College of Allied Health Sciences

Course Title
Introduction to Statistics, Health Data Analytics, & Business Intelligence

Course Number
IS7033

Instructor
David Rapien

Term Spring 2016
Table of Contents

Faculty & Staff.............................................................................................................................................. 1

Syllabus.......................................................................................................................................................... 2
  Program Name .......................................................................................................................................... 2
  Course Title: ........................................................................................................................................... 2
  Course Description.................................................................................................................................. 2
  Textbook(s), Readings and Assignments: ................................................................................................. 2
  Course Objectives: .................................................................................................................................. 2
  Learning Activities ................................................................................................................................... 3
  Quizzes ..................................................................................................................................................... 3
  Grading Policy ......................................................................................................................................... 3
  Grading Scale .......................................................................................................................................... 4
  Course Policies ....................................................................................................................................... 4

Course Schedule ......................................................................................................................................... 7
  Module 1 .................................................................................................................................................. 8
  Module 2 .................................................................................................................................................. 10
  Module 3 .................................................................................................................................................. 12
  Module 4 .................................................................................................................................................. 15
  Module 5 .................................................................................................................................................. 17
  Module 6 .................................................................................................................................................. 19
  Module 7 .................................................................................................................................................. 21
Faculty & Staff

Online Material Voice/Designer: David Rapien
Online Course Moderator: David Rapien
Syllabus

Program Name:  Masters in Heath Informatics

Course Title:  Introduction to Statistics, Health Data Analytics, & Business Intelligence

Course Number:  IS7033  Semester:  Spring (First Half)

Course Description  This is a two-part course. The first part introduces students to key introductory statistical approaches, applying those approaches to health and healthcare data. Qualitative and quantitative analytic methods, evaluation and research design are introduced. In the second part, students learn best practices of strategically managing large amounts of healthcare data. Data warehousing principles and concepts are taught. Students design a data warehouse relevant to a health or healthcare scenario and learn online analytic processing and data mining. Topics such as enterprise data architecture, data integration and data management are introduced.

Prerequisites:  IS 7031 - Database Modeling and Design

Course Credits:  3

Textbook(s), Readings and Assignments:
Required:

Statistics, Data Analytics and Decision Modeling 5th Edition by James Evans


Additional Resources:
Microsoft Excel 2013

Toolbox:
All materials contained in the course toolbox are supplementary to requirements; designed to aid the student, but not mandatory.

Course Objectives:
Students who successfully complete this course will be able to:

1.  Apply introductory statistical methods;
2. Apply evidence based knowledge systems for literature searches;
3. Develop a research question on a health informatics topic;
4. Describe approaches for effectively managing data as a major resource for health and healthcare organizations;
5. Design a data warehouse for a health or healthcare scenario;
6. Apply data and dimensional modeling, and extract, transform, load (ETL).

Learning Activities
A variety of learning activities are designed to support the course objectives, facilitate different learning styles, and build a community of learners. Learning activities for the modules include the following:

1. Reading textbook assignments
2. Reading other articles as assigned
3. Viewing and listening to PowerPoint lectures
4. Participating in discussions by responding to assigned topics/questions and replying to comments posted by faculty or other students
5. Participating in Toolbox activities (optional)
6. Completing writing assignments by individualized exploration of topics
7. Completing Weekly Quizzes

Quizzes
Instructions:

<table>
<thead>
<tr>
<th>Assessment</th>
<th>Open in Blackboard</th>
<th>Closes in Blackboard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Module 1 Quiz</td>
<td>1/14/16 @ 5:00pm</td>
<td>1/17/16 @ 11:59pm</td>
</tr>
<tr>
<td>Module 2 Quiz</td>
<td>1/21/16 @ 5:00pm</td>
<td>1/24/16 @ 11:59pm</td>
</tr>
<tr>
<td>Module 3 Quiz</td>
<td>1/28/16 @ 5:00pm</td>
<td>1/31/16 @ 11:59pm</td>
</tr>
<tr>
<td>Module 4 Quiz</td>
<td>2/04/16 @ 5:00pm</td>
<td>2/07/16 @ 11:59pm</td>
</tr>
<tr>
<td>Module 5 Quiz</td>
<td>2/11/16 @ 5:00pm</td>
<td>2/14/16 @ 11:59pm</td>
</tr>
<tr>
<td>Module 6 Quiz</td>
<td>2/18/16 @ 5:00pm</td>
<td>2/21/16 @ 11:59pm</td>
</tr>
<tr>
<td>Module 7 Quiz</td>
<td>2/25/16 @ 5:00pm</td>
<td>2/28/16 @ 11:59pm</td>
</tr>
</tbody>
</table>

Grading Policy
Method for Calculation of Course Grade

<table>
<thead>
<tr>
<th>Assignment</th>
<th>Module One</th>
<th>Module Two</th>
<th>Module Three</th>
<th>Module Four</th>
<th>Module Five</th>
<th>Module Six</th>
<th>Module Seven</th>
<th>Total Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participation in Discussion Board</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>30</td>
<td>30</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Project</td>
<td>10</td>
<td>30</td>
<td></td>
<td>30</td>
<td>30</td>
<td>100</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quizzes</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>70</td>
</tr>
<tr>
<td>Total Points</td>
<td>20</td>
<td>20</td>
<td>50</td>
<td>10</td>
<td>50</td>
<td>10</td>
<td>40</td>
<td>200</td>
</tr>
</tbody>
</table>
Grading Scale
- 94% and above = A
- 90% = A-
- 87% = B+
- 84% = B
- 80% = B-
- 77% = C+
- 74% = C
- 70% = C-
- 60% = D
- Below 60% = F

Grade Center: All grades will be maintained in Blackboard’s online Grade Center. Students are responsible to track their progress by referring to the online grade book. Email me with any questions on this.

Course Policies

Participation Policies: Students are expected to actively participate in class and in the Blackboard learning environment and to complete all assignments in a timely manner. Infrequent and inconsistent participation and work completion will negatively influence the benefits that may be obtained from the course as well as lead to a lower grade.

Make-Up Policy: Assignments are due by 11:59 PM, Sunday of the week they are due. All work will generally be graded and returned within 7 calendar days. If you have a SERIOUS problem that can be documented/verified and that keeps you from participating on time, please contact me immediately. I will determine if the seriousness of your problem warrants an exception to the late assignment rule. If you are not passing the class at any point due to missing work, you might be asked to drop the class.

You are responsible for timely assignment submission. Should your personal computer system or network go down, you must still turn in your work in a timely manner. Don’t wait until the last minute. Plan ahead by seeking alternative means for submitting your work before you need to. Local libraries and the University of Cincinnati campuses can serve as alternative resources. Not having access to the required software on your home or work computer is NOT a legitimate excuse for turning in assignments late.

Academic Integrity Policy: The University Rules, including the Student Code of Conduct, and other policies of the department, college, and university related to academic integrity will be enforced. Any violation of these regulations, including acts of plagiarism, cheating, or falsifying field work will be dealt with according to the severity of the misconduct. Dishonesty in any form may result in a failing grade in a course and/or suspension or dismissal from a program (e.g., graduate or undergraduate).
Electronic Communication / Email Policy: Students are required to use a University of Cincinnati email address for all program activity. The primary reasons for the new policy relate to issues of confidentiality, security and the receipt of information from the University of Cincinnati and the Educational Leadership Program.

The University of Cincinnati is now sending many official notices, announcements and important information to students via email. For example, student bills are now sent by email and are no longer mailed through the U.S. Postal Service. To ensure this type of confidential information is sent to the correct individual, items are only sent to student UC email accounts. A UC email address is clearly identified with the student’s name and only a student can register for her/his own email account. Therefore, there is a high level of confidence by the University the student will receive the email.

