

# Assignment 1

[Re-submit Assignment](#)

---

**Due** 1 Aug by 23:59      **Points** 20      **Submitting** a file upload  
**Available** after 30 Jul at 13:00

---

## Assignment 1 – Due date: Aug 1, 11.59pm

Log in to Marvin.

Write down each command line you execute to implement the following steps:

1. In your home directory, create a directory called "Assignment\_1".
2. Within Assignment 1, create two other directories called INPUT and OUTPUT.
3. Download the data from the following link to the INPUT directory.

[http://ssgac.org/documents/SWB\\_Full.txt.gz](http://ssgac.org/documents/SWB_Full.txt.gz) [\(http://ssgac.org/documents/SWB\\_Full.txt.gz\)](http://ssgac.org/documents/SWB_Full.txt.gz)

4. Create a copy of SWB\_Full.txt.gz in the INPUT directory, name it Mydata.txt.gz
5. Display the first 20 lines of Mydata.txt.gz without decompressing the file.
6. Decompress the file.
7. Display how many lines (observations) there are in the data.
8. Display the lines containing the string "rs10830206".
9. Extract the first column from the data and write it into another file in the OUTPUT directory. Name the file "SNPID.txt".
10. In a single command line (using a pipe), extract the last 100 lines from SNPID.txt and sort them. Write the output into another file in the OUTPUT directory named "100\_SNPs.txt"
11. Display the following sentence: "My name is YourName".
12. Write your command line history into a file in the OUTPUT directory, named "History.txt".
13. Display the contents of "History.txt", one screen at a time.
14. Go into the OUTPUT directory using its relative pathname. Create a tarball (tar archive) containing the files "History.txt" and "100\_SNPs.txt", named "YourName.tar"

Congratulations! You are done with the first assignment!

Upload your answers to questions 1-13 above and the "YourName\_assignment1.tar" file to Canvas.