

University of Southern California
MARSHALL SCHOOL OF BUSINESS
Summer 2015-PM CLASS– First Draft

Course Guidelines & Syllabus

DSO 528 – DATA WAREHOUSING, BUSINESS INTELLIGENCE AND DATA MINING -16274R
(AKA – Big Data Analytics)
Tuesday 6:00-10:00 p.m. in JKP104

Instructor: Dr. Arif Ansari
Office: BRI 401 R (Bridge Hall 401R)
Office Hours: Tuesday 5:00-6:00 p.m. and by appointment
Office phone: (213) 821-5521
Email: aansari@usc.edu
Emergency Contact number: 213-740-0172

COURSE OBJECTIVES

- Introduce Data Warehousing (DW) , Big Data (BD) and Business Intelligence
- To develop data mining skills to monetize data
- To develop desktop level data mining skills using SAS JMP software and introduce industry level “SAS-Enterprise Miner” tool.
- To provide hands-on skills of building predictive models
- To develop Critical thinking and Modeling Skill Set.
- Introduce “MAGIC” framework for building efficient predictive models.

EXPECTED LEARNING OUTCOMES

- Ability to visualize Business Analytical problems, formulate and solve
- Ability to build predictive models
- Hands on data mining skills to monetize data
- Familiarity with SAS JMP software and introduce industry level “SAS-Enterprise Miner” tool.
- Understand the Big Data Framework
- Familiarity with “MAGIC” framework for building efficient predictive models.
- **For MANAGERS – BE WELL INFORMED ABOUT ANALYTICS**

COURSE STRUCTURE

- **70% of the class will be focused on Data Mining**
- **10% on Business Intelligence tools**
- **20% on Data warehousing and Big Data Tools**

Overview:

Big Data has made available a lot of Structured and Unstructured data along with its new Paradigms for Monetizing data. Many of these new paradigms are ill-defined problems and businesses are struggling to develop methods and models to leverage the “Big Data Opportunity.” Most entry level Analytics professionals are “man-in-the-middle” between data and business requirements; they don’t know how to fully utilize all the available data and are not fully aware of the domain expertise needed to build efficient business savvy models. This course is for students who want to be “Standing on the Shoulders of Giants (Big Data Analytics)” and have great vision on the data side and on the business side, understand Big Data - its potential and drawbacks, Statistics - its usefulness and limitations, data mining - its usefulness and limitations, Business needs and available opportunity. In short this class is about how to monetize Data in customer facing applications.

Companies have huge amount of data in their data warehouse and have access to Big Data through 3rd party APIs. Companies want to leverage data for decision making by building “Data Driven Decision Making Models” and they want to monetize big data using data mining (DM) and Business Intelligence.

The access to social, demographic, transactional, click-stream, web usage etc., data has made companies “data rich” and now they want new ways to monetize data as well as enhance the traditional predictive models using Enriched data. For example, Fortune 500 companies such as American Express, Wells Fargo and Wal-mart have accumulated a great deal of data from their day to day business now they want to monetize the data by providing value to customers and sell their products and services through Omni channels in an efficient manner.

In the Big Data Analytics space what are really critical are information, knowledge, insight and monetization. Some of the questions are: what is the utility of the data? How can one use data in managing customer relationship and empowering employees? How can one uncover patterns and relationships hidden in databases? How can one creatively find ways to monetize data through analytical models? How can one enhance the performance of existing models?

In summary, managers need to understand the strategic values of their company's information assets, be capable of building analytical models to monetize data, understand the models built by third party companies, be able to extract insights for the models and be able to visualize data and insights.

COURSE GOALS:

After taking this class, students should be able to:

- Understand the basic terms that are used in DW, BD, On Line Analytical Process (OLAP), BI and DM
- Choose appropriate tools for specific purposes of storing, integrating and analyzing data (business consideration).
- Identify a Business Analytics Project, collect Data, White Board, Story Board the problem, Build Analytical Models, evaluate performance and monetize it.

