
Introduction to MySQL

Julie Butler

What is MySQL?

- ▶ MySQL is a software which allows us to interact with a database to retrieve and modify data using a set list of commands.
 - ▶ We will use MySQL through a Python library called pymysql, which allows us to combine MySQL commands with Python functionality (like Pandas!)

What is a Database?

- ▶ **Database:** structured collection of organized information
- ▶ Databases allows us to store data in tables (like a Pandas Dataframe or csv file) BUT there can be multiple tables in a single database (like there can be multiple sheets in a single Excel workbook).

Why would you use a database?

- ▶ More than one person needs to access and modify a data set (useful in companies and for group projects).
- ▶ A dataset contains data which is organized into more than one dataframe/table. A database let's you keep this together
- ▶ The data set is very large and you do not need to access it all at once to perform your analysis.

MySQL Commands

- ▶ "SHOW TABLES": lists the tables in the dataframe.
- ▶ "SHOW COLUMNS FROM <table>": lists the columns from a table.
- ▶ "SELECT * FROM <table>": selects an entire table.
- ▶ "SELECT <column>,<column> from <table>": selects certain columns from a table.
- ▶ "SELECT <column> FROM <table> WHERE <column>=<value>": selects rows from a certain column that equals a certain value.
- ▶ "SELECT t.<column> FROM <table> t INNER JOIN <table2> u ON t.<column> = u.<column>": pulls a column from a table, but only rows where there is overlap with another column from a different table.

Notes

- ▶ **Note 1:** The command words can be all uppercase or all lowercase but using all uppercase helps the commands stick out from the other words.
- ▶ **Note 2:** The SQL commands must be typed inside double quotation marks (“ ”)