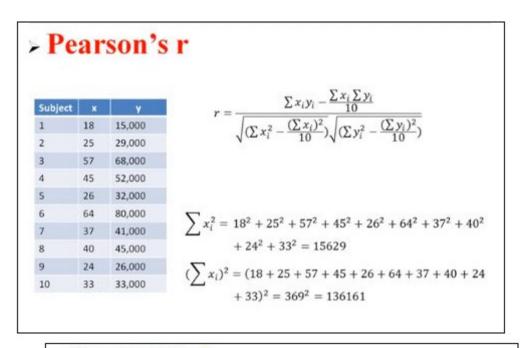
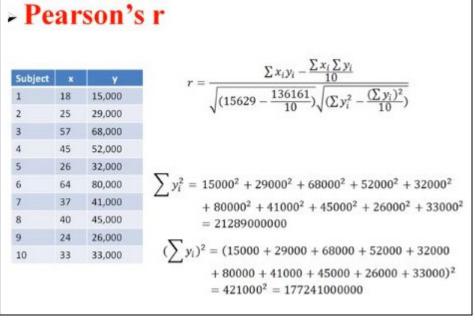
PEARSON CORRELATION

When two sets of data are strongly linked together we say they have a **High Correlation**.

For Example, consider two entities called X and Y. Where X represents the **age** and Y represents the **Salary**. Now let's see, If there is any correlation between the Age and salary is existed or not using the Pearson correlation equation.





> Pearson's r

Subject	×	У
1	18	15,000
2	25	29,000
3	57	68,000
4	45	52,000
5	26	32,000
6	64	80,000
7	37	41,000
8	40	45,000
9	24	26,000
10	33	33,000

$$r = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{10}}{\sqrt{(15629 - \frac{136161}{10})} \sqrt{(14889640000 - \frac{177241000000}{10})}}$$

> Pearson's r

Subject	×	Y	(x)(y)
1	18	15,000	270,000
2	25	29,000	725,000
3	57	68,000	3,876,000
4	45	52,000	2,340,000
5	26	32,000	832,000
6	64	80,000	5,120,000
7	37	41,000	1,517,000
8	40	45,000	1,800,000
9	24	26,000	624,000
10	33	33,000	1,089,000

$$r = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{10}}{\sqrt{(15629 - \frac{136161}{10})} \sqrt{(21289000000 - \frac{177241000000}{10})}}$$

> Pearson's r

Subject			(x)(y)
1	18	15,000	270,000
2	25	29,000	725,000
3	57	68,000	3,876,000
4	45	52,000	2,340,000
5	26	32,000	832,000
6	64	80,000	5,120,000
7	37	41,000	1,517,000
8	40	45,000	1,800,000
9	24	26,000	624,000
10	33	33,000	1,089,000
Sum	369	421,000	18,193,000

$$r = \frac{\sum x_i y_i - \frac{\sum x_i \sum y_i}{10}}{\sqrt{(15629 - \frac{136161}{10})} \sqrt{(21289000000 - \frac{177241000000}{10})}}$$

> Pearson's r

Subject			(x)(y)
1	18	15,000	270,000
2	25	29,000	725,000
3	57	68,000	3,876,000
4	45	52,000	2,340,000
5	26	32,000	832,000
6	64	80,000	5,120,000
7	37	41,000	1,517,000
8	40	45,000	1,800,000
9	24	26,000	624,000
10	33	33,000	1,089,000
Sum	369	421,000	18,193,000

$$r = \frac{18193000 - \frac{(369)(421000)}{10}}{\sqrt{(15629 - \frac{136161}{10})}\sqrt{(21289000000 - \frac{177241000000}{10})}}$$

 $r = \frac{18193000 - 15534900}{(44.865)(59706.78)} = 0.99$

x(Age) and y(Yearly Income) have a strong positive relationship.