Homework 0: Environment Setup
COMS W4111 Introduction to Databases
Fall 2021

Introduction and Overview:

Homework 0 provides step-by-step instructions on how to install and setup all the necessary tools this course will require. We start the semester with a HW 0 environment setup prior to HW 1. This allows us to resolve configuration and set up issues before students begin working on graded assignments.

Tests and screenshots are required to show that the environment was set up successfully. Submission for this assignment is the accompanying W4111_HW0.ipynb file with completed tests and screenshots converted to PDF format.

This assignment is for both Programming and Non-Programming sections. All steps will be required for both tracks unless specifically marked for Programming only.

A points earned score between 0 and 100 determines the final grade for the semester. We will explain the grading, HW and exams in the first lecture. HW 0 is not worth any points, i.e. the points earned for submission is 0. Failing to submit HW 0 will result in a points deduction from HW 1, however. We will determine and publish the exact deduction by the start of the first class.

Note:

You can find some motivation, explanations, instructions and tutorials for Homework 0 in the homework's section on the website's [exams and assignments](#) page.

Due date: September 17, 10:00am ET on GradeScope
Please note: You may NOT use late days for the submission of this assignment. Check Courseworks for GradeScope access.

The following tools will need to be installed and or setup:
1. Anaconda/Python 3/Jupyter Notebook
2. PyCharm (Programming Track only, but strongly recommended for all students).
3. MySQL Server Community Edition.
4. DataGrip
5. Postman (Programming Track Only)

In the zip file accompanying this assignment you will find:

- HW0 Test Folder
- W4111_HW0.ipynb (what you will submit as a pdf)
- DDL.sql
- smallRelationsInsertFile.sql
- lahman-mysql-dump.sql

Anaconda/Python 3/Jupyter Notebook

Follow the online instructions to download the Anaconda Individual Edition here. It should default to installing Python 3.8. If you have previously installed Anaconda and have your environment on a different Python 3 version that is OK. If you have previously installed Anaconda and have a Python2 environment please follow the instructions at the end of this section to create a new Python3 Environment.

Anaconda Navigator is automatically installed with Anaconda. On the Home page, there should be several different applications that are available to install. Click install on the Jupyter Notebook application.

Launch jupyter notebook from Anaconda Navigator once it is finished installing and open the .ipynb file provided. Complete the associated python version tests (test details for this section will be in the notebook).

If you have previously installed Python 2 on Anaconda:

On the Anaconda Navigator, click the "Environments" tab on the left. On the bottom left click Create. Give your new environment a name and select Python 3.8. Once your environment is set up, you will need to select the environment to launch Jupyter Notebook on. Select the Python 3.8 environment from the "Applications on" dropdown menu at the top of the home page. You will need to reinstall Jupyter notebook and any python packages like pandas and numpy for this new environment.

PyCharm (Programming Track only, recommended for all)

Next, download the professional version of PyCharm here. You need to apply for an education license.

PyCharm requires configuration for each project. In the test for this section you will practice configuring a Python Environment.

In PyCharm, open the HW0 Test subfolder of the homework folder, it will open up a new project window. Click on the HW0PyCharmTest.py file. The Anaconda installation should make Python 3.8 the default interpreter for your projects. If you get a message about missing an interpreter or cannot find the run option, please follow online instructions to configure a Python interpreter for the project.

Follow the instructions and replace your uni in the Python file where specified. Right click on the HW0PyCharmTest.py file, either on the tab or anywhere in the file's body. Click "Run HW0PyCharmTest" to execute. Take a screenshot of your window and embed it into the test ipynb file provided where specified. There is a Markdown cell. You just need to replace the name of the file in Image("./PyCharmScreenshot.png").

If your file will not run because it is missing a configuration or interpreter follow the instructions below, otherwise move onto the next section.

To configure the interpreter, click "Add Configuration" in the top right corner then the "+" button. Scroll down and select Python. You can name this configuration how you choose, I suggest naming it the Python version you use. Select the appropriate Python version in the Python Interpreter field then click OK.