- Use Enterprise Miner and JMP to perform DM activities on moderately large data sets.
 - Articulate and present the results of their analyses and the business implications of these results
 - Gain inference from your analysis, from Business and Statistical point of view.
 - Solve ill-defined business analytics problems as well as improve existing Predictive Analytics Models.
- 1) In Data Mining you will develop in-depth skill set to do desktop Data Mining and learn the industry level Data Mining tool.
 - 2) In Data Warehousing/Big Data (DW/BD) part you will learn, why companies need DW/BD, advantages of DW/BD and how to create a DW schema that an executive will understand, I will not teach the hands-on programming for DW part, DW programming part is made available through Teradata student network and you can learn on your own.
 - 3) In Business Intelligence you will learn what current BI can do, how to develop the requirements of a BI system for a company. I will not teach the hands-on programming part, programming part is made available through Teradata student network (and Tableau) and you can learn on your own.
 - 4) You will learn how the 3 parts are interconnected and integrated to form the basis of corporate knowledge system. How to leverage them to convert your company to near real-time corporation. How to monetize data.
 - 5) **Identify, Conceive, Formulate and Solve Predictive Analytics Problems.**

Structure of lectures:

DSO 528 will be organized in a way that includes some combination of the following: lectures, case-based class discussion, group project, computer lab work, and guest lectures.

This class is designed in such a way that only limited mathematical and statistical (Descriptive Statistics, Hypothesis testing and Regression) background is required. I will give a brief review on the above mentioned topics. Learning and understanding underlying DW/BD concepts, studying cases, applying DM/BD ideas and methods to business data, and communicating ideas and solutions will be our main theme. Technical details of selected DM methods will be discussed. Students are expected to use Data Mining software for various cases in class.

COURSE REQUIREMENTS

1. **Class Attendance & Participation.** I strongly suggest that you attend all classes. I strongly encourage, as well as expect, questions during the lectures. I am always accessible by e-mail, and will be more than happy to speak with you before or after class or during office hours.

2. **Class Work (Cases).** We will analyze 5 cases during the semester (You will turn in only 4 cases, case 3 is optional). The 4 cases will be evaluated and will be counted towards the class work points. The Cases can be done in groups of two or three.
3. **Topic Presentation (Optional – Strongly recommended students): Each student (or a team of 3 or 4) in the class will be making a 10-minute presentation on a topic related to “Big Data Analytics” – the topic has to be approved by the Professor.**
4. **Group Project.** I strongly believe the students learn the most during the project. The Group will consist of 4 or 5 students. Learning to work in teams is essential and to get different perspective and it will greatly enhance your learning. The project points will be based on the following criteria:
 - a) **Selection of the project and approval of the proposal - 15% (Recursive Process)**
 - b) **Submission of the Data set and descriptive statistics - 15%.**
 - c) **Preliminary report with Analysis and further direction of the project – 30%**
 - d) **Final Project report, Poster and Presentation – 40%**

A word/PowerPoint document of the Final project report is required as well as a hard copy of the Final project report. The groups will also do peer evaluation of the group. The final report will include an **Executive Summary** write-up that translates the quantitative findings into a real-world analysis. You will be expected to participate in the discussion of your project during the semester to share your methodologies and interesting findings in class.

5. **Midterm and Final Exam.** The midterm will also take place at the beginning of class approximately one hour and 45 minutes. You may bring **two sheets (four pages)** containing formulas, definitions etc., to the midterm except solved problems and solved multiple choice questions. For the final, you may bring **four sheets (eight pages)** containing formulas, definitions etc., except solved problems and solved multiple choice questions. ***No make-ups of mid-term or the final will be given.*** You will receive a grade of zero for each missed exam unless you have a written excuse from your doctor or the professor. In case of emergency or approved absence, the professor may decide to give a make-up exam or redistribute the points.

There will be 1 midterm and 1 final exam. They are closed-book.

Grading	Percentage
Class Work	25
Midterm	25
Final	25
Project	25
Total	100

Hands on Cases

We will use cases built on dataset from the following companies,

Kaggle is a platform for data prediction competitions. Companies, organizations and researchers post their data and have it scrutinized by the world's best data scientists. They compete with each other to solve complex data science problems, and the top competitors are invited to work on the most interesting and sensitive business problems from some of the world's biggest companies through Masters Competitions. The students will select a dataset to build the best model.

SAS is the leader in business analytics software and services. SAS Platform for Analytics helps you turn data into insights.

Case 1 – Prof. Ansari's Smart Partyware – “Find the high propensity customers for Celebrating American Arts product”

Learn Key Concepts like God Marketing, Decision Tree, Beating Computer Models, Understanding Data Mining Metrics, KPIs, Lift and Monetization

The Smart Partyware Company's business model is direct-to-consumer marketing. Over the years they have gained dedicated upscale customers and currently have 500,000 members in their database.

In the direct-marketing industry, the response rate is measured as a percentage of customers who buy the directly mailed product. Smart Partyware's historical response rate for direct mail to selected members is approximately 10%—far above the industry average. SPW was using RFM (Recency-Frequency-Monetary) analysis to target customers. Smart Partyware wants to increase the response rate well beyond the 10% rate.

