

- 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, tzn.:

A.Sakowicz@mini.pw.edu.pl lub 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. (14pkt) Załóżmy, że w bibliotece o nazwie `bibl` znajduje się pewna liczba zbiorów sasowych ze zmienną numeryczną `zmienna`. Napisać makro o parametrach `bibl` i `zmienna`, które stworzy zbiór sasowy zawierający (tylko) zmienną tekstową `zbiór`, której wartościami będą nazwy tych zbiorów sasowych z biblioteki `bibl`, w których w zmiennej `zmienna` nie ma braków danych.
2. (10pkt) Dane są trzy zbiory sasowe o strukturze takiej jak `KOLO.klienci`, `KOLO.auta` i `KOLO.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 nazwiska tych klientów, którzy wypożyczali samochody tylko jednej marki.
3. (8pkt) Załóżmy, że dane są zbiory sasowe: **dane** (z grupującą zmienną tekstową  $x$  i pewną zmienną numeryczną  $y$ ) i **ktore** (ze zmienną  $x$  taką, jak w zbiorze **dane** oraz z numeryczną zmienną  $k$ ). Zakładamy, że zbiór wartości zmiennej  $x$  ze zbioru **ktore** jest podzbiorem (właściwym albo niewłaściwym) zbioru wartości zmiennej  $x$  ze zbioru **dane**. Zakładamy także, że obydwa zbiory są posortowane rosnąco względem zmiennej  $x$  oraz że w zbiorze **ktore** wartości zmiennej  $x$  występują bez powtórzeń. Napisać jeden step (albo DATA step albo PROC step), który obliczy i wypisze do okienka Log średnią arytmetyczną tych wartości zmiennej  $y$  ze zbioru **dane**, które są na  $k$ -tej pozycji w grupie wyznaczonej przez zmienną  $x$ . Przykładowo, dla zbioru **dane** postaci

$x$	$y$
A	5
A	2
A	0
B	1
B	8
C	1
C	3

oraz zbioru **ktore** postaci

$x$	$k$
A	3
C	2

, szukana średnia wynosi  $\frac{0+3}{2} = 1.5$ .

4. (8pkt) Załóżmy, że macierz kwadratowa  $A \in \mathbb{R}^{5 \times 5}$  jest dana jako zbiór sasowy **WORK.MACIERZ**. Napisać program w języku IML, który obliczy i wypisze do zbioru sasowego **WORK.NORMY** (z jedną zmienną `norma` i trzema obserwacjami) normy macierzowe  $A$ :  $\|A\|_1$ ,  $\|A\|_2$  i  $\|A\|_\infty$ . Przypomnienie: jeśli

$$A = (a_{i,j})_{i,j=1}^n \in \mathbb{R}^n,$$

to

$$\begin{aligned} \|A\|_1 &= \max_j \sum_{i=1}^n |a_{i,j}|, \\ \|A\|_2 &= \max\{\sqrt{\lambda} : \lambda \text{ jest wartością własną } A^T A\}, \\ \|A\|_\infty &= \max_i \sum_{j=1}^n |a_{i,j}|. \end{aligned}$$