

# MIS 150 Introduction to Data Science Spring 2016 Syllabus

**Instructor:** Dr. Angela Berardinelli

**Office:** Old Main Tower 403

**Office Phone:** 814-824-2421

**E-mail:** aberardinelli@mercyhurst.edu

**Class Times:** MWF 9:15-10:20 AM (Section 1); MWF 10:30-11:35 AM (Section 2)

**Class Location:** Old Main Main Lab

**Office Hours:** MW 3:30-5:00PM, Tu 9:30AM-12:30PM, or e-mail to request an appointment

**Course Webpage:** <http://math.mercyhurst.edu/~aberardine/classes/MIS150S16.php> and  
<http://math.mercyhurst.edu/moodle/>

**This syllabus is a contract.** It is meant to tell you what you can expect of me, and what I will expect of you. It is a binding document you should read and understand thoroughly.

## 1 Course Description

An introduction to Microsoft Excel and Access in a data science context. The focus will be on machine learning techniques, particularly cluster analysis, Naïve Bayes, and ensemble methods. 3 credits.

**Prerequisites:** None.

## 2 Required Resources

**Textbook:** *Data Smart: Using Data Science to Transform Information into Insight*, 1st Edition, by John W Foreman.

**Software:** We will be using Microsoft Excel 2013 and Microsoft Access 2013 in class. This software is already available on the computers in the lab. (Note: Everything we do in Microsoft Excel can also be done in Excel for Mac and LibreOffice.)

**Data Storage:** The lab computers are “wiped” every time they restart. This means **any files you save on the lab computers will be deleted when you log off**. Therefore, in order to follow along with our work in class, you will need an alternative way of saving your files so that you can refer back to them later (in the next class period when we are continuing with the same file, or outside of class when you are studying). Here are a few options:

1. You can bring a USB drive to class every day and use that to save your files for class.
2. You can use cloud storage (Google Drive, Dropbox, etc). In fact, your Mercyhurst account includes OneDrive access (Microsoft’s cloud storage system).
3. You can choose to “map” your Mercyhurst network drive before the start of each class; I will provide you with instructions on how to do this if you prefer.

- 
4. You can save the file on the lab computer then e-mail it to yourself at the end of class. This is very inefficient and I do not recommend it, but you can do it if you want to.

### 3 Course Objectives

A student who successfully completes this course will be able to:

- import Excel worksheets into Access;
- import Access data into Excel;
- use SQL SELECT, ORDER BY, WHERE, GROUP BY, HAVING, and JOIN;
- create calculated fields with SQL;
- create frequency distributions, column charts, line charts, and scatter plots in Excel;
- create pivot tables in Excel;
- use formulas, including array formulas, in Excel;
- employ relative and absolute cell references in Excel;
- use solver for optimization in Excel;
- compute basic statistics in Excel;
- model and solve linear programming problems with Excel;
- implement Naive Bayes classification in Excel;
- implement clustering algorithms in Excel, including K-means clustering and network clustering;
- compute Euclidean distance, Manhattan distance, and cosine similarity between/of two points;
- implement user-based and item-based collaborative filtering in Excel.

### 4 Course Topics

**Unit 1: Introduction to Excel.** Covers Chapter 1 in the textbook.

**Unit 2: Introduction to Databases.** Supplemental material.

**Unit 3: Linear Programming.** Covers Chapter 4 in the textbook.

**Unit 4: Naive Bayes Classification.** Covers Chapter 3 in the textbook.

**Unit 5: Clustering Methods.** Covers Chapters 2 and 5 in the textbook.

**Unit 6: Collaborative Filtering.** Supplemental material.

## 5 Grading

### Final Grade Calculation:

	Percentage of Final Grade
Homework Projects	40%
Quizzes	10%
Exam 1	15%
Exam 2	15%
Final Exam	20%

### Letter Grade Scale:

If you have a weighted average of at least:	94%	90%	84%	78%	70%	65%	60%	0%
then you will earn a(n):	A	B+	B	C+	C	D+	D	F

**Grade Assignment:** Student grades will be determined based solely on the evaluation criteria listed in this section of the syllabus. Grades reflect proficiency in the course content as demonstrated on the graded evaluation criteria. In particular, if you want to earn an A, you need to demonstrate consistent excellence over the course of the entire term; an A on the final is not equivalent to an A in the course.

