

人大经济论坛

统计软件应用培训班报名通知

在社会科学研究中定量研究趋势越来越明显的时代，高校、研究机构在从事数据处理和分析的工作中，统计软件具有非常重要的作用，掌握统计分析软件的应用，无疑有助于学术研究水平和能力的提升，有益于更有效率的工作。

为了帮助一线教师、数据分析人员更从好地进行教学、研究工作，人大经济论坛组织了统计软件应用系列现场班。此培训班将就社会科学中的定量研究与数据处理、统计软件的基础知识与统计应用等内容展开。培训包含常用软件及理论：

科目	讲师	时间	课程详情
应用计量经济学理论与应用 基础+高级	方匡南	7月3-11日	详情
Stata 暑期特训班 基础+应用+论文攻略	连玉君 王存同	7月6-19日	详情
R 从入门到精通 基础+高级	方匡南	7月20-28日	详情
数据挖掘 Clementine 及应用	李御玺	7月25-28日	详情
Matlab 基础及金融数量分析技术	谢中华 郑志勇	7月31—8月4日	详情
SAS 基础及统计分析	杨群	8月8-11日	详情
结构方程 AMOS\PLS\Lisrel 基础+论文完全攻略	张伟豪	8月3-16日	详情
购买优惠： 1. 同报2门及以上9折优惠 2. 三人及以上同时报名9折优惠 3. 论坛现场班老学员9折优惠 4. 现场班学员赠送论坛币每班次1500个			

***欢迎将此通知转告给您的同事和朋友 *盖有红色印章的培训通知来电索取**

参加培训的人员可以通过以下任意一种方式报名：

1. 电子邮件方式。请填写下页回执（请务必写明姓名、单位、联系电话），发送至以下邮箱 619492407@qq.com
2. 传真方式。请将回执内容填写完整后，发传真至 010—68478566（邓荣收）
3. 在线报名：<http://baoming.pinggu.org/PostMyInfo.aspx?id=153>

回 执（可以复印）

姓名：_____

工作单位：_____

电话：_____

邮箱：_____

参加人数：_____

参加课程名称：_____

课程详细介绍如下

讲师简介.....	3
《王存同—应用计量经济学（初阶班）》现场班大纲.....	4
《方匡南—应用计量经济学（基础班）》现场班大纲.....	5
《方匡南—应用计量经济学（提升班）》现场班大纲.....	6
《连玉君—Stata 基础班》现场班大纲.....	7
《连玉君—Stata 高级班》现场班大纲.....	8
《王存同—Stata 高级（社科论文专题）》现场班大纲.....	9
《方匡南—R 基础班》现场班大纲.....	10
《方匡南—R 提升班》现场班大纲.....	11
《李御玺—Clementine 数据挖掘及应用》现场班大纲.....	12
《张伟豪—AMOS 初中级》现场班大纲.....	14
《张伟豪—AMOS 高级》现场班大纲.....	15
《张伟豪—PLS 多变量统计分析技巧》现场班大纲.....	16
《张伟豪—Lisrel 应用班》现场班大纲.....	17
《谢中华—MATLAB 初中级》现场班大纲.....	18
《郑智勇—MATLAB 高级（金融数量分析技术）》现场班大纲.....	19
《杨群—SAS 统计分析及应用》现场班大纲.....	20

讲师简介

1. 应用计量经济学/R 讲师

方老师，厦门大学统计学副教授，耶鲁大学访问学者。曾就职于广发期货、广发基金，担任金融衍生品研究员。曾先后在《Statistics in Medicine》、《经济研究》、《统计研究》、《Communication in statistics-Simulation and Computation》、《Annals of Opreation Reaserch》、《Energy Policy》、《Journal of Applied Statistics》、《PLoS ONE》《Mathematical and Computer Modelling》等期刊发表论文 50 多篇，出版专著一本。先后主持了国家自然科学基金、国家统计局重大项目、中央高校基本业务项目、福建省社科等多个项目。

2. Stata 讲师

(1) 连玉君，经济学博士，副教授。2007 年 7 月毕业于西安交通大学金禾经济研究中心，现任教于中山大学岭南学院金融系。主讲课程为“金融计量”、“计量分析与 Stata 应用”、“实证金融”等。已在《Global Finance Journal》、《经济研究》、《管理世界》、《金融研究》、《会计研究》、《世界经济》、《统计研究》、《经济学(季刊)》等期刊发表论文 40 余篇，出版专著一部。连玉君副教授主持国家自然科学基金项目、教育部人文社科基金项目、广东自然科学基金项目各一项，并参与了多项国家自科和社科基金项目的研究工作。目前已完成 Panel VAR (1800 余行)、Panel Threshold (1200 余行)、Two-tier Stochastic Frontier (500 余行) 等计量模型的 STATA 实现程序，并编写过几十个小程序，如 xtbalance.ado、bdiff.ado 等。

(2) 王存同，副教授，毕业于北京大学（与 University of Michigan 合作培养），现工作于中央财经大学社会发展学院。兼任国家人口计生委计划生育综合改革专家组专家、北京大学社会科学定量方法培训暑期班教员、美国 PA (the Population Association)、国际 IUSSP (the International Union for the Scientific Study of Population) 会员及 IUSSP 社会科学定量方法培训专家组成员。主要从事社会定量分析、人口社会学等领域的研究与教学，曾在 SSCI/SCI 收录期刊发表定量研究论文 4 篇、在《中国社会科学》、《社会学研究》、《中国人口科学》等国内核心期刊发表定量研究论文 20 余篇，出版著作 2 部，主持国家社科基金项目 1 项、横向课题 4 项；以子课题负责人身份参与国家重大自然科学基金 2 项、国家部委横向课题 10 余项。

