

Load Data into MYSQL RDS Instance

Author:	Sri Govind Gutala
Environment:	Amazon Web Services
Date:	September 5, 2016
Document:	MySQL_RDS_Creation_In_AWS.docx

Table of Contents

1. Required Prior Knowledge.....	3
2. Create MYSQL RDS Instance on AWS	3
Step 1: Open AWS Console	3
Step 2: Launch DB Instance	3
Step 3: Specify DB Details.....	4
Step 4: Configure Advanced Settings.....	5
3. Connect to MySQL RDS Instance using MySQL Workbench.....	8
Step 1: Create a connection	8
Step 2: Provide Connection details.....	9
Step 3: Open MySQL RDS Instance	11
Step 4: Download Dataset	12
Step 5: Table Creation	13
Step 6: Load data into table	14
Step 7: Verify data in table	15

1. Required Prior Knowledge

You are required to have a basic knowledge on AWS console and its basic components such as IAM, RDS, VPC, creating roles, policies, VPC security groups, etc.

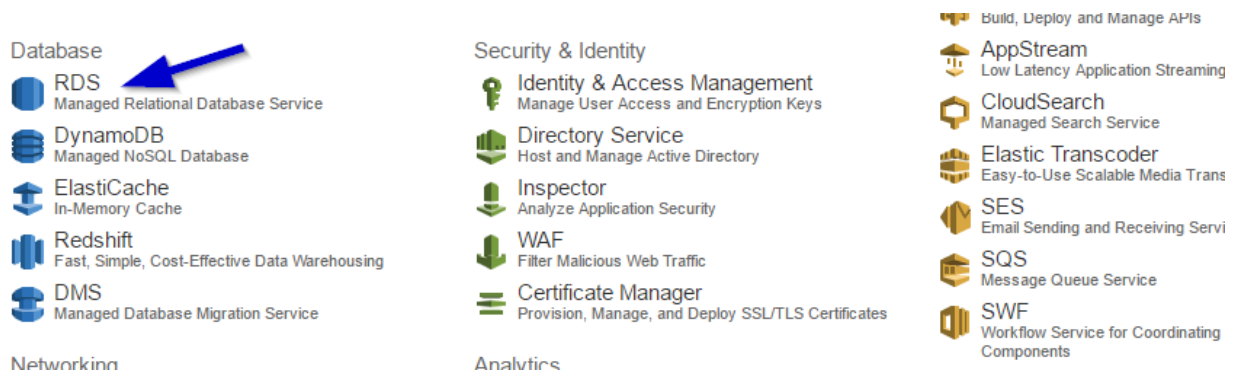
- RDS is a cloud based Relational Database Service delivered by AWS. It is easy to set up, operate and scale.
- In the case of failure, it ensures a safe and quick recovery.
- You can launch a DB instance and get access to a fully featured relational database without worrying about DBA tasks like backups, patch management and security management.
- Using AWS RDS, you can easily achieve high availability and high redundancy.

This document provides information on how to set up your RDS Instance and load data into it. I'm using AWS free tier membership and using micro instances, which are available for free tier. Lets get started!

