

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**.

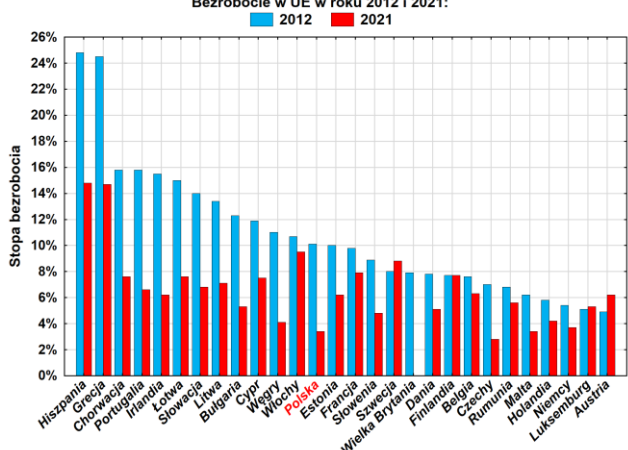
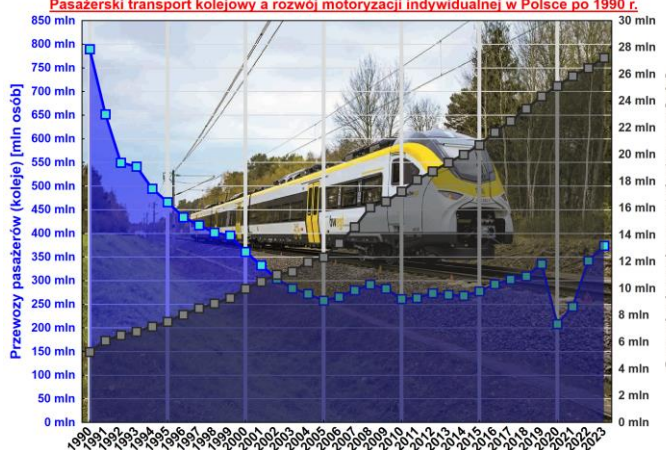
Na dzisiejszych zajęciach zapoznamy się z wybranymi wykresami prezentacyjnymi – jest to określenie, którego używać będziemy do wykresów odzwierciedlających w „dokładny” sposób strukturę danych. Będzie na nich tyle elementów (słupków, znaczników, linii), ile przypadków liczy zbiór danych.

