Problem 2. In a class with $n > 2$ students the teacher wants to assign to each student some topics to work on in such a way that any two students have a unique common topic assigned, each topic is given to more than one student but no topic is assigned to all the students. Show that the teacher has to use at least n topics.

Solution. Let S be the set of students and T the set of topics. Thus $|S| = n > 2$. Let $m = |T|$. Our task is to prove that $m \geq n$. For any topic $t \in T$ let S_t be the set of all students who have been assigned topic t and for any student $s \in S$ let T_s be the set of all topics assigned to s. Finally, for any two students s, z let $t(s,z)$ be the unique topic assigned to both of them.

We start with the following observation:

Lemma 1. If a topic t is not assigned to a student s (i.e. $s \notin S_t$) then $|T_s| \geq |S_t|$.

To prove the lemma note that for any student $z \in S_t$ the topic $t(s,z)$ is in T_s and $t(s,z) \neq t$. If x is another student in S_t then $t(x,z) = t$ and therefore $t(s,x) \neq t(s,z)$ (the equality $t(s,x) = t(s,z)$ would mean that $t = t(x,z) = t(s,x) = t(s,z) \neq t$, which is false). This shows that the assignment $z \mapsto t(s,z)$ is an injective function from S_t to T_s and therefore $|T_s| \geq |S_t|$.

Our second observation is the following counting formula.

Lemma 2. $\sum_{t\in T} |S_t| = \sum_{s\in S} |T_s|$.

To justify Lemma 2 note that both sides of the equality in Lemma 2 count the number of pairs (s,t) such that $s \in S_t$ (or, equivalently, $t \in T_s$).

We are ready now to start our proof that $m \geq n$. Let k be largest positive integer such that there exists k different topics $t_1, \ldots, t_k$ and k different students $s_1, \ldots, s_k$ such that t_i is not assigned to s_i (i.e. $s_i \notin S_{t_i}$) for $i = 1, 2, \ldots, k$.

Suppose that $k = m$. Then by lemma 1 we get

$$\sum_{t\in T} |S_t| = \sum_{j=1}^{m} |S_{t_j}| \leq \sum_{j=1}^{m} |T_{s_j}| \leq \sum_{s\in S} |T_s|.$$

By Lemma 2 we conclude that the weak inequalities in the above formula are actually equalities. Since $|T_s| > 0$ for every s, we conclude that $S = \{s_1, \ldots, s_m\}$ and $|T_{s_i}| = |S_{t_i}|$ for every i. In particular, we have $m = n$ in this case.

Suppose now that $k < m$. If $k = n$ then clearly $n < m$. Suppose then that $k < n$. Since $k < m$ there is a topic t different from any of the topics $t_1, \ldots, t_k$. If there is a student s different from students $s_1, \ldots, s_k$ which does not have topic t assigned then we could take $t_{k+1} = t$, $s_{k+1} = s$ and contradict the definition of k. Therefore topic t is assigned to every student different from students $s_1, \ldots, s_k$. Therefore t cannot be assigned to each of the students $s_1, \ldots, s_k$ (as no topic is assigned to all the students). After renumbering the students and topics if necessary, we may assume that for some $r > 0$ the topic t is not assigned to students $s_1, \ldots, s_r$ and it is assigned to the remaining students. Now consider any student s different from $s_1, \ldots, s_k$. If t_1 is not assigned to s then we can take $t_{k+1} = t$, $s_{k+1} = s_1$ and replace s_1 with s to contradict the definition of k. Thus t_1 is assigned to every student different from $s_1, \ldots, s_k$. We see that both t_1 and t are assigned to every student different from $s_1, \ldots, s_k$. Since $t_1 \neq t$, there can be at most one student different from $s_1, \ldots, s_k$ (otherwise two students would share more than one topic). This means that $n = k + 1$. It follows that $m \geq n$ (as $m > k$). This completes our argument.

Exercise. Show that k actually must be equal to n by extending the last part of our proof to argue that the assumption $k < n$ leads to a contradiction.

Second solution (after Emmett Wyman). Suppose that the set of students is $S = \{s_1, \ldots, s_n\}$ and the set of topics is $T = \{t_1, \ldots, t_m\}$. Consider the $n \times m$ matrix $M = (m_{i,j})$, where $m_{i,j} = 1$ if topic t_j is assigned to student s_i, and $m_{i,j} = 0$ otherwise. Consider any two rows of M, say row i and row k. Since students s_i and s_k have exactly one common topic, there is exactly one column in which rows i and k have both entry 1. This means that the dot product of row i and row k is 1. Note that every student must have at least two topics assigned: if a student had just one topic assigned, this topics would be shared with all the other students and therefore we would have a topic assigned to all students. It follows that the dot product of any row with itself is at least 2. These observations can be rephrased as follows: the $n \times n$ matrix $T = MM^t$ has all non-diagonal entries equal to 1 and all diagonal entries greater than or equal to 2 (here M^t denotes the transpose of the matrix M). Suppose that the diagonal entries of T are $a_1, \ldots, a_n$. Then for any vector $(x_1, \ldots, x_n) \in \mathbb{R}^n$ we have

$$T(x_1, \ldots, x_n)^t = (l + (a_1 - 1)x_1, \ldots, l + (a_n - 1)x_n),$$

where $l = x_1 + \ldots + x_n$. If $T(x_1, \ldots, x_n)^t = 0$ then $l = (1 - a_i)x_i$ for $i = 1, \ldots, n$. Thus $lx_i = (1 - a_i)x_i^2 \leq -x_i^2$ (recall that $a_i \geq 2$). Adding these inequalities yields $l^2 \leq -(x_1^2 + \ldots + x_n^2) \leq 0$. Thus $l = 0$ and therefore $x_1 = \ldots = x_n = 0$. This shows that the matrix T is non-singular. Now we can easily see that $m \geq n$. Indeed, if $m < n$ then there would exists a non-zero vector $(x_1, \ldots, x_n) \in \mathbb{R}^n$ such that $M^t(x_1, \ldots, x_n)^t = 0$, so also $T(x_1, \ldots, x_n)^t = 0$, contrary to our observation and T is non-singular.

Remark. The result of Problem 2 is actually a well know theorem due to de Bruijn and Erdös. It is often stated as follows:

Theorem (de Bruijn - Erdös). In any finite incidence geometry the number of lines is greater or equal than the number of points.

Recall that an incidence geometry is a set Π with at least 3 elements, whose elements are called points, together with a collection of proper subsets of Π, called lines, such that any two distinct points belong to a unique line and any line has at least two points.

Problem. Suppose that n in Problem 2 is not of the form $k^2 + k + 1$ for any positive integer k. Prove that in order to use exactly n topics, the teacher has to pick one student, say s, and assign to s all but one topics. Each of the other students is then assigned the remaining topic and exactly one other topic.

Remark. It is known that if $n = k^2 + k + 1$ and k is a power of a prime number, then it is possible for the teacher to use exactly n topics in such a way that exactly $k + 1$ topics are assigned to each student. It is an open problem whether there is a k which is not a power of a prime such that for $n = k^2 + k + 1$ the teacher can use exactly n topics in a way different than the one described in the last problem. It is known that any such assignment, if it exists, would need to assign to each student exactly $k + 1$ topics.