**CPRE 419**

**Lab 5**

**Eranda Sooriyarachchi & Jyothi Prasanth**

**Experiment 1**

In experiment 1 we use 2 reducers. We are making key, value pairs of integer 1. We are reducing the value by key to produce total count of each row. Then key and value pairs are swapped, and they are sorted by the order of key. This will give us the frequencies of words in the decreasing order. In this case key = frequency, value = word.

Graphical user interface

Description automatically generated

Graphical user interface

Description automatically generated

**Experiment 2**

In experiment 2, you read ip\_trace and make a key value pair of connection id and value is the complete string. Then it reads the raw\_block and makes key value pairs of connection id and status respectively. You then filter the blocked status connection ids. Then you join the key value pairs with each other. So now you will have connection id, complete string, and status. Then print this file as the firewall file.

Next you traverse firewall file and make key value pairs where key is source id and value are integer 1. Then reduce it so that you get the count of each source id. Then you swap the key value pairs and sort in descending order. This gives the most frequently blocked IP address.

Firewall file containing details of blocked connections.

Text

Description automatically generated

List of all the unique IP addresses along with their frequencies, that were blocked.

Text

Description automatically generated