

STATYSTYKA OPISOWA (lab. 10)

GRAFICZNA ANALIZA DANYCH – część 2.

Na dzisiejszych zajęciach zapoznamy się z wybranymi **wykresami statystycznymi** – jest to nieformalne określenie, które dotyczyć będzie wykresów, które pokazują wyniki pewnych analiz statystycznych, a nie dane surowe. Najczęściej są to proste analizy, dotyczące grupowania danych (odpowiadające im wykresy to: **histogramy** i **wykresy kołowe**) albo wyliczania statystyk opisowych (**wykresy ramka-wąsy**).

Wykresy „statystyczne” nie są „wrażliwe” na zmiany struktury zbioru danych – w szczególności sortowanie.

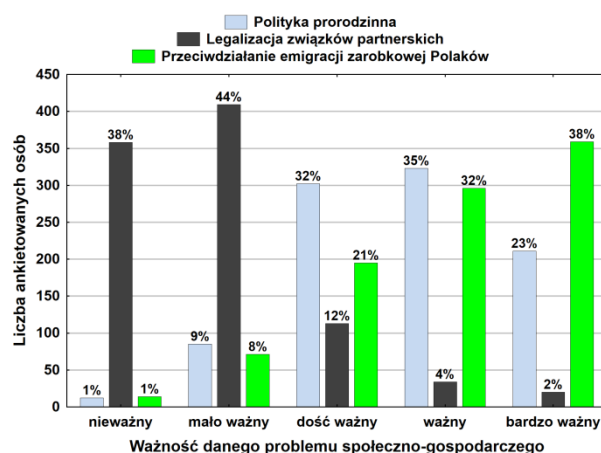
UWAGA: Wszystkie wykorzystywane wykresy są dostępne po wybraniu polecenia **Wykresy / Wykresy 2W**.

Histogram w wersji wielokrotnej Dane: Ankieta studencka 2013-2016

Histogramy są to wykresy pozwalające przedstawić częstość występowania wartości pewnej cechy w badanej zbiorowości. Oczywiście, w zależności od potrzeb użytkownika, program może zliczać wszystkie różne wartości bądź grupować je w pewne przedziały.

Celem analizy będzie jednoczesna wizualizacja odpowiedzi ankietowanych na trzy pytania, dotyczące ważności trzech problemów społecznych w opinii respondentów: *Polityki prorodzinnej*, *Legalizacji związków partnerskich* i *Przeciwdziałania migracji zarobkowej Polaków*.

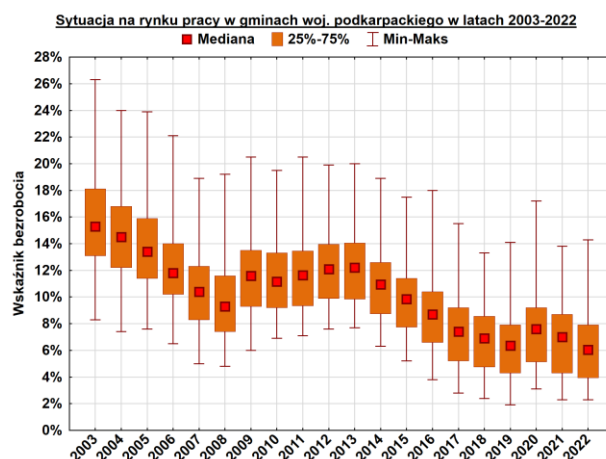
W tym celu ustalamy rodzaj wykresu na **Wielokrotny**. Następnie w zakładce **Więcej** włączamy pewne dodatkowe opcje: **Pokaż procenty**, **Odstępy między kolumnami** i wyłączamy opcję: **Dopasuj**, aby zwiększyć czytelność wykresu. Wykonujemy wykres i starannie go formatujemy, wykorzystując zdobyte wcześniej umiejętności.



Wykres ramka-wąsy w wersji wielokrotnej Dane: Bezrobocie w gminach 2003-2022

Wykres ramka-wąsy służy do wizualizacji statystyk opisowych, a więc miar przeciętnego poziomu i zmienności danej cech liczbowych. Celem analizy będzie prezentacja **bezrobocia** ogółem w 160 gminach Podkarpacia w latach 2003-2022:

- 1) znajdujemy w grupie **Wykresy 2W / Wykres ramka-wąsy**;
- 2) na liście zmiennych wskazujemy wskaźniki bezrobocia **ogółem** z lat 2003-2022 i zmieniamy rodzaj wykresu na **Wielokrotny**;
- 3) włączamy warunek selekcji **Województwo = „podkarpackie”** (prościej v3=9);
- 4) w zakładce **Więcej** ustawiamy **Punkt środkowy** jako **Medianę**, **Ramkę** jako **Centyl 25** i **Wąs** jako **Minimum-maksimum** oraz [wartości] **Odstające** jako **Wyłączone**;
- 5) na wykresie dokonujemy formatowania – zmieniamy położenie legendy, dodajemy tytuł osi i główny tytuł;
- 6) opisy lat na osi poziomej są niewygodne (program wykorzystał nazwy zmiennych „nie wiedząc”, że w zupełności wystarczą same lata), aby je zmienić klikamy dwa razy etykiety osi poziomej i w zakładce **Wartości skali/Edytuj etykiety użytkownika** wprowadzamy własne etykiety (w polu **Pozycja** należy podać 1, a w polu **Tekst** – 2003; następnie 2 i 2004, 3 i 2005 i tak dalej). Można tu wykorzystać arkusz **Excelsa** i tam przygotować tabelę, a potem wkleić ją do programu **STATISTICA**.



STATYSTYKA OPISOWA (lab. 10)

GRAFICZNA ANALIZA DANYCH – część 2.

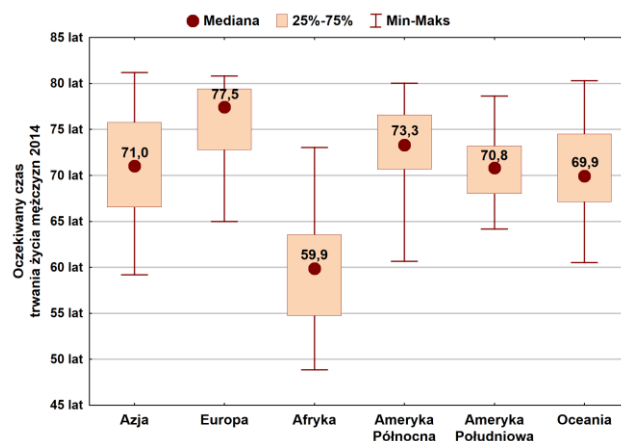
Wykres ramka wąsy ze zmienną grupującą Dane: Czas trwania życia na świecie 1960-2022

Celem analizy będzie porównanie poziomu oczekiwanego czasu trwania życia mężczyzn w państwach świata w roku 2014 według kontynentów. W tym celu ponownie wykorzystamy **wykres ramka-wąsy**. W oknie wyboru zmiennych na pierwszej liście wskazujemy *Oczekiwany czas trwania życia mężczyzn 2014*, a na drugiej *Kontynent*.

W zakładce *Więcej* ustawiamy *Punkt środkowy* jako *Mediana*, *Wąs* jako *Minimum-maksimum* i *Odstające* jako *Wylączone*. Wykonujemy wykres. Dokonujemy formatowania wykresu, tak by uzyskać jego wygląd zgodny z załączonym rysunkiem.

Szczegółowe wskazówki:

- w zakładce *Ramka-wąsy* ustalamy szerokość ramki na 60%;
- w zakładce *Etykiety punktów* włączamy *Etykiety punktów* i zaznaczamy *Współrzędne Y*, odpowiednio je formatując;
- formatujemy tytuły (pogrubiamy i powiększamy czcionkę), przenosimy legendę do tytułu, etc.



Wykres kołowy ze zmienną grupującą Dane: Opinie o integracji z UE 2004

Celem analizy będzie pokazanie sposobu głosowania w referendum akcesyjnym ankietowanych osób, w zależności od płci i miejsca zamieszkania.

W oknie wyboru zmiennych **Wykresu kołowego** wskazujemy *Sposób głosowania w referendum europejskim*, a następnie w zakładce *Skategoryzowane* włączamy *Kategorię X i Y*, wskazując na listach odpowiednio zmienne: *Płeć* i *Miejsce zamieszkania*. Tworzymy wykres.

Na gotowym wykresie włączamy wyświetlanie procentów (pod dwukrotnym kliknięciu etykiet przy wykresie), w zakładce *Kategoryzacja* ustawiamy *Nazwy zmiennych i klas* na *Tylko nazwy klas*.

Formatujemy tytuł wykresu, usuwamy zbędne opisy.

Pozostaje problem **sposobu wyliczania** procentów – jak widać, na wykresie pokazane są procenty względem całej zbiorowości, a powinny się one sumować do 100 na każdym wykresie. Aby zmienić sposób wyświetlania procentów, znajdujemy zakładkę *Histogram* i ustawiamy *Procent na wykresach skategoryzowanych* jako *Kategoria*.

