzu wird die univariate Zeitreihenanalyse angewandt und sehr gute Prognosen für den Zeitraum von 2022 bis 2024 mithilfe der AUTO-ARIMA-Funktion erstellt. Die AUTO-ARIMA-Funktion befindet sich im package 'forecast' und wird durch die Sprache R implementiert.

Keywords: AUTO-ARIMA, Prognose, Zeitreihenanalyse, univariat, R, forecast

1. Einleitung

Allgemein sind Senioren mit 65 Jahren noch erwerbstätig oder bereits in Frührente. Spätestens im Alter von 67 Jahren erhalten Senioren ihre monatliche gesetzliche Rente, welche in der Regel nicht für alle Fixkosten ausreicht. Aufgrund der aktuellen Inflation und den weltweiten wirtschaftlichen Krisen ist anzunehmen, dass die gesetzliche Rente in der Bundesrepublik Deutschland nicht ausreichend angepasst und die Altersarmut mit ihren Folgen wachsen wird. Dadurch wird sowohl die Sozialempfängerzahl als auch die Zahl der Grundsicherungsempfänger durch Senioren weiter steigen [1].

Zusätzlich kommt es in unserer modernen Welt immer häufiger vor, dass die Todesursachen von Senioren ab 65 Jahren durch psychische Erkrankungen verursacht werden. Nach dem BMFSJ (Bundesministerium für Familien, Senioren, Frauen und Jugend) sind über 40% der Selbstmordtoden über 60 Jahre alt [2]. Aus diesem Grund wird für die Zukunft unter anderem eine Steigerung der Selbstmordrate der über 60-jährigen Senioren erwartet [3]. Dieses Paper dient als belegte Mahnung für die Politik, damit die stetig fortschreitende Verelendung der Senioren durch Prognosen begründet beleuchtet wird. Zudem soll es auch als Ansporn dafür dienen, das Rentensystem an sich zu reformieren [4].

Anhand der aktuellen gesellschaftlichen Gegebenheiten werden folgende Thesen zu den Lebensumständen von Senioren bis 2024 behauptet:

- Steigerung der Sozialhilfeempfänger von weiblichen und männlichen Senioren ab 65 Jahren bis 2024
- Steigerung der psychischen Erkrankungen von weiblichen und männlichen Senioren ab 65 Jahren bis 2024

- Daraus resultierende Unterschiede zwischen weiblichen und männlichen Senioren ab 65 Jahren

Weiterführend sind nun überzeugende Prognosen notwendig, um diesem Trend erfassen sowie entgegen wirken zu können. Um nun durch stichhaltige Prognosen diese Thesen belegen zu können, bedient man sich der Auto-ARIMA-Funktion also `auto.arima()` aus dem Package 'forecast', welches im Paper 'Automatic Time Series Forecasting : The forecast Package for R' erstmals vorgestellt wurde [5]. Dieses Package wird speziell für univariate Regressionsmodelle und deren Zeitreihenanalyse verwendet.

Dieses Paper teilt die feste Überzeugung, dass durch die AUTO-ARIMA Funktion sehr gute und stichhaltige Prognosen erstellt werden können. Die Zuverlässigkeit dieser Prognosen wird durch die vielversprechenden MAPE-Werte belegt [6], die vergleichsweise bessere Prognosen liefert als durch die herkömmliche Bestimmung von ARIMA-Modellen. Das heißt, man kann durch die AUTO-ARIMA Funktion eine hohe Prognosegenauigkeit generieren [7], die sowohl mit saisonalen, nicht-saisonalen als auch mit nicht-stationären Daten umgehen kann. Ausgenommen sind aber Prognosen über Aktienwerte, dessen ARIMA-Modelle am Besten auf die klassische Weise bestimmt werden sollten [6]. Dementsprechend wird folgende Arbeitsweise umgesetzt:

- 1) Die betrachteten Datensätze werden auf missing values beleuchtet.
- 2) Die Daten werden in Intervallen für Alterbegrenzungen unterteilt, die jeweils ein Regressionsmodell wiedergeben.
- 3) Mithilfe der AUTO-ARIMA-Funktion wird ein passendes ARIMA-Modell für die Datenwerte erstellt.

4) Nach Hyndman wird dann die Autokorrelation dieses Modells durch den Ljung-Box Test überprüft.

5) Wenn daraufhin eine statistische Signifikanz vorherrscht, wird eine Prognose mithilfe der forecast-Funktion (forecast ()) für die Jahre 2022 bis 2024 produziert.

6) Die MAPE-Werte der einzelnen Prognosen werden dann auf Genauigkeit begutachtet.

2. Zeitreihen

Durch Daten aus Zeitreihen ist es möglich, Prognosen für zukünftige Daten zu ermitteln. Somit sind spätere Entwicklungen frühzeitig erkennbar und präventiv steuerbar [8]. Eine Zeitreihe beruht auf dem Verständnis des stochastischen Prozesses, die als Menge von Zufallsvariablen X_t definiert wird. Hierbei gilt die Zufallsvariable X_t als Umsetzung der Beobachtung zu einem Zeitpunkt $t \in T$ [9].

„Eine Variable oder ein Merkmal X , dessen Werte oder Ausprägungen die Ergebnisse eines Zufallsvorgangs sind, heißt Zufallsvariable X . Die Zahl $x \in \mathbf{R}$, die X bei einer Durchführung des Zufallsvorgangs annimmt, heißt Realisierung oder Wert von X .“ [10]

Generell spricht man von abhängigen Zufallsvariablen, wenn sich in einer zufälligen Stichprobe mehrere Zufallsvariablen gegenseitig beeinflussen. Hingegen gilt bei unabhängigen Zufallsvariablen, dass sich in einer zufälligen Stichprobe mehrere Zufallsvariablen nicht gegenseitig beeinflussen [10].

Seien der Index $t \in \mathbf{N}, \mathbf{N}_0$ oder \mathbf{Z} als Zeit und X_t als Folge von Zufallsvariablen bzw. als stochastischer Prozess gegeben. So gilt folgendes für die Umsetzung von X_t , die Trajektorie und die Zeitreihe [9, 11]:

$$y_1, \dots, y_N$$

Aufgrund der Anwendungsorientiertheit werden im Folgenden nur diskrete Messwerte für univariate Zeitreihen untersucht. Somit gilt $t \in \mathbf{Z}$ für die Menge der diskreten Messwerte für univariate Zeitreihen [9]. Als univariate Zeitreihen werden jene Zeitreihen bezeichnet, die nur eine abhängige Variable bzw. Merkmal betrachten [12]. Allgemein beinhalten diskrete Variablen also nur abzählbar unendlich viele Werte [13].

