Chapter 1

Seven Is More Than Six. The Pigeon-Hole Principle

1.1 The Basic Pigeon-Hole Principle

Seven is more than six. Four is more than three. Two is more than one. These statements do not seem to be too interesting, exciting, or deep. We will see, however, that the famous Pigeon-hole Principle makes excellent use of them. We choose to start our walk through combinatorics by discussing the Pigeon-hole Principle because it epitomizes one of the most attractive treats of this field: the possibility of obtaining very strong results by very simple means.

Theorem 1.1 (Pigeon-hole Principle). *Let n and k be positive integers, and let $n > k$. Suppose we have to place n identical balls into k identical boxes. Then there will be at least one box in which we place at least two balls.*

Proof. While the statement seems intuitively obvious, we are going to give a formal proof because proofs of this nature will be used throughout this book.

We prove our statement in an *indirect way*, that is, we assume its contrary is true, and deduce a contradiction from that assumption. This is a very common strategy in mathematics; in fact, if we have no idea how to prove something, we can always try an indirect proof.

Let us assume there is no box with at least two balls. Then each of the k boxes has either 0 or 1 ball in it. Denote by m the number of boxes that have zero balls in them; then certainly $m \geq 0$. Then, of course, there are $k - m$ boxes that have one. However, that would mean that the total number of balls placed into the k boxes is $k - m$ which is a contradiction because we had to place n balls into the boxes, and $k - m \leq k < n$. Therefore, our assumption that there is no box with at least two balls must have been false. $\square$

In what follows, we will present several applications that show that this innocuous statement is in fact a *very* powerful tool.

Example 1.2. There is an element in the sequence $7, 77, 777, 7777, \cdots$, that is divisible by 2003.

Solution. We prove that an even stronger statement is true, in fact, one of the first 2003 elements of the sequence is divisible by 2003. Let us assume that the contrary is true. Then take the first 2003 elements of the sequence and divide each of them by 2003. As none of them is divisible by 2003, they will all have a remainder that is at least 1 and at most 2002. As there are 2003 remainders (one for each of the first 2003 elements of the sequence), and only 2002 possible values for these remainders, it follows by the Pigeon-hole Principle that there are two elements out of the first 2003 that have the same remainder. Let us say that the ith and the jth elements of the sequence, a_i and a_j, have this property, and let $i < j$.

	7777777777777777777777777	j digits
−	777777777777777777	i digits
	7777777000000000000000000	j-i digits equal to 7, i digits equal to 0

Fig. 1.1 The difference of a_j and a_i.

As a_i and a_j have the same remainder when divided by 2003, there exist non-negative integers k_i, k_j, and r so that $r \leq 2002$, and $a_i = 2003k_i + r$, and $a_j = 2003k_j + r$. This shows that $a_j - a_i = 2003(k_j - k_i)$, so in particular, $a_j - a_i$ is divisible by 2003.

This is nice, but we need to show that there is an element in our sequence that is divisible by 2003, and $a_j - a_i$ is not an element in our sequence. Figure 1.1 helps understand why the information that $a_j - a_i$ is divisible by 2003 is nevertheless very useful.

Indeed, $a_j - a_i$ consists of $j - i$ digits equal to 7, then i digits equal to

0. In other words,

$$a_j - a_i = a_{j-i} \cdot 10^i,$$

and the proof follows as 10^i is relatively prime to 2003, so a_{j-i} must be divisible by 2003.

In this example, the possible values of the remainders were the boxes, all 2002 of them, while the first 2003 elements of the sequence played the role of the balls. There were more balls than boxes, so the Pigeon-hole Principle applied.

Example 1.3. A chess tournament has n participants, and any two players play one game against each other. Then it is true that in any given point of time, there are two players who have finished the same number of games.

Solution. First we could think that the Pigeon-hole Principle will not be applicable here as the number of players ("balls") is n, and the number of possibilities for the number of games finished by any one of them ("boxes") is also n. Indeed, a player could finish either no games, or one game, or two games, and so on, up to and including $n-1$ games.

The fact, however, that two players play their games *against* each other, provides the missing piece of our proof. If there is a player A who has completed all his $n-1$ games, then there cannot be any player who completed zero games because at the very least, everyone has played with A. Therefore, the values 0 and $n-1$ cannot both occur among the numbers of games finished by the players at any one time. So the number of possibilities for these numbers ("boxes") is at most $n-1$ at any given point of time, and the proof follows.

Quick Check

(1) Prove that among eight integers, there are always two whose difference is divisible by seven.

(2) A student wrote six distinct positive integers on the board, and pointed out that none of them had a prime factor larger than 10. Prove that there are two integers on the board that have a common prime divisor. Could we make the same conclusion if in the first sentence we replaced "six" by "five"?

(3) A bicycle path is 30 miles long, with four rest areas. Prove that either there are two rest areas that are at most six miles from each other,

or there is a rest area that is at most six miles away from one of the endpoints of the path.

1.2 The Generalized Pigeon-Hole Principle

It is easy to generalize the Pigeon-hole Principle in the following way.

