Sebastian Smoliński

**Wydział Elektroniki i Technik Informacyjnych Politechnika Warszawska**

LAB WDEC 5 Sprawozdanie Zestaw danych: t\_dane\_2.txt

1. Utwórz bibliotekę o nazwie LIB5.

/\* Zadanie 1 \*/   
libname LIB5 "/folders/myfolders/Analiza";

1. Wczytaj plik tekstowy t\_dane\_[x].txt do pliku sas o nazwie t\_dane\_lib5 w bibliotece LIB5.

/\* import txt file – opcja nr 1 \*/   
**proc** **import** datafile =

folders/myfolders/Analiza/t\_dane\_2.txt"   
 out = LIB5.t\_dane\_lib5;

delimiter = ";"; getnames = NO;

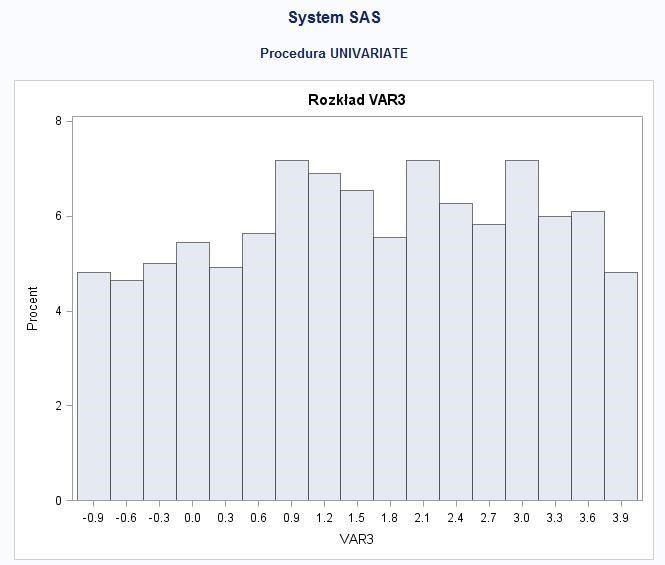
**run**;

/\* import txt file – opcja nr 2 \*/   
**data** LIB5.t\_dane\_lib5\_2;   
infile

"/folders/myfolders/Analiza/t\_dane\_2.txt"  
dlm = ";"; input VAR1 VAR2 $ VAR3 VAR4 VAR5; **run**;

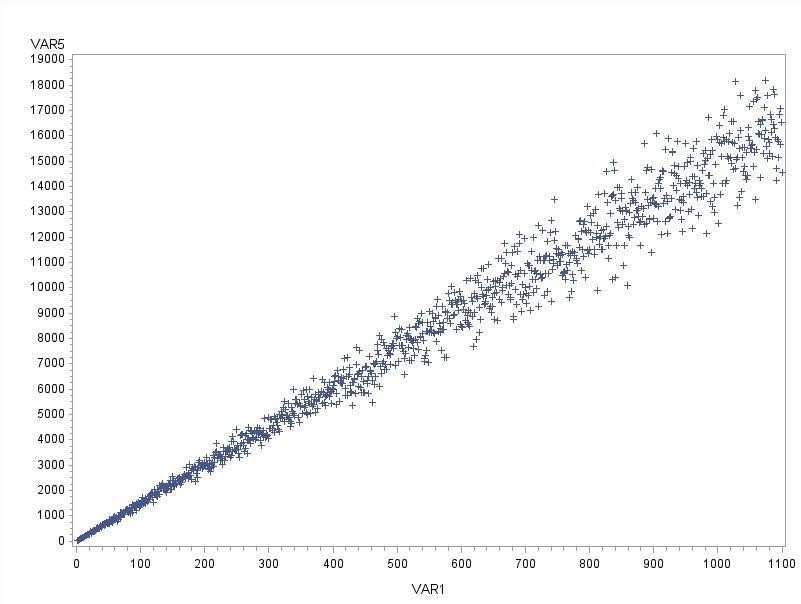
1. Narysuj histogramy dla zmiennych numerycznych.

