西部人民的生活

中国西部省份社会与经济发展地图集

本地图集仅用于课题内部交流，不得公开引用

中国西部省份社会与经济发展地图集

张化枫 Jon Pedersen

NRCSTD Fafo AIS
中国科学技术促进发展研究中心
挪威Fafo应用国际研究所

本书编委会名单：

主任：

王元　Jon Pedersen

副主任：

王奋宇　赵延东　Kristin Dalen

编委：

邓大胜　马缨　薛姝　何光喜　谢桂华　张化枫　樊立宏　王俊峰　邓雪明

课题组成员：

中方：

组长：王奋宇 中国科技促进发展研究中心

成员：赵延东 中国科技促进发展研究中心

马缨 中国科技促进发展研究中心

何光喜 中国科技促进发展研究中心

邓大胜 中国科技促进发展研究中心

薛姝 中国科技促进发展研究中心

樊立宏 中国科技促进发展研究中心

王俊峰 中国科技促进发展研究中心

邓雪明 中国科技促进发展研究中心

刘冬梅 中国科技促进发展研究中心

谢桂华 中国人民大学

朱斌 北京大学

钟元俊 南开大学云南研究院

陈文江 兰州大学

白雪锋 北京美兰德社会经济调查信息研究所

挪方：组长：**Jon Pedersen**

成员：**Jon Hanssen-Bauer**

Guri Tyldum

Kristin Dalen

张化枫

Anne Hatløy

Christophe Gironde

Ole Fossgaard

Amsale Temesgen

Linn-Kaia Rogstad

Gro Hasselknippe

David Dury

Sara Randall

科技部研究中心和 Fafo 在此感谢所有西部地区与我们合作开展此次调查项目的所有大学的教职员工和学生，所有接受调查的家户和个人，所有中央和地方帮助我们开展此次调查的机构和部门，以及提供我们此次调查项目经费的挪威外交部。

目录

概述 1

中国西部地区人口分布 5

人口 6
西部地区人口结构示意图 7
性别比例和人口结构 7
少数民族分布 8
出生性别比 9
家户成员的流动 9
家户成员数 10
城市化进程 11

健康和卫生 12
总生育率 12
熟练保健人员接生的分娩(城市户口) 13
熟练保健人员接生的分娩(农村户口) 13
家中分娩（城市户口） 14
家中分娩（农村户口） 14
产前保健 15
产后保健 15
艾滋病毒/艾滋病的认识 16
儿童接种疫苗 16
抽烟和喝酒（男性） 17
抽烟和喝酒（女性） 17
生病求医 18
医疗保险/保障 18
医疗保险/保障（城市户口） 19
医疗保险/保障（农村户口） 19
公共医疗卫生服务（城市） 20
公共医疗卫生服务（农村） 20

劳动人口和就业 21
劳动参与率 21
失业率（男性） 22
失业率（女性） 22
失业率：详细分布图 23
年轻人的失业率 24
失业率按年龄和性别的分布 24

失业者求职途径 25
获得工作的方式 25
第二职业（男性） 26
第二职业（女性） 26
非农业部门就业 27
职业和非农业部门就业（男性） 27
职业和非农业部门就业（女性） 28
雇主类型（男性） 28
雇主类型（女性） 29
行业和妇女的份额 29
妇女在非农业部门就业 30
劳动合同 31
带薪产假 31
雇主提供的正式培训 32
雇主拖欠工资 32

教育 33
总体入学率 33
学到五年级的学生 33
7至14岁儿童小学的总体入学率（城市） 34
7至14岁儿童小学的总体入学率（农村） 34
识字率（男性） 35
识字率（女性） 35
15至24岁人口识字率（男性） 36
15至24岁人口识字率（女性） 36
教育和妇女 37
完成的教育程度 37
学校教育（男性） 38
学校教育（女性） 38
学杂费和上学总支出 39
优惠政策 39
电脑的普及 40
母语和其他语言的掌握 40
小学和初中的覆盖率（城市） 41
小学和初中的覆盖率（农村） 41

住房和基础建设 42
饮用水煮沸 42
饮用水的安全性和可持续性（城市） 43
饮用水的安全性和可持续性（农村） 43
饮用水源的距离（城市） 44

饮用水源的距离（农村） 44
居住拥挤度 45
居住拥挤度：详细分布图 45
改良厕所和生活垃圾处理（城市） 46
改良厕所和生活垃圾处理（农村） 46
固定电话 47
移动电话 47
住宅的结构类型 48
农村地区的道路 48

收入和财富 49
年家庭总收入和资产指数 49
资产指数：详细分布图 50
家户的耐用消费品 51
耐用消费品 51
收入分配不均 52
收入的基尼系数和分配 52
贫困 53
经济状态自评 54
经济状态展望 54
负担基本经济支出的能力 55
最低生活保障（MLG） 55
政府转移支付 56
贫困救济 56

农业 57
土地分配 57
耕地面积的基尼系数和分配 57
耕地面积和类别 58
土地灌溉 58
草场 59
围栏草场 59
农户从事的其他农业活动 60
农业机械化 60
租赁土地 61
出租土地 61
雇佣农业劳动力和劳动力成本 62
农业收入 62
农作物 63
耕地和农作物种类 63
好收成 64

坏收成 64
家畜和家禽 65

环境..... 66
自然灾害 66
自然灾害的分布 66
洪涝灾害 67
政府自然灾害援助 67
由于征地或者自然灾害而失去土地 68
失去土地的补偿 68
失去房屋（城市） 69
失去房屋（农村） 69
对房屋损失的补偿 70
草地退化 70
参与退耕还林项目 71
经济损失和退耕还林的补偿 71
炉灶和能源（城市） 72
炉灶和能源（农村） 72
冬天取暖能源（城市） 73
冬天取暖能源（农村） 73
小汽车和卡车 74
拖拉机 74

附录：抽样方案说明..... 75

概述

“西部省份社会与经济发展监测研究”（Monitoring on Social and Economic Development in the Western Regions of China，简称MEDOW）是专门针对我国西部省份城乡居民的生活和生产状况开展的一项研究活动，它主要是基于人口抽样调查开展的一项社会与经济发展政策研究。这里的西部省份包括：甘肃、青海、宁夏、陕西、内蒙古、新疆、云南、贵州、四川、广西、重庆等11个省（市、区），鉴于西藏的特殊性，我们在本项目中没有把西藏包括进来。

本地图集是MEDOW项目的成果之一，所有的统计指标和数据均来自于该项目的实地调查。该项研究由挪威外交部提供主要资助，科技部中国科技促进发展研究中心（以下简称：科技部研究中心）提供相应的配套资金和条件支持，由科技部研究中心负责组织实施，由挪威Fafo应用国际研究所（Fafo Institute for Applied International Studies 以下简称：Fafo）派出的若干专家提供方法与技术上的支持。

科技部研究中心是中国科学技术部直属的综合性政策研究机构。1982年10月由国务院批准建立。同年，邓小平先生亲笔为研究中心题名。该机构主要从事国家科学技术发展战略、政策、体制、管理、预测、评价以及科技促进经济社会发展等方面的研究，为国家科技、经济、社会发展的宏观决策提供咨询和建议。

科技部研究中心与Fafo研究所的合作关系已有10余年，在上世纪90年代，科技部研究中心与Fafo研究所开展了“中国劳动力流动与就业及其政策应用”的合作研究，这项研究围绕我国改革开放以来城市劳动就业领域出现的新情况、新变化，从劳动力的从业模式、职业生涯和城市新移民等方面，对我国的劳动就业市场建设以及相关的社会政策进行了深入的调查研究，并对劳动力市场的培育和发展提出了有价值的政策建议（王奋宇，李路路, 2001）。

此后，中挪双方决定就中国西部省份的社会与经济发展状况进行监测研究。下文将简要介绍本项目的实施过程、主要特点和本报告的若干技术说明。

一、项目界定

由于自然、历史、社会等原因，长期以来西部地区经济发展相对落后。从2000年起，中央政府开始正式实施西部大开发战略，动员中央部委和地方各级政府共同投入力量，来积极应对西部地区面临的巨大挑战：比如，污染、水资源枯竭、土地退化等环境问题，产业结构调整和国有企业改革带来的经济阵痛，农业增产和农民增收的现实困难，以及建立和完善社会保障体系的迫切需要。

我们首先对该项研究进行了定位——该项研究是为了配合我国的西部开发战略开展的一项基于人口抽样调查的社会与经济发展政策研究。

这一定位意味着我们必须注重三个方面的问题：一是作为一项政策研究，必须密切联系我国在西部发展方面已有的政策环境，以及政府各部门在西部开发战略部署中面对的政策问题和信息需求，调查结果应当尽可能地回答中央和地方政府相关部门关于政策制定过程中的问题，并最大限度地满足他们的信息需求；二是该项研究尽管是通过国际合作的方式开展的，但是我们在注重研究内容的国际比较视角之外，还要注重密切联系西部各省份城乡居民社会和经济发展的实际情况，调查内容的设计必须是西部居民生产生活状况的真实反映；三是严格遵循当今社会科学研究的科学规范，在抽样方法、调查实地组织、质量控制以及后期的数据分析等阶段要做到科学和客观。

后面的两个问题是属于项目设计过程中的任务，而第一个问题则与我们的项目研究内容界定密切相关，因此，我们在项目界定阶段走访了中央政府的10余个与西部大开发政策制定密切相关的部委（局、办），主要了解目前针对西部大开发的政策环境和现状，了解他们在制定政策过程中所遇到的主要问题和政策需求，并请他们对我们的项目提出意见和建议。

通过走访，我们发现在西部大开发政策制定过程中当时主要面临两方面的问题；一是关于西部（乃至全国）的发展状况的信息数据主要反映的是省、市、县等层次上的情况，在村（社区）、家户和个人层次上的数据信息相当缺乏，这大大影响了政策制定的精细程度；二是部门之间的已有数据信息要么不全要么难以整合，各部门一般侧重于与本部门工作直接相关的数据收集，而忽略相关因素的综合分析，由此带来的主要问题是政策的一致性和关联性不足。现实生活的复杂要求我们在考虑某项单一问题时必须关照其他看似并不相关的因素，并将这些众多的因素统合考量，例如，居民健康的问题不仅仅是医疗基础设施全不全的问题，它同时还与居民的生活方式、居住方式、教育水平、收入水平以及自然环境等众多因素相关，但是这些相关因素之间又是什么样的关系呢？它们之间的关联程度有多大？各因素中最为关键的因素又是什么呢？对此类问题的回答要求我们把所有这些相关因素的信息放到同一个分析数据库中进行分析。

根据走访过程中发现的主要问题，我们对研究内容的界定进行了大量的讨论，并同时成立了一个由15人组成的专家顾问组参与讨论，该顾问组包括了来自大学、社科院、国务院发展研究中心、政府部门政策研究机构的专家学者，涉及到社会学、经济学、农村发展、区域发展研究等若干学科。最终的研究内容界定就是在这些讨论之后逐步形成的。

二、项目设计

项目设计工作包括：调查内容（问卷）设计、抽样方法设计、调查和研究组织设计等。为了使设计的问题能

够契合西部城乡居民的日常生活和生产的实际情况，课题组在2002年上半年用了近两个月的时间走访了青海、内蒙古、新疆、宁夏、云南、四川、贵州7个省（自治区）的200多个城乡居民家庭，进行设计前的探索性研究。与此同时，我们还走访了若干村、居委会或社区，对于城乡居民的居住方式，居住地的居民管理方式进行了考察，这些探索性的研究为我们后来的问卷设计和抽样方法设计提供了大量有益的帮助。

问卷设计、抽样设计和实地组织设计是整个设计工作中的三个重点环节。

问卷设计工作既要保证所要收集信息的准确性，又要能够使受访者在理解上不产生歧义，而且我们所关注的问题又十分广泛。虽然过去15年里Fafo在居民生活状况调查方面积累的经验（Hanssen-Bauer, Pedersen and Tiltnes 1997）给我们提供了很好的参考，但是中国的情况有着自己的一些特点，调查设计也需要作相应的调整。在开始实地调查之前，我们对调查工具进行了精心设计和完善。从2001年开始，科技部研究中心和Fafo的研究人员多次到西部地区进行探索性研究和试调查，以此来确定调查哪些主题内容才可以最好地概括西部地区人们的生活状况。

在问卷设计和项目设计的过程中，我们还咨询了国内外的许多机构、组织和专家，他们提出了非常有价值的意见和建议。

我们一共设计了12个大的问题模块，包括：家户成员人口学特征、受教育与培训状况、劳动就业情况、迁移流动情况、营养健康和享受医疗服务情况、农业生产情况、家户收入和经济状况、住房和生活基础设施、育龄妇女生育历史与产前产后照顾情况、社会参与、社会交往、政策评价等。

最终形成的调查工具由一组调查问卷组成，共分四大部分：

- 家户问卷

  家户问卷由该家户中了解情况的成年人来回答。家户问卷又分成两部分：一部分是关于家户整体的情况，另一部分是关于家户成员个人的情况。家户整体的情况主要包括住房条件、基础设施、家庭经济、农业活动、土地退还、自然灾害等。成员个人的情况主要包括每个人的性别、年龄、婚姻、迁移、教育、健康和劳动就业。另外，还设计了一个人体测量表用于评估每个人的营养健康状况。

- 妇女儿童问卷

  我们规定家户成员中所有年龄介于15至54岁的妇女都要回答此问卷。如果家户中有5岁以下的小孩而该小孩的母亲又不住在当前住所，那么由这个小孩的主要照顾者来回答相应部分的问题。这样，我们可以收集到所有适龄妇女和儿童的信息。对于妇女，主要调查她们的怀孕和生育经历；对于儿童，主要调查儿童免疫情况。

- RSI问卷

  每户中将随机抽取一位15岁以上的成年人（Randomly Selected Individual, RSI）回答RSI问卷。这些问题主要包括更为详细的个人就业情况、身体健康和生活习惯、对家人教育和公共话题的看法。我们仅选取一个人回答这些问题，避免了逐一询问所有家庭成员带来的困难。不过，本报告仅有少量表格使用了RSI问卷的调查数据。

- 社区问卷

  社区问卷用于了解家户所生活的社区环境状况。包括道路、电力、通信等基础条件，也包括学校、诊所等公共服务设施，还包括自然条件、自然灾害、自然资源管理等方面的问题。这份问卷由居委会或村委会的主要负责人来回答。

抽样设计工作是保障调查所收集的数据代表性的关键环节。我们希望所收集的数据能够推论的目标人口为所有目前居住在中国西部省份（西藏除外）的家户——无论他们的户口是否在那里或是否在那里进行了登记，包括所有流动人口，但不包括现役军人。我们在每一个省（市、自治区、直辖市）采用了“概率与规模成比例抽样”（PPS）的方法，以2000年第五次人口普查的数据为基础进行抽样，共抽取44,000户居民，每个省抽取4,000户。在具体抽取家户的方法设计中，我们采用了列表与制图相结合的方法，将列表时点上所有居住在被抽中的初级抽样单位范围内的居民全部列入表中，如果基层居民组织没有现成的居民登记表或登记表不完整，抽样人员则通过画图的方法，将区域内所有有人居住的建筑物标记在图上，并同时将建筑物中的所有居民按户列出，以保证完全没有遗漏，然后再在列出的表中随机抽取一定数量的家户作为访问对象。

本次调查的抽样采取的是相对标准的家户抽样方案。抽样框以中国2000年人口普查分省的街道、乡镇人口为基础，首先按照概率与规模成比例（PPS）的方法抽取了2772个单元，然后从每一个选中的街道、乡镇中再抽取一个居委会或行政村。为了便于实地工作的开展，有些居委会或行政村又被进一步划分成一定规模的“片”（Segments）。在被抽中的村（居）委会或片，列表员负责抄录居住在该辖区范围内的所有家户的花名册，或者通过制图和逐一走访的方式记录下所有的住户清单。

在列表过程中，我们特别强调要把在调查时点在辖区范围内住的所有家户都包括进来，尤其是那些短期的流动人口。比如在村子里修路的工人，尽管他们在调查期间只是临时住在这里，我们也要求列表员必须列在表中。在实地调查时，我们又要求调查员收集那些不在家的成员信息，比如到城里打工的农民，到外地上学的孩子，以及流动人口他们老家的那些家人。

通过强调这些操作规程，我们获得了一个包括流动人口的完整样本。但是，这对如何给“家户”下定义增加了难度。首先，属于某个家户的个人并不一定住在这个家（调查地的当前住所）中，可能在外上学或者工作。其次，从理论上讲，在抽样过程中一个人有可能被列为两个家户的成员，一个是他目前真实居住的那个家，一个是他的“老家”。在附录的抽样设计方案中，将详细讨论这个问题。

组织设计则是一个更为庞大的系统工程。由于涉及的省份有11个，我们在甘肃的兰州大学和位于昆明的南开大学云南研究院设置了两个区域办公室，分别负责协调西北和西南地区的实地调查和协调工作。在每一个省又分别设置了两个相互独立的工作系统，一个是实地调查系统，一个是质量控制系统。同时，由于调查中需使用全球定位测量仪（GPS）、身高测量仪和体重测量仪，为合理调配和使用这些设备，同时考虑到抽样列表的从实地到省级办公室再到实地的往返以及问卷的运送，我们还建立了相应的

物流控制系统，以保证从抽样列表到实地访谈到数据录入等环节的衔接和顺畅。

为了检验我们的设计工作情况，课题组在2004年初在甘肃和云南两省分别进行了试调查，对项目的问卷设计、抽样设计和组织设计做了全面的测试。此后，我们又对先前的设计做了进一步的修改和调整，并撰写了详细的《抽样列表员工作手册》、《调查员工作手册》、《督导员工作手册》、《数据录入员工作手册》等文件。

三、实地工作

实地工作是本项研究的数据信息收集工作的主体。由于涉及地域广泛、人口众多。我们在各省主要依靠了当地高校师生的力量进行数据收集工作。在11个省我们先后动员了近20所高校的师生，共有近2,000人参与了该项调查的实地工作。

实地调查工作从2004年7月10日开始，到2005年2月28日结束。由于各种客观条件的限制，各省的实地调查无法同时展开，所以最早和最晚开始调查的两个省份之间间隔了大约半年（如图0所示）。

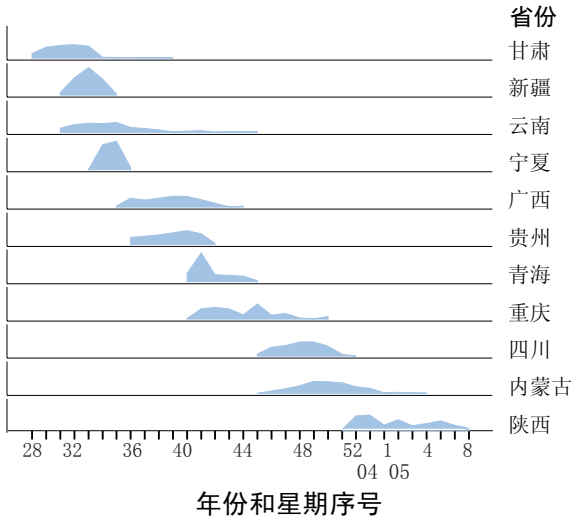


图0：各省的调查进度：每周完成的问卷数量

实地工作分为培训、抽样列表、入户访谈和数据录入几个环节。

首先我们分别对各省招募到的师生进行了一周到10天不等的抽样列表培训，又分别对各省的调查员进行了两到三周的调查员入户访谈培训，然后分别对各省的督导员、编辑检查员、数据录入员等工作人员进行了专门培训。

抽样列表和入户访谈工作同时展开，这是整个实地调查工作的核心内容。所有抽样列表员按照要求进入抽中的社区（村或居委会），进行列表和制图，省级办公室在列出的表上抽取样本户，调查员小组再进入社区进行入户访问，在访谈过程中既要访问家户中最为了解家户情况的成员，又要对家户中的所有育龄妇女、5岁以下儿童的照顾者、以及随机抽取的家户成员个人（Randomly Selected Individual, RSI）进行分别访谈，还要对家户中的所有成员测量身高和体重，与此同时，小组督导员将通过访问村或居委会的负责人，填写一份社区问卷，以了解本社区的基础设施、公共服务、社区发展等基本情况。

所有调查问卷填写完成后将及时迅速送达省级办公室，以对所有问卷进行编辑检查，一旦发现问题，立即予以解决。在所有问卷编辑检查完成后，被迅速送达区域办公室，进行数据录入。

由于问卷涉及的调查内容比较多，因此每次访问所需时间也较长。所有完成的访问中，每个家户平均的访问时间为1小时35分钟，中位值是1小时30分钟。

总体来说，大部分的受访者都非常配合我们的调查。本次调查共抽中44,738个家户，其中具备受访条件的共41,695个家户，实际完成调查的共41,222个家户，收集的信息涉及人口总计167,000余，分布在西部11个省（市、自治区）的128个市（地、州）805个县（区、市）2336个乡、镇、街道的2,707个村（居）委会。在所有抽中的44,738个家户中，有880户被确定为无法访问，因为调查员发现这些住所没有人居住或者住所已经被拆除。

所有被抽中且真实存在的家户中，应答率为94%，拒访率仅有0.96%，绝大部分的无应答情况是由其他因素造成的。这些问卷没有完成的原因有多种，比如，调查员无法找到正确的地址，或者连续三次上门都发现家中没有人或没有适合受访的，或者访问被打断而无法继续（这也可以算是拒访的一种形式），或者调查员确信受访者提供的信息是虚假的或错误的。

四、数据录入、清理与分析

数据录入工作与实地访谈工作同步展开，在2004年12月下旬，其中10个省的实地工作已基本完成，只有最后一个省因开始时间较晚，又加连续的大雪导致无法到达调查地而耽搁到2005年2月底完成，全部数据录入工作随后在3月份完成。课题组又用了大约6个月的时间进行数据清理（data cleaning）工作，并形成了最终的工作数据库。

本报告是应用该数据库完成的数据报告，我们还将陆续出版分析报告和相关的专题报告。

五、本研究特点

本项目在研究内容和方法上具有这样几个特点：

第一，研究内容强调多模块指标的综合，强调对模块问题的多因素分析，这一点有助于改善目前政府部门分部门收集数据、分部门决策的现状。在政府决策过程中，由于信息资源难以共享导致决策的部门化特征，其中一个很重要的原因是因为我们目前收集数据的方式往往是部门化的。

第二，强调信息的多层次分析。该项研究获得的数据结果既可以在个人层次上进行分析，也可以在家户层面上分析，同时还可以在社区层次上分析，11个省的数据可以分别对11个省的情况在省的层次上进行推论，进而提出各省的分析研究报告。分析单位的多层次对于我们认识西部地区的总体情况提供了更为丰富的分析视角。

第三，强调数据信息的政策价值。如果项目的数据不能为政府政策所用，那么就失去了进行该项研究的意义。在项目完成后，已经有若干政府部门提出希望使用我们的数据，以支持其进行规划战略和政策研究。

第四，指标使用的国际化。在挪威专家的帮助下，我们在很多模块中都使用了可以直接进行国际比较的指标口径，如，失业率的指标使用了国际劳工组织（ILO）的指标口径，关于文盲的定义我们使用了联合国教科文组织（

UNESCO）的口径，关于儿童计划免疫的情况我们采用了世界卫生组织（WHO）的描述框架等等。通过使用国际比较指标，可以对某些领域更方便地进行国际比较研究和国际对话。

第四，抽样方法上是以居住地而非户口登记地为基础的抽样，将全部流动人口包括了进来。我国传统的抽样调查方法通常是按照人们的户口所在地进行抽样的，但随着人口流动的加速，户籍制度对于人们迁移行为的约束越来越弱。流动人口在流入地的生产和生活更多地影响流入地（而非流出地）的经济与社会发展，因此流入地政府在制定公共政策时必须将流动人口的问题考虑进来，而传统的抽样方法则不能满足这一要求。可喜的是，我国统计部门也已开始在抽样方法上开展了对这一方法的尝试。

第五，城乡统一的指标、抽样和分析框架。我国传统的抽样社会调查通常将城市和农村分别对待，对城市和农村的研究分别采用不同的抽样样本，不同的问卷和不同的分析框架。但随着城市化进程的加快，这一方法的弊端已十分明显，在传统意义上的农村社会，可能存在着大量的非农生产活动，因此，在我们的指标和抽样设计中采用了城乡统一的设计方法。实际上，由于我们收集了足够的信息，因此在分析中可以用户籍、居住地、职业等多种指标来进行“农”或“非农”的分类。统一的抽样指标和框架，为我们认识和分析城乡之间的异同提供了可能，避免再出现城乡二元分割的经济和社会发展政策研究结论。

第六，使用了全球定位测量仪（GPS）测量各调查点的基本地理信息，这一方法的应用可以帮助我们进行地理信息系统（GIS）的分析，将收集到的数据信息进行地理空间描述和分析，这对于政府部门制定区域规划具有重要的参考价值。

第七，调查中少数民族语言的使用。因为西部地区是我国少数民族分布最为广泛的区域，在某些民族地区，民间通用的语言是本民族语言，为了调查沟通的顺畅和准确，我们还将汉语问卷分别译成了藏语和维吾尔语，为调查带来了很大的益处。

六、地图集

本册地图集所收集的地图和图表建立在MEDOW项目实地调查所获得的统计指标和数据的基础上，覆盖了项目所涉及的11个省（市、自治区）。大多数地图中，每个省（市、自治区）都被赋予某种单一的颜色来表示其各项指标的具体数值。而有部分地图使用一系列连贯的阴影在整个调查的西部省市区域内表示某些指标的分布情况。这些地图通过首先计算各个抽样组中某项指标特征的比例，然后加权平均所有邻近的抽样组中该指标的比例（使用距离反比例加权），最终估计出整个地区的分布情况。使用这个方法进行估计，抽样组的密度在一个地区越大，估计的精确度就越高。如果一个地区没有或只有很少的抽样组，就无法获得有用的估计。特别值得注意的是，人口密度会影响某些地区指标估计的准确性，例如11页的城市化分布图。特别是青海省的西部，内蒙古自治区的北部和新疆维吾尔自治区的东南部，由于人口分布较稀疏，这几个地区指标分布估计的准确性相对较差。

七、指标和定义

如前所述，我们在设计问卷的时候，努力保证问卷的设计反映西部省（市、自治区）的实际情况，而同时又考虑到社会经济发展指标的国际标准，以及与中国国家统计局现有的统计标准具有可比性。但是有的时候很难同时兼顾这三个目标，而且在很多情况下，并没有国际标准或者现有的中国国家统计局的统计标准与我们的数据相对应。

因此，我们在这里罗列出在地图集中所使用的一些比较重要的指数：

西部省（市、自治区）：本地图集所使用的11个省（市、自治区）来自中国国家统计局2000年人口普查报告。地图边界和各省（市、自治区）的分界线取自“2000年中国县/区人口普查资料汇编大全”，由授权对外发布中国国家统计局可公开数据的北京华通人市场信息有限责任公司出版。地图的边界线仅用于数据描述的便利性，中国科学技术促进发展研究中心（NRCSTD）和挪威FAFO应用国际研究所不对地图上所使用的区域和边界的准确性负任何法律责任。

城乡地区：地区指的是一个家户或者个人是否居住在城市或者农村地区。家户所在地是根据受访的家户接受调查时的住所所在社区类型来划分的：如果该社区是村委会，则定义为农村；如果该社区不是村委会（通常叫做居委会、家委会或社区），则定义为城镇。城乡的划分主要依据中国国家统计局2000年人口普查的分类方法，同时根据实地的观察对2000年以后的变化做出相应的调整。而个人居住地的城乡类型 同样依据上述标准来划分。需要说明的是，家户成员可能主要居住在（调查地的当前住所以外的）其他住所，所以个人层次的城乡类型与家户层次的城乡类型并不一定等同。

户口类型：根据我国的户口登记制度规定，户口一般分成农业户口和非农业户口两种，习惯上也称为农村户口和城镇户口。调查中以被访者自报的为准。

年龄：根据受访者的出生日期和调查日期的间隔来计算，所有的年龄均是指周岁。基于地图和图表的不同目的，会根据年龄将受访者分为不同的年龄组。

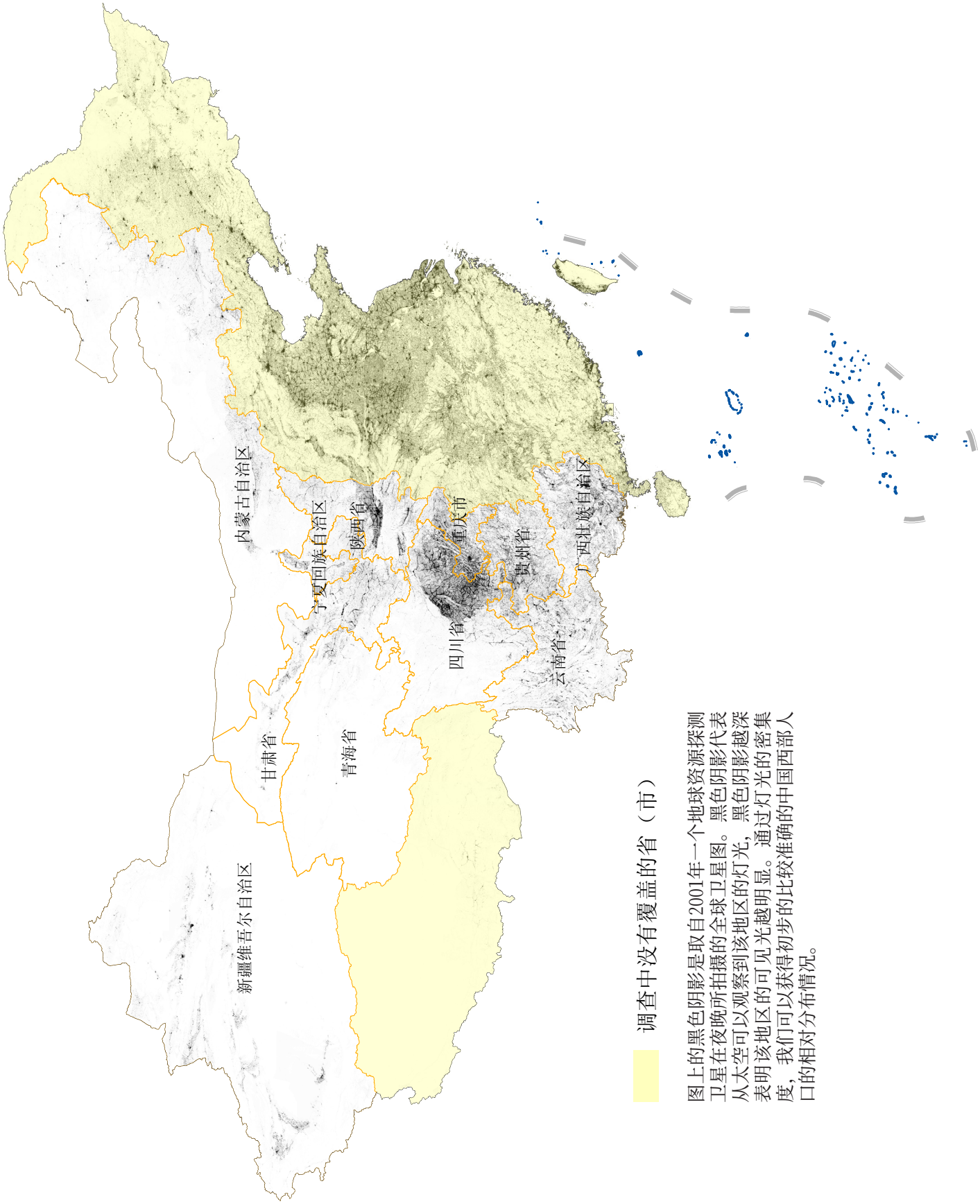
少数民族：问卷中列出了25个主要民族和“其他少数民族”作为选项。由于篇幅所限，地图中仅列出汉、壮、回、维、蒙、藏等6个人口较多的民族，其他选项合并为“其他少数民族”。

参考文献

王奋宇，李路路等著《中国城市劳动力流动——从业模
· · 2001

Hanssen-Bauer, J., J. Pedersen and A. Tiltnes [eds.] 1997, Jordanian Society, Oslo: Fafo.

中国西部地区人口分布



图上的黑色阴影是取自2001年一个地球资源探测卫星在夜晚所拍摄的全球卫星图。黑色阴影代表从太空可以观察到该地区的灯光，黑色阴影越深表明该地区的可见光越明显。通过灯光的密集度，我们可以获得初步的比较准确的中国西部人口的相对分布情况。

人口

中国人口

年份	总人口（万人）	自然增长率（‰）	男（%）	女（%）	城镇人口（%）	乡村人口（%）	少数民族（%）
1953*	59435		51.8	48.2	14.0	86.0	6.1
1964*	69458		51.3	48.7	18.3	81.7	5.8
1978	96259	12.0	51.5	48.5	17.9	82.1	
1980	98705	11.9	51.5	48.6	19.4	80.6	
1982*	100818	15.7	51.5	48.5	20.9	79.1	6.7
1985	105851	14.3	51.7	48.3	23.7	76.3	
1989	112704	15.0	51.6	48.5	26.2	73.8	
1990*	113368	14.4	51.6	48.4	26.4	73.6	8.0
1991	115823	13.0	51.3	48.7	26.9	73.1	
1992	117171	11.6	51.1	49.0	27.5	72.5	
1993	118517	11.5	51.0	49.0	28.0	72.0	
1994	119850	11.2	51.1	48.9	28.5	71.5	
1995	121121	10.6	51.0	49.0	29.0	71.0	
1996	122389	10.4	50.8	49.2	30.5	69.5	
1997	123626	10.1	51.1	48.9	31.9	68.1	
1998	124761	9.1	51.3	48.8	33.4	66.7	
1999	125786	8.2	51.4	48.6	34.8	65.2	
2000*	126583	7.6	51.8	48.2	36.2	63.8	8.4
2001	127627	7.0	51.5	48.5	37.7	62.3	
2002	128453	6.5	51.5	48.5	39.1	60.9	
2003	129227	6.0	51.5	48.5	40.5	59.5	
2004	129988	5.9	51.5	48.5	41.8	58.2	

数据来源：中国统计年鉴2005。本表各年人口未包括香港、澳门特别行政区和台湾省的人口数据。

*1953，1964，1982，1990和2000年的数据为全国人口普查数据

2004年西部各地区人口（万人）

地区	年底总人口	自然增长率（‰）
广西	4857.0	7.3
重庆	3130.0	2.7
四川	8700.4	3.1
贵州	3869.7	9.0
云南	4375.6	9.8
内蒙古	2379.6	3.1
陕西	3689.5	4.3
甘肃	2603.3	6.1
青海	533.8	10.9
宁夏	580.3	11.0
新疆	1934.0	10.8

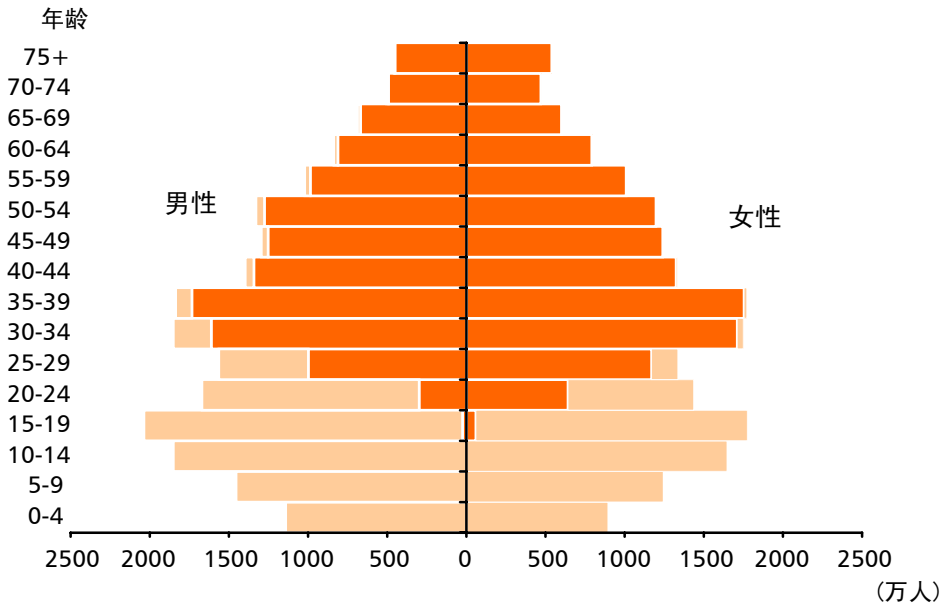
数据来源：中国统计年鉴2005

本表各年人口未包括香港、澳门特别行政区和台湾省的人口数据。

如上表所示，中国人口的自然增长率在过去的几十年中呈现递减的状态。然而相对来说，少数民族人口增长率的增长速度明显高于总人口的增长速度。此外，男性在总人口中所占的比例一直明显高于女性。最后，城市化和城镇人口的迅速增加也是中国近年来另一个非常重要的发展方向。

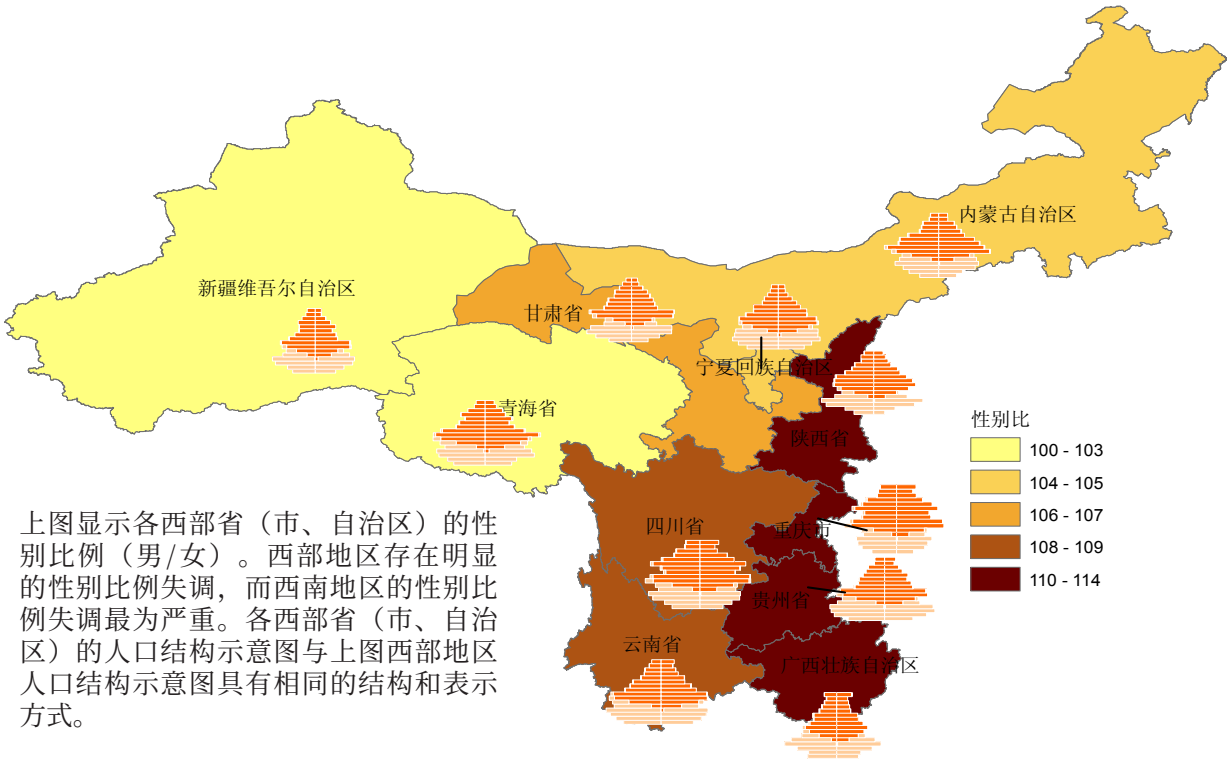
左表显示大多数西部地区的人口自然增长率都远高于中国总体的人口增长率。

西部地区人口结构示意图



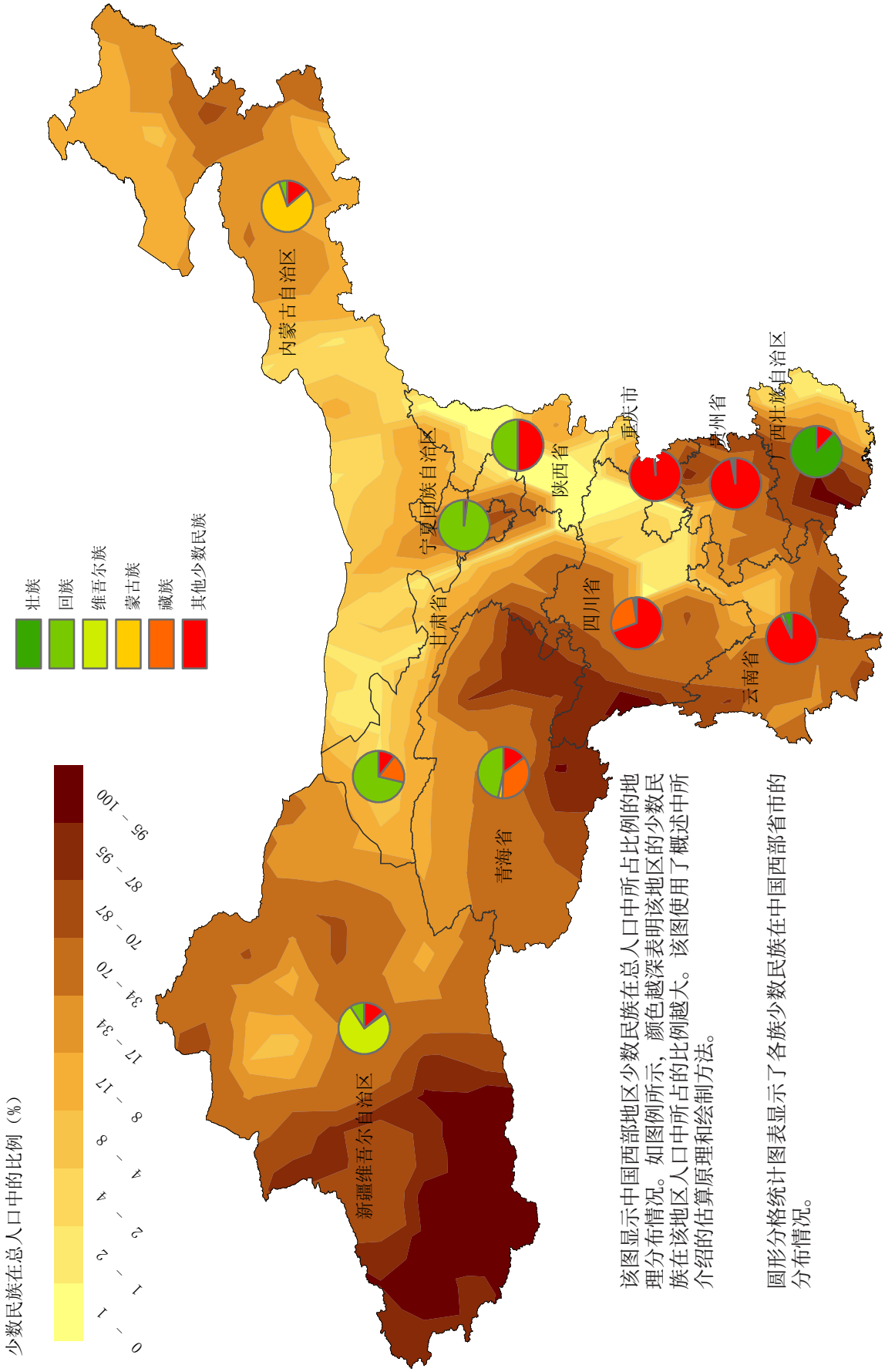
左图的人口结构金字塔示意图表现出中国西部地区人口结构历史变迁的一些重要特征：三年自然灾害；计划生育以及人们对儿子的偏好；特定性别和年龄组人群的选择性移民等。已婚或曾经结过婚的人口在图中用深色代表，而从未结过婚的人口用浅色表示。

性别比例和人口结构



上图显示各西部省（市、自治区）的性别比例（男/女）。西部地区存在明显的性别比例失调，而西南地区的性别比例失调最为严重。各西部省（市、自治区）的人口结构示意图与上图西部地区人口结构示意图具有相同的结构和表示方式。

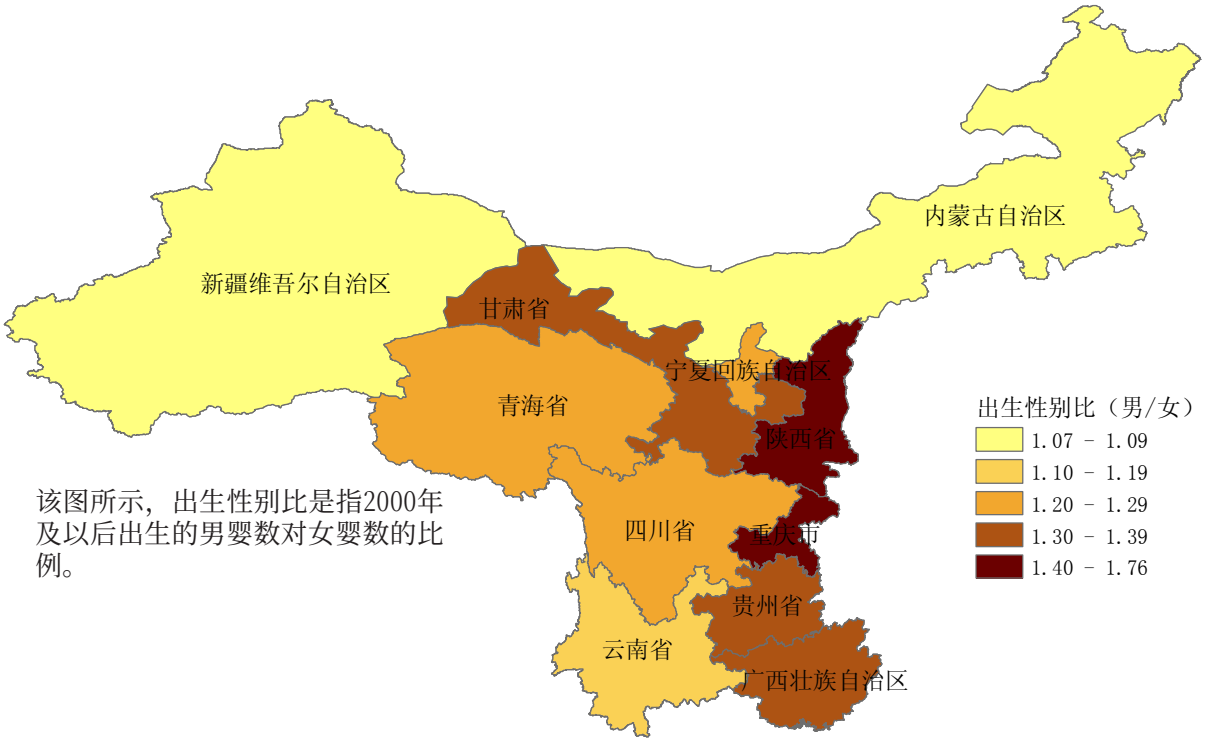
少数民族分布



该图显示中国西部地区少数民族在总人口中所占比例的地理分布情况。如图例所示，颜色越深表明该地区的少数民族在该地区人口中所占的比例越大。该图使用了概述中所介绍的估算原理和绘制方法。

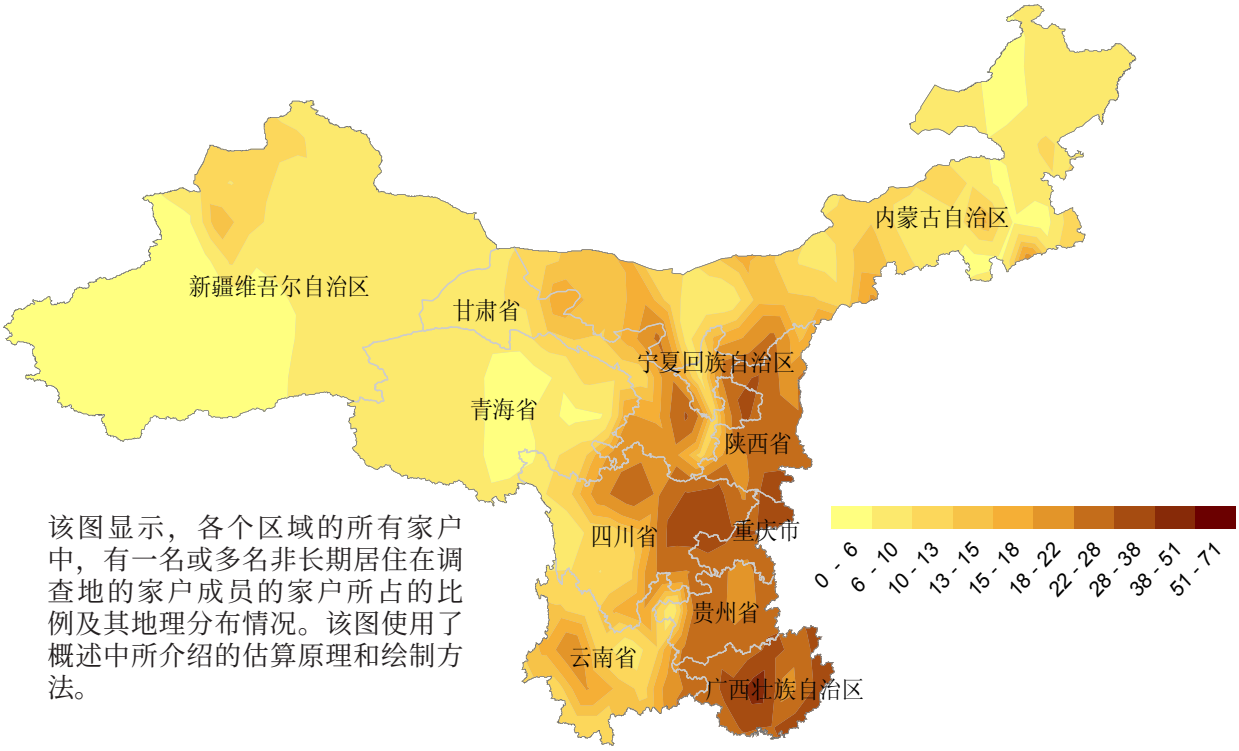
圆形分格统计图表显示了各族少数民族在中国西部省市的分布情况。

出生性别比



该图所示，出生性别比是指2000年及以后出生的男婴数对女婴数的比例。

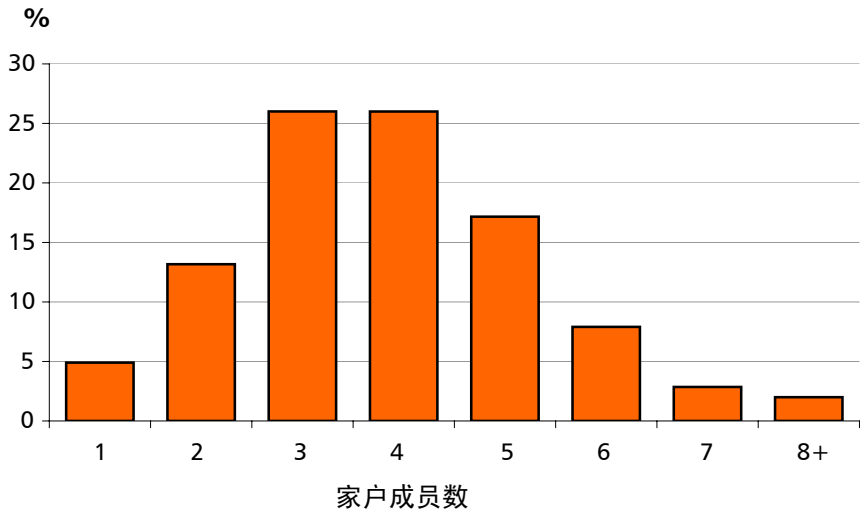
家户成员的流动

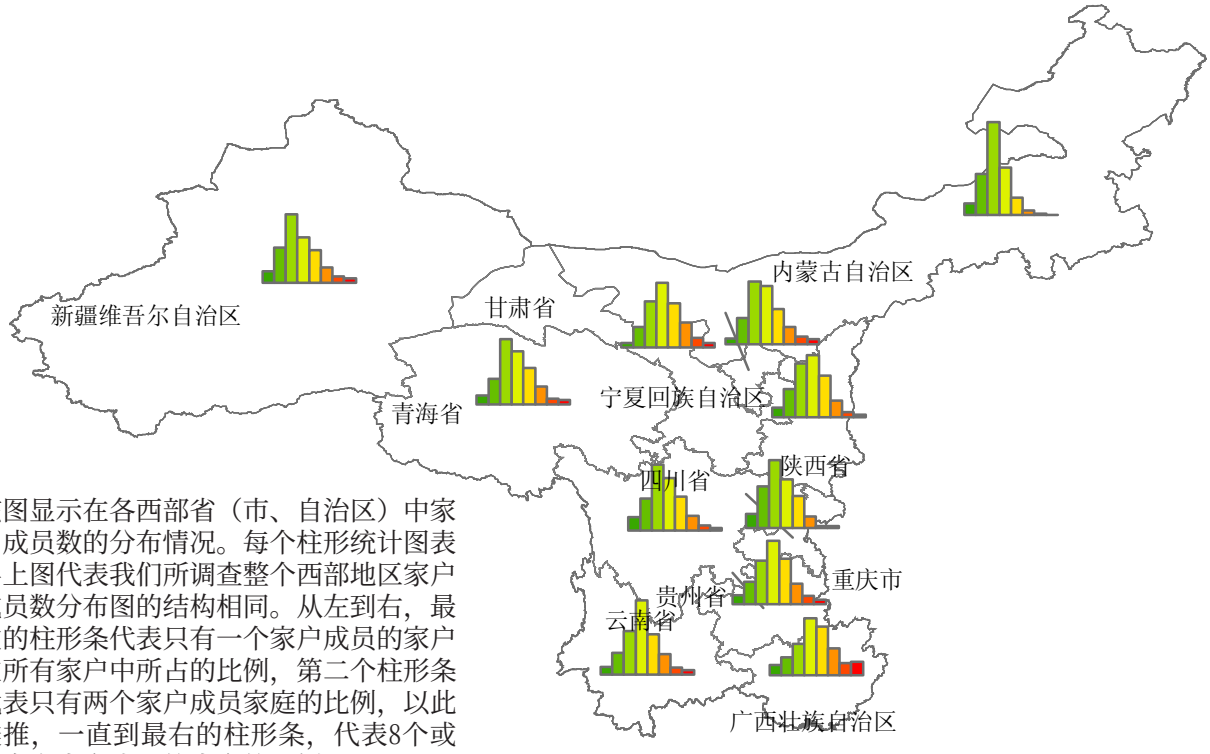


该图显示，各个区域的所有家户中，有一名或多名非长期居住在调查地的家户成员的家户所占的比例及其地理分布情况。该图使用了概述中所介绍的估算原理和绘制方法。

家户成员数

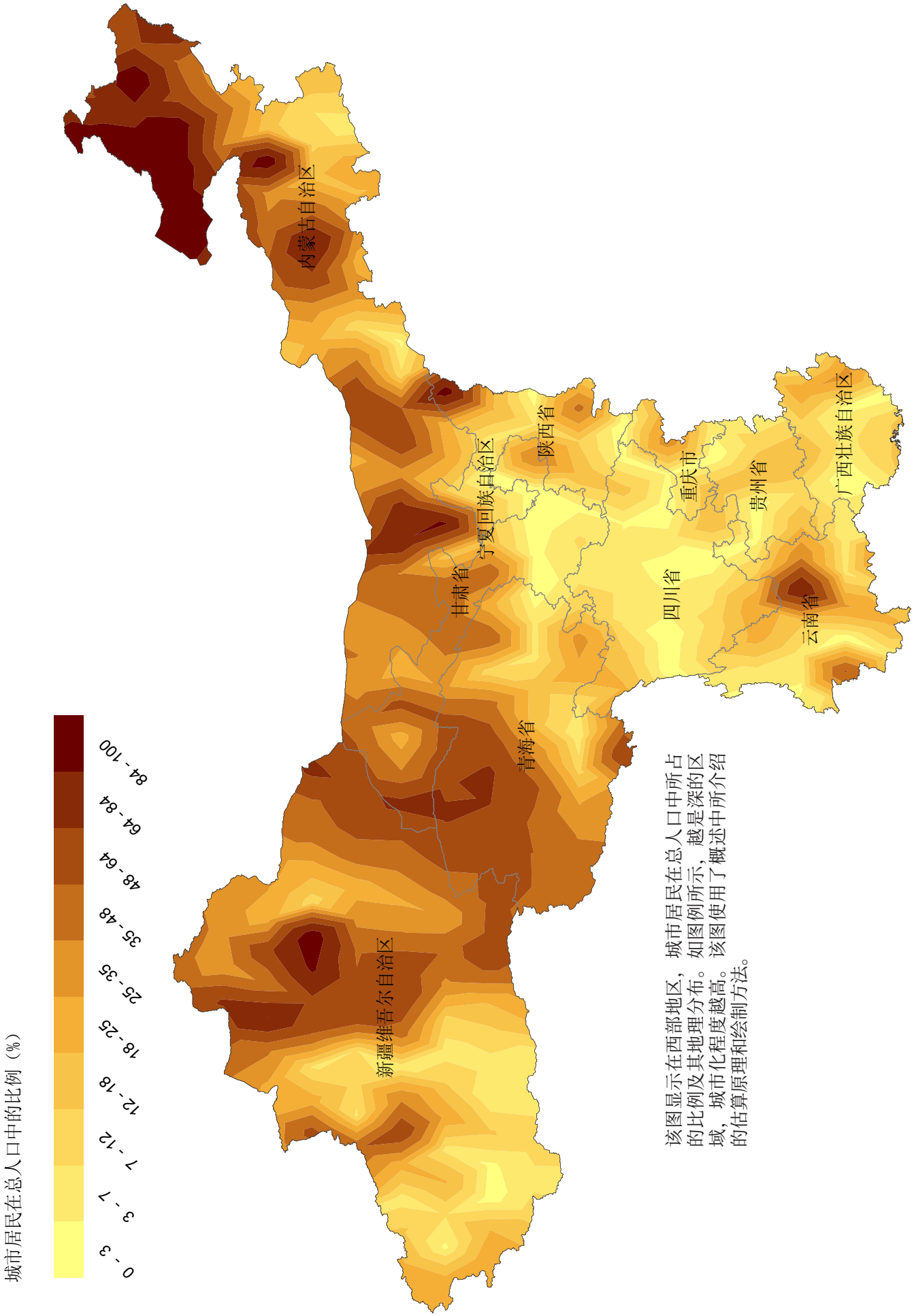
这里的平均家户成员数包括非长期居住的家户成员。平均家户成员数为3.8名成员，中位户是4名成员。