3. Univariate Zeitreihenanalyse

Um die Prognose einer Zeitreihe effektiv umsetzen zu können, erfordert es die Nutzung eines Modells [13]. Seien x_t die beobachteten Werte und t die Zeit mit $t = 1, 2, \dots, n$ gegeben. Zudem seien m_t die Trendkomponente, s_t die Saisonkomponente und u_t die Restkomponente. Dann gilt für das additive Komponentenmodell der Zeitreihen folgendes [13]:

$$x_t = m_t + s_t + u_t$$

Insbesondere sind Trends in einer Zeitreihe hinreichend für die Erstellung einer Prognose [13].

„... m_t stellt die glatte Komponente oder den Trend dar. Damit wird die langfristige Veränderung des Mittels erfasst. s_t ist die Saisonkomponente, eine jahreszeitlich bedingte, relativ regelmäßige Struktur. In der Restkomponente u_t sind die unregelmäßigen Abweichungen von den beiden Komponenten subsumiert.“ [13]

3.1. ARIMA-Modell

Die Autoregression bzw. das AR-Modell nutzt die Werte von vergangenen Zufallsvariablen, um eine Prognose zu erstellen. Hierbei wird eine Linearkombination von Prädiktoren verwendet, die auf einer multiplen Regression basiert [14]. Dem Gegenüber gestellt nutzt das Moving-Average-Modell die vergangenen Prognosefehler, um somit zukünftige Messwerte ermitteln zu können [14].

Sei die Ordnung p für den Autoregressive Abschnitt, der Grad der ersten Ableitung d und die Ordnung des Moving-Average-Bereichs q gegeben. Dann gilt für das nicht-saisonale ARIMA (p, d, q) - Modell folgendes [14]:

$$y'_t = c + \phi_1 y'_{t-1} + \dots + \phi_p y'_{t-p} + \theta_1 e_{t-1} + \dots + \theta_q e_{t-q} + e_t$$

3.2. Stationarität

Um für eine Zeitreihe eine aussagekräftige Prognose erstellen zu können, gilt als Grundvoraussetzung die Stationarität. Ohne diese Eigenschaft sind die Verfahren Autoregression und Moving-Averages jeweils nicht nutzbar [8]. Ein stochastischer Prozess $(Y_t)_{t \in T}$ heißt [15]:

1) mittelwertstationär, wenn μ_t mit $\mu_t := \mu$ für alle $t \in T$ konstant ist.

2) varianzstationär, wenn σ_t^2 mit $\sigma_t^2 := \sigma^2$ für alle $t \in T$ konstant ist.

3) kovarianzstationär, wenn die Kovarianzfunktion $\gamma(s, t)$ des Prozesses nur von der Entfernung $s-t$ mit $\gamma(s, t) := \gamma(s-t)$ für alle $t \in T$ abhängt.

4) schwachstationär, wenn er mittelwert- und kovarianzstationär ist.

Kurzum bedeutet Stationarität, dass die Korrelation zwischen den Zufallsvariablen konstant bleibt [8]. In der Praxis ist ein schwachstationärer stochastischer Prozess vollkommen ausreichend. Aus diesem Grund gilt die Stationarität als Sammelbegriff für alle stationäre Formen [15].

4. Daten und Auto-ARIMA - Funktion

Die Datensätze für diese Untersuchung sind auf der Seite des statistischen Bundesamtes (DESTATIS) öffentlich zugänglich und geben jeweils absolute Zahlen wieder. Kurzum beruht die Untersuchung auf metrischen Werten. Die Auto-ARIMA Funktion in R ist in der Bibliothek forecast enthalten und erstellt anhand der Daten das bestmögliche ARIMA-Modell.

4.1. Daten

Hierzu werden folgende Datensätze untersucht:

- 23211-0004: Gestorbene: Deutschland, Jahre, Todesursachen, Geschlecht, Altersgruppen [17]
- 22121-0001: Empfänger von Hilfe zum Lebensunterhalt: Deutschland, Stichtag, Ort der Leistungserbringung, Geschlecht, Altersgruppen [18]

Durch die Datensätze wird somit ein Einblick in die psychischen und ökonomischen Lebensumstände der Senioren ab 65 Jahren gewährt. Hierzu wurden vorab beide Datensätze nach den Datums- und Alterintervallen geordnet und betrachtet. Durch die vorherige Ordnung und die stetige Kontinuität der Daten sind missing values vorab beseitigt worden. Um die Daten für die Prognose vorzubereiten, wurden diese in einzelnen univariaten Regressionsmodellen zusammengetragen. Insbesondere wird zwischen weiblichen und männlichen Senioren unterschieden. Zudem werden nur die Daten von Senioren ab einschließlich 65 Jahren untersucht. Als Zeitintervall werden die erhobenen Daten von 2005 bis 2021 (beides einschließlich) betrachtet.

Der Datensatz 23211-0004 beinhaltet sowohl mehrere Todesursachen als auch eine Zeitspanne von 1980 bis 2021. Der Untersuchungsfokus liegt hierbei auf die durch psychische Verhaltensstörungen bedingten Todesursachen, die in natürlichen Zahlen wiedergegeben sind. Diesbezüglich existieren pro Jahr, Geschlecht und Altersspanne jeweils ein univariates Regressionsmodell. Dementsprechend liegen für folgende Altersspannen Regressionsmodelle vor:

- 65 bis unter 70 Jahre
- 70 bis unter 75 Jahre
- 75 bis unter 80 Jahre
- 80 bis unter 85 Jahre
- 85 Jahre und mehr.

Entsprechend der ökonomischen Lebensumstände werden nur die Empfängerzahlen der Lebensunterhaltshilfe des Datensatzes 22121-0001 betrachtet. Die Empfängerzahlen werden in natürlichen Zahlen wiedergegeben. Zudem wird zwischen Leistungen in- und außerhalb von Einrichtungen unterschieden, wobei der Unterschied nicht Gegenstand der Untersuchung ist. Das heißt, somit werden jeweils die gesamten Daten von weiblichen sowie männlichen Senioren im Gesamten begutachtet. Demnach existieren pro Jahr und je Geschlecht ein univariates Regressionsmodell, welches als Zielvariable die

Empfängerzahlen darlegt. Die Prognose wird für die Jahre 2022 bis einschließlich 2024 erstellt.

4.2. Auto-ARIMA-Funktion

Die Auto-ARIMA-Funktion wählt das ideale Modell für eine Zeitreihe aus, indem es vorab mithilfe der KPSS-Tests differenziert und den kleinst möglichen AIC-Wert bestimmt. Hierbei ist es irrelevant, ob die Zeitreihe nicht stationär ist, da dieses durch die Auto-ARIMA-Funktion ausgeglichen wird [14]. Nachdem das ideale ARIMA-Modell durch die Auto-ARIMA-Funktion berechnet wurde, muss die Autokorrelation zwischen den Werten überprüft werden. Dieses wird mithilfe der Funktion ACF() und mit dem Portmanteau bzw. Ljung Box-Test validiert. In diesem Paper liegt der Untersuchungsfokus auf dem Ljung Box-Test, da dieser qualitativ besser ist [14].

