

Kolokwium nr 2 - 28.01.2013

- 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 kolokwium, tzn.:

K.Szpojankowski@mini.pw.edu.pl lub matysiak@mini.pw.edu.pl

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. (15 pkt.) Załóżmy, że w pewnej bibliotece *bibl* znajduje się niezerowa liczba zbiorów, które zawierają zmienną numeryczną o nazwie *data* (być może w *bibl* znajdują się także zbiory bez zmiennej *data*). Każdy ze zbiorów ze zmienną *data* niekoniecznie jest posortowany rosnąco względem tej zmiennej, ale nie zawiera powtarzających się wartości *data*. Można zakładać, że każdy z tych zbiorów, który zawiera *data*, ma w sumie dwie zmienne i że dodatkowa zmienna jest numeryczna (nie zakłada się znajomości nazwy tej zmiennej). Napisać makro z pojedynczym parametrem *bibl*, które stworzy zbiór *razem*, ze zmienną *data* oraz zmiennymi numerycznymi o nazwach takich, jak nazwy zbiorów, z których pochodzą. Wartości zmiennej *data* w *razem* mają być wszystkimi wartościami zmiennej *data* w rozważanych zbiorach i mają się nie powtarzać w *razem*. Wartości pozostałych zmiennych mają być równe odpowiednim wartościom dodatkowej zmiennej w odpowiednich zbiorach. Przykładowo, jeśli w danej bibliotece znajdowałyby się zbiory **w1**, **ke3** i **rt8**, to makro powinno stworzyć zbiór **razem** (patrz spakowany plik).
2. (5 pkt.) Załóżmy, że dane w postaci dwóch zbiorów sasowych o nazwach *m1* i *m2* są dwie macierze kwadratowe wymiaru 3x3. Za pomocą PROC IML stworzyć macierz będącą iloczynem danych macierzy i zapisać ją do zbioru sasowego *m1m2*.
3. (10 pkt.) Dane są trzy zbiory sasowe, o strukturze takiej jak **klienci**, **auta** i **wypozyczenia** – patrz spakowany plik. 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” (zarówno wielkiej, jak i małej, bo może się zdarzyć, że w bazie danych nazwiska są błędnie zapisane małymi literami).
4. (10 pkt.) Załóżmy, że dany jest zbiór zawierający jedynie zmienne tekstowe (ich liczba i nazwy nie są znane; wiadomo, że zmiennych jest więcej niż dwie). Napisać program, składający się z jak najmniejszej liczby kroków języka 4GL, który dla każdej obserwacji z danego zbioru wypisze do okienka Log nazwy tych dwóch zmiennych, które zawierają dwa najdłuższe (liczące najwięcej liter) napisy (dla danej obserwacji). Przyjmujemy założenie, że w każdej obserwacji są dokładnie dwie zmienne zawierające dwa najdłuższe napisy; napisy te nie muszą, ale mogą, mieć tyle samo liter. Do testowania można użyć zbioru **tekst** (patrz spakowany plik).