CS3204 - Assignment 2 Jamie Lynes 117498036

Cloud computing is the on-demand access, via the internet, to computing resources such as data storage, development tools, applications, networking capabilities, etc., hosted at a remote data centre managed by a cloud services provider.

A cloud instance is a virtual server instance in a cloud computing environment. It is built, hosted, and delivered using a cloud computing platform. A cloud computing platform is a shared pool of resources and services that can be provisioned with great speed and little effort.

AWS, provided by Amazon is an example of a cloud-computing platform.

Elastic Beanstalk is a platform within AWS that is used for deploying and scaling web applications. This platform as a service (PaaS) allows for simple deployment of web applications while provisioning the supporting architecture and compute resources required for the application to work. Elastic Beanstalk will be used as part of this assignment.

The aim of this assignment was to learn how to use a cloud instance to host a web server, a database and perform analysis on large sets of data.

Assignment Task 1: Web application:

After becoming familiar with the AWS management console and having set up several Environments on which to run sample applications, I began creating my Web Application.

Prior to this module I had no experience with Web Development and no experience coding in languages such as HTML, CSS, and JavaScript. I wanted to write the code for the application using a language I had experience with and was comfortable using. I decided to use Dash to create my Web Application. Dash is an open-source Python framework often used for building analytical web applications and interactive dashboards which allow for the sharing of insights gained from data. I felt this was an appropriate choice given my studies of Data Science and Analytics.

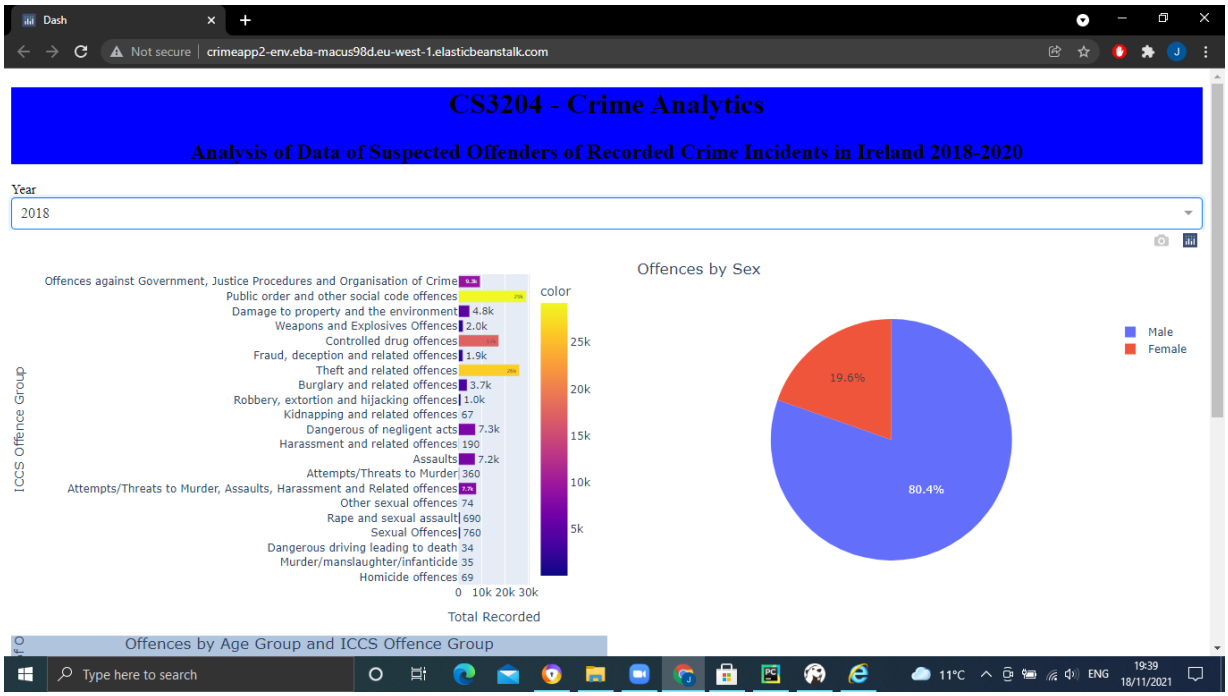
My Web Application is a very simple interactive dashboard with analyses and summarizes data of Recorded Crimes in Ireland between 2018 and 2020. The application processes three datasets taken from the Central Statistics Office:

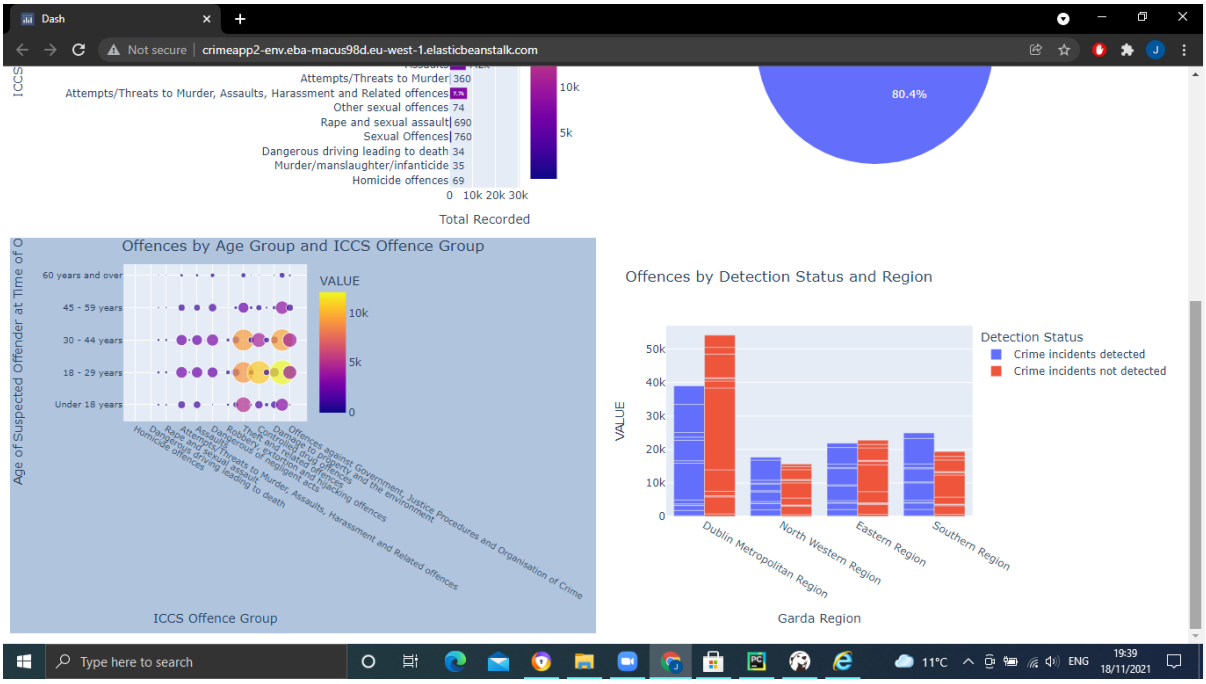
CVA08 - Suspected offenders of Recorded Crime Incidents - Under Reservation – Focus on Sex of Suspected Offender.

CVA09 - Suspected offenders of Recorded Crime Incidents - Under Reservation – Focus on Age of Suspected Offender.

RCD03 - Recorded Crime Incidents (Under Reservation) – Focuses on Detection Status and Location of Offences.

