电子科技大学信息与软件工程学院

标 准 实 验 报 告

**（实验）课程名称 计量经济学**

**电子科技大学教务处制表**

**电 子 科 技 大 学**

**实 验 报 告**

**学生姓名：** **\*\*\*** **学 号：202109120\*\*\*\*** **指导教师：李亚静**

**实验地点： 信软楼西303**  **实验学时**：**2学时**

**一、实验室名称：软件工程实验室**

**二、实验项目名称：基于最小二乘法对身高体重进行简单回归**

**三、实验原理：**

**（一）普通最小二乘法**

普通最小二乘法（OLS）通过最小化残差平方和来估计回归方程中的参数，是计量经济学中最经典、应用最广泛的估计方法。该方法的使用需满足一系列古典假设，包括：

1. 回归模型是线性的，模型设定无误且含有误差项；
2. 误差项总体均值为零；
3. 所有解释变量与误差项都不相关；
4. 误差项观测值互不相关（不存在序列相关性）；
5. 误差项具有同方差(不存在异方差性)；
6. 任何一个解释变量都不是其他解释变量的完全线性函数（不存在完全多重共线性）

在上述古典假设成立的情况下，普通最小二乘估计量是最佳线性无偏估计量。所谓无偏估计量，是指OLS参数估计值的均值等于总体参数的真实值。

**四、实验目的：**

掌握使用EViews软件进行普通最小二乘估计；加深对普通最小二乘估计方法的认识。

**五、实验内容：**

在自行收集相关人员身高、体重、性别数据的基础上，采用EViews软件对收集到的样本数据进行分析、查看。同时通过最小二乘法估计被解释变量WEIGHT和解释变量HEIGHT之间的关系，并解释各个回归参数的大小和意义，最后将MALE变量也纳入考量，探究性别的影响。

通过最小二乘法估计得到的回归参数可以告诉我们解释变量对被解释变量的影响程度和方向。例如，身高（HEIGHT）的回归参数表示单位身高变化对体重（WEIGHT）的平均影响，正值表示两者正相关，负值表示负相关。在实验中将对各个回归参数的意义进行详细的分析。

同时，探究解释调整的判定系数的实际意义。

通过实验，掌握使用EViews软件进行样本数据的导入、样本数据描述的查看、普通最小二乘估计等基本操作，学会EViews的基本命令和使用方法，并加深对普通最小二乘估计方法的认识和理解。

**六、实验器材（设备、元器件）：**

计算机、EViews软件

**七、实验步骤：**

1. **收集相关数据**：记录五个同性别的身高（单位英寸），以及体重（单位是磅），同时报告实验者本人的身高和体重，数据要求为：
   1. HEIGHT是为5英尺以上的身高变量
   2. WEIGHT是以磅为单位的体重变量
   3. MALE是一个虚拟变量，性别男性用1表示，女性用0表示。
   4. 在Excel表录入全部的6条数据
2. **利用Eviews计算统计数据：**利用Eviews软件对收集到的数据进行整理、查看。具体内容包括其
3. 最小值、最大值
4. 样本大小
5. 样本变化趋势
6. **使用Eviews进行回归拟合：**利用最小二乘法进行回归拟合，解释WEIGHT和HEIGHT之间的关系，确定各项回归参数的值和意义，并在电脑上进行输出展示。具体步骤包括——
   1. 导入数据
   2. 输入指令“LS WEIGHT C HEIGHT”
   3. 按下回车，结果显示在屏幕上
7. **解释估计系数：**基于Eviews计算得出的回归系数，解释每个回归系数的意义，并于预期的值和符号进行比较。具体包括——
   1. 回归系数估计值的符号表示了解释变量对被解释变量的的正负影响关系。
   2. 回归系数估计值的大小表示了单位变化对被被解释变量的影响程度大小
8. **解释调整的判定系数：**在Eviews中的调整结果中找到调整的判定系数，解释他的意义，评判回归方程的拟合效果
9. **新增一个参数，再估计一个方程：**将性别MALE作为第二个解释变量加入到方程中，并重复步骤2-4，基于Eviews的计算结果解释“性别和体重是否没有关系”这个论断。

**八、实验数据及结果分析：**

1. 数据采集

首先，我们采集了身边同性别的五位男同学和自己的身高体重数据。我们将采集到的数据汇总到表格，可得到如下数据——

表 1：6名同学的身高体重数据

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 身高（英寸） | 体重（磅） | 性别 |
| 1 | 4 | 126 | 1 |
| 2 | 9 | 154 | 1 |
| 3 | 6 | 131 | 1 |
| 4 | 2 | 120 | 1 |
| 5 | 6 | 116 | 1 |
| 6 | 3 | 118 | 1 |

