Experimen t 6

1. Two judges gave the following rank to a series of eight one act plays in drama competition. Examine the relationship between their judgments.

Judge A	8	7	6	3	2	1	5	4
Judge B	7	5	4	1	3	2	6	8

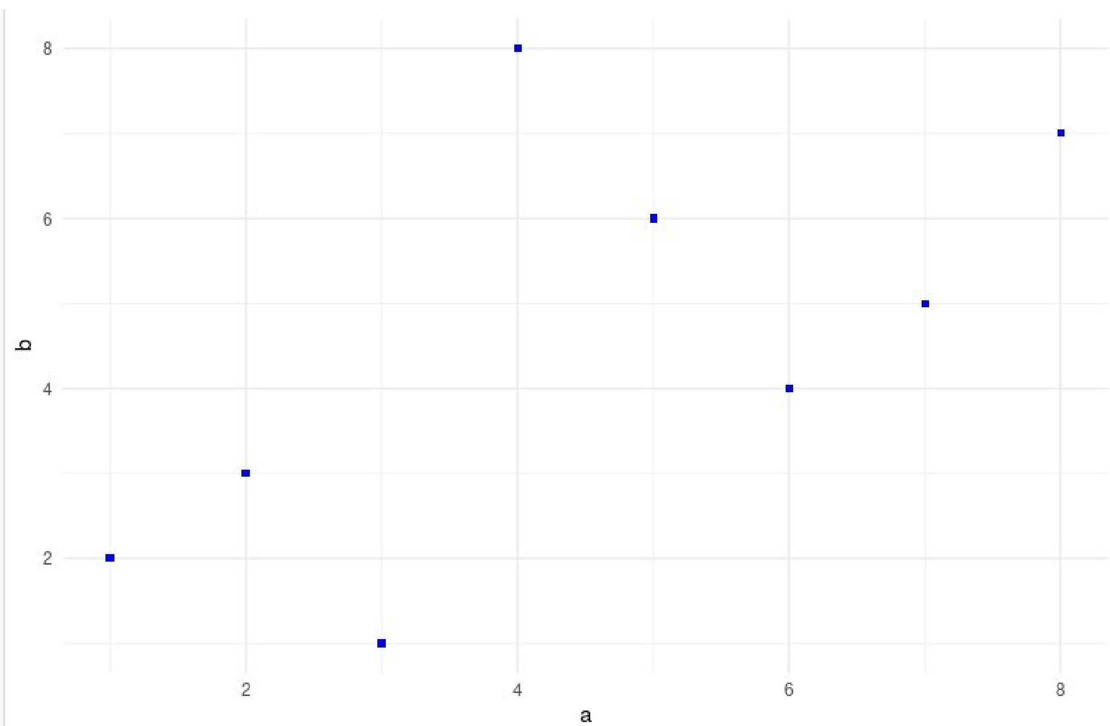
Write a R program for above problem.

Code:

```
#Q1
a=c(8,7,6,3,2,1,5,
4)
b=c(7,5,4,1,3,2,6,8)
cor(a,b,method="spearman") ggplot()+aes(x=a,y=b)
+geom_point(shape=15,colour="blue")+theme_minimal()
```

