

ITP 349 – Enterprise Data Analytics

Units: 4

Fall 2024, Noon-1:50PM MW

Location: SGM 101

Instructor: Mike Lee

Contact Info: mikelee@usc.edu
Office Hours: bit.ly/professorlee

Learning Assistants:

Lead: Leilani Ventura (<u>lcventur@usc.edu</u>)
 Section: Allen Mercado (<u>allenjos@usc.edu</u>)

 Section: Pumarie Madden (pamadden@usc.edu)

• Section: Cindy Xia (cindyx@usc.edu)

Section: TBD

See bit.ly/professorlee for latest info

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. Digital transformation using data analytics techniques 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 the following areas:

- Enterprise Organizational Structure and Decision Making
- Enterprise Data Warehouses
- Data Analytics used by Enterprises
- USC Applied Data Analytics Methodology (ADAM)

Learning Objectives

After completing the course, students will be able to

- Understand the organizational structure of enterprises (large organizations)
- Understand how enterprises make major technology decisions
- Define enterprise data analytics and its drivers
- 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
- Explore any data set and apply a repeatable approach to data analytics to gain relevant insights
- Apply data analytics techniques that is in demand by enterprises

Page 1 of 14

This content is protected and may not be shared, uploaded, or distributed.

Prerequisite(s): ITP 320 or ITP 249

Remote Attendance

This course does not support remote attendance. Lectures will not be recorded or available on Zoom, there are short in-person individual/group activities during many class meetings and exams are in-person.

Course Notes

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.

USC Technology Support Links

Zoom information for students
Blackboard help for students
Software available to USC Campus

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 (phones/tablets are not sufficient) to lecture sessions to participate in hands-on activities. Students will be given tutorials to gain familiarity with software tools.

Most of the enterprise software required for the class is Windows based or delivered via the cloud. The software will be provisioned through the Viterbi Virtual Lab, Amazon Web Services, Google Cloud, SAP Cloud and/or installed your computer. Specifically, students will be using:

- SAP BW/4HANA and/or SAP Analytics Cloud
- Eclipse for SAP BW Modeling (installed locally or Viterbi Virtual Lab)
- SAP Analysis for Microsoft Excel (installed locally or Viterbi Virtual Lab)
- Amazon Web Services/RDS (cloud)
- ChatGPT and other Generative AI tools (semester specific)
- Relational Database/SQL (cloud)
- Google Colab/Python/Pandas (cloud)
- Github (cloud)
- Python/Pandas
- SOI
- * Microsoft Power BI (optional windows only)
- * Google Big Query/Shopify (optional)

VITERBI VIRTUAL LAB – VMWARE VDI

^{*} bonus exercises that students may use to add to resume skill sets

Some software can also be accessed via Virtual Desktop by logging into the General Desktop at: http://mydesktop.vlabs.usc.edu as the VDI server. See blackboard for additional instructions on installing.

Alternatively, you can install the required software on your Windows machine (no support will be provided). Instructions will be posted on Blackboard.

Readings and Supplementary Materials

Reading and supplementary 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. Grading will be based on completeness, accuracy, and timeliness. Feedback will be provided through Blackboard. These are individual effort assignments. One homework assignment will be dropped (lowest score) from your grade calculation.

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 team or individual exercises. The score used for grading is the percentage of in-class exercises completed and turned in in-class. Two in-class exercises will be dropped (lowest scores) from your grade calculation.

Exams: Each exam will be comprised of 1) in-person and in-class multiple-choice part during class-time and 2) take-home project that you will have several days to complete. Details will be posted on Blackboard.

Grading Breakdown

Homework	30%
In-Class Exercises	10%
Midterm	25%
Quiz	10%
Final	25%
TOTAL	100%

Grading Scale

Course final grades will be determined using the following scale:

Α	95-100
A-	90-94
B+	87-89
В	83-86
B-	80-82
C+	77-79
С	73-76
C-	70-72
D+	67-69
D	63-66
D-	60-62
F	59 and below

Grading Timeline

Grading will typically be completed 7 days after submission. Any variations will be announced in class or on Brightspace. *Regrade requests must be submitted within a week of the grades being published unless otherwise communicated in class.*

Generative AI Policy

Use of Generative AI technologies, including ChatGPT, are encouraged and allowed unless explicitly stated otherwise. YOU MUST CITE THAT YOU USED THE TECHNOLOGY AND INCLUDE ALL PROMPTS THAT YOU HAVE USED.