/\* Zadanie 3 \*/   
**proc** **univariate** data = LIB5.t\_dane\_lib5; var VAR3; histogram; **run**;



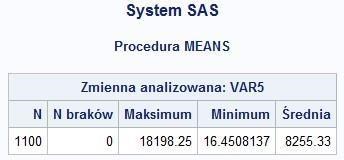
1. Narysuj wykresy zależności zmiennych numerycznych od id.

/\* Zadanie 4 \*/   
**proc** **sgplot**   
data = LIB5.t\_dane\_lib5;   
scatter x=VAR1 y=VAR5; **run**;



1. Policz za pomocą procedury MEANS statystyki: N, MAX, MIN, NMISS, MEAN.

/\* Zadanie 5 \*/   
**proc** **means** data = LIB5.t\_dane\_lib5 N NMISS MAX MIN MEAN; var VAR5; **run**;



1. Policz statystyki oddzielnie dla id parzystych i nieparzystych z wykorzystaniem CLASS i BY.

/\* Zadanie 6 z użyciem class \*/

**proc means** data=LIB5.t\_dane\_lib5 N NMISS MAX MIN MEAN;

var VAR5;class VAR2; **run**;



/\* Zadanie 6 z użyciem by \*/

**proc sort** data=LIB5.t\_dane\_lib5 out=LIB5.t\_dane\_lib5\_mean;

by VAR2;

**run**;

**proc means** data=LIB5.t\_dane\_lib5\_mean N NMISS MAX MIN MEAN;

by VAR2; var VAR5; **run**;



1. Znajdź parametry a i b modelu regresji y = a + b\*id dla id parzystych.

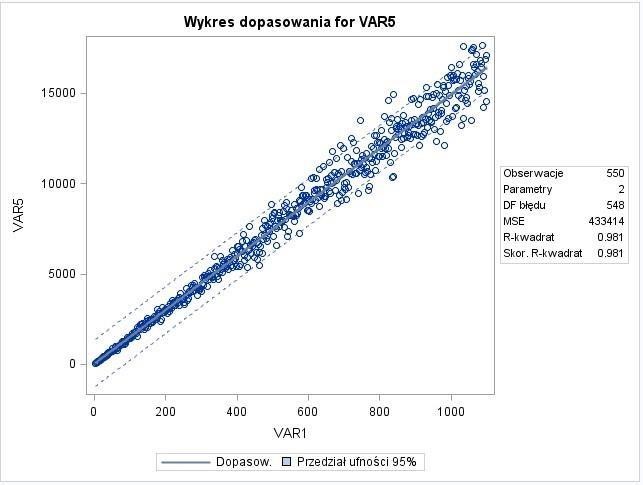
/\* Zad 7 \*/

**proc reg** data=LIB5.t\_dane\_lib5;

model VAR5=VAR1;

where (VAR2="s4\_parz");

**run**;   
Wynik: /\* a = 40.97727, b = 14.86533 \*/



1. Dodaj kolumnę o nazwie ye zawierającą obliczona wartość z modelu regresji dla id nieparzystych.

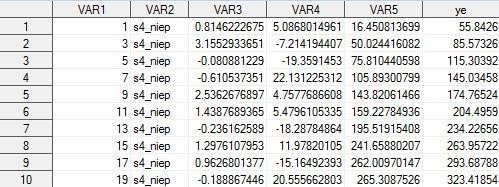
/\* Zadanie 8. \*/

**data** LIB5.reg;

set LIB5.t\_dane\_lib5;

ye=**40.97727**+ **14.86533** \* VAR1;

where (VAR2="s4\_niep");**run**;



9.

1. Narysuj zależność y i ye od id nieparzystych dla 10, 100 i wszystkich obserwacji (nieparzystych).

/\* Zadanie 9. \*/

**proc sgplot** data=LIB5.reg (obs=**10**);

/\*plot\*/

scatter x=VAR1 y=ye;

**run**;

**proc sgplot** data=LIB5.reg (obs=**10**);

scatter x=VAR1 y=VAR5;

**run**;

**proc sgplot** data=LIB5.reg (obs=**100**);

/\*plot\*/

scatter x=VAR1 y=ye;

**run**;

**proc sgplot** data=LIB5.reg (obs=**100**);

scatter x=VAR1 y=VAR5;

**run**;

**proc sgplot** data=LIB5.reg;

/\*plot\*/

scatter x=VAR1 y=ye;

