Spring 2019 Wales Study Group

Director: Ahmet Ay,
Department of Biology and Mathematics

For more information, please email aay@colgate.edu

Information Sessions:
Monday, October 23rd at 7pm
in 101 McGregory Hall and
Tuesday, October 24th at 4pm
in 301 Olin Hall

Application Deadline: Wednesday, November 15, 2017
www.colgate.edu/OCS
Spring 2019 Wales Study Group

Director: Ahmet Ay

Associate Professor of Biology and Mathematics

Purpose:
The Wales Study Group is intended for majors in the Division of Natural Sciences and Mathematics who wish to live and study abroad. Although the program is designed primarily for students who will be juniors during the spring term of 2019, students who will be sophomores in spring 2019 may also be considered.

The program is hosted at Cardiff University during that college’s spring semester. The Colgate study group’s orientation is tentatively scheduled to begin Monday, 21 January 2019. The last final exam day is Friday, 14 June 2019 with the official program end date Saturday, 15 June, 2019.

Site:

As in the past, the study group will be based at Cardiff University, which, with its student body of thirty thousand students, is the preeminent research university of Wales. Colgate students will register as non-matriculated students of the University of Cardiff and have access to all of the University’s academic and athletic facilities, including libraries, gymnasiums, and student union.

Wales is part of the United Kingdom and only in 1998 formed its own National Assembly. Its people consider Wales to be a separate and distinct nation; they are passionately conscious of their cultural heritage. While the dominant language of southern Wales is English, Wales is a bilingual nation. The Welsh language, a Celtic tongue, is alive and well as a second language and is strongly promoted in the Welsh educational system. Cardiff is an ideal locale for exploring this fascinating culture, serving as a center for the renaissance of Welsh language, literature, and national pride.

The city of Cardiff, the capital of Wales, has a population of approximately 350,000. The Welsh are justifiably proud of their artistic and cultural accomplishments, as exemplified by city theaters for opera, drama and ballet — the Welsh National Opera and the BBC Welsh Symphony Orchestra are located in Cardiff — and by the Welsh National Museum. In addition, Cardiff University, which is located near the center of Cardiff in Cathays Park, maintains a number of musical and drama theaters. The principal sport of Wales is rugby, and Millennium Park, an internationally known rugby ground, is within walking distance of the University’s campus.
Courses:

All students will enroll in four courses: two elective courses (at least one normally must be in the student’s concentration) and in two required courses, which are designed specifically to reflect the study group’s unique location.

**AHUM 360 - Welsh Culture (required Cardiff course taught by the Welsh Dept.)**
This course examines Welsh literature, language, and culture, and in so doing, augments the cross-cultural experience that is an integral part of any study abroad program. The specific course content varies with the instructors, all of who are Cardiff University faculty. Before the semester begins, a member of the Welsh Department provides an introduction to the Welsh language. This course has a number of required field trips, including excursions to the Welsh Folk Museum, the Dylan Thomas home site, and the National Assembly of Wales. This course has been approved for GE credit.

**Director’s Course (required course)**
**MATH 169 – Dive into Data**: Data mining is the science of discovering patterns, and making predictions in large-scale data sets. This course provides an introduction to this field, and focuses on real life applications. In this course, we will study data mining methods for a wide range of tasks including visualization of data, classification and clustering. The students will be exposed to numerous industrial and scientific applications drawn from a wide range of topics including disease classification through whole-genome expression data sets, and personality prediction from social media interactions. No background in programming is presumed. Computer software that is designed for data mining will be introduced throughout the course. Prerequisite: Three years of secondary school mathematics.

**Two Elective Courses**
Of the two elective courses, at least one must be in the student’s concentration. Students must have satisfied the pre-requisites for the elective courses before departure for Wales. Students will consult with their academic departments and the Registrar’s Office to ensure that elective courses taken in Cardiff will be approved as Colgate credit.

Field Trips:

Students are required to participate in all of the study group trips. Tentatively planned are several day trips for the Welsh Culture course and the Director’s course, a two-day trip to Stonehenge and Stratford Upon Avon, a trip to western Wales and the Pembrokeshire National Park region, and a trip to the culturally distinct north of Wales.

Costs:

For details of student expenses on this study group, please see the student cost sheet on the Off-Campus Study/International Programs website.

Deadlines:

The deadline for applications to the Spring 2019 Wales Study Group is **Wednesday, November 15, 2017**. Applications are on the Off-Campus Study/International Programs (www.colgate.edu/ocs) study groups’ websites and are submitted online. Interviews of applicants will take place after that date and will be arranged by email. Student notification of selections will take place late December 2017.

