Q1

If $L_1 \leq_p L_2$, then there exists a function f such that:

$$x \in L_1 \iff f(x) \in L_2$$

and if $L_2 \leq_p L_3$, then there exists a function g such that:

$$x \in L_2 \iff f(x) \in L_3$$

where f and g are polynomial-time computable functions. Since both are polynomial-time computable, we can combine f and g (i.e. perform a composition of f and g) to reduce L_1 to L_3, done as follows:

$$x \in L_1 \iff g(f(x)) \in L_3$$

$g(f(x))$ is polynomial-time computable. Thus, $L_1 \leq_p L_3$.

Q2

If a graph has a clique of size n, then each vertex in the clique is adjacent to each other $n-1$ vertices in the clique. A vertex chosen at random has a color, and the other $n-1$ vertices each have a different color since they are all adjacent to one another. See Figure 1.
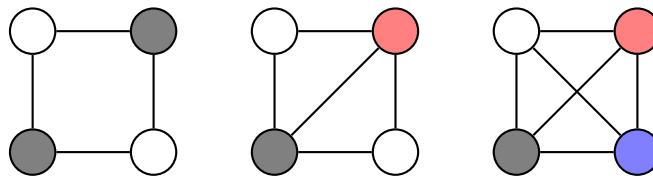

Figure 1: The left square is clique size two and requires two colors. The middle square is clique size three and requires three colors. The right square is clique size four and requires four colors.

Thus, at least n colors are required. A graph cannot be more connected than in the case of a clique, and thus n is the minimum number of colors required, and not $n+1$, $n+2$, etc. Of course, there can be more colors required than the size of the maximal clique n, as shown in Figure 2.
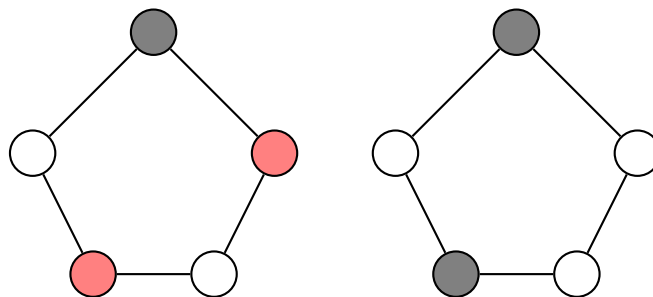

Figure 2: The left pentagon has a maximal clique size of two, but requires three colors. The right pentagon does not satisfy the requirements.

Q3

Here we show that the Efficient Recruiting problem (ER) is NP-complete (NPC). To show this, we must show that ER is NP, and that ER is NP-hard.

First, we show that ER is NP by verifying a certificate in polynomial time. This is easy to do - simply check that m coaches is less than or equal to k, and then iterate through n sports and m coaches to see if all sports are covered. This is polynomial time $O(mn)$ and so ER is NP.

Next, we show that ER is NP-hard by showing that $SET\text{-}COVER \leq_p ER$, since $SET\text{-}COVER$ is NPC. This can be done in a manner similar to $HAM\text{-}CYCLE \leq_p TSP$ from the text.

A is an instance of $SET\text{-}COVER$ where there are subsets of a set $S = \{x_1, x_2, \ldots, x_n\}$.

A' is an instance of ER where each subset is a coach and each element of a set S is a sport. We can show that ER has a subset cover if and only if there are k or fewer subsets in A' (counselors) that cover the items (sports) in set S. Since ER has a subset cover and can be shown in polynomial time, we can say that ER is NP-hard.

Since ER is NP and NP-hard, ER is also NPC.

Q4

Here we show that the Diverse Subset problem DS is NPC.

First, verify a certificate in polynomial time to show that DS is NP. This can be easily achieved by simply checking that there are k or more elements in the certificate subset and then verifying that no customers in the subset ever bought the same product. This can be achieved in $O(mn)$ time, where m is the number of items and n is the number of customers in the subset. $O(mn)$ is polynomial time, and so DS is NP.

Next, we need to show that DS is NP-hard by showing $IS \leq_p DS$.
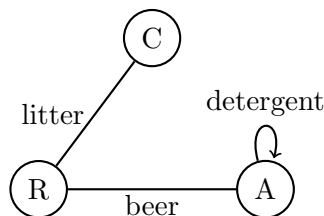

Figure 3: Graph representation of Table 1 from prompt.

To show the reduction, let's first phrase the problem as a decision question. "Given an $m \times n$ array as defined in the prompt and a number $k \leq m$, is there a subset of at least k customers that is diverse?"

Construct a graph G' which is an instance of G from IS. Each vertex in G' represents a customer and each edge in G' represents a product. Note that there can be multiple edges between vertices, and also edges pointing to the same vertex. We can show that G' has a diverse subset of size k if and only if there is an independent set of at least size k. Since there is an independent set, we can say that DS is NP-hard.

Since *DS* is NP and NP-hard, *DS* is also NPC.

From the example in Table 1 in the prompt, we construct a graph as shown in Figure 3. Alanis and Chelsea form an independent set, as they have never purchased the same product and there is no edge between them. Thus, Alanis and Chelsea form a diverse set.