

CAPSTONE PROJECT

Description

This is a group project. You will form a group of two to three members, and pursue an independent research and analysis on some big data of your choice.

The goal of the project is to give you the opportunity to take what you've learned throughout the class and apply it to solve a real world problem.

Choose your subject carefully and wisely. It should b

1. Be based on unstructured or semi-structured data
2. Be attainable within allotted timeframe.
3. Use technologies we have covered in the course, i.e., Java, Python, MapReduce and Spark.

Deliverables

1. An introduction to the project and discussion of the approach you take to perform data capturing and analysis.
2. Source code, make file, and output
3. A summary of the working experience – where you've got stuck, how you resolved it, the limitation of the system used, lessons learned.

Presentation

In the last meeting of the class, your group will be presenting to the class your project, and take Q&A.

Grading

The grading will be calculated, based on:

1. Self-evaluation, 33%
2. Peer reviews (by three other groups), 33%
3. The professor, 33%