Theorem 1.4 (Pigeon-hole Principle, general version). *Let n, m, and r be positive integers so that $n > rm$, and let us distribute n identical balls into m identical boxes. Then there will be at least one box into which we place at least $r + 1$ balls.*

Proof. Just as in the proof of Theorem 1.1, we assume the contrary statement. Then each of the m boxes can hold at most r balls, so all the boxes can hold at most $rm < n$ balls, which contradicts the requirement that we distribute n balls. $\square$

It is certainly not only in number theory that the Pigeon-hole Principle proves to be very useful. The following example provides a geometric application.

Example 1.5. Ten points are given within a square of unit size. Then there are two of them that are closer to each other than 0.48, and there are three of them that can be covered by a disk of radius 0.5.

Solution. Let us split our unit square into nine equal squares by straight lines as shown in Figure 1.2. As there are ten points given inside the nine small squares, Theorem 1.1 implies that there will be at least one small square containing two of our ten points. The longest distance within a square of side length $1/3$ is that of two opposite endpoints of a diagonal. By the Pythagorean theorem, that distance is $\frac{\sqrt{2}}{3} < 0.48$, so the first part of the statement follows.

To prove the second statement, divide our square into four equal parts by its two diagonals as shown in Figure 1.3. Theorem 1.4 then implies that at least one of these triangles will contain three of our points. The proof again follows as the radius of the circumcircle of these triangles is shorter than 0.5.

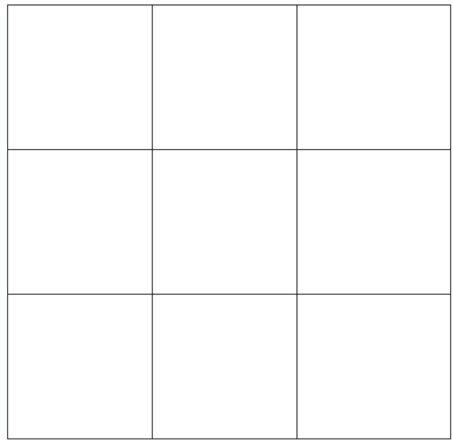

Fig. 1.2 Nine small squares for ten points.

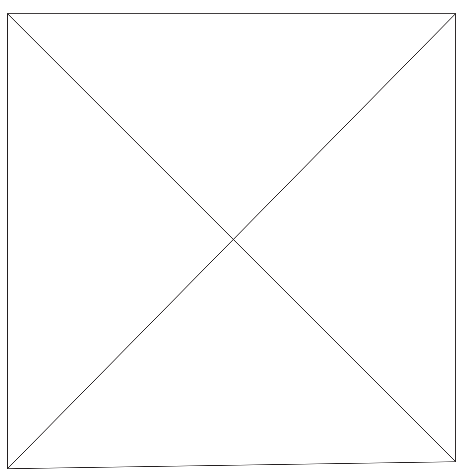

Fig. 1.3 Four triangles for ten points.

We finish our discussion of the Pigeon-hole Principle by two highly surprising applications. What is striking in our first example is that it is valid for *everybody*, not just say, the majority of people. So we might as well discuss our example choosing the reader herself for its subject.

Example 1.6. During the last 1000 years, the reader had an ancestor A such that there was a person P who was an ancestor of *both* the father and the mother of A.

Solution. Again, we prove our statement in an indirect way: we assume its contrary, and deduce a contradiction. We will use some rough estimates for the sake of shortness, but they will not make our argument any less valid.

Take the family tree of the reader. This tree is shown in Figure 1.4.

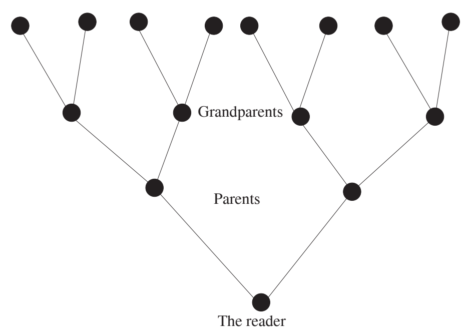

Fig. 1.4 The first few levels of the family tree of the reader.

The root of this tree is the reader herself. On the first level of the tree, we see the two parents of the reader, on the second level we find her four grandparents, and so on. Assume (for shortness) that one generation takes 25 years to produce offspring. That means that 1000 years was sufficient time for 40 generations to grow up, yielding $1 + 2 + 2^2 + \cdots + 2^{40} = 2^{41} - 1$ nodes in the family tree. If any two nodes of this tree are associated to the same person B, then we are done as B can play the role of P.

Now assume that no two nodes of the first 40 levels of the family tree coincide. Then all the $2^{41} - 1$ nodes of the family tree must be distinct. That would mean $2^{41} - 1$ distinct people, and that is a lot more than the number of all people who have lived in our planet during the last 1000 years. Indeed, the current population of our planet is less than 10^{10}, and was much less at any earlier point of time. Therefore, the cumulative population of the last 1000 years, or 40 generations, was less than $40 \cdot 10^{10} < 2^{41} - 1$, and the proof follows by contradiction.

