

第六章 数据的收集与整理

1 数据的收集

(1) **收集数据的方法**：调查、试验、查阅资料等，调查又分为实地调查、问卷调查和访问调查等。

(2) **统计活动的过程**：明确调查目的和问题；确定调查对象；选择调查方法；展开调查；收集并整理数据；分析数据，得出结论。

2 普查与抽样调查

(1) 为了特定目的对全部考察对象进行的全面调查，叫做普查。

(2) 其中被考察对象的全体叫做总体。

(3) 组成总体的每一个被考察对象称为个体。

(4) 从总体中抽取部分个体进行调查，这种调查称为抽样调查。

(5) 其中从总体抽取的一部分个体叫做总体的一个样本。

(6) 抽样调查样本应具有广泛性和代表性。

3 扇形统计图

(1) **扇形统计图**：利用圆与扇形来表示总体与部分的关系，扇形的大小反映部分占总体的百分比的大小，这样的统计图叫做扇形统计图。（各个扇形所占的百分比之和为1）

(2) 扇形统计图的意义 $\text{圆心角度数} = 360^\circ \times \text{该项所占的百分比}$ 。（各个部分的圆心角度数之和为 360° ）

(3) 扇形统计图的优缺点：

优点是易于显示每组数据相对于总数的大小；

缺点是在不知道总体数量的条件下，无法知道每组数据的具体数量。

(4) 绘制扇形统计图的步骤：求出全体（即总量）；

计算各部分占总体的百分比；

求出各部分所对扇形圆心角的度数；

画出扇形统计图。

4 频数直方图

(1) 频数直方图是一种特殊的条形统计图，它将统计对象的数据进行了分组画在横轴上，纵轴表示各组数据的频数。

(2) 频数直方图与条形统计图

联系：频数直方图本质上是一种条形统计图，也可以认为是一种以频数为纵向指标的条形统计图。

区别：如果样本中数据多、差距大，频数直方图能更清晰、更直观地反映数据的分布情况。

(3) 频数直方图的优缺点

频数直方图的优点能够显示每组中的具体数据，易于比较数据之间的差别；缺点是无法显示每组数据占总体的百分比的多少。

(4) 绘制频数直方图的步骤：求出最大值与最小值的差；

确定组数和组距并进行分组；

统计每组中数据的频数；

画出频数直方图

5 各种统计图的特点

条形统计图：能清楚地表示出每个项目的具体数目。

折线统计图：能清楚地反映事物的变化情况。

扇形统计图：能清楚地表示出各部分在总体中所占的百分比。

6 全面调查是考查全体对象的调查，不能有遗漏，抽样调查是对考察对象选取一部分进行的调查。

采取哪种方式，一定要依据具体的问题，使得调查有可靠的结果，又不能造成太大的损失或付出较大的代价。

考点 1 样本与总体

例 1. 下列调查中，最适合采用抽样调查的是（ ）

- A. 乘坐飞机时对旅客行李的检查
- B. 了解小明一家三口人对端午节来历的了解程度
- C. 了解我校初 2016 级 1 班全体同学的视力情况
- D. 了解某批灯泡的使用寿命

【答案】D

【解析】A、乘坐飞机时对旅客行李的检查适合采用全面调查方式；

B、了解小明一家三口人对端午节来历的了解程度适合采用全面调查方式；

C、了解我校初 2016 级 1 班全体同学的视力情况的调查适合采用全面调查方式；

D、了解某批灯泡的使用寿命适合采用用抽样调方式；

故选：D.

考点 2 调查方式的选择

例 2. 下列调查中，适合采用抽样调查的是（ ）

- A. 调查 18 中非毕业年级学生对“社团课”的满意程度
- B. 调查本班同学的身高
- C. 为保证某种新研发的战斗机成功试飞，对其零部件进行检查
- D. 对乘坐高铁的乘客进行安检

【答案】A

【解析】A、调查 18 中非毕业年级学生对“社团课”的满意程度适合采用抽样调查；

B、调查本班同学的身高适合采用全面调查；

C、为保证某种新研发的战斗机成功试飞，对其零部件进行检查调查本班同学的身高适合采用全面调查；

D、对乘坐高铁的乘客进行安检调查本班同学的身高适合采用全面调查；

故选：A.

考点 3 数据的收集

例 3.为了了解南京市八年级学生的身高情况，请你运用所学的统计知识，将解决上述问题要经历的几个重要步骤进行排序. ①用样本估计总体；②整理数据；③设计调查问卷；④分析数据；⑤收集数据. 则正确的排序为（ ）

- A. ⑤③②④① B. ③⑤②①④ C. ③⑤②④① D. ③⑤④②①

【答案】解：将解决上述问题要经历的几个重要步骤进行排序，③设计调查问卷；⑤收集数据；②整理数据；④分析数据；①用样本估计总体；

∴正确的排序为③⑤②④①，

故选：C.

