King Fahd University of Petroleum and Minerals
Department of Mathematics
MATH-587: Advanced Applied Regression (Term 231)

Instructor: Dr. Muhammad Riaz   **Office:** 5-332
Phone: 013-860-7622   **E-mail:** riazm@kfupm.edu.sa
Office Hours: will be announced later

Course Description: Least square method and properties. Simple and multiple linear regression with matrix approach. Development of liner models. Residual analysis. Polynomial models. Use of dummy variables in multiple linear regression. Analysis of variance approach. Selection of 'best' regression equation. Concepts of mathematical model building. Non-linear regression and estimation. Extensive use of computer packages.

Course Objectives: MATH587 is intended to be an advance level course in applied regression analysis. The emphasis is on understanding how to use regression analysis techniques to solve real-world problems. Upon completion of this course, students should be able to:

- ❖ understand the method of least square estimation and interpret the fitted model
- ❖ analyse the simple and multiple linear regression models in connection with ANOVA
- ❖ validate the assumptions of regression model through residual analysis
- ❖ develop the hypothesis test and confidence intervals for regression models
- ❖ demonstrate the model building techniques and analyze the model adequacy
- ❖ emphasise on the role of polynomial models in regression analysis
- ❖ choose the best regression model using different variable selection techniques
- ❖ highlight the role of indicator variable(s) in regression models
- ❖ discuss the non-linear regression models and their estimation approaches
- ❖ implement various regression techniques on some relevant datasets using R packages

Textbook: *Applied Regression Analysis and Generalized Linear Models* by John Fox, 3rd edition, SAGE Publications 2015.
Software Packages: R Language + RStudio

Reference Books:
- *Applied Linear Statistical Models* by M.H. Kutner, C.J. Nachtsheim, J. Neter and W. Li, 5th edition, McGraw-Hill International 2005.
- *Introduction to Linear Regression Analysis* by Montgomery, Peck and Vinning, 5th edition, Wiley (2012).

Assessment

Activity	Weight
Classwork (quizzes, assignments, attendance, etc.)	15%
Midterm Exam(s)	30%
Project	15%
Final Exam (Comprehensive)	40%

R Language and RStudio: All R commands, procedures and packages will be explained in the class and the student are expected to practice them during and after the class.

Project Description
The project should be based on a real problem (with complete description) and a detailed analysis using the skills developed in the course. All results of the project should be made available numerically with the software/packages used in class. There should be some concluding remarks that refer to the real implications of your chosen problem. You may use online sources in your project with proper citation/reference.

Project Requirements:
- Each group should contain a maximum of 5 students.
- Each group should submit the following:
  - a formal report (pdf)
  - a power point presentation
- Deadline: The end of semester (before the last day of classes)

Weekly Schedule (Tentative)