2. Create MYSQL RDS Instance on AWS

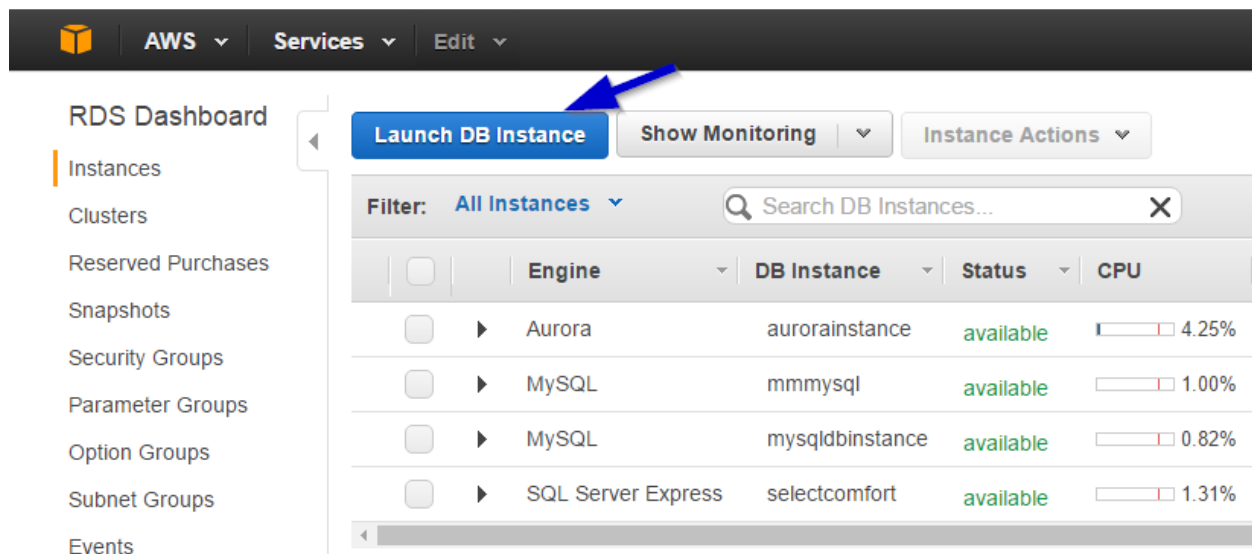
Step 1: Open AWS Console

Login to AWS account with your credentials and select RDS under Database Section

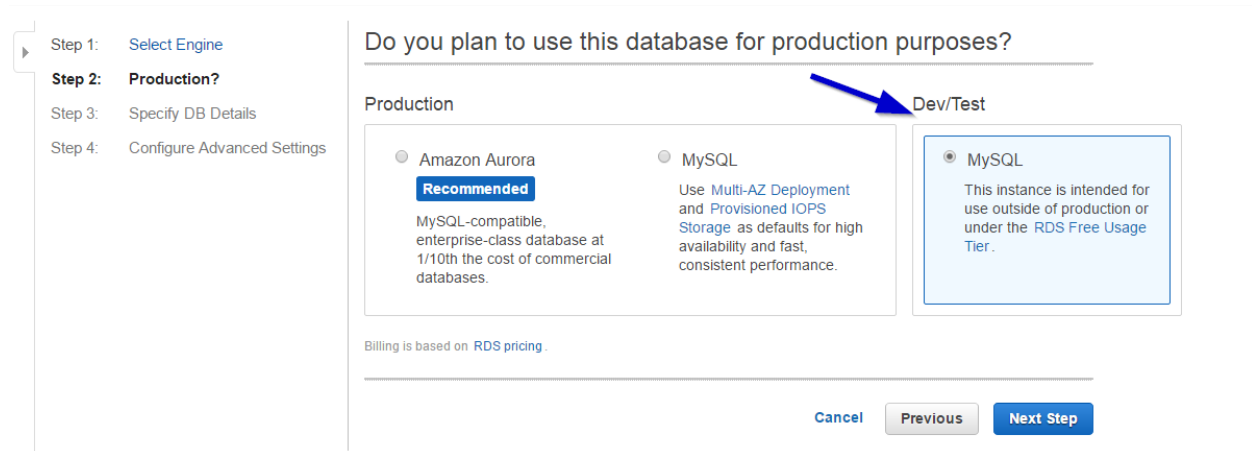


Step 2: Launch DB Instance

Click on **Instances** that is on left side of the page and click **Launch DB Instance**.



Select MySQL as engine and click on “**Select**” button. Select Database for **Dev/Test** as shown in below figure and click on “**Next Step**”.



Step 3: Specify DB Details

Specify the DB Details as shown in below figures and click on **Next Step**.

Figure 1

Step 1: [Select Engine](#)

Step 2: [Production?](#)

Step 3: Specify DB Details

Step 4: [Configure Advanced Settings](#)

i Your current selection is eligible for the free tier.

[Learn More.](#)

i Estimate your monthly costs for the DB Instance using the [RDS Instance Cost Calculator](#).

Specify DB Details

Free Tier

The Amazon RDS Free Tier provides a single db.t2.micro instance as well as up to 20 GB of storage, allowing new AWS customers to gain hands-on experience with Amazon RDS. Learn more about the RDS Free Tier and the instance restrictions [here](#).

☐ Only show options that are eligible for RDS Free Tier

Instance Specifications

DB Engine

mysql

License Model

general-public-license ▼

DB Engine Version

5.6.27 ▼

... Review the **Known Issues/Limitations** to learn about potential compatibility issues with specific database versions.

DB Instance Class

db.t2.micro — 1 vCPU, 1 GiB RAM ▼

Multi-AZ Deployment

No ▼

Storage Type

General Purpose (SSD) ▼

Allocated Storage*

5

GB

Figure 2

Settings

DB Instance Identifier*

MySQLRDSWebinar

Master Username*

rdsuser

Master Password*

.....

Confirm Password*

.....

Retype the value you specified for Master Password.

* Required

[Cancel](#)
[Previous](#)
[Next Step](#)


Step 4: Configure Advanced Settings

Provide VPC and corresponding VPC security group that you have created already. And select a database name as “**trafficdb**”.

- Step 1: [Select Engine](#)
Step 2: [Production?](#)
Step 3: [Specify DB Details](#)
Step 4: Configure Advanced Settings

Configure Advanced Settings


Network & Security

 VPC*


Subnet Group

Publicly Accessible

Availability Zone

VPC Security Group(s) 

Database Options

 Database Name

Note: if no database name is specified then no initial MySQL database will be created on the DB Instance.

