

Wydział Elektroniki i Technik Informacyjnych Politechnika Warszawska

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

1. Utwórz bibliotekę o nazwie LIB5.

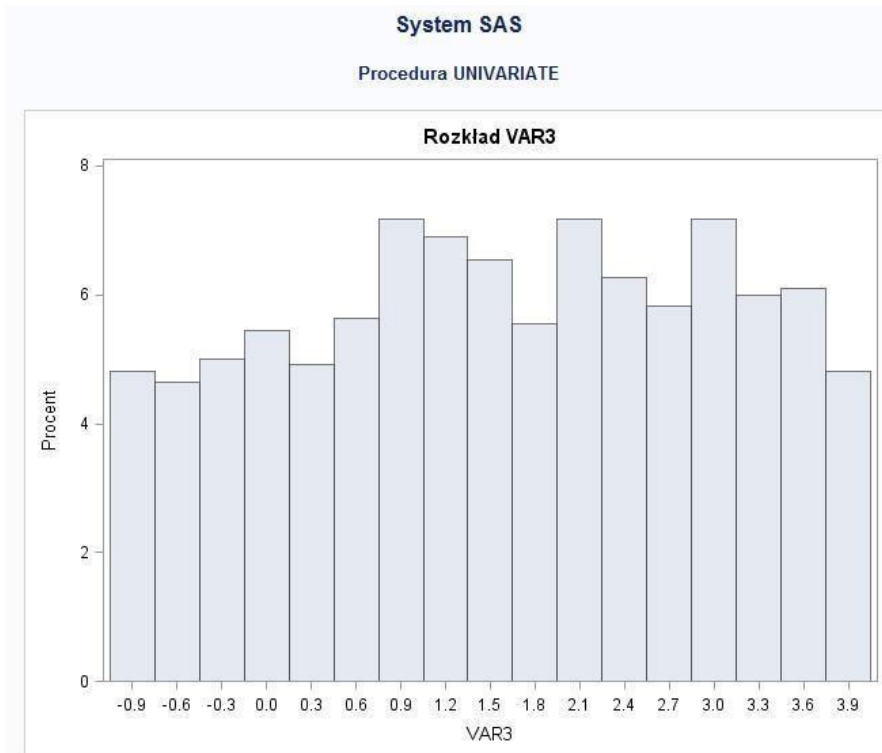
```
/* Zadanie 1 */  
libname LIB5 "/folders/myfolders/Analiza";
```

