**Course Abstract**

***If you need accommodations due to a disability, contact Disability Services in***

***Edison Hall Room 100, 732.906.2546.***

***To foster a productive learning environment, the College requires that all students adhere to the Code of Student Conduct which is published in the college catalog and website.***

**Course ID and Name: DSA 220 – Big Data Fundamentals**

**Department: Business and Computer Science**

Chairperson or Course Coordinator: Dr. Aslihan Cakmak

Office Location: ED123

E-mail Address: [ACakmak@Middlesexcc.edu](mailto:ACakmak@Middlesexcc.edu)

Telephone: 732-906-2526

**Prerequisites:** DSA 120

**Co-requisites:** MAT 285

**Course Description:**

Providing an overview of Big Data, the processes, tools, and techniques used to work with it, this course will introduce students to the methods used to collect, integrate, and extract large volumes of data from multiple file formats and database platforms. Students will explore common data analysis tools, software, and techniques for processing large volumes of data will be studied. Real-time and batch processing techniques for data aggregation and presentation will also be presented. Large data file software tools will be integrated throughout the course.

**General Education Status:** N/A

**Credits: 4 Lecture Hours: 2 Lab Hours: 4**

**Learning Outcomes:**

1. Describe the characteristics and challenges of structured and unstructured big data
2. Identify the existing big data processing platforms/ tools
3. Explain big data collection methods, integration and storage formats
4. Demonstrate big data analysis techniques
5. Explain common data analysis techniques
6. Identify core big data processing techniques
7. Build queries using big data
8. Generate graphs to demonstrate foal points
9. Explain how business intelligence drives the methods used to process big data

**Upon successful completion of this course, a student will be able to:**

1. Describe the characteristics and challenges of structured and unstructured big data
2. Identify the existing big data processing platforms/ tools
3. Explain big data collection methods, integration and storage formats
4. Demonstrate big data analysis techniques
5. Explain common data analysis techniques
6. Identify core big data processing techniques
7. Build queries using big data
8. Generate graphs to demonstrate foal points
9. Explain how business intelligence drives the methods used to process big data

**Course Content Areas:**

1. Structured and unstructured big data characteristics and challenges
2. Big data processing platforms and programming tools
3. Big data collection and extraction
4. Big data integration
5. Big data business intelligence
6. Big data storage devices, file systems, and databases
7. Big data analysis techniques
8. Real-time and batch processing
9. Big data-parallel processing and distributed processing
10. Big data query techniques
11. Large scale graph processing techniques
12. Big data stream techniques and algorithms
13. Big data visualization tools and techniques
14. Big data privacy