Database Port

DB Parameter Group

Option Group

Copy Tags To Snapshots ☐

Specify a string of up to 64 alpha-numeric characters that define the name given to a database that Amazon RDS creates when it creates the DB instance, as in "mydb". If you do not specify a database name, Amazon RDS does not create a database when it creates the DB instance.

Leave other parameters with default values and click on **“Launch DB Instance”**

Copy Tags To Snapshots ☐

Enable Encryption

Backup

Please note that automated backups are currently supported for InnoDB storage engine only. If you are using MyISAM, refer to detail [here](#).

Backup Retention Period days

Backup Window

Monitoring

Enable Enhanced Monitoring

Maintenance

Auto Minor Version Upgrade

Maintenance Window

* Required

[Cancel](#) [Previous](#) [Launch DB Instance](#)

Once Launch DB Instance is clicked, it starts creating DB. Click on View Your DB Instances to view the DB Instance that was created in above step.

✔ **Your DB Instance is being created.**

Note: Your instance may take a few minutes to launch.

Connecting to your DB Instance

You will be unable to connect to your database instance unless you have previously authorized access on your chosen security group.

[Go to the Security Groups Page](#)

Related AWS Services

Amazon ElastiCache
Add a managed Memcached or Redis-compatible in-memory cache to speed up your database access.

[Click here to learn more and launch your Cache Cluster](#)

View Your DB Instances

The below figure shows that the “**mysqlrdswebinar**” DB Instance is being created. The status changes to “**available**” once database is created completely.

Launch DB Instance

Show Monitoring

Instance Actions

Filter: All Instances

Search DB Instances...

Viewing 5 of 5 DB Instances

	Engine	DB Instance	Status	CPU	Current Activity	Maintenance	Class	VPC	Multi-AZ	Replication
<input type="checkbox"/>	Aurora	aurorainstance	available	4.25%	2 Selects/sec	None	db.r3.large	vpc-a1a817c5	No	writer
<input type="checkbox"/>	MySQL	mmmysql	available	1.36%	0 Connections	None	db.t2.micro	vpc-a1a817c5	No	
<input type="checkbox"/>	MySQL	mysqldbinstance	available	1.00%	0 Connections	None	db.t2.micro	vpc-a1a817c5	No	
<input checked="" type="radio"/>	MySQL	mysqlrdswebinar	backing-up			None	db.t2.micro	vpc-a1a817c5	No	
<input type="checkbox"/>	SQL Server Express	selectcomfort	available	1.36%	0 Connections	None	db.t2.small	SelectComfortDND	N/A	

Click on the DB Instance to view connection details (Endpoint). Endpoint is used to connect to MySQL RDS instance through MySQL workbench. In the below figure, one can find database name, username and Endpoint.

▼

MySQL

mysqlrdswebinar

available

1.00%

0 Connections

None

db.t2.micro

vpc-a1a817c5

No

Endpoint: mysqlrdswebinar.cs1iw4nkkips.us-west-2.rds.amazonaws.com:3306 (authorized) ⓘ

Configuration Details

Security and Network

Instance and IOPS

ARN

arn:aws:rds:us-west-2:167270772459:db:mysqlrdswebinar

Engine

MySQL 5.6.27

License Model

General Public License

Created Time

September 13, 2016 at 12:39:23 PM UTC-7

DB Name

trafficdb

Username

rdsuser

Option Group

default:mysql-5-6 (in-sync)

Parameter Group

default:mysql5.6 (in-sync)

Copy Tags To Snapshots

No

Resource ID

db-43YJNPXTH4ICFSTNSCD33MBZCQ

Availability Zone

us-west-2b

VPC

vpc-a1a817c5

Subnet Group

default (Complete)

Subnets

subnet-00f71558
subnet-7a8a311e
subnet-2af12c5c

Security Groups

Security Group Redshift (sg-ff102199)
(active)

Publicly Accessible

Yes

Writer Endpoint

mysqlrdswebinar.cs1iw4nkkips.us-west-2.rds.amazonaws.com

Port

3306

Certificate Authority

rds-ca-2015 (Mar 5, 2020)

Instance Class

db.t2.micro ⓘ

Storage Type

General Purpose (SSD)

IOPS

disabled

Storage

5 GB

Encryption Details

Availability and Durability

Maintenance Details

Encryption Enabled

No

DB Instance Status

available

Multi AZ

No

Automated Backups

Enabled (7 Days)