2. 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"  
dlim = ";"; input VAR1 VAR2 $ VAR3 VAR4 VAR5; run;
```

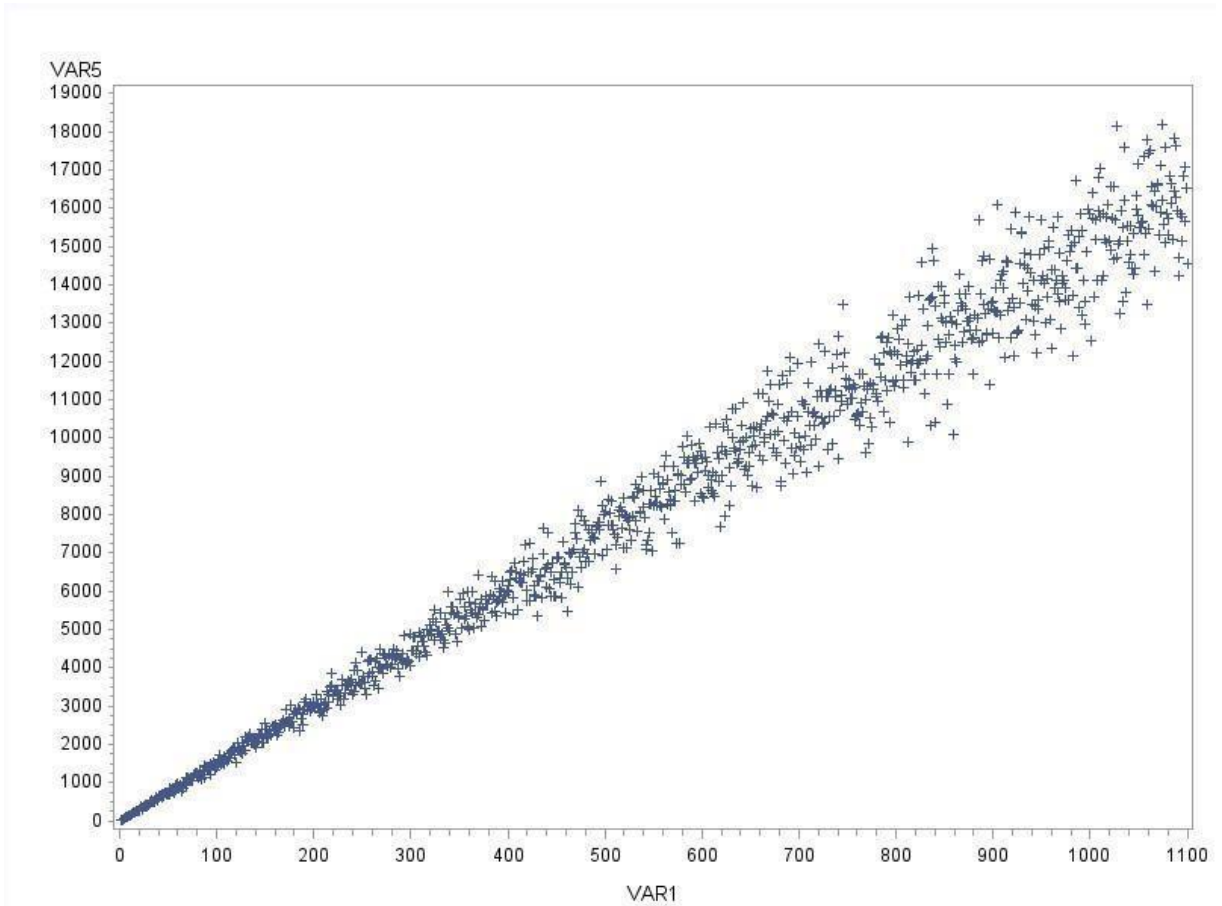
3. Narysuj histogramy dla zmiennych numerycznych.

```
/* Zadanie 3 */  
proc univariate data = LIB5.t_dane_lib5; var VAR3; histogram; run;
```



4. Narysuj wykresy zależności zmiennych numerycznych od id.

```
/* Zadanie 4 */  
proc sgplot  
data = LIB5.t_dane_lib5;  
scatter x=VAR1 y=VAR5; run;
```



5. 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;
```

System SAS				
Procedura MEANS				
Zmienna analizowana: VAR5				
N	N braków	Maksimum	Minimum	Średnia
1100	0	18198.25	16.4508137	8255.33

6. 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;
```

## System SAS

### Procedura MEANS

Zmienna analizowana: VAR5						
VAR2	N obs.	N	N braków	Maksimum	Minimum	Średnia
s4_niep	550	550	0	18198.25	16.4508137	8278.88
s4_parz	550	550	0	17625.39	31.1938689	8231.77

/\* 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;
```

## System SAS

### Procedura MEANS

VAR2=s4\_niep

Zmienna analizowana: VAR5				
N	N braków	Maksimum	Minimum	Średnia
550	0	18198.25	16.4508137	8278.88

VAR2=s4\_parz

Zmienna analizowana: VAR5				
N	N braków	Maksimum	Minimum	Średnia
550	0	17625.39	31.1938689	8231.77

