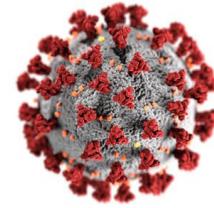


Corona und das exponentielle Wachstum Eine Analyse mit R



In diesem Aufgabenblatt soll es um die Analyse der weltweiten Corona Fallzahlen und deren Vorhersage für die Zukunft gehen, falls keine eindämmende Maßnahmen getroffen werden.

Am 31. März 2020 wurden in Deutschland seit Ende Januar insgesamt ~62'000 Covid19 Erkrankte gemeldet, in Italien ~102'000 (Quelle: WHO: <https://covid19.who.int/table>).

Fragen die sich hieraus z.B. ergeben:

- Welches Wachstumsmodell liegt vor?
- Können basierend auf diesem Modell Vorhersagen gemacht werden?
- Wie unterschiedlich verlaufen die Fallzahlen in verschiedenen Ländern?
- Kann man aus den Daten z.B. schließen, dass der Verlauf in Italien grundsätzlich problematischer ist als in Deutschland?
- Wie verhält es sich mit den Todeszahlen?
- ...

1. Aufgabe (Datenaufbereitung):

- a) Download der aktuellen Corona Falldaten von der WHO (<https://covid19.who.int/table>)
- b) ggf. Datensatz in UTF-8 umkodieren, Sonderzeichen entfernen... (mit Texteditor, z.B. Notepad⁺⁺ für Windows)
- c) Erstellen Sie ein neues Skript "Covid.r" und installieren Sie folgendes Package:

```
install.packages("data.table")  
library(data.table)
```
- d) Lesen Sie die Daten mit `read.csv` in R ein (Name: `COVID`) und wandeln diese in eine `data.table` Tabelle um.
- e) Prüfen Sie die Datenbasis (insbesondere die Datumsspalte) mit: `str(COVID)`
- f) Überführen Sie das Datumsformat der Datumsspalte in das Datumsformat von R
- g) Prüfen Sie die Datenbasis erneut mit: `str(COVID)` und `head(COVID)`
- h) Land auswählen und die Datensätze BRD und ITA erstellen
(`Country_code=="DE", "IT"`)
- i) Merkmal für die kummulierte Anzahl von Infizierten auswählen und Datensätze `BRD_Infektionen` und `ITA_Infektionen` erstellen
- j) Monat März auswählen und Datensatz `BRD_Infektionen_Maerz` und `ITA_Infektionen_Maerz` erstellen
- k) Prüfen Sie die Datensätze wieder z.B. mit:

```
str(BRD_Infektionen_Maerz)  
head(BRD_Infektionen_Maerz)  
View(BRD_Infektionen_Maerz)
```

2. Aufgabe (Datenanalyse):

- a) Erstellen Sie für Deutschland im März (`BRD_Infektionen_Maerz` und `ITA_Infektionen_Maerz`) ein einfaches Punktdiagramm mit dem Befehl: `plot`.

Welcher allgemeinen Funktion folgen die Fallzahlen rein visuell gesehen am ehesten, vor allem in der ersten Märzhälfte?

Was kann man bezüglich den Infiziertenzahlen nach der TV-Ansprache von Angela Merkel mit Lockdown am 19.3.2020 erkennen?

Erstellen Sie die Diagramme ebenfalls für `BRD_Infektionen` und `ITA_Infektionen`.

- b) Wie korrelieren die beiden Datensätze im März und insgesamt?
Stellen Sie dazu die Infiziertenzahlen von Deutschland und Italien gegenüber und bestimmen Sie die Korrelationskoeffizienten nach Pearson mit dem Befehl: `cor.test`.

3. Aufgabe (Modellbildung):

Um das genaue Modelle zu erhalten, stellen wir ein Exponential-Modell (decay-model) nach folgender Gleichung auf:

$$N(t) = N_0 a^t \quad \text{oder gleichwertig}$$

$$N(t) = N_0 e^{\lambda t} \quad \text{mit } \lambda = \ln(a)$$

Einfaches Modell mit Schätzung von a (Wahrscheinlichkeitsrate) bzw. λ (Wahrscheinlichkeitskonstante) mit zwei Werten:

- a) Bestimmen Sie allgemein den Parameter a und λ .
- b) Bestimmen Sie den Wert für a , indem Sie zwei Werte aus dem Datensatz (`BRD_Infektionen_Maerz`) einsetzen. Nehmen Sie den Startzeitpunkt 2020-03-01 (Tag 0) an und den zweiten Wert am Tag 2020-03-19 (Tag 18).
- c) Erstellen Sie ein Schaubild mit dem Modell für die BRD im Monat März und vergleichen Sie dies mit den Originaldaten im selben Plotfenster.
- d) Erstellen Sie abschließend noch ein Blockdiagramm mit: `barplot`.