

第

5

数据分组与频数统计

章

数据分组就是根据统计研究任务，按照一定的要求，把研究的社会现象总体划分为若干个性质相同的组。而一组测量值中，落在各组内的数据个数就称为频数。

数据分组与频数统计的目的是为了区分现象的不同类型，研究总体的内部结构，分析现象间的依存关系。

- 离散型数据与连续型数据的特点
- 离散型数据的分组
- 连续型数据的分组
- 利用函数统计频数
- 利用直方图统计频数
- 运用直方图统计图表

5.1 数据分组



在统计学中，数据按变量值是否连续，可分为离散数据与连续数据两种。分清数据性质才能选择合适的分组方式。

- 离散数据：其数值只能用自然数或整数单位计算。例如，企业个数、职工人数、设备台数等，只能按计量单位数计数。
- 连续数据：在一定区间中可以任意取值，其数值是连续不断的，相邻两个数值可作无限分割，即可取无限个数值。例如，生产零件的规格尺寸、身高的测量值、体重的测量值等。

5.1.1 离散型数据分组——单项式分组

对于离散型的变量，如果变量值的变动幅度小，就可以一个变量值对应一组，称单项式分组。如居民家庭按儿童数分组、毕业生按年龄分组等，均可采用单项式分组。

1. COUNTIF 分组统计

如图 5-1 所示，记录了某企业中在相同条件下不同工人的生产数量数据（篇幅限制，只显示部分记录），要求对工人生产数量的水平进行分析。由于抽样的生产数量变化幅度不大，此时可以使用单项式分组，一个数据分为一组，即每个生产数量都是离散型数据分组的界限。

	A	B	C
1	抽样编号	操作人	生产数量
2	1	李为洋	58
3	2	杨依娜	59
4	3	朱子进	60
5	4	曹正	59
6	5	郭丽	61
7	6	王雪峰	60
8	7	吴东梅	61
9	8	张以军	60
10	9	孙倩	62
11	10	简志能	60
12	11	李军	61
13	12	顾源	60
14	13	刘维	61
15	14	李婷婷	62
16	15	周玉杰	58
17	16	华新伟	59
18	17	邹志志	60
19	18	韩志	59
20	19	吴伟云	61
21	20	杨清	60
22	21	李欣	58
23	22	金鑫	61
24	23	华涵涵	62
25	24	张玮	60
26	25	聂新余	60

图 5-1

① 在工作表空白位置建立生产数量分布表格并建立人数统计标识，如图 5-2 所示。

	A	B	C	D	E	F
1	抽样编号	操作人	生产数量		生产数量	人数
2	1	李为洋	58		58	
3	2	杨依娜	59		59	
4	3	朱子进	60		60	
5	4	曹正	59		61	
6	5	郭丽	61		62	
7	6	王雪峰	60			
8	7	吴东梅	61			
9	8	张以军	60			
10	9	孙倩	62			
11	10	简志能	60			
12	11	李军	61			
13	12	顾源	60			
14	13	刘维	61			
15	14	李婷婷	62			
16	15	周玉杰	58			

图 5-2

② 选中 F2 单元格，在编辑栏中输入公式：
=COUNTIF(\$C\$2:\$C\$31,E2)

按 Enter 键即可计算出生产数量为 58 件的人数，如图 5-3 所示。

	A	B	C	D	E	F
F2						3
1	抽样编号	操作人	生产数量		生产数量	人数
2	1	李为洋	58		58	3
3	2	杨依娜	59		59	
4	3	朱子进	60		60	
5	4	曹正	59		61	
6	5	郭丽	61		62	
7	6	王雪峰	60			
8	7	吴东梅	61			
9	8	张以军	60			
10	9	孙倩	62			

图 5-3

③ 将光标定位到 F2 单元格右下角，向下填充公式至 F6 单元格，即可计算出其他生产数量的人数，如图 5-4 所示。

1	A	B	C	D	E	F
2	抽样编号	操作人	生产数量		生产数量	人数
3	1	李为洋	58		58	3
4	2	杨依娜	59		59	4
5	3	朱子进	60		60	10
6	4	曹正	59		61	9
7	5	郭丽	61		62	4
8	6	王雪峰	60			
9	7	吴东梅	61			
10	8	张以军	60			
11	9	孙倩	62			
12	10	简志能	60			
13	11	李军	61			
14	12	顾源	60			

图 5-4

由统计结果看到，生产数量为 60 和 61 的人最多。

知识扩展

COUNTIF 函数用于统计指定区域中符合指定条件的单元格数目。本例中公式“=COUNTIF(\$C\$2:\$C\$31,E2)”表示统计出 \$C\$2:\$C\$31 单元格区域中等于 E2 单元格数据的记录条数，即统计 \$C\$2:\$C\$31 单元格区域中“58”的个数。