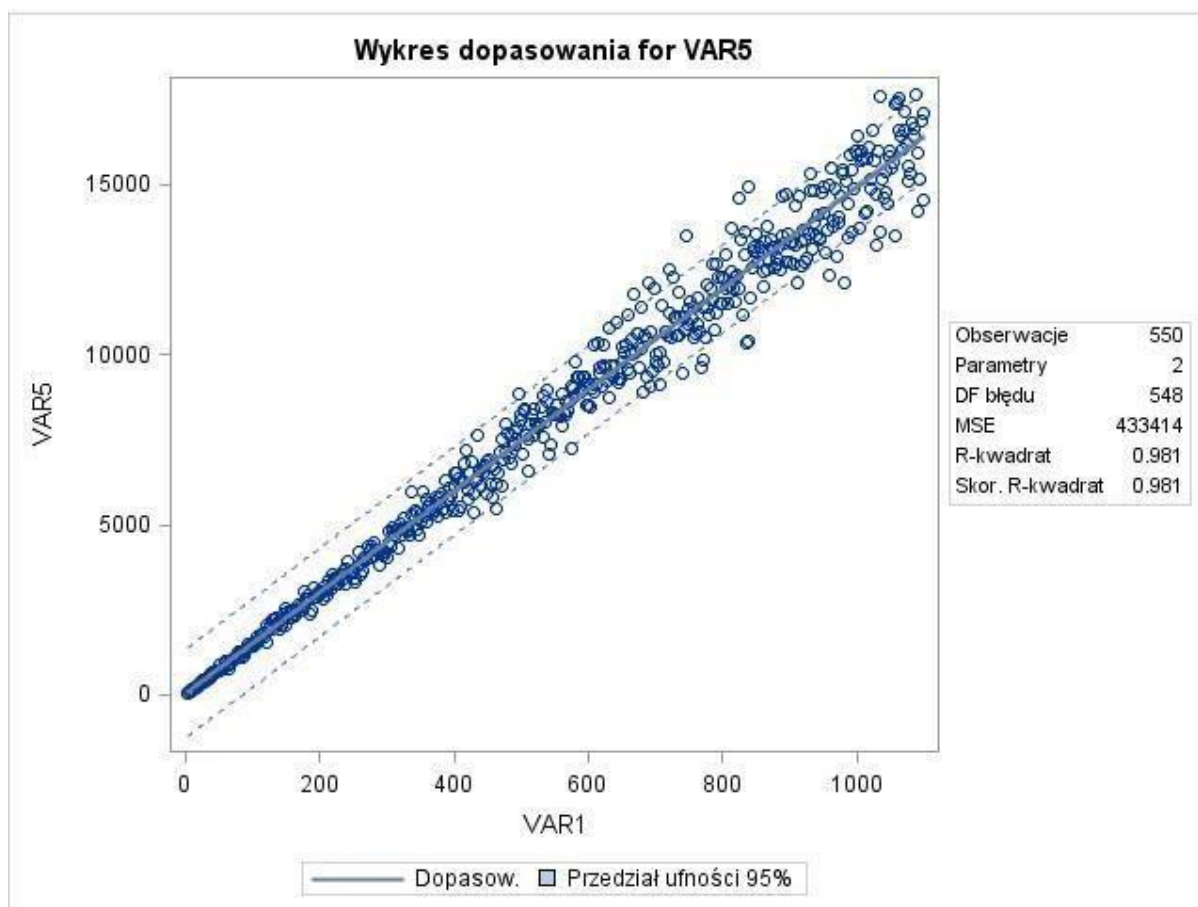
7. Znajdź parametry a i b modelu regresji  $y = a + b \cdot id$  dla id parzystych.

/\* Zad 7 \*/

```
proc reg data=LIB5.t_dane_lib5;  
model VAR5=VAR1;  
where (VAR2="s4_parz");  
run;
```

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

Oceny parametrów					
Zmienna	DF	Ocena parametru	Błąd standardowy	Wartość t	Pr. >  t
Intercept	1	40.97727	56.22020	0.73	0.4664
VAR1	1	14.86533	0.08840	168.15	<.0001



