

King Fahd University of Petroleum and Minerals
Department of Mathematics
STAT-513: Statistical Modelling (Term 241)

Instructor: Dr. Nasir Abbas

Phone: 013-860-4485

Office Hours: will be announced later

Office: 5-333

E-mail: nasirabbas@kfupm.edu.sa

Course Description: Statistical tools for learning from the data by doing statistical analysis on the data with an emphasis on the implementation using various software, toolboxes, and libraries like R, Scikit-Learn, and Stats models. Topics include simple and Multiple Linear Regression, Polynomial Regression, Splines, Generalized Additive Models; Hierarchical and Mixed Effects Models; Bayesian Modeling; Logistic Regression, Generalized Linear Models, Discriminant Analysis; Model Selection.

Pre-requisite: STAT 503

Not to be taken for credit with: STAT 413

Course Main Objectives:

- Introduce statistical tools for modeling;
- Develop models that learn from the observed data;
- Implement statistical models based on the statistical analysis.

Course Learning Outcomes:

- Develop statistical models to describe the observed data using computational tools;
- Describe different statistical tools to analyze data;
- Measure the effectiveness of models;
- Present effectively through oral presentation and written reports outcome of the models;
- Interpret the statistical models.

Textbook: A. Agresti. *Foundations of Linear and Generalized Linear Models*, Wiley (2015).

Reference Books:

- *Linear Regression Models - Applications in R* by John P. Hoffman, Chapman & Hall/CRC (2021).
- *Introduction to Linear Regression Analysis* by Montgomery, Peck and Vinning, 6th edition, Wiley (2021).
- *An Introduction to Statistical Learning with Applications in R* by G. James, D. Witten, T. Hastie and R. Tibshirani, 2nd Edition, Springerlink 2021.

Software: R Language and RStudio

Assessment*

Activity	Weight
Classwork (quizzes, assignments, attendance, bonuses, etc.)	15%
Project	20%
Midterm Exam	30%
Final Exam (Comprehensive)	35%

Important Notes:

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

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

Excuse: There will be no make-up for the quizzes, homework, major exams, and the final for students who miss any of them except for those who have an extreme case (they need to provide an official excuse issued by Deanship of Student Affairs).

Cheating and Plagiarism: This course is composed of individual assignments. It is important that your individual assignment be completed with your own efforts instead of copying it from your fellow student. KFUPM instructors follow “*zero tolerance*” approach with regard to cheating and plagiarism. During examinations (quizzes, major exams, lab tests) cheating or any attempt of cheating by use of illegal activities, techniques and forms of fraud will result in a *grade of F* in the course along with reporting the incident to the higher university administration.

Mobiles: The use of mobile is *strictly banned* during the class. Students are required to keep their phones off/silent and placed inside their pockets during the class timings.

Project: Project Guidelines will be uploaded to Blackboard, and guidelines for the report will also be posted there. The instructor will form project groups.

Letter Grades: The letter grades will follow a grading curve, which depends on the average of all students enrolled in the course.

Attendance Notes:

- If a student misses a class, he/she is responsible for any announcement made in that class/lab.
- 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% (6 classes) unexcused absences or more than 33% (10 classes) excused and unexcused absences.

Student Responsibilities:

- You will be encouraged to participate in the class.
- Keep up with the material presented in class. If you get behind it, it will not be easy to recover.
- Submit assignments on time.
- You are expected to attend all lectures on time.
- Do not leave before the end of class. Attendance is marked as “present” only if you are in class for 80% of the duration of the class.
- Teamwork will be encouraged for Project-related activities.

No	List of Topics	Contact hours
1	Statistical Learning <ul style="list-style-type: none"> • What is Statistical Learning • Descriptive Analysis • Probability and Sampling Distributions • Inferential Analysis • Introduction to R language 	3
2	Simple Linear Regression <ul style="list-style-type: none"> • Estimating the Coefficients • Assessing the Accuracy of the Coefficient Estimates • Assessing the Accuracy of the Model 	3
3	Multiple Linear Regression <ul style="list-style-type: none"> • Estimating the Regression Coefficients • Some Important Questions • Other Considerations in the Regression Model • Qualitative Predictors • Extensions of the Linear Model • Model fitting using R 	6
4	Polynomial Regression <ul style="list-style-type: none"> • Introduction to Polynomial Regression • Step Functions • Basis Functions • Polynomial Regression Fitting using R 	3
5	Linear and Non-Linear Spline Fitting <ul style="list-style-type: none"> • Regression Splines • Piecewise Polynomials • Constraints and Splines • Smoothing Splines • Spline Fitting using R 	3
6	LOGIT and PROBIT models <ul style="list-style-type: none"> • Models with a Binary Response Variable • Logistic Regression Models • Probit Regression Models • Estimating the Parameters in a Logistic Regression Model • Interpretation of the Parameters in a Logistic Regression Model • Fitting LOGIT and PROBIT Models using R 	6

7	Poisson Regression <ul style="list-style-type: none"> • Introduction to Poisson Regression • Application of Poisson Regression using R 	3
8	Generalized Linear Models <ul style="list-style-type: none"> • Link Functions and Linear Predictors • Parameter Estimation and Inference in the GLM • Prediction and Estimation with the GLM • Residual Analysis in the GLM • Using R to Perform GLM Analysis 	6
9	Model Building Techniques <ul style="list-style-type: none"> • Best Subset Selection • Stepwise Selection • Choosing the Optimal Model • Model Selection Using R 	6
10	Bayesian Modeling <ul style="list-style-type: none"> • Testing and Comparing Models • Making the Most Accurate Predictions • Interpreting Results Quickly & Effectively 	3
11	Non-Linear Regression <ul style="list-style-type: none"> • Linear and Nonlinear Regression Models • Origins of Nonlinear Models • Nonlinear Least Squares • Transformation to a Linear Model • Parameter Estimation in a Nonlinear System • Statistical Inference in Nonlinear Regression • Examples of Nonlinear Regression Models Using R 	3
Total		45