

Projektleitung: Dr. Tobias Siebenlist, Christine Meschede

Interaktive Statistik mit R – ein problembasierter Ansatz

Ziel des Projekts

- Zielgruppe: Masterstudierende des Studiengangs „Informationswissenschaft und Sprachtechnologie“
- Ziel: Neugestaltung der statistischen Grundlagenveranstaltung
- Erwerb statistischer Grundkenntnisse zeit- und ortsunabhängig mit der freien Software R

Inhaltliche Ausrichtung

- Grundlagen der Statistik
- Typische Anwendungsfälle im Bereich Informationswissenschaft und verwandten Disziplinen
- Erhebung von Daten, Bereinigung, Verarbeitung, Aufbereitung dieser
- Auswertung und Interpretation von erhobenen Daten, statistische Auswertungen
- Reale Datensätze und Beispiele
- Vorbereitung für Abschlussarbeiten und das Berufsleben

Konzeption des Projekts

Hauptziel des Projekts war die Bereitstellung von umfangreichen Lernmaterialien zur Vor- und Nachbereitung des Seminars mit einem starken Fokus auf problemorientiertem Lernen (POL). Kern der Materialien bilden die folgenden Elemente:

- Videos / Screencasts
- Swirl-Module / DataCamp-Kurs

Die Lehrveranstaltung wurde durch die eLearning-Angebote nicht ersetzt, sondern sinnvoll ergänzt und unter Einbezug der eLearning-Inhalte komplett neu geplant und strukturiert.

Ergebnisse und Evaluation

- Evaluation der Videos durch Umfragen im Seminar
 - Videos eignen sich gut, die grundlegenden Inhalte und Zusammenhänge noch einmal zu erfahren
 - Verbindung von Grundlagen und Umsetzung in R kann noch einmal gut nachvollzogen werden
 - Vor- und Zurückschalten möglich, daher eigenes Tempo einstellbar
 - Umfang jedoch auf wenige Themen begrenzt
- Evaluation der interaktiven Lektionen qualitativ durch Befragung einzelner Studierender
 - Interaktive Arbeit an Aufgaben direkt im Browser hat gut gefallen
 - Hinweise auf der Plattform und schrittweise Erarbeitung von Inhalten sind gut gelöst
 - Plattform wirkt modern, Gamification-Elemente spornen bei der Nutzung an

Videos / Screencasts

- Theoretische Grundlagen der Statistik und Inhalte des Seminars werden aufgearbeitet (**Abbildung 1**)
- Screencasts zur Benutzung der Software RStudio und praktische Anwendung der Theorie (**Abbildung 2**)

Skalenniveau	Modus	Median	Arithmetischer Mittelwert	Geometrischer Mittelwert
Nominal	✓	✗	✗	✗
Ordinal	✓	✓	✗	✗
Intervall	✓	✓	✓	✗
Verhältnis	✓	✓	✓	✓

