

# Faktoren für Lebenserwartung

Hannes Schwanzer

Juli 2022

## Inhaltsverzeichnis

|                                                            |   |
|------------------------------------------------------------|---|
| 1 Herkunft der Daten                                       | 2 |
| 2 Beschreibung des Datensatzes                             | 2 |
| 3 Thesen                                                   | 2 |
| 4 Benutzte Tools                                           | 3 |
| 5 Bereinigung der Daten                                    | 3 |
| 6 Auswertung                                               | 4 |
| 6.1 Korrelationen                                          | 4 |
| 6.2 Vergleich von Industriestaaten und Entwicklungsländern | 5 |
| 7 Fazit                                                    | 5 |

# 1 Herkunft der Daten

Die Daten wurden von der Global Health Observatory (GHO) bereitgestellt, welche der World Health Organisation (WHO) angehören. Diese beinhalten Informationen über Lebenserwartung und gesundheitliche Faktoren, wie z.B. Impfungen gegen Viren, Gesundheitsausgaben, etc. gruppiert nach Jahr und Land.

Die Daten wurden für die Jahre 2000 bis 2015 auf der Seite Kaggle vom Nutzer Kumar Rajarshi zusammengetragen.

## 2 Beschreibung des Datensatzes

Der Datensatz umfasst 2938 Daten mit 22 Attributen.

Die Attribute sind:

- Country: Von welchem Land die Daten stammen
- Year: Aus welchem Jahr die Daten sind
- Status: Ob das Land ein Entwicklungsland ist oder nicht
- Life expectancy: Die Lebenserwartung in Jahren
- Adult Mortality: Die Sterblichkeit von Erwachsenen (15-60 Jahre) pro 1000 Einwohner
- Infant deaths: Anzahl der Säuglings Sterbefälle pro 1000 Einwohner
- Under-five deaths: Anzahl der Sterbefälle von Kinder unter 5 Jahren pro 1000 Einwohner
- Percentage expenditure: Gesundheitsausgaben in Prozent des Bruttoinlandsprodukt pro Kopf
- Total expenditure: Staatliche Gesundheitsausgaben in Prozent der gesamten Staatsausgaben
- GDP: Bruttoinlandsprodukt pro Kopf in USD
- Schooling: Anzahl der Schuljahre
- Population: Bevölkerungszahl

Es gibt noch weitere Attribute wie Alkoholkonsum, Impfschutz gegen spezifische Krankheiten, sowie Untergewicht bei Kinder und Body Mass Index (BMI), welche hier nicht betrachtet werden.

Im Folgenden werden die Attribute im Zusammenhang mit der Lebenserwartung analysiert.

## 3 Thesen

Aufgrund des medizinischen Fortschritts, sowie steigender Lebensqualität steigt die Lebenserwartung im Laufe der Jahre an.

In Entwicklungsländern wird die Lebenserwartung pro Jahr wahrscheinlich linear ansteigen. In Industriestaaten hingegen wird die Lebenserwartung zwar steigen, aber nicht so stark im Vergleich zu Entwicklungsländern und eventuell sogar abflachen, aufgrund steigender Fettleibigkeit und wenig Bewegung.

Zudem wird die Lebenserwartung mit dem Bruttoinlandsprodukt stark zusammenhängen, da Geld ein wesentlicher Faktor für medizinische Versorgung und ausreichend Nahrung ist.

## 4 Benutzte Tools

Die komplette Datenverarbeitung wurde mit Python gemacht.

Dazu wurden die Bibliotheken pandas und numpy zur Datenverarbeitung und matplotlib und seaborn für die Datenvisualisierung benutzt.

