King Fahd University of Petroleum and Minerals
Department of Mathematics & Statistics
MATH 506 Syllabus, Term 241

Code: MATH 506
Title: Fundamentals of Data Science
Credit Hours: 3-0-3
Prerequisite: Graduate Standing

Instructor: Dr. Jamal Al-Smail **E-mail:** jamalhas@kfupm.edu.sa
Office Hours:
Sundays & Tuesdays: 3:00 pm – 4:50 pm
Sundays & Tuesdays: 7:40 pm – 8:10 pm
Office: Building 5-407

Course Objectives: The main objective of the course is to

- Introduce the mechanism of the learning process,
- Implement solutions using data scientific software, toolboxes, and libraries.

Description: All aspects of the data science pipeline using the software, toolboxes, and libraries like NumPy, SciPy, Pandas, SymPy, Matplotlib, and Seaborn: Data acquisition, cleaning, handling missing data, EDA, visualization, feature engineering, modeling, model evaluation, bias-variance tradeoff, sampling, training, testing, experimenting with a classical model.

Learning Outcomes: Upon completion of the course, students should be able to:

- Distinguish data science tasks
- Prepare data for analysis
- Describe the learning process
- Build a model in a computer environment

Textbook [TB]: Data Science using Python and R by C. Larose and D. Larose, Wiley, 2019.

Supplementary Material:

1. [R1] A Hands-On Introduction to Data Science, by Chirag Shah, Cambridge University Press
2. [R2] Introduction to Data Science: A Python Approach to Concepts, Techniques and Applications by Igual, Laura, Seguí, Santi, Springer

Grading Policy:
Group Class Activities (10%), Group Assignments (10%)
Team Projects and Poster Sessions (15%),
Two IBM Certifications and Presentations (10%),
Attendance (5%), Exam1 (10%), Exam2 (10%), Final Exam (30%)

Attendance: Attendance is a University Requirement. A DN grade is rewarded after accumulating 6 unexcused absences.
Academic Integrity: All KFUPM policies regarding ethics apply to this course.

Course Outline:

Weeks	Topics and Class Activities	References
1	**Introduction to Data Science** Data Science Methodologies and Tasks **Active Learning Classes:** Factors for Successful Team Projects Data Science Applications [Group Search and Presentations]	Ch 1 [TB], [R1],[R2]
2-3	**Toolboxes for Data Scientists** Python Introduction and Data Acquisition **Active Learning Classes:** Hands-on Python Training	Ch 2 [TB] Ch 5.1-5.3 [R1] Ch 2.1-2.6 [R1]
4-5	**Data Preparation** Types, Sources, Formats, Pre-Processing **Active Learning Classes:** Hands-on Python Training Statistical Topics [Group Search and Presentations]	Ch 2 [R1] Ch 3 [TB]
6-7	**Data Analysis Techniques** Descriptive, Multivariate Analysis, Feature Engineering **Active Learning Classes:** Hands-on Python Training Statistical Topics [Group Search and Presentations]	Ch 4 [TB] Ch 3 [R1] Ch 3 [R2]
8	**Data Visualization** **Active Learning Classes:** Hands-on Python Training Advanced Data-Visualization [Group Search and Presentations]	External Notes
9-10-11	**Introduction to Modeling** Datasets, Machine Learning, Modeling, Training-Testing-Validation, Regression, Classification **Active Learning Classes:** Hands-on Python Training Data Science Applications using IBM SPSS Statistics Data Science Applications using IBM SPSS Modeler	Ch 5, Ch 11[TB] Ch 8.1 – 8.3, Ch 9.4 [R1] Ch 6.1 [R2]
12-13	**Evaluating Models** Metrics, Cross-Validation, Hyperparameters **Active Learning Classes:** Evaluation Metrics [Group Search and Presentations] Modeling using Python and IBM SPSS Statistics	Ch 7 [TB] Ch 12.4 [R1]
14	**Automating Models** Building Pipelines, Joining Pipelines, Saving Models **Active Learning Classes:** Building Pipelines using Python Building Pipelines using IBM SPSS Statistics Building Pipelines using IBM SPSS Modeler	External Notes
15	**Project Presentations**	

Important Dates:

- **Exam1**: Week 6 ; **Exam2**: Week 12
- **Data Science Project Proposal**: Week 7
- **IBM Certification-1:** Week 8
- **IBM Certification-2:** Week 10
- **Project Report/Notebook Submission:** Week 14
- **Project Presentations:** Week 15
- **Final Exam**: Posted on the registrar's website