**Zadanie 1** W celu analizy porównawczej trzech grup społeczno-ekonomicznych gospodarstw domowych (pracownicy - P, rolnicy - R, emeryci - E) ze względu na poziom zamożności zastosowano analizę korespondencji. Zamożność grup społeczno-ekonomicznych gospodarstw domowych (bardzo zamożni - BZ, zamożni - Z, niezamożni - NZ). W wyniku badania otrzymano m. in. informacje zawarte w tabeli 1 oraz na wykresie 1.

**Tabl. 1. Wartości własne i stopień wyjaśniania inercji przez czynniki.**

|  |  |  |  |
| --- | --- | --- | --- |
| **Czynniki** | **Wartości własne** | **Stopień wyjaśniania inercji** | **Skumulowany stopień wyjaśniania inercji** |
| F1 | 0,115 | 92,443 | 92,443 |
| **F2** | 0,010 | 7,557 | 100,0 |

**Wykres 1. Wykres konfiguracji punktów reprezentujących grupy społeczno-ekonomiczne oraz poziom zamożności w dwuwymiarowej przestrzeni czynnikowej.**

 **F2**

 +R

 οN

 οBZ

 **F1**

 οZ +E

 +P

1. ocenić znaczenie poszczególnych czynników w wyjaśnianiu zróżnicowania grup pracowniczych ze względu na poziom zamożności.
2. przeprowadzić analizę porównawczą grup pracowniczych ze względu na poziom zamożności.

**Zadanie 2**

Dla dziesięciu krajów Afrykańskich dane są wartości wybranych zmiennych opisujących ich stopień rozwoju społeczno-ekonomicznego (kolejno: średnie zbiory zbóż z ha, wskaźnik percepcji korupcji, ogólny wskaźnik umieralności, liczba serwerów internetowych, odsetek pracujących z wykształceniem wyższym, oczekiwana długość życia, umieralność niemowląt oraz odsetek osób korzystających z Internetu). Zmienne zostały wystandaryzowane. Dana jest również macierz korelacji pomiędzy zmiennymi.

**Tab. 1. Wystandaryzowane wartości wskaźników.**



**Tab. 2. Macierz współczynników korelacji.**



Dokonaj wyboru zmiennych reprezentantów dowolną metodą taksonomiczną. Określając graniczną wartość współczynnika korelacji liniowej użyj formuły minimaksowej.

**Zadanie 3**

Wykorzystując dane z poprzedniego zadania dokonaj porządkowania liniowego krajów dowolną metodą wzorcową. W procedurze porządkowania wykorzystuj tylko zmienne reprezentanty wybrane w zad. 2.

**Zadanie 4**

Wykorzystując wyniki porządkowania krajów z poprzedniego zadania dokonaj ich grupowania dowolną metodą. Porównaj utworzone grupy państw ze względu na przeciętne wartości wskaźników reprezentantów wybranych w zad. 2.

# Zadanie 5

Poddano analizie głównych składowych 5 cech charakteryzujących 100 marek samochodów, a mianowicie:

liczba cylindrów, moc, przyspieszenie, cena, waga.

Uzyskano m.in. następujące informacje przedstawione w formie tabelarycznej:

|  |  |  |
| --- | --- | --- |
| **Numer głównej składowej** | **% wyjaśnionej wariancji** | Ładunki dwóch pierwszych głównych składowych |
| 1 | 54.60660 | (1,1) 0.149508 | (1,2) 0,9016535 |
| 2 | 23.18331 | (2,1) 0.174869 | (2,2) -0.926433 |
| 3 | 17.11052 | (3,1) -0.240566 | (3,2) 0.862204 |
| 4 | 3.75920 | (4,1) 0.888968 | (4,2) 0.306372 |
| 5 | 1.34036 | (5,1) 0.909356 | (5,2) 0.169553 |

1. oceń zasoby informacyjne poszczególnych głównych składowych,
2. wyodrębnij grupy cech reprezentowanych przez dwie pierwsze główne składowe oraz oceń znaczenie tych składowych w opisie wyróżnionych w badaniu cech,
3. na podstawie znanych kryteriów zdecyduj o liczbie głównych składowych, jakie należy zachować w analizie.

# Zadanie 6

Poddano analizie wartości dwóch grup zmiennych opisujących 150 marek samochodów (4 charakteryzujące parametry techniczne samochodów oraz 2 charakteryzujące koszty ich nabycia i eksploatacji). Uzyskano m. in. następujące informacje przedstawione w formie tabelarycznej.