I can be reached via email and will try respond to all emails within 48 hours. If something is urgent, you can call me at the numbers listed.

NETIQUETTE

1--- Be Friendly, Positive and Self--- Reflective

When people cannot see you, and also do not know you, feelings can be hurt if you are not careful in how you express yourself. The old saying, think before you speak is important here. Think before you write. One word of advice is, do not respond when you feel angry. Wait. Write it down somewhere and come back to it. When you do, you may find that you no longer feel the same way as you did when you wrote it, because you have had time to reflect about the situation. Last, if you still feel the need to be heard, then edit before you post, and write it in terms that are easily embraced. This is also true when you feel a critique is necessary; say it in a positive tone. Reread what you have written to be sure it is positive.

2--- Use Proper Language and Titles

Do not use slang or even profane words in an online education environment, even if they are words you consider, "not so bad," as they will sound offensive to the reader. Do not refer to your professor as "Doc" or by his or her first name, unless it is acceptable with him or her to do so. Also, do not use caps lock when typing. It will insinuate yelling. That would hurt someone's feelings and possibly give him (or her) the wrong impression of you.

3--- Use Effective Communication

Say what you mean to say. This takes practice and thoughtful writing. Try to speak and write clearly at all times. Again, reread before you respond. Define and restate your words when necessary. Correct a misunderstanding right away. Chances are, if one
person felt a certain way about what you said, another may have as well. Likewise, be mindful of chosen words and joking. Let's say for example, I write, "get out!" This slang term can be interpreted in several ways, either positively or negatively.

4--- Professionalism

Leave the characters like smiley faces, and instant message abbreviations out. Your friends may like it, but chances are your professor will not. Save it for personal conversations or definitely ask for permission before using them. They may be interpreted as childish or too casual for the online education environment. Last, always say please and thank you.

5--- Ask for Clarification

If you are unsure of what was said, or the instructor's directive, or are trying to interpret a person's expressions, then ask again. Do not sit in silence either misunderstanding or feeling offended. Do not interrupt though; wait until there is a break in the conversation, or until the open interaction occurs. Your instructor will appreciate your responsiveness and maturity. A simple way to do this is to say (or write), "I did not understand...", which will always keep the onus for the misunderstanding on yourself.

The Golden Rule of Netiquette

With these top five netiquette rules, you are on your way to a great grade in your online course. Most importantly, when speaking in an online course or in any online environment for that matter, the same rules apply for etiquette as in real--time. The golden rule of netiquette in an online class or environment is, do not do or say online what you would not do or say offline. (Patsalides, Laurie. http://www.brighthub.com/education/online---learning/articles/26946.aspx
## Course Schedule

### Dates

<table>
<thead>
<tr>
<th>Dates/Modules</th>
<th>Readings</th>
<th>Lectures</th>
<th>Discussions</th>
<th>Assignments</th>
<th>Quizzes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Module 1</strong></td>
<td>1/11/16-1/17/16</td>
<td>Complete Course Readings</td>
<td>View Lectures</td>
<td>Discussion Board I: Introduction</td>
<td>Excel Practice I</td>
</tr>
<tr>
<td><strong>Module 2</strong></td>
<td>1/18/16-1/24/16</td>
<td>Complete Course Readings</td>
<td>View Lectures</td>
<td></td>
<td>Project Part I</td>
</tr>
<tr>
<td><strong>Module 3</strong></td>
<td>1/25/16-1/31/16</td>
<td>Complete Course Readings</td>
<td>View Lectures</td>
<td>Discussion Board II</td>
<td>Project Part II</td>
</tr>
<tr>
<td><strong>Module 4</strong></td>
<td>2/1/16-2/7/16</td>
<td>Complete Course Readings</td>
<td>View Lectures</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Module 5</strong></td>
<td>2/8/16-2/15/16</td>
<td>Complete Course Readings</td>
<td>View Lectures</td>
<td>Discussion Board III</td>
<td>Project Part III</td>
</tr>
<tr>
<td><strong>Module 6</strong></td>
<td>2/15/16-2/21/16</td>
<td>Complete Course Readings</td>
<td>View Lectures</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Module 7</strong></td>
<td>2/22/16-2/28/16</td>
<td>Complete Course Readings</td>
<td>View Lectures</td>
<td></td>
<td>Final Project</td>
</tr>
</tbody>
</table>

