Eksploracja danych

Ćwiczenia nr 9

Plik ***karp.csv*** zawiera dane z baz GUS odnośnie przeciętnych cen detalicznych żywności w Polsce w latach 1999-2019. Poszczególne zmienne zawierają informację o cenie następujących produktów:

* *ryz* – ryż za 1 kg,
* *maka* - mąka pszenna za 1 kg,
* *kurczak* – kurczęta patroszone za 1 kg,
* *kielbasa* – kiełbasa wędzona za 1 kg,
* *jaja* – jaja kurze świeże za 1 szt.,
* *maslo* – masło świeże o zawartości tłuszczu ok. 82,5%, za 200 g,
* *olej* – olej rzepakowy produkcji krajowej za 1 l,
* *cukier* – cukier biały kryształ za 1 kg,
* *miod* – miód pszczeli za 400 g,
* *kawa* – kawa naturalna mielona za 250 g,
* *herbata* – herbata czarna liściasta za 100 g,
* *karp* – karp świeży za 1 kg.
  1. Oblicz korelacje między cenami wszystkich produktów. Z ceną jakiego produktu jest najbardziej skorelowana cena karpia? Ceny których innych produktów są ze sobą skorelowane?
  2. Wykonaj wykresy rozrzutu lub macierz wykresów, by zwizualizować zależność ceny karpia od cen pozostałych produktów.
  3. Wykonaj analizę regresji liniowej. Zwróć uwagę na założenia: niezależność, normalność i homoskedastyczność reszt oraz występowanie obserwacji odstających.
  4. Wykonaj trójwymiarowy wykres rozrzutu, na którym przedstawisz zależność ceny karpia od jej dwóch najważniejszych predyktorów. Czy obserwacje leżą na jednej płaszczyźnie?