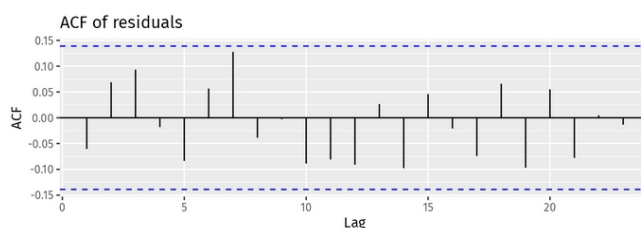


Bild 1: Ausgabe eines idealen ACF-Testes

Die acf()-Funktion betrachtet die Autokorrelation zwischen den gemessenen Werten, wobei der Ljung Box-Test eine direkte Prüfung anwendet. Hierbei muss der Test einen P-Wert ausgeben, der größer als 0.05 sein muss. Dementsprechend gilt dann die Nullhypothese und die unabhängige Verteilung der gemessenen Werte ist dadurch statistisch signifikant [16].

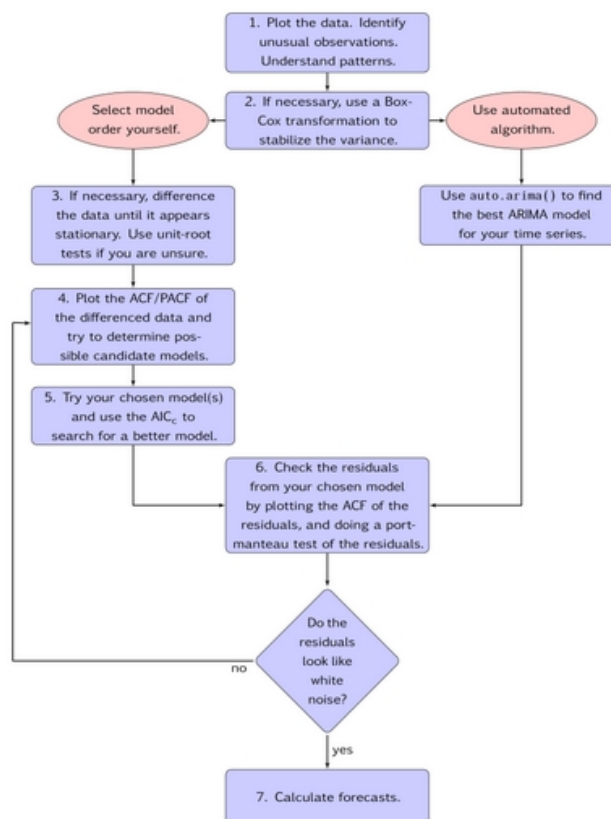


Bild 2: Prozess zur Prognose-Erstellung durch ein ARIMA-Modell [14]

5. Implementierung in R

Die Implementierung in R erfolgt über die Bibliothek `forecast` und insbesondere durch folgende Funktionen:

`install.packages("...")` :

Diese Funktion installiert das nötige Package bzw. Bibliothek, um die benötigten Funktionen abrufen zu können.

`file.choose()` :

Diese Funktion verhilft zur direkten Auswahl des Datensatzes auf dem lokalen Speicher.

`read_excel()` :

Diese Funktion importiert die Datensätze in die R-Umgebung.

`auto.arima()` :

Diese Funktion erstellt das best möglichste ARIMA-Modell.

`ts(..., frequency= 1, start= c(2005,1))` :

Diese Funktion gibt die Werte als Zeitreihe in einem Koordinatensystem wieder. Hierbei werden die Werte pro Jahr von 2005 bis 2021 betrachtet.

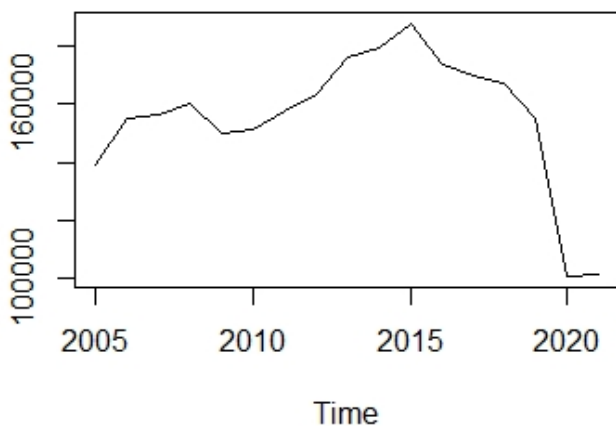


Bild 3: Zeitreihendarstellung durch die `ts`-Funktion vom Datensatz 001_wUeber65

`Box.test(...$residuals, type="Ljung-Box")` :

Diese Funktion prüft mithilfe des Ljung-Box Testes, ob die Werte im Auto-ARIMA-Modell unabhängig voneinander sind. Die statistische Signifikanz wird durch den p-Wert identifiziert.

`forecast(..., h=12)` :

Wenn die statistische Signifikanz für das Auto-ARIMA-Modell durch den Ljung-Box-Test festgestellt wurde, so ist dann erst eine aussagekräftige Prognose erstellbar. In dessen Plot wird hierzu das automatisch erstellte ARIMA-Modell angegeben und die Prognose ab 2022 für die nächsten 12 Jahre dargestellt. Dadurch wird eine bessere Übersicht der Prognose ermöglicht.

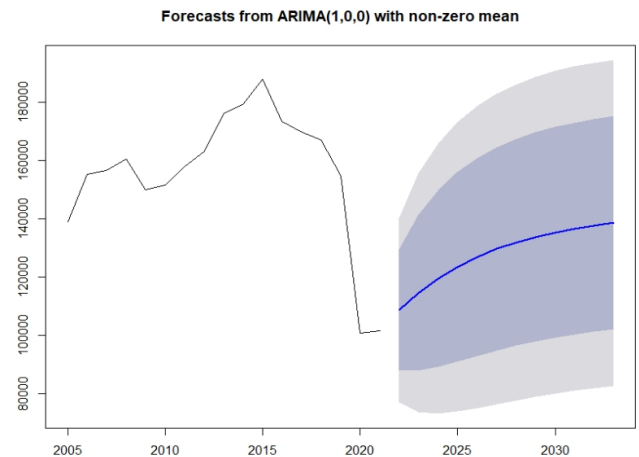


Bild 4: Übersichtliche Prognose zum Datensatz 001 der über 65-jährigen weiblichen Senioren von 2022 bis 2030

`plot(...,xlim=c(2015,2024))` :

Hierbei wird die erstellte Prognose durch die Funktion `forecast()` zwischen den Jahren 2015 bis 2024 als Plot vergrößert dargelegt. Gleichzeitig gibt der Plot das beste ARIMA-Modell für diese Werte wieder.

