

第三章 建模过程

在第 2 章中,相信您对 ArcView 在很多领域中的分析功能已经有所实践和操作,对空间分析模块解决问题的潜在能力也有了一定程度上的认识,但是空间分析模块本身并不能解决所有问题。空间分析模块可以回答您诸如“某东西有多大”“事物之间距离有多远”等问题,但是要得到“这是否是最佳地点”之类问题的答案,您就需要自己提出合适恰当的标准。本章将逐步引导您来建立自己的模型,并帮助您分解问题和选用合适的工具处理相应问题。

在本章中,您将学到:

- 如何将空间问题、数据以及操作等概念化;
- 如何为找到合适的地点创建模型;
- 如何将问题分解为软件所能解决的各个部分;
- 如何定量化“什么是好的”;
- 什么时候选用不同类型的空间数据;

1 如何将空间问题模型化

所谓模型是指一种描述不能直接观察的事物的方式。因为在我们的生活环境中的所有现象都不能同时直接观测到,因而我们通过使用模型以创建真实世界的简化表达。具体到地理信息系统(Geographic Information System, GIS)中,是通过使用一系列的地图图层以及它们之间的关系来表达真实世界的。在 ArcView 中,这些图层即为所谓的主题,而图层之间的关系则用空间分析工具模型化。

