

ISM541: Data Warehousing in Enterprise Environments

Credit Hours: 3

Contact Hours: This is a 3-credit course, offered in accelerated format. This means that 16 weeks of material is covered in 8 weeks. The exact number of hours per week that you can expect to spend on each course will vary based upon the weekly coursework, as well as your study style and preferences. You should plan to spend 14-20 hours per week in each course reading material, interacting on the discussion boards, writing papers, completing projects, and doing research.

Course Description and Outcomes



Course Description:

This course provides students with an overview of data warehouses in an enterprise and how data is captured, analyzed, and translated into information to assist executives with streamlined information from disparate systems for effective decision making based on data for strategic and operational objectives.

Course Overview:

This course provides students with an in-depth study of the architecture of enterprise data warehouses customized to meet the unique business requirements of a variety of different vertical markets. To that end, a series of case studies are presented. All enterprise data warehouses are based on a set of dimension tables related to one or more fact tables, thus the considerations for dimension and fact table columns and rows is investigated.

The Kimball approach to creating a successful enterprise data warehouse, using the business intelligence (BI) system development lifecycle, is followed for all case studies. This approach, assembled from years of practical experience, is based on the data warehouse bus architecture and a matrix of business processes versus dimension tables. The importance of selecting a suitable fact table grain is presented. Important data warehouse issues such as slowly changing dimensions are covered in detail.

The course concludes with a high-level introduction to big data storage, business analytical processing, and best practices for handling big data.

Course Learning Outcomes:

1. Formulate strategic and effective management decisions that demonstrate a high level of knowledge of strategic planning, competitive intelligence, and ethical standards in a culturally diverse global marketplace.
2. Demonstrate written communication skills to express ideals, research findings, data analysis, and recommendations effectively and persuasively with all organizational stakeholders using a variety of communications skills and tools.

3. Develop comprehensive and cutting edge solutions to business problems by researching, analyzing, synthesizing and evaluating information using qualitative and quantitative analytical reasoning.
4. Direct the effective use of technology in an organization to achieve superior organizational and individual performance, competitive advantage, and operational effectiveness within the global marketplace.
5. Apply proven management theories and practices to resolve a wide range of organizational issues.
6. Advance and foster a culture of innovation that uses new approaches, theoretical frameworks, and technologies to drive the economic performance of organizations.

Participation & Attendance



Prompt and consistent attendance in your online courses is essential for your success at CSU-Global Campus. Failure to verify your attendance within the first 7 days of this course may result in your withdrawal. If for some reason you would like to drop a course, please contact your advisor.

Online classes have deadlines, assignments, and participation requirements just like on-campus classes. Budget your time carefully and keep an open line of communication with your instructor. If you are having technical problems, problems with your assignments, or other problems that are impeding your progress, let your instructor know as soon as possible.

Course Materials



Textbook Information is located in the CSU-Global Booklist on the Student Portal.

Course Schedule



Due Dates

The Academic Week at CSU-Global begins on Monday and ends the following Sunday.

- Discussion Boards: The original post must be completed by Thursday at 11:59 p.m. MT and Peer Responses posted by Sunday 11:59 p.m. MT. Late posts may not be awarded points.
- Critical Thinking: Assignments are due Sunday at 11:59 p.m. MT.

Week #	Readings	Assignments
1	<p>Required:</p> <ul style="list-style-type: none"> • Chapters 1 & 2 in <i>The Data Warehouse Toolkit</i> <p>Recommended:</p> <ul style="list-style-type: none"> • DWBI Concepts Team. (2014, September 15). [Blog post]. Top 5 challenges of data warehousing. Retrieved from http://dwbi.org/data-modelling/dw-design/11-top-5-challenges-of-data-warehousing 	<ul style="list-style-type: none"> • Discussion (25 points)