Our assumption that every generation takes 25 years to produce offspring was a realistic one. Given that by all available data, the average life expectancy of humans is longer today than ever before, 25 seems to be a high-end estimate. The reader should spend a little time thinking about how (and if) the argument would have to be modified if 25 were to be replaced by a smaller or larger number.

Our last example comes from the theory of graphs, an extensive and important area of combinatorics to which we will devote several chapters later.

Example 1.7. Mr. and Mrs. Smith invited four couples to their home. Some guests were friends of Mr. Smith, and some others were friends of Mrs. Smith. When the guests arrived, people who knew each other beforehand shook hands, those who did not know each other just greeted each other. After all this took place, the observant Mr. Smith said "How interesting. If you disregard me, there are no two people present who shook hands the same number of times".

How many times did Mrs. Smith shake hands?

Solution. The reader may well think that this question cannot be answered from the given information any better than say, a question about the age of the second cousin of Mr. Smith. However, using the Pigeonhole Principle and a very handy model called a *graph*, this question *can* be answered.

To start, let us represent each person by a node, and let us write the number of handshakes carried out by each person except Mr. Smith next to the corresponding vertex. This way we must write down nine *different* nonnegative integers. All these integers must be smaller than nine as nobody shook hands with himself/herself or his/her spouse. So the numbers we wrote down are between 0 and 8, and since there are nine of them, we must have written down each of the numbers $0, 1, 2, 3, 4, 5, 6, 7, 8$ exactly once. The diagram we have constructed so far can be seen in Figure 1.5.

Now let us join two nodes by a line if the corresponding two people shook each other's hands. Such a diagram is called a *graph*, the nodes are called the *vertices* of the graph, and the lines are called the *edges* of the graph. So our diagram will be a graph with ten vertices.

Let us denote the person with i handshakes by Y_i. (Mr. Smith is not assigned any additional notation.) Who can be the spouse of the person Y_8? We know that Y_8 did not shake the hand of only one other person,

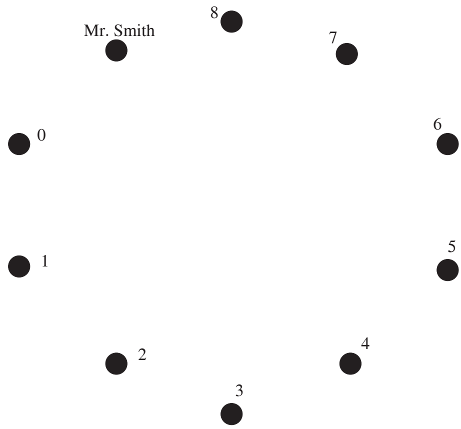

Fig. 1.5 The participants of the party.

so that person must have been his or her spouse. On the other hand, Y_8 certainly did not shake the hand Y_0 as nobody did that. Therefore, Y_8 and Y_0 are married, and Y_8 shook everyone's hand except for Y_0. We represent this by joining his vertex to all vertices other than Y_0. We also encircle Y_8 and Y_0 together, to express that they are married.

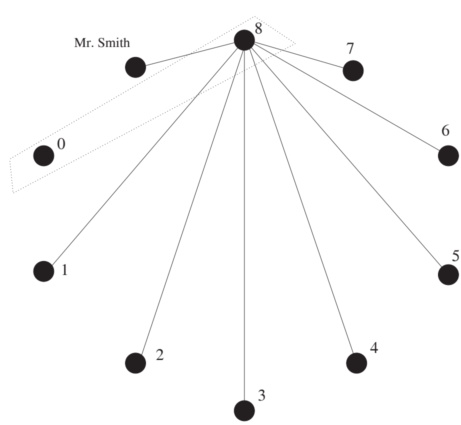

Fig. 1.6 Y_8 and Y_0 are married.

Now try to find the spouse of Y_7, the person with seven handshakes. This person did not shake the hands of two people, one of whom was his/her spouse. Looking at Figure 1.6, we can tell who these two people are. One of them is Y_0 as he or she did not shake anyone's hand, and the other one is Y_1 as he or she had only one handshake, and that was with Y_8. As spouses do not shake hands, this implies that the spouse of Y_7 is either Y_0 or Y_1. However, Y_0 is married to Y_8, so Y_1 must be married to Y_7.

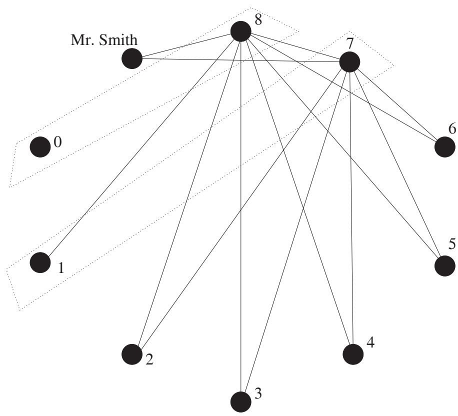

Fig. 1.7 Y_1 and Y_7 are married.

By a similar argument that the reader should be able to complete, Y_6 and Y_2 must be married, and also, Y_5 and Y_3 must be married. That implies that by exclusion, Y_4 is Mrs. Smith, therefore Mrs. Smith shook hands four times.

