

Statisztikai programcsomagok gyakorlat

Pót zárthelyi dolgozat

A feladatok megoldásához használandó adatállományok: *potzh* és *potolando* (weboldalon található) Az állományok kiterjesztése sas7bdat, a beolvasáskor a fájl típusát erre kell állítani, hogy látszódjanak a fájlok.

Potzh állomány változói:

Változók	Megnevezés
ID	Azonosító
Monogram	Adott egyed monogramja
Nem	Egyed neme
Kor	Kora
Magassag	Magassága (lbs – 1 kg = 2,2046 lbs)
Suly	Súlya (incsben – 1 incs = 2,54 cm)
heti_nikotin_bevitel	Heti nikotin bevitel
diasztoles_vernyomas	Diasztolés vérnyomás érték – A mérés kori második érték (Szisztolés / Diasztolés – még normális érték 140 / 90 mm Hg)

A *potolando* állomány ugyanezeket a változókat tartalmazza, azonban a diasztolés vérnyomás értékek hiányoznak.

Adattranszformáció (10 pont)

1. Olvassa be a potzh adatállományt és az alábbi táblázatban található *pótolandó értékek* oszlopban olvasható értékekkel pótolja az adott változó hiányzó értékeit.

Változó neve	Pótolandó érték
Magassag	Mintaátlag
Suly	Mintaátlag
Heti_nikotin_bevitel	Leggyakoribb elem (Modusz)
diasztoles_vernyomas	Leggyakoribb elem (Modusz)

A megoldás lépéseit szövegesen is dokumentálja!

A mintaátlaggal való pótlás. (3 pont)

A leggyakoribb értékkel való pótlás (7 pont)

Segítség: első lépésben határozza meg a leggyakoribb értéket, majd kódolja át egy új változóba az adott változót úgy, hogy a hiányzó értékek a leggyakoribb értéket vegyék fel, míg a többi megfigyelés esetén másoljuk be a régi értékeket. (Copy old values)

Adatelemzés (15 pont)

1. A továbbiakban az azonos változók közül a pótolatott adatokkal dolgozzon. Határozza meg, hogy melyik változó magyarázza legjobban a diasztolés vérnyomás értékét. (3 pont) - Válaszát szövegesen is indokolja!
2. A kiválasztott változót valamint a diasztolés vérnyomás értékeket ábrázolja pontdiagramon. Mit tapasztal? (2 pont)
3. Határozza meg a regressziós egyenest annak a változónak a segítségével, amelyik a legjobban magyarázza a diasztolés vérnyomás értékét. (5 pont) - Válaszát szövegesen is indokolja, valamint mellékelje a kapott táblázatokat is!
4. A regressziós egyenes meghatározásával pótolja a *potolando* állományban a diasztolés vérnyomás értékeket. (5 pont)

Segítség: A számított regressziós egyenes adatait felhasználva létre kell hozni egy változót (Compute Variable...), ami kiszámítja a hiányzó értékeket. **Ezek után a pótolatott állományt mentse el és a végén ezt is csatolja a megoldásokhoz!!!**

Egyszerű lekérdezések (10 pont)

Válaszolja meg a következő kérdéseket a potzh adatállomány pótolatott adatait használva. A számítások menetét is mellékelje.

1. Mekkora az eltérés a férfiak és nők átlagos diasztolés vérnyomás értékei között? (3 pont)
2. Adja meg a diasztolés vérnyomás percentiliseinek az értékét.(Q1 Median Q3) (3 pont)
3. Átlagosan milyen magasak, milyen súlyosak és mennyi nikotint visznek be a szervezetükbe a magas diasztolés vérnyomásértékkel (>90 mm Hg) rendelkező egyedek? (4 pont)

SPSS programozás (5 pont)

Az alábbi SPSS programkód számos hibát tartalmaz. Adjon meg 5 hibát, ami a kód futása közben jelentkezhet és feltételezve, hogy a nem képzett változók a kívánt formában rendelkezésre állnak. (5 pont) **A hibákat helyes formában és szövegesen is adja meg!**

```
RECODE nem ('Male'=1) ('Female'='2')
INTO nem_recode.
EXECUTE.
```

```
COMPUTE valtozo='suly' -'magassag'.
EXECUTE.
```

```
VALUE LABELS nem_recode 1 'Férfi' 2 'Nő'.
EXECUTE.
```

```
RECODE valtozo
( MISSING = COPY )
( LO THRU 0 ='1' )
( LO THRU HI = 2 )
( ELSE = SYSMIS ) INTO valtozo_kat.
EXECUTE.
FREQUENCIES VARIABLES=nem
BARChart.
```

Mentse el a feladatok megoldása során kapott OUTPUT-t [sajátnév].SPV. A többi eredményt és indoklást egy Word állományba mentse el. A Word, az OUTPUT és a *potolando* nevű adatállományt küldje el a következő e-mail címre: csicsman@calculus.hu