

Reference for mapper1.py code:

1.

GeeksforGeeks. (2020, April 28). *Take input from stdin in Python*. GeeksforGeeks. <https://www.geeksforgeeks.org/take-input-from-stdin-in-python/>

2.

Dari Obukhova. (2019, May 21). *Reading CSV file from stdin in Python and modifying it*. Stack Overflow. <https://stackoverflow.com/questions/56234034/reading-csv-file-from-stdin-in-python-and-modifying-it>

3.

Python Software Foundation. (2020). *csv — CSV File Reading and Writing — Python 3.8.1 documentation*. Python.org. <https://docs.python.org/3/library/csv.html>

4.

Writing An Hadoop MapReduce Program In Python. (n.d.). Wwww.michael-Noll.com. <https://www.michael-noll.com/tutorials/writing-an-hadoop-mapreduce-program-in-python/>

5.

GeeksforGeeks. (2020, April 28). *Take input from stdin in Python*. GeeksforGeeks. <https://www.geeksforgeeks.org/take-input-from-stdin-in-python/>

6.

How to read one single line of csv data in Python? (n.d.). Stack Overflow. <https://stackoverflow.com/questions/17262256/how-to-read-one-single-line-of-csv-data-in-python>

7.

Reading CSV files in Python. (2019, December 16). GeeksforGeeks. <https://www.geeksforgeeks.org/reading-csv-files-in-python/>

8.

W3 Schools. (2024). *Python Try Except*. Wwww.w3schools.com.
https://www.w3schools.com/python/python_try_except.asp

9.

H.Choi. (2012, August 10). *I want to exception handle “list index out of range.”* Stack Overflow.
<https://stackoverflow.com/questions/11902458/i-want-to-exception-handle-list-index-out-of-range>

10.

geeksforgeeks. (2016, March 12). *Python Exception Handling*. GeeksforGeeks.
<https://www.geeksforgeeks.org/python-exception-handling/>

Reference for reducer.py code:

1.

GeeksforGeeks. (2019, December 6). *Declare an empty List in Python*. GeeksforGeeks.
<https://www.geeksforgeeks.org/declare-an-empty-list-in-python/>

2.

Syntax behind sorted(key=lambda: ...). (n.d.). Stack Overflow.
<https://stackoverflow.com/questions/8966538/syntax-behind-sortedkey-lambda>

Reference for run_pokemon_mapreduce.sh code:

1.

user7649922. (2017, March 7). *How can I delete a file only if it exists?* Stack Overflow.
<https://stackoverflow.com/questions/42655177/how-can-i-delete-a-file-only-if-it-exists>

2.

Bash Reference Manual. (2022). Gnu.org.

<https://www.gnu.org/software/bash/manual/bash.html#Special-Parameters>

3.

Apache Hadoop 3.3.1 – HDFS Architecture. (n.d.). Hadoop.apache.org.

<https://hadoop.apache.org/docs/stable/hadoop-project-dist/hadoop-hdfs/HdfsDesign.html>

4.

Apache Hadoop 3.4.1 – Hadoop Commands Guide. (2024). Apache.org.

<https://hadoop.apache.org/docs/stable/hadoop-project-dist/hadoop-common/CommandsManual.html>

5.

Bash Reference Manual. (2022). Gnu.org.

<https://www.gnu.org/software/bash/manual/bash.html#Special-Parameters>

6.

File System Shell Guide. (2025). Apache.org.

https://hadoop.apache.org/docs/r1.2.1/file_system_shell.html#test

7.

Liondancer. (2016, January 29). *using hdfs dfs -test to see if directory exists*. Stack Overflow.

<https://stackoverflow.com/questions/35094078/using-hdfs-dfs-test-to-see-if-directory-exists>

8.

Apache Hadoop MapReduce Streaming – Hadoop Streaming. (n.d.). Hadoop.apache.org.

<https://hadoop.apache.org/docs/current/hadoop-streaming/HadoopStreaming.html>

9.

Writing An Hadoop MapReduce Program In Python. (n.d.). Wwww.michael-Noll.com.
<https://www.michael-noll.com/tutorials/writing-an-hadoop-mapreduce-program-in-python/>

10. The following reference was used to configure all the Hadoop files that were required:

References Hadoop Streaming Documentation:

Apache Hadoop MapReduce Streaming – Hadoop Streaming. (n.d.). Hadoop.apache.org.
<https://hadoop.apache.org/docs/current/hadoop-streaming/HadoopStreaming.html>

References for Setup Hadoop on Ubuntu: Arjun Krish's Guide:

Arjun Krishna K. (2023, November 7). *Step-by-Step Guide to Setting Up Hadoop on Ubuntu: Installation and Configuration Walkthrough.* Medium. <https://arjunkrish.medium.com/step-by-step-guide-to-setting-up-hadoop-on-ubuntu-installation-and-configuration-walkthrough-60e493e9370d>