**实验报告**

**实验项目名称**  美国king county地区房价的回归分析

**实　验　室**

**所属课程名称**  计量经济学

**实 验 类 型**

**实 验 日 期**  2019.5.3

**班 级**  16数据科学

**学 号**  2015110501

**姓 名**  梁力

|  |
| --- |
| **实验概述：** |
| **【实验目的】**  **使用计量实验课上所学的知识，对king county地区房价做带有虚拟变量的回归分析。**  **【实验原理】**  **Do-file Editor - toolbox.do**  **【实验内容】(详细列示)** |
| **实验步骤及结论：(记录详细分析过程)** |
| **【实验过程】（实验步骤、记录、数据、分析）**  数据集介绍  **来源[https://www.kaggle.com/harlfoxem/housesalesprediction/](https://www.kaggle.com/harlfoxem/housesalesprediction/downloads/housesalesprediction.zip/1)**  **图片包含 屏幕截图  描述已自动生成**    数据集基本统计信息    目标变量的统计信息    可以看到目标变量没有缺失值。  准备工作   1. **把日期转化为stata可以识别的日期，再用stata内置函数构建新变量year。** 2. **使用样本所在年份减去建造日期，我们可以得到构建的新变量age** 3. **从上一部分可以知道，sqftliving15和sqftlot15是针对于2015年而言的居住面积和停车场面积，所以我们需要构建一个新变量five代表是否是2015的住房价格，再构建一个变量four代表是否是2014年的住房价格。根据新变量five和four以及面积信息构建样本所在年限的面积，方便之后进行回归分析。** 4. **构建虚拟变量。从之前的数据summary中可以看出，我们已经有一个虚拟变量waterfront代表房屋是否在滨水地区。为了获得更充分的信息，我们再构建一个关于grade的虚拟变量good。以8作为分界线，大于等于8用1表示代表评分高，小于8用0表示代表评分一般。根据数据介绍中的信息我们可以知道grade是king county房屋评分系统给出的评分。我们通过构建虚拟变量可以知道房屋评分系统对房屋价格影响有多大。**     **数据预处理完毕后单个样本如图所示**  **图片包含 屏幕截图  描述已自动生成**  建模   |  |  | | --- | --- | | **解释变量** | **目标变量** | | **Bedrooms,bathrooms,floors,i.waterfront,view,condition,sqft\_above,sqft\_basement,age,living,lot,i.good**  **由于没有直觉上认为有交互效应的变量组，所以在这里不引入交互效应。** | **price** |   **【结论】（总结主要分析结果，注重经济及管理意义的分析与解释）**    **从调整后r2可以看出，解释变量能解释60%的价格变化，并且所有变量都显著。出人意料的是bedrooms的个数对于价格呈负相关，这是违反直觉的。其他系数的符号都在常识范围内。**    **两个虚拟变量都通过了显著性检验。**    **两个虚拟变量的平均边际效应的值等于建模结果中的系数。可以看到good变量对价格的边际效应很强烈，可见king county当地的房屋评分系统对房价有很大的影响。但是相比而言，是否是滨水地区的边际效应比评分好坏的边际效应大得多。**    **我们进一步分析，滨水地区评分高的房屋和非滨水地区评分低的房屋之间的价格差别。使用lincom检验。可以看到两类房屋之间有63w的价格差别。** |
| **【小结】(简要描述实验的重难点及收获)**  **通过自行寻找数据进行分析巩固了计量课上学到的对虚拟变量进行回归分析的方法。在这过程中，通过查文档的手段解决了数据预处理过程中格式转化和年份提取两个难点。** |
| **指导教师评语及成绩：**  **成绩： 指导教师签名：**  **批阅日期：** |

**附件：**

实验报告说明

**1．实验项目名称：**要用最简练的语言反映实验的内容。

**2．实验类型：**一般需说明是验证型实验还是设计型实验、综合型实验或其他实验。

**3．实验目的**：目的要明确，要抓住重点。

**4．实验原理：**简要说明本实验项目所涉及的理论和方法。

**6．实验内容（方案设计）**：实验报告极其重要的内容，概括整个实验的主要内容及过程。

**7．实验步骤（实验中涉及的记录、数据、分析）：**写明具体上述实验内容（方案）的具体实施，包括实验过程中的记录、数据和相应的分析。

**8．结论（结果）：**对实验过程中所得到主要结果进行总结，指明实验的应用价值(审视宏观政策，指导市场企业及个人行为，预测经济重要指标及风险)

**9．小结：**对本次实验的心得体会和思考。

**10．指导教师评语及成绩：**指导教师依据学生的实际报告内容，用简练语言给出本次实验报告的评价和价值。