

**Meets:** Tuesdays and Thursdays, 4 pm to 7 pm

**Lectures:** PHO 202 (first part of each class)

**Lab:** LSE B03 (last part of each class)

**Instructor:** Joe Mellor ([mellor@bu.edu](mailto:mellor@bu.edu), 617.233.3785)

**Office Hours:** by appointment

**Teaching Assistant:** Jiggy Parikh ([jparikh@bu.edu](mailto:jparikh@bu.edu))

## **Syllabus**

This course introduces basic principles and techniques of relational database systems, especially as applied to research and data storage problems encountered in biology. Lectures will review theoretical and practical concepts that underlie the design and implementation of robust biological database systems. The laboratory section of the course will cover practical exercises using Python and PostgreSQL in the creation and usage of databases. Later stages of the course will focus on the development and deployment of collaborative database projects related to current biological research problems.

Lecture topics will include:

### **Part I. Relational Data Models**

- Entity-Relationship (ER) Data Model
- Common Object Models for Biological Data
- Tables and Keys
- Data Redundancy
- Normal Forms

### **Part II. The Database Management System (DBMS)**

- Data Storage
- Survey of DBMS
- Structured Query Language (SQL)
- Indexing Strategies and Optimization
- Biology Namespaces and Integration Issues
- Query Optimization
- Data Input/Output, Versioning

### **Part III. Database Interface Design**

- Remote APIs
- Programmatic Interfaces
- SQL and Python
- Web/DB Interfaces
- Data Input/Output, XML
- Web Services (e.g SOAP, RSS, AJAX)
- Human Computer Interface Issues

## **Grading:**

Midterm Exam	20%
Homework	30%
Final Project	50%

## **Midterm Exam**

An exam that covers Parts I and II of the lectures will be held approximately half-way through the semester. Material for the exam will be drawn from the lectures and lecture notes posted on the class website.

## **Homework**

Homework exercises will be assigned periodically during the course. There will be approximately six assignments, combined worth 30% of the final grade. Assignments will be turned in electronically through the BE768 Blackboard site. Late work will not be accepted unless prior arrangement has been made.

## **Final Project**

The second half of the semester will be heavily focused on designing and building a novel biological database and accompanying interface. Projects will be done in small groups (of three students each), with milestones during the semester for preliminary proposals, implementation proposals, final report and presentation to the class. Detailed guidelines and deadlines for the project will be distributed in late January.

## **Course Materials**

Lecture notes, homework assignments and other reading materials will be posted on Blackboard at ([http://blackboard.bu.edu/bin/common/course.pl?course\\_id=5449\\_1](http://blackboard.bu.edu/bin/common/course.pl?course_id=5449_1)) (Kerberos login required).

When necessary, lab material and programming examples used in the course (database scripts, python/SQL examples, etc.) will be accessible on the class web/database server at <http://bioed.bu.edu>.

## **Laboratory Session**

Programming tutorials and exercises will be designed around the Python programming language and PostgreSQL database system.

## Resources and Tutorials (also updated on class website):

### HTML

<http://www.w3schools.com/html/>

<http://www.yourhtmlsource.com/starthere/>

### PostgreSQL

<http://www.postgresql.org/>

<http://www.postgresql.org/docs/8.0/static/tutorial.html>

<http://www.petefreitag.com/cheatsheets/postgresql/>

<http://www.faqs.org/docs/ppbook/book1.htm>

### SQL

<http://www.w3schools.com/sql/>

<http://sqlcourse.com/intro.html>

<http://sqlzoo.net/>

<http://philip.greenspun.com/sql/>

### Python

<http://docs.python.org/>

### Python CGI

<http://docs.python.org/library/cgi.html>

[http://www.cs.virginia.edu/~lab2q/lesson\\_1/](http://www.cs.virginia.edu/~lab2q/lesson_1/)

<http://www.python.org/doc/essays/ppt/sd99east/index.htm>

<http://www.algonet.se/~ug/html+pycgi/>

<http://www.upriss.org.uk/python/PythonCourse.html>

<http://www.devshed.com/c/a/Python/Writing-CGI-Programs-in-Python/>

[http://gnosis.cx/publish/programming/feature\\_5min\\_python.html](http://gnosis.cx/publish/programming/feature_5min_python.html)

### Python and PostgreSQL/DbAPI

<http://python.about.com/od/pythonanddatabases/ss/postgresqlread.htm>

<http://python.about.com/od/pythonanddatabases/ss/dbpostgresql.htm>

<http://wiki.python.org/moin/DbApiCheatSheet>

<http://wiki.python.org/moin/UsingDbApiWithPostgres>

<http://initd.org/>