



“十二五”职业教育国家规划教材
经全国职业教育教材审定委员会审定

财经商贸类职业技能大赛参考教材

基础统计

第二版

总主编 丁增稳

主 编 朱兆军 吴 适

基础统计
(第二版)

总主编 丁增稳
主 编 朱兆军 吴 适

北京出版集团
北京出版社

北京出版集团
北京出版社



图书在版编目 (CIP) 数据

基础统计 / 朱兆军, 吴适主编. — 2 版. — 北京:
北京出版社, 2020.8 (2023 重印)

高职十二五规划教材: 2014 版

ISBN 978-7-200-15811-3

I. ①基… II. ①朱… ②吴… III. ①统计学—高等
职业教育—教材 IV. ① C8

中国版本图书馆 CIP 数据核字 (2020) 第 151475 号

基础统计 (第二版)

JICHU TONGJI (DI-ER BAN)

主 编: 朱兆军 吴 适

出 版: 北京出版集团
北京出版社

地 址: 北京北三环中路 6 号

邮 编: 100120

网 址: www.bph.com.cn

总 发 行: 北京出版集团

经 销: 新华书店

印 刷: 定州启航印刷有限公司

版 印 次: 2020 年 8 月第 2 版 2023 年 6 月修订 2023 年 7 月第 3 次印刷

成品尺寸: 185 毫米 × 260 毫米

印 张: 18

字 数: 405 千字

书 号: ISBN 978-7-200-15811-3

定 价: 54.00 元

教材意见建议接收方式: 010-58572162 邮箱: jiaocai@bphg.com.cn

如有印装质量问题, 由本社负责调换

质量监督电话: 010-82685218 010-58572162 010-58572393

目录

第一章 绪论 / 1

- 第一节 统计的产生与发展 / 3
- 第二节 统计学的研究对象和特点 / 8
- 第三节 统计研究的基本方法 / 15
- 第四节 统计学中的基本概念 / 19

第二章 统计调查与统计整理 / 28

- 第一节 统计调查的意义和种类 / 30
- 第二节 调查方案的设计 / 36
- 第三节 统计调查的组织形式 / 39
- 第四节 统计整理的意义与步骤 / 48
- 第五节 统计分组与分配数列 / 52
- 第六节 统计表与统计图 / 63
- 第七节 Excel 在统计数据处理中的应用 / 69

第三章 综合指标 / 84

- 第一节 总量指标 / 85
- 第二节 相对指标 / 91
- 第三节 平均指标 / 101
- 第四节 标志变异指标 / 115
- 第五节 Excel 在统计指标计算中的应用 / 123

第四章 抽样推断 / 135

- 第一节 抽样推断中几个基本概念 / 137
- 第二节 抽样误差 / 141

- 第三节 抽样估计方法 / 147
- 第四节 假设检验 / 153
- 第五节 抽样的组织形式 / 157
- 第六节 Excel 在抽样推断中的应用 / 160
- 第七节 Excel 在假设检验中的应用 / 164

第五章 相关分析与回归分析 / 172

- 第一节 相关关系的意义和种类 / 173
- 第二节 相关分析 / 178
- 第三节 回归分析 / 182
- 第四节 Excel 在相关与回归分析中的应用 / 188

第六章 时间数列分析 / 200

- 第一节 时间数列概述 / 201
- 第二节 时间数列的水平分析 / 205
- 第三节 时间数列的速度分析 / 213
- 第四节 长期趋势分析 / 217
- 第五节 季节变动分析 / 224
- 第六节 Excel 在时间数列分析中的应用 / 228

第七章 统计指数 / 240

- 第一节 统计指数概述 / 242
- 第二节 综合指数的编制和应用 / 246
- 第三节 平均指数的编制和应用 / 253
- 第四节 指数体系与因素分析 / 258
- 第五节 几种常见的经济指数 / 265
- 第六节 Excel 在指数分析中的应用 / 272

附录 / 278

参考文献 / 282

第二章

统计调查与统计整理

教学目标

【知识目标】

1. 理解统计调查在整个统计工作中的地位和作用。
2. 掌握搜集统计资料的方法和技术。
3. 掌握统计调查方案的基本框架和内容。
4. 理解各个统计调查组织方式的特点和应用范围。
5. 了解统计整理的概念和内容。
6. 掌握统计分组的方法并在分组的基础上进行次数分配数列的编制。
7. 学会用统计表来表示统计资料。
8. 熟悉 Excel 在编制统计表、统计图等方面的应用。

【能力目标】

1. 能针对具体问题正确选择统计调查方法。
2. 调查问卷初步设计及应用。
3. 掌握统计调查方案设计，并指导实践。
4. 能正确运用统计调查技术获取统计信息。
5. 能对调查资料按调查目的进行科学整理。
6. 能根据研究目的对统计资料进行有效分组，形成反映现象特征的分配数列。
7. 能根据某一实际问题设计统计调查问卷。
8. 锻炼在统计数据处理中灵活使用 Excel 的能力。

【素养目标】

1. 养成科学严谨的工作作风。
2. 具备爱岗敬业的职业素养。



关于亚太地区知名商学院的有关情况的案例分析

3. 具备初步的辩证思维能力。
4. 养成认真负责的工作态度。
5. 实事求是、公正无私。
6. 态度谦逊平和，平易近人，为对方着想。
7. 要树立起终身学习的观念。

案例导入

企业员工基本情况研究分析

某企业人事部门为了制订各类管理、技术人员的职业发展规划和调整员工的薪酬待遇，造就一支业务精干的高素质、高境界、具有高度凝聚力和团队精神的人才队伍，并形成以绩效考核为核心导向的人才管理机制，对全体员工做了全面调查。其中，利用调查表取得以下统计数据，如表 2-1 所示。

表 2-1 某企业员工基本情况研究分析

员工编号	姓名	性别	年龄(岁)	工龄(年)	文化程度	专业技术职称	职务或岗位	月收入(元)
001	丁志军	男	48	24	本科	经济师	部门经理	8 600
002	万珊珊	女	28	2	硕士研究生	助理工程师	职员	7 420
003	王优丽	女	33	5	硕士研究生	会计师	职员	7 860
004	方莉莉	女	25	3	大专	无	职员	6 180
005	孙伟	男	38	14	本科	工程师	职员	8 200
006	马小可	女	40	16	大专	无	职员	7 800
007	毛晓波	男	46	20	博士研究生	高级工程师	副总经理	12 260
008	吕央	女	33	10	大专	无	职员	7 200
009	宋万其	女	29	4	本科	无	职员	6 600
010	姚婕	女	31	7	本科	经济师	职员	7 600
011	白刚	男	55	30	博士研究生	高级工程师	总经理	12 600
012	洪生	男	37	12	本科	工程师	部门经理	8 600
013	叶俊雅	女	33	8	硕士研究生	工程师	职员	7 340
014	叶美娟	女	44	20	本科	会计师	部门经理	8 260
015	朱根发	男	32	10	大专	无	职员	7 520
016	何晴	女	27	4	大专	统计员	职员	7 520
017	潘晴	女	30	6	本科	经济师	职员	6 120
018	钟东	男	38	13	大专	会计师	职员	6 180
019	钟丽	女	42	16	硕士研究生	工程师	业务经理	8 500
020	蔡俊	男	35	10	本科	经济师	职员	6 120

如果你是这个企业的管理人员，现在需要对这些数据进行分组整理并用统计图表予以展示，以便分析揭示企业员工的性别构成、年龄结构、文化素质、专业技术构成以及各变量之间的内在关系，应从何入手？



一、统计调查的意义和要求

统计调查就是根据统计研究的目的、任务和要求,运用科学的调查方法,有计划、有组织地搜集客观事物原始资料的过程。统计调查是调查活动的一种,是统计工作的基础环节,是正确认识社会的前提,是制定方针政策的依据。

统计调查有别于一般社会调查,其主要特征是:统计调查是对社会经济现象总体中全部或足够多的单位进行调查,搜集大量以数字为主的信息资料,借以反映总体的数量特征。统计调查离不开统计数据,统计数据的来源渠道很多,不同的统计数据可通过不同渠道获得。我们把来源于直接反映被调查者的个体原始资料(又称为初级资料)和科学实验的统计数据称为第一手统计数据;把来源于别人调查和科学实验,或已经加工、整理出来的,说明现象总体的第二手资料(又称为次级资料)的数据称为第二手统计数据。统计调查的基本任务是搜集社会经济现象的原始资料(包括次级资料)。

统计调查在统计工作的整个过程中,担负着提供基础资料的任务,所有的统计计算和统计研究都是在原始资料搜集的基础上建立起来的。因此,统计调查是统计工作的基础环节,也是决定着整个统计工作质量好坏的重要环节,又是统计整理和分析的前提。如果统计调查工作做得好,能准确、及时、全面、系统地搜集到丰富的统计资料,那么就为统计整理和分析打好了一个坚实的基础,从而有利于正确认识被研究现象的本质及其规律性,还可以检查和监督政策的贯彻执行情况,为制定政策提供依据,并据此预测未来;反之,若调查工作做得不好,所得到的资料不完整、不真实或不及时,即使是经过科学的整理、严谨周密的分析,也是没有意义的。因此,可以说统计工作的质量,在相当程度上取决于统计调查所搜集到的原始资料的质量。

为了更好地完成统计工作的任务,发挥统计调查的作用,在统计调查过程中必须达到如下基本要求。

(一) 准确性

准确性是指搜集的资料要符合客观实际情况,保证各项统计资料真实可靠。这是保证统计资料质量的首要环节,是统计工作的生命。因此,在统计调查工作中要求所有统计调查人员和社会各界必须按照《统计法》及统计制度的规定,搜集或提供真实准确的数据资料。

请思考

统计调查资料的准确性是统计工作的生命线。你是如何理解的?

（二）及时性

及时性是指统计资料的时效性，要求及时上报各项统计资料以满足各种需要。如果统计资料提供的不及时，失去时效的资料反映不了实际的情况，犹如“雨后送伞”起不到统计应有的作用。值得注意的是：统计调查的及时性和准确性是辩证统一的关系，不能顾此失彼。及时性是在准确性的前提下才有意义的，但也不能强调准确性而不顾及时性要求，时过境迁的统计资料也是没有多大意义的。

（三）完整性

完整性也即全面性，是指按统计调查方案所规定的项目和内容，将所涉及的资料毫无遗漏地搜集起来。片面、零碎的资料不可能反映所研究对象的全貌和社会经济现象总体的特征，正如“盲人摸象”的故事所描绘的那样，最终也就难以对社会经济现象的规律性做出准确的判断，甚至会得出错误的结论。

（四）系统性

系统性是指搜集的统计资料要符合事物的逻辑，不能杂乱无章，所提供的统计资料，应该是便于整理、便于汇总的资料。

二、统计调查的种类

统计研究对象的复杂性和统计研究目的的针对性，决定了统计调查方法的多样性，根据不同的调查对象和调查目的，有必要采取不同的统计调查方式与方法，而不同的统计调查方法，又具有不同的特点和作用，因此，应该从不同的角度将众多统计调查方法进行分类。

（一）全面调查和非全面调查

统计调查按调查对象包括的范围不同，分为全面调查和非全面调查，如图 2-1 所示。

全面调查，是对构成调查对象的所有总体单位，一一进行调查登记的一种调查方法。全面统计报表和普查，都是全面调查。例如，为了研究我国人口数量、性别比例、不同民族的人口构成、年龄结构、地区结构、受教育程度、收入水平等问题而进行的第六次人口普查，就属于全面调查。全面调查能够掌握比较全面的、完整的统计资料，了解总体单位的全貌，但它需要花费较多的人力、物力和财力，操作难度较大，一般应慎重进行。

非全面调查，是对被研究对象中的一部分单位进行调查登记的一种调查方法。重点调查、抽样调查、典型调查及非全面统计报表等均属于非全面调查。例如，为

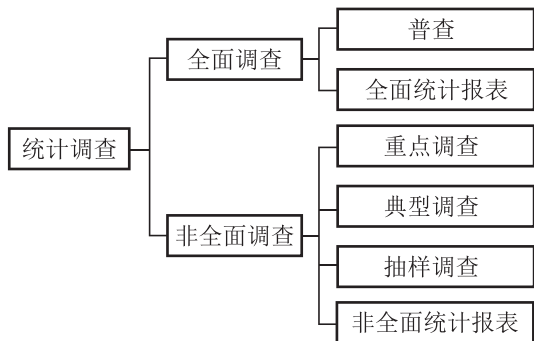


图 2-1 统计调查的种类图

了解某城市居民的收入水平,并不需要对该城市的全部居民进行调查,只需搜集该城市中各个收入层次的一部分居民的有关资料,就可以得出该城市居民的生活水平了。再如,要掌握某批大量生产的产品的质量,也不需要对此批产品进行逐一的质量检验,只需抽出必要的一部分进行检验即可。非全面调查的调查单位少,可以用较少的时间和人力,调查较多的内容,并能推算和说明全面情况,收到事半功倍的效果。其缺点是掌握的资料不够齐全,调查结果有时不够准确。

(二) 经常性调查和一次性调查

统计调查按调查登记的时间是否连续,分为经常性调查和一次性调查。

经常性调查,是指随着调查对象的发展变化,连续不断地进行调查登记的方法。通过经常性调查可以了解事物在一定时期内发生、发展的全部过程。例如,商业企业每天对销售量、销售额的登记,单位每天对员工的出勤进行统计,气象部门每天对气象数据的测量等,都属于经常性调查。经常性调查并不要求每天进行调查,有些现象定期或不定期地登记(时间间隔不超过一年)也属于经常性调查。经常性调查都属于定期调查。

一次性调查,是指间隔一定时间,一般是相当长一段时间进行一次的统计调查方法。该调查方式可用以了解经济现象在一定时点上的状态。一次性调查可以定期进行,如企业每年年末都要统计其资金占用额,这属于定期调查,一次性调查也可以不定期进行。我国进行的六次人口普查,就属于不定期的一次性调查。

(三) 定期统计报表制度和专门调查

统计调查按调查的组织形式分类,分为定期统计报表制度和专门调查。

定期统计报表制度是一种按国家有关法规的规定,自上而下统一布置、自下而上提供统计资料的一种统计调查方法。统计报表制度属于经常性调查。它以一定的原始资料为基础,按照统一的表式、统一的指标、统一的报送时间和报送程序进行填报,是一种严格的报告制度。统计报表在我国的统计工作中占有重要地位,它是我国统计部门和各业务管理部门获取全面而系统的统计资料的一种重要方式,是一种基层单位和各级组织向上级和国家报告工作情况的报告制度。统计报表主要以定期报表为主,如工业、商业、交通等部门的统计报表。

专门调查,是为了研究某些专门问题,由进行调查的单位专门组织的登记和调查。专门调查属于一次性调查,包括普查、重点调查、抽样调查和典型调查。专门调查灵活多样、适应性强,既可针对某专项内容进行,又可弥补统计报表的不足。

(四) 直接观察法、报告法、采访法和通信法

统计调查按搜集资料方式不同,分为直接观察法、报告法、采访法、通信法。

1. 直接观察法

直接观察法,是指调查人员亲自到调查现场对调查对象进行观察和计量,并加

以登记,以取得原始资料的一种调查方法。例如,为准确、及时地了解农作物的收获量,调查人员亲自深入田间、地块参加农作物的收割和称量;为了解工业产品质量情况,调查人员深入到车间亲自进行产品质量的观察和现场检测等。这种方法能够得到统计资料的高度准确性,但需要花费大量的人力、物力和时间,而且对历史资料的搜集来说,不可能采用此方法。

2. 报告法

报告法,是指报告单位以原始记录和核算资料为基础,按照隶属系统逐级向上级单位提供统计资料的一种方法。我国现行的统计报表制度就是采用报告法搜集资料并逐级上报的。这种调查方法是各地方、各部门、各单位按照《中华人民共和国统计法》的规定,必须向国家履行的一种义务。报告法由于采用统一的项目、统一的表式、统一的上报时间、统一的上报要求、统一的上报程序,适宜进行大量单位的调查。这种方法可以促进被调查单位建立、健全原始资料的登记和审核的工作制度。缺点是当被调查单位的自身利益受到影响时,可能出现虚报和瞒报的现象,且对资料的了解不够生动、具体。

3. 采访法

采访法,是指由调查人员以调查表和有关资料为依据,逐项向被调查者询问有关情况,并将答案记录下来的一种调查方法。它又分为个别询问法、开调查会法和被调查者自填法。

个别询问法,是由调查人员根据对被调查者逐一询问的情况来取得资料的一种调查方法。这种方法由于调查人员对调查项目有统一、正确的理解,能够向被调查者逐项询问,因而取得的资料准确、可靠。其缺点是需要花费较多的时间和人力。