3. Clementine 讲师

李御玺 (Yue-Shi Lee), 国立台湾大学计算机工程博士，铭传大学计算机工程学系教授兼系主任暨所长，铭传大学数据库与数据挖掘实验室召集人，中国厦门大学数据挖掘中心顾问，中国人民大学数据挖掘中心顾问。其研究领域专注于数据仓储、数据挖掘、与数据库设计。

在其相关研究领域已发表超过 240 篇以上的研究论文，同时也是国科会与教育部多个相关研究计划的主持人。服务过的客户包括：台新银行、联邦银行、新光银行、新竹国际商业银行（现已并入渣打银行）、第一银行、永丰银行、美商大都会人寿、嘉义基督教医院、微软、全国意向顾问股份有限公司、零售业如赫莲娜 (Helena Rubinstein) 化妆品公司、特立和乐 (HOLA) 公司、航空公司如东方航空公司、汽车行业如福特 (Ford) 公司；政府行业如国税局、台北市国税局、台湾省北区国税局、台湾省南区国税局、高雄市国税局等。

4. 结构方程讲师

张伟豪，外号“Amos 亚洲一哥”，曾在台湾 SPSS 公司担任多年的统计资深顾问，有著作《SEM 论文写作不求人》与《与 SEM 共舞-首部曲》。现职为全台湾结构方程模型最佳培训机构—三星统计服务有限公司的执行长兼首席顾问。讲学过台湾和大陆 100 多所高校，包括台湾大学、政治大学、台湾师范大学、交通大学、中央大学、中兴大学、中国医药大学、中正大学、成功大学、中山大学、北京师范大学等。亦受邀于台湾析数与中国析数有限公司特约讲师，多年来所培训之学者及硕、博士生已逾五千人。学员除了台湾学术界之外，亦有来自香港、日本、澳洲等知名大学之学者，堪称结构方程模型教学领域的人气天王。

5. MATLAB 讲师

(1) 谢中华，副教授，资深 MATLAB 培训师，十多年 MATLAB 编程经验，已出版书籍《MATLAB 统计分析与应用：40 个案例分析》和《MATLAB 从零到进阶》。现于天津科技大学数学系任教，长期从事 MATLAB 相关课程的教学与培训。精通 MATLAB、SAS、R 语言等软件，有着扎实的理论基础和丰富的实战经验。

(2) 郑志勇，资深 Matlab 讲师，10 年 Matlab 编程经验，金融工程师（产品设计），编著书籍《运筹学与最优化 MATLAB 编程》《金融数量分析：基于 MATLAB 编程》

6. SAS 讲师

杨群，资深 SAS 培训师，商业银行信用评分模型专家，毕业于北京大学光华管理学院商务统计及经济计量系，现供职于平安银行总行零售信贷管理部，负责个人贷款风险评分卡开发应用和新资本协议实施项目。曾在国家核心期刊上发表文章。业务特长：开发应用风险评分卡、营销响应模型、以及新资本协议风险计量及策略应用。

课程大纲简介

王存同一应用计量经济学（初阶班）

第一部分 社会科学研究导论

第一讲 统计方法在社会科学研究中的应用

第 1 节 社会现象的性质与社会科学研究的任务 第 2 节 社会科学研究的逻辑

第二讲 测量与抽样

第 1 节 变量测量与测量层次 第 2 节 抽样方法 第 3 节 定量分析方法的分类框架

第三讲 社会研究中的数据

第 1 节 横截面数据 第 2 节 纵贯数据 第 3 节 数据管理

第四讲 描述性统计

第 1 节 统计表与统计图 第 2 节 集中趋势的测量 第 3 节 离散趋势的测量

第二部分 基本统计学基础

第五讲 概率分布

第 1 节 概率与概率分布 第 2 节 几种常见的概率分布

第 3 节 随机过程与抽样分布 第 4 节 中心极限定理与统计推论

第六讲 统计推论：参数估计

第 1 节 点估计 第 2 节 置信区间的选择 第 3 节 样本量的选择

第七讲 统计推论：假设检验

第 1 节 假设检验的要素

第 2 节 假设检验的步骤

第 3 节 小样本的假设检验

第 4 节 假设检验中的错误类型

第八讲 从单变量分析到双变量分析

第 1 节 为什么需要双变量分析 第 2 节 组间比较和 t 检验 第 3 节 列联表分析

第九讲 相关系数与简单线性回归模型

第 1 节 散点图与相关系数 第 2 节 线性模型：估计与解释 第 3 节 线性模型的统计推论

方匡南—应用计量经济学（基础班）

专题名称	授课内容
第 1 讲 （3 小时） 计量经济学入门介绍	计量经济学入门介绍 经典计量经济学的步骤与应用 计量软件 Eivews 介绍
第 2 讲 （3 小时） 一元线性回归模型	一元线性回归的假设 一元线性回归的参数估计 一元线性回归的统计检验 回归的预测问题 案例分析：我国人均消费支出影响因素
第 3 讲 （3 小时） 多元线性回归模型	多元线性回归的假定 多元线性回归的参数估计 多元线性回归的统计检验 多元线性回归的预测问题 案例分析：中国税收收入增长因素
第 4 讲 （3 小时） 多重共线性问题 异方差性问题	多重共线性和异方差的后果 多重共线性和异方差的检验 多重共线性和异方差的克服 逐步回归分析 岭回归分析 案例分析： 中国电信业务总量模型 中国省市农作物产值与播种面积 中国消费函数案例分析
第 5 讲 （3 小时） 序列相关性问题 虚拟变量模型	序列相关的后果 序列相关的检验 序列相关的克服 虚拟变量模型 案例分析：中国农村居民消费支出模型
第 6 讲 （3 小时） 动态宏观模型 综合案例与论文写作指导	分布滞后模型 自回归模型 Granger 因果关系检验 综合案例分析 经济实证论文写作指导