因为这个公式建立后需要向下复制，因此设计公式时要注意对数据源的引用方式，不想变化的需要使用绝对引用方式。本例中的第一个参数为用于统计判断的区域，是始终不变的，所以使用绝对引用方式；而当公式向下复制时，需要改变的是第二个参数（即被统计的对象），所以得到相对引用方式。

2. 数据透视表分组统计

离散型数据的单项式分组也可以使用数据透视表功能快速实现。

① 选中表格任意单元格区域，在“插入”选项卡的“表格”组中单击“数据透视表”按钮，如图 5-5 所示。

② 打开“创建数据透视表”对话框，保持默认选项，如图 5-6 所示。

③ 单击“确定”按钮创建数据透视表。设置“生产数量”字段为行标签，再设置“操作人”字段为值标签，如图 5-7 所示。



图 5-5



图 5-6



图 5-7

④ 在“数据透视表工具”“设计”选项卡的“布局”组中单击“报表布局”按钮，在下拉列表中单击“以表格形式显示”，如图 5-8 所示。此时可以让“生产数量”这个字段名称显示出来，如图 5-9 所示。



图 5-8

	A	B
1		
2		
3	生产数量	计数项: 操作人
4	58	3
5	59	4
6	60	10
7	61	9
8	62	4
9	总计	30
10		

图 5-9

⑤ 将 B3 单元格中的名称更改为“人数”(选中单元格，在编辑栏中去输入)，如图 5-10 所示。

	A	B	C
1			
2			
3	生产数量	人数	
4	58	3	
5	59	4	
6	60	10	
7	61	9	
8	62	4	
9	总计	30	
10			

图 5-10

5.1.2 离散型数据分组——组距式分组

如果离散变量值的变动幅度很大，个数很多，则可以把整个变量值依次划分为几个区间，各变量值按其大小确定所归并的区间，区间的距离称为组距，这样的分组称为组距式分组。在组距式分组中，相邻组既可以有确定的上下限，也可以将相邻组的组限重叠。

如图 5-11 所示，表格中记录了某种技能考试的成绩数据（篇幅限制，只显示部分记录），显然调查的数据是离散型的，最低分是 68 分，最高分是 99 分，变化幅度很大，此时应该将性质相似的数据分为同组，性质悬殊的分为不同的组。将数据分为 0~70、70~80、80~90 以及 90~100 这 4 个区间，因此 70、80 和 90 这 3 个数字就成了离散型数据的分组界限。

① 在表格空白处将数据分为“ ≤ 70 ”、“70-80”、“80-90”、“90-100”4 组数据，如图 5-11 所示。

	A	B	C	D	E	F	G
1	某技能考试成绩数据分析					区间	数量
2	1组	2组	3组	4组		≤ 70	
3	89	82	78	98		70-80	
4	98	87	90	99		80-90	
5	69	80	77	96		90-100	
6	87	73	85	94			
7	85	85	82	96			
8	85	90	91	88			
9	95	70	90	94			
10	68	89	87	96			
11	78	87	82	98			
12	82	78	86	96			
13	85	81	91	82			
14	98	96	94	98			
15	99	98	96	92			
16	91	91	97	87			
17	96	94	87	82			
18	96	97	96	94			

图 5-11

② 选中 G2 单元格，在编辑栏中输入公式：
=COUNTIF(A3:D18," ≤ 70 ")

按 Enter 键，即可统计出小于等于 70 分的人数，如图 5-12 所示。

③ 选中 G3 单元格，在编辑栏中输入公式：
=COUNTIFS(A3:D18," >70 ",A3:D18," ≤ 80 ")

按 Enter 键，即可计算出分数在 70 ~ 80 分的人数，如图 5-13 所示。

某技能考试成绩数据分析						区间	数量
1							
2	1组	2组	3组	4组	<=70	3	
3	89	82	78	98	70-80		
4	98	87	90	99	80-90		
5	69	80	77	96	90-100		
6	87	73	85	94			
7	85	85	82	96			
8	85	90	91	88			
9	95	70	90	94			
10	68	89	87	96			
11	78	87	82	98			
12	82	78	86	96			
13	85	81	91	82			
14	98	96	94	98			
15	99	98	96	92			
16	91	91	97	87			
17	96	94	87	82			
18	96	97	96	94			

图 5-12

某技能考试成绩数据分析						区间	数量
1							
2	1组	2组	3组	4组	<=70	3	
3	89	82	78	98	70-80	6	
4	98	87	90	99	80-90		
5	69	80	77	96	90-100		
6	87	73	85	94			
7	85	85	82	96			
8	85	90	91	88			
9	95	70	90	94			
10	68	89	87	96			
11	78	87	82	98			
12	82	78	86	96			
13	85	81	91	82			
14	98	96	94	98			
15	99	98	96	92			
16	91	91	97	87			
17	96	94	87	82			
18	96	97	96	94			

