DEPARTMENT OF MATHEMATICS
EAST STROUDSBURG UNIVERSITY

MATH 416: Linear Statistical Modeling Methods with SAS – Spring 2024

Instructor: Xuemao Zhang	**Email:** xzhang2@esu.edu	**Telephone:** 570-422-3441
Office Location: SciTech Building room 128		
Office Hours **Monday** 1:00PM – 3:00PM **Wednesday** 1:00PM –3:00PM **Thursday** 10:00AM – 11:00AM and other times by appointment		
Secretary: Christine Getz	**Email:** cgetz@esu.edu	
Office Location: SciTech Building room 118	**Telephone:** 570-422-3447	

Course Prerequisites: MATH 110 or MATH 311.
Course Description: This course is intended for advanced undergraduate students, graduate students and working professionals who engage in applied research. Statistical linear modeling methods are used in conjunction with SAS computer software to analyze correlated data, applications in supervised machine learning and data analysis for experimental designs. Topics include linear regression analysis, logistic regression, generalized linear models, supervised machine learning and ANOVA for design of experiments.

TextBook:

- *Applied Linear Statistical Models* (5th Edition) by Michael Kutner, Christopher Nachtsheim, John Neter and William Li, McGraw-Hill, 2005.
- *An Introduction to Statistical Learning: with Applications in R* by Gareth James, Daniela Witten, Trevor Hastie and Robert Tibshirani 2017 (For eTextbook and R/Python code, go to https://www.statlearning.com/).

Required:

1. **Laptop/Desktop and Internet**: Technology needs of a Laptop/Desktop computer, and 24/7 reliable Internet Access are required for our class.
2. **Desire2Learn(D2L):** https://esu.desire2learn.com/ D2L is the application we use to help manage classes. Relevant course material and grades will be made available on the course website D2L. Important dates or sudden announcements will also be made via the course main page and news section. It is important to visit the course page regularly.
3. **Lecture Notes:** Typed lecture notes will be available on D2L. It is recommended to bring a printed copy of the lecture notes to each class. Then you can simply add your own solutions/thoughts/comments/reminders as needed and have more time to digest the lecture.
4. **WeBWork Homework**: Homework in this course are assigned through WeBWork https://postulate.klnpa.org/webwork2/Math416_Linear_Models/ Your username for the

WeBWork is your MyESU/D2L username and your default password will be your student number without the leading 0's (**will be available to use soon**).

5. **SAS:** One of the most comprehensive and widely used statistics programs is SAS (Statistical Analysis System. SAS OnDemand for Academics in the cloud is free for us to use. Please create an account https://www.sas.com/profile/ui/#/create first. For registration, go to https://welcome.oda.sas.com/enroll with course code **820ecb9c-8fc2-48af-83e6-d67504f0c3df**. Running SAS programs will be illustrated in the lectures. Other resources for learning SAS include:
   - The UCLA Academic Technology Services site https://stats.oarc.ucla.edu/sas/ is highly recommended!
   - https://www.tutorialspoint.com/sas/index.htm
   - The Little SAS Book: A Primer, 5th Ed, Lora Delwiche and Susan Slaughter, SAS Institute.

SAS Global Certification program: SAS is widely used in industry, and it is advantageous to know SAS to be a data analyst. The SAS Institute offers a SAS Global Certification program (https://www.sas.com/en_us/certification.html) and SAS Certification credentials are globally recognized as the premier means to validate SAS knowledge. You need to pass (must achieve score of 72.5% correct to pass) the SAS Base Programming for SAS 9.4 exam to obtain the first certificate, SAS 9.4 Base Programming - Performance-Based Exam (https://www.sas.com/en_us/certification/credentials/foundation-tools/base-programming-specialist.html#exam-details). This course will cover most topics required by the exam. However, it does not prepare you for the exam since we focus on the applications of linear statistical models and analysis of data from experimental designs. You need to spend extra time (like 1 to 2 months, at least 1 hour each day) to learn and practice to be prepared for the exam.

6. **AI to Assist Mathematical Reasoning**: (1) Though ChatGPT (https://chat.openai.com/) is not a specialized mathematical software, it has been trained on a wide range of texts and can handle mathematical concepts and reasoning up to a certain level of complexity. If you encounter any limitations or errors in its responses when you use ChatGPT, double-check its answers with other sources or consult with your instructor. (2) ChatGPT can also be used to **debug your SAS programming code**.

Course Format: Three weekly lecture hours will be used for the presentation of course material, recitations, and exams.

