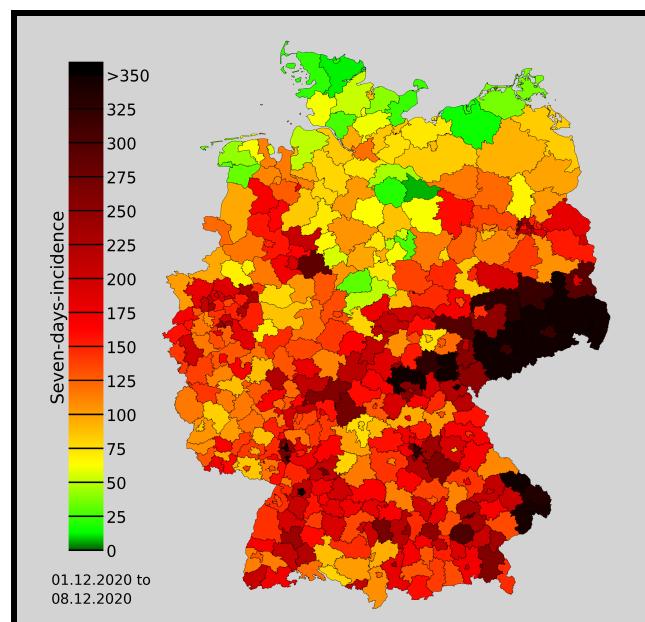


Analyse der Covid-19-Daten der deutschen Landkreise

Leander Marius Bürkin

Gutachter: Dr. Andreas Greiner



Professur für Simulation
Institut für Mikrosystemtechnik - IMTEK
Albert-Ludwigs-Universität, Freiburg im Breisgau

Bearbeitungszeit

16. 06. 2021 – 16. 09. 2021

Gutachter

Dr. Andreas Greiner

Zweitgutachter

Prof. Dr. Moritz Mathias Diehl

Betreuer

Dr. Andreas Greiner

Erklärung

Hiermit erkläre ich, dass ich diese Abschlussarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen/Hilfsmittel verwendet habe und alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen wurden, als solche kenntlich gemacht habe. Darüber hinaus erkläre ich, dass diese Abschlussarbeit nicht, auch nicht auszugsweise, bereits für eine andere Prüfung angefertigt wurde.

Ort, Datum

Unterschrift

Abstract

2020 and 2021 have been marked by COVID-19 and the ensuing uncertainty: How does the virus spread? Does the virus spread from the cities to rural areas? Can the counties whose case numbers skyrocketed earlier be categorized in any way? Or more generally: is the spread of Corona correlated with certain characteristics of a given area?

An attempt to answer these questions and to summarize the given Corona data is described in this thesis. In particular, the project uses data provided by the Robert Koch Institute (RKI) via ArcGIS.

On the basis of the processed data, an SIR model for Germany is created and several correlation analyses of Corona data from various administrative districts are carried out. Finally, the results are plotted and interpreted in order to make the elusive coronavirus infection patterns accessible.

Kurzfassung

Die Jahre 2020 und 2021 waren vom Coronavirus und der einhergehenden Unsicherheit geprägt: Wie verbreitet sich das Virus? Breitet sich das Virus von den Städten ausgehend in die ländlichen Gebiete aus? Lassen sich die Stadt- und Landkreise, deren Fallzahlen früher in die Höhe schießen, irgendwie kategorisieren? Oder ganz allgemein: Korreliert die Ausbreitung von Corona mit bestimmten Eigenschaften der Gebiete?

Ein Versuch, diese Fragen zu beantworten und die gegebenen Corona-Daten zusammenzufassen, ist in dieser Arbeit beschrieben. Insbesondere wird mit Daten, welche das Robert Koch-Institut (RKI) über ArcGIS zur Verfügung stellt, gearbeitet.

Auf Basis der aufgearbeiteten Daten wird ein SIR-Modell für Deutschland erstellt und mehrere Korrelationsanalysen von Corona-Daten die kreisfreien Städte, Landkreise und Regierungsbezirke durchgeführt. Zuletzt werden graphische Darstellungen der Ergebnisse erstellt und interpretiert, um das schwer greifbare Infektionsgeschehen des Coronavirus zugänglich zu machen.

*Those who cannot remember the past
are condemned to repeat it.* - George Santayana [7]

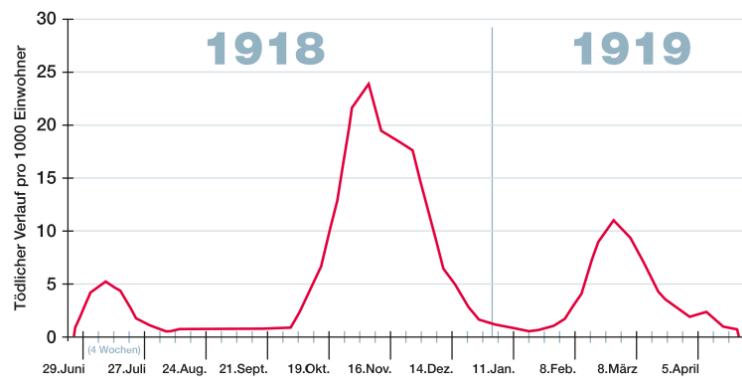


Abbildung 1: Die drei Wellen der „Spanischen Grippe“ in Großbritannien, wöchentliche kombinierte Grippe- und Lungenentzündungssterblichkeit von Juni 1918 bis Mai 1919. [1]

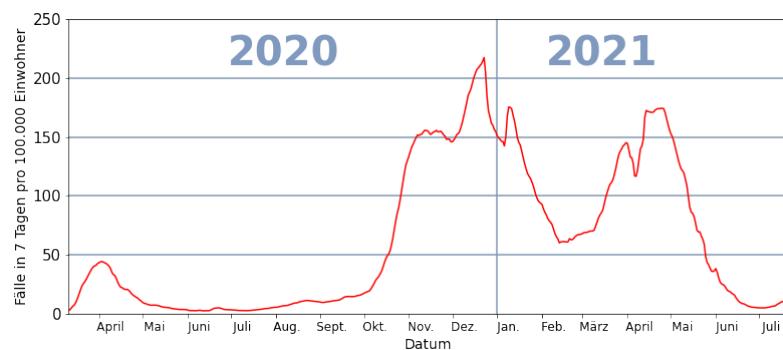


Abbildung 2: Die drei Wellen der COVID-19-Pandemie in Deutschland, dargestellt durch die sieben Tage Inzidenz jeden Tages vom 01.03.2020 bis 21.07.2021. Aus den Daten des Robert-Koch-Instituts erstellt.

Inhaltsverzeichnis

1 Motivation	1
2 Grundlagen	3
2.1 Susceptible-Infectious-Removed-Modell	3
2.2 Bevölkerungsdichte	4
2.3 7-Tages Inzidenz	4
2.4 Korrelationsanalyse mithilfe einer Faltung	5
2.4.1 Berechnung der Korrelationswerte	5
2.4.2 Korrelation zwischen zwei Gebieten am Beispiel von Flensburg und Kiel	8
2.4.3 Komprimierung und Darstellung als Matrizen	12
2.5 Durchschnitt, Farbgebung und Skalierung	13
3 Vorgehensweise	15
3.1 Datenquellen - Ursprung und Abspeicherung	15
3.2 Datenaufbereitung	17
3.3 Datenformat	19
3.4 Datendarstellung	23
3.4.1 SIR-Modell	23
3.4.2 Korrelationsanalyse	23
3.4.3 Korrelation einzelner ausgewählte Städte und Landkreise . . .	24

4 Resultate	28
4.1 Bevölkerungsdichten der Landkreise und Regierungsbezirke	28
4.2 Die Landkreise und Regierungsbezirke sortiert nach dem Gemeindeschlüssel	30
4.3 Summe der 7-Tages Inzidenzen der Landkreise und Regierungsbezirke	31
4.4 SIR-Modell für Deutschland	33
4.5 Korrelationen zwischen Städten und ihrem Umland	34
4.6 Korrelationsmatrizen sortiert nach der Bevölkerungsdichte	37
4.6.1 Korrelationsmatrizen mit den nach Bevölkerungsdichten sortierten Landkreisen	37
4.6.2 Korrelationsmatrizen mit den nach Bevölkerungsdichten sortierten Regierungsbezirken	39
4.7 Korrelationsmatrizen sortiert nach dem Gemeindeschlüssel	41
4.7.1 Korrelationsmatrizen mit den nach Gemeindeschlüsseln sortierten Landkreise	41
4.7.2 Korrelationsmatrizen mit den nach Gemeindeschlüsseln sortierten Regierungsbezirken	43
4.8 Durchschnittliche Verschiebung im nationalen Vergleich	45
4.8.1 Verteilung der durchschnittlichen Verschiebung unter den Landkreise	45
4.8.2 Verteilung der durchschnittlichen Verschiebung unter den Regierungsbezirken	46
5 Diskussion	49
5.1 Aussagekraft des SIR-Modells	49
5.2 Zusammenhänge von Bevölkerungsdichten und Corona-Zahlen der Landkreisen	50
5.3 Zusammenhänge von Bevölkerungsdichten und Corona-Zahlen der Regierungsbezirke	51

5.4	Zusammenhänge von Gemeindeschlüssel und Corona-Zahlen der Landkreise und Regierungsbezirke	54
5.5	Durchschnittliche Verschiebung im nationalen Vergleich Landkreise .	55
6	Zusammenfassung	56
7	Danksagung	58
8	Anhang	60
8.1	Die deutschen Landkreise sortiert nach ihrer Bevölkerungsdichte . . .	64
8.2	Die deutschen Landkreise lexikographisch sortiert nach ihren Gemeindeschlüsseln	83
8.3	Die Regierungsbezirke sortiert nach ihrer Bevölkerungsdichte	103
8.4	Die Regierungsbezirke lexikographisch sortiert nach den ersten beiden Teilen der Gemeindeschlüssel ihrer Landkreise	105
	Literaturverzeichnis	110

Abbildungsverzeichnis

1	Die drei Wellen der „Spanischen Grippe“ in Großbritannien, wöchentliche kombinierte Grippe- und Lungenentzündungssterblichkeit von Juni 1918 bis Mai 1919. [1]	iv
2	Die drei Wellen der COVID-19-Pandemie in Deutschland, dargestellt durch die sieben Tage Inzidenz jeden Tages vom 01.03.2020 bis 21.07.2021. Aus den Daten des Robert-Koch-Instituts erstellt.	iv
3	In Deutschland neu gemeldete COVID-19 Fälle am jeweiligen Wochentag im Vergleich zu den am Montag der selben Woche gemeldeten Fällen.	5
4	Beispielhafte Darstellung einer Faltung anhand der Folgen [1, 2, 3, 2] und [5, 7, 5, 1]. Auf der linken Seite hinter dem Gleichheitszeichen von oben nach unten die Korrelationswahrscheinlichkeiten für die negativen Verschiebungen $\tau = -3$, $\tau = -2$ und $\tau = -1$. Entsprechend in der Mitte bei keine Verschiebung ($\tau = 0$) und rechts für die positiven Verschiebungen $\tau = 1$, $\tau = 2$ und $\tau = 3$	6
5	Der Verlauf der 7-Tages Inzidenz des Stadtkreises Flensburg (Gemeindeschlüssel 1001).	9
6	Der Verlauf der 7-Tages Inzidenz des Stadtkreises Kiel (Gemeindeschlüssel 1002).	9
7	Caption	10
8	Caption	10

9	Caption	11
10	Das Farbspektrum, in welchem sich die Darstellungen bewegen. Von links nach rechts steigen die eingegebenen Werte konstant. Der angegebene Wert wird jeweils anhand des ersten und des letzten Werten einer Referenzliste linear in diesem Spektrum verortet.	14
11	Die Dictionaries non_county_specific_data, counties_geography und covid19 mit ihren jeweiligen Schlüsseln. Für die Dictionaries covid19 und counties_geography wurde jeweils der Landkreis Kiel (Gemeindeschlüssel 1002) zur Veranschaulichung verwendet.	22
12	Die ausgewählten Landkreise mit den Stadtkreisen, die sie jeweils umgeben.	26
13	Verteilung der Bevölkerungsdichten unter den deutschen Landkreisen.	29
14	Verteilung der Bevölkerungsdichten unter den deutschen Regierungsbezirken. Die Skalierung entspricht der Farbgebung in Abbildung 13.	30
15	Die Landkreise und Regierungsbezirke eingefärbt nach der lexikographischen Größe ihres Gemeindeschlüssels. Bei blau beginnend werden den nach ihrem Gemeindeschlüssel sortierten Landkreise die Farbfolge des Regenbogens zugeteilt.	31
16	Verteilung der Summe der 7-Tages Inzidenzen unter den deutschen Landkreisen.	32
17	Verteilung der Summe der 7-Tages Inzidenzen unter den deutschen Regierungsbezirken.	32
18	Die drei Kennzahlen des SIR-Modells für Deutschland in drei Graphen.	33
19	Die Korrelationswahrscheinlichkeiten der 7-Tages INzidenzen der ausgewählten Landkreise mit den Städten, die sie umgeben. Die Kurven sind je nach ihrem Hochpunkt eingefärbt: Rot bei positiven x-Werte, Blau bei negativen und grün bei 0.	35

- 22 Verschiedene Korrelationsmatrizen berechnet aus den Korrelationen jedes deutschen Regierungsbezirks mit sich selbst und jedem anderen deutschen Regierungsbezirks. Jeder Zeile und Spalte sind Regierungsbezirke zugeteilt, die vollständige Auflistung befindet sich im Anhang. Die Farben der Zellen auf der linken Seite entsprechen der Differenz zwischen der durchschnittlichen Korrelationswahrscheinlichkeit für negative Verschiebungen des Regierungsbezirks der Spalte in Relation zum Regierungsbezirk der Zeile und der durchschnittlichen Korrelationswahrscheinlichkeit für die zugehörigen positiven Verschiebungen. Auf der rechten Seite ist die Zelle entsprechend der Verschiebung der Zeitserie des Regierungsbezirks der Spalte entgegen der Zeitserie der Zeile mit der höchsten Korrelationswahrscheinlichkeit eingefärbt. Beide Vorgehensweise werden für alle ganzzahligen Verschiebungen $\tau \in [-14, 14]$, $\tau \in [-30, 30]$ und $\tau \in [-50, 50]$ durchgeführt und in dieser Reihenfolge von oben nach unten dargestellt.

24	Verschiedene Korrelationsmatrizen berechnet aus den Korrelationen jedes deutschen Regierungsbezirks mit sich selbst und jedem anderen deutschen Regierungsbezirks. Jeder Zeile und Spalte sind Regierungsbezirke zugeteilt, die vollständige Auflistung befindet sich im Anhang. Die Farben der Zellen auf der linken Seite entsprechen der Differenz zwischen der durchschnittlichen Korrelationswahrscheinlichkeit für negative Verschiebungen des Regierungsbezirks der Spalte in Relation zum Regierungsbezirk der Zeile und der durchschnittlichen Korrelationswahrscheinlichkeit für die zugehörigen positiven Verschiebungen. Auf der rechten Seite wird die Zelle entsprechend der Verschiebung der Zeitserie des Regierungsbezirks der Spalte entgegen der Zeitserie der Zeile mit der höchsten Korrelationswahrscheinlichkeit eingefärbt. Beide Vorgehensweise werden für alle ganzzahligen Verschiebungen $\tau \in [-14, 14]$, $\tau \in [-30, 30]$ und $\tau \in [-50, 50]$ durchgeführt und in dieser Reihenfolge von oben nach unten dargestellt.	44
25	Auf der rechten Seite die deutschen Landkreise eingefärbt im Durchschnitt der Maximalwerte der Korrelationen mit allen Landkreisen und Verschiebungen $\tau \in [-50, 50]$. Auf der linken Seite die Verteilung dieser Werte.	45
26	Auf der rechten Seite die deutschen Landkreise eingefärbt im Durchschnitt der Differenz der Korrelationswahrscheinlichkeiten der positiven Verschiebungen zu den Korrelationswahrscheinlichkeiten der negativen Verschiebungen der Korrelationen dieses Landkreises mit allen Landkreisen und Verschiebungen $\tau \in [-50, 50]$. Auf der linken Seite die Verteilung dieser Werte.	46
27	Auf der rechten Seite die deutschen Regierungsbezirke eingefärbt im Durchschnitt der Maximalwerte der Korrelationen mit allen Regierungsbezirken und Verschiebungen $\tau \in [-50, 50]$. Auf der linken Seite die Verteilung dieser Werte.	47

28	Auf der rechten Seite die deutschen Regierungsbezirke eingefärbt im Durchschnitt der Differenz der Korrelationswahrscheinlichkeiten der positiven Verschiebungen zu den Korrelationswahrscheinlichkeiten der negativen Verschiebungen der Korrelationen dieses Regierungsbezirks mit allen Regierungsbezirken und Verschiebungen $\tau \in [-50, 50]$. Auf der linken Seite die Verteilung dieser Werte.	48
29	Die Deutschlandkarte, rot eingefärbt jeweils die Landkreise mit einer 7-Tages Inzidenz am 06.10.2020 beziehungsweise 07.10.2020 über 50.	52
30	Die 7-Tages Inzidenzen der deutschen Landkreise, in Blau die 7-Tages Inzidenz von Cloppenburg.	53
31	Der Verlauf der 7-Tages Inzidenzen der deutschen Regierungsbezirke. Der Verlauf des Regierungsbezirks Trier ist blau hervorgehoben. . . .	54

Tabellenverzeichnis

1	Landkreise mit Name und Gemeindeschlüssel (AGS), die den jeweils mit Name und Gemeindeschlüssel gekennzeichneten Stadtkreis komplett umgeben.	25
---	---	----

1 Motivation

Viele große Zivilisationen hatten bereits mit Pandemien ähnlich zur aktuellen COVID-19 Pandemie zu kämpfen.

Schon aus dem Alten Rom haben wir Zeugnisse der Antoninischen Pest im zweiten Jahrhundert nach Christus: Vermutlich ein Pocken Ausbruch, dessen 7-10 Millionen Tote das Römische Reich destabilisierten ([5], S. 255). Im Mittelalter kamen 200 Millionen Menschen durch die Pest um [4]. Auch die Armeen des ersten Weltkriegs wurden durch eine Pandemie stark geschwächt: die Spanische Grippe forderte circa 100 Millionen Opfer weltweit [6].

Daten in Relation zur Bevölkerung haben wir beispielsweise aus Island: nach 1707 starben dort innerhalb von zwei Jahren ein Drittel, also circa 18.000 der 50.000 Einwohner, an Pocken ([3], Seite 325 Zeile 20ff). Die Auswirkungen auf die Azteken, die Inka und die anderen Völker der Neuen Welt, welche vermutlich ein ähnlich vernichtender Pockenausbruch traf, und ihr anschließender Niedergang sind uns ebenfalls wohl bekannt [3].

Um derartige Schicksale in unserer globalisierten Welt zu verhindern ist es unabdingbar, diese Bedrohungen bestmöglich zu untersuchen. Jedoch wissen wir über diese vergangenen Pandemien recht wenig, da nicht nach heutigen wissenschaftlichen Standards Daten erhoben und Schlussfolgerungen gezogen wurden. Beispielsweise wurden die Pandemien durch den Zorn Gottes erklärt ([3], Seite 324 Zeile 7) und die Zahlen der Zeitzeugen spiegeln eher die subjektive Wahrnehmung der Zeitzeugen wieder als wissenschaftliche Daten zu liefern ([3], Seite 324 Zeile 20-24). Mit dem heutigen Stand

der Datenerfassung und -übermittlung lässt sich die COVID-19 Pandemie sehr viel einfacher untersuchen als ihre Vorgänger.

Die meisten von uns haben genau diese Zahlen des Robert Koch-Instituts (www.rki.de) oder der Weltgesundheitsorganisation (www.who.int) beobachtet und versucht zu verstehen, wie man sich zu Verhalten hat. Viele Hypothesen wurden aufgestellt und man musste sich als Laie auf einmal mit Fragen und Begriffen der Epidemiologie beschäftigen, wobei man meist nur ein Puzzleteil vor Augen hatte.

Daher soll diese Arbeit zum Verständnis von Pandemien beitragen, damit die Ausbreitung zukünftiger Erreger verlangsamt wird und der Kollaps des Gesundheitssystems oder der gesamten Zivilisation gar nicht erst im Bereich des Möglichen liegen.

2 Grundlagen

2.1 Susceptible-Infectious-Removed-Modell

Ein Weg zur Beschreibung der Pandemie bietet das „Susceptible-Infectious-Removed-Modell“ (SIR-Modell) [8]. Dieses Modell teilt die Mitglieder einer Menschengruppe in eine der drei folgenden Kategorien ein und ermöglicht es, die zeitliche Entwicklung einer Pandemie übersichtlich darzustellen:

- „Susceptible“: Menschen, welche angesteckt werden können.
- „Infectious“: Infizierte Menschen, welche weitere Menschen anstecken können.
Werden auch als „die aktiven Fälle“ bezeichnet.
- „Removed“: Menschen, welche in die Kategorie „Infectious“ fielen, nun keine weiteren Menschen mehr anstecken können und aus dem Infektionsgeschehen entfernt wurden, in diesem Fall Genesene und Verstorbene.

In dem hier verwendeten SIR-Modell wird weder die Geburten- noch die Sterberate beachtet, die angenommen Bevölkerung bleibt durchgehend konstant.

Zudem wird angenommen, dass jedes Individuum nur einmal infiziert werden kann und die Infektion die einzige Möglichkeit darstellt, von der Kategorie „Susceptible“ in die Kategorie „Removed“ zu wechseln. Somit fällt die Zahl der Menschen in der Kategorie „Susceptible“ monoton und die Zahl der Menschen in der Kategorie „Removed“ steigt monoton an.

2.2 Bevölkerungsdichte

Die Bevölkerungsdichte eines Gebietes ρ_{Gebiet} wird berechnet, indem die Anzahl der Einwohner im Gebiet $n_{Einwohner\ Gebiet}$ durch die Fläche des Gebiets A_{Gebiet} geteilt wird:

$$\rho_{Gebiet} = \frac{n_{Einwohner\ Gebiet}}{A_{Gebiet}} \quad (1)$$

2.3 7-Tages Inzidenz

Die 7-Tages Inzidenz i_t wird für jeden Tag t seit Beginn der Pandemie berechnet. Zuerst wird von der akkumulierte Zahl der Fälle am gewählten Tag f_t die akkumulierte Zahl der Fälle sieben Tage zuvor f_{t-7} abgezogen, dies ergibt die neu hinzugekommenen Fälle innerhalb von sieben Tagen.

Schlussendlich wird diese Zahl durch die Anzahl der Bewohner des Gebiets p_{Gebiet} geteilt und mit 100.000 multipliziert. Dies ergibt Gleichung 2.

$$i_t = \frac{f_t - f_{t-7}}{p_{Gebiet}} \cdot 100.000 \quad (2)$$

Die 7-Tages Inzidenz bietet sich im Vergleich zur täglichen Fallzahl aus mehreren Gründen als Kennzahl für die Verbreitung eines Virus in einem Gebiet an:

Wie in Abbildung 3 klar zu sehen, werden an Wochenenden im Schnitt deutlich weniger neue Fälle registriert als an den anderen Wochentagen. Daher werden immer sieben Tage in der 7-Tages Inzidenz zusammengefasst.

