

## CS-3200 Homework 4

### Creating a database from a collection of .csv files

This assignment gives you an opportunity to design a data schema from scratch from 4 input .csv (comma separated values) files. The files are an excerpt of a data base that describes character appearance in the Original Star War's trilogy. The provided files are:

- **Characters.csv** – data on the characters in the movie. In particular each row in the file is information on a specific character. The fields for each character are: the character's name, the character's race (if known), the character's home world (if known) and the character's affiliation (either rebel, empire, neutral, free-lancer)
- **Movies.csv** – the names of the movies you are tracking. In particular each row in the file is information on a specific movie. The fields for each movie are: the movie's id, the movie's name, the number of scenes that have been entered into the database, and the total number of scenes for the movie.
- **Planets.csv** – data on the planets appearing in the movies. In particular each row in the file is information on a specific planet. The fields for each planet are: the planet's name the planet type (either gas, swamp, forest, handmade, ice, desert) and the planet's affiliation (either rebel, empire, neutral, free-lancer)
- **Timetable.csv** – the appearance of a character on a specific planet in a specific movie during a specific time interval. Each movie is broken into 12 chunks or time interval, where data for the first 10 time intervals has been entered into the movie. Timetable tracks the movement of characters from place to place. Each row contains information on a character's location for scene (time interval) numbers of the movie. Each row contains the character's name, the planet's name, the movie id in which the character visited the planet and the time of arrival on that planet ( a scene number) and the time of departure from that planet ( an equal or later scene number).

Design a schema that supports the relationships in the data. Name the schema *starwarslastnamefirstinitial*, where *lastname* is your last name and *firstinitial* is the first letter of your first name. Please submit a zip file following the same naming terminology containing both the EER diagram/model and the dump of your schema. Please follow these steps to complete this assignment:

1. Identify the tables needed to support this data.
2. Identify legal MySQL field and table names for the data. Please make the names descriptive.
3. Identify data types for each of the fields.
4. Identify primary keys for the tables.
5. Identify all foreign key constraints and determine the action that should take place in the database if a constraint would be violated.
6. Create your tables using the InnoDB data engine.
7. Use the MySQL model tool to create an EER diagram (available from the home tab in the MySQL workbench) to help with the design process listed above or to verify your final design.
8. Submit either a picture of the EER diagram of your schema or the .mwb file of your schema from the model tool.
9. Once you have completed your schema, import the provided data using the 'table data import wizard'. You may need to modify the data in the provided files in order for the import to be

successful. The provided files are small and can easily be modified using a simple text editor. You may also have to order the MySQL file imports due to foreign key constraints.

10. Generate a self-contained extract of your database to blackboard using the 'data export' tool from the 'Server' menu. We must be able to import your schema so please ensure the extracted file works with import.
11. Submit a zip file containing the self-contained extracted data and the EER.