How did we obtain such a strong result from "almost no data"? The truth is that the data we had, that is, that all people except Mr. Smith shook hands a different number of times, is quite restrictive. Indeed, consider Example 1.3 again. An obvious reformulation of that Example shows that it is simply impossible to have a party at which no two people shake hands the same number of times (as long as no two people shake hands more than once). Example 1.7 relaxes the "all-different-numbers" requirement a little bit, by waiving it for Mr. Smith. Our argument then shows that with that extra level of freedom, we can indeed have a party satisfying the new, weaker conditions, but only in one way. That way is described by the graph shown in Figure 1.8.

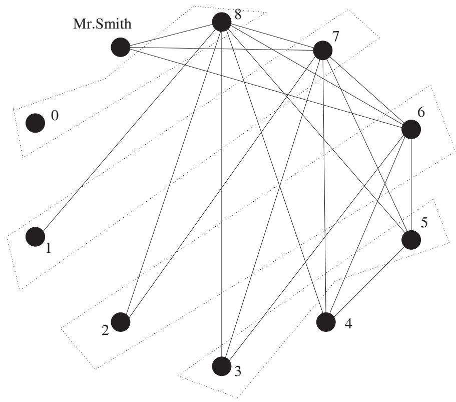

Fig. 1.8 Mrs. Smith shook hands four times.

Quick Check

(1) The product of five given polynomials is a polynomial of degree 21. Prove that we can choose two of those polynomials so that the degree of their product is at least nine.
(2) A college has 39 departments, and a total of 261 faculty members in those departments. Prove that there are three departments in this college that have a total of at least 21 faculty members.
(3) Let n be a positive integer that has exactly three prime divisors, and at least seven divisors of the form p^k, where p is a prime, and k is a positive integer. Prove that n must be divisible by the cube of an integer that is larger than 1.

Exercises

(1) A busy airport sees 1500 takeoffs per day. Prove that there are two planes that must take off within a minute of each other.
(2) Find all triples of positive integers $a < b < c$ for which

$$\frac{1}{a} + \frac{1}{b} + \frac{1}{c} = 1$$

holds.

(3) We have selected 169 points inside a regular triangle of side length 100 at random. Prove that there will be three among the selected points that span a triangle of area at most 68.

(4) (+) We have distributed two hundred balls into one hundred boxes with the restrictions that no box got more than one hundred balls, and each box got at least one. Prove that it is possible to find some boxes that together contain exactly one hundred balls.

(5) (+) Last year, the Division One basketball teams played against an average of eighteen different opponents. Is it possible to find a group of teams so that each of them played against at least ten other teams of the group?

(6)(a) The set M consists of nine positive integers, none of which has a prime divisor larger than six. Prove that M has two elements whose product is the square of an integer.

(b) (+) (Some knowledge of linear algebra and abstract algebra required.) The set A consists of $n+1$ positive integers, none of which has a prime divisor that is larger than the nth smallest prime number. Prove that there exists a non-empty subset $B \subseteq A$ so that the product of the elements of B is a perfect square.

(7) (++) The set L consists of 2003 integers, none of which has a prime divisor larger than 24. Prove that L has four elements, the product of which is equal to the fourth power of an integer.

(8) (+) The sum of one hundred given real numbers is zero. Prove that at least 99 of the pairwise sums of these hundred numbers are non-negative. Is this result the best possible one?

(9) (+) We colored all points of R^2 with integer coordinates by one of six colors. Prove that there is a rectangle whose vertices are monochromatic. Can we make the statement stronger by limiting the size of the purported monochromatic rectangle?

(10) Prove that among 502 positive integers, there are always two integers so that either their sum or their difference is divisible by 1000.

(11) (+) We chose $n+2$ numbers from the set $1, 2, \cdots, 3n$. Prove that there are always two among the chosen numbers whose difference is more than n but less than $2n$.

(12) There are four heaps of stones in our backyard. We rearrange them into five heaps. Prove that at least two stones are placed into a smaller heap.

(13) There are infinitely many pieces of paper in a basket, and there is a positive integer written on each of them. We know that no matter how

we choose infinitely many pieces, there will always be two of them so that the difference of the numbers written on them is at most ten million. Prove that there is an integer that has been written on infinitely many pieces of paper.

(14) (+)

(a) A soccer team played 30 games this year, and scored a total of 53 goals, scoring at least one goal in each game. Prove that there was a sequence of consecutive games in which the team scored exactly six goals.

(b) Prove that the claim of part (a) does not hold for a team that scored 60 goals, with the other parameters unchanged.

(c) Prove that the claim of part (a) does hold for a team that scored 59 goals, with the other parameters unchanged.

(15) (+) The set of all positive integers is partitioned into several arithmetic progressions. Show that there is at least one among these arithmetic progressions whose initial term is divisible by its difference.

(16) Sixty-five points are given inside a square of side length 1. Prove that there are three of them that span a triangle of area at most $1/32$.

