



二、数值型数据的整理与显示



二、数值型数据的整理与显示

知识点：数据的分组★★★

知识点：数值型数据的图示★★



二、数值型数据的整理与显示

知识点：数据的分组★★★

分组的方法有**单变量值分组**和**组距分组**两种。

1.单变量值分组：把每一个变量值作为一组，通常只适合于离散变量且变量值较少的情况。

2.组距分组：将全部变量值依次划分为若干个区间，并将这一区间的变量值作为一组。在连续变量或变量值较多的情况下，通常采用组距分组。



二、数值型数据的整理与显示

采用组距分组需要经过以下几个步骤：

【举例】某高中一年级一班共有55名学生，高一语文考试中成绩分别为：

59 73 87 65 89 85 77 94 69 97

56 80 68 95 96 50 63 88 91 90

96 92 93 79 74 65 74 89 83 51

74 79 94 67 92 92 93 70 87 86

54 87 86 54 62 76 86 73 86 70

100 110 108 102 112



二、数值型数据的整理与显示

第一步，确定分組組數。

確定分組組數的要求是：

①劃分的組數，既不應太多也不應太少量組數過多。

②組數的確定，要盡量保證組間資料的差異性與組內資料的同質性。

③採用的分組辦法，要能夠充分顯示客觀現象本身存在的狀態。



二、数值型数据的整理与显示

第二步，对原始资料进行排序。

结果如下：

50 51 54 54 56 59 62 63 65 65
67 68 69 70 70 73 73 74 74 74
76 77 79 79 80 83 85 86 86 86
86 87 87 87 88 89 89 90 91 92
92 92 93 93 94 94 95 96 96 97
100 102 108 110 112



二、数值型数据的整理与显示

第三步，求极差。

将最大的观察值与最小的观察值相减便得到极差。

此例中，极差值为 $112-50=62$ 。



二、数值型数据的整理与显示

第四步，确定各组组距。

在实行等距分组的情况下，组距的确定办法为：

组距=极差/组数。

根据上式计算出来的组距，可能带有小数，为了编表和计算方便，也是审美习惯使然，最好把它取成接近于能被5除尽的一个数。本例中，组距=62/7 $\approx$ 8.9，组距可取10。

组距（每组观察值的最大差）：某组的上限值—该组的下限值

组距与组数成反比关系，组数越多，组距越小，组数越少，组距越大。



二、数值型数据的整理与显示

第五步，确定组限。

组限是组与组之间的界限，或者说是每组观察值变化的范围。

组限有上限与下限之分，在组距分组中，一个组的最小值称为下限，最大值称为上限；上限与下限的差值称为组距；上限值与下限值的平均数称为组中值。

组中值的代表性如何，取决于组中观察值的变化是否呈对称分布状态。



二、数值型数据的整理与显示

组中值=（上限值+下限值）/2

确定组限时应注意：

- ①第一组的下限值应比最小的观察值小一点，最后一组的上限值应比最大的观察值大一点。
- ②特别需要或不得已的情况除外，最好不要使用开口组。
- ③组限应取得美观些，按数字偏好，组限值应能被5除尽，且一般要用整数表示。



二、数值型数据的整理与显示

本例中，我们把第一组的下限值定为50，那么各组的组限依次为：

50~60， 60~70， 70~80， 80~90， 90~100， 100~110，
110~120。



二、数值型数据的整理与显示

第六步，确定各组观察值出现的频数。

遵循“不重不漏”的原则。

“不重”是指一项观察值只能分在其中的某一组，不能在其他组重复出现；

“不漏”是指组别能够穷尽，即在所分的全部组别中每项数据都能分在其中的某一组，不能遗漏。



二、数值型数据的整理与显示

为解决“不重”的问题，统计分组时习惯上规定“上组限不在内”，即当相邻两组的上下限重叠时，恰好等于某一组上限的观察值不算在本组内，而计算在下一组内。

在本例中，70这一数值不计算在“60~70”这一组中，而计算在“70~80”这一组中。



二、数值型数据的整理与显示

第七步，制作频数分布表，并填上相关的内容，以及其他需要说明的事项。



二、数值型数据的整理与显示

本例中的频数分布表：

观察值区间	组中值	频数	频率 (%)
50~60	55	6	10.9
60~70	65	7	12.7
70~80	75	11	20.0
80~90	85	13	23.6
90~100	95	13	23.6
100~110	105	3	5.5
110~120	115	2	3.7
合计	—	55	100.0