图 5-13

知识扩展

COUNTIF 函数用于统计指定区域中满足多个条件的单元格数目。本例中公式“=COUNTIFS(A3:D18,>70,A3:D18,<=80)”将“>70”和“<=80”作为两个条件写入参数中,表示统计出 A3:D18 单元格区域中同时满足“>70”和“<=80”这两个条件的单元格个数。

当只有一个条件时使用 COUNTIF 函数,有两个条件时使用 COUNTIFS 函数。

④ 选中 G4 单元格,在编辑栏中输入公式:

=COUNTIFS(A3:D18,>80,A3:D18,<=90)

按 Enter 键,即可计算出分数在 80 ~ 90 分的人数,如图 5-14 所示。

某技能考试成绩数据分析						区间	数量
1							
2	1组	2组	3组	4组	<=70	3	
3	89	82	78	98	70-80	6	
4	98	87	90	99	80-90	25	
5	69	80	77	96	90-100		
6	87	73	85	94			
7	85	85	82	96			
8	85	90	91	88			
9	95	70	90	94			
10	68	89	87	96			
11	78	87	82	98			
12	82	78	86	96			
13	85	81	91	82			
14	98	96	94	98			
15	99	98	96	92			
16	91	91	97	87			
17	96	94	87	82			
18	96	97	96	94			

图 5-14

⑤ 选中 G5 单元格,在编辑栏中输入公式:

=COUNTIF(A3:D18,>90)

按 Enter 键,即可计算出分数在 90 ~ 100 分的人数,如图 5-15 所示。

某技能考试成绩数据分析						区间	数量
1							
2	1组	2组	3组	4组	<=70	3	
3	89	82	78	98	70-80	6	
4	98	87	90	99	80-90	25	
5	69	80	77	96	90-100	30	
6	87	73	85	94			
7	85	85	82	96			
8	85	90	91	88			
9	95	70	90	94			
10	68	89	87	96			
11	78	87	82	98			
12	82	78	86	96			
13	85	81	91	82			
14	98	96	94	98			
15	99	98	96	92			
16	91	91	97	87			
17	96	94	87	82			
18	96	97	96	94			

图 5-15

由统计结果看到,分数在 90 ~ 100 分的人数是最多的。

5.1.3 连续型数据分组界限

鉴于连续型变量的特性,连续型数据无法全部列举其数值,其分组只能是组距式分组。但在按数量标志分组时,各个分组的数量界限的选择必须能反映各个样本的本质差异,还需要根据被研究的现象总体的数量特征,采用适

当的组数，确定适合的组距。

如图 5-16 所示，登记了东三省主要城市的海拔高度，根据地貌特征，海拔在 200 m 以下为平原，海拔介于 200~500 m 为丘陵地带，海拔高于 500 m 的面积广大地带为高原地带。根据这种地理学常识来对各个城市的海拔数据进行分组，以反映数据本质特征，而 200、500 则是该连续型分组的科学界限。

东三省海拔高度表		
省份	城市	海拔高度(m)
黑龙江省	贵阳	1071.2
	思南	416.3
	遵义	843.9
	毕节	1510.6
	威宁	2237.5
	安顺	1392.9
	独山	972.2
	兴仁	1378.5
	鹤岗	228
	海拉尔	613
吉林省	博克图	739
	长春	237
	吉林	184
	四平	164
	通化	403
	通辽	180
	开鲁	235
辽宁省	阜新	138
	抚顺	82
	沈阳	42
	锦州	66
	鞍山	22
	营口	4
	丹东	15
大连	62	

图 5-16

① 在表格空白处将数据分为 ≤200、200~500、>500 这 3 组数据，并添加分析表格，如图 5-17 所示。

东三省海拔高度表					
省份	城市	海拔高度(m)	地貌	组段	数量
黑龙江省	贵阳	1071.2	平原	≤200	
	思南	416.3	丘陵	200~500	
	遵义	843.9	高原	>500	
	毕节	1510.6			
	威宁	2237.5			
	安顺	1392.9			
	独山	972.2			
	兴仁	1378.5			
	鹤岗	228			
	海拉尔	613			
吉林省	博克图	739			
	长春	237			
	吉林	184			
	四平	164			
	通化	403			
	通辽	180			
	开鲁	235			
辽宁省	阜新	138			
	抚顺	82			
	沈阳	42			
	锦州	66			
	鞍山	22			
	营口	4			
	丹东	15			
大连	62				

图 5-17

② 选中 G3 单元格，在编辑栏中输入公式：

=COUNTIF(C3:C27,"<=200")

