

The Pearson correlation similarity of two users  $x, y$  is defined as

$$simil(x, y) = \frac{\sum_{i \in I_{xy}} (r_{x,i} - \bar{r}_x)(r_{y,i} - \bar{r}_y)}{\sqrt{\sum_{i \in I_{xy}} (r_{x,i} - \bar{r}_x)^2 \sum_{i \in I_{xy}} (r_{y,i} - \bar{r}_y)^2}}$$

$$\text{sim}(x,y) = \frac{\sum(i \text{ in co-rated-items, } (r_{x,i} - r_{x\text{-avg}}) * (r_{y,i} - r_{y\text{-avg}}))}{\sqrt{(\sum(i \text{ in co-rated items, } (r_{x,i} - r_{x\text{-avg}})^2) * \sum(i \text{ in co-rated items, } (r_{y,i} - r_{y\text{-avg}})^2))}}$$

### Example

	Blues Traveler	Norah Jones	Phoenix	The Strokes	Weird Al
Clara	4.75	4.5	5	4.25	4
Robert	4	3	5	2	1

$r_{x\text{-avg}}$  is  $(4.75 + 4.5 + 5 + 4.25 + 4) / 5 = 4.5$

$r_{y\text{-avg}}$  is  $(4 + 3 + 5 + 2 + 1) / 5 = 3.0$

**numerator** =  $(4.75 - 4.5) * (4 - 3.0) +$   
 $(4.5 - 4.5) * (3 - 3) +$   
 $(5 - 4.5) * (5 - 3) +$   
 $(4.25 - 4.5) * (2 - 3) +$   
 $(4 - 4.5) * (1 - 3) =$   
 $(.25 * 1) + (0 * 0) + (.5 * 2) + (-.25 * -1) + (-.5 * -2) = \mathbf{2.5}$

**first sum in denominator** =  
 $(4.75 - 4.5)^2 +$   
 $(4.5 - 4.5)^2 +$   
 $(5 - 4.5)^2 +$   
 $(4.25 - 4.5)^2 +$   
 $(4 - 4.5)^2 =$   
 $.25^2 + 0^2 + 0.5^2 + (-.25)^2 + (-.5)^2 = \mathbf{0.625}$

**second sum in denominator** =  
 $(4 - 3.0)^2 +$   
 $(3 - 3)^2 +$   
 $(5 - 3)^2 +$   
 $(2 - 3)^2 +$   
 $(1 - 3)^2 =$   
 $1^2 + 0^2 + 2^2 + (-1)^2 + (-2)^2 = \mathbf{10}$

**denominator** =  $\sqrt{0.625 * 10} = \mathbf{2.5}$

**sim** =  $2.5 / 2.5 = \mathbf{1}$