STAT 3910: Applied Regression Analysis

2022 Spring Session	
Total Class Sessions: 25 **Class Sessions Per Week: 5** **Total Weeks: 5** **Class Session Length (Minutes): 145** **Credit Hours: 4**	**Instructor: Staff** **Classroom: TBA** **Office Hours: TBA** **Language: English**

Course Description:

This course introduces regression analysis techniques through hand-on data analysis, with focuses on the application of linear regression models in practice. The fundamentals of simple linear regression, multiple linear regression and nonlinear regression will be introduced in class, involving the use of standard statistical software. The following topics will be included: statistical review, Regression Inference, Inference on Mean Response and Prediction, ANOVA and General Linear Test, Coefficient of Determination, Residual Graphics and Diagnostics, Residual Diagnostic Tests, Lack of Fit Test, Remedial Measures, Simultaneous Inference, Regression in Matrix Form, Multiple Regression (including dummy variables), Model Selection, Influence, Leverage, and Multicollinearity, Remedial Measures Influence, Leverage, and Multicollinearity, Nonlinear (Logistic) Regression, Logistic Regression, Poisson Regression, and Generalized Linear Models.

Prerequisite: STAT 2110 Applied Statistics is required. MATH 1220 Calculus 2 and MATH 2160 Linear Algebra is recommend.

Learning objectives:

Upon successful completion of this course, students will be able to:

- Develop a thorough understanding of regression analysis techniques and its applications
- Become familiar with the assumption related to different statistical models
- Use statistical methods and build proper regression models to analyze data
- Evaluate the results of analyses and make use of remedial measures
- Validate the modeling assumptions with formal tests and visual diagnostic tools
- Select appropriate predictor variables
- Apply indicator variables in regression models
- Construct and validate regression models
- Apply the knowledge and techniques learned in this course in your own data and problems

Course Materials:

Textbook:

Michael H Kutner, Christopher J. Nachtsheim, John Neter, William Li, *Applied Linear Statistical Models,* 5th Edition, McGraw-Hill Irwin (2004)

Recommended Software:

R(Current Version), RStudio (Desktop, Open Source Edition), SAS or SPASS (Any available Edition)

Course Format and Requirement:

This course is a combination of lectures, in-class discussion, and problem solving through software. Students are expected to finish assigned readings before class and participate in all in-class discussion.

Students are required to prepare a laptop to the class which is capable of connecting to Internet and running statistic software.

Attendance

Attendance at all class meetings is required. You should notify the instructor as far in advance as possible of any class meeting for which you need to be absent, late, or leave early because of illness or other serious extenuating circumstances. You will be evaluated on your participation in class activities. Your participation also includes coming to class prepared to discuss the readings and related material and with all completed assignments due.

Course Assignments and Assessment:

Project

A data-analysis class project will consist of a combination of a written portion and oral presentation. Students will be required to use of the statistical software. Suggested project topics and detailed rubrics will be provided by the instructor.

Quizzes:

There will be seven quizzes administered through the whole semester. Quizzes will always be completed in the beginning of class. The quiz problems will be similar to textbook topics and examples on slides. There will be no make-up quizzes. The lowest one will be dropped.

Exams:

Midterm Exam

There will be two midterm exams in this course. The midterm exam will be based on concepts covered in class. It will be in-class, close-book and non-cumulative.

Final Exam

The final will be cumulative and close-book. Note that the final will not be taken during the normal class times. Exact time and location for final will be announced later.

Attendance	5%
Project	10%
Quizzes	15%
Midterm exam 1	20%
Midterm Exam 2	20%
Final exam	30%
Total	**100%**

Grading Scale (percentage):

A+	A	A-	B+	B	B-	C+	C	C-	D+	D	D-	F
98-100	**93-97**	**90-92**	**88-89**	**83-87**	**80-82**	**78-79**	**73-77**	**70-72**	**68-69**	**63-67**	**60-62**	**<60**

Academic Integrity:

Students are encouraged to study together, and to discuss lecture topics with one another, but all other work should be completed independently.
Students are expected to adhere to the standards of academic honesty and integrity that are described in the Chengdu University of Technology's *Academic Conduct Code.* Any work suspected of violating the standards of the *Academic Conduct Code* will be reported to the Dean's Office. Penalties for violating the *Academic Conduct Code* may include dismissal from the program. All students have an individual responsibility to know and understand the provisions of the *Academic Conduct Code*.

Special Needs or Assistance:

Please contact the Administrative Office immediately if you have a learning disability, a medical issue, or any other type of problem that prevents professors from seeing you have learned the course material. Our goal is to help you learn, not to penalize you for issues which mask your learning.

