

牌号 6 · 单位 8 · 评估指南

数据集和分布

家长或教师指南——观察要点

8 单元旨在了解分布——中心、分布和形状。观察学生是否可以针对具体情况选择正确的衡量标准（平均值与中位数）并阐明原因。

问题 1 — 统计与否

- (a) 非统计数据。单一固定答案（~330 m）。
- (b) 统计。不同的学生→不同的身高。
- (c) 非统计数据。一只宠物，一个名字。
- (d) 统计。不同的六年级学生有不同的就寝时间。
- 强：解释统计问题预示着可变性。
- 弱：认为任何带有数字的问题都是统计问题。

问题 2 — 平均值、中位数、极差、MAD

- 平均值： $(8+7+9+6+10+8)/6 = 48/6 = 8$ 。
- 中位数： $(7,8,8,9 \rightarrow \text{中间两个数为 } 8,8) = 8$ 。
- 范围： $10 - 6 = 4$ 。
- MAD：与8的偏差为0、1、1、2、2、0。偏差算术平均数 $= 6/6 = 1$ 。所以 $MAD = 1$ 。
- 强：清晰显示偏差表；计算偏差的算术平均数。
- 弱：将偏差带符号（包括负偏差），给出 $MAD = 0$ ；或者错误地计算范围或方差。

问题 3 — 通过 MAD 比较两个类

- A 类更加一致（MAD 较低= 分数簇接近平均值）。
- B 类更加分散。
- 强：将MAD 解释为与平均值的典型距离。
- 弱：认为高MAD = 高分。

问题 4 — 中位数、Q1、Q3、IQR

- 数据（已排序）：4、7、8、8、10、12、13、15、18。
- $n=9$ ，中位数= 5，第一个值= 10。
- 下半部分(4,7,8,8)：中位数= 7.5 (Q1)。
- 上半部分(12,13,15,18)：中位数= 14 (Q3)。
- $IQR = 14 - 7.5 = 6.5$ 。
- 注意：将中位数包含在两半中的学生（给出 $Q1=8$ ， $Q3=13$ ）。

问题 5 — 相同范围，不同 IQR

- IQR 小但范围大意味着大多数数据都聚集在一起，尾部有一些异常值。具有相同范围的大 IQR 意味着中间的 50% 也被分散。
- 示例：{1、5、5、5、5、5、9}（范围 8，IQR 小）与 {1、2、4、5、6、8、9}（范围 8、IQR 较大）。
- 强：清晰、具体的例子。
- 弱：重述问题；没有具体的例子。

问题 6 — 猫体重、箱线图

- 最小值=7、Q1=9、中位数=10、Q3=12、最大值=15。
- 箱线图：从7到9和12到15的须线；盒子从9到12；10线。
- 强：正确识别五数摘要并绘制与比例成比例的箱线图。
- 弱：混淆Q1和Q3；画框不平整；使用点图而不是箱线图。

问题 7 — 直方图与箱线图

- 直方图：用于显示连续或分组数据的分布形状（偏斜、众数）。
- 箱线图：对于比较分布或总结分布（尤其是异常值）非常有用。
- 强：给出具体示例（“直方图表示年龄；箱线图比较男孩和女孩的身高”）。
- 弱：仅说明“当您想要查看数据时”。

问题 8 — Han 的均值与中位数

- 排序：4、7、7、12、50。中位数 = 7。
- 平均值 = $80/5 = 16$ 。
- 中位数 (7) 是更好的中心度量，因为50是拉高均值的异常值。
- 强：将50识别为异常值，并阐明中位数对异常值具有鲁棒性。
- 弱：捍卫均值，因为“它使用了所有数据”。
- ★ 离群值驱动的均值和中位数之间的选择是单位8判断的关键。

问题 9 — 向所有数据添加常量

- 迭戈错了。 $\{2,4,6\}$ 的算术平均数 = 4。添加5： $\{7,9,11\}$ 后，平均值 = 9。
- 平均值增加5 — 正是添加的常数。
- 强：用数据进行测试并阐明均值按常数变化。
- 弱：未经测试就同意；或计算出错误的平均值。

问题 10 — 原始调查

- 强：真正的统计问题（预期存在可变性）；合理的数据收集计划；深思熟虑的衡量标准选择（“我会使用中位数来衡量屏幕时间，因为有些孩子可能是异常值”）。
- 弱：问题不是统计性的；数据计划模糊；没有采取措施的理由。
- ★ 反思综合题——统计推理的最深入测试。