8. Dodaj kolumnę o nazwie ye zawierającą obliczoną 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**;

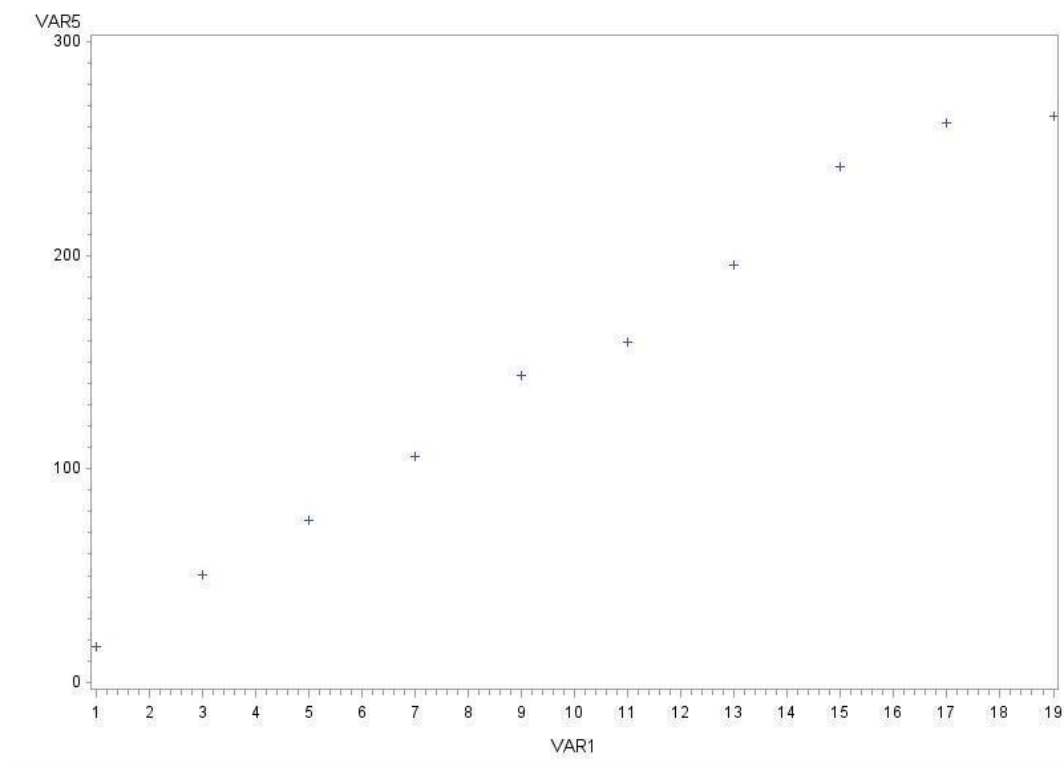
	VAR1	VAR2	VAR3	VAR4	VAR5	ye
1	1	s4_niep	0.8146222675	5.0868014961	16.450813699	55.8426
2	3	s4_niep	3.1552933651	-7.214194407	50.024416082	85.57326
3	5	s4_niep	-0.080881229	-19.3591453	75.810440598	115.30392
4	7	s4_niep	-0.610537351	22.131225312	105.89300799	145.03458
5	9	s4_niep	2.5362676897	4.7577686608	143.82061466	174.76524
6	11	s4_niep	1.4387689365	5.4796105335	159.22784936	204.4959
7	13	s4_niep	-0.236162589	-18.28784864	195.51915408	234.22656
8	15	s4_niep	1.2976107953	11.97820105	241.65880207	263.95722
9	17	s4_niep	0.9626801377	-15.16492393	262.00970147	293.68788
10	19	s4_niep	-0.188867446	20.555662803	265.3087526	323.41854

9.