开调查会法,是由调查人员根据调查题目,邀请与调查问题有关的专家或有经验的人员进行座谈讨论来搜集资料的一种调查方法。这种方法的优点是,在座谈会的环境中,参加人员之间相互影响、相互启发、相互补充,并在座谈的过程中不断地修正自己的观点,从而有利于取得较为广泛的、深入的意见和想法,可以取得比较正确的资料和信息。

被调查者自填法,是由调查人员将调查表送交被调查者,说明填表的要求和方法,由被调查者填写,然后由调查人员审核收回来取得资料的一种调查方法。这种调查方法比较节省人力和时间,其缺点是要求被调查者具有一定的文化水平和较强的责任心,否则难以保证资料的准确性。

4. 通信法

通信法又分为电话调查法、邮寄调查法和网络调查法三种。

电话调查法,实际上是一种间接采访法。它是调查人员利用电话同被调查者进行语言交流,从而取得资料和信息的一种调查方法。电话调查的时效快、费用低,随着电话在我国的普及,电话调查的应用也越来越广泛。电话调查可以事先按照设计好的问题进行,也可以针对某一专门问题进行。应当注意的是,电话调查的问题

要明确，问题数量不宜太多。

邮寄调查法，是通过邮寄或其他方式将调查问卷、调查表送至被调查者手中，由被调查者填写，然后将调查表或调查问卷寄回或投放到指定搜集点的一种调查方法。这种调查方法，调查范围广、节省人力财力，但时效性较差，回收率也不稳定，使得样本对涉及总体的代表性难以预先准确控制。

网络调查法，是传统调查方式在新的信息传播媒体上的应用。它是由调查人员将调查表或问卷通过互联网送至被调查者手中，由被调查者在网上自主选择填写作答来取得资料的一种调查方法。随着计算机和互联网技术应用的普及，网络调查的应用也会越来越广泛。

互联网作为一种信息沟通渠道，具有开放性、自由性、平等性、广泛性和直接性的特点。由于这些特点，网络调查具有传统调查所不可比拟的优势，具有成本低、速度快、隐匿性好（可以保证资料具有较强的真实性）和互动性强等特点。

虽然网络调查与传统调查方式相比具有较明显的优势，但由于我国目前的网络管理还处于探索阶段，这就使得网络调查的应用技术不尽成熟，仍然存在调查的客观性与识别资料的真实性、可靠性的问题。随着网络管理及网络应用技术的日渐完善和成熟，这些问题会逐步得到解决。

三、统计调查问卷

调查问卷是调查者根据调查目的和要求设计的，由一系列问题、备选答案、说明以及代码表组成的一种书面文件。它在统计调查中扮演着非常重要的角色，绝大多数统计数据的搜集都是依赖于调查问卷而完成的。一份完整的调查问卷通常包括如下几个部分。

（一）开头部分

开头部分一般具有以下五项功能。

（1）对调研人员的验证表明调研人员的身份。表明身份不仅仅是出于礼貌，更希望能尽快使调查活动进入状态，消除被调查者的疑惑。

（2）说明调查目的。调查目的应该清晰简洁地表达出来，应答者并不会关注调研者的调研目标，但他们会就提问的目的进行推测，这种推测可能会干扰他们的真实回答，所以要说明调查目的。

（3）解释如何选择调研对象。要回答被访问者的问题：“为什么是我？”往往只需要告诉他们，我们是随机选样的，如果选择不是随机的，你应该告诉他们所采用的选择方法。

（4）鼓励激发参与兴趣。有很多方式可以帮助做到这一点，给予一定报酬或是赠送一些纪念品是一种常见方式，另一种方法是让他们了解他们的参与将是十分重要的。

（5）填写说明。在自填式问卷中应有详细的填写说明，让被调查者知道如何填写问卷，如何将问卷返回到调查者手中，这部分内容可以集中放在问卷的前面，也

可以分散在各有关问题的前面。

（二）主体部分

1. 问题形式

（1）开放式问题。

开放式问题又称无结构问题的问答题。例如，在网上购物进行问卷调查时，可以问一些开放性问题：你在网上购物多久了？你喜欢的购物网站是什么？

（2）封闭式问题。

封闭式问题又称有结构的问答题。例如，你购买××咖啡的主要原因是什么？（选择最主要的两种）

A. 味道不错 B. 价格适中 C. 任何商店都有出售 D. 朋友介绍

2. 问题设计

（1）事实性问题。

事实性问题主要是要求应答者回答一些有关事实的问题。例如，你有网上购物的经验吗？设计事实性的主要目的在于求取事实资料，因此问题中的字眼定义必须清楚，让应答者了解后能正确回答。

（2）意见性问题。

在问卷中往往会询问应答者一些有关意见或态度的问题。例如，你是否喜欢××电视节目？意见性问题实际上是态度调查问题。

（3）困窘性问题。

困窘性问题是指应答者不愿意在调查者面前作答的某些问题，如关于私人的问题，或不为一般社会道德所接纳的行为、态度，或属于有碍声誉的问题。例如：

- ① 平均来说，你每个月打几次麻将？
- ② 如果你的汽车是分期购买的，一共分多少期？
- ③ 你是否想在银行抵押贷款购买股票？
- ④ 除了你的工作收入外，还有其他收入吗？

如果一定想要获得困窘性问题的答案，又避免应答者做不真实的回答，可采用以下方法：

间接问题法。不直接询问应答者对事项的观点，而改问他对其他有关事项的看法如何。例如，间接问题法旨在让应答者回答对旁人的观点进行评价的问题。所以介绍了旁人的观点后，应立即再加上问题：“你同意他们的看法吗？”

卡片整理法。将困窘性问题的答案分为“是”与“否”两类，调查员可暂时走开，让应答者自己取卡片投入箱中，以避免困窘气氛。应答者在没有被调查者注视的情况下，选取真实的答案的可能性会提高不少。

（4）假设性问题。

有许多问题是先假定一种情况，然后询问应答者在该种情况下，他会采取什么

行动。例如:

- ① 如果《××晚报》涨价至2元,你是否将转而购买另一种未涨价的晚报?
- ② 如果××牌洗衣粉跌价1元,你是否愿意用它?
- ③ 你是否愿意加薪?
- ④ 你是否赞成公交公司改善服务?

以上皆属于假设性问题,应答者对这种问题的回答多数为“是”。这种探测应答者未来行为的问题,应答者的答案事实上是没有多大的意义的,因为多数人都愿意尝试一种新事物,或获得一些新经验。

3. 问题顺序(表2-2)

表 2-2 问题顺序

问题类型	问题位置	例句	理论
过滤	最早提出	“过去一个月中您是否在家世界买过东西?” “这是您第一次来这里吗?”	用于选择符合要求的被访者
热身	过滤之后	“您经常购物吗?” “您通常在一周的哪几天购物?”	问题较为简单、有趣、轻松
过渡	在主要问题前或变换一下提问方式	“以下几个问题中,我要问一些有关您在家中看电视习惯的问题。” “下面我会读一些句子,要求您在我读完每一句后告诉我同意或不同意该种说法。”	将被访者引入正题,或者提醒他接下来要变换一下提问方式
较复杂或难以回答的问题	一般在中间,接近结束	“请按您喜欢的程度将下列这些品牌从1~7排列。” “在未来三个月中您可能会花多少时间来做什么事情?”	此时被访者已同意完成调研,并且所剩问题已经不多了
分类和人文统计	最后	“您的最高学历是什么?” “您是否在单亲家庭中长大?”	一些关于“个人”的问题,应将这些可能令人不快的话题放在最后

(三) 背景资料部分

背景资料部分包括被访者的基本资料,如性别、年龄、民族、文化程度、婚姻状况、职业、家庭人数、收入、单位等。

(四) 结束语

结束语的任务是要告诉被访者访问完毕,表示对被访者的谢意。

第二节

调查方案的设计

在统计调查工作正式开始之前,首先应当明确为什么要进行调查?向谁调查?在何时、何地调查?调查的内容是什么?怎样进行?也就是说应当事先设计一个切

实可行、周密细致的数据搜集方案，它是调查工作的依据，是保证调查顺利进行的前提。数据搜集方案又称调查方案，它是指导整个调查过程的纲领性文件，包括确定调查目的，确定调查对象和调查单位，确定调查项目，确定调查时间、调查期限，确定调查地点和调查方法，拟订调查的组织实施计划等。

一、确定调查目的与任务

确定调查目的与任务，是制定统计调查方案的首要问题。所谓调查目的与任务，就是指为什么要进行调查，调查要解决什么问题。只有调查目的与任务确定后，才能据此确定调查对象、调查单位和应采用的调查方式方法，才能做到有的放矢，节约人力，缩短调查时间，提高调查资料的时效性。例如，根据国家统计局《第六次全国人口普查方案》，第六次全国人口普查的目的是查清 2000 年以来我国人口数量、结构、分布和居住环境等方面的变化情况，按照“全国统一领导、部门分工协作、地方分级负责、各方共同参与”的原则组织实施。

二、确定调查对象和调查单位

调查对象是指根据调查目的和任务，需要进行调查的社会经济现象的总体，也就是说统计总体在统计调查阶段称为调查对象。它是由性质相同的许多调查单位所构成的。确定调查对象，首先需要根据调查目的，对研究对象进行认真分析，要明确规定调查对象总体的范围和界限，这样才能避免登记的重复或遗漏，保证统计资料的准确。

调查单位是指调查对象中所要调查的具体单位，说明由谁来提供资料的问题。它是指统计总体中的每一个个体（即总体单位），是调查项目和标志的承担者和载体，是在调查过程中应该登记各种标志的那些具体单位。“调查单位”这个词的含义在这里是指计量事物的基本量，不能理解为从事调查工作的“部门”。调查单位的确定取决于调查目的和调查对象。

例如，调查目的是为了搜集广东省国有工业企业的生产情况的资料时，调查对象是广东省所有的国有工业企业。产值、利税、职工人数等调查项目的承担者是广东省的每一个国有工业企业，所以调查单位就是广东省的每一个国有工业企业。又如，当调查的目的是为了了解广东省国有工业企业的机械设备的状况时，则调查对象是该省国有工业企业的所有设备。该省国有工业企业的每一台机械设备都是设备类型、价值、使用时间等调查项目的承担者，调查单位就是该省国有工业企业的每一台机械设备。再如，要了解 2019 年应届大学毕业生的就业情况，那么当年所有的大学应届毕业生是调查对象，调查单位就是当年每一位大学应届毕业生。

报告单位也称填报单位，它是负责向上汇报调查内容、提交统计资料的单位。实际工作中，需要注意的是不要把调查单位和填报单位相混淆。调查单位是调查项目的承担者，而填报单位则是负责上报调查资料的单位。填报单位一般指在行政上、经济上具有一定独立性的单位，而调查单位既可以是人，也可以是一个部门（或机

构),还可以是物。根据调查目的的不同,调查单位与报告单位有时一致有时不一致。在前面的三个例子中:对国有工业企业调查,每个国有工业企业既是调查单位又是填报单位,两者是相同的;调查企业设备情况时,调查单位是企业的每一台设备,而填报单位则是企业,两者是不同的;毕业生就业情况调查,如果规定向每一位毕业生调查,那么调查单位和报告单位都是每一位毕业生,如果规定向各院校调查,则调查单位是每一位毕业生,而报告单位是学校。

三、确定调查项目和调查表

调查项目是所要调查的具体内容,也就是调查单位所承担的基本标志。它可以是调查单位的数量特征,也可以是调查单位的某种属性或品质特性。因为统计调查的内容、范围、目的和要求都是通过调查项目来反映的,所以确定调查项目要解决向被调查者调查什么,需要被调查者回答什么的问题。调查项目应根据调查目的和调查单位的特点而定。

例如,2020年第七次全国人口普查根据调查目的拟定了姓名、公民身份证号码、性别、年龄、民族、受教育程度、行业、职业、迁移流动、婚姻生育、死亡、住房情况等多个记录调查项目。

在确定调查项目时须注意:调查项目应该是为实现调查目的所需要的项目,要少而精;各调查项目内容含义要明确、具体,必须是可行的,能够取得确切资料的项目,有些虽属需要,但实际上没有条件取得实际资料的项目,不要列入;调查项目要注意彼此衔接,避免重复和相互矛盾。各调查项目之间尽可能做到相互联系,彼此衔接,以便于相互核对和印证。

调查项目确定之后,接下来要考虑怎样得到这些项目的具体资料,那就要制定调查表了。调查表又称问卷(或询问表),是调查项目按照一定的顺序排列起来形成的一定的表式,是系统地记载调查内容的一种印件。调查表是搜集原始资料常用的基本工具,它有单一表和一览表两种形式。

在一张表上只登记一个调查单位资料的叫单一表,它可以容纳较多的项目(标志),便于整理和分类。在一张表上可以登记若干个调查单位资料的叫一览表,它便于合计和核对差错,但它容纳的调查项目很少。使用单一表还是一览表,应根据项目的多少,以及调查单位是否集中来确定。

四、确定调查时间、调查期限和调查地点

确定调查时间、期限和地点,是要解决在何时、何地调查的问题。

调查时间,是指调查资料所属的时间。从资料的性质来看,有的资料反映现象在某一时刻上的状态(时点现象),而有的资料反映现象在一段时期内发展过程的结果(时期现象)。如果所要调查的是时期现象,即调查的资料是某一时期内的总量,调查时间就是资料的起止的时间范围;如果所要调查的是时点现象,即调查的资料在某一时刻的水平,调查时间就是规定统一的标准时间。明确规定调查的时期或时点,

是保证调查资料的准确性、可比性及可汇总的重要前提条件。

调查期限即整个调查工作的时限，包括搜集资料及报送资料的整个工作过程所需要的时间。如某企业进行职工人数调查，从9月1日起进行登记，至9月20日结束，则9月1日就是调查的标准时点，9月1日至20日的20天就是调查期限。为了保证资料的及时性，对调查期限的规定，要在保证资料准确性的前提下尽可能缩短。

调查地点，是指调查对象所在的地点，也就是统计资料所属的空间范围。明确统计资料的空间界限，也是为了防止统计资料发生错漏。

请思考

我国第六次人口普查规定标准时间为2020年11月1日0时，资料呈报时间为12月10日，则调查时间为（ ），调查期限为（ ）。

五、拟订调查的组织实施计划

为了保证整个统计调查工作的顺利进行，在调查方案中还应该有一个周密的组织实施计划，使调查工作的进行有组织上、措施上强有力的保证。它包括：明确调查机构；实施调查的步骤；调查人员的选择、组织和培训；调查表格、问卷等必备工具的准备等；资料报送方法、检验方法；调查经费来源和开支预算等。

调查机构是进行调查并对该项调查工作负责的组织或部门。有常设机构，如国家、省（市）、市、县统计局，业务主管部门的统计机构，各社会团体等。也有临时性的非常设机构，如人口普查办公室、工业普查领导小组等。在实施计划中要明确各调查机构的职责和彼此之间的关系。

调查组织工作计划要明确实施调查的步骤，对于调查进行的每一步骤和每一时间段应该完成的工作做出具体的部署。除了调查人员的选择、组织和培训，印制调查表格、问卷等必备工具的准备外，还要对资料报送方法、检验方法、调查经费来源和开支预算等均应作具体的规定。

选择恰当的调查方法和合适的组织形式，也是实施统计调查的重要问题。

根据统计调查的目的和任务的要求，以及调查对象的特点，选择一种合适的调查方式，然后根据调查资料的全面性、可靠性的要求确定采取哪一种调查方法。

第三节

统计调查的组织形式



统计调查的组织形式

一、普查

（一）普查的含义

普查是为了某一特定的目的专门组织的一次性全面调查。一般用来调查属于一

定时点状态的重要社会经济现象,如人口普查、经济普查等。有些客观现象不需要或不可能进行经常性调查,但需要掌握它的准确情况,这时就可以采用普查的方式搜集资料。普查可以取得某些社会经济现象不宜或不需要通过统计报表去搜集的较为准确的全面统计资料,以搞清一个国家主要的国情国力,作为制定重要政策和长期发展规划的依据。普查是一种重要的调查方式,世界各国在进行本国的国情国力调查时,都采用普查的方式来完成。

(二) 普查的特点

1. 一次性

普查一般用来搜集属于一定时点的现象的总量,这些时点现象的数量在短期内往往变动不大,不必做连续登记,只需间隔一段较长时间进行一次性的调查。而普查的规模大、指标多、任务重,耗费大量人力、物力和时间,不可能进行经常性调查,只能采用一次性调查。例如,我国是一个人口大国,我国的人口普查工作,不可能年年组织,更不可能月月组织。我国第七次人口普查与第六次之间就间隔了10年时间。

2. 时点性

普查的对象主要是时点现象。每次普查都有标准时点。例如,2020年全国第七次人口普查的标准时点就确定为2020年11月1日0时。当然,普查也不排斥搜集某些时期现象的资料。

3. 全面性

普查对象范围广,调查内容详细,比其他任何方式的调查更能掌握全面、详尽的统计资料,具有重要的分析价值。例如,我国在2004年进行的全国第一次经济普查仅不同种类的普查表就有近50种,涉及的调查指标近千个。内容包括了单位基本属性、从业人员、财务状况、生产经营情况、生产能力、原材料和能源消耗、科技活动情况等,为国家制定国民经济和社会发展规划,为各级部门提高决策和管理水平提供了重要的信息。

(三) 普查的组织形式

普查的组织形式有两种:一种是自上而下成立专门普查机构,并由这个机构组织普查队伍对调查单位进行直接登记,例如人口普查等;另一种是在各单位的会计统计和业务核算资料、报表资料的基础上,结合实际盘点和实际观察进行调查登记,自下而上由被调查单位自行填报调查表格,并逐级上报来实施普查的形式,例如我国的物资库存普查和牲畜普查,便是后一种方式。这种方式也仍需要组织普查的领导机构,配备专门人员对普查工作进行组织领导。

(四) 普查的实施原则

普查因涉及面广,工作量大,需要动员大量的人力、物力和财力,所以普查不宜经常进行。只有在研究对于国民经济和社会发展具有重大决定意义的问题时,才

有可能和必要组织普查。

普查的组织原则如下。

(1) 规定统一的普查标准时间。

即普查资料的所属时间,以避免搜集资料的重复或遗漏。例如我国第六次全国人口普查规定的标准时间是2010年11月1日0时。

(2) 确定统一的普查期限。

在普查范围内的各调查单位应尽可能同时进行调查,并尽可能在最短期内完成,以便在方法上、步调上取得一致,以保证调查资料的真实性和及时性。

(3) 统一规定普查的项目和指标。

普查项目一经统一规定,就不能任意改变或增减,以免影响汇总综合,降低调查资料的质量。

(4) 同类普查应尽量按照一定的周期进行。

普查可以不定期进行,但某些重要的普查,应尽可能按照一定的周期进行,这样便于对历次调查资料进行动态对比,也便于尽早做好普查的各项准备工作。例如,我国对目前法定的几项全国性的普查的规定是:人口普查和农业普查均每十年进行一次,调查的标准时间分别为逢0和6的年份;2004年开展了第一次全国经济普查,它是以前的工业普查、第三产业普查和基本单位普查的合并,并且纳入了建筑业普查,以后经济普查每五年一次,调查的标准时间为逢3和8的年份。

请思考

普查为什么要规定标准时间?

