

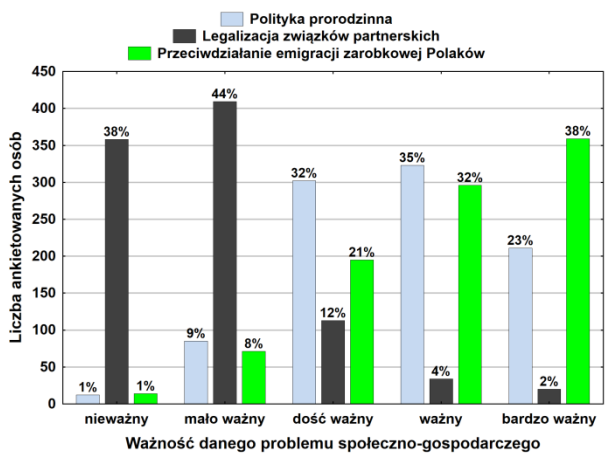
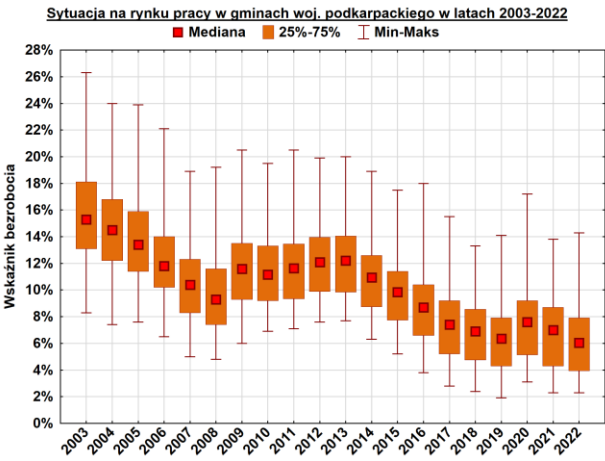
# STATYSTYKA OPISOWA (lab. 10)

