

**UNIwersytet Ekonomiczny w Katowicach**

**Przedmiot:**

Przetwarzanie, analiza i wizualizacja danych w systemie SAS

**Temat:**

Analiza Noblistów na przestrzeni lat 1901-2016

**Prowadzący:**

dr Żytniewski Mariusz

**Anna Krzyżowska – 139503**

**Szkoła Studiów I Stopnia, Stacjonarne,**

**Informatyka i Ekonometria,**

**Analityka Danych,**

**Rok 2, semestr 3**

**Katowice 2020**

## Spis treści

<b>Wstęp .....</b>	<b>3</b>
<i>Opis zmiennych.....</i>	<i>4</i>
<b>Analiza kategorii.....</b>	<b>6</b>
<b>Analiza Noblistów ze względu na płeć .....</b>	<b>8</b>
<i>Liczba kobiet i mężczyzn, którzy zdobyli Nobla.....</i>	<i>8</i>
<i>Laureatki Nagrody Nobla.....</i>	<i>9</i>
Pierwsze Noblistki.....	10
<i>Nobliści.....</i>	<i>12</i>
<i>Porównanie udziału kobiet i mężczyzn w poszczególnych kategoriach .....</i>	<i>13</i>
<b>Analiza podziału nagród .....</b>	<b>14</b>
<i>Dzielenie Nagrody .....</i>	<i>14</i>
<i>Potrójna nagroda .....</i>	<i>15</i>
<b>Analiza czasowa .....</b>	<b>17</b>
<i>Korelacja wieku do roku zdobycia nagrody .....</i>	<i>17</i>
<i>Średnia wieku, a kategoria .....</i>	<i>19</i>
<i>Ciągłość przyznawania nagród .....</i>	<i>20</i>
<i>Analiza wieku.....</i>	<i>21</i>
<i>Długość życia.....</i>	<i>21</i>
Ile lat żyli nobliści po otrzymaniu nagrody? .....	23
<b>Analiza ze względu na kraj urodzenia.....</b>	<b>26</b>
<i>W Stanach.....</i>	<i>27</i>
<i>Pokojowa Nagroda Nobla .....</i>	<i>28</i>
<b>Analiza argumentacji .....</b>	<b>30</b>
<b>Wnioski.....</b>	<b>33</b>
<b>Spis ilustracji.....</b>	<b>34</b>
<b>Bibliografia.....</b>	<b>35</b>

## Wstęp

Nagroda Nobla jak powszechnie wiadomo jest przyznawana co roku za wybitne osiągnięcia w trzech dziedzinach a mianowicie za osiągnięcia: naukowe, literackie oraz za zasługi dla społeczeństwa. Fundatorem nagrody jest oczywiście Alfred Nobel, szwedzki przemysłowiec, który wynalazł dynamit. Pierwszy raz dnia 10 grudnia 1901r. w Sztokholmie, ww. Pokojową Nagrodę otrzymali Jean Henri Dunant, oraz Frédéric Passy, za założenie ruchu humanitarnego o nazwie Czerwony Krzyż.

Laureatem owej nagrody może zostać człowiek jak i organizacja. Słowo "laureat" podkreśla znaczenie otrzymania tegoż wyróżnienia, ponieważ w starożytnej Grecji wieńce laurowe przyznawano ludziom jako znak honoru i zwycięstwa.

Nagroda Nobla jest najbardziej znaczącą i prestiżową nagrodą na świecie, ze względu na rygorystyczne zasady jej przyznawania. Sprawia to, że nagrodę mogą otrzymać tylko nieliczni, zasłużeni swoimi niesamowitymi osiągnięciami.

Ten zestaw danych zawiera zapis dla każdej osoby lub organizacji, która otrzymała Nagrodę Nobla od 1901 do 2016 roku. Poniższa analiza została wykonana w oprogramowaniu SAS Base. Baza danych została pobrana ze strony kaggle.com - <https://www.kaggle.com/nobelfoundation/nobel-laureates?fbclid=IwAR2rdSMtiqCnyMsGtAYR6GF5cePIBKoXs24CjoVjEcHUFGAffPzWKnZzh1o>.

Cel główny: Jakie czynniki muszą być spełnione, aby zwiększyć szansę na otrzymania Nagrody Nobla.

Cele szczegółowe:

- Analiza kategorii
- Analiza płci
- Analiza podziału nagród
- Analiza czasowa
- Analiza kraju urodzenia

## Import zmiennych

Analizę zaczniemy od zaimportowania danych. Bazę wprowadzamy do systemu za pomocą poniższego kodu:

```
proc import out=sas.tab1
  datafile='C:\Users\Ania\Desktop\sas\archive.csv'
  dbms=csv replace;
  delimiter=',';
  getnames=yes;
  datarow=2;
run;
```

Rysunek 1 Import bazy

Podgląd części zaimportowanej tabeli:

	Year	Category	Prize	Motivation	Prize_Share	Laureate_ID	Laureate_Type	Full_Name
1	1901	Chemistry	The Nobel Prize in Chemistry 1901	"In recognition of the extraordinary services he has rendered by the discovery of the laws of chemical dynamics and osmotic pressure in solutions"	1/1	160	Individual	Jacobus Henricus van t Hoff
2	1901	Literature	The Nobel Prize in Literature 1901	"In special recognition of his poetic composition, which gives evidence of lofty idealism, artistic perfection and a rare combination of the qualities of both heart and intellect"	1/1	569	Individual	Sully Prudhomme
				"For his work on serum therapy, especially its application against				

Rysunek 2 Baza

Za pomocą poniższego polecenia możemy sprawdzić liczbę obserwacji i zmiennych

```
proc contents data=Sas.Tab1;
run;
```

Rysunek 3 Procedura zawartości

<b>Obserwacje</b>	<b>969</b>
<b>Zmienne</b>	<b>18</b>

Rysunek 4 Wynik zawartości

W tabeli głównej znajduje się 969 obserwacji i 18 zmiennych.

## Opis zmiennych

Za pomocą tego samego polecenia otrzymujemy również alfabetyczną listę zmiennych i atrybutów.

Alfabetyczna lista zmiennych i atrybutów					
#	Zmienna	Typ	Dług.	Format	Informat
10	Birth_City	Znak	18	\$18.	\$18.
11	Birth_Country	Znak	23	\$23.	\$23.
9	Birth_Date	Num	8	YYMMDD10.	YYMMDD10.
2	Category	Znak	10	\$10.	\$10.
17	Death_City	Znak	14	\$14.	\$14.
18	Death_Country	Znak	14	\$14.	\$14.
16	Death_Date	Num	8	YYMMDD10.	YYMMDD10.
8	Full_Name	Znak	34	\$34.	\$34.
6	Laureate_ID	Num	8	BEST12.	BEST32.
7	Laureate_Type	Znak	10	\$10.	\$10.
4	Motivation	Znak	246	\$246.	\$246.
14	Organization_City	Znak	10	\$10.	\$10.
15	Organization_Country	Znak	14	\$14.	\$14.
13	Organization_Name	Znak	111	\$111.	\$111.
3	Prize	Znak	46	\$46.	\$46.
5	Prize_Share	Znak	3	\$3.	\$3.
12	Sex	Znak	4	\$4.	\$4.
1	Year	Num	8	BEST12.	BEST32.

*Rysunek 5 Alfabetyczna lista zmiennych i atrybutów*

Uważam, że łatwiej jest przeprowadzać analizę na angielskich nazwach, ponieważ polskie nie są aż tak dokładne, dlatego poniżej napiszę ich tłumaczenie:

- Birth\_City=Miejsce urodzenia
- Birth\_Country=Kraj urodzenia
- Birth\_Date>Data urodzenia
- Category=Kategoria
- Death\_City=Miejsce śmierci
- Death\_Country=Kraj śmierci
- Death\_Date>Data śmierci
- Full\_Name=Pełna nazwa
- Laureate\_ID=ID laureata
- Laureate\_Type=Typ laureata
- Motivation=Argumentacja
- Organization\_City=Miejsce Organizacji
- Organization\_Country=Kraj Organizacji
- Organization\_Name=Nazwa Organizacji
- Prize=Nagroda
- Prize\_Share=Dzielenie nagrody
- Sex=Płeć
- Year=Rok

## Analiza Kategorii

Nagrody Nobla są przyznawane corocznie i w następujących dziedzinach:

- fizyki, przez Królewską Szwedzką Akademię Nauk, za „najważniejsze odkrycie lub wynalazek w dziedzinie fizyki”  
chemii, przez Królewską Szwedzką Akademię Nauk, za „najważniejsze odkrycie lub postęp w dziedzinie chemii”
- fizjologii lub medycyny, przez Instytut Karolinska, za „najważniejsze odkrycie w dziedzinie fizjologii lub medycyny”
- literatury, przez Akademię Szwedzką, za „wybitną pracę na rzecz idealistycznych tendencji”
- Pokojowa Nagroda Nobla, przez Norweski Komitet Noblowski, za „najlepszą pracę na rzecz braterstwa między narodami, likwidacji lub redukcji stałych armii oraz za udział i promocję stowarzyszeń pokojowych”.

