



Korea University International Summer Campus (KU ISC) 2024

Embark on a unique summer

June 26, 2024 - August 1, 2024

ISC356B – Big Data Analytics

I . Instructor

Professor	:	Q B. Chung, Ph.D.
E-mail	:	q.chung@villanova.edu
Home Institution	:	Villanova University
Class Time	:	Period 3 (13:10 - 14:50 KST)
Class Format	:	Blended
Office	:	TBA
Office Hours	:	By appointment

II. Textbook

Required Textbook	:	Minelli, M., Chambers, M. & Dhiraj, A. (2013). Big Data Big Analytics . Wiley CIO Series. Hoboken, NJ: John Wiley & Sons, Inc. (ISBN: 978-1-118-14760-3)
Recommended Additional Readings	:	Additional reading materials will be available on Blackboard.

III. Course Description and Objectives

Typically characterized by volume, velocity, and variety – dubbed the *three V's of big data* – the big data phenomenon is undeniably real and has been growing in enthusiasm, broad attention, corporate spending, and even new educational program launches. No matter how unscientific 'big data' may sound, the term has now gained sufficient respect from both academics and practitioners, granting legitimacy to a rapidly increasing number of big data initiatives.

For big data, just like any other enabling technologies, the volatile process of setting standards in hardware, software, and service provision will see many giants rise and fall, and once the stability of a steady state is obtained – however brief it may be – all technologies surrounding big data are bound to return to their 'true' calling of analytics.

This course will examine big data from both theoretical and practical perspectives by blending lecture, quick-and-dirty research, group discussion and presentation, and hands-on exercises.

By completing this course, students will be able to:

- Define and characterize the big data phenomenon;
- Describe the *three V's of big data*;

- Recognize the role of data played in decision making;
- Discuss big data practices and trends in various industries;
- Identify and describe the technology and techniques behind the big data movement;
- Identify and describe Apache Hadoop and the related software platforms;
- Discuss managerial issues of big data deployment;
- Install and use MySQL;
- Install and use RapidMiner;
- Construct SQL queries for structured data using MySQL;
- Perform predictive analytics using RapidMiner for text mining; and
- Make presentations using bootstrap research on technical topics.

IV. Grading

Participation	:	10%
Assignments	:	20%
Group Presentations	:	20%
Midterm Exam	:	20%
Final Exam	:	30%

V. Class Outline

Date	Topic	Chapter	Remarks
June 26 (Wed)	Orientation Day (No Class)		
June 27 (Thu)	Introduction, Group formation, Course overview		
June 28 (Fri)	What is Big Data?	Ch.1	
July 1 (Mon)	Data and Decisions	Ch.1	
July 2 (Tue)	Basic Concepts of Big and Traditional Data		Handout
July 3 (Wed)	Assignment #1 Group Session		Handout
July 4 (Thu)	Group Presentations on Basic Concepts		
July 5 (Fri)	Big Data Industry Snapshot	Ch.2	
July 8 (Mon)	Apache Hadoop and the Evolving Hadoop Ecosystem	Ch.4	
July 9 (Tue)	Big Data Analytics Technologies and Techniques	Chs.3,5	
July 10 (Wed)	Moving Forward with Big Data	Ch.6	
July 11 (Thu)	Midterm Exam		
July 15 (Mon)	Big Data & Data Ethics	Ch.7	
July 16 (Tue)	Big Data Analytics in Action with Structured Data (1)		Handout
July 17 (Wed)	Big Data Analytics in Action with Structured Data (2)		Handout
July 18 (Thu)	Big Data Analytics in Action with Structured Data (3)		Handout
July 22 (Mon)	Big Data Analytics in Action with Unstructured Data (1)		Handout
July 23 (Tue)	Big Data Analytics in Action with Unstructured Data (2)		Handout
July 24 (Wed)	Big Data Analytics in Action with Unstructured Data (3)		Handout
July 25 (Thu)	Big Data Analytics in Action with Unstructured Data (4)		Handout
July 29 (Mon)	Assignment #2 Group Session		Handout
July 30 (Tue)	Final Exam		
July 31 (Wed)	Review & Advising Session		
Aug 1 (Thu)	Graduation Day		