考点 4 频数直方图

例 4.某校有 500 名学生参加体育测试，其成绩在 25 - 30 分之间的有 300 人，则在 25 - 30 分之间的频率是（ ）

- A. 0.6 B. 0.5 C. 0.3 D. 0.1

【答案】A

【解析】根据题意，得：在 25 - 30 分之间的频率是 $300 \div 500 = 0.6$.

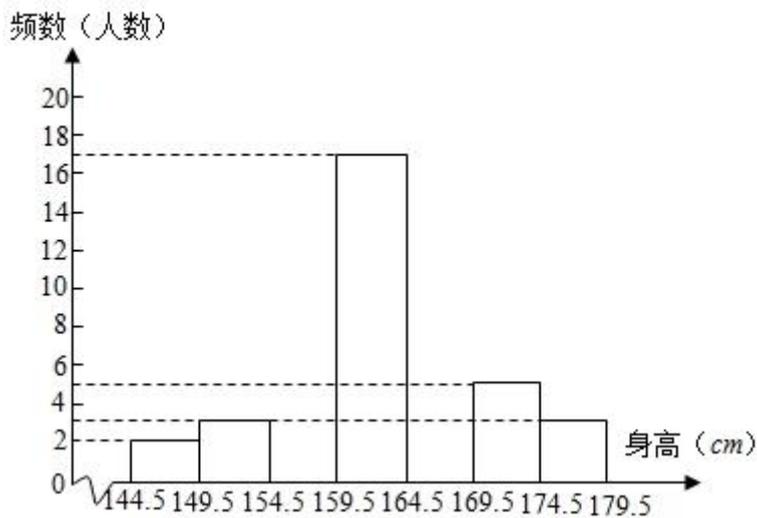
故选：A.

例 5.为了解同学们的身体发育情况，学校体卫办公室对七年级全体学生进行了身高测量（精确到 $1cm$ ），并从中抽取了部分数据进行统计，请根据尚未完成的频数分布表和频数分布直方图解答下列问题：

频数分布表

分组	频数	百分比
144.5~149.5	2	4%
149.5~154.5	3	6%
154.5~159.5	a	16%
159.5~164.5	17	34%
164.5~169.5	b	$n\%$
169.5~174.5	5	10%
174.5~179.5	3	6%

频数分布直方图



(1) 填空： $a=8$ 、 $b=12$ 、 $n=24$ 。

(2) 补全频数分布直方图；

(3) 学校准备从七年级学生中选拔护旗手，要求身高不低于 170cm 且低于 175cm ，如果七年级有学生 350 人，护旗手的候选人大概有多少？

【答案】解：(1) 本次抽取的学生有： $2 \div 4\% = 50$ (人)，

$$a = 50 \times 16\% = 8, \quad b = 50 - 2 - 3 - 8 - 17 - 5 - 3 = 12, \quad n\% = \frac{12}{50} \times 100\% = 24\%,$$