Auto Minor Version Upgrade

Yes

Maintenance Window

thu: 13:14-thu: 13:44

Backup Window

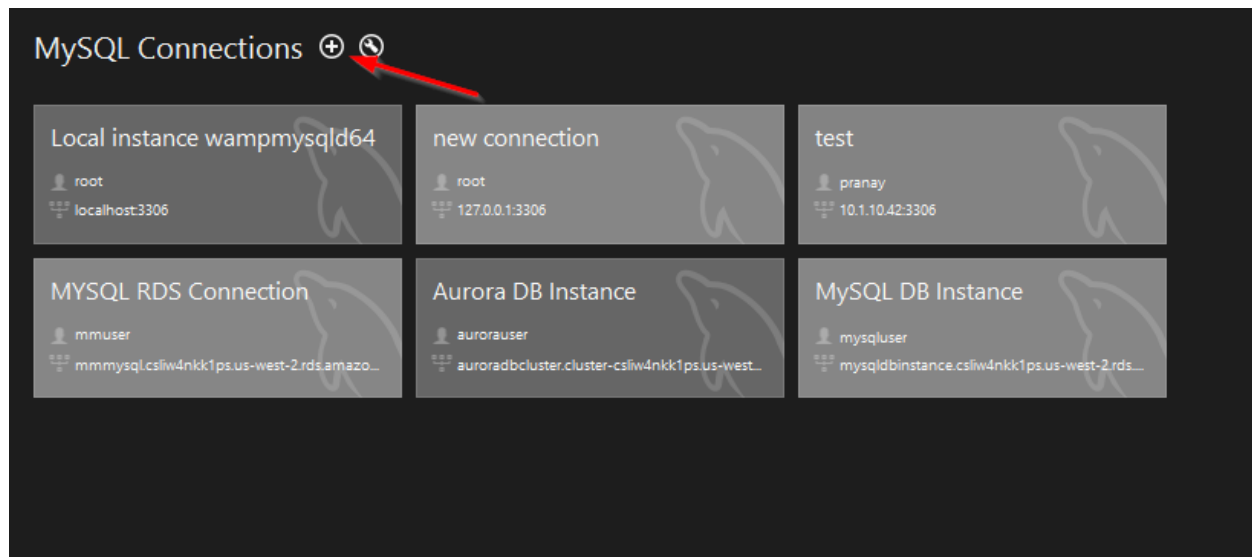
10:52-11:22

3. Connect to MySQL RDS Instance using MySQL Workbench

Step 1: Create a connection

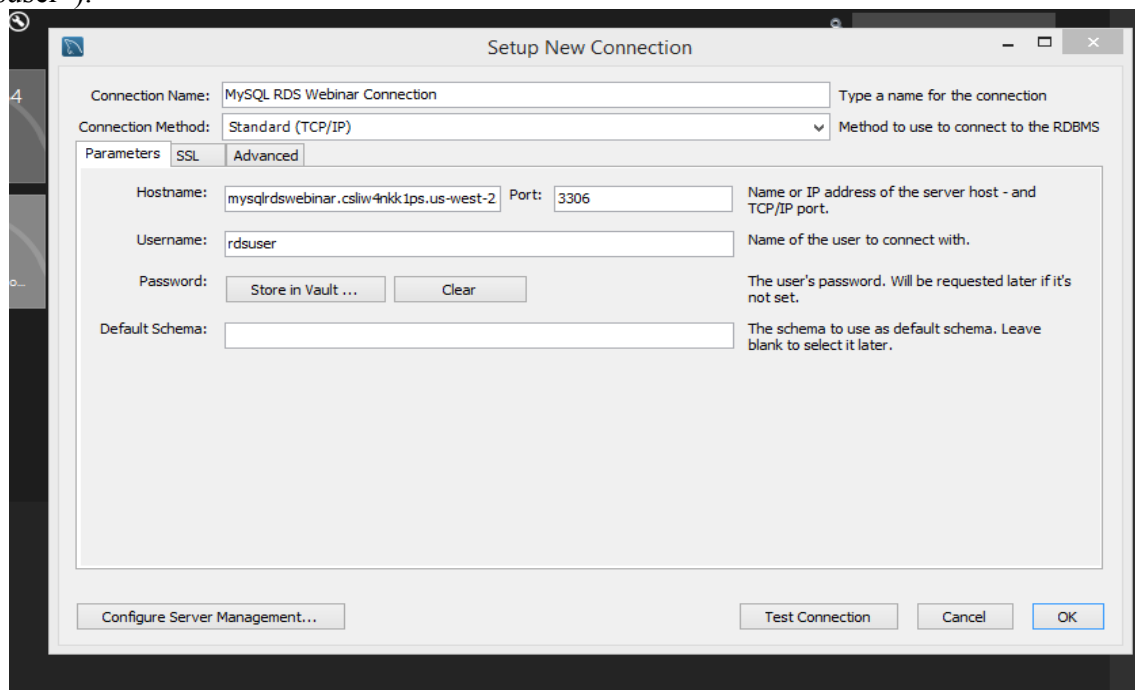
Download MySQL workbench from <https://dev.mysql.com/downloads/workbench/> .

Open MySQL workbench and click on plus button to Set up a new connection.