二、数值型数据的整理与显示

【例题·单选】将某单位在职职工年龄数据分为4组各组的组限依次为：“20~30” “30~40” “40~50” 和 “50~60”。按照统计分组的习惯规定50这一数值应计算在（ ）这一组中。

- A. “20~30”
- B. “30~40”
- C. “40~50”
- D. “50~60”



二、数值型数据的整理与显示

答案：D

解析：采用组距分组，需要遵循“不重不漏”的原则。“不重”是指一项观察值只能分在其中的某一组，不能在其他组重复出现。“不漏”是指组别能够穷尽，即在所分的全部组别中每项数据都能分在其中的某一组，不能遗漏。为解决“不重”的问题，统计分组时习惯上规定“上组限不在内”，即当相邻两组的上下限重叠时，恰好等于某一组上限的观察值不算在本组内，而计算在下一组内。在本题中，50这一数值不计算在“40~50”这一组中，而计算在“50~60”这一组中。



二、数值型数据的整理与显示

知识点：数值型数据的图示★★

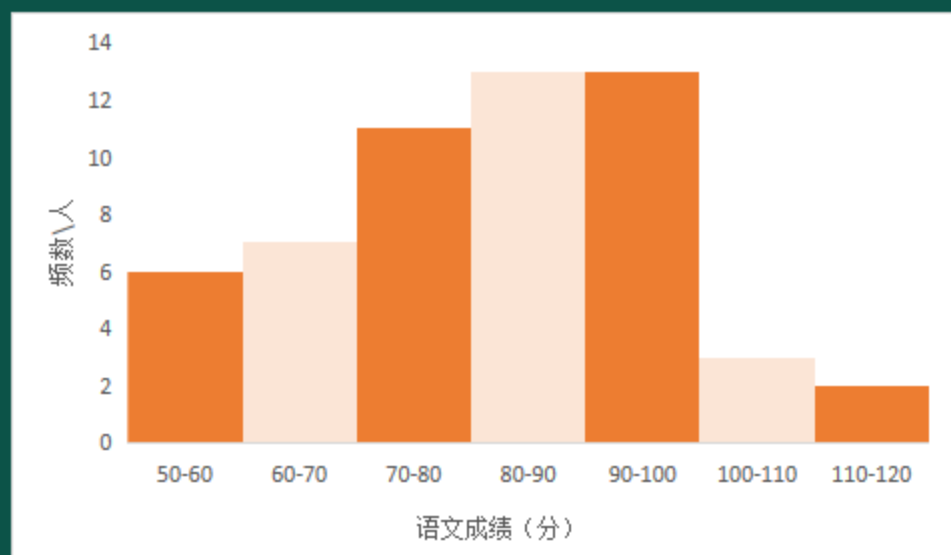
显示分组数据频数分布特征的图形有直方图、折线图等，上面介绍的条形图、圆形图等也都适用于显示数值型数据。