(17) Let A be an $n \times n$ matrix with 0 and 1 entries only. Let us assume that $n \geq 2$, and that at least $2n$ entries are equal to 1. Prove that A contains two entries equal to 1 so that one of them is strictly above and strictly on the right of the other.

(18) A state has seven counties. In one year, three candidates run in a statewide election. Is it possible that in each county the same number of people vote, and the candidate who gets the highest number of votes statewide does not get the highest number of votes in any county?

Supplementary Exercises

(19) (-) Prove that every year contains at least four and at most five months that contain five Sundays.

(20) (-) A soccer league features 17 games for today. Let us assume that no team will score more than three goals. Prove that there will be a result that will occur more than once. (A result consists of the number of goals scored by the home team, followed by the number of goals scored by the visiting team. So 3-2 and 2-3 are considered different scores.)

(21) (-) A group of seven co-workers are trying to predict the total number of points scored in a given basketball game. The first six people al-

ready took their guesses, and, curiously, they all picked distinct even numbers. Mr. Slow is the last person to guess, and he knows all previous guesses. Is there a strategy for him that assures that his guess will be better than the guesses of half of his colleagues?

(22) (-) A soccer team scored a total of 40 goals this season. Nine players scored at least one of those goals. Prove that there are two players among those nine who scored the same number of goals.

(23)(a) In the month of April, Ms. Consistent went to the swimming pool 26 times, though she never went more than once on the same day. Is it true that there were six consecutive days when she went to the swimming pool?

(b) Same as (a), but for the month of May instead of April.

(24)(a) We select 11 positive integers that are less than 29 at random. Prove that there will always be two integers selected that have a common divisor larger than 1.

(b) Is the statement of part (a) true if we only select ten integers that are less than 29?

(25) Prove that there exists a positive integer n so that $44^n - 1$ is divisible by 7.

(26) The sum of five positive real numbers is 100. Prove that there are two numbers among them whose difference is at most 10.

(27) Find all 4-tuples (a, b, c, d) of distinct positive integers so that $a < b < c < d$ and

$$\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d} = 1.$$

(28) Complete the following sentence, that is a generalization of the Pigeonhole Principle to real numbers. "If the sum of k real numbers is n, then there must be one of them which is...". Prove your claim.

(29) We are given 17 points inside a regular triangle of side length one. Prove that there are two points among them whose distance is not more than 1/4.

(30) Prove that the sequence 1967, 19671967, 196719671967, $\cdots$, contains an element that is divisible by 1969.

(31) A teacher receives a paycheck every two weeks, always the same day of the week. Is it true that in any six consecutive calendar months she receives exactly 13 paychecks?

(32) (+) Let T be a triangle with angles of 30, 60 and 90 degrees whose hypotenuse is of length 1. We choose ten points inside T at random. Prove that there will be four points among them that can be covered

by a half-circle of radius 0.42.

(33) We select $n+1$ different integers from the set $\{1, 2, \cdots, 2n\}$. Prove that there will always be two among the selected integers whose largest common divisor is 1.

(34)(a) Let $n \geq 2$. We select $n+1$ different integers from the set $\{1, 2, \cdots, 2n\}$. Is it true that there will always be two among the selected integers so that one of them is equal to twice the other?

(b) Is it true that there will always be two among the selected integers so that one is a multiple of the other?

(35) One afternoon, a mathematics library had several visitors. A librarian noticed that it was impossible to find three visitors so that no two of them met in the library that afternoon. Prove that then it was possible to find two moments of time that afternoon so that each visitor was in the library at one of those two moments.

(36) (+) Let r be any irrational real number. Prove that there exists a positive integer n so that the distance of nr from the closest integer is less than 10^{-10}.

(37) Let p and q be two positive integers so that the largest common divisor of p and q is 1. Prove that for any non-negative integers $s \leq p-1$ and $t \leq q-1$, there exists a non-negative integer $m \leq pq$ so that if we divide m by p, the remainder is s, and if we divide m by q, the remainder is t.

(38) Does the statement of Exercise 17 remains true if we only assume that A has at least $2n-1$ entries equal to 1?

(39) (++) Let K denote the 1000 points in the three-dimensional space whose coordinates are all integers in the interval $[1, 10]$. Let S be a subset of K that has at least 272 points. Prove that S contains two points u and v so that each coordinate of v is strictly larger than the corresponding coordinate of u.

(40) Six points are given on the perimeter of a circle of radius 1. Prove that there are two among the given points whose distance from each other is at most 1.

Solutions to Exercises

(1) There are 1440 minutes per day. If our 1440 minutes are the boxes, and our 1500 planes are the balls, the Pigeon-hole Principle says that there are two balls in the same box, that is, there are two planes that

take off within a minute of each other.

(2) It is clear that $a = 2$. Indeed, $a = 1$ is impossible because then the left-hand side would be larger than 1, and $a \geq 3$ is impossible as $a < b < c$ implies $\frac{1}{a} > \frac{1}{b} > \frac{1}{c}$, so $a = 3$ would imply that the left-hand side is smaller than 1. Thus we only have to solve

$$\frac{1}{b} + \frac{1}{c} = \frac{1}{2},$$

with $3 \leq b < c$. We claim that b must take its smallest possible value, 3. Indeed, if $b \geq 4$, then $c \geq 5$, and so $\frac{1}{b} + \frac{1}{c} \leq \frac{1}{4} + \frac{1}{5} < \frac{1}{2}$. Thus $b = 3$, and therefore, $c = 6$.