Lecture Hours		Classroom
Monday	9:00 – 9:50Am	Science & Technology Center 145
Wednesday	9:00 – 9:50Am	Science & Technology Center 145
Friday	9:00 – 9:50Am	Science & Technology Center 145

- ❖ In the event of inclement weather and the closure of the university, the class will be conducted via Zoom https://esu-online.zoom.us/j/4509134271

Topics covered: The course gives a thorough exposition of regression analysis, generalized linear models, supervised learning, and ANOVA for experimental designs. It covers in depth simple linear regression models, multiple linear regression models and generalized linear models, transformations, inferences in regression and correlation analyses, simultaneous prediction intervals, diagnostics and remedial measures, and model selections in machine learning; analysis of variance for completely randomized design, randomized complete block design, and factorial design. The computer package SAS is used for data analysis.

Student Learning Outcomes:

After completing the course students will demonstrate the ability to:

(1) **Review:** Random variables and some typical continuous distributions: Normal, Student's t, Chi-square and F distribution; Statistical inferences about a population mean. If you took MATH 311 and MATH 411, please be patient for the first two weeks.

(2) **Simple Linear Regression Models (Chapter 1):** Use appropriate graphs (scatterplots) and summary measures (means, standard deviations and correlation) to summarize the relationship between two numerical variables. Be able to specify, and fit (using least squares) a suitable simple linear regression model.

(3) **Inferences in Regression and Correlation Analysis (Chapter 2):** Be able to conduct ANOVA (Analysis of Variance) and assess quality of the fit, make statistical inferences about the intercept, slope, mean response and individual response.

(4) **Diagnostics and Remedial Measures (Chapter 3):** Be able to conduct analysis of residuals to check model assumptions.

(5) **Simultaneous Inferences (Chapter 4):** Be able to jointly estimate the intercept and slope, fit and make inferences about a simple linear regression model with fixed intercept.

(6) **Matrix Approach to Simple Linear Regression Models (Chapter 5):** Be able to use matrix approach to analyze simple linear regression models.

(7) **Multiple Linear Regression Models (Chapter 6 and 7):** Be able to fit and interpret a multiple linear regression model. Conduct ANOVA (Analysis of Variance) and assess quality of the fit, make statistical inferences about the regression coefficients, mean response and individual response models. Recognize the phenomenon of multicollinearity.

(8) **Regression Models for Quantitative and Qualitative Predictors (Chapter 8):** Understand and apply transformations and include qualitative predictors in a polynomial regression model.

(9) **Building the Regression Model (Chapter 9, 10 and 11):**
- Be able to apply some influence diagnostics: Cooks D, Dffits, Dbetas, and multicollinearity diagnostics: tolerance, variance inflation factor, condition indices, condition number, weighted least squares, and ridge regression to remedial regression models;
- Be able to use several selection criteria for "best" regression model selection: the stepwise, forward, and backward methods to select and validate "best" regression model(s).

(10) **Generalized Linear Models (14)**: Be able to fit and make inferences about generalized linear models including logistic regression and Poisson regression models.

(11) **Regression methods in statistical learning**: understand and calculate training errors and test errors, deal with categorical variables in regression models, logistic regression, and model selection by cross-validation.

(12) **Linear model selection and regularization**: understand and apply the shrinkage methods, Ridge regression and LASSO, principal components regression and partial least squares.

(13) **Classification**: solve classification problems using K-Nearest neighbors, logistic regression, linear discriminant analysis, quadratic discriminant analysis and naive Bayes.

(14) **SAS Viya**: Use SAS Viya for visual statistics and data mining.

(15) **Data Analysis for Design of Experiments:** For experiments with a single factor and factorial design, be able to (a) fit, interpret and check the model aptness; (b) test for equality of and conduct multiple comparisons of population means and contrasts; (c) conduct ANOVA, estimation of model parameters, and check model adequacy.

Course Evaluations

Class Attendance	6%
Projects (6)	24%
WeBWork Homework (8 assignments)	16%
Two Midterm Exams (March 20 and April 17)	18%+20%
Final Project (8:00-10:00 AM Friday)	16%

- Attendance will be recorded at all classes using the class attendance sheet. 80% or more is required to get the full 6% credit of the course participation. Please sign the attendance sheet for every class!
- There will be 7 projects and the lowest grade will be dropped, and these projects require you to conduct data analysis using SAS.
- There will be 8 or more online homework assignments posted on WeBWork (https://postulate.klnpa.org/webwork2/Math416_Linear_Models/). Only the **highest 8** assignments will be included when calculating the overall course average. **Unlimited attempts** are allowed for each online homework, and you are not allowed to write the homework once the due date & time have passed (The due dates will be announced as the course progresses). After due date, the online homework score and solutions will be available.
- The two midterm exams are tentatively scheduled as shown above. The material which each exam will cover will be announced in the days before the exam. Each exam will run for about 50 minutes. All exams are timed, close-book and proctored. Communicating with others is not allowed!
- The evaluation of the final project contains two parts: oral presentation (6%), and SAS code and written report (10%).

