

## Mini Challenge: Steuerdaten einer Gemeinde im Kanton Solothurn

### *spz – Stochastische Prozesse und Zeitreihen*

Ihr löst Probleme aus Eurem eigenen Projekt unter Zuhilfenahme wesentlicher Werkzeuge, die Ihr in *spz* kennengelernt habt. Ihr seid bei den Fragestellungen frei, sie sollen aber sinnvoll zur Projektdomäne passen. Eure Antworten sollen mit Daten hinterlegt sein. Nach Möglichkeit soll die statistische Signifikanz der Ergebnisse abgeschätzt werden. Die zu lösenden Fragen sollen explizit gemacht werden. Aus der Datenexploration dürfen sich auch weitere Fragestellungen ergeben.

Spezifisch zu den Steuerdaten einer Gemeinde im Kanton Solothurn beantwortet Ihr folgende Fragen:

- Erstellt eine Prognose der Steuererträge absolut und pro Person für die nächsten 10 Jahre unter Berücksichtigung allfälliger zyklischer Schwankungen aufgrund der historischen Steuererträge, ggf. in einer geeigneten Partition der Bevölkerung
- Bestimmt, wie stark BIP- und Demographiezahlen als vorlaufende Indikatoren verwendbar sind, wie viele Jahre Vorlauf sie Euch gewähren und
- Berechnet, welches Steuerloch im schlechtesten Fall und welcher Überschuss im besten Fall alle 50 Jahre bei den natürlichen Personen zu erwarten sind.

#### Vorgaben zu den Methoden:

- **Zeitreihen:** Ihr charakterisiert die Zeitreihen und zerlegt sie in relevante Komponenten. Ihr charakterisiert die Komponenten bezüglich Stationarität und anderer Zeitreiheneigenschaften. Ihr minimiert oder modelliert die Residuen und bestimmt relevante Auto- und Kreuz-Korrelationsstrukturen.
- **Stochastische Prozesse:** Ihr modelliert die passenden Komponenten als stochastische Prozesse, macht die Zusammenhänge transparent und überprüft Eure Prognosen mittels Back Test.
- **Extremwerte:** Ihr definiert für die Problemstellung relevante extreme und/oder seltene Ereignisse, z.B. weil sie stören und daher besondere Aufmerksamkeit brauchen. Ihr begründet die verwendeten Verteilungen sowohl statistisch (im Rahmen des Möglichen) als auch von der Struktur des datengenerierenden Prozesses her.
- **Interpretationen:** Ihr interpretiert Eure Ergebnisse und hinterfragt sie.
- **Simulationen** (optional): Ihr simuliert mögliche Realisierungen Eurer Modelle, um die Ergebnisse zu illustrieren und einzuordnen.

#### Abgaben:

- Eure Datenanalyse und Eure Modelle werden in einem R-Studio-Notebook (R) oder einem Jupyter-Notebook (Python) erstellt und nachvollziehbar dokumentiert (s. unten).

- Ihr bereitet einen 10-minütigen Vortrag vor oder ein kurzes Video (max. 10 Minuten) vor. Dabei werden die untenstehenden Aspekte Eurer Arbeit erläutert. Der Vortrag wird aufgezeichnet.
- Bonus-Fragestellungen und weitere Analysen sind natürlich willkommen.

Im Notebook und im Video / Vortrag wird eine **nachvollziehbare Dokumentation** zu folgenden Teilen Eurer Arbeit erwartet:

- **Fragestellungen:** Welche Fragestellungen beantwortet Ihr?
- **Zeitreihen:** Woher kommen die entsprechenden Zeitreihen? Wie habt Ihr sie bearbeitet?
- **Modelle:** Als welche stochastischen Prozesse modelliert Ihr die Zeitreihen, mit welchen Methoden parametrisiert Ihr sie, und was sind Eure Begründungen?
- **Vorhersage:** Welche Annahmen habt Ihr gemacht? Helfen Szenarien? Was sind die Grenzen der Vorhersagbarkeit?
- **Extremwertanalyse:** Über welche extremen und/oder seltenen Ereignisse soll eine Aussage gemacht werden?
- **Fazit:** Welche Empfehlungen gebt Ihr dem Kunden mit?
- **Ausblick:** Sehr zu empfehlen ist ein Ausblick, nicht zuletzt, um Euren Aufwand zu begrenzen.

### *Mögliche Themen aus anderen Bereichen*

Euer Beratungsbüro bearbeitet auch analoge Fragestellungen für andere Kunden, die in anderen Gebieten tätig sind. Wählt eine alternative Anwendung aus und formuliert die entsprechende Aufgabenstellung.

Eure Kunden haben Analysebedarf in den folgenden Gebieten:

- Betriebswirtschaft: Pufferlager für «just-in-time» Produktion
- Naturwissenschaft-Technik: Ausfall von Infrastruktur aufgrund starker Sonnenaktivität
- Volkswirtschaft: Konjunkturvorhersage mit Handlungsempfehlung
- Biomedizin: Herzrhythmen, Stoffwechselvorgänge, Hirnaktivitäten, biologische Zyklen
- Ein Gebiet Eurer Wahl