1. Sign in to AWS RDS & create a new MYSQL DB instance
2. Instance identifier-mydbinstance1
3. Master Username:karthik
4. Pwd:k12345678
5. Username:root
6. Pwd: (no root pwd)
7. DBName-mycitydb
8. Security group-create new security group
9. Default values
10. Link: mydbinstance1.com52r4wudge.us-west-2.rds.amazonaws.com:3306
11. Use putty to connect to RDS(EC2-Karthik)
12. sudo apt-get install mysql-client-core-5.6
13. https://www.youtube.com/watch?v=vE2E7cYJ0IQ
14. Change RDS security group to connect from anywhere
15. ubuntu@ip-172-31-24-20:~$ mysql -h mydbinstance1.com52r4wudge.us-west-2.rds.amazonaws.com -P 3306 -u karthik –p
16. ubuntu@ip-172-31-24-20:~$ mysql -u karthik -p -h mydbinstance1.com52r4wudge.us-west-2.rds.amazonaws.com
17. https://forums.aws.amazon.com/message.jspa?messageID=346105
18. enter password: k12345678
19. mysql> show databases;
20. mysql> use mycitydb;
21. mysql>create table cityinfo(

country varchar(100),

city varchar(100),

accentcity varchar(100),

region int,

population int,

latitude Decimal(9,6),

longtitude Decimal(9,6)

);

1. <https://www.maxmind.com/en/free-world-cities-database>
2. Download worldcitypop.txt
3. rename worldcitypop.txt to cityinfo.csv & ftp via filezilla to EC2 /home/Ubuntu
4. In EC2 enable file permissions

The my.cnf file you should edit is the /etc/mysql/my.cnf file. Just:

sudo nano /etc/mysql/my.cnf

Then add:

[mysqld]

local-infile

[mysql]

local-infile

<http://stackoverflow.com/questions/10762239/mysql-enable-load-data-local-infile>

1. Import the csv file into mycityinfo table

LOAD DATA LOCAL INFILE '/home/ubuntu/cityinfo.csv' INTO TABLE mycitydb.cityinfo FIELDS TERMINATED BY ',' ENCLOSED BY '"' LINES TERMINATED BY '\n' IGNORE 1 ROWS;

3170000+ imported

1. select \* from cityinfo limit 1000;select \* from cityinfo limit 1000;

<http://stackoverflow.com/questions/19827388/mysql-select-top-n-max-values>

1. connect also via mysql workbench
2. Add index to the table

alter TABLE cityinfo ADD INDEX (city);

SHOW INDEXES FROM cityinfo

alter TABLE cityinfo ADD INDEX (city);

SHOW INDEXES FROM cityinfo

alter TABLE cityinfo ADD INDEX (latitude);

SHOW INDEXES FROM cityinfo

alter TABLE cityinfo ADD INDEX (longtitude);

alter TABLE cityinfo ADD INDEX (country);

drop index country on cityinfo

# EC2 instance setup

1. Link: http://ec2-52-35-186-45.us-west-2.compute.amazonaws.com/
2. Sudo pip install pymysql
3. Sudo pip install python-memcached

References:

1. <https://www.maxmind.com/en/free-world-cities-database>
2. <http://memcached.org/>
3. <https://en.wikipedia.org/wiki/Memcached#Example_code>
4. <https://www.datacamp.com/community/tutorials/machine-learning-in-r>
5. <http://stackoverflow.com/questions/3635166/how-to-import-csv-file-to-mysql-table>
6. <http://stackoverflow.com/questions/159255/what-is-the-ideal-data-type-to-use-when-storing-latitude-longitudes-in-a-mysql>

SELECT city, latitude, longitude, 3956 \* 2 \*

ASIN(SQRT( POWER(SIN(((select latitude from city.CityInfo2 where city='mumbai') - latitude)\*pi()/180/2),2)

+COS((select longitude from city.CityInfo2 where city='mumbai')\*pi()/180 )\*COS(latitude\*pi()/180)

\*POWER(SIN(((select longitude from city.CityInfo2 where city='mumbai')-longitude)\*pi()/180/2),2)))

as distance FROM city.CityInfo2 WHERE

longitude between ((select longitude from city.CityInfo2 where city='mumbai')-10/cos(radians((select latitude from city.CityInfo2 where city='mumbai')))\*69)

and ((select longitude from city.CityInfo2 where city='mumbai')+10/cos(radians((select latitude from city.CityInfo2 where city='mumbai')))\*69)

and latitude between ((select latitude from city.CityInfo2 where city='mumbai')-(10/69))

and ((select latitude from city.CityInfo2 where city='mumbai')+(10/69))

having distance < 10 ORDER BY distance limit 100