Grading Scheme

<table>
<tr><td rowspan="3">A = [94, 100]
A- = [90, 93]</td><td>B+ = [87, 89]</td><td>C+ = [77, 79]</td><td rowspan="3">D=[60, 69]</td><td rowspan="3">E = [0, 59]</td></tr>
<tr><td>B = [83, 86]</td><td>C = [73, 76]</td></tr>
<tr><td>B- = [80, 82]</td><td>C- = [70, 72]</td></tr>
</table>

Some Important Dates:

- The final date for withdrawal from the course without academic record is Tuesday January 23, 2024. The last day to withdraw the course is **Monday, April 1, 2024**. After this date students remain registered will receive final grades.
- March break: March 4, 2024 – March 8, 2024

Course Policies:

- **Attendance**: If you miss a class, you are **responsible** for getting any missed information from the recorded lecture or another student.
- Any missed online homework will receive an automatic grade of 0 after the due date and time.
- Students are required to be available for all examinations during the periods for which they are scheduled. **A makeup Quiz/Exam will not be given unless arrangements have been made with the instructor prior** to the Quiz/Exam, and arrangements are granted only in extreme **documented** circumstances.

- Extension of due dates for completion of Homework/Exam or final examinations may be granted to students on the basis of illness, accident or other extreme and unanticipated legitimate circumstances beyond the student's control, with supporting documentation. Supporting documentation will be required and must be submitted before deferrals are approved. For illness or accident, supporting documentation will take the form of:
  - A certificate or letter from the attending physician clearly indicating the start and end dates of the illness and the student's inability to write an examination.

  For other circumstances, students should consult the instructor about acceptable forms of documentation.

Communication Policy: I expect to be able to contact you via the email address listed in your MyESU account, which is your **ESU email** by default. If you haven't maintained this address lately, log on and fix it. Empty your deleted messages folder and your sent messages folder then clean up your inbox messages. No excuses. Trouble with email? Call the help desk @3789 or 570-422-3789 from an outside phone.

IMPORTANT DISCLAIMER: All information in this syllabus is provided as a guide for the course. It is possible that specific items and dates may change based on learning needs, emergencies, illness, school closings, other conditions that may arise, etc. These possible changes are at the discretion of the instructor.

University Policies

Mental Health: As a student you may experience a range of issues that can cause barriers to your learning. These might include strained relationships, anxiety, high levels of stress, alcohol/drug problems, feeling down, or loss of motivation. University Counseling Services can help with these or other issues you may experience. Help is always available. You can learn about free, confidential Counseling and Psychological Services (CAPS) are available to you; call 570-422-3277 or visit website https://www.esu.edu/counseling/index.cfm.

If you, or someone you know, is in Emotional Distress, in a Crisis, or thinking about Suicide, please call the National Suicide Prevention Lifeline number at 988. The Lifeline provides 24-hour, confidential support to anyone in suicidal crisis or emotional distress.

Title IX: East Stroudsburg University and its faculty are committed to assuring a safe and productive educational environment for all students. In order to meet this commitment and to comply with Title IX of the Education Amendments of 1972 and guidance from the Office for Civil Rights, the University requires faculty members to report incidents of sexual violence shared by students to the University's Title IX Coordinator. The only exceptions to the faculty member's reporting obligation are when incidents of sexual violence are communicated by a student during a classroom discussion, in a writing assignment for a class, or as part of a University--approved research project. Faculty members are obligated to report to the person designated in the University protection of minors policy any sexual violence or any other abuse of a student who was, or is, a child (a person under 18 years of age) when the abuse allegedly occurred. Information regarding the reporting of sexual violence and the resources that are available to victims of sexual violence is set forth at: Title IX: Sexual Harassment and Sexual Violence See. http://www4.esu.edu/titleix/

ESU Diversity, Equity and Inclusion Statement (DEI) : East Stroudsburg University recognizes that achieving academic and inclusive excellence requires challenging bias with a focus on equity. ESU is committed to creating a supportive campus climate where all members of the ESU community feel a sense of responsibility to grow and contribute positively to a just, global and diverse society. See https://www.esu.edu/about/history_beliefs/diversity-equity-inclusion.cfm

ESU's DEI Statement was created to be a foundational recognition of the institutional commitment to diversity, equity, inclusion, and social justice. This statement sets a campus-wide purposeful tone for how ESU will navigate issues related DEI and continue its growth and development in this area.

