Day 2 Fellowship program in AI/ML - Problems for Regression

(1) A sample of 6 person was selected the value of their age (x variable) and their weight is demonstrated in the following table. Find the regression equation and what is the predicted weight when age is 8.5 years

x	y	x^2	y^2
7	12	84	49	144
6	8	48	36	64
8	12	96	64	144
5	10	50	25	100
6	11	66	36	121
9	13	117	81	169
41	66	461	291	742

$$\theta_1 = \frac{n\sum xy - \sum x \sum y}{n\sum x^2 - (\sum x)^2}$$

$$\theta_1 = \frac{6 \times 461 - 41 \times 66}{6 \times 291 - 41 \times 41} = .92$$

$$\theta_2 = \frac{66 \times 291 - 41 \times 461}{6 \times 291 - 41 \times 41} = 4.692$$

$$y = .92x(x) + 4.692$$

$$\Rightarrow 4.692 + .92 \times 8.5$$

$$\Rightarrow 12.510$$

(2) The following are the age (in years) and systolic blood pressure of 20 apparently healthy adults. Find the regression equation. What is the blood pressure for a man aging 25 years

Age

70
63
61
26
63
71
53
60
53

Age(x)	B.P(y)	xy	x^2	y^2
20	120	2400	400	14200
43	128	5504	1849	16384
63	141	8742	3969	19881
26	126	3276	676	15876
53	134	7102	2809	17956
31	128	3968	961	16384
58	136	7888	3364	18496
46	132	6072	2116	17424
58	140	8120	3364	19600
70	144	10080	4900	20736
46	128	5888	2116	16384
53	136	7208	2809	18496
60	146	8760	3600	21316
20	124	2480	400	15376
63	143	9009	3969	20449
43	130	5590	1849	16900
26	124	3224	676	15376
19	121	2299	361	14641
31	126	3906	961	15876
23	123	2829	529	15129
852	2630	114486	41678	347080

$$\theta_2 = \frac{\Sigma y \Sigma x^2 - \Sigma x \Sigma xy}{n\Sigma(x^2) - (\Sigma x)^2} \Rightarrow \frac{109284390 - 97307595}{831060 - 849}$$

$$\Rightarrow 14.4267$$

$$\theta_1 = \frac{n\Sigma xy - \Sigma x \Sigma y}{n(\Sigma x^2) - (\Sigma x)^2} = 228690$$

$$\theta_1 = \frac{n\Sigma xy - \Sigma x \cdot \Sigma y}{n(\Sigma x^2) - (\Sigma x)^2} = \frac{2289720 - 2240760}{6 \times 41678 - 725904}$$

$$= 105895 \cdot 454781$$

$$\theta_2 = \frac{\Sigma y \Sigma x^2 - \Sigma x \cdot \Sigma xy}{n(\Sigma x^2) - (\Sigma x)^2} = \frac{10961340 - 97542077}{833560 - 725904}$$

$$= 112.1262$$

$$y = 105895 + 14496 \cdot 454(25) + 112.1262$$

$$B.P = 123.4958$$

Equation $= \cdot 454(\theta_1) + 112.1262$

(03) The time x in years that an employee spent at a company and the employee's hourly pay, y for 5 employees are listed in the table below. Calculate and interpret the correlation coefficient r and also find the equation of regression line. Include plot of data in your discussion.

x	y	x^2	y^2	xy
5	25	25	625	125
3	20	9	400	60
4	21	16	441	84
10	35	100	1225	350
15	38	225	1444	570
37	139	375	4135	1189

$$\theta_1 = \frac{n\Sigma xy - \Sigma x \cdot \Sigma y}{n\Sigma x^2 - (\Sigma x)^2}$$

$$\theta_2 = \frac{\Sigma y \Sigma x^2 - \Sigma x \cdot \Sigma xy}{n\Sigma x^2 - (\Sigma x)^2}$$

$$r = \frac{N\Sigma xy - \Sigma x \cdot \Sigma y}{\sqrt{[N\Sigma x^2 - (\Sigma x)^2][N\Sigma y^2 - (\Sigma y)^2]}}$$