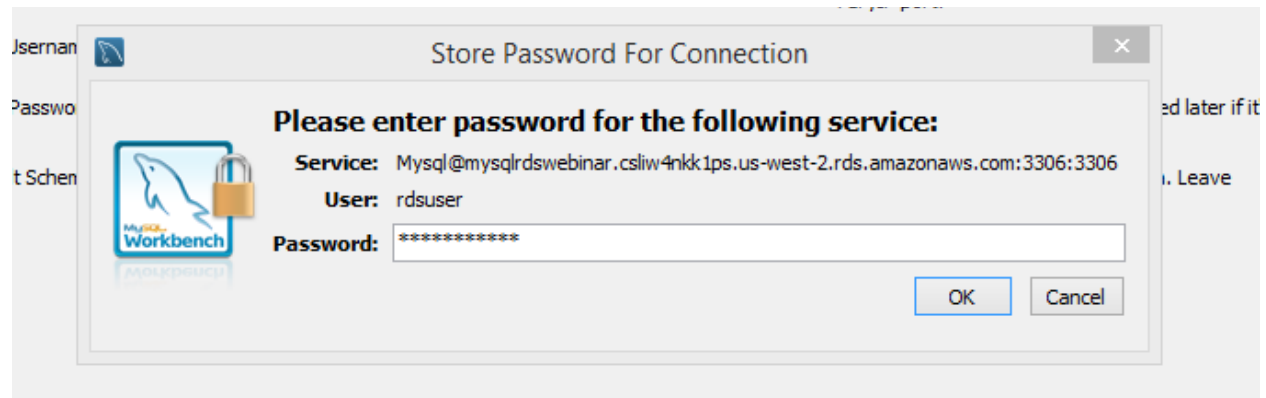


Step 2: Provide Connection details

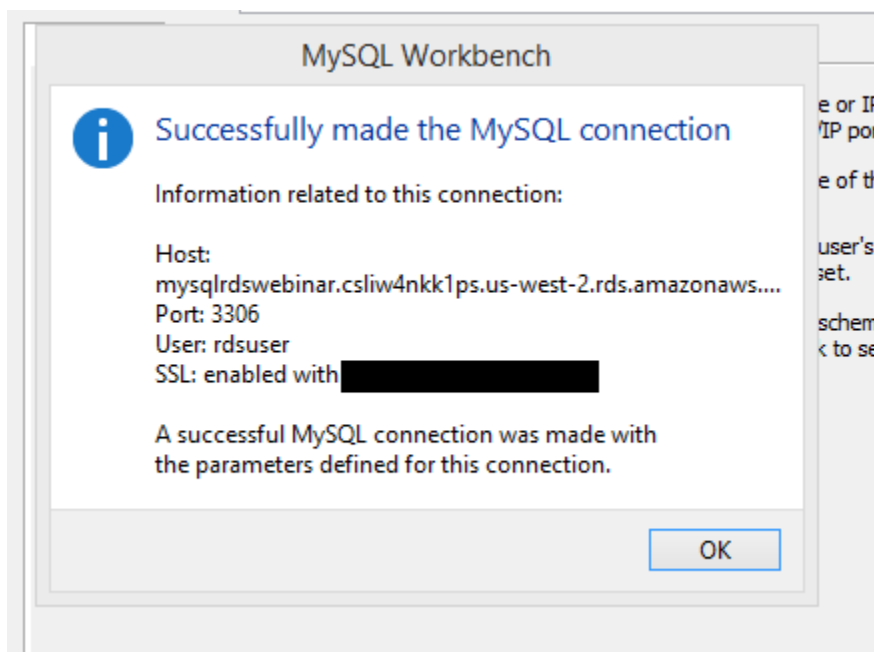
Specify endpoint for Hostname (remove port number “:3306” at the end of endpoint), Port number as 3306, username and password as provided while creating DB Instance (here it is “rdsuser”).



Click on **Store in Vault** for password. It will prompt to enter the password. Specify the password and click on **OK**.