方匡南—应用计量经济学（提升班）

专题名称	授课内容
第1讲（3小时） 非线性回归模型 及其应用	非线性回归假定 非线性回归的参数估计（非线性最小二乘法） 非线性回归的检验 非线性优化求解 案例分析：城镇居民人均食品消费支出非线性模型
第2讲（3小时） 离散选与限制因变量模型 及其应用	Probit 回归分析；logistic 回归分析 Poisson 回归分析；Tobit 回归分析 案例分析： 1. logistic 回归在教学方法效果建模应用 2. Poisson 回归在事故发生数建模应用 3. Tobit 回归在婚外性行为的建模应用
第3讲（3小时） 联立方程及其应用	联立方程的概念与分类 联立方程的识别 联立方程的参数估计（ILS, 2SLS、IV） 案例分析： 凯恩斯收入决定模型的 ILS 估计 货币供给函数的 2SLS 估计 恰好识别时 ILS 和 2SLS 等价性的验证
第4讲（3小时） VAR 模型与协整模型 及其应用	VAR 模型的含义及特点；VAR 模型的估计 脉冲响应函数；方差分解 基于 VAR 模型的 granger 检验 非平稳时间序列与单位根检验 协整检验；误差校正模型 案例分析： 宏观经济 VAR 模型的实例分析 协整与误差校正模型案例分析
第5讲（3小时） 面板数据模型及其应用	面板数据的概念与分类 静态面板的固定效应分析 静态面板的随机效应分析 Hausman 检验 动态面板模型的估计 案例分析： 航空公司成本函数的静态面板分析 宏观经济的动态面板分析
第6讲（3小时） 分位数回归及其应用 综合案例与论文写作指导	分位数回归含义及特点 分位数回归的估计 分位数回归的注意事项 案例分析： 分位数回归在收入分配中的应用 综合案例与论文写作指导

连玉君—Stata 基础班

专题名称	授课内容
第 1 讲 （3 小时） Stata 简介和数据处理	数据的导入和导出 数据的横向合并和纵向追加 执行指令和基本统计分析 do 文件和 log 文件的使用 帮助文件的使用和外部命令的获取 重复样本值、缺漏值和离群值的处理 文字变量的处理
第 2 讲 （3 小时） Stata 程序	局域暂元和全局暂元（local, global） 控制语句（条件语句、循环语句） Stata 中的各类函数 分组回归分析 范例：盈余管理程度的估算、现金持有调整系数的估算
第 3 讲 （3 小时） 模型的设定和解释	线性回归模型回顾（OLS、GLS） 虚拟变量的使用和解释 交乘项的使用和边际效果解释 组间系数差异检验 模型的设定和筛选 Bootstrap、Jackknife 及稳健性标准误的获取 估计结果的呈现和分析
第 4 讲 （3 小时） 内生性问题及估计方法	工具变量法（IV）和广义矩估计法（GMM）简介 倍分法（Difference in Difference） 处理效应模型（Treatment Effect Model）
第 5 讲 （3 小时） 静态面板数据模型	静态面板模型：固定效应和随机效应 异方差和序列相关（Bootstrap 稳健型标准误的获取） 包含内生变量的固定效应模型 实证分析中的常见问题
第 6 讲 （3 小时） 动态面板数据模型 面板 VAR 模型	一阶差分 GMM 估计量（FD-GMM） 序列相关检验和过度识别检验（Sargan 检验） 面板 VAR 模型

连玉君—Stata 高级班

专题名称	授课内容
第 1 讲 (3 小时) 门槛面板和联立方程	面板门槛模型 (Panel Threshold Model) 面板联立方程组模型 (Panel Simultaneous Equations Model)
第 2 讲 (3 小时) 空间面板模型	空间面板固定效应模型 (Spatial Panel Fixed Effects Models) 空间动态面板模型 (Spatial Dynamic Panel Data Models)
第 3 讲 (3 小时) 随机边界分析 I	传统的 SFA 模型 (Traditional SFA) SFA 的模型设定和估计方法 异质性 SFA 模型 (Heterogeneity SFA)
第 4 讲 (3 小时) 随机边界分析 II	面板 SFA 模型 (Panel Data SFA) 双边 SFA 模型 (Two-tier SFA) 范例: 卢洪友, 连玉君, 卢盛峰 (2011). 中国医疗服务市场中的信息不对称程度测算. 经济研究, (4): 94-106.
第 5 讲 (3 小时) 离散选择模型	Logit 模型的设定和估计方法 多元 Logit 模型 (Multinomial Logit) 有序 Logit 模型 (Ordered Logit)
第 6 讲 (3 小时) 自抽样和蒙特卡洛模拟	Bootstrap 的原理和 Stata 实现 Bootstrap 组间系数差异检验 Bootstrap 获取复杂统计量的临界值 Monte Carlo 的基本原理 Monte Carlo 应用实例: 内生性偏误的后果
第 7 讲 (3 小时) 内生性专题 I	Heckman 选择模型 (Heckman Selection Models) 处理效应模型模型 (Treatment Effect Models) 多元处理效应模型 (Multinomial Treatment Effects Models)
第 8 讲 (3 小时) 内生性专题 II	倾向得分匹配分析 (Propensity Score Matching, PSM) 断点回归 (Regression Discontinuity Designs, RDD)

