

数据科学： 搭建原理与应用的桥梁



提升数据素养，增强决策信心

Berkeley ExecEd | **Berkeley Haas**

导言



对颠覆的恐惧是企业界比较常见的一种心态：害怕更具智能化、更具灵敏度的企业利用数据抢占先机，获得竞争优势。但在许多高瞻远瞩领导者以及加州大学伯克利分校高管教育看来，战胜恐惧的最佳方式就是重新思考企业的定位，将企业定位为数据型企业——因为数据素养技能是能真正提升企业业绩的重要保障。

目前，对大数据和相关技术的投资空前高涨，然而New Vantage Partners的调查研究显示，仅48%的企业认为他们在数据和分析方面有竞争优势；仅31%的企业是数字驱动型企业；而建成数字文化的企业仅有28%。实际上，阻碍实现数据的价值和巨大可能性的原因，不在于技术本身，而在于人才转型、文化转型、流程转型。

本课程在严格基于学科内容的基础上突出实践性，学员通过亲身经历在实践过程中学习。加州大学伯克利分校高管教育《数据科学：搭建原理与应用的桥梁》不仅能帮助各行各业在职专业人士深度学习和理解数据科学和数据分析，还注重实践，学员将在亲身经历实践过程中学习体验各种方法，以达到灵活运用数据分析的目标。完成学业后，学员将能够与数据科学团队、数据分析团队高效合作，帮助和指导数据科学家、数据分析家为企业更好地服务，实现更为可观的商业价值。

66

“本《数据科学》课程能够帮助您真正提升数据素养。我们的重点不是数学，而是准备比《谷歌分析》报告更深入地挖掘数据。”

- Steve Tadelis (史蒂夫·塔德利斯)， 哈斯商学院 James J. and Marianne B. Lowrey 商学讲座教授

您适合本课程吗？

本课程在严格基于学科内容的基础上突出实践性，学员在亲身经历实践过程中学习现代数据分析方法，适合个人贡献者和国企、民企的中高层管理人员。

典型受益行业和群体：

- 肩负或即将肩负管理数据科学家团队或供货商重任的经理
- 绩效营销专家
- 产品工程师、产品经理、研发经理
- 商业/技术策略师和顾问
- 人力资源专家
- 以数据分析为抓手的技术驱动型企业，包括：零售业、信息技术业、电子商务业、金融服务业、金融科技业、制造业、医疗保健业等。

数据科学素养需要的知识储备

不要求学员在上课前具备编码等专业知识，但学员应当有量化的概念，该概念非常重要。第一周为热身周，将带领学员回顾数学与统计学的基础概念，包括平均值、标准差、图、直方图、线性函数、对数函数等。此外，每周设有“预备课”，专门介绍下一模块中涉及的重点概念，供学员参考。但真正养成数据科学素养，离不开学员在亲身实践体验中学习数据、数学、统计学。课程始终，无论是理论还是实践，我们会全力支持学员。

课程重点：

养成数据驱动理念

- 学会提出正确的问题
- 学习最常用的数据技能，从数据中获得洞见
- 为具体的商业问题找到正确的解决方法

学习沟通和解释数据

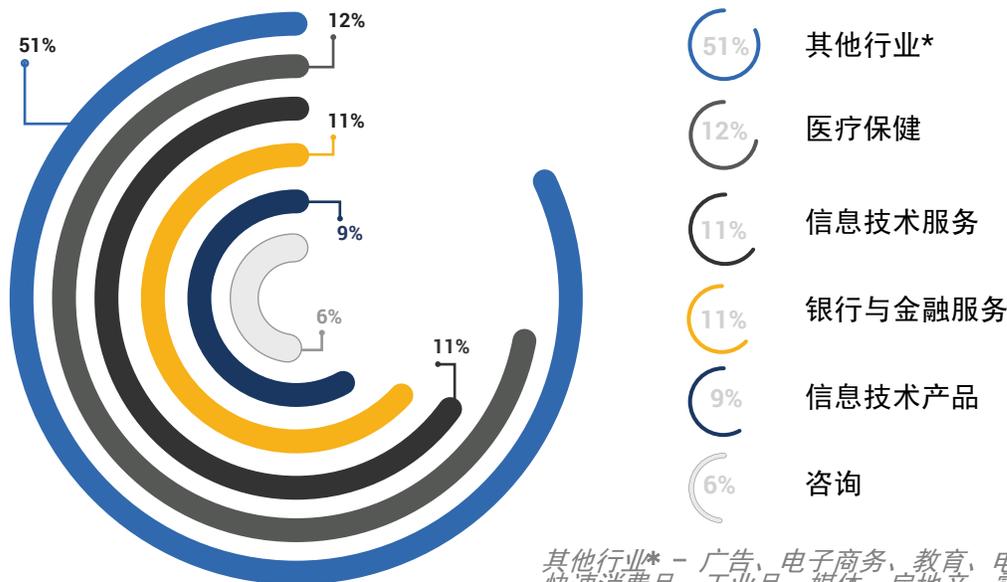
- 卓有成效地呈现数据
- 学习与数据科学家沟通的语言
- 通过学习最常用的数据技能，更有效地解释数据

