

Sophie Bailard/Seth Friman

Professor Rachlin

DS 4300

02/02/2022

Twitter_RDB

Software Stack

- 1) RDB and Version – PostgreSQL v. 14, using psql
- 2) Programming language – python
- 3) Python Packages used
 - a. Main.py: datetime, sys, time, random, pandas, psycopg2
 - b. dbConnect.py: psycopg2
 - c. TweetAPI.py: abc
- 4) Other files
 - a. Csv files included = data and sample data
 - b. twitterDBPostgres.txt = database username and password

Hardware Configuration

- 1) CPU speed
 - a. 2.3 GHz
- 2) Number of cores
 - a. 2
- 3) Ram
 - a. 8 GB
- 4) Disk

- a. 41.97 GB available out of 250.69 GB total

API calls per seconds

API Method	API Calls/Second
postTweet	6297.5
getHomeTimeline	363.8

Describe any factors you think might impact your results:

We noticed that our code method seems to have a greater number of API Calls/Second compared to other classmates -- even those with similar hardware configurations. One reason for this could be that our team used PostgreSQL as opposed to MySQL or Microsoft SQL Server. We also used indexing which significantly sped up our results.