Project

Purpose
To analyze the requirements, design, implement, document and test a database application for Elections Department of the country of Wonderland. The User Requirements of the database application are given in **Appendix A** below.

Materials to submit
You will be asked to submit reports at various milestones as major phases of the project are completed. Further, you will also need to demonstrate the database application you designed and developed. Details about the milestones of the project, the phases they include and their due dates will be posted on the class webpage.

Project Report/Phases
The project consists of the following phases. In addition, you will develop and maintain a Project Report. Your Project report should consist of a section for each phase, as well as an executive summary.

A. Executive summary that includes a brief introduction to the project, high level description of the requirements, and a succinct discussion of your accomplishments (as well as any significant limitations).

B. Conceptual Design.
   This stage involves the following
   - Develop an Entity-Relationship model detailing the relations involved.
   - Identify the attributes of the entities and the relations along with the primary key for each entity.
   - List the constraints for each relation and entity.
   - You should be able to explain the reasons for the particular design approach you have chosen.

C. Logical Design
   This stage involves the mapping of your conceptual design above onto the relational data model. In this stage, you will design the tables for all your entities and relations. You should apply all normalizations you find useful and/or necessary. Ensure that your design still satisfies the user requirements. Justify your design choices.

D. Physical Design
   This stage involves the following:
   - Design a MySQL database based on the design developed in section B.
   - Implement tables for the relations and the constraints. Maintain scripts for the creation and deletion of tables.
   - Maintain scripts for loading data into your tables.
   - Ensure that your design still satisfies the user requirements.
   - Justify your design choices.

E. Prototype, Development, and Testing.
   This consists of the following:
   - Develop (an) appropriate applications and/or user interface(s), using Python Jupyter Notebooks, which satisfy all functional user requirements.
   - SQL scripts for creating indices for the database application. Justify the reasons for creating any such indices.

   Your notebooks should have appropriate cells acquiring input from the user and providing output to the user (as needed).

F. Make a user's guide for the database application. The user guide could be integrated within your Jupyter notebook.

Project Demonstration

- Populate the tables with data as described in the Appendix.
- Default values for each query/report demonstrating the corresponding functionality.

Miscellaneous

The final project report should document all the activities with appropriate E-R diagrams, relation schema, etc. It should also give a list of the limitations of the application and give possibilities for improvement.

Features and functions other than specified in the document can also be added but should be documented clearly and demonstrated as well.

Appendix A

WVT22 Elections Champaign
User Requirements

Due to an upcoming elections season in the country of Wonderland, its Elections Department is interested in developing a database application to help manage the ballots cast by folks across the country.

System Scope

The users of the application are the folks residing in Wonderland and staff of the Elections Department. It is assumed that all users have network computers capable of running Web browsers.

Data Requirements

Folks

The information stored about each individual in Wonderland includes a 16-digit personal identification number, name (first name, last name, and a nickname), date of birth, two 10-digit telephone numbers (primary, and secondary), and one or more email addresses. Identification numbers are unique across Wonderland. All folks of Wonderland reside at a place, and some folks may reside at the same place.

Elections Staff

Some folks are staff members of the Elections Department. Staff can be clerks or monitors. Elections staff are scheduled to work at a voting center for some datetime interval. Some staff may not be scheduled at all.

The format for all dates is YYYY-MM-DD, the format for all times is HH:MM:SS, and the format for all datetimes is YYYY-MM-DD HH:MM:SS. Date, time, or datetime intervals consist of a start and end date, time, or datetime values respectively.

Places

Each place in Wonderland has an address (street number, street name, city, state, and zipcode) and XY-coordinates (a pair of floats; see below). There are no two distinct places in Wonderland with the same address or coordinates. Places can be either residences or voting centers, but not both.

Voting centers have unique acronyms (with at most 4 alphanumeric characters [A-Z0-9]) and one or more operating periods, i.e., a collection of datetime intervals. Changes to a center's operating period(s) should be allowed only if it would not invalidate the (previously) valid registration of an already cast vote.

