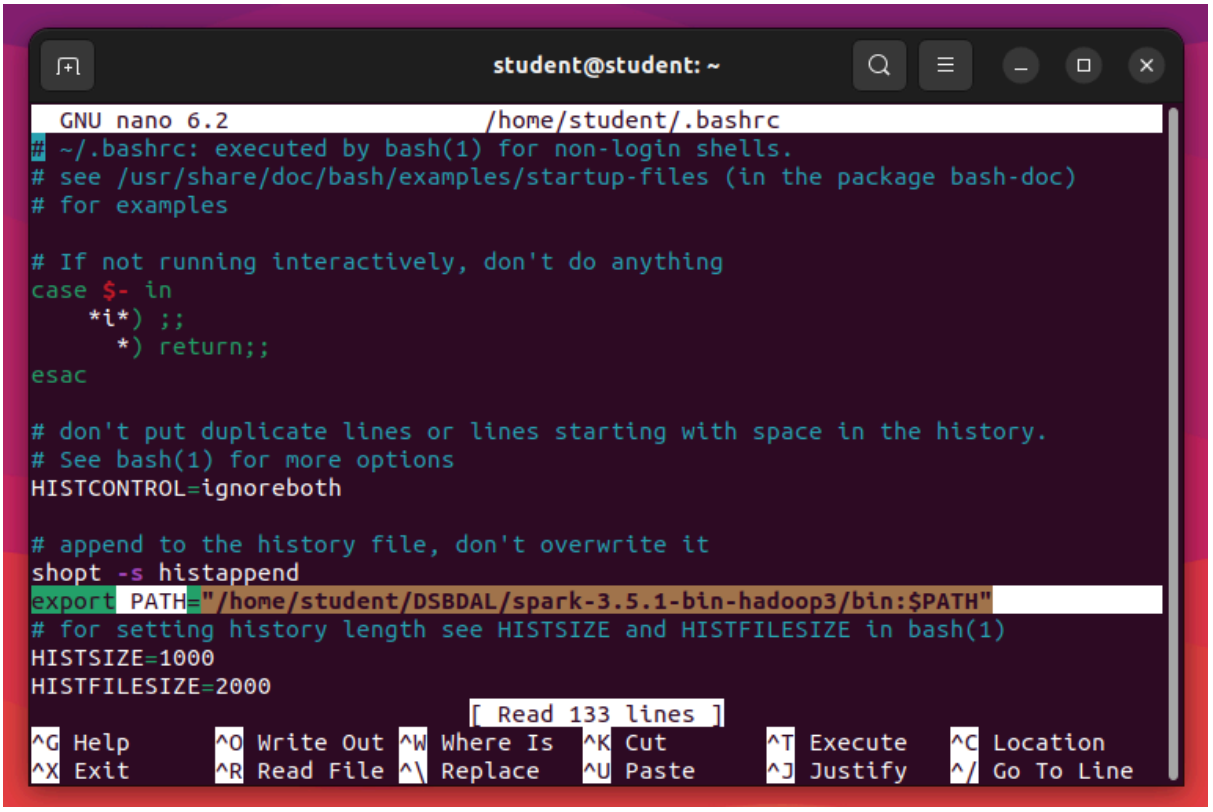
Step 1: Install Apache Spark

1. **Download Apache Spark**: Go to the Apache Spark website (https://spark.apache.org/downloads.html) and download the latest version of Apache Spark.
2. **Extract Spark**: After downloading, extract the contents of the downloaded file to a directory of your choice. (Create DSBDAL folder in home directory and extract the contents)
3. **Set up Environment Variables**: Add Spark's `bin` directory to your `PATH` environment variable. You can do this by modifying your shell profile file (e.g., `.bashrc`, `.bash_profile`, `.zshrc`, etc.) (Our is bash)
    a. **Open your Terminal**: Launch your terminal application. This process may differ slightly depending on your operating system (e.g., Terminal on macOS, Command Prompt or PowerShell on Windows, Terminal on Linux).
    b. **Determine your Shell**: Before proceeding, determine which shell you're using. Common shells include Bash, Zsh, and Fish. You can typically find out your current shell by running the following command:

    ```
    echo $SHELL
    ```

    c. **Edit the Shell Profile File**: Based on your shell, you'll edit the corresponding profile file:
        i. For Bash (`~/.bashrc`):
        ```
        nano ~/.bashrc
        ```
    d. **Add Spark's bin Directory to PATH**: Inside the opened file, add the following line at the end:

    ```
    export
    PATH="/home/student/DSBDAL/spark-3.5.1-bin-hadoop3/bin:$P
    ATH"
    ```



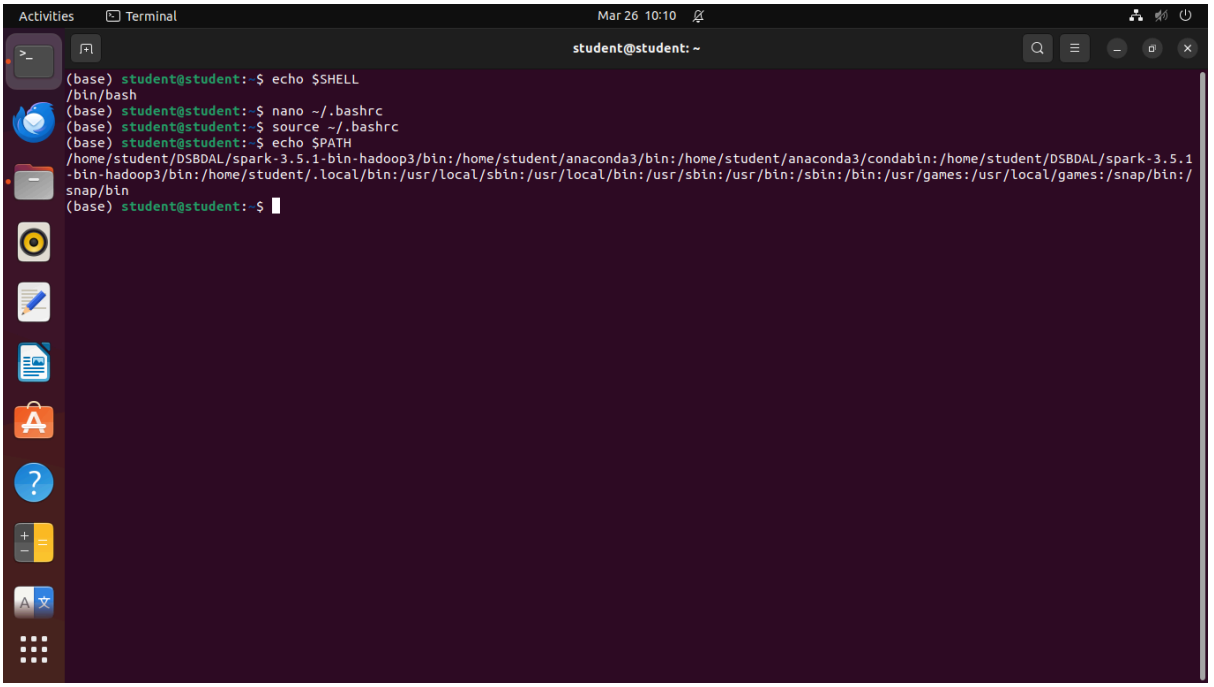
e. **Save and Exit**: After adding the line, save the file and exit the editor. In Nano, you can do this by pressing `Ctrl + O` to write the file and `Ctrl + X` to exit.
f. **Apply Changes**: To apply the changes to your current terminal session, either close and reopen the terminal or run:

   ```
   source ~/.bashrc
   ```

g. **Verify**: You can verify that Spark's `bin` directory has been added to your `PATH` by running:

   ```
   echo $PATH
   ```

   You should see the path to Spark's `bin` directory listed in the output.



4. **To execute Code:**

```
cd DSBDAL

spark-shell<WebLog_Processing.scala
```