第一讲 线性回归知识点回顾 (Linear Regression)

1.1 线性回归基础

基础统计学知识回顾；一元线性回归·模型；多元线性回归模型

1.2 线性回归高级

引进虚拟变量的回归；因变量与自变量间违背基本假设条件；
奇异点或重要影响点的处理；测量误差对回归的影响；
缺失值的处理；|内生性和工具变量估计·

第二讲 二分类 Logit/Probit 模型 (Binary Logit/Probit Models)

2.1 一般线性模型·

2.2 二分因变量 logit 模型 2.3 二分因变量 Probit 模型 2.4 模型回归系数的解释

第三讲 多分类 Logit/Probit 模型及条件 Logit 模型

(Multinomial Logit / Probit Models, Conditional Logit Models)

3.1 多分类 Logit 模型

3.2 条件 Logit 模型 3.3 多分类 Logit 模型与条件 Logit 模型比较

第四讲 定序 Logit / Probit 模型 (Ordered Logit / Probit Models)

4.1 定序（积累）Logit 模型 4.2 定序（积累）Probit 模型·

4.3 定序模型的等比例发生比假设与检验 4.4 连续比率模型·

第五讲 计数回归（泊松回归，负二项回归及零膨胀模型）

(Poisson regression,
Negative Binary regression, Zero Inflated Poisson regression)

5.1 Poisson 回归模型· 5.2 过离散和负二项回归模型· 5.3 零膨胀模型·

第六讲 删截数据回归

(Censored and Truncated Regression Models, including Tobit models, Heckman models)

6.1 Tobit 模型· 6.2 样本选择模型

第七讲 社会事件史分析 (Event History Analysis)

7.1 事件史分析的数据结构及建立

7.2 生存函数和风险函数 7.3 事件史分析中的参数模型·

7.4 比例风险模型 7.5 离散时间风险模型 7.6 Piecewise Exponential 模型

第八讲 面板数据模型

8.1 整群抽样 8.2 追踪/面板数据模型

第九讲 分层模型 (Multilevel linear/logit models)

9.1 分层模型的原理

9.2 一般模型及其简单子模型

9.3 应用 HLM 软件进行两层线性模型操作示范

9.4 一般化线性分层模型 9.5 如何报告 HLM 的结果 9.6 Z 检验与 t 检验

方匡南—R 基础班

专题名称	授课内容
第 1 讲 （3 小时） R 语言入门介绍	R 语言介绍 编辑软件 Rstudio 使用 R 程序包的载入与使用 数据对象及运算 (向量、矩阵、数组、列表与数据框处理)
第 2 讲 （3 小时） 数据读入、读出与 R 基本编程	R 数据读入与读出 (读入 txt、xls、SPSS、SAS、stata 以及数据库文件) R 函数编写 R 的条件与循环函数 高效编程技巧介绍 利用 R 做迭代优化求解 案例分析：非线性回归
第 3 讲 （3 小时） 数据预处理与缺失值处理	数据预处理 缺失值处理
第 4 讲 （3 小时） 统计模拟与随机抽样	随机数生成 常用统计方法的蒙特卡洛模拟 中心极限定理的模拟 随机抽样 案例：蒙特卡罗模拟的应用
第 5 讲 （3 小时） 探索性分析与作图	单变量数据分析与作图 双变量数据分析与作图（列联表） 多变量数据分析与作图（多变量相关系数矩阵） 案例分析： 1. 统计作图在调查数据中的应用 2. 统计作图在临床医学中的应用
第 6 讲 （3 小时） 假设检验	参数假设检验 非参数假设检验 案例分析： 1. 假设检验在医学临床试验中的应用 2. 假设检验在市场调查中的应用
第 7 讲 （3 小时） 相关与线性回归分析	相关分析 一元线性回归 多元线性回归 逐步回归 案例分析：中国税收收入增长案例分析
第 8 讲 （3 小时） 综合案例分析与答疑	综合案例分析 答疑与互动

方匡南—R 提升班

专题名称	授课内容
第 1 讲 (3 小时) R 高级作图	R 做地图 Ggplot 的使用 R 做三维立体图 R 做词云分析
第 2 讲 (3 小时) 聚类分析与分类分析	相异度的度量 系统聚类 K-means 聚类 决策树分类 KNN 分类方法 案例分析： 我国城市宏观经济发展的聚类分析 Fisher 鸢尾花(Iris)案例分析
第 3 讲 (3 小时) 支持向量机分析 随机森林分析 神经网络分析	支持向量分类 支持向量回归 随机森林分类 随机森林回归 神经网络分析 案例分析：信用卡违约分析的案例分析
第 4 讲 (3 小时) 主成分分析 因子分析	主成分分析 利用主成分进行综合评价 因子分析 案例分析： 主成分分析在教育中的应用 因子分析在经济中的应用
第 5 讲 (3 小时) 非线性回归与分位数回归	非线性的迭代优化求解 非线性最小二乘法 非线性极大似然估计
第 6 讲 (3 小时) 广义线性模型	Probit 回归分析 logistic 回归分析 Poisson 回归分析 Tobit 回归分析 案例分析： 1. logistic 回归在教学方法效果建模应用 2. Poisson 回归在事故发生数建模应用 3. Tobit 回归在婚外性行为的建模应用