## 5 Bereinigung der Daten

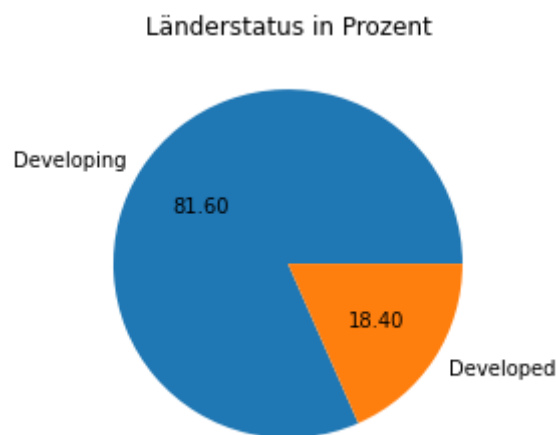
Die nicht genutzten Attribute werden verworfen um Probleme mit leeren Feldern vorzubeugen.

Von Zeichenketten wurden doppelte Leerzeichen, sowie Leerzeichen am Anfang und am Ende entfernt. Die Attributnamen wurden zu Kleinbuchstaben verändert und mit einem Unterstrich konkateniert, um die Verarbeitung der Daten zu erleichtern.

Zudem wurden doppelte Daten, sowie Daten mit leeren Feldern gelöscht, da man diese mit keinen sinnvollen Werten ersetzen konnte und Fehler verursachen können.

Dabei sind vor allem die Attribute GDP und Population herausgestochen, da dort sehr viele Werte gefehlt haben.

Insgesamt wurden 818 Daten gelöscht, welche 27% des Datensatzes ausmachen.

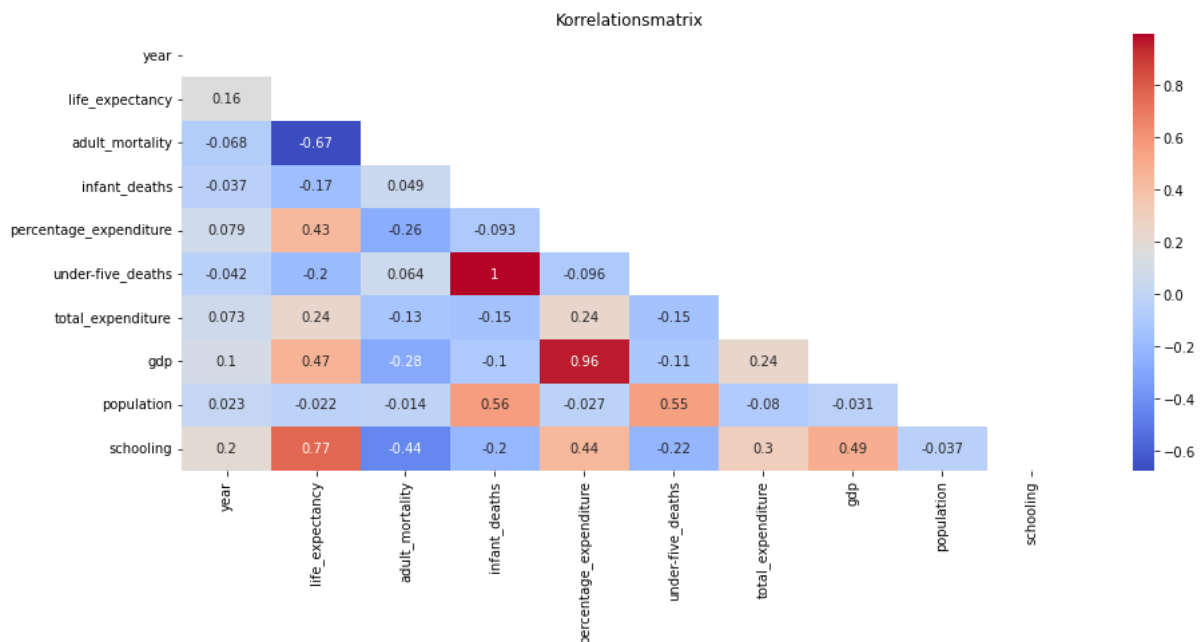


Nach dem Löschen der Daten machen die Entwicklungsländern 81,60% der Daten aus und die Industriestaaten 18,40%.

## 6 Auswertung

### 6.1 Korrelationen

Zunächst wird die Korrelationen der verschiedenen Attribute betrachtet.



