

The Term Project is a group assignment that is based on the content and sequence of weekly lectures, individual assignments, and quizzes. Students will form teams (max. 5 per team) and work with a chosen Big Data set. They will have to formulate a data analytics use case around the dataset and engineer a data processing system using at least 3 technologies studied in the course. An example term project may involve designing an information retrieval system with a front-end API service and back-end SQL and NoSQL data stores.

## Details

The Term Project deliverables include:

1. Data store and ETL process designs, as well as the Python implementation of the end-to-end data management application
2. Live presentation with a demo of an API or an interface for interaction with the application.

The presentation must contain:

- Background and definition of the data analytics problem
- Data source specification and procurement details
- Implemented design choices and the rationale for using the selected technologies
- Applicable data quality dimensions, licensing, scalability, and any cost implications
- Conclusion and future recommendations

During the last 2 class sessions, each student group will be presenting their slide deck with participation of every group member.

## Assessment

The project will be assessed per the grading rubric below.

## Submission

1. Jupyter Notebook(s) (.ipynb files) containing Python code implementation
  2. A PowerPoint slide deck (maximum 10 pages) of the final presentation
- 
1. Click the blue **Submit Assignment** button at the top of this page.
  2. Click the **Choose File** button, and locate your submission(s).
  3. Feel free to include a comment with your submission(s).
  4. Finally, click the blue **Submit Assignment** button.

