Solution 1: ROC metrics

a) First, sort the table:

ID	True class	Score	Predicted class
6	1	0.63	1
7	1	0.62	1
10	0	0.57	1
4	1	0.38	0
1	0	0.33	0
8	1	0.33	0
2	0	0.27	0
5	1	0.17	0
9	0	0.15	0
3	0	0.11	0

	True 1	True 0
Predicted 1	2	1
Predicted 0	3	4

so we get

#FN	#FP	#TN	#TP
3	1	4	2

b)

$$\rho_{\text{PPV}} = \frac{\#\text{TP}}{\#\text{TP} + \#\text{FP}} = \frac{2}{3}$$

$$\rho_{\text{NPV}} = \frac{\#\text{TN}}{\#\text{TN} + \#\text{FN}} = \frac{4}{7}$$

$$\rho_{\text{TPR}} = \frac{\#\text{TP}}{\#\text{TP} + \#\text{FN}} = \frac{2}{5}$$

$$\rho_{\text{FPR}} = \frac{\#\text{FP}}{\#\text{TN} + \#\text{FP}} = \frac{1}{5}$$

$$\rho_{\text{ACC}} = \frac{\#\text{TP} + \#\text{TN}}{\#\text{TP} + \#\text{TN} + \#\text{FP} + \#\text{FN}} = \frac{6}{10}$$

$$\rho_{\text{MCE}} = \frac{\#\text{FP} + \#\text{FN}}{\#\text{TP} + \#\text{TN} + \#\text{FP} + \#\text{FN}} = \frac{4}{10}$$

$$\rho_{F1} = \frac{2 \cdot \rho_{\text{PPV}} \cdot \rho_{\text{TPR}}}{\rho_{\text{PPV}} + \rho_{\text{TPR}}} = 0.5$$

c) First we sort the results by score:

	true_labels	scores
6	1	0.63
7	1	0.62
10	0	0.57
4	1	0.38
1	0	0.33
8	1	0.33
2	0	0.27
5	1	0.17
9	0	0.15
3	0	0.11

Here we see that $\frac{1}{n_+} = \frac{1}{n_-} = 0.2$. Now we follow the algorithm as described in the lecture slides:

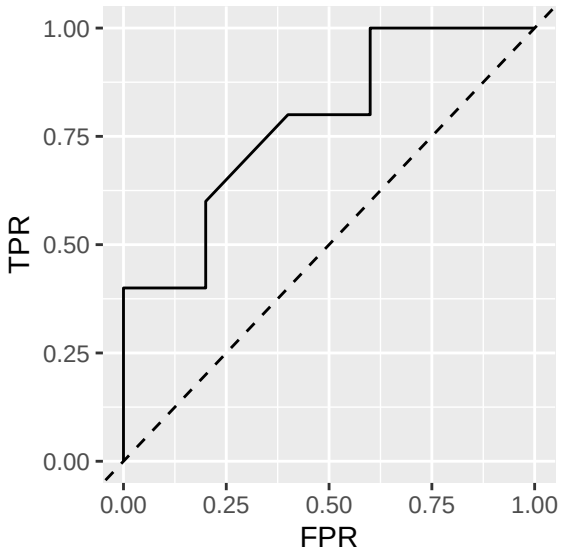
1) $c = 1 \implies$ we start in $(0,0)$ and predict everything as negative, so TPR 0 and FPR 0.
2) $c = 0.625 \implies$ TPR $0 + \frac{1}{n_+} = 0.2$ and FPR 0 (obs 6 correctly classified).
3) $c = 0.6 \implies$ TPR $0.2 + \frac{1}{n_+} = 0.4$ and FPR 0 (obs 7 correctly classified).
4) $c = 0.5 \implies$ TPR 0.4 and FPR $0 + \frac{1}{n_-} = 0.2$ (obs 10 misclassified).
5) $c = 0.35 \implies$ TPR $0.4 + \frac{1}{n_+} = 0.6$ and FPR 0.2 (obs 4 correctly classified).
6) $c = 0.3 \implies$ TPR $0.6 + \frac{1}{n_+} = 0.8$ and FPR $0.2 + \frac{1}{n_-} = 0.4$ (obs 8 correct but obs 1 misclassified).
7) $c = 0.2 \implies$ TPR 0.8 and FPR $0.4 + \frac{1}{n_-} = 0.6$ (obs 2 misclassified).
8) $c = 0.16 \implies$ TPR $0.8 + \frac{1}{n_+} = 1$ and FPR 0.6 (obs 5 correctly classified).
9) $c = 0.14 \implies$ TPR 1 and FPR $0.6 + \frac{1}{n_-} = 0.8$ (obs 9 misclassified)
10) $c = 0.09 \implies$ TPR 1 and FPR 1 (obs 3 misclassified).