Week	Topics	Chapter Text Book
Week 1 27/8 – 31/8	**An introductory review** • A review of basic terminologies • A review of fundamental statistical procedures • Introductory session using R and RStudio (libraries and packages in RStudio) • Exporting Excel and text files to RStudio	Chapters 1&3
Week 2 03/9 – 07/9	**Introduction to regression** • An introduction to regression modelling • Exploring R and RStudio including libraries and packages, exporting Excel and text files, performing basic statistical analysis in RStudio	Chapter 2
Week 3 10/9 – 14/9	**Simple linear regression models (SLR)** • Simple linear regression model and its estimation using least-square method • Hypothesis testing and confidence intervals for regression parameters Analysis of variance technique and its role in regression analysis	Chapters 5&8
Week 4 17/9 – 21/9	**Simple linear regression models (SLR) (cont.)** • Prediction of new observations for individual and mean values • Regression through the origin as a special case • Implementation of simple linear regression using RStudio	Chapter 6
Week 5 25/9 – 28/9 (National Day 24/9)	**Multiple linear regression models (MLR)** • Multiple linear regression model and its estimation using least-square method • Individual hypothesis testing and confidence intervals • Prediction of individual and mean values	Chapter 5
Week 6 01/10 – 05/10	**Multiple linear regression models (MLR) (cont.)** • Measures of variation with multiple independent variables • Overall F-test for significance • Partial F-test: Testing for the significance of portions of the model	Chapters 6&8
Week 7 08/10 – 12/10	**Multiple linear regression models (MLR) (cont.)** • Pitfalls and issues in multiple linear regression (hidden extrapolation) • Standardized regression coefficient • Implementation of multiple linear regression using RStudio	Chapters 6&8
Week 8 15/10 – 19/10	**Variable selection and model building** • Approaches to model building • Strategy for variable selection • Implementation of feature selection using RStudio	Chapter22
Week 9 22/10 – 26/10	**Variable selection and model building** • feature selection procedures • computational mechanism for selection of best model • Implementation of feature selection using RStudio	Chapter 22
Week 10 29/10 – 02/11	**Model adequacy checking** • Residual analysis to test assumptions • Lack of fit of the regression model • Detection and treatment of outliers • Implementation of model adequacy using RStudio	Chapter 12
Week 11 05/11 – 09/11	**Diagnostics for leverage and influence** • Detecting influential observations • Measures of influence (cook's D, DFFITS, DFBETAS) • Detection and treatment of influential observations • Implementation of unusual and influential data using RStudio	Chapter 11
Week 12 12/11 – 16/11	**Dummy Variables** • What are dummy variables and when it should be added to regression models? • Models with dummy variables that possess more than 2 levels • Implementation of indicator variables in multiple regression using RStudio	Chapter 7
19/11 – 23/11	**Midterm Break**	
Week 13 26/11 – 30/11	**Polynomial Regression Models** • Polynomial models with one variable • Models in two or more variables • Implementation of polynomial models using RStudio	Chapter 17
Week 14 03/12 – 07/12	**Polynomial and Non-Linear Regression Models** • Non-linear regression models • Estimations of non-linear regression models • Implementation of non-linear regression using RStudio	Chapter 17
Week 15 10/12 – 14/12	**An introduction to some selective topics in Regression analysis (depending on time availability)** • Non-Parametric regression • Logit and Probit models • Time series models • Implementation of each topic using RStudio	Chapters 14, 16, 18
Week 16 17/12 (Normal Thursday)	**Review**	-

Important Notes:

Blackboard: All contacts or announcements between the instructor and the students are supposed to be through Blackboard, so the student must check Blackboard at least once a day.

Academic Integrity: All KFUPM policies regarding ethics and academic honesty apply to this course.

Important Rules
1- Student is not allowed to enter the exam hall without either KFUPM ID cards or Saudi ID/ Iqama ID cards.
2- Students are not allowed to carry mobile phones and smart watches to the exam halls.
3- Students need to strictly adhere to the attendance policy of the university.
4- DN-Grade will be assigned to the eligible students after their instructors have warned them twice.

Cheating in Exams
Cheating or any attempt of cheating by use of illegal activities, techniques and forms of fraud will result in a grade of **DN** in the course along with reporting the incident to the higher university administration for further action. Cheating in exams includes (but is not restricted to):
➢ looking at the papers of other students

➢ talking to other students

➢ using mobiles or any other electronic devices.

Missing an Exam
In case a student misses an exam (Exam I, Exam II, or the Final Exam) for a legitimate reason (such as medical emergencies), he must bring an official excuse from Students Affairs. Otherwise, he will get zero in the missed exam.

Attendance
➢ Students are expected to attend all lecture classes.
➢ If a student misses a class, he is responsible for any announcement made in that class.
➢ Attendance on time is very important. Mostly, attendance will be checked within the first five minutes of the class. Entering the class after that, is considered as one late, and every two times late equals to one absence.
➢ A DN grade will be awarded to any student who accumulates more than 20% unexcused absences (6 lectures) or 33.3% excused and unexcused absences (10 lectures).

The usage of mobile phones and apple watches
• Students are not allowed to use mobiles for any purpose during class time unless given permission.
• Violations of these rules will result in a penalty decided by the instructor.
• Academic Integrity: All KFUPM policies regarding ethics apply to this course. See the Undergraduate Bulletin.