**Unless otherwise specified, all assignments, quizzes, discussion board postings due by 11:59 PM (EDT) of specified date.**
Module 1
Basic Terms of Statistics
1/11/16-1/17/16

Topic Overview
In this module, you will explore the basic meanings and uses of common statistical terms. You will learn how to calculate them by hand and by Excel. You will learn how and when to apply these terms as well as what is being avoided when certain terms are left out of an analysis

Learning Outcomes
By the end of this module, students will be able to:

1. Define Basic Statistical Terms and when to use them;
2. Calculate the Mean, Mode, and Median;
3. Describe what a distribution is;
4. Illustrate the different types of Charts;
5. Apply Standard Deviation.

Module Checklist
Your tasks for this module are:

1. Complete Module 1 Course Readings
2. View Module 1 Lectures & Other Media
3. Complete Module 1 Discussion Board (Initial Post due by: Thursday, January 14th at 11:59pm, Response to Peers due by: Sunday, January 17th at 11:59pm)
4. Complete Module 1 Learning Activity 1-1
5. Complete Module 1 Quiz (Opens: Thursday, January 14th at 5:00pm, Closes: Sunday, January 17th at 11:59pm)
6. Complete Module 1 Learning Activity 1-2

Required Readings:

Statistics, Data Analysis, and Decision Modeling, 5th Edition:
1. Chapter 1: Data and Business Decisions
2. Chapter 2: Displaying and Summarizing Data
3. Chapter 3: Probability Distributions and Applications

Media

1. L1-01 through L1-5

Assessment

Discussion Board Topics

By Thursday at 11:59pm, introduce yourself (in 150-200 words) to the other people in the class. In your posting, be sure to include the following:

1. Describe your current stance and base understanding of Statistics, Charting, Data mining, and Data Warehousing with regards to your education and work experience.
2. Describe what you want or expect to get out of this course.

By Sunday at 11:59pm, make a brief POSITIVE comment or personal understanding on one other person’s “Description of what they want to get out of the course”. No more than 2 may comment on another’s Description.

Assignments

Learning Activity 1-1:

How the Standard Deviation is calculated

Learning Activity 1-2:

Formulate a list of problems in your workplace that you want to be able to better understand or to answer as you progress through this course.

Quiz/Test

Quiz 1: Understanding Descriptive Statistics and Using Excel to find them
Module 2
Correlation and Regression
1/18/15-1/24/15

Topic Overview
Correlation and Regression: In this module, you will learn about how two variables that are both part of the same individual can be used together for prediction. You will learn how to calculate, understand and use to predict one value if you are given the other value. You will understand how there is always variation in nature and how to calculate, understand and work with this variation.

Learning Outcomes
By the end of this module, students will be able to:

- Illustrate the Normal Distribution;
- Demonstrate the basic concept of Correlation;
- Setup and dissect a regression formula;
- Use Excel to complete and graph a regression function.

Module Checklist
Your tasks for this module are:

1. Complete Module 2 Course Readings
2. View Module 2 Lectures & Other Media
3. Complete Learning Activity 2-1
4. Submit Assignment 2-1 (Due by: Sunday, Janauary 24th at 11:59pm)
5. Complete Module 2 Quiz (Opens: Thursday, January 21nd at 5:00pm, Closes: Sunday, January 24th at 11:59pm)

Readings
Statistics, Data Analysis, and Decision Modeling, 5/E

1. Chapter 6 Regression Analysis
2. Chapter 7 Forecasting

Media
1. L2-01 through L2-03
Assignments
Learning Activity 2-1: Regression Calculated

Project Part I
By Sunday, January 24th at 11:59pm, submit a research question you would like to have answered that relates to a situation or need at your work or in your field of study. This will be the basis of the last 3 parts of the final course project.