(3) Split the original regular triangle into 64 congruent triangles, by repeatedly using the method of midlines. Each of these small triangles will have area 67.658. On the other hand, by the Pigeon-hole Principle, at least one of these triangles must contain at least three of the selected points.

(4) Arrange our boxes in a line so that the first two boxes do not have the same number of balls in them. We can always do this unless all boxes have two balls, in which case the statement is certainly true.
Let a_i denote the number of balls in box i, for all positive integers $1 \leq i \leq 100$. Now look at the following sums: a_1, $a_1 + a_2$, $a_1 + a_2 + a_3$, $\cdots$, $a_1 + a_2 + \cdots + a_{100}$. If two of them yield the same remainder when divided by 100, then take the difference of those two sums. That will yield a sum of type $a_i + a_{i+1} + \cdots + a_j$ that is divisible by 100, is smaller than 200, and is positive. In other words, $a_i + a_{i+1} + \cdots + a_j = 100$, so the total content of boxes $i, i+1, \cdots, j$ is exactly 100 balls.
Now assume this does not happen, that is, all sums $a_1 + a_2 + \cdots + a_k$ yield different remainders when divided by 100. Attach the one-element sum a_2 to our list of sums. Now we have 101 sums, so by Theorem 1.1, two of them must have the same remainder when divided by 100. Since we assumed this did not happen before a_2 joined the list, we know that there is a sum S on our list that has the same remainder as a_2. As we know that $a_1 \neq a_2$, we also know that $S \neq a_1$, and we are done as in the previous paragraph, since $S - a_2 = a_1 + a_3 + \cdots + a_t = 100$.
We note that this argument works in general with $2n$ boxes and $4n$ balls. We also note that we in fact proved a stronger statement as our chosen boxes are almost consecutive.

(5) Yes. Take a team T that played against at most nine opponents. If there is no such team, then the group of all Division One teams has the

required property, and we are done. Omit T; we claim that this will not decrease the average number of opponents. Indeed, as we are only interested in the number of opponents played (and not games), we can assume that any two teams played each other at most once. The 18-game-average means that all the m Division One teams together played $9m$ games as a game involves two teams. Omitting T, we are left with $m-1$ teams, who played a grand total of at least $9m-9$ games. This means that the remaining teams still played at least 18 games on average against other remaining teams.

Now iterate this procedure- look for a team from the remaining group that has only played nine games and omit it. As the number of teams is finite, this elimination procedure has to come to an end. The only way that can happen is that there will be a group of which we cannot eliminate any team, that is, in which every team has played at least ten games against the other teams of the group.

(6)(a) Each element of M can be written as $2^i3^j5^k$ for some non-negative integers i, j, k. Therefore, we can divide the elements of M into eight classes according to the *parity* of their exponents i, j, k. By the Pigeon-hole Principle, there will be two elements of M, say x and y, that are in the same class. As the sum of two integers of the same parity is even, this implies that $x \cdot y = 2^{2a}3^{2b}5^{2c}$ for some non-negative integers a, b, c, therefore, $xy = (2^a3^b5^c)^2$.

(b) The $n+1$ elements of A can be considered as elements of an n-dimensional vector space over the binary field. Let B be a linearly dependent subset of A, then the product of all elements of B is a perfect square since all prime factors must occur in that product an even number of times in that product.

(7) If we try to copy the exact method of the previous problem, we may run into difficulties. Indeed, the elements of L can have nine different prime divisors, $2, 3, 5, 7, 11, 13, 17, 19, 23$. If we classify them according to the remainder of the exponents of these prime divisors modulo four, we get a classification into $4^9 > 2003$ classes. So it seems that it is not even sure that there will be a class containing two elements of L, let alone four.

The reason this attempt did not work is that it tried to prove too much. For the product of four integers to be a fourth power, it is not necessary that the exponents of each prime divisor have the same remainder modulo four in each of the four integers. For example, 1,1,2,8 do not have that property, but their product is $16 = 2^4$.

A more gradual approach is more successful. Let us classify the elements of L again just by the parity of the exponents of the nine possible prime divisors in them. This classification creates just $2^9 = 512$ classes. Now pick two elements of L that are in the same class, and remove them from L. Put their product into a new set L'. This procedure clearly decreased the size of L by 2. Then repeat this same procedure, that is, pick two elements of L that are in the same class, remove them, and put their product into L'. Note that all elements of L' will be squares as they will contain all their prime divisors with even exponents. Do this until you can, that is, until there are no two elements of L in the same class. Stop when that happens. Then L has at most 511 elements left, so we have removed at least 1492 elements from L. Therefore L' has at least 746 elements, all of which are squares of integers.

