Diversität von Nachnamen als Maß für soziale Offenheit – Was sagt uns die Verteilung von Nachnamen in einer deutschen Großstadt?

Manuel Diaz Garcia - Jonas Elis · 19. Februar 2024
Über Politik aus der Wissenschaft

Manuel Diaz Garcia und Jonas Elis

Lässt sich an unseren Nachnamen erkennen, ob wir in einer sozial offenen Region leben? Also in einer Gesellschaft, die Außenstehenden relativ einfachen Zugang zur Gesellschaft gewährt? Die Antwort lautet: Ja! Zumindest laut einer Studie von Paolo Buonanno und Paolo Vanin. Sie gingen dem Zusammenhang von sozialer Offenheit und Kriminalitätsraten sowie Steuerhinterziehung in italienischen Gemeinden nach. Soziale Offenheit wurde dabei ermittelt durch die Unterschiedlichkeit von Nachnamen in den jeweiligen Gemeinden. Aber warum? Und wie gut funktioniert dieser Ansatz, wenn wir uns dieselbe Methode zu eigen machen und versuchen, die soziale Offenheit in einer deutschen Großstadt zu messen?

Was uns Nachnamen wirklich erzählen

Um zu erklären, warum genau Nachnamen eine relevante Information über soziale Offenheit sind muss man verstehen, wie Nachnamen funktionieren. In vielen Gesellschaften, so wie Italien, ist es üblich, dass Nachnamen vom Vater an die Kinder weitergegeben werden. Gepaart mit dem Fakt, dass sozial geschlossene Gesellschaften eine geringere Wahrscheinlichkeit sowohl für Migration in diese Gesellschaften als auch für eine Heirat über die Gesellschaften hinweg haben, kommt es zu einer interessanten Entwicklung. Die Verteilung von Nachnamen wird mit der Zeit diverser, wenn neue

Männer in eine Region ziehen. Und sie wird homogener durch zwei weitere Prozesse. Zum einen, wenn Männer die Region verlassen und wegziehen. Zum anderen, wenn Männer Nachkommen mit Frauen aus derselben Region bekommen. Dies ist der Fall da die Wahrscheinlichkeit für einen männlichen Nachkommen geringer ist, der den Nachnamen bei einer zukünftigen Heirat behalten würde. Dabei muss es sich nicht zwingend um transnationale Migrationsbewegungen handeln, auch Binnenmigration spielt hierbei eine entscheidende Rolle. Damit geht die Unterschiedlichkeit von Nachnamen über übliche Kontextmerkmale wie Ausländeranteile hinaus.

Die Idee hinter der Messung von sozialer Offenheit über die Verteilung von Nachnamen in Gemeinden ist also möglicherweise ein guter Indikator für (a) die Geschichte von Migration und (b) die Muster von sozialen Interaktionen in der Gemeinde. Aber wie genau lässt sich die Diversität von Nachnamen in einer Region bestimmen?

Aus dem Tierreich in die Sozialwissenschaften

Buonanno und Vanin machen sich den Shannon Diversity Index zunutze. Ursprünglich wurde dieser genutzt, um die Diversität von Tierarten in unterschiedlichen Habitaten oder über die Zeit hinweg zu bestimmen. Dabei berücksichtigt der Index sowohl die Anzahl der Lebewesen in einem Habitat als auch die Anzahl der unterschiedlichen Tierarten. Der Index nimmt den Wert 0 an, wenn alle Tiere in einem Habitat derselben Spezies angehören. Maximale Diversität ist dann erreicht, wenn die Anzahl der Tiere gleich der Anzahl der unterschiedlichen Tierarten ist. Für die mathematisch Interessierten sieht das Ganze wie folgt aus:

$$H = -\sum p_i * ln\,(p_i)$$

Der Index (H) berechnet den Anteil jeder einzelnen Spezies an der Gesamtzahl der Lebewesen im Habitat (p_i), multipliziert dies mit dem natürlichen Logarithmus desselben

Anteils und nimmt die negative Summe aus den Werten jeder Spezies.

