

1.

a.

	Age	Age stand	Sex	Survived
A	27	0	M	0
B	30	0.056604	F	1
C	80	1	F	1
D	50	0.433962	M	0
E	60	0.622642	F	0
F	70	0.811321	F	1

Calculating the distances:

$$AG = 1.131$$

$$BG = 0.472$$

$$CG = 0.472$$

$$DG = 1.004$$

$$EG = 0.094$$

$$FG = 0.283$$

EG is the closest neighbor, thus I predict that G does not survive.

If 3NN is used, EG, FG, and either one of BG or CG are used. In this case, the majority outcome is survival, thus I predict that G survives.

b.

The distances are the same, so the nearest neighbors are the same.

For 1NN, the predicted salary is 10k.

For 3NN, the predicted salary, either B or C can be chosen as the last neighbor. If B is chosen, the average and predicted salary is 60k. If C is chosen, the salary is instead 66.67k.

2.

a.

With uniform weights

$$\text{Prediction of } E = \frac{1(Y_A) + 1(Y_B) + 1(Y_C)}{3} = \frac{1 + 0 + 1}{3} = \frac{2}{3}$$

Since the prediction is greater than 0.5, the prediction of E is X.

b.

With distance weights

$$\text{Prediction} = \frac{\frac{1}{DA}(Y_A) + \frac{1}{DB}(Y_B) + \frac{1}{DC}(Y_C)}{\frac{1}{DA} + \frac{1}{DB} + \frac{1}{DC}} = \frac{\frac{1}{3}(1) + \frac{1}{3}(1)}{\frac{1}{3} + \frac{1}{1.4} + \frac{1}{3}} = 0.483$$

Since the prediction is less than 0.5, the prediction of E is O

c.

Distance weights, 4NN

$$\text{Prediction} = \frac{\frac{1}{DA}(Y_A) + \frac{1}{DB}(Y_B) + \frac{1}{DC}(Y_C) + \frac{1}{DD}(Y_D)}{\frac{1}{DA} + \frac{1}{DB} + \frac{1}{DC} + \frac{1}{DD}} = \frac{\frac{1}{3}(1) + \frac{1}{3}(1) + \frac{1}{4}(1)}{\frac{1}{3} + \frac{1}{1.4} + \frac{1}{3} + \frac{1}{4}} = 0.562$$

Since the prediction is greater than 0.5, the prediction of E is X.

3.

a.

Using Manhattan distance,

$$d_{A1} = 6$$

$$d_{A2} = 5$$

$$d_{A3} = 4$$

$$d_{A4} = 8$$

The two nearest neighbors are user 3 and user 2. The average rating between these two for item 5 is 4.5.

Thus the item is recommended.

b.

Using cosine similarity and 3NN

$$s_{51} = \frac{(3 * 3) + (5 * 2) + (4 * 3) + (2 * 1)}{\sqrt{3^2 + 5^2 + 4^2 + 2^2} * \sqrt{3^2 + 2^2 + 3^2 + 1^2}} = 0.936$$

$$s_{52} = 0.821$$

$$s_{53} = 0.803$$

$$s_{54} = 0.924$$

$s_{51}$ ,  $s_{52}$ , and  $s_{54}$  are the three nearest neighbors. The rating on item 5 then is Alice's average ratings for item 1, 2, and 4, which is  $\frac{5+3+4}{3} = 4$ . Thus, item 5 should be recommended to Alice.