按 Enter 键，即可计算出平原地形城市个数，如图 5-18 所示。

东三省海拔高度表					
省份	城市	海拔高度(m)	地貌	组段	数量
黑龙江省	贵阳	1071.2	平原	≤200	11
	思南	416.3	丘陵	200~500	
	遵义	843.9	高原	>500	
	毕节	1510.6			
	威宁	2237.5			
	安顺	1392.9			
	独山	972.2			
	兴仁	1378.5			
	鹤岗	228			
	海拉尔	613			
吉林省	博克图	739			
	长春	237			
	吉林	184			
	四平	164			
	通化	403			
	通辽	180			
	开鲁	235			
辽宁省	阜新	138			
	抚顺	82			
	沈阳	42			
	锦州	66			
	鞍山	22			
	营口	4			
	丹东	15			
大连	62				

图 5-18

③ 选中 G4 单元格，在编辑栏中输入公式：

=COUNTIFS(C3:C27,">200",C3:C27,"<=500")

按 Enter 键，即可计算出丘陵地形城市个数，如图 5-19 所示。

东三省海拔高度表					
省份	城市	海拔高度(m)	地貌	组段	数量
黑龙江省	贵阳	1071.2	平原	≤200	11
	思南	416.3	丘陵	200~500	5
	遵义	843.9	高原	>500	
	毕节	1510.6			
	威宁	2237.5			
	安顺	1392.9			
	独山	972.2			
	兴仁	1378.5			
	鹤岗	228			
	海拉尔	613			
吉林省	博克图	739			
	长春	237			
	吉林	184			
	四平	164			
	通化	403			
	通辽	180			
	开鲁	235			
辽宁省	阜新	138			
	抚顺	82			
	沈阳	42			
	锦州	66			
	鞍山	22			
	营口	4			
	丹东	15			
大连	62				

图 5-19

④ 选中 G5 单元格，在编辑栏中输入公式：

=COUNTIF(C3:C27,">500")

按 Enter 键，即可计算出高原地形城市个数，如图 5-20 所示。

东三省海拔高度表					
省份	城市	海拔高度(m)	地貌	组段	数量
黑龙江省	贵阳	1071.2	平原	≤200	11
	思南	416.3	丘陵	200~500	5
	遵义	843.9	高原	>500	9
	毕节	1510.6			
	威宁	2237.5			
	安顺	1392.9			
	独山	972.2			
	兴仁	1378.5			
	鹤岗	228			
	海拉尔	613			
吉林省	博克图	739			
	长春	237			
	吉林	184			
	四平	164			
	通化	403			
	通辽	180			
	开鲁	235			
辽宁省	阜新	138			
	抚顺	82			
	沈阳	42			
	锦州	66			
	鞍山	22			
	营口	4			
	丹东	15			
大连	62				

图 5-20

5.2 频数统计



频数又称次数。在一组测量值中，当按一定的组距将其分组时出现在各组内的数据个数就称为频数。按分组依次排列的频数构成频数数列，用来说明各组标志值对全体标志值所起作用的强

度。各组频数的总和等于总体的全部单位数。

频数的表示方法既可以是表，也可以是图形。

5.2.1 单项式分组的频数统计

单项式分组的频数统计是针对离散型数据中变量值变动幅度较小，可用于单项式分组的数据，因此其频数统计方法实际就是 5.1.1 小节介绍的操作，可以使用 COUNTIF 函数或数据透视表来求解。

如图 5-21 所示，记录了 2019 年某地区对每个家庭儿童数量的抽样数据（共 100 个数据），可以统计出频数，并与之前年份的儿童数量进行比较，以分析二胎政策之后的人口增加情况。

① 在工作表空白部分建立分组表格。选中 G3 单元格，在编辑栏中输入公式：

```
=COUNTIF(SA$2:$D$26,F3)
```

按 Enter 键，即可计算出儿童数为“1”的家庭数，如图 5-21 所示。

G3		=COUNTIF(SA\$2:\$D\$26,F3)					
	A	B	C	D	E	F	G
1	某地区每个家庭儿童数量抽样						
2	1	2	2	3		儿童数	家庭数
3	2	2	1	1		1	56
4	1	2	2	1		2	
5	4	2	1	1		3	
6	1	2	2	1		4	
7	1	1	1	2			
8	2	1	1	2			
9	1	1	1	1			
10	3	2	1	2			
11	1	1	2	1			
12	2	1	1	2			
13	3	1	3	3			
14	2	2	1	2			
15	1	1	1	1			
16	2	2	1	1			
17	2	2	2	2			
18	1	2	2	1			
19	1	1	1	1			
20	2	2	2	2			
21	1	2	1	2			
22	1	1	1	1			
23	1	1	1	1			
24	1	1	2	3			
25	2	2	1	1			
26	1	1	1	1			

