data reshaping

De Farmer

4/30/2021

**Data frame using a vector.**

distance time  
 1 500 5.5  
 2 353 6.0  
 3 167 4.0  
 4 120 2.0  
 5 553 6.3  
 6 340 3.0  
 7 223 4.5  
 8 880 9.0

Adding column to data frame

Speed=distance/time

distance time speed  
 1 500 5.5 90.90909  
 2 353 6.0 58.83333  
 3 167 4.0 41.75000  
 4 120 2.0 60.00000  
 5 553 6.3 87.77778  
 6 340 3.0 113.33333  
 7 223 4.5 49.55556  
 8 880 9.0 97.77778

Add let speed >=70 be fast

distance time speed fast\_slow  
 1 500 5.5 90.90909 Fast  
 2 353 6.0 58.83333 Slow  
 3 167 4.0 41.75000 Slow  
 4 120 2.0 60.00000 Slow  
 5 553 6.3 87.77778 Fast  
 6 340 3.0 113.33333 Fast  
 7 223 4.5 49.55556 Slow  
 8 880 9.0 97.77778 Fast

Class of fast\_slow:

[1] "character"

more columns

distance time speed fast\_slow gender  
 1 500 5.5 90.90909 Fast male  
 2 353 6.0 58.83333 Slow female  
 3 167 4.0 41.75000 Slow male  
 4 120 2.0 60.00000 Slow male  
 5 553 6.3 87.77778 Fast male  
 6 340 3.0 113.33333 Fast female  
 7 223 4.5 49.55556 Slow female  
 8 880 9.0 97.77778 Fast male

Working on data

i.e data transformation

distance time speed fast\_slow gender  
 1 500 5.5 90.90909 Fast 0  
 2 353 6.0 58.83333 Slow 1  
 3 167 4.0 41.75000 Slow 0  
 4 120 2.0 60.00000 Slow 0  
 5 553 6.3 87.77778 Fast 0  
 6 340 3.0 113.33333 Fast 1  
 7 223 4.5 49.55556 Slow 1  
 8 880 9.0 97.77778 Fast 0

Data transformation

distance time speed fast\_slow gender  
 1 500 5.5 90.90909 Fast 0  
 2 353 6.0 58.83333 Slow 1  
 3 167 4.0 41.75000 Slow 0  
 4 120 2.0 60.00000 Slow 0  
 5 553 6.3 87.77778 Fast 0  
 6 340 3.0 113.33333 Fast 1  
 7 223 4.5 49.55556 Slow 1  
 8 880 9.0 97.77778 Fast 0

**Analysing Factor Data**

Gender  
 0 1   
 5 3

regression

Call:  
 lm(formula = time ~ distance + gender, data = df)  
   
 Coefficients:  
 (Intercept) distance gender1   
 1.801699 0.008014 0.251301

Call:  
lm(formula = time ~ distance + gender, data = df)  
Residuals:  
 1 2 3 4 5 6 7 8   
-0.30879 1.11799 0.85993 -0.76340 0.06645 -1.77783 0.65984 0.14581   
   
Coefficients:  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) 1.801699 0.935981 1.925 0.11222   
distance 0.008014 0.001780 4.503 0.00638 \*\*  
gender1 0.251301 0.855433 0.294 0.78074   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 1.122 on 5 degrees of freedom  
Multiple R-squared: 0.8104, Adjusted R-squared: 0.7346

F-statistic: 10.69 on 2 and 5 DF, p-value: 0.01564

Reducing decimal places

1:**Round**

distance time speed fast\_slow gender  
 1 500 5.5 90.91 Fast 0  
 2 353 6.0 58.83 Slow 1  
 3 167 4.0 41.75 Slow 0  
 4 120 2.0 60.00 Slow 0  
 5 553 6.3 87.78 Fast 0  
 6 340 3.0 113.33 Fast 1  
 7 223 4.5 49.56 Slow 1  
 8 880 9.0 97.78 Fast 0