For this assignment, I used MySQL 8.0.28 and Python 3.10. I used the default storage engine. My computer is a 2020 Macbook air with the Apple M1 chip and 16 GB RAM. I batched tweets in sets of 5 and added a secondary index in the tweets table of my database on user_id. The only specific library I used was numpy.

API Method	**API Calls Per Second**
read_load_tweets	28247.0
load_timeline	346.9

There are several factors that I believe contributed to the large number of tweets I was able to load per second. First, I did not commit the tweets until the very end. This is not the best strategy for isolating separate tweets, but it greatly increased the efficiency of loading in tweets. Additionally, I found that my code was most efficient when I batched 5 tweets together for each insert.

There are a few factors that contributed to the decent number of home timelines I was able to load per second. I had a secondary index in the tweets table on user_id, which greatly increased the efficiency of querying for tweets by a set of users. Also, to avoid using a join, I separated my original query into two queries: one to get user ids of users the input user follows and one to get the ten most recent tweets by those users. Separating my original query improved the efficiency of my code; although, I do not think this is the best strategy for if the database had significantly more users.

On a more general note, I tried to use list comprehensions anywhere I could to improve efficiency.

Finally, to get my code to run, you have to update the authenticate.txt file to have your user on the first line and their password on the second line.