MySQL Server Community Edition

You must install MySQL (Server) Community. There is a download link and you can select your operating system. The installation is relatively easy. There are
- Instruction for Windows. Use the MySQL Installer method and choose the Developer Default.
- Instructions for Mac.
- Linux.

At some point, you will be promoted for/have to set login/authentication options.
- Write down and remember the *root* user ID and *password.*
- Choose the Legacy Authentication method.

Installing MySQL registers MySQL Server as a service. It should start automatically. If you are ever unsure if MySQL Server is running, there are online OS specific instructions for determining status, starting and stopping the server.

DataGrip

DataGrip will be the GUI (Graphical User Interface) so you can visualize and modify the data tables on your AWS server. You need to apply for an education license from JetBrains if you didn't already do so in the PyCharm setup.

Download DataGrip and setup. You will need the education license so you can use DataGrip beyond your free trial period.

You will need to set up a connection in DataGrip to your local MySQL Sever. There are online instructions for configuring a connection. In addition to the general instructions, there are MySQL specific instructions.

Once DataGrip is installed, you need to install two databases. You do this by running 3 SQL scripts that are in the HW 0 folder:
- DDL.sql
- smallRelationsInsertFile.sql
- lahmans-mysql-dump.sql

You will run the SQL scripts in that order. The DataGrip documentation provides instructions on how to run an SQL script from a file (hard drive). Follow the instructions for each of the SQL files.

When your Lahman database has been successfully loaded, copy, paste, then run the following query in the console:

SELECT * FROM lahmansbaseballdb.People LIMIT 100;

Don't worry about what this query means for now, we will cover this in class.

You should get an output similar to the screenshot below. Take a screenshot of your window and insert your screenshot in the appropriate section of the .ipynb provided. You need to save the screenshot to a file in the HW 0 directory and replace the file name in the DataGrip section on the notebook.

Example screenshot:
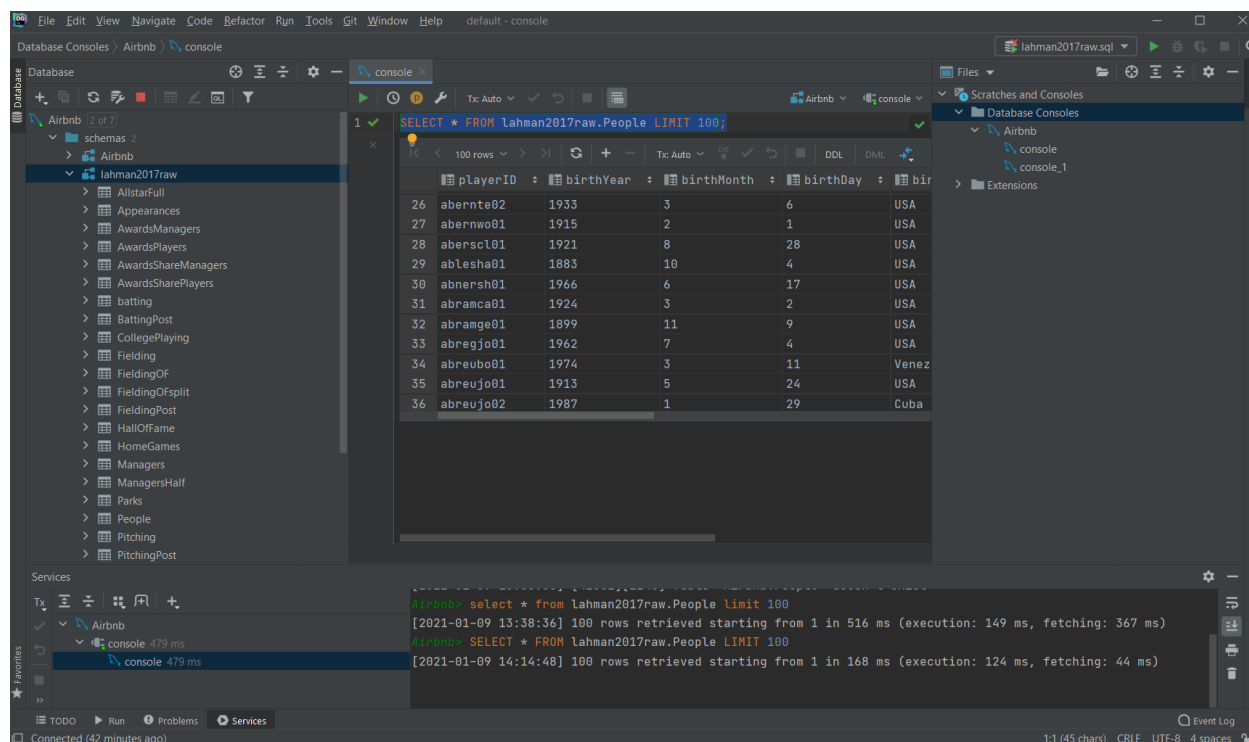
You will also need to run some code cells that connect your jupyter notebook to your Lahman database.