其中，身高（HEIGHT）是为5英尺以上的身高变量，体重(WEIGHT)是以磅为单位的体重变量，性别（MALE）是一个虚拟变量，性别男性用1表示，女性用0表示。

1. 数据导入

我们将数据集导入到Excel表格中，导入结果如图 1所示

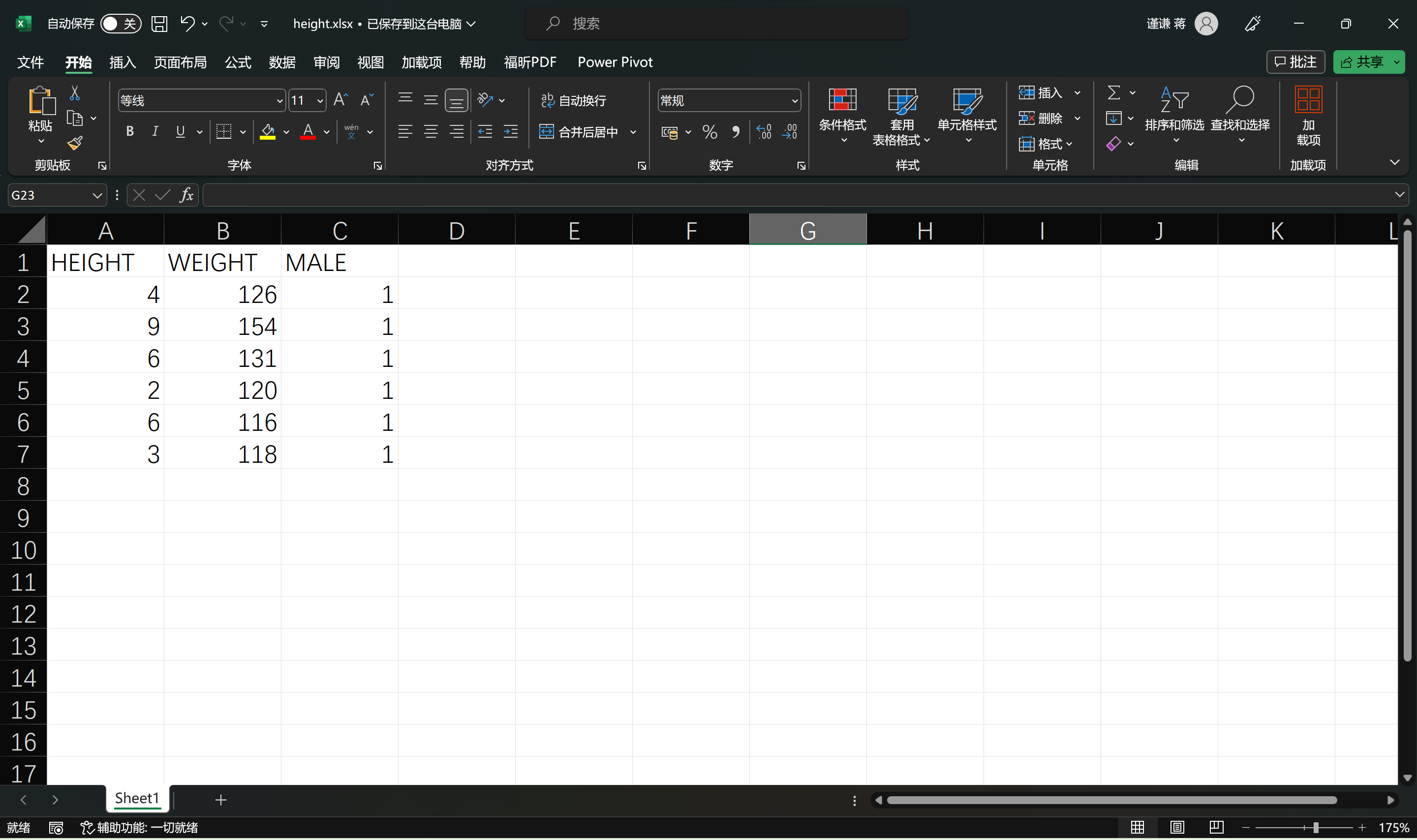


图 1：数据导入表格

随后，我们打开Eviews软件，依次选择“File->Import ->Import From File”将Excel表格数据导入Eviews软件中。

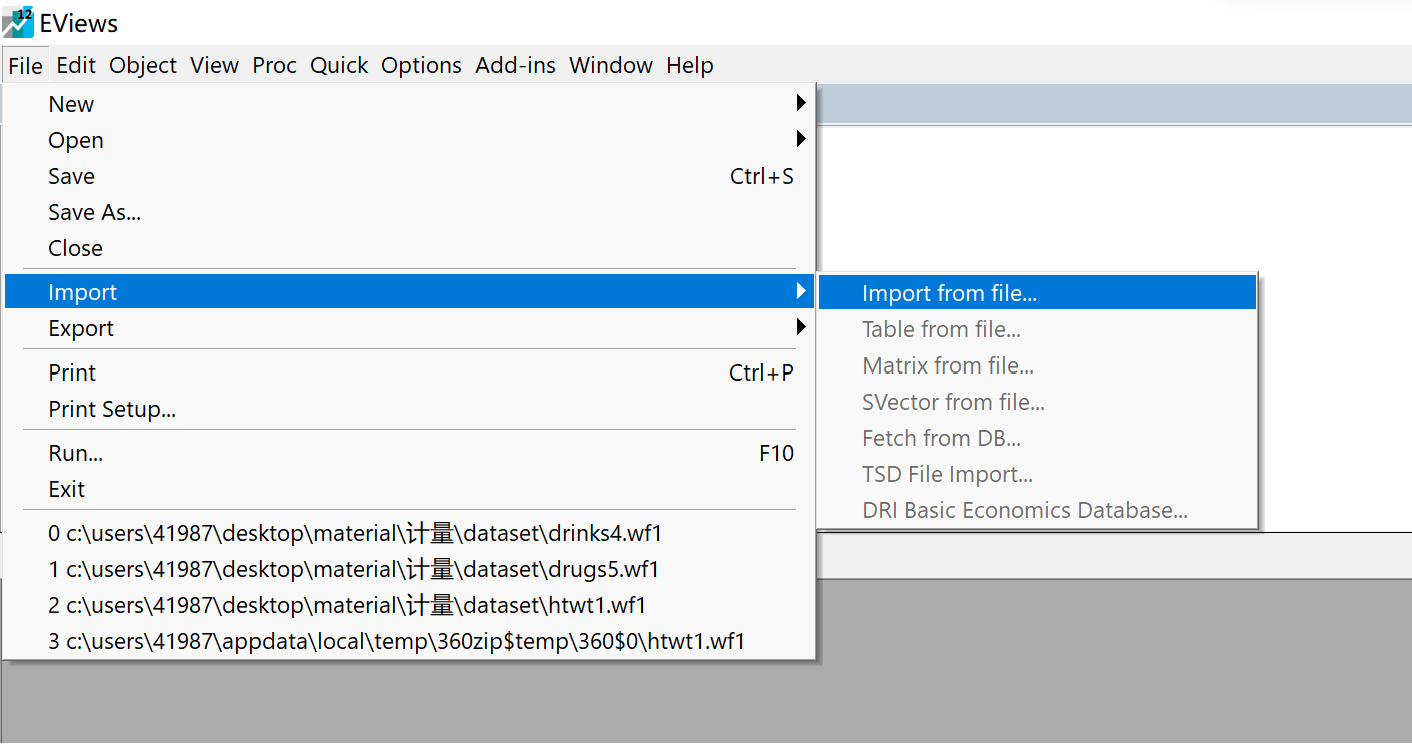


图 2: 将数据导入表格

紧接着我们为导入的数据设定数据类型等相关内容。这几类数据均为数值型数据，且无特殊结构，相关选择如图 3所示。

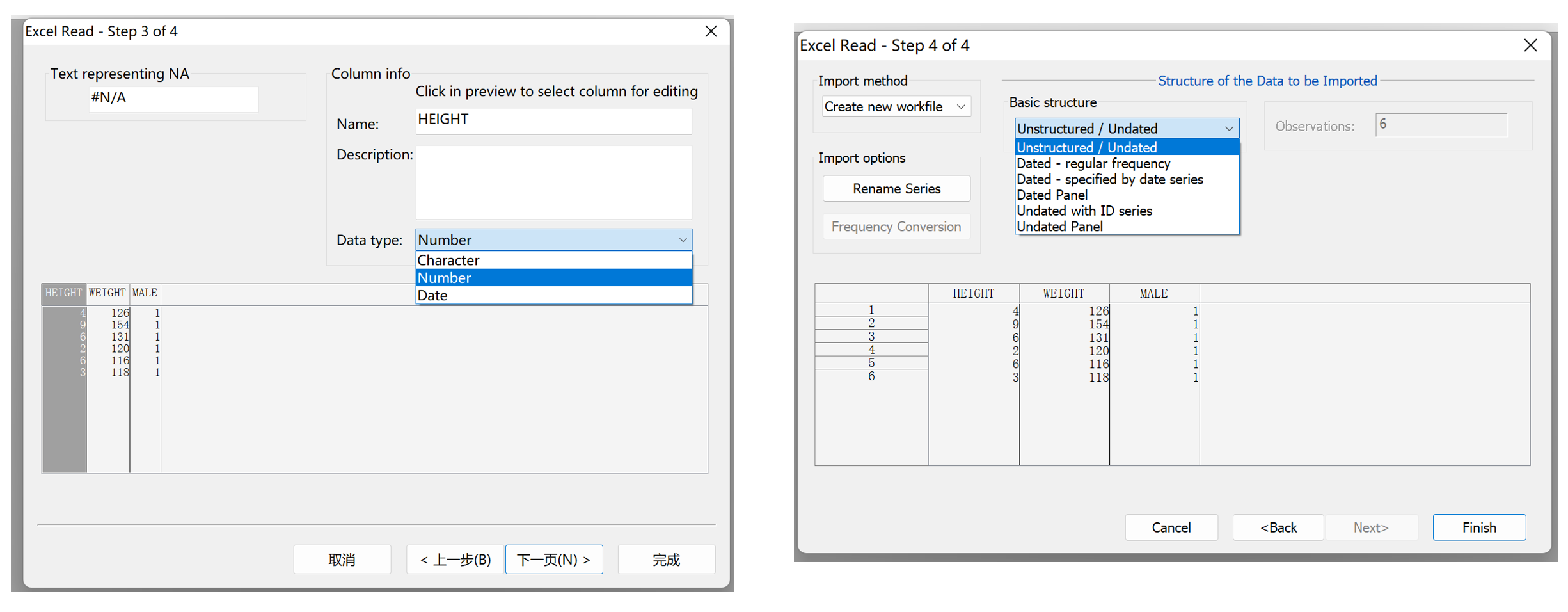


图 3：数据类型选择

点击“Finish”后即表示导入成功完成。

1. 数据特征查看

接下来，我们对导入的数据进行相关数据特性的查看

* **查看导入的全部样本数据**

首先将要查阅的数据选中（包括WEIGHT、HEIGHT），右击选择“OPEN->as GROUP”，即可展示导入的完整数据。

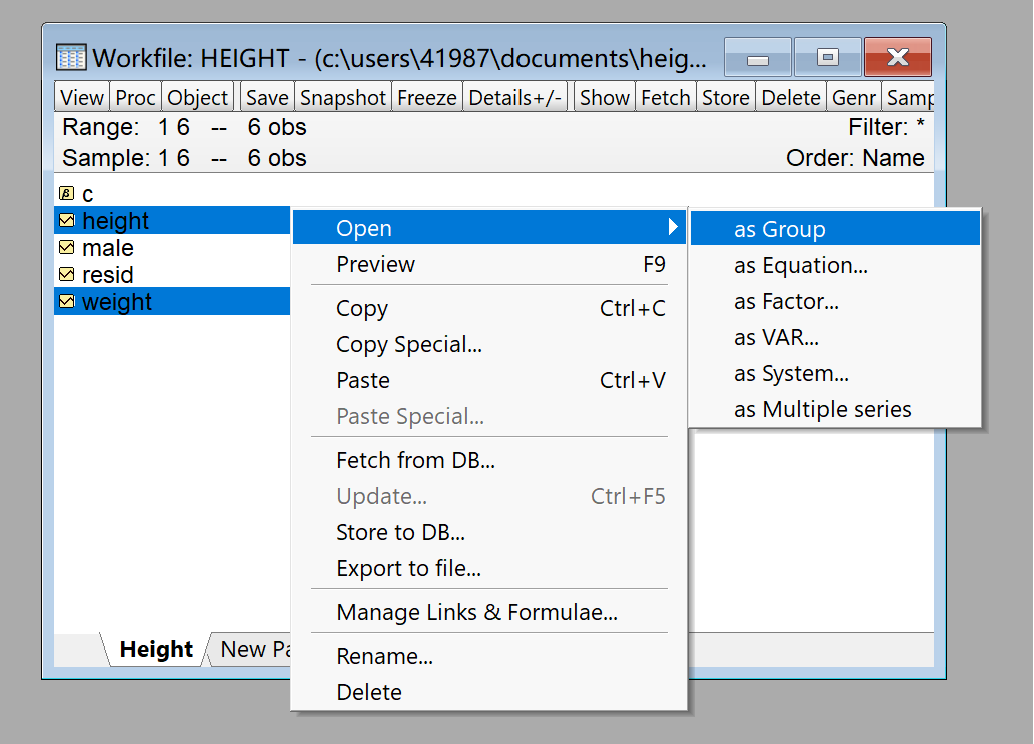


图 4：打开数据展示详情

点击后即可得到如图 5所示的数据列表——

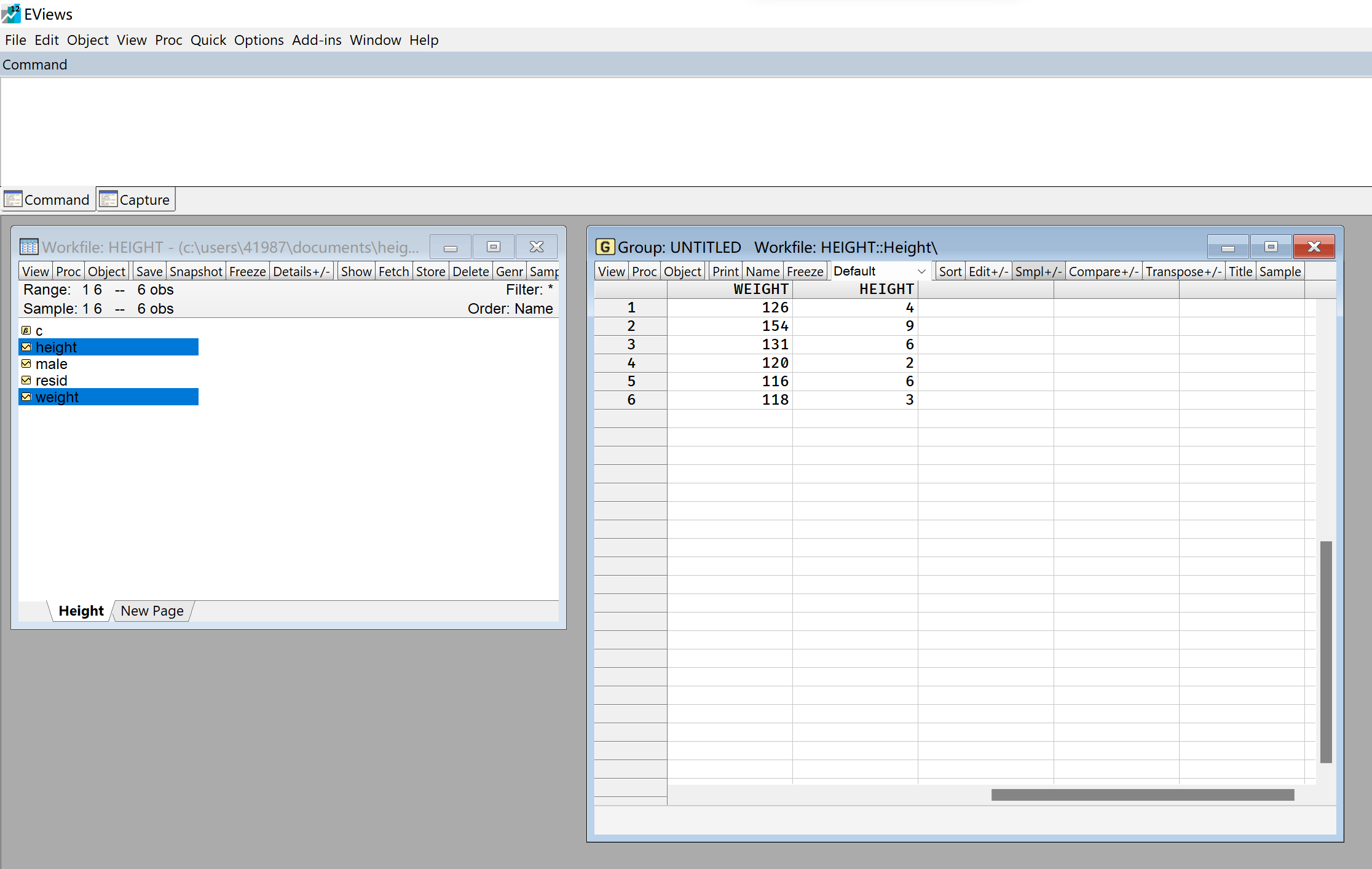


图 5：查看导入的数据

可以看到左上方写明了“Sample……. 6 obs”即表明样本大小为6个，即我们**一共导入了六个样本数据**。验证正确

* **查看样本数据的最大值、最小值**

在Group视窗中，选择“View->Descriptive Stats->Common Sample”即可查看数据的详细特征（包括最值、标准差、方差等），相关步骤截图如图 6所示

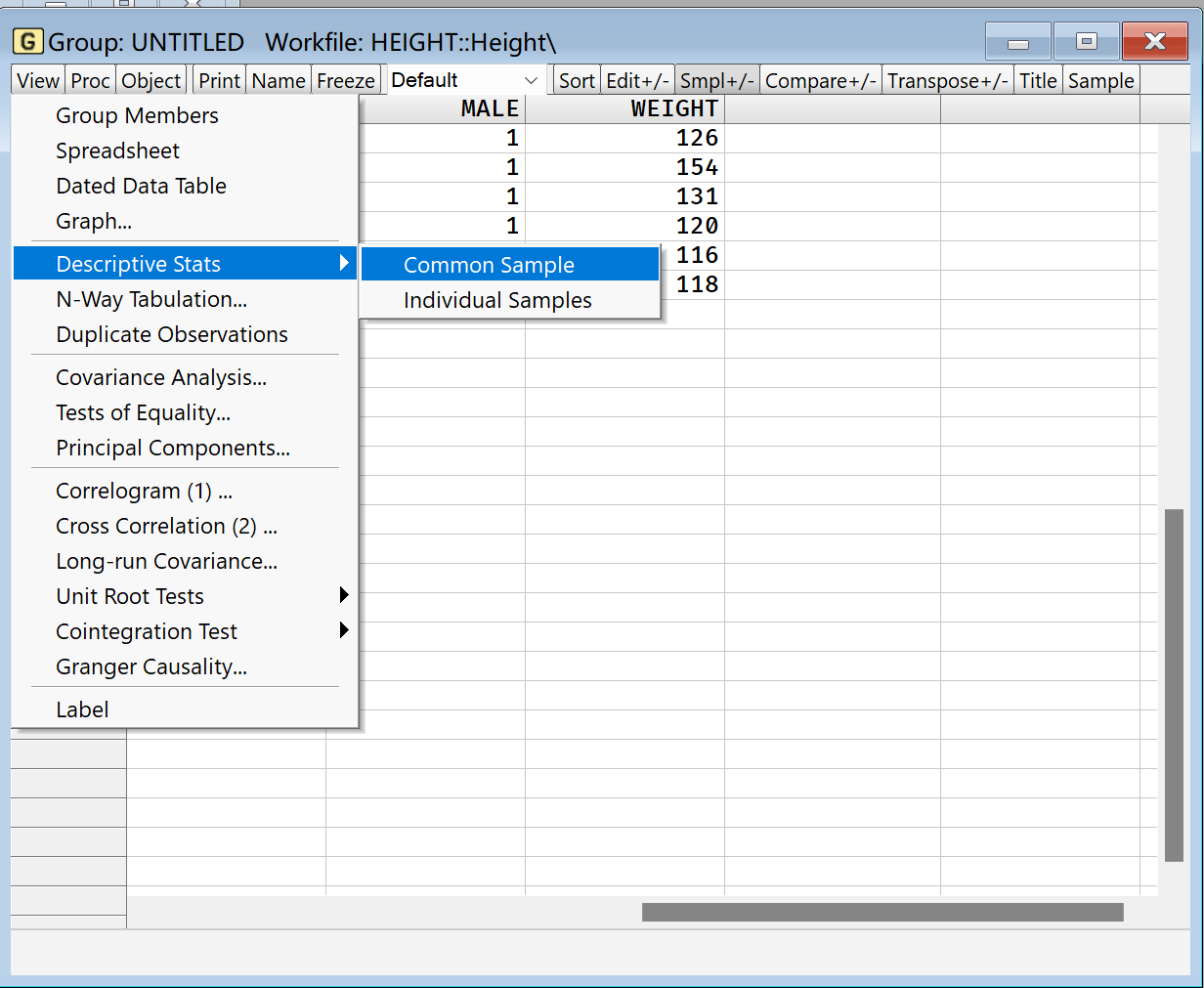


图 6：查看样本数据的特征步骤

得到的数据特征如下——

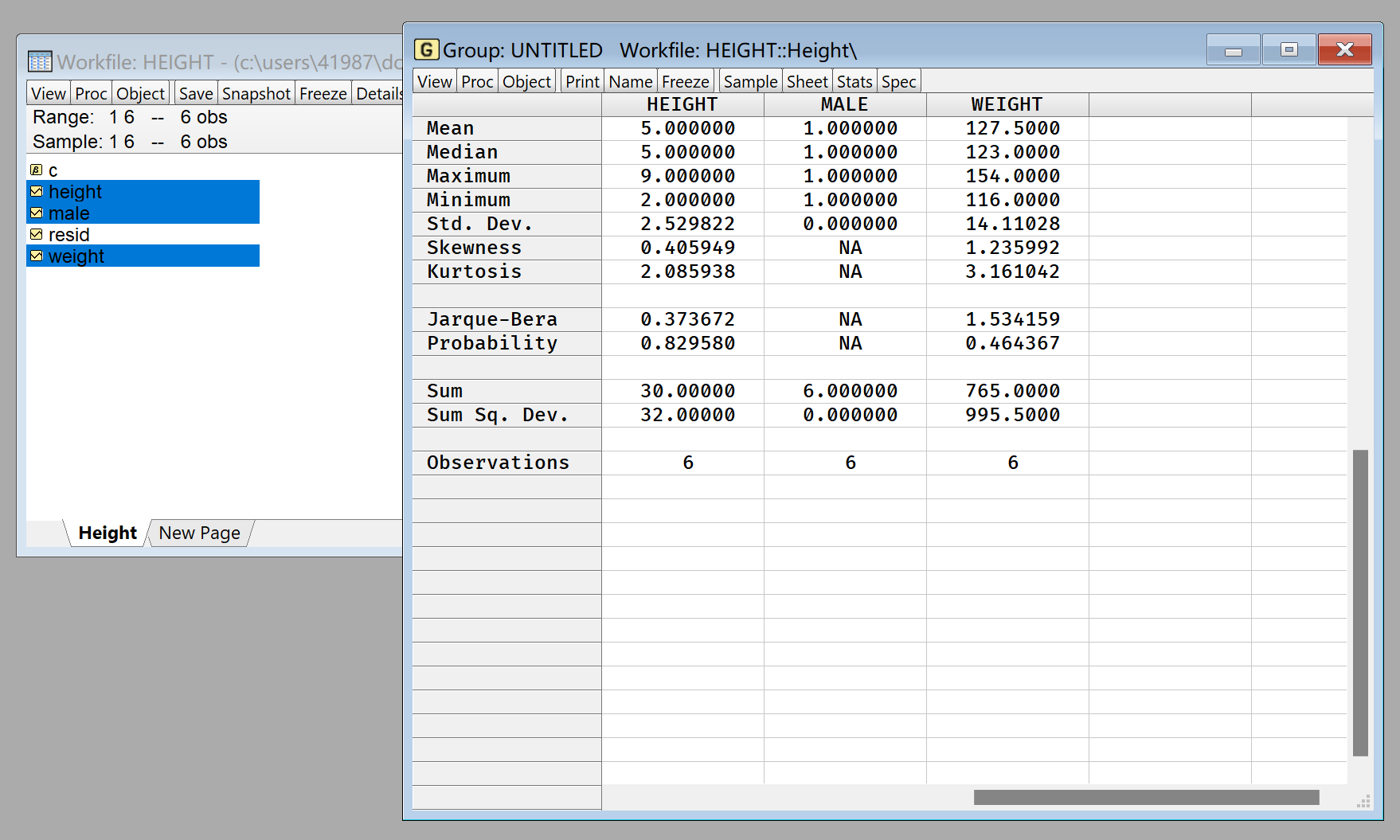


图 7：样本数据的特征

可以看到**HEIGHT的最大值为9，最小值为2，标准差2.53；WEIGHT的最大值为154，最小值为116，标准差为14。**数据分布和导入一致，验证正确，且没有异常情况发生。

* **查看样本数据的趋势分布**

我们以“离散点图”的形式展示WEIGHT和HEIGHT 的趋势分布。在GROUP视窗中，点击Group Options，选择散点图（Scatter）并添加辅助的回归线（Regression Line），相关步骤截图见图 8

