**Project Task 1**

Once this was completed, I needed to import my data into 3 RDBMS of my choosing. I choose the standard MySQL, as well as PostgreSQL and Microsoft SQL Server (also known as SQL Server). MySQL was already installed and so I was required to install the other two RDBMS. To install PostgreSQL, I made a notepad document with all the instructions. Next, I needed to install SQL Server, which proved to be a much more challenging task, and a lengthier process.

Now that all three of my RDBMS were installed and up and running, the next task was to import my data, which were in CSV format. For MySQL, it was a simple task. I created a database and then used the database. The next step was to create a table in MySQL for my data, I used different data types. Using this table, I imported by data, and made a notepad document with all the instructions. Next was PostgreSQL, which I also created a database, and then ensured it was selected and used. Similarly, I created a table, and then imported the data into the table, and I made a notepad document with all the instructions. Although PostgreSQL and MySQL share very similar syntax, there are differences for the naming of the data types, summarised in the table below. Lastly, I did the exact same process for Microsoft SQL Server. I created a database, selected and used it, and then imported data, and I made a notepad document with all the instructions. Again, the data types were different and I had to do additional research.

Now that I had all my data in all 3 of my RDBMS, I needed to import them into Hadoop HDFS, in order for analysis. Conventional methods are not suitable to handle such big data. To import the data, the tool SQOOP was used but in order to use it, I required JDBC drivers for each of my RDBMS. The MySQL JDBC driver was already installed, and so I searched online for the correct JDBC drivers for PostgreSQL and SQL Server. A variety of JDBC drivers exist for SQL Server, and the one that is required is dependent on the version of Java running in Ubuntu and HDFS. The install process was not as straightforward as one would expect, and they needed to be moved to specific directories. With all three JDBC drivers installed, I could use SQOOP to import the data into HDFS. Some errors did occur, which required me to do further research into what might cause the error and the solution for it.