Passports and Visas:

You must confirm that your passport is valid through December 2019. All students participating on the Spring 2019 Wales Study Group will be required to obtain a U.K. student visa. With participation on this study group comes the responsibility of understanding and complying with U.K. visa requirements. If you will not be traveling on a U.S. passport it is imperative that you contact an adviser in Off-Campus Study/International Programs, 101 McGregor (Center for International Programs), and International Student Services, 103C Lathrop, to learn as much as you can about the visa requirements. For some students there are significant requirements to be met that take time, advance planning, and incur extra costs.
GOALS:
Data mining is the science of discovering patterns, and making predictions in large-scale data sets. This course provides an introduction to this field, and focuses on real life applications. In this course, we will study data mining methods for a wide range of tasks including visualization of data, classification and clustering. The students will be exposed to numerous industrial and scientific applications drawn from a wide range of topics including disease classification through whole-genome expression data sets, and personality prediction from social media interactions. No background in programming is presumed. Computer software that is designed for data mining will be introduced throughout the course. Prerequisite: Three years of secondary school mathematics.

By the end of this course you are expected to:

1. Understand basic data processing and organization methods.
2. Know how to apply data mining techniques such as Regression, Support Vector Machines, and Clustering to industrial and scientific problems.
3. Know how to use data mining computer software such as R and WEKA. Data mining analysis could include gathering data, setting up an experiment, and performing the analysis.
4. Formulate interesting questions that could be answered using techniques from data mining, and reasonable plans for answering them. Students should be able to offer justifications for the selection of specific data mining techniques, and suggest circumstances under which other techniques would be better applied instead.
5. Be able to accurately and completely interpret and present results from a data mining study. Students should communicate their results using effective, and appropriately chosen analysis, and visualizations.

COURSE RESOURCES:

Textbook: An Introduction to Statistical Learning with Applications in R by Gareth James, Daniela Witten, Trevor Hastie, Robert Tibshirani

Course Website: The syllabus, homework assignments and other important communications will be conveyed via Moodle (http://moodle.colgate.edu/). I will also post outlines of the lecture notes for each class. I strongly encourage you to read them before each class.

COURSE POLICIES:
Class Time/Office Hours: This course will be mainly lecture based. The class meetings will be devoted to the introduction of concepts. The primary time to assist you with difficulties associated with the course work will be during visits to my office.

Classroom Policy: In order to ensure an atmosphere conducive to learning, students must be considerate for one another in the classroom. The following rules should be observed during class.

➢ Every student is expected to be in class on time.
➢ No chatting during lecture.
➢ Usage of cell phones during the class is STRICTLY forbidden. Please put your cellphone on silent mode before the class starts.
➢ Usage of computers during the class is STRICTLY forbidden, unless you have my permission.
➢ No leaving early without prior notification and approval from me.

GRADING:

Grades in this course will be based upon the point total of the scores from a midterm exam (30%), a final exam (30%), homework assignments (30%), participation / active learning exercises (10%).

On each of the assignments and exams I will list the mean and standard deviation, which will provide an indication regarding your performance on that particular assignment or exam. Your course grade will be determined by a subjective curve based upon the accumulation of the points on all assignments. Comparing your grades with the class averages will give an indication for your tentative letter grade for the course.

Participation/Active Learning Exercises: You are expected to attend the class and urged to ask questions in class whenever you believe this will enhance your understanding. Attendance/tardiness, and class participation will contribute to your overall assessment in the course. During the class we will do active learning exercises time to time. These assignments will be collected for grading a few times, and will constitute your grade for the active learning exercises.

Homework Assignments:

➢ Approximately every week a problem set that includes written or programming assignments using real world data will be assigned. These assignments will be posted on Moodle and are due at the beginning of the lecture of the date it is scheduled.
➢ Late assignments will not be accepted unless a valid excuse is presented. For the students with an excuse I will accept homework up to three days after the initial due date. Ten points will be deducted daily.
➢ Please note that, throughout the semester, a few times, you may be asked to explain your solutions for the submitted homework assignment of the week.
➢ Please do not submit your assignments by email unless you have made prior arrangements.
➢ You may work together on the homework, but the submitted work should be yours and yours alone. If you have questions concerning the grading write them on your
assignment paper and hand it in again. It will be returned with your next assignment.

**Exams:** Exam questions will require you to independently integrate material from topics presented across multiple lectures. Exams will also require you to apply the concepts you have learned to situations not covered directly in class. The exams will consist of two parts: a take-home part and an in class part. Note that the final exam will NOT be a comprehensive exam.

**Midterm Exam** covers class sections through Friday, March 8th. Take home part of the exam (25 pts) should be submitted at the beginning of the class on Friday, March 15th. The in class part of the exam (75 pts) will be given at 7pm-9pm on Monday, March 11th in our classroom.

**Final Exam** covers class sections March 11th through May 10th. Take home part of the exam (25 pts) should be submitted at the beginning of the final exam on May 17th. The in class part of the exam (75 pts) will be given at our classroom on Friday, May 17th (7pm-9pm).

