

ZH kérdések a Statisztika gyakorlati alkalmazásai tantárgyból rendszerekből

- 1. Határozza meg az alábbi informatikai rendszerek kapcsolatát:
Metainformációs rendszerek, Tranzakciós rendszerek, Vezetői információs
rendszerek, Adatbányászati technológiák, Adatraktárok.
2 pont**
- 2. Ismertesse a felsorolt informatikai rendszerek fő funkcióit!
6 pont**
- 3. Soroljon fel az integrált informatikai rendszerek megvalósítására szolgáló
eszközöket!
3 pont**
- 4. Végezhetünk-e aritmetikai műveletet a Metainformációs rendszer Mutatóival?
1 pont, jó indoklással 2 pont**
- 5. Meghatározhat-e Metainformációs rendszer nomenklatúrája egy statisztikai
táblában sor, oszlop vagy harmadik dimenziót?
1 pont, jó példák a nomenklatúrákra 2 pont**
- 6. Egy statisztikai tábla egy cellájába, vagy a többdimenziós adatkocka egy elemébe
hogyan számol be a mutató és a nomenklatúra?
2 pont**
- 7. Sorolja fel a legfontosabb statisztikai mutatókat az adatállományok első
megismerésekor!
1 pont definíciókkal 3 pont**
- 8. Soroljon fel a változók és az objektumok megismerésére szolgáló adatbányászati
módszereket!
1 pont definíciókkal 4 pont**

Gyakorlati feladatok:

Adott két SAS állomány. Az egyik állomány egy reprezentatív felmérés lehetséges alanyainak megjelölő adatait tartalmazza (TELJES állomány). A másik a ténylegesen megkérdezetteket, ill. azoknak a kérdőíven szereplő kérdésekre adott kódolt válaszait tartalmazza (MINTA állomány).

Általános feladatleírás: a TELJES állományon szereplő TEA_CSOP mező megoszlása alapján alakítson ki súlyokat (N/n, ahol N = db-szám az adott kategóriában a TELJES állományon, n = db szám az adott kategóriában a MINTA állományon).

A súlyokat írja rá a MINTA állományra, hogy a MINTA állomány mutatóit ezekkel felszorozva majdan következtetni lehessen a teljes sokaságra (*teljeskörűsítés*). Végül jelenítse meg a teljeskörűsített árbevételt TEAOR főcsoportonként az egyes létszám kategóriákban. .

Egy lehetséges megoldás folyamata részletesen:

1. Először a TELJES és a MINTA nevű SAS állományokat töltsse le a saját munkakönyvtárába és ismerkedjen meg velük. (Elég a SAS Explorer | Properties, és a SAS Tableviewer eszköze.)
2. Hozza létre a súly kiszámításához szükséges aggregátumot, vagyis adja össze a TEA_CSOP mező értékei szerint a KICSICC (n) és NAGYCC (N) mezőket. (Ezeket segítségképpen már korábban rátettük a TELJES állományra. A NAGYCC mező fixen 1-et tartalmaz, a KICSICC mezőben jelöltük 1-essel azt, aki a MINTA állományban szerepel.)
Ha az aggregálás előtt nem rendez, BY helyett a CLASS utasítással is megadhatja, hogy mely változó szerint szeretne aggregálni.
Az aggregálás eredményét tegye file-ba, neve legyen TEA_SULY, s csak a szükséges változókat tartalmazza (TEA_CSOP,KICSICC,NAGYCC);

Segítségül egy kis szintaktika:

```
PROC MEANS SUM DATA=aggregálandó állomány
              (KEEP=aggregálási szempont
               és az aggregálandó változók felsorolása)
              noprint; /* hogy az eredmény ne jelenjen meg az output ablakban. */
CLASS aggregálási szempont;
OUTPUT OUT=output SAS állomány(DROP=_TYPE_ _FREQ_)
       /* _TYPE_ és _FREQ_ az aggregálás során keletkező rendszerváltozók, amelyekre jelen esetben
       nincs szükség */
       SUM=;
RUN;
/* Magyarázat: NOPRINT opció= az eredmény ne jelenjen meg az output ablakban.
_Type_ és _FREQ_ az aggregálás során keletkező rendszerváltozók, amelyekre most nincs szükség */
```

(A használandó eljárás szintaktikája megtalálható a

3. A 2. feladatban létrehozott állományon hozza létre a SULY változót a fentebb leírt módon. (Vigyázat, ha a nevező hiányzó érték, elszáll a program!)
4. Az így kapott súlyokkal (és csak a súlyokkal!) egészítse ki a MINTA állományt (SORT...,MERGE... BY TEA_CSOP.....). Ugyancsak ebben a lépésben teljeskörűsítse az ARBEVETEL változót (SULY*ARBEVETEL).
5. Mutassa ki a teljeskörűsített árbevételt TEAOR főcsoportonként az egyes létszám kategóriákban.
6. „Csinosítsa ki” az eredménytáblát, hogy publikus formát öltjön.

Pl.

Adjon táblájának címet.

A tábla értékeit három számjegyenként tagolt egész számként írassa ki.

Használja a következő kódleírásokat:

TEA_CSOP:

'A'='MEZŐGAZDASÁG, VADGAZDÁLKODÁS, ERDŐGAZDÁLKODÁS'
'B'='HALGAZDÁLKODÁS'
'C'='BÁNYÁSZAT'
'D'='FELDOLGOZÓIPAR'
'E'='VILLAMOSENERGIA-, GÁZ-, GŐZ-, VÍZELLÁTÁS'
'F'='ÉPÍTŐIPAR'
'G'='KERESKEDELEM, JAVÍTÁS'
'H'='SZÁLLÁSHELY-SZOLGÁLTATÁS, VENDÉGLÁTÁS'
'I'='SZÁLLÍTÁS, RAKTÁROZÁS, POSTA, TÁVKÖZLÉS'
'J'='PÉNZÜGYI KÖZVETÍTÉS'
'K'='INGATLANÜGYLETEK, GAZDASÁGI SZOLGÁLTATÁS'
'L'='KÖZIGAZGATÁS, VÉDELEM; KÖTELEZŐ TÁRSADALOMBIZTOSÍTÁS'
'M'='OKTATÁS'
'N'='EGÉSZSÉGÜGYI, SZOCIÁLIS ELLÁTÁS'
'O'='EGYÉB KÖZÖSSÉGI, SZEMÉLYI SZOLGÁLTATÁS'
'P'='HÁZTARTÁSOK TEVÉKENYSÉGE'
'Q'='TERÜLETEN KÍVÜLI SZERVEZET'
'X'='ISMERETLEN'

Mentse el az egész futtatott programot [sajátnév].SAS, a kapott eredményt pedig [sajátnév].HTML néven. Az elméleti kérdésekre adott válaszokat tartalmazó Word-t és ezt a két file-t küldje el a következő e-mail címre: csicsman@calculus.hu