	<ul style="list-style-type: none"> Mead, K. (2009, September 26). [Blog post]. A layman's understanding of star schemas. Retrieved from http://www.orafaq.com/node/2286 	
2	<p>Required:</p> <ul style="list-style-type: none"> Chapters 3, 4, 5, & 6 in <i>The Data Warehouse Toolkit</i> <p>Recommended:</p> <ul style="list-style-type: none"> 1keydata.com. (2016). Fact table granularity. Retrieved from http://www.1keydata.com/datawarehousing/fact-table-granularity.html What are slowly changing dimensions? (2015, July 19). [Blog post]. Retrieved from http://www.learn-datamodeling.com/sl_dim.php#.UbUax5xkn5g 	<ul style="list-style-type: none"> Discussion (25 points)
3	<p>Required:</p> <ul style="list-style-type: none"> Chapters 7, 8, 9, & 10 in <i>The Data Warehouse Toolkit</i> <p>Recommended:</p> <ul style="list-style-type: none"> Microsoft. (2016). Recursive hierarchies (master data services). Retrieved from http://msdn.microsoft.com/en-us/library/ff487006 	<ul style="list-style-type: none"> Discussion (25 points) Critical Thinking (90 points) Portfolio Milestone (50 points)
4	<p>Required:</p> <ul style="list-style-type: none"> Chapters 11, 12, & 13 in <i>The Data Warehouse Toolkit</i> <p>Recommended:</p> <ul style="list-style-type: none"> Oakton Solution LLC. (2011). Factless fact tables. Retrieved from http://blog.oaktonsoftware.com/2011/09/factless-fact-tables.html 	<ul style="list-style-type: none"> Discussion (25 points) Critical Thinking (115 points)
5	<p>Required:</p> <ul style="list-style-type: none"> Chapters 14, 15 & 16 in <i>The Data Warehouse Toolkit</i> <p>Recommended:</p> <ul style="list-style-type: none"> Sukhwani, S., Garla, S., & Chakraborty, G. (2012). Analysis of clickstream data using SAS®. Retrieved from http://support.sas.com/resources/papers/proceedings12/100-2012.pdf 	<ul style="list-style-type: none"> Discussion (25 points) Critical Thinking (115 points) Portfolio Milestone (50 points)
6	<p>Required:</p> <ul style="list-style-type: none"> Chapters 17 & 18 in <i>The Data Warehouse Toolkit</i> <p>Recommended:</p> <ul style="list-style-type: none"> Haertzen, D. (2016, April 4.) Business intelligence and data warehousing require project management know how [Blog post]. Retrieved from http://infogoal.com/datawarehousing/project_management.htm 	<ul style="list-style-type: none"> Discussion (25 points) Critical Thinking (130 points)
7	<p>Required:</p> <ul style="list-style-type: none"> Chapters 19 & 20 in <i>The Data Warehouse Toolkit</i> <p>Recommended:</p> <ul style="list-style-type: none"> Javlin - Data Solutions. (2015). ETL (Extract-Transform-Load). Retrieved from http://www.dataintegration.info/etl 	<ul style="list-style-type: none"> Discussion (25 points)
8	<p>Required:</p> <ul style="list-style-type: none"> Chapter 21 of <i>The Data Warehouse Toolkit</i> <p>Recommended:</p> <ul style="list-style-type: none"> The Apache Software Foundation. (2016). MapReduce tutorial. Retrieved from 	<ul style="list-style-type: none"> Discussion (25 points) Portfolio (250 points)

Assignment Details



This course includes the following assignments/projects:

Module 1

PORTFOLIO PROJECT REMINDER (0 points)

Review the Portfolio Project description and rubric in the Module 8 folder. This project, due in Week 8, constitutes approximately one third of your grade for the course.

Begin thinking about the proposal you will write for the Portfolio Project now. Contact your instructor by email if you have any questions.

Module 3

CRITICAL THINKING ASSIGNMENT (90 points)

Option #1: Star Schema Modeling

Write a response addressing this question: Why is there no one-size-fits-all star schema model for all types of industries?