Quiz/Test
Quiz 2: Understanding Calculating, reading and using Correlation and Regression.
Module 3
Use of Standard Deviation in Practice
1/25/2016 – 1/31/2016

Topic Overview
In this module, you will explore how the Normal Distribution applies to any situation and can be used to help predict or explain variations in situations and between people. You will see how to use the scientific method to help prove your point or suggest quality improvements that support proper reasonable business decisions.

Learning Outcomes
By the end of this module, students will be able to:

- Apply the Law of Large Numbers;
- Explain the Central Limit Theorem;
- Graph and comprehend a Normal Distribution;
- Develop a Hypothesis;
- Test and use the Normal Distribution to make statistical Inferences.

Module Checklist
Your tasks for this module are:

1. Complete Module 3 Course Readings
2. View Module 3 Lectures and Other Media
3. Complete Learning Activity 3.1
4. Review Module 3 Toolbox
5. Complete Module 3 Discussion Board (Initial Post due by: Thursday, January 28th at 11:59pm, Response to Peers due by: Sunday, January 31st at 11:59pm)
6. Submit Module 3 Assignment 1 (Due by: Sunday, January 31st at 11:59pm)
7. Complete Module 3 Quiz (Opens: Thursday, January 28th at 5:00pm, Closes: Sunday, January 31st at 11:59pm)

Readings

Required Readings:

Statistics, Data Analysis, and Decision Modeling, 5th Edition:
1. Chapter 4: Sampling and Estimation
2. Chapter 5: Hypothesis Testing and Statistical Inference
3. Chapter 8: Statistical Quality Control

### Media
1. L3-01 through L3-08

### Toolbox
1. Pile Foundation.xls
2. Consumer Transportation Survey.xls
3. Customer Support Survey.xls

### Assessment
Learning Activity 3-1: Standard Deviation & Histograms

#### Discussion Board Topics

**Step 1: Read the following scenario:**

You have found that there are formulas to properly schedule amounts of time considering groups of samples taken “n” at a time. These sample means follow a normal distribution based on the information presented so far. This model COULD be used to forecast the amount of time it would take to complete a set of routine maintenance tasks nurses are scheduled to perform each week.

The hospital has nurses perform “routine tasks” that are defined in their union contract. The task could be anything from patient care to facility setup. We will just call it a “task” for the purpose of this discussion. You may dwell further in your discussion board defense.

For new nurses making $15 per hour on average, each task takes 1 hour on average to complete with a standard deviation of 30 minutes. For experienced nurses making $20 per hour, each task takes 50 minutes on average to complete with a standard deviation of 15 minutes.
Each nurse, no matter the experience level, must perform each task 20 times a week as part of their duty. In general, your hospital claims a baseline of 99% quality.

For this scenario, you are in contract negotiations and are trying to defend an amount of time allowed for the nurses to complete these tasks each week.

**Step 2:** Post a response to the discussion board link above by **Thursday at 11:59pm** regarding the questions and criteria below:

- What amount of time should be allocated to nurses to perform these tasks?
- Should we try to allocate more tasks to new or experienced nurses with the understanding that nurses need experience with the tasks to improve their time?
- We are looking to present a fair solution for the hospital and the nurses.

- Students with the last name A-M, take the position of a Hospital Administrator in these contracts.
- Students with the last name of N-Z, represent the Nurses Union in your posting.

**Step 3:** Using statistics to defend your position, comment on one other opposing person’s position with a negotiation reply by **Sunday at 11:59pm**. No more than 2 people can reply per position.

**Part II of Final Course Project**

**Research Search Results**

For this assignment, complete an initial search on the subject you chose to learn more about or answer. Find at least 4 other sources of similar issues, some statistics and some data about others who have faced a similar problem. Report on your literary search, some base findings and your cited resources to the link above no later than **Sunday, January 31st at 11:59pm**.