思政之窗

国务院第七次全国人口普查领导小组办公室负责人就《全国人口普查条例》 施行十周年答记者问(节选)

2010年6月1日《全国人口普查条例》正式实施,2020年是施行十周年,当前第七次全国人口普查准备工作正在各地如火如荼进行中。十年来,《全国人口普查条例》施行情况如何,在保障人口普查顺利进行方面发挥了怎样的作用?日前,就《全国人口普查条例》施行十周年相关问题,国务院第七次全国人口普查领导小组办公室负责人接受了《中国信息报》记者专访。

记者:2020年是《全国人口普查条例》施行十周年,请问在人口普查实施过程中,《全国人口普查条例》发挥了怎样的作用?

答:人口普查是重大的国情国力调查,是和平时期最大的社会动员,涉及每一个人、

每一个家庭以及社会的各个方面,需要社会各界及公众的理解、支持与配合。2020年在疫情防控常态化下开展人口普查,更需要调查对象和调查人员共同努力。同时,人口普查工作也是一项庞大的社会系统工程,需要强有力的领导和严密的组织,普查机构的设立、普查人员的选调和培训、普查经费的安排、普查的宣传动员、普查数据的采集、汇总、处理和发布等,都需要有关部门和单位的积极参与和配合。随着我国社会主义市场经济体制的逐步完善,特别是依法行政观念逐步深入人心,公民对自身权益的保护意识也不断增强。

《全国人口普查条例》(以下简称《条例》)对普查目的、普查原则、普查任务、普查对象、普查范围、普查的组织实施等做出了明确规定。对于科学、有效地组织实施全国人口普查,依法保障人口普查的有序开展,维护人口普查对象的合法权益,赢得普查对象的信任、支持与合作,保障人口普查数据的真实性、准确性、完整性、及时性,发挥了重要作用。

记者:《全国人口普查条例》是如何促进人口普查对象积极配合第七次全国人口普查的?

答:《条例》对人口普查的对象、内容、方法、组织实施、普查资料管理和公布、法律责任等做出明确规定,对普查对象、普查机构和普查人员在人口普查活动中的权利、义务以及违法行为所应承担的法律责任等做出具有法律效力的严格界定。

2020年实施的第七次全国人口普查将采取电子化方式开展普查登记,可以由普查员使用PAD或智能手机入户登记数据直接上报,也可以由普查对象通过互联网自主填报。同时,此次普查还将广泛应用部门行政记录,通过在普查指标设置中增加公民身份号码,实现与公安、卫健等部门行政记录的比对核查。为了充分保护人口普查对象的合法权益,赢得普查对象对普查工作的理解、支持和配合,《条例》在保护人口普查对象的权利方面做了严格规定:一是人口普查对象提供的资料,依法予以保密;二是人口普查中获得的能够识别或者推断单个普查对象身份的资料,任何单位和个人不得对外提供、泄露,不得作为对人口普查对象做出具体行政行为的依据,不得用于普查以外的目的;三是人口普查中获得的原始普查资料,按照国家有关规定保存、销毁;四是对泄露或者向他人提供能够识别或者推断单个普查对象身份的资料的行为,依法追究法律责任。

同时,对于人口普查对象需要履行的普查义务,《条例》规定:人口普查对象应当按照《中华人民共和国统计法》和本条例的规定,真实、准确、完整、及时地提供人口普查所需的资料。人口普查对象应当如实回答相关问题,不得隐瞒有关情况,不得提供虚假信息,不得拒绝或者阻碍人口普查工作。

记者:在人口普查中,人们会担心个人信息泄漏,《全国人口普查条例》如何规定人口普查中获得的个人信息必须严格保密?

答:刚才也提到,《条例》对保护人口普查对象的合法权益做了非常明确的规定。人口普查中获得的能够识别或者推断单个普查对象身份的资料,必须严格保密,任何

单位和个人不得对外提供、泄露。人口普查资料不得作为对普查对象做出具体行政行为的依据，也不得作为对地方人民政府进行政绩考核和责任追究的依据。

同时，《国务院关于开展第七次全国人口普查的通知》也明确提出，“全流程加强对公民个人信息的保护，各级普查机构及其工作人员必须严格履行保密义务，严禁向任何机构、单位、个人泄露或出售公民个人信息。”第七次全国人口普查将采取电子化方式开展普查登记，登记方式的改变，减少了数据搜集上报的中间环节，提高了普查数据质量，也加强了对个人信息安全的保障。

二、抽样调查

抽样调查是按照随机原则，从调查对象中抽取一部分单位作为样本进行观察，并根据其观察的结果来推断总体数量特征的一种非全面调查方法。抽样调查是实际中应用最广泛的一种调查方法，它虽然是非全面调查，但也可以取得反映总体情况的信息资料，因此，在一定意义上说，它可以起到全面调查的作用。

根据抽选样本的方法，抽样调查可以分为概率抽样和非概率抽样两类。概率抽样是按照概率论和数理统计的原理从调查研究的总体中，按照随机原则来抽选样本，并从数量上对总体的某些特征做出估计推断，对推断中可能出现的误差可以从概率意义上加以控制。在我国，习惯上将概率抽样称为随机抽样调查。非概率抽样也称非随机抽样，它是依人们的主观判断来选择样本单位。这种取样方式对于总体中的每一个个体被抽取的可能性无法用概率表示，其主要特点是简单方便，所以在实践中也有一定的应用。如：偶遇抽样、主观抽样、滚雪球抽样、配额抽样等都是非概率抽样。

关于抽样调查的详细内容，将在第四章介绍。

三、重点调查

重点调查是一种非全面调查，它是在调查对象中选择一部分对全局具有决定性作用的重点单位进行调查。所谓重点单位，是指在总体中具有举足轻重的地位的那些单位，这些单位的数量在总体中虽然不多、所占比重不大，但其标志总量在被研究总体的全部标志总量中占有很大比重。因此，对这些重点单位的标志进行调查，就可以在数量方面说明总体在该标志总量方面的基本情况。重点调查适用于调查任务只要求掌握调查总体的基本情况，调查标志比较单一，调查标志表现在数量上集中于少数单位，而这些少数单位的标志值之和在总体中又占绝对优势的情况。例如，要了解我国钢铁生产的基本情况，由于宝钢、马钢、鞍钢、武钢等几个大型钢铁公司的钢铁产量占了全国钢铁产量的大部分，对这些重点企业进行调查，便能省时省力而且及时地了解全国钢铁生产的基本情况，满足调查任务的要求。因此，重点调查可以节省大量的人力、物力和财力，并能使调查工作做得更加细致、及时。此外，重点调查的组织方式也较灵活，既可以组织专门调查，也可以运用统计报表形式进

行调查。

重点单位的选择要客观。选取重点单位，应遵循两个原则：一是要根据调查任务的要求和调查对象的基本情况而确定选取重点单位及数量，着眼于这些单位的标志值在总体标志总量中的比重，而不是这些单位在技术管理或其他方面是否有特定意义，更不要带有主观因素。一般来讲，要求重点单位应尽可能少，而其标志值在总体中所占的比重应尽可能大，以保证有足够的代表性；二是要注意选取那些管理比较健全、业务能力较强、统计工作基础较好的单位作为重点单位，这样才能准确、及时地取得资料。

选择重点单位时，要注意重点是可以变动的。一个单位在某一问题是重点，在另一问题上不一定是重点；在某一调查总体中是重点，在另一调查总体中不一定是重点；在这个时期是重点，在另一个时期不一定是重点。因此，对不同问题的重点调查，或同一问题不同时期的重点调查，要随着情况的变化而随时调整重点单位。

重点调查的主要特点是：调查单位少，可以调查较多的项目。由于投入少、调查速度快、所反映的主要情况或基本趋势比较准确，所以当调查任务只要求掌握总体的基本情况，而且总体中确实存在重点单位时，应该采用重点调查。但必须指出，由于重点单位与一般单位的差别较大，通常不能由重点调查的结果来推算整个调查对象的总体指标。

重点调查的组织方式有两种：一种是专门组织的一次性调查，适用于临时调查任务；另一种是利用定期统计报表经常性地对一些重点单位进行调查。其优点是花费较少人力、物力，在较少时间内及时取得有关的基本情况。定期调查适用于经济性调查任务，可以颁发定期报表，由选择的重点单位填报，定期观察一些重点单位的主要技术经济指标的完成情况及其变动。

请思考

如何理解重点调查中的“重点单位”？“重点单位”是客观存在的，这种说法对吗？

四、典型调查

典型调查是根据调查的目的与要求，在对被调查对象进行全面分析的基础上，有意识地选择若干具有典型意义或有代表性的单位进行调查的一种调查方式。典型调查也是一种专门组织的非全面调查，进行典型调查的主要目的不在于取得社会经济现象的总体数值，而在于了解与有关数字相关事物的具体情况，借以认识事物的本质特征、因果关系和发展变化的趋势。

就人类认识的顺序来说，总是遵循由特殊到一般的认识规律。因此，从被研究对象中选择有代表性的典型单位，对其进行深入细致的调查研究，就可以了解事物的本质及其发展过程，也就可以认识同类事物的本质和发展规律。“麻雀虽小，五脏俱全”，通过解剖“麻雀”，可以认识总体的内部构成、一般水平和发展变化规律。

典型调查的中心问题是如何正确选择典型单位。选择典型单位必须依据正确的理论进行全面的分析, 综合比较, 了解被研究总体的全面情况和一般水平, 然后从各个可供选择的单位中挑选富有代表性的典型单位。切忌主观片面性和随意性, 它不仅要求调查者有客观的、正确的态度, 而且要有科学的方法。根据不同的研究目的和要求, 有以下三种选典方法。

(1) “划类选典”的方法: 在统计调查工作中有时为了近似地估算总体的指标数值, 当总体内部差异明显时, 可以在了解总体概况的基础上, 按某种标志把总体划分为若干类型, 根据每一类型在总体中所占比例, 选出被调查的若干典型单位。

(2) “解剖麻雀”的方法: 这种选典方法适用于总体内各单位差别不太大的情况。通过对个别代表性单位的调查, 即可估计总体的一般情况。

(3) “抓两头”的方法: 挑选先进、落后单位或新生事物典型进行调查的方法。这是为了总结成功的经验, 找出失败的教训, 带动中间状态的单位, 推动整体的发展而进行的调查。

典型调查具有如下特点: ①典型调查是对调查单位中个别或某些单位进行的调查, 因此, 这些单位必须具有代表性, 研究人员一般要对调查范围内的所有调查单位进行初步了解, 在此基础上, 划分类型, 选取少量典型。②调查单位是根据调查目的有意识选择出来的, 这便于从典型入手, 逐步认识事物的一般性和普遍性。③调查范围小、调查单位少、灵活机动、具体深入、节省人力、财力和物力等。有利于提高调查效率, 探索事物发展变化的规律性。

典型调查有两种类型: 一种是一般的典型调查, 即对个别典型单位的调查研究。在这种典型调查中, 只需在总体中选出少数几个典型单位, 通过对这几个典型单位的调查研究, 用以说明事物的一般情况或事物发展的一般规律; 另一种是具有统计特征的划类选典的典型调查, 即将调查总体划分为若干个类, 再从每类中选择若干个典型进行调查, 以说明各类的情况。

典型调查的作用主要有以下两点。

(1) 典型调查可以对问题做具体深入的分析, 它可以具体地了解事物发生和发展的全过程及同各方面的联系, 特别是有利于研究新情况、新问题。

(2) 典型调查可以补充全面调查和其他非全面调查的不足。典型调查可以搜集全面调查或其他形式的非全面调查搜集不到的统计资料或情况, 因为调查的单位较少, 可以对典型单位进行深入调查, 从而把问题搞清楚。

以上介绍的这几种不同的统计调查方式, 各有其特点和作用。在实际的统计工作中, 往往需要多种统计调查方式结合运用。这不仅因为国民经济和社会发展情况复杂, 门类众多, 变化较快, 只有采用多种统计调查方式, 才能搜集到丰富的统计资料, 还因为任何一种统计调查方法, 都有其优越性与局限性以及各自不同的实施条件, 只用一种统计调查方式, 往往是不能满足多种需要的。

请思考

人们经常提到的“解剖麻雀”式的工作方法，属于哪种调查方式？

五、统计报表

统计报表是按照国家或上级部门统一规定的表格形式、统一规定的指标内容、统一规定的报送程序和报送时间，由填报单位自下而上逐级提供统计资料的一种统计调查方式。它是一种以全面调查为主的调查方式，是国家取得调查资料的一种重要的组织形式。它已形成一种制度即统计报表制度。执行统计报表制度，是各地区、各部门、各基层单位必须向国家履行的一种义务。它的任务是经常地、定期地搜集反映国民经济和社会发展基本情况的资料，为各级政府和有关部门制订国民经济和社会发展计划、检查计划执行情况服务。

统计报表制度的内容包括以下几个方面。

(1) 表式：统计报表的具体格式。基本上由三部分组成，即表头（报表标题、表号、填报单位、制表单位、计量单位等）、表身（具体填报的数据、资料）、表脚（备注，填表人、审核人签章等）。它是由国家统计局根据研究的任务与目的而专门设计制定的统计报表表格，用于搜集统计资料，是统计报表制度的主体。

(2) 填表说明：它是对统计报表的统计范围、指标等做出的规定，具体如下。

① 填报范围：即统计报表的范围，规定每种统计报表的报告单位和填报单位，各级统计部门与主管部门的范围等。

② 指标解释：对列入表的统计指标的口径、计算方法以及其他有关问题的具体说明。

③ 分类目标：对有关统计报表主栏中应进行填报的有关项目的分类。

④ 其他有关事项的规定。除了以上各项规定以外的一些注意事项，如：报送日期，报送方式，报送份数等。

统计报表的资料来源，主要是基层的原始记录、台账及基层的内部报表。

统计报表按其性质和要求不同，有如下几种类别。

(1) 按调查范围不同，统计报表可分为全面统计报表和非全面统计报表。

(2) 按报表内容和实施范围不同，分为国家统计报表、部门统计报表和地方统计报表。

(3) 按报送周期长短不同，分为日报、周报、旬报、月报、季报、半年报和年报。

(4) 按填报单位不同，分为基层统计报表和综合统计报表。

(5) 按报送方式不同，可分为电讯统计报表和邮递统计报表。

统计报表主要用于搜集全面的基本情况，此外，也常为重点调查等非全面调查所采用。统计报表具有以下特点。

(1) 由于统计报表的指标体系、表格形式、报送时间及报送程序都是按照国家

统计局规定实施的，保证了统计资料的统一性。

(2) 在统计报表的实施范围内被调查单位都进行填报，并且经过部门和各个地区的乃至全国的层层汇总，各个部门、各个地区及国家可以获得相应的统计资料。能够满足各个层次对统计资料的需求，保证了统计资料的全面性。

