

Bericht zur asp-Methodenfortbildung R (Nov. / Dez. 2022)

Zum Ausklang des Jahres 2022 hat die asp-Methodenfortbildung zur Statistik-Programmiersprache *R* unter der Leitung von Prof. Dr. Till Utesch und Dr. Katharina Geukes von der Universität Münster stattgefunden.

Die gut besuchte, exzellent vorbereitete und atmosphärisch sehr gelungene Online-Fortbildung wurde organisatorisch in zwei Blöcke aufgeteilt. Uns wurde dadurch die Möglichkeit geboten, die grundlegenden Inhalte des ersten Blocks *Einführung, Visualisierung und Projektmanagement in R* am 24. und 25. November 2022 zu wiederholen und zu verinnerlichen, bevor weitere Inhalte im zweiten Block *Datenanalyse in R* am 08. und 09. Dezember 2022 weiter gelehrt und vertieft wurden. Tag für Tag hat sich das Verhältnis von Vortrag und Vorgabe zu eigenständigem Erarbeiten zugunsten des praktischen Ausprobierens verschoben.

Direkt zu Beginn des ersten Blocks der Fortbildung wurde uns von Prof. Utesch gezeigt, welche Möglichkeiten *R* bietet. Wir haben eine Befragung in *formr*, einer kostenlosen Umfrage-Umgebung z.B. zur Umsetzung von komplexen Studien direkt mit *R*, zu Geschlecht, Alter, Bundesland, Fachrichtung, Promotionsthema sowie unseren Statistik- und Programmierkenntnissen ausgefüllt. Durch die Visualisierung der Ergebnisse über ein direkt aus den Daten erstelltes *RMarkdown* haben wir einen schnellen Überblick über die Gruppe der Teilnehmenden erhalten. Wir haben zum Beispiel erfahren, dass wir im Durchschnitt 29.3 Jahre alt sind, Baden-Württemberg das Bundesland mit den meisten Teilnehmenden im Workshop war und die Selbsteinschätzung unserer Statistik- und Programmierkenntnisse mit 5.1 respektive 2.7 von 10 sehr demütig ausfällt.

Von dieser fortgeschrittenen Umsetzung der Befragung sind wir anschließend zurück zum rohen Anfang von *R* gesprungen und wurden erstmal mit der Oberfläche der Open-Source-Software *RStudio*, einem Software-Interface von *R*, vertraut gemacht. Hier wurde uns die grundlegendste Nutzungsmöglichkeit von *R* gezeigt – die Nutzung als Taschenrechner. Diese Funktionsweise ist sozusagen in der Grundausstattung der Software enthalten. Die hervorragend aufeinander abgestimmten und aufeinander aufbauenden weiteren Inhalte der Fortbildung folgten. Hierbei haben wir gelernt, dass es eine Notwendigkeit in *R* ist, zu Beginn eines Projekts sogenannte Pakete, wie z.B. *tidyverse*, *readxl*, *pacman* und *ggplot*, zu installieren und laden, je nachdem welche Funktionen man benötigt. In diesem Punkt befindet sich auch der Vorteil von *R*. Neben der freien Zugänglichkeit der Software, wird sie durch die zahlreichen Pakete ständig verbessert und erweitert. Darauf aufbauend haben wir uns anschließend mit den Funktionen der Pakete *dplyr* zur Transformation von Datensätzen und *ggplot* zur Entwicklung vielfältiger Plots auseinandergesetzt und ausprobiert.

Der zweite Tag begann mit der Wiederholung und Fortführung der Datenvisualisierung in sog. Plots. Die zahlreichen Visualisierungsmöglichkeiten von *R* sind sehr umfangreich, mit die bekanntesten und beliebtesten Funktionen und oftmals der ausschlaggebende Grund für die Nutzung von *R*. Anschließend

hat Dr. Katharina Geukes einen inspirierenden Vortrag zum Thema *Projektmanagement und Open Science* gehalten, der im Wesentlichen den offenen Umgang mit Material, Code und Daten, sowie die Präregistrierung von Forschungsvorhaben thematisiert und illustriert hat. Besonders das Vorgehen, zu publizierende Paper bzw. sogar bereits die Studien in ihrer vollendeten Planungsphase schon zur Präregistrierung beim Journal einzureichen, war wahrscheinlich für fast alle neu und besonders für uns junge Nachwuchswissenschaftler_Innen sehr inspirierend. Im restlichen Teil des ersten Blocks wurden wir in *RMarkdown* eingeführt. Diese Output-Form von *R* ermöglicht den Einsatz als Projektmanagement-Tool, inklusive der Möglichkeit ein Paper direkt in *R* zu schreiben. Wir haben hierzu eine Aufgabe erhalten, um uns eigenständig und fokussiert mit dieser Möglichkeit auseinanderzusetzen.

