

专题 01 统计初步 难点分析

题型归纳

题型一：已知一组数据的统计量，求另一组数据的统计量

题型二：判断统计量的变化I

题型三：判断统计量的变化II；统计量之间的关系

题型四：算错、遮盖等问题

题型五：统计图表的综合应用

题型六：统计量的综合应用

题型七：解答题

题型专练

题型一：已知一组数据的统计量，求另一组数据的统计量

1. 已知 $2, 3, 5, m, n$ 五个数据的方差是 2，那么 $4, 5, 7, m+2, n+2$ 五个数据的方差是_____.

2. 已知 $x_1, x_2, x_3, \dots, x_{20}$ 的平均数是 5，方差是 2，则 $3x_1+2, 3x_2+2, 3x_3+2, \dots, 3x_{20}+2$ 的平均数是_____，方差是_____.

3. 一组数据 $x_1, x_2, \dots, x_n$ 的平均数为 5，方差为 16， n 是正整数，则另一组数据 $3x_1+2, 3x_2+2, \dots, 3x_n+2$ 的标准差是_____.

4. 已知数据 $x_1, x_2, \dots, x_n$ 的平均数为 m ，方差为 s^2 ，则数据 $kx_1+b, kx_2+b, \dots, kx_n+b$ 的平均数为_____，方差为_____，标准差为_____.

题型二：判断统计量的变化I

5. 当一组非零数据每一个数字乘以 n 倍 ($n \neq 1$) 后，下列说法正确的个数是 ()

(1) 方差一定改变 (2) 平均数一定改变 (3) 中位数一定改变 (4) 众数一定改变

A. 1

B. 2

C. 3

D. 4

6. 已知一组数据 x_1, x_2, x_3, x_4 ，如果这组数据中的每一个数都减去常数 $a(a \neq 0)$ ，得到新的一组数据，

那么下列描述这组新数据的信息中正确的是 ()

- A. 平均数改变，方差不变； B. 平均数改变，方差改变；
C. 平均数不变，方差不变； D. 平均数不变，方差改变。

7. 在一组数据 4, 6, 2, 4 中，如果再添加一个数据 4，那么发生变化的统计量是 ()

- A. 平均数 B. 中位数 C. 众数 D. 方差

8. 将一组数据中的每一个数据都加上 3，那么所得的新数据组与原数据组相比，没有改变大小的统计量是 ()

- A. 平均数 B. 中位数 C. 众数 D. 方差

9. 已知两组数据： x_1, x_2, x_3 和 x_1+2, x_2+2, x_3+2 ，下列说法正确的是 ()

- A. 平均数相等，方差不相等 B. 中位数相等，方差不相等
C. 平均数不相等，方差相等 D. 中位数不相等，众数相等

题型三：判断统计量的变化II；统计量之间的关系

10. 某校在“阅读之星”的评选活动中，5 位评委给小王同学的综合表现打分，分别是：9.2、9.3、8.8、9.3、

9.1. 如果每位评委的打分都提高 0.1，那么比较前后两组数据，统计量一定不会发生改变的是 () .

- A. 中位数 B. 众数 C. 方差 D. 平均数

11. 为了解某校九年级学生中长跑的成绩情况，随机抽取 30 名学生的中长跑成绩 (满分 20 分) 绘制成表：

成绩 / 分	15	16	17	18	19	20
人数 / 人	x	y	6	8	5	4

关于中长跑成绩的统计量中，一定不随 x, y 的变化而变化的是 ()

- A. 众数，中位数 B. 中位数，方差 C. 平均数，方差 D. 平均数，众数

12. 某兴趣小组有 5 名成员，身高 (厘米) 分别为：161,165,169,163,167. 增加一名身高为 165 厘米的成员后，现兴趣小组成员的身高与原来相比，下列说法正确的是 ()

- A. 平均数不变，方差不变 B. 平均数不变，方差变小
C. 平均数不变，方差变大 D. 平均数变小，方差不变。

13. 设有 n 个数 $x_1, x_2, x_3, \dots, x_n$ ，其标准差为 S_1 。另有 n 个数 $y_1, y_2, y_3, \dots, y_n$ ，其标准差为 S_2 。其中 $y_k = 2x_k + 3 (k=1, 2, 3, \dots, n)$ ，则下列说法正确的是 ()

