



School of Engineering
*Information
Technology Program*

ITP 487 – Enterprise Data Analytics

Units: 4

Spring 2020

Section 31829R: Mon & Wed noon-1:50pm

Section 31855R: Mon & Wed 2pm-3:50pm

Location: KAP 163

Professor: Mike Lee

Office Hours: See <http://bit.ly/professorlee>

Contact Info: mikelee@usc.edu

Teaching Assistant: Brian Plotnick

Office Hours: See <http://bit.ly/professorlee>

Contact Info: bplotnik@usc.edu

Teaching Assistant: Victor Qiu

Office Hours: See <http://bit.ly/professorlee>

Contact Info: weikaiqi@usc.edu

IT Help:

USC IT (ITS): <https://itservices.usc.edu/contact/>

Viterbi IT: <https://viterbi.usc.edu/resources/vit/contact-us.htm>

Course Description

While the increased capacity and availability of data gathering and storage systems have allowed enterprises to store more information than ever before, most **organizations still lack the ability to effectively consolidate, arrange, and analyze** this vast amount of data. Analyzing large data sets to forecast and predict future events has become a highly sought-after skill in business, engineering, services, science, health, and other industries.

This course will explore the theory and practice of three major areas:

- Data Warehouses for Enterprises
- Business Intelligence for Enterprise Resource Planning Systems (ERP)
- Business Forecasting
- Cloud Analytics for Enterprises

Learning Objectives

After completing the course, students will be able to

- Describe the components of an enterprise data warehouse
- Model the relational database required for an enterprise data warehouse
- Extract, cleanse, consolidated, and transform heterogeneous data into a single enterprise data warehouse
- Analyze data to generate information and knowledge that lead to informed decisions for businesses
- Show how ERP business intelligence can be derived from data warehouses
- Describe the various forecasting techniques
- Describe components of the cloud analytics and how they compare to enterprise data warehouses

Prerequisite(s): ITP 320 or ITP 249

Course Notes

Lectures are delivered face to face in classroom. Lectures are not recorded so attendance is strongly recommended. All course materials will be made available through Blackboard. These include:

- Lecture slides
- In-class exercises
- Homework assignments
- Readings
- Software details and instructions for accessing Viterbi Virtual Lab
- Grades and feedback
- Office hours
- Online discussion forums will be used for out-of-class discussions

Announcements made in class and content posted in Blackboard will supersede the contents of this syllabus.

Technological Proficiency and Hardware/Software Required

The assignments for this class will include both reading assignments as well as hands-on computer assignments. ***Students must bring their laptop computers to lecture*** sessions to participate in hands-on activities. Students will be given tutorials to gain familiarity with software tools.

Most of the SAP software required for the class is Windows based. The software will be provisioned through the Viterbi Virtual Lab. Specifically, students will be using: (more tools may be added)

- Eclipse with SAP BW Modeling Tools
- SAP BW/4HANA (Business Warehouse)
- SAP GUI
- SAP Analysis for Microsoft Excel
- SAP Predictive Analytics
- Microsoft Excel and Access
- Amazon Web Services (AWS) Data Lakes and Analytics
- Additional tools as needed

VITERBI VIRTUAL LAB – VMWARE VDI

All software can also be accessed into Virtual Desktop by logging in at: <http://mydesktop.vlabs.usc.edu>. See blackboard for additional instructions on installing.

Required Readings and Supplementary Materials

Practical Analytics, Nitin Kale and Nancy Jones, Second Edition, Epistemy Press 2020
<http://store.epistemypress.com/books/analytics.html>

In addition to the required reading and supplementary materials listed in the weekly breakdown section of this syllabus, additional materials will be announced in class and published on Blackboard.

Description and Assessment of Assignments

Homework: Most homework is computer based. Homework should be turned in to Blackboard on time. Grading will be based on completeness, accuracy, and timeliness. Feedback will be provided through Blackboard. These are individual effort assignments.

In-Class Exercises: are guided Q&A and hands-on exercises that are used to spark additional discussion and deeper understanding of the materials and concepts before the student leaves the class. Announcement of in-class exercises may or may not be given prior to the class. In-class exercises can be a team or individual exercises. The score used for grading is the percentage of in-class exercises completed and turned in in-class vs what was assigned in the semester.

Exams: are in-class tests with multiple choice, free response, and/or live tool sections. As an example, students may be required to create an SAP query and analyze the results on a live enterprise data warehouse system as part of the “live tool” portion of the exam.