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 Stopy bezrobocia 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>. <p>WYKONANIE I EDYCJA WYKRESU:</p> <ol style="list-style-type: none">2) Wykonujemy Wykres słupkowo/kolumnowy 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: Zamień na pływający tekst, następnie ponownie klikamy prawym przyciskiem myszki i wskazujemy Zamień na tytuł...) i edytujemy.5) Nadajemy format i tytuł osi Y.		<p>Bezrobocie w UE w roku 2012 i 2021:</p>  <table><caption>Przybliżone dane o stopie bezrobocia (%)</caption><tr><th>Kraj</th><th>2012</th><th>2021</th></tr><tr><td>Hiszpania</td><td>24.5</td><td>14.5</td></tr><tr><td>Grecja</td><td>24.5</td><td>14.5</td></tr><tr><td>Chorwacja</td><td>16.0</td><td>14.5</td></tr><tr><td>Portugalia</td><td>16.0</td><td>14.5</td></tr><tr><td>Irlandia</td><td>16.0</td><td>14.5</td></tr><tr><td>Łowia</td><td>15.0</td><td>14.5</td></tr><tr><td>Słowacja</td><td>14.0</td><td>14.5</td></tr><tr><td>Litwa</td><td>13.0</td><td>14.5</td></tr><tr><td>Bulgaria</td><td>12.0</td><td>14.5</td></tr><tr><td>Cypr</td><td>11.0</td><td>14.5</td></tr><tr><td>Węgry</td><td>10.0</td><td>14.5</td></tr><tr><td>Polska</td><td>10.0</td><td>14.5</td></tr><tr><td>Włochy</td><td>10.0</td><td>14.5</td></tr><tr><td>Estonia</td><td>10.0</td><td>14.5</td></tr><tr><td>Francja</td><td>10.0</td><td>14.5</td></tr><tr><td>Szwecja</td><td>10.0</td><td>14.5</td></tr><tr><td>Słowenia</td><td>10.0</td><td>14.5</td></tr><tr><td>Wielka Brytania</td><td>10.0</td><td>14.5</td></tr><tr><td>Dania</td><td>10.0</td><td>14.5</td></tr><tr><td>Finlandia</td><td>10.0</td><td>14.5</td></tr><tr><td>Belgia</td><td>10.0</td><td>14.5</td></tr><tr><td>Czechy</td><td>10.0</td><td>14.5</td></tr><tr><td>Rumunia</td><td>10.0</td><td>14.5</td></tr><tr><td>Malta</td><td>10.0</td><td>14.5</td></tr><tr><td>Holandia</td><td>10.0</td><td>14.5</td></tr><tr><td>Niemcy</td><td>10.0</td><td>14.5</td></tr><tr><td>Luksemburg</td><td>10.0</td><td>14.5</td></tr><tr><td>Austria</td><td>10.0</td><td>14.5</td></tr></table>		Kraj	2012	2021	Hiszpania	24.5	14.5	Grecja	24.5	14.5	Chorwacja	16.0	14.5	Portugalia	16.0	14.5	Irlandia	16.0	14.5	Łowia	15.0	14.5	Słowacja	14.0	14.5	Litwa	13.0	14.5	Bulgaria	12.0	14.5	Cypr	11.0	14.5	Węgry	10.0	14.5	Polska	10.0	14.5	Włochy	10.0	14.5	Estonia	10.0	14.5	Francja	10.0	14.5	Szwecja	10.0	14.5	Słowenia	10.0	14.5	Wielka Brytania	10.0	14.5	Dania	10.0	14.5	Finlandia	10.0	14.5	Belgia	10.0	14.5	Czechy	10.0	14.5	Rumunia	10.0	14.5	Malta	10.0	14.5	Holandia	10.0	14.5	Niemcy	10.0	14.5	Luksemburg	10.0	14.5	Austria	10.0	14.5
Kraj	2012	2021																																																																																								
Hiszpania	24.5	14.5																																																																																								
Grecja	24.5	14.5																																																																																								
Chorwacja	16.0	14.5																																																																																								
Portugalia	16.0	14.5																																																																																								
Irlandia	16.0	14.5																																																																																								
Łowia	15.0	14.5																																																																																								
Słowacja	14.0	14.5																																																																																								
Litwa	13.0	14.5																																																																																								
Bulgaria	12.0	14.5																																																																																								
Cypr	11.0	14.5																																																																																								
Węgry	10.0	14.5																																																																																								
Polska	10.0	14.5																																																																																								
Włochy	10.0	14.5																																																																																								
Estonia	10.0	14.5																																																																																								
Francja	10.0	14.5																																																																																								
Szwecja	10.0	14.5																																																																																								
Słowenia	10.0	14.5																																																																																								
Wielka Brytania	10.0	14.5																																																																																								
Dania	10.0	14.5																																																																																								
Finlandia	10.0	14.5																																																																																								
Belgia	10.0	14.5																																																																																								
Czechy	10.0	14.5																																																																																								
Rumunia	10.0	14.5																																																																																								
Malta	10.0	14.5																																																																																								
Holandia	10.0	14.5																																																																																								
Niemcy	10.0	14.5																																																																																								
Luksemburg	10.0	14.5																																																																																								
Austria	10.0	14.5																																																																																								
Wykres liniowy zmiennych (w wersji podwójnej)		Dane: Transport w Polsce 1990-2023 (R)																																																																																								
<p>Celem analizy jest prezentacja zmian liczby pasażerów przewożonych koleją w latach 1990-2023 w kontekście rozwoju motoryzacji, opisanego za pomocą liczby zarejestrowanych samochodów osobowych.</p> <p>W tym celu proszę wykorzystać wykres Liniowy (zmiennych) w wersji Podwójnej (zakładka: Więcej).</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 znalezioną w Internecie.</p>		<p>Pasażerski transport kolejowy a rozwój motoryzacji indywidualnej w Polsce po 1990 r.</p>  <table><caption>Przybliżone dane o transporcie w Polsce</caption><tr><th>Rok</th><th>Przewozy pasażerów (mln osób)</th><th>Samochody osobowe (mln)</th></tr><tr><td>1990</td><td>850</td><td>2.0</td></tr><tr><td>1995</td><td>750</td><td>3.0</td></tr><tr><td>2000</td><td>650</td><td>4.0</td></tr><tr><td>2005</td><td>550</td><td>5.0</td></tr><tr><td>2010</td><td>450</td><td>6.0</td></tr><tr><td>2015</td><td>350</td><td>7.0</td></tr><tr><td>2020</td><td>250</td><td>8.0</td></tr><tr><td>2023</td><td>200</td><td>9.0</td></tr></table> <p>Źródło danych: Główny Urząd Statystyczny (www.stat.gov.pl)</p>		Rok	Przewozy pasażerów (mln osób)	Samochody osobowe (mln)	1990	850	2.0	1995	750	3.0	2000	650	4.0	2005	550	5.0	2010	450	6.0	2015	350	7.0	2020	250	8.0	2023	200	9.0																																																												
Rok	Przewozy pasażerów (mln osób)	Samochody osobowe (mln)																																																																																								
1990	850	2.0																																																																																								
1995	750	3.0																																																																																								
2000	650	4.0																																																																																								
2005	550	5.0																																																																																								
2010	450	6.0																																																																																								
2015	350	7.0																																																																																								
2020	250	8.0																																																																																								
2023	200	9.0																																																																																								