Now classify the elements of L' according to the remainders of the exponents of their prime divisors modulo four. As the elements of L' are all squares, all these exponents are even numbers, so their remainders modulo four are either 0 or 2. So again, this classification creates only 512 classes, and therefore, there will be two elements of L' in the same class, say u and v. Then uv is the fourth power of an integer, and since both u and v are products of two integers in L, our claim is proved.

(8) **First Solution**: Let $a_1 \leq a_2 \leq \cdots \leq a_{100}$ denote our one hundred numbers. We will show 99 non-negative sums. We have to distinguish two cases, according to the sign of $a_{50} + a_{99}$. Assume first that $a_{50} + a_{99} \geq 0$. Then we have

$$0 \leq a_{50} + a_{99} \leq a_{51} + a_{99} \leq a_{52} + a_{99} \leq \cdots \leq a_{100} + a_{99},$$

providing 51 non-negative sums. On the other hand, for any i so that $50 \leq i \leq 100$, we now have

$$0 \leq a_i + a_{99} \leq a_i + a_{100},$$

providing the new non-negative sums $a_{50} + a_{100}, a_{51} + a_{100}, \cdots, a_{98} + a_{100}$, which is 49 new sums, so we have found 100 non-negative sums.

Now assume that $a_{50} + a_{99} < 0$. Then necessarily

$$a_1 + a_2 + \cdots + a_{49} + a_{51} + \cdots + a_{98} + a_{100} > 0. \tag{1.1}$$

In this case we claim that all sums $a_i + a_{100}$ are non-negative. To see this, it suffices to show that the smallest of them, $a_1 + a_{100}$ is non-negative. And that is true as

$$0 > a_{50} + a_{99} \geq a_{49} + a_{98} \geq a_{48} + a_{97} \geq \cdots \geq a_2 + a_{51},$$

and therefore the left-hand side of (1.1) can be decomposed as the sum of $a_1 + a_{100}$, and 48 negative numbers. So $a_1 + a_{100}$ is positive, and the proof follows.

Second Solution: It is well known from everyday life that one can organize a round robin tournament for $2n$ teams in $2n-1$ rounds, so that each round consists of n games, and that each team plays a different team each round. A rigorous proof of this fact can be found in Chapter 2, Exercise 4. Now take such a round robin tournament, and replace the teams with the numbers $a_1, a_2, \cdots, a_{100}$. So the fifty games of each round are replaced by fifty pairs of type $a_i + a_j$. As each team plays in each round, the sum of the 100 numbers, or 50 pairs, in any given round is zero. Therefore, at least one pair must have a non-negative sum in any given row, otherwise that row would have a negative sum.

This result is the best possible one: if $a_{100} = 99$, and $a_i = -1$ if $1 \leq i \leq 99$, then there will be exactly 99 non-negative two-element sums.

(9) There is only a finite number of choices for the color of each point, so there is only a finite number F of choices to color the integer points of a 7×7 square. Now take a column built up from $F+1$ squares of size 7×7 that have the same x coordinates. (They are "above one another".) By the Pigeon-hole Principle, two of them must have the very same coloring. This means that if the first one has two points of the same color in the ith and jth positions, then so does the second, and a monochromatic rectangle is formed. The Pigeon-hole Principle ensures that such i and j always exist, and the proof follows. In fact, we also proved that there will always be a monochromatic rectangle whose shorter side contains at most 7 points with integer coordinates.

(10) Consider the remainders of each of the given integers modulo 1000, and the opposites of these remainders modulo 1000. Note that if an integer is not congruent to 0 or 500 modulo 1000, then its remainder and opposite remainder modulo 1000 are two different integers.

We distinguish two cases. First, if at least two of our integers are divisible by 1000, or if at least two of our integers have remainder 500 modulo 1000, then the difference and sum of these two integers are both divisible by 1000, and we are done.

If there is at most one among our integers that is divisible by 1000, and there is at most one among our integers that has remainder 500 modulo 1000, then we have at least 500 integers that do not fall into either category. Consider their remainders and opposite remainders modulo 1000, altogether 1000 numbers. They cannot be equal to 0 or 500, so there are only 998 possibilities for them. Therefore, the Pigeon-hole Principle implies that there must be two equal among them, and the proof follows.

(11) Denote $3n - a$ the largest chosen number (it could be that $a = 0$). Let us add a to all our chosen numbers; this clearly does not change their pairwise differences. So now $3n$ is the largest chosen number. Therefore, if any number from the interval $[n+1, 2n-1]$ is chosen, we are done. Otherwise, we had to choose a total of $n+1$ numbers from the intervals $[1, n]$ and $[2n, 3n-1]$. Consider the n pairs

$$(1, 2n); (2, 2n+1); \cdots ; (i, i+2n-1), \cdots ; (n, 3n-1).$$

As there are n such pairs, and we chose $n+1$ integers, there is one pair with two chosen elements. The difference of those two chosen elements is $2n-1$, and our claim is proved.