建设数字文化

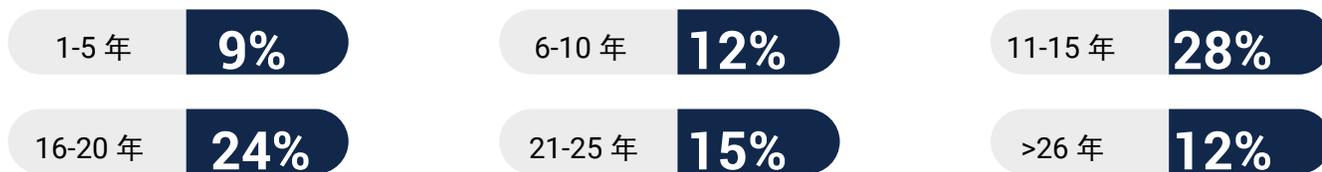
- 在数据影响战略、决策、执行的企业，通过技术和流程驱动文化转型
- 了解建设优秀数据科学家团队的必备能力

学员一览

《数据科学》课程的学习历程中，学员有益友相伴。以往的学员来自各行各业，职位各异，管理阶层不同。



工作经验



学员来自以下企业

部分学员代表来自：

- Adobe
- 拜耳
- 欧睿国际
- 宝洁公司
- 亚马逊
- 思科系统
- 埃克森美孚
- 盛世长城
- 苹果
- 星展银行
- 微软
- Visa
- 美国银行
- 道琼斯
- 甲骨文公司
- 富国银行

学员在企业中的角色

部分学员代表在企业中的角色：

- 首席执行官和董事总经理
- 预测分析总监
- 战略分析高级经理
- 首席信息官
- 首席数据官
- 业务发展高级经理
- 全球数据仓库和企业应用总监
- 数字化战略副总裁
- 数据分析师
- 企业架构总监
- 机器学习和人工智能主管
- 数字策略营销顾问

学员分享



“本课程让我受益最深的是：使用Jupyter Notebook 动手实践任务以及在线实战模拟。”

- 人才管理分析师**Bita Luliano**



“我喜欢课程集视频课、动手实践任务和Jupyter Notebook应用练习于一体。”

- **Digital Platforms & Products**执行总监**Michael Wolff**



“学习本课程后，我对数据科学有了深入了解，深度掌握了各种语言、模型、算法和价值，并了解它们各自在实际应用中具备的优点和存在的问题。”

- 应用科学家**Saloni Sonawala**



“我接触到好用的在线平台Jupyter Notebook，学习到现实世界的案例，课程还提供学习支持，这些都为精彩的学习之旅画上浓重的一笔。”