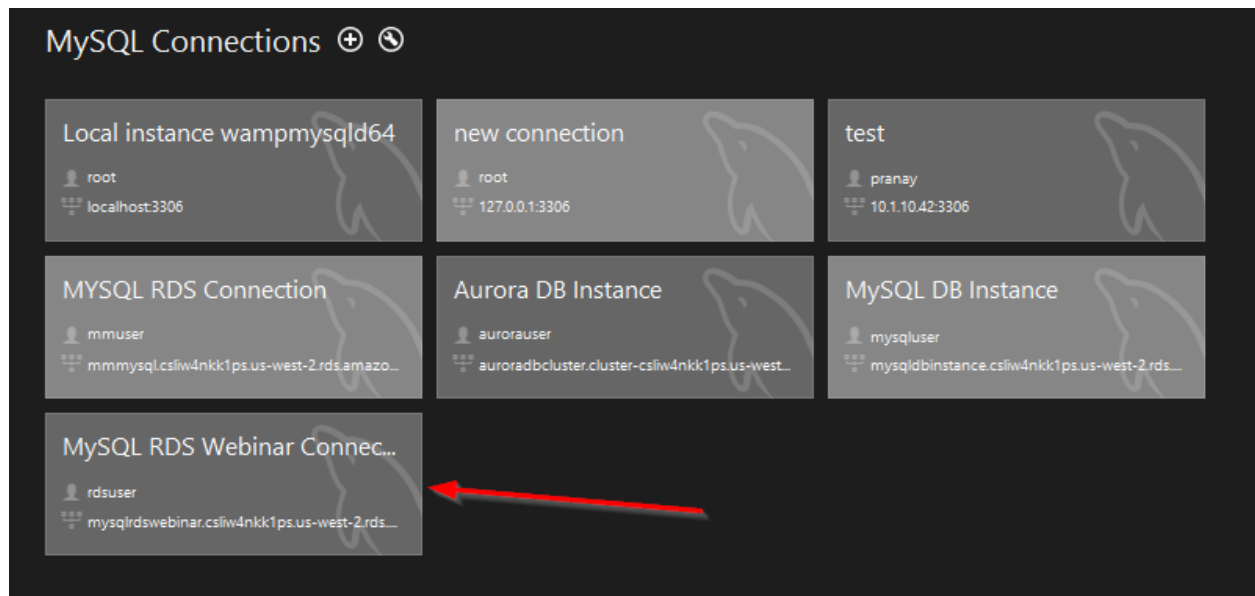


If the connection is successful, it shows the below message.

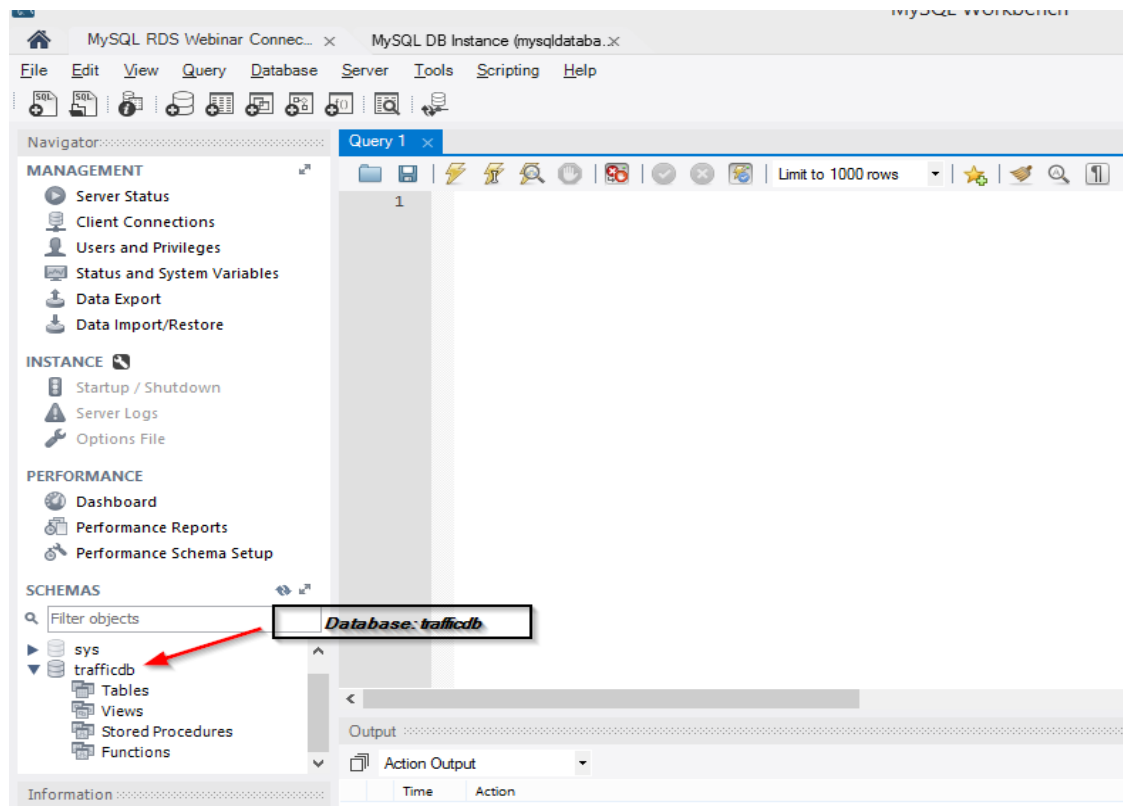


Step 3: Open MySQL RDS Instance

Open on DB Connection that was created in above step.



It opens a SQL Editor to execute sql queries.



Step 4: Download Dataset

Click on <https://catalog.data.gov/dataset/traffic-violations-56dda> and click on Download for CSV dataset.

