Folder named RUC13 contains some data that is available about 26% of the entire data.

We need to match the training data to the data we have collected. A file called Training\_Data.csv contains a list of dust events with the given date and spacial location (lat, long).

**File description:**

Training\_Data.csv :

Contains 4 columns

For example:



We need to find compressed file of 2002.03.01 and extract the specific data at (-107.40643,31.594833) from 317 layers in one grb2 file

**#of instances:**

548 rows of Training\_Data.csv \* num of grb2 we choose

**# of features:**

317 layers + date + latitude + longitude

\*Now, each compressed file contains 30~50 grb2 files, which file do we need to choose is a problem

|  |  |  |  |
| --- | --- | --- | --- |
|  |  |  |  |

**Programs to write:**

1.) A program that checks to see if we have data for a given date based on the training data.

2.) A program that extracts all the variables and their values from the specified location (lat, long) this program should also export data into a csv file.

**Format of training data:**

date(1) **||** lat **||** long **||** variable 1 **||** variable 2 **||** . . . . . **||** variable n-1 **||** variable n **||**

date(2) **||** lat **||** long **||** variable 1 **||** variable 2 **||** . . . . . **||** variable n-1 **||** variable n **||**

date(3) **||** lat **||** long **||** variable 1 **||** variable 2 **||** . . . . . **||** variable n-1 **||** variable n **||**

date(4) **||** lat **||** long **||** variable 1 **||** variable 2 **||** . . . . . **||** variable n-1 **||** variable n **||**

date(5) **||** lat **||** long **||** variable 1 **||** variable 2 **||** . . . . . **||** variable n-1 **||** variable n **||**

**.**

**.**

**.**

date(m-1) **||** lat **||** long **||** variable 1 **||** variable 2 **||** . . . . . **||** variable n-1 **||** variable n **||**

date(m) **||** lat **||** long **||** variable 1 **||** variable 2 **||** . . . . . **||** variable n-1 **||** variable n **||**