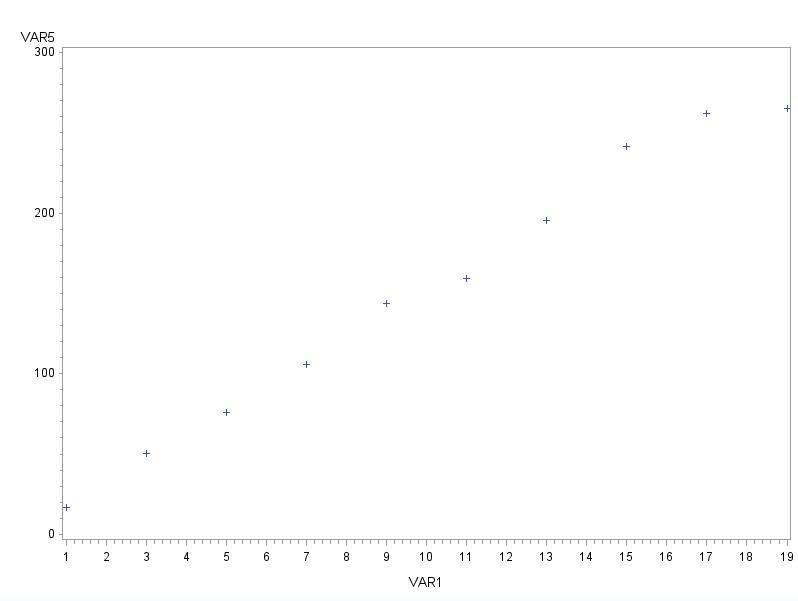
**run**;

**proc sgplot** data=LIB5.reg;

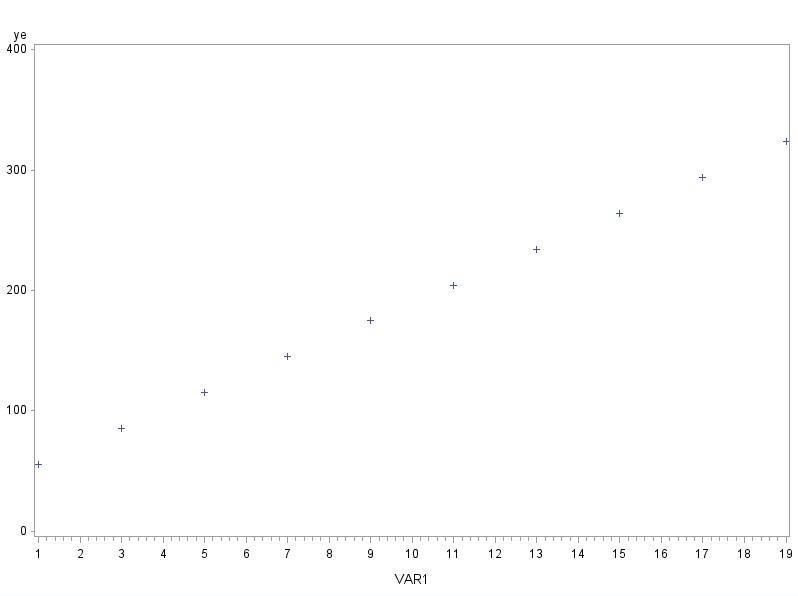
scatter x=VAR1 y=VAR5;

**run**;

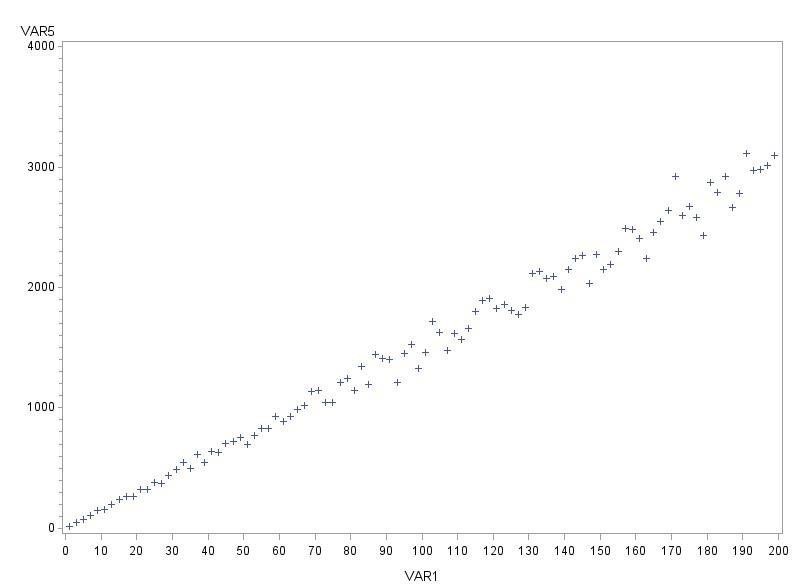
Wykres 1. Zależność VAR5 od id dla 10 obserwacji.