$\theta_1 = 1.5849$

$\theta_2 = 16.0711$

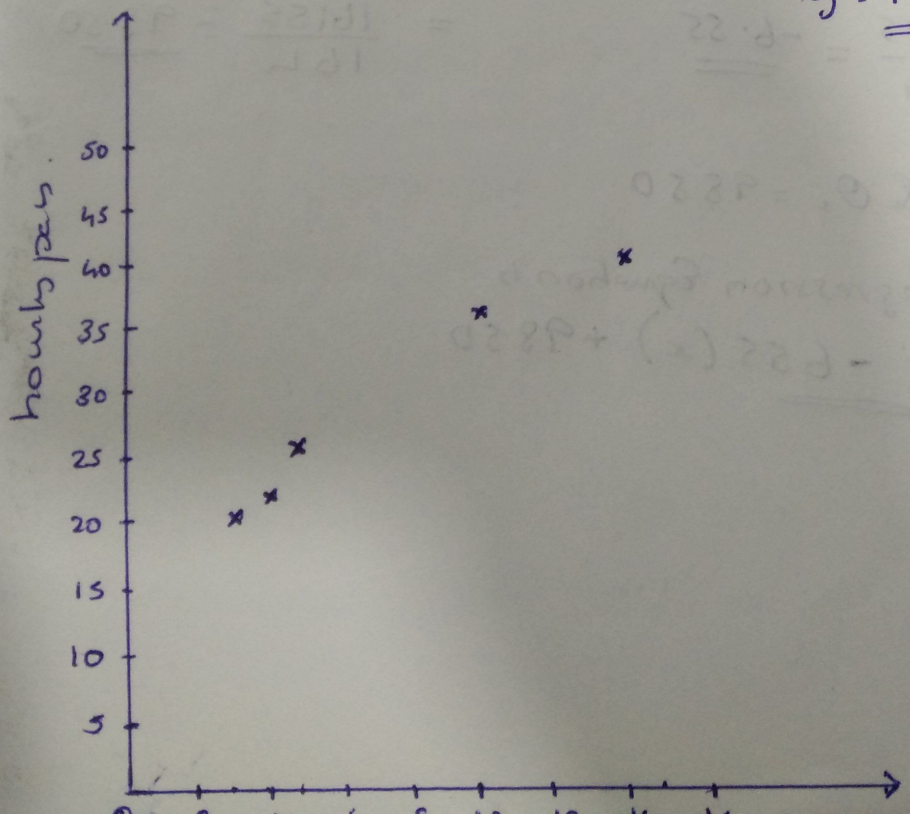
$r = .9889$ (high)

$$\theta_1 = \frac{5945 - 5143}{1875 - 1369} = \frac{802}{506} \Rightarrow 1.58$$

$$\theta_2 = \frac{52125 - 43993}{1875 - 1369} = \frac{8132}{506} = 16.0711$$

$$r = \frac{5 \times 375 - 37 \times 139}{\sqrt{(5 \times 375 - 37^2)(5 \times 4135 - (139)^2}} \Rightarrow \frac{802}{827.72} = .9689$$

$$y = 1.58(x_i) + 16.071$$



The table shows number of absences, x in a calculus course and final exam grade for 7 students. Find the correlation coefficients also find the equation of Regression line and interpret the result.

x	1	0	2	6	4	3	3
y	95	90	90	55	70	80	85

$\Sigma x = 19$ $\quad \Sigma y = 565$

$\Sigma x^2 = 75$ $\quad \Sigma y^2 = 46775$

$\Sigma xy = 1380$

$n = 7$

$$\text{Correlation coefficient} = \frac{n\Sigma xy - \Sigma x \cdot \Sigma y}{\sqrt{(n\Sigma x^2 - (\Sigma x)^2) * (n\Sigma y^2 - \Sigma y^2)}}$$

$$= \frac{7 \times 1380 - 19 \times 565}{\sqrt{(7 \times 75 - (19)^2) \times (7 \times 46775 - (565)^2)}}$$

$$\Rightarrow \frac{-1075}{115965} = -0.92 \text{ (strong negative)}$$

Equation of line

$$\theta_1 = \frac{n\Sigma xy - \Sigma x \cdot \Sigma y}{n\Sigma x^2 - (\Sigma x)^2}$$

$$\theta_1 = \frac{7 \times 1380 - 19 \times 565}{7 \times 75 - 19^2}$$

$$= \frac{9660 - 10735}{525 - 361}$$

$$= \frac{-1075}{164} = -6.55$$

$$\theta_2 = \frac{\Sigma y \Sigma x^2 - \Sigma x \cdot \Sigma xy}{n\Sigma x^2 - (\Sigma x)^2}$$

$$\theta_2 = \frac{565 \times 75 - 19 \times 1380}{7 \times 75 - 19^2}$$

$$= \frac{42375 - 26220}{525 - 361}$$

$$= \frac{16155}{164} = 98.50$$

$\theta_1 = -6.55$ & $\theta_2 = 98.50$

hence Regression Equation is

$$y = -6.55(x) + 98.50$$