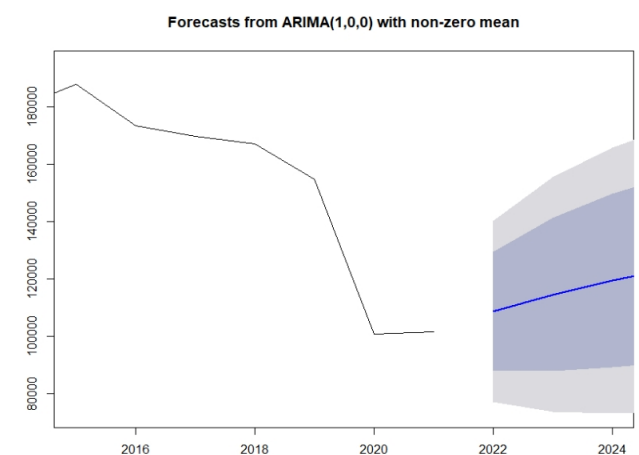


Bild 5: Vergrößerte Perspektive der Prognose des Datensatzes 001 der über 65-jährigen weiblichen Senioren

Sowohl die Struktur als auch die genutzten Funktionen gelten für alle untersuchenden Regressionsmodelle. Aus diesem Grund wird der folgende Code für alle anderen Regressionsmodelle integriert und analog wahrgenommen.

```

5 install.packages("readxl")
6 install.packages("forecast")
7 library(readxl)
8 data0001_wueber65<-read_excel(file.choose(), na="NA")
9 dataFrame_0001_wueber65<-
10   data.frame(data0001_wueber65$`ort der Leistungserbringung - Insgesamt weiblich`)
11
12 library(forecast)
13 'Erstellung der Zeitreihe'
14 tsdata<- ts(dataFrame_0001_wueber65, frequency = 1,start=c(2005,1))
15 plot (tsdata)
16
17 AutoArima_0001_wueber65<-auto.arima(tsdata)
18
19 'Prüfung durch Ljung-Box Test, ob man Prognose erstellen kann '
20 Box.test(AutoArima_0001_wueber65$residuals,type="Ljung-Box")
21
22 'Erstellung der Prognose'
23 forecast_0001_wueber65<-forecast(AutoArima_0001_wueber65,h=12)
24
25 plot(forecast_0001_wueber65)
26 plot(forecast_0001_wueber65,xlim=c(2015,2024))
27 summary(forecast_0001_wueber65) 'MAPE-Wert erstellen und ablesen'

```

Bild 5: Code zum Datensatz 001 bezüglich der über 65-jährigen weiblichen Senioren

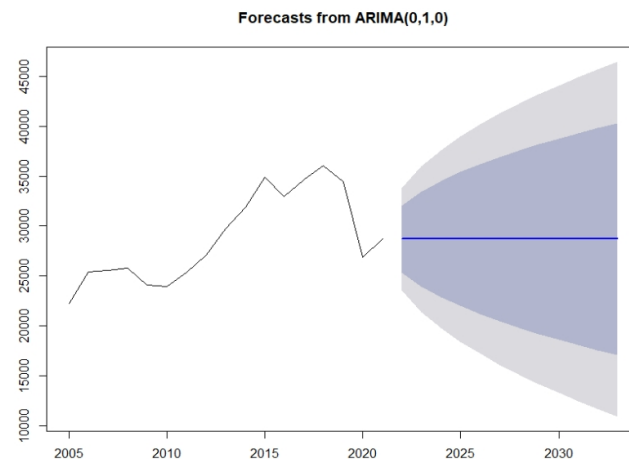


Bild 6: Prognose aus dem Datensatz 001 bezüglich der über 65-jährigen männlichen Senioren

6. Ergebnisse und Diskussion

Die Interpretation der erstellten Prognosen sollen Handlungsempfehlungen für die Politik und andere Institutionen im Gesundheitswesen vermitteln. Die Diskussion bezieht sich auf die Stichhaltigkeit der ARIMA-Modelle, die durch die Auto-ARIMA-Funktion erstellt wurden. Hierzu wird neben dem Ljung-Box Test noch der MAPE-Wert als Güte für die Prognosen genutzt, um die Prognose-Genauigkeit noch weiter zu untermauern [19]. In R wird dafür der Befehl `summary()` verwendet, welches den MAPE-Wert der Prognose ausgibt.

6.1. Ergebnisse

Zunächst ist anzumerken, dass alle untersuchten Regressionsmodelle die statistische Signifikanz des Ljung-Box Testes erfüllen. Somit beinhalten alle erstellten Prognosen eine ausreichende Aussagekraft für den betrachteten Zeitraum von 2022 bis 2024.

6.1.1. Datensatz 22121-0001 für weibliche Senioren

Das Bild 5 aus dem Kapitel 5 gibt die Prognose bezüglich diesem Datensatz wieder. Hierbei ist zu erkennen, dass die Lebensunterhaltsempfänger der weiblichen Senioren von ca. 100.000 im Jahre 2020 bis ca. 120.000 im Jahre 2024 steigen wird. Da für das ARIMA-Modell (1,0,0) gilt, werden die Fehler aus der ersten Ableitung verwendet. Zudem liegt der MAPE-Wert bei 7.37965, was eine hohe Prognose-Genauigkeit widerspiegelt.

6.1.2. Datensatz 22121-0001 für männliche Senioren

Diese Prognose besagt, dass sich die Empfängerzahl der männlichen Senioren bei über 25.000 konstant halten wird. Hierbei handelt es sich um einen einfachen gleitenden Durchschnitt (Single Moving Average) und einem Randomwalk-Modell.

Das Randomwalk-Modell beinhaltet die Tatsache, dass die Werte zufällig aber unabhängig voneinander entstanden sind. Außerdem liegt der MAPE-Wert bei 6.590725, was eine hohe Prognose-Genauigkeit widerspiegelt [7,15].

6.1.3. Datensatz 23211-0004 für weibliche Senioren 65 bis unter 70 Jahre:

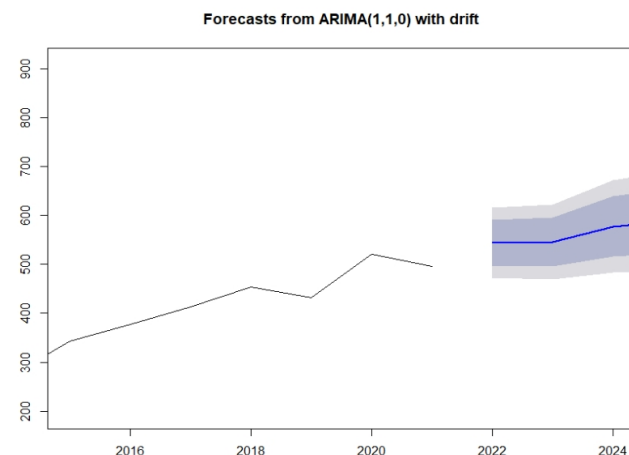


Bild : Prognose aus dem Datensatz 004 der 65 bis unter 70-jährige weibliche Senioren

Hier sieht man einen Anstieg der psychischen Erkrankungen von ca. 480 bis ca. 580 Personen, was durch den Drift beschrieben wird [20]. Somit wird also die Steigerung durch einen Trend begründet, welcher unter anderem durch die historischen Werte hervorgesagt wurden. Zudem liegt der MAPE-Wert bei 8.218079, was eine hohe Prognose-Genauigkeit widerspiegelt.