李御玺—Clementine 数据挖掘及应用

时间	主题
3 小时	<p>课程目标: 说明企业如何利用顾客关系管理来进行营销活动</p> <p>数据挖掘基本介绍:</p> <p>1. 数据挖掘技术的功能分类 2. 数据挖掘技术的绩效增益 3. 数据挖掘技术的产业标准 4. 数据挖掘基本观念与实际应用解说</p> <p>实务操作:</p> <p>1. 环境介绍 IBM SPSS Modeler 2. 基本操作 Getting Started with IBM SPSS Modeler 3. 创建项目图标 Creating a Modeler Project and Diagram</p>
3 小时	<p>课程目标: 说明企业如何利用数据前处理技术来进行资料的重整, 建置数据仓储, 以做为数据挖掘之输入</p> <p>数据预处理:</p> <p>1. 数据挖掘技术的流程- CRISP DM 2. 数据前处理(Data Preprocessing)技术 3. 字段选择(Attribute Selection) *数据整合(Data Integration) 4. 资料清理(Data Cleansing): *错误值(Wrong Value), *离群值(Outlier), *遗失值(Null Value) 5. 字段扩充(Attribute Enrichment): *内/外部数据的扩充方法 6. 数据编码(Data Coding): *数据转换(Data Transformation), *资料精简(Data Reduction), *记录精简(Record Reduction), *字段精简(Attribute Reduction) *域值精简(Attribute Value Reduction),</p> <p>实务操作: (使用银行数据集为实例)</p> <p>Defining a Data Source Exploring a Data Source Managing Wrong Values/Outliers/Null Values</p>
3 小时	<p>课程目标: 讲解企业如何利用关键变量发掘技术来发掘对项目目标有效之关键变量, 以做为数据挖掘之输入变量</p> <p>1. 训练数据与测试数据的产生方法 2. 关键变量(Key Attributes)发掘技术 *利用检定(Test)选择关键变量 *利用决策树(Decision Tree)选择关键变量</p> <p>实务操作: (使用电信数据集为实例)</p> <p>Creating Training & Test Dataset Using Test for Variable Selection Using the Decision Tree for Variable Selection</p>

3 小时	<p>课程目标: 说明企业如何利用分类技术之决策树及简单贝式来建立交叉销售(Cross-Selling)模型, 以提升公司获利</p> <p>1.分类之简单贝式(Naive Bayes) 2.分类之决策树(Decision Tree)</p> <p>实务操作: (使用电信数据集为实例)</p> <p>Constructing a Decision Tree Model</p> <p>Assessing a Decision Tree</p> <p>Manual Tree Growth</p> <p>Create a Score Data Source</p> <p>Exploring a Scored Table</p>
3 小时	<p>课程目标: 说明企业如何利用分类技术之罗吉斯回归及类神经网络来建立信用评分(Credit Scoring)模型, 以降低公司损失</p> <p>1.分类之类神经网络(Neural Network) 2.分类之罗吉斯回归(Logistic Regression)</p> <p>实务操作: (使用银行数据集为实例)</p> <p>Running the Regression Node; Interpreting a Regression Model</p> <p>Training a Neural Network; Selecting Neural Network Inputs</p> <p>Comparing Models with Summary Statistics; Comparing Models with ROC Charts</p>
3 小时	<p>课程目标: 说明企业如何利用分群技术之阶层式及 k-平均分群来建立客群分析模型, 使银行针对不同客群, 采用不同之营销策略, 让银行获利最大化</p> <p>1.分群(Clustering)之 k-平均分群(K-Means Clustering)</p> <p>2.分群(Clustering)之两阶段分群(Two-Step Clustering)</p> <p>实务操作: (使用银行数据集为实例)</p> <p>Exploring & Filtering Analysis Data</p> <p>Setting Cluster Options</p> <p>Constructing a Clustering Model</p> <p>Specifying the Segment Count</p> <p>Exploring & Profiling Segments</p>
3 小时	<p>课程目标: 说明企业如何利用关联及序列分析技术来建立交叉销售(Cross-Selling)及提升销售(Up-Selling)模型, 以提升公司获利</p> <p>1.关联分析(Association Analysis) 2.序列分析(Sequence Analysis)</p> <p>实务操作: (使用零售数据集为实例)</p> <p>Consolidating Categorical Inputs</p> <p>Market Basket Analysis</p> <p>Sequence Analysis</p>
3 小时	<p>课程目标: 说明企业如何利用预测(Prediction)技术之线性回归、回归树及类神经网络来建立数值预测模型-如预测客户之年收入, 以利公司设计营销活动</p> <p>1.线性回归(Linear Regression)</p> <p>2.回归树(Regression Tree)</p> <p>3 类神经网络(Neural Network)-非线性回归</p> <p>实务操作: (使用银行数据集为实例)</p> <p>Review and Set the Regression Node</p> <p>Review and Set the Decision Tree Node; Review and Set the Neural Network Node</p>