- A. $S_2 = 2S_1 + 3$ B. $S_2 = 2S_1$ C. $S_2 = \sqrt{2}S_1 + 3$ D. $S_2 = \sqrt{2}S_1$

题型四：算错、遮盖等问题

14. 已知某篮球队有 15 名队员，他们身高的平均数和中位数都是 185 厘米，后来发现在登记身高时，将一名队员的身高由 174 厘米误写成 184 厘米，再经过重新计算后，正确的身高平均数为 m 厘米，中位数为 n 厘米，那么下列结论中正确的是 ()

- A. $m < 185, n < 185$ B. $m < 185, n = 185$ C. $m > 185, n > 185$ D. $m > 185, n = 185$

15. 运动会 200 米赛跑，5 位运动员成绩如下表所示，其中有两个数据被遮盖，那么被遮盖的两个数据依次是 ()

运动员	A	B	C	D	E	平均成绩	标准差
时间 (秒)	32	34		36	33	33	

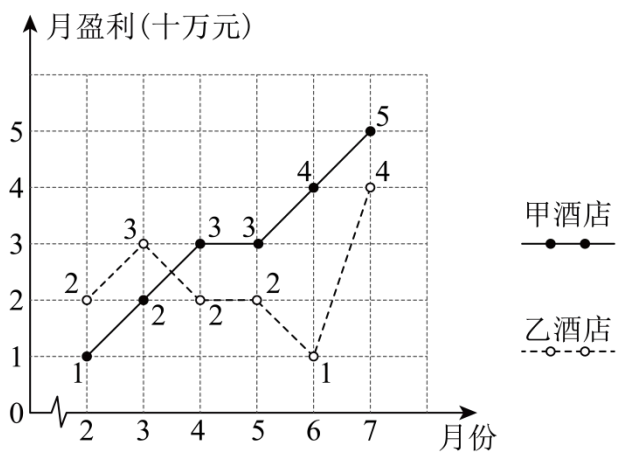
- A. 30.4 B. 30.2 C. 32.4 D. 32.2

16. 【平均数】李老师在黑板上写上若干个从 1 开始的连续自然数 1, 2, 3, …, 后来擦掉其中的一个，剩下的数的平均数是 10.9，则擦掉的这个自然数是_____。

题型五：统计图表的综合应用

17. 甲、乙两家酒店规模相当，去年 2~7 月的月盈利折线统计图如图所示。下列说法中，不正确的是 ()

甲、乙两家酒店 2~7 月的月盈利折线统计图



- A. 甲酒店每月盈利呈现不断增长的趋势
- B. 乙酒店经营状况有可能很快超过甲酒店
- C. 甲酒店月盈利的平均数大于乙酒店月盈利的平均数
- D. 甲酒店月盈利的方差小于乙酒店月盈利的方差

18. 某校七年级运动队为了备战校运动会需要购置一批运动鞋. 已知该队伍有 20 名同学, 统计表如下表.

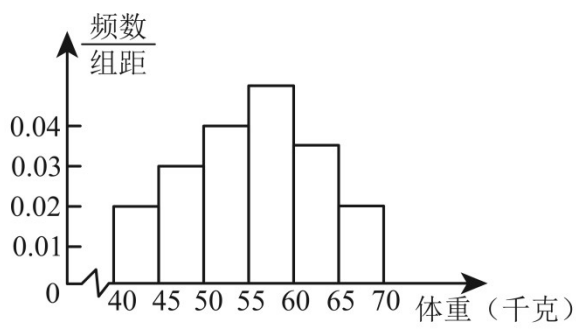
由于不小心弄脏了表格, 有两个数据看不到.

鞋 码	38	39	40	41	42
人 数	5			3	2

下列关于鞋码说法中正确的是 ()

- A. 中位数是 40, 众数是 39
- B. 中位数与众数一定相等
- C. 平均数 $\bar{x}$ 满足 $39 < \bar{x} < 40$
- D. 平均数可能为 39

19. 为了解全区 5000 名初中毕业生的体重情况, 随机抽测了 200 名学生的体重, 频率分布如图所示 (每小组数据可含最小值, 不含最大值), 其中从左至右前六个小长方形的高依次为 0.02、0.03、0.04、0.05、0.035、0.025, 由此可估计全区初中毕业生的体重在 50 到 55 千克的学生人数约为_____人.