Abb. 1. Wiederholung der theoretischen Grundlagen in den Videos

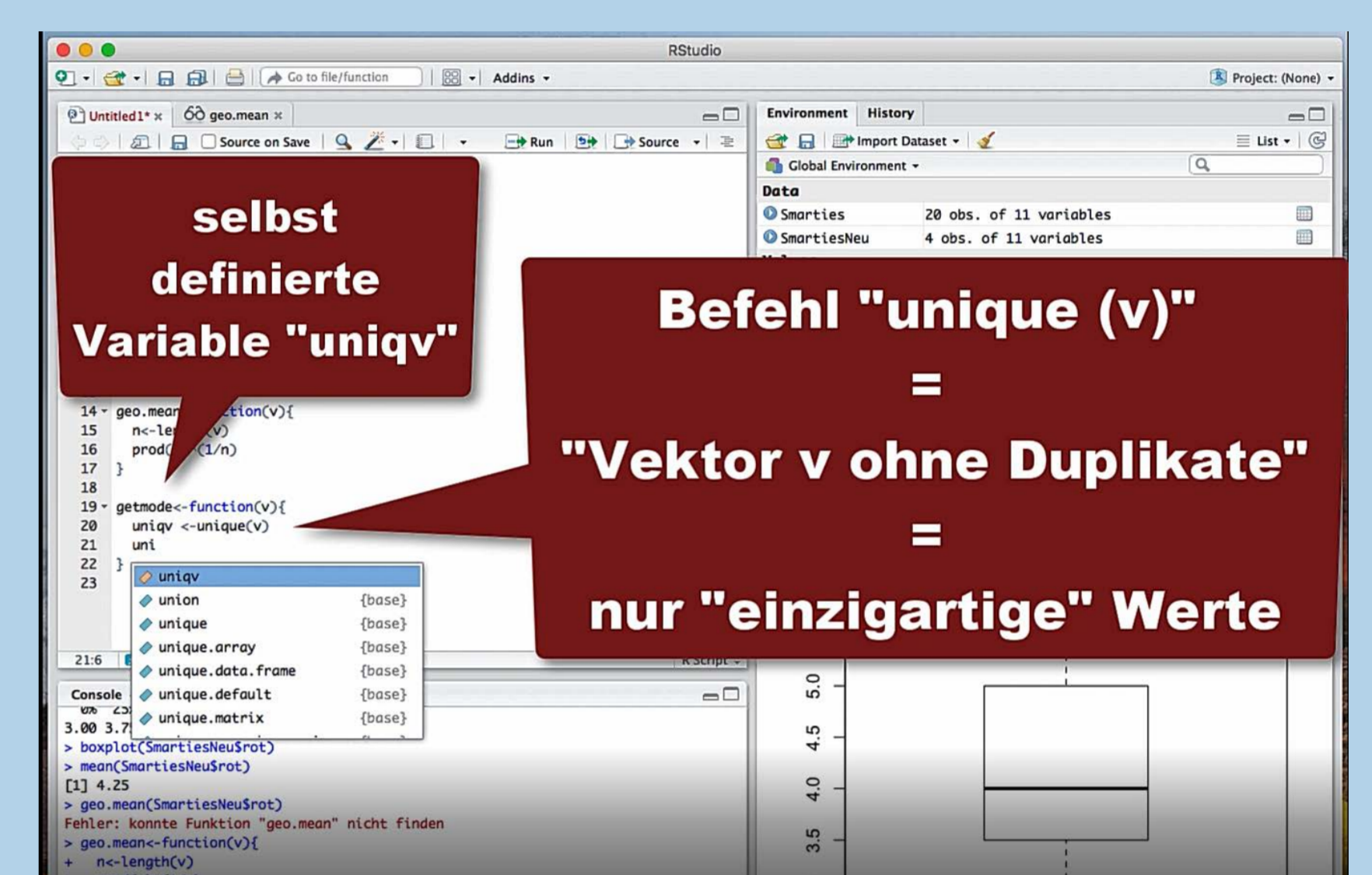


Abb. 2. Screencasts zur Benutzung der Software RStudio

Swirl / DataCamp

Swirl ist eine Erweiterung für R, welche das Durchlaufen interaktiver Module ermöglicht.

- Im Rahmen des Projekts wurden eigene swirl-Module erstellt, angepasst an die Themen der Informationswissenschaft
- Die Abläufe der Seminarinhalte konnten so noch einmal Schritt für Schritt (inklusive Prüfung der Ergebnisse) durchgeführt werden
- Probleme bei der Verarbeitung von Eingaben machten einen Wechsel nötig; die Wahl fiel auf DataCamp

DataCamp.com ist eine Plattform für interaktive Online-Kurse, auf der auch eigene Inhalte erstellt werden können (**Abbildung 3**).

- Der interaktive Online-Kurs kann im Browser belegt werden, in dem ein interaktiver Editor, Beschreibungstexte, Datensätze und Lösungshinweise zur Verfügung stehen (**Abbildung 4**)
- Gamification-Elemente der Plattform können zu einer verstärkten Motivation und damit besseren Beteiligung beitragen (**Abbildung 5**).

