

Podstawy statystyki praktycznej

Lista nr 1

1. Zbiór danych *table1_6.txt* zawiera dane 78 uczniów siódmej klasy szkoły amerykańskiej. Dane zawierają średnią ocen, wynik standardowego testu IQ, płeć i wynik testu psychologicznego samo-oceny Piersa-Harrisa.
 - a) Narysuj histogramy wszystkich zmiennych ilościowych. Opisz kształt, położenie i rozrzut ich rozkładów - podaj wartości odpowiednich statystyk.
 - b) Za pomocą wykresów ramkowych porównaj rozkłady zmiennych ilościowych dla chłopców i dziewczynek. Podsumuj wyniki.
 - c) Narysuj wykresy rozrzutu średniej z ocen w funkcji IQ i samo-oceny. Podsumuj wyniki.
2. Zbiór danych *individuals.dat* zawiera dane 55 899 ludzi, zebrane w oparciu o ankietę z 2000 roku. Dane zawierają wiek (w latach), wykształcenie: 1= podstawowe, 2=zawodowe, 3=średnie, 4=szkoła pomyślna, 5=licencjat, 6=wyższe, płeć o wartościach: 1=mężczyzna, 2=kobieta, dochód i sektor zatrudnienia o wartościach :5=sektor prywatny, 6=sektor rządowy, 7=samo-zatrudnienie.
 - a) Narysuj wykresy kołowe reprezentujące rozkład wykształcenia i sektora zatrudnienia u kobiet i mężczyzn. Podsumuj wyniki.
 - b) Narysuj wykresy słupkowe przedstawiające średni dochód jako funkcję wykształcenia i sektora zatrudnienia, osobno dla kobiet i mężczyzn. Podsumuj wyniki.
 - c) Narysuj histogram dochodów. Opisz jego kształt, położenie i rozrzut - podaj wartości odpowiednich statystyk. Porównaj średnią i medianę - wyznacz iloraz średniej i mediany.
 - d) Utwórz nową zmienną - logarytm z dochodów (wyklucz osoby z dochodem równym zero) i wykreśl jej histogram. Porównaj jego kształt z histogramem uzyskanym w punkcie c). Porównaj średnią i medianę dla przekształconego dochodu - wyznacz iloraz średniej i mediany.
 - e) Narysuj wykres rozrzutu dochodu w funkcji wieku - osobno dla mężczyzn i kobiet. Podsumuj wyniki.

Zadania do policzenia ręcznego

3. Zbiór danych zawiera siedem wartości: 6.8 5.3 6.0 5.9 6.8 7.4 6.2
 - a) Oblicz średnią, odchylenie standardowe i współczynnik zmienności.
 - b) Oblicz medianę i rozstęp międzykwartyłowy.
 - c) Zastąp obserwację 7.4 przez 15.4 i powtórz punkty a) i b). Które z obliczonych statystyk są odporne a które nie ?

4. Średnia i odchylenie standardowe 47 pomiarów temperatury w stopniach Celsjusza są równe: $\bar{y} = 36.497, s = 0.172$.

Oblicz średnią, odchylenie standardowe i współczynnik zmienności tych samych pomiarów wyrażonych w stopniach Fahrenheita.

Małgorzata Bogdan