## Przetwarzanie i analiza danych w systemie SAS®

## Laboratorium nr 2

Zbiory danych potrzebne do rozwiązania poniższych zadań znajdują się na stronie WWW http://www.mini.pw.edu.pl/~matysiak

- 2.1 Wygenerować zbiór sasowy z dwiema zmiennymi numerycznymi x i y, o wartościach pochodzących z rozkładu jednostajnego na przedziale [-1,1]. Znaleźć średnie i maksymalne wartości zmiennych x i y. Utworzyć zbiór zawierający maksima i średnie zmiennych x i y w kolejnych obserwacjach.
- 2.2 Na podstawie zbioru **a** napisać program, który dla każdej wartości zmiennej id utworzy x kopii obserwacji ją zawierającej (patrz zbiór **b**).
- 2.3 Sklasyfikować wartości zmiennej *id* ze zbioru **rank** ze względu na malejące wartości zmiennej x. Jednakowym wartościom x należy przypisać tę samą wartość stworzonej przez siebie zmiennej klasyfikującej. Zakładamy, że **rank** jest posortowany malejąco ze względu na x.
- 2.4 Na podstawie zbioru cb znaleźć liczbę wzrostów i spadków wartości zmiennej cb.
- 2.5 Dany jest zbiór sasowy **zbior**. Napisać programy, w wyniku których w okienku Log pojawiają się komunikaty takie jak w plikach k1.txt, k2.txt i k3.txt. (Zmienne v3 i v4 są sumami zmiennych v1 i v2; nie można im bezpośrednio przypisać braków danych. Można korzystać z instrukcji SET, RETAIN i PUT.)
- 2.6 Na podstawie zbioru **cb** znaleźć liczbę lokalnych maksimów wartości zmiennej *cb*.
- 2.7 Zbiór sasowy **brak** zawiera zmienną numeryczną x, w której są izolowane braki danych (dodatkowo zakładamy, że w pierwszej i ostatniej obserwacji nie ma braków). Uzupełnij braki danych średnią arytmetyczną obserwacji poprzedniej i następnej w zbiorze.
- 2.8 Wartości ze zbioru **bezkropek** poprzeplatać losowa liczba braków danych (patrz zbiór **kropki**).
- 2.9 Wygenerować zbiór sasowy definiujący dystrybuantę rozkładu dyskretnego; następnie wczytując go odtworzyć rozkład i zapisać go w postaci zbioru sasowego.
- 2.10 Dla dwóch niezależnych trajektorii standardowych błądzeń losowych (startujących z 0) znaleźć
  - (a) ich maksymalne odchylenia od zera,
  - (b) liczbę wspólnych wierzchołków,
  - (c) liczbę przecięć.