(3) 统计报表是按照一定周期进行报送的（如月报、季报），可以获得周期性的统计资料，保证了统计资料的动态性。

(4) 由于统计资料是建立在原始记录和统计台账的基础之上，可以获得较为准确的统计资料，保证了统计资料的可靠性。

统计是计划的基础，没有科学的统计，就不可能有真正的计划。我国国民经济和社会发展计划的编制离不开统计报表提供的统计数字。至于检查监督计划的执行情况，更需要通过统计报表，及时了解计划进程及存在的问题和问题发生的原因，以便采取有效措施，保证计划实现。

普查和全面统计报表都是全面调查，但统计报表属于经常性调查，报表的内容主要是经常调查的项目。而普查属于一次性调查，主要用于调查有关国情国力的重要资料在一定时点状态下的数量。有些社会现象不可能也不需要经常调查，但又需要掌握比较全面的统计资料，这就可以通过普查来解决。普查比一般调查规模要大，而且调查内容比较详细，可以得到完整的统计资料。而统计报表则不可能像普查那样掌握如此详尽的全面资料。

资料卡

统计报表的资料来源和填表说明

统计报表的资料来源于基层单位的原始记录和统计台账。原始记录是基层单位通过各种表格、票据、单证、卡片等形式，对生产经营活动所做的最初记载，如发票、材料入库单、产品出库单、固定资产卡片等。统计台账是基层单位根据填报统计报表和经营管理的需要，按时间顺序分门别类地登记原始记录的一种表册，它是原始记录到统计报表的中间环节，如产品产量台账、原材料领用台账等。

填表说明是为了准确、及时、全面、系统地填报统计报表而对相关事项做出的统一规定，包括对报表的填报范围、分组体系、指标解释、分类目录、报送日期、报送方式等所做的具体说明。

请思考

举例说明统计报表产生的过程。

一、统计整理的意义

通过统计调查可以获得第一手资料,但这种资料是分散的、零碎的、表面的,只能反映各单位的具体情况。要说明总体情况,揭示出总体的内在特征,还需要对这些资料进行加工整理,使之系统化,以便通过综合指标对总体做出概括性的说明。

统计整理,就是根据统计研究的目的,对所搜集到的初始数据进行审核、分组、汇总,使之系统化、条理化,变成能反映总体特征的综合数据的工作过程。统计整理既包括对统计调查所得到的原始资料进行整理,也包括对加工过的综合资料,即次级资料进行再整理。

统计整理在整个统计研究中占有十分重要的地位。它是整个统计工作和研究过程的中间环节,起着承前启后的作用。统计整理是统计调查的继续,又是统计分析的基础。统计整理的正确与否,将直接影响和决定着能否完成整个统计研究的任务。如果采用不科学、不完整的整理方法,即使搜集到准确、全面的统计资料,也往往会使这些资料失去应用价值。统计调查所搜集到的资料,只有通过科学的审核、分类、汇总等整理工作,才能使统计在认识社会的作用过程中,实现由个别到全体、由特殊到一般、由现象到本质、由感性到理性的转化,才能从整体上反映出事物的数量特征。否则统计调查所得的资料再丰富、再完备,其作用也发挥不出来,统计调查就会失去意义,统计分析也将无法进行。因此,必须十分重视统计整理工作。

统计整理从广义上讲,包括两种整理。第一种是对统计调查所搜集到的各种数据进行分类和汇总,称为汇总性整理。第二种是对现成的综合统计资料进行整理。本节所说的统计整理,指的是第一种整理。从狭义上来讲统计整理的类型有:定期统计报表数据的整理、专题性统计数据的整理、历史统计数据的整理。对已有的统计资料进行筛选,以及按历史的口径对现有的统计资料重新调整、分类和汇总等,都必须通过统计整理工作来完成。

统计整理要坚持以下三个原则。

(1) 准确性原则:整理后的资料,事实要清楚,数据要准确。如果整理出来的资料事实含混不清,数据相互矛盾,那么就不可能得出科学的结论。

(2) 完整性原则:整理资料应尽可能地全面、完整,既要有历史的资料又要有现实的资料;既要有正面的、肯定的资料,又要有负面的、否定的资料;既要有主体的、主流的资料,又要有辅助的、支流的数据等。只有这样才能真实地反映调查对象的全貌。

(3) 简明性原则:整理所得的资料应该尽可能地系统化、条理化,并以简明的方式反映调查对象的总体情况。如果整理出来的资料仍然凌乱、没有条理,使人们

对调查对象难以形成一个完整、清晰的概念，那么就会给以后的研究分析工作增加许多困难。

二、统计整理的步骤

统计整理的全过程包括对统计资料的审核、分组、汇总和编制统计图表、资料积累、存档等几个环节，需要按照一定的步骤进行。

第一步，设计和制定统计整理方案。

第二步，对原始资料进行审核。

第三步，对经过审核的资料进行分组并综合汇总，计算出总体总量指标。

第四步，将汇总计算的结果，以统计表或统计图的形式表现出来。

第五步，对统计资料妥善保存，系统积累。

以下分别对以上几个步骤进行阐述、说明。本节主要介绍资料的审核和汇总，对于统计分组和编制统计表是后面两节的内容，这里不再赘述。

（一）统计整理方案设计

统计整理方案是整理工作的指导性文件，体现了整理工作的具体规划和要求，是保证统计资料准确、及时、全面的重要前提。统计整理方案应包括对总体资料的处理方法和确定用哪些具体的统计指标来说明总体两方面的设计。

（二）统计资料的审核

在对统计数据进行整理时，首先要对搜集到的资料进行全面审核，以确保统计资料符合统计研究目的的要求，保证数据的质量，为进一步的整理与分析打下基础。从不同渠道取得的统计数据，其审核内容和方法有所不同。

对于通过直接调查取得的原始数据，应主要从完整性和准确性两个方面去审核。它包括以下内容。

1. 审核资料的完整性和及时性

审核资料的完整性，就是看调查单位或填报单位是否齐全，应调查的单位或个体是否有遗漏，所有的调查项目或指标是否填写齐全，规定的项目是否都有答案，应报资料的份数是否符合规定等。

审核资料的及时性，是看填报单位是否按时报送了有关资料。对不报、漏报或迟报的现象都要及时查清。

2. 审核资料的正确性

审核资料的正确性，是检查所填报的资料是否准确可靠。它主要包括两个方面：一是检查数据资料是否真实地反映了客观实际情况，内容是否符合实际；二是检查数据是否有错误，计算是否正确等。

常用的审核方法有两种。

(1) 逻辑检查。

逻辑检查主要用于对定类数据和定序数据的审核。它主要是从定性角度审核统计资料内容是否合理,各项目或数字之间有无相互矛盾的现象,有无不切实际或不符合逻辑的地方。比如一张调查表中,年龄是9岁,婚姻状况却是已婚,其中必有一个是错误的;小学文化程度的人所填的职业是大学教师,对于这些违背逻辑的项目应予以纠正。

(2) 计算检查。

计算检查是检查调查表中的各项数据在计算结果和计算方法上是否有误。比如各分项数字之和是否等于相应的合计数,各结构比例之和是否等于1或100%,出现在不同表格上的同一指标数值是否相同。还要检查各项指标的计算口径、计量单位是否符合规定,各指标间的数字是否相互衔接,等等。

对于通过其他渠道取得的第二手数据,除了对其完整性和准确性进行审核外,还应着重审核数据的适用性和时效性。第二手数据可以来自多种渠道,有些数据可能是为特定目的通过专门调查而取得的,或者是已经按特定目的的需要做了加工整理。对于使用者来说,首先应弄清楚数据的来源、数据的口径以及有关的背景材料,以便确定这些数据是否符合分析研究的需要,是否需要重新加工整理等,不能盲目地生搬硬套。在利用历史资料(或其他间接资料)时,应审核资料的可靠程度、指标含义、所属时间与空间范围、计算方法和分组条件与规定的要求是否一致。一般可以从调查资料的历史背景、调查者搜集资料的目的以及资料来源等,来判断资料的可靠程度。也可以从指标间的相互关系以及指标的变动趋势来检查它的正确性。对不能满足现在要求、缺漏或有疑问的资料,要进行有科学根据的推算、弥补和订正。

此外,还要对数据的时效性进行审核,有些时效性较强的问题,如果所取得的数据过于滞后,就失去了研究的意义。一般来说,应尽可能使用最新的统计数据。数据经过审核后,确认适合实际需要,才有必要做进一步的加工整理。

3. 资料审核后的订正

通过上述审核,如发现有缺报、缺份和缺项等情况,应及时催报、补报;如有不正确之处,则应分不同情况做如下处理。

(1) 对于可以肯定的一般错误,应及时代为更正,并通知原报单位。

(2) 对于可疑之处或无法代为更正的错误,应要求原单位复查更正。

(3) 如果所发现的差错在其他单位也可能发生时,应将错误情况通报所有单位,以免发生类似错误。

(4) 对于严重的错误,应发还重新填报,并查明发生错误的原因,若属于违法行为,则应依法严肃处理。

对审核过程中发现的错误应尽可能予以纠正。调查结束后,当在数据中发现的错误不能予以纠正,或者有些数据不符合调查的要求而又无法弥补时,就需要对数据进行筛选。数据筛选包括两方面内容:一是将某些不符合要求的数据或有明显错

误的数据予以剔除；二是将符合某种特定条件的数据筛选出来，对不符合特定条件的数据予以剔除。数据的筛选在市场调研中是十分重要的。

（三）对经过审核的资料进行分组、汇总

先根据研究目的和统计分析的需要，选择整理的标志，并进行划类分组（第五节详述）。统计整理的关键是统计分组，只有正确的分组才能整理出有科学价值的综合指标，并借助这些指标来揭示现象的本质与规律。然后根据研究目的设计整理汇总方案。统计汇总方案的设计包括两个方面：一是对于总体的处理方法，主要是如何进行统计分组；二是确定用哪些统计指标来说明总体。根据汇总方案，对各个调查项目的资料进行汇总，得出反映各组 and 总体数量特征的各种指标。

统计数据的汇总主要是针对直接调查取得的原始数据。统计资料的汇总组织形式可分为逐级汇总和集中汇总两种。逐级汇总是按照一定的统计管理制度，自下而上对调查资料进行逐级汇总，然后上报。我国的统计报表制一般都是这种形式。逐级汇总便于满足各级部门领导对统计资料的需要，但所需时间较长，产生汇总差错的可能性较大。集中汇总是将全部资料集中到组织调查的最高一级机关或由它指定的机构进行汇总。一般对时效性强的快速普查和对汇总要求很高的调查常用集中汇总形式。这种形式时效性强，可以减少汇总差错，但对于原始资料的差错难以更正，也不能满足各级部门对资料的需要。

由于统计调查搜集的原始数据，很大部分以问卷或调查表格形式存在，往往分散、不系统，不易表现出总体的数量特征。通过将问卷或调查表格记录的各单位信息汇总成一个更大的数据表，作为进一步整理的基础。选择恰当的汇总技术，对提高汇总速度，保证汇总质量具有重要意义。

统计汇总的技术主要有两种，即手工汇总和电子计算机汇总。

1. 手工汇总

手工汇总就是用算盘和小型计算器进行汇总，采用这种汇总技术其方法有划记法、过录法、折叠法、卡片法。手工汇总适合于总体单位数量和调查项目较少的调查研究，在总体单位数量和调查项目较多的调查研究中，手工汇总花费时间长，而且容易出错。

2. 电子计算机汇总

电子计算机大大提高了数据汇总的速度和精确度，其目前已经成为我国统计工作的重要工具。电子计算机汇总大体分为以下几个步骤：①选择统计软件，目前比较常用的是 Excel、SPSS、SAS 等统计软件；②设置变量，编码；③将问卷或调查表的数据录入；④资料错误信息的逻辑检查和修正；⑤运行程序，输出汇总结果，存为数据表。

本书将以 Excel 统计软件为例，介绍原始数据的电子汇总程序，参见统计实验部分。

(四) 统计资料的显示

通过统计表或统计图的形式,将整理出的资料简洁明了、系统有序地显示出来(第六节详述)。

(五) 统计资料系统积累、存档

统计研究中经常要用到动态分析,这就需要有长期积累的历史资料,而根据积累资料的要求,应对统计数据分门别类地系统积累、存档。所以统计整理也是积累历史资料的必要手段。

第五节

统计分组与分配数列

一、统计分组的意义与作用

(一) 统计分组的概念

根据统计研究的需要和现象总体的特点,将统计总体按照一定的标志分为若干个不同性质的组成部分,这样的一种统计方法称为统计分组。比如,要了解我国人口状况,只知道总人口数量是不够的,而应将人口总体按照年龄、性别、民族、城乡、文化程度等分组,才能进一步地深入地了解我国人口总体的年龄结构、性别比例、民族构成等。统计分组的对象是总体,统计分组的标志可以是品质标志,也可以是数量标志。

资料卡

统计分组的原则

需要注意的是,为了保证分组的科学性,必须遵循穷尽原则和互斥原则。其中,穷尽原则是指全部分组必须容纳所有的总体单位,即总体中的每个总体单位都必须有组的归属。例如,企业职工按文化程度分组,若只分为小学毕业、中学毕业和大学毕业三组,则未上过小学的以及大学以上文化程度的员工就无组可归了,这种分组就没有做到穷尽。互斥原则是指在特定的分组标志下,总体中的任何一个单位只能归于某一组,不能同时归于几个组。例如,将鞋子分为男鞋、女鞋、童鞋三类,就不符合互斥原则,因为童鞋也有男鞋与女鞋之分。

例1 我国第七次人口普查结果显示,我国的人口共1 411 778 724人(未包括港澳台地区),分别按性别和年龄分组如表2-3、表2-4所示。

表 2-3 2020 年我国人口（未包括港澳台地区）性别构成情况

按性别分组	人口数（人）	占总人口的比重（%）
男	723 339 956	51.24
女	688 438 768	48.76
合计	1 411 778 724	100

表 2-4 2020 年我国人口（未包括港澳台地区）年龄构成情况

按年龄分组	人口数（人）	占总人口的比重（%）
0～14 岁	253 383 938	17.95
15～59 岁	894 376 020	63.35
60 岁及以上	264 018 766	18.70
其中：65 岁及以上	190 635 280	13.50
合计	1 411 778 724	100

（二）统计分组的作用

（1）区分现象的不同类型：社会经济现象千差万别，要了解各种社会经济现象的性质、特点及其相互关系，必须根据某种标志把它们划分为性质不同的类型，以便揭示不同社会经济现象的质的差异。

例 2 表 2-5 是 2019 年我国居民人均消费支出构成情况。

表 2-5 2019 年我国居民人均消费支出构成情况

指标	支出（元）	占比（%）
总消费	21 559	100
食品烟酒消费	6 084	28.2
衣着消费	1 338	6.2
居住消费	5 055	23.4
生活用品及服务消费	1 281	5.9
交通通信消费	2 862	13.3
教育文化娱乐消费	2 513	11.7
医疗保健消费	1 902	8.8
其他用品及服务消费	524	2.4

资料来源：中华人民共和国国家统计局. 中国统计年鉴 [M]. 北京：中国统计出版社，2013.

