

2. SzRM gyakorlat

(szept 26, kedd 2-4, 4-6; szept 28, csütörtök 4-6)

Hatvan amerikai nagyváros statisztikai adatai

Y	halálozási ráta	Total age-adjusted mortality rate per 100,000	X8	városi népsűrűség	Population per sq. mile in urbanized areas, 1960
X1	csapadék	Average annual precipitation in inches	X9	nem-európaiak aránya	% non-white population in urbanized areas, 1960
X2	januári hőmérséklet	Average January temperature in degrees	X10	fehérgallérosok aránya	% employed in white collar occupations
X3	júliusi hőmérséklet	Same for July	X11	bevétel 3000 dollár alatt	% of families with income less than \$3000
X4	idősek aránya	% of 1960 SMSA population aged 65 or older	X12	CO2 szennyezés	Relative hydrocarbon pollution potential
X5	háztartások átlagos mérete	Average household size	X13	NO szennyezés	Same for nitric oxides
X6	iskolázottság	Median school years completed by those over 22	X14	SO2 szennyezés	Same for sulphur dioxide
X7	összkomfort aránya	% of housing units which are sound & with all facilities	X15	páratartalom	Annual average % relative humidity at 1pm

Az adatok alapján a halálozási ráta modellezésére egy lineáris regresszió modellt építünk.

1. Döntse el, milyen erősségű lineáris összefüggés mutatható ki a 3000 \$ alatti bevétellel rendelkezők aránya és a halálozási ráta között. Készítsen pontdiagramot a városokról a két változó segítségével! Illesszen lineáris regressziót úgy, hogy a magyarázott a halálozási adat legyen. Számszerűsítse a modell jóságát. Értelmezze magyarázó változó együtthatóját!

2.a. A légszennyeződés hatása a halálozási rátára.

Válassza ki azokat a változókat, amelyek a légszennyeződést írják le! Számoljon korreláció mátrixot és értelmezze értékeit!

Illesszen megfelelő, többdimenziós lineáris regresszió modellt. Értékelje a modell és az együtthatók szignifikanciáját!

Melyek a vizsgálható hipotézisek? Melyek azok a változók, amelyeket a modelltől kihagyna?

2.b. A fölöslegesnek ítélt változók kihagyása után illesszen új modellt és értelmezze azt!

3. A szociális és gazdasági tényezők hatása a halálozási rátára.

Válassza ki azokat a változókat, amelyek a gazdasági, szociális tényezők csoportjába tartoznak! Számolja ki a korreláció mátrixot és értelmezze!

Illesszen megfelelő, többdimenziós lineáris regresszió függvényt. Értékelje a modell és az együtthatók szignifikanciáját.

Melyek azok a változók, amelyeket a modelltől kihagyna? Elemezze az új modellt!

4. Elemezze lineáris regresszió segítségével az időjárás hatásokat a halálozási rátára nézve!

5. Készítsen egy összesített modellt, figyelembe véve az előző eredményeket!