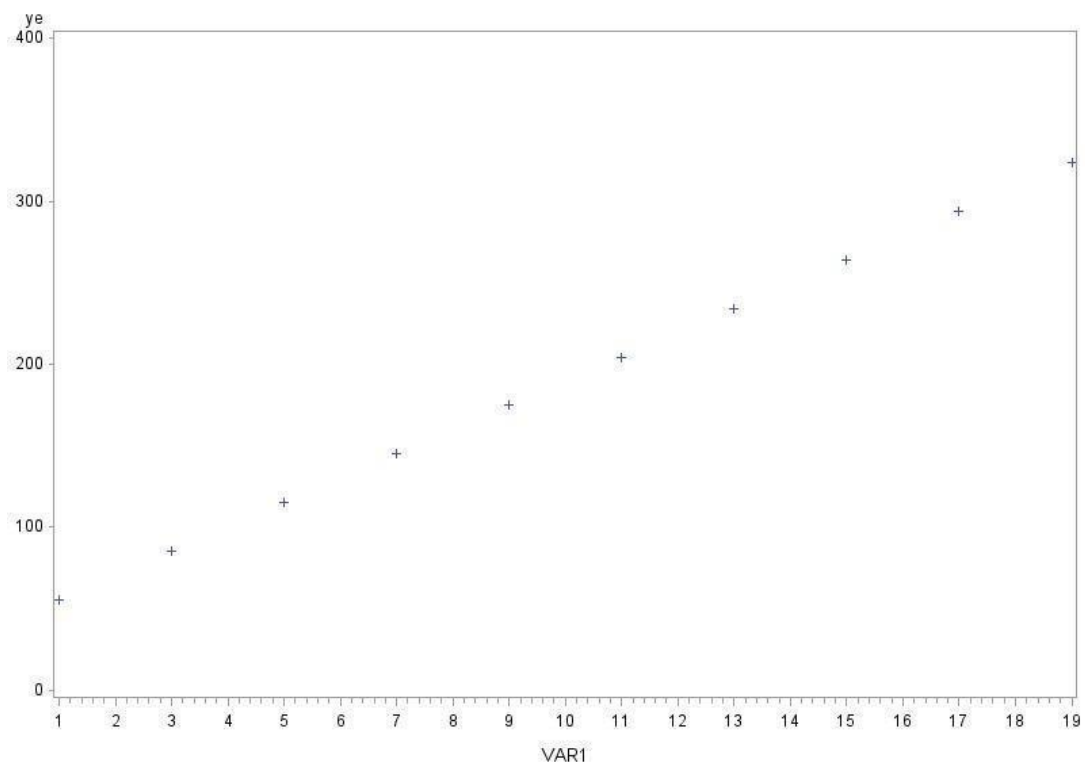
9. Narysuj zależność y i ye od id nieparzystych dla 10, 100 i wszystkich obserwacji (nieparzystych).

```
/* Zadanie 9. */  
proc sgplot data=LIB5.reg (obs=10);  
/*plot*/  
scatter x=VAR1 y=ye;  
run;  
proc sgplot data=LIB5.reg (obs=100);  
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;
```

