BRAC UNIVERSITY

STA201

Elements of Statistics and Probability

Assignment 3

Student Information:
Name: Md. Danial Islam ID: 20101534
Section: 01

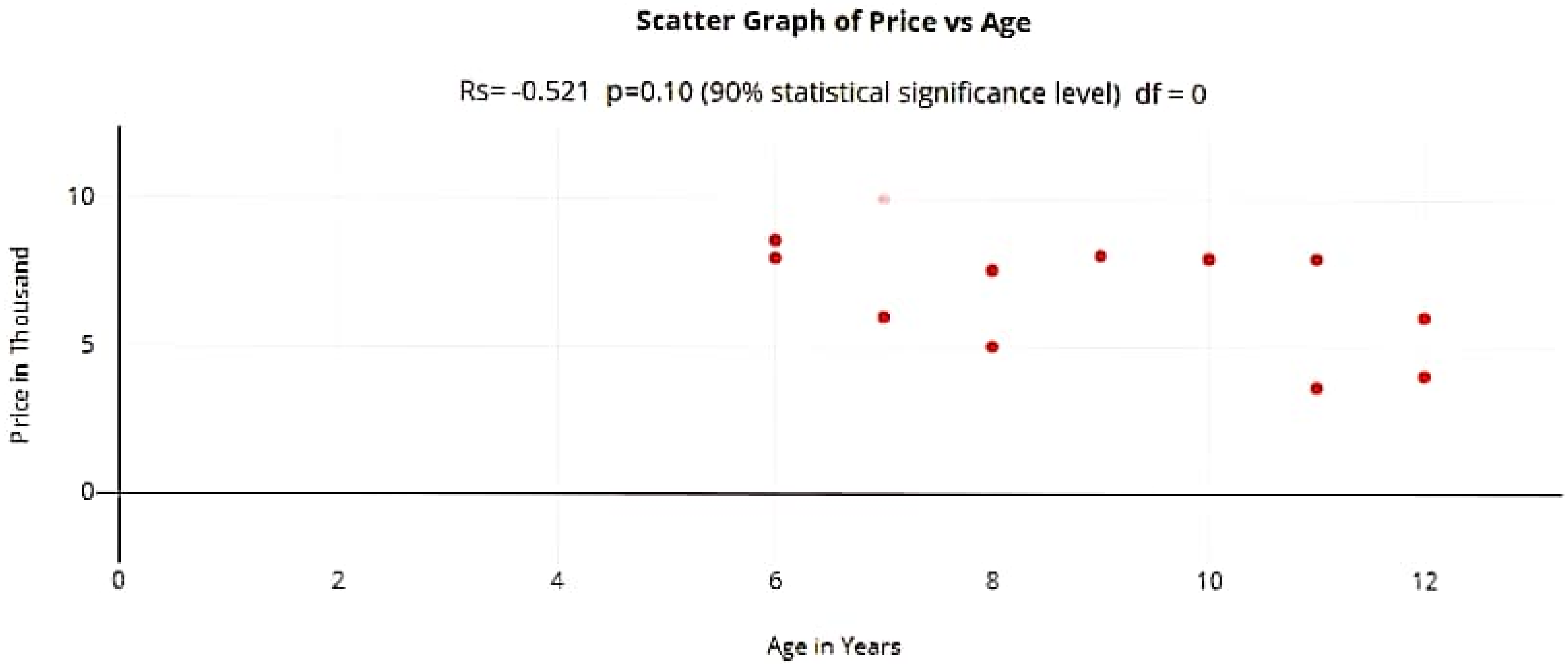
Date: 19 November 2021

Date :......../........

STA201 Assignment 3

Problem 1

(a)



Scatter Graph of Price vs Age
Rs= -0.521 p=0.10 (90% statistical significance level) df = 0
Price in Thousand
10
5
0
0
2
4
6
8
10
12
Age in Years

from the diagram we can see that there is a negative correlation between age and price.

~~from the~~

Date : ..../........../..........

b)

Age (years) (x)	Price (\$) (Y)	x^2	y^2	xy
9	8.1	81	65.61	72.9
7	6	49	36	42
11	3.6	121	12.96	39.6
12	4	144	16	48
8	5	64	25	40
7	10	49	100	70
8	7.6	64	57.76	60.8
11	8	121	64	88
10	8	100	64	80
12	6	144	36	72
6	8.6	36	73.96	51.6
6	8	36	64	48
Total 107	82.9	1009	615.29	712.9

$$\bar{x} = \frac{107}{12} = 8.9167$$

$$\bar{y} = \frac{82.9}{12} = 6.9083$$

$$r = \frac{\sum_{i=1}^{12} (x_i y_i) - n\,\bar{x}\,\bar{y}}{\sqrt{\left(\sum_{i=1}^{12} x_i^2 - n\bar{x}^2\right)\left(\sum_{i=1}^{12} y_i^2 - n\bar{y}^2\right)}} = \frac{712.9 - (12 \times 8.917 \times 6.908)}{\sqrt{\left(1009 - (12 \times 8.917^2)\right) \times \left(615.29 - 12 \times 6.9083^2\right)}}$$

Date :........../........../..........

$\therefore r = -0.5436$.

as $r = -0.5436$ there is a moderate negative correlation between age and price

$r^2 = 0.2956 = 0.2956 \times \frac{100}{100}$

$= 29.56\ \%$

only 29.56 % of the total variation in the price of car can be explained by the age of the car.

Date : ..../......../........