1. ocenić znaczenie par zmiennych kanonicznych w opisie wyodrębnionych grup zmiennych,
2. zapisać równania dla pierwszej pary zmiennych kanonicznych oraz dokonać ich interpretacji.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Numer pary zmiennych kanonicznych** | **Wartość** **własna** | **Korelacja kanoniczna** | **Statystyka Lambda** | **Statystyka χ2** | **Liczba stopni swobody** | **Poziom istotności** |
| 1 | 0,7738 | 0,8797 | 0,2089 | 227,83 | 8 | 0,0000 |
| 2 | 0,0763 | 0,2762 | 0,9237 | 11,55 | 3 | 0,0091 |
| Współczynniki zmiennych kanonicznych pierwszej grupy zmiennych |
| CAR. liczba cylindrów | 0,27214 | -1,26413 |  |
| CAR. moc | -0,20923 | -1,11814 |  |
| CAR. przyspieszenie | 0,13753 | -0,17611 |  |
| CAR. waga | -1,02008 | 1,98415 |  |
| Współczynniki zmiennych kanonicznych drugiej grupy zmiennych |
| CAR. zużycie paliwa\*) | 0,97283 | 0,23178 |  |
| CAR. cena | -0,24218 | 0,97029 |  |

\*) liczba przejechanych mil na 1galonie paliwa

**Zadanie 7**

Przeprowadzono ocenę zdolności dyskryminacyjnej krajów, ze względu na ich położenie (Europa, Ameryka Południowa, Azja), 11 wybranych wskaźników poziomu rozwoju społeczno-ekonomicznego wykorzystując metodę krokową postępującą. W kroku 5 analizy uzyskano poniższe wyniki:

Charakterystyki zmiennych dyskryminujących wprowadzonych do modelu oraz poza modelem (metoda krokowa postępująca) w kroku 5.

**Tab. 1. Zmienne poza modelem.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Zmienna** | **Lambda Wilksa** | **Lambda Wilksa Cząstkowa** | **F wprowadzenia** | **df 1** | **df 2** | **Pr>F** |
| WSK\_1 | 0,6717 | 0,9987 | 1,56 | 2 | 2406 | 0,2114 |
| WSK\_2 | 0,6756 | 1,0000 | 0,00 | 2 | 2406 | 0,9990 |
| WSK\_5 | 0,6625 | 0,9858 | 12,76 | 2 | 2406 | 0,0000 |
| WSK\_7 | 0,6767 | 1,0000 | 0,02 | 2 | 2406 | 0,9760 |
| WSK\_8 | 0,6766 | 0,9996 | 0,54 | 2 | 2406 | 0,5855 |
| WSK\_11 | 0,6728 | 0,9999 | 0,08 | 2 | 2406 | 0,9237 |

**Tab. 2. Zmienne w modelu.**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Zmienna** | **Lambda Wilksa** | **Lambda Wilksa Cząstkowa** | **F wprowadzenia** | **df 1** | **df 2** | **Pr>F** |
| WSK\_3 | 0,6769 | 0,7632 | 373,49 | 2 | 2407 | 0,0000 |
| WSK\_6 | 0,6769 | 0,8491 | 231,88 | 2 | 2407 | 0,0000 |
| WSK\_10 | 0,6769 | 0,9484 | 65,44 | 2 | 2407 | 0,0000 |
| WSK\_9 | 0,6769 | 0,9575 | 53,40 | 2 | 2407 | 0,0000 |
| WSK\_4 | 0,6769 | 0,9749 | 30,95 | 2 | 2407 | 0,0000 |

1. Czy w kolejnym kroku do modelu wejdzie kolejna zmienna? Jeżeli „tak”, to która z nich i dlaczego?
2. Porównać moc dyskryminacyjną zmiennych znajdujących się już w modelu.

**Zadanie 8**

W wyniku klasyfikacji do typów gmin (gminy miejskie, gminy miejsko-wiejskie, gminy wiejskie), za pomocą funkcji klasyfikacyjnych, 100 gmin otrzymano następujące wyniki w postaci macierzy klasyfikacji.

|  |  |  |
| --- | --- | --- |
| **Typy gmin** | **Klasyfikacja gmin** | **Ogółem** |
| **miejskie** | **miejsko-wiejskie** | **wiejskie** |
| miejskiemiejsko-wiejskiewiejskie | 2051 | 8109 | 2540 | 302050 |
| ogółem | 26 | 27 | 47 | 100 |

Ponadto wartość współczynnika *tau* będącego miarą poprawności klasyfikacji wyniosła 0,56.

1. zinterpretować wyniki klasyfikacji uzyskane w postaci macierzy klasyfikacji,
2. zinterpretować wartość współczynnika *tau*.