Policies

Students are expected to attend and participate in lecture discussions, in-class exercises, and team meetings.

Assignments turned in late will have 25% of the total points deducted from the graded score for each late day.

No make-up exams (except for documented medical or family emergencies) will be offered. If they will not be able to attend an exam due to an athletic game or other valid reason, then they must coordinate with the instructor before the exam is given. They may arrange to take the exam before they leave, with an approved university personnel during the time they are gone, or within the week the exam is given. If students do not take an exam, then they will receive a 0 for the exam. Accommodations religious observance must be arranged with the Professor at least two weeks before the exam.

If students need accommodations authorized by OSAS (Office of Student Accessibility Services), notify the instructor at least two weeks before the exam. This will allow time for arrangements to be made.

Sharing of course materials outside of the learning environment *SCampus Section* 11.12(*B*)

Distribution or use of notes or recordings based on university classes or lectures without the express permission of the instructor for purposes other than individual or group study is a violation of the USC Student Conduct Code. This includes, but is not limited to, providing materials for distribution by services publishing class notes. This restriction on unauthorized use also applies to all information, which had been distributed to students or in any way had been displayed for use in relationship to the class, whether obtained in class, via email, on the Internet or via any other media. (See Section C.1 Class Notes Policy).

Week	Date	Lecture Topic(s)	See Blackboard for Due Dates
1	8/26	Course Introduction	
		What is an Enterprise?	
		What is Enterprise Data Analytics?	
		 USC Applied Data Analytics Methodology (ADAM) 	
		Use Cases	
		Career Relevance	
		Course Overview	
	8/28	Enterprise Structure	
		Structure of enterprises	
		CIO reporting structure	
		Technology decision making	
		Components of a strategy	
		 Strategy tools – Gartner Magic Quadrant 	
2	9/2	NO CLASS – LABOR DAY	
	9/4	Data Concepts	HW1: ER Diagram – you will
		Structured vs Unstructured Data	create a data dictionary and ER
		Type vs Instance	diagram from narratives
		Physical vs Virtual	
		Data Storage Component	TAKE HOME: Install MySQL
		Master Data vs Transactional Data	Workbench on Laptop
		Big Data vs Enterprise Data	
		Relational Database for Analytics	
		 Relational Database for Analytics vs non-Analytics 	
		 Data sets: Narratives, Data dictionary & ER 	
		diagrams	
		 Entity relationship diagram (ERD or ER Diagram) 	
		Crows foot notation	
		Entity relation key attributes relationships cardinality	
		Entity, relation, key, attributes, relationships, cardinality,	

	Т	T	T
3	9/9	Data Modeling for Analytics – Lucid Chart Online ER Diagram Creation Entities, Key, Field, Type Relationships & Crows Foot Notation Exporting ERD to DDL In-Class: ER Diagram (LucidChart) – design ER diagram in LucidChart and export to SQL DDL. Amazon Web Services (AWS) / RDS Overview Amazon Web Services & Relational Data Service Creating a Database Creating and Modifying Schemas Loading Data Querying In-Class: Build Your Lab (AWS/RDS) – you will be creating a	
		database in AWS/RDS, building a simple schema, and testing connectivity from MySQL workbench to your	
	9/11	database SQL for Analytics (part 1) SQL for Analytics vs non-Analytics Data Manipulation Commands CREATE, DROP, INSERT, UPDATE, & DELETE Creating and Modifying Schemas: CREATE SCHEMA Loading Data: USE & INSERT Querying: SELECT SQL for Analytics (part 2) Data Set: Narrative, Data dictionary & ER diagram SELECT & JOIN Loading Data Querying	HW#2 - AWS/RDS – you will be implementing the ER diagram that you previously designed in LucidChart in the lab the you previously built int AWS/RDS. You will also load and query the data using SQL.
4	9/16	 Enterprise Use of Generative AI Enterprise Use of Generative AI/ChatGPT Gartner Access ChatGPT Access Image AI Access 	
	9/18	Hands on Generative AI/ChatGPT	HW#3 – Generative AI/ChatGPT – you will leverage generative AI tools to launch a new product