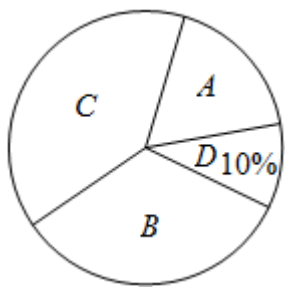


20．2021 年 7 月，中共中央办公厅、国务院办公厅印发《关于进一步减轻义务教育阶段学生作业负担和校外培训负担的意见》，明确要求初中生每天的书面作业时间不得超过 90 分钟．某校随机抽取部分学生进行问卷调查，并将调查结果制成如下不完整的统计图表．则下列说法不正确的是（ ）

作业时间频数分布

组别	作业时间（单位：分钟）	频数
<i>A</i>	$60 < t \leq 70$	8
<i>B</i>	$70 < t \leq 80$	17
<i>C</i>	$80 < t \leq 90$	<i>m</i>
<i>D</i>	$t > 90$	5

作业时间扇形统计图



- A．调查的样本容量是为 50
- B．频数分布表中 *m* 的值为 20
- C．若该校有 1000 名学生，作业完成的时间超过 90 分钟的约 100 人
- D．在扇形统计图中 *B* 组所对的圆心角是 144°

题型六：统计量的综合应用

21．有 11 个正整数，平均数是 10，中位数是 9，众数只有一个 8，问最大的正整数最大为（ ）

- A . 25 B . 30 C . 35 D . 40

22 . 若五个整数由小到大排列后，中位数为 4，唯一的众数为 2，则这组数据之和的最小值是_____ .

23 . 六名学生投篮，规定每人投 20 次，统计他们投中的次数，得到 6 个数据 . 若这六个数据的中位数是 8.5，唯一的众数是 9，则他们投中次数的平均数可能是 ()

- A . 5 B . 6 C . 10 D . 11

24 . 已知 5 个数据：8，8，x，10，10 . 如果这组数据的某个众数与平均数相等，那么这组数据的中位数是 _____ .

25 . 已知样本数据 $x_1, x_2, \dots, x_n$ 的平均数与方差 s^2 满足如下关系式：
$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n} = \frac{\sum_{i=1}^n (x_i^2) - n \cdot (\bar{x})^2}{n} .$$

若已知 15 个数 $x_1, x_2, \dots, x_{15}$ 的平均数为 6，方差为 9，现从原 15 个数中剔除 x_1, x_2, x_3, x_4, x_5 这 5 个数，且剔除的这 5 个数的平均数为 8，方差为 5，则剩下的 10 个数 $x_6, x_7, \dots, x_{15}$ 的方差为_____ .

(参考公式： $\sum_{i=1}^n x_i^2 = x_1^2 + x_2^2 + x_3^2 + \dots + x_n^2$)

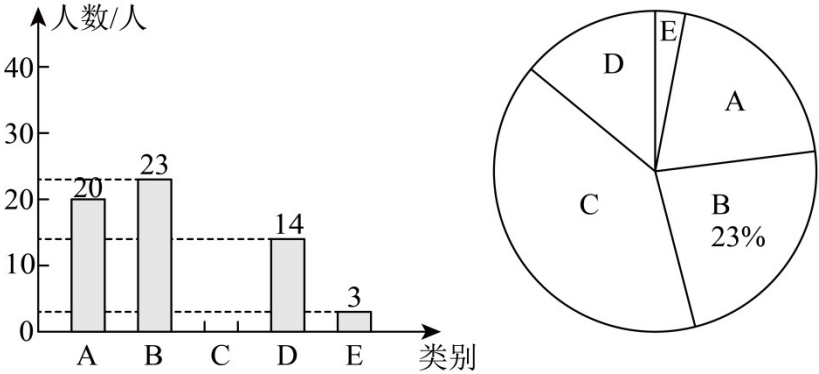
26 . 定义：一组数据 $x_1, x_2, \dots, x_n$ 的平均数为 $\bar{x}$ ，那么称这 n 个数据与平均数 $\bar{x}$ 的差的平方和叫做这 n 个数据的离差平方和，记作 $S^2 = (x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \dots + (x_n - \bar{x})^2$. 那么 100，101，99，98，102 的离差平方和是_____ .