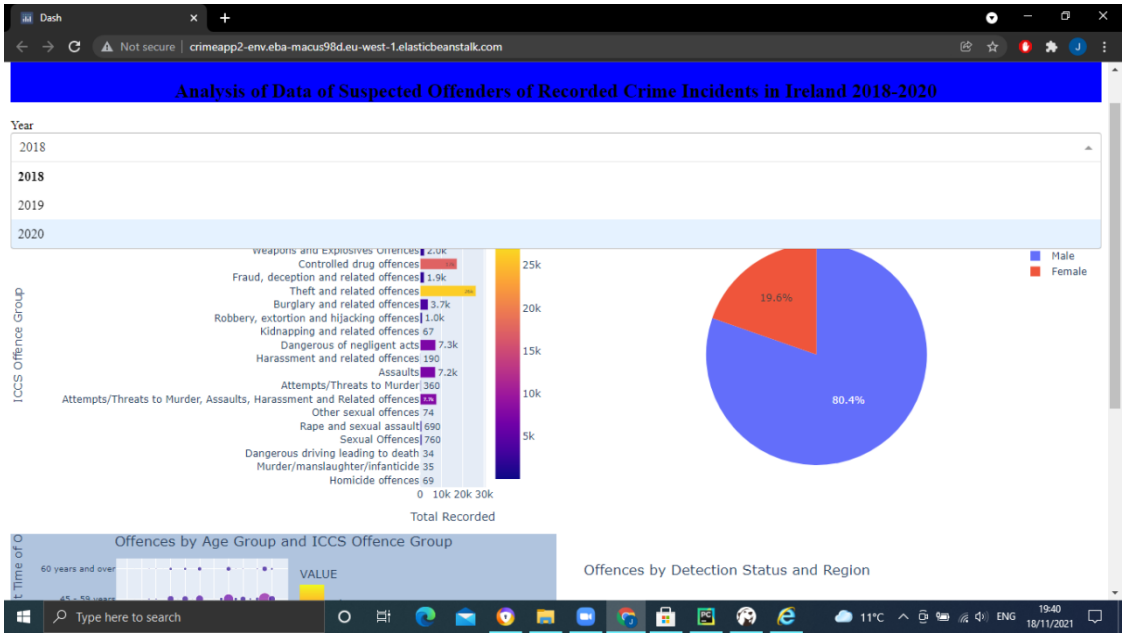
My Web Application – Analysis of Crime in Ireland 2018-2020





The Dashboard presents 4 Charts, each presenting a different insight from the data:

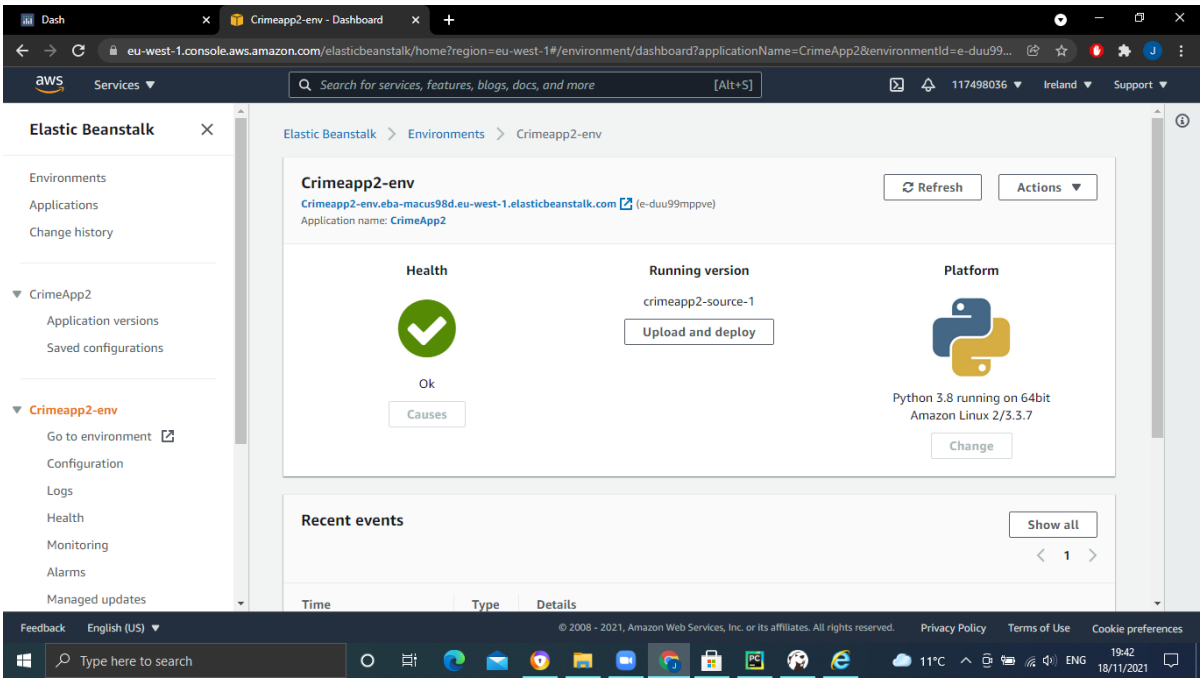
- Graph 1: Bar-chart examining Recorded Offences and their respective occurences.
- Graph 2: Pie-chart presenting the percentage of offences attributed to each sex.
- Graph 3: Scatterplot examining occurrences of Recorded Offences among each age group.
- Graph 4: Grouped Bar-chart looking at the number of offences by region, and the proportions of these offences that were Detected/resulted in a prosecution.