二、数值型数据的整理与显示

1. 直方图

直方图是用面积表示各组频数的多少，矩形的高度表示每一组的频数或百分比，宽度则表示各组的组距，因此其高度与宽度均有意义。直方图的各矩形通常是连续排列。





二、数值型数据的整理与显示

条形图：

是用条形的长度（横置时）表示各类别频数的多少，其宽度（表示类别）则是固定的。

条形图是分开排列。





二、数值型数据的整理与显示

2.折线图

折线图也称频数多边形图，它是在直方图的基础上，把直方图顶部的中点（即组中值）用直线连接起来，再把原来的直方图抹掉就是折线图。

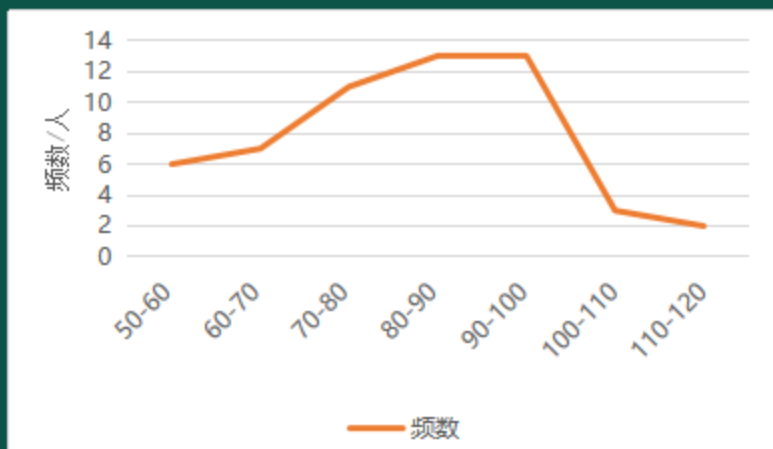


图20-5 某班高一语文成绩分布的折线图



二、数值型数据的整理与显示

【例题·单选】在直方图中，矩形的宽度表示（ ）。

- A. 各组的组距
- B. 各组所属的类别
- C. 各组的百分比
- D. 各组的频数



二、数值型数据的整理与显示

答案：A

解析：直方图是用矩形的宽度和高度来表示频数分布的图形。在平面直角坐标中，用横轴表示数据分组，纵轴表示频数或频率，所形成的矩形的面积表示各组频数的多少，矩形的高度表示每一组的频数或百分比，宽度则表示各组的组距，因此其高度与宽度均有意义。



二、数值型数据的整理与显示

数值型数据的整理与显示	数据的分组
	数值型数据的图示：直方图、折线图。



三、统计表



三、统计表

知识点：统计表的构成★

知识点：统计表的设计★★



三、统计表

知识点：统计表的构成★

统计表**一般由四个主要部分组成**，即**表头、行标题、列标题和数字资料**，**必要时**可以在统计表的下方加上表外**附加**。

表头应放在表的上方，它所说明的是统计表的主要内容。

行标题和列标题通常安排在统计表的第一列和第一行，它所表示的主要是所研究问题的类别名称和指标名称，通常也被称为“类”。

表外附加通常放在统计表的下方，主要包括资料来源、指标的注释和必要的说明等内容。



三、统计表

【例题·多选】统计表一般应包括（ ）。

- A. 附加
- B. 表头
- C. 行标题
- D. 数字资料
- E. 图例



三、统计表

答案：BCD

解析：统计表一般由四个主要部分组成，即表头（B）、行标题（C）、列标题和数字资料（D），必要时可以在统计表的下方加上表外附加。图例仅在统计图中出现。

A、E选项都不属于“一般”应包括的部分，因此本题不选。



三、统计表

知识点：统计表的设计★★

1.合理安排统计表的结构，如行标题、列标题、数字资料的位置应安排合理。

2.表头一般应包括表号、总标题和表中数据的单位等内容。

总标题应简明确切地概括出统计表的内容，一般需要表明统计数据的时间（When）、地点（Where）以及何种数据（What），即标题内容应满足“3W”要求。



三、统计表

3.表中的上下两条横线一般用粗线，中间的其他线要用细线，这样使人看起来清楚、醒目。通常情况下，统计表的左右两边不封口，列标题之间一般用竖线隔开，而行标题之间通常不必用横线隔开。总之，表中尽量少用横竖线。表中的数据一般是右对齐，有小数点时应以小数点对齐，而且小数点的位数应统一。对于没有数字的表格单元，一般用“—”表示，一张填好的统计表不应出现空白单元格。



三、统计表

4. 在使用统计表时，必要时可在表的下方加上注释，特别要注意注明资料来源，以表示对他人劳动成果和知识产权的尊重，方便读者查阅使用。



三、统计表

统计表的构成	表头、行标题、列标题和数字资料
统计表的设计	合理安排统计表的结构 表头包括表号、总标题和表中数据的单位等内容 表中的上下两条横线用粗线，中间的其他线要用细线 可在表的下方加上注释，注明资料来源

谢谢 观看

THANK YOU