**Quiz/Test**

*Quiz 3: Hypothesis Testing and using the Normal Distribution to make reasonable business decisions*
Module 4
Terminology of Data Warehouses and DataMarts
2/1/15-2/7/15

In this module, you will explore the terminology of Data Warehouses and Datamarts. You should better understand what is meant by different Architectures as well as be able to see some of the advantages and difficulties of setting up a data warehouse or datamart.

Learning Outcomes

By the end of this module, students will be able to:

- Explain reasons for need of data warehousing;
- Describe the basic definitions and concepts of data warehouse;
- List the different types of data warehousing architectures; their comparative advantages and disadvantage;
- Compare the differences between Inmon and Kimball’s models for DW development;
- Describe the processes used in developing and managing data warehouse;
- Describe real-time (a.k.a. right-time and/or active) data warehousing.

Module Checklist

Your tasks for this module include:

1. Complete Module 4 Course Readings
2. View Module 4 Lectures and Other Media
3. Complete Module 4 Quiz (Opens: Thursday, February 4th at 5:00pm, Closes: Sunday, February 7th at 11:59pm)

Readings

The books and/or online materials you need for this module are:

- Data Warehousing and Online Analytical Processing
  - Read Sections 4.1 & 4.3 (Data Warehouse Design & Usage)
Media
Lectures L4-1 through L4-6

Assessment

Quiz/Test
Quiz 4: Details and terminology of the Data warehouse and datamart.
Module 5
Planning out a Data Warehouse
2/8/15-2/14/15

In this module, you will explore the intricacies of planning out a data warehouse. You will look at different methods and models for building, implementing and combining data from disparate sources to deliver the knowledge your decision makers need to run your business.

... 

By the end of this module you will be able to:

- Suggest a basic plan for the building and implementation of a data warehouse;
- Avoid common pitfalls or at least prepare for them;
- Aide with the modeling of a schema for your data;
- Plan for and set up different dimension tables and different fact tables for expected scheduled and adhoc queries.

Your tasks for this module include:

1. Complete Module 5 Course Readings
2. View Module 5 Lectures & Other Media
3. Complete Module 5 Discussion Board (Initial Post due by: Thursday, February 11th at 11:59pm, Response to Peers due by: Sunday, February 14th at 11:59pm)
4. Complete Module 5 Quiz (Opens: Thursday, February 11th at 5:00pm, Closes: Sunday, February 14th at 11:59pm)

The books and/or online materials you need for this module are:

- Data Warehousing and Online Analytical Processing
  - Read Sections 4.2 (Data Warehouse Modeling: Data Cube and OLAP) & 4.4 (Data Warehouse Implementation)
Lectures 5-1 through 5-5

Assessment

Discussion Board Topics

Step 1: Read the following scenario:

Your company has decided to implement a data warehouse in some form. The CIO of your company has appointed you to manage the project.

Step 2: Post a response to the discussion board by Thursday at 11:59pm regarding the questions and criteria below:

1. Give a brief description of your company and industry.
2. State what Operational Systems, Transactional Systems and other systems are you going to use to extract data to your operational data store(s). Make sure to include current datamarts or data warehouses if they exist. State what model/Architecture you plan to use as a basis for your data warehouse (datamarts).
3. Give a timeframe, scope, initial expected budget and plan of attack from an “organizational” standpoint to begin your project.

Step 3: By Sunday at 11:59pm, comment on one other person’s plan with constructive suggestions or possible pitfalls of their strategy so that their chances of success are increased. No more than 2 people can reply per position.

Quiz/Test

Quiz 5: Planning a Data warehouse.
Module 6
Data Mapping, ETL
2/15/15-2/21/15

Topic Overview
In this module, you will explore the importance of data governance, data quality, and programs for improving data quality. We will discuss three approaches for data integration including processes and procedures for Extracting, Transforming and Loading data from transactional and operational systems to an operational datastore.

