SEOULTECH
Seoul National University of Science & Technology

The Project Overview for ITM 527

➢ Description

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

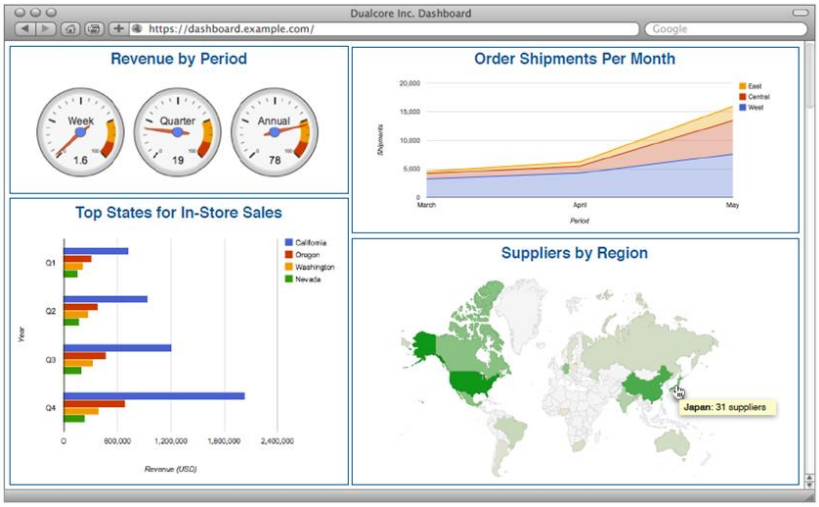
1) Intermediate presentation: Idea Proposal
2) Final result presentation

The requirements are twofold: 1) to use big data management systems for storing and managing data efficiently in a big data management environment, and 2) to conduct meaningful analysis based on the managed data. The analysis should move beyond simple summaries to generate valuable insights (e.g., trend detection, predictive modeling, comparative analysis). The final results should be presented in a clear and effective manner (e.g., interactive Web services with visualization).

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



Both big data management and analysis are emphasized equally. Intermediate presentations will be assessed not only on creative ideas and feasibility, but also on clarity of planned analysis methods. Final presentations will place emphasis on the completeness of the implemented management system and the quality of analysis results derived from it.

➢ Grading

The total portion of 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 to 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 feedback will be given during the presentation. Some more detailed comments could be provided through e-class system.

SEOULTECH
Seoul National University of Science & Technology

Intermediate Presentation: Idea Proposal

➢ Description

- Present idea proposal in the given wide topic
- Prepare the presentation within 10 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 demonstrated as you preferred way, for example, using Web-based applications. The details could be ignored during idea proposal.
- Design the overall architecture for big data management

➢ 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 the idea be 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

- The presentation slide includes the following:
  - 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 project management software, e.g., bitrix24, Scrum)

- You can choose one of presentation methods: 1) online or 2) offline (**WHEN: 13 Oct. 2025 at 14:00 PM)**
  - If you choose online manners, refer to the following directions: Build your own video presentation. When we record the video, please note that your face needs to be included.

For easy recording for this configuration, Zoom is recommended.

- Submit 1) video contents, 2) the presentation slide, and 3) supplementary materials (e.g., source codes) to Google Drive (a specific location will be noticed in the e-class) before the presentation time
- If you choose an offline manner, you need to submit 2) and 3) above to e-class and then present your results in an offline session below.

➢ Submission Deadline

- Submit 1) the presentation slide and 2) supplementary materials (e.g., source codes) to e-class
- **WHEN: 13 Oct. 2025 at 14:00 PM**
- Evaluate 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: 15 Oct. 2025 at 23:59 PM**
- Late submissions are not allowed

SEOULTECH
Seoul National University of Science & Technology

Final Presentation

➢ Description

- Present the results of final implementation of the proposed idea
- Prepare the presentation within 15 minutes for each team

➢ Environments

- You can use **big data management systems** such as HDFS, Sqoop, Impala, and Hive for collecting, storing, and processing the data sets (i.e., raw and intermediate data sets).
- You can use any kind of analysis tool or programming languages (e.g., Python).
- Depth of data management
  - Efficiency and appropriateness of chosen big data management tools (databases, Hadoop framework, etc.)
- Depth of analysis
  - Appropriateness of analysis methods for the dataset
  - Ability to derive meaningful or novel insights from the data
  - Consideration of limitations and exceptional cases

➢ Evaluation Criteria

- Well-designed to satisfy the proposed idea
- Completeness of the results
  - Are all the cases considered?
- Analyzing the problems that occurred during the implementations
- Debugging the problems
- Presentation

➢ Presentation and Submissions

- The presentation slide includes the following:
  - The final result of implementation
  - Detailed descriptions of the implemented components
  - The steps of analyzing the problems and debugging them
- A compressed file for the supplementary materials including all the source codes, how to obtain the data, a manual document describing how to deploy the source codes and data. Based on this manual, the implemented services could be reproducible in any environment.
- Evaluate 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)

- You can choose one of presentation methods: 1) online or 2) offline (**WHEN: 1 Dec. 2025 at 14:00 PM)**
  - If you choose online manners, refer to the following directions: Build your own video presentation. When we record the video, please note that your face needs to be included. For easy recording for this configuration, Zoom is recommended.

- Submit 1) video contents, 2) the presentation slide, and 3) supplementary materials (e.g., source codes) to Google Drive (a specific location will be noticed in the e-class) before the presentation time
- If you choose an offline manner, you need to submit 2) and 3) above to e-class and then present your results in an offline session.

➢ Submission Deadline

- Submit 1) recorded video contents, 2) the presentation slide, and 3) supplementary materials (e.g., source codes) to **Google Drive (a specific location will be noticed in the e-class)**
- If the presentation is done in an offline manner, you need to submit 2) and 3) above to e-class
- **WHEN: 1 Dec. 2025 at 14:00 PM (Before the presentation)**
- **Due date for the peer-review: 3 Dec. 2025 at 23:59 PM**
- Late submissions are not allowed