Od 1968 Szwedzki Bank Narodowy przyznaje w dziedzinie ekonomii nagrodę imienia Alfreda Nobla: Nagroda Banku Szwecji im. Alfreda Nobla w dziedzinie ekonomii.

Zliczamy liczbę przyznanych Nobli dla kategorii:

```
proc sql;  
select count(Category) as liczba, Category as Kategoria  
from Sas.Tab1  
group by Category  
order by liczba desc;  
quit;
```

Rysunek 6 Polecenie – lista przyznanych Nobli w kategoriach

Rezultat jest następujący:

liczba	Kategoria
227	Medicine
222	Physics
194	Chemistry
130	Peace
113	Literature
83	Economics

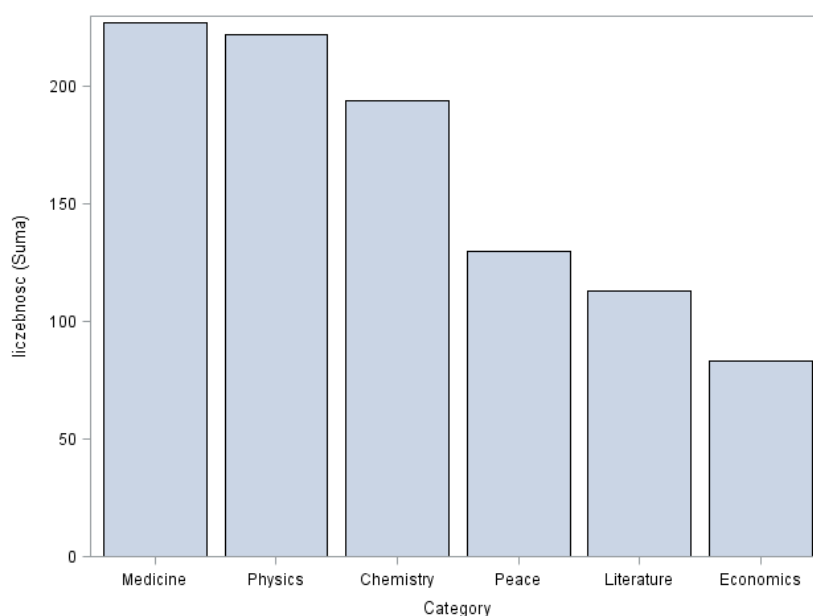
Rysunek 7 Wynik – lista przyznanych Nobli w kategoriach

Najwięcej Nagród zostało przyznanych w medycynie, następnie fizyce, chemii, pokoju i literatury. Ze względu na późniejsze pojawienie się dziedziny ekonomii jest ona ostatnia w tabelce. Wykres dla rozdanych nagród w zależności od kategorii:

```
proc sgplot data=sas.tab1;  
  vbar category / response = liczebnosc categoryorder = respdesc;  
run;
```

*Rysunek 8 Polecenie wykres słupkowy – liczba przyznanych nagród w kategoriach*

Rezultat:



*Rysunek 9 Wykres słupkowy – liczba przyznanych nagród w kategoriach*

Wyraźnie widać, że ekonomia jest wschodzącą kategorią w Nagrodach Nobla.

## Analiza Noblistów ze względu na płeć

W kolejnym etapie analizy warto przyjrzeć się podziału ze względu na płeć.

### *Liczba kobiet i mężczyzn, którzy zdobyli Nobla*

Badając liczbę kobiet i mężczyzn wykorzystujemy procedurę freq od frequency czyli częstotliwość. Tables w poniższym kodzie oznacza na której zmiennej chcemy się skupić. W naszym przypadku jest to płeć.

```
proc freq data=Sas.Tab1;  
tables Sex;  
run;
```

*Rysunek 10 Polecenie - częstotliwość występowania płci*

Sex	Liczebność	Procent	Liczebność skumulowana	Procent skumulowany
Fema	50	5.30	50	5.30
Male	893	94.70	943	100.00

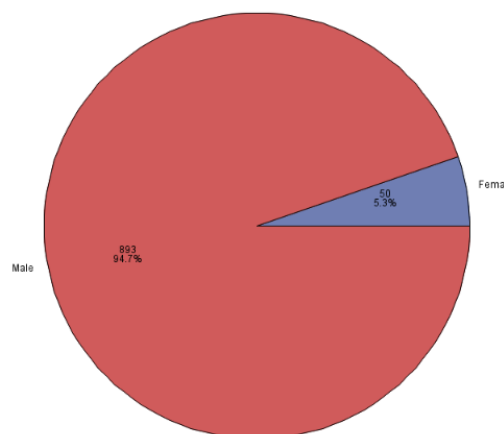
*Rysunek 11 Wynik - częstotliwość występowania płci*

Tylko 50 razy kobiety otrzymały Nobla w latach 1901-2016. Stanowią one tylko 5,30%. Z kolei liczba mężczyzn, których nagrodzono wynosi aż 893, czyli 94,70%. Zobaczmy, jak to przedstawia się na wykresie kołowym:

```
proc gchart data=Sas.tab1;  
pie Sex / DISCRETE VALUE=INSIDE  
PERCENT=INSIDE SLICE=OUTSIDE;  
run;
```

*Rysunek 12 Polecenie wykres kołowy - rozkład płci*





Rysunek 13 Wykres kołowy - rozkład płci

Wyraźnie widać dużą przewagę mężczyzn w stosunku do kobiet.

### *Laureatki Nagrody Nobla*

Wiele kobiet również zasłużyło na tę prestiżową nagrodę, jednak na początku XX wieku, nie były traktowane na równi z mężczyznami, co znacznie utrudniało im pracę na wysokim szczeblu stanowisku np. w laboratoriach.

Dobrym przykładem jest Lise Meitner (1878-1968), swego czasu studentka fizyki pochodzenia Austriackiego. Po ukończeniu studiów we Wiedniu, Lise przeniosła się do Berlina, gdzie prowadziła badania nad promieniotwórczością wraz z chemikiem Ottonem Hahnem. Ich laboratorium mieściło się w piwnicy Instytutu Chemii w Berlinie. Kobieta została tam zatrudniona w drodze wyjątku przez dyrekcję, która jednak postawiła jej takie ultimatum jak wchodzenie osobnym wejściem do laboratorium. Nie mogła także wchodzić na wyższe piętra budynku oraz musiała wychodzić do toalety do hotelu naprzeciw, żeby nie rzucać się w oczy.

W 1945r. Lise wraz z współpracownikiem dokonała rozszczepienia uranu, jednak nagrodę otrzymał tylko Otton Hahnem, ona sama została natomiast całkowicie pominięta, mimo znacznego wkładu w owe osiągnięcie.

W przeszłości było wiele przypadków pomijania kobiet jako naukowców w pracach badawczo-naukowych i ich osiągnięcia naukowe przypisywano mężczyznom. Ze względu na ilość takich przypadków przypisano im nazwę jako tzw. efekt Matyldy.

## Pierwsze Noblistki

Sprawdźmy w którym roku i w jakich dziedzinach kobiety otrzymały pierwsze Noble.

```
Proc sort data=Sas.Tab4;  
by Year ;  
where Sex='Fema';  
proc print;  
run;
```

Rysunek 14 Polecenie sortujące z wydrukiem

Obs.	Year	Category	Prize	Motivation	Prize_Share	Laureate_ID	Laureate_Type	Full_Name	Birth_Date	Birth_City	Birth_Country	Sex
1	1903	Physics	The Nobel Prize in Physics 1903	"in recognition of the extraordinary services they have rendered by their joint researches on the radiation phenomena discovered by Professor Henri Becquerel"	1/4	6	Individual	Marie Curie, née Skłodowska	1867-11-07	Warsaw	Russian Empire (Poland)	Fema
2	1905	Peace	The Nobel Peace Prize 1905		1/1	468	Individual	Baroness Bertha Sophie Felicitas von Suttner	1843-06-09	Prague	Austrian Empire (Czech)	Fema
3	1909	Literature	The Nobel Prize in Literature 1909	"in appreciation of the lofty idealism, vivid imagination and spiritual perception that characterize her writings"	1/1	579	Individual	Selma Ottilia Lovisa Lagerlöf	1858-11-20	Malmö	Sweden	Fema

Rysunek 15 Wynik polecenia sortującego - wydruk

Pierwszą nagrodę jako kobieta otrzymała Maria Curie Skłodowska (dzieląc nagrodę Prize\_Share=1/4) w Fizyce za badania nad zjawiskiem promieniotwórczości, następnie Pokojową Nagrodę Nobla w 1905 roku otrzymała Bertha von Suttner za osiągnięcia w Międzynarodowym Stałym Biurze Pokoju. Z kolei 3 kobieta to Selma Lagerlöf dostała Nobla w dziedzinie literatury „w hołdzie wielkiemu idealizmowi, błyskotliwej wyobraźni i duchowej przenikliwości, które wyróżniają jej utwory”.