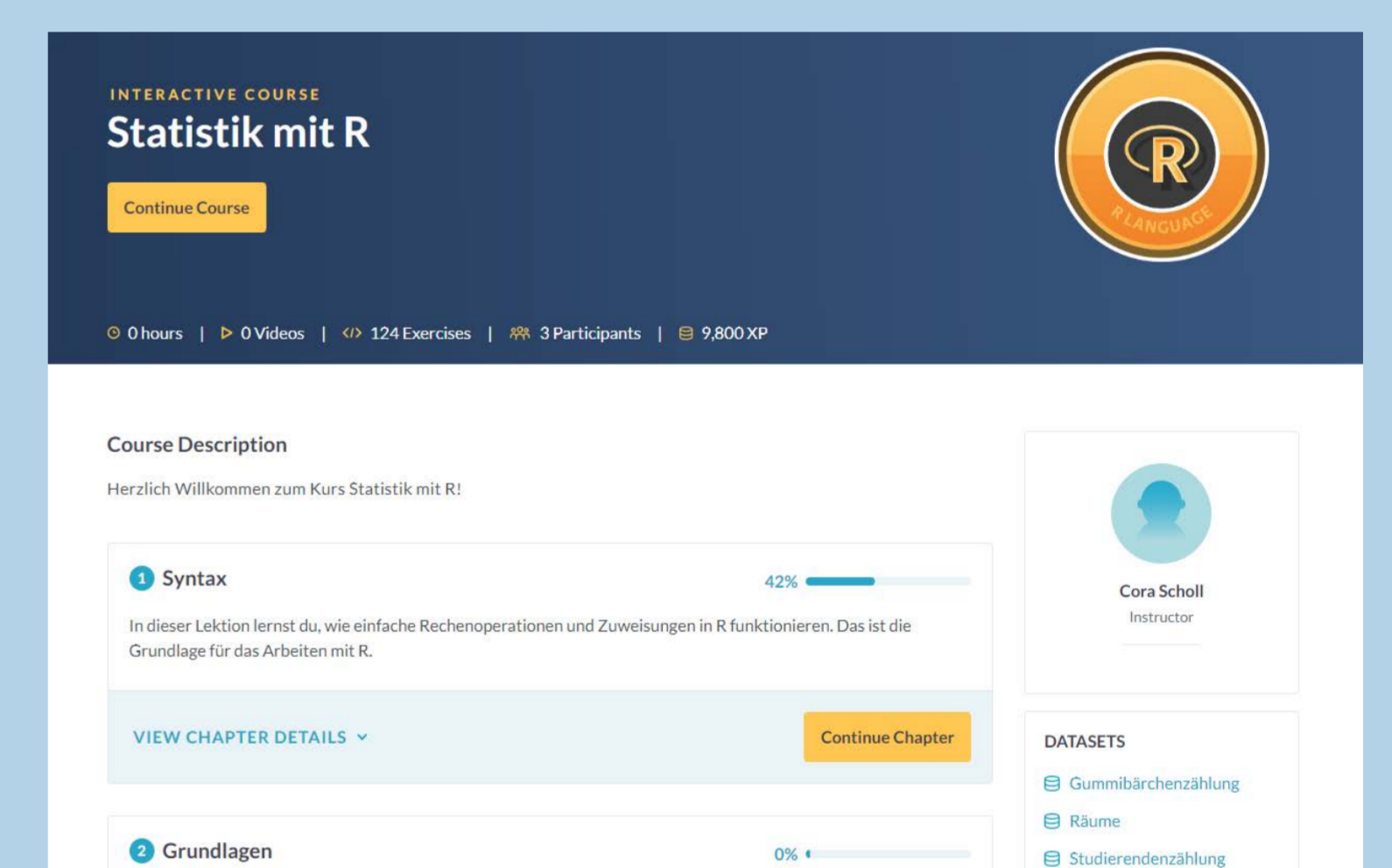


Abb. 3. DataCamp-Kurs zum Projekt

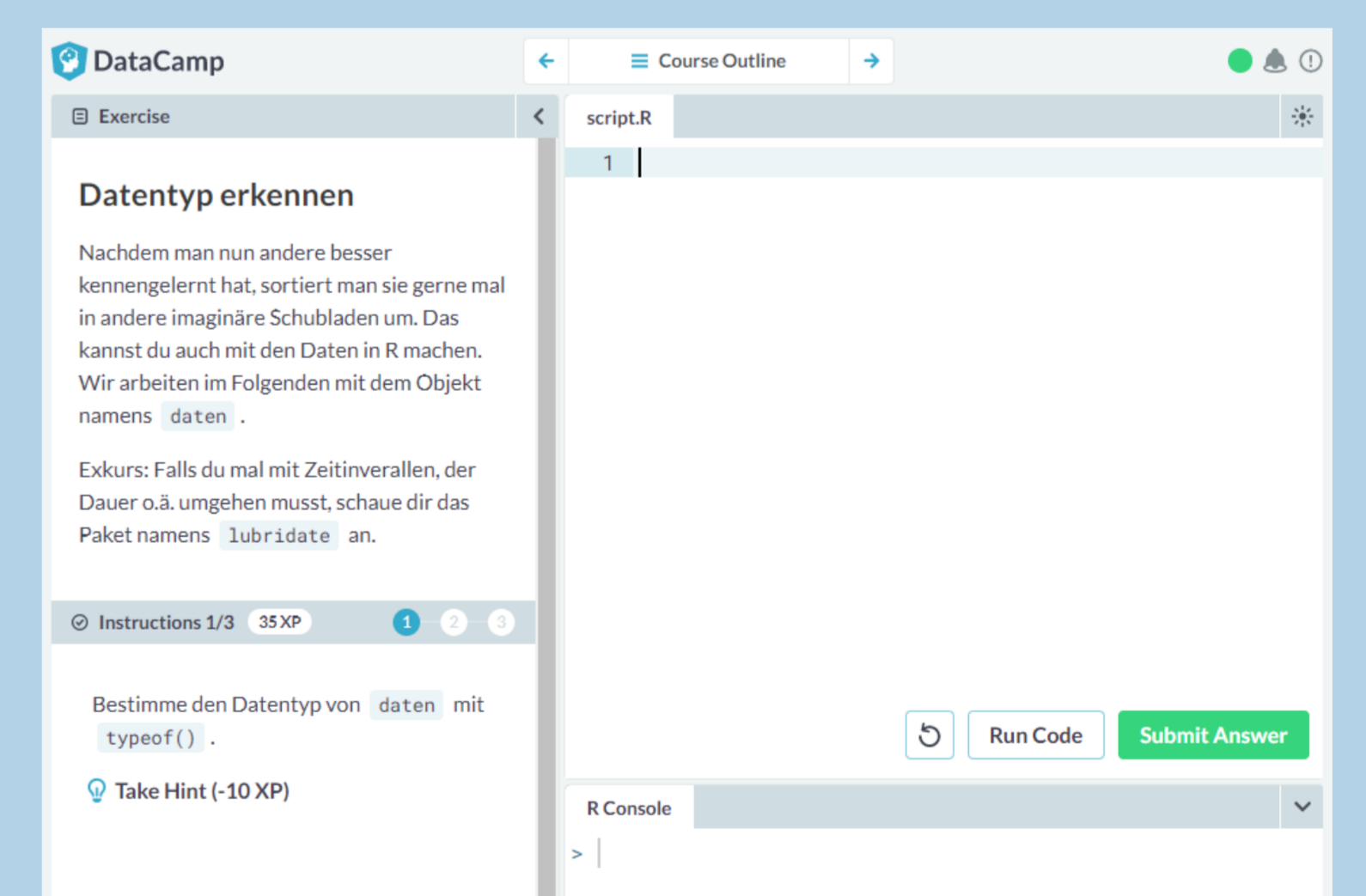


Abb. 4. Interaktiver Editor bei DataCamp

Datenstrukturen 17% 100 XP

In dieser Lektion geht es um Persönlichkeitstypen... ääh, Datentypen. Da gibt es die pünktlichen, gewissenhaften oder die mit der Einstellung 'wir nehmen's wie es kommt'. Und nicht nur im echten Leben entwickelt man ein Gespür, wie mit verschiedenen Persönlichkeitstypen... äh Datentypen umzugehen ist. Eine große Quelle von Missverständnissen ist schon das Kennenlernen. Beim Einlesen von Daten legt die Funktion, die zum Einlesen benutzt wird, die Datentypen erstmal automatisch fest. Da können schon mal 'Vorurteile' entstehen.

- Datentypen anzeigen 100 XP
- Verschiedene Datentypen 50 XP
- Datentyp erkennen 100 XP
- Datentypen verändern 100 XP
- Objekttyp 100 XP
- Faktoren 100 XP

Übersicht über gesammelte XP

Instructions 100 XP

Konstruiere analog zu dem Beispiel ein neues Beispiel, bei dem die Spalte `alter` ausgewählt wird und ein Jahr zu dem Alter addiert wird. Kleiner Tipp: Die Addition von z.B. eins funktioniert in magrittr mit der Funktion `~+*(1)`. Benutze auch das Pipe-Symbol `%>%` um die Anweisungen zu verketten.

Take Hint (-30 XP) **XP-Abzug bei Lösungshinweis**

Instructions 100 XP

Wende diese Funktion auf `id` an und überschreibe `id`.

Show Answer (-100 XP) **XP-Abzug bei Anzeige der Lösung**

Abb. 5. Gamification-Elemente auf der Plattform: Für das Lösen der Aufgabe werden Erfahrungspunkte (XP) gesammelt, für Lösungshinweise werden XP abgezogen