

STATYSTYKA OPISOWA (lab. 9)

GRAFICZNA ANALIZA DANYCH – część 1.

Celem zajęć nr 9 i 10 jest zapoznanie się z wybranymi metodami graficznej analizy danych dostępnymi w programie STATISTICA. „Wybranymi”, gdyż samych tylko podstawowych typów wykresów w programie jest kilkadziesiąt, a uwzględniając ich szczegółowe wersje – nawet kilkaset. Dobór typu wykresu zależy od rodzaju analizowanego zbioru danych i analizowanych cech. Inaczej będziemy wizualizować dane przekrojowe, inaczej ankietowe, inne specyfikacje wykresów będą odpowiednie dla danych liczbowych a inne dla danych nominalnych. Sposób wykonania wykresu może także zależeć od liczby analizowanych przypadków, no i oczywiście od celu analizy.

Większość wykresów w programie STATISTICA jest dostępna poprzez polecenie *Wykresy / Wykresy 2W*.

Na dzisiejszych zajęciach zapoznamy się z wybranymi wykresami prezentacyjnymi – jest to określenie, którego używać będziemy do wykresów odzwierciedlających w „dokładny” sposób strukturę danych. Będzie na nich tyle elementów (słupków, znaczników, linii), ile przypadków liczy zbiór danych.

Wymienić tu można (według nomenklatury programu STATISTICA):

- wykresy słupkowo/kolumnowe;
- wykresy liniowe;
- wykresy rozrzutu;
- wykresy obrazkowe.

Istotne jest to, że większość z tych wykresów jest „wrażliwa” na zmiany struktury zbioru danych – w szczególności sortowanie. Wyjątek to wykres rozrzutu. Dlatego przed przystąpieniem do sporządzania niektórych wykresów prezentacyjnych trzeba czasem odpowiednio przygotować arkusz (np. dane posortować).

Uwaga: Więcej szczegółów dotyczących możliwości modyfikacji wykresów omówionych będzie na wykładach.

Wykres słupkowo/kolumnowy (w wersji wielokrotnej)	Dane: Informacje o krajach UE-28
<p>Celem analizy jest prezentacja danych dotyczących <i>Stopy bezrobocia</i> w państwach UE w roku 2012 i 2021.</p> <p>PRZYGOTOWANIE DANYCH:</p> <ol style="list-style-type: none"> 1) Dane sortujemy malejąco według poziomu bezrobocia z 2012 r., w arkuszu zmieniamy na czerwony kolor nazwy przypadku <i>Polska</i>. 2) Wykonujemy <i>Wykres słupkowo/kolumnowy</i> i aktualizujemy go w celu wczytania koloru dla Polski. 3) Włączamy wszystkie nazwy krajów (!). 4) Przesuwamy legendę do górnego tytułu (aby przesunąć legendę odblokowujemy ją klikając prawym przyciskiem myszki i wskazując: <i>Zamień na pływający tekst</i>, następnie ponownie klikamy prawym przyciskiem myszki i wskazujemy <i>Zamień na tytuł...</i>) i edytujemy. 5) Nadajemy format i tytuł osi Y. 	<p>Bezrobocie w UE w roku 2012 i 2021:</p> <p>2012 2021</p>
Wykres liniowy zmiennych (w wersji podwójnej)	Dane: Transport w Polsce 1990-2022 (R)
<p>Celem analizy jest prezentacja zmian <i>liczby pasażerów przewożonych koleją w latach 1990-2022</i> w kontekście rozwoju motoryzacji, opisanego za pomocą <i>liczby zarejestrowanych samochodów osobowych</i>.</p> <p>W tym celu proszę wykorzystać wykres <i>Liniowy (zmiennych)</i> w wersji <i>Podwójnej</i> (zakładka: <i>Więcej</i>).</p> <p>Po sporządzeniu wykresu proszę dokonać formatowania wykresu, tak by wyglądał on jak ten zamieszczony na rysunku. Jeżeli tytuł i wartości osi ustawimy w tych samych kolorach, co odpowiadające im linie i znaczniki, legendę będzie można całkowicie usunąć.</p> <p>Aby uatrakcyjnić prezentację można włączyć pole wykresów i odpowiedni stopień przezroczystości oraz wypełnić tło przykładową grafiką znaną w Internecie.</p>	<p>Pasażerski transport kolejowy a rozwój motoryzacji indywidualnej w Polsce po 1990 r.</p> <p>Przewozy pasażerów (koleje) [mln osób] Samochody osobowe zarejestrowane [mln]</p> <p>Źródło danych: Główny Urząd Statystyczny (www.stat.gov.pl)</p>