表中将全部消费品分为八大类，通过分类，可以反映出我国居民消费中不同类别的商品所占的地位。

（2）研究总体的内部结构：从数量上反映总体内部的结构是统计研究的重要任务。总体的内部结构可体现部分与整体的关系以及各部分之间存在的差别和相互联系，反映事物从量变到质变的过程，帮助人们掌握事物的特征，认识事物的性质。

如例 1 中的表 2-3 和表 2-4, 就反映出 2020 年 11 月 1 日我国人口(未包括港澳台地区)性别、年龄构成情况。

(3) 分析现象之间的依存关系: 社会经济现象之间存在着广泛的相互依存关系, 如农作物的合理密植与产量之间、家庭的工资收入与生活费支出之间等。在研究过程中, 往往是先通过分组分析法观察出现象之间的依存关系, 然后在此基础上再应用其他方法做进一步深入分析。

例 3 表 2-6 是 2015—2019 年我国居民人均可支配收入及居民人均消费支出。

表 2-6 2015—2019 年我国居民人均可支配收入及居民人均消费支出

年份	我国居民人均可支配收入(元)	居民人均消费支出(元)
2015	21 966	15 712
2016	23 821	17 111
2017	25 974	18 322
2018	28 228	19 853
2019	30 733	21 559

从表 2-6 中可以看出我国居民人均可支配收入及居民人均消费支出之间的依存关系。

二、统计分组的方法

统计分组的关键问题是正确地选择分组标志与划分各组界限。

(一) 分组标志选择的原则

选择统计分组的标志应依据研究的目的与任务。如研究人民生活水平变动情况时, 可供选择的分组标志有: 家庭人口数、每户就业人数、家庭总收入、平均每人月生活费用等。而其中最能反映人民生活水平变动的标志是平均每人月生活费, 故应选择这一标志作为分组标志。选择分组的标志应根据现象所处的历史条件, 且能够反映事物本质或主要特征。例如, 同是划分企业规模, 在劳动力密集型的行业或地区, 可采用职工人数作为分组标志; 而在技术密集型的行业或地区, 则应选择固定资产价值或生产能力作为分组标志。

(二) 统计分组的种类

(1) 按分组的作用或目的不同, 分为类型分组、结构分组和分析分组。

类型分组是将复杂的现象总体, 划分为若干个不同性质的部分。如企业按所有制性质分组。

结构分组是在对总体分组的基础上计算出各组对总体的比重, 以此来研究总体各部分的结构。类型分组和结构分组往往紧密联系在一起。结构分组的目的是研究同质总体的构成。如人口按性别分组或按年龄分组。

分析分组是为研究现象之间的依存关系而进行的统计分组, 一般用在分析现象

的两个或两个以上标志之间的关系。选择用来分组的标志称为原因标志，与原因标志相对应的标志称为结果标志。如，分析家庭平均月收入水平与消费水平的关系，如果按家庭平均月收入分组，那么它就是原因标志，对应的消费水平就是结果标志。当然也可以按消费水平来分组。分析分组的目的是研究现象总体内部诸标志间的依从和制约关系。

(2) 按分组标志的多少，可分为简单分组、复合分组和并列分组。

简单分组是将总体按一个标志进行分组。复合分组是将总体按两个或两个以上的标志重叠起来进行分组，即在按某一标志分组的基础上再按另一标志进一步分组。并列分组是对总体采用两个或两个以上标志分别进行简单分组。

例如，例 1 中的表 2-3 和表 2-4 表示的 2020 年我国人口（未包括港澳台地区）分别按性别和年龄的简单分组情况。这两个简单分组合并在一起，就是并列分组。如果在按性别分组的基础上，再分别按接受教育程度分组，那就是复合分组了。

(3) 按分组标志的性质分为品质分组和数量分组。

品质分组是将总体按品质标志进行分组。品质分组比较简单，分组标志一经确定，一个品质标志表现即为一组，组的名称和组数也随之确定。例如，企业按经济类型、行业分组；人口按性别、民族分组；大学生按专业分组等。这种分组可以反映总体的构成和不同属性事物在总体中的地位和作用。如例 1 中的表 2-3，人口按性别分组就属于品质分组。

数量分组是将总体按数量标志进行的分组。例如，地区经济按国内生产总值分组、企业按销售收入分组等。这种分组的目的在于通过事物在数量上的差异来反映事物在性质上的区别。如例 1 中的表 2-4，人口按年龄分组就是一种数量分组。

品质分组和数量分组是一对重要的统计分组，统计分组方法主要是围绕这两种分组来阐述的。

三、次数分布表的编制

(一) 分配数列的意义

在统计分组的基础上，总体中的所有单位按其所属的组别归类整理，列出各组对应的单位数，并且按照一定的顺序排列，所形成的数列称为分配数列，又称次数分配或次数分布。

分配数列中，分布在各个组的总体单位数叫次数，又称频数。

如果将分组标志序列与各组相对应的频率按照一定的顺序排列，就形成频率分布数列。

分配数列有两个组成要素：一是总体按某标志所分的组；另一个是次数或频率。

(二) 分配数列的种类

根据分组标志的性质不同，有品质数列与变量数列。

1. 品质数列

品质数列是按品质标志分组形成的数列,它由各组名称和各组的单位数构成,用来观察总体单位中不同属性的单位分布情况。如表 2-7 所示。

表 2-7 2018 年我国人口性别构成情况

按性别分组	人口数(万人)	占总人口的比重(%)
男	71 351	51.13
女	68 187	48.87
合计	139 538	100
(各组名称)	(次数)	(频率)

品质数列的编制比较简单,组限是自然形成的或比较明显的,但要注意分组时,应包括分组标志的所有表现,不能有遗漏,各种表现相互独立,不得相容。如果存在属性之间的过渡形式,使分组界限难以确定,如企业按行业分组,产品按经济用途分组等,这种比较复杂的属性分组,国家有关部门都制定有标准的分类目录,分组时可以依据分类目录来确定组限。

2. 变量数列

变量数列是将总体按数量标志分组,将分组后形成的各组变量值与该组中所分配的单位次数或频数,按照一定的顺序相对应排列所形成的分配数列。

例 4 某班级 50 名学生的统计学成绩按考试分数分组形成的分配数列,如表 2-8 所示。

表 2-8 某班级统计学成绩分布表

考试分数	人数(人)	频率(%)
60 以下	3	6.0
60 ~ 70	12	24.0
70 ~ 80	15	30.0
80 ~ 90	12	24.0
90 以上	8	16.0
合计	50	100.0
(各组变量值)	(次数)	(频率)

变量数列按分组的形式不同又分为单项式变量数列与组距式变量数列。

单项式变量数列是用一个变量值作为一组(组名就是该变量值)所形成的分配数列。对离散型变量,如果变量值的变动幅度不大,就可用一个变量值对应一组,单项式分组一般适用于离散型变量且变量变动范围不大的情形。如居民家庭按儿童数或人口数分组,均可采用单项式分组,育龄妇女按其生育子女存活数分组等。

组距式变量数列是把整个变量值依次划分为几个区间,把一段区间内的所有变量值作为一组所形成的分配数列。对于连续型变量由于不能一一列举其变量值,只能采用组距式的分组方式,且相邻的组限必须重叠。为了遵循统计分组穷尽和互斥原则,统计上规定“上限不在内”的原则,即每一组的变量值区间包含下限在内,

但不包含上限（也就是数学上所定义的左闭右开区间）。对于变动范围较大的，变量值的个数很多的离散型变量，也适宜采用组距式分组。离散型变量在组距式分组中，其组限可以间断，也可以重叠。

在组距式变量数列中，需要明确以下概念。

（1）组限。

组距式变量数列中，各组两端的数值称组限。每一组的两个组限中，较大者叫上限，较小者叫下限。如果各组的组限都齐全，称为闭口组；组限不齐全，即最小组缺下限或最大组缺上限，称为开口组。

分组时各组组限的确定必须能反映事物质的差别。例如，学生按学习成绩分组，不能把 55 分和 65 分合为一组，因为这样的分组未区分及格与不及格的质的差别。一般地，第一组的下限应小于或等于实际变量值的最小值，最后一组的上限必须略大于实际变量值的最大值，并尽可能使各单位的标志值在组内分布比较均匀。

（2）组距。

每组下限与上限之间的距离称为组距，表示各组标志值变动的范围。即：

$$\text{组距} = \text{上限} - \text{下限}$$

对于间断式分组的组距大小的计算，应该采用如下公式：

$$\text{组距} = \text{上组下限} - \text{本组下限}$$

组距式变量数列，有等距数列和不等距（异距）数列之分。统计分组时采用等距分组还是不等距分组，取决于研究对象的性质特点。在标志值变动比较均匀的情况下宜采用等距分组。等距分组便于各组单位数和标志值直接比较，便于绘制统计图，也便于计算各项综合指标。在标志值变动很不均匀的情况下采用不等距分组，有时更能说明现象的本质特征。

（3）组中值。

上下限之间的中点数值称为组中值，以代表各组标志值的一般水平。组中值并不是各组标志值的实际平均数，各组标志数的平均数在统计分组后很难计算出来，常以组中值近似代替。组中值仅存在于组距式分组数列中，单项式分组中不存在组中值。组中值的计算是假定各组标志值的变化是均匀的（与等距式分组的假定条件相同）。一般情况下，

$$\text{组中值} = \frac{\text{上组的下限} + \text{本组的下限}}{2} \quad (\text{适用所用闭口组})$$

对于开口组的组中值的计算方式可以利用如下公式：

$$\text{缺下限组的组中值} = \text{本组上限} - \frac{\text{邻组组距}}{2}$$

$$\text{缺上限组的组中值} = \text{本组下限} + \frac{\text{邻组组距}}{2}$$

(4) 组数。

组数的多少直接取决于两个因素，一个是总体的全距，另一个是组距。在等距分组的条件下，组数等于全距除以组距。全距是总体中最大的标志值与最小的标志值之差。

决定组数的多少，并无规则可言，必须凭借经验和所研究问题的性质做出判断。美国学者斯特杰斯（Sturges）提出一种确定组数和组距的经验公式，称为斯特杰斯经验公式，即 $n=1+3.322 \lg N$ ，则有 $i = \frac{R}{n} = \frac{R}{1+3.322 \lg N}$ ，公式中， n 为组数， N 为总体单位数， i 为组距， R 为全距。

在全距一定的条件下，组距大则组数少，组距小则组数多。全距是客观存在的事实，不以人的意志为转移，所以，确定组数的关键是确定组距。

（三）分配数列的编制**1. 单项式变量数列的编制**

用一个变量值作为一组进行分组，编制的分配数列就是单项式变量数列。其编制方法，用以下例题说明。

例 5 某棉纺厂织布车间有 27 名工人，分别看管机器台数如下：

5、4、2、4、3、4、3、4、4
2、4、3、4、3、2、6、4、4
2、2、3、4、5、3、2、4、3

编制单项式变量数列。

解：首先将每一个工人的标志值（看管机器台数）按大小顺序排列：2、2、2、2、2、2、3、3、3、3、3、3、3、3、4、4、4、4、4、4、4、4、4、4、4、4、4、5、5、6。然后，以总体各单位的标志值（看管机器台数）为各组的标志值，对应每一个标志值出现的次数为各组的次数或频数，计算出各组的频率并列表，如表 2-9 所示。

表 2-9 工人看管机器台数分布表

工人看管机器台数（台）	工人数（人）	频率（%）
2	6	22.2
3	7	25.9
4	11	40.8
5	2	7.4
6	1	3.7
合计	27	100

注意到，工人看管机器台数是标志，工人数是总体单位数，分组是按标志值进行的，然后统计对应各组的总体单位数就是频数（或次数）。

2. 组距式变量数列的编制

组距式变量分配数列编制的步骤如下。

① 将原始资料按其数值从小到大顺序重新排列。② 确定全距：全距是变量值中

最大值和最小值的差数。③确定组距和组数：在实际应用中，组距应是整数，最好是5或10的整数倍。在确定组距时，必须考虑原始资料的分布状况和集中程度。④确定组限：组限要根据变量的性质来确定。如果变量值相对集中，无特大或特小的极端数值时，则采用闭口组。采用闭口组时，应做到最小组的下限小于或等于最小变量值，最大组的上限必须大于最大变量值。如果变量值相对比较分散，则采用开口式，使最小组只有上限（用“××以下”表示），最大组只有下限（用“××以上”表示）。⑤编制变量数列：确定了组限以后，就可以把变量值归类排列，最后把各组单位数综合后填入表格相应的各组次数栏中。

以下举例说明。

例6 根据抽样调查，某市50户居民月购买消费品支出额资料如表2-10所示。

表2-10 某市50户居民月购买消费品支出额

单位：元

830	880	1 230	1 100	1 180	1 580	1 210	1 460	1 170	1 080
1 050	1 100	1 070	1 370	1 200	1 630	1 250	1 360	1 270	1 420
1 180	1 030	870	1 150	1 410	1 170	1 230	1 260	1 380	1 510
1 010	860	810	1 130	1 140	1 190	1 260	1 350	930	1 420
1 080	1 010	1 050	1 250	1 160	1 320	1 380	1 310	1 270	1 250

编制组距式分配数列。

解：

第一步：计算全距。

将各变量值由小到大排序（略），找出最大值、最小值并计算全距。变量的最大值是1 630，最小值是810。

$$\text{全距} = \text{最大值} - \text{最小值} = 1\ 630 - 810 = 820$$

第二步：确定组数和组距。

若取组距为100，由组数 $=820 \div 100 = 8.2$ ，则需要分成9个组。考虑到数据分布的特点，偏小和偏大的数据相对比较分散，则采用开口式，最小组设为1 000以下，最大组设为1 500以上，这样分成7个组就行了。

第三步：确定组限。

由于最小变量值是810，最小组的下限（起点值）应低于最小变量值，故可以取最小组的下限为800。考虑到900~1 000只有一个数据，较为分散，故将该两组合并为一组并采用开口式，设为“1 000以下”。其余按组距100分别确定各组的上下限（这里采用组限重叠方式）。

第四步：编制频数（频率）分布表，如表2-11所示。

表 2-11 某市 50 户居民月购买消费品支出情况表

按户月消费品支出额分组(元)	频数(户)	频率(%)
1 000 以下	6	12
1 000 ~ 1 100	8	16
1 100 ~ 1 200	11	22
1 200 ~ 1 300	11	22
1 300 ~ 1 400	7	14
1 400 ~ 1 500	4	8
1 500 以上	3	6
合计	50	100

第五步：计算累计频数和累计频率。

为了更详细地认识变量的分布特征，还可以计算累计频数和累计频率，编制累计频数和累计频率数列。累计频数和累计频率有向上累计频数（频率）和向下累计频数（频率）两种。

以变量值大小为依据，由变量值小的组向变量值大的组累计频数和频率，成为向上累计频数和向上累计频率。

向上累计的意义是：计算小于该组上限的各组的频数或频率之和。一般用于低端情况的分析。计算时先由标志值低的组向标志值高的组依次累计频数（或频率），直到该组的上限截止。

相反，由变量值大的组向变量值小的组累计各组的频数（或频率），称为向下累计频数（或频率）。

向下累计的意义是：计算大于及等于该组下限的各组的频数或频率之和。一般用于高端情况的分析。计算时，先由标志值高的组向标志值低的组依次累计频数（或频率），直到该组下限截止。

以根据表 2-11 所得 50 户居民月购买消费品支出额的资料为例，分别进行向上和向下累计，其结果如表 2-12 所示。

表 2-12 某市 50 户居民月购买消费品支出累计表

居民月消费品支出额分组(元)	频数(户)	频率(%)	向上累计		向下累计	
			累计频数(户)	累计频率(%)	累计频数(户)	累计频率(%)
1 000 以下	6	12	6	12	50	100
1 000 ~ 1 100	8	16	14	28	44	88
1 100 ~ 1 200	11	22	25	50	36	72
1 200 ~ 1 300	11	22	36	72	25	50
1 300 ~ 1 400	7	14	43	86	14	28
1 400 ~ 1 500	4	8	47	94	7	14
1 500 以上	3	6	50	100	3	6
合计	50	100	—	—	—	—

以上资料表明：居民月消费品支出额在 1 100 元以下的有 14 户，占总数的 28%；月消费品支出额在 1 200 元以下的有 25 户，占总数的 50%；依此类推。

居民月消费品支出额在 1 400 元以上的有 7 户，占总数的 14%；月消费品支出

额在 1 300 元以上的有 14 户，占总数的 28%；依此类推。

资料卡

构建定量数据频数分布应注意的问题

实际中构建定量数据频数分布时，应注意以下两点。

(1) 尽量避免设置开口组。因为对于开口组，有时无法确定其组中值，容易给一些统计量的计算带来困难。

(2) 灵活运用不等距分组。如果不同组的组距不相等，则这种分组称为不等距分组。不等距分组常常出现在某个或某些组的频数过低的情形。例如，有以下 20 个数据：

1 1 1 2 2 2 3 3 3 11 18 18 19 19 19 19 21 21 21

若按照惯例，则形成的频数分布如表 2-13 所示。

表 2-13 组距为 4 的等距分组

组	频数
0 ~ 4	9
4 ~ 8	0
8 ~ 12	1
12 ~ 16	0
16 ~ 20	7
20 ~ 24	3
合计	20

通过表 2-13 可以看到。在这个频数分布中，第二组、第三组和第四组的频数非常小，其中第二组和第四组的频数为零。如果数据量很大，而且数据在某些区间上的分布很稀疏，则空白组或稀疏组的数目就会较多。对于这种情形，可以将多个相邻组合并，形成组距不等的结果。例如，表 2-13 的频数分布经过合并后变为表 2-14。

表 2-14 不等距分组

组	频数
0 ~ 4	9
4 ~ 16	1
16 ~ 20	7
20 ~ 24	3
合计	20

四、次数分布的类型

由于社会经济现象性质不同，各种统计总体有不同的次数分配，呈现出各种不同的分布特征。描述统计总体的分布特征，还可采用直方图或曲线图进行描述。通过这些图形，可以更直观地表明现象的分布特征。概括起来，曲线形状的特点，大