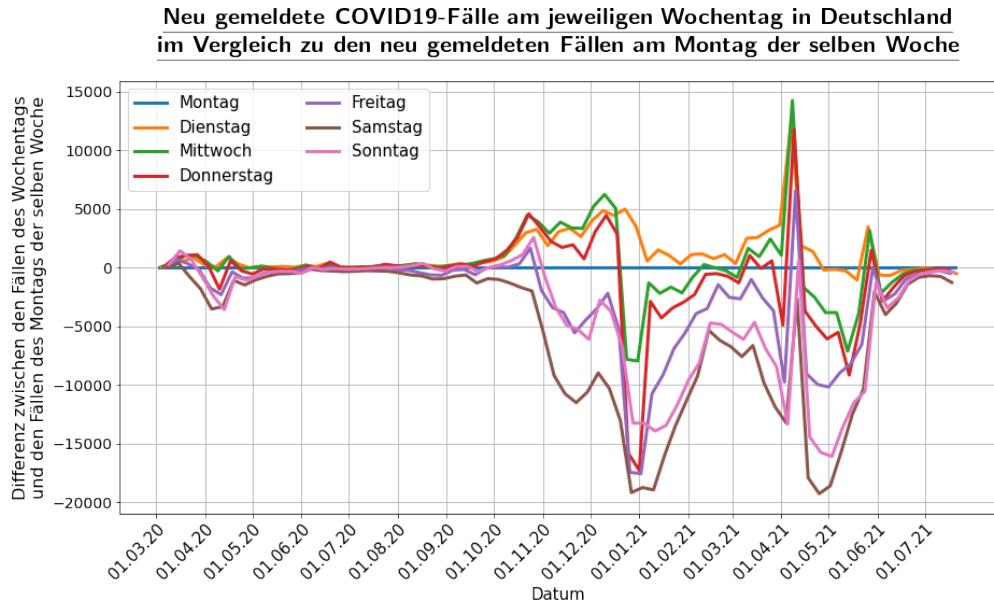


Abbildung 3: In Deutschland neu gemeldete COVID-19 Fälle am jeweiligen Wochentag im Vergleich zu den am Montag der selben Woche gemeldeten Fälle.

Da die Bevölkerung der deutschen Landkreise und Regierungsbezirke nicht identisch ist, wird jeweils durch die Bevölkerungszahl geteilt, um die einzelnen Gebiete miteinander vergleichen zu können.

Aufgrund der ansonsten sehr kleinen Zahlen, bietet es sich zudem an, das Ergebnis mit 100.000 zu multiplizieren.

Dies sind die drei Begründungen für die drei Schritte in der Berechnung der 7-Tages Inzidenz nach Gleichung 2.

2.4 Korrelationsanalyse mithilfe einer Faltung

2.4.1 Berechnung der Korrelationswerte

Um festzustellen, ob die 7-Tages Inzidenzen einiger Landkreise im Vergleich zu anderen Landkreisen eher voraus- oder nacheilen, wird die Korrelationsfunktion verwendet [2].

Bei diskreten Werten, aufgeteilt in zwei Zeitserien X und Y , wie in diesem Fall, lässt sich eine Korrelationsanalyse sehr einfach umsetzen: Für eine zeitliche Verschiebung τ wird mit jedem Wert x_i zum jeweiligen Zeitpunkt t_i aus der ersten Zeitserie mit dem zugehörigen Wert y_i aus der zweiten Zeitserie ein Produkt gebildet. Der zugehörige Wert aus der zweiten Zeitserie entspricht hierbei dem Zeitpunkt t des Wertes der ersten Zeitserie plus die gewählte Verschiebung τ . Sollte dieser zweite Wert nicht existieren, wird kein Produkt gebildet.

Für jede zeitliche Verschiebung τ , für die mindestens ein Produkt gebildet wird, werden alle möglichen Produkte aufsummiert.

Somit ergibt sich Gleichung 3, mit $n :=$ Länge von X und wenn $y_{i+\tau} \notin Y$, dann $y_{i+\tau} := 0$:

$$c(\tau) = \sum_{i=1}^n x_i \cdot y_{i+\tau} \quad (3)$$

Bildlich gesprochen wird die zweite Zeitserie an der ersten Zeitserie vorbeigeschoben, beginnend an dem Punkt, an dem ausschließlich das erste Element der ersten Zeitserie mit dem letzten Element der zweiten Zeitserie multipliziert wird. Dies ist beispielhaft mit den Folgen $[1, 2, 3, 2]$ und $[5, 7, 5, 1]$ in Abbildung 4 dargestellt.

$\begin{matrix} [1 & , & 2 & , & 3 & , & 2] \\ * & & & & & & \\ [5 & , & 7 & , & 5 & , & 1] \\ 1 = 1 \end{matrix}$	$\begin{matrix} [1 & , & 2 & , & 3 & , & 2] \\ * & * & * & * & * \\ [5 & , & 7 & , & 5 & , & 1] \\ 10 + 21 + 10 = 41 \end{matrix}$	$\begin{matrix} [1 & , & 2 & , & 3 & , & 2] \\ * & * & * & * \\ [5 & , & 7 & , & 5 & , & 1] \\ 15 + 14 + 2 = 29 \end{matrix}$
$\begin{matrix} [1 & , & 2 & , & 3 & , & 2] \\ * & * & * \\ [5 & , & 7 & , & 5 & , & 1] \\ 5 + 2 = 7 \end{matrix}$	$\begin{matrix} [1 & , & 2 & , & 3 & , & 2] \\ * & * & * & * \\ [5 & , & 7 & , & 5 & , & 1] \\ 5 + 14 + 15 + 2 = 36 \end{matrix}$	$\begin{matrix} [1 & , & 2 & , & 3 & , & 2] \\ * & * \\ [5 & , & 7 & , & 5 & , & 1] \\ 7 + 10 + 3 = 20 \end{matrix}$
		$\begin{matrix} [1 & , & 2 & , & 3 & , & 2] \\ * \\ [5 & , & 7 & , & 5 & , & 1] \\ 10 = 10 \end{matrix}$

Abbildung 4: Beispielhafte Darstellung einer Faltung anhand der Folgen $[1, 2, 3, 2]$ und $[5, 7, 5, 1]$. Auf der linken Seite hinter dem Gleichheitszeichen von oben nach unten die Korrelationswahrscheinlichkeiten für die negativen Verschiebungen $\tau = -3, \tau = -2$ und $\tau = -1$. Entsprechend in der Mitte bei keine Verschiebung ($\tau = 0$) und rechts für die positiven Verschiebungen $\tau = 1, \tau = 2$ und $\tau = 3$

Da bei einer Korrelationsanalyse nach Gleichung 3 Summen aus mehr Produkten übergewichtet werden und Zeitserien mit größeren Werten größere Korrelationswerte erzeugen, müssen die Korrelationswerte noch skaliert werden.

Um zum einen die unterschiedliche Anzahl der Produkte auszugleichen, werden die Summen durch die Anzahl ihrer Summanden geteilt, wie in Gleichung 4 gezeigt. Ohne diese Gewichtung würden zwei Zeitserien mit konstanten Werten größer null bei einer Verschiebung $\tau = 0$ die größte Korrelation aufweisen und die Korrelation bei betragsmäßig größeren Verschiebungen abnehmen, was nicht gewünscht ist.

Da jede Zeitserie die gleiche Länge n hat, erhält man die Anzahl der Summanden indem man den Betrag der Verschiebung $|\tau|$ von der Länge der Zeitserie n abzieht:

$$c(\tau) = \frac{1}{n - |\tau|} \sum_{i=1}^n x_i \cdot y_{i+\tau} \quad (4)$$

Um zum anderen die tendenziell größeren Werte mancher Zeitserien auszugleichen, werden die Werte aller Zeitserien mithilfe des sogenannten „Autokorrelationswerts für die Verschiebung $\tau = 0$ “ gewichtet: Er beschreibt den Wert der Korrelation einer Zeitserie mit sich selbst bei keiner zeitlichen Verschiebung. Wie in Gleichung 5 zu sehen, werden die Werte einer Zeitserie gewichtet, indem sie durch die Wurzel des Autokorrelationswerts für die Verschiebung $\tau = 0$ dieser Zeitserie geteilt werden.

$$c(\tau) = \frac{1}{n - |\tau|} \sum_{i=1}^n \frac{x_i}{\sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}} \cdot \frac{y_{i+\tau}}{\sqrt{\frac{1}{n} \sum_{i=1}^n y_i^2}} = \frac{n}{n - |\tau|} \frac{\sum_{i=1}^n x_i \cdot y_{i+\tau}}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}} \quad (5)$$

Zudem wird von jedem Wert der Zeitserie x_i der Mittelwert $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ abgezogen, dadurch lassen sich Antikorrelationen feststellen: Wenn beispielsweise die Anzahl der COVID-19 Infektionen eines Landkreises zu einem Zeitpunkt überdurchschnittlich wächst, also die 7-Tages-Inzidenz minus dem Mittelwert der 7-Tages-Inzidenzen positiv

ist, und das andere Edukt aus der anderen Zeitserie negativ ist, also in dem anderen Landkreis die Anzahl der COVID-19 Infektionen unterdurchschnittlich wächst, ergibt sich ein negatives Produkt, da sich die Situation in dem einen Landkreis schneller als üblich verschlechtert, während sich die Situation im anderen Landkreis verbessert oder langsamer als üblich verschlechtert.

Ergibt die Summe aus all den Produkten einer Korrelation eine negative Zahl, scheint die 7-Tages Inzidenz des einen Landkreises zu fallen oder langsam zu steigen, während die 7-Tages Inzidenz des anderen Landkreises steigt oder langsam fällt, dies wird hier Antikorrelation genannt.

Mit allen Ergänzungen berechnet sich die Korrelation $c(\tau)$ zwischen Zeitserie X : $|X| = n$ mit den Werten $x_i \in X$ und Zeitserie Y : $|Y| = n$ mit den Werten $y_i \in Y$ und einer Verschiebung τ relativ zu X mithilfe der Mittelwerte der Zeitserien \bar{x} , \bar{y} wie in Gleichung 6 beschrieben.

$$c(\tau) = \frac{n}{n - |\tau|} \frac{\sum_{i=1}^n (x_i - \bar{x}) \cdot (y_{i+\tau} - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (6)$$

2.4.2 Korrelation zwischen zwei Gebieten am Beispiel von Flensburg und Kiel

Um die hergeleiteten Gleichungen und die weiterhin verwendete Terminologie zugänglicher zu machen, wird eine Korrelationsanalyse mit dem Verlauf der 7-Tages Inzidenz der Stadtkreise Flensburg und Kiel durchgeführt. Der Verlauf der 7-Tages Inzidenz der beiden Stadtkreise sind in Abbildung 5 und Abbildung 6 einmal im Original und einmal um die mittlere 7-Tages Inzidenz in x-Richtung verschoben abgebildet.

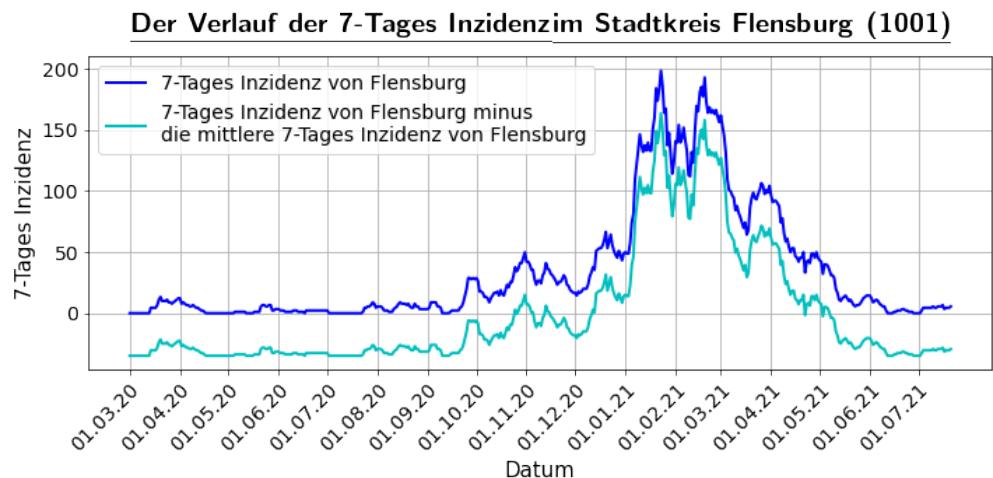


Abbildung 5: Der Verlauf der 7-Tages Inzidenz des Stadtkreises Flensburg (Gemeindeschlüssel 1001).

In Abbildung 5, Abbildung 6, Abbildung 7, Abbildung 8 und Abbildung 9 ist das Ergebnis mit den originalen 7-Tages Inzidenzen blau dargestellt und das Ergebnis mit den 7-Tages Inzidenzen, von denen der Mittelwert der 7-Tages Inzidenzen des jeweiligen Landkreises abgezogen wurde, hellblau dargestellt.

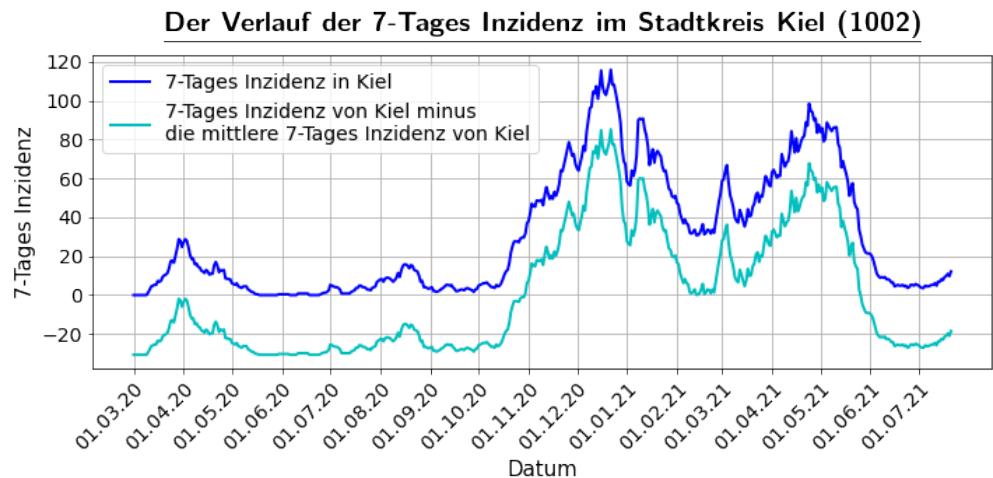


Abbildung 6: Der Verlauf der 7-Tages Inzidenz des Stadtkreises Kiel (Gemeindeschlüssel 1002).

In Abbildung 7 ist die unskalierte Faltung, wie sie in Gleichung 3 definiert ist, abgebildet.

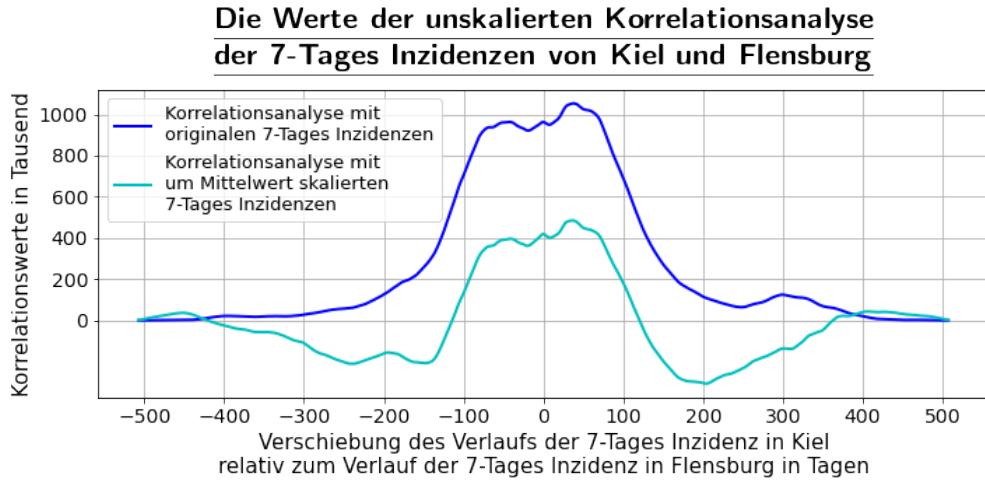


Abbildung 7: Caption

In Abbildung 8 ist die nach Gleichung 5 um die Autokorrelation skalierte Faltung dargestellt.

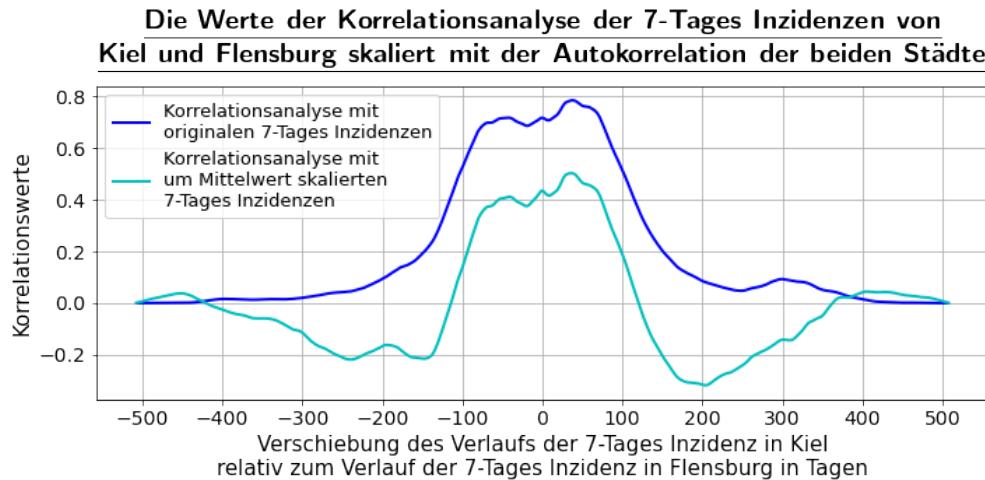


Abbildung 8: Caption

Wenn nach Gleichung 4 die Summen jeweils durch die Anzahl ihrer Summanden (den

Produkten) geteilt werden, ergibt sich Abbildung 9.

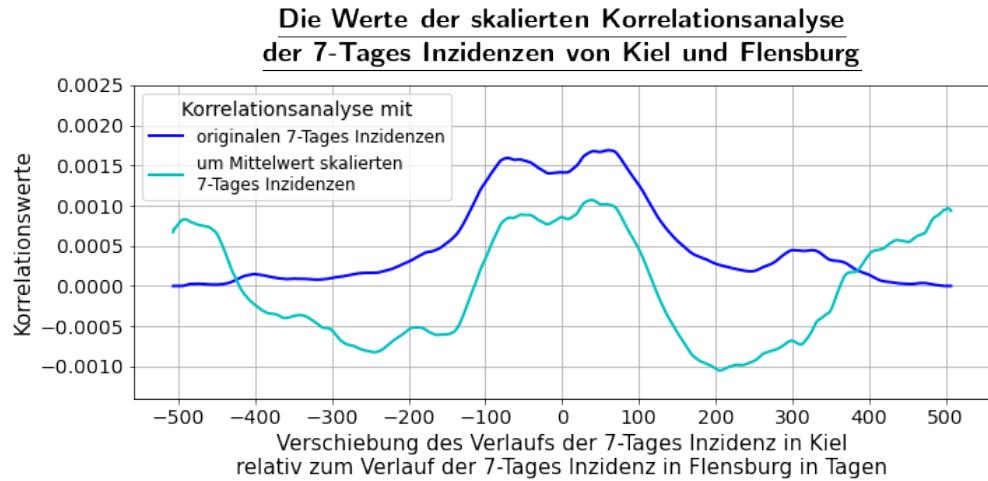


Abbildung 9: Caption

In Abbildung 9 ist schön zu erkennen, dass die Korrelationsanalyse von zwei Zeitserien mit 508 7-Tages Inzidenzen 1015 Korrelationswerte ergibt. Für jeden der 412 Landkreis ergeben sich daher aus der Korrelationsanalyse mit sich selbst und jedem anderen Landkreise 418.180 Werte, ebenso ergeben sich für jeden der 38 Regierungsbezirk 38.570. Dies ergibt in Summe $1.465.660 + 172.290.160 = 173.755.820$ Werte.

Jeder Wert gibt die Wahrscheinlichkeit einer Korrelation bei der jeweiligen Verschiebung an, wobei die zugeordneten Verschiebungen von links nach rechts bei jedem Schritt um eins zunehmen und der mittlerste der Korrelationswerte der Verschiebung $\tau = 0$ zugeordnet ist. Die Verschiebung ist im Kontext dieser Arbeit stets in ganzen Tagen zwischen -507 und 507 angegeben. Der Wert an Position 501 in der Liste der Korrelationswerte gibt also an, wie wahrscheinlich der Verlauf der 7-Tages Inzidenz des ersten Gebiets mit dem um eine Woche nach links verschobenen Verlauf der 7-Tages Inzidenz des zweiten Gebiets korreliert.

2.4.3 Komprimierung und Darstellung als Matrizen

Das Ziel der nachfolgend beschriebenen Schritte ist, die Korrelation von zwei Gebieten auf einen Wert zusammenzuführen, sodass man einfach und schnell alle Korrelationen zwischen zwei Gebieten vergleichen kann.

In dieser Arbeit werden nur Korrelationen mit einem zeitlichen Versatz zwischen $\tau = -50$ und $\tau = 50$ betrachtet, da zum einen eine Interpretation für eine Verschiebung von mehr als 4 Wochen sehr schwierig ist und zum anderen einzelne Ausreißer bei größeren Verschiebungen stärker ins Gewicht fallen, weil immer weniger Produkte aufsummiert werden. (**TODO: anderes Wort für schwierig in „Interpretation für eine Verschiebung von mehr als 4 Wochen sehr schwierig ist“**) Um ein detailliertes Bild zu erhalten, werden zudem Korrelationsanalysen für die Verschiebungen $\tau \in [-30, 30]$ und $\tau \in [-14, 14]$ durchgeführt.

Somit sind jeder Kombination aus zwei Landkreisen nur noch höchstens 101 Werte zugeordnet. Diese Werte werden in zwei verschiedenen Varianten zu einem Wert zusammengefasst:

- Die Verschiebung mit dem maximalen Korrelationswert: Der maximale Wert wird herausgesucht und die zugehörige zeitliche Verschiebung wird als Maß dafür angegebenen, bei welcher Verschiebung in Tagen am wahrscheinlichsten eine Korrelation vorliegt.
- Die Verschiebungstendenz: Der Durchschnitt der ersten Hälfte der Liste wird vom Durchschnitt der zweiten Hälfte abgezogen. Das Resultat ermöglicht eine grobe Einschätzung, ob der maximale Wert nur ein Ausreißer ist oder nicht.

Die erste Variante generiert Werte welche sich leicht interpretieren lassen. Doch um die kompletten Informationen aus der Liste zu verarbeiten und nicht nur einen Wert herauszunehmen, bietet es sich an, das Ergebnis der ersten Variante mit dem der zweiten Variante zu vergleichen. Jedoch lässt sich bei der zweiten Variante nur noch feststellen, ob die Werte der Zeitserie vor, nach oder mit den Werten der anderen

wachsen, jedoch nicht mehr wie viele Tage früher/später - dafür braucht es das Ergebnis der ersten Variante.

Die Werte lassen sich jeweils in zwei Matrizen darstellen: Jeder Zeile und Spalte wird der Index für ein Gebiet zugeordnet und der Wert in einer Zelle entspricht dem Ergebnis der Korrelationsanalyse der Gebiete, denen jeweils der Index der Zeile und der Spalte zugeordnet ist. Die Matrizen sind (mit verkehrtem Vorzeichen) symmetrisch an den Diagonalen von links oben nach rechts unten. Die Diagonalen sind mit Nullen besetzt, da die positive Verschiebung symmetrisch zur negativen Verschiebung ist und die Korrelation mit sich selbst trivialerweise bei einer Verschiebung von $\tau = 0$ am größten ist, da die Werte und Trends komplett identisch sind.

Um den Landkreisen Werte zuzuordnen und nicht der Kombination aus Landkreisen, wird sowohl bei den Maximalwerten wie auch den Verschiebungstendenzen der Mittelwert gebildet, indem die Zeilen der Matrizen aufsummiert werden und durch die Anzahl der Spalten geteilt werden.