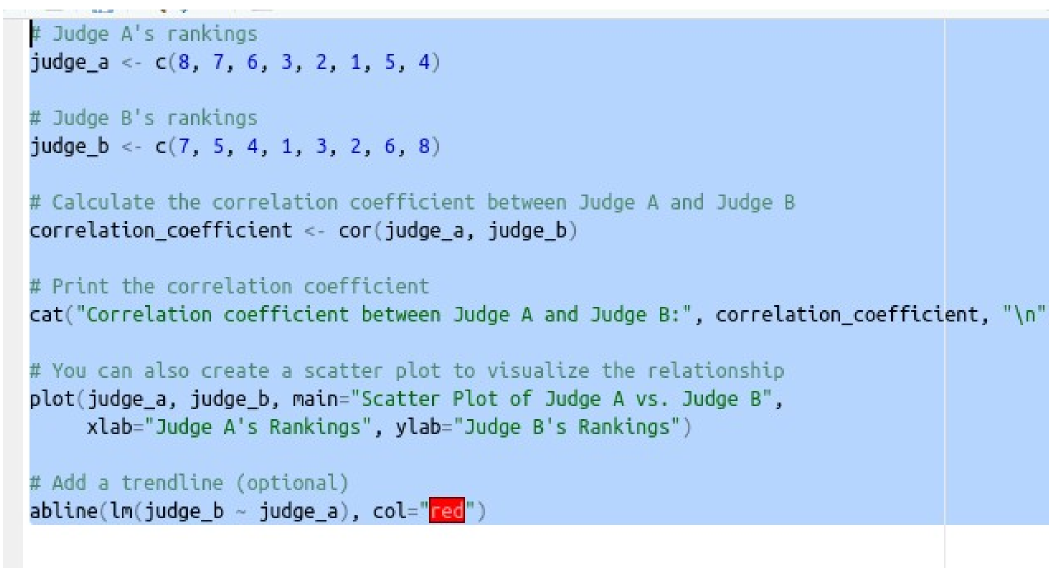
Output:

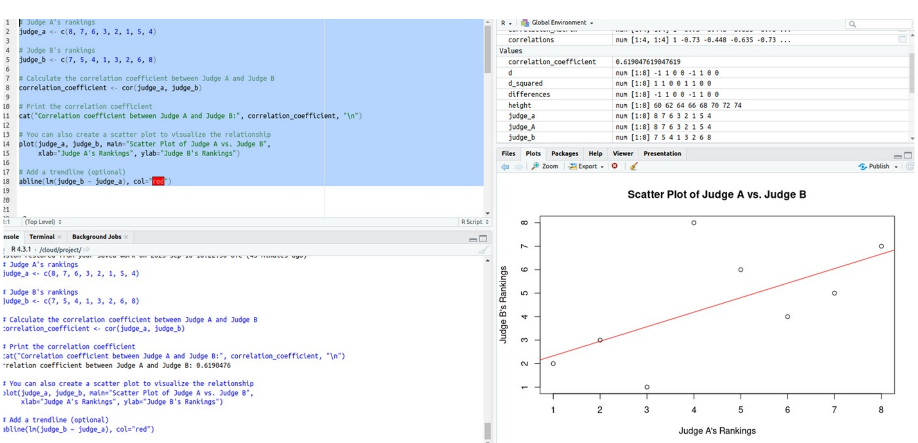
```
> #Q1
> a=c(8,7,6,3,2,1,5,4)
> b=c(7,5,4,1,3,2,6,8)
> cor(a,b,method="spearman")
[1] 0.6190476
> ggplot()+aes(x=a,y=b)+geom_point(shape=15,colour="blue")+theme_minimal()
> p
```



Approach 2

```
# Judge A's rankings
judge_a <- c(8, 7, 6, 3, 2, 1, 5, 4)
# Judge B's rankings
judge_b <- c(7, 5, 4, 1, 3, 2, 6, 8)

# Calculate the correlation coefficient between Judge A and Judge B
correlation_coefficient <- cor(judge_a, judge_b)


# Print the correlation coefficient
cat("Correlation coefficient between Judge A and Judge B:", correlation_coefficient, "\
n")

# You can also create a scatter plot to visualize the relationship
plot(judge_a, judge_b, main="Scatter Plot of Judge A vs. Judge B",
    xlab="Judge A's Rankings", ylab="Judge B's Rankings")

# Add a trendline (optional)
abline(lm(judge_b ~ judge_a), col="red")
```





Solved manually:

PS Exp6 Saakshi Jain
B033 Btech CE(B)

(Q1) Judge A 8 7 6 3 2 1 5 4

Judge B 7 5 4 1 3 2 6 8

$$R = 1 - \frac{6 \sum d_i^2}{n^3 - n}$$

$$= 1 - \frac{6\left[(1)^2+(2)^2+(2)^2+(2)^2+(1) + (-1)^2+(-1)^2+(-4)^2\right)}{8^3 - 8}$$

$$= 1 - \frac{6(1+12+3+16)}{504}$$

$$= 1 - \frac{6(32)}{8(64-1)}$$

$$= 1 - \frac{24}{63}$$

$$= \boxed{0.6190}$$

2. Calculate the rank correlation coefficient from the following data.

Height	60	62	64	66	68	70	72	74
Weight	92	83	101	110	128	119	137	146

Write a R program for above problem.