Course Schedule:

Class	Topics	Assignments

Class 1~5 ➢ Simple Linear Regression - Part 1(Most of Chapter 1, 2, 4)	**Introduction to the course** **Linear Regression with One Predictor Variable** Relations between Variables Regression Models and Their Uses Simple Linear Regression Model with Distribution of Error Terms Unspecified Data for Regression Analysis Estimation of Regression Function Estimation of Error Terms Variance σ^2 Normal Error Regression Model **Inferences in Regression and Correlation Analysis** Inferences Concerning β_1 Inferences Concerning β_0 Some Considerations on Making Inferences Concerning β_1 and β_0 Interval Estimation of $E\{Y_h\}$ Prediction of New Observation Confidence Band for Regression Line Analysis of Variance Approach to Regression Analysis General Linear Test Approach **Simultaneous Inferences and Other Topics in Regression Analysis** Joint Estimation of β_1 and β_0 Simultaneous Estimation of Mean Responses Simultaneous Prediction Intervals for New Observations	• Quiz 1 • Quiz 2

Class 6~9 ➢ Simple Linear Regression - Part 2 (Most of Chapter 3, 5)	**Diagnostics and Remedial Measures** Diagnostics for Predictor Variable Residuals Diagnostics for Residuals Overview of Tests Involving Residuals Correlation Test for Normality Tests for Constancy of Error Variance *F* Test for Lack of Fit Overview of Remedial Measures **Matrix Approach to Simple Linear Regression Analysis** Matrices Matrix Addition and Subtraction Matrix Multiplication Special Types of Matrices Inverse of a Matrix Random Vectors and Matrices Simple Linear Regression Model in Matrix Terms Least Squares Estimation of Regression Parameters Fitted Values and Residuals Analysis of Variance Results Inferences in Regression Analysis	• Quiz 3 • Midterm exam 1

Class 10- 15 Multiple Linear Regression - Part 1 (Most of Chapter 6, 7, 8)	**Multiple Regression I** Multiple Regression Models General Linear Regression Model in Matrix Terms Estimation of Regression Coefficients Fitted Values and Residuals Analysis of Variance Results Inferences about Regression Parameters Estimation of Mean Response and Prediction of New Observation Diagnostics and Remedial Measures **Multiple Regression II** Extra Sums of Squares, Uses of Extra Sums of Squares in Tests for Regression Coefficients Summary of Tests Concerning Regression Coefficients Coefficients of Partial Determination Standardized Multiple Regression Multicollinearity and Its Effects **Regression Models for Quantitative and Qualitative Predictors** Polynomial Regression Models Interaction Regression Models Qualitative Predictors Some Considerations in Using Indicator Variables Modeling Interactions between Quantitative and Qualitative Predictors Comparison of Two or More Regression Functions	• Quiz 4 • Quiz 5

Class 16~20 ➢ Multiple Linear Regression - Part 2 (Most of Chapter 9, 10, 11)	**Building the Regression Model I: Model Selection and Validation** Overview of Model-Building Process Criteria for Model Selection Automatic Search Procedures for Model Selection Model Validation **Building the Regression Model II: Diagnostics** Model Adequacy for a Predictor Variable-Added-Variable Plots Identifying Outlying *Y* Observations- Studentized Deleted Residuals Identifying Outlying X Observations-Hat Matrix Leverage Values Identifying Influential *Cases-DFFITS,* Cook's Distance, and *DFBETAS* Measures Multicollinearity Diagnostics-Variance Inflation Factor **Building the Regression Model III: Remedial Measures** Unequal Error Variances Remedial Measures-Weighted Least Squares Multicollinearity Remedial Measures-Ridge Regression Remedial Measures for Influential Cases-Robust Regression Nonparametric Regression: Lowess Method and Regression Trees Remedial Measures for Evaluating Precision in Nonstandard Situations-Bootstrapping	• Quiz 6 • Midterm exam 2

Class 21~25 ➢ Nonlinear Regression (Chapter 13 and 14)	**Introduction to Nonlinear Regression and Neural Networks** Linear and Nonlinear Regression Models Least Squares Estimation in Nonlinear Regression Model Building and Diagnostics Inferences about Nonlinear Regression Parameters Learning Curve Example Introduction to Neural Network Modeling **Logistic Regression, Poisson Regression, and Generalized Linear Models** Regression Models with Binary Response Variable Sigmoidal Response Functions for Binary Responses Simple Logistic Regression Multiple Logistic Regression Inferences about Regression Parameters Automatic Model Selection Methods Tests for Goodness of Fit Logistic Regression Diagnostics Inferences about Mean Response Polytomous Logistic Regression for Nominal Response Polytomous Logistic Regression for Ordinal Response Poisson Regression **Review For Final Exam**	• Quiz 7 • Project Presentation and Final Delivery • Final exam