该图显示在各西部省（市、自治区）中家户成员数的分布情况。每个柱形统计图表与上图代表我们所调查整个西部地区家户成员数分布图的结构相同。从左到右，最左的柱形条代表只有一个家户成员的家户在所有家户中所占的比例，第二个柱形条代表只有两个家户成员家庭的比例，以此类推，一直到最右的柱形条，代表8个或更多个家户成员的家户的比例。

城市化进程



该图显示在西部地区，城市居民在总人口中所占的比例及其地理分布。如图例所示，越是深的区域，城市化程度越高。该图使用了概述中所介绍的估算原理和绘制方法。

健康和卫生

联合国千年发展目标

目标五：改善产妇保健

具体目标 6：从1990年到2015年，将产妇的死亡率降低四分之三

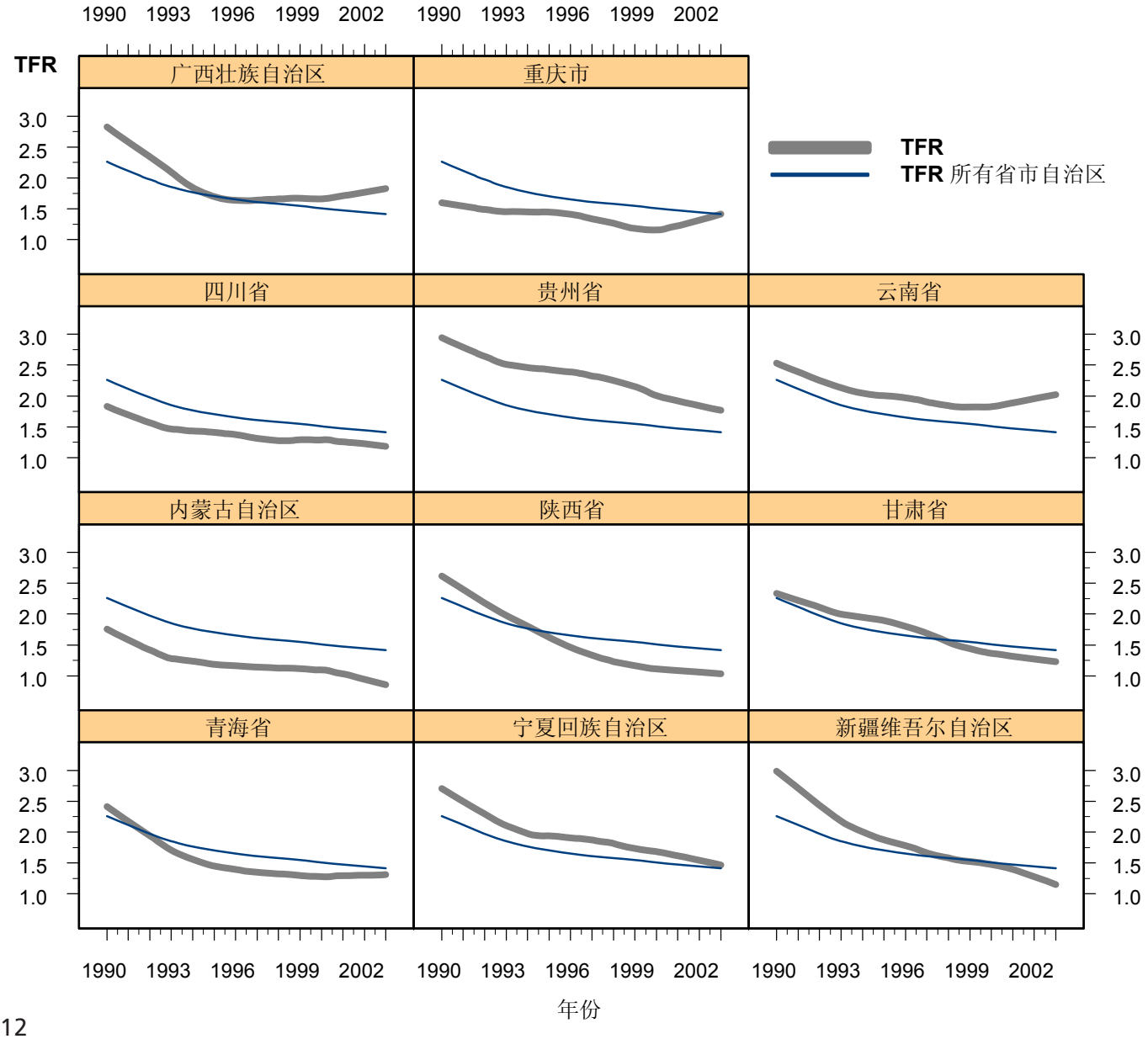
指标 17： 由熟练的保健人员接生的比例

中国西部地区熟练保健人员接生的比例：（调查前5年期间内分娩次数的百分比）	61.8%
城市户口妇女由熟练保健人员接生的比例：	94.5%
农村户口妇女由熟练保健人员接生的比例：	55.8%

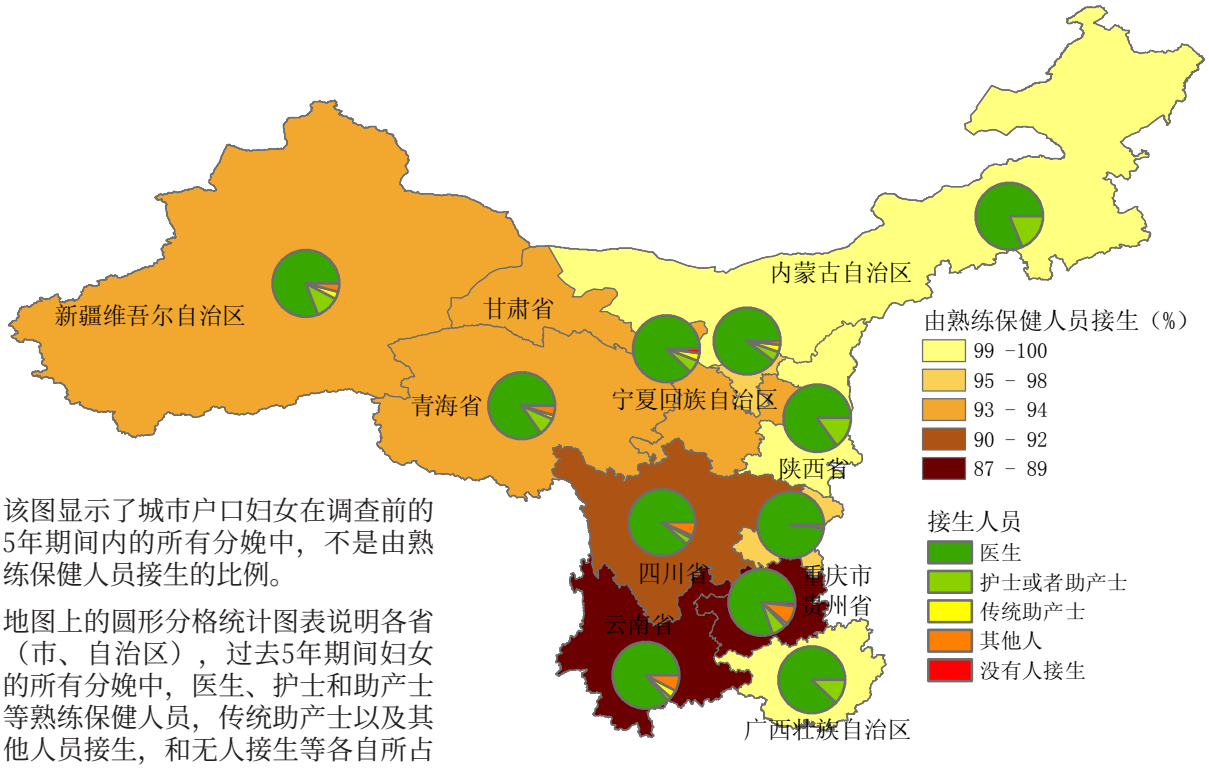
“熟练保健人员接生的分娩”是指由受过专业训练的医生，护士或者助产士接生的分娩。中国传统的助产士或接生婆不能定义为受过专业训练的熟练的保健人员。

总生育率

总生育率 (TFR) 是指在一段时期内，如某一年，计算当年每个年龄段的妇女生的婴儿数量与该年龄段妇女的数量的比率，然后平均各个年龄段的比率所得出来的该年总的妇女的生育比例。



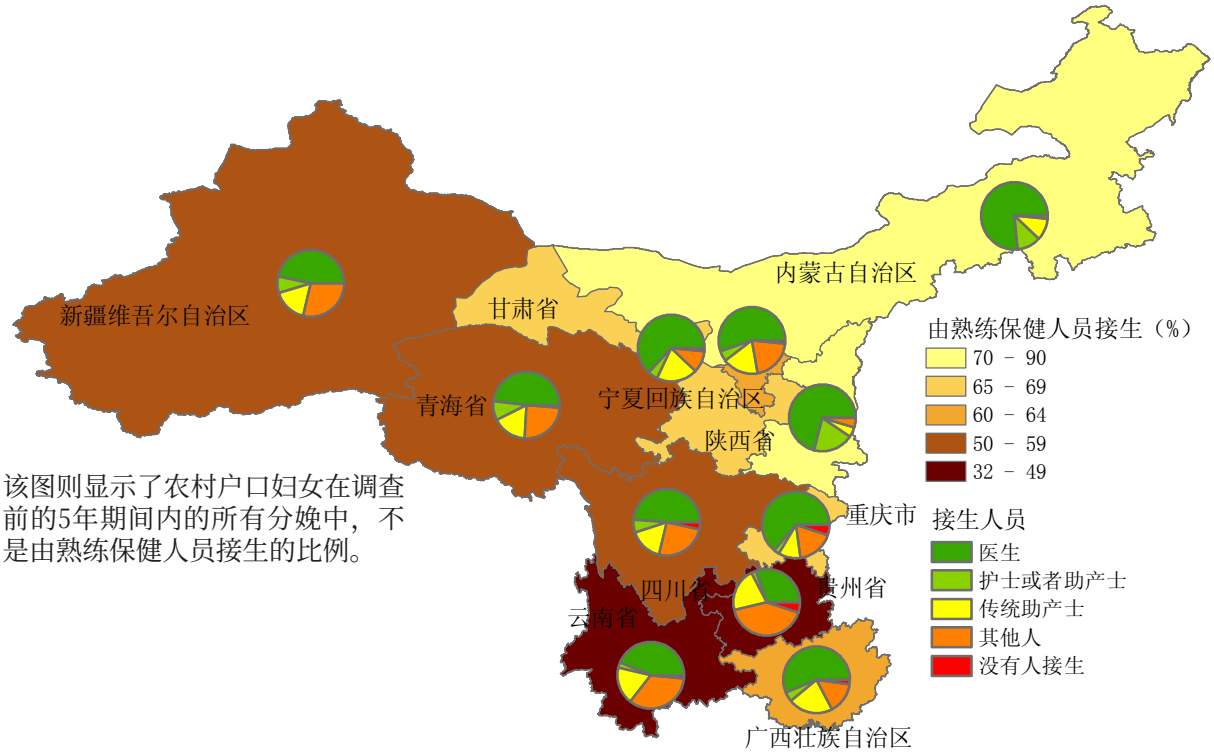
熟练保健人员接生的分娩(城市户口)



该图显示了城市户口妇女在调查前的5年期间内的所有分娩中，不是由熟练保健人员接生的比例。

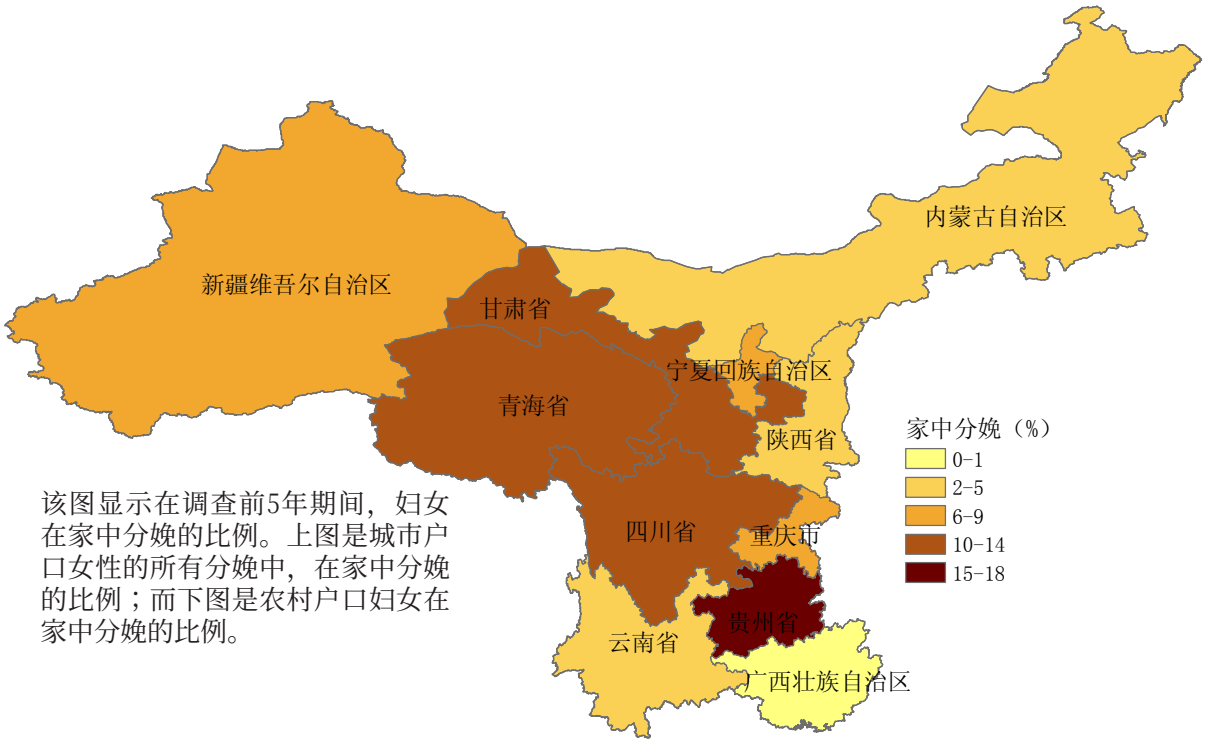
地图上的圆形分格统计图表说明各省（市、自治区），过去5年期间妇女的所有分娩中，医生、护士和助产士等熟练保健人员，传统助产士以及其他人员接生，和无人接生等各自所占的比例。

熟练保健人员接生的分娩(农村户口)



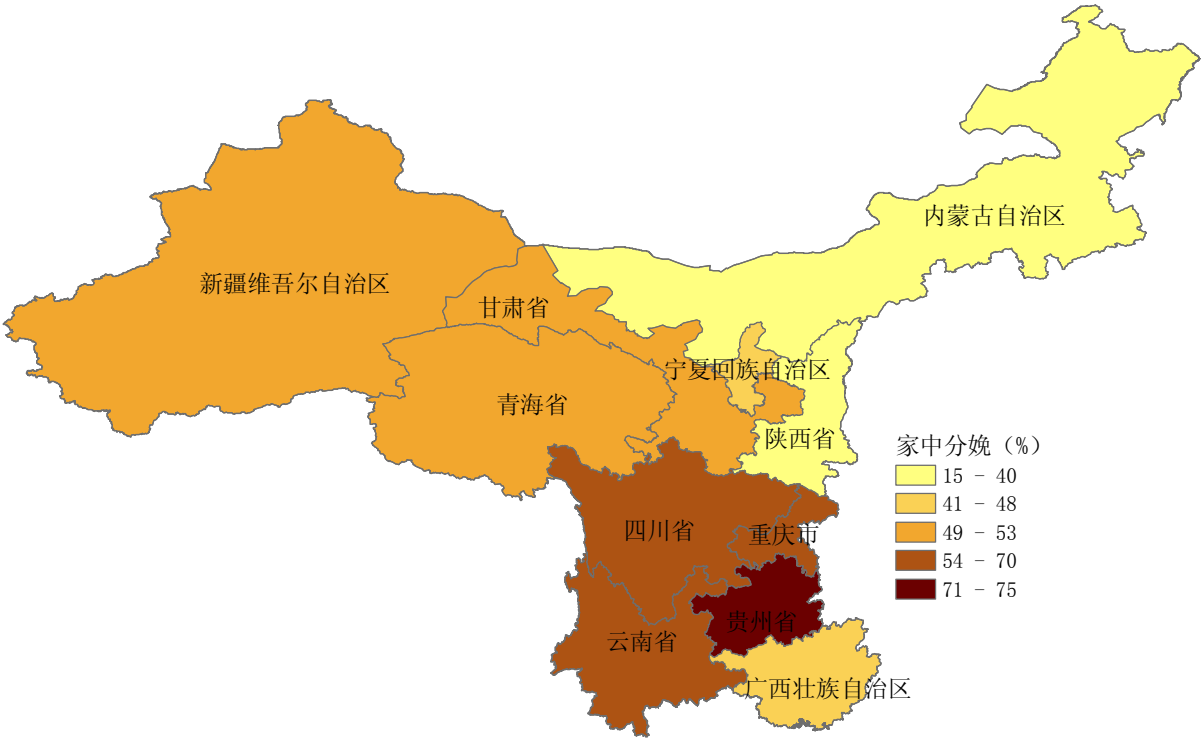
该图则显示了农村户口妇女在调查前的5年期间内的所有分娩中，不是由熟练保健人员接生的比例。

家中分娩（城市户口）

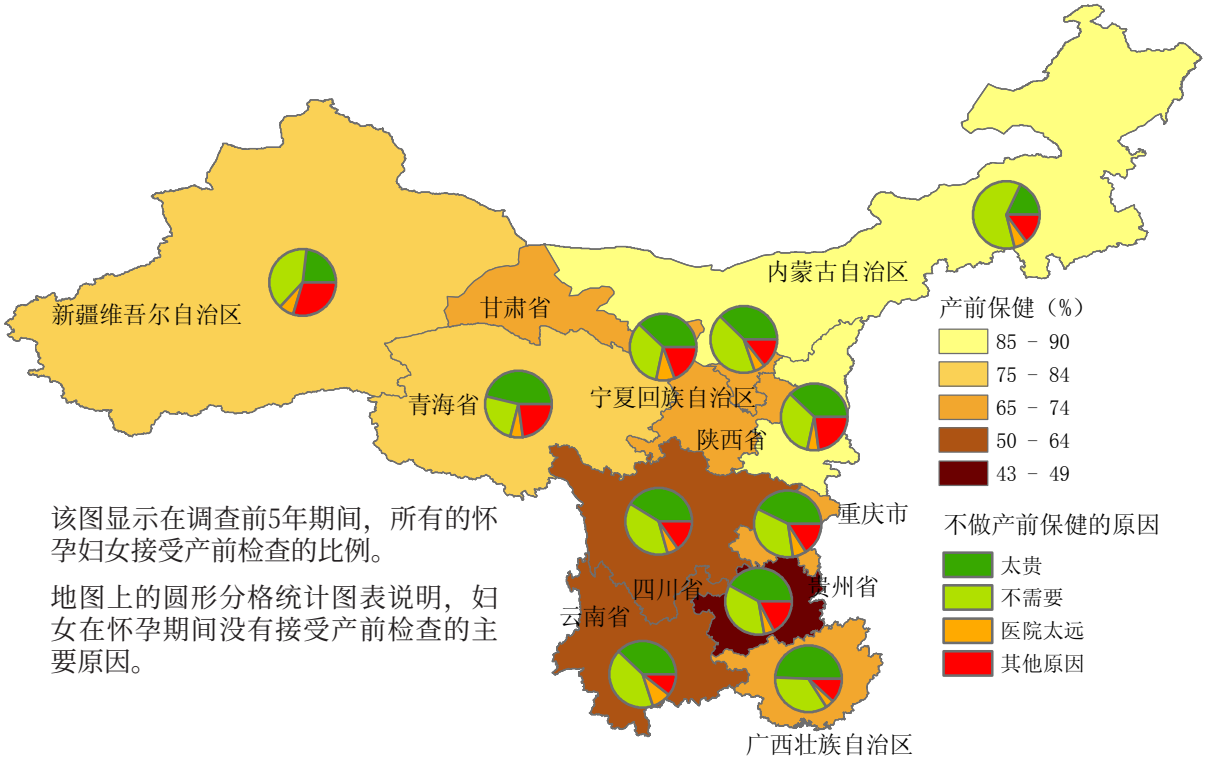


该图显示在调查前5年期间，妇女在家中分娩的比例。上图是城市户口女性的所有分娩中，在家中分娩的比例；而下图是农村户口妇女在家中分娩的比例。

家中分娩（农村户口）



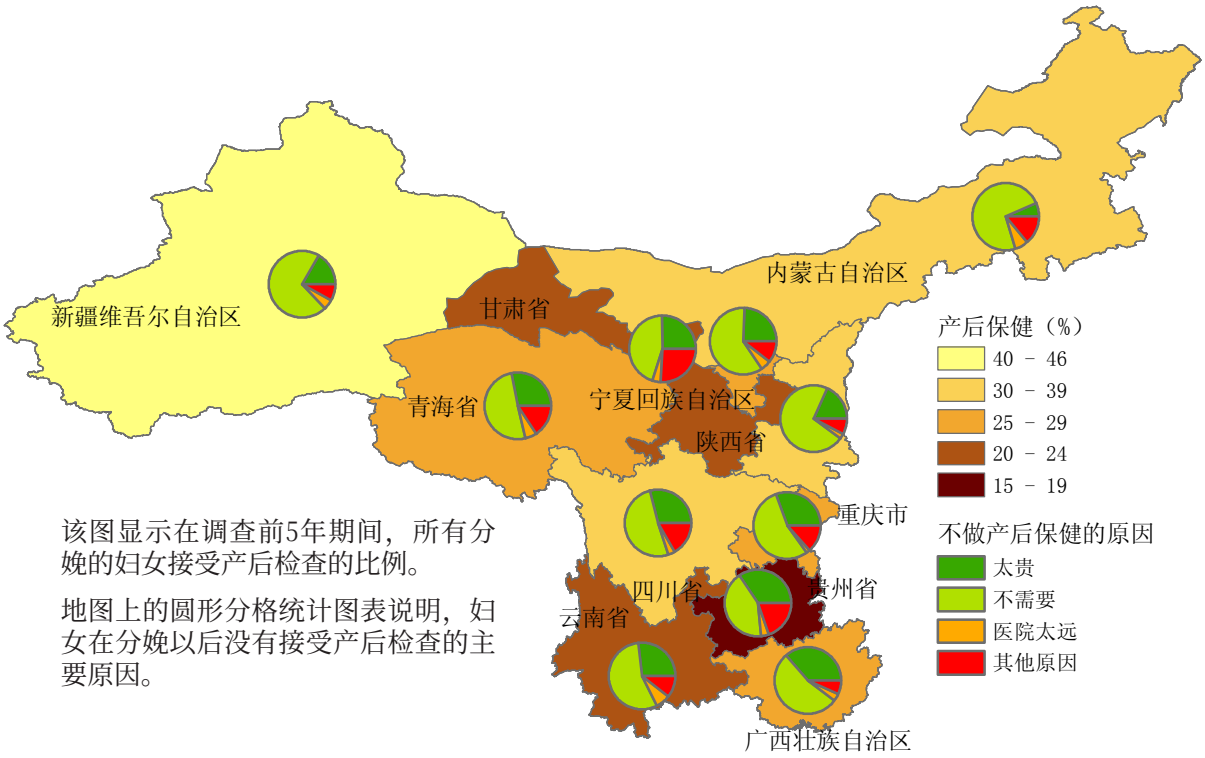
产前保健



该图显示在调查前5年期间，所有的怀孕妇女接受产前检查的比例。

地图上的圆形分格统计图表说明，妇女在怀孕期间没有接受产前检查的主要原因。

产后保健



该图显示在调查前5年期间，所有分娩的妇女接受产后检查的比例。

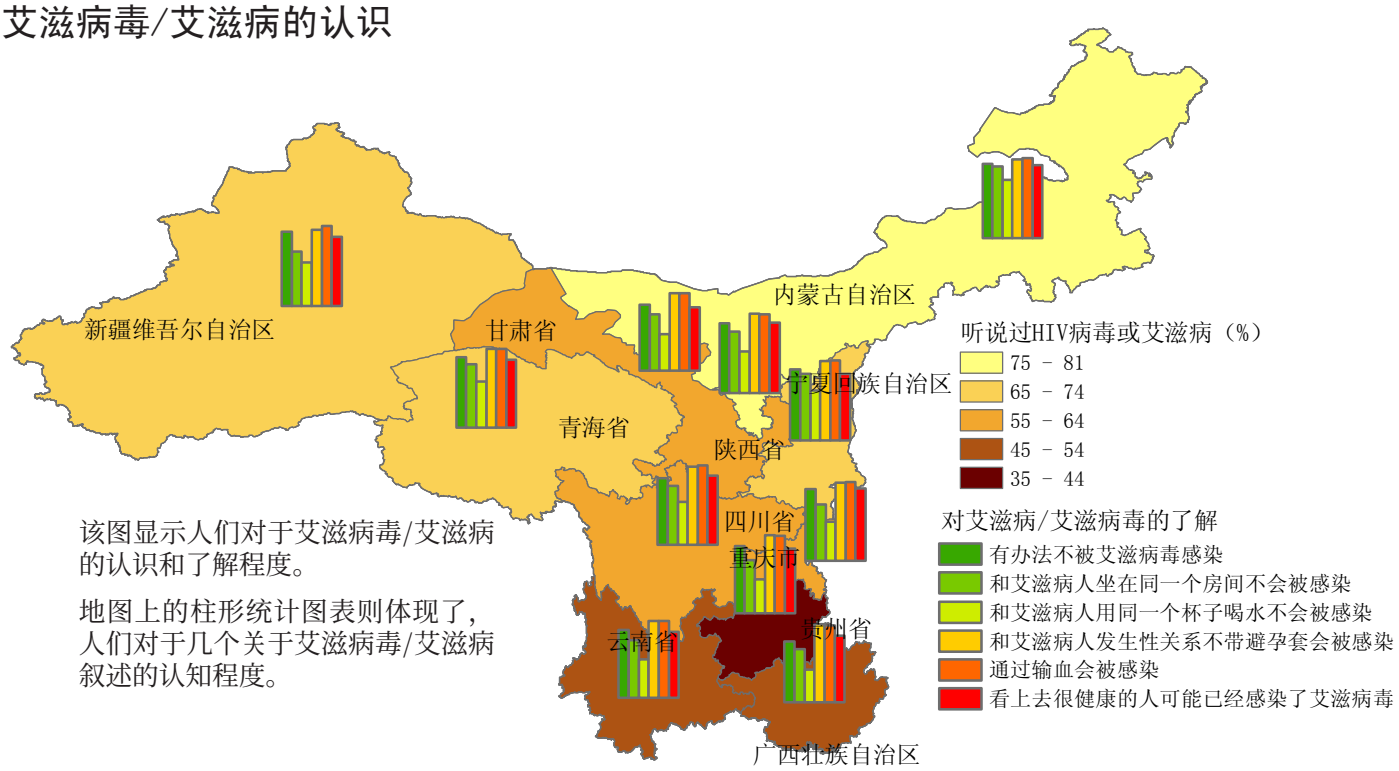
地图上的圆形分格统计图表说明，妇女在分娩以后没有接受产后检查的主要原因。

联合国千年发展目标

目标六：与艾滋病毒/艾滋病、疟疾和其他疾病作斗争

具体目标 7：到2015年制止并开始扭转艾滋病毒/艾滋病的蔓延
具体目标 8：到2015年制止并开始扭转疟疾和其他主要疾病的发病率

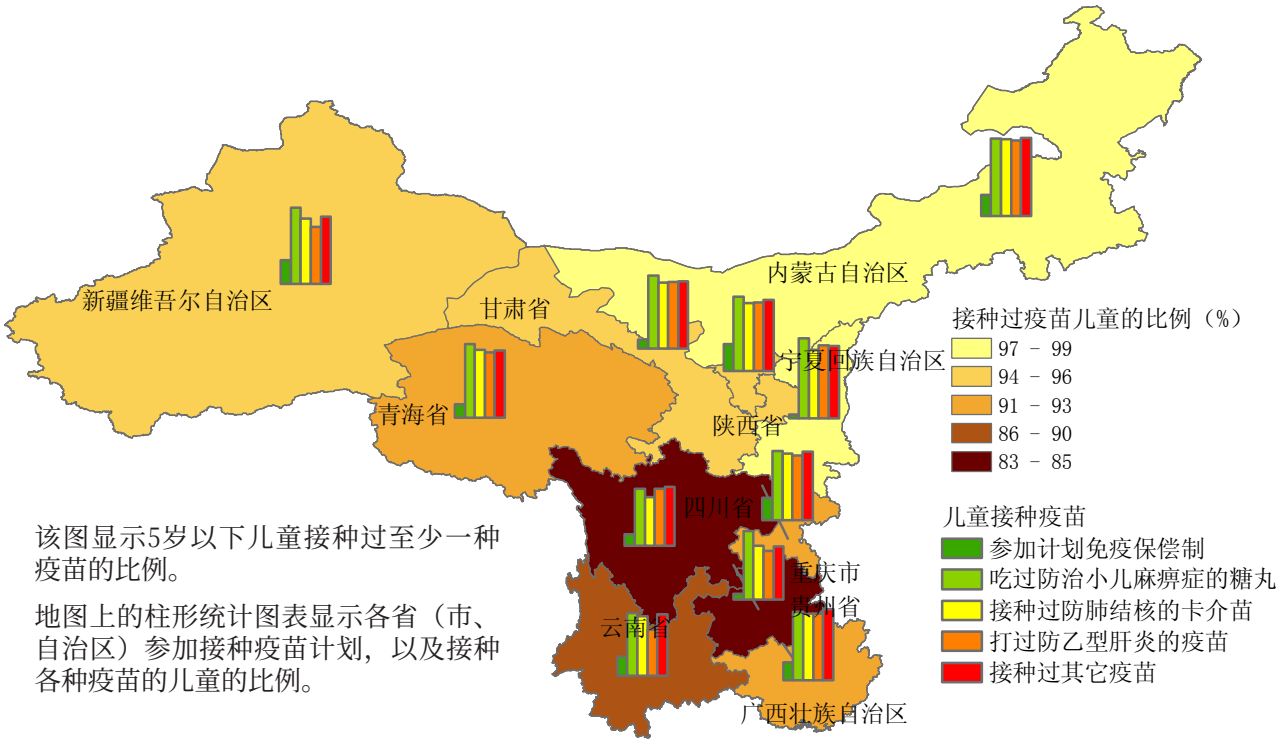
艾滋病毒/艾滋病的认识



该图显示人们对于艾滋病毒/艾滋病的认识和了解程度。

地图上的柱形统计图表则体现了，人们对于几个关于艾滋病毒/艾滋病叙述的认知程度。

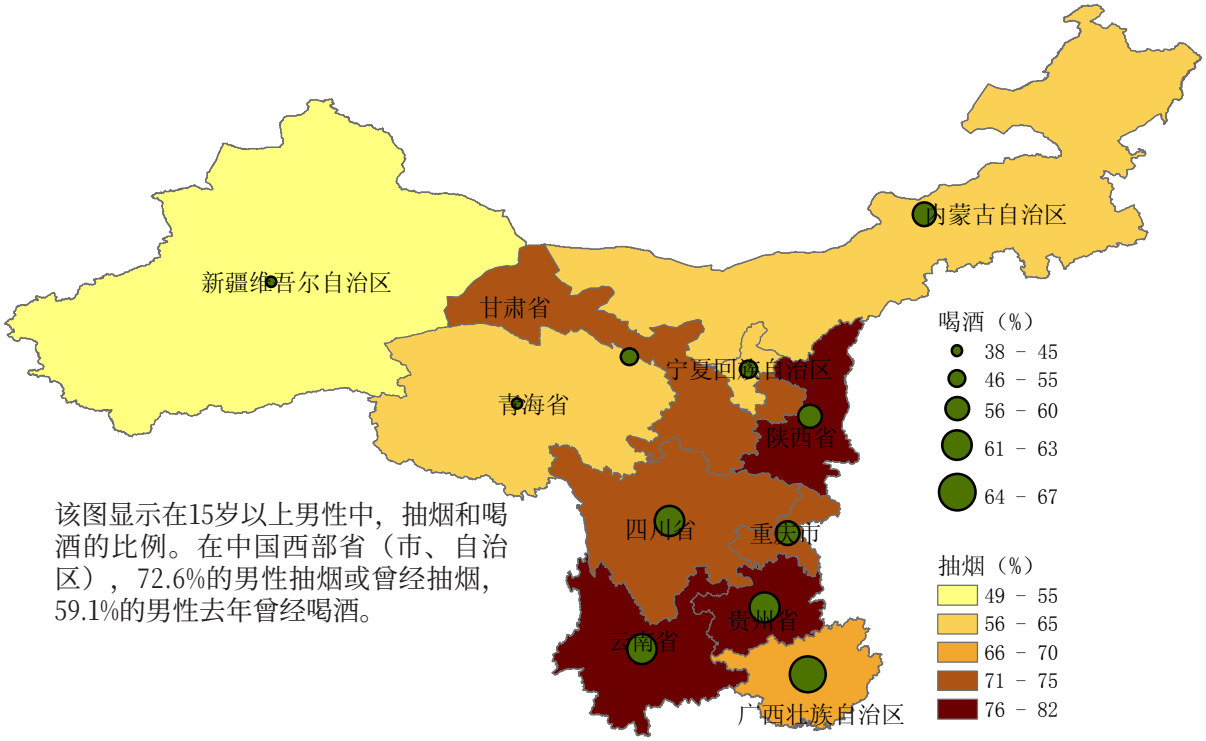
儿童接种疫苗



该图显示5岁以下儿童接种过至少一种疫苗的比例。

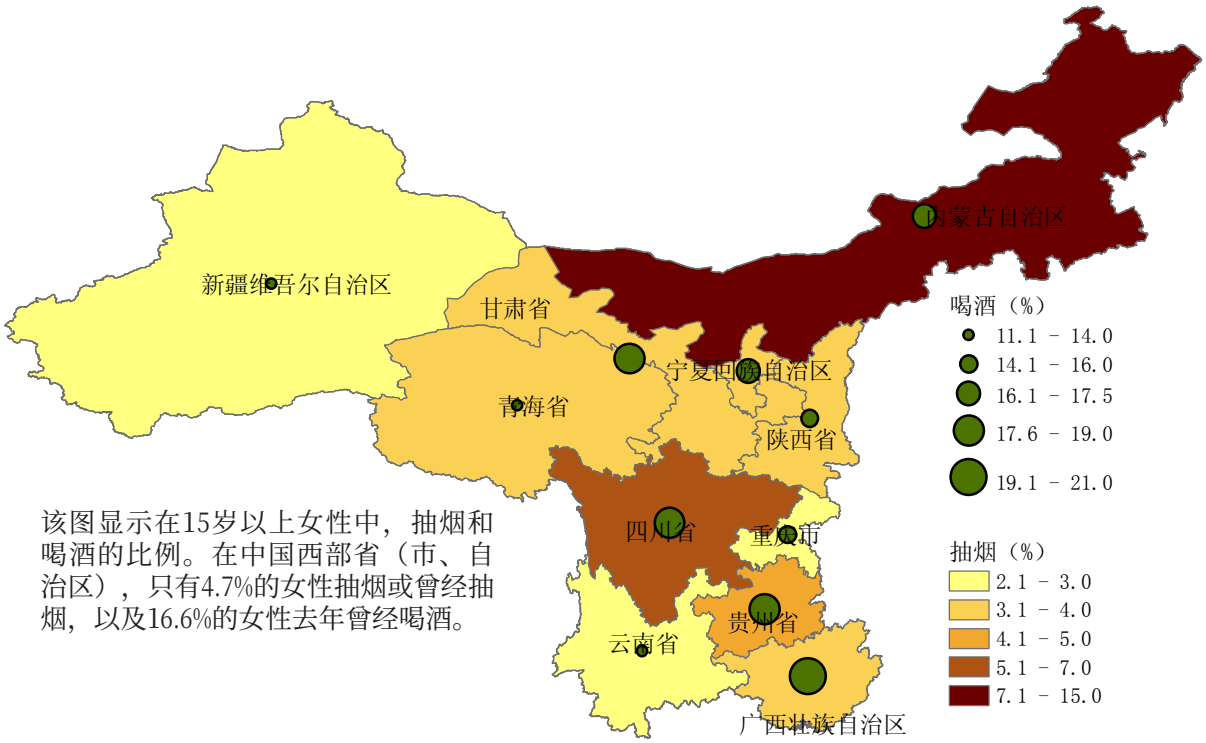
地图上的柱形统计图表显示各省（市、自治区）参加接种疫苗计划，以及接种各种疫苗的儿童的比例。

抽烟和喝酒（男性）



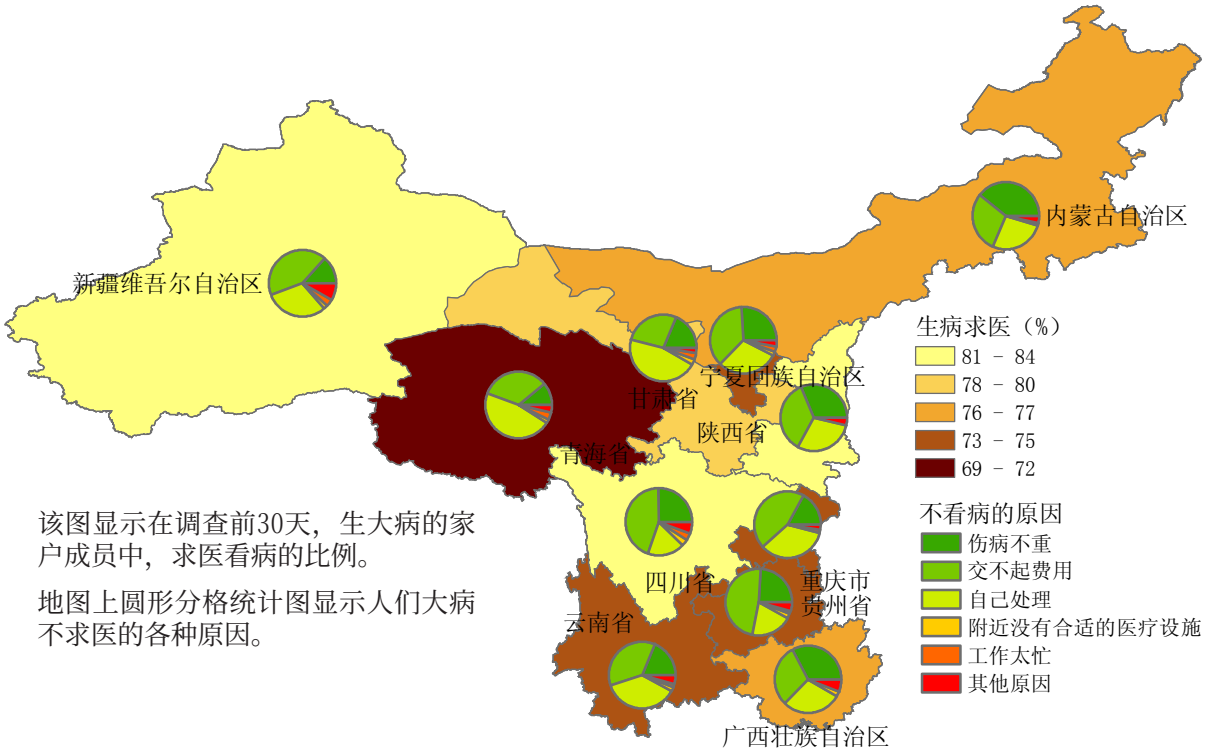
该图显示在15岁以上男性中，抽烟和喝酒的比例。在中国西部省（市、自治区），72.6%的男性抽烟或曾经抽烟，59.1%的男性去年曾经喝酒。

抽烟和喝酒（女性）



该图显示在15岁以上女性中，抽烟和喝酒的比例。在中国西部省（市、自治区），只有4.7%的女性抽烟或曾经抽烟，以及16.6%的女性去年曾经喝酒。

生病求医

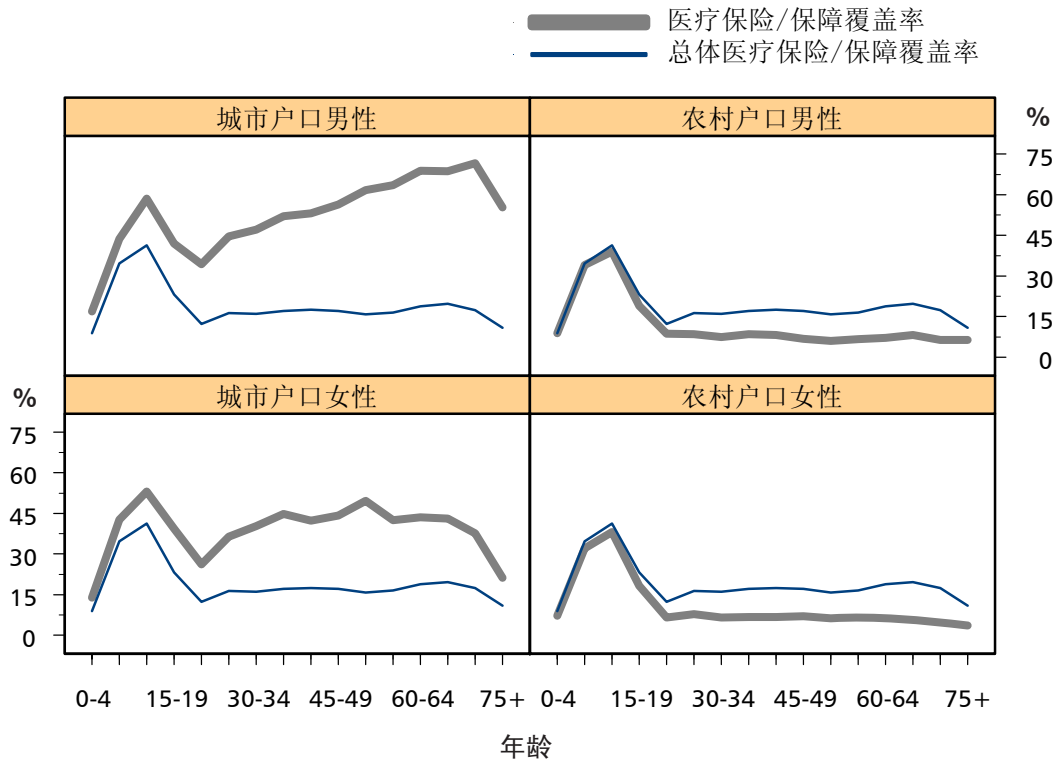


该图显示在调查前30天，生大病的家户成员中，求医看病的比例。

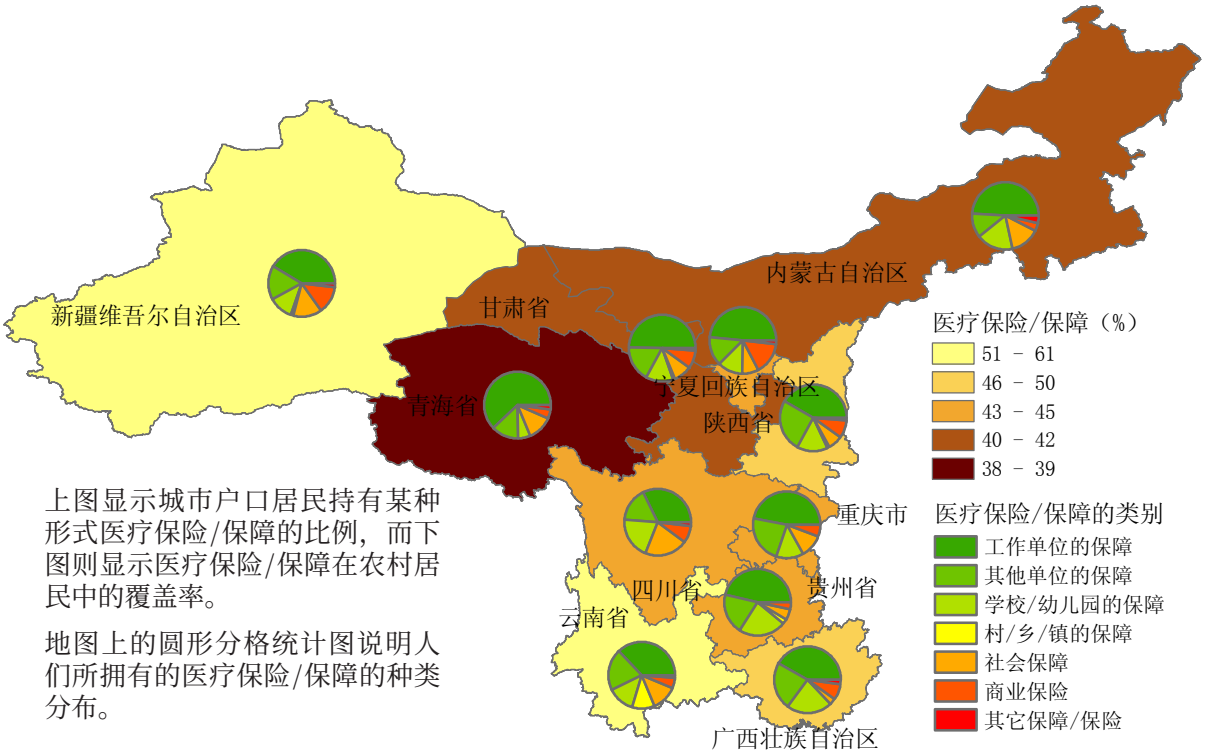
地图上圆形分格统计图显示人们大病不求医的各种原因。

医疗保险/保障

中国西部地区医疗保险/保障的覆盖率非常的低，平均只有20%的居民拥有某种形式的医疗保险/保障。农村户口居民拥有医疗保险/保障的比例特别低（12.7%），而将近一半的城市居民有医疗保险/保障（46.9%）。从图表可知，城市户口的男性拥有某种形式医疗保险/保障的比例最高。



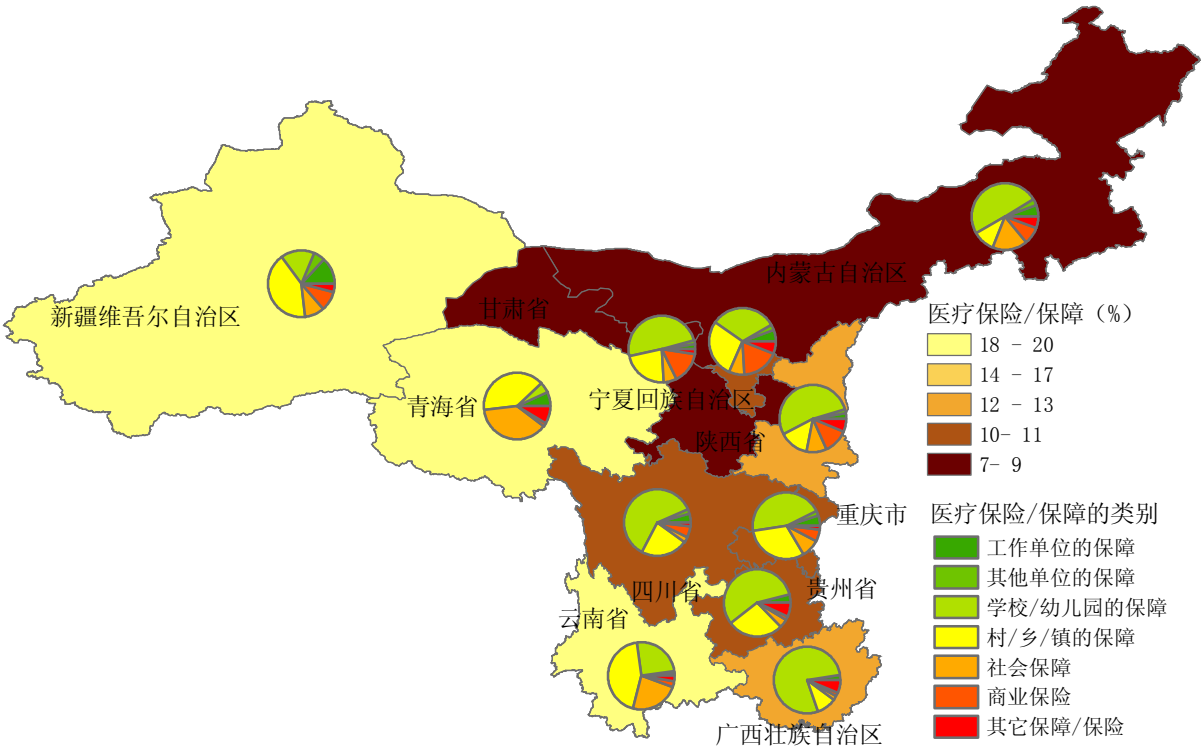
医疗保险/保障（城市户口）



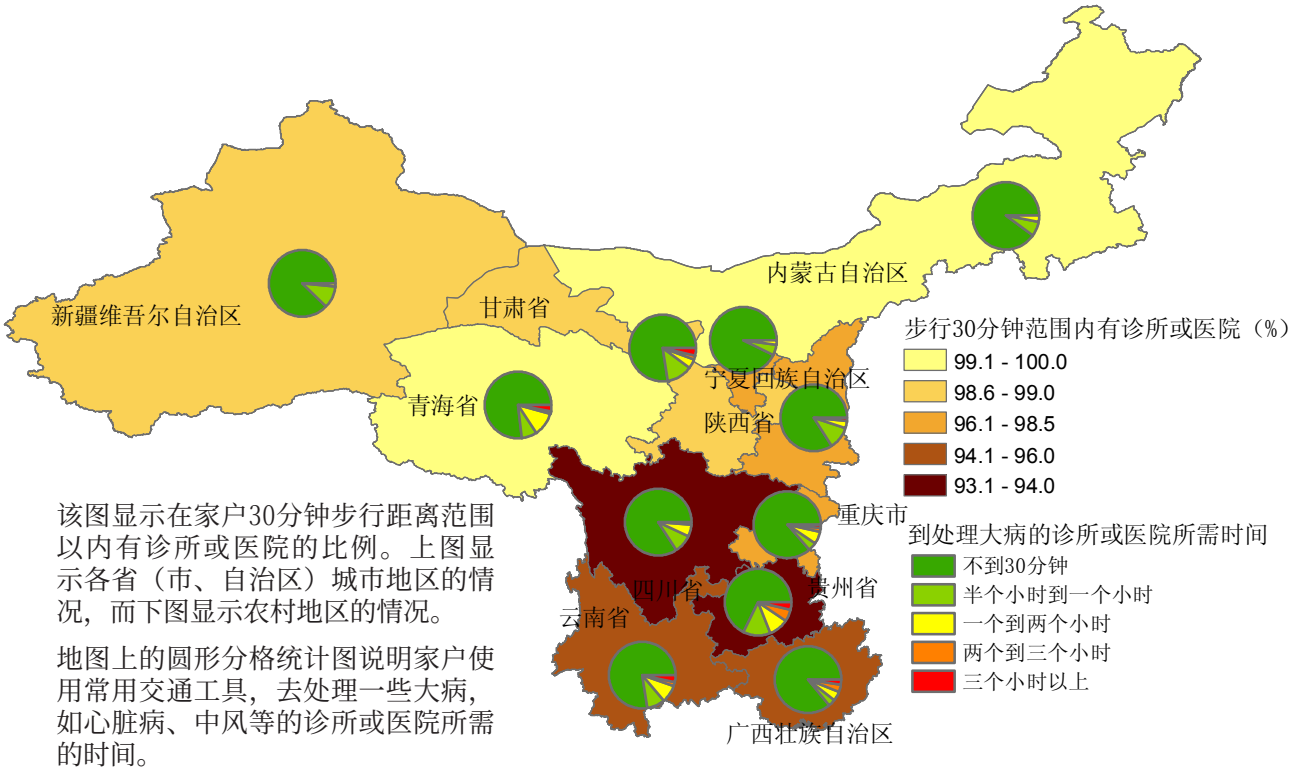
上图显示城市户口居民持有某种形式医疗保险/保障的比例，而下图则显示医疗保险/保障在农村居民中的覆盖率。