27 . 五个互不相等的正偶数 x_1, x_2, x_3, x_4, x_5 的平均数和中位数都是 A ，且六个数 $x_1, x_2, x_3, x_4, x_5, m$ 的众数是 6，平均数还是 A ，则这五个互不相等的正偶数 x_1, x_2, x_3, x_4, x_5 的方差为_____ .

题型七：解答题

28 . 电视剧《山花烂漫时》以“七一勋章”、“时代楷模”的获得者张桂梅老师为原型，描绘了她在云南华坪女子高级中学辛勤耕耘的画面，展现了英模人物的非凡力量. 为了解初中部学生对“张桂梅老师事迹”的了解程度，我校随机抽取了部分初中学生进行调查，并将调查结果分为了五类： A . 非常了解； B . 比较了解； C . 了解； D . 不太了解； E . 不了解 . 根据调查结果，绘制出如图所示的两幅不完全统计图，请根据图中信息，解答下列问题：

抽取学生的调查结果条形统计图 抽取学生的调查结果扇形统计图

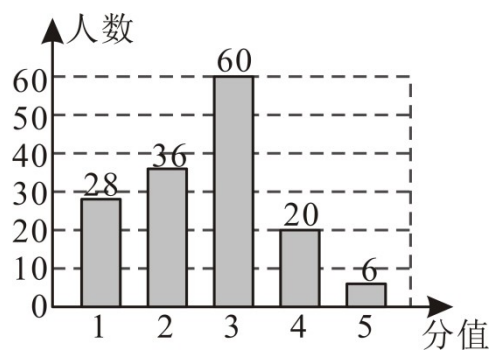


- (1)本次被抽查的学生共有_名；在扇形统计图中， A 类所对应的圆心角度数为_；
- (2)请补全条形统计图；
- (3)若我校共有 3000 名初中学生，估计我校初中学生对“张桂梅老师事迹”的了解程度为“不太了解及不了解”的人数。

29. 某校七年级要举行“阅读之星”评选活动，设计评选方案时考虑如下几个指标因素：①书籍的数量 A ；②书籍的总页数 B ；③书籍的类别 C ；④网络评分 D 。根据以上指标因素的重要程度赋以不同的系数，建立“阅读之星”的得分公式 $x = x_1A + x_2B + x_3C + x_4D$ ，其中 x_1 、 x_2 、 x_3 、 x_4 是各项指标因素的系数。假如小海同学一学期读了 4 本书，总页数 1350 页，涉及 3 个类别，4 本书的网络评分的平均分为 5.5 分，那么小海的得分计为 $x = 4x_1 + 1350x_2 + 3x_3 + 5.5x_4$ 。如果各项指标因素的系数一旦确定，那么他的“阅读之星”的得分也就确定。

评选小组通过向七年级学生和教师发放“阅读之星”评选指标因素重要程度的问卷调查，分别对上述四个指标因素打分，每个指标因素的分值范围为 0~10 分，四个指标因素分值的和必须为 10 分，指标因素的分值越高表示该指标因素越重要，然后将得到的每一个指标因素的所有分值取平均数作为该指标因素的系数。

评选小组对调查问卷的数据进行整理，得到“书籍的数量 A ”指标因素的得分情况统计图（如图）及各指标因素的系数表（如表 1）。



指标因素	系数
书籍的数量 A	m
书籍的总页数 B	2.4
书籍的类别 C	3.5
网络评分 D	n

表 1

- (1)指标因素“书籍的数量 A ”的系数 m 的值为_____；
- (2)确定各指标因素的系数后，“阅读之星”的得分公式为 $x =$ _____；
- (3)表 2 是该校七年级甲、乙两位同学“阅读之星”各项指标因素的数值．