- 业务发展经理**Dmitry Karablinov**



学习历程

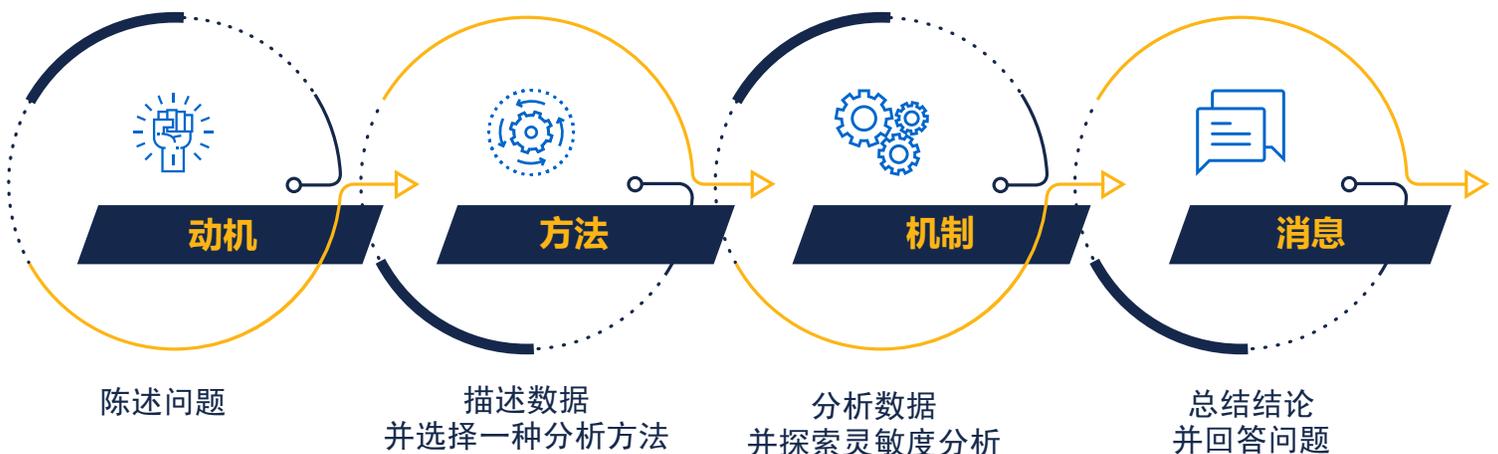
本在线课程为期10周，学员将与加州大学伯克利分校高管教育的教授、行业领袖以及世界各地的同行建立直接联系。课程在严格基于学科内容的基础上突出实践性，学员将在亲身经历实践过程中学习使用开源交互式平台Jupyter Notebook计算、分析数组。课程内容虽为事先预定，我们仍会致力于提供生动的体验式学习，学员将在解决一大批现实世界问题的同时，发现新的可能性。

- ✓ 从对谷歌、the Oakland A's、优步等数据驱动型企业大咖的访谈中学习
- ✓ 每周设实时在线“预备课”，专门介绍下一模块涉及的技术概念；以及在线实战模拟，在线作业指导
- ✓ 与加州大学伯克利分校高管教育的教授实时在线研讨，研讨会设答疑环节
- ✓ 设有两次为期一周的实验室学习，学员亲身实践，神挖数据
- ✓ 应用练习：在Jupyter Notebook平台上使用Python完成（完整的和不完整的）数据分析、可视化展示。
- ✓ 主持讨论板

数据科学与数据分析的步骤

数据科学和数学分析的四大步骤：

1. 动机：清晰地陈述问题。
2. 方法：描述使用的数据以及需要使用的分析方法。
3. 机制：与其他团队成员一起动手实践、真抓实干、干出成效。
4. 消息：总结结论，回答最初提出的问题时言简意赅、清晰明了，便于企业领导一听就明白。



行业实例和企业实例

企业实例：

由于加州大学伯克利分校高管教育的教授与硅谷及其周边地区众多具有行业代表性的顶级科技公司联系密切，这些公司的科研及应用直接或间接地成为了本课程内容的来源：



亚马逊



优步



易趣



盖洛普



StubHub

应用练习：



火箭燃料

作为一家数字广告公司，火箭燃料使用大数据和人工智能来优化数字营销。使用Python运营广告宣传数据，使数据可视化，实施A/B测试，评判广告的成功。



首都共享单车

作为一家共享单车公司，首都共享单车致力于预测共享单车的单日使用总量。使用Python生成并分析摘要统计信息，创建回归模型，对共享单车的单日使用总量进行预测。



IBM员工流失

使用IBM创建的虚拟数组对员工的离职率做预测。使用交互式平台Jupyter Notebook中的Scikit-Learn软件（一款免费的开源的Python机器学习软件），创建、训练并评估其它机器学习模型。

行业实例：

体现在通过提升数据素养技能而受益的数据敏感性企业中。典型行业包括：



金融科技/金融



医疗保健



信息技术制造业



零售业



服务业

注：本课程提到的企业名称或产品名称均为商标™或注册®商标名称。使用商标名称不代表和商标所有人之间存在任何附属关系或赞助关系。



课程模块

模块 1

概率决策

介绍数据科学和数据分析的基础概念，然后探索数据的基本知识

- 区分分类数据和数值型数据
- 探索数据揭示信息的基本方式
- 类别变量和数值变量的相关性分析
- 医疗保健业实例：大数据管理式医疗(HMO) 模式
- Jupyter Notebook, Python和Panda简介