SPW designs new party ware for every campaign, gives a new name to its party ware, and broadly classifies the party ware under one of its many party themes. Most of the designs cut across many themes but are classified into a particular category based on the main design theme in the party ware. The recent product to be marketed is Celebrating American Arts. It has famous American art works printed in the party ware and the objective is to find high propensity customers for the current marketing campaign. In this case the students will be using the standard business analytics methods like, Decision Tree, Logistics Regression and Neural Network.

Case 2 – SAS's “Home Equity”

Learn Key Concepts like Logistic Regression, Profiler, Odds, KPIs, Lift and Monetization

The consumer credit department of a bank wants to automate the decision-making process for approval of home equity lines of credit. To do this, they will follow the recommendations of the Equal Credit Opportunity Act to create an empirically derived and statistically sound credit scoring model. The model will be based on data collected from recent applicants granted credit through the current process of loan underwriting. The model will be built from predictive modeling tools, but the created model must be sufficiently interpretable so as to provide a reason for any adverse actions (rejections). The HMEQ data set contains baseline and loan performance information for 1000 recent home equity loans. The target (BAD) is a binary variable that indicates if an applicant eventually defaulted or was seriously delinquent. This adverse outcome occurred in approximately 20% of the cases. For each applicant, 12 input variables were recorded. The challenge of this case is to predict the BAD loans so the consumer credit department will be able to prevent default and make appropriate changes to the home equity line of credit.

Case 3 (Optional or Makeup) – Kaggle’s “Don’t Get Kicked” – Or Another case using Neural Network

Learn Key Concepts like Neural Network, Transformation, Profiler and Leveraging Multiple DM methods

One of the biggest challenges of an auto dealership purchasing a used car at an auto auction is the risk of that [sic] the vehicle might have serious issues that prevent it from being sold to customers. The auto community calls these unfortunate purchases "kicks". Kicked cars often result when there are tampered odometers, mechanical issues the dealer is not able to address, issues with getting the vehicle title from the seller, or some other unforeseen problem. Kick cars can be very costly to dealers after transportation cost, throw-away repair work, and market losses in reselling the vehicle. Modelers who can figure out which cars have a higher risk of being kick can provide real value to dealerships trying to provide the best inventory selection possible to their customers. The challenge of this case is to predict if the car purchased at the Auction is a Kick (bad buy).

Case 4 – Prof. Ansari’s Smart Partyware – Big Data Approach – Data Blending and Ensemble Modeling

Learn Key Concepts like Big Data, Data Blending, Ensemble Modeling, SAS-Enterprise Miner Tools

The Smart Partyware Company’s business model is direct-to-consumer marketing. Over the years they have gained dedicated upscale customers and currently have 500,000 members in their database.

In the direct-marketing industry, the response rate is measured as a percentage of customers who buy the directly mailed product. Smart Partyware’s historical response

rate for direct mail to selected members is approximately 10%—far above the industry average. SPW was using RFM (Recency-Frequency-Monetary) analysis to target customers. Smart Partyware wants to increase the response rate well beyond the 10% rate.

SPW designs new party ware for every campaign, gives a new name to its party ware, and broadly classifies the party ware under one of its many party themes. Most of the designs cut across many themes but are classified into a particular category based on the main design theme in the party ware. The recent product to be marketed is Celebrating American Arts. It has famous American art works printed in the party ware and the objective is to find high propensity customers for the current marketing campaign. In this case the students will be using the standard business analytics methods like, Decision Tree, Logistics Regression and Neural Network.

Case 5 – Prof. Ansari’s Search Engine Marketing for Smart Partyware – “Selecting the right set of keywords for search engine marketing campaign.”

Learn Key Concepts like Search Engine Marketing, Dash Boards, Clustering and Campaign Management

The Smart Partyware (SPW) Company’s business model is direct-to-consumer marketing. Over the years they have gained dedicated upscale customers and currently have 500,000 members in their database.

Applichem has signed a Memorandum of Understanding (MOU) with SPW. They will acquire 10 percent of SPW for an undisclosed sum and have an option to buy up to a total of 49% in the following year at current valuation determined by independent evaluators.

John Runner one of the founders of SPW has a vested interest—he wanted to increase the revenue and profit of SPW so that the valuation of SPW in a year will be high and Applichem will have to pay more for the shares of SPW. John and other executives’ contracts with SPW allowed them to sell up to 25% of their shares as part of the deal with Applichem. John Runner was sure his prodigy Vijay would be able to do his magic once again and would be able to increase revenue and profit.

