Домашна работа #2

Диляна Пламенова Тодорова

Информационни системи, 2 курс, 71128

1зад

Таблица за оценките, получени при провеждане на студентски изпити

А) въвеждане на данните под формата на вектори

Първи начин за въвеждане на вектори, чрез директна заявка

Q1=c(3,3,3,4,3,4,3,4,3,4)

Втори начин, чрез използване на функцията scan(), която чете и записва данните до въвеждането на два ентъра

Q2=scan()

5 2 5 5 2 2 5 5 4 2

Трети начин, чрез функцията data.entry(), която извиква таблица за редактирането на данните.

За да може да се използва посочената функция за Q3 се задават стойности от едно до десет, при самото му дефиниране, след което се извиква функцията за редактиране и в таблицата се записват нужните данни.

Q3=1:10

Data.entry(Q3)

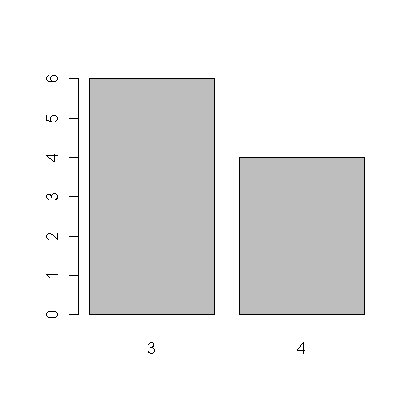
Направянето на таблици с въведените данни

tq1 = table(Q1)

tq2 = table(Q2)

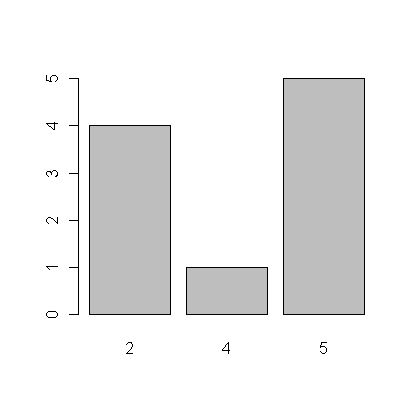
tq1q2 = table(Q1, Q2)

Б) Извикването на barplot и stacked barplot

barplot(tq1) 

Ясно се забелязва, че при първия тест имаме повече 3ки, отколкото 4ки, при представянето на учениците

barplot(tq2)

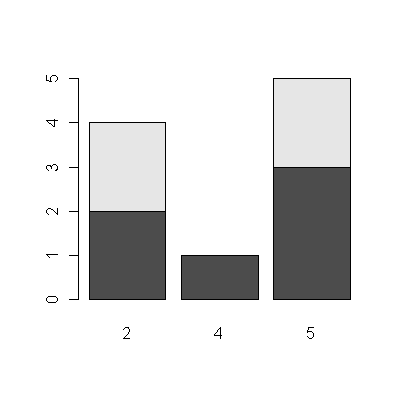


При втория проведен тест имаме най-много петици и двойки, докато средната оценка четири е най-слабо застъпена. Получените резултати не съответстват на нормалното разпределение, при което имаме най-много от средните стойности и малко количество на крайните стойности (каквито са двойките и петиците).

tq1q2 = table(Q1, Q2)

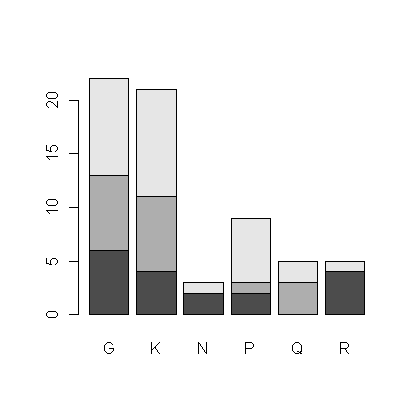
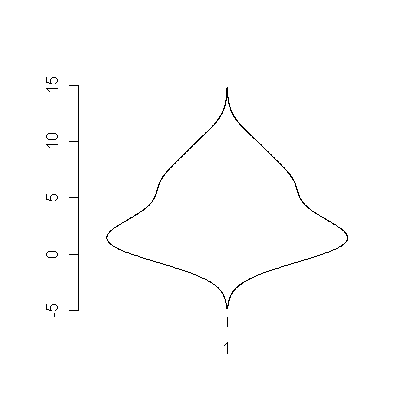
barplot(tq1q2)

Изчертаване на обща таблица, показваща отношението между двата теста. Забелязва се, че имаме повече от силните оценки, както и на слабите, но лепсват средни.