70 bis unter 75 Jahre:

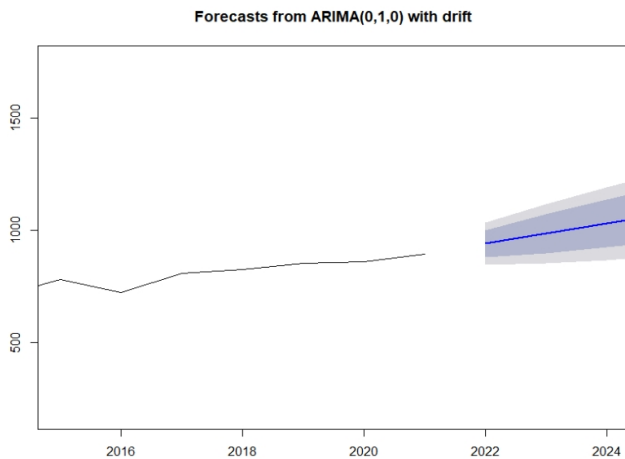


Bild : Prognose aus dem Datensatz 004 der 70 bis unter 75-jährige weibliche Senioren

Hier sieht man einen Anstieg der psychischen Erkrankungen von ca. 800 bis über 1000 Personen, was durch den Drift beschrieben wird. Somit wird also die Steigerung durch einen Trend begründet, welcher unter anderem durch die historischen Werte hervorgesagt wurde. Zudem liegt der MAPE-Wert bei 6.916871, was eine hohe Prognose-Genauigkeit widerspiegelt.

75 bis unter 80 Jahre:

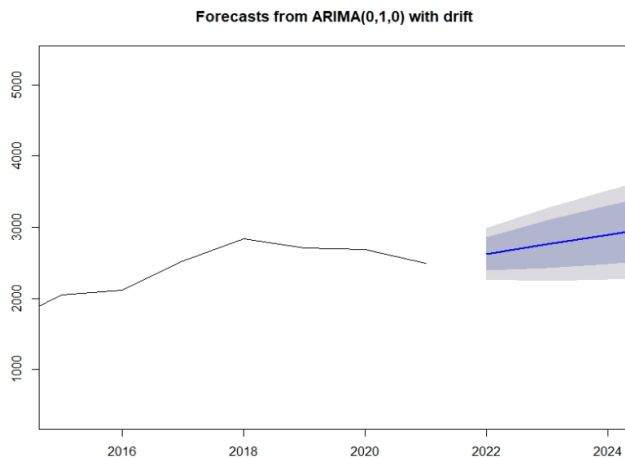


Bild : Prognose aus dem Datensatz 004 der 75 bis unter 80-jährige weibliche Senioren

Hier sieht man einen Anstieg der psychischen Erkrankungen von ca. 2500 bis ca. 3000 Personen, was durch den Drift beschrieben wird. Somit wird also die Steigerung durch einen Trend begründet, welcher unter anderem durch die historischen Werte hervorgesagt wurde. Zudem liegt der MAPE-Wert bei 7.798129, was eine hohe Prognose-Genauigkeit widerspiegelt.

80 bis unter 85 Jahre:

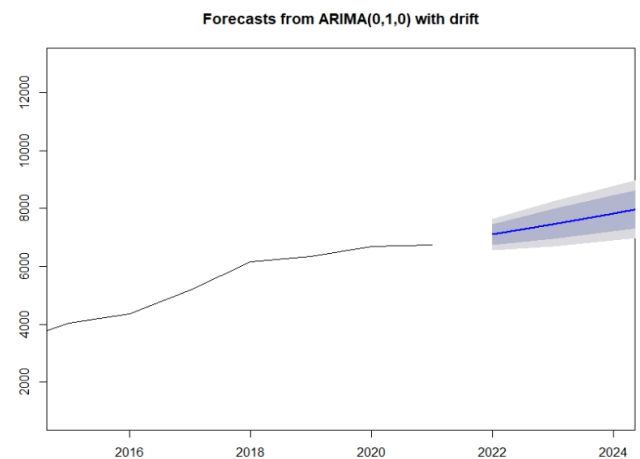


Bild : Prognose aus dem Datensatz 004 der 80 bis unter 85-jährige weibliche Senioren

Hier sieht man einen Anstieg der psychischen Erkrankungen von ca. 7000 bis ca. 8000 Personen, was durch den Drift beschrieben wird. Somit wird also die Steigerung durch einen Trend begründet, welcher unter anderem durch die historischen Werte hervorgesagt wurde. Zudem liegt der MAPE-Wert bei 5.951762, was eine hohe Prognose-Genauigkeit widerspiegelt.

85 Jahre und mehr:

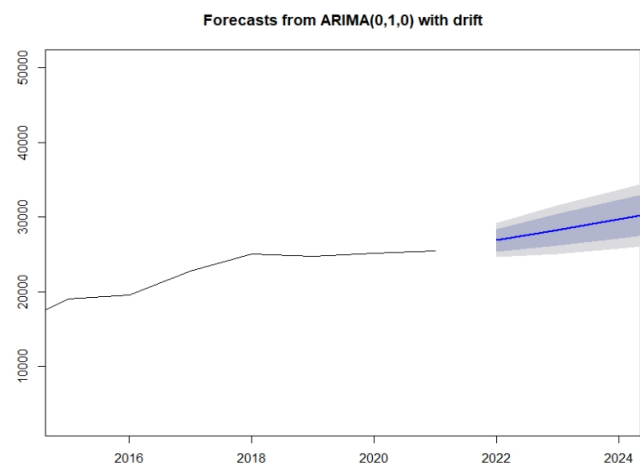


Bild : Prognose aus dem Datensatz 004 der 85-jährige und ältere weibliche Senioren

Hier sieht man einen Anstieg der psychischen Erkrankungen von über 25.000 bis ca. 30.000 Personen, was durch den Drift beschrieben wird. Somit wird also die Steigerung durch einen Trend begründet, welcher unter anderem durch die historischen Werte hervorgesagt wurde. Der MAPE-Wert liegt hier bei 5.462946, was eine hohe Prognose-Genauigkeit wiedergibt.