STATYSTYKA OPISOWA (lab. 9)

GRAFICZNA ANALIZA DANYCH – część 1.

Wykres liniowy zmiennych Dane: Wybrane wskaźniki (Europa)

Celem analizy będzie porównanie poziomu życia w krajach europejskich w podziale na dwie grupy – państwa postkomunistyczne i pozostałe.

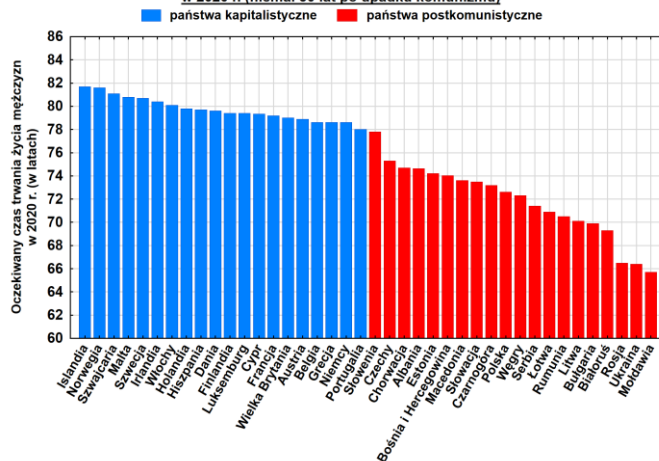
Sortujemy dane malejąco względem *Oczekiwanego czasu trwania życia mężczyzn* w 2010 roku.

Wykresem wyjściowym będzie *Wykres liniowy*, w którym jako główną zmienną wskazujemy *Oczekiwany czas trwania życia mężczyzn (2010)*. W zakładce *Skategoryzowane* włączamy *Kategorię X* i wybieramy zmienną *Historia*. Ponadto włączamy *Układ* w wersji *Nalozone*.

Zmieniamy wygląd wykresów wykonując następujące czynności:

- 1) umieszczamy legendę wykresu w tytule;
- 2) wyłączamy linie i znaczniki a włączamy słupki;
- 3) starannie formatujemy wykres.

Porównanie poziomu życia w państwach postkomunistycznych i kapitalistycznych w 2020 r. (niemal 30 lat po upadku komunizmu)



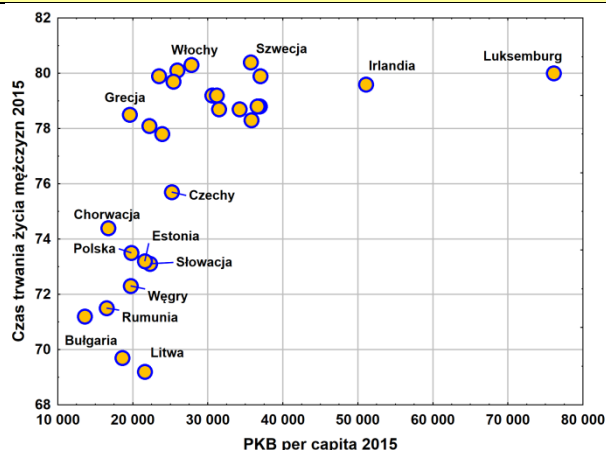
Wykres rozrzutu Dane: Informacje o krajach UE-28

Wykres rozrzutu pozwala analizować zależność pomiędzy dwiema wielkościami liczbowymi. Przedmiotem analizy jest wpływ *PKB per capita (2015)* na *Oczekiwany czas trwania życia (2015)*. Zmienne do wykresu wybieramy w takiej kolejności jak wymieniono je powyżej. W zakładce *Więcej* wyłączamy opcję *Dopasuj*.