地图上的圆形分格统计图说明人们所拥有的医疗保险/保障的种类分布。

医疗保险/保障（农村户口）



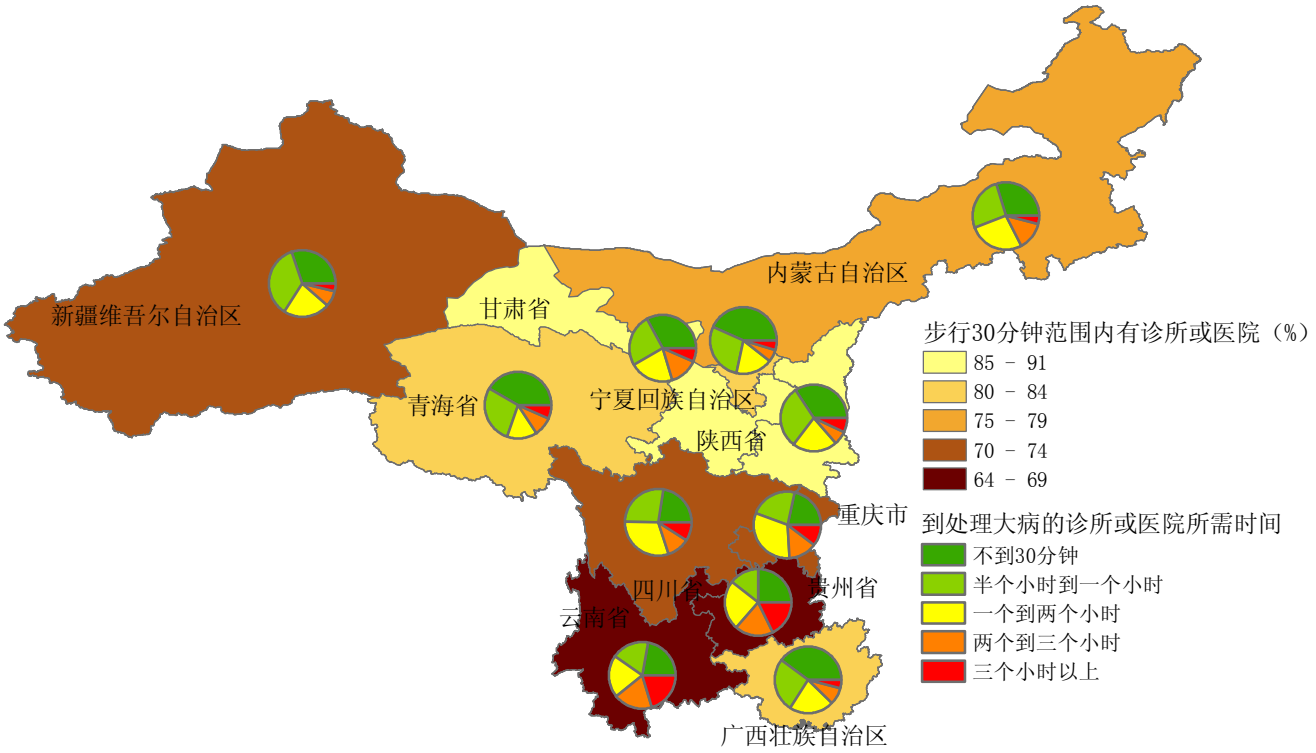
公共医疗卫生服务（城市）



该图显示在家户30分钟步行距离范围以内有诊所或医院的比例。上图显示各省（市、自治区）城市地区的情况，而下图显示农村地区的情况。

地图上的圆形分格统计图说明家户使用常用交通工具，去处理一些大病，如心脏病、中风等的诊所或医院所需的时间。

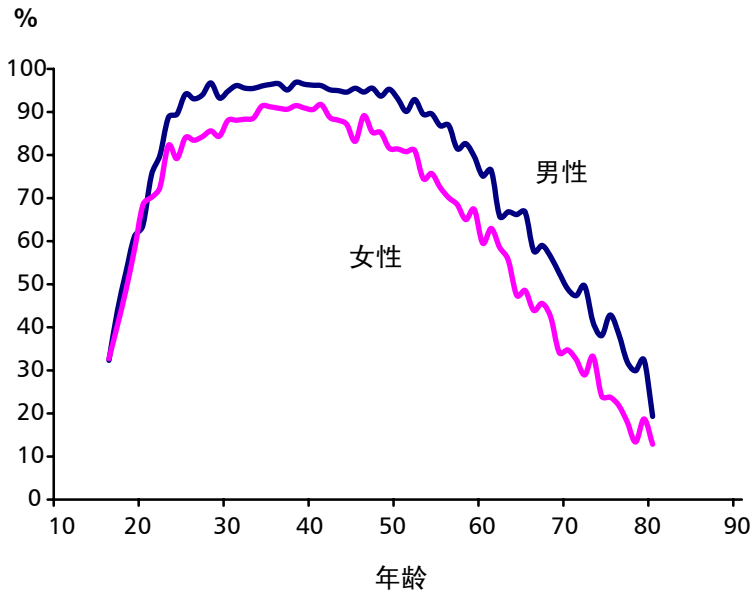
公共医疗卫生服务（农村）



劳动人口和就业

就业状态			
西部地区总人口　3.69亿			
劳动年龄人口　2.8亿			不到劳动年龄人口　8900万
劳动力　2.15亿		不在劳动力市场　6500万	
就业人口　2.11亿	失业人口　400万		

劳动参与率

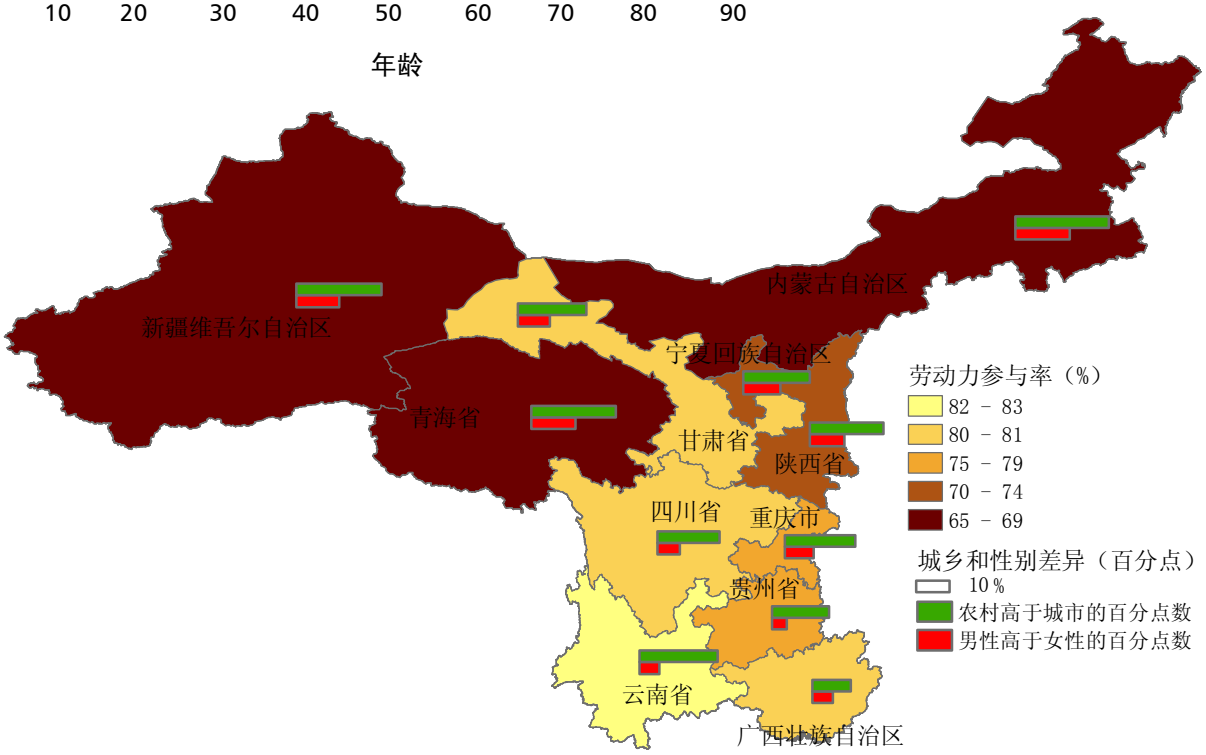


根据国际劳工组织（ILO）的定义，失业人员就是那些在受调查的前一个星期内，没有工作过至少一个小时，并且在积极寻找工作而且如果获得机会随时可以开始投入工作的人。

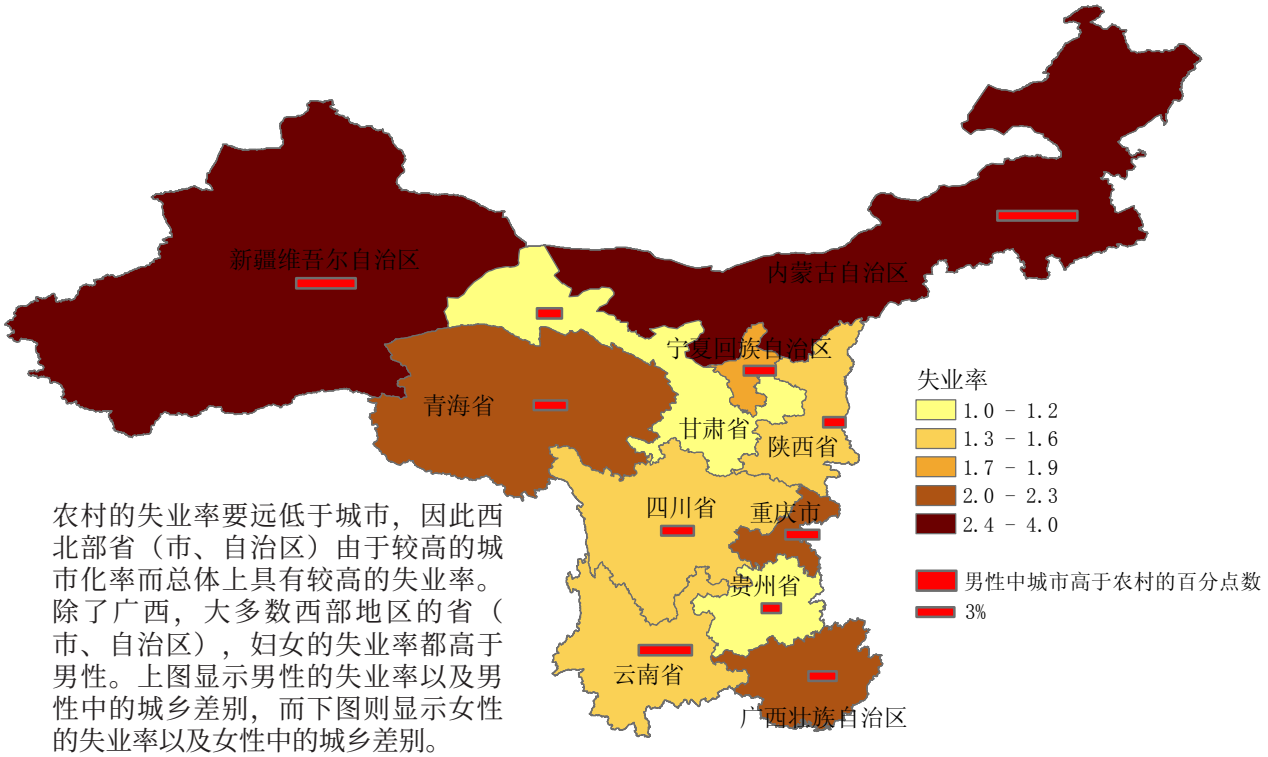
不到劳动年龄人口在这里被定义为不满16周岁的人口。

左图显示西部地区不同年龄和性别组群的劳动参与率。妇女劳动年龄人口中的不参与劳动力市场的比例略高于男性。

下图说明西北地区省（市、自治区）的劳动参与率普遍低于西南地区，而农村地区的劳动参与率要远高于城市地区。最后，各地区的男性总体的劳动参与率都略高于女性。

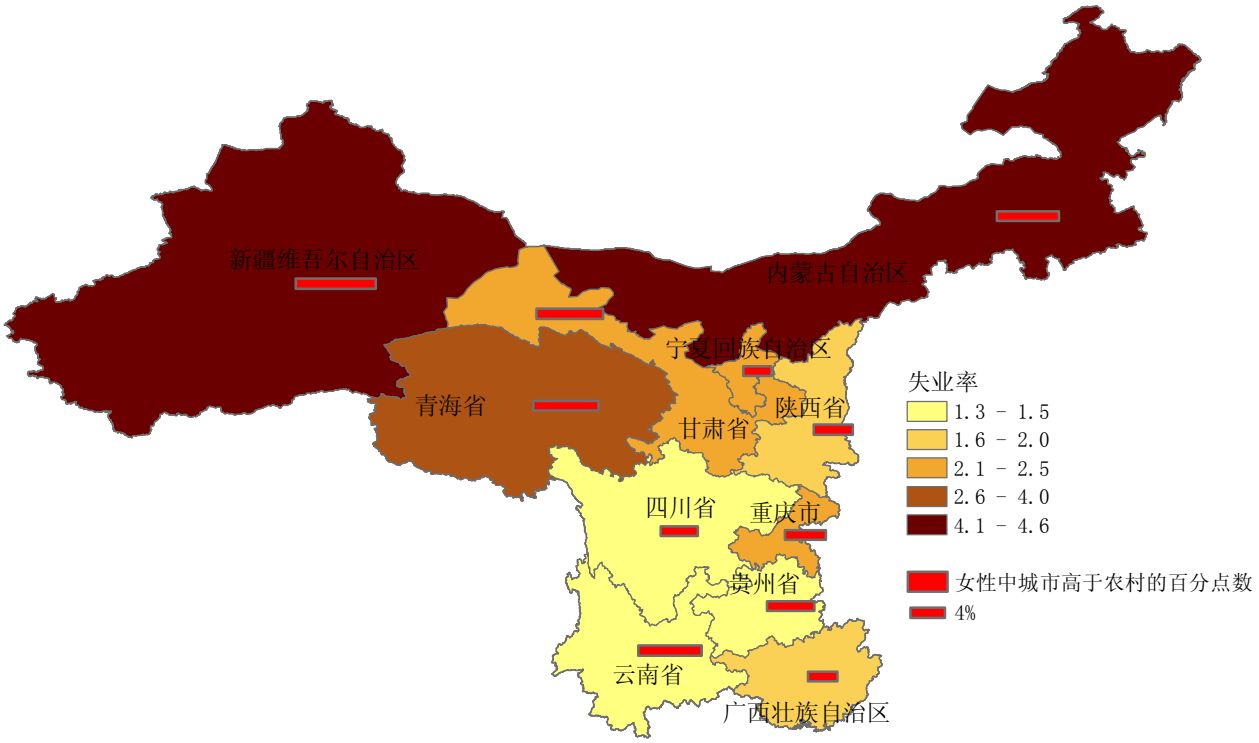


失业率（男性）

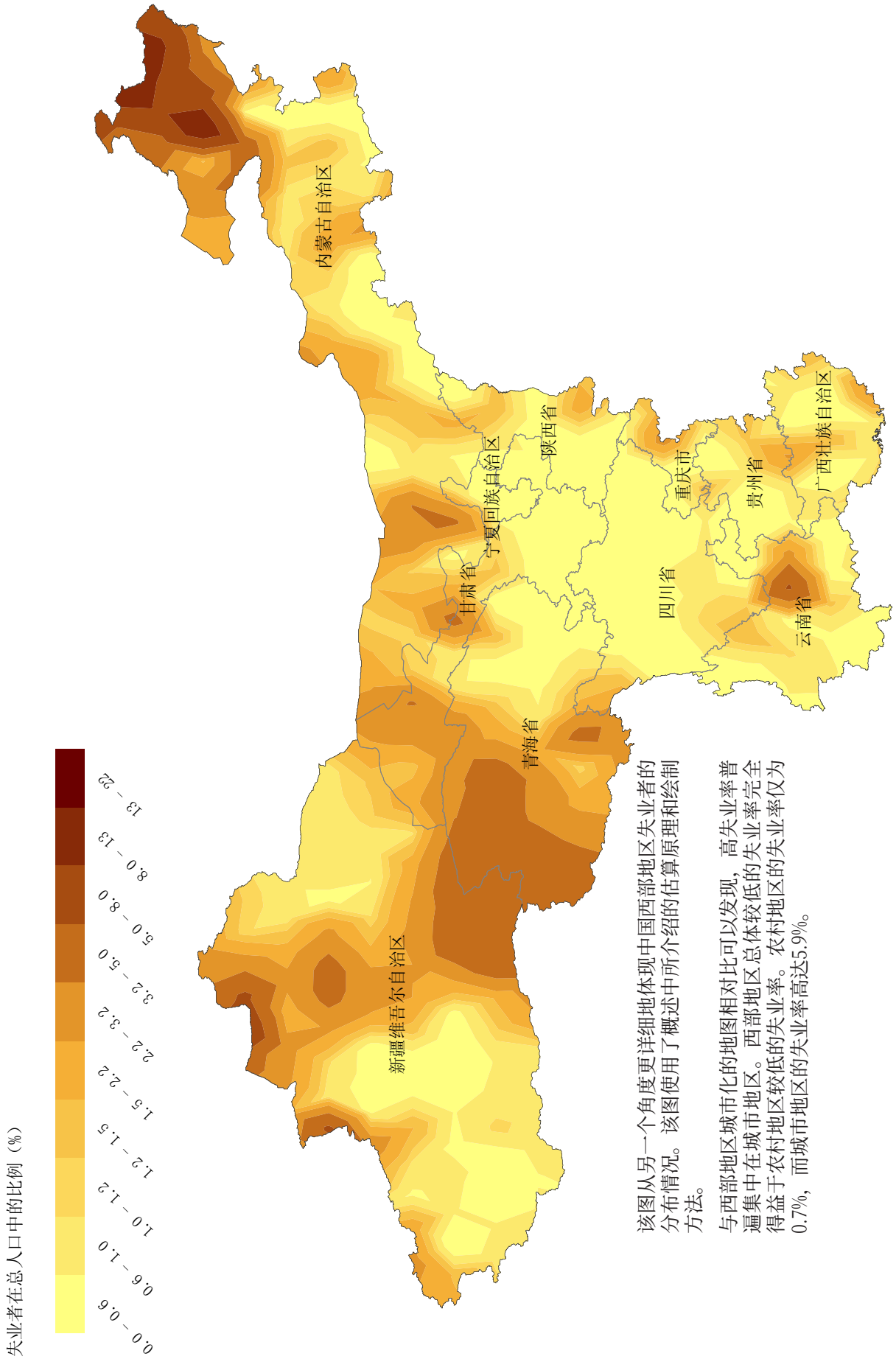


农村的失业率要远低于城市，因此西北部省（市、自治区）由于较高的城市化率而总体上具有较高的失业率。除了广西，大多数西部地区的省（市、自治区），妇女的失业率都高于男性。上图显示男性的失业率以及男性中的城乡差别，而下图则显示女性的失业率以及女性中的城乡差别。

失业率（女性）



失业率：详细分布图



该图从另一个角度更详细地体现中国西部地区失业者的分布情况。该图使用了概述中所介绍的估算原理和绘制方法。

与西部地区城市化的地图相对比可以发现，高失业率普遍集中在城市地区。西部地区总体较低的失业率完全得益于农村地区较低的失业率。农村地区的失业率仅为0.7%，而城市地区的失业率高达5.9%。

年轻人的失业率

联合国千年发展目标

目标八：全球合作促进发展

具体目标16：与发展中国家合作，拟订和实施为青年创造体面的生产性就业机会的战略

指标45：15至24岁青年的失业率

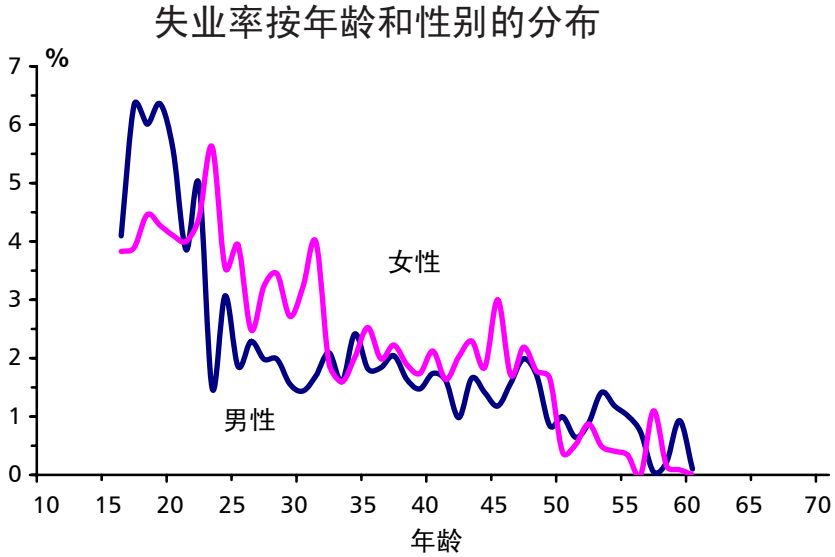
中国西部地区15至24岁青年的失业率

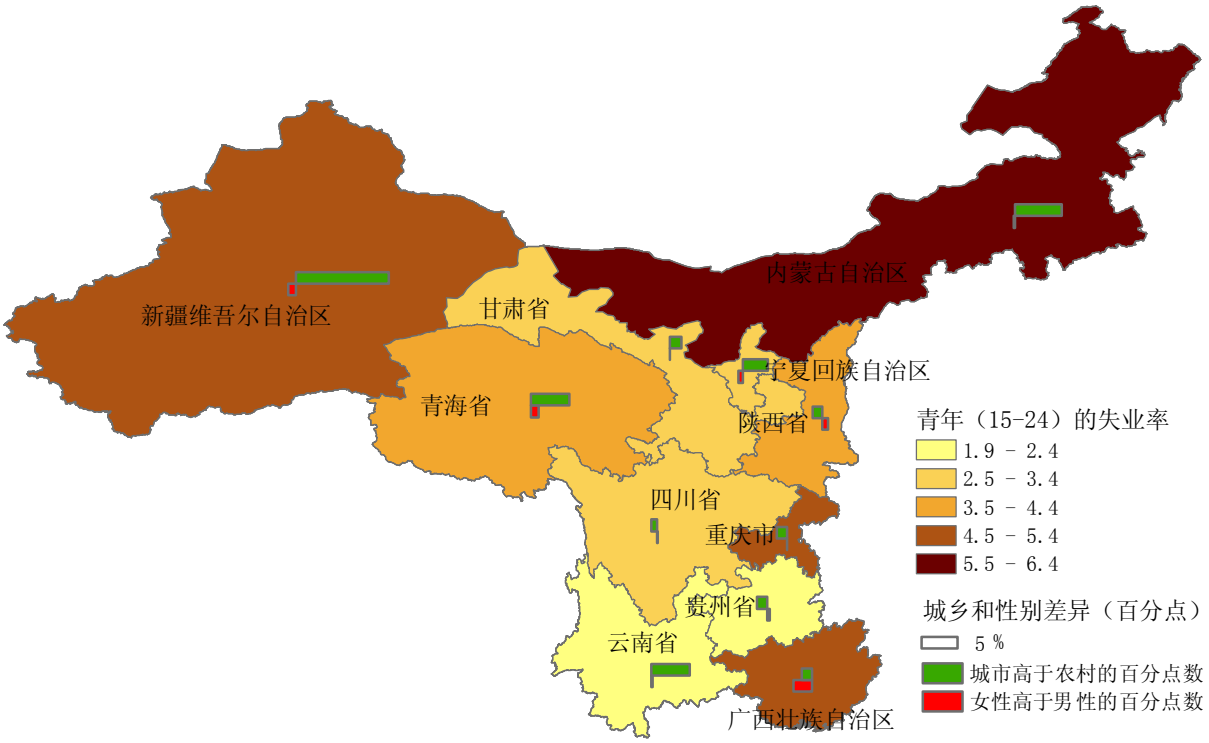
男性：　4.4%
女性：　4.2%
总体：　4.3%

在中国西部地区3.69亿人口中，4700万是15至24岁的年轻人。

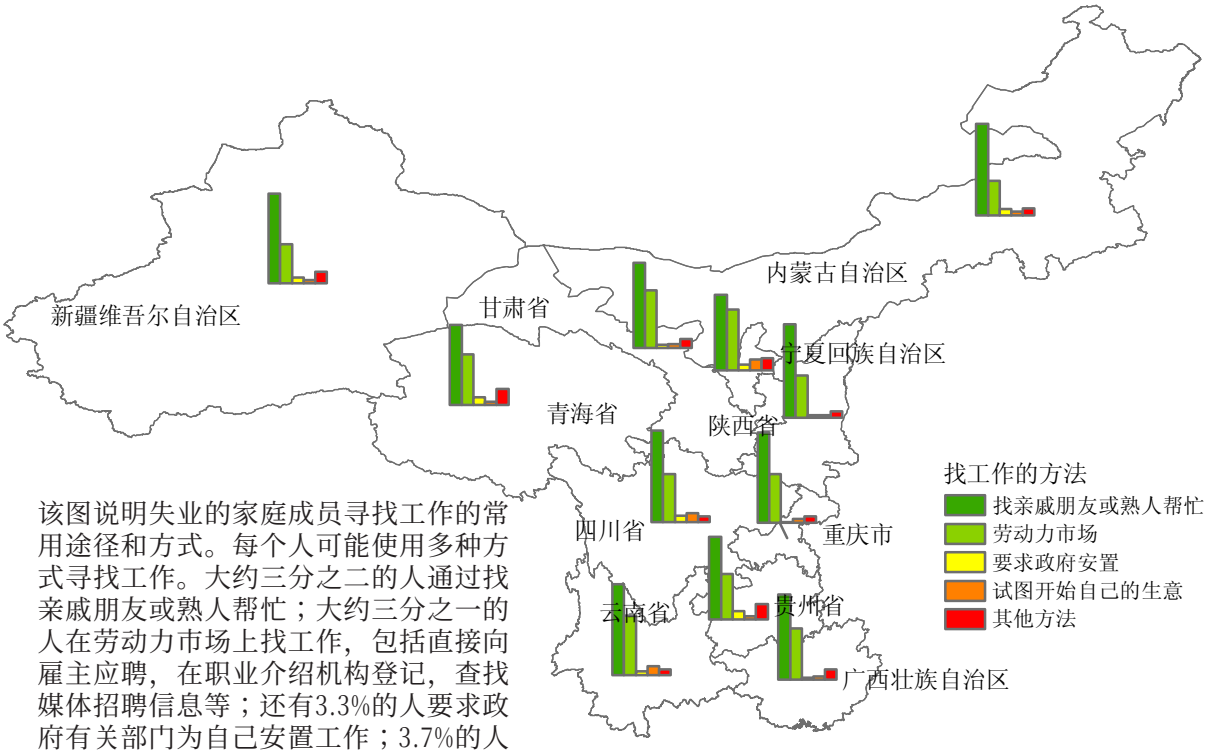
在这些年轻人中，有2800万（60%）参与劳动力市场的就业，而其中有124万正处于失业状态。年轻人中的失业率（4%）远高于西部地区总体的失业率（1.9%），也就是说大约有三分之一的失业者是15到24岁的年轻人。

右图展示了目前15至24岁年轻人的失业率，男性和女性之间的性别差异很小，在某些省（市、自治区），如广西，宁夏等，男性的失业率要明显高于女性。



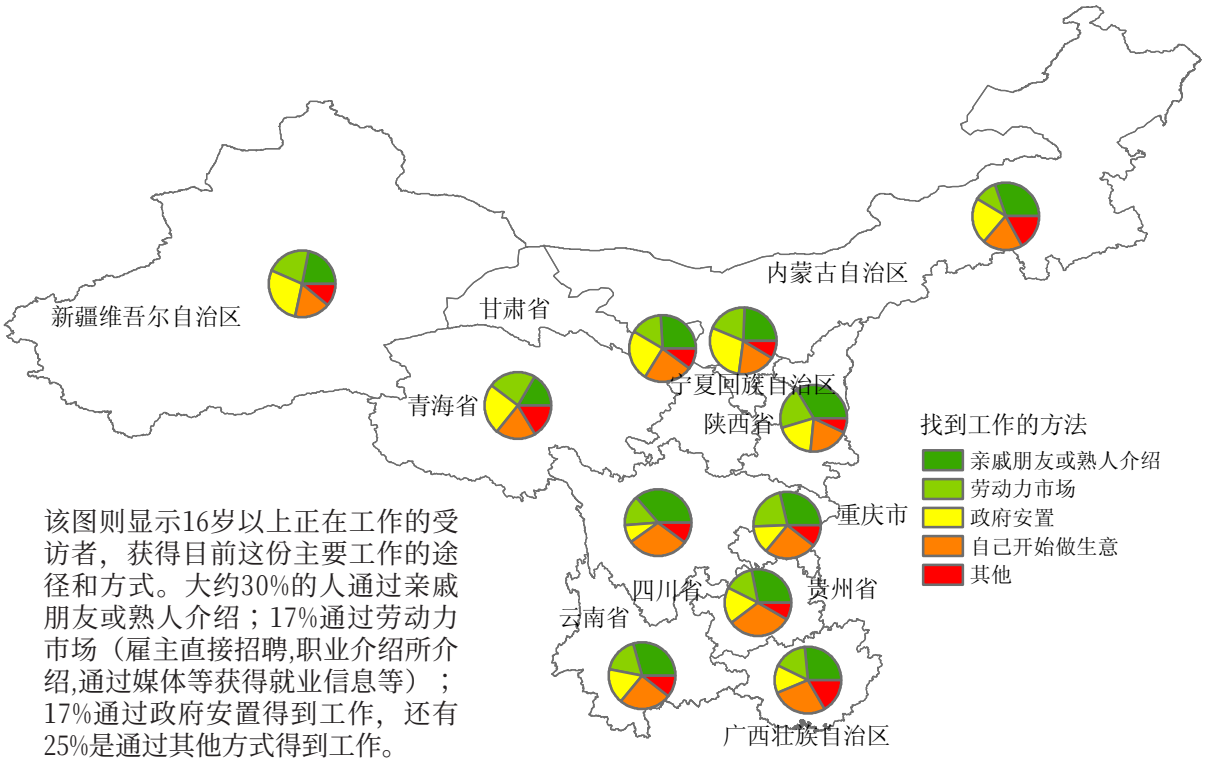


失业者求职途径



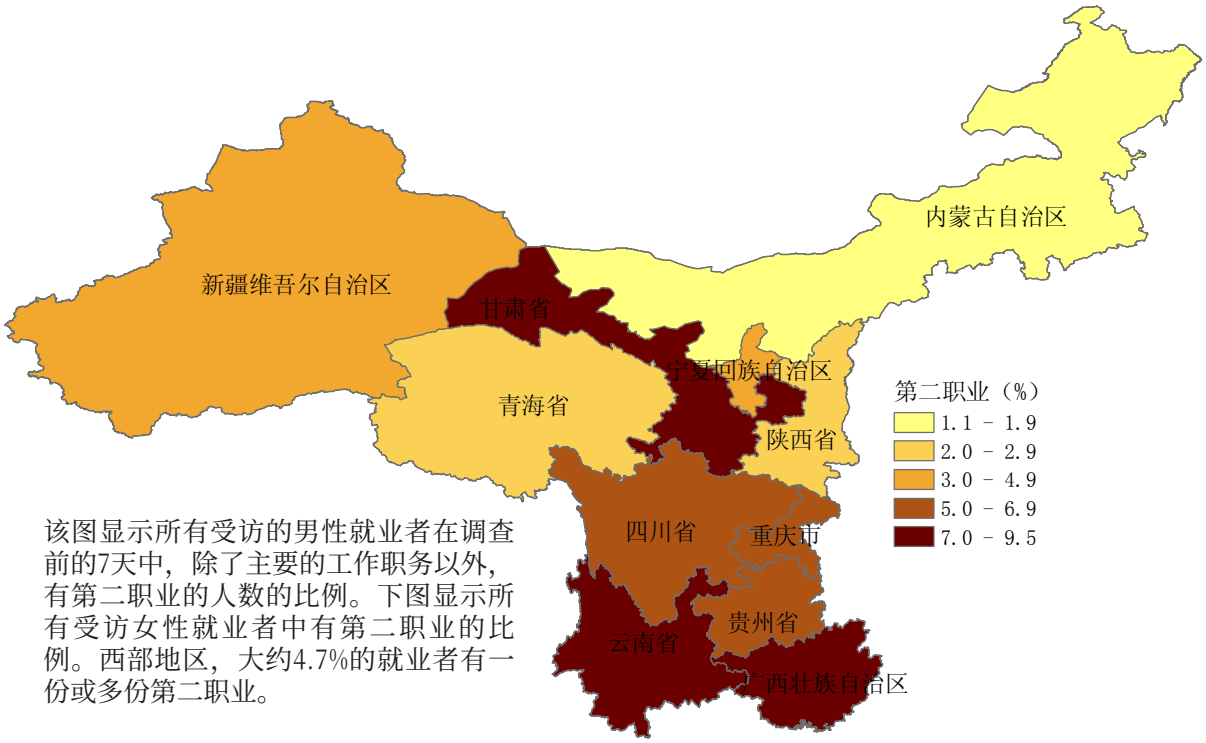
该图说明失业的家庭成员寻找工作的常用途径和方式。每个人可能使用多种方式寻找工作。大约三分之二的人通过找亲戚朋友或熟人帮忙；大约三分之一的人在劳动力市场上找工作，包括直接向雇主应聘，在职业介绍机构登记，查找媒体招聘信息等；还有3.3%的人要求政府有关部门为自己安置工作；3.7%的人试图开始自己的生意。

获得工作的方式



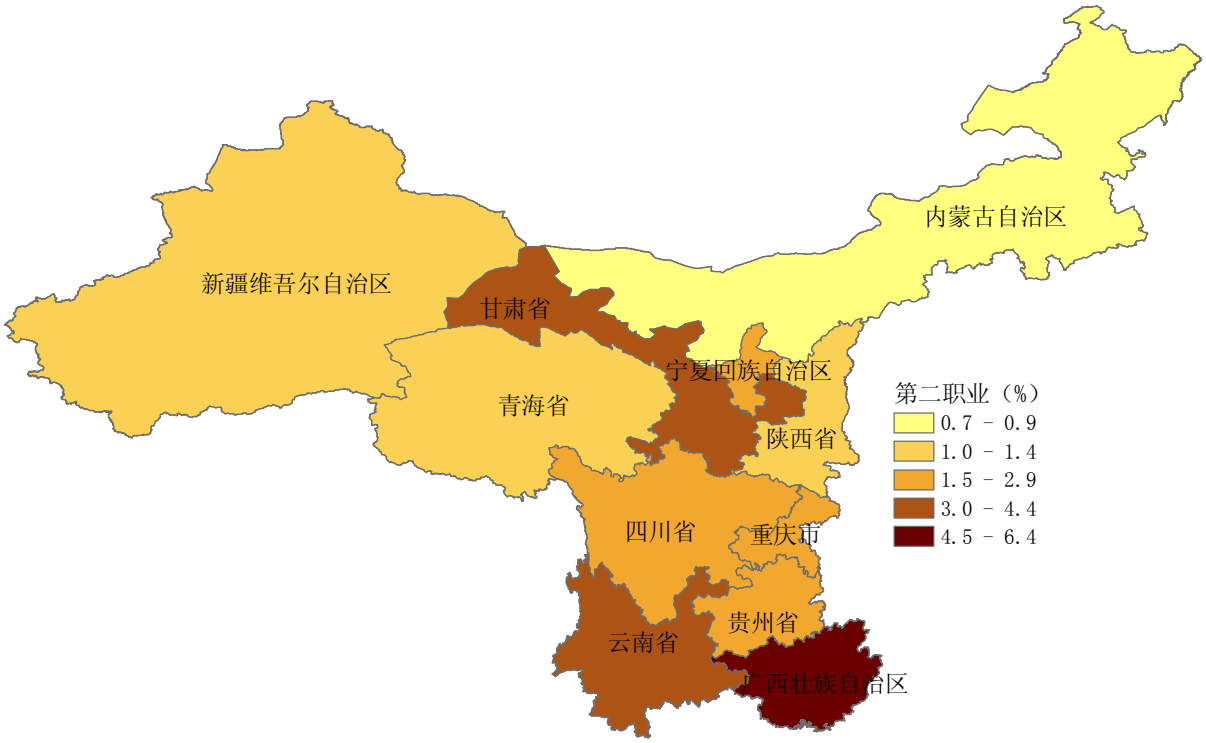
该图则显示16岁以上正在工作的受访者，获得目前这份主要工作的途径和方式。大约30%的人通过亲戚朋友或熟人介绍；17%通过劳动力市场（雇主直接招聘,职业介绍所介绍,通过媒体等获得就业信息等）；17%通过政府安置得到工作，还有25%是通过其他方式得到工作。

第二职业（男性）

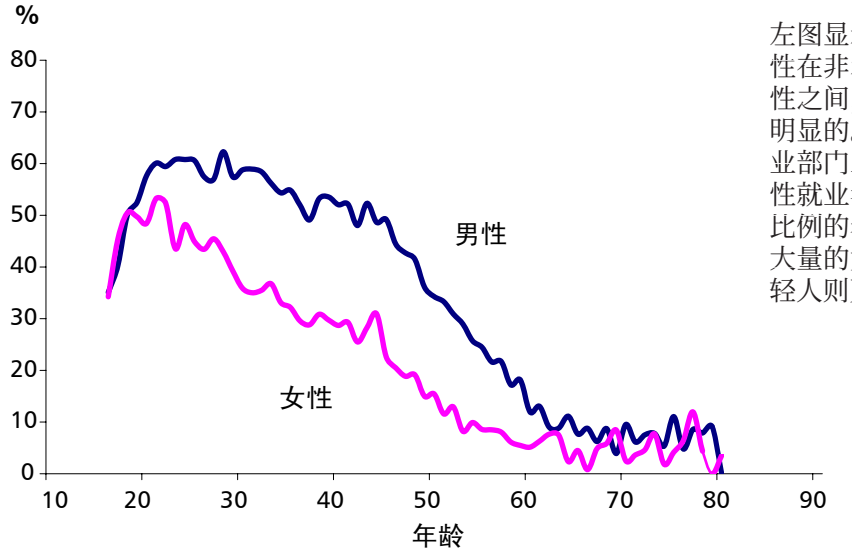


该图显示所有受访的男性就业者在调查前的7天中，除了主要的工作职务以外，有第二职业的人数的比例。下图显示所有受访女性就业者中有第二职业的比例。西部地区，大约4.7%的就业者有一份或多份第二职业。

第二职业（女性）

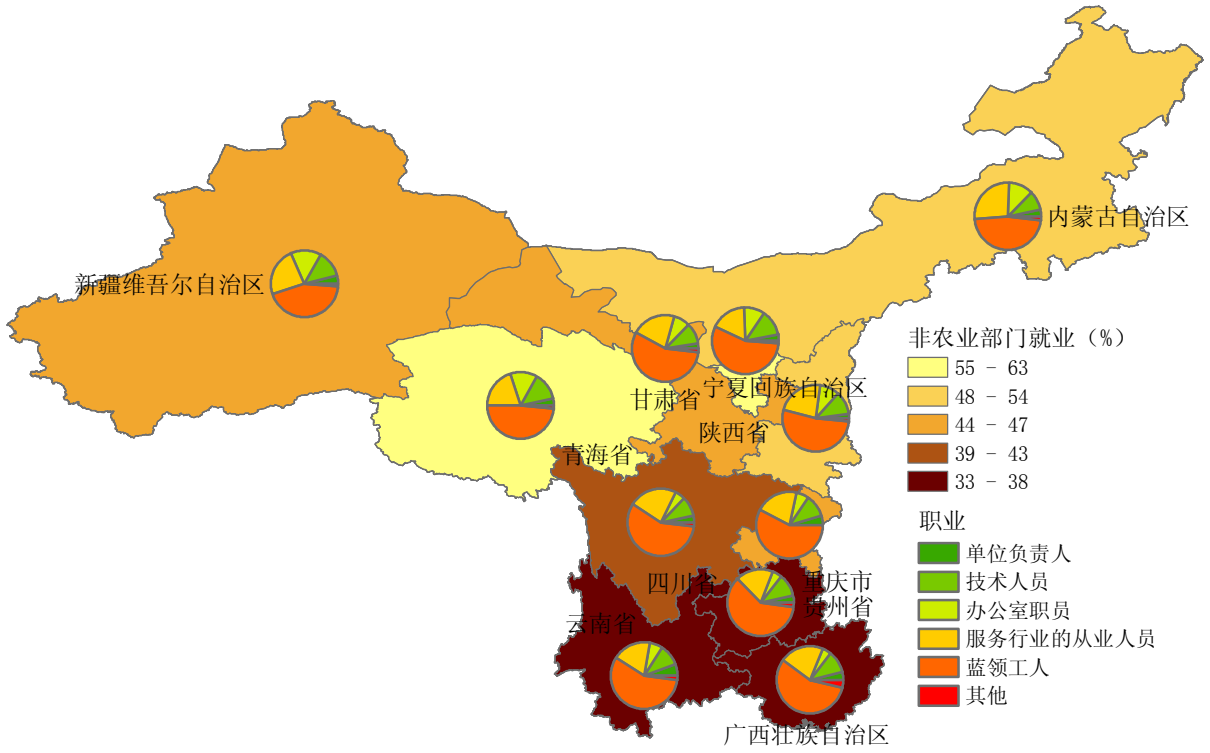


非农业部门就业

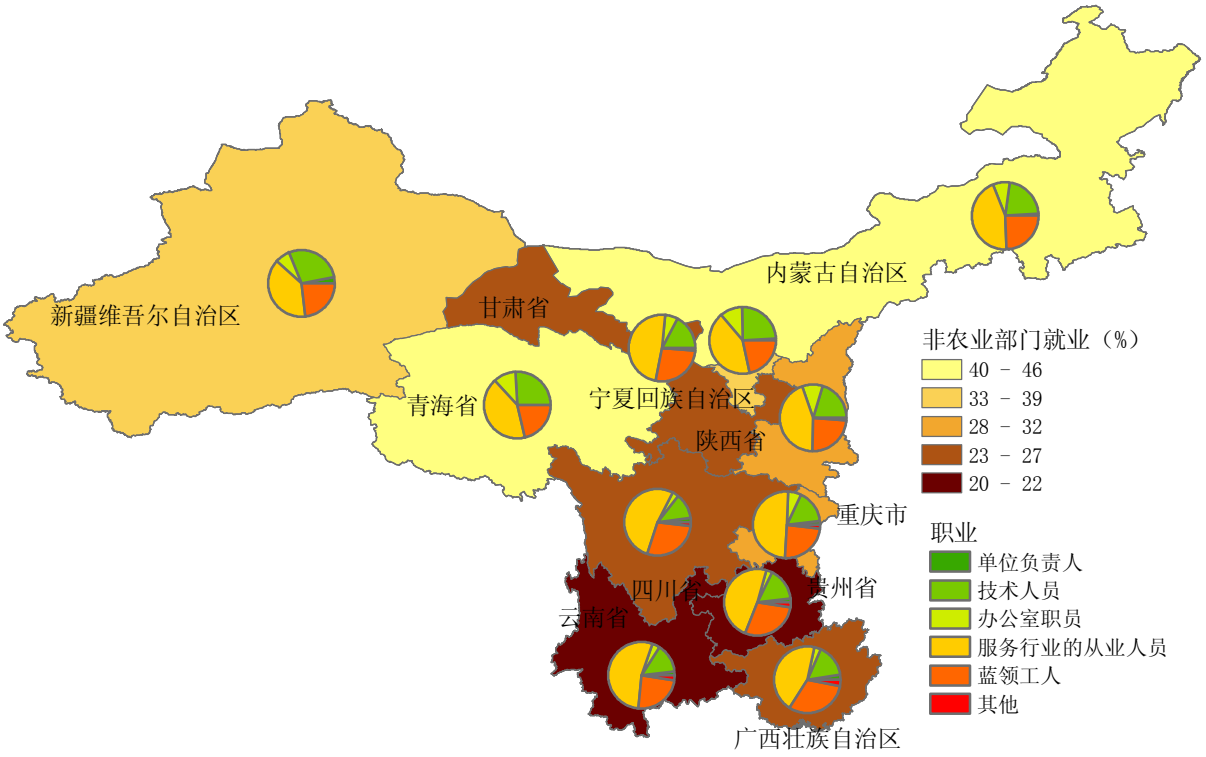


左图显示在所有的就业者中，男性和女性在非农业部门就业的比例。男性和女性之间，不同年龄层之间的差距是非常明显的。大约一半的男性就业者在非农业部门工作，而只有大约三分之一的女性就业者在非农业部门就业，而且极少比例的老年就业者在非农业部门工作。大量的女性和老年人在农村耕地，而年轻人则更多在非农业工薪部门工作。

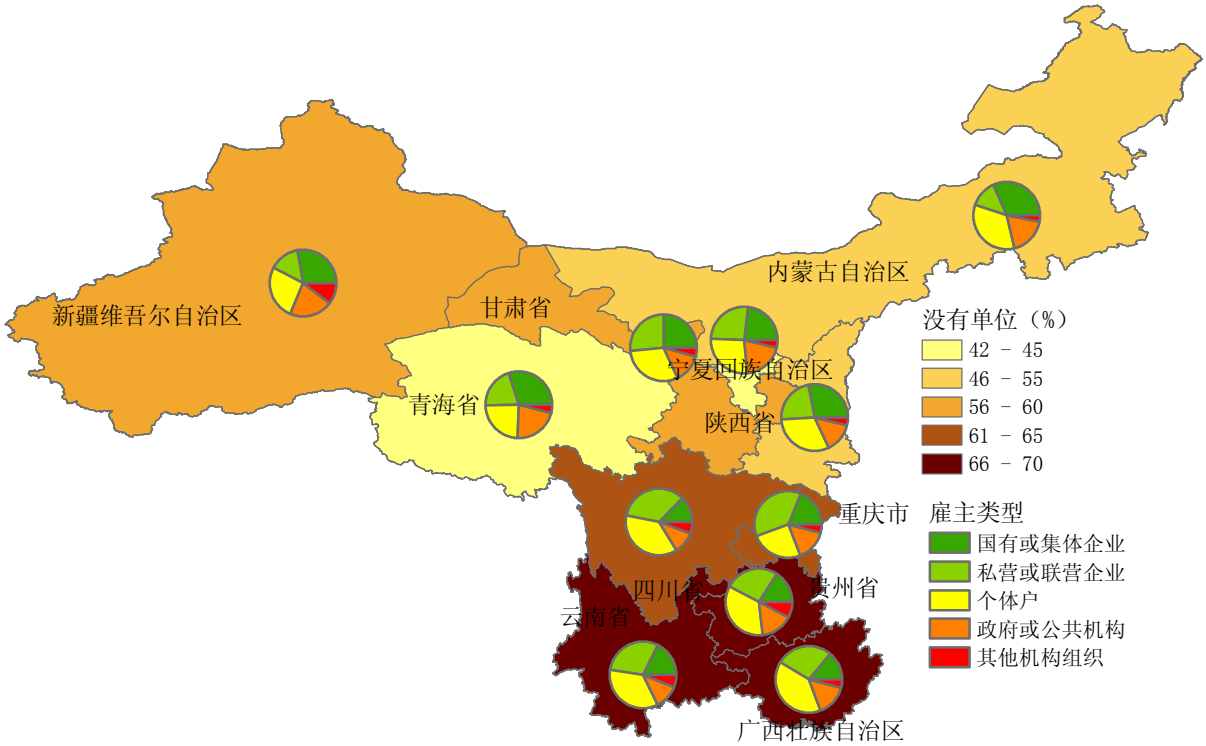
职业和非农业部门就业（男性）



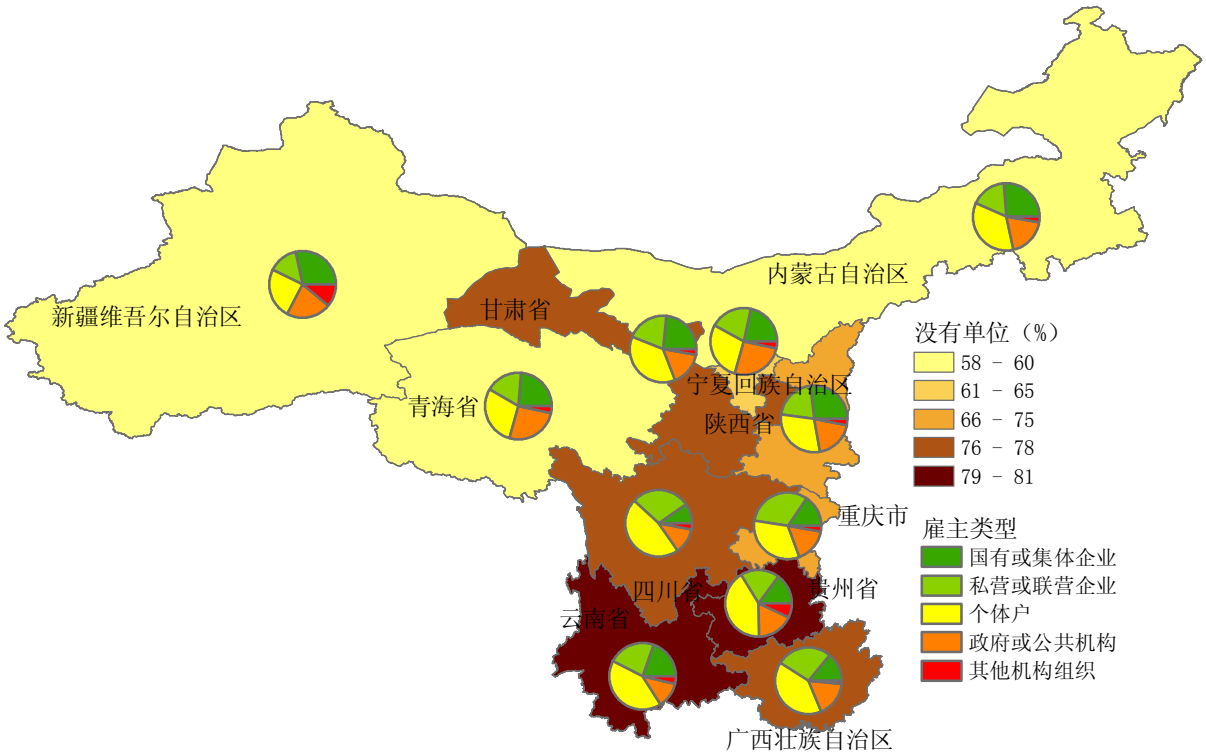
职业和非农业部门就业（女性）



雇主类型（男性）



雇主类型（女性）



行业和妇女的份额

行业	就业（%）			妇女在行业中所占的份额（%）
	男性	女性	总共	
农业	57.9	73.3	65.1	53
采矿业	2.5	0.5	1.6	15
制造业	8.0	5.5	6.8	37
能源和水资源	0.9	0.4	0.6	27
建筑	8.5	1.3	5.1	12
运输	4.6	0.9	2.9	14
信息技术	0.3	0.2	0.3	35
批发和零售	6.0	7.5	6.7	52
旅馆业	2.0	2.9	2.4	57
金融	0.3	0.4	0.3	53
房地产	0.2	0.2	0.2	42
租赁和商业服务	0.8	0.4	0.6	32
研究	0.2	0.1	0.2	29
环境保护	0.2	0.3	0.2	53
服务业	1.4	1.5	1.5	47
教育	2.1	2.3	2.2	48
健康和社会保险	0.8	0.9	0.9	49
文化，体育和娱乐	0.4	0.3	0.4	38
公共管理和社会组织	2.8	1.3	2.1	28