Postman (Programming Track only)

Postman is an API development tool that can be downloaded here. We will not be getting too deep into API development as this is a Databases course, but some homeworks will require the use of HTTP requests and Postman is a much friendlier environment to be working in to do such a thing.

After downloading and creating an account, create a new request by either clicking the "+" button in the top left corner or by selecting "Create a request" on the Launchpad tab.

Postman allows you to easily see the raw html from a webpage. In the screenshot below, a GET request was sent to the www.columbia.edu webpage.

Example screenshot:
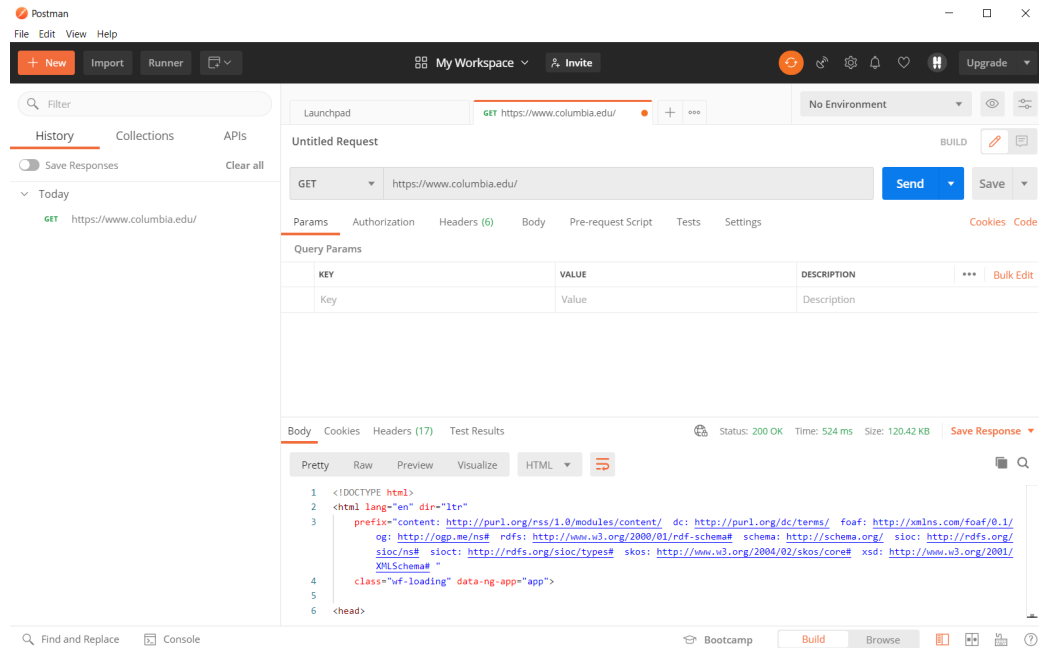
Take a screenshot like in the example of the results of a GET request to a website of your choice (school appropriate and different from www.columbia.edu). Insert this screenshot into the appropriate place in the .ipynb.

Once you have completed these steps and tests you are done!