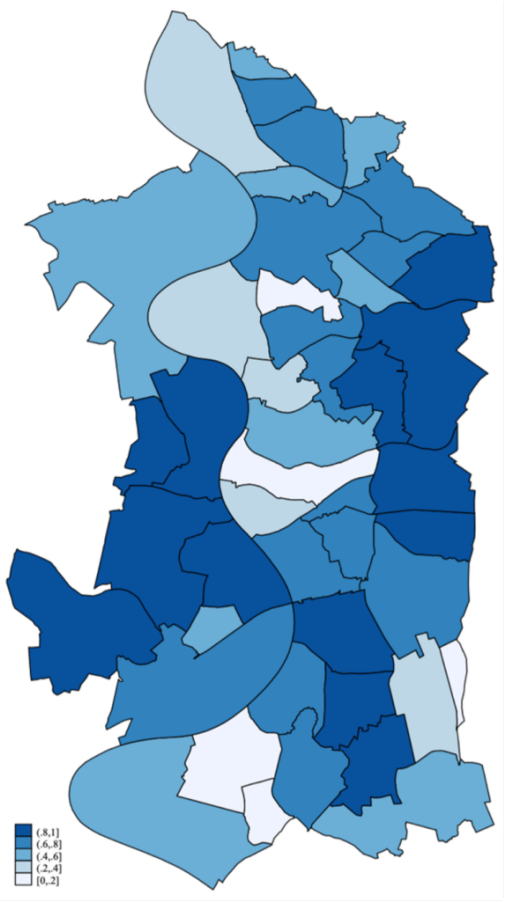
Ersetzt man Tiere in einem Habitat jetzt durch Menschen in einer Gemeinde und die verschiedenen Spezies durch verschiedene Nachnamen lässt sich auf dieselbe Art die Diversität von Nachnamen bestimmen. Eine Maßzahl für die soziale Offenheit ist entstanden und lässt sich so eben auch in den Sozialwissenschaften anwenden. Zur Reproduktion unserer Berechnungen des Shannon Diversity Index gibt es hier ein R-Skript.

Die Diversität von Nachnamen in Duisburg

Um sich die Diversität von Nachnamen als Variable zunutze machen zu können benötigt man zunächst eine Liste von Nachnamen. Diese lag während der Stichprobenziehung der Immigrant German Election Study II vor. Es handelt sich um eine einfache Zufallsstichprobe, die ausreichend groß ist um zwischen Stadtteilen differenzieren zu können.

Auf der Karte betrachtet, zeigen sich einige interessante Muster über das Duisburger Stadtgebiet hinweg. Erstens sehen wir hohe Werte der Namensdiversität in mehreren Stadtteilen links des Rheins. Eine Erklärung dafür könnte die Eingemeindung unter anderem von Rheinhausen, Homberg und Baerl nach Duisburg 1975 sein. Ähnlich sieht es in den östlichen Stadtteilen um Neudorf-Nord, Duissern bis in den Norden nach Neumühl aus. Ob dafür der Standort der Universität (Neudorf-Nord) oder der Zuzug Russlanddeutscher Zuwanderer*Innen in die Siedlung Hagenshof (Neumühl) ausschlaggebend ist, lässt sich nur vermuten. In all diesen Fällen wären sehr unterschiedliche historische Ereignisse der Grund für die hohe Varianz der Namensdiversität und das Maß würde eine Dynamik erfassen, die andernfalls unsichtbar wäre.