Sprawdźmy, jak wygląda podział nagród przyznanych kobietom w zależności od kategorii:

```

proc sql;
  select count(Sex) as Kobiety, Category as Kategoria
  from Sas.Tab1
  where Sex="Fema"
  group by Category;
quit;

```

Rysunek 16 Polecenie podział nagród między kobiety – kategorie

Kobiety	Kategoria
4	Chemistry
2	Economics
14	Literature
12	Medicine
16	Peace
2	Physics

Rysunek 17 Wynik podział nagród między kobiety – kategorie

W Fizyce i Ekonomii zostały wyróżnione tylko po dwie kobiety, następnie cztery w Chemii, 12 w Medycynie, 14 w Literaturze i najwięcej w ‘Pokoju’ - 16.

Utwórzmy tabelę, gdzie będą znajdowały się tylko kobiety:

```

data Sas.Tab1;
  set Sas.Tab5;
  Where Sex='Fema';
run;

```

Rysunek 18 Polecenie tworzące tabelę - kobiety

Zobaczmy rozkład kategorii na wykresie kołowym

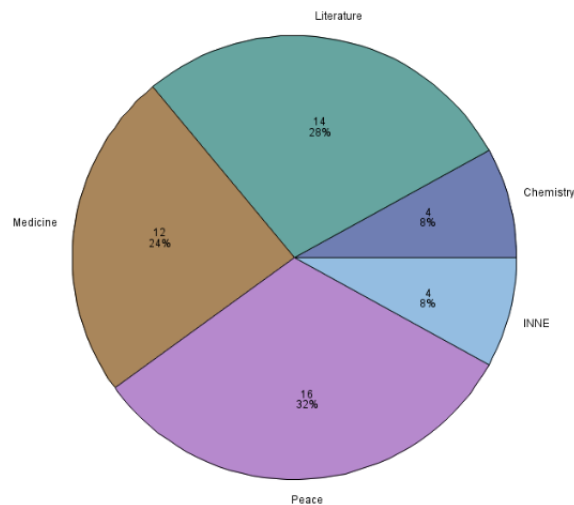
```

proc gchart data=Sas.Tab5;
  pie Category / DISCRETE VALUE=INSIDE
  PERCENT=INSIDE SLICE=OUTSIDE;
run;

```

Rysunek 19 Polecenie tworzące wykres kołowy udział procentowy kobiet w kategoriach

Rezultat:



Rysunek 20 Wykres kołowy udział procentowy kobiet w poszczególnych kategoriach

### Nobliści

Elitarne grono noblistów liczy niemal tysiąc osób i organizacji z czego większość stanowią mężczyźni. Sprawdźmy, jak rozkładają się nagrody, które dostali mężczyźni w zależności od kategorii.

```
proc sql;  
select count(Sex) as Mezczyzni, Category as Kategoria  
from Sas.Tab3  
where Sex='Male'  
group by Category;  
quit;
```

Rysunek 21 Polecenie - podział nagród pomiędzy mężczyznami

Mezczyzni	Kategoria
190	Chemistry
81	Economics
99	Literature
215	Medicine
88	Peace
220	Physics

Rysunek 22 Wynik polecenia - podział nagród pomiędzy mężczyznami

Najwięcej nagród przyznano mężczyznom w dziedzinie Fizyki - 220 i Medycyny - 215, troszkę mniej, bo 190 w Chemii, następnie 99 w Literaturze i kolejno 88 w 'Pokoju' i najmniej 81 w Ekonomii. Tworzymy tabelę, gdzie będą znajdować się tylko mężczyźni

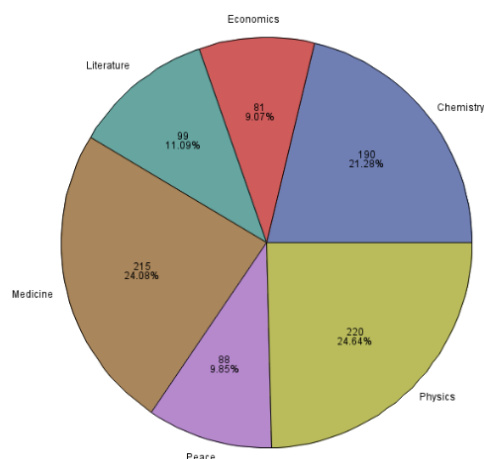
```
data Sas.Tab6;
set Sas.Tab1;
where Sex='Male';
run;
```

Rysunek 23 Polecenie tworzące tabelę – mężczyźni

Zobaczmy rozkład kategorii na wykresie kołowym:

```
proc gchart data=Sas.Tab6;
pie Category / DISCRETE VALUE=INSIDE
PERCENT=INSIDE SLICE=OUTSIDE;
run;
```

Rysunek 24 Polecenie tworzące wykres kołowy udział procentowy mężczyzn w kategoriach



Rysunek 25 Wykres kołowy udział procentowy mężczyzn w poszczególnych kategoriach

### Porównanie udziału kobiet i mężczyzn w poszczególnych kategoriach

Porównując wykres kołowy Rysunek 24 i Rysunek 25 można zauważyć, że procentowy udział mężczyzn jest większy niż kobiet w Fizyce i Chemii i Ekonomii, natomiast procentowy udział kobiet jest większy niż mężczyzn w dziedzinie Pokoju, jednak jest porównywalny udział procentowy w medycynie.

## Analiza podziału nagród

### Dzielenie Nagrody

Nagrody Nobla można udostępnić jednemu lub większej liczbie odbiorców lub organizacji za współpracę w tej samej dziedzinie czy temacie. Nagrody nie można podzielić między więcej niż trzy osoby, chociaż Pokojową Nagrodę Nobla można przyznać organizacjom złożonym z więcej niż trzech osób. Sortujemy tabelę po kolumnie Nagroda\_dzielona;

```
proc sort data=Sas.Tab1;  
  by Prize_Share;  
run;
```

Rysunek 26 Polecenie sortowania tabeli po dzielonych nagrodach

Zliczamy ilość nagród dzielonych:

```
data Sas.Tab10;  
  set Sas.Tab1;  
  retain suma_dzielonych;  
  if first.Prize_Share then suma_dzielonych=1;  
  else suma_dzielonych=suma_dzielonych+1;  
  by Prize_Share; if last.Prize_Share then output;  
  keep suma_dzielonych Prize_Share;  
run;
```

Rysunek 27 Polecenie zliczające ilość nagród dzielonych

Rezultat:

	Prize_Share	suma_dzielonych
1	1/1	351
2	1/2	328
3	1/3	225
4	1/4	64

Rysunek 28 Tabela - zliczona ilość nagród dzielonych

Trzeba wziąć pod uwagę, że program zliczył występujące pozycje. Jednak, żeby się dowiedzieć, ile nagród za jedną pracę zostało podzielonych trzeba podzielić przez liczbę członków zespołu:

```

data Sas.Tab11;
set sas.Tab10;
if Prize_Share='1/2' then liczba=suma_dzielonych/2;
if Prize_Share='1/3' then liczba=suma_dzielonych/3;
if Prize_Share='1/4' then liczba=suma_dzielonych/4;
run;

```

Rysunek 29 Polecenie tworzące tabele z pojedynczymi nagrodami

Rezultat:

	Prize_Share	suma_dzielonych	liczba
1	1/1	351	.
2	1/2	328	164
3	1/3	225	75
4	1/4	64	16

Rysunek 30 Tabela z liczbą nagród za projekt

Z tabelki możemy odczytać, że najwięcej nagród zostało wręczonych indywidualnie, następnie osoby pracujące w parach, zespoły trzyosobowe, i na końcu zespoły czteroosobowe.

### Potrójna nagroda

Laureaci / organizacje mogą otrzymać Nagrodę Nobla więcej niż raz za pracę w tej samej lub innej dziedzinie / temacie. Sprawdźmy, czy jest ktoś kto dostał nagrodę więcej niż 2 razy:

```

proc sql;
select count(Laureate_ID) as liczba, Full_Name as Imie
from Sas.Tab3
group by Full_Name
having liczba > 2
order by liczba asc;
quit;

```

Rysunek 31 Polecenie sprawdzające czy ktoś uzyskał więcej niż 2 nagrody

Rezultat:

3	Comité international de la Croix
---	----------------------------------

Rysunek 32 Rezultat polecenia - zdobywca więcej niż 2 nagród

Zdobywcą jest - Comité international de la croix rouge – inaczej Międzynarodowy Komitet Czerwonego Krzyża. Jest to niezależna, neutralna organizacja zapewniająca ochronę i pomoc humanitarną ofiarom wojny i przemocy. Nie jest organizacją rządową, mimo iż współpracują z nią rządy państw. Nie jest również organizacją pozarządową, ponieważ oprócz działalności społecznej i indywidualnej obejmuje także zakres działalności państwowej.