ESU has a variety of resources available for our students who need support or want to engage in activities to promote diversity, equity, and inclusion on campus. We encourage students to utilize the resources available through departments such as the Center of Multicultural Affairs and Inclusive Education, the Gender and Sexuality Center, OASIS, and the Veterans Center. For more information, contact the Office of Campus Life and Inclusive Excellence via Phone: 570-422-3463 or Email: DEI@esu.edu

East Stroudsburg University of Pennsylvania does not discriminate on the basis of race, color, national origin, religion, sex, disability, age, sexual orientation, gender identity or veteran's status in its programs and activities in accordance with applicable federal (Titles VI, VII and IX of Civil Rights Act) and state laws (43 P.S §. 953) and regulations. For more information, visit https://www.esu.edu/about/notices.cfm

It is everyone's responsibility to create an environment where we all feel safe and welcome at ESU. If you experience or witness a bias incident, discrimination or harassment, you are encouraged to complete an incident report. For more information or submit a report visit https://www.esu.edu/diversity/community-restoration-team.cfm

Accommodations: Accessibility and Accommodations "It is the University's goal that learning experiences be as accessible as possible. If you anticipate or experience physical or academic barriers based on a disability, please let me know immediately so that we can discuss options. You are also welcome to contact the Office of Accessible Services Individualized for Students, or OASIS, on campus (Sycamore Suites, Lower Level, 17, 304 Normal Street) at 570-422-3954 (**Fax**: 570-422-3268; **Email**: oasis@esu.edu; Webpage: https://www.esu.edu/oasis/index.cfm) for information concerning the process to request reasonable accommodations." Please be aware that the accessible table and chairs in this room should remain available for students who find that standard classroom seating is not usable.

Academic Integrity: It is expected that students will embrace and practice the principles of academic honesty and integrity (see page. 35 "VI. Statement of Community Standards Academic Misconduct" in the ESU Student Handbook, 2023-2024 for academic misconduct violations). Appropriate collaboration is encouraged, but cheating and plagiarism will result in an E (failing grade) for the test or assignment and the possibility of more severe penalties based upon proceedings facilitated by the ESU Office of Student Conduct and Community Standards. For additional information about the requirements and consequences involved, contact the Office of Student Conduct and Community Standards at https://www.esu.edu/student_conduct/index.cfm or at 570-422-3461.

All students are expected to engage in academically honest work. In an academic environment, everything we do is based on the integrity of the ideas we exchange. Therefore, academic integrity is the foundation of academic work and it's vital that all students uphold this principle-- whether using AI-based tools such as ChatGPT or otherwise. Academic honesty means precisely and accurately identifying facts and accurately reporting the sources of ideas we use in our work. Academic dishonesty is representing someone else's work as your own and can sometimes occur unintentionally when a student fails to understand the proper methods of documentation. You may not submit any work generated by an AI program as your own. If you include material generated by an AI program, it must be cited like any other reference material (with due consideration for the quality of the reference, which may be poor). Let me know if you have any questions about how to document outside sources or visit the Writing Studio in Kemp Library.

Willful academic dishonesty occurs when a student purposefully tries to represent another's ideas as his/her own in order to avoid doing assigned work, for example cutting and pasting material from an internet site, AI-based tools, copying text from a print source, purchasing or stealing a paper, presentation or other work written by someone else, cheating on exams or quizzes, etc. You will find a complete list in the Student Handbook. These standards apply to all work done for this class, not just major projects. All cases of suspected academic dishonesty will be referred to the Office of Student Conduct and Community Standards. I reserve the right to fail any student found responsible for violations of academic integrity.

Student Procedure for Extended Absence Notification: The Dean of Students provides assistance and support for students who miss a minimum of three class days due to illness, personal or family emergencies. A note will be circulated to professors advising of the absence. Students are still required to contact their professors for guidance related to any missed work during the absence.

To utilize this service, the student or family member completes the Request Instructor Notification for Extended Absences Form including supporting documentation related to the absence. In some cases, if a student is unable to return to one or more of their classes, the Dean of Students can discuss enrollment options. Students or family members should contact the office at 570-422-3461 to discuss options. For form, please go to https://www.esu.edu/registrar/extended_absence.cfm

Writing Three or More Exams on the Same Day

This course has an in-class final exam during the regularly scheduled final exam period for this class. If you have a conflict with taking an exam on the date scheduled, you must give your instructor two-weeks' notice, then a mutually agreeable time can be rescheduled. When an alternative testing schedule cannot be agreed upon by the student and faculty, the determination shall be made by Provost. You should email the form (available on https://www.esu.edu/registrar/documents/22-23/Final-Examination-Policy_V2_va.pdf) to the Assistant to the Provost at least three weeks (21 calendar days) prior to the start of final exams.