**数据分析和计量经济学编程——教学大纲**

**2018年春季**

**基本信息：**

上课时间：周四18:40-20:30

任课教师：黄卓 办公室：国家发展研究院623 电话：010-62751424

E-mail：[zhuohuang@nsd.pku.edu.cn](mailto:zhuohuang@nsd.pku.edu.cn)

陈赟 yunchen1221@gmail.com

童晨 [tongchen1119@163.com](mailto:tongchen1119@163.com)

邱智敏 [qiuzhimin94@126.com](mailto:qiuzhimin94@126.com)

**课程简介：**

本课程主要讲授数据处理、统计分析和计量经济学编程的基本方法和工具。通过本课程学习，学生应熟悉几种常见的计量编程语言（Stata，Python，Matlab等）的基本命令，并熟练掌握其中一种语言，对常用的计量经济学模型进行编程实现。

**总评成绩：**平时作业 （60%） + 课程报告（40%）

**参考教材：**

1. Stata18讲

作者：陈传波

1. 高级计量经济学及Stata应用

作者：陈强

<http://product.dangdang.com/25205781.html>

1. Python基础教程（第二版）

作者：（挪）海特兰德

<http://product.dangdang.com/23473514.html>

1. 利用Python进行数据分析（中文版）

作者：Wes Mckinney

<http://product.dangdang.com/23368089.html>

1. Econometric and Statistical Analysis in MATLAB (2016,3rd)

作者：Kevin Sheppard

1. Applied Econometrics using MATLAB (1999, Matlab Econometrics Toolbox Documentation)

作者：James P. LeSage

**课程安排：**

|  |  |  |
| --- | --- | --- |
| **日期** | **课程** | **具体内容** |
| 第一周 | **经济学中的常见数据和计量软件**  **黄卓主讲** |  |
| **Stata模块 （邱智敏主讲）** | | |
| 第二周 | **Stata基本介绍及命令语句** | 1. Stata的安装，界面介绍，命令示例 2. 命令语句格式及分解介绍 3. 运算符: 代数运算、字符运算、关系运算、逻辑运算 |
| 第三周 | **数据处理** | 1. 数据处理: 打开／录入／导入数据，数据类型转化，数据显示格式，标签数据 2. 矩阵: 生成矩阵，矩阵运算，矩阵函数 |
| 第四周 | **绘图与基本描述统计** | 1. 绘图命令:命令结构、曲线类型、标题与坐标轴、图例 2. 几种常用的图：散点图、连线图、条形图、直方图、箱图、函数图等 3. 同时做多个图 4. 描述统计的基本命令 |
| 第五周 | **数据整理** | 1. 拆分与连接数据 2. 数据重整 3. 字符运算 |
| 第六周 | **函数与流程语句** | 1. 函数function：数学函数、字符函数 2. 分类操作by 3. 循环语句while／forvalues／foreach，嵌套循环 4. 条件语句 |
| **Python 模块 （陈赟主讲）** | | |
| 第七周 | **Python的安装与基本概念** | Python安装，算法、变量、语句、函数、字符串等基本概念的介绍，列表和元组 |
| 第八周 | **条件、循环和其他语句** | 字符串、字典的介绍，赋值语句，条件语句，循环语句 |
| 第九周 | **numpy和pandas包** | python数据分析包，numpy、pandas的介绍 |
| 第十周 | **Python爬虫** | 一个简单的爬虫，抓取网页数据。  BeautifulSoup，正则表达式 |
| 第十一周 | **数据分析** | 数据加载、存储，数据规整化（清理转换合并重塑），数据聚合与分组运算 |
| **Matlab模块 （童晨主讲）** | | |
| 第十二周 | **Matlab简介与矩阵运算** | 1. Matlab简介 2. 数据类型：数值、逻辑、字符、函数 3. 矩阵基本操作：矩阵生成、矩阵属性、矩阵索引 4. 矩阵基本运算：元素查找、排序、加减乘除、基本数学函数、基本矩阵函数 |
| 第十三周 | **Matlab编程基础** | 1. 条件控制语句 2. 循环控制语句 3. 控制程序流的其他常用命令 4. 脚本和函数构造 |
| 第十四周 | **数据可视化与OLS回归** | 1. 数据导入与导出：.xls，.csv，.mat 2. 折线图、散点图、柱状图、饼图 3. 最小二乘回归相关的计算。 |
| 第十五周 | **数值优化和参数估计** | 1. 基本的参数形式转换技巧 2. 无约束条件优化 3. 有约束条件优化：线性约束、非线性约束 4. NLS/MLE参数估计 |
| 第十六周 | **数值计算与概率统计** | 1. 概率统计：Econometrics toolbox 2. 数值计算 |