## 6 Homework Projects

There will be eight (8) homework projects assigned throughout the semester. Each homework project will involve analyzing a given data set using Excel and/or Access. The average of your 8 scores on these projects will constitute your “Homework Project” grade for the course. These projects will be posted on the course Moodle page, and you will submit your files through Moodle as well.

It is expected that you will spend approximately 6-9 hours per week studying the material outside of class meetings according to the typical 2-3 hours per credit rule of thumb. **In addition to time spent out of class, every other Monday that we have class will be a “Lab Day.”** On Lab Days, you will be expected to work independently on the assigned homework projects. I will be available to give assistance and answer questions.

**Late Homework:** For every day that an assignment is late, there will be a 10% penalty. If you turn in your assignment between 1 second late and 24 hours late, the highest grade you can receive is 90%; if you turn in your assignment between 24 hours plus 1 second late and 48 hours late, the highest grade you can receive is 80%; and so on. No assignments will be accepted more than 7 days late (168 hours after the deadline).

---

**Late Homework Example Grading:** The late penalty will be calculated as a scale, not a direct subtraction. For example, if your assignment scores 88 out of 100 possible points, but you submitted it two days late (for a maximum grade of 80%), your actual recorded grade on the assignment will be 70.4 out of 100 ( $88 \times 0.8 = 70.4$ ) - NOT 68 out of 100 ( $88 - 20 = 68$ ).

## 7 Quizzes

Your quiz score, which accounts for 10% of your final course grade, will be based on your average score of all quiz assignments. Some quiz scores will come from completion of out of class tasks (such as the first day of class survey). The rest of your quiz scores will come from traditional short content-based questions administered in class. All such assignments/quizzes will be announced in class in advance. There will be no makeup quizzes or extra credit.

## 8 Exams

There are three exams scheduled for this semester, two in-class exams and a final exam. Exam 1 will cover Units 1-4 and will be held in class on Wednesday, March 16. Exam 2 will cover Unit 5 and will be held in class on Wednesday, April 20. The final exam will be cumulative (Units 1-6). Your final exam day and time slot depends on your section.

Section 1 (9:15AM): Final Exam will be held **Friday, May 20 from 8:00AM-10:00AM**

Section 2 (10:30AM): Final Exam will be held **Wednesday, May 18 from 10:30AM-12:30PM**

**Exam Format and Expectations:** Each exam may be made up of two components: an Excel component and a written component. The written component will be “on paper” and will not be multiple choice. You will not be permitted to use technology of any kind during this portion of the exam (no calculator, no computer, no cell phone, no tablet, etc). The Excel component will be submitted through Moodle and must be completed in the classroom during the time allotted. Your activity on the computer will be monitored during exam periods. During the exam period, if you open any program other than a web browser or Microsoft Excel, and/or if you visit any site other than the Mercyhurst Moodle site in your web browser, you will receive a zero on the exam for cheating.

**Exam Preparation:** Throughout the class notes, there will be “on your own” practice problems. These problems, as well as homework and quiz problems, are the types of problems that may appear on exams. You should use these as practice for the exams.

**Make-Up Exams:** No make-up exams will be given. If you know you are going to miss a scheduled exam for a pre-scheduled event (examples: Mercyhurst-operated sporting event, academic event for another Mercyhurst course, doctor’s appointment, wedding, etc.), you must contact me via e-mail or in office hours at least seven days before the exam is scheduled to take place to arrange to take the exam early. That is, you may arrange to take the exam prior to the scheduled date and time, but you may not make an exam up after it has been administered in class. If you miss an in-class exam, you will receive a zero.