*Rysunek 33 Logo Międzynarodowego Komitetu Czerwonego Krzyża*



## Analiza czasowa

### Korelacja wieku do roku zdobycia nagrody

Czy istnieje związek pomiędzy wiekiem a rokiem uzyskania nagrody? Żeby się tego dowiedzieć musimy stworzyć nową zmienną 'Age' czyli wiek. Możemy ją uzyskać dzięki odjęciu roku, w którym noblista uzyskał nagrodę od roku urodzenia.

```
data Sas.Tab4;  
set Sas.Tab1;  
Age=Year-year(Birth_Date);  
format Age 3.;  
run;
```

Rysunek 34 Tworzenie zmiennej Age

Jeśli chcemy uzyskać współczynnik korelacji liniowej Pearsona, musimy sprawdzić, czy nasze zmienne posiadają rozkład normalny. Na początek weźmiemy pod uwagę zmienną Age. Można to sprawdzić za pomocą procedury univariate:

```
proc univariate data= sas.tab7 normal;  
var Age;  
run;
```

Rysunek 35 Polecenie sprawdzające czy zmienna Age ma rozkład normalny

Otrzymany rezultat:

Testy normalności				
Test	Statystyka		Wartość p	
Shapiro-Wilka	W	0.993171	Pr. < W	0.0067
Kołmogorowa-Smirnowa	D	0.042091	Pr. > D	<0.0100
Cramera-von Misesa	W-kwadr.	0.142836	Pr. > W-kwadr.	0.0307
Andersona-Darlinga	A-kwadr.	0.929167	Pr. > A-kwadr.	0.0197

Rysunek 36 Rezultat proc univariate

Żeby zmienna miała rozkład normalny wartość p musi przekraczać 0,05. Jak widać dla żadnego testu ta zależność nie jest spełniona. Dlatego można wybrać korelację Spearmana, dla której zmienne nie muszą mieć rozkładu normalnego. Sprawdźmy jaką wartość przyjmuje współczynnik:

```
proc corr data=sas.tab7
  outs=sperman outp=pearson;
  var Age Year;
run;
```

Rysunek 37 Polecenie korelacja Spearmana

Współczynniki korelacji Spearmana, N = 616 Prawd. >  r  przy H0: rho=0		
	Age	Year
Age	1.00000	0.43440 <.0001
Year	0.43440 <.0001	1.00000

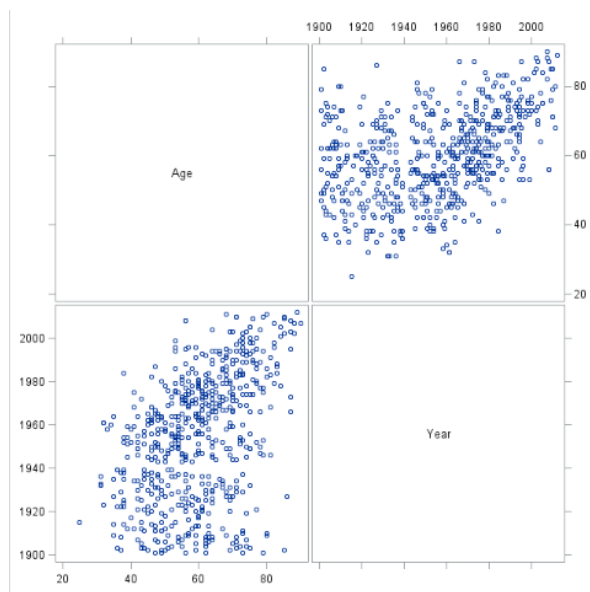
Rysunek 38 Współczynnik korelacji Spearmana

Wartość współczynnika korelacji Spearmana, który mierzy dowolną monotoniczną zależność wynosi 0,4344. Jest on ogólniejszy niż współczynnik korelacji Pearsona. Korelacja jest dodatnia, czyli gdy wartość wieku rośnie to lata też. Siła korelacji zawiera się w przedziale (0.4;0.7), czyli jest to umiarkowana zależność. Zobaczmy, jak to prezentuje się na wykresie:

```
proc sgscatter data=sas.tab7;
  matrix Age Year ;
run;
```

Rysunek 39 Polecenie tworzące wykres macierzowy

Rezultat:



Rysunek 40 Wykres macierzowy Age i Year

### Średnia wieku, a kategoria

W kolejnym kroku warto się przyjrzeć średniej wieku Noblistów w poszczególnych kategoriach. W tym celu użyjemy poniższego kodu i od razu posortujemy wiek rosnąco.

```
proc sql;
  select avg(Age) as srednia_wieku, Category
  from Sas.Tab4
  group by Category
  order by srednia_wieku asc;
quit;
```

Rysunek 41 Polecenie liczące średnią wieku

Otrzymujemy następującą tabelę:

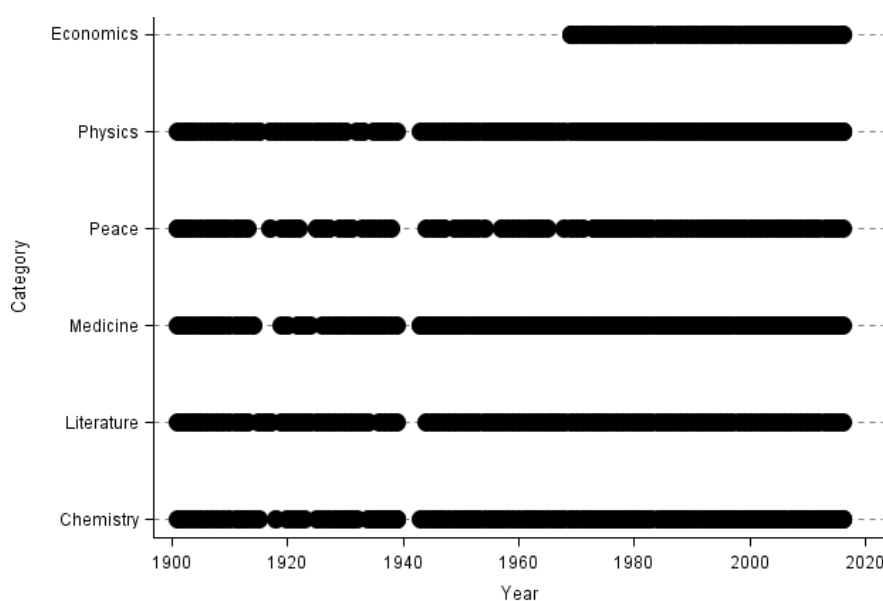
srednia_wieku	Category
55.85	Physics
57.93392	Medicine
58.13472	Chemistry
61.38835	Peace
64.67257	Literature
67.15854	Economics

Rysunek 42 Średnia wieku w kategoriach

Z tabeli otrzymujemy dowiadujemy się, że najmniejszą średnią wieku mają Nobliści w kategorii – fizyka, a największą w kategorii Ekonomia. Różnią się one o 12 lat.

### *Ciągłość przyznawania nagród*

Przyjrzyjmy się teraz ciągłości przyznawania nagród. W tym celu posłużyłam się narzędziem ODS Graphic Designer dostępnym w module SAS BASE. Na osi x znalazły się kategorie, a na osi y daty.



*Rysunek 43 Wykres - ciągłość przyznawania nagród*

Możemy zignorować w tym wykresie kategorię – Ekonomię, ponieważ trzeba wziąć pod uwagę, że w tej dziedzinie Noble zaczęło wręczać w 1969 roku. Jednak jako jedyna zachowała ona swoją ciągłość i od początku jej przyznawania nie została pominięta ani jedna nagroda. Najbardziej nieciągłą kategorią jest – ‘Pokój’. Wyraźnie widać przerwę ciągłości około roku 1940. Z danych historycznych wiemy, że trwała wtedy II Wojna Światowa. Tak samo można zauważyć przerwę ciągłości przed rokiem 1920, kiedy trwała I Wojna Światowa. Noble nie były wtedy przyznawane z uwagi na globalny zasięg wojny.

## Analiza wieku

Chcąc dowiedzieć się jaki był najmłodszy i najstarszy Noblista, oraz jaka była mediana wieku, stosujemy poniższy kod.

```
Proc Means Data=Sas.Tab4 Min Median Max;  
Var Age;  
Run;
```

Rysunek 44 Polecenie procedury means - Age

Otrzymujemy następujący rezultat:

Zmienna analizowana: Age		
Minimum	Mediana	Maksimum
17.0000000	60.0000000	90.0000000

Rysunek 45 Rezultat polecenia means - Age

Wynika z niego, że najmłodszy Noblista miał tylko 17 lat, a najstarszy 90. Czyli na nic nie jest za późno, ani za wcześnie 😊 . Z kolei 50% Noblistów jest starsza lub ma równe 60 lat, a pozostałe 50% ma wiek równy bądź mniejszy niż 60 lat.