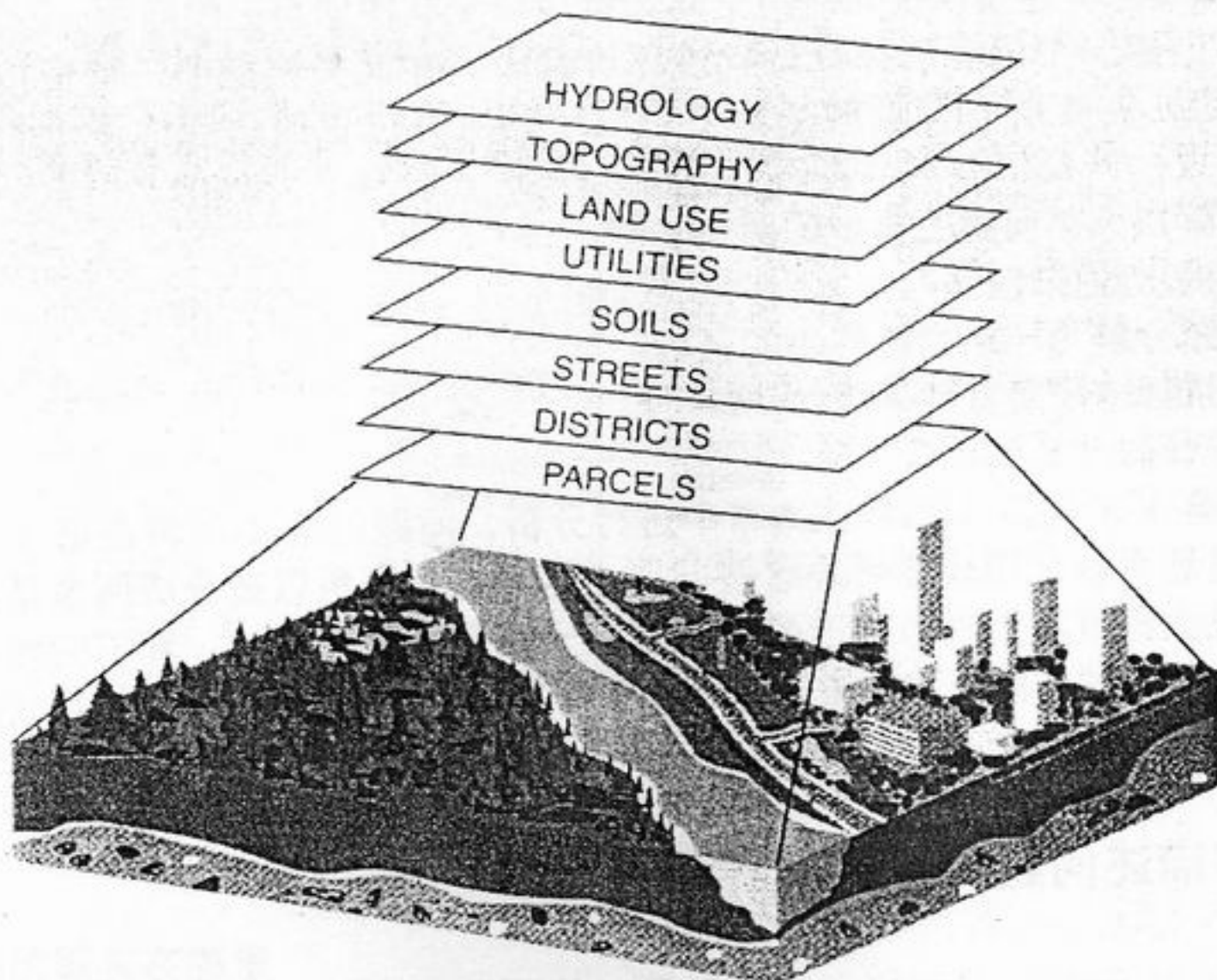


图 1-16. GIS 中的地图图层

空间建模是指从着眼于每一个位置在许多图层中的特征开始到问题得到解决的整个过程。在这个过程中，空间分析模块在被选择图层上加上一个矩形网格或栅格，每一个栅格单元代表一个位置，且在每一个图层中均有一个值。不同图层的网格单元堆积起来，描述每一位置的众多属性（图 1-17）。