张伟豪—AMOS 初中级

时间	主题	内容
第一天: Amos 操作完全攻略		
3 小时	SEM 基本注意事项 1. SEM 的基本原理 2. SEM 基本流程; 3. SEM 术语 4. 合理的样本量; 5. SEM 问卷设计要领	了解 SEM 基本的注意事项及要求, 才不会犯了不该犯的错误, 并在一开始的构念及问卷设计上不会犯错, 之后分析才会顺利。
3 小时	Amos 操作环境与模式建立 1. Amos 环境介绍 2. 建立路径图 3. 数据读取 4. 数据处理(分析)	Amos 软件完整的操作训练, 带领学员将一步步的将所有的指令完整的走一遍, 并提示各项快速的操作及绘图技巧, 让学员在 6 小时后, 操作软件毫无障碍。
第二天: SEM 论文分析标准流程		
3 小时	1. 验证式因素分析(CFA) 2. 多元常态及例外值检定 3. 一阶验证性因素分析 4. 二阶验证性因素分析 5. 组成信度 (CR); 6. 变异数萃取量 AVE	CFA 为 SEM 的基本元素, 本内容以实务的方式探讨 CFA 分析常碰到的问题, 包括 CFA 非正定、模型拟合度不佳、模型修正及所代表意涵等等的原因及解决之道。
3 小时	1. 检定 SEM 模型适合度 2. 评估整体模型拟合 3. 违犯估计的检 AMOS 拟合指标意义及应用 1. 绝对拟合指标; 2. 增值拟合指标 3. 精简拟合指标; 4. 竞争拟合指标	SEM 分析拟合度指标所代表的意义及解读。造成拟合度不佳的原因及处理因应之道。Amos 所有报表输出的意义及了解如何正确的在论文中报告应有的拟合度指标。
第三天: 中介变量		
3 小时	中介变量的原理说明 中介变量评估方法的比较 因果法; 直接与间接效果法	探讨目前最新的中介效果分析方法, 从 old school 到 state of art,
3 小时	系数差异法(Difference in coefficients) 系数乘法; 赖区间法	加以探讨并分析介绍其优劣
第四天: 调节变量		
3 小时	调节变量的定义与实务意涵, 观察调节变量的验证方法剖析 1. two-way ANOVA 2. 分群回归及费雪 Z 转换 3. 阶层式回归(Hierarchical regression)	SEM 潜在变量的调节大概是最难检定的部分, 本内容将介绍多个分析
3 小时	潜在调节变量验证方法剖析 1. 调节为类别变量 2. 调节为连续潜在变量 Ping' s single indicator and Ping' s indicant product 3. 中介的调节效果 vs. 调节的中介效果	分法, 并带学员一步步的完成分析。

张伟豪—AMOS 高级

时间	主题	内容
3 小时	<ol style="list-style-type: none"> 1. 遗漏值的处理 2. 分析数据的多元常态与例外值检定 3. 模型配适度显著的检定与修正 (Bollen-Stine 检定) 	多元常态及例外值的检定，在 SEM 分析中希望检定结果为不显著 ($P > 0.05$)，但由于 SEM 为大样本分析，P 值往往都会显著，介绍 Bollen-Stine 修正方法。
3 小时	<ol style="list-style-type: none"> 1. 数据非常态的处理 -Bootstrap 的介绍与使用 2. 结构模型的二阶段准则 (Anderson and Gerbing, 1988) 	Bootstrap 的使用时机及解读 SEM 模型估计配适度好坏的评估及处理
3 小时	<ol style="list-style-type: none"> 1. SEM 共线性的判断及处 2. 模型中潜在变项单一指标测量 3. SEM 处理顺序尺度及类别尺度的研究设计 4. 最佳模型的搜索 	共线性的检查及处理方式 非连续型尺度量表的处理 小样本的处理方式 找出最佳模型
3 小时	<p>结构方程模型检定力的计算及样本数的决定</p> <ol style="list-style-type: none"> 1. 检定力 (Power) 的介绍及重要性 2. SEM 检定力与样本数的计算 3. 检定每个估计值的 Power (Satorra & Sarris, 1985) 4. 检定 SEM 整体模型的 Power <p>结构方程模型的关键报告</p> <ol style="list-style-type: none"> 1. 概念模型及统计模型的叙述 2. 样本数据的细节 3. 结果说明； 4. 事后修正 5. 模型配适的额外信息 6. 解释与推论 	<p>了解何谓检定统计力，避免犯型 II 错误，样本数除了可依经验法则外，如何计算本节说分明。</p> <p>论文写作期间，那些重点是一定要报告的，本节会以检核表的方式介绍呈现。</p>
3 小时	<ol style="list-style-type: none"> 1. 研究设计 潜在构面探讨； 问卷设计 2. 预试 资料检查 3. 资料预检 叙述性统计； 遗漏值处理 交叉分析； 适合度检定； 同构型检定 	问卷的设计必须先了解可以做何种分析再进行设计，接下来问卷回收后做一些资料的处理，以确保数据的严谨与完整性，并且透过叙述性统计了解样本特性。
3 小时	<ol style="list-style-type: none"> 1. 验证式因素分析 一阶及二阶验证式因素分析 多元常态及例外值检定 SEM 二阶段准则 违反多元常态的因应 2. 共同方法变异检定 哈门氏单因子检定 	在 SEM 分析前先有 CFA 最常出现的问题一一解决，以奠定结构模型的基础。此外为了检测测量工具是否有误差，介绍共同方法变异以提高严谨性。