Vijay knew he had fully leveraged the power of data mining; increasing the efficiency of the algorithms would not increase the revenue and profit by 50%. His first approach was to buy a potential member list from data brokers to increase the number of members at SPW. This approach was not successful; the additional revenue from new members was not substantial. In fact, the profit from new members was negligible after taking into account the amount of money paid for the data acquisition and the cost of phone-based marketing to enroll them as new members. The second approach was revamping the site and doing Search Engine Marketing (SEM).

SPW signed up with Google AdSense and created an account with Google. Based on “Partyware” search wording Google AdSense gave a list of nearly 800 keywords and

phrases that people normally search, along with the level of competition, the number of local monthly searches, and the approximate cost per click (CPC). The total amount spent per month on the “partyware” keyword was approximately \$250,000. SPW agreed to allocate \$20,000 for ad budget the first month, and based on the success or failure the next month’s budget would be decided. Based on the keyword bidding SPW wants to sign-up as many visitors to its website as members and increase its membership base.

The challenge of this case is to find the right cluster(s) of keywords for SEM campaign so that many prospects will visit the website and join as members.

Course Materials. The following items will be necessary for completion of reading assignments and homework and successful completion of the course.

1. Online Resources

➤ Sign up with Teradata University Network

Teradata University Network (www.TeradataUniversityNetwork.com) is a **free** learning portal designed to help faculty to teach, learn, and connect with others in the fields of data warehousing, DSS/Business Intelligence, and database.

Teradata University offers web-based courses and related web sites on data warehousing, DSS/BI and database. They have a library of Teradata white papers. Students can become Teradata certified. We will use their material and software in the class particularly for the Business Intelligence and Data warehousing part of the class.

Students register for www.TeradataStudentNetwork.com and login using the current password: **UnifiedDataArchitecture**

➤ Sign up with IBM → Currently working with IBM to get their Academic Resources.

2. Text Books and Class notes

Problems:

- 1. Big Data Analytics is a wide area to be covered in a single textbook.**
- 2. In general the text books are either business oriented (less emphasis on techniques and methods) or subject oriented (less emphasis on real world problems).**
- 3. The text books tend to solve standard problems not the ill-defined problems that are common in Big Data Analytics world.**

Solution:

- 1. I will cover most of the topics in my PowerPoint slides. In addition, I have given access to Teradata University which has a lot of material on Data**

- warehousing, Databases, BI etc. I will post information from IBM to cover Big Data material.**
- 2. The Minicases will give you hands on experience in building models and using tools.**
 - 3. The Project will help you to learn how to solve ill-defined Analytics Problems.**

The information below is additional resource for you.

- The first book is a standard book for Data Mining, the book talks about the various techniques and it is written from computer science perspective. **(Recommended)**

Data Mining: Concepts and Techniques, Second Edition by Jiawei Han and Micheline Kamber, Morgan Kaufmann Publishers, ISBN 13: 978-1-55860-901-3, ISBN-10: 1-55860-901-6, website: www.mkp.com

Note: The book is written from a Computer Science perspective and it will help you to understand the data mining techniques but it does not have real world business application – Buy the book if you want to understand Data Mining Algorithms. My PowerPoint slides will cover the data mining topics but not in depth.

- The second book is a standard book for Direct Marketing, the book talks about the various techniques and it is written from Marketing perspective. **(Recommended)**

The New Direct Marketing: How to Implement A Profit-Driven Database Marketing Strategy
Hardcover: 736 pages
Publisher: McGraw-Hill; 3 edition
ISBN-10: 0070580561
ISBN-13: 978-0070580565

Note: The book is written from a Marketing perspective and it will help you to understand the competing techniques– Buy the book if you want to understand new methods and current methods used in Marketing.

- The third book is from SAS – The world's leading Data mining software company. This book introduces you to industry level Data mining software – SAS Enterprise Miner. **(Recommended)**

Data Mining Using SAS Enterprise Miner – A Case Study Approach, Second Edition. ISBN 1-59047-190-3, SAS publishing
website: www.sas.com

Note: The book helps you to get hands on experience in real-world tool and teaches how to solve well-defined problems.