联合国千年发展目标

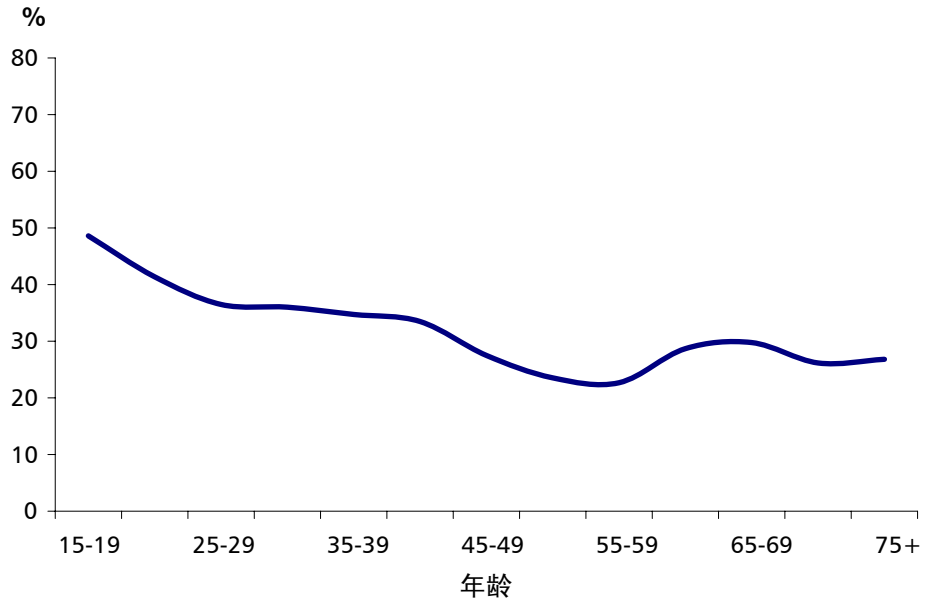
目标3：促进男女平等并赋予妇女权利

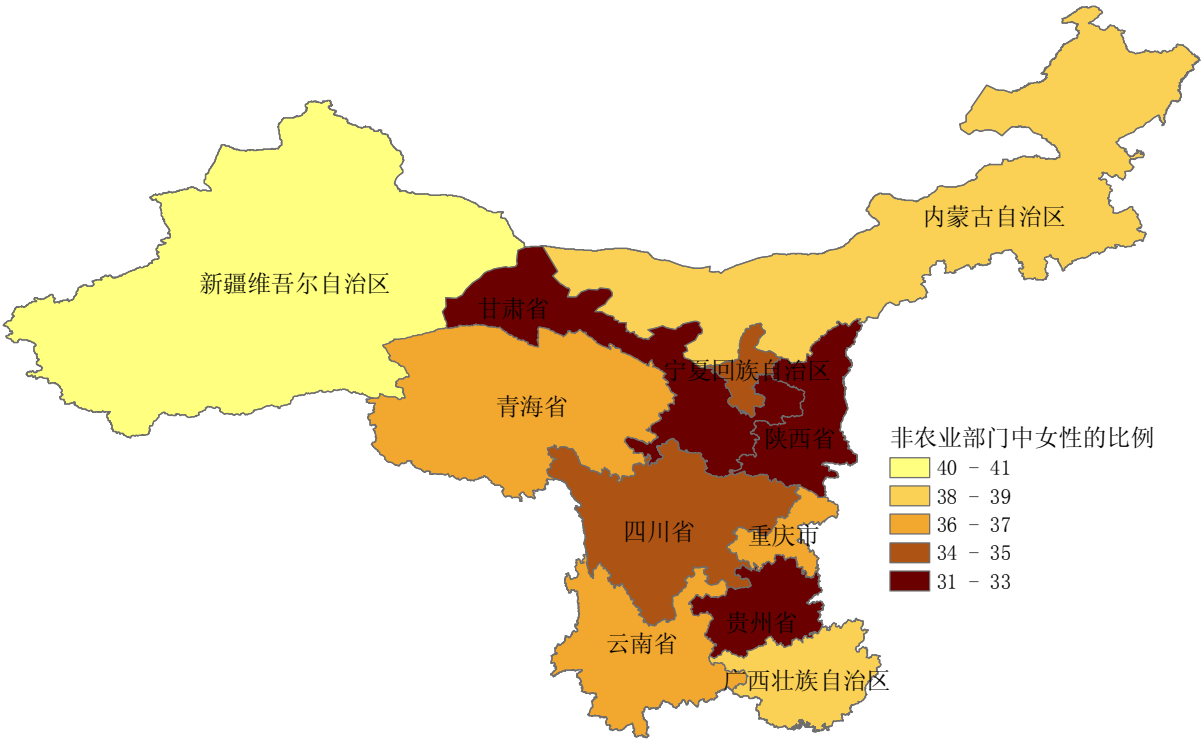
具体目标4：到2005年基本消除初等教育和中等教育中的男女差距，并且最迟于2015年消除各级教育中的男女差距

指标11：妇女在非农业工薪部门中所占的份额

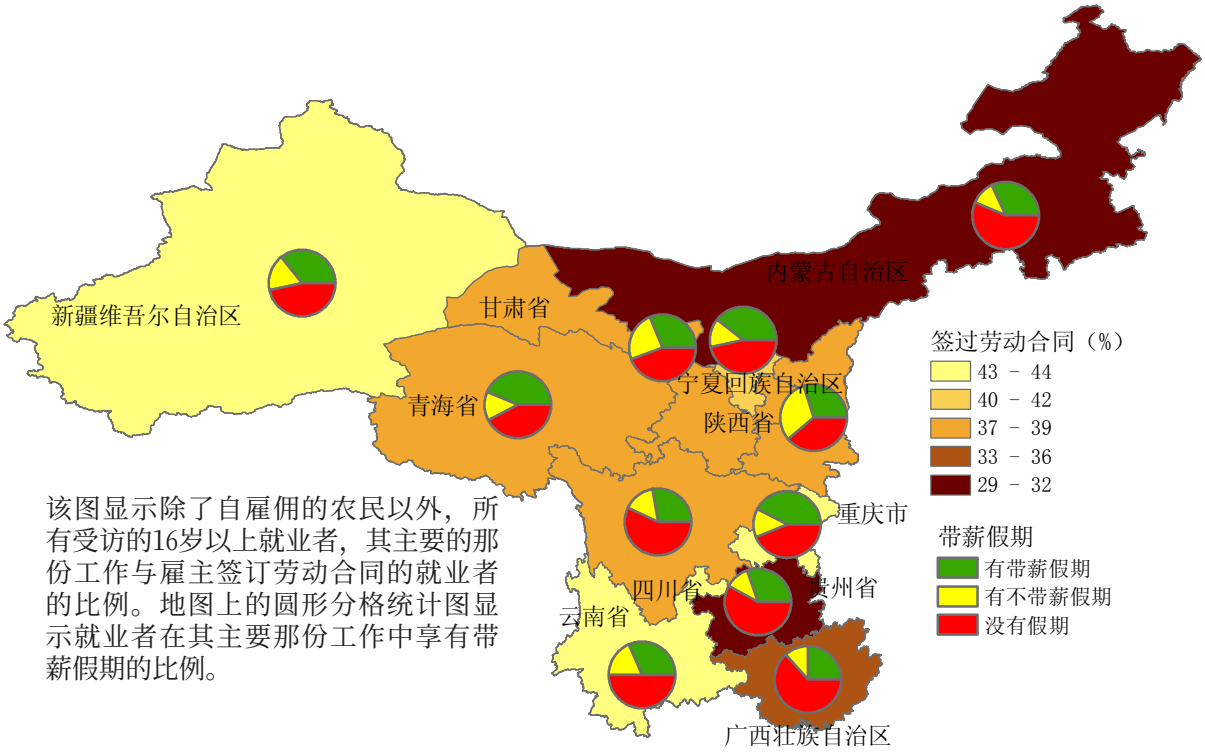
中国西部地区妇女在非农业部门工薪阶层中所占的份额：35.5%

妇女在非农业部门就业



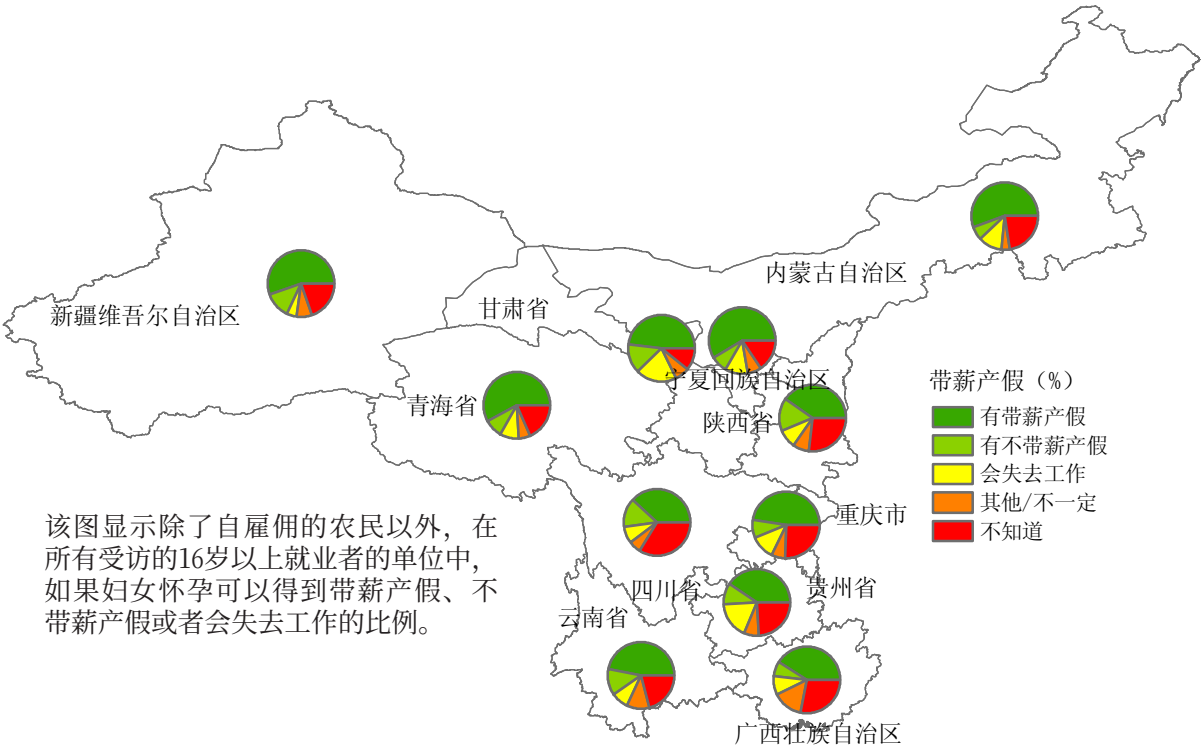


劳动合同



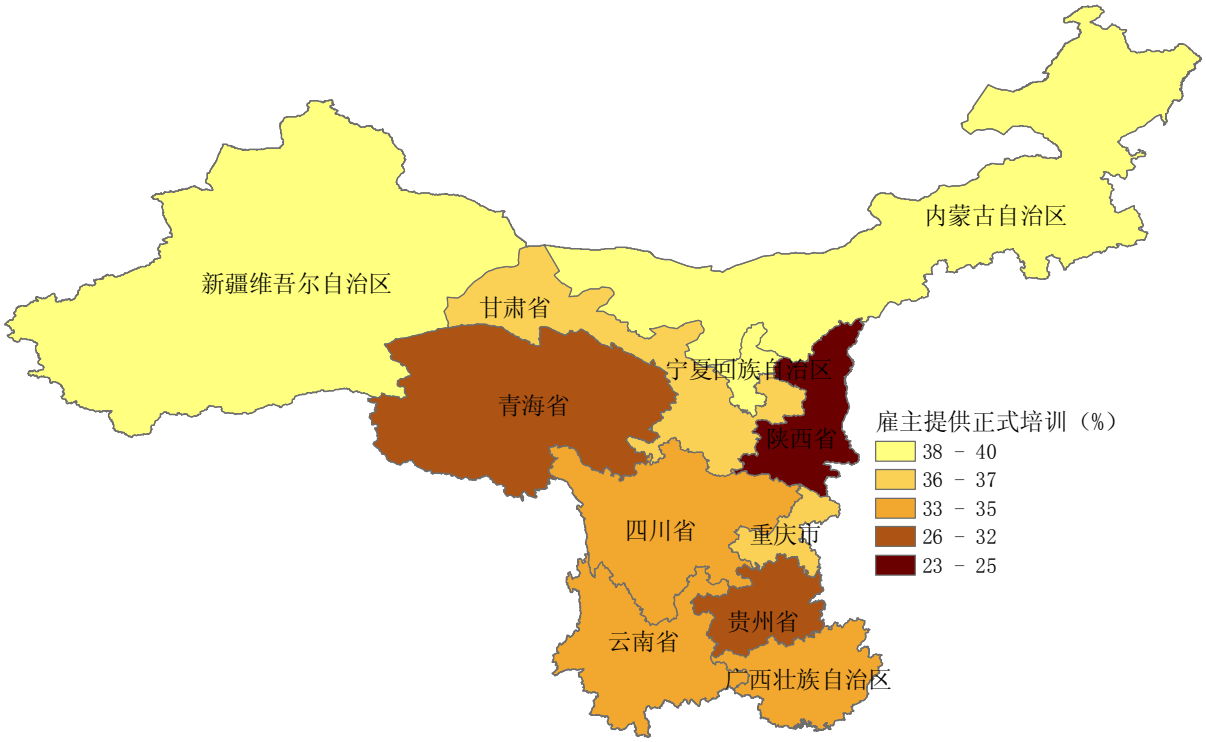
该图显示除了自雇佣的农民以外，所有受访的16岁以上就业者，其主要的那份工作与雇主签订劳动合同的就业者的比例。地图上的圆形分格统计图显示就业者在其主要那份工作中享有带薪假期的比例。

带薪产假

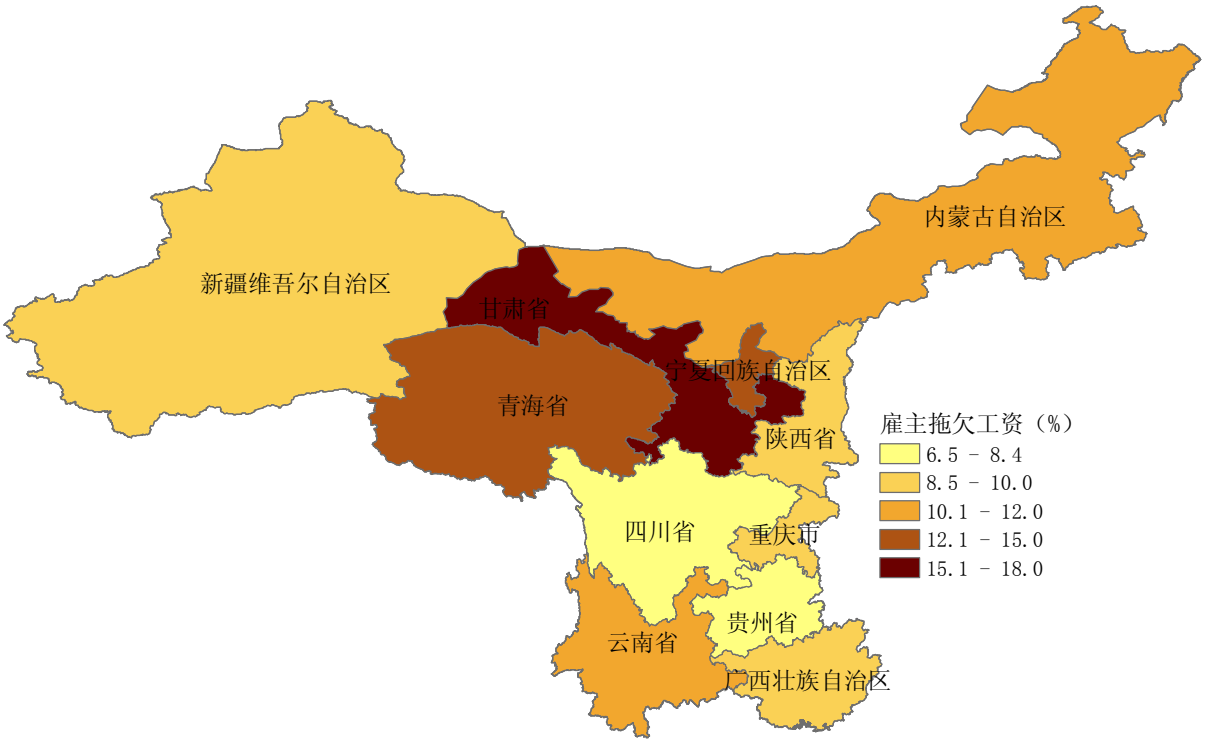


该图显示除了自雇佣的农民以外，在所有受访的16岁以上就业者的单位中，如果妇女怀孕可以得到带薪产假、不带薪产假或者会失去工作的比例。

雇主提供的正式培训



雇主拖欠工资



教育

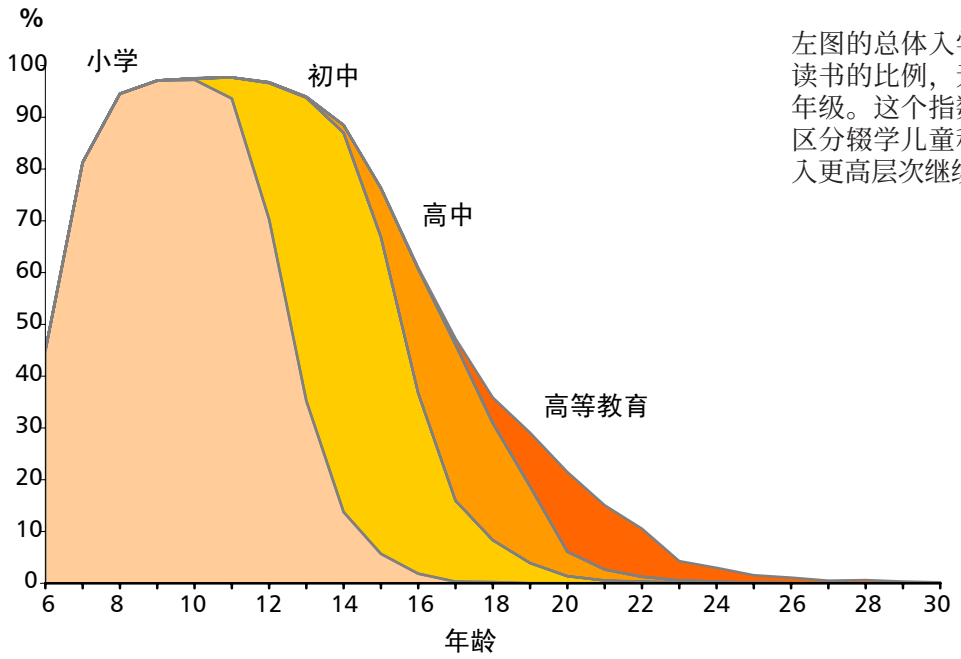
联合国千年发展目标

目标二：普及初等教育

具体目标3：确保到2015年各地儿童不论男女都能完成全部初等教育课程

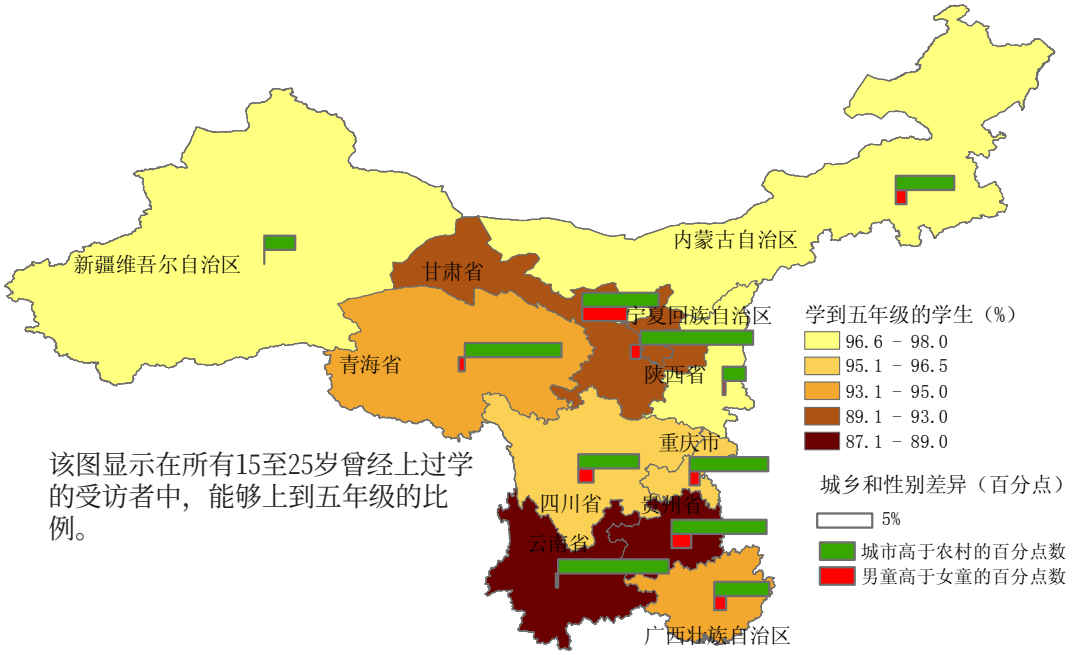
指标7：能够完成五年级学习的学生在所有曾经开始一年级学习的学生中所占的比例	指标8：15至24岁人口的识字率
男孩：94.0% 女孩：93.2% 总共：93.6%	男性：86.8% 女性：84.0% 总共：85.5%

总体入学率



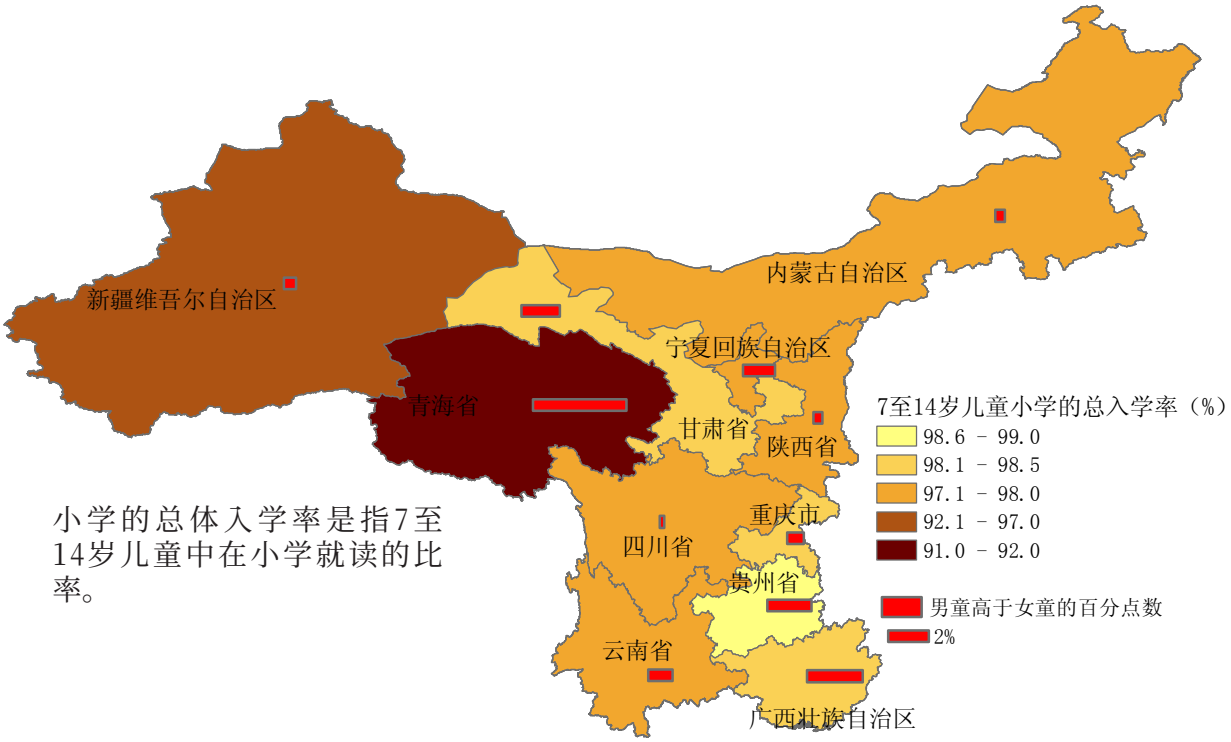
左图的总体入学率计算儿童总体在校读书的比例，无论他们在哪个层次和年级。这个指数使得我们可以有效地区分辍学儿童和完成某个层次学业进入更高层次继续学习的儿童。

学到五年级的学生



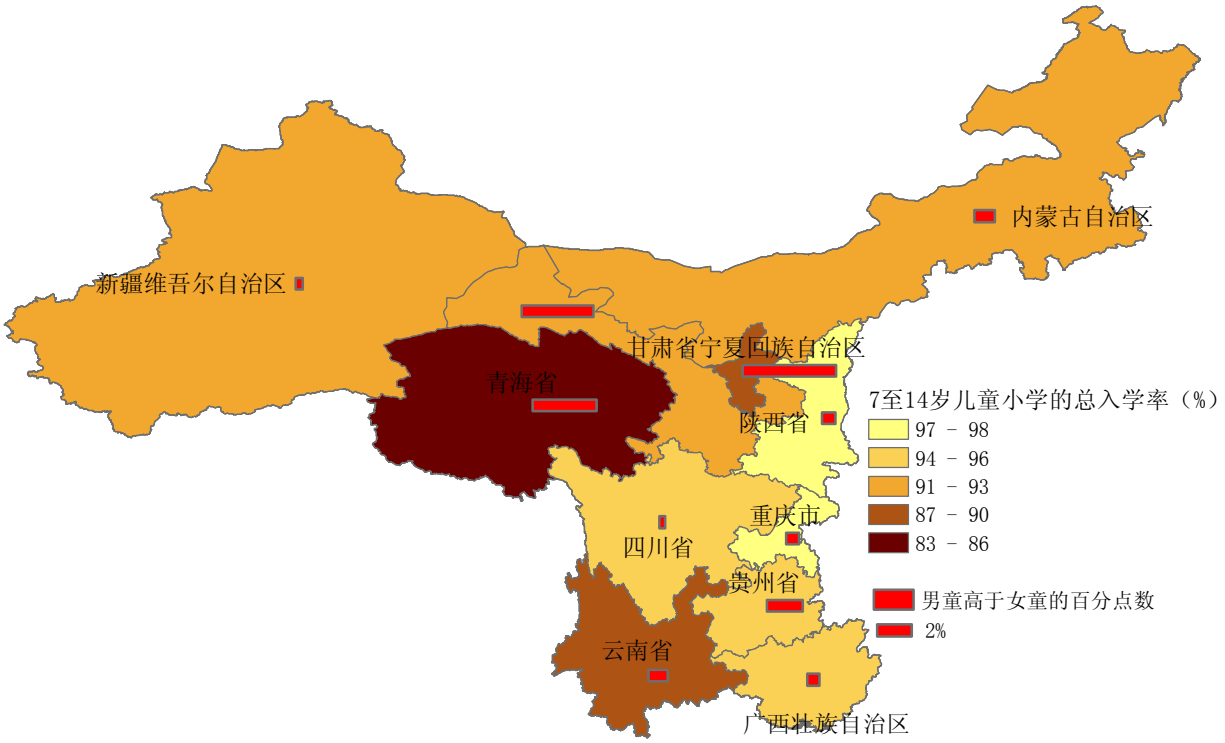
该图显示在所有15至25岁曾经上过学的受访者中，能够上到五年级的比例。

7至14岁儿童小学的总体入学率（城市）

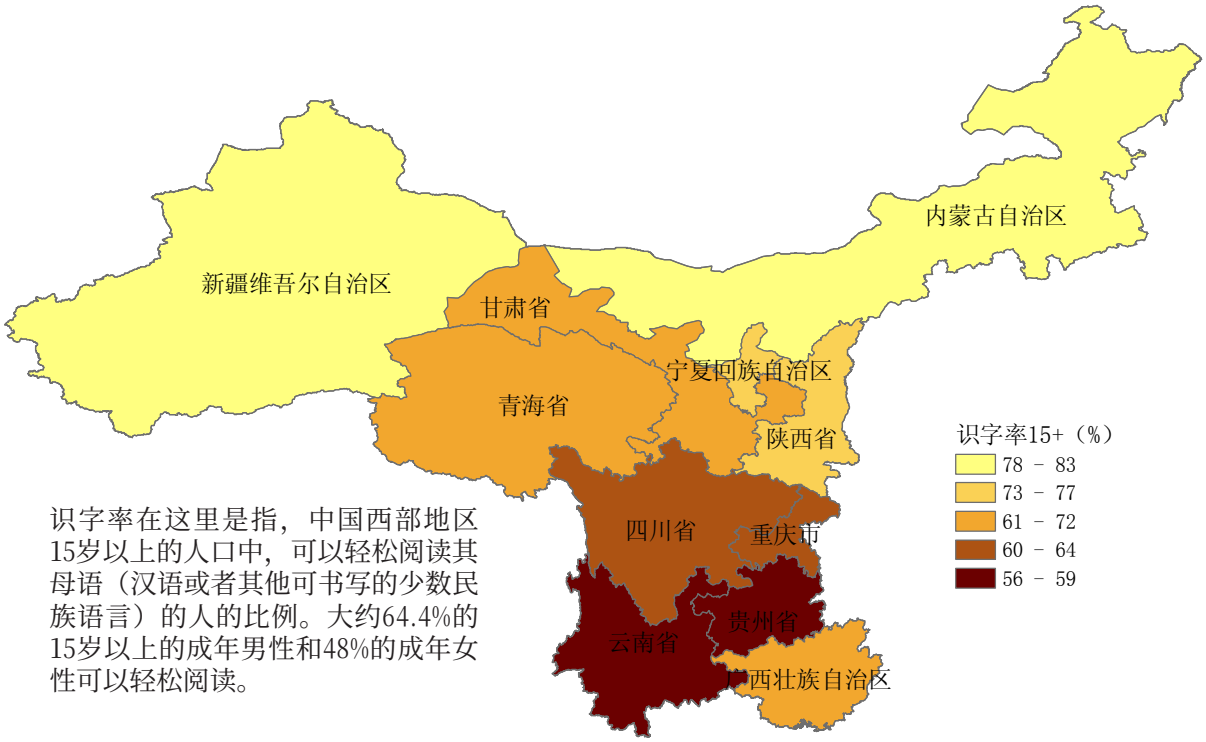


小学的总体入学率是指7至14岁儿童中在小学就读的比率。

7至14岁儿童小学的总体入学率（农村）

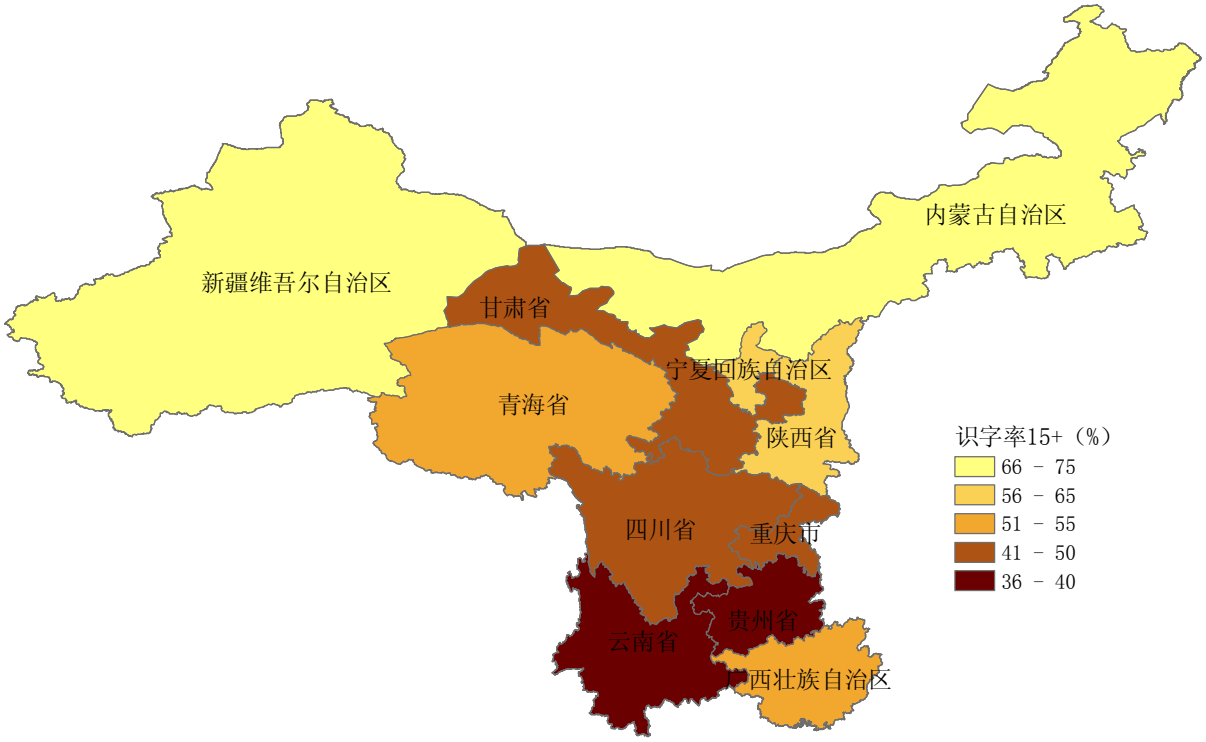


识字率（男性）

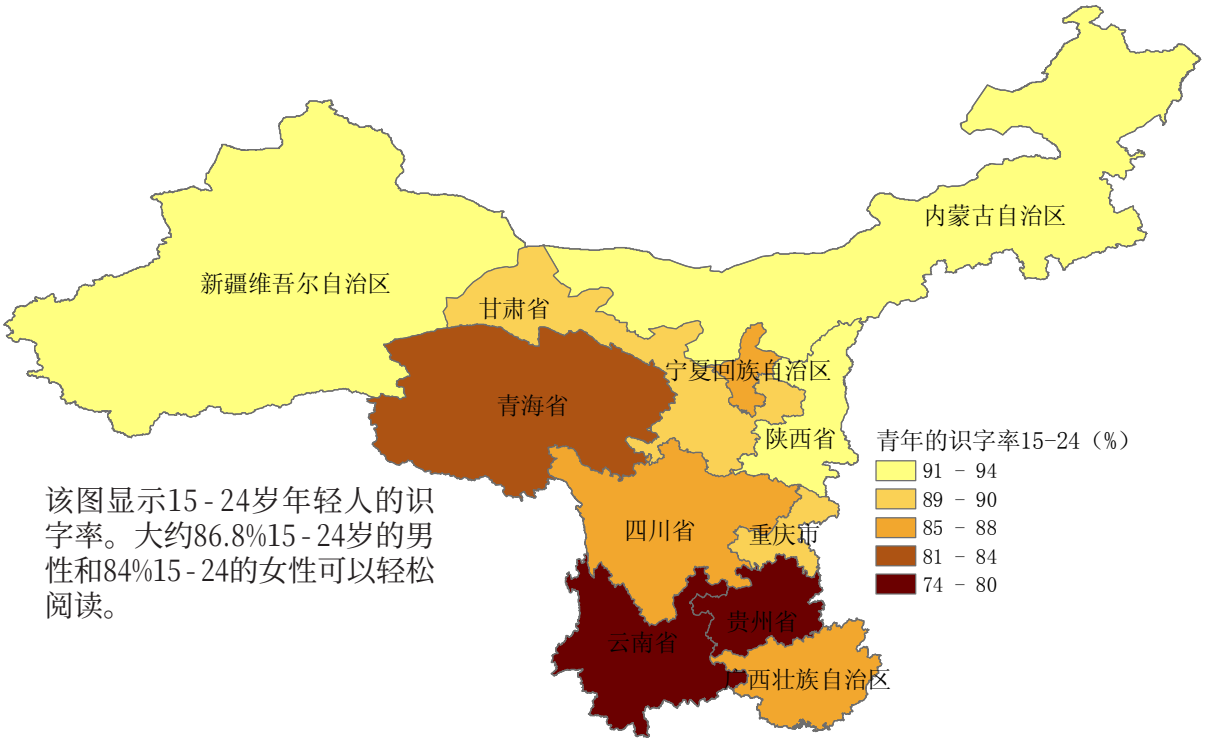


识字率在这里是指，中国西部地区15岁以上的人口中，可以轻松阅读其母语（汉语或者其他可书写的少数民族语言）的人的比例。大约64.4%的15岁以上的成年男性和48%的成年女性可以轻松阅读。

识字率（女性）

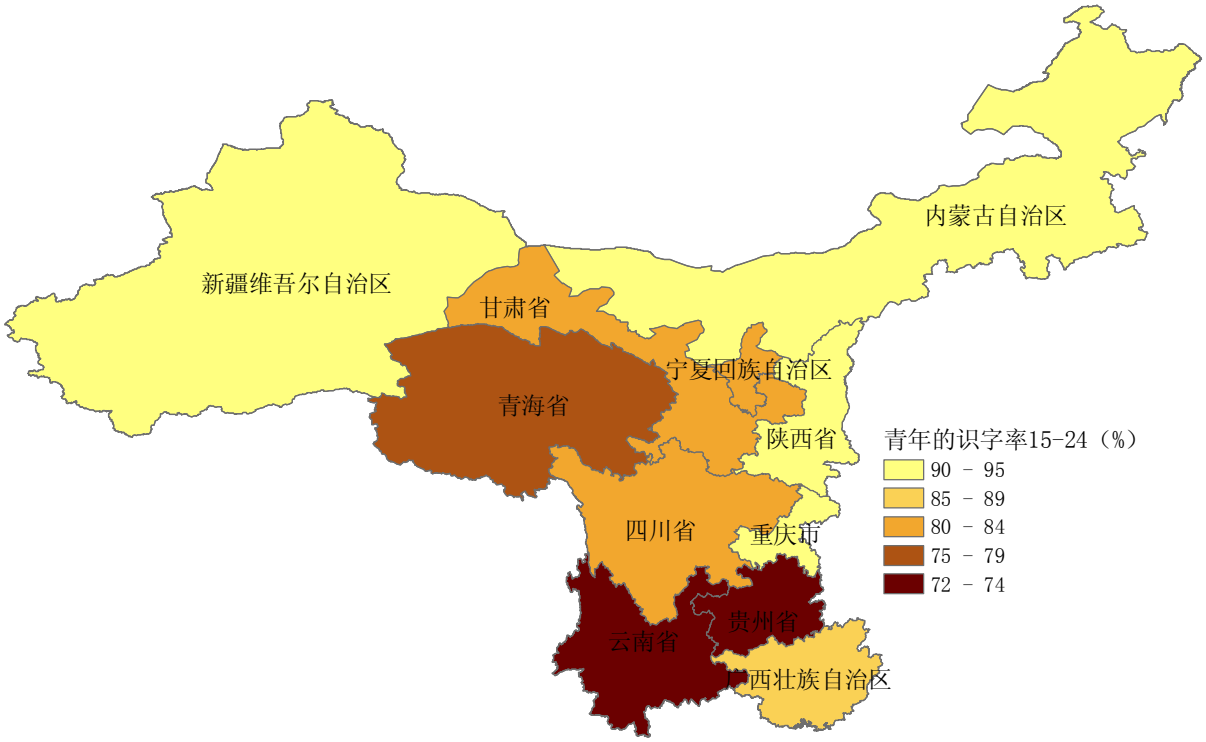


15至24岁人口识字率（男性）



该图显示15-24岁年轻人的识字率。大约86.8%15-24岁的男性和84%15-24的女性可以轻松阅读。

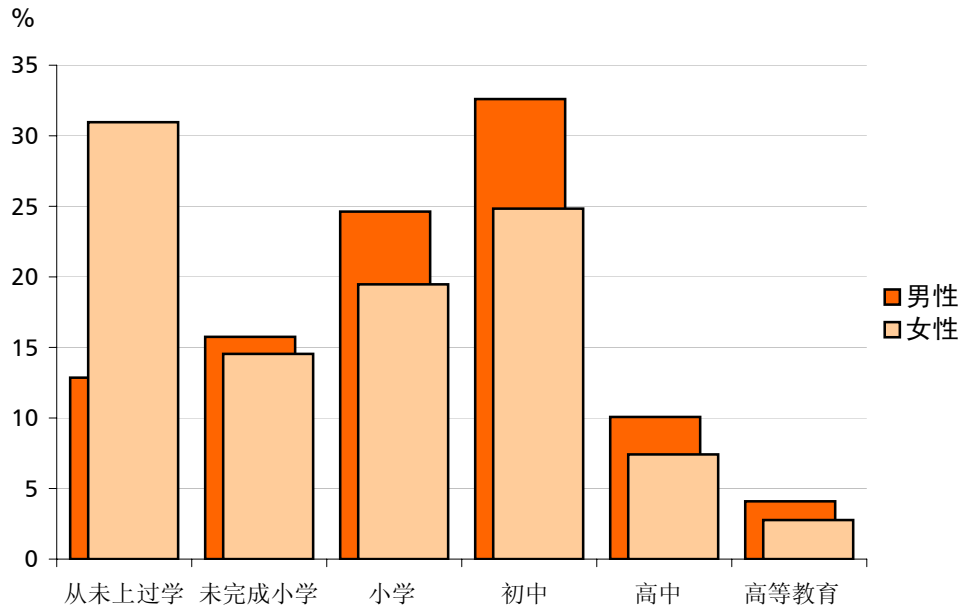
15至24岁人口识字率（女性）



教育和妇女

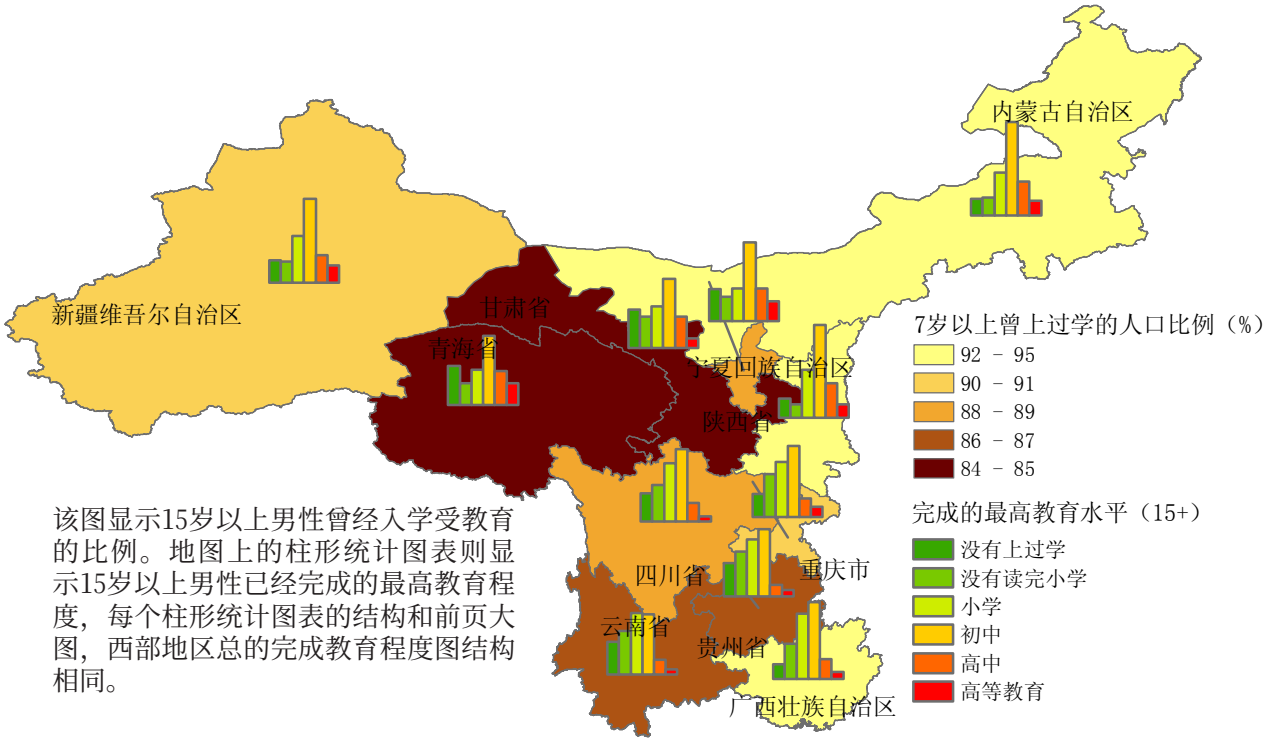
联合国千年发展目标	
目标三：促进男女平等并赋予妇女权力	
具体目标4：最好到2015年消除初等教育和中等教育中的男女差距，并最迟于2015年消除各级教育中的男女差距	
指标9：初等、中等和高等教育中女生相对于男生（100%）的比例	指标10：15至24岁女性相对于男性的识字比例
小学：85.2% 初中：90.3% 高中：80.1% 高等教育：79.0%	比例：85.9%

完成的教育程度



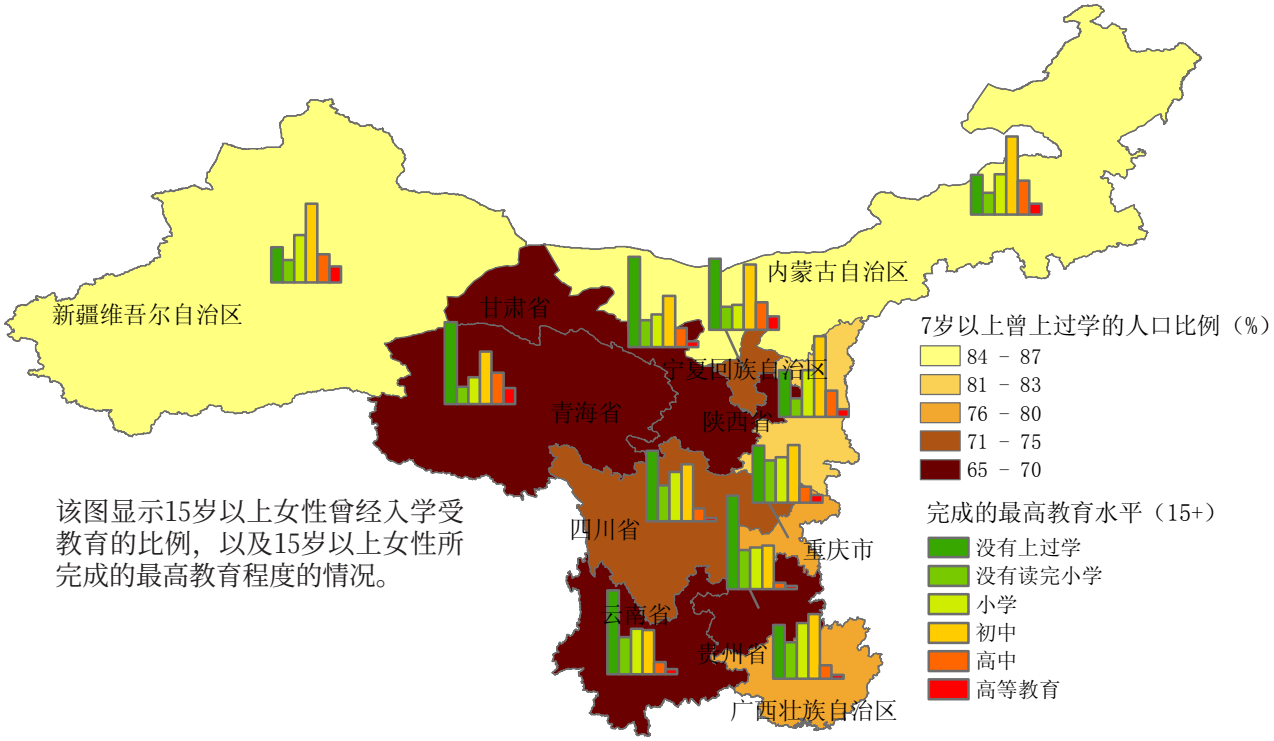
识字率（%）			
年龄组	男性	女性	总共
15-24	87	84	86
25-34	74	63	68
35-44	70	54	62
45-54	56	31	44
55-64	47	23	35
65+	32	9	21

学校教育（男性）



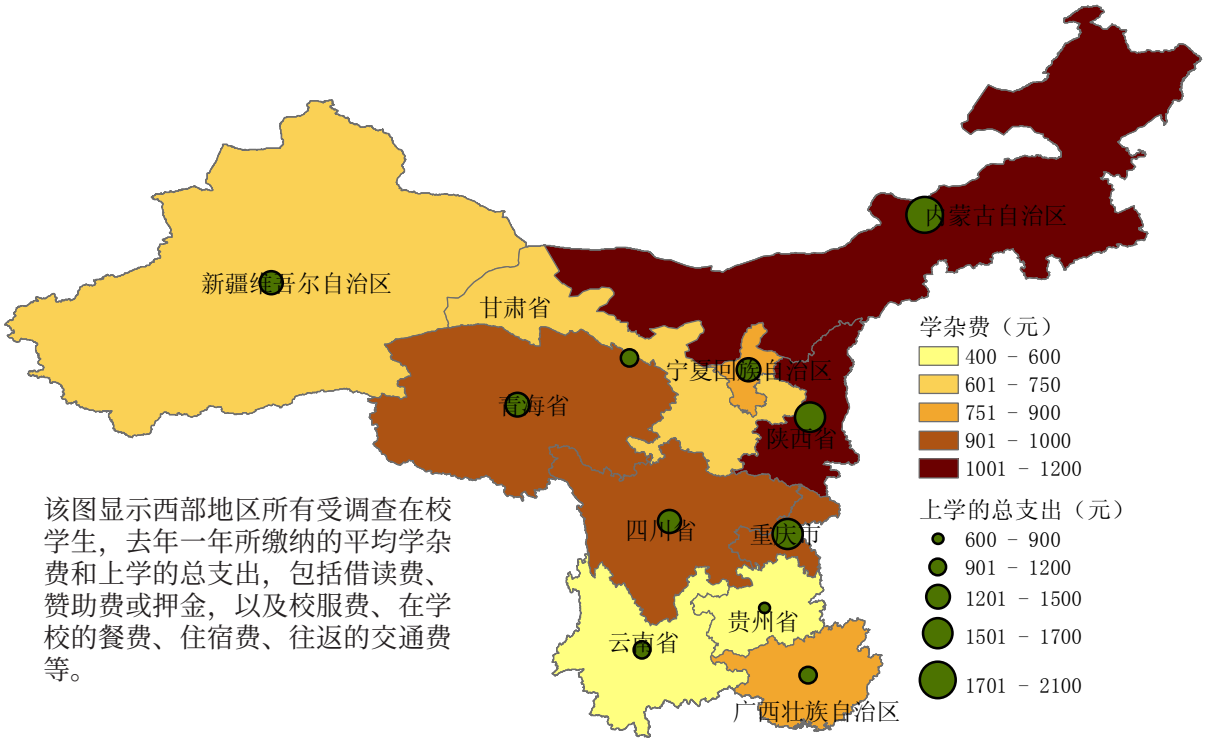
该图显示15岁以上男性曾经入学受教育的比例。地图上的柱形统计图表则显示15岁以上男性已经完成的最高教育程度，每个柱形统计图表的结构和前页大图，西部地区总的完成教育程度图结构相同。

学校教育（女性）



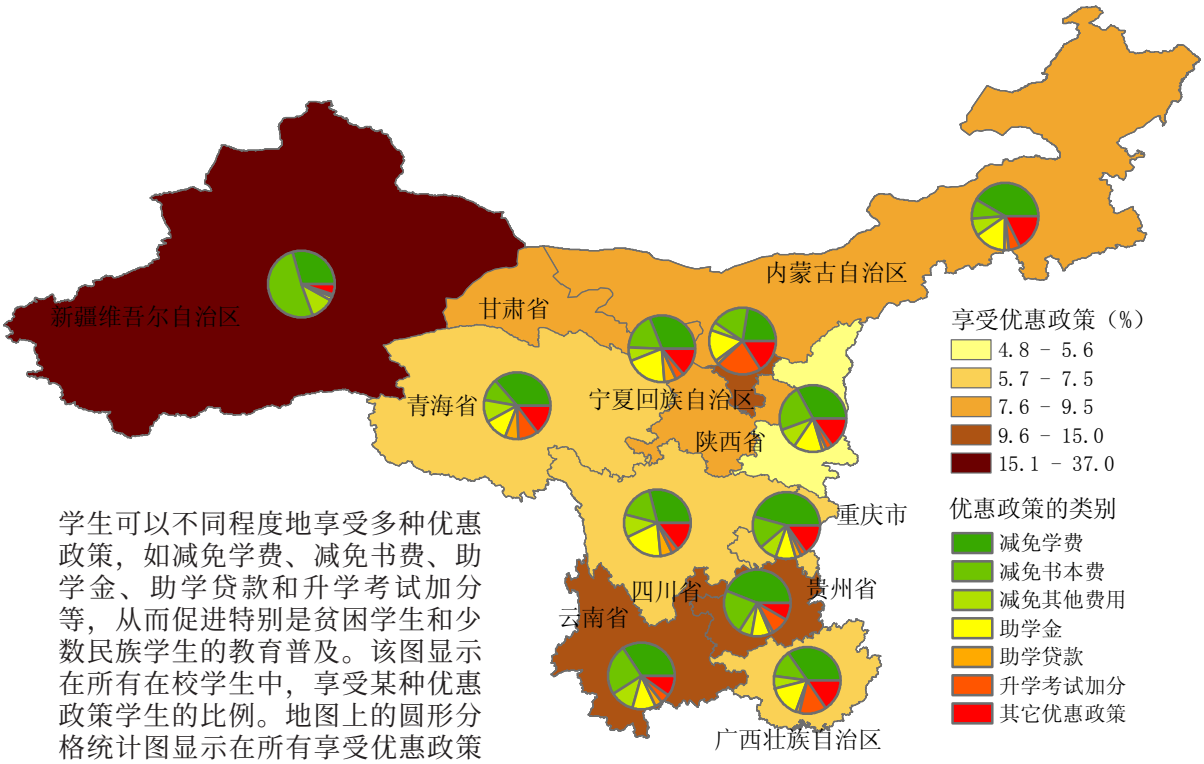
该图显示15岁以上女性曾经入学受教育的比例，以及15岁以上女性所完成的最高教育程度的情况。

学杂费和上学总支出



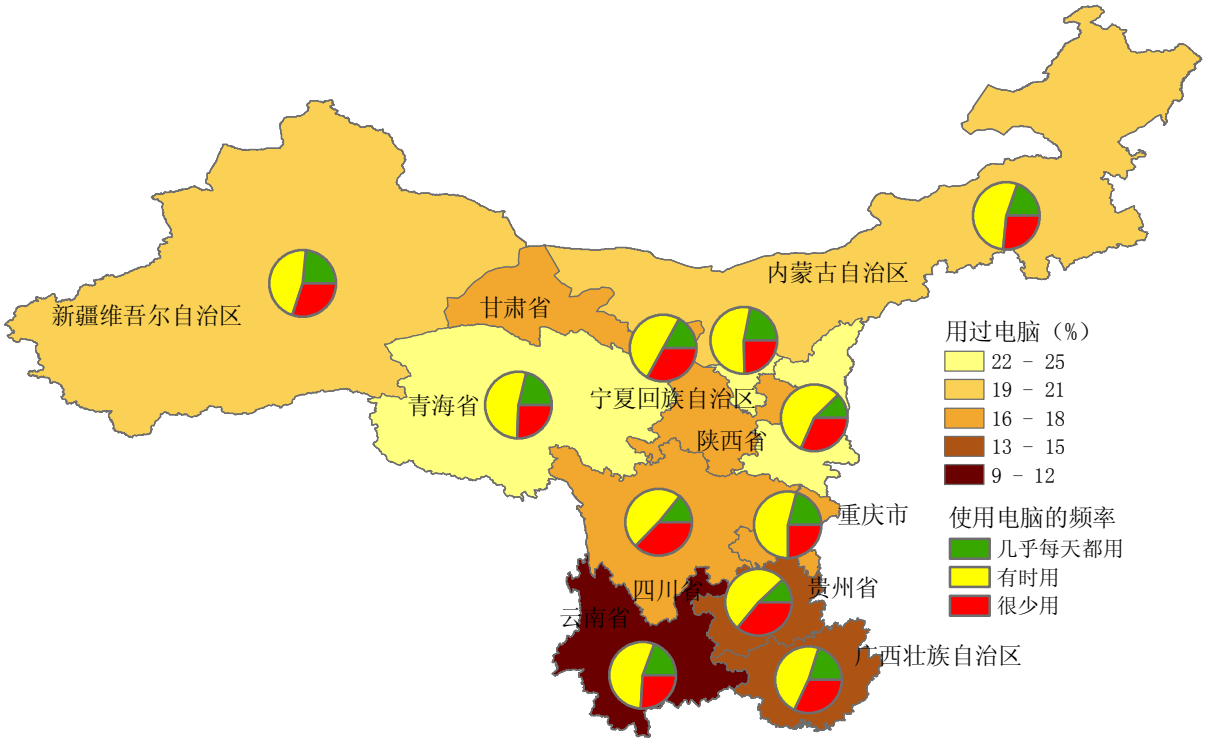
该图显示西部地区所有受调查在校学生，去年一年所缴纳的平均学杂费和上学的总支出，包括借读费、赞助费或押金，以及校服费、在学校的餐费、住宿费、往返的交通费等。

