Problem.

For a given data set
x 5 15 25 35 45 50
y 10 18 20 25 32 45
i. Find the value of , and
ii. Write the normal equations and estimate the equation of line of regression of Y on x.
iii. Find the value of sample correlation coefficient for the above data.

Remark: The part of question is missed. I suppose that the correct statement is
" For a given data set

x	5	15	25	35	45	50
y	10	18	20	25	32	45

i. Find the value of $\sum x^2$, $\sum y^2$ and $\sum xy$
ii. Write the normal equations and estimate the equation of line of regression of Y on x.
iii. Find the value of sample correlation coefficient for the above data."
(see http://www.bits-pilani.ac.in/uploads/ReadPDFDOC/AAOC_C111_515_C_2009_2.doc)

Solution:

i.

							$\sum$
x	5	15	25	35	45	50	175
y	10	18	20	25	32	45	150
x^2	25	225	625	1225	2025	2500	6625
y^2	100	324	400	625	1024	2025	4498
xy	50	270	500	875	1440	2250	5385

Then $\sum x^2 = 6625, \sum y^2 = 4498$ and $\sum xy = 5385$.

ii. The normal equation for a and b are $a\sum x + nb = \sum y$ and $a\sum x^2 + b\sum x = \sum xy$. Hence
$175a + 6b = 150$ and $6625a + 175b = 5385$. Then $a = \frac{1212}{1825}$ and $b = \frac{411}{73}$. Hence the equation of
the regression line is $y = ax + b = \frac{1212}{1825}x + \frac{411}{73}$.
Answer: $y = \frac{1212}{1825}x + \frac{411}{73}$.

iii. The sample correlation coefficient
$$r = \frac{\sum_{i=i}^{6}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=i}^{6}(x_i - \bar{x})^2}\sqrt{\sum_{i=i}^{6}(y_i - \bar{y})^2}} \approx 0.947.$$
Answer: $r \approx 0.947$.