The geographic extent of Wonderland is a flat plane. Wonderland uses a standard XY Cartesian 2D orthogonal system of axes with the origin at its capital city, Megapolis, and units of miles. Distances between places are computed using the standard Pythagorean Theorem.

Ballots

The Elections Department maintains a collection of ballots it offers at all the voting centers. Each ballot has a unique short name (with at most 4 alphanumeric characters [A-Z0-9]), a question text, and an availability period (i.e., the datetime interval when folks may cast a vote for the ballot). The set of possible choices as answers to each ballot question is "YES", "NO", "ABSTAIN", or NULL.

A ballot may be modified only if there are not any cast votes for the ballot. Ballots with no registrations may be deleted.

Voting Registry

Wonderland forks indicate their intent to cast a ballot at a voting center on some date by registering with the Elections department. To register, they provide their identifier, the voting center identifier, and their desired voting date. In order to choose their desired voting center and voting date, folks ask to first see the number of folks currently registered to vote at each voting center in their city on their desired date, and then they quickly complete their registration.

A folk can register to vote at a center on a date only if that date is contained in an availability period of that center; such registrations are deemed valid (registrations those that don't satisfy this condition are deemed invalid). A folk can have multiple valid registrations.

Casted Votes

A folk can cast a vote for a ballot on the date and center of one of their valid registrations, and during the operating hours of that center on that date.

When a folk casts a vote for the ballot, the Elections Department maintains their answer to the question of the ballot together with the datetime and center where they casted the vote.

A folk can cast at most one vote for any single ballot. Attempts to vote more than once or without a currently valid registration are rejected.

For each casted vote, the Elections department wishes to record the voter's choice (vote), the datetime when the vote was cast, and information uniquely identifying the corresponding valid voting registration (that enabled the vote).

Changes to ballots or centers that may affect the integrity of a casted vote should be rejected (e.g., invalidating of a cast vote's certified valid voting registration, changing the ballot question).

Functional Requirements

Populate your database with at least

- 6 staff
- 12 folks
- 6 places, distributed among 2 cities and 2 states.
- 3 voting centers, each with 4 operating periods
- 4 ballots, each with 3 possible answers
- 24 voting registrations, distributed among 3 ballots, 3 centers, and 2 months
- 18 cast votes, distributed among 2 ballots

and appropriate relationship instances among them.

Develop a Jupyter App with a section for each of the following:

A. Activities:
   1. a clerk creating a new ballot
   2. a folk registering to vote at a center
   3. a clerk modifying the availability period of a current ballot
   4. a voter casting a ballot while confirming a valid voting registration
      - Use a SQL transaction to ensure ACID properties (serializability) in the face of other concurrent operations or failures.
   5. a staff removing a folk (and all their associated information)

B. Queries/Reports:
   1. List the name, city, and email (any single email suffices) of all folks.
   2. List the city, state, and the number of residents of each city in Wonderland (skip cities with no residents) in decreasing order of number of residents.
   3. List each center together with its number of currently registered folks (include states with no inhabited places) in increasing alphabetical order of their zipcode.
   4. Find the distinct identifiers and names of folks registered at a given voting center within a given time period.
   5. Find for a given month, the number of unique registrations at any voting center which is within 5 miles from the center of Megapolis, except for voting centers in a given (exclusion) list of voting centers.
   6. Find the most popular voting center(s) (in terms of total number of registrations) in a given time period among those in a given city.
   7. Find the unique folk that have valid registrations with every voting center on a given state.
   8. Find folks that registered at a voting center that is farther away than the voting center closest to their residence (break ties alphabetically by center's acronym).

9. Write a SQL function that returns the acronym of the voting center closest to the residence of a given folk among those whose operating period(s) contain a given date (return NULL if no such center exists; break ties at alphabetically by acronym).
10. For a given ballot, construct a cross-tabulation of voting centers (acronym) as rows, unique ballot answers (options) as columns, and cells indicating number of cast votes at the row's center with the column's option.

Populate your database with sufficient data to demonstrate successfully execution of the 15 operations/queries/reports above.

For any of these operations/queries/reports, you may assume that the default start and end date of a time period is the current date of the system when the operation is invoked, e.g. NOW().