	A	B	C	D	得分
甲	4	1500	3	7	
乙	3	1800	2	4	

表 2

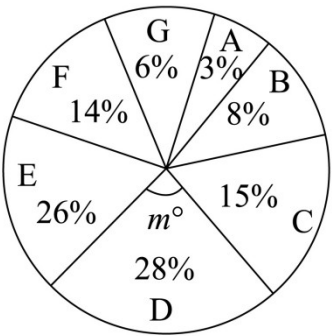
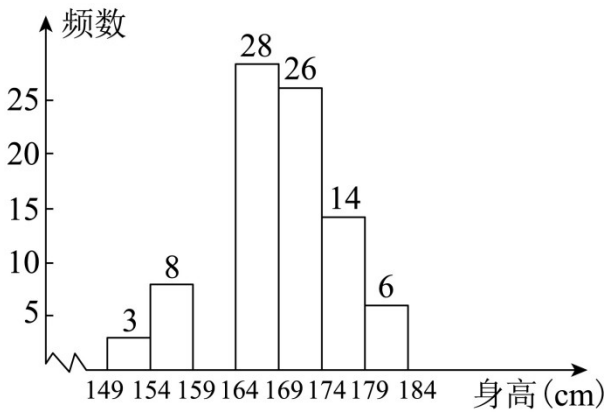
- ①请计算甲、乙两人“阅读之星”的得分．
- 甲得分为_____，乙得分为_____；
- ②根据两人的得分情况，请提出一条优化“阅读之星”评选方案的建议：_____．

30．某校喜迎国庆，七年级准备排练舞蹈《我和我的祖国》，为使舞蹈演员的身高比较整齐，需了解学生的身高分布情况，现从12个班级中任取两个班级的学生，收集他们的身高数据，并整理出如下的频数分布表、频数分布直方图和扇形统计图（部分信息未给出）

组别	身高范围（单位：厘米）	划记	频数	频率
A	$149 \leq x < 154$	下	3	0.03
B	$154 \leq x < 159$	正下	8	0.08
C	$159 \leq x < 164$		a	0.15
D	$164 \leq x < 169$	正正正正正下	28	0.28
E	$169 \leq x < 174$	正正正正正一	26	0.26
F	$174 \leq x < 179$	正正下	14	0.14
G	$179 \leq x < 184$	正一	6	0.06

频数分布直方图

扇形统计图

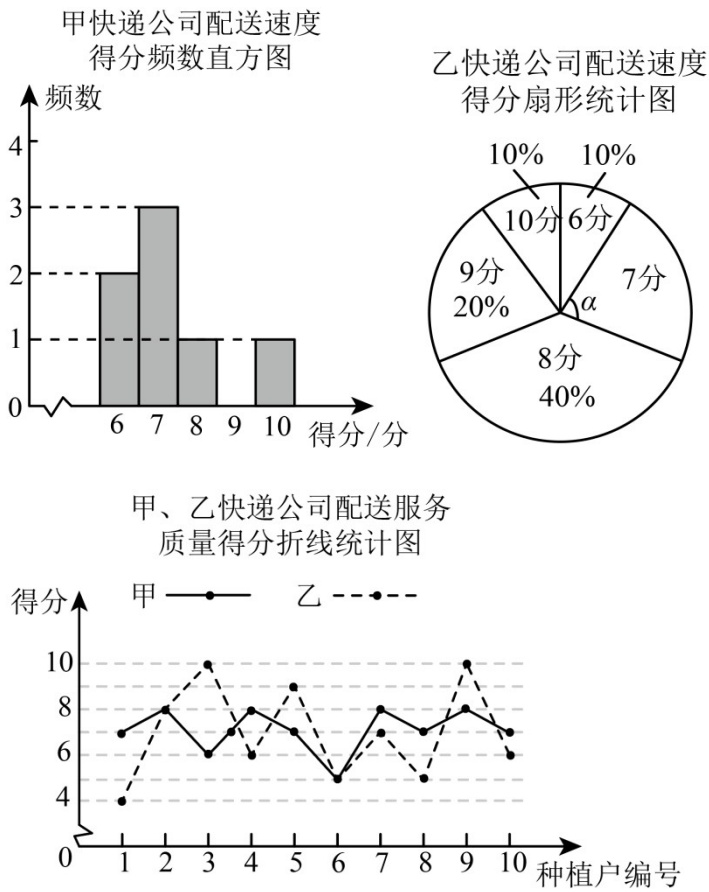


请根据统计图表提供的信息，解答下列问题：

- 本次抽样调查的样本容量是_____。
- $a =$ _____， $m =$ _____°。
- 请补全频数分布直方图
- 若七年级共有 600 名学生，请估计身高在 D 组的学生的人数。

