1.某地经连续六年观察，发现其年度货运量与年度工业总产值和年度基建投资是线性相关的，6组统计数据如下表所示。现有该地下一年度的多组年度工业总产值和年度基建投资预测值（其中一组为39.71亿元和1.1025亿元），相关部门希望通过建立货运量预测决策支持系统，利用二元线性回归模型来帮助预测下一年度的年度货运量（百万吨）。

|  |  |  |  |
| --- | --- | --- | --- |
| **年度序号** | **年度货运量（百万吨）** | **年度工业总产值（亿元）** | **年度基建投资（亿元）** |
| 1 | 7.9 | 18.2 | 0.45 |
| 2 | 8.4 | 20.3 | 0.63 |
| 3 | 8.9 | 21.8 | 0.76 |
| 4 | 10.4 | 25.2 | 0.83 |
| 5 | 11.6 | 30.7 | 1.00 |
| 6 | 13.9 | 36.1 | 1.05 |

2. 二元线性回归模型的回归方程为：。其中，是因变量，和是自变量，a是方程的常数项，和是方程的回归系数（a，和统称为参数）。当已知因变量和自变量的多组数值、和时，可以根据下面给出的公式1、公式2和公式3（最小二乘法）分别计算出常数项a，回归系数和。而在获得常数项a，回归系数和的值后，给出一组自变量和的值，就可以根据回归方程计算出相应的因变量的值，即预测值。

 （公式1）

 （公式2）

 （公式3）

1. 分析：

1.1 决策问题

根据某地6年的年度货运量与年度工业总产值和年度基建投资数据，发现其存在二元的线性关系。数据如下所示：



图 1 六年的相关数据

相关部门期望能从这六年的数据中总结出规律，建立货运量预测决策支持系统，根据年度工业总产值和年度基建投资，预测下一年度的年度货运量。

并且，用于预测年度工业总产值和年度基建投资已给出，分别为39.71亿元和1.1025亿元。

1.2 决策目标

1.2.1建模：根据图1的六年相关数据，使用相关公式得到年度工业总产值、年度基建投资与年度货运量之间的二元线性关系：

年度货运量（y） = 常数（a）+b1\*年度工业总产值（x1）+b2\*年度基建投资（x2）。

1.2.2 预测：根据给出的年度工业总产值（x1）和年度基建投资（x2）来，预测年度货运量（y）。

2. 设计：

2.1 初步设计

2.1.1 问题分解

首先，可以将复杂的货运量预测问题分解为两个子问题：建立模型寻找年度货运量与年度工业总产值、年度基建投资之间的规律；使用得到的二元线性回归方程模型，依照输入的数据计算出预测的年度货运量。

2.1.2 问题综合

根据各个子问题之间的处理顺序进行综合设计：

图示

描述已自动生成

图 2 问题综合设计

首先计算出二元线性回归方程中的常数项a和回归系数b1、b2，得出二元线性回归方程；然后根据决策者输入的年度工业总产值数据和年度基建投资数据，计算并输出年度货运量Y的预测值。

2.1.3 子问题模型设计

2.1.3.1 寻找规律，建立模型

构建模型所需的公式如下所示：

文本, 信件

描述已自动生成

基于最小二乘法模型，使用公式1、2、3可以用来计算得到模型中的相关常量及系数，即计算出年度货运量与年度工业总产值和年度基建投资之间的规律。

解决问题1：根据已有的相关数据计算得到预测模型。

2.1.3.2 根据数据，输出预测



公式4根据第一个问题解决得到的a、b1、b2 建立二元线性回归方程。解决第二个问题：根据输入的年度工业总产值、年度基建投资，计算并输出年度货运量预测值。

2.1.4 子问题数据设计

在使用最小二乘法模型计算二元线性回归方程y=a+b1x1+b2x2时，需要使用图1 给出的6年的6组相关年度数据；在使用该模型进行预测年度货运量的时候，需要使用到决策者输入的年度工业总产值、年度基建投资。