	CFA 单因子及多因子检定	
第四天:		
3 小时	1. 信度与效度 组成信度 (CR) 平均变异数萃取量 (AVE) 收敛效度 (Convergent validity) 区别效度 (Discriminated validity)	信效度的评估是 SSCI 期刊中必备的分析报表。如何呈现与解释是 SEM 论文不可或缺的内容。
3 小时	1. SEM 模型估计 配适度指标的 参数的解读 2. 模型不变性检定 交叉效度 检定统计力及样本数计算	结构模型的好坏攸关研究架构与样本的配适情形。因此论文最后阶段在于 SEM 模型的评估。交叉效度可以显示模型的稳定性。

张伟豪—PLS 多变量统计分析技巧

时间	主题
3 小时	Introduction to SEM, PLS and SmartPLS 1. SEM, PLS 的基本原理相似与相异; 2. SmartPLS 的介绍 SmartPLS 实例操作 1. Create Project and Import Data 2. Draw the Model; 3. Estimate the Model 4. Examine the Results; 5. How to Handle Missing Data
3 小时	Basic Linear Path Modeling using SmartPLS 1. PLS Inner model and Outer Model 2. Reflective VS. Formative Key Issues in Formative and Reflective Construct Measurement 1. Uni-dimensionality; 2. single indicator reliability 3. 信度 4. 收敛效度 5. 区别效度
3 小时	二阶 CFA 的 PLS 模型及结构模型的评估 1. 二阶 CFA 在 PLS 的应用 2. PLS Model Explanatory Power; 3. PLS Model Predictive Power 连续型干扰及控制变量 1. 干扰效果于 PLS 的应用; 2. 控制变量于 PLS 的应用
3 小时	Multiple Groups Comparison 1. 1. 数据的输入; 2. 2. 图形的调整; 3. 3. 系数的比较 结论与讨论 PLS 论文的回顾

张伟豪—Lisrel 应用班

时间	主题
第一天	
3 小时	<p>Introduction to SEM, PLS and SmartPLS</p> <ol style="list-style-type: none"> 1. SEM, PLS 的基本原理相似与相异 2. SmartPLS 的介绍 <p>SmartPLS 实例操作</p> <ol style="list-style-type: none"> 1. Create Project and Import Data 2. Draw the Model 3. Estimate the Model 4. Examine the Results 5. How to Handle Missing Data
3 小时	<p>Basic Linear Path Modeling using SmartPLS</p> <ol style="list-style-type: none"> 1. PLS Inner model and Outer Model 2. Reflective VS. Formative <p>Key Issues in Formative and Reflective Construct Measurement</p> <ol style="list-style-type: none"> 1. Uni-dimensionality 2. single indicator reliability 3. 信度 4. 收敛效度 5. 区别效度
第二天	
3 小时	<p>二阶 CFA 的 PLS 模型及结构模型的评估</p> <ol style="list-style-type: none"> 1. 二阶 CFA 在 PLS 的应用 2. PLS Model Explanatory Power 3. PLS Model Predictive Power <p>连续型干扰及控制变量</p> <ol style="list-style-type: none"> 1. 干扰效果于 PLS 的应用 2. 控制变量于 PLS 的应用
3 小时	<p>Multiple Groups Comparison</p> <ol style="list-style-type: none"> 1. 1. 数据的输入 2. 2. 图形的调整 3. 3. 系数的比较 <p>结论与讨论</p> <p>PLS 论文的回顾</p>

谢中华—MATLAB 初中级

第一天 MATLAB 编程基础	
MATLAB 简介	课程设置, 教材介绍; MATLAB 的安装与启动
MATLAB 用户界面	介绍 MATLAB 开发环境和用户界面的主要功能.
变量、表达式、数组运算	MATLAB 指令, 创建和读取变量中的数据, 如何进行数组运算。
MATLAB 程序设计	先介绍脚本文件和函数文件; 之后介绍 MATLAB 的程序流程控制方法以及程序调试方法; 介绍匿名函数、子函数和嵌套函数的概念。
模块的开发与封装发布	介绍 MATLAB 中程序的封装与发布; 制作 .exe 可执行文件; 制作 dll 动态链接库文件; 制作 Excel 插件

第二天 MATLAB 数据读写、可视化和统计	
从 TXT 文件中读取数据	本节以 TXT 文件为例, 介绍从文本文件中读取数据的方法。包括界面操作和编程操作: 利用界面操作导入 TXT 文件; 调用函数读取数据
把数据写入 TXT 文件	介绍把数据写入文本文件的方法。 调用 dlmwrite 函数写入数据; 调用 fprintf 函数写入数据
从 EXCEL 文件中读取数据	介绍从 Excel 文件中读取数据的方法。 利用界面操作导入 Excel 文件; 调用 xlsread 函数读取数据
把数据写入 EXCEL 文件	介绍把数据写入 Excel 文件的方法
图形对象与图形对象句柄	本节将对图形对象与图形对象句柄、二维绘图、三维绘图、图形的打印和输出和动画制作等内容作详细介绍。 1. 图形对象与图形对象句柄 2. 二维图形绘制 3. 三维图形绘制 4. 图形的打印和输出 5. 动画制作
数据预处理	数据导入到 MATLAB 之后, 通常需要对数据进行一些预处理, 本节介绍数据预处理的 MATLAB 实现。 1. 数据平滑处理 2. 数据标准化变换 3. 数据极差归一化变换
分布、随机数及蒙特卡洛模拟	主要介绍统计工具箱中产生不同概率分布的函数功能, 以及产生各种分布随机数, 以及基于随机数的蒙特卡洛模拟。 概率分布; 统计工具箱中的分布; 产生随机数 随机数流; 任意分布的随机数; 蒙特卡洛模拟
参数估计与假设检验	以案例形式介绍参数估计和假设检验这两种重要的统计推断形式。 1. 常见分布的参数估计 2. 正态总体参数检验 3. 计算描述性统计量 4. 统计图 5. 分布的拟合与检验 6. 核密度 (kernel density) 估计。