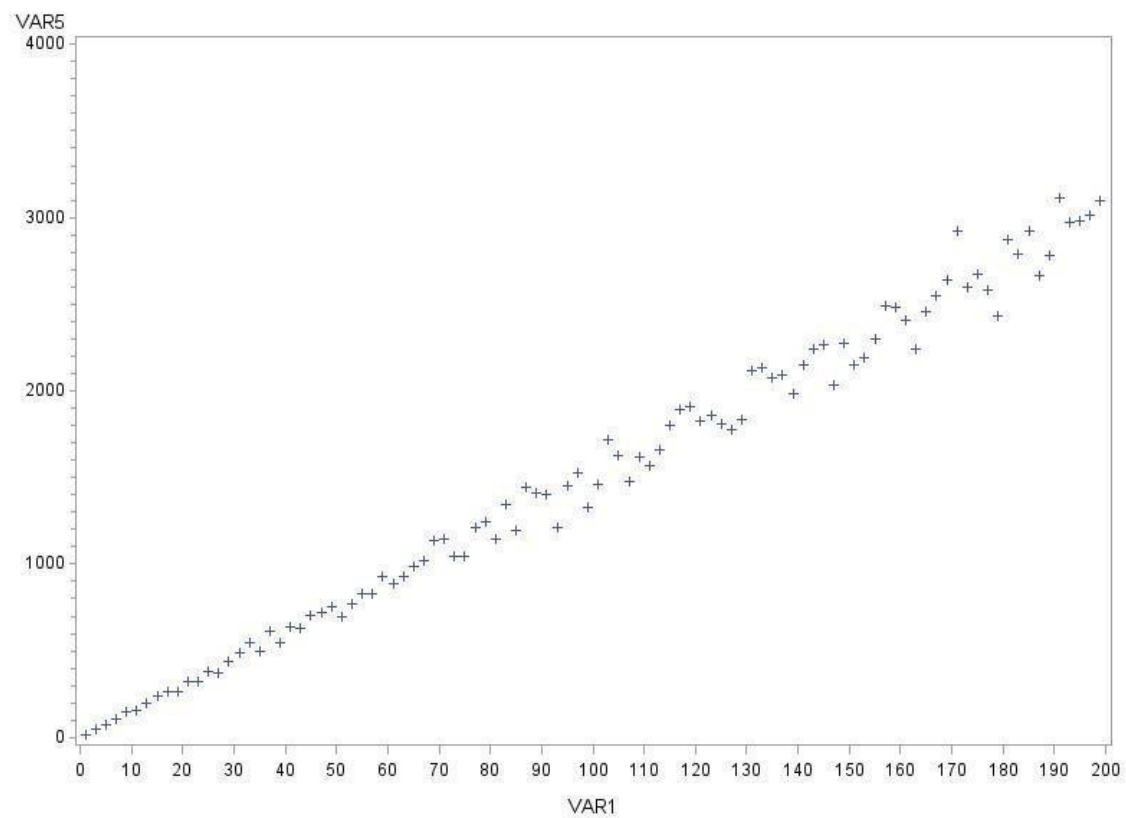
Wykres 1. Zależność VAR5 od id dla 10 obserwacji.



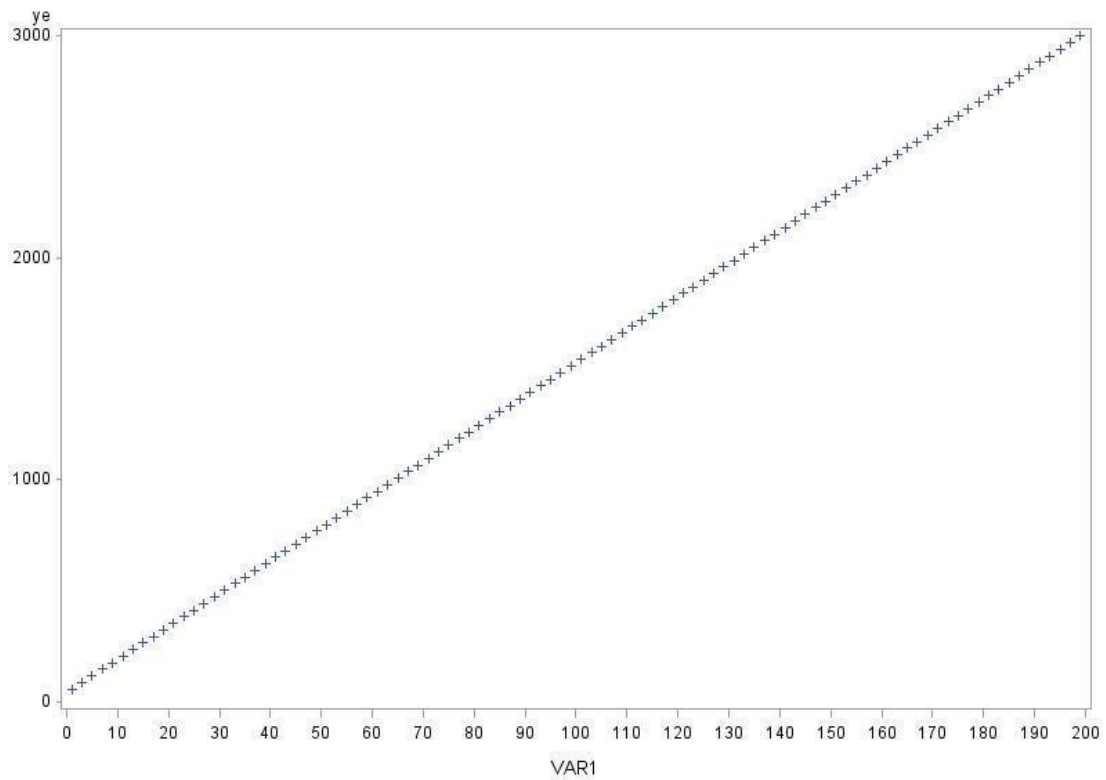
Wykres 2. Zależność  $y_e$  od  $id$  dla 10 obserwacji.