2.5 Durchschnitt, Farbgebung und Skalierung

Um schnell verständliche Abbildungen bereitzustellen zu können, werden die Werte skaliert und die Farbgebung der Deutschlandkarten derart angepasst, dass das gesamte Farbspektrum abgedeckt wird. Das Farbspektrum reicht gemäß den Farben des Regenbogens von blau über grün zu gelb zu rot, wie in Abbildung 10 demonstriert. Die niedrigsten Werte werden blau gefärbt. Da manche dieser Farbwerte im Kontrast zu einem weißen Hintergrund schwer zu erkennen sind, ist der Hintergrund der meisten Abbildungen grau.

Die Matrizen werden durch die verwendete Programmbibliothek „Matplotlib“ automatisch eingefärbt.

Um die Daten zusammenfassen wird das arithmetische Mittel benutzt. Der Mittelwert entspricht nach Gleichung 7 der Summe der einzelnen Werte $x_1, x_2, x_3 \dots x_n$ geteilt



Abbildung 10: Das Farbspektrum, in welchem sich die Darstellungen bewegen. Von links nach rechts steigen die eingegebenen Werte konstant. Der angegebene Wert wird jeweils anhand des ersten und des letzten Werten einer Referenzliste linear in diesem Spektrum verortet.

durch ihre Anzahl n .

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (7)$$

3 Vorgehensweise

Um dem im Kapitel „1 Motivation“ formulierten Ziel zu folgen, muss zunächst eine Datenquelle gewählt werden. Anschließend werden deren Daten in eine nutzbare Form übertragen. Schlussendlich können die Informationen aus den Daten verknüpft, interpretiert und graphisch dargestellt werden.

Der entsprechende Programmcode findet sich in den Programmdateien im Anhang der Bachelorarbeit. (**(TODO: Wo sind die Programmdateien)**)

Diese Bachelorarbeit basiert teilweise auf den Vorleistungen des Bachelorprojekts von Leander Marius Bürkin. Ziel dieses Bachelorprojekts war eine Reihe an Deutschlandkarten in einem Video zusammenzufassen, welches die Ausbreitung der COVID-19 Pandemie in Deutschland vom ersten März 2020 bis zum letzten Tag, für den die API des Robert-Koch-Instituts Daten liefert, darstellt. Das fertige Video ist verfügbar unter (**(TODO: Video verlinken)**).

3.1 Datenquellen - Ursprung und Abspeicherung

Das Bachelorprojekt sowie diese Bachelorarbeit verwenden Informationen zur COVID-19 Pandemie und den geographischen Daten von 412 deutschen Landkreisen. Alle Daten stammen von der Programmierschnittstelle (API) „COVID-19 Datenhub“ (npgeo-corona-npgeo-de.hub.arcgis.com) oder wurden aus den daher stammenden Daten generiert. Diese Datenquelle wurde gewählt, weil sie vom Robert-Koch-Institut (RKI, www.rki.de) und dem deutschen Staat referenziert wird.

(TODO: wie verlinke ich die URL korrekt?)

(TODO: Staat referenzieren)

Insgesamt werden drei verschiedene Datenpakete der API verwendet, zum einen die geographischen Daten der Landkreise, zum anderen die Summe aller aufgetretenen COVID-19 Fälle seit Beginn der Pandemie für jeden Landkreis und jeden Tag vom 01.03.2020 bis zum 21.07.2021 sowie eine Auflistung aller Meldungen der Gesundheitsämter, in welchen das Referenzdatum, das Melde datum, die Anzahl der betroffenen Menschen und deren Zustand (entweder genesen, verstorben oder noch infektiös) angegeben werden. Das Referenzdatum kann als Tag der Infektion interpretiert werden, das Melde datum als Genesungsdatum beziehungsweise Sterbedatum.

(TODO: RKI Paper verlinken)

Die Landkreise, welche das RKI angibt stimmen nicht mit den Landkreisen des Statistischen Bundesamtes überein: In den Daten des RKIs gibt es 118 Landkreise mehr, beziehungsweise wenn man die kreisfreien Städte abzieht, zwölf Landkreise mehr. Diese Diskrepanz kommt durch die Aufteilung von Berlin in seine 12 Bezirke und Eisenach wird in den Daten des RKIs als eigene kreisfreie Stadt gewertet und nicht zum Wartburgkreis hinzugezählt, der Stadtkreis Eisenach ist mit dem Gemeindeschlüssel 16056 versehen. Trotz dieser Unterschiede und obwohl auch kreisfreie Städte miteinbezogen sind, wird nach Vorbild des RKIs im folgenden der Begriff „Landkreise“ für alle 412 Gebiete verwendet. (TODO: Verlinken: <https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnisse-regierungsbezirke.html>) (TODO: <https://www.destatis.de/DE/Themen/Laender-Regionen/Regionales/Gemeindeverzeichnis/Administrativ/04-kreise.html>)
(TODO: Liste der NUTS-Codes Gemeindeschlüssel leiten sich daraus ab. URL siehe Kommentar) Die Daten werden abgespeichert, damit sie nicht bei jeder Ausführung erneut angefordert und aufbereitet werden müssen.

3.2 Datenaufbereitung

Bevor die Daten genutzt werden können, müssen sie verifiziert werden, überflüssige Informationen entfernt werden und neue Kennzahlen aus den gegebenen Informationen berechnet werden. Daten vor der Datenaufbereitung, welche direkt von der API oder einem Backup davon stammen, werden als „unmodifizierte“ Daten bezeichnet.

Daten, die rudimentär auf Vollständigkeit überprüft wurden, aus den Daten generierte Daten enthalten und bei welchen überflüssige Informationen entfernt wurden, werden „modifizierte“ Daten genannt.

Zunächst werden die unmodifizierten Daten in eine übersichtlichere Form übertragen und es wird sichergestellt, dass gleich viele Werte für jeden Landkreis vorhanden sind. Zudem werden die Umrisse der 100 Landkreise von Hand geprüft, welche mehrere Polygone enthalten, diese werden entweder als Ausschnitt oder als reale Fläche interpretiert. Würde man einfach alle Polygone zeichnen, kann es passieren, dass ein Bereich, welcher komplett von einem Landkreis umgeben ist von einem seiner Polygone übermalt wird, welches genau diese Fläche aus dem Landkreis ausschneiden sollte.

Nachdem die unmodifizierten Daten in eine übersichtlichere Form übertragen und überprüft wurden, werden aus den Daten weitere Werte berechnet.

Die Bevölkerungsdichte wird berechnet, indem die Anzahl der Einwohner durch die Fläche in Quadratmetern geteilt wird (siehe Gleichung 1). Beide Informationen werden von der API bereitgestellt. Die Bevölkerungsdichte wird auch für die Regierungsbezirke berechnet. Nicht alle Bundesländer wurden in Regierungsbezirke unterteilt, in diesem Fall wird die Bevölkerungsdichte des Bundeslandes gewählt. Im Folgenden wird dennoch von „den Regierungsbezirken“ gesprochen. (**(TODO: Wo werden die regierungsbezirke definiert?)**) Sowohl die Bevölkerungsdichten der Landkreise wie auch die Bevölkerungsdichten der Regierungsbezirke werden zur Einordnung der Korrelationsanalysen auf einer Deutschlandkarte dargestellt, wobei die Farbe die

Bevölkerungsgröße repräsentiert, wie in Abschnitt 2.5 beschrieben.

Für jeden Landkreis und jeden Regierungsbezirk wird eine Zeitserie mit den 7-Tages Inzidenzen angefertigt, welche die 7-Tages Inzidenz nach Gleichung 2 für jeden Tag enthält, für den eine Fallzahlen angegeben ist.

Um die Auflistung aller Meldungen der Gesundheitsämter nutzen zu können, müssen diese erst für jeden Landkreis und jeden Tag gesammelt und aufsummiert werden: Das Referenzdatum wird als Tag der Infektion interpretiert. Die Anzahl der betroffenen Individuen wird für jeden folgenden Tag zur akkumulierten Anzahl der Fälle hinzugefügt. Das MeldeDatum wird als Tag der Genesung beziehungsweise Tag des Todes interpretiert, je nachdem ob die Meldung Genesung oder Tod angibt. Die Anzahl der betroffenen Individuen wird hier zum einen für jeden folgenden Tag zur akkumulierten Anzahl der Genesenen/Verstorbenen hinzugefügt. Zum anderen wird die Anzahl der betroffenen Individuen für jeden Tag zwischen dem Referenzdatum und dem MeldeDatum zu den aktiven Fällen hinzugefügt.

Mithilfe der Gesamtbevölkerung eines Landkreises werden aus diesen Daten für jeden Tag und jeden Landkreis alle drei Kategorien des SIR-Modells berechnet:

- „susceptible“: Die Gesamtbevölkerung des Landkreises minus die akkumulierte Anzahl der Fälle.
- „infectious“: Liegt als aktive Fälle bereits vor.
- „recovered“: Die Summe aus der akkumulierten Anzahl der Genesenen und der Verstorbenen. Oder die akkumulierte Anzahl der Fälle minus die aktiven Fälle, solange zu jeder Meldung ein Referenzdatum angegeben ist, entspricht dies der ersten Variante.

Um ein Gefühl dafür zu bekommen, welcher Landkreis in welchem Ausmaß von der Corona Pandemie getroffen wurde, wird die akkumulierte Anzahl der COVID-19 Fälle des letzten Tages durch die Bevölkerung des Landkreises geteilt und farblich in einer

Deutschlandkarte dargestellt.

3.3 Datenformat

Die Daten liegen in einer Kombination aus sogenannten Python Dictionaries und sogenannten Python Listen vor.

Ein Dictionary ist dadurch charakterisiert, dass man über den Namen eines Elements („Key“) das Element („Value“) bekommt, ähnlich wie man früher in Telefonbüchern anhand des Namens (Key) die Telefonnummer (Value) erhalten hat.

Listen sind dadurch charakterisiert, dass sie Elemente in einer fixen Reihenfolge enthalten und sich diese durch ihren Index leicht auffindbar machen. So sind beispielsweise die Anzahl der COVID-19 Fälle der einzelnen Tage chronologisch in einer Liste gespeichert, sodass man daraus direkt eine Abbildung erstellen kann.

Die Programmdateien erstellen drei verschiedene sogenannte globale Dictionaries, welche (in untergeordneten Dictionaries und Listen) alle Daten enthalten. Globale Dictionaries sind in allen Dateien verfügbar und unterscheiden sich insofern von lokalen Dictionaries/Listen, welche nur in einzelnen Programmabschnitten vorliegen und daher nicht derart zentral sind. Die drei Dictionaries heißen(**TODO: vervollständigen, districts**)

- covid19
- counties_geography
- non_county_specific_data

covid19

Das Dictionary covid19 speichert die COVID-19 Fälle jedes deutschen Landkreises und die daraus berechnete sieben Tages Inzidenz.

Mit dem Gemeindeschlüssel eines Landkreises als Key und dem zusätzlichen Key „cases“ erhält man die akkumulierte Anzahl an COVID-19 Fällen pro Tag des Landkreises als Liste.

Mit dem Gemeindeschlüssel eines Landkreises als Key und dem Key „incidences“ erhält man die sieben Tages Inzidenz pro Tag des Landkreises als Liste.

Beide Listen enthalten für jeden Tag seit dem 01.03.2020 bis zu dem Tag der aktuellsten Daten jeweils einen Eintrag.

Die Anzahl an COVID-19 Fällen pro Tag stammt aus dem „COVID-19 Datenhub“. Die sieben Tages Inzidenz wird wie in Abschnitt 2.3 beschrieben aus den Daten vom COVID-19 Datenhub berechnet.

(TODO: Aufbau des Gemeindeschlüssels erklären) counties_geography

Das Dictionary counties_geography enthält zu jedem Landkreis ein Dictionary mit den folgenden Elementen:

name: Der Name des Landkreises

population: Die Einwohnerzahl des Landkreis aus einer offiziellen Schätzung (nähtere Informationen finden sich im COVID-19 Datenhub). Diese Zahlen werden auch für die offizielle Berechnung der sieben Tages Inzidenz verwendet.

area_in_m2: Die Fläche des Landkreises in Quadratmetern

geometry: Die Form des Landkreises in einer für die Darstellung angepassten Form

raw_geometry: Die Form des Landkreises in der originalen Form

population_density: Die Bevölkerungsdichte des Landkreises, berechnet aus der Fläche des Landkreises und seiner Einwohnerzahl

Außer die Bevölkerungsdichte, welche berechnet wird, und die angepassten Formen

der Landkreise, welche aus den originalen Formen generiert werden, stammen alle Daten direkt aus dem „COVID-19 Datenhub“.

non_county_specific_data Das Dictionary non_county_specific_data enthält die folgenden Elemente:

unixtime: Die Unixzeit, wie sie vom COVID-19 Datenhub zur Verfügung gestellt wird:

Die Zahl der Millisekunden seit dem 01.01.1970 00:00 Uhr UTC. Die Zahl der COVID-19 Fälle und die sieben Tages Inzidenz eines entsprechenden Tages befinden sich an der selben Stelle in Ihrer jeweiligen Liste, wie der Tag in dieser Liste.

states: Die Namen und Nummern der deutschen Bundesländer. Die ersten beiden oder die erste Zahl des Gemeindeschlüssel eines Landkreises geben das Bundesland an, in welchem der Landkreis liegt. Die ersten zwei bzw. drei Zahlen geben den Regierungsbezirk an, in dem der Landkreis. Sollte es im jeweiligen Bundesland keine Regierungsbezirke geben, entspricht dies dem Bundesland.

highest_case_number: Die höchste akkumulierte Fallzahl unter allen Landkreisen.

lowest_case_number: Die niedrigste akkumulierte Fallzahl unter allen Landkreisen.

highest_incidence: Die höchste sieben Tages Inzidenz eines Tages unter allen Landkreisen.

lowest_incidence: Die niedrigste sieben Tages Inzidenz eines Tages unter allen Landkreisen.

UTC: Die Unixzeit konvertiert in die Koordinierte Weltzeit im Format „DD.MM.YYYY“.

UTC+7days: Die Liste der Zeiten gespeichert in UTC plus sieben weitere Tage vor dem erstem Datum in der Liste. Diese werden benutzt, um den Zeitraum der 7-Tages Inzidenz der ersten sieben Tage darstellen zu können: die hierfür verwendeten Zeiträume beginnen jeweils vor dem ersten hier dokumentierten Fall.

(TODO: itemize reparieren) Die Unixzeit, die Namen und die Nummern der Bundesländer stammen aus dem „COVID-19 Datenhub“. Alle anderen Informationen wurden gesammelt oder berechnet.

In Abbildung 11 sind die drei Dictionaries dargestellt, wie sie auch im Programm verwendet werden.

```
Input:      list(covid19["1002"].keys())
Output:     ['cases', 'incidences']

Input:      list(counties_geography["1002"].keys())
Output:     ['name',
            'population',
            'area_in_m2',
            'geometry',
            'raw_geometry',
            'population_density']

Input:      list(non_county_specific_data.keys())
Output:     ['unixtime',
            'states',
            'highest_case_number',
            'lowest_case_number',
            'highest_incidence',
            'lowest_incidence',
            'UTC',
            'UTC+7days']
```

Abbildung 11: Die Dictionaries non_county_specific_data, counties_geography und covid19 mit ihren jeweiligen Schlüsseln. Für die Dictionaries covid19 und counties_geography wurde jeweils der Landkreis Kiel (Gemeindeschlüssel 1002) zur Veranschaulichung verwendet.

(TODO: Daten aus Meldungen einpflegen)

3.4 Datendarstellung

Zuerst werden die Daten der Landkreise und Regierungsbezirke dargestellt, welche auf einfachen Gleichungen beruhen und lediglich zur Einordnung der anderen Ergebnisse dienen:

- Die Bevölkerungsdichte, um Schlussfolgerungen ziehen zu können, ob diese in irgendwie mit dem Infektionsverhalten zusammenhängt
- Die Anordnung der Gebiete, wenn man sie nach ihrem Gemeindeschlüssel sortiert - Die Sortierung erfolgt hier lexikographisch, das heißt, dass die Länge des Gemeindeschlüssels keine Rolle spielt, sondern immer das erste Zeichen verglichen wird und bei Übereinstimmung das nächste (Dadurch ergibt sich eine einheitlichere Nord-Süd-Aufteilung als die Sortierung nach der Größe)
- Die Summe der 7-Tages Inzidenzen als Maß wie stark ein Gebiet betroffen war

3.4.1 SIR-Modell

Da die Zahl der aktiven Fälle sehr niedrig und dementsprechend sensibel gegenüber kleinen Anomalien ist, wird nur auf Bundesebene ein SIR-Modell erstellt und dargestellt. Hierfür werden die drei Kennzahlen des SIR-Modells in drei Abbildungen dargestellt. (**TODO: SIR-Modell einfügen und mit Grundlagen abgleichen**)

3.4.2 Korrelationsanalyse

Die 7-Tages Inzidenzen der folgenden Gebiete werden nach den in Abschnitt 2.4 beschriebenen Methoden untersucht:

- Die Korrelation einzelner ausgewählter Städte, welche komplett von einem Landkreis umgeben sind, zu eben diesem Landkreis
- Die Korrelationen aller Landkreise untereinander

- Die Korrelationen aller Regierungsbezirke untereinander

3.4.3 Korrelation einzelner ausgewählte Städte und Landkreise

Zum Einstieg werden die Korrelationen zwischen den Landkreisen aus Tabelle 1 und den Stadtkreisen, die sie umgeben, berechnet. An ihnen lässt sich besonders gut testen, ob sich in den Korrelationswahrscheinlichkeiten zwischen einer Stadt und ihrem Umland eine zeitliche Verschiebung feststellen lässt. Um die festgestellten zeitlichen Verschiebungen einzuordnen, wird zudem ermittelt, bei wieviel Prozent der Korrelationen aller deutschen Landkreise untereinander die höchste Korrelationswahrscheinlichkeit bei einer Verschiebung von $\tau = 0$ zu finden ist.

Wie in Abschnitt 2.4 werden für die ausgewählten Landkreis-Stadtkreis-Paare Korrelationsanalysen durchgeführt. Da es sich um eine übersichtliche Datenmenge handelt und für jeden Landkreis nur eine Korrelation geprüft wird, werden die Werte, welche normalerweise in Matrixform dargestellt werden, explizit ausgegeben.

In Abbildung 12 sind die ausgewählten Landkreise mit den Stadtkreisen, die sie umgeben abbgebildet.

Name des Landkreises (LK)	AGS Landkreis	Umgebener Stadtkreis (SK)	AGS Stadtkreis
Kassel	6633	Kassel	6611
Trier-Saarburg	7235	Trier	7211
Südliche Weinstraße	7337	Landau i.d.Pfalz	7313
Südwestpfalz	7340	Pirmasens	7317
Heilbronn	8125	Heilbronn	8121
Rastatt	8216	Baden-Baden	8211
Rosenheim	9187	Rosenheim	9163
Landshut	9274	Landshut	9261
Straubing-Bogen	9278	Straubing	9263
Amberg-Sulzbach	9371	Amberg	9361
Neustadt a.d.Waldnaab	9374	Weiden i.d.OPf	9363
Regensburg	9375	Regensburg	9362
Bamberg	9471	Bamberg	9461
Oberallgäu	9780	Kempten	9763
Potsdam-Mittelmark	12069	Brandenburg a.d. Havel	12051
Spree-Neiße	12071	Cottbus	12052
Saalekreis	15088	Halle (Saale)	15002
Weimarer Land	16071	Weimar	16055

Tabelle 1: Landkreise mit Name und Gemeindeschlüssel (AGS), die den jeweils mit Name und Gemeindeschlüssel gekennzeichneten Stadtkreis komplett umgeben.

Die ausgewählten Stadt- und Landkreise

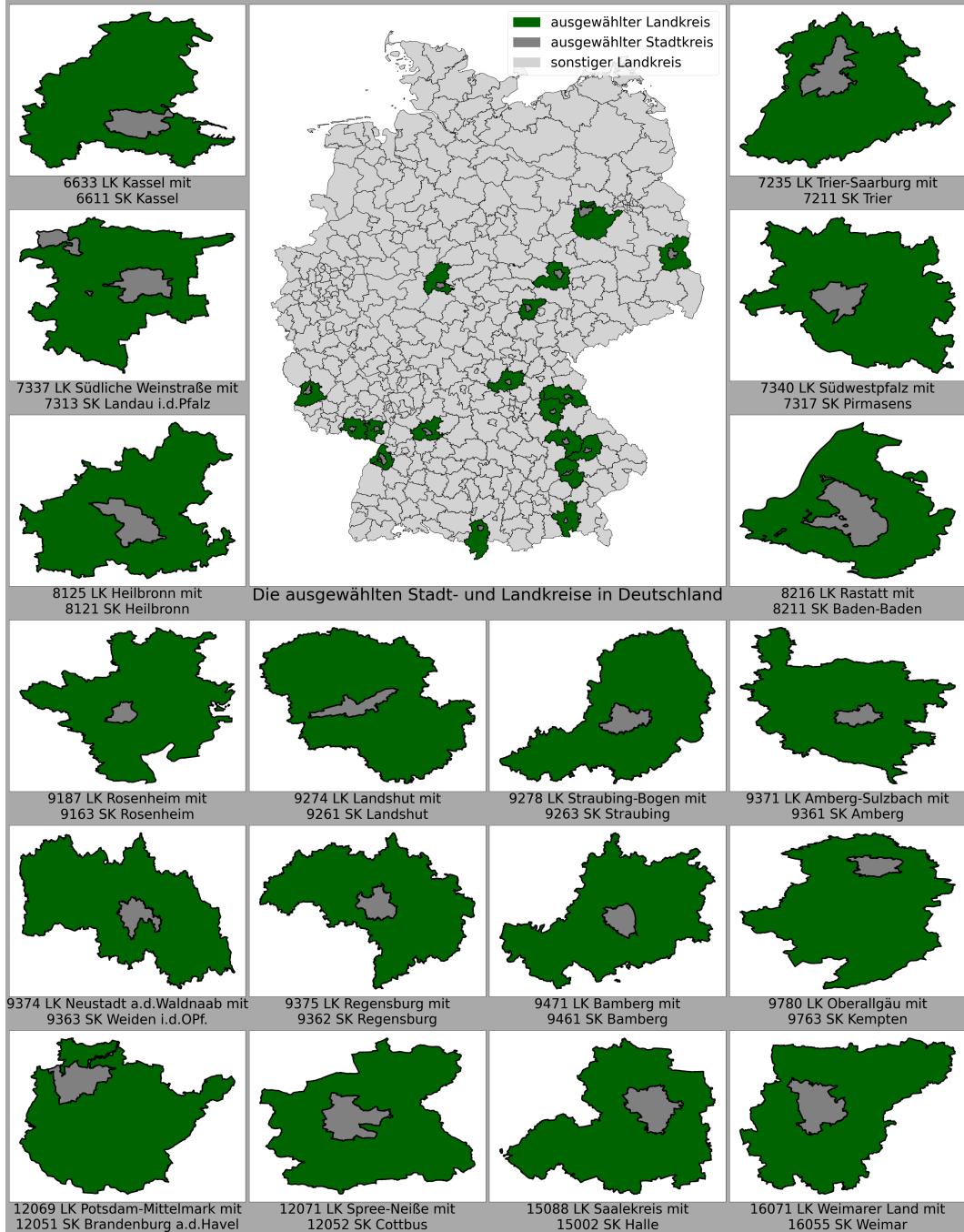


Abbildung 12: Die ausgewählten Landkreise mit den Stadtkreisen, die sie jeweils umgeben.

Matrix der maximalen Wahrscheinlichkeiten

(TODO: farbliche Darstellung der Matrix hier erklären) (TODO: Grundlagen referenzieren)

So werden für jede Kombination aus zwei Gebieten sechs Werte ermittelt: Jeweils für Verschiebungen zwischen $\tau = -50$ und $\tau = 50$, $\tau = -30$ und $\tau = 30$ und $\tau = -14$ und $\tau = 14$ zwei Werte aus den beiden erwähnten Varianten. Diese Matrizen werden jeweils mit beiden Methoden für alle Landkreise und alle Regierungsbezirke erstellt.