6.1.4. Datensatz 23211-0004 für männliche Senioren

65 bis unter 70 Jahre:

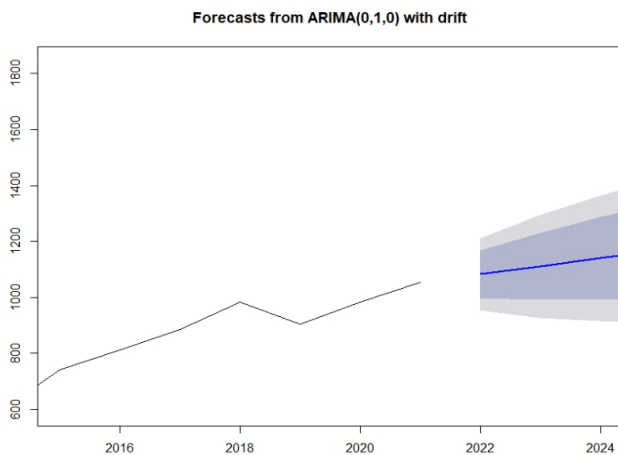


Bild : Prognose aus dem Datensatz 004 der 65 bis unter 70-jährigen männlichen Senioren

Bei dieser Prognose kann man einen Anstieg der psychischen Erkrankungen von ca. 1000 bis ca. 1200 Personen beobachten. Der Drift beschreibt also einen Trend, der unter anderem auf den historischen Daten basiert. Zudem liegt der MAPE-Wert bei 7.031265, was eine hohe Prognose-Genauigkeit widerspiegelt.

70 bis unter 75 Jahre:

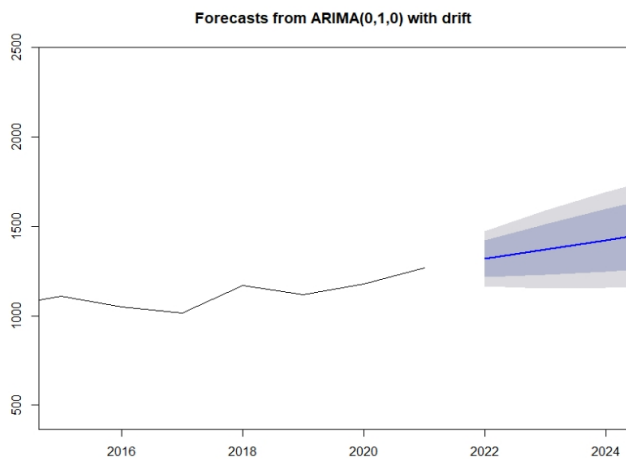


Bild : Prognose aus dem Datensatz 004 der 70 bis unter 75-jährigen männlichen Senioren

Bei dieser Prognose kann man einen Anstieg der psychischen Erkrankungen von ca. 1200 bis ca. 1500 Personen beobachten. Der Drift beschreibt also einen Trend, der unter anderem auf den historischen Daten basiert. Zudem liegt der MAPE-Wert bei 5.683828, was eine hohe Prognose-Genauigkeit widerspiegelt.

75 bis unter 80 Jahre:

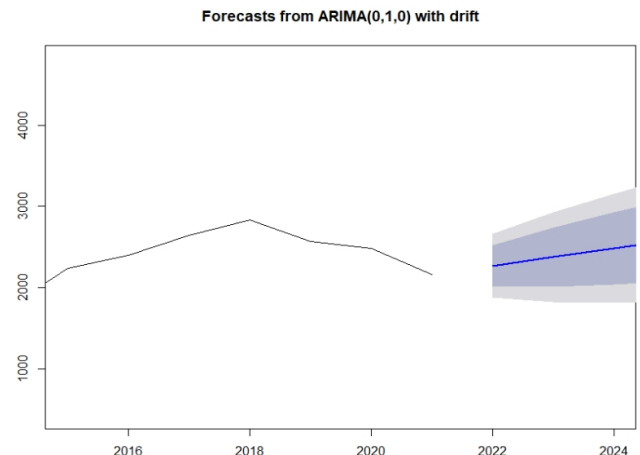


Bild : Prognose aus dem Datensatz 004 der 75 bis unter 80-jährigen männlichen Senioren

Bei dieser Prognose kann man einen Anstieg der psychischen Erkrankungen von ca. 2200 bis über 2500 Personen beobachten. Der Drift beschreibt also einen Trend, der unter anderem auf den historischen Daten basiert. Zudem liegt der MAPE-Wert bei 6.228493, was eine hohe Prognose-Genauigkeit widerspiegelt.

80 bis unter 85 Jahre:

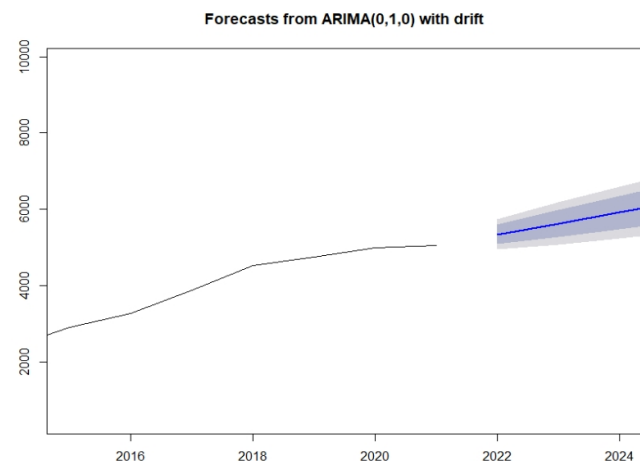


Bild : Prognose aus dem Datensatz 004 der 80 bis unter 85-jährigen männlichen Senioren

Bei dieser Prognose kann man einen Anstieg der psychischen Erkrankungen von ca. über 5000 bis ca. über 6000 Personen beobachten. Der Drift beschreibt also einen Trend, der unter anderem auf den historischen Daten basiert. Zudem liegt der MAPE-Wert bei 6.454541, was eine hohe Prognose-Genauigkeit widerspiegelt.

85 Jahre und mehr:

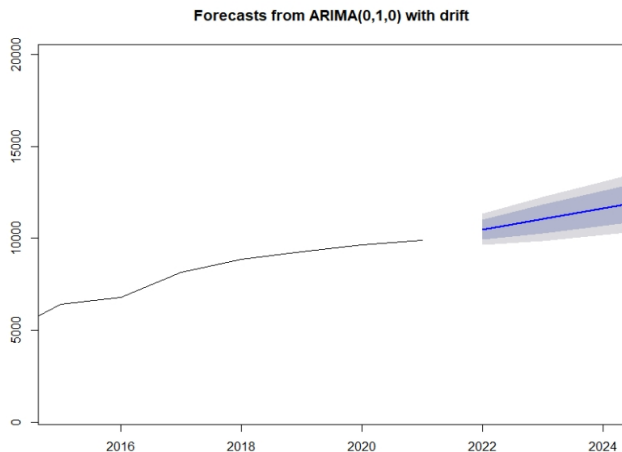


Bild : Prognose aus dem Datensatz 004 der 85 jährigen und älteren männlichen Senioren

Bei dieser Prognose kann man einen Anstieg der psychischen Erkrankungen von ca. über 10.000 bis ca. über 12.000 Personen beobachten. Der Drift beschreibt also einen Trend, der unter anderem auf den historischen Daten basiert. Zudem liegt der MAPE-Wert bei 7.123139, was eine hohe Prognosegenauigkeit widerspiegelt.