## Długość życia

Przy pomocy zmiennej Death\_Date, czyli dacie śmierci, możemy dowiedzieć się ile żyli Nobliści. Na początek musimy usunąć z tabeli Noblistów, którzy jeszcze żyją:

```
data Sas.Tab5;  
set Sas.Tab4;  
if Death_Date=. then delete;  
run;
```

Rysunek 46 Polecenie tworzące tabelę, tylko z zmarłymi Noblistami

Dzięki poniższemu kodowi uzyskamy nową zmienną - długość życia.

```
data Sas.Tab6;  
set Sas.Tab5;  
dlugosc_zycia=ceil(yrdif(Birth_Date,Death_Date , 'dlugosc_zycia'));  
format dlugosc_zycia 3.;  
run;
```

Rysunek 47 Polecenie tworzące zmienną długość\_życia

Rezultat:

	Organization_Country	Death_Date	Death_City	Death_Country	Age	dlugosc_zycia
1	Germany	1911-03-01	Berlin	Germany	49	59
2		1907-09-07	Châtenay	France	62	69
3	Germany	1917-03-31	Marburg	Germany	47	64
4		1910-10-30	Heiden	Switzerland	73	83
5		1912-06-12	Paris	France	79	91
6	Germany	1923-02-10	Munich	Germany	56	78
7	Germany	1919-07-15	Berlin	Germany	50	67
8		1903-11-01	Charlottenburg	Germany	85	86
9	United Kingdom	1932-09-16	Putney Heath	United Kingdom	45	76

Rysunek 48 Tabela z zmarłymi Noblistami

Następnie przeanalizujemy naszą nową zmienną:

```
proc means Data=Sas.Tab6 Min Median Max;  
var dlugosc_zycia;  
run;
```

Rysunek 49 Polecenie procedury means - długość życia

Tabela wynikowa:

Zmienna analizowana: dlugosc_zycia		
Minimum	Mediana	Maksimum
40.0000000	82.0000000	104.0000000

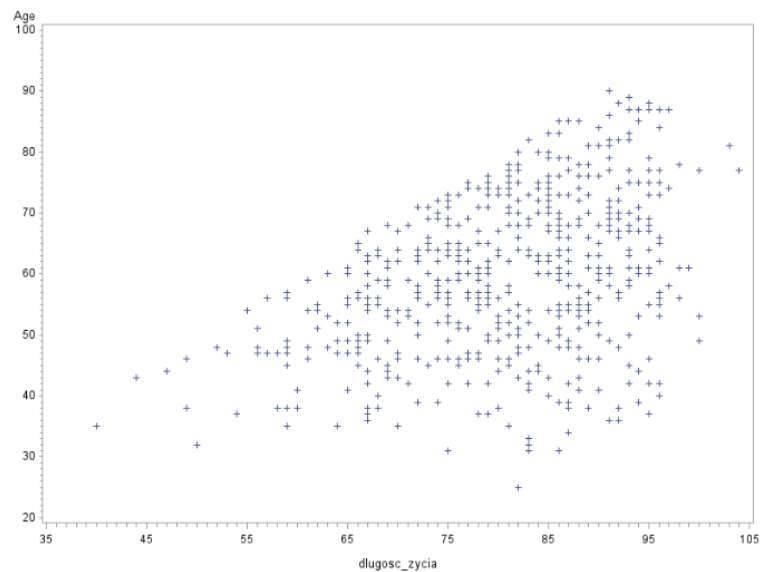
Rysunek 50 Rezultat procedury means - długość życia

Najdłużej żyjący Noblista żył 104 lata, a najkrócej 40 lat. Z kolei 50% Noblistów żyło mniej lub równe 82 lata, a kolejne 50% żyło więcej. Przedstawmy w interpretacji geometrycznej na wykresie punktowym, wiek, kiedy uzyskali Nobla i ile żyli Laureaci za pomocą poniższego kodu:

```
goptions reset=all border;  
  
proc gplot data=Sas.Tab7;  
plot Age*dlugosc_zycia / haxis=35 to 110 by 10  
vaxis=20 to 100 by 10;  
run;
```

Rysunek 51 Procedura tworząca wykres punktowy

Rezultat:



Rysunek 52 Wykres punktowy – Age / długość życia

Można zauważyć zagęszczenie w przedziale długości życia <65,95> i w przedziale wieku uzyskania nagrody <40,80>, kiedy Nobliści zdobyli swoje nagrody. Nie widać wyraźnej zależności pomiędzy tymi zmiennymi.

#### Ile lat żyli nobliści po otrzymaniu nagrody?

Żeby sprawdzić jaka jest długość przeżycia po wręczeniu Nobla, musimy odjąć zmienną długość\_zycia od 'Age'-wieku. W ten sposób otrzymujemy nową kolumnę z nazwą, ile\_po.

```
data Sas.Tab7;  
set sas.Tab6;  
ile_po=dlugosc_zycia-Age;  
format ile_po 3.;  
run;
```

Rysunek 53 Procedura tworząca tabelę z zmienną, ile\_po

Rezultat:

	Organization_City	Organization_Country	Death_Date	Death_City	Death_Country	Age	dlugosc_zycia	ile_po
1	Berlin	Germany	1911-03-01	Berlin	Germany	49	59	10
2			1907-09-07	ChÃˆtenay	France	62	69	7
3	Marburg	Germany	1917-03-31	Marburg	Germany	47	64	17
4			1910-10-30	Heiden	Switzerland	73	83	10
5			1912-06-12	Paris	France	79	91	12
6	Munich	Germany	1923-02-10	Munich	Germany	56	78	22
7	Berlin	Germany	1919-07-15	Berlin	Germany	50	67	17
8			1903-11-01	Charlottenburg	Germany	85	86	1
9	Liverpool	United Kingdom	1932-09-16	Putney Heath	United Kingdom	45	76	31

Rysunek 54 Tabela ze zmiennq, ile\_po

Posortujemy teraz wyniki, ile\_po rosnqco:

```
Proc sort data=Sas.Tab7;
  by ile_po;
run;
```

Rysunek 55 Polecenie sortujqce po ile\_po

Ostatnie rekordy:

dlugosc_zycia	ile_po
94	55
92	56
92	56
96	56
82	57
95	58

Rysunek 56 Ostatnie rekordy posortowanej tabeli

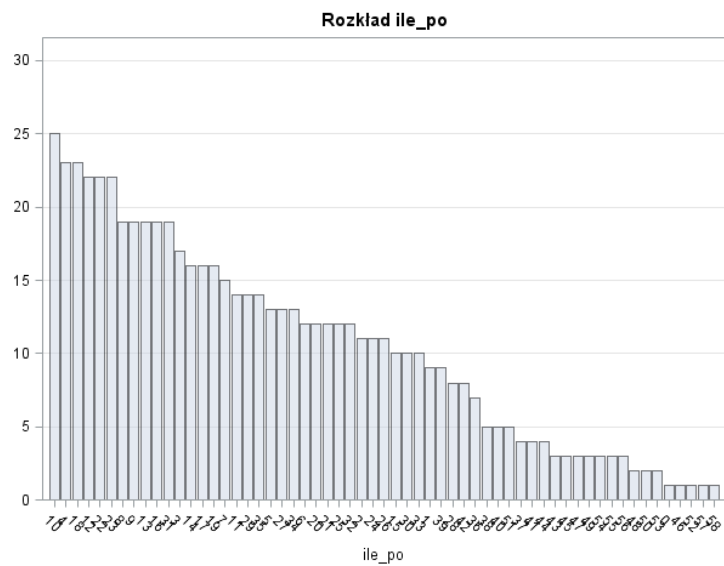
Sprawdźmy, jak prezentujq siq dane na wykresie słupkowym:

```
proc freq data=sas.tab7 order=freq;
  table ile_po / plots=FreqPlot;
run;
```

Rysunek 57 Polecenie tworzcqce wykres słupkowy

Rezultat:





*Rysunek 58 Wykres słupkowy, ile\_po*

Najwięcej osób żyło 10 lat po otrzymaniu Nagrody Nobla. Po przyjęciu wygranej długość przeżycia zawierała się w przedziale  $<0, 58>$  lat.