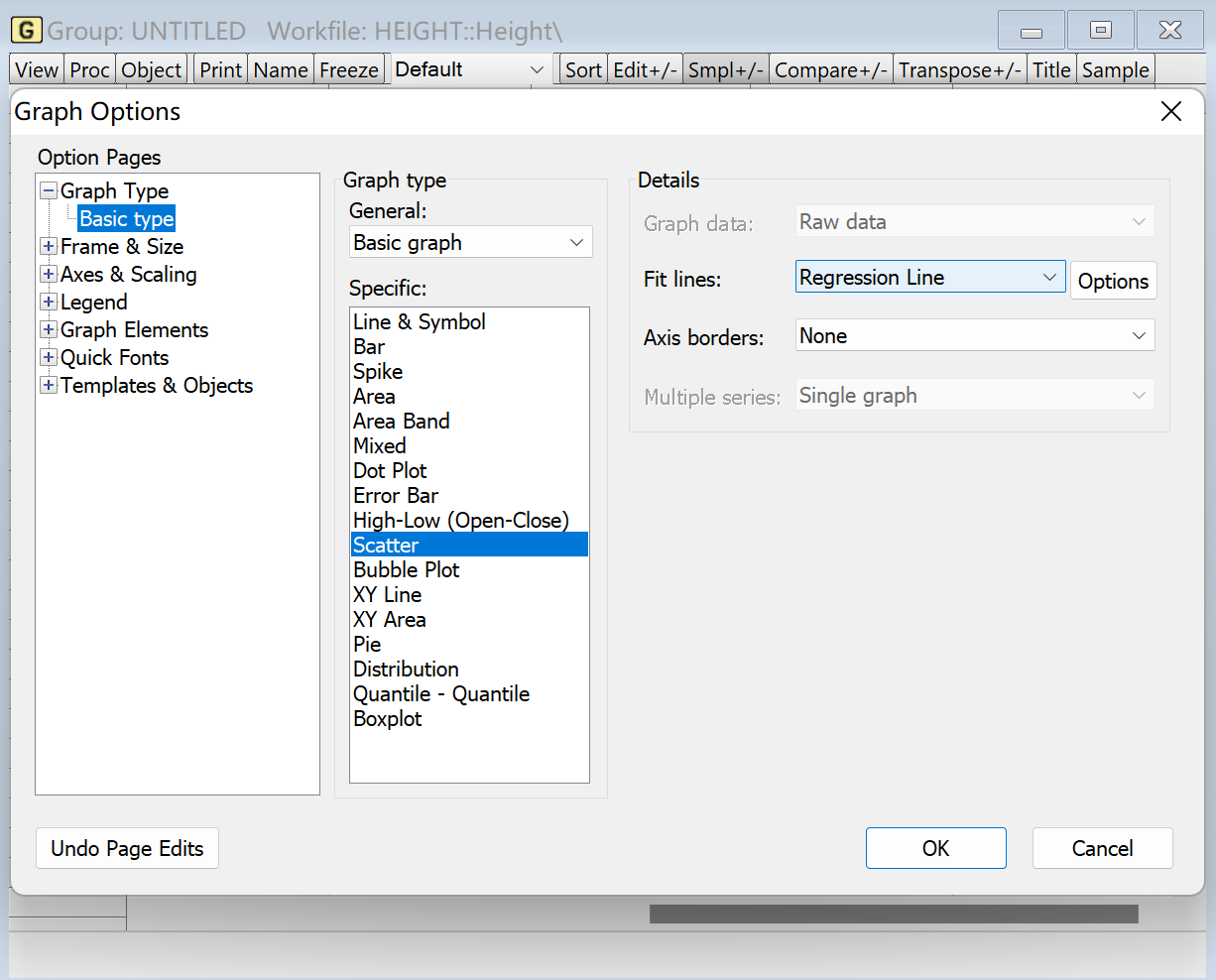


图 8：绘制散点图

绘制的散点图结果如下——

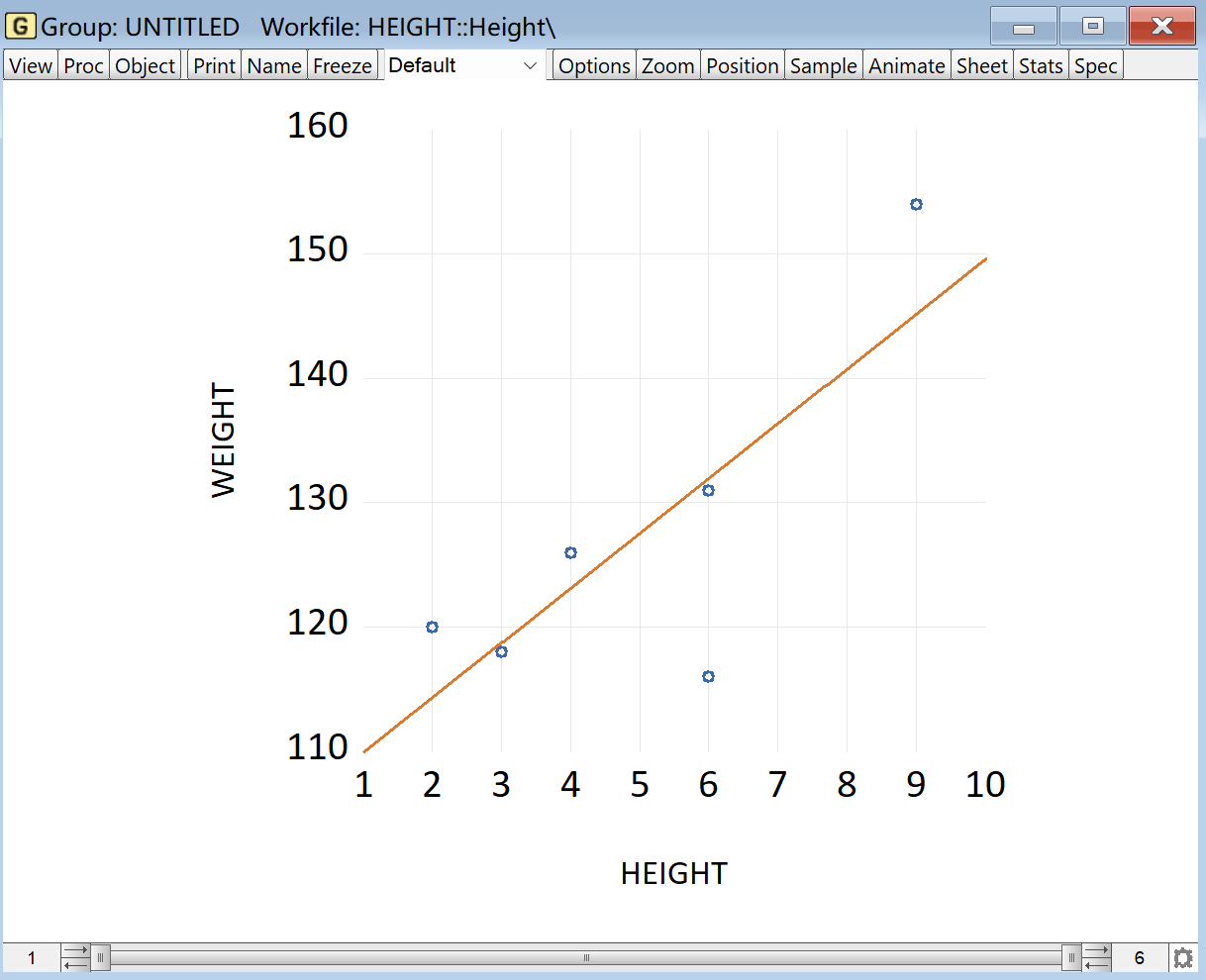


图 9：散点图的分布趋势

从图中可以看出，其大致趋势为：随着身高的增加体重也随之增长。符合我们的预期。

1. 回归拟合

接下来利用Eviews进行回归拟合

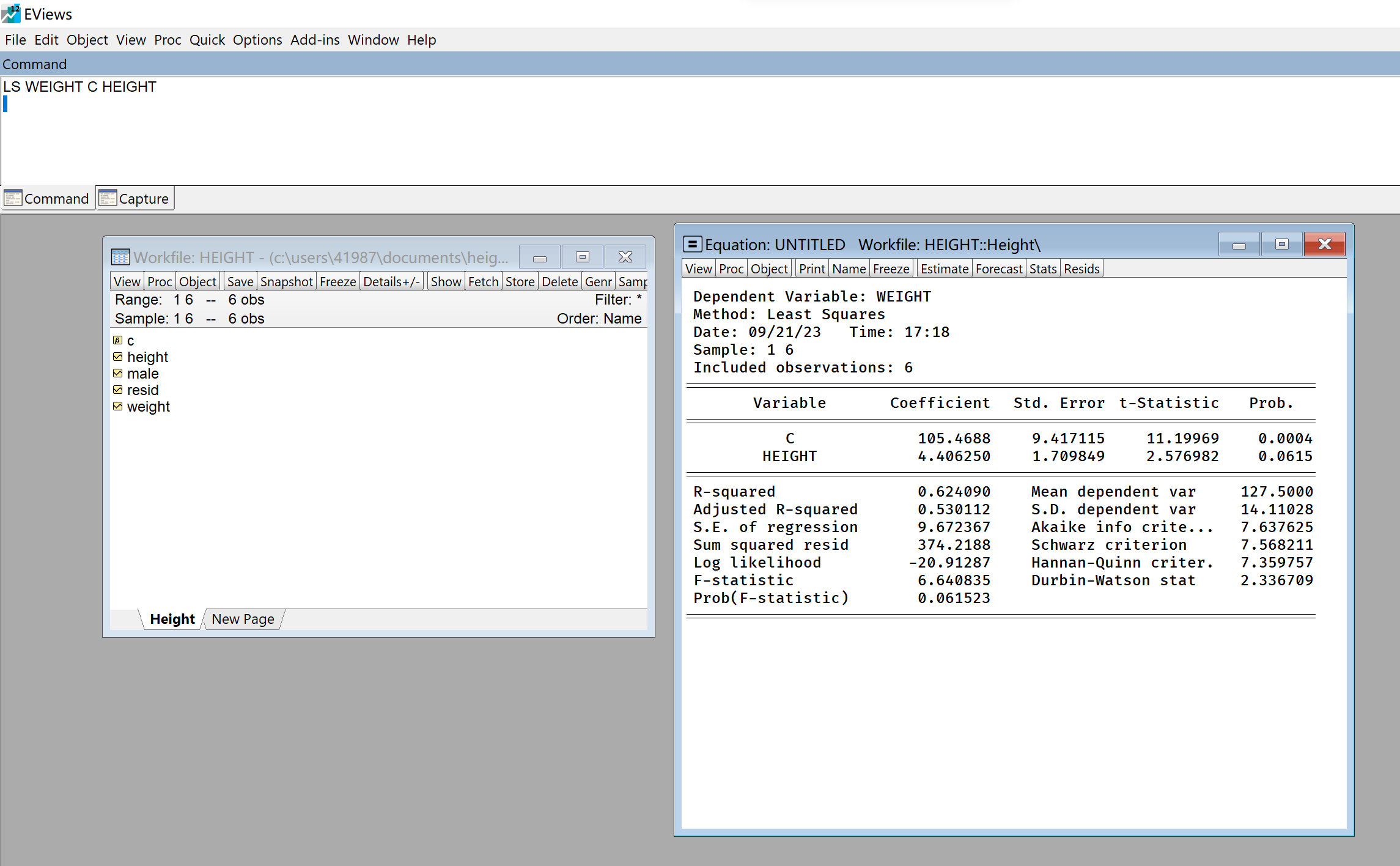


图 10：输入指令进行拟合

如图 10所示，我们输入指令“**LS WEIGHT C HEIGHT**”用最小二乘法进行回归拟合。其中，指令中的LS代表使用最小二乘法的方法进行拟合，WEIGHT为被解释变量体重，C为回归方程的常数项，HEIGHT为解释变量身高。

点击View切换到“Estimation视图”，我们可以看到更为直观的回归方程（如图 11所示）

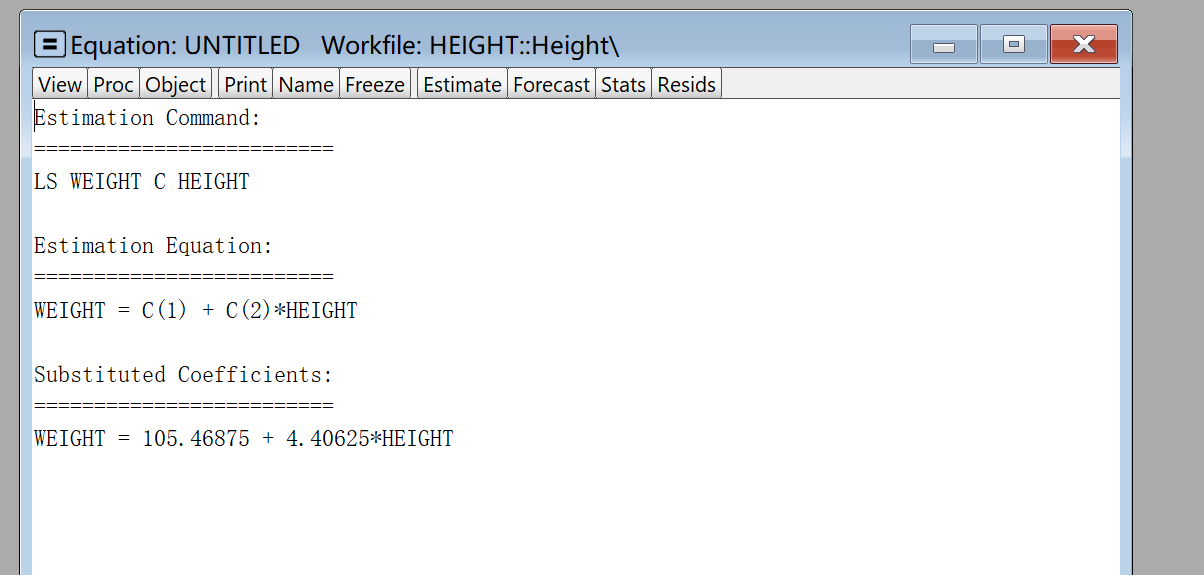


图 11：使用Eviews得出的回归方程

可以看到，我们得出的回归拟合方程为——

1. 解释系数的意义

下面对每个系数的意义和符号预期进行说明——

首先，105.46875是对方程的截距的估计，代表当HEIGHT为零时（即身高刚好为5英寸时），这个人的体重的值约为105.46875磅。其符号为正，符合正常的情况（体重一定是正数）。

而4.40625是对HEIGHT的斜率参数的估计，其代表这个人身高每增加一英寸，那么体重会增加4.40625磅。

我们预期身高和体重是正相关的关系，因为从生物学来讲，高的人具有更大的体表面积，这使得他们的身体可以容纳更多的脂肪组织。因此，相对于身高较矮的个体，身高较高的个体通常具有更高的体重。其符号为正，符合预期。

这个估计值和教材的方程（2-6）有些许差别，是因为选取的样本值得不同，从而导致估计的值也就有差异。这是正常的。

1. 解释调整的判定系数

在介绍调整的判定系数（）之前，我们需要了解什么是判定系数

判定系数反映了回归平方和在总平方和中所占的比例，其越大反应这个方程的拟合效果更好。其公式为——

但是判定系数是一个对解释变量的非减函数，增加解释变量会使得判定系数增大或者不变。因此判定系数不能很好的反应解释变量的个数对方程拟合效果的影响，增加无关的解释变量，很可能使模型失去经济含义！

因此，我们提出了调整的判定系数（），其公式为——

其中，k代表方程中的自变量个数当模型中新增解释变量时，调整的判定系数（）变化的方向依赖于新变量对回归拟合的贡献是否超过对损失一个额外自由度所作修正的补偿。

在本样本拟合的方程中，调整的判定系数（）为0.530112（图 12）。**其表征该模型解释了53.01%的平均高度变化。**这是一个可以接受的值，其表明方程**具有一定的拟合效果**。理想的值应该为接近1。究其原因是因为在HEIGHT=6时，有一个偏差较大的点（详见图 9）。

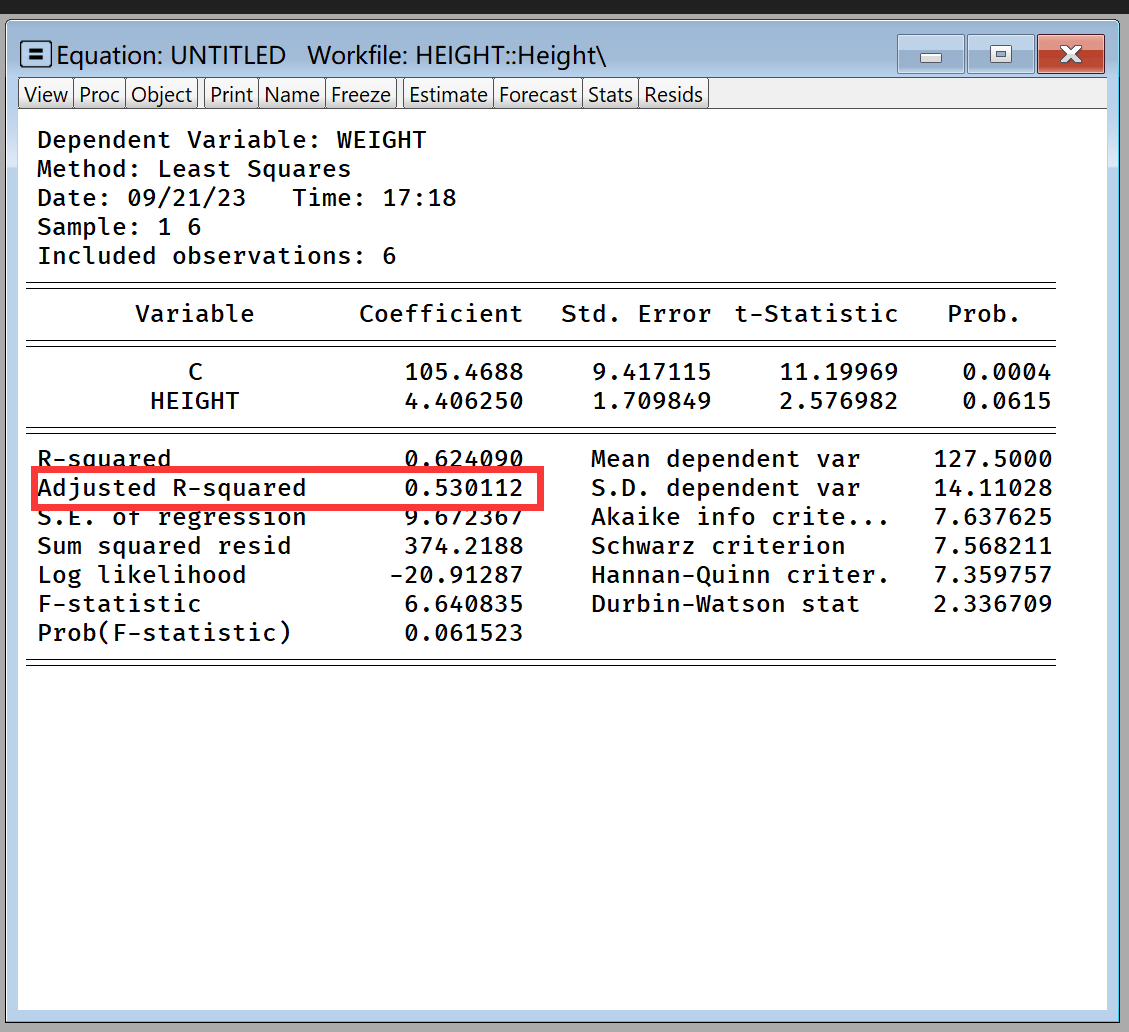


图 12：拟合结果的调整的判定系数

1. 新增变量

根据书本的要求，现在我们尝试把MALE加入到回归方程中。值得说明的是，因为要求统计的是同性别的数据，因此数据所有的性别虚拟变量均相同。因此预期MALE对WEIGHT没有任何影响，可能的估计值为0。

我们输入指令“**LS WEIGHT C HEIGHT MALE**”尝试进行拟合回归。

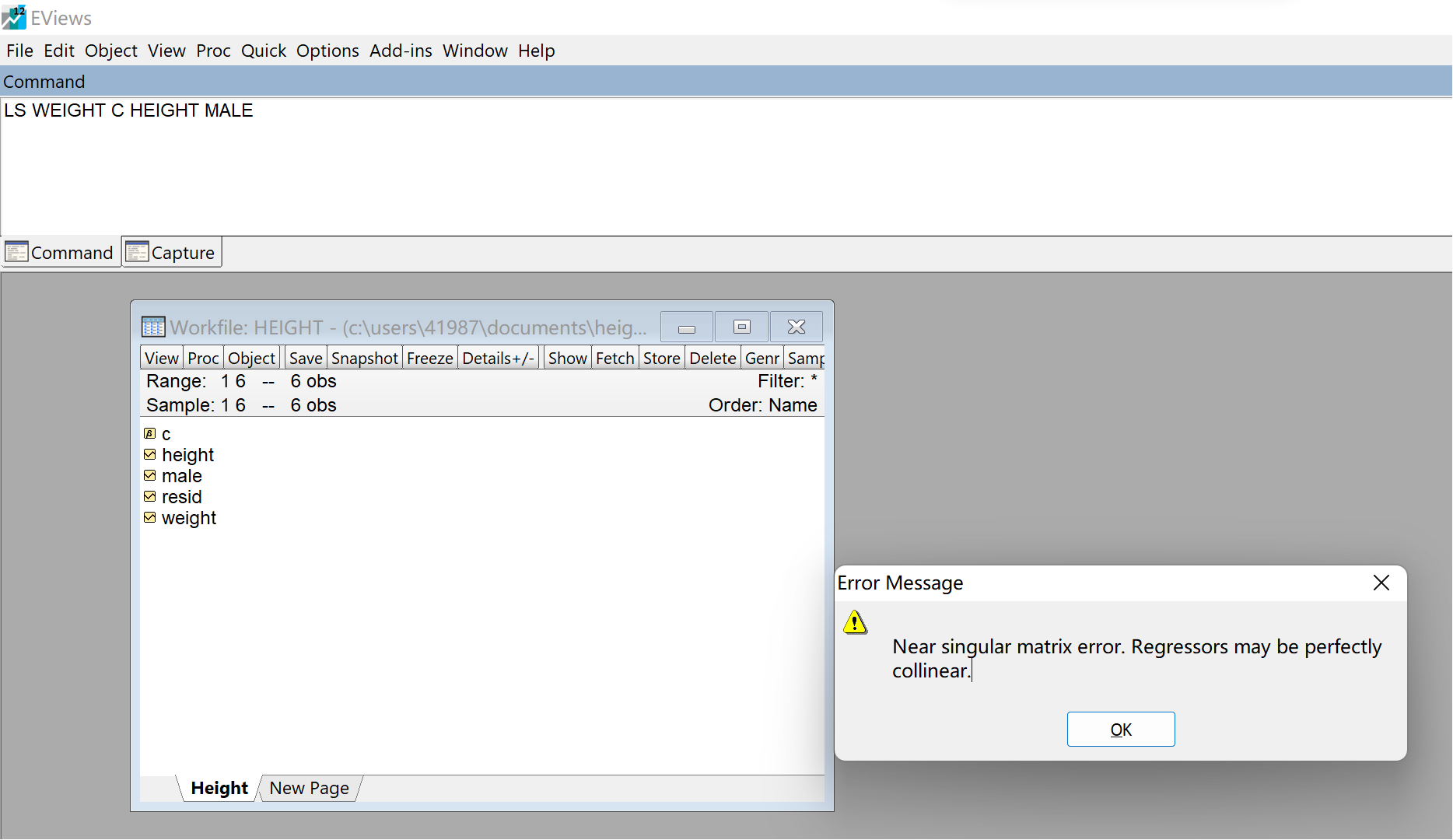


图 13：尝试直接加入变量MALE

从图 13可以看到，程序报错了，这是因为我们设定的MALE是完全一致的数值（性别均为男），回归模型需要解释变量和因变量之间存在一定的变异性才能进行有意义的分析。如果某个变量的值全部相同，那么它对于解释因变量的能力为0。程序无法判定其他性别对WEIGHT的影响。因此报错。