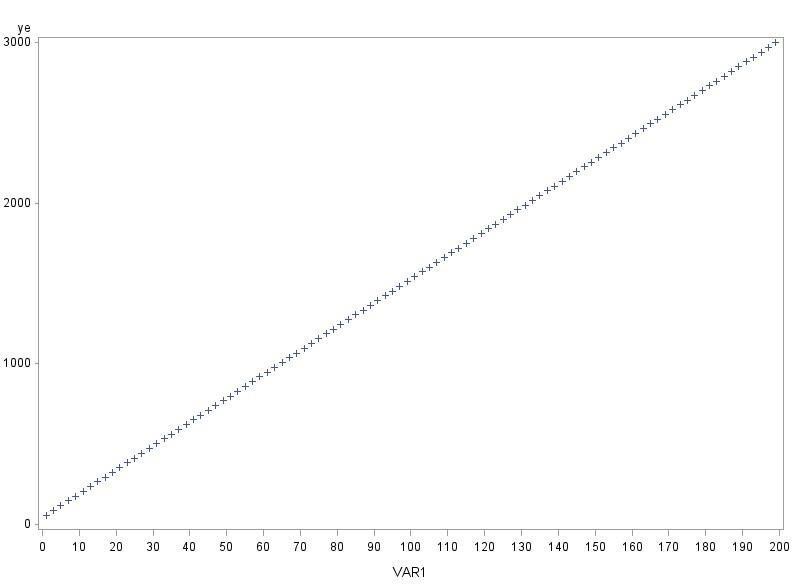
Wykres 2. Zależność ye od id dla 10 obserwacji.



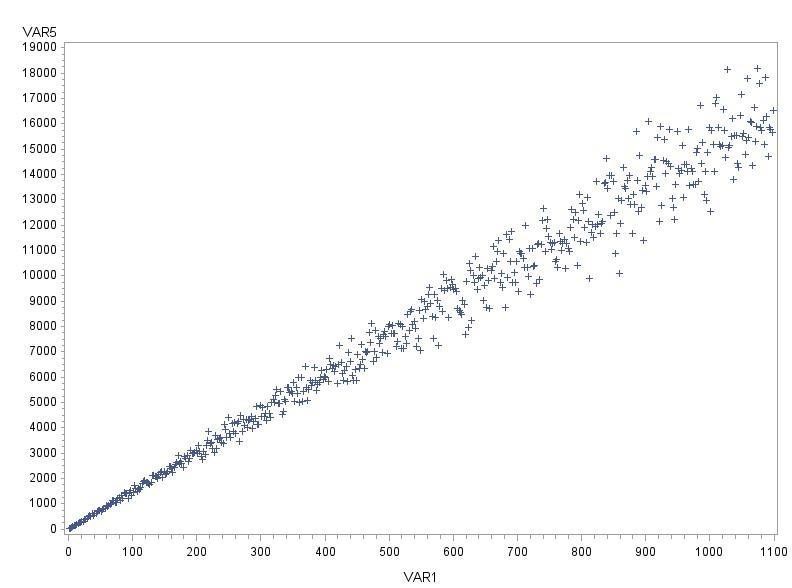
Wykres 3. Zależność VAR5 od id dla 100 obserwacji.



Wykres 4. Zależność ye od id dla 100 obserwacji.



Wykres 5. Zależność VAR5 od id dla wszystkich obserwacji.



Wykres 6. Zależność ye od id dla wszystkich obserwacji.