A dropdown menu is included that allows the user to switch between the datasets for each year. When a different year is selected, all graphs will update accordingly through the applications 'callback' methods.

After Creating the Web Application, I started the process of deploying it on Elastic Beanstalk.

I Created a text file called requirements.txt with a list of programs required by my virtual environment to run my application. I then created a Zip folder containing this text file along with my application.py file. I then created a new environment on Elastic Beanstalk (Crimeapp2-env) and uploaded my zip file to deploy my application.

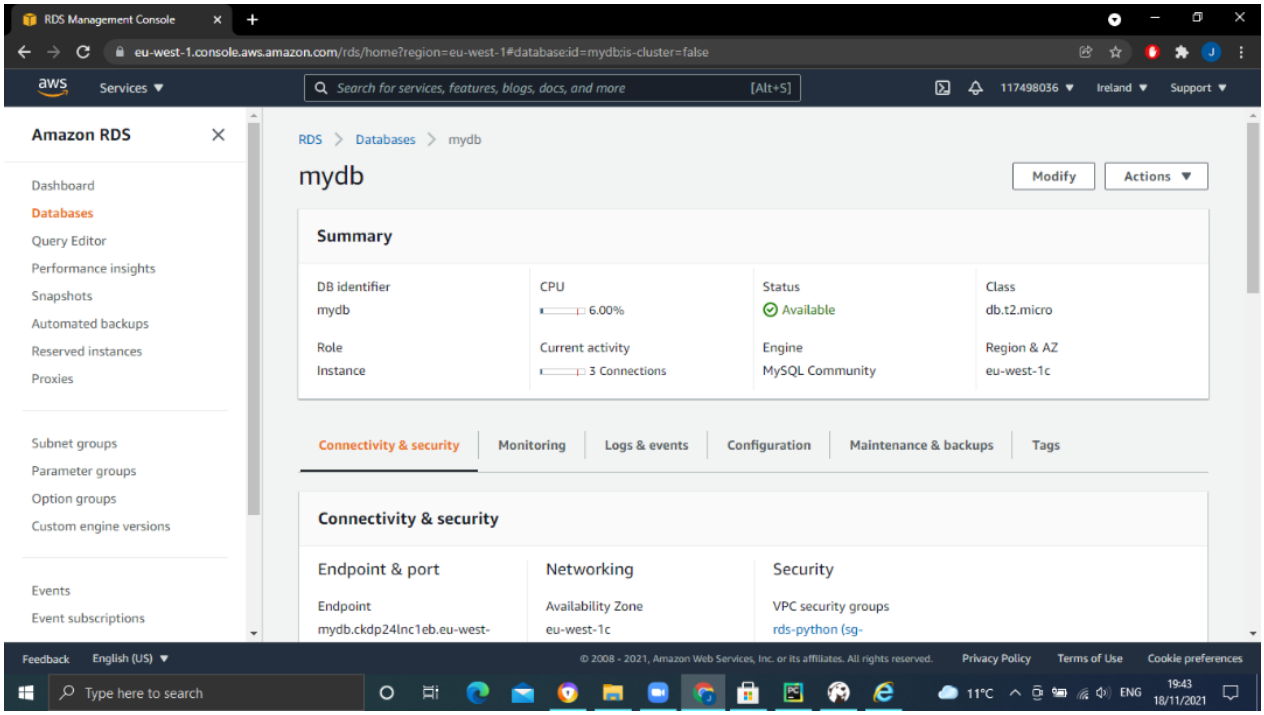


Deployment was successful, and after clicking on the URL I was able to access my Web App which I have shown already. The screenshots of my Web App earlier in the report were taken after clicking on the URL shown.

