DATE OF LAST REVIEW: 02/2013

CIP CODE: 15.1201

SEMESTER: Departmental Syllabus

COURSE TITLE: Data Cleansing

COURSE NUMBER: CIST-0204

CREDIT HOURS: 3

INSTRUCTOR: Departmental Syllabus

OFFICE LOCATION: Departmental Syllabus

OFFICE HOURS: Departmental Syllabus

TELEPHONE: 913-334-1100

PREREQUISITE(S): CIST-0154 (Principles of Data Warehousing)

REQUIRED TEXT(S):
Please see bookstore for current textbook.

COURSE DESCRIPTION:
This course will provide the student with an understanding of a critical dimension of data performance, data quality. The student will gain an understanding of how to cope with dirty data and the causes of it.

METHOD OF INSTRUCTION:
A variety of instructional methods may be used depending on content area. These include but are not limited to: lecture, multimedia, cooperative/collaborative learning, labs and demonstrations, projects and presentations, speeches, debates, and panels, conferencing, performances, and learning experiences outside the classroom. Methodology will be selected to best meet student needs.

COURSE OUTLINE:
Course content may vary, but will generally include the following:

I. How Dirty Data Gets In
   A. Legacy Operation
   B. Improper Integration
   C. Data Warehouse Aging
   D. Changing User Requirements

II. Cleansing Dirty Data
   A. Cleansing the Legacy Environment
   B. Cleansing at the Point of Integration
   C. Cleansing after loading
   D. Cleansing Data Over Time
III. Auditing Data
   A. Single Row Sequential Audit
   B. Multi Row Audit

EXPECTED LEARNER OUTCOMES:
A. Upon completion of the course the student will be able to demonstrate knowledge of how dirty data gets into a data warehouse.
B. Upon completion of the course the student will be able to demonstrate knowledge of how to cleanse dirty data.
C. Upon completion of the course the student will be able to demonstrate knowledge of how to audit data.

COURSE COMPETENCIES:
Upon completion of the course the student will be able to demonstrate knowledge of how dirty data gets into a data warehouse.

1. Upon completion of the course the student will be able to define legacy operation.
2. Upon completion of the course the student will be able to identify methods by which dirty data enters a warehouse through legacy.
3. Upon completion of the course the student will be able to define moment of integration as it relates to a data warehouse.
4. Upon completion of the course the student will be able to identify methods by which dirty data enters a warehouse through integration.
5. Upon completion of the course the student will be able to define aging as it relates to a data warehouse.
6. Upon completion of the course the student will be able to identify methods by which dirty data enters a warehouse through warehouse aging.
7. Upon completion of the course the student will be able to define changing user requirements as it relates to a data warehouse.
8. Upon completion of the course the student will be able to identify methods by which dirty data enters a warehouse through changing user requirements.

Upon completion of the course the student will be able to demonstrate knowledge of how to cleanse dirty data.

9. Upon completion of the course the student will be able to identify Legacy Environment as it relates to a data warehouse.
10. Upon completion of the course the student will be able to identify at least six reasons why cleansing a legacy environment is difficult.
11. Upon completion of the course the student will be able to interpret data and cleanse a Legacy Environment.
12. Upon completion of the course the student will be able to identify Point of Integration as it relates to a data warehouse.
13. Upon completion of the course the student will be able to interpret data and cleanse data at the Point of Integration by reshaping it to fit a data model.
14. Upon completion of the course the student will be able to interpret data and cleanse data at the Point of Integration by scrubbing.
15. Upon completion of the course the student will be able to identify at least twelve methods by which data may be cleansed at the Point of Integration by scrubbing.
16. Upon completion of the course the student will be able to identify cleansing after loading as it relates to a data warehouse.
17. Upon completion of the course the student will be able to identify the two steps by which data may be cleansed after loading.
18. Upon completion of the course the student will be able to interpret data and cleanse data after loading.
19. Upon completion of the course the student will be able to identify cleansing data over time as it relates to a data warehouse.
20. Upon completion of the course the student will be able to interpret data and cleanse data over time.

**Upon completion of the course the student will be able to demonstrate knowledge of how to audit data.**

21. Upon completion of the course the student will be able to define audit data as it relates to a data warehouse.
22. Upon completion of the course the student will be able to identify the elements of a single row audit.
23. Upon completion of the course the student will be able to identify the elements of a multi row audit.
24. Upon completion of the course the student will be able to interpret data and perform a single row audit.
25. Upon completion of the course the student will be able to interpret data and perform a multi row audit.

**ASSESSMENT OF LEARNER OUTCOMES:**

Assessment methods may include, but are not limited to, the following: Homework, Assignments, Quizzes, Class Participation, Chapter Tests, and Final Exam. The grading scale and the process for calculating the course grades are to be determined by the individual instructors. This information will be included in each instructor’s syllabus.

**SPECIAL NOTES:**

This syllabus is subject to change at the discretion of the instructor. Material included is intended to provide an outline of the course and rules that the instructor will adhere to in evaluating the student’s progress. However, this syllabus is not intended to be a legal contract. Questions regarding the syllabus are welcome any time.

Kansas City Kansas Community College is committed to an appreciation of diversity with respect for the differences among the diverse groups comprising our students, faculty, and staff that is free of bigotry and discrimination. Kansas City Kansas Community College is committed to providing a multicultural education and environment that reflects and respects diversity and that seeks to increase understanding.

Kansas City Kansas Community College offers equal educational opportunity to all students as well as serving as an equal opportunity employer for all personnel. Various laws, including Title IX of the Educational Amendments of 1972, require the college’s policy on non-discrimination be administered without regard to race, color, age, sex, religion, national origin, physical handicap, or veteran status and that such policy be made known.

Kansas City Kansas Community College complies with the Americans with Disabilities Act. If you need accommodations due to a documented disability, please contact the Director of Academic Resources, in Rm. 3354 or call at: 288-7670 V/TDD.