6.2. Auswertung

In diesem Abschnitt werden nun die Ergebnisse aus den Prognosen der Regressionsmodelle ausgewertet. Die Prognosegenauigkeit ist bei allen Regressionsmodellen der jeweiligen Datensätze sehr hoch, da der MAPE-Wert jeweils kleiner als 10 ist.

6.2.1. Prognosen aus dem Datensatz 22121-001

Bei den weiblichen Senioren ab 65 Jahren ist aufgefallen, dass eine Steigerung der Lebensunterhaltsempfänger von 20% bis 2024 vorherrschen wird. Kurzum sind davon also 120.000 weibliche Senioren betroffen. Bei den männlichen Senioren stagniert die Anzahl der Lebensunterhaltsempfänger bei 25.000 und bleibt somit konstant. Im Allgemeinen kann man also hier nicht erkennen, in welchem Alter die Senioren besonders von der finanziellen Unterstützung abhängig sind.

6.2.2. Prognosen aus dem Datensatz 23211-0004

Generell ist bei den weiblichen Senioren festzustellen, dass je älter die Senioren werden desto stärker steigt die Anzahl der psychisch erkrankten weiblichen Senioren an. In fast allen Regressionsmodellen existiert jeweils eine Zunahme von ca. 20%. Hierbei ist im Alter von 80 bis unter 85 Jahren die geringste Zunahme mit 14,29% verzeichnet. Gleichwohl ist im Alter von 70 bis unter 75 Jahren der größte Zuwachs mit ca. 25% vorhanden, wobei dieser ab 80 Jahren wieder auf 20% abfällt.

Insbesondere nimmt der MAPE-Wert der Regressionsmodelle regelmäßig ab, je älter die betroffenen Senioren sind. Dadurch wird also auch die Genauigkeit der Erkenntnis weiter bestärkt, dass die Anzahl der psychisch erkrankten Senioren

mit dem Alter konstant zunimmt.

Bei den männlichen Senioren ist ein identisches Verhalten zu beobachten, wobei die Anzahl der psychisch erkrankten männlichen Senioren nicht so steil ansteigt wie die der weiblichen Senioren. Ab dem Alter von 75 Jahren überholen die Erkrankungswerte der weiblichen die der männlichen Senioren, wobei sich der Unterschied zunächst auf ca. 14%, dann auf ca. 30% und dann schließlich auf über 100% beläuft. Analog zu den weiblichen Senioren existiert die gemeinsame Auffälligkeit, dass im Alter von 75 bis unter 80 Jahren ein kurzweiliger Anstieg der psychischen Erkrankungen um 25% erfolgt.

Der größtenteils stetig absinkende MAPE-Wert ist aus den Prognosen der männlichen Senioren nicht zu abzulesen.

7. Related work

Nicht nur dieses Paper bestärkt die Tatsache, dass die AUTO-ARIMA Funktion eine schnelle und effektive Methode zur Prognosenerstellung darstellt. Insbesondere zeigt der aktuelle Forschungsstand eine stärkere Tendenz zur Nutzung der AUTO-ARIMA Funktion, da dadurch der Prognoseprozess stark vereinfacht wird. So muss der Anwender keine tiefen Kenntnisse über die ARIMA-Familie oder die Box-Jenkins - Methoden mitbringen und kann so einen praktischen Überblick über die zukünftigen Werte erhalten. Ergänzend lässt sich aber noch anmerken, dass eine zusätzliche Prognosequalität durch nicht automatisierte Funktionen in R verbessert werden kann. Hierzu kann man Funktionen wie `arima()`, `ucm()` und insbesondere das `rucm`-package dafür verwenden [22].

8. Diskussion und Fazit

In diesem Abschnitt gehen wir auf das endgültige Fazit ein und wie die zuvor aufgestellten Thesen nun zu beurteilen sind.

8.1. These 1 - Steigerung der Sozialhilfeempfänger von weiblichen und männlichen Senioren ab 65 Jahren bis 2024

Die Steigerung der Sozialhilfeempfänger der weiblichen Senioren ab 65 Jahren ist bis zum Jahr 2024 gegeben und wird durch die hohe Prognosegenauigkeit bestätigt. Die Prognose der männlichen Senioren ab 65 Jahren besagt, dass eine Stagnation der Sozialhilfeempfänger bis 2024 vorherrscht. Somit konnte also nur die Steigerung der weiblichen Senioren als Sozialhilfeempfänger belegt werden. Für die männlichen Senioren kann diese These nicht belegt werden.

8.2. These 2 - Steigerung der psychischen Erkrankungen von weiblichen und männlichen Senioren ab 65 Jahren bis 2024

Eine Steigerung der psychischen Erkrankungen ist für weibliche als auch für männliche Senioren ab 65 Jahren bis 2024 gegeben und wird durch die hohe Genauigkeit aller Prognosen untermauert. Hierbei ist anzumerken, dass weibliche Senioren stärker von psychischen Erkrankungen betroffen sind.

8.3. These 3 - Nachvollziehbare Unterschiede zwischen weiblichen und männlichen Senioren ab 65 Jahre

Die Unterschiede zwischen weiblichen und männlichen Senioren sind vor allem bei den psychischen Erkrankungen klar zu erkennen. Insbesondere steigt ab dem Alter von 75 Jahren die Anzahl der psychischen Erkrankungen bei weiblichen Senioren rasant an und überholt die der männlichen. Bei den Sozialempfängerzahlen ist nur jener Unterschied festzuhalten, dass eine Steigerung der weiblichen Sozialhilfeempfänger bis 2024 gegeben ist. Durch die Stichhaltigkeit und der hohen Prognosegenauigkeit sind somit diese Unterschiede verifiziert.

Nach einer Studie des SOEP entsteht eine natürliche Persönlichkeitsänderung von Senioren ab 70 Jahren, welche nachweislich nicht durch Gesundheitsveränderungen, den Rentenbeginn oder die Großelternschaft hervorgerufen werden [21]. Durch die Forschung ist außerdem bekannt, dass die soziale Interaktion mit Menschen ein Schutz für die mentale Gesundheit darstellt. Die fehlende gesellschaftliche Akzeptanz von Senioren, die erzwungende Abhängigkeit im Alter und auch die körperlichen Beeinträchtigungen können mit hoher Wahrscheinlichkeit trotz allem zu psychischen Erkrankungen führen [23].