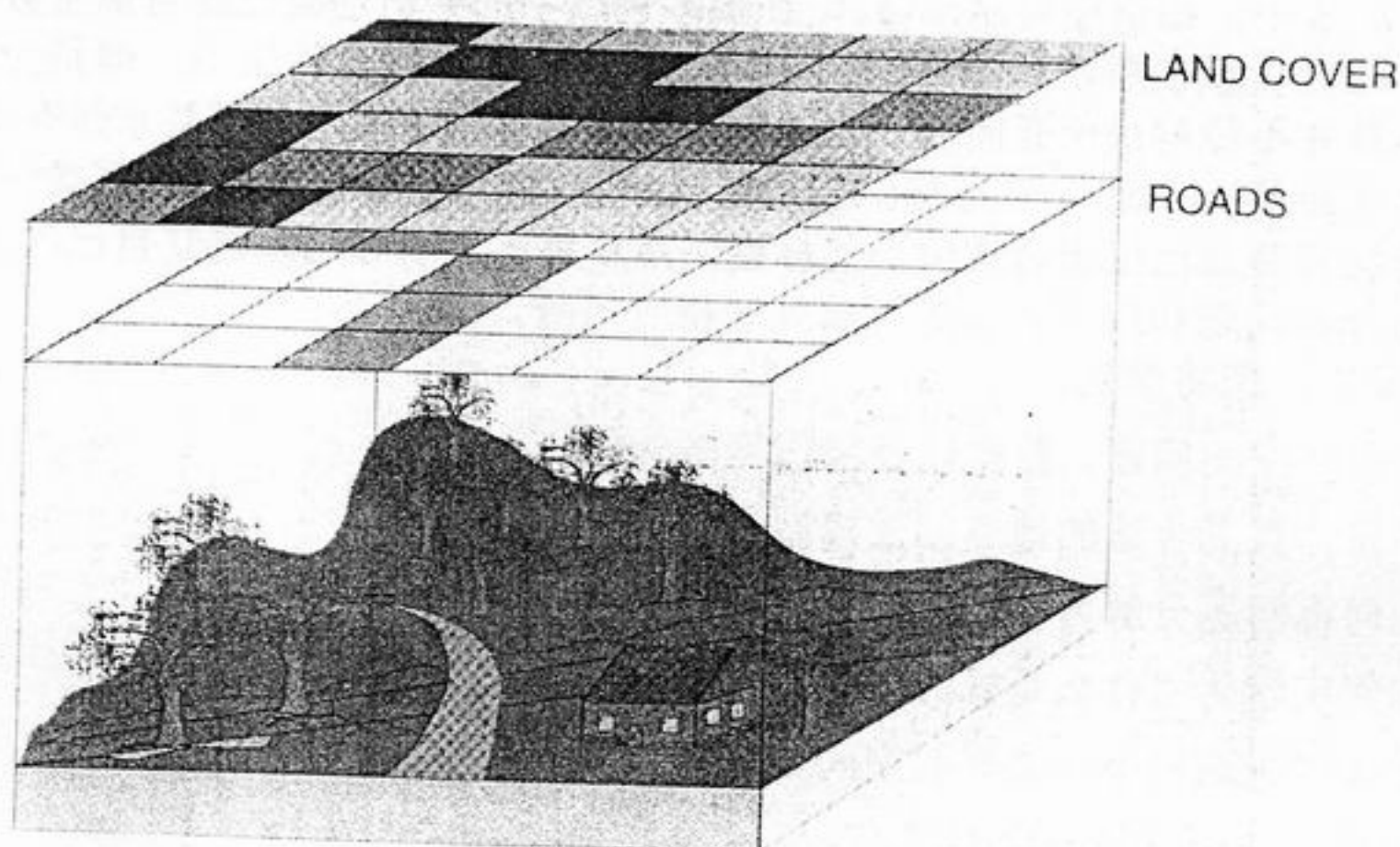


图 1-17. 空间建模的栅格单元

大多数的空间分析模型包含有找出最佳位置的问题，这些模型是位置选择或适宜性评价型模型，其目的是为诸如种植杂交作物或是钻探石油或是掩埋垃圾或是新建学校等问题找到最好的地点。虽然这些问题在尺度上和数据要求上有很大的不同，但它们的解决方法是一致的。在本章和接下来的一章中，我们将一步一步地为一个新开的商店进行选址。从头到尾认真的分析和学习这个过程将有助于您更好的理解怎样建模以及如何用空间分析模块来解决您自己的问题。

您将要创建一个位置选择或适宜性评价模型。在此类模型中，特征按适宜程度进行分级，并在所有变量的基础上生成一幅所有位置的适宜性评价的合成地图。

如下是解决该类问题的典型的四个步骤：

1. 问题提出和描述；
2. 将问题分解为一系列待解决的目标；
3. 为各目标分配适宜性等级；
4. 解决问题。

问题的第 1 步到第 3 步将在本章中进行分析，问题的第 4 步将在第 4 章中分析，在那里您可以使用软件解决这些问题。这一过程之所以被分成两章是因为建模过程的很大程度在于思考与准备，理解这一点是非常重要的，真正的使用软件仅仅是解决问题的一部分。如果前面三步做得很好，第 4 步用软件解决问题将会非常的容易。

2 提出和描述问题

解决问题的第 1 步是提出和描述问题。从研究分析预计的输出概念着手；形

为找出优秀商店销售区内的人口信息，您可以察看调查数据，生活方式数据，或者政府人口普查数据等。在这个案例中，您可以在调查结果的基础上使用生活方式数据找出符合您的“潜在顾客”定义的群体。在调查和普查信息的基础上，生活方式数据将具有相似的社会背景和经济地位背景的人群聚集成组。对于定义消费者的行为，这样的数据比原始的官方户口调查数据更加容易使用和更为有效。