4 Resultate

4.1 Bevölkerungsdichten der Landkreise und Regierungsbezirke

In Abbildung 13 sind die Bevölkerungsdichten der einzelnen Landkreise dargestellt. Auf der linken Seite befindet sich die Verteilung und auf der rechten Seite die räumliche Anordnung.

Die Bezirke Berlins sind einzeln gelistet, daher entsprechen die sechs höchsten Bevölkerungsdichten den Berliner Bezirken Friedrichshain-Kreuzberg, Mitte, Neukölln, Tempelhof-Schöneberg, Lichtenberg, Charlottenburg-Wilmersdorf, obwohl die Bevölkerungsdichte des gesamten Berliner Stadtkreises niedriger ist als die Bevölkerungsdichte Münchens (in dieser Auflistung Platz 7, ohne Berliner Bezirke Platz 1).

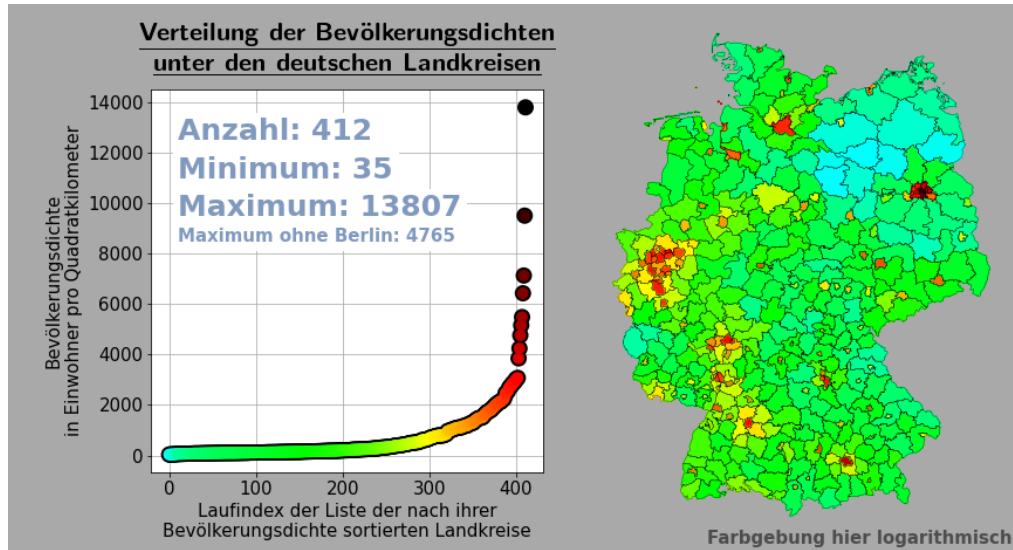


Abbildung 13: Verteilung der Bevölkerungsdichten unter den deutschen Landkreisen.

(TODO: klarstellen, das die Landkreise nicht den Landkreisen entsprechen (Berlin ist aufgeteilt) Aus Wikipedia "Landkreis": In Deutschland gibt es 294 Landkreise. Zusammen mit den 106 kreisfreien Städten bilden sie die insgesamt 400 Gebietskörperschaften auf Kreisebene. Wir haben 412.)

In Abbildung 14 sind die Bevölkerungsdichten der einzelnen Regierungsbezirke dargestellt. Auf der linken Seite befindet sich die Verteilung und auf der rechten Seite die räumliche Anordnung.

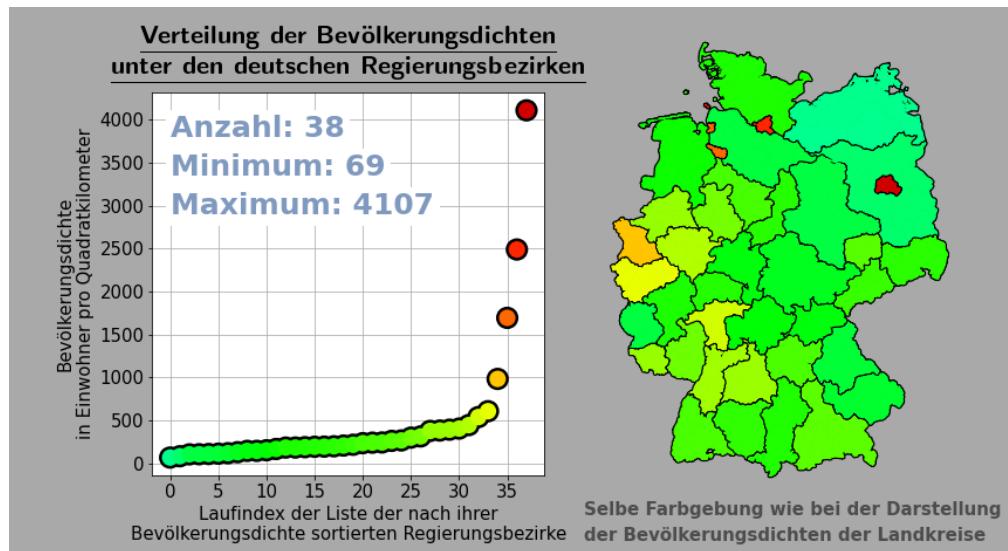


Abbildung 14: Verteilung der Bevölkerungsdichten unter den deutschen Regierungsbezirken. Die Skalierung entspricht der Farbgebung in Abbildung 13.

4.2 Die Landkreise und Regierungsbezirke sortiert nach dem Gemeindeschlüssel

In Abbildung 15 lässt sich an der Farbe der Landkreise/Regierungsbezirke grob erkennen, in welcher Reihenfolge diese behandelt werden, wenn sie in der Reihenfolge ihrer Gemeindeschlüssel benutzt werden.

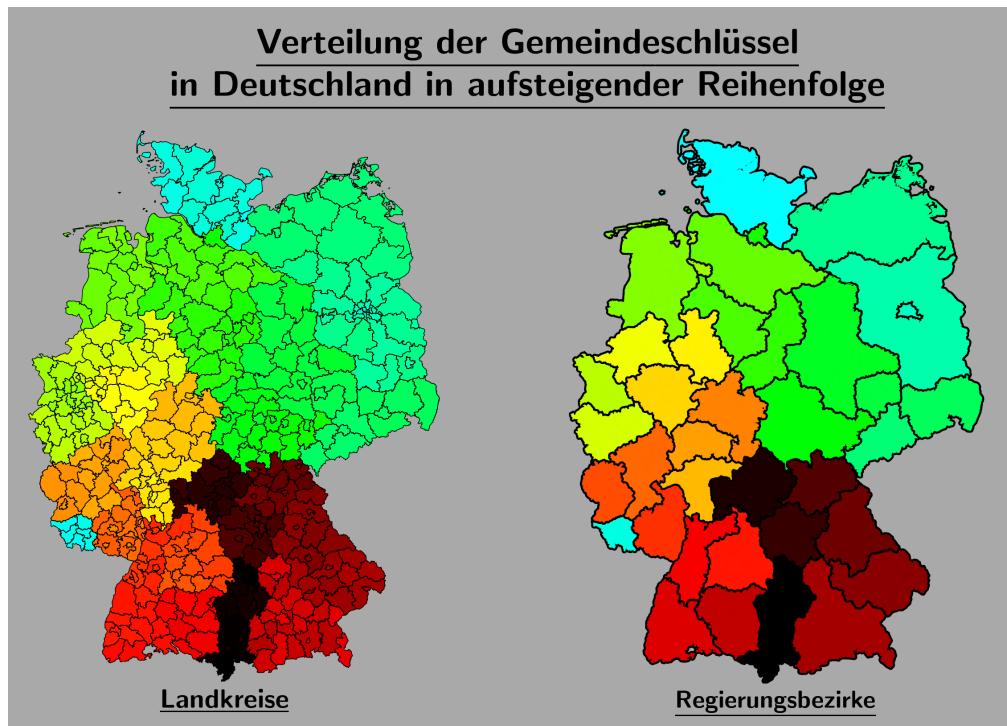


Abbildung 15: Die Landkreise und Regierungsbezirke eingefärbt nach der lexikographischen Größe ihres Gemeindeschlüssels. Bei blau beginnend werden den nach ihrem Gemeindeschlüssel sortierten Landkreise die Farbfolge des Regenbogens zugeteilt.

4.3 Summe der 7-Tages Inzidenzen der Landkreise und Regierungsbezirke

In Abbildung 16 sind die Summe der 7-Tages Inzidenzen der einzelnen Landkreise dargestellt. Auf der linken Seite befindet sich die Verteilung und auf der rechten Seite die räumliche Anordnung.

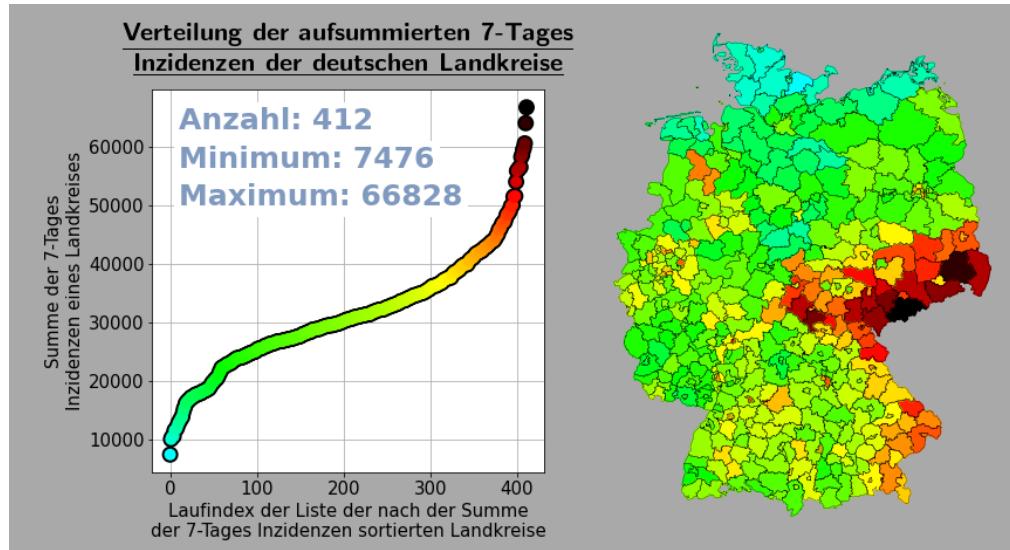


Abbildung 16: Verteilung der Summe der 7-Tages Inzidenzen unter den deutschen Landkreisen.

In Abbildung 17 sind die Summe der 7-Tages Inzidenzen der einzelnen Regierungsbezirke dargestellt. Auf der linken Seite befindet sich die Verteilung und auf der rechten Seite die räumliche Anordnung.

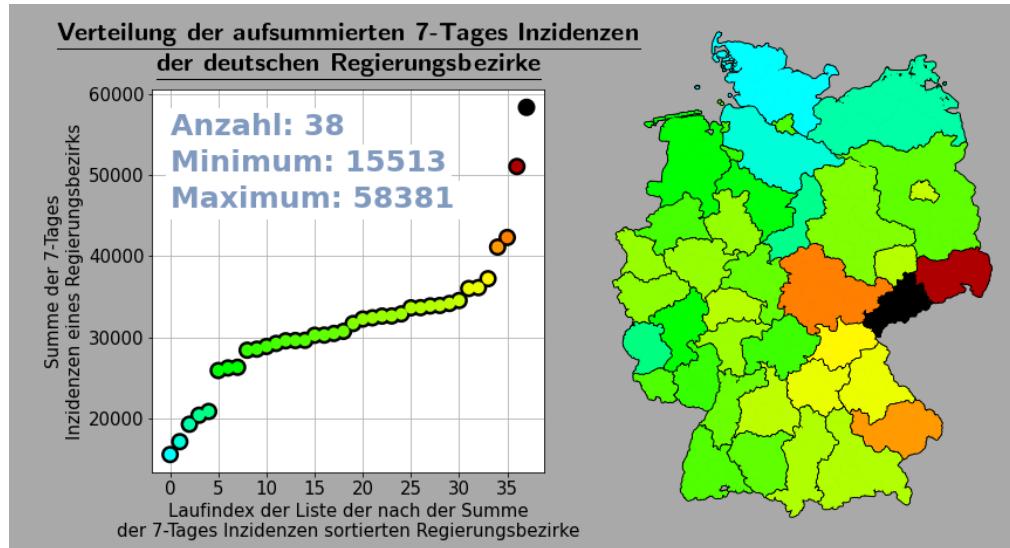


Abbildung 17: Verteilung der Summe der 7-Tages Inzidenzen unter den deutschen Regierungsbezirken.

4.4 SIR-Modell für Deutschland

In Abbildung 18 sind die drei Kennzahlen des SIR-Modells für Deutschland in drei Graphen dargestellt.

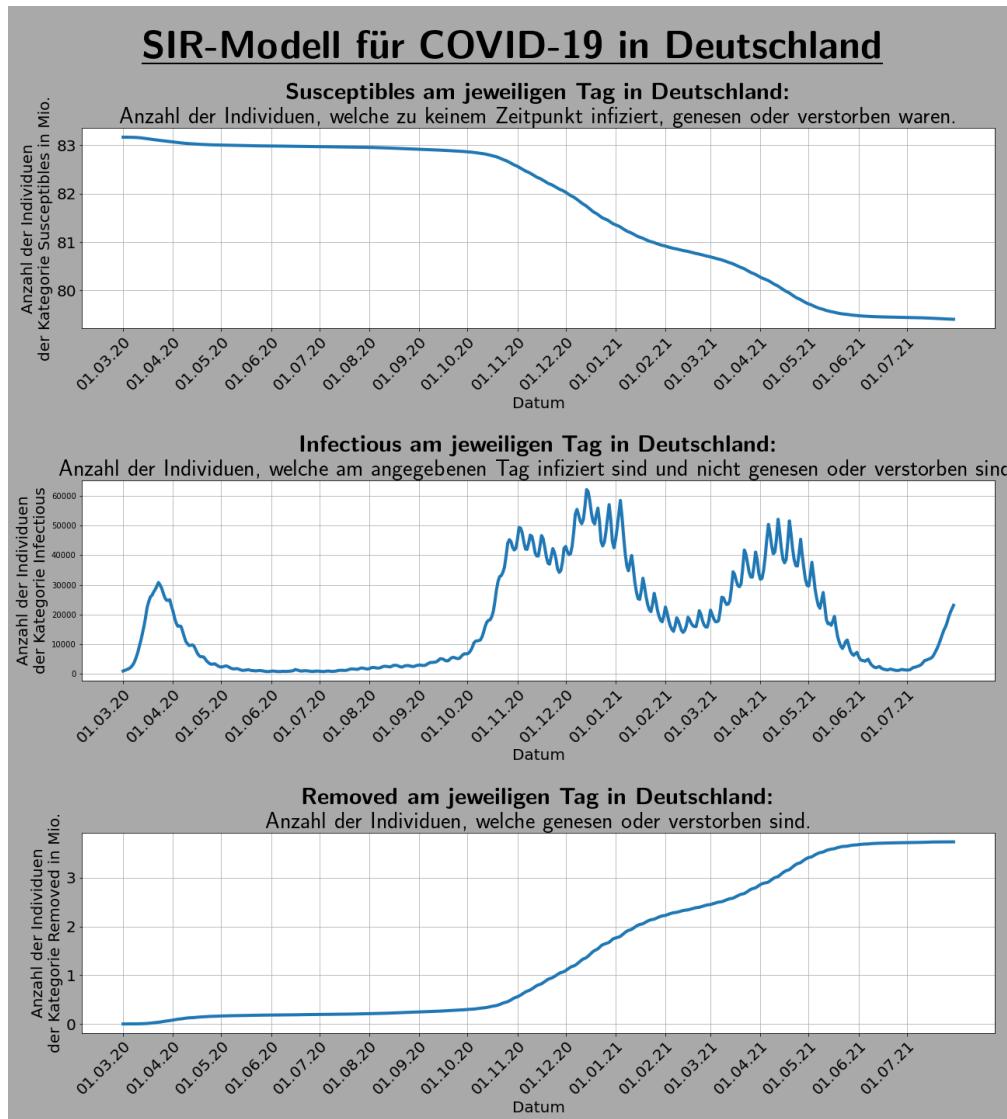


Abbildung 18: Die drei Kennzahlen des SIR-Modells für Deutschland in drei Graphen.

Von den 3.748.038 gemeldeten Fällen weisen 46.033 Meldungen (1,2%) Inkonsistenzen auf. Dies entspricht fast der höchsten Zahl an Menschen in der Kategorie „infectious“

von ca. 60.000. In 37757 Fällen liegt das Meldedatum vor dem Referenzdatum, das heißt die Person ist nach der hier verwendeten Interpretation genesen bzw. gestorben bevor sie sich angesteckt hat. In 8276 Fällen liegt das Referenzdatum mehr als 30 Tage vor dem Referenzdatum, die Person ist also mehr als 30 Tage krank, was laut RKI sehr unwahrscheinlich ist. (**TODO: QQuelle**)

1409 Fälle wurden vor dem 01. März an das RKI gemeldet. Teilweise handelt es sich um authentische Meldungen, teilweise um fragwürdige Meldungen, da die Gesamtzahl jedoch unter einem Prozent liegt, werden die Daten, welche vor dem 01. März liegen, in Abbildung 18 auf den 01. März gesetzt.

4.5 Korrelationen zwischen Städten und ihrem Umland

In Abbildung 19 sind die Korrelationswahrscheinlichkeiten zwischen den in Abschnitt ?? beschriebenen Städten und den Landkreisen für eine zeitliche Verschiebung zwischen $\tau = -50$ und $\tau = 50$ als Kurven dargestellt. (**TODO: Städte und Landkreisnamen verlinken**) Die Kurven, deren Höhepunkt rechts der null sind, sind rot eingefärbt. Die Kurven mit dem höchsten Punkt links der null sind blau gefärbt. Die restlichen Kurven, welche am Nullpunkt maximal werden, sind grün gefärbt.

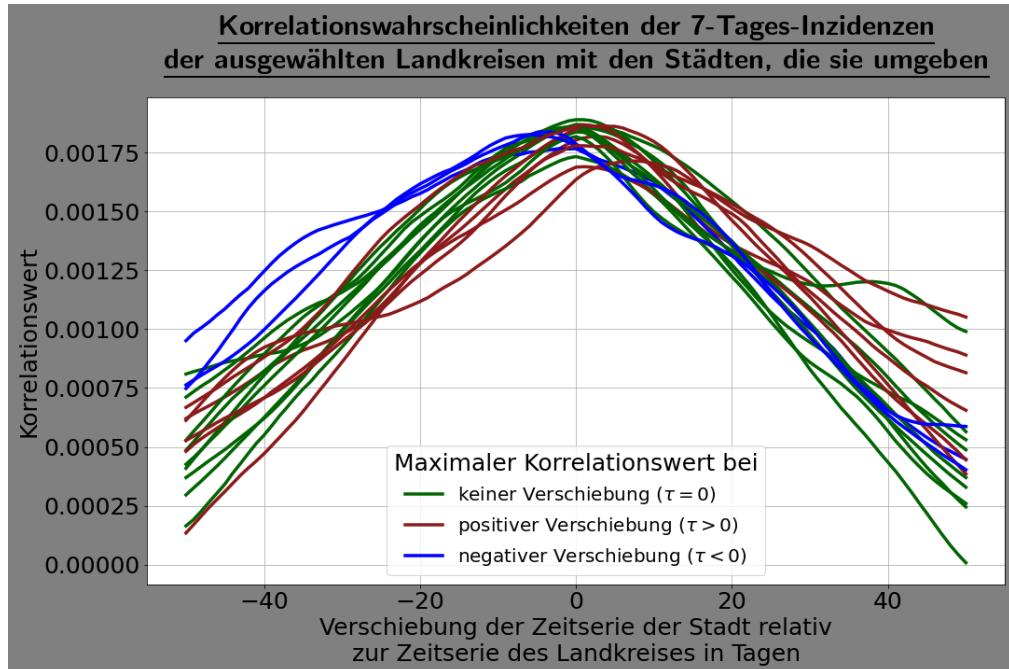


Abbildung 19: Die Korrelationswahrscheinlichkeiten der 7-Tages INzidenzen der ausgewählten Landkreise mit den Städten, die sie umgeben. Die Kurven sind je nach ihrem Hochpunkt eingefärbt: Rot bei positiven x-Werten, Blau bei negativen und grün bei 0.

In dieser sehr kleinen Teilmenge weisen 50 % der Korrelationen die höchste Korrelationswahrscheinlichkeit bei einer Verschiebung von $\tau = 0$ auf. Bei den Korrelationen aller deutschen Landkreise untereinander liegt bei 8,7 % die höchste Korrelationswahrscheinlichkeit bei einer Verschiebung von $\tau = 0$ (8,5 % wenn man die Autokorrelationen herausrechnet).

Abbildung 20 zeigt die Differenz zwischen der durchschnittlichen Korrelationswahrscheinlichkeit für negative Verschiebungen und der durchschnittlichen Korrelationswahrscheinlichkeit für positive Verschiebungen.

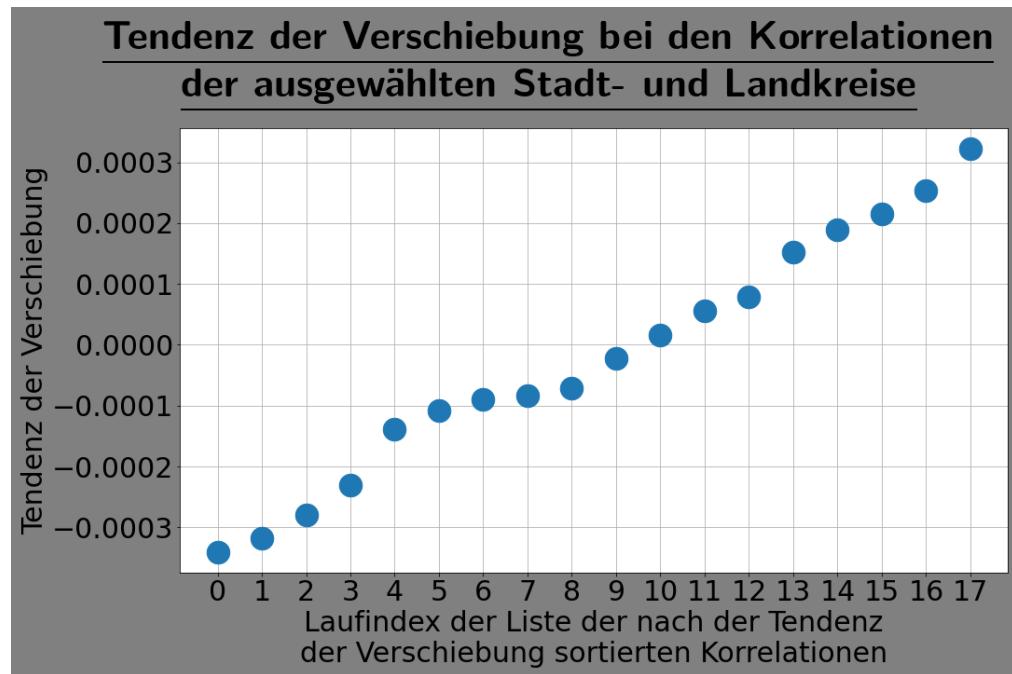


Abbildung 20: Die Differenz zwischen der durchschnittlichen Korrelationswahrscheinlichkeit aller negativen Verschiebungen und der durchschnittlichen Korrelationswahrscheinlichkeit aller positiven Verschiebungen für die ausgewählten Korrelationen zwischen Städten und Landkreisen.