致有三种类型：钟形分布、U形分布和J形分布。

(一) 钟形分布

钟形分布的特征是“两头小、中间大”，即靠近中间的变量值分布的次数多，靠近两端的变量值分布的次数少，如果将变量值与其对应的频数在直角坐标系中对应的点连接起来绘制成曲线图，宛如一口钟，所以称之为钟形分布。

在自然或社会经济现象中，有许多次数分布是属于钟形分布的。例如，人的体重、身高，学生的成绩，居民的货币收入，单位面积的农产品产量，市场价格等现象都属于钟形分布。

如图 2-2 (a) 所示，其分布特征是以标志变量中间值为对称轴，左右两侧对称，两侧变量值分布的次数随着与变量中间值距离的增大而渐次减少。在统计学中，称这种分布为对称分布（也叫正态分布）。而图 2-2 (b) 为偏态分布，它们各有不同方向的偏态，即左偏态分布和右偏态分布。人均收入分布的曲线就是左偏曲线，即低收入的人数较多，在收入较低的左边形成高峰，高收入的人数较少，且收入越高的人越少，在右边形成了一个细长的尾巴。

客观实际中，许多社会现象统计总体的分布都趋于钟形分布中的正态分布。正态分布是描述统计中的一种主要分布，它在社会经济统计分析中具有重要的意义。

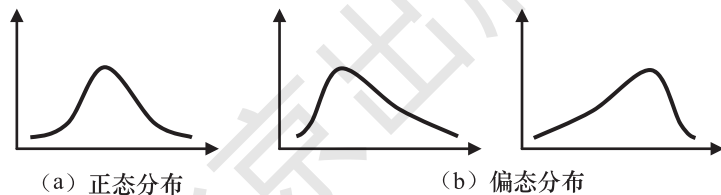


图 2-2 钟形分布示意图

(二) U形分布

U形分布的特征与钟形分布刚好相反：靠近中间的变量值分布的次数少，靠近两端的变量值分布的次数多，形成“两头大，中间小”的分布特征。将这种分布绘成曲线，像英文字母“U”的形状，故称U形分布。如图 2-3 (a) 所示。

例如，人口死亡率的分布，一般是婴幼儿死亡率和老年人死亡率均较高，而中年人死亡率最低，所以人口年龄分组的死亡率是呈U形分布的。另外，失业人口按年龄的分布也呈U形分布。

(三) J形分布

J形分布的特征是“一边小，一边大”，呈现单边上升或单边下降的趋势。有正J形分布和反J形分布两种类型。

正J形分布是次数随着变量值的增大而增多，如图 2-3 (b) 所示。如投资额按利润率大小分布，供给量按价格的分布一般都是正J形分布。

反J形分布是次数随着变量值的增大而减小，如图 2-3 (c) 所示。如成年人人数

量按年龄大小分组，表现出年龄越高，人数越少。还有随着产品产量的增加，产品单位成本下降等。

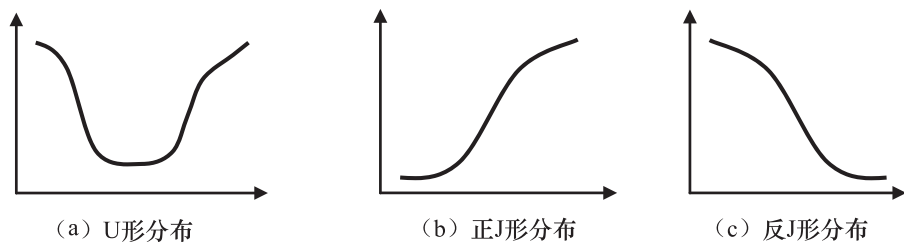


图 2-3 U 形分布与 J 形分布示意图

(四) 研究变量数列次数分布类型的作用

可以利用各种类型的次数分布特征，以检验统计整理资料的准确性；可利用次数分布类型的特征，分析现象变化的原因，如果现象总体发生了异常变化，会产生通过整理所得到的分布类型与社会现象的分布特征不相符的情况。

请思考

1. 将大学老师总体进行如下分组：按一个品质标志分组、按两个品质标志的复合分组、按一个品质标志和一个数量标志的复合分组、按两个数量标志的复合分组。
2. 如果要对某地区的家庭按家庭拥有的儿童数进行分组，请问应编制单项式数列还是组距式数列？为什么？如果对某地区的所有工业企业按产值分组，又应编制什么样的数列呢？为什么？

第六节

统计表与统计图

经过整理的统计数据，常常用统计表或用统计图将其展示出来，便于人们进一步的分析。统计表条理分明，集中醒目，减少大量的文字叙述，便于对比分析与积累；统计图能够形象、鲜明、直观地显示现象之间的相互关系。

一、统计表

(一) 统计表的定义和结构

统计调查所得来的原始资料，经过整理，得到说明社会现象及其发展过程的数据，将这些数据按一定的顺序排列在表格上，就形成了统计表。

统计表可分为广义统计表和狭义统计表两种。广义统计表包括统计工作各阶段中所用的一切表格。它包括调查阶段的调查表、统计报表，整理阶段的分组表、汇总表，

以及分析阶段的分析表和公布统计资料的表式等。狭义统计表也就是通常所说的统计表,主要是指记载汇总结果和公布统计资料的表式。本节所讲的统计表是指狭义统计表。

利用统计表,可以合理地安排统计资料,清晰、简明地反映出现象总体的特征,因而是统计分析的重要工具。它清楚地、有条理地显示统计资料,并能直观地反映统计分布特征。统计表还是汇总和积累统计资料,进行统计分析的重要手段。

统计表的结构,可以从表式和内容两个方面来认识。

(1) 从外表形式上看,统计表由总标题、横行标题、纵栏标题和指标数值四个部分组成。如图 2-4 所示。

总标题是统计表的名称,它扼要地说明表的基本内容,并指明时间和范围。它置于统计表格的正上方。

横行标题是横行各组的名称,一般放在表格的左方;纵栏标题是纵行的名称,它是分组标志或指标的名称,说明纵行所列各项资料的内容,一般放在表格的上方。横行标题和纵行标题共同说明填入表格中的统计数字所表达的内容。

指标数值是列在横行和纵栏的交叉处,即表格中的数字就是指标数值,用来说明总体及其组成部分的数量特征,它是统计表格的具体内容。

总标题			
20××年规模以上工业增加值及其增长速度统计表			
指标	增加值 (亿元)	比上年 增长(%)	纵栏 标题
规模以上工业	79 752	16.6	指标 数值
其中: 国有及国有控股企业	28 396	12.6	
其中: 集体企业	2 558	11.6	
股份制企业	39 918	17.8	
外商及港澳台投资企业	22 502	16.9	
其中: 私营企业	15 547	24.4	
其中: 轻工业	24 314	13.8	
重工业	55 438	17.9	
	主词栏	宾词栏	

图 2-4 统计表的结构

(2) 从统计表的内容来看,由主词和宾词两个部分组成。主词是统计表所说明的总体,总体的各组或各组的名称。宾词是用于说明主词的各个统计指标及其数值。通常,统计表的主词列在表的左方,宾词列在表的右方,如图 2-4 所示。

（二）统计表的种类

统计表按照总体分组情况不同（根据主词是否分组和分组的情况），可分为简单表、分组表和复合表三类。

（1）简单表。

主词未经过任何分组的统计表称为简单表，也称一览表。它是罗列出总体各单位的名称或按时间顺序简单排列的统计表。

（2）分组表。

分组表是主词只按照一个标志分组的统计表，也称简单分组表。它可以揭示出现象的不同类型的特征，研究现象的内部结构，分析现象的依存关系，如表2-15所示。

表 2-15 某年某地区固定资产投资情况统计表

固定资产投资按经济类型分组	投资额（亿元）	所占比率（%）
国有经济投资	85.03	42.0
集体经济投资	109.67	54.1
居民个人投资	2.33	1.1
其他类型投资	5.61	2.8
合计	202.64	100

（3）复合表。

复合表是主词按照两个或两个以上的标志层叠分组所形成的统计表。在复合分组表中设计横行标题时，第二次分组的组别应在第一次分组的各组组别下退一或二字，此时第一次分组的组别就成为第二次分组的各组小计。若需再进行第三、四次分组，均可按此类推。如图2-4中2016年规模以上工业增加值及其增长速度的统计表。

（三）统计表的设计

统计表设计总的要求是：简练、明确、实用、美观、便于比较。统计表设计应注意的事项如下。

（1）标题设计。统计表的总标题，横行、纵栏标题应简明扼要，以简练而又准确的文字表述统计资料的内容、资料所属的空间和时间范围。

（2）线条的绘制。表的上下端应以粗线绘制，表内纵横线以细线绘制。表格的左右两端一般不画线，采用“开口式”。

（3）合计栏的设置。统计表各纵列若需合计时，一般应将合计列在最后一行，各横行若需要合计时，可将合计列在最前一栏或最后一栏。

（4）指标数值。表中数字应该填写整齐，对准位数，同一列应精确到小数点后同一位。当数字小到可忽略不计时，可写上“0”；当缺某项数字资料时，可用符号“…”表示；不应有数字时，用符号“—”表示。

（5）计量单位。统计表必须注明数字资料的计量单位。当全表只有一种计量单

位时,可以把它写在表头的右上方。如果表中各行的指标数值计量单位不同,可在横行标题后添一列计量单位。

(6) 栏数的编号。如果栏数较多,应当按顺序编号,习惯上主词栏部分分别编以“甲,乙,丙,丁……”序号,宾词栏编以“(1)(2)(3)……”序号。

(7) 注解或资料来源。为保证统计资料的科学性与严肃性,在统计表下应加注解,以便查考。必要时,在统计表下加注说明。

二、统计图

统计图是统计资料的另一种表现形式。它可以简洁直观地展示统计表中的数据关系,具有鲜明醒目、易于理解的特点,可以更迅速、更有效地传递信息,揭示现象的内部结构和依存关系,显示现象的发展趋势和分布状况,帮助我们发现规律,以便进一步地统计分析与研究。随着计算机技术不断发展,电脑制图功能日益强大,使得统计图的制作更加方便和精确。

统计图一般包括以下几个基本要素。

(1) 标题。有图标题、数值轴(横、纵轴)标题。图标题是统计图的名称,它简明扼要地说明全图的基本内容,一般置于统计图的上端(或下端)。

(2) 绘图区。将统计表中的数据或原始数据绘制出图形的区域。

(3) 图例。图例用来表明图中的数据系列,一般置于图的右上方或右方。如果只有一个数据系列则不需要图例。

统计图的种类很多,常用的有圆形图、条形图、直方图、折线图等。

(一) 圆形图

圆形图又称为饼图,是用圆形和圆内扇形的面积来表示数值大小的图形,主要用于表示总体中各组成部分所占的比例,对研究结构性问题十分有用。

例 7 2020 年 AppStore 游戏畅销榜 TOP1000 游戏类型分布,如图 2-5 所示。

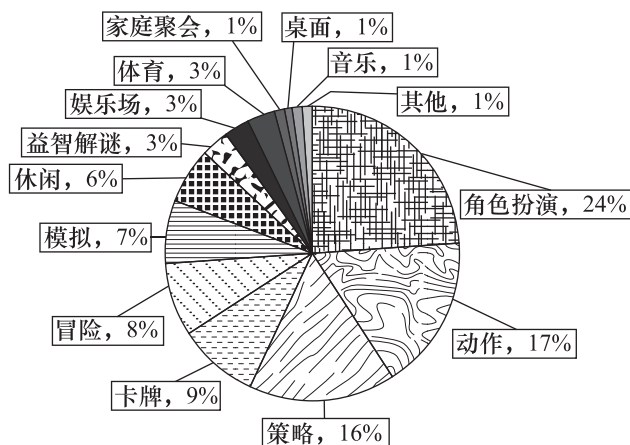
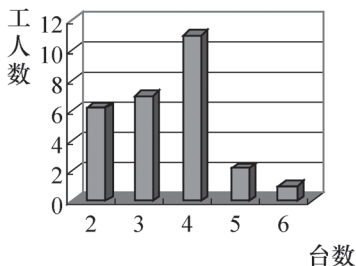


图 2-5 2020 年 AppStore 游戏畅销榜 TOP1000 游戏类型分布

（二）条形图

条形图可用于显示离散型变量的次数分布。最主要是显示顺序数据和分类数据的频数分布。条形图中条形的宽度是相同的，而用大小不等的高度或长度表示各类别数据的频数或频率。条形图可以横置或纵置，纵置时也称为柱形图。此外，条形图有单式、复式等形式。



例 8 根据表 2-9 绘制的工人看管机器台数分布情况条形图，如图 2-6 所示。

图 2-6 工人看管机器台数分布情况条形图

（三）直方图

直方图是用长方形的宽度和高度来表示次数分布的图形。绘制直方图时，一般情况下，横轴表示各组组限，纵轴表示次数（一般标在左方）和比率（或频率，一般标在右方），若没有比率，直方图只保留左侧的次数。依据各组的组距的宽度与次数的高度绘成长方形。直方图与条形图区别在于，直方图的长条形紧密地排列在一起，而条形图的长条形是分散地排列，原因就在于离散数据与连续数据之间的差别；还有，条形图是用条形的长度表示各类别频数的多少，其宽度（表示类别）是固定的；直方图是用面积表示各组频数的多少，长方形的高度表示每一组的频数密度，宽度则表示各组的组距，在不等距分组时其矩形的宽度是不同的，因此其高度与宽度均有意义。对于不等组距式变量数列，则通常按次数密度（频数密度）绘制直方图以表示其分布。

例 9 根据表 2-10，某市 50 户居民月购买消费品支出额的资料绘制直方图，如图 2-7 所示。

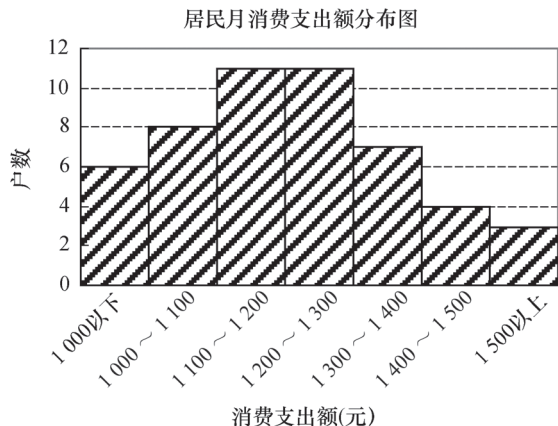


图 2-7 某市 50 户居民月购买消费品支出额直方图



直方图与条形图的区别

资料卡

直方图与条形图的区别

直方图和条形图外观比较相似,但两者却有着本质上的差异。具体表现为以下几点。

(1) 条形图的横轴只表示类的差异,却没有具体的坐标;而直方图的横轴则有真正的坐标,矩形的位置不可以随意移动。因此,条形图中矩形的宽度没有意义,而直方图中矩形的宽度则代表每一组的组距。

(2) 条形图中矩形之间既可以分离,也可以相接,其排列方式由绘图者决定;而直方图中矩形之间有无间隔则不是随意决定的,而是由每一组的区间位置决定的。

(3) 条形图的纵轴表示频数;而直方图的纵轴只有在等距分组的情形下才表示频数,在非等距分组的情形下表示的则是频数(或频率)密度。

(四) 折线图

折线图是以线段的起伏表示其数量分布的特征,绘图时,可以在直方图的基础上,用折线将各组次数高度的坐标连接而形成折线图,也可以用组中值与次数求坐标点连接而成。

需要注意,折线图的两个终点要与横轴相交,具体的做法是将第一个矩形的顶部中点通过竖边中点(即该组频数一半的位置)连接到横轴,最后一个矩形顶部中点与其竖边中点连接到横轴。这样才会使折线图下所围成的面积与直方图的面积相等,从而使二者所表示的频数分布一致。

对数据所分的组数很多时,组距会很小,这时所绘制的折线图就会越来越光滑,逐渐形成一条平滑的曲线,这就是频数分布曲线。

例 10 某企业加工车间 40 名工人日产量(件)资料如表 2-16 所示,用图示法描述工人日产量的分布状况。

表 2-16 40 名工人日产量次数分布表

按日产量分组(件)	次数	频率(%)
50 ~ 60	2	5.0
60 ~ 70	7	17.5
70 ~ 80	11	27.5
80 ~ 90	12	30.0
90 ~ 100	8	20.0
合计	40	100.0