The screenshot shows a web browser window with the URL <https://catalog.data.gov/dataset/traffic-violations-56dda>. The page is titled "Access & Use Information" and "Downloads & Resources". Under "Access & Use Information", it states: "Public: This dataset is intended for public access and use.", "Non-Federal: This dataset is covered by different Terms of Use than Data.gov.", and "License: No license information was provided." Under "Downloads & Resources", there are four download options: "Comma Separated Values File" (1721 views), "RDF File" (104 views), "JSON File" (640 views), and "XML File" (308 views). Each option has a "Download" button. A red arrow points to the "Download" button for the "Comma Separated Values File". There is also a "Landing Page" link with a "Visit page" button.

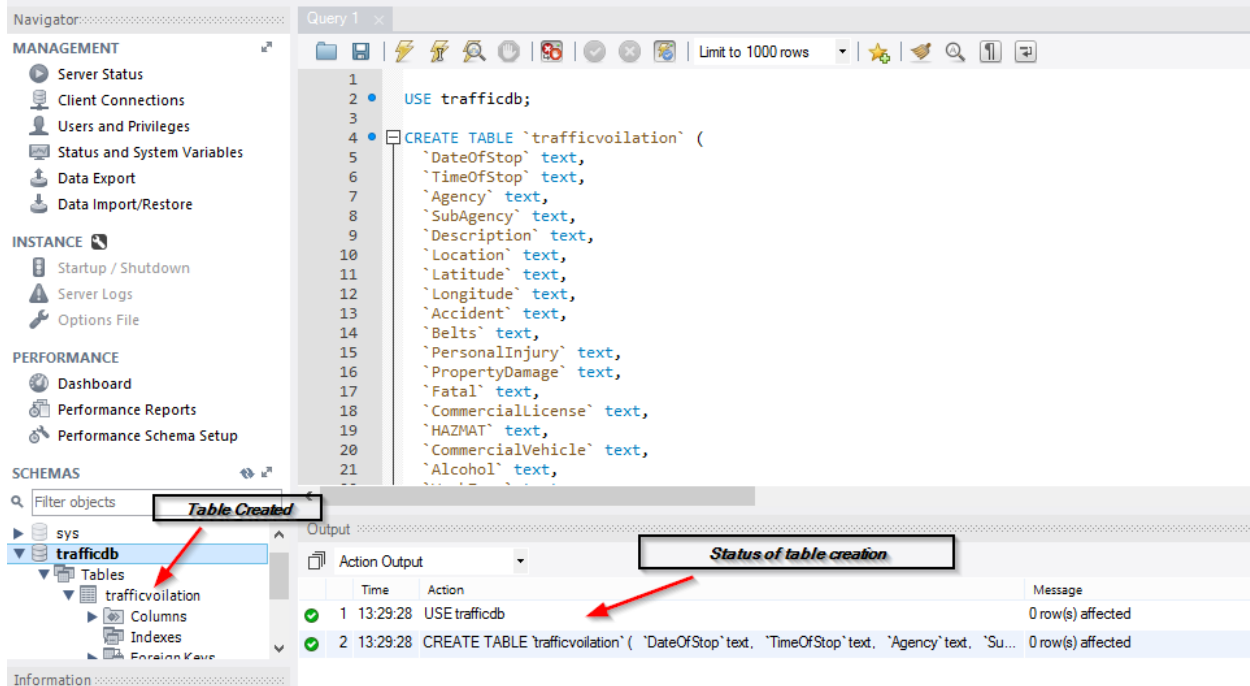
File Type	Views	Action
Comma Separated Values File	1721 views	Download
RDF File	104 views	Download
JSON File	640 views	Download
XML File	308 views	Download
Landing Page		Visit page

Step 5: Table Creation

Create table with below command on MySQL Workbench.

```
USE trafficdb;

CREATE TABLE `trafficviolation` (
  `DateOfStop` text,
  `TimeOfStop` text,
  `Agency` text,
  `SubAgency` text,
  `Description` text,
  `Location` text,
  `Latitude` text,
  `Longitude` text,
  `Accident` text,
  `Belts` text,
  `PersonalInjury` text,
  `PropertyDamage` text,
  `Fatal` text,
  `CommercialLicense` text,
  `HAZMAT` text,
  `CommercialVehicle` text,
  `Alcohol` text,
  `WorkZone` text,
  `State` text,
  `VehicleType` text,
  `Year` int(11) DEFAULT NULL,
  `Make` text,
  `Model` text,
  `Color` text,
  `ViolationType` text,
  `Charge` text,
  `Article` text,
  `ContributedToAccident` text,
  `Race` text,
  `Gender` text,
  `DriverCity` text,
  `DriverState` text,
  `DLState` text,
  `ArrestType` text,
  `Geolocation` text
) ENGINE=InnoDB DEFAULT CHARSET=latin1;
```