Your response should include:

- A discussion of the physical, structural variations that exist in the star schema's dimension and fact tables.
- A discussion of fact table grains.
- A description of the expected types of user analyses of the output data.

Your paper should be 2-4 pages in length, organized, and well-written in conformity with the [CSU-Global Guide to Writing & APA](#).

In support of your response, cite and integrate at least two credible outside sources, as well as your textbook, for a total of three or more sources. Place these in a reference page formatted in accordance with APA style. This reference page does not count toward the total page requirement. The CSU-Global Library is a great place to find these resources.

Examine the Critical Thinking grading rubric, found in the Week 3 folder, to see how you will be graded for this assignment.

Option #2: Star Schema Modeling

Write a justification memo to your Chief Information Officer (CIO) explaining why your organization needs a data warehouse.

Your response should include:

- A discussion of the differences between online transaction processing system (OLTP) databases and Business Intelligence (BI) systems in terms of the types of results they produce.
- The problems with using an OLTP database for actionable BI for strategic decision making.
- The relationship between OLTP data from production systems and historical data in a data warehouse.

Your memo should be 2-4 pages in length, organized, and well-written in conformity with the [CSU-Global Guide to Writing & APA](#).

In support of your response, cite and integrate at least two credible outside sources, as well as your textbook, for a total of three or more sources. Place these in a reference page formatted in accordance with APA style. This reference page does not count toward the total page requirement. The CSU-Global Library is a great place to find these resources.

Examine the Critical Thinking grading rubric, found in the Week 3 folder, to see how you will be graded for this assignment.

PORTFOLIO PROJECT MILESTONE (50 points)

Submit a list of the data elements to be presented in your Portfolio Project data model. See the description of the Portfolio Project found in the Week 8 folder for more information.

Begin thinking about the outline that you will write for your Portfolio Project; the outline will be due in Week 5. Contact your instructor by email if you have any questions.

Module 4

CRITICAL THINKING ASSIGNMENT (115 points)

Option #1: Case Study 1: Data Warehouse Application in an Industry

Perform research to locate a data warehouse system that has been applied to the telecommunication, transportation, or education industries. Be sure to select one that has been developed and deployed to successfully produce actionable business intelligence.

Write a case study paper that describes:

- The architecture of this data warehouse.
- The source input systems.
- ETL processing that is required.
- The types of output reporting.
- The types of users.

Be sure to also discuss the unique challenges that were encountered for this data warehouse project and how they were overcome. Describe the success/outcome of this data warehouse application and any future expansion plans for the system.

Your paper should be 3-5 pages in length, organized, and well-written in conformity with the [CSU-Global Guide to Writing & APA](#).

In support of your response, cite and integrate at least two credible outside sources, as well as your textbook, for a total of three or more sources. Place these in a reference page formatted in accordance with APA style. This reference page does not count toward the total page requirement. The CSU-Global Library is a great place to find these resources.

Examine the Critical Thinking grading rubric, found in the Week 4 folder, to see how you will be graded for this assignment.

Option #2: Case Study 1: Data Warehouse Design for Your Workplace

Propose a data warehouse architecture for your workplace. Make sure that it includes the following elements:

- A drawing of the complete star scheme of dimension and fact tables.
- A complete description about how the attributes of all tables would meet the requirements for your organization's use of actionable Business Intelligence (BI) data.
- A detailed description of the ETL processing to include source production systems from which data will be extracted, the types of data cleansing activities that would be required due to overlapping data sources

and content, and the proposed data warehouse loading plan to provide adequately refreshed data for effective decision making by your organization's executive and senior level business analysts.

- A detailed description of the challenges that might occur due to slowly changing dimensions and your strategy to handle them.

Your paper should be 3-5 pages in length, organized, and well-written in conformity with the [CSU-Global Guide to Writing & APA](#).

