King Fahd University of Petroleum and Minerals
Department of Mathematics
STAT-513: Statistical Modelling (Term 231)

Instructor: Dr. Nasir Abbas   **Office:** 5-333
Phone: 013-860-4485   **E-mail:** nasirabbas@kfupm.edu.sa
Office Hours: will be announced later

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).

Software: R

Reference Books:
- *An Introduction to Statistical Learning with Applications in R* by G. James, D. Witten, T. Hastie and R. Tibshirani, 2nd Edition, Springerlink 2021.
- M.H. Kutner, C.J. Nachtsheim, J. Neter and W. Li (2005). *Applied Linear Statistical Models*. 5th edition, McGraw-Hill International.
- *Introduction to Linear Regression Analysis* by Montgomery, Peck and Vinning, 6th edition, Wiley (2021).

Assessment*

Activity	Weight
Classwork (quizzes, assignments, attendance, bonuses, etc.)	10%
Midterm Exam I	20%
Midterm Exam II	20%
Project	20%
Final Exam (Comprehensive)	30%

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.

Project Description:
The project should be targeted at statistical modeling using parametric/non-parametric estimates. Classical regression analysis needs to be extended to more advanced modeling areas e.g. splines, generalized linear models. logistic regression, mixed effect models, etc. All results of the project should be made available numerically with R packages.

Attendance Notes:
➢ In accordance with University rules, 20% unexcused absences will automatically result in a grade of DN.
➢ 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.

No	List of Topics	Contact hours
1	Statistical Learning • What is Statistical Learning • Descriptive Analysis • Probability and Sampling Distributions • Inferential Analysis • Introduction to R language	3
2	Simple Linear Regression • Estimating the Coefficients • Assessing the Accuracy of the Coefficient Estimates • Assessing the Accuracy of the Model	3
3	Multiple Linear Regression • 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 • Introduction to Polynomial Regression • Step Functions • Basis Functions • Polynomial Regression Fitting using R	3
5	Linear and Non-Linear Spline Fitting • Regression Splines • Piecewise Polynomials • Constraints and Splines • Smoothing Splines • Spline Fitting using R	3

6	**LOGIT and PROBIT models** • 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** • Introduction to Poisson Regression • Application of Poisson Regression using R	3
8	**Generalized Linear Models** • 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** • Best Subset Selection • Stepwise Selection • Choosing the Optimal Model • Model Selection Using R	6
10	**Bayesian Modeling** • Testing and Comparing Models • Making the Most Accurate Predictions • Interpreting Results Quickly & Effectively	3
11	**Non-Linear Regression** • 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