Final Project: Final project is an individual summative assignment where you will be applying most of the skills that you have learned through the semester. Each semester may have a different final project, but will include modeling the relational database required for an enterprise data warehouse, extract/cleanse/consolidate/transform heterogeneous data into a single enterprise data warehouse, and analyze data.

Project Timeline: (may differ slightly based on the individual pace of the class)

- Week 13: Project assigned
- Week 14: Solution approach and design
- Week 16: Final project due

Grading Breakdown

Homework	30%
In-Class Exercises	10%
Exam I	25%
Exam II	25%
Final Project	10%
TOTAL	100%

Grading Scale

Final grades represent how you perform in the class relative to other students. Historically, the average grade for this class is about a 3.4.

Assignment Submission Policy

It is the responsibility of the student to make sure problem solution and assignment are turned in on time. Make sure you follow the procedures outlined in each assignment (Blackboard submissions).

Students are encouraged to work with their classmates. However, students must turn in their own, original work. Late homework submissions will be subject to a late penalty. The penalty is 25% per day. No assignments will be accepted later than four days from after the due date.

Grading Timeline

Assignments Grading will typically be completed 7 days after submission. Any variations will be announced in class or on blackboard.

Additional Policies

No make-up exams (except for documented medical or family emergencies) will be offered nor will there be any changes made to the Final Exam schedule, except as permitted by university rules. The use of mobile devices, books, notes or computers is not permitted during the exam unless explicitly specified by the professor – e.g. use of a computer for the live tool portion of an exam.

Course Schedule:

	Topics/Daily Activities	READING	ASSIGNMENT
Week 1 Jan 13	Course Introduction <ul style="list-style-type: none"> • Course objectives and outcomes Enterprise Data Analytics <ul style="list-style-type: none"> • Why do enterprises need data analytics? • What is a data warehouse? • Various types of data repositories Log into VMware (bring laptops)	Lecture 0 Lecture 1	
Week 2 Jan 20	NO CLASS ON MONDAY – MLK DAY Data & Relational Database Concepts <ul style="list-style-type: none"> • Types & Instances • Tabular vs Multi-Dimensional Data • Relations, attributes, relationships • ER Diagrams • Database Normalization, normal forms • Denormalization of tables IC #1: Data Modeling (ER Diagram & Access)	Lecture 2	
Week 3 Jan 27	SQL Review <ul style="list-style-type: none"> • SQL • JOINS • Introduction to in-memory databases IC#2: SQL (SQL on Access & Teradata)	Lecture 3	HW #1 (SQL on Teradata)
Week 4 Feb 3	Data Warehousing Concepts <ul style="list-style-type: none"> • Transactional vs Master data • Types and sources of data • Transactional databases vs. data warehouses • Enterprise data warehouses • Data store objects • Multidimensional Model for data warehouses Star Schema <ul style="list-style-type: none"> • Dimension and fact tables Snowflake Schema <ul style="list-style-type: none"> • Difference between star schema and snowflake schema • Master data tables <ul style="list-style-type: none"> o Attributes – Display, Navigational o Texts o Hierarchies (e.g. Geo) IC #3: Schemas	Lecture 4	HW #2 (Star Schema)
Week 5 Feb 10	Data Warehousing Implementation <ul style="list-style-type: none"> • SAP HANA & BW/4HANA • InfoObjects • Source Systems • Data Source 	Lecture 5	HW #3 (Modeling Tool)