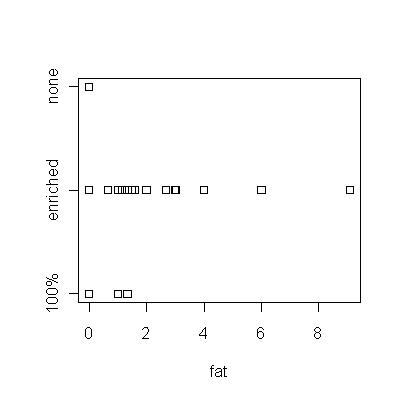
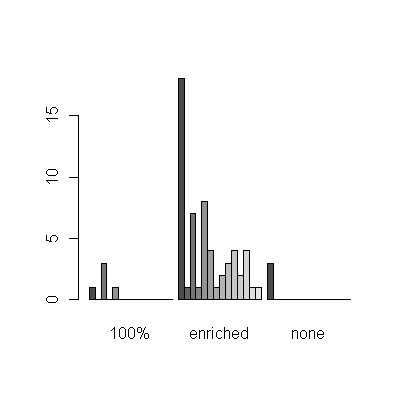


2 зад

Отношението на производство на зърнени закуски спрямо продажбите

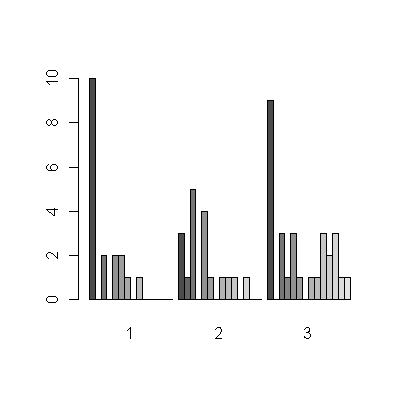


Компаниите, покриващи и трите категории продукти, реализират най-големи продажби. Най-ниски са на компанията предлагаща само нисък и среден клас (N). Това наподобява нормално разпределение.



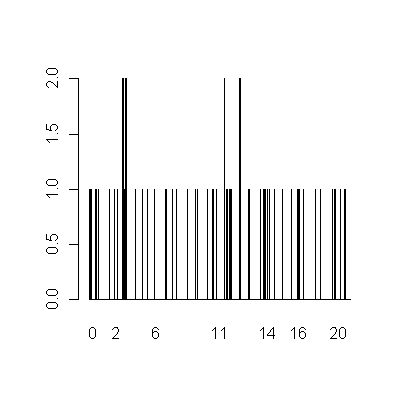
stripchart(fat~vitamins)

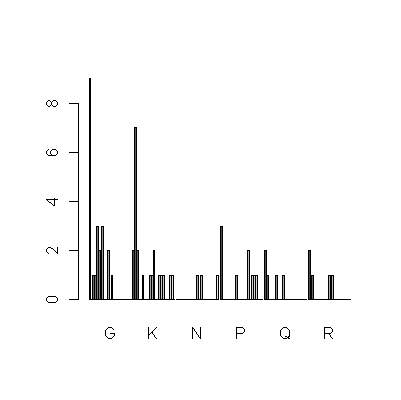
Виждаме, че обагатените съдържат в определени случай най-много мазнини, а в други продукти най-много витамини.

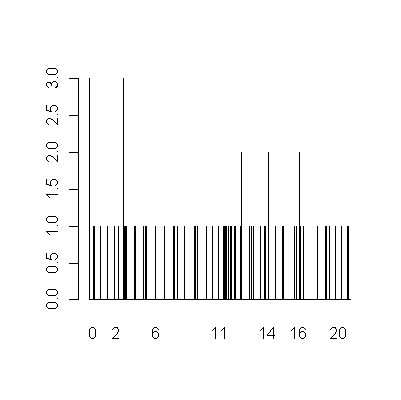
Мазнини спрямо рафт

barplot(table(fat,shelf), beside=T)

Най-продаваните и най-слабо продаваните са със високо съдържание на мазнини.

въглехидрати спрямо захари

фибри спрямо производство

натрий спрямо захари

3 зад

> cor(age,wt)

[1] 0.02904064

> cor(age,wt)^2

[1] 0.000843359

Коефицентите на корелация ρ и ρ2 са със стойности близки до нулата, следователно са слабо зависими, тоест не се установява връзка между теглото и годините на бебетата.

> cor(ht,wt)

[1] 0.1255413

> cor(ht,wt)^2

[1] 0.01576061

Между височината и теглото се установява слаба зависимост. Стойностите не са съвсем близко до 1, така че не може да твърдим че та са силно зависими.

> cor(rank(ht),rank(wt))

[1] 0.2147450

> cor(rank(ht),rank(wt))^2

[1] 0.0461154

Коефицентът Спиърман е спрямо векторите подредени изравнени по големината им. Отново може да заключим че са слабо свързани, тъй като са близки повече до нулата.

