

# Databases

## Lab 04

Andrés Calderón, Ph.D.

February 26, 2025

### 1 Introduction

In this lab, we will explore SQL queries using the **superhero** database, a sample dataset available from <https://www.databasestars.com/>. This database contains structured information about superheroes, their attributes, and their affiliations, providing an excellent opportunity to practice SQL operations such as **JOIN** variants, aggregation, and filtering.

The objective of this lab is to reinforce database querying skills by executing and analyzing SQL queries that extract meaningful insights from the dataset. Students will work individually to execute a series of predefined queries, capture their results, and document their findings. Additionally, students will be required to propose two original queries using different types of **JOIN** operations.

Throughout the lab, we will:

- Load the **superhero** database into PostgreSQL.
- Practice SQL queries using **JOIN** operations.
- Analyze the dataset to answer specific questions about superhero characteristics.
- Submit a structured report including SQL queries, results, and a ZIP file containing the SQL code.

By the end of this lab, students will have strengthened their understanding of SQL joins and database querying techniques while working with real-world-style data.

### 2 W3Schools Interactive SQL Tutorial

We will revisit the interactive SQL tutorial on the W3Schools website, but this time, we will cover new topics. Remember that the tutorial starts [here](#). From the lessons in the left panel, you need to cover the lessons from **SQL Joins** up to (and including) **SQL Null Functions**. In total, there are 15 lessons, so make sure to cover all of them.

Again, you will capture a screenshot when you execute the SQL statements (just one per lesson is enough) so you can attach them later to your report. Try to include your username if possible. This part of the lab is individual, so all members of the group must show their work.

### 3 The superhero Database

The website <https://www.databasestars.com/> provides a rich collection of sample databases. One of them is the “*Superhero*” database. You will read the documentation about it at <https://www.databasestars.com/sample-database-superheroes/>. See Figure 1 for the E-R diagram of this database. To load the data into PostgreSQL, you will need to download three files:

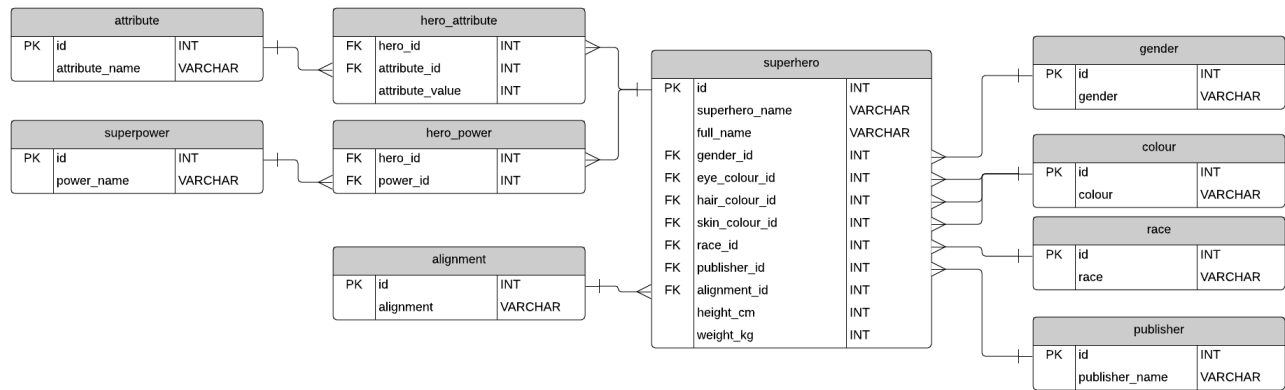


Figure 1: The Superhero Database E-R Diagram.

1. [01\\_reference\\_data.sql](#).
2. [02\\_hero\\_attribute.sql](#).
3. [03\\_hero\\_power.sql](#).

Download the files and place them in the same folder. You can load the files in order, one after the other, in the same way we did in the last lab.

First, we will create a database named **heroes**:

```
createdb heroes
```

Next, we will connect to the database:

```
psql heroes
```

Now that we are in the PostgreSQL prompt, we will use the `\i` command to load the files into our database. You will need to update the file paths accordingly.

First, we will load the `01_reference_data.sql` file:

```
\i /path_to_files/01_reference_data.sql
```

Then, we will load the `02_hero_attribute.sql` file:

```
\i /path_to_files/02_hero_attribute.sql
```

Finally, we will load the `03_hero_power.sql` file:

```
\i /path_to_files/03_hero_power.sql
```

And you're done! Let's read a transcript of a podcast discussing the database. You can read it in [English](#) or [Spanish](#), and try out the SQL examples found there.

## 4 Independent Work

Let's see what we can learn from the **superhero** database. You are asked to answer the following queries<sup>1</sup>:

1. What is Magneto's full name?
2. Display the name, full name, and gender of *supervillains*. We define a *supervillain* as a superhero with a "Bad" alignment.
3. Based on gender, which group has more superpowers? Show the total for each category.
4. List the superheroes who are the most diverse, meaning those who have different eye, hair, and skin colors. Display all colors along with the superhero's name.
5. Display the list of superheroes who do not have any attributes at all. Hint: Use a `LEFT JOIN` and look for **null** values.
6. True or False: Does 'Marvel Comics' have more "Good" superheroes than 'DC Comics'?
7. What is the average height and maximum weight of characters published by 'Dark Horse Comics'? Do not take into account any **null** or 0 values in the calculation.
8. Who is stronger: Superman or Goku? We will measure strength using the following formula:

$$\text{strength} = \text{count\_of\_attributes} + 2 \times \text{count\_of\_powers}$$

9. This time, you are asked to propose **TWO** queries using any (different) variant of `JOIN` and solve them using `SQL` on your own.

For each query, you must present the `SQL` code and the query result (either copy-pasted or as a screenshot). We expect you to submit a well-structured report in PDF format, containing the requested information from the previous sections. Additionally, include the `SQL` code with the answers to the queries in a ZIP file. The submission deadline is **March 12th, 2025**.

Happy Hacking ☺!

---

<sup>1</sup>Use only `JOIN` variants in your answers.