**Steps to execute Spark programs:**

1. Create a new file named **Bash.sh** inside your **spark-3.3.0-bin-hadoop3** folder by typing **vi Bash.sh** in terminal
2. Copy below code and paste inside **Bash.sh** file

**export JAVA\_HOME=$(readlink -f $(which javac) | awk 'BEGIN {FS="/bin"} {print $1}')**

**if ! command -v spark-shell --version &> /dev/null**

**then**

**export PATH=$(echo $PATH):$(pwd)/bin**

**fi**

1. In terminal execute Bash.sh file using **source Bash.sh**
2. Create your program directory and write program in <program\_name>.py (names may not be the same for the files),
3. Execute Program using below commnad

**spark-submit <python\_file\_name>.py <input\_file> <output\_folder\_name>**

1. To see output

**cat <output\_folder\_name>/\***