2.2 数据库设计

使用ACCESS数据库创建货运量预测（hylyc）数据库

文本

描述已自动生成

图 3 货运量预测数据库

2.2.1 将使用最小二乘法模型计算二元线性回归方程y=a+b1x1+b2x2的常数项a和回归系数b1、b2时需要使用的6组统计数据放入货运量预测数据库（hylyc）的统计数据（tjsj）表中；

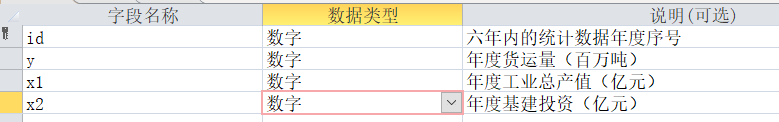


图 4 统计数据表结构

表格

描述已自动生成

图 5 字段为双精度型

图形用户界面, 应用程序, 表格

描述已自动生成

图 6 六年的统计数据

2.2.2 将得到的二元线性回归方程常数项a和回归系数b1、b2的值放入货运量预测数据库的参数（cs）表中；

图形用户界面, 应用程序

中度可信度描述已自动生成

图 7 参数表数据结构



图 8 参数表数据记录

2.2.3 将决策者输入的年度工业总产值和年度基建投资，以及利用二元线性回归方程模型计算得到的年度货运量放入货运量预测数据库的预测（yc）表中。

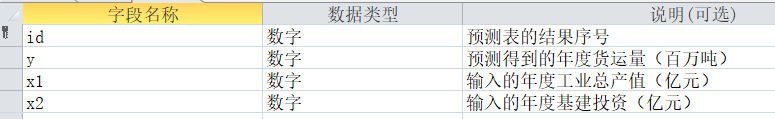


图 9 预测表数据结构



图 10 预测表数据记录

2.3 模型库设计

在货运量预测数据库中建立字典（zd）表，用于储存最小二乘法模型和二元线性回归方程模型的模型名称、模型文件名，以及模型文件路径。



图 11 字典表数据结构



图 12 字典表数据记录

2.4 运行结构图

图示, 示意图

描述已自动生成

运行结构图中展示了控制流和数据流，分别用实线和虚线表示。

2.4.1 控制流详细设计

在系统界面展现问题相关介绍，开始使用系统预测年度货运量；

系统从模型库调用最小二乘法模型；

系统显示常数项a和回归系数b1、b2的值

决策者输入年度工业总产值x1、年度基建投资x2的值；

系统调用二元线性回归方程模型；

系统显示年度货运量y的值；

决策者选择是否预测；如果是，则系统为其展示常数项a和回归系数b1、b2的值；如果否，则退出系统。

2.4.2 数据流详细设计

数据流1表示六年的统计数据被最小二乘法模型调用；

数据流2表示由最小二乘法模型计算出的常数项a和回归系数bl、b2被放入参数（cs）表中；

数据流3表示常数项a和回归系数b1、b2的值被总控程序调用，在窗体中显示；

数据流4表示输入的年度工业总产值x1和年度基建投资x2被存入预测（yc）表中；

数据流5表示年度工业总产值x1和年度基建投资x2的值被二元线性回归方程模型调用；

数据流6表示常数项a和回归系数b1、b2的值传入二元线性回归方程模型中来；

数据流7表示由二元线性回归方程模型计算出的年度货运量被存入预测（yc）表中；

数据流8表示年度货运量y被总控程序调用，在窗体中显示。

**【实验方案设计】**

**实验步骤**：

1. 根据实验要求，严格遵循DSS开发步骤，首先进行DSS的分析和设计。
2. 根据分析和设计结果编制程序，实现符合实验要求的DSS。
3. 演示运行系统。
4. 编写实验报告书。