但这并不意味着性别和体重完全无关，因为我们还没有考虑其他性别对WEIGHT的影响。接下来，我们新增几个女同学的例子——

表 2：新增女同学后的身高体重数据

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 身高（英寸） | 体重（磅） | 性别 |
| 1 | 4 | 126 | 1 |
| 2 | 9 | 154 | 1 |
| 3 | 6 | 131 | 1 |
| 4 | 2 | 120 | 1 |
| 5 | 6 | 116 | 1 |
| 6 | 3 | 118 | 1 |
| 7 | 2 | 108 | 0 |
| 8 | 2 | 110 | 0 |
| 9 | 3 | 113 | 0 |

重新输入指令“**LS WEIGHT C HEIGHT MALE**”进行回归拟合。

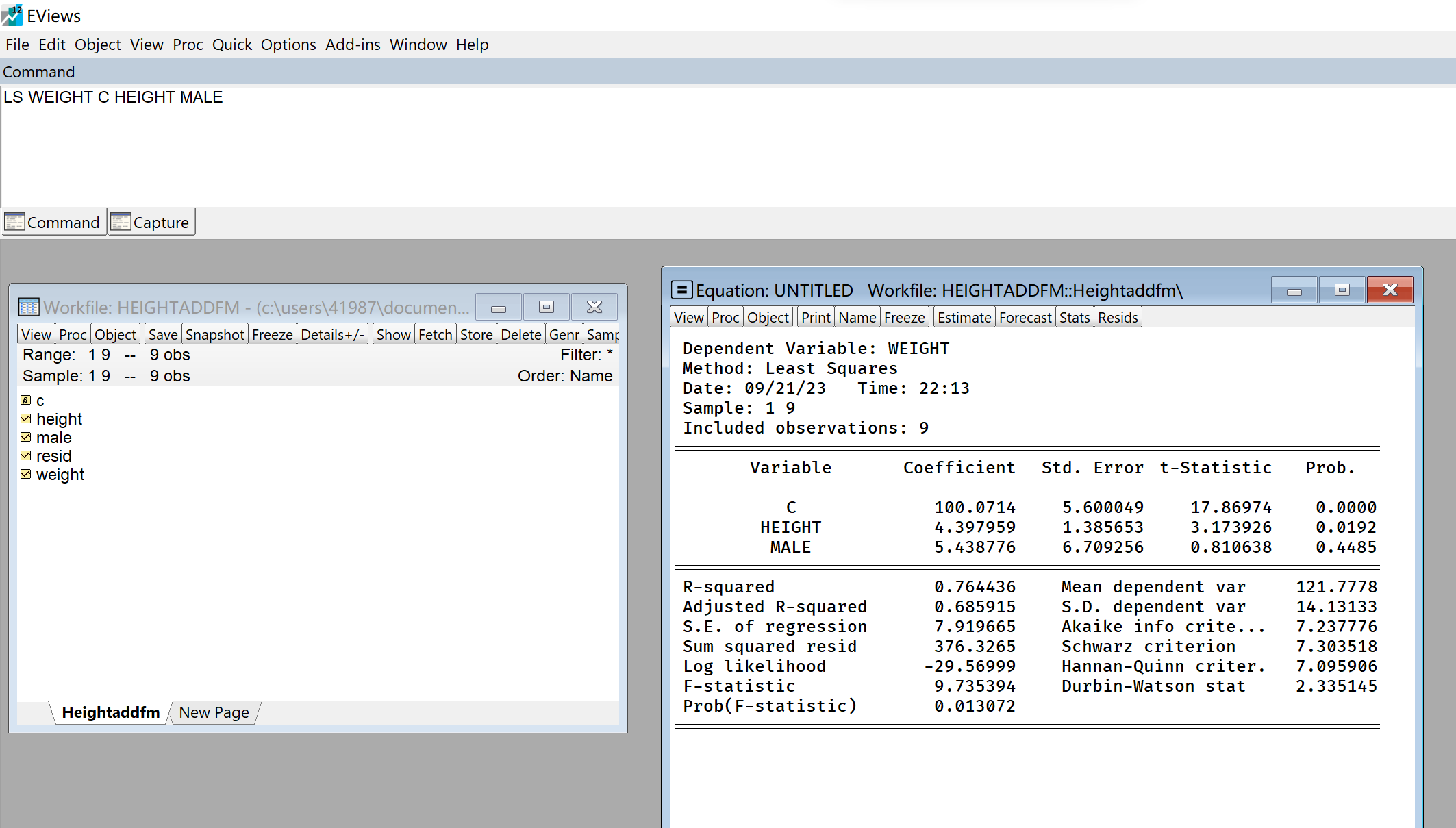


图 14：新增女同学的样本后的回归数据

可以看到程序完成了输出，结果显示，MALE对结果有着正向影响。符合预期，男生的身高普遍更高。拟合回归的方程为

可以看到增加了MALE变量后的拟合回归方程的其他两个参数也相较之前未加入MALE的发生了改变。说明新增参数对其他参数的估计值也有一定的影响。我们也看到到增加了MALE变量后的调整的判定系数（）也变大了，说明其拟合效果更好。

这告诉我们，如果要将一个变量纳入回归方程中，**那么这个变量的值一定要具有变异性（差异性），否则无法衡量这个变量对被解释变量的影响。**

**九、实验结论：**

通过本次实验，我们基于Eviews软件完成了简单的回归方程分析，以身边的同学为样本，进行数据收集，用最小二乘法拟合回归出了身高对体重的影响。在实验中，通过实践掌握了Eviews软件的基本操作，包括样本的查看、样本趋势分析（绘制了散点图）、样本特征（如最值、标准差、均值）的查看，并对样本数据进行尝试拟合。

在未考虑性别影响时，拟合的结果为——

当新增了女同学的样本数据，考虑了性别影响时，拟合的结果为——

同时，我们还发现了以下结论：

* 回归方程的斜率参数的意义在于：当其它变量相同（保持其他变量不变）时，特定变量对被解释变量的边际影响（贡献）
* 本次拟合结果仅反应本次收集到的样本的参数估计，对于不同的样本，参数的大小可能有所不同，但是符号应该基本一致。
* 判定系数反映了回归平方和在总平方和中所占的比例，但不能很好的反应解释变量的个数对方程拟合效果的影响
* 调整的判定系数（）变化的方向依赖于新变量对回归拟合的贡献是否超过对损失一个额外自由度所作修正的补偿。一般来说，其越接近1方程的拟合效果越好。
* 如果要将一个变量纳入回归方程中，那么这个变量的值一定要具有变异性（差异性），否则无法衡量这个变量对被解释变量的影响。

本次实验收获颇丰。

**报告评分：**

**指导教师签字：**

**电 子 科 技 大 学**

**实 验 报 告**

**学生姓名：** **\*\*\*** **学 号：202109120\*\*\*\*** **指导教师：李亚静**

**实验地点： 信软楼西303**  **实验学时**：**2学时**

**一、实验室名称：软件工程实验室**

**二、实验项目名称：估计1945~2014年美国经济总消费函数**

**三、实验原理：**

**（一）普通最小二乘法**

普通最小二乘法（OLS）是计量经济学中最经典且应用最广泛的回归分析方法。该方法通过最小化残差平方和来估计回归方程中的参数。在OLS中，我们试图找到一条线性回归线，使得观测数据点与该回归线的残差平方和最小。残差是指观测值与估计值之间的差异。

具体而言，OLS方法通过计算观测数据点到回归线的垂直距离（即残差），将这些距离平方化，并将它们求和，得到残差平方和。然后，根据最小二乘法原理，我们寻找能够使残差平方和最小的回归系数。

通过OLS方法，我们可以得到回归方程中的参数估计值，以及相关的标准误差、置信区间和显著性检验等统计指标。OLS的优点在于它是一种有效且易于实施的估计方法，可以用于各种回归模型，包括简单线性回归和多元回归。由于其广泛适用性和解释性好，OLS成为了计量经济学研究和实证分析中最常用的方法之一。该方法的使用需满足一系列古典假设，包括：

1. 回归模型是线性的，模型设定无误且含有误差项；
2. 误差项总体均值为零；
3. 所有解释变量与误差项都不相关；
4. 误差项观测值互不相关（不存在序列相关性）；
5. 误差项具有同方差(不存在异方差性)；
6. 任何一个解释变量都不是其他解释变量的完全线性函数（不存在完全多重共线性）

在上述古典假设成立的情况下，普通最小二乘估计量是最佳线性无偏估计量。所谓无偏估计量，是指OLS参数估计值的均值等于总体参数的真实值。

**四、实验目的：**

1. 学会通过查找文献寻找方程中可能的参数；
2. 掌握使用EViews软件进行普通最小二乘估计；
3. 加深对普通最小二乘估计方法的认识。

**五、实验内容：**

消费函数是宏观经济学中描述个人或家庭消费行为的数学模型。它将总消费与个人可支配收入之间的关系进行建模，用来解释个人在不同收入水平下的消费行为。本次实验的主要内容即是估计一个线性消费函数。

这个线性消费函数的被解释变量是实际个人消费支出，解释变量为个人实际可支配收入和利率。我们会通过查阅文献寻找变量的预期符号，并对拟合结果做出可能的预期假设，然后通过最小二乘法估计得到的回归参数可以告诉我们解释变量对被解释变量的影响程度和方向。在实验中将对各个回归参数的意义进行详细的分析。

同时，探究解释调整的判定系数和判定系数的实际意义和区别联系。

通过实验，掌握使用EViews软件进行样本数据的导入、样本数据描述的查看、普通最小二乘估计等基本操作，学会EViews的基本命令和使用方法，并加深对普通最小二乘估计方法的认识和理解。同时也需要掌握计量经济学文献收集的基本方法。

**六、实验器材（设备、元器件）：**

计算机、EViews软件

**七、实验步骤：**

1. **收集相关文献**：首先对于消费函数这一基本经济学概念进行初步了解，通过查阅文献、网页资源搜索等相关资料学习其基本概念和基本的表示方法。了解典型的消费函数。具体内容包括——
   1. 了解“消费函数”这一基本概念内容以及典型的消费函数（如凯恩斯消费函数）
   2. 了解“边际消费倾向”的实际定义
   3. 查阅文献，具体了解关于消费函数的实例、内涵、基本概念等相关信息，并进行整理。
2. **确定基本模型和预期符号：**根据题目的信息，确定基本的模型（包括解释变量和函数形式），题目已经限定该消费函数为一个线性函数，解释变量包括个人可支配收入和利率。具体包括——
3. 确定变量的名称
   1. :第t年的实际个人消费支出
   2. :第t年的个人实际可支配收入；
   3. :第t年的实际利率。
4. 利用书本方程(3-1)的形式，把个人可支配收入和利率的函数表达出消费方程，然后用适当的变量名来代替Y和X。

同时，基于文献调查的结果和给出的相关方程模型，确定每一个系数的预期符号，并解释原因。

1. **使用Eviews检测、整理数据：**从课程网站中下载相关的数据样本，并将其导入到Eviews软件中，利用Eviews查询其基本的数据特征，具体包括——
   1. 最值
   2. 平均值
   3. 标准差

并检查上述数据特征有无异常情况发生，如有无不可能的最大值或最小值（如负消费）或不合理的高值（如超过100%的利率）。若有则需要进行错误数据的修正。

1. **使用Eviews进行回归拟合：**利用最小二乘法进行回归拟合，解释个人可支配收入、利率和实际个人消费支出之间的关系，确定各项回归参数的值和意义，并在电脑上进行输出展示。具体步骤包括——
   1. 导入数据
   2. 输入指令“LS CONS C PYD AAA”
   3. 按下回车，结果显示在屏幕上
2. **解释并评价估计方程：**基于Eviews计算得出的回归系数，解释每个回归系数的意义，并于预期的值和符号进行比较。具体包括——
   1. 回归系数估计值的符号表示了解释变量对被解释变量的的正负影响关系。
   2. 回归系数估计值的大小表示了单位变化对被被解释变量的影响程度大小

并在Eviews中的调整结果中找到调整的判定系数和判定系数，解释他们的意义、联系和区别，评判回归方程的拟合效果，回答以下几个问题——

1. 边际消费价值是多少（保留三位小数）？
2. 保持企业债券AAA利率不变，如果人均可支配收入减少100万美元，边际消费价值会变化多少？
3. 如果企业债券AAA利率上升三个百分点（个人可支配收入不变）边际消费价值将变化多少？
4. 基于上面部分的答案，评价回归结果是否合理？或者是有某种错误吗？
5. **报告结果：**整理最终得出的回归结果并放在书本方程(3-4)给出的标准格式里。

**八、实验数据及结果分析：**

1. 文献收集

首先，我们对于消费函数、边际消费倾向这一核心概念进行文献收集。

* **边际消费倾向**

边际消费倾向（Marginal Propensity to Consume，MPC）是经济学中用来衡量个体或整个经济体在收入变动下，新增收入中用于消费支出的比例。它表示当可支配收入增加一美元时，消费支出的增长幅度。

边际消费倾向在宏观经济学中具有重要意义。通过研究边际消费倾向，我们可以了解个体或整个经济体对于经济政策、收入变动等因素的反应。当经济面临衰退时，通过刺激消费支出来提振经济，需要了解边际消费倾向的大小，以便预测和评估政策的影响。此外，边际消费倾向也对理解储蓄行为、收入分配、消费模式等方面具有重要启示作用。

* **消费函数**

消费函数是宏观经济学中描述个人或家庭消费行为的数学模型，最先由**J·凯恩斯**1936年在 **《就业、利息和货币通论》**一书中提出。它将总消费与个人可支配收入之间的关系进行建模，用来解释个人在不同收入水平下的消费行为。消费函数的一般形式可以表示为：。其中，C表示总消费，Y表示个人或家庭的可支配收入，f表示函数关系。

消费函数可以有多种具体形式，其中较常见的是线性消费函数

线性消费函数假设每增加一个单位的可支配收入，消费也增加一个固定的数量。它的一般形式为：。其中，a表示固定消费，b表示边际消费倾向度（即每增加一个单位的可支配收入，消费增加的数量）。

消费函数的具体形式可以根据实际情况进行拟合和估计。经济学家使用消费函数来分析个人或家庭在不同收入水平下的消费行为，评估政策对消费的影响以及预测经济增长等方面的变化。消费函数在宏观经济模型中起着重要的作用，对经济政策制定和宏观经济预测有着重要的参考价值。

按照题目要求，查阅资料涉及到的参考文献如下：

1. 李柱锡. 消费函数[J]. 外国经济与管理,1981(11):11-14.
2. 史玉伟. 消费函数理论主要假说述评[J]. 经济经纬,2005(3):17-19. DOI:10.3969/j.issn.1006-1096.2005.03.006.
3. 确定基本模型

根据题目提供的信息，我们可以将个人可支配收入（PYD）和实际利率（AAA）作为解释变量，实际个人消费支出（）作为被解释变量。

根据题目要求，我们要拟合的消费函数为一个线性函数形式。那么消费方程可以表示为：

其中，

* 表示第t年的实际个人消费支出
* 表示第t年的个人实际可支配收入；
* 表示第t年的AAA公司的债券实际利率。

是截距项，表示当个人可支配收入和利率为零时的基础消费水平；是对应于个人可支配收入和利率的回归系数。

1. 符号假定

根据经济学相关原理和日常经验，不难对的符号做出预期假设——

我们认为**的符号为正**：因为一般来讲，个人实际可支配收入越高，边际消费倾向也就较高。随着个人可支配收入的增加，个人的财务状况得到改善，个人的信心也会得到增强，从而使得他们更愿意购买更昂贵的产品或服务。因此，实际个人消费支出也越高。即个人实际可支配收入对实际个人消费支出呈正向的影响。

而**的符号为负：**因为一般来讲，利率越高，人们越愿意将钱存入银行以换取高额的利息，这样留存下来可消费的金额变少。因此，实际个人消费支出也随之降低；相反，如果利率处于较低水平，把钱存入银行是个不明智地选择（因为利息很低），因此人们倾向于将钱用来购买其他物资或者服务。因此银行利息对实际个人消费支出呈负相关的影响。