Code:

```
#Q2
height=c(60,62,64,66,68,70,7
2,74)
weight=c(92,83,101,110,128,119,137,146)
cor(height,weight,  method="spearman")
```

Output:

```
> #Q2
> height=c(60,62,64,66,68,70,72,74)
> weight=c(92,83,101,110,128,119,137,146)
> cor(height,weight, method="spearman")
[1] 0.952381
```

Approach 2

q2

```
# Height and Weight data
height <- c(60, 62, 64, 66, 68, 70, 72, 74)
weight <- c(92, 83, 101, 110, 128, 119, 137, 146)

# Calculate the rank of height and weight
rank_height <- rank(height)
rank_weight <- rank(weight)

# Calculate the difference in ranks
d <- rank_height - rank_weight

# Calculate the squared difference
d_squared <- d^2

# Calculate the sum of squared differences
sum_d_squared <- sum(d_squared)

# Calculate the number of observations
n <- length(height)

# Calculate the rank correlation coefficient (Spearman's rho)
rho <- 1 - (6 * sum_d_squared) / (n * (n^2 - 1))

# Print the rank correlation coefficient
cat("Rank Correlation Coefficient (Spearman's rho):", rho, "\n")
```

```
>
> # Calculate the rank of height and weight
> rank_height <- rank(height)
> rank_weight <- rank(weight)
>
> # Calculate the difference in ranks
> d <- rank_height - rank_weight
>
> # Calculate the squared difference
> d_squared <- d^2
>
> # Calculate the sum of squared differences
> sum_d_squared <- sum(d_squared)
>
> # Calculate the number of observations
> n <- length(height)
>
> # Calculate the rank correlation coefficient (Spearman's rho)
> rho <- 1 - (6 * sum_d_squared) / (n * (n^2 - 1))
>
> # Print the rank correlation coefficient
> cat("Rank Correlation Coefficient (Spearman's rho):", rho, "\n")
Rank Correlation Coefficient (Spearman's rho): 0.952381
>
```

Solved manually:

Create a new dataframe, auto_num, that contains only columns with numeric values from the auto dataset. You can do this using the Filter function. Calculate correlation for all pairs of numeric variables .

(Q2) Height 60 62 64 66 68 70 72 74

Weight 92 83 101 110 128 119 137 146

7 8 6 5 3 4 2 1

Rank of X	Rank of Y
8	7
7	8
6	6
5	5
4	3
3	4
2	2
1	1

$$R = 1 - \frac{6 \sum d_i^2}{n^3 - n}$$

$$= 1 - \frac{6(1^2+1^2+0^2+0^2+1^2+1^2+0^2+0^2)}{8(63)}$$

$$= 1 - \frac{6 * 4}{8 * 63}$$

$$= 1 - \frac{3}{63}$$

$$= 1 - 0.004761$$

$$= \boxed{0.952381}$$

Teacher's Signature:............

Code:

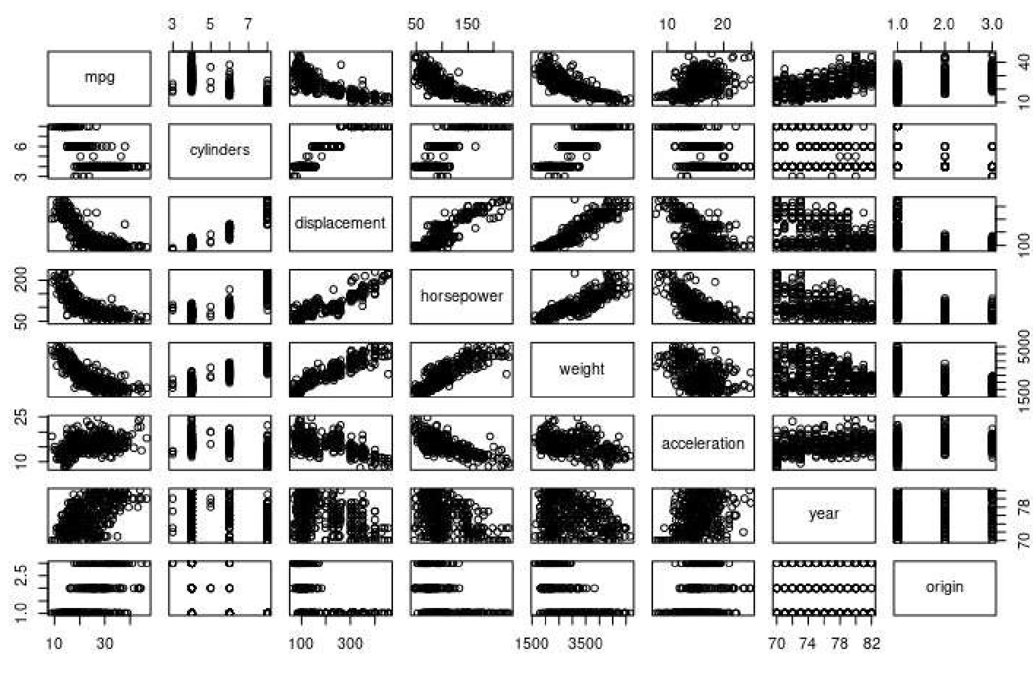
```
#Q3
library(ISL
R)
data("Aut
o")
head(Aut
o)
auto_num=data.frame(Auto[c(1,2,3,4,5,6,7,8)]
) auto_num
pairs(auto_num)
```

Output:

```
q3

# Assuming 'auto' is your original dataframe
# Create a new dataframe 'auto_num' with only numeric columns
auto_num <- Filter(is.numeric, auto)

# Calculate correlations for all pairs of numeric variables
correlations <- cor(auto_num)

# Print the correlation matrix
print(correlations)
```

```
> # Assuming 'auto' is your original dataframe
> # Create a new dataframe 'auto_num' with only numeric columns
> auto_num <- Filter(is.numeric, auto)
>
> # Calculate correlations for all pairs of numeric variables
> correlations <- cor(auto_num)
>
> # Print the correlation matrix
> print(correlations)
                  MPG  Cylinders Horsepower     Weight
MPG         1.0000000 -0.7298004 -0.4483129 -0.6346757
Cylinders  -0.7298004  1.0000000  0.9214427  0.9421279
Horsepower -0.4483129  0.9214427  1.0000000  0.9170876
Weight     -0.6346757  0.9421279  0.9170876  1.0000000
```

	mpg	cylinders	displacement	horsepower	weight	acceleration	year	origin
1	18	8	307.0	130	3504	12.0	70	1
2	15	8	350.0	165	3693	11.5	70	1
3	18	8	318.0	150	3436	11.0	70	1
4	16	8	304.0	150	3433	12.0	70	1
5	17	8	302.0	140	3449	10.5	70	1
6	15	8	429.0	198	4341	10.0	70	1
7	14	8	454.0	220	4354	9.0	70	1
8	14	8	440.0	215	4312	8.5	70	1
9	14	8	455.0	225	4425	10.0	70	1
10	15	8	390.0	190	3850	8.5	70	1
11	15	8	383.0	170	3563	10.0	70	1
12	14	8	340.0	160	3609	8.0	70	1
13	15	8	400.0	150	3761	9.5	70	1
14	14	8	455.0	225	3086	10.0	70	1
15	24	4	113.0	95	2372	15.0	70	3



4. Use the cor function to create a matrix of correlation coefficients for variables in the auto_num dataframe.

Code:

#Q4

cor(auto_num)

