

STATYSTYKA OPISOWA (lab. 5)

CHARAKTERYSTYKA DANYCH LICZBOWYCH – STATYSTYKI OPISOWE

Przykład 1 (Bezrobocie w gminach Podkarpacia 2003-2022)

WYZNACZANIE I INTERPRETACJA STATYSTYK OPISOWYCH

Przedmiotem analizy będzie poziom bezrobocia w gminach woj. podkarpackiego w latach 2003-2022 – z podziałem względem płci. Proszę podać liczbę analizowanych przypadków – gmin w woj. podkarpackim: _____

Ponieważ dane mają charakter liczbowy do ich analizy wykorzystamy statystyki opisowe.

Za pomocą poleceń *Statystyka / Statystyki podstawowe i tabele / Statystyki opisowe* wyznaczmy średnią (\bar{x}), wartość środkową (medianę – Me), odchylenie standardowe (s), minimum i maksimum, kwartyl dolny i kwartyl górny (Q_1 i Q_3) dla stóp bezrobocia w woj. podkarpackim. W oknie wyboru zmiennej wskazujemy wszystkie zmienne ze wskaźnikami bezrobocia (od 2 do 40), a następnie w zakładce *Więcej* wybieramy wymienione powyżej statystyki.

Na podstawie uzyskanych rezultatów proszę uzupełnić luki w raporcie (wskazówka: żeby odpowiedzieć na część pytań trzeba będzie wykorzystać także inne, poznane już wcześniej sposoby analizy – na przykład sortowanie danych):

BEZROBOCIE OGÓŁEM

W roku 2003 średni wskaźnik bezrobocia w gminach woj. podkarpackiego wynosił _____%, w roku 2013 _____%, zaś w roku 2022 _____%.

Najniższy poziom bezrobocia w roku 2003 to _____%, a najwyższy _____%. W roku 2022 wartości skrajne wynosiły odpowiednio: _____ i _____%. Gmina o najwyższym poziomie bezrobocia w roku 2022 to: _____ a gmina o najniższym bezrobociu to _____.

W 2022 roku w co drugiej gminie wskaźnik bezrobocia był wyższy niż _____%, a w co czwartej wyższy niż _____%. W co czwartej gminie wskaźnik bezrobocia nie przekraczał _____%.

BEZROBOCIE WŚRÓD KOBIET I MĘŻCZYZN

W 2022 średni poziom bezrobocia wśród kobiet był o _____% wyższy niż wśród mężczyzn, w 2015 różnica ta wynosiła _____%, a w 2008 roku _____%. Czy nierówność kobiet i mężczyzn na rynku pracy zwiększa się czy zmniejsza w ostatnich latach?

Przykład 2 (Wynagrodzenia w powiatach 2002-2023)

MIARY ZMIENNOŚCI I ASYMETRII

Celem analizy jest opis przeciętnego poziomu płac w powiatach w roku 2003, 2008, 2013, 2018 i 2023. Dodatkowo wyznaczone zostaną także miary pokazujące zróżnicowanie regionalne płac w Polsce – współczynnik zmienności i współczynnik asymetrii i skośności.

A oto etapy rozwiązania zadania:

- Po wywołaniu okna analizy *Statystyki opisowe* wybieramy jako zmienne wskazane wyżej lata, a następnie w zakładce *Więcej* zaznaczamy statystyki takie jak w tabeli poniżej, po czym wywołujemy wyniki. Proszę uzupełnić poniższą tabelę.

Rok	\bar{x}	Me	min	max	s	V	A
2003							
2008							
2013							
2018							
2023							

\bar{x} – średnia arytmetyczna, s – odchylenie standardowe, V – współczynnik zmienności, A – skośność

- Proszę opisać zmiany w wynagrodzeniach w powiatach w latach 2002-2017 ze szczególnym uwzględnieniem regionalnego zróżnicowania, czyli współczynników V i A. O czym świadczą dodatnie wartości skośności?

.....
.....
.....
.....

STATYSTYKA OPISOWA (lab. 5)

CHARAKTERYSTYKA DANYCH LICZBOWYCH – STATYSTYKI OPISOWE

Przykład 3 (Stopa bezrobocia w wybranych krajach 2000-2022 (M))

WIZUALIZACJA STATYSTYK OPISOWYCH

Celem analizy będzie pokazanie zmian w poziomie bezrobocia w wybranych krajach europejskich w latach 2000-2022 ze szczególnym uwzględnieniem kryzysu „epidemicznego” w 2020 i 2021 roku.