这些数据在研究中非常有用，因为您已经有了目标顾客的一些标准。这里您可以根据顾客生活方式的分类定义谁是最佳顾客以及他们居住在哪里。您可以通过研究每个销售区的生活方式类别的统计汇总以及寻求它们之间的相似性和关系实现。

研究的结果将是生活方式类别的一个列表，这些类别在成功商店的销售区内最为共同。

生活方式数据

国家统计和决策系统等机构会把人口按其社会地位和经济地位分成统计群，因而从这些机构可以获得人口统计数据。这些数据可以用在基于“具有相似的社会地位和经济背景的人群倾向于比邻而居”假设的基础上的市场分析中。假定某人的居住地很大程度上影响其想法、行为及购买倾向，同样，您也可以通过人们的衣着和交通工具判断他们的可能居住地，通过其居住地判断他们可能的购买倾向。

绘制潜在顾客的百分比地图

现在您已经知道您所寻求的顾客，接下来就需要找出他们，最好的办法就是创建一个潜在顾客的百分比图。生活方式数据代表着一个地理单元内的人口统计抽样，而不是该地区内所有人口数目。绘制潜在顾客的百分比地图将允许您在不同的抽样密度区里进行比较。创建百分比图是一个简单的表操作，在此，您要总计每个需要的类别的人口数量，除以被抽样的人口总数，然后再乘以 100。

顾客是否充足？

为确定顾客是否充足，您需要创建一幅商店销售区内的人口数量图。从调查的数据中可以得知，大多数的顾客旅行一般不超过 3 公里，因此该图应显示在研究区内每个位置 3 公里以内的人口数量。

从人口的点数据集，利用“Calculate Density(计算密度)”功能可以创建一个人口密度图。

顾客距现有的商店足够远吗？

从调查的数据中得知，几乎没人愿意到离家 3 公里以上的商店购买商品。为防止现有商店的客源流失，您可以使用空间分析模块创建一个表示到现有商店的距离的地图。任何到现有商店超过 3 公里距离以上的地方即为理想地方。您可以使用软件的“Find Distance(发现距离)”功能完成工作。