解:根据表 2-16 的资料绘制的直方图和折线图(图 2-8、图 2-9)如下。

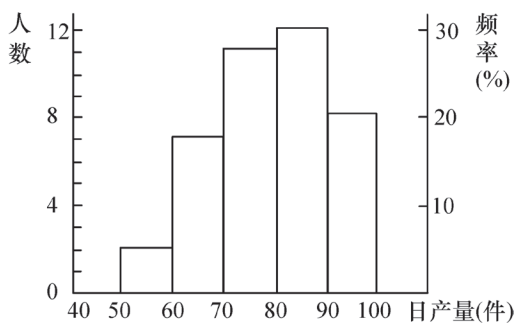


图 2-8 工人日产量次数分布直方图

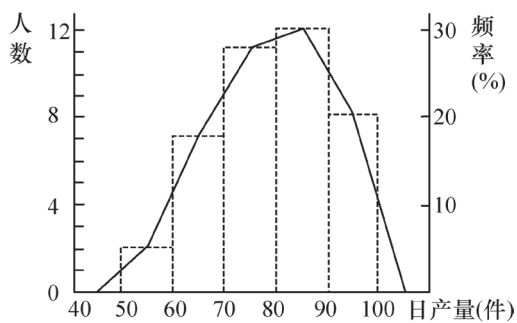


图 2-9 工人日产量次数分布折线图

第七节 Excel 在统计数据处理中的应用

Excel 是微软公司出品的 Office 系列办公软件中的一个组件, 利用它可以进行表格处理、图形分析、数据管理等。Excel 操作简单, 具有强大的统计功能, 是进行统计实践的重要工具。本教材主要以 Excel 2003 为示范软件。

一、使用 Excel 编制统计表

例 11 一家文具店, 2016—2020 年的销售收入分别为 40 000 元、50 000 元、65 000 元、82 000 元和 96 000 元; 销售支出分别为 16 000 元、21 000 元、28 000 元、34 000 元和 42 000 元。使用 Excel 对上述数据进行处理, 编制统计表。

解: 具体步骤如下。

(1) 命名工作表。工作表的命名有两种方法。第一种是启动 Excel 之后, 系统会自动创建一个空白工作簿“Book1”, 此工作簿包含 3 个工作表, 分别为 Sheet1、Sheet2 和 Sheet3。在 Sheet1 上单击鼠标右键, 弹出快捷菜单, 选择“重命名”选项, 输入名称“2016—2020 年文具店经营状况统计表”。另一种方法是启动 Excel 之后, 执行“格式”—“工作表”—“重命名”命令, 输入名称“2016—2020 年文具店经营状况统计表”。结果如图 2-10 所示。

(2) 创建统计表标题。统计表文本内容的输入非常简单, 可在单元格中直接输入文字。打开已经创建的统计表, 如果需要在第一行输入统计表名称, 首先合并单元格, 然后输入名称“2016—2020 年文具店经营状况统计表”, 在下一行输入计量单位“元”。接着, 在单元格 A3 中输入第一列标题“年份”, 移动光标至 B3 输入第二列标题“销售收入”, 移动光标至 C3 输入第三列标题“销售支出”。结果如图 2-11 所示。

在单元格 B4 ~ B8 依次输入 40 000、50 000、65 000、82 000、96 000，在单元格 C4 ~ C8 依次输入 16 000、21 000、28 000、34 000、42 000，编制成如图 2-12 所示的统计表。



年份	销售收入	销售支出
2016	40000	16000
2017	50000	21000
2018	65000	28000
2019	82000	34000
2020	96000	42000

图 2-12 输入统计数据

此外，在录入数据时，还可灵活使用 Ctrl+C 和 Ctrl+V 快捷键，迅速复制数据并粘贴到相应的单元格中。

二、使用 Excel 绘制统计图

随着计算机技术的迅猛发展，计算机制图功能日益提高。Excel 软件提供了 100 多种基本图形类型，以下介绍其中常用的几种统计图的绘制方法。

（一）条形图的绘制

例 12 （续例 11）使用 Excel 对文具店 2016—2020 年的销售收入绘制条形图。

解：具体步骤如下。

（1）打开 Excel 工作表，在工作表的 A1、B1 单元格中分别输入“年份”“销售收入”字样，在其下方区域（A2：B6）分别输入相应的数值。

（2）执行“插入”——“图表”命令（或单击工具栏中的“图标向导”按钮），弹出“图表向导—4 步骤之 1—图表类型”对话框。在“标准类型”选项中选择“条形图”，再选择其中的子图表类型（本例选择第一个）。具体如图 2-13 所示。

（3）单击“下一步”按钮，弹出“图表向导—4 步骤之 2—图表源数据”对话框。在“数据区域”栏中输入数据所在单元格区域，本例中输入 B1：B6，也可以单击“数据区域”栏尾端的小图标，再用鼠标选定该数据区域，如图 2-14 所示。在“系列”

选项卡中,于“分类(X)轴标志”栏中输入A2:A6,如图2-15所示。

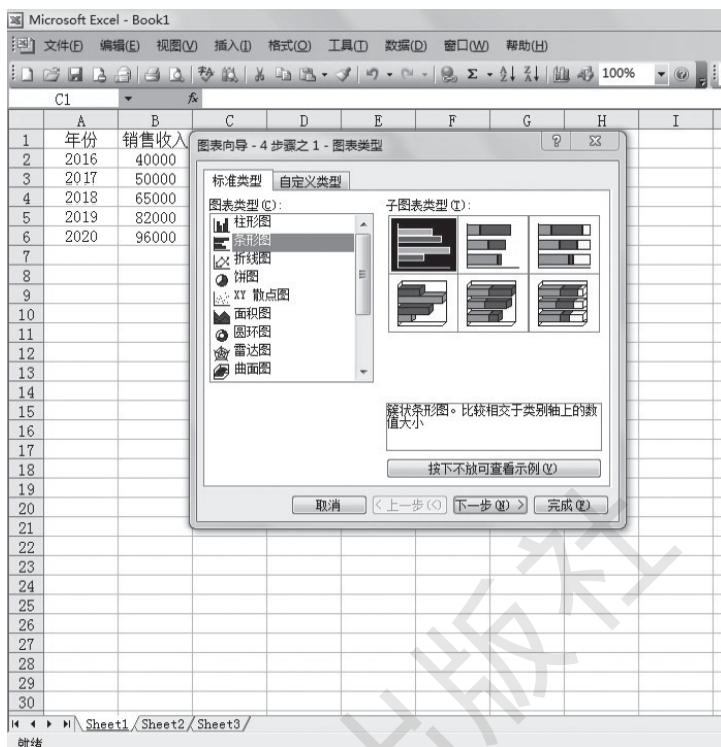


图 2-13 选择条形图类型

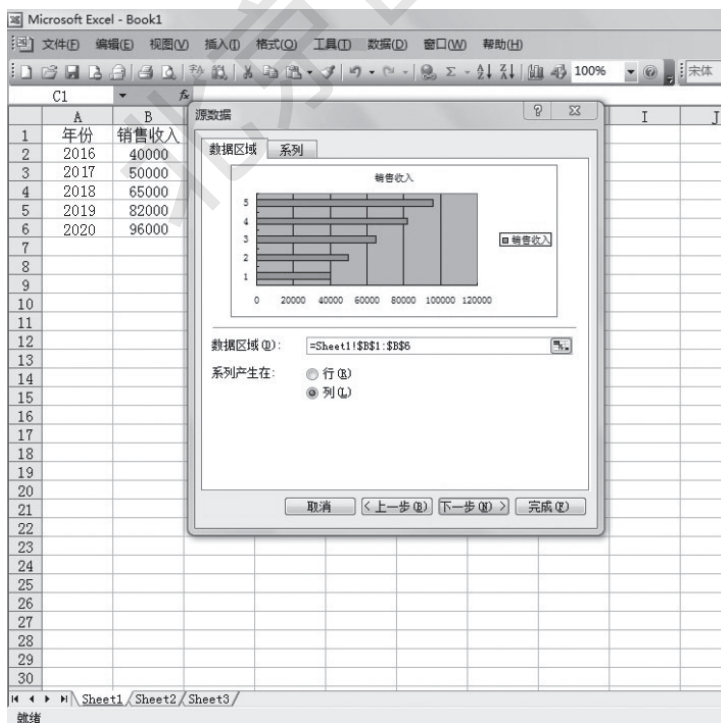


图 2-14 输入数据区域

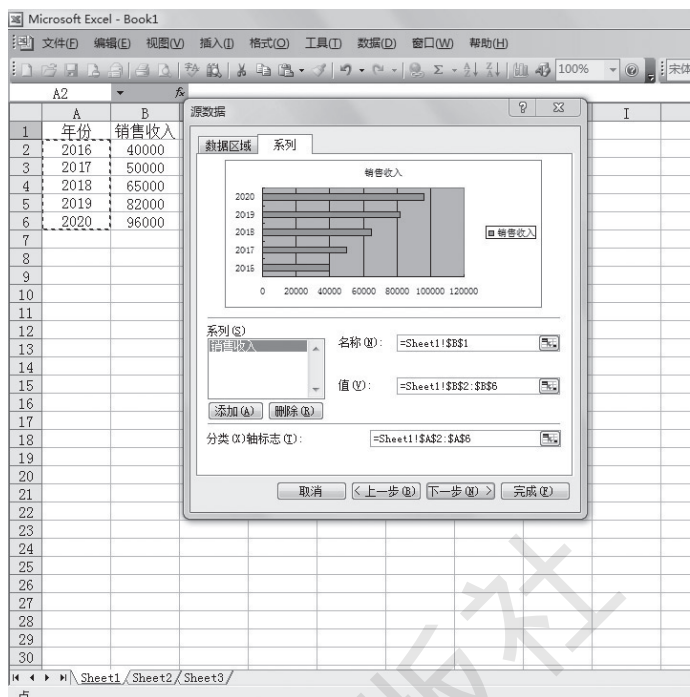


图 2-15 输入“分类 (X) 轴标志”

(4) 单击“下一步”按钮，弹出“图表向导—4 步骤之 3—图表选项”对话框。根据需要，分别在“标题”“坐标轴”“网格线”“图例”“数据标志”“数据表”等选项中做出选择和修改，如图 2-16 所示。最后单击“完成”，便可得到图 2-17。

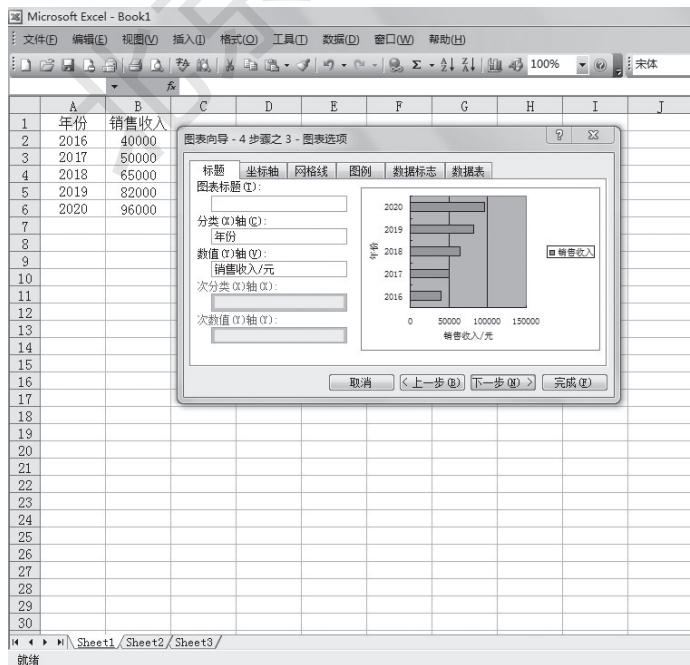


图 2-16 选择和修改图表选项

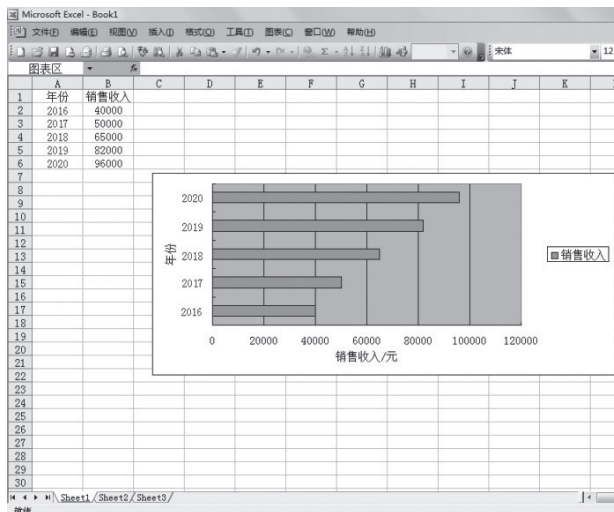


图 2-17 最终条形图

(二) 饼图的绘制

例 13 使用 Excel 对数据绘制饼图。

解：具体步骤如下。

(1) 打开 Excel 工作表，在工作表的 A1、B1 单元格中分别输入“兴趣”“人数”字样，在其下方区域 (A2: B6) 分别输入相应的数值。

(2) 执行“插入”——“图表”命令，弹出“图表向导—4 步骤之 1—图表类型”对话框。在“标准类型”选项中选择“饼图”，然后选择其中的子图表类型 (本例选择第二个)。具体如图 2-18 所示。

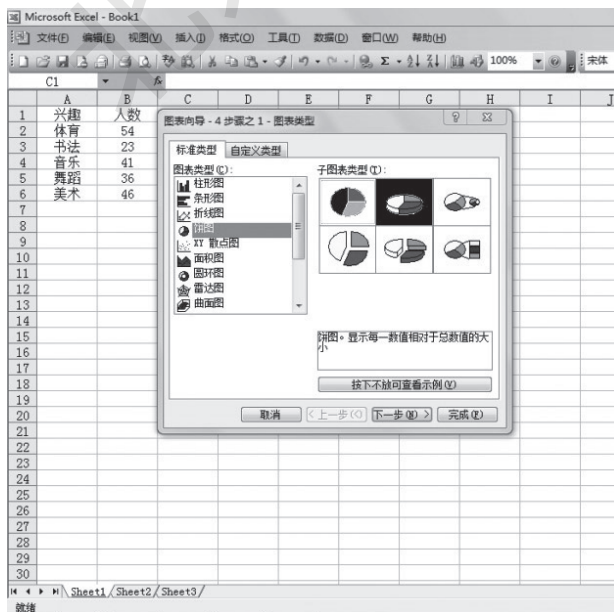


图 2-18 选择饼图类型

(3) 单击“下一步”按钮，弹出“图表向导—4 步骤之 2—图表源数据”对话框。在“数据区域”栏中输入数据所在单元格区域，本例中输入 B1: B6；在“系列”选项卡中，于“分类 (X) 轴标志”栏中输入 A2: A6。

(4) 单击“下一步”按钮，弹出“图表向导—4 步骤之 3—图表选项”对话框。根据需要，分别在“标题”“图例”“数据标志”等选项中做出选择和修改。最后单击“完成”，便可得到图 2-19 所示的饼图。

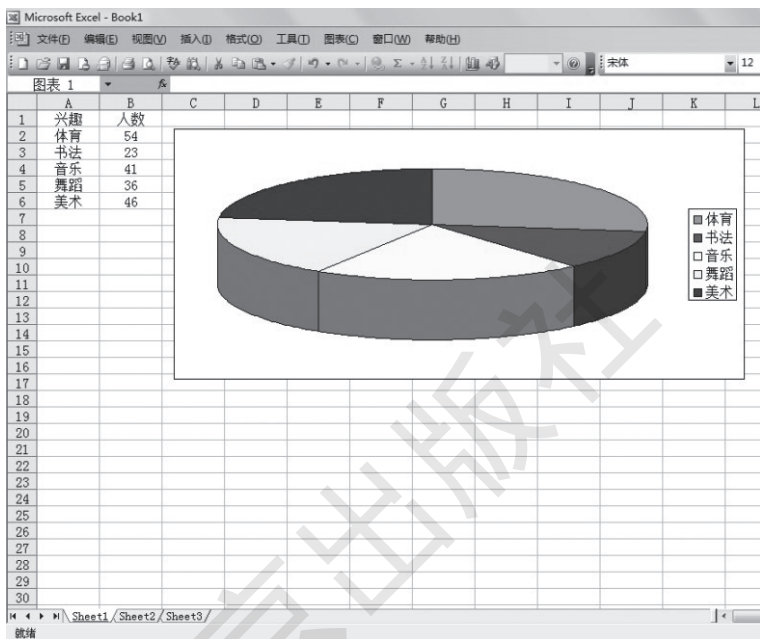


图 2-19 学生兴趣爱好调查饼图

(三) 折线图的绘制

例 14 使用 Excel 绘制折线图。

解：具体步骤如下。

(1) 打开 Excel 工作表，在工作表的 A1、B1 单元格中分别输入“年份”“国内生产总值”字样，在其下方区域 (A2: B6) 分别输入相应的数值。

(2) 执行“插入”—“图表”命令，弹出“图表向导—4 步骤之 1—图表类型”对话框。在“标准类型”选项中选择“折线图”，再选择其中的子图表类型 (本例选择第四个)。具体如图 2-20 所示。

(3) 单击“下一步”按钮，弹出“图表向导—4 步骤之 2—图表源数据”对话框。在“数据区域”栏中输入数据所在单元格区域，本例中输入 B1: B6；在“系列”选项卡中，于“分类 (X) 轴标志”栏中输入 A2: A6。