4.6 Korrelationsmatrizen sortiert nach der Bevölkerungsdichte

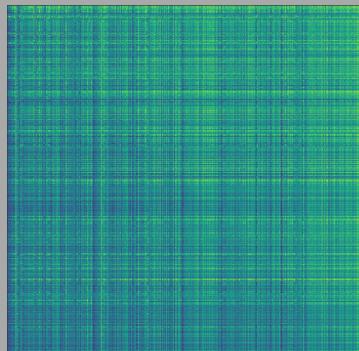
4.6.1 Korrelationsmatrizen mit den nach Bevölkerungsdichten sortierten Landkreisen

In Abbildung 21 finden sich die sechs Matrizen mit den Werten für die Korrelationen zwischen den Landkreisen. Die Zeilen und Spalten sind nach der Bevölkerungsdichte der Landkreise sortiert, eine vollständige Auflistung befindet sich im Anhang in Abschnitt 8.1.

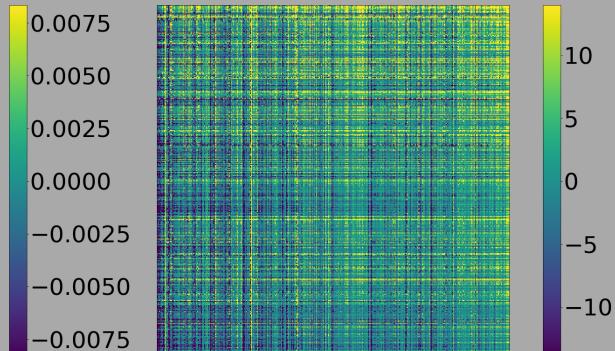
Korrelationsmatrizen der Landkreise

Zeilen und Spalten entsprechen jeweils der Liste der Landkreise sortiert nach ihren Bevölkerungsdichten

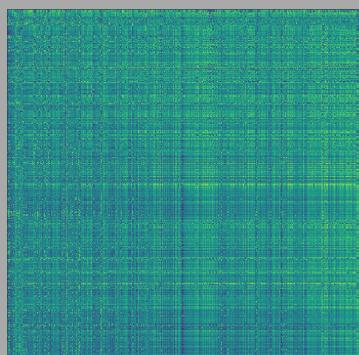
Mittel der Korrelationswerte aller Verschiebungen zwischen -14 und 14.



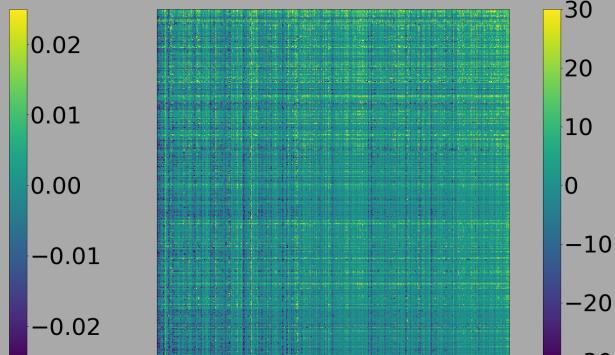
Die Verschiebung mit dem höchsten Korrelationswert zwischen -14 und 14.



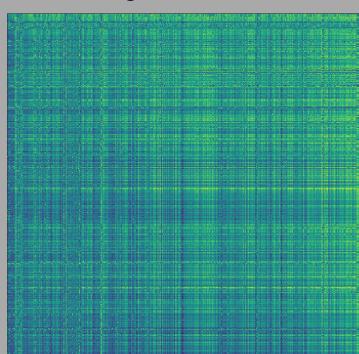
Mittel der Korrelationswerte aller Verschiebungen zwischen -30 und 30.



Die Verschiebung mit dem höchsten Korrelationswert zwischen -30 und 30.



Mittel der Korrelationswerte aller Verschiebungen zwischen -50 und 50.



Die Verschiebung mit dem höchsten Korrelationswert zwischen -50 und 50.

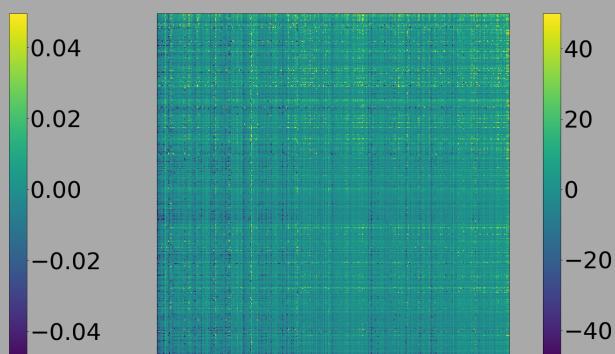


Abbildung 21: Verschiedene Korrelationsmatrizen berechnet aus den Korrelationen jedes deutschen Landkreises mit sich selbst und jedem anderen deutschen Landkreis. Jeder Zeile und Spalte sind Landkreise zugeordnet, die vollständige Auflistung befindet sich im Anhang. Die Farben der Zellen auf der linken Seite entsprechen der Differenz zwischen der durchschnittlichen Korrelationswahrscheinlichkeit für negative Verschiebungen des Landkreises der Spalte in Relation zum Landkreis der Zeile und der durchschnittlichen Korrelationswahrscheinlichkeit für die zugehörigen positiven Verschiebungen. Auf der rechten Seite wird die Zelle entsprechend der Verschiebung der Zeitserie des Landkreises der Spalte entgegen der Zeitserie der Zeile mit der höchsten Korrelationswahrscheinlichkeit eingefärbt. Beide Vorgehensweisen werden für alle ganzzahligen Verschiebungen

(TODO: Richtige Richtung? Nachprüfen! (Spalte und Zeile evtl. tauschen))

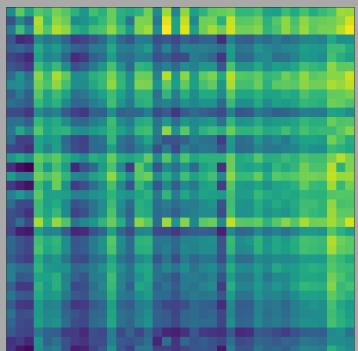
4.6.2 Korrelationsmatrizen mit den nach Bevölkerungsdichten sortierten Regierungsbezirken

In Abbildung 22 finden sich die sechs Matrizen mit den Werten für die Korrelationen zwischen den Bezirken. Die Zeilen und Spalten sind nach der Bevölkerungsdichte der Regierungsbezirke sortiert, eine vollständige Auflistung befindet sich im Anhang in Abschnitt 8.3.

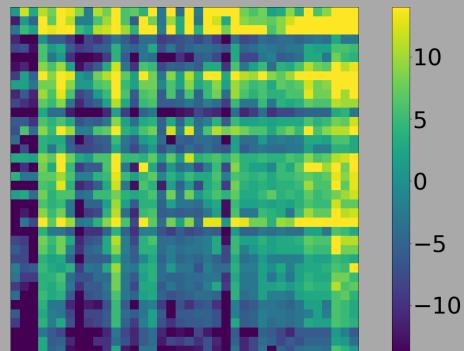
Korrelationsmatrizen der Regierungsbezirke

Zeilen und Spalten entsprechen jeweils der Liste der Regierungsbezirke sortiert nach ihren Bevölkerungsdichten

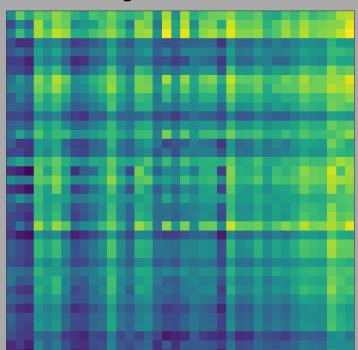
Mittel der Korrelationswerte aller Verschiebungen zwischen -14 und 14.



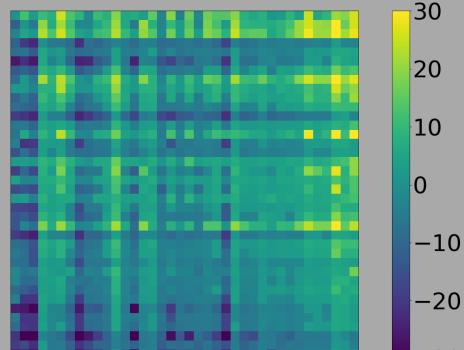
Die Verschiebung mit dem höchsten Korrelationswert zwischen -14 und 14.



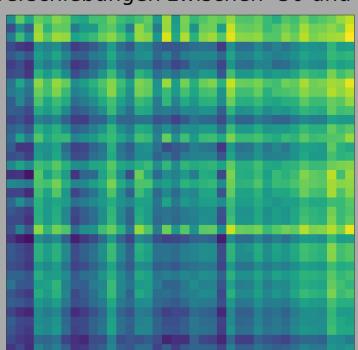
Mittel der Korrelationswerte aller Verschiebungen zwischen -30 und 30.



Die Verschiebung mit dem höchsten Korrelationswert zwischen -30 und 30.



Mittel der Korrelationswerte aller Verschiebungen zwischen -50 und 50.



Die Verschiebung mit dem höchsten Korrelationswert zwischen -50 und 50.

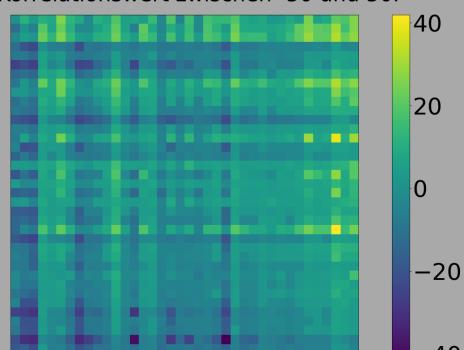


Abbildung 22: Verschiedene Korrelationsmatrizen berechnet aus den Korrelationen jedes deutschen Regierungsbezirks mit sich selbst und jedem anderen deutschen Regierungsbezirk. Jeder Zeile und Spalte sind Regierungsbezirke zugewiesen, die vollständige Auflistung befindet sich im Anhang. Die Farben der Zellen auf der linken Seite entsprechen der Differenz zwischen der durchschnittlichen Korrelationswahrscheinlichkeit für negative Verschiebungen des Regierungsbezirks der Spalte in Relation zum Regierungsbezirk der Zeile und der durchschnittlichen Korrelationswahrscheinlichkeit für die zugehörigen positiven Verschiebungen. Auf der rechten Seite wird die Zelle entsprechend der Verschiebung der Zeitserie des Regierungsbezirks der Spalte entgegen der Zeitserie der Zeile mit der höchsten Korrelationswahrscheinlichkeit eingefärbt. Beide Vorgehensweisen werden

(TODO: Richtige Richtung? Nachprüfen! (Spalte und Zeile evtl. tauschen))

4.7 Korrelationsmatrizen sortiert nach dem Gemeindeschlüssel

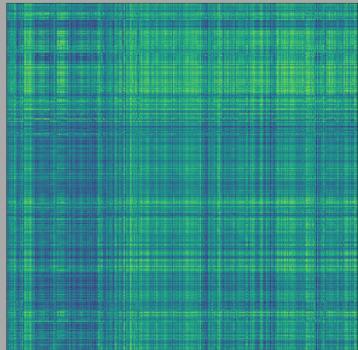
4.7.1 Korrelationsmatrizen mit den nach Gemeindeschlüsseln sortierten Landkreise

In Abbildung 23 finden sich die sechs Matrizen mit den Werten für die Korrelationen zwischen den Landkreisen. Die Zeilen und Spalten sind lexikographisch nach den Gemeindeschlüsseln der Landkreise sortiert, dies entspricht in etwa einem Nord-Süd-Verlauf (siehe (TODO: Nord-Süd-verlauf Bild verlinken)). Eine vollständige Auflistung befindet sich im Anhang in Abschnitt 8.2.

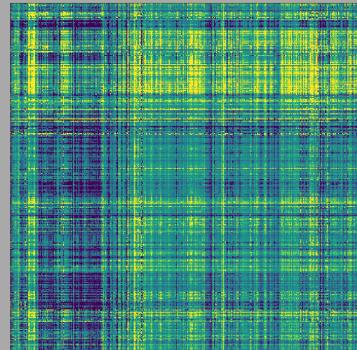
Korrelationsmatrizen der Landkreise

Zeilen und Spalten entsprechen jeweils der Liste der Landkreise sortiert nach ihren Gemeindeschlüsseln

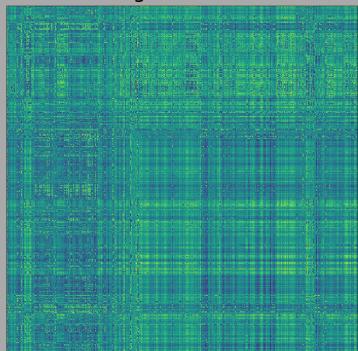
Mittel der Korrelationswerte aller Verschiebungen zwischen -14 und 14.



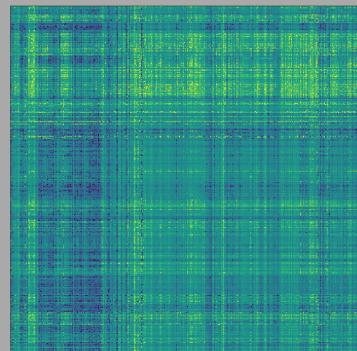
Die Verschiebung mit dem höchsten Korrelationswert zwischen -14 und 14.



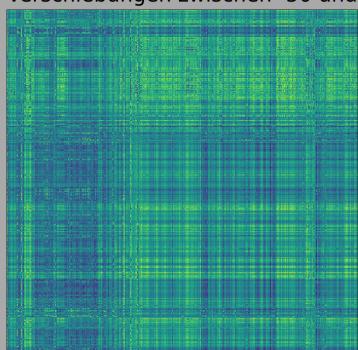
Mittel der Korrelationswerte aller Verschiebungen zwischen -30 und 30.



Die Verschiebung mit dem höchsten Korrelationswert zwischen -30 und 30.



Mittel der Korrelationswerte aller Verschiebungen zwischen -50 und 50.



Die Verschiebung mit dem höchsten Korrelationswert zwischen -50 und 50.

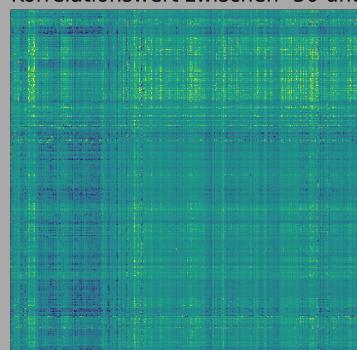


Abbildung 23: Verschiedene Korrelationsmatrizen berechnet aus den Korrelationen jedes deutschen Landkreises mit sich selbst und jedem anderen deutschen Landkreis. Jeder Zeile und Spalte sind Landkreise zugeordnet, die vollständige Auflistung befindet sich im Anhang. Die Farben der Zellen auf der linken Seite entsprechen der Differenz zwischen der durchschnittlichen Korrelationswahrscheinlichkeit für negative Verschiebungen des Landkreises der Spalte in Relation zum Landkreis der Zeile und der durchschnittlichen Korrelationswahrscheinlichkeit für die zugehörigen positiven Verschiebungen. Auf der rechten Seite wird die Zelle entsprechend der Verschiebung der Zeitserie des Landkreises der Spalte entgegen der Zeitserie der Zeile mit der höchsten Korrelationswahrscheinlichkeit eingefärbt. Beide Vorgehensweisen werden für alle ganzzahligen Verschiebungen

(TODO: Richtige Richtung? Nachprüfen! (Spalte und Zeile evtl. tauschen))

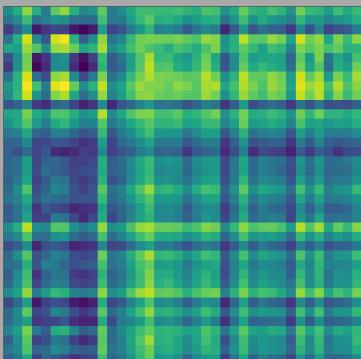
4.7.2 Korrelationsmatrizen mit den nach Gemeindeschlüsseln sortierten Regierungsbezirken

In Abbildung 24 finden sich die sechs Matrizen mit den Werten für die Korrelationen zwischen den Bezirken. Die Zeilen und Spalten sind lexikographisch nach den Gemeindeschlüsseln der Landkreise in den Regierungsbezirken sortiert, eine vollständige Auflistung befindet sich im Anhang in Abschnitt 8.4

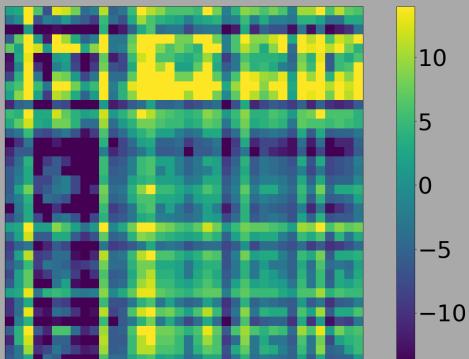
Korrelationsmatrizen der Regierungsbezirke

Zeilen und Spalten entsprechen jeweils der Liste der Regierungsbezirke sortiert nach ihren Gemeindeschlüssen

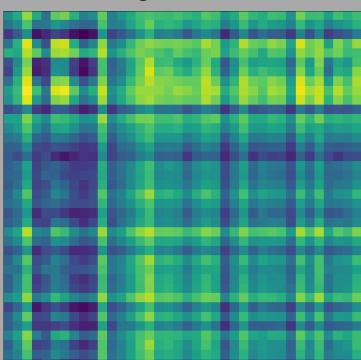
Mittel der Korrelationswerte aller Verschiebungen zwischen -14 und 14.



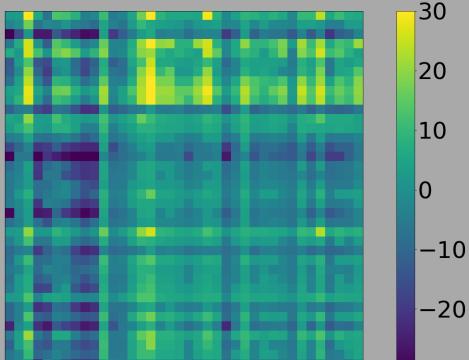
Die Verschiebung mit dem höchsten Korrelationswert zwischen -14 und 14.



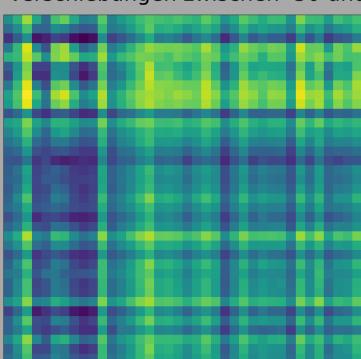
Mittel der Korrelationswerte aller Verschiebungen zwischen -30 und 30.



Die Verschiebung mit dem höchsten Korrelationswert zwischen -30 und 30.



Mittel der Korrelationswerte aller Verschiebungen zwischen -50 und 50.



Die Verschiebung mit dem höchsten Korrelationswert zwischen -50 und 50.

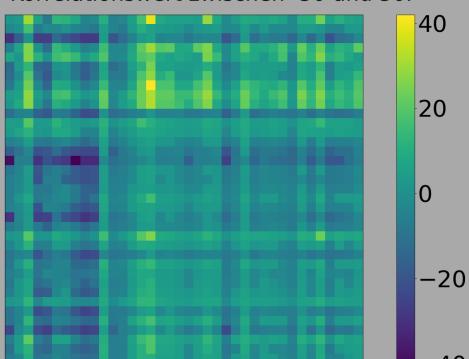


Abbildung 24: Verschiedene Korrelationsmatrizen berechnet aus den Korrelationen jedes deutschen Regierungsbezirks mit sich selbst und jedem anderen deutschen Regierungsbezirk. Jeder Zeile und Spalte sind Regierungsbezirke zugewiesen, die vollständige Auflistung befindet sich im Anhang. Die Farben der Zellen auf der linken Seite entsprechen der Differenz zwischen der durchschnittlichen Korrelationswahrscheinlichkeit für negative Verschiebungen des Regierungsbezirks der Spalte in Relation zum Regierungsbezirk der Zeile und der durchschnittlichen Korrelationswahrscheinlichkeit für die zugehörigen positiven Verschiebungen. Auf der rechten Seite wird die Zelle entsprechend der Verschiebung der Zeitserie des Regierungsbezirks der Spalte entgegen der Zeitserie der Zeile mit der höchsten Korrelationswahrscheinlichkeit eingefärbt. Beide Vorgehensweisen werden

(TODO: Richtige Richtung? Nachprüfen! (Spalte und Zeile evtl. tauschen))

4.8 Durchschnittliche Verschiebung im nationalen Vergleich

4.8.1 Verteilung der durchschnittlichen Verschiebung unter den Landkreise

In den folgenden sind die deutschen Landkreise entsprechend dem Durchschnitt der Werte ihrer Zeile in den obigen Matrizen eingefärbt.

In Abbildung 25 sind die Landkreise entsprechend der Durchschnitte ihrer Zeilen in der rechten unteren Matrix aus Abbildung 21 bzw. Abbildung 23 eingefärbt. Links daneben ist die Verteilung dargestellt.

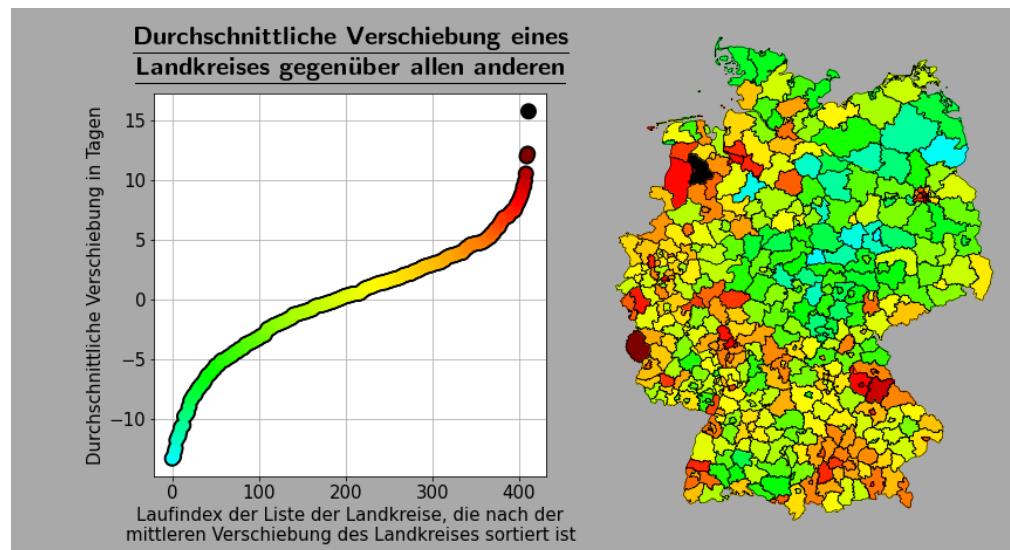


Abbildung 25: Auf der rechten Seite die deutschen Landkreise eingefärbt im Durchschnitt der Maximalwerte der Korrelationen mit allen Landkreisen und Verschiebungen $\tau \in [-50, 50]$. Auf der linken Seite die Verteilung dieser Werte.

In Abbildung 26 sind die Landkreise entsprechend der Durchschnitte ihrer Zeilen in der linken unteren Matrix aus Abbildung 21 bzw. Abbildung 23 eingefärbt. Links daneben ist die Verteilung dargestellt.

