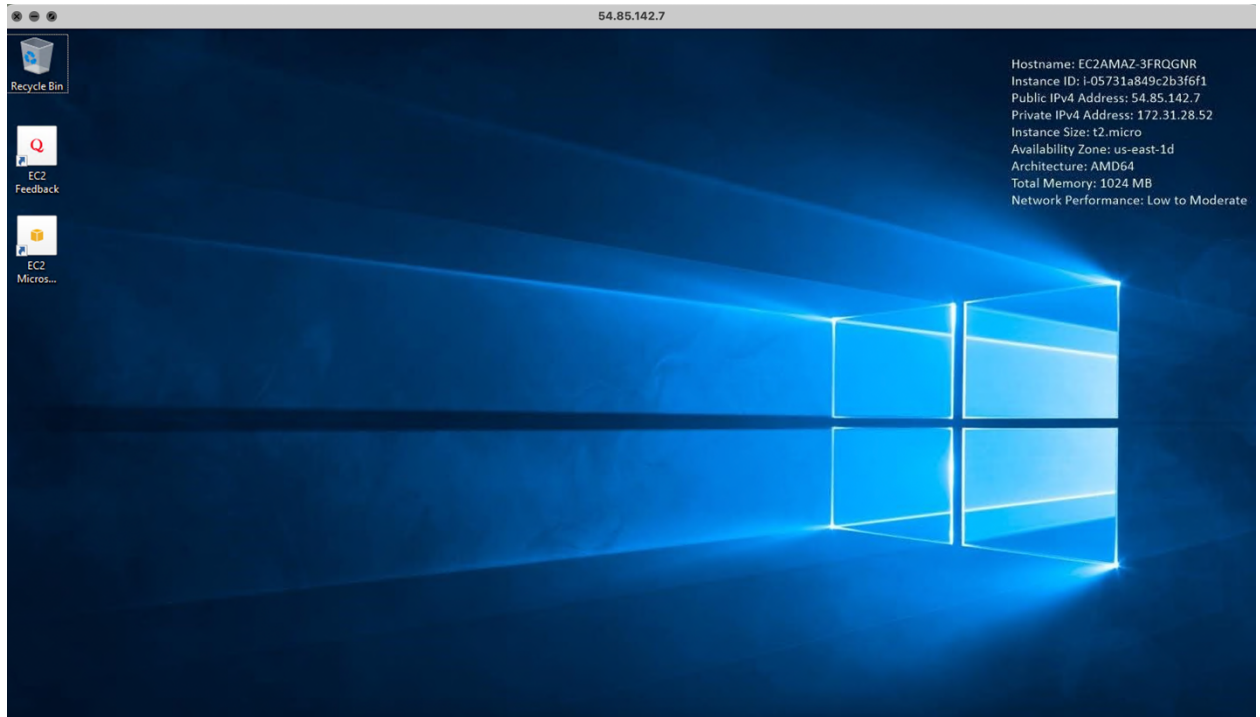


Lab-4 Answers

1. The advantages of EC2 include being able to utilize a window's desktop on a mac. It also allows for faster processing of big data and resulting in efficient results. Along with this, it allows for an automatic backup of your work. The disadvantages I can recognize so far would be the costs of AWS.

























2.



3. Amazon EBS (Elastic Block Store) allows for storage in large volumes for EC2 instances. It allows for storage with databases and log processing. S3 works well for storing images, analytics and object-level data. However, the differences arise in the accessibility of this data. Through EBS, this is not accessible through the internet or multiple devices. S3 is accessible through the internet and on any machine. Creating an additional EBS in your virtual machine can be useful as they are replicated in servers and are very safe to use.
4. SQL is typically used for databases which are stored into tables, NoSQL is usually fitted with key-value pairs to provide organized databases such as, JSON documents. NoSQL is implemented in MongoDB while SQL can be used in Oracle, MySQL, etc. NoSQL can allow for the use of large volumes of data with queries that can be used to find said data more easily.
5. We utilized tweepy and pymongo. [Tweepy](#) is a python library that allows use of the Twitter API. It allows for sets of classes that give access to twitter models through classes within python. [Pymongo](#) is a python library which allows and includes tools to work with mongoDB. It basically serves as a driver to allow access to databases through python.
6. MongoDB is a database program that works directly with NoSQL. MongoDB works with large volumes of data which sorts data into documents such as JSON files. It is known to

be very flexible and easy to learn. 3 features that I really like about MongoDB is that it is very dynamic and doesn't need a defined schema before and can be created instantly. It also utilizes real-time insights and allows for flexibilities within deployment. What I don't like about it is that it does not implement transactions so it can lead to corrupted data, there are no relational databases so it can be difficult to link data, and upon further research there could be duplications of data.

7. My selected keyword was "bunny" and it returned 50 tweets.

 633b3acd88442...	 1577020868978...	 RT @sanbenitot: cuando bad bunny dijo "mi mundo esta jodi...	 1533673035533...	 1533673035533...	 imkriistvl	 @	 { 4 fields }
 633b3acd88442...	 1577020866705...	 RT @tourbadbunny: Bad Bunny Day is now OCT 1st 🇵🇷🇵🇷🇵🇷 ...	 1528049017304...	 1528049017304...	 Mr_Diego69	 Diego	 { 4 fields }
 633b3acd88442...	 1577020866084...	 RT @painlandpalace: whos excited for live action tiger &...	 1305345270490...	 1305345270490...	 MilesMaroon	 noah @ octob...	 { 4 fields }

Unfortunately, my results had a few strange results, I'm not sure how to filter that but perhaps using another keyword would be better. I knew there was a bad bunny concert in Los Angeles so it would probably return a lot of tweets, but it hadn't!