

### **Qualität KI-gestützter Datenanalyse: Potenziale und Probleme**

Im Zuge der quantitativen Datenanalyse können verschiedene Fehler (z.B. Wahl ungeeigneter Analyseverfahren) auftreten. Es ist davon auszugehen, dass vor allem mangelndes Wissen das Risiko fehlerhafter Ergebnisse erhöht. Verschärft wird dieses Risikopotenzial nun durch neue KI-gestützte Möglichkeiten der Datenanalyse, da man nun in der Lage ist, sowohl einfache als auch komplexe Datenanalysen mit nur begrenztem Wissen durchzuführen. Unsere Forschungsfrage lautet daher: Was sind die Fallstricke bei der Nutzung von KI-gestützter Datenanalyse?

Unsere Untersuchung veranschaulicht das Fehlerpotenzial der KI-basierten Datenanalyse am Beispiel deskriptiver Statistik, Regressionsanalyse und Multidimensionale Skalierung. Wir verwenden dazu Replikationsdatensätze und vergleichen die veröffentlichten Ergebnisse mit den KI-basierten. Wir verwenden KI-Werkzeuge, um die erforderliche Syntax für SPSS, Stata und R zu generieren und stellen der KI die Ergebnisse zur Interpretation zur Verfügung.

Die Ergebnisse zeigen: die Verwendung einer von Chat-GPT generierten Syntax ist vielversprechend, erfordert jedoch Fachwissen (zur Überprüfung und Anpassung verschiedener Spezifikationen). Je komplexer jedoch das Verfahren ist, das für die Anwendung in Betracht gezogen wird, desto höher ist das Risiko, dass die Syntax nicht funktioniert bzw. fehlerhaft ist.

Der Nutzen der KI-gestützten Programmierung für die Datenanalyse liegt sicherlich in der Zeitersparnis (z.B. durch das Wegfallen einfacher Programmierfähigkeit). Die Nutzung von KI sollte allerdings mit Vorsicht erfolgen und aufgrund der kontinuierlichen, schnellen Entwicklung von KI-Tools für die Datenanalyse stetig zu „aktualisieren“.