**【实验过程】（实验步骤、记录、数据、分析）**

**实验数据：**（系统分析、设计说明）

1. 分析：决策问题（1分）和决策目标（1分）

1.1 决策问题

根据某地6年的年度货运量与年度工业总产值和年度基建投资数据，发现其存在二元的线性关系。数据如下所示：



图 1 六年的相关数据

相关部门期望能从这六年的数据中总结出规律，建立货运量预测决策支持系统，根据年度工业总产值和年度基建投资，预测下一年度的年度货运量。

并且，用于预测年度工业总产值和年度基建投资已给出，分别为39.71亿元和1.1025亿元。

1.2 决策目标

1.2.1建模：根据图1的六年相关数据，使用相关公式得到年度工业总产值、年度基建投资与年度货运量之间的二元线性关系：

年度货运量（y） = 常数（a）+b1\*年度工业总产值（x1）+b2\*年度基建投资（x2）。

1.2.2 预测：根据给出的年度工业总产值（x1）和年度基建投资（x2）来，预测年度货运量（y）。

2. 设计：初步设计（4分），数据库设计（4分），模型库设计（2分），运行结构图（8分）

2.1 初步设计

2.1.1 问题分解

首先，可以将复杂的货运量预测问题分解为两个子问题：建立模型寻找年度货运量与年度工业总产值、年度基建投资之间的规律；使用得到的二元线性回归方程模型，依照输入的数据计算出预测的年度货运量。

2.1.2 问题综合

根据各个子问题之间的处理顺序进行综合设计：

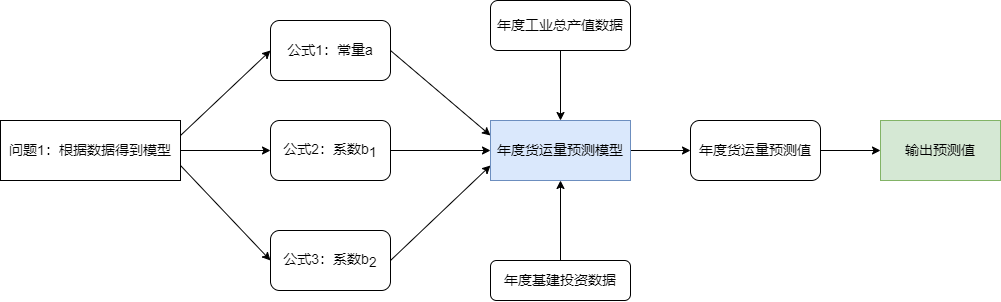


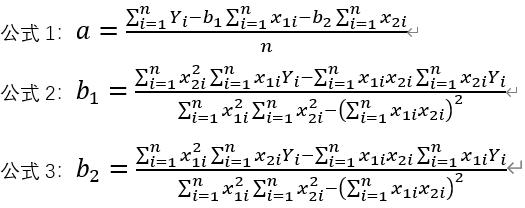
图 2 问题综合设计

首先计算出二元线性回归方程中的常数项a和回归系数b1、b2，得出二元线性回归方程；然后根据决策者输入的年度工业总产值数据和年度基建投资数据，计算并输出年度货运量Y的预测值。

2.1.3 子问题模型设计

2.1.3.1 寻找规律，建立模型

构建模型所需的公式如下所示：



基于最小二乘法模型，使用公式1、2、3可以用来计算得到模型中的相关常量及系数，即计算出年度货运量与年度工业总产值和年度基建投资之间的规律。

解决问题1：根据已有的相关数据计算得到预测模型。

2.1.3.2 根据数据，输出预测



公式4根据第一个问题解决得到的a、b1、b2 建立二元线性回归方程。解决第二个问题：根据输入的年度工业总产值、年度基建投资，计算并输出年度货运量预测值。

2.1.4 子问题数据设计

在使用最小二乘法模型计算二元线性回归方程y=a+b1x1+b2x2时，需要使用图1 给出的6年的6组相关年度数据；在使用该模型进行预测年度货运量的时候，需要使用到决策者输入的年度工业总产值、年度基建投资。

