

Przetwarzanie i analiza danych w systemie SAS®

Kolokwium nr 2 - 20.01.2010

- Zapisać plik `szablon.sas` pod nazwą `nazwisko.sas` (gdzie `nazwisko` to nazwisko piszącego kolokwium). Na początku pliku należy wpisać w komentarzu własne imię i nazwisko.
 - Rozwiązania zadań należy wpisywać do pliku `nazwisko.sas`. Plik należy we własnym interesie często zapisywać.
 - Rozpakowane pliki z danymi wejściowym mają być umieszczone w bibliotece o nazwie *KOLO*.
 - Wszelkie tworzone przez siebie zbiory sasowe należy umieszczać w bibliotece *WORK*.
 - Tworzone programy powinny działać poprawnie bez żadnych zmian dla dowolnych zbiorów o takiej samej strukturze (tj. o takich samych zmiennych i ich atrybutach), jak wymienione w treści zadań. W szczególności, rozwiązania będą testowane na zbiorach różnych od podanych.
 - Rozwiązania (tylko zawartość pliku `nazwisko.sas`, bez tworzonych zbiorów sasowych) należy przesłać na adres prowadzącego laboratorium. List ma nie zawierać żadnych załączników – innymi słowy kod program musi być bezpośrednio wklejony do listu. We własnym interesie należy poczekać na potwierdzenie odbioru.
 - Powodzenia!
-
1. Dane są trzy zbiory sasowe, o strukturze takiej jak **klienci**, **auta** i **wypozyczenia** (w szczególności zakłada się, że zmienne o nazwach zaczynających się od *id* jednoznacznie identyfikują klientów, samochody i wypożyczenia). Używając języka SQL znaleźć marki samochodów wypożyczonych choć raz przez klienta o nazwisku zaczynającym się od litery „N”.
 2. Zakładamy, że w pewnej bibliotece są zbiory sasowe o co najmniej dwóch wspólnych zmiennych numerycznych. Napisać makro z jednym parametrem *bib*, które stworzy zbiór zawierający wszystkie wspólne zmienne numeryczne (wraz ze wszystkimi ich wartościami ze wszystkich zbiorów), a następnie na jego podstawie wypisze do okienka Log wartość kwartyła dolnego tej zmiennej (tych zmiennych), która ma najmniejszy kwartył dolny spośród wszystkich wspólnych zmiennych.
 3. Powiedzmy, że dany jest zbiór sasowy z dwoma zmiennymi numerycznymi x i y i liczba naturalna n (dana np. w postaci makroziennej globalnej). Niech x_{min} i x_{max} oznaczają odpowiednio najmniejszą i największą wartość x w zbiorze. Używając jak najmniejszej liczby kroków (dopuszczalne są zarówno DATA, jak i PROC Step), znaleźć grupę bądź grupy wyznaczone przez zmienną x , w której suma różnych wartości zmiennej y jest największa (granice grup powstają z podziału odcinka $[x_{min}, x_{max}]$ na n równych części; grupy numerowane są liczbami od 1 do n).
 4. Za pomocą procedury IML znaleźć w zadanej w postaci zbioru sasowego macierzy kwadratowej jej maksymalny minor główny.