## Analiza ze względu na kraj urodzenia

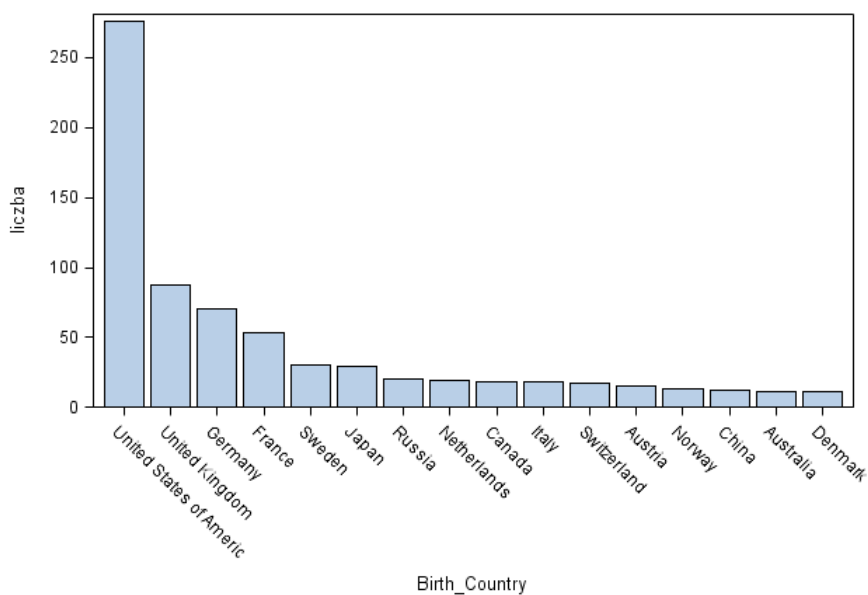
Możemy również przeprowadzić analizę ze względu na kraj urodzenia. Można przyjąć, że właśnie tam trafiały zdobyte Noble. W tym celu grupujemy po kraju urodzenia i bierzemy pod uwagę tylko te kraje które mają ponad 10 zdobytych nagród.

```
proc sql;  
  select count(Birth_Country) as liczba, Birth_Country  
  from Sas.tab1  
  group by Birth_Country  
  having liczba>10  
  order by liczba desc;  
quit;
```

Rysunek 59 Polecenie zliczające kraje urodzenia

Rezultat:

liczba	Birth_Country
276	United States of Americ
88	United Kingdom
70	Germany
53	France
30	Sweden
29	Japan
20	Russia
19	Netherlands
18	Canada
18	Italy
17	Switzerland
15	Austria
13	Norway
12	China
11	Australia
11	Denmark



Rysunek 60 Tabela ze zliczonymi krajami urodzenia

Rysunek 61 Rezultat polecenia zliczającego

Można odczytać, że największą ilością zdobytych Nobli cieszą się Stany Zjednoczone, następnie daleko w tyle jest Wielka Brytania i gonią ich Niemcy i Francja.

Przejdźmy teraz do rozkładu nagród w Stanach.

## W Stanach

Stworzmy nową tabelę, w której zliczymy ilość zdobytych nagród w stosunku do roku:

```
proc sql;  
  create table Sas.Tabl0 as select count(Year) as liczba, Year  
  from Sas.Tab9  
  where Birth_Country='United States of Americ'  
  group by Year  
  order by Year asc;  
quit;
```

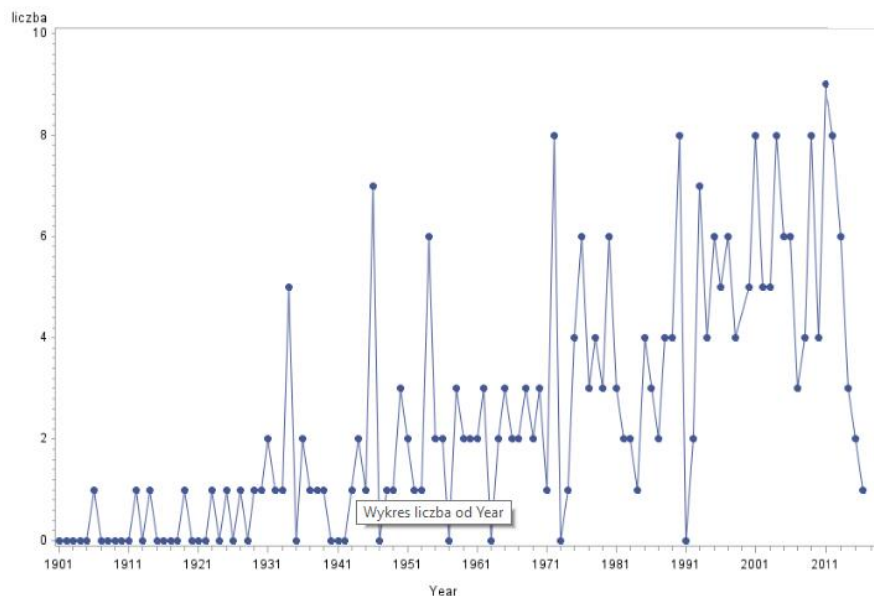
Rysunek 62 Procedura zliczenia nagród zdobytych w ciągu lat w Stanach

Na podstawie utworzonej tabeli tworzymy wykres ze zmienną na osi x=jako liczbą zdobytych Nobli i osi y=jako roku.

```
symbol interpol=join  
value=dot;  
proc gplot data=sas.tabl0;  
  plot liczba*Year / haxis=1901 to 2016 by 10  
  vaxis=0 to 10 by 2;  
  run;  
quit;
```

Rysunek 63 Procedura tworzenia wykresu punktowego łączonego

Rezultat:



Rysunek 64 Wykres punktowy łączony

Można zauważyć na przestrzeni lat tendencje wzrostową ilości zdobytych Nobli. W początkowym okresie na przełomie lat 1906-1930 Ameryka była dość bierna pod względem uzyskanych nagród. Zdarzały się lata, że zdobyła jeden. W 2011 Stany zdobyły aż 9 laurów, po czym w kolejnych latach zaliczali coraz to większe spadki.

### *Pokojowa Nagroda Nobla*

Warto rozważyć również przyznanie Pokojowej Nagrody Nobla ze względu na pochodzenie Laureatów, skąd się wywodzili. Tworzymy nową tabelę, gdzie kategorią ma być Pokój.

```
data Sas.Tab11;  
set Sas.Tab1;  
if Category='Peace' then output Sas.Tab11;  
run;
```

*Rysunek 65 Polecenie tworzące tabelę - Pokojowa Nagroda Nobla*

Następnie zliczamy liczebność tych krajów:

```
proc freq data=Sas.Tab11 noprint;  
tables Birth_Country / out=Sas.Tab12;  
run;
```

*Rysunek 66 Polecenie tworzące tabelę zliczającą kraje urodzenia*

I finalnie sortujemy rosnąco po liczebności uzyskanych Nobli:

```
proc sort data=sas.tab12;  
by descending Count ;  
run;
```

*Rysunek 67 Polecenie sortowania malejąco liczebności*

Otrzymujemy tabelę:

	Birth_Country	Liczebność	Procent całości
1	United States of Americ	19	18.269230769
2	France	9	8.6538461538
3	Germany	5	4.8076923077
4	Sweden	5	4.8076923077
5	United Kingdom	5	4.8076923077
6	Northern Ireland	4	3.8461538462
7	Belgium	3	2.8846153846
8	Egypt	3	2.8846153846
9	South Africa	3	2.8846153846
10	Switzerland	3	2.8846153846
11	Argentina	2	1.9230769231
12	East Timor	2	1.9230769231
13	Liberia	2	1.9230769231
14	Norway	2	1.9230769231
15	Scotland	2	1.9230769231
16	Austria	1	0.9615384615

*Rysunek 68 Tabela Kraje urodzenia – liczebność*

Najwięcej nagród – aż 19 pochodzi z Stanów Zjednoczonych Ameryki, można to tłumaczyć tym, że USA jest jednym z najbardziej liczących się krajów na arenie międzynarodowej, oraz to że mają największe zasięgi w każdym kraju. Następną pozycją jest Francja z 9 Noblami. Dalej znajdują się Niemcy, Szwecja i Wielka Brytania z wynikiem 5 nagród. Trzeba zauważyć, że te kraje znajdują się w wysoko rozwiniętej czołówce. Na tej podstawie można stwierdzić, że państwa wysoko usytuowane mają największy wpływ na zachowanie pokoju na świecie.

## Analiza argumentacji

Bardzo ważne jest za co przyznawane są nagrody. Przeanalizujemy teraz kolumnę 'Motivation'. Do tego celu wykorzystamy poniższy kod, który rozłoży tekst na wyrazy, a następnie zliczy ich występowanie:

```
data Sas.nowa;
  set Sas.tab1;
  do _n_=1 to countw(Motivation);
    new_var=scan(Motivation,_n_);
    output;
  end;
run;

ods output Sas.nowa=Sas.nowa2 (keep=Motivation new_var frequency);
PROC FREQ DATA = Sas.nowa
  ORDER=FREQ
;
  TABLES Motivation * new_var /
    NOCOL
    NOPERCENT
    NOCUM
    SCORES=TABLE
    LIST
    ALPHA=0.05;
RUN; QUIT;

ods all close;
ods listing;
```

Rysunek 69 Procedura zliczania ilości wyrazów w Motivation

Otrzymujemy tabelę:

Procedura FREQ		
Motivation	new_var	Liczebność
"for their contributions to the theory of the unified weak and electromagnetic interaction between elementary particles, including, inter alia, the prediction of the weak neutral current"	of	8
"for their contributions to the theory of the unified weak and electromagnetic interaction between elementary particles, including, inter alia, the prediction of the weak neutral current"	the	16
"for their contributions to the theory of the unified weak and electromagnetic interaction between elementary particles, including, inter alia, the prediction of the weak neutral current"	"for	4
"for their contributions to the theory of the unified weak and electromagnetic interaction between elementary particles, including, inter alia, the prediction of the weak neutral current"	and	4
"for their contributions to the theory of the unified weak and electromagnetic interaction between elementary particles, including, inter alia, the prediction of the weak neutral current"	their	4
"for their contributions to the theory of the unified weak and electromagnetic interaction between elementary particles, including, inter alia, the prediction of the weak neutral current"	to	4
"for their contributions to the theory of the unified weak and electromagnetic interaction between elementary particles, including, inter alia, the prediction of the weak neutral current"	contributions	4
"for their contributions to the theory of the unified weak and electromagnetic interaction between elementary particles, including, inter alia, the prediction of the weak neutral current"	theory	4

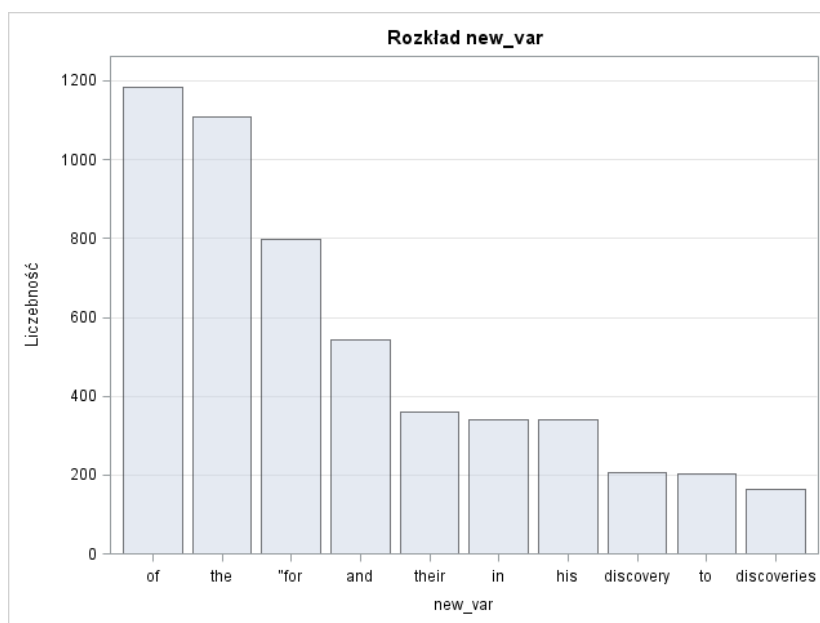
Rysunek 70 Tabela - liczebność występowania wyrazów

Zobaczmy, jak wygląda top 10 najczęściej występujących słów na wykresie:

```
%let TopN=10;  
proc freq data=Sas.Nowa order=freq;  
tables new_var/maxilevels=&TopN Plot=FreqPlot;  
run;
```

Rysunek 71 Polecenie ograniczenia do 10 najczęściej występujących wyrazów

Rezultat:



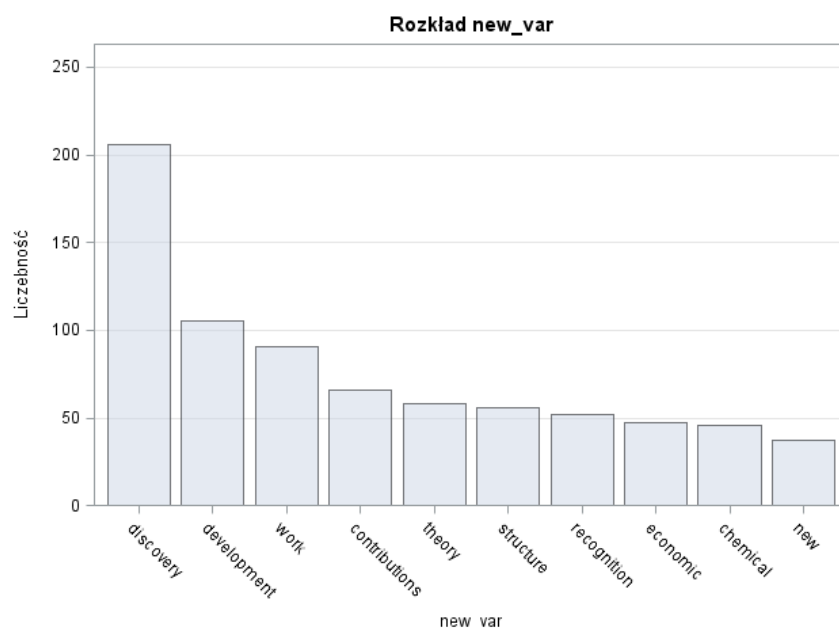
Rysunek 72 Wykres kolumnowy - najczęściej występujące słowa

Widzimy, że na wykresie są słowa (przedimki, zaimki) takie jak „of” czy „the”, które nie dostarczają nam informacji, więc usuńmy przeszkadzające wyrazy:

```
data Sas.Nowa4;  
set Sas.Nowa3;  
if new_var='by' then delete;  
if new_var='in' then delete;  
if new_var='he' then delete;  
if new_var='has' then delete;  
if new_var='and' then delete;  
if new_var='their' then delete;  
if new_var='to' then delete;  
run;
```

Rysunek 73 Polecenie tworzące nową tabelę bez zbędnych wyrazów

Finalny wykres ze słowami mającymi znaczenie dla analizy prezentuje się następująco: (użyty został ponownie poprzedni kod)



Rysunek 74 Wykres kolumnowy - najczęściej występujące słowa 2

Widzimy więc, że najczęściej występującymi słowami są:

Discovery - odkrycie

Development - rozwój

Work - praca

Contributions - wkład

Theory - teoria

Structure - struktura

Recognition - uznanie

Economic - gospodarczy

Chemical – chemiczny

New – nowy

Największą sławą cieszy się „Discovery” czyli odkrycie. Wynika to z tego, że Nobla przyznaje się głównie za odkrycia naukowe. Kolejnym najpopularniejszym słowem jest „Development” oznaczającym rozwój. Postęp w różnych dziedzinach ułatwia, a nawet ratuje życie. Na trzecim miejscu znajduje się praca – bez pracy nie ma żadnych efektów, a Nobla się za coś przyznaje.



## **Wnioski**

Z przeprowadzonej analizy można wywnioskować kilka rzeczy. Po pierwsze najczęściej nagradza się wybitnych ludzi w kategoriach takich jak: medycyna, fizyka i chemia.

Kolejnym spostrzeżeniem jest analiza płci. Wynika z niej to, że najczęściej nagradzani są mężczyźni. Kobiety najwięcej nagród zdobyły w literaturze i medycynie, a mężczyźni w fizyce. Pierwszą noblistką była Polka – Maria Skłodowska-Curie. Płeć żeńska jest w wyróżnieniach często pomijana.

Najwięcej nagród przyznawanych jest pojedynczo. Mimo, że bardzo często pracuje się w zespole, to nagroda jest przyznawana osobie, która włożyła największy wkład w daną pracę. Warto wspomnieć, że jedynym potrójny Nobel trafił do organizacji - Międzynarodowy Komitet Czerwonego Krzyża.

Nie ma dużej zależności pomiędzy wiekiem w jakim zdobyło się nagrodę, a rokiem w jakim się ją uzyskało. Najniższa średnia wieku jest w dziedzinie fizyki, a najwyższa w ekonomii. Wojny nie sprzyjają przyznawaniu nagród, następuje wtedy ich wstrzymanie. Wartością środkową wieku uzyskania nobla jest 60 lat. Z kolei wartość środkowa przeżycia Noblistów wynosi 82 lata.

Najwięcej osób, które dostały Nobla urodziło się w Stanach Zjednoczonych Ameryki, w 2011 roku zdobyły aż 9 nagród. Przodują również w zdobywaniu Pokojowych Nagród Nobla – 19.

Po analizie argumentacji, za co przyznawany jest Nobel, widzimy, że najczęściej słów to odkrycie, rozwój i praca. Słowa te sugerują nam, że osoby, które zostały nagrodzone, nie dążyły do zdobycia sławy czy pieniędzy. Ich kluczem do sukcesu była ciężka praca i dążenie do osiągnięcia przełomu. Nagroda była tylko uhonorowaniem ich działań.

Podsumowując największe szanse na Nagrodę Nobla mają mężczyźni, w wieku 60 lat, pracujących indywidualnie, w dziedzinie fizyki, urodzonych w Stanach Zjednoczonych Ameryki. Żeby przyznali im nagrodę, musieli by coś odkryć albo rozwinąć istniejącą już technologię.