第三天 MATLAB 统计和最优化建模	
数据拟合——回归分析	本节讨论如何对多变量数据集进行线性和非线性拟合。散点图; 相关系数和协方差; 线性模型; 非线性模型
线性规划问题求解	本节讨论线性规划问题求解。模型建立; 模型求解
非线性规划问题求解	本节讨论非线性规划问题求解。模型建立; 模型求解
神经网络方法	神经网络原理; 基于神经网络方法的案例分析 (包含建模与求解)

郑智勇—MATLAB 高级（金融数量分析技术）

专题名称	授课内容
第 1 讲 随机模拟 —概率分布与随机数	1. 各类型随机数生成 2. 简单蒙特卡罗模拟 3. 案例：期权定价（欧式、美式、奇异期权）
第 2 讲 数据拟合与资产收益率分布检验	1. 随机拟合工具介绍 2. 案例：沪深 300 指数收益率分布检测(2007-2012 年)
第 3 讲 KMV 模型求解—方程与方程组的数值解	1. 方程组的求解 2. 案例： 根据公司股票市值与波动率计算公司价值与波动率并计算违约率
第 4 讲 策略模拟—组合保险策略分析	1. 样本路径选取：随机、历史抽样、复合方法 2. 案例：CPPI 策略与 TIPP 策略模拟与参数调优
第 5 讲 BS 公式与二叉树模型—期权定价与分析	1. 二叉树模型 2. BS 公式 Greeks 计算 3. 案例：期权价值计算\隐含收益率计算
第 6 讲 编程经验分享	1. 量化中的疑惑与理性 2. 编程逻辑与思考 3. 理想与现实的矛盾
第 7 讲 线形与非线性最优化算法	1. 线形最优化算法 2. 最小二乘（线形与非线性） 3. 非线性最优化 4. 案例：使用 N 个股票跟踪沪深 300 指数
第 8 讲 马柯维茨模型与投资组合绩效	1. 均值-方差模型 2. 有效前沿面计算（含约束条件，例如分行业） 3. 案例：CAPM 模型\基金与指数的投资组合绩效 Beta\alpha\sharp\最大回撤
第 9 讲 分形技术—移动平均 Hurst 指数计算	1. 分形理论 2. R/S 方法 3. 案例：上证指数与沪深 300 指数的 Hurst 指数计算
第 10 讲 固定收益证券的久期与凸度计算	1. 债券价格与收益率计算 2. 久期与凸度计算 3. 案例：债券组合久期免疫策略
第 11 讲 编程实用技巧	1. 定时触发程序运行 2. Matlab 发邮件 3. 坐标轴过原点实现 4. 案例：自动化办公—分级基金数据提取与分析

杨群—SAS 统计分析及应用

时间	主题	说明
第一天：数据步攻略		
1 小时	1. 数据读写 a) Data 步读入数据 b) PROC IMPORT 读入数据 c) Data 步输出数据 d) PROC EXPORT 输出数据 e) 案例：投票统计	帮助学员熟悉 SAS 软件基本用法，理解数据分析基本流程
3 小时	2. Data 步数据处理 a) 数据属性 b) 列操作 c) 行操作 d) 匹配连接 e) 循环/数组 f) 案例：分期付款计划计算表/手机客户通话统计	使用 DATA 步进行数据编辑汇总，根据需要定制各种统计指标
2 小时	3. Data 步函数应用 a) 数值型 b) 字符型	详解常用 SAS 函数，结合实际运用
第二天：SQL/巧用 FORMAT/常用过程步介绍		
2 小时	4. PROC SQL a) 分类汇总 b) 匹配连接 c) 案例：客服黑名单排查/营销数据汇总统计	使用 SQL 进行描述性统计分析，创建各种统计指标
1 小时	5. FORMAT 创建应用 a) PROC FORMAT 方法 b) Data 步方法 c) 案例：身份证号码解析统计/年龄分组频数统计	使用 format 进行数据转换、赋值、分组
3 小时	6. 常用过程步 a) PROC MEANS/SUMMARY：案例 iris 数据各种统计 b) PROC UNIVARIATE：案例：连续变量分位点定位 c) PROC FREQ：案例：卡方检验解析 d) PROC SORT：案例：长名单去重 e) PROC SURVEYSELECT：案例：特征样本分层抽样 f) PROC APPEND g) PROC PRINTTO h) PROC TRANSPOSE	理解各种常用过程步的功能和应用，根据参数定制输出所需的分析结果
第三天：宏程序 线性回归		
3 小时	7. 宏语句 a) 宏变量 b) 宏程序 c) 案例：销售业绩月度报表制作	使用宏程序循环处理多个数据集，提高工作效率
3 小时	8. 线性回归解析 PROC REG 案例：CEO 薪水解释模型	用 excel 计算表详细解析 reg 算法逻辑，结合应用
第四天：逻辑回归 主成分分析		
3 小时	9. 逻辑回归解析 PROC LOGISTIC 案例：上市公司 ST 处理预测模型	用 excel 计算表详细解析 logistic 算法逻辑，结合应用
3 小时	10. 主成分分析解析 PROC PRINCOMP 案例：图像压缩（适合工科科研）	使用工科案例介绍多元分析方法，拓展视野