2.2 数据库设计

使用ACCESS数据库创建货运量预测（hylyc）数据库



图 3 货运量预测数据库

2.2.1 将使用最小二乘法模型计算二元线性回归方程y=a+b1x1+b2x2的常数项a和回归系数b1、b2时需要使用的6组统计数据放入货运量预测数据库（hylyc）的统计数据（tjsj）表中；

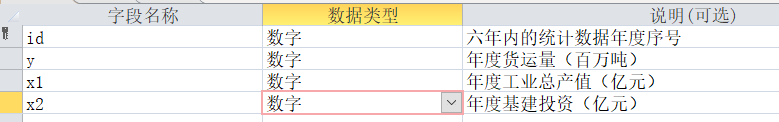


图 4 统计数据表结构

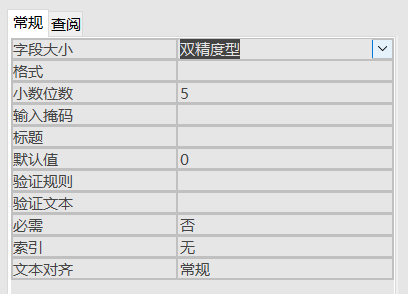


图 5 字段为双精度型

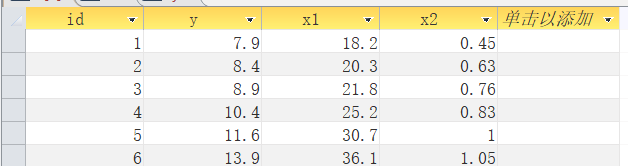


图 6 六年的统计数据

2.2.2 将得到的二元线性回归方程常数项a和回归系数b1、b2的值放入货运量预测数据库的参数（cs）表中；

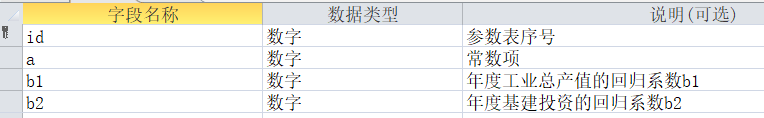


图 7 参数表数据结构



图 8 参数表数据记录

2.2.3 将决策者输入的年度工业总产值和年度基建投资，以及利用二元线性回归方程模型计算得到的年度货运量放入货运量预测数据库的预测（yc）表中。

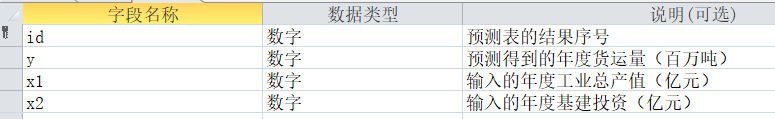


图 9 预测表数据结构



图 10 预测表数据记录

2.3 模型库设计

在货运量预测数据库中建立字典（zd）表，用于储存最小二乘法模型和二元线性回归方程模型的模型名称、模型文件名，以及模型文件路径。

