

Упатство за изработка на проектот по предметот Бизнис статистика

Вовед

Во пакетот „R“ да се изработат барањата наведени во деловите А и Б.

Секој студент избира доволно големо податочно множество (најмалку неколку стотици редици) кое содржи најмалку две непрекинати квантитативни обележја. Изборот го прави од предлог изворите на податочни множества објавени на Courses или од друга релевантна база на податочни множества.

При изработка на секоја од задачите, студентот исклучиво ги користи само податоците од податочното множество што го избрал, освен во задача 2 од првиот дел.

По изработката на семинарската работа, сите добиени резултати (вредности и визуелизации) треба да се вклучат во една *.pdf* датотека (зачувана како *broj_na_indeks.pdf* со наслов „Семинарска работа по предметот Бизнис статистика“. На самиот почеток од оваа датотека, пред сите резултати, треба да се наведе изворот на конкретното податочното множество, и да се даде линкот до истото. Потоа, истото накратко да се објасни (колку единици и обележја има, и кои се тие). Секој резултат треба детално да биде објаснет, кои податоци од податочното множество се употребени (тука студентите имаат право на избор) и за секој резултат да се даде толкување што претставува. Со еден збор, семинарската работа треба да претставува една осмислена целина.

Финално, секој студент треба да прикачи една *.zip* архива која ќе содржи 2 датотеки, *.pdf* датотеката, како и R скриптата (датотека со *.R* екстензија) со вашиот код.

А. Прв дел

За податоците за соодветно избрани две обележја од Вашето податочно множество:

1. Да се изработи табела со распределба на честоти. Да се определат средните точки на интервалите, релативните фреквенции и кумулативните фреквенции. Потоа, добиените вредности графички да се претстават со хистограми и полигони.
2. Користејќи избрани податоци од примерокот (или ако тие не соодветствуваат изберете други, соодветни податоци) конструирајте стебло-лист дијаграм.
3. Да се нацрта графика на расејување за податоците од двете обележја и да се дискутира односот (врската) помеѓу двете обележја.
4. Да се определи мода, медијана и просек на податоците.

5. Да се определат кварталите, опсегот и интеркварталниот распон на податоците.
6. Да се пресмета дисперзија и стандардна девијација.
7. Да се определи коефициент на корелација.

Б. Втор дел

Врз основа на податоците за едно обележје (ако е неопходно за две обележја) од избраното податочно множество да се:

1. Одреди интервал на доверба за параметар што е најсоодветен;
2. Поставаат и тестираат соодветни хипотези за тестираниот параметар;
3. Направи тест за распределба;
4. Тестираат хипотези за независност. Во случај тоа да не е можно, да се образложи зошто не е можно;
5. Направи регресиона анализа.