按照书本（3-1）方程的形式，可以将这个消费函数写成如下的形式——

+

1. 收集检查整理数据

我们打开Eviews软件，依次选择“File->Import ->Import From File”将相关表格数据导入Eviews软件中。

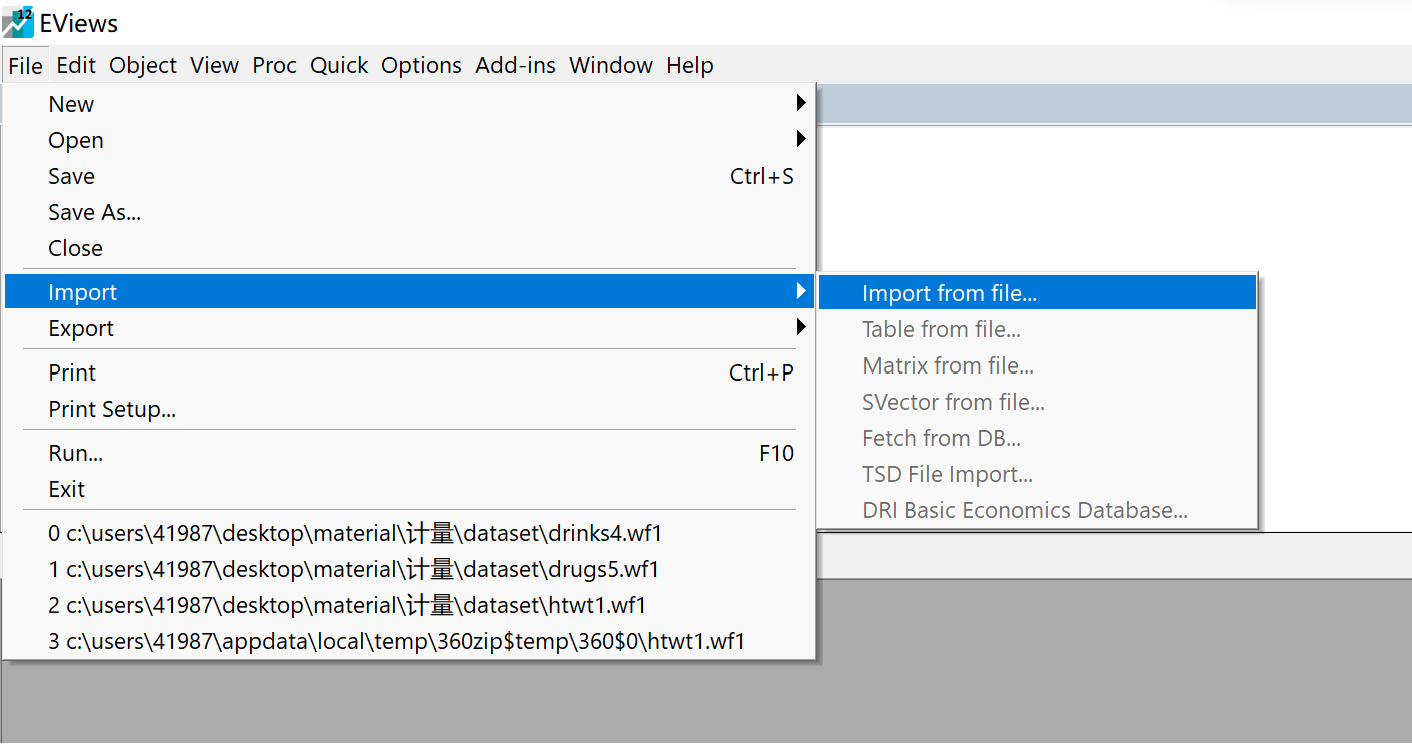


图 15: 将数据导入表格

使用FRED的数据核对数据集，数据集基本无误

接下来，我们对导入的数据进行相关数据特性的查看、检查

* **查看样本容量**

首先将要查阅的数据选中，右击选择“OPEN->as GROUP”，即可展示导入的完整数据。

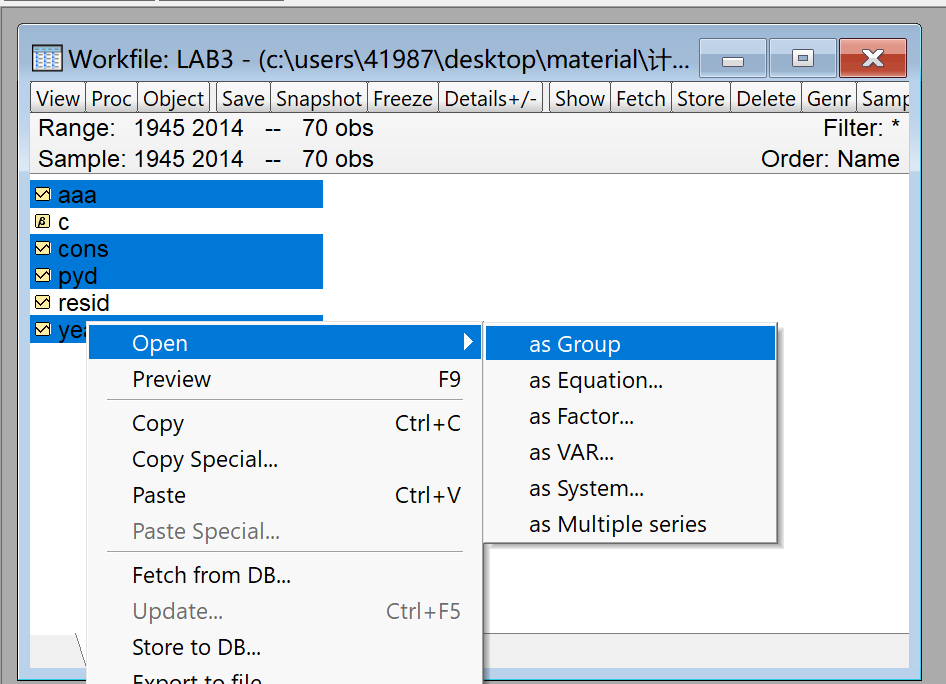


图 16：打开数据展示详情

点击后即可得到如图 5所示的数据列表——

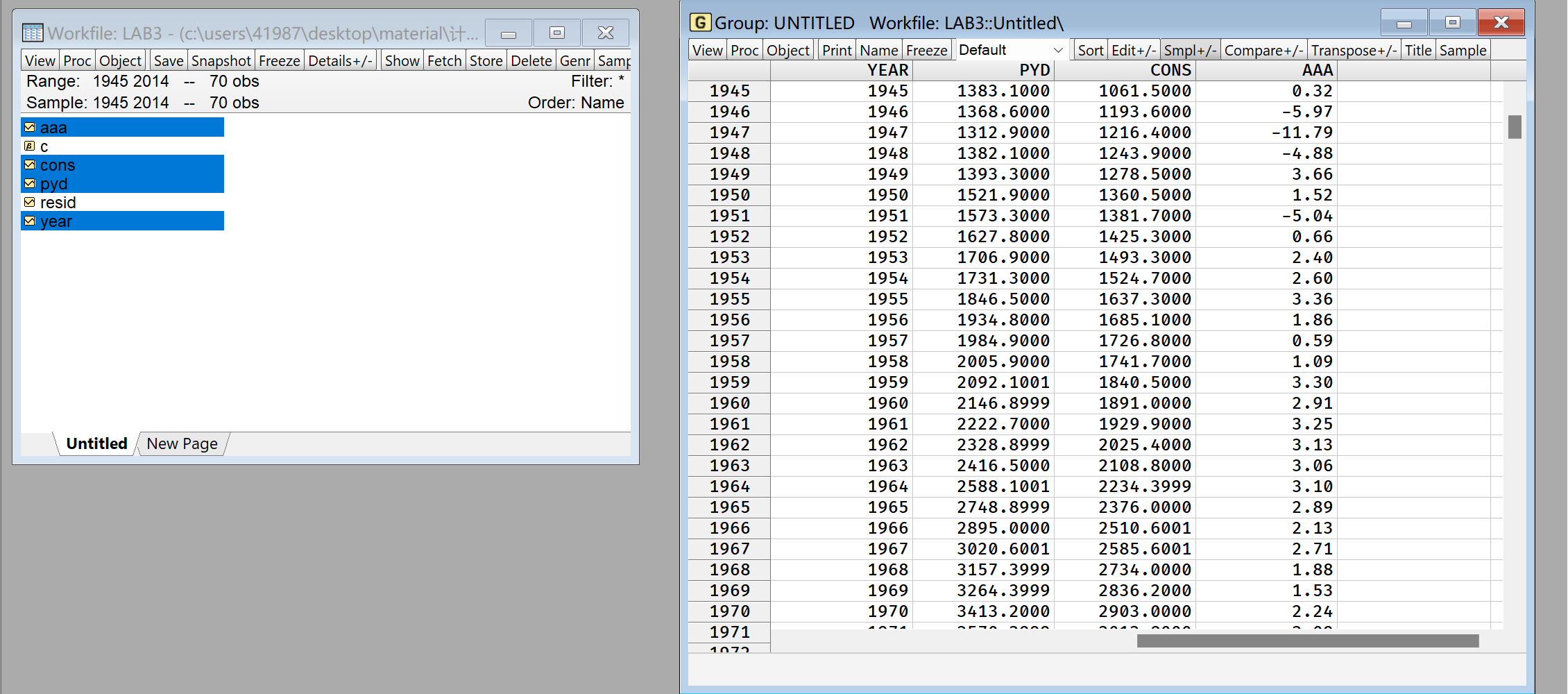


图 17：查看导入的数据

可以看到左上方写明了“Sample……. 70 obs”即表明样本大小为70个，即我们**一共导入了70个样本数据**。验证正确

* **查看样本数据的最大值、最小值**

在Group视窗中，选择“View->Descriptive Stats->Common Sample”即可查看数据的详细特征（包括最值、标准差、方差等），相关步骤截图如图 6所示

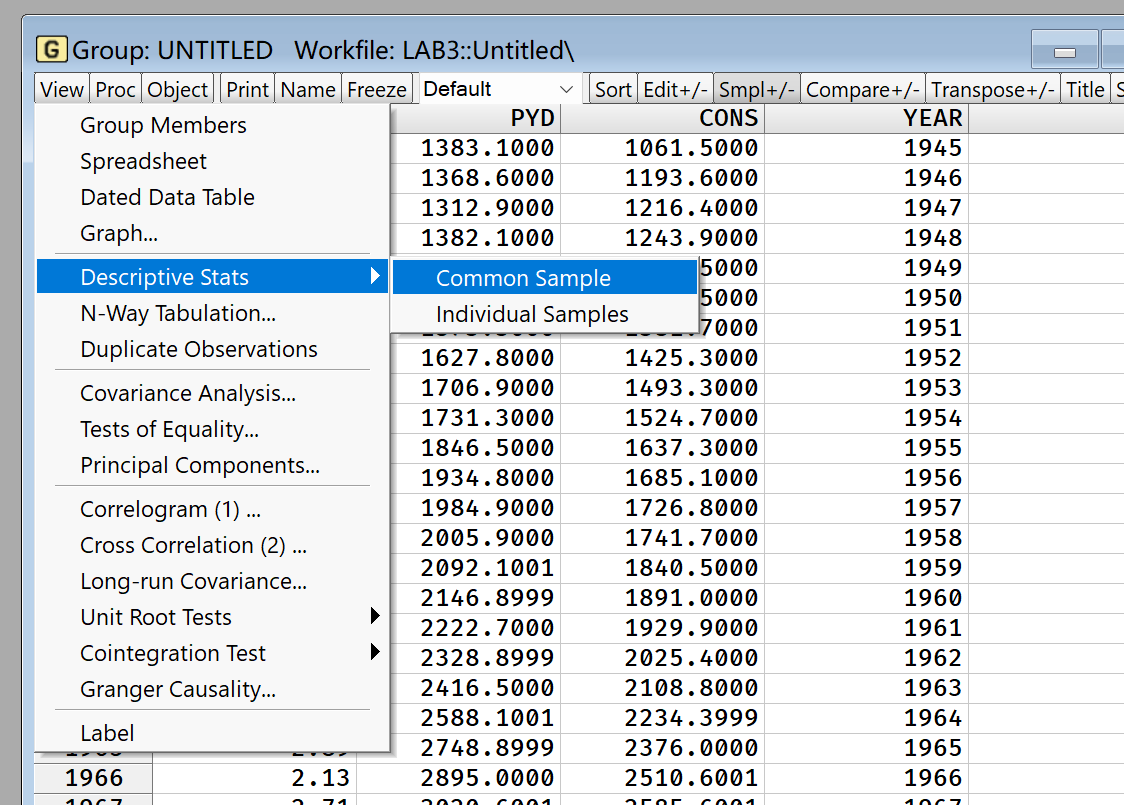


图 18：查看样本数据的特征步骤

得到的数据特征如下——

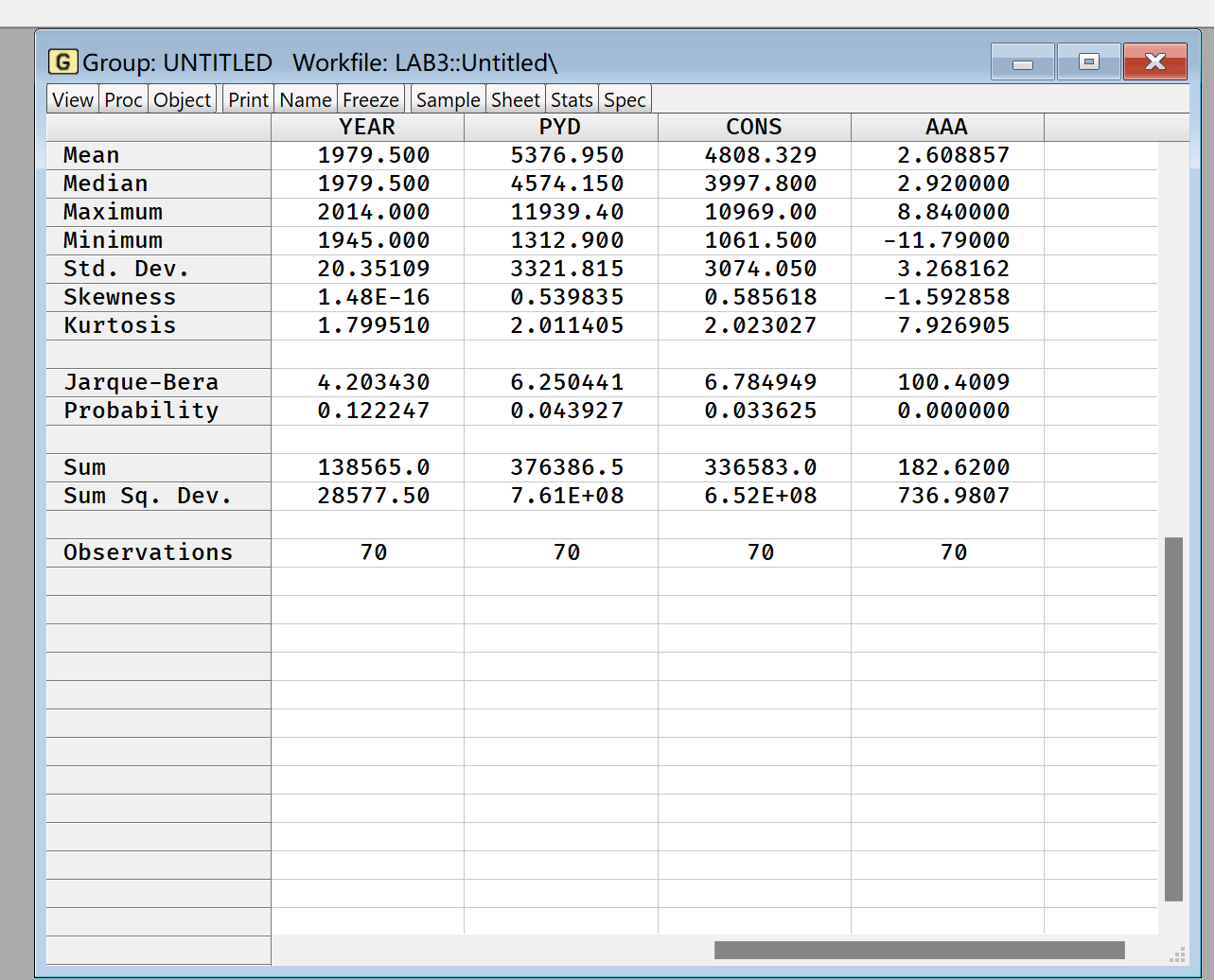


图 19：样本数据的特征

可以看到

* **PYD的最大值为11939.4，最小值为1312.900，标准差3321.815；**
* **CONS的最大值为10969.00，最小值为1061.500，标准差为3074.050；**
* **AAA的最大值为8.840000，最小值为-11.79000，标准差为3.268162。**

