PROJECT REQUIREMENTS
WORKSHOP IN DATA MANAGEMENT (FALL 2021-2022)

SUBMISSION DEADLINE

29/1/25

FIRST STAGE

10 minutes presentation of

- Application idea (remember the theme!)
- A bit about the data you plan to use (which dataset, where the data was taken from, which part of the data within)
- Use slides for the presentation, show what you are going to use (snippets of the data, initial app design etc.)
- Show initial schema design

Submit handout – at moodle:

- The slide you use in your presentation
- 2 pages with your
    - Application idea
    - Data
    - Initial schema design
    - Work plan: steps of the project, who does what and when

Choose a slot for presentation at moodle

REQUIREMENTS

- You must use an external dataset as the basis for your application
    - Open source, free, no special registration
    - You must filter the data and change the DB design to suit your application
- DB design according to the principles we have learned
    - See DB design grade guide below
- At least 100,000 rows in you DB in total
    - Can and should be much more!
    - All must be really used in the application
- Application using the DB
    - Should be interesting and fun to use!
    - See application and source code requirements below
- At least 5 “complex” queries used by your application
    - Nesting and/or
    - Union and/or
    - Group by and aggregation
    - Make sure the complexity is justified, i.e. the query is not equivalent to a simple query
    - Make sure it makes sense in the context of your application
- Application relies on the DB where possible for data processing

GRADE GUIDE

The final grade is composed of 5 factors:

- DB design
  - **Structure**
    - Schema design suits the application
    - Complexity – rich schema
    - Meaningful tables and table names
    - Meaningful fields and field names
    - Field types and lengths
    - Avoiding data redundancy
  - **Queries and Indexing**
    - Efficient and correct queries
    - Efficient and correct updates, including transaction usage if needed
    - Complexity – complex queries/updates
    - Use of indexes where needed
  - **Keys, FKs**
    - Use of primary keys for “logical” keys
    - Use of foreign keys for “logical” connections between tables
    - Others (e.g., unique), if needed
- Application
  - Idea
  - Usability
  - Features
  - Interesting use of data
- Source code
  - Modularization – separate DB access, logic and UI
  - Using the DB where possible instead of application code or alternative storage
  - Wrapping the DB connection effectively
  - Error handling
- Project Management
  - Steps of the project
  - User manual
  - Software documentation
- Bonus (should be something really exceptional)
  - Interesting algorithms
  - Exceptional UI
  - …

Tips: The most important aspects of the project are what you do with the data, i.e., your DB design and querying / manipulation of data by the application.

The two documentation files are *extremely* important (user manual and software documentation) – without them the project cannot be checked, and if they are well-written this ensures that the grader is aware of everything you have worked so hard on.
The grader may fail the project if due to lack of documentation it cannot be graded, executed or the DB cannot be inspected.

GRADE GUIDE

Will be announced.

- You will not be allowed to store any part of your application on a remote server on your computer or an external host. Plan for installing your project including the DB on our VM or server, so that if you are writing a Web application, you may need to install the Web server or ensure such a server is available.
- If you are developing an Android or any other embedded application, note that
    - It will have to run in an emulator
    - You will have to connect to an external DB server