故答案为：8，12，24；

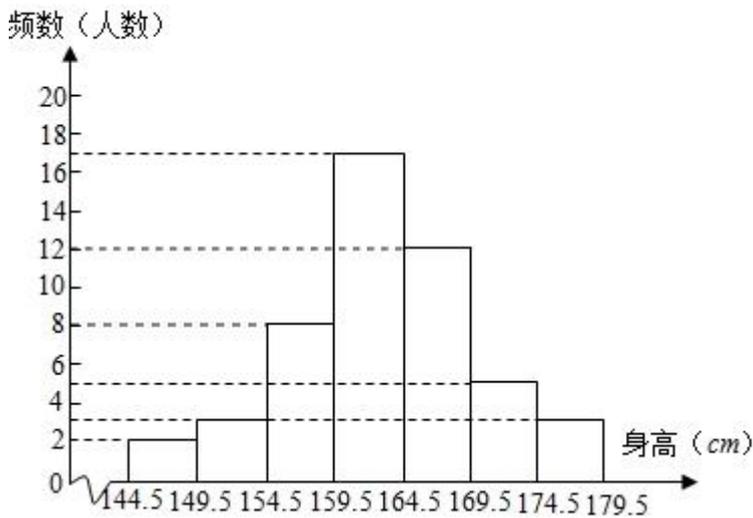
(2) 由 (1) 知 $a=8$ ， $b=12$ ，

补全的频数分布直方图如右图所示；

(3) $350 \times 10\% = 35$ (人)，

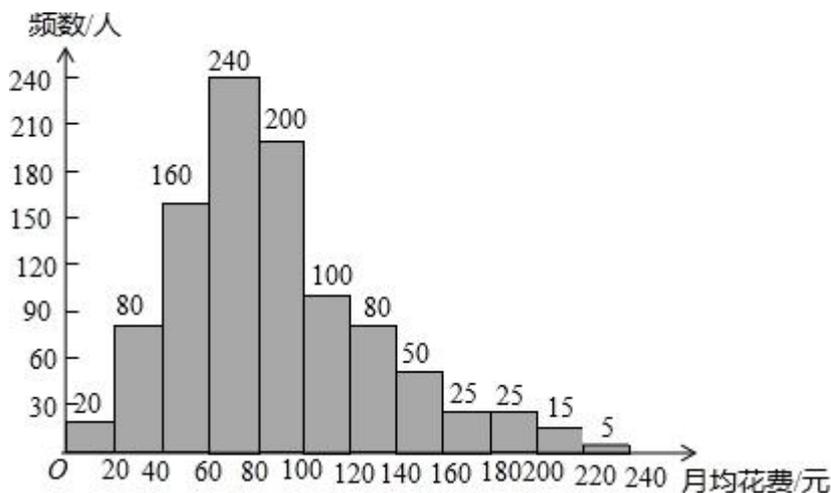
答：护旗手的候选人大概有 35 人。

频数分布直方图



考点 5 统计图的应用

例 6. 为了了解 2019 年北京市乘坐地铁的每个人的月均花费情况，相关部门随机调查了 1000 人乘坐地铁的月均花费（单位：元），绘制了如下频数分布直方图，根据图中信息，下面三个推断中，合理的是（ ）



- ① 小明乘坐地铁的月均花费是 75 元，那么在所调查的 1000 人中一定有超过一半的人月均花费超过小明；
 ② 估计平均每人乘坐地铁的月均花费的不低于 60 元；
 ③ 如果规定消费达到一定数额可以享受折扣优惠，并且享受折扣优惠的人数控制在 20% 左右，那么乘坐地铁的月均花费达到 120 元的人可享受折扣。

A. ①② B. ①③ C. ②③ D. ①②③

【答案】C

【解析】① $\because 200+100+80+50+25+25+15+5=500$ ，而 75~80 元的人数不能确定，

\therefore 在所调查的 1000 人中一定有一半或超过一半的人月均花费超过小明，此结论错误；

② 根据图中信息，可得大多数人乘坐地铁的月均花费在 60~120 之间，

估计平均每人乘坐地铁的月均花费的范围是 60~120，

所以估计平均每人乘坐地铁的月均花费的不低于 60 元，此结论正确；

③ $\because 1000 \times 20\% = 200$ ，而 $80 + 50 + 25 + 25 + 15 + 5 = 200$ ，

\therefore 乘坐地铁的月均花费达到 120 元的人可以享受折扣。此结论正确；

综上，正确的结论为②③，

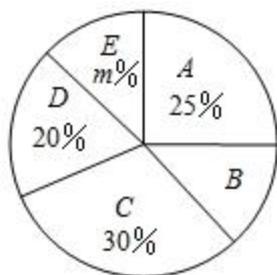
故选：C。

例 7.共享经济与我们的生活息息相关，其中，共享单车的使用给我们的生活带来了许多便利，但在使用过程中出现一些不文明现象。某市记者为了解“使用共享单车时的不文明行为”，随机抽查了该市部分市民，并对调查结果进行了整理，绘制了如下两幅尚不完整的统计图表（每个市民仅持有一种观点）。

调查结果分组统计表

组别	观点	频数（人数）
A	损坏零件	50
B	破译密码	20
C	乱停乱放	a
D	私锁共享单车， 归为己用	b
E	其他	30

调查结果扇形图



请根据以上信息，答案下列问题：

- （1）填空： $a = \underline{60}$ ； $b = \underline{40}$ ； $m = \underline{15}$ ；
- （2）求扇形图中 B 组所在扇形的圆心角度数；
- （3）若该市约有 100 万人，请你估计其中持有 D 组观点的市民人数。

【答案】解：（1） $50 \div 25\% = 200$ 人， $c = 200 \times 30\% = 60$ 人， $b = 200 \times 20\% = 40$ 人， $30 \div 200 = 15\%$ ，

