

METODY EKSPLOKACJI DANYCH
Laboratorium. Analiza regresji
Dobór zmiennych do modelu wieloczynnikowego
Budowa modelu na podstawie danych (ogólnie dostępnych)

Zadanie

Zebrano pomiary 250 mężczyzn w różnym wieku (źródło: <http://www.byu.edu/chhp>). Zebrane dane zawierają poprawnie zmierzony % tłuszczu „Pct.BF” (%bodyfat) oraz inne zweryfikowane wartości zmiennych charakteryzujących sylwetkę. Tab. 1 zawiera tylko fragment danych i stanowi tylko ich ilustrację. Pełny zestaw danych znajduje się w pliku o nazwie „MED-Lab-1-Zad-Dobor zmiennych-bodyfat-dane-i-opis.txt”

Tabela 1. Fragment danych z pliku „MED-Lab-1-Zad-Dobor zmiennych-bodyfat-dane-i-opis.txt”

Gęstość	Pct.BF	Wiek	Waga	Wzrost	Szyja	Klatka	Brzuch	Talia
Density	Pct.BF	Age	Weight	Height	Neck	Chest	Abdomen	Waist
1.0708	12.3	23	154.25	67.75	36.2	93.1	85.2	33.543307
1.0853	6.1	22	173.25	72.25	38.5	93.6	83	32.677165
1.0414	25.3	22	154	66.25	34	95.8	87.9	34.606299
1.0751	10.4	26	184.75	72.25	37.4	101.8	86.4	34.015748
1.034	28.7	24	184.25	71.25	34.4	97.3	100	39.370079
1.0502	20.9	24	210.25	74.75	39	104.5	94.4	37.165354
1.0549	19.2	26	181	69.75	36.4	105.1	90.7	35.708661
1.0704	12.4	25	176	72.5	37.8	99.6	88.5	34.842520
1.09	4.1	25	191	74	38.1	100.9	82.5	32.480315
1.0722	11.7	23	198.25	73.5	42.1	99.6	88.6	34.881890
1.083	7.1	26	186.25	74.5	38.5	101.5	83.6	32.913386
1.0812	7.8	27	216	76	39.4	103.6	90.9	35.787402
1.0513	20.8	32	180.5	69.5	38.4	102	91.6	36.062992
1.0505	21.2	30	205.25	71.25	39.4	104.1	101.8	40.078740
1.0484	22.1	35	187.75	69.5	40.5	101.3	96.4	37.952756
1.0512	20.9	35	162.75	66	36.4	99.1	92.8	36.535433
1.0333	29	34	195.75	71	38.9	101.9	96.4	37.952756
1.0468	22.9	32	209.25	71	42.1	107.6	97.5	38.385827
1.0622	16	28	183.75	67.75	38	106.8	89.6	35.275591

Proszę:

- 1) zbudować model pozwalający przewidzieć %bodyfat na podstawie innych zmiennych. Procent tłuszczu ciała każdego (%bodyfat, PctBF) znajduje się w drugiej kolumnie danych.
- 2) przed budową modelu proszę zaproponować/wybrać procedurę eliminacji zmiennych wraz z uzasadnieniem.

- 3) dodatkowo, na podstawie danych zaproponować model/-e dla innej/-ych zmiennej/-ych objaśnianej/-ących.
- 4) przygotować w formie pisemnej wyczerpujące sprawozdanie z wykonania zadania (!), które wraz z innymi plikami (wymaganymi do weryfikacji rozwiązania zadania) należy umieścić w prywatnym notesie zespołu MS Teams w sekcji „Laboratorium” w odpowiednio nazwanej stronie, np. Sprawozdanie z lab.1