

## Визуелизација и анализа података

Други поправни колоквијум

12. јануар 2021.

1. Написати  $R$  скрипту која учитава матрицу реалних бројева  $M$ , димензије  $4 \times 4$ , а затим учитану матрицу претвара у листу  $T$ . Затим одредити:

- Индексе свих елемената листе  $T$  који су већи од средње вредности листе  $T$ .
- Направити листу  $T_2$  од елемената листе  $T$  који задовољавају услов из претходне тачке задатка. Минимални елемент листе  $T_2$  заменити медијаном листе  $T_2$ .
- Решити систем једначина

$$3a + 4b - 5c + d = 10$$

$$2a + 2b + 2c - d = 5$$

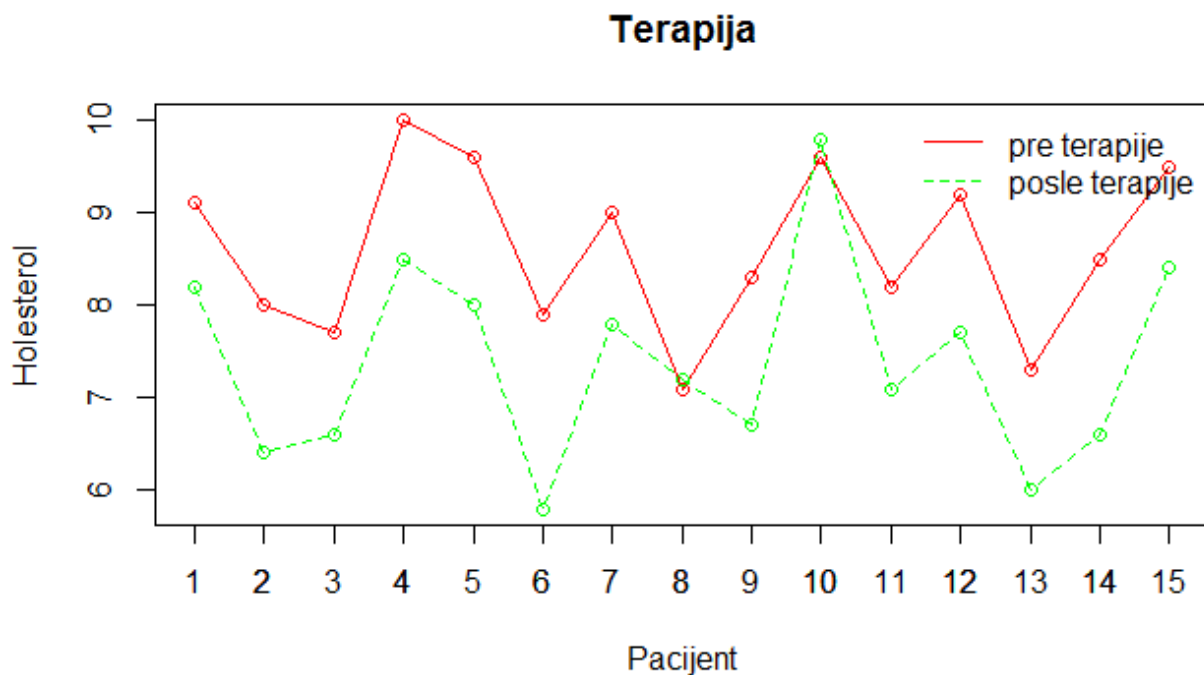
$$a - b + 5c - 5d = 7$$

$$5a + d = 4.$$

2. Спроведена је терапија за снижавање холестерола за 15 пацијената. Забележени ниво холестерола пре и после терапије дат је у табели:

пацијент	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
пре терапије	9.1	8.0	7.7	10.0	9.6	7.9	9.0	7.1	8.3	9.6	8.2	9.2	7.3	8.5	9.5
после терапије	8.2	6.4	6.6	8.5	8.0	5.8	7.8	7.2	6.7	9.8	7.1	7.7	6.0	6.6	8.4

- За колико се разликују медијане ова два мерења? У ком од два мерења је расипање око средње вредности веће?
  - Формирати вектор који представља разлику мерења пре и мерења после терапије. Одредити минимум, максимум и квилите за дати вектор. Формирати хистограм за тај вектор и на основу њега оценити да ли се расподела вредности значајно разликује од нормалне, уз образложење.
  - Да ли за дати вектор из дела под б. постоје нетипичне вредности? Уколико постоје, издвојити их из тог вектора.
3. Податке пре и после терапије из задатка 2 приказати графички по угледу на дату слику.



4. У фајлу Employee data.csv дати су подаци о запосленима једне компаније.

- а. Одредити Спирманов коефицијент корелације између почетне (salbegin) и тренутне плате (salary) **запослених мушкараца**. Одредити линеарни регресиони модел којим се помоћу вредности почетне плате рачунају вредности тренутне плате. На графику представити податке и регресиону праву. Графику додати наслов.
- б. Помоћу тог модела одредити предвиђену вредност тренутне плате за мушкарца чија је почетна плата 70000.
- в. Одредити јачину и смер линеарне корелације између почетне плате (salbegin) и претходног радног искуства (prevexp) код **запослених жена** помоћу Пирсоновог коефицијента корелације. На основу тога закључити има ли смисла формирати линеарни модел којим се почетна плата за жене одређује према претходном радном искуству.