Output:

```
> #Q4
> cor(auto_num)
                    mpg  cylinders displacement horsepower     weight acceleration       year     origin
mpg           1.0000000 -0.7776175   -0.8051269 -0.7784268 -0.8322442    0.4233285  0.5805410  0.5652088
cylinders    -0.7776175  1.0000000    0.9508233  0.8429834  0.8975273   -0.5046834 -0.3456474 -0.5689316
displacement -0.8051269  0.9508233    1.0000000  0.8972570  0.9329944   -0.5438005 -0.3698552 -0.6145351
horsepower   -0.7784268  0.8429834    0.8972570  1.0000000  0.8645377   -0.6891955 -0.4163615 -0.4551715
weight       -0.8322442  0.8975273    0.9329944  0.8645377  1.0000000   -0.4168392 -0.3091199 -0.5850054
acceleration  0.4233285 -0.5046834   -0.5438005 -0.6891955 -0.4168392    1.0000000  0.2903161  0.2127458
year          0.5805410 -0.3456474   -0.3698552 -0.4163615 -0.3091199    0.2903161  1.0000000  0.1815277
origin        0.5652088 -0.5689316   -0.6145351 -0.4551715 -0.5850054    0.2127458  0.1815277  1.0000000
```

5. Find Karl Pearson's coefficient of correlation for the following

X	62	64	65	69	70	71	72	74
Y	126	125	139	145	165	152	180	208

Write a R program for the above problem.

Code:

#Q5

x=c(62,64,65,69,70,71,72

,74)

y=c(126,125,139,145,165,152,180,208)

cor(x,y, method="pearson")

Output:

```
> #Q5
> x=c(62,64,65,69,70,71,72,74)
> y=c(126,125,139,145,165,152,180,208)
> cor(x,y, method="pearson")
[1] 0.9031822
```

Solved manually:

(Q5) X= 62 64 65 64 70 71 72 74

Y= 126 125 139 145 165 152 180 208

$$r = \frac{\sum xy - \frac{(\sum x \sum y)}{n}}{\sqrt{\left[\sum x^2 - \frac{(\sum x)^2}{n}\right]\left[\sum y^2 - \frac{(\sum y)^2}{n}\right]}}$$

X	y	xy	x^2	y^2
62	126	7812	3844	15876
64	125	8000	4096	15625
65	139	9035	4225	19321
69	145	10005	4761	21025
70	165	11550	4900	27225
71	152	10792	5041	23104
72	780	12960	5184	32400
74	208	15392	5476	43264

$\Sigma x = 547, \Sigma y = 1240, \Sigma x^2 = 37529$ $\Sigma y^2 = 197840, \Sigma xy = 85546$

$$\bar{x} = \frac{1}{n}\Sigma x_i = \frac{547}{8} = 68.375$$

$$\bar{y} = \frac{1}{n}\Sigma y_i = \frac{1240}{8} = 155$$

$$r_{xy} = \frac{\frac{1}{n}\Sigma x_i y_i - \bar{x}\bar{y}}{\sqrt{\left(\frac{1}{n}\Sigma x_i^2 - \bar{x}^2\right)\left(\frac{1}{n}\Sigma y_i^2 - \bar{y}^2\right)}}$$

$$= \frac{\frac{1}{8} * 85546 - (68.375 * 155)}{\sqrt{\left(\frac{1}{8} * 37527 - (68.375)^2\right)\left(\frac{1}{8} * 197840 - (135)^2\right)}}$$

$$\boxed{r = 0.903}$$

6. At Hogwarts School of Witchcraft and Wizardry, students often have a lot of homework. The table below indicates the number of hours students studied, and how they performed on an exam in two of their classes.

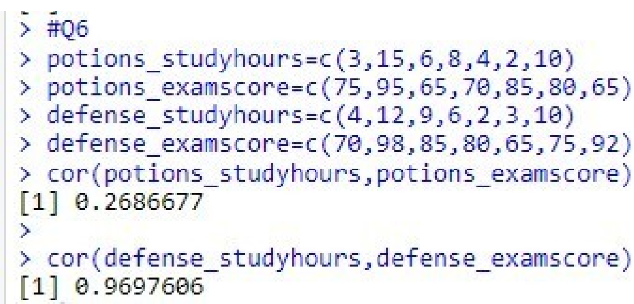
Student	Potions		Defense against the dark arts	
	study hours	exam score	study hours	exam score
1	3	75	4	70
2	15	95	12	98
3	6	65	9	85
4	8	70	6	80
5	4	85	2	65
6	2	80	3	75
7	10	65	10	92

Find the correlations between hours spent studying and how students performed in their potions and defense against the dark arts classes.