模块 3

假设检验

基于数据制定企业决策还依赖于表达清晰的假设，只有这样的假设才有助于统计检验。本模块会介绍假设检验的基础知识，包括统计比较、置信区间和误差范围。

- 实验设计的基本原则
- 识别并区分数据的单、双侧检验
- 使用4M模型解决问题 (动机、方法、机制、消息)
- 实例：24 Hour Fitness测试新专用饮食方案—测试对照组和治疗组

模块 2

创建样本数据

了解数据分析报告常用的专业术语，学习样本分析的方法；借助样本分析，检验一类错误和二类错误和极限控制来分析企业决策的利弊。

- 了解样本类型、样本差异、样本质量
- 了解样本数据的基本概念及分类
- 识别并缩小样本数据中的偏差
- 实例展示：联合概率、边际概率、条件概率：Comcast、谷歌、Nextag

模块 4

利用样本数据推断

学习统计学中最常见的直线模型和曲线模型。学习拟合线性模型的操作方法及适用范围：市场需求、定价、弹性。

- 了解使用直线模型和曲线模型的条件，学习解读直线模型和曲线模型
- 研究曲线（非线性）模型在车辆重量和燃油效率方面的应用
- 使用4M模型解决问题：美国费城的信用卡问题、犯罪问题、房价问题

模块 5

简单回归模型

简单回归模型是数据驱动型企业制定决策的核心模型。本模块的重点是了解简单回归模型的使用方法，模型有效的假设条件，以及如何使用模型帮助企业做出更佳决策。

- 了解简单回归模型的定义和应用；明确简单回归模型的适用范围
- 应用预测间隔并解释
- 确定影响回归模型的三大问题：数据变化、离群值和观测值之间的依存关系
- 零售业实例：使用回归模型解决连锁店的选址问题

模块 6

多元回归模型

在基本概念的基础上进一步深入，定义多元回归模型，探索该模型的适用范围。

- 区分边坡斜率和偏斜率
- 利用多元回归模型分析：推断
- 总结多元回归模型的建模步骤
- 金融业实例：建立多元回归模型，解释索尼公司的股票投资回报
- 人力资源行业实例：使用修正回归模型分析工资数据，判断是否有同工不同酬的现象

模块 7

预测机器学习

学习基本概念，研究不同的应用，解密机器学习。

- 辨识机器学习的三种学习方法：有监督学习、半监督学习、无监督学习
- 研究机器学习的方法，包括有监督学习使用的“词袋模型”法。
- 使用时间序列回归预测
- 网络安全实例：机器学习检测垃圾邮件

模块8

建立高效的数据科学家团队

掌握基本概念和常用工具后，进一步深挖高效数据科学家团队的必备能力，学习建设企业数字文化。

本模块会指出需要避免的常见误区，并提供策略，帮助企业与数据科学家高效沟通。

- 学习建立高效的数据科学家团队的必要条件
- 继续探索数字文化建设
- 广告业实例：火箭燃料：转化率、优势、投资回报率、机会成本、A/B测试

注：为了帮助学员通过亲身体会实践学习数据技术，学会用数据制定更好的决策，课程设置了2次各为期1周的实验室学习课程，学员将有机会深挖数据。该实践活动是10周课程总体设计的一部分。

授课导师

STEVE TADELIS

史蒂夫·塔德利斯教授

哈斯商学院James J. and Marianne B. Lowrey
商学讲座教授



史蒂夫·塔德利斯教授在斯坦福大学任助理教授八年，后来到加州大学伯克利分校哈斯商学院任教。

曾在eBay研发实验室任总监、杰出经济学家（2011 - 13）。在亚马逊担任经济学与市场设计副总裁期间（2016 - 17），他与技术人员、机器学习科学家、业务领导携手合作，实现了将经济研究工具应用于多款产品和业务。至今仍是亚马逊经济学家团队一员，是亚马逊的兼职顾问。

史蒂夫·塔德利斯教授目前的研究领域：电子商务、组织经济学、采购合同、企业理论和产业组织理论、合同理论、博弈论。

部分荣誉和成就：

- 荣誉称号——全日制工商管理硕士课程Cheit教学奖，2010-11
- 豪蒂亚斯奖——《比较经济学》期刊最佳文章，2010-11
- 加州大学伯克利分校哈斯商学院Barbara and Gerson Bakar教职研究员，2008-15
- 斯坦福大学Phi Beta Kappa 本科教学奖，2005
- 斯坦福大学经济系优秀导师奖，2002