图 5-21

② 选中 G3 单元格，向下填充公式至 G7 单元格中，即可计算出其他儿童数量对应的家庭数，如图 5-22 所示。

	A	B	C	D	E	F	G
1	某地区每个家庭儿童数量抽样						
2	1	2	2	3		儿童数	家庭数
3	2	2	1	1		1	56
4	1	2	2	1		2	37
5	4	2	1	1		3	6
6	1	2	2	1		4	1
7	1	1	1	2			
8	2	1	1	2			
9	1	1	1	1			
10	3	2	1	2			
11	1	1	2	1			
12	2	1	1	2			
13	3	1	3	3			
14	2	2	1	2			
15	1	1	1	1			
16	2	2	1	1			
17	2	2	2	2			
18	1	2	2	1			
19	1	1	1	1			
20	2	2	2	2			
21	1	2	1	2			
22	1	1	1	1			
23	1	1	1	1			
24	1	1	2	3			
25	2	2	1	1			
26	1	1	1	1			

图 5-22

从统计结果可以看到，频数最高的为“1”和“2”，与 2019 年之前的年份相比较，可以确定二孩家庭逐渐增多。

5.2.2 组距式分组的频数统计

组距式分组的频数统计先要确定全距，然后根据全距确定组数和组距，最后根据分组的情况来确定组项。确定组项时要注意以下几点：

- 最小值的下限要低于最小值变量，最大值的上限应高于最大值变量；
- 组限的确定有利于表现出总体分布的特点，应反映出事物的变化；
- 组限尽可能选取整数。

如图 5-23 所示，登记了全国 50 个城市的房价数据，现在对这些数据进行分组，并计算出频数。在 Excel 2019 中，组距式分组的频数统计一般使用函数 FREQUENCY 来实现，而且非常方便快捷。

① 在表格空白处创建分组过程表格和分组结果表格，如图 5-24 所示。

② 选中 H2 单元格，在编辑栏中输入公式：

```
=MAX(C2:C51)
```


7 选中 I11:J17 单元格区域, 在编辑栏中输入公式:

```
=FREQUENCY(C2:C51, G11:G17)
```

按 Shift+Ctrl+Enter 组合键, 即可计算出各个区间对应的频数, 如图 5-30 所示。

排名	城市	单价(元/m ²)	环比	同比	分组过程	频数	
1	深圳	69129	-0.0001	0.1867	最大值	69129	
2	北京	62212	-0.0166	-0.0188	最小值	9857	
3	上海	56256	-0.0004	0.0665	全距	59272	
4	厦门	52583	0.0112	0.0394	组数	6	
5	广州	51729	-0.0195	0.1295	组距	9878.66667	
6	三亚	41075	0.0074	0.1383	组距选取	10000	
7	南京	30947	0.0032	0.0076			
8	杭州	31524	0.0309	-0.0371			
9	福州	25838	-0.003	-0.0254	分组结果		
10	天津	24719	-0.0137	-0.0529	界限	区间	
11	宁波	23871	-0.0038	0.1603	10000	<=10000	1
12	苏州	22640	-0.0067	0.0034	30000	10000-20000	34
13	珠海	22309	0.0169	-0.0447	40000	20000-30000	7
14	温州	21382	-0.0018	0.02	50000	30000-40000	2
15	青岛	20535	0.0263	-0.0463	60000	40000-50000	1
16	武汉	18952	-0.0457	0.0563	70000	50000-60000	3
17	丽水	18209	-0.017	0.0227	80000	60000-70000	2

图 5-30

5.3 频数统计直方图

在 Excel 的高级分析工具有直方图工具, 使用此工具可以快速进行频数统计并生成图表。这是一项非常实用的功能。

5.3.1 加载直方图分析工具

要想使用分析工具对表格数据分析, 首先需要安装分析工具库加载项, 加载步骤如下。

1 打开表格, 单击“文件”选项卡, 在打开的面板中单击“选项”命令(如图 5-31 所示), 打开“Excel 选项”对话框。单击“加载项”右侧面板的“转到”按钮(如图 5-32 所示), 打开“加载项”对话框。

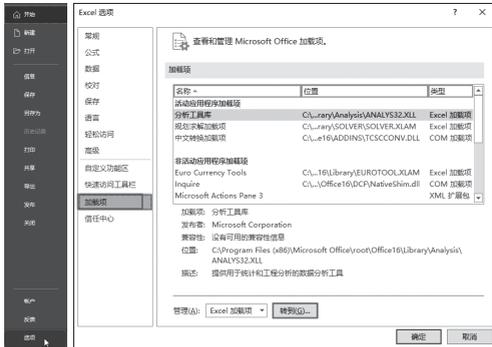


图 5-31



图 5-32

2 选中“分析工具库”复选框, 单击“确定”

