Mathematisch-Naturwissenschaftliche Fakultät
Wilhelm-Schickard-Institut für Informatik
Datenbanksysteme · Prof. Dr. Grust

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Datenbanksysteme I

WS 2019/20
Torsten Grust, Christian Duta

Assignment #2

Submission Deadline: November 05, 2019 - 10:00

Exercise 1: Types and Tables (10 Points)

So far, we considered an untyped "tabular" version of the Twitter data. We now transform this untyped data into a typed relational table.

Consider the files `tweets.csv` and `users.csv` provided to you in this assignment.

1. For each of the files `tweets.csv` and `users.csv`, construct a `CREATE TABLE` statement and a `\copy` statement to load the CSV data into two tables named `tweets` and `users`, respectively. For each column, choose an *appropriate* data type (`INTEGER`, `BIGINT`, `TEXT`, `TIMESTAMP`, ...). Please hand in a `.sql` file which contains your solution.

   Note: The PostgreSQL documentation provides additional information on the `CREATE TABLE`[1] and `\copy`[2] commands as well as the available data types[3]. The `\copy` meta-command may only be used from within the `psql`-shell.

2. Now is also the perfect time to install PostgreSQL 12 on your system and use it to test your `.sql` script:

   - The given CSV files still contain header lines with column names. If these remain unchanged the import using `\copy` will fail. Explain why!
   - Remove the header lines to import the data only.
   - Use the SQL `TABLE` command to list the contents of your new tables `tweets` and `users`. See the PostgreSQL documentation `here`[4] for more information.

[1] https://www.postgresql.org/docs/12/static/sql-createtable.html
[2] https://www.postgresql.org/docs/12/app-psql.html#APP-PSQL-META-COMMANDS-COPY
[3] https://www.postgresql.org/docs/12/static/datatype.html
[4] https://www.postgresql.org/docs/12/sql-select.html#SQL-TABLE