Przygotowanie danych do analizy

Analiza dotyczyć będzie danych rocznych, zaś w pliku *Stopa bezrobocia w krajach UE 2000-2022 (M)* dane są zawarte w układzie miesięcznym, dlatego też najpierw wyliczymy średnie roczne stopy bezrobocia i dopiero te dane poddamy dalszej analizie. Czasem samo przygotowanie danych do analizy to pracochłonny, niełatwy ale niezbędny etap w pracy analityka danych.

Proszę wykonać następujące polecenia:

1. W arkuszu danych po zmiennej *DATA* proszę dodać dwie nowe zmienne i nazwać je *Rok* oraz *Miesiąc*.
2. Proszę zaznaczyć kolumnę *DATA* i za pomocą przycisku *Zmienne* znaleźć polecenie *Operacje na danych*, a następnie dokonać „rozbicia” dat na kolumnę miesiąc i rok.
3. Do wyliczenia średnich rocznych wykorzystamy analizę *Statystyka / Statystyki podstawowe i tabele / Przekroje (prosta ANOVA)*, w której jako zmienną grupującą wybieramy *Rok*, a jako zmienne zależne informacje o stopie bezrobocia ogółem w krajach UE i stowarzyszonych (zmienne od 4 do 34).
4. Klikamy *OK* i przechodzimy do okna *Statystyki w grupach – wyniki*, gdzie w zakładce *Statystyki opisowe*, odznaczamy wszystkie statystyki (średnia i tak jest liczona obligatoryjnie) i za pomocą przycisku *Podsumowanie* otrzymujemy tabelę ze średnimi rocznymi dla każdego kraju.
5. Aby zapisać uzyskane wyniki, jako plik danych, proszę w skoroszycie wyników wywołać menu podręczne (prawy przycisk myszki) i wybrać polecenie *Pobierz do oddzielnego okna / Przenieś*.
6. Teraz proszę wybrać polecenie *Dane / Arkusz wejściowy* żeby program „rozpoznał” arkusz wyników, jako wejściowy plik danych do przyszłych analiz.
7. Proszę zapisać ten plik pod nazwą ***Bezrobocie w krajach w latach 2000-2022***.

Aby zakończyć przygotowania pliku, należy jeszcze wykonać kilka czynności:

8. Usuujemy ostatni wiersz (przypadek) w arkuszu danych.
9. Za pomocą polecenia *Dane / Menedżer nazw przypadków* tworzymy nazwy przypadków kopiując lata z 1. zmiennej – następnie usuwamy zmienną z latami z arkusza.
10. Ponieważ przedmiotem analiz będą kraje w poszczególnych latach, musimy zamienić miejscami przypadki i zmienne w arkuszu danych – w tym celu wybieramy polecenia *Dane / Transponuj / Plik*.
11. Wartości w arkuszu formatujemy, w tym celu zaznaczamy cały arkusz, a następnie za pomocą przycisku  ustalamy format wyświetlania na jedno miejsce po przecinku.
12. Proszę usunąć z nazw przypadków dopisek „(O)”, pozostawiając tam tylko nazwę kraju – w tym celu proszę wykorzystać odpowiednie polecenie (patrz skróty klawiaturowe z zajęć nr 1).

Analiza statystyk opisowych

Za pomocą analizy *Statystyki opisowe* proszę wyznaczyć wartości statystyk, tak by uzupełnić poniższą tabelę.

Rok	\bar{x}	Me	min	max	s	V	A
2019							
2020							
2021							

Proszę wznowić analizę (Ctrl+R) i w zakładce *Opcje* ustawić typ wykresu ramka-wąsy na *Mediana / kwartyle / rozstęp* po czym w zakładce *Podstawowe* proszę wykonać ten wykres (i sformatować wykres, by był zbliżony do wzorca na rysunku obok – oczywiście aktualny wykres ma dotyczyć wszystkich lat z przedziału 2000-2022).

Umiejętności do sprawdzenia na kolokwium:

1. Wyliczanie statystyk opisowych.
2. Interpretacja wartości takich statystyk jak: N , \bar{x} , s , min, max, Me, Q_1 , Q_3 , V , A .
3. Wizualizacja statystyk opisowych za pomocą wykresów typu ramka-wąsy.