象化您想要的地图类型。

您的问题是为新开商店选择最佳的区域范围。假设您拥有邻近一些商店的经销权，这些商店的主要服务对象是高层次的职业人员，您将在上午出售给他们顶级的咖啡和新月面包，而晚上则出售给他们各种各样的可口的外带食品。其中有些商品做得非常的好，而且您现在想寻找最佳的房产以增开新的店。这样您所寻求的结果可表示为一张地图，它为新商店的区域范围从最好到最差分出了适宜性等级。因为这样的图显示了一个地区的适宜程度的度量值，所以称之为适宜性分级图。这种地图的种类及其生成过程可用于空间分析和 GIS 的各个领域。

从提出和描述问题开始，画出一个模型框图。随着问题的处理过程，您将一步步扩展该框图以显示数据和操作的流程。



3 分解问题

问题被提出之后，将它分解成越来越小的部分，直到您非常清楚的知道解决它需要哪些数据和步骤（图 1-18）。这些步骤是您为了度量各个位置适宜性所需解决的目标。

定义这些目标的一个考虑是，它们必须是可度量的。您如何度量“哪里是一个新店的最好地点”？毫无疑问，一个成功的商店需要充足的顾客。您有一些成功的商店，这些商店有着充足的顾客。这些顾客最重要的特征就是其他他们对您的商品有购买欲和距您商店的合适的距离。

接下来您的问题将是：“潜在的顾客(称之为 good customers)在哪里？他们的数量足够多吗？他们离现有的商店足够远吗？”这些问题即为寻找答案的三个目标或问题，它们就将成为数据层或主题。



图 1-18. 分解问题

潜在的顾客在哪里

您关于潜在顾客的特征是建立在现有商店的调查和信用卡收益的基础上的，

但是您需要将它们加以量化以便图示出潜在的商店位置。

为创建潜在顾客的地图，您需要为喜欢您商品的顾客定义一些特征。您可以从调查数据中了解这些。从商店的位置和属性数据集中，可以选择出成功的商店，图示它们的销售区，并且与人口统计的数据相结合，证实调查中的最佳顾客和居住在成功商店附近的人群是一致的。然后，分析这些统计人口的特征，做出相应的人口地图。

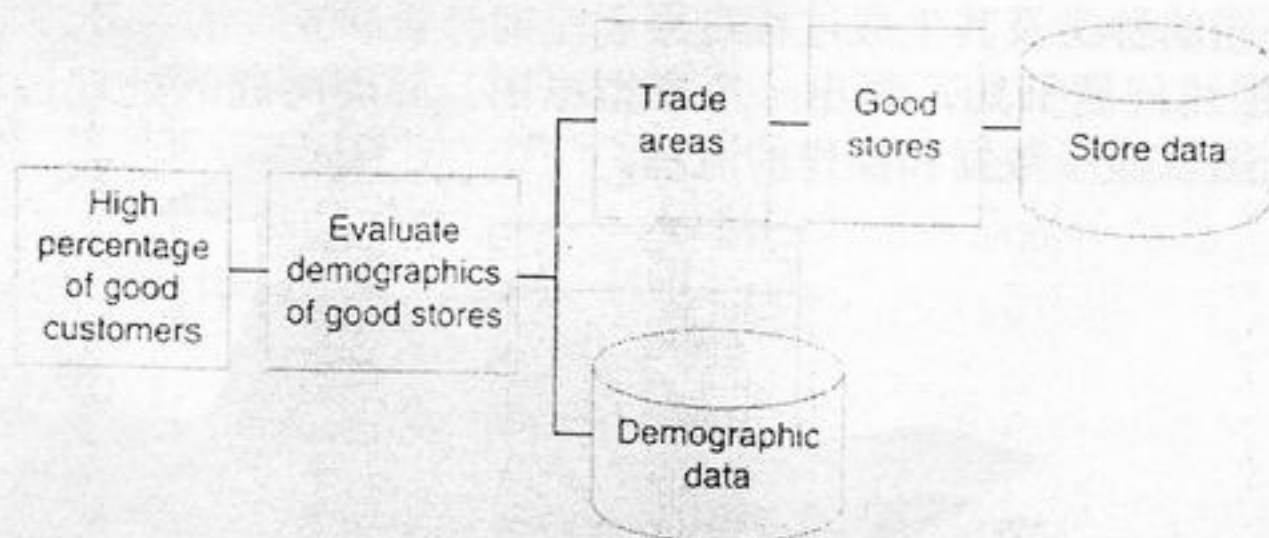


图 1-19. 潜在顾客问题分解图

度量目标

找到度量目标的理想的数据和工具并不总是可能的。这时，您可能需要创造度量目标的替代的方法。例如：为了防止流失顾客，您可能想知道顾客愿意走多远才会选择您的商店。“多远”究竟意味着什么？它是顾客愿意付出的距离或旅行时间吗？度量一个人愿意走“多远”的最直接的尺度是旅行时间。但是，要测量旅行时间，您需要公路网络、限速以及阻碍因素等属性，就象 ArcView 网络分析模块一样。一天中不同的时段，一周中不同的天以及其它的变量都将改变同一路径上的旅行时间。

