

Executive Summary – World Bank Datenanalyse

Diese Executive Summary bietet einen Überblick über die Datenanalyse der offen zugänglichen Weltentwicklungsdatenbank der World Bank Group. Der für diese Analyse zur Verfügung stehende Auszug beinhaltet gesellschaftliche und sozioökonomische Daten im Rahmen von jeweils 18 Indikatoren für 25 Länder der Welt und umfasst eine jährliche Datenerhebung von 2000 bis 2021.

Fragestellungen

Das Ziel des Projekts ist, Zusammenhänge zwischen den folgenden Indikatoren zu untersuchen: (i) die HIV-Prävalenz, der Alkoholkonsum pro Kopf und welche Bedeutung dem Anteil der Erwerbsbevölkerung mit Grundbildung auf die HIV-Prävalenz zufällt, (ii) der Anteil der Erwerbsbevölkerung mit Grundbildung, die Zentralstaatsverschuldung und die Schüler-Lehrer-Relation, (iii) das BIP pro Kopf und die Tabakkonsum-Prävalenz sowie (iv) der Anteil an landwirtschaftlicher Nutzfläche und die CO₂ Emissionen pro Kopf und welche Rolle die Landesfläche in dieser Beziehung spielt.

Problematik und Methodik

Die Längsschnittdaten sehen sich mit Herausforderungen wie Datenlücken konfrontiert. Außerdem liegen für die CO₂ Emissionen bei rund zehn Prozent der vorliegenden Daten fehlerhafte Werte vor, die um das Zehn- oder Hundertfache niedriger ausfallen. Nach Rücksprache wurden diese zwar in die Analyse mitaufgenommen, allerdings als Fehler berücksichtigt. Die Zusammenhänge zwischen den Indikatoren werden mit einer Vielzahl von Visualisierungen (z.B. Säulen-, Linien-, Punkt- und Boxdiagramme) ausgewertet. Aufgrund der vielfältigen Eigenschaften und begrenzten Vergleichbarkeit der Länder wird im Laufe der Analyse innerhalb der Fragestellungen zudem eine Einteilung der Länder in Quantile nach thematisch gewählten Indikatoren unternommen. Abhängig von der verfügbaren Datenmenge werden die Länder in drei oder fünf Kategorien eingeteilt, um aussagekräftigere Datengruppen zu generieren.

Ergebnisse

In der ungruppierten Betrachtung der Beziehungen zwischen den Indikatoren konnten jeweils tendenziell positive Zusammenhänge im schwachen bis mittelstarken Intensitätsbereich festgestellt werden. Durch die Gruppierung haben sich folgende zusätzliche Anmerkungen ergeben. Zu (i): Die HIV-Werte für Länder mit sehr großem relativem Alkoholkonsum pro Kopf weisen viel größere Streuung auf, aber nicht unbedingt höhere HIV-Prävalenzen. Zu (ii): Der Zusammenhang zwischen Zentralstaatsverschuldung und Grundbildung variiert mit Einkommen und Bevölkerungsgröße. In dicht besiedelten Ländern bewirkt eine niedrigere Schüler-Lehrer-Relation eine verbreitete Grundbildung und ein höheres Bildungsniveau – ein Effekt, der in dünn besiedelten Ländern nicht auftritt. Zu (iii): Länder mit sehr geringem relativem Tabakkonsum weisen ein sehr niedriges BIP pro Kopf auf, während sich für die anderen Kategorien ein negativer Zusammenhang zeigt. Zu (iv): Einzig bei Betrachtung der relativen Änderungen über den untersuchten Zeitraum haben sehr große Länder eine gegenläufige Entwicklung in CO₂ Emissionen pro Kopf im Vergleich zum Anteil an landwirtschaftlicher Nutzfläche.

Ausblick

Methodische Herausforderungen in der Analyse umfassen Datenlücken in den Längsschnittdaten sowie die Vergleichbarkeit der inkongruenten Länder. Die Ergebnisse betonen die Wichtigkeit von präziseren Datenerfassungsmethoden, um dem geringwertigen Datenaufkommen entgegenzuwirken. Zudem werden für einen zukünftigen Rückschluss aus der Stichprobe auf die restlichen Länder der Welt repräsentativere Untersuchungseinheiten für die weltweite Lage benötigt, um außerhalb eines vakuumierten Umfelds schlussfolgern zu können. Informationen zu weiteren Indikatoren können dabei helfen, gewonnene Erkenntnisse zu bestätigen oder zu widerlegen.