优惠政策

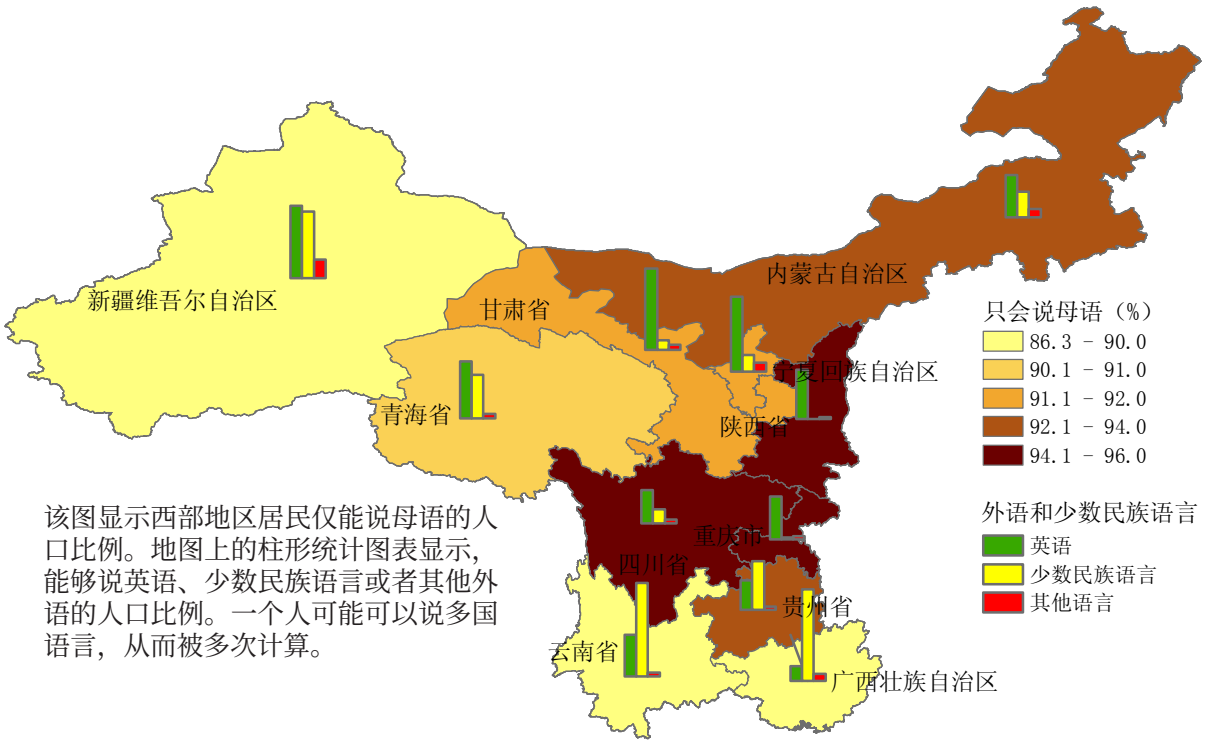


学生可以不同程度地享受多种优惠政策，如减免学费、减免书费、助学金、助学贷款和升学考试加分等，从而促进特别是贫困学生和少数民族学生的教育普及。该图显示在所有在校学生中，享受某种优惠政策学生的比例。地图上的圆形分格统计图显示在所有享受优惠政策的学生中，各种优惠政策的分布情况。

电脑的普及

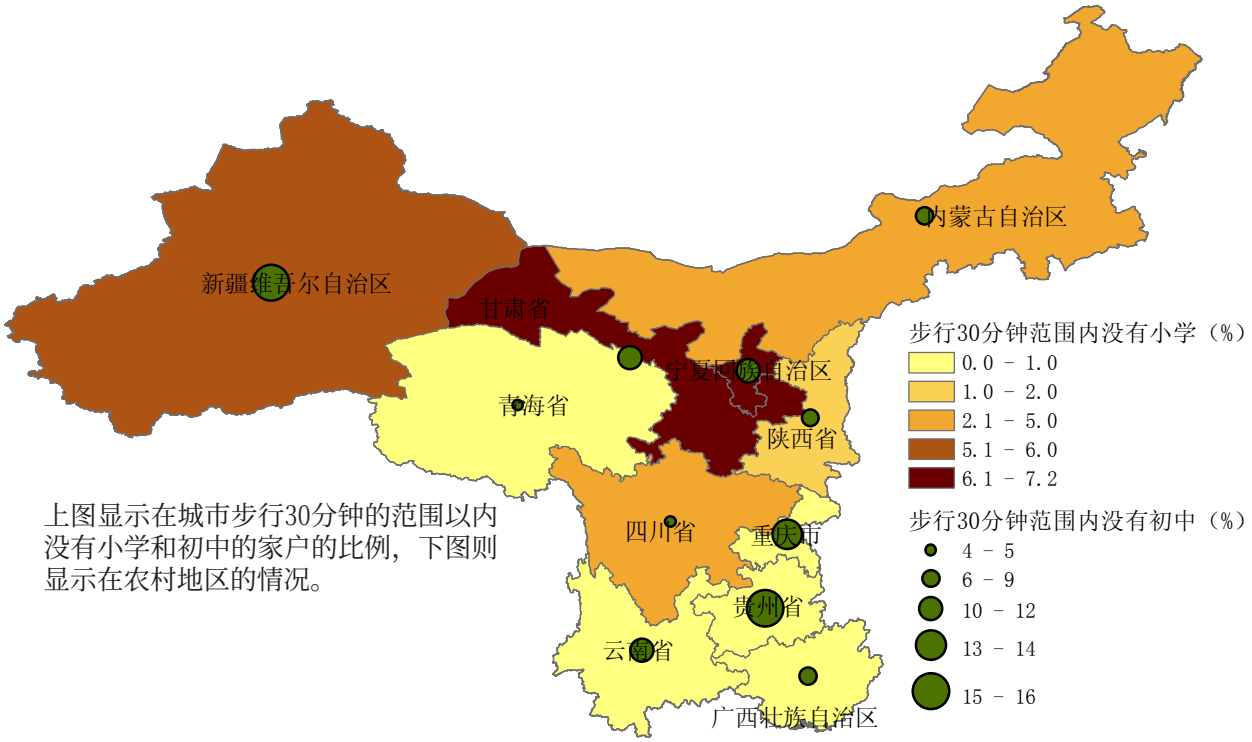


母语和其他语言的掌握



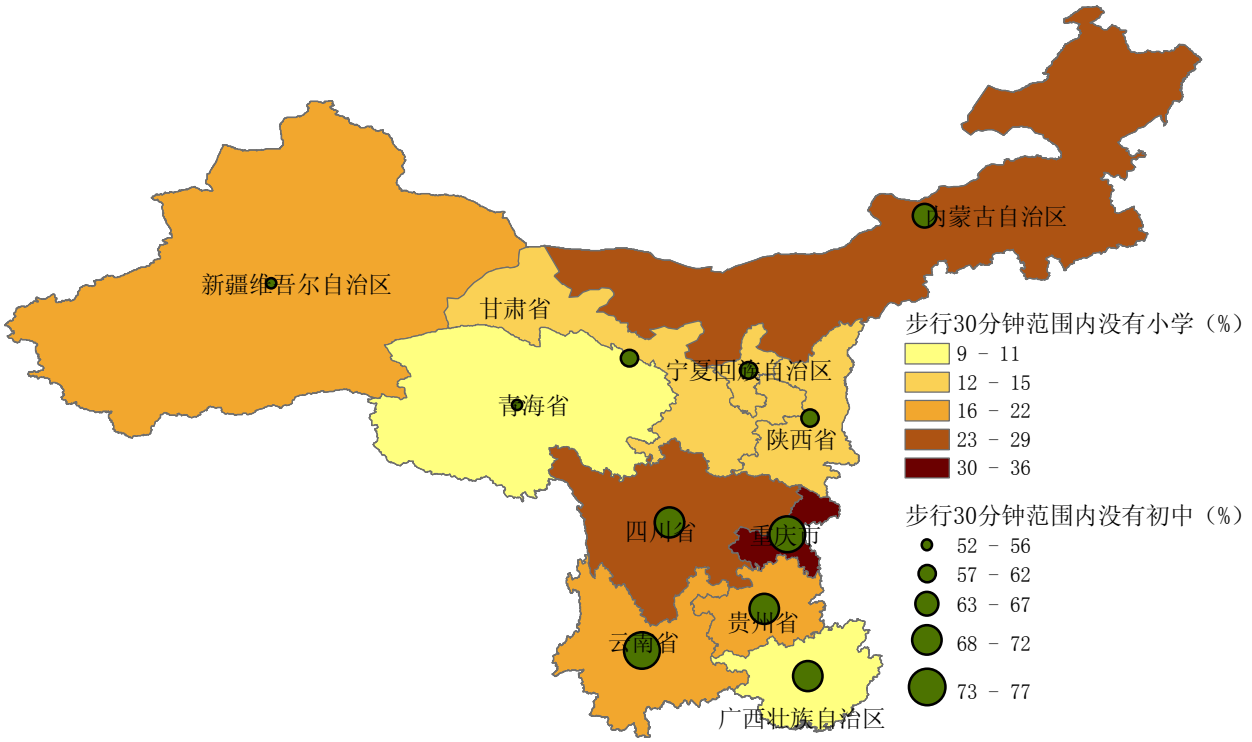
该图显示西部地区居民仅能说母语的人口比例。地图上的柱形统计图表显示，能够说英语、少数民族语言或者其他外语的人口比例。一个人可能可以说多国语言，从而被多次计算。

小学和初中的覆盖率（城市）



上图显示在城市步行30分钟的范围以内没有小学和初中的家户的比例，下图则显示在农村地区的情况。

小学和初中的覆盖率（农村）



住房和基础建设

联合国千年发展目标

目标七：确保环境的可持续能力

具体目标10：到2015年将无法持续获得安全饮用水的人口比例减半

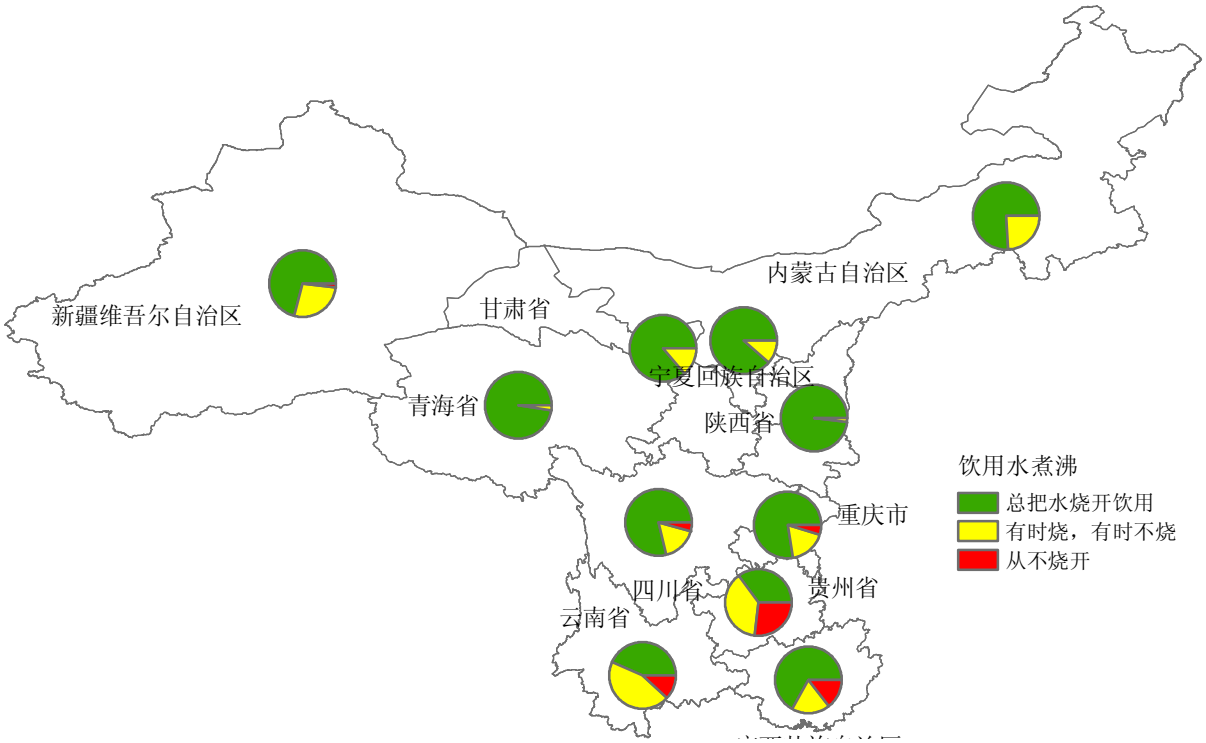
指标29：城市和农村可以持续获得安全水源的人口比例

总体： 55%
城市： 89%
农村： 44%

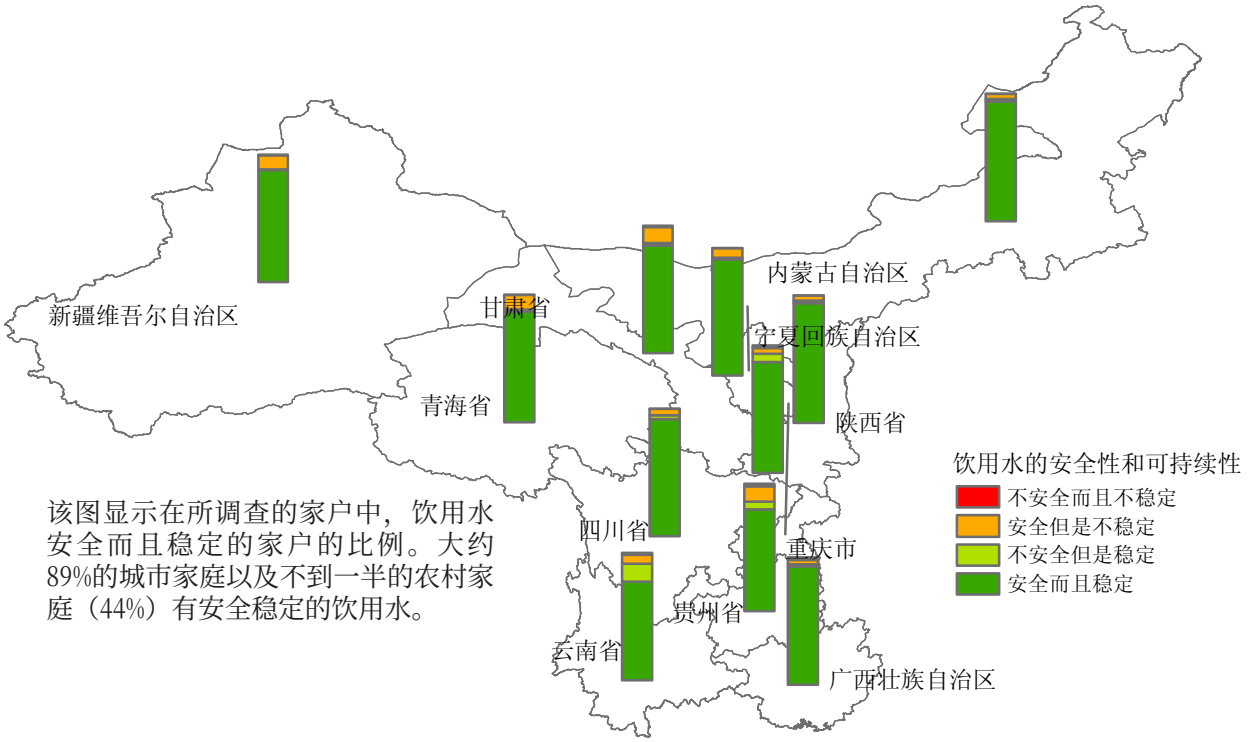
联合国千年发展目标要求饮用水不仅要安全，而且要具有可持续性。“安全性”是指水的来源，来自“安全”水源的水被认为是“安全”的。而所谓的“安全”水源是指，水厂的管道水，有井盖的井，压力井，瓶装水或者桶装水，或者水车的运输水。

水的“可持续性”是指家户的水供给没有出现过一整天以上的断水情况。

饮用水煮沸

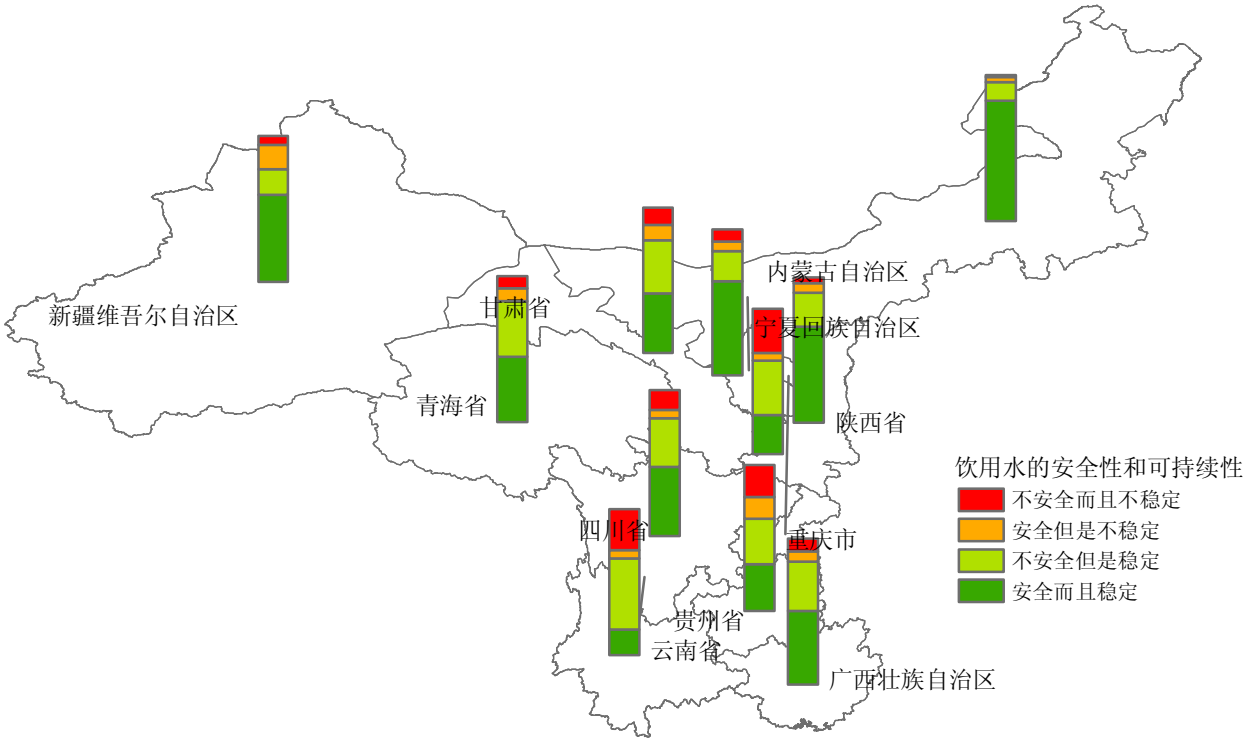


饮用水的安全性和可持续性（城市）

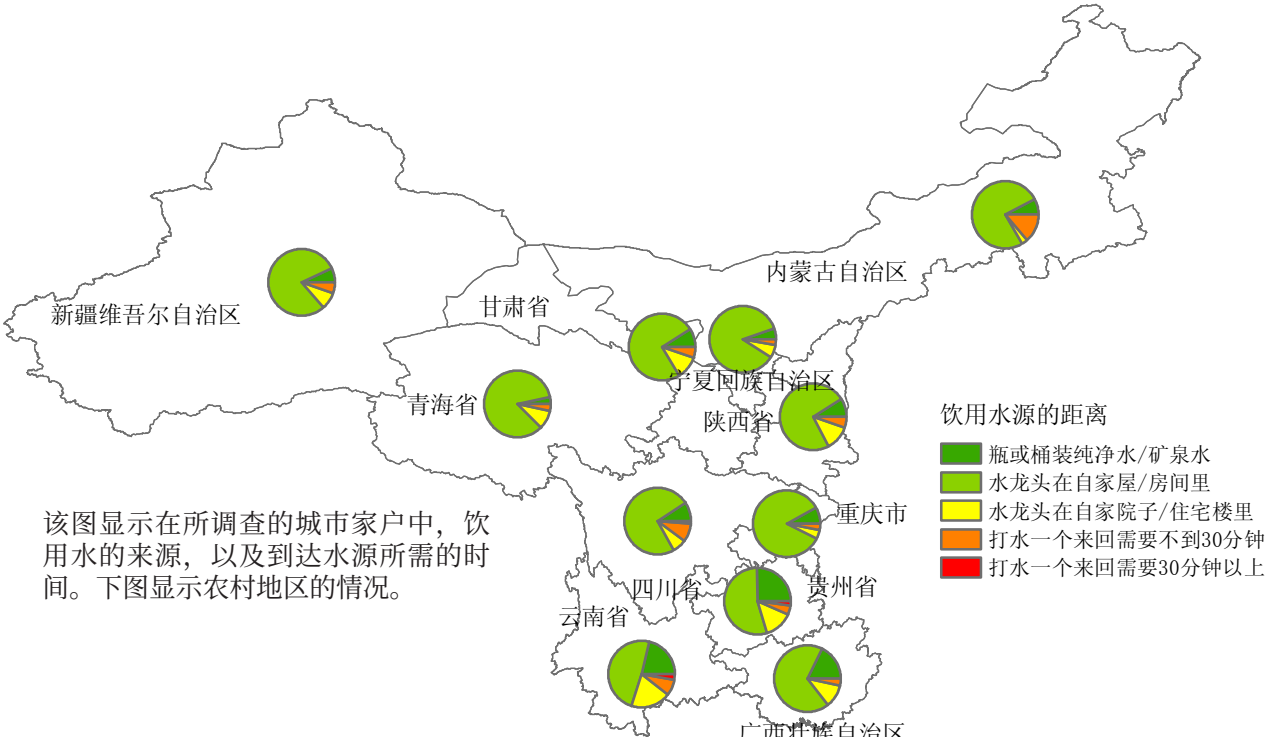


该图显示在所调查的家户中，饮用水安全而且稳定的家户的比例。大约89%的城市家庭以及不到一半的农村家庭（44%）有安全稳定的饮用水。

饮用水的安全性和可持续性（农村）

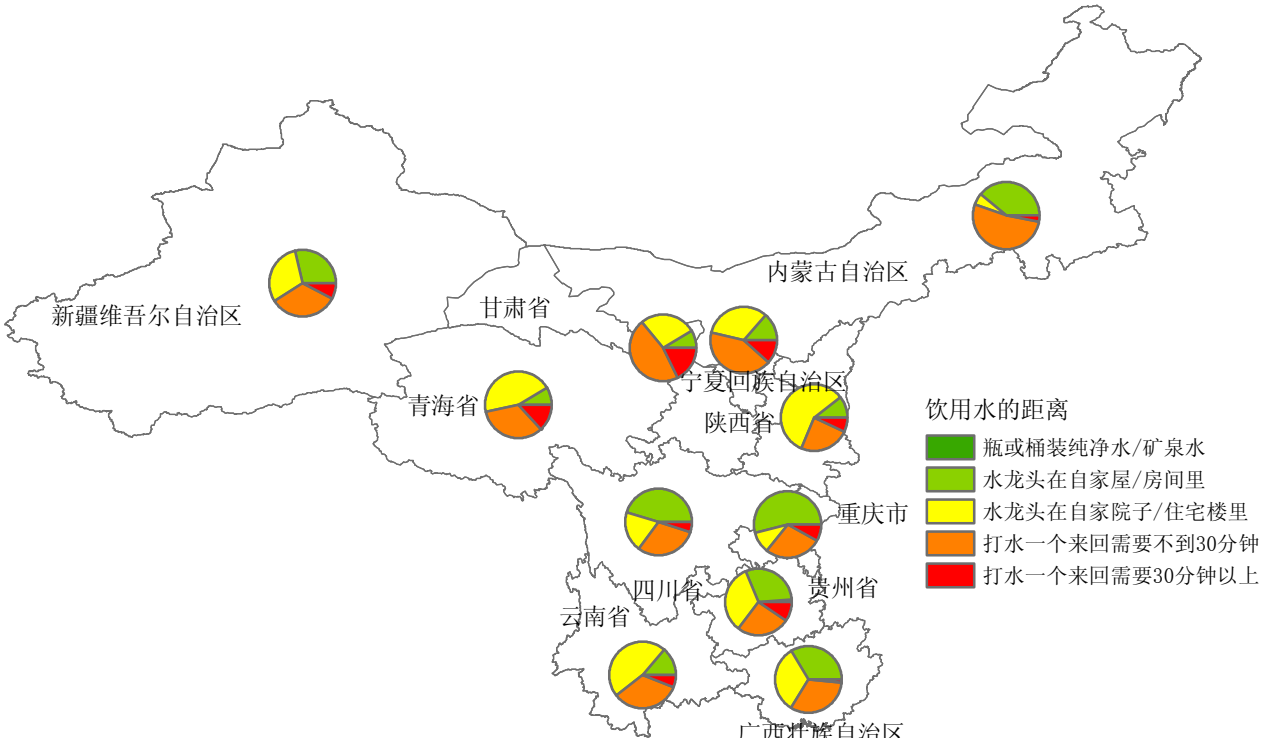


饮用水源的距离（城市）

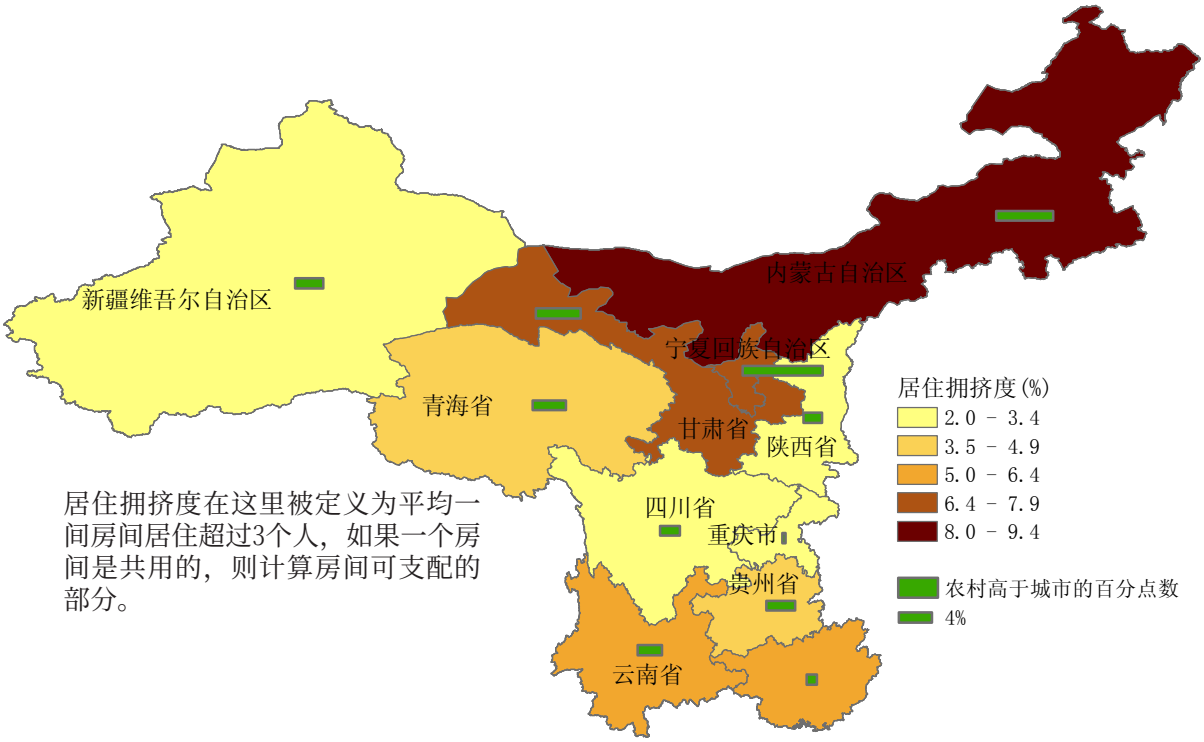


该图显示在所调查的城市家户中，饮用水的来源，以及到达水源所需的时间。下图显示农村地区的情况。

饮用水源的距离（农村）

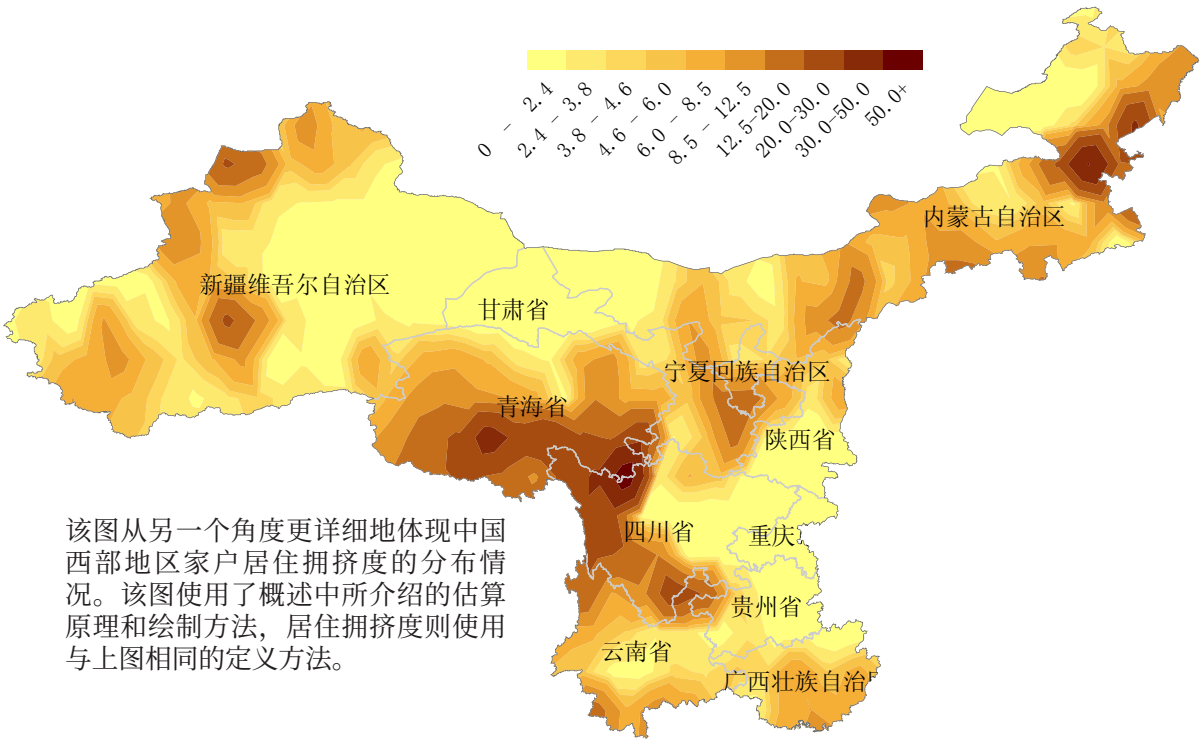


居住拥挤度



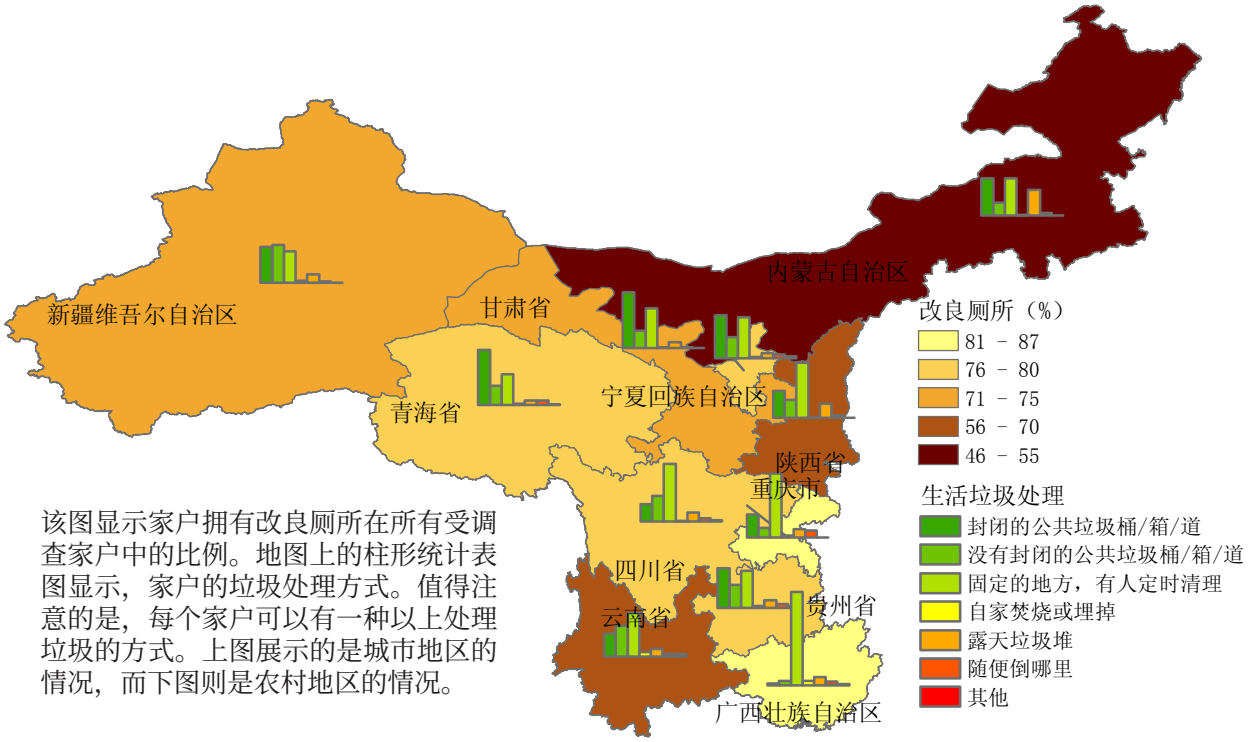
居住拥挤度在这里被定义为平均一间房间居住超过3个人，如果一个房间是共用的，则计算房间可支配的部分。

居住拥挤度：详细分布图



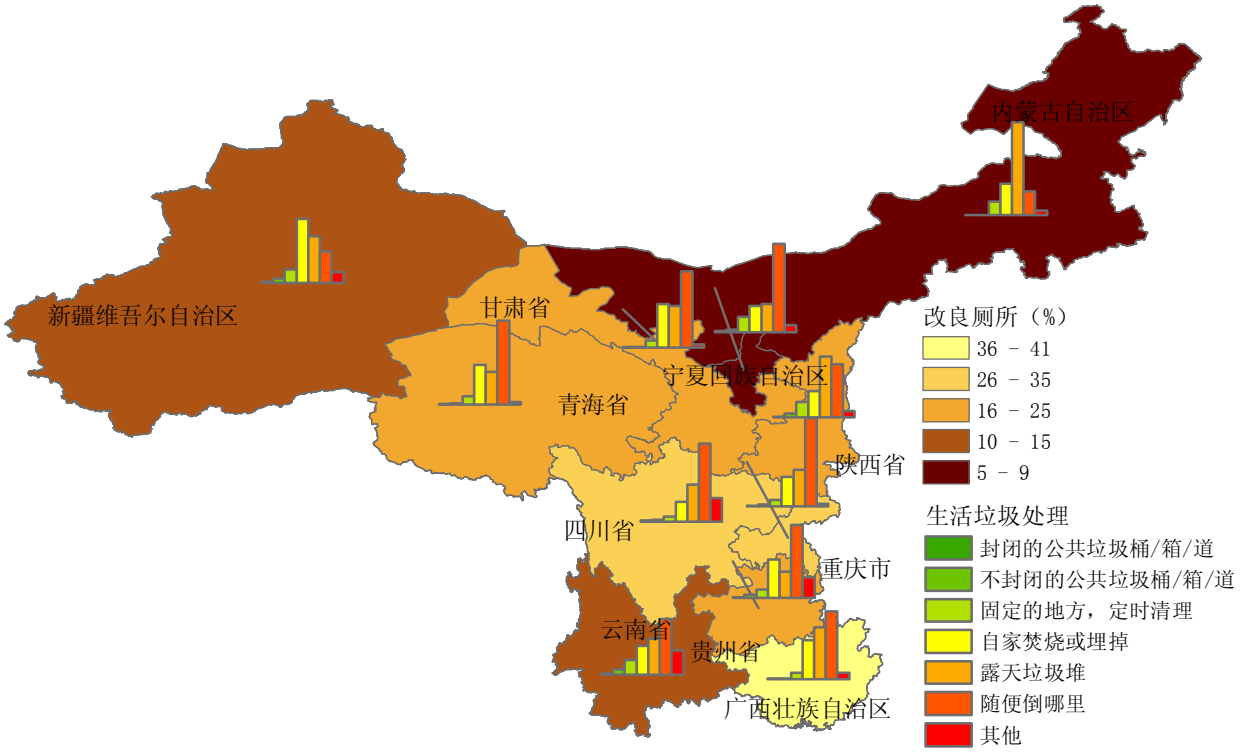
该图从另一个角度更详细地体现中国西部地区家户居住拥挤度的分布情况。该图使用了概述中所介绍的估算原理和绘制方法，居住拥挤度则使用与上图相同的定义方法。

改良厕所和生活垃圾处理（城市）

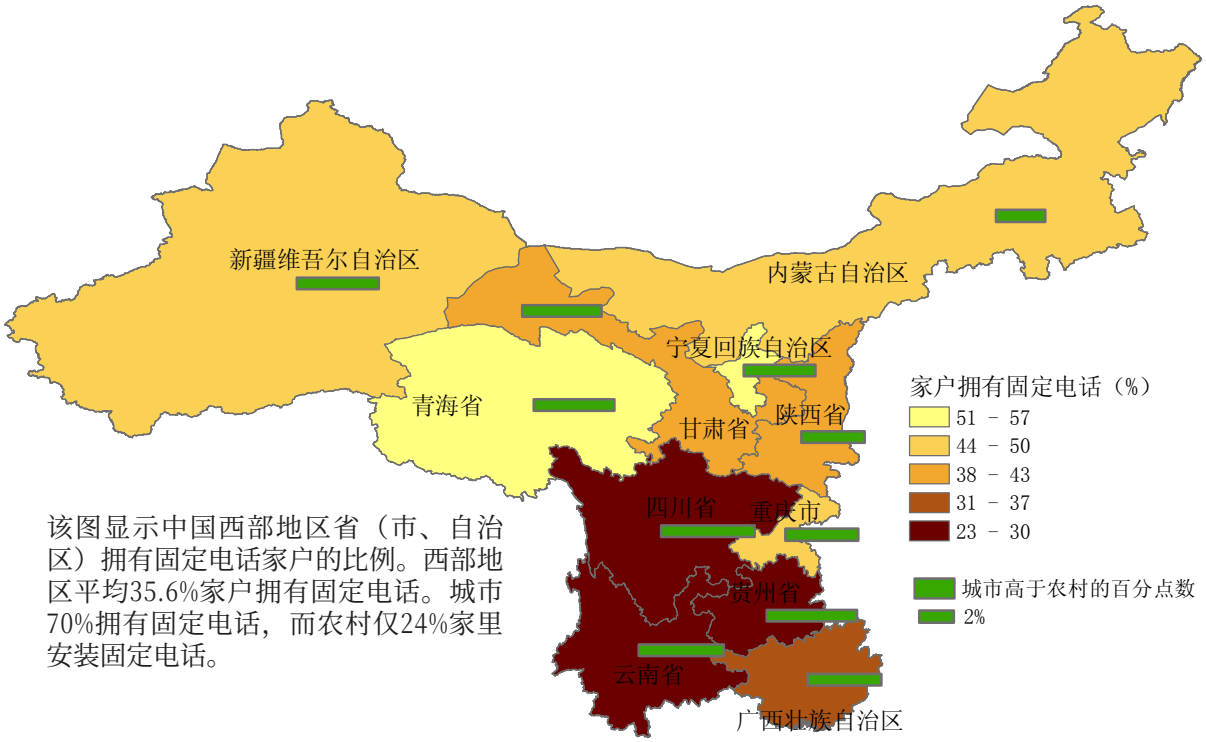


该图显示家户拥有改良厕所在所有受调查家户中的比例。地图上的柱形统计表图显示，家户的垃圾处理方式。值得注意的是，每个家户可以有一种以上处理垃圾的方式。上图展示的是城市地区的情况，而下图则是农村地区的情况。

改良厕所和生活垃圾处理（农村）

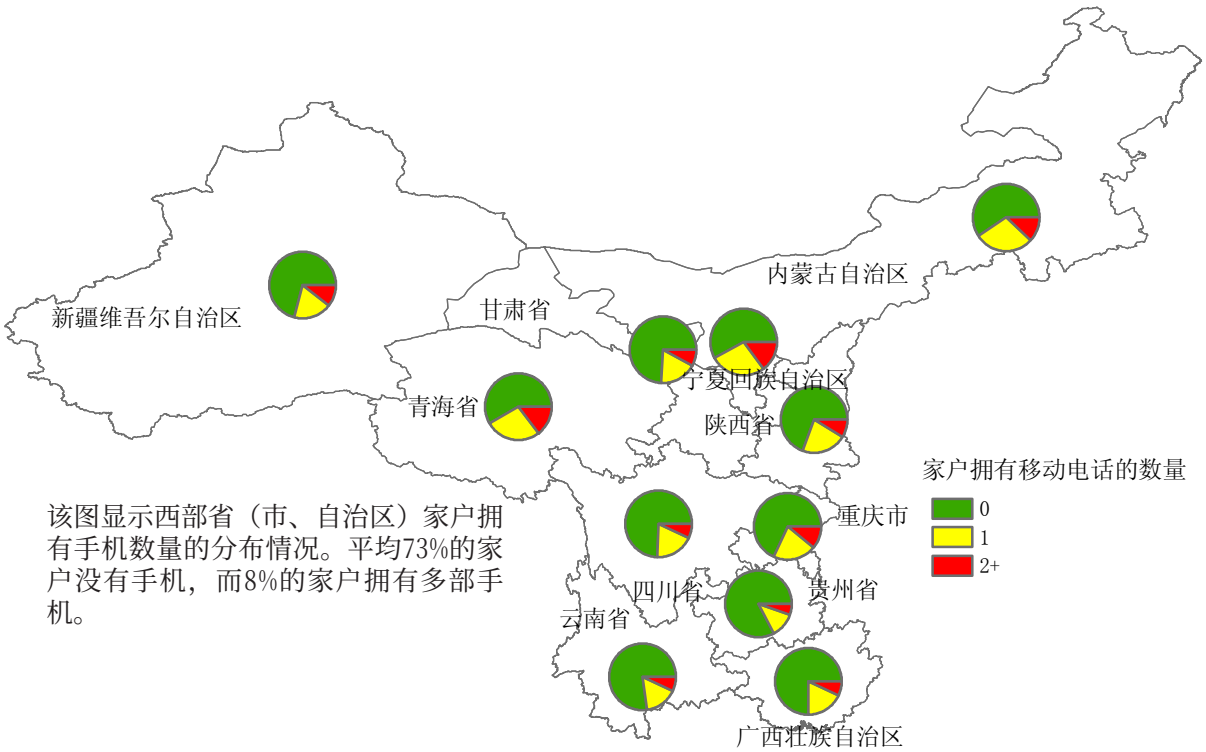


固定电话



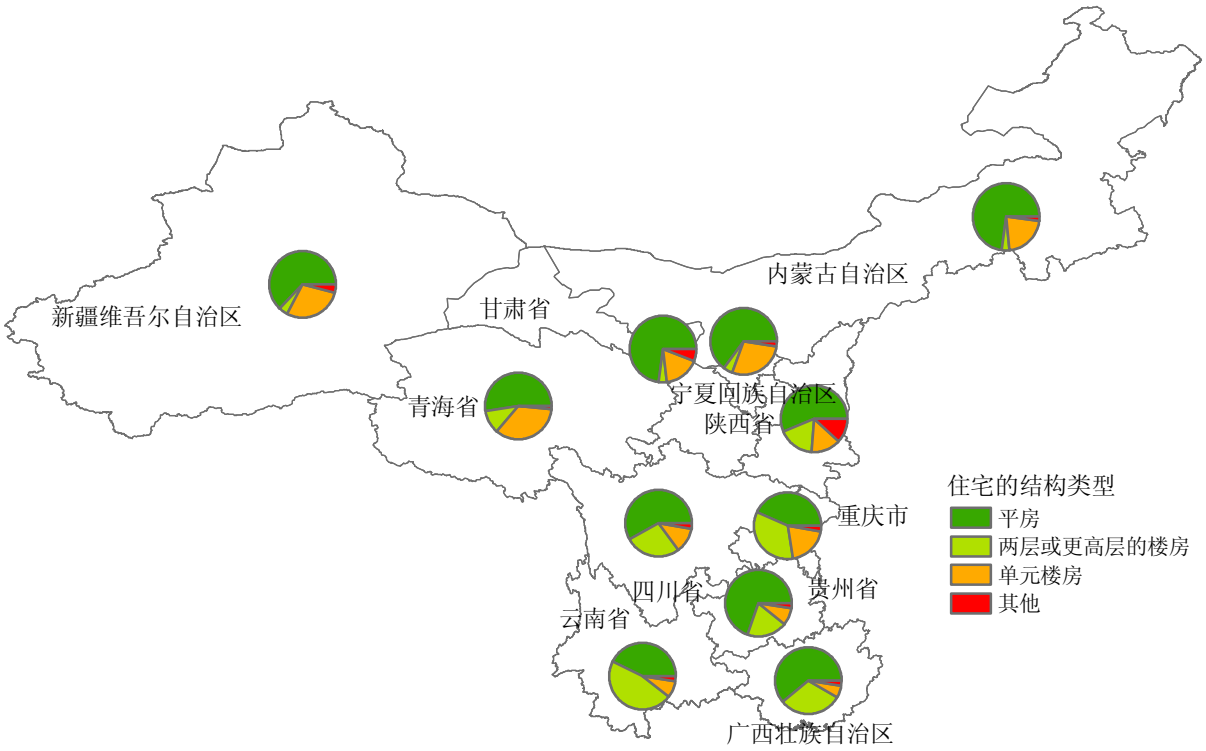
该图显示中国西部地区省（市、自治区）拥有固定电话家户的比例。西部地区平均35.6%家户拥有固定电话。城市70%拥有固定电话，而农村仅24%家里安装固定电话。

移动电话

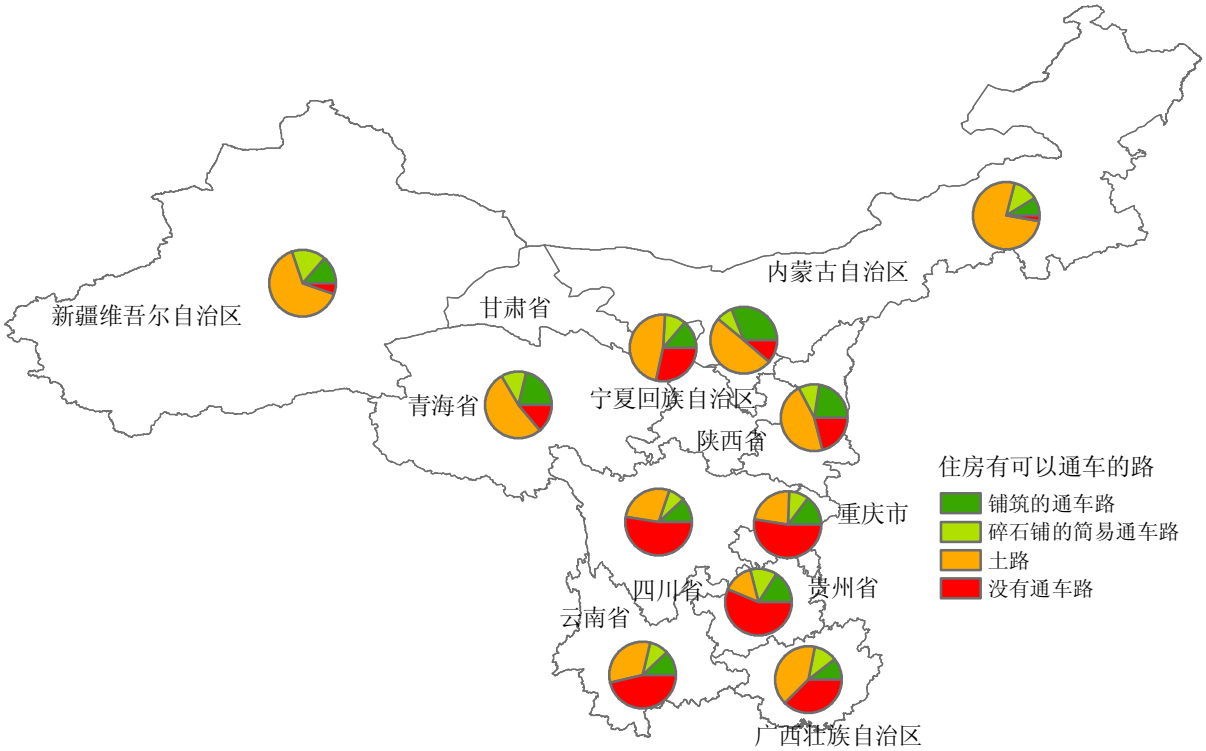


该图显示西部省（市、自治区）家户拥有手机数量的分布情况。平均73%的家户没有手机，而8%的家户拥有多部手机。

住宅的结构类型



农村地区的道路



收入和财富

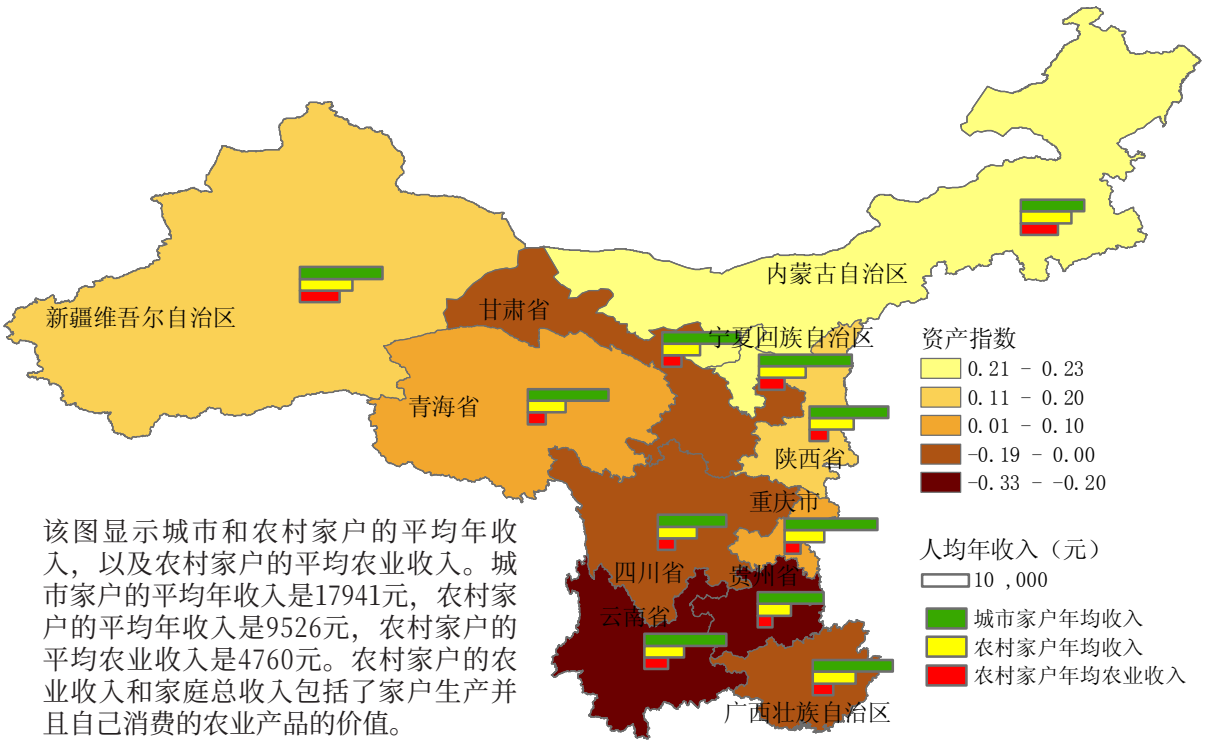
所有受调查的家户都被要求列举受访前一年时间内各种家庭和个人的具体收入。图中所描绘的家庭总收入是由各种调查中所记录的收入进行加总所得的。

农村家户的总收入和农村家户的农业收入也包括了家户本身消费的农业产品的价值。

家户也被问及是否拥有各种家庭耐用消费品。资产指数就是一个线性的以家庭所拥有的资产为基础计算的一种财富指数。资产指数的计算是基于家户所拥有的耐用消费品和其他与房屋、水供给和卫生设备以及其他提高家户生活满意度的设施和装备。