	0/22	B . W . L . C	
5	9/23	Data Warehouse Concepts What is a Data Warehouse? Transactional (OLTP) vs Data Warehouse Systems (OLAP) Need for Data Warehouse Importance of Master Data Data Warehouse Components Data Warehouse Process Multi-Dimensional Data Multi-Dimensional Modeling Multi-Dimensional Data Storage Tabular vs multi-dimensional data Star Schema Fact tables (transactional data) Dimension tables (master data) Design Star Schema from Transactional Systems Data Warehouse Implementation Identify, Collect, & Clean: Narrative, Data dictionary & ER diagram	
		 Model: Dimension tables (master data) Model: Fact tables (transactional data) 	
6	9/30	Data Warehouse Implementation – Continued In-Class: Design and load Star Schema Datawarehouse (Lucid Chart, AWS RDS, MySQL Workbench)	HW#4 – Star Schema Warehouse – you will load transactional DDL files, design star schema, load the star schema from tables from the transactional tables using staging tables.
	10/2	Midterm Review	<u> </u>
	10/7	NO CLASS – PRESIDENT'S DAY	

7	10/9	Midterm - Part 1	Assigned Video Lecture
	10,0	Take Home Assigned	T. S. B. G. T. G. S. E. C. G.
		Enterprise Data Warehouse Fundamentals	
		 What is an Enterprise Data Warehouse? 	
		Additional Enterprise Needs	
		 Historization (Time Dependency) 	
		 Language Dependency 	
		 Multi-currency/unit of measure 	
		Snowflake Schema	
		o Fact Tables	
		 Dimensional Tables 	
		 Master Data Tables 	
		Master data tables	
		 Attributes – Display, Navigational 	
		o Texts	
		o Hierarchies	
		Difference between star schema and snowflake .	
	40/44	schema	
8	10/14	Midterm – Part 2	
	10/16	Multiple Choice - taken in class Data Warehouse Implementation	
	10/16	Enterprise Data Warehouse Implementation	
		Why SAP?SAP HANA Database	
		SAP BW/4HANA SAP BW/6W Components	
_	10/21	SAP BW Key Components InfoObjects - SAP's Data Warehouse Catalog	LIVA/#E InfoObjects
9	10/21	Enterprise-Wide Definition	HW#5 - InfoObjects
		Characteristics	
		Master Data	
		o Physical Storage	
		Creating Characteristics	
		 Time Dependency (Historization) 	
		 Language Dependency 	
		How Used	
		Key Figures	
		 Creating Key Figures 	
		 Multi-Currency 	
		 Unit of Measure 	
		 Standard and Exception Aggregation 	
		How Used	
	10/23	Loading Master Data: Characteristics	
		 Moving data through the data warehouse 	
		• ETL	
		Process chains	
		 Master data loading into characteristics 	

10	10/28	InfoProviders: Data Containers/Views	HW#6 – InfoProviders
		Advanced Data Stores (ADSOs)	
		Defining an ADSO	
		Composite providers	
		 Defining a composite provider 	
		Changing output	
		Loading Transactional Data: Fact Table Loading	
		Source systems	
		Data Sources	
		Extractors for data (APIs etc.)	
		Mapping of fields	
		Transformation rules	
		Data cleansing and harmonization	
		Transactional data loading into ADSOs	
	10/30	Data Analyst: Queries	HW#7 - Queries
		 Enterprise Analyst Roles & Organization 	
		Query basics	
		Query designer	
		 Sheet definition 	
		o Filter	
		 Free characteristics 	
		 Conditions 	
		 Exceptions 	
		 Calculated key figures / formulas 	
		 Navigational and display attributes 	
		Currency conversion	
		In-Class: Queries	
11	11/4	Business Analyst: Analysis for Excel	HW#8- Analysis for Excel
		Enterprise software nomenclature	
		 Enterprise analyst roles & data organizations 	
		Analysis basics	
		SAP BW Query vs Analysis for Excel	
		Dimensions	
		Members	
		 Hierarchy 	
		 Measures 	
		Filter by member	
		Filter by measure	
		Conditional formatting	
		In-Class: Analysis for Excel	

