# Answer Key

| | | | | |
|---|---|---|---|---|
| 1. F | 3. I | 5. G | 7. E | 9. B |
| 2. J | 4. D | 6. H | 8. A | 10. C |
| 11. B | 15. A | 19. D | 23. B | 27. A |
| 12. A | 16. C | 20. B | 24. D | 28. B |
| 13. C | 17. C | 21. C | 25. B | 29. B |
| 14. D | 18. C | 22. A | 26. D | 30. B |

31. (a) 90% confidence interval yields a z-value of 1.645. The interval is then given by $\mu \pm \frac{z\sigma}{\sqrt{n}}$. Final answer: $37.7 \pm 1.5$ or 36.2 to 39.2 (2pt). We are 90% sure that the true mean number of players lies within this interval (2pt).

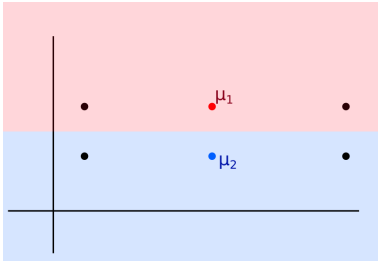
    (b) 145 more samples (245 total). Will accept $\pm 5$ samples.

32. (a)
$$L = \frac{1}{n}\sum^{n}(y - (mx + b))^2$$

    (b) $m = 0.8 \pm 0.2$ (2pt), $b = 20 \pm 10$ (2pt)

    (c) Square it (1pt). It estimates the fraction of the variance in Y that is explained by its relation to X (2pt).

33. Many answers accepted. Sample shown below. The blue and red regions are the final clusters, the $\mu_i$ are the means at their converged values. Because the algorithm minimizes distances within clusters, this situation is at a local minimum and cannot converge to the global minimum (i.e. clustering the left points together and the right points together):




34. (a) Supervised (1pt). It's performing a regression task, so it needs labelled data (1pt).

    (b) L2 loss (mean squared error) or L1 loss (mean absolute error)

    (c) i. Independent variables: learned parameters. Dependent variable: loss.

    ii. A possible state of the neural network.

    iii. Since the loss landscape of the training set is only an approximation of the "true" loss landscape, the model can converge to a minimum in the training landscape which isn't a minimum in the "true" landscape.

    (d) No, CNNs are effective for vision tasks (since they preserve locality and involve learning filters) and other tasks where the features have geometric structure. Here, the features don't have geometric structure.