为了简化数据获取和分析，有时候应采用间接的手段。例如在本例子中，您可以使用距离商店一定距离的、半径不变的尺度。

在研究要求的范围内，度量应该尽可能地直接。直接和间接两种方式应该协调使用。直接度量由于有很多诸如不同时段和旅行路线等不确定因素而很难精确完成；间接测量虽然很容易实现，但它不够准确，为结果的解释留下了太多的不确定性。

确定优秀商店的销售区

人们往往只愿在一定距离范围内购物，所以只应该考虑商店销售区以内的人口统计数据。要考察成功商店的人口因素，首先需要定义它们的销售区。您可以使用“Find Distance(发现距离)”功能。

销售区的确定

“Assign Proximity(指定邻近)”的功能也可以用来创建销售区地图，但它不允许您限制顾客愿意旅行的最大距离。如果结合费用表面(Cost surface)和最大分配(maximum allocation)，如创建一个成功的销售区域范围所需的总人口，则可以用费用距离请求(Costdistance request)来为商店分配销售区。

评价优秀商店的人口统计

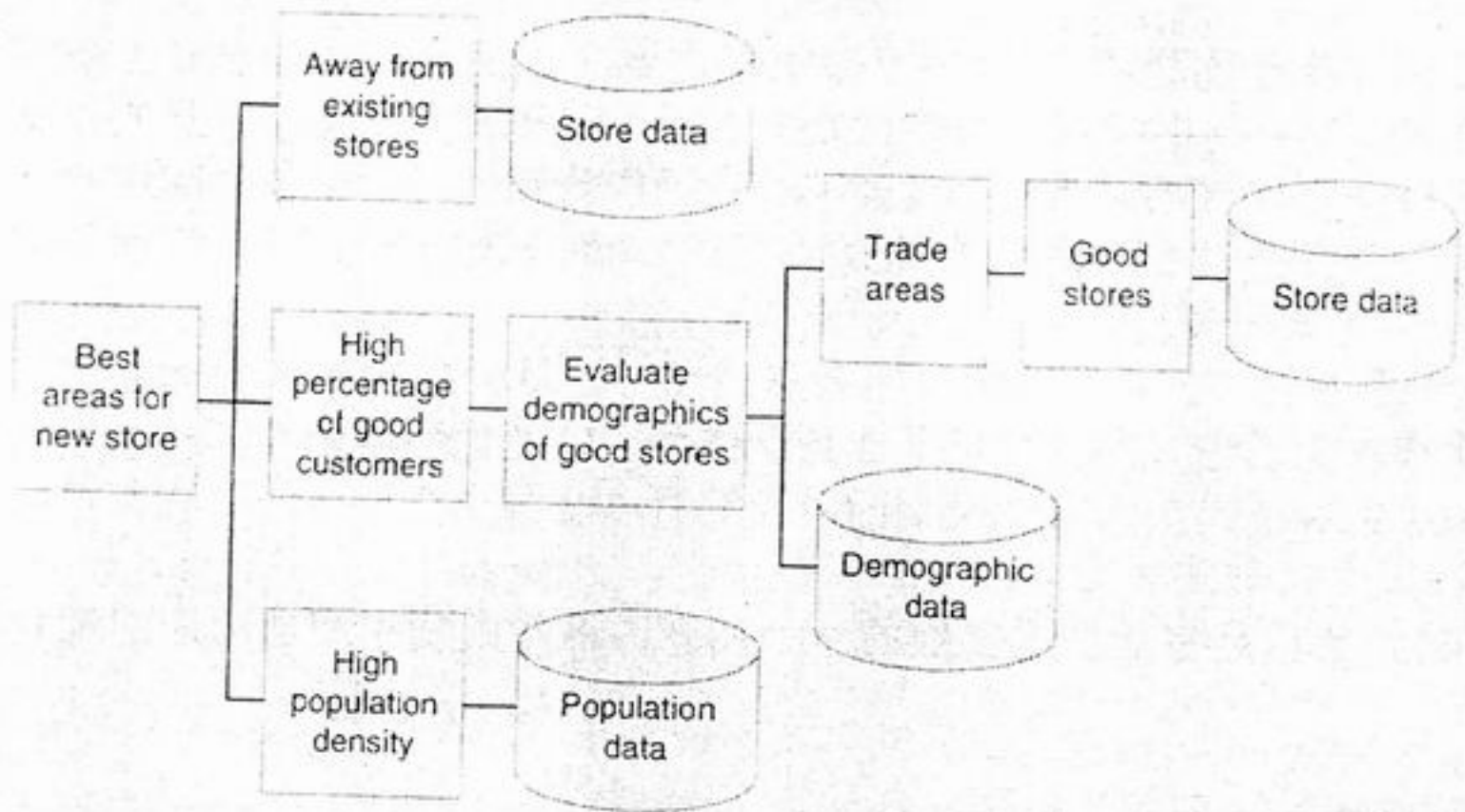


图 1-20. 进一步分解问题

当问题被进一步分解和图示时，示意图中的每一个分支的终端应该是一个数据源（图 1-20）。这些分支指定了目标度量所需的数据。

4 目标赋值

现在每个目标有一个图层，怎样将它们组合起来以创建一个潜在商店位置的单一分级的地图呢？您将需要一种不同类之间的值进行比较的方法。为此，可以对在每个图层或主题中的各类赋予数值。

回忆被提出的问题“为新商店选址”。每个目标根据在一个位置建商店的合适程度进行分级，给每个目标赋一个 1-10 之间的值，10 表示最合适，这通常被称为适宜性分级。

“没有数据(No Data)”或“不适宜”的地区通常在图中予以屏蔽而不被标出，对它们不予考虑。所有的度量因子应有同样的数值尺度，表示在决定最适宜地点时，它们的重要性相同。本模型开始就以此方式构建；当用于其它情况时，可以加入图层或度量因子的权重因素，以进一步地研究数据及其它它们之间的关系。

在对目标的适宜性分配级数的步骤中，您将需要更详细地定义度量目标和分配级数所需要的工具和步骤。

潜在顾客百分比分级

从分析您最好商店的顾客的统计特征中，您已经创建了一幅潜在顾客的百分比地图。所做的假定是如果找到一些地方，这些地方和您成功的商店附近居住着相类似的人群，这将会增加您的新商店成功的机会。因为这是一个抽象的概念化量算，因而您可以将最低百分比定为 1，最高百分比定为 10，然后线性分配中间级别的值。

有充足的客源吗？

为确保您的新商店有充足的客源，您需要一个人口密度图。“Calculate Density(计算密度)”功能将帮您创建一幅显示某一特定半径之内人口数量的地

图。从顾客调查和已有商店的经验中可知，很少有顾客愿意旅行 3 公里以上到您的商店购物。为模型化此变量，您需要创建一幅 3 公里半径以内的人口密度图。从已有商店可知，支持一个商店至少需要 1000 个人，任何人口密度低于 1000 的地点均被分级为 1，任何超过 1900 的地区最为合适，被标为 10。您可以使用的分级为：