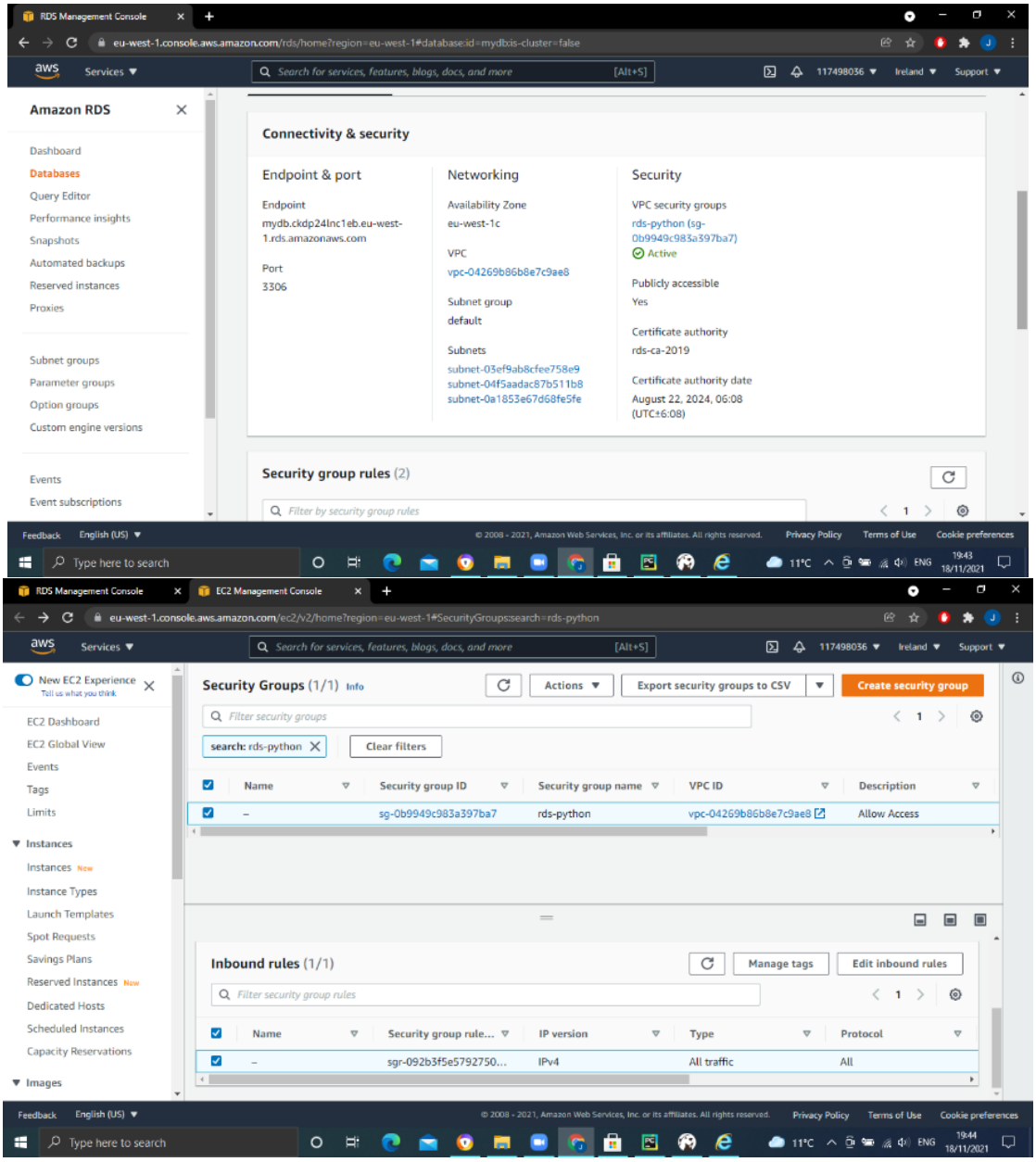
Assignment Task 2: Amazon Relational Database Service:

I edited the configuration of my Crimeapp2 environment in order to attach to it a RDS instance.

I followed the steps provided in the Lab. Guide to create this RDS instance.



