**296A-02. Data Science and Analytics
Thought Leaders in Big Data Analytics**

**Course Description:**

The seminar explores
- Leading-edge trends in Data Science and Analytics at Silicon Valley: Tech and VC firms, and Startups. The speakers will include executives, entrepreneurs, VCs, and researchers from leading firms.

The topics covered will include (a subset of):
- “Big Data”: Landscape
- Big Data Architecture, including Streaming, Real-time and CEP
- “Big Data Analytics” Landscape
- Data as a Product and service, and Data Products
- Distributed and Scalable Machine Learning and Statistics
- Internet Analytics including
 Online Advertising and Marketing: Targeting, Attribution, Exchanges
 Display Advertising Analytic
 Search analytics and Information extraction
 Social Media and Networks Analytics
 Mobile, Web, App, and IT Analytics
 Measurement and Audit Analytics
-Business Analytics including
 Financial Services and Risk Analytics
 Operational and Service Analytics
 Healthcare Analytics
 Energy Analytics

The seminar will cover the following aspects:
- The types of problems being addressed in data science and analytics, the component methods and technologies being developed, and fruitful areas for research and entrepreneurial efforts
 - This requires attendance and participation in the seminar series, which is open to the broader student and faculty community

URL: <http://courses.ischool.berkeley.edu/i296a-dsa/s12/>
URL: <http://courses.ischool.berkeley.edu/i296-dsa/s13/>

Contact Prof. Ram Akella, akella@ischool.berkeley.edu, Cell 650-279-3078, 422, Sutardja Dai Hall

Prerequisites:

None

Units: 2-3

**Course History**Spring 2013 | Instructor(s): [Ramakrishna Akella](http://www.ischool.berkeley.edu/people/faculty/ramakrishnaakella). Ray Larson, Lee Fleming
Industry Experts: Jake Flomenberg, Jimi Shanahan

**Office Hours:**
By appointment in the 4-6 pm window (or variants) on Wednesday (or by phone).
Location: South Hall 205; on January

**Contact Information:**

422 SDH (Sutardja Dai Hall)
akella@ischool.berkeley.edu
Cell: 650-279-3078
Skype: ramakella1

**Course Objectives**

1. Learn about and understand landscape of Data Science and Analytics including a subset of
* Implications of Big Data for Analytics
* The impact of Data Analytics on Business Analytics
* Data Analytics **including Machine Learning, Data Mining, Statistics, Text/Image/Video Mining and Analytics, Search - Information Retrieval, Social Networks, Web Analytics, Online Advertising and Computational Marketing**
1. Develop
* A perspective on the business needs, state of the art research and technology, gaps, and emerging and novel mathematical and other techniques and approaches to address these
* A demonstrated capability to identify enterprise or consumer needs, and develop solutions to meet there

**Expected Diversity of Participants**

Given the impact of Big Data Analytics, and the range of areas it spans, it is anticipated that participants will include 3 groups, with overlap:

* Doctoral and Masters students with strong mathematical/statistical training, such as those in Machine Learning, Control, and Signal Processing (EECS), IEOR, Finance, Marketing, Operations, etc.
* Masters or doctoral students who are familiar with software, e.g. CS
* Masters of doctoral students who have a more general background,

e.g. ISchool, MBA

**Projects and Evaluation**

* Teams of 3 (or smaller if desired)
* Research Paper or Project on Big Data Analytics
* These can be fairly diverse, depending on the team
* Could be
	+ Hands on, working with data identified by student or faculty,

and/or working with firm or VCs and faculty

* + More theoretical, large-scale data analysis (whether mathematical, technical, political, legal, social, etc.)
	+ Or a combination
* Reports should address a topic of interest, whether it be an exploration of a particular gap within the Big Data landscape, a business proposal, or something else, such as a research paper for a conference or journal. The goal is to delve deeply into a particular area as opposed to surveying a broad landscape. You should be sure to spend time understanding and articulating the problem/pain that you are addressing to ensure that you are addressing an issue that truly exists.
* 2 Units: participation. 3 Units: As below
* 10% class participation – Each team writes 4 reports, each summarizing a presentation; monthly progress report (30%): 1. Define problem and background, 1/28 2. Analyze problem, 3/31 3. Analyze problem in depth, 4/30, and final report – 1/8 & 1/15 (50%), and PPT presentation (10%). Crisp, insightful, analytic reports will receive higher points.

**More detail on Project**

* A report summarizing perspective, including:

Research approaches, and/or product/service or entrepreneurial opportunities in some detail

* Detail for research would include problem description, model formulation, and developed solution approaches, with validation on real data, where feasible
* Detail for product/service or entrepreneurial business opportunities would include the details of a business case, business model, competitive analysis, positioning, market analysis (and study where feasible etc.)
* Monthly 5 page updates, and a final 20 (15-25) page report and 15 PPT VC presentation (with 15 page Appendix); Team size of 3.
* Weekly rotating team coverage of talks and detailed write up

**Big Data Analytics Background Resources**

[**http://datascienc.es/**](http://datascienc.es/)

<http://courses.ischool.berkeley.edu/i290-dma/s12/>

https://blogs.ischool.berkeley.edu/i290-abdt-s12/author/hearst/

[**http://www.cs.berkeley.edu/~jordan/courses/294-fall09/**](http://www.cs.berkeley.edu/~jordan/courses/294-fall09/)

**http://alex.smola.org/teaching/berkeley2012/**