



# SIGMA KWADRAT

CZWARTY LUBELSKI KONKURS STATYSTYCZNO-DEMOGRAFICZNY

## Wykorzystanie programu MS Excel do opracowań statystycznych

**NBP**

Narodowy Bank Polski

PROJEKT DOFINANSOWANY ZE ŚRODKÓW  
NARODOWEGO BANKU POLSKIEGO



URZĄD STATYSTYCZNY  
W LUBLINIE



WYŻSZA SZKOŁA ZARZĄDZANIA  
I ADMINISTRACJI W ZAMOŚCIU



POLSKIE TOWARZYSTWO  
STATYSTYCZNE

# Wprowadzenie

Każdy badacz staje przed problemem wyboru odpowiedniego narzędzia do wprowadzenia danych uzyskanych w badaniu ankietowym i przeprowadzenia analiz i obliczeń. Narzędziem, które bardzo dobrze sprawdza się do prostych jak i bardziej zaawansowanych analiz nawet na stosunkowo dużej ilości danych jest arkusz kalkulacyjny.

W materiale korzystać będziemy z programu Excel z pakietu MS Office (z uwagi na zawarty w nim dodatek Analysis ToolPak, który stanowi zbiór narzędzi przydatnych zwłaszcza w statystyce). Jednakże osoby nie posiadające licencji na ww. program mogą z równie dobrym skutkiem korzystać z oprogramowania dystrybuowanego na zasadach wolnego oprogramowania jak np. wbudowany w pakiet OpenOffice program Calc bądź program Gnumeric.

# Przygotowanie danych do późniejszych analiz

Po zebraniu materiału statystycznego w postaci ankiet należy dobrze przemyśleć i zaplanować układ arkusza do dalszych analiz. Kilka zasad jakich należy przestrzegać:

- Zanim zaczniemy wprowadzać wyniki do arkusza należy dokładnie przemyśleć sposób kodowania odpowiedzi np. w pytaniach zamkniętych jednokrotnego wyboru warto przyporządkować każdej odpowiedzi o ile to możliwe odpowiednie i czytelne kody liczbowe bądź literowe, zaś w pytaniach wielokrotnego wyboru na każdą odpowiedź powinno się przeznaczyć osobną kolumnę w arkuszu.
- Odpowiedzi z jednej ankiety powinny znajdować się jednym wierszem (nie kolumnie, ponieważ kolumn jest mniej w arkuszu aniżeli wierszy).
- Ankietom można przydzielić unikalne identyfikatory (np. numerując je) tak, aby łatwo można było odnaleźć z której ankiety pochodzą dane odpowiedzi.
- Pytania powinny posiadać unikalne i łatwe do rozpoznania etykiety, treść pytań można umieścić w komentarzach do komórki, nie ma sensu umieszczać jej w komórkach.
- Przy dużych ilościach odpowiedzi arkusz z danymi powinien być oddzielony od arkusza z analizami, dzięki czemu zyskuje się lepszą czytelność, można filtrować bądź sortować dane bez potrzeby zmiany pozostałych komórek.

## Przykład

Przykład prostej ankiety i sposobu poprawnego zaprojektowania arkusza do wprowadzania odpowiedzi znajduje się w pliku *w20-przyklad-ankiety.xls* (odpowiednio *arkusze: Ankieta i Odpowiedzi*)

Po poprawnym wprowadzeniu danych następuje etap analiz. Najprostsze z nich polegają na dokonaniu podsumowań kolumn, wyliczeniu liczby odpowiedzi na każdy z wariantów cechy, bądź obliczeniu udziałów (*wartości procentowych*).

# Tabelaryczna i graficzna prezentacja wyników

Dane jednostkowe (wyniki ankiet) są zwykle zbyt obszerne i nieczytelne do bezpośredniej prezentacji w opracowaniu. Można je co najwyżej umieścić w załączniku do pracy (bądź na płycie w formie elektronicznej). Kategorycznie nie powinno się ich umieszczać bezpośrednio w treści naszego opracowania (gdyż po pierwsze nic nie mówią czytającemu pracę, po drugie wręcz zaciemniają pracę, sprawiają, że jest ona mało przejrzysta). W opracowaniu powinno się umieszczać wyniki analizy a co za tym idzie dane przetworzone lub odpowiednio zagregowane. Najczytelniejsze są formy najprostsze do odkodowania takie jak tabela (tablica) bądź wykres (będący graficznym odwzorowaniem tabeli).

# Wyniki

W arkuszu *Wyniki* znajdują się przykłady przetworzenia danych jednostkowych do formy tabeli i wykresu.

Odpowiedź	Liczba odpowiedzi	Udział procentowy
Tak	16	59,26%
Nie	7	25,93%
Nie wiem	4	14,81%
<b>Suma</b>	<b>27</b>	<b>100,00%</b>