数据分布和导入一致，验证正确。

且没有不可能的最大值或最小值（如负消费）或不合理的高值（如超过100%的利率）。数据无需修正。

1. 回归拟合

接下来利用Eviews进行回归拟合

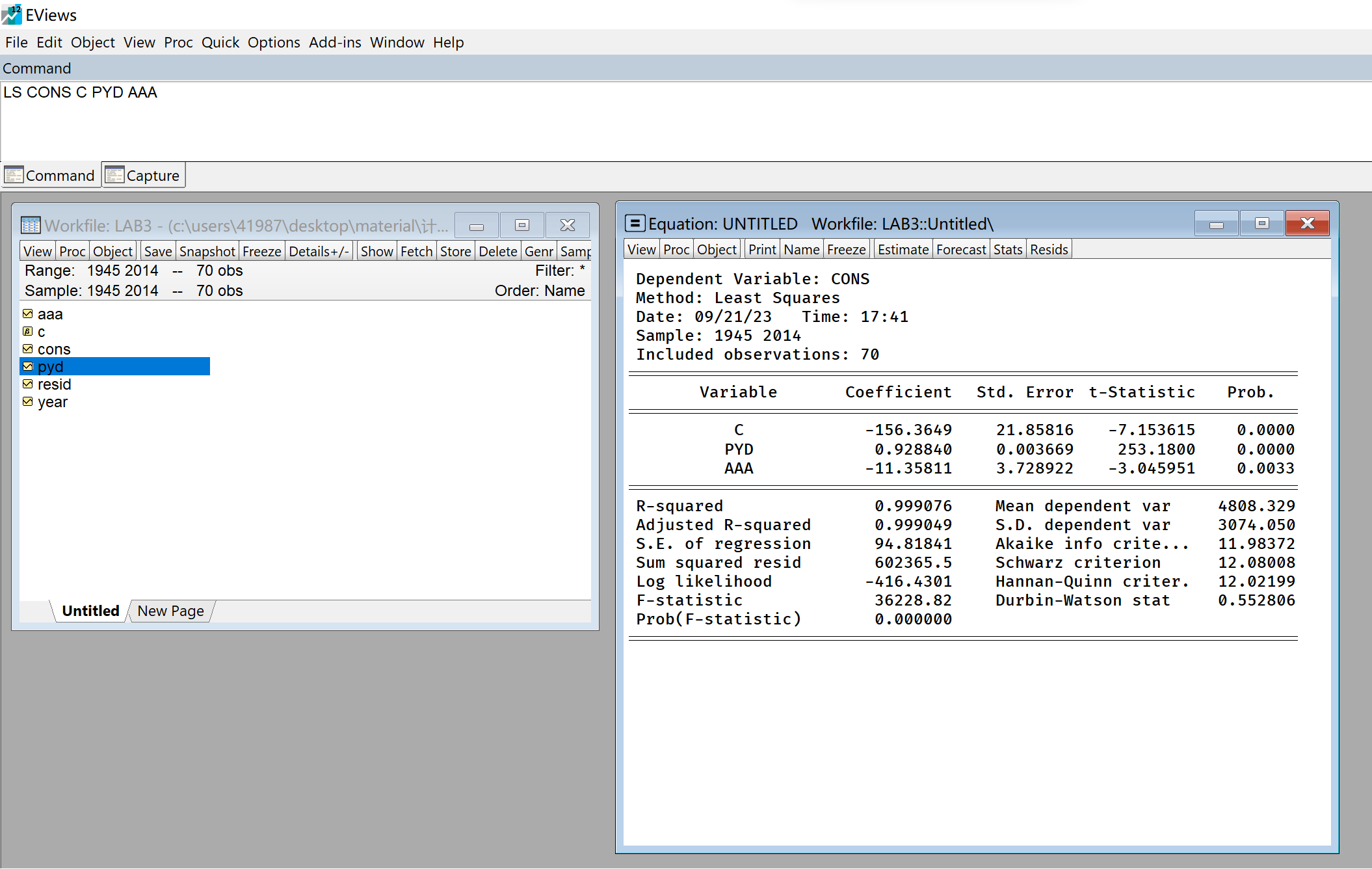


图 20：输入指令进行拟合

如图 10所示，我们输入指令“**LS CONS C PYD AAA**”用最小二乘法进行回归拟合。其中，指令中的LS代表使用最小二乘法的方法进行拟合，CONS为被解释变量实际个人消费支出，C为回归方程的常数项，PYD为解释变量个人实际可支配收入,AAA为实际利率。

点击View切换到“Estimation视图”，我们可以看到更为直观的回归方程（如图 11所示）

可以看到，我们得出的回归拟合方程为——

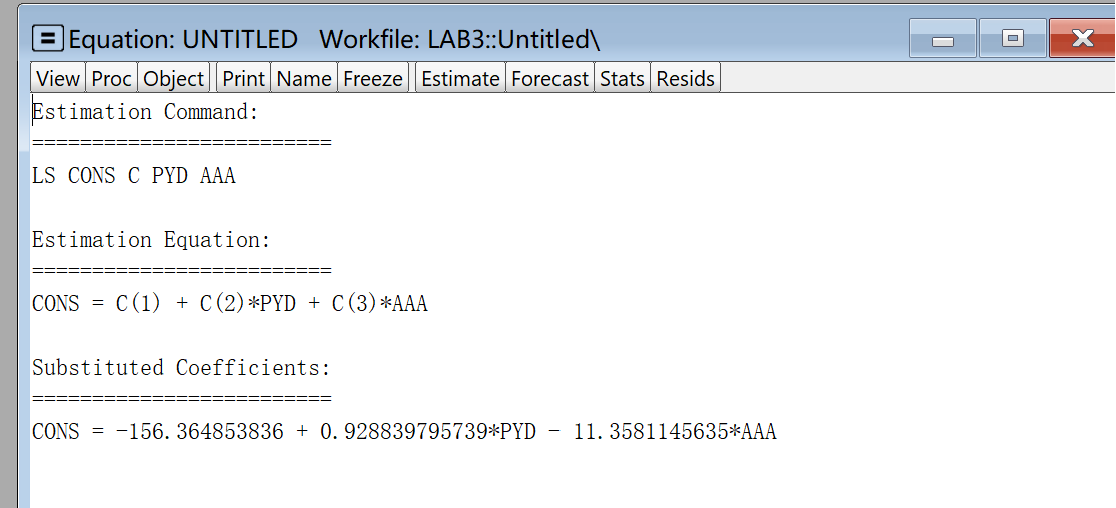


图 21：使用Eviews得出的回归方程

1. 解释并评价估计方程

下面对Eviews计算得出的回归方程进行分析和评价

* **符号分析**

在给出的回归方程

中，的符号分别为正、负，其**与我们在第三步中对方程的符号估计完全一致**。其表明随着个人可支配收入的增加，实际个人消费支出也会增长；随着实际利率的增长，实际个人消费支出降低。

* **方程的拟合效果**

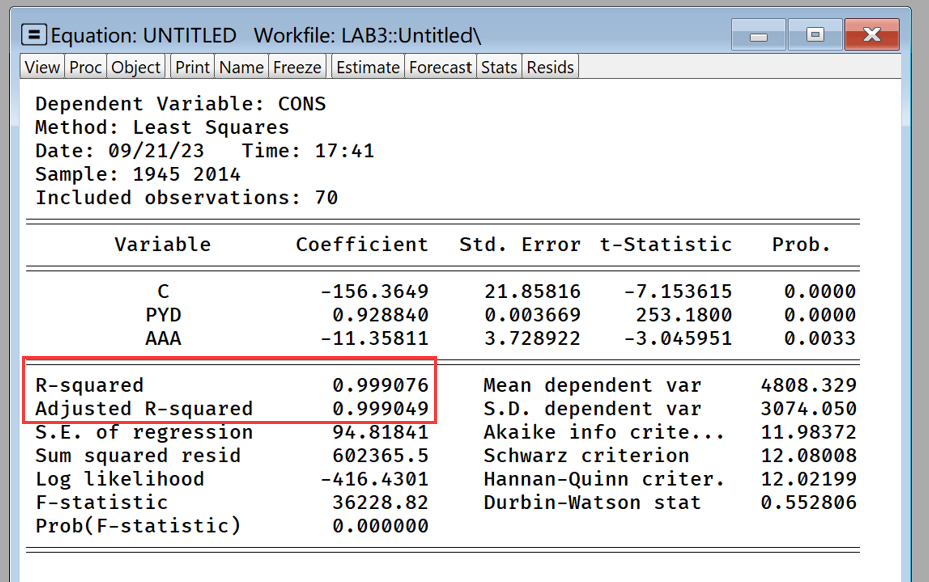


图 22：方程拟合的结果

在图 22的红圈部分我们可以看到，Eviews软件给我们列举了两个值——判定系数（R-squared）和调整的判定系数（Adjusted R-squared）。

在本题中判定系数（R-squared）为0.999076；调整的判定系数（Adjusted R-squared）为0.999049，均极接近1，表明**方程拟合效果非常好**。

判定系数反映了回归平方和在总平方和中所占的比例，其越大反应这个方程的拟合效果更好。其公式为——

而调整的判定系数（），其公式为——

其中，k代表方程中的自变量个数当模型中新增解释变量时，调整的判定系数（）变化的方向依赖于新变量对回归拟合的贡献是否超过对损失一个额外自由度所作修正的补偿。

这两者的**区别即在于能否很好的反应解释变量的个数对方程拟合效果的影响。**若增加无关的解释变量，往往会使得判定系数增大，但增加的变量可能使得原方程失去经济意义，那么判定系数就不能很好的反应解释变量的个数对方程拟合效果的影响。

调整的判定系数考虑了自变量的数量和样本量，通过惩罚自变量的增加和样本量的减少来修正模型复杂度，因此相比于判定系数，调整的判定系数更具有一般性和稳健性。

在比较不同模型时，调整的判定系数更适合作为参考指标，以确保模型的拟合程度是基于合理的自由度调整而得出的。

* **关于方程系数的解释**（对应问题c~d）

边际消费倾向是衡量新增收入中用于消费支出的比例。

在本题中，PYD的系数0.92884即表示保持方程中其他解释变量不变的前提下，个人实际可支配收入每增加1美元，那么个人实际消费支出就会增加 0.929美元；AAA的系数-11.358即表示保持方程中其他解释变量不变的前提下，实际利息每上升1，那么个人实际消费支出就会减少11.358美元

**因此，边际消费倾向（即新增收入中用于消费支出的比例）即是0.929**。

在利率不变的情况下，当个人实际可支配收入降低100万美元，个人实际消费支出会减少0.929×100=**92.9万美元。**

在个人实际可支配收入不变的情况下，当利率上升三个百分点，个人实际消费支出会减少11.358×3=**34.074万美元。**

基于以上结果，我们看出，拟合的结果和我们预期的结果基本一致，没有异常情况出现，因此，可以认为拟合过程基本没有错误。

1. 形成报告

根据书本的要求，我们将最终的拟合结果反应成与书本方程（3-4）的标准形式。

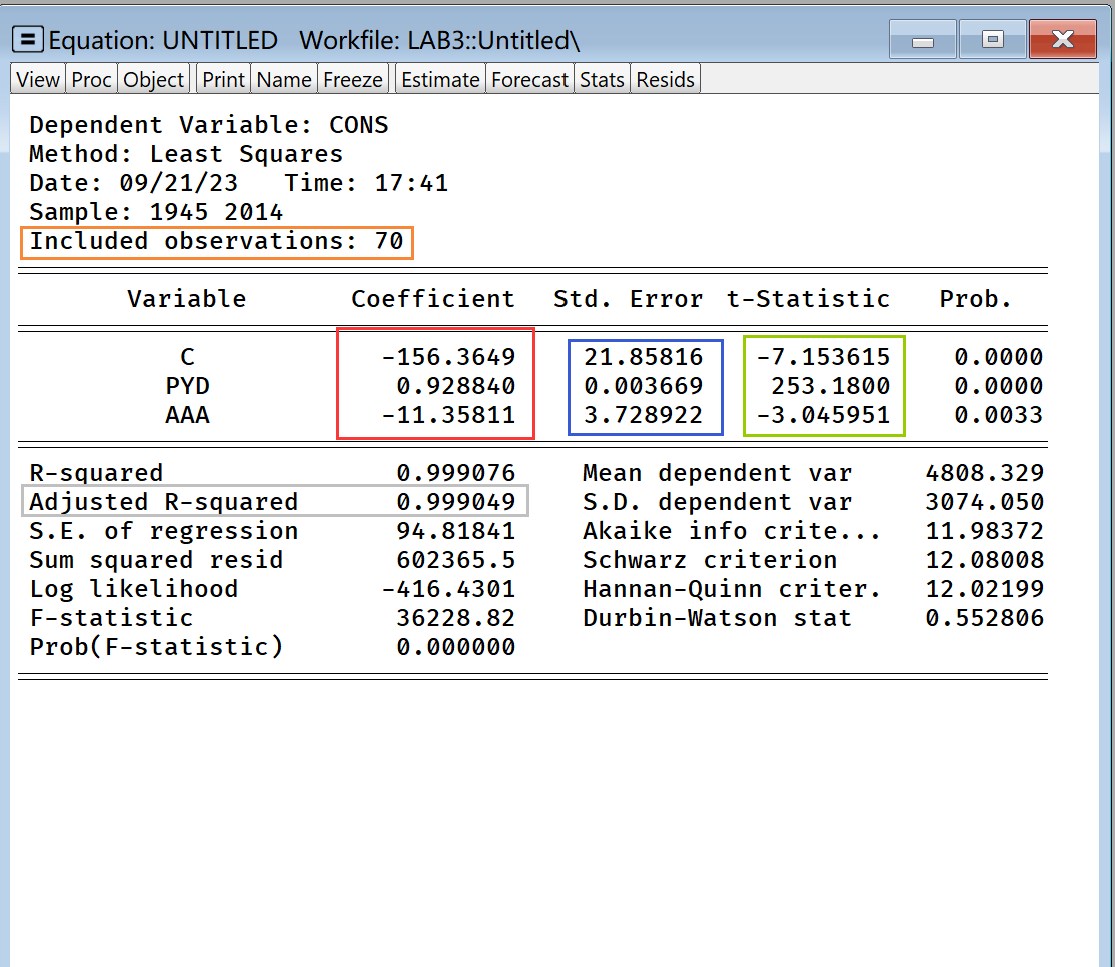


图 23：拟合结果

在Eviews给出的拟合结果中，我们能够很清除的找出书本方程（3-4）的标准形式所需要的各个参数。其中红色框的数据为系数、蓝色框的数据为标准差、绿色框的数据为t值、上方橙色框的数据为样本容量、下方灰色框的数据为调整的判定系数。

综上，我们将结果汇总，可以写出如下的标准形式——

**九、实验结论：**

通过本次实验，我们完成了一个完整的线性方程回归模型的拟合分析。

通过查阅资料，了解到“消费函数”的基本信息，基于文献确定模型的形式并做出参数估计。随后基于Eviews软件完成了对样本数据回归方程分析，用最小二乘法拟合回归出了个人可支配收入、利息与实际个人消费之间的关系。在实验中，通过实践掌握了Eviews软件的基本操作，包括样本的查看本特征（如最值、标准差、均值）的查看，并对样本数据进行尝试拟合。

最后，汇总数据，得出了完整的拟合结果，并反应为标准形式——

同时，我们还发现了以下结论：

* 拟合的消费函数反映出——利率对个人消费支出呈负相关，个人可支配收入对个人消费支出呈正相关。
* 回归方程的斜率参数的意义在于：当其它变量相同（保持其他变量不变）时，特定变量对被解释变量的边际影响（贡献），它告诉我们，当特定自变量增加（或减少）一个单位时，被解释变量平均会增加（或减少）多少个单位。
* 判定系数反映回归平方和在总平方和中所占的比例，但不能很好的反应解释变量的个数对方程拟合效果的影响；而调整的判定系数（）考虑了自变量的数量和样本量，通过惩罚自变量的增加和样本量的减少来修正模型复杂度，因此相比于判定系数，调整的判定系数更具有一般性和稳健性。

**报告评分：**

**指导教师签字：**

**电 子 科 技 大 学**

**实 验 报 告**

**学生姓名：** **\*\*\*** **学 号：202109120\*\*\*\*** **指导教师：李亚静**

**实验地点： 信软楼西303**  **实验学时**：**2学时**

**一、实验室名称：软件工程实验室**

**二、实验项目名称：估计地区平均寿命的模型与假设检验**

**三、实验原理：**

**（一）普通最小二乘法**

普通最小二乘法（OLS）是计量经济学中最经典且应用最广泛的回归分析方法。该方法通过最小化残差平方和来估计回归方程中的参数。在OLS中，我们试图找到一条线性回归线，使得观测数据点与该回归线的残差平方和最小。

通过OLS方法，我们可以得到回归方程中的参数估计值，以及相关的标准误差、置信区间和显著性检验等统计指标。OLS的优点在于它是一种有效且易于实施的估计方法，可以用于各种回归模型，包括简单线性回归和多元回归。由于其广泛适用性和解释性好，OLS成为了计量经济学研究和实证分析中最常用的方法之一。该方法的使用需满足一系列古典假设，包括：

1. 回归模型是线性的，模型设定无误且含有误差项；
2. 误差项总体均值为零；
3. 所有解释变量与误差项都不相关；
4. 误差项观测值互不相关（不存在序列相关性）；
5. 误差项具有同方差(不存在异方差性)；
6. 任何一个解释变量都不是其他解释变量的完全线性函数（不存在完全多重共线性）

在上述古典假设成立的情况下，普通最小二乘估计量是最佳线性无偏估计量。所谓无偏估计量，是指OLS参数估计值的均值等于总体参数的真实值。

**（二）假设检验**

假设检验是统计推断的一个重要方法，用于对一个或多个关于参数的假设进行验证。**可以用于验证经济理论模型的参数值的显著性水平高低**。

在假设检验中，我们提出一个假设和一个备择假设。原假设通常是和我们预期相反否定的假设，而备择假设则是对原假设的否定或替代。通过对收集到的样本数据进行统计分析，我们根据统计量的结果来判断是支持原假设还是支持备择假设。

其中，**t检验**是一种常用的假设检验方法，它基于样本均值、样本标准差和样本大小，计算得到一个t统计量。通过比较该统计量与t分布的临界值，可以判断两组样本均值是否存在显著差异。如果计算得到的t统计量超过了临界值，就可以拒绝零假设，认为样本均值之间存在显著差异。

