

Projekt "Einführung in die Datenanalyse mit R"

Stackoverflow Developer Survey 2019

Erstellung eines Reports zur Auswertung der *Entwicklerumfrage von Stackoverflow 2019*

Im Archiv *Stackoverflow_Developer_Survey_2019.zip* finden Sie folgende Dateien:

- *README_2019.txt* → Beschreibung
- *so_survey_2019.pdf* → Fragebogen
- *survey_results_public.csv* → Umfrage-Daten
- *survey_results_schema.csv* → Daten-Schema

Aufgabenstellungen

- Erstellung eines "durchgehenden" Reports (mit Erklärungen zwischen den Code- bzw. "Analyse"-Blöcken)
- Aufarbeiten der Rohdaten
- Geeignetes "zur Verfügung stellen" der Rohdaten für den Report, z.B. in Form einer Datenbank als Datenquelle → Abfrage über in R eingebettetes SQL
- Kurze Beschreibung der Daten
- Explorative Auswertung der Daten mittels aussagekräftiger Grafiken und Statistiken, z.B. folgende Fragestellungen wären möglich:
 - Wer hat alles an der Umfrage teilgenommen - Verteilung nach Region (*Geo-Visualisierung*)
 - Job-Zufriedenheit
 - insgesamt?
 - In Abhängigkeit vom Ausbildungsgrad?
 - Zufriedenheit mit der Karriere
 - In Abhängigkeit von Ausbildungsgrad?
 - Gehalt
 - Welches Lohnniveau in welchen Ländern?
 - In Abhängigkeit von der Berufserfahrung?
 - Gender-Pay-Gap:
 - Spiegelt sich die Gender-Pay-Gap in den Daten wieder?
 - Wo ist sie am größten - wo am kleinsten? (Region, Land, ... *Geo-Visualisierung*)
 - Hat Sie Einfluß auf die Abhängigkeit des Gehalts von den Berufsjahren?
 - ...
- Vergessen Sie auch nicht Quellen anzugeben, falls Sie bestimmte andere Quellen heranziehen oder auf besondere (nicht-standard) Packages zurückgreifen, z.B. für besondere grafische Darstellungen

Abgabe

Datum: bis 18.01.2021

Die Abgabe sollte ein Zip-Archiv mit den von Ihnen verwendeten Files, wie z.B. R-Skripte, R-Markdown, Datenbank-File oder sonstige für den Report notwendige Ressourcen - sowie das gerenderte Endergebnis enthalten.