Die Lebenserwartung hängt stark mit dem Bruttoinlandsprodukt, der prozentualen Gesundheitsausgaben und der Schulbildung zusammen.

Dazu sieht man das die Lebenserwartung stark negativ mit der Erwachsenen Sterblichkeit korreliert ist, was darauf zurückzuführen ist, dass ein großer Teil der Bevölkerung schon im Erwachsenenalter stirbt und somit die Lebenserwartung im Land insgesamt sinken lässt.

Es ist zudem zu sehen, dass die Lebenserwartung kaum mit der Bevölkerungszahl zusammenhängt.

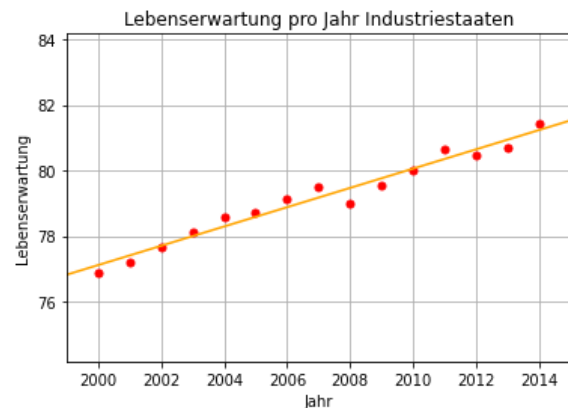
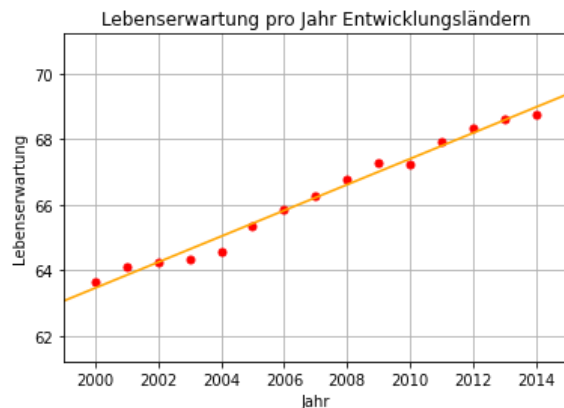
Weitere herausstechende Werte sind:

- Kindersterblichkeit und Säuglingssterblichkeit sind stark korreliert
- Prozentuale Gesundheitsausgaben und das Bruttoinlandsprodukt sind stark korreliert

## 6.2 Vergleich von Industriestaaten und Entwicklungsländern

In den folgenden Diagrammen wurde mit Hilfe der Least Squares Methode eine Ausgleichsgerade berechnet.

Da nur 2 Werte in dem Jahr 2015 vorhanden sind, wird dieses Jahr nicht mit betrachtet, da es nicht aussagekräftig ist und die Ausgleichsgerade beeinflusst, wodurch das Ergebnis verfälscht wird.



In den Diagrammen kann man erkennen, dass die Lebenserwartung fast linear ist, mit einer Steigung von ca. 0,39.

Die Lebenserwartung von Industriestaaten steigt auch linear, mit einer Steigung von ca. 0,29. Anders als erwartet flacht diese auch nicht ab, sondern steigt fast stetig an.

Im Vergleich kann man erkennen, dass die Lebenserwartung in Entwicklungsländern stärker ansteigt.

## 7 Fazit

Generell stimmen die Thesen mit den Ergebnissen überein, außer dass die Lebenserwartung in Industriestaaten immer noch stetig ansteigt.

Bei der Auswertung ist jedoch zu beachten, dass ein Teil der Daten, aufgrund von fehlenden Werten nicht beachtet wurde, wodurch einige Länder überhaupt nicht betrachtet werden.

Zudem umfasst der Datensatz der WHO nicht alle Länder der Welt, wodurch die Ergebnisse verfälscht werden können.

## 8 Quellen

Datensatz:

<https://www.kaggle.com/datasets/kumarajarshi/life-expectancy-who>