通过计算结果, 可以看到房价分布在哪个价格区间的数量最多。

知识扩展

FREQUENCY 函数计算数值在某个区域内的出现频率, 然后返回一个垂直数组。由于函数 FREQUENCY 返回一个数组, 所以它必须以数组公式的形式输入。函数语法如下:

```
FREQUENCY(data_array, bins_array)
```

- data_array: 是一个数组或对一组数值的引用, 要为其计算频率。

- bins_array: 是一个区间数组或对区间的引用, 该区间用于对 data_array 中的数值进行分组。



按钮, 如图 5-33 所示。完成加载后, 在“数据”选项卡的“分析”组中单击“数据分析”按钮, 如图 5-34 所示。单击就可以打开“数据分析”对话框。

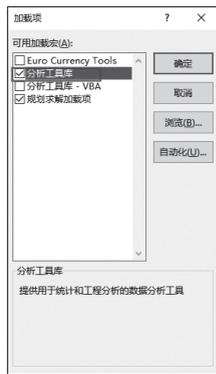


图 5-33



图 5-34

5.3.2 应用直方图分析工具

在加载了直方图分析工具后，可以一次性对频数进行统计并自动生成图表。例如当前表格登记了100份调查问卷对某产品的评分情况，现在对这些数据以10为组距进行分组，计算出频数，并生成频数统计图表。

① 在数据旁建立组限（组距以10为区间）。在“数据”选项卡的“分析”组中单击“数据分析”按钮（如图5-35所示），打开“数据分析”对话框。

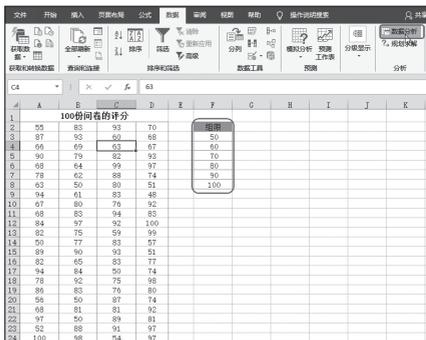


图 5-35

② 选择“直方图”工具（如图5-36所示），单击“确定”按钮进入“直方图”对话框，按如图5-37所示设置各项参数，并选中“图表输出”复选框。注意“输入区域”为整个数据区域，“接收区域”为设置的组限。