Zusammenfassend ist nun zu sagen, dass vor allem eine soziale Anbindung der Senioren ab 65 Jahren und insbesondere für die weibliche Senioren sehr wichtig ist. Hierbei sollte nicht nur die Altenpflege, Seniorenvereine oder die Rente ausgebaut werden. Ergänzend sollte auch eine weitreichende ethische Aufklärung bereits in der Schule stattfinden, um eine angemessene Sensibilisierung für die Belange von Senioren einzuleiten. Denn der vergangene Beitrag der älteren Mitbürger für die deutsche Gesellschaft sollte gerecht und mit Respekt zurück gezollt werden.

References

- [1] **Destatis Statistisches Bundesamt** (2023), URL: <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Soziales/Sozialhilfe/empfaenger-hilfe-pflege.html> , Sozialhilfe - Knapp 400 000 Personen erhielten im Laufe des Jahres 2021 Hilfe zur Pflege, zuletzt zugegriffen am 12.09.2023.
- [2] **Bundesministerium für Familien, Senioren, Frauen und Jugend - BMFSFJ** (2020), <https://www.bmfsfj.de/bmfsfj/service/publikationen/wenn-das-altwerden-zur-last-wird-96012>, Wenn das Altwerden zur Last wird - Suizidprävention im Alter, zuletzt zugegriffen am 12.09.2023.
- [3] **Hammer, Mirja** (2015), URL: <https://www.stern.de/gesundheit/suizid-im-alter-warum-bringen-sich-vor-allem-aeltere-um-6452916.html>, Suizid im Alter- Warum nehmen sich so viele ältere Menschen das Leben?, zuletzt zugegriffen am 12.09.2023.
- [4] **Korfmacher, Carsten** (2023), URL: <https://www.nordkurier.de/politik/wie-bricht-das-rentensystem-zusammen-1592870> , Wie bricht das Rentensystem zusammen?, zuletzt zugegriffen am 16.09.2023.
- [5] **Hyndman, Rob J., et al.** (2008) *Automatic Time Series Forecasting: The forecast Package for R* ,Journal of Statistical Software, July 2008, Volume 27, Issue 3..
- [6] **Yermal, Lakshmi, et al.** (2017), *Application of Auto ARIMA Model for Forecasting Returns on Minute Wise Amalgamated Data in NSE* ,2017 IEEE International Conference on Computational Intelligence and Computing Research.
- [7] **Moreno, Juan Jose Montano, et al.** (2013), *Using the R-MAPE index as a resistant measure of forecast accuracy*, *Psicothema* 2013, Vol. 25, No. 4, 500-506, Psicothema.
- [8] **Hirschle, Jochen** (2021), *Machine Learning für Zeitreihen: Einstieg in Regressions-, ARIMA- und Deep-Learning-Verfahren mit Python*, Carl Hanser Verlag München.
- [9] **Neusser, Klaus** (2009), *Zeitreihenanalyse in den Wirtschaftswissenschaften* , 2.aktualisierte Auflage, Vieweg+Teubner.
- [10] **Fahrmeir, Ludwig, et al.** (2016), *Statistik: Der Weg zur Datenanalyse* , 8.überarbeitete und ergänzte Auflage, Springer Verlag Berlin Heidelberg.
- [11] **Schlittgen, Rainer, et al.** (2020), *Angewandte Zeitreihenanalyse mit R* , 4., erweiterte und aktualisierte Auflage, Walter de Gruyter GmbH.
- [12] **Novustat-Webseite** (2019), URL: <https://novustat.com/statistik-blog/multivariate-statistik-beispiele.html> , Multivariate Statistik richtig nutzen: Einsatzmöglichkeiten und Anwendungsbeispiele , zuletzt zugegriffen am 21.05.2023.
- [13] **Schlittgen, Rainer** (2008), *Einführung in die Statistik: Analyse und Modellierung von Daten*, 11.vollständig überarbeitete und neu gestaltete Auflage, Oldenbourg Verlag München.
- [14] **Hyndman, Rob J., et al.** (2014), *Forecasting: Principles and Practice* , Published by Otexts.com .
- [15] **Schlittgen, Rainer, et al.** (2001), *Zeitreihenanalyse* , 9.unwesentlich veränderte Auflage, R. Oldenbourg Verlag München Wien.
- [16] **Statologie-Webseite** (2020), URL: <https://statologie.de/ljung-box-test/google-ignette-ljung-box-test-definition-beispiel>, zuletzt zugegriffen am 17.09.2023.
- [17] **Destatis statistisches Bundesamt-Webseite** (2023), URL: <https://www-genesis.destatis.de/genesis/online?operation=tablecode=23211-0004bypass=truelevelindex=0levelid=1695090245479abreadcrumb,23211-0004:Gestorbene:Deutschland,Jahre,Todesursachen,Geschlecht,Altersgruppen>, zuletzt zugegriffen am 17.09.2023.
- [18] **Destatis statistisches Bundesamt-Webseite** (2023) ,URL: <https://www-genesis.destatis.de/genesis/online?operation=tablecode=22121-0001bypass=truelevelindex=0levelid=1694992097715abreadcrumb,22121-0001:Empfänger von Hilfe zum Lebensunterhalt: Deutschland,Stichtag,Ort der Leistungserbringung,Geschlecht,Altersgruppen>, zuletzt zugegriffen am 17.09.2023.
- [19] **Hoegel, Bianca** (2023), URL: https://www.biancahoegel.de/mathe/stochastik/random_walk.html, Random Walk, zuletzt zugegriffen am 25.08.2023
- [20] **Prof. Dr. W. Zucchini, et al.** (2007), URL: https://www.uni-goettingen.de/de/document/download/bced8d2cf092e3487839c365955c85f2.pdf/ZRSkript_eil20kr09.pdf, Zeitreihenanalyse Teil II: Lineare Zeitreihenmodelle , zuletzt zugegriffen am 01.09.2023.
- [21] **Jule Specht, et al.** (2014), *On the Consistency of Personality Types Across Adulthood: Latent Profile Analyses in Two Large-Scale Panel Studies, SOEP — The German Socio-Economic Panel Study at DIW Berlin*.
- [22] **Rahardja, Dewi, et al.** (2023), *Univariate Time-Series Forecast Computing via R 'auto.arima()' Function*, *Journal of Research in Applied Mathematics* Volume 9, Issue 5 (2023) pp: 46-50.
- [23] **Landeszentrum Gesundheit Nordrhein-Westfalen** (2022), URL: www.lzg.nrw.de/10256159 , Psychische Erkrankungen im hohen Erwachsenenalter, zuletzt zugegriffen am 08.09.2023.