Project 2
(Due on Tuesday, ~~Dec 3rd at 2:00pm~~)

April 14th due date

14 points towards your final grade (Graded out of 100)

In this project, we will focus on using a dataset for creating an application. This exercise is to be completed and submitted as a group (Max 2 people).

You as a group can choose a dataset of your preference. The only condition is that the data set should be substantial.

Basic requirements
1. You must use a substantial data set (on the order of the voter data, but not necessarily the voter data itself). Data size should not be less than 200MB. Must have at least 50 attributes in total.
2. You must use MySQL as a back-end data source (explicitly, directly, live connecting / using the MySQL service) for some application function. This could be, for example, a web application (e.g., register to vote) or using the database as a live source for an interactive visualization package like Tableau. The application may be either analytical or transactional.
3. As part of the implementation, you must incorporate some functionality that uses concepts from the second half of the term (e.g., transactions, views, triggers, indexing, etc.).
4. Populate your database with a significant amount of data. You will want this both for testing and for demonstration of your system.
5. Develop an application interface - that is, connect your database functionality to an application interface, enabling user interaction with the underlying data. This may be for a subset of the overall functionality, but it should be significant.

The final project deliverable consists of five main parts.

1. **In-Class Project Demonstration** during the final week of classes.

2. **Project report document**. It should be submitted as a PDF document. This should be on the order of 5 pages, but not more than 10 pages, maximum. This should not include any large blocks of code or SQL (those will be submitted separately). Screenshots should not be larger than half-page (Max 2 pages of screenshots). Preferred font is Cambria/Arial. Font size shouldn't be greater than 12 and line space shouldn't exceed more than 1.3. 5-10 page limit doesn't include cover page.

   Report content should consist of:

- o Introduction -- About application, dataset. Why this application? Why this dataset? What you wanted to do with the dataset? What you wanted to do in your application?
- o Final Implementation
    - ▪ Description of final implementation: use cases, platform, etc.
    - ▪ Final database design (description, conceptual model, business rules).
        1. Explain the normalization process with functional dependencies.
        2. Explain the final database design.
        3. Include UML with cardinality. (No hand sketched UML)
    - ▪ Description of data used for testing (how created / acquired, numbers of items entered / needed for testing)
    - ▪ Summary of implemented use cases
- o Illustration of Functionality
    - ▪ Walkthroughs for primary use cases using sample data (should illustrate all of the main functionality of the system, should have enough detail that course staff can do them with your setup and data).
- o Summary Discussion
    - ▪ Explanation of status and stopping point, if incomplete.
    - ▪ Discuss the challenging parts of the assignment. How did you overcome all or some of the challenges?
    - ▪ Discussion of division of labor, team roles and responsibilities / work breakdown, specifying who did what and why this is a fair and equal split.

3. **Database code**. This should consist of SQL files containing the SQL statements that are the core of the project. Running these files, we should be able to recreate exactly the database tables and data that you have used for the project. Separate following parts in your SQL file with clear headers and some white space:
   - o Database creation
   - o Data insertion (**not full dataset** - load statements only, but with reference / pointer to data source)
   - o Analytics and/or application functionality (these may be illustrative of particular uses).

4. **Application code**. This should consist of all the application files that would be needed to set up and run your application. This would be an archive of the project, including code, libraries, supplemental data (e.g., image and other assets for web pages), and this MUST include all code source files. This must also include a README file documenting the basic environment setup needed in order to re-create the running project (application framework, DBMS, versions, standard framework libraries, etc.).

5. **Demonstration Video**. Teams will need to demonstrate the functionality of their project. This will essentially be a brief introduction to the application and team's

original functionality, but mostly running through the walkthroughs of the report to show that the project actually works. The demonstration should illustrate all of the main functionality / use cases. Teams can complete the demonstration either by:

- o Creating a screencast video (max. 5 minutes duration). There are a number of software packages, such as Quicktime, freely available that will allow you to record video from your screen along with audio. Record the video illustrating the use cases, then upload to a service such as Youtube or Vimeo, and submit the link to the online video. It is strongly suggested that you do some practice runs before creating a video.

6. Complete the individual submission on teamwork, noting effort breakdown and roles. This will be posted in Canvas and must be completed individually, NOT in the presence of other team members.

Final submission should consist of:
If the submission is not in the following format then there is a 10% deduction.

- Single team submission of a zip file archive, called "**lastname1-lastname2-Project2.zip**" (So if I am working with John Doe then our zip file will be named **singh-doe-Project2.zip**) containing:
  - o "report" directory containing a PDF document version of the report file
  - o "db" directory containing the SQL files for database creation, population, and use.
  - o "application" directory containing the specified project / code / documented SQL files for the project application as a whole
  - o "demo" directory containing a text file with the URL/link to the video (note, **this is not the full video file**)
- Text in the assignment submission, that provides either the URL/link to the video
- Individual completion of the teamwork assessment posted in Brightspace.

In the very unlikely event that the final project zip file is too large to submit to Brightspace (upload limit is 500MB), (1) you should notify your instructor/grader well in advance of the deadline; (2) you must submit the non-gigantic parts through Brightspace (report, demo link, etc); (3) you will be required to submit the same information on external media such as Google drive by the deadline.

Every member from your team must make the final submission of same files. Once uploaded to Brightspace, you should verify your submission by re-downloading the file, opening the archive and verifying your content. This is how we will grade the work, and if it is not a SQL file that can be opened and viewed in this way, it will receive no credit.

Grading Distribution:

1. **Database code – 60%**
   a. Your database code must be a self contained db file. It means that if I execute your sql file it should be able to create the database and tables.
   b. Normalized tables in at least 3NF
   c. Use of correct data type
   d. Use of constraints on tables
   e. Use of advanced features – At least three features (with 2 instances of each selected) from the following.
      i. Views
      ii. Transactions
      iii. Stored Procedures
      iv. Triggers – Use of audit tables
      v. Indexing etc. (Indexing on Primary Key not counted)
   f. Commented code
   g. Other DB usage including DML and DDL
   h. Detailed UML diagram (in report)
   i. Functional dependencies (in report)

2. **Project Report – 15%**
   a. Clarity of content in the report
   b. Structure of the report
   c. Completeness of the report

3. **Application and In class demo – 15%**
   a. How complete is the application?
   b. The application is connected to the data and interacts with it.
   c. How many different functionalities working?
   d. Class presentation coordination
   e. Peer review

4. **Video walkthrough – 5%**
   a. A working demo video submitted
   b. Video demonstrates application's functionalities
   c. Audio commentary

5. **Project Progress Report -5%**
   a. Submitted all progress reports.
   b. Adequate progress shown on each report.