PET 328E GEO-ENERGY DATA ANALYTICS

Course Title		Geo-Energy Data Analytics				
				Course Implementation, Hours/Week		
Code	**Semester**	**Local Credits**	**ECTS Credits**	**Theoretical**	**Tutorial**	**Laboratory**
PET 328E	8	3	6	3	0	0
Department		Petroleum and Natural Gas Engineering				
Course Type		Elective		**Course Language**		English
Course Prerequisites						

Course Category By Content, %	Math & Basic Sciences	Engineering Topics; Check if Contains Significant Design	Other
	-	100	-

Course Description	Overview of data science concepts. Data types for subsurface energy resources. Basic principles of descriptive and inferential statistics. Exploratory data analysis and data mining as applied to subsurface data types. Data visualization. Introduction to supervised and unsupervised machine learning. Applications with modern computational tools and packages. Case studies for oil, natural gas and geothermal engineering.
Course Objectives	1. Familiarize students with subsurface data types collected in oil, natural gas and geothermal engineering 2. Develop students' ability to apply exploratory data analysis, data mining concepts to subsurface data with appropriate visualization and analysis techniques 3. Develop students' ability to deal with large data sets through the use of modern computational tools and packages 4. Introduce students supervised and unsupervised machine learning algorithms for the development of prediction and classification models for subsurface data
Course Learning Outcomes	Students who pass the course will be able to: 1. Define and classify different types of data collected for subsurface energy resources 2. Clean and process subsurface data using modern statistical computational packages for further analysis and visualization 3. Perform exploratory data analysis by creating and interpreting visual representations and statistical summaries of data related to subsurface energy resources 4. Design and train machine learning models for prediction and classification using supervised and unsupervised algorithms 5. Report data analytics and machine learning projects in an organized way and in reproducible formats
Textbook	1. Datta-Gupta, A., Mishra, S. (2017). *Applied Statistical Modeling and Data Analytics: A Practical Guide for the Petroleum Geosciences*, Elsevier.
Other References	1. James, G., Witten, D., Hastie, T., Tibshirani, R. (2021). *An Introduction to Statistical Learning: With Applications in R*, 2nd Edition, Springer. 2. Tang, P., Steinbach, M., Karpatne, A., Kumar, V. (2018). *Introduction to Data Mining*, 2nd Edition, Pearson. 3. Iliinsky, N., Steele, J. (2011) *Designing Data Visualizations*, O'Reilly.
Homework & Projects	-
Laboratory work	-
Computer Use	-

Other Activities	-		
Assessment Criteria	**Activities**	**Quantity**	**Effects on Grading, %**
	Midterms	1	30
	Quizzes	-	-
	Homework	5	30
	Projects	-	-
	Term Paper/Projects	-	-
	Laboratory Work	-	-
	Other Activities	-	-
	Final Exam	1	40

Weeks	**Course Plan (lectures)**	**Course Outcomes**
1	Introduction to data science and analytics for subsurface energy sources	1
2	Subsurface data sources, types, scales (seismic, logs, production, pressure, PVT)	1
3	Review of descriptive and inferential statistics concepts	3
4	Data wrangling: filtering/grouping and dealing with missing values	2
5	Data wrangling: statistical summaries	2
6	Data visualization: Numerical/categorical variables (reservoir / time-series data)	3
7	Data visualization: Text analytics (drilling reports)	3
8	Multivariate data analysis: applications to subsurface data	3
9	Analytics reporting: reproducibility, version control, reporting	5
10	Introduction to machine learning	4
11	Supervised learning - prediction problems	4
12	Supervised learning - classification problems	4
13	Unsupervised learning	4
14	Term project presentations and discussions	5

Prepared by	**Date**
Assoc. Prof. Dr. Emre Artun	01/05/2022