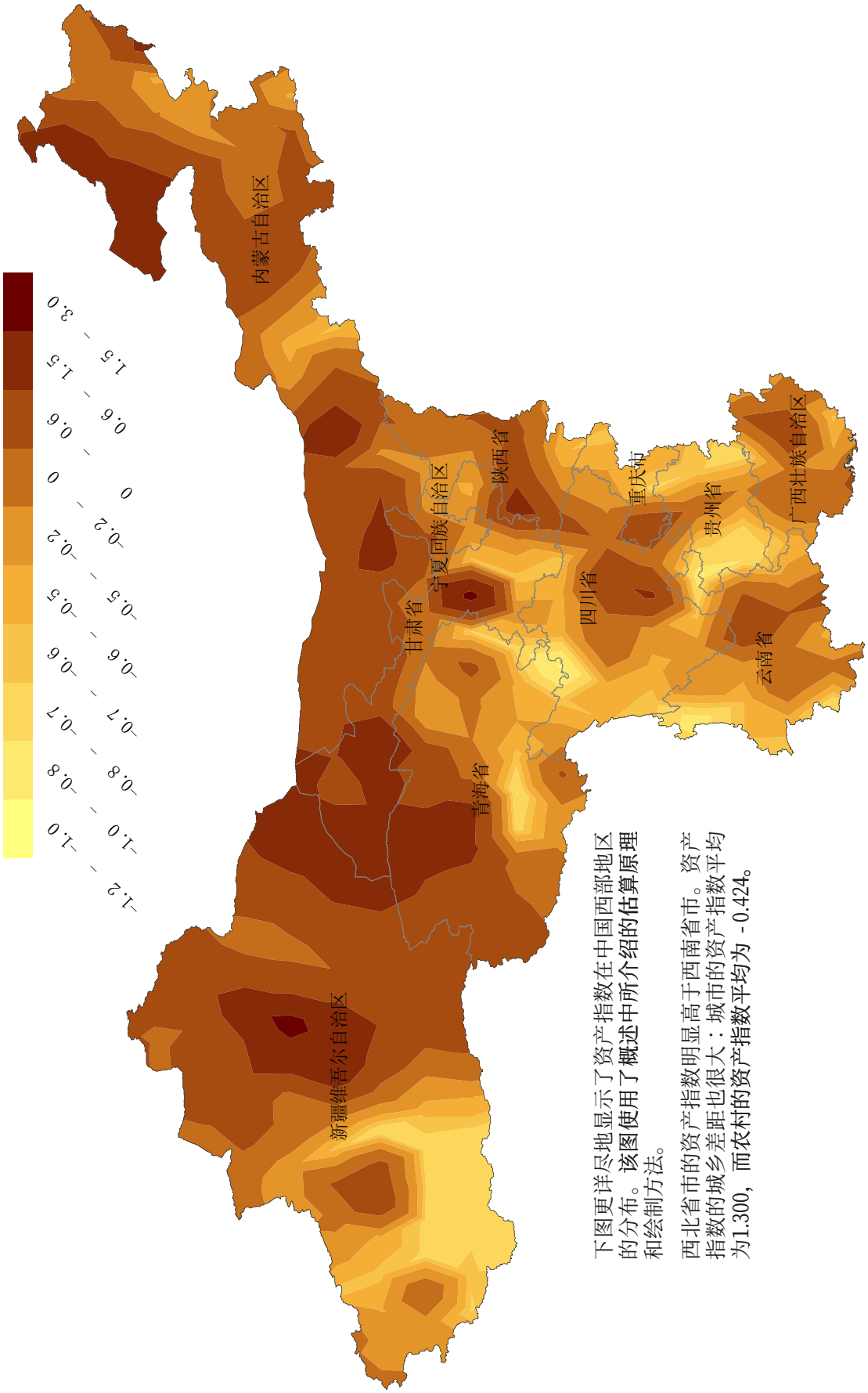
资产指数是一个非常好的代表家庭财富的指数。在计算资产指数的时候，我们应用主要资产分析方法（PCA）来推导每个资产所使用的权重，并把他们加权加总成一个代表家户经济状态的指数。西部地区所有家户的资产指数的加总为0。

年家庭总收入和资产指数



该图显示城市和农村家户的平均年收入，以及农村家户的平均农业收入。城市家户的平均年收入是17941元，农村家户的平均年收入是9526元，农村家户的平均农业收入是4760元。农村家户的农业收入和家庭总收入包括了家户生产并且自己消费的农业产品的价值。

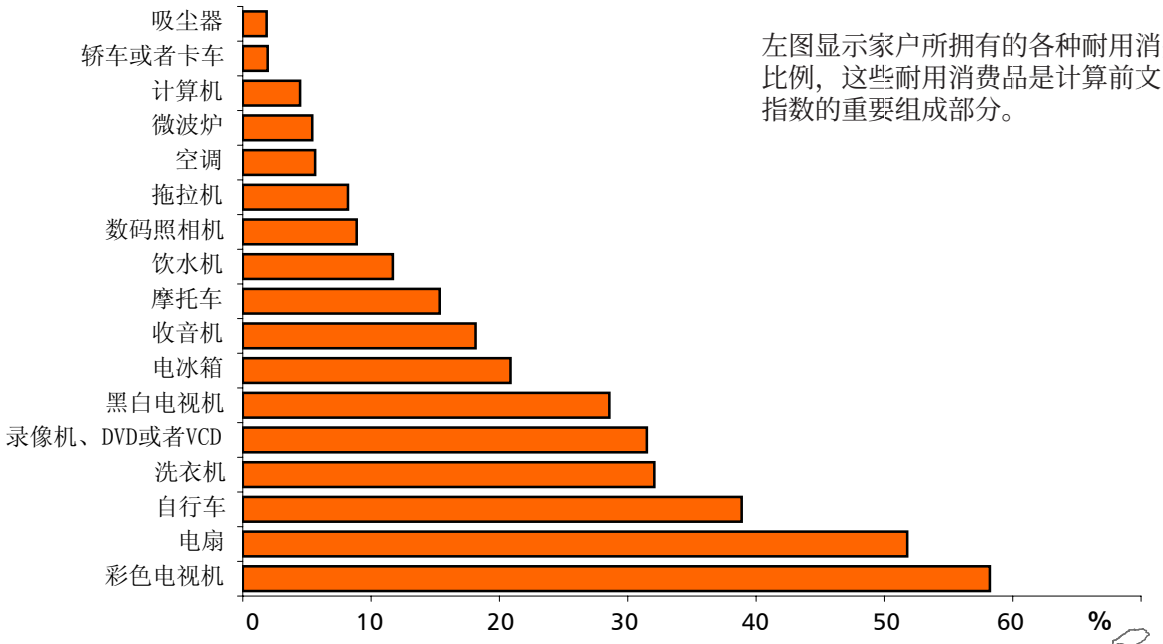
资产指数：详细分布图



下图更详尽地显示了资产指数在中国西部地区的分布。该图使用了概述中所介绍的估算原理和绘制方法。

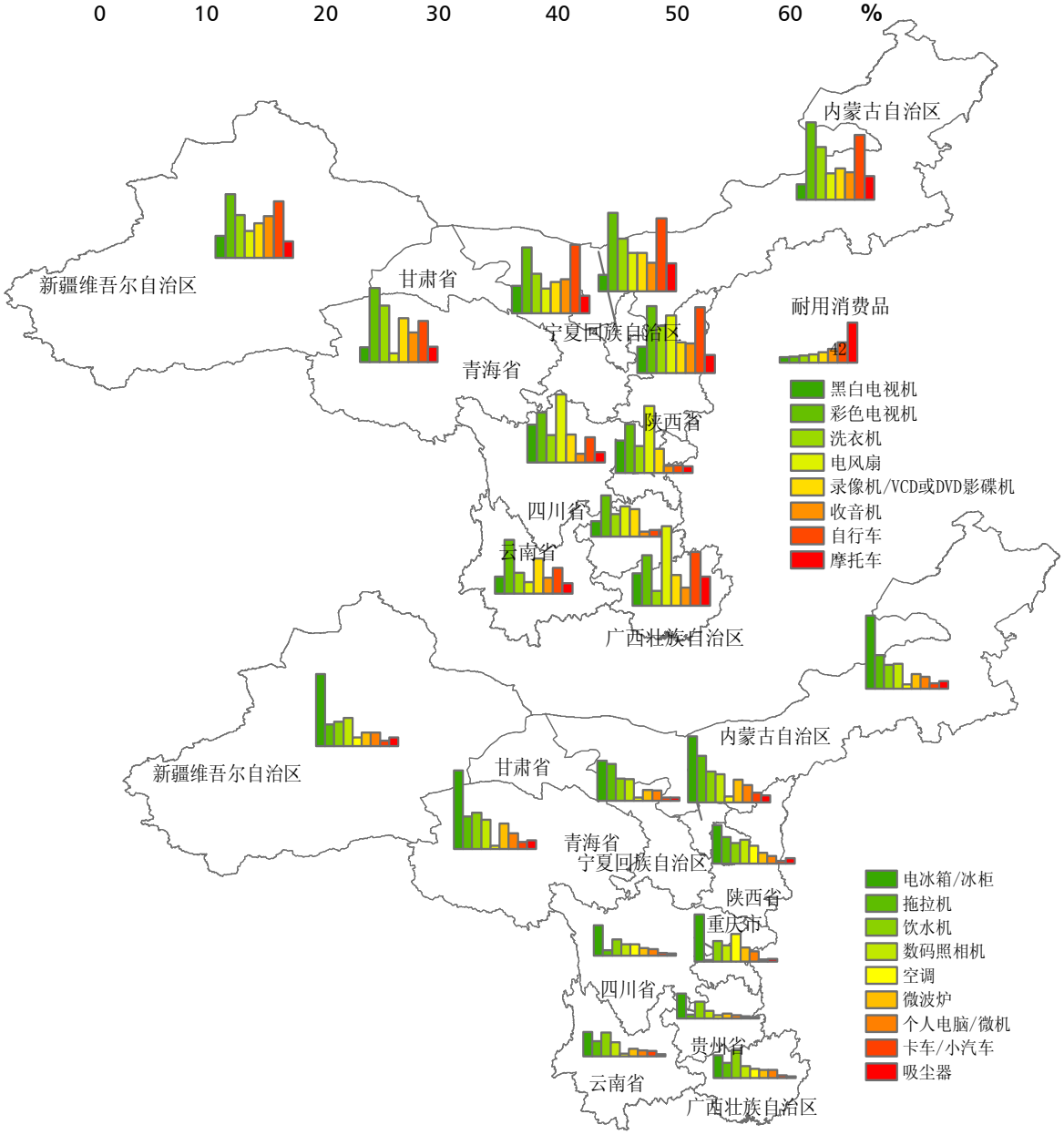
西北省市的资产指数明显高于西南省市。资产指数的城乡差距也很大：城市的资产指数平均为1.300，而农村的资产指数平均为 -0.424。

家户的耐用消费品



左图显示家户所拥有的各种耐用消费品的比例，这些耐用消费品是计算前文的资产指数的重要组成部分。

耐用消费品

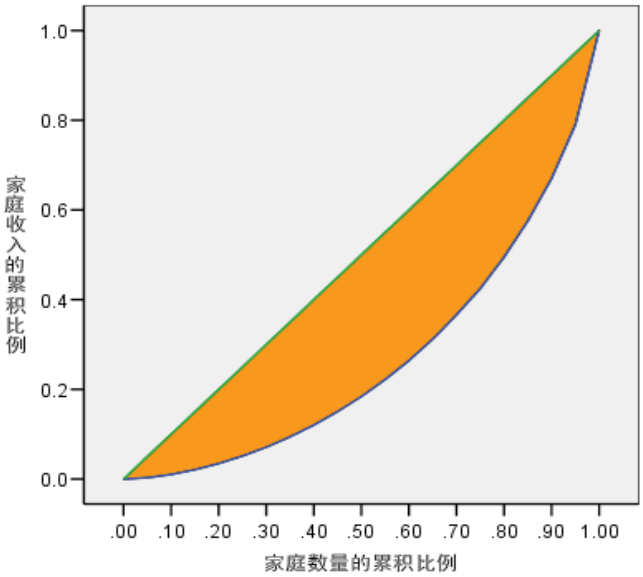


收入分配不均

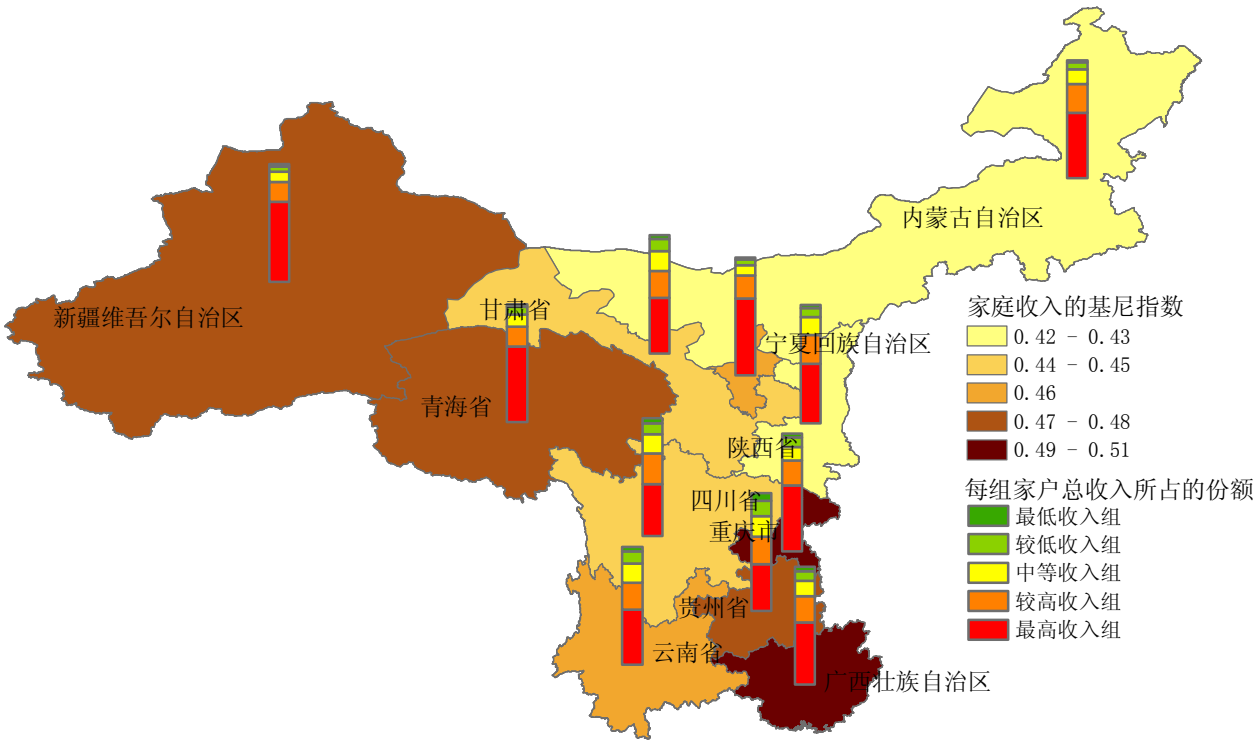
中国西部地区的收入分配不均比较严重。基尼系数是研究收入分配不均的一个重要的指数，如果基尼系数为1，代表一个家户拥有国家所有的收入；如果基尼系数为0，则代表国家的所有收入均匀地分配给所有家户。调查显示，中国西部地区的基尼系数为0.47。这个基尼系数是通过加总每一个家户的年收入，计算累积家户总收入在所有家户中的分配情况计算所得的。

右图显示的是罗伦兹曲线，代表调查前的一年期间内，家户数的累积比相对于这些家庭总收入的累积比。罗伦兹曲线和平均分配线之间的面积越大，说明收入的分配越不公。

下图显示的是各个省（市、自治区）的基尼系数，以及在每个省（市、自治区）将所有家户按照收入分成相同数目的五组，每组家户的收入之和在全部总收入中所占的份额。



收入的基尼系数和分配

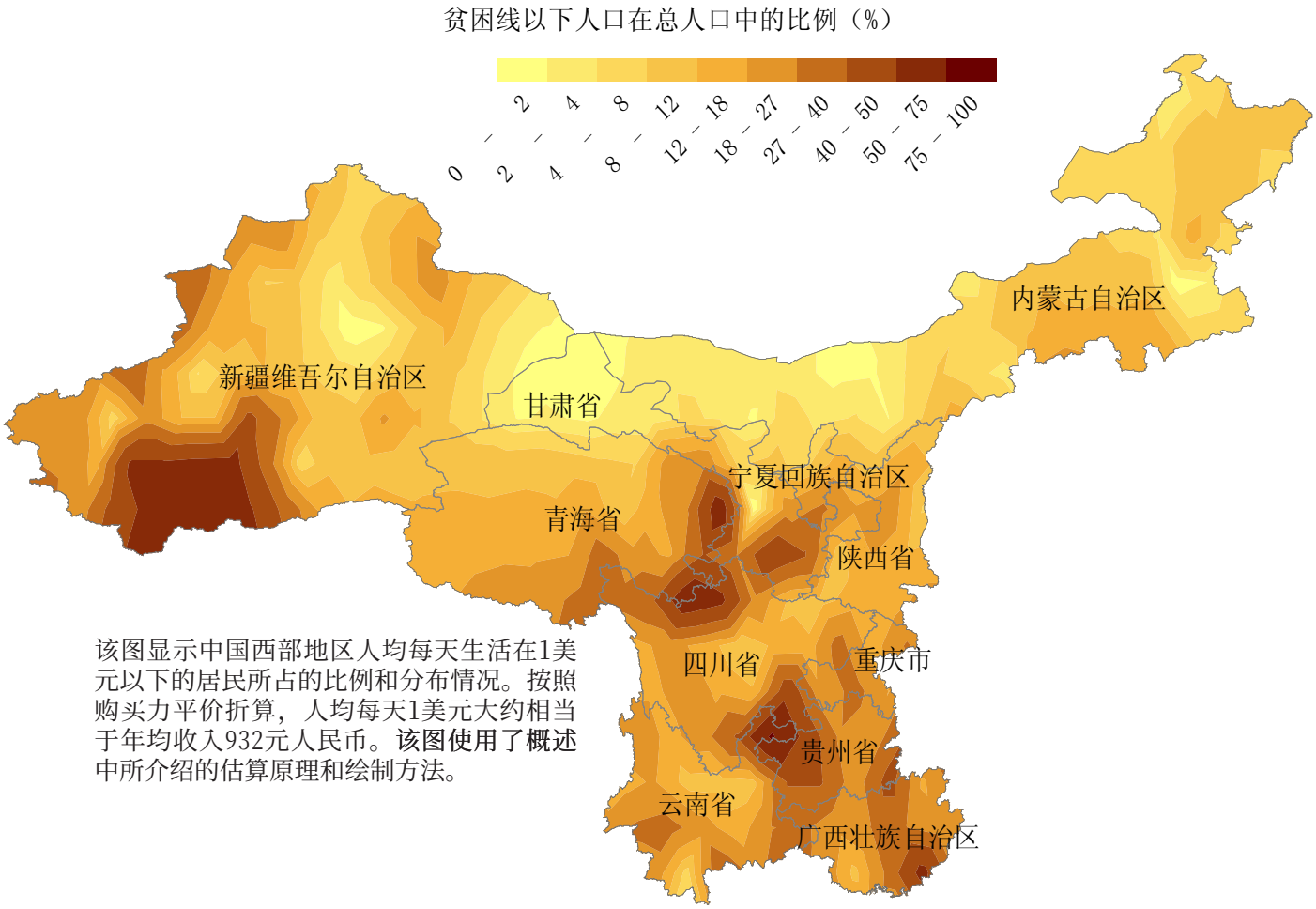


联合国千年发展目标

目标一：消除极端贫穷和饥饿

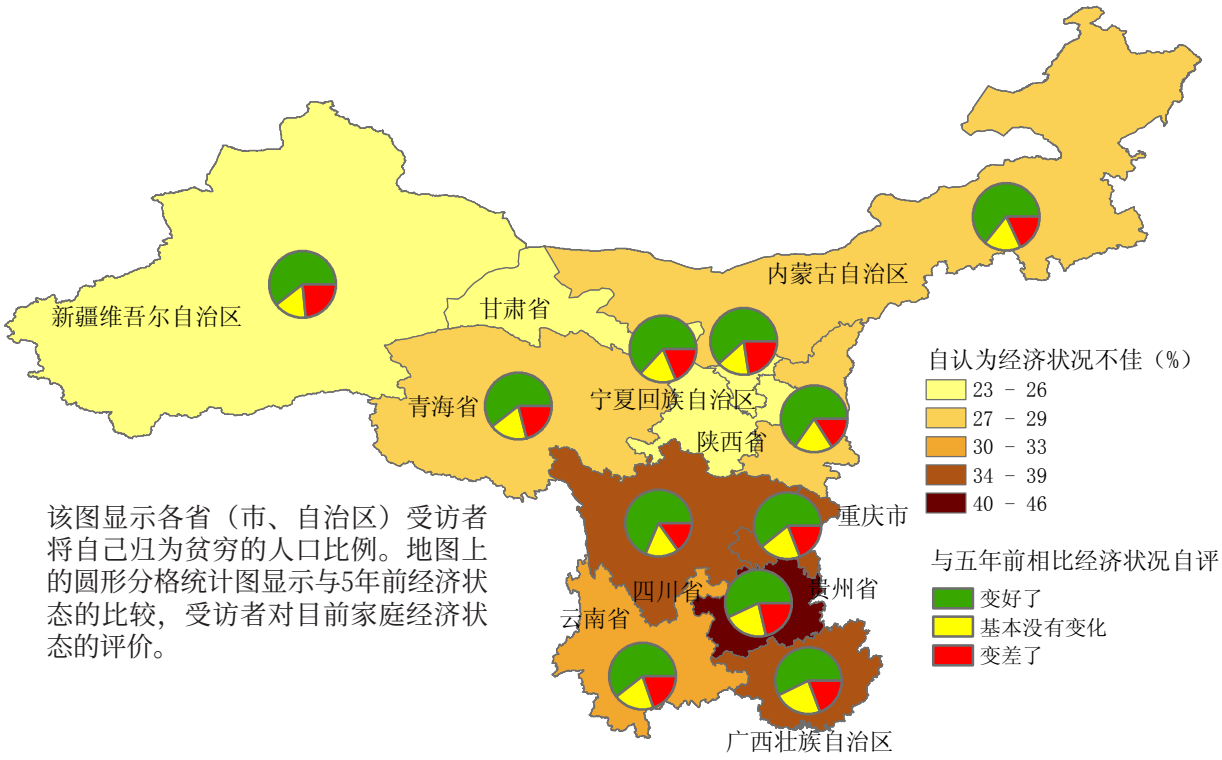
具体目标1：在1990-2015年期间将每日收入低于1美元的人口比例减半

指标1：每日收入低于1美元（按照购买力平价折算）的人口比例		指标2：一天1美元以下贫穷人口收入差距比		指标3：最贫穷的五分之一人口在国民收入或者消费中所占的份额	
总共：	17.8%	总共：	7.0%		
城市：	5.9%	城市：	2.9%	份额：	4.3%
农村：	21.8%	农村：	8.4%		



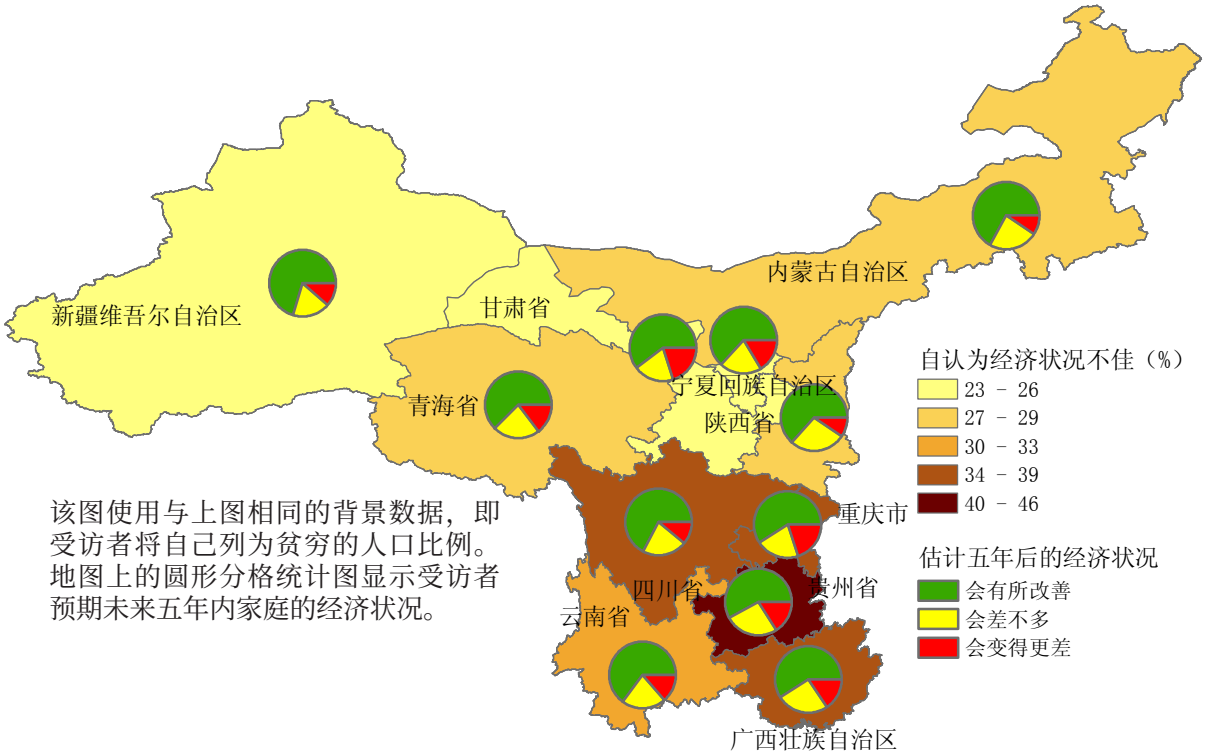
该图显示中国西部地区人均每天生活在1美元以下的居民所占的比例和分布情况。按照购买力平价折算，人均每天1美元大约相当于年均收入932元人民币。**该图使用了概述**中所介绍的估算原理和绘制方法。

经济状态自评



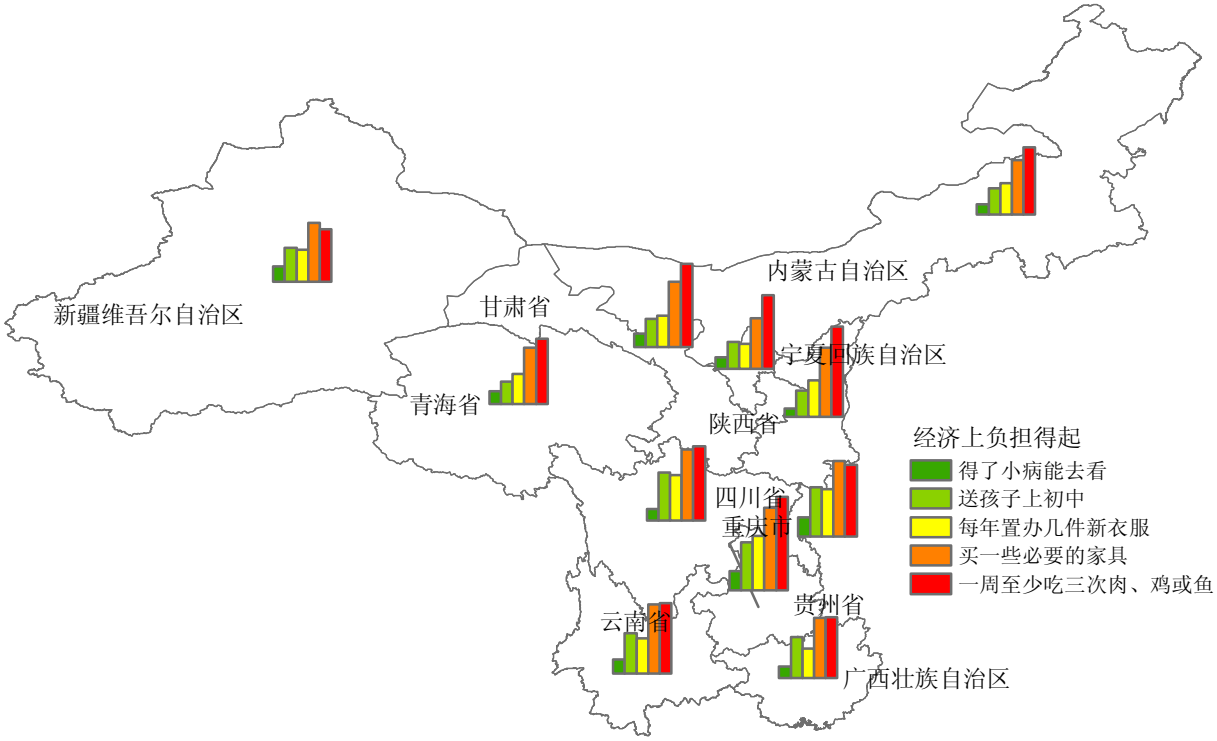
该图显示各省（市、自治区）受访者将自己归为贫穷的人口比例。地图上的圆形分格统计图显示与5年前经济状态的比较，受访者对目前家庭经济状态的评价。

经济状态展望

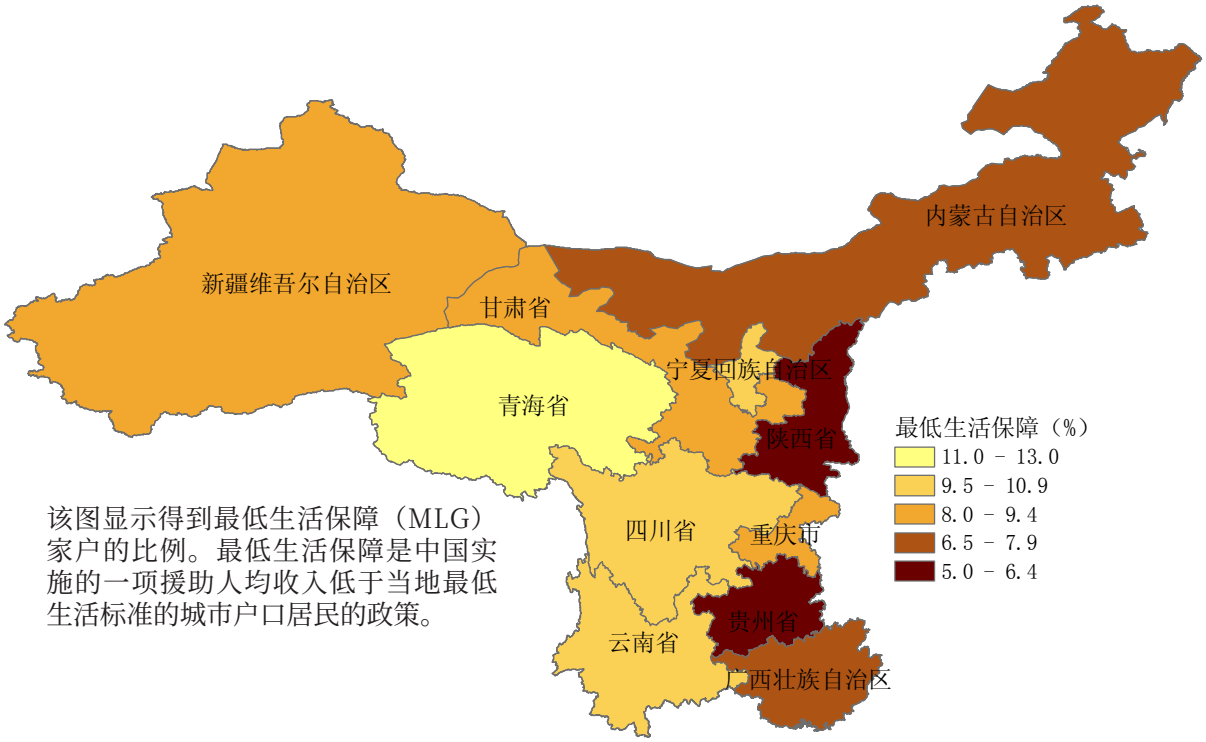


该图使用与上图相同的背景数据，即受访者将自己列为贫穷的人口比例。地图上的圆形分格统计图显示受访者预期未来五年内家庭的经济状况。

负担基本经济支出的能力

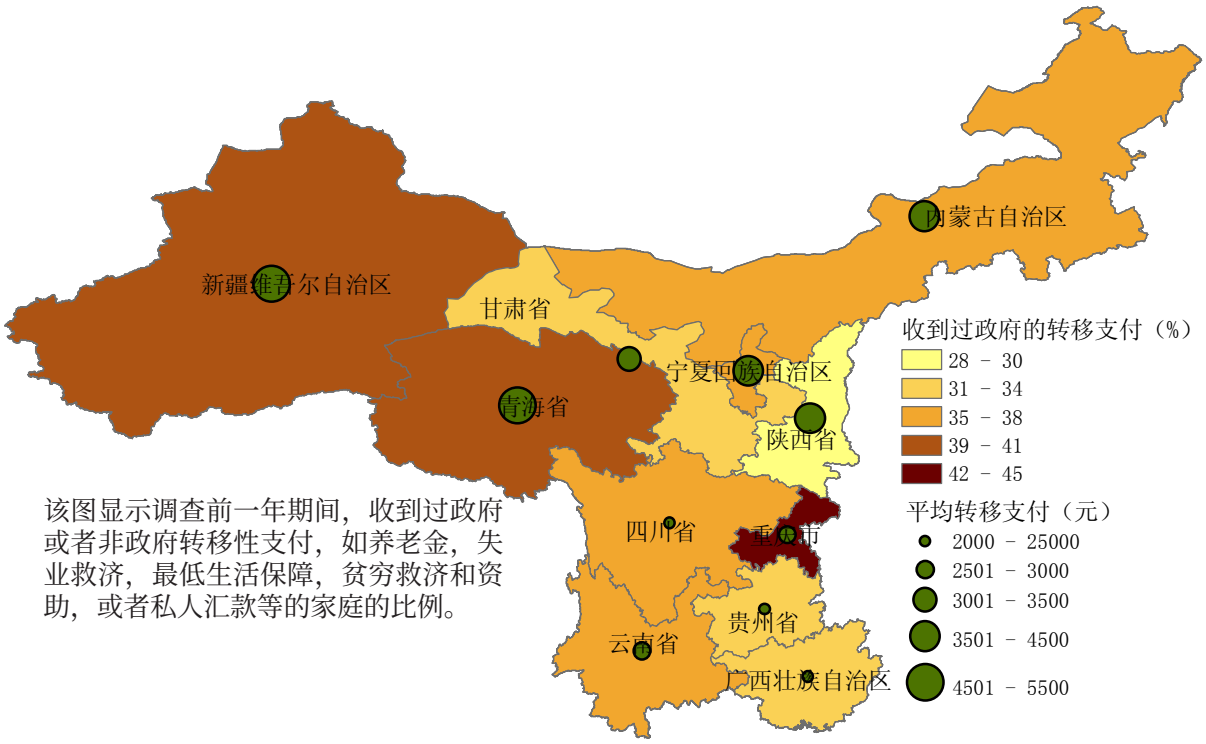


最低生活保障（MLG）



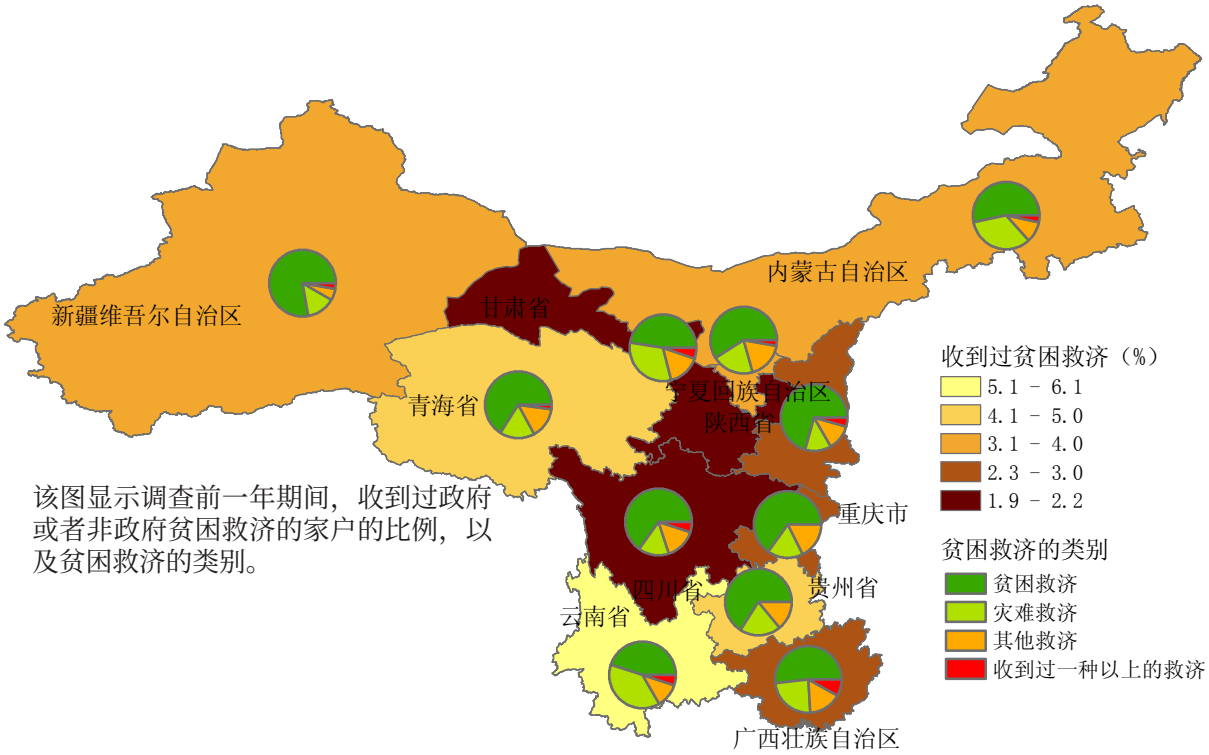
该图显示得到最低生活保障（MLG）家户的比例。最低生活保障是中国实施的一项援助人均收入低于当地最低生活标准的城市户口居民的政策。

政府转移支付



该图显示调查前一年期间，收到过政府或者非政府转移性支付，如养老金，失业救济，最低生活保障，贫穷救济和资助，或者私人汇款等的家庭的比例。

贫困救济



该图显示调查前一年期间，收到过政府或者非政府贫困救济的家户的比例，以及贫困救济的类别。

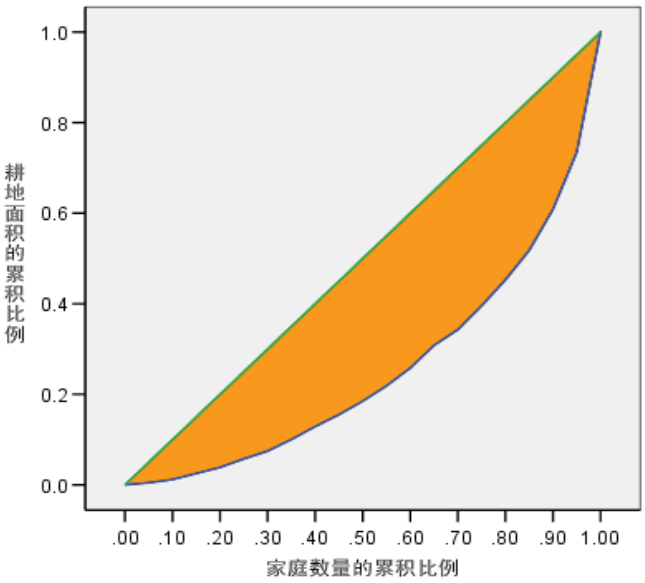
农业

土地分配

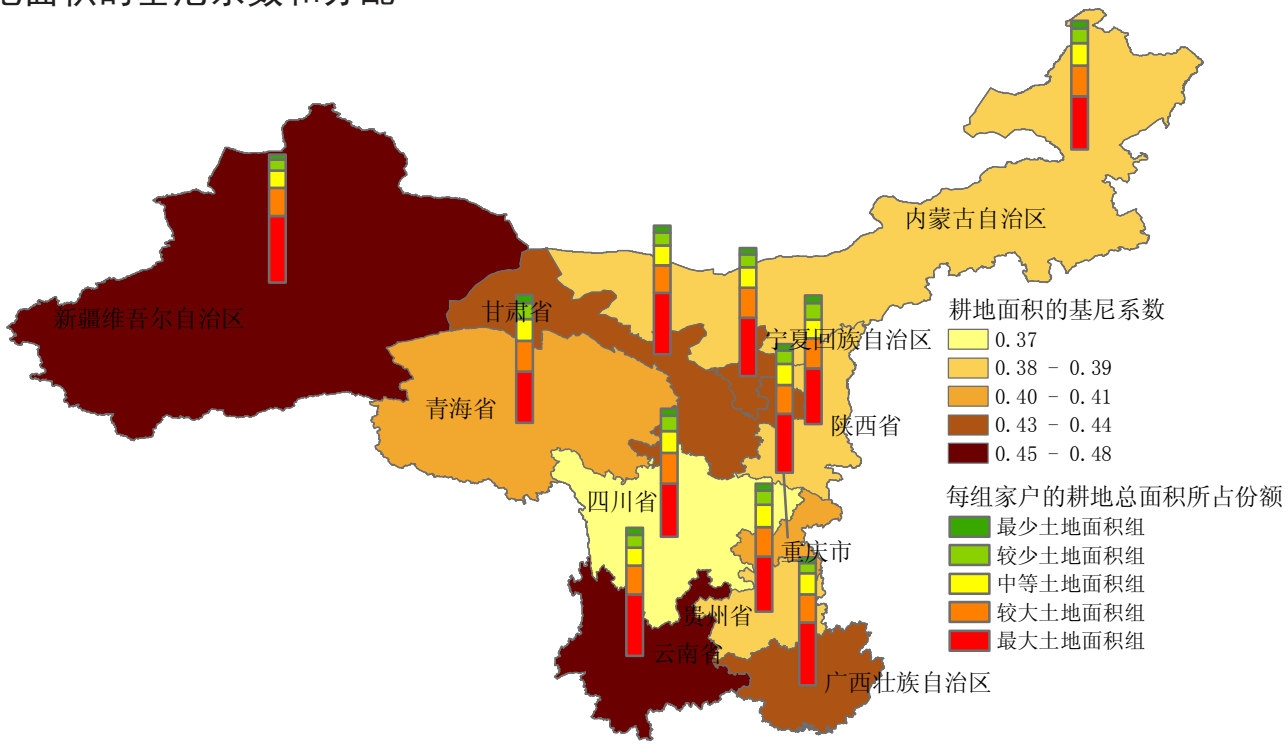
中国西部地区的土地分配是不平均的。西部地区土地分配的基尼系数是0.495，这个系数是基于农户拥有的总的耕地面积在所有拥有耕地的农户中的分配情况计算所得的。农户所拥有总的的耕地面积包括了农户租赁的耕地面积。

右图所示的罗伦兹曲线，代表调查前的一年期间内，农户数的累积比相对于这些家庭所拥有耕地的总面积的累积比。罗伦兹曲线和平均分配线之间的面积越大，说明土地的分配越不公。

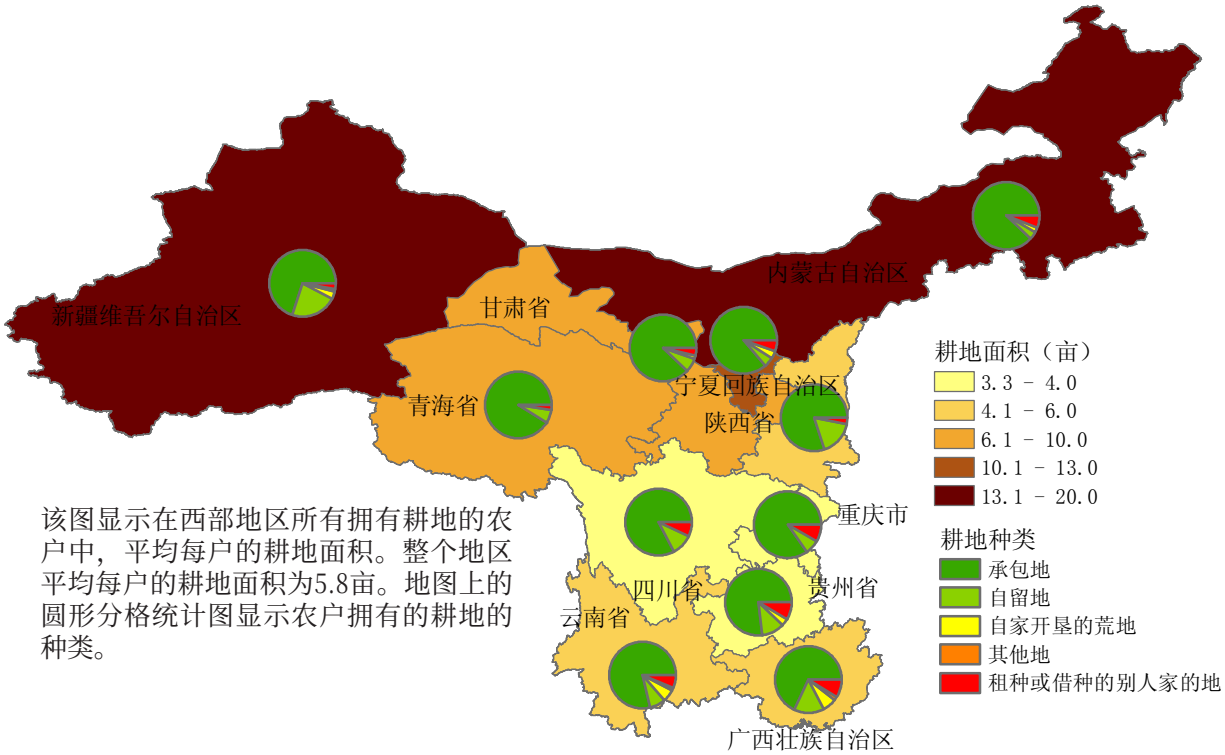
下图表示各个省（市、自治区）耕地面积的基尼系数，以及在每个省（市、自治区）将所有农户按照拥有的耕地面积分成相同数目的五组，每组农户的耕地面积之和在全部总耕地面积中所占的份额。最低土地面积组农户的总耕地面积仅为西部地区总耕地面积的5.8%，而最高拥有土地面积组拥有总共40%的耕



耕地面积的基尼系数和分配

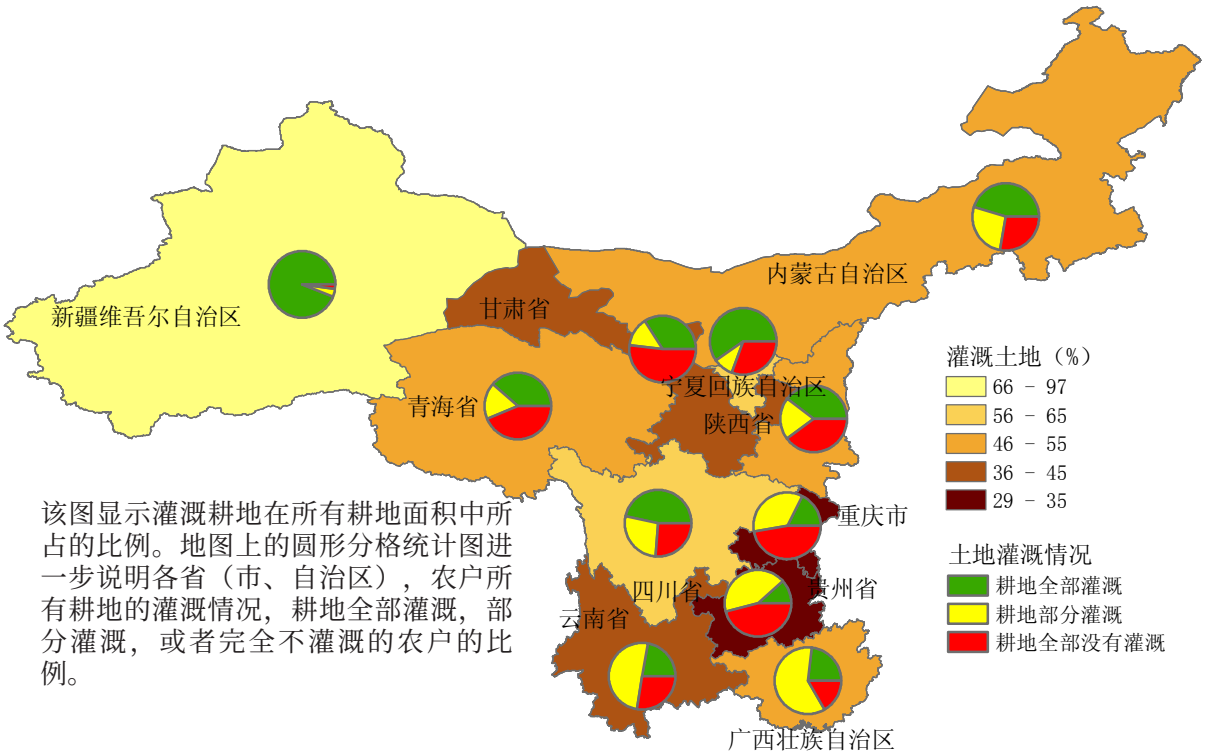


耕地面积和类别



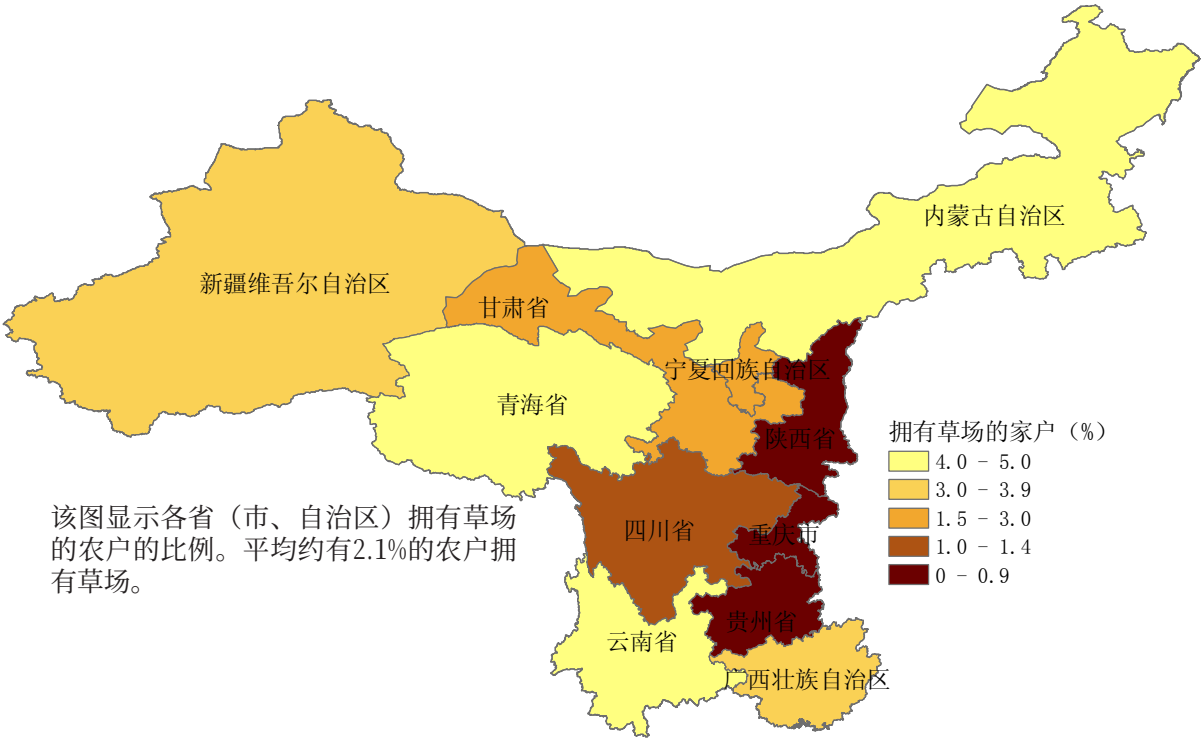
该图显示在西部地区所有拥有耕地的农户中，平均每户的耕地面积。整个地区平均每户的耕地面积为5.8亩。地图上的圆形分格统计图显示农户拥有的耕地的种类。

土地灌溉



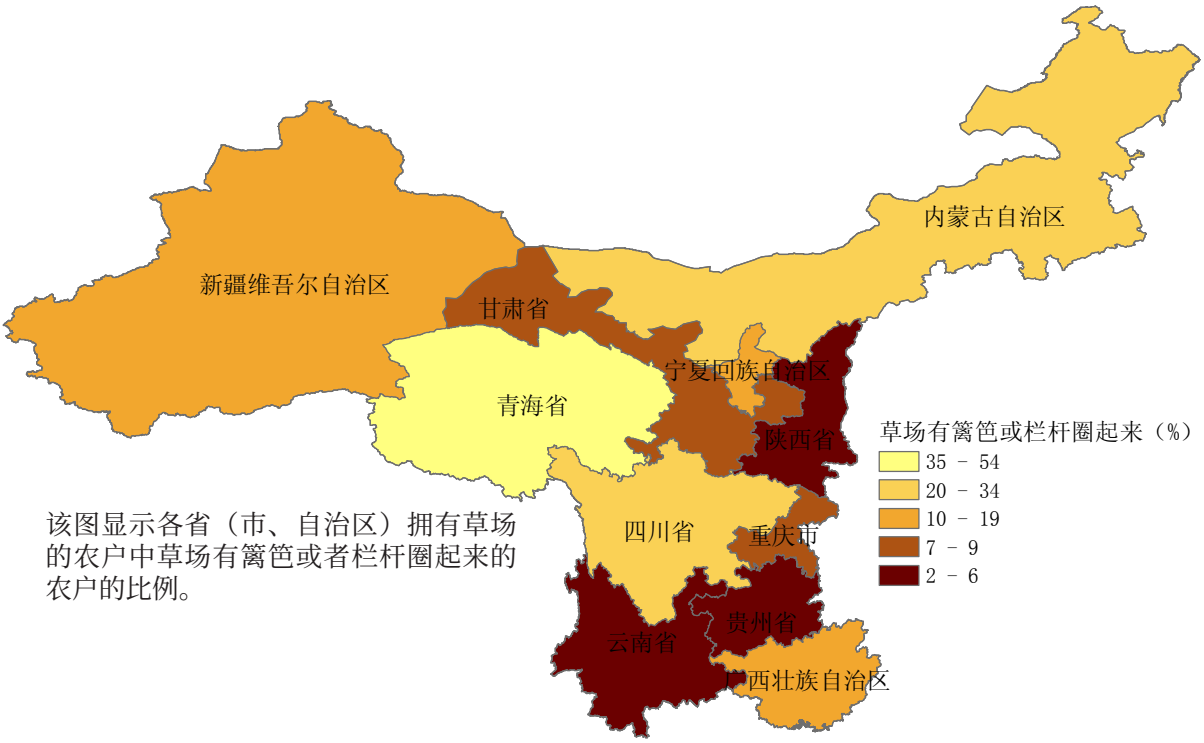
该图显示灌溉耕地在所有耕地面积中所占的比例。地图上的圆形分格统计图进一步说明各省（市、自治区），农户所有耕地的灌溉情况，耕地全部灌溉，部分灌溉，或者完全不灌溉的农户的比例。