故答案为：60；40；15；

$$(2) 360^\circ \times (1 - 25\% - 30\% - 20\% - 15\%) = 36^\circ;$$

答：扇形图中 B 组所在扇形的圆心角度数为 36° 。

$$(3) 100 \times 20\% = 20 \text{ (万人)}$$

答：持有 D 组观点的市民人数大约为 20 万人，

考点 6 统计图的选择

例 8. 我国五座名山的海拔高度如下表：

山名	泰山	华山	黄山	庐山	峨眉山
海拔/米	1 545	2 155	1 864	1 474	3 099

若想根据表中数据绘制统计图，以便更清楚地比较五座山的高度，最合适的是()

- A. 条形统计图 B. 折线统计图 C. 扇形统计图 D. 以上都可以

【答案】A

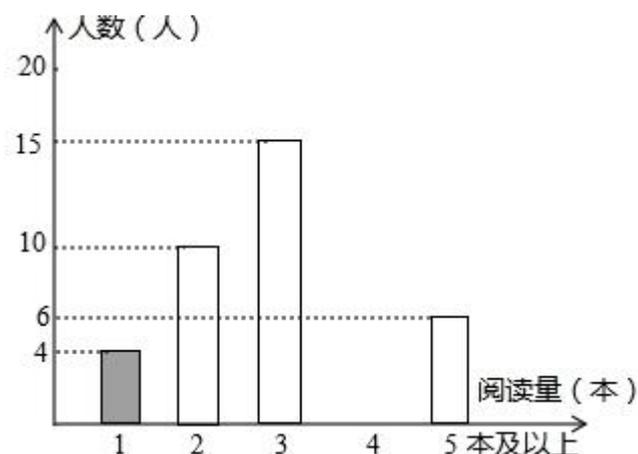
【解析】根据统计图的特点选择合适的统计图和从统计图中获取信息。本题选 A。

例 9. 某学校为了解学生的课外阅读情况，随机抽查部分学生，并对其寒假期间的课外阅读量进行统计分析，绘制成如图所示但不完整的统计图。已知抽查的学生在寒假期间阅读量为 2 本的人数占抽查总人数的 20%，根据所给出信息，答案下列问题：

(1) 求被抽查学生人数；

(2) 将条形统计图补充完整；

(3) 若规定：假期阅读 3 本及 3 本以上课外书者为完成假期作业，据此估计该校 1800 名学生中，完成假期作业的有多少人？



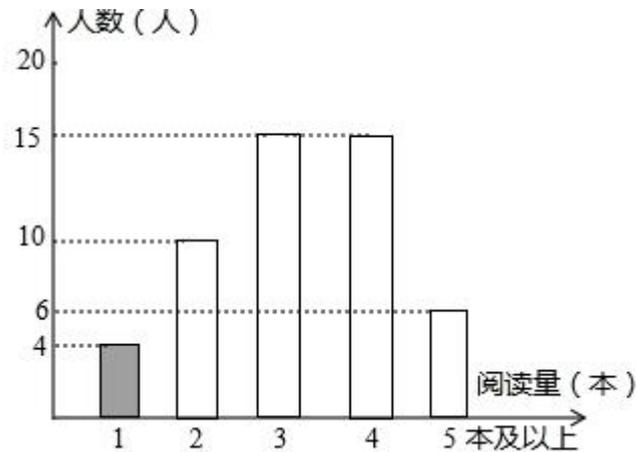
【答案】解：(1) $10 \div 20\% = 50$ 人，

答：被抽查人数为 50 人。

(2) $50 - 4 - 10 - 15 - 6 = 15$ 人，补全条形统计图如图所示：

$$(3) 1800 \times \frac{15+15+6}{50} = 1296 \text{ 人,}$$

答：该校 1800 名学生中，完成假期作业的有 1296 人。



一、选择题

- 一次数学测试后，某班 60 名学生的成绩被分为 5 组，第一至第四组的频数分别为 8、10、16、14，则第五组的频率是 ()
A. 0.1 B. 0.2 C. 0.3 D. 0.4
- 为了考察某初中 4500 名毕业生的数学成绩，从中抽出 25 份试卷，每份 30 张。在这个问题中，样本容量是 ()
A. 4500 B. 25 C. 30 D. 750
- 下列调查中适合采用全面调查（普查）的方式的是 ()
A. 了解“中国达人秀第六季”节目的收视率
B. 调查重庆实验外语学校某班学生喜欢上数学课的情况
C. 调查我国民众对“香港近期暴力”行为的看法
D. 调查我国目前“垃圾分类”推广情况
- 要了解一所中学七年级学生的每周课外阅读情况，以下方法中比较合理的是 ()
A. 调查七年级全体学生的每周课外阅读情况
B. 调查其中一个班的学生每周课外阅读情况
C. 调查七年级全体男生的每周课外阅读情况
D. 调查七年级每班学号为 3 的倍数的学生的每周课外阅读情况
- 某区为了解 5600 名初中生的身高情况，抽取了 300 名学生进行身高测量在这个问题中，样本是 ()
A. 300 B. 300 名学生

