Data Analysis and Modeling
CSE 5301

Jean Gao
Dept. of Computer Science and Engineering
gao@uta.edu

CSE5301: Data Analysis and Modeling

- Instructor: Jean Gao
gao@uta.edu
- Office hours: Tues & Thus, 12:30pm-1:30pm, or by appointment
- Office Location: Engineering Research Building (ERB), Room 538
- Course Website:
  
  http://crystal.uta.edu/~gao/CSE5301_DataAnalysis/Syllabus.html
  Or from Blackboard.

The lecture notes if there is any will be posted on-line the day before class. Please have them ready when coming to class.
Course TA Info

- TA: Fariba Khoshghalbvash
- Email: fariba.khoshghalbvash@mavs.uta.edu
- Office: ERB 545
- Office Hour: Tue & Thu, 3:30pm - 5pm or by appointment

Grading Scale

- Homework Assignments (~6) 10%
- Quizzes (3) 15%
- Three Exams 75%
Grading Scale (cont’d)

• Class attendance is **required** which will help you to keep pace with the class material and understand course content.

• Hard copies of hand-written or typed HW assignments are collected physically in or before class on the due date. **Late submissions will not be accepted and extensions will only be granted in extreme situations.**

• Homework assignments will be graded based on the completeness. So please try your best to finish them.

• There will be **3** fifteen-minutes quizzes. Each quiz will cover materials since last quiz day. (Closed book, closed notes, no internet/phone access, no calculator usage.)

Grading Scale (cont’d)

• Discussions on homework are encouraged for understanding the techniques, however you are supposed to provide the solutions independently.

• Exams are closed book, closed notes (No phones or any other internet access. Only scientific calculators are allowed and NO programmable ones. Two double-sided cheat sheets will be permitted.). All the exams will be one hour and 20 minutes long. None of them will be comprehensive.

• It is the student’s responsibility to read and review the assigned sections in the textbook.

• **There will be NO make-up exams or quizzes unless** the instructor has been notified in advance.
Final letter grade rubric

<table>
<thead>
<tr>
<th>Weighted Final Grade</th>
<th>Final Letter Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>≥ 85</td>
<td>A</td>
</tr>
<tr>
<td>≥ 70</td>
<td>B</td>
</tr>
<tr>
<td>≥ 55</td>
<td>C</td>
</tr>
<tr>
<td>≥ 40</td>
<td>D</td>
</tr>
<tr>
<td>Otherwise</td>
<td>F</td>
</tr>
</tbody>
</table>

- The instructor reserves the right to lower these thresholds, based on the distribution of scores and the difficulty degree of quizzes and exams.

More on grading policy

- *Any request for re-grading or reviewing of exams/quiz/Hws must be made within 5 days of receipt of that grade.* Re-grading can lead to a higher or lower grade, depending on grading errors that are discovered.

- There will be no extra credit assignments.

- No grade bumps will be offered; 84.44 is a B if 85 is the cut-off.
Course Readings

Required Textbook:


This week read:
Chap. 2

References:
Additional course materials will be available electronically through the course website.

Honesty

- No collaboration on tests or quizzes.
- Academic dishonesty will not be tolerated and violations will result in severe penalties for all parties involved.
- You are expected to know the University's definitions and policies regarding academic dishonesty.
Goal of this Course

• To provide students the basic data analysis and modeling concepts and methodologies using probability theory.

• Students successfully completing this course will have formed an important basis solving practical statistics and data analysis related problems.

Tentative Topics to be Covered:

• Basic probability:
  Discrete and continuous random variables, independence, covariance, central limit theorem, Chebyshev inequality, diverse continuous and discrete distributions.

• Statistics, Parameter Estimation, and Fitting a Distribution:
  Descriptive statistics, method of moments, maximum likelihood estimation

• Computer Simulation and Stochastic Processes:
  Sampling of continuous distributions, Monte Carlo methods, Markov Processes, Poisson Process
Tentative Topics to be Covered: (cont’d)

• Hypothesis Testing:
  Type I and II errors, rejection regions; Z-test, T-test, F-test, Chi-square test, Bayesian test, Confidence Intervals
• Data Modeling:
  Hidden Markov Models, Bayesian network if time allows.