(12) Let the numbers of stones in the original four heaps be $a_1 \geq a_2 \geq a_3 \geq a_4$, and let the numbers of stones in the five new heaps be $b_1 \geq b_2 \geq b_3 \geq b_4 \geq b_5$. Then $a_1 + a_2 + a_3 + a_4 > b_1 + b_2 + b_3 + b_4$. Let k be the smallest index so that $a_1 + \cdots + a_k > b_1 + \cdots + b_k$. (It follows from the previous sentence that there is such an index.) This implies that $a_k > b_k$. Then the stones from the k largest old heaps could not all go to the k largest new heaps. (Indeed, there are too many of them.) In fact, note that $a_1 + \cdots + a_k > b_1 + \cdots + b_{k-1} + 1$. So at least two of these stones had to go to a heap with b_k stones or less, and we are done as $a_1 \geq \cdots \geq a_k > b_k \geq b_{k+1} \geq \cdots \geq b_5$.

(13) Assume the contrary, that is, that each positive integer appears on a finite number of pieces only. As we have an infinite number of pieces, this means that there is an infinite sequence of different positive integers $a_1 < a_2 < a_3 < \cdots$ so that each a_i appears on at least one piece of paper. Then the subsequence $a_1, a_{10^7+1}, a_{2\cdot 10^7+1}, a_{3\cdot 10^7+1}, \cdots$, is an infinite set in which any two elements differ by at least ten million. As all elements of this subsequence appear on some pieces of paper, we have reached a contradiction.

(14)(a) Let a_i denote the number of goals the team scored in the ith game. Consider the 30 numbers $b_i = a_1 + a_2 + \cdots + a_i$ for all i satisfying $1 \leq i \leq 30$, and the 30 numbers $b_i + 6$ for $1 \leq i \leq 30$. This is a collection of 60 numbers, each of which is a positive integer, and none of which is larger than $53 + 6 = 59$. So by the Pigeon-hole

Principle, two of these numbers are equal. One of them must be b_i and the other must be $b_j + 6$ for some $j < i$, since all the b_i are different. Then the team scored exactly six goals total in games $j+1, j+2, \cdots, i$.

(b) A counterexample is given by the sequence 2, 1, 2, 2, 3, 2, repeated four more times, for the numbers $a_1, a_2, \cdots$ as defined in the solution of part (a). Another counterexample is given by the sequence 1, 1, 1, 1, 1, 7 repeated four more times.

(c) Let the numbers a_i and b_i be defined as in the solution of part (a). Let us assume that our claim does not hold. Consider the sequence of the ten integers $F = \{1, 7, 13, \cdots, 55\}$. Let B denote the sequence $b_1, b_2, \cdots, b_{30}$.
At most five elements of F can be elements of B since no two *consecutive* elements of F can be part B. Similarly, at most five elements of the sequence $2, 8, \cdots, 56$ can be part of B. The same goes for the sequence $3, 9, \cdots, 57$, the seqeunce $4, 10, \cdots, 58$, and the sequence $5, 11, \cdots, 59$. Therefore, since B consists of 30 positive integers, the largest of which is 59, the sequence of the remaining positive integers not larger than 60, that is, the sequence $6, 12, \cdots, 54$ must contain at least five elements of B. If our claim does not hold, then $6 \notin B$, so the eight-element sequence $S = \{12, 18, \cdots, 54\}$ contains at least five numbers b_i. That means that the there are two *consecutive* elements of S that are part of B, which is a contradiction.

(15) Let $a_1, a_2, \cdots, a_k$ be the initial terms of our k progressions, and let $d_1, d_2, \cdots, d_k$ be their differences. The number $d_1 d_2 \cdots d_k$ is an element of one of these progressions, say, the ith one. Therefore, there is a positive integer m so that

$$d_1 d_2 \cdots d_k = a_i + m d_i,$$

$$d_1 d_2 \cdots d_k - m d_i = a_i.$$

So a_i is divisible by d_i. This problem had nothing to do with the Pigeon-hole Principle. We included it to warn the reader that not all that glitters is gold. Just because we have to prove that one of many objects has a given property, we cannot necessarily use the Pigeon-hole Principle.

(16) Let us cut the square into two congruent triangles using one of its diagonals; then cut each of these triangles into 16 congruent triangles using the method of midlines. This yields 32 congruent triangles of

area $1/32$ each. As we have $65 > 2 \cdot 32$ points, by the generalized version of the Pigeon-hole principle, at least one of these 32 triangles must contain at least three of our points.

(17) Let us call a set of entries of A a *diagonal* if they form the intersection of A with a line of slope 1. There are $2n-1$ such diagonals in A, and each entry belongs to exactly one of them. (Indeed, the diagonals are obtained as intersections of the lines $y = x + a$ with A, where $a \in [-(n-1), n-1]$ is an integer.) As there are $2n-1$ diagonals and at least $2n$ entries equal to 1, at least one diagonal contains at least two entries 1, and the statement is proved.

(18) Yes. Here is an example. Let us assume that in each county, 100 people vote. Candidate A gets 40 votes in each county. Candidate B gets 50 votes in three counties, and 10 votes in the remaining four counties, while candidate C gets 10 votes in the first three counties, and 50 votes in the remaining four counties. This means that statewide, candidate A gets 280 votes, candidate B gets 190 votes, and candidate C gets 230 votes.