Der zweite Block der Methodenfortbildung zwei Wochen später widmete sich dann der *Datenanalyse mit R*. Anschließend an die Fragerunde zu Beginn, haben wir die für die Datenanalyse notwendigen Grundlagen des ersten Blocks aufgegriffen. Dazu gehörten das Einlesen und die Umwandlung eines Datensatzes sowie die Rekodierung einzelner Variablen. Hier wurden die Aufgaben für (fast) alle Teilnehmenden schon deutlich anspruchsvoller. Auch in diesem Teil der Fortbildung wurde konsequent ein sehr guter Mix aus Fragen in der großen Gruppe und einzelnen Break-out-Sessions für individuelle Anfragen gefunden. Mit den transformierten Daten konnten wir einfache und komplexere statistische Analysen in *R* umsetzen, angefangen mit Korrelations- und Regressionsanalysen sowie zugehöriger Plot-Visualisierungen. Dazu wurde uns hilfreiches Hintergrundwissen über spezifische Bezeichnungen, Grundlagen sowie eine Auffrischung unseres statistischen Verständnisses vermittelt. Der abschließende vierte Fortbildungstag beinhaltete eine intensive Übung zur Aufbereitung eines Datensatzes, um mit diesem in *R* arbeiten zu können. Darüber hinaus wurden wir in einem kurzweiligen Vortrag von Prof. Dr. Till Utesch über Datenanalyse im Forschungsprozess, Hypothesenbildung und Effektstärken in Multilevel-Modelle eingeführt. Diese haben wir anschließend auch noch in *R* umgesetzt, da wir uns diese Form der Datenanalyse im Vorfeld des zweiten Blocks gewünscht hatten und von den Dozierenden immer wieder die Artikulation von auch spezielleren Fragen und Wünschen zu Inhalten gefordert und initiiert wurde.

Kurzfristig hat sich diese inhaltlich gehaltvolle und spannend umgesetzte Methodenfortbildung äußerst bereichernd auf die bereits angesprochene geringe Vorerfahrung der Gruppe hinsichtlich der Statistik- und Programmierkenntnisse ausgewirkt. Die Aufrechterhaltung dessen, liegt nun aber in unserer eigenen Hand, auch wenn wir bestimmt alle die beiden Dozierenden gerne dauerhaft neben uns sitzen hätten. Bei all den neuen und vielfältigen Inhalten, die didaktisch ausgezeichnet aufbereitet wurden, haben wir uns alle verständlicherweise hin und wieder etwas überfordert gefühlt. Wir wurden

aber stets von beiden Dozierenden aufgefangen und bestärkt. Die Hilfsbereitschaft und Offenheit der beiden für Fragen sowohl vor, während als auch im Nachhinein des Workshops, sowie das sehr gut aufbereitete und hilfreiche Material, das uns zur Verfügung gestellt wurde, haben die Fortbildung umso lohnender gemacht. Und für uns steht fest: langfristig wird sich die Teilnahme an der asp-Methodenfortbildung *R* äußerst positiv auf die Statistik- und Programmierkenntnisse jeder Person unserer Gruppe auswirken!

Wir möchten uns noch einmal herzlichst bei Prof. Dr. Till Utesch und Dr. Katharina Geukes, die durchweg ihr umfassendes und detailliertes Wissen in R-Programmierung, Statistik, Projektmanagement und Open Science bewiesen, für die sehr gelungene Veranstaltung und die großartige Atmosphäre bedanken! Abschließend möchten wir nachfolgende Workshops der Dozierenden jedem empfehlen.

Melinda Herfet, Uni Bayreuth

Frederik Hellermann, Uni Paderborn