0-1000=1, 1000-1400=5, 1400-1900=7, 1900-2600=10

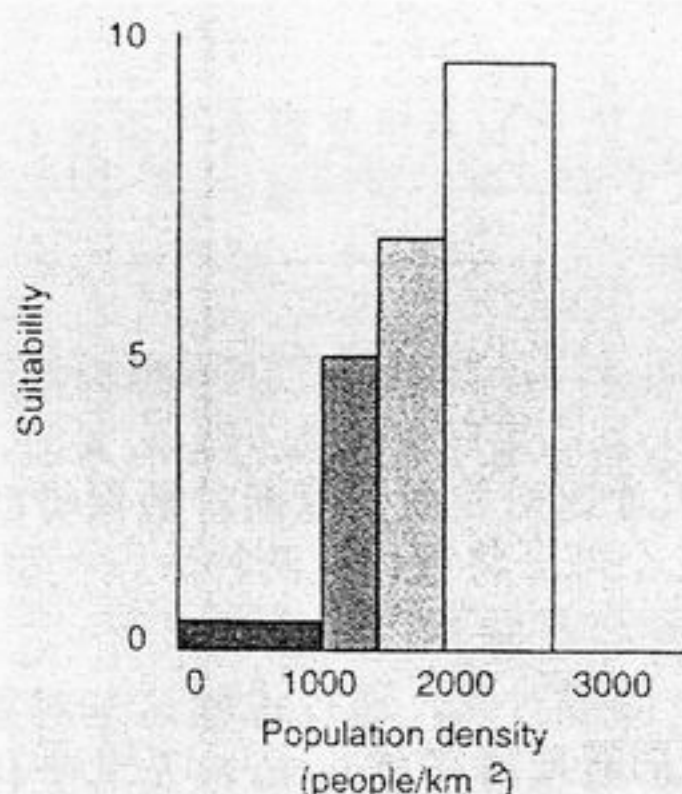


图 1-21. 潜在顾客百分比分级图

距离现有商店区域的分级

为避免和您其他的商店竞争顾客，您需要知道距现有商店的距离。使用“Find Distance(发现距离)”创建一个到商店的距离图。结果是一个栅格主题，图上每个单元表示了它到最近的商店的距离。您从调查中得知，多数顾客分布在商店的 3 公里以内，而所有顾客都在 5 公里以内。使用“Reclassify(重分类)”功能，将所有在现有商店 3 公里以内的地区标注为 1，因为它们最不合适；任何超过 6 公里的地点，标注为 10，这些地点不会造成现有商店的客源流失。您可以使用的分级为：

0-3000=1, 3000-4000=4, 4000-5000=5, 5000-6000=7, 大于 6000=10。(注意：在您将要创建的地图中的单位为“米”而不是“公里”)。

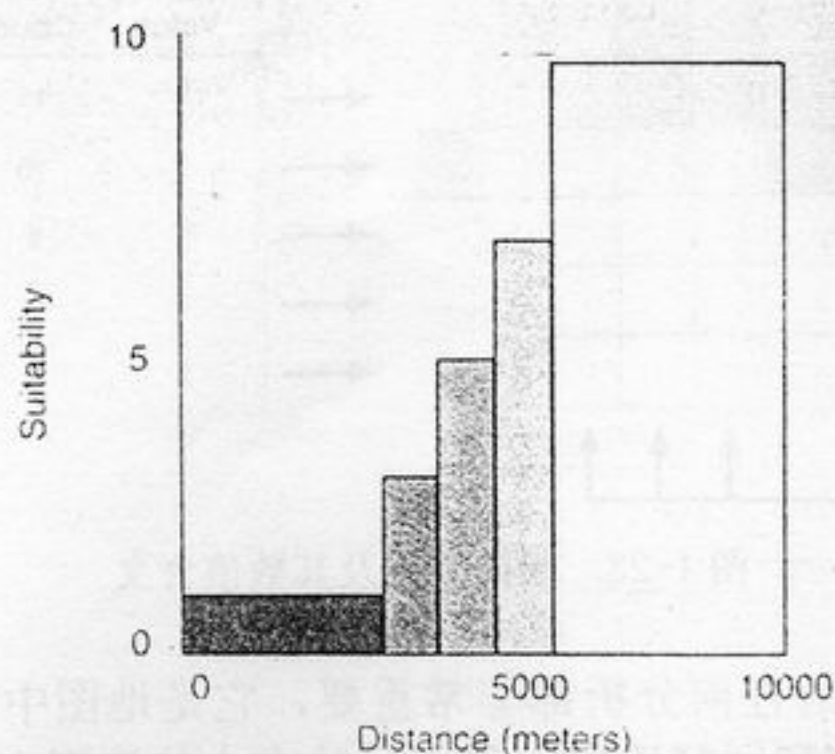


图 1-22. 到商店距离的分级图

5 创建合成适宜性评价图

建模的最后一步是将潜在顾客、人口、距离的适宜性地图组合成“新商店最佳选址”的单一度量因子。为此可将三幅地图相加再除以 3 即可。

在第 4 章中将讲解上述的模型中的每一步如何用空间分析模块的功能来解决。

6 GIS 数据及其考虑

要想理解在 GIS 中怎样解决上述的模型，了解 GIS 数据的一些基本概念是有益的。

在一个模型中，图层或主题的特征决定了其最佳用途。在过去的研究中常常不得不确定为使用栅格空间数据还是矢量空间数据而进行分析，而有了空间分析模块，您就可以同时处理矢量数据和栅格数据。必要时，两种数据可通过简单的一步处理实现转换。以下几个部分将探讨这两种数据类型及其它们的适用范围。

矢量数据及其应用

矢量数据使用点、线、面等对象，通过说明这些对象的地理位置和地理边界并且存储对象的信息来表示地理特征(feature)及其变化。矢量数据适宜用于：存储精确的位置；模型化线性网络之间的连通性，如：哪条公路连通纽约和芝加哥；或者处理准确的边界，例如两国之间的边界。