Darstellung 1: Verteilung von Nachnamen nach Shannon Diversity Index in 46 Stadtteilen der Stadt Duisburg



Da der Shannon Diversity Index die Diversität nicht in einer inhaltlichen Einheit wie Anteilen angibt, haben wir die errechneten Werte standardisiert, sodass sie zwischen 0 und 1 variieren. Den Wert 0 bekommt dabei der Stadtteil mit der geringsten Diversität und den Wert 1 bekommt der Stadtteil mit der höchsten Diversität. Dies ist möglich, da wir lediglich an der Varianz zwischen den Stadtteilen und nicht den absoluten Werten interessiert sind. Für die exakten Werte nach Stadtteilen, klicke hier.

Anwendbarkeit in sozialwissenschaftlichen Studien

Die Diversität von Namen drückt soziale Offenheit aufgrund von Zuzügen oder Weggang von Menschen in ein Gebiet aus. Der Index ist damit etwas anderes als Kontextvariablen wie beispielsweise dem Anteil der Ausländer*Innen pro Stadtteil. Zwar gibt es einen positiven Zusammenhang zwischen den beiden, dieser ist allerdings nicht sonderlich stark (Pearson’s Korrelationskoeffizient: 0,104)

Während die Berechnung mithilfe statistischer Analysesoftware anhand der oben gegebenen Formel recht einfach ist, besteht jedoch eine Hürde: Es braucht einen

möglichst unverzerrten Datensatz mit vielen Personennamen. In der Regel stehen diese nur über Melderegisterauszüge zur Verfügung.

Je nach lokalem Kontext liegt auch die Vermutung nahe, dass die Namensdiversität stellenweise überschätzt wird, wenn sehr ähnliche oder Doppelnamen nicht gekennzeichnet werden. Beispielsweise würden die Nachnamen “Müller” und “Müller-Schmidt” sowie “Mueller” in einem Datensatz zu einer Vergrößerung der Diversität führen, obwohl sie sich stark ähneln. Ein Faktor der Ähnlichkeit von Substrings kann hier zu einer weiteren Verbesserung der Messung führen. Ein Python-Skript zur Anwendung dieser Verfeinerung stellen wir hier zur Verfügung.

Sofern eine entsprechende Datengrundlage zur Verfügung steht, kann Namensdiversität also ein aufschlussreiches Maß für die Beantwortung sozialwissenschaftlicher Fragestellungen sein, welche Dynamiken von Bevölkerungsentwicklung betreffen, die sonst nicht betrachtet werden können. Beispielsweise können Gebiete mit hoher relativer Namensdiversität für Parteien interessant sein, weil Parteipräferenzen in der Bevölkerung gegebenenfalls weniger gefestigt sind.

Literatur

Paolo Buonanno und Paolo Vanin (2017): Social closure, surnames and crime In Journal of Economic Behavior & Organization (137), S. 160-175. DOI: https://doi.org/10.1016/j.jebo.2017.03.002

Achim Goerres, Dennis C. Spies, Sabrina S. Mayer, Jonas Elis, Manuel Diaz Garcia und Phillip Hoffmann (2016-2021): Immigrant German Election Study I & II. https://achimgoerres.de/imges

---

Zitation:

Manuel Diaz Garcia und Jonas Elis: Diversität von Nachnamen als Maß für soziale Offenheit – Was sagt uns die Verteilung von Nachnamen in einer deutschen Großstadt?, Erschienen in „Über Politik aus der Wissenschaft“, Herausgeber Achim Goerres, 10. April 2025, abrufbar unter: https://www.politik-wissenschaft.org/2024/02/19/diversitat-

nachnamen/, DOI: https://doi.org/10.17185/politik-wissenschaft/20250410-2024-03

Manuel Diaz Garcia - Jonas Elis (2024): Diversität von Nachnamen als Maß für soziale Offenheit – Was sagt uns die Verteilung von Nachnamen in einer deutschen Großstadt? . Empirische Politik UDE.
https://www.politik-wissenschaft.org/2024/02/19/diversitat-nachnamen/
Lesefassung aus dem Onlinebeitrag. Interaktive Inhalte sind in der Onlinefassung verfügbar.