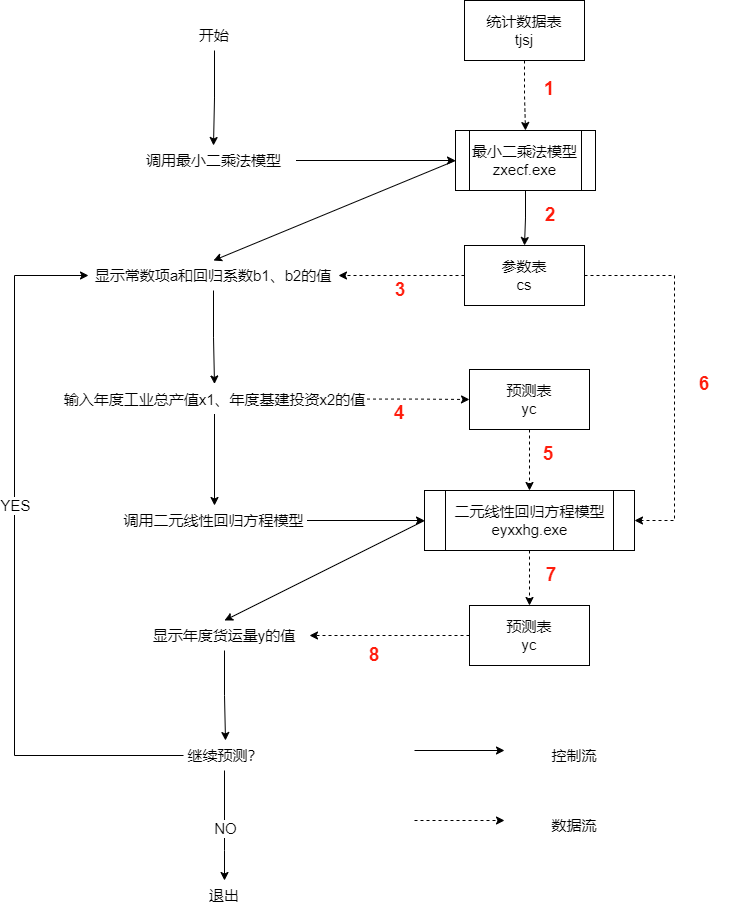


图 11 字典表数据结构



图 12 字典表数据记录

2.4 运行结构图



运行结构图中展示了控制流和数据流，分别用实线和虚线表示。

2.4.1 控制流详细设计

在系统界面展现问题相关介绍，开始使用系统预测年度货运量；

系统从模型库调用最小二乘法模型；

系统显示常数项a和回归系数b1、b2的值

决策者输入年度工业总产值x1、年度基建投资x2的值；

系统调用二元线性回归方程模型；

系统显示年度货运量y的值；

决策者选择是否预测；如果是，则系统为其展示常数项a和回归系数b1、b2的值；如果否，则退出系统。

2.4.2 数据流详细设计

数据流1表示六年的统计数据被最小二乘法模型调用；

数据流2表示由最小二乘法模型计算出的常数项a和回归系数bl、b2被放入参数（cs）表中；

数据流3表示常数项a和回归系数b1、b2的值被总控程序调用，在窗体中显示；

数据流4表示输入的年度工业总产值x1和年度基建投资x2被存入预测（yc）表中；

数据流5表示年度工业总产值x1和年度基建投资x2的值被二元线性回归方程模型调用；

数据流6表示常数项a和回归系数b1、b2的值传入二元线性回归方程模型中来；

数据流7表示由二元线性回归方程模型计算出的年度货运量被存入预测（yc）表中；

数据流8表示年度货运量y被总控程序调用，在窗体中显示。

3.1 计算常数项、系数、预测值

