

Übungsblatt 5

Aufgabe 1 (9 Punkte)

Dass wir gut mit DataFrames umgehen können, ist absolut zentral. Und da hilft nur Übung, Übung, Übung. Deshalb geht es hier nun in die nächste Runde Datenanalyse...

- (a) Im GitHub-repository unter **data** befindet sich die Datei **fish.csv**.
Ladet den Datensatz in Julia als **DataFrame** ein.
- (b) Macht Euch mit der Struktur des Datensatzes vertraut. Gebt eine Tabelle mit dem Gewicht der Fische je Spezies aus (für die Funktion **mean** braucht ihr wieder das Modul **Statistics**).
- (c) Überlegt euch kurz, wozu wir **CategoricalArrays.jl** brauchen.
- (d) Die Spalte **:Spezies** soll nun kategorische Variablen beinhalten; die levels sollen nach Anzahl an Observationen geordnet sein (Tipp: Erstellt euch einen Array mit sortierten Spezies.).
- (e) Erstellt eine lineare Regression, um die **Laenge_diag** anhand von **Laenge_vert** vorherzusagen. Bestimmt auch die Kennzahl R^2 für die Regression.
- (f) **select**: Gebt alle Spalten bis auf die letzten beiden aus. Lasst Euch darüber hinaus alle Spalten anzeigen, die den Buchstaben „a“ enthalten (Tipp: **names**).

Aufgabe 2 (5 + 2 Punkte)

In dieser Aufgabe geht es um Korrelationen. Wenn noch gar kein Vorwissen über Korrelation vorhanden ist, bitte etwas in das Thema einlesen.

- (a) Arbeitet mit dem Datensatz aus Aufgabe 1 weiter. Berechnet mithilfe des **cor()** Befehls die Korrelation zwischen **Laenge_vert** und **Gewicht**. Erscheint ein Zusammenhang dieser Variablen sinnvoll?
- (b) Plottet die Spalten **Laenge_vert** und **Gewicht** gegeneinander ab und fügt eine Ausgleichsgerade hinzu. Wie gut passt die Ausgleichsgerade zu den Daten? (Tipp: **predict**)
- (c) Zusatzaufgabe: Erstellt eine quadratische Regression (wieder mit hinzugefügtem Linienplot) für die vorherige Teilaufgabe.

Aufgabe 3 (2 Punkte, Zusatzaufgabe)

Überlege Dir, wie man Folgendes elegant hinbekommt:

Für $x = [[0, 1], [0, 1, 2], \text{ usw.}]$ hätten wir gerne einen Operator \oplus , sodass gilt
 $\exp^\oplus(x) = [[1.0, 2.71828\dots], [1.0, 2.71828\dots, 7.38905\dots] \text{ usw.}]$

Viel Erfolg!