stmw2024

Study repository for Uni Bremen course "Search Technology for Media & Web (Winter Semester 2024/2025)"

How to Run Batch File

- Open docker container
  `docker start -ai pa1_container`
- Go to directory where the batch file stored
  `cd student_workspace`
- Restart MySQL server inside docker (somehow it's needed to resolve some MySQL server error)
  `service mysql restart`
- Setup local-infile inside MySQL
  `mysql`

  `SET GLOBAL local_infile=1;`

```
root@4be5133fa12a:/stmw/student_workspace# mysql
Welcome to the MySQL monitor.  Commands end with ; or \g.
Your MySQL connection id is 38
Server version: 8.0.40-0ubuntu0.24.04.1 (Ubuntu)

Copyright (c) 2000, 2024, Oracle and/or its affiliates.

Oracle is a registered trademark of Oracle Corporation and/or its
affiliates. Other names may be trademarks of their respective
owners.

Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

mysql> SET GLOBAL local_infile=1;
Query OK, 0 rows affected (0.00 sec)

mysql>
```

- Run the bash script
  `./CustomRunLoad.sh`
- Example process:

root@4be5133fa12a: /stmw/student_workspace

```
root@4be5133fa12a:/stmw/student_workspace# ./CustomRunLoad.sh
Dropping existing tables...
Tables dropped successfully.

Creating new tables...
Tables created successfully.

Create new directory for storing output files from conversion XML to CSV...
New Directory 'ebay-data-csv' created successfully.

Compile and run the convertor...
Java XMLToCSV ran successfully.

Loading data into tables...
Data loaded successfully.

Remove all temporary files...
All temporary files removed successfully...

Database setup and data loading completed successfully.
root@4be5133fa12a:/stmw/student_workspace#
```

##

Known Limitations

- There is an error in `Bids` table as I missed `itemID` column

Design Relational Schema

Step 1: List Your Relations

Items:

- `ItemID` (Primary Key)
- `Name`
- `Currently`
- `Buy_Price`
- `First_Bid`
- `Number_of_Bids`
- `Location`
- `Country`
- `Started`
- `Ends`
- `Description`

ItemCategory:

- `ItemID` (Foreign Key)
- `Category` (Composite Primary Key with `ItemID`)

Bids:

- `ItemID` (Foreign Key)
- `BidderID` (Composite Primary Key with `ItemID` and `Time`)
- `Time`
- `Amount`

Bidder:

- `UserID` (Primary Key)
- `Rating`
- `Location`
- `Country`

Seller:

- `UserID` (Primary Key)
- `Rating`

Step 2: List Functional Dependencies

Let's list the non-trivial functional dependencies for each relation:

- **Items**:
  - `ItemID → Name, Currently, Buy_Price, First_Bid, Number_of_Bids, Location, Country, Started, Ends, Description`
- **ItemCategory**:
  - `ItemID, Category → (Item is identified by the combination of ItemID and Category)`
- **Bids**:
  - `ItemID, BidderID, Time → Amount`
- **Bidder**:
  - `UserID → Rating, Location, Country`
- **Seller**:
  - `UserID → Rating`

These dependencies capture the relationships and constraints in your data.

Step 3: Boyce-Codd Normal Form (BCNF)

Checking if the relations are in BCNF:

- **Items**: Already in BCNF, as all non-trivial FDs have a superkey as their determinant.
- **ItemCategory**: In BCNF, as the combination of ItemID and Category forms a composite key.
- **Bids**: In BCNF, as all non-trivial FDs have a superkey as their determinant.
- **Bidder**: In BCNF, as all non-trivial FDs have a superkey as their determinant.
- **Seller**: In BCNF, as all non-trivial FDs have a superkey as their determinant.

Step 4: Fourth Normal Form (4NF)

There are no multi-valued dependencies for all five columns

Key components

1. **Items** table:
   - Contains all essential attributes of an item.
   - `ItemID` as the primary key.
2. **ItemCategory** table:
   - Handles the relationship between items and categories.
   - Composite primary key consisting of `ItemID` and `Category`.
3. **Bids** table:
   - Records bid information.
   - Composite primary key of `ItemID`, `BidderID`, and `Time`.

4. **Bidder** table:
    - Details about each bidder.
    - `UserID` as the primary key.
5. **Seller** table:
    - Details about each seller.
    - `UserID` as the primary key.

Troubleshooting

- How to use `dos2unix` in Windows
  https://unix.stackexchange.com/questions/721844/linux-bash-shell-script-error-cannot-execute-required-file-not-found
    - Convert using this command:

      ```
      dos2unix ./<shell_file_name>.sh
      ```

    - Example usage:

      ```
      root@4be5133fa12a:/stmw/student_workspace# dos2unix ./CustomRunLoad.sh
      dos2unix: converting file ./CustomRunLoad.sh to Unix format...
      root@4be5133fa12a:/stmw/student_workspace# ./CustomRunLoad.sh
      test
      ```

- Got this error
  `ERROR 2002 (HY000): Can't connect to local MySQL server through socket '/var/run/mysqld/mysqld.sock' (2)`
    - Restart `mysql` inside docker container:
      `service mysql restart` and try to run `mysql` again
    - Example:

      ```
      root@4be5133fa12a:/stmw/student_workspace# service mysql restart
       * Stopping MySQL database server mysqld                                [ OK ]
       * Starting MySQL database server mysqld
      su: warning: cannot change directory to /nonexistent: No such file or directory
                                                                             [ OK ]
      root@4be5133fa12a:/stmw/student_workspace# mysql
      Welcome to the MySQL monitor.  Commands end with ; or \g.
      Your MySQL connection id is 10
      Server version: 8.0.40-0ubuntu0.24.04.1 (Ubuntu)

      Copyright (c) 2000, 2024, Oracle and/or its affiliates.

      Oracle is a registered trademark of Oracle Corporation and/or its
      affiliates. Other names may be trademarks of their respective
      owners.

      Type 'help;' or '\h' for help. Type '\c' to clear the current input statement.

      mysql> exit
      Bye
      ```

    - Source: https://forum.hestiacp.com/t/error-2002-hy000-cant-connect-to-local-mysql-server-through-socket-run-mysqld-mysqld-sock-2/10239
- MySQL - ERROR: Loading local data is disabled - this must be enabled on both the client and server sides
    - After searching online, I fixed it by these steps:

- set the global variables by using this command:

```
mysql> SET GLOBAL local_infile=1;
Query OK, 0 rows affected (0.00 sec)
```

- quit current server:

```
mysql> quit
Bye
```

- connect to the server with local-infile system variable :

```
mysql --local-infile=1 -u root -p1
```