(a) $R(U_1, U_2) = \frac{(3)(1)+(3)(2)+(3)(2)}{\sqrt{3^2+3^2+3^2}\ \sqrt{1^2+2^2+2^2}}$

$= 0.9623$

$R(U_1, U_3)$ = No same rating

$R(U_1, U_4) = \frac{(3)(3)+(3)(3)}{\sqrt{3^2+3^2}\ \sqrt{3^2+3^2}} = 1$

$R(U_1, U_5) = \frac{(3)(5)+(3)(5)+(3)(5)}{\sqrt{3^2+3^2+3^2}\ \sqrt{5^2+5^2+5^2}} = 1$

$R(U_4, U_8) = \frac{(3)(5)+(3)(5)+(4)(5)}{\sqrt{3^2+3^2+4^2}\ \sqrt{5^2+5^2+5^2}} = 0.9901$

$R(U_4, U_2) = \frac{(3)(1)+(3)(2)+(4)(1)}{\sqrt{3^2+3^2+4^2}\ \sqrt{1^2+2^2+1^2}}$

$= 0.9$

=) for user 1 → either user 4 or user 5

=) For user 4 → user 1

(a)

(b)

	Item 1	Item 2	Item 3	Item 4	Mean
User 1	0	0	?	0	3
User 2	-0.5	0.5	-0.5	0.5	1.5
User 3	?	?	0	?	3
User 4	-0.333	-0.333	0.667	?	3.333
User 5	0	0	0	0	5

(b)

$$\text{Adj cosine}(I_3, I_1) = \frac{(-0.5)(0.5) + (-0.333)(0.667) + 0}{\sqrt{0.5^2 + 0.333^2} \cdot \sqrt{0.5^2 + 0.667^2}}$$

$$= \frac{0.028}{0.601 \times 0.8336}$$

$$= 0.05567$$

$$\text{Adj cosine}(I_3, I_2) = \frac{(-0.5)(0.5) + (-0.333)(0.667)}{\sqrt{0.5^2 + 0.6667^2} \cdot \sqrt{0.5^2 + 0.333^2}}$$

$$= \frac{0.47211}{0.8336 \times 0.601}$$

$$= 0.9424$$

$$\text{Adj cosine}(I_3, I_4) = \frac{(-0.5)(0.5)}{\sqrt{0.5^2}\sqrt{0.5^2}} = -1$$

$$R(U_1, I_3) = \frac{(0.05567)(3) + (0.9424)(3)}{|0.05567| + |0.9424|}$$

$R(U_1, I_3) = 3$

Adj cosine $(I, I_2) = \dfrac{0 + (0.5)(0.5) + (0.333)(0.333)}{\sqrt{0.5^2 + 0.333^2}\ \sqrt{0.5 + 0.333^2}}$

only rating given by User 3 is 3 which is for item 3.

So,

$R(U_3, I_1) = 3$

Adj cosine $(I_4, I_1) = \dfrac{(0.5)(-0.5)}{\sqrt{0.5^2}\ \sqrt{0.5^2}} = -1$

Adj cosine $(I_4, I_2) = \dfrac{(0.5)(0.5)}{\sqrt{0.5^2}\sqrt{0.5^2}} = 1$

Adj cosine $(I_4\ I_3) = -1$

$R(U_4, I_4) = \dfrac{(1)(3)}{1} = 3$

(c) User 3 may be least reliable in terms of their provided rating. Also it has so much unknow rating. User 2 is the best choice because it rated all the item and also it has different type of rating based on user's likeness of item.

(d) Using the median we can make prediction more robust and accurate in certain case. There are many missing value in the table, when calculating the mean these values will be treated as 0 but in median we can simply ignore these values. The median is better choice where the dataset has outliers or missing values that could significantly skew the mean.