

amazon_analysis

PR

11/05/2020

```
llibres <- read.csv("llibres.csv", header = TRUE, sep = "|")
```

```
llibres <- llibres[,0:7]
```

```
head(llibres)
```

```
##
## 1 Aprendiendo a repasar: L neas Formas Letras: Libro de actividades para ni os: +3 a os: Un cuadern
## 2
## 3
## 4
## 5
## 6
##
##              Autor              Puntuacio Num_resenyes
## 1   June & Lucy Kids 4,2 de un m ximo de 5 estrellas      91.000
## 2   Javier Castillo 4,4 de un m ximo de 5 estrellas      95.000
## 3 Pegasus Coloring Book 4,5 de un m ximo de 5 estrellas     184.000
## 4   Teresa Blanch 4,6 de un m ximo de 5 estrellas     437.000
## 5   Juan G mez-Jurado 4,5 de un m ximo de 5 estrellas       2.935
## 6   Papeterie Bleu 4,5 de un m ximo de 5 estrellas     181.000
##
##      Categoria      Preu      Paginas
## 1 Tapa blanda  5,99     92 p ginas
## 2 Tapa blanda 17,95    512 p ginas
## 3 Tapa blanda  6,99    168 p ginas
## 4 Tapa blanda  5,65     96 p ginas
## 5 Tapa dura   19,85    568 p ginas
## 6 Tapa blanda  6,59     66 p ginas
```

```
summary(llibres)
```

```
##
## 100 Animal Mandalas Para Colorear: Libro para colorear para adultos con patrones de animales y mand
## 101 casos extraordinarios para resolver en 5 minutos (Serie   Qu  n ha sido? 1)
## 1984 (Contempor  nea)
##      qu   sabe la luna? (libros para so  ar)
##    Puedo mirar tu pa  al? (Libros de cart  n)
## Adivinanzas para leer antes de dormir (Adivina adivinanza)
## (Other)
##
##              Autor              Puntuacio
## Papeterie Bleu      : 9   4,6 de un m ximo de 5 estrellas:19
## June & Lucy Kids     : 5   4,8 de un m ximo de 5 estrellas:19
## Roberto Santiago    : 5   4,5 de un m ximo de 5 estrellas:18
## Jeff Kinney         : 4   4,7 de un m ximo de 5 estrellas:17
## El sabet Benavent   : 3   4,4 de un m ximo de 5 estrellas: 5
## Ram  n Rubio Silvestre: 3   4,9 de un m ximo de 5 estrellas: 5
```

```
## (Other) :71 (Other) :17
## Num_resenyes Categoria Preu
## Min. : 1.079 Artículos de papelería: 2 9,45â : 8
## 1st Qu.: 20.500 CÃ³mic : 1 11,35â : 7
## Median :141.000 Libro de bolsillo : 1 14,25â : 6
## Mean :215.220 Tapa blanda :69 21,75â : 4
## 3rd Qu.:314.000 Tapa dura :27 5,99â : 4
## Max. :871.000 6,59â : 4
## NA's :1 (Other) :67
## Paginas
## 32 pÃ¡ginas : 7
## 224 pÃ¡ginas: 6
## 96 pÃ¡ginas : 5
## : 4
## 256 pÃ¡ginas: 4
## 128 pÃ¡ginas: 3
## (Other) :71
```

```
str(llibres)
```

```
## 'data.frame': 100 obs. of 7 variables:
## $ Nom : Factor w/ 100 levels "100 Animal Mandalas Para Colorear: Libro para colorear para a
## $ Autor : Factor w/ 67 levels "Ãlvarez Bilbao",...: 38 30 47 61 35 46 50 67 20 46 ...
## $ Puntuacio : Factor w/ 13 levels "", "3,9 de un mÃ¡ximo de 5 estrellas",...: 5 7 8 9 8 8 12 8 6 8
## $ Num_resenyes: num 91 95 184 437 2.94 ...
## $ Categoria : Factor w/ 5 levels "Artículos de papelería",...: 4 4 4 4 5 4 4 4 4 4 ...
## $ Preu : Factor w/ 47 levels "1,85â", "1,89â",...: 32 17 38 30 20 35 22 18 14 36 ...
## $ Paginas : Factor w/ 62 levels "", "10 pÃ¡ginas",...: 61 46 13 62 50 53 17 45 51 56 ...
```

```
llibres <- llibres[,0:7]
```

```
llibres$Puntuacio <- as.factor(substring(llibres$Puntuacio, 0, 3))
llibres$Preu <- (gsub('[^0-9,]', '\\1', llibres$Preu))
llibres$Preu <- gsub(',', '.', llibres$Preu)
llibres$Preu <- as.numeric(llibres$Preu)
llibres$Paginas <- as.numeric(gsub('[^0-9]', '\\1', llibres$Paginas))
str(llibres)
```

```
## 'data.frame': 100 obs. of 7 variables:
## $ Nom : Factor w/ 100 levels "100 Animal Mandalas Para Colorear: Libro para colorear para a
## $ Autor : Factor w/ 67 levels "Ãlvarez Bilbao",...: 38 30 47 61 35 46 50 67 20 46 ...
## $ Puntuacio : Factor w/ 13 levels "", "3,9", "4,0",...: 5 7 8 9 8 8 12 8 6 8 ...
## $ Num_resenyes: num 91 95 184 437 2.94 ...
## $ Categoria : Factor w/ 5 levels "Artículos de papelería",...: 4 4 4 4 5 4 4 4 4 4 ...
## $ Preu : num 5.99 17.95 6.99 5.65 19.85 ...
## $ Paginas : num 92 512 168 96 568 66 20 496 640 70 ...
```

```
llibres_resum <- llibres[3:7]
pairs(llibres_resum)
```