Następnie formatujemy wykres (zwiększamy czcionkę i wielkość znaczników).

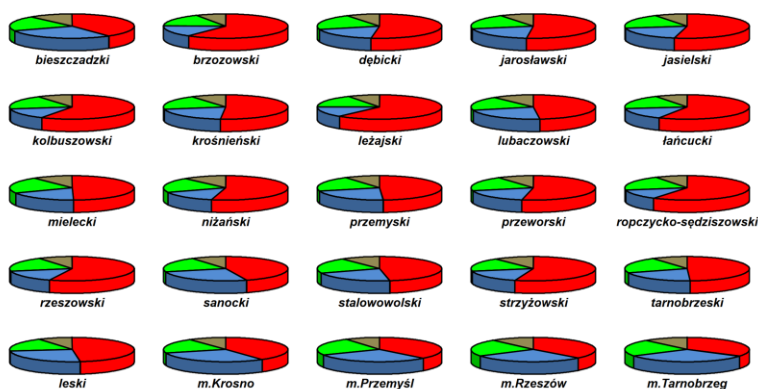
Na gotowym wykresie istnieje możliwość etykietowania (a także wyróżniania, usuwania, etc.) dowolnych znaczników za pomocą *Narzędzia wyróżniania*.

Narzędzie wyróżniania można włączyć po kliknięciu wykresu prawym przyciskiem myszki i wskazaniu opcji *Pokaż wyróżnianie*. Wybieramy opcje *Etykietuj* i *Zastosuj automatycznie*, po czym interaktywnie dowiadujemy się jakie kraje w Europie charakteryzują się pewnymi wartościami obu zmiennych.



Wykres obrazkowy Dane: Wybory 2015

Wyniki I tury wyborów prezydenckich 2015



Celem analizy jest prezentacja wyników I tury wyborów prezydenckich w powiatach woj. podkarpackiego trzech kandydatów, którzy osiągnęli najwyższy wynik (*Andrzej Duda*, *Bronisław Komorowski* i *Paweł Kukiz*) i skumulowanego wyniku pozostałych osób. Po pierwsze, na końcu arkusza dodajemy nową zmienną i nazywamy ją *Pozostali kandydaci*. Następnie wyliczamy jej wartości za pomocą odpowiedniej formuły. Zanim przystąpimy do sporządzania wykresu ze wszystkich powiatów musimy wyselekcjonować powiaty woj. podkarpackiego – kod województwa stanowi **2 i 3 cyfra kodu powiatu** (kod Podkarpacia to 18). Naciskamy klawisz *F8* i w oknie *Selekcji przypadków* wpisujemy warunek $\text{Mid}(v1;2;2)=^{18}$.

Teraz wybieramy polecenia *Wykresy / Wykresy obrazkowe* i wybieramy odpowiednie zmienne, ustalamy rodzaj wykresu na *Koła* i wykonujemy roboczą wersję wykresu. Wykres formatujemy: przenosimy legendę do tytułu, w zakładce *Okno* zmieniamy rozmiar na: *szerokość* – 20 cm, *wysokość* – 12 cm, w zakładce *Układ* ustawiamy liczbę *wierszy* – 5 i *kolumn* – 5. W zakładce *Etykiety punktów* włączamy podpisy powiatów.

W których powiatach Andrzej Duda otrzymał relatywnie najmniej głosów, w których ponad 50%?