31．随着快递行业在农村的深入发展，全国各地的特色农产品有了更广阔的销售空间．不同的快递公司在

配送、服务、收费和投递范围等方面各具优势，某农产品种植户经过前期调研，打算从甲、乙两家快递公司中选择一家合作。为此，该种植户收集了 10 家农产品种植户对两家公司的相关评价，并整理、描述、分析如下：



配送速度和服务质量得分统计表

项目统计量快递公司	配送速度得分		服务质量得分	
	平均数	中位数	平均数	方差
甲	7.8	m	7	$s_{\text{甲}}^2$
乙	8	8	7	$s_{\text{乙}}^2$

(1)补全频数分布直方图，扇形统计图中圆心角 α 的度数是_；

(2)表格中的 $m =$ _； $s_{\text{甲}}^2$ _ $s_{\text{乙}}^2$ (填“>”“=”或“<”)；

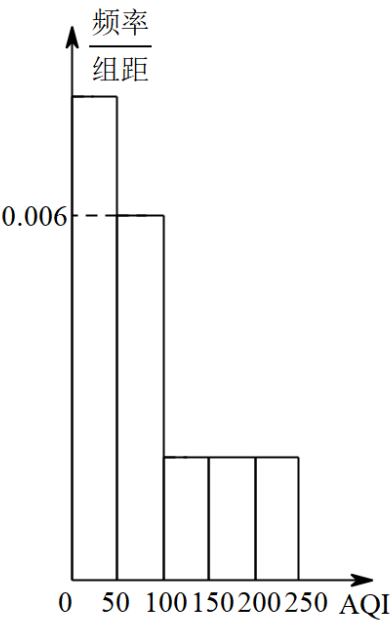
(3)如果 A, B, C 三家农产品种植户分别从甲、乙两个快递公司中任选一个公司合作，求三家种植户选择同一快递公司的概率。

32. 空气质量指数 (*Air Quality Index* , 缩写 *AQI*) 是定量描述空气状况的非线性无量纲指数。其数值越大、级别和类别越高，说明空气污染状况越严重，对人体的健康危害也就越大，适用于表示某地区的短期空气质量状况和变化趋势。(空气污染指数为 $0\sim 50$ 是优；空气污染指数为 $50\sim 100$ 是良好；空气污染指数为 $100\sim 150$ 是轻度污染；空气污染指数为 $150\sim 200$ 是中度污染；空气污染指数为 $200\sim 250$ 是重度污染。)

如图表示的是某地区 2022 年 11 月份 30 天日均 *AQI* 指数的频率分布直方图。

空气质量指数 (<i>AQI</i>)	0~50	50~100	100~150	150~200	200~250
天数	a	b	3	3	3
频率	c	d	0.1	0.1	0.1

(注：每组数据可含最高值，不含最低值)



(1)请你根据上述频率分布直方图及表格完成下面的填空：

这个地区 11 月份空气为轻度污染的天数是_____天。 $a =$ _____； $b =$ _____； $c =$ _____； $d =$ _____。

(2)为了进一步改善生活环境和空气质量，提高人民的生活质量，当地政府计划从 2023 年开始增加绿化面积。已知 2022 年底该地区的绿化面积为 20 万亩，如果到 2024 年底，该地区的绿化面积比为 2022 年的绿

化面积增加了 50%，假设这两年绿化面积的年增长率相同，求这两年中绿化面积每年的增长率（精确到 0.01）（参考数据： $\sqrt{2} \approx 1.414$ ， $\sqrt{3} \approx 1.732$ ， $\sqrt{5} \approx 2.236$ ， $\sqrt{6} \approx 2.449$ ）

33. 某社区为了加强居民对新型冠状病毒肺炎防护知识的了解，鼓励社区居民在线参与作答《2020 年新型冠状病毒肺炎的防护全国统一考试（全国卷）》试卷（满分 100 分），社区管理员随机从该社区抽取 40 名居民的答卷，并对他们的成绩（单位：分）进行整理、分析，过程如下：

收集数据

85 65 95 100 90 95 85 65 75 85 100 90 70 90 100 80 80 100 95 75 80 100 80 95 65 100 90 95 85 80 100 75 60 90 70 80 95 75 100 90