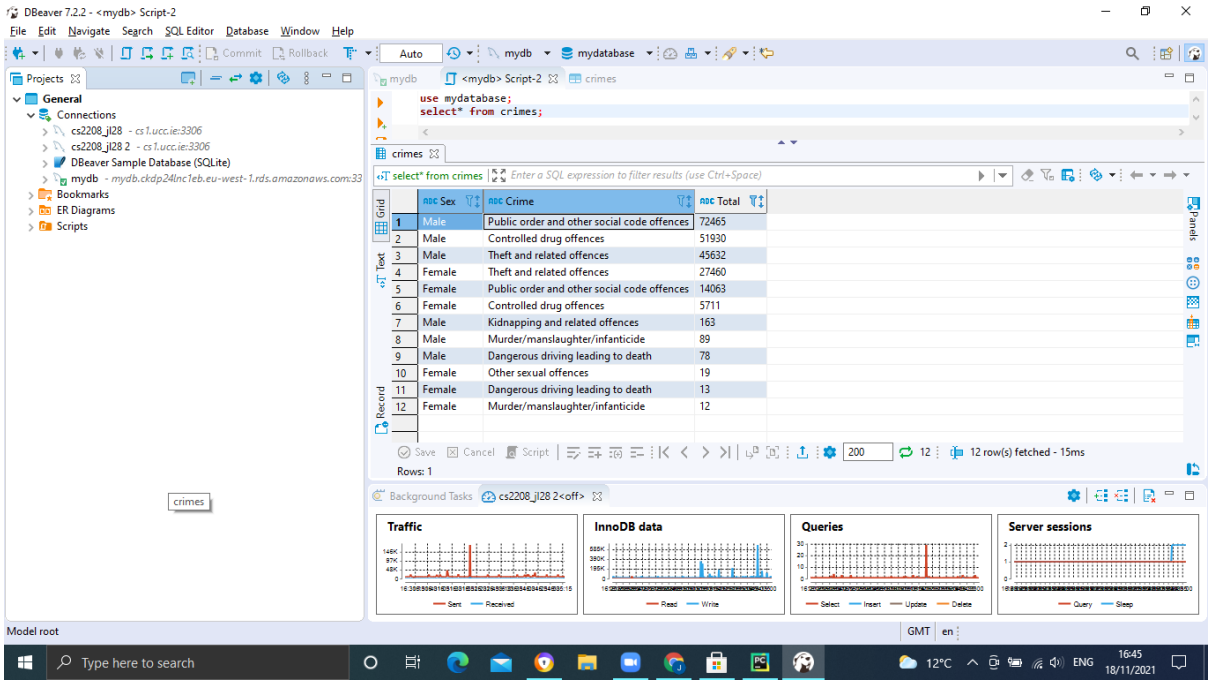
Having renamed by database to 'mydb' and having took note of the Endpoint, Port and Master Username and Password for my database, I then made come additional changes to my RDS instance to that it would be accessible to my application. For example, I deleted some VPC security groups, and changed the Inbound Rules for one Security Group so that all Inbound traffic from my IP address and my IP address only would be allowed.

