

Note:

Course content may be changed, term to term, without notice. The information below is provided as a guide for course selection and is not binding in any form, and should not be used to purchase course materials.

COURSE SYLLABUS

CSIS 657

STATISTICAL ANALYSIS AND DATA MINING

COURSE DESCRIPTION

This course provides an in-depth study of the field of statistical analysis and data mining as it relates to real-world applications. It explores the complexities of data mining algorithms, software tools, and techniques employed in modern analytics and massive data sets. The selection, application, and evaluation of statistical approaches are examined in the context of data mining.

RATIONALE

Technology's role in society continues to expand in application and influence. The data generated through this digital frontier is growing exponentially, creating new challenges as well as exciting opportunities. The ability to sift through the vast amount of information requires a skillset that is part engineer and part artist. Finding data, making appropriate associations between data, constructing ways to communicate relationships of data, and applying business intelligence and analytics are fundamental to business in the digital age. Never before has there been so much information available for companies to strategize with, analyze, consume, and even market. The effective exploitation of data mining and predictive analytics technologies will be an invaluable skillset for the company looking for opportunities to stay competitive while maintaining lean and efficient organizations.

I. PREREQUISITE

For information regarding prerequisites for this course, please refer to the [Academic Course Catalog](#).

II. REQUIRED RESOURCE PURCHASE

Click on the following link to view the required resource(s) for the term in which you are registered: <http://bookstore.mbsdirect.net/liberty.htm>

III. RECOMMENDED RESOURCE

American Psychological Association. *Publication manual of the American Psychological Association* (Current ed.). Washington, DC: Author.

IV. ADDITIONAL MATERIALS FOR LEARNING

- A. Computer with basic audio/video output equipment
- B. Internet access (broadband recommended)
- C. Blackboard [recommended browsers](#)

- D. Microsoft Office
- E. *CRISP-DM 1.0* (available as a free PDF in Blackboard)

V. MEASURABLE LEARNING OUTCOMES

Upon successful completion of this course, the student will be able to:

- A. Define the data mining process in relation to Christian principles and ethics. (PLO: 1)
- B. Compare data mining algorithms, tools, and techniques. (PLO: 3)
- C. Analyze data and text mining case studies. (PLO: 3)
- D. Select appropriate statistical data mining models for various scenarios. (PLO: 2)
- E. Apply data and text mining algorithms to real-world applications. (PLO: 2)

VI. COURSE REQUIREMENTS AND ASSIGNMENTS

- A. Textbook readings and lecture presentations
- B. Course Requirements Checklist

After reading the Course Syllabus and [Student Expectations](#), the student will complete the related checklist found in Module/Week 1.

- C. Discussion Board Forums (4)

Discussion boards are collaborative learning experiences. Therefore, the student is required to create a thread in response to the provided prompt for each forum. Each thread must be 400–500 words and demonstrate course-related knowledge. In addition to the thread, the student is required to reply to the threads of at least 2 classmates. Each reply must be 200–300 words. Each thread and reply must include at least 2 scholarly sources other than the Reading & Study material and 2 contextually appropriate scriptural references. Each thread and reply must follow current APA format.

- D. Tutorials (7)

The student will complete practical exercises (tutorials) designed to (1) create experience with the software used in the course, (2) provide real-world examples of problems facing a variety of business sectors, (3) build understanding of how to methodically approach solving a data mining hypothesis, and (4) foster a greater understanding of the potential value of corporate data and the impact of big data.

- E. Briefs (4)

The student will research and write a 3–5-page brief that focuses on a topic provided by the instructor. The paper must be in current APA style and contain at least 3 scholarly sources other than the course textbook.

F. Project

The student will incorporate all aspects of the course into an integrated holistic project on data mining. This effort will follow the scientific method outlined in the course textbook and be completed in the following phases. Prior to beginning the project, the student will submit a 2-page summary of the issue being addressed.

Phase I

The student will submit a paper of at least 5 pages that outlines the problem, identifies the sources of data, and explains the student's initial hypothesis. The paper must be in current APA style and include at least 8 scholarly sources other than the course textbook.