(4) 单击“下一步”按钮，弹出“图表向导—4 步骤之 3—图表选项”对话框。根据需要，分别在“坐标轴”“网格线”“图例”“数据标志”“数据表”等选项中做出选择和修改。最后单击“完成”，便可得到图 2-21 所示的折线图。

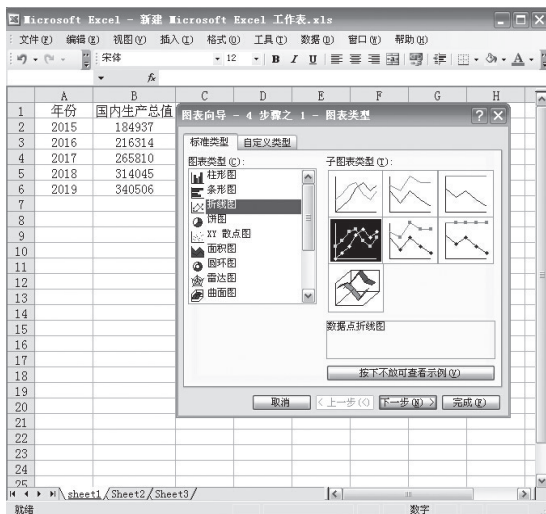


图 2-20 选择折线图类型

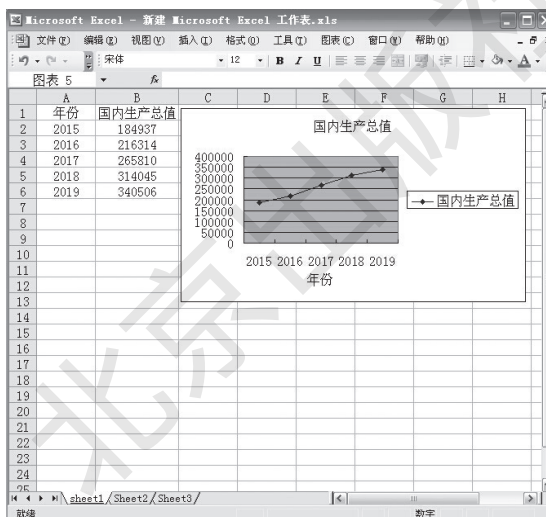


图 2-21 2015—2019 年国内生产总值折线图

实际上，饼图和折线图的绘制步骤与条形图基本相同，只是选择的图表类型不同而已。

(四) 直方图的绘制

例 15 Excel 对某超市 45 种食品的重量的测量结果进行统计分组，并得到频数分布表和直方图。

解：具体步骤如下。

(1) 打开 Excel 工作表，在 A1: E9 单元格中输入 45 种食品的重量。在 A11 单元格输入“各组上限”字样，然后在 A12: A18 单元格分别指定各组上限数值，本例中 7 个组的上限分别是 89、99、109、119、129、139、149。

(2) 选择“工具”菜单栏，执行“数据分析”命令，在弹出的“数据分析”对话框中选择“直方图”，如图 2-22 所示。

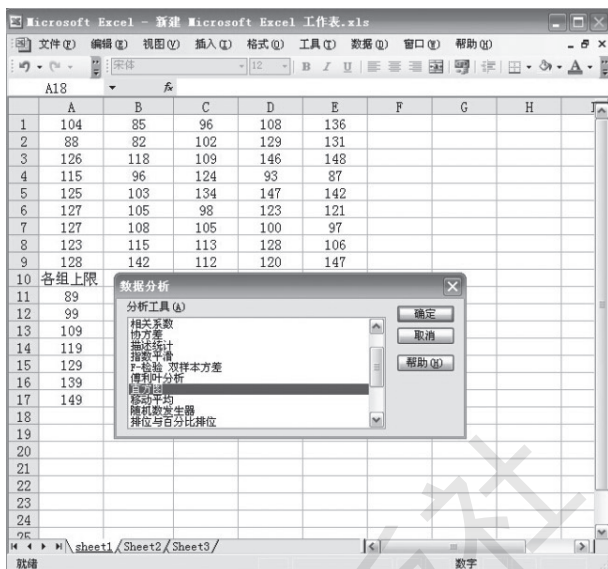


图 2-22 “数据分析”对话框

(3) 单击“确定”按钮，弹出“直方图”对话框。在“输入区域”栏中输入原始数据所在区域，本例为 A1: E9；在“接收区域”栏中输入指定的各组上限所在区域，本例 A12: A18（如果引用范围包括标志值的名称，须勾选“标志”选项）；在“输出区域”栏中指定频数分布表输出区域的起点单元格，本例为 B11；再勾选“图表输出”选项，如图 2-23 所示。最后，单击“确定”按钮，即可得到频数分布表和直方图，如图 2-24 所示。

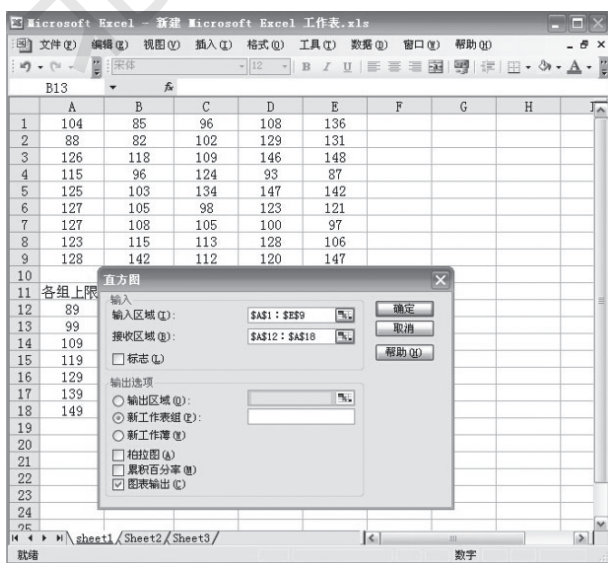


图 2-23 “直方图”对话框

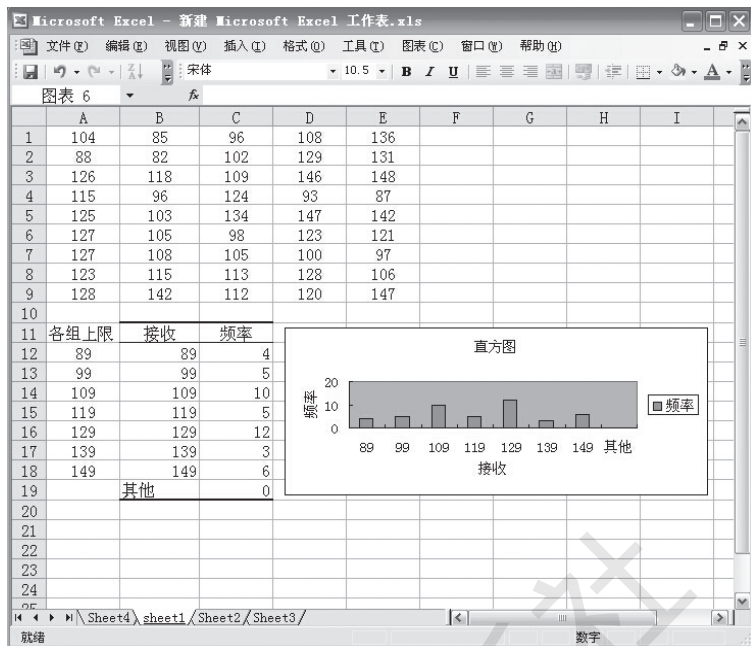


图 2-24 频数分布表和直方图

基本能力拓展

根据提供的训练资料和相应的训练要求,用已学过的统计调查理论和相关的方法技巧,分析一些具体的社会经济问题。

拓展目标 1

掌握统计调查方案的制定、调查问卷的设计和统计资料的搜集过程。

拓展资料

本班同学的生活消费状况,包括学习、餐饮、衣着、日用品、社会交往、文体活动等方面的消费。

拓展要求

1. 制定在校大学生消费的调查方案。
2. 设计在校大学生消费状况调查表。
3. 在本校范围内组织问卷调查。

拓展目标 2

掌握统计资料整理的整个过程。

拓展资料

拓展目标 1 中在校大学生消费状况问卷调查的原始数据。

拓展要求

1. 按照统计整理的要求, 进行原始资料的审核、分组、汇总。
2. 运用适当的统计表和统计图显示统计整理的结果。
3. 分析大学生消费水平、构成、指标间相互依存关系等基本状况。

课后习题

一、单项选择题

- (1) 统计调查按调查单位在时间上进行登记的连续性不同, 可分为()。
 - A. 全面调查和非全面调查
 - B. 经常性调查和一次性调查
 - C. 统计报表和专门调查
 - D. 定期调查和不定期调查
- (2) 调查几个主要的大型发电厂, 就可以了解我国电力生产的基本情况, 这种调查属于()。
 - A. 普查
 - B. 重点调查
 - C. 典型调查
 - D. 抽样调查
- (3) 统计调查按其组织形式分类, 可分为()。
 - A. 普查和典型调查
 - B. 重点调查和抽样调查
 - C. 统计报表和专门调查
 - D. 经常性调查和一次性调查
- (4) 调查项目的承担者是()。
 - A. 调查对象
 - B. 调查项目
 - C. 调查单位
 - D. 填报单位
- (5) 所选择单位的标志总量占全部总体标志总量的比例较大, 这些单位就是()。
 - A. 调查单位
 - B. 代表性单位
 - C. 重点单位
 - D. 典型单位
- (6) 调查项目是指()。
 - A. 调查方案
 - B. 调查对象
 - C. 调查内容
 - D. 调查单位
- (7) 商店库存盘点, 调查人员亲自搜集资料的方法是()。
 - A. 大量观察法
 - B. 采访法
 - C. 报告法
 - D. 直接观察法
- (8) 对某地工业设备普查时, 调查对象是()。
 - A. 该地全部工业设备
 - B. 全部工业企业
 - C. 每个工业企业
 - D. 每台工业设备
- (9) 确定调查时间是()。
 - A. 确定调查资料所属时间
 - B. 确定调查登记的时间
 - C. 确定调查期限
 - D. 确定进行调查的时间
- (10) 抽样调查的目的是()。
 - A. 了解总体的全面情况
 - B. 掌握总体的基本情况
 - C. 由样本指标推断总体指标
 - D. 由个别推断总体

- (11) 由 12 名工人分别看管机器台数资料如下: 2、5、4、4、3、4、3、4、4、2、2、4。按以上资料编制变量数列, 应采用()。
- A. 单项变量数列 B. 组距变量数列
C. 等距变量数列 D. 不等距变量数列
- (12) 对连续型组距数列, 凡是某单位的标志值刚好等于相邻两组上下限数值时, 一般是()。
- A. 将此值归入上限所在组
B. 将此值归入下限所在组
C. 将此值归入上限所在组或下限所在组
D. 另立一组
- (13) 把一个班级学生先按性别分组, 再按年龄分组, 这种分组是()。
- A. 按品质标志分组 B. 按数量标志分组
C. 简单分组 D. 复合分组
- (14) 按某一标志分组的结果, 表现出()。
- A. 组内同质性和组间差异性 B. 组内差异性和组间差异性
C. 组内同质性和组间同质性 D. 组内差异性和组间同质性
- (15) 组距、组限和组中值之间的关系是()。
- A. 组距 = (上限 - 下限) ÷ 2 B. 组中值 = (上限 + 下限) ÷ 2
C. 组中值 = (上限 - 下限) ÷ 2 D. 组限 = 组中值 ÷ 2
- (16) 某连续变量数列, 其末组为开口组, 下限为 500, 又知其邻组组中值为 480, 则末组组中值为()。
- A. 490 B. 500 C. 510 D. 520
- (17) 统计分组的关键在于()。
- A. 正确选择不同特征的品质标志和数量标志
B. 确定组距
C. 选择统计指标和统计指标体系
D. 选择分组标志和划分各组界限
- (18) 某同学考试成绩为 80 分, 应将其计入()。
- A. 成绩为 80 分以下的人数中 B. 成绩为 70 ~ 80 分的人数中
C. 成绩为 80 ~ 90 分的人数中 D. 根据具体情况来具体确定
- (19) 将统计表分为总标题、横行标题、纵栏标题和指标数值四部分是()。
- A. 从表式结构看 B. 从内容上看
C. 从作用上看 D. 从性质上看
- (20) 统计分组是指对()。
- A. 总体单位的分组 B. 总体的分组
C. 总体与总体单位的分组 D. 一个分组标志的分组

二、多项选择题

- (1) 统计调查的基本要求是()。
- A. 准确性 B. 具体性 C. 及时性
D. 全面性 E. 系统性
- (2) 在工业企业设备调查中()。
- A. 工业企业是调查对象 B. 每台设备是调查单位
C. 每台设备是填报单位 D. 每个企业是填报单位
E. 工业企业的设备是调查对象
- (3) 下列是非全面调查方式的有()。
- A. 普查 B. 重点调查 C. 典型调查
D. 抽样调查 E. 统计报表
- (4) 按搜集调查资料的方式,可分为()。
- A. 直接观察法 B. 报告法 C. 采访法
D. 通讯法 E. 估计法
- (5) 一个完整的统计调查方案应包括()。
- A. 确定调查目的 B. 确定调查对象
C. 确定调查项目 D. 确定调查时间
E. 确定调查的组织实施计划
- (6) 在组距式数列中,组中值()。
- A. 是上限和下限之间的中点数
B. 是用来代表各组标志值的平均水平
C. 在开口组中无法确定
D. 在开口组中,可以参照邻组的组距来确定
E. 就是组平均数
- (7) 统计分组的作用在于()。
- A. 反映总体的基本情况 B. 说明总体的数量特征
C. 区分现象的类型 D. 研究现象的结构
E. 分析现象之间的依存关系
- (8) 统计调查表有()。
- A. 简单表 B. 简单分组表 C. 单一表
D. 一览表 E. 复合分组表

三、填空题

(1) 统计调查按搜集资料的组织方式不同可分为_____和_____；按调查对象包括的单位多少不同可分为_____和_____；按调查单位在时间上登记的连续性不同可分为_____和_____。

- (2) 普查从调查对象的范围上看是_____。
- (3) 调查时间是_____, 调查期限是_____。
- (4) 统计分组同时具有两方面的含义: 对总体而言是_____, 即将总体区分为性质_____的若干部分; 对总体单位而言是_____, 即将性质相同的总体单位合在一起。
- (5) 变量数列的两个构成要素是_____和_____。
- (6) 从内容结构上看, 统计表由主词和_____两部分组成。
- (7) 对于开口组的组距与组中值, 原则上是以_____的组距来确定。
- (8) 统计分组是根据统计研究的需要, 将统计总体按照一个或几个标志划分_____的一种统计方法。
- (9) 对于连续型变量只能进行组距式分组, 并且相邻组的组限必须_____, 采用_____的原则。
- (10) 统计分组的关键在于选择_____。

四、计算题

- (1) 某班 40 名学生的考试成绩如下(表 2-17)。

表 2-17 某班 40 名学生的考试成绩

89	88	76	99	74	60	82	60	93	99	94
82	77	79	97	78	87	84	79	65	98	67
59	72	56	81	77	73	65	66	83	63	89
86	95	92	84	85	79	70				

学校规定: 60 分以下不及格; 60 ~ 75 分为中; 76 ~ 89 分为良; 90 ~ 100 分为优。试将该班学生分为不及格、中、良、优四组, 编制一张次数分布表。

- (2) 某班学生统计学原理考试成绩次数分布如表 2-18 所示。

表 2-18 某班学生统计学原理考试成绩资料

考分(分)	人数(人)	比率(%)	向上累计		向下累计	
			人数(人)	比率(%)	人数(人)	比率(%)
60 以下	2					
60 ~ 70	7					
70 ~ 80	11					
80 ~ 90	12					
90 以上	8					
合计	40					

要求: 根据上表资料, 计算相应的数字, 填入表中空格, 并说明各指标的意义。

- (3) 某公司所属 20 家企业某月工业增加值资料如表 2-19 所示。

表 2-19 20 家企业某月工业增加值资料

企业编号	工业增加值 (万元)	企业编号	工业增加值 (万元)
A	46	K	24
B	68	L	78
C	118	M	92
D	33	N	57
E	79	O	40
F	50	P	60
G	89	Q	72
H	27	R	58
I	127	S	66
J	99	T	74

要求：进行汇总，编制组距数列。

(4) 某百货公司连续 40 天的商品销售额如表 2-20 所示。

表 2-20 某百货公司连续 40 天的商品销售额

单位：万元

41	25	29	47	38	34	30	38	43	40
46	36	45	37	37	36	45	43	33	44
35	28	46	34	30	37	44	26	38	44
42	36	37	37	49	39	42	32	36	35

要求：根据上面的数据进行适当分组，编制频数（频率）分布表，并绘制直方图。