1. The following table gives the daily wages in rupees in a commercial organisation:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Daily wages (Rs.) | 30-32 | 32-34 | 34-36 | 36-38 | 38-40 | 40-42 | 42-44 | 44-46 | 46-48 | 48-50 |
| No. of persons | 3 | 8 | 24 | 31 | 50 | 61 | 38 | 21 | 12 | 2 |

Find the mean, median, mode, standard deviation and quartile deviation.

> low = seq(30, 48, 2)

> high = seq(32, 50, 2)

> x = (high + low) / 2

> f = c(3, 8, 24, 31, 50, 61, 38, 21, 12, 2)

> data = data.frame(x, f)

> mean = mean(rep(x, f))

>

> cf = cumsum(f)

> n = sum(f)

> mc = min(which(cf >= n/2))

> h = 2

> fr = f[mc]

> c = cf[mc - 1]

> l = x[mc] - h/2

> median = l + ((n/2 - c) / fr) \* h

>

> m = which(f == max(f))

> fm = f[m]

> f1 = f[m - 1]

> f2 = f[m + 1]

> l = x[m] - h/2

> mode = l + (fm - f1)/(2\*fm - f1 - f2) \* h

>

> sd = sd(rep(x, f))

>

> q = c()

> cr = c()

> h = c()

> l = c()

> qdata = c()

> for(i in c(1, 2, 3)) {

+ q = c(q, min(which(cf >= i\*n/4)))

+ cr = c(cr, cf[q[i] - 1])

+ h = c(h, high[q[i]] - low[q[i]])

+ l = c(l, x[q[i]] - h[i]/2)

+ qdata = c(qdata, l[i] + (h[i] / f[q[i]]) \* ((i\*n/4) - cr[i]))

+ }

>

> qd = (qdata[3] - qdata[1]) / 2

>

> data

x f

1 31 3

2 33 8

3 35 24

4 37 31

5 39 50

6 41 61

7 43 38

8 45 21

9 47 12

10 49 2

> mean

[1] 40.144

> median

[1] 40.29508

> mode

[1] 40.64706

> sd

[1] 3.605449

> qd

[1] 2.389219

1. The following table gives the weight(x) (in 1000 lbs.) and highway fuel efficiency (y) (in miles/gallon) for a sample of 13 cars. Calculate the correlation coefficient between weight and fuel efficiency of the vehicles.





> x = c(3.545, 2.6, 3.245, 3.93, 3.995, 3.115, 3.235, 3.225, 2.44, 3.24, 2.29, 2.5, 4.02)

> y = c(30, 32, 30, 24, 26, 30, 33, 27, 37, 32, 37, 34, 26)

> data = data.frame(x, y)

> varx = var(x)

> vary = var(y)

> varxy = var(x, y)

> cor = varxy / sqrt(varx \* vary)

> data

x y

1 3.545 30

2 2.600 32

3 3.245 30

4 3.930 24

5 3.995 26

6 3.115 30

7 3.235 33

8 3.225 27

9 2.440 37

10 3.240 32

11 2.290 37

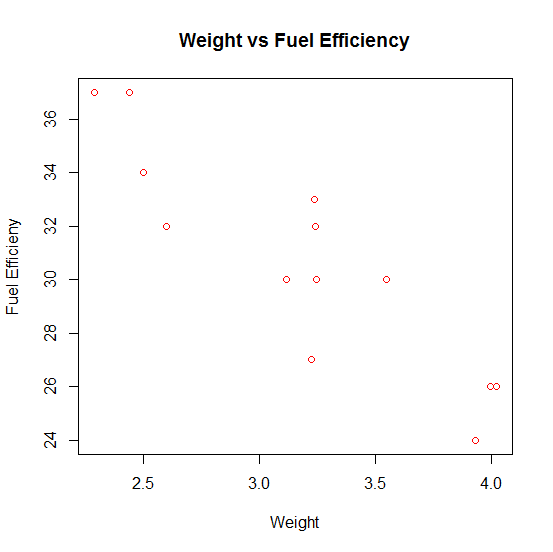
12 2.500 34

13 4.020 26

> cor

[1] -0.8977642

plot(x, y, main="Weight vs Fuel Efficiency", xlab="Weight", ylab="Fuel Efficieny", col="red")



1. 



> x = c(0.414, 0.383, 0.399, 0.402, 0.442, 0.422, 0.466, 0.5, 0.514, 0.530, 0.569, 0.558, 0.577, 0.572, 0.548, 0.581, 0.557, 0.550, 0.531, 0.55, 0.556, 0.523, 0.602, 0.569, 0.544, 0.557, 0.530, 0.547, 0.585)

> y = c(26186, 29266, 26215, 30162, 38867, 37831, 44576, 46097, 59698, 67705, 66088, 78486, 89869, 77369, 67095, 85156, 69571, 84160, 73466, 78610, 67657, 74017, 87291, 86836, 82540, 81699, 82096, 75657, 80490)

> data = data.frame(x, y)

> data

x y

1 0.414 26186

2 0.383 29266

3 0.399 26215

4 0.402 30162

5 0.442 38867

6 0.422 37831

7 0.466 44576

8 0.500 46097

9 0.514 59698

10 0.530 67705

11 0.569 66088

12 0.558 78486

13 0.577 89869

14 0.572 77369

15 0.548 67095

16 0.581 85156

17 0.557 69571

18 0.550 84160

19 0.531 73466

20 0.550 78610

21 0.556 67657

22 0.523 74017

23 0.602 87291

24 0.569 86836

25 0.544 82540

26 0.557 81699

27 0.530 82096

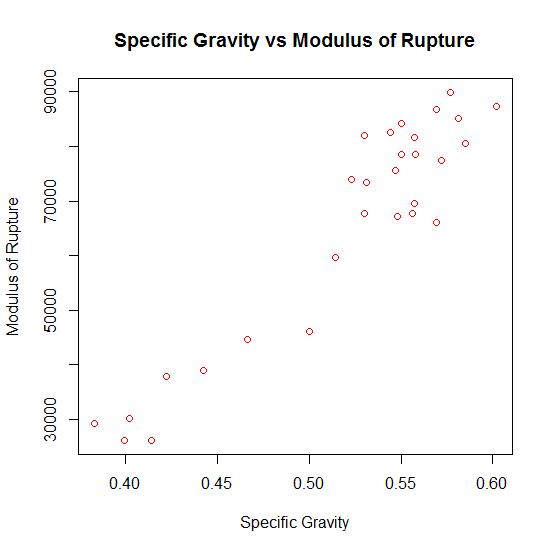
28 0.547 75657

29 0.585 80490

> cor(x, y)

[1] 0.9432149

plot(x, y, main="Specific Gravity vs Modulus of Rupture", xlab="Specific Gravity", ylab="Modulus of Rupture", col="red")



1. Calculate the Spearman’s rank correlation coefficient between advertisement cost and sales from the following data:

|  |  |
| --- | --- |
| Advertisement Cost(Rs.in 1000) | 39 65 62 90 82 75 25 98 36 78 |
| Sales (Rs. in lakhs ) | 47 53 58 86 62 68 60 91 51 84 |

> x = c(39, 65, 62, 90, 82, 75, 25, 98, 36, 78)

> y = c(47, 53, 58, 86, 62, 68, 60, 91, 51, 84)

> cor.test(x, y, method="spearman")

Spearman's rank correlation rho

data: x and y

S = 30, p-value = 0.006811

alternative hypothesis: true rho is not equal to 0

sample estimates:

rho

0.8181818