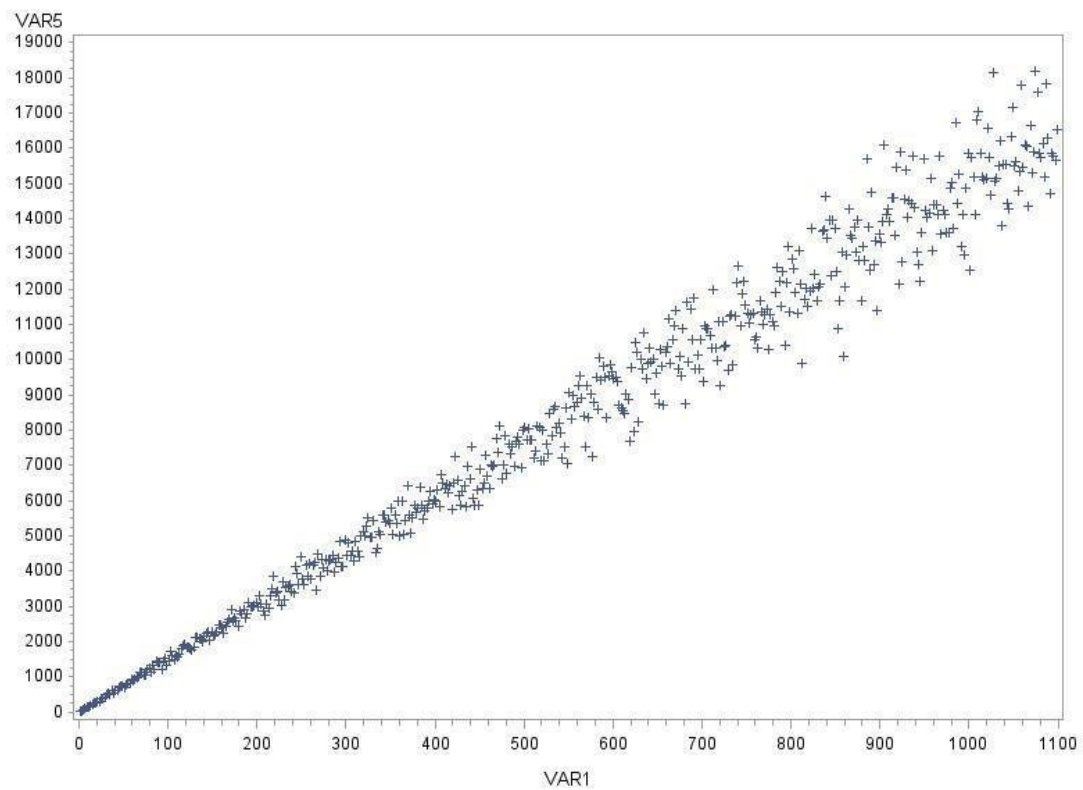
Wykres 3. Zależność VAR5 od  $id$  dla 100 obserwacji.



Wykres 4. Zależność ye od id dla 100 obserwacji.



Wykres 5. Zależność VAR5 od id dla wszystkich obserwacji.



Wykres 6. Zależność  $y_e$  od  $id$  dla wszystkich obserwacji.