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 2020 roku.

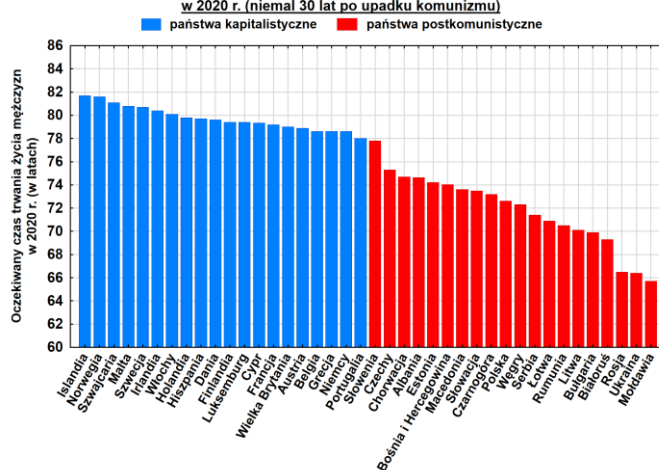
Wykresem wyjściowym będzie **Wykres liniowy**, w którym jako główną zmienną wskazujemy *Oczekiwany czas trwania życia mężczyzn (2010)*.

W zakładce **Skategoryzowane** włączamy **Kategorię X** i wybieramy zmienną *Historia*. Ponadto włączamy **Układ** w wersji **Nalozzone**.

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)



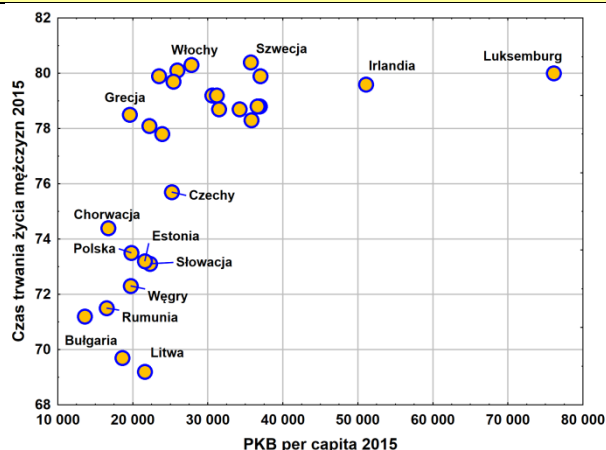
Wykres rozrzutu Dane: Informacje o krajach UE-28

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).

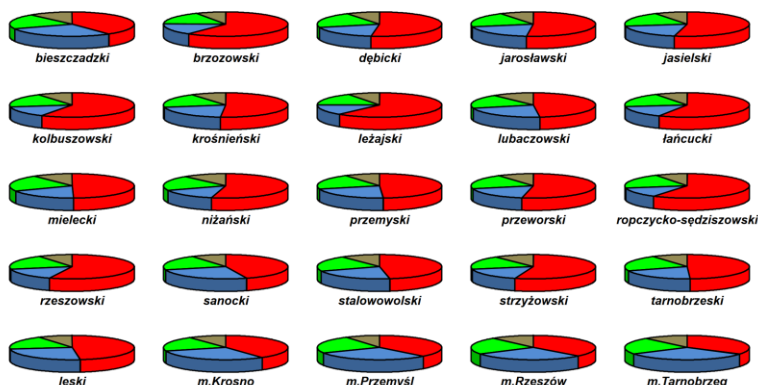
Na gotowym wykresie istnieje możliwość etykietowania (a także wyróżniania, usuwania, etc.) dowolnych znaczników za pomocą **Narzędzia wyróżniania**.

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
 ■ Andrzej Sebastian Duda ■ Bronisław Maria Komorowski
 ■ Paweł Piotr Kukiz ■ Pozostali kandydaci



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 **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%?