**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 pro Land 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 Erhebungsschwierigkeiten und daraus resultierenden Datenlücken konfrontiert. Außerdem liegen für die CO2 Emissionen bei rund zehn Prozent der vorliegenden Daten fehlerhafte Werte vor, die um das Zehnfache oder Hundertfache niedriger ausfallen als in der Realität. Nach entsprechender Rücksprache wurden diese zwar in die Analyse mitaufgenommen, allerdings als Fehler berücksichtigt und damit die Aussagekraft auf die wahrhaftigen Daten beschränkt. Die Zusammenhänge zwischen den einzelnen Indikatoren werden mit einer Vielzahl von unterschiedlichen Visualisierungen (z.B. Punkt-, Linien-, Säulen- und Boxdiagramme) ausgewertet. Aufgrund der mannigfaltigen Charakteristika und geringen Vergleichbarkeit der untersuchten Länder wird im Laufe der Analyse innerhalb der Fragestellungen zudem eine Kategorisierung der Länder in äquidistante Quantile nach thematischen gewählten Indikatoren unternommen, um aussagekräftigere Datengruppen zu generieren - je nach vorliegender Datenmenge in drei oder fünf Gruppen.

**Ergebnisse**

[VGL. BEISPIELABGABE UND FÜGE ERGEBNISSE EIN]

**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. Auch Informationen zu weiteren Indikatoren können dabei helfen, gewonnene Erkenntnisse zu bestätigen oder zu widerlegen.