## Spis ilustracji

Rysunek 1 Import bazy.....	4
Rysunek 2 Baza .....	4
Rysunek 3 Procedura zawartości.....	4
Rysunek 4 Wynik zawartości .....	4
Rysunek 5 Alfabetyczna lista zmiennych i atrybutów .....	5
Rysunek 6 Polecenie – lista przyznanych Nobli w kategoriach.....	6
Rysunek 7 Wynik – lista przyznanych Nobli w kategoriach .....	6
Rysunek 8 Polecenie wykres słupkowy – liczba przyznanych nagród w kategoriach.....	7
Rysunek 9 Wykres słupkowy – liczba przyznanych nagród w kategoriach .....	7
Rysunek 10 Polecenie - częstotliwość występowania płci.....	8
Rysunek 11 Wynik - częstotliwość występowania płci .....	8
Rysunek 12 Polecenie wykres kołowy - rozkład płci.....	8
Rysunek 13 Wykres kołowy - rozkład płci .....	9
Rysunek 14 Polecenie sortujące z wydrukiem .....	10
Rysunek 15 Wynik polecenia sortującego - wydruk.....	10
Rysunek 16 Polecenie podział nagród między kobiety – kategorie .....	11
Rysunek 17 Wynik podział nagród między kobiety – kategorie.....	11
Rysunek 18 Polecenie tworzące tabele - kobiety .....	11
Rysunek 19 Polecenie tworzące wykres kołowy udział procentowy kobiet w kategoriach ....	11
Rysunek 20 Wykres kołowy udział procentowy kobiet w poszczególnych kategoriach.....	12
Rysunek 21 Polecenie - podział nagród pomiędzy mężczyznami.....	12
Rysunek 22 Wynik polecenia - podział nagród pomiędzy mężczyznami.....	12
Rysunek 23 Polecenie tworzące tabele – mężczyźni .....	13
Rysunek 24 Polecenie tworzące wykres kołowy udział procentowy mężczyzn w kategoriach .....	13
Rysunek 25 Wykres kołowy udział procentowy mężczyzn w poszczególnych kategoriach ...	13
Rysunek 26 Polecenie sortowania tabeli po dzielonych nagrodach .....	14
Rysunek 27 Polecenie zliczające ilość nagród dzielonych.....	14
Rysunek 28 Tabela - zliczona ilość nagród dzielonych .....	14
Rysunek 29 Polecenie tworzące tabele z pojedynczymi nagrodami .....	15
Rysunek 30 Tabela z liczbą nagród za projekt .....	15
Rysunek 31 Polecenie sprawdzające czy ktoś uzyskał więcej niż 2 nagrody .....	15
Rysunek 32 Rezultat polecenia - zdobywca więcej niż 2 nagród .....	15
Rysunek 33 Logo Międzynarodowego Komitetu Czerwonego Krzyża.....	16
Rysunek 34 Tworzenie zmiennej Age.....	17
Rysunek 35 Polecenie sprawdzające czy zmienna Age ma rozkład normalny .....	17
Rysunek 36 Rezultat proc univariate.....	17
Rysunek 37 Polecenie korelacja Spearmana .....	18
Rysunek 38 Współczynnik korelacji Spearmana .....	18
Rysunek 39 Polecenie tworzące wykres macierzowy .....	18
Rysunek 40 Wykres macierzowy Age i Year.....	19
Rysunek 41 Polecenie liczące średnią wieku .....	19
Rysunek 42 Średnia wieku w kategoriach.....	19
Rysunek 43 Wykres - ciągłość przyznawania nagród.....	20
Rysunek 44 Polecenie procedury means - Age .....	21
Rysunek 45 Rezultat polecenia means - Age .....	21
Rysunek 46 Polecenie tworzące tabelę, tylko z zmarłymi Noblistami .....	21

Rysunek 47 Polecenie tworzące zmienną długość_życia.....	21
Rysunek 48 Tabela z zmarłymi Noblistami .....	22
Rysunek 49 Polecenie procedury means - długość_życia .....	22
Rysunek 50 Rezultat procedury means - długość_życia .....	22
Rysunek 51 Procedura tworząca wykres punktowy .....	22
Rysunek 52 Wykres punktowy – Age / długość życia .....	23
Rysunek 53 Procedura tworząca tabelę z zmienną, ile_po.....	23
Rysunek 54 Tabela ze zmienną, ile_po .....	24
Rysunek 55 Polecenie sortujące po ile_po .....	24
Rysunek 56 Ostatnie rekordy posortowanej tabeli .....	24
Rysunek 57 Polecenie tworzące wykres słupkowy .....	24
Rysunek 58 Wykres słupkowy, ile_po .....	25
Rysunek 59 Polecenie zliczające kraje urodzenia .....	26
Rysunek 60 Tabela ze zliczonymi krajami urodzenia .....	26
Rysunek 61 Rezultat polecenia zliczającego .....	26
Rysunek 62 Procedura zliczenia nagród zdobytych w ciągu lat w Stanach .....	27
Rysunek 63 Procedura tworzenia wykresu punkowego łączonego .....	27
Rysunek 64 Wykres punktowy łączony .....	27
Rysunek 65 Polecenie tworzące tabele - Pokojowa Nagroda Nobla.....	28
Rysunek 66 Polecenie tworzące tabele zliczające kraje urodzenia .....	28
Rysunek 67 Polecenie sortowania malejąco liczebności.....	28
Rysunek 68 Tabela Kraje urodzenia – liczebność .....	29
Rysunek 69 Procedura zliczania ilości wyrazów w Motivation.....	30
Rysunek 70 Tabela - liczebność występowania wyrazów .....	30
Rysunek 71 Polecenie ograniczenia do 10 najczęściej występujących wyrazów .....	31
Rysunek 72 Wykres kolumnowy - najczęściej występujące słowa.....	31
Rysunek 73 Polecenie tworzące nową tabelę bez zbędnych wyrazów .....	31
Rysunek 74 Wykres kolumnowy - najczęściej występujące słowa 2.....	32

## Bibliografia

„Elementy wnioskowania statystycznego z wykorzystaniem pakietu SAS w przykładach” – Wojciech Gamrot, 2012 Wydawnictwo Uniwersytetu Ekonomicznego

<https://documentation.sas.com/?docsetId=graphref&docsetTarget=p12onmnsmc7hfn1lgi55vvh9lv2.htm&docsetVersion=9.4&locale=en>

<https://documentation.sas.com/?docsetId=graphref&docsetTarget=n1e0onjgevom4kn1hbiibc6w34oe.htm&docsetVersion=9.4&locale=en>

[http://support.sas.com/rnd/papers/sugi30/p90-30\\_020105.pdf](http://support.sas.com/rnd/papers/sugi30/p90-30_020105.pdf)

<https://communities.sas.com/t5/SAS-Procedures/Proc-freq-output-in-descending-order-of-variable-values/td-p/62396>

<https://communities.sas.com/t5/SAS-Programming/Descending-order-frequencies-in-proc-freq-order-freq/td-p/347599>

<http://support.sas.com/kb/25/510.html>

[https://support.sas.com/sassamples/graphgallery/PROC\\_GCHART.html](https://support.sas.com/sassamples/graphgallery/PROC_GCHART.html)

<https://blogs.sas.com/content/graphicallyspeaking/2012/03/30/cluster-groups/>

<https://support.sas.com/rnd/datavisualization/papers/GraphBasics.pdf>

[https://documentation.sas.com/?docsetId=bivwrand&docsetTarget=view-data\\_sort\\_charts.htm&docsetVersion=8.32&locale=en](https://documentation.sas.com/?docsetId=bivwrand&docsetTarget=view-data_sort_charts.htm&docsetVersion=8.32&locale=en)

[https://documentation.sas.com/?docsetId=procstat&docsetTarget=procstat\\_corr\\_details03.htm&docsetVersion=9.4&locale=en](https://documentation.sas.com/?docsetId=procstat&docsetTarget=procstat_corr_details03.htm&docsetVersion=9.4&locale=en)

<http://support.sas.com/kb/37/000.html>

<https://documentation.sas.com/?docsetId=egamotasks&docsetTarget=p1dedws8689lp2n0znbz1so8s5jo.htm&docsetVersion=8.1&locale=pl>

<https://quizlet.com/180005563/statystyka-w-sas-flash-cards/>

<http://statystyka.org.pl/wprowadzenie-do-programu-sas/>

<https://communities.sas.com/t5/Statistical-Procedures/How-to-test-if-a-sample-has-a-Normal-Distribution/td-p/92080>

<https://stats.idre.ucla.edu/sas/output/proc-univariate/>

<https://data-flair.training/blogs/sas-proc-sort/>

<https://blogs.sas.com/content/iml/2014/12/01/max-and-min-rows-and-cols.html?fbclid=IwAR3Wuf0uIH-pHWbQhAeLOCwst27VZQ8igbvOKs2ti5uk0eeQbigtTaRd8I>

<https://blogs.sas.com/content/iml/2018/02/26/how-to-use-first-variable-and-last-variable-in-a-by-group-analysis-in-sas.html>

<https://wyborcza.pl/1,75400,20779745,dlaczego-kobiety-dostaly-tak-malo-nagrod-nobla-to-zapewne-efekt.html?disableRedirects=true>

[https://pl.wikipedia.org/wiki/Nagroda\\_Nobla](https://pl.wikipedia.org/wiki/Nagroda_Nobla)

<https://www.statystyka-zadania.pl/wspolczynnik-korelacji-spearmana/>