A）标题：收入数据  
  
B）相关信息：  
  
营销数据库。  
  
资料来源：俄亥俄州哥伦布的影响资源公司（1987年）。  
总共有N = 9409份问卷，包含502个问题  
由旧金山湾区的购物中心顾客填写。  
  
数据集income.data是本调查的摘录。它包括  
14个人口属性。数据集是分类的好混合物  
以及大量缺失数据的连续变量。这是特点  
用于数据挖掘应用。  
  
C）目标是预测其他家庭的年收入13  
人口统计特征。  
  
D）属性信息  
   
 1家庭年收入（单身个人收入）  
             1.不到1万美元  
             2. 10,000美元到14,999美元  
             3. 15,000美元到19,999美元  
             4. 20,000到24,999美元  
             5. 25,000美元到29,999美元  
             6. $ 30,000至$ 39,999  
             7. $ 40,000至$ 49,999  
             8. 50,000美元至74,999美元  
             9. 75,000美元或更多  
  
  
 2 SEX  
             1.男  
             2.女  
   
 3婚姻状况  
             已婚  
             2.一起生活，不结婚  
             3.离婚或分居  
             4.丧偶  
             5.单身，从未结婚  
   
 4年龄  
             1. 14日至17日  
             2. 18日至24日  
             3. 25至34  
             4. 35到44  
             5. 45到54  
             6. 55到64  
             7. 65以上  
   
 5教育  
             1.八年级或以下  
             2. 9至11年级  
             3.高中毕业  
             4.大学1至3年  
             5.大学毕业生  
             6.毕业研究  
   
 6职业  
             1.专业/管理  
             2.销售人员  
             3.工厂工人/劳工/司机  
             4.文书/服务工作者  
             5.家庭主妇  
             6.学生，HS或大学  
             7.军事  
             8.退休  
             9.失业  
   
   
   
 7您在SAN FRAN./OAKLAND/SAN JOSE地区居住多久？  
             1.不到一年  
             2.一至三年  
             3.四到六年  
             4.七至十年  
             5.十多年  
   
 8双重收入（如果已婚）  
             1.未结婚  
             是的  
             3.没有  
   
 9户人家  
             1.一个  
             2.两个  
             三  
             4.四  
             5.五  
             6.六  
             七  
             8.八  
             9.九或更多  
   
 10名18岁以下的家庭成员  
             0.无  
             1.一个  
             2.两个  
             三  
             4.四  
             5.五  
             6.六  
             七  
             8.八  
             9.九或更多  
   
 11家庭状况  
             1.拥有  
             2.租金  
             3.与父母/家人同住  
   
 12种家庭  
             1.众议院  
             2.共管公寓  
             3.公寓  
             4.移动房屋  
             5.其他  
   
 13种族分类  
             美洲印第安人  
             2.亚洲人  
             3.黑色  
             4.东印度群岛  
             5.西班牙裔  
             太平洋岛民  
             7.白色  
             8.其他  
   
 14你家里最常见的语言是什么？  
             1.英语  
             2.西班牙语  
             3.其他  
   
  
E）实例数量：8993。  
  
   这些从9409个实例的原始数据集中获得，  
通过消除这些观察与回应（年收入）  
失踪。  
  
  
F）缺失值标志是NA。  
       皮克。 352。