Zadanie analityczne

Wiosna 2022

Plik *zadnie\_analityczne.xlsx* zawiera dane z lat 2018-2020 dotyczące przychodów i kosztów przedsiębiorstw działających w trzch sektorach gospodarki. Twoim zadaniem jest analiza rentowności przedsiębiorstw w poszczególnych sektorach.

W tym celu:

1. wczytaj dane z pliku *zadanie\_analityczne.xlsx*;
2. dla każdej obserwacji empirycznej (tj. danego przedsiębiorstwa w danym roku) oblicz wskaźnik rentowności (udział dochodów w przychodach);
3. dla każdego przedsiębiorstwa oblicz średnią rentowność w latach 2018-2020 (pamiętaj o obsłudze braków danych);
4. przy użyciu wybranej metody dokonaj detekcji – osobno dla poszczególnych sektorów – obserwacji nietypowych;
5. oczyść dane z obserwacji nietypowych (poprzez winsoryzację lub ich usunięcie);
6. dokonaj podsumowania rozkładu rentowności w poszczególnych sektorach: oblicz średnią, średnią winsorowską (jeżeli zdecydowałaś/zdecydowałeś się na imputację obserwacji odstających), średnią trymowaną (jeżeli zdecydowałeś się na usunięcie obserwacji odstajacych), medianę, wybrane precentyle, odchylenie standardowe itd.
7. odpowiedz na pytanie: w jakim stopniu rozkłady wskaźników rentowności w poszczególnych sektorach są zbliżone do rozkładu normalnego?
8. odpowiedź na pytanie: w którym sektorze powinna zainwestować osoba, biorąca pod uwagę średnią wartość wskaźników rentowności (z uwzględnieniem obserwacji nietypowych oraz przy ich winsoryzacji/usunięciu)?

Obliczenia do raportu należy przeprowadzić w pakiecie Stata lub środowisku R. Do raportu należy dodać wydruk skryptu .do lub .r. Raport wraz z wydrukiem skryptu należy przesłać w formacie pdf.