Code:

```
#Q6
potions_studyhours=c(3,15,6,8,4,2,
10)
potions_examscore=c(75,95,65,70,
85,80,65)
defense_studyhours=c(4,12,9,6,2,3
,10)
defense_examscore=c(70,98,85,80,6
5,75,92)
cor(potions_studyhours,potions_exa
mscore)

cor(defense_studyhours,defense_examscore)
```

Output:



```
> #Q6
> potions_studyhours=c(3,15,6,8,4,2,10)
> potions_examscore=c(75,95,65,70,85,80,65)
> defense_studyhours=c(4,12,9,6,2,3,10)
> defense_examscore=c(70,98,85,80,65,75,92)
> cor(potions_studyhours,potions_examscore)
[1] 0.2686677
>
> cor(defense_studyhours,defense_examscore)
[1] 0.9697606
```

Solved manually:

Saakshi Jain B02 09/09/20

(Q6) Potions

Study hours(x)	Exam score(Y)	X·X	YY	XY
3	75	9	5625	225
15	95	225	9025	1425
6	65	36	4225	390
8	70	64	4900	560
4	85	16	7225	340
2	80	4	6400	160
10	65	100	4225	650

$\Sigma x = 48$, $\Sigma y = 535$, $\Sigma x^2 = 454$, $\Sigma y^2 = 41625$
$\Sigma xy = 3750$

$\bar{x} = \frac{1}{n}\Sigma x_i = \frac{1}{7} \times 48 = 6.857$

$\bar{y} = \frac{1}{n}\Sigma y_i = \frac{1}{7} \times 535 = 76.429$

$$r_{xy} = \frac{\frac{1}{n}\Sigma x_i y_i - \bar{x}\bar{y}}{\sqrt{\left(\frac{1}{n}\Sigma x_i^2 - \bar{x}_i^2\right)\left(\frac{1}{n}\Sigma y_i^2 - \bar{y}^2\right)}}$$

$$= \frac{\frac{1}{7} \times 3750 - (6.857)(76.429)}{\sqrt{\left(\frac{1}{7} \times 454 - (6.857)^2\right)\left(\frac{1}{7} \times 41625 - (76.429)^2\right)}}$$

$\boxed{r = 0.269}$

Defence

Study hours(x)	Exam Score(Y)	x^2	y^2	XY
4	70	16	4900	280
12	98	144	9604	1176
9	85	81	7225	765
6	80	36	6400	480
2	65	4	4225	130
3	75	9	5625	225
10	92	100	8464	920

$\Sigma x = 46$, $\Sigma y = 565$, $\Sigma x^2 = 390$, $\Sigma y^2 = 46443$, $\Sigma xy = 3976$

$\bar{x} = \frac{1}{n}\Sigma x_i^2 = 46/7 = 6.571$

$\bar{y} = \frac{1}{n}\Sigma y_i^2 = \frac{565}{7} = \cancel{800}\ 80.714$

$$r_{xy} = \frac{\frac{1}{n}\Sigma x_i y_i - \bar{x}\bar{y}}{\sqrt{\left(\frac{1}{n}\Sigma x_i^2 - \bar{x}^2\right)\left(\frac{1}{n}\Sigma y_i^2 - \bar{y}^2\right)}}$$

$$= \frac{\left(\frac{1}{7} \times 3976\right) - \left(\frac{46}{7} \times \frac{565}{7}\right)}{\sqrt{\left(\frac{1}{7} \times 390 - \left(\frac{46}{7}\right)^2\right)\left(\frac{1}{7} \times 46443 - \left(\frac{565}{7}\right)^2\right)}}$$

$\boxed{r = 0.96976}$