Step 6: Load data into table

Execute below commands to load the dataset into table created in Step 5.

```
SET autocommit=0;
SET unique_checks=0;
SET foreign-key_checks=0;
SET session character_set_database=latin1;
```

```
LOAD DATA LOCAL INFILE 'C:\\Users\\sri\\Downloads\\Traffics_Violations.csv'
INTO TABLE trafficviolation
FIELDS TERMINATED BY ',' ENCLOSED BY '"'
LINES TERMINATED BY '\n'
IGNORE 1 LINES;
```

```
SET unique_checks=1;
SET foreign_key_checks=1;
COMMIT;
```

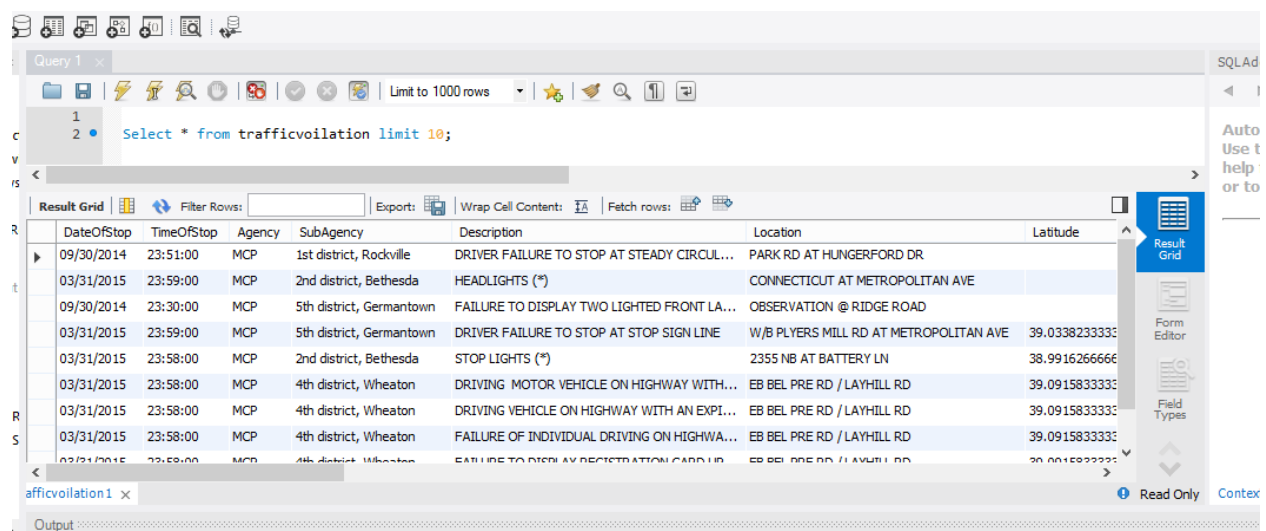
It takes about 1 to 2 minutes to load 300 MB traffic violation file.

Step 7: Verify data in table

Execute any query to view the data loaded in table.

```
Select * from trafficviolation limit 10;
```

The result of query can be seen in below image.



DateOfStop	TimeOfStop	Agency	SubAgency	Description	Location	Latitude
09/30/2014	23:51:00	MCP	1st district, Rockville	DRIVER FAILURE TO STOP AT STEADY CIRCUL...	PARK RD AT HUNGERFORD DR	
03/31/2015	23:59:00	MCP	2nd district, Bethesda	HEADLIGHTS (*)	CONNECTICUT AT METROPOLITAN AVE	
09/30/2014	23:30:00	MCP	5th district, Germantown	FAILURE TO DISPLAY TWO LIGHTED FRONT LA...	OBSERVATION @ RIDGE ROAD	
03/31/2015	23:59:00	MCP	5th district, Germantown	DRIVER FAILURE TO STOP AT STOP SIGN LINE	W/B PLYERS MILL RD AT METROPOLITAN AVE	39.0338233333
03/31/2015	23:58:00	MCP	2nd district, Bethesda	STOP LIGHTS (*)	2355 NB AT BATTERY LN	38.9916266666
03/31/2015	23:58:00	MCP	4th district, Wheaton	DRIVING MOTOR VEHICLE ON HIGHWAY WITH...	EB BEL PRE RD / LAYHILL RD	39.0915833333
03/31/2015	23:58:00	MCP	4th district, Wheaton	DRIVING VEHICLE ON HIGHWAY WITH AN EXPI...	EB BEL PRE RD / LAYHILL RD	39.0915833333
03/31/2015	23:58:00	MCP	4th district, Wheaton	FAILURE OF INDIVIDUAL DRIVING ON HIGHWA...	EB BEL PRE RD / LAYHILL RD	39.0915833333
03/31/2015	23:58:00	MCP	4th district, Wheaton	FAILURE TO DISPLAY REGISTRATION CARD UP...	EB BEL PRE RD / LAYHILL RD	39.0915833333