The University of Jordan

School of Engineering

Computer Engineering Department

Fall Term 2021/2022

Course	Big Data Analysis – 0917546 (3 Cr. – Elective Course)
Catalog Description	Definitions and applications; Market trends; Data analytics lifecycle; Data exploration and preprocessing; Data visualization; Theory, tools and methods; Introduction to Big data management, warehousing and processing; Ethical issues.
Prerequisites by Course	Introduction to Database Systems (1902223)
Prerequisites by Topic	Students are assumed to have good background in mathematics, particularly, calculus, linear algebra, and statistics. Additionally, the students should have good programming skills using Python.
Textbooks	1. Arshdeep Bahga and Vijay Madisetti, Big Data Analytics: A Hands-On Approach, 2019. 2. Wes McKinney, Python for Data Analysis: Data Wrangling with Pandas, NumPy, and Ipython, O'Reilly Media, 2nd Edition, 2018. 3. Course web page: http://www.abandah.com/gheith/?page_id=2759
References	1. Jake VanderPlas, A Whirlwind Tour of Python, O'Reilly Media, 2016. 2. Joel Gurs, Data Science from Scratch, O'Reilly Media, 2015. 3. Aurélien Géron, Hands-On Machine Learning with Scikit-Learn, Keras and TensorFlow: Concepts: Tools, and Techniques to Build Intelligent Systems, 2nd Edition, O'Reilly Media, Oct 2019.
Course Website	http://www.abandah.com/gheith/?page_id=2759
Microsoft Team	Link
Schedule & Duration	15 weeks; 45 lectures, 60 minutes each (including exams)
Student Material	Textbook, class handouts, some instructor keynotes, selected YouTube videos, and access to a personal computer and the internet.
College Facilities	Classroom with whiteboard and projection display facilities with speakers, library, and computer laboratory.
Course Objectives	The objectives of this course are: 1. Introduce students to the basic concepts and techniques in big data. 2. Introduce students to the practical techniques used in data analytics including loading, cleaning, preparation, wrangling, visualization, and analysis.
Course Outcomes and Relation to ABET Program Outcomes	Upon successful completion of this course, a student should be able to: 1. Know the main concepts and techniques used in handling big data and performing data analytics [1]. 2. Use Python and its specialized libraries to gain insight from data [2].

Course Topics

No	Topic	Hours	Reference
1.1	Course Introduction	1.5	3
1.2	Introduction to Python, Anaconda and PyCharm	2*	3
1.3	Python Basics, IPython and Jupyter Notebooks	3*	3
1.4	Built-in Data Structures, Functions and Files	3*	3
1.5	Object Oriented Programming in Python	2*	3
1.6	NumPy Basics: Arrays and Vectorized Computation	6*	3
1.7	Pandas Data Structures, Essential Functionality & Descriptive Statistics	6	3
2.1	Introduction to Big Data	3	1
2.2	Big Data Architectures and Patterns	3	1
2.3	MapReduce Patterns	3	1
3.1	Data Loading, Storage and File Formats	6	2
3.2	Data Cleaning and Preparation	6	2
3.3	Data Wrangling: Join, Combine and Reshape	3	2
3.4	Plotting and Visualization with Matplotlib and Seaborn	6	2
3.5	Data Aggregation and Group Operations	3	2
3.6	Time Series	3	2

* Offline review

Computer Usage   Practical aspects of the course are covered in class and through the assignments.

Important Dates

Date	Event
Mon 11 Oct, 2021	First Lecture
Wed 1 Dec, 2021	Midterm Exam
Wed 12 Jan, 2022	Last Lecture
Jan 18 – 30, 2022	Final Exam Period

Policies
- Attendance is required. Class attendance will be taken every class and the university polices will be enforced in this regard.
- All submitted work must be yours
- Cheating will not be tolerated
- Open-book exams
- Covid-19 vaccination or valid PCR test is required to join the classes
- Check department announcements at: http://www.facebook.com/pages/Computer-Engineering-Department/369639656466107 for general department announcements.

Assessments   Reports, participation, and exams

Grading policy
Assignments and quizzes   20%
Midterm Exam   30%
Final Exam   50%

Instructors   **Prof. Gheith Abandah**

Email: abandah@ju.edu.jo
Homepage: http://www.abandah.com/gheith
Office Hours: Sun through Wed: 12:30 am – 1:30 pm

Time and Location   Section 1:   Mon and Wed: 10:00–11:30, CPE 001, Microsoft Teams

Last Updated   Oct 7, 2021

Program Outcomes (PO)

1	an ability to identify, formulate, and solve complex engineering problems by applying principles of engineering, science, and mathematics
2	an ability to apply engineering design to produce solutions that meet specified needs with consideration of public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors
3	an ability to communicate effectively with a range of audiences
4	an ability to recognize ethical and professional responsibilities in engineering situations and make informed judgments, which must consider the impact of engineering solutions in global, economic, environmental, and societal contexts
5	an ability to function effectively on a team whose members together provide leadership, create a collaborative and inclusive environment, establish goals, plan tasks, and meet objectives
6	an ability to develop and conduct appropriate experimentation, analyze and interpret data, and use engineering judgment to draw conclusions
7	an ability to acquire and apply new knowledge as needed, using appropriate learning strategies.