Problem 2

(a)

Judge 1(x)	Judge 2(Y)	Rank x	Rank Y	RX-RY = d	d^2
650	900	8	4	4	16
760	720	2	9	-7	49
740	690	3	11.5	-8.5	72.25
700	850	5	6	-1	1
590	920	11	2.5	8.5	72.25
620	800	9	7	2	4
700	890	5	5	0	0
690	920	7	2.5	4.5	20.25
900	1000	1	1	0	0
500	690	12	11.5	0.5	0.25
610	700	10	10	0	0
700	760	5	8	-3	9
				Total	244

$$r_s = 1 - \frac{6\sum_{i=1}^{12} d_i^2}{n(n^2-1)} = 1 - \frac{6\times 244}{12\times 143} = 0.1469$$

b) as $r_s = 0.1469$ so the scores of judges have a very week (+ve) correlation

— o —

Problem 3 :

(a)

Numbers of Rooms	Energy Consumption (thousand kWh)	x^2	y^2	xy
12	9	144	81	108
9	7	81	49	63
14	10	196	100	140
6	5	36	25	30
10	8	100	64	80
8	6	64	36	48
10	8	100	64	80
10	10	100	100	100
5	4	25	16	20
7	7	49	49	49
Total 91	74	895	584	718

$$b_1 = \frac{n\left(\sum_{i=1}^{10} x_i y_i\right) - \left(\sum_{i=1}^{10} x_i\right)\left(\sum_{i=1}^{10} y_i\right)}{n\left(\sum_{i=1}^{10} x_i^2\right) - \left(\sum_{i=1}^{10} x_i\right)^2}$$

$$= \frac{(10 * 718) - (91 * 74)}{(10 \times 895) - 91^2}$$

$$= 0.6667$$

$$b_0 = \bar{y} - b_1 \bar{x} = 7.4 - (0.6667 \times 9.1)$$

$$= 1.3333$$

$$\hat{y} = b_0 + b_1 x$$

$$= 1.3333 + 0.6667x \quad \text{(Ans)}$$

(b) b_0 indecates the initial total energy consumption when the numbers of rooms are 0, so, when number of rooms are 0 energy consumption will be 1.333.

Date :........../........../..........

b_1 means the slope or, the overall energy consumption will increase by b_1 or 0.6667 (thousand kWh) for increasing of 1 room.

— o —

c). if $x = 6$

$\hat{y} = 1.3333 + (0.6667 \times 6)$

$= 5.335$

if room $= 6$ then energy consumption will be 5.335 (thousand kWh)

— o —

Protivas

Date :........../........../..........

d) we know,

$$r^2 = 1 - \frac{SSE}{SST}$$

$$= 1 - \cancel{\frac{\sum y_i^2 - \frac{(\sum y_i)^2}{n}}{\quad}}$$

$$= 1 - \frac{\sum y_i^2 - b_0 \sum y_i - b_1 \sum x_i y_i}{\sum y_i^2 - \frac{(\sum y_i)^2}{n}}$$

$$= 1 - \frac{6.6667}{36.4}$$

$$= 0.8168$$

$$r^2 = 0.8168 = 0.8168 \times \frac{100}{100}$$

$$= 81.68\%$$

So, total 81.68% of the variation in energy consumption can be explained by the number of rooms.

—— o ——

Problem 4

a) from the given data

$$b_0 = 356.12083$$

$$b_1 = -0.09874$$

$$b_2 = 122.86721$$

$$\hat{y} = b_0 + b_1 x_1 + b_2 x_2$$

$$= 356.12083 - 0.09874 x_1 + 122.86721 x_2$$

b) b_0 or 356.12083 indecates that the price of a backpack will be 356.12083 when both the capasity and comfort rating are 0.

$b_1 = -0.09874$ Indecates that the price of a backpack will be -0.09874 less than the previous when the

Date :........../........../..........

capasity will increase 1 cf but comfort rating is fixed

b_2 or 122.86721 indecates that, the price of the backpack will be 122.8672 more than previous unit cost when comfort rating increase by one unit but capasity is fixed.

— o —

c) from (a)

$$\hat{y} = 356.12083 - 0.09874\, x_1 + 122.86721\, y_1$$

$$= 356.12083 - (0.09874 \cdot 4500) + (122.86721 \times 4)$$

$$= 403.25967 \text{ (Ans)}$$

— o —

Date : ..../........../..........

d) from the data

Adjusted R-squared $= 0.7838$

$= 0.7838 \times \frac{100}{100}$

$= 78.38\%$

$\therefore$ only 78.38% of the variation in the price of a backpack can be explained by taking into account of capasity and comfort rating.

Problem 5

a) from the data

$\beta_0 = -471.441$

$\beta_1 = 6.394$

$\beta_2 = 1.347$

Protivas

Date :........../........../..........

$$\hat{y} = \frac{e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2)}}$$

$$= \frac{e^{(-471.441 + 6.394 x_1 + 1.347 x_2)}}{1 + e^{(-471.441 + 6.394 x_1 + 1.347 x_2)}}$$

b) from the data,

$$b_1 = 6.394, \; b_2 = 1.347$$

so, odds ratio $= e^{6.394}$

$= 598.2448$

and odds Ration $= e^{1.347}$

$= 3.8459$

for b_1, the odds of having another heart attack within one year is increased 598.2448 for every unit increase in the age.

Trotivas

Date :........../........../..........

for, b_2,

the odds of having a second heart attack within 1 year in increased 3,8459 for every unit increase in the score of anxiety.

— c —