	1 .		T
	11/6	USC Applied Data Analytics Methodology (ADAM)	
		Importance of an Approach	
		 Methodology: Identify, Collect, Clean, Model, 	
		Analyze, Publish	
		 Toolkits 	
		Use Cases	
		Date Sets*	
		Wordle	
		Electric Bill	
		Student Survey	
		NBA Regular Season Stats	
		* Data sets may change	
12	11/11	Quiz – SAP BW/4HANA (see video for prep)	
	11/13	USC ADAM + Python Toolkit	
	11,10	Google Colab Overview	
		Notebook vs Runtime	
		Connecting to a Runtime	
		Temporary Modules / Files	
		• Github	*
		Python & Libraries	
		Other sources of data	
		o the sources of data	
		In-Class: Build Your Lab (Google Colab & Github) - you will	
		be loading data from Github into Google Colab to perform	
		simple analysis.	
13	11/18	Pandas for Analytics (part 1)	HW#9 – Pandas for Analytics
13	11/10	What is Pandas for Analytics?	Tive 5 - Failuas for Allarytics
		Data Set Overview	
		Data Storage Components	
		Data Storage Components Data Frames: Tabular Data	
		o Series: Column of Data	
		Loading/saving data	
		Viewing data	
		SelectionCleaning data	
		o Remove	
		o Format	
		Model Calculated columns	
		Calculated columns Marga	
		o Merge	
		o Grouping	
	11/20	Analysis	
	11/20	USC ADAM – Hands On	
		In Class Evarging, NRA Statistics	
İ	1	In-Class Exercise: NBA Statistics	1

14	11/25	Pandas for Analytics (part 2)	HW#10 – Data Visualization
14	11/23	Lambda	
			(Video will be assigned, no in-
		Data Visualization	person class)
		 Matplotlib vs Alternatives 	
		o Histogram	
		 Bar & Stacked Bar 	
		o Line	
		 Area & Area Stacked 	
		o Pie & Donut	
		 Scatter & Scattered Bubble 	
		In-Class Exercise: Data Set TBD	
	11/27	NO CLASS – THANKSGIVING HOLIDAY	
15	12/2	Analysis Techniques - beyond simple visualization	HW#11 – Analysis Techniques
		USC ADAM Helpers	
		Correlation Analysis	
		Cohort Analysis	
		Data Visualization Advanced Commands	
		In-Class Exercise: Data Set TBD	
	12/4	Final Exam Review	·
		Course Wrap Up & Questions	
		Industry & Discussions	
		Bonus Projects	
Finals		Final Exam – In-Class Coding Exam	
Exam	1		

^{*} Data sets change each semester. Data set listed is the possible data set that will be used.

Statement on Academic Conduct and Support Systems

Academic Integrity:

The University of Southern California is a learning community committed to developing successful scholars and researchers dedicated to the pursuit of knowledge and the dissemination of ideas. Academic misconduct, which includes any act of dishonesty in the production or submission of academic work, comprises the integrity of the person who commits the act and can impugn the perceived integrity of the entire university community. It stands in opposition to the university's mission to research, educate, and contribute productively to our community and the world.

All students are expected to submit assignments that represent their own original work, and that have been prepared specifically for the course or section for which they have been submitted. You may not submit work written by others or "recycle" work prepared for other courses without obtaining written permission from the instructor(s).

Other violations of academic integrity include, but are not limited to, cheating, plagiarism, fabrication (e.g., falsifying data), collusion, knowingly assisting others in acts of academic dishonesty, and any act that gains or is intended to gain an unfair academic advantage.

The impact of academic dishonesty is far-reaching and is considered a serious offense against the university. All incidences of academic misconduct will be reported to the Office of Academic Integrity and could result in outcomes such as failure on the assignment, failure in the course, suspension, or even expulsion from the university.

For more information about academic integrity see <u>the student handbook</u> or the <u>Office of Academic Integrity's website</u>, and university policies on <u>Research and Scholarship Misconduct</u>.

Please ask your instructor if you are unsure what constitutes unauthorized assistance on an exam or assignment, or what information requires citation and/or attribution.

Course Content Distribution and Synchronous Session Recordings Policies

USC has policies that prohibit recording and distribution of any synchronous and asynchronous course content outside of the learning environment.

Recording a university class without the express permission of the instructor and announcement to the class, or unless conducted pursuant to an Office of Student Accessibility Services (OSAS) accommodation. Recording can inhibit free discussion in the future, and thus infringe on the academic freedom of other students as well as the instructor. (Living our Unifying Values: The USC Student Handbook, page 13).