## 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*.

<b>Histogram w wersji wielokrotnej</b>	<b>Dane: Ankieta studencka 2013-2016</b>																								
<p><b>Histogramy</b> 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.</p> <p>Celem analizy będzie jednoczesna wizualizacja odpowiedzi ankietowanych na trzy pytania, dotyczące ważności trzech problemów społecznych w opinii respondentów: <i>Polityki prorodzinnej</i>, <i>Legalizacji związków partnerskich</i> i <i>Przeciwdziałania migracji zarobkowej Polaków</i>.</p> <p>W tym celu ustalamy rodzaj wykresu na <b>Wielokrotny</b>. Następnie w zakładce <b>Więcej</b> włączamy pewne dodatkowe opcje: <b>Pokaż procenty</b>, <b>Odstępy między kolumnami</b> i wyłączamy opcję: <b>Dopasuj</b>, aby zwiększyć czytelność wykresu. Wykonujemy wykres i starannie go formatujemy, wykorzystując zdobyte wcześniej umiejętności.</p>	 <table border="1" style="margin-left: auto; margin-right: auto; border-collapse: collapse; text-align: center;"> <caption>Ważność danego problemu społeczno-gospodarczego</caption> <thead> <tr> <th>Ważność</th> <th>Polityka prorodzinna</th> <th>Legalizacja związków partnerskich</th> <th>Przeciwdziałanie emigracji zarobkowej Polaków</th> </tr> </thead> <tbody> <tr> <td>nieważny</td> <td>1%</td> <td>38%</td> <td>1%</td> </tr> <tr> <td>mało ważny</td> <td>9%</td> <td>44%</td> <td>8%</td> </tr> <tr> <td>dość ważny</td> <td>32%</td> <td>12%</td> <td>21%</td> </tr> <tr> <td>ważny</td> <td>35%</td> <td>4%</td> <td>32%</td> </tr> <tr> <td>bardzo ważny</td> <td>23%</td> <td>2%</td> <td>38%</td> </tr> </tbody> </table>	Ważność	Polityka prorodzinna	Legalizacja związków partnerskich	Przeciwdziałanie emigracji zarobkowej Polaków	nieważny	1%	38%	1%	mało ważny	9%	44%	8%	dość ważny	32%	12%	21%	ważny	35%	4%	32%	bardzo ważny	23%	2%	38%
Ważność	Polityka prorodzinna	Legalizacja związków partnerskich	Przeciwdziałanie emigracji zarobkowej Polaków																						
nieważny	1%	38%	1%																						
mało ważny	9%	44%	8%																						
dość ważny	32%	12%	21%																						
ważny	35%	4%	32%																						
bardzo ważny	23%	2%	38%																						
<b>Wykres ramka-wąsy w wersji wielokrotnej</b>	<b>Dane: Bezrobocie w gminach 2003-2022</b>																								
<p><b>Wykres ramka-wąsy</b> służy do wizualizacji statystyk opisowych, a więc miar przeciętnego poziomu i zmienności danej cech liczbowych. Celem analizy będzie prezentacja <b>bezrobocia</b> ogółem w 160 gminach Podkarpacia w latach 2003-2022:</p> <ol style="list-style-type: none"> <li>1) znajdujemy w grupie <i>Wykresy 2W / Wykres ramka-wąsy</i>;</li> <li>2) na liście zmiennych wskazujemy wskaźniki bezrobocia <b>ogółem</b> z lat 2003-2022 i zmieniamy rodzaj wykresu na <b>Wielokrotny</b>;</li> <li>3) włączamy warunek selekcji <b>Województwo = „podkarpackie”</b> (prościej <math>v3=9</math>);</li> <li>4) w zakładce <b>Więcej</b> ustawiamy <b>Punkt środkowy</b> jako <b>Medianę</b>, <b>Ramkę</b> jako <b>Centyl 25</b> i <b>Wąs</b> jako <b>Minimum-maksimum</b> oraz [wartości] <b>Odstające</b> jako <b>Wyłączone</b>;</li> <li>5) na wykresie dokonujemy formatowania – zmieniamy położenie legendy, dodajemy tytuł osi i główny tytuł;</li> <li>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 <b>Wartości skali//Edytuj etykiety użytkownika</b> wprowadzamy własne etykiety (w polu <i>Pozycja</i> należy podać 1, a w polu <i>Tekst</i> – 2003; następnie 2 i 2004, 3 i 2005 i tak dalej). Można tu wykorzystać arkusz <i>Excelsa</i> i tam przygotować tabelę, a potem wkleić ją do programu <i>STATISTICA</i>.</li> </ol>	 <p style="text-align: center;"><b>Sytuacja na rynku pracy w gminach woj. podkarpackiego w latach 2003-2022</b></p> <p style="text-align: center;">■ Mediana    ■ 25%-75%    ┆ Min-Maks</p>																								

## STATYSTYKA OPISOWA (lab. 10)

### GRAFICZNA ANALIZA DANYCH – część 2.

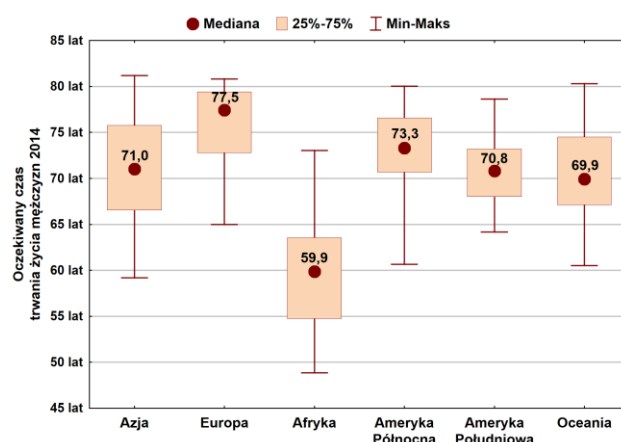
#### Wykres ramka wąsy ze zmienną grupującą Dane: Czas trwania życia na świecie 1960-2014

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 *Wyłączone*. 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*.