**Make Up Exam Policy:** There will be no “make–up” exams. Unless a valid excuse is presented in advance, a missed exam will receive the score 0. Travel schedules (business or personal), a desire for a cheaper airfare, etc. are NOT acceptable excuses. Students must look at this syllabus carefully and plan well ahead.

**KEY POINTS FOR SUCCESS:**

**Attendance/Class Participation:** This course is highly dependent on class discussions. For this reason your attendance will be highly correlated to your success in this course. Any changes in this syllabus or in the scheduling of exams, assignments, etc. will be announced during class meetings. Please note that I rely heavily on material covered in class when I write exams. If you miss a class meeting, you should copy a classmate’s notes for that meeting. Please ask questions in class whenever you believe this will enhance your understanding.

**Study Groups:** I encourage you to study with others in the same course, and meet regularly to go over the concepts. Many students have reported phenomenal success in their math courses after they have joined a study group. You may freely discuss concepts, and solutions with classmates. However, you should provide your own solution for the assignments.

**STUDENTS WITH DISABILITIES:**

Any student who feels s/he may need an accommodation based on the impact of a disability should contact me privately to discuss her/his specific needs. Please provide a letter from Disability Services no later than a week before the first test.

**HONOR CODE POLICY:**

You are reminded of the Academic Honor Code in place at Colgate University. All tests and other graded assignments fall under the auspices of this policy. Unless otherwise indicated all graded materials must contain only work that you have completed yourself. Appropriate citation of the work or ideas of others is expected in all writing assignments. Cases of plagiarism will be referred to the Colgate Student Conduct Board. Students are reminded that
convictions of violations of the Academic Honesty Policy of Colgate result in the grade of F for
the entire course, not for the individual assignment.

**CAALENDAR (Subject to Modification):**

<table>
<thead>
<tr>
<th>Chapter 2</th>
<th>Jan 28</th>
<th>Chapter 2</th>
<th>Jan 30</th>
<th>Chapter 2</th>
<th>Feb 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chapter 2</td>
<td>Feb 4</td>
<td>Chapter 3</td>
<td>Feb 6</td>
<td>Chapter 3</td>
<td>Feb 8</td>
</tr>
<tr>
<td>Chapter 3</td>
<td>Feb 11</td>
<td>Chapter 3</td>
<td>Feb 13</td>
<td>Chapter 4</td>
<td>Feb 15</td>
</tr>
<tr>
<td>Chapter 4</td>
<td>Feb 18</td>
<td>Chapter 4</td>
<td>Feb 20</td>
<td>Chapter 4</td>
<td>Feb 22</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>Feb 25</td>
<td>Chapter 5</td>
<td>Feb 27</td>
<td>Chapter 5</td>
<td>Mar 1</td>
</tr>
<tr>
<td>Chapter 5</td>
<td>Mar 4</td>
<td>REVIEW</td>
<td>Mar 6</td>
<td>REVIEW</td>
<td>Mar 8</td>
</tr>
<tr>
<td>Mar 11</td>
<td>Chapter 8</td>
<td>Mar 13</td>
<td>Chapter 8</td>
<td>Mar 15</td>
<td></td>
</tr>
<tr>
<td>Mar 18</td>
<td>Chapter 8</td>
<td>Mar 20</td>
<td>Chapter 8</td>
<td>Mar 22</td>
<td></td>
</tr>
<tr>
<td>Mar 25</td>
<td>Chapter 9</td>
<td>Mar 27</td>
<td>Chapter 9</td>
<td>Mar 29</td>
<td></td>
</tr>
<tr>
<td>Chapter 9</td>
<td>Apr 1</td>
<td>Chapter 9</td>
<td>Apr 3</td>
<td>Chapter 10</td>
<td>Apr 5</td>
</tr>
<tr>
<td>Chapter 10</td>
<td>Apr 8</td>
<td>Chapter 10</td>
<td>Apr 10</td>
<td>Chapter 10</td>
<td>Apr 12</td>
</tr>
<tr>
<td>Apr 15</td>
<td>NO CLASS</td>
<td>Apr 17</td>
<td>NO CLASS</td>
<td>Apr 19</td>
<td></td>
</tr>
<tr>
<td>Apr 22</td>
<td>NO CLASS</td>
<td>Apr 24</td>
<td>NO CLASS</td>
<td>Apr 26</td>
<td></td>
</tr>
<tr>
<td>Apr 29</td>
<td>NO CLASS</td>
<td>May 1</td>
<td>NO CLASS</td>
<td>May 3</td>
<td></td>
</tr>
<tr>
<td>May 6</td>
<td>Chapter 10</td>
<td>REVIEW</td>
<td>May 8</td>
<td>REVIEW</td>
<td></td>
</tr>
<tr>
<td>May 13</td>
<td>May 15</td>
<td>May 17</td>
<td>FINAL 12PM-2PM</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>