In support of your response, cite and integrate at least two credible outside sources, as well as your textbook, for a total of three or more sources. Place these in a reference page formatted in accordance with APA style. This reference page does not count toward the total page requirement. The CSU-Global Library is a great place to find these resources.

Examine the Critical Thinking grading rubric, found in the Week 4 folder, to see how you will be graded for this assignment.

Module 5

CRITICAL THINKING ASSIGNMENT (115 points)

Option #1: Case Study 2: Data Warehouse Application in an Industry

Perform research to locate a data warehouse system that has been applied to the health care, electronic commerce, or insurance industries and has been successfully developed and deployed to successfully produce actionable business intelligence.

Write a case study paper that describes:

- The architecture of this data warehouse.
- The source input systems.
- ETL processing that is required.
- The types of output reporting.
- The types of users.

Be sure to also discuss the unique challenges that were encountered for this data warehouse project and how they were overcome. Describe the success/outcome of this data warehouse application and any future expansion plans for the system.

Your paper should be 3-5 pages in length, organized, and well-written in conformity with the [CSU-Global Guide to Writing & APA](#).

In support of your response, cite and integrate at least two credible outside sources, as well as your textbook, for a total of three or more sources. Place these in a reference page formatted in accordance with APA style. This reference page does not count toward the total page requirement. The CSU-Global Library is a great place to find these resources.

Examine the Critical Thinking grading rubric, found in the Week 5 folder, to see how you will be graded for this assignment.

Option #2: Case Study 2: Data Warehouse Application in an Industry

Propose a data warehouse architecture for Amazon.com, including:

- A drawing of the complete star scheme of dimension and fact tables.
- A complete description about how the attributes of all tables would meet the requirements for your organization's use of actionable Business Intelligence (BI) data.
- A detailed description of the ETL processing to include source production systems from which data will be extracted, the types of data cleansing activities that would be required due to overlapping data sources

and content, and the proposed data warehouse loading plan to provide adequately refreshed data for effective decision making by Amazon.com.

- A detailed description of the challenges that might occur due to slowly changing dimensions and the strategy to handle them.

Your paper should be 3-5 pages in length, organized, and well-written in conformity with the [CSU-Global Guide to Writing & APA](#).

In support of your response, cite and integrate at least two credible outside sources, as well as your textbook, for a total of three or more sources. Place these in a reference page formatted in accordance with APA style. This reference page does not count toward the total page requirement. The CSU-Global Library is a great place to find these resources.

Examine the Critical Thinking grading rubric, found in the Week 5 folder, to see how you will be graded for this assignment.

PORTFOLIO PROJECT MILESTONE (50 points)

Submit an outline of your Portfolio Project data model. Describe the challenges you have encountered, and may expect to encounter, in collecting, integrating, and transforming the data in preparation for loading to the data warehouse. Submit your outline to the Week 5 folder.

Module 6

CRITICAL THINKING ASSIGNMENT (130 points)

Option #1: Importance of a Data Warehouse System in the Entertainment Industry

Imagine that you are the CIO of a large organization in the entertainment industry. Your organization has been suffering an alarming decline in ticket sales profits for the past several years.

In your role as CIO, you realize the importance of a data warehouse system and how this can help your organization regain profits from ticket sales. But your boss, the CEO, and the Board of Directors do not understand why a new database system is needed at this time. Resources are tight and their focus is on the immediate bottom line. They point to the company's mainframe system that is already handling ticket sales transactions and ask, "Why can't you just assign some COBOL programmers or something? Make it a quick project, figure out how to increase sales and get profits back up using the big computers we already have!"

Why not do what they suggest? Develop a proposal to your CEO and Board of Directors explaining:

- The importance of a data warehouse system, introducing it in terms they can understand.
- How a data warehouse system is not a waste of time and money.
- Why the present system is inadequate.
- How a data warehouse system will increase profits from ticket sales by providing strategic information not currently available.

