Data Mining and Visualization
PA 756
Spring 2017

Stephen Harris, Lecturer

email address:
sharrissfo@sfsu.edu

Class meetings: Saturday 9:00 am-5:00pm
February 4 and February 18, 2017

Phone: 415-734-8461

Office Hours: by appointment 5:00 - 6:00pm

Course Description

This course examines data mining and visualization concepts, tools, and techniques with analysis of practical application by government (Public Sector) and industry (Private Sector). The course focuses on how to concepts to mine and transform data into visual representations so that key decision makers and data consumers can effectively prepare and present data for analytical insights. These concepts are also referred to as components of "Big Data & Analytics". The topics covered in this course include 1) Establishing a common business and technology Lexicon/Taxonomy, 2) An introduction to visual representations and interaction techniques, 3) Introduction and overview of Data mining, visualization and transformation tools, 4) Introduction and overview of key techniques to support the preparation, presentation and dissemination of analytic results. 5) Industry best practices for leveraging data mining and visualization (Big Data & Analytics) tools and techniques. This course will blend various theoretical and applied concepts

Learning Goals & Objectives

At the completion of the course, students will be able to:

• Knowledgeable of the various data mining and visualization tools available in the market
• Converse with Private & Public sector professionals leveraging leveraging the common taxonomy/lexicon that bridges business, government and technology
• Apply the major concepts in data mining & visualization design
   • Identify the market tools & processes available to transform large datasets for analysis
• Critique the private & public sectors application of best practices
• Design a data mining and visualization case study outline that applies the concepts learned in this course

Course Methodology and Grading Requirements

Key Readings	5%
Class Participation	20%
Written Assignments	75%

Refer to class schedule & grading chart for additional details

Learning Materials: (Key Readings, Case Studies & White Papers)

The table below outlines Key Pre-Read Materials in preparation for the course. Additionally each student should download, review and be prepared to discuss the referenced case studies. For those case studies where links are not included students should locate them on the web by searching for the exact title provided. The purchase of software is not required for this course*

Required Pre- Readings	Case Studies	Technology White Papers
What is Data Mining http://www.anderson.ucla.edu/facultv/ iason.frand/tcachcr/tcchnologics/ palacc/datam mine, h tm	Detection of Anomalous Particles from Deepwater Horizon Oil Spill Using SIPPER3 Underwater Imaging Platform*	Magic Quadrant for Business Intelligence and Analytics https: //www. ₂artner.com/doc/ reprints? id=l -2XXF.T8P&ct= 160204&st=sb
8 Big Trends in Big Data & Analytics - http://www.computerworld.com/article/ 2690856/big-data/8-big-trends-in-big- data-analvtics.html	Click Fraud Detection: Adversarial Pattem Recognition over 5 Years at Microsoft*	Magic Quadrant for Advanced Analytics httDs://vww. gartncr.com/doc/ reprints⁹. ct= 160210&id= 1 -2 YEIILW&st= șb
Big Data Analytics - Eating an Elephant http://www.forbes.com/sites/kevinomarah/ 2016/12/22/big-data-analvtics-eating-an- elephant/#23 76aa9feb8d	Mining Medical Images*	
Harvard Business Review - Analytics 3.0 Thomas Davenport httns://hbr.org/ 2013/12/analvtics-30	Detecting and Tracking Political Abuse in Social Media http: //www. aaai..ora/ocs/ index.php/IGWSM/ 10WSM11 /paper/viewFile/ 2850/3274/	
Big Data Wikipedia -https:// en.wikipedia.org/wiki/Big data	MobileMiner: A Real World Case Study of Data Mining in Mobile Communication*	
Top 10 Big Data Trends for 2017 - Tableau*	A Review: Application of Data Mining Tools in CRM for Selected Banks.*	
Predictive Analytics in the Public Sector: Using Data Mining to Assist Better Target Selection for Audit https ://www. google .com/url? sa=t&rct=j &q=&esrc=s&source=web &cd=&ved=OahUKEwiHocKW INvR AhVUcAKHbnGBCoOFgguMAI&url =http%3 A%2F%2Fwww.ejeg.com %2Fissue%2Fdownload.html %3FidArticle %3D232&usg=AF0jCNE1015Ypvtgj 2kMfhS2OMoyRzqQ	Data Mining at FDA - http://www.fda.gov/ ScienceResearch/ DataMiningatFDA/ ucm446239.htm	

**__Additional Resources:__ **

It is recommended that each student consider a student membership to become part of the The Data Warehouse Institute (TDWI), more information can be found at TDWI.org - . TDWI provides individuals and teams with a comprehensive portfolio of business and technical education and research to acquire the knowledge and skills they need, when and where they need them. The in-depth, best-practices-based information TDWI offers can be quickly applied to develop world-class talent across your organization's business and Information Technology functions to enhance analytical, data-driven decision making and performance.

TDWI advances the art and science of realizing business value from data by providing an objective forum where industry experts, solution providers, and practitioners can explore and enhance data competencies, practices, and technologies. ***Membership to TDWI is <u>not</u> a required to successfully complete this course. ****

<u>Class Schedule - Assignment Due Dates & Grading</u>

Pre Course Read	Key Readings listed above should be read prior to Feb 4th	Key Pre-Readings	Prior to 2/4 - 5%
Feb 4	Introduction Common Taxonomy of terms Big Data & Analytics concepts Tools & Techniques Best Practices	Individual Exercises & Group Discussion	Due Same Day 10%
	Write a 1-2 pager summarizing the course content covered and highlight 1 technology and Best Practice of your choice as part of the summary	1-2 pager learning summary	Due 2/18 25% of grade
Feb 18	Application of data mining and Visualization tools & techniques Analysis of Industry Application	Written Exercises and Group Presentations	Due Same Day 10%

Final Assignment	Completion of Case Study Exercise	Design a case study using the instructor provided template to outline the details of a case study of personal interest applying the concepts learned in this course	Due 2/25 NLT than 6pm to instructors email referenced above - 50 % of grade

The requirements of this class include participation in several class exercises (independent and small group). You are expected to demonstrate your knowledge through participation in class and small group discussions as well as completion of all exercise assignments. There are only two sessions so you must attend both in order to receive full credit for the course which includes the submission of an individual written assignments.