Learning Outcomes
By the end of this module, students will be able to:

- Identify characteristics of quality data;
- Describe steps for data integration;
- Sort through some basic pitfalls of transforming data;
- Work with Excel to manipulate data.

Module Checklist
Your tasks for this module are:

1. Complete Module 6 Course Readings (Suggested)
2. View Module 6 Lectures & Other Media
3. Complete Learning Activity 6.1
4. Submit Part III of Final Project (Due by: Sunday, February 21st at 11:59pm)
5. Complete Module 6 Quiz (Opens: Thursday, February 18th at 5:00pm, Closes: Sunday, February 21st at 11:59pm)

Readings
1. (SUGGESTED!!...not Mandatory) Measuring Quality Improvement in Healthcare
   o Chapter 3 – Data Collection
   o Chapter 4 – Understanding Variation
   o Chapter 5 – Using Run and Control Charts to Analyze Process Variation
   o Chapter 6 – Control Chart Case Studies

Media
Lectures 6-1 through 6-3
Learning Activity 6-1: Extract Transform Load Example

Project Part III: Due Sunday, February 21st at 11:59pm

With regards to your “question you would like answered within your company”, complete the questions below and submit them to the link above no later than Sunday, February 21st at 11:59pm.

1. State what Operational Systems, Transactional Systems and other systems are you going to use to extract data to your operational data store(s). Make sure to include current datamarts or data warehouses if they exist.

2. State what model/Architecture you plan to use (or are using) as a basis for storing your data to be able to answer your project question.

3. Give a timeframe, scope, initial expected budget and plan of attack from an “organizational” standpoint to begin your project.

4. Describe the approach you would plan to use for Data Integration (Consolidation, Federation or Propagation)

5. Give an example of at least one star schema fact table that you might need, along with the fact fields and foreign key fields, to answer your question.

6. Give explanations of at least two dimensional tables that would support the fact table.

Quiz/Test

Quiz 6: Data Quality, Data Mapping, and Data Integration
Module 7
OLAP and Data Mining
2/22/16– 2/28/16

Topic Overview
In this module, you will explore Online Analytical Processing models and data mining techniques. You will explore what it means to Cube data in multiple dimensions, Slice, Dice, Roll up and drill down. You will explore Pivot tables and reporting methodologies.

Learning Outcomes
By the end of this module you will be able to:

- Define Online Analytical processing;
- Define Data Mining;
- Apply pivoting to a dataset and return answers quickly to what could be difficult questions.

Module Checklist
Your tasks for this module are:

1. Complete Module 7 Course Readings
2. View Module 7 Lectures & Other Media
3. Submit Final Project (Due by: Sunday, February 28th at 11:59pm)
4. Complete Module 7 Quiz (Opens: Thursday, February 25th at 5:00pm, Closes: Sunday, February 28th at 11:59pm)

Readings
The books and/or online materials you need for this module are:

- Data Warehousing and Online Analytical Processing
  - Section 4.5 (Data Generalization by Attribution-Oriented Induction)
  - Section 4.6 (Summary)
- Mining Frequent Patterns, Associations, and Correlations: Basic Concepts and Methods
Learning Activity 7-1: Pivot Tables and Multiple Correction Analysis

Final Course Project Assignment: Due by **Sunday, February 28th at 11:59pm**

Using the link above, submit your Final Course Project by Sunday, March 1st at 11:59pm. The following criteria must be included in your submission:

1. From a multi-dimensional standpoint, present at least 2 OLAP cube setups which you would expect to have access to in order to answer your research question.
2. Explain an example of a standard report that you would want your system to return including data and graph type(s) that would need to be shown on this report.
3. Explain what time frames this report would need to be generated with regards to Active or Historical Information as well as timing of report generation.

**Quiz/Test**

*Quiz 7: OLAP and Data Mining*