栅格数据及其应用

栅格数据通过将世界分隔成离散的方块(称为单元)来表示地理特征及其变化，每个单元中存储一个值。每个单元通过原点的位置以及单元位置和原点之间的相对关系隐含的表达其位置属性，而并不具体的存储其位置，存储的只是原点、单元的尺寸、距离原点的单元数等。影象(image)是最简单的栅格数据形式；在灰度影象中，每个单元仅存储该位置的一个单一值。空间分析模块中的栅格(grid)是栅格(raster)的一种特殊形式，其中被存储的值是表中的一个记录，该表存储了单元的更多附加描述信息(图 1-23)。

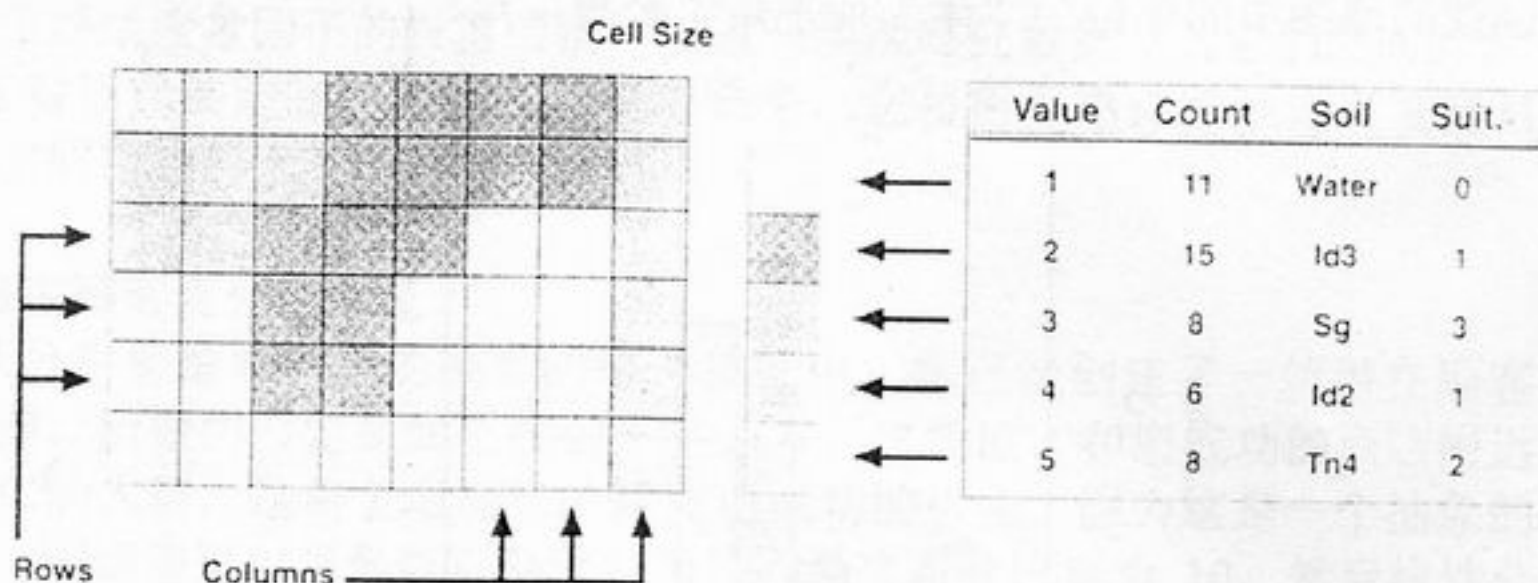


图 1-23. 栅格数据及其数值含义

所使用的单元尺寸对任何分析都非常重要，它是地图中您所感兴趣的最小单位并且限制了您分析的空间精确度。单元尺寸大小的选择由您分析结果要求的精

确度所决定。单元尺寸越小则结果越精确，但同时将产生更大的数据集和较慢的处理速度。

空间规则的正方形划分可以表现连续的现象，例如高度信息或到商店的距离信息等。它同时也提供了另外一种观察世界的方式。以往总是需要以“鸟瞰”的方式向下俯视，而您现在可以“定点透视”或者从下向上观察事物，这将允许您提出关于邻接或相邻的问题。栅格数据在存储和进行数据层的叠加(overlay)操作中十分有效。

整型和浮点栅格

空间分析模块可以在网格中存储整型和浮点型数据。如果您要绘制地图的变量是整型的，将创建一个整型栅格及其一个与之相关联的值属性表(VAT)。该表为栅格中的每个唯一值存储一个记录，并存储具有该值的单元数以及您愿意为其附加的任何属性。如果您要显示的变量是连续的小数，创建的将是浮点栅格，而没有 VAT 表。一般情况下，整型栅格比浮点栅格要小且操作起来速度要快。