- The fourth book is from SAS (I am currently reviewing this detailed technical book). This book introduces you to industry level Data mining software – SAS Enterprise Miner. **(Recommended)**

Data Mining Using SAS Enterprise Miner by **Randall Matignon**, John Wiley and Sons
website: www.sasenterpriseminer.com

Note: The book helps you to learn SAS model building methods and get hands on experience in real-world tool and teaches how to solve well-defined problems.

- *Class notes.* Class notes for this class will be available on blackboard. You should familiarize yourself with these notes before they are covered in class.

Recommended (If you want to concentrate on Data Warehousing)

- Building the Data Warehouse 3rd Edition, W.H. Inmon, Wiley, ISBN 0-471-08130-2
- Data Warehouse: Practical Advice from the experts, Joyce Bischoff and Ted Alexander, Prentice hall, ISBN 0-13-577370-9
- Recommended: Data Warehousing: using the Wal-mart model. *Paul Westerman*, Morgan Kauffman publishers.

Recommended (If you want to concentrate on Analytics, Business Intelligence and Forecasting)

1. Information Dashboard Design: The Effective Visual Communication of Data (Paperback) –by Stephen Few (Author)

http://www.amazon.com/Information-Dashboard-Design-Effective-Communication/dp/0596100167/ref=sr_1_1?ie=UTF8&s=books&qid=1240440281&sr=1-1

2. Balanced Scorecards & Operational Dashboards with Microsoft Excel (Paperback)
by Ron Person (Author)

http://www.amazon.com/Balanced-Scorecards-Operational-Dashboards-Microsoft/dp/0470386819/ref=sr_1_1?ie=UTF8&s=books&qid=1240440389&sr=1-1

3. Competing on Analytics: The New Science of Winning (Hardcover)
by Thomas H. Davenport (Author), Jeanne G. Harris (Author)

http://www.amazon.com/Competing-Analytics-New-Science-Winning/dp/1422103323/ref=sr_1_1?ie=UTF8&s=books&qid=1240442477&sr=1-1

4. Introduction to Time Series and Forecasting (Hardcover)
by Peter J. Brockwell (Author), Richard A. Davis (Author)

http://www.amazon.com/Introduction-Time-Forecasting-Peter-Brockwell/dp/0387953515/ref=sr_1_1?ie=UTF8&s=books&qid=1240442944&sr=1-1

Important dates: (Refer to Schedule of classes for up-to-date information)

Midterm exam:

July 14th, 2015

Final Exams:

August 11th, 2015

MARSHALL GUIDELINES

Learning Goals

In this class, emphasis will be placed on the USC Marshall School of Business learning goals as follows:

Goal	Description	Course Emphasis
1	Our graduates will understand types of markets and key business areas and their interaction <i>to effectively manage different types of enterprises</i>	High
2	Our graduates will develop a global business perspective. They will understand how local, regional, and international markets, and economic, social and cultural issues impact business decisions <i>so as to anticipate new opportunities in any marketplace.</i>	Moderate
3	Our graduates will demonstrate critical thinking skills <i>so as to become future-oriented decision makers, problem solvers and innovators</i>	High
4	Our graduates will develop people and leadership skills to promote their effectiveness as <i>business managers and leaders</i>	Low
5	Our graduates will demonstrate ethical reasoning skills, understand social, civic, and professional responsibilities <i>and aspire to add value to society</i>	Low
6	Our graduates will be effective communicators <i>to facilitate information capture and flow in organizational, social, and intercultural contexts</i>	Moderate

Email: The following email protocol is required for effective administration of the class.

- The subject part of the email should start with DSO528, example, DSO528 – question on today’s class.
- When you email the grader the professor should be carbon copied.
- Don’t email from non-USC accounts, most of the time my spam filter will filter out your message. Try to use USC email as much as possible
- Don’t expect a reply for the email on weekends; I will try my best to answer your emails in a timely manner.
- Don’t send many emails one after the other, collect your thoughts and compose the email.

Academic Integrity. Academic dishonesty of any type will not be tolerated in this class. Students who find this statement ambiguous should consult the Student Conduct Code, page 83, of the USC *SCampus* handbook.

A comment about writing the assignments up individually and working in teams: You can work together in teams to discuss the problems and concepts. However, you are required to write up the assignments individually. This means that all the words in your assignments are your own, and you generate all of your own computer output and graphs.

Now, while correct solutions will have very similar or even the same computer output, no two answers should be phrased the same way. If I find two or more assignments that are highly similar, I will at a minimum give the homework a zero, and may refer the incident to the Dean. *Do not test me on this policy.*

STUDENTS WITH DISABILITIES

Any student requesting academic accommodations based on a disability is required to register with Disability Services and Programs (DSP) each semester. A letter of verification for approved accommodations can be obtained from DSP. Please be sure the letter is delivered to me as early in the semester as possible. DSP is located in STU 301 and is open 8:30 am - 5:00 pm, Monday through Friday. The phone number for DSP is 213 740-0776.

