

Pro účely této analýzy jsem zvolil data s názvem Supermodel.sav z dat Andyho Fielda (2009). Tato matice zachycuje informace o **213** supermodelkách , přičemž u každé z nich jsou zaznamenány 4 proměnné, a to:

- **denní plat uveden v librách (m=11,34, min. hodnota=0, max. hodnota=95,34, s=16,03)**
- **věk (m=18,07, min. hodnota=12,17, max. hodnota=25,29, s=2,42)**
- **počet let v modelingovém byznysu (m=4,59, min. hodnota=0,43, max. hodnota=9,93, s=1,58)**
- **krása (atraktivita) uvedená v procentech (m=75,94, min. hodnota=58,13, max. hodnota=99,22, s=6,77)**

Jako závislá pro mnohonásobnou lineární regresi byla vybrána proměnná **denní plat**. Jako predikátory tedy sloužily proměnné věk, atraktivita a počet let v byznysu. Snažili jsme se tudíž zjistit která z proměnných má pozitivnější souvisí s platem. Regrese byla provedena v jednom kroku v spss.

| koeficient | B | SE B | β | Kolinearita (VIF) |
|---------------------|--------|------|---------|-------------------|
| konstanty | -60,89 | 16,5 | | |
| Krása | -5,56 | 0,15 | -0,08 | 1,15 |
| věk | 6,23 | 1,41 | 0,94 | 12,16 |
| Počet let v byznysu | -0,2 | 2,12 | -0,55 | 12,65 |

$$R^2 = 0,18, \text{ sig} = 0$$

Jelikož se však hodnota největšího VIF pohybuje nad 10 regrese je velice pravděpodobně zkreslená (Myers, Bowerman and O-Connell , cit. Dle Field, 2005). A to z důvodu, že proměnné počet let v byznysu a věk spolu velice silně korelují ($r=0,96$).

Proto uděláme regresi pouze s dvěma prediktory, krásou a věkem.

| koeficient | B | SE B | β | Kolinearita (VIF) |
|------------|--------|-------|---------|-------------------|
| Konstanty | -30,47 | 11,87 | | |
| Krása | -0,09 | 0,15 | -0,04 | 1,07 |
| věk | 2,7 | 0,42 | 0,41 | 1,07 |

$$R^2 = 0,16, \text{ sig} = 0$$

V této regresi byly kromě absence vysoké kolinearity naplněny také další předpoklady regrese, jako homoskedasticita , linearita nebo nezávislost reziduí (Durbin-Watson=2,07) a další.

Výše zmíněná data lze pomocí hodnoty **B** interpretovat takto:

- Pokud se **atraktivita** zvýší o procento tak plat klesne o 0,09 liber za den.
- Pokud se **věk** zvýší o rok tak plat stoupne o 2,7 liber za den.

Výše zmíněné interpretace jsou platné pouze pokud zbylé proměnné zůstanou stejné.

Jak je z regrese vidno, v případě platového ohodnocení modelky překvapivě nejde ani tak o krásu jako spíš o její věk.

Nevidím tu další předpoklady lineární regrese. Toto je bohužel nedostatečná oprava, kterou ale přijímám z důvodu, že jste na to byl tentokrát sám. U dalších úkolů a u hlavně u zkoušky by toto nebylo považováno za splnění.

Zdroje:

Field, A.(2005). *Discovering Statistics using SPSS*. London: Sage publications

SPSS Datafiles. (2009). Retrieved from:
<http://www.uk.sagepub.com/field3e/SPSSdata.htm>