## 9 Course Policies

**Note-taking:** I highly recommend taking notes in this class. You can do this on the computer (in the Excel document we are using in class that day, or in a separate text file, or in a note-taking app on your own phone, tablet, or laptop) but I sincerely recommend using a good old-fashioned notebook and pen/pencil. This will be a way of keeping all of your notes for the class in one place, instead of spread over a bunch of different files. Good note-taking is an essential skill to develop and perfect during your college career. For this class in particular, following along step-by-step in class is great, but filling in the gaps with helpful notes to yourself will be invaluable when you are studying for exams, or if you have to look back at the course material in 5 years after your boss says "Hey can you put together a spreadsheet that does A B and C?" and you say "Yes! You can count on me!".

**Attendance:** Students are responsible for all information (notes, announcements, etc.) given in class, regardless of attendance.

**E-mail:** You can always e-mail me with course-related questions or to request an appointment outside of office hours. However, you should allow up to 2 days for a reply to your e-mail. Also, you should not e-mail me with questions about your grade; to discuss your grade please meet with me in person in my office. E-mail is not a substitute for class attendance. Sometimes, I will need to send out e-mail communications to the class. These communications will be sent to your Mercyhurst account. I will not send to any other e-mail account you may use, so be sure you have access to your Mercyhurst account and check it often enough to receive these important announcements in a timely manner.

**Classroom Etiquette:** Please be courteous to the instructor and your fellow students and silence your cell phone before class and do not send or receive calls or text messages during class time. Take off your headphones; do not read the newspaper or other books. Avoid disrupting the instructor and your classmates by arriving to class late or leaving class early unless absolutely necessary.

**Academic Integrity:** Cheating and plagiarism in any form are serious offenses and will be dealt with as such. Please see the course website for details about "What is cheating in a computer course?" and the consequences of such an offense. Ignorance of standards will not be an acceptable excuse. University policy related to this issue may be found in the Student Handbook under Academic Affairs. <https://my.mercyhurst.edu/handbook/academic-affairs/>

**Regarding Learning Differences:** In keeping with college policy, any student with a disability who needs academic accommodations must call Learning Differences Program secretary at 824-3017, to arrange a confidential appointment with the director of the Learning Differences Program during the first week of classes.

**Support of the Mercy Mission:** This course supports the mission of Mercyhurst University by creating students who are intellectually creative. Students will foster this creativity by: applying critical thinking and qualitative reasoning techniques to new disciplines; developing, analyzing, and synthesizing scientific ideas; and engaging in innovative problem solving strategies.

## 10 Course Schedule

Week of	Class Dates	Material to be Covered	Special Events
2/1	2/3, 2/5	Unit 1	None
2/8	2/8, 2/10, 2/12	Units 1, 2	Add/drop deadline Wednesday
2/15	2/15, 2/17, 2/19	Units 2, 3	Lab Day
2/22	2/22, 2/24, 2/26	Unit 3	None
2/29	2/29, 3/2, 3/4	Unit 4	Lab Day
3/7	3/7, 3/9, 3/11	Unit 4	None
3/14	3/14, 3/16, 3/18	Unit 5	Lab Day on Monday; Exam on Wednesday
3/21	N/A	N/A	Spring Break
3/28	3/30, 4/1	Unit 5	Easter Break Monday
4/4	4/4, 4/6, 4/8	Unit 5	Midsemester grades due Friday
4/11	4/11, 4/13, 4/15	Unit 5	Lab Day
4/18	4/18, 4/20	Units 5	Exam on Wednesday; Spring Break Friday
4/25	4/25, 4/27, 4/29	Unit 6	Lab Day
5/2	5/2, 5/4, 5/6	Unit 6	None
5/9	5/9, 5/11, 5/13	Review	Lab Day and Review for Final Last day to withdraw Friday
5/16	N/A	N/A	Final Exam

**Final Note:** This syllabus is subject to change if deemed necessary. Any syllabus changes or addendum will be communicated in class.