3.1 常数项a ，回归系数b1 、b2的调用计算

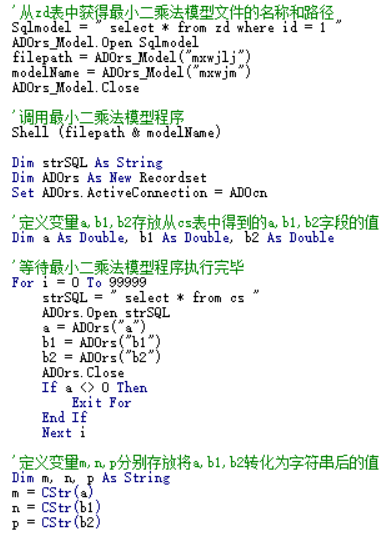


图 2 计算相关系数

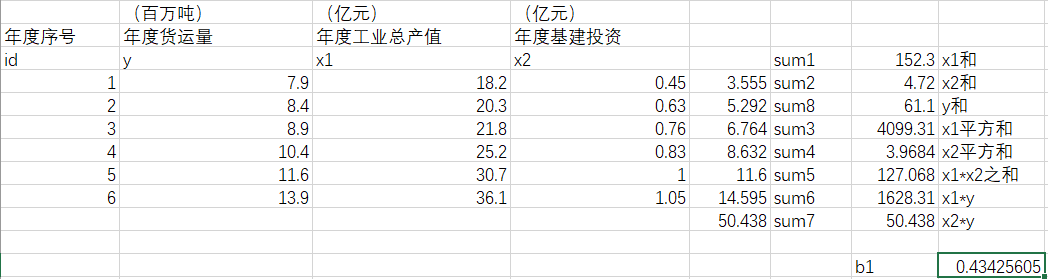


图 3 人工计算回归系数用于验证



图 4 验证得到的回归系数

3.2 人机交互设计

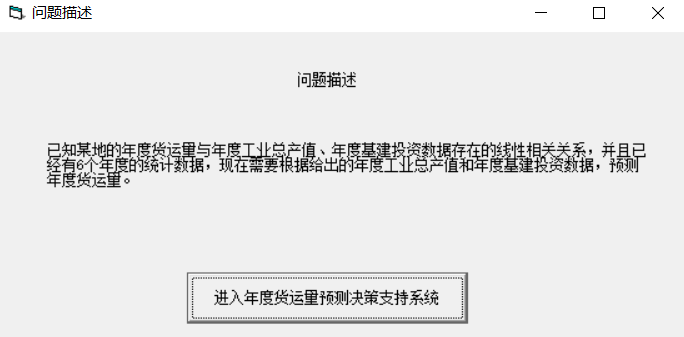


图 5 问题描述



图 17 总控页面

4. 根据意见后的相关改进

4.1 改进人机交互，删除显示参数按钮

将显示参数项的Button删除，在form\_load块中增加代码确保参数能正常显示。

图 6 删除显示参数的按钮

4.2在未预测前，将继续预测的Button的enable属性设置为false

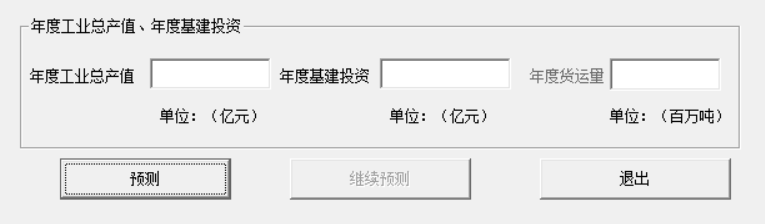
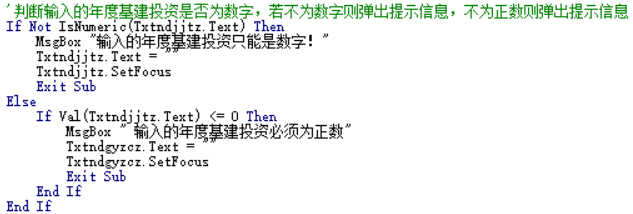
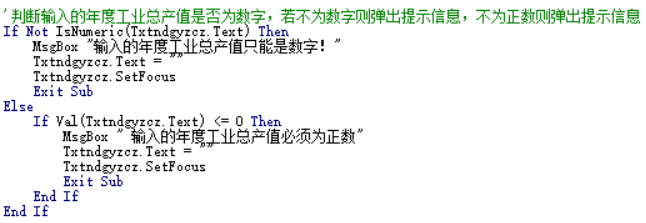


图 7 继续预测按钮不可用

4.3增加对输入的年度工业总产值和年度基建投资的检验，检验其是否为负数



图 8 年度工业总产值为负数



4.4点击继续预测后，年度货运量为空

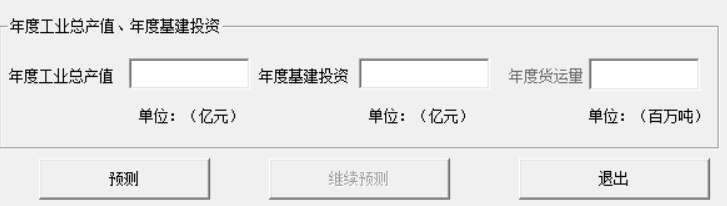


图 9点击继续预测后清除数据