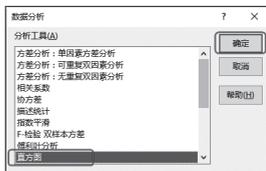


图 5-36



图 5-37

③ 单击“确定”按钮即可快速统计频数并生成直方图，如图5-38所示。

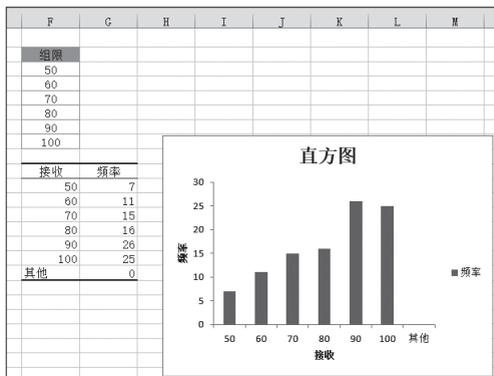


图 5-38

④ 由于图表中默认包含有一个“其他”分类，可以通过设置取消此分类。选中图表并单击鼠标右键，在弹出的右键菜单中单击“选择数据”命令，如图5-39所示。

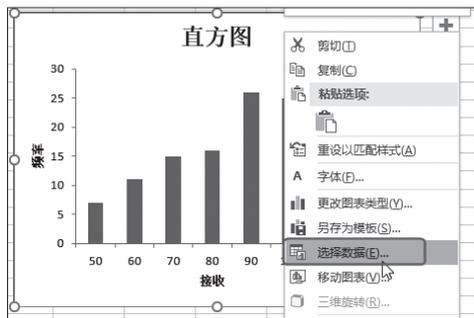


图 5-39

⑤ 打开“选择数据源”对话框，重新设置图表数据区域为“=直方图!\$F\$11:\$G\$16”，如图5-40所示。



图 5-40

⑥ 在“水平(分类)轴标签”区域单击“编辑”按钮,打开“轴标签”对话框。拖动选择“F11:F16”单元格区域为轴标签,如图 5-41 所示。

⑦ 单击“确定”按钮回到“选择数据源”对话框,选中“系列1”系列,然后单击“删除”按钮,如图 5-42 所示。

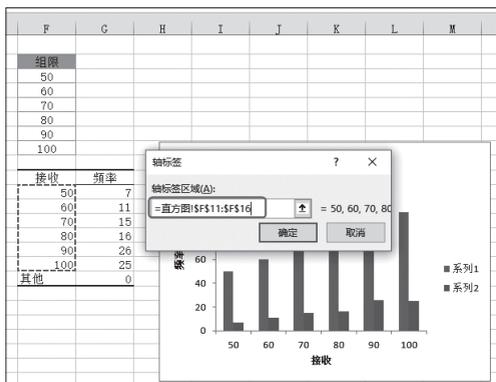


图 5-41



图 5-42

⑧ 单击“确定”按钮得到调整好的直方图,如图 5-43 所示。

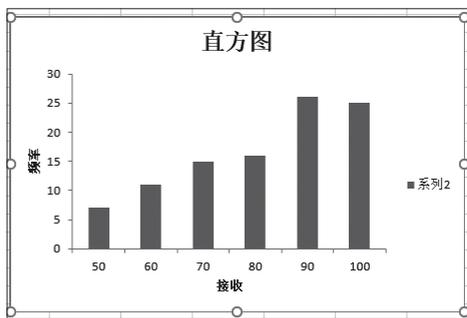


图 5-43

⑨ 在数据系列上单击鼠标右键,在打开的右键

菜单中执行“添加数据标签”“添加数据标签”命令(如图 5-44 所示),即可在图表上显示出“值”数据标签,即每个分组区间的频数值是多少,如图 5-45 所示。

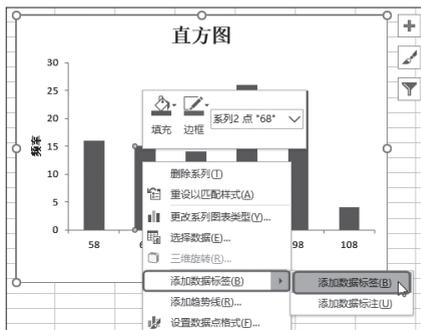


图 5-44

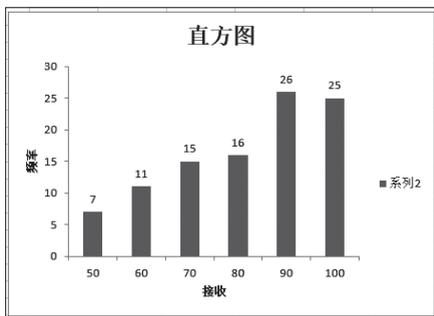


图 5-45

⑩ 选中图表,单击图表右上角出现的“图表样式”按钮,在展开的列表中找到想使用的样式,单击即可套用(鼠标指针指向即时预览),如图 5-46 所示。利用此方法可以达到一键美化图表的目的。

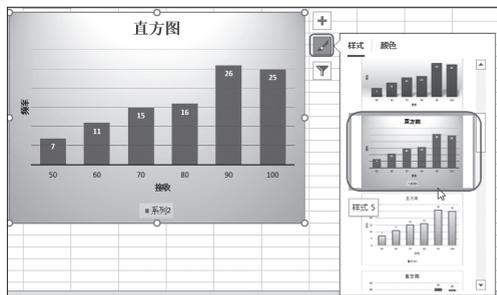


图 5-46

最后为图表添加上能表达主题的标题文字,效果如图 5-47 所示。

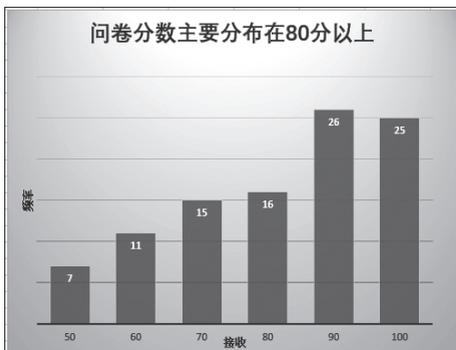


图 5-47

5.3.3 应用直方图统计图表

从 Excel 2016 开始, Excel 在图表类别中提供了“直方图”统计图表类型, 可对同一列数据进行频数分布统计。针对 5.3.2 小节中的数据, 为了能顺利创建图表, 还需要对源数据进行处理, 下面来介绍具体操作步骤。

① 将源数据改为到一列显示 (可以复制工作表来完成此操作), 如图 5-48 所示。

	A	B	C	D	E
1	100份问卷的评分				
2	55	83	93	70	
3	87	93	60	68	
4	66	69	63	67	
5	90	79	82	93	
6	68	64	99	97	
7	78	62	88	74	
8	63	50	80	51	
9	94	61	83	48	
10	67	80	76	92	
11	68	83	94	83	
12	84	97	92	100	
13	82	75	59	99	
14	50	77	83	57	
15	89	90	93	51	
16	82	65	83	77	
17	94	84	50	74	
18	78	92	75	98	
19	86	83	76	80	
20	56	50	87	74	
21	68	81	81	92	
22	97	50	89	81	
23	52	88	91	97	
24	100	98	54	97	
25	92	75	88	50	
26	98	56	87	53	
27	83				
28	93				
29	69				
30	79				
31	64				
32	62				
33	50				
34	61				
35	80				
36	83				
37	97				
38	75				
39	77				

