For my hardware configuration, I am currently running on a Windows 11 device. The CPU I have is an AMD Ryzen 7 7700X 8-Core Processor. My computer has 32GB of RAM and has a 1TB SSD.

For this homework assignment the technology stack that I decided to use was Python and MySQL. I created the schema and the tables with the required attributes for the ‘tweets’ and ‘followings’ tables inside of MySQL. One important thing that I implemented was a trigger that would add the current datetime of when a tweet would be inserted into the tweet table.

In the main class is where I did the performance testing for the given data, which I stored inside of a pandas dataframe. My approach to inserting it into the tweet table was iterating each row of the dataframe, then creating a tweet object and passing it to the postTweet funciton and I used the profiler to calculate the calls per second. My approach to measuring the calls per second for getting a random user’s timeline was by counting the number calls that were made in 5 seconds and dividing to get calls per second.

My results for the performance of my code would not be able to keep up with the demand for twitter as it was only able to insert 356 tweets per second into the database. When recieving a random user’s timeline, my program was able to make 55 calls to the getTimeline function per second which also would not be able to handle the amount of worldwide refereshes of twitter.

Some factors that I believe impacted my results are the software stack that I decided to use, and possibly even the way that I implemented the code. One change I would’ve made if there was no paramter constraint for postTweet was instead of doing self.con.comit() every iteration, we only do it once at the end, since we will have to call comit a million times versus one time.