草场



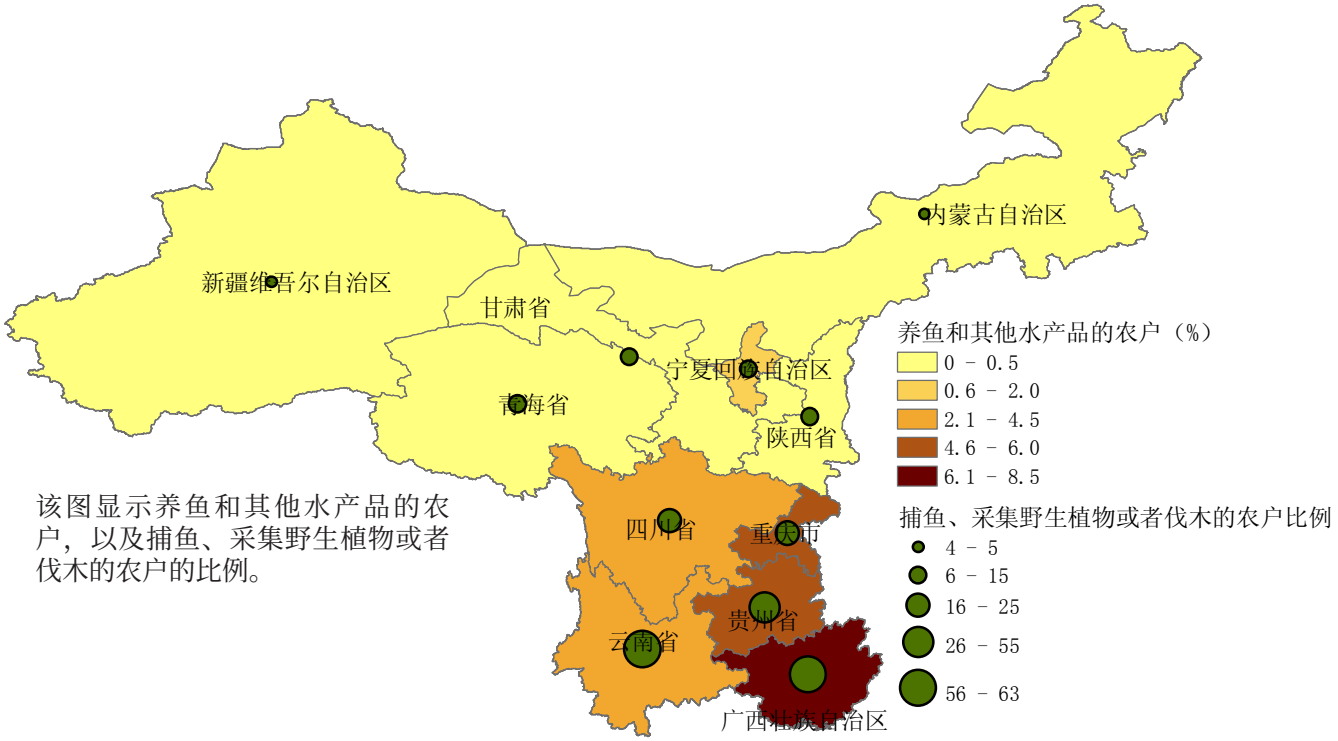
该图显示各省（市、自治区）拥有草场的农户的比例。平均约有2.1%的农户拥有草场。

围栏草场



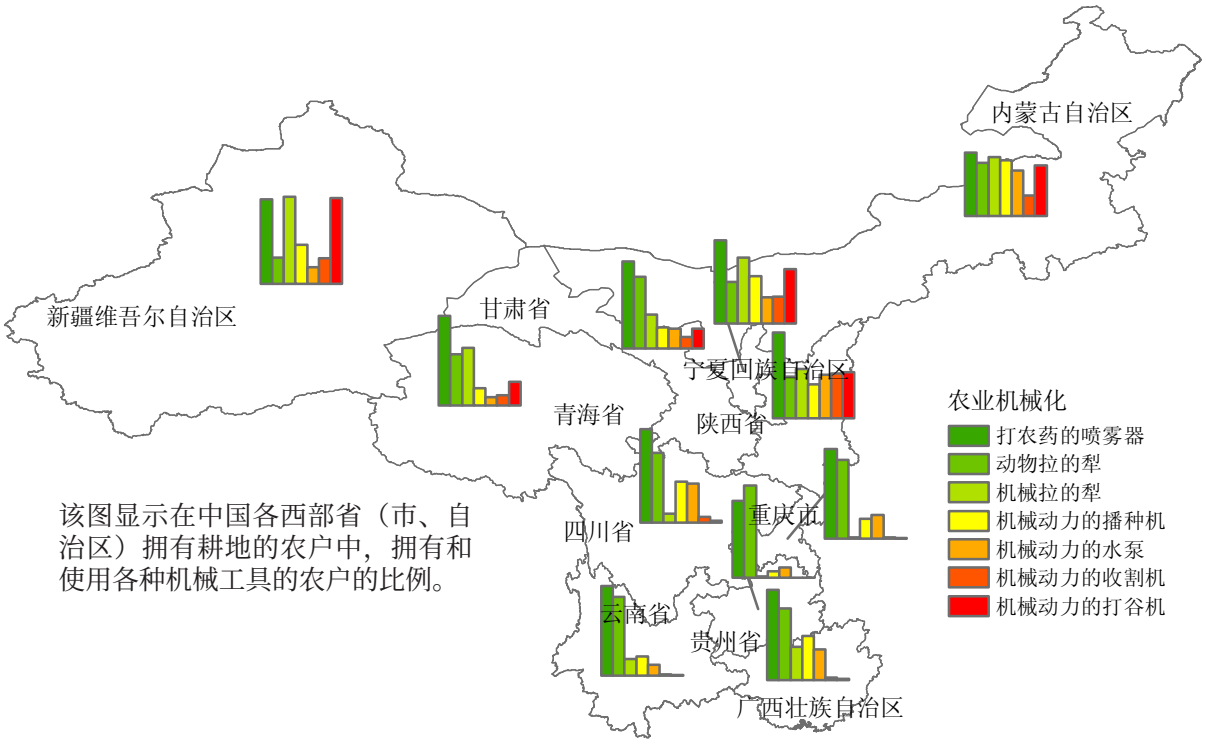
该图显示各省（市、自治区）拥有草场的农户中草场有篱笆或者栏杆圈起来的农户的比例。

农户从事的其他农业活动



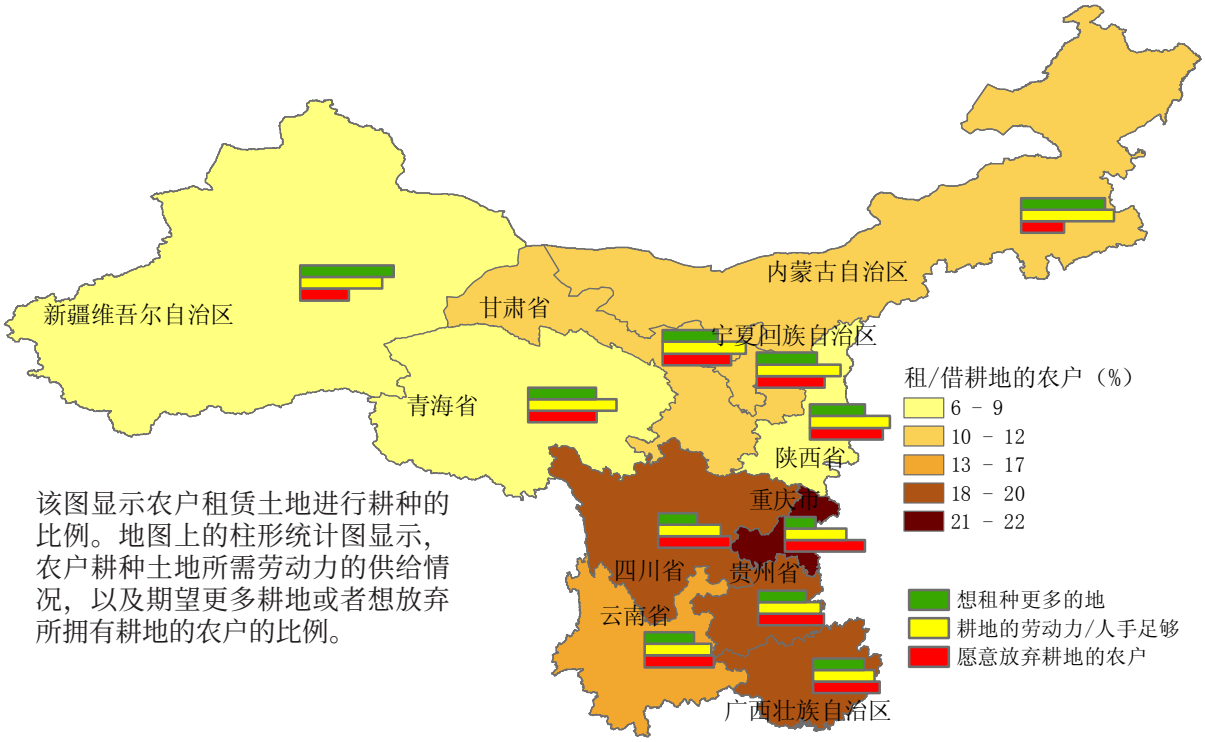
该图显示养鱼和其他水产品的农户，以及捕鱼、采集野生植物或者伐木的农户的比例。

农业机械化



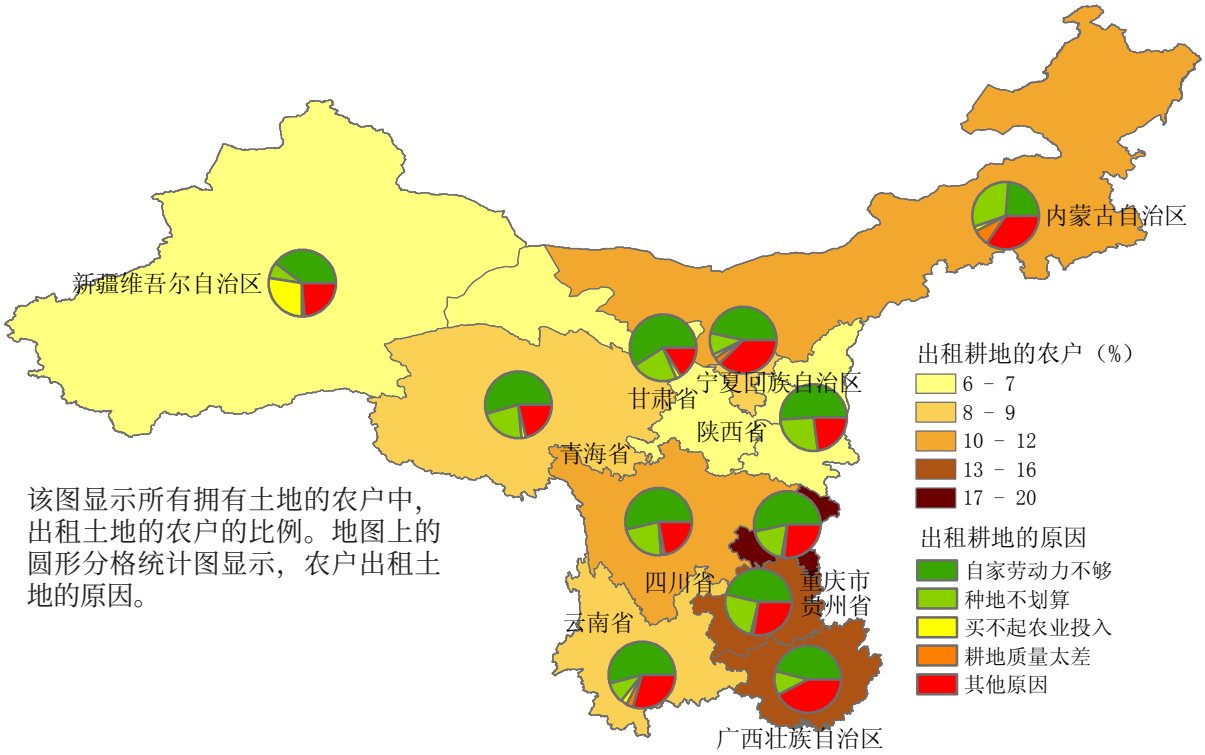
该图显示在中国各西部省（市、自治区）拥有耕地的农户中，拥有和使用各种机械工具的农户的比例。

租赁土地



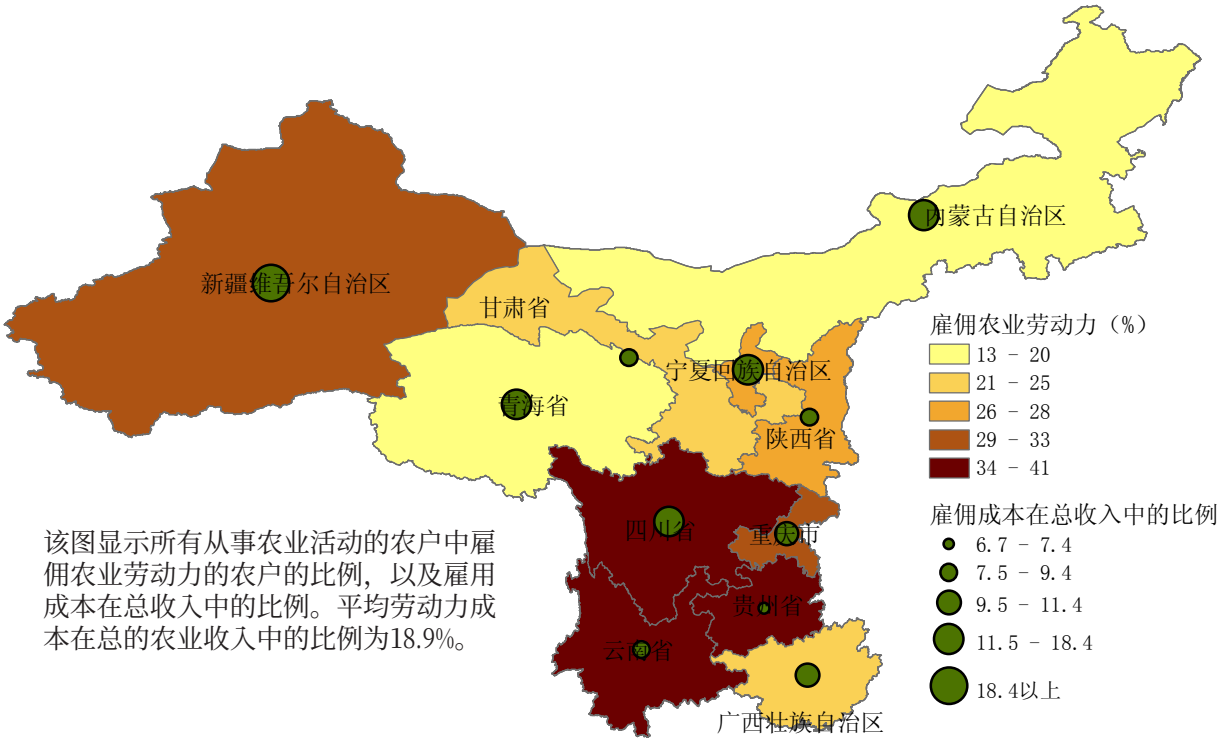
该图显示农户租赁土地进行耕种的比例。地图上的柱形统计图显示，农户耕种土地所需劳动力的供给情况，以及期望更多耕地或者想放弃所拥有耕地的农户的比例。

出租土地



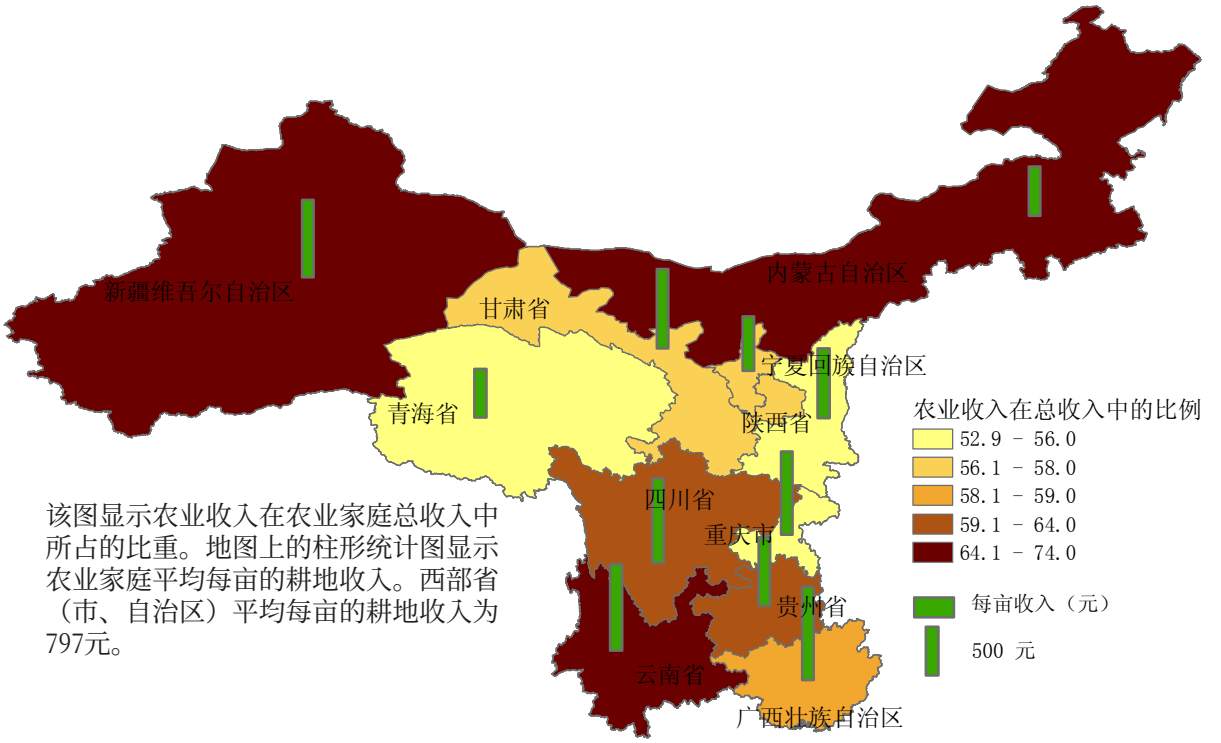
该图显示所有拥有土地的农户中，出租土地的农户的比例。地图上的圆形分格统计图显示，农户出租土地的原因。

雇佣农业劳动力和劳动力成本



该图显示所有从事农业活动的农户中雇佣农业劳动力的农户的比例，以及雇用成本在总收入中的比例。平均劳动力成本在总的农业收入中的比例为18.9%。

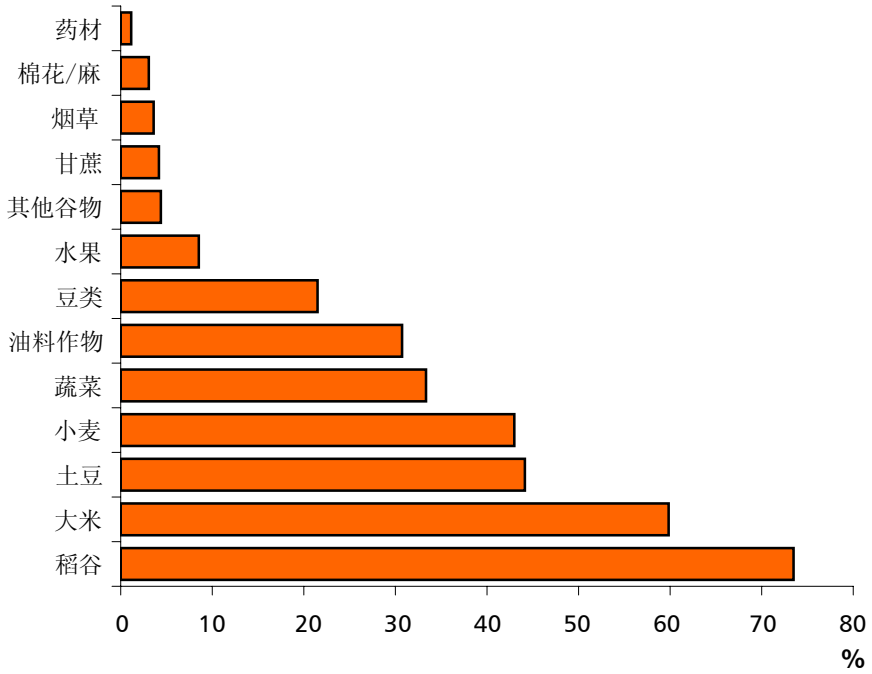
农业收入



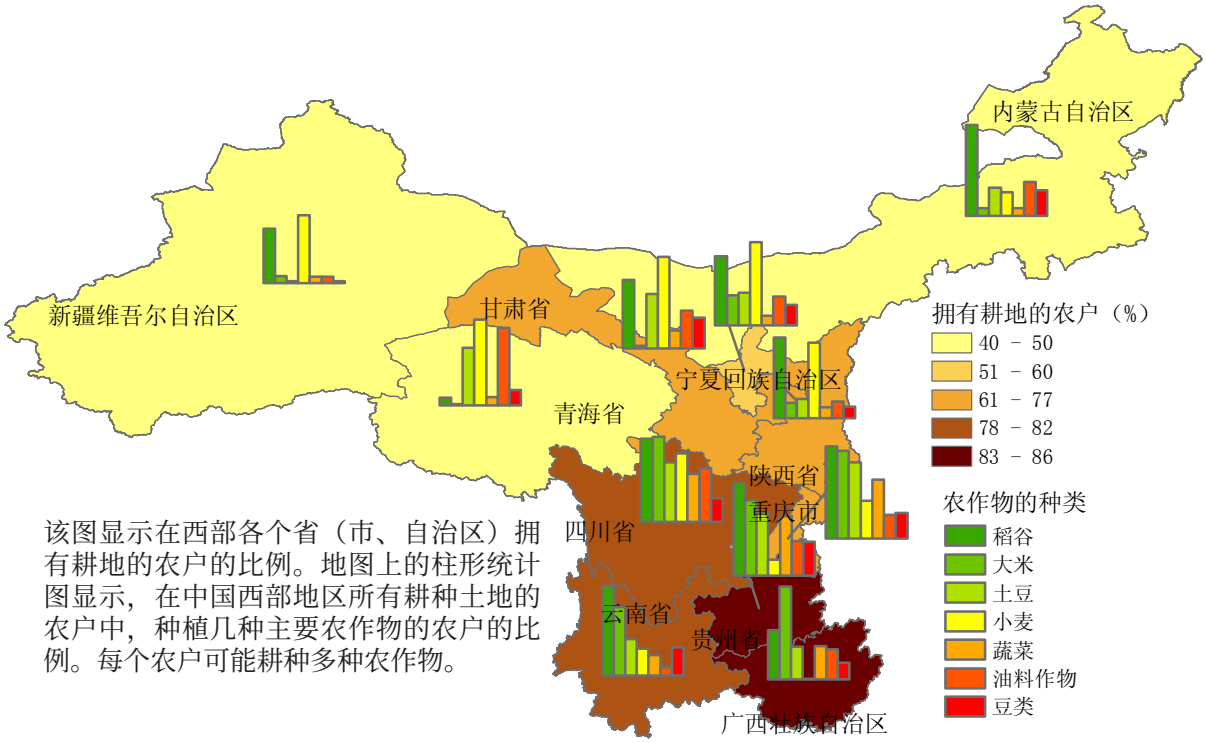
该图显示农业收入在农业家庭总收入中所占的比重。地图上的柱形统计图显示农业家庭平均每亩的耕地收入。西部省（市、自治区）平均每亩的耕地收入为797元。

农作物

该图表显示西部地区所有耕种土地的农户中，种植各种农作物的农户的比例。每个农户可能耕种多种农作物，因而种植各种农作物的农户的比例之和大于100。

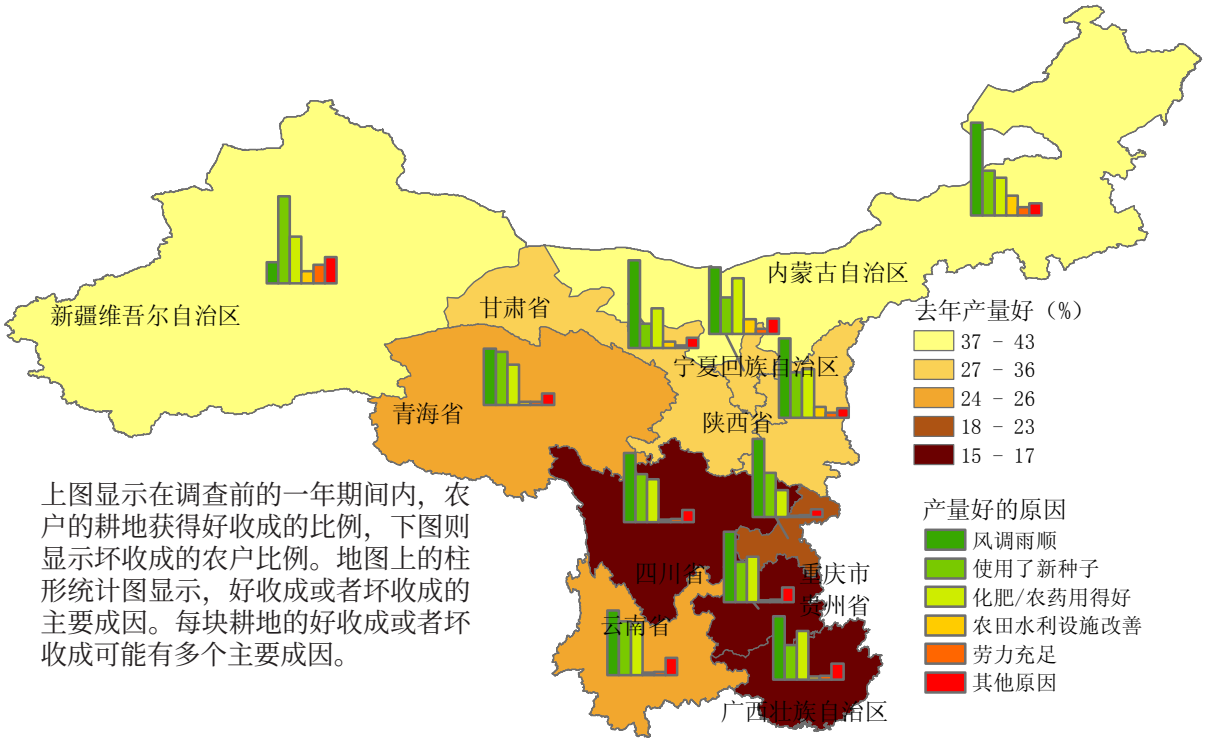


耕地和农作物种类



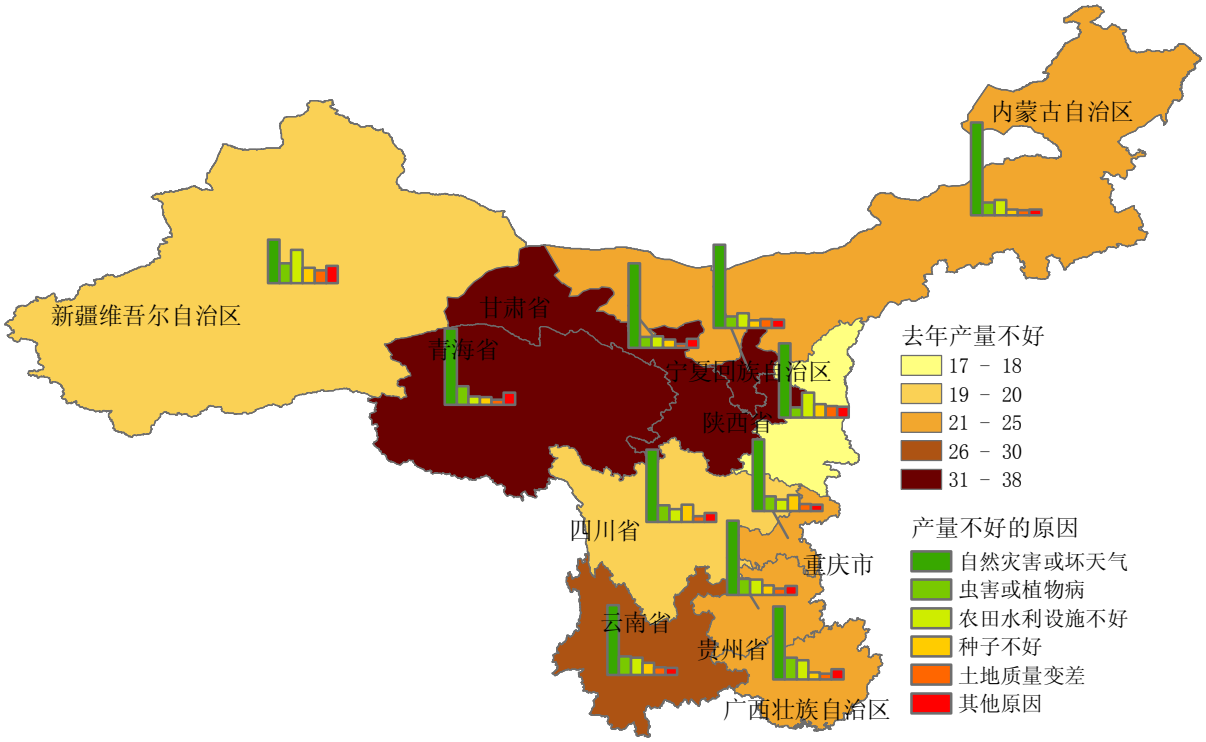
该图显示在西部各个省（市、自治区）拥有耕地的农户的比例。地图上的柱形统计图显示，在中国西部地区所有耕种土地的农户中，种植几种主要农作物的农户的比例。每个农户可能耕种多种农作物。

好收成



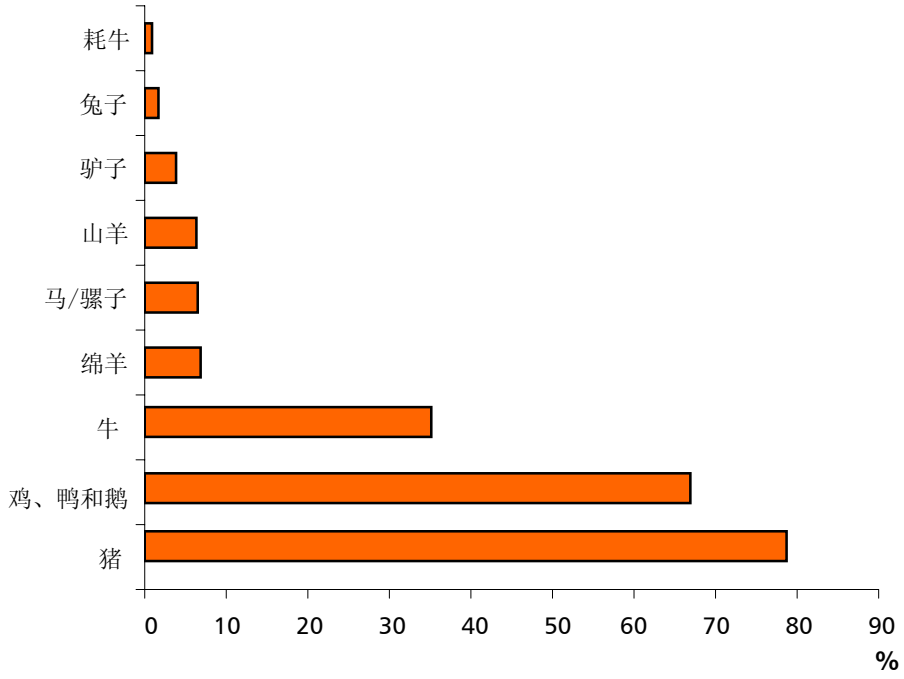
上图显示在调查前的一年期间内，农户的耕地获得好收成的比例，下图则显示坏收成的农户比例。地图上的柱形统计图显示，好收成或者坏收成的主要成因。每块耕地的好收成或者坏收成可能有多个主要成因。

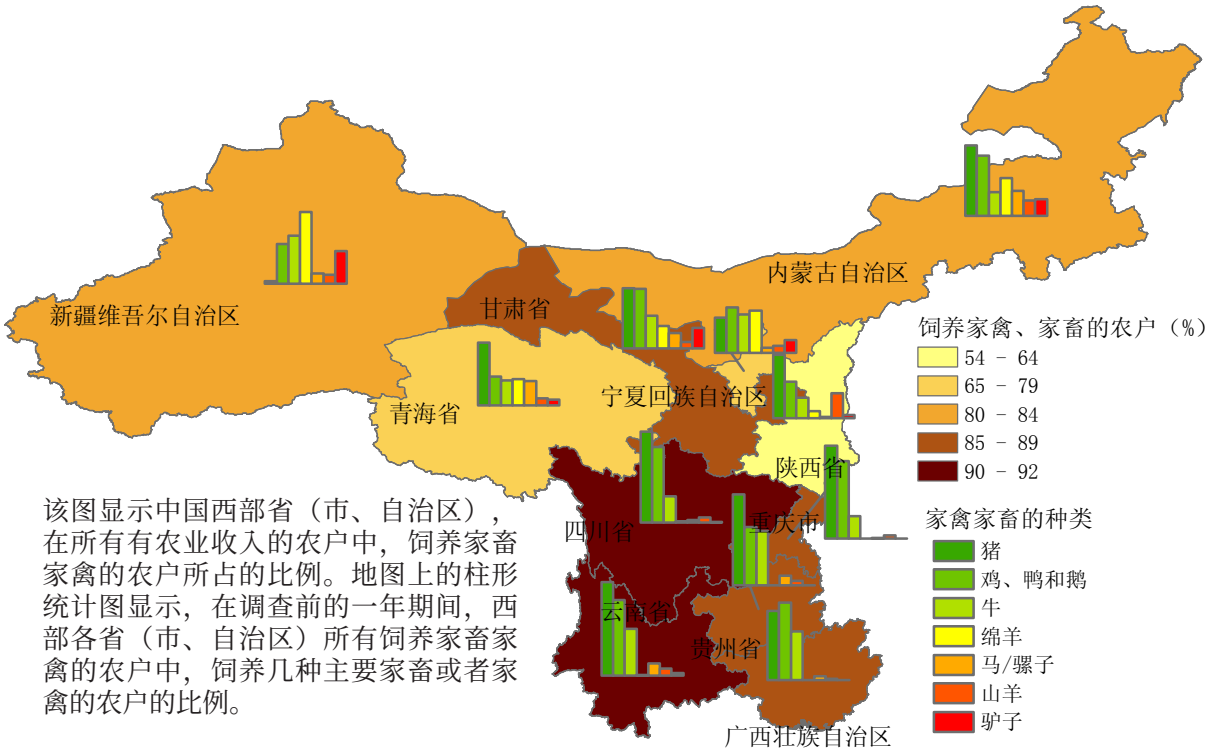
坏收成



家畜和家禽

该图表显示，在调查前的一年期间，西部地区所有饲养家畜家禽的农户中，饲养各种家畜或者家禽的农户的比例。在所有的从事农业活动的农户中，85%的农户饲养家禽家畜。



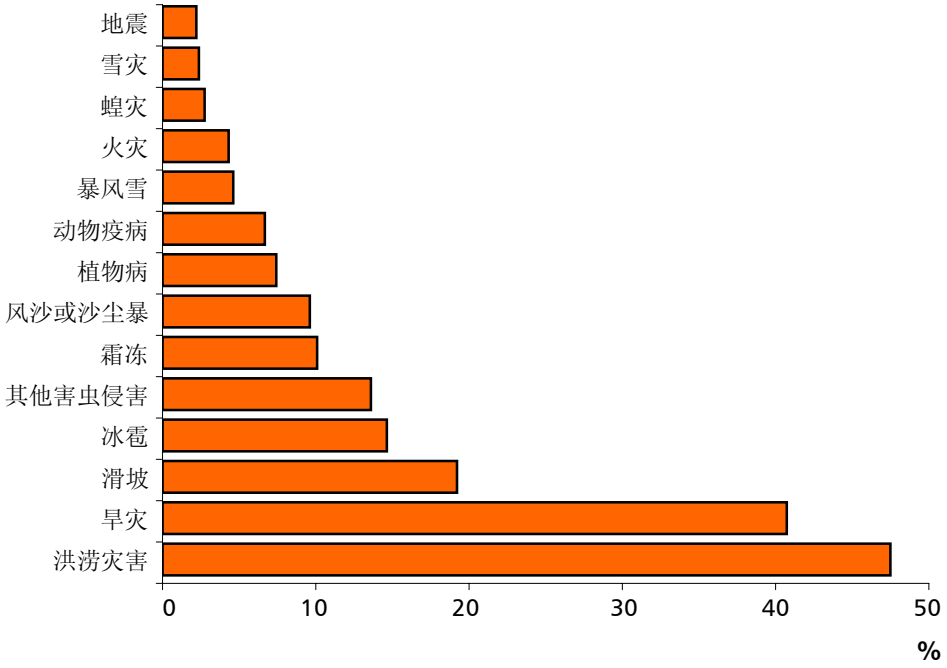


该图显示中国西部省（市、自治区），在所有有农业收入的农户中，饲养家畜家禽的农户所占的比例。地图上的柱形统计图显示，在调查前的一年期间，西部各省（市、自治区）所有饲养家畜家禽的农户中，饲养几种主要家畜或者家禽的农户的比例。

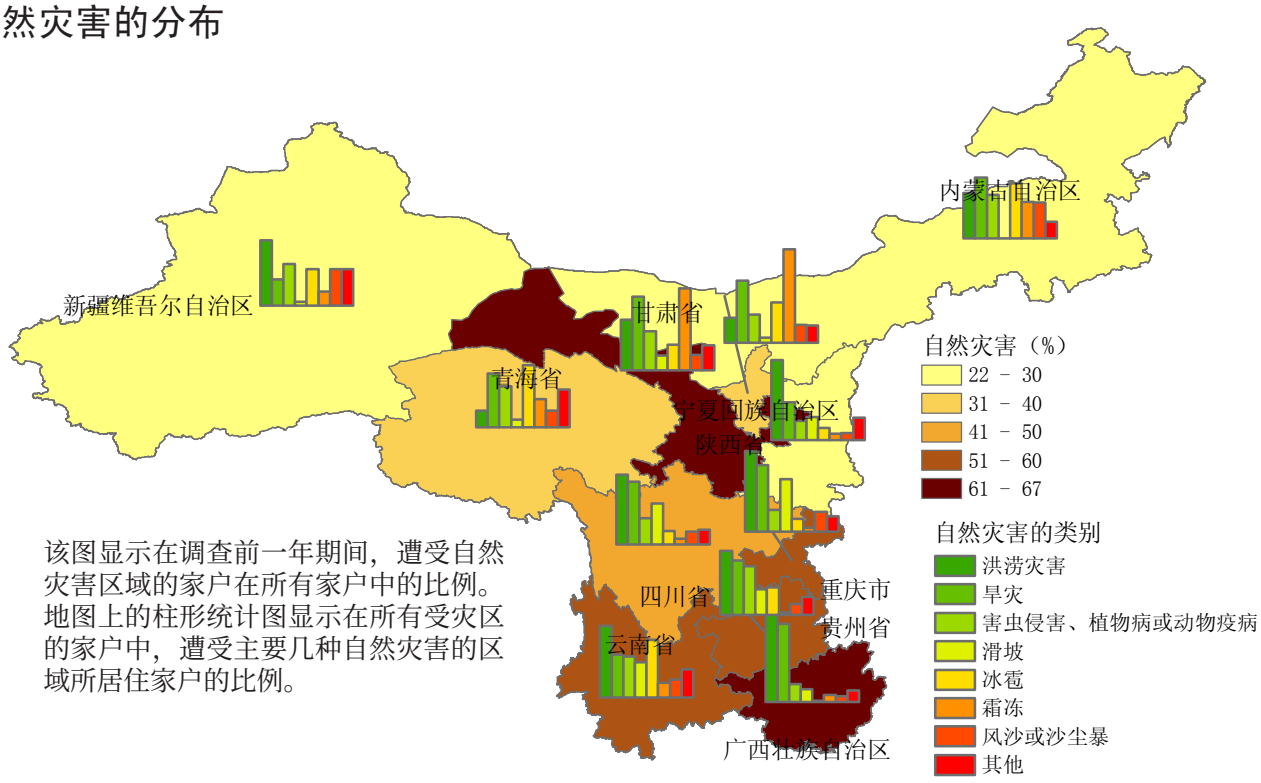
环境

调查前的一年期间，西部地区大约有将近一半的家户（47%）居住在报告有自然灾害的地区。下图表显示，在所有遭受自然灾害区域的家户中，各种类型自然灾害区域所居住的家户的比例。

自然灾害

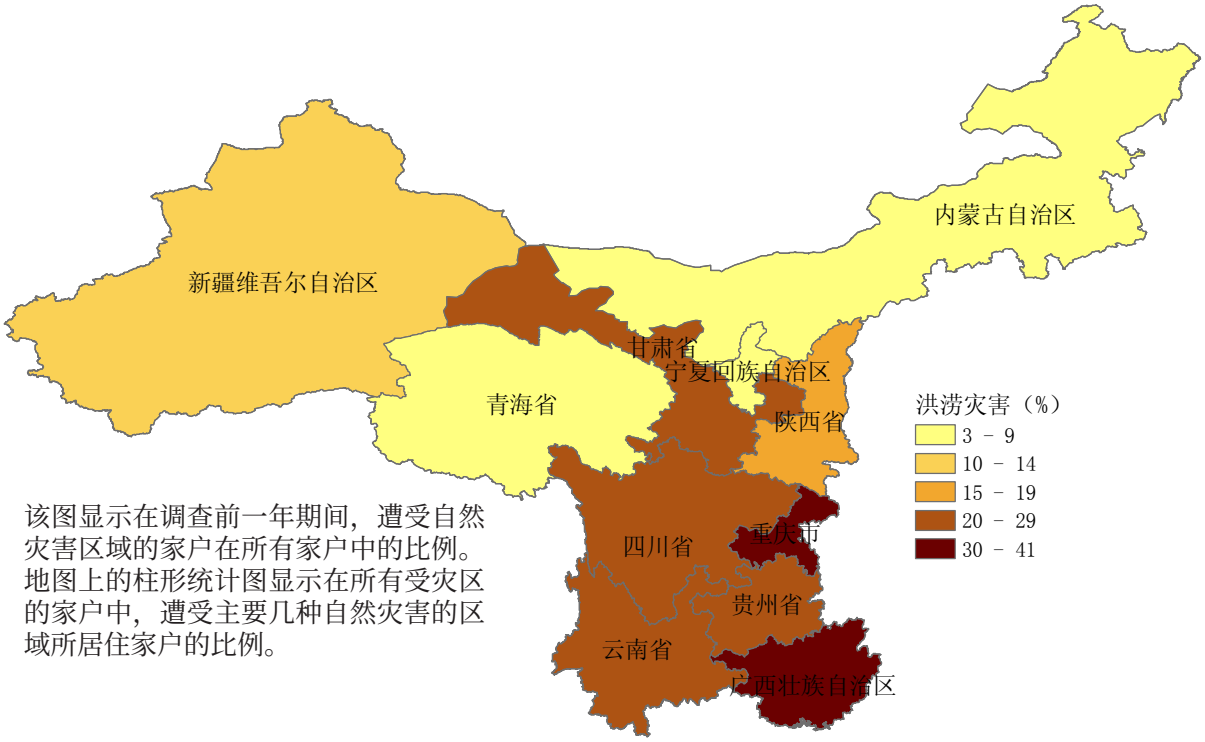


自然灾害的分布



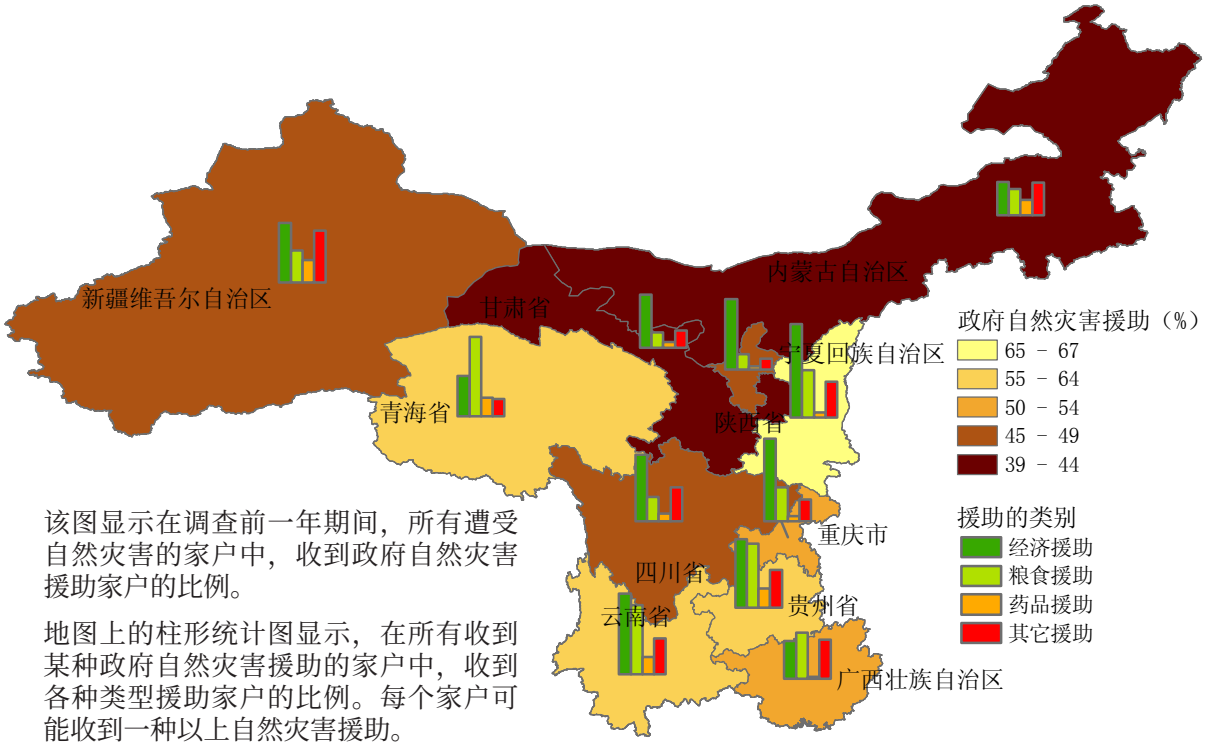
该图显示在调查前一年期间，遭受自然灾害区域的家户在所有家户中的比例。地图上的柱形统计图显示在所有受灾区的家户中，遭受主要几种自然灾害的区域所居住家户的比例。

洪涝灾害



该图显示在调查前一年期间，遭受自然灾害区域的家户在所有家户中的比例。地图上的柱形统计图显示在所有受灾区的家户中，遭受主要几种自然灾害的区域所居住家户的比例。

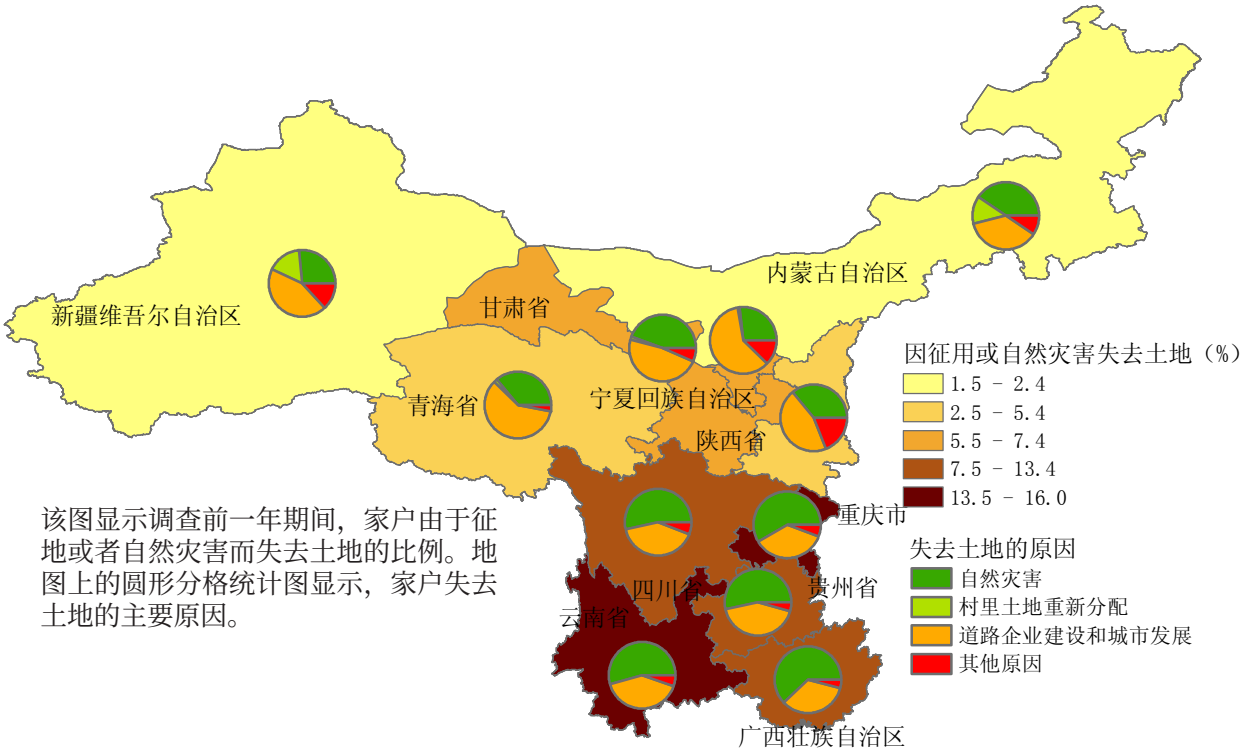
政府自然灾害援助



该图显示在调查前一年期间，所有遭受自然灾害的家户中，收到政府自然灾害援助家户的比例。

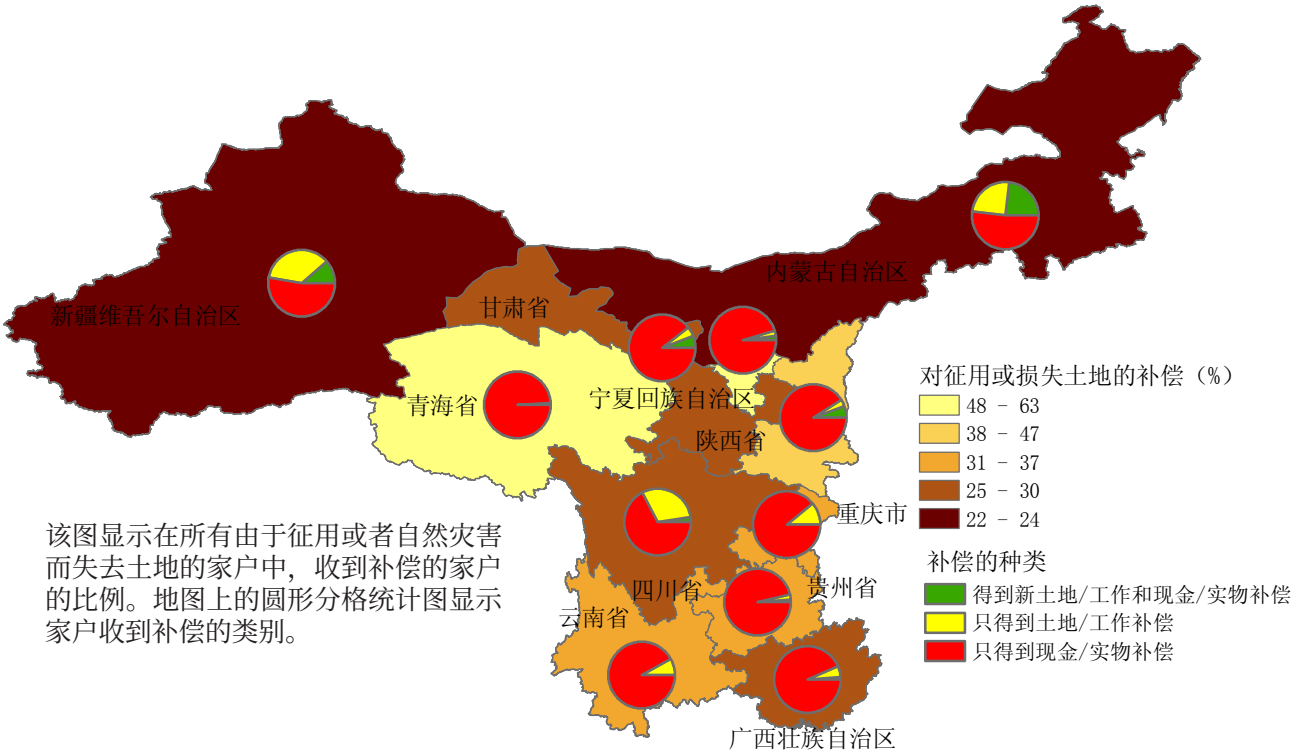
地图上的柱形统计图显示，在所有收到某种政府自然灾害援助的家户中，收到各种类型援助家户的比例。每个家户可能收到一种以上自然灾害援助。

由于征地或者自然灾害而失去土地



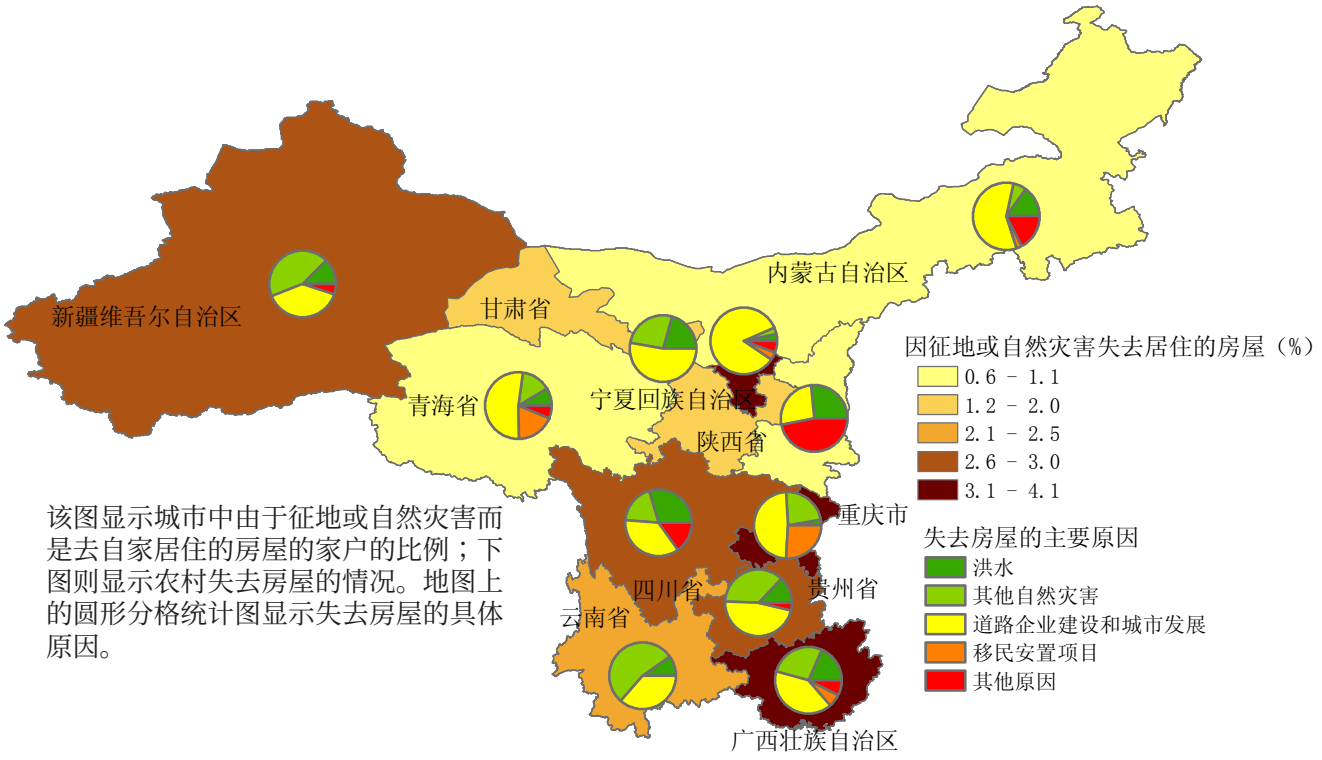
该图显示调查前一年期间，家户由于征地或者自然灾害而失去土地的比例。地图上的圆形分格统计图显示，家户失去土地的主要原因。