整理数据（每组数据可含最低值，不含最高值）

分组（分）	频数	频率
60 ~ 70	4	0.1
70 ~ 80	a	b
80 ~ 90	10	0.25
90 ~ 100	c	d
100 ~ 110	8	0.2

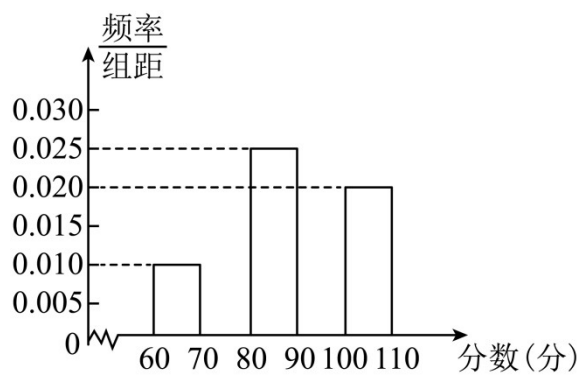
分析数据

（1）填空： $a = \underline{\hspace{1cm}}$ ， $b = \underline{\hspace{1cm}}$ ， $c = \underline{\hspace{1cm}}$ ， $d = \underline{\hspace{1cm}}$ ；

（2）补全频率分布直方图；

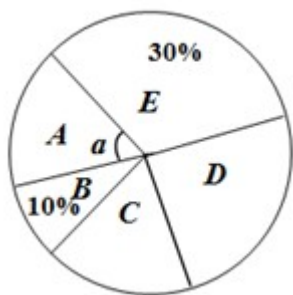
（3）由此估计该社区居民在线答卷成绩在 $\underline{\hspace{1cm}}$ （分）范围内的人数最多；

（4）如果该社区共有 800 人参与答卷，那么可估计该社区成绩在 90 分及以上约为 $\underline{\hspace{1cm}}$ 人。

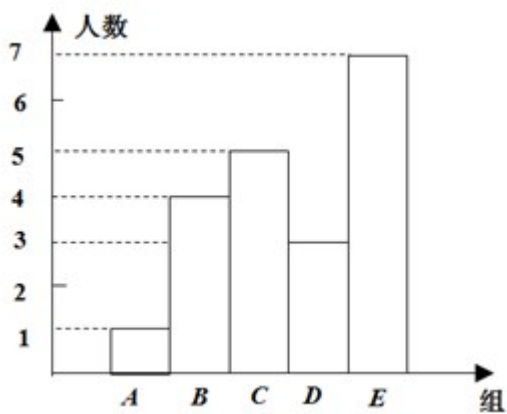


34. 9月16日，2020线上智博会举行西部（重庆）科学城新闻发布会。会上透露，西部（重庆）科学城是“科学家的家、创业者的城”，力争到2035年，全面建成具有全国影响力的科技创新中心核心区。为了解民众对科学城相关知识的知晓程度，某公司派甲、乙两人各随机调查20名群众，填写了对科学城相关知识的调查问卷（满分为10分），得分用 x 表示（ x 为整数），数据分组为A： $0 \leq x < 2$ ，B： $2 \leq x < 4$ ，C： $4 \leq x < 6$ ，D： $6 \leq x < 8$ ，E： $8 \leq x \leq 10$ ）。对问卷得分进行整理分析，给出了下面部分信息：

甲问卷得分的扇形统计图



乙问卷得分频数分布直方图（人数）



两组问卷得分的平均数，中位数，众数，满分率如下表：

	平均数	中位数（分）	众数	满分率
--	-----	--------	----	-----

	(分)		(分)	
甲公司	5.15	n	6	5%
乙公司	5.55	6	P	5%

甲公司 B 组占 10%，E 组占 30%，A 圆心角度数 α ；

甲公司分数在 C、D 组的数据为：6，4，4，6，6，7，6，5；乙公司 E 组所有数据之和为 58。

根据以上信息，解答下列问题：

- (1) 扇形统计图中 $\alpha =$ _度，信息表中的中位数 $n =$ _分，众数 $P =$ _分；
- (2) 通过以上数据分析，你认为 _公司问卷调查的成绩更好，理由是_；(写一条即可)
- (3) 若分数大于等于 6 即为合格，请估计问卷调查 1600 名群众中合格的人数是多少？