	<ul style="list-style-type: none"> Transformation / Data Transfer Process Advanced DataStore (ADSO) Composite Provider Query Analysis Tools IC #4: SAP BW/4HANA Components (Eclipse) Exam I Review		
Week 6 Feb 17	NO CLASS ON MONDAY – PRESIDENT’S DAY Exam I – Feb 19 in CLASS		
Week 7 Feb 24	InfoObjects: Key Figures & Characteristics <ul style="list-style-type: none"> InfoObjects: Characteristics and Key Figures Creating InfoObjects Handling aggregations – standard and exception Handling time dependency Handling language dependency IC #5: Characteristics	Lecture 5	HW #4 (Key Figures & Characteristics) HW #5 (Master Data Load)
Week 8 Mar 2	InfoProviders: ADSO & Composite Providers <ul style="list-style-type: none"> ADSO Composite Provider Defining ADSO Defining a Composite Provider IC #6: InfoProviders (ADSO & Composite Providers)	Lecture 6	HW #6 (ADSO & Composite Providers)
Week 9 Mar 9	Extraction, Transformation, and Loading (ETL) <ul style="list-style-type: none"> Source systems Data Sources Extractors for data (APIs etc.) Mapping of fields Transformation rules Data cleansing and harmonization Data flow objects Fact Table Loading IC #7: Data Cleansing IC #8: Fact Table Loading	Lecture 7	HW #7 (Data Source & Fact Table Load)
Mar 16	SPRING BREAK – NO CLASS		
Week 10 Mar 23	Building queries IC #9: Data Analyst: Queries	Lecture 8	HW# 8 (Build Query in Eclipse)
Week 11 Mar 30	Slicing and Dicing <ul style="list-style-type: none"> Basics of slicing and dicing Pivot tables Working with aggregation functions, hierarchies Slicing and dicing multidimensional data IC #10: Business Analyst: SAP Analysis for Excel	Lecture 9	HW# 8 (Analyze Data using SAP Analysis)
Week 12 Apr 6	Slicing and dicing contd. <ul style="list-style-type: none"> Currency conversion 	Lecture 9 (cont)	

	<ul style="list-style-type: none"> • Hierarchies • Charting Exam II Review		
Week 13 Apr 13	Business Forecasting <ul style="list-style-type: none"> • Time series analysis • Forecasting Final Project Overview Exam II – APRIL 15 in CLASS	Lecture 10	HW #10 (Forecasting) Final Project
Week 14 Apr 20	Cloud Analytics <ul style="list-style-type: none"> • Overview • Comparison to Enterprise Data Warehouse • Tool Mappings • Use Cases IC #10: AWS Part 1	Lecture 11	HW #11 (AWS)
Week 15 Apr 27	Cloud Analytics - Design <ul style="list-style-type: none"> • Data Catalog • Data Movement • Queries IC #11: AWS Part 2		
Week 16 Finals Week	Final Project Due		

Statement on Academic Conduct and Support Systems

Academic Conduct

Plagiarism – presenting someone else’s ideas as your own, either verbatim or recast in your own words – is a serious academic offense with serious consequences. Please familiarize yourself with the discussion of plagiarism in *SCampus* in Part B, Section 11, “Behavior Violating University Standards” <https://policy.usc.edu/scampus-part-b/>. Other forms of academic dishonesty are equally unacceptable. See additional information in *SCampus* and university policies on scientific misconduct, <http://policy.usc.edu/scientific-misconduct>.

Support Systems

Student Counseling Services (SCS) - (213) 740-7711 – 24/7 on call

Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention. <https://engemannshc.usc.edu/counseling/>

National Suicide Prevention Lifeline - 1-800-273-8255

Provides free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week. <http://www.suicidepreventionlifeline.org>

Relationship and Sexual Violence Prevention Services (RSVP) - (213) 740-4900 - 24/7 on call

Free and confidential therapy services, workshops, and training for situations related to gender-based harm. <https://engemannshc.usc.edu/rsvp/>

Sexual Assault Resource Center

For more information about how to get help or help a survivor, rights, reporting options, and additional resources, visit the website: <http://sarc.usc.edu/>

Office of Equity and Diversity (OED)/Title IX Compliance – (213) 740-5086

Works with faculty, staff, visitors, applicants, and students around issues of protected class. <https://equity.usc.edu/>

Bias Assessment Response and Support

Incidents of bias, hate crimes and microaggressions need to be reported allowing for appropriate investigation and response. <https://studentaffairs.usc.edu/bias-assessment-response-support/>

The Office of Disability Services and Programs

Provides certification for students with disabilities and helps arrange relevant accommodations. <http://dsp.usc.edu>

Student Support and Advocacy – (213) 821-4710

Assists students and families in resolving complex issues adversely affecting their success as a student EX: personal, financial, and academic. <https://studentaffairs.usc.edu/ssa/>

Diversity at USC

Information on events, programs and training, the Diversity Task Force (including representatives for each school), chronology, participation, and various resources for students. <https://diversity.usc.edu/>

USC Emergency Information

Provides safety and other updates, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible, <http://emergency.usc.edu>

USC Department of Public Safety – 213-740-4321 (UPC) and 323-442-1000 (HSC) for 24-hour emergency assistance or to report a crime.

Provides overall safety to USC community. <http://dps.usc.edu>