Phase II

The student will submit a 5–8-page paper that outlines the processes implemented for collecting and preparing the data for examination. This phase must be a natural continuation of Phase I, describing in detail how the student prepared the data sets used as well as the analytics selected and reasoning for those selections. Identification of key variables and significant descriptive statistics must be included. The paper must be in current APA style and include at least 8 scholarly sources other than the course textbook.

Phase III

The student will submit a paper of at least 5 pages that builds off the work conducted in Phases I and II. The submitted document will be comprehensive and include the findings, analysis, and next steps recommended as a result of the information generated by the data mining model(s) utilized. The student's analysis of the data must be presented in a manner that would be appropriate for repeatability by a fellow data miner. The paper must be in current APA style and include at least 5 scholarly sources other than the course textbook.

Phase IV

The student will create a persuasive slide presentation for the purpose of communicating the project (Phases I–III) and the recommendations coming out of this effort. The target audience for this information must be consistent with C-level management and other organization decision makers. Details might include cost, schedule, impact, implementation planning, change management issues, risks, benefits, etc. This presentation must include at least 15 slides. The student will also finalize the report coming from the work from Phases I–III into a cohesive product. The report must be at least 15 pages. The paper must be in current APA style and include at least 10 scholarly sources other than the course textbook.

VII. COURSE GRADING AND POLICIES

A. Points

Course Requirements Checklist	10
Discussion Board Forums (4 at 30 pts ea)	120
Tutorials (7 at 40 pts ea)	280
Briefs (4 at 45 pts ea)	180
Project	
Phase I	100
Phase II	150
Phase III	100
Phase IV	70
Total	1010

B. Scale

A = 940–1010 A- = 920–939 B+ = 900–919 B = 860–899 B- = 840–859
 C+ = 820–839 C = 780–819 C- = 760–779 F = 0–759

C. Disability Assistance

Students with a documented disability may contact Liberty University Online’s Office of Disability Academic Support (ODAS) at LUOODAS@liberty.edu to make arrangements for academic accommodations. Further information can be found at www.liberty.edu/disabilitysupport.

COURSE SCHEDULE

CSIS 657

MODULE/ WEEK	READING & STUDY	ASSIGNMENTS	POINTS
1	<i>CRISP-DM 1.0</i> : pp. 1–12 Joyner et al.: chs. 3–5 Nisbet et al.: chs. 2–3 1 presentation	Course Requirements Checklist Class Introductions Tutorial – A 367–375 Brief 1	10 0 40 45
2	Joyner et al.: ch. 8 Nisbet et al.: chs. 4–5 1 presentation	DB Forum 1 Tutorial – S DVD Files Project – Phase I	30 40 100
3	<i>CRISP-DM 1.0</i> : pp. 13–18 Joyner et al.: ch. 9 Nisbet et al.: chs. 6–7 1 presentation	DB Forum 2 Tutorial – N 681 (691)–703	30 40
4	<i>CRISP-DM 1.0</i> : pp. 19–24 Joyner et al.: ch. 11 Nisbet et al.: chs. 8–9 1 presentation	Tutorial – E 459–470 Brief 2	40 45
5	<i>CRISP-DM 1.0</i> : pp. 25–31 Nisbet et al.: ch. 10 1 presentation	DB Forum 3 Tutorial – G 481–512 Project – Phase II	30 40 150
6	<i>CRISP-DM 1.0</i> : pp. 32–34 Joyner et al.: ch. 15 Nisbet et al.: ch. 11 1 presentation	Tutorial – CC DVD Files Brief 3	40 45
7	<i>CRISP-DM 1.0</i> : pp. 35–39 Joyner et al.: ch. 16 Nisbet et al.: ch. 12 1 presentation	DB Forum 4 Tutorial – T DVD Files Project – Phase III	30 40 100
8	Nisbet et al.: ch. 13 1 presentation	Brief 4 Project – Phase IV	45 70
TOTAL			1010

DB = Discussion Board

NOTE: Each course module/week begins on Monday morning at 12:00 a.m. (ET) and ends on Sunday night at 11:59 p.m. (ET). The final module/week ends at 11:59 p.m. (ET) on **Friday**.