失去土地的补偿



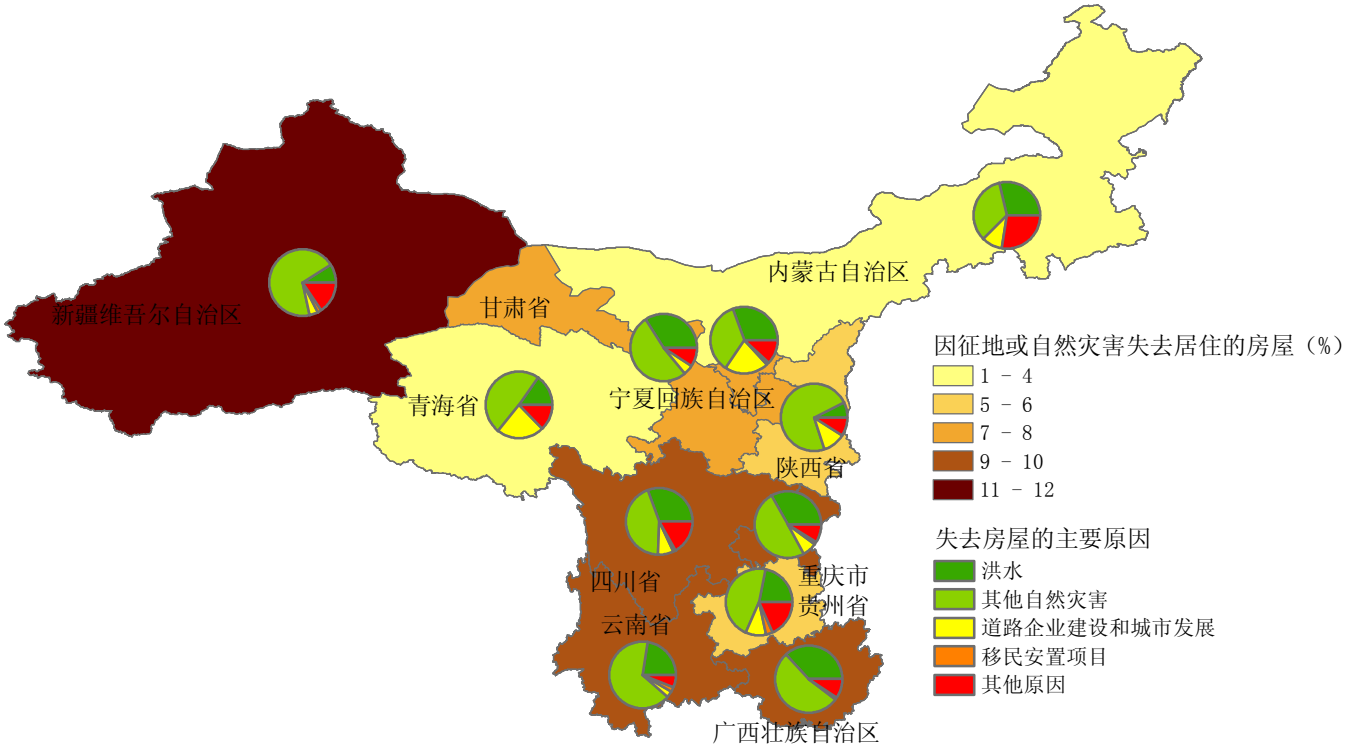
该图显示在所有由于征用或者自然灾害而失去土地的家户中，收到补偿的家户的比例。地图上的圆形分格统计图显示家户收到补偿的类别。

失去房屋（城市）

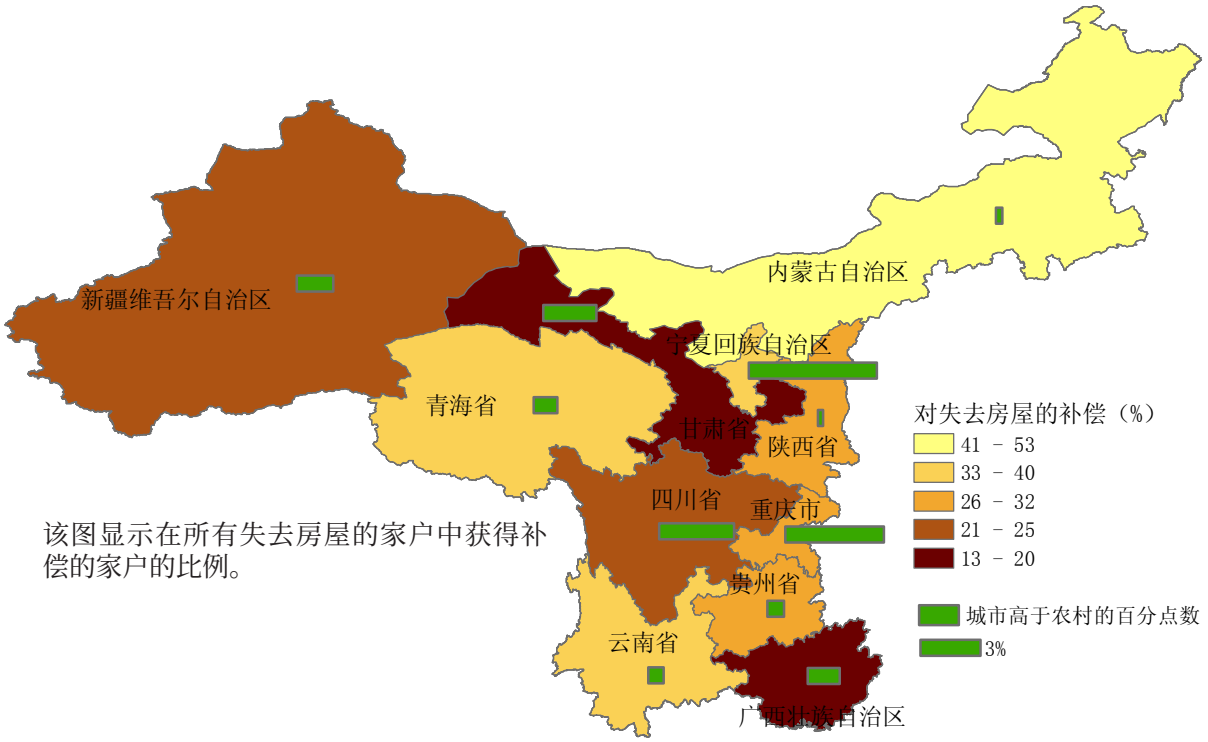


该图显示城市中由于征地或自然灾害而是去自家居住的房屋的家户的比例；下图则显示农村失去房屋的情况。地图上的圆形分格统计图显示失去房屋的具体原因。

失去房屋（农村）

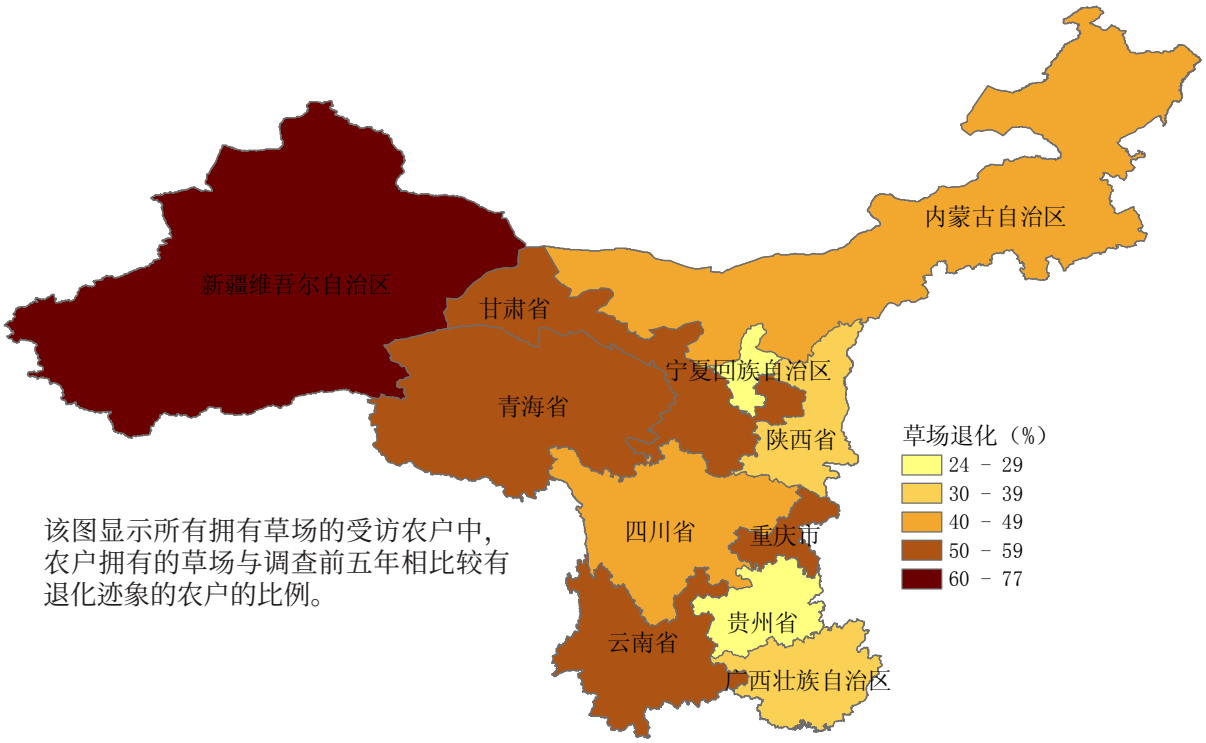


对房屋损失的补偿



该图显示在所有失去房屋的家户中获得补偿的家户的比例。

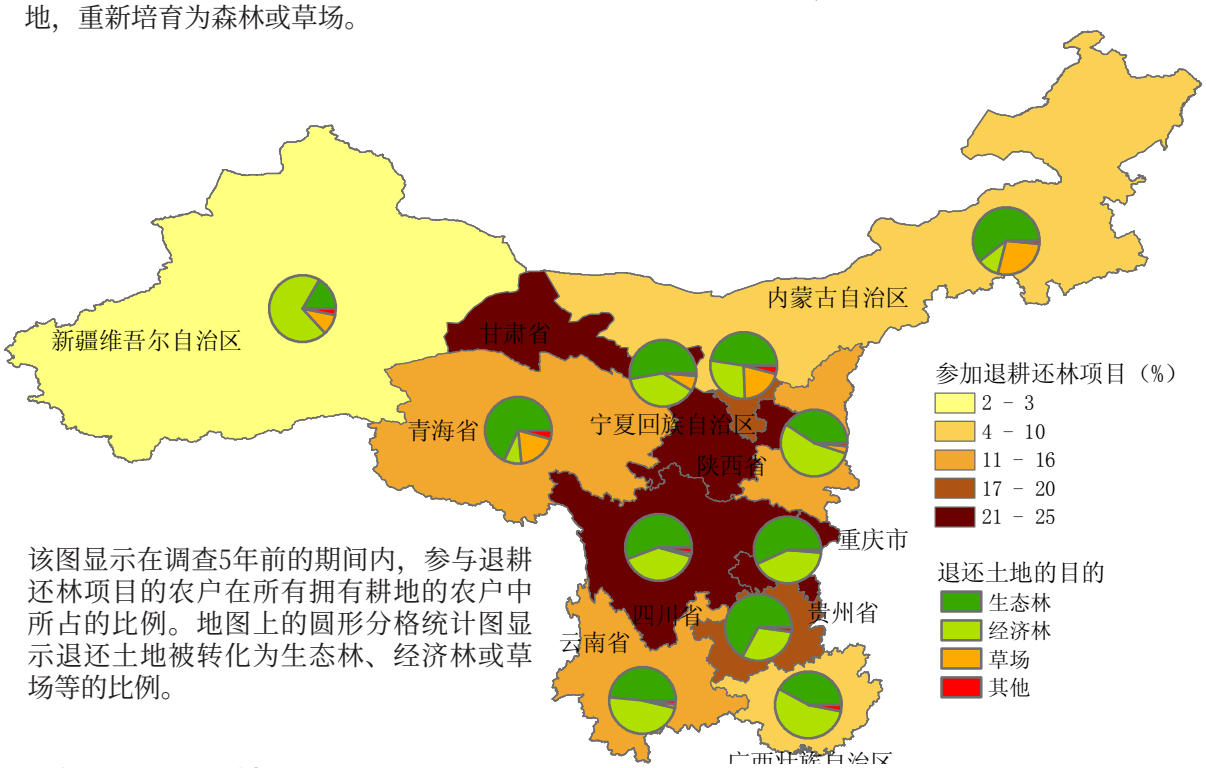
草地退化



该图显示所有拥有草场的受访农户中，农户拥有的草场与调查前五年相比较有退化迹象的农户的比例。

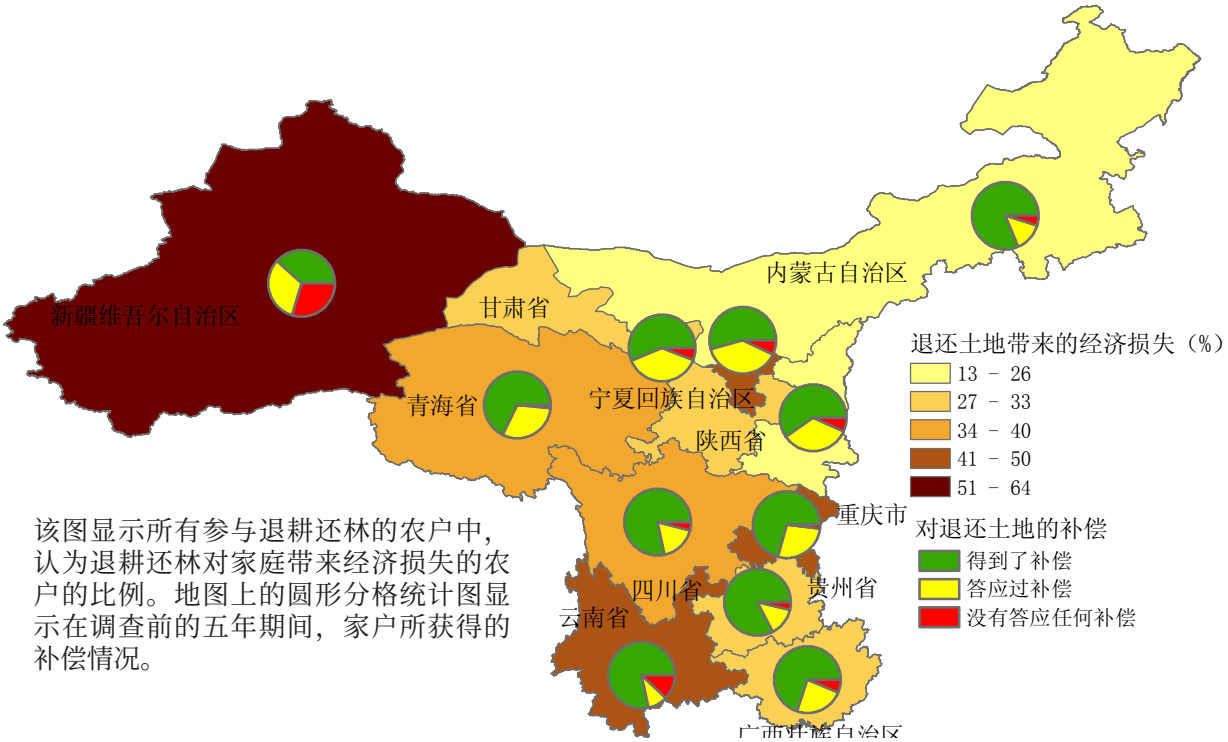
参与退耕还林项目

退耕还林项目包括退耕还林/还草、退牧还林/还草或退渔还湖项目，最初在1999年在四川、陕西和甘肃推行，到2003年有23个省（市、自治区）参与退耕还林项目，大多为西部的省（市、自治区）。退耕还林项目旨在帮助中国的西部地区恢复生态平衡。采取用粮食补偿的形式，鼓励农民退还他们生产效益较差的耕地，重新培育为森林或草场。



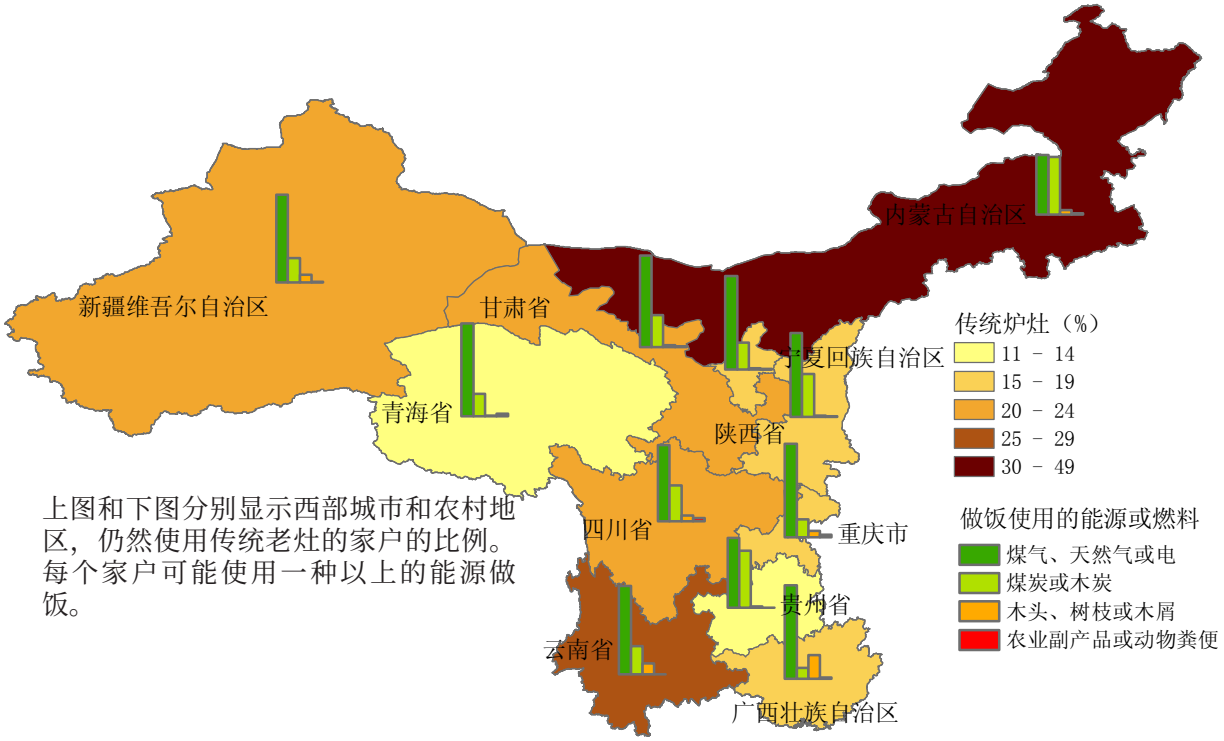
该图显示在调查5年前的期间内，参与退耕还林项目的农户在所有拥有耕地的农户中所占的比例。地图上的圆形分格统计图显示退还土地被转化为生态林、经济林或草场等的比例。

经济损失和退耕还林的补偿



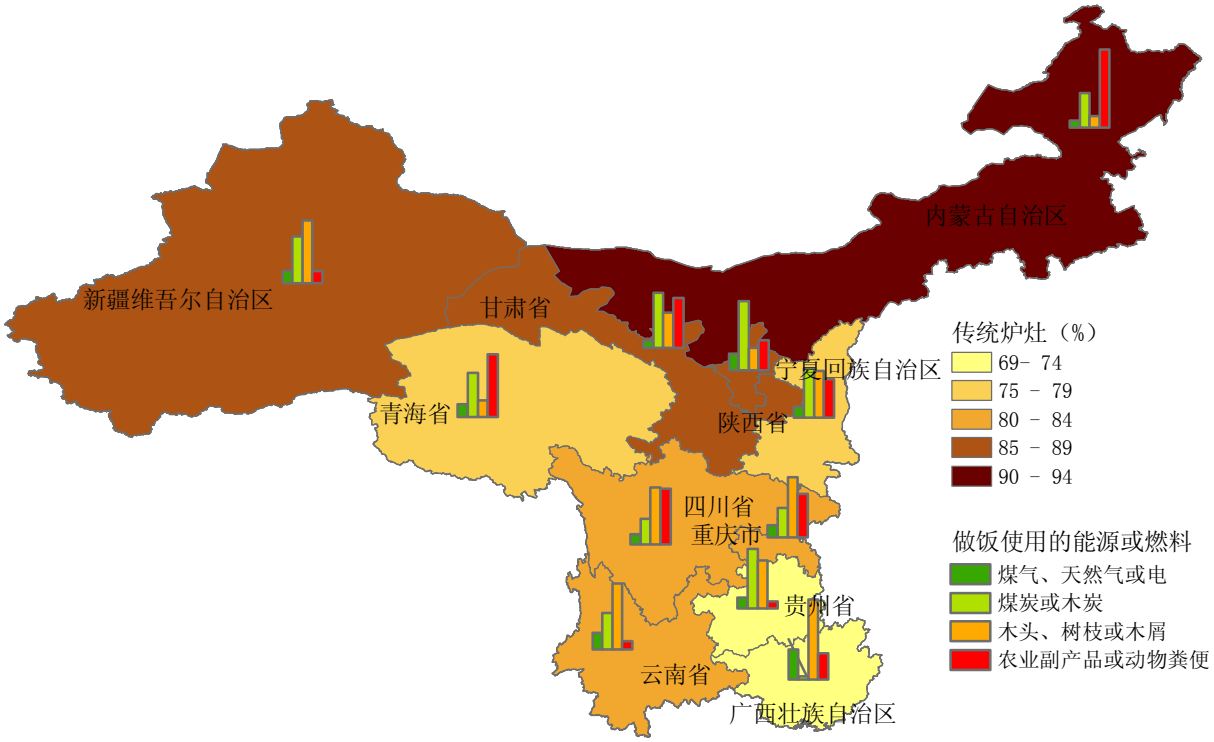
该图显示所有参与退耕还林的农户中，认为退耕还林对家庭带来经济损失的农户的比例。地图上的圆形分格统计图显示在调查前的五年期间，家户所获得的补偿情况。

炉灶和能源（城市）

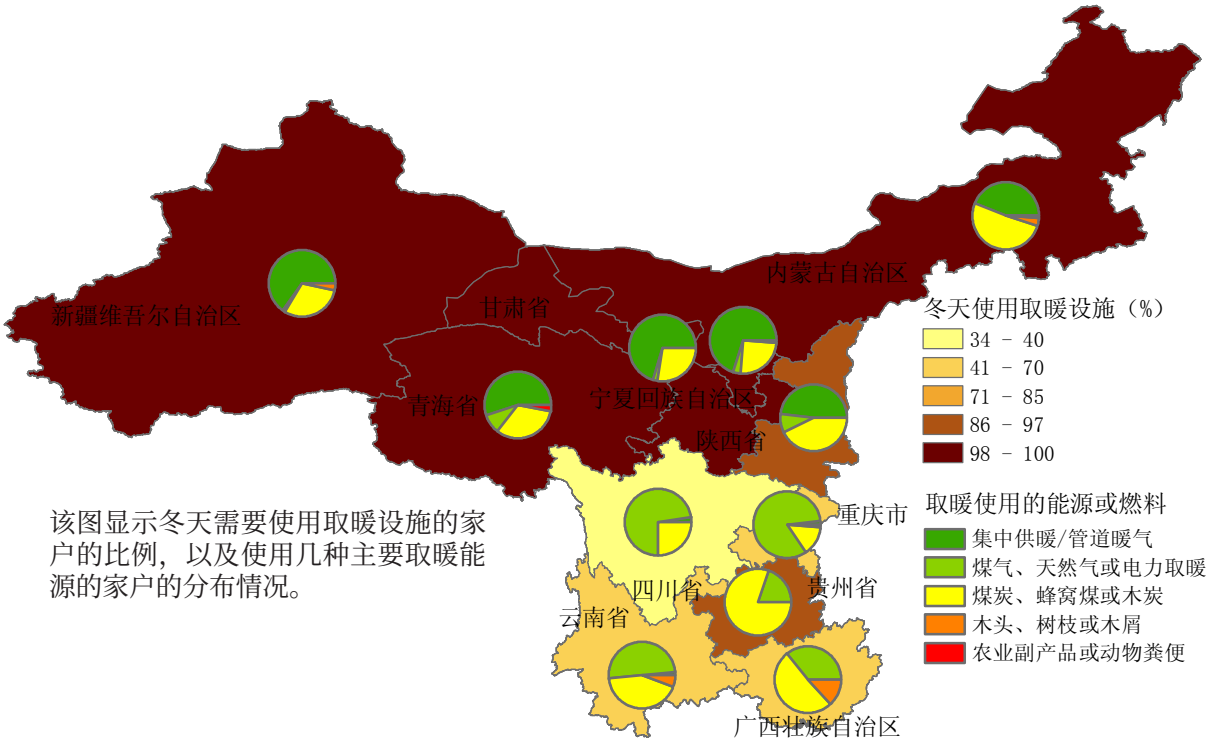


上图和下图分别显示西部城市和农村地区，仍然使用传统老灶的家户的比例。每个家户可能使用一种以上的能源做饭。

炉灶和能源（农村）

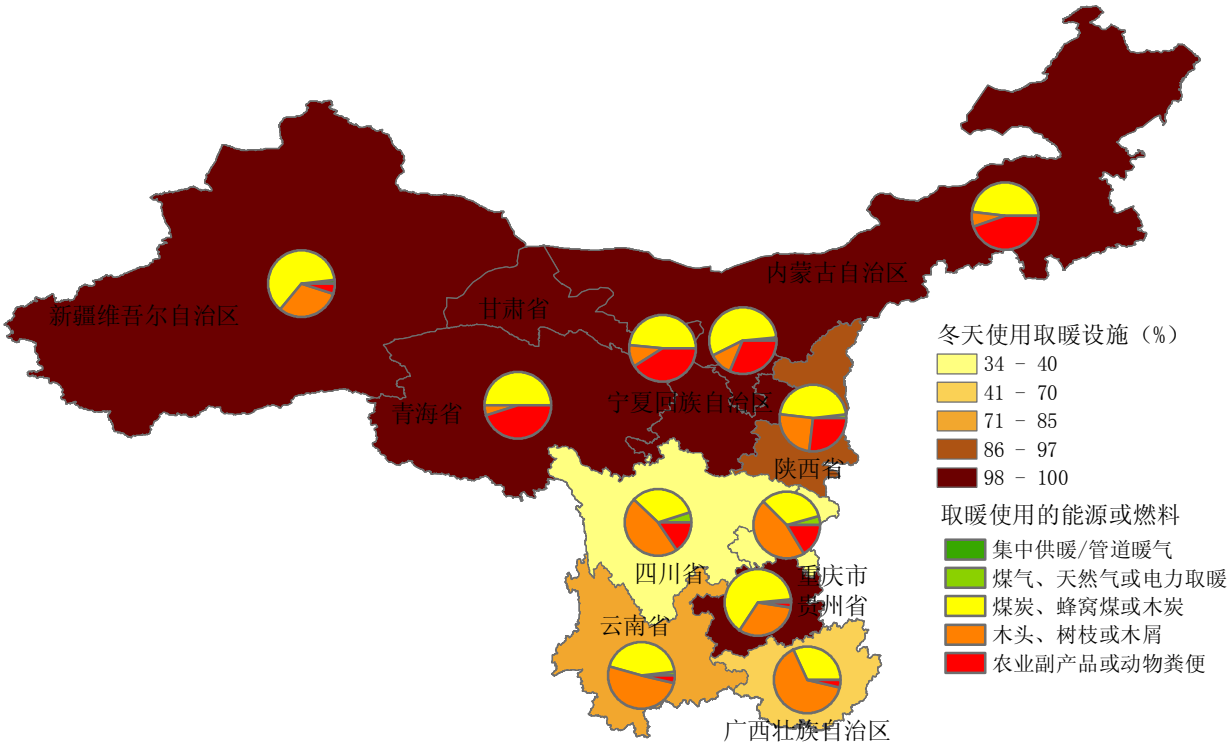


冬天取暖能源（城市）

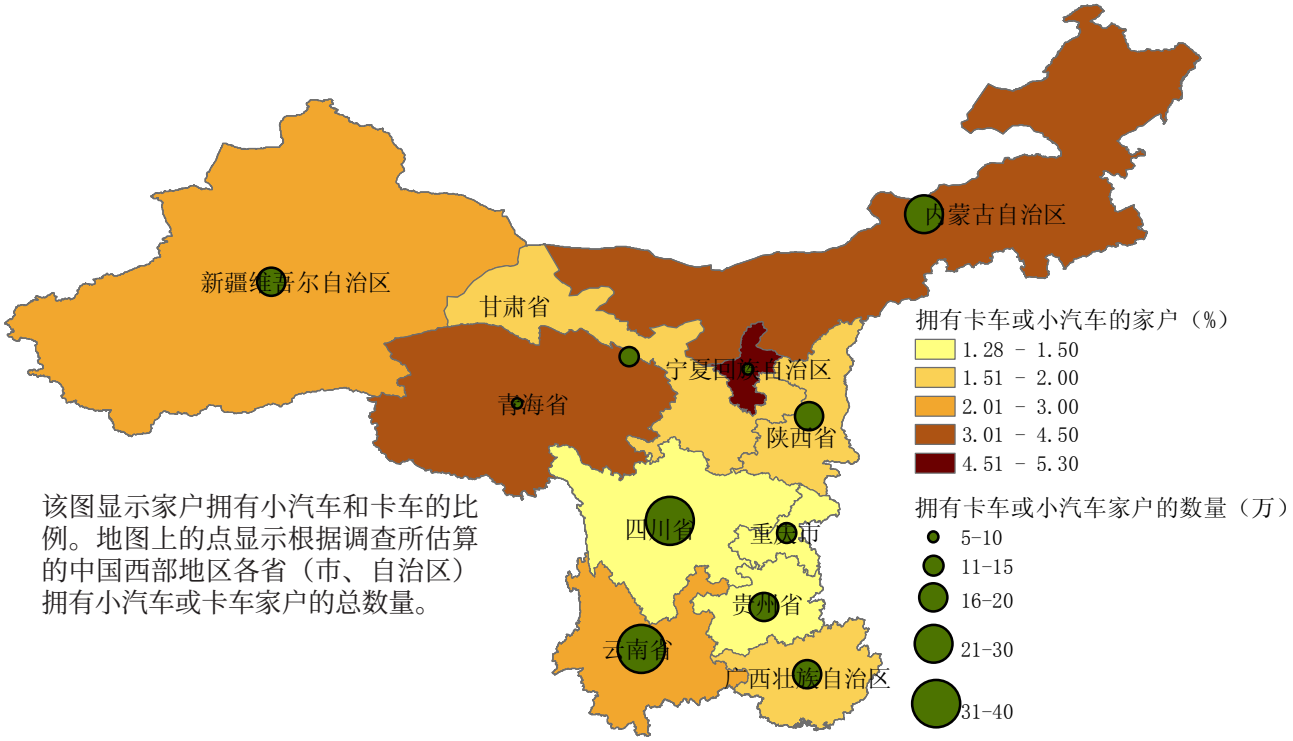


该图显示冬天需要使用取暖设施的家户的比例，以及使用几种主要取暖能源的家户的分布情况。

冬天取暖能源（农村）

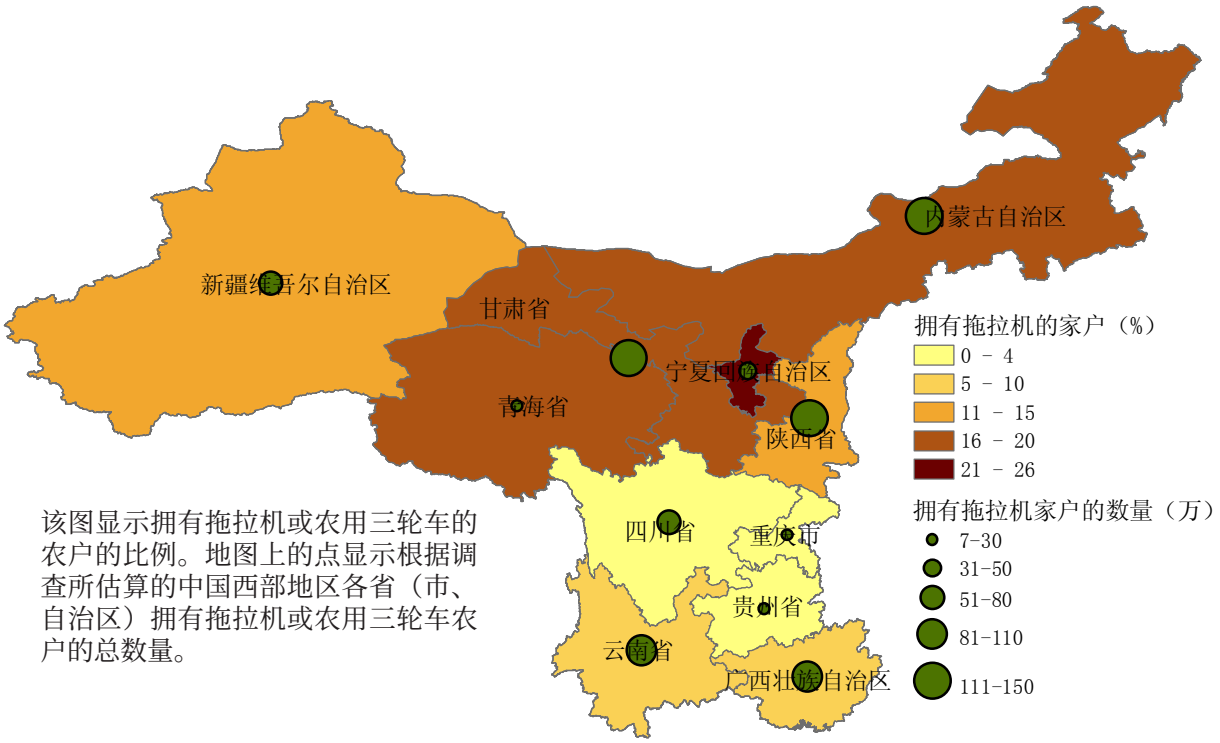


小汽车和卡车



该图显示家户拥有小汽车和卡车的比例。地图上的点显示根据调查所估算的中国西部地区各省（市、自治区）拥有小汽车或卡车家户的总数量。

拖拉机



该图显示拥有拖拉机或农用三轮车的农户的比例。地图上的点显示根据调查所估算的中国西部地区各省（市、自治区）拥有拖拉机或农用三轮车农户的总数量。

附录：抽样方案说明

MEDOW项目采用了标准的家户调查设计方案，下面我们简要介绍一下抽样的方法和步骤。读者如果有兴趣了解更详细的信息，可以参考“Living conditions Survey for the Western Provinces of China: Sample design”(Pedersen and Tyldum 2003)，以及“MEDOW Mapping and listing manual”(Pedersen and Tyldum 2004)。

西部地区幅员宽广、人口众多，因此抽取的样本应具有足够的代表性。但是，与其他调查一样，MEDOW的样本量也受到经费预算的制约。在综合考虑各方面因素以后，我们确定计划的样本量为44000户。同时为了能够推断每一个省的情况，我们在11个省（区、市）平分样本量，即每个省4000户。

多阶段抽样

样本是分阶段抽取出来的。第一阶段的基础抽样框来自于国家统计局第五次人口普查的（城市）街道和（农村）乡镇名单。在城市的行政管理体系中，街道是最低一级的行政单位，但是通常每个街道又由若干个基层群众性自治组织（即居委会）组成，每个居委会作为相对独立单元存在；农村的乡镇与城市的街道对应，村委会（即行政村）与居委会对应。

各个街道、乡镇的人口规模差异很大。所有选中的街道、乡镇中，人口数从491到260535不等，中位数是24201人；如果以家户数来计，最少的只有144户，最多的却有76628户，中位数是7118户。村（居）委会也是大小不一，最少的只有18户，最多的达到14961户，中位数为474户。

乡镇、街道的选取采用了概率与规模成比例（PPS）的直线等距抽样方法。抽样框首先依据省份进行显分层（explicit stratification），然后按照城乡类型和地理位置进行隐分层（implicit stratification）。抽取样本时我们将每个省的样本分成18个独立、重复抽取的子样本。也就是说，从同一个抽样框中重复18次抽样获得的18个独立样本构成总样本〔关于重复抽样（replicated sampling），可以参考Koop 1988〕。这种方法有两个优点：一方面，使得方差估计变得相对容易；另一方面，这个样本可以用作“母样本（master sample）”，因为从总样本中可以很容易地选出一个子样本。尽管这种重复抽样一般要求事先进行比较复杂的分层以使得各层之间的差异明显，但事实证明采用隐分层的方法可以非常有效地解决这个问题。

每一个子样本包括14个初级抽样单位（PSU），每个省合计252个PSU，西部11个省共有2772个PSU。每个PSU计划抽取16个家户样本，计划的总样本量为44352户。由于有些样本被抽中两次，还有些无应答情况，所以实际调查的样本量要略微少些。

一个街道或乡镇被抽中后，我们接着依据统计局提供的村（居）委会名单来抽取次级抽样单元（2SU）。该阶段的抽样仍遵循概率与规模成比例（PPS）的原则，并使用抽取PSU时的选择号码（具体操作可参考Pedersen and Tyldum 2003）。由于上述抽样过程采取了分层和重复抽样的方法，这使得在绝大多数情况下，每个PSU仅有一个2SU被抽中。如果从方差角度来评价这个抽样设计的效应，那么可以发现这个设计近似于一个二阶抽样，而不是通常的三阶抽样。

列表

负责制图和列表的工作小组到达一个村（居）委会后，首先要解决两个问题：一是应该采取何种方法和步骤来获取家户名单，二是这个PSU是否需要分片。

从村（居）委会能否获得完整有效的家户名单，将决定列表员采用直接抄表还是重新制图列表。按常规来讲，所有的村（居）委会都应该有一份囊括所有常住人口和暂住人口的完整名单（花名册），但实际情况并不一定如此。有的名单虽然包括短期流动人口，但更新日期可能是一周前、一月前甚至更早；有的名单列出了户口在当地的人口，以及长期居住的外来人口，但是没有短期流动人口；有的名单甚至更加不完整。我们希望获得一份最完整的家户名单，尤其是要包括所有短期的流动人口，这样才能尽量避免国内很多调查的一个共同缺点：即仅仅依靠村（居）委会提供的花名册，而往往遗漏那些流动人口。

因此，列表员的第一个任务就是要判断所拿到花名册的质量。如果村（居）委会的负责人说这个花名册是不完整的话，列表员就必须制图和重新列表，即画出该区域内所有建筑的草图，并通过实地走访列出每一住所内的家户名单。如果该负责人说花名册是完整的，那么列表员首先要通过一组甄别问题来确认其完整性。然后，列表员要按照手册规定方法在该村（居）委会辖区内，选择20个住所逐一列出所有住户名单，然后逐一核对这些住户是否已经包含在村（居）委会提供的花名册上。其中，列表员要特别核对6个“可能有问题的”居所，包括刚刚建成的房屋、已经被部分拆除或毁损的房屋、工人集体宿舍、店铺、街上或桥下等可能住人的地方。在所有检查的住所中，一旦发现有任何家户不在花名册上，那么列表员就必须制图和重新列表。

对所有村（居）委会列表的结果统计，使用已有名单（花名册）的占85%，重新制图列表的占9%，另外6%则是部分制图列表、部分使用已有名单。

分片与抽取家户

列表员需要解决的第二个问题就是：这个村（居）委会是否要分片，就是将村（居）委会按照一定规则分成一些小的部分？在进行分片后，最终的样本户将从某一个分片中选出，而不是从村（居）委会整个辖区中选出。

可以分片的情况有以下几种：当村（居）委会很大，家户数量超过一定规模，以至于难以进行列表或制图；当一个行政村所属的各个自然村相距遥远，如果每个村子都有样本，那么实施调查的成本和时间就会增加。总的来说，我们的目标就是尽量使每片的家户数量在200户以下。当然列表员把片分得大一些也是可以的，只要能从村（居）委会获得符合要求的名单，并且没有什么实施上的困难。分片时，通常依据容易辨识的地理边界来划分。片的

抽取过程遵循概率与规模成比例的原则。所有实施调查的村（居）委会中，有分片的比例为55%。

列表员完成上述工作以后，将列表资料及时地寄回到省级办公室，交由抽样督导员进行下一项工作：按直线等距抽样方法从选中的片中抽取16个家户。考虑到城镇地区的无应答率会更高一些，所以部分居委会（占总样本的12%）抽取的家户数量调整为18户。另外，也发现个别村（居）委会的家户数量不到16户。最后完成调查的样本共有44738个家户。

根据研究设计，调查员入户调查时还要依据Kish表（Kish 1966）随机抽取出一位15岁及以上的家户成员完成RSI部分的问卷。首先要将所有符合条件的家户成员按年龄和性别排序编号，然后调查员根据问卷封面上事先印好的代码和Kish表找到应选的对象。

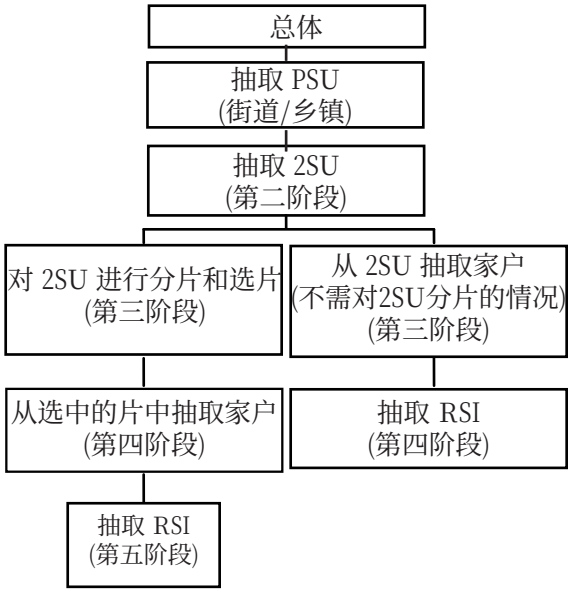
MEDOW 的样本说明

总结起来，MEDOW的样本有下列特征：

1. 首先以省份作为显分层标准，然后又依据城乡类型和地理位置进行隐分层；
2. 由18次有放回重复抽样获得的18个独立子样本构成；
3. 每个省份有4000个家户样本，而且每个省内的样本近似地自加权；
4. 依据2000年人口普查的汇总数据和每个行政单位的估计人口数来选取PSU；
5. 在每个省，村（居）委会的抽取都遵循概率与规模成比例（PPS）的原则。
6. 过半数的村（居）委会进行了分片，并按PPS的原则选取了其中一片；
7. 通常直接抄录已有的花名册即可获得该村（居）委会或者选中的片的家户列表；而一旦这个花名册不符合要求，则必须画出草图并列出该辖区内所有住户清单。
8. 在确认后的名单中，按照直线等距抽样的方法抽取家户；
9. 每一家户中，用Kish表随机选出一位15岁及以上的成员回答RSI部分的问卷。

入样概率和权数

根据上文知道，MEDOW的抽样方案属于一种多阶段抽样。具体来说，又分成两种情况（如下图所示）：



可以看到，家户样本和随机选取的个人（RSI）样本的抽取都存在两种情况。如果以家户（或家户中的所有成员）作为最终抽样单元，这是一个三阶段或四阶段的抽样；如果以随机选取的个人（RSI）作为最终抽样单元，这就成了一个四阶段或五阶段的抽样。

符号

为了便于描述和计算入样概率，这里引入一些计算符号（见表 1）。

表 1：符号说明

符号	解释
N	家户数或人数 （初始估计值）
N^l	列表获得的家户数
N^q	列表中快速估计的家户数（仅用于分片和选片）
N	样本数（指定的）
m	应抽取 的PSU 抽样单位个数
K	应抽取的2SU个数
P	入样概率
h	层号
c	抽样单位编号 （c后面 2,3,4表示抽样的阶段，第一阶段省略）
k	PSU编号
f and I	家户编号 (f用于标识不同阶段的家户, i用于标识样本中从1 到n的所有家户)
d	家户成员编号

抽取PSU、2SU和家户

对于18个重复抽样的子样本中的任何一个，其中入样概率的计算方法如下：

第一阶段：

对于h 层中PSU c的入样概率为：

公式 1：

$$p_1 = p_{s,c} = \frac{N_{h,c}}{N_h} \cdot m_h$$

该公式适用于计算所有第一阶段的入样概率。

第二阶段：

第二阶段的公式与第一阶段类似，不同的是这里m 等于常数1，这样公式中就省略了它：

公式 2：

$$p_2 = p_{s,c,c_2} = \frac{N_{h,c,c_2}}{N_{h,c}}$$

（可选的）第三阶段：

当遇到分片的情况时，所选片的入样概率为：

公式 3：

$$p_3 = p_{h,c,c_2,c_3} = \frac{N^q_{h,c,c_2,c_3}}{N^q_{h,c,c_2}}$$

其中，　q表示这里使用的是列表员快速估计的家户数。

这也是计算概率与规模成比例的入样概率标准公式，其中省略了等于常数1的m 。

第四/第三阶段：

在每一个PSU都抽取固定数量的家户。这样，在h层中的某个PSU或其中一片中（无论经过几个阶段）家户f的入样概率为：

公式 4：

$$p_4 = p_f = \frac{n_{h,k}}{N^l_{h,k}}$$

这里使用的是由列表员在实地获得的家户数，而不是人口普查或者统计局提供的原始数据。

确定家户：

综合上述公式，家户i的入样概率为：

公式 5：

$$p_i = p_1 p_2 p_f \quad \text{（没有分片的情况），或}$$

$$p_i = p_1 p_2 p_3 p_f \quad \text{（有分片的情况）}$$

个人的入样概率

在MEDOW调查中，计算家户成员个人的入样概率要相对复杂一些。一般情况下，个人的入样概率与家户的入样概率是相等的，因为个人是否被抽中取决于所在的家户是否被抽中。不过有两种例外：一是一个人属于多个家户的情况，二是要讨论每个家户中随机抽选的个人（RSI）。下面我们分别来讨论。

MEDOW调查中之所以会出现一个人属于多个家户的情况，是因为我们强调要包括流动人口，即一部分住在（调查地当前住所以外的）其他住所的家户成员，比如常年住在城里、只是偶尔回农村老家的那些流动人口。此外，还包括那些一年中的某些时间与家人分离的游牧者或短期流动者。调查结果显示，约有20%的人口属于各种类型的短期居民。

因为被访者可能居住在多个不同的住所，那么只要这两个住所都在我们所研究的这11个西部省份范围内，从理论上讲就有可能被抽中不止一次。这时，个人在当前样本中的入样概率仍然等于家户的入样概率；在另一个住所的入样概率则无法直接估计，只能假定等于这个省的平均入样概率。所以对于有多个住所的那些个人，总的入样概率就由上述两个概率决定：

公式 6：

$$p_{individual} = \begin{cases} p_i & \text{（唯一住所）} \\ 1-((1-p_i)(1-p_p)) & \text{（两个或多个住所）} \end{cases}$$

这里 p_p 指另一住处所在省份的平均入样概率。

每个家户只抽取一个RSI，所以有N个成年人的家户f中RSI d的入样概率为：

公式 7：

$$p_r = p_d = \frac{1}{N^{\geq 5}_{f,d}}$$

这样，RSI的总入样概率就是$p_{individual}\, p_r$。

至于其他的个人（比如妇女、儿童），他们并不是从家户成员中抽取的，所以个人的入样概率就等于家户的入样概率。

样本权数

样本权数分成两种：绝对权数（expansion weights）和相对权数（relative　weights）。前者可以用于估计总体的规模，数值上等于入样概率的倒数。而后者保持样本规模不变，只是调整了每一个分析单位（家户或个人）的相对分布，它等于相应的绝对权数除以全体绝对权数的均值。

家户i的绝对权数的计算公式为：

公式 8：

$$W_i^e = \frac{1}{p_i}$$

当然，上述公式计算的权数还不能直接用于估计调查的结果。因为，最终的权数还要根据下文讨论的无应答情

况做相应调整。

计算个人以及RSI的样本权数方法与家户的样本权数相类似。这样，我们得到三组样本权数：分别是家户权数、个人权数和RSI权数。

无应答

我们可以通过对问卷无应答的详细分类来分析调查完成情况（参考Hidiroglou, Drew and Gray 1993）。表 2列出了应答情况类别和对应的频数分布。

调查无法成功进行的原因可能有两种：第一，所选中的家户不属于抽样框范围，比如在华的外交人员；第二，所选中的家户存在也符合条件，但是拒绝接受访问或者家里找不到人。

另外，还要考虑到调查员无法确定家户是否存在的一些情况。即使调查员能找到受访家户，也可能无法获得任何有效信息，因为受访者生病或者其他原因导致无法回答问题。

在所有抽中的44738个家户中，调查员能够实施调查的有41695户，其中41222户完成了全部问卷。详细的情况如下表所示：

表 2: 无应答情况

应答情况分类	频数	百分比
1 - 访问顺利并全部完成	40915	91.5
2 - 受访者拒绝，但被说服后完成全部问卷	307	0.7
3 - 部分完成	473	1.1
4 - 住户情况不明	70	0.2
5 - 没有有用的信息	132	0.3
6 - 住所不存在	78	0.2
7 - 没有人住在这里	538	1.2
8 - 住房正在建设中	5	0.0
9 - 不符合受访条件	124	0.3
10 - 无法与受访户取得联系	1,689	3.8
11 - 拒绝访问	407	0.9
合计	44,738	100.0

我们把应答率定义为：最终接受调查的家户数占所有符合条件的受访家户数的比例。MEDOW调查中，43858个符合条件的受访家户中有41222户完成了全部问卷，应答率为94%。拒访率是指拒访的家户数占所有能联系上的家户数的比例。MEDOW调查中仅有0.96%的拒访率表明导致无应答的主要原因不是拒访，而是列表错误或者调查员无法与受访户取得联系。

权数校正

当存在无应答情况时，直接使用抽样的样本权数会导致有偏估计。这种偏误一般表现为两种情形：第一，无应答的样本单元被当作缺失处理后，用抽样权数估计的总和可能太小；第二，无应答的样本单元可能具有某种特征，这导致估计出现偏误。

要消除无应答引起的偏误，就要对样本权数进行校正。MEDOW项目中，针对无应答情况的权数校正采用了所谓的“调整单元方法”（adjustment cell method，参考 Lehtonen and Pahkinen 1995; Little and Rubin 1987）。这种方法将那些非常相似的家户指定为一组家户（即调整单元，adjustment cell），然后计算每组中的无应答率。按照这种方法计算的无应答率只考虑那些本可以应答而因为某种原因没有应答的那些无应答情况。

然后，我们用每个调整单元的无应答率来调整每个家户的样本权数（包括绝对权数和相对权数），得出所谓的估计权数（estimation weights）。用估计权数加权后，样本规模就等于总体大小，同时还增加了那些与缺失问卷相似的样本的代表性。

在MEDOW项目中，调整单元组（adjustment cell groups）的划分依据三个指标：省份、地市、以及在村（居）委会中心用GPS测量的海拔高度。调整单元按照海拔高低分成规模相等的三个组。如果调整单元只包含一个或两个村（居）委会，那么将它们归并到相邻的组。

使用表 3中的符号，无应答权数校正系数可以用公式9来表示.

表 3: 校正无应答的符号说明

记号	说明
C	校正系数
a	调整单元的编号
h^r	应答的家户数
h^f	无应答的家户数

所有可能应答的家户数（即无应答率的分母）等于表 2中1，2，3，5，10 以及 11类样本数的总和。而无应答的样本数等于3，5，10以及11这四类总和。“住户情况不明”一类则不在讨论之列。

公式 9：

$$C_a = \frac{1}{\dfrac{h_a^r}{h_a^r + h_a^f}}$$

所以，调整后的权数就是：

公式 10：

$$W_i^{adjusted} = C_i W_i$$

校正带来的效应就是绝对系数增大，最终的估计权数之和等于抽样框的单元总数（不包括少量不存在或不符合条件的单元）。

国家统计局公布了每个省份年底估计人口数（国家统计局 2005）。为了使得调查估计的人口数与国家统计局的估计数一致，我们对权数做进一步调整，公式如下所示：

公式 11：

$$W_{p,i}^{estimation} = W_{p,i}^{adjusted} \frac{P_p}{\sum W_{p,i}^{adjusted}}$$

其中，P_p是样本i所在省的人口数。

参考文献

Hanssen-Bauer, J., J. Pedersen and A. Tiltnes [eds.] 1997, Jordanian Society, Oslo: Fafo.

Hidiroglou, M., J. Drew and G. Gray 1993, "The measurement of non-response in surveys": Survey Methodology 19: 81-94.

Kish, L. 1965, Survey sampling. New York: Wiley.

Koop, J.C. 1988, "The Technique of Replicated or Interpenetrating Samples" in P.R.Krishnaiah and C.R. Rao [eds.]: Handbook of Statistics, Vol 6: 333-368, New York: Elsevier.

Lehtonen, R. and E. J. Pahkinen 1995, Practical methods for design and analysis of complex surveys. Chicester: Wiley.

Little, R. and D. Rubin. 1987, Statistical analysis with missing data. New York: Wiley.

Pedersen, J and G. Tyldum 2003, "Living conditions Survey for the Western Provinces of China: Sample design", Oslo: Fafo

Pedersen, J and G. Tyldum 2004, "MEDOW Mapping and listing manual", Oslo: Fafo

Wolter, K.1985, Introduction to Variance Estimation, New York: Springer Verlag.

国家统计局 2005,《中国统计摘要2005》，北京，中国统计出版社

西部人民的生活

中国西部省份社会与经济发展地图集

本地图集收集了150多张地图和图表，比较全面深入地展示给读者中国西部地区社会和经济的发展状况。基于对44，000多户家户的调查，这本地图集覆盖了对11个省（市、自治区）各个方面的研究信息，包括了人口，居民健康，教育，家庭经济，劳动力市场，农业，房屋和基础建设以及环境等问题的研究。

中国西部省份社会与经济发展地图集是“中国西部省份社会与经济发展监测研究”项目的成果之一，另外还包括 “中国西部省份社会与经济发展监测研究” 数据报告和分析报告。

NRCSTD Fafo AIS

中国科学技术促进发展研究中心
挪威Fafo应用国际研究所

本地图集仅用于课题内部交流，不得公开引用

西部人民的生活 中国西部省份社会与经济发展地图集