Therefore we get the polygonal path consisting of the ordered list of vertices

$$(0,0), (0.2,0), (0.4,0), (0.4,0.2), (0.6,0.2), (0.8,0.4), (0.8,0.6), (1,0.6), (1,0.8), (1,1).$$

```
library(ggplot2)
roc_data <- data.frame(
  TPR = c(0, 0.2, 0.4, 0.4, 0.6, 0.8, 0.8, 1, 1, 1),
  FPR = c(0, 0, 0, 0.2, 0.2, 0.4, 0.6, 0.6, 0.8, 1))

ggplot(roc_data, aes(x = FPR, y = TPR)) + geom_line() +
  geom_abline(slope = 1, intercept = 0, linetype = 'dashed')
```



We see that the resulting ROC curve is distinct from the diagonal marking a purely random classifier, but also not too great. The step function character is clearly visible for so few observations (the non-axis-parallel part in the middle is due to the fact that we have two observations with the same score but different true class, so both TPR and FPR go up when we move from $c = 0.35$ to $c = 0.3$).

d) We can compute the AUC by adding rectangular and triangular areas, s.t.

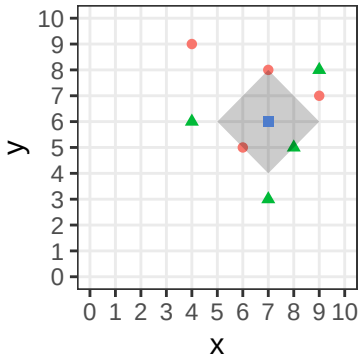
$$\rho_{\text{AUC}} = 0.2 \cdot 0.4 + 0.2 \cdot 0.6 + \tfrac{1}{2} \cdot 0.2 \cdot 0.2 + 0.2 \cdot 0.8 + 0.4 \cdot 1 = 0.78.$$

e) Not at all, because the ROC curve is drawn by iterating through *all* thresholds, and the corresponding AUC does not depend on a particular choice of c.

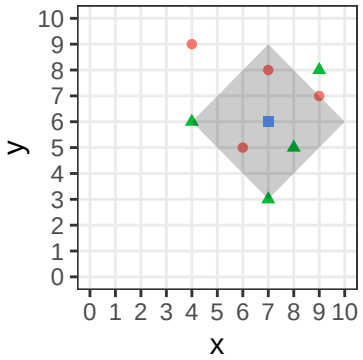
Solution 2: k-NN

a) k-NN classification

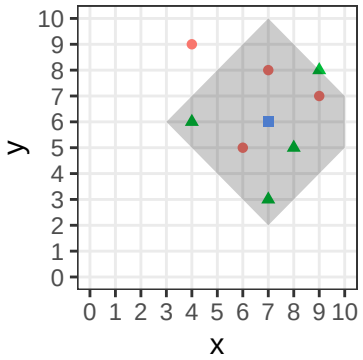
i) $k = 3 \implies$ 2 circles and 1 triangle, so we predict "circle".



b) $k = 5 \implies$ 3 circles and 3 triangles, we have to specify beforehand what to do in case of a tie.



c) $k = 7 \implies$ 3 circles and 4 triangles, so we predict "triangle".



b) k-NN regression

We now consider both unweighted and weighted predictions. Recall that weights are computed based on the distance between the point of interest and its respective neighbors. With the Manhattan, or "city block" metric, the distance can be read from the plot by walking along the grid lines (shortest way). For example, in the 3-neighborhood, all points have a distance of 2 from our square, so all get weights $\frac{1}{2}$.

i) $k = 3$

$$\begin{aligned}\hat{y} &= \frac{2+2+4}{3} = \frac{8}{3} \approx 2.67 \\ \hat{y}_{\text{weighted}} &= \frac{\frac{1}{2}\cdot 2 + \frac{1}{2}\cdot 2 + \frac{1}{2}\cdot 4}{\frac{3}{2}} = \frac{8}{3} \approx 2.67\end{aligned}$$

ii) $k = 5$

$$\begin{aligned}\hat{y} &= \frac{3\cdot 2 + 3\cdot 4}{6} = 3 \\ \hat{y}_{\text{weighted}} &= \frac{\frac{1}{2}\cdot 2 + \frac{1}{2}\cdot 2 + \frac{1}{3}\cdot 2 + \frac{1}{2}\cdot 4 + \frac{1}{3}\cdot 4 + \frac{1}{3}\cdot 4}{\frac{5}{2}} = \frac{44}{15} \approx 2.93\end{aligned}$$

iii) $k = 7$

$$\begin{aligned}\hat{y} &= \frac{3\cdot 2 + 4\cdot 4}{7} = \frac{22}{7} \approx 3.14 \\ \hat{y}_{\text{weighted}} &= \frac{\frac{1}{2}\cdot 2 + \frac{1}{2}\cdot 2 + \frac{1}{3}\cdot 2 + \frac{1}{2}\cdot 4 + \frac{1}{3}\cdot 4 + \frac{1}{3}\cdot 4 + \frac{1}{4}\cdot 4}{\frac{11}{4}} = \frac{100}{33} \approx 3.03\end{aligned}$$