



Prof. Hanspeter Herzel
Prof. Nils Blüthgen
Dr. Manuela Benary
Institute for Theoretical Biology

www.sys-bio.net/teaching
h.herzel@biologie.hu-berlin.de
nils.bluthgen@charite.de
manuela.benary@biologie.hu-berlin.de

BIOINFORMATIK SS 2017

Gib deine Lösung bitte am 22.5. in der Vorlesung ab. Alternativ kannst du die Lösung auch per E-mail an nils.bluthgen@charite.de schicken.

1. Normalisierung (Die Aufgabe kann entweder mit Hilfe von einem Tabellenkalkulationsprogramm oder einem Statistikprogramm gelöst werden)

Es werden folgende zwei Expressionsdatensätze gemessen:

	Datensatz 1	Datensatz 2
Gen 1	86	74
Gen 2	93	87
Gen 3	67	45
Gen 4	25	6
Gen 5	42	18
Gen 6	50	16
Gen 7	62	39
Gen 8	15	2
Gen 9	53	28
Gen 10	26	12
Gen 11	61	37
Gen 12	32	10
Gen 13	43	19
Gen 14	29	9
Gen 15	70	49
Gen 16	90	81
Gen 17	73	53
Gen 18	80	64
Gen 19	17	3

- (a): Zeichne einen Scatterplot (gerne auch mit dem Computer)!
- (b): Welches Gen zeigt die größte Änderung in der Expression?
- (c): Führe eine Z-Normalisierung durch und zeichne einen Scatterplot!
- (d): Führe eine Quantilnormalisierung durch und zeichne einen Scatterplot!
- (e): Welches Gen ist nach der Quantilnormalisierung am stärksten reguliert?



2. Korrelation

(a): Berechne die Pearson-Korrelation der Daten aus Aufgabe 1!

(b): Berechne die Rangkorrelation der Daten aus Aufgabe 1!

(c): Wird ändert sich die Pearson-Korrelation und die Rangkorrelation durch die Quantilnormalisierung (muss nicht ausgerechnet werden. Antworten: wird größer, kleiner, bleibt gleich)? Bitte Begründung!