Your paper should be 5-7 pages in length, organized, and well-written in conformity with the [CSU-Global Guide to Writing & APA](#).

In support of your responses, cite and integrate at least two credible outside sources, as well as your textbook, for a total of three or more sources. Place these in a reference page formatted in accordance with APA style. This reference page does not count toward the total page requirement. The CSU-Global Library is a great place to find these resources.

Examine the Critical Thinking grading rubric, found in the Week 6 folder, to see how you will be graded for this assignment.

Option #2: Proposed Design for a Data Warehouse System in the Entertainment Industry

Imagine that you are the Chief Technology Officer (CTO) of a large organization in the entertainment industry. Your organization has been suffering an alarming decline in ticket sales profits for the past several years.

The Board of Directors and CIO have tasked you to develop a data warehouse to determine why sales have slumped and how best to increase them in the future.

You need to identify the dimension tables and fact tables that will assist business analysts perform online analytical processing (OLAP) and data mining analyses. Note any slowly changing dimension tables, why data for these tables may occur, and how you will handle it. You also need to determine the best user front end tools to use, and possible statistical models to use to analyze the data.

Your paper should be 5-7 pages in length, organized, and well-written in conformity with the [CSU-Global Guide to Writing & APA](#).

In support of your responses, cite and integrate at least two credible outside sources, as well as your textbook, for a total of three or more sources. Place these in a reference page formatted in accordance with APA style. This reference page does not count toward the total page requirement. The CSU-Global Library is a great place to find these resources.

Examine the Critical Thinking grading rubric, found in the Week 6 folder, to see how you will be graded for this assignment.

Module 8

PORTFOLIO PROJECT (250 points)

Option #1: K-12 Student Data Warehouse

You are a senior data warehouse architect for LECAL Corporation. Your company has been contracted by U.S. Department of Education to develop a data warehouse to track student performance on a series of planned nationwide tests as part of a new STEM (Science, Technology, Engineering, and Math) initiative.

Specifically, new tests will be created for algebra, geometry, calculus, chemistry, and physics.

The data warehouse will produce annual average test scores for graduating seniors for each of these tests. All graduating seniors, regardless of the type of K-12 school they attended (public, private, parochial, home schooled, for example) are to be included in the data warehouse.

The system must include all relevant courses in the students' course history and the final course grades. Prerequisite courses and grades are to be included as well. Course names and numbers are non-standard among the nation's school districts and educational institutions and tend to change each year-to-year. Also, course grades are non-standard.

The reporting system for this data warehouse will produce tabular statistics and graphical color-coded charts of year-to-year progress by student course preparation, race, gender, special education status, ESOL (English Speakers of Other Languages), and family income level. Note that these student demographics can change year-to-year and even within a school year.

Your Portfolio Project is to prepare a proposal for a star schema data model to meet the requirements of this system. Determine what dimension and fact tables will be required, what data they will contain, what data is to be collected, and how the source data is to be uniformly integrated into the data warehouse.

Additionally:

- Note any slowly changing dimension tables, why data for these tables may occur, and how you will handle it.
- Describe the opportunities to capture, store, and analyze big data for this application and how this data might be processed with Hadoop Map/Reduce or alternate techniques using a NoSQL database.
- Describe the ETL challenges with data collection, transformation, and loading that may occur. Describe the challenges to keep your data model flexible enough to handle the vast number of different evolving course.

Your proposal should be 15-20 pages in length, organized, and well written in conformity with the [CSU-Global Guide to Writing & APA](#).

Option #2: GM Vehicle Sales Forecasting

You are a senior data warehouse architect for MARDER Corporation. Your company has been contracted by the Eastern Region of General Motors to develop a data warehouse of sales of GM vehicles over the last 5 years to determine how best to increase sales for the next 5 years in this region.

The actionable business intelligence (BI) outputs of the system are intended to help the manufacturer predict what makes, models, and features historically have been most desired in the region in order to focus on these for the future.