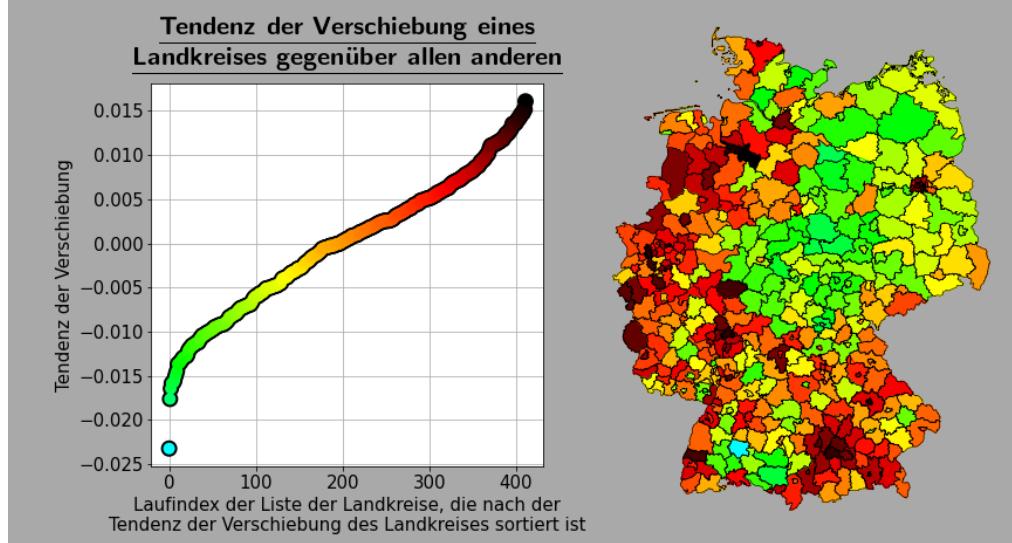


Abbildung 26: Auf der rechten Seite die deutschen Landkreise eingefärbt im Durchschnitt der Differenz der Korrelationswahrscheinlichkeiten der positiven Verschiebungen zu den Korrelationswahrscheinlichkeiten der negativen Verschiebungen der Korrelationen dieses Landkreises mit allen Landkreisen und Verschiebungen $\tau \in [-50, 50]$. Auf der linken Seite die Verteilung dieser Werte.

4.8.2 Verteilung der durchschnittlichen Verschiebung unter den Regierungsbezirken

In den folgenden sind die deutschen Regierungsbezirke entsprechend dem Durchschnitt der Werte ihrer Zeile in den obigen Matrizen eingefärbt.

In Abbildung 27 sind die Regierungsbezirke entsprechend der Durchschnitte ihrer Zeilen in der rechten unteren Matrix aus Abbildung 22 bzw. Abbildung 24 eingefärbt. Links daneben ist die Verteilung dargestellt.

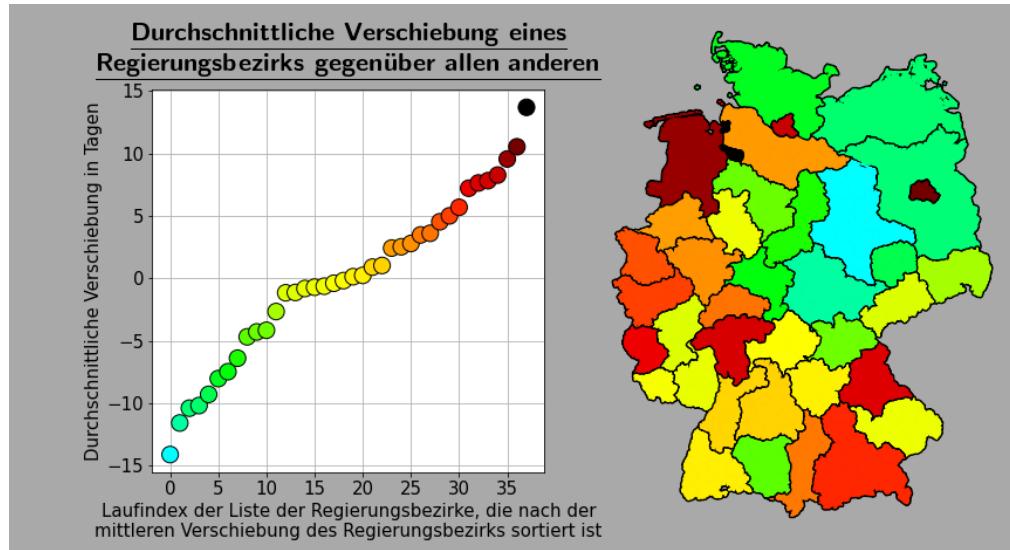


Abbildung 27: Auf der rechten Seite die deutschen Regierungsbezirke eingefärbt im Durchschnitt der Maximalwerte der Korrelationen mit allen Regierungsbezirken und Verschiebungen $\tau \in [-50, 50]$. Auf der linken Seite die Verteilung dieser Werte.

In Abbildung 28 sind die Regierungsbezirke entsprechend der Durchschnitte ihrer Zeilen in der linken unteren Matrix aus Abbildung 22 bzw. Abbildung 24 eingefärbt. Links daneben ist die Verteilung dargestellt.

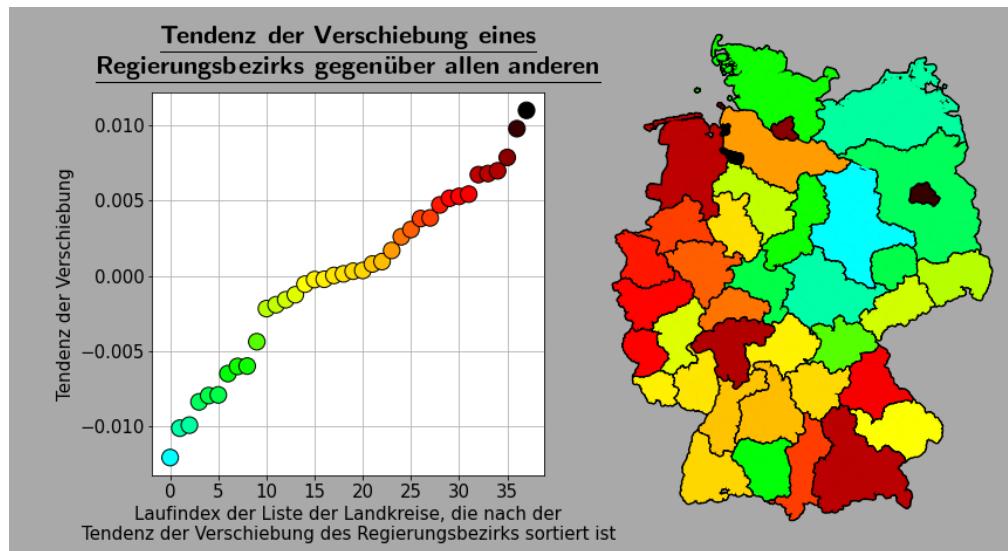


Abbildung 28: Auf der rechten Seite die deutschen Regierungsbezirke eingefärbt im Durchschnitt der Differenz der Korrelationswahrscheinlichkeiten der positiven Verschiebungen zu den Korrelationswahrscheinlichkeiten der negativen Verschiebungen der Korrelationen dieses Regierungsbezirks mit allen Regierungsbezirken und Verschiebungen $\tau \in [-50, 50]$. Auf der linken Seite die Verteilung dieser Werte.

5 Diskussion

5.1 Aussagekraft des SIR-Modells

In Abbildung 18 sind die drei Kennzahlen für das SIR-Modell dargestellt. Die R-Wert Berechnung des Robert-Koch-Instituts basiert auf den selben Daten [], jedoch werden hierfür noch einige weitere Maßnahmen und Annahmen getroffen, welche das Maß dieser Bachelorarbeit sprengen würden. Wie auch in den Veröffentlichungen und Abschnitt 4.4 beschrieben, weisen die Meldungen der Daten überraschend viele Fehler auf:

Und dies sind lediglich die automatisch festgestellten Fehler. Inwiefern die Daten nicht das Infektionsverhalten wiederspiegeln, lässt sich anhand der wöchentlichen Schwankungen der Infizierendenzahl abschätzen: Vermutlich werden am Wochenende weniger Tests durchgeführt oder weniger Meldungen abgegeben. Zudem spiegeln die Zahlen nur die getesteten Fälle wieder. Somit kann zum einen verstärktes Testen zu späteren Zeiträumen in der Pandemie die Zahlen verfälschen und zum anderen machen symptomfreie Krankheitsverläufe es schwer, festzustellen ob und wie lange eine Person infektiös war.

Da die Anzahl der Infizierenden in einem SIR-Modell sehr klein ist und daher kleine Unregelmäßigkeiten eine große Auswirkung haben, lässt sich anhand dieses simplen SIR-Modells keine stichfeste Aussage treffen. Im Rahmen dieser Arbeit dient es als Mahnmal, dass die erwähnten Probleme stets beachtet werden müssen: Die täglichen Schwankungen werden dadurch ausgeglichen, dass die 7-Tages Inzidenz verwendet

wird. Die Verwendung von den täglichen Fallzahlen des RKIs als offizielle Stelle sorgt für eine Vorsortierung der Meldungen. Zudem werden im weiteren Verlauf ausschließlich die Fallzahlen benutzt, da diese höchste Größenordnungen erreichen und daher Inkonsistenzen von einigen 100 Fällen pro Tag weniger verheerend sind. (**TODO: Vergleiche mit (kommentar siehe Code) Quelle einpflegen**) (**TODO: Eventuell anregung aus dem Video zu SIR einbinden**)

5.2 Zusammenhänge von Bevölkerungsdichten und Corona-Zahlen der Landkreisen

Klar zu erkennen in Abbildung 13 sind die vielen kleinen roten Punkte, welche die Stadtkreise kennzeichnen, wie auch im Anhang in Abschnitt 8.1 klar zu erkennen. Diese sind jedoch oftmals von Landkreisen mit vergleichbar niedrigen Bevölkerungsdichten (außer im Rhein-Ruhrgebiet) umgeben. Daher lässt sich systematisch nur sehr schwer zwischen Landkreisen mit wirklich niedriger Bevölkerungsdichte unterscheiden und Landkreisen mit niedriger Bevölkerungsdichte, welche eine Stadt umschließen und eine deutlich höhere Bevölkerungsdichte aufweisen würden, wenn die Stadt zum Landkreis gehören würde. Da somit manche mittelmäßig dicht bevölkerte Gebiete nicht trivial in einen stark bevölkerten Stadtkreis und einen dünn bevölkerten Landkreis zerlegt sind, lassen sich keine Schlussfolgerungen ziehen, inwiefern dünn bevölkerte Gebiete stärker betroffen sind als dichter bevölkerte Gebiete, dies spiegelt sich auch in den im Vergleich zu den restlichen Matrizen recht homogenen Matrizen in Abbildung 21 wieder. Jedoch lässt sich die Aufteilung vorzüglich nutzen, um festzustellen, ob die 7-Tages Inzidenzen einer Stadt vor den 7-Tages Inzidenzen im Umland steigen. Abbildung 19 zeigt jedoch eindeutig, dass die 7-Tages Inzidenzen der gewählten Städte und Landkreise außergewöhnlich zeitgleich steigen und fallen: Die höchste Korrelationswahrscheinlichkeit bei einer Verschiebung von $\tau = 0$ bei 50% ist derart über dem nationalen Durchschnitt von ca. 8,7%, das davon ausgegangen werden kann, dass sich ein Landkreis, der eine Stadt komplett umgibt einen sehr ähnlichen Verlauf

der 7-Tages Inzidenz aufweisen wird. Auch die sehr kleinen in ?? eingetragenen Differenzen zwischen dem Durchschnitt der Korrelationswahrscheinlichkeiten der positiven Verschiebungen und dem Durchschnitt der Korrelationswahrscheinlichkeiten der negativen Verschiebungen: die maximalsten Differenzen entsprechen lediglich 10% des Maximalwerts und sind ausgeglichen oberhalb und unterhalb der null angeordnet.

Da die Inzidenzwerte jedoch zuvor normiert wurden und wie in Abbildung 2 zu sehen manche Städte deutlich stärker als ihr Umland betroffen waren - oder umgekehrt - lässt sich dies nicht über die absoluten Werte sagen, sondern lediglich über die Anstiege und Abnahmen sagen. Zudem ist die geringe Größe der Teilmenge mit 18 untersuchten Paaren stets zu berücksichtigen, auch wenn kein trivialer Zusammenhang zwischen dem Ergebnis und der Auswahlmethode besteht. Es wurden die Städte gewählt, welche vollständig von einem Landkreis umgeben sind,

5.3 Zusammenhänge von Bevölkerungsdichten und Corona-Zahlen der Regierungsbezirke

Um das oben beschriebene Problem von nicht trivial aufgeteilten Gebieten zu beheben und einen Zusammenhang zwischen Bevölkerungsdichten und Corona-Zahlen herstellen zu können, werden statt der Landkreise die Regierungsbezirke verwendet.

Alle Matrizen in Abbildung 22 scheinen einen Farbverlauf von links unten nach rechts oben zu besitzen, in der Matrix rechts oben ist es am deutlichsten ersichtlich. Dies spricht stark dafür, dass die 7-Tages Inzidenz der bevölkerungsdichten Regierungsbezirke früher ansteigt. Der Farbverlauf wirkt relativ kontinuierlich, was den Verdacht auf einen Zusammenhang noch weiter verstärkt.

Dieser Effekt entspricht durchaus der ersten Intuition, da sich, wie der Name schon sagt, in dichter bevölkerten Gebieten mehr Menschen auf engerem Raum bewegen und dadurch von mehr Infizierenden schneller angesteckt werden. Zudem sind die Treffpunkte in Städten vermutlich deutlich größer: Größere Unternehmen, Universitäten,

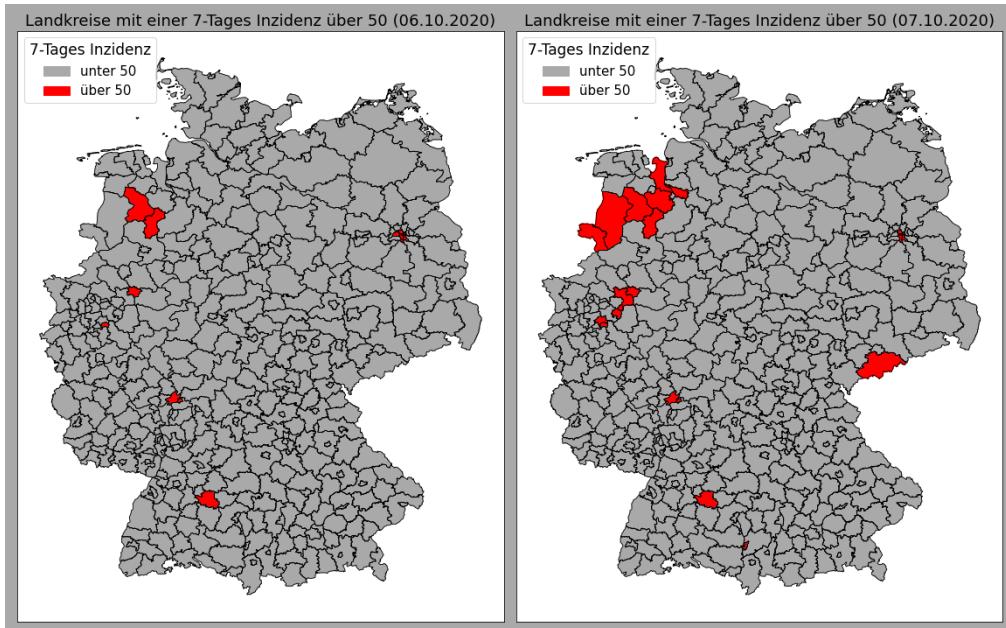


Abbildung 29: Die Deutschlandkarte, rot eingefärbt jeweils die Landkreise mit einer 7-Tages Inzidenz am 06.10.2020 beziehungsweise 07.10.2020 über 50.

größere Diskotheken, größere Wohneinheiten, stärker frequentierte Knotenpunkte des Nah- und Fernverkehrs etc. Dadurch kann es zu mehr und stärkeren Superspreading-Events, Ereignisse, an denen sich eine Vielzahl an Menschen an wenigen Infizierten ansteckt, kommen.

Die Superlative dieser Superspreading-Events stören jedoch auch die Regelmäßigkeit des angesprochenen Farbverlaufs in den Matrizen: Beispielsweise kam es laut Zeit [] im Oktober 2020 im Landkreis Cloppenburg zu einem Corona Ausbruch in einem Schlachthof. Erstellt man für den 06.10.2020 und den 07.10.2020 eine Karte mit allen Landkreisen mit einer 7-Tages Inzidenz über 50 rot ein, ergibt sich die Karte in Abbildung 29.

(**TODO: Zeit Artikel referenzieren, siehe hier Kommentar:**) Klar erkennbar: Die extrem starke Ausbreitung des Corona-Viruses in diesem Gebiet. Vermutlich sorgte der Corona Ausbruch in Cloppenburg für einen rapiden Anstieg in diesem

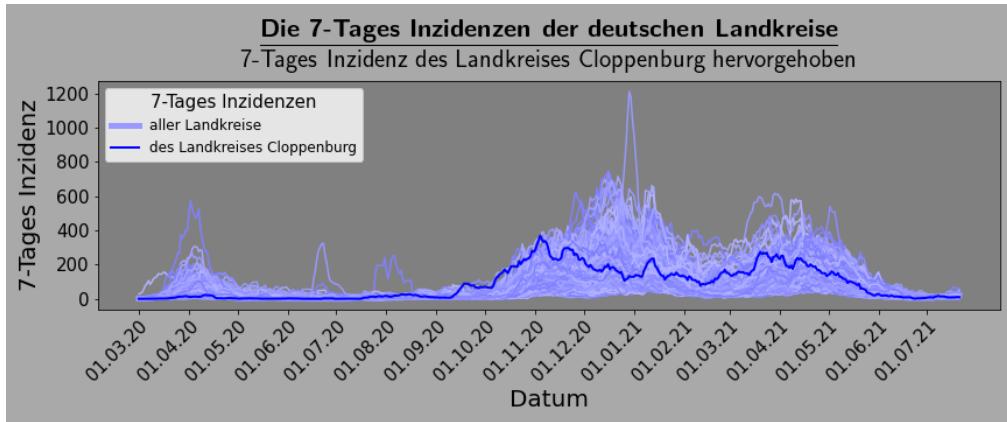


Abbildung 30: Die 7-Tages Inzidenzen der deutschen Landkreise, in Blau die 7-Tages Inzidenz von Cloppenburg.

Gebiet in diesem Zeitraum, in welchem die Anomalie in der Matrix begründet liegt. In Abbildung 30 ist zudem klar ersichtlich, dass die 7-Tages Inzidenz in Cloppenburg nur in diesem Zeitraum zu den höchsten zählte und sich anschließend eher im Mittelfeld bewegte.

Jedoch muss nicht immer ein außerordentlicher Corona Ausbruch für eine auffällige Färbung in der Matrix sorgen: Die vierte Zeile in den Matrizen, welche dem Regierungsbezirk Trier zugeordnet ist, ist erstaunlich dunkel. Vergleicht man den Inzidenzverlauf des Regierungsbezirks mit den Inzidenzverläufen der restlichen Regierungsbezirke, scheint die 7-Tages Inzidenz des Regierungsbezirkes Trier zu Beginn durchschnittlich früh und stark angewachsen zu sein wie die restlichen, was gegen ein markantes Superspreading-Event spricht. Im Vergleich zu den anderen Regierungsbezirke blieb die 7-Tages Inzidenz jedoch ungewöhnlich stabil, wie in Abbildung ?? zu sehen. Die insgesamt niedrige Summe aller täglichen 7-Tages Inzidenzen des Regierungsbezirks, welche in Abbildung 17 erkennbar ist, stützt diese These, das die 7-Tages Inzidenz im nationalen Vergleich zwar früh stieg, aber durchgehend bei relativ niedrigen Werten blieb.

Die Schlussfolgerungen, welche anhand der Matrizen aufgestellt wurden, werden auch von Abbildung 27 und Abbildung 28 gestützt: Bis auf die beiden erwähnten

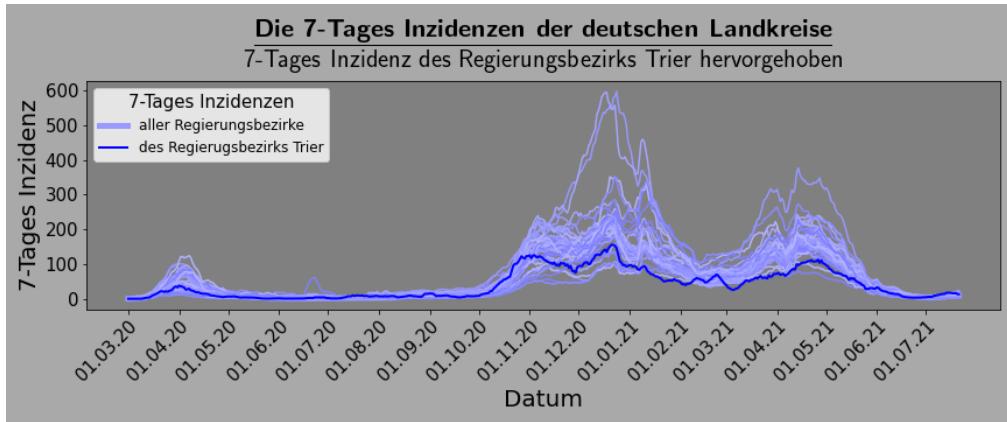


Abbildung 31: Der Verlauf der 7-Tages Inzidenzen der deutschen Regierungsbezirke.
Der Verlauf des Regierungsbezirks Trier ist blau hervorgehoben.

Ausnahmen, ähneln die beiden Deutschlandkarten Abbildung 14.

5.4 Zusammenhänge von Gemeindeschlüssel und Corona-Zahlen der Landkreise und Regierungsbezirke

Die Matrizen aus Abbildung 23 und Abbildung 24 zeigen keinen konstanten Farbverlauf: Viel eher sind die obersten Zeilen bis auf wenige Ausnahmen sehr hell, sie sind den neuen Bundesländern, Schleswig-Holstein und dem Saarland zugeordnet. Der mittlere und untere Teil, welche grob den Gebieten der alten Bundesländer zugeordnet sind, sind eher dunkler.

Der dunkle Strich in der dritten Zeile der Matrizen in Abbildung 24 ist Berlin zugeordnet. Die Interpretation dieser Zeile findet sich in Abschnitt 5.2. Hamburg, welches für den dunklen Strich in den selben Matrizen in Zeile 11 verantwortlich ist, fällt in die selbe Kategorie.

Das Saarland, welches in Abbildung 24 laut Abschnitt 8.4 der zweiten Zeile zugeordnet ist, ist in allen Matrizen in der oberen Hälfte eine der dunkleren Linien, im unteren Teil würde sie vermutlich nicht auffallen. Dies ist wenig verwunderlich, da das Saarland

bei der Sortierung nach dem Gemeindeschlüssel am auffälligsten aus dem Nord-Süd- und Ost-West-Raster herausfällt, wie in Abbildung 15 zu sehen.

Somit scheinen die 7-Tages Inzidenzen in den alten Bundesländern vor den 7-Tages Inzidenzen in den neuen Bundesländern zu steigen.

5.5 Durchschnittliche Verschiebung im nationalen Vergleich Landkreise

Abbildung 25 und Abbildung 26 weisen die selben Merkmale auf wie die analysierten Abbildungen und stützen daher die genannten Thesen.

6 Zusammenfassung

Auch mit diesen simplen Methoden ließ sich ein wenig des Mysteriums lüften, wie sich das Virus in Deutschland verhält: Nach den Erkenntnissen dieser Arbeit scheinen Anstiege in den 7-Tages Inzidenzen der deutschen Stadtstaaten bis jetzt als Vorauswarnung gut geeignet gewesen zu sein. Die vielen Anomalien wie beispielsweise das Superspreading-Event in Cloppenburg im Oktober 2020 zeigen jedoch auf, dass sich diese Ausbrüche nicht an trivialen Charakteristiken der Landkreise, Regierungsbezirke oder Bundesländer festmachen lässt.

Die Meldungen der einzelnen Landkreise, welche das Robert-Koch-Institut (RKI) zur Verfügung stellt, enthalten leider ähnlich viele fragwürdige Meldungen (Genesungsdatum liegt vor dem Ansteckungsdatum oder außergewöhnlich lange Infektionszeiten) wie die Zahl der aktiven Fälle maximal annimmt: 46.033 von 3.748.038 Fälle stammen aus fragwürdigen Meldungen, die maximale Zahl der aktiven Fälle liegt bei circa 60.000. Da zudem die Anzahl der aktiven Fälle innerhalb einer Woche beachtlich schwankt, lässt sich aus dem SIR-Modell nicht ohne weitere Annahmen und Maßnahmen, wie sie das RKI bei seinen Berechnungen benutzt, Wissen generieren.