F检验(F-test)是在原假设中包含多个假设或涉及关于若干参数的单个假设时所进行的检验。像这种“联合”或“复合”假设适用于模型所隐含的经济理论中特定的数值针对多个参数的情况。

**四、实验目的：**

1. 掌握使用EViews软件进行普通最小二乘估计以及假设检验；
2. 加深对普假设检验基本步骤的认识。
3. 熟悉t检验、p值、F检验这几类典型的检验方法

**五、实验内容：**

假设检验是用于检测经济理论模型的参数值的显著性水平高低的一种方法。在本次实验中，我们首先会基于美国51个地区的平均寿命，推理、拟合一个寿命模型，以回顾之前的最小二乘法。

此后，我们将详细对每个参数进行t检验、p值检验，以验证这些参数值的显著性水平高低，并做出合理的解释。

随后，我们会使用F检验对整个方程进行一个总体检验。

这些统计方法可以帮助我们更加精确地评估参数的显著性，从而为数据分析提供更有力的支持。

**六、实验器材（设备、元器件）：**

计算机、EViews软件

**七、实验步骤：**

1. **指定模型**：运用书本方程(5-1)的格式，用样本数据中全部的七个独立变量指定一个关于预期寿命的线性回归方程，并用适当的希腊字母和标识来表示他们。
2. **确定预期符号：**根据题目的信息，基于日常生活经验和经济学知识，对假定的模型，确定每一个系数的预期符号，并解释原因。
3. **使用Eviews汇总、整理数据：**从课程网站中下载相关的数据样本，并将其导入到Eviews软件中，利用Eviews查询其基本的数据特征，具体包括——
   1. 最值
   2. 平均值
   3. 标准差

并检查上述数据特征有无异常情况发生，如有无不可能的最大值或最小值或不合理的高值。若有则需要进行错误数据的修正。

1. **使用Eviews进行回归拟合、估计：**利用最小二乘法进行回归拟合，解释个人可支配收入、利率和实际个人消费支出之间的关系，确定各项回归参数的值和意义，并在电脑上进行输出展示。具体步骤包括——
   1. 导入数据
   2. 输入指令“LS LIFEEXPECT C MEDINC UNINSURED SMOKE OBESITY TEENBIRTH GUNLAW METRO”
   3. 按下回车，结果显示在屏幕上
2. **假设检验（t检验）：**检验MEDINC、UNINSURED、SMOKE、TEENBIRTH 这四个参数。具体步骤为——
   1. 写出原始假设和备择假设。
   2. 根据样本数量求得自由度以及t统计量的临界值
   3. 比较计算得到的t统计量与临界值的大小，得出结论
3. **假设检验（p值）：**检验OBESITY、GUNLAW、METRO 这3个参数。具体步骤为——
   1. 写出原始假设和备择假设。
   2. 根据样本数量求得自由度以及p值
   3. 得出结论
4. **总体假设检验（F检验）：**在5%的显著性水平下，用整体F统计量检验回归模型是否有效。具体步骤为——
   1. 写出原始假设和备择假设。
   2. 根据样本数量求得自由度以及F统计量的临界值
   3. 比较计算得到的F统计量与临界值的大小，得出结论
5. **得出结论：**整理最终得出的假设检验结果，并解释“gunlaw系数的绝对值远大于smoke的系数绝对值。这是否意味着通过枪支法来保护儿童在预期寿命上的影响比减少吸烟的影响大3%呢？”

**八、实验数据及结果分析：**

1. 指定模型

根据题目提供的信息，因为我们的目的是求得50个州和哥伦比亚特区的平均寿命的模型，因此我们以预期寿命（）作为被解释变量，其余变量（包括家庭收入、有无健康保险、吸烟情况、肥胖情况、早产情况、有无保护儿童的枪支法案、都市人口）作为解释变量

根据题目要求，我们要拟合的消费函数为一个线性函数形式。那么平均寿命方程可以表示为：

其中各个变量的含义如下表所示——

表 3：各个变量的意义

|  |  |
| --- | --- |
| 变量 | 意义 |
|  | 在时期i出生时的预期寿命，2010年 |
|  | 时期i家庭收人的中位数，2010年 |
|  | 时期i没上健康保险责任范围的人数百分比，2008~2010年 |
|  | i时期州内吸烟人口所占百分比，2006~2010年 |
|  | 时期i肥胖成年人占比（身体质量指数比大于等于30），2006~2012年 |
|  | 时期i每1000名女性中在15到19岁生产的母亲的数量，2010年 |
|  | 为虚拟变量，若有时期i时有保护儿童的枪支法案为1，否则为0,2010年 |
|  | 时期i在住在都市统计区域的人口百分比，2010年 |

方程中，是截距项，表示当其他变量的值为0的情况下对应的平均寿命；是对应于各个解释变量的回归系数。

1. 符号假定

根据经济学相关原理和日常经验，不难对各个斜率参数的符号做出预期的假设——

表 4：各个变量的预期符号

|  |  |  |
| --- | --- | --- |
| 变量 | 预期符号 | 理由 |
|  | **+** | 家庭收入越高，可能受到的医疗条件、健康服务更好 |
|  | **-** | 未上保险可能对自己的健康不太在意，拥有保险的人可能对自己的健康更加关注 |
|  | **-** | 吸烟有害健康 |
|  | **-** | 肥胖也会对健康造成影响 |
|  | **-** | 母亲早产也对自己的身体有着一定伤害 |
|  | **?** | 无法确定，因为这个枪支法只对儿童有效。 |
|  | **?** | 无法确定，城市的人口越高，可能会造成污染等城市化原因导致寿命减少；但是城市也可能具有更好的医疗条件，因此无法确定 |

1. 汇总数据

我们打开Eviews软件，依次选择“File->Import ->Import From File”将相关表格数据导入Eviews软件中。

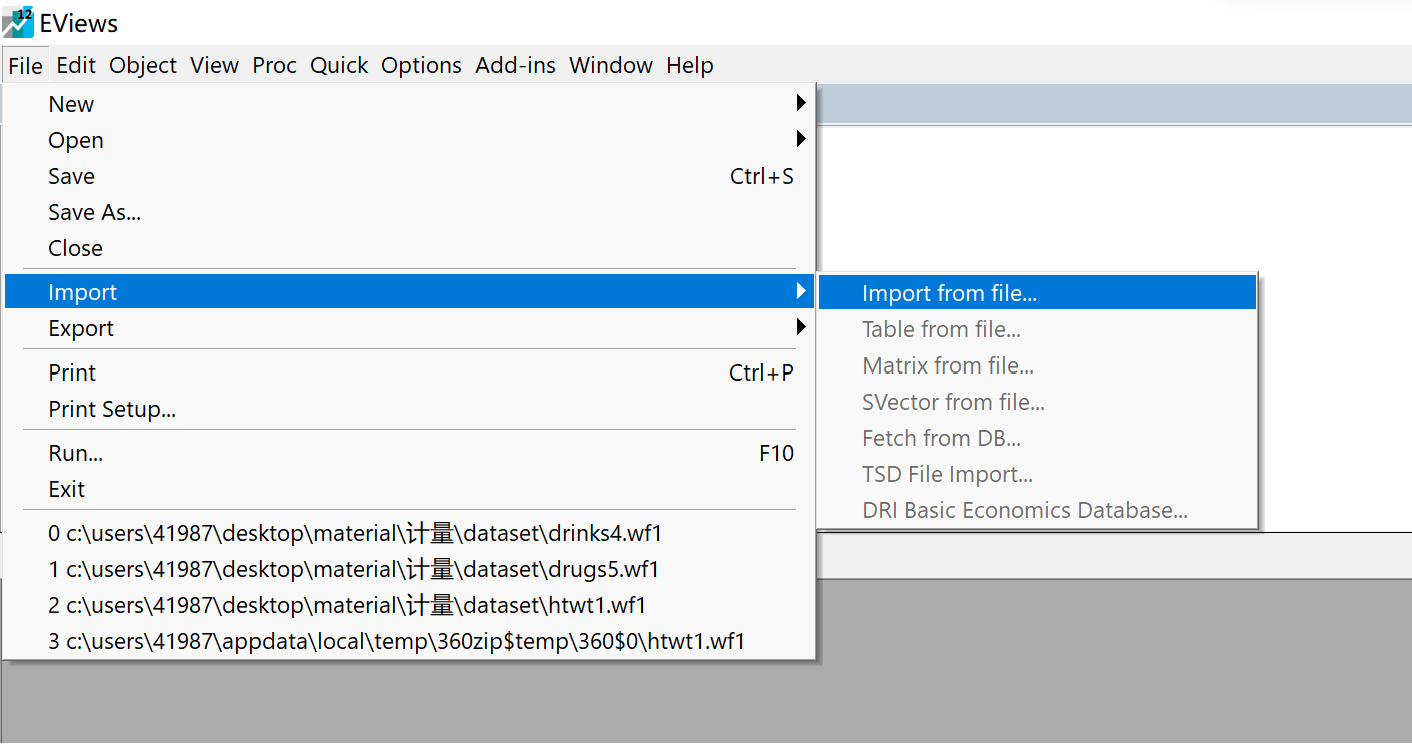


图 24: 将数据导入表格

接下来，我们对导入的数据进行相关数据特性的查看、检查

首先将要查阅的数据选中，右击选择“OPEN->as GROUP”，即可展示导入的完整数据。

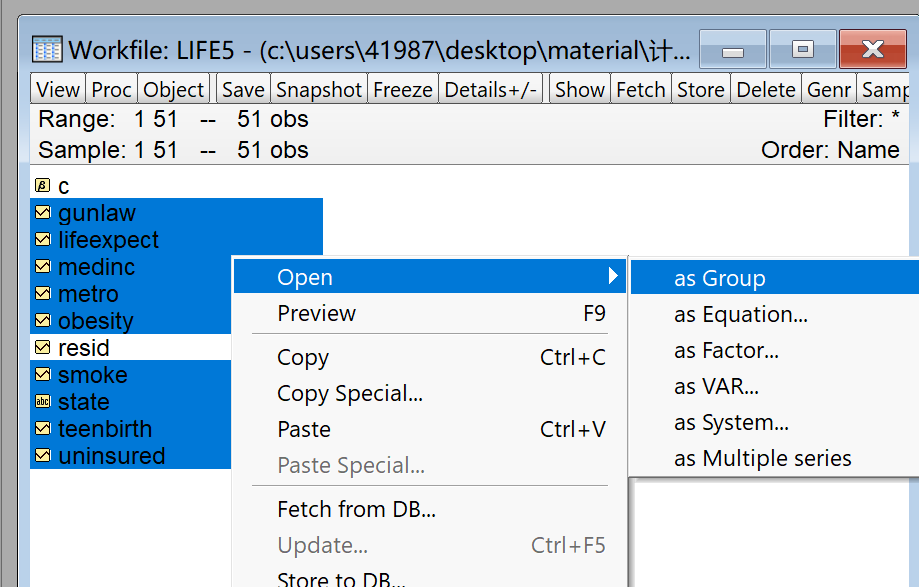


图 25：打开数据展示详情

点击后即可得到如图 5所示的数据列表——

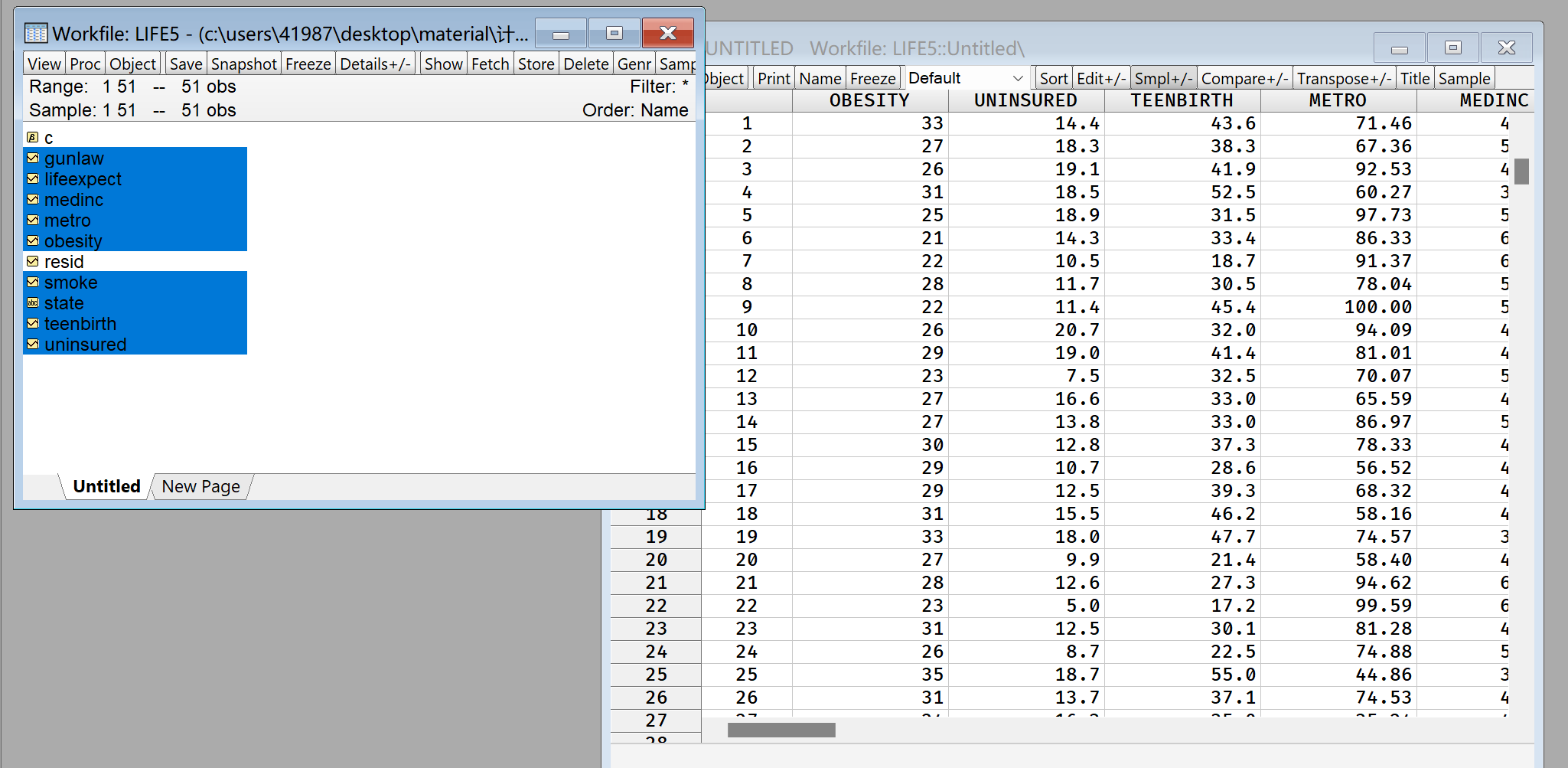


图 26：查看导入的数据

在Group视窗中，选择“View->Descriptive Stats->Common Sample”即可查看数据的详细特征（包括最值、标准差、方差等），相关步骤截图如图 6所示

