

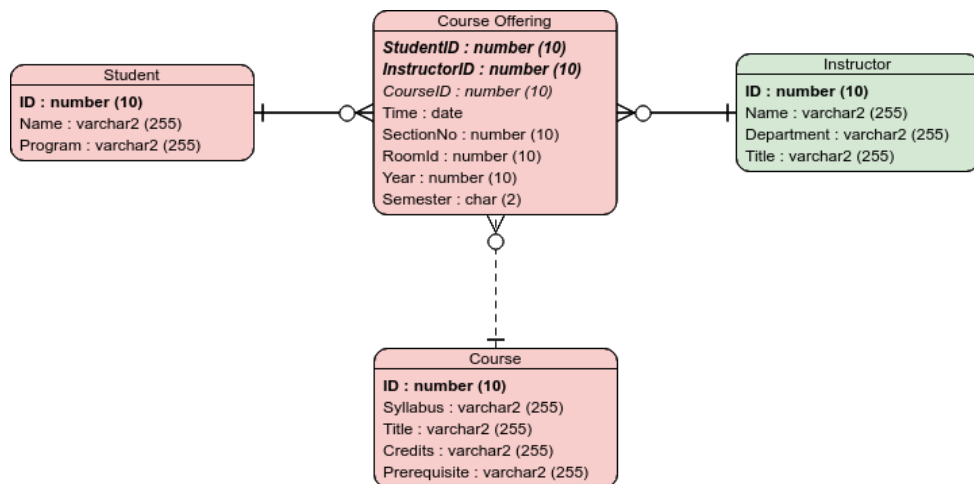
BINUS University

Academic Career: <i>Undergraduate / Master / Doctoral *)</i>	Class Program: <i>International/Regular/Smart Program/Global Class*)</i>						
<input checked="" type="checkbox"/> Mid Exam <input type="checkbox"/> Final Exam <input type="checkbox"/> Short Term Exam <input type="checkbox"/> Others Exam : _____	Term : Odd/Even/Short *)						
<input checked="" type="checkbox"/> Kemanggisan <input checked="" type="checkbox"/> Alam Sutera <input checked="" type="checkbox"/> Bekasi <input type="checkbox"/> Senayan <input type="checkbox"/> Bandung <input type="checkbox"/> Malang	Academic Year : 2020 / 2021						
Faculty / Dept. : School of Computer Science	<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <td style="width: 10%;">Deadline</td> <td style="width: 10%;">Day / Date</td> <td style="width: 80%;">Friday/ Nov 13th, 2020</td> </tr> <tr> <td></td> <td>Time</td> <td>13:00</td> </tr> </table>	Deadline	Day / Date	Friday/ Nov 13 th , 2020		Time	13:00
Deadline	Day / Date	Friday/ Nov 13 th , 2020					
	Time	13:00					
Code - Course : COMP6579 - Big Data Processing	Class : All Classes						
Lecturer : Team	Exam Type : Online						
*) <i>Strikethrough the unnecessary items</i>							
<i>The penalty for CHEATING is DROP OUT!!!</i>							

1. **(10%)** Big Data has been successfully applied in many different domains such as Web, Financial, Healthcare, Environment, Logistic and Transportation, Industry, and Retail. Explore the possibilities Big Data implementation in Higher Education, give a use case of Big Data implementation in a University.
2. **(10%)** There are four types of analytics in Big Data i.e. Descriptive, Diagnostic, Predictive, and Prescriptive analysis. What do you know about those type of analytics, give an example for each!
3. **(10%)** Distributed Computing and Distributed File System are the core technologies behind Big Data. Explain how Hadoop Distributed File System (HDFS) works in the Big Data Cluster.
4. **(15%)** There are two programming models in Big Data Ecosystem i.e. Map Reduce and Spark, what are the similarities and differences between them? Why Spark outperform Map Reduce in the execution time?
5. **(10%)** Cloud computing already became the most popular infrastructure for data center, including for Big Data Cluster. We could prepare a Big Data Cluster with required number of nodes just in several minutes. Explain what are the benefits and drawbacks of Big Data in the cloud?
6. **(20%)** One of the benefits from Graph Database is significantly faster than traditional Relational Database System (RDBMS), especially for a database with many relations. Create a graph database of the following RDBMS tables with 3 records for each table. The Entity Relationship Diagram represents an Online Course Database. Use a graph diagram that consists of vertices and edges to represents the graph database. Explain why a database query would be executed significantly faster in the Graph Database than the RDBMS based on the example database.

Verified by,

Kristien Margi Suryaningrum - D6414 and sent to Program on Oct 20, 2020



Case study:

(25%) Suggest a Big Data Architecture for a company that provides a flight information service to their customers (for example <https://flightstats.com/>). Read more about flightstats here:

- <https://onemileatatime.com/best-website-for-flight-status/>
- <https://www.dropbox.com/s/yqf2u4nmi5sbya2/Session%2001-03%20Example%20-%20Flight%20data%20management.pdf?dl=0>

Verified by,

Kristien Margi Suryaningrum - D6414 and sent to Program on Oct 20, 2020