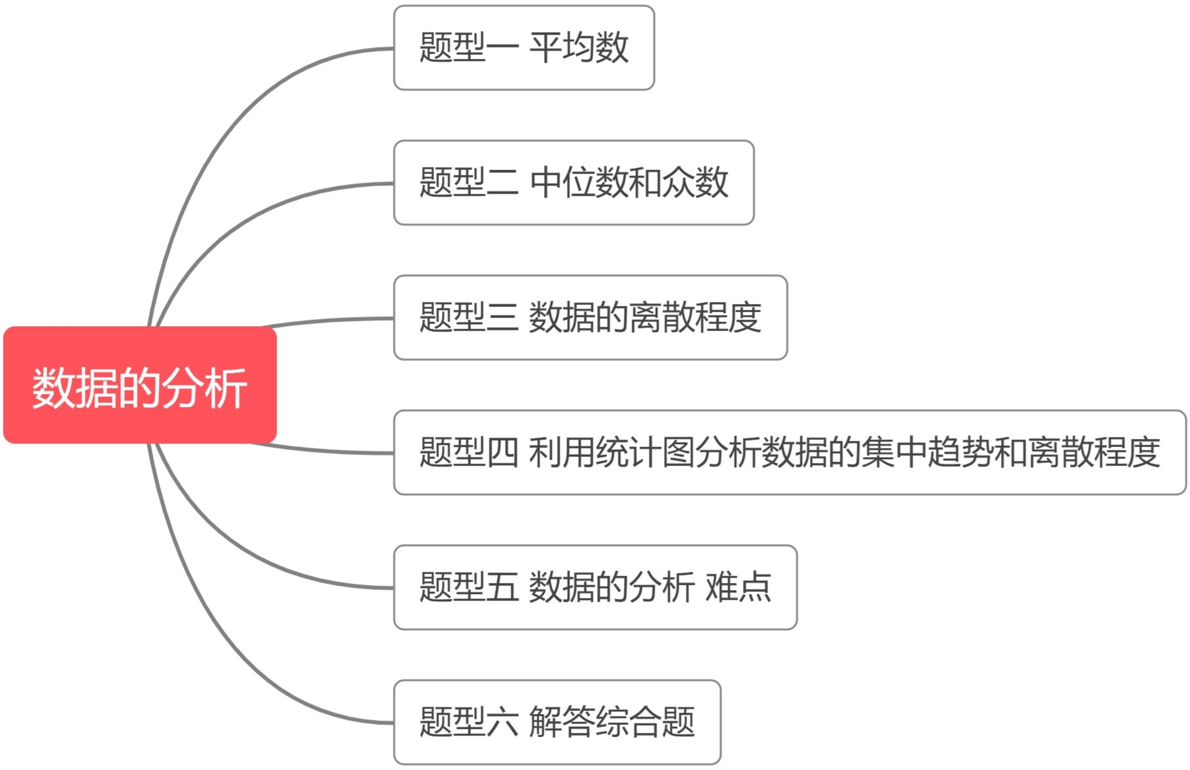
第六章 数据的分析 知识归纳与题型突破（六类题型清单）

01 思维导图



02 知识速记

一、算术平均数和加权平均数

一般地，对于 n 个数 x_1、x_2、x_3、… x_n，我们把 $\frac{1}{n}\left(x_1+x_2+x_3+\cdots+x_n\right)$ 叫做这 n 个数的**算术平均数**，简称平均数，记作 $\overline{x}$. 计算公式为 $\overline{x}=\frac{1}{n}\left(x_1+x_2+x_3+\cdots+x_n\right)$.

要点：

平均数表示一组数据的“平均水平”，反映了一组数据的集中趋势.

（1）当一组数据较大时，并且这些数据都在某一常数 a 附近上、下波动时，一般选用简化计算公式 $\overline{x}=\overline{x}'+a$. 其中 $\overline{x}'$ 为新数据的平均数，a 为取定的接近这组数据的平均数的较“整”的数.

（2）平均数的大小与一组数据里的每个数据均有关系，其中任一数据的变动都会相应引起平均数的变动. 所以平均数容易受到个别特殊值的影响.

若 n 个数 x_1、x_2、… x_n 的权分别是 w_1、w_2、…、w_n，则 $\frac{x_1w_1+x_2w_2+...+x_nw_n}{w_1+w_2+...+w_n}$ 叫做这 n 个数的**加权平均数**.

要点：

（1）相同数据 x_i 的个数 w_i 叫做权，w_i 越大，表示 x_i 的个数越多，“权”就越重. 数据的权能够反映数据的相对“重要程度”.

（2）加权平均数实际上是算术平均数的另一种表现形式，是平均数的简便运算.

二、中位数和众数

1.中位数

一般地，n 个数据按照大小顺序排列，处于最中间位置的一个数据（或最中间两个数据的平均数）叫做这组数据的**中位数**.

要点：

（1）一组数据的中位数是唯一的；一组数据的中位数不一定出现在这组数据中.

（2）由一组数据的中位数可以知道中位数以上和以下数据各占一半.

2.众数

一组数据中出现次数最多的那个数据叫做这组数据的**众数**.

要点：

（1）一组数据的众数一定出现在这组数据中；一组数据的众数可能不止一个.

（2）众数是一组数据中出现次数最多的数据而不是数据出现的次数.

三、平均数、中位数与众数的联系与区别

联系：平均数、众数、中位数都是用来描述数据集中趋势的量，其中以平均数最为重要.

区别：平均数的大小与每一个数据都有关，任何一个数的波动都会引起平均数的波动，当一组数据中有个别数据太高或太低，用平均数来描述整体趋势则不合适，用中位数或众数则较合适. 中位数与数据排列位置有关，个别数据的波动对中位数没影响；众数主要研究各数据出现的频数，当一组数据中不少数据多次重复出现时，可用众数来描述.

四、极差、方差和标准差

1.极差

一组数据中最大数据与最小数据的差，称为**极差**，极差＝最大数据－最小数据.

要点：

极差是最简单的一种度量数据波动情况的量，它受极端值的影响较大. 一组数据极差越小，这组数据就越稳定.

2.方差

方差是各个数据与平均数差的平方的平均数. 方差 s^2 的计算公式是：

$S^2=\frac{1}{n}\left[(x_1-\overline{x})^2+(x_2-\overline{x})^2+...+(x_n-\overline{x})^2\right]$，其中，$\overline{x}$ 是 x_1，x_2，… x_n 的平均数.

要点：

（1）方差反映的是一组数据偏离平均值的情况. 方差越大，数据的波动越大；方差越小，数据的波动越小.

（2）一组数据的每一个数都加上（或减去）同一个常数，所得的一组新数据的方差不变.

（3）一组数据的每一个数据都变为原来的 k 倍，则所得的一组新数据的方差变为原来的 k^2 倍.