

Prüfungsaufgabe „Grundlagen Data Science“

Prof. Dr. Martin Lachmair

22.11.2022

Anmerkungen:

- Abgabe ist am **16.12.2022** von jedem Einzelnen mit Namen
- **50 Punkte** sind maximal zu erreichen.
- Die zu verwendende Datenquelle ist auf Moodle.
- Datenmanipulationen nur in und mit R erlaubt.
- **Abzugeben:**
- Ein grundsätzlich lauffähiges R-Skript (Shiny-App)
- Ein PDF mit den Abbildungen und dem Storytelling (Erstellen Sie dazu 3-5 Folien, die gute Argumente liefern für Open-Data und einen Ansatz für einen guten Einstieg)

Aufgabe:

Auf dem Open-Data Portal von Berlin gibt es Nutzerdaten, die eine Statistik über die Nutzung des Portals erlauben. Ziel ist es, eine Präsentation zu erstellen, die gute Argumente für mehr und besseres Open Data in Stuttgart liefern.

- a) Erstellen Sie eine Grafik über die 10 am häufigsten verwendeten Dienste mit den Page impressions und visits. Beschreiben Sie diese beiden Merkmale in Worten. (20 Punkte)
- b) Geben Sie auch eine Auskunft über die 10 am wenigsten verwendeten Dienste im Portal an. (10 Punkte)
- c) Erstellen Sie eine **Shiny-App** die den monatlichen Verlauf der Nutzung dieser Dienste darstellen kann, für beide Merkmale: „Page impressions“ und „visits“. (10 Punkte)
- d) Was kann man über das Verhältnis zwischen „Page impressions“ und „visits“ sagen? (5 Punkte)
- e) Ermitteln Sie ebenfalls, mit welchen Diensten im Rahmen des in den Daten gegebenen Zeitraums gestartet worden ist. (5 Punkte)
- f) Bonus: Ermitteln Sie weitere Fakten, die für Ihre Argumentation wertvoll sein könnten (5 Extra-Punkte).