Distribution or use of notes, recordings, exams, or other intellectual property, based on university classes or lectures without the express permission of the instructor for purposes other than individual or group study. This includes but is not limited to providing materials for distribution by services publishing course materials. This restriction on unauthorized use also applies to all information, which had been distributed to students or in any way had been displayed for use in relationship to the class, whether obtained in class, via email, on the internet, or via any other media. (Living our Unifying Values: The USC Student Handbook, page 13).

Students and Disability Accommodations:

USC welcomes students with disabilities into all of the University's educational programs. The Office of Student Accessibility Services (OSAS) is responsible for the determination of appropriate accommodations for students who encounter disability-related barriers. Once a student has completed the OSAS process (registration, initial appointment, and submitted documentation) and accommodations are determined to

be reasonable and appropriate, a Letter of Accommodation (LOA) will be available to generate for each course. The LOA must be given to each course instructor by the student and followed up with a discussion. This should be done as early in the semester as possible as accommodations are not retroactive. More information can be found at osas.usc.edu. You may contact OSAS at (213) 740-0776 or via email at osasfrontdesk@usc.edu.

Support Systems:

Counseling and Mental Health - (213) 740-9355 - 24/7 on call

Free and confidential mental health treatment for students, including short-term psychotherapy, group counseling, stress fitness workshops, and crisis intervention.

988 Suicide and Crisis Lifeline - 988 for both calls and text messages – 24/7 on call

The 988 Suicide and Crisis Lifeline (formerly known as the National Suicide Prevention Lifeline) provides free and confidential emotional support to people in suicidal crisis or emotional distress 24 hours a day, 7 days a week, across the United States. The Lifeline is comprised of a national network of over 200 local crisis centers, combining custom local care and resources with national standards and best practices. The new, shorter phone number makes it easier for people to remember and access mental health crisis services (though the previous 1 (800) 273-8255 number will continue to function indefinitely) and represents a continued commitment to those in crisis.

<u>Relationship and Sexual Violence Prevention Services (RSVP)</u> - (213) 740-9355(WELL) – 24/7 on call Free and confidential therapy services, workshops, and training for situations related to gender- and power-based harm (including sexual assault, intimate partner violence, and stalking).

Office for Equity, Equal Opportunity, and Title IX (EEO-TIX) - (213) 740-5086

Information about how to get help or help someone affected by harassment or discrimination, rights of protected classes, reporting options, and additional resources for students, faculty, staff, visitors, and applicants.

Reporting Incidents of Bias or Harassment - (213) 740-5086 or (213) 821-8298

Avenue to report incidents of bias, hate crimes, and microaggressions to the Office for Equity, Equal Opportunity, and Title for appropriate investigation, supportive measures, and response.

The Office of Student Accessibility Services (OSAS) - (213) 740-0776

OSAS ensures equal access for students with disabilities through providing academic accommodations and auxiliary aids in accordance with federal laws and university policy.

USC Campus Support and Intervention - (213) 740-0411

Assists students and families in resolving complex personal, financial, and academic issues adversely affecting their success as a student.

Diversity, Equity and Inclusion - (213) 740-2101

Information on events, programs and training, the Provost's Diversity and Inclusion Council, Diversity Liaisons for each academic school, chronology, participation, and various resources for students.

<u>USC Emergency</u> - UPC: (213) 740-4321, HSC: (323) 442-1000 – 24/7 on call

Emergency assistance and avenue to report a crime. Latest updates regarding safety, including ways in which instruction will be continued if an officially declared emergency makes travel to campus infeasible.

<u>USC Department of Public Safety</u> - UPC: (213) 740-6000, HSC: (323) 442-1200 – 24/7 on call Non-emergency assistance or information.

Page 13 of 14

This content is protected and may not be shared, uploaded, or distributed.

Office of the Ombuds - (213) 821-9556 (UPC) / (323-442-0382 (HSC)

A safe and confidential place to share your USC-related issues with a University Ombuds who will work with you to explore options or paths to manage your concern.

Occupational Therapy Faculty Practice - (323) 442-2850 or otfp@med.usc.edu

Confidential Lifestyle Redesign services for USC students to support health promoting habits and routines that enhance quality of life and academic performance.