Die Zahl der Ansteckungen, welche täglich vom RKI herausgegeben wird eignet sich jedoch sehr gut zur Berechnung der 7-Tages Inzidenz. Bei der Korrelation der 7-Tages Inzidenzen der deutschen Landkreise und Regierungsbezirke ergeben sich einige interessante Erkenntnisse:

- Die 7-Tages Inzidenzen von 18 Städten, welche von einem Landkreis komplett

umgeben sind, steigen nahezu zeitgleich mit den 7-Tages Inzidenzen der Landkreise. Die Auswahl der Städte hängt ausschließlich mit der geographischen Eigenheit zusammen, dass sie von komplett von einem Landkreis umschlossen sind.

- Die Inzidenzwerte scheinen in dichter besiedelten Gebieten tendenziell früher zu steigen als in dünner besiedelten Gebieten.
- Einzelne Superspreading-Events können für einen signifikanten früheren Auschlag der 7-Tages Inzidenz eines ganzen Regierungsbezirks sorgen.
- die 7-Tages Inzidenzen in den neuen Bundesländern scheinen vor den 7-Tages Inzidenzen in den alten Bundesländern zu steigen, ansonsten lässt sich jedoch kein Nord-West-Verlauf oder ähnliches feststellen.

7 Danksagung

Vielen herzlichen Dank den zahlreichen Helfern, welche mich mental und inhaltlich unterstützt haben!

Zuallererst danke ich meinem Betreuer Dr. Andreas Greiner, welcher mich vor einem Jahr mit der Idee für diese Arbeit infiziert hat und seither dieses Abenteuer mit mir gewagt hat. Er hat mich durch seine Motivation und sein Wissen auch nach tiefen Rückschlägen immer wieder dazu gebracht, weiterzumachen und neue Möglichkeiten zu finden. Vielen Dank für deine Expertise und deine höchst sympathische Art!

Zudem möchte ich mich bei Prof. Moritz Mathias Diehl bedanken, dass er sich mit seiner immensen Fachexpertise, in die ich durch seine Vorlesungen einen winzigen Einblick erhaschen durfte, mit meiner Bachelorarbeit befasst und sie begutachtet. Vielen Dank!

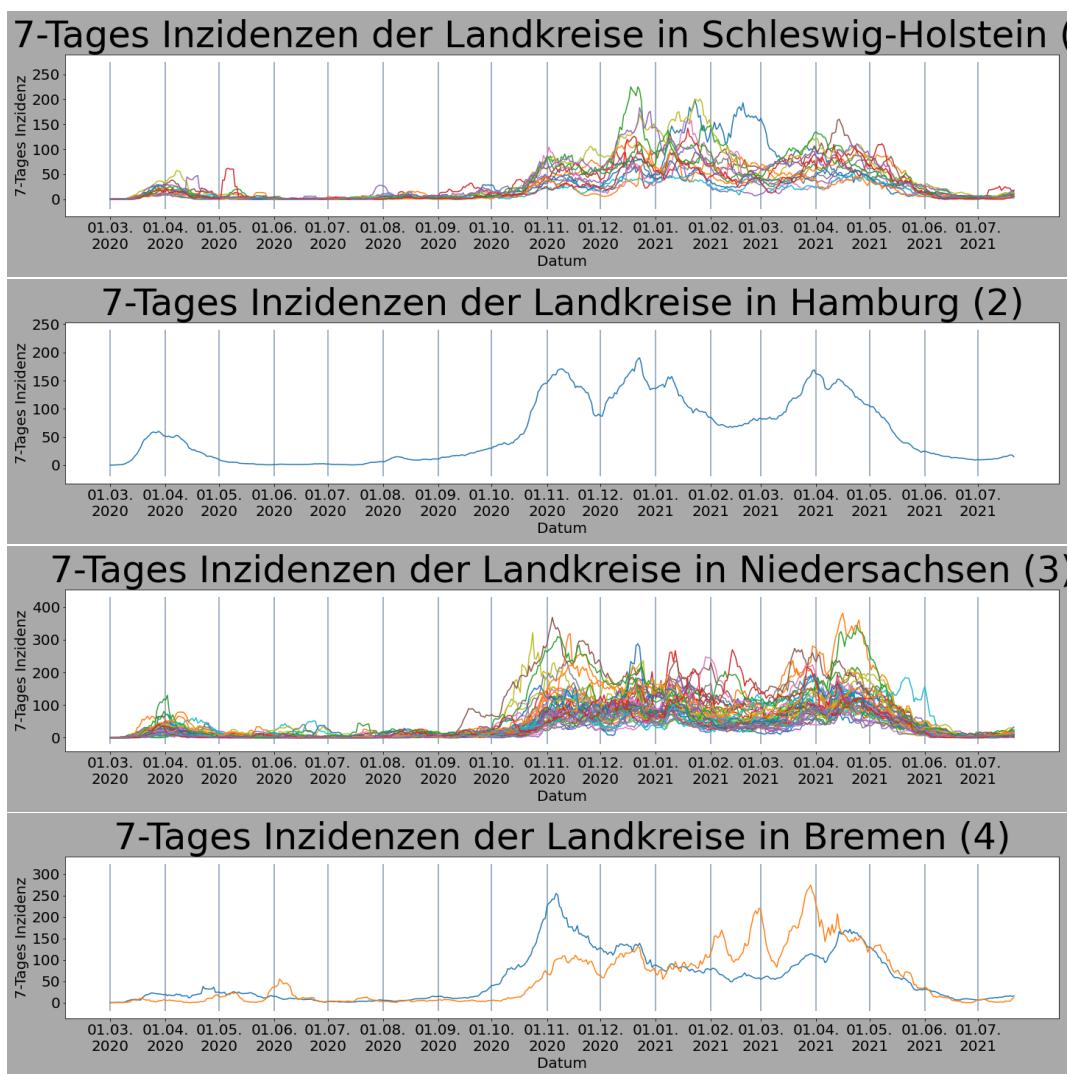
An zweiter Stelle möchte ich meinen Mitbewohnern Sarah Weitz, Sebastian Rauser und Nibischan Raveendran danken, welchen ich meine Programme und Gedankengänge an langen Abenden erklären durfte und ich so logische Fehler korrigieren oder schwer verständliche Dinge zugänglicher machen konnte. Vielen Dank für eure Geduld, eure Kochkünste und eure Wissbegierde!

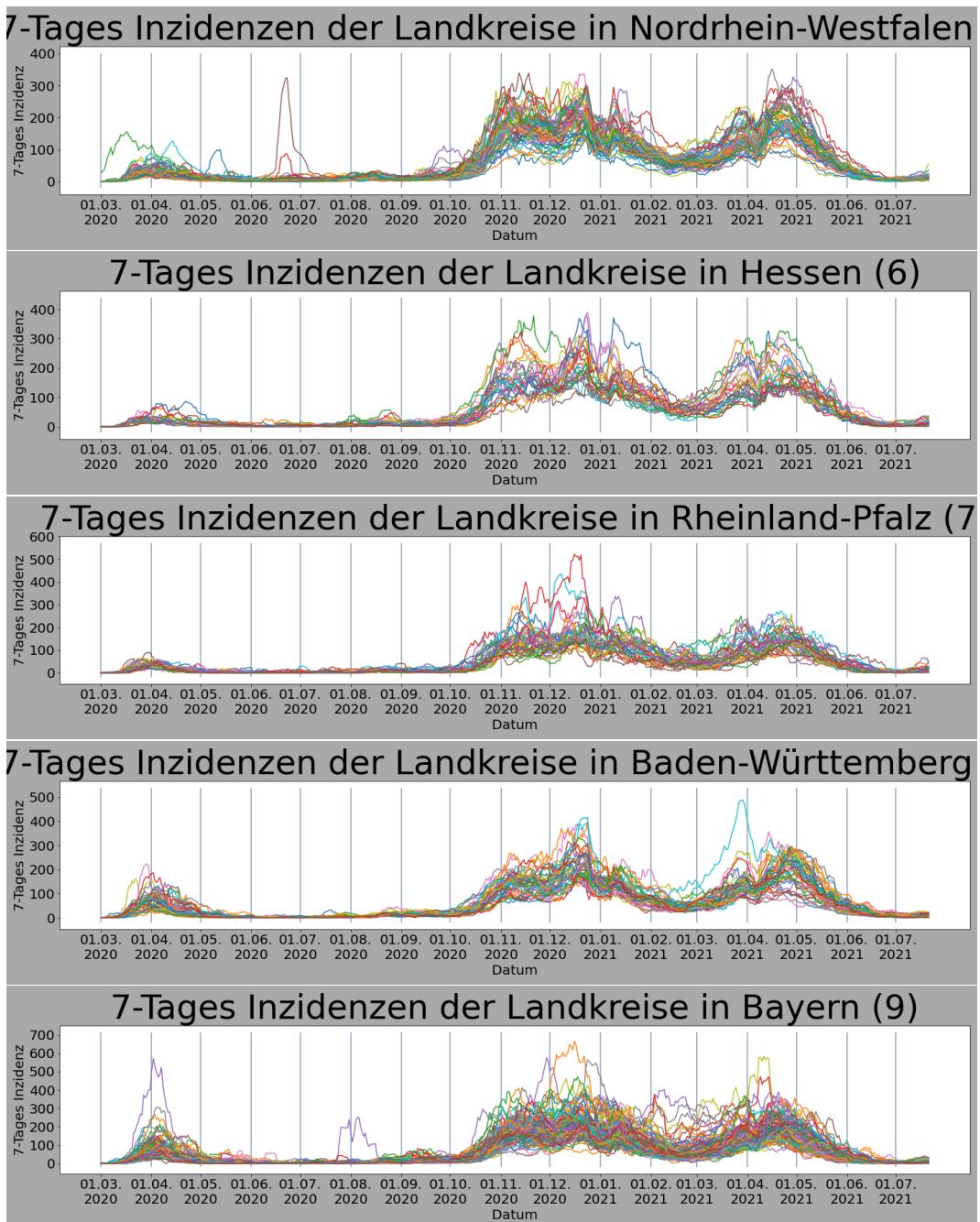
Nicht zu vernachlässigen ist zudem der Beitrag meiner Familie: Im gesamten Studium boten mir meine Geschwister Elena und Fabian Bürkin Orientierung und Halt, wenn ich nicht mehr weiter wußte oder auf einem falschen Weg war. Beide konnten mir durch ihre Erfahrungen aus dem Mikrosystemtechnik- bzw. Mathematik-Master einen

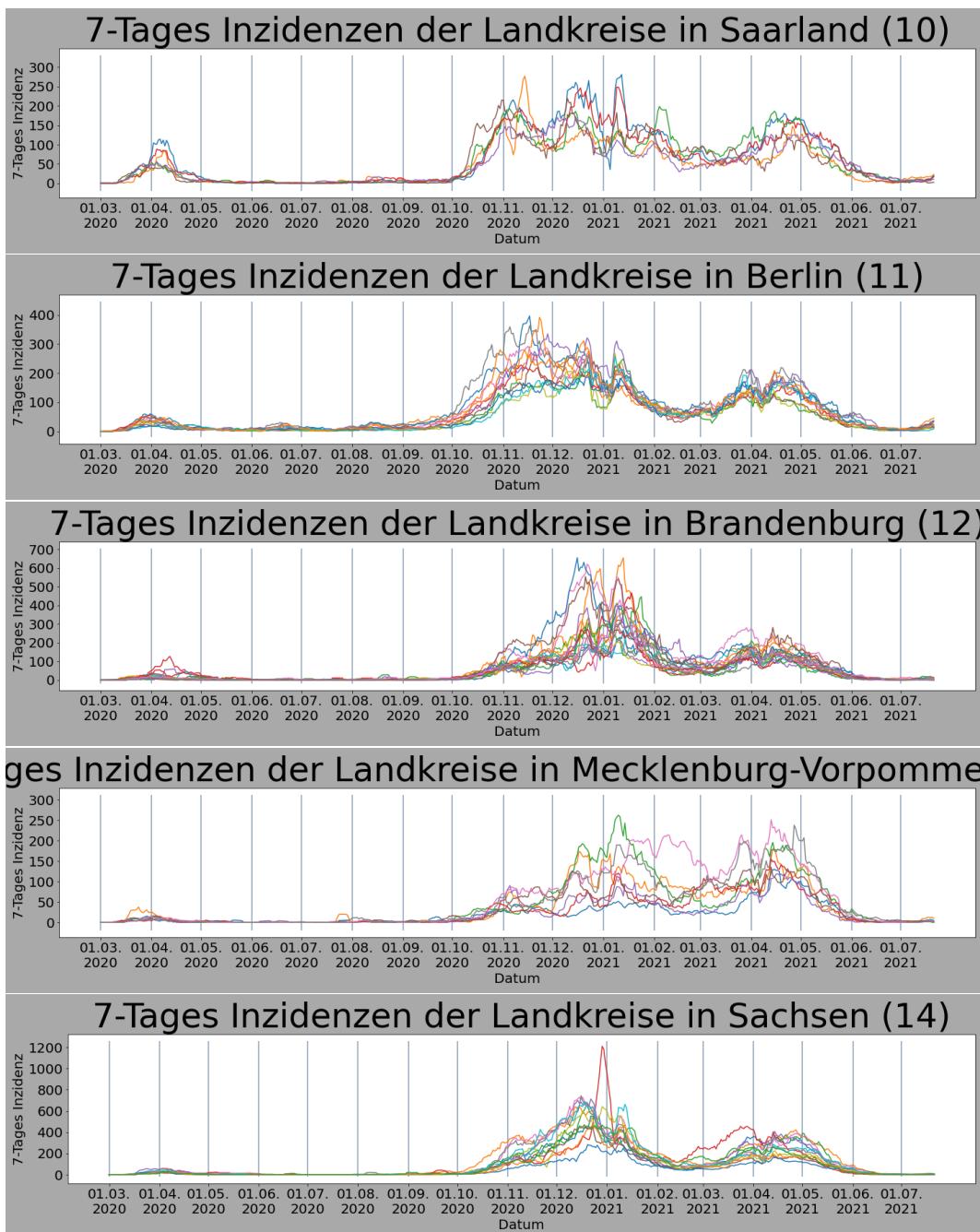
Strauß an Tipps und Tricks mitgeben, mit denen ich theoretisch schon an meinem Master wäre, wenn ich sie richtig umgesetzt hätte. Vielen herzlichen Dank für eure Geduld mit eurem kleinen Bruder und eurer Hilfsbereitschaft zu jeder Zeit!

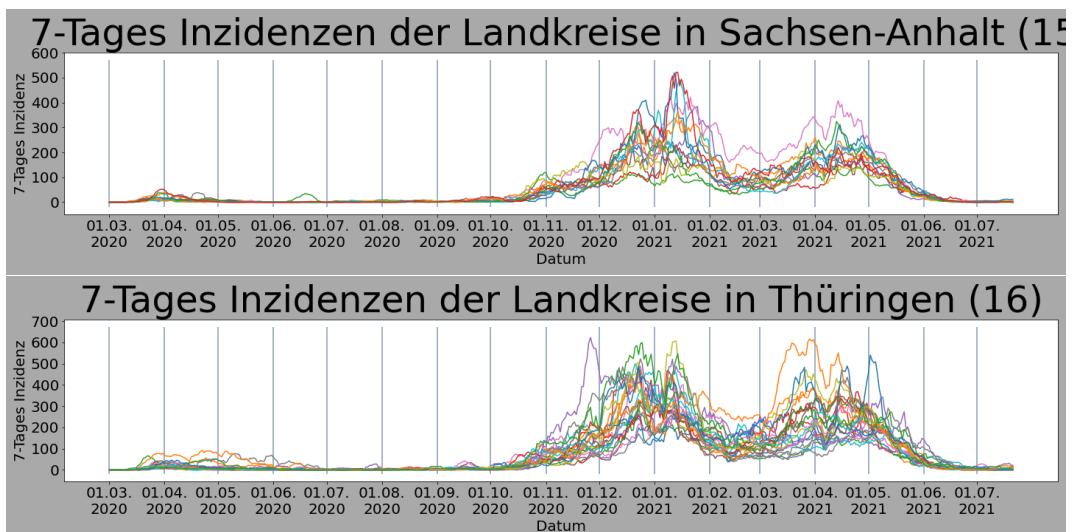
Zu guter Letzt möchte ich meinen Kollegen, Komilitonen, Korrekturlesern und vor allem Freunden Andreas Philipp, Franz Kostelezky, Lea Hohl und Ellen Hermle bedanken, die sich durch diese Arbeit gequält haben und die mir bei Dr. Andreas Greiners teils schwierigen mathematischen Ausführungen Hoffnung gegeben haben, weil sie auch nicht alles beim ersten Mal verstanden haben. Auf einen Kaffee und vielen Dank!

8 Anhang









8.1 Die deutschen Landkreise sortiert nach ihrer Bevölkerungsdichte

12070 LK Prignitz

15081 LK Altmarkkreis Salzwedel

12073 LK Uckermark

12068 LK Ostprignitz-Ruppin

3354 LK Lüchow-Dannenberg

13076 LK Ludwigslust-Parchim

15090 LK Stendal

13071 LK Mecklenburgische Seenplatte

12062 LK Elbe-Elster

15086 LK Jerichower Land

7232 LK Bitburg-Prüm

13075 LK Vorpommern-Greifswald

13072 LK Rostock

3360 LK Uelzen

15091 LK Wittenberg

9374 LK Neustadt a.d.Waldnaab

9377 LK Tirschenreuth

7233 LK Vulkaneifel

16069 LK Hildburghausen

12071 LK Spree-Neiße

16075 LK Saale-Orla-Kreis

13073 LK Vorpommern-Rügen

16065 LK Kyffhäuserkreis

15083 LK Börde

6535 LK Vogelsbergkreis

13074 LK Nordwestmecklenburg

3358 LK Heidekreis

12061 LK Dahme-Spreewald

9673 LK Rhön-Grabfeld

12067 LK Oder-Spree

3357 LK Rotenburg (Wümme)

9276 LK Regen

1054 LK Nordfriesland

9272 LK Freyung-Grafenau

9575 LK Neustadt a.d.Aisch-Bad Windsheim

12072 LK Teltow-Fläming

9472 LK Bayreuth

9371 LK Amberg-Sulzbach

12069 LK Potsdam-Mittelmark

9372 LK Cham

9278 LK Straubing-Bogen
6635 LK Waldeck-Frankenberg
16068 LK Sömmerda
3462 LK Wittmund
3256 LK Nienburg (Weser)
9180 LK Garmisch-Partenkirchen
9674 LK Haßberge
7135 LK Cochem-Zell
12066 LK Oberspreewald-Lausitz
12064 LK Märkisch-Oderland
9672 LK Bad Kissingen
16063 LK Wartburgkreis
15087 LK Mansfeld-Südharz
1051 LK Dithmarschen
9571 LK Ansbach
12063 LK Havelland
9277 LK Rottal-Inn
9677 LK Main-Spessart
3352 LK Cuxhaven
7231 LK Bernkastel-Wittlich
14730 LK Nordsachsen

9577 LK Weißenburg-Gunzenhausen

6636 LK Werra-Meißner-Kreis

7340 LK Südwestpfalz

16073 LK Saalfeld-Rudolstadt

9373 LK Neumarkt i.d.OPf.

1059 LK Schleswig-Flensburg

15085 LK Harz

9376 LK Schwandorf

9777 LK Ostallgäu

3255 LK Holzminden

8128 LK Main-Tauber-Kreis

16074 LK Saale-Holzland-Kreis

9780 LK Oberallgäu

16071 LK Weimarer Land

9476 LK Kronach

16066 LK Schmalkalden-Meiningen

7140 LK Rhein-Hunsrück-Kreis

7134 LK Birkenfeld

3155 LK Northeim

16064 LK Unstrut-Hainich-Kreis

9779 LK Donau-Ries

9475 LK Hof

16061 LK Eichsfeld

3461 LK Wesermarsch

15082 LK Anhalt-Bitterfeld

8437 LK Sigmaringen

9477 LK Kulmbach

3251 LK Diepholz

9176 LK Eichstätt

6632 LK Hersfeld-Rotenburg

9279 LK Dingolfing-Landau

3151 LK Gifhorn

3454 LK Emsland

16076 LK Greiz

9173 LK Bad Tölz-Wolfratshausen

9182 LK Miesbach

9273 LK Kelheim

3351 LK Celle

9189 LK Traunstein

6634 LK Schwalm-Eder-Kreis

16062 LK Nordhausen

5762 LK Höxter

7333 LK Donnersbergkreis
12065 LK Oberhavel
9778 LK Unterallgäu
9274 LK Landshut
1057 LK Plön
14626 LK Görlitz
9479 LK Wunsiedel i.Fichtelgebirge
3453 LK Cloppenburg
9773 LK Dillingen a.d.Donau
7336 LK Kusel
3458 LK Oldenburg
14625 LK Bautzen
12060 LK Barnim
9275 LK Passau
9172 LK Berchtesgadener Land
16070 LK Ilm-Kreis
15084 LK Burgenlandkreis
9471 LK Bamberg
1058 LK Rendsburg-Eckernförde
15088 LK Saalekreis
8225 LK Neckar-Odenwald-Kreis

9478 LK Lichtenfels
1061 LK Steinburg
9185 LK Neuburg-Schrobenhausen
15089 LK Salzlandkreis
5958 LK Hochsauerlandkreis
8127 LK Schwäbisch Hall
16072 LK Sonneberg
9675 LK Kitzingen
7235 LK Trier-Saarburg
3154 LK Helmstedt
8237 LK Freudenstadt
9678 LK Schweinfurt
9271 LK Deggendorf
3355 LK Lüneburg
9375 LK Regensburg
3456 LK Grafschaft Bentheim
9190 LK Weilheim-Schongau
3153 LK Goslar
9576 LK Roth
8426 LK Biberach
14522 LK Mittelsachsen

9183 LK Mühldorf a.Inn

16067 LK Gotha

1055 LK Ostholstein

8126 LK Hohenlohekreis

8425 LK Alb-Donau-Kreis

3452 LK Aurich

9473 LK Coburg

14628 LK Sächsische Schweiz-Osterzgebirge

9181 LK Landsberg a.Lech

8337 LK Waldshut

6437 LK Odenwaldkreis

5366 LK Euskirchen

14729 LK Leipzig

7141 LK Rhein-Lahn-Kreis

16077 LK Altenburger Land

1053 LK Herzogtum Lauenburg

9177 LK Erding

3455 LK Friesland

14523 LK Vogtlandkreis

3457 LK Leer

6631 LK Fulda

3158 LK Wolfenbüttel

7131 LK Ahrweiler

7335 LK Kaiserslautern

14627 LK Meißen

9774 LK Günzburg

9679 LK Würzburg

9186 LK Pfaffenhofen a.d.Ilm

3459 LK Osnabrück

3359 LK Stade

3451 LK Ammerland

9771 LK Aichach-Friedberg

7337 LK Südliche Weinstraße

3361 LK Verden

3356 LK Osterholz

8436 LK Ravensburg

3460 LK Vechta

9676 LK Miltenberg

9474 LK Forchheim

9187 LK Rosenheim

8325 LK Rottweil

10046 LK Sankt Wendel

6633 LK Kassel
14521 LK Erzgebirgskreis
7133 LK Bad Kreuznach
10042 LK Merzig-Wadern
3159 LK Göttingen
3252 LK Hameln-Pyrmont
5966 LK Olpe
8315 LK Breisgau-Hochschwarzwald
8327 LK Tuttlingen
9171 LK Altötting
6534 LK Marburg-Biedenkopf
5558 LK Coesfeld
8235 LK Calw
7132 LK Altenkirchen
3353 LK Harburg
7143 LK Westerwaldkreis
1060 LK Segeberg
8417 LK Zollernalbkreis
8326 LK Schwarzwald-Baar-Kreis
8136 LK Ostalbkreis
5570 LK Warendorf