**Zadanie 9**

1. Na podstawie danych o charakterystykach samochodów z polskiego rynku motoryzacyjnego (po standaryzacji) zbudowano macierz odległości dla 9 samochodów. Przeprowadź grupowanie nieliniowe samochodów rysując odpowiedni dendrogram (drzewko połączeń).



Wyjaśnij które z metod definiowania odległości pomiędzy grupami obiektów przy tworzeniu dendrogramu można zastosować dla powyższych danych, a zastosowanie których z tych metod wymagałoby posiadania dodatkowych informacji.

**Zadanie 10**

Poniższa tabela zawiera pewne charakterystyki samochodów z polskiego rynku motoryzacyjnego.



1. Przeprowadź standaryzację dowolnie wybranej zmiennej (charakterystyki samochodów).
2. Zbiór zmiennych podziel na stymulanty i destymulanty z punktu widzenia nabywcy samochodu i uzasadnij ten podział.

**Zadanie 11**

Przeprowadzono analizę korespondencji pomiędzy preferencjami politycznymi (preferowana partia polityczna) oraz poziomem wykształcenia (podstawowe, zasadnicze zawodowe, średnie, wyższe). Wyniki w formie graficznej przedstawiono poniżej. Przeprowadź analizę porównawczą poparcia dla partii politycznych w zależności od poziomu wykształcenia.



**Zadanie 12**

Dane są następujące zmienne opisujące stopień adaptacji technologii informatycznych w społeczeństwach i gospodarkach państw europejskich:

X1 – odsetek osób korzystających z komputera,

X2 – odsetek osób korzystających z Internetu,

X3 – odsetek populacji zatrudnionej, jako specjaliści IT,

X4 – odsetek firm zatrudniających specjalistów IT,

X5 – odsetek firm mających system informatyczny wspomagający bieżącą działalność,

X6 – udział handlu w Internecie w całym handlu detalicznym,

X7 – odsetek osób korzystających z Internetu mobilnego.

Poniższa tabela zawiera współczynniki korelacji powyższych zmiennych dla państw Europy Zachodniej:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|   | x1 | x2 | x3 | x4 | x5 | x6 | x7  |
| x1 | 1,00 | 0,97 | 0,65 | -0,36 | 0,65 | -0,51 | 0,37 |
| x2 | 0,97 | 1,00 | 0,69 | -0,18 | 0,57 | -0,35 | 0,39 |
| x3 | 0,65 | 0,69 | 1,00 | 0,01 | 0,21 | -0,46 | 0,61 |
| x4 | -0,36 | -0,18 | 0,01 | 1,00 | -0,07 | 0,71 | -0,32 |
| x5 | 0,65 | 0,57 | 0,21 | -0,07 | 1,00 | -0,35 | -0,10 |
| x6 | -0,51 | -0,35 | -0,46 | 0,71 | -0,35 | 1,00 | -0,47 |
| x7 | 0,37 | 0,39 | 0,61 | -0,32 | -0,10 | -0,47 | 1,00 |

Dane są również współczynniki zmienności:

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| stats | x1 | x2 | x3 | x4 | x5 | x6 | x7 |
| cv | 0,07 | 0,07 | 0,19 | 0,22 | 0,22 | 0,32 | 0,60 |

Kolejna tabela zawiera wartości zunitaryzowanych zmiennych dla poszczególnych państw Europy Zachodniej.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  Zmienna | x1uni | x2uni | x3uni | x4uni | x5uni | x6uni | x7uni |
| Holandia | 0,94 | 0,89 | 0,76 | 0,75 | 1,00 | 0,43 | 0,56 |
| Wlk\_Brytania | 0,90 | 0,86 | 0,76 | 0,54 | 0,43 | 0,54 | 0,58 |
| Luksemburg | 0,96 | 0,95 | 0,55 | 0,63 | 0,94 | 0,71 | 0,15 |
| Austria | 0,78 | 0,65 | 0,61 | 0,58 | 0,73 | 0,39 | 0,13 |
| Francja | 0,73 | 0,59 | 0,42 | 0,33 | 0,61 | 0,46 | 0,44 |
| Irlandia | 0,59 | 0,49 | 0,47 | 1,00 | 0,45 | 1,00 | 0,17 |
| Niemcy | 0,86 | 0,70 | 0,55 | 0,42 | 0,78 | 0,54 | 0,21 |
| Belgia | 0,78 | 0,68 | 0,49 | 0,83 | 0,73 | 0,71 | 0,06 |

1. Dokonaj wyboru zmiennych reprezentantów metodą taksonomiczną.
2. Na podstawie wybranego zbioru zmiennych uporządkuj liniowo państwa dowolną metodą.
3. Wykonaj grupowanie państw dowolną metodą. Porównaj otrzymane grupy ze względu na stopień adaptacji technologii informatycznych.