双教授实地科研项目-教授大纲

大数据时代下的商业分析：利用数据可视化与推断统计制定决策

Data Visualization for Business

R语言，管理学，统计分析，数据科学，商业决策，数据隐私

南加利福尼亚大学 (美国) x 北京航空航天大学

Session Plan

Schedule	Topics
Professor Warm Up	Course Introduction and Greetings
TA Preview	Course Preview and Introduction
Session 1	**Objective: Introduction to Data visualization and descriptive statistics** Description: Basic principles of effective visualization & Descriptive statistics **学习目标：数据可视化导论&描述性统计** 学习描述：有效数据可视化的基本原则，描述性统计的概论及基本原则
中方教授课1	主题：统计学基础 描述：回归分析基础知识与实战
Session 2	**Objective: Introduction to R** Description: What is R and How to use R **学习目标：R语言介绍** 学习描述：统计软件R语言的介绍与使用说明
中方教授课2	主题：联合分析大数据方法（理论基础） 描述：对联合分析大数据方法的基本概念和理论进行系统性学习，为后面的数据分析打下基础。
Session 3	**Objective: Data visualization, data types and formats, and data cleaning** Description: Using software for visualization **学习目标：数据可视化，数据类型与格式，以及数据清除** 学习描述：学习如何应用软件实现可视化
中方教授课3	主题：切分分析大数据方法（理论基础） 描述：对切分分析大数据方法的基本概念和理论进行系统性学习，为后面的数据分析打下基础。
Session 4	**Objective: Descriptive data mining and privacy** **学习目标：描述性数据挖掘与隐私**
中方教授课4	主题：大数据科学研究1 描述：对联合分析大数据进行1个科学研究项目
Session 5	**Objective: Principles of good visualization design & Inferential statistics I** Description: Using software for visualization, designing visuals, and useful statistical tests **学习目标：良好可视化的设计原则，以及推断统计(一)** 学习描述：使用软件进行可视化，设计视觉效果和有效的统计检验
中方教授课5	主题：大数据科学研究2 描述：对切分分析大数据进行1个科学研究项目
	Objective: Predictive data mining

Session 6	**学习目标：预测性数据挖掘**
Session 7	**Objective: Exploratory data analysis, and inferential statistics II** Description: Examining data and more statistical tests **学习目标：探索性数据分析 & 推断统计(二)** 学习描述：检验数据和统计测试相关知识
Session 8	**Objective: Imputation and data ethics** **学习目标：归因与数据伦理**
Session 9	**Objective: Storytelling and inferential statistics III** Description: How does data tell a story & More statistical tests **学习目标：数据展示说明 & 推断统计(三)** 学习描述：如何用数据进行说明 & 基于推断统计相关知识进行多次统计测试
Final	**Presentations and wrap-up**

Reading Materials

	Session 2: Intro to R Agrawala, M., Li, W., & Berthouzoz, F. (2011). Design principles for visual communication. Communications of the ACM, 54(4), 60-69. doi :10.1145/1924421.1924439 Byrne, J. (2015, July 12). Consulting: Why so many MBAs do it. Poets & Quants. . https://poetsandquants.com/2015/07/12/consulting-why-so-many-m Heer, J., Bostock, M., & Ogievetsky, V. (2010). A tour through the visualization zoo. Communications of the ACM, 53(6), 59-67. https://doi :10.1145/1743546.1743567 **Session 3: Data visualization, data types and formats, and data cleaning** Ali, S. M., Gupta, N., Nayak, G. K., Lenka, R. K. (2016). Big data visualization: Tools and challenges, Computer Science, 2nd International Conference on Contemporary Computing and Informatics (IC3I) Heer, J. (2010). A conversation with Jeff Heer, Martin Wattenberg, and Fernanda Viégas. Queue, 8(3), 10 link: https://queue.acm.org/detail.cfm?id=1744741 **Session 4: Descriptive data mining and privacy** Elegido, J. (2011). The ethics of price discrimination. Business Ethics Quarterly, 21(4), 633-660. Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. Proceedings of the National Academy of Sciences, 110(15), 5802–5805. doi:10.1073/pnas.1218772110.

| 外方教授课题阅读材料 | Narayanan, A., & and Shmatikov, V. (2010). Privacy and security: Myths and fallacies of "Personally Identifiable Information". Communications of the ACM, 53(6), 24-26.
Zang, J., Dummit, K., Graves, J., Lisker, P., & Sweeney, L. (2015). Who knows what about me? A survey of behind the scenes personal data sharing to third parties by mobile apps. Journal of Technology Science…
http://techscience.org/a/2015103001
Session 5: Data visualization, principles of good visualization design, and inferential statistics I
Parlapiano, A. (2016, November 1). There are many ways to map election results. We've tried most of them. The New York Times.
https://www.nytimes.com/interactive/2016/11/01/upshot/many-ways
Syed Fiaz, A. S., Asha, N. Sumathi , D., & Syed Navaz, A. S. (2016). Data visualization: Enhancing big data more adaptable and valuable. International Journal of Applied Engineering Research, 11(4), 2801-2804
Talasek, J. D. (2015). Science and Culture: Data visualization nurtures an artistic movement. Science and Culture, 112(8), 2295
Session 6: Predictive data mining
Visual vocabulary (n.d.)
http://ft-interactive.github.io/visual-vocabulary/
Session 7: Data visualization, exploratory data analysis, and inferential statistics II
Mackinlay, J., Kosara, R., & Wallace, M. (n.d.) Data Storytelling: Using visualization to share the human impact of numbers. Tableau.
https://www.tableau.com/data-storytelling
Murrary, E. (2017, February 26). Tableau, travel and triathlons. Tri My Data.
https://trimydata.com/2017/02/26/your-tableau-dreamjob/
Session 8: Imputation and data ethics
ACM (2018, June 22). Code of ethics. Association for Computing Machinery. DOI 10.1145/3274591
http://www.acm.org/binaries/content/assets/membership/images2/
Barocas, S., & Boyd, D. (2017) Computing ethics: Engaging the ethics of data science in practice. Communications of the ACM, 60(11), 23-25.
https://doi:10.1145/3144172
Bishop, L. (2017, February). Big data and data sharing: Ethical issues. UK Data Service, UK Data Archive.
https://ukdataservice.ac.uk/media/604711/big-data-and-data-sharing
Lerman, J. (2013, September) Big data and its exclusions. Stanford Law Review Online, 66, 55-57.
Session 9: Storytelling and inferential statistics III
Kosara, R. (2014, July 23). Putting data into context. |

	Eagereyes. https://eagereyes.org/basics/putting-data-into-context Oetting, J. (2018, November 23). Data visualization 101: How to choose the right chart or graph for your data. Hubspot. https://blog.hubspot.com/marketing/types-of-graphs-for-data-visual Schwabish, J. A. (2014). An economist's guide to visualizing data. Journal of Economic Perspectives, 28(1), 209–234.
中方教授课题阅读材料	1. **数据科学与商业分析 用机器学习与统计学优化商业决策（图灵出品）** 2. **斯坦福数据挖掘教程（图灵出品）**