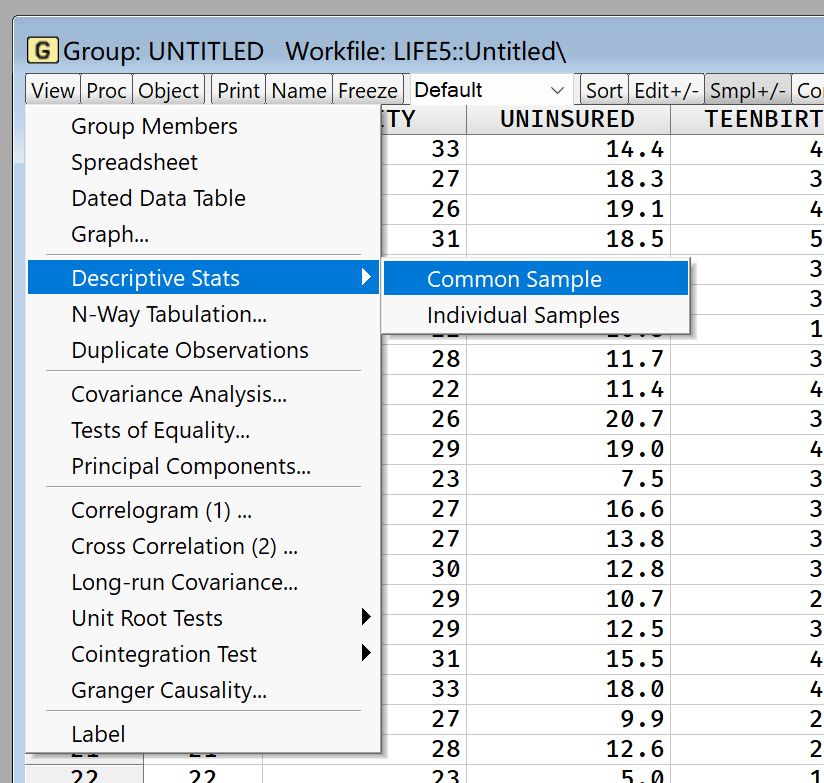


图 27：查看样本数据的特征步骤

得到的数据特征如下——

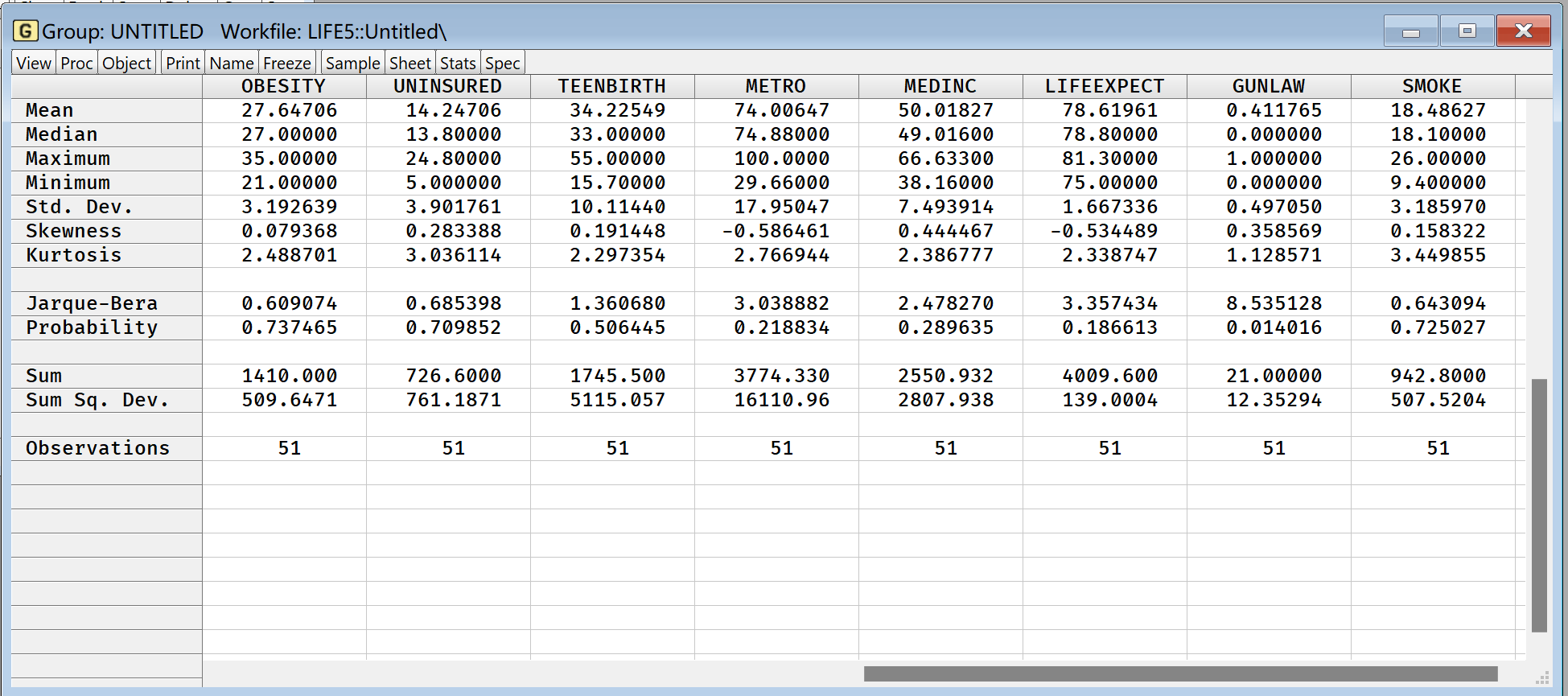


图 28：样本数据的特征

因为State变量为地区名称，与方程无关，因此在后续步骤将其剔除。

我们以前三列为例，说明数据的特性，可以看到

* 肥胖率OBESITY的最大值为35，最小值为21，标准差3.19；
* 未保险率UNINSURED的最大值为24，最小值为5，标准差为3.90；
* 早产比率TEENBIRTH的最大值为55最小值为15.7，标准差为10.11

其余各类变量的取值范围均合乎可选值内，不再赘述。数据正常，没有不可能的最大值或最小值（如负年龄）或不合理的高值（如超过100%的比率）。

检查结论：数据无明显异常。

1. 回归拟合

接下来利用Eviews进行回归拟合

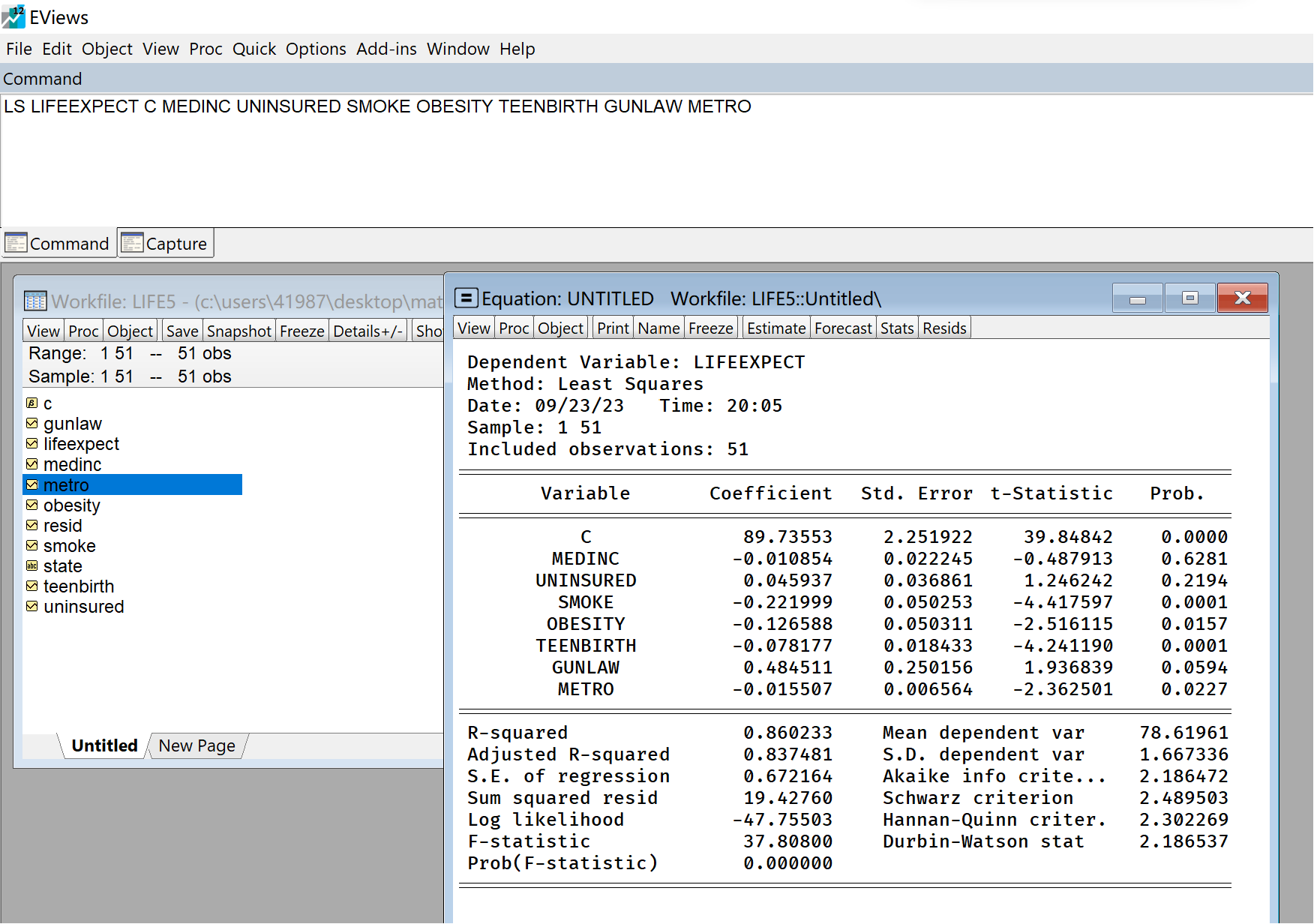


图 29：输入指令进行拟合

如图 10所示，我们输入指令“**LS LIFEEXPECT C MEDINC UNINSURED SMOKE OBESITY TEENBIRTH GUNLAW METRO**”用最小二乘法进行回归拟合。其中，指令中的LS代表使用最小二乘法的方法进行拟合，LIFEEXPECT为被解释变量平均寿命，C为回归方程的常数项，MEDINC、UNINSURED、SMOKE、OBESITY、TEENBIRTH、GUNLAW、METRO为其他的解释变量。

可以看到，我们得出的回归拟合方程为——

1. 假设检验（t检验）

本部分检验MEDINC、UNINSURED、SMOKE、TEENBIRTH 这四个参数。他们的预期符号分别是正、负、负、负。样本容量51，因此自由度为。显著性水平5%，查表得自由度为40的临界值为1.684，查表得自由度为60的临界值为1.671，那么

根据Eviews计算得出的结果（图 30红框）

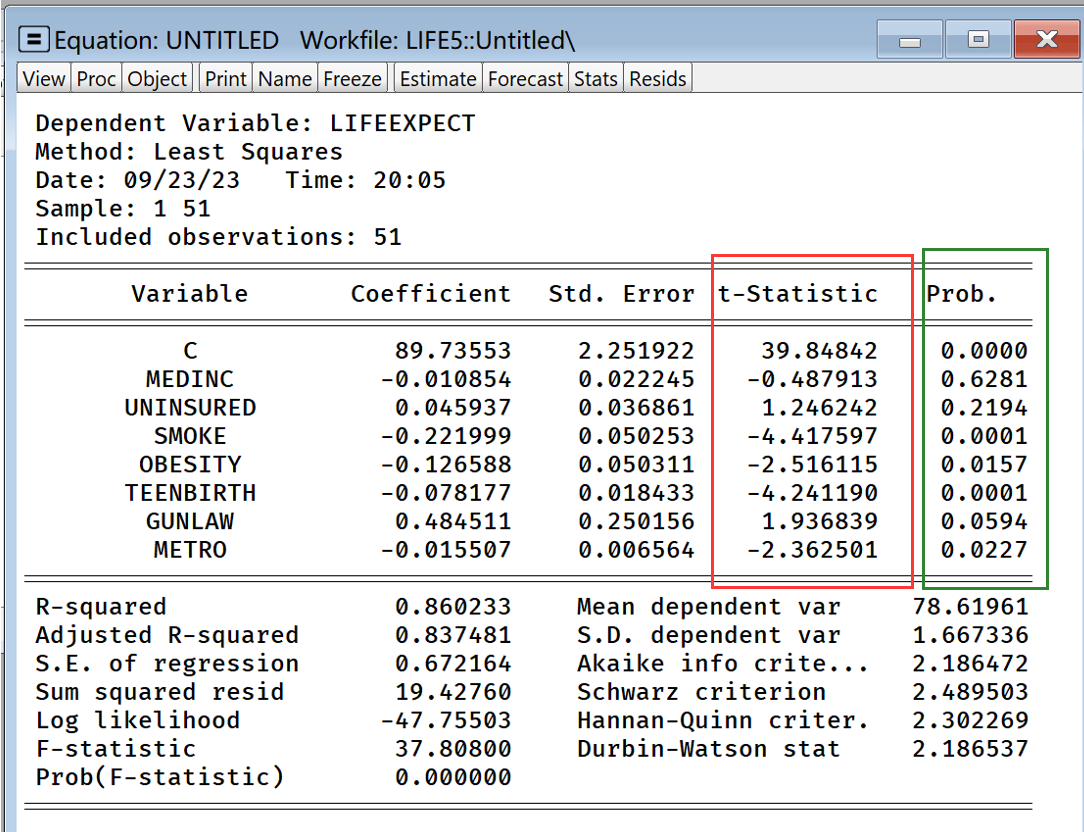


图 30：拟合结果与t值、p值

可以看到

* **对MEDINC参数**：做出假设，而，可知，因此，**不能拒绝**
* **对UNINSURED参数**：做出假设，而，可知，因此，**不能拒绝**
* **对SMOKE参数**：做出假设，而，可知，且备择假设的符号和预期符号相同（均为负数），因此**拒绝**。
* **对TEENBIRTH参数**：做出假设，而，可知，且备择假设的符号和预期符号相同（均为负数），因此**拒绝**。

1. 假设检验（p值）

本部分检验剩下的OBESITY、GUNLAW、METRO 这3个参数，相关p值已经由Eviews计算得出（图 30绿框），这三个参数的符号分别是：负、不确定、不确定

在显著性水平为5%的情况下，可得——

* **对于OBESITY参数：**做出假设，而，那么，且备择假设的符号与预期相同（都为负），因此**拒绝**。
* **对于GUNLAW参数：**做出假设，而，那么，因此**不能拒绝**。
* **对于METRO参数：**做出假设，而，那么，因此**拒绝**。

1. 总体假设检验（F检验）

本部分是在5%的显著性水平下，用整体F统计量检验回归模型是否有效。

样本容量51，因此自由度为。显著性水平5%，查表得分子自由度为7（共7个受线性约束的变量），分母自由度为40的F临界值为2.25。且对F统计量，分子自由度不变的情况下，分母自由度增大，F临界值变小，因此。

做出假设

在Eviews的拟合结果中，我们很容易得到F统计量的值

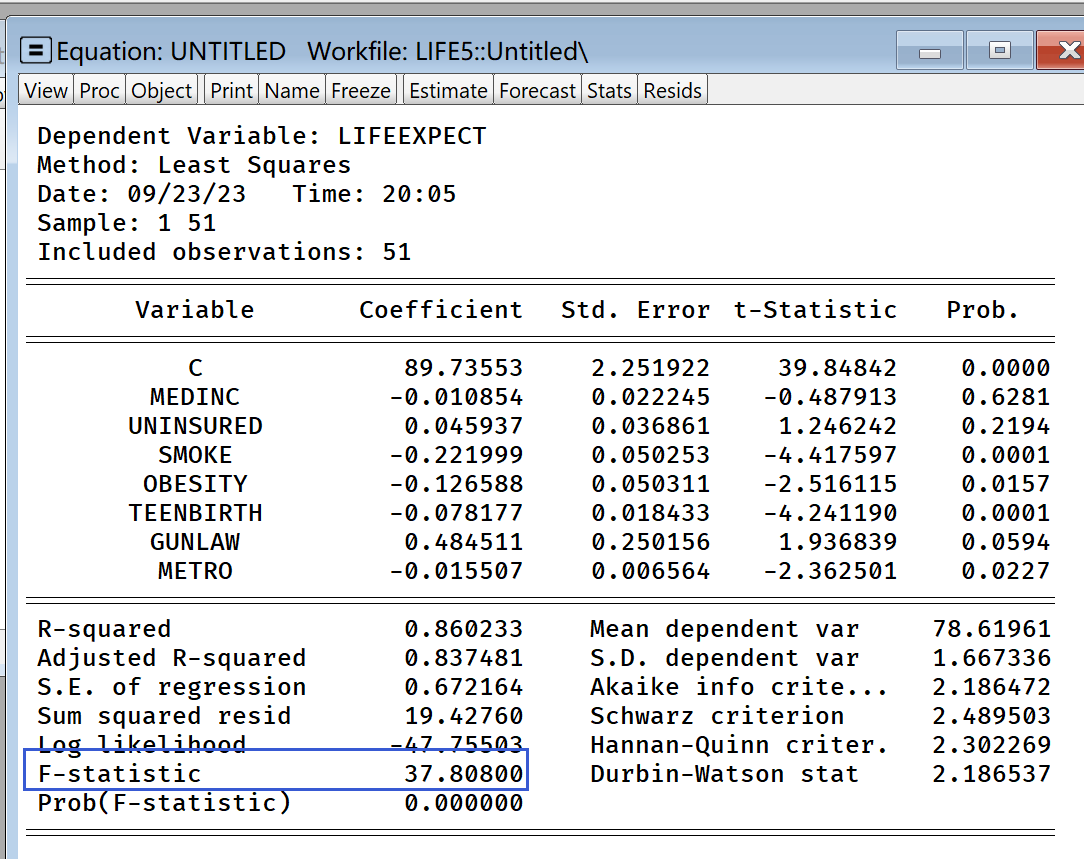


图 31：拟合结果与F值

可以看到F值为37.808，，因此，**可以拒绝**

1. 得出结论

通过第5、6、7步骤，我们已经对全部的参数以及方程的整体进行了假设检验，总体来讲拟合得到的参数基本具备一定的显著性。只有GUNLAW、UNINSURED、MEDINC显著性较差。

同时我们发现gunlaw系数的绝对值远大于smoke的系数绝对值。这是否意味着通过枪支法来保护儿童在预期寿命上的影响比减少吸烟的影响大3%呢？**显然不是**。

吸烟减少三个百分点将使预期寿命增加约0.22×3=0.66岁，比枪支法规定的0.48岁效果更大。

**九、实验结论：**

通过本次实验，我们完成了一个有关平均寿命的线性方程回归模型的拟合分析，并重点对其进行了参数检验。

我们基于Eviews软件完成了对样本数据回归方程分析，用最小二乘法拟合回归出了包括吸烟情况、枪支法保护、肥胖率、保险情况等共计7个变量与平均寿命之间的关系。在实验中，我们对样本数据进行尝试拟合，得出了完整的拟合结果——

同时，我们还发现了以下结论：

* 拟合的平均寿命回归函数的符号情况与我们假定的情况有一些不同，可以利用假设检验的方法来对其进行显著性水平的检验
* t检验和p值主要是对某一个参数的显著性情况进行检验，通过对单个参数的t值计算和对应的p值判断，我们可以评估该参数是否具有统计学上的显著差异。
* F检验主要针对拟合方程整体或者多个参数的显著性情况进行检验，如果模型中的自变量对因变量的解释能力较强，引起的回归平方和将会较大，而残差平方和相对较小。此时，F统计量的值就会较大，以达到（高于）统计显著水平，我们就可以拒绝原假设，即认为模型整体具有统计显著性。
* 假设检验不能说明参数的“重要性”
* 在求得的拟合结果中，大部分参数基本具备一定的显著性，但是GUNLAW、UNINSURED、MEDINC显著性较差。

**报告评分：**

**指导教师签字：**