Students are expected to attend all classes, be on time for class and stay for the entire length of class unless cleared with the professor in advance due to professional or other responsibilities elsewhere. Assignments are expected to be completed and to be turned in on time.

✍ I.N OF December 21, 2016

THIS INFORMATION IS SUBJECT TO CHANGE

College of Health & Social Sciences

Standard Wording for All Syllabi

Spring 2017

From: Dean Alvin Alvarez

B. CHSS Policy

Syllabi are to incorporate the Academic Senate Policy regarding finals week, the deadlines or withdrawals, late adds, CR/NC option and information related to course registration and Disability Programs and Resource Center. A sample listing of the required College policies are shown below.

Final Exam: According to Academic Senate policy F76-12 a time period is set aside at the end of each semester for a formal examination period. All classes are expected to meet during the final examination period whether an examination is given or not. The final examination schedule is published each semester in the Class Schedule.

http : //www, sf su. edu/~acadre s/final_exam s/final s 17. htm

CHSS Withdrawal Policy: The last day to drop a class is February 10ᵗʰ, 2017 until 11:59pm. Administrative drops will no longer be permitted or approved after February 10ᵗʰ. Starting February 11ᵗʰ - April 24ᵗʰ you must submit a paper withdrawal petition. Withdrawal from a class starting **February 11ᵗʰ will** be considered for *serious and compelling* reasons only and **must have accompanying documentation.** The following reasons are **not** considered *serious and compelling'.* Changing your major, poor performance, class not required for graduation/major, lacking the pre-requisite, instructor forgot to drop me, not attending class or more time needed for other classes. If you wish to withdraw from class due to unexpected changes in your work schedule, illness or family emergencies, **documentation will be required,** along with a copy of unofficial transcripts. Submit your petition **within a reasonable timeframe (e.g., within 2 weeks of a change in work hours, illness, etc.).** From **April 25ᵗʰ - May 16ᵗʰ, 2017, you** may not withdraw from a class or the University, except only in the case of a **documented** serious illness or verified accident.

Withdrawals **cannot** be initiated electronically and must be submitted using a paper application. All electronic submissions of withdrawals will be denied automatically by the Associate Dean. **You are only allowed to withdraw from a maximum of 18 units and take a class no more than 2 times at SF State.** Approval from the instructor and/or Chair does not constitute automatic approval from the Associate Dean so continue attending class until a decision is made. Please refer to the following website for further information on withdrawal polices: http://chss.sfsu. edu/ sr c.

CR/NC Option: The last day to request CR/NC option is March 17ᵗʰ, 2017 until 11:59pm. The Associate Dean will not approve requests for changes if you miss this deadline.

Late Add Policy: The period to add classes via permission numbers is January 23ᵗʰ - February 10ᵗʰ. 2017 until 11:59. It is your responsibility to procure a late permission number from your instructor and add the class. Faculty cannot add you into a class. Starting **February 11th.** a Waiver of College Regulations petition must be signed by your instructor, Chair and CHSS Associate Dean to add. This will be approved only if there was an administrative error.

Check your registration through *SF State Gateway* : Sign up for CR/NC, drop and add classes by the appropriate deadline online through *SF State Gateway.* ALWAYS check your registration after making any changes and BEFORE deadlines to be sure you are registered properly for your classes. Deadlines for all registration procedures, including withdrawals and requests for credit/no credit, are listed in the class schedule and will be strictly adhered to by the instructor, the Department Chair and the Associate Dean of College of Health & Social Sciences. **It is ALWAYS the student's responsibility to ensure their schedule is correct, even if the instructor indicates they will drop you.**

The information contained here can be viewed on the Registration Calendar at the following website:

http://www.sfsu.edu/~admisrec/reg/regsched2173.html

Disability Programs and Resource Center: Students with disabilities who need reasonable accommodations are encouraged to contact the instructor. The Disability Programs and Resource Center (DPRC) is available to facilitate the reasonable accommodations process. The DPRC is located in the Student Service Building and can be reached by telephone (voice/TTY 415-338-2472) or by email tdprc@sfsu.edut.

Student Disclosures of Sexual Violence: SF State fosters a campus free of sexual violence including sexual harassment, domestic violence, dating violence, stalking, and/or any form of sex or gender discrimination. If you disclose a personal experience as an SF State student, the course instmctor is required to notify the Dean of Students. To disclose any such violence confidentially, contact:

The SAFE **Place** - (415) 338-2208; http://www.sfsu.edu/~safe_plc/

Counseling and Psychological Services Center - (415) 338-2208; http://psyservs.sfsu.edu/

For more information on your rights and available resources: http://titleix.sfsu.edu

Note on Plagiarism

Whenever words are taken directly from another author without quotation marks and direct attribution to the author, it constitutes plagiarism which is a serious and punishable offense at this university. Please ensure that your papers are written in your own words, that ideas and facts taken from others are clearly attributed to them, and that anything taken verbatim from another source is enclosed in quotation marks. Citations can take the form of footnotes or in-text references, but there should be a full bibliographic reference somewhere in the document. In the case of a direct quotation, the citation should include the page number where the quote was found as well. Direct quotes from another author should be used sparingly, and only when the point you are making is stated best by using someone else's words.