Specifically, sales of Cadillacs, GMCs, and Buicks are to be gathered, stored, and analyzed.

The goal of this project is to determine which types of vehicles and which features are most desired by consumers.

The data warehouse also needs to keep a record of repairs to know which vehicle components fail most often, since this might influence buyers' decisions to stay with GM vehicles or switch to a competitor.

Another requirement is to record which previous vehicle type each customer in a sale owned, and how long it has been since the sale, in an attempt to monitor customer churn.

Note any slowly changing dimension tables, why data for these tables may occur, and how you will handle it.

The reporting system for this data warehouse will produce tabular statistics and graphical color-coded charts of year-to-year sales by vehicle make and model, as well as all possible feature packages for each model.

Your Portfolio Project is to prepare a proposal for a star schema data model to meet the requirements of this system.

Additionally, you will:

- Determine what dimension and fact tables will be required, what data they will contain, what data is to be collected, and how the source data is to be uniformly integrated into the data warehouse.
- Describe the challenges with data collection, transformation, and loading that may occur.
- Describe the challenges to keep your data model flexible enough to handle the vast number of different evolving vehicle models and features.
- Describe the opportunities to capture, store, and analyze big data for this application and how this data might be processed with Hadoop Map/Reduce or alternate techniques using a NoSQL database.

Your proposal should be 15-20 pages in length, organized, and well written in conformity with the [CSU-Global Guide to Writing & APA](#).

Course Policies



Course Grading

20% Discussion Participation
45% Critical Thinking Assignments
35% Portfolio Project

Grading Scale and Policies

A	95.0 – 100
A-	90.0 – 94.9
B+	86.7 – 89.9
B	83.3 – 86.6
B-	80.0 – 83.2
C+	75.0 – 79.9
C	70.0 – 74.9
D	60.0 – 69.9
F	59.9 or below

In-Classroom Policies

For information on late work and incomplete grade policies, please refer to our [In-Classroom Student Policies and Guidelines](#) or the Academic Catalog for comprehensive documentation of CSU-Global institutional policies.

Academic Integrity

Students must assume responsibility for maintaining honesty in all work submitted for credit and in any other work designated by the instructor of the course. Academic dishonesty includes cheating, fabrication, facilitating academic dishonesty, plagiarism, reusing /re-purposing your own work (see *CSU-Global Guide to Writing and APA Requirements* for percentage of repurposed work that can be used in an assignment), unauthorized possession of academic materials, and unauthorized collaboration. The CSU-Global Library provides information on how students can avoid plagiarism by understanding what it is and how to use the Library and Internet resources.

Citing Sources with APA Style

All students are expected to follow the *CSU-Global Guide to Writing and APA Requirements* when citing in APA (based on the APA Style Manual, 6th edition) for all assignments. For details on CSU-Global APA style, please review the APA resources within the CSU-Global Library under the “APA Guide & Resources” link. A link to this document should also be provided within most assignment descriptions in your course.

Disability Services Statement

CSU-Global is committed to providing reasonable accommodations for all persons with disabilities. Any student with a documented disability requesting academic accommodations should contact the Disability Resource Coordinator at 720-279-0650 and/or email ada@CSUGlobal.edu for additional information to coordinate reasonable accommodations for students with documented disabilities.

Netiquette

Respect the diversity of opinions among the instructor and classmates and engage with them in a courteous, respectful, and professional manner. All posts and classroom communication must be conducted in accordance with the student code of conduct. Think before you push the Send button. Did you say just what you meant? How will the person on the other end read the words?

Maintain an environment free of harassment, stalking, threats, abuse, insults or humiliation toward the instructor and classmates. This includes, but is not limited to, demeaning written or oral comments of an ethnic, religious, age, disability, sexist (or sexual orientation), or racist nature; and the unwanted sexual advances or intimidations by email, or on discussion boards and other postings within or connected to the online classroom.

If you have concerns about something that has been said, please let your instructor know.