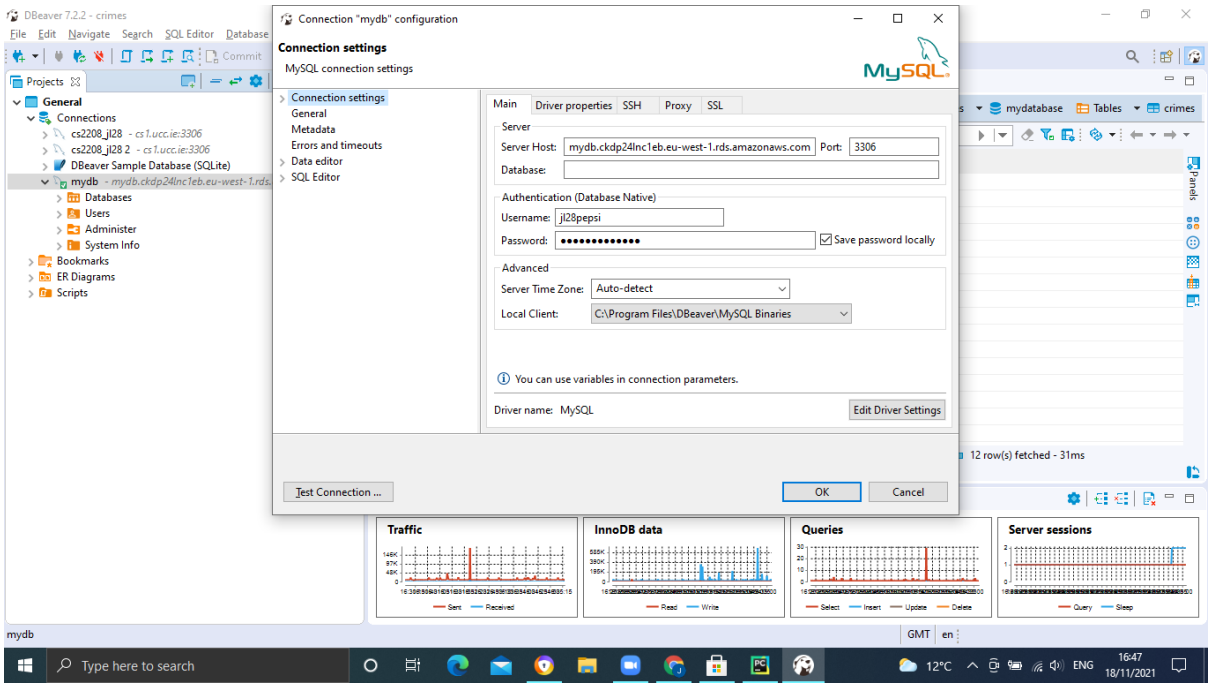


To answer Part 2 of the assignment, I processed the following dataset:

CVA08 - Suspected offenders of Recorded Crime Incidents - Under Reservation

Using the ‘Pandas’ module, I then found the three most common offences committed by each Sex in the period 2018-2020, along with the least common offences. I used a Python module named ‘PyMySQL’ in order to connect to my RDS instance from my application. I then inserted the results of my analysis from my application into a SQL table named Crimes, which is hosted on my RDS instance. Here is the SQL table created with 12 records, most common offences at the top, when viewed with DBeaver after deploying my application through Elastic Beanstalk.


DBeaver 7.2.2 - <mydb> Script-2
File Edit Navigate Search SQL Editor Database Window Help
Projects
General
Connections
cs2208_jl28 - cs1.ucc.ie:3306
cs2208_jl28 2 - cs1.ucc.ie:3306
DBeaver Sample Database (SQLite)
mydb - mydb.ckdp24lnc1eb.eu-west-1.rds.amazonaws.com:33
Bookmarks
ER Diagrams
Scripts
use mydatabase;
select* from crimes;
select* from crimes
Enter a SQL expression to filter results (use Ctrl+Space)
Sex
Crime
Total
1 Male Public order and other social code offences 72465
2 Male Controlled drug offences 51930
3 Male Theft and related offences 45632
4 Female Theft and related offences 27460
5 Female Public order and other social code offences 14063
6 Female Controlled drug offences 5711
7 Male Kidnapping and related offences 163
8 Male Murder/manslaughter/infanticide 89
9 Male Dangerous driving leading to death 78
10 Female Other sexual offences 19
11 Female Dangerous driving leading to death 13
12 Female Murder/manslaughter/infanticide 12
Save Cancel Script
200
12 row(s) fetched - 15ms
Rows: 1
crimes
Traffic
InnoDB data
Queries
Server sessions
Model root
GMT en
Type here to search
12°C ENG 16:45 18/11/2021


DBeaver 7.2.2 - crimes
Connection "mydb" configuration
Connection settings
MySQL connection settings
MySQL
Connection settings
General
Metadata
Errors and timeouts
Data editor
SQL Editor
Main
Driver properties
SSH
Proxy
SSL
Server
Server Host: mydb.ckdp24lnc1eb.eu-west-1.rds.amazonaws.com
Port: 3306
Database:
Authentication (Database Native)
Username: jl28pepsi
Password:
Save password locally
Advanced
Server Time Zone: Auto-detect
Local Client: C:\Program Files\DBeaver\MySQL Binaries
You can use variables in connection parameters.
Driver name: MySQL
Edit Driver Settings
Test Connection ...
OK
Cancel
12 row(s) fetched - 31ms
Traffic
InnoDB data
Queries
Server sessions
mydb
GMT en
Type here to search
12°C ENG 16:47 18/11/2021

Learning Outcomes/Conclusion:

- Prior to this lab, I had not only little experience with cloud computing and cloud instances, but also no experience in web development and web applications. I feel I have learned a lot through completing this assignment and would feel a lot more comfortable using some of these services and languages used in my assignment in the future, along with Web Development in general.

- I was surprised at the ease at which one could create a remote server environment and deploy an application without the worry of resources and supporting infrastructure. I gained an appreciation of the potential value that a cloud computing platform like AWS can offer a business or company.

- I also feel the experience I gained of creating an Interactive Dashboard and of processing large data sets from the Central Statistics Office website will be vital going forward in my Data Science studies and has reaffirmed my interest in pursuing a career in Data Science.