You may work in groups, but you must write solutions yourself. List collaborators on your submission.

To prove that a problem $X$ is NP-complete, don’t forget to:

- Show that $Y \leq_P X$ where $Y$ is a known NP-complete problem
- Prove that the reduction is correct (it outputs Yes on a “yes” instance, and No on a “no” instance).
- Prove that $X$ belongs to NP.

Most of the work is done after you give the reduction and prove it is correct, which you already know how to do. In Problems 2 and 3, the hint tells you which NP-complete problem $Y$ to use. After Monday, you will understand how to argue that $X$ is in NP—this is very short.

Submission instructions. This assignment is due by noon on Thursday, December 6 in Gradescope (as a pdf file). Please review the course policies on the course home page about Gradescope submissions.

1. (10 points) Interval Scheduling. For each of the two questions below, decide whether the answer is (i) “Yes”, (ii) “No” or (iii) “Unknown, because it would resolve the question of whether $P = \text{NP}$”. Explain your answer. (Hint: don’t use answer (ii) “No”.)

(a) Let’s define the decision version of the Interval Scheduling Problem from Chapter 4 as follows: Given a collection of intervals on a time-line, and a bound $k$, does the collection contain a subset of nonoverlapping intervals of size at least $k$?

Question: Is it the case that Interval Scheduling $\leq_P \text{Vertex Cover}$?

(b) Question: Is it the case that Independent Set $\leq_P \text{Interval Scheduling}$

2. (10 points) Diverse Subset. A store trying to analyze the behavior of its customers will often maintain a two-dimensional array $A$, where the rows correspond to its customers and the columns correspond to the products it sells. The entry $A[i,j]$ specifies the quantity of product $j$ that has been purchased by customer $i$.

Here’s a tiny example of such an array $A$.

<table>
<thead>
<tr>
<th></th>
<th>detergent</th>
<th>beer</th>
<th>diapers</th>
<th>cat litter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raj</td>
<td>0</td>
<td>6</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>Alanis</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Chelsea</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>7</td>
</tr>
</tbody>
</table>

One thing that a store might want to do with this data is the following. Let us say that a subset $S$ of the customers is diverse if no two of the of the customers in $S$ have ever bought the same product (i.e., for each product, at most one of the customers in $S$ has ever bought it). A diverse set of customers can be useful, for example, as a target pool for market research.

We can now define the DIVERSE-SUBSET Problem as follows: Given an $m \times n$ array $A$ as defined above, and a number $k \leq m$, is there a subset of at least $k$ of customers that is diverse?

Show that DIVERSE-SUBSET is NP-complete. (Hint: reduce from INDEPENDENT-SET.)
3. **(10 points) Hitting Set.** Consider a set $A = \{a_1, \ldots, a_n\}$ and a collection $B_1, \ldots, B_m$ of subsets of $A$ (i.e., $B_i \subseteq A$ for all $i$). We say that $H \subseteq A$ is a hitting set if $H$ contains at least one element from each $B_i$, that is $H \cap B_i$ is non-empty for all $i$ (so $H$ “hits” all the sets $B_i$).

The **Hitting-Set** problem is the following: Given a set $A = \{a_1, \ldots, a_n\}$, subsets $B_1, \ldots, B_m \subseteq A$, and a number $k$, is there a hitting set $H \subseteq A$ of size at most $k$?

Prove that **Hitting-Set** is NP-Complete. (Hint: reduce from **Vertex-Cover.**)