



HANYANG UNIVERSITY

INTERNATIONAL SUMMER SCHOOL

* Please fill out the form completely in English in detail.

Name	Jongwook Woo
E-mail	jongwook@alumni.usc.edu
Home University	California State University Los Angeles
Department	Computer Information Systems

Course Title	INTRODUCTION TO BIG DATA ANALYSIS
Field of Study	Big Data, Data Analysis, Large Scale Data, Cloud Computing, Hive
Credits	3
Contact Hours	45
Course Code/Number	* In case it was opened at Hanyang University previously
Course Description	Big Data is a solution to process and store large scale data. In this course, the students understand the practical knowledge and architecture of Big Data, identify how to process and store large scale data using Big Data solution with MapReduce, HDFS, Yarn, and Hive, and develop hands-on examples in Big Data Analysis.
Course Objective	<ul style="list-style-type: none">• Understand the genesis of Big Data Systems• Understand practical knowledge of Big Data Analysis, Text Processing, Sentiment Analysis using Hive• Provide the student with a detailed understanding of effective behavioral and technical techniques in Cloud Computing on Big Data• Demonstrate knowledge of Big Data in industry and its Architecture• Learn data analysis, modeling and visualization in Big Data systems
Preparations (Pre-Knowledge)	<ol style="list-style-type: none">a. Mastery over MS-Windows File Management (Windows Explorer) facilities.b. Students are expected to attend every class session. Since Cloud computing and Big Data concepts are presented during class time, class attendance is essential for successful completion of assignments and tests.c. As a large part of the course involves work on cloud computing, it is essential that you utilize the time in class for discussion and exercises on the computer. If attendance is not possible for one of the class meetings, please contact the instructor beforehand.



HANYANG UNIVERSITY

INTERNATIONAL SUMMER SCHOOL

	d. Students are expected to use the equipment of the computer labs on campus if you do not have a personal computer nor internet.
Materials (Textbook/Websites link)	<ol style="list-style-type: none"> 1. Instructional materials (Lecture and Lab) from the instructor 2. Hadoop: The Definitive Guide by Tom White 3. https://hadoop.apache.org/ 4. https://www.cloudera.com/tutorials.html

Lesson Plan: Fill out the topic for each class in detail		
Week 1	1st Day	Orientation & Opening Ceremony
	Class 1	Lecture 1 An Introduction to Big Data and Cloud Computing Systems
	Class 2	Lecture 2 Big Data system development <ol style="list-style-type: none"> a. Introduction to Hadoop b. Motivation for Hadoop
	Class 3	Lab 1: set up cloud computing accounts such as Big Data Compute Edition
Week 2	Class 4	Lecture 3 Basic Concepts: HDFS, MapReduce, Hive and Lab
	Class 5	Lab 2: HDFS and Hive in Oracle Big Data
	Class 6	Lecture 4 Initiating Phase: Basic Concepts: MR cont'd; Cluster; Ecosystems, Hive
	Class 7	Midterm Exam
Week 3	Class 8	Lab 3: Hive Web Log Analysis in Oracle Big Data
	Class 9	Lecture 5 Hive Data Processing (Join, Union)
	Class 10	Lab 4: IoT Sensor Data Analysis using Hive in Oracle Big Data
	Class 11	Lecture 6 Text Analysis in Hive
Week 4	Class 12	Lab 5: Twitter Data Text Analysis using Hive in Oracle Big Data
	Class 13	Lecture 7 NGram, Text Processing Functions in Hive
	Class 14	Lecture 8 Big Data Trend with Data Science
	Class 15	Final Exam



HANYANG UNIVERSITY

INTERNATIONAL SUMMER SCHOOL

Evaluation (%)

* Total sum of percentages should be 100%

* Only below options are available, please do not change the form (fill out the given form)

Assignments	Attendance	Final	Group Project	Mid-term	Participation	Presentation	Quiz	Total
30	10	35	0	25	0	0	0	100 %