图 5-48

② 选中 A2:A101 单元格区域, 单击“插入”选项卡, 在“图表”选项组中单击“插入统计图表”下拉按钮, 在其下拉菜单中选择“直方图”子图表类型, 如图 5-49 所示。

③ 执行上述命令后即可立即插入默认图表, 如图 5-50 所示。

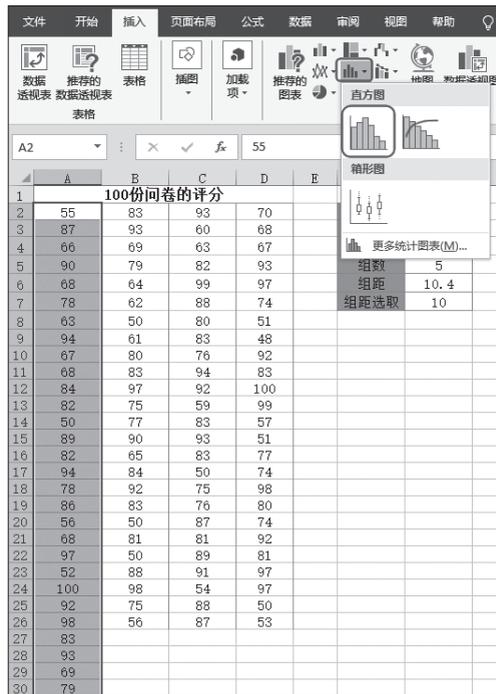


图 5-49

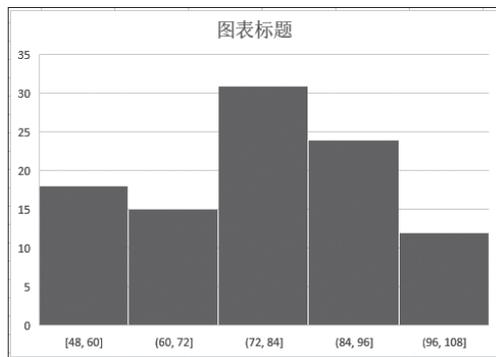


图 5-50

④ 对于默认图表需要根据实际情况更改箱体宽度, 即各个箱体以多少为区间进行分组。这时就要

考虑当前的组距了。在水平轴的标题上双击鼠标, 打开“设置坐标轴格式”右侧窗格, 单击“坐标轴选项”标签, 在“箱宽度”框中设置值为“10”(即组距值), 设置“溢出的箱”为“100”, “下溢箱”为“50”, 如图 5-51 所示。设置后可以看见图表效果, 如图 5-52 所示。



图 5-51

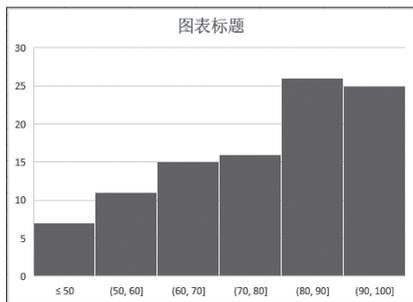


图 5-52

⑤ 接着再在箱体上单击鼠标右键, 在打开的右键菜单中执行“添加数据标签”“添加数据标签”命令, 如图 5-53 所示。

⑥ 添加数据标签后再为图表重新输入标题, 得到的直方图如图 5-54 所示。从直方图中可以清晰地看到数据的分组情况以及频数统计值。

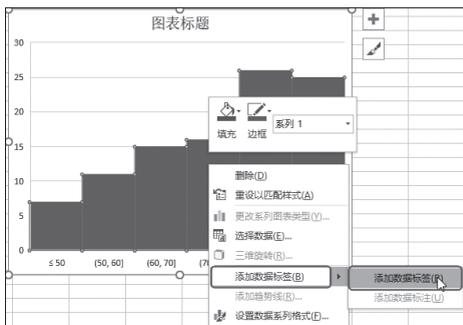


图 5-53

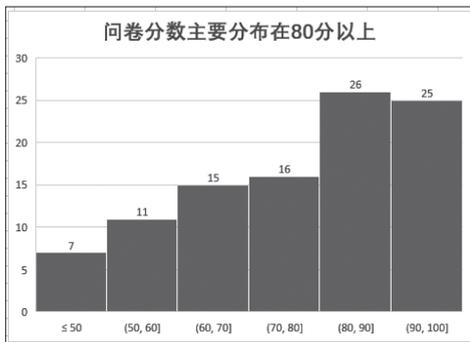


图 5-54