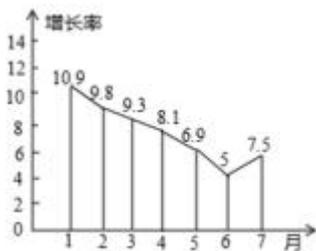
C. 300 名学生的身高情况 D. 5600 名学生的身高情况

6. 下列调查中适合全面调查的是 ()

- A. 考察某种灯泡的使用寿命 B. 了解一锅汤烧得是否鲜美可口
C. 某公司对退休员工进行健康检查 D. 了解我国农民的年人均收入情况

7. 某种海产品在七个月之内的价格增长变化情况如图所示, 则下列说法中正确的个数是 ()

- ①2-6 月海产品价格增长率逐月减少
②7 月份海产品价格开始回升
③这 7 个月中, 海产品价格不断上涨
④这 7 个月中, 海产品价格有上涨有下跌



A. 1 B. 2 C. 3 D. 4

8. 下列说法错误的是 ()

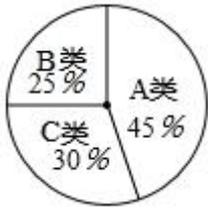
- A. “打开电视, 正在播放新闻节目”是随机事件.
B. 为了解某种节能灯的使用寿命, 选择全面调查.
C. 频数折线图能清楚的反映事物的变化情况, 显示数据变化趋势.
D. 今年我市有 5.6 万名初中毕业生参加升学考试, 为了了解这 5.6 万名考生的数学成绩, 从中抽取 200 名考生的数学成绩进行统计, 在这个问题中样本是这 200 名考生的数学成绩.

二、填空题

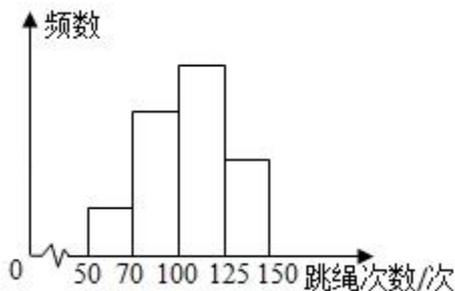
9. 已知某组数据的频数为 32, 频率为 0.8, 则这组数据个数为 40.

10. 已知数据: 25, 21, 23, 25, 29, 27, 28, 25, 27, 30, 22, 26, 25, 24, 26, 28, 26, 25, 24, 27. 若取组距为 2, 则列频数表时, 应将数据分成的组数是 5.

11. 某图书馆有 A、B、C 三类图书, 它们的数量用如图所示的扇形统计图表示, 若 B 类图书有 100 万册, 则 C 类图书有 120 万册.



12. 为了解小学生的体能情况，抽取了某小学同年级 50 名学生进行 1 分钟跳绳测试，将所得数据整理后，画出如图所示的频数分布直方图（各组只含最小值，不含最大值），已知图中从左到右各组的频率分别 a , 0.3, 0.4, 0.2，设跳绳次不低于 100 次的学生有 b 人，则 a, b 的值分别是 0.1, 30。

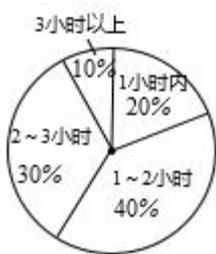


13. 若某校有学生 4000 名，从中随机抽取了 40 名学生，调查他们每天做作业的时间，结果如下表：

每天做作业时间 t (时)	$0 < t \leq 1$	$1 < t \leq 2$	$2 < t \leq 3$	$3 < t \leq 4$	$t > 4$
人数	3	16	9	8	4

则全校学生每天做作业超过 3 小时的人数约有 1200 人。

14. 如图是某校学生家庭作业完成时间情况的统计图，若该校作业完成时间在 1 小时内的学生有 300 人，则该校作业完成时间在 2~3 小时的学生有 人。



15. 某高中自主招生考试只考数学和物理，数学与物理成绩按 6: 4 计入综合成绩。已知小明数学成绩为 95 分，综合成绩为 93 分，那么小明的物理成绩为 分。

16. 某中学为了了解学生数学课程的学习情况，在 3000 名学生中随机抽取 200 名，并统计这 200 名学生的某次数学考试成绩，得到了样本的频率分布直方图(如图)。根据频率分布直方图推测，这 3000 名学生在该次数学考试中成绩小于 60 分的学生数是 。

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/808120106027006036>