8135 LK Heidenheim

9574 LK Nürnberger Land

7331 LK Alzey-Worms

7332 LK Bad Dürkheim

9178 LK Freising

5974 LK Soest

3254 LK Hildesheim

6439 LK Rheingau-Taunus-Kreis

8317 LK Ortenaukreis

6533 LK Limburg-Weilburg

3257 LK Schaumburg

9772 LK Augsburg

6532 LK Lahn-Dill-Kreis

9572 LK Erlangen-Höchstadt

5970 LK Siegen-Wittgenstein

8316 LK Emmendingen

5774 LK Paderborn

9671 LK Aschaffenburg

5566 LK Steinfurt

3157 LK Peine

5154 LK Kleve

9776 LK Lindau

9175 LK Ebersberg

5554 LK Borken

7137 LK Mayen-Koblenz

8415 LK Reutlingen

9174 LK Dachau

5770 LK Minden-Lübbecke

7334 LK Germersheim

5766 LK Lippe

6440 LK Wetteraukreis

9188 LK Starnberg

5358 LK Düren

8336 LK Lörrach

7138 LK Neuwied

5374 LK Oberbergischer Kreis

6435 LK Main-Kinzig-Kreis

8216 LK Rastatt

8125 LK Heilbronn

12051 SK Brandenburg a.d.Havel

6531 LK Gießen

1062 LK Stormarn

15001 SK Dessau-Roßlau

8435 LK Bodenseekreis

14524 LK Zwickau

10045 LK Saarpfalz-Kreis

9775 LK Neu-Ulm

8236 LK Enzkreis

7339 LK Mainz-Bingen

8335 LK Konstanz

16054 SK Suhl

6431 LK Bergstraße

5754 LK Gütersloh

9573 LK Fürth

5962 LK Märkischer Kreis

12053 SK Frankfurt (Oder)

8211 SK Baden-Baden

8117 LK Göppingen

16056 SK Eisenach

5370 LK Heinsberg

8215 LK Karlsruhe

9561 SK Ansbach

10044 LK Saarlouis

8416 LK Tübingen

5170 LK Wesel

3402 SK Emden

6432 LK Darmstadt-Dieburg

7316 SK Neustadt a.d.Weinstraße

3102 SK Salzgitter

7320 SK Zweibrücken

6434 LK Hochtaunuskreis

1056 LK Pinneberg

8119 LK Rems-Murr-Kreis

3241 Region Hannover

9179 LK Fürstenfeldbruck

7338 LK Rhein-Pfalz-Kreis

8226 LK Rhein-Neckar-Kreis

5382 LK Rhein-Sieg-Kreis

9184 LK München

10043 LK Neunkirchen

5166 LK Viersen

5758 LK Herford

7313 SK Landau i.d.Pfalz

12052 SK Cottbus

9363 SK Weiden i.d.OPf.
3103 SK Wolfsburg
6433 LK Groß-Gerau
16052 SK Gera
9764 SK Memmingen
8115 LK Böblingen
5378 LK Rheinisch-Bergischer Kreis
7317 SK Pirmasens
5362 LK Rhein-Erft-Kreis
9263 SK Straubing
3405 SK Wilhelmshaven
7312 SK Kaiserslautern
5978 LK Unna
13004 SK Schwerin
9262 SK Passau
7319 SK Worms
16055 SK Weimar
5162 LK Rhein-Kreis Neuss
9464 SK Hof
5954 LK Ennepe-Ruhr-Kreis
5334 StädteRegion Aachen

5915 SK Hamm

16051 SK Erfurt

8118 LK Ludwigsburg

10041 LK Stadtverband Saarbrücken

5562 LK Recklinghausen

9361 SK Amberg

8116 LK Esslingen

9463 SK Coburg

12054 SK Potsdam

7211 SK Trier

16053 SK Jena

6438 LK Offenbach

9565 SK Schwabach

1003 SK Lübeck

9161 SK Ingolstadt

5515 SK Münster

8421 SK Ulm

7111 SK Koblenz

6436 LK Main-Taunus-Kreis

9763 SK Kempten

9762 SK Kaufbeuren

9261 SK Landshut

14511 SK Chemnitz

7311 SK Frankenthal

9462 SK Bayreuth

1004 SK Neumünster

9661 SK Aschaffenburg

5512 SK Bottrop

5914 SK Hagen

7318 SK Speyer

15003 SK Magdeburg

5158 LK Mettmann

13003 SK Rostock

3401 SK Delmenhorst

8121 SK Heilbronn

8231 SK Pforzheim

5711 SK Bielefeld

3101 SK Braunschweig

6411 SK Darmstadt

6414 SK Wiesbaden

3404 SK Osnabrück

9461 SK Bamberg

9562 SK Erlangen

9663 SK Würzburg

8221 SK Heidelberg

4012 SK Bremerhaven

8311 SK Freiburg i.Breisgau

5120 SK Remscheid

9662 SK Schweinfurt

5116 SK Mönchengladbach

11009 SK Berlin Treptow-Köpenick

3403 SK Oldenburg

5114 SK Krefeld

14612 SK Dresden

9163 SK Rosenheim

4011 SK Bremen

15002 SK Halle

5122 SK Solingen

8212 SK Karlsruhe

1001 SK Flensburg

5117 SK Mülheim a.d.Ruhr

9362 SK Regensburg

6611 SK Kassel

14713 SK Leipzig

9761 SK Augsburg

9563 SK Fürth

5316 SK Leverkusen

5913 SK Dortmund

5124 SK Wuppertal

5112 SK Duisburg

8222 SK Mannheim

1002 SK Kiel

7314 SK Ludwigshafen

7315 SK Mainz

5314 SK Bonn

5513 SK Gelsenkirchen

2000 SK Hamburg

5911 SK Bochum

11005 SK Berlin Spandau

5315 SK Köln

5119 SK Oberhausen

9564 SK Nürnberg

5113 SK Essen

5111 SK Düsseldorf

6413 SK Offenbach

11012 SK Berlin Reinickendorf

11006 SK Berlin Steglitz-Zehlendorf

8111 SK Stuttgart

5916 SK Herne

6412 SK Frankfurt am Main

11003 SK Berlin Pankow

11010 SK Berlin Marzahn-Hellersdorf

9162 SK München

11004 SK Berlin Charlottenburg-Wilmersdorf

11011 SK Berlin Lichtenberg

11007 SK Berlin Tempelhof-Schöneberg

11008 SK Berlin Neukölln

11001 SK Berlin Mitte

11002 SK Berlin Friedrichshain-Kreuzberg

8.2 Die deutschen Landkreise lexikographisch sortiert

nach ihren Gemeindeschlüsseln

1001 SK Flensburg

1002 SK Kiel

1003 SK Lübeck

1004 SK Neumünster

10041 LK Stadtverband Saarbrücken

10042 LK Merzig-Wadern

10043 LK Neunkirchen

10044 LK Saarlouis

10045 LK Saarpfalz-Kreis

10046 LK Sankt Wendel

1051 LK Dithmarschen

1053 LK Herzogtum Lauenburg

1054 LK Nordfriesland

1055 LK Ostholstein

1056 LK Pinneberg

1057 LK Plön

1058 LK Rendsburg-Eckernförde

1059 LK Schleswig-Flensburg

1060 LK Segeberg

1061 LK Steinburg

1062 LK Stormarn

11001 SK Berlin Mitte

11002 SK Berlin Friedrichshain-Kreuzberg

11003 SK Berlin Pankow

11004 SK Berlin Charlottenburg-Wilmersdorf

11005 SK Berlin Spandau

11006 SK Berlin Steglitz-Zehlendorf

11007 SK Berlin Tempelhof-Schöneberg

11008 SK Berlin Neukölln

11009 SK Berlin Treptow-Köpenick

11010 SK Berlin Marzahn-Hellersdorf

11011 SK Berlin Lichtenberg

11012 SK Berlin Reinickendorf

12051 SK Brandenburg a.d.Havel

12052 SK Cottbus

12053 SK Frankfurt (Oder)

12054 SK Potsdam

12060 LK Barnim

12061 LK Dahme-Spreewald

12062 LK Elbe-Elster

12063 LK Havelland

12064 LK Märkisch-Oderland

12065 LK Oberhavel

12066 LK Oberspreewald-Lausitz

12067 LK Oder-Spree

12068 LK Ostprignitz-Ruppin

12069 LK Potsdam-Mittelmark

12070 LK Prignitz

12071 LK Spree-Neiße

12072 LK Teltow-Fläming

12073 LK Uckermark

13003 SK Rostock

13004 SK Schwerin

13071 LK Mecklenburgische Seenplatte

13072 LK Rostock

13073 LK Vorpommern-Rügen

13074 LK Nordwestmecklenburg

13075 LK Vorpommern-Greifswald

13076 LK Ludwigslust-Parchim

14511 SK Chemnitz

14521 LK Erzgebirgskreis

14522 LK Mittelsachsen

14523 LK Vogtlandkreis

14524 LK Zwickau

14612 SK Dresden

14625 LK Bautzen

- 14626 LK Görlitz
- 14627 LK Meißen
- 14628 LK Sächsische Schweiz-Osterzgebirge
- 14713 SK Leipzig
- 14729 LK Leipzig
- 14730 LK Nordsachsen
- 15001 SK Dessau-Roßlau
- 15002 SK Halle
- 15003 SK Magdeburg
- 15081 LK Altmarkkreis Salzwedel
- 15082 LK Anhalt-Bitterfeld
- 15083 LK Börde
- 15084 LK Burgenlandkreis
- 15085 LK Harz
- 15086 LK Jerichower Land
- 15087 LK Mansfeld-Südharz
- 15088 LK Saalekreis
- 15089 LK Salzlandkreis
- 15090 LK Stendal
- 15091 LK Wittenberg
- 16051 SK Erfurt

- 16052 SK Gera
- 16053 SK Jena
- 16054 SK Suhl
- 16055 SK Weimar
- 16056 SK Eisenach
- 16061 LK Eichsfeld
- 16062 LK Nordhausen
- 16063 LK Wartburgkreis
- 16064 LK Unstrut-Hainich-Kreis
- 16065 LK Kyffhäuserkreis
- 16066 LK Schmalkalden-Meiningen
- 16067 LK Gotha
- 16068 LK Sömmerda
- 16069 LK Hildburghausen
- 16070 LK Ilm-Kreis
- 16071 LK Weimarer Land
- 16072 LK Sonneberg
- 16073 LK Saalfeld-Rudolstadt
- 16074 LK Saale-Holzland-Kreis
- 16075 LK Saale-Orla-Kreis
- 16076 LK Greiz

16077 LK Altenburger Land

2000 SK Hamburg

3101 SK Braunschweig

3102 SK Salzgitter

3103 SK Wolfsburg

3151 LK Gifhorn

3153 LK Goslar

3154 LK Helmstedt

3155 LK Northeim

3157 LK Peine

3158 LK Wolfenbüttel

3159 LK Göttingen

3241 Region Hannover

3251 LK Diepholz

3252 LK Hameln-Pyrmont

3254 LK Hildesheim

3255 LK Holzminden

3256 LK Nienburg (Weser)

3257 LK Schaumburg

3351 LK Celle

3352 LK Cuxhaven

- 3353 LK Harburg
- 3354 LK Lüchow-Dannenberg
- 3355 LK Lüneburg
- 3356 LK Osterholz
- 3357 LK Rotenburg (Wümme)
- 3358 LK Heidekreis
- 3359 LK Stade
- 3360 LK Uelzen
- 3361 LK Verden
- 3401 SK Delmenhorst
- 3402 SK Emden
- 3403 SK Oldenburg
- 3404 SK Osnabrück
- 3405 SK Wilhelmshaven
- 3451 LK Ammerland
- 3452 LK Aurich
- 3453 LK Cloppenburg
- 3454 LK Emsland
- 3455 LK Friesland
- 3456 LK Grafschaft Bentheim
- 3457 LK Leer

3458 LK Oldenburg

3459 LK Osnabrück

3460 LK Vechta

3461 LK Wesermarsch

3462 LK Wittmund

4011 SK Bremen

4012 SK Bremerhaven

5111 SK Düsseldorf

5112 SK Duisburg

5113 SK Essen

5114 SK Krefeld

5116 SK Mönchengladbach

5117 SK Mülheim a.d.Ruhr

5119 SK Oberhausen

5120 SK Remscheid

5122 SK Solingen

5124 SK Wuppertal

5154 LK Kleve

5158 LK Mettmann

5162 LK Rhein-Kreis Neuss

5166 LK Viersen

5170 LK Wesel

5314 SK Bonn

5315 SK Köln

5316 SK Leverkusen

5334 StädteRegion Aachen

5358 LK Düren

5362 LK Rhein-Erft-Kreis

5366 LK Euskirchen

5370 LK Heinsberg

5374 LK Oberbergischer Kreis

5378 LK Rheinisch-Bergischer Kreis

5382 LK Rhein-Sieg-Kreis

5512 SK Bottrop

5513 SK Gelsenkirchen

5515 SK Münster

5554 LK Borken

5558 LK Coesfeld

5562 LK Recklinghausen

5566 LK Steinfurt

5570 LK Warendorf

5711 SK Bielefeld

5754 LK Gütersloh

5758 LK Herford

5762 LK Höxter

5766 LK Lippe

5770 LK Minden-Lübbecke

5774 LK Paderborn

5911 SK Bochum

5913 SK Dortmund

5914 SK Hagen

5915 SK Hamm

5916 SK Herne

5954 LK Ennepe-Ruhr-Kreis

5958 LK Hochsauerlandkreis

5962 LK Märkischer Kreis

5966 LK Olpe

5970 LK Siegen-Wittgenstein

5974 LK Soest

5978 LK Unna

6411 SK Darmstadt

6412 SK Frankfurt am Main

6413 SK Offenbach

6414 SK Wiesbaden

6431 LK Bergstraße

6432 LK Darmstadt-Dieburg

6433 LK Groß-Gerau

6434 LK Hochtaunuskreis

6435 LK Main-Kinzig-Kreis

6436 LK Main-Taunus-Kreis

6437 LK Odenwaldkreis

6438 LK Offenbach

6439 LK Rheingau-Taunus-Kreis

6440 LK Wetteraukreis

6531 LK Gießen

6532 LK Lahn-Dill-Kreis

6533 LK Limburg-Weilburg

6534 LK Marburg-Biedenkopf

6535 LK Vogelsbergkreis

6611 SK Kassel

6631 LK Fulda

6632 LK Hersfeld-Rotenburg

6633 LK Kassel

6634 LK Schwalm-Eder-Kreis

6635 LK Waldeck-Frankenberg

6636 LK Werra-Meißner-Kreis

7111 SK Koblenz

7131 LK Ahrweiler

7132 LK Altenkirchen

7133 LK Bad Kreuznach

7134 LK Birkenfeld

7135 LK Cochem-Zell

7137 LK Mayen-Koblenz

7138 LK Neuwied

7140 LK Rhein-Hunsrück-Kreis

7141 LK Rhein-Lahn-Kreis

7143 LK Westerwaldkreis

7211 SK Trier

7231 LK Bernkastel-Wittlich

7232 LK Bitburg-Prüm

7233 LK Vulkaneifel

7235 LK Trier-Saarburg

7311 SK Frankenthal

7312 SK Kaiserslautern

7313 SK Landau i.d.Pfalz

7314 SK Ludwigshafen

7315 SK Mainz

7316 SK Neustadt a.d.Weinstraße

7317 SK Pirmasens

7318 SK Speyer

7319 SK Worms

7320 SK Zweibrücken

7331 LK Alzey-Worms

7332 LK Bad Dürkheim

7333 LK Donnersbergkreis

7334 LK Germersheim

7335 LK Kaiserslautern

7336 LK Kusel

7337 LK Südliche Weinstraße

7338 LK Rhein-Pfalz-Kreis

7339 LK Mainz-Bingen

7340 LK Südwestpfalz

8111 SK Stuttgart

8115 LK Böblingen

8116 LK Esslingen

8117 LK Göppingen

8118 LK Ludwigsburg

8119 LK Rems-Murr-Kreis

8121 SK Heilbronn

8125 LK Heilbronn

8126 LK Hohenlohekreis

8127 LK Schwäbisch Hall

8128 LK Main-Tauber-Kreis

8135 LK Heidenheim

8136 LK Ostalbkreis

8211 SK Baden-Baden

8212 SK Karlsruhe

8215 LK Karlsruhe

8216 LK Rastatt

8221 SK Heidelberg

8222 SK Mannheim

8225 LK Neckar-Odenwald-Kreis

8226 LK Rhein-Neckar-Kreis

8231 SK Pforzheim

8235 LK Calw

8236 LK Enzkreis

8237 LK Freudenstadt

8311 SK Freiburg i.Breisgau

8315 LK Breisgau-Hochschwarzwald

8316 LK Emmendingen

8317 LK Ortenaukreis

8325 LK Rottweil

8326 LK Schwarzwald-Baar-Kreis

8327 LK Tuttlingen

8335 LK Konstanz

8336 LK Lörrach

8337 LK Waldshut

8415 LK Reutlingen

8416 LK Tübingen

8417 LK Zollernalbkreis

8421 SK Ulm

8425 LK Alb-Donau-Kreis

8426 LK Biberach

8435 LK Bodenseekreis

8436 LK Ravensburg

8437 LK Sigmaringen

9161 SK Ingolstadt

9162 SK München

9163 SK Rosenheim

9171 LK Altötting

9172 LK Berchtesgadener Land

9173 LK Bad Tölz-Wolfratshausen

9174 LK Dachau

9175 LK Ebersberg

9176 LK Eichstätt

9177 LK Erding

9178 LK Freising

9179 LK Fürstenfeldbruck

9180 LK Garmisch-Partenkirchen

9181 LK Landsberg a.Lech

9182 LK Miesbach

9183 LK Mühldorf a.Inn

9184 LK München

9185 LK Neuburg-Schrobenhausen

9186 LK Pfaffenhofen a.d.Ilm

9187 LK Rosenheim

9188 LK Starnberg

9189 LK Traunstein

9190 LK Weilheim-Schongau

9261 SK Landshut

9262 SK Passau

9263 SK Straubing

9271 LK Deggendorf

9272 LK Freyung-Grafenau

9273 LK Kelheim

9274 LK Landshut

9275 LK Passau

9276 LK Regen

9277 LK Rottal-Inn

9278 LK Straubing-Bogen

9279 LK Dingolfing-Landau

9361 SK Amberg

9362 SK Regensburg

9363 SK Weiden i.d.OPf.

9371 LK Amberg-Sulzbach

9372 LK Cham

9373 LK Neumarkt i.d.OPf.

9374 LK Neustadt a.d.Waldnaab

9375 LK Regensburg

9376 LK Schwandorf

9377 LK Tirschenreuth

9461 SK Bamberg

9462 SK Bayreuth

9463 SK Coburg

9464 SK Hof

9471 LK Bamberg

9472 LK Bayreuth

9473 LK Coburg

9474 LK Forchheim

9475 LK Hof

9476 LK Kronach

9477 LK Kulmbach

9478 LK Lichtenfels

9479 LK Wunsiedel i.Fichtelgebirge

9561 SK Ansbach

9562 SK Erlangen

9563 SK Fürth

9564 SK Nürnberg

9565 SK Schwabach

9571 LK Ansbach

9572 LK Erlangen-Höchstadt

- 9573 LK Fürth
- 9574 LK Nürnberger Land
- 9575 LK Neustadt a.d.Aisch-Bad Windsheim
- 9576 LK Roth
- 9577 LK Weißenburg-Gunzenhausen
- 9661 SK Aschaffenburg
- 9662 SK Schweinfurt
- 9663 SK Würzburg
- 9671 LK Aschaffenburg
- 9672 LK Bad Kissingen
- 9673 LK Rhön-Grabfeld
- 9674 LK Haßberge
- 9675 LK Kitzingen
- 9676 LK Miltenberg
- 9677 LK Main-Spessart
- 9678 LK Schweinfurt
- 9679 LK Würzburg
- 9761 SK Augsburg
- 9762 SK Kaufbeuren
- 9763 SK Kempten
- 9764 SK Memmingen

9771 LK Aichach-Friedberg

9772 LK Augsburg

9773 LK Dillingen a.d.Donau

9774 LK Günzburg

9775 LK Neu-Ulm

9776 LK Lindau

9777 LK Ostallgäu

9778 LK Unterallgäu

9779 LK Donau-Ries

9780 LK Oberallgäu

8.3 Die Regierungsbezirke sortiert nach ihrer Bevölkerungsdichte

13 Mecklenburg-Vorpommern

12 Brandenburg

15 Sachsen-Anhalt

72 Trier

33 Lüneburg

93 Oberpfalz

92 Niederbayern

16 Thüringen

- 66 Kassel
94 Oberfranken
96 Unterfranken
34 Weser-Ems
71 Koblenz
1 Schleswig-Holstein
97 Schwaben
65 Gießen
31 Braunschweig
146 Dresden
84 Tübingen
145 Chemnitz
32 Hannover
83 Freiburg
95 Mittelfranken
147 Leipzig
91 Oberbayern
73 Rheinhessen-Pfalz
57 Detmold
55 Münster
10 Saarland

81 Stuttgart

82 Karlsruhe

59 Arnsberg

64 Darmstadt

53 Köln

51 Düsseldorf

4 Bremen

2 Hamburg

11 Berlin

8.4 Die Regierungsbezirke lexikographisch sortiert nach den ersten beiden Teilen der Gemeindeschlüssel ihrer Landkreise

1 Schleswig-Holstein

10 Saarland

11 Berlin

12 Brandenburg

13 Mecklenburg-Vorpommern

145 Chemnitz

146 Dresden

147 Leipzig

15 Sachsen-Anhalt

16 Thüringen

2 Hamburg

31 Braunschweig

32 Hannover

33 Lüneburg

34 Weser-Ems

4 Bremen

51 Düsseldorf

53 Köln

55 Münster

57 Detmold

59 Arnsberg

64 Darmstadt

65 Gießen

66 Kassel

71 Koblenz

72 Trier

73 Rheinhessen-Pfalz

81 Stuttgart

82 Karlsruhe

83 Freiburg

84 Tübingen

91 Oberbayern

92 Niederbayern

93 Oberpfalz

94 Oberfranken

95 Mittelfranken

96 Unterfranken

97 Schwaben

ToDo Counters

To Dos: 28; 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28

Parts to extend: 0;

Draft parts: 0;

Literatur

- [1] Auge=mit. *Spanische Grippe 1918 1919 GB*. [Online; abgerufen am 07.06.2021]. 2021. URL: https://commons.wikimedia.org/wiki/File:Spanische_Grippe_1918_1919_GB.svg.
- [2] Ronald N. Bracewell. „The Fourier Transform and Its Applications“. In: Singapore: McGraw-Hill Book Co, 2000, S. 24–54.
- [3] Alfred W. Crosby. „Conquistador y Pestilencia: The First New World Pandemic and the Fall of the Great Indian Empires“. In: *Hispanic American Historical Review* 47. Duke University Press, Aug. 1967, S. 321–337. DOI: 10.1215/00182168-47.3.321.
- [4] F.R. DeLeo und B.J. Hinnebusch. „A plague upon the phagocytes“. In: *Nature Medicine* 11.9. 9 (2005), S. 927–928.
- [5] Robert Littman und M Littman. „Galen and the Antonine Plague“. In: *American journal of philology* 94 (Feb. 1973), S. 243–255. DOI: 10.2307/293979.
- [6] M. Kroneman P. Spreeuwenberg und J. Paget. „Reassessing the Global Mortality Burden of the 1918 Influenza Pandemic“. In: *American Journal of Epidemiology* 187.12 (2018), S. 2561–2567.
- [7] George Santayana. *The Life of Reason: Introduction and Reason in Common Sense*. 1905.
- [8] F.S. Lobo T. Harko und M. Mak. „Exact analytical solutions of the Susceptible-Infected-Recovered (SIR) epidemic model and of the SIR model with equal

death and birth rates“. In: *Applied Mathematics and Computation* 236 (2014), S. 184–194.