Tentative Schedule:

- I. The course will start with Data Mining. The Data Mining part of the class will be quantitative and the following topics will be covered in it.

1. Standard Data Mining techniques:
 - a. Classification
 - b. Clustering
 - c. Association
 - d. Prediction (Similar to what is covered in core MBA class)
 - e. **Text Mining, Link Analysis, Visual Data mining (if time permits)**

Using various appropriate techniques,

- i) Bayesian Estimation
- ii) Neural Networks
- iii) Decision Tree
- iv) Similarity Measures
- v) **Other techniques like Ensemble Modeling (if time permits)**

2. Statistical Model Building using Data Mining methods and Logistic Regression.

Depending on the project other topics may be covered.

- II. The second part of the course will be Business Intelligence Software. You will be introduced to software used as Business Intelligence software.

- III. The third part of the course will be Data Warehousing/Big Data. You will be introduced to Data Warehousing from business perspective, how to create Data Warehouse Architecture. Big Data Platforms will be discussed.
- IV. In addition to the above, if time permits we will discuss the following.
 - a. Hadoop
 - b. MOBI
 - c. Advanced Data Mining Concepts

Approximate Schedule of class**TUN – Teradata University Network, SAS – Enterprise Miner Text book****JM – Data Mining textbook by Jiawei Han and Micheline Kamber****Bring Your Laptop to class with JMP software**

Date		Topic	Reading from textbooks	Reading	Due/Other
June 2	1	Introduction to Big Data Analytics Classification - Distance Based Algorithms / Evaluating Classification Critical Thinking Exercise / Project Ideas	JM 1-26, JM 36-40, JM 359-362, JM372-375 JM285-290, JM 310-318, JM 347-350, SAS 39-67JM 291-306	Dr. Ansari Notes	
June 9	2	This will be Short Class By Dr. Abbass Descriptive Statistics/Classification Methods – Decision Tree Based Methods/Case 1 –	JM 291-306 SAS 19-36SAS 39-67, SAS 67-81	Dr. Ansari Notes	Group List Due
June 16	3	Tuple Discussion/Decision Tree Strategies/Review of Hypothesis Testing/Logistic Regression/ Project Discussion	JM 358-359, JM 327-336 JM 358-359, JM 327-336	Dr. Ansari Notes	Project Proposal Due
June 23	4	Logistic Regression / Naïve Bayesian/ Project Refinement	JM 384-414, JM 227-234 JM 384-414, JM 227-234, SAS 91-104, SAS 105-109	Dr. Ansari Notes	Turn in your Class work 1 (Case1)
June 30	5	Introduction to SAS Miner/ Neural Network	TUN relevant information	Dr. Ansari Notes	Turn in your Project Data set and Descriptive Stats
July 7	6	Neural Network/ Data Enrichment via Big Data	JM 227-234, SAS 91-104,	Dr. Ansari Notes	Turn in your Class work 2 (Case2) Turn in your Data set and Descriptive Stats HW1-optional
July 14	7	Midterm/Business Intelligence/ Search Engine Marketing	TUN: Microstrategy BI Information/Continental Airlines Takes Off with Real-time Business Intelligence	Dr. Ansari Notes	Turn in your Class work 3 (Case3) Optional or Makeup
July 21	8	Clustering and Association/ Google Analytics and Adwords	JM 114-123 and TUN relevant information JM 123-126 and TUN relevant information	Dr. Ansari Notes	Turn in your Class work 4 (Case4) Video Record your Topic Presentation - Optional
July 28	9	Lecture DW1: Data Warehousing(I): Strategic View Lecture DW2/BD: What is Big Data Lecture DW3: Dimensionally Designed DW (I&II)	JM 114-123 and TUN relevant information JM 123-126 and TUN relevant information JM 135-137, JM 144-152 and TUN relevant information	Dr. Ansari Notes	Video Record your Topic Presentation – Optional - Turn in your Preliminary report HW2-optional
Aug 4	10	Lecture DW5: OLAP,/More on Big Data / Review for Final Exam/Poster Session ?			Turn in your Class work 5 (Case5)
Aug 11	11	Final Exam/Project Presentation			Oral Presentation of Project and Turn in your Final Report