

## The Project Overview for ITM 511

### ➤ Description

The projects will be conducted by team projects. The projects consist of two components:

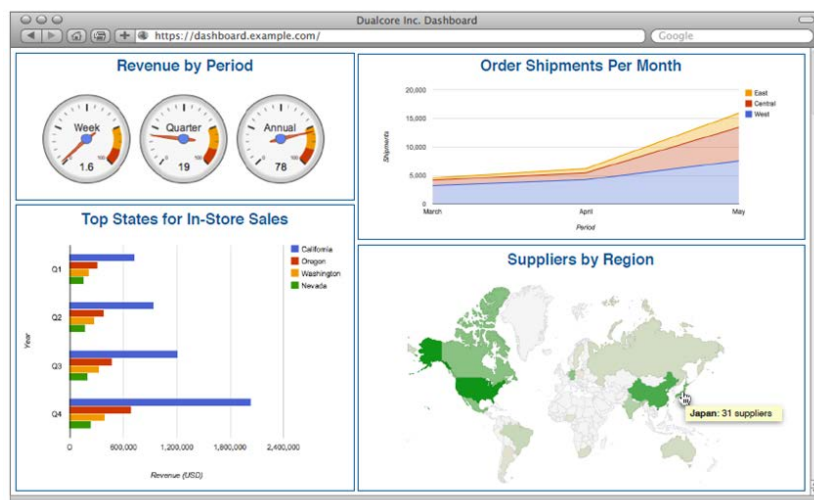
- 1) Intermediate presentation: Idea Proposal
- 2) Final result presentation

The wide scope for the project is **“big data management and analysis”**. It includes data collection, data storing into the data management system, and finding new interesting results from the data. Only requirement is to use databases for storing and managing data and to make Web services to show the results.

Under the given scope, specific requirements for the results need to be investigated by each team with creative ideas.

### ➤ Example: Trend Capturing from Social Media

- A social media research firm wants to know the trends of topics discussed on Twitter
- For easy analysis, it calculates the relevant counts based on the selected keywords in the following categories:
  - Apparel (clothes, shoes, watches, ...)
  - Art (book, DVD, music, ...)
  - Event (travel, concert, ...)
  - Health (beauty, spa, ...)
  - Tech (computer, laptop, tablet, ...)
- The final results: recent trends by the categories in Twitter
- Show the results using Web services in different multiple perspectives



### ➤ Grading

The total portion for the projects is 30%. The specific portions are as follows:

- 1) Intermediate presentation: 10%
- 2) Final result presentation: 20%

Basically, the evaluations are conducted by the teams. However, individual contributions on each team will be considered as well. For this, each student should submit the evaluation to the team members for each presentation in the e-class system.

### ➤ Team formation

- One team will be formed with 3 or 4 students
- The results will be noticed in e-class

### ➤ Feedback strategy

The feedbacks will be given during the presentation. Some more detailed comments could be provided through e-class system

## Intermediate Presentation: Idea Proposal

### ➤ Description

- Present idea proposal in the given wide topic
- Prepare the presentation within 15 minutes for each team

### ➤ Requirements

- Choose data sets to analyze
- Present how to collect, store, and analyze the data sets in a big data management system (i.e., Hadoop framework)
- Present the expected results, which will be shown using the Web services
- Design the overall architecture that connecting big data management systems and Web services.
  - You don't need to consider the details of Web services in this step

### ➤ Evaluation criteria

- Creativity of the idea
  - Choosing a good topic based on the idea is one of important evaluation criteria
- Feasibility of the idea
  - Validity of the idea. Can be the idea developed in the semester?
- Completeness of the idea
  - Include all the possible cases? Consider exceptional cases?
- Planning the project
  - Valid plan to follow
- Presentation
  - Organization of the presentation
  - Well understanding

### ➤ Presentation

- WHEN: 3 Apr. 2023 at 14:00 PM ~
- The presentation slide including the followings:
  - Basic idea (data sets, final target)
  - System overview and architecture (components for overall service and their relationships; components could be the user interfaces, big data management systems, databases, or Web servers)
  - Detailed functionalities (details for each component)
  - Intermediate results (using big data management systems, which will be connected to the Web services in the final result)
  - Project plan (you can optionally use a project management software, e.g., bitrix24,

Scrum)

➤ **Submission Deadline**

- Submit 1) the presentation slide and 2) supplementary materials (e.g., source codes) to e-class
- WHEN: 3 Apr. 2023 at 14:00 PM (Before the presentation)
- Evaluate to the team members. Use the evaluation form in the team project section of the e-class. (The evaluation results are not shown by the others)
  - Due date for the peer-review: 4 Apr. 2023 at 23:59 PM
- Late submissions are not allowed