

The problem is in \mathcal{NP} because we can exhibit a set of k customers, and in polynomial time it can be checked that no two bought any product in common.

We now show that *Independent Set* \leq_P *Diverse Subset*. Given a graph G and a number k , we construct a customer for each node of G , and a product for each edge of G . We then build an array that says customer v bought product e if edge e is incident to node v . Finally, we ask whether this array has a diverse subset of size k .

We claim that this holds if and only if G has an independent set of size k . If there is a diverse subset of size k , then the corresponding set of nodes has the property that no two are incident to the same edge — so it is an independent set of size k . Conversely, if there is an independent set of size k , then the corresponding set of customers has the property that no two bought the same product, so it is diverse.

¹ex640.690.659