SHACHAR KARIV

沙查· 卡里夫教授

哈斯商学院 Benjamin N. Ward 经济学教授



沙查·卡里夫教授是实验社会科学实验室 (Xlab) 的前系主任兼教务主任。其对行为经济学和实验经济学的研究，为理解个人偏好、风险态度、时间态度等问题提供了新工具，为决策提供几近全面的信息。

学术方面，教授曾在斯坦福大学、普林斯顿大学、牛津大学、剑桥大学、欧洲大学学院、挪威经济学院等

学府担任客座教授。教授还是Capital Preferences公司的联合创始人、首席科学家，这家公司变革了企业投资建议、贷款、保险、人才市场解决方案。

部分荣誉和成就：

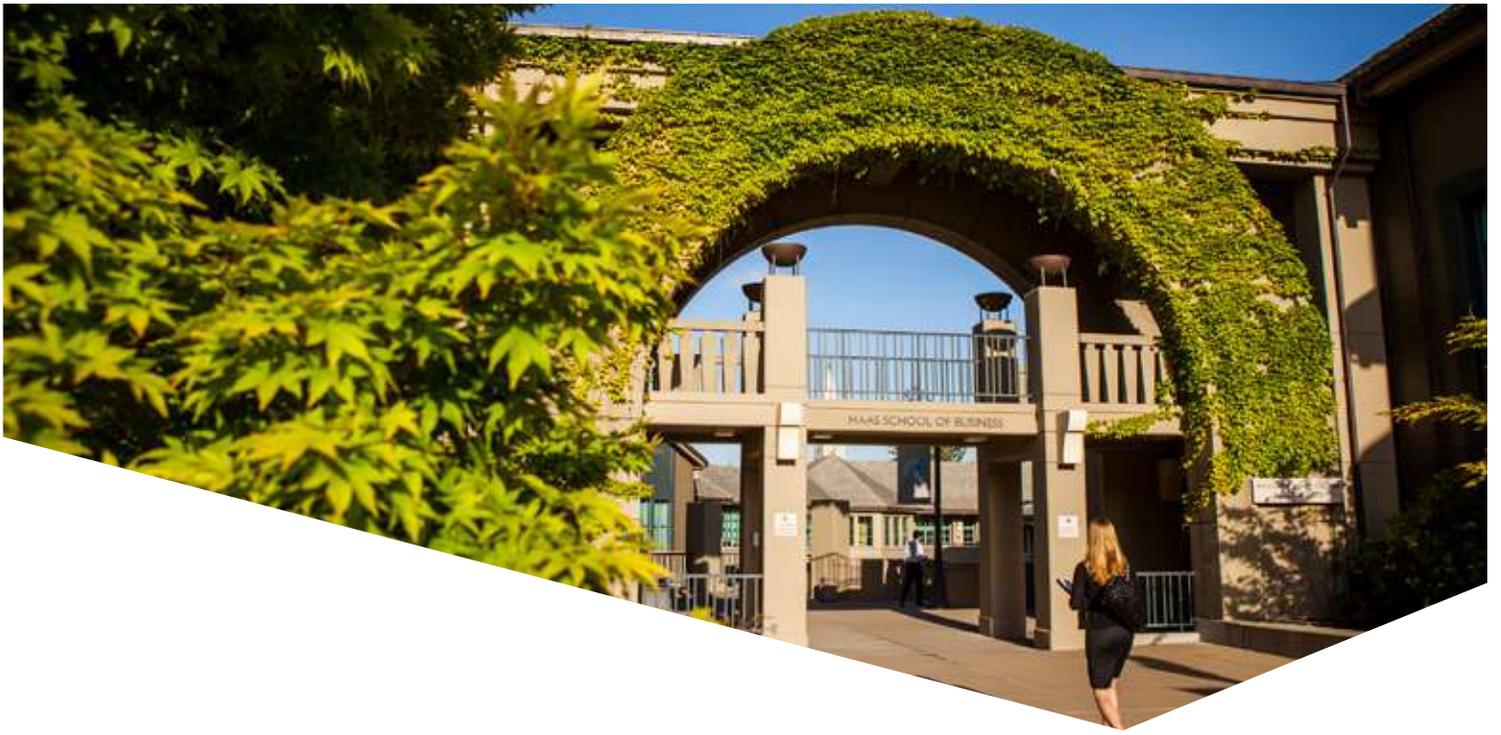
- 加州大学伯克利分校哈斯商学院 Earl F. Cheit 教学杰出奖，2012
- 获美国斯隆基金会研究资助，2009
- 加州大学伯克利分校社会科学系杰出教学贡献奖，2006-07
- 加州大学伯克利分校经济系研究生会优秀导师奖，2006-07

认证证书

成功完成本课程的学习后，学员将被授予由加州大学伯克利分校高管教育认证的电子版证书。本课程成绩分合格和不合格两等，80分合格，合格方可获得认证证书。



注：成功学完本课程后，我们将按照您报名时登记的姓名签发电子版认证证书，并使用电子邮件发送到您的邮箱。证书中使用的图像仅用于说明目的，加州大学伯克利分校高管教育可酌情更改。



关于优仕研习社

加州大学伯克利分校高管教育与知名在线教育优仕研习社合作，提供一系列高影响力的在线课程。课程凭借加州大学伯克利分校高管教育多年教研与实践中的经验，传达技术实践形成的思想领导力。

通过与优仕研习社的合作，我们能够以相互协作和富有吸引力的方式，将校内提供的课程扩展到更大的范围，并保持加州大学伯克利分校的质量。在课程设计方面，优仕研习社基于班级学习社区的理念设计课程，最大化调动学员之间互动分享，同时邀请世界一流知名教授实时授课，结合基于项目的学习增强实践性。去年，30,000多名来自120多个国家/地区的学生因优仕研习社而受益。

***学习模块、完成作业和参加实时学习支持会议所需的时间长度取决于学员的技术能力和背景。此为预估顺利学完课程平均需要花费的时间。**

学习体验



课程在设计时，充分考虑学员个性化的学习需求，同时也考虑学员之间相互沟通、相互交流、相互学习的积极作用。为此，我们采用用户友好型学习平台，帮助学员清晰准确地把握课程结构设置，确保学员完成学业。

突出实践性

采用将理论融入生活的教学方法，包括：

- 系统知识碎片分解
- 真实世界中的应用
- 学员之间沟通讨论
- 实时在线互动教学



提高便利性

轻松获得课程内容，支持多种登录设备，便于在职专业人士灵活安排时间远程学习——无论何时，无论何地。报名学员收到的学习资料以模块为单位，学习内容每周更新，课程模块涵盖多种教学手段，包括：

- 视频讲座
- 讨论
- 课堂材料: 文章、案例
- 测验
- 调查
- 作业



采用实时在线教学研讨会的方式，进行实时在线互动答疑，充分满足学员对模块教学的个性化学习需求。未能参加实时在线教学研讨会的学员将收到研讨会录播视频，因此，课程进度不会落下。学习平台设计行业领先，功能多元，学员能够在线创建档案，与其他学员交流沟通，协作互动，与包括课程负责人和教学助理在内的学术/业界专家畅所欲言、直抒胸臆。课程注重实践，作业设置常常以学员现实的问题为切入点，增强理论的实践效果。

增强趣味性

国际化的班级文化有利于学员之间无缝互动沟通，合作完成任务，携手共进，直到完成学业——全程充实、文化体验丰富。

课程要求

学员学习课程需要具备以下条件：

- 有效的电子邮箱
- 计算机联网设备：台式计算机/笔记本电脑、平板电脑、智能手机
- 课堂材料：文章、案例
- 登录学习平台任务需要使用最新版本的浏览器
- 需要使用微软办公软件以及PDF浏览器浏览word文档、数据表、PPT展示、PDF文件、字幕。

其他要求

课程可能需要使用多种软件、工具和应用。如须使用这些，学员会在注册阶段或者课程开始时提前得到通知。如需咨询相关事宜，我们的课程顾问非常乐意回答这些问题。



联系课程顾问

Email:

电 话:

* “EDS Executive Education”、“ERS Institute of Management”和“优仕研习社”均为EDS Group 旗下品牌。
“优仕研习社”的主体为上海择藤教育科技有限公司，在法律允许的范围内择藤享有以上资料的解释权。

Berkeley ExecEd | **Berkeley Haas**