Factors

Practice

1. Suppose you ran a survey where you asked 10 respondents about their favorite sports shoes. The possible answers were Adidas, Nike and Reebok, coded with 1, 2 and 3, respectively. You got the following answers:

   3, 2, 2, 1, 3, 2, 1, 1, 3, 1

   Create a factor with the levels above and the corresponding sports shoes brands as labels

2. Recode the shoes brands at #1 as follows: 1 – Reebok, 2 – Adidas, 3 – Nike.

3. Add a new level to the factor created at #1: the brand Puma, coded with 4. Afterwards add two components with this value.

4. Suppose you ran a survey where you asked 15 respondents about their satisfaction with hotel services. The satisfaction was measured on four levels: not satisfied, somewhat satisfied, satisfied and very satisfied, coded with 1, 2, 3 and 4, respectively. You got the following answers:

   4, 2, 2, 3, 1, 3, 4, 3, 2, 1, 4, 4, 3, 2, 2

   Create an ordered factor that contains these answers and their labels.

5. Add a new level to the factor created at #4: extremely satisfied, coded with 5. Afterwards add two components with this value.

6. Create a vector x of 30 discrete random numbers between 1 and 100, with or without replacement (your choice). Next, create a vector of 30 discrete random numbers between 1 and 4 (with replacement, of course) and convert this vector into a factor f. Afterwards do the following operations:

   - get the values in the vector x for each level of the factor f

- get the mean and standard deviation of the values in x for each level of the factor f, using the tapply() function
- get the mean and standard deviation of the values in x for each level of the factor f, using the by() function

7. Create another vector of 30 discrete random